DNA barcoding in *Momordica* spp.

By

GIRME AOUDUMBAR RAMESH
(2012-11-104)

THESIS

Submitted in partial fulfillment of the requirement

forthe degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR – 680656

KERALA, INDIA

2014

DECLARATION

I hereby declare that the thesis entitled **"DNA barcoding in *Momordica* spp."** is a bonafide record of research work done by me during the course of research and that it has not been previously formed the basis for the award to me of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara
Date: 28.09.2014

Girme Aoudumbar Ramesh
(2012-11-104)

CERTIFICATE

Certified that the thesis entitled **"DNA barcoding in *Momordica* spp."** is a record of research work done independently by **Mr. Girme Aoudumbar Ramesh** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, associateship, fellowship to him.

Vellanikkara
Date: 23·09·2014

Dr. Deepu Mathew
(Chairman, Advisory Committee)
Assistant Professor
Centre for Plant Biotechnology
and Molecular Biology
College of Horticulture
Vellanikkara, Thrissur

CERTIFICATE

We, the undersigned members of the advisory committee of **Mr. Girme Aoudumbar Ramesh** a candidate for the degree of **Master of Science in Agriculture** with major field in **Plant Biotechnology,** agree that the thesis entitled **"DNA barcoding in *Momordica* spp."** may be submitted by **Mr. Girme Aoudumbar Ramesh** in partial fulfilment of the requirement for the degree.

28/9/2014

Dr. Deepu Mathew
(Chairman, Advisory Committee)
Assistant Professor
Centre for Plant Biotechnology
and Molecular Biology
College of Horticulture

23/9/14

Dr. P. A. Valsala
(Member, Advisory committee)
Professor and Head
Centre for Plant Biotechnology
and Molecular Biology
College of Horticulture

Dr. P. A. Nazeem
(Member, Advisory committee)
Professor and Coordinator (DIC)
Centre for Plant Biotechnology
and Molecular Biology
College of Horticulture

Dr. P. S. Abida
(Member, Advisory committee)
Associate Professor
Centre for Plant Biotechnology
and Molecular Biology
College of Horticulture

23.9.14

Dr. D. Girija
(Member, Advisory committee)
Professor and Head
Department of Agricultural
Microbiology
College of Horticulture

23.9.14

Dr. Joseph John K.
(Member, Advisory committee)
Principle Scientist
NBPGR Regional Station
Thrissur

External Examiner
Dr. Sobha S.
Senior Scientist
Rubber Research Institute of India, Kottayam

23/9/14

ACKNOWLEDGEMENT

I bow to the lotus feet of ***Almighty*** *whose grace had endowed me the inner strength and confidence and blessed me with a helping hand at each step during my long sojourn at Kerala.*

At this moment of accomplishment, my heart is overwhelmed with gratitude and I wish if these words could convey the subtle feelings.

I am ineffable in expressing my deep sense of gratitude to my Major Advisor ***Dr. Deepu Mathew,*** *Assistant Professor, CPBMB, College of Horticulture, Vellanikkara, for his inspiring guidance, critical comments, constant supervision, support and encouragement throughout the course of my research work. I am really grateful for the keen interest taken by him in the prompt correction of the manuscript.*

My sincere thanks are due to, ***Dr. P. A. Valsala,*** *Professor and Head, CPBMB, College of Horticulture, Vellanikkara, member of my advisory committee for the valuable guidance, timely suggestions and help rendered in the conduct of the experiment.*

I also owe my deep sense of gratitude and sincere thanks to ***Dr. P. A. Nazeem,*** *Professor and coordinator (DIC), CPBMB, College of Horticulture, Vellanikkara, member of my advisory committee for the her timely help, critical scrutiny of the manuscript and valuable guidance for the successful completion of my research work.*

I am extremely thankful to ***Dr. D. Girija,*** *Professor and Head, Department of Agricultural Microbiology College of Horticulture, Vellanikkara, member of my advisory committee for her support, suggestions and guidance throughout the study.*

I take this opportunity to express my indebtedness with veneration and devotion to ***Dr. Joseph John K.,*** *Principal Scientist, National Bureau of Plant Genetic Resources, Regional Station, Thrissur, member of my advisory committee for his constant support, suggestions, valuable guidance and timely help rendered in the conduct of the experiment.*

I am thankful to ***Dr. P. S. Abida,*** *Associate Professor, CPBMB, College of Horticulture, Vellanikkara, member of my advisory committee for her support, suggestions and guidance throughout the study.*

Words fall short as I place on record my indebtedness to ***Dr. M. R. Shylaja*** *of CPBMB for their untiring support and valuable guidance during my course of study.*

My sincere thanks to ***Shailaja, Simi, Anu, Githu, Bincy Reji, Aishwary, Nayana, Sandhya, Teena, Manila, Divya, Lidya, Sinjula, Shruti*** *and* ***Vipin*** *of CPBMB, who helped me in several ways for the completion of this venture.*

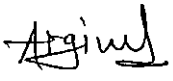
I wish to express my sincere thanks to all the non-teaching staff members and labourers of CPBMB for their whole-hearted cooperation and timely assistance. I also appreciate and acknowledge the facilities provided at ***CPBMB*** *and* ***Bioinformatics Centre*** *for the successful completion of my research work.*

My heartfelt gratitude cannot be captured in words for the unflinching support, constant encouragement, warm concern, patience and valuable advice of my friends ***Shriram, Sachin, Yogesh, Purushottam, Datta, Yashwant, Dheemant, Arun, Vaishali, Renuka, Shital, Vilas and Basavraj*** *whose prayers, love and affection rendered me a successful path which propped up my career all along. I wish to express my sincere gratitude to my senior friends* ***Ashok, Amar, Pawan, Parimal, Ravindra, Sharda, Subhashini, Pranita, Manibala, Pujaita, Agatha, Rohini, Sreeja, Sidhesh, Vikram, Ranjith*** *and* ***Ajay*** *for their whole hearted support.*

My duty is incomplete if I forget my junior friends ***Dolagabinda, Nabarun, Naresh, Ramesh, Rahul, Donjo, Manikesh, Irfana*** *whose helping hands, love and affection fetched a remarkable place in my days in Kerala.*

The award of ***DBT fellowship*** *is greatly acknowledged.*

Above all, I am forever beholden to ***my beloved parents, my brothers, my sisters and my family members*** *for their unfathomable love, boundless affection, personal sacrifice, incessant inspiration and constant prayers, which supported me to stay at tough tracks.*



Girme Aoudumbar Ramesh

Dedicated to my Beloved Parents and Teacher

TABLE OF CONTENTS

LIST OF TABLES

LIST OF PLATE

LIST OF FIGURES

LIST OF ANNEXURES

ANNEXURE NO.	TITLE
I	List of laboratory equipments used for the study
II	Reagents required for DNA isolation
III	Composition of buffers and dyes used for gel electrophoresis

ABBREVIATIONS

%	Percentage
@	At the rate
<	Less than
=	Equal to
>	Greater than
µg	Microgram
µl	Microlitre
AFLP	Amplified Fragment Length Polymorphism
BLAST	Basic Local Alignment Search Tool
bp	Base pair
cm	Centimetre
CPBMB	Centre for Plant Biotechnology and Molecular Biology
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribo Nucleoside Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
g	Gram
ha	Hectare
ISSR	Inter Simple Sequence Repeat
Kb	Kilo basepairs
L	Litre
M	Molar
Mb	Mega bytes
mg	Milligram
ml	Millilitre
mM	Milli mole

NCBI	National Centre for Biotechnology Information
ng	Nanogram
°C	Degree Celsius
OD	Optical Density
PCR	Polymerase Chain Reaction
pH	Hydrogen ion concentration
PVP	Poly vinyl pyrolidone
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
SCAR	Sequence Characterized Amplified Region
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeats
TAE	Tris Acetate EDTA
TE	Tris EDTA
U	Unit
UV	Ultra violet
V	Volts
β	Beta
BOLD	Barcode of Life Database System
CBOL	Consortium for the Barcode of Life

Introduction

1. INTRODUCTION

The genus *Momordica* derived from Latin name 'mordeo' (momordi = to bite) to mention the jagged seeds, is comprised of 59 species. *Momordica charantia* (common name: bitter gourd, karela or balsam pear or bitter melon) is a vegetable with many culinary uses especially in Asia and Africa. It is also grown as an ornamental and has been used for centuries in ancient traditional Indian, Chinese, and African pharmacologist. Among the cucurbits, *M. charantia* and *M. dioica* are considered prized vegetables because of their high nutritive values especially ascorbic acid and iron and medicinal properties (Chakravarty, 1959; Behera *et al.*, 2008). In India, all the *Momordica* species except bitter gourd (*M. charantia*), are being gathered from wild and eaten, besides being used in genetic improvement of the species.

Taxonomic confusion exists in *Momordica* spp. and details of their floral biology, system of evolution and inheritance are poorly understood. The botanical names and common names are often used incorrectly or interchangeably (Joseph *et al.*, 2007). Different taxonomic classification approaches have resulted in controversies about the number of species that exist and the phylogenetic relationships among these species. According to the latest revision of *Momordica* spp. of Indian origin (Joseph, 2005), there are six well identified species of which four are dioecious and two are monoecious. The monoecious taxa are *M. charantia* L. (2n = 22) and *M. balsamina* L. (2n = 22). The dioecious taxa are *M. dioica* Roxb. ex Willd. (2n = 28), *M. sahyadrica* Joseph *et* Antony (2n = 28), *M. cochinchinensis* (Lour.) Spreng. (2n = 28) and *M. subangulata* Blume subsp. *renigera* (G. Don) W.J.J de Wilde (2n = 56). However, *M. cymbalaria* (Hook. Fenzl ex Naud.), which is expected to be under *Momordica* has not been included under the class of *Momordica* of Indian origin. Though a number of varieties belonging to different *Momordica* species have been developed in India, no information is available on their genetic base.

Advancements in DNA technology have resulted in an array of tools for DNA polymorphism assays. DNA based molecular markers are used to provide a relative estimation of genetic diversity and establish genetic relationships. But, DNA molecular marker analysis is time-consuming, required quality DNA samples, costly and laborious and depends largely on the experience of observers. Further, the molecular marker analysis in many cases will end up in genetic distance estimation only but fails to differentiate the species (Hebert *et al.*, 2004). Accordingly, an accurate, sensitive and simple alternate is needed.

DNA barcoding is a novel system designed to provide rapid, accurate, and automatable species identification using short, standardized gene regions as internal species tags. As a consequence, it will make the Linnaean taxonomic system more accessible, with benefits to ecologists, conservationists, and the diversity. In addition to assigning specimens to known species, DNA barcoding will accelerate the pace of species discovery by allowing taxonomists to rapidly sort specimens and by highlighting divergent taxa that may represent new species (Hebert *et al.*, 2004).

Understanding the extent of natural variation and phylogenetic relationship at molecular level is essential to develop DNA barcode for *Momordica* spp. At present, *rbcL, matK, psbA-trnH, rpoC1, rpoC1, ITS2, atpF–atpH* spacer and *psbK–psbI* spacer have been popularly used as DNA barcodes in plants worldwide (Janzen, 2009).

Previous work in the taxonomic line in this genus was restricted to the morphology and to the molecular markers such as RAPD to assess the genetic diversity (Dey *et al.*, 2006b). ISSR markers have also been employed in genetic diversity analysis of different cucurbits (Dje *et al.*, 2006; Levi *et al.*, 2004; Ritschel *et al.*, 2004). Phylogenetic relationship among the different monoecious and dioecious *Momordica* species has also been studied using plastid and mitochondrial DNA based markers (Schaefer and Renner, 2009).

No systematic effort at molecular level has yet been made to confirm the enormous variability reported and studied in this genus in the cultivated Indian species varieties and wild species. With the well established use of species discrimination, DNA barcode have to be employed in *Momordica* since high level of controversy persists in the species allocation of the Indian *Momordica* accessions.

With these points under consideration, the present study was undertaken with the objective to assess the genetic diversity and inter species relationship among 25 Indian *Momordica* lines belonging to seven different species and to develop the DNA barcode for various Indian *Momordica* species.

Review of Literature

2. REVIEW OF LITERATURE

The present study on "DNA barcoding in *Momordica* spp." has been executed through the amplification of chloroplast *matK* gene from different *Momordica* accession, belonging to seven species. Then amplified chloroplast *matK* gene fragment were sequenced and the sequence information was further used to identify the species relation. The relevant literatures available on various aspects of this study were collected and are reviewed in this chapter, under different heads.

2.1 Origin and History

The genus *Momordica,* best known as bitter gourd, derived its name from Latin word 'mordeo' (momordi=to bite). The genus *Momordica* belongs to the tribe Joliffieae (Schrad.) of the family Cucurbitaceae and is a native of the Palaeotropics (Robinson and Decker-Walters, 1997). This family comprise of about 40 species (Chakravarty, 1959), distributed in warm tropics of both hemispheres, chiefly in Africa and with about 10 species in South-East Asia (De Wilde and Duyfjes, 2002). Bitter gourd is believed to have its centre of origin at Eastern India and Southern China (Sands, 1928; Walters, 1989).

The *Momordica* in India consists about seven species and each have importance as crop wild relative of bitter gourd, besides their utility as nutritious vegetable. As their commercial importance is largely local, the extent of production and area under cultivation are often under estimated. However, the average yield is proposed to be around 20 tonnes per hectare (Lingaiah *et al*, 1993) and in Kerala, it's beyond 50 tonnes.

These widely consumed nutritious vegetables have been relatively neglected in research. Local people in India use wild, forms such as teasle gourd, sweet gourd and spin gourd as vegetables, yet there is no data to show the extent of genetic diversity. Often the taxonomic descriptions on *Momordica* are based on insufficient data and scientific support, leading to confusion on species categories.

So it is necessary to make taxonomic accounts of *Momordica* species more comprehensively.

2.2 Species diversity

The family Cucurbitaceae consist of 118 genera and 825 species (Jeffery, 1980a; 1980b) mostly tropical and subtropical. The family is further divided in to two well defined subfamilies, eight tribes representing varying degree of circumscriptive cohesiveness. The genus *Momordica* belongs to the subtribe Thaladianthinae pax, tribe Joliffeae Schrad, and subfamily Cucurbitoideae. In India, the family Cucurbitaceae is represented by 34 genera and 108 species of which 38 species are endemic (Pandit and Acharya, 2008). These indicate existence of rich genetic diversity and variability which need of the breeder for crop improvements.

Inter specific diversity in the genus *Momordica* has been estimated differentially by different scientist. The maximum number of species so far reported in *Momordica* is 35. In the Flora of Tropical Africa (Oliver, 1979), 14 species of *Momordica* are described. They are *M. cardiospermoides, M. cissoids, M. pterocarpa, M. angiosantha, M. trifoliate, M. balsamina, M. charantia, M. welwitcschii, M. cucullata, M. morkorra, M. mannii, M. corymbifera, M. multifera* and *M. cymbalaria*. Out of these 14 only *M. charantia, M. balsamina, M. cymbalaria* occurs in India. The first five species are distinct by 3-7 foliate compound leaves. About five species reported from Kenya (Nijoroge and Newton, 2002) of which *M. calantha* A. Zimm. do not figure in the above list. The Tanzanian endemic *M. glabra* Zimm, *M. leiocaroa* Gilg and *M. pycantha* Hams also do not figure in the list. In south East Asia 10 species are reported, of which six each occur in Malaysia (deWilde and Duyfies 2002) and India, where *M. charantia, M. subangulata* and *M. cochinchiesis* are common and Sri Lanka has three species of which *M. charantia* and *M. dioica* are represented in India too.

2.3 Species available in India

The genus *Momordica* L. in India includes several important vegetable cum medicinal plants which has the centre of diversity and distribution in the Indo-Malayan region extending to China, Myanmar, Nepal and Bangladesh. All India Coordinated Project on Underutilized Plants has also taken *Momordica dioica* as priority crop for development as a vegetable (Joshi *et al.*, 2002)

(i) *M. charantia*: cultivated and wild vegetable in Asia, monoecious in nature, also called bitter gourd, n= 11, germination epigeal, annual, non-tuberous, muricate-tubercled, seed sides- rectangular, leaf shape- angular, two sub species are available in India, the cultivated *M. charantia* var. *charatia* and *M. charantia* var. *muricate.*

(ii) *M. c.* var. *muricate:* wild gathered, high value vegetable, multipurpose, medicinal tuber, and ornamental value, n= 11, germination epigeal, annual, non-tuberous, muricate-tubercled.

(iii) *M. dioica*: wild gathered, anthesis in the evening, flower small, pale yellow, intensely musky scented, male calyx whitish yellow, sepals of male flower narrow acute.

(iv) *M. cochinchinesis*: under exploited vegetable, leaf unlobed or shallowly 3 lobed, margins undulate, male calyx green, broad, tip triangular, fruit with short conical projections, seeds large, smooth on surface.

(v) *M. subangulata*: also called teasle gourd wild and cultivated vegetable, Germination hypogeal, perennial, taproot tuberous, plant dioecious, nectar of the male flower closed with prominat scales, fruit echinate, petal with black purpal bloth, male calyx- hypanthium saucer shaped, leaf cordate, unlobed, margin dentate.

(vi) *M. balasmina*: ornamental, wild vegetable-famine food, medicinal, camel fodder, pickle, flavouring, monoecious in nature, also called bitter gourd, n= 11, germination epigeal, annual, non-tuberous, muricate-tubercled, seed sides- rectangular, leaf shape- angular.

(vii) *M. sahydrica*: wild fruit and leafy vegetable, petals without purpale blotch, male calyx hypanthium cup shaped, flower large showy, bright yellow, feeble scented, male calyx blackish purple.

(viii) *M. cymbalaria*: monoecious in nature, perennial, anthesis late in morning, fruit surface ribbed and seeds were smooth.

2.4 Taxonomy and distribution

The taxonomic treatment of the genus *Momordica* is extensive. Generic and species description with keys are found to varying degree in various publication during pre Indian independence period. Rhodes' (1678) description and illustration of 'Paval' is the first printed record. Linnaeus (1800), DeCandolle (1828), Roxburgh (1832), Clarke (1879), Cooke (1908), Gamble and Fischer (1919), Blatter (1919) and Kanjilal (1938) have dealt with the systematic of the genus extensively. During the post independence period, Santapau (1953), Saldhana and Nicolson (1976), Chakravarthy (1982) and Mathew (1981, 1983) have treated the genera in their floristic work. Many of the regional and district floras also mention and gives small description of various *Momordica* species (Srivastava, 1976; Oommachan, 1977; Bhandari, 1978; Naik, 1979; Rao, 1985; Shetty and Singh, 1987; Ramachandran and Nair, 1988; Vajravelu, 1990; Narasimhan and Sharma, 1991; Sasidharan and Sivarajan, 1996; Sivarajan and Mathew, 1997; Pallithanam, 2001; Singh *et al.*, 2002; and Bhat, 2003). Occurrence along with detailed technical description of *M. dioica and M. charantia* from Nallamalais (Ellis, 1987) and Kurnool (Raju and Pulliah, 1995) of Andhra Pradesh are reported. Saldhana (1976) gave detailed technical description along with key and occurrence of *M. charantia, M. dioica, M. subangulata* and *M. cochinchinesis* from Karnataka. *Momordica subangulata* is reported as rare in semi evergreen and evergreen forest of Belgaum, Ester Karnataka and *M. dioica* as frequent in wet forest of Western Ghats and Deccan (Saldhana, 1985).

2.5 Cytological studies

According to the description of Indian *Momordica* spp. by Joseph (2005) there are six species identified in which four are dioecious and two monoecious. The monoecious are *M. charantia* L. (2n = 22) and *M. balsamina* L. (2n = 22). The dioecious taxa are *M. dioica* Roxb. (2n = 28), *M. sahyadrica* (2n = 28) (Joseph and Antony, 2007a), *M. cochinchinensis* (Lour.) Spreng. (2n = 28) and *M. subangulata* Blume subsp. *renigera* (G. Don) (De Wilde and Duyfjes, 2002) (2n = 56).

In a study entitled "Cytotaxonomical analysis of *Momordica* L. (Cucurbitaceae) species of Indian occurrence", Bharathi *et al.* (2011) have documented the chromosome number and karyotype analysis in all six Indian species to reconfirm the chromosome status and to study the cytotaxonomic relationship of *M. cymbalaria* with other *Momordica.*

Bharathi *et al.* (2010) have provided the cyto-morphological evidence for the segmental allopolyploid origin of teasle gourd (*Momordica subangulata* ssp. *renigera*) in which it shows that teasle gourd (2n = 56) gives the same morphological characters which were present in both *M. dioica* (2n = 28) and *M. cochinchinensis* (2n = 28). Morphological analysis of *M. subangulata* subsp. *renigera* suggests an allopolyploid origin. A metaphase study also had shown that the genomes of these species are partially homologous. A higher number of bivalents in the triploid hybrids suggest that *M. subangulata* subsp. *renigera* is a segmental allopolyploid of *M. dioica* and *M. cochinchinensis* and that its genome has diverged from the parental genomes.

A crossing programme by Bharathi *et al.* (2010), in which seven species of *Momordica* and two varieties of *Momordica charantia* were used. In inter-subspecies cross between *M. c.* var. *charantia* × *M. c.* var. *muricata,* high crossability and pollen fertility were observed, whereas it was poor in inter-specific cross *M. charantia* × *M. balsamina.* The species *cochinchinensis* was found uncrossable, except for *M. cochinchinensis* × *M. dioica* and *M.*

cochinchinensis × *M. sahyadrica*. *M. dioica* and *M. sahyadrica* showed higher crossability with *M. subangulata* subsp. *renigera* in both directions and *M. cochinchinensis* in unidirectional. *M. s.* subsp. *renigera* had reproductive compatibility with *M. cochinchinensis* in both directions (Bharathi *et al.*, 2012). *M. sahyadrica* and *M. dioica* showed high crossability in both directions and produced fertile hybrids.

Cross between the sect. *Momordica* and *cochinchinensis* yielded parthenocarpic fruits. *M. cymbalaria* was not found crossable with sect. *Momordica* and also with sect. *cochinchinensis*. Based on crossability, relationship was found to be closer in two varieties var. *Charantia* and var. *muricata* and also between *M. charantia* and *M. balsamina* (Bharathi *et al.*, 2012). All dioecious species included in this study appear to be closely related. The result supported that recent taxonomic revision of the genus and the gene pool classification provides a base for improvement of *Momordica* species.

2.6 Genetic markers in plant taxonomy

Genetic markers are measurable inherited genetic variations, used to understand genetic components. There are different types of genetic markers with different properties, each having its own characteristics to assess the genetic variations among populations. Presently, the most commonly used genetic markers are molecular markers. Due to the rapid developments in the field of molecular genetics, during the last few decades a variety of techniques have emerged to analyze genetic variation (Whitkus *et al.*, 1994; Karp *et al.*, 1996; 1997; Parker *et al.*, 1998; Schlötterer, 2004). These genetic markers may differ with respect to important features, such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and financial investment. No marker is superior to all others for a wide range of applications. The most appropriate and important genetic marker will depend on its specific application, the presumed level of polymorphism, the presence of sufficient technical facilities or know-how, time constraints and financial

limitations. Generally markers are divided in to three broad classes: those based on visually assessable traits (morphological and agronomic traits), those based on gene product (biochemical markers), and those relying on a DNA assay (molecular markers). Genetic markers are widely used by breeders and conservationists to study genetic diversity.

2.6.1 Morphological markers

The genetic characterization and divergence are based on morphological markers such as qualitative and quantitative traits (Arriel *et al.,* 2007). These markers are inexpensive and simple to score and are based on distinct phenotypes such as plant colour, plant height, seed characteristics, etc. The apparent disadvantage of such markers is prominent in studies of genetic diversity where the expression of the phenotype is highly influenced by environmental conditions. Lack of adequate genome coverage because of limitation of the number of markers and problems of dominance can also be mentioned as weaknesses of morphological markers (Brown, 1978). Furthermore, the expression of these characters which are influenced by the environment may require that plants be grown to a suitable stage before certain characters can be scored.

2.2.2 Biochemical markers

Biochemical markers are markers derived from study of the chemical products of gene expressions. These are also termed as isozyme/allozyme markers or protein markers. According to Karp *et al.* (1997), the agro-morphological markers resulted in the development of biochemical markers that complement its drawbacks. According to Gottlieb (1981) the oldest biochemical technique used to study variations is isozyme analysis. It has the power to reveal polymorphism of alleles at particular locus on the basis of protein mobility (Brown, 1978). Isozyme technique is fast, cheap and simple. However, isozyme markers are not as plentiful as DNA markers; these underestimates the level of genetic diversity (Dudnikov, 2003) and sometimes interpretation of bands become difficult due to complex banding profiles arising from polyploidy or duplicate genes. In addition,

proteins with identical electrophoretic mobility (co-migration) may not be homologous (Morell *et al.*, 1995). Isozyme studies in plants have demonstrated differences in pattern and band intensities by tissue types and developmental stages (Montarroyos *et al.*, 2003). Although isozymes have limitations but the technique was used for genetic diversity analysis in many species (Dudnikov, 2003) and it appeared to be more informative at lower taxonomic levels, particularly at species and population level characterization (Brown, 1990). Isozyme markers were experimented to characterise *Coffea arabica* (Berthaud and Charrier, 1988). However, their use on wild coffea species characterization from Ethiopia had failed to reveal polymorphism, indicating that isozyme technique is not appropriate for diversity study in *Coffea* spp. due to the small number of isozyme systems available (Berthaud and Charrier, 1988). Paillard *et al.* (1996) also tried to construct isozyme based genetic map for coffee and were unsuccessful due to the low polymorphism level.

2.6.3 Molecular markers

Molecular markers are fragments of nuclear, mitochondrial or chloroplast DNA, with sequences specified by the marker system such as RFLP, RAPD, AFLP, ISSR, SSR etc. These technologies based on polymorphism in DNA, can be considered as objective measures of variations and have catalyzed research in a variety of disciplines such as phylogeny, taxonomy, ecology, genetics and plant breeding. Markers are informative only if, they are polymorphic in populations. Level of polymorphism is an important determinant of what a marker is useful for. Different types of molecular markers with different properties exist, each with its own advantages and disadvantages (Karp *et al.*, 1997; Weising *et al.*, 2005). However, it is extremely difficult to find molecular markers which could adequately hold all the ideal properties. Depending upon the type of the study to be undertaken, one can identify between varieties of marker systems that could fulfill the objective of the study (Weising *et al.*, 2005). Many authors also suggest the use of more than one type of molecular marker in a single experiment (Karp *et al.*, 1997). The DNA based marker systems are generally classified as

hybridization-based (non-PCR) markers and PCR based markers (Joshi *et al.*, 2000).

2.6.3.1 PCR- based molecular marker techniques

DNA fingerprinting is a technique, which has been widely adapted to differentiate organisms at the species and subspecies levels (McClean *et al.*, 1994). The techniques used for cultivar identification are designed to detect the presence of specific DNA sequences or combination of sequences that uniquely identify the plant. Cultivar identification can be achieved more accurately using DNA fingerprinting data, especially in materials characterized by high genetic variations between cultivars. The most closely related cultivars are usually distinguished with the DNA fingerprinting methods (Beckman and Soller, 1986). The application of DNA fingerprinting could be very valuable in the identification of cultivars and species and could help to create more efficient breeding programs through the detection of genetic linkages between DNA fingerprinting bands and agriculturally important quantitative trait loci (QTL). The high variability of DNA marker described in humans, animals and plants allows the identification of different individuals, genotypes, and species (Lin *et al.*, 1993).

The Polymerase Chain Reaction (PCR) technique has been the basis of a growing range of new techniques for genome analysis based on the selective amplification of genomic DNA fragments (Saiki *et al.*, 1988). Williams *et al.* (1990) reported the use of PCR with short oligonucleotide primers of arbitrary (random) sequence to generate markers, the basis of the Random Amplified Polymorphic DNA (RAPD). Welsh and McClelland (1990) also reported on Arbitrarily Primed Polymerase Chain Reaction (AP-PCR). DNA Amplification Fingerprinting (DAF) was also reported as another technique of PCR used in various experiments (Caetano-Anolles *et al.*, 1991). The PCR reaction requires deoxynucleotides, DNA polymerase, primer, template and buffer containing magnesium. Typical PCR amplification utilises oligonucleotide primers which hybridise to complementary strands. The product of DNA synthesis of one primer

serves as template for another primer. The PCR process requires repeated cycles of DNA denaturation, annealing and extension with DNA polymerase enzyme, leading to amplification of the target sequence. This results in an exponential increase in the number of copies of the region amplified by the primer (Saiki *et al.*, 1988). The technique can be applied to detect polymorphism in various plants, animals, bacterial species and fungi.

The introduction of the PCR technique has revolutionized standard molecular techniques and has allowed for the proliferation of new tools for detecting DNA polymorphism (Hu and Quiros, 1991). The electrophoresis pattern of fragments generated by each primer for one isolate can be used as DNA fingerprints for assaying diversity (Tommerup *et al.*, 1998). Insertion can change the size of a DNA fragment without preventing its amplification (Williams *et al.*, 1990). PCR is simple, fast, specific, sensitive and the main advantage of this technique over others is its inherent simplistic analysis and the ability to amplify extremely, small quantities of DNA (Welsch *et al.*, 1991).

2.6.3.2 Molecular markers in taxonomy

Markers based on differences in DNA sequences between individuals generally detect more polymorphism than morphological and protein based markers and thus constitutes a new generation of genetic markers (Bostein *et al.*, 1980; Tanksley *et al.*, 1989). Hence, varietal profiling methods that directly utilize DNA could potentially address all the limitations associated with morphological and biochemical data. Cultivar identification using DNA marker is currently being investigated in a number of laboratories using different methods (Weising *et al.*, 1995).

Polymerase chain reaction (PCR) based techniques. These include Randomly Amplified Polymorphic DNA (RAPD) (Williams *et al.*, 1990), Amplified Fragment Length Polymorphism (AFLP) (Vos *et al.*, 1995), Simple Sequence Repeats (SSR) (Tautz, 1989), Inter Simple Sequence Repeats (ISSR), DNA Amplification Fingerprinting (DAF), Sequence Tagged Sites (STS),

Sequence Characterized Amplified Regions (SCAR), Expressed Sequence Tags (EST) etc. Of these, SSR is co-dominant markers, while others are dominant. In almost all the known plants at least few of these markers are employed so far, for one or the other purpose.

2.7 Molecular marker analysis in *Momordica* spp.

Molecular markers like RAPD, AFLP and ISSR have been employed in *Momordica* spp. for identification of species and their relationships with other species.

As the taxonomic confusion exists in *Momordica* spp. and details of their floral biology, system of evolution and inheritance are poorly understood and also the botanical names and common names are often used incorrectly or interchangeably (Joseph *et al.*, 2007), all these studies are highly relevant and the reported species specific fingerprint have to be augmented into taxonomy.

Bharathi *et al.* (2012) have given the molecular diversity and relationship of *Momordica* spp. of Indian occurrence. They used twenty-one RAPD and twelve ISSR primers for assessment of genetic diversity and establishing phenetic relationships among 35 genotypes belonging to six Indian cultivated *Momordica* species and five genotypes of two *Luffa* species. The level of inter-specific diversity was maximum (90%) between *M. charantia* and *M. cochinchinensis* whereas the extent of intra-specific diversity was highest particularly in dioecious species (51%) as compared to monoecious species like *M. charantia* (38%). Wider divergence of the taxon of controversial identity, *M. cymbalaria* from the other Indian cultivated *Momordica* species and their evolutionary closeness with *Luffa* species was evident. The clustering pattern obtained among the genotypes belonging to different *Momordica* and *Luffa* species corresponded well with their morphological, cytological and taxonomic classification.

Dey *et al.* (2006b) used RAPD assay for assessing genetic diversity among the species of bitter gourd. The 38 genotypes of *M. charantia* which included the commercially cultivars were analysed for diversity.

Rasul *et al.* (2007) have studied the genetic diversity and identified the 29 cultivars of *Momordica dioica* Roxb. and 1 *Momordica cochinchinensis* accession using RAPD marker. From the cluster analysis using UPGMA method, *M. cochinchinensis* was out grouped as single accession, while others showing relatively weak grouping formed four groups.

RAPD assay has also been used in the related genera such as cucumber (Horejsi and Staub 1999), pumpkin (Gwanama *et al.*, 2000), watermelon (Lee *et al.*, 1996) and ash gourd (Sureja *et al.*, 2006) for understanding the extent of natural variation and phylogenetic relationship at molecular level. Besides, ISSR markers have been used in genetic diversity analysis of other cucurbits such as *Citrullus, Cucumis* and *Praecitrullus fistulosus* (Dje *et al.*, 2006; Levi *et al.*, 2004; Ritschel *et al.*, 2004; Levi *et al.*, 2005).

Dalamu *et al.* (2012) studied the genetic diversity of fifty indigenous and exotic genotypes of Asian bitter gourd (*Momordica charantia* L.) with RAPD and ISSR markers. The results had shown large genetic variability among the Asian bitter gourd genotypes which indicates that they should be considered as a valuable gene pool for bitter gourd breeding programs.

Gaikwad *et al.* (2008) used AFLP markers for analysis 38 Indian bitter gourd varieties. Six primer combinations were used and similarities among cultivars ranged between 0.44 and 0.88, indicating that the bitter gourd examined were genetically diverse. The cultivars were grouped with respect to geographical region, in which cultivars within a group and subgroups possessed high degrees of genetic similarity.

Bootprom *et al.* (2012) studied the molecular diversity among 25 selected *Momordica cochinchinensis* accessions of Thailand and Vietnam using RAPD markers. The coefficients of genetic similarity varied from 0.63 to 0.90, which suggested that it has a wide genetic base for the genotypes. The accessions were grouped into eight major clusters clearly differentiating with each other and clearly identifying male and female genotypes.

Paul *et al.* (2010) studied the genetic diversity of twelve accessions of *Momordica charantia* L. using morphological, RAPD and SCAR markers, in which 6 were the *Momordica charantia* var. *muricata* and remaining *Momordica charantia* var. *charantia*. RAPD analysis using 23 primers has generated two clusters, one showing *Momordica charantia* var. *muricata* and another *Momordica charantia* var. *charantia*.

2.8 DNA barcoding in plant systematics

DNA barcoding is a method for identifying the living organisms at species levels, in which a short universal gene sequence taken from a standardized portion of the genome is used. Dayrat (2005) opined that, 'delineating species boundaries correctly and also identifying species – are crucial to the discovery of life's diversity because it determines whether different individual organisms are members of the same entity or not'. The identification of species depends on the knowledge held by taxonomists whose work cannot cover all taxon identification requested by non specialists. To deal with these difficulties, the 'DNA Barcode of Life' project was started, which aims to develop a standardized, rapid and inexpensive species identification method accessible even to non-specialists.

Though several universal systems have been employed for molecular-based identification of lower taxa (Floyd *et al.*, 2002), they are not successfully implemented for higher taxa. The Barcode of Life project aims to create a universal system or inventory for eukaryotic species, based on standard molecular approaches. It has been initiated in 2003 by researchers at the University of Guelph, Ontario, Canada (http://www.barcoding.si.edu) and was promoted in 2004 by the international initiative 'Consortium for the Barcode of Life' (CBOL). The DNA barcode project attempts to produce a simple diagnostic tool based on strong taxonomic knowledge that is collated in the DNA barcode reference library (Schindel and Miller, 2005; Gregory, 2005). The DNA Barcode of Life Data System (BOLD, http://www.boldsystems.org) has progressively been developed since 2004 and was officially established in 2007 (Ratnasingham and Hebert,

2007). This data system enables the acquisition, storage, analysis and publication of DNA barcode records.

The DNA barcode project was initially conceived as a standard system for fast and accurate identification of animal species and presently all eukaryotic species are within its scope (Hebert *et al.*, 2003; Miller, 2007). DNA barcoding is based on the postulate that every species will most likely have a unique DNA barcode among the 4^{650} possible ATGC-combinations (Wilson, 2004) and that genetic variation between species exceeds variation within species (Hebert *et al.*, 2003; Hebert *et al.*, 2004). The two main ambitions of DNA barcoding are to (i) assign unknown specimens to species and (ii) enhance the discovery of new species and facilitate identification, particularly in cryptic, microscopic and other organisms with complex or inaccessible morphology (Hebert *et al.*, 2003).

DNA barcoding acquired the sequences of identified barcoding locus and based on this, the species delimitation will be computed, in compression with the existing DNA barcodes. Furthermore, in three important situations, relevant species identification must necessarily be molecular-based. First, in determining the taxonomic identity of damaged organisms or fragments. The DNA barcoding tool is thus potentially useful in the food industry, diet analyses, forensic sciences and in preventing illegal trade and poaching of endangered species (e.g. fisheries, trees, bush meat). Second, molecular-based identification is necessary when there are no obvious means to match adults with immature specimens. The third case is when morphological traits do not clearly discriminate species, especially when size precludes visual identification (Blaxter *et al.*, 2005; Webb *et al.*, 2006).

2.8.1 DNA barcoding in plants

Chase *et al.*, (2005) had given the standardised protocol to barcode all land plants using the loci *rpoC1+rpoB+matK* and *rpoC1+matK+psbA-trnH*. The protein encoding plastid gene *rbcL* has been proposed as a potential plant barcode by several researchers (Chase *et al.*, 2005; Newmaster *et al.*, 2006). Size of *rbcL* comes up to 1300 bp and this high length of DNA may not be available from

degraded samples and hence, another plastid DNA region proposed is the non-coding *psbA-trnH* spacer (Kress *et al.*, 2005; Shaw *et al.*, 2007).

A DNA barcode for land plants has also given by Janzen (2009). Where, 7 leading candidate plastid DNA regions *atpF–atpH* spacer, *matK* gene, *rbcL* gene, *rpoB* gene, *rpoC1* gene, *psbK–psbI* spacer, and *trnH–psbA* spacer were used. Based on recoverability, sequence quality, and levels of species discrimination, 2-locus combination of *rbcL* and *matK* was recommended as the plant barcode. This core 2-locus barcode will provide a universal framework for the routine use of DNA sequence data to identify the specimens.

Universality of single primer pair for each locus in angiosperms has been checked and it resulted in 90-98 % success in PCR and sequencing, for 6 regions out of 7 tried. Success for the seventh region, *psbK–psbI*, was 77 % (Janzen, 2009).

All 7 loci were sequenced from 397 sample and, species discrimination for single-locus barcodes ranged from 43% (*rpoC1*) to 68-69% (*psbK–psbI* and *trnH–psbA*), with *rbcL* and *matK* providing 61 and 66% discrimination respectively Two-locus combinations gave 59-75% resolution, and 3-locus combinations 65-76% discrimination.

DNA barcoding is widely used for species identification but there is not yet any universally accepted barcode for land plants. Lahaye (2007) has collected >1,600 samples from two biodiversity hotspots Mesoamerica and southern Africa. *trnH-psbA*, *matK*, *ycf5*, *rbcL* , *rpoB*, *ndhJ*, *accD* and *rpoC1* were identified as potential barcodes. With adequate rate of variation, easy amplification, and alignment, a portion of the plastid *matK* gene wasidentified as a universal DNA barcode for flowering plants.

Hollingsworth *et al.* (2009) studied the DNA barcoding using a divergent group of land plants *Inga* (angiosperm), *Araucaria* (gymnosperm) and *Asterella* (liverwort). In this study, seven main candidate plastid regions *rpoC1*, *rpoB*, *rbcL*, *matK*, *trnH-psbA*, *atpF-atpH* and *psbK-psbI* were evaluated.

Diverse genera of land plants were discriminated with the plastid multiloci by Fazekas *et al.* (2008), in which 92 species in 32 genera has been used. The plastid coding five loci *rpoB*, *rpoC1*, *rbcL* , *matK* and *23S rDNA* and three non-coding *trnH-psbA*, *atpF–atpH,* and *psbK–psbI* loci used. The regions differed in their ability to discriminate species and in ease of retrieval, in terms of amplification. In single locus, resolution ranged from 7% (23S rDNA) to 59% (*trnH-psbA*) of species, with well supported monophyly. Sequence recovery rates were related primarily to amplification success (85–100% for plastid loci), with *matK* requiring the greatest effort to achieve reasonable recovery (88% using 10 primer pairs). Several loci *matK*, *psbK–psbI* and *trnH-psbA* were problematic for generating fully bidirectional sequences.

The internal transcribed spacer 2 (*ITS2*) region of nuclear ribosomal DNA is valuable characteristics, such as the availability of conserved regions for designing universal primers, the ease of its amplification, and sufficient variability to distinguish even closely related species. Yao *et al.,* (2010) discriminated the plants and animals species with *ITS2* barcodes. In this study, 50,790 plants and 12,221 animals *ITS2* sequences were downloaded from GenBank and it were evaluated according to sequence length, GC content, intra- and inter-specific divergence, and efficiency of identification. The results show that the inter-specific divergence of congeneric species in plants and animals was greater than *ITS* corresponding intra-specific variations. The success rates for using the *ITS2* region to identify dicotyledons, monocotyledons, gymnosperms, ferns, mosses, and animals were 76.1, 74.2, 67.1, 88.1, 77.4, and 91.7 percent at the species level, respectively.

Across these groups, no single locus showed high levels of universality and resolvability. Interspecific sharing of sequences from individual loci was common and when multiple loci were combined, fewer barcodes were shared among species. Minor improvements on these were obtained by various new three-locus combinations involving *rpoC1*, *rbcL*, *matK* and *trnH-psbA*, but no single combination clearly outperformed all others (Yao *et al.*, 2010)

DNA barcoding study was employed in Panama forest to identify the species. Woody trees, shrubs, and palms found within the 50-ha of forest (296 species) has yielded >98 per cent correct identification. In this study, three loci *rbcL*, *matK*, and *trnH-psbA*, have identified the species at 75, 99, and 95 percent respectively.

The sixth largest monocot family Arecaceae contain about 2400 species which is difficult to resolve using only morphological characters and hence DNA barcoding was employed (Jeanson *et al.,* 2011). In this study, three plastid encoded (*matK*, *rbcL* and *psbA-trnH*) and one nuclear encoded (nr*ITS2*) were used. The combination of three loci *matK*, *rbcL* and *ITS2* gives 92 per cent species discrimination. The two core loci suggested by the CBOL-PWG, *rbcL* and *matK*, had a low species discrimination rate and hence be supplemented by another locus.

Liu *et al.* (2012) studied the sampling strategy and potential utility of InDels for DNA barcoding of closely related plant species in *Taxus*. In this study, 39 populations of nine *Taxus* species have been taken and DNA barcode *matK* and three supplementary regions *trnH-psbA*, *trn*L-*trn*F and *ITS* were employed. One individual per population was adequate to represent thc variation in the species and also found that sequence occurring in the chloroplast *trn*L-*trn*F and *trnH-psbA* regions were informative to differentiate among closely related taxa.

The chloroplast maturase K gene (*matK*) is one of the most variable coding genes of angiosperms and has been suggested to be a "barcode" for land plants. However, *matK* exhibits low amplification and sequencing rates due to low universality of currently available primers and mononucleotide repeats. To resolve these technical problems, and evaluated the entire *matK* region was evaluated to find a region of 600–800 bp that is highly variable and after careful evaluation, a region in the middle was chosen and a pair of primers were designed to amplify and sequence the *matK* fragment of approximately 776 bp. This region encompasses the most variable sites, represents the entire *matK* region best, and

also exhibits high amplification rates and quality of sequences. The universality of this primer pair was tested using 58 species from 47 families of angiosperm plants. The primers showed a strong amplification (93.1%) and sequencing (92.6%) successes in the species tested (Jing *et al.*, 2011)

DNA barcoding offers identification of a species even if only a small fragment of the organism at any stage of development is available. This could be of great utility in scrutinizing the illegal trade of both endangered plant and animal species. Therefore, DNA barcodes of Indian species of *Paphiopedilum* were developed along with their three natural hybrids, using loci from both the chloroplast and nuclear genomes (Parveen *et al.*, 2012). The five loci tested for their potential as effective barcodes were RNA polymerase-b subunit (*rpoB*), RNA polymerase-b' subunit (*rpoC1*), Rubisco large subunit (*rbcL*) and maturase K (*matK*) from the chloroplast genome and nuclear ribosomal internal transcribed spacer (nr*ITS*) from the nuclear genome. The *matK* with 0.9% average inter-specific divergence value yielded 100% species resolution thus could distinguish all the eight species of *Paphiopedilum* unequivocally.

An experiment was conducted to identify three species of Allium using five barcode regions including *ndhJ, rpoC1, rpoB, YCF5* and *rbcL* (Anvarkhah *et al.*, 2013). Barcode region of *YCF5* had the highest Species Rate of Sequencing Percentage (SRSP) with 93.33%. The highest interspecific diversity and intraspecific divergence with 0.8% and 0.4% were observed in region of *ndhJ* which was the most efficient barcode region for this species identification.

Phylogenetic relationships of different members of the family Cucurbitaceae were estimated from sequences of the internal transcribed spacer (*ITS1* and *ITS2*) regions of the nuclear ribosomal RNA genes (Jobst *et al.*, 1998). Twenty-six species of different genera belonging to different tribes and several subtribes were analyzed. The whole *ITS* regions were amplified by PCR technique and cloned, and three to five different clones of each species were sequenced.

*ITS*1 and *ITS2* regions are slightly variable in length, with each length appearing genus-specific.

The aquatic monocot family Lemnaceae or duckweeds represent the smallest and fastest growing flowering plants. Their highly reduced morphology and infrequent flowering result in a lack of characters for distinguishing between the nearly 38 species that exhibit these tiny, closely-related and often morphologically similar features within the same family of plants. Wang *et al.* (2010) developed a DNA barcoding system for the Lemnaceae. A Lemnaceae specific primers were designed for four coding plastid genes *rpoB*, *rpoC1*, *rbcL* and *matK* and three noncoding spacers *atpF-atpH, psbK-psbI* and *trnH-psbA*, based on the Lemnaceae minor chloroplast genome sequence and the sequences generated have successfully discriminated different species.

Bucklin *et al.* (2009) developed DNA barcoding in Arctic Ocean holozooplankton for species identification and recognition. A 700 bp region of the mitochondrial cytochromeoxidase I (*mtCOI*) gene was amplified and sequenced. Kimura-2-Parameter (K-2-P) distances, with 1000-fold boots trapping among individuals of the same species ranged from 0.0 to 0.2; distances between species ranged widely from 0.1 to 0.7.

Wide systematic range of Bryophyta (Mosses) has been evaluated by Liu *et al.* (2010), using eight previously proposed plant barcoding regions *atpF-atpH*, *ITS2*, *matK*, psbK-*psbI*, *rbcL* , *rpoB*, *rpoC1*, and *trnH-psbA* and two popular phylogenetic markers *rps4* and *trnL-trnF* of cpDNA. The *ITS2*, *rbcL* , *rpoC1*, *rps4*, *trnH-psbA* and *trnL*-trnF regions showed good universality, and therefore the efficacy of these loci as DNA barcodes was further evaluated in 36 mosses and 2 liverworts, each of which included two to three individuals per taxa. The five loci, *rbcL*, *rpoC1*, *rps4*, *trnH-psbA* and *trnL-trnF*, were easy to amplify and sequence and showed significant inter-specific genetic variability, making them potentially useful DNA barcodes for mosses.

Ebihara *et al.* (2010) described DNA barcodes for pteridophyte with two plastid DNA regions *rbcL* and *trnH-psbA*. DNA sequences were obtained from each of 689 (94.0%) taxa for *rbcL* and 617 (84.2%) taxa for *trnH-psbA*. Mean interspecific divergence values across all taxon did not reveal a significant difference in rate between *trnH-psbA* and *rbcL* , but mean distances of each genus showed significant heterogeneity according to systematic position.

NW-European Ferns were discriminated by DeGroot *et al.* (2011) by using two-loci DNA barcode, *rbcL* and *trnL-F*. A regional approach was adopted, by creating a reference database of trusted *rbcL* and *trnL-F* sequences for the wild-occurring homosporous ferns.

2.8.2 Bio security for medicinal plants

Herbal medicinal materials have been used worldwide for centuries to maintain health and to treat disease. However, adulteration of herbal medicines remains to be a major concern of users and industry, for the reasons of safety and efficacy. Identification of herbal medicinal materials by DNA technology has been widely applied from the mid 1990s. Recently Li *et al.* (2011) had given the DNA barcodes for identification of herbal medicinal materials using four standard loci *rbcL*, *matK*, *trnH–psbA* and *ITS*. These DNA barcodes can also be used as reliable tools to facilitate the identification of herbal medicinal materials for the safe use of herbs, quality control, and forensic investigation.

Chen *et al.* (2010) have compared seven candidate DNA barcoding loci *psbA-trnH*, *matK*, *rbcL*, *rpoC1*, ycf5, *ITS2*, and *ITS*. With respect to amplification efficiency, differential intra- and inter-specific divergences, and the DNA barcoding gap, the second internal transcribed spacer (*ITS2*) of nuclear ribosomal DNA was suggested as it represents the most suitable region.

Four barcoding loci *rbcL* , *matK*, *trnH-psbA*, and internal transcribed spacer (*ITS*) were used to test their ability to distinguish the possible contamination in *Tetrastigma* which is extensively used in Chinese medicine (Fu

et al., 2011). The results indicated that the best barcode was *ITS*, which showed significant inter-specific genetic variability. Multiple loci provided a greater ability to distinguish species than single loci and use of combined barcode *rbcL* +*matK*+*ITS* was recommended.

Liu *et al.* (2012) applied DNA barcodes for identification of plant species in the family Araliaceae using five DNA loci *matK*, *rbcL* , *ITS2*, *psbA-trnH* and *ycf5*. *ITS2* showed superiority in species discrimination with an accurate identification of 85.23 and 97.29 percent at the species and genus levels, respectively, in plant samples from the 589 sequences derived from Araliaceae.

To conclude, the phylogenetic relation in the genus *Momordica* are highly complicated and molecular technology to delineate the species relation in the genus is necessary. DNA barcoding in *Momordica* is not reported so far. The DNA barcoding in Cucurbitaceae is also rare. The universal primers for the common barcode loci have to be validated for using *Momordica*. The barcoding in animal is less implicated than that in plants, since the locus CO1 itself is found to be enough variation at species level; whereas, in plants various loci such as *matK*, *rbcL*, *trnH-psbA* and their combinations reported in different genus by different authors. Thus for arriving at species relation in this genus, the barcoding using the candidate loci using the universal primers has to be done and the specific barcoding primers for the loci has to be identified for *Momordica*.

Material and Methods

3. MATERIALS AND METHODS

The study on "DNA barcoding in *Momordica* spp." was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Kerala Agricultural University during the period 2012-2014. The materials used and methodologies adopted are being presented in this chapter.

3.1 Materials

3.1.1 Plant materials

A total of 25 accessions belonging to 7 species of *Momordica* present in India were used in this study. Additionally, 2 *Luffa* accession were also employed to confirm the species relation.

The *Momordica* species selected for this study were *Momordica charantia* cv. *charantia*, *Momordica charantia* cv. *muricata* (Wild) Chakrav, *Momordica dioica* Roxb., *Momordica sahyadrica* Joseph and Antony, *Momordica balsamina*, *Momordica cochinchinesis* Spreng (Gac), *Momordica subangulata* Blume ssp. *renigera* (G. Don) de wild and *Momordica cymbalaria* Hook (syn. *M. tubersa*). All species leaf samples were obtained from different locations like National Bureau of Plant Genetic Resources (NBPGR) Thrissur, Krishi Vighyan Kendra (Kerala Agricultural University) and College of Horticulture and Research Institute, Periyakulam (Tamil Nadu Agricultural University).

3.1.2 Laboratory chemicals, glass wares and plastic wares

The chemicals used in the study were of good quality (AR grade), procured from Merck India Ltd., HIMEDIA and SISCO Research Laboratories. The *Taq* DNA polymerase, dNTPs, Taq buffer and molecular markers were supplied by Bangalore Genei Ltd. All the plastic wares used were obtained from Axygen and Tarson India Ltd. Different primers were obtained from Sigma Aldrich Chemicals Pvt. Ltd.

3.1.3 Equipments and machinery

The present research work was carried out using molecular biology facilities and equipments available at CPBMB, College of Horticulture. Centrifugation was done in high speed refrigerated centrifuge Kubota 6500. NanoDrop^{R} ND-1000 spectrophotometer was used for the estimation of quality and quantity of DNA. The DNA amplification was carried out in Veriti thermal cycler (Applied Biosystem, USA) and SureCycler 8800 thermal cycler (Agilent Technology). Horizontal gel electrophoresis system (BIORAD, USA) was used for agarose gel electrophoresis. Gel Doc XR+ (BIORAD, USA) was used for imaging and documenting the agarose gel and the images were analysed using the Quantity One software.

3.2 METHODS

3.2.1 Morphological analysis

Morphological characters of *Momordica* spp. were recorded as per the minimal descriptors of developed by Joseph and Antony (2011). The characters observed were the following:

a) Plant growth habit
b) Leaf colour
c) Petiole colour
d) Leaf margin
e) Leaf shape
f) Leaf size
g) Leaf pubescence
h) Flower colour
i) Fruit skin colour (Immature)
j) Fruit surface
k) Fruit shape
l) Fruit size

3.2.2 Molecular analysis

Molecular analyses of *Momordica* spp. were carried out with different primers sets.

3.2.2.1 Genomic DNA isolation

Young tender, pale green leaves (first to third from the tip) were collected on ice from individual plants in each accession and brought to the laboratory. The surface was cleaned by washing with sterile water and wiping with 70 per cent ethanol. The fresh leaves were ground using ice-cold mortar and pestle to fine powder in liquid nitrogen along with β -mercaptoethanol and PVP. β -mercaptoethanol and PVP have phenol oxidase activity.

CTAB method detailed by Rogers and Bendich (1994) was used for the extraction of total genomic DNA. The reagents used for DNA isolation are presented in Annexure II.

Procedure:

- One gram of clean leaf tissue was ground in pre-chilled mortar and pestle in the presence of liquid nitrogen.
- 4ml of extraction buffer (2x), 50µl of β-mercaptoethanol and a pinch of Poly Vinyl Pyrolidone (PVP) were added to the mortar.
- The homogenized sample was transferred into an autoclaved 50ml centrifuge tube and 3ml of pre-warmed extraction buffer was added (total 7ml).
- The contents were mixed well and incubated at 65°C for 20 to 30 minutes with occasional mixing by gentle inversion.
- Equal volume (7ml) of chloroform: isoamyl alcohol (24:1) was added and mixed by inversion to emulsify. Spun at 10,000 rpm for 15 minutes at 4°C

- After centrifugation, the contents got separated into three distinct phases.

Aqueous topmost layer	-	DNA and RNA
Middle layer	-	fine particles and proteins
Lower layer	-	Chloroform, pigments and cell debris

- The top aqueous layer was transferred to a clean centrifuge tube and $1/10^{th}$ volume of 10 per cent CTAB solution and equal volume of chloroform: isoamyl alcohol (24:1) were added and mixed by gently inversion and subsequently centrifuged at 10,000 rpm for 15 minutes at 4°C.

- The aqueous phase was transferred into a clean centrifuge tube and 0.6 volume of chilled isopropanol was added and mixed by quick gentle inversion till the DNA got precipitated. The contents were kept at -20°C for half an hour for complete precipitation.

- After incubation, tubes were again centrifuged at 10,000 rpm for 15 minutes at 4°C and the supernatant was gently poured off.

- The DNA pellet was washed with 70 per cent ethanol followed by 100 per cent ethanol and subsequently spun for 5 min at 10,000 rpm and the ethanol was decanted.

- The pellet was air dried, dissolved in 50µl of sterilized water and stored at -20°C.

3.2.2.2 Assessing the quality of DNA by electrophoresis

The quality of the isolated DNA was evaluated through agarose gel electrophoresis on 0.8 % agarose gel (Sambrook *et al.*, 1989).

Reagents and equipments

The reagent and equipment used in electrophoresis (AGE) are presented in Annexure III.

Procedure for gel casting and AGE

- For casting the gel, the gel tray was prepared by sealing the ends with tape. Comb was placed in gel tray about 1 inch from one end of the tray and positioned the comb vertically such that the teeth are about 1 to 2 mm above the surface of the tray.

- Prepared 0.8 per cent agarose in a glass conical flask by dissolving 0.8g agarose in 100 ml 1X TAE buffer (2ml from 50X TAE buffer stock was made up to 100ml). The contents were heated for 45 to 60 seconds until agarose was dissolved and solution was clear.

- Solution was allowed to cool to 42 to 45^0C under room condition and at this point 4µl ethidium bromide was added to a concentration of 10µl/ml and mixed well.

- This warm gel solution was poured into the tray to a depth of about 5 mm and the gel was allowed to solidify for about 30 to 45 minutes at room temperature.

- After the expiry of time, the comb and the tape used for sealing the gel tray along with the gel were gently removed and the tray was placed in electrophoresis chamber, and covered (just until wells are submerged) with 1X electrophoresis buffer.

- Samples for electrophoresis were prepared by adding 1 µl of 6X gel loading dye for every 5µl of DNA sample and by mixing them well. Loaded 6µl DNA sample per well. The molecular weight marker was loaded (λDNA *Eco*RI/ *Hind*III double digest) in first lane.

- The Electrophoresis was carried at 70 volts until dye has migrated two third of the length of the gel.

- The gel profile was examined for intactness, clarity of DNA band, presence of RNA and protein. The intact DNA has appeared as orange fluorescent bands and the degraded one appears as a smear, because of the presence of a large number of bands, which differed in few base length. The presence of protein was observed as a thick white patches which got trapped in well itself. The RNA was observed as a thick band with size less than 100bp.

3.2.2.3 Gel documentation

Gel documentation was done with BioRad Gel Documentation System using PDQuest™ software. PDQuest is a software package for imaging, analyzing, and databasing the electrophoresed gels. The gel containing DNA was viewed under UV transilluminator for presence of DNA. The DNA fluoresces under UV light due to ethidium bromide dye. The image of a gel is captured using the PDQuest controls in the imaging device window and subjected to further analyses.

3.2.2.4 Assessing the quality and quantity of DNA using spectrophotometer (NanoDrop ND-1000)

The purity of DNA was further checked using NanoDrop ND-1000 spectrophometer. Nucleic acid shows absorption maxima at 260 nm whereas proteins show peak absorbance at 280nm. Absorbance has been recorded at both wavelengths and the purity was indicated by the ratio OD_{260}/OD_{280}. A value 1.8 indicated that the DNA is pure and free from proteins and RNA. When the ratio was <1.8 it meant that the sample is RNA contaminated and >1.8 had shown that the sample is protein contaminated. The quantity of DNA in the pure sample was calculated using the relation.

1 OD at 260 nm = 50 µg DNA/ml

Therefore OD_{260} X 50 gives the quantity of DNA in µg/ml.

Procedure

- The Nanodrop spectrophotometer was connected to the computer installed with ND-1000 software.
- The option 'Nucleic acid' was selected in the software.
- With the sampling arm open, pipetted 1µl distilled water onto the lower measurement pedestal.
- The sampling arm was closed and spectral measurement was initiated using the operating software. The sample column was automatically drawn between the upper and lower measurement pedestals and the spectral measurement was made.
- The reading was set to zero with sample blank.
- Subsequently 1µl of sample was pipette out onto the measurement pedestal and selected the 'Measure' option.
- When the measurement has been completed, the sampling arm was opened and the sample was wiped from both upper and lower pedestals using a soft laboratory tissue paper. Simple wiping has prevented the sample carryover in successive measurements for the samples varying by more than 1000 fold in concentration.

3.2.2.5 Purification of DNA

The DNA which had RNA as contaminant (as observed from the electrophoresis) was purified by RNase treatment and subsequent precipitation (Sambrook *et al.,* 1989).

Reagents used for RNase treatment of DNA were

I. Phenol: chloroform mixture (24:1 v/v)
II. Chilled isopropanol (100%)

III. Ethanol (70%)

IV. 50 X TAE buffer

V. Chloroform: Isoamyl alcohol (24:1 v/v)

VI. RNase (Sigma, USA) at 1:9 ratio (1 µl RNase and 9 µl water)

Procedure

- For 50 µl DNA sample, 5 µl of diluted RNase was used and incubated at 37^0C in dry bath for 40 minutes.
- The total volume was made up to 250 µl with distilled water.
- Equal volume of chloroform: isoamyl alcohol (24: 1) mixture was added and mixed gently.
- Content were centrifuged at 12000 rpm for 15 minutes at 4^0C.
- The aqueous phase was transferred into a fresh micro centrifuge tube and equal volume of chloroform: isoamyl alcohol (24: 1) was added.
- Centrifuged at 12,000 rpm for 15 minutes at 4^0C. The aqueous phase was transferred into a clean centrifuge tube and 0.6 volume of chilled isopropanol was added and mixed by 2-3 gentle inversion till the DNA precipitated. For complete precipitation samples were kept at -20°C for half an hour.
- The mixture was incubated at 20^0C for 30 minutes and centrifuged at 10,000 rpm for 15 minutes at 4^0C.
- DNA pellet was washed with 70 per cent ethanol by subsequent centrifugation at 10,000 rpm for 10 minutes at 4^0C.
- The pellet was air dried and dissolved in 50 µl sterile distilled water.
- The samples were loaded on 0.8 per cent agarose gel at constant voltage of 70V to test the quality and to find whether there is any shearing during RNAse treatment.

3.2.3 DNA barcoding in *Momordica*

Primer sets for 3 common barcoding loci were used for the initial screening to assess their compatibility to generate the banding patterns in various *Momordica* accession. These loci were *matK* (Maturase K), *ITS2* (Internal transcribed spacer) and *trnH-psbA* intergenic spacer. Amplification of DNA from all the lines of *Momordica* spp. was attempted separately with all the selected primers.

3.2.3.1 DNA amplification conditions

The PCR conditions required for effective amplification include appropriate proportions of components of the reaction mixture. Including template DNA, assay buffer A, Taq DNA polymerase, dNTPs and primers. The aliquot of this master mix were dispensed into 0.2 ml PCR tubes. The thermal cycling was carried out in Veriti Thermal Cycler (Applied Biosystems, USA).

Another important factor, which affected the amplification rate was the temperature profile of thermal cycles. The thermocycler was programmed for desired number of cycles and temperatures for denaturation, annealing and polymerization.

Thermal cycling – Composition and thermal profile

Total genomic DNA was diluted up to 25-30 ng/µl and subjected to PCR. DNA barcoding primers which are already reported used for amplification of DNA (Table 1).

PCR amplification was performed in a 20 µl reaction mixture and the composition of the reaction mixture consisted of,

a) Genomic DNA (30 ng)	- 1 µl
b) 10X Taq assay buffer A	- 2 µl
c) dNTP mix (10 mM each)	- 1.5 µl
d) Taq DNA polymerase (3U)	- 0.3 µl
e) Primer (10 pM)	- 0.75 µl
f) Autoclaved distilled water	- 13.7 µl
Total volume	- 20.0 µl

The PCR amplification was carried out with the following thermal profile. (Protocol issued by CBOL, Ivanova *et al.*, 2006)

Initial denaturation - 94^{0}C for 1 minute

Denaturation - 94^{0}C for 30 second

Primer annealing - Primer annealing temperature as per Table 2 for 40 sec } 40 Cycles

Primer extension - 72^{0}C for 40 second

Final extension - 72^{0}C for 5 minutes

4^{0}C for infinity to hold the sample

3.2.3.2 Screening of primers and analysis

Since, so far no DNA barcoding results are published in Cucurbits, the universal primers recommended by previous researchers were initially tried in this study. Eight primer sets, of which 6 for *matK* and 1 for *ITS2* and 1 for *trnH-psbA* were initially used for amplifying the barcode loci in *Momordica*. Based on the initial studies it was found that *matK* is giving a better amplification response and hence it was decided to choose *matK* as the candidate loci for this study. The list of primers used in the initial screening is presented in Table 1. Subsequently, 18 primer combinations using the forward and reverse primers for *matK* loci were

attempted to amplify the loci, from 25 *Momordica* and 2 *Luffa* accessions (Table 2).

The amplified products were electrophoresed on 1.5 per cent agarose gel using 1X TAE buffer stained with ethidium bromide. The profile was visualized under UV (312 nm) transilluminator (BioRad) and documented for further analysis. The documented profiles were carefully examined for amplification of bands in comparison with 100 bp- 3 kb ladder. The average size of the expected band in all these combination of primer was around 900 bp.

Table 1: List and sequence of DNA barcoding primers used in this study

Sl No.	Primer	Nucleotide Sequences	Reference
1	*matK* F1	CCTATCCATCTGGAAATCTTAG	Jarret, 2008 ; Dunning and Savolainen, 2010
	matK R1	GTTCTAGCACAAGAAAGTCG	
2	*matK* F2	ATCCATCTGGAAATCTTAGTTC	
	matK R2	CTTCCTCTGTAAAGAATTC	
3	*matK* F3	CGATCTATTCATTCAATATTTC	
	matK R3	TCTAGCACACGAAAGTCGAAGT	Lahaye *et al.*, 2007
4	*matK* F4	(T)AATTTACGATCAATTCATTC	Jarret, 2008 ; Dunning and Savolainen, 2010
	matK R4	TCTAGCACACGAGTCGAAGT	Lahaye *et al.*, 2007
5	*matK* F5	CGATCTATTCATTCAATATTTC	Jing *et al.*, 2011
	matK R5	GTTCTAGCACAAGAAAGTCG	
6	*matK* F6	ACCCAGTCCATCTGGAAATCTTGGTTC	Jarret, 2008 Dunning and Savolainen, 2010
	matK R6	CGTACAGTACTTTTGTGTTTACGAG	
7	*ITS2* F	GCGATACTTGGTGTGAAT	Jing *et al.*, 2011
	ITS2 R	GACGCTTCTCCAGACTACAAT	
8	*trnH-psbA* F	CGCGCATGGTGGATTCACAATCC	Jarret, 2008
	trnH-psbA R	GTTATGCATGAACGTAATGCT C	

Table 2: List of combination of forward and reverse primer used for amplifying *matK* loci of *Momordica*

Sl. No.	Primer	Annealing Temp. (^{0}C)	Combination
1	*matK* 1	51.7	*matK* F1
			matK R1
2	*matK* 2	45.2	*matK* F2
			matK R2
3	*matK* 3	49.5	*matK* F3
			matK R3
4	*matK* 4	45.2	*matK* F1
			matK R2
5	*matK* 5	53.2	*matK* F1
			matK R3
6	*matK* 6	51.7	*matK* F2
			matK R1
7	*matK* 7	52.7	*matK* F2
			matK R3
8	*matK* 8	49.5	*matK* F3
			matK R1
9	*matK* 9	45.2	*matK* F3
			matK R2
10	*matK* 10	48.4	*matK* F4
			matK R4
11	*matK* 11	49.5	*matK* F5
			matK R5
12	*matK* 12	55.6	*matK* F6
			matK R6
13	*matK* 13	48.4	*matK* F4
			matK R5

14	*matK* 14	48.4	*matK* F4
			matK R6
15	*matK* 15	49.5	*matK* F5
			matK R4
16	*matK* 16	49.5	*matK* F5
			matK R6
17	*matK* 17	57.9	*matK* 6
			matK R4
18	*matK* 18	51.7	*matK* F6
			matK R5

3.2.3.3 DNA extraction from agarose gel (NucleoSpun® Gel and PCR Clean-up kit protocol)

The specific primers used in the DNA barcoding are supported to yield only one band, on PCR assay. The band at the expected size was used to sequencing

(i) Excising DNA fragment

A clean scalpel was used to excise the DNA fragment from an agarose gel. Excess agarose was removed. Determined the weight of the gel slice and transfered it to a clean tube. For each 100 mg of agarose gel (< 2 %), 200 µl Buffer NTI was added. For gels containing > 2 % agarose, double the volume of Buffer NTI. Samples were incubated for 5–10 min at 50 °C. The samples were vortexed briefly, every 2–3 min until the gel slice was completely dissolved.

(ii) Binding the DNA to column

Placed a NucleoSpun® Gel and PCR Clean-up Column into a Collection Tube (2 mL) and loaded up to 700 µl sample. Centrifuged for 30 s at 11,000 Xg. Discarded the flow-through and placed the column back into the

collection tube. Loaded the remaining sample and repeated the centrifugation step.

(iii) Washing silica membrane

700 µl Buffer NT3 was added to the NucleoSpun® Gel and PCR Clean-up Column. Centrifuged for 30 s at 11,000X g. Discarded flow-through and placed the column back into the collection tube.

(iv) Drying silica membrane

The silica member was centrifuged for 1 min at 11,000X g to remove Buffer NT3 completely. While removing it from the centrifuge and the collection tube it was made sure the spun column does not come in contact with the flow-through. Total removel of ethanol was achieved by incubating the columns for 2–5 min at 70 °C prior to elution.

(v) DNA elution

Placed the NucleoSpun® Gel and PCR Clean-up Column into a new 1.5 mL micro centrifuge tube. Added 15–30 µl Buffer NE and incubated at room temperature (18-25°C) for 1 min. Centrifuged for 1 min at 11,000 X g.

3.2.3.4 Confirmation of single band after reamplification of eluted DNA

The specific primers are supposed to yield only one band at the expected product size. The direct sequencing of DNA could be done only if the band conforming product from a single genomic region. To confirm single genomic locus, the eluted DNA was subjected for one more PCR. After confirmation, the DNA eluted from the expected band was used for sequencing.

3.2.3.5 PCR product sequencing

A total of 25 sequences of the *Momordica* species and 2 *Luffa acutangula* were amplified with the selected primers, generated band were eluted, cleaned by reamplification and the products giving single band on PCR assay were used for sequencing.

If multiple bands were obtained after amplification, only the band having molecular weight nearest to that of the targeted region was eluted and purified. Sequencing was done by outsourcing (SciGenome Lab. Pvt. Ltd., Cochin).

3.2.3.6 Submission to NCBI

A total of 25 sequences of the *Momordica* species and 2 *Luffa* have been submitted to the NCBI GenBank. The procedure involved was

(i) Logged in to the password account in MyNCBI at http://www.ncbi.nlm.nih.gov/guide/howto/submit-sequence-data/ website
(ii) Sequence with simple annotation was submitted through BankIt.
(iii) Additional information like date for public release (immediate or at a specified future date), basic information (authors and a working title) for a corresponding reference paper, name of the organism or plant from which the sequence data were isolated and any other related descriptive data were also be provided.
(iv) The allocation of the accession number is awaited from NCBI

3.2.4 Data analysis

3.2.4.1 Species identification based on phylogeny analysis

The sequences were used to construct the pylogenetic tree of the accession under study using the software clustalW Omega (http://www.ebi.ac.uk/clustalw). The sequence in FASTA format were pasted in the interaction box in the webpage

The software has generated the phylogenetic tree, similarity matrix, sequence alignment etc. for the sequence under study.

3.2.4.2 Genetic divergence within and between species

The inter specific distances were calculated using K2P (Kimura 2 parameter) model as recommended by the Consortium of Barcode of Life (CBOL, http:// www.barcoding.si.edu/protocols.html) using MEGA 4.0 software.

4.0 software available at http://www.megasoftware.net/mega4/mega.html was used for the analysis. The 'Distance|Compute Pair-wise' command generates the pair wise distances.

3.2.4.3 Barcoding gap assessment

The distributions of intraversus inter-specific variability were compared using DNA barcoding gap. All the 27 sequences have aligned with the software ClutalW and with the mark like "*" had studied. Each gap has been selected which were unique in each species.

3.2.4.4 Identification efficiency of the DNA barcodes

The species identification success rate was calculated using genetic distance and BLAST methods at http://blast.ncbi.nlm.nih.gov/Blast.cgi, using all *matK* sequences as query sequences.

Graphical display had shown where the query is similar to the sequence, hit list, sequence accession number and the name, description, the bit score, the e-value (the expectation value), the alignment and length.

3.2.4.5 Identification of Conserved Region

The conserved sequence in the generated sequence was identified using the Software ClustalW.

3.2.4.6 Designing Primer

The primer for the genus *Momordica* was designed using Primer3 software. Care was taken to design the forward and reverse primers to yield the product size of approximately 1 Kb.

3.2.5 Submission to Barcode of Life Data system

Barcode data generated was further submitted to BOLD http://www.boldsystems.org, after furnishing the necessary information such as project title, project code and selected locus as primary marker.

Results

4. RESULTS

The study on DNA barcoding in *Momordica* species has been done at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Kerala Agricultural University. The objective of the study was to develop a DNA barcode which can easily differentiate the seven Indian species with each other. The results of various aspects of the investigations are presented in this chapter.

4.1 Morphological characterisation

Morphological characters were recorded from the fully grown up plants at fruiting stage. The plants available at NBPGR Thrissur and KVK Thrissur were directly used for recording. Whereas the characters of *M. cymbalaria* were recorded as given from Horticulture College and Research Institute, Periyakulam. Morphological parameters observed for the different species are listed in Table 3. Among the species the main morphological variations were observed in fruit shape (Plate 1).

4.1.1 Characters

1. There are three type of growth habit in *Momordica* genotype *viz.*, short viny, medium viny and long viny. Most of the genotypes had intermediate growth habit.
2. In petiole, only one color was observed among all the genotypes.
3. Flowers varied in color from white to yellow with the dark color inside.
4. Two main leaf colors observed were light green and dark green.
5. Regarding fruit shape all genotypes varied widely with their surface ridges and spines. The fruits in *charantia* pendulous, ovoid and soft tubeculuate with broken or continous ridges. As in *balsamina* fruit was bulged in middle, in other species soft spines were present except in *cymbalaria* which had clear ridges.

Table 3: Morphological characters of different *Momordica* species used for DNA barcoding

Morphological characters	*M. charantia* Var. Prrethi	*M. charantia* Var. Kurupantara	*M. charantia* Var. JNM 7	*M.charantia* Var. Vadakara	*M. charantia* Var. V53	*M. charantia muricata*	*M. dioica* Kerala 1	*M. dioica* Kerala 1	*M. dioica* Kerala 1
Plant growth habit	Medium viny	Medium viny	Medium viny	Medium viny	Medium viny	Medium viny	Long viny	Long viny	Long viny
Leaf colour	Green	Green	Green	Green	Green	Green	Light Green	Light Green	Light Green
Petiole colour	Light green	Light green	Light green	Light green	Light green	Light green	Green	Green	Green
Leaf margin	Entire	Entire	Entire	Entire	Entire	Entire	Entire, undulated	Entire, undulated	Entire, undulated
Leaf shape	Trianular-Lobed	Trianular-Lobed	Trianular-Lobed	Trianular-Lobed	Trianular-Lobed	Trianular-Lobed	Triangular - ovate cordate	Triangular - ovate cordate	Triangular - ovate cordate
Leaf Size	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
Leaf pubescence	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
Flower colour	White yellowish	White yellowish	White yellowish	White yellowish	White yellowish	White yellowish	Whitish yellow	Whitish yellow	Whitish yellow
Fruit skin colour	Light green	Light green	Light green	Light green	Light green	Light green	Dark green	Dark green	Dark green
Fruit surface	Clearly Ridges	Clearly Ridges	Clearly Ridges	Clearly Ridges	Clearly Ridges	Clearly Ridges	Clearly ridged and soft short spines	Clearly ridged and soft short spines	Clearly ridged and soft short spines
Fruit shape	Ellipsoid	Ellipsoid	Ellipsoid	Ellipsoid	Ellipsoid	Short Ellipsoid	Ovoid oblong	Ovoid oblong	Ovoid oblong
Fruit size	15 to 20 cm	15 to 20 cm	15 to 20 cm	15 to 20 cm	15 to 20 cm	8 to 10 cm	4 to 5 cm	4 to 5 cm	4 to 5 cm

Morphological characters	*M. dioica* Kerala 4	*M. dioica* Odisa	Wild 2 (*M. charantia cv. muricata*),	Wild 1 (*M. sahyadrica*)	Wild 2 (*M. sahyadrica*)	*M. sahyadrica* Annamalai type	*M. balsamina*	*M. cochinchinensis* Andaman	*M. cochinchinensis* Northeast
Plant growth habit	Long viny	Long viny	Medium viny	Long viny	Long viny	Long viny	Medium viny	Long viny	Long viny
Leaf colour	Light Green	Light Green	Green	Green	Green	Green	Light green	Dark green	Dark green
Petiole colour	Green	Green	Light green	Dark green	Dark green	Dark green	Green	Green	Green
Leaf margin	Entire, undulated	Entire, undulated	Entire	Variable entire, undulate	Variable – entire, undulate	Variable – entire, undulate	Entire	Entire, dentate	Entire, dentate
Leaf shape	Triangular - ovate cordate	Triangular - ovate cordate	Trianular-Lobed	Ovate, triangular	Ovate, triangular	Ovate, triangular	Angular	Ovate , sub orbicular outline	Ovate , sub orbicular outline
Leaf Size	Medium	Medium	Medium	Medium	Medium	Medium	Small	Medium	Medium
Leaf pubescence	Medium	Medium	Medium	Sparse	Sparse	Sparse	Glabrous	Medium	Medium
Flower colour	Whitish yellow	Whitish yellow	White yellowish	Yellowish white purpal at base	Yellowish white purpal at base	Yellowish white purpal at base	Whitish yellow	Yellowish blackish inside	Yellowish blackish inside
Fruit skin colour	Dark green	Dark green	Light green	Dark green	Dark green	Dark green	Green	Dark green	Dark green
Fruit surface	Clearly ridged with soft short spines	Clearly ridged with soft short spines	Clearly Ridges	Soft short spins	Soft short spins	Soft short spins	Obscurely ridged (feeble)	Soft spines all over fruit	Soft spines all over fruit
Fruit shape	Ovoid oblong	Ovoid oblong	Short Ellipsoid	Ellipsoid oblong	Ellipsoid oblong	Ellipsoid oblong	Ovoid-ellipsoid	Ovoid	Ovoid
Fruit size	4 to 5 cm	4 to 5 cm	8 to 10 cm	5 to 7.5 cm	5 to 7.5 cm	5 to 7.5 cm	4 cm	5 to 6 cm	5 to 6 cm

Morphological characters	*M. subangulata* Arka Gurav	*M. subangulata* 1	*M. subangulata* 2	*M. subangulata* 3	*M. subangulata* 4	*M. subangulata renigera*	*M. cymbalaria*
Plant growth habit	Long viny	Long viny	Long viny	Long viny	Long viny	Long viny	Short Viny
Leaf colour	Light green	Light green	Light green	Light green	Light green	Light green	Dark green
Petiole colour	Green	Green	Green	Green	Green	Green	Green
Leaf margin	Undulate , coarsely denticulate with fine bristle	Undulate , coarsely denticulate with fine bristle	Undulate , coarsely denticulate with fine bristle	Undulate , coarsely denticulate with fine bristle	Undulate , coarsely denticulate with fine bristle	Undulate , coarsely denticulate with fine bristle	Entire
Leaf shape	Ovate cordate, unlobed	Ovate cordate, unlobed	Ovate cordate, unlobed	Ovate cordate, unlobed	Ovate cordate, unlobed	Ovate cordate, unlobed	Ovate, Round
Leaf size	Large	Large	Large	Large	Large	Large	Small
Leaf pubescence	Glabrous	Glabrous	Glabrous	Glabrous	Glabrous	Glabrous	Glabrous
Flower colour	Cremish yellow	Cremish yellow	Cremish yellow	Cremish yellow	Cremish yellow	Cremish yellow	Yellow
Fruit skin colour	Light green	Light green	Light green	Light green	Light green	Light green	Dark green
Fruit surface	Soft echinate, ridges at base	Soft echinate, ridges at base	Soft echinate, ridges at base	Soft echinate, ridges at base	Soft echinate, ridges at base	Irregular ridges	Clearly soft ridges
Fruit shape	Ovoid ellipsoid	Ovoid ellipsoid	Ovoid ellipsoid	Ovoid ellipsoid	Ovoid ellipsoid	Ovoid ellipsoid	Short ellipsoid
Fruit size	3 to 5 cm	3 to 5 cm	3 to 5 cm	3 to 5 cm	3 to 5 cm	3 to 5 cm	3 to 4 cm

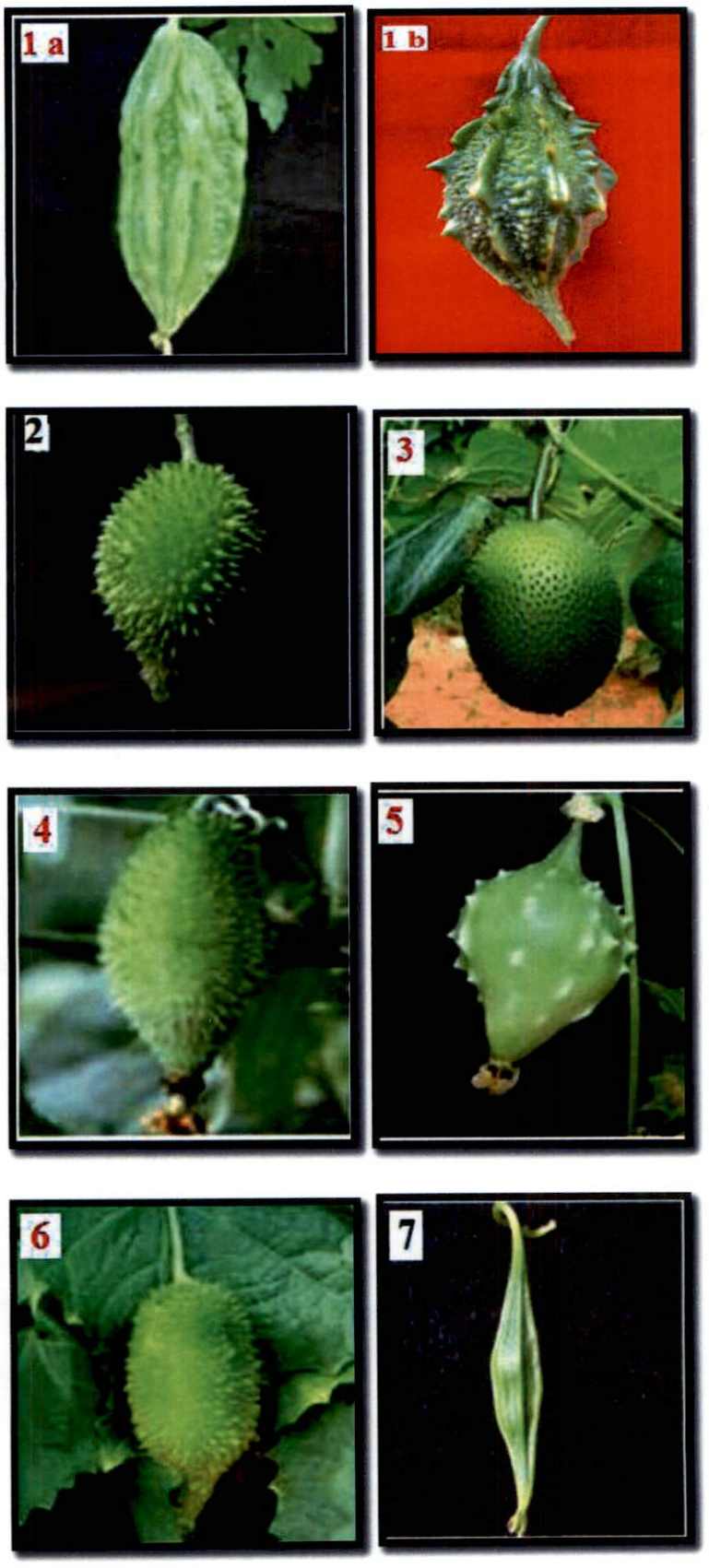


Plate 1: Morphological characters of fruit **1a**- *M. charantia*, **1b**- *M. charantia* var. *muricata*, **2**- *M. dioica*, **3**- *M. cochinchinensis*, **4**- *M. sahyadrica*, **5**- *M. balsamina*, **6**- *M. subangulata*, **7**-*M. cymbalaria*

4.2 DNA barcoding

4.2.1 DNA isolation

Very tender, fresh, pale green leaves (0.5 to 1g) were found to yielded good quality DNA in sufficient quantity. Since *Momordica* is incomparativly poor in phenols, browning was not a problem and use of antioxidant such as PVP was not neccessory.

Genomic DNA was isolated using the CTAB method (Rogers and Bendich, 1994). The isolated DNA was found contaminated with RNA (Fig. 1). RNase treatment followed by isopropanol precipitation has yielded sufficient quantity of good quality DNA. The subsequent agarose gel electrophoresis has indicated clear discrete bands with no RNA contamination (Fig.2) and spectrophotometric analysis gave ratio of UV absorbance ratio ($A_{260/280}$) around 1.8 (Table 4).

4.2.2 Screening of primers for amplifying the barcode loci

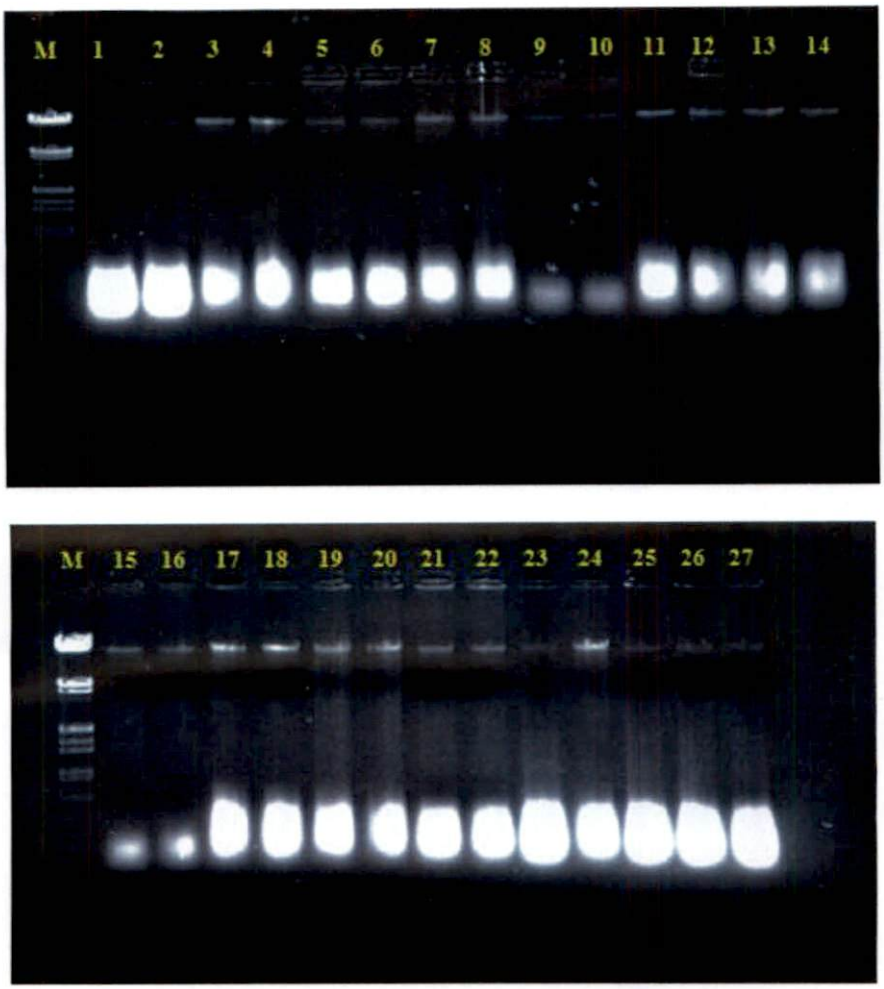
Good quality total genomic DNA was used for the amplification of the barcode loci, using the previously reported universal primers. Total genomic DNA was used since this contain nuclear, mitochondrial and chloroplast DNAs.

Twenty DNA barcoding primers reported for chloroplast *matK*, nuclear internal transcribed spacer (*ITS*2) and *trnH-psbA* spacer were tested for amplification.

The PCR assay using the standardized thermal settings gave satisfactory results. Though the selected primer were reported to be universal, the amplification pattern obtained was highly varying in each primer sets.

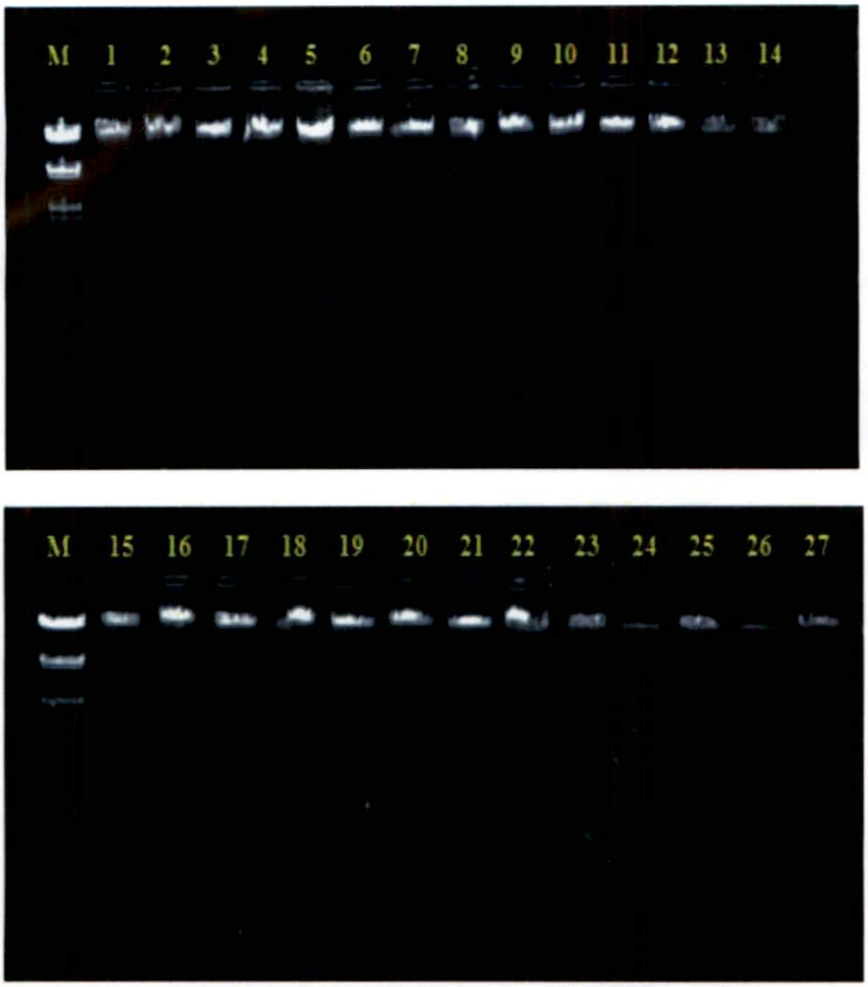
4.2.3 Amplification of barcode loci with selected primers

The universal barccoding primers used in this study has mostly amplified a clear distinct band at the size range of 800 to 1000 bp.



M- Lamda/Hind III DNA Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa,**13**- Wild 2 (*M. charantia* cv. *muricata*),**14**- Wild 3 *(M. sahyadrica)*, **15**- Wild 1 *(M. sahyadrica)*, **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman ,**19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5, **25**- *M. cymbalaria,* **26**- *Luffa acutangula* (Haritam), **27**- *Luffa acutangula* (Arka Sumit)

Fig. 1 : DNA isolated from tender leaves of *Momordica* genotypes



M- Lamda/Hind III DNA Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa,**13**- Wild 2 (*M. charantia* cv. *muricata*),**14**- Wild 1 *(M. sahyadrica)*, **15**- Wild 2 *(M. sahyadrica)*, **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina* L., **18**- *M. cochinchinensis* Andaman ,**19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5, **25**- *M. cymbalaria,* **26**- *Luffa acutangula* (Haritam), **27**- *Luffa acutangula* (Arka Sumit)

Fig. 2 : DNA isolated from tender leaves of *Momordica* genotypes after RNase treatment

Table 4: Quality and quantity of DNA isolated from *Momordica* species as assessed by NanoDrop spectrophotometer method

Sl. No.	Species Sample	UV absorbance at 260 nm (A_{260})	UV absorbance at 280 nm (A_{280})	$A_{260}/_{280}$	Quantity (ng/µl)
1	*M. charantia* var. Preethi	51.28	28.70	1.79	2574
2	*M. charantia* var. Kurupantara	50.28	27.36	1.88	576
3	*M. charantia* var. Vadakara	12.33	6.54	1.80	616
4	*M. charantia* var. V53	85.23	42.22	1.76	4261
5	*M. charantia* var. JNM7	9.0	4.7	1.9	452
6	*M. subangulata renigera*1	43.979	21.75	1.9	2198
7	*M. subangulata renigera* 2	57.86	28.09	1.89	2893
8	*M. subangulata renigera* 3	60.11	30.16	1.99	3005
9	*M. subangulata renigera* 4	50.45	25.20	1.82	2522
10	*M. subangulata renigera* 5	30.88	15.00	1.82	1544
11	*M. subangulata renigera* var. Arka Gurav	22.52	10.65	1.77	1126
12	*M. c. muricata*	11.18	6.19	1.80	559
13	Wild 2 similar to *M. c. muricata*	39.21	19.11	1.89	1960
14	*M. dioica* KL 1	10.381	5.2	1.87	519
15	*M. dioica* KL 2	8.320	4.546	1.83	416
16	*M. dioica* KL 3	9.27	4.8	1.93	463
17	*M. dioica* KL 4	9.012	4.97	1.81	450
18	*M. dioica* Odisha	11.30	6.46	1.75	565

19	Wild 3 similar to *M. sahyadrica*	30.05	14.62	1.79	1502
20	Wild 1 similar to *M. sahyadrica*	58.24	29.76	1.89	2912
21	*M. sahyadrica* Annamalai type	11.88	5.007	1.87	94
22	*M. balsamina*	83.45	48.73	1.7	4172
23	*M. cochinchinensis* NE	84.34	48.59	1.74	4217
24	*M. cochinchinensis* Andaman	37.57	18.84	1.99	1878
25	*M. cymbalaria*	57.7	28.96	1.88	2889
26	*Luffa acutangula* var. Haritam	5.22	2.8	1.82	260
27	*Luffa acutangula* var. Arka Sumit	9.5	4.9	1.94	475

I. *matK* 1 (F1-R1)

The agarose gel profile for the DNA amplification pattern observed in the selected 25 *Momordica* and 2 *Luffa* genotypes with primer *matK* 1 (F1-R1) is presented in Fig 3. *matK* loci of 3 genotypes were amplified Wild 2 (*M. charantia* cv. *muricata*), Wild 1 *(M. sahyadrica)* and *M. subangulata*. The primer *matK* 1 was shown multiple bands and the amplicons ranged in size from 1200bp to 100 bp. When multiple bands were obtained after amplification, only the band having molecular weight nearest to the target region (1000) bp was eluted and purified using a DNA purification kit.

II. *matK* 2 (F2-R2)

A total of 7 genotypes were amplified *M. dioica* Kerala 4, *M. dioica* Odisa, *M. cochinchinensis* Andaman, *M. cochinchinensis* Northeast, *M. subangulata* 3, *M. subangulata* 4 and *M. subangulata* 5 with *matK* 2 (F2-R2) primer. The primer *matK* 2 had given multiple bands and the amplicons ranged in size from 1200bp to 100 bp. Among seven *M. dioica* Kerala 4, *M. dioica* Odisa, *M. cochinchinensis* Northeast, *M. subangulata* 4 and *M. subangulata* 5 shown distinct single band. When multiple bands were obtained after amplification, only the band having molecular weight nearest to the targeted region 1000 bp was selected and purified using a DNA purification kit (Fig. 4).

III. *matK* 3 (F3-R3)

Amplification with this primer was obtained in Wild 2 (*M. charantia* cv. *muricata*) only. Thie primer had shown two bands at 1000bp and 800bp. The 1000bp band has been purified and used for sequencing (Fig. 5).

IV. *matK* 4 (F1-R2)

A total of 4 genotypes have been amplified JNM 7, Vadakara, Arka Gaurav and *M. dioica* Kerala 2. The band with the molecular weight of 1000 bp was purified using a DNA purification kit and sequenced (Fig. 6).

V. *matK* 6 (F2-R1)

A total of 6 genotypes have been amplified using the F2-R1 combination *M. balsamina*, *M. cochinchinensis* Andaman, *M. cochinchinensis* Northeast and *M. subangulata*. 1 were shown single distinct band of 1000 bp range and Vadakara and Wild 2 *(M. sahyadrica)* showed multiple band. However in multiple bands 1000 bp band has been selected, eluted and sequenced (Fig.7).

VI. *matK* 7 (F2-R3)

A total of 3 genotypes have been amplified and they were Arka Gaurav, *M. dioica* Kerala 1 and Wild 2 (*M. charantia* cv. *muricata*). The primer *matK* 7 had shown only one distinct band at 1000 bp and these band was purified using a DNA purification kit and used for sequencing (Fig.8).

VII. *matK* 9 (F3- R2)

A total of 4 genotypes have been amplified using the primer combination F3-R2 and they were V 53, *M. dioica* Kerala 2, *M. dioica* Kerala 3 and *M. dioica* Kerala 4. The primer *matK* 9 was showed only one distinct band at 1000 bp ranges in all four amplified genotype. The band having molecular weight 1000 bp was purified using a DNA purification kit and used for sequencing (Fig.9).

VIII. *matK* 10 (F4-R4)

The agarose gel profile for the amplification pattern were observed in the all 25 *Momordica* genotype with primer *matK* 10 (F4-R4) are shown in Fig 10. This primer was shown amplification only in *M. charantia muricata* genotype. The primer was showed only one distinct band at 1000 bp ranges. The band having molecular weight 1000 bp was purified using a DNA purification kit (Fig. 10)

IX. *matK* 13 (F4-R5)

The two genotypes have been amplified using this primer combination and they were Preethi and Kurupantara. The primer *matK* 13 had shown distinct band

at 1000 bp range. This band was eluted and purified using DNA purification kit and used for sequencing (Fig. 11).

X. *matK* 17 (F6-R4)

A total of 3 genotypes have been amplified *M. cymbalaria, Luffa acutangula* (Haritam) and *Luffa acutangula* (Arka Sumit). The primer *matK* 17 had shown clear distinct band at 1000 bp range. The targeted 1000 bp band was eluted and purified using a DNA purification kit and used for sequencing (Fig. 12).

XI. *ITS 2*

A total of 6 genotypes have been amplified viz, *M. dioica* Kerala 1, *M. dioica* Kerala 2, *M. dioica* Kerala 3, *M. dioica* Kerala 4, *M. dioica* Odisa and *M. subangulata* 2. The primer *ITS 2* had shown multiple bands from 1000 to 550 bp in range (Fig. 13).

XII. *trnH-psbA*

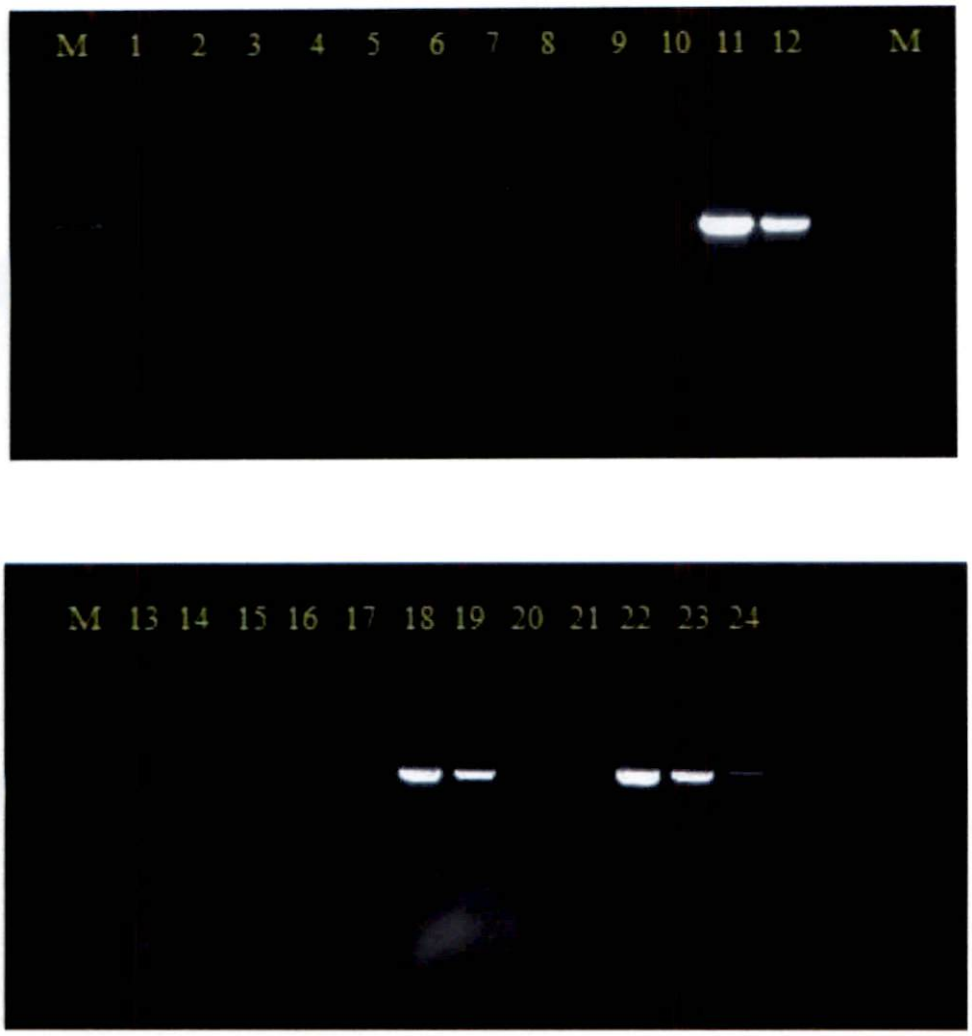
A total of 6 genotypes have been amplified viz, Arka Gaurav, *M. charantia* cv. *Muricata, M. dioica* Kerala 2, *M. dioica* Kerala 3, *M. dioica* Kerala 4 and *M. dioica* Odisa. The primer *trnH-psbA* was showed single distinct band at 200 bp in range (Fig.14). Since this size is much lower than the expected size, there were not used for sequencing.

Among this 18 combination of *matK* primer, 10 had shown amplification whereas remaining 8 viz, *matK* 5 (F1-R3), *matK* 8 (F3-R1), *matK* 11(F5-R5), *matK* 12 (F6-R6), *matK* 14 (F4-R6), *matK* 15 (F5-R4), *matK* 16 (F5-R6) and *matK* 18 (F6-R5) had not shown any type of amplification.

It was found that none of the reported primers, for all the three barcode loci are universal. Further, *matK* loci was found to be more amplifiable with the reported primer (Table 5). Only very few universal primer reported for *ITS2* and *trnH-psbA*, composed to *matK* and hence it was decided to consider *matK* as the candidate locus in this study.

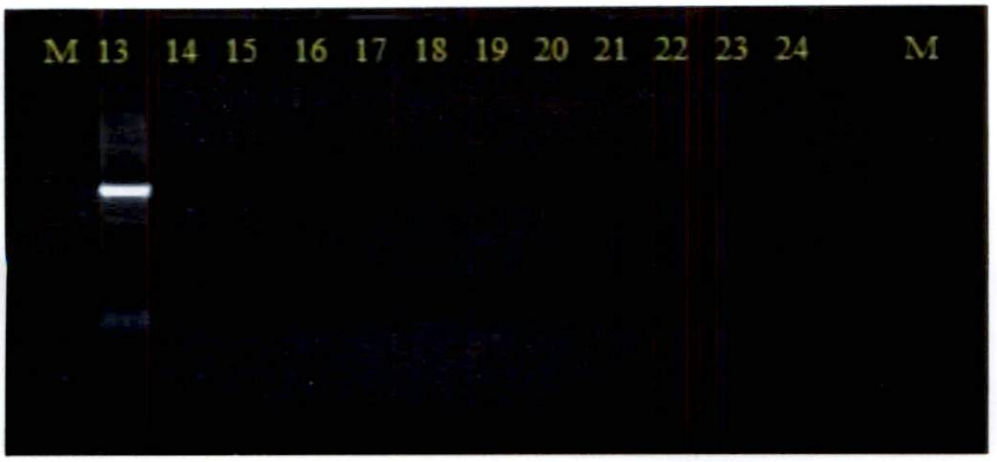
M- 1000 bp Ladder, **13**- Wild 2 (*M. charantia* cv. *muricata*),**14**- Wild 1 *(M. sahyadrica)*, **15**- Wild 2 *(M. sahyadrica)*, **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman ,**19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5

Fig. 3: DNA amplification pattern of the *Momordica* genotypes with barcode primer F1-R1



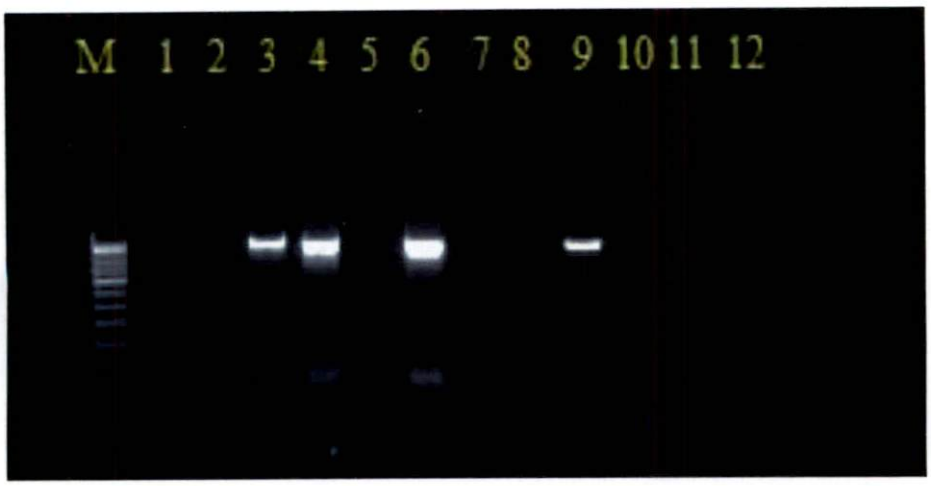
M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa,**13**- Wild 2 (*M. charantia* cv. *muricata*),**14**- Wild 1 *(M. sahyadrica)*, **15**- Wild 2 *(M. sahyadrica)*, **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman ,**19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5

Fig. 4: DNA amplification pattern of the *Momordica* genotypes with barcode primer F2-R2



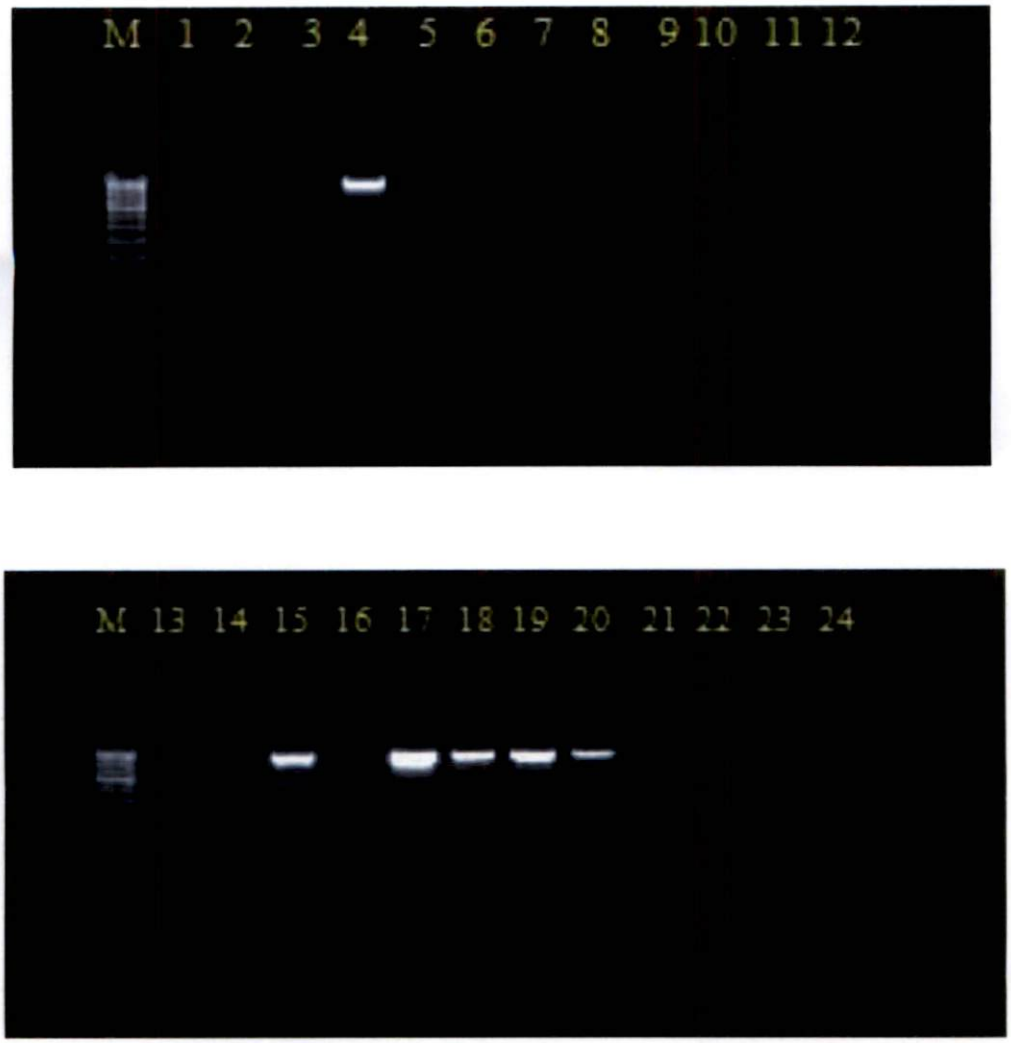
M- 1000 bp Ladder, **13**- Wild 2 (*M. charantia* cv. *muricata*),**14**- Wild 1 (*M. sahyadrica*), **15**- Wild 2 (*M. sahyadrica*), **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman ,**19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5

Fig. 5: DNA amplification pattern of the *Momordica* genotypes with barcode primer F3-R3



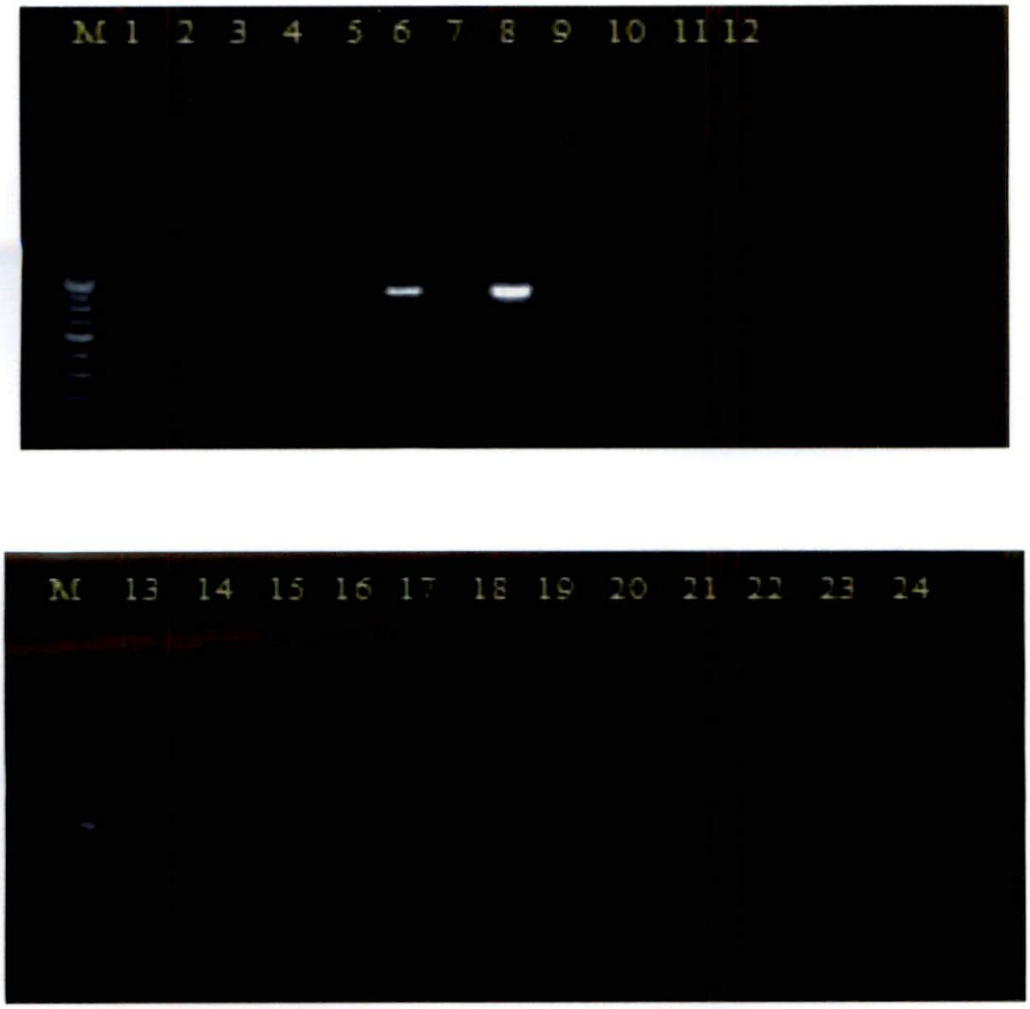
M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa

Fig. 6: DNA amplification pattern of the *Momordica* genotypes with barcode primer F1-R2



M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa, **13**- Wild 2 (*M. charantia* cv. *muricata*), **14**- Wild 1 *(M. sahyadrica)*, **15**- Wild 2 *(M. sahyadrica)*, **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman, **19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5

Fig. 7: DNA amplification pattern of the *Momordica* genotypes with barcode primer F2-R1

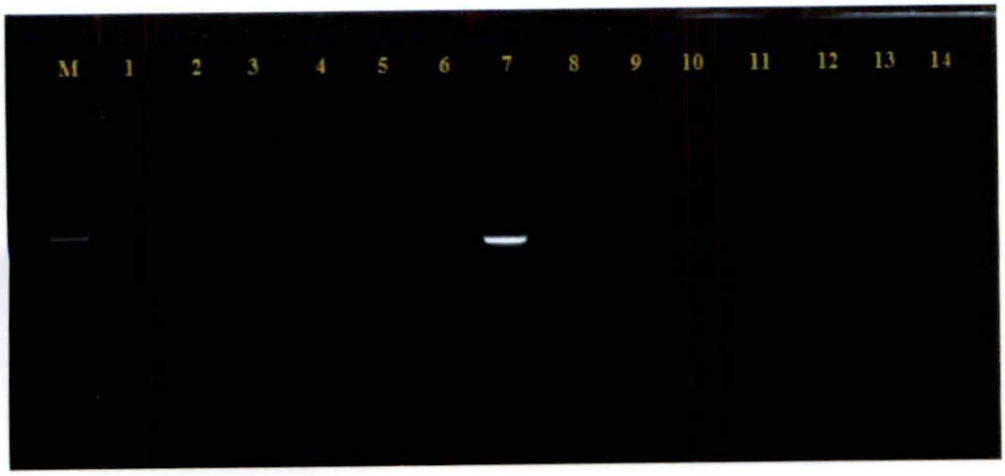


M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa, **13**- Wild 2 (*M. charantia* cv. *muricata*), **14**- Wild 1 *(M. sahyadrica)*, **15**- Wild 2 *(M. sahyadrica)*, **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman ,**19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5

Fig. 8: DNA amplification pattern of the *Momordica* genotypes with barcode primer F2- R3

M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata,* **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa

Fig. 9: DNA amplification pattern of the *Momordica* genotypes with barcode primer F3-R2

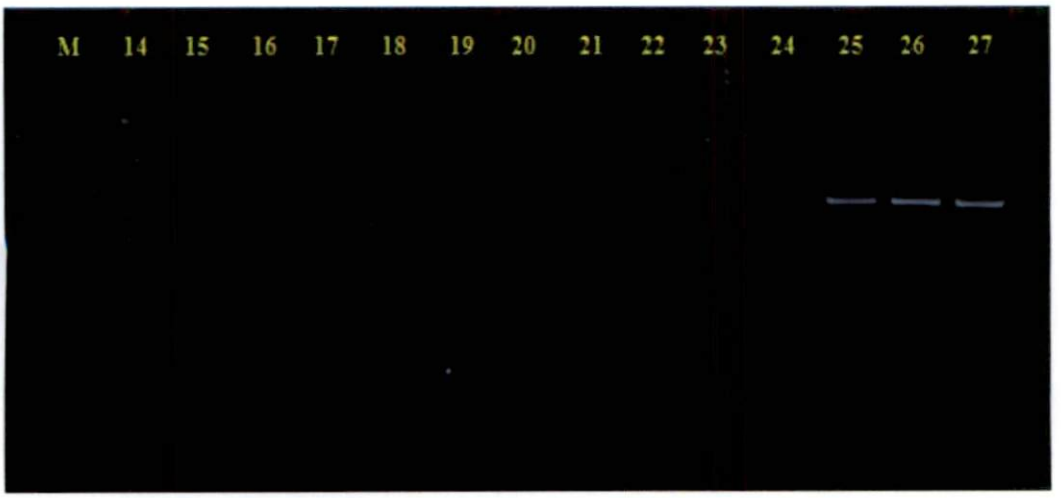


M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, 7- *M. charantia* cv. *muricata,* **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa, **13**- Wild 2 (*M. charantia* cv. *muricata*), **14**- Wild 1 (*M. sahyadrica*)

Fig. 10: DNA amplification pattern of the *Momordica* genotype with barcode primer F4-R4

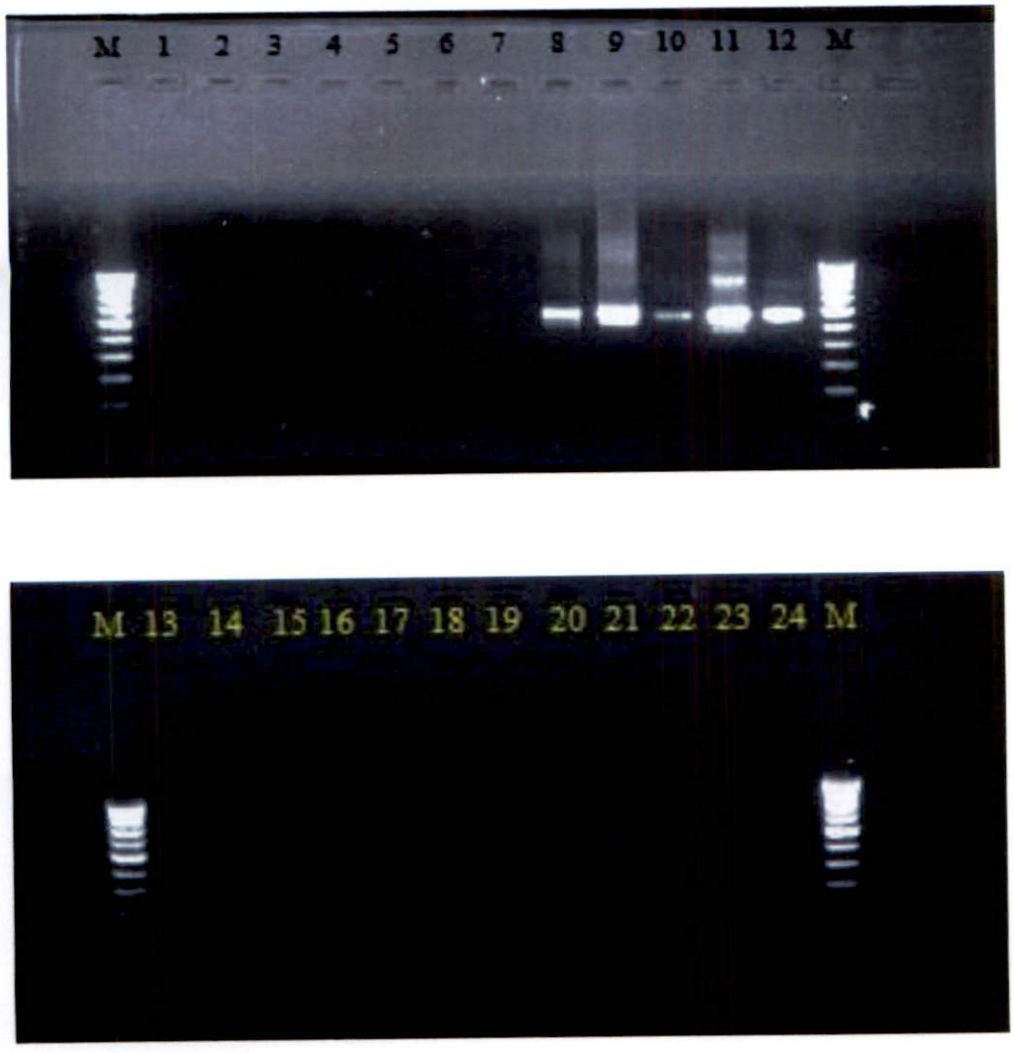
M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa, **13**- Wild 2 (*M. charantia* cv. *muricata*)

Fig. 11: DNA amplification pattern of the *Momordica* genotype with barcode primer F4-R5



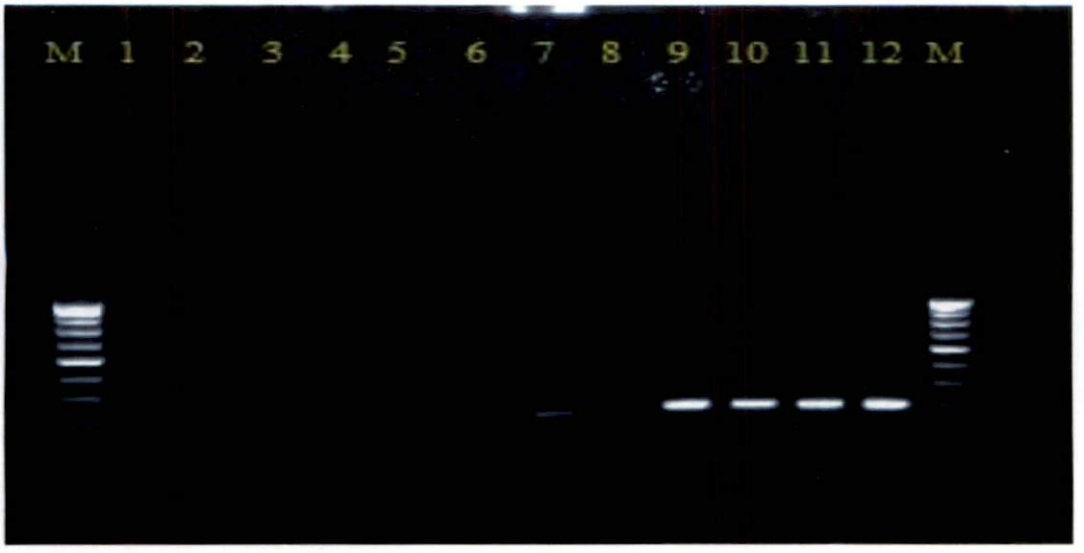
M- 1000 bp Ladder, **14**- Wild 1 (*M. sahyadrica*), **15**- Wild 2 (*M. sahyadrica*), **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman, **19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5, **25**- *M. cymbalaria,* **26**- *Luffa acutangula* (Haritam), **27**- *Luffa acutangula* (Arka Sumit)

Fig. 12: DNA amplification pattern of the *Momordica* genotype with barcode primer F6-R5



M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa, **13**- Wild 2 (*M. charantia* cv. *muricata*), **14**- Wild 1 (*M. sahyadrica*), **15**- Wild 2 (*M. sahyadrica*), **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman, **19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5

Fig. 13: DNA amplification pattern of the *Momordica* genotype with barcode primer *ITS2*



M- 1000 bp Ladder, **1**- Preethi , **2**- Kurupantara , **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata,* **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa

Fig. 14: DNA amplification pattern of the *Momordica* genotype with barcode primers *trnH-psbA*

Table 5: The amplification assay for the 25 *Momordica* and 2 *Luffa* accession using 18 primer combinations for *matK* locus and one primer each for *ITS2* and *trnH-psbA* loci

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	*ITS*2	trnH-psbA
M. charantia var. Preethi													✓							
*M. charantia*var.Kuruppantara													✓							
M. charantia var. Vadakara				✓																
M. charantia var. V53									✓											
M. charantia var. JNM7				✓																
M. subangulata renigera1						✓														
M. subangulata renigera2	✓			✓															✓	
M. subangulata renigera3	✓	✓																		✓
M. subangulata renigera4		✓																		
M. subangulata renigera5		✓																		
M. subangulata arka gurav				✓			✓												✓	✓
M.c. muricata										✓										✓
Wild2similarto*M.c.muricata*	✓		✓				✓													
M. dioica KL1							✓		✓										✓	
M. dioica KL2				✓					✓										✓	✓
M. dioica KL3	✓								✓										✓	✓
M. dioica KL4		✓																	✓	✓
M. dioica odisha		✓																	✓	✓
Wild3similar to *sahyadrica*	✓		✓				✓													
Wild1similar to *sahyadrica*	✓																			
M. sahyadrica Annamalai						✓														
M. balsamina						✓														
M. cochinchinensis NE		✓				✓														
M. cochinchinensis Andaman		✓				✓														
M. cymbalaria													✓				✓			
L. acutangula Haritam													✓				✓			
*L. acutangula*ArkaSumit													✓				✓			

4.3.2 PCR product sequencing

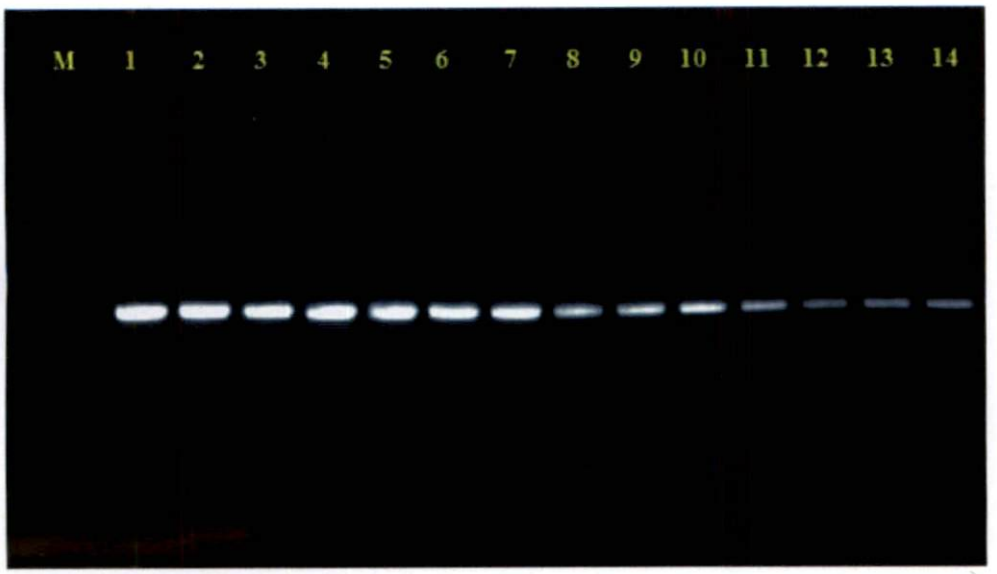
A total of 25 sequences of the *Momordica* species and 2 sequences of the *Luffa* species were amplified with the primer sets selected on the above mentioned screening. The amplified products were run on 1.5 per cent agarose gel. If multiple bands were obtained after amplification with any primer, only the band having molecular weight nearest to that of the targeted region that was 1000 bp was eluted and purified using a DNA purification kit (NucleoSpun® Gel Clean-up kit). The quality and quantity of the eluted and purified DNA were confirmed using agarose gel electrophoresis and was further subjected to PCR with same primer sets (Fig.15) to confirm that the band contains the amplified products from a single region of the genome. If the DNA eluted from a single band showing multiple band, then it was considered as non specific amplification, and was not used for direct sequencing on NGS platform.

The PCR product, confirmed to yield only single band on electrophoresis, was further sent for sequencing with specific set of primer. The sequencing was carried out by outsource (Sci Genome Lab. Pvt. Ltd., Cochin).

The sequences generated from each accession are furnished here under:

1. *M. charantia* var. Preethi

TTTAAATTATGTGTCAGATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCT
TCGCTCCTGGGTGAAAGATGCCTCTTCTTTTTTTATTACGGTTCTTTTTTCACGAGTATTGTAATT
GGAATAGTCTTAGACTTCAAAAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCT
TGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTACTTTTTCTACGTAACCAATCTT
CTCATATACGATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAA
CATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATG
CATTATGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAA
GTGGAAATATTACCTTGTCAATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGA
TCTATATAAACCAATTATCCAAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGACTA
AATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGA
AACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTAGCTAAAGCTAAATTTTGTAAC
GTATTAGGGCATCCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCG
ATTTGTGCGTATATCTAGAAATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAGAATTTGT
ATCGAATCAAATATATACTTCGCCTTTCTTGT



M- 1000 bp DNA Ladder, **1-14-** Re-amplified PCR sample

Fig. 15: Reamplification of the DNA eluted from *matK* specific bands to confirm the locus specificity

2. *M. charantia* var. Kurupantara

CTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCAA
AAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATG
TATGTGAATACGAATCCATTTTACTTTTTCTACGTAACCAATCTTCTCATATACGATTAACTTCTT
ATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCT
AATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGG
AAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAATATTACCTTGTCA
ATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC
AAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAG
TCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCT
ATTATTACTCTGCTTGGATCATTAGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAG
TAAGACGACCTGATCGAT

3. *M. charantia* var. JNM 7

TCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCA
AAAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTAT
GTATGTGAATACGAATCCATTTTACTTTTTCTACGTAACCAATCTTCTCATATACGATTAACTTCT
TATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGC
TAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGG
AAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAATATTACCTTGTCA
ATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC
AAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAG
TCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCT
ATTATTACTCTGCTTGGATCATTAGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAG
TAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCTAGAA
ATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAGAATTTGTATCGAATCAAATATATACTT
CGCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGCGCTCTTTTTAAAAG
GTTAAAT

4. *M. charantia* var. Vadakara

TTCTTTTCTTTATTACGGTTCTTTTTTCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCAAAA
AAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTA
TGTGAATACGAATCCATTTTACTTTTTCTACGTAACCAATCTTCTCATATACGATTAACTTCTTAT
AGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAA
TTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAA
AATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAATATTACCTTGTCAAT
TTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAA

GCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTC
AGATGCTAGAA

5. *M. charantia* var. V 53

CTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCAA
AAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATG
TATGTGAATACGAATCCATTTTACTTTTTCTACGTAACCAATCTTCTCATATACGATTAACTTCTT
ATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCT
AATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGG
AAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAATATTACCTTGTCA
ATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC
AAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAG
TCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCT
ATTATTACTCTGCTTGGATCATTAGCTAAAGCTAAATTTTGTAACGTATTAGGCATCCCATTAGT
AAGACGACCTGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCTAGAAAT
CTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCG
CCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTG

6. *M. subangulata* var. Arka Gaurav

CGAATCTTAGTTCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTAAGGTTT
TTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAAAAAATTTATTTCTTTTTTTTCAAAAAGA
AATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTT
CTACGTAACCAATCTTCTCATATACTATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTC
TATGGAAAAATAGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTT
TAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATCAATTCTGGTTTCAAAAGATACGC
CACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGG
TCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATAT
TTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAG
ATAATGCTACAAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAA
GCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTT
TGATATTATTGATCGATTTGTGCGTATATCCAGAAATCTTTCTCATTATTACGGAGGATCCTCAA
AAAAAAAGAATTTGTATCGAATCAAATATATACTTCGCCTTTCTTGTGTTAAAACTTTGGCTCGT
AAACACAAAAGTACTGTACG

7. *M. charantia* cv. *muricata*

TCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTTATTAAGGTTCTTTTTTCACG
AGTATTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAA
AAAAATTGATTTATTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTAT
GTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCT

TATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGC
TAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGG
AAAATCTATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCA
ATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC
AAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAG
TCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCT
ATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAG
TAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGAA
ATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAGAATTTGTATCGAATCAATATATACTTCG
CCTTTCTTGTGTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGCGCTC

8. *M. dioica* Kerala 1

TGTGTCAGATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCTCCTG
GGTGAAAGATGCCTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAG
TACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGCAGATTAGTCTTGTTCCTATATA
ATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTAT
TAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAA
GTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGA
TATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATT
ACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGAAACGATCTATATAAAC
CAATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGT
GGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACA
CTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCA
TCCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTA
TATCCAGGAAATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCA
AATATATACTTCGCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGCGC

9. *M. dioica* Kerala 2

CTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAAAAAAT
TGATTTATTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTG
AATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCTTATAGGG
GCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAATTAT
TTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATC
TATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTAT
GGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCG
TTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGA
TGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATT
ACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGAC
GACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGAAATCTTT
CTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCGCCTT

TCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGCGCTCTTTTTAAAAGGTTAAA
TTC

10. *M. dioica* Kerala 3

TGTCAGATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCTCCTGG
GTGAAAGATGCCTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGT
ACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGCAGATTAGTCTTGTTCCTATATAA
TTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATT
AACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAG
TGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGAT
ATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTA
CCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGAAACGATCTATATAAACC
AATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTG
GTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACAC
TAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCAT
CCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTAT
ATCCAGAAATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAA
TATATACTTCGCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACG

11. *M. dioica* Kerala 4

CTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAAAAAAT
TGATTTATTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTG
AATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCTTATAGGG
GCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAATTAT
TTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATC
TATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTAT
GGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCG
TTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGA
TGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATT
ACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGAC
GACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGAAATCTTT
CTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCGCCTT
TCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGCGCTCTTTTTAAAA

12. *M. dioica* Odisha

CTTTATTAAGGTTTTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAAAAAATTTATTTCTT
TTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAA
TCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCTTATAGGGGCCTTTTTG
AGCGAATATATTTCTATGGAAAAATAGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCT

ATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATCAATTCTGGT
TTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATC
ATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGAC
TTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAA
TTCATTTCTAATAGATAATGCTACAAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTG
GTCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGACGAACCTGGAT
CGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGAAATCTTTCTCATTATTA
CGGAGGATCCTCAAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCGCCTTTCTTGTGTTA
AAACTTTGGCTCGTAAACACAAAAGTACTGTAC

13. Wild 2 (*M. charantia* cv. *muricata*)

AGATGCTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCA
AAAAAATTGATTTCTTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTAT
GTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCT
TATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGC
TAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGG
AAAATCTATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCA
ATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC
AAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAG
TCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCT
ATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAG
TAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCAGAAA
TCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAATATATACTTC
GCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACG

14. Wild 3 *(M. sahyadrica)*

CAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAG
TATTGTAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGATT
AGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGCAACCA
ATCTTCTCATAAACTATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAA
TCGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTT
TCATGCATTATGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTATGATG
AATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGCCACAACCAGA
AAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGGGCGATATTTCGAGGGGCG
ACTAAATCCTTCAGTGGGAGGGAGGCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACG
AAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAAT

15. Wild 2 *(M. sahyadrica)*

CATTTAAATTATGTGTCAGATGTATTAATACCTATTAATACCCTATCCCCTCCATCTGGAAATTTT
AGTTCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTAAGGTTCTTTTGTCA
CGAGTATTGTAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAA
GATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTA
ACCAATCTTCTCATATACTATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGA
AAAATCGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGA
TCCTTTCATGCATTATGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTA
TGATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGTCACAA
CCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAG
TGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAAT
GCTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAA
ATTTTGTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGATCGATTCGTCGGATTTTGATAT
TATTGATCGATTTGTGCGTATATCCAGAAATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAA
AAGAATTTGTATCG

16. *M. sahyadrica* Anamyla type

ATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTAC
GGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAA
GATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTA
TGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGG
GCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTT
CTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATT
GGCTAAAAGCTAAATTTTGTAACGTATTTAGGGCATCCCATTAATAAGACGACCCTGGATCGAT
TCGTCGGATTTTGATATTATTGA

17. *M. balsamina*

TCTTAGTTCAAATCCTTCTCTCCTGGGTGAAAAGATGCCTCTTCTTTTCATTTATTACGGTTTTTTT
TCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAA
GAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTT
TTCTACGTAACCAATCTTCTCATATACGATTAACTTCTTATAGGGGCCTTTTTAAGCGAATATATT
TCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGATATCTTACGGGTC
TTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATAC
GCCACTTCTGATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATGTCATTTTTATGTGT
GGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTTGAATTTTTGGGCTAT
ATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGTTAGAAAATTCATTTCTAAT
AGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTA
AAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGGATCGATTCGTCGGA
TTTTGATATTATTGATCGATTTGTGCGTATATCCATAAATCTTTCTCATTATTACAGAGGATCCTC
AAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCGACTTTCTGTGGC

18. *M. cochinchinensis* Andaman

CAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAG
TATTGTAATAGTCTTAGTACTTCAAAAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATT
AGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCA
ATCTTCTCATATACGATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAA
TCGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTT
TCATGCATTATGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTATGATG
AATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGA
AAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGC
GACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTAC
GAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTT
GTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTTTTATATTATT
GATCGATTTGTGCGTATATCCAGAAATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAG
AATTTGTATCGAATCAAATA

19. *M. cochinchinensis* North East

TCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAAAAAA
TTTATTTCTTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGT
GAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACGATTAACTTCTTATAGG
GGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAATTA
TTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAAT
CTATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTA
TGGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGC
GTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAG
ATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTAT
TACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGA
CGACCTGGATCGATTCGTCGGATTTTTATATTATTGATCGATTTGTGCGTATATCCAGAAATCTTT
CTCATTATTACAGAGGATCCTCAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCGCCTTT
CTTGTGTTAAAACTTTGGCTCGTAAACACAAAGTACTGTACGCGCTCTTTTTAAAGG

20. *M. subangulata* 1

TTCAAATCCTTCGCTCCTGGGTGAAAAGATGCCTCTTCTTTTCATTTATTAAGAGTTTTTTTTTCA
CGAGTATTGTAATAGTCTTAGTACTTCAAAAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAA
GATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTA
ACCAATCTTCTCATATACTATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGA
AAAATAGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGA
TCCTTTCATGCATTATGTTAGATATCAAGGAAAATCAATTCTGGTTTCAAAAGATACGCCACTTA
TGATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGTCACAA

CCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAG
TGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAAT
GCTACAAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAA
ATTTTGTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGATCGATTCGTCGGATTTTGATAT
TATTGATCGATTTGTGCGTATATCCAGAAATCTTTCTCATTATTACGGAGGATCCTCAAAAAAAA
AAGAATTTG

21. *M. subangulata* 2

GATGCCTCTTCTTTTCATTTATTAAGGTTTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAA
AAAAATTTATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAAATTCTTAT
GTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCT
TATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGTCAAAGTGTTTGC
TAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGG
AAAATCAATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTC
AATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATC
CAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGA
GTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACAAAGAAACTCGATACACTAGTTCC
TATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTA
GTAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGA
AATCTTTCTCATTATTACGGGAGGATCCTCAAAAAAAAAAGAATTTGG

22. *M. subangulata* 3

CTCTTCTTTTCTTTATTAAGGTTTTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAAAAAA
TTTATTTCTTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGT
GAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCTTATAGG
GGCCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGTCAAAGTGTTTGCTAATTA
TTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAAT
CAATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCAATTTA
TGGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGC
GTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAG
ATGCTAGAAAATTCATTTCTAATAGATAATGCTACAAAGAAACTCGATACACTAGTTCCTATTAT
TACTCTGCTTGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGAC
GACCTGGGATCGATTCGTCGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGAAATCTTT
CTCATTATTACGGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAATATATACTTCGCCTT
TCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGCGC

23. *M. subangulata* 4

TTCAAATCCTTCGCTCCTGGGTGAAAGCATGCCTCTTCTTTTCTTTATTAAGGTTTTTTTTTCACG
AGTATTGTAATAGTCTTAGTACTTCAAAAAAATTTATTTCTTTTTTTTTCAAAAAGAAATCGAAGA

TTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAAC
CAATCTTCTCATATACTATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAA
AATAGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATC
CTTTCATGCATTATGTTAGATATCAAGGAAAATCAATTCTGGTTTCAAAAGATACGCCACTTATG
ATGAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATATCATTTTTATGTGTGGTCACAACC
AGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTG
TGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCT
ACAAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATT
TTGTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGATCGATTCGTCGGATTTTGATATTAT
TGATCGATTTGTGCGTATATCCAGAAATCTTTCTCATTATTACGGAGGATCCTCAAAAAAAAAG
AATTTGTATCGAATCAAATATATACTTCGCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAA
AAGTACTGTACGCGCT

24. *M. subangulata* 5

TGCCTCTTCTTTTCATTTATTAAGGTTCTTTTTTCACGAGTATTGTAATAGTCTTAGTACTTCAAA
AAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGT
ATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACTATTAACTTCTTA
TAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGTCAAAGTGTTTGCTA
ATTATTTTTCGGCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGA
AAATCAATTCTGGTTTCAAAAGATACGCCACTTATGATGAATAAGTGGAAATATTACCTTGTCA
ATTTATGGCAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC
AAGCGTTCTCTGGACTTTTTGGGCTATATTTCGAGTGTGCGACTAAATCCTTCAGTGGTATGGAG
TCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACAAAGAAACTCGATACACTAGTTCCT
ATTATTACTCTGCTTGGATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAG
TAAGACGACCTGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCGTATATCCAGAAA
TCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAATATATACTT
CGCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGC

25. *M. cymbalaria*

TCATTCAATATTTCCTTTTTTGAGGATAAATTCCACATTTTAATTATGTGTCAGATGTATTAATAC
CCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTT
TTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCAAAAAAAT
TTATTTATTTTTTTTTCAAAAATAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTG
AATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATACGATTAACTTCTTATAGGG
GCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTCAAAGTGTTTGCTAATTAT
TTTTCGCCTATCTTACGGGTCTTTAAGGATCCTTTCATGCATTATGTTAGATATCAAGGAAAATC
TATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAATATTACCTTGTCAATTTAT
GGAAATATCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATACAAGCG
TTCTCTTGACTTTTTTGGGCTATATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGA
TGCTAGAAAATTCATTTCTAATAGATAATGCTACGAATAAACTCGATACACTAGTTCCTATTATT

ACTCTGCTTGGATCATTGGCTAAAGCGAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGAC
GACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGCATATATCAAGAAATCTTT
CTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAATCAAA

26. *Luffa acutangula* var. Haritam

ACATTTAAATTATGTGTCAGATGTATTAATACCCTATCCCCTTCATCTGGCAATTTTAGTTCAAAT
CCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTG
TAATTGGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGAT
TAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACC
AATCTTCTCATATACGATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAA
ATCGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATCCT
TTCATGCATTATATTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGAT
GAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAG
AAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTG
GGACTAAATCCTTCAGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTA
CGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGTCTAAAGCGAAATTT
TGTAACGTGTTGGGGCATCCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTTTGATATTAT
TGATCGATTTGTGCGTATATCCAGAAATCTTTT

27. *Luffa acutangula* var. Arka Sumit

GAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAATTGGAATAGTCT
TAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTAT
ATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAACCAATCTTCTCATATAC
GATTAACTTCTTATAGGGGCCTTTTTGAGCGAATATATTTCTATGGAAAAATCGAACATCTTGTC
AAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATCCTTTCATGCATTATATT
AGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAAT
ATTACCTTGTCAATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATA
AACCAATTATCCAAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGGGACTAAATCCTTC
AGTGGTATGGAGTCAGATGCTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGAT
ACACTAGTTCCTATTATTACTCTGCTTGGATCATTGTCTAAAGCGAAATTTTGTAACGTGTTGGG
GCATCCCATTAGTAAGACGACCTGGATCGATTCGTCGGATTTTGATATTATTGATCGATTTGTGC
GTATATCCAGAAATCTTTCTCATTATTACAGAGGATCCTCAAAAAAAAAAGAATTTGTATCGAAT
CAAATATATACTTCGCCTTTCTTGTGTTAAAACTTTGGCTCGTAAACACAAAAGTACTGTACGC

4.4 Sequence analysis

4.4.1 Submission to NCBI

A total of 25 DNA sequences from *Momordica* species and 2 from *Luffa* species were deposited in the NCBI GenBank.

Deposition was carried out with the submission tool BankIt. Annotation was done and sequence information was furnished in 5-column tab-delimited table. This table included the gene name, gene product, coding region, stop codon, start codon etc.

4.4.2 Cluster analysis

The genetic dissimilarity was determined for the genotypes based on the *matK* sequence (Fig. 16). The phenogram generated had shown III clusters. Cluster I had *M. balsamina*, *M. cymbalaria*, *M. charantia*, *M. sahyadrica*, *M. dioica*, *M. subangulata* and *Luff* spp.

Subcluster Ia had *M. balsamina* and Subcluster it was closely associated with *M. cymbalaria* and *M. charantia* cv. *charantia* accession. Interestingly, the *Luffa* accession had fallen within the *M. cymbalaria*. Subcluster Ia represented the monocious plants. Subcluster Ia had *M. sahyadrica* Annamalai type and also wild relative of *M. sahydrica*. The two accession of *M. dioica* had fallen within subcluster Ia. The variability within *M. dioica* was clear since two accession of *M. dioica*, accession 2 and 4 kerala had fallen in cluster III.

Cluster Ib was distinctly a *M. subangulata* cluster with the exception that one accession of *M. dioica* for odisa had fallen into this. This had clearly shown the power of *matK* locus in species discrimination.

In cluster II had two accessions belonging to *M. cochinchinesis* and one wild accession which is supported to belong to *M. sahyadrica*. Cluster III appeared like a miscellaneous cluster where the accession which did not get the exact allocation has fallen.

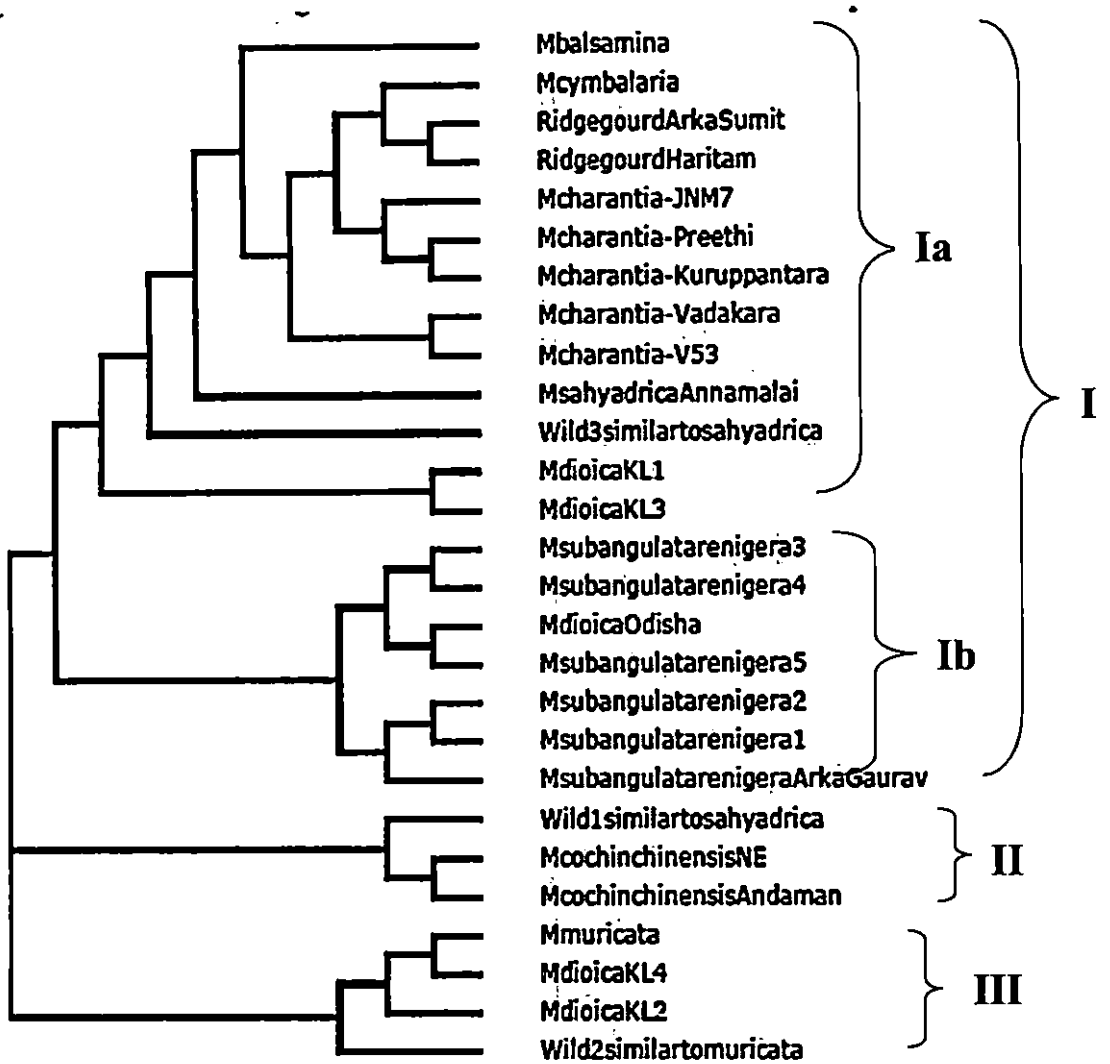


Fig. 16: The pylogenetic tree of 25 *Momordica* and 2 *Luffa* accession based on dissimilarity coefficients using the *matK* sequence, as generated using ClustalW omega.

The cluster III had two lines of *M. charantia* ssp. *muricata* and two accession of *M. dioica*. The sequence had shown that though *M. c. muricata* was close to *M. c. charantia* by floral morphology, at genome level, they are very distinct.

4.4.3 Genetic divergence within and between species

The inter specific distances were calculated using K2P (Kimura 2 parameter) model as recommended by the Consortium of Barcode of Life (CBOL, http:// www.barcoding.si.edu/protocols.html) using MEGA 4.0 software.

The range of pair wise dissimilarity was 0.0032 to 0.04 with an average of 0.0134 (Table.6). The similarity between *Luffa* and *Momordica* ranged from 0.0096 to 0.0458 (*Luffa* and *M. sahyadrica* Wild3). Among the *Momordica* species, maximum dissimilarity was observed between wild *M. sahyadrica* and *M. balsamina* (0.0458) followed by between *M. subangulata* and *M. balsamina* (0.0291) and *M. subangulata* ana *M. cymbalaria* (0.0259), and minimum between *M. dioica* and *M. cochinchinesis* (0.0032).

4.4.4 Barcoding gap assessment and barcode generation

The distribution of inter-specific variability were compared using DNA barcoding gap. This DNA barcode had separate and non-overlapping variations among all the species. The results had shown that the species variation of *matK* exhibited distinct gaps.

The *Momordica* and *Luffa* sequences were aligned with software clustalW (Fig. 17). The conserved region was shown in "*" in the aligned sequence. The variability has been checked manually. The nucleotide sequence and variability have been observed specific to the each species (Fig.18).

By aligning, all the sequences, the maximum length of *matK* gene developed from this study was 982 bp. After identifying the gaps with position for each species, the barcode were developed.

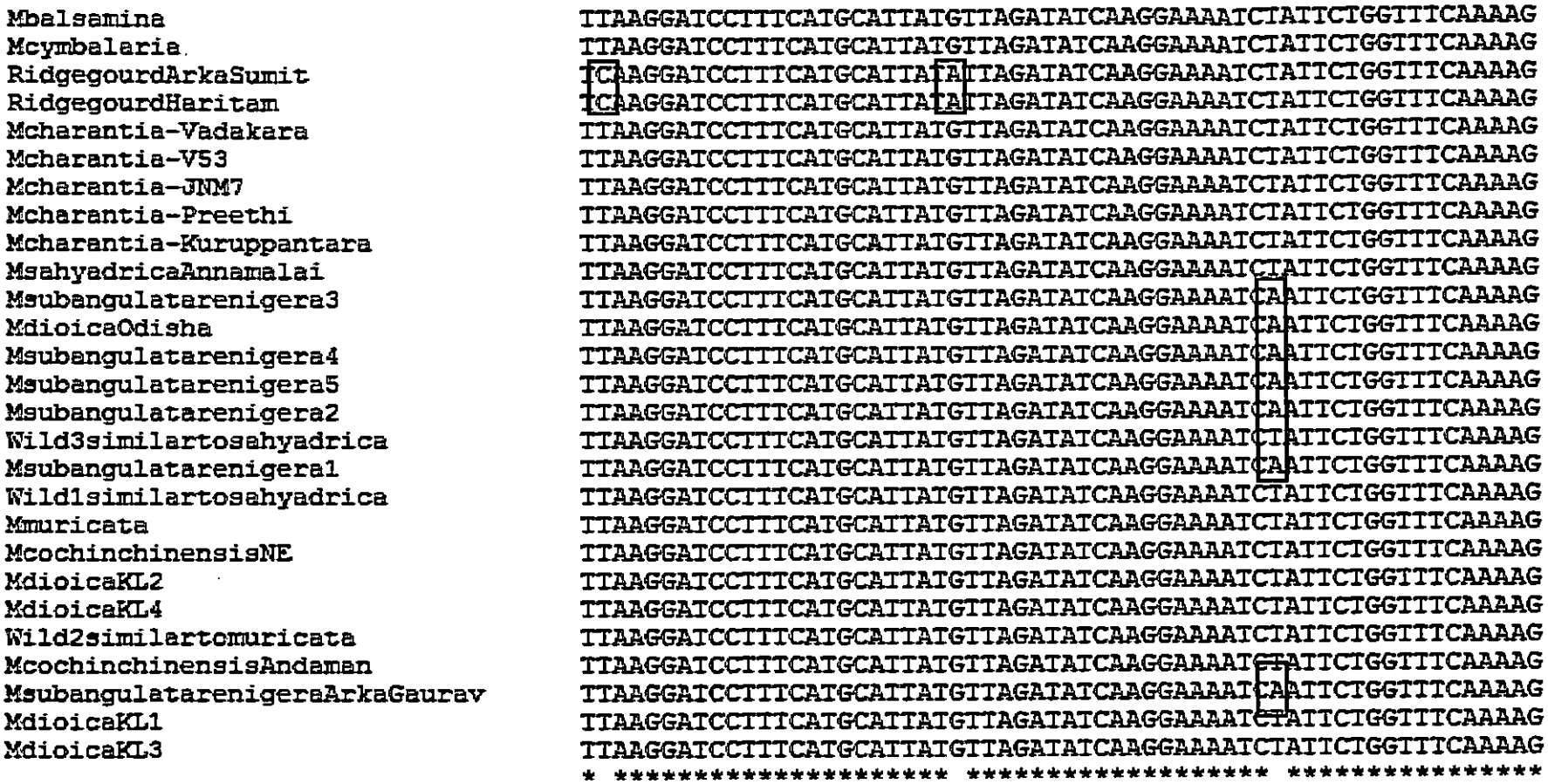


Fig 17: The *matK* sequences of *Momordica* and *Luffa* accession aligned using ClustalW software to isolate the barcode gaps

Absence of * shows the presence of gaps

Table 6 : The distance among the accession based on the *matK* sequence, as derived using MEGA 4.0

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1. Mbelasma																											
2. Mcymbalella	1.0225																										
3. RidgepurdAkaSumit	1.0193	0.0226																									
4. Ridgepurd-Hariari	1.0193	0.0226	0.0000																								
5. Mcrataria-Kacakara	1.0098	0.0128	0.0036	0.0036																							
6. Mcrataria-M5C	1.0098	0.0128	0.0036	0.0036	0.0000																						
7. Mcrataria-JNM7	1.0098	0.0128	0.0036	0.0036	0.0000	1.0000																					
8. Mcrataria-Pleethi	1.0098	0.0128	0.0036	0.0036	0.0000	1.0000	0.0000																				
9. Mcrataria-Kuruppantara	1.0098	0.0128	0.0036	0.0036	0.0000	1.0000	0.0000	0.0000																			
10. MsahyadricaAmanula	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128																		
11. Msubangulatarerigera3	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064																	
12. MziricaOcisra	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064	1.0000																
13. Msubangulatarerigera4	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064	1.0000	1.0000															
14. Msubangulatarerigera5	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064	1.0000	1.0000	0.0000														
15. Msubangulatarerigera2	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064	1.0000	1.0000	0.0000	0.0000													
16. Wid3siriatisahyadrica	1.0456	0.0425	0.0458	0.0458	0.0358	1.0358	0.0358	0.0358	0.0358	0.0226	1.0292	1.0292	0.0292	0.0292	0.0292												
17. Msubangulatarerigera"	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064	1.0000	1.0000	0.0000	0.0000	0.0000	0.0292											
18. Wid1siriatisahyadrica	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064										
19. Mmuricata	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064	0.0000									
20. MtocrinchnensisNE	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064	0.0000	0.0000								
21. MziricaK_2	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064	0.0000	0.0000	0.0000							
22. MziricaK_4	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064	0.0000	0.0000	0.0000	0.0000						
23. Wid2siriatsmuricata	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064	0.0000	0.0000	0.0000	0.0000	1.0000					
24. MtocrinchnensisAncamran	1.0225	0.0193	0.0226	0.0226	0.0128	1.0125	0.0128	0.0128	0.0128	0.0000	1.0064	1.0064	0.0064	0.0064	0.0064	0.0225	1.0064	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000				
25. Msubangulatarerigera ArkaGaura	1.0291	0.0259	0.0291	0.0291	0.0193	1.0193	0.0193	0.0193	0.0193	0.0064	1.0000	1.0000	0.0000	0.0000	0.0000	0.0292	1.0000	0.0064	0.0064	0.0064	0.0064	1.0064	0.0064	0.0064			
26. MziricaK_1	1.0125	0.0226	0.0259	0.0259	0.0160	1.0161	0.0160	0.0160	0.0160	0.0032	1.0096	1.0096	0.0096	0.0096	0.0096	0.0253	1.0096	0.0032	0.0032	0.0032	0.0032	1.0032	0.0032	0.0032	0.0096		
27. MziricaK_3	1.0125	0.0226	0.0259	0.0259	0.0160	1.0161	0.0160	0.0160	0.0160	0.0032	1.0096	1.0096	0.0096	0.0096	0.0096	0.0253	1.0096	0.0032	0.0032	0.0032	0.0032	1.0032	0.0032	0.0032	0.0096	0.0000	

4.4.5 BLAST for species discrimination

The sequences were further subjected to BLAST to discreminat the species. Results indicated that *matK* is powerful to yield up to 95 % efficiency to differentiate the species. At the genus level, *matK* yielded 100% efficiency (Table 7, Fig. 18).

4.4.6 Identification of conserved region and design of barcode primer for *Momordica*

The conserved sequence in the generated sequence was identified using the clutalW software. A total of 27 sequences were aligned and the conserved region across the total length of 980 base pairs was identified.

The conserved sequence in the generated sequence was identified using the clustalW Software. A total 27 sequences was align it shows the conserver region in "*" symbol (Fig. 17). Maximum stretch has been taken for conserved region.

In the present study, previously reported barcode primers were used to amplify the *matK* locus. As discussed earlier, here primers were not successful to amplify the locus across all the accession; hence it was mandatory to identify the barcode primers for *Momordica*. Based on the aligned sequences, the primers were designed using Primer3 software. The identified conserved region has been taken as a query sequence. The parameter has been chosen to get the product size of 900bp. The parameters for designing good primers were taken care.

The 982 bp sequence has been generated for this study, which was without any barcoding gaps.

5'TCATTCAATATTTCCTTTTTTGAGGATAAATTCCACATTTTAATTATGTGTCAGATGTATTAATACCCTA
TCCCCTCCATCTGGAAATCTTAGTTCAAATCCTTCTCTCCTGGGTGAAAAGATGCCTCTTCTTTTCATTTATTA
CGGTTTTTTTTTCACGAGTATTGTAATTGGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAA
AGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACG
TAACCAATCTTCTCATATACGATTAACTTCTTATAGGGGCCTTTTTAAGCGAATATATTTCTATGGAAAAATC
GAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGATATCTTACGGGTCTTTAAGGATCCTTTCATGCATTA
TGTTAGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATGAATAAGTGGAAATATTA
CCTTGTCAATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCC

AAGCGTTCTCTTGAATTTTTGGGCTATATTTCAAGTGTGCGACTAAATCCTTCAGTGGTATGGAGTCAGATGT
TAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGG
ATCATTGGCTAAAGCTAAATTTTGTAACGTATTAGGGCATCCCATTAGTAAGACGACCTGGATCGATTCGTCG
GATTTTGATATTATTGATCGATTTGTGCGTATATCCATAAATCTTTCTCATTATTACAGAGGATCCTCAAAAA
AAAAGAATTTGTATCGAATCAAATATATACTTCGACTTTCTGTGGCTAAAACTTTGGCTCGTAAACACAAAA
GTACTGTACGCGCTCTTTTTAAAAGGTTAAATTC3'

Primer set for DNA barcoding in Momordica using matK locus

5' AGGGTTTGGAGTCATTGTGG 3' (59.82°C) - Forward F1

5' GAATCGATCCAGGTCGTCTT 3' (59.09°C) - Reverse R1

Product size: 897

A total 25 accession of *Momordica* has amplified with both the primer, among this Primer set 1 has shown result in all accession (Fig. 20).

Table 7: The length of the *matK* generated in the accession under study

Sl. No.	Species name	Length of *matK* (bp) generated by sequencing	Maximum identity (%) in BLAST analysis
1	*M. charantia* var. Preethi	883	99
2	*M. charantia* var. Kuruppantara	670	99
3	*M. charantia* var. Vadakara	534	99
4	*M. charantia* var. V53	826	99
5	*M. charantia* var. JNM7	855	95
6	*M. subangulata renigera1*	790	90
7	*M. subangulata renigera2*	765	98
8	*M. subangulata renigera3*	829	92
9	*M. subangulata renigera4*	862	92
10	*M. subangulata renigera5*	832	93
11	*M. subangulata* Arka gurav	869	96
12	*M. c. muricata*	900	88
13	Wild 2 similar to *muricata*	832	97
14	*M. dioica* KL1	910	93
15	*M. dioica* KL2	850	91
16	*M. dioica* KL3	904	94
17	*M. dioica* KL4	840	92
18	*M. dioica* Odisha	816	92
19	Wild 3 similar to *sahyadrica*	648	-
20	Wild 1 similar to *sahyadrica*	861	-
21	*M. sahyadrica* Annamalai type	477	-
22	*M. balsamina*	835	97
23	*M. cochinchinensis* NE	840	91
24	*M. cochinchinensis* Andaman	801	99
25	*M. cymbalaria*	899	100
26	Ridge gourd Haritam	816	99
27	Ridge gourd Arka Sumit	844	95

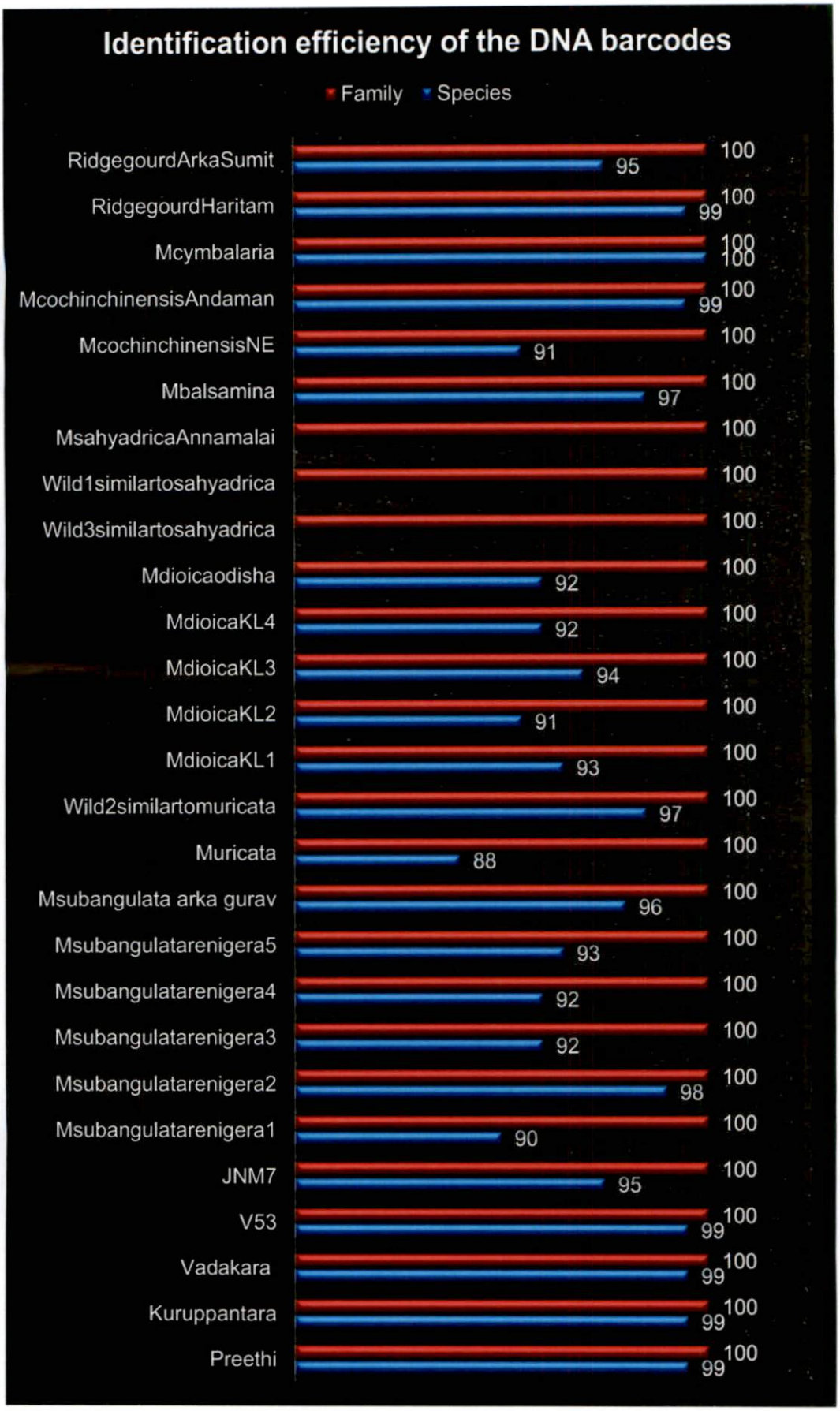


Fig 19: Identification of efficiency of the DNA barcode

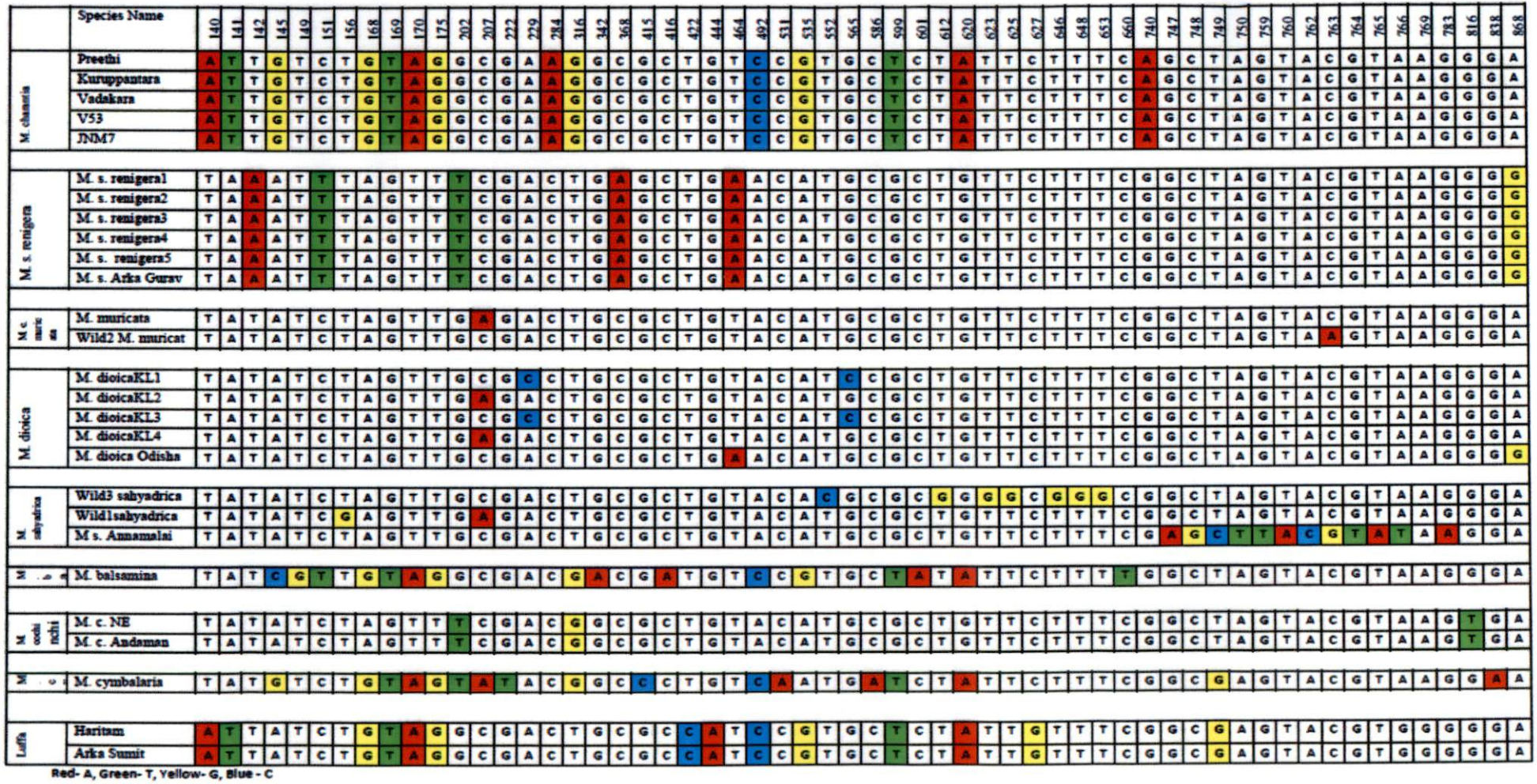
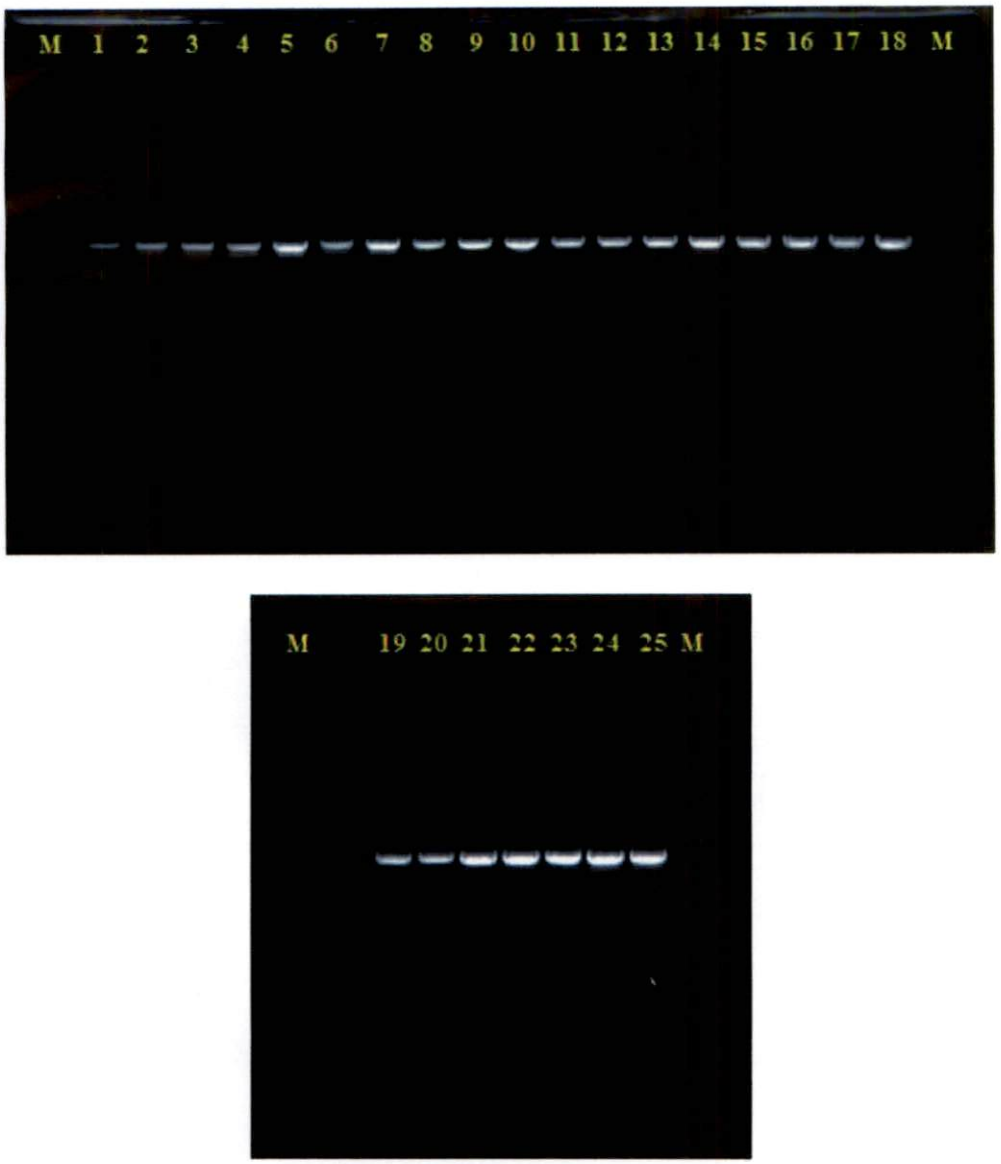


Fig 18: DNA barcode generated for the 7 *Momordica* spp. under study the distinct barcode gap of *M. c.* ssp *charantia*, *M. subangulata*, *M. balsamina*, *M. cymbalaria* and *Luffa* may be noted



M- 1000 bp Ladder, **1**- Preethi, **2**- Kurupantara, **3**- JNM 7, **4**- Vadakara, **5**- V 53, **6**- Arka Gaurav, **7**- *M. charantia* cv. *muricata* , **8**- *M. dioica* Kerala 1, **9**- *M. dioica* Kerala 2, **10**- *M. dioica* Kerala 3, **11**- *M. dioica* Kerala 4, **12**- *M. dioica* Odisa, **13**- Wild 2 (*M. charantia* cv. *muricata*),**14**- Wild 1 (*M. sahyadrica*), **15**- Wild 2 (*M. sahyadrica*), **16**- *M. sahyadrica* Annamalai type, **17**- *M. balsamina*, **18**- *M. cochinchinensis* Andaman, **19**- *M. cochinchinensis* Northeast, **20**- *M. subangulata* 1, **21**- *M. subangulata* 2, **22**- *M. subangulata* 3, **23**- *M. subangulata* 4, **24**- *M. subangulata* 5, **25**- *M. cymbalaria*

Fig. 19: DNA amplification pattern by *Momordica* species specific F1- R1 primer

Discussion

5. DISCUSSION

The genus *Momordica* L. (Cucurbitaceae) comprises 59 species distributed in Africa, with about ten species in Southeast Asia, out of that seven species *M. charantia, M. balsamina, M. dioica, M. cymbalaria, M. subangulata, M. cochinchinensis* and *M. sahyadrica* are reported in India. The genus *Momordica* is unique in India as all the species in this country hold edible fruits which are important for their taste and medicinal properties. Bitter melon (*M. charantia*) is the most widely cultivated species of this genus in India, Sri Lanka, Philippines, Thailand, Malaysia, China, Japan, Australia, tropical Africa, South America and the Caribbean. Teasle gourd (*M. s.* ssp. *renigera*) is a semi-domesticated vegetable crop, which is grown commercially in several states of India such as West Bengal, Odisha, Assam, Tripura and Andaman and Nicobar Islands and Bangladesh.

However, information on identification and characterization in the genus *Momordica* was misleading due to the taxonomic misidentification, especially in the Indian region (Bharathi *et al.*, 2012). The existing diversity pattern and phonetic relationships among the cultivated varieties included under Indian *Momordica* spp. were not studied for their taxonomic characterization (Bharathi *et al.*, 2011).

The DNA barcoding was used for correctly identifying unknown species, with no taxonomic expertise (Hebert *et al.*, 2003). This study was done for the assessment of genetic diversity and establishing phenetic relationships among different genotypes of six currently cultivated Indian *Momordica* spp. and a taxon of controversial identity of *Momordica cymbalaria.*

5.1 Morphological characterization

The morphological characters were observed with the descriptor of *Momordica* which was given by Joseph *et al.*, in 2011. As Bharathi *et al.* (2011), Bharathi *et al.* (2012) had observed morphology, in this study also all *Momordica*

spp. shows the annual habitat, with entire or lobed leaves, monoecious in case of *charantia* and *balsamina* all others were dioecious, fruits were in medium to big size and green when unripe and red or orange on ripening and on surface soft spines.

5.2 Molecular characterization

The seven Indian *Momordica* species with their 25 genotypes, were characterised using the *matK* locus. The assessment of species variation using the reported universal barcoding primers has been done. At present, *rbcL, matK, psbA-trnH, rpoC1, rpoC1, ITS2, atpF–atpH* spacer and *psbK–psbI* spacer have been popularly used as DNA barcodes in plant worldwide (Janzen *et al.*, 2009) among that *matK* has been seen good for differentiation.

5.2.1 DNA isolation

Among the three loci *matK*, *ITS2* and *trnH-psbA* used in this study, *matK* lies in the chloroplast DNA. The *ITS2* falls in the nuclear DNA. Though separate DNA extraction protocol is available for the extraction of chloroplast DNA, since the total genomic DNA contains nuclear DNA, mitochondrial DNA and chloroplast DNA, isolation of total genomic DNA was followed in this study.

Pale green tender leaves (first or second leaf from tip) were used for DNA extraction from all 27 genotypes of *Momordica* and *Luffa.* The protocol suggested by Rogers and Bendich (1994) yielded good quality DNA. The electrophoresis showed RNA contamination which was eliminated with the RNase treatment. The major problem encountered in the isolation and purification of DNA from plants species is the poor quality of DNA due to protein contamination, endonuclease, polyphenols and other secondary metabolites that directly and indirectly interfere with subsequent enzymatic reactions as reported by Weising *et al.* (1995) and Matasyoh *et al.* (2008). On grinding, the leaves of *Momordica* form the perfect fine powder but at end, DNA turned brown due to phenolic oxidation. The secondary metabolite present in leaf samples were known to be powerful agents,

which damage DNA and proteins (Weising *et al.*, 1995). The addition of antioxidant such as β-mercaptoethanol and sodium metabisulfite in the extraction buffer was found effective by stopping the oxidation of secondary metabolites. Similar results were previously reported in coffee, which is a crop with high amount of phenols (Ram and Sreenath, 1999).

5.2.2 Purification and quantification of DNA

The quality of the DNA was tested by subjecting it to agarose gel electrophoresis as well as spectophotometric methods. DNA was visualized on 0.8 per cent agarose gel under UV light by ethidium bromide staining. A DNA sample was reported as high quality if it had a band of high molecular weight with little smearing and a low amount of RNA (Wettasingf and Peffley, 1998). The DNA extracted showed high amount of RNA and to remove RNA, RNase A was used. Use of RNase A is recommended by several workers (Raval *et al.*, 1998; Wettasingf and Peffley, 1998; Gallego and Martinez, 1996). In the present investigation, the RNase treated DNA sample on electrophoresis showed a high molecular weight DNA by single intact band.

The spectrophotometric method is the recent technology for the precise estimation of DNA quantity. In spectrophotometer method, the ratio of optical density at 260 and 280 nm was worked out to test the quality. The absorbance ratio was calculated as OD at 260/280, for the various samples. Those samples with ratio between 1.8 and 2.0 were considered to be of high quality. Majority of the samples recorded a ratio between 1.8 and 2.0.

5.3 DNA barcoding

An ideal DNA barcode should be routinely retrievable with a single primer pair so the different barcoding primer has been screened. Screening for single or multiple regions appropriate for DNA barcoding studies has been an important research focus of species identification (Liu *et al.*, 2012). The DNA barcodes gives the inter-specific and intra-specific divergence so it must have good PCR recovery sequence with the single primer pair.

Cytochrome c oxidase subunit 1 (*CO1* gene) has been proved distinctive advantages in animal species discrimination. However, it is not favourable in plant species due to the low amount of variation in the genes and the variable structure of mitochondrial genome so for plant, in this study three barcode primer have been chosen *ITS2, trnH-psbA* and *matK* which were already reported in plant but not specific to species (Luo *et al.*, 2010).

The *ITS2* is referred as one of the most important markers in molecular systematic and evolution (Luo *et al.*, 2010). Chloroplast intergenic *psbA-trnH* spacer has also used in plant molecular phylogenetic studies at low taxonomic level. The *psbA-trnH* has good priming sites because it is one of the most rapidly evolving spacers in chloroplast DNA with 75 bp conserved fragments at the ends. In some species it was shown good universality and high amplification but they were not in broad species (Shaw *et al.*, 2005; Yao *et al.*, 2010). The 5′-region of the spacer situated directly after the *psb*A gene is more conserved in length compared to the 3′-region, which has greater sequence variation. However, a comprehensive evaluation with careful sequence preprocessing and statistical testing on the utility of *trnH–psbA* and its combinations as DNA barcodes is lacking. Another primer which has been chosen for this study was *matK,* which was found in chloroplast and it has been identified as a rapidly evolving gene at nucleotide and corresponding amino acid levels (Barthet, 2006). The high number of nucleotide substitutions and length mutations in *matK* has provided a strong phylogenetic result for resolving plant phylogenies at various taxonomic levels. *matK* is the only proposed chloroplast-encoded group II intron maturase. There are 15 genes in the chloroplast that would require a maturase for RNA splicing (Barthet, 2006).

When the primer for *trnH-psbA, ITS2* and *matK* were used to amplify the total genome DNA, it was found that the different combination of *matK* primers used were successful to amplify the locus in all the accessions under study. Hence, it was decided to choose *matK* as the candidate locus for this study.

Different *matK* primer sets, which have been given by Jarret (2008) were employed in this study. All species has been shown amplification with different combinational pairs. The agarose gel profiles for the amplification pattern were observed in the all 25 *Momordica* genotypes and 2 *Luffa* accessions. Some primer was showed multiple bands, in that case only the band having molecular weight nearest to the targeted region 1000 bp was purified using a DNA purification kit. As *matK* length is variable in each genus but in the ORF of *matK* was approximately 1500 bp in most angiosperms (Hilu *et al.*, 1999) and in *Momordica* it is approximately 1200 bp as on today the maximum length of *matK* locus of *Momordica* available in GeneBank is 1207 bp (GenBank: DQ491019.1). This is the reason why the primers described universally or specific for *Momordica matK* will be successful to amplify the bands of a maximum size of 1000bp only. However, in this locus 600–800 bp is variable (Jing *et al.*, 2011) and, this reariability could be well utilized as barcode.

Jing *et al.* (2011) has studied *matK* primers for barcoding in few angiosperms. He had evaluated the entire *matK* around 600 to 800 bp in length and designed universal primers and named matK472F and matK1248R. This region has given the most variable sites, representing the entire *matK* region and also exhibited high amplification rates and quality of sequences. The universality of this primer pair was tested using 58 species from 47 families of angiosperm plants. The primers showed a strong amplification (93.1%) and sequencing (92.6%) successes in the species tested.

Muller *et al.* (2006) also studied the *matK* region for angiosperms evolution. Result has shown the robustness of *matK* in phylogenetic trees than those based on *rbcL* and other primer. Also in different genera like *Tetrastigma* (Fu *et al.*, 2011), Cucurbitales (Zhang *et al.*, 2006), Taxus (Liu *et al.*, 2012), almost all flower plant (Lahaye *et al.*, 2007), medicinal plant (Li *et al.*, 2011), Lemnaceae family plant (Wang *et al.*, 2010), palms (Jeanson *et al.*, 2011) has been studied and *matK* considered as best for the *Momordica* study.

It was recognised that levels of species discrimination in plants with standard DNA barcoding loci are generally lower than those obtained by CO1 (Cytochrome C oxidase subunit 1) in many animal groups (Fazekas *et al.*, 2009). The ability to discriminate between species using barcoding loci has seen more difficult in plants than animals, raising the possibility that plant species boundaries are less well defined. The use of different genetic markers, analytical methods and depths of taxon sampling may lead to complicated comparisons, and the number of species supported as monophyletic, using barcoding markers is higher in animals (>90%) than plants (~70%). This suggests that more than a simple lack of variability limits the species discrimination in plants. Both animal and plant species pairs have variable size gaps between intra- and interspecific genetic distances, but animal species tend to have larger gaps than plants, even in relatively densely sampled genera.

An analysis of different plant genus suggests that hybridization contributes significantly to variation in genetic discontinuity. Barcoding success may be improved in some plant groups by careful choice of markers and appropriate sampling. However, overall fine-scale species discrimination in plants relative to animals may be inherently more difficult because of greater levels of gene-tree paraphyly. This is in the part due to the lower rate of nucleotide substitution in the plastid genome, but also due to other reasons including hybridization, polyploidy, speciation via breeding system transitions, species defined on very narrow taxon concepts, large ancestral population sizes, and low levels of intraspecific gene flow for plastid markers (Kress *et al.*, 2009; Hollingsworth *et al.*, 2011).

These issues are not evenly distributed among all plant groups; therefore, it is expected that resolution at the species level will be reasonably good in some groups and quite poor in others. In floristic contexts where geographical limitation usually restricts the number of closely related species, rates of species discrimination are expected to be greater (Ivanova *et al.*, 2006).

5.3.1 PCR product sequencing

A total of 25 sequences of the *Momordica* species and 2 sequences of *Luffa* species were amplified with selected primers and sequenced by outsourcing (Sci Genome Lab. Pvt. Ltd., Cochin). With forward and reveres sequencing was done for all the samples. The average length of sequence was 806 which show good sequencing success. The lower length is the major disadvantage of the next generation sequencing platform. A good rate of recovery (67% as compared to average sequence length) was obtained. In the accession where the reveres reads have failed the total sequence length generated was low.

5.3.2 Species identification based on phylogenetic tree analysis

The genetic similarity was determined for all of 27 genotypes belonging to seven *Momordica* spp. and two *Luffa* spp., based on *matK* sequences. The phenogram based on 27 genotypes was grouped into III clusters. Cluster I contain *M. balsamina*, *M. cymbalaria*, *M. charantia*, *M. sahydrica*, *M. dioica* and *Luffa* spp.

The sequence information provides clear genetic variation among the *M. balsamina*, *M. cymbalaria*, *M. charantia*, *M. sahydrica*, *M. dioica* and *Luff* spp. *Momordica* (n = 11) and *Luffa* (n = 13) are different in their chromosome number and morphology, which was evident in this study also. *M. charantia* and *M. balsamina* which are monocious in nature, clustered together in sub group as both have high bivalent frequency with normal meotic cycle in the hybrid progeny (Singh, 1990).

The *M. cymbalaria* and *Luffa* spp. came in one sub cluster as both were monocious in nature (Bharathi *et al.*, 2011), Few scientist consider *M. cymbalaria* as a sub species of *Momordica* (De Wilde and Duyfjes, 2002). The morphological and cytological study (Bharathi *et al.*, 2011) shows its distinctness from other *Momordica* species. Also *M. cymbalaria* is different from others by chromosome number (Bharathi *et al.*, 2011). The study by Schaefer and Renner (2009) had

shown the close relationships among African species such as *M. humilis* and *M. bovinii*, but the clustering of a *M. cymbalaria* was different.

The *M. sahydrica* Annamalai type and its wild relative had fallen in two different sub clusters and *M. dioica* were forming different sub cluster. These results are in accordance with the previous reports by Bharathi *et al*, (2011).

The *M. subangulata* and *M. dioica* have been grouped in separate clusters, showing the accurate diversity among the groups. As *M. subangulata* is an amphidiploids it clusters with *M. dioica* and this look perfectly because *M. dioica* is act as putative parents for *M. subangulata* (Bharathi *et al*., 2010).

M. sahyadrica alone has constituted the cluster I and in cluster II *M. dioica, M. charantia muricata* and *M. sahyadrica* were seen. Within *Momordica*, the highest degree of diversity was observed among *M. charantia* and *M. cochinchinesis* and this is already reported by Schaefer and Renner (2009) in which they studied plastid, mitochondria and nuclear DNA.

Although *matK* showed good identification rate at the genus level in phylogenetic tree, some species such as *M. sahyadrica* remained unresolved.

5.3.3 Genetic divergence within and between species

The inter specific distances were calculated using K2P (Kimura 2 parameter). The range of pair wise dissimilarity was 0.0032 to 0.04 with an average of 0.0134. The dissimilarity between *Luffa* and *Momordica* ranged from 0.0096 to 0.0458 (*Luffa* and *M. sahyadrica*) which had shown maximum distance. Among the *Momordica* species, maximum dissimilarity was observed between wild *M. sahydrica* and *M. balsamina* (0.0458) followed by *M. subangulata* and *M. balsamina* (0.0291). Same results were reported by Bharathi *et al.* (2011).

5.3.4 Barcoding gap and efficiency assessment

A robust DNA barcode should have separate and non-overlapping genetic variations among species. If all species share the identical sequences, as

straightforward failures, but when all individuals of a species show change in some sequence, it is treated as successful discrimination.

A total of 27 sequences were aligned and the nucleotide change has been observed. The distributions of intra- versus inter- specific and genotype divergence in the *matK* barcode was examined. In Fig. 18 the results demonstrated that the intra/inter-specific variation of *matK* exhibited distinct gaps. The inter-specific divergences for *matK* were significantly higher than their corresponding intra-specific variations.

As the sequences from each species were unique except with some conserved region, *matK* is found to be a potentially successful locus for barcoding. The *matK* locus is previously reported as the most successful locus in many plant (Jing *et al.*, 2011).

BLAST was used for assessment of correct discrimination using *matK* barcode. The results based on BLAST have indicated that *matK* give a high identification efficiency of 91 % at the species level, followed by 100% in genus level. These analyses are further supported by the work by Jing *et al.* (2011).

5.3.5 Identification of conserved region and designing primers for *Momordica*

The *matK* gene sequences were aligned using ClustalOmega (Sievers *et al.*, 2011). 982-bp data subsets were generated by examining the entire length and analyzed with primer 3 software (Untergasser *et al.*, 2012). The primers designed have been successful to amplify the *matK* locus from the *Momordica* accession under study. Due to the lack of barcoding studies in *Momordica*, the primers were not available and the strategy followed in this study was to attempt the capability of universal primers. The identification of new *matK* barcoding primers will quicken the process in this genus. The size of the amplified product was found to be mainly 1000 bp which is a reasonably good length to identify the barcoding gaps in this genus since the maximum length of *matK* in *Momordica*, reported so far in 1207 bp.

5.3.6 Utilization of the developed barcodes in *Momordica* species identification

The species specific barcodes generated from this study could be immediately used for the species allocation of any *Momordica* accession with taxonomical controversy. *matK* locus may be amplified using the primers and sequence could be generated. Based on the BLAST with the basic sequence of 982 bp generated and deposited in NCBI, the barcoding gaps could be determined in the sample and from this, the species allocation could be easily done.

5.3.7 Future line of work

The superiority of other barcoding loci, such as *rbc L, trnH-psbA, ITS2* etc. and their combination have to be worked out in *Momordica*. In the present study, the species *M. c. muricata, M. sahydrica* and *M. dioica* were not successfully differentiated using *matK* locus. Using other loci, barcoding for this species should be generated.

Summary

6. SUMMARY

The study entitled "DNA barcoding in *Momordica* spp." was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Vellanikkara, during the period 2012-2014. The objective of the study was to barcode the different member of the genus *Momordica,* available in India, using the identified locus in the genome. India has 7 species of *Momordica* species and the species selected for this study included *Momordica charantia* L. cv. *charantia, Momordica charantia* cv. *muricata* (wild) Chakrav, *Momordica dioica* Roxb., *Momordica sahyadrica* Joseph and Antony, *Momordica balsamina, Momordica cochinchinesis* Spreng. (Gac.), *Momordica subangulata* Blume ssp. *renigera* (G. Don) de wild and *Momordica cymbalaria.* Two lines of *Luffa acutangula* were also included in the study. Leaf samples were collected from different locations like National Bureau of Plant Genetic Resources Regional Station (NBPGR) Thrissur, Agricultural Research Station, Munnuty, College of Horticulture and Research Institute, Periyakulum, and Krishi Vigyan Kendra Thrissur.

The salient findings of the study are stated below:

1. The protocol for genomic DNA isolation was standardised. The protocol suggested by Rogers and Bendich (1994) with slight modification (high β mercaptoethenol) was found to be the most appropriate for isolation of DNA from different *Momordica* genotypes. The RNA contamination was completely removed through RNase treatment.

2. The quality and quantity of DNA was analyses by NanoDrop^{R} ND-1000 spectrophometer. The absorbance ratio ranged from 1.80-1.89, which indicated good quality DNA and the recovery was high. This DNA was suitable for PCR amplification analysis.

3. Primer sets for 3 common barcoding loci *matK, ITS2* and *trnH-psbH* were used in this study. All the primer sets were seen to be differentially

compatible to generate the bands in various *Momordica* and *Luffa* accessions.

4. *matK* was found to be good for the barcoding in *Momordica* since the size of the locus was 1.2 kb against 600 kb in *ITS2* and 300 bp in *trnH-psbA*.

5. Only few barcode primers have yielded single and distinct bands, representing the *matK*. When more than one bands were produced, distinct band at 1 kb size was eluted and used for sequencing.

6. The *matK* gene was amplified using various primer combinations and sequenced by outsourcing.

7. A total of 25 sequences of the *Momordica* species and 2 *Luffa* species have been submitted to the NCBI GenBank.

8. The 27 sequences were used to construct the pylogenetic tree using the software ClustalW. The pylogenetic tree generated in the study clearly discriminated all the 27 genotypes and resulted in definitive grouping among different species of *Momordica* and *Luffa*. The results corresponded well with their known phenetic relationships as well as morphological, cytological and taxonomic classifications.

9. The genetic distance was computed for all combinations of 27 genotypes belonging to seven *Momordica* spp. and two *Luffa* spp. based on K2P model using MEGA 4.0 software which was recommended by the Consortium of Barcode of Life.

10. The distributions of intra-versus inter-specific variability were compared using DNA barcoding gap, which differentiated the *Momordica sahyadrica*, *Momordica dioica* and *Momordica charantia* cv. *muricata* species clearly.

11. The species identification success rate was calculated using BLAST using all *matK* sequences as query sequences, which indicated that *matK* is powerful to yield up to 95 per cent efficiency to differentiate the species and at genus level with 100 per cent efficiency.

12. The conserved sequences in the generated base sequence was identified using the Software ClustalW and two primer for the genus *Momordica* was designed using Primer 3 software, among which primer set 1 had shown perfect amplification in all genotype.

Future line of work includes; a universal barcode system for *Momordica* species should be experimented with other barcoding loci, such as *rbcL*, *trnH-psbA*, *atpF–atpH* spacer, *rpoB* gene, *rpoC1* gene, *psbK–psbI* spacer, *trnH–psbA* spacer, *ITS2* etc. The barcodes for *M. c. muricata*, *M. sahydrica* and *M. dioica* have to be developed, and has study with more genotypes. The primer developed species specific should have to be used for barcoding the remaining 52 species.

References

7. REFERENCES

Anvarkhah, S., Khajeh-Hosseini, Hassan Rashed, M., Panah, A. E. and Hashemi, H. 2013. Identification of three species of genus Allium using DNA barcoding. *Int. J. Agril. Crop Sci.* **5**(11):1195-1203.

Arriel, N. H. C., Mauro, A. O. D., Arriel, E. F., Unêda-Treviso, S. H., Costa, M. M., Bárbaro, I. M. and Muniz, F. R. S. 2007. Genetic divergence in sesame based on morphological and agronomic traits. *Crop Breed. Appl. Biotechnol.* 7: 253-261.

Barthet, M. M. 2006. Expression and function of the chloroplast-encoded gene *matK*. Ph. D thesis, Virginia Polytechnic Institute and State University, Virginia, 213p.

Beckman, J. S. and Soller, M. 1986. Restriction fragment length polymorphisms in plant genetic improvement. In: Miflin, B. J. (ed.), *Oxford Surveys of Plant Molecular and Cell Biology*. Vol III, Oxford Press, Oxford, pp.196-250.

Behera, T. K., Singh, A.K. and Staub, J. E. 2008. Comparative analysis of genetic diversity in Indian bitter gourd (*Momordica charantia* L.) using RAPD and ISSR markers for developing crop improvement strategies. *Hort. Sci.* **115**: 209-217.

Berthaud, J. and Charrier, A. 1988. Genetic resources of Coffea. In: Clarke, R. J. and Macrae, R. (eds.), *Coffee Agronomy*. Elsevier Applied Science, London & New York, pp.1-42.

Bhandari, M. M. 1978. *The flora of the Indian desert*. Scientific Publishers, Jodhpur, 435p.

Bharathi, L. K., Munshi, A. D., Behera, T. K., Chandrashekaran, S., Kattukunnel, J. J., Das, A. B. and Vishalnath. 2010. Cytomorphological evidence for

segmental allopolyploid origin of Teasle gourd (*Momordica subangulata* subsp. *renigera*). *Euphytica.***176**: 79-85.

Bharathi, L. K., Munshi, A. D., Behera, T. K., Vinod, Joseph, J. K., Das, A. B., Bhat, K. V. and Sidhu, A. S. 2012. Production and preliminary characterization of inter-specific hybrids derived from *Momordica* species. *Current Sci.* **103**(2):178-184.

Bharathi, L. K., Munshi, A. D., Vinod, Chandrashekaran, S., Behera, T. K., Das, A. B., Joseph, J. K. and Vishalnath. 2011. Cyto-taxonomical analysis of *Momordica* L. (Cucurbitaceae) species of Indian occurrence. *J. Genet.* **90**: 21-30.

Bhat, K. G. 2003. *The Flora of Udipi.* Index Naturalist, Udipi, 914p.

Blatter, E. 1919. *Flora Arabica.* Bishen Singh Mahendrapal Singh Publishers, Dehradun, 200p.

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R. and Abebe, E. 2005. Defining operational taxonomic units using DNA barcode data. *Phil. Trans. Roy. Soc. Lond. B.* **360**: 1935-1943.

Bootprom, N., Songsri. P., Suriharn, B., Chareonsap, P., Sanitchon, J. and Lertrat, K. 2012. Molecular Diversity Among Selected *Momordica Cochinchinensis* (Lour.) Spreng Accessions Using RAPD Markers. *J. Breed. Gen.* **44** (2) 406-417.

Bostein, D., Raymond, W. L., Skolnick, M. and Ronald, W. D. 1980.Construction of a genetic linkage map in man using Restriction Fragment Length Polymorphisms. *Am. J. Hum. Genet.* **32**: 314-331.

Brown, A. H. D. 1978. Isozymes plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* **52**:145-157.

Brown, A. H. D. 1990. The role of isozyme studies in molecular systematics. In: Ladiges, P. Y. and Martinelli, L. W. (ed.), *Plant Systematics in the Age of Molecular Biology*. CSIRO Publisher, Melbourne, pp. 39-46.

Bucklin, A., Hopcroft, R. R., Kosobokova, K. N., Lisa, M. N., Brian, D. O., Robert, M. J. and Christopher, J. S. 2009. DNA barcoding of Arctic Ocean holozooplankton for species identification and recognition. *Deep-Sea Res. Plant II* **57(1)**: 40-48.

Caetano-Anolles, G., Bassam, B. J. and Gresshoff, P. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Mol. Biotechnol.* **9**: 553-557.

Chakravarthy, H. L. 1982. *Fascicles of Flora of India : Cucurbitaceae*. Botanical Survey of India, Kolkata, 136p.

Chakravarty, H. L. 1959. *Monograph of Indian Cucurbitaceae*. Botanical Survey of India, Kolkata, 81p.

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haider, N. and Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**: 1889-1895.

Chen, S., Yao, H., Han, J., Liu, C. and Song, J. 2010. Validation of the *ITS2* region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE* **5(1)**: 8613.

Clarke, C. B. 1879. Cucurbitaceae. In: Hooker, J. D. (ed.), *The flora of British India*. L. Reeve & Co., London, pp. 604- 635.

Cooke, T. 1908. *The Flora of the Presidency of Bombay*. Taylor and Francis, London, 400p.

Dalamu, Behera, T. K., Gaikwad, A. B., Saxena, S., Bharadwaj, C. and Munshi, A. D. 2012. Morphological and molecular analyses define the genetic diversity of Asian bitter gourd (*Momordica charantia* L.). *Aust. J. of Crop Sci.* **6**(2): 261-267.

Dayrat, B. 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc.* **85**: 407-415.

De Candolle. 1828. Prodromus systematis naturalis regni vegetabilis. Treuttel, Paris, pp. 1824-1873.

De Groot, G. A., During, H. J., Maas, J. W., Schneider, H. and Vogel, J. C. 2011. Use of *rbcL* and *trnL-F* as a two-locus DNA barcode for identification of European Ferns: An Ecological Perspective. *PLoS ONE* **6(1)**: 16371.

De Wilde, W. J. J. O. and Duyfjes, B. E. E. 2002. Synopsis of Momordica (Cucurbitaceae) in South East Asia and Malaysia. *Bot. Zhurn.* **57**(3):132-148.

Dey, S. S., Behera, T. K., Munshi, A. D., Sirohi, P. S. 2006a. Studies on genetic divergence in bitter gourd (*Momordica charantia* L.). *Indian J. Hort.* **64**: 53–57.

Dey, S. S., Singh, A. K., Chandel, D. and Behera, T. K. 2006b. Genetic diversity of bitter gourd (*Momordica charantia* L.) genotypes revealed by RAPD markers and agronomic traits. *Sci. Hort.* **109**: 21-28.

Dje, Y., Tahi, G. C., Zoro, B. I. A., Malice, M., Baudoin, J. P. and Bertin, P. 2006. Optimization of ISSR marker for African edible seeded *Cucurbitaceae* species genetic diversity analysis. *Afr. J. Biotechnol.* **5**: 83-87.

Dudnikov, A. J. 2003. Allozymes and growth habit of *Aegilops tauschii*: genetic control and linkage patterns. *Euphytica.* **129**: 89-97.

Dunning, L.T. and Savolainen, V. 2010. Broad-scale amplification of *matK* for DNA barcoding plants, a technical note. *Bot. J. Linnean Soc.* **164**: 1–9.

Ebihara, A., Nitta, J. H. and Ito, M. 2010. Molecular species identification with rich floristic sampling: DNA Barcoding the Pteridophyte Flora of Japan. *PLoS ONE* **5**(12): 15136.

Ellis, J. L. 1987. *Flora of Nallamalai.* Botanical Survey of India, Kolkata, pp. 193.

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M. and Barrett, S. C. H. 2008. Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. *PLoS ONE* **3**(7): 2802-2814.

Fazekas, A. J., Kesanakurti, P. R., Burgess, K. S., Percy, D. M., Graham, S. W., Barrett, S. C., Newmaster, S. G., Hajibabaei, M. and Husband, B. C. 2009. Are plant species inherently harder todiscriminate than animal species using DNA barcoding markers? *Mol. Ecol. Res.* **9**:130–139.

Floyd, R., Abebe, E., Papert, A. and Blaxter, M. 2002. Molecular barcodes for soil nematode identification. *Mol. Ecol.* **11**(4): 839-850.

Fu, Y., Jiang, W. and Fu, C. 2011. Identification of species within *Tetrastigma* (Miq.) Planch.(Vitaceae) based on DNA barcoding techniques. *J. Syste. Evol.* **49** (3): 237–245.

Gaikwad, A. B., Behera, T. K., Singh, A. K., Chandel, D., Karihaloo, J. L. and Staub, J. E. 2008. AFLP analysis provides strategies for improvement of bitter gourd (*Momordica charantia*). *Hort. Sci.* **43**:127-133.

Gallego, F. J. and Martinez, I. 1996. Molecular typing of rose cultivars using RAPD. *J. Hortic. Sci.* **71**(6): 901-908.

Gamble, J. S. and Fischer, C. E. C. 1919. *Flora of the Presidentry of Madras.* Botanical Survey of India, Kolkata, pp. 410-942.

Gottlieb, L. D. 1981. Electrophoretic evidence and plant populations. *Prog. Phytochem.* 7: 1-46.

Gregory, T. R. 2005. DNA barcoding does not compete with taxonomy. *Nature* **434**: 1067-1067.

Gwanama, C., Labuschagne, M. T. and Botha, A. M. 2000. Analysis of genetic variation in *Cucurbita moschata* by random amplified polymorphic DNA (RAPD) markers. *Euphytica.* **113**: 19-24.

Hebert, P. D. N, Ratnasingham, S. and de Waard, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. Biol. Sci.* **270**: 96-99.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. and Francis, C. M. 2004. Identification of birds through DNA barcodes. *PLoS Biol.* **2**(10): 1657-1663.

Hilu, K. W. and Alice, L. A. 1999. Evolutionary implications of *matK* indels in Poaceae. *Am. J. of Bot.* **86**:1735-1741.

Hollingsworth, M., Clark, A. A., Forrest, L., Richardson, J., Pennington, R. T., Long, D. G., Cowan, R., Chase, M. W., Gaudeul, M. and 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.Res.* **9**:439–457.

Hollingsworth, P. M., Graham, S. W. and Little, D. P. 2011. Choosing and using a plant DNA barcode. *PLoS ONE* **6**(5):19254- 19267

Horejsi, T. and Staub, J. E. 1999. Genetic variation in cucumber (*Cucumis sativus* L.) as assessed by random amplified polymorphic DNA. *Genet. Resour. Crop Evol.* **46**: 337-350.

Hu, J. and Quiros, C. F. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. *Pl. Cell Reporter* **10**: 505-511.

Ivanova, N. V., deWaard, J. R., Hajibabaei, M. and Hebert, P. D. N. 2006. *Protocols for high volume DNA barcode analysis, Draft submission to DNA working group*. Consortium for the Barcode of Life, Canada, 24p.

Janzen, D. H. 2009. A DNA barcode for land plants. *PNAS*. **106**(31): 12794-12797.

Jarret, R. L. 2008, DNA Barcoding in a crop genebank: The *Capsicum annuum* species complex. *The Open Biol. J.* **1**: 35-42.

Jeanson, M. L., Labat, J. N. and Damon, P. L. 2011. DNA barcoding: a new tool for palm taxonomists. *Annals. Bot.* 1-7.

Jeffery, C. 1980a. The Cucurbitaceae of the Indian Subcontinent. *Rew. Bull.* **34**(4): 789-809.

Jeffrey, C. 1980b. A Review of the Cucurbitaceae. *J. Linn. Soc. Bot.* **81**:233-247.

Jing, Y. U., Jian-Hua, X. and Shi-Liang, Z. 2011. New universal *matK* primers for DNA barcoding angiosperms. *J. System. Evol.* **49**(3): 176-181.

Jobst, J., Klaus, K. and Vera, H. 1998. Molecular evolution of the internal transcribed spacers (*ITS1* and *ITS2*) and phylogenetic relationships among species of the family Cucurbitaceae. *Mol. Phylogen. Evol.* **9**: 204-219.

Joseph, J. K. 2005. Studies on ecogeography and genetic diversity of the genus *Momordica* L. in India. Ph.D. thesis, Department of Botany, Mahatma Gandhi University, Kottayam, 315p.

Joseph, J. K. and Antony, V. T. 2007. *Momordica sahyadrica* sp. nov. (Cucucrbitaceae) an endemic species of Western Ghats of India. *Nord. J. Bot.* **24**: 539-542.

Joseph, J. K. and Antony, V. T. 2011. Developing Descriptors for Dioecious *Momordica* Species for Germplasm Characterization and Evaluation. *Indian J. Plant Genet. Resour.* **24**(1):.31-39

Joseph, J. K., Antony, V. T. and Roy, Y. C. 2007. On the occurrence, distribution and taxonomy of *Momordica subangulata* Blume ssp. *renigera* (G. Don) de Wilde in India. *Genet. Resour. Crop Evol.* **54**: 1327-1332.

Joshi, S. P., Gupta, V. S., Aggarwal, R. K., Ranjekar, P. K. and Brar, D. S. 2000. Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus Oryza. *Theor. Appl. Genet.* **100**:1311-1320.

Joshi, V., Gautham, P. L., Bhabmal, G., Sharma, D. and Kochhar, S. 2002. Conservation and use of underutilization crops an Indian perspective. In: Egels, J. M. M., RamanathaRao, V., Brown, A. H. D. and Jackson, M. T. (Ed.), *Managing Plant Genetic Diversity*. IPGRI, Rome.

Kanjilal, U. 1938. *Flora of Assam*. Omsons Publications, Assam, 2334 p.

Karp, A. S., Kresovich, K. V., Bhat, W. G. and Ayad, T. H. 1997. *Molecular tools in plant genetic resources conservation: A Guide to the technologies.* IPGRI Technical Bulletin No.2, International Plant Genetic Resources Institute, Rome, Italy. 47p.

Karp, A., Seberg, O. and Buiatti, M. 1996. Molecular techniques in the assessment of botanical diversity. *Ann. of Bot.* **78**:143-149.

Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A. and Janzen, D. H. 2005. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci.* **102**: 8369-8374.

Kressa, W. J., Ericksona, D. L., Jonesb, F. A., Swensond, N. G., Perezb, R., Sanjurb, O. And Berminghamb, E. 2009. Plant DNA barcodes and a

community phylogeny of a tropical forest dynamics plot in Panama. *Proc. Nat. Acad. Sci.* **106**:18621–18626

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T. G. and Savolainen, V. 2007. DNA barcoding the floras of biodiversity hotspots. *PNAS* **105**(8): 2923-2928

Lee, S. J., Shin, J. S., Park, K. W. and Hong, Y. P. 1996. Detection of genetic diversity using RAPD-PCR and sugar analysis in watermelon [*Citrullus lanatus* (Thumb) Mansf.] germplasm. *Theor. Appl. Genet.* **92**: 719-725.

Levi, A., Thomas, C. E., Newman, M., Reddy, O. U. K., Zhang, X. and Xi, Y. 2004. ISSR and AFLP markers differ among American watermelon cultivars with limited genetic diversity. *J. Am. Soc. Hort. Sci.* **129**: 553-558.

Levi, A., Thomas, C. E., Simmons, A. M. and Thies, J. A. 2005. Analysis based on RAPD and ISSR markers reveals closer similarities among *Citrullus* and *Cucumis* species than with *Praecitrullus fistulosus* (Stocks) Pangalo. *Genet. Resour. Crop Evol.* **52**: 465-472.

Li, M., Cao, H., But, P. P. and Shaw, P. 2011. Identification of herbal medicinal materials using DNA barcodes. *J. Sys. Evol.* **49** (3): 271–283

Lin, D., Hubbes, M. and Zsuffa, L. 1993. Differentiation of poplar and willow clones using RAPD fingerprints. *Tree physiol.* **14**:1097-1105.

Lingaiah, H. B., Uthaiah, B. C. and Herle. 1993. Performance of bitter gourd cultivars in costal Karnataka. *Curr. Res.* **22**(1):16.

Linnaeus, C. 1800. *Species Plantarum.* Laurentius Salvius, Sweden, 1200p.

Liu, Y., Yan, H. and Xue-Jun, G. E. 2010. Evaluation of 10 plant barcodes in Bryophyta (Mosses). *J. of Syst. and Evol.* **48** (1): 36–46.

Liu, Z., Zeng, X., Yang, D., Chu, G., Yuan, Z. and Chen, S. 2012. Applying DNA barcodes for identification of plant species in the family Araliaceae. *Gene*. **499**:76-80.

Luo, K., Chen, S., Chen, K., Song, J., Yao, H., Ma, X., Zhu, Y., Pang, X., Yu, H., Li, X. and Liu, Z. 2010. Assessment of candidate plant DNA barcodes using the Rutaceae family. *Sci. China Life Sci.* **53**:701–708.

Matasyoh, G. L., Francis, N. W., Miriam, G. K., Anne, W., Thairu, M. and Titus, K. M. 2008. Leaf storage conditions and genomic DNA isolation efficiency in *Ocimum gratissimum* L. from Kenya. *Afr. J. Biotechnol.* **7(5)**: 557-564.

Mathew, K. M. 1981. *The Flora of the Tamil Nadu Carnatic.* Vol.3. Repr. Herbarium, St. Joseph's College, Tiruchirappally, 648p.

Mathew, K. M. 1983. *Illustration on the Flora of the Tamil Nadu Carnatic*. Vol.3. Repr Herbarium, St. Joseph's College, Tiruchirappally, 301p.

McClean, P. L., Ewing, J.,Lince, M. and Grafton, K. 1994. Developments of RAPD map of *Phaseoolus vulgaris* L. *Theor. Appl. Genet.* **37**:79-80.

Miller, S. E. 2007. DNA barcoding and the renaissance of taxonomy. *Proc. Natl. Acad. Sci.* **104**: 4775-4776.

Montarroyos, A. V. V., de Lima, M. A. G., dos Santos, E. O. and de França, J. G. E. 2003. Isozyme analysis of an active cassava germplasm bank collection. *Euphytica.* **130**:101-110.

Morell, M. K., Peakall, R., Appels, L. R. and Preston, H. L. 1995. DNA profiling techniques for plant variety identification. *Aust. J. Exp. Agri.* **35**: 807-819.

Müller, K. F., Borsch, T., Hilu, K. W. 2006. Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: contrasting *matK*, *trnT*-F, and *rbcL* in basal angiosperms. *Mol. Phylogenet. Evol.* **41**(1): 99-117.

Naik, V.N. 1979. *Flora of Osmanabad.* Venus Publishers, Aurangabad, 146 p.

Narasimhan, P. L. and Sharma, B. D. 1991. *Flora of Nashik District*, Botanical Survey of India, Calcutta, 229p.

Newmaster, S. G., Fazekas, A. J. and Ragupathy, S. 2006. DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Canad. J. Bot.* **84**: 335–341.

Nijoroge, G. N. and Newton, L. E. 2002. Ethnobotany and distribution of wild genetic resources of the family Cucurbitceae in the Central highlands of Kenya. *Pl. Genet. Resour. Newsletter.* **132**:10-16.

Oliver, D. 1979. *Flora of Tropical Africa.* vol3. AJ Reprint Agency, New Delhi, 596 p.

Oommachan, M. 1977. *Flora of Bhopal (Angiosperms)*. J. K. Jain Brothers Publishers, Bhopal, 187p.

Paillard, M., Lashermes, P. and Petiard, V. 1996. Construction of a molecular linkage map in coffee. *Theor. Appl. Genet.* **93**: 41-47.

Pallithanam, J. M. 2001. *A Pocket Flora of the Sirumalai Hills South India.* Rapinat Herbarium, St. Joseph's college, Tiruchirappally, 360p.

Pandit, M. K. and Acharya, S. 2008. Potential of minor cucurbits for sustainability and livelihood support in west Bengal. *J. Agrofor. Environ.* **2**(2): 209-212.

Parker, P. G., Snow, A. A., Schug, M. D., Booton, G. C. and Fuerst, P. A. 1998. What molecules can tell us about populations: choosing and using molecular markers. *Ecol.* **79**: 361–382.

Parveen, I., Singh, H. K., Raghuvanshi, S., Pradhan, U. C. and Babbar, S. B. 2012. DNA barcoding of endangered Indian Paphiopedilum species. *Mol. Ecol. Res.***12**: 82-90.

Paul, S., Bandyopadhyay, S., Acharyya, P. and Raychaudhuri, S. S. 2010. Studies on genetic diversity of twelve accession of *Momordica charantia* L. using morphological, RAPD and SCAR markers. *Asian J. of Plant Sci.* **9**(8): 471-478.

Raju, R. R. and Pullaiah, T. 1995. *Flora of Kurnool (A.P.).* Bishenpal Mahendepal Singh, New Delhi, 223p.

Ram, A. S. and Sreenath, H. C. 1999. A method for isolation and amplification of coffee DNA using random octamer and decamer primers. *J. plantn. crops.* **27**(2): 125-130.

Ramachandran, V. S. and Nair, V. J. 1988. *Flora of Cannanore*. Botanical Survey of India, Kolkatta, 600p.

Rao, R. S. 1985. *Flora of Goa, Damen, Dadra and Nagarhaveli*. Vol.I, Botanical Survey of India, Kolkatta, 546p.

Rasul, M. G., Hiramatsu, M. and Okubo, H. M. 2007. Genetic relatedness (diversity) and cultivar identification by randomly amplified polymorphic DNA (RAPD) markers in teasle gourd (*Momordica dioica* Roxb.). *Sci. Hort.* **111**:271–279

Ratnasingham, S., Hebert, P. D. N. 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Mol. Ecol.*7:355-364.

Raval, R. R., Lau, K. W., Tran, M. G., Sowter, H. M., Mandriota, S. J., Li, J. L., Pugh, C. W., Maxwell, P. H., Harris, A. L. and Ratcliffe, P. J. 1998. Contrasting properties of Hypoxia-Inducible Factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma. *Mol. Cell Biol.* **25**: 5675-5686.

Rhodes, M. J. C. 1678. The climacteric and ripening of fruits In: Hulme, A.C. (ed.), *The Biochemistry of Fruits and their Products.* Academic Press, New York, pp.521-533.

Ritschel, P. S., Lins, T. C. D. L., Tristan, R. L., Buso, G. S. C., Buso, J. A. and Ferreira, M. E. 2004. Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). *BMC Plant Biol.* 4:9–23.

Robinson, R. W. and Decker-Walters, D. S. 1997. *Cucurbits.* CAB International-Wallingford, Oxford, UK, 226p.

Rogers, S. O. and Bendich, A. J. 1994. Extraction of total cellular DNA from plants, algae and fungi. *Plant Mol. Biol. Manual.* **D1**: 1-8

Roxburgh, W. 1832. *Indica or Descriptions of Indian Plants.* Today and Tomorrow Publishers, New Delhi, 875p.

Saiki, R. K., Gelfand, D. H., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermo stable DNA polymerase. *Sci.* **239**: 487-491.

Saldhana, C. J. 1985. *Flora of Karnataka.* Oxford & IBH publishers, New Delhi, 535p.

Saldhana, C. J. and Nicolson, D. H. 1976. *Flora of Hassan District, Karnataka.* India Amerind Publishing Co.Pvt.Ltd. , New Delhi, 915p.

Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. *Molecular Cloning A Laboratory Mannual.* Academic Press, New York, USA, 1322p.

Sands, W. N. 1928. The Bitter Cucumber of Paris. *Malayan Agric. J.* **16**:32-39.

Santapau, H. 1953. *Flora of Khandala on the Western Ghats of India.* Botanical Survey of India, Kolkata, 372p.

Sasidharan, N. and Sivarajan,V. V. 1996. *Flowering Plants of Thrissur Forest (Western Ghats, Kerla,India).* Scientific Publishers, Jodhpur, 204p.

Schaefer, H., Renner, S. S. 2009. A three-genome phylogeny of *Momordica* (Cucurbitaceae) suggests seven returns from dioecy to monoecy and recent long-distance dispersal to Asia. *Mol. Phylogenet. Evol.* **54**: 553–560.

Schindel, D. E., Miller, S. E. 2005. DNA barcoding a useful tool for taxonomists. *Nat.* **435**:17-17.

Schlötterer, C. 2004. The evolution of molecular markers-just a matter of fashion. *Nat. Rev. Genet.* **5**: 63–69.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Kunsiri C. Siripun, Charles T. Winder, Edward E. Schilling and Randall L. Small 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92: 142–166.

Shaw, J., Lickey, E. B., Schilling, E. E. and Small, R. L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Amer. J. Bot.* **94**: 275–288.

Shetty, B. V. and Singh, V. 1987. *Flora of Rajastan.* Botanical Survey of India, Kolkatta, 348p.

Sievers, F., Wilm, A., Dineen, D. G., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. and Higgins, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539

Singh, A. K. 1990. Cytogenetics and evolution in the Cucurbitaceae. In: Bates, D. M., Robinson, R. W. and Jeffrey, C. (eds.), *Biology and utilization of*

Cucurbitaceae. Comstock Publishing Associates, Cornell University Press, Ithaca, New York and London, pp 10–28.

Singh, N. P., Chauhan, A. S. and Mandal, M. S. 2002. *Flora of Manipur*. Botanical Survey of India, Kolkatta, pp.417-418.

Sivarajan, V. V. and Mathew, P. 1997. *Flora of Nilambur*, Bishen Singh Mahendra Pal Singh, pp.293-294.

Srivastava, T.N. 1976. *Flora Gorakpurensis*. Today & Tomorrow Publisher, New Delhi, 149p.

Sureja, A. K., Sirohi, P. S., Behera, T. K. and Mohapatra, T. 2006. Molecular diversity and its relationship with hybrid performance and heterosis in ash gourd [Benincasahispida (Thumb.)Cogn.]. *J. Hort. Sci. Biotchnol.* **81**: 33-38.

Tanksley, S. D., Young, N. D., Paterson, A. H. and Bonierbale, M. W. 1989. RFLP mapping in plant breeding. *Nucleic Acids Res.* 7: 257-264.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* 17:6463-6471.

Tommerup, I. C., Barton, J. E. and O'brein, P. A. 1998. Reliability of RAPD fingerprinting of three basidiomycetes fungi, *Laccaria*, *Hydnangium* and *Rhizoctonia*. *Mycol. Res.* **99**:179-186.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G. 2012. Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* **40**(15): 115.

Vajravelu, E. 1990. *Flora of Palghat District*. Botanical Survey of India, Kolkatta, 646p.

Vos, R., Hogers, R., Bleeker, M., Reijans, M., Homes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: A new concept for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407-4414.

Walters, T, W. 1989. Historical overview on Domesticated plants in China with special emphasis on the Cucurbitaceae. *Econ. Bot.* **43**(3):297-313.

Wang, W., Wu, Y., Yan, Y., Ermakova, M., Kerstetter, R. and Messing, J. 2010. DNA barcoding of the Lemnaceae, a family of aquatic monocots. *BMC Plant Biol.* **10**: 205-216.

Webb, K. E., Barnes, D. K. A., Clark, M. S. and Bowden, D. A. 2006. DNA barcoding: A molecular tool to identify Antarctic marine larvae. *Deep-Sea Res. Part II* **53**:1053-1060.

Weising, K., Atkinson, G. and Gardner, C. 1995. Genomic fingerprinting by microsatellite-primed PCR: a critical evolution. *PCR Methods Appl.* 4(5): 249-255.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. 2005. *DNA Fingerprinting in plants: principles, methods and applications* (2nd Ed). CRC Press, Taylor and Francis Group, New York, 472p.

Welsch, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18**:7213-7218.

Welsch, J., Honeycutt, R. J., McClelland, M. and Sobral, B. W. S. 1991. Parentage Determination in maize hybrids using the arbitrarily primed polymerase chain reaction (AP-PCR). *Theor. Appl. Genet.* **82**:473-476.

Wettasingf, R. and Peffley, E. B. 1998. A rapid and efficient extraction method for onion DNA. *Plant Breed.* **117**: 588-589.

Whitkus, R., Doebley, J. and Wendel, J. F. 1994. Nuclear DNA markers in systematics and evolution. In: Phillips, R. L. and Vasil, I. K. (ed.), *DNA-*

based Markers in Plants (Advances in Cellular and Molecular Biology of Plants. Kluwer Academic Publishers, Dordrecht, Netherland. pp. 116-141.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **18**: 6531-6535.

Wilson, O. E. 2004. The encyclopedia of life. *Trends Ecol. Evol.* **18**(2):77-80.

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., Pang, X., Xu, H., Zhu, Y., Xiao, P. and Chen, S. 2010. Use of *ITS2* region as the universal DNA barcode for plants and animals. PLoS ONE **5**(10): 13102

Zhang, L. B., Simmons, M. P., Kocyan, A., Renner, S. S. 2006. Phylogeny of the Cucurbitales based on DNA sequences of nine loci from three genomes: implications for morphological and sexual system evolution. *Mol Phylogenet. Evol.* **39**(2): 305-322.

Annexure

ANNEXURE I

List of laboratory equipments used for the study

Refrigerated centrifuge: Kubota, Japan

Horizontal electrophoresis System: BioRad, USA

Thermal cycle : Veriti Thermal Cycler (Applied Biosystem, USA)

Gel documentation system : BioRad, USA

Nanodrop® ND-1000 spectrophotometer : Nanodrop®Technologies Inc. USA

ANNEXURE II

Reagents required for DNA isolation

Reagents:

1. 2x CTAB extraction buffer (100 ml)

CTAB	:	2g
(Cetyl trimethyl ammonium bromide)		
Tris base	:	1.21 g
EDTA	:	0.745 g
NaCl	:	8.18 g
PVP	:	1.0 g

Adjusted the pH to 8 and made up final volume up to 100 ml.

2. Chloroform- Isoamyl alcohol (24:1 v/v)

To chloroform (24 parts), isoamyl alcohol (1 part) was added and mixed properly.

3. Chilled isopropanol

Isopropanol was stored in refrigerator at 0°C and was used for the study.

4. Ethanol (70 %)

To the 70 parts of absolute ethanol (100 %), 30 parts of sterile distilled water was added to make 70 per cent ethanol.

ANNEXURE III

Composition of buffers and dyes used for gel electrophoresis

1. TAE Buffer 50X

Tris base	:	242 g
Glacial acetic acid	:	57.1 ml
0.5 M EDTA (pH 8.0)	:	100 ml

2. Loading Dye (6X)

0.25% Bromophenol blue

0.25% Xylene cyanol

30% Glycerol in water

3. Ethidium bromide

The dye was prepared as a stock solution of Ethidium bromide (stock 10 mg/ml; working concentration 0.5 µg/ml)(SRL) and was stored at room temperature in a dark bottle.

4. Agarose - 0.8 per cent (Promega)

- 1.5 per cent (for PCR samples)

5. Electrophoresis unit- BioRad power PAC 1000, gel casting tray, comb

6. UV translluminator- (Herolab ^{R})

7. Gel documentation and analysis system- BioRad Gel DOC-It^{TM} imaging system

DNA barcoding in *Momordica* spp.

By

GIRME AOUDUMBAR RAMESH

(2012-11-104)

ABSTRACT OF THE THESIS

Submitted in partial fulfillment of the requirement

for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR – 680656

KERALA, INDIA

2014

ABSTRACT

The genus *Momordica* comprises of 59 species, among which 7 are of Indian origin. Though the vegetables belonging to this genus are nutritionally rich with medicinal properties, their taxonomy remains confusive. The botanical names are often used incorrectly and interchangeably. Different taxonomic classification approaches have resulted in controversies about the number of species that exist and the phylogenetic relationships among these species. This situation necessitates an accurate, sensitive and simple alternate for the traditional taxonomy.

DNA barcoding is a novel system designed to provide rapid, accurate, and automatable species identification using short, standardized genomic regions as internal species tags. DNA barcoding is based on the variation on the sequences identified genomic regions, which can distinguish individuals of a species because genetic variation between species exceeds that within species. Species identification through barcoding is usually achieved by the retrieval of a short DNA sequence i.e. the 'barcode' from a standard part of the genome (i.e. a specific gene region either from chloroplast, mitochondria or nuclear genome). The barcode sequence from each unknown specimen is then compared with a library of reference barcode sequences derived from individuals of known identity.

The study entitled "DNA barcoding in *Momordica* spp." was done at the Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara. The objective of this study was to develop the barcodes for the 7 *Momordica* spp. which were found in India. For this study, 25 accessions of seven *Momordica* species *M. charantia*, *M. balsamina*, *M. dioica*, *M. sahyadrica*, *M. cochinchinensis*, *M. subangulata* subsp. *renigera* and *M. cymbalaria* and two accession of *Luffa* has been taken.

Total genomic DNA isolated using CTAB method (Rogers and Bendich, 1994) was subjected to PCR assay using various combinations of universal barcode primers for three loci *matK*, *ITS2* and *trnH-psbA*. As *ITS2* and *trnH-psbA*

gave the bands of 600 and 300bp, respectively, and which are not sufficient to develop the complete barcodes, these loci were not used in the study. Thus the bands generated using the *matK* primers from all the 25 *Momordica* and 2 *Luffa* were used for sequencing.

Phylogenetic analysis using ClustalW has discriminated the various species under *Momordica*, except *dioica* and *sahyadrica*. A wide level of molecular diversity detected with both the method which shows the high level of genetic variation among the species of 25 accessions of *Momordica* and *Luffa*.

Barcoding gap, a position in the sequence at which unique nucleotide is present in all the members of a particular species, was assessed for all the *Momordica* species and the *Luffa*, and these gaps were used to generate the barcodes for that species. From this study, barcodes were successfully generated for *M. charantia*, *M. subangulata*, *M. cochinchinensis*, *M. cymbalaria*, *M. balsamina* and *Luffa acutangula*.

The BLAST analysis had shown that *matK* is 95 per cent efficient for species discrimination in *Momordica*. From the base sequence of *matK* generated from this study, barcoding primers were designed for *Momordica* and were successful in all the 25 accessions. The *matK* barcodes developed in this study could be successfully used to solve the taxonomic confusion in *Momordica*.