DNA BARCODING IN GENERA *BENINCASA* AND *PRAECITRULLUS*

By

PRIYA SONKAMBLE

(2015-11-104)

THESIS

Submitted in partial fulfilment 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

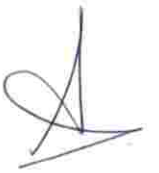
VELLANIKKARA, THRISSUR – 680 656

KERALA, INDIA

2019

DECLARATION

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



Priya Sonkamble

(2015-11-104)

Vellanikkara

Date: 09/05/2019

CERTIFICATE

Certified that the thesis entitled **'DNA barcoding in genera *Benincasa* and *Praecitrullus*'** is a record of research work done independently by **Ms. Priya Sonkamble** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to him.

Vellanikkara

Date:

Dr. C. Narayanankutty

(Major advisor)

Associate Director of Research (Seeds)

Directorate of Research

Kerala Agricultural University, Vellanikkara

CERTIFICATE

We, the undersigned members of the Advisory Committee of Ms. Priya Ashok Sonkamble, 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 genera *Benincasa* and *Praecitrullus*" may be submitted by Ms Priya Ashok Sonkamble. in partial fulfillment of the requirement for the degree.

9/5/19

Dr. C. Narayanankutty
(Major Advisor, Advisory Committee)
Associate Director of Research (Seeds)
Directorate of Research
Kerala Agricultural University, Vellanikkara

9/5/19

Dr. Shylaja M. R
(Member, Advisory Committee)
Professor and Head
CPBMB
College of Horticulture,
Vellanikkara

9/5/19

Dr. Deepu Mathew
(Member, Advisory Committee)
Assistant Professor
CPBMB
College of Horticulture,
Vellanikkara

Dr. Joseph John K.
Principal Scientist
National Bureau of Plant
Genetic Resources
Regional Station-Thrissur
KAU Post, Vellanikkara

9/5/19

EXTERNAL EXAMINER

Dr. SHEEJA T. E.
PRINCIPAL SCIENTIST & MEMBER SEC., ITM-BPD UNIT
DIVISION OF CROP IMPROVEMENT & BIOTECHNOLOGY
ICAR-INDIAN INSTITUTE OF SPICES RESEARCH
MARIKUNNU.P.O., KOZHIKODE-673 012, KERALA

ACKNOWLEDGEMENT

First and foremost I bow my head before the ***Almighty God*** *for enlightening and making me confident and optimistic throughout my life and enabled me to successfully complete the thesis work in time.*

It is with immense pleasure I avail this opportunity to express my deep sense of whole hearted gratitude and indebtedness to my major advisor ***Dr. C. Narayanankutty*** *for his expert advice, inspiring guidance, valuable suggestions, constructive criticisms, constant encouragement, affectionate advice and above all, the extreme patience, understanding and wholehearted co-operation rendered throughout the course of my study. I really consider it my greatest fortune in having his guidance for my research work.*

I consider it as my privilege to express my deep-felt gratitude to ***Dr. Shylaja M. R****, Professor and Head, CPBMB, CoH, Vellanikkara and member of my advisory committee for her constant support, valuable suggestions and critical scrutiny of the manuscript.*

I express my gratitude to ***Dr. Deepu Mathew****, Assistant Professor, CPBMB, CoH, Vellanikkara and member of my advisory committee for his immense help and assistance provided for constituting the manuscript.*

I sincerely thank ***Dr. Joseph John K****., Principal Scientist, NBPGR, CoH, Vellanikkara my advisory committee for his expert advice, constant inspiration, precious suggestions, generous support and constructive criticisms during my entire study which helped in successful completion of this work.*

I express my sincere thanks to ***Dr.Valsala P. A.****, retired Professor of CPBMB for her for her support, enthusiasm and relevant suggestions.*

I am deeply obliged to the teaching assistants, research assistants and non-teaching staffs at CoH, Vellanikkara as well as CPBMB, Vellanikkara especially ***Dr. Beenu*** *and* ***Dr. Mariet*** *for their invaluable help, guidance and critical assessment throughout the period of work. I thank them for all the help and cooperation they has extended to me.*

I also appreciate and acknowledge the facilities provided at ***Bioinformatics Centre****, the teaching staff* ***Ms. Priyanka*** *for their valuable help and guidance for the successful completion of my research work.*

I am deeply obliged to the teaching assistants, research assistants and non-teaching staffs at COH, Vellanikkara as well as CPBMB, Vellanikkara especially ***Dr. Beenu*** *and* ***Dr. Mariet*** *for their invaluable help, guidance and critical assessment throughout the period of work. I thank them for all the help and cooperation they has extended to me.*

I am in dearth of words to express my love towards ***my beloved*** *Parentsand my lovely sister* ***Pallavi Tayade*** *and brother* ***Vikrant Sonkamble*** *for their boundless affection, moral support, eternal love, deep concern, prayers and personal sacrifices which sustains peace in my life.*

I consider it as my privilege to express my deep-felt gratitude to ***Dr. Shylaja M. R,*** *Professor and Head, CPBMB, CoH, Vellanikkara and member of my advisory committee for her constant support, valuable suggestions and critical scrutiny of the manuscript.*

I duly acknowledge the encouragement, moral support, precious suggestions and timely persuasions by my dear seniors, not only in my research work but also throughout my PG programme. I express my sincere thanks to my classmates ***Marjan, Arunima, Geethu, Giridhari, Deepali, Debashish, Basil, Ashwin, Pramod, Kishore*** *and* ***Swapnil*** *for their affection and kind help offered during my thesis work. I have infinite pleasure to express whole hearted thanks to my loving juniors for their love, innumerable help and support.*

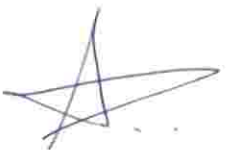
I express my deep sense of gratitude to Kerala Agricultural University for financial and technical support for persuasion of my study and research work. It would be impossible to list out all those who have helped me in one way or another in the successful completion of this work. I once again express my heartful thanks to all those who helped me in completing this venture.

I owe special thanks to Librarian, College of Horticulture, ***Dr. A.T. Francis*** *and all other staff members of Library, who guided me in several ways, which immensely helped for collection of literature for writing my thesis. Iam thankful for the service* ***Mr. Aravind*** *had done all along the academic period.*

*I thank my dear friends**Akhil,Deepa, Aparna, Nusrath, Nadhika, Selpriya, Rosemary, Preethy, Zita, Nagendra, Sunil, Chijina, Nishida**and junior friends at COH, Vellanikkrafor the unconditional support, help, timely valuable suggestions and encouragement which gave me enough mental strength and perseverance to get through all odds and tedious circumstances and immense thanks to all M.Sc. classmates for their moral support and encouragement.*

*It is with immense pleasure I avail this opportunity to express my deep sense of whole hearted gratitude and indebtedness to my major advisor **Dr. C. Narayanankutty,** for his expert advice, inspiring guidance, valuable suggestions, constructive criticisms, constant encouragement, affectionate advice and above all, the extreme patience, understanding and wholehearted co-operation rendered throughout the course of my study. I really consider it my greatest fortune in having his guidance for my research work.*

My word of apology to those I have not mentioned in person for the successful completion of this endeavor.



Priya Ashok Sonkamble

ABBREVIATIONS

%	Percentage
@	At the rate
<	Less than
=	Equal to
>	Greater than
µg	Microgram
µl	Microlitre
BLAST	Basic Local Alignment Search Tool
bp	Base pair
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
kg	Kilogram
L	Litre
M	Molar
mg	Milligram
ml	Millilitre
mM	Millimolar
NCBI	National Centre for Biotechnology Information
ng	Nanogram
°C	Degree Celsius
OD	Optical Density
PCR	Polymerase Chain Reaction
pH	Hydrogen ion Concentration

mat	maturase K
ITS2	Second Internal Transcribed Spacer
COI	Mitochondrial Cytochrome Oxidase subunit I
RFLP	Restriction Fragment Length Polymorphism
RAPD	Random Amplified Polymorphic DNA
DAF	DNA Amplification Fingerprinting
SSR	Simple Sequence Repeats
ISSR	Inter Simple Sequence Repeats
STS	Sequence Tagged Sites
SCAR	Sequence Characterized Amplified Regions
EST	Expressed Sequence Tags
DAMD	Direct Amplification of Minisatellite DNA
RNA	Ribonucleic Acid
rpm	Revolutions per minute
TAE	Tris Acetate EDTA
TE	Tris EDTA
V	Volts
BOLD	Barcode of Life Database System

TABLE OF CONTENTS

LIST OF TABLES

LIST OF PLATE

LIST OF FIGURES

LIST OF ANNEXURES

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

Introduction

1. INTRODUCTION

The genus *Benincasa* is a monotypic genus that belongs to family Cucurbitaceae with a single cultivated species, *B. hispida* (Thunb.) Cogn. It is named after Italian count, Guisepe *Benincasa* and is commonly known as ashgourd. The ashgourd is also called as white gourd, tallow gourd, pumpkin gourd (Nayar and More, 1998). In India, it is known as Petha, Safed Kolu and Kohla in vernacular language. Indo-China is considered as the main centre of diversity of ashgourd (Rubatazky and Yamaguchi, 1999).This annual vine is reported to be native to the Asian tropics like Japan, Indonesia and Malayasia (including India) and is grown primarily in the old world tropics.Since the existence of wild population of ash gourd is unconfirmed, the centre of origin still remains uncertain. It is probably a native of Indo-Malayasia (Burkill, 1935). It is supposed to be indigenous to Asian subtropics. It is widely grown in countries like India, China, Japan and Indonesia. (Pandey *et al.*, 2015).

In India, *Benincasa* exhibits rich diversity. There is substantial variation in vegetative traits and fruits characters and it is difficult to distinguish ashgourd genotypes based on their external morphology alone. Ashgourd is still classified as a cucurbit of minor importance in the world although it is widely grown in all parts of India. In India, it is mostly grown in Kerala, Tamil Nadu, Andhra Pradesh, Karnataka, Rajasthan, Haryana, Bihar, Uttar Pradesh and West Bengal.

*Praecitrullus fistulous*is a herbaceous annual belonging to Cucurbitaceae. It is native to India and distributed widely among the tropics. It is commonly known as Tinda or Indian baby pumpkin, Apple gourd or Indian round gourd (Sankar *et al.*, 2011). It is a monoecious vegetable grown in irrigated and arid areas of Indian subcontinent, and is a rich source of minerals and vitamins. In India, 75% of tinda production is confined to the state of Punjab. *P. fistulosus* is used as antioxidant with antihelmintic activity. It is effective in increasing the urinary flow and excretion of toxins from kidney. Tinda contains antioxidants like carotenoids and saponins. Carotenoid acts as an anti-inflammatory agent effective for controlling blood pressure, heart diseases and stroke. Saponins possess anti-tumorous activity which lowers the risk of human cancers by preventing growth of cancer cells (Khan, 2016).

Benincasa exhibits substantial variation for morphological characters especially growth habit, size, shape, size and thickness of fruits (Sundararajan and Muthukrishnan 1982, Peter *et al.*, 1991, Mandal, *et al.*, 2002). In spite of having enormous morphological variability all members are accommodated at present into one single species. It is difficult to distinguish genotypes based on their external morphological characters alone as they are greatly influenced by environmental factors and development stage of plant. Although there is substantial variation, it is difficult to differentiate genotypes based on their external morphology alone (Pandey *et al.*, 2008).

Levi *et al.*, (2010) evaluated that taxonomic classification of *P. fistulosus* is incomplete; it has been considered as a close relative of watermelon and hence botanically named as *Citrullus lanatus* subsp. *fistulosus* (Stocks). In order to explore the genetic relatedness of *P. fistulosus* in comparison with other cucurbit species, two sets of DNA based (Expressed Sequence Tag-Simple Sequence Repeat, Expressed Sequence Tag-Polymerase Chain Reaction and Simple Sequence Repeats) markers were used to conduct phylogenetic analysis. The marker and pollen morphology study revealed genetic relatedness of *P. fistulosus* and *B. hispida* when compared to other cucurbit species.

At present, to detect phylogenetic relationship in plant, DNA barcodes such as *rbcL*, *matK*, *psabA-trnH*, *rpoC1*, *ITS2*, *atpF-atpH* spacer and *psbK-psbI* spacer have been popularly employed worldwide (Janzen, 2009).

DNA barcoding is an accepted methodology for species level identification of living beings, even if not a substitute for conventional taxonomy. It is based on sequencing a short-standardized genome of target specimen, later comparing to sequence library from known species. It provides higher flexibility for species identification in large taxonomic assemblages. It is based on obtaining sequence information which is used to link unknown haplotypes to known species. It is helpful to fill the conceptual gap between traditional taxonomy and molecular systematics as a mature field of biodiversity sciences (Hubert and Hanner, 2006).

The protocols for barcode of few horticultural crops such as capsicum, citrus and *Momordica* (Girme, 2014) are established. In correlation, the *ITS2* (Internal Transcribed

Spacer) nuclear ribosomal DNA and *matK* (chloroplast DNA) is recommended, based on previous work done (Girme, 2014).

In *Benincasa* and *Praecitrullus*, very limited work is done when consideredat the molecular level to confirm enormous variability in the cultivated Indian species. The physiological and DNA marker studies revealed genetic similarity among the species. Thus, the members have to be allocated to different species on the basis of their cross compatibility. The suggestion for reallocation of non-compatible accession within the species could be supported with the DNA barcoding definition.

The present study was undertaken with the objective to assess genetic diversity and interspecies relationship among *Benincasa* and *Praecitrullus* genotypes and to check the species status of *Benincasa* using barcodes.

Review of literature

REVIEW OF LITERATURE

The present study on "DNA barcoding in genera *Benincasa* and *Praecitrullus*" has been executed through the amplification of chloroplast *matK* gene and *ITS2* locus from different accessions of ashgourd and tinda. The amplified chloroplast *matK* gene and nuclear gene *ITS2*were sequenced and the sequence information was further used to identify the species relation. The relevant literature available on various aspects of this study were collected and are reviewed in this chapter, under different heads.

2.1 Ashgourd (*Benincasa hispida* Thunb. Cogn.)

2.1.1 Origin and history

Benincasa is a monotypic genus belonging to the family Cucurbitaecae with a single species, *Benincasa hispida* (Thunb.) Cogn. This annual vine is reported to be native to the Asian tropics like Japan, Indonesia and Malysia (including India) and is grown primarily in the old-world tropics. Ashgourd occupies a prominent position among the tropical cucurbitaceous vegetables.

According to Burkill (1935), ashgourd was "probably a native to Malysia" and occurred in the wild Java. According to Gopalkrishnan (1957), a variety of *Benincasa hispida*, not previously recorded, known as medicinal ashgourd or *vaidyakumbalam/neikumbalam* was found in cultivation in Malabar.

De Candolle (1882) found it on the sea shore of Java and later it spread northwards to Japan, Central America and West Indies. It is grown for its fruit, which is used in preparing confectionaries and in Ayurvedic medicinal formulations (Indira and Peter, 1987). Walters and Decker-Walters (1989) reported that this monoecious crop is a native of Southern Asia and has been widely cultivated since fifth century in tropical Asia, China and India.

It is also used to treat a variety of ailments in ayurvedic and naturopathy systems of medicine (Ramesh *et al.*, 1989).

Benincasa hispida is commonly known as ashgourd. In Sanskrit it is known as "KumbhaPhala" means "water pot fruit" (Decker-Walters, 1998). Whistler (1990), reported that the *Benincasa hispida* was present in Polynesia before the advent of

Europeans in 1767. The ashgourd seems to have been domesticated in India during pre-historic times. The crop is cultivated in both tropical and subtropical regions. It is having wide variability in morphological and quality parameters. Ashgourd is considered as a potential crop of America and Africa (Srivastava 1993). Walters, (1989) reported that the ashgourd cultivars which were introduced to Europe in early days were actually Indian ashgourd cultivars. Some of the reports from Ceylon and India, suggested that the fruit was dominant in Asian region (Zaini *et al.*, 2011).

Ashgourd is a good source of carbohydrate, vitamin A, vitamin C and minerals like iron and zinc (Sureja *et al.*, 2006).

The different names of ashgourd are Kundur in Malay language, BhuruKolu or SafedKolu in Gujarat, Petha in Hindi, Kohla in Marathi, Kusumanda in Sanskrit, donggua in Chinese, Fakkio in Thailand, Calabaza china or Calabaza blanca in Spanish, Kondol in Phillipines and Bleego in Indonesia. The other names of *B. hispida* are winter melon, ash gourd, ash pumpkin, white gourd, white pumpkin, gourd melon, tallow gourd, wax gourd and Chinese water melon or Chinese preserving melon. In several mythological stories, rituals and idioms ashgourd is described as "gourds" perhaps due to its life-giving properties (Marr *et al.*, 2007).

Ashgourd is considered as a prized vegetable among the cucurbits because of its high nutritional value, long storage life and good transport qualities, besides its medicinal property (Gangopadhay *et al.*, 1983).

2.1.2 Diversity

The family Cucurbitaceae is considered one of the most important and genetically diverse groups of plants in the plant kingdom. It consisted of about 118 genera and 825 species (Robinson and Decker-Walters, 1999). Out of this about 20 species belonging to 9 genera are in cultivation. *Benincasa* is named after Count *Benincasa*, an Italian patron of botany.

Wild populations of the small fruited type, classified as *Benincasa hispida* var *pruriens* (Parkinson) Whistler, are found in several islands in the South Pacific (Morton, 1971).

Two types of ash gourd, viz. purple green coloured and green coloured were described by Chauhan (1989).Four recognized cultivars of ash gourd have been identified by Walters and Decker-Walters (1989).The specific characteristic of ashgourd is the presence of dense pubescence on foliage and immature fruits an 'epithet' used, 'hispida' refers to hirsute pubescence on the foliage. The variation even can be observed in fruit colour from light-green to dark green and speckled green. Five-categories of seed size were observed: 1) super small seed (90-95 seed/g), 2) very small seed (60-65 seeds/g), small seed (35-40 seeds/g), medium seed (20-25 seeds/g) and large seed (10-12 seeds/g) (Meusel *et al.*, 1994).

Bates and Robinson (1995), identified four groups of ashgourd cultivars namely, i) unridged winter melon, ii) ridged winter melon iii) fuzzy gourd and iv) wax gourd. Later, sixteen cultivars of *B. hispida* were reported by Marr *et al.*, (2007) based on the shape, length and width and skin color of the fruit. Two cultivars of ash gourd, i.e. round shaped and elongated shaped are grown in Malaysia. The round shaped fruit is the cultivated type and is used for commercial purposes.

Ashgourd belongs to the tribe Benincaseae (Jeffrey, 1990). Ashgourd is a monotypic genus, no wild or any other related species are reported (Gangopadhay, 1983).In northeast region of India, landraces are genetically divergent from those originating from other regions of the country (Pandey *et al.*, 2015). The cucurbit is grown upto altitude of 1500 m above sea level (Pandey *et al.*, 2015). Ashgourd was earlier classified under genus Cucurbita but on the basis of the observation of the 3 stamens which are separate, whereas in Cucurbita it is united, therefore it was placed under Benincasa (Pandey *et al.*, 2015). In India, a few major categories have been recognized

1) unridged winter melon group

2) ridged winter melon group

3) fussy gourd group

4) waxgourd group and

5) waxless gourd group (Pandey *et al.*, 2015).

Indian consumers, mostly prefer round or oblong fruits (6-8 kg) of light green to dark-green colour, whereas long cylindrical fruits (10-12 kg) with dark-green colour and white specks are mostly preferred by consumers in Vietnam. The region especially Nagaland, Garo Hills of northern side of country has rich diversity of *Benincasa*. It is observed that there are various varieties and hybrids of ashgourd having different traits, which exist at a particular time in southern parts of India (Manikandan *et al.*, 2017).

2.2 Tinda (*Praecitrullus fistulosus*)

2.2.1 Origin and History

Praecitrullus fistulosus, is commonly known as tinda. The origin of *Praecitrullus* is probably northwestern India and adjoining Pakistan (Samadia, 2007). The are several names for *Praecitrullus* like Tendu in Punjab, Dhemse in Marathi, Kovaikkai in Tamil, Tinda Kaya in Telugu and 'Indian Round Gourd', 'Indian Baby Pumpkin' and 'Apple Gourd' in English language.

P. fistulosus is considered as a minor crop in Africa. It has been cultivated in Asia since ancient times and is still an underexploited crop in Western world. The 'Arka' and 'Dilpas' are two cultivars introduced from India to Kenya. It is extensively grown in northwestern India (Samadia, 2007). In India, tinda cultivars with dark green fruits and pale green fruits are cultivated.

Only very little reports are available on genetic diversity for quantitative traits in round gourd (Dahiya *et al.*, 2001). In the last decade, it has been reported that among the various cucurbit species, the waxgourd or ashgourd (*Benincasa hispida*) is the closest to the round gourd (*P. fistulosus*) on the basis of DNA markers and pollen morphology (Koycan *et al.*, 2007) (Schaefer, Heibl and Renner 2009), (Levi *et al.*, 2010).

The chromosome number is 2n=24. Tinda is mainly cultivated as a market vegetable. It is a monoecious, creeping climber, with stout vine and annual with round fruits which is pale or dark green in colour. The fruit is called a pepo, a type of berry (Tyagi *et al.*, 2012).

It is grown mostly in the countries of Africa as an export commodity for countries in Asia and United Kingdom. Tinda is grown in dry season (February to end of April) or in rainy season (mid June to end of July) (Tyagi *et al.*, 2012). Garg (2017) classified fifteen accessions of *P. fistulosus* for quality characters like leaf blade, fruit shape, fruit skin colour, flesh colour, fruit surface hairs and fruit skin texture.

2.3 Cytological studies of *B. hispida* and *P. fistulosus*

The chromosome number of *P. fistulosus* is (2n=24) (Tyagi *et al.*, 2012). *Benincasa hispida* synonym *Benincasa cerifera Savi.* is also having the diploid number is (2n=24). The meiotic preparations from root tip cells have 24 chromosomes. The karyotype analysis of chromosomes revealed that length of chromosomes was 2.00-3.00 µm. There are eight pairs of median and four pairs of sub terminal chromosomes in complement. The twelve pairs of chromosomes can be arranged into six groups according to their comparative length and arm ratio. The total length of chromosomes of the complement and mean length of chromosomes are estimated to be 57.36 µm and 2.39 µm respectively. All stages of meiosis were normal as per the observation. Pollen fertility estimated to be 97.8% (Pandey *et al.*, 2015).

2.4 Genetic assessment among *Praecitrullus fistulosus* and *B*enincasa *hispida*

In fact, both the species *P. fistulosus* and *B. hispida* have same diploid chromosome number (2n=24). They are both considered as underexploited in the western hemisphere, although they are grown in some areas in the US (Morton 1971). *Benincasa hispida* is having tremendous economic importance and still no systematic efforts have been made to assess the genetic diversity of its germplasm (Gangopadhay *et al.*, 2008). *P. fistulosus* has small flowers which are similar to size of watermelon and *B. hispida* has large flowers similar to the size of *Laegeneria siceraria. P. fistulosus* and *B. hispida* both the species evolved in Asia and cultivated there too, it is considered as less exploited crop in western world (Levi *et al.*, 2010). *B. hispida* has become an industrial crop, relatively less attention was made towards its morpho-molecular characterization of its strains (Resmi and Sreelathakumary, 2011).

The morphological characters of *P. fistulosus* like flower, seed and leaf are similar to watermelon (*Citrullus* spp.), therefore previously it was considered as a close relative of the watermelon and for this reason it was called as *Citrullus vulgaris* var. *fistulosus* (Pangalo 1938; Sujata and Seshadri, 1989). The taxonomic classification of *P. fistulosus* is incomplete and its relation to *Citrullus* spp. and *Cucumis* spp., is still unclear (Levi *et al.*, 2010). As it is considered as a distant relative of watermelon, various cross pollination attempts were carried out to cross *Fistulous* with *Citrullus lanatus* var. lanatus (watermelon) and with other *Citrullus 99*spp, which was not successful. Thus, to detect the genetic relatedness among *P. fistulosus* and *Citrullus lanatus* var. *lanatus*(watermelon) phylogenetic analysis was carried out including species *C. lanatus* var. *citroides* and wild *Citrullus colocynthis* based on DNA based marker RAPD (Random Amplified Polymorphic DNA) and ISSR (Inter Simple Sequence Repeats) markers. It was concluded that there is no genetic similarity between *P. fistulosus* and *Cucumis* or *Citrullus* groups. *P. fistulosus* appeared to be distant from both the species (Levi *et al.*, 2005).

The chromosome numbers of these two species are different, *P. fistulosus* (2n=24), *Citrullus lanatus* (2n=22). According to Shimutsoma (1963), *P. fistulosus* should be classified with *Cucumis melo* as these two species shows same monoploid chromosome number (2n=24) (Beevy and Kuriachan, 1996). Trivedi and Roy (1969), suggests separate species status should be given to *P. fistulosus* in genus *Citrullus*, on the basis of chromosome number.

As, it is confirmed that *P. fistulosus* is not crossable with either watermelon or either muskmelon, DNA marker analysis also confirmed that *P. fistulosus* is totally distinct from both the species. The isozymes GOT (Glutamate Oxaloacetate Transaminase) and Peroxidase analysis also provide additional evidence for distinctness among the three species. When *P. fistulosus* was compared with *Cucumis melo* no common isozyme was found between them. Thus, it is an argumental dispute to classify *P. fistulosus* along with *Cucumis melo* (Shimotsuma, 1963).

Chloroplast genome has been used widely as a tool for phylogenetic analysis in plants. It has conserved gene order, its widespread availability of primers, a general lack

of recombination and heteroplasmy. Its uniparental mode of inheritance in case of gymnosperms, makes it possible to elucidate relative contributions of seed and pollen flow to genetic structure by comparing nuclear and chloroplasts markers in natural populations (Olmstead and Palmer, 1994) and (Ennos *et al.*, 1999).

Dane and Lang (2004) conducted studies in watermelon (*Citrullus lanatus*) utilizing cpDNA to investigate the phylogenetic relationship among *Citrullus* haplotypes in order to estimate divergence and deduce colonization. In plants, the phylogenetic relationship study is dominated by investigation of cp (chloroplast) genome by many researchers (Schaal *et al.*, 1998). As, cpDNA is evolutionary and conserved in terms of genome size, structure of gene content and linear order of genes among plant lineages. However, cucurbit chloroplasts genome contains multiple variable sites. So, there is a wide range of possibilities for resolving relationship from the level of species, genus up to family, since different portions of chloroplast are evolved at different rates. In *Citrullus* most of the cpDNA variation is found as a result of large indels and nucleotide substitutions. The investigation is carried out among *Citrullus* accessions and *Praecitrullus* accessions. *Praecitrullus fistulosus* (Stocks) Pang., chosen as an out group, because this species once considered as *Citrullus* species, but it has given its own taxonomic category (Sujata and Seshadri, 1989). There are three indels and four restriction enzyme site differences at four of the cpDNA regions (*ndhA, trnS-trnsfM*, *ndhf* and *trnC-trnD*) which separates *C. lanatus* and *P. fistulosus*. The fragment/ restriction site differences are detected at almost all of cpDNA region in *Citrullus* and *P. fistulosus*, indicating a wide divergence in both of these genera.

In plants, the phylogenetic relationship study is dominated by investigation of cp (chloroplast) genome by many researcher's (Schaal *et al.*, 1998).

Koycan *et al.*, (2007) conducted studies in phylogeny of Cucurbitaceae for 171 species haplotypes 123 genera based on chloroplast DNA sequences from two genes, two spacers and one intron. The study revealed that the *Cucumis* spp., *Laegenaria siceraria*, the *Citrullus* spp., *P. fistulosus* and *Benincasa hispida* belong to single clade the *Benincasea*. Schaefer *et al.*, (2009) utilized chloroplast DNA based on standardized

protocols and primers described by Koycan *et al.*, (2007) and found that *Benincasa hispida* and *P. fistulosus* could be classified under a single clade.

The taxonomic classification of *Praecitrullus* is still uncertain. To work out the taxonomic classification with respect to other cucurbits, additional information of other cucurbits is also needed. Therefore, assessment of phylogenetic relationships of *Praecitrullus* with *Citrullus* spp., *Cucumis* spp., *Cucurbits* spp., *Laegeneria siceraria* and *Benincasa. hispida*(ashgourd) was done with two sets of DNA markers (Levi *et al.*, 2010). They are EST-PCR, EST-SSR and SRAP primers, which were derived from watermelon genome sequences. The marker analysis showed that *P. fistulosus* exhibit close genetic similarity with *B. hispida.* The pollen morphology study of above-mentioned cucurbits was carried out and it was found *B. hispida* and *P. fistulosus* have same pollen grains feature with spherical or semispherical structure compared to other curcurbits. This feature was in agreement with the phylogenetic relationship based on two sets of DNA markers (Levi *et al.*, 2009; Levi *et al.*, 2010).

The *P. fistulosus* was found genetically similar to *B. hispida* in findings given by Koycan *et al.*, (2007) and Schaefer *et al.*, (2009) for chloroplast DNA and Levi *et al.*, (2010) for nuclear gene sequence.

Echt *et al.*, (1998), examined the genetic diversity in *Pinus resinosa* using chloroplast genome and successfully detected the genetic divergence detecting the common haplotypes. The genetic divergence study in this species is limited RAPD and even no diversity could be detected using allozyme markers. Thus, the chloroplast marker proved to be helpful to reveal genetic diversity in *Pinus halepnsis* as well in case of *Torrey pine* (*Piruster reyana*), where detection is not possible using RFLP techniques, choloroplast genome detection is applied to reveal genetic divergence.

2.7 Morphological markers

Morphological markers are those genetic variations that can be observed by naked eye. The important morphological traits are plant height, sensitivity to environment, disease response, shape or colour of flowers, fruits or seeds or protein too, etc. Although they are generally easily scored, quick and simple without the use of laboratory

equipments, such markers are not always beneficial in use because of some reasons: genotypes are ascertained generally as a whole plant and the mature plant is used frequently. Such markers cause major differences in breeding programme. The dominant and recessive interaction frequently prevents differentiating genotypes associated with morphological traits. This markers mask effect of linked minor gene, making it completely impossible to identify desirable linkages for selection. The number of such markers are further limited in numbers and also specifically influenced by environment at certain stages (Bhat *et al.*, 2010). Sax (1923) proposed idea identification and selection of minor genes of interest by linkage with major genes, which could be scored more easily. This idea has surfaced many times and extended by many coworkers. Unfortunately, there are constraints in the use of morphological markers: 1) they cause such large effects on phenotype that are undesirable in breeding programme 2) they mask the effects of linked minor genes, making it nearly impossible to identify desirable linkages for selection and they are highly influenced by environment. The morphological/phenotype is the result of genetic constitution and its interaction with environment. Due to varying levels of GxE interaction, it is not appropriate to compare the morphological data of varieties that have been collected across different years and/ or locations.

2.8 Biochemical markers

Biochemical markers are proteins which code the gene expression. Isozymes have been used successfully as biochemical markers in plant breeding and geneticts as nearly as genetic markers. They are the products of various alleles of one or more gene. Isozymes are generally codominant. It should be noted that in most cases the polymorphism of isozyme is sometimes poor within a cultivated species. Isozymes techniques were historically, the first application of markers. Isozyme techniques as biochemical markers have some advantages over DNA markers such as 1) they are less costly, 2) interpretation of results is easier due to less noise and 3) technique is simpler and still being used effectively (Karaca, 2013).

Sujata and Seshadri (1989), utilized biochemical markers glutamate oxaloacetate transaminase (GOT) and peroxidase enzyme (PRX) for detecting genetic relatedness among *P. fistulosus*, *Citrullus lanatus* and *Cucumis melo* as they are not crossable.

Isozyme analysis could provide additional evidence for comparison of this species. In zymogram analysis, three isozymes found in *P. fistulosus*, which were different from *C. lanatus* and Cucumis *melo*. There was no similarity for *P. fistulosus* with *Cucumis lanatus* or *Cucumis melo* for Peroxidase (PRX) and glutamate oxaloacetate transaminase (GOT) enzyme. The zymogram analysis indicates dissimilarity between *P. fistulosus*, *Cucumis lanatus* and *Cucumis melo* (Levi *et al.*, 2010).

2.9 Molecular markers

A molecular marker is a DNA sequence that is readily detected and whose inheritance can easily be monitored. The use of molecular markers is based on naturally occurring DNA polymorphism, which forms the basis for designing strategies to exploit for applied purposes. A marker must be polymorphic; that it must exist in different forms so that chromosomes carrying the mutant gene can be distinguished from the chromosome with the normal gene by a marker it also carries. Genetic polymorphism is defined as the simultaneous occurrence of a trait in the same population of two or more discontinuous variants or genotypes. The first such DNA marker to be utilized was the restriction fragment length polymorphism (RFLP).

2.6 Molecular characterization

Molecular markers are genotypic markers. Unlike morphological markers, molecular markers characterize diversity at the molecular level and therefore are environmentally independent. The use of these markers provide a potential effective selection technique for crop improvement and has advantage over selection based on phenotype alone.

Molecular markers have been widely used in genetic analysis and literally assessment in a number of plant species (Waugh and Powell, 1992; Bretting and Widerlencher, 1995; Staub *et al.*, 2002).

Molecular markers that reveal polymorphism at the DNA level are known as DNA markers. They provide an opportunity to characterize genotypes and to measure relationships more precisely than other markers (Soller and Beckman, 1983). Various

types of molecular markers are utilized to evaluate DNA polymorphism and among them, the most important is polymerase chain reaction based.

2.7 PCR-based molecular marker

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 plants characterized by high genetic variations between cultivars.

2.8 Polymerase Chain Reaction (PCR)

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 Arbitarily Primed Polymerase Chain Reaction (AP-PCR). DNA Amplification Fingerprinting (DAF) was also reported as another technique of PCR used in various experiments (Caeteno-Anolles and Brant *et al.*, 1991). The PCR reaction requires deoxynucleotides, DNA polymerase, primer, template and buffer containing magnesium (Bej *et al.*, 1991). Typical PCR amplification utilizes oligonucleotide primers which hybridize 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 is applied to detect polymorphisms in various plants, animal, bacterial species and fungi.

The introduction of the PCR technique has revolutionized standard molecular techniques and allotted for the proliferation of new tools for detecting DNA

polymorphism (Hu and Quiros, 1991). 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 simple analysis and the ability to amplify extremely, small quantities of DNA (Welsch *et al.*, 1991).

2.9 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 advantages and disadvantages to assess the genetic variations among natural populations. Currently, the most commonly used genetic markers are molecular markers. Due to the rapid developments in the field of molecular genetics, a variety of different techniques have emerged to analyze genetic variation during the last few decades (Whitkus *et al.*, 1994; Karp *et al.*, 1996, 1997a, b; 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 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 study diversity and to assist in crop improvement. (Singh and Sarkar, 2014)

2.10 Molecular markers in taxonomy

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

2.11 Molecular marker analysis

The marker system like RAPD, AFLP, SSR and ISSR have been tested in cucurbits. Molecular markers have been used for species identification and their relationship with other species.

RAPD is now being applied to a wide range of research activities including genome fingerprinting (Welsh and McClelland, 1990), classification of genome specific markers (Williams *et al.*, 1990; Erlich *et al.*, 1991), population biology studies

Verma and Behera (2007), Pandey *et al.*, (2008) and Resmi and Sreelathakumary, (2011) utilized RAPD markers to analyzed the genetic diversity in ashgourd. Rashid *et al.*, (2016) analyzed genetic diversity in Snake gourd germplasm in Bangladesh using RAPD markers. The study is helpful in carrying out crop improvement by stressing the variation in genotypes (Resmi and Sreelathakumary, 2011).

The sex type of plant species can be detected using a DNA-based test, but such tests can only be possible if a sex-associated marker is available. RAPD markers being linked to a gene or genomic region of interest is mainly dependent on chromosome number, genomic size, type of gene or genomic region, which can used to determine the sex and even which type of population was used for marker analysis(Trivedi *et al.*, 2015).

Kumar *et al.*, (2012) identified two and Singh *et al.*, (2002) identified three RAPD markers associated with sex expression in pointed gourds. However, the above markers were not reliable to detect sex expression in pointed gourds from different geographical regions. Nanda *et al.*, (2013), identified a male specific STS marker (TdSTSM), which is an ISSR based marker. It is highly reliable and reproducible which shows effective distinctness of male and female plant of pointed gourd.

11 species of cucurbitaceae family were classified based on the analysis of three different kinds of markers. Isozyme, RAPD and ISSR marker system were applied for

studying the genetic diversity between the cucurbits species. The 11 species are *Lagenaria siceraria*, *Cucurbita maxima*, *Cucumis sativus*, *Benincasa hispida*, *Luffa acutangula*, *Luffa cylindrica*, *Trichosanthes cucumeria*, *Trichosanthes diocia*, *Momordica charantia*, *Momordica diocia* and *Coccinia cordifolia*. The resolving power of markers compared for the applicability in phylogenetic analysis using Mantel's test. The Isozyme markers fail to distinguish species of *Momordica* and *Luffa*, ISSR could not distinguish between Momordica species and RAPD fails to distinguish between Tricosanthes species (Sikdar *et al.*, 2010).

The AFLP system is a PCR based method. It is an effective method for genome analysis, detecting more than 50 independent bands in a single PCR reaction. This markers reveal polymorphism between related and unrelated genotypes like mutants, NIL, RIL and DH. It is highly reproducible and specific (Witkowicz, *et al.*, 2003).

The three markers viz., RAPD, AFLP and RFLP were used to measure genetic diversity in six genotypes of *Cucumis melo* L. The three of the marker system differentiate *Cucumis melo* L. genotypes into two different kinds and it suggests all three markers are highly informative, while AFLP showed higher efficiency (Mas *et al.*, 2000).

SSR markers are also termed as microsatellite markers consists of a variable number of tandem repeats. It is of two to five nucleotides that form highly informative and locus-specific genetic markers. They are abundant and evenly dispersed in eukaryotic genome. They can be analyzed efficiently through PCR reaction by applying the primer specific to their flanking region. Variation in PCR product length is a function of SSR unit numbers. They have been recommended as markers for constructing highly saturated genetic map. SSR marker were constructed and developed for characterization and diversity studies in *Cucumis* (Poleg *et al.*, 2001).

The DAMD markers were used to study the genetic relationship in 19 cucumber genotypes. The (DAMD) Direct Amplification of Minisatellite DNA markers are minisatellite core sequences which is used as primers, it is used to amplify the minisatellite rich DNA region. The minisatellite core sequences were obtained from human, rice and M13 phage, their utilization results in four groups of cucumber germplasm which demonstrate relationship among them. This marker shows rapid

resolution of intervarietal variations and rapid specific markers in cultivars of *Cucumis sativus*. This is the first case study of utilization of such markers in Cucurbits (Hu *et al.*, 2011).

2.13 DNA Barcoding in Plants

DNA barcoding is a technique, in which species identification achieved by retrieval of short DNA sequence from the "barcode" of a standard part of genome. The barcode sequence is then compared with library of reference barcode sequences which is obtained from individuals of known identity. Thus the specimen is identified if its sequence closely matches with one in the barcode library. According to the CBOL view, the definition fits with the identification of species level using a single standardized DNA fragment (Valentini *et al.*, 2009).

DNA barcoding is in fact based on the concept of distinguishing an individual based on a short stretch of standardized sequence which distinguish the individual of a species, due to the reason that genetic variation between species exceeds those within the species. The concept has recently become popular. The barcoding is characterized by using one or few DNA fragments to identify the different species. In molecular systematics, DNA sequencing method takes more than 20 years to reveal evolutionary relationship (Kress *et al.*, 2005).

The CBOL (Consortium for Barcode of Life) initiates a new database totally dedicated to DNA barcode. Barcode of Life Data Systems (BOLD) (http://www.barcodinglife.org). BOLD has been designed to retrieve not only DNA sequence but also the complete taxonomic information, place and date of collection, with the image of specimen from several individual species. BOLD is an interactive collaboration where deposited sequences can be replaced taxonomically and adjusted (Frezal and Leblois, 2008).

The DNA barcodes is similar to the black stripes of Universal Product Code which are used to recognize commercial products. The DNA barcoding can be mere useful when an assembled public library of sequence is linked to name specimens together with faster and cheaper sequencing. The primary goals of barcoding are species

identification of already known specimens and as well discovery of overlooked species for improving quality for future fulfillment for benefit of science and society (Kress and Erickson, 2007). Barcoding helps to preserve Earth's biodiversity by strengthening ongoing efforts and contribute to further actions of botanic gardens and ex-situ conservation collections. The goal of DNA barcoding is simple to find one or few regions of DNA that will differentiate among the majority of overall world species and sequence it to produce a reference library of life on earth. This tool proved to be highly significant for species identification and also led to help in discovery of new species (Kress, 2005).

The plant barcode search has focused on the chloroplasts genome and several candidate genes such *accD*, *psbK-psbI*, *rbcL*, *rpoB*, *atpF-atpH*, *nhdJ*, *rpoC1* and *trnH-psbA* and *matK*. Chase *et al.*, (2005); Kress *et al.*, (2005); Newmaster *et al.*, (2006); Yoo *et al.*, (2006). Unfortunately, very few from these loci are variable enough to distinguish plant species, when used alone. This idea lead to the use of several combinations of loci, such as *matK*+ *atpF-H* + *trnH-psbA*, *rpoC1*+ *rpoB* + *matK*, *rbcL* + *trnH-psbA*, *rpoC1* + *matK* + *trnH- psbA* (Yoo *et al.*, 2006), (Chase *et al.*,2007) and (Kress and Erickson, 2007) CBOL Plant Working Group, (2009). Among these, *matK* barcode is one of the promising candidate for a plant barcoding work (Chase *et al.*, 2005).

The *matK* gene codes for maturase protein approximately 1570 bp in length. The coding region for maturase K is generally located within an intron of the chloroplast trnK gene, while in ferns it encodes tRNA Lys (UUU) (Neuhaus and Link, 1987). The *matK* being a coding region has high evolutionary rate and thus it made it easy and usable for phylogenetic reconstruction at higher taxonomic level such as Genus or Species, Order or Family. *matK* proves to be one of the useful markers for DNA barcoding in identification of plant families although the extensive divergence for higher taxonomic categories led to questionable positions and relationship of some phylogenetic clades (Yu *et al.*, 2011).

matK acts as an essential splicing factor with high nucleotide substitution rate and low structural conservation gene. Thus the significance of this gene is to bring about the phylogenetic relationship within families and also within genera in higher plants. Therefore, genetic relationships were examined with the help of *matK* sequence variation among 37 genotypes of oil-bearing rose's cultivars. As far as genetic diversity is

concerned, the genus Rosa is a genus of complicated phylogenetic relationship due to interbreeding and long cultivation history. Thus the *matK* gene sequence study brought a great importance to fragrant variety identification, development and breeding (Wang *et al.*, 2012).

Cucurbitaceae has high genetic diversity with 120 genera and 825 species. The chloroplast *matK* gene is a rapidly evolving plastid coding region, due to its high substitution rates. Reddy (2009) retrieve *matK* sequences of Cucurbitaceae from GenBank to examine inter and intragenous phylogeny. This study confirmed that *matK* sequence can be utilized to resolve the relationships unambiguously and it provides a major indication of supra-generic groupings among the selected members of Cucurbitaceae.

The *ITS* region is a part of transcriptional unit of 18S-26S nrDNA. The function of *ITS1* and *ITS2* regions of nrDNA transcript appear to be least in maturation of rRNAs. The *ITS* region includes three components: the 5.8S subunit, evolutionary highly conserved region and most important the two spacers designated as *ITS-1* and *ITS-2* as per the above image given. The two spacers of the region *ITS-1* and *ITS-2* each of < 300 bp sequence can be readily amplified by PCR with usage of universal primer present, even from DNAs of herbarium specimens. (Baldwin *et al.*, 1995).The internal transcribed spacer can be *ITS2* has been potentially used a standard barcode. The nuclear ribosomal DNA represents most of the perfect region for DNA and barcoding applications in higher plants. (Chen *et al.*,2010). The general features of *ITS* which promotes phylogenetic analysis in angiosperms is, first the *ITS* is a highly repeated unit in plant nuclear genome along with other nrDNA multigene family. Second, this gene family undergoes rapid concerted evolution which is important in phylogeny reconstruction. Third, small size (> 700 bp) and highly conserved sequence flanking each of the spacers make this region easy to amplify. Despite the high copy number of both spacers, it allows direct sequencing of pooled PCR products due to the near uniformity of *ITS* paralogues which attribute to the rapid concerted evolution in many species. In some of the reported studies, the variation between ITS sequences are mostly due to point mutation (Baldwin *et al.*, 1995).

ITS1 and *ITS2* (Internal Transcribe Spacer) have been widely used for DNA barcoding analysis of various different biological groups, *ITS* has been reported very rarely in DNA barcode evaluation and or employed as DNA barcode for species identification in some plants. Moreover the *ITS/ITS2* proposal is confusing, because *ITS2* is a sub region of *ITS1* and cannot be treated as different as *ITS1*. The efficiency of *ITS1* is more compared to *ITS2* in terms of presence of DNA barcoding gaps, sequence length distribution, species discrimination efficiency, GC content distribution, primer universality across major eukaryotic group which includes fungi, plants and animals. The analysis is carried out using the sequence available from public databases. The analysis shows that *ITS1* has better and higher discriminating power overall and represents a better barcode and prove guideline for strategies like designing analyzing particular taxonomic group (Wang *et al.*, 2015).

Chen *et al.*, (2010) tested 7 potential barcodes (*ITS1*, *ITS2*, *psbA-trnH*, *matK*, *rbcL* and *ycf5*) in a group of medicinal plants and closely related samples belonging to different species, genera and family in 7 phyla (Gymnosperms, Angiosperms, Mosses, Ferns, Liverworts, Fungi and Algae). According to BLAST test the *ITS2* level were at 92.7% and 99.8% at species and genus level. It shows high interspecific divergence among the loci used. This finding suggests that as mitochondrial gene "*CO1*" is favorable region for use as DNA barcode for most of the animal and fungi species, *ITS2* region is suitable for authenticating taxa at different taxonomic levels. *ITS2* sequences are potential phylogenetic markers, therefore considered as a potential barcode. *ITS2* loci possess many advantages over plastid genome fragments and other genome nuclear regions including ITS. *ITS2* is considered as a gold standard barcode for identifying plants as well as fungi.

Girme(2014) utilized three different barcode primers namely *matK*, *ITS2* and *trnH-psbA* for evaluating genetic relationship between 25 genotypes of *Momordica* and 2 genotypes of *Luffa*. It was studied for establishing phenetic relationship between six currently cultivated Indian *Momordica spp.*, a taxon of controversial identity of *Momordica cymbalaria* and finding the barcode gaps for the *Momordica* genotypes. The primers *matK*, *trnH-psbA* and *ITS2* were used to amplify the total genomic DNA obtained

from both *Momordica* and *Luffa* genotypes. It was found that different combinations of *matK* primers used were successful to amplify the locus in all accessions. There is clear genetic variation observed in the sequences between *Momordica* spp. and *Luffa* spp. which was further confirmed with the evidence of varying chromosomal number and morphological characteristics.

Materials and methods

3. MATERIALS AND METHODS

The study on "DNA barcoding in genera *Benincasa* and *Praecitrullus*" was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Kerala Agricultural University during 2015-2017. The materials used and methodologies adopted are being described/detailed in this chapter.

3.1 Materials

3.1.1 Plant materials

A total of 26 accessions of ashgourd present in India were used in this study. Additionally, 5 accessions of tinda were also employed to confirm the species relation. All accessions of ashgourd and tinda were obtained from different parts of India. Out of 26 ashgourd collections, 6 are from North-East India, one from Maharashtra and remaning 19 are from various parts of Kerala. Out of 19 Kerala collections, the accessions Kottamuri, Vellarikundu, Areeparambu, Nedungadapalli, Mallapally, North Paravur, Perumbilisserry, Kannara and Thanniam are medicinal types.

Table 3.1 Accessions used in this study

S.L. No.	Accessions	Region
1	Kottamuri	Chalakuddy, Kerala
2	Vellarikundu	Kasargod, Kerala
3	West Eleri	Kasragod, Kerala
4	Areeparambu	Alapuzha, Kerala
5	Kanjirampoyil	Kasargod, Kerala
6	Nedungdapalli	Kottayam, Kerala
7	Mallapally	Pathanamthitta, Kerala
8	North Paravur	Ernakulam, Kerala
9	Perumbilisserry	Thrissur, Kerala
10	Cheramangalam	Cherthala, Kerala

11	Wayanad	Kerala
12	Kannara	Kerala
13	Nashik	Nashik, Maharashtra
14	Acc. No.5	Kerala(ARS, Mannuthy)
15	Acc. No.9	Kerala(ARS, Mannuthy)
16	Acc. No.10	Kerala(ARS, Mannuthy)
17	Acc. No.11	Kerala(ARS, Mannuthy)
18	IC 0596986	North-East(NBPGR)
19	IC 0596985	North-East(NBPGR)
20	IC 0596987	North-East(NBPGR)
21	IC 0596989	North-East(NBPGR)
22	IC 0596995	North-East(NBPGR)
23	IC 05969692	North-East(NBPGR)
24	Elavancheri	Kerala
25	Thanniam	Kerala
26	KAU local	Kerala
27	Ankur tinda	Maharashtra
28	Nirmal tinda	Maharashtra
29	Komal tinda	Maharashtra
30	Mahyco tinda	Maharashtra
31	Kailash tinda	Maharashtra

3.2 METHODS

3.2.1 Experimental methods

The experiment was laid out in Centre for Hi-tech Horticulture and Precision farming during June to October 2017. The 26 ashgourd and 5 tinda genotypes, were raised on beds. 5 plants of each genotype of ashgourd and tinda were raised maintaining a distance of 2m between

each plant in a row and a distance of 4m between two adjacent beds. Total 15 plants of KAU local was raised as it's a female cultivar for convenient harvesting of hybrids.

3.2.2 Morphological analysis

Morphological characters of ashgourd (*Benincasa hispida*) and tinda (*P. fistulosus*) were recorded as per the minimal descriptors developed by NBPGR, New Delhi. The characters observed from the initial stage of plant growth up to the harvesting stage of plant and separation of seeds. The characters are divided into two group's viz., a) qualitative and b) quantitative characters. The descriptor/ character studied or recorded are as follows:-

Table. 3.2 Plant descriptor and its scoring

S.I. No.	Plant descriptor/ character	Score
A.	**Qualitative characters**	
1	Early plant vigour	(0-poor, 1-good, 2-very good)
2	Plant growth	(1-short, 2-medium, 3-long)
3	Stem pubescence	(1-hispid, 2-downy, 3-tomentose, 4-puberulent, 5-pilose, 6-villous, 7-hirsute)
4	Stem shape	(1-rounded, 2-angular)
5	Tendril	(0-absent,1-present)
6	Tendril type	(1-coiled, 2-straight)
7	Tendril branching	(1-unbranched, 2-branched)
8	Leaf margin	(1-serrate, 2-double serrated, 3-incised)
9	Leaf shape	(1-cordate, 2-oblong, 3-ovate, 4-obvoate, 5-orbicular)
10	Leaf size	(1-small, 2-medium, 3-large)
11	Leaf pubescence	(0-no hairs, 1-sparse, 3-intermediate, 4-dense, 5-hispid, 6-puberulent)

	nature	
12	Leaf pubescence density	(0-no hairs, 1-sparse, 2-intermediate, 3-dense)
13	Sex type	(1-monoecious, gynomonoecious, 3-andromonoecious, 4-hermaphrodite, 5-androeciuos or dioecious)
14	Peduncle shape	(1-round, 2-smoothly angular, 3-sharply angular)
15	Peduncle attachment	(1-hard and flared, 2-hard not flared, 3-soft and flared, 4-soft not flared)
16	Fruit shape	(1-cylindrical,2-club shaped, 3-spindle shape)
17	Fruit skin colour	(1-light green, 2-intermediate, 3-dark green)
18	Fruit skin lustre	(1-matt, 2-intermediate, 3-glossy)
19	Fruit pubescence density	(1-no hairs, 2-sparse, 3-intermediate, 4-dense)
20	Fruit ridge shape	(1-superficial, 2-grooved, 3-intermediate, 4-deep grooved)
21	Stem-end fruit shape	(1-depressed, 2-flattened, 3-rounded, 4-pointed)
22	Blossom-end fruit shape	(1-depressed, 2-flattened, 3-rounded, 4-pointed)
23	Flesh texture	(1-smooth, 2-spongy, 3-fibrous-gelatinous, 4-fibrous-dry, 5-grainy)
24	Flesh colour	(1-white, 2-creamish, 3-yellow)
25	Seed lustre	(1-matt, 2-intermediate, 3-glossy)
26	Seediness	(1-low, 2-medium,3-high)
B.	**Quantitative characters**	
27	Petiole length (cm)	

28	Peduncle length (cm)
29	Node number at which first female flower appears
30	Days to 50 per cent flowering
31	Number of ridges per fruit
32	Fruit length (cm)
33	Fruit width (cm)
34	Fruit weight (g)
35	Number of primary branches (recorded at the end of flowering stage)
36	Days to first fruit harvest
37	Days to last fruit harvest
38	Number of fruits per plant

3.2.3 Pollination works

Ashgourd is predominantly a monoecious crop. The anthesis of ashgourd flowers takes place at 4:30 am to 7:30 am, with anther dehiscence at 3:00 to 5:00 am. The fruit set on hand pollination of flowers is 65 per cent to 75 per cent. It is reported that pollen fertility is maximum at most 95 per cent at anthesis. The stigma receptivity was observed 12 h before anthesis and 12 h after anthesis.

The pollens from 25 genotypes of ashgourd were dusted on the female flowers of ashgourd variety KAU local. The pollination was done between 6:30 am to 8:30 am. The female flower and male flower are covered with butter paper cover on the previous day of anthesis. The male flowers were carefully harvested from male genotypes and pollen carefully dusted on stigma of the female flowers after this the flowers were covered with butter paper and tagged.

3.2.4 Evaluation of fruit set

The fruit set was calculated based on number of flowers pollinated cross pollination and days counted for first set after dusting of pollen grains on female ovaries. The pollination was done in the early morning hours.

Plate 1: Field view of experimental plot

The female flower was selected and bagged on previous day and on the very next day cross pollination was carried out. The day, date and time was recorded and female flowers were with tag written on it and further counting of days were done with the help of secured data. The days and cross pollination attempts were recorded, from the day on which the pollens are dusted on female ovaries.

The fruits were harvested at maturity and seed set in mature fruit was also observed.

3.2.5 Laboratory chemicals, glassware and plastic ware

The chemicals used in this 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 Banglore Genei Ltd. All the plastic wares were obtained from Sigma Aldrich Chemicals Pvt. Ltd.

3.2.6 Equipment and machinery

The present research work was carried out using molecular biology facilities and equipments available at CPBMB, College of Horticulture, Vellanikkara. Centrifugation was done in high speed refrigerated centrifuge Kubota 6500. Nanodrop® ND-1000 spectrophotometer was used for the estimation of quality and quantity of DNA. The DNA amplification was carried out in Proflex thermal cycler (Applied Biosystem, USA). 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.

3.2.7 DNA barcoding

DNA barcoding of *Benincasa* and *Praecitrullus* genera were carried out with two different primer sets.

3.2.7.1 Total Genomic DNA isolation

Young, pale and tender green leaves (first to third from the tip) were collected in the early morning period and cover with aluminium foil, avoiding exposure to sunlight and

maintaining cool temperature with the help of ice from each individual plants in each accession and brought to the laboratory. The leaf surface was cleaned by washing with sterile water wiping with 70 per cent ethanol. CTAB method detailed by Rogers and Benedich (1994) was used for the extraction of total genomic DNA. The reagents used for DNA isolation are presented in Annexure II.

Procedure:

- The leaf sample was collected at initial stage of growth, during fruiting stage starch accumulates and there is fewer chances to get purified DNA. At this time CTAB (5-10 percent) concentration can increase to purify DNA from homogenized sample.
- One gram of cleaned leaf tissue was placed in a chilled mortar and liquid nitrogen was added to obtain crystallized and fine powder form of leaf sample.
- 1ml of extraction buffer (5x), 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 2 ml centrifuge tube and 1ml of pre-warmed extraction buffer was added (total 700µl).
- The contents were mixed well and incubated at 65^0C for 20 to 30 minutes with occasional mixing by gentle inversion.
- Equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed by inversion to emulsify. Spun at 10,000 rpm for 15 minutes at 4^0C.
- 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 5 per cent CTAB solution and equal volume of chloroform: isoamyl alcohol (24:1) were added and mixed by gentle inversion and subsequently centrifuged at 10,000 rpm for 15 minutes at 4^0C.

- The supernatant 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^{0}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^{0}C.

3.2.7.2 Assessing the quality of DNA by electrophoresis

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

Reagents and equipment

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

Procedure for gel casting and AGE

1. For casting the gel, the 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.
2. Prepared 0.8 per cent agarose in a glass conical flask by dissolving 0.8g agarose in 100 ml 1X TAE buffer. The contents were heated until agarose get dissolved and clear crystal solution obtained.
3. After dissolving agarose, 4µl ethidium bromide was added to the concentration and mixed well.
4. This warm gel solution was poured into tray to a depth of about 5 mm and the gel was allowed to solidify for about 30 to 45 minutes at room temperature.
5. After, the gel is cast, comb is gently removed and the gel is placed in electrophoresis chamber along with the tray maintaining enough distance from both sides, pouring

enough electrophoresis buffer (1X TAE buffer), just enough to cover the wells formed in the agarose gel.

6. Samples for electrophoresis were prepared by adding 1µl of 6X gel loading dye for every 2µl of DNA sample by mixing and were loaded with 3µl DNA sample per well. The molecular weight marker was loaded (λDNA *EcoRI*/ *Hind*III double digest) in first lane.
7. The electrophoresis was carried at 70 volts until dye has migrated two third length of the gel.
8. The gel profile was examined under U.V. transilluminator for checking 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.

3.2.7.3 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 image of a gel was captured using the PDQuest controls in the imaging device window and subjected to further analyses.

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

The purity of DNA was further checked using NanoDrop ND-1000 spectrophotometer. Nucleic acid shows absorption maxima at 260 nm whereas, proteins show peak absorbance at 280 nm. 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 protein contaminated and 2.0 reading shows that it consist of RNA contamination. The quantity of DNA in the pure sample was calculated using the relation.

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 pedestal.
- The reading was set to zero with sample blank.
- Subsequently 1 µl of sample was pipetted out onto the measurement pedestal and selected the 'Measure' option.
- When the measurement was 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.

3.2.3 DNA barcoding in *Benincasa* and *Praecitrullus*

Primer sets for 2 common barcoding loci were used to assess their compatibility and to generate the banding patterns in various *Benincasa* and *Praecitrullus* accession. These primer sets were *matK* (Maturase K) and *ITS*2 (Internal Transcribed Spacer). Amplification of DNA from all the lines of *Benincasa* and *P. fistulosus was* attempted separately with the two selected primers.

3.2.3.1 DNA amplification

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 Proflex Thermal Cycler (Applied Biosystems,USA).

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 were used for amplification of DNA (Table 3.3).

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

a) Genomic DNA
b) 10X taq assay buffer A
c) dNTP mix (10 mM each)
d) Taq DNA polymerase (3U)
e) Primer (10 pM)
f) Autoclaved distilled water

The PCR amplification was carried out with the following thermal profile

Initial denaturation - 94^0C for 1 minute

Denaturation - 94^0C for 30 second

Primer annealing - 59.5^0 C for 40 sec

Primer extension - 72^0C for 1 minute and 30 second

40 cycles

Final extension - 72^0C for 5 minutes

4^0C for infinity to hold the sample

3.2.3.2 Screening of primers and analysis

The universal primer (*matK* and *ITS2* loci) was already screened by Girme (2014) for cucurbit species and thus used in this study which gives better amplification. The subsequent 2 primer using the forward and reverse primers for *matK* loci and *ITS2* loci were attempted to amplify the loci, from 9 *Benincasa* and 3 *P. fistulosus* accessions.

The amplified products were electrophoresed on 1.5 to 2 per cent agarose gel using 1 X TAE buffer stained with ethidium bromide. The profile was visualized under gel

documentation system and documented for further analysis. The documented profiles were carefully examined for amplification of bands in comparison with 100 bp-3kb ladder. The average size of the expected band in all these combination of primer was around 900 bp for matK and 799 for *ITS2* loci.

Table 3.3: List and sequence of DNA barcoding primer used in the study

SI SNo.	Primer	Nucleotide Sequences	Reference
1	*matK* F1	5'AGGTTTGGAGTCATTGTGG 3'	Girme, 2014
2	*matK* R1	5'GAATCGATCCAGGTCGTCTT 3'	
3	*ITS2 F*	5' GCGATACTTGGTGTGAAT 3'	Jing *et al.*, 2011
4	*ITS2 R*	5' GACGCTTCTCCAGACTACAAT 3'	

3.2.3.3 PCR product sequencing

A total of 12 genotypes were selected on the basis of distinct a morphological character which was further amplified with two different primer *matK* and *ITS* loci. The product having single band assay with standard volume was sent for sequencing. The sequencing is carried out at SciGenome Pvt. Ltd., Cochin, Kerala.

3.2.4 Data analysis

3.2.4.1 Morphological data analysis

The morphological characters were recorded with the help of descriptor developed by NBPGR, New Delhi. The character specifically helps to differentiate the ashgourd and tinda genotypes inter-specifically and intra-specifically. The plant characters are grading from fruit shape, leaf and stem pubescence density as well seed characters, which shows variation in ashgourd genotypes.

The morphological data analysis is carried out by a software "Minitab" developed by Pennsylvania University by researchers in 1972. It is a command-and menu-driven package for statistical analysis which is easily available for windows.

3.2.5 Data analysis using *In-silico* tools

3.2.5.1. Sequence analysis and annotation

The forward and reverse sequences obtained for each accession were merged to form contigs using CAP3 sequence assembly programme (http://doua.prabi.fr/software/cap3). The *matK* and *ITS2* sequences which are protein coding sequences were analysed for the presence of stop codons. For this, the merged *matK* and *ITS2* sequences were aligned using "align by muscle" option in MEGA 7. These aligned sequences were translated and were compared to the invertebrate genetic code table to assess the presence of stop codons. The stop codons if present were removed using Bio-Edit software.

3.2.5.2. Analysis of sequence homology

Basic Local Alignment Search Tool (BLAST), a sequence similarity search tool provided by NCBI was utilized to assess the homology of twenty-four sequences to the sequences present in NCBI database. For this the individual accessions were subjected to nucleotide BLAST (BLASTn). The sequences from the database showing maximum identity, query coverage and with least expected value (E value) to the query sequence were identified.

3.2.5.3 Phylogenetic analysis

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

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

3.2.5.4 Genetic divergence within and between species

The inter specific distances among the genotypes 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 7 software. It is available at http://www.megasoftware.net/mega4/mega.html was used for analysis. The 'Distance | Compute Pair wise' command generates the pair wise distances in MEGA.

3.2.5.5 Barcoding gap assessment

The distributions of inter-specific variability were compared using DNA barcoding gap. All the selected 12 genotypes for two different loci, each were sequenced and have aligned with the software ClustalW Omega. This mark "*" indicated the conserved sequences and absence of this mark in the aligned sequences is considered as gap. The barcode gap represents the uniqueness of the genotype.

3.2.5.6 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 *matK* and *ITS2* sequences as query sequences.

Graphical display had shown where the query which 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.

3.2.5.7 Submission to NCBI

An account was created in NCBI BankIt (http://www.ncbi.nlm.nih.gov/WebSub/?tool=genbank) and by logging into the account, twenty-four sequences were submitted. In step wise procedure the details regarding authors, title of the work, address for correspondence, sequence information, name of organism, source modifier details like country, date of collection of sample, sample description were also provided while submission of the data.

3.2.5.8 Submission to Barcode of Life Data (BOLD) system

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

Results

4. RESULTS

The study on DNA barcoding in *Benincasa* and *Praecitrullus* genera was undertaken at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Vellanikkara. The objective of the study was to develop DNA barcodes which can easily differentiate the genus *Benincasa* and *Praecitrullus* with each other. The results of various aspects of the investigations are presented in this chapter.

4.1 Morphological characterization

Morphological characters were recorded in plants grown in the field at various stages of crop growth and were directly used for data analysis. The morphological parameters observed for the different genotypes are listed in Table.4.1.

4.1.1 Characters

The ashgourd and tinda genotypes can be successfully differentiated based on the morphological characters such as leaf size, leaf pubescence nature, leaf pubescence density, leaf margin, stem pubescence, peduncle attachment, fruit shape, stem end shape and blossom end fruit shape, fruit ridge shape and flower size., etc.

1. Early plant vigour

The early plant vigour was recorded at 30 days after sowing. Twenty-one genotypes of ashgourd had very good plant vigour and only five genotypes had good plant vigour. All tinda genotypes had very good plant vigour.

2. Plant growth habit

The growth habit was recorded at fully grown plant. There are three kind of growth habit observed in ashgourd genotypes, medium viny, long viny and short viny. Eleven of the ashgourd genotypes had long vines, nine genotypes had medium vines and six genotypes had short vines. All the tinda genotypes had short vines.

3. Stem pubescence

The stem pubescence was observed and recorded at peak fruiting stage. In ashgourd genotypes seven different types of stem pubescence was observed. In sixteen

genotypes hispid kind of stem pubescence was observed, three genotypes West Eleri, Areeparambu and Nashik genotypes had downy stem pubescence and Kanjirampoyil had tomentose kind of stem pubescence. The genotype Mallapally had puberulent stem pubescence nature, two genotypes Acc. No.5 and Acc. No.10 had pilose kind of stem pubescence, while Acc No.9 and IC 0596992 had villous kind of stem pubescence. Only one genotype (Elavancheri) had hirsute kind of stem pubescence. All the tinda genotypes had hispid kind of stem pubescence nature.

4. Stem shape

The stem shape was observed and recorded at peak fruiting stage. The stem shape is observed to be angular in all ashgourd and tinda genotypes.

5. Tendril

The tendril was found to be present in all genotypes of ashgourd and tinda.

6. Tendril type

The tendril type was found to be coiled in all genotypes of ashgourd and tinda.

7. Tendril branching

The tendril was found to be branched in all genotypes of ashgourd and tinda.

8. Leaf margin

The leaf margin observed in ashgourd genotypes are serrated, double serrated and deep serrated. Serrated leaf margin was observed only in genotypes Kottamuri, Kanjirampoyil, Kannara and Thanniam double serrated leaf margin was observed in thirteen genotypes and deep serrated leaf margin was observed in nine genotypes. All tinda genotypes had incised leaf margin.

9. Leaf shape

The leaf shape was recorded at full foliage stage. The leaf shape of ashgourd and tinda genotypes was cordate.

10. Leaf size

The leaf size was observed as small, medium and large in all ashgourd genotypes. Majority of the ashgourd genotypes had medium size leaf (6.2-8.5 cm), Kottamuri genotype had smallest leaf size (5.0 cm), and all other ashgourd genotypes had large sized leaves (10-11.25 cm). The leaf size of tinda genotypes were small ranging from 4.1-4.5 cm.

11. Leaf pubescence nature

The leaf pubescence was observed to be soft, intermediate, hispid and puberulent in ashgourd genotypes. Soft leaf pubescence was observed in six genotypes, intermediate kind of pubescence was observed in fifteen genotypes, hispid pubescence was observed in four genotypes and only one genotype Mallapally had puberulent kind of stem pubescence. The sparse kind of leaf pubescence nature was observed in all tinda genotypes.

12. Leaf pubescence density

The leaf pubescence density was observed to be sparse, intermediate and dense in ashgourd genotypes. In fifteen ashgourd genotypes the pubescence was observed as sparse, dense kind of leaf pubescence was observed in ten genotype and only a single genotype Kottamuri had intermediate kind of leaf pubescence density. All the tinda genotypes had intermediate kind of leaf pubescence density.

13. Peduncle attachment

In ashgourd genotype, two types of peduncle attachment was observed viz., soft and flared, hard and flared. Three ashgourd genotypes Kottamuri, Nedungadapalli, Mallapally had soft and flared kind of peduncle attachment and remaining twenty-three genotypes had hard and flared kind of peduncle attachment. In tinda genotypes, hard and flared type of peduncle attachment was observed.

14. Peduncle shape

All the ashgourd and tinda genotypes had angular kind of peduncle shape.

15. Peduncle length

In ashgourd genotypes, the peduncle length ranged from 4.6 cm to 13.8 cm. All the tinda genotypes had a peduncle length of only 2.5 cm.

16. Fruit shape

Fruit shape is considered as an important feature for differentiating genotypes. Out of this twenty-six ashgourd genotypes, twenty-one genotypes were observed to have club shape of fruit, for one genotype Vellarikundu had spindle shaped fruit and two genotypes Kanjirampoyil and Elavancheri had cylindrical shaped fruits. The genotypes Kannara and Areeparambu had oblong shaped fruit. Ridged shape fruit was observed only in the genotype IC 0596989. All the tinda genotypes had small and round shaped fruits.

17. Stem-end fruit shape

The stem-end fruit shape in ashgourd was observed in three different forms viz. depressed, flattened, and rounded. Eleven of the ashgourd genotypes had rounded shaped stem-end fruit shape, four genotypes had depressed stem-end fruit shape and the remaining eleven genotypes had flattened stem-end fruit shape. All the tinda genotypes had rounded stem-end fruit shape.

18. Blossom-end fruit shape

The blossom-end fruit shape in ashgourd was observed in three different forms viz. depressed, flattened and rounded. Rounded blossom-end fruit shape was observed in eleven genotypes, flattened blossom-end fruit shape was observed in twelve genotypes and depressed shape was observed in three varities. All the tinda genotypes had rounded blossom-end fruit shape.

19. Fruit length

Fruit length differed significantly with fruit size in ashgourd genotypes. The fruit length ranged from 4.6 cm to 50.4 cm. Only one genotype Thanniam measured 4.6 cm.

Sixteen genotypes had medium size fruit ranging from 13.0-29.3 cm. Eight genotypes had long fruits ranging from 31.5 cm to 40.0 cm. The genotype West Eleri had the largest fruit (50.4 cm). All the tinda genotype had fruit length in the range of 6.0-8.1 cm.

20. Fruit width

Fruit width also differed significantly among genotypes. In ashgourd genotypes, the fruit width range from 10.3 to 60.1 cm. The genotype Thanniam had only 10.3 cm fruit width. Twenty-one genotypes fruit width between 15.0 cm to 49.5 cm. Four of genotypes measured had high fruit width in the range from 50.3 cm to 60.1 cm. West Eleri had maximum width at 60.1 cm. All the tinda genotypes had fruit width in the range of 8.0 cm to 9.2 cm.

21. Fruit weight

The ashgourd genotypes had fruit weight ranging from 0.191 kg to 12.0 kg. The genotype Thanniam had the lowest fruit weight. Nine genotypes had fruit weight ranging from 0.307 kg to 0.979 kg. Thirteen genotypes had fruit weight ranging from 4.500 kg to 5.983 kg. The genotype West Eleri had the maximum fruit weight (12.0 kg). All the tinda genotype had average fruit weight of 250 kg.

22. Flower size

Ashgourd is a monoecious crop, male and flower appears on same plant and both the flowers appear in same yellow colour. The male flowers are 5.0-15 cm long with pedicel. Female flowers are 2.0-4.0 cm long. Majority of the ashgourd genotypes had female flowers with inferior shape. In tinda flowers are actinomorphic and unisexual. The flowers are small, 3.0 cm in diameter.

23. Node number at which first female flower appear

The number of days for first female flower appearance was counted from the date of planting to the date when first female flower appeared on the lowest node. The ashgourd genotype IC 0596992 recorded it first female flower appearance at node number four and genotype Nedungadapalli recorded its first female appearance at node

number eighteen. In tinda genotypes, MAHYCO tinda recorded its first female flower at node number first.

24. Days to 50% flowering

The number of days for 50% flowering of plants was counted from the date of planting. The number of days recorded for 50 per cent flowering was 22-30 days in ashgourd genotypes. The genotype IC 0596986 took 22 counts of days and genotype Kanjirampoyil took more which is 30 counts of days to 50 per cent flowering. In tinda genotypes, 21-24 days was recorded for 50 per cent flowering.

25. Petiole length

In ashgourd genotypes, the petiole length ranged from 8.1 cm to 24.5 cm. The genotype Kottamuri had the minimum length (8.1 cm). The maximum length of petiole was measured in variety Cheramangalam (24.5 cm). In tinda genotypes, the petiole length ranged from 1.5-2.0 cm.

26. Fruit skin colour

The fruit skin colour in ashgourd genotypes varied from light green to dark green. Light green skin coloured fruits were observed in genotypes Nashik, Acc. No.5, Elavancheri, Thanniam and KAU local. The remaining twenty one genotypes had dark green skin coloured fruit. All the tinda genotypes had light green fruit skin colour.

27. Fruit skin lustre

The fruit skin was classified as matt type in the twenty-six ashgourd genotypes. The matt appearance of fruit skin is due to formation of ashy coating at maturity. In tinda genotypes, all the fruits had glossy fruit skin.

28. Fruit pubescence density

The fruit pubescence was absent in all the twenty-six genotypes of ashgourd. The tinda genotypes had sparse pubescence on fruit skin.

29. Fruit ridge shape

The fruit ridge shape was observed as superficial in twenty-five genotypes of ashgourd, only the genotype IC 0596989 had deep grooved ridged shape fruit. In tinda genotypes, no ridge was observed on fruits.

30. Number of ridges per fruit

There was no ridges observed on fruits of twenty-five genotypes of ashgourd, only the genotype IC 0596989 had twelve ridges on the fruit.

31. Flesh texture

The flesh is observed to be spongy in all ashgourd genotypes. In tinda varities, the flesh texture was firm.

32. Flesh colour

All the ashgourd genotypes had creamish flesh colour while all the tinda genotypes had whitish coloured flesh.

33. Seed lustre

The seeds lustre was observed on mature and dried seeds. In ashgourd genotypes, three kind of seed lustreness was observed viz., glossy, intermediate and matt. The glossy seed lustre was observed in ninteen ashgourd genotypes, intermediate kind of seed lustreness was observed only in the genotypes Areeparambu and Kannara. Matt seeds lustreness was observed in five genotypes of ashgourd. All the tinda genotypes had seed with matt lustre.

34. Seediness

The seediness was observed as low, medium, high and very high based on seed counts. Majority of the ashgourd genotypes had high seed count, some of them had medium count of seeds and only Perumbilisherry genotype had low seed count numbers. All the tinda genotypes had high seed count numbers.

35. Number of primary branches

Majority of the ashgourd genotypes had three primary branches, while ten of ashgourd genotypes had four primary branches, one ashgourd genotype Areeparambu had two primary branches, the genotype IC 0596995 had five primary branches. The maximum number of seven primary branches was observed in genotype Mallapally. All the tinda genotypes had three primary branches.

36. Days to first harvest

It was recorded as number of days from date of planting to the date of first marketable fruit harvest. All the ashgourd genotypes took period of 44 days to 48 days for its first harvest. The genotype IC 0596987 took 44 days for its first harvest, while three genotypes took a count of 48 days for its first harvest of fruits. All the tinda genotypes took 35- 37 days of period for first harvest.

37. Days to last harvest

It was recorded as number of days from the date of planting to the date of last marketable fruit harvest. All the ashgourd genotypes took a period of 46 days to 60 counts of days for last harvest. Three genotypes of ashgourd recorded a period of 48 days for its last harvest of fruits, while majority of ashgourd genotypes required 49 days for its last harvest of fruits, only genotype KAU local took a maximum period of 60 days for its last harvest. All the tinda genotypes took 34 to 38 counts of days for its last harvest.

38. Number of fruits per plant

The total numbers of fruits were counted during each fruit picking. The number of fruits counted for ashgourd genotypes varied between 2 to 10. The maximum number of 10 fruit was counted for KAU local, while the genotype IC 0596992 counted had only 2 fruits. All the tinda genotypes had 4 to 6 number of fruits counted per plant.

4.2 Cross compatibility

For testing cross compatibility the twenty-five ashgourd genotypes were used as a male parent and they were crossed with variety KAU local as the female cultivar. All the crosses were successful and there was good fruit set. Reciprocal crosses were also carried

out with twenty-five genotypes as female cultivars and KAU local variety as male parent. The reciprocal crosses were also successful and good fruit set was observed. None of the ashgourd genotypes produced any successful cross with tinda either as male or female parent.

4.3 Morphological analysis

The morphological characters were analysed with the 'Minitab' software. The accessions were classified on the basis of the morphological observations using the descriptor developed by NBPGR, New Delhi.

Ashgourd accessions exhibited enormous diversity in morphological characters, which results into formation of clusters and subclusters.

The tinda genotypes were less diverse, thus formed a single and separate cluster.

The ashgourd, genotypes, could be grouped into three clusters. Cluster I represented five genotypes viz., Kottamuri, Nedunagapalli, North Paravur, Perumbilisherry and Mallapally within the cluster Kottamuri and Nedungadapalli showed similarity. North Paravur and Perumbilisserry were similar, only one genotype Mallapally formed a single clade, showing no similarity with any of the other genotypes in the cluster.

Cluster II had two subclusters IIa and IIb, the sub clsuter IIa represents four groups. In group one genotypes Vellarikundu and Thanniam were included due to their similarity. Group two was represented by genotypes Areeparambu, Kanjirampoyil, Kannara and Nashik. In group three there were genotypes shows West Eleri, IC 0596985, Acc No.11 and IC 0596986. In group four, there were five genotypes IC 0596995, KAU Local, IC 0596987, IC 059698 and Wayanad which showed no similarity with any of their groups.

Subcluster IIb, formed two clades, one clade represented by ashgourd genotype Cheramangalam, while another clade consists of all tinda genotypes, showing no similarity with any of the ashgourd genotypes. Cluster III, represented five ashgourd genotypes, genotype Acc No.9 showed similarity with Elavancheri, and genotype Acc

No.10 showed similarity with IC 0596992, a single clade represented the genotype Acc No.5 with no similarity with any of the other genotypes.

4.4 Selection of genotypes for barcoding

The nine genotypes of ashgourd and three genotypes of tinda were selected on the basis of the each representing a cluster formed through Euclidean distance and difference in their fruit shape. All the tinda genotypes formed a single cluster however; randomly three tinda genotypes were selected. All the selected genotypes were further sequenced for purpose of barcode formation.

4.5 DNA barcoding

4.5.1 DNA isolation

To obtain good quality and sufficient quantity of DNA, very fresh tender, pale and tender leaves were used (0.5 to 1 g). PVP as an antioxidant and β-mercaptoethanol becomes essential, while DNA isolation.

In molecular biology work, isolation of high-quality DNA is very much important, because contaminants such as proteins, polysaccharides and polyphenols may interfere with enzymes such restriction enzymes (in case of blotting techniques) and *Taq* polymerase (in PCR) (Ausubel *et al.*, 1994).

The tender, pale and green leaves from 26 genotypes of *Benincasa* and 5 genotypes of *Praecitrullus* were used for fine quality DNA extraction. The protocol suggested by (Rogers and Benedich, 1994) is followed for DNA isolation and gives good quality DNA. The DNA isolation is the separation of genomic DNA from polysaccharides, carbohydrates and histones, non-histones protein so far. The leaf of ashgourd contains polysaccharides and phenols and secondary metabolites, such as phenols, alkaloids, flavonoids and terpenes. Due to these compounds, co-isolation of highly viscous polysaccharides degradation of DNA due to endonucleases is inhibited. The inhibitors like polyphenols and other secondary metabolites directly or indirectly interfered with enzymatic reactions. The contaminating RNA precipitates along with DNA causes suppression of PCR amplification (Das *et al.*, 2009). The common procedure is to grind plant tissue in liquid nitrogen and transferred it to preheated

extraction buffer (Dellaporta *et al.*, 1983 and Mohapatra *et al.*, 1992). The DNA quality is observed by using gel electrophoresis unit and examining the gel under U.V transilluminator. However, major threat to obtain purified DNA is protein contamination and shearing of DNA i.e. RNA contamination. The RNA contaminated sample can be treated back by RNase to obtain a good quality DNA.

The quality of extracted DNA has to be analyzed before it could be utilized for PCR reaction. The analysis was done by agarose gel electrophoresis at 0.8% agarose, followed by ethidium bromide staining under U.V. (Sharma *et al.*, 2008). A DNA sample reported as a high quality if it had a band of high molecular weight with little amount of RNA contamination with little smear (Wettasinf and Peffey, 1998). For isolation of DNA using the protocol, good quality DNA was obtained. All the lines had good quality and it was represented in the (Plate. 4.1.2) showing that they are good for PCR analysis. RNA contamination was present in accessions 2, 3, 4, 6, 9, 11, 13, 17, 21, 22 and 29. However, these DNA are good enough for PCR analysis. $A_{260/280}$ was nearly in the range of 1.8-2.0 in some of the lines whereas the for the RNA contamination RNase treatment can be done. After RNase treatment, contamination free DNA was obtained. (Plate.4.1.3.) Thus, this DNA was utilized for further amplification with primers.

4.6 Purification and quantification of DNA

The DNA concentration was measured using a U.V. spectrophotometer at 260 nm. It is a recent and precise method for estimation of DNA quantity. The purity of DNA was determined by the ratio estimated by absorbance at 260 nm to that of 280 nm. The absorbance ratio was calculated as OD at 260/280nm for the samples. The samples with ratio below 1.8 are considered as protein contaminated and ratio above 2.0 considered as RNA contaminated. The majority of the sample measurement ranges between 1.8 and 2.0 were considered as a high quality DNA (Table.4.2).

Fig. 4.1.1 Genetic similarity among the accessions under the study elucidated through Euclidean distances derived from morphological parameters

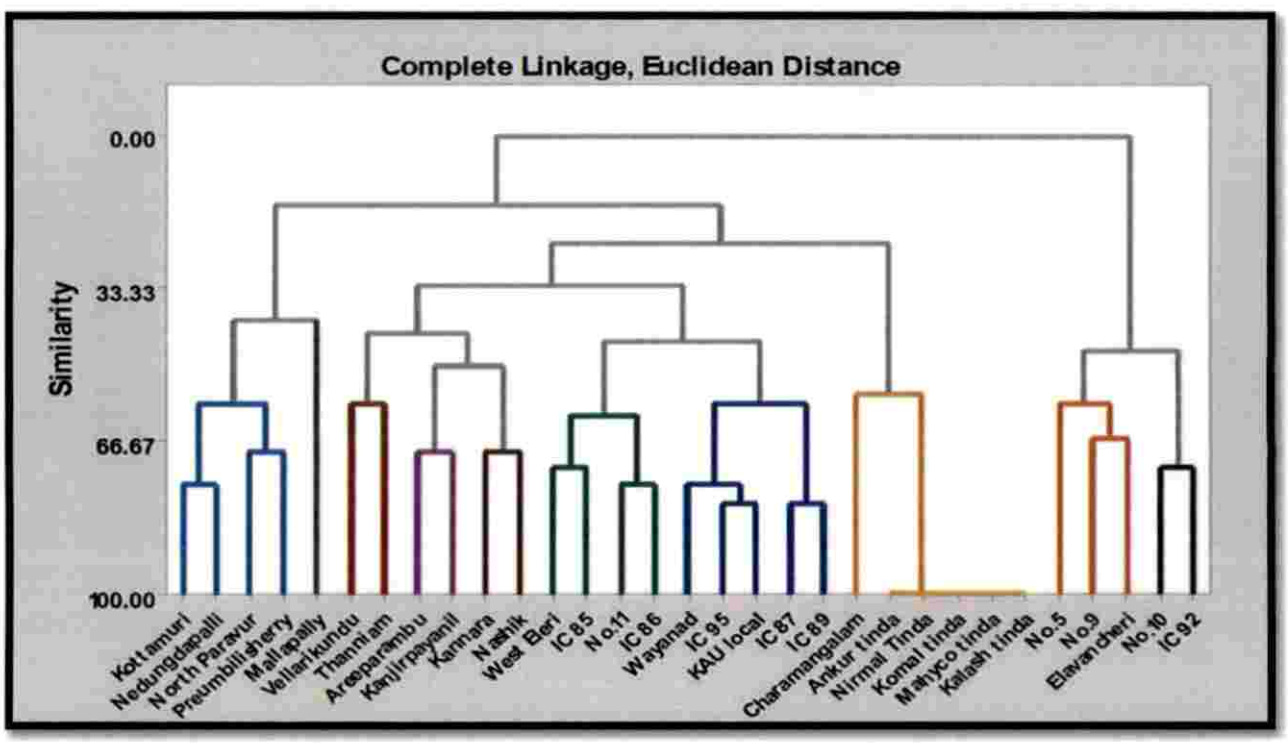


Plate 2: Comparison of ashgourd genotypes Kottamuri, IC 0596995, IC 0596992, Kannara, Nashik and Cheramangalam flowers cross section with KAU Local flower

Plate 3: Comparison of ashgourd genotypes West Eleri, Acc. No.10, North Paravur, Perumbilisserry, IC 0596985 and IC 0596986 flower cross section with KAU Local flower

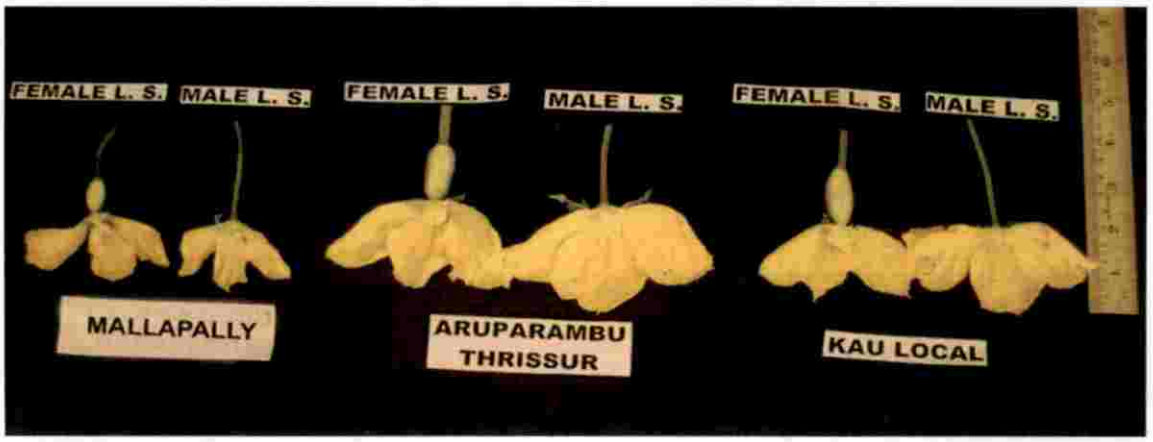


Plate 4: Comparison of ashgourd genotypes Mallapally, Areeparambu, IC 0596987, IC 0596989, Wayanad, Acc. No.5 and Elavanchery cross section with KAU Local flower

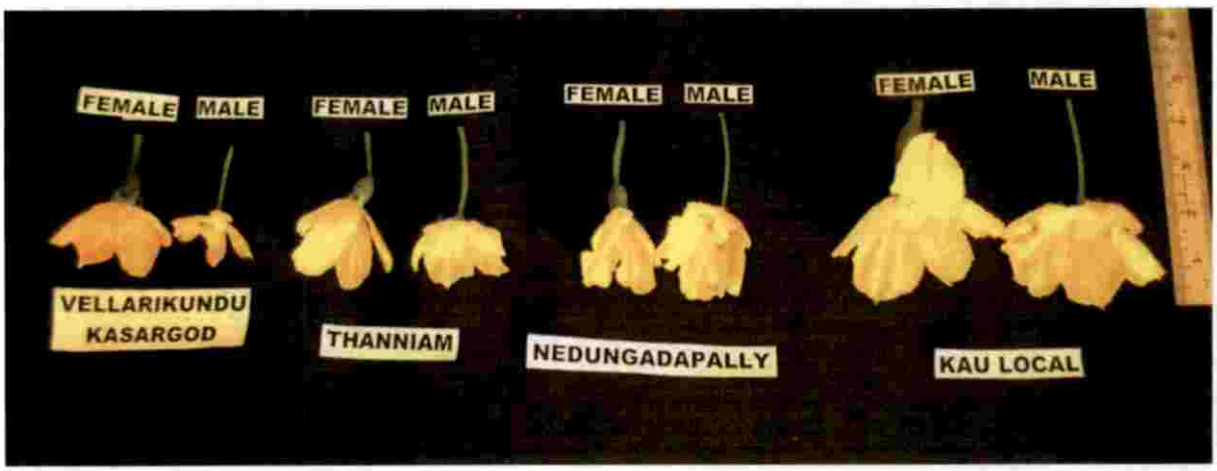


Plate 5: Comparison of ashgourd genotypes Vellarikundu, Thanniam, Nedungadapally, Kanjirampoyil and Acc. No.11 and Acc. No.9 cross section with KAU Local flower

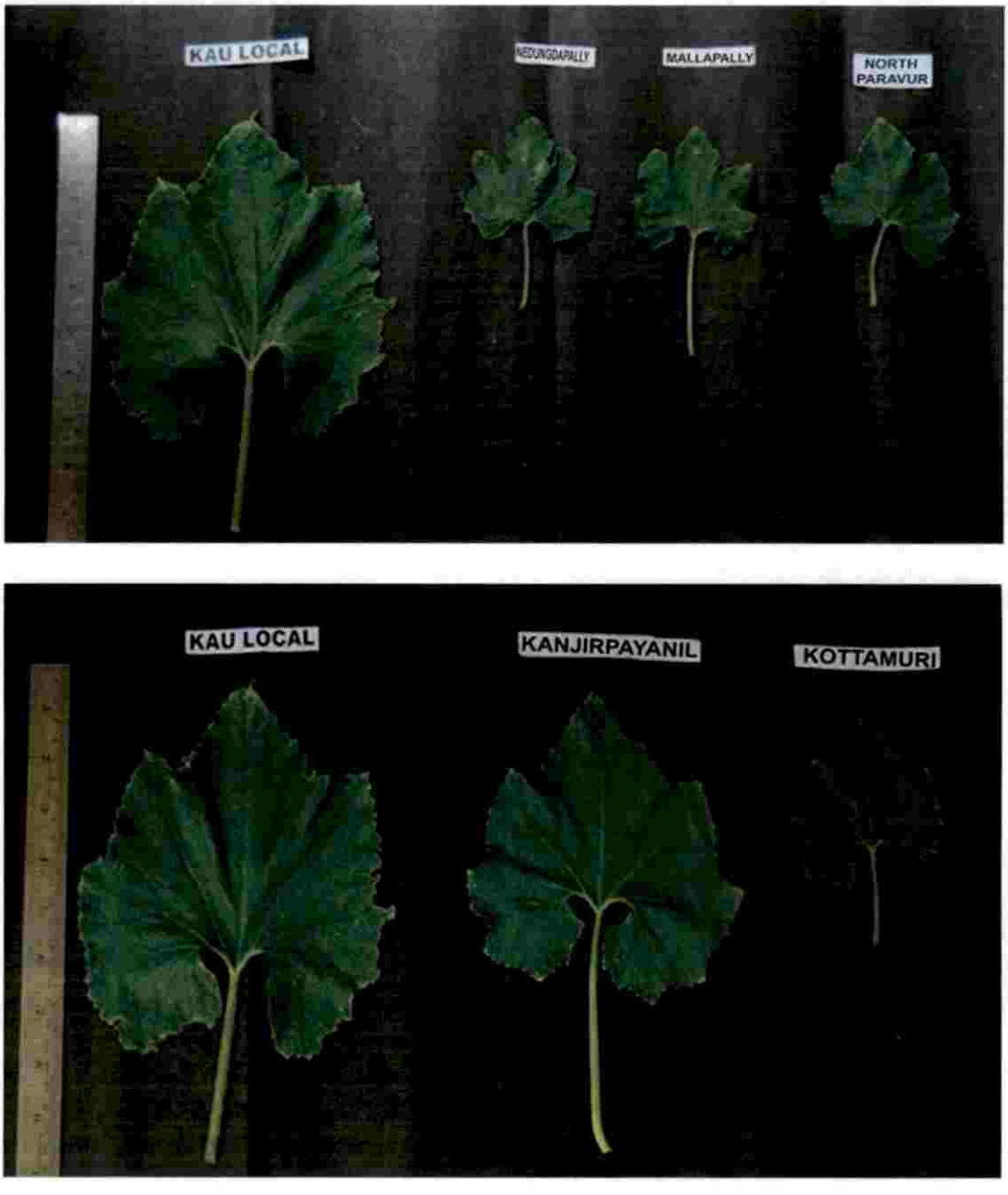


Plate 6: Leaf images of ashgourd genotypes Nedungadapally, Mallapally, North Paravur, Kanjirampoyil, Kottamuri compare to KAU Local

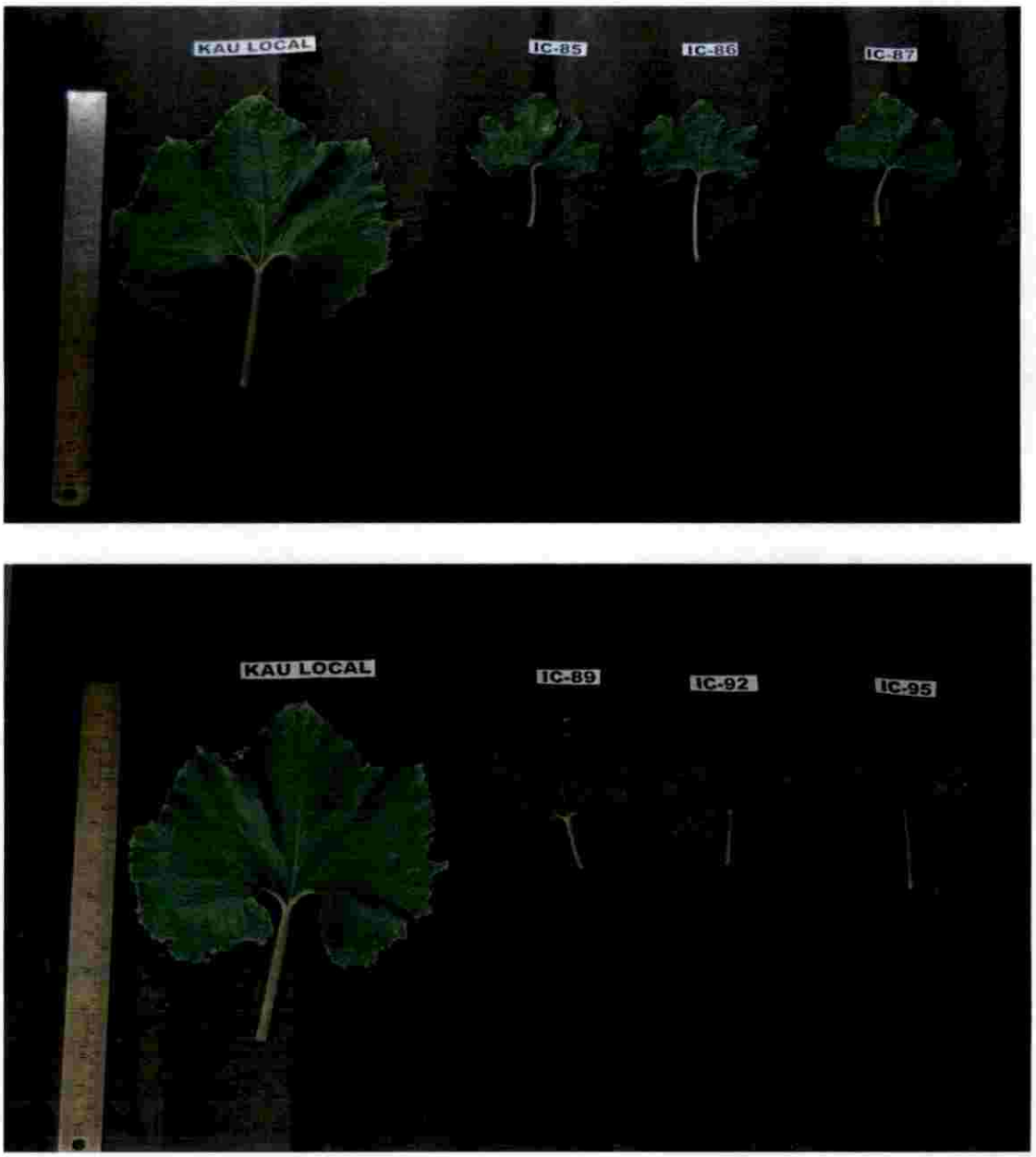
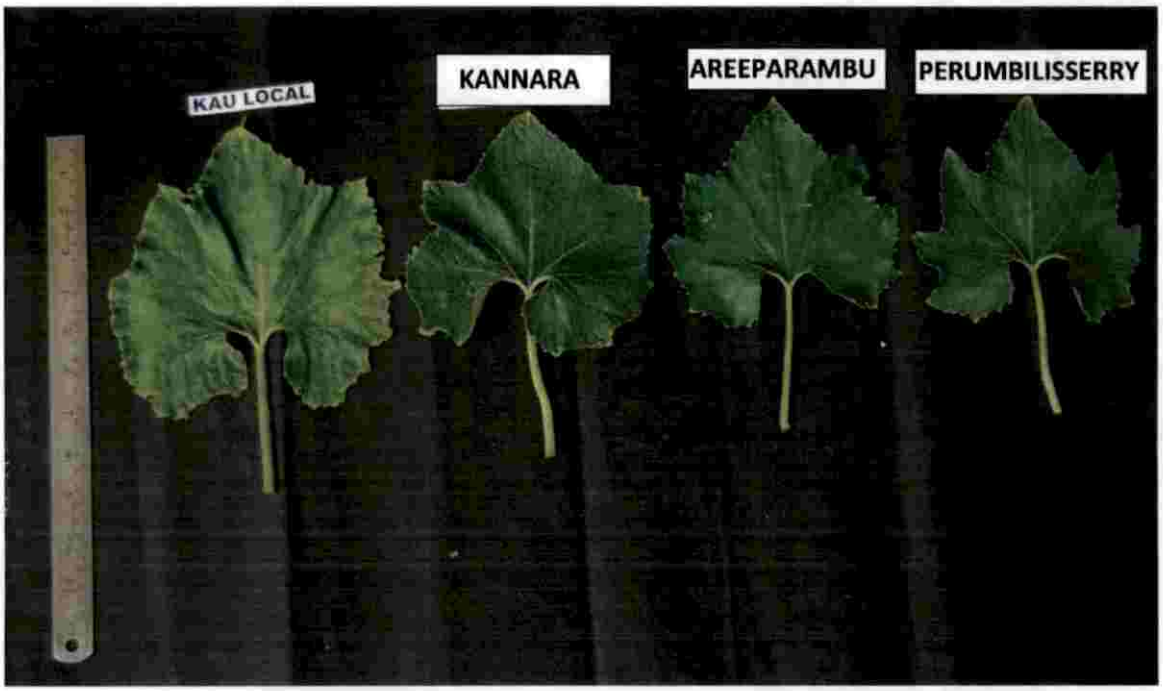


Plate 7: Leaf images of ashgourd genotypes IC 0596986, IC 0596986, IC 0596987, IC 0596989, IC 0596992 and IC 0596995 compare to KAU Local

Plate 8: Leaf images of ashgourd genotypes Elavancheri, Thanniam, Acc. No.9, Acc. No.10 and Acc. No.11 compare to KAU Local

Plate 9: Leaf images of ashgourd genotypes Nashik, West Eleri, Wayand, Charamangalam and Vellarikundu compare to KAU Local



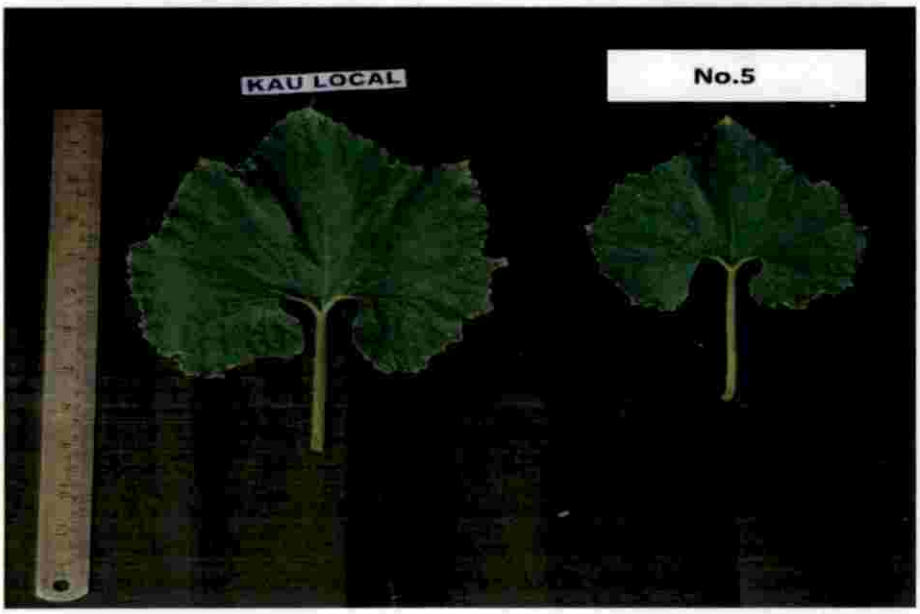
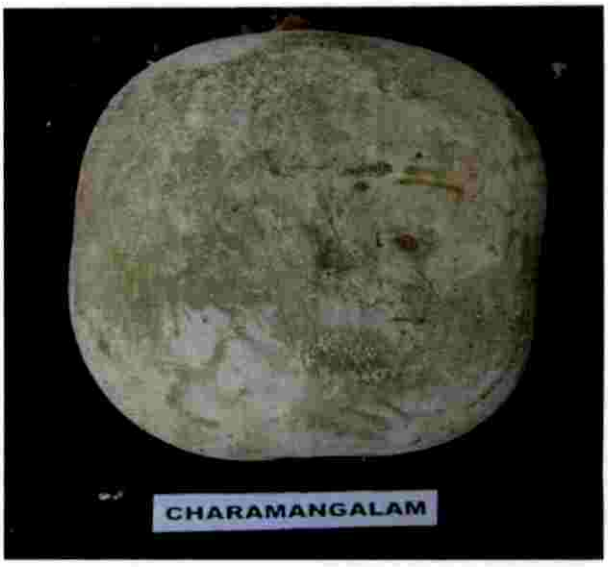


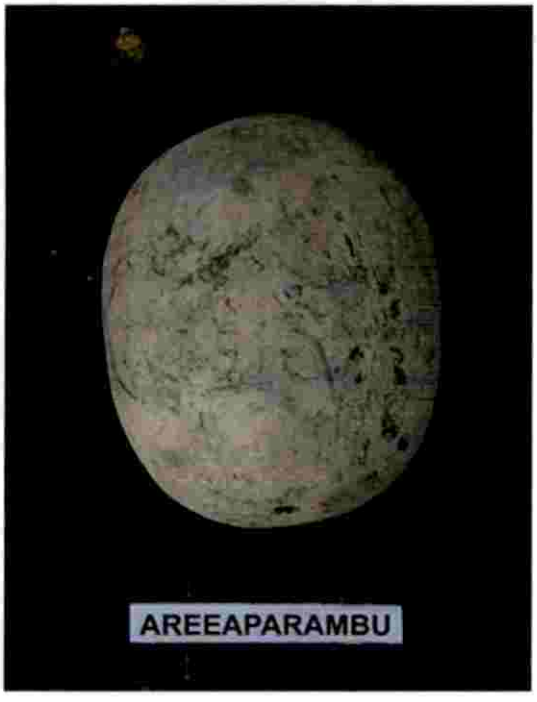
Plate 10: Leaf images of ashgourd genotypes Kannara, Areeparambu and Perumbilisserry and Acc. No.5 compare to KAU Local

Plate 11: Leaf image of tinda genotype

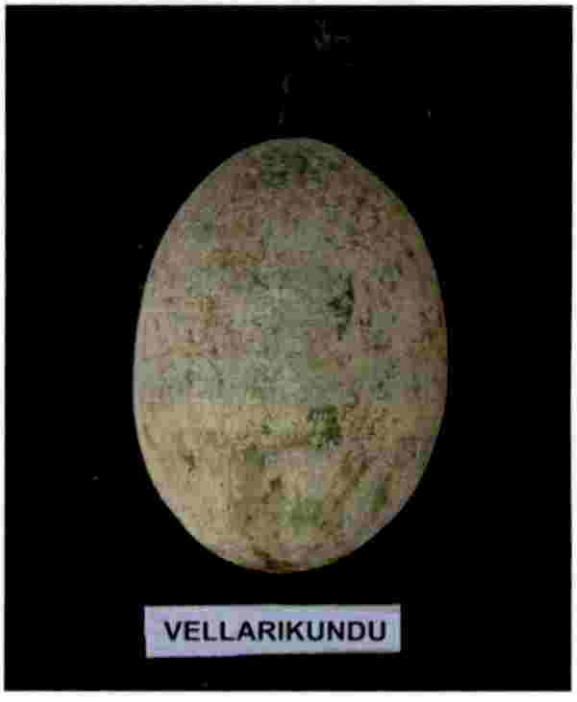
Kottamuri



Cheramangalam

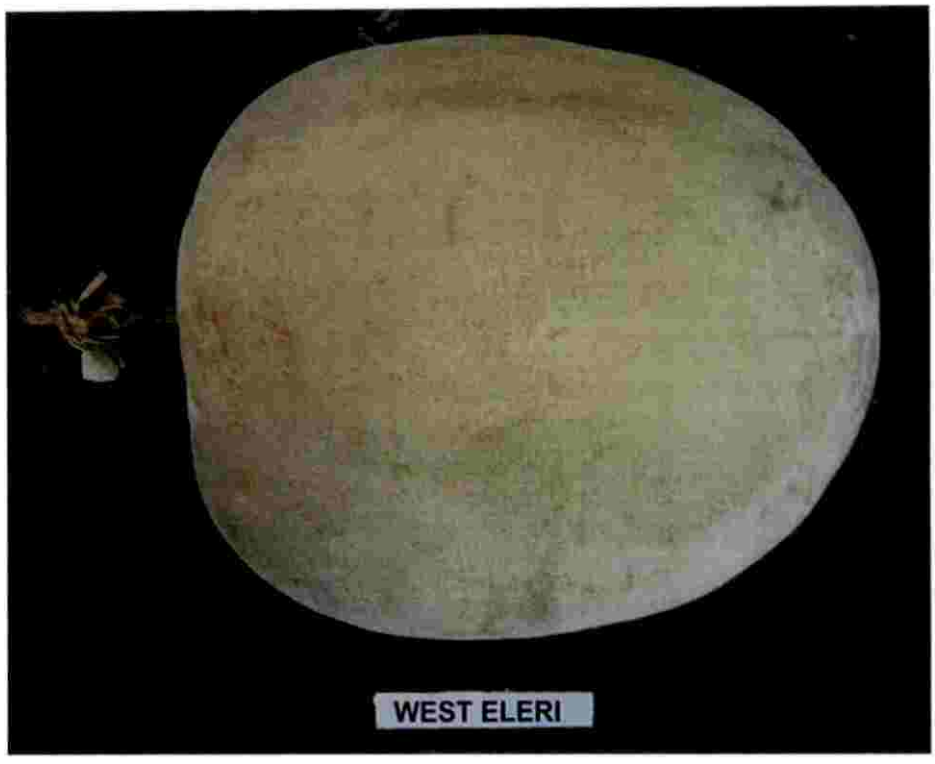


Areeparambu

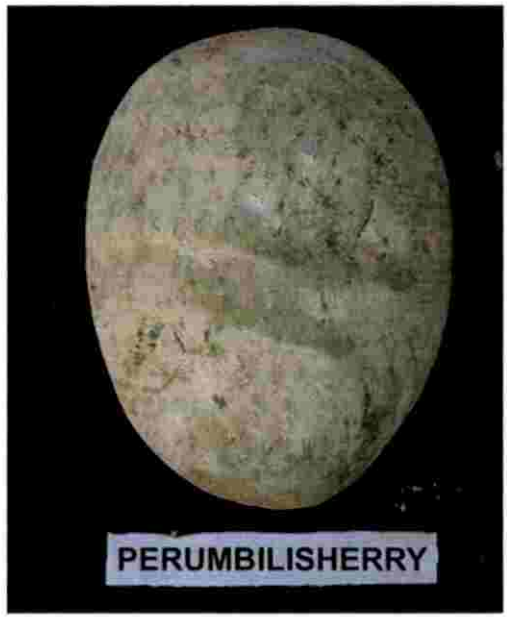


Vellarikundu

Plate 12: Mature fruits of ashgourd genotypes Kottamuri, Cheramangalam, Areeparambu and Vellarikundu



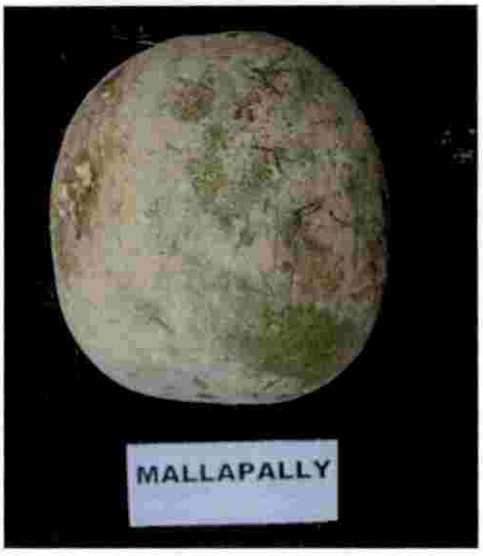
West Eleri



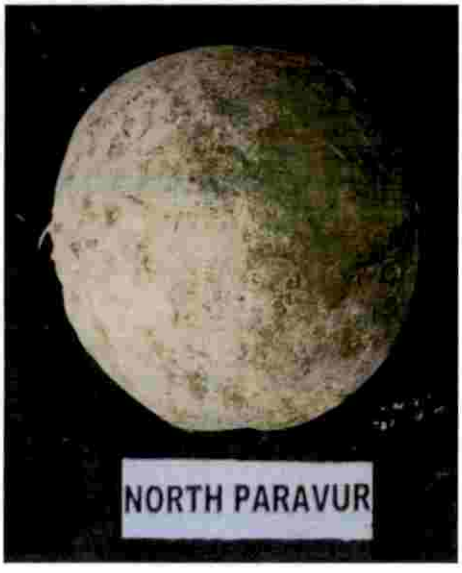
Perumbilisserry

Plate 13: Mature fruits of ashgourd genotypes West Eleri and Perumbilisserry

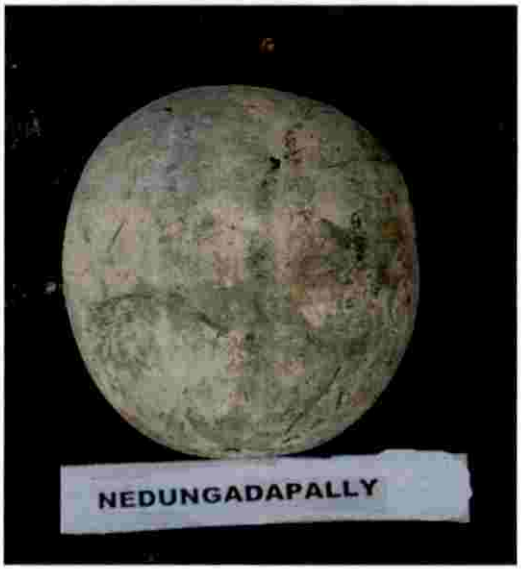
Kanjirampoyil



Mallapally

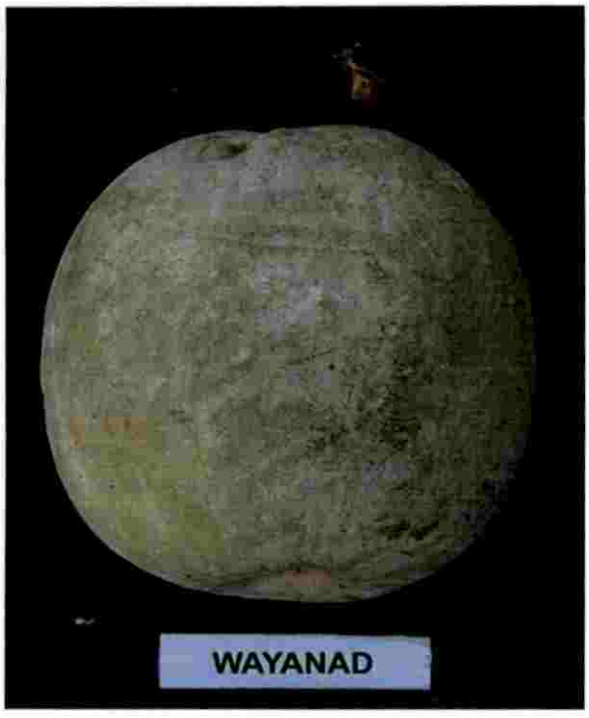


North Paravur

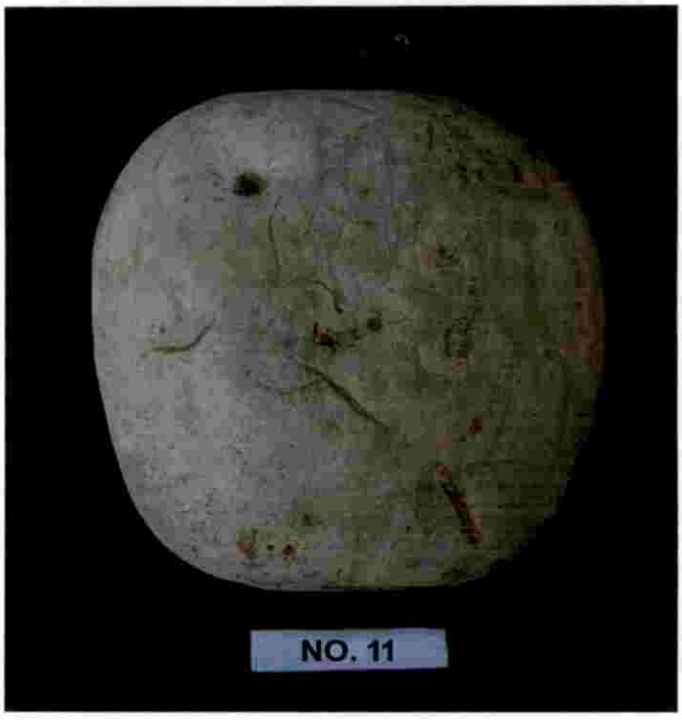


Nedungadapally

Plate 14: Mature fruits of ashgourd genotypes Kanjirampoyil, Mallapally, North Paravur and Nedungadapally

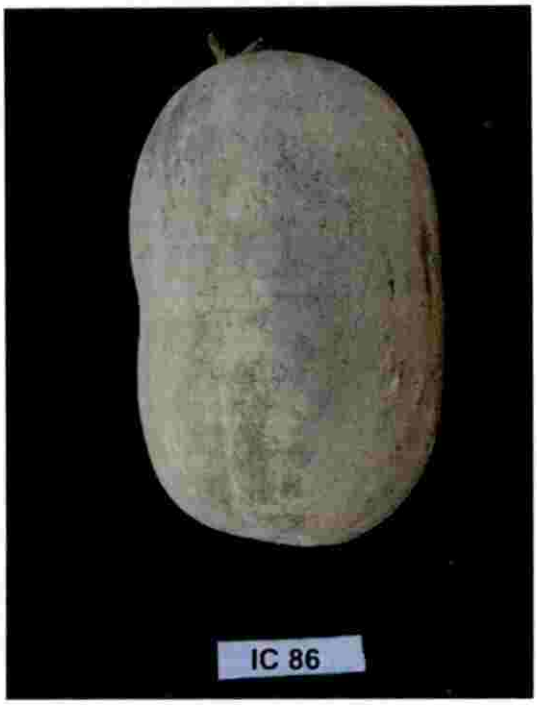


Wayanad

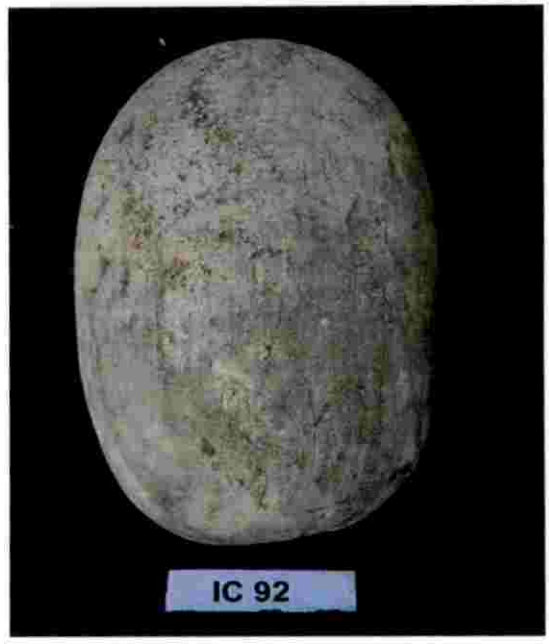


No.11

Plate 15: Mature fruits of ashgourd genotypes Wayanad and Acc. No.11

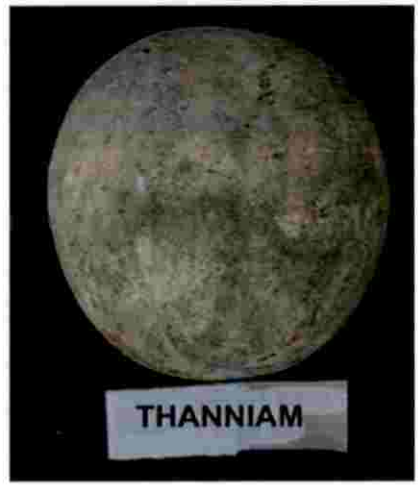


IC 0596986



IC 0596992

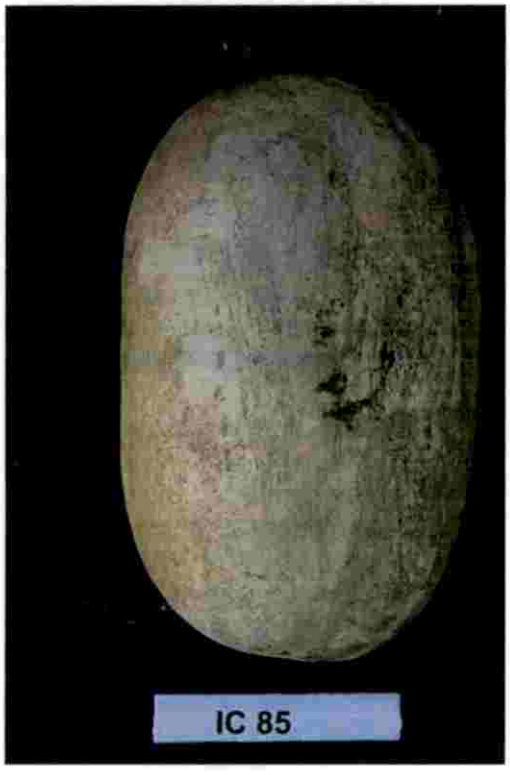
IC 0596987



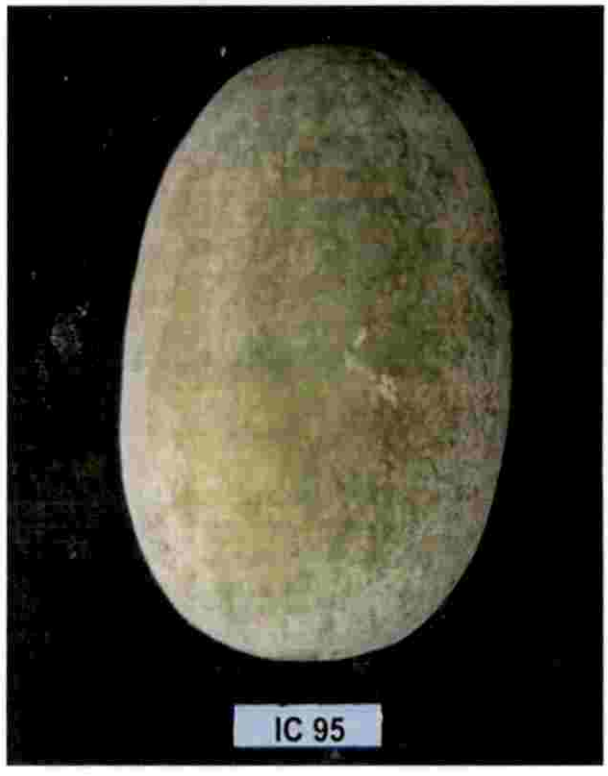
Thanniam

Plate 16: Mature fruits of ashgourd genotypes IC 0596986, IC 0596992, IC 0596987 and Thanniam

Elavancheri

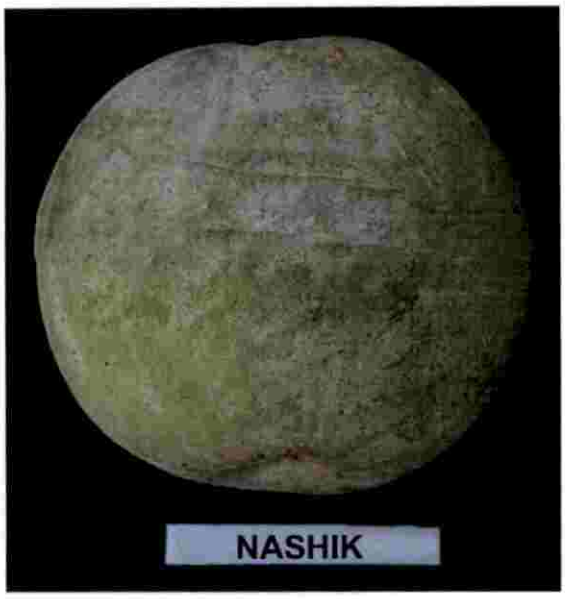


IC 0596985

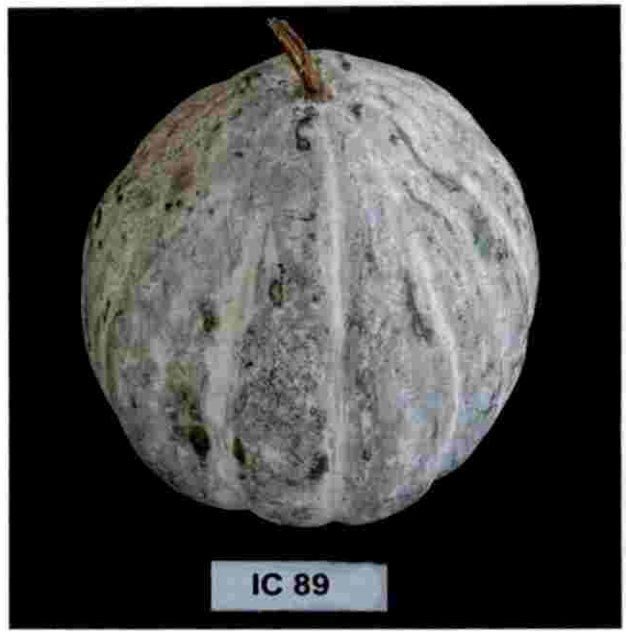


IC 0596995

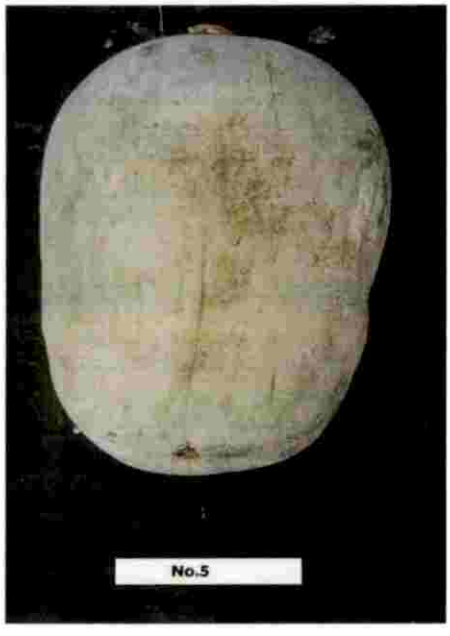
Plate 17: Mature fruits of ashgourd genotypes Elavancheri, IC 0596985, IC 0596995



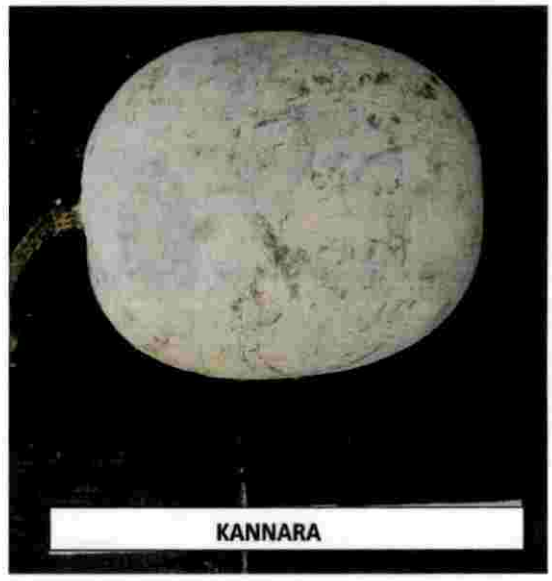
Nashik



IC 0596989



Acc. No.5



Kannara

Plate 18: Mature fruits of ashgourd genotypes Nashik, IC 0596989, Acc. No.5 and Kannara

Acc. No.10

KAU local

Acc. No.9

Plate 19: Mature fruit of ashgourd genotype Acc. No.10, Acc. No.9 and KAU local

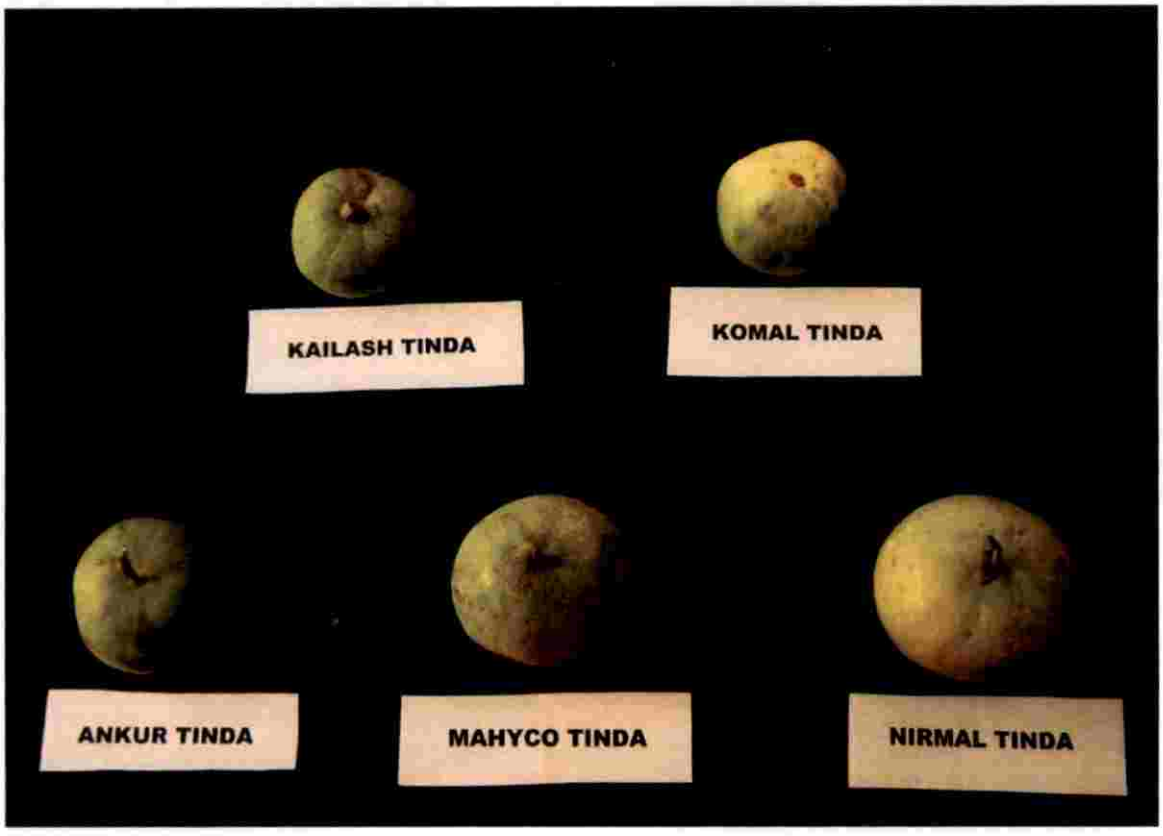


Plate 20: Mature fruits of tinda genotypes Kailash tinda, Komal tinda, Ankur tinda, Mahyco tinda and Nirmal tinda

Table 4.2 Quality and quantity of DNA isolated from *Benincasa* and *Praecitrullus* species as assessed by Nanodrop spectrophotometer method

S.I No.	Species sample	UV absorbance at 260 UV nm A_{260}	UV aborbance at 280 nm A_{280}	$A_{260/280}$	Quantity ng/µl
1	Kottamuri Chalakudy	11.30	4.546	1.75	4285
2	Vellarikundu	9.012	4.7	1.91	1100.5
3	West Eleri	51.28	27.70	1.82	548.8
4	Areeparambu	12.33	6.19	1.92	1549.6
5	Kanjirampoyil	60.11	8.2	1.92	419.0
6	Nedungdapalli	11.18	7.99	1.93	548.8
7	Mallapally Pathanamthittha	39.21	8.19	1.92	938.0
8	North Paravur	9.0	25.20	1.76	938.9
9	Perumbilisserry Thrissur	9.27	5.2	1.87	
10	Cheramangalam Cherthala	8.320	15.00	1.80	487.5
11	Wayanad	11.30	6.46	1.75	2808.9
12	Kannara	57.86	21.75	2.09	1063.1
12	Nashik	50.45	28.70	1.97	1890.8
14	Acc. No. 5	22.52	42.22	1.87	2027.8
15	Acc. No. 9	43.979	6.54	1.84	1042.7
16	Acc. No. 10	57.86	10.67	1.94	339.9
17	Acc. No. 11	85.23	4.97	1.85	477.7
18	IC 0596986	9.5	14.62	1.87	395.2

19	IC 0596986	30.05	2.8	1.99	1139.6
20	IC 0596987	5.22	3.33	1.98	563.5
21	IC 0596989	4.6	5.007	1.97	1681.5
22	IC 0596995	57.7	28.96	1.94	11231
23	IC 0596992	83.45	18.84	1.72	1335.2
24	Elavancheri	37.7	17.22	1.78	1042.5
25	Thanniam	11.88	4.56	1.91	1845
26	KAU local variety	58.24	16.20	1.98	477.5
27	Ankur Tinda	84.34	23.36	1.93	820.0
28	Nirmal Tinda	25.20	27.1	1.70	858.7
29	Komal Tinda	4.979	23.20	1.96	391.1
30	Kailash Tinda	10.381	11.21	1.98	3455.5
31	Mahyco Tinda	9.0	13	1.99	260.7

4.5 Amplification of barcode loci with selected primers

The amplification was obtained at 59.5^{0}C for both the *matK* and *ITS2* locus. The period for primer extension of 40 sec has been extended upto 1 minute and 30 seconds. The improved protocol can be used for amplification of PCR (Plate 23 and Plate 24).

The universal barcoding primers used in this study has mostly amplified a clear distinct band at the range of 980 bp.

I. matK

The primer *matK* shows amplification across all the accessions in the range of 950 bp.

II. ITS2

In PCR, *ITS2* had given bands across all the accessions at 799 bp.

4.6 PCR product sequencing

A total of 9 sequences of the ashgourd and 3 sequences of tinda genotypes were selected based on distinct morphological observation using Minitab software. They are amplified with the primer sets mentioned above. The amplified products were run on 2

per cent agarsoe gel. The amplification was obtained with both of the primer at region near 1000 bp range. It is confirmed that the amplification is specific, as single band appears.

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).

MatK

A total of 26 genotypes *Benincasa* and 5 *Praecitrullus* genotypes were amplified using the *matK* loci. All the genotypes got amplification at the range of 1000 bp. The primer *matK* shows clear distinct band at 1000 bp range.

ITS2 loci

A total of 26 *Benincasa* genotypes and 5 genotypes of *Praecitrullus* were amplified with another locus *Internal Transcribe Spacer* and this primer also performed well to give a better amplification pattern with a clear distinct band at the range of 800 bp.

It is confirmed that both the reported primers (*matK* and *ITS*), are amplifiable and gives a clear distinct band pattern and both the locus can be considered as candidate loci in future study of *Benincasa* and *Praecitrullus*.

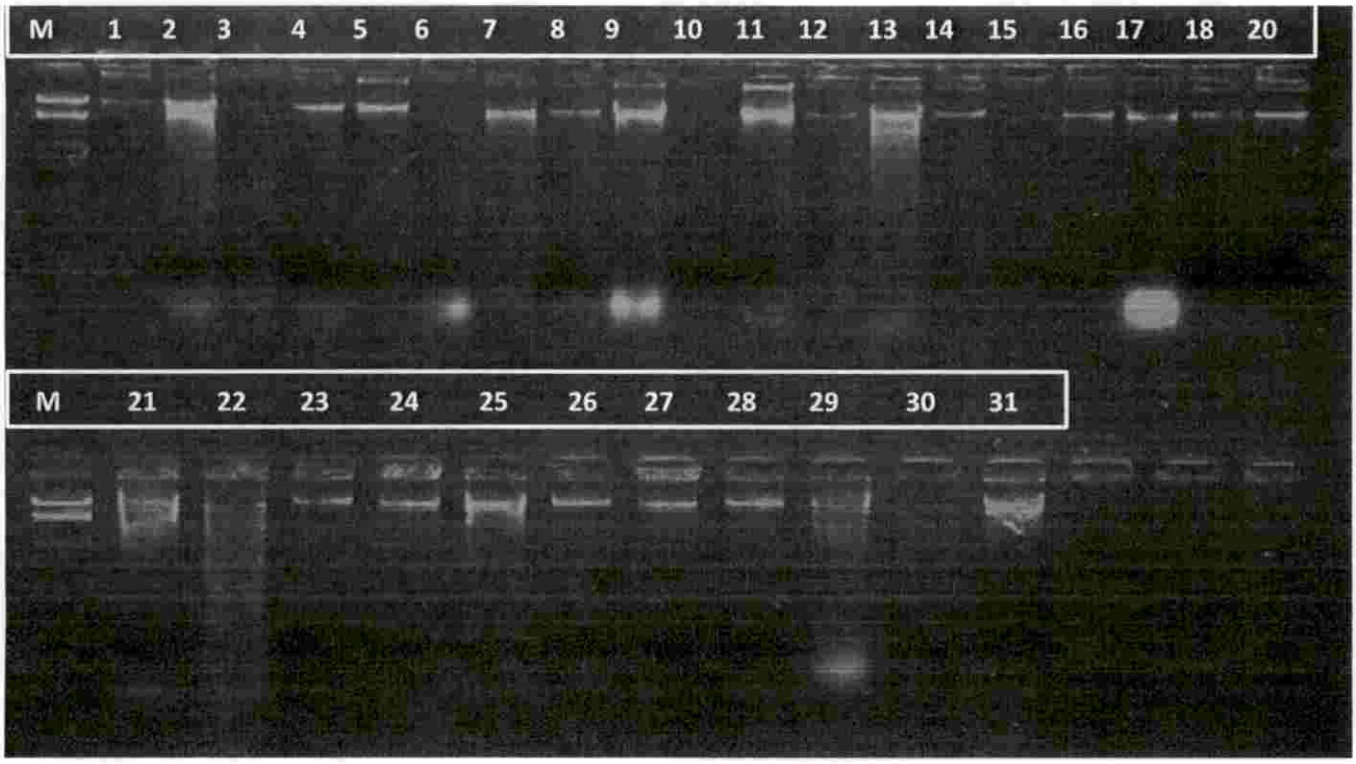


Plate 21: DNA isolated from tender leaves of *Benincasa* and *Praecitrullus* genotypes

M-λ marker, 1-Kottamuri, 2-Vellarikundu, 3-West Eleri, 4-Areeparambu, 5-Kanjirampoyil, 6-Nedungadapalli, 7-Mallapally, 8-North Paravur, 9-Perumbilisherry, 10-Cheramangalam, 11- Wayanad, 12-Kannara, 13-Nashik, 14-No.5, 15-No.9, 16-No.10, 17-No.11, 18- IC 0596986, 19-IC 0596986, 20-IC 0596987, 21-IC 0596989, 22-IC 0596995, 23-IC 0596992, 24-Elavancheri, 25-Thanniam, 26-KAU Local, 27-Ankur tinda, 28-Komal tinda, 29-Nirmal tinda, 30-Mahyco tinda, 31- Kailash tinda

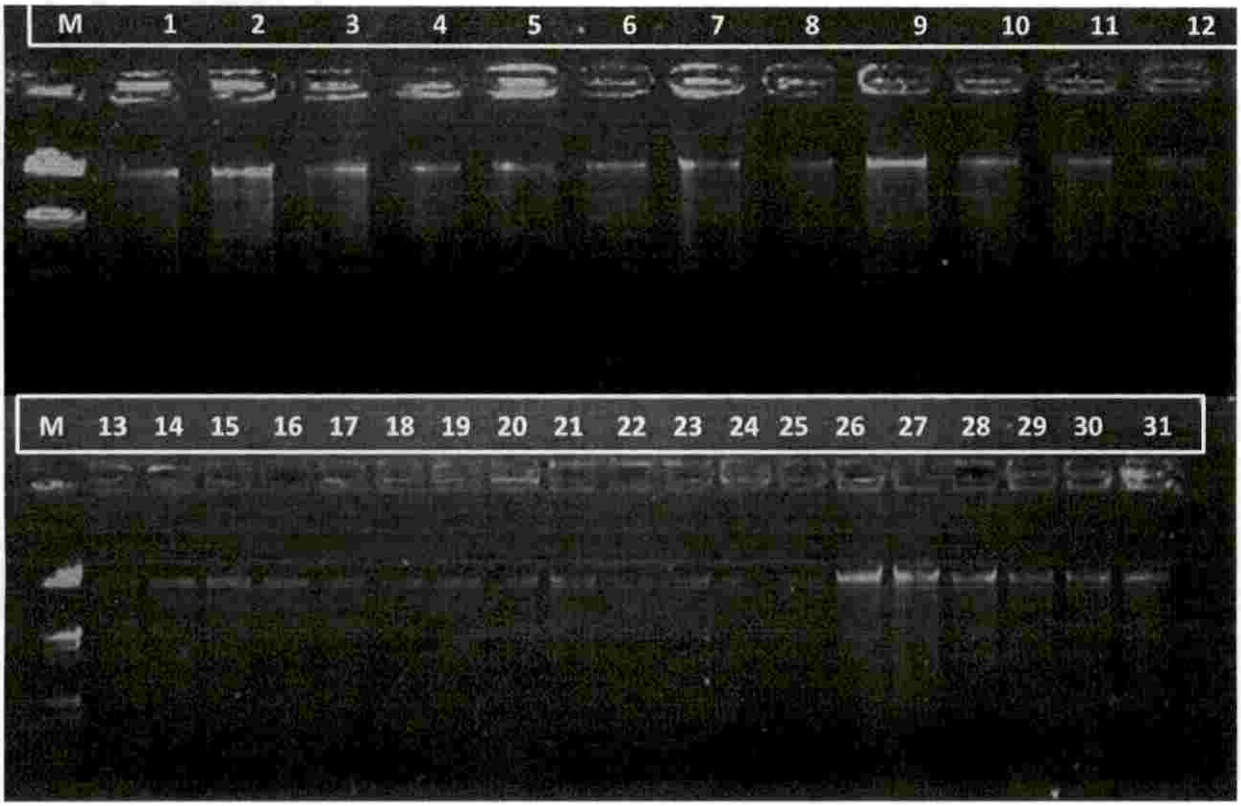


Plate 22: DNA isolated from tender leaves of *Benincasa* and *Praecitrullus* genotypes after RNase treatment

M-λ marker, 1-Kottamuri, 2-Vellarikundu, 3-West Eleri, 4-Areeparambu, 5-Kanjirampoyil, 6-Nedungadapalli, 7-Mallapally, 8-North Paravur, 9-Perumbilisherry, 10-Cheramangalam, 11- Wayanad, 12-Kannara, 13-Nashik, 14-No.5, 15-No.9, 16-No.10, 17-No.11, 18- IC 0596986, 19-IC 0596986, 20-IC 0596987, 21-IC 0596989, 22-IC 0596995, 23-IC 0596992, 24-Elavancheri, 25-Thanniam, 26- KAU Local, 27-Ankur tinda, 28-Komal tinda, 29-Nirmal tinda, 30-Mahyco tinda,31- Kailash tinda

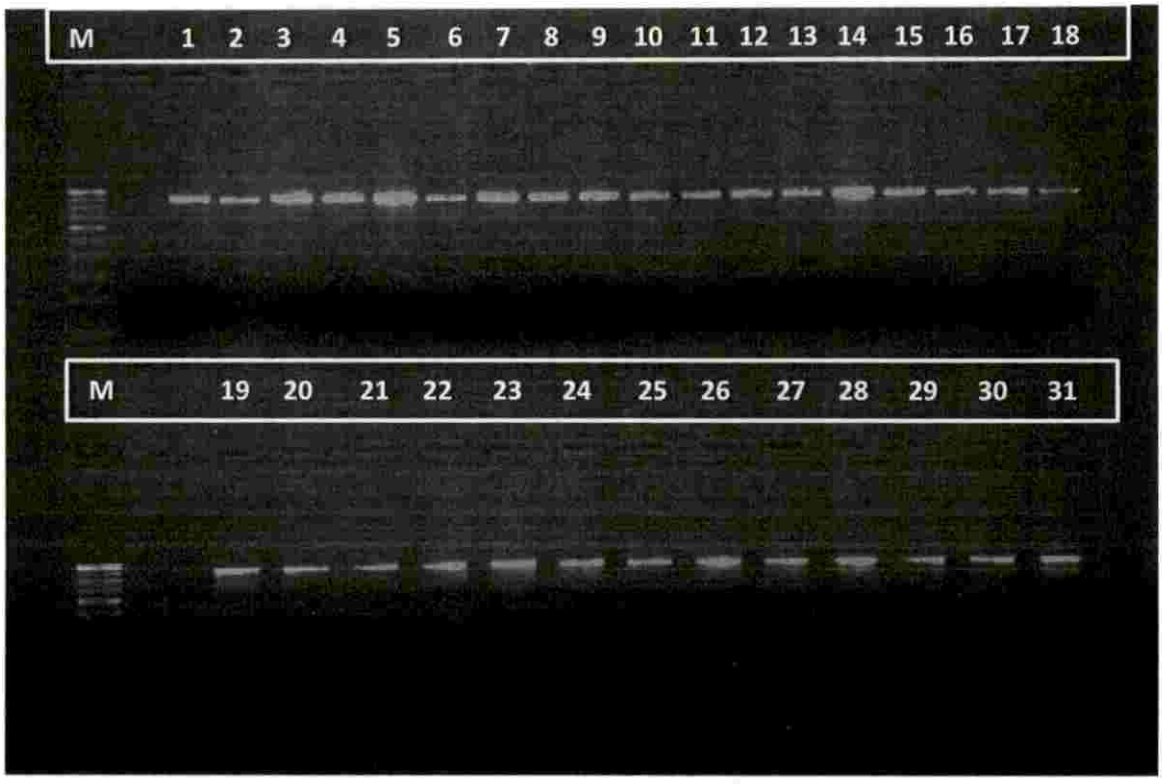


Plate 23: DNA amplification patteren of *Benincasa* and *Praecitrullus* genotypes with barcode primer *matK*

M- *matK*, 1-Kottamuri, 2-Vellarikundu, 3-West Eleri, 4-Areeparambu, 5-Kanjirampoyil, 6-Nedungadapalli, 7-Mallapally, 8-North Paravur, 9-Perumbilisherry, 10-Cheramangalam, 11- Wayanad, 12-Kannara, 13-Nashik, 14-No.5, 15-No.9, 16-No.10, 17-No.11, 18- IC 0596986, 19-IC 0596986, 20-IC 0596987, 21-IC 0596989, 22-IC 0596995, 23-IC 0596992, 24-Elavancheri, 25-Thanniam, 26- KAU Local, 27-Ankur tinda, 28-Komal tinda, 29-Nirmal tinda, 30-Mahyco tinda, 31- Kailash tinda

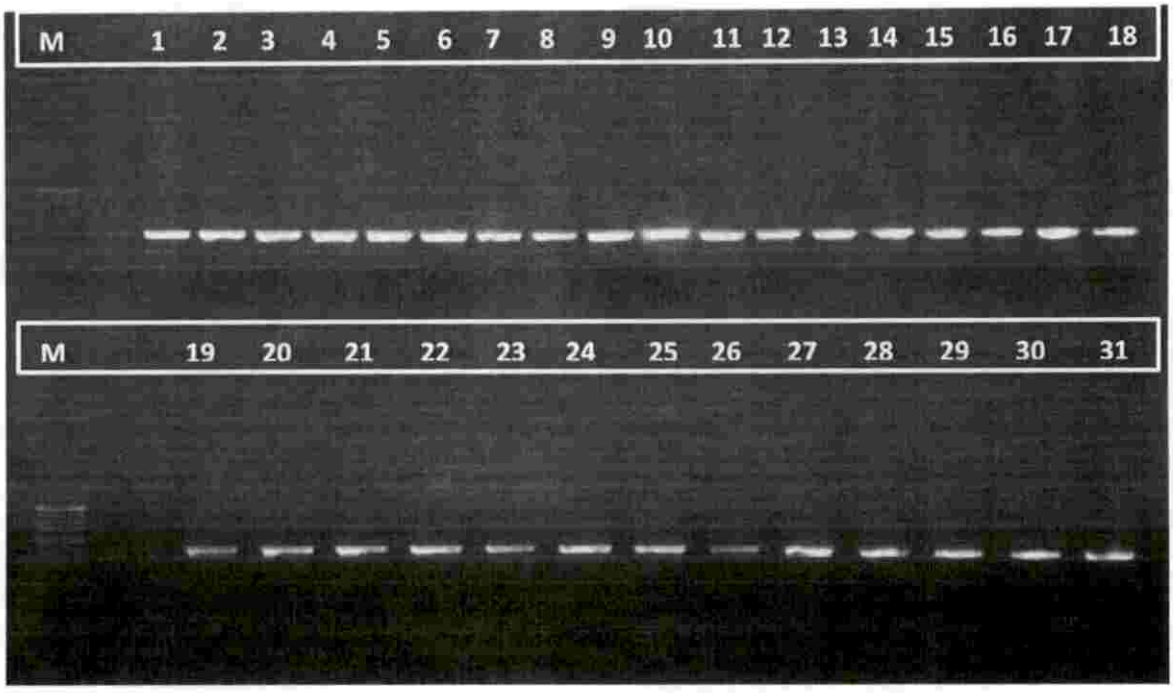


Plate 24: DNA amplification pattern of the *Benincasa* and *Praecitrullus* genotypes with barcode *ITS2*

M- *ITS2*, 1-Kottamuri, 2-Vellarikundu, 3-West Eleri, 4-Areeparambu, 5-Kanjirampoyil, 6-Nedungadapalli, 7-Mallapally, 8-North Paravur, 9-Perumbilisherry, 10-Cheramangalam, 11- Wayanad, 12-Kannara, 13-Nashik, 14-No.5, 15-No.9, 16-No.10, 17-No.11, 18- IC 0596986, 19-IC 0596986, 20-IC 0596987, 21-IC 0596989, 22-IC 0596995, 23-IC 0596992, 24-Elavancheri, 25-Thanniam, 26- KAU Local, 27-Ankur tinda, 28-Komal tinda, 29-Nirmal tinda, 30-Mahyco tinda,31- Kailash tinda

4.8 Sequencing of PCR products

Twelve sequences were obtained after the sequencing by Sanger dideoxy method carried out at AgriGenome Labs. Pvt. Ltd., Cochin. List of *matK* sequences and *ITS2* sequences obtained from each accession.

4.9 Data analysis using In-*silico tools*

4.9.1. Sequence analysis and annotation

On merging of the forward and reverse sequences with CAP3 sequence assembler, proper contigs were formed for 12 accessions. Length of forward, reverse and contig sequences (in bp) obtained for the accessions are presented in Table 4.6.

The sequences generated for *matK* loci

1. Kottamuri

5'TTTATGTGTTTTGGAGAGCATTGTGGAAATTCCATTTTCCCTACGATTAGTA
TCTTCCTTAGAGGAAGAAGAAATCGAAAAATCTTATAATTTACGATCACGTC
ATTCAATATTTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGTCA
GATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCG
CTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACG
AGTATTGTAATTTGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTT
TCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATACAATTCTTATGTATGTG
AATACGAATCCATTTTCCTTTTTCTACGTAATCAATCTTCTCATATACGATTA
ACTTCTTATAGGGACCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAAC
ATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAG
GATCCTTTCATGCATTATGTTCGATATCAAGGAAAATCTATTCTGGTTTCAAA
AGATACGCCACTTCTGATTAATAAGTGGAAATATTACCTTGTCCATTTATGG
CAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAAT
TATCCAAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGATTAAAT
CCTTCAGTGGTATGGAGTCAGATGTTAGAAAATTCATTTCTAATAGATAATG
CTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTG
GCTAAAGCGAAATTTTGTAACGTGTTAGGGCATCCCATTAGTAAGACGACCA
GGATTCGATTTCA'3

2. Vellarikundu

5'TTTTAGAGGGGTTTGGGGATACATGTGAAATTCCATTTTTTCTTACGATTAG
TATCTCTTAGAGAAGAAGAAATCGAAAAATCTTATAATTTACGATCACGTCA
TTCAATATTTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGTCAGA
TGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCTC
CTGGGTGAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTA
TTGTAATTTGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAA
AAGAAATCGAAGATTAGTCTTGTTCCTATACAATTCTTATGTATGTGAATACG
AATCCATTTTCCTTATTCTACGTAATCAATCTTCTCATATACGATTAACTTCTT
ATAGGGACCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGTC
AAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATCCTTTC
ATGCATTATGTTCGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCC
ACTTCTGATTAATAAGTGGAAATATTACCTTGTCCATTTATGGCAATGTCATT
TTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCG
TTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGATTAAATCCTTCAGTGGT
ATGGAGTCAGATGTTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAA
CTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAGCGAA
ATTTTGTAACGTGTTAGGGCATCCCATTAGTAAGACGTGGGAGTCTTCTCCA'3

3. West Eleri

5'TTAGATGTTTTGGGGAGAAACATGTGGAAATTCCATTTTCCCTACGATTAGT
ATCTTCTTAGAGGAAGAAGAAATCGAAAAATCTTATAATTTACGATCACGTC
ATTCAATATTTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGTCAG
ATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCT
CCTGGGTGAAAGATGCCTCTTAGATGTTTTGGGGAGAAACATGTGGAAATTC
CATTTTCCCTACGATTAGTATCTTCTTAGAGGAAGAAGAAATCGAAAAATCTT
ATAATTTACGATCACGTCATTCAATATTTCCTTTTTTAGAGGATAAATTCCCA
CATTTAAATTATGTGTCAGATGTATTAATACCCTATCCCCTCCATCTGGAAAT
TTTAGTTCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATT
ACGGTTCTTTTTTCACGAGTATTGTAATTTGAATAGTCTTAGTACTTCAAAAA
AATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATAC
AATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAATCAATCT
TCTCATATACGATTAACTTCTTATAGGGACCTTTTTGAGCGAATATATTTCTAT
GGAAAAATAGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTT
ACGGGTCTTCAAGGATCCTTTCATGCATTATGTTCGATATCAAGGAAAATCTA
TTCTGGTTTCAAAAGATACGCCACTTCTGATTAATAAGTGGAAATATTACCTT
GTCCATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGATCTA
TATAAACCAATTATCCAAGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGT
GCGATTAAATCCTTCAGTGGTATGGAGTCAGATGTTAGAAAATTCATTTCTAA
TAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCTGCTT
GGATCATTGGCTAAAGCGAAATTTTGTAACGTGTTAGGGCATCCCATTAGTA
AGCGTGGGG'3

4. Areeparambu

5'CTGTGAATGAGTCCAGTCGTCTTACTAATGGGAGAAAGTCGTACCGAAACG
AATTTCGCTTTAGCCATGATCCAAGCAGAGTAATAATAGGAACTAGTGTATC
GAGTTTCTTCGTAGCATTATCTATTAGAAATGAATTTTCTAACATCTGACTCC
ATACCACTGAAGGATTTAATCGCACACTTGAAATATAGCCCAAAAAGTCAAG
AGAACGCTTGGATAATTGGTTTATATAGATCCTTTCTGGTTGTGACCACACAT
AAAAATGACATTGCCATAAATGGACAAGGTAATATTTCCACTTATTAATCAG
AAGTGGCGTATCTTTTGAAACCAGAATAGATTTTCCTTGATATCGAACATAAT
GCATGAAAGGATCCTTGAAGACCCGTAAGATAGCCGAAAAATAATTAGCAA
ACACTTTGACAAGATGTTCTATTTTTCCATAGAAATATATTCGCTCAAAAAGG
TCCCTATAAGAAGTTAATCGTATATGAGAAGATTGATTACGTAGAAAAAGGA
AAATGGATTCGTATTCACATACATAAGAATTGTATAGGAACAAGACTAATCT
TCGATTTCTTTTTGAAAAAAAAGAAATCAATTTTTTTGAAGTACTAAGACTAT
TCAAATTACAATACTCGTGAAAAAAGAACCGTAATAAATGAAAAGAAGAGG
CATCTTTCACCCAGGAGCGAAGGATTTGAACTAAAATTTCCAGATGGAGGGG
ATAGGGTATTAATACATCTGACACATAATTTAAATGTGGGAATTTATCCTCTA
AAAAAGGAAATATTGAATGACGTGATCGTAAATTATAAGATTTTTCGATTTCT

TCTTCCTCTAAGGAAGATACTAATCGTAGGGAAATGGAATTTCCACATGTTCC
CCCAAAACCATCATAGGGGGTTACCTGTGGAACACGATTGCA'3

5. Nashik

GGGGTTGCCGGACCTGATCGATCCTTCTCGTCATCCATCGATCCAAGCAGAGT
AATAATAGGAACTAGTGTATCGAGTTTCTTCGTAGCATTATCTATTAGAAATG
AATTTTCTAACATCTGACTCCATACCACTGAAGGATTTAATCGCACACTTGAA
ATATAGCCCAAAAAGTCAAGAGAACGCTTGGATAATTGGTTTATATAGATCC
TTTCTGGTTGTGACCACACATAAAAATGACATTGCCATAAATGGACAAGGTA
ATATTTCCACTTATTAATCAGAAGTGGCGTATCTTTTGAAACCAGAATAGATT
TTCCTTGATATCGAACATAATGCATGAAAGGATCCTTGAAGACCCGTAAGAT
AGCCGAAAAATAATTAGCAAACACTTTGACAAGATGTTCTATTTTTCCATAGA
AATATATTCGCTCAAAAAGGTCCCTATAAGAAGTTAATCGTATATGAGAAGA
TTGATTACGTAGAAAAGGAAAATGGATTCGTATTCACATACATAAGAATTGT
ATAGGAACAAGACTAATCTTCGATTTCTTTTTGAAAAAAAAGAAGGGGTTGC
CGGACCTGATCGATCCTTCTCGTCATCCATCGATCCAAGCAGAGTAATAATAG
GAACTAGTGTATCGAGTTTCTTCGTAGCATTATCTATTAGAAATGAATTTTCT
AACATCTGACTCCATACCACTGAAGGATTTAATCGCACACTTGAAATATAGC
CCAAAAAGTCAAGAGAACGCTTGGATAATTGGTTTATATAGATCCTTTCTGGT
TGTGACCACACATAAAAATGACATTGCCATAAATGGACAAGGTAATATTTCC
ACTTATTAATCAGAAGTGGCGTATCTTTTGAAACCAGAATAGATTTTCCTTGA
TATCGAACATAATGCATGAAAGGATCCTTGAAGACCCGTAAGATAGCCGAAA
AATAATTAGCAAACACTTTGACAAGATGTTCTATTTTTCCATAGAAATATATT
CGCTCAAAAAGGTCCCTATAAGAAGTTAATCGTATATGAGAAGATTGATTAC
GTAGAAAAGGAAAATGGATTCGTATTCACATACATAAGAATTGTATAGGAAC
AAGACTAATCTTCGATTTCTTTTTGAAAAAAAAGAAATCAATTTTTTTGAAGT
ACTAAGACTATTCAAATTACAATACTCGTGAAAAAAGAACCGTAATAAATGA
AAAGAAGAGGCATCTTTCACCCAGGAGCGAAGGATTTGAACTAAAATTTCCA
GATGGAGGGGATAGGGTATTAATACATCTGACACATAATTAAAATGTGGGAA
TTTATCCTCTAAAAAAGGAAATATTGAATGACGTGATCGTAAATTATAAGATT
TTTCGATTTCTTCTTCCTCTAAGGAAGATACTAATCGTAGTGAAAATGGAAAT
TTCCACATGCCCCCCAACCCTAAAACGGCTGCGGACTTATGTATGAACACTC'
3

6. IC 0596989

5'ATAGTGTGTGGTTGGGGGAGTATGTGAATTCATTTCCCTACGATTAGTATCT
CTTAGAGAGAGAAATCGAAAAATCTATAATTTACGATCACGTCATTCAATAT
TTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGTCAGATGTATTA
ATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCTCCTGGGT
GAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAA
TTTGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAA
ATCGAAGATTAGTCTTGTTCCTATACAATTCTTATGTATGTGAATACGAATCC
ATTTTCCTTTTTCTACGTAATCAATCTTCTCATATACGATTAACTTCTTATAGG

GACCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGTCAAAGT
GTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATCCTTTCATGCA
TTATGTTCGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTC
TGATTAATAAGTGGAAATATTACCTTGTCCATTTATGGCAATGTCATTTTTAT
GTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTC
TTGACTTTTTGGGCTATATTTCAAGTGTGCGATTAAATCCTTCAGTGGTATGG
AGTCAGATGTTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTCG
ATACACTAGTTCCTATTATTACTCTGCTTGGATCATGGCTAAAGCGAAATTTT
GTAACGTGTTACGGCATCCCATTAGTAAGACGACTGAAGATACCGAATACAG
ACGCCTGAT'3

7. IC 0596992

5'TTTTTATGTGTGTTGGGGGATACATGTGGAAATTCCATTTTCCCTACGATTA
GTATCTCCTTAGAGAAGAAGAAATCGAAAAATCTATAATTTACGATCACGTC
ATTCAATATTTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGTCAG
ATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCT
CCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGT
ATTGTAATTTGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAA
AAAGAAATCGAAGATTAGTCTTGTTCCTATACAATTCTTATGTATGTGAATAC
GAATCCATTTTCCTTTTTCTACGTAATCAATCTTCTCATATACGATTAACTTCT
TATAGGGACCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGT
CAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATCCTTT
CATGCATTATGTTCGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGC
CACTTCTGATTAATAAGTGGAAATATTACCTTGTCCATTTATGGCAATGTCAT
TTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGC
GTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGATTAAATCCTTCAGTGG
TATGGAGTCAGATGTTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAA
ACTCGATACACTAGTTCCTATTATTACTCTGCTTGATCATTGCTAAAGCGAAA
TTTTGTAACGTGTTAGGGCATCCCA'3

8. Elavancheri

5'ACCAGGTGATTGATCAAATTTGTCCTTCTAGGGGTGCCGTAACACGTGCAA
ATTCGCCTTGCCAATGATCCAAGCAGAGTAATAATAGGAACTAGTGTATCGA
GTTTCTTCGTAGCATTATCTATTAGAAATGAATTTTCTAACATCTGACTCCATA
CCACTGAAGGATTTAATCGCACACTTGAAATATAGCCCAAAAAGTCAAGAGA
ACGCTTGGATAATTGGTTTATATAGATCCTTTCTGGTTGTGACCACACATAAA
AATGACATTGCCATAAATGGACAAGGTAATATTTCCACTTATTAATCAGAAG
TGGCGTATCTTTTGAAACCAGAATAGATTTTCCTTGATATCGAACATAATGCA
TGAAAGGATCCTTGAAGACCCGTAAGATAGCCGAAAAATAATTAGCAAACAC
TTTGACAAGATGTTCTATTTTTCCATAGAAATATATTCGCTCAAAAAGGTCCC
TATAAGAAGTTAATCGTATATGAGAAGATTGATTACGTAGAAAAAGGAAAAT
GGATTCGTATTCACATACATAAGAATTGTATAGGAACAAGACTAATCTTCGA
TTTCTTTTTGAAAAAAAAGAAATCAATTTTTTTGAAGTACTAAGACTATTCAA

ATTACAATACTCGTGAAAAAAGAACCGTAATAAATGAAAAGAAGAGGCATC
TTTCACCCAGGAGCGAAGGATTTGAACTAAAATTTCCAGATGGAGGGGATAG
GGTATTAATACATCTGACACATAATTTAAATGTGGGAATTTATCCTCTAAAAA
AGGAAATATTGAATGACGTGATCGTAAATTATAAGATTTTTCGATTTCTTCTT
CCTCTAAGGAAGATACTAATCGTAGGGAAAATGGAATTTCCACAATGACTCC
CAAACCACTCAA'3

9. KAU local

5'AGTATGTGAAATTCATTTCCCTACGATTAGTATCTTCTTAGAGAAGAAGAA
ATCGAAAAATCTTATAATTTACGATCACGTCATTCAATATTTCCTTTTTTAGA
GGATAAATTCCCACATTTAAATTATGTGTCAGATGTATTAATACCCTATCCCC
TCCATCTGGAAATTTTAGTTCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCT
TCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGTAATTTGAATAGTCTT
AGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGT
CTTGTTCCTATACAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCT
ACGTAATCAATCTTCTCATATACGATTAACTTCTTATAGGGACCTTTTTGAGC
GAATATATTTCTATGGAAAAATAGAACATCTTGTCAAAGTGTTTGCTAATTAT
TTTTCGGCTATCTTACGGGTCTTCAAGGATCCTTTCATGCATTATGTTCGATAT
CAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACTTCTGATTAATAAGTG
GAAATATTACCTTGTCCATTTATGGCAATGTCATTTTTATGTGTGGTCACAAC
CAGAAAGGATCTATATAAACCAATTATCCAAGCGTTCTCTTGACTTTTTGGGC
TATATTTCAAGTGTGCGATTAAATCCTTCAGTGGTATGGAGTCAGATGTTAGA
AAATTCATTTCTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTA
TTATTACTCTGCTTGGATCATTGGCTAAAGCGAAATTTTGTAACGTGTTAGGG
CATCCCATTAGTAAGCGTGTAAGTGATAGCATATCAAAAA

10. Nirmal Tinda

ATAGGGTGGGGGGATATGTGGAAATTCCATTTTCCCTACGATTAGTATCTTCC
TTAGAGGAAGAAGAAATCGCAAAATCTTATAATTTACGATCACGTCATTCAA
TATTTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGTCAGATGTAT
TAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTTCGCTCCTGG
GTGAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCACGAGTATTGG
AATTTGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTTTCAAAAAG
AAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTGAATACGAAT
CCATTTTCCTTTTTCTACGTAATCAATCTTCTCATATACGATTAACTTCTTATA
GGGACCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATCTTGTCAAA
GTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATCCTTTCATG
CATTATGTTCGATATCAAGGAAAATCTATTCTGGTTTCAAAAGATACGCCACT
TCTGATTAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATGTCATTTTT
ATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCAAGCGTTC
TCTTGACTTTTTGGGCTATATTTCAAGTGTGCGATTAAATCCTTCAGTGGTATG
GAGTCAGATGTTAGAAAATTCATTTCTAATAGATAATGCTACGAAGAAACTC
GATACACTAGTTCCTATTATTACTCTGCTTGGATCATGGCTAACGCGAAATTT

TGTAACGTGTTAGGGCAGCACACTAGTAAGACGTCTGAATCCGACTCACCGG
ATGTCCTT'3

11. Komal tinda

5'TAGATATGTGGTTTTTGAGAGTACTGTGAAATTCCATTTTCCCTACGATTAG
TATCTTCCTTAGAGGAAGAAGAAATCGCAAAATCTTATAATTATACGATCAC
GTCATTCAATATTTCCTTTTTTAGAGGATAAATTCCCACATTTAAATTATGTGT
CAGATGTATTAATACCCTATCCCCTCCATCTGGAAATTTTAGTTCAAATCCTT
CGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTTATTACGGTTCTTTTTTCAC
GAGTATTGGAATTTGAATAGTCTTAGTACTTCAAAAAAATTGATTTCTTTTTTT
TCAAAAAGAAATCGAAGATTAGTCTTGTTCCTATATAATTCTTATGTATGTGA
ATACGAATCCATTTTCCTTTTTCTACGTAATCAATCTTCTCATATACGATTAAC
TTCTTATAGGGACCTTTTTGAGCGAATATATTTCTATGGAAAAATAGAACATC
TTGTCAAAGTGTTTGCTAATTATTTTTCGGCTATCTTACGGGTCTTCAAGGATC
CTTTCATGCATTATGTTCGATATCAAGGAAAATCTATTCTGGTTTCAAAAGAT
ACGCCACTTCTGATTAATAAGTGGAAATATTACCTTGTCAATTTATGGCAATG
TCATTTTTATGTGTGGTCACAACCAGAAAGGATCTATATAAACCAATTATCCA
AGCGTTCTCTTGACTTTTTGGGCTATATTTCAAGTGTGCGATTAAATCCTTCAG
TGGTATGGAGTCAGATGTTAGAAAATTCATTTCTAATAGATAATGCTACGAA
GAAACTCGATACACTAGTTCCTATTATTACTCTGCTTGGATCATTGGCTAAAG
CGAAATTTTGTAACGTGTTAGGGCATCCCATTAGTAGCGTGAG'3

12. Ankur tinda

5'GCTTCTAGTCAACTCCCTGATATATATACAAGACCTCTTGATGTATCAGACG
CTTGACCTCTACAAGTCGCCGGATCTTTTGATGTGTGTGGGGGATCATGTGGA
AATTCCATTTTCCCTACGATTAGTATCTCCTTAGAGAAGAAGAAATCGCAAAA
TCTTATAATTTACGATCACGTCATTCAATATTTCCTTTTTTAGAGGATAAATTC
CCACATTTAAATTATGTGTCAGATGTATTAATACCCTATCCCCTCCATCTGGA
AATTTTAGTTCAAATCCTTCGCTCCTGGGTGAAAGATGCCTCTTCTTTTCATTT
ATTACGGTTCTTTTTTCACGAGTATTGGAATTTGAATAGTCTTAGTACTTCAA
AAAAATTGATTTCTTTTTTTTCAAAAAGAAATCGAAGATTAGTCTTGTTCCTA
TATAATTCTTATGTATGTGAATACGAATCCATTTTCCTTTTTCTACGTAATCAA
TCTTCTCATATACGATTAACTTCTTATAGGGACCTTTTTGAGCGAATATATTTC
TATGGAAAAATAGAACATCTTGTCAAAGTGTTTGCTAATTATTTTTCGGCTAT
CTTACGGGTCTTCAAGGATCCTTTCATGCATTATGTTCGATATCAAGGAAAAT
CTATTCTGGTTTCAAAAGATACGCCACTTCTGATTAATAAGTGGAAATATTAC
CTTGTCAATTTATGGCAATGTCATTTTTATGTGTGGTCACAACCAGAAAGGAT
CTATATAAACCAATTATCCAAGCGTTCTCTTGACTTTTTGGGCTATATTTCAA
GTGTGCGATTAAATCCTTCAGTGGTATGGAGTCAGATGTTAGAAAATTCATTT
CTAATAGATAATGCTACGAAGAAACTCGATACACTAGTTCCTATTATTACTCT
GCTTGATCATTGGCTAAAGCGAAATTTTGTAACGTGTTAGGGCATCCCATTAG
TAAGACGCGTG'3

The sequences for *ITS2* loci

1. Kottamuri

5'CCATATGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTG
CGCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCT
GCCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCG
TGCGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCG
CGACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCC
GCCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGG
CGCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCAT
ATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCCTAGTAACGGCG
AGCGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAAT
TGTAGTCTGAACTAAACCGTCA'3

2. Vellarikundu

5'TTTCTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTT
GCGCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGC
TGCCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCC
GTGCGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTC
GCGACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCC
CGCCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATG
GCGCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCA
TATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGA
GCGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATT
GTAGTTA'3

3. West Eleri

5'TTTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTGC
GCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCTG
CCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGT
GCGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGC
GACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCCG
CCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGGC
GCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATA
TCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAG
CGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATTG
TAGTCTAA'3

4. Areeparambu

5'CTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTGCG
CCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCTGC
CCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTG
CGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCG

ACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCCGC
CTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGGCG
CTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATAT
CAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAGC
GAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATTGT
AGTCAAAA'3

5. Nashik

5'TTTGTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTT
GCGCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGC
TGCCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCC
GTGCGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTC
GCGACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCC
CGCCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATG
GCGCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCA
TATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGA
GCGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATT
GTAGTCTGAAAAAGCGTCA'3

6. IC 0596989

5'TTCTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTG
CGCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCT
GCCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCG
TGCGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCG
CGACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCC
GCCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGG
CGCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCAT
ATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGA
GCGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATT
GTAGTCTAAAAAAACCGTCA'3

7. IC 0596992

5'GGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTGCGC
CCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCTGCC
CCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGC
GCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGA
CACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCCGCCT
CCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGGCGCT
CTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATATCA
ATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAGCGA
ACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATTGTAG
TCTAA'3

8. Elavancheri

5'TTCGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTGC
GCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCTG
CCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGT
GCGCATCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGC
GACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCCG
CCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGGC
GCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATA
TCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAG
CGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATTG
TAGTCAAA'3

9. KAU local

5'TTCTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTG
CGCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCT
GCCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCG
TGCGCACCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCG
CGACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCC
GCCTCCACGGACTCATGCATTGACCCTCTGAGCGTTGTACCCGAAAGGATGG
CGCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCAT
ATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGA
GCGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATT
GTAGTCAAAATAAACCGTC'3

10. Nirmal Tinda

5'TTTTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTGCG
CCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCTGC
CCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTG
CGCATCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGTGA
CACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCCGCCT
CCACGGACTCATGCACCGACCCTCCGAGCGTTGTACCCGAAAGGATGGCGCT
CTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATATCA
ATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAGCGA
ACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATTGTAG
TTA'3

11. Komal Tinda

5'CTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACGCAAGTTGCG
CCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACGCATCGCTGC
CCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTG
CGCATCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGTGA
CACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCGACCCCGCCT

CCACGGACTCATGCACCGACCCTCCGAGCGTTGTACCCGAAAGGATGGCGCT
CTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATATCA
ATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAGCGA
ACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCCGAATTGTAG
TCTGAATAAAGCGTC'3

12. Ankur Tinda

5'GCGATTTTTTTTGTGAATTGCAGGATCCCGCGAACCACCGAGTCTTTGAACG
CAAGTTGCGCCCGGAGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCACG
CATCGCTGCCCCCCTCCACACAACTCGTTGTGCAGGCGGGGGCACATGTTGG
CCTCCCGTGCGCATCGTCGTGCGGATGGCTTAAATTCGAGTCCTCGGCGCACG
TCGTCGTGACACTACGGTGGTTGATCCAACCTCAGTACCATGTCGCGGCCTCG
ACCCCGCCTCCACGGACTCATGCACCGACCCTCCGAGCGTTGTACCCGAAAG
GATGGCGCTCTCGACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTA
AGCATATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACG
GCGAGCGAACCGGGAAGAGCCCAGCTTGAGAATCGGGCGTCCTCGACGTCC
GAATTGTAGTCTAAAAAAAACCGTC'3

Table 4.3 Sequence length of forward, reverse and contig sequences

S.I. No	Accession	Species and locality	Loci	Length of sequence obtained after sequencing(bp)		Length of contig formed (bp)
				F	R	
1	Kottamuri	*Benincasa hispida*, Kerala	*matK*	877	874	906
2	Vellarikundu	*Benincasa hispida*, Kerala	*matK*	870	875	903
3	West Eleri	*Benincasa hispida*, Kerala	*matK*	855	873	893
4	Areeparambu	*Benincasa hispida*, Kerala	*matK*	870	900	931
5	Nashik	*Benincasa hispida*, Maharashtra	*matK*	900	900	900
6	IC 0596989	*Benincasa hispida*, North-east	*matK*	882	866	908
7	IC 0596992	*Benincasa hispida*, North-east	*matK*	854	872	876
8	Elavancheri	*Benincasa hispida*, Kerala	*matK*	875	872	902
9	KAU local	*Benincasa hispida*, Kerala	*matK*	875	852	889
10	Ankur tinda	*Praecitrullus fistulosus*, Maharshtra	*matK*	881	868	910
11	Komal tinda	*Praecitrullus fistulosus*, Maharshtra	*matK*	857	875	894
12	Kalash tinda	*Praecitrullus fistulosus*, Maharshtra	*matK*	857	955	968
13	Kottamuri	*Benincasa hispida*, Kerala	*ITS2*	463	459	489

14	Vellarikundu	*Benincasa hispida*, Kerala	*ITS2*	448	453	474
15	West Eleri	*Benincasa hispida*, Kerala	*ITS2*	449	453	474
16	Areeparambu	*Benincasa hispida*, Kerala	*ITS2*	452	449	474
17	Nashik	*Benincasa hispida*, Maharastra	*ITS2*	460	456	486
18	IC 0596989	*Benincasa hispida*, North-east	*ITS2*	461	452	486
19	IC 0596992	*Benincasa hispida*, North-east	*ITS2*	451	455	472
20	Elavancheri	*Benincasa hispida*, Kerala	*ITS2*	450	456	474
21	KAU local	*Benincasa hispida*, Kerala	*ITS2*	461	453	485
22	Ankur tinda	*Praecitrullus fistulosus*, Maharshtra	*ITS2*	452	454	471
23	Komal tinda	*Praecitrullus fistulosus*, Maharshtra	*ITS2*	460	450	483
24	Kalash tinda	*Praecitrullus fistulosus*, Maharshtra	*ITS2*	463	461	494

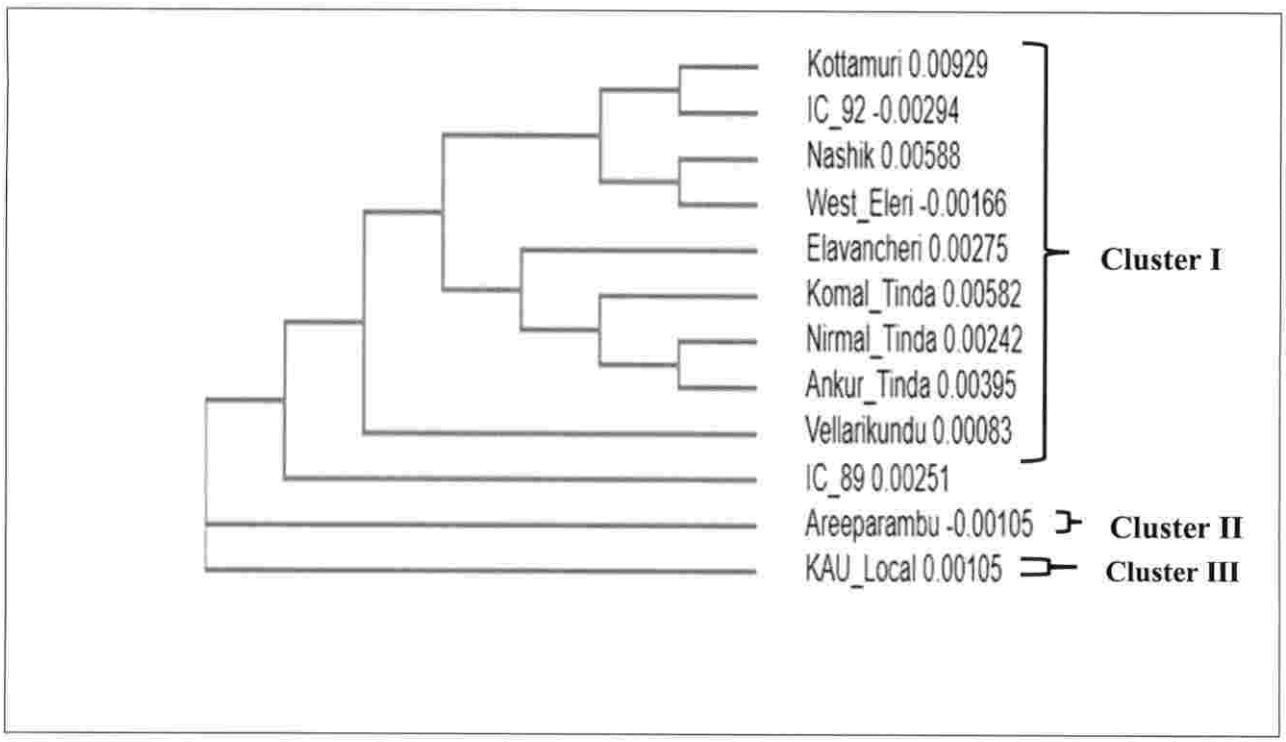


Fig. 4.1.2 The phylogenetic tree of 9 *Benincasa hispida* and 3 *Praecitrullus fistulosus* accessions based on dissimilarity coefficient using the *ITS2* sequence, as generated using ClustalW omega

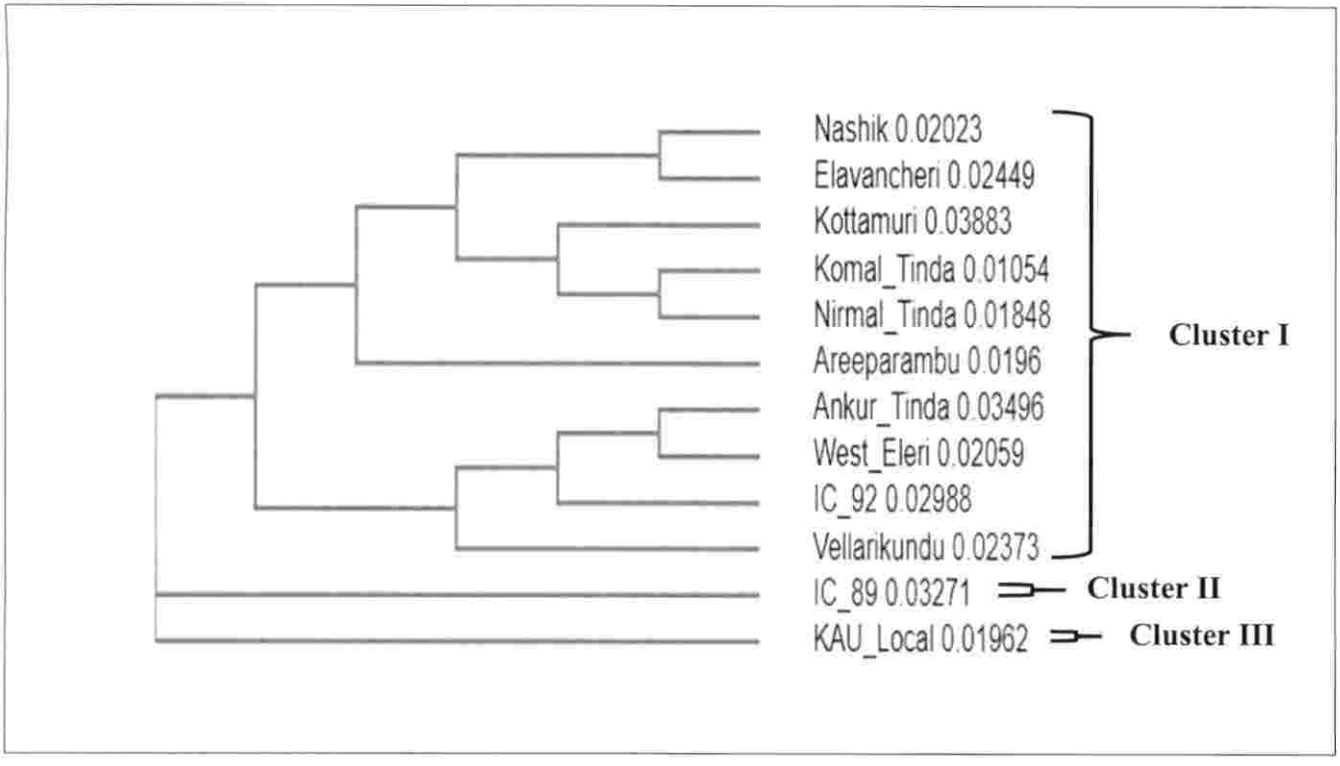


Fig.4.1.3 The phylogenetic tree of 9 *Benincasa hispida* and 3 *Praecitrullus fistulosus* accessions based on dissimilarity coefficient using the *matk* sequence, as generated using ClustalW omega

4.10 Cluster analysis

The genetic dissimilarity was determined for the genotypes based on *matK* and *ITS2* sequence. The phenogram generated showed III clusters.

For *ITS2* loci, three clusters were formed, cluster I is divided into two subclusters. Subcluster Ia represents the accessions Kottamuri, IC 0596992, Nashik, West Eleri, Elavancheri and tinda accessions Komal tinda, Ankur tinda and Nirmal tinda. Subcluster Ib consists of only accession of ashgourd Vellarikundu. Subcluster IIIc represents only an accession of ashgourd IC 0596989, forming a distinct clade from all those accessions. Cluster II is represented by only a single ashgourd accession Areeparambu and Cluster III is formed with consisted of only KAU Local. (Fig.4.1.2)

For, *matk* loci, three clusters were formed with cluster I divided into three subclusters. Subcluster Ia represented the accession Nashik, Elavancheri, Kottamuri and tinda accessions Komal tinda and Nirmal tinda. Subcluster IIb represented a single accession of ashgourd Areeparambu. Subcluster IIIc had a single accession Ankur tinda which shares similarity with ashgourd genotype West Eleri. The ashgourd genotypes IC 0596992 forms a single clade and genotype Vellarikundu forme a single and distinct clade.

Cluster I consists of only a single ashgourd genotype IC 0596989. Cluster III also had a single ashgourd genotype KAU Local. (Fig.4.1.3)

Although, on morphological basis both the species are different, the *matk* loci and *ITS2* loci represents genetic relatedness among the *Benincasa* and *Praecitrullus*.

4.11 Phylogeny based on phylogenetic tree

The phylogenetic tree analysis carried out using MEGA 7 software with the neighbor joining method. The phylogenetic tree formed for both the loci *matK* and *ITS*2 loci formed two different clusters.

For *matk* loci, three clades were formed. Clade I consisted of accessions Kottamuri, West Eleri, IC 0596992 and tinda accessions Komal tinda and Ankur tinda it represents similarity among them. In clade II, had Nashik, Elavancheri, IC 0596989 and a

single accession of tinda Nirmal tinda, sharing similarity in between them. Clade III represents three accessions of ashgourd KAU Local and Areeparambu share similarity and Vellarikundu is distinctly related. All the above accession has a common node with a bootstrap value of 100, indicating common ancestory Fig.4.1.4.

For *ITS2* loci, two clades are formed, one clade consists of ashgourd accessions West Eleri, Elavancheri closely related to tinda accession Ankur tinda. The accession IC 0596992 is closely related to IC 0596989 and KAU Local. The tinda accessions Komal tinda and Nirmal tinda observed to be closely related. The accessions Vellarikundu and Areeparambu shares similarity with Nashik accession. The second clade formed with only a single accession of ashgourd named Kottamuri. All the above accessions have common node with a bootstrap value of 100, indicating common ancestory Fig.4.1.5.

The findings suggest that *Benincasa hispida* and *Praecitrullus fistulosus* exhibit close genetic similarity.

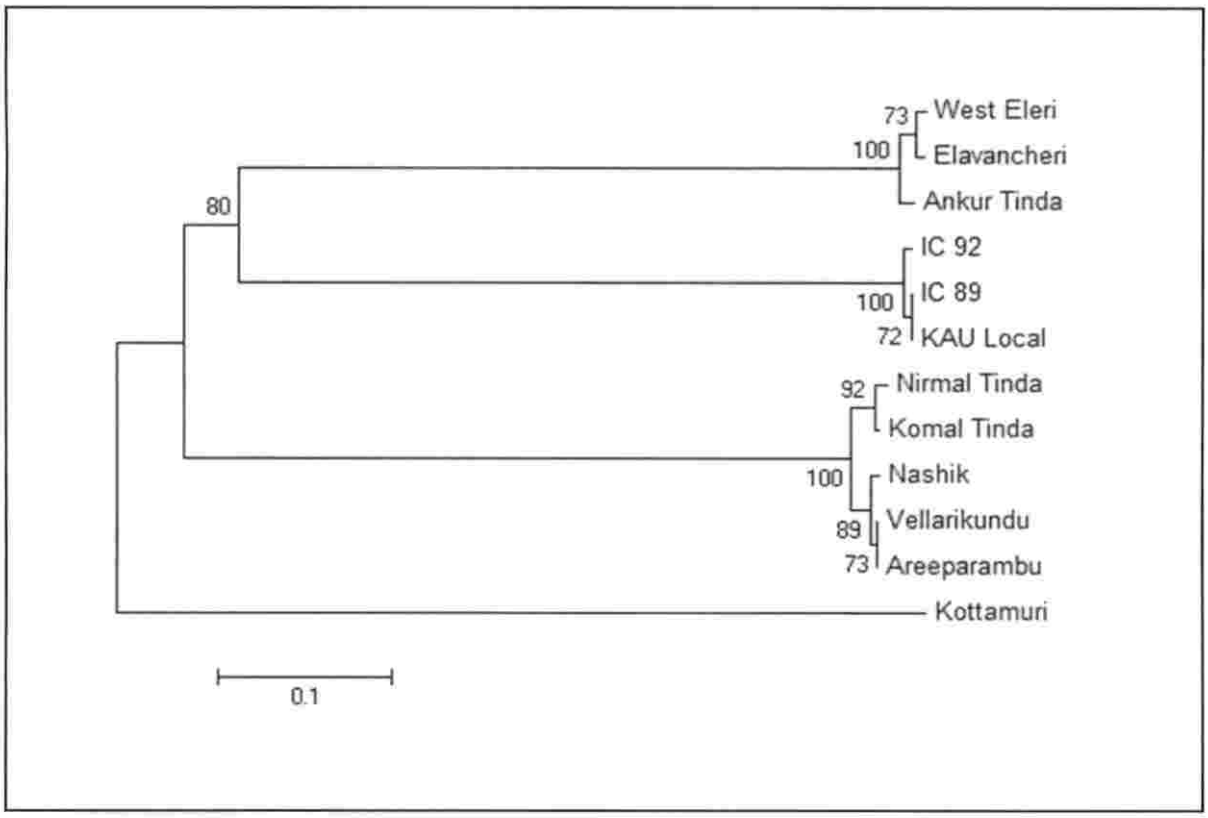


Fig.4.1.4 The phylogenetic tree of 9 *Benincasa hispida* and 3 *Praecitrullus fistulosus* accessions based on dissimilarity coefficient using the *matK* sequence, as generated by MEGA software

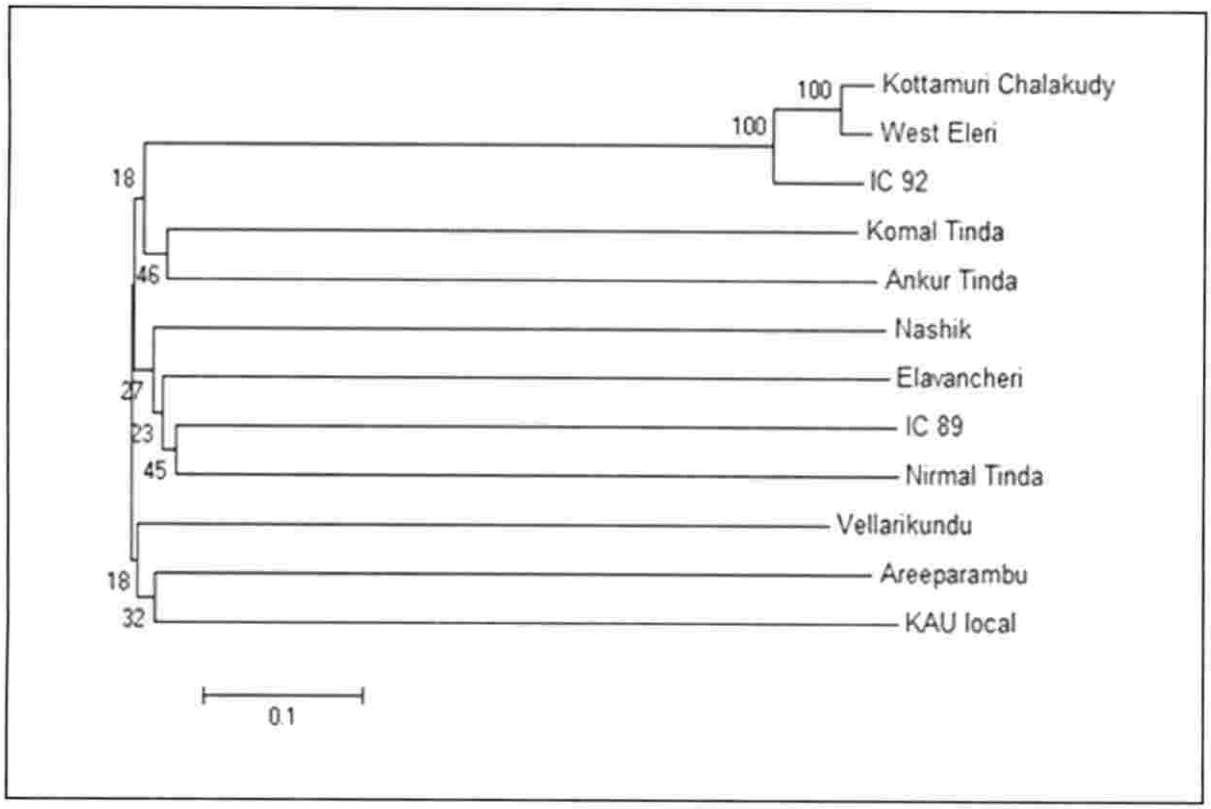


Fig.4.1.5 The phylogenetic tree of 9 *Benincasa hispida* and 3 *Praecitrullus fistulosus* accessions based on dissimilarity coefficient using the *matK* sequence, as generated by MEGA software

Table 4.4 The distance among accessions based on *matK sequence*

	1	2	3	4	5	6	7	8	9	10	11	12
1. Kottamuri Chalakudy												
2. Vellarikundu	0.899											
3. West Eleri	0.039	0.888										
4. Areeparambu	0.961	0.894	0.961									
5. Nashik	0.950	0.911	0.955	0.922								
6. IC 89	0.961	0.933	0.955	0.961	0.916							
7. IC 92	0.117	0.894	0.117	0.955	0.939	0.950						
8. Elavancheri	0.922	0.939	0.922	0.927	0.933	0.927	0.933					
9. KAU local	0.933	0.922	0.939	0.922	0.978	0.961	0.933	0.955				
10. Nirmal Tinda	0.944	0.922	0.939	0.939	0.933	0.911	0.939	0.916	0.983			
11. Komal Tinda	0.922	0.883	0.922	0.955	0.939	0.927	0.911	0.899	0.939	0.944		
12. Ankur Tinda	0.916	0.933	0.916	0.933	0.933	0.933	0.899	0.966	0.927	0.966	0.883	

Table. 4.5 The distance among accessions based on *ITS2* sequence

	1	2	3	4	5	6	7	8	9	10	11	12
1. Kottamuri												
2. Vellarikundu	0.901											
3. West Eleri	0.919	0.829										
4. Areeparambu	0.901	0.000	0.829									
5. Nashik	0.901	0.009	0.820	0.009								
6. IC 89	0.928	0.811	0.775	0.811	0.811							
7. IC 92	0.928	0.811	0.775	0.811	0.811	0.009						
8. Elavancheri	0.919	0.829	0.009	0.829	0.820	0.775	0.775					
9. KAU Local	0.928	0.811	0.775	0.811	0.811	0.000	0.009	0.775				
10. Nirmal Tinda	0.901	0.036	0.829	0.036	0.036	0.820	0.820	0.829	0.820			
11. Komal Tinda	0.901	0.027	0.829	0.027	0.036	0.820	0.820	0.829	0.820	0.009		
12. Ankur Tinda	0.910	0.820	0.027	0.820	0.811	0.775	0.775	0.018	0.775	0.820	0.820	

4.13 Genetic divergence within and between the species

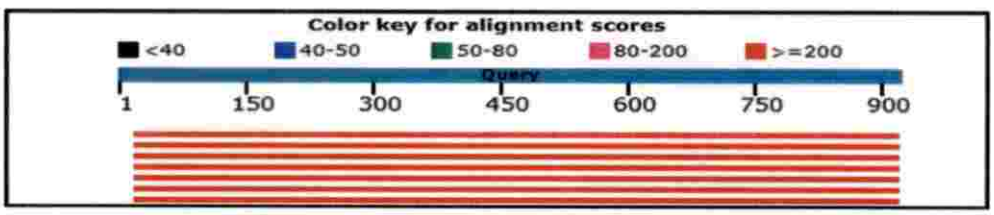
The inter specific distances were calculated using K2P (Kimura 2 parameter) model as recommended by the Consortium of Barcode of Life (CBOL, http://www.barcoding.si.edu./protocols.html) using MEGA 7 software.

For *matk* loci, the range of pairwise dissimilarity was between 0.899 and 0.833. The minimum pairwise distance among ashgourd accession is observed in West Eleri at 0. 039.Table 4.4.

For *ITS2* loci, the range of pairwise dissimilarity was between 0.901 and 0.820. The maximum similarity is observed in the range of 0.928 in ashgourd accessions. The minimum dissimilarity is observed at 0.901 among the ashgourd accession and tinda accessions. Table.4.5.

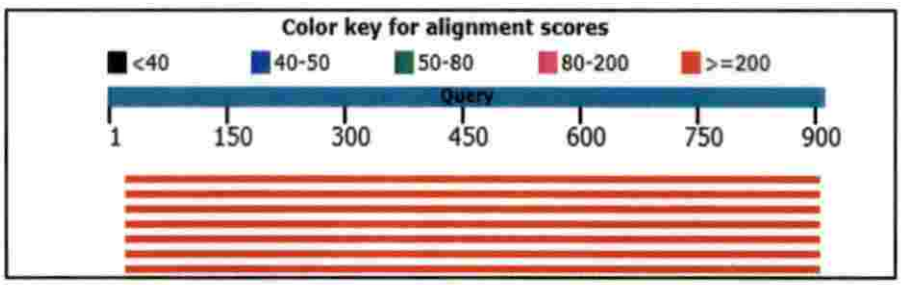
BLAST for *matK* sequences

Fig.4.1.6 BLASTn expression for Kottamuri accession



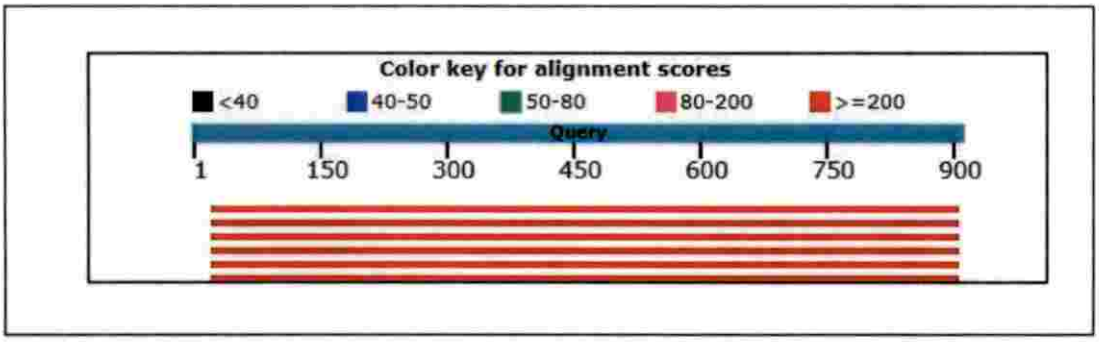
Description	Max score	Total score	Query cover	E value	Ident	Accession
[redacted]	1483	1483	97%	0.0	97%	[redacted]

Fig.4.1.7 BLASTn expression for Vellarikundu accession



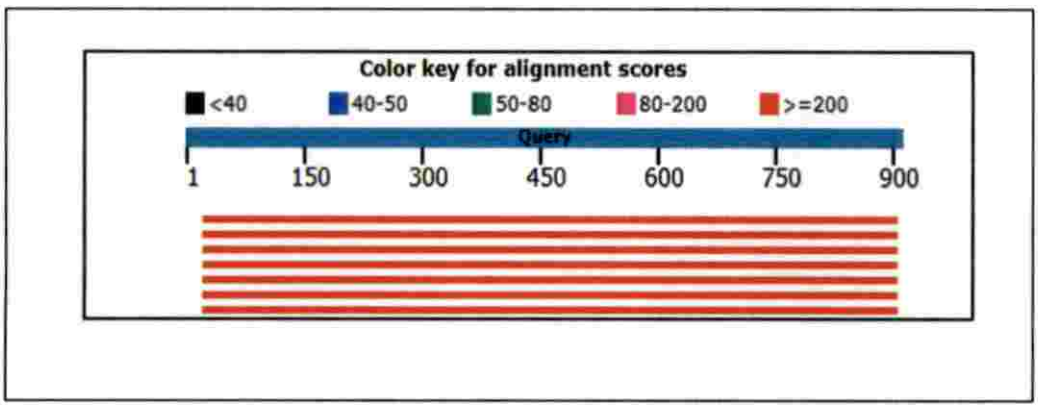
Description	Max score	Total score	Query cover	E value	Ident	Accession
Beninca forpida voucher S. S. Renner et al. 2790 (M). cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1465	1465	96%	0.0	97%	DQ536836.1
Beninca fotiulosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds; chloroplast	1452	1452	96%	0.0	97%	DQ536719.1

Fig.4.1.8 BLASTn expression for West Eleri accession



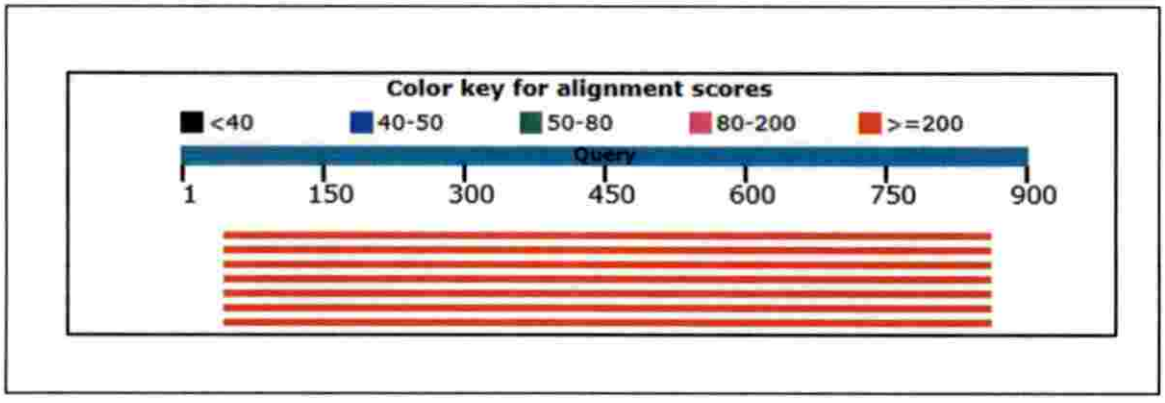
Description	Max score	Total score	Query cover	E value	Ident	Accession
Berkheya hispida voucher S. S. Renner et al. 2760 (MO, cult. Mainz BG) maturase K (matK) gene, partial cds; chloroplast	1465	1465	96%	0.0	97%	DQ536636.1
Berkheya fistulosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds; chloroplast	1452	1452	96%	0.0	97%	DQ536719.1
Coccinia trilobata maturase K-like (matK) gene, partial sequence; chloroplast	1435	1435	96%	0.0	96%	HQ609271.1

Fig.4.1.9 BLASTn for Areeparambu accession



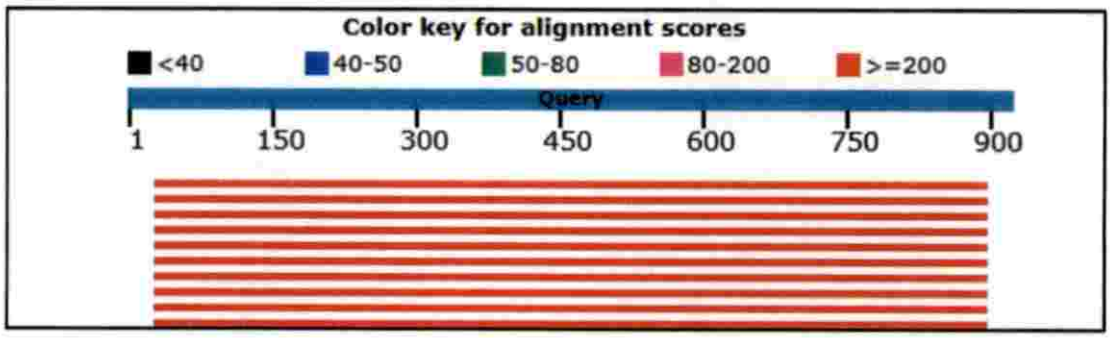
Description	Max score	Total score	Query cover	E value	Ident	Accession
Beninciasa hispida voucher S. S. Renner et al. 2700 (M), cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1465	1465	96%	0.0	97%	DQ536936.1
Beninciasa flabuliosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds; chloroplast	1452	1452	96%	0.0	97%	DQ536716.1

Fig.4.1.10 BLASTn for Nashik accession



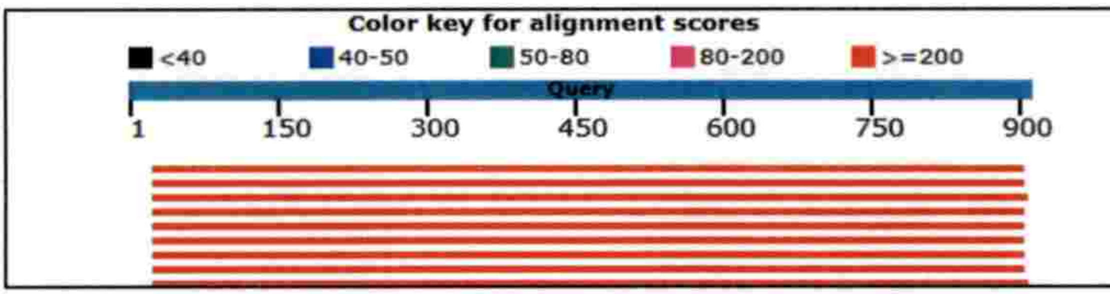
Description	Max score	Total score	Query cover	E value	Ident	Accession
Berlinicarpa hispida voucher S. S. Renner et al. 2760 (M). cult. Mainz B65 maturase K (matK) gene, partial cds; chloroplast	1480	1480	98%	0.0	99%	DQ536936.1
Berlinicarpa floridana voucher D. Decker-Walters 803 (FTG) maturase K (matK) gene, partial cds; chloroplast	1463	1463	98%	0.0	99%	DQ536719.1

Fig.4.1.11 BLASTn for accession IC 0596989



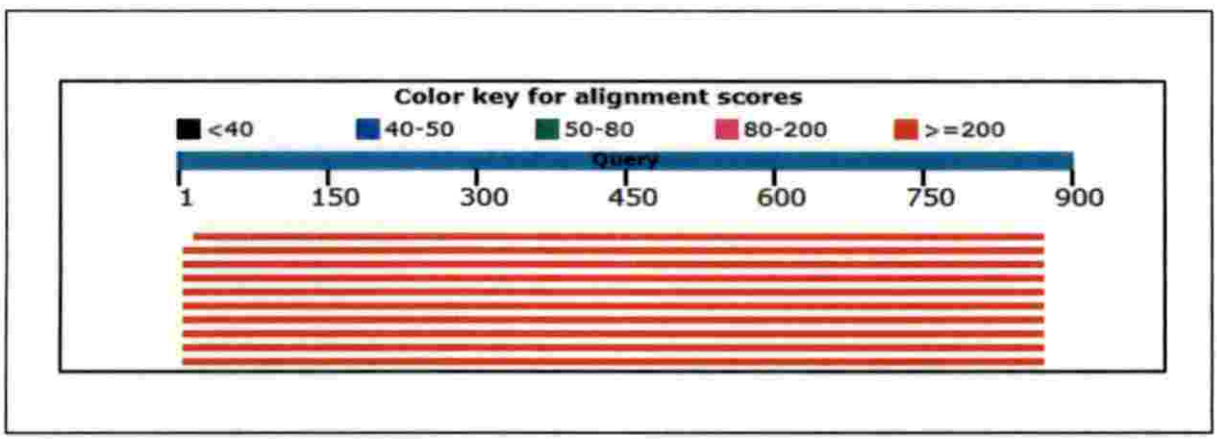
Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa hispida voucher S. S. Renner et al. 2769 (M). cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1452	1452	94%	0.0	97%	DQ536836.1
Benincasa flavidissa voucher D. Decker-Walters 893 (FTG) maturase K (matK) gene, partial cds; chloroplast	1439	1439	94%	0.0	97%	DQ536719.1

Fig.4.1.12 BLASTn for accession IC 0596992



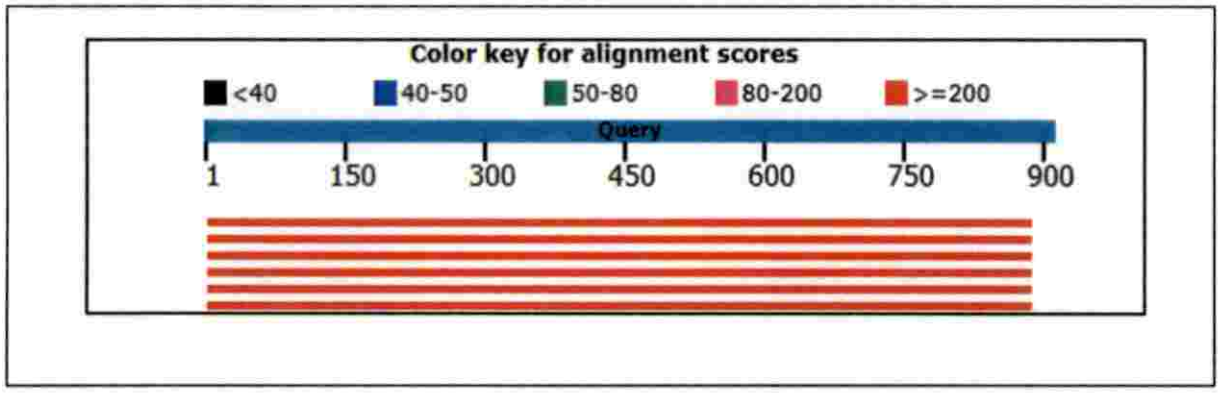
	Description	Max score	Total score	Query cover	E value	Ident	Accession
	Benincasa hispida voucher S. S. Renner et al. 2760 (M), cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1454	1454	96%	0.0	97%	DQ536636.1
	Benincasa fistulosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds; chloroplast	1443	1443	96%	0.0	97%	DQ536719.1

Fig.4.1.13 BLASTn for accession Elavancheri



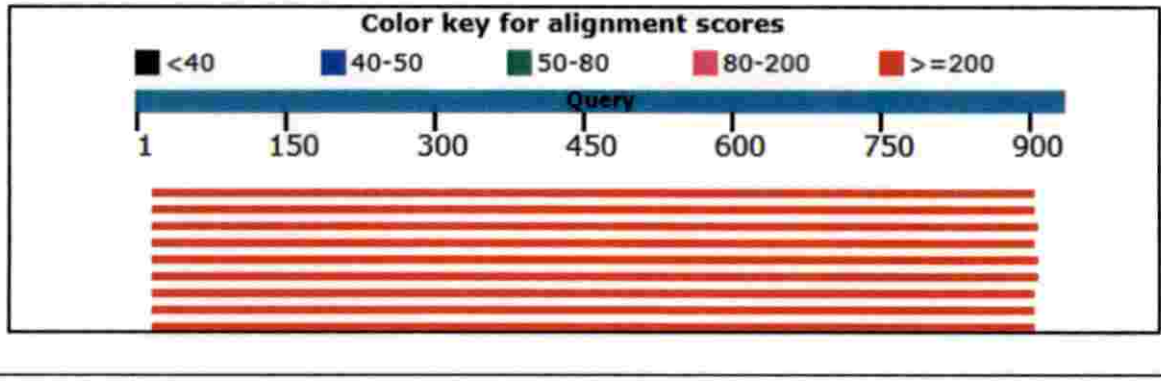
Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa hispida voucher S. S. Renner et al. 2760 (M), cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1531	1531	94%	0.0	99%	DQ536636.1

Fig.4.1.14 BLASTn for accession KAU local



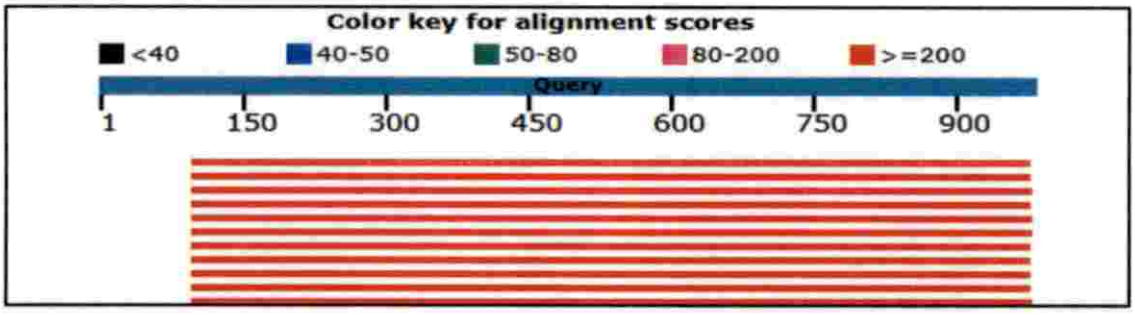
Description	Max score	Total score	Query cover	E value	Ident	Accession
Beloniciana hirsutifa voucher S. S. Renner et al. 2760 (M). cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1491	1491	96%	0.0	98%	DQ539836.1

Fig.4.1.15 BLASTn for Nirmal Tinda accession



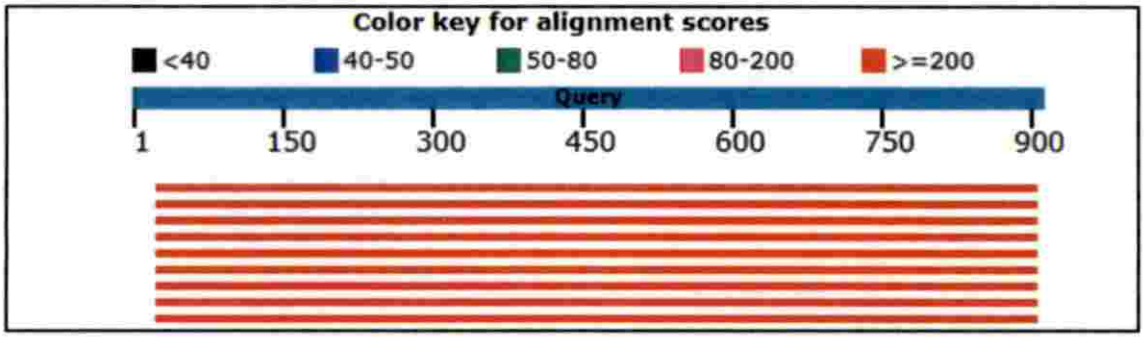
Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa fistulosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds; chloroplast	1502	1502	94%	0.0	98%	DQ539719.1
Benincasa hispida voucher S. S. Renner et al. 2756 (M), cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1471	1471	94%	0.0	97%	DQ539936.1

Fig.4.1.16 BLASTn for Ankur Tinda accession



Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa fistulosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds: chloroplast	1502	1502	94%	0.0	98%	DQ536719.1
Benincasa hispida voucher S. S. Renner et al. 2760 (MI, cult. Mainz BG) maturase K (matK) gene, partial cds: chloroplast	1471	1471	94%	0.0	97%	DQ536636.1

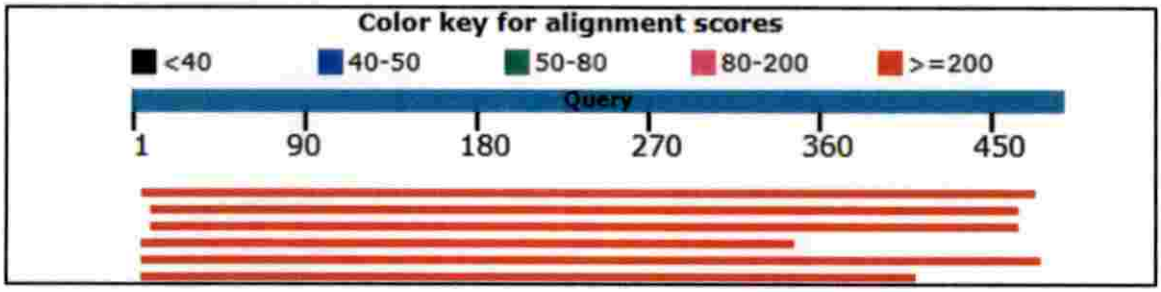
Fig.4.1.17 BLASTn for Komal tinda accession



Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa fistulosa voucher D. Decker-Walters 883 (FTG) maturase K (matK) gene, partial cds; chloroplast	1519	1519	96%	0.0	98%	DQ536719.1
Benincasa hispida voucher S. S. Renner et al. 2760 (M), cult. Mainz BG maturase K (matK) gene, partial cds; chloroplast	1487	1487	96%	0.0	97%	DQ536636.1

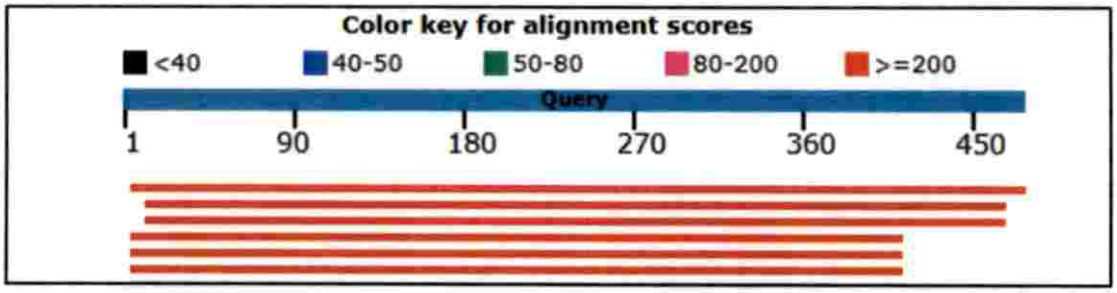
ITS2 BLAST

Fig.4.1.18 BLASTn for Kottamuri accession



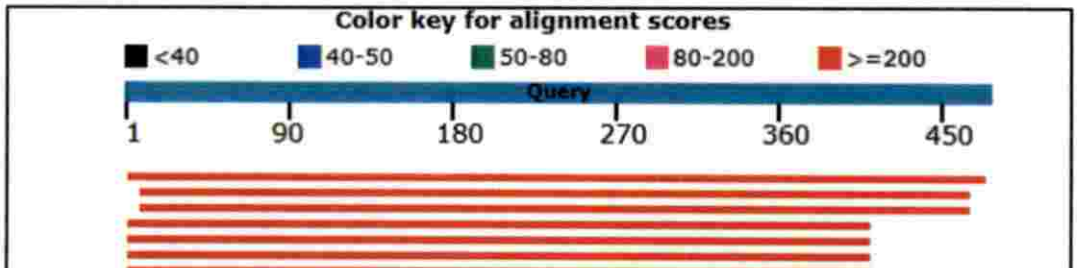
Description	Max score	Total score	Query cover	E value	Ident	Accession
Berlinciera triopida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	856	856	95%	0.0	99%	KR210320.1
Berlinciera triopida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	832	832	93%	0.0	99%	KX674775.1
Berlinciera triopida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	832	832	93%	0.0	99%	KX674774.1
Berlinciera triopida isolate F internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence	630	630	69%	1e-176	100%	KF835486.1

Fig.4.1.19 BLASTn for Vellarikundu accession



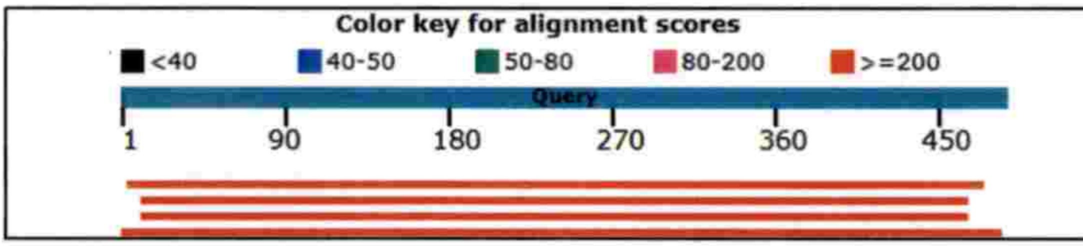
Description	score	score	cover	value	Ident	Accession
Benincasa hispida voucher 20130320106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	869	869	99%	0.0	100%	KM210320.1
Benincasa hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	95%	0.0	100%	KX674775.1
Benincasa hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	95%	0.0	100%	KX674774.1

Fig.4.1.20 BLASTn for West ELeri accession



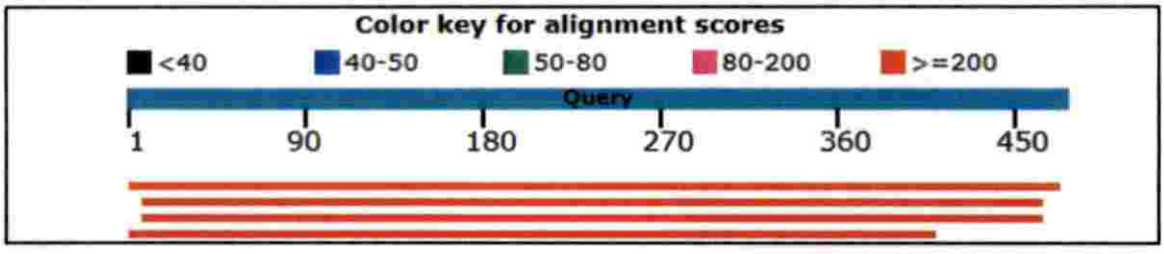
Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa hispida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	865	865	98%	0.0	100%	KM218320.1
Benincasa hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequer	837	837	95%	0.0	100%	KX674775.1
Benincasa hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequer	837	837	95%	0.0	100%	KX674774.1

Fig.4.1.22 BLASTn for Nashik accession



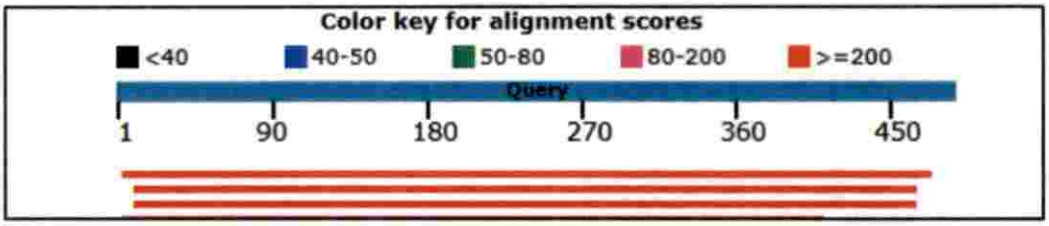
	Description	Max score	Total score	Query cover	E value	Ident	Accession
	Benincasa hispida voucher [illegible] 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	867	867	96%	0.0	100%	[illegible]
	Benincasa hispida voucher [illegible] 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	93%	0.0	100%	[illegible]
	Benincasa hispida voucher [illegible] 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	93%	0.0	100%	[illegible]

Fig.4.1.23 BLASTn for Areeparambu accession



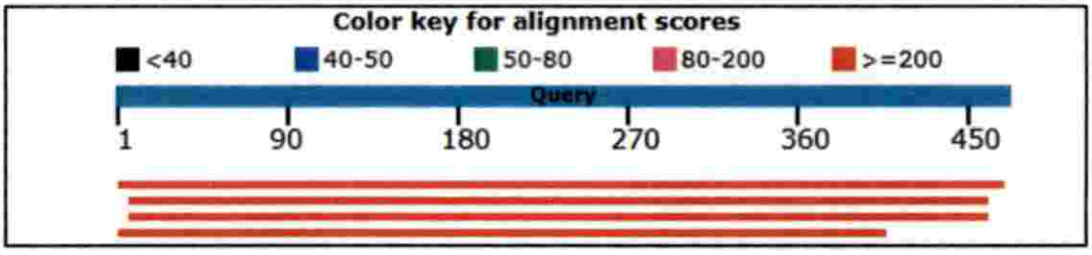
Description	Max score	Total score	Query cover	E value	Ident	Accession
Bennicasa hispida voucher 20130320106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	865	865	98%	0.0	100%	KM210320.1
Bennicasa hispida voucher PAN01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	95%	0.0	100%	KX674775.1

Fig.4.1.24 BLASTn for IC 0596989 accession



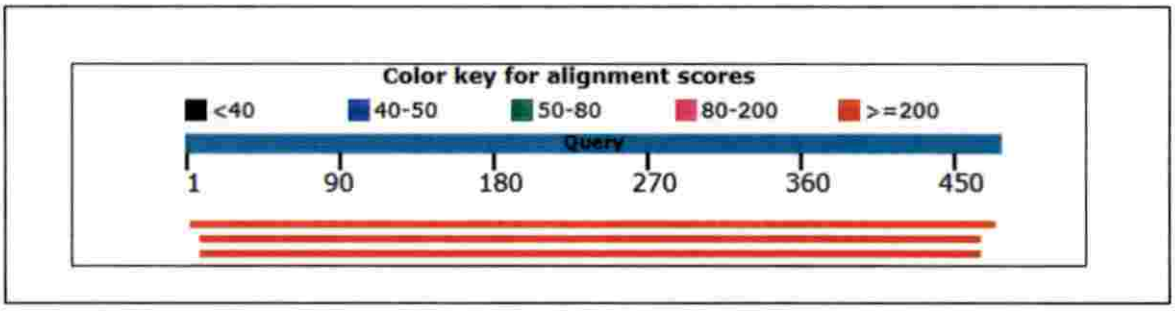
Description	Max score	Total score	Query cover	E value	Ident	Accession
Bennicasa hispida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	865	865	96%	0.0	100%	KM210328.1
Bennicasa hispida voucher PHR1264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequer	837	837	93%	0.0	100%	KX674775.1
Bennicasa hispida voucher PHR1263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequer	837	837	93%	0.0	100%	KX674774.1

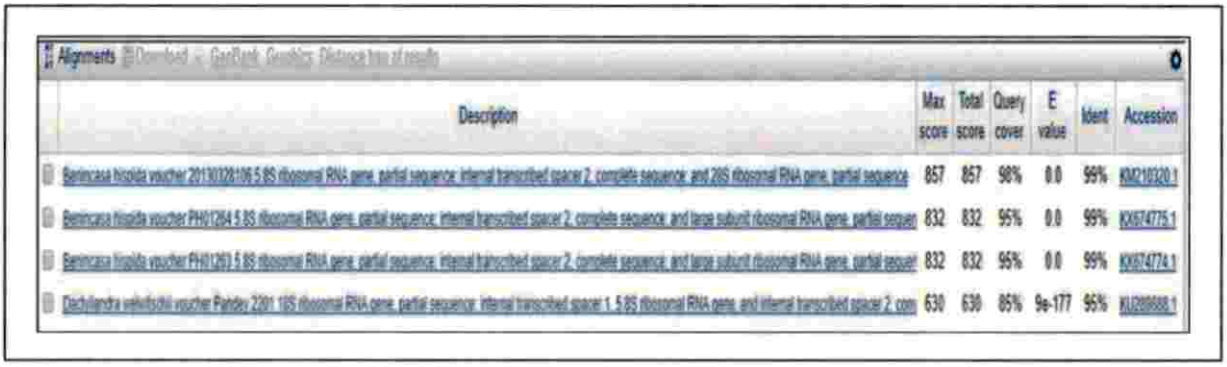
Fig.4.1.25 BLASTn for IC 0596992 accession



Description	Max score	Total score	Query cover	E value	Ident	Accession
Bemiscia hispida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	863	863	98%	0.0	100%	KM210320.1
Bemiscia hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	95%	0.0	100%	KX074775.1
Bemiscia hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	95%	0.0	100%	KX074774.1

Fig. 4.1.26 BLASTn for Elavancheri accession

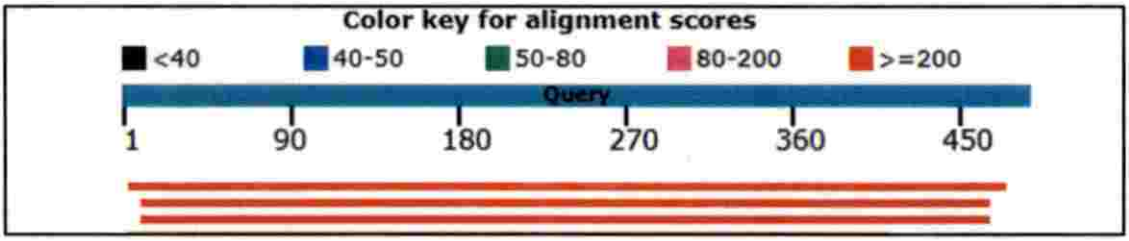




Alignments | Download | GenBank Graphics Distance tree of results

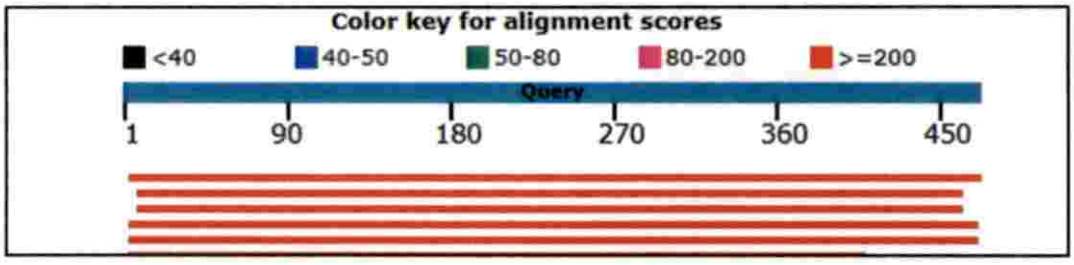
Description	Max score	Total score	Query cover	E value	Ident	Accession
Bennicasa hispida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	857	857	98%	0.0	99%	KM218320.1
Bennicasa hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequer	832	832	95%	0.0	99%	KX674775.1
Bennicasa hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequer	832	832	95%	0.0	99%	KX674774.1
Dichilandra wellsfitchii voucher Pandey 2201 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, com	630	630	85%	9e-177	96%	KU289688.1

Fig.4.1.27 BLASTn for KAU local



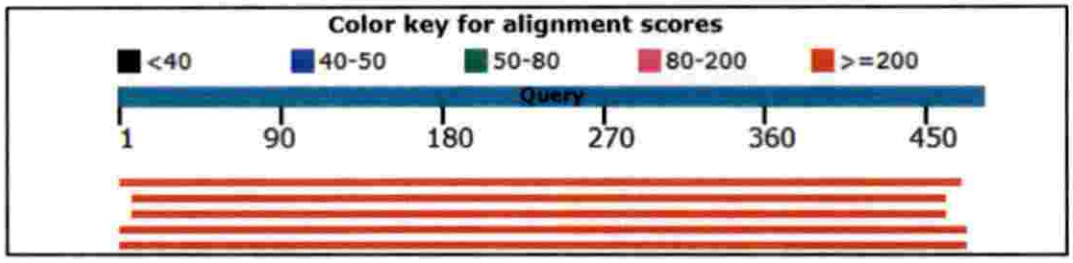
Description	Max score	Total score	Query cover	E value	Ident	Accession
Bemicasia trispila voucher 20130320106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	865	865	96%	0.0	100%	KM210320.1
Bemicasia trispila voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	93%	0.0	100%	KM674775.1
Bemicasia trispila voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	837	837	93%	0.0	100%	KM674774.1

Fig.4.1.28 BLASTn for Nirmal Tinda accession



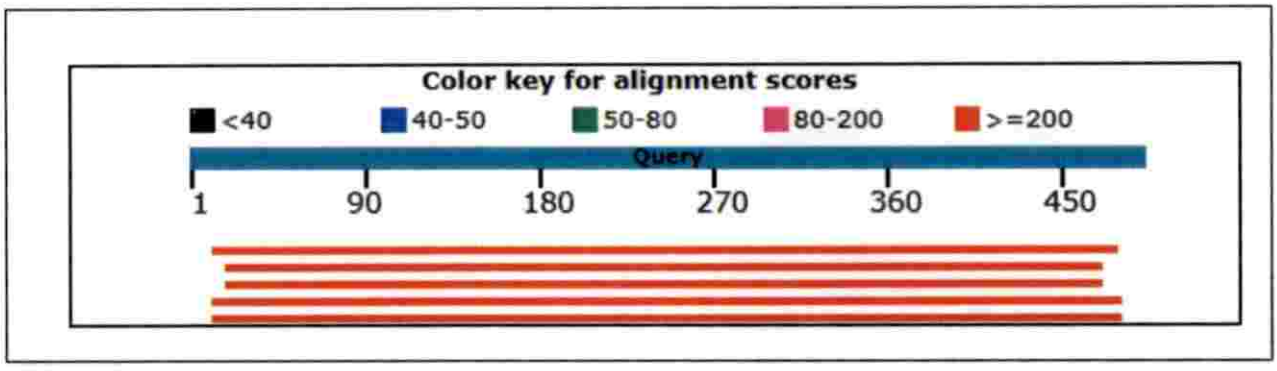
Description	Max score	Total score	Query cover	E value	Ident	Accessio
Benincasa hispida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	837	837	95%	0.0	99%	KM210328
Benincasa hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequenc	809	809	92%	0.0	99%	KX674775
Benincasa hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequenc	809	809	92%	0.0	99%	KX674774

Fig.4.1.29 BLASTn for Komal tinda accession



Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa hispida voucher 20130328106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	837	837	96%	0.0	99%	KM210328.1
Benincasa hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	809	809	93%	0.0	99%	KX674775.1
Benincasa hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	809	809	93%	0.0	99%	KX674774.1

Fig.4.1.30 BLASTn for Ankur Tinda



Description	Max score	Total score	Query cover	E value	Ident	Accession
Benincasa hispida voucher 20130326106 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	837	837	94%	0.0	99%	KM210320.1
Benincasa hispida voucher PH01264 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	809	809	91%	0.0	99%	KX674775.1
Benincasa hispida voucher PH01263 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequen	809	809	91%	0.0	99%	KX674774.1

Table.4.6 Homology of sequences – BLASTn analysis

S.I	Accession	Locus	Query coverage	Identity (%)	E value
1	Kottamuri	*matK*	97	97	0.0
2	Vellarikundu	*matK*	96	97	0.0
3	West Eleri	*matK*	96	97	0.0
4	Areeparambu	*matK*	90	99	0.0
5	Nashik	*matK*	96	97	0.0
6	IC 056989	*matK*	94	97	0.0
7	IC 056992	*matK*	96	97	0.0
8	Elavancheri	*matK*	94	99	0.0
9	KAU local	*matK*	96	96	0.0
10	Ankur tinda	*matK*	94	97	0.0
11	Komal tinda	*matK*	94	98	0.0
12	Kalash tinda	*matK*	94	98	0.0
13	Kottamuri	*ITS2*	95	99	0.0
14	Vellarikundu	*ITS2*	99	100	0.0
15	West Eleri	*ITS2*	98	100	0.0
16	Areeparambu	*ITS2*	96	100	0.0
17	Nashik	*ITS2*	96	100	0.0
18	IC 056989	*ITS2*	96	100	0.0

19	IC 056992	*ITS2*	96	100	0.0
20	Elavancheri	*ITS2*	98	99	0.0
21	KAU local	*ITS2*	96	100	0.0
22	Ankur tinda	*ITS2*	96	99	0.0
23	Komal tinda	*ITS2*	96	99	0.0
24	Kalash tinda	*ITS2*	94	99	0.0

4.14 Analysis of sequence homology

The sequence was assessed using MEGA BLASTn for all twelve accession for two different locus *matK* and *ITS2*.

For *matK* sequences, the accession Areeparambu, Nashik and Elavancheri shows 99% similarity towards NCBI accession *Benincasa hispida* on MEGA BLAST and query coverage and Identity value is 90-94% and 97-99% respectively. The accession KAU local shows 98% similarity towards NCBI accession *Benincasa hispida* with the query coverage of 96% and identity value 98%. The accessions Kottamuri, Vellarikundu, West Eleri, IC 0596989, IC 0596992 shows 97% similarity towards NCBI accession *Benincasa hispida* with the query coverage in the range of 94-96% and identity value in the range of 97-99%. For the tinda accessions Ankur tinda, Nirmal tinda and Komal tinda they shows 98% similarity towards NCBI accession *Benincasa hispida* and query coverage and identity value is 94-96% and 98% respectively.

For *ITS2* sequences, the accession Kottamuri, Vellarikundu, West Eleri, Areeparambu, Nashik, IC 0596989, IC 0596992, Elavancheri, KAU local shows 100% similarity towards NCBI accession *Benincasa hispida* on MEGA BLAST and query coverage and identity value is 95-98% and 99-100% respectively. For tinda accessions Ankur tinda, Nirmal tinda and Komal tinda shows 99% similarity towards NCBI accession *Benincasa hispida* and query coverage and identity value is 94-96% and 99% respectively (Table. 4.5 & Fig.4.1.18-4.1.30).

4.15 Barcoding gap assessment and barcode generation

The distribution of inter-specific variability was compared using DNA barcoding gap. This DNA barcode had separate and non-overlapping variations among all the species. The results had shown that species variation of *ITS2 loci* exhibited distinct gaps among ashgourd genotypes and tinda genotypes, separating both the species.

The ashgourd and tinda genotypes were aligned with software clustalW omega. The conserved region was shown '*' in the aligned sequence. The variability have been check manually. The nucleotide sequence and variability have been observed specific to each of the species.

All the sequences are aligned and the maximum length of *matK* gene developed from this study was 931 bp and *ITS2* gene was 489 bp. After identifying the gaps with position for each species, the barcodes were developed (Fig.4.1.33 & 4.1.34).

BLAST was used for assessment of correct discrimination using *matK* and *ITS2* barcode. The results identified on BLAST have identified that matK has given high efficiency of 99% at the genus level in *Benincasa* and *ITS2* loci gave an identification efficiency of 100% at the genus level in *Praecitrullus*.

In this study, barcode gaps are observed only for tinda accessions, for *matK* locus they are observed at position 47(G→T), 127 (T→C) and 232 (T→C) for accessions Komal tinda, Nirmal tinda and Ankur tinda respectively and for ITS2 loci, barcode gaps are observed at position 71(T→C), 207(T→C), 278 &279 (C→T) and 288(C→T) for all of tinda accessions included in this work. It is concluded, that *Praecitrullus fistulosus* is totally different species from *Benincasa hispida* (Levi, 2005) and (Levi, 2009) So *P. fistulosus*, cannot be included in *B. hispida* even though, there is wide availability in morphology, all the genotypes studied belongs to *B. hispida* species only.

4.16 BLAST for species discrimination

The sequences were further subjected to BLAST to discriminate the species. Results indicate that *ITS2* loci is powerful to yield upto 95% efficiency in differentiating the species. At the genus level, *ITS2* loci and *matK* both yielded 100 % efficiency.

4.17 Identification of conserved region

The conserved sequence in the generated sequence was identified using ClustalW software. A total of 12 sequences were aligned and the conserved region across the total length of 980 base pairs identified.

In the present study, the universally reported barcodes and primers were used to amplify the *matK* locus and *ITS2* loci. The selected primers were successful to amplify the locus across the accessions.

For *matK* loci, (Fig.4.1.31), the accession showed more barcode gaps. The maximum length of matk sequences were 926 base pairs and the barcode gaps were identified at alignment positions; for accessions Komal tinda 47(G→T), Nirmal tinda 47(G→T), Ankur tinda 47(G→T), Komal tinda 127(T→C), Nirmal tinda 127(T→C), Ankur tinda 127 (T→C), Komal tinda 232(T→C), Nirmal tinda 232(T→C), Ankur tinda 232(T→C).

For *ITS* loci, (Fig.4.1.32), the accession showed barcode gap at three instances, where ashgourd genotypes can easily differentiated from tinda genotypes. The maximum length of *ITS* sequences 496 base pairs and the barcode gaps were identified at alignment positions; 162 for accession for Komal tinda (T→C), for accession Nirmal tinda 162(T→C) and Ankur tinda 162(T→C). For accession Komal tinda 207(T→C), Ankur tinda 207(→C) and accession Nirmal tinda 207(T→C). For accession Komal tinda 288(C→T), for acession Ankur tinda 288(C→T) and Nirmal tinda 288(C→T). For accession Ankur tinda, 279(C→T), for accession Komal tinda 279(C→T), Nirmal tinda 279(C→T). For accession Ankur tinda 278(C→T), Komal tinda 278(C→T), Nirmal tinda 278(C→T).

In this analysis, both loci proved to be more accurate to differentiate the various species.

Primer set for DNA barcoding in ashgourd and tinda using *matK* and *ITS2*

Product size: 897

Primer set for DNA barcoding in ashgourd and tinda using *ITS2*

Product size: 799

A total 26 genotypes of ashgourd and 5 genotypes of tinda were amplified with both the primers and both of them have shown result in all accessions (Fig.4.1.6. & 4.1.7).

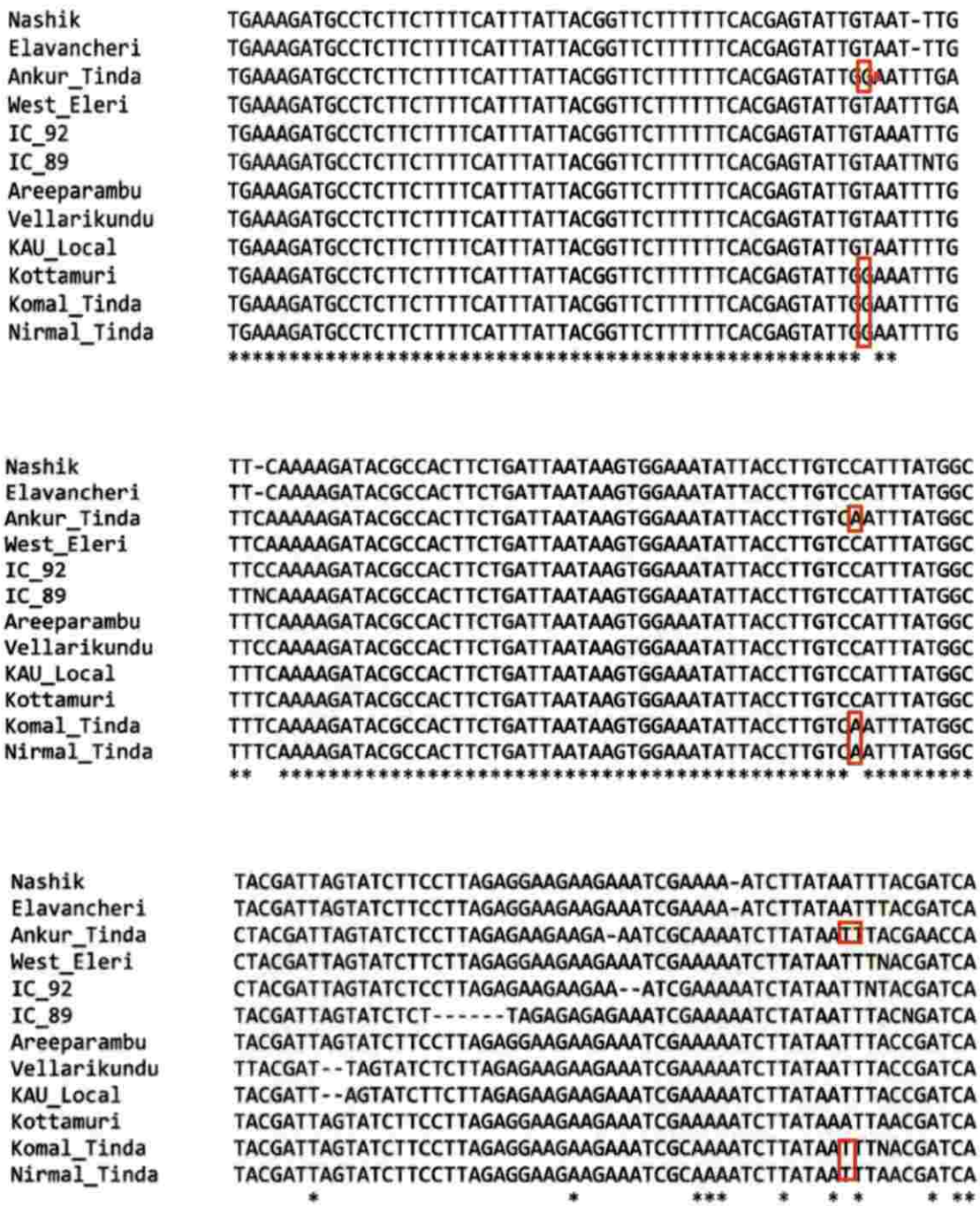
Table 4.7 Alignment position of barcode gaps and type of substitution using *matK* sequence

S.I No.	Alignment position	Substitution	Types of substitution	Species
1	47	G→T	Transition	Komal tinda
2	47	G→T	Transition	Nirmal tinda
3	47	G→T	Transition	Ankur tinda
4	127	T→C	Transition	Nirmal tinda
5	127	T→C	Transition	Komal tinda
6	127	T→C	Transition	Ankur tinda
7	232	A→C	Transition	Nirmal tinda
8	232	A→C	Transition	Komal tinda
9	232	A→C	Transition	Ankur tinda

Table 4.8 Alignment position of barcode gaps and types of substitution using *ITS*2 sequence

S.I. No	Alignment position	Substitution	Type of substitution	Species
1	162	T→C	Transition	Komal tinda
2	162	T→C	Transition	Nirmal tinda
3	162	T→C	Transition	Ankur tinda
4	207	T→C	Transition	Komal tinda
5	207	T→C	Transition	Nirmal tinda
6	207	T→C	Transition	Ankur tinda
7	279	C→T	Transition	Komal tinda
8	279	C→T	Transition	Ankur tinda
9	279	C→T	Transition	Nirmal tinda
10	278	C→T	Transition	Komal tinda
11	278	C→T	Transition	Ankur tinda
12	278	C→T	Transition	Nirmal tinda
13	288	C→T	Transition	Komal tinda
14	288	C→T	Transition	Ankur tinda
15	288	C→T	Transition	Nirmal tinda

Fig. 4.1.31 Multiple sequence alignment for *matK* sequences



4.1.32 Multiple sequence alignment for *ITS2* sequences

```
Kottamuri          CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
Vellarikundu       CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
Nashik             CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
Elavancheri        CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCATCGTCGTGCGG
IC_89              CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
Areeparambu        CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
KAU_Local          CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
West_Eleri         CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
IC_92              CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCACCGTCGTGCGG
Komal_Tinda        CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCATCGTCGTGCGG
Nirmal_Tinda       CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCATCGTCGTGCGG
Ankur_Tinda        CACACAACTCGTTGTGCAGGCGGGGGCACATGTTGGCCTCCCGTGCGCATCGTCGTGCGG
                   ************************************************** **********

Kottamuri          ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
Vellarikundu       ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
Nashik             ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
Elavancheri        ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
IC_89              ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
Areeparambu        ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
KAU_Local          ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
West_Eleri         ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
IC_92              ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGCGACACTACGGTGGTTGATCCAACC
Komal_Tinda        ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGTGACACTACGGTGGTTGATCCAACC
Nirmal_Tinda       ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGTGACACTACGGTGGTTGATCCAACC
Ankur_Tinda        ATGGCTTAAATTCGAGTCCTCGGCGCACGTCGTCGTGACACTACGGTGGTTGATCCAACC
                   ************************************ ************************

Kottamuri          TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
Vellarikundu       TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
Nashik             TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
Elavancheri        TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
IC_89              TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
Areeparambu        TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
KAU_Local          TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
West_Eleri         TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
IC_92              TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCATTGACCCTCTGAGC
Komal_Tinda        TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCACCGACCCTCCGAGC
Nirmal_Tinda       TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCACCGACCCTCCGAGC
Ankur_Tinda        TCAGTACCATGTCGCGGCCTCGACCCCGCCTCCACGGACTCATGCACCGACCCTCCGAGC
                   ************************************************** ********* ****
```

4.18 Submission to NCBI

A total of 9 sequences of the *Benincasa* species and 3 of *Praecitrullus* were 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/how to submit-sequence-data/website sequence with simple annotation was submitted through BankIt.

(ii) 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 or plant from which the sequence data were isolated and any other related descriptive data were also be provided.

(iii) The allocation of accession number is awaited.

Fig.4.1.33 DNA barcode gaps generated for *Benincasa*and *Praecitrullus* species under the study the distinct barcode gaps are formed for *ITS2* loci

Species	50	55	60	71	89	125	162	163	164	165	168	188	190	200	205	206	207	209	210	229	230	234	255	260	277	278	279	278	279	280	281	286	288	290	291	300	305	380	450	494
Kottamuri	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
Vellarikundu	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
West Eleri	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
Areeparambu	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
Nashik	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
IC 89	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
IC 92	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
Elavancheri	C	A	C	A	C	A	T	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
KAU local	C	A	C	A	C	A	C	C	T	C	G	G	G	G	C	A	C	A	T	G	G	C	C	T	C	C	T	T	T	G	C	G	T	C	C	A	C	G	G	G
Ankur tinda	C	A	C	A	C	A	T	C	T	C	G	G	G	G	C	A	T	A	T	G	G	C	C	T	C	C	C	C	T	G	C	G	C	C	C	A	C	G	G	G
Nirmal tinda	C	A	C	A	C	A	T	C	T	C	G	G	G	G	C	A	T	A	T	G	G	C	C	T	C	C	C	C	T	G	C	G	C	C	C	A	C	G	G	G
Komal tinda	C	A	C	A	C	A	T	C	T	C	G	G	G	G	C	A	T	A	T	G	G	C	C	T	C	C	C	C	T	G	C	G	C	C	C	A	C	G	G	G

Fig.4.1.34 DNA barcode gaps generated for *Benincasa* and *Praecitrullus* species under the study the distinct barcode gaps are formed for *matK* loci

Species	22	24	30	41	46	47	50	68	69	75	78	85	90	102	127	130	131	135	140	155	160	165	232	235	240	245	340	400	445	500	553	780	825	926
Kottamuri	T	A	G	T	T	G	A	A	C	T	C	T	C	T	C	T	A	G	A	G	T	A	C	A	C	C	T	T	T	G	A	A	T	G
Vellarikundu	T	A	G	T	T	T	C	A	T	T	C	T	C	T	C	T	C	T	T	G	T	C	C	C	C	C	T	T	T	T	A	A	T	G
West Eleri	C	T	G	A	G	T	A	T	T	T	T	T	C	T	C	T	A	T	T	A	T	A	C	A	C	C	T	A	A	T	A	A	T	G
Areeparambu	T	A	G	T	T	T	C	A	C	T	C	T	C	T	C	T	C	T	T	G	A	C	C	C	C	A	T	A	T	T	A	A	C	C
Nashik	T	A	G	T	T	T	A	A	C	T	C	A	T	C	C	T	A	T	T	G	T	C	C	C	C	T	T	T	A	T	C	T	T	C
IC 89	T	A	G	T	T	T	C	A	T	C	T	A	T	C	C	G	C	T	T	G	T	A	C	C	T	C	T	T	T	T	A	A	T	C
IC 92	C	T	G	A	G	T	A	T	T	T	C	A	T	C	C	T	A	T	T	G	T	C	C	A	C	C	T	T	T	G	A	A	T	G
Elavancheri	T	A	G	T	T	T	A	A	C	T	C	A	T	C	C	T	A	T	T	G	G	C	C	C	C	C	T	A	A	T	A	T	A	G
KAU local	T	A	G	T	T	T	C	G	A	C	T	T	C	T	C	T	A	T	T	G	T	C	C	C	C	C	C	T	T	T	T	A	A	T	G
Ankur tinda	C	T	G	A	G	G	A	T	T	T	C	T	C	T	T	T	C	T	T	A	T	A	A	C	C	C	T	T	T	T	A	A	T	G
Nirmal tinda	T	A	G	T	T	G	A	A	C	T	C	T	C	T	T	G	A	T	T	G	T	C	A	A	C	T	T	T	T	T	A	A	T	G
Komal tinda	T	A	G	T	T	G	A	A	C	T	C	T	C	T	T	T	A	T	T	G	T	C	A	A	T	C	T	T	T	T	A	A	T	G

Table. 4.1 Morphological evaluation of ashgourd and tinda accessions

Morphological characters	Kottamuri	Vellarikundu	West Eleri	Areeparambu	Kanjirampoyil	Nedungadapalli	Mallapally
Early plant vigour	Very good	Very good	Very good	Very good	Very good	Very good	Very good
Plant growth habit	Medium viny	Long viny	Medium viny	Long viny	Long viny	Long viny	Long viny
Stem pubescence	Hispid	Hispid	Downy	Downy	Tomentose	Hispid	Puberulent
Stem shape	Angular	Angular	Angular	Angular	Angular	Angular	Angular
Tendril	Present	Present	Present	Present	Present	Present	Present
Tendril type	Coiled	Coiled	Coiled	Coiled	Coiled	Coiled	Coiled
Tendril branching	Branched	Branched	Branched	Branched	Branched	Branched	Branched
Leaf margin	Serrated	Deep serrated	Double serrated	Double serrated	Serrated	Deep serrated	Deep serrated
Leaf shape	Cordate	Cordate	Cordate	Cordate	Cordate	Cordate	Cordate
Leaf size	Small	Medium	Medium	Medium	Large	Small	Small
Leaf pubescence nature	Intermediate	Soft	Soft	Soft	Soft	Hispid	Puberulent
Leaf pubescence density	Intermediate	Sparse	Sparse	Sparse	Sparse	Sparse	Sparse
Node number at which female flower appears	11	14	15	13	17	18	13
Days to 50% flowering	25	26	25	27	30	28	28
Petiole length(cm)	8.1	20.4	23.1	18	14.5	18.2	15
Sex type	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious
Peduncle length(cm)	5.2	6.4	10.7	5.8	6.5	6.8	6.5
Peduncle shape	Round	Angular	Angular	Angular	Round	Round	Round
Peduncle attachment	Soft and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Soft and flared	Soft and flared

Fruit shape	Club	Spindle	Club	Club	Cylindrical	Round	Round
Fruit skin colour	Dark green	Dark green	Dark green	Dark green	Dark green	Dark green	Dark green
Fruit skin lustre	Matt	Matt	Matt	Matt	Matt	Matt	Matt
Fruit pubescence density	No hairs	No hairs	No hairs	No hairs	No hairs	No hairs	No hairs
Fruit ridge shape	Superficial	Superficial	Superficial	Superficial	Superficial	Superficial	Superficial
Number of ridges fruits per fruit	0	0	0	0	0	0	0
Stem end fruit shape	Rounded	Rounded	Depressed	Rounded	Flattened	Rounded	Rounded
Blossom end fruit shape	Rounded	Rounded	Flattened	Rounded	Flattened	Rounded	Rounded
Fruit length (cm)	16.5	22.0	50.4	19.0	40.0	14.0	13.0
Fruit width (cm)	31.0	34.3	60.1	33.5	58.9	26.6	26.2
Fruit weight (Kg)	0.565	0.968	12.0	0.828	5.983	0.335	0.307
Flesh texture	Smooth	Smooth	Smooth	Smooth	Smooth	Smooth	Smooth
Flesh colour	Creamish	Creamish	Creamish	Creamish	Creamish	Creamish	Creamish
Seed lustre	Matt	Glossy	Glossy	Intermediate	Matt	Matt	Matt
Seediness	Medium	Medium	Very high	Very high	High	Medium	Medium
Number of primary branches	3	3	3	2	3	4	7
Days to first harvest	45	46	47	48	45	46	45
Days to last harvest	46	48	47	49	48	47	47
Number of fruits per plant	8	5	8	8	7	6	9

Morphological characters	**North Paravur**	**Perumbilishe-rry**	**Charamang-alam**	**Wayanad**	**Kannara**	**Nashik**	**No.5**
Early plant vigour	Very good	Very good	Good	Good	Very good	Very good	Very good

Plant growth habit	Long viny	Long viny	Medium viny	Medium viny	Long viny	Long viny	Long viny
Stem pubescence	Hispid	Hispid	Hispid	Hispid	Hispid	Downy	Pilose
Stem shape	Angular	Angular	Angular	Angular	Angular	Angular	Angular
Tendril	Present	Present	Present	Present	Present	Present	Present
Tendril type	Coiled	Coiled	Coiled	Coiled	Coiled	Coiled	Coiled
Tendril branching	Branched	Branched	Branched	Branched	Branched	Branched	Branched
Leaf margin	Deep serrated	Deep serrated	Deep serrated	Deep serrated	Serrate	Double serrated	Double serrated
Leaf shape	Cordate	Cordate	Cordate	Cordate	Cordate	Cordate	Cordate
Leaf size	Small	Large	Medium	Medium	Medium	Medium	Medium
Leaf pubescence nature	Hispid	Hispid	Hispid	Intermediate	Intermediate	Intermediate	Intermediate
Leaf pubescence density	Sparse	Densed	Densed	Densed	Densed	Densed	Densed
Node number at which female flower appears	15	9	8	7	8	10	11
Days to 50% flowering	26	29	27	25	26	28	29
Petiole length(cm)	11.5	16.3	24.5	15.0	18.0	17.8	21.3
Sex type	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious
Peduncle length	6.1	4.6	2.8	6.0	13.8	7.0	6.5
Peduncle shape	Angular	Angular	Angular	Angular	Angular	Angular	Angular
Peduncle attachment	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared
Fruit shape	Round	Round	Round	Round	Oblong medium	Round	Club
Fruit skin colour	Dark green	Dark green	Dark green	Dark green	Dark green	Light green	Light green
Fruit skin lustre	Matt	Matt	Matt	Matt	Matt	Matt	Matt
Fruit pubescence density	No hairs	No hairs	No hairs	No hairs	No hairs	No hairs	No hairs
Fruit ridge shape	Superficial	Superficial	Superficial	Superficial	Superficial	Superficial	Superficial

Number of ridges fruits per fruit	0	0	0	0	0	0	0
Stem end fruit shape	Rounded	Rounded	Flattened	Flattened	Rounded	Flattened	Flattened
Blossom end fruit shape	Rounded	Rounded	Flattened	Flattened	Rounded	Flattened	Flattened
Fruit length (cm)	15.6	22.4	34.2	32.0	22.0	24.3	26.0
Fruit width (cm)	29.0	29.3	45.1	45.0	36.2	35.4	42.0
Fruit weight (kg)	0.443	0.390	0.565	4.500	0.979	4.335	2.307
Flesh texture	Smooth	Smooth	Smooth	Smooth	Smooth	Smooth	Smooth
Flesh colour	Creamish	Creamish	Creamish	Creamish	Creamish	Creamish	Creamish
Seed lustre	Matt	Matt	Glossy	Glossy	Intermediate	Glossy	Glossy
Seediness	Medium	Low	Medium	Very high	Very high	High	High
Number of primary branches	2	3	4	4	4	3	3
Days to first harvest	48	47	46	45	46	45	46
Days to last harvest	48	48	49	49	46	47	47
Number of fruits per plant	8	6	5	8	4	7	6

Morphological characters	**AccNo.9**	**AccNo.10**	**AccNo.11**	**IC 0596986**	**IC 0596985**	**IC 0596987**	**IC 0596989**
Early plant vigour	Very good	Good	Very good	Very good	Good	Good	Very good
Plant growth habit	Medium viny	Medium viny	Medium viny	Short viny	Short viny	Short viny	Short viny
Stem pubescence	Villlous	Pilose	Hispid	Hispid	Hispid	Hispid	Hispid
Stem shape	Angular	Angular	Angular	Angular	Angular	Angular	Angular
Tendril	Present	Present	Present	Present	Present	Present	Present
Tendril type	Coiled	Coiled	Coiled	Coiled	Coiled	Coiled	Coiled

Tendril branching	Branched	Branched	Branched	Branched	Branched	Branched	Branched
Leaf margin	Double serrated	Double serrated	Double serrated	Deep serrated	Deep serrated	Deep serrated	Deep serrated
Leaf shape	Cordate	Cordate	Cordate	Cordate	Cordate	Cordate	Cordate
Leaf size	Large	Large	Large	Small	Small	Small	Small
Leaf pubescence nature	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Leaf pubescence density	Sparse	Densed	Densed	Sparse	Sparse	Sparse	Sparse
Node number at which female flower appears	5	13	10	11	10	10	10
Days to 50% flowering	27	26	28	22	26	25	28
Petiole length (cm)	11.5	22.5	16.2	15.2	12.5	17.0	17.5 cm
Sex type	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious
Peduncle length (cm)	5.2	5.5	10.0	7.0	8.2	7.5	7.2
Peduncle shape	Angular	Angular	Angular	Angular	Angular	Angular	Angular
Peduncle attachment	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared
Fruit shape	Club	Club	Club	Club	Club	Club	Ridge
Fruit skin colour	Dark green	Dark green	Dark green	Dark green	Dark green	Dark green	Dark green
Fruit skin lustre	Matt	Matt	Matt	Matt	Matt	Matt	Matt
Fruit pubescence density	No hairs	No hairs	No hairs	No hairs	No hairs	No hairs	No hairs
Fruit ridge shape	Superficial	Superficial	Superficial	Superficial	Superficial	Superficial	Deep groove

Number of ridges fruits per fruit	0	0	0	0	0	0	12
Stem end fruit shape	Flattened	Flattened	Flattened	Depressed	Depressed	Depressed	Flattened
Blossom end fruit shape	Flattened	Flattened	Flattened	Depressed	Depressed	Depressed	Flattened
Fruit length (cm)	32.0	37.0	31.5	26.3	32.3	24.3	21.4
Fruit width (cm)	45.0	52.1	59.0	42.8	50.3	48.3	49.5
Fruit weight (kg)	4.500	4.410	3.749	3.677	4.203	1.760	6.890
Flesh texture	Smooth	Smooth	Smooth	Smooth	Smooth	Smooth	Smooth
Flesh colour	Creamish	Creamish	Creamish	Creamish	Creamish	Creamish	Creamish
Seed lustre	Glossy	Glossy	Glossy	Glossy	Glossy	Glossy	Glossy
Seediness	Very high	Very high	Very high	Very high	Very high	Very high	Very high
Number of primary branches	3	3	3	4	4	4	4
Days to first harvest	47	47	45	45	46	44	45
Days to last harvest	47	49	48	49	47	48	48
Number of fruits per plant	4	3	3	4	4	4	9

Morphological characters	IC 0596992	IC 0596995	Elavancheri	Thanniam	KAU local
Early plant vigour	Very good	Very good	Very good	Very good	Very good
Plant growth habit	Short viny	Short viny	Medium viny	Long viny	Medium viny
Stem pubescence	Villous	Hispid	Hirsute	Hispid	Hispid
Stem shape	Angular	Angular	Angular	Angular	Angular
Tendril	Present	Present	Present	Present	Present
Tendril type	Coiled	Coiled	Coiled	Coiled	Branched
Tendril branching	Branched	Branched	Branched	Branched	Coiled
Leaf margin	Deep serrate	Deep serrate	Double serrated	Serrate	Double serrated
Leaf shape	Cordate	Cordate	Cordate	Cordate	Cordate
Leaf size	Small	Small	Large	Small	Large
Leaf pubescence nature	Intermediate	Intermediate	Intermediate	Soft	Soft
Leaf pubescence density	Sparse	Sparse	Densed	Sparse	Densed
Node number at which female flower appears	4	8	8	9	8

Days to 50% flowering	28	24	27	26	24
Petiole length (cm)	22.5	14.7	22.8	13.5	23.5
Sex type	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious
Peduncle length (cm)	3.2 cm	6.0	6.5	6.0	6.5
Peduncle shape	Angular	Angular	Angular	Angular	Angular
Peduncle attachment	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared
Fruit shape	Club	Club	Cylindrical	Round	Club
Fruit skin colour	Dark green	Dark green	Light green	Light green	Light green
Fruit skin lustre	Matt	Matt	Matt	Matt	Matt
Fruit pubescence density	No hairs	No hairs	No hairs	No hairs	No hairs
Fruit ridge shape	Superficial	Superficial	Superficial	Superficial	Superficial
Number of ridges fruits per fruit	0	0	0	0	0
Stem end fruit shape	Flattened	Flattened	Rounded	Rounded	Rounded
Blossom end fruit shape	Flattened	Flattened	Rounded	Rounded	Rounded
Fruit length (cm)	29.3	24.2	37.0	4.6	15.0
Fruit width (cm)	49.4	48.5	39.3	10.3	28.0
Fruit weight (kg)	2.233	1.998	2.730	0.191	3.500
Flesh texture	Smooth	Smooth	Smooth	Smooth	Smooth
Flesh colour	Creamish	Creamish	Creamish	Creamish	Creamish
Seed lustre	Glossy	Glossy	Glossy	Glossy	Glossy
Seediness	Very high	Very high	High	Very high	Very high
Number of primary branches	4	5	3	3	3
Days to first harvest	48	46	47	45	45
Days to last harvest	49	48	48	49	60
Number of fruits per plant	2	3	3	7	10

Morphological characters	Ankurtinda	Nirmaltinda	Komaltinda	MAHYCOtinda	Kailashtinda
Early plant vigour	Good	Very good	Very good	Very good	Very good
Plant growth habit	Short viny	Short viny	Short viny	Short viny	Short viny
Stem pubescence	Hispid	Hispid	Hispid	Hispid	Hispid
Stem shape	Angular	Angular	Angular	Angular	Angular
Tendril	Present	Present	Present	Present	Present
Tendril type	Branched	Branched	Branched	Branched	Branched
Tendril branching	Coiled	Coiled	Coiled	Coiled	Coiled
Leaf margin	Incised	Incised	Incised	Incised	Incised
Leaf shape	Cordate	Cordate	Cordate	Cordate	Cordate
Leaf size	Small	Small	Small	Small	Small
Leaf pubescence nature	Sparse	Sparse	Sparse	Sparse	Sparse
Leaf pubescence density	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Node number at which female flower appears	5	6	3	1	2
Days to 50% flowering	22	23	24	24	21
Sex type	Monoecious	Monoecious	Monoecious	Monoecious	Monoecious
Peduncle length (cm)	2.5	2.5	2.5	2.5	2.5
Peduncle attachment	Hard and flared	Hard and flared	Hard and flared	Hard and flared	Hard and flared
Fruit shape	Round	Round	Round	Round	Round
Fruit skin colour	Light green	Light green	Light green	Light green	Light green
Fruit skin lustre	Glossy	Glossy	Glossy	Glossy	Glossy
Fruit pubescence	Sparse	Sparse	Sparse	Sparse	

density					Sparse
Fruit ridge (rib) shape	Circular	Circular	Circular	Circular	Circular
Number of ridges per fruit	0	0	0	0	0
Stem-end fruit shape	Rounded	Rounded	Rounded	Rounded	Rounded
Blossom-end fruit shape	Rounded	Rounded	Rounded	Rounded	Rounded
Fruit length (cm)	7.1	6.0	7.9	8.0	8.1
Fruit width (cm)	8.0	9.0	9.1	9.0	9.2
Fruit weight (kg)	0.250	0.250	0.250	0.250	0.250
Flesh texture	Firmed	Firmed	Firmed	Firmed	Firmed
Flesh colour	White	White	White	White	White
Seed lustre	Matt	Matt	Matt	Matt	Matt
Seediness	High	High	High	High	High
Number of primary branches	3	3	3	3	3
Days to first fruit harvest	35	34	35	37	35
Days to last fruit harvest	36	34	36	38	36
Number of fruits per plant	6	5	4	4	3

Discussion

5. DISCUSSION

There is substantial variation in ashgourd with respect to vegetative traits, and it is difficult to distinguish ashgourd genotypes based on their external morphology alone. Even after accounting this enormous variability all members are accommodated into a single species.

Praecitrullus is a monotypic genus with creepy, climbing stout herb. In the genus *Praecitrullus* there is the only one cultivated species.

Benincasa hispida and *Praecitrullus fistulosus* are both considered as minor cucurbits, although it is cultivated all over the world. *P. fistulosus* share morphological similarity with watermelon (*C. lanatus* spp.) thus it is considered as close relative and was named earlier as *Citrullus vulgaris* var. *fistulosus* (tinda). (Pangalo, 1938; Sujata and Seshadri, 1989). But the chromosome number of *C. lanatus* (n=11) and *P. fistulosus*(n=12) does not match. Shimotsuma (1963) suggested that *P. fistulosus* should be classified with *Cucumis melo* as both the species carries same monoploid chromosome number (n=12). The taxonomic classification of *P. fistulosus* is still not complete and it was suggested that this genus should be considered distinct and placed differently (Pangalo, 1938).

In order to confirm the genetic relationship among *P. fistulosus*, *C. lanatus* and *C. melo*. Isozyme and molecular analysis were carried out among the Cucurbits including *B. hispida* (Levi *et al.*, 2005). Isozyme analysis (Glutamate oxaloacetate transaminase and Peroxidase) indicated no similarity between *P. fistulosus*, *C. melo* and *Citrullus lanatus* subsp. *vulgaris* (Sujata and Seshadri 1989), supporting the suggestion given by Panglao (1983) that it should be placed or classified as distinct genus. DNA marker set (EST-SSR/PCR and SRAP), confirmed that *B. hispida* is more closely related to *P. fistulosus* than to *C. lanatus* or *C. melo* (Levi *et al.,* 2009).The pollen morphological studies among *B. hispida* and *P. fistulosus* shared the same morphological features and this is in agreement with the phylogenetic analysis carried out by DNA markers.

B. hispida and *Praecitrullus fistulosus* share the same chromosome number (2n=24). Several researchers and studies have supported the idea that these two species

are similar at genomic level and *P. fistulosus* genus should be involved in genus *Benincasa hispida.* Still, one cannot totally rely on molecular studies as DNA markers because it sometimes failed to show the clear distinctness. DNA barcoding could solve this problem at species level as it utilizes locus specific marker.

The DNA barcoding is a technique which utilizes DNA sequences for species identification, this speculation is fuelled by accurate species identification without knowledge of taxonomic expertise (Waugh, 2007).

5.1 Morphological characterization

The morphological characters were recorded using the descriptor of ashgourd developed by NBPGR, New Delhi. In this study the ashgourd and tinda genotypes showed monoecious plant characteristics, leaf shape of both the genotypes is cordate with with wide variation in shape. The ashgourd is a large viny creeper having large, dark yellow flowers with gradiation of fruit's shape and size and even leaf shape. In case of ashgourd, on maturation fruits form a coating of ash. In India, a wide range of variability is available for different component characters in ash gourd (Mandal *et al.*, 2002).

The twenty-six genotypes of ashgourd collected from different regions of country showed wide morphological variation in terms of vegetative characters, leaf size, leaf pubescence, plant habit, fruit shape and size, seed lustre. Gangopadhay (1983), Lovely (2001) and Reshmi (2004) also reported distinct variation in morphological characters of ashgourd landraces in terms of vegetative, flower, fruits and seed characters.

Plant growth habit is observed as long viny, medium viny and short viny. Menon (1998), Latif *et al.*, (2007) and Mohan *et al.*, (2012) studied plant growth habits in different ashgourd genotypes and observed similar variations. A higher percentage of genotypes were observed as long viny, long viny genotypes had smaller fruit size with less pubescence observed in stem and leaves. The leaf shape was also small. Medium viny growth was observed in smaller number of genotypes, they had heavy fruits with club shape. They had dense pubescence in leaves and stem. Short viny growth is observed in genotypes which were collected from northeast region. They had more dense pubescence in leaves and stem, female flowers are medium in size and unique in shape.

In ashgourd, short vine length (bushy type) with greater flesh thickness and fruit yield are considered as desirable characters (Sanwal *et al.*, 2011). Menon (1998) also observed different plant growth habits in different ashgourd genotypes.

Stem pubescence is observed as a varying trait in ashgourd varieties, seven different kinds of stem pubescence like hispid, tomentose, downy, puberulent, pilose, villous and hirsute were observed in twenty-six ashgourd genotypes. Till now, no such kind of stem pubescence variation has been reported in ashgourd genotypes.

Leaf margin was observed as serrated, deep serrated and double serrated. Menon (1998) and Reshmi (2004) also observed similar variations in leaf margin in ashgourd genotypes. The genotypes with small sized leaves had deep serration which was observed in all the accessions from northeast. Double serrated margin was observed in genotypes with large leaves and serrated leaf margins were observed in medium sized leaf.

Leaf shape was measured as from varying from small, medium and large in different genotypes varying from 5.0cm to 11.25cm length. Gangopadhay (1983), Menon (1998), Lovely (2001) and Reshmi (2004) observed such variation in leaf sizes in ashgourd genotypes.

Leaf pubescence nature observed were soft, intermediate, hispid and puberulent such variations in leaf pubescence nature is not yet recorded in any other genotypes of ashgourd.

Leaf pubescence density was observed were sparse, intermediate and dense in twenty-six ashgourd genotypes. No such kind of variation in leaf pubescence density has been recorded so far in ashgourd genotypes.

Peduncle attachment was observed in two different forms soft-flared, hard-flared. 80 per cent of genotypes had hard and flared kind of peduncle attachment. No such observations have been reported in ashgourd genotypes so far.

Peduncle length variation was recorded and observed in all of the genotypes which significantly vary among the ashgourd genotypes.

A wide variability for fruit shape was observed in the twenty-six ashgourd genotypes. The fruits were classified as club shaped, oblong, cylindrical and ridged. Majority of ashgourd genotypes had club shaped fruits, cylindrical, oblong fruits were also observed in ashgourd genotypes, but ridged shape fruits were observed only in northeast accessions. Gangopadhay (1983), Menon (1998), Lovely (2001) and Reshmi (2004) also reported wide variability in fruit shape in ashgourd. Pandey *et al.*, (2008) observed wide variation in fruit shape and size in collected germplasms of ashgourd.

Blossom end fruit shape and stem end fruit shape were recorded as round, depressed and flattened. Menon (1998) also observed and recorded variation in fruit end shapes in ashgourd genotypes.

Fruit length was minimum in Thanniam (4.6cm) and maximum in West Eleri (50.4cm) variety. Such variation in fruit length was reported by Hamid *et al.*, (1989), Menon (1998), Lovely (2001) and Reshmi (2004), Latif *et al.*, (2007) and Mohan *et al.*, (2012).

Fruit width varied among the different genotypes. The genotype Thanniam had minimum fruit width (10.3 cm), while maximum fruit width was observed in genotype West Eleri (60.1cm). Hamid *et al.*, (1989), Menon (1998), Lovely (2001) Reshmi (2004) and Mohan *et al.*, (2012) had also recorded such variations in fruit width.

Fruit weight is a significant character for characterizing genotypes. A wide range of fruit weight observed in the ashgourd genotypes. The minimum weight of fruit was recorded in genotype Thanniam (0.191kg) and maximum weight was recorded for genotype West Eleri (12.0 kg). Such variations were also reported by Gangopadhay (1983), Hamid *et al.*, (1989), Menon (1998), Lovely (2001) and Reshmi (2004), Latif *et al.*, (2007) and Mohan *et al.*, (2012). Sanwal *et al.*, (2011) reported fruit weight in ashgourd genotype as 24.84 kg.

Flower size also shows wide variation. Such kind of variations in flower size was earlier reported by Menon (1998) and Reshmi (2004).

The node number at which first female flower appeared was recorded in twenty-six ashgourd genotypes and five genotypes of tinda. The northeast collection had

recorded lowest node of first female flower appearance, while in tinda, Mahyco genotype recorded lowest node at number one for its first female flower appearance. Hamid *et al.*, (1998), Menon (1998), Lovely (2001), Reshmi (2004) and Latif *et al.*, (2007) observed and recorded the lowest node number at which female flower appears.In, cucurbitaceous crops it is desirable to have female flowers coming before male flowers because it can indicate an early-maturing genotypes (Sanwal *et al.*, 2011).

Days to 50 per cent flowering is observed and recorded in ashgourd and tinda genotypes. 22-30 count of days took place in twenty-six ashgourd genotypes, the northeast collection IC 0596986 had 50% per cent of flowering after 22 days of planting. The least count of 24 days for 50% flowering are recorded by Menon (1998), Lovely (2001) and Reshmi (2004).

Petiole length is recorded and observed in all genotypes of ashgourd and tinda.

Fruit skin colour varies in different genotypes from light green colour to dark green colour. 90 per cent of ashgourd varities had dark green coloured fruits. Tinda genotypes had light green coloured fruit skin.

Fruit pubescence was absent in all the 26 ashgourd genotypes at mature stage. In tinda genotypes sparse kind of pubescence was observed at mature stage.

Fruit ridges were superficial in all 25 varities of ashgourd. Only the genotype IC 0596989 had deep grooved ridges. Menon (1998) observed and recorded fruit ridge shape.

Among the twenty-six ashgourd genotypes only single collection (IC 0596989) from northeast had ridged fruit. In tinda genotype fruits no ridges were observed.

Flesh texture was observed as spongy in ashgourd genotypes. All the tinda genotypes smooth fleshed. Menon (1998) observed and recorded flesh texture of ashgourd genotypes.

The flesh colour of all ashgourd genotypes were creamish. Tinda genotypes had white flesh colour.

Seed lustre is observed as glossy in majority of ashgourd genotypes. Gangopadhay (1983), Menon (1998), Lovely (2001) and Reshmi (2004) observed seed lustreness in ashgourd genotypes. Tinda had matt seed lustre.

Seediness was high in majority of ashgourd genotypes, low seediness is observed as desirable character in ashgourd. Gangopadhay (1983), Menon (1998), Lovely (2001) and Reshmi (2004) observed high seediness in different ashgourd genotypes.

Number of primary branches is a significant characteristic for differentiating ashgourd genotypes. Majority of ashgourd genotypes had three primary branches, four, five, two and seven primary branches are also observed in ashgourd genotypes. Such variation in number of primary branches are reported by Reshmi (2004) and Mohan *et al.*, (2012) in different ashgourd genotypes.

A total count of 44-48 days was recorded for its first harvest of fruits in ashgourd genotypes. Only one genotype counted 44 days for its first harvest, in other ashgourd genotypes harvest period counted as 45-48 days. Such observations for first fruit harvest day are recorded by Gangopadhay (1983), Menon (1998), Lovely (2001), Reshmi (2004), Latif *et al.*, (2007) in different ashgourd genotypes.

For the days to last fruit harvest a total of 46-60 count of days took place. Only three genotypes took 46 days for its last harvest and the genotype KAU local took maximum period of 60 days. Such observations for days count for last harvest of fruits were also recorded by Gangopadhay (1983), Menon (1998), Lovely (2001), Reshmi (2004) Latif *et al.*, (2007) in different genotypes.

Number of fruits per plant counted from 2 per plant to 10 fruits per plant. Majority of ashgourd genotypes had 5 number of fruits per plant. Such kind of wide variation in number of fruits per plant was observed and recorded by Menon (1998), Lovely (2001) in different ashgourd genotypes.

The stem shape was observed as angular in all twenty-six ashgourd genotypes and five tinda genotypes. Lovely (2001) also reported as angular stem shape in other ashgourd genotypes.

The tendril, tendril type, tendril branching is present in all twenty-six genotypes in ashgourd and five genotypes of tinda. Gangopadhay (1983), Menon (1998), Lovely (2001) and Reshmi (2004) also reported presence of tendril, tendril type and branching in different ashgourd genotypes.

Dahiya *et al.*, (2001), Samadia (2007), Tyagi *et al.*, (2012), Munawar *et al.*, (2015), Khan (2016) and Garg *et al.*, (2017) observed and recorded the parameters like leaf area, stem diameter, fruit weight, fruit length, fruit diameter, number of vines per plant, number of fruits per plant, yield per plant in tinda genotypes.

5.2 Cross compatibility

In ashgourd, majority of the morphological characters are greatly influenced by environmental factors at the development stage of plant and thus assessment of genetic diversity in ashgourd based on phenotypic observations alone has limitation (Pandey *et al.*, 2008).

Pandey *et al.*, (2008) reported quite big diversity on the basis of molecular marker studies. Sureja *et al.*, (2006) also reported diversity in ashgourd germplasm.

The ashgourd accessions are collected from different regions of country including Kerala, Maharashtra and North-East region. All the twenty-six genotypes showed variation in vegetative characters like plant habit, fruit shape and size, leaf size, leaf pubescence, stem pubescence, number of primary branches but they were crossable in all directions. Cross-pollination was carried out with KAU local as female cultivar and twenty-five genotypes as male parents, and crosses were successful. Reciprocal crosses were also successful. The cross ability among different accessions of ashgourd is also reported by Gowda (1983), Sureja (2006), Latif *et al.*, (2007) and Verma *et al.*, (2007). Ashgourd genotypes are crossable in all directions and all the crosses can be done successfully, thus it is confirmed that they are crossable (Latif *et al.*, 2007). There is no such report of crossing of *Praecitrullus fistulousus* species with *Benincasa hispida* species.

5.2 Selection of genotypes for barcoding purpose

The nine ashgourd genotypes and three tinda genotypes were selected based on morphological data generated Euclidean distance. The Euclidean distance formed nine clusters for twenty-six ashgourd genotypes and single cluster for all the tinda genotypes. For ashgourd genotypes, one representative from each cluster was selected based on their unique fruit shape. While all the tinda genotypes formed a different cluster, three genotypes were selected at random. Thus, nine genotypes of ashgourd and three genotypes of tinda were finalized for barcoding.

5.3 DNA barcoding in plants

The two species with their 12 genotypes were characterized using the *matK* and *ITS* loci. The genetic assessment has been done with the reported barcode primers. There are several reported pair primers available for barcoding of land plants. Unfortunately, when only a single pair is used a very few loci are variable to distinguish plant species, a very few. This idea lead to the use of several combinations of loci, such as *matK+ atpF-H + trnH-psbA*, *rpoC1+ rpoB + matK*, *rbcL + trnH-psbA*, *rpoC1 + matK + trnH- psbA* (Yoo *et al.*, 2006; Chase *et al.*, 2007; Kress and Erickson, 2007; CBOL Plant Working Group, 2009). Among all of them, *matK* barcode is one of the promising candidates for plant barcoding work (Chase *et al.*, 2005).

5.3.1 DNA isolation

The total genomic DNA was isolated from each genotype using standardized DNA isolation protocol. The genotypes were characterized using the two loci. One is chloroplast-based marker *matK* and another is *ITS2* locus. The protocol for isolating DNA for each locus is available, but total genomic DNA isolation is carried out by CTAB based protocol (Rogers and Benedich, 1994). To study the relationships among various plant species DNA sequencing system is utilized by various plant systematists. To obtain accurate data from sequence information, it is necessary to isolate the DNA free from many contaminants found in plant cells. Many of the DNA isolation protocol yields actually greater amount of RNA, especially 18S and 25S rRNA. To remove RNA,

RNAse treatment is generally accepted. The contaminating RNA which was precipitate with DNA may cause many problems like non-reproducibility and suppression of PCR amplification (Jobes *et al.*, 1995).

5.4 DNA barcoding

The principal element of DNA barcoding is: 1) the ability to assign unknown sample to a known species 2) ability to detect previouslyunstudied samples (Meyer and Paulay, 2005).

An ideal DNA barcoding region should be short enough to get amplified from degraded DNA and analyzed *via* single pass sequencing. The accuracy of barcodes depends upon the extent of separation between intraspecific marker and interspecific divergence in selected marker. More the overlap between genetic variation within species and divergence separating sister species, the less effective the barcode. The accuracy is critically dependent upon species delineation. An additional trait which is desirable for a potential barcode is having a reading frame so that the presence of nonsense substitution could be used as a criterion to evaluate at what extend sequencing reactions have been done well. It is viewed as desirable trait for universal plant barcoding primers, so that alignment could be easily accomplished, thus making the selected markers useful for studies of phylogenetic and molecular evolutionary relationship. In case of non-coding region (intron or intergenic spacers) to represent as viable alternative, it should have 1) universal pair primer and standardized PCR protocol 2) high variation than coding regions consistently have and a 3) an easy, non-complicated pattern of molecular evolution (Chase *et al.*, 2007).

The selection of DNA barcode for land plants are more complicated. This is due to naturally available differences in observed nucleotide substitution rates at almost all the barcode loci proposed for land plants in comparison to *COI* in animals. It may also be due to ambiguity in the selection criteria used for choosing a barcode primer. The *COI* has been excluded as a candidate plant barcode, due to the limitations of mitochondrial gene, thus it provides a relatively low standard in comparison with COI barcode. In plants, nearly all chloroplast and nuclear markers have been proven more diagnostic than *COI*.

The gene *CoxI* (or COI) is selected as a standard barcode for animals by zoologist, it is clear that an alternative solution should be required to distinguish plants prior to animal barcodes. The equivalent and closest source of a plant barcoding region is plastid genome. The plastid genome shares many attributable features with animal mitochondrial DNA, such as conserved gene order and high copy number enabling easy retrieval of DNA for sequencing (Chase *et al.*, 2007). However, one of the problems with plastid DNA is generally slow rate of evolution and it is a challenge to find a plastid region that is sufficiently variable for DNA barcoding. A suitable region should have ideally show enough variation within to discriminate among species. However, this is a significant problem, finding a marker for which primer binding sites are conserved but shows high levels of variability across all groups of land plants. If markers have highly conserved primer binding sites, they tend to be internally more conservative (Chase *et al.*, 2007).

matK being one of the promising candidate regions for phylogenetic reconstruction among early diverging plants. It is the plastid gene region trnK$_{UUU}$–*matK* that is successfully used among various angiosperm lineages. This region is present universally in land plants, only few exceptions of a secondary loss and reorganizations are known for this. The genic region in this coding for Lysine transfer RNA, is divided into two exons, which is separated by a group II intron. The *matK* region is the only intact ORF within plastid group II introns. It has been speculated, to play an essential role in RNA processing, acting as a putative general maturase for plastid introns, although its exact function remains unclear or uncertain. In case of Embryophytes of synapomorphy, an open residing frame which is inserted into domain IV of the intron which is commonly known as *matK*. An exceptional feature of *matK* coding region that it shows equal distribution of first, second and third codon position. The *matK* gene evolves more rapidly in comparison to other plastid genes, despite having functional and underlying transcription constraints. With the combination of fast evolving nature and constrained trnK$_{UUU}$ intron region it acts as a suitable marker to address both low and deep level phylogenetic queries. The entire region ranges from 2.2kb to 2.6kb approximately in size. The *matK* is a fast-evolving region which provides highest phylogenetic structure with a desired phylogenetic information at deeper nodes. (Wickle and Quandt, 2009).

Apart from this three regions of plastid DNA, the more variable and most widely used region that has been proposed as a barcoding region is (*ITS*) internal transcribe spacer portions of nuclear ribosomal DNA or nrITS. *ITS* has long record of use in most groups of flowering plants and performed well as phylogenetic marker. The result obtained from ITS usage is nearly found similar to that of plastid DNA, but have 3-4 times more variable sites and evolve up to four times more rapidly. In nr*ITS*, presence of multiple and divergent copies makes it unacceptable as a standard barcode in all land plants. There are still several groups of angiosperms for which nrITS make a valuable contribution as a "local barcode", in case when low levels of plastid DNA variation is encountered. The problem related to its molecular evolution in many groups makes it undesirable (Chase et al., 2007).

The internal transcribe spacer (ITS) region of 18S-5.8S-26S nuclear ribosomal was extensively around the globe and has been first utilized a decade ago by Baldwin et al., (1995). The ITS based phylogenetic methodology has been dominant in plant molecular phylogenetic analysis apparently due to two reasons:1) advantageous properties for phylogenetic analysis and 2) a powerful bandwagon effect. Baldwin et al., (1995) have reported of ITS loci as: 1) Biparental inheritance: The ITS sequence is bi-parentally inherited since 18S-26S rDNA arrays reside in nuclear genome. Thus, distinguished from cpDNA loci. This property is determined to be useful for revealing past cases of reticulation, parentage of polyploids and hybrid speciation.2) Universality: The ITS sequence data can be readily obtained from perhaps any other nuclear marker. White *et al.*, (1990) also described a set of primers which is useful for amplifying *ITS* sequences from most of the fungal and phyla. 3) Simplicity: The nuclear DNA size is in range of 10 kb. This tandemly repeated sequence is present at one or more chromosomal loci per haploid chromosome complement. It is easy to isolate the gene, because there are hundreds to thousands of nuclear rDNA repeats in plant genome. 4) Intragenomic uniformity: Concerted evolution occurs when there is sequence difference among repeated copies in genome, accumulating their own distinct mutation and becomes homogenized to same sequence type by mechanism of high-frequency crossing over or gene conversion. On completion of this process, it eliminates sequence variation within genome and potentially mixed up variation, leaving species and clade specific character-

state changed to form phylogenetic construction effortful. 5) Intergenomic variability: *ITS* sequence variation levels for phylogenetic interference are suitable at specific, generic and family levels. The variation at hierarchical levels attribute mostly to nucleotide polymorphisms or indels (insertion or deletion) which is common. 6) Low function constraints: The *ITS* sequence is removed during transcription via splicing, it is subject to reasonably mild constraints which offers a preponderance of nucleotide sites that evolve neutrally but essentially. *ITS* is related to specific cleavage of primary transcript within *ITS1* and *ITS2* during maturation of small subunit (SSU) 5.8S and large subunit (LSU) ribosomal RNAs. The splicing and maturation process depend upon secondary structure of *ITS*, which implies some degree of conservation at sequence or structure level (Alvarez and Wendel, 2003).

5.5.1 PCR product sequencing

A total of 9 sequences of *Benincasa hispida* species and 3 sequences of *P. fistulosus* species were amplified with selected primers and sequenced by outsourcing (Sci Genome Lab. Pvt. Ltd., Cochin). The average length of sequence for *matK* sequence is 875 bp and for *ITS2*is 454 bp which show good sequencing process.

The recent advancement in genome sequencing through development of second and third generation technologies provide opportunities for rapid identification of candidate genes through bioinformatics analysis.

The highthroughput sequencing termed as next or second-generation sequencing platform able to produce highthroughput sequence data is advanced by increasingly high throughput technology. Second generation sequencing produce large amount of short DNA sequences of 25 to 400 bp. The first approach to second generation sequencing was pyrosequencing, commercialized by Roche (Basel, Switzerland) as the GS20. The current Roche 454 GS FLX Titanium system produces read lengths on average of 300-500 bp, produces over 400 Mbp of sequence with a single read accuracy of >99.5%. The two others advanced hightroughput sequencing platforms are the SOLiD system from Applied Biosystems (ABI) (Carlsbald, California, USA), the AB SOLiD System (currently version 3) used for tag-based applications such as gene expression and ChIPSeq. The AB SOLiD system used predominantly used for re-sequencing where identification and

removal of errors are possible when compared with references and Solexa Genome Analyser technology commercialized by Illummina (San Diego, California, USA). The IlluminaSolexa Genome Analyzer (currently the GAllx) uses reversible terminator chemistry to generate up to 50 Gbp of data per run with read lengths over 100 bp.

The different kinds of technologies are currently brought to market by several companies termed as third generation sequencing taking DNA sequencing production to further level of scale and for reducing the costs. The first of the technology that came to market uses single-molecule sequencing commercialized by Helicos Bioscience (Cambridge, MA, USA) termed as true single-molecule sequencing (tSMS). This approach differs from existing second-generation systems where sequencing is possible without requirement of DNA amplification, which has been used to sequence the genome of virus M13. The Pacific bioscience developed a single-molecule real-time sequencing system which produces several Gbp of long reads (>1 kbp). The third-generation system is likely to be developed as more advanced sequencing technology in relatively near future (Edward and Batley, 2010).

5.5.2 Species identification on the basis of Euclidean distance

The Euclidean distance similarity is based on morphological observation recorded according to the descriptors for Ashgourd developed by NBPGR, New Delhi. The *Benincasa hispida* genotypes shows the gradation from leaf and stem pubescence density to fruit size and shape. The observation is recorded from the initial stage of plant growth up to harvesting and separation of seeds. The Euclidean distance similarity based on description points like leaf pubescence, leaf and stem pubescence density, fruit shape, rib shape, seediness and seed texture. On the basis of Euclidean similarity, the genotypes are grouped in different clusters, from each cluster a single genotype is picked and further sent for sequencing purpose. The selected nine ashgourd genotypes are Kottamuri, Vellarikundu, West Eleri, Areeparambu, Nashik, IC 0565989, IC 0565992, Elavancheri, KAU Local (female parent) and three of tinda named as Nirmal Tinda, Ankur Tinda and Komal Tinda genotypes.

5.5.3 Species identification based on phylogenetic tree analysis

The genetic similarity was determined for all the genotypes of 9 *Benincasa* and 3 *Praecitrullus* based on *matK* and *ITS2* sequence. The phenogram based on 12 selected genotypes was grouped into III clustered.

For *matK* sequence, all the three genotypes of tinda (Nirmal tinda, Komal tinda and Ankur tinda) are grouped with ashgourd (Kottamuri, West Eleri, Vellarikundu, Nashik, Elavancheri, IC 0596989 and IC 0596992) in Cluster I. In Cluster I, a single genotype of ashgourd (Areeparambu) is present and Cluster III consists of (ashgourd) female parent cultivar KAU local variety.

For *ITS* sequence, the results are comparatively similar in case of tinda genotypes. Cluster I consist of three tinda genotypes (Nirmal tinda, Ankur tinda and Komal tinda) and ashgourd genotypes (Kottamuri, Nashik, Elavancheri, IC 0596992, West Eleri, Vellarikundu). Cluster II consists of single genotype IC 0596989 and Cluster III carries single genotype (ashgourd) KAU Local variety (female parent cultivar).

The phylogenetic similarity analysis is matched with the results estimated by researchers Koycan *et al.*, (2007) estimated close genetic similarity among *P. fistulosus* and *B. hispida* on the basis of chloroplasts DNA sequences. Schaefer *et al.*, (2009) utilizes *ITS2* sequences to reveal the genetic similarity between *P. fistulosus* and *B. hispida*.

5.5.4 Species identification based on phylogenetic generated by MEGA software

The phylogenetic tree for *matK* and *ITS2* sequences were obtained from MEGA software. For *matK* sequences, three clusters are formed, Cluster I form two groups first group represents accessions Kottamuri, West Eleri, IC 0596992 with a common bootstrap value 100 and group two represents Ankur tinda and Komal tinda accessions. Cluster II represents accessions Nashik, Elavancheri, IC 0596989 and Nirmal tinda. Cluster III represents accessions Vellarikundu, Areeparambu and KAU local.

For *ITS2* loci, cluster I represents accessions West Eleri, Elavancheri, Ankurtinda, IC 0596992, IC 0596989, KAU local, Nirmaltinda, Komaltinda, Nashik, Vellarikundu and Areeparambu. Cluster II represents single accession Kottamuri.

The phylogenetic analysis using MEGA software is applied in Indian *Berberis* L. and two other genera *Ficus* and *Gossypium*. The NJ, MP and UPGMA methods were used for each single locus and combination of locus *ITS*, *matK* and *trnH-psbA*. UPGMA, NJ and NP method is applied for phylogenetic analysis, where NJ method proved to be better for species recovery in case of combination of some locus (Roy *et al.*, 2010).

5.5.5 Genetic divergence within and between species

The inter specific distances were calculated using K2P (Kimura 2 parameter) available in the MEGA software.

For *matK* loci, the range of pairwise dissimilarity was between 0.899 and 0.833. The minimum pairwise distance among ashgourd accession is observed in West Eleri at 0.039.

For *ITS2* loci, the range of pairwise dissimilarity was between 0.901 and 0.820. The maximum similarity is observed in the range of 0.928 in ashgourd accessions. The minimum dissimilarity is observed at 0.901 among the ashgourd accession and tinda accessions.

Roy *et al.*, (2010) utilized the locus *ITS* and *trnH-psbA* to determine the genetic the interspecific and intraspecific divergences in Indian *Berberis* L. and two other genera *Ficus* and *Gossypium*.

5.5.6 Barcoding gap and efficiency assessment

The DNA barcode should be robust and have separate and non-overlapping genetic variations among the species. It will be treated as a useful barcode primer and successful discrimination, if it is showing some changes in sequences carried by the individual species.

A total of 9 sequences were aligned for both the locus *matK* and *ITS2* and the nucleotide change has been observed. As the sequences from each genotype were uniqued except with some conserved region in tinda genotypes, the *matK* and *ITS2* loci were found to be a potentially successful for barcoding. The *matK* locus is previously reported as the most successful locus in much plant (Jing *et al.*, 2011).

5.5.7 Identification of conserved region

The identification of conserved region is carried out by Clustal W software. The size of amplified product for *matK* loci was found to be mainly 1000 and *ITS2* was loci found to be mainly 799 bp, which is a reasonably good length to identify the barcoding gaps in the genus *Benincasa* and *Praecitrullus* both. The sequences obtained from both *matK* and *ITS2* loci were aligned separately. *ITS2* and *matK* both loci successfully differentiated the *Benincasa hispida* species from *Praecitrullus fistulosus* species.

5.5.8 Utilization of developed barcodes in *Benincasa* and *Praecitrullus* species

The species-specific barcodes generated from this study could be immediately used for species allocation of *Benincasa hispida* and *Praecitrullus fistulosus* species having a taxonomical controversy (Levi, 2010). On the basis of BLAST, the basic sequence is generated and deposited in NCBI. The barcoding gaps could be determined in the sample and from this, the species allocation could be easily done.

5.5.9 Future line of work

The barcodes *matK* and *ITS2* loci, successfully worked in both of the species. The barcodes can be used to work in different cucurbit species, as the results obtained from the above-mentioned barcode gives an unambiguous and precise output. In the present study, the barcode gap generated by *matK* loci gives a perfect picture of difference among both the species compared to *ITS2* locus, as the constant nucleotide substitution was observed only in *Praecitrullus* species, this concluded that barcode gap only exists in tinda genotypes and making the *Praecitrullus* species totally different from *Benincasa* species. Further study of *Benincasa* genotypes Kottamuri, Vellarikundu, West Eleri, Areeparambu, Nashik, Elavancheri, IC 0596989 and IC 0596992 should be done, the barcode gap exists within the species, showing similarity between them, although they had different shape and represent different locality. Other locus study should be done than *matK* and *ITS2* locus for species discrimination in cucurbits. Thus, morphological parameters studies and barcode gap finding suggests that both the species are totally different from each other.

Summary

6. SUMMARY

The study entitled "DNA barcoding in genera *Benincasa* and *Praecitrullus*" was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Vellanikkara, during the period 2015-2017. The present study was undertaken with objective to assess genetic diversity and interspecies relationship among *Benincasa* and *Praecitrullus* genotypes and to check the species status of *Benincasa* using barcodes. In India, *Benincasa hispida* species exhibits variation in its morphological characters. The previous works reported genetic similarity between the two species based on molecular marker analysis and pollen morphology studies. Although, very limited molecular work is done to confirm the genetic relatedness among the two species. Twenty-six genotypes of ashgourd collected from different parts of the country and five genotypes of tinda from different regions of India were included in the present study

The salient findings of this study are stated below:

1. The Euclidean distances derived from morphological characters based on the descriptor developed by NBPGR, New Delhi for both the species shows morphological dissimilarity among the species. All the tinda formed a separate distinct cluster showing their dissimilarity with ashgourd genotypes.
2. A total of 9 genotype of ashgourd and 3 genotype of tinda are selected on the basis of clusters formed through Euclidean distances for barcoding purpose. Selection of ashgourd genotype is based on formation of 9 different cluster, each ashgourd genotype is selected representing each cluster, with uniqueness in their fruit shape. A random of 3 tinda genotypes were selected as all tinda genotypes together formed a single cluster.
3. The protocol for genomic DNA isolation was standarised. The protocol suggested by Rogers and Benedich (1994) with slight modification (increase in CTAB percentage, double wash of C: I) was found to be the most appropriate for isolation of DNA from different ashgourd and tinda genotypes. The RNA contamination was completely removed through RNase treatment.
4. The quality and quantity of DNA was analysed by NanoDrop® ND-1000 spectrophotometer. The absorbance ratio ranged from 1.81-1.94, which indicated

good quality DNA and the recovery was high. This DNA was suitable for PCR amplification analysis.

5. Primer sets for 2 barcoding loci *matK* and *ITS2* loci were used in this study. Both the locus are compatible to generate the bands in both the species.
6. A total of 9 sequences of *Benincasa* species and 3 *Praecitrullus* species have been submitted to the NCBI GenBank.
7. The 12 sequences for each locus separately were used to construct the phylogenetic tree using the software ClustalW Omega. The phylogenetic tree generated clearly showed the interspecies diversity in both species.
8. The genetic distances was computed for all combinations of 12 genotypes belonging to 9 *Benincasa* and 3 *Praecitrullus* spp. based on K2P model using MEGA 7 software which was recommended by the Consortium of Barcode of Life.
9. The species identification success rate was calculated using BLAST using *matK* and *ITS2* sequences, in which *matK* yield up to 95 per cent similarity and *ITS2* yield 100 per cent similarity at species level with NCBI accessions.
10. The conserved sequences in the generated base sequence was identified using the software ClustalW omega.
11. No subspecies of *Benincasa* was reported.
12. The genotypes Kottamuri, West Eleri, Vellarikundu, Areeparambu, Elavancheri, IC 0596989, IC 0596992, KAU local and Nashik can be considered as unique cultivars as they represents unique shape of fruits.

Future line of work; other than *matK* and *ITS2 l*oci, another locus should also be studied for species discrimination in cucurbits. The ashgourd genotypes Nashik, IC 0596989, IC 0596992 and Elavancheri generated distinct barcodes within *Benincasa* species, and so has to be studied in detail.

7. REFERENCES

Álvarez, I. and Wendel, J.F. 2003. Ribosomal *ITS* sequences and plant phylogenetic inference. *Mol. Phylogenetics Evol.* 29(3): 417-434.

Ausubel, F.M. 1994. Preparation of genomic DNA from bacteria. *Current protocols in molecular biology. Wiley Interscience*, 2-4p.

Baldwin, B.G., Sanderson, M.J., Porter, M., Wojciechowski, M.F., Campbell, C.S., and Donoghue, M.J. 1995. The Region of Nuclear Ribosomal DNA: A valuable sources of evidence on Angiosperm phylogeny. *Ann. Missouri Bot. Garden*. 82(2): 247-277.

Bates and Robinson 1995. Appendix: an outline classification of the *Cucurbitaceae*. In: Bates, D.M., Robinson, R.W., and Jeffrey, C. (eds.), *Biology and Utilization of the Cucurbitaceae*. Comstock Publication Associates, Cornell University Press, Ithaca, New York, pp. 449-463.

Beevy, S.S. and Kuriachan, P. 1996. Chromosome numbers of south Indian *Cucurbitacea* and a note on the cytological evolution in the family. *J. Cytol. Genet.* 31(1): 65-71.

Bej, A.K., Mahbubani, M.H. and Atlas, R.M. 1991. Amplification of nucleic acids by polymerase chain reaction (PCR) and other methods and their applications. *Critical Rev. Biochem Mol. Biol.* 26(3-4): 301-334.

Bhat, Z.A., Dhillon, W.S., Rashid, R., Bhat, J.A., Dar, W.A., and Ganaie, M.Y. 2010. The role of molecular markers in improvement of fruit crops. *Notulae Sci. Biologicae* 2(2): 22-27.

Bretting, P.K., and Widrlechner, M.P. 1995. Genetic markers and horticultural germplasm management. *Hort. Sci.* 30(7): 1349-1356.

Burkill, I.H. 1935. *Dictionary of the Economic Products of the Malay Peninsula*. London: Crown Agents for the Colonies, 160p.

Caetano-Anolles, G. and Brant, B. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Nat. Biotechnol.* 9(6): 553-561.

Chase, M.W., Cowan, R.S., Hollingsworth, P.M., Petersen, G., Seberg, O., Jorgensen, T., Cameron, K.M., and Carine, M. 2007. A proposal for a standardized protocol to barcode all land plants. *Taxonomy.* 56: 295-299.

Chase, M.W., Salamin, N., Wilkinson, M., Dunwell, J.M., Kesanakurthi. R.P., Haidar, N., and Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Trans. R. Soc. Biol. Sci.* 360: 1889-1985.

Chauhan, D.V.S. 1989. *Vegetables Production in India*. Ramprasad and Sons, Agra, India, 152p.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Linchun, S., Yingjie, Z., Xinjye, M., Gao, T., Pang, X., Luo, K., Li, Y., Xiwen, L., Jia, X., Lin, Y., and Leone, C. 2010. Validation of the *ITS2* region as a novel DNA barcode for identifying medicinal plant species. *PloS ONE.* 5(1): 8613-8623.

Dahiya, M.S., Baswana, K.S., and Tehlan, S.K. 2001. Genetic variability studies in round melon (*Praecitrullus fistulosus*). *Haryana J. Hortic. Sci.* 30(1): 81-83.

Dane, F. and Lang, P. 2004. Sequence variation at cpDNA regions of watermelon and related wild species: implications for the evolution of *Citrullus* haplotypes. *Am. J. Bot.* 91(11): 1922-1929.

Das, B.K., Jena, R.C., and Samal, K.C. Optimization of DNA isolation and PCR protocol for RAPD analysis of banana/plantain (*Mussa* spp.). 2009. *Int. J. Agric. Sci.* 1(2): 21-25.

De Candolle, A. 1882. *Origin of Cultivated Plants*. Hafner Publishing Co., New York, 106p.

Decker-Walters, D., 1998. *Underexploited germplasm resources of the Cucurbitaceae. In Cucurbitaceae*. 98: 17-30p.

Decker-Walters, D.S. 1999. Cucurbits, Sanskrit, and the Indo-Aryas. *Econ. Bot.* 53(1): 98-112.

Echt, C.S., DeVerno, L.L., Anzidei, M., and Vendramin, G.G. 1998. Chloroplast microsatellites reveal population genetic diversity in red pine (*Pinus resinosa* Ait.). *Mol. Ecol.* 7(3): 307-316.

Edwards, D. and Batley, J. 2010. Plant genome sequencing: applications for crop improvement. *Plant Biotechnol. J.* 8(1): 2-9.

Ennos, R.A. 1999. Using organelle markers to elucidate the history, ecology and evolution of plant populations. *Mol. Syst. Plant. Evol.* 7: 1-19.

Erlich, H.A., Gelfand, D., and Sninsky, J.J. 1991. Recent advances in the polymerase chain reaction. *Science* 252(5013): 1643-1651.

Frézal, L. and Leblois, R. 2008. Four years of DNA barcoding: current advances and prospects. *Infection Genet. Evol.* 8(5): 727-736.

Gangopadhay, K.K., Kumar, G., Meena, B.L., and Bisht, I.S. 2008. Genetic diversity studies in ashgourd (*Benincasa hispida* (Thunb.) Cogn.) from Northern India. *J. Plant Genet. Resour.* 21(3): 206-212.

Gangopadhay, C. 1983. Biometric studies in ashgourd. Msc. (Plant Biotechnology) Thesis, Kerala Agricultural University, Thrissur, 150p.

Garg, N. 2017. Genetic diversity in round gourd [*Praecitrullus fistulosus* (Stocks) Pangalo.] accessions introduced from USDA for various qualitative and quantitative traits. *J. Crop Improv.* 31(6): 801-815.

Girme, A.R. 2014. DNA barcoding in momordica spp. M.Sc. (Plant Biotechology) Thesis, Kerala Agricultural University, Thrissur, 150p.

Gopalkrishnan, R. 1957. Medicinal ashgourd (A new variety). *Madras Agric. J.* 44: 345-346.

Gowda, R.V. (1983). Genetic Analysis of Yield and its Components in Ash Gourd [*Benincasa hispida* (Thunb.) Cogn.]. M.Sc. Thesis, University of Agricultural Sciences. Bangalore, India. 110 p.

Hamid, M.M., Sana, M. C., Begum, R. A., and Hussain, S.M.M. 1989. Physiso-morphology and yield of different ash gourd lines. *J. Agric. Sci.* 14: 51-55.

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

Hu, J.B., Li, J.W., Wang, L.J., Liu, L.J., and Si, S.W. 2011. Utilization of a set of high-polymorphism DAMD markers for genetic analysis of a cucumber germplasm collection. *Acta Physiol. Plant* 33(1): 227-231.

Hubert, N. and Hanner, R. 2006. DNA barcoding, species delineation and taxonomy: a historical perspective. *DNA barcodes* 3(1): 44-58.

Indira, P. and Peter, K.V. 1987. *Underexploited Tropical Vegetables*. Kerala Agricultural University, Vellanikkara, Kerala.115p.

Ivanova, N.V., Dewaard, J.R. and Hebert, P.D. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular ecology notes*. 6(4): 998-1002.

Janzen, D.H., Hallwachs, W., Blandin, P., Burns, J.M., Cadiou, J.M., Chacon, I., Dapkey, T., Deans, A.R., Epstein, M.E., Espinoza, B., and Franclemont, J.G. 2009. Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Mol. Ecol. Resour.* 9: 1-26.

Jeffrey, C. 1990. Systematics of the *Cucurbitaceae*: an overview. In: Bates, D.M., Robinson, R.W., and Jeffrey, C. (eds.), *Biology and utilization of the Cucurbitaceae*. Cornell University Press, Ithaca, 221p.

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

Jobes, D.V., Hurley, L.D., and Thein, L.B. 1995. Plant DNA isolation: a method to efficiently remove polyphenolics, polysaccharides and RNA. *Int. Assoc. Plant Taxon*. 44(3): 379-386.

Karaca, M. 2013. Isozymes as biochemical markers in plant genetics. *Int. J. Agriscience* 3(11): 851-861.

Karp, A.S., Kresovich, K.V., Bhatt, W.G., and Ayad, T.H. 1997. Molecular tools. In: *Plant Genetic Resources Conservation*: *A Guide to the Technologies*. IPGRI No. 2, International Plant Genetic Resources Institute, Rome, Italy, 47p.

Khan, M. 2016. Assesment of biochemical attributes of *Praecitrullus fistulosus* treated with mutagens. *J. Plant Breed. Genet.* 3(2): 39-47.

Koycan, A., Zhang, B.L., Schaefer, H., and Renner, S.S. 2007. A multi-locus chloroplast phylogeny for the *Cucurbitacea* and its implications for the character evolution and classification. *Mol. Phylogenet. Evol.* 44: 553-557.

Kress, W.J. and Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL*gene complements the non-coding *trnH-psbA* spacer region. *PLOS One* 2: 508-515.

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

Kumar, S., Singh, D.B., and Sinha, P.D. 2012. RAPD markers for identification for sex in pointed gourd (*Trichsanthes diocia* Roxb.) *Indian J. Biotechnol.* 11: 251-256.

Latif, M.A., Alam, M.M., Alam, M.K., and Rahman, M.M. 2007. Floral biology of different exotic and indigenous ash gourd genotypes (*Benincasa hispida*). *Int. J. Plant Breed. Genet.* 1(2): 67-74.

Levi, A., Harris, R.K., Wechter, P.W., Koushik, S.C., and Thies, J.A. 2010. DNA markers and pollen morphology reveal that *Praecitrullus fistulosus* is more closely related to *Benincasa hispida* than to *Citrullus* spp. *Genet. Resour. Crop Evol.* 57: 1191-1205.

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

Levi, A., Wechter, P., and Davis, A. 2009. EST-PCR markers representing watermelon fruit genes are polymorphic among watermelon heirloom cultivars sharing a narrow genetic base. *Plant. Genet. Res.* 7: 16-32.

Lovely, B. 2001. Evaluation of genetic divergence in ashgourd. M.Sc. (Ag.) Thesis, Kerala Agricultural University, Thrissur, 76p.

Mandal, J., Sirohi, P.S., and Behera, T. K. 2002. Genetic studies on flowering and fruit maturity in ash gourd [*Benincasa hispida* (Thunb.) Cogn.]. *Orissa J. Hort.* 30(1): 40-42.

Manikandan, M., Yassin, G.M., Kanthaswamy, V., and Kamaladevi, S. 2017. Correlation and path coefficient analysis in ash gourd [*Benincasa hispida* (Thunb.) Cogn.] for yield and yield attributing traits. *Chem. Sci. Rev. Lett.* 6(23): 1399-1403.

Marr, K.L., Xia, Y., and Bhattarai, N.K. 2007. Allozymic, morphological, phenological, linguistic, plant use and nutritional data of *Benincasa hispida* (Cucurbitaceae). *Econ. Bot.* 61: 44-59

Mas, G.J., Oliver, M., Paniagua, G.H., and Vicente, M.C. 2000. Comparing AFLP, RAPD and RFLP markers for measuring genetic diversity in melon. *Theor. Appl. Genet.* 101: 860-864.

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

Menon, M.P. 1998. Cataloguing and identification of promising ash gourd ecotypes in relation to season and maturity. M.Sc. (Hort.) Thesis, Kerala Agricultural University, Thrissur, 132p.

Meusel, I., Lestner, E., and Wilhelm, B.W. 1994. Chemistry and micro-morphology of compound epicuticular wax crystalloids (Strelitzia type). *Plant. Syst. Evol.* 7: 115-123.

Meyer, C. and Paulay, G. 2005. DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biol.* 3(12): 422.

Mohan, N.B., Madalageri, M.B., Mulge, R., Adiger, S., and Kencharaddi, H.G. 2012. Genetics of vine yield and yield contributing traits in ash gourd (wax gourd). *Plant Arch.* 12(2): 731-736.

Mohapatra, T., Sharma, R.P., and Chopra, V.L. 1992. Cloning and use of low copy sequence genomic DNA for RFLP analysis of somaclones in mustard (*Brassica juncea* (L.) Czern and Coss). *Curr. Sci.* 62(6): 482-484.

Morton, J.F. 1971. The wax gourd, a year-round Florida vegetable with unusual keeping quality. *Proc. Florida State Hortic. Soc.* 84: 104-109.

Munawar, M., Hammad, G., Nadeem, K., Raza, M.M. and Saleem, M. 2015. Assessment of genetic diversity in tinda gourd through multivariate analysis. *Int. J. Vegetable Sci.* 21(2): 157-166.

Nanda, S., Kar, B., Nayak, S., Jha, S., and Joshi, R.K. 2013. Development of an ISSR based STS marker for sex identification in pointed gourd (*Trichosanthes dioica* Roxb.). *Sci. Hortic.*150: 11-15.

Nayar, N.M. and More, T.A. 1998. *Curcurbits*. Oxford and IBH Publishing Co. Pvt. Ltd., New Delhi, 340pp.

Neuhaus, H. and Link, G. 1987. The chloroplast tRNA Lys (UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. *Curr. Genet*. 11: 251-257.

Newmaster, S.G., Fazekas. A.J., and Raghupaty, S. 2006. DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Botany* 84: 335-341.

Olmstead, R.G. and Palmer, J.D. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *Am. J. Bot.* 81(9): 1205-1224.

Pandey, A.K., Bhardawaj, D.R., Dubey, K.R., Singh, V., and Pandey, S. 2015. Botany, diversity, utilization and improvement of ash gourd (*Benincasa Hispida* Thunb Ex Murray Cogn.)-a review. *Ann. Hortic.* 8(1): 1-15.

Pandey, S., Kumar, S., Mishra, U., Rai, A., Singh, M., and Rai, M. 2008. Genetic diversity in Indian ashgourd (*Benincasa hispida*) accessions as revealed by quantitative traits and RAPD markers. *Sci. Hortic*. 118: 80-86.

Pangalo, K.I. 1938. Living ancestors of cultivated watermelon. *Acad. Sci. URSS* 10: 599-600

Parker, P.G., Snow, A.A., Schug, M.D., Botton, G.C., and Furest, P.A. 1998. What molecules can tell us about populations: choosing and using molecular markers. *Ecology* 79: 361-382.

Peter, K.V., Kumar, S., Raj, P.G.M., and Prasanna, K.P. 1991. Genetic diversity in cucurbits ashgourd. *Indian Hort*. 36-39.

Poleg, D.Y., Reis, N., Tzuri, G., and Katzir, N. 2001. Development and characterization of microsatellite markers in *Cucumis*. *Theor. Appl. Genet.* 102(1): 61-72.

Ramesh, M., Gayathri, V., Rao, A., Prabhakar, M.C., and Rao, C.S. 1989. Pharmacological actions of fruit juice of *Benincasa hispida*. *Fitoterpia* 60(3): 241-247.

Rashid, H.M., Khan, R.M.A., Yasmin, N., Ishtiaquie, S., and Chaki, K.A. 2016. Genetic diversity in snake gourd genotypes revealed by RAPD markers. *J. Agric. Technol.* 12(1): 125-144.

Reddy, U.B. 2009. A phylogenetic analysis of the *Cucurbitaceae*: evidences from *mat*K nucleotide sequence. *Inter. J. Bioinforma. Res.* 1(2): 47-53.

Reshmi, J. 2004. Characterization of landraces of ashgourd (*Benincasa hispida* (Thunb.) Cogn.). M.Sc. Thesis, Kerala Agricultural University, Vellanikkara, 97p.

Resmi, J. and Sreelathakumary, I. 2011. RAPD markers for genetic variability studies in ash gourd *Benincasa hispida* (Thunb.) Cogn. *J. Agric. Technol.* 7(4): 1097-1106.

Robinson, R.W. and Decker-Walters, D.S. 1999. *Cucurbits*. CAB International, Wallingford, Oxon (GB), 226p.

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

Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U.M., Chaudhary, L.B., Datt, B., Bag, S.K., Singh, P.K., Nair, N.K., and Husain, T. 2010. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian *Berberis* species. *PloS ONE.* 5(10): 13674.

Rubatazky, V.E. and Yamaguchi, M. 1999. *World Vegetables*. Chapman and Hall, New York, USA, 43p.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839): 487-491.

Samadia, K.D. 2007. Studies on genetic variability and scope of improvement in round melon under hot arid conditions. *Indian J. Hortic.* 64(1): 58-62.

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

Sankar, N.R., Devamma, M.N., Kumar, V.K., and Giridhar, D. 2011. First report of *Fusarium moniliforme* causing fruit rot of tinda (*Praecitrullus fistulosus*) in India. *New Disease Reports*. 24p.

Sanwal, S.K., Kozak, M., and Verma, M.R. 2011. Searching for the optimum multivariate profile of an ash gourd [*Benincasa hispida* (Thunb.) Cogn.] genotype. *Acta. Physiol. Plant* 33(2): 463-468.

Sax, K. 1923. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8: 552-560.

Schaal, B.A., Hayworth, D.A., Olsen, K.M., Rauscher, J.T., and Smith, W.A. 1998. Phylogeographic studies in plants: problems and prospects. *Mol. Ecol.* 7(4): 465-474.

Schaefer, H., Heibl, C. and Renner, S.S. 2009. Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1658), pp.843-851.

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

Sharma, K., Mishra, A.K., and Misra, R.K. 2008. A simple and efficient method for extraction of genomic DNA from tropical tuber crops. *Afr. J. Biotechnol.* 7(8): 1018-1022.

Shimotsuma, M. 1963.Cytogenetics and evolutionary studies in the genus *Citrullus*. *Seiken Ziho* 15: 23-24.

Sikdar, B., Bhattacharya, M., Mukherjee, A., Banerjee, A., Ghosh, E., Ghosh, B., and Roy, C.S. 2010. *Biologia Plant* 54(1): 135-140.

Singh, M., Kumar, S., Singh, A.K., Ram, D. and Kaloo, G. 2002. Female sex associated RAPD marker in pointed gourd (*Tricosanthes diocia* Roxb.). *Curr. Sci.* 82 (2): 131-132.

Singh, D.K. and Sarkar, M. 2014. *Biodiversity of cucurbitaceous vegetable crops in India*. Cucurbitaceae. 12: 64p.

Soller, M. and Beckmann, J.S. 1983. Genetic polymorphism in varietal identification and genetic improvement. *Theor. Appl. Genet.* 67(1): 25-33.

Srivastava, U.C. 1993. Exploration and evaluation of cucurbitaceous crops. In: Rana, R.S., Singh, B., Srivastava, U., Mathur, P.N., Rai, Singh, B.N., Kochhar, S., and Bisht, I.S. (eds.), *Exploration and Conservation of Plant Genetic Resources [Lecture notes]*. NBPGR, New Delhi, India. pp.57-67.

Staub, J.E., Robbins, M.D., Chung, S., and Lopez-Sese, A.I. 2002. Molecular methodologies for improved genetic diversity assessment in cucumber and melon. *Acta Horti.* 637: 41-47.

Sujata, V.S. and Seshadri, V.S. 1989. Taxonomic position of round melon (*Praecitrullus fistulosus*). *Cucurbit Genet. Coop. Rep.* 12: 86-88.

Sundararajan, S. and Muthukrishnan, C.R. 1982. A high yielding ash gourd. *Indian Hortic.* 27: 15-20.

Sureja, K.A., Sirohi, S.P., Behera, K.T., and Mohapatra, T. 2006. Molecular diversity and its relationship with hybrid performance and heterosis in ash gourd [*Benincasa hispida* (Thunb.) Cogn.]. *J. Hortic. Sci. Biotechnol.* 81(1): 33-38.

Tautz, D. 1989. Hyper variability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* 17(16): 6463-6471.

Trivedi, M.K., Branton, A., Trivedi, D., Nayak, G., Gangwar, M., and Jana, S. 2015. Morphological and molecular analysis using RAPD in biofield treated sponge and bitter gourd. *Am. J. Agric. For.* 3(6): 264-270.

Trivedi, R.N and Roy, R.P. 1969. Cytological studies in *Cucumis* and *Citrullus*. Department of Botany, Patna University.

Tyagi, N., Sharma, N.G., and Vikas, H. 2012. Medicinal value of *Praecitrullus fistulous*: an overview. *Int. J. Phama. Therap*. 3(2): 136-142.

Valentini, A., Pompanon, F., and Taberlet, P. 2009. DNA barcoding for ecologists. *Trends Ecol. Evol.* 24(2): 110-117.

Verma, V.S. and Behera, K.T. 2007. Heterosis studies in ash gourd [*Benincasa hispida* (Thunb.) Cogn.] for yield and related traits. *J. Trop. Agri*. 45(1-2): 51-54.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T.V.D., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23(21): 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.

Walters, T.W. and Decker-Walters, D.S. 1989. Systematic re-evaluation of *Benincasa hispida* (Cucurbitaceae). *Econ. Bot.* 43(2): 274-278.

Wang, C.X., Liu, C., Huang, L., Palme, J.B., Chen, H., Zhang, J.H., Cai, D., and Li, J.Q. 2014. *ITS1*: a DNA barcode better than *ITS2* in eukaryotes. *Mol. Ecol. Res*. 15(3): 573-586.

Wang, H., Yao, L., Cai, R., Pan, J., and Chen, X. 2012. Genetic relationship analysis of oil-bearing roses in China using *matK* sequences. *Sci. Hortic.* 137: 121-124.

Wang, X.C., Liu, C., Huang, L., Bengtsson-Palme, J., Chen, H., Zhang, J.H., Cai, D., and Li, J.Q. 2015. *ITS1*: a DNA barcode better than *ITS2* in eukaryotes. *Mol. Ecol. Res.* 15(3): 573-586.

Waugh, J. 2007. DNA barcoding in animal species: progress, potential and pitfalls. *Bio Essays* 29(2): 188-197.

Waugh, R. and Powell, W. 1992. Using RAPD markers for crop improvement. *Trends Biotechnol.* 10: 186-191.

Welsch, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 18(24): 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(4): 473-476.

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

Whistler, W.A. 1990. The other Polynesian Gourd. *Pac. Sci.* 44(2): 115-122.

White, T.J., Bruns, T., Lee, S.J.W.T., and Taylor, J.L. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: a Guide to Methods and Applications.* 18(1): 315-322.

Whitkus, R., Doebley, J., and Wendel, J.F. 1994. Nuclear DNA markers in systematics and evolution. In: Phillips, R.L. and Vasil, I.K. (eds.), *DNA-based Markers in Plants: Advances in Cellular and Molecular Biology of Plants*, Dordrecht, Netherland, pp.116-141.

Wicke, S. and Quandt, D. 2009. Universal primers for the amplification of the plastid trnK/matK region in land plants. *Anales del Jardín Botánico de Madrid* 66: 2.

Williams, J.G., 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(22): 6531-6535.

Witkowicz, J., Wochniak, E.A., and Przybecki, Z. 2003. AFLP marker polymorphism in cucumber (*Cucumis sativus* L.) Near isogenic lines differing in sex expression. *Cell Mol. Biol. Lett.* 8: 375-381.

Yoo, H.S., Eah., J., Kim, Y., Min, M., Paek, W.K., Lee, H, and Kim, C. 2006. DNA barcoding Korean birds. *Mol. Cells* 22: 323-327.

Yu. J., Xue. H., and Zhou, S. 2011. New universal *matK* primers for DNA barcoding angiosperms. *J. Syst. Evol.* 49: 176-181.

Zaini, N.A.M., Anwar, F., Hamid, A.A., and Saari, N. 2011. Kundur [*Benincasa hispida* (Thunb.) Cogn.]: A potential source for valuable nutrients and functional foods. *Food Res. Int.* 44(7): 2368-2376.

Abstract

Benincasa is a monotypic genus with only one cultivated species *hispida*. In India, *Benincasa* exhibits rich diversity. There is substantial variation in vegetative traits and fruitcharacters and it is difficult to distinguish ashgourd genotypes based on their external morphology alone. In spite of having enormous morphological variability, all members at present are accommodated into asingle species. The taxonomic classification of *Praecitrullus* is still uncertain. The marker and pollen morphology studieshave revealed that *P. fistulosus*is more related to *B. hispida*, than compared to other cucurbit species.

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 characteristic variations on the sequences of identified genomic regions, which can distinguish individuals of a species. Species identification through barcoding is usually achieved by the retrieval of a short DNA sequence from a standard part of the genome (i.e. a specific gene region either from chloroplast, mitochondria or nuclear genome) and identifying the barcode gaps for each species. 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 genera *Benincasa* and *Praecitrullus* was done at Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, with the objectives to assess genetic diversity and interspecies relationship among *Benincasa* and *Praecitrullus* genotypes and to check the species status of *Benincasa* using barcodes.

Twenty-six ashgourd and five tinda accessions were morphologically screened in the field using the standard descriptor and basedaccessions were clustered, based on morphological characters recorded. The ashgourd accessions varied widely for their morphological traits,fruit size varied between0.191 to 12.0 kg. Thefruit shape, varied from club, spindle, cylindrical, oblong to ridged shape. The growth habit varied from long viny to short viny.Serrated, deep serrated and double serrated leaf margin was observed. For stem pubescencevariations recorded were hispid, downy, tomentose, puberulent, pilose,villous and hirsute. The cross compatibility of the ash gourd accessions were also

checked to establish the species status. All the accessions of ash gourd were cross compatiblein both the directions, showing that they belong to the same species. Based on the morphological data of ashgourd accessions nine cluster were generated. One representative accession from each cluster and three representative accessions of tinda selected for the molecular studies. Total genomic was isolated and subjected to PCR assay. The *matK* and *ITS2* gave the bands of 950 bp and 799 bp respectively, which were sufficient to identify barcode gaps.The markers were sequenced from 9 genotypes of ashgourd and 3 tinda genotypes.The BLAST analysis had shown that *matK* and *ITS2* both loci is 99 per cent efficient for species discrimination in *Benincasa* and *Praecitrullus*.

Barcode gap, a position in the sequence at which a particular nucleotide shows characteristic variation in all the members of a particular species, was identified for all the members of *Benincasa* and *Praecitrullus* species. Barcode gaps were identified for tinda, in comparison with *Benincasa* for both the loci. The barcodes in tinda were observed in *matK* locus at 47, 127 and 232bp. Similarly, for the locus *ITS2*, the barcode gaps in tinda were identified at 162, 207, 278, 279 and 288 bp. However, within *Benincasa*, there were no such characteristic barcode gaps, indicating that the null hypothesis of monotypic status of *Benincasa* is true.

Phylogenetic analysis using Clustal Omega showed the variation within *Benincasa* species and *Praecitrullus* species.The barcodes developed in this study could be successfully used to solve the genetic relatedness of *Benincasa* and *Praecitrullus* species.

ANNEXURE I

Reagents used for DNA isolation

1. CTAB extraction buffer (100ml)

CTAB (Cetyl Trimethyl ammonium bromide)	: 2g
Tris HCl (1M, pH- 8)	: 10ml
EDTA (0.5 M, pH- 8)	: 2ml
NaCl (5M)	: 30ml
Distilled water	: 54 ml

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

To twenty four parts of chloroform, one part of isoamyl alcohol was added and mixed properly.

3) Ethanol (70%)

To seventy parts of absolute ethanol thirty parts of distilled water was added and mixed well. 70% alcohol stored at 9° C.

ANNEXURE II

Composition of buffers and dyes used for gel electrophoresis

1. TAE buffer (50X)

Tris base : 242g

Glacial acetic acid : 57.1ml

0.5M EDTA (pH- 8) : 100ml

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 10mg/ml; working concentration 0.5ug/ml (Genie) and was stored at room temperature in a dark bottle.

4. Agarose gel

Gels with two different compositions were made: 1.5 per cent for Genomic DNA and 2 per cent for PCR samples.

ANNEXURE III

List of laboratory equipment used for the study

SI NO.	Equipment	Stage used	Company
1	Vortexer	DNA isolation	GeNei ™
2	High speed refrigerated centrifuge	DNA isolation	Kubota 6500
3	Nanodrop® spectopotometer ND- 1000	Qualitative assessment of nucleic acids	GeNei
4	Laminar Air Flow Cabinet	Preparation of PCR reaction mixture	Rotek, B&C
5	Proflex PCR system	Polymerase chain reaction	Applied Biosystems
6	Electrophoresis unit	Agarose Gel Electrophoresis (AGE)	GeNei
7	Gel Doc ™ XR+	Gel documentation	BIO-RAD
8	Ultra low temperature freezer	Storage of DNA samples	Haier BIO-MEDICAL