MOLECULAR CHARACTERIZATION OF GYNOECY IN BITTER GOURD (*Momordica charantia* L.)

by

NIVETHITHA B.

(2018-11-007)

DEPARTMENT OF PLANT BIOTECHNOLOGY

COLLEGE OF AGRICULTURE

VELLANIKKARA, THRISSUR-680656

KERALA, INDIA

2021

MOLECULAR CHARACTERIZATION OF GYNOECY IN BITTER GOURD (*Momordica charantia* L.)

by

NIVETHITHA B.

(2018-11-007)

THESIS

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Agriculture

(Plant Biotechnology)

Faculty of Agriculture

Kerala Agricultural University

DEPARTMENT OF PLANT BIOTECHNOLOGY

COLLEGE OF AGRICULTURE

VELLANIKKARA, THRISSUR - 680656

KERALA, INDIA

2021

DECLARATION

I, hereby declare that this thesis entitled **"Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)"** 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.

B. Nivetha

Place : Vellanikkara

Date : 17.04.2021

NIVETHITHA B.

(2018-11-007)

CERTIFICATE

Certified that this thesis entitled **"Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)"** is a record of research work done independently by **Ms. Nivethitha B.** 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 her.

Place: Vellanikkara
Date: 17.04.2021

Dr. T. Pradeepkumar
(Major Advisor, Advisory Committee)
Professor and Head
Department of Vegetable science
College of Agriculture, Vellanikkara

CERTIFICATE

We, the undersigned members of the advisory committee of **Ms. Nivethitha B.** for the degree of **Master of Science in Agriculture** with major in **Plant Biotechnology**, agree that the thesis entitled **"Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)"** may be submitted by **Ms. Nivethitha B.** in partial fulfillment of the requirement for the degree.

Dr. T. Pradeepkumar
(Chairman, Advisory Committee)
Professor and Head
Department of Vegetable science
College of Agriculture, Vellanikkara
Thrissur

17/04/2021

Dr. Abida P. S.
(Member, Advisory Committee)
Professor and Head
Department of Plant Biotechnology
College of Agriculture, Vellanikkara
Thrissur

Dr. Deepu Mathew
(Member, Advisory Committee)
Assistant Professor
Department of Plant Biotechnology
College of Agriculture, Vellanikkara
Thrissur

17/4/2021

Dr. Minimol J. S.
(Member, Advisory Committee)
Associate Professor (Plant Breeding and Genetics)
Cocoa Research Centre
Kerala Agricultural University, Vellanikkara
Thrissur

Acknowledgement

*First and foremost, I bow my head before **Almighty** for giving me the opportunity to undertake this research and bestowing me with inner strength, perseverance and helping hands during the entire course of my PG Research work.*

*With great respect and devotion, I express my heartfelt gratitude to Chairman of my advisory committee, **Dr. Pradeepkumar**, Professor and Head, Department of Vegetable Science, CoA, Vellanikkara and Director of Planning, KAU for his valuable guidance, constructive suggestions and ample support at each and every step of my research work. His lively encouragement and words of energy have motivated me in keep on going and bringing out my best throughout this endeavour. It has been a proud privilege to have had an opportunity to work under his guidance. His generousity and empathetic nature will be remembered forever.*

*My special thanks to **Dr. Deepu Mathew**, Assistant Professor, Department of Plant Biotechnology (CPBMB), CoA, Vellanikkara and member of my advisory committee for his expert advice, valuable suggestions, critical evaluation, endless support and generous guidance rendered during the whole venture. I am highly indebted to him for his involvement and constant encouragement throughout.*

*My diction doesn't seem too rich enough to provide suitable words to articulate my sincere thanks to **Dr. Abida P.S.**, Professor and Head, Department of Plant Biotechnology (CPBMB), CoA, Vellanikkara, and **Dr. M. R. Shylaja**, former Professor (Retd.) and Head, Department of Plant Biotechnology (CPBMB), CoA, Vellanikkara, for their generous support, ever willing help and encouragement during my entire study.*

*I have immense pleasure to express my sincere thanks to **Dr. Minimol J. S.**, Associate Professor **(**Plant Breeding and Genetics) Cocoa Research Centre, KAU, Vellanikkara and member of advisory committee for her critical evaluation and valuable suggestions. Her*

keen observations in detecting errors and correcting manuscript have helped me a lot for the improvement of the thesis.

My heartfelt thanks to my beloved teachers ***Dr. Binu****,* ***Dr. Rehna Augustine****,* ***Dr. Smitha Nair****,* ***Dr. Donald*** *and* ***Dr. Mariet*** *who have introduced me to this beautiful subject, been really amazing teachers and have managed to inculcate passion and interest for this subject. I also thank* ***Ravishankar sir*** *who helped me with Bioinformatics during this study.*

I feel fortunate for being blessed with bunch of wonderful classmates ***Sreekutty, Ashwini, Athira, Sharat, Pranali, Rasha, Aswathy, Sunny, Vaibhav*** *and* ***Ribin*** *for making this journey a smoother and memorable one. Without their timely help and mental support, this work would be incomplete. I would also like to extend hearty thanks to my batch friends* ***Anusha, Sravanthi, Apeksha, Mili*** *and* ***Alphy Mathew*** *and matron,* ***Malliga chechi*** *for their care, help and support.*

I owe my deepest and whole-hearted gratitude to my beloved senior friends, ***Shivaji sir*** *and* ***Vipul sir*** *for their critical suggestions, encouragement, advice and timely help in laboratory as well as for completion of the thesis. I extend my heartfelt thanks to all my lovable seniors,* ***Architha akka****,* ***Feba chechi, Anju chechi, Veerabhadra sir, Deepak sir*** *and* ***Nayana akka*** *for their kind advice, and support.*

I convey my gratefulness to ***Nimmy chechi****,* ***Asha chechi*** *and other Research Assistants of CPBMB for introducing the laboratory to us and lending a helping hand whenever any kind of support, guidance and advice in the laboratory were required. My sincere thanks to* ***Veni chechi*** *and* ***Varun chettan*** *of Hi-tech seed testing Laboratory, Department of Vegetable science, CoA, for their technical advice and guidance.*

I am grateful to all the ***non-teaching staffs*** *and* ***labourers*** *of the Department of Plant Biotechnology as well as the Department of Vegetable Science, CoA for the help rendered by them during the course of my study.*

I am deeply obliged to the members of my Tamil family ***Siva, Mounisha****,* ***Pavithrakumar****,* ***Dr. Saravana kumar****,* ***Dr. Rajkumar*** *and* ***Dr. Indhusree*** *for their jovial companionship, affectionate advice, timely help and encouragement. My special thanks to all my lovable junior friends* ***Gokul, Nithya, Naga, Mullai, Thenmozhi*** *and* ***Manoj*** *for their affection, ever willing help and mental support without which this venture would not have completed this much smoothly.*

The financial support by ***Department of Biotechnology (DBT)****, Ministry of Science and Technology, Government of India is greatly acknowledged.*

On my personal ground, I wish to express my genuflect love and affectionate gratitude to my beloved father ***Mr. S. Baluchamy,*** *my beloved mother* ***Mrs. M. Vasuki*** *and my dearest brother* ***Mr. B. Kabilan*** *for their unconditional love, warm blessings and prayers, encouragement, everlasting support and fidelity, which have made this venture easier. I can barely find words to express all their personal sacrifices and selflessness in giving me the best out of everything for which I am eternally grateful.*

With deep sense of gratitude, I avail this opportunity to thank my beloved aunt, ***Dr. A. Anandhi,*** *for being a well-wisher and shaping my childhood days by constant motivation and assistance, which definitely have made a positive impact in my academic life. Special words of thanks to* ***Harikuttan*** *for his empathy and moral support,* ***Susmitha*** *for her motivation, my buddy* ***Sarath Kumar*** *and Uncle* ***Jeyaprakash*** *for lending hands in tougher times of this venture.*

I wish to thank all the others who had helped me in one way or the other in completing the study.

Errors and omissions are entirely unintentional and a word of apology for those I have missed to mention.

Nivethitha B.

CONTENTS

LIST OF TABLES

LIST OF PLATES

LIST OF FIGURES

LIST OF ANNEXURES

ABBREVIATIONS

%	Percentage
>	Greater than
β	Beta
µg	Microgram
AR	Analytical Reagent
AFLP	Amplified Fragment Length Polymorphism
bp	Base pair
cm	Centimetre
cM	Centimorgan
CPBMB	Centre for Plant Biotechnology and Molecular Biology
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxyribo Nucleoside Triphosphate
EBI	European Bioinformatics Institute
ISSR	Inter Simple Sequence Repeat
In/Del	Insertion/Deletion
KAU	Kerala Agricultural University
Kb	Kilo base pairs
L	Litre
M	Molar
mg	Milligram

mL	Millilitre
mM	Milli Molar
NCBI	National Center for Biotechnology Information
ng	Nano gram
°C	Degree Celsius
OD	Optical Density
PCR	Polymerase Chain Reaction
pH	Hydrogen ion concentration
PVP	Poly Vinyl Pyrrolidone
QV	Quality Value
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
SCAR	Sequence Characterized Amplified Polymorphism
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeats
TAE	Tris Acetate EDTA
TE	Tris EDTA
TF	Transcription Factor
UV	Ultra Violet
UTR	Untranslated region
V	Volts

Introduction

1. INTRODUCTION

Bitter gourd (*Momordica charantia* L.; 2n=2x=22) is an important vegetable crop belonging to the family Cucurbitaceae. It is originated in tropical Africa and the genus *Momordica* comprises 47 species whereas only 12 species are distributed in Asia and Australia (Schaefer and Renner, 2010). Two main botanical varieties of the genus *Momordica* includes*, M. charantia* var. *charantia*, which produces large fruits and *M. charantia* var. *muricata,* a wild variety with small and round fruits (Chakravarthy, 1990; John and Antony, 2008).

The various names of bitter gourd in different parts of the world include bitter melon, balsam pear, bitter cucumber, African cucumber and *Karela* (Heiser, 1979). India produces about 1204 thousand metric tons of bitter gourd from an area of 99,000 hectare (NHB database, 2018-19). The leading states of bitter gourd cultivation in India are Tamil Nadu, Kerala, Maharashtra, Uttar Pradesh, Andhra Pradesh, Gujarat *etc*.

It is a fast growing trailing or climbing vine crop and having immense nutritional and medicinal values. The immature fruits of bitter gourd are widely consumed as a vegetable and are well known mainly for its anti-diabetic and other medicinal properties (Robinson and Decker-Walters, 1997). It is considered as the most nutritive among cucurbits as its fruits are rich in vitamin A, vitamin C, iron, phosphorus and carbohydrates (Miniraj *et al*., 1993; Desai and Musmade, 1998). Ascorbic acid and iron content are higher in bitter gourd compared to other cucurbits (Behera, 2004). Its immense medicinal properties are due to the presence of beneficial phytochemicals which are known to have anti-diabetic, antibiotic, anti-mutagenic, anti-cancer, antioxidant, antiviral, antihelmenthic and immune enhancing properties. Phytochemicals like charantin and momordin in fruits are clinically proven for its hypoglycemic activity (Grover and Yadav, 2004).

The major sex form in bitter gourd is monoecious where both male and female flowers occurs seperately in a same plant. However, gynoecious sex type bearing only

female flowers has also been reported from India, Japan and China (Ram *et al*., 2002; Behera *et al*., 2006; Iwamoto and Ishida, 2006).

Heterosis breeding is the best option for obtaining higher yield in bitter gourd, since most yield contributing traits such as number of fruits per plant, fruit length, fruit weight *etc*., are governed by non-additive gene action (Mishra *et al*., 2015). However, it involves labour-intensive emasculation and controlled mating, which makes development of hybrids so expensive.

Utilization of gynoecy is economical and easier for exploiting hybrid vigour in bitter gourd as it facilitates hybrid seed production under open field conditions eliminating the need for emasculation and assisted pollination. Use of gynoecious parental lines that produce only pistillate flowers, reduces the labour requirement, increases seed yield and helps to maintain hybrid genetic purity. These lines also have better genetic combining ability and gynoecious x monoecious hybrids mature early having higher yield potential (Dey *et al.,* 2010).

Sex expression in cucurbits is highly influenced by the environmental conditions in which bitter gourd seedlings grow (Wang and Zeng, 1997) which makes the early phenotypic detection of gynoecious lines a challenging task. Marker-Assisted Selection (MAS) using molecular markers closely linked to gynoecious locus helps in distinguishing gynoecious genotypes from monoecious bitter gourd types. However in India, molecular breeding in bitter gourd is in its infancy and the available conventional polymorphic markers for gynoecism in bitter gourd is very limited.

Hence, the present study was undertaken with the objective of developing molecular marker that can distinguish gynoecious genotypes from the monoecious genotypes of bitter gourd through characterization of putative candidate genes responsible for gynoecy.

Review of literature

2. REVIEW OF LITERATURE

2.1. General background

Momordica charantia L., commonly known as bitter gourd is a common vegetable of tropics and sub-tropics of Asia. The immature fruits are valued for its nutritional and medicinal uses. It belongs to the family Cucurbitaceae and it has a diploid chromosome number of 2n=22.

The significance of gynoecious lines in increasing crop yield is well known and studied extensively in cucurbits especially in cucumber and melon. Understanding the genetics and the molecular mechanism behind the gynoecious sex determination will pave the way for improvement of the trait. A great deal of experimental evidence is there in literature showing the influence of various plant growth hormones in regulating plant sex determination and the complex network of genes and pathways related to it. The pertinent literature on various aspects investigated during the course of this study is reviewed here under different topics.

2.2. Plant sex determination

Most of the angiosperms species produce hermaphroditic flowers in which both male and female reproductive organs are present. To prevent self- fertilization that may lead to loss of fitness due to inbreeding depression, the flowering plants have evolved some strategies for sex determination (Darwin, 1876). The production of unisexual flowers is one such strategy, in which the gametes of male and female are segregated on different flowers of the same plant (monoecious species) or on separate individuals (dioecious species). Thus the unisexual flower will have either stamens or pistils. There are two types of unisexual plants: monoecious with both male and female flowers on the same plant and dioecious with male and female flowers on different plants. Outcrossing promoted by this strategy leads to enhanced genetic diversity.

Flowers that are initially perfect undergo stage-specific and selective arrest of preformed organ primordia to attain unisexuality; but the difference arises in the manner in which and stage of floral morphogenesis at which the arrest of reproductive organs occurs (Dellaporta and Calderon-Urrea, 1994). For example, the sex organ primordia abort completely in maize, whereas in cucumber, the primordial arrest occur in their growth stage (Malepszy and Niemirowcz-Szcytt, 1991; Grant *et al*., 1994).

Sex determination, an important phenomenon in cucurbits is a developmentally regulated way culminating in plants having different sexual flowers. Many studies have been conducted for understanding what determines plant sex differentiation (Ainsworth, 1999; Tanurdzic and Banks, 2004; Perl-Treves and Rajagopalan, 2006; Boualem *et al*., 2008; 2009; Li *et al*., 2009; Martin *et al*., 2009). Yet the knowledge on what triggers the development decision process and how it could be carried out in an organ-specific manner remains vague.

2.3. Sex expression in bitter gourd

Cucurbits are not only the crops of economic importance, but also represented as model species for the studying plant sex determination, since they are having the ability to harbour various sex forms from primitive hermaphrodite to advanced monoecious type (Robinson and Decker-Walters, 1999). A definite gender of individuals in a plant population is termed as sex expression (Cruden and Lloyd, 1995; Neal and Anderson, 2005). The evolutionary changes over generations and the effects of dominant mutations give rise to modification of sex expression from perfect to the range of intermediate sex forms.

In bitter gourd, the sex expression is categorized into two main types, monoecious and gynoecious, whereas other sex forms like androecy, andromonoecy and hermaphrodite were also found rarely in nature. Further, plants with both female flowers and hermaphrodite were also detected (El-Shawaf and Baker, 1981).

Usual flowering habit in a monoecious bitter gourd plant is the formation of flowers in a preset, developmental sequence along the main stem, with an initial male flowers followed by an alternate phase of female and male flowers and terminated by a female flowering phase (Shifriss, 1961). Gynoecism is a condition in which all flower nodes bears female flowers. Gynoecious plants with only pistillate flowers or with predominate expression of pistillate flowers were identified in different places of world.

Although sex expression is controlled by sex determination genes, various factors such as phytohormones, chemical inhibitors and environmental factors like day length, humidity and temperature also alter the sex expression by modified expression of such genes (Meagher, 2007). Growing *M. charantia* seedlings under short photoperiod with low temperature (*e.g.* 25 ºC day/15 ºC night) promotes rapid development of gynoecy. High frequency of pistillate flowers can be encouraged by beginning the short day treatment from seedling emergence till sixth – leaf stage (Behera *et al*., 2011).

Similar to cucumber and melon, the plant hormone ethylene has feminizing effect in bitter gourd. Hence, chemical application of ethylene inhibitors like GA_3, silver nitrate and silver thiosulfate is adopted after the first female flower to induce hermaphroditic flowers (Matsumura *et al*., 2014). It helps in maintanance of gynoecious plants population.

2.4. Sources of gynoecious bitter gourd

The predominant sex type in the bitter gourd is monoecious but gynoecious sex type also has been randomly reported from India, China and Japan (Behera *et al*., 2006).

Although the gynoecy in bitter gourd occurs very rare in nature, Zhou *et al.* (1998) have identified a population with high proportion of pistillate flowers in bitter gourd and utilized it for hybrid development.

The gynoecious plants of bitter gourd with absolute expression (only pistillate flowers) of gynoecism was identified by Ram (2002) and it was the first report on its preliminary characterization. Three gynoecious plants, viz., Gy23, Gy63 and Gy263B were obtained in three germplasm populations and each plant is crossed with its

corresponding monoecious population (Ram *et al*., 2002). Finally they developed five populations with high proportion of pistillate flowers from the absolute gynoecious plant through selfing and sib-mating. These five populations were characterized by the emergence of pistillate flowers at lower nodes which makes them promising for yield improvement and hybrid production.

DBGy-201 and DBGY-202 are the two gynoecious lines isolated from indigenous source, from its related wild form, *M. charantia var. muricata* L. (Behera *et al*., 2006). In 2006, Iwamoto and Ishida developed gynoecious inbred from the cv. Aochu-naga plants with a large percentage of female flowers chosen from the breeding material line (LCJ980120) by selfing or crossing to same pedigree.

Matsummra *et al*. (2014) also identified a gynoecious bitter gourd line OHB61-5 from Okinawa, Japan in which only female flowers were present and it was supposed to be a spontaneous mutant. Two another gynoecious bitter gourd lines, IIHRBTGy-491 and IIHRBTGy-492 were reported by Varalakshmi *et al*. (2014).

Kerala Agricultural University identified a gynoecious line (KAU-MCGy-101) and tested it for stability of the gynoecious sex expression. The gynoecious sex expression in the inbred (KAU-MCGy-101) was stable, as only female flowers were produced throughout the growth phase in all the plants resulted from sib-mated population (Minnu Ann, 2019). Compared to the early reported gynoecious lines, the inbred was superior in terms of average fruit weight, fruit length and girth, thus holding enormous potential for earliness and yield in the upcoming breeding programmes.

Jadhav *et al*. (2018) isolated four gynoecious lines (Gy-1, Gy-14, Gy-15 and Gy-34) with 100 per cent gynoecy nature from the bitter gourd cultivar, NBGH-167 expressing predominately gynoecious sex type, through plant by plant selection of its transgressive segregants in progeny row method.

2.5. Inheritance of gynoecy in bitter gourd

In 2006, Ram *et al*. performed genetic analysis to study the inheritance of gynoecism in bitter gourd using a 100 per cent gynoecious line (Gy263B) and an inbred monoecious line Pusa Do Mosami. Both the lines were crossed, the obtained F_1 was crossed with Gy263B to produce test cross progeny. Both the F_2 and test cross progeny segregated for gynoecism in the ratio of 3:1 (3 monoecious: 1 gynoecious plants) and 1:1 respectively which revealed that gynoecy is under the control of a single recessive gene. A gene symbol *gy-1* is proposed for the expression of gynoecism in bitter gourd according to cucurbits gene nomenclature (Robinson *et al*., 1976). However, in contrast to their finding, Iwamoto and Ishida (2006) proposed the partial dominant nature of gynoecism in bitter gourd.

Behera *et al*. (2009) crossed two gynoecious lines, DBGY-201 and DBGY-202 with two monoecious cultivars, Pusa Do Mausami and Pusa Vishesh and observed the segregation of gynoecism in F_2 populations in the ratio of 3:1. Their findings confirmed the single recessive (*gy-1*) control of gynoecious sex form in bitter gourd reported earlier by Ram *et al*. in 2006. The same pattern of inheritance of gynoecism was also reported early by Poole and Grimball in 1939 in melon.

In the same way, the gynoecious line OHB61-5 was crossed with a monoecious line, OHB95-1A in which the frequency of female flowers per plant is approximately 5 per cent and the resulting F_2 population were analyzed. The results of Matsummra *et al*. (2014) were also on par with the earlier reports indicating the monogenic recessive nature of gynoecy trait, since they observed the segregation ratio of 3:1 for gynoecism. In 2015, Mishra *et al*. also substantiated the same in the cross DBGy 201 × S-2 and DBGy 201 × Pusa Do Mausami of bitter gourd.

Two cultivated inbred lines, the gynoecious line, K44 and the monoecious line, Dali-11 were crossed and the derived F_2 population was studied for determining the inheritance of gynoecy. The population was divided and grown in two different environment. Though their data showed *gy* as a qualitative character, Cui *et al*. (2018)

noticed a change in segregation pattern between these two populations grown at different environments. One of the population segregated in the 3:1 ratio for gynoecy, however in contrary to the previous findings, the other population showed two-genic segregation (15:1), showing that gynoecy may be controlled by interaction of atleast two genes.

In 2018, Rao *et al.* crossed the bitter gourd lines DBGy-201 with 100 per cent pistillate flower frequency and Pusa Do Mousami with 6 per cent pistillate flower frequency to investigate the inheritance pattern of gynoecy. All the F_1 population were monoecious with 33 per cent pistillate frequency showing it as a dominant character. F_2 population fitted well in the expected ratio of 3 monoecious: 1 gynoecious and BC_1P_1 population also segregated in 1: 1 ratio explaining the monogenic recessive inheritance nature of gynoecy in bitter gourd.

Overall, there are different opinions among the researchers regarding the inheritance of gynoecism in bitter gourd. The majority confimed that a monogenic recessive gene (*gy-1*) governs gynoecy (Ram *et al*., 2006; Behera *et al*., 2009; Matsummura *et al*., 2014) and the rest claims that gynoecism is partially dominant (Iwamoto and Ishida, 2006) or semi-dominant. Two pairs of genes governing gynoecy were also reported by Cui *et al*. (2018).

2.6. Candidate genes for gynoecy

Plant sex determination which is associated with yield and quality, is a research hotspot in plant science. In cucurbits, the genes involved in sex determination is orthologous and highly conserved. Cucumber is considered as a model system for sex determination studies due to its diverse floral sex types (Tanurdzic and Banks, 2004). Cloning of several major sex-determining genes in cucurbit species have paved the way for significant progresses over years in elucidating the mechanisms of this fundamentally important development process in flowering plants (Boualem *et al*., 2008; Li *et al*., 2009; Martin *et al*., 2009). The process is regulated on many omics level at several points during floral morphogenesis by a complex regulatory network of interacting genes (Pawełkowicz *et al*., 2019).

In 2010, Guo *et al*. performed digital expression profiling to compare the transcriptome dynamics between the two near-isogenic lines, a gynoecious line (WI1983G) and a hermaphroditic line (WI1983H) of cucumber. 214 differentially expressed genes were identified out of which, 90 and 124 genes were upregulated in gynoecious and bisexual flowers respectively. They also reported some transcription factors (TFs) like zinc finger TFs, auxin response TFs that have higher expression only in hermaphrodite flowers.

Gunniah *et al.* (2014) performed an *in-silico* differential gene expression analysis between a gynoecious (Gy323) and monoecious (DRAR-1) line to identify potential candidate genes governing gynoecy in bitter gourd. A total of 7865 genes were differentially regulated between two lines, 4131 genes were up regulated and 3734 genes were down-regulated in gynoecious line Gy-323. Among the up regulated, genes having significant role in ethylene biosynthesis or flower organ development might be the potential candidate genes of gynoecy. Thus a list of putative candidate genes for gynoecy have been identified which includes BTB/POZ domain-containing protein, ACS, ACO, ETR and ERFs. The maximum up regulation for gynoecious line was observed for an Arabidopsis homologue BTB/POZ domain-containing protein that determines the floral organ development with seven fold increase than the monoecious line.

Similarly, in bitter gourd transcriptome dynamics between female flowers of a gynoecious line (DBGY-201H) and hermaphrodite flowers (DBGY-201H) of the same line was compared using RNA-seq technology (Behera *et al*., 2016). A significant difference in gene expression was obtained for 477 annotated unigenes, out of which 237 were down-regulated and 59 were up-regulated in the DBGY-201H sample. They also reported differentially expressed ESTs between two flower types, which will provide valuable information for understanding the molecular mechanism of sex determination and accelerates gene discovery.

In 2018, Cui *et al*. constructed restriction site associated DNA (RAD)-based genetic map for bitter gourd using F_2 population between two cultivated inbred lines, the gynoecious line, K44 and the monoecious line, Dali-11. The map comprised 11 linkage

groups out of which they have identified QTL/gene gy/fffn/ffn, controlling sex expression involved in gynoecy, first female flower node, and female flower number in the first linkage group. They have also reported the list of candidate genes controlling the sex expression in bitter gourd by mapping the flanking RAD tags to the OHB3-1 reference genome. They also confirmed the location of *gy* locus determining the male to female flowers ratio at the distal end of the linkage group 1 which was previously reported by Matsumura *et al.* (2014).

In 2020, Matsummura *et al.* identified a gene *NUA* (*Nuclear Pore Anchor*) associated with the QTL determining gynoecy in bitter gourd. The 60-kb full-length gene comprises mostly of introns and the coding sequence of only 6,240 bp. Upon BLAST search they found homolog in *Arabidopsis thaliana* having 6,345 bp coding sequence which is reported to greatly influence stamen length and anther size, thus showing *NUA* as a potential candidate gene for gynoecy in bitter gourd.

It is well known that the plant hormones like auxin, ethylene and gibberellin play a major role in modifying plant sex. Studies have shown that majority of the differentially regulated genes between monoecious and gynoecious bitter gourd were related to plant hormone signal transduction such as auxin and ethylene biosynthesis which will be discussed below under separate topics.

2.6.1. Ethylene mediated sex determination

The phytohormone ethylene is involved in various aspects of the plant life cycle starting from germination of seed, root hair development, root nodulation, flower senescence, abscission to ripening of fruit (Abeles, 1973; Lieberman, 1979; Yang, 1985; Yang and Hoffman, 1984; Mattoo and Suttle, 1991). It is recognized as a potent modulator of growth and development in plants (Ecker, 1995).

The precursors of ethylene are *S*-adenosylmethionine (*S*-AdoMet) and ACC (1-aminocyclopropane-1-carboxylic acid). The initial step of ethylene biosynthesis comprises of converting *S*-AdoMet to ACC by the enzyme ACC synthase. Eventually, ACC is converted to ethylene by the enzyme ACC oxidase (Yang and Hoffman, 1984; Kende,

1993). Apart from this, ethylene signaling pathway genes like ETR, ERS, EIN and ERF also plays important role in regulation of ethylene in plants.

Ethylene is considered as the primary sex determination factor and it has been highly correlated with femaleness (Rudich *et al*., 1972a, 1976; Yin and Quinn, 1995). It is produced in high quantities in gynoecious cucumber plant than in monoecious (Rudich *et al.*, 1972b; Trebitsh *et al*., 1997). Furthermore, when ethylene biosynthesis and action were inhibited, female flower development was found to be suppressed (Beyer, 1976; Atsmon and Tabbak, 1979; Takahashi and Suge, 1980; Takahashi and Jaffe, 1984).

2.6.1.1. Ethylene related genes

The facilitators of ethylene receptor for the development of unisexual flowers of cucumber include two ACS genes, *CsACS1G* and *CsACS2* (Trebitsh *et al*., 1997; Kamachi *et al*., 1997; Yamasaki *et al*., 2001, 2003a, 2003b; Saito *et al*., 2007), two ACC oxidase genes, *CsACO2* and *CsACO3* (Kahana *et al*., 1999), three ethylene receptor genes, *CsETR1, CsETR2* and *CsERS* (Yamasaki *et al*., 2000), and a MADS-box gene, *ERAF17* (Ando *et al*., 2001) suggesting the role of ethylene in induction of female flower.

Ethylene response factor (ERFs) are transcription factors characterized by the presence of an AP2/ERF-type DNA-binding domain specific to plants. (Lin *et al.*, 2009; Klee and Giovannoni, 2011). They promote initiating the expression of downstream ethylene-responsive genes (Zhang *et al.*, 2009; Klee and Giovannoni, 2011; Liu *et al.*, 2016). ERF proteins takes part a major role in ethylene biosynthesis by controlling the transcription of *ACS* and *ACO* genes (Xiao *et al*., 2013 and Li *et al*., 2016; Tao *et al*., 2018).

2.6.1.1.1. ACS genes

It has been proposed that ethylene biosynthesis is primarily regulated at the ACC synthase level (Yang and Hoffman, 1984). A multi-gene family encodes this enzyme and its member shows differential expression as a response to many factors (Huang *et al*., 1991, Nakagawa *et al*., 1991, Olson *et al*., 1991, Rottmann *et al*., 1991, Liang *et al*., 1992, Zarembinski and Theologis, 1993, Destefano-Beltran *et al*., 1995). Molecular regulation

of sex expression by ethylene can be understood by identification of specific ACC synthase (ACS) gene(s) involved in pistillate flower development.

The enzyme ACC synthase activity was observed more in the apex of gynoecious plants than in monoecious plants (Trebitsh *et al.,* 1987). The gene ACS was found to be differentially expressed between gynoecious and monoecious cucumber when genome-wide transcriptome profiling was done. The differential expression is highly attributable to the mutation in ACS and auxin related genes (Wu *et al*., 2010).

High expression of *CsACS1*, *CsACS2*, *CsACS11* accompanied with elevated endogenous ethylene concentration was correlated with gynoecy and female flower formation in cucumber (Knopf and Trebitsh, 2006; Saito *et al.*, 2007; Boualem *et al.*, 2016; Chen *et al.*, 2016).

A single copy of Cs-ACS1 gene was obtained by Trebitsh *et al*. (1997) in monoecious (cv. Marketmore76; MMff) cucumber, whereas, at least one additional copy of this gene was obtained from gynoecious (cv. Marketmore 76F; MMFF). The accelerated rate of sex conversion may be the reason for the occurrence of additional Cs-ACS1G in gynoecious line. The monoecious lines contained Cs-ACS1G in their genome which is auxin-inducible. So, it can be inferred that female flowering potential of monoecious cucumber plant is retained and manipulation of auxin and ethylene can promotes female flowering. They suggested that the *Cs-ACS1G* is closely linked to F locus and may play pivotal role in sex differentiation in cucumber.

Wu *et al*. (2010) also confirmed the presence of one additional copy of CsACS1G in gynoecious cucumber. Li *et al*. (2020) confirmed the refined structure of F locus in gynoecious cucumber, and elaborated that out of three genes (*CsACS1*, *CsACS1G*, and *CsMYB*) in the F locus, only *CsACS1G,* a duplicated copy of *CsACS1* with recombinantdistal promoter is responsible for gynoecy. Further, loss of this additional copy due to unequal crossing over during meiosis caused gynoecy instabilty in cucumber.

Among the ACC synthase encoding genes, only *Cs-ACS2* mRNA was detected at the apices of gynoecious cucumber in which female flowers were developing. It was also

found co-segregating with M locus. Thus *Cs-ACS2* gene was reported to be potentially involved in regulating the development of female flowers at the apex of the gynoecious cucumber plants (Kamachi *et al*., 1997). Involvement of *CsACS4* gene in the diurnal rhythm of ethylene evolution was suggested by Yamasaki *et al*. (2003b).

A modified model for sex determination in cucumber was proposed by Li *et al*. (2012) in which *CsACS2* and *CsACS1G* encode ACC synthase enzymes; their cooperation produces male, female or perfect flowers. In female flowers bud primordia, at first F gene activated and female flower development will be promoted by ethylene. It also activates M which, then starts its positive feedback expression. The accumulation of ethylene continuously beneath the region of the pistil primordia by translation of the M gene mRNA constantly arrests development of stamen primordia, resulting in female flowers.

Boualem *et al*. (2009) found the specific expression of *CmACS-7* in carpel primordia of buds which are determined to develop into carpels of melon. An inter-organ communication is likely responsible for staminal development inhibition after ACS expression. Development of female flowers in cucumber is influenced by ethylene through organ specific induction of DNA damage in the primordial anther of female flowers. This was demonstrated by Wang *et al*. in 2010 where they observed the down regulation of *CsETR1,* the component gene in ethylene signaling pathway, temporally and spatially in the stamens of stage-6 female flowers. In contrast to that in water melon, the development of male flowers is promoted by ethylene. Both water melon ACS genes, *CitACS1 and CitACS3* were expressed in floral tissues but the former one was also detected in the vegetative tissues showing its involvement in plant growth and development. The expression of *CitACS3* was found in in open flowers and staminate floral buds (male or hermaphrodite), but absent in the female flowers. Up regulation of *CitACS3* by ACC and its role of ethylene-regulated anther development were explained by Minkov *et al*. in 2008.

Bouleam *et al*. (2008) revealed that a single-nucleotide mutation in the highly conserved region of the *CmACS-7* affected the protein function and resulted in the reduced enzymatic activity of ACC synthase in melon. The active enzyme does not leads to the

carpel development directly but suppresses the stamen development in female flowers. Thus, failure of the enzyme to act results in the hermaphrodite flowers, leading to an andromonoecious plant. The stamen inhibiting 'A' locus in melon encodes this gene. *CmACS7* is the orthologous gene of *CS-ACS2* in cucumber (Boualem *et al*., 2009). Similarly in water melon, the active *ClACS7* enzyme leads to female flower development in monoecious lines, whereas a hermaphrodite flowers will be resulted when the enzymatic activity is decreased (Bouleam *et al*., 2016).

Manzano *et al*. (2016) isolated and characterized the watermelon gene *CitACS4*, which is highly similar to the previously reported ethylene biosynthesis genes, *CmACS7, CsACS2* and *CpACS27A* known to be associated with andromonoecy. ACS type III enzyme encoded by this gene has its predominant expression in pistillate flowers. A mis-sense mutation in the conserved region (C364W) of this gene was detected in andromonoecious lines, with concomitant reduction in the ethylene levels. Thus they elucidated that the mutation interrupted the ethylene biosynthesis required for arrest of stamen in flower buds destined to become female, resulting in a hermaphrodite flower for andromonoecy.

In 2015, Bouleam *et al*. illustrated that the gene *CmACS11* in melon is responsible for determination of female flower by repressing the male promoting gene, *CmWIP1*. So, loss of *ACS11* function causes stamen development leading andromonoecy. Similarly in cucumber, *CsACS11* was suggested to be connected with 'A' locus. ERF protein genes in cucumber and melon, *CsERF11 and CmERF11* mediates ethylene regulated transcription of *CsACS11* or *CmACS11* gene respectively, as a response to ethylene signaling (Tao *et al*., 2018).

Silver Thiosulphate (STS) is an ethylene inhibitor that blocks the perception of ethylene by plants thereby modifiying its sex. In bitter gourd, spraying Silver Thiosulphate (STS) @ 6mM can be manipulated for conversion of female flower buds to hermaphrodite ones. In 2015, Kumari took up a study to understand the expression of *McACS* genes during four different stages of conversion of female buds to perfect buds. The expression of *McACS2* gene was seen reduced after 1st spray and increment in expression was observed during the formation of hermaphrodite flowers. Thus, it is hypothesized that ethylene

perception will be reduced by silver ions which in turn resulted in the reduced accumulation *of McACS* gene mRNA in the flower buds of gynoecious line (DBGy-201). This provides the strong evidence that *McACS2* mediated biosynthesis of ethylene is associated with sex determination in bitter gourd.

In 2014, Gunniah *et al*. identified a list of putative candidate genes responsible for gynoecy in bitter gourd through *insilico* differential gene expression analysis between monoecious and gynoecious plants. Ethylene related genes like ACS, ACO, ETR and ERS predominated the list.

Synteny mapping and phylogenetic analysis of conserved genes done by, Urasaki *et al*. in 2017, revealed that bitter gourd was closely related to water melon (*Citrullus lanatus*) than Cucumber (*Cucumis sativus*) or melon (*C. melo*). Since the genes of sex determination were identified in cucumber and water melon, they were used for comparison in bitter gourd genome to find the orthologous genes. The protein MOMC3_649 and two proteins MOMC46_189, MOMC518_1 were the presumed orthologs of *CmAcs11* and *CmAcs-7* in bitter gourd respectively. On validation, elevated expression of MOMC3_ 649g (encoding CmAcs11-like protein) and MOMC518_1g (encoding CmAcs-7-like protein) in pistillate flowers than staminate ones was observed. The results were consistent with previous findings which reported the high expression of *ACS11, ACS7 and ACS2* in female flower buds of cucumber and melon.

2.6.1.1.2. ACO genes

High ACC oxidase activity and transcript levels were associated with actively dividing tissues in pea and sunflower seedlings, respectively (Liu *et al*., 1997). Organ-specific ACO expression patterns were demonstrated in developing tomato flowers (Barry *et al*., 1996). In Petunia, two ACO genes were specifically expressed in developing pistil tissue (Tang *et al*., 1993).

In 1999, Kahana *et al*. isolated three full length cDNAs of ACO genes, *CsACO1, CsACO2* and *CsACO3* from different sex type and phases in cucumber plants. Differential

response to etherel treatment was observed in *CsACO2* and *CsACO3* in different floral types and organs. Further, they demonstrated that the mechanism controlling the ethylene production during flower development is a complex process and the expression of *CsACO2* in the shoot apex is irrelevant to the process of sex determination.

Chen *et al*. (2016) provided an evidence of a single nucleotide mutation in ACO gene conferring andromonoecy to cucumber. The enzymatic activity of *CsACO2* gene will be disrupted resulting in the decreased ethylene emission from shoot tips during the critical stage of sex determination. The development of carpel was blocked, turning the plant to andromonoecy. The orthologous gene of *CsACO2* in melon, *CmACO3* also showed similar expression pattern in the region of carpel development. Though, it is contradictory to the previous report by Kahana *et al*. (1999), they elucidated that their conclusion was not drawn from the buds of critical stage of sex determination.

2.6.2. Transcription factors (TFs) in sex determination

Sequence-specific DNA-binding factor, otherwise called transcription factor (TF) is a protein that regulate the transcription of DNA to mRNA by attaching to specific DNA sequences. Transcription factor regulates various genes involved in organ differentiation (Singh *et al*., 2002). Sharma *et al*. (2013) also demonstrated that neo-functionalization and sub-functionalization of transcription factor plays a major role in plant morphology differentiation.

2.6.2.1. MADS box genes

In flowering plants, sex determination is a complex process involving several genes expressed differentially at various tissues and developmental stages (Charlesworth and Mank, 2010). One among those is the genes encoding the MADS box family transcription factors which plays a prominent role in the reproductive stage of the plants.

The *Arabidopsis* AGAMOUS (AG) gene which encodes the floral homeotic genes is one of the first isolated MADS box genes (Yanofsky *et al*., 1990). AG is a class C

homeotic gene that play a key regulatory role in determining the identity of stamens and carpels, which is demonstrated using overexpression and mutant phenotypes. It encodes for MADS box transcription factor and its interaction with other MADS-box protein plays an essential role in induction of floral organ in *Arabidopsis* (Honma and Goto, 2001). Kater *et al*. (2001) performed a study in monoecious cucumber which provides strong evidence that the position dependent arrest of particular reproductive organ, a basic mechanism in unisexual flower development is highly controlled by class C homeotic genes.

The transcription factors encoded by these genes contains a highly conserved DNA binding domain, the MADS box and a moderately conserved domain called the K-box, which has the ability to form coiled-coil structures that allows dimerization of these transcription factors (Ma *et al*., 1991). The amino acid sequences of MADS box genes from different species was compared which has led to the classification of the MADS box gene family in B, C and other groups of MADS box genes as *AGAMOUS (AG), APETALA2 (AP2), APETALA3 (AP3)/ PISTILLATA, APETALA1/AGL9* (Purugganan *et al*., 1995). The uniqueness of these MADS box is their ability to bind to the specific CArG box in the promoter region, thereby regulating gene expression.

The homologs of the AG gene have been isolated from many other plants like antirrhinum, petunia, tobacco, tomato, rapeseed, rice, wheat, maize and cucumber and their functional analyses by ectopic expression demonstrated their essential role in establishing stamen and and carpel identity (Mandel *et al*., 1992; Kempin *et al*., 1993; Tsuchimoto *et al*., 1993; Pnueli *et al*., 1994; Saedler and Huijser, 1994; Kang *et al*., 1995).

AGAMOUS-LIKE6 (AGL6)-like MADS-box genes are ancient and are presented in wide range range of eukaryotic genomes including gymnosperms and angiosperms (Becker and Theissen, 2003). The critical role of AGL6-like genes in flower development of angiosperms have been shown by several evidences. When the *AGL6* orthologs from orchid (Hsu, 2003) and from hyacinth (Fan, 2007) were overexpressed in Arabidopsis they have been shown to promote flowering and cause flower homeotic transformation.

Findings of Perl-Treves *et al*. (1998) imply that, the pathway which leads to the arrest of reproductive organ in cucumber unisexual buds acts independently of MADS-box gene expression. In contrast to their findings, Ando *et al*. in 2001 demonstrated that induction of female flowers by ethylene might be regulated by the expression of a MADS-box gene, *ERAF17* in apical floral buds of cucumber.

Table 1. List of MADS box family genes involved in floral regulation

S. No.	Gene	Crop	Reference
1	*Cucumber MADS box gene 1 (CUM1)* and *CUM10*	*Cucumis sativus*	Kater *et al*. (1998)
2	*OsMADS6*	*Oryza sativa*	Li *et al*. (2010)
3	*TaAP3, TaAGAMOUS* and *TaMADS13*	*Triticum aestivum*	Su *et al*. (2019)

2.6.2.1.1. MADS-box genes in *Momordica*

Homologue of *AG*-like gene in bitter gourd, named BAG was isolated using conserved MADS box prolcin family region. High sequence similarity of amino acid was found with *CUM10,* cucumber MADS-box protein (95%) and *AGL11* in *Arabidopsis thaliana.*

In 2005, Peng *et al*. isolated two MADS box gene, *MCAG2* and *MCAG6* from female flower buds of *Momordica charantia* L. on the basis of the conserved MADS box sequences. They had cloned and sequenced these genes to full length for undergoing phylogenetic analysis and sequence comparison. Amino acid sequence of *MCAG2* and *MCAG6* were highly identical to *AG*-like and *AGL6*-like genes, respectively. High similarity was also observed between proteins of *MCAG2* and *Cucumber MADS box gene 1, CUM1*. Maximum expression of the genes *MCAG2* and *MCAG6* were observed in carpel and shoot apex respectively. *MCAG6* was the first *AGL6*-like gene in Cucurbitaceae and its high level of expression in vegetative tissues of bitter gourd was interesting as it is not

common in the previously studied plants. Pattern of gene expression in different tissues clearly indicated the role of two genes in regulation of floral organ development in *Momordica charantia* L.

In 2019, Hseih and Yin-yin Do performed genome-wide analysis of MADS-box genes in bitter gourd for identification of different members of that family and characterization of those genes. Phylogentical analyses revealed that the MADS box protein from bitter gourd were closely related to watermelon. Three putative genes related to sex determination mechanism *McMADS1, McMADS10, McMADS22* were identified according to gene expression clustering of the transcripts. They also stated that hormones like GA_3 and different environmental factors may have influence on these genes, which seems contradictory to the previous findings of Perl-Treves *et al*. (1998) in cucumber.

Mohanty and Joshi (2019) have undertaken a study to explore the functions of MADS-box family transcription factors (TFs) in *Momordica dioica* Roxb. with respect to floral organogenesis. 17 *MdMADS* genes were cloned and characterized from the flower buds of *M. dioica* Roxb. They clustered the genes into three sub-groups based on the motifs as $MIKC^C$, MIKC* and Mα (type I). From the gene expression analysis of the transcriptome, they found the association of 11 $MIKC^C$ with floral homeotic functions, four MIKC* with male gametophyte development and two M-type genes with female gametophytic development significantly. These genes may be considered as a candidate MADS box genes responsible for sexual dimorphism in *M. dioica* Roxb.

2.6.2.2. Auxin mediated sex determination

Auxin, the first phyto-hormone known has wide-ranging effects on the growth and development of plant from embryogenesis to senescence. It stimulates the production of ethylene and induces female flower formation (Takahashi and Jaffe, 1984) particularly in vegetables (Galun *et al*. 1962; Atsmon and Tabbak, 1979). Expression of specific ACS genes were increased by auxin (Trebitsh *et al*., 1987).

The genes exerting direct control on the biogenesis of sexual organs can be induced by auxin. The transcriptional regulation of gene families, such as auxin/indole-3-acetic acid (*Aux/IAA*), Gretchen Hagen3 (*GH3*), small auxin up RNA (*SAUR*) and *auxin response factor* (*ARF*) by auxin coordinates the development of plants (Shen *et al*., 2014; Zouine *et al*., 2014).

2.6.2.2.1. ARF genes

Auxin response factors (*ARFs*) is the family of functionally distinct DNA-binding factor that offers specificity to auxin response by selecting target genes as transcription factors. They are the important members of auxin signaling pathway and have the tendency to either activate or repress transcription of several auxin response genes by binding to the conserved auxin response element (AuxRE) in their promoter region (Tiwari *et al*., 2003).

In 2015, Liu *et al*. identified various ARF genes in papaya (*Carica papaya*) and analyzed their expression pattern at different stages of development. The genes *CpARF1, CpARF2*, *CpARF4*, *CpARF5*, and *CpARF10* were expressed highly at the very initial stage in developing of flower and declined during the following stages. *Auxin responsive factor* genes (*arf* genes), one of the auxin regulated genes controls the development of pistillate flowers in tomato, arabidopsis and papaya (Li *et al*., 2016).

The interaction between Aux/IAAs and auxin response factor (ARF) has significant influence in auxin signaling transduction pathway. Low concentration of auxin inhibits the activity of ARF protein by dimerization with Aux/IAAs (Ulmasov *et al.,* 1997; Hagen and Guilfoyle, 2002) and in high concentration of auxin, they will be activated.

De novo transcriptome sequencing performed by Shukla *et al*. (2015) derived a total of 65,540 transcripts for gynoecious line (Gy323) and a monoecious line (DRAR1) of bitter gourd. Digital gene expression analysis was done to compare the transcript count and associated differential expression pattern between both the lines. Among that, highly differentially expressed gene were related to hormone signaling and response. Based on

available transcripts they have identified 56 ARF genes which play a significant role in auxin-regulated gene expression and development.

2.6.3. WRKY transcription factors (TFs)

WRKY proteins forms the large family of transcription factor in plants. Members of this family contains one or two conserved WRKY domains, in which N-terminus contains conserved WRKYGQK heptapeptide sequence and C terminus contains C_2HC or C_2H_2 zinc-finger protein motif (Ishiguro *et al*., 1994). The WRKY family will be divided into three groups (I, II and III) on the basis of C-terminal zinc-finger motif structure and the number of WRKY domains (Rushton *et al*., 2010). The members of this family bind to W-box [(T)TGAC(C/T)] found in the promoter sequences of various target genes.

WRKY DNA-binding protein has been identified one among the putative transcription factor of sex determination, when transcriptome profiling was done between monoecious and gynoecious cucumber (Guo *et al*., 2010). Though its role in stress response and various processes of plant development were well known, this is the first report to correlate WRKY factors with plant sex determination.

WRKY TFs plays a crucial role in plant signal regulation particularly in ABA response pathway (Rushton *et al*., 2012). Rudich *et al*. (1974) suggested that ABA promotes female flowering tendency and inhibits the GA_3 action which increases male flowers. Regulation of flowering by WRKY transcription factors have been studied. In Arabidopsis, Zhang *et al*. (2018) identified WRKY75 which positively regulated floral initiation by interaction with DELLA proteins, which acts a feminizing factor by negative regulation of GA signaling pathway (Peng *et al*., 1997; Silverstone *et al*., 1998 and Pysh *et al*., 1999).

Shukla *et al*. (2015) performed *de novo* transcriptome sequencing of monoecious (DRAR1) and gynoecious lines (Gy323) of bitter gourd. Comparison between transcription factors of both of these lines was adopted to identify candidate genes of sex determination. They reported differentially expressed transcripts sets involved in floral differentiation

Momordica transcripts were functionally annotated to *Arabidopsis thaliana* transcription factor sequences using UniProt non-redundant protein and TAIR (The Arabidopsis Information Resource) data sets to obtain classes of transcription factors mainly associated in the differential pattern formation. They have identified 80 WRKY transcription factors associated genes based on the available transcriptome.

2.7. Molecular markers associated with gynoecy

Selection for high and stable yield using conventional phenotypic method is very time consuming and expensive, since it requires yield evaluation multiple environments over many seasons (Yuan *et al*., 2002). Marker assisted selection (MAS) in contrast definitely accelerates the breeding progress and powerful tool for selecting traits such as gynoecism. The marker patterns of the two bulks shows polymorphism only when they show genetic linkage to underlying gene (Giovannoni *et al*., 1991). The markers that are tightly linked to the trait of interest at less than 5cM can be used for marker assisted breeding (Tanksley, 1983b).

Identifying molecular marker(s) associated with the trait of interest has enormous potential for the crop improvement through molecular breeding (Collard *et al*., 2005). But, it is very important to ensure that the markers used are polymorphic in the population so that it can be successfully applied in marker-assisted selection in a breeding programme. The polymorphic molecular markers are scarce in public database which hinders genetic mapping and its usage in molecular breeding of bitter gourd.

Identification of gynoecious sex type in very earlier stage using marker-assisted breeding in bitter gourd holds immense potential in introgressing the gynoecious trait. Since flowering occurs only 35 - 40 days after sowing, the molecular markers plays a significant role in ascertaining the purity of the gynoecious line at an early stage for a cost effective hybrid seed production. The commonly used conventional molecular markers include RAPD (Random Amplified Polymorphic DNA), AFLP (Amplified Fragment Length Polymorphism), RFLP (Restriction Fragment Length Polymorphism), SSR (Simple

sequence repeats), ISSR (Inter Simple sequence repeats), SCAR (Sequence Characterized Amplified Polymorphism) and SNPs.

Mishra *et al*. (2014) performed a study to identify molecular markers that can be used for differentiating monoecious and gynoecious line in bitter gourd. The gynoecious DBGy-201 and monoecious Pusa Do Mausami were the parents used in the study. A total of 154 primers (67 SSR, 16 ISSR and 71 RAPD) were used in screening the parents and F_2 population derived out of it by the use of bulked segregant analysis (BSA). Among them, only one RAPD primer OPZ 13 (3′GACTAAGCCC 5′) produced a polymorphic band of 700 bp band which is specific only to the gynoecious parent and bulks. In addition, the authors confirmed that, the RAPD marker, OPZ 13_{700bp} is linked to the gynoecy locus (*gy*) at a distance of 22cM, which was the first report on targeting the gene responsible for gynoecious trait in bitter gourd.

In a study conducted by Gaikwad *et al.* in 2014 for identification of molecular markers linked to gynoecy in bitter gourd, an ISSR marker showing polymorphism between monoecious and gynoecious plants was found, out of 200 RAPD and 28 ISSR markers screened. The gynoecious plants used for screening were the F_5 generation plants of gynoecious line DBGy-201. The marker was a 1000 bp fragment amplified by the primer $[AC]_8T$, which was conspicuously absent in monoecious plants.

2.7.1. SSR markers for gynoecy

Simple sequence repeats (SSRs) are ideal genetic marker type among other DNA markers. Due to its desirable attributes like multi-allelic nature and co-dominant inheritance, significant importance have been gained by SSR markers in plant genetics and breeding (Akkaya *et al*., 1992). The advantages includes very low requirement of DNA, ease of detection by PCR and capillary electrophoresis (Morgante and Olivieri, 1992; Peakall *et al*., 1998), highly polymorphic (Gupta *et al*., 1996) and extensive genome coverage. SSRs are highly informative than dominant DNA markers (Powell *et al*., 1996) and shows more heterozygosity.

Guo *et al*., 2010 screened cucumber unigene data sets to identify SSR markers between isogenic lines WI1983G (gynoecious) and WI1983H (hermaphrodite) of cucumber. They have identified set of 2,860 EST-derived SSR markers between the both lines, out of which 1,679 had sufficient flanking sequences for primer design. Miao *et al*. (2011) identified two SSR markers, SSR13251 and CSWCT28 flanking the *F* locus at the distance of 1.2cM and 1.7cM, respectively. This is contradictory to the previous report of Fazio *et al*. (2003) which reported the linkage of SSR marker CSWCT28 to the F locus at a distance of 5cM in cucumber.

PPC-2, a gynoecious and parthenocarpic line and Pusa Uday, monoecious and non-parthenocarpic cucumber cultivar were screened using 179 SSR markers, out of which 39 SSR markers differentiated them. When their F_2 mapping population were screened with these markers, only 17 out of those found segregated along with the population. Those markers were used for genotyping and linkage analysis. Spanning a total distance of 100.9cM on chromosome 6, these markers were placed along with the F locus. Among those two SSR markers, SSR13251 and UW020605 were found to be closely linked to gynoecious (F) locus at 1.0 and 4.5cM, respectively which can be used in marker assisted selection of gynoecy in cucumber (Jat *et al*., 2019).

Bommesh *et al.* (2019) used four gynoecious cucumber lines belonging to the F_4 families, namely IIHR-434, IIHR435, IIHR-436 and IIHR-437 for SSR marker validation. Twelve previously reported SSR primers were used to validate the gynoecious cucumber lines of F_4 population's along with monoecious variety Swarna Agethi. Two primers, SSR-02021 and SSR-18718 among the twelve showed polymorphism differentiating the gynoecious plants from the monoecious plants and the remaining markers showed monomorphic bands.

2,440 SSR primers for gynoecious and 2,404 for monoecious line of bitter gourd were designed by Shukla *et al*. (2015) based on transcriptome sequencing information. Comparison of SSR variation pattern between both the lines will aid in screening and developing of markers for gynoecism.

2.7.2. Candidate gene based marker

Win *et al*. (2015) identified F (female) locus-specific co-dominant marker based on In/Dels polymorphism which distinguished perfectly homozygous and heterozygous gynoecious cucumber lines. 50-kb genomic sequence of five gynoecious and five monoecious inbred lines were analyzed to detect the polymorphism linked to F locus. They designed a specific pair of primers Cs-BCAT-F/CsBCAT-R based on 56-bp In/Del at 3'-UTR of branched-chain amino acid transaminase (BCAT) gene to analyze parents and F_1 hybrid. Polymorphism was found between parents, with an approximately 216 bp band in the monoecious and 160 bp band in the two gynoecious lines while both the bands were present in F_1 plants amplified both bands.

2.7.3. SNP markers for gynoecy

High quality genotyping markers can be developed using SNP from coding regions. It also provides insights into functional changes in protein coding domains (Baird *et al*., 2008).

In 2010, Guo *et al*. compared the ESTs generated using the Roche-454 massive parallel pyrosequencing technology, for the near isogenic lines of cucumber, a gynoecious plant WI1983G (MMFF) and a hermaphroditic plant WI1983H (mmFF) using Roche-454 massive parallel pyrosequencing technology. They have identified a total of 114 SNPs between the two lines, which comprises 42 transitions, 16 transversions and 56 were In/Dels.

In 2014, Matsumura *et al*. conducted a study with an objective of genetically mapping the locus for gynoecy (*Mcgy*) in OHB61-5, a commercial gynoecious line of bitter gourd used as a maternal parent in F_1 cultivar production. To reveal genome-wide DNA polymorphisms a RAD-seq (restriction-associated DNA tag sequencing) analysis was employed, which is a high-throughput tool for DNA polymorphism detection and genotyping. F_2 individuals of the cross were analyzed using RAD-seq and based on the results, a linkage map was constructed using 552 co-dominant markers. Pooled genomic

DNA from monoecious or gynoecious F_2 plants were analyzed and it resulted in the identification of several SNP loci genetically linked to gynoecy. The closest SNP locus, GTFL-1, was located 5.46cM from the putative gynoecy locus, *Mcgy*. To apply it practically for marker-assisted selection of gynoecy in bitter gourd, it was converted to a conventional DNA marker using invader assay technology, the SNP-typing tool. The draft genome sequence positioned GTFL1 sequence in the end region of scaffold 326.

A RAD-based (restricted site associated DNA) genetic map for anchoring scaffold sequences was developed by Cui *et al*. (2018) using F_2 mapping population. The map spanned a total genetic distance of 2,203.95cM comprising of 11 linkage groups and 1009 SNP markers. Transcripts of monoecious and gynoecious line were compared by Shukla *et al*. (2015) which yielded SNP variations between those two lines.

Rao *et al*. (2018) employed GBS technique for identifying molecular markers in bitter gourd. A high-density, high-resolution genetic map was constructed with 2013 high quality SNP markers across 20 linkage groups (LG) spanning a cumulative distance of 2329.2. A total of 22 QTLs were identified and mapped on 20LGs for four traits *viz.* (gynoecy, sex ratio, node and days at first female flower appearance). The SNP markers markers TP_54865 and TP_54890 flanked gynoecious (*gy-1*) locus on LG 12 with TP_54890 closest at a distance of 3.04cM and TP_54865 at 7.06cM.

.

Materials and Methods

3. MATERIALS AND METHODS

The study entitled 'Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)' was carried out at Centre for Plant Biotechnology and Molecular Biology and Department of Vegetable Science, College of Agriculture, Vellanikkara, Thrissur, during 2018- 2021.

3. 1. Materials

3. 1. 1. Plant materials

The two bitter gourd genotypes contrasting for sex expression *i.e.* the monoecious, MC-136 and the gynoecious line, KAU-MCGy-101 identified from Department of Vegetable science, College of Agriculture, Vellanikkara were used as base material to study the sequence variation between the two genotypes and develop molecular marker showing polymorphism between them. Three other monoecious lines Priya, Priyanka and the wild-type *M. charantia* var. *muricata* were used in this study for screening of the developed molecular markers.

The seeds were pre-treated with pseudomonas and kept overnight. Next day, the seeds were wrapped in germination paper with water for facilitating germination. All the germinated monoecious and gynoecious seeds were grown separately in Department of Vegetable Science, College of Agriculture, Vellanikkara for obtaining leaves for DNA isolation.

3. 1. 2. Laboratory chemicals, glass wares and plastic wares

The chemicals used in the study were purchased from Merck India Ltd., HIMEDIA and SISCO Research laboratories and were of AR grade. The PCR reagents *viz. Taq* DNA polymerase, dNTPs, $MgCl_2$, *Taq* assay buffer and DNA ladders were ordered from Bangalore Genei Ltd. The Plastic wares like tips, tubes and beakers used in the study were obtained from Axygen and Tarson India Ltd. The glass wares used in the study were obtained from Borosil Ltd. Primers pairs were synthesized and bought from Sigma Aldrich

Chemicals Pvt. Ltd.

3. 1. 3. Equipments and machinery

The research work was done using molecular biology facilities and equipments available at Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Agriculture. High speed refrigerated centrifuge Kubota 6500 was used for carrying out centrifugation. Estimation of quality and quantity of DNA was done using NanoDrop^{R} ND-1000 spectrophotometer. The Polymerase Chain Reaction was carried out in Sure Cycler 8800 thermal cycler (Agilent Technology), Veriti thermal cycler (Applied Biosystems, USA) and ProFlex PCR system (Applied Biosystems, USA). Agarose gel electrophoresis was performed using Horizontal gel electrophoresis system (BIORAD, USA). Imaging and documentation of the agarose gel was done in Gel Doc XR^{+} (BIORAD, USA). Quantity One software was used for analyzing the gel images (Annexure I).

3. 2. Methods

3. 2. 1. Candidate gene selection and sequence retrieval

20 Putative candidate genes for sex expression, particularly for gynoecy were identified from the information supported by different literatures. Genes of comparatively shorter length were given preference while selecting. The list of candidate genes selected were given in Table 2.

Table 2. List of candidate genes selected

S. No.	Candidate gene	Species/ crop	Reference
1	*ACO1*	*Momordica charantia*	Gunniah *et al. (*2014)
2	*ACS2*	*Momordica charantia*	Gunniah *et al. (*2014) and Kumari (2015)
3	*ACS3 (Acs1/ Acs4)*	*Momordica charantia*	Gunniah *et al. (*2014) and Shukla *et al. (*2015)
4	*ACS7*	*Momordica charantia*	Gunniah *et al. (*2014) and Shukla *et al. (*2015)
5	*AG6*	*Momordica charantia*	Peng *et al.* (2008)
6	*CsARF5*	*Cucumis sativus*	Shukla *et al*. (2015)
7	*MADS box23*	*Momordica charantia*	Peng *et al.* (2008) and Shukla *et al*. (2015)
8	*McAG2*	*Momordica charantia*	Peng *et al.* (2008) and Shukla *et al*. (2015)
9	*WRKY16*	*Cucumis sativus*	Shukla *et al*. (2015)
10	*WRKY37*	*Cucumis sativus*	Shukla *et al*. (2015)
11	*WRKY4*	*Cucumis sativus*	Shukla *et al*. (2015)
12	*WRKY9 WRKY10*	*Cucumis sativus*	Shukla *et al*. (2015)
13	*WRKY21*	*Cucumis sativus*	Shukla *et al*. (2015)
14	*WRKY23*	*Cucumis sativus*	Shukla *et al*. (2015)
15	*WRKY25*	*Cucumis sativus*	Shukla *et al*. (2015)
16	*WRKY49*	*Cucumis sativus*	Shukla *et al*. (2015)
17	*WRKY50*	*Cucumis sativus*	Shukla *et al*. (2015)
18	*WRKY51*	*Cucumis sativus*	Shukla *et al*. (2015)
19	*WRKY54*	*Cucumis sativus*	Shukla *et al*. (2015)
20	*WRKY56*	*Cucumis sativus*	Shukla *et al*. (2015)

3. 2. 2. Sequence retrieval of the candidate genes

The genomic nucleotide sequences of each of the candidate genes from their respective species were retrieved from the open access database by NCBI, GenBank. A search was done in the database with the gene name and the respective species name or using accession number. The gene sequence of each of the candidate genes selected from bitter gourd and cucumber were downloaded in FASTA format.

3. 2. 3. Primer designing

Primers for the downloaded sequences were designed using 'Primer3', the free online primer designing software (Untergasser *et al.*, 2012). All the primer designing parameters were kept default as provided by the software except the product size which was given in range of 850 - 1000 bp. Primer set for each of the genes was selected based on the optimum product size and length of the primers, melting temperatures, GC content and least dimerization capacity. The primer binding regions in the gene sequence were noted down and highlighted to identify the target region to be amplified by the each of the primer sets.

3. 2. 4. Isolation of total genomic DNA from bitter gourd

Young tender leaves of pale green colour (first to third leaves from the tip) from a typical plant of each of the genotypes were collected separately in a labeled aluminium foil and brought to laboratory by keeping it in an ice box. The samples were kept in -80ºC freezer when intended to use for a long time. The leaves were cleaned using sterile water and wiped using 70 per cent ethanol.

Extraction of total genomic DNA was performed by following CTAB method detailed by Rogers and Bendich (1994). The reagents used for DNA isolation are given in Annexure II.

Procedure

- 0.1g of clean leaf tissue was ground in a pre-chilled mortar and pestle in the

presence of liquid nitrogen

- Pre-warmed one mL CTAB (extraction) buffer along with a pinch of PVP (Poly Vinyl Pyrolidone) and 50 µL of β-mercapto ethanol were added to the mortar
- The homogenized sample was transferred into an autoclaved two mL centrifuge tube and 500 µL of CTAB buffer was added to the tube again
- All the added contents were mixed properly and incubated at 65 ºC for 20 to 35 minutes with occasional mixing and gentle inversion
- Equal volume of Chloroform: Isoamyl alcohol (24:1) was added and gently mixed by inversion to emulsify. The tubes were centrifuged at 10,000 rpm for 15 minutes at 4ºC
- The contents got separated into three distinct layers as the centrifugation was finished. The top-most aqueous layer contains DNA and RNA. The middle layer has proteins and fine particles. The bottom layer contains chloroform, pigments and cell debris
- A clean centrifuge tube was taken and the top aqueous layer was carefully pipetted into it. $1/10^{th}$ volume of 10 per cent CTAB solution and equal volume of chloroform: isoamyl alcohol (24:1) was added. The contents were mixed gently and centrifuged at 10,000 rpm for 15 minutes at 4 ºC
- The aqueous phase was transferred to a sterile centrifuge tube (two mL) and 2-3 µL of RNase was added, mixed well and incubated at 37 ºC for 1.5 hours. Equal volume of chloroform: isoamyl alcohol (24:1) was added and centrifuged at 10,000 rpm for 10 minutes at 4 ºC
- The supernatant was pipetted out carefully and transferred into a clean centrifuge tube (1.5µL) and 0.6 volume of chilled isopropanol was added. The tubes were gently inverted in horizontal position to observe the formation of threads and the contents were allowed to precipitate. For complete precipitation to happen, the tubes were kept in -20 ºC freezer for 30 minutes to one hour

- Once the precipitation is completed, the tubes were again centrifuged at 4 ºC, 10,000 rpm for 15 minutes
- The pellet obtained was washed with 70 per cent ethanol (200 µL) for 5 minutes at 10,000 rpm and the ethanol was decanted
- Subsequently, washing the pellet with 100 per cent ethanol (100 µL) for 3 minutes at 10,000 rpm was carried and decanted
- The pellet was air dried, dissolved completely in autoclaved distilled water (50 µL) and finally stored in –20 ºC for storage

3. 2. 5. Quality assessment by agarose gel electrophoresis

The quality of the isolated DNA was assessed by gel electrophoresis on a 0.8 per cent agarose gel (Sambrook *et al.,* 1989). The reagents and equipment used in agarose gel electrophoresis are presented in Annexure III.

Procedure for gel casting and electrophoresis

- For casting gel, the 60 ml gel casting tray was fixed in a gel casting apparatus and the comb was placed vertically at about one inch from one end of the tray such that gap of one to two mm will be present between the teeth of the comb and surface of the tray
- 0.8 per cent agarose gel was prepared by adding 0.48 g of agarose in 60 mL of 1X TAE in a conical flask and heated using a microwave oven for 45-60 seconds till it was dissolved completely and the solution becomes clear
- After the solution was allowed to cool to about 42-45 ºC at room temperature, 5 µL of ethidium bromide dye from a working concentration of 0.5µg/mL was added to the solution and mixed properly for uniform distribution. Formation of bubbles was avoided by carrying out gentle mixing
- The warm solution was carefully poured inside the casting tray to a depth of about 5-6 mm depth and allowed to cool and solidify for about 30-40 minutes
- The casting apparatus and the comb was removed cautiously after solidification of

gel. The gel tray was placed inside the gel tank of the horizontal electrophoresis unit with the wells of the gel towards cathode. The gel tank was filled with 1X TAE buffer solution to a point where all the wells were submerged

- Suitable molecular marker of 4 µL was loaded into the first well
- Samples to be loaded for electrophoresis was prepared by mixing one µL of 6X gel loading dye and 5 µL of DNA sample and the total 6 µL of the mixture was loaded into the each of the wells
- Tank was closed, anode and cathode were connected to the power pack and electrophoresis was carried out at a constant voltage of 70 V until the tracking dye reached two thirds of length of the gel
- Once the tracking dye reached two third the length of the gel, it was taken out of the unit and viewed under a UV trans-illuminator for the presence of DNA

3. 2. 6. Gel documentation

The gel containing the electrophoresed DNA was observed under a UV trans-illuminator for presence of DNA. The DNA fluoresces under UV light due to ethiduim bromide dye. Ethidium bromide binds to the hydrogen bonds between the double helix and this dye gets excited at a radiation of 260 nm and in the range of 590 nm wavelength, it re-emits yellowish orange coloured light which causes flourescence.

The electrophoresed gel was further analyzed and documented in the gel documentation system (BioRad Gel DOC-It™ imaging system) using Quantity one software. The gel was observed for intactness and clarity of DNA band and for the presence of RNA and protein. RNA will be present as a thick band with the size less than 100 bp. Protein will get stuck in the wells. Smear represents the presence of degraded DNA. The intact single band represents the good quality DNA.

3. 2. 7. Assessment of quality and quantity of DNA

The quantity and quality of the isolated DNA was analyzed using spectrophotometer **(**NanoDrop® ND-1000). Nucleic acid shows maximum absorbance at

the wavelength of 260 nm, whereas proteins displays highest absorbance at 280 nm. The absorbance were recorded at both the wavelengths and the purity was determined from A_{260}/A_{280}. A ratio value in between 1.8 and 2.0 indicates that DNA is of good quality and relatively free from proteins and RNA. A ratio less than 1.8 is indicative of protein contamination and a value of more than 2.0 is indicates the contamination with RNA. The following relation was used to calculate the quantity of DNA present in a pure sample.

$$1 \text{ OD at } 260 \text{ nm} = 50 \text{ µg DNA mL}^{-1}$$

Hence, the OD_{260} obtained from the sample X50 gives the total quantity of DNA in µg/mL.

Procedure

- NanoDrop spectrophotometer was connected to the system installed with the software, ND-1000
- The option 'nucleic acid' was selected. With the sampling arm lifted, the upper and lower measurement pedestal was wiped carefully and one µl of distilled water was pipetted out and placed onto the lower measurement pedestal
- Sampling arm was closed and the spectral measurement was initiated in the operating software
- The reading was set to zero with the respective blank sample
- The sampling arm was once again opened. Both the pedestals were wiped clean
- One µl of the DNA sample was pipetted and placed onto the measurement pedestal and the option 'measure' was selected
- Upon completion of measurement, the sampling arm and the upper and lower pedestals were thoroughly wiped using a soft laboratory wipe and the next sample was measured
- The results of all the sample were saved in the software, ND-1000

3. 2. 8. Optimization of amplifying conditions for the designed primers

The successful amplification of any primer depends on the proportion of the different components of a polymerase chain reaction in the reaction mixture. The reaction mixture consists of the DNA, which acts as the template for the amplification, an assay buffer which is necessary for maintaining optimum pH, $MgCl_2$, which acts as an important co-factor for the enzymatic activity of the polymerase, dNTPs, which are the building blocks of the PCR product, the primer, short stretch of oligonucleotides which determines the region to be amplified and the finally the enzyme, *Taq* polymerase which helps in extension of the primers by synthesizing long chains of nucleic acid.

Other important factors determining the effective amplification of DNA are the temperature and duration of various stages of a Polymerase Chain Reaction. Three main steps in PCR includes denaturation, annealing and extension. The samples were heated up to 92-98 ºC for separating two strands of the double helix in denaturation. During annealing the temperature of the samples was dropped down to about 48-78 ºC for the binding of primers to the complementary regions on the template DNA. Normally, the annealing temperature will be less than 68. In the final step of extension, the temperature of samples will be brought up to 72 ºC so that the polymerase can extend the primers for forming PCR products. The process will be repeated multiple times and the term 'cycle' denotes it.

3. 2. 8. 1. Gradient PCR for obtaining optimum annealing temperature

The nature of bases in the each primer sequences and its length decides the annealing temperature (T_a) of the primers in the thermal cycling conditions. Though T_a can be calculated from the melting temperature (T_m), it can only be an approximate value. The optimum annealing temperature provide effective amplification which could be achieved only by standardizing each newly designed primers.

A variant of the normal PCR, known as gradient PCR was used in optimizing the annealing temperature. In this, gradient of temperatures were applied for the samples in different tubes to find which temperature provides only the intended product with

maximum amplification with no non-specific products and minimum primer-dimers.

Gradient PCR was carried out as the first step of optimization for all the designed primer sets in Agilent SureCycler8800. The PCR amplification was carried out in a 20 µL reaction mixture in individual 0.2 mL PCR tubes. The standard composition of the mixture used are listed out below in the Table 3. The program used to perform PCR for most of the primer pairs are shown in the Table 4.

Table 3. Standard composition of PCR reaction mixture used for optimization

S. No	PCR components	Volume (µL)
1	Genomic DNA (50 ngµL^{-1})	2.0
2	10X *Taq* assay buffer A (with $MgCl_2$)	2.0
3	dNTP mix (2.5mM each)	1.8
4	Forward primer (100 pM)	1
5	Reverse primer (100 pM)	1
6	*Taq* DNA polymerase (3U/ µL)	0.4
7	Autoclaved ultrapure distilled water	11.8
	Total volume	20

Table 4. Standard PCR program used for optimization of primers

S. No.	Steps	Temperature	Time	Cycle
1	Initial denaturation	95 °C	5 minutes	-
2	Denaturation	95 °C	30 seconds	30
3	Primer annealing	Gradient range	30 seconds	
4	Primer extension	72 °C	1 minute	
5	Final extension	72 °C	8 minutes	
6	Hold	4 °C	Infinity	-

A gradient temperature range was set between the +5 ^{o}C and -5 ^{o}C of the average of calculated annealing temperature of the forward and reverse primers for a primer set.

Once the reaction was over, the entire 20 µL reaction mixture was mixed with 3 µL of 6X gel loading dye and was loaded into a 1.4 per cent agarose gel and electrophoresed. The optimum temperature for the reaction was obtained by analyzing the quality of the desired bands on the gel.

3. 2. 8. 2. Approaches adopted to increase specificity of primers

Once the gradient PCR approach failed to produce only the intended band of expected product size, many approaches were adopted for increasing the specificity of primers. To eliminate the formation of primer-dimer, the concentration of forward and reverse primers was reduced gradually up to 0.5 µM. The avoidance of non-specific binding was tried using changes in the volume of different PCR components. If *Taq* assay buffer B was used, the specificity was increased by gradually reducing the concentration of $MgCl_2$ from 1.8 µL. To improve the yield of the obtained band, the volume of *Taq* polymerse was increased gradually up to 0.6 µL. Use of touchdown PCR was also attempted for solving the specificity problem. Gel elution and purification after PCR were kept as a last option since only a low concentration of PCR product can be obtained.

3. 2. 8. 3. Touchdown PCR for optimizing primers

Touchdown PCR is a method known to increase specificity, sensitivity and the yield of the PCR products when the optimization becomes very difficult in some cases. It uses a cycling program where the initial annealing temperature will be several degrees more than the estimated (T_m) of the primers and it will be gradually reduced (1-2 °C/every second cycle), till the calculated T_a is reached or some degrees below. Amplification is then continued using this annealing temperature (Korbie and Mattick, 2008).

For standardizing the PCR conditions of the gene-specific primers ACS2, ACS3, ACS7 and AG6 touchdown PCR was used. The temperature increment of -0.5°C per cycle was set for the following program. Table 5 depicts the program used for touchdown PCR.

Table 5. Standard touchdown PCR program used in the study

S. No.	Steps	Temperature	Time	Cycles
1	Initial denaturation	95 ºC	3 minutes	-
Cycle wizard 1				15
2	Denaturation	94 ºC	20 seconds	
3	Primer annealing	65 ºC	20 seconds	
4	Elongation	72 ºC	30 seconds	
Cycle wizard 2				25
5	Denaturation	94 ºC	20 seconds	
6	Primer annealing	62 ºC	25 seconds	
7	Elongation	72 ºC	30 seconds	
9	Hold	4 ºC	Infinity	-

3. 2. 9. Gel elution and purification

Once all the standardization procedures failed to produce the required amplification for the given set of primers, gel elution and purification was adopted using QIAquick® gel extraction kit for obtaining only the desired band for sequencing using the protocol given in manual given with the kit.

The reagents in the kit used for carrying out gel extraction were listed in Annexure IV. Centrifugation (13,000 rpm) needed for gel elution was done in tabletop Eppendorf Centrifuge 5418R. The Rotek water bath used in the study was set at 50 ºC for incubation.

Procedure

- **Gel excision:** The surface of the UV-trans-illuminator where the gel was placed were wiped thoroughly using 70 per cent alcohol. The UV light was switched on once the electrophoresed agarose gel was placed. The band of desired size was marked cautiously using a sterile sharp scalpel

- After demarcation, the UV light was switched off, gels excised and transferred to fresh two mL centrifuge tubes for weighing. Three volumes of QG buffer was added to 1 volume gel (100 mg gel ~100 µL)
- The tubes were incubate at 50 °C for 10 min or until the gel slice has completely dissolved. Gentle inversion was given every 2–3 min to help dissolve gel
- One gel volume of isopropanol was added to the sample and mixed
- QIAquick spin column was placed in a provided two ml collection tube. The sample was applied to the QIAquick column to bind DNA and centrifuged for one min. The flow-through was discarded and the column was placed back into the same tube. If the sample volumes was more than 800 µL, the process can be repeated by splitting up the volume
- 500 µL QG buffer was added to the QIAquick column and centrifuged for one min or apply vacuum. Again the flow-through was discarded and column placed back into the same tube
- For washing, 750 µL Buffer PE (ethanol mixed) was added to column, incubated for 2-5 minutes, centrifuged for one minute for the residual wash buffer to be completely washed off from the column
- The column was placed into a clean 1.5 mL centrifuge tube for elution. 50µL Buffer EB (in split quantities) was added to the center of the QIAquick membrane, allowed to stand for 3 minutes and centrifuged for 1 min to collect the eluted product

The concentration of the eluted product was checked in NanoDrop^{R} ND-1000 spectrophotometer. If the concentration was found lower than 50 ng/µL, one more PCR was carried out using the eluted product as a template with the same amplifying conditions for the primer set and confirmed again using electrophoresis.

For *ACO1* gene, 1.4 per cent agarose gel, where the PCR product ran was kept in UV trans-illuminator to excise the gel containing 850 bp sized band. The excised gel

weighing 380 mg was used for gel elution using QIAquick® gel extraction kit. QIAquick column provided in the kit was used for eluting the DNA from solution of gel dissolved QG buffer. The eluted product is used as a template for another PCR for improved concentration. Annealing temperature was set at 62.5 ºC in the same PCR program with the same composition of reaction mixture, except the primer concentration of 0.7 µM per 20 µL reaction (1.225 µL) as mentioned in Table 6.

3. 2. 10. Sample preparation for sequencing

Once the optimum PCR amplifying conditions of each primers were standardized, the same conditions were applied in a separate PCR reaction for sequencing withincreased volume. Each reaction of PCR was carried out in a 0.2 mL tube. For proper mixing of components in reaction mixture a short spin was given before starting the run in thermal cycler. The reaction volume of 35 µL was prepared and the reaction mixture has the following composition as listed in Table 6.

Table 6. Composition for 35 µL volume reaction mixture

S. No	PCR components	Volume (µL)
1	Genomic DNA (50 ngμL^{-1})	3.5
2	10X *Taq* assay buffer A (with $MgCl_2$)	3.5
3	dNTP mix (2.5 mM)	3.15
4	Forward primer (100 pM)	1.75
5	Reverse primer (100 pM)	1.75
6	*Taq* DNA polymerase (3U/ µL)	0.7
7	Autoclaved ultrapure distilled water	20.65
	Total volume	35

The following PCR program mentioned in Table 7 was applied for the primers those optimum amplifying conditions were standardized and known.

Table 7. Standard PCR program used in this study

S. No.	Steps	Temperature	Time	Cycle
1	Initial denaturation	95 ºC	5 minutes	-
2	Denaturation	95 ºC	30 seconds	30
3	Primer annealing	Optimum temperature	30 seconds	
4	Primer extension	72 ºC	1 minute	
5	Final extension	72 ºC	8 minutes	
6	Hold	4 ºC	Infinity	-

The proper amplification in PCR samples was confirmed by loading 15 µL of the PCR product in a 1.4 per cent agarose gel, when the cycle gets completed. This was followed by electrophoresis at 80 V for one hour. The gels were analyzed in UV trans-illuminator for the presence of single intended band of expected size. If the desired amplification was achieved, the samples were stored in -20 ˚C till sequencing.

3. 2. 11. Sequencing, data analysis and SNP search

3. 2. 11. 1. Sequencing of PCR product

PCR products of all the primer sets for the monoecious and gynoecious samples were sequenced by outsourcing to AgriGenome Labs Pvt. Ltd., Kochi. Column purification was done to remove the chemical debris left before sequencing of all the PCR samples except for the eluted ones. Sequencing of the samples was carried out using Sanger sequencing method (Sanger *et al.*, 1977).

3. 2. 11. 2. Sequence data analysis

3. 2. 11. 2. 1. Assembly of contiguous sequence

The forward and reverse sequence of a primer set was assembled into one single contiguous sequence to represent sequence of a sample. 'CAP3' is an online based sequence assembly program, maintained by PRABI-Doua used for building contiguous sequences from a set of overlapping sequences (http://doua.prabi.fr/software/cap3) (Huang and

Madan, 1999). Contigs of all the 13 out of 14 samples sequenced were made using CAP3. The reverse complementary sequence of the obtained reverse sequence was inputed in the software along with the forward sequence to obtain correct contig. Manual assembly was done for the gynoecious sample sequence of the primer set AG6 by alignment and comparison with the reference sequence in NCBI.

3. 2. 11. 2. 2. Sequence variation analysis

The contigs made separately for the monoecious and gynoecious samples of each of the primer sets were compared along with the corresponding primer targeted region of the gene from reference genome given in the NCBI GenBank database. The reference sequences for the genes selected from *M. charantia* belongs to the genome assembly of a monoecious line, OHB3-1 and those from *C. sativus* belongs to the genome assembly of a monoecious cultivar**,** 9930. The comparison was done using alignment of the sequences under study using online multiple sequence alignment tool, Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), maintained by European Bioinformatics Institute (Goujon *et al.,* 2010). The software aligns the sequences based on the identity shared between the sequences.

The multiple sequence alignment of these three sequences for each set of gene-specific primers was also used for computationally confirming whether only the intended band was amplified. An asterisk (*) symbol was used to denote identical bases at any position and gaps ( ) denotes a non-identicals bases or any In/Dels, which were considered as variations. The monoecious and gynoecious sequences were compared with primer target region of reference gene sequence for each of the gene and were carefully observed for finding variations (gaps) and noted down by numbers. The deviations of the obtained gynoecious sample from the monoecious and reference sequence under consideration were only accounted for variation.

3. 2. 11. 2. 3. Confirmation of sequence variations

Sequencing error is common in the extreme starting and ending region of the trimmed sequence due to noise arise during signal pick up. To confirm whether the variations found by sequence alignment were valid SNPs or In/Dels, the chromatogram peaks of the sequence given in AB1 file provided within the sequencing data folder was carefully checked. The flanking regions (5-10 bases) of the variation was copied and searched in the AB1 file of respective run for the sample. When the variation was found towards the end of the sequence, the reverse complementary sequence of the flanking region was taken and searched in the reverse run chromatogram of the sample. Once confirmed the SNPs and In/Dels were given numbers for easy identification.

3. 2. 12. Designing primers for SNPs and In/Dels

The variations obtained in the putative candidate genes can be used as a molecular markers for differentiating monoecious line from gynoecy. This can be achieved by designing separate primers for each of the variations.

3. 2. 12. 1. SSR primers

The sequence having In/Dels repeats of 48 bp was used to design SSR primers. The online primer designing software 'Primer3' was used to design primers from the flanking regions of the repeats. The target region of AG6 gene-specific primers from NCBIsequence was used as input for designing SSR primers. The repeats in the input sequencewere given in bracket to denote it as must flank target region. All the primer designing parameters were kept default except product size range which was given as 250-300 bp.

3. 2. 12. 2. PCR amplification and validation

The PCR amplification with the designed SSR primer was carried with composition of reaction mixture mention in Table 3 and the program was ran as mentioned in Table 7 for both monoecious and gynoecious samples (base population). The PCR samples were ran on 2.5 per cent agarose gel. The SSR marker was finally screened with three other monoecious lines, Priya, Priyanka and the wild-type *M. charantia* var. *muricata.*

Results

4. RESULTS

The study entitled 'Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)' was carried out at Centre for Plant Biotechnology and Molecular Biology and Department of Vegetable Science, College of Agriculture, Vellanikkara, Thrissur, during 2018- 2021. The results of the study are presented in this chapter.

4.1. Planting material

The study was done using two bitter gourd genotypes *viz.* the monoecious, MC-136 which bears both male and female flowers (Reshmikka *et al.,* 2019) and the gynoecious line KAU-MCGy-101 (Minnu Ann, 2019), which bears only female flowers. The seeds started germinating after 5 days in water soaked germination paper. The pre-germinated seeds were sown on separate portrays inside a shade net and watered regularly. The true leaves started appearing after one week of sowing. And the third leaf stage was attained after 17 days of sowing. Leaves for DNA isolation were collected from these plants (Plate 1a).

4.2. Retrieval of gene sequences and primer designing

The DNA sequences of the putative candidate genes reported previously for their role in sex determination and female flower regulation were retrieved from the NCBI GenBank and downloaded in FASTA format.

4.2.1.a. Candidate gene sequences retrieval from bitter gourd

All the nucleotide sequences of genes corresponding to bitter gourd were retrieved from the previous scaffold level draft genome assembly of *Momordica charantia,* ASM19950v1. To get the location of the gene in chromosome, each of the sequences was subjected to BLAST with the recent scaffold level genome assembly, Mcharantia_2.0.

The gene sequence of *1-aminocyclopropane-1-carboxylate oxidase (ACO1)* retrieved was 2019 bp in length and located in between the region of 886071 – 888089 containing four exons in the scaffold 59. Subjecting to BLAST with the recent chromosome level assembly, Mcharantia_2.0 revealed, the location of the gene in chromosome 2.

The gene sequence of *1-aminocyclopropane-1-carboxylate synthase CMA101-like (ACS2)* retrieved was 1994 bp in length. It was present in between the region of 4277806-4279799 containing four exons in the scaffold 3. The gene was located in chromosome 8 in the new assembly.

The gene sequence of *1-aminocyclopropane-1-carboxylate synthase (ACS3*) retrieved was 2737 bp in length. It was present in between the region of 1273465-1276201 consisting five exons in the scaffold 8. The gene can also be represented as *ACS1 or ACS4*. It was located in chromosome 10 in new chromosome level assembly, Mcharantia_2.0.

The gene sequence of *1-aminocyclopropane-1-carboxylate synthase 7-like (ACS7*) was 2541 bp in length and was retrieved using accession number, NW_019105006.1. It was present in between the region of 13130-15670 consisting four exons in the scaffold 518. Location of this gene in chromosome 6 was revealed when subjected to BLAST with the recent chromosome level assembly, Mcharantia_2.0.

Agamous-like MADS-box protein AGL6 (AG6) gene sequences was retrieved using GenBank accession number DQ431247 and of 4443 bp in length. It was present in between the region of 321525-325967 consisting five exons in the scaffold 1.

The gene sequence of *MADS-box transcription factor 23 (MADSbox-23)* retrieved was 7624 bp in length. The gene was present in between the region of 1086555-1094178 consisting eight exons in the scaffold 55. The gene was found located in chromosome 10 of new genome assembly.

Floral homeotic protein AGAMOUS-like (McAG2) gene sequences was retrieved using GenBank accession number, DQ299943, It was 6969 bp in length and present in between the region of 539262-546230 consisting nine exons in the scaffold 20. The gene was located in chromosome 1 of recent genome assembly.

4.2.1.b. Candidate gene sequences retrieval from cucumber

All the nucleotide sequences of putative genes corresponding to cucumber were retrieved from the previous chromosome level genome assembly of *Cucumis sativus,* ASM407v2.

The gene sequence of *Auxin response factor 4 (CsARF5)* was 7263 bp in length and present in chromosome 6 in between the region of 14094436-14087174 consisting of 12 exons.

The gene sequence of *Probable WRKY transcription factor 2* (*WRKY16)* was retrieved using GenBank accession number E7CEW5 and was 4274 bp in length. The sequence was present in chromosome 3 in between the region of 3091238-3086515 consisting of 6 exons.

The gene sequence of *probable WRKY transcription factor 3-like (WRKY37)* was retrieved using GenBank accession number E7CEX8 and 4068 bp long. The sequence was present in chromosome 5 in between the region of 26517598-26513531 consisting of 4 exons.

The gene sequence of *probable WRKY transcription factor 32-like (WRKY4)* was 2000 bp in length and present in chromosome 1 in between the region of 12529832-12533031 consisting of 6 exonic regions.

Probable WRKY transcription factor 21 (WRKY9 WRKY10) sequence was retrieved using GenBank accession number, E7CEW1 and the length of the sequence was 2066 bp. It was present in chromosome 2 in between the region of 10906612-10904547 consisting of 3 exons.

The gene sequence of *probable WRKY transcription factor 40 (WRKY21)* was retrieved using GenBank accession number, E7CEW9 and the length of the sequence was

1798 bp. It was present in chromosome 3 in between the region of 26996761-26994964 comprising 5 exons.

The gene sequence of *probable WRKY transcription factor 26 (WRKY23)* was retrieved using GenBank accession number, E7CEX1 and the length of the sequence was 2948 bp. It was noted in chromosome 3 in between the region of 27354489-27351542 comprising 5 exons.

The gene sequence of *probable WRKY transcription factor 11-like* (*WRKY25*) was retrieved using GenBank accession number, E7CEX3 and the length of the sequence was 1614 bp. It was present in chromosome 3 in between the region of 37367541-37365928 consisting of 3 exons.

Probable WRKY transcription factor 20-like (*WRKY49*) sequence was retrieved using GenBank accession number, E7CEY3. The length of the sequence was 3781 bp and was found located on chromosome 7 in the region between 1865310-1861530 consisting 7 exons.

The gene sequence of *probable WRKY transcription factor 46-like (WRKY50*) was retrieved using GenBank accession number, E7CEY4 and the length of the sequence was 1577 bp. It was present in chromosome 7 in between the region of 11601148-11602724 consisting of 3 exons.

The gene sequence of probable WRKY transcription factor 21 (WRKY51) was retrieved using GenBank accession number, E7CEY5 and the length of the sequence was 2130 bp. It was present in chromosome 7 in between the region of 13868903-13871032 comprising 3 exons.

Probable WRKY transcription factor 51 (*WRKY54*) sequence was 3453 bp long and was located in chromosome 2 in between the region of 14138761-14135309 consisting of 3 exons.

Probable WRKY transcription factor 75 (*WRKY56*) sequence was retrieved by GenBank accession number, E7CEZ0. The total length of the sequence was 2556 bp and it was found located in chromosome 3 comprising 2 exons in between the region of 22244268-22247112.

The details of entire sequences of all putative candidate gene sequence retrieved were present in the Table 8.

4.2.2. Primer designing

Each of the retrieved gene sequences was inputted into Primer3. Since the processivity of *Taq* polymerase is limited to maximum of about 1500 bp, the product size range was entered as 850-1000 bp. The software had provided the details of the designed primers including the position at which the primer binds, the length of the primer, GC content, melting temperature, self-complementarity as well as the sequence of each of the primer in a set. First out of 5 primers sets provided by Primer3 was chosen for all the genes. Total of 20 primer pairs of optimal length, GC content, melting temperature, product size and least primer dimerization were designed, using Primer3 software (Table 9).

4. 3. Molecular characterization

4.3.1. Isolation of genomic DNA

Young tender, pale green leaves (1st to 3rd leaves from tip) were collected for DNA isolation (Plate 1b). CTAB method of DNA isolation by Rogers and Bendich (1994) was the protocol adopted to isolate total genomic DNA from bitter gourd. RNase treatment was given to remove RNA contamination in the samples.

Colourless pellet was obtained after precipitation. After washing thoroughly with ethanol and air drying, the pellet was dissolved in sterile distilled water of 50 µL by gentle tapping. The tube was kept in room temperature, till the pellet got fully dissolved in the added water.

Table 8. **Details of gene sequences retrieved from NCBI**

Sl. No.	Candidate Gene	Chromosome	Number of exons	Gene length (bp)
1	*ACO1*	2	4	2019
2	*ACS2*	8	4	1994
3	*ACS3*	10	5	2737
4	*ACS7*	6	4	2541
5	*AG6*	7	5	4443
6	*CsARF5*	6	12	7263
7	*MADS box23*	10	8	7624
8	*McAG2*	1	9	6969
9	*WRKY16*	3	6	4724
10	*WRKY37*	5	4	4068
11	*WRKY4*	1	6	2000
12	*WRKY9 WRKY10*	2	3	2066
13	*WRKY21*	3	5	1798
14	*WRKY23*	3	5	2948
15	*WRKY25*	3	3	1614
16	*WRKY49*	7	7	3781
17	*WRKY50*	7	3	1577
18	*WRKY51*	7	3	2130
19	*WRKY54*	2	3	3453
20	*WRKY56*	3	2	2556

Table 9. List of 20 gene-specific primer pairs designed

S. No.	Primer name	Forward primer	Tm (°C)	Product Size (bp)
1	*ACO1*	5'GGCATCATCCTTCTCTTCC3'	60.16	850
		5'CACAGCTTTCATGGCTTCAA3'	59.99	
2	*ACS2*	5'AAGAACCCCTACCACCCAAC3'	60.09	984
		5'ATCCGTTTGAACTCGGACAC3'	59.97	
3	*ACS3*	5'CGTCTACAGCCTCTCCAAGG3'	60.01	975
		5'GCAACTACACACAGCCTCCA3'	59.90	
4	*ACS7*	5'GTGGAGCTCTCCAAAGTTGC3'	60.00	920
		5CAGAGCCGGAGTAGATTTCG3'	59.97	
5	*AG6*	5'AAGATCCCCTCTCCTTCCAA3'	60.01	856
		5'GACCGACACATTTCGAACCT3'	59.97	
6	*CsARF5*	5'CAAAGAGCTGGAGGAGTTGG3'	59.98	981
		5'CTTTCACCACACGAGCAAGA3'	60.02	
7	*MADS box23*	5'GGCAGCTACCTTGAAGCAAC3'	60.02	1000
		5'AGGCCAGTGTTTTCATTTGG3'	59.97	
8	*McAG2*	5'TCCTCCCTCAGCTGCTTTTA3'	60.09	951
		5'ACGAAAAGGCACAACCAAAC3'	60.02	
9	*WRKY16*	5'CACCATCCTCAAGTGGGACT3'	59.96	888
		5'GGCCCCACTCATCTCTGTTA3'	60.07	
10	*WRKY37*	5'TCCTCGGGTCATCATTTTTC3'	59.87	996
		5'CAACATGCTTTCGGACATTG3'	60.11	
11	*WRKY4*	5'TTGGATTCATGGGTGAGGAT3'	60.13	953
		5'TTCAGCCTGCAAAAGAAGGT3'	59.99	
12	*WRKY9 WRKY10*	5'GCCATTGCTCAAAGAAGAGG3'	59.96	888
		5'CAATAAGCATCGACGGGTCT3'	60.10	
13	*WRKY21*	5'TCTGGTGAACCTCCGGTAAG3'	60.10	971
		5'TCTGAGATGGGTGTGGATGA3'	60.05	
14	*WRKY23*	5'TTCTGGTTTAAGCCCCACTG3'	60.10	909
		5'AACAATCCTCCTGCAACACC3'	59.97	
15	*WRKY25*	5'TCAGCCACAGACAAAAGTCG3'	60.02	877
		5'TCGGATGGTCTTCTTCATCC3'	60.01	
16	*WRKY49*	5'TCTTTGTTCGCCTCGAAGAT3'	59.96	851
		5'TCGATGCATTAGTTCCACCA3'	60.07	
17	*WRKY50*	5'CCTCCCATGGCATTCTCTTA3'	60.03	922
		5'GTTCGGAATGCAACCAAAGT3'	59.98	
18	*WRKY51*	5'GTGAAACGGTCGATCAAGGT3'	59.97	965
		5'CAAGGAAACCCACCTCTCAA3'	60.08	
19	*WRKY54*	5'TGTTTGGCGATGACAACAAT3'	59.97	852
		5'ACTACTCAAGCCGCTCCA3'	60.95	
20	*WRKY56*	5'TGGGGATGAAAAGAGAGGTG3'	60.04	946
		5'CTAACCAAACCCAAGCCAAA3'	59.97	

4.3.2. Assessment of quality and quantity of DNA

The quality and quantity of the isolated DNA were estimated using electrophoresis and NanoDrop^{R} ND-1000 spectrophotometer respectively. The DNA samples were loaded along with gel loading dye in 0.8 per cent agarose gel for electrophoresis. The samples yielded intact clear bright bands of high molecular weight indicating the good quality DNA (Plate 2a and 2b). $A_{260/280}$ absorbance ratio of all the DNA samples checked in spectrophotometer ranged in between 1.8 to 2.0, showing the quality of isolated DNA was good (Table 10).

Table 10. The quality and quantity of DNA estimated using NanoDrop^{R} spectrophotometer

Sample	Absorbance at 260nm (A_{260})	Absorbance at 280nm (A_{280})	Optical Density ($A_{260/280}$)	Quantity of DNA (ng/µl)
MC 136	11.44	5.90	1.92	763.69
KAUMCGy-101	10.85	5.87	1.84	666.98
Priya	76.83	41.64	1.84	2134.78
Priyanka	41.38	21.74	1.90	1528.66
M. charantia var. *muricata*	34.35	17.61	1.95	1023.15

4.3.3. Optimization of amplifying conditions for the designed primers

4.3.3.1. ACO1

When gradient temperature range of 55-65 ºC was set for annealing of primer set *ACOF/ACOR*, two clear distinct bands along with primer dimer was obtained in all temperatures till 63 ºC, after which no bands appeared. One of the band was of 850 bp (expected product size) and the extra non-intended band also appeared with same intensity

Plate 1(a). Bitter gourd plants grown for DNA isolation

Plate 1(b). Bitter gourd leaves samples collected for DNA isolation

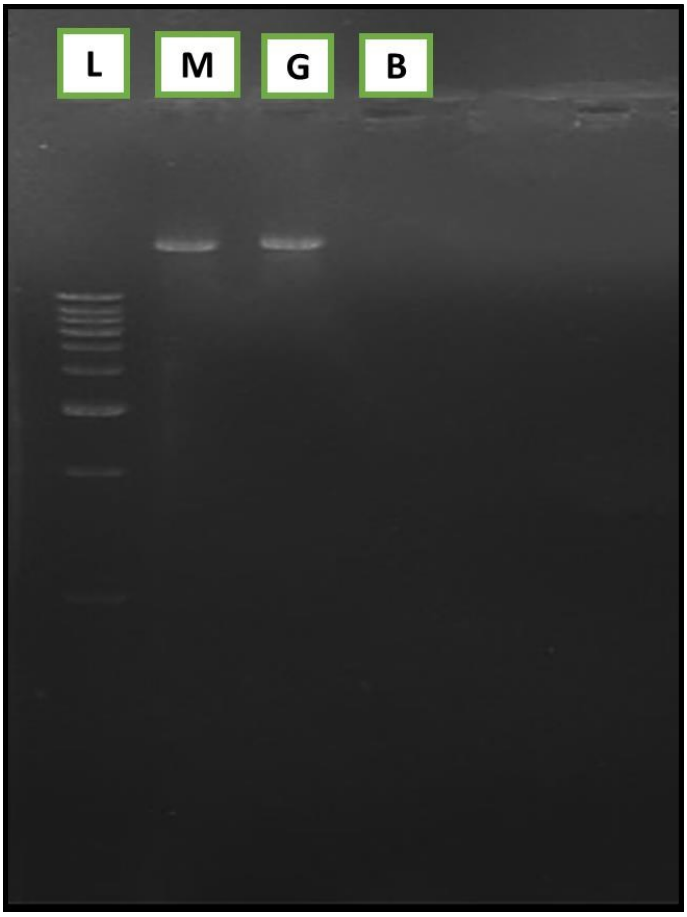


Plate 2(a). Genomic DNA isolated from bitter gourd

L - 1 kb DNA ladder

M - DNA isolated from monoecious genotype (MC-136)

G - DNA isolated from gynoecious genotype (KAUMCGy-101)

B - Blank

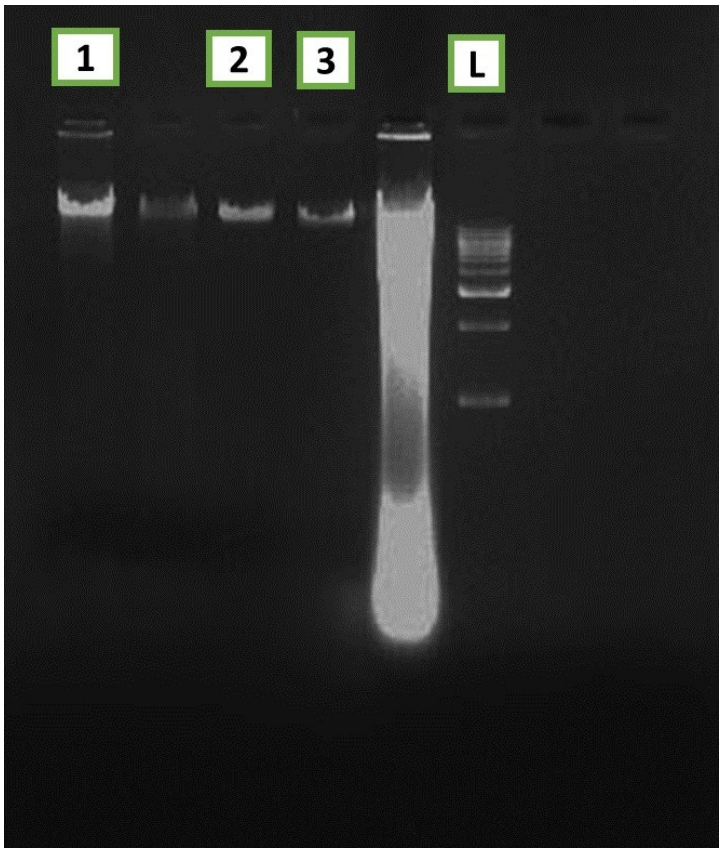


Plate 2(b). Genomic DNA isolated from bitter gourd for validation of marker

L - 1 kb DNA ladder

1 - Priya

2 - Priyanka

3 - *M. charantia* var. *muricata*

as that of intended one. The concentration of each primer was gradually reduced up to 0.7 µM where no primer-dimer was observed. To prevent non-specific binding of the primers, the template volume was increased to 2.5 µL which also failed in elimination of the extra band. The $MgCl_2$ concentration in the reaction mixture was reduced up to 1.3 µM when *Taq* assay buffer B was used, which had only decreased the intensity of both the bands.

Gel elution was adopted as a last approach to extract only the band of expected size. The QG buffer remained as the same yellow coloured clear solution after dissolving of the gel, indicating optimum pH for elution. QIAquick column provided in the kit was used for eluting the DNA from solution of gel dissolved QG buffer.

The concentration of the final eluted product was analyzed using NanoDrop^R ND-1000 spectrophotometer, which showed 35 ng/ µL. Since the minimum concentration of eluted product required for sequencing was 50 ng/ µL, another one PCR was carried out with the eluted product as DNA template. It had resulted in a clear single band of 850 bp with no primer dimer for both monoecious and gynoecious samples (Plate 3).

4.3.3.2. ACS2

As the first step of optimization, gradient PCR with the temperature range of 55-65 ºC was set for annealing of the gene-specific primer set *ACS2F/ACS2R* to the template. The gel image have shown the clear distinct band of 984 bp along with two non-specific bands and primer-dimers only up to 62 ºC after which faint intended band were observed. 61.5 ºC was found as the optimum annealing temperature. The primer-dimers was completely prevented upon gradual reduction of primer concentration to 0.6 µM.

Touchdown PCR was carried on with the annealing temperature of the first cycle kept at 65 ºC, the temperature increment of -0.5 ºC/ cycle and the second cycle at 61.5 ºC using the program mentioned in Table 5. The gel image of these samples showed a clear distinct band of 984 bp with no non-specific products and primer-dimers in both the monoecious and gynoecious genotypes (Plate 4).

4.3.3.3. ACS3

The gradient temperature range of 55-65 ºC was set for annealing of the gene-specific primer set *ACS3F/ACS3R* to the template. The intended band of 975 bp was observed as a clear distinct band along with primer dimer and non-specific bands up to 62.5 ºC, after which the band started fainting. The annealing temperature of 62 ºC was found optimum. The optimum concentration of primer was found to be 0.8 µM for complete elimination of the primer-dimers.

Touchdown PCR by using the PCR program mentioned in Table 5 was carried out. 1.4 per cent gel image of these PCR products showed a clear intact band of 975 bp alone for both the monooecious and gynoecious samples (Plate 5).

4.3.3.4. ACS7

The gradient temperature range of 55-65 ºC was set for annealing of the gene-specific primer set *ACS7F/ACS7R* to the template. The intended band of 920 bp was observed as a clear distinct band along with two closely located non-intended bands and primer dimer up to 62.5 ºC, after which the band started fainting. The annealing temperature of 62 ºC was found to be optimum for the band of expected size. The optimum primer concentration was found to be 0.8 µM for the complete elimination of primer-dimers.

Touchdown PCR by using the PCR program mentioned in Table 5 was carried out. The gel image of these PCR products showed a clear intact band of 920 bp alone for both the monooecious and gynoecious samples (Plate 6).

4.3.3.5. AG6

Gradient PCR was carried out with the annealing temperature range of 55-65 ºC for both the monoecious and gynoecious samples with primer set *AG6F/ AG6R*. Light bands of 856 bp were observed upto 62 ºC after which no bands were seen. 62 ºC was considered optimum since, no mis-priming products were observed and the intended band was of good intensity comparatively. When the number of cycle was increased to 32 to improve the

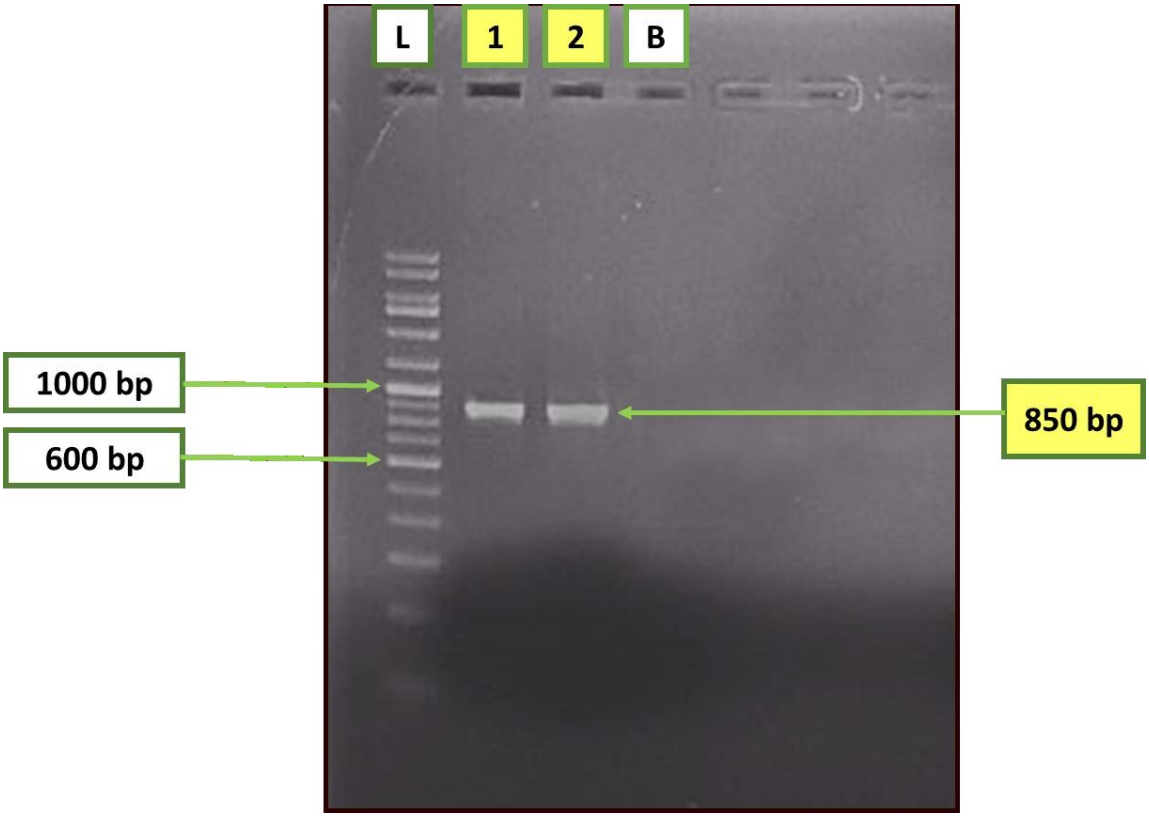


Plate 3. Amplification of *ACO1* gene

L - 1 Kb plus DNA ladder

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

B - Blank

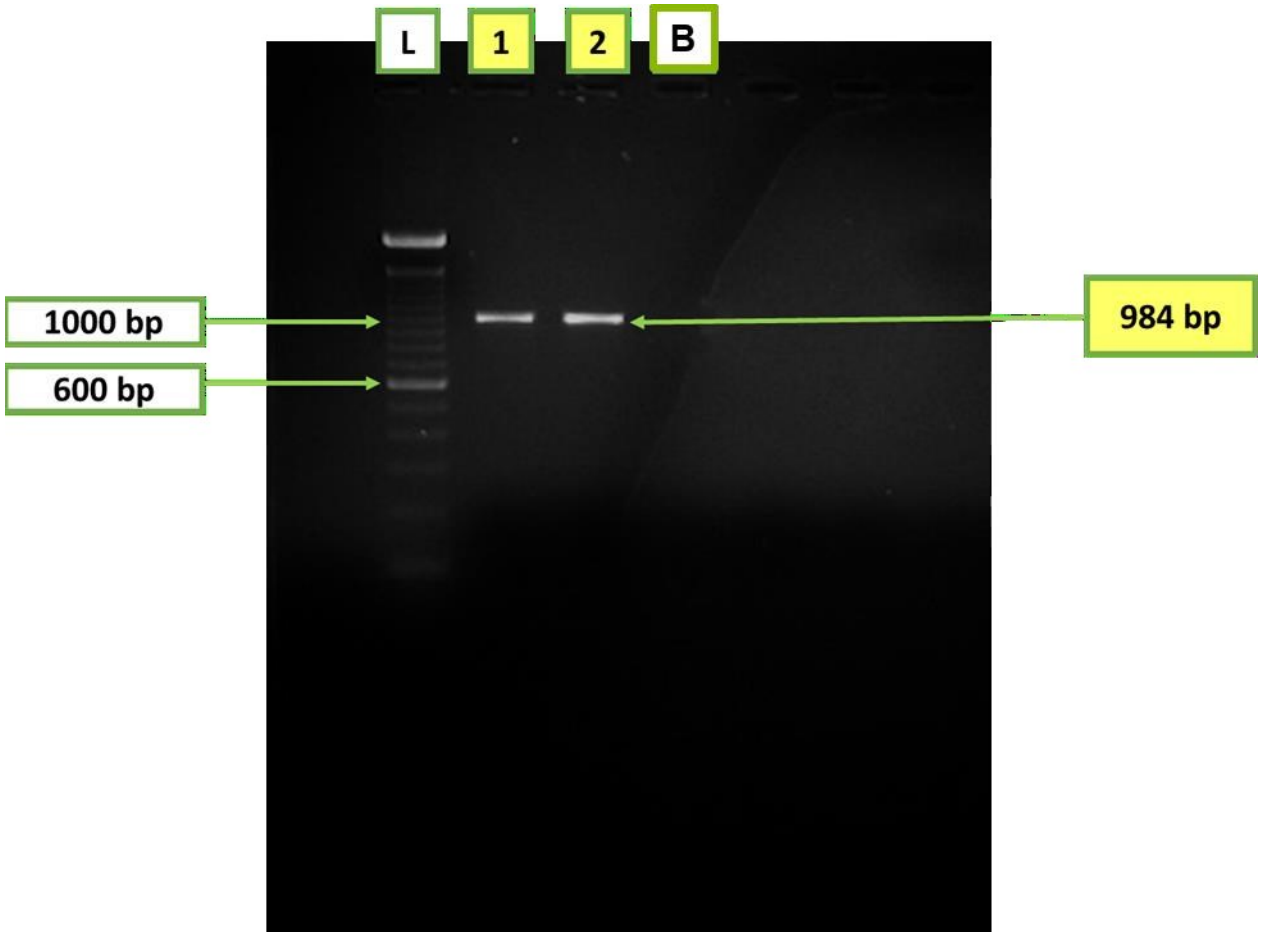


Plate 4. Amplification of *ACS2* gene

L - 1 Kb low range DNA ruler

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

B - Blank

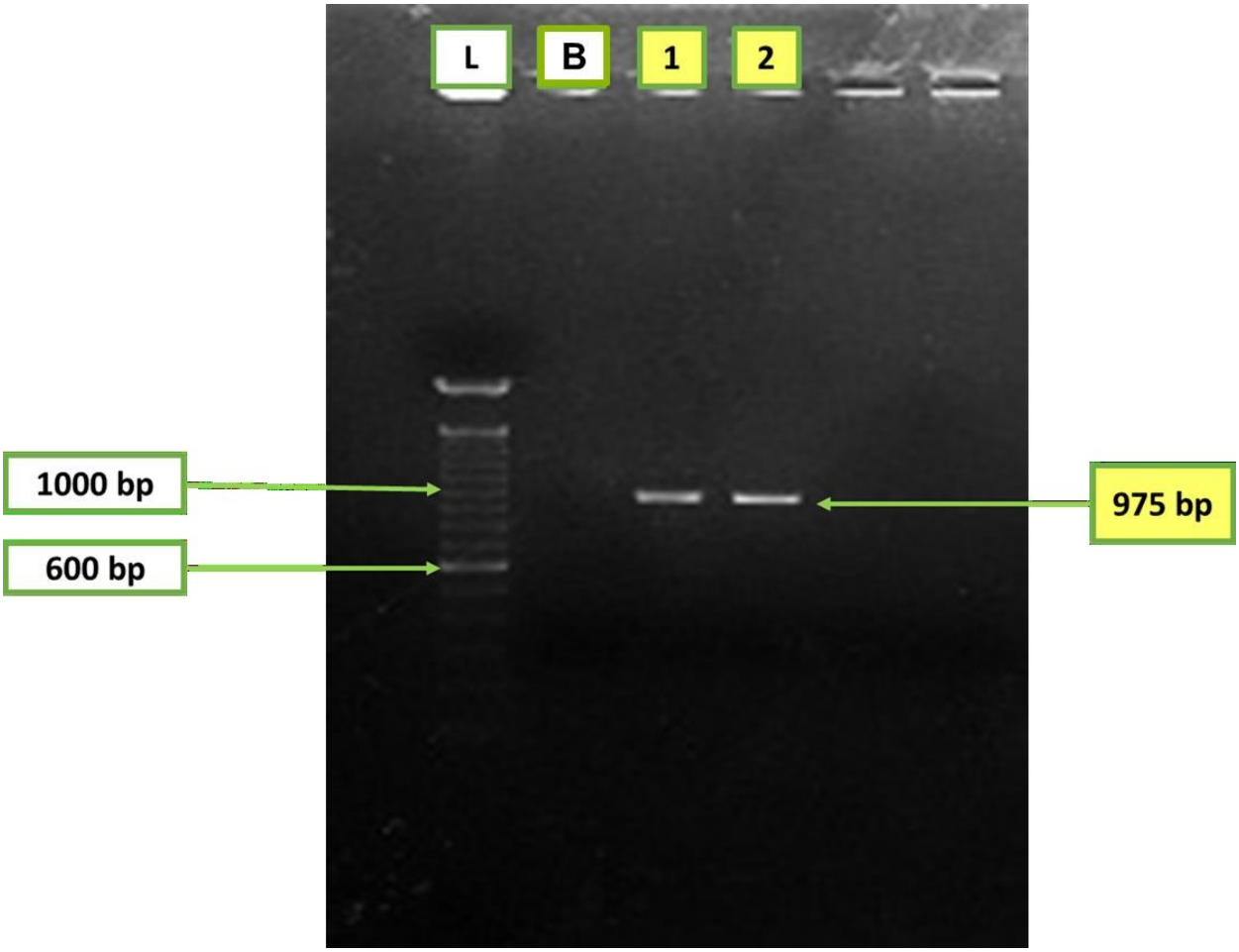


Plate 5. Amplification of *ACS3* gene

L - 1 Kb low range DNA ruler

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

B - Blank

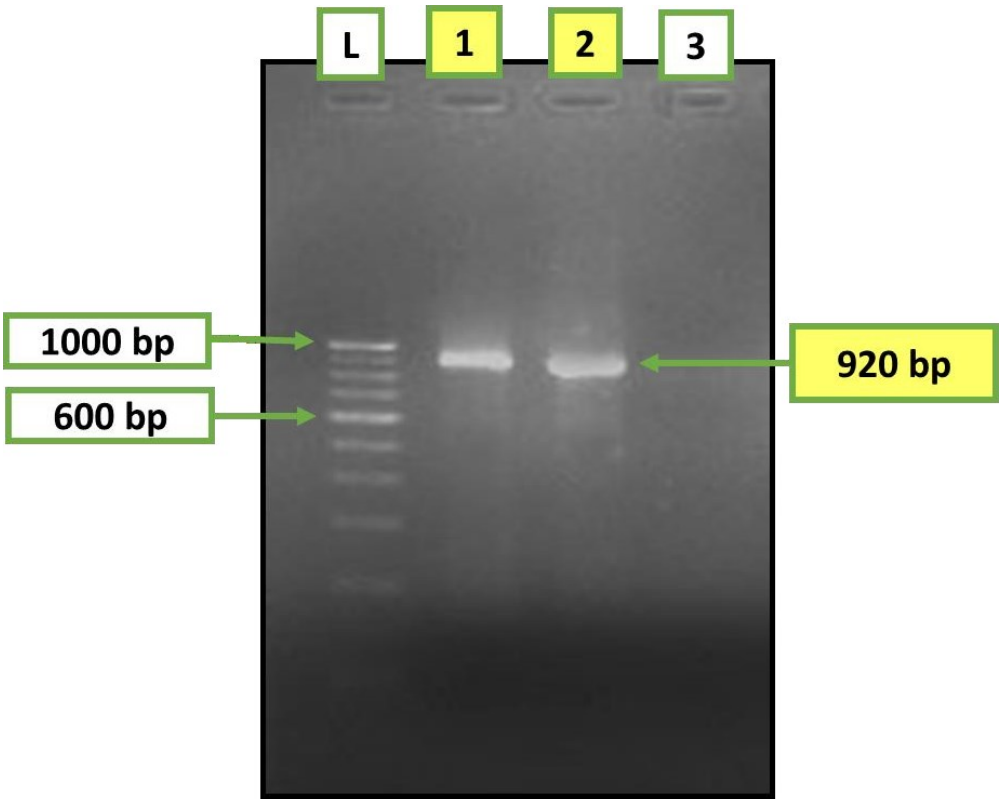


Plate 6. Amplification of *ACS7* gene

L - 100 bp DNA ladder

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

3 - Blank

thickness of bands, many faint non-reproducible bands close to the intended band occurred in both the samples, especially in gynoecious line sample.

Touchdown PCR with the annealing temperature of the first cycle set at 69 ºC (temperature increment of -0.5 ºC/ cycle) and the second one set at 62 ºC following the PCR program in Table 5 was performed. The clear intense intact bands of 856 bp was observed without any primer-dimer or mis-priming products for both samples (Plate 7).

4.3.3.6. MADS-box 23

To obtain the optimum annealing temperature for the primer sets *MADS-box 23F/ MADS-box 23R,* gradient temperature range of 55-65 ºC was set using the PCR program mentioned in Table 3. Optimum temperature found was 61 ºC in all aspects without non-specific and primer-dimer product with good yield of intended band of 1000 bp. The exact composition of reaction mixture mentioned in Table 3 worked perfectly for the samples of both sex types under investigation (Plate 8).

4.3.3.7. McAG2

To obtain the optimum annealing temperature for the primer sets *McAG2F/ McAG2R,* the gradient range of temperature, 55-65 ºC was set using the PCR program mentioned in Table 3. Optimum temperature to obtain the specific band was 61 ºC, but the presence of faint primer-dimer was observed. The exact composition of reaction mixture mentioned in Table 3 except the primer concentration of 0.8 µM worked well for the samples of both sex types under study (Plate 9).

4.3.3.8. Gene-specific primers designed from cucumber

All the gene-specific primer sets designed from the gene sequences of*Cucumis sativus* failed to amplify the bands of intended size even with varied range of temperatures, different composition of reaction mixture and the wide range of PCRamplification conditions. *CsARF5F/ CsARF5R* have shown very faint band, which was notin the range of expected product size.

4.3.4. Amplification, sequencing and sequence analysis of the genes

4.3.4.1. PCR amplification and sequencing of the genes

Out of 20 sets of gene-specific primers designed for 20 putative candidate genes selected, 7 sets designed from *Momordica charantia* amplified properly. 13 primer sets designed from the gene sequences of *Cucumis sativus* did not showed any amplification. The details of PCR conditions of primer pairs optimized and given for sequencing are represented in Table 11.

Each of the seven primer sets was used for amplification of monoecious (MC 136) and gynoecious (KAUMCGy-101) genotypes of bitter gourd. Thus total of 14 samples were sequenced.

Table 11. PCR conditions for optimized primer sets

Gene	Annealing temperature (^{o}C)	Optimum primer concentration (µM)	Product size (bp)
ACO1	62.5	0.7	850
ACS2	61.5	0.6	984
ACS3	62	0.8	975
ACS7	62	0.8	920
AG6	62	1	856
MADS-box 23	61	1	1000
McAG2	61	0.8	951

4.3.4.2. Sequence analysis

Sequencing of all the 14 samples were done by using Sanger method (Sanger *et al.*, 1977). The length and quality of the all the sequences obtained were thoroughly analyzed. The trimmed length of both forward sequence and reverse sequence were observed. All the sequences obtained had Quality Value (QV) score more than 20 indicated that sequences were of good quality and can be relied. The forward and reverse sequences obtained for each sample were assembled into a single contiguous sequence using 'CAP3'software. Multiple sequence alignment for the monoecious, gynoecious and the

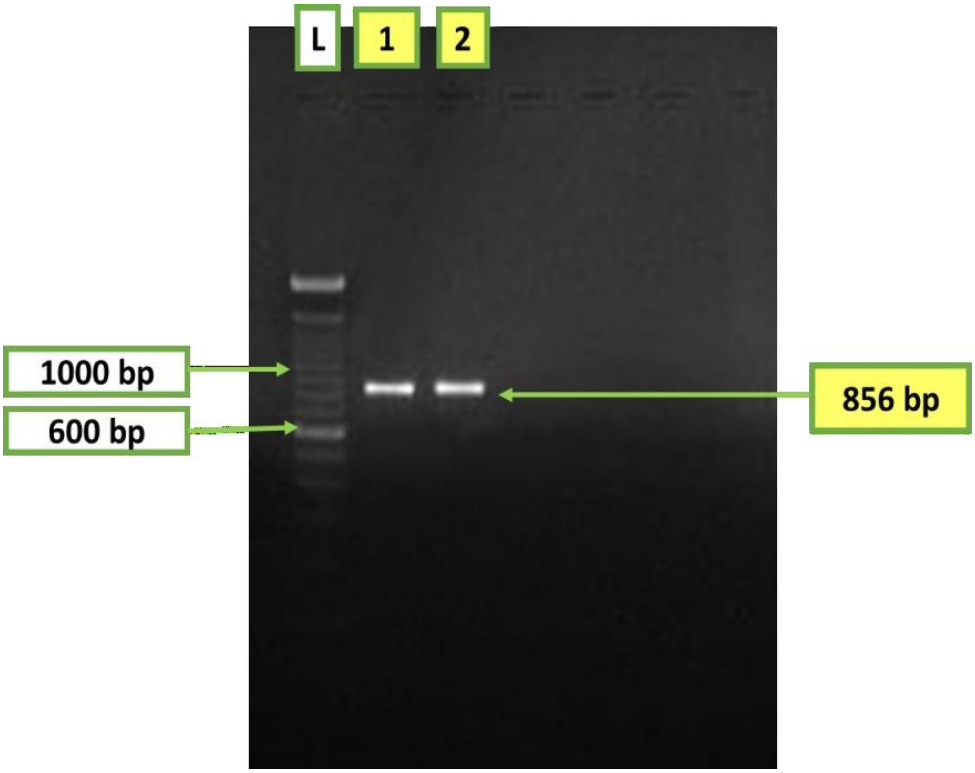


Plate 7. Amplification of *AG6* gene

L - 1kb low range DNA ruler

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

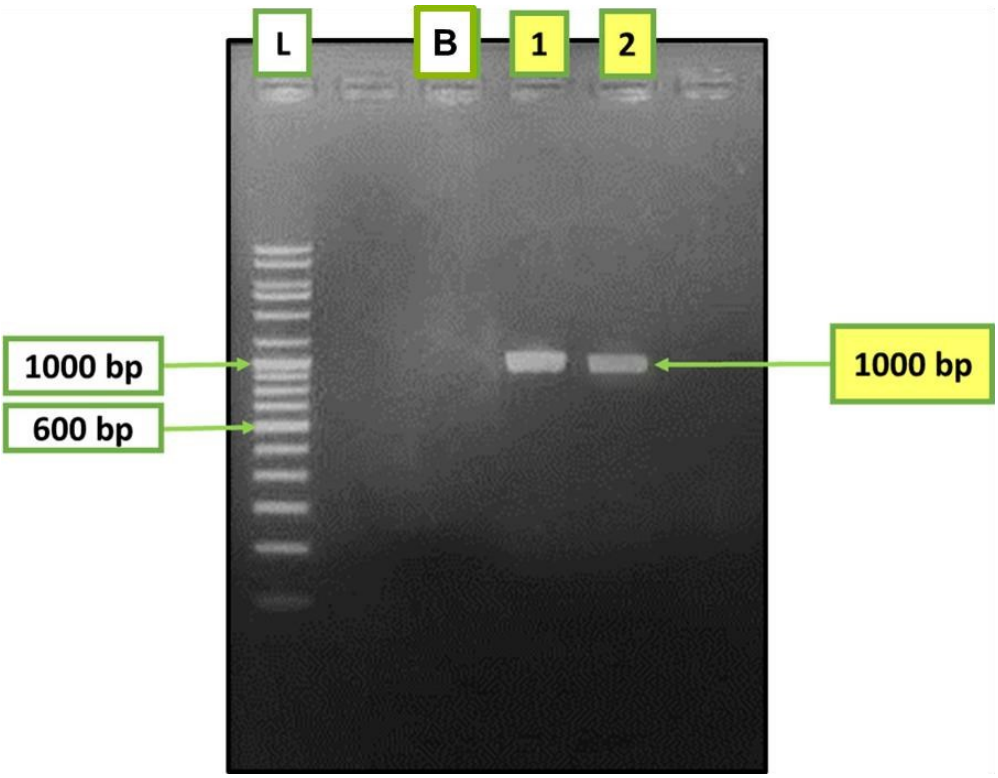


Plate 8. Amplification of *MADS-Box TF23* gene

L - 1 Kb plus DNA Ladder

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

B - Blank

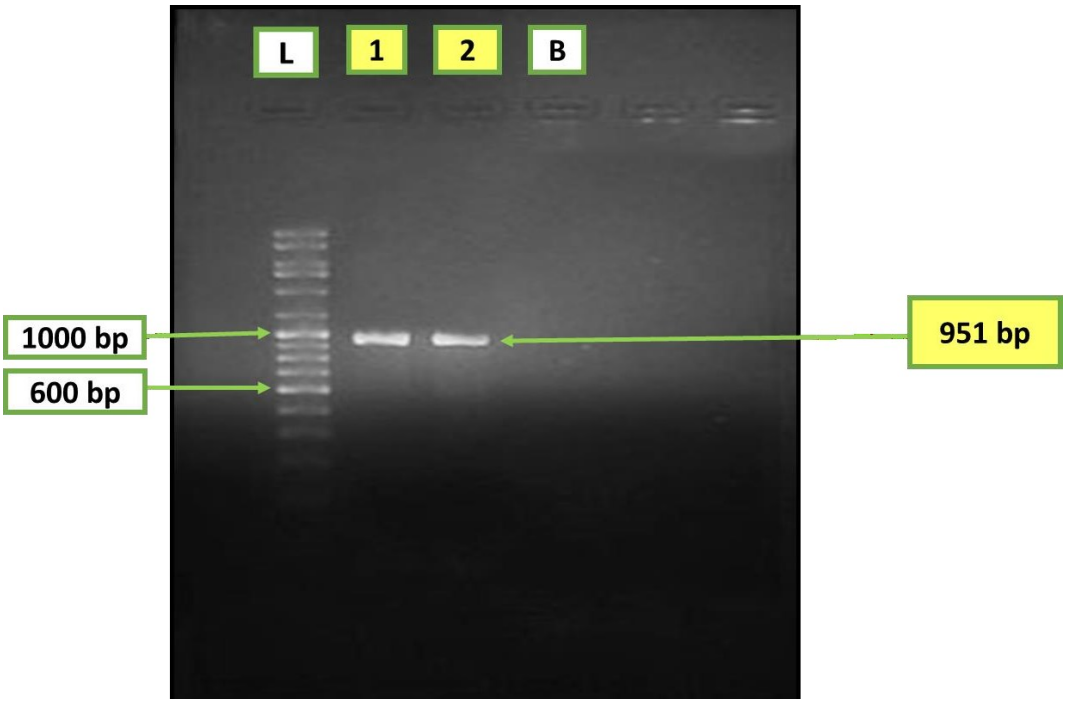


Plate 9. Amplification of *McAG2* gene

L - 1 kb plus DNA ladder

1 - Amplification in monoecious line (MC-136)

2 - Amplification in gynoecious line (KAUMCGy-101)

B - Blank

reference sequence for each primer set were done using 'clustal omega' for computational confirmation of the obtained sequences and for analyzing the variations.

4.3.5. Candidate genes characterized

4.3.5.1. 1-aminocyclopropane-1-carboxylate oxidase (ACO1)

4.3.5.1.1. Amplification of the primer

The primer set (*ACO1*) was used to amplify the region of 846 to 1695 in the gene, targeting 850 bp. An annealing temperature of 62.5 ºC was set to run PCR. The optimum concentration of forward and reverse primers of reaction mixture was found to be 0.7 µM. Agarose gel electrophoresis in 1.4 per cent have shown a single band of 850 bp after a gel elution and a PCR was ran again with the eluted product as template (Table 11). The gel image is portrayed in Plate 3. The PCR product was sequenced.

4.3.5.1.2. Analysis of the sequences

In monoecious sample, trimmed sequence of the forward run was of 636 bp and had a QV score of 48. The length of the reverse sequence was of 701 bp and the QV score was 48 (Table 12). In gynoecious sample, trimmed sequence of the forward run was of 755 bp and had a QV score of 48. The length of the reverse sequences was of 753 bp and QV score was 50 (Table 12).

4.3.5.1.3. Contig assembly

The contig obtained after assembling the forward and reverse sequence of*ACO 1* primers was 746 bp long for monoecious sample and 780 bp for gynoecious sample (Table 12).

Table 12. Details of gene sequences obtained

Gene	Sample	Sequence	Length (bp)	Q.V Score	Contig length (bp)
ACO1	Monoecious	Forward	636	48	746
		Reverse	701	48	
	Gynoecious	Forward	755	48	780
		Reverse	753	50	
ACS2	Monoecious	Forward	586	42	-
		Reverse	0	0	
	Gynoecious	Forward	724	55	-
		Reverse	25	39	
ACS3	Monoecious	Forward	674	49	855
		Reverse	384	38	
	Gynoecious	Forward	841	54	841
		Reverse	753	56	
ACS7	Monoecious	Forward	785	52	820
		Reverse	819	53	
	Gynoecious	Forward	803	51	833
		Reverse	832	52	
AG6	Monoecious	Forward	589	50	724
		Reverse	597	32	
	Gynoecious	Forward	561	50	720
		Reverse	411	36	
MADS - box TF23	Monoecious	Forward	913	52	914
		Reverse	865	52	
	Gynoecious	Forward	889	52	896
		Reverse	838	51	
McAG2	Monoecious	Forward	800	48	847
		Reverse	812	53	
	Gynoecious	Forward	853	49	855
		Reverse	845	52	

4.3.5.1.4. Multiple sequence alignment

Sequences of monoecious and gynoecious samples were aligned with the primer target region from reference sequence of NCBI (monoecious) using Clustal Omega. No variation was found between the sequences of monoecious and gynoecious samples (Fig.1a, 1b and 1c).

4.3.5.2. 1-aminocyclopropane-1-carboxylate synthase CMA101-like (ACS2)

4.3.5.2.1. Amplification of the gene

The primer set (*ACS2*) was used to amplify the region of 82 to 1065 in the gene, targeting 984 bp. An optimum annealing temperature of 61.5 ºC was set in the touchdown PCR program ran. The optimum concentration of forward and reverse primer in the reaction mixture was found to be 0.6 µM. A single intact band of 984 bp was observed upon electrophoresis. The gel image is portrayed in the Plate 4. The PCR product was sequenced.

4.3.5.2.2. Analysis of the sequences

In monoecious sample, trimmed sequence of the forward run was of 586 bp and had a QV score of 42. The reverse sequence obtained was nil (Table 12). In gynoecious sample, trimmed sequence of the forward run was of 724 bp and had a QV score of 55. The length of the reverse sequence was of 25 bp and the QV score was 39 (Table 12).

4.3.5.2.3. Assembly of contiguous sequence

No contig was obtained for both monoecious and gynoecious sample of *ACS2* primer, as reverse run of the former was failed and the latter yielded only 25 bp after trimming.

4.3.5.2.4. Multiple Sequence alignment

The forward sequences of both the monoecious and gynoecious samples were aligned with the reference sequence from NCBI (monoecious) to find out variations. No valid variation was found in the gynoecious line compared to both monoecious genotype (Fig. 2a, 2b and 2c).

4.3.5.3. *1-aminocyclopropane-1-carboxylate synthase (ACS3*)

4.3.5.3.1. Amplification of the gene

The primer set (*ACS2*) was used to amplify the region of 1576 to 2550 in the gene, targeting 975 bp. An optimum annealing temperature of 62 ºC was set in the touchdown PCR program ran. The optimum concentration of forward and reverse primer in the reaction mixture was found to be 0.8 µM. A single intact band of 975 bp was observed upon electrophoresis in 1.4 per cent agarose gel. The gel image was portrayed in the Plate **5.** The PCR product was sequenced.

4.3.5.3.2. Analysis of the sequences

In the monoecious sample the trimmed sequence of the forward run was of 674 bp and had a QV score of 49. The length of reverse sequence was 384 bp with the QV score of 38. In the gynoecious sample the trimmed sequence of the forward run was of 841 bp and had a QV score of 54. The length of reverse sequence was 753 bp with the QV score of 56.

4.3.5.3.3. Assembly of contiguous sequence

The length of contig obtained after assembling the forward and reverse run of *ACS3* primer for monoecious sample was of 855 bp. The length of contig obtained after assembling the forward and reverse run of *ACS3* primer for gynoecious sample was of 841 bp.

4.3.5.3.4. Multiple Sequence alignment

The contiguous sequences of both the monoecious and gynoecious samples for *ACS3* primer sets were aligned with the reference sequence from NCBI (monoecious) to find out variations. No variation was found in the sequence of gynoecious line compared to the sequences of both monoecious genotype (Fig. 3a, 3b and 3c).

Figure 1(a). Multiple sequence alignment of *ACO1*

Figure 1 (b). Multiple sequence alignment of *ACO1*

Figure 1 (c). Multiple sequence alignment of *ACO1*

Figure 2 (a). Multiple sequence alignment of *ACS2*

Figure 2 (b). Multiple sequence alignment of *ACS2*

Figure 2 (c). Multiple sequence alignment of *ACS2*

Figure 3 (a). Multiple sequence alignment of *ACS3*

Figure 3 (b). Multiple sequence alignment of *ACS3*

Figure 3 (c). Multiple sequence alignment of *ACS3*

4.3.5.4. *1-aminocyclopropane-1-carboxylate synthase 7-like (ACS7*)

4.3.5.4.1. Amplification of the gene

The primer set (*ACS7*) was used to amplify the region of 91 to 1010 in the gene, targeting 920 bp. An optimum annealing temperature of 62 ºC was set in the touchdown PCR program ran. The optimum concentration of forward and reverse primer in the reaction mixture was found to be 0.8 µM. A single intact band of 920 bp size was observed upon electrophoresis in 1.4 per cent agarose gel (Table 12). The gel image was portrayed in the Plate 6. The PCR product was sequenced.

4.3.5.4.2. Analysis of the sequences

In the monoecious sample, the trimmed length of the forward run was of 785 bp and had a QV score of 52. The length of reverse sequence was 819 bp with the QV score of 53. In the gynoecious sample the trimmed sequence length of the forward run was of 803 bp and had a QV score of 51. The length of reverse sequence was 832 bp with the QV score of 52 (Table 12).

4.3.5.4.3. Assembly of contiguous sequence

The length of the contig obtained after assembling the forward and reverse run of the monoecious sample for *ACS7* primer was of 820 bp and of gynoecious sample was 833 bp.

4.3.5.4.4. Multiple sequence alignment

The obtained monoecious and gynoecious sequences for the primer *ACS7* was compared with the reference sequence from NCBI which also belongs to a monoecious sample of bitter gourd. One base variation was found in the sequence of gynoecious sample (G) compared to both monoecious (C) (Fig. 4a, 4b and 4c). On confirmation with chromatogramic peaks, it was found as an error (Fig. 4d). No valid variations were found in the obtained gynoecious line sequence compared to both the monoecious genotype.

4.3.5.5. Agamous-like MADS-box protein AGL6 (AG6)

4.3.5.5.1. Amplification of the gene

The primer set (*AG6*) was used to amplify the region of 22 to 882 in the gene, targeting 856 bp. An optimum annealing temperature of 62 ºC was set in the touchdown PCR program ran. The optimum concentration of forward and reverse primer in the reaction mixture was found to be 1 µM. A single intact band of 856 bp size was observed upon electrophoresis in 1.4 per cent agarose gel (Table. 12). The gel image is portrayed in the Plate 7. The PCR product was sequenced.

4.3.5.5.2. Analysis of the sequences

In the monoecious sample, the trimmed length of the forward run was of 589 bp and had a QV score of 50. The length of reverse sequence was 597 bp with the QV score of 32. In the gynoecious sample the trimmed sequence length of the forward run was of 561 bp and had a QV score of 50. The length of reverse sequence was 411 bp with the QV score of 36 (Table 12).

4.3.5.5.3. Assembly of contiguous sequence

The length of the contig obtained by assembling the forward and reverse run of the monoecious sample for *AG6* primer in CAP3 was of 724 bp. The software failed to provide contiguous sequence for the forward and reverse run of the gynoecious sample, as the overlapping region between the two sequences was too low to be assembled. So the assembly was done manually by aligning the forward and reverse run sequence with the reference sequence in NCBI (monoecious).

4.3.5.5.4. Multiple sequence alignment

The obtained monoecious and gynoecious sequences for the primer *AG6* was compared with the reference sequence from NCBI which also belongs to a monoecious genotype of bitter gourd (Fig.5a, 5b and 5c). Total of three SNPs and two In/Del variations were found in the obtained gynoecious sample sequence compared to the sequence of both monoecious genotype (Table 13)**.** Each of these variations were compared with the

Figure 4 (a). Multiple sequence alignment of *ACS7*

Figure 4 (b). Multiple sequence alignment of *ACS7*

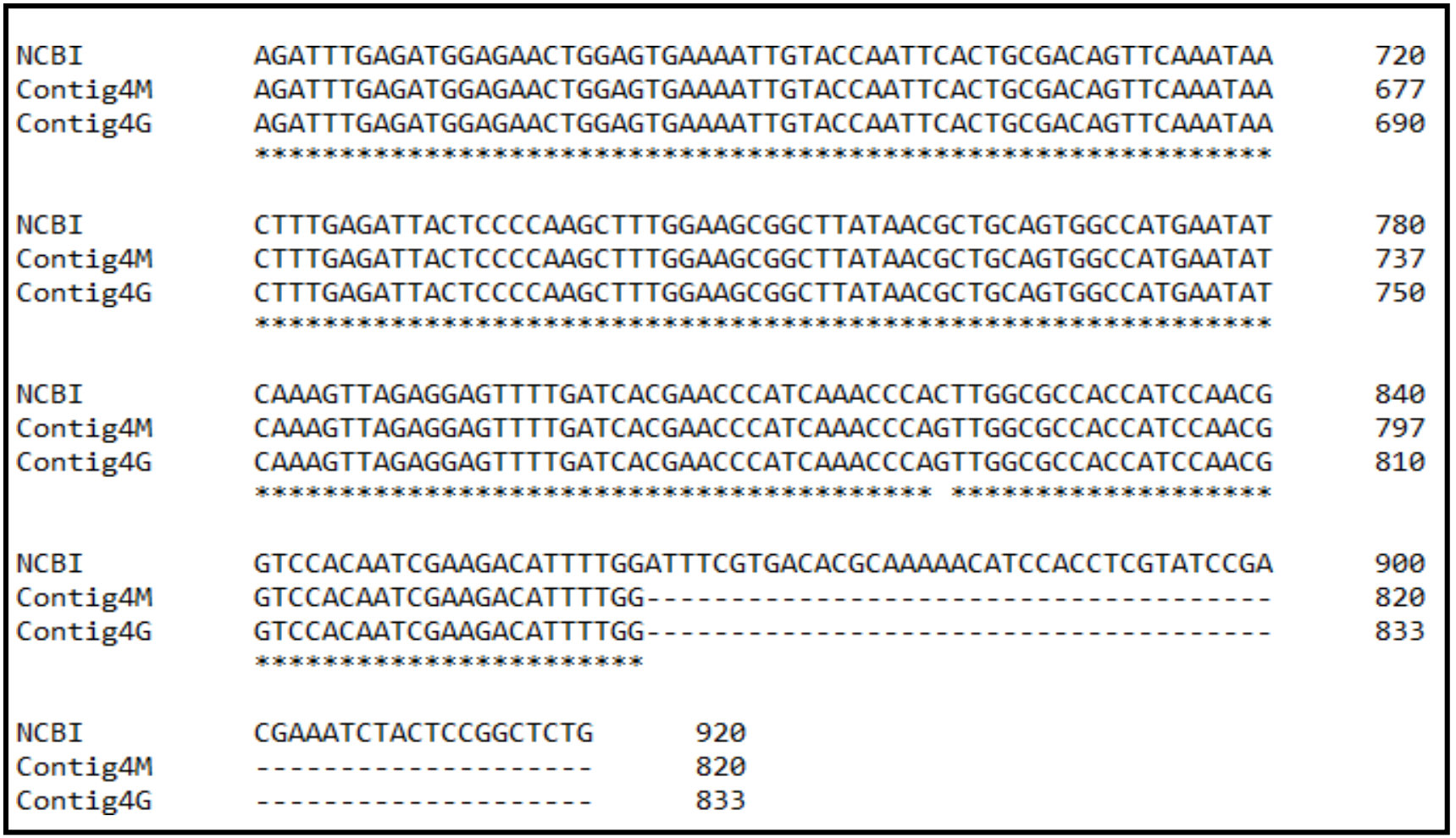
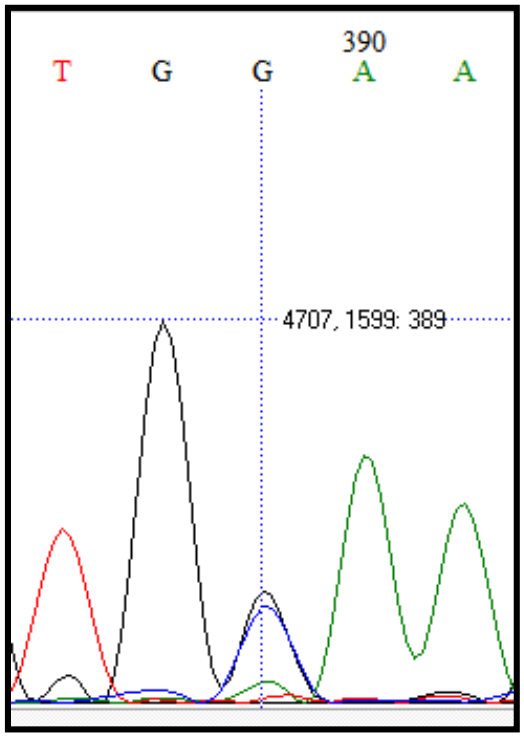
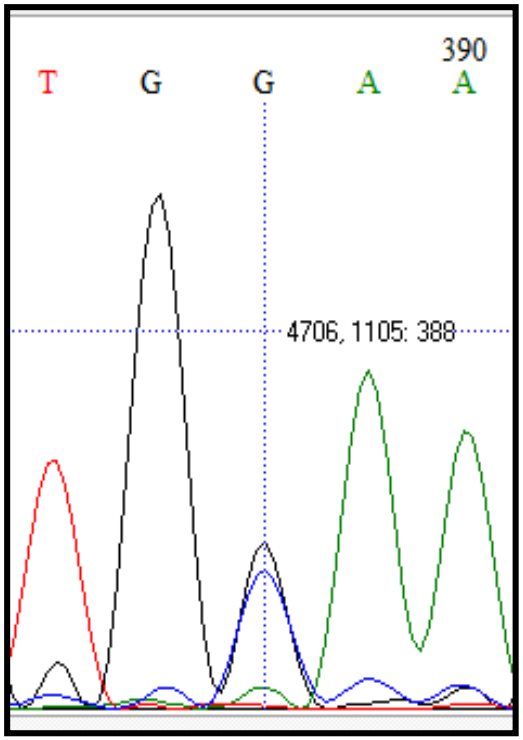


Figure 4 (c). Multiple sequence alignment of *ACS7*



i) Monoecious sequence



ii) Gynoecious sequence

Figure 4 (d). Chromotogram peak images for Var1

Figure 5 (a). Multiple sequence alignment of *AG6*

Figure 5 (b). Multiple sequence alignment of *AG6*

Figure 5 (c). Multiple sequence alignment of *AG6*

corresponding chromatogram peaks to ensure that the variation was not due to sequencing error. To further understand and rely on these variations, multiple sequence alignment was done using Clustal Omega including the homologous sequences available from the related cucurbits like *Cucurbita maxima, Cucurbita pepo* and *C. moschata.* Those sequences were obtained by subjecting the primer target region of *AG6* sequence to BLASTn (Fig. 9).

4.3.5.5.5. Characterization of sequence variations

4.3.5.5.5.1. Characterization of SNPs

SNP1

Multiple sequence alignment had revealed the presence of guanine (G) in 59th position of contig assembled for gynoecious sample whereas in both monoecious sequences, the same aligned contained adenine (A). This was also confirmed by the obtained chromatogram peaks of the sequences which clearly depicts the base calling of guanine (G) (single peak for G at 148th base) in gynoecious sample and Adenine (A) (single peak for A at 141th) in the monoecious sample (Fig. 6). The SNP1 was found to be located within the complete mRNA sequence of the gene (Fig. 15). But, the SNP region was not found to directly code for a protein.

Alignment with orthologous sequences of other related cucurbits (Fig.10) had shown the presence of guanine (G) in particular aligned position of all the sequences except monoecious sequence obtained and the reference one. The SNP1 and its flanking region were not found to be conserved.

SNP2

Multiple sequence alignment had revealed the presence of the base guanine (G) in contig of gynoecious sample, where Adenine (A) was occupying in both the monoecious sequences in the aligned region (Fig. 5b). Upon close observation of the chromatogram peaks, a single peak was found for gynoecious sample at 390th base whereas, a single peak for Adenine (A) was found in the monoecious sample at 383rd base (Fig. 7). The SNP was not comprised within the mRNA sequence of the gene.

Alignment result with orthologous sequences of other related cucurbits (Fig. 11) had shown the presence of Thymine (T) in particular of the three cucurbits, where 'G' and 'A' was present in the gynoecious and monoecious sequence respectively. Flanking region was not found to be conserved.

SNP3

Multiple sequence alignment had shown the presence of the base Guanine (G) in an aligned region where Thymine (T) was found in the mooecious and reference sequence. Since this SNP was found towards the end of the sequence, the chromatogram peaks for this flanking sequence was not found in the AB1 file of the forward run. 5-7 bases flanking the SNP was converted to reverse complementary sequence and search of this sequence in AB1file of reverse run, enabled finding the chromatogram peaks of this sequence. A single peak was observed in the 314^{th} position of reverse monoecious sequence where the Adenine (A) was base called. A peak of Cytosine (C) was observed in 392^{nd} position along with some other background noises in the reverse run of gynoecious sample (Fig. 8). The flanking region also had some gaps attributed only due to the changes in monoecious sequence from gynoecious and reference. The SNP was not comprised within the mRNA sequence of the gene.

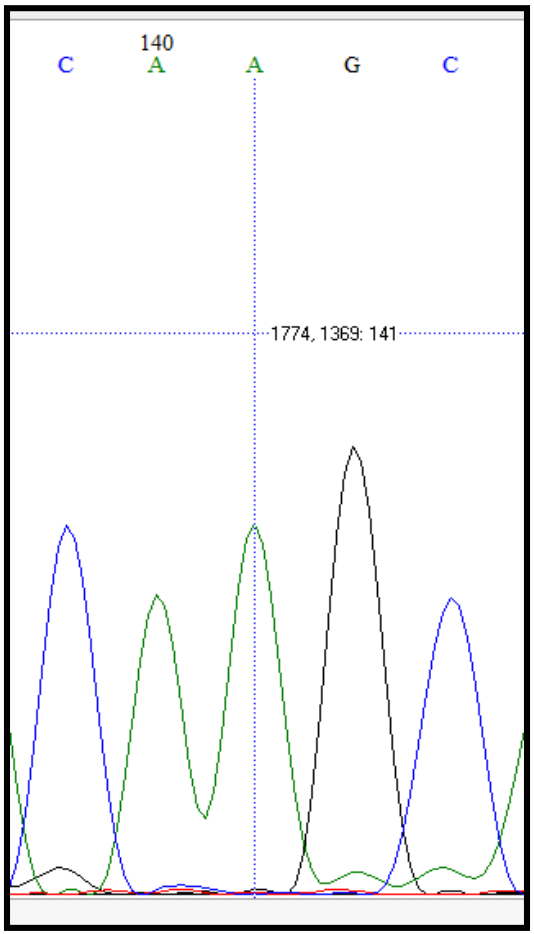
Alignment with orthologous sequences of other related cucurbits (Fig. 12) had shown the presence of Thymine (T) in particular position of all the sequences except the gynoecious sample sequence of the bitter gourd. Flanking region was not found to be conserved.

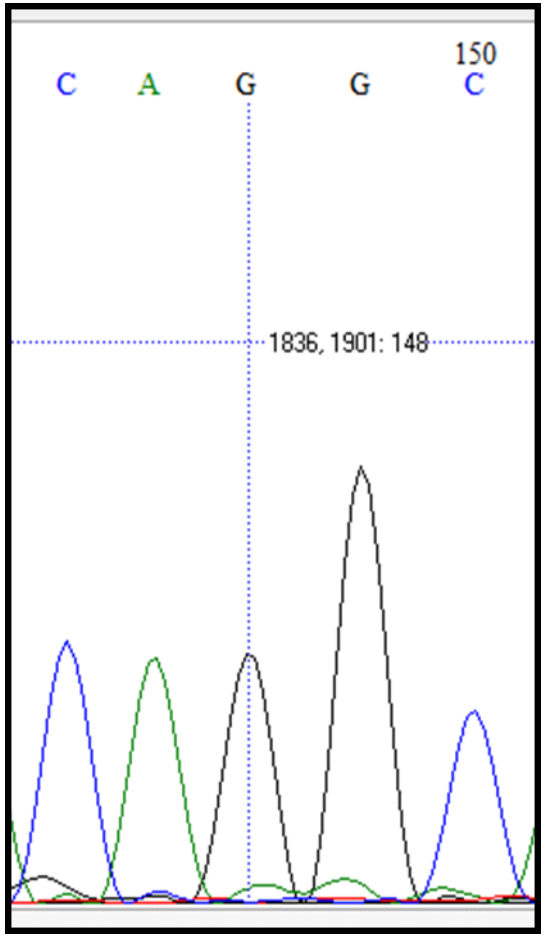
4.3.5.5.5.2. Characterization of In/Dels

In/Del variation 1

Multiple sequence alignment presented a deletion of one base, thymine (T), from gynoecious sequence compared to monoecious sequences (Fig. 5b). It was also confirmed by observing chromatogram peaks. The region was not comprised within mRNA sequence of the gene.

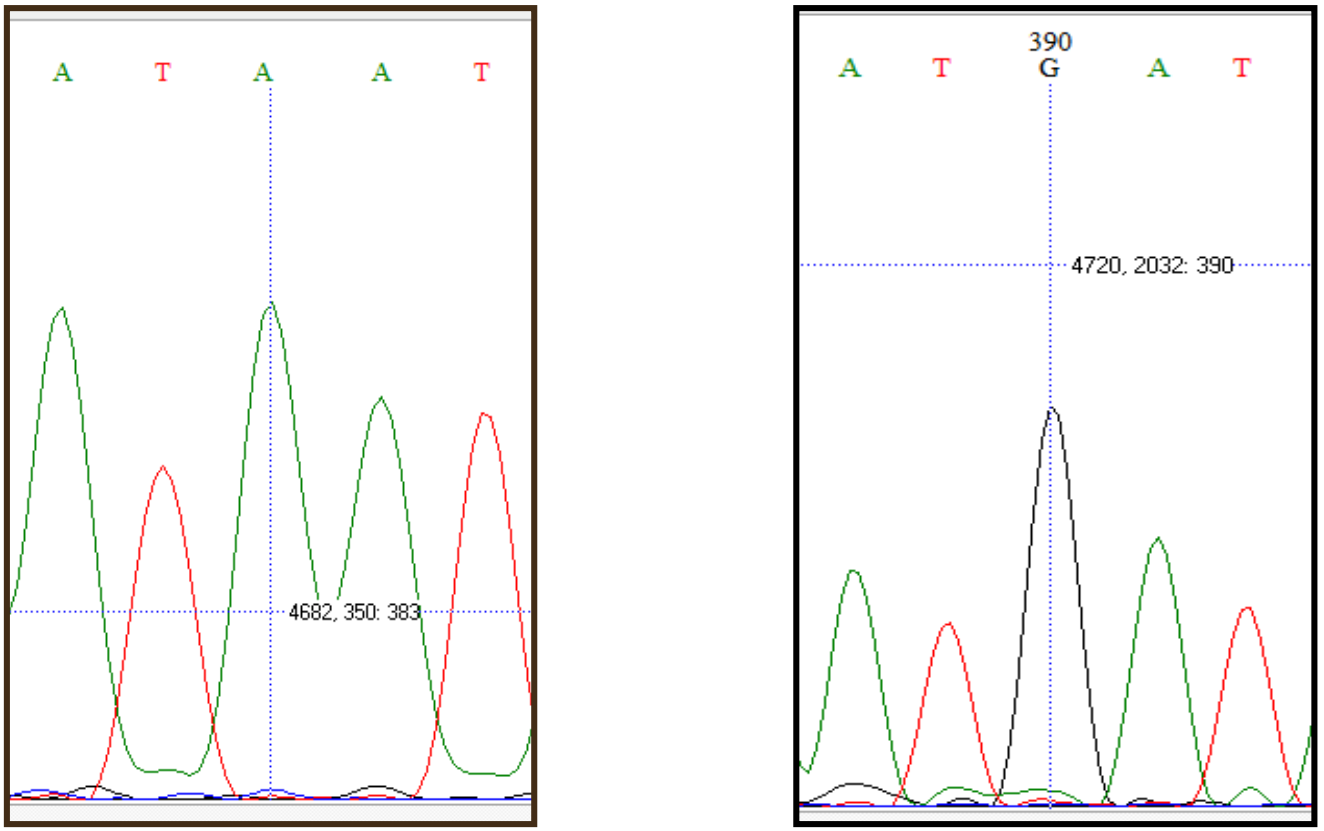
Alignment with orthologous gene sequences had shown the presence of thymine





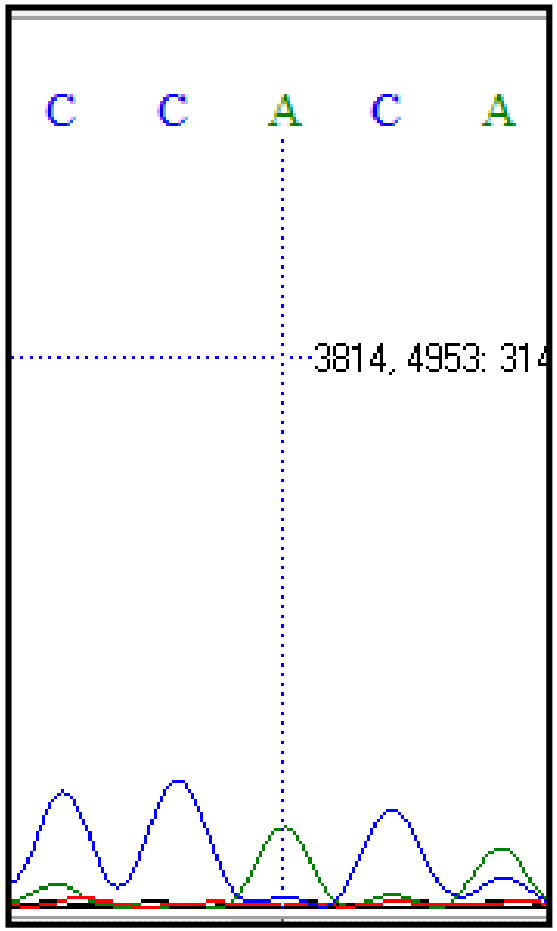
a) Monoecious sequence **b) Gynoecious sequence**

Figure 6. Chromatogram peaks of monoecious and gynoecious sequences for SNP1

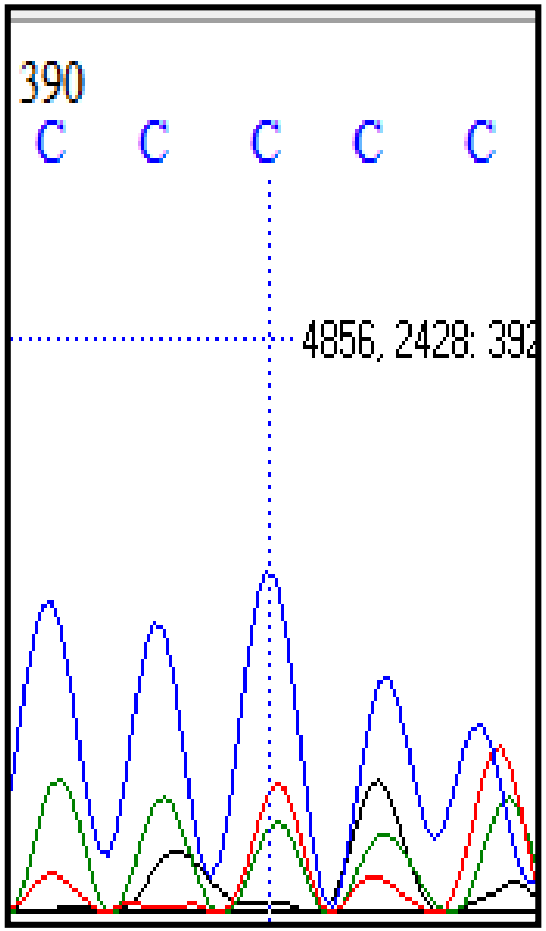


a) Monoecious sequence **b) Gynoecious sequence**

Figure 7. Chromatogram peaks of monoecious and gynoecious sequences for SNP2

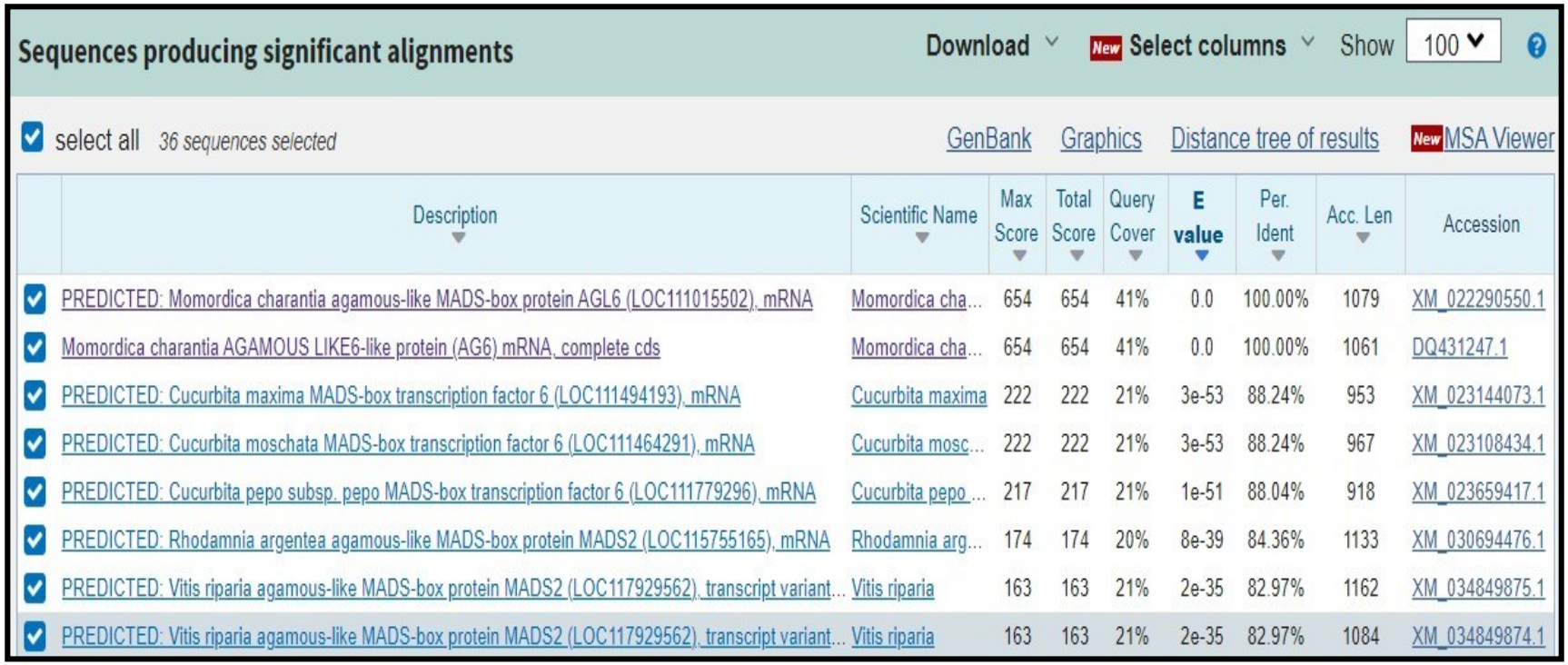


a) Monoecious sequence



b) Gynoecious sequence

Figure 8. Chromatogram peaks of monoecious and gynoecious sequences for SNP3



Sequences producing significant alignments Download New Select columns Show 100

select all *36 sequences selected* GenBank Graphics Distance tree of results New MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
PREDICTED: Momordica charantia agamous-like MADS-box protein AGL6 (LOC111015502), mRNA	Momordica cha...	654	654	41%	0.0	100.00%	1079	XM_022290550.1
Momordica charantia AGAMOUS LIKE6-like protein (AG6) mRNA, complete cds	Momordica cha...	654	654	41%	0.0	100.00%	1061	DQ431247.1
PREDICTED: Cucurbita maxima MADS-box transcription factor 6 (LOC111494193), mRNA	Cucurbita maxima	222	222	21%	3e-53	88.24%	953	XM_023144073.1
PREDICTED: Cucurbita moschata MADS-box transcription factor 6 (LOC111464291), mRNA	Cucurbita mosc...	222	222	21%	3e-53	88.24%	967	XM_023108434.1
PREDICTED: Cucurbita pepo subsp. pepo MADS-box transcription factor 6 (LOC111779296), mRNA	Cucurbita pepo ...	217	217	21%	1e-51	88.04%	918	XM_023659417.1
PREDICTED: Rhodamnia argentea agamous-like MADS-box protein MADS2 (LOC115755165), mRNA	Rhodamnia arg...	174	174	20%	8e-39	84.36%	1133	XM_030694476.1
PREDICTED: Vitis riparia agamous-like MADS-box protein MADS2 (LOC117929562), transcript variant...	Vitis riparia	163	163	21%	2e-35	82.97%	1162	XM_034849875.1
PREDICTED: Vitis riparia agamous-like MADS-box protein MADS2 (LOC117929562), transcript variant...	Vitis riparia	163	163	21%	2e-35	82.97%	1084	XM_034849874.1

Figure 9. BLASTn results of target region in *AG6* gene sequence

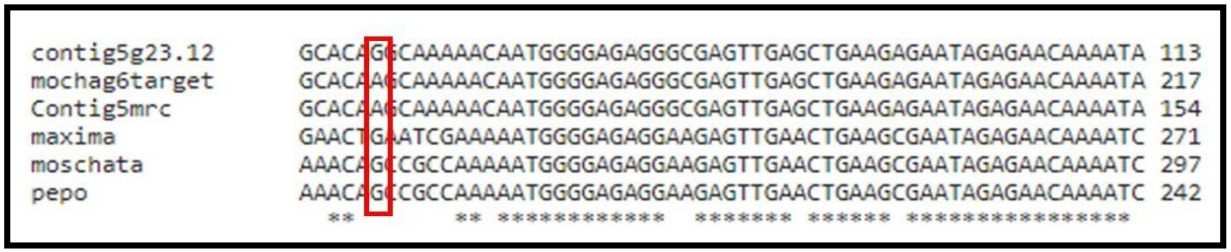


Figure 10. Multiple sequence alignment with orthologous sequences for SNP1

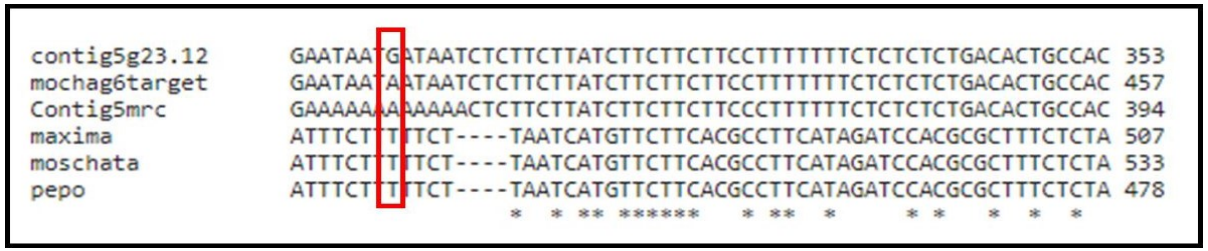


Figure 11. Multiple sequence alignment with orthologous sequences for SNP2

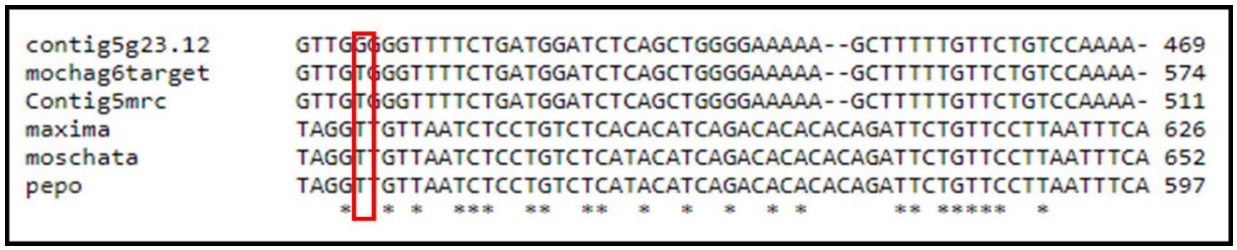


Figure 12. Multiple sequence alignment with orthologous sequences for SNP3

Figure 13. Multiple sequence alignment with orthologous sequences for In/Del 1

Figure 14. Multiple sequence alignment with orthologous sequences for In/Del 2

Figure 15. SNP1 within the mRNA region of *AG6*

(T), in the particular position of all the gene sequence except in the obtained gynoecious sequence (Fig. 13). Four bases flanking right side of the variation was found to be conserved in all the sequence.

In/Del variation 2

Multiple sequence alignment presented a insertion of 48 bases of thymine (T) and cytosine (C) repeats $(TC)_{24}$ from gynoecious sequence compared to monoecious sequences (Fig. 5b). It was also confirmed by observing chromatogram peaks. The region was not comprised within mRNA sequence of the gene.

Alignment with orthologous gene sequences had shown the presence of some other base in the particular position of the gene sequences from other but no bases in both monoecious sequences of bitter gourd, depicting it as deletion from monoecious genotype (Fig. 14).

Table 13. List of Sequence variation in *AG6*

Variations	Gynoecious	Monoecious	Reference
SNP1	G	A	A
SNP2	G	A	A
SNP3	G	T	T
In/Del 1	-	T	T
In/Del 2	$(TC)_{24}$	-	-

4.3.5.6. MADS-box transcription factor 23 (MADS-box23)

4.3.5.6.1. Amplification of the gene

The primer set *(MADSbox-23)* was used to amplify the region of 4898 to 5897 in the gene, targeting 1000 bp. An optimum annealing temperature of 61 ºC was set to run the PCR program. The optimum concentration of forward and reverse primer in the reaction

mixture was found to be 1 µM. A single intact band of 1000 bp size was observed upon electrophoresis in 1.4 per cent agarose gel (Table 12). The gel image was portrayed in the Plate 8. The PCR product was sequenced.

4.3.5.6.2. Analysis of the sequences

In the monoecious sample, the trimmed length of the forward run was of 913 bp and had a QV score of 52. The length of reverse sequence was 865 bp with the QV score of 52. In the gynoecious sample the trimmed sequence length of the forward run was of 889 bp and had a QV score of 52. The length of reverse sequence was 838 bp with the QV score of 51 (Table 12).

4.3.5.6.3. Assembly of contiguous sequence

The length of the contig obtained after assembling the forward and reverse run of the monoecious sample for *MADS-box23* primer was of 914 bp and of gynoecious sample was 896 bp.

4.3.5.6.4. Multiple sequence alignment

The obtained monoecious and gynoecious sequences for the primer was compared with the reference sequence from NCBI which also belongs to a monoecious sample of bitter gourd. No valid variations were found in the obtained gynoecious line sequence compared to both the monoecious genotype (Fig. 16a, 16b and 16c).

4.3.5.7. Floral homeotic protein AGAMOUS-like (McAG2)

4.3.5.7.1. Amplification of the gene

The primer set (*McAG2)* was used to amplify the region of 1153 to 2103 in the gene, targeting 951 bp. An optimum annealing temperature of 61 ºC was set to run the PCR program. The optimum concentration of forward and reverse primer in the reaction mixture was found to be 0.8 µM. A single intact band of 951 bp size was observed upon electrophoresis in 1.4 per cent agarose gel (Table 12). The gel image was portrayed in the Plate 9. The PCR product was sequenced.

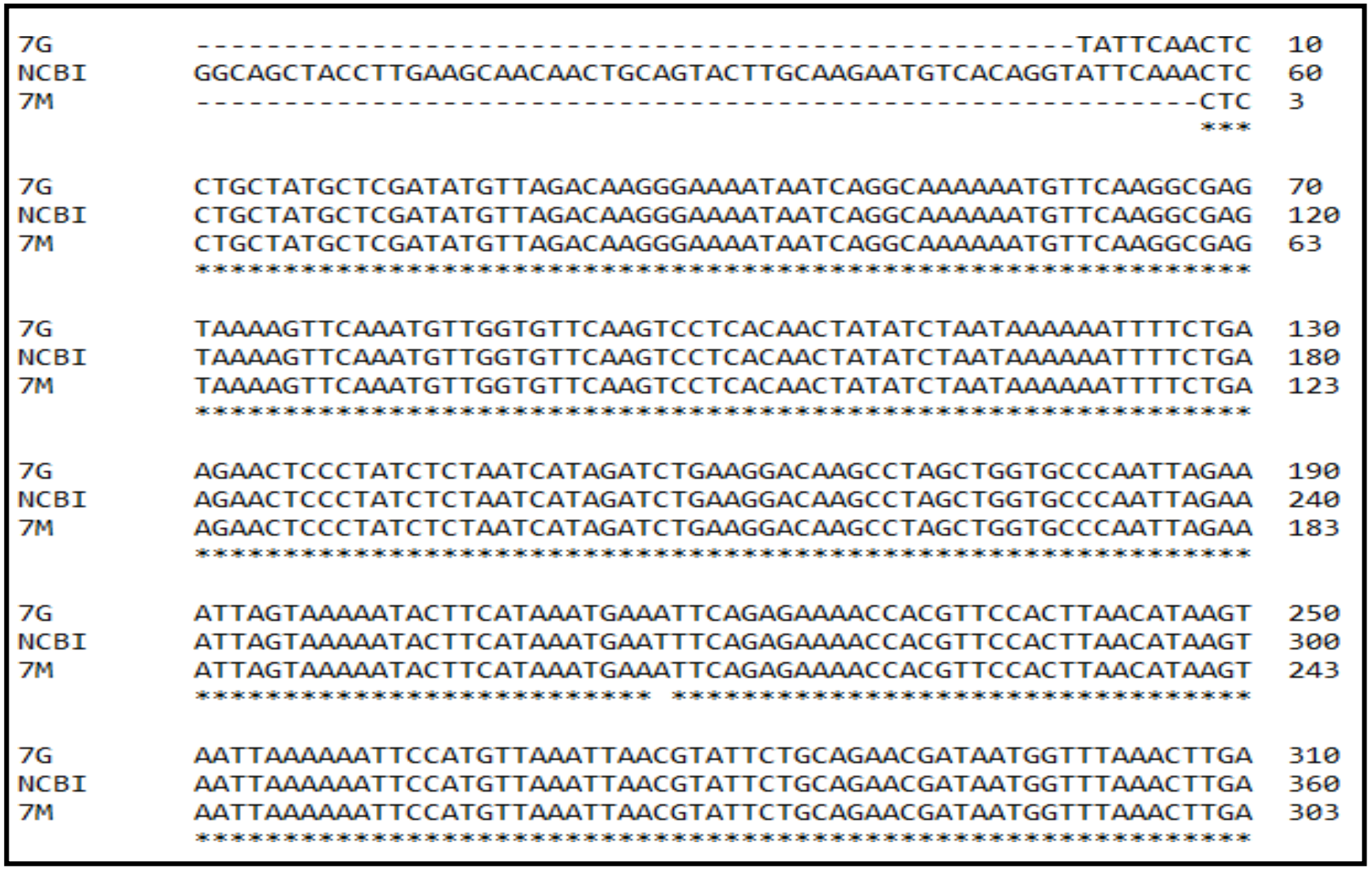


Figure 16(a). Multiple sequence alignment of *MADS-box23* sequence

Figure 16(b). Multiple sequence alignment of *MADS-box23* sequence

Figure 16(c). Multiple sequence alignment of *MADS-box23* sequence

4.3.5.7.2. Analysis of the sequences

In the monoecious sample, the trimmed length of the forward run was of 800 bp and had a QV score of 48. The length of reverse sequence was 812 bp with the QV score of 53. In the gynoecious sample, the trimmed sequence length of the forward run was of 853 bp and had a QV score of 49. The length of reverse sequence was 845 bp with the QV score of 52 (Table 12).

4.3.5.7.3. Assembly of contiguous sequence

The length of the contig obtained after assembling the forward and reverse run of the monoecious sample for *MADS box-23* primer was of 847 bp and of gynoecious sample was 855 bp.

4.3.5.7.4. Multiple sequence alignment

The obtained monoecious and gynoecious sequences for the primer was compared with the reference sequence from NCBI which also belongs to a monoecious line of bitter gourd. One In/Del variation was found in the obtained gynoecious line sequence compared to the both monoecious genotype. A base Guanine (G) was found inserted in gynoecious lines /deleted in monoecious genotype (Fig. 17c). But, the variation was not lying within coding region of the gene.

4.3.6. Marker development with variations

The variation obtained in the putative candidate gene can be used as a molecular markers for differentiating monoecious lines from gynoecy.

4.3.6.2. SSR primer designing

The In/Del2 variation contains a dinucleotide [TC] repeated 24 times more in gynoecious sample compared to both obtained monoecious sequence and the reference sequence. MC136 sequence had deletion of one TC repeats. The total number of TC repeats consecutively in gynoceious line (KAUMCGy-101) was 33, whereas in obtained monoecious sequence (MC 136) and reference (OHB3-1) it was 8 and 9 respectively. This significant variation in number of TC repeats between the monoecious and gynoecious samples were utilized for the development of SSR marker.

Primers were designed from the flanking region of TC repeats using Primer3 software. The details of SSR primers designed were given in Table 14. The SSR marker was named as BGAG6. The SSR primer was designed targeting the region of 470-755 in *AG6* gene sequence (NCBI). The primer binding regions were conserved in both monoecious and gynoecious samples which ensure the binding of primer to the template.

Table 14. Details of SSR primer designed

Primer	Sequence	Length (bp)	T_m(°C)	Product size (bp)
BGAG6F	5'TCTGACACTGCCACAGGTTC3'	20	59.87	286
BGAG6R	5'AAACAGACAAATGGGGTTGC3'	20	59.84	

4.3.6.3. SSR amplification and validation

4.3.6.3.1. PCR amplification using BGAG6

PCR amplification with the primer set BGAG6F/ BGAG6R yielded a clear band of 286 bp in monoecious sample (MC 136) and 334 bp in gynoecious sample (KAU-MCGy-101). The gel image showing polymorphic bands in genotypes used for marker development is portrayed in Plate 10.

4.3.6.3.2. Validation of marker BGAG6

The SSR marker was screened in three other monoecious genotypes *viz.* Priya, Priyanka and wild type *M. charantia var. muricata* to check for polymorphic bands. The marker had yielded approximate band size of 286 bp in all the screened monoecious varieties. In the variety Priya, one additional band of 334 bp was also observed. The gel image showing amplification pattern of SSR in different lines is portrayed in Plate 11.

Figure 17(a). Multiple sequence alignment of *McAG2* sequence

Figure 17(b). Multiple sequence alignment of *McAG2* sequence

Figure 17(c). Multiple sequence alignment of *McAG2* sequence

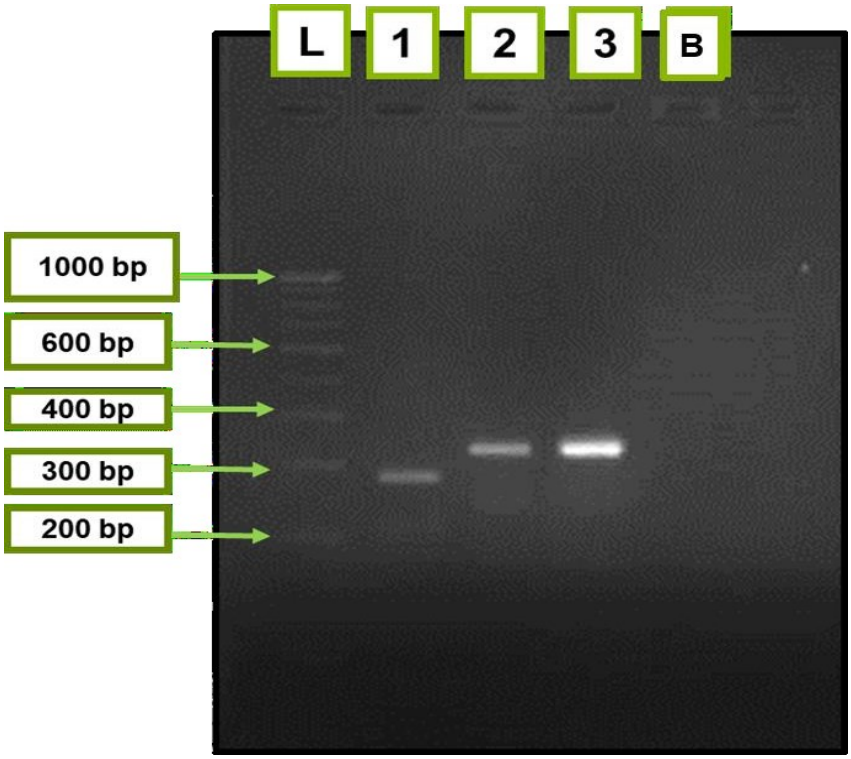


Plate 10. Amplification using SSR primer set BGAG6

L - 100 bp ladder

1 - Monoecious line (MC-136)

2 - Gynoecious line (KAUMCGy-101)

3 - Gynoecious line (KAUMCGy-101)

B - Blank

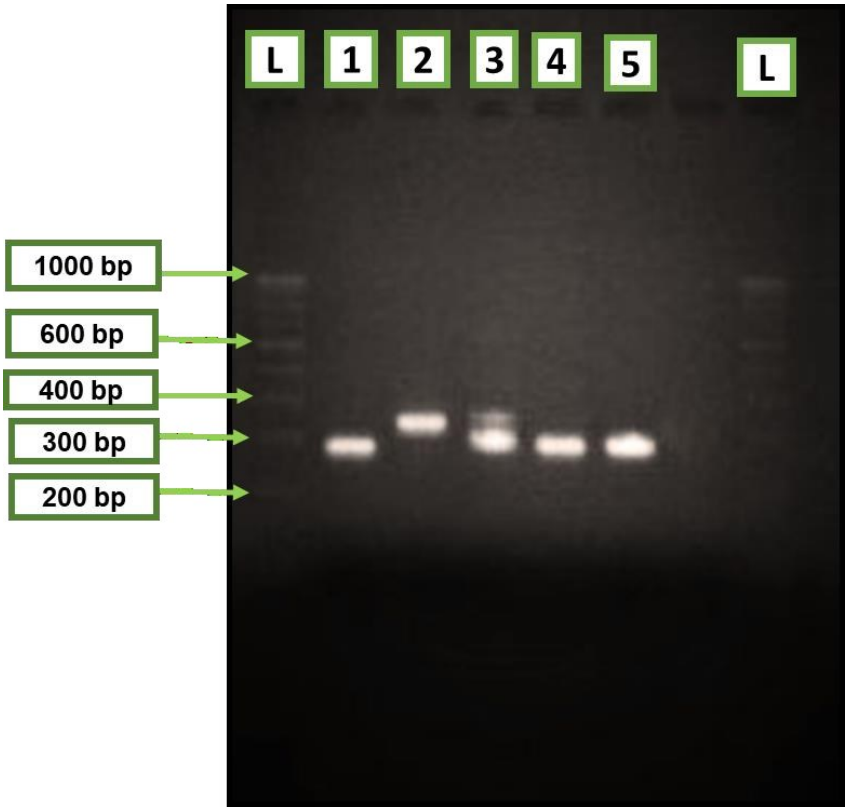


Plate 11. Validation of SSR marker BGAG6

L - 100 bp DNA ladder

1 - Monoecious line (MC-136)

2 - Gynoecious line (KAUMCGy-101)

3 - Monoecious line (Priya)

4 - Monoecious line (Priyanka)

5 - Monoecious line *(M. charantia* var. *muricata*)

Discussion

5. DISCUSSION

Bitter gourd (*Momordica charantia* L.) is an economically and medicinally important vegetable crop belonging to the family Cucurbitaceae. Hybrid vigour has been exploited well in the crop for earliness and yield related traits. However, hand emasculation and pollination makes the production of hybrids labour intensive and costly. Use of gynoecious line as female parent in hybrid production eliminates this problem. Since gynoecy is reported to be highly influenced by environment (Dey *et al*., 2012), the early detection of the gynoecious lines will be difficult. Molecular markers linked to gynoecy aids in the early detection of this economically important trait**.** Identification of candidate genes for the gynoecism may help in development of molecular markers for detecting gynoecious line.

The study entitled 'Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)' was carried out at Centre for Plant Biotechnology and Molecular Biology (CPBMB) and Department of Vegetable Science, College of Agriculture, Vellanikkara, Thrissur, during 2018- 2021. The findings of this study a will be discussed in this chapter.

53.1. Candidate gene analysis for gynoecy

53.1.1. Selection of putative candidate genes

Putative candidate genes for gynoecy selected in this study can be grouped as ethylene and auxin related genes, MADS-box transcription factor genes and WRKY transcription factor genes**.**

53.1.1.1. Ethylene and auxin related genes

Potential candidate genes governing gynoecy in bitter gourd were identified by Gunniah *et al.,* (2014) using *in-silico* differential gene expression analysis between a gynoecious (Gy323) and monoecious (DRAR-1). They identified a list of putative

candidate genes which includes ethylene biosynthesis genes, ACOs and ACS. A study by Kumari (2015) provided a strong evidence for role of *McACS2* mediated biosynthesis of ethylene in sex determination of bitter gourd.

The candidate genes related to ethylene biosynthesis mediated sex determination namely, *1-aminocyclopropane-1-carboxylate oxidase (ACO1)***,** *1-aminocyclopropane-1-carboxylate synthase CMA101-like (ACS2), 1-aminocyclopropane-1-carboxylate synthase (ACS3*) and *1-aminocyclopropane-1-carboxylate synthase 7-like (ACS7*) were selected and each of their gene sequences were retrieved from previous scaffold level draft genome assembly of *M. charantia,* ASM19950v1 in NCBI GenBank database.

Auxin related transcripts were one among the highly differentially expressed genes between monoecious and gynoecious lines (Shukla *et al.,* 2015). *Auxin response factor 4 (CsARF5)* from cucumber was selected and retrieved from chromosome level genome assembly of *Cucumis sativus,* ASM407v2.

53.1.1.2. *MADS-box* transcription factor genes

AG (*Agamous*) is a class C homeotic gene which encodes for MADS box transcription factor. Its interaction with other MADS-box protein plays an essential role in induction of floral organ in *Arabidopsis* (Honma and Goto, 2001). Peng *et al.* (2008) studied gene expression pattern of *McAG2* and *McAG6* in different tissues of bitter gourd which clearly indicated the role of two genes in regulation of floral organ development in *Momordica charantia* L.

A great number of MADS-box genes have expression pattern that is different between male and female sexes during genetic and molecular basis of analyses in many plants (Gramzou and Theissen, 2010). Three MADS box genes *McMADS1, McMADS10* and *McMADS22* were identified as putative genes related to sex determination mechanism in bitter gourd (Hseih and Yin-yin Do, 2019).

Three *MADS-box* transcription factor genes *viz. Agamous-like MADS-box protein AGL6 (AG6)*, *MADS-box transcription factor 23* and *Floral homeotic protein AGAMOUS-*

like (McAG2) from bitter gourd were selected in this study as putative candidate genes for gynoecy.

53.1.1.3. WRKY transcription factor genes

WRKY DNA-binding protein has been identified as one among the putative transcription factors of sex determination, when transcriptome profiling was done between monoecious and gynoecious cucumber (Guo *et al*., 2010). Rudich *et al*. (1974) suggested that ABA promotes female flowering tendency wherein WRKY TFs plays a crucial role in ABA response pathway (Rushton *et al.,* 2010).

A list of 12 WRKY transcription factor genes were selected from a report by Shukla *et al.* (2015) from cucumber for candidate analysis in the study (Table 2). The genes of smaller size were preferred during selection.

53.1.2. Total genomic DNA isolation

Pale green, tender leaves of bitter gourd (first or second from the tip) were used for isolation of DNA. The CTAB method of DNA isolation protocol suggested by Rogers and Bendich (1994) with slight modification yielded good quality and quantity of DNA. The modification adopted was adding of PVP directly to the CTAB buffer while pre-warming in water bath.

The main problem encountered during isolation of DNA from plants is the accumulation of protein, polyphenols and secondary metabolites (Weising *et al*., 1995 and Matasyoh *et al*., 2008). Collection of samples in ice box was practiced to prevent the oxidation of polyphenols to polyquinones, which binds to the DNA, making the extraction difficult. Grinding the samples with liquid nitrogen, makes the cell wall brittle and hold the metabolic activities, which also helped in reducing phenol oxidation. Addition of 50 µL of β-mercapto ethanol helped in controlling protein contamination by denaturation of proteins. RNAse A (Ribonuclease A) was used in eliminating RNA contamination(Gallego and Martinez, 1996; Wettasinghe and Peffley 1998). Over drying of DNA after

ethanol wash caused problem in dissolving pellets. This is due to the conversion of B form of DNA to D form.

53.1.3. Assessment of quality and quantity of DNA isolated

During molecular analysis, the quality and quantity of DNA used plays a crucial role in determining the success of further steps like PCR and sequencing. Agarose gel electrophoresis and spectrophotometric methods were used to determine the quality and quantity respectively.

The isolated DNA samples were ran on 0.8 per cent agarose gel. The EtBr added in the gel enables the visualization of DNA when observed under UV light. This is due to the intercalation of dye between the base pairs of DNA. Good quality DNA will be indicated by the presence of an intact, high molecular weight band, with less smearing and minimum contamination by RNA (Wettasinghe and Peffley, 1998). The good quality DNA isolated in the present study is portrayed in Plate 2.

The quantity analysis by NanoDrop^{R} ND-1000 spectrophotometer was found efficient, since it measures absorbance from one µL sample with high accuracy and reproducibility. It eliminates the use of cuvettes and other sample contaminating devices as in the conventional UV spectrophotometer. Maximum absorbance of nucleic acids will be at the wavelength of 260 nm and that of protein will be 280 nm. A_{260}/A_{280} ratio will be within the range of 1.8-2.0 for good quality DNA samples (Meena, 2014; Thakur e*t al*., 2014). The absorbance ratio of the isolated samples fell within this range. (Table 10).

After assessing the quantity, the DNA samples should be diluted to 50 µL for carrying out optimum PCR. The fitness of the diluted sample for further molecular analyses can be indicated by good amplification when ran with a RAPD marker.

53.1.4. PCR amplification using gene-specific primers

Candidate gene analysis involves complete sequencing of the genes under study in all the genotypes compared and studying of variations in the sequences for its association with particular trait. In this study, a particular region of each of the candidate genes were

sequenced by designing gene-specific primer pairs. The product size of designed primers were set in the range of 850-1000 bp for all genes considering the processivity of *taq* polymerase used in this study.

53.1.4.1. PCR optimization

The newly designed primers have to be standardized for optimum amplification of the desired product. Primer dimers are a potential threat to the quality of sequence obtained *via* Sanger sequencing (Brownie *et al.,* 1997; Poritz and Ririe, 2014). The optimum concentration of primers required in reaction were standardized for each set of newly designed gene- specific primer. Low annealing temperature caused non-specific binding of primers, yielding extra band along with the intended ones. Gradient PCR was found useful in obtaining optimum temperature of annealing. Touchdown (TD) PCR offers advantage of fast optimization, increased specificity, sensitivity and yield (Korbie and Mattick, 2008). TouchDown PCR was applied in this study for primers whose annealing temperature were known, but specificity and yield was a problem. No binding was found in some primer pairs designed from gene sequences of cucumber, because the primer binding region of those genes were not conserved in cucumber and bitter gourd.

53.1.4.2. Gel elution of desired band

Gel elution was useful in the case where all the optimization approaches failed to eliminate the non-specific PCR products that reduces the quality of sequencing. Incubating the samples up to 5 minutes in the spin column where DNA tends to bind, before each centrifugation step contributed in eluting a good concentration of product. Running one more PCR with the eluted product as a template with same composition of reaction mixture and amplifying conditions helped solving the issue of very low concentration in gel elution and purification.

53.1.5. Sequencing of PCR products

The quality of sequencing data will be determined based on established metric known as Quality Value (QV). The quality sequencing data having QV greater than 20,

means the probability that the base was miscalled is not greater than one per cent, is the acceptable standard for a good sequence reaction (http://www3.appliedbiosystems.com). The obtained sequences are of good quality and reliable, since all the sequences had QV score more than 20.

Errors in sequencing can be found by scanning chromatogram peaks. It can be identified by the presence of unevenly spaced peaks and more background noise to the signal due to non-specific binding of primers. Errors in the extreme ends are acceptable and common. This will be due to loss of resolution of signal as the gel progresses, but the sequencer still attempts to read the signals, causing errors. Therefore, only trimmed sequence obtained after removal of poor quality reads were used for assembling contigs.

Failure of only reverse run for the sample with the forward one extremely good, is one of the common problems encountered in sequencing of PCR products. Before sequencing, the samples will be purified to remove excess primers and residual salts already present. If both the primers were still present in the sample, the forward primer could dominate the reaction causing the failure of reverse run or shortening length of reverse read excessively (Goraichuk, 2018). The failure of reverse run sequence with the primer ACS2R was inferred due to improper purification of the PCR product.

53.1.6. Sequence variation analysis

Variations in the DNA sequences of plant genome can be deployed as molecular breeding tools for trait association or molecular breeding (Stokes, 2002 and Gupta, 2002). Spotting valid variations between contrasting genotypes requires very good quality sequences. Scanning chromatogram peaks for all the variations found in the multiple sequence alignment helped in avoiding misinterpretations of base calling and finding out only valid variations.

No variations were found in sequence of gynoecious sample compared to monoecious genotypes in the amplified region of the following genes- *ACO1, ACS2, ACS3, ACS7* and *MADS-box 23.*

A single nucleotide mutation in *ACO* gene conferring andromonoecy to cucumber was studied by Chen *et al.* (2016). The enzymatic activity of *CsACO2* gene will be disrupted resulting in the decreased ethylene emission from shoot tips during the critical stage of sex determination. The development of carpel was blocked, turning the plant to andromonoecy. The orthologous gene of *CsACO2* in melon, *CmACO3* also showed similar expression pattern in the region of carpel development.

Since only 780 bp of *ACO1* gene in the gynoecious line of bitter gourd was sequenced in this study, and no mutation was found, full gene sequencing have to be done to establish its influence on sex expression.

Cs-ACS2 gene was reported to be potentially involved in regulating the development of female flowers at the apex of the gynoecious cucumber plants (Kamachi *et al*., 1997). Role of *McACS2* in ethylene mediated gynoecious sex expression in bitter gourd was elucidated by Kumari (2015).

Bouleam *et al*. (2008) revealed that a single-nucleotide mutation in the highly conserved region of the *CmACS-7* affected the protein function and resulted in the reduced enzymatic activity of ACC synthase in melon leading to andromonoecy.

Martin *et al.* (2009) delineated that a transposon insertion which led to methylation of the WIP1 (G) gene promoter resulted gynoecious plants in melon. WIP1 encodes a transcription factor whose expression prevents development of female organ.

Since only a selected region of the genes *ACS2, ACS3, ACS7* and *MADS-box 23* in bitter gourd were sequenced and no mutation was observed, full characterization or at least sequencing only the exonic region helps in understanding their role in gynoecious sex expression.

This explains that the region of these genes amplified and sequenced in this study have no direct role in influencing the gynoecious sex expression of bitter gourd.

Agamous-like MADS-box protein AGL6 (AG6) gene accounts for total of 3 SNPs and 2 In/Del variations within the region amplified and sequenced in this study. *Floral*

homeotic protein AGAMOUS-like (McAG2) gene accounts for one valid In/Del variation within the sequenced region of gynoecious genotype used in this study.

53.1.6.1. Characterization of VariationsSNP1

SNP1 in *Agamous-like MADS-box protein AGL6 (AG6)* was identified as a transistion mutation [G/A] which is present in the coding region of the gene 10 bp downstream from protein coding ORF, but not actually translated to protein. Therefore, its presence in the 3'UTR (untranslated) region of the gene can be understood, which may influence gynoecy trait.

The non-coding regions downstream of coding sequences (CDS) is usually termed terminators or 3'UTRs. The 3' regulatory regions have a many types of *cis*-regulatory elements which is directly involved in polyadenylation, stability, transport and mRNA translation, essential to achieve the desired levels of gene expression (Bernardes and Menossi, 2020). '3'UTR mediated control of flowering time in Arabidopsis is established by Kim *et al.* (2013).

Maximum expression of *McAG6* in the shoot apex of bitter gourd along with expression in both male and female flowers buds was revealed by Peng *et al*. (2008). They also established that *MCAG6* have a role in regulating bitter gourd floral development.

SNP2

This is the transition mutation [G/A] in the non-coding region of the gene *Agamous-like MADS-box protein AGL6 (AG6).*

SNP3

This is identified as a transversion mutation [G/T] found in the non-coding region of the gene. Since T is found conserved in a monoecious cucurbits like *C. moschata, C. pepo, C. maxima* (Fig. 12), this SNP may influence the gynoecy trait in bitter gourd.

In/Dels

In/Del 1 in non-coding region of *Agamous-like MADS-box protein AGL6 (AG6)* gene may have influence on gynoecious trait, since the base pair thymine [T] was found deleted in gynoecious line, which is present in all the orthologous genes from cucurbits compared. In/Del mutation of 56 bp at 3'-UTR of branched-chain amino acid transaminase (BCAT) gene between monoecious and gynoecious cucumber lines was reported by Win *et al*. (2015).

In/Del 2 mutation of 48 bp dinucleotide repeats $[TC]_{24}$ present in the non-coding sequence of gynoecious line was observed in none of the monoecious cucurbits orthologs.

To overcome low pollination success in dioecious plants, at least seven transition events from dioecy to monoecy might have followed in the genus Momordica (Schaefer and Renner, 2010). Since the possible ancestral form of sex expression was dioecy, the In/Dels in the gene *AG6* which was also found conserved in other cucurbits indicates that those were deletion in monoecious lines and not insertion in gynoecious line.

In/Del 1 variation [T] found in non-coding region of *Floral homeotic protein AGAMOUS-like (McAG2)* gene may influence gynoecious sex expression in bitter gourd indirectly.

53.2. Potential candidate genes of gynoecy in bitter gourd

Agamous-like MADS-box protein AGL6 (AG6) gene have valid SNPs and In/Dels having potential association with gynoecism. But, its location in chromosome was not clear, since BLAST with new genome assembly have shown the sequence in both the chromosome 2 and 7 in the hit list with 100 per cent identity. *Floral homeotic protein AGAMOUS-like (McAG2)* is located in chromosome 1 and have strong evidences of flowering regulation.

Previous studies have mentioned the presence of gynoecy trait governing QTL in the distal end of chromosome 1/linkage group 1 (Matsumura *et al*., 2014 and Matsumura *et al*., 2020). Cui *et al*. (2018) have proposed that gynoecy in bitter gourd may be governed

by interaction of two genes located in linkage groups MC01 and MC02. QTLs for other sex expression related traits that may have influence in gynoecious nature such as female flower number and first female flower node were also reported to be present in linkage group 1 (Cui *et al.,* 2018).

Thus, the above said evidences suggest that these two genes can be considered as the potential candidate genes of sex expression in bitter gourd, especially *McAG2* as potential candidate for gynoecy. However, full gene characterization of both the genes are required to reveal its potential as candidate genes in governing gynoecy.

53.3. Development of molecular markers for gynoecy

Marker assisted selection (MAS) certainly accelerates the breeding process and powerful tool for selecting traits such as gynoecism. (Rao, 2018). The polymorphic molecular markers are scarce in public database which hinders genetic mapping and its usage in molecular breeding of bitter gourd. Due to limited molecular resources, India molecular breeding of bitter gourd in India is in its infancy, particularly on functional markers for traits such as gynoecy (Saha, 2018). Identification of gynoecious sex type in bitter gourd using marker-assisted breeding in the early stage holds immense potential in introgression of gynoecious trait.

53.3.1. Candidate gene based SSR marker

Simple sequence repeats (SSRs) are the preferred genetic marker having significant importance in plant genetics and breeding over other marker types (Akkaya *et al*., 1992). This is attributed to its advantages like very low requirement of DNA, ease of detection by PCR (Morgante and Olivieri, 1992; Peakall *et al*., 1998), highly polymorphic (Gupta *et al*., 1996) and extensive genome coverage.

F (female) locus-specific co-dominant marker based on In/Dels polymorphism was identified by Win *et al*. (2015) which perfectly distinguished homozygous and heterozygous gynoecious cucumber lines. The specific pair of primer Cs-BCAT-F/CsBCAT-R was designed based on 56-bp In/Del at 3'-UTR of branched-chain amino

acid transaminase (BCAT) gene to analyze parents and F_1 hybrid. Polymorphism was found between parents, with an approximately 216 bp band in the monoecious line and 160bp band in the two gynoecious lines while both the bands were present in F_1 plants.

Agamous-like MADS-box protein AGL6 (AG6) gene based SSR marker with primers designed specific to 48 bp In/Del $(TC)_{24}$ variation in the gene was developed on this study. The developed candidate gene based SSR marker have shown polymorphism between the monoecious (MC-136) and gynoecious (KAU-MCGy-101) genotype.

High yielding monoecious varieties of bitter gourd released from Kerala Agricultural University *viz*. Priya and Priyanka were used in screening the developed SSR marker. Priya is a green fruited variety whereas, Priyanka is white-fruited bitter gourd variety. The wild type *M. charantia* var. *muricata* included in the screening is also of green fruited type.

The molecular marker developed in this study, BGAG6 have shown monomorphic band of 286 bp in the monoecious genotypes, MC136, Priyanka and *muricata* and polymorphic band of 334 bp in gynoecious genotype on validation. The monoecious variety Priya possesses both size of bands showing heterozygous condition. Since the gynoecy is said to be governed by recessive gene, the variety Priya expresses monoecious character.

Thus the SSR marker, BGAG6 have yielded the expected amplification pattern in all the three varieties screened and differentiated all the three monoecious genotypes (Priya, Priyanka and *M. charantia* var. *muricata)* included in the study, from the gynoecious genotype (KAU-MCGy-101). This indicates that the SSR marker developed in this study is a polymorphic marker and it can be used for identification of the gynoecious line (KAU-MCGy-101) from monoecious lines. The marker needs to be validated in large population of monoecious lines, Priya, Priyanka and *muricata*, to confirm it as a reliable marker for distinguishing these monoecious lines from the gynoecious line, KAU-MCGy-101. Also, it is required to screen more number monoecious and gynoecious varieties for unravelling its potential in marker-assisted selection of gynoecious genotypes in bitter gourd.

Molecular marker for SNPs have to be designed such that the 3' end of the primer falls complementary to the base on the allele that is present only in monoecious or gynoecious genotype. Hence, amplification will occur only in either of the genotypes. The designed marker also needs validation with large number of population.

The gene *AG6* from which the SSR was developed may have role in fruit colour development, since *TAG1,* an *Agamous* homolog from tomato have been reported to regulate genes of chloroplast synthesis involved in fruit colour development (Liu *et al.*, 2020). Kole *et al*., (2012) have mapped the QTL governing fruit colour in linkage group 7. Hence, it can be hypothesized that *AG6* may also governs fruit colour in bitter gourd and polygenic control of fruit colour in bitter gourd may be related to sex expression genes. The marker have to be screened in more number of white fruited monoecious genotypes to to unravel its potential in marker assisted selection based on immature fruit colour.

Full gene characterization of 60 kb *NUA* (*Nuclear Pore Anchor*) gene present in the region associated with the QTL determining gynoecy (chromosome 1) in bitter gourd may be done in future for development of large number of molecular markers by using sequence variations. This gene which comprises mostly of introns with coding sequence of only 6,240 bp was identified by Matsumura *et al*. (2020) as a potential candidate gene governing gynoecious sex expression in bitter gourd.

Summary

6. SUMMARY

The study entitled 'Molecular characterization of gynoecy in bitter gourd (*Momordica charantia* L.)' was carried out at Centre for Plant Biotechnology and Molecular Biology (CPBMB) and Department of Vegetable Science, College of Agriculture, Vellanikkara, Thrissur, during 2018- 2021. The objective of the study was to characterize the gynoecious line of bitter gourd at molecular level through identification of molecular markers that are closely linked with the sex expression. The findings of this study are summarized as follows:

- Total genomic DNA was isolated from pale, tender leaves of one monoecious and one gynoecious genotype. The isolated DNA was quantified using NanoDrop® ND-1000 spectrophotometer. The absorbance ratio in the range of 1.80 – 2.0 was obtained for all the samples, indicating the presence of good quality DNA for carrying out molecular characterization
- 20 putative candidate genes of gynoecy were selected for this study and were retrieved from bitter gourd and cucumber genome assembly in NCBI GenBank database
- Eight putative candidate genes sequences were retrieved from bitter gourd genome assembly, out of which four were ethylene biosynthesis genes (*ACO1, ACS2, ACS3, ACS4* and *ACS7*) and three were MADS-box transcription factor encoding genes (*AG6, MADS-box TF23* and *McAG2*)
- 13 putative candidate genes sequences were retrieved from cucumber genome assembly, out of which one belongs to auxin signalling pathway (*CsARF5*) and the rest 12 were WRKY transcription factor encoding genes (*WRKY16, WRKY37, WRKY4, WRKY9 WRKY10, WRKY21, WRKY23, WRKY25, WRKY49, WRKY50, WRKY51, WRKY54* and *WRKY56)*
- Primer sets were designed for a selected region for each of the genes. Totally 20 gene-specific primer pairs were designed. Seven primers were PCR amplified and

sequenced. Rest of the primer sets designed from cucumber sequences failed to amplify in bitter gourd template

- Sequence variation were analyzed by comparing each of the obtained monoecious and gynoecious bitter gourd sequences with the NCBI reference sequence which also belongs to a monoecious bitter gourd line
- A total of three SNPs and three In/Dels were obtained in gynoecious sequence in comparison with both monoecious genotypes sequences. The sequenced region of the genes *AG6* had three SNPs and two In/Dels and *McAG2* had 1 In/Del mutation. Thus both genes were identified as potential candidates for gynoecism in bitter gourd
- Except SNP1 of *AG6*, all other variations were from non-coding region. The lengthiest variation was found in In/Del2 from *AG6* gene which was of 48 bp, a di-nucleotide repeat of 24 times $[TC]_{24}$
- Significant difference in the number of di-nucleotide repeats [TC] of monoecious and gynoecious sequence was used to develop SSR marker BGAG6. PCR amplification with this marker yielded polymorphic band between the monoecious and gynoecious genotypes
- Validation of this SSR marker BGAG6 was done using isolated DNA samples of three other monoecious lines – Priya, Priyanka and the wild type *M. charantia* var. *muricata.* Expected banding pattern of polymorphism was obtained for all the three monoecious varieties screened. The SSR marker developed needs to be validated in large number of population and varieties to reveal its use as a potential marker in marker-assisted selection of gynoecious lines of bitter gourd

Reference

REFERENCE

Abeles, F. B. 1973. *Ethylene in Plant Biology*. Academic press, New York, 302p.

Abeles, F. B., Morgan, P. W., and Saltveit Jr, M. E., 2012. *Ethylene in plant biology*. Academic press. pp. 302

Ainsworth, C. 1999. *Sex Determination in Plants*. Oxford press, UK, BIOS scientific publishers.

Akkaya, M. S., Bhagwat, A. A., Cregan, P. B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132(4):1131–1139.

Angenent, G.C. and Colombo, L. 1996. Molecular control of ovule development. *Trends Plant Sci.* 1(7): 228-232.

Atsmon, D. and Tabbak, C. 1979. Comparative effects of gibberellin, silver nitrate and amino ethoxyvinyl glycine on sexual tendency and ethylene evolution in the cucumber plant *(Cucumis sativus* L.). *Plant Cell Physiol.* 20: 1547-1555.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., and Johnson, E. A. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*. *3*(10):.e3376.

Barry, C.S., Blume, B., Bouzayen, M., Cooper, W., Hamilton, A.J. and Grierson, D. 1996. Differential expression of the 1-aminocyclopropane-1-carboxylate oxidase gene family of tomato. *Plant J. 9*(4): 525-535.

Becker, A. and Theissen, G. 2003. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* 29(3): 464-489.

Behera, T. K., Rao, A. R., Amarnath, R., and Kumar, R. R. 2016. Comparative transcriptome analysis of female and hermaphrodite flower buds in bitter gourd

(*Momordica charantia* L.) by RNA sequencing. *J. Hortic. Sci. Biotechnol.* 91(3): 250-257.

Behera, T.K., 2005. Heterosis in bittergourd. *Journal of New Seeds*, *6*(2-3), pp.217-221.

Behera, T.K., Dey, S.S. and Sirohi, P.S. 2006. DBGy-201 and DBGy-202: two gynoecious lines in bitter gourd (*Momordica charantia* L.) isolated from indigenous source. *Indian J Genet*. *66*(1): 61-62.

Behera, T.K., Dey, S.S., Munshi, A.D., Gaikwad, A.B., Pal, A. and Singh, I. 2009. Sex inheritance and development of gynoecious hybrids in bitter gourd (*Momordica charantia* L.). *Scientia horticulturae 120*(1):130-133.

Behera, T.K., John, K.J., Bharathi, L.K., and Karuppaiyan, R. 2011. Momordica. In *Wild crop relatives: genomic and breeding resources* Springer. Berlin, Heidelberg, pp. 217-246.

Bernardes, W. S., & Menossi, M. 2020. Plant 3'Regulatory regions from mRNA encoding genes and their uses to modulate expression. *Front. Plant Sci*. *11.1252*p.

Beyer, E. J. 1976. Silver ion: a potent antiethylene agent in cucumber and tomato. *Hortic. Sci.* 11: 195-196.

Bommesh, J.C., Pitchaimuthu, M., Ravishankar, K.V. and Varalakshmi, B. 2019. Development and Maintenance of Tropical Gynoecious Inbred Lines in Cucumber (*Cucumis sativus*) and Validation by DNA Markers. *Agrl. Res.* pp.1-8.

Boualem, A., Fergany, M., and Fernandez, R. 2008. A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. *Science* 321:836–883.

Boualem, A., Lemhemdi, A., Sari, M. A., Pignoly, S., Troadec, C., Abou Choucha, F., Solmaz, I., Sari, N., Dogimont, C., and Bendahmane, A. 2016. The andromonoecious sex determination gene predates the separation of Cucumis and Citrullus genera. *PloS one 11*(5):e0155444.

Boualem, A., Troadec, C., Camps, C., Lemhemdi, A., Morin, H., Sari, M.A., Fraenkel-Zagouri, R., Kovalski, I., Dogimont, C., Perl-Treves, R. and Bendahmane, A. 2015. A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science 350*(6261):688-691.

Boualem, A., Troadec, C., Kovalski, I., Sari, M. A., and Perl-Treves, R. 2009. A conserved ethylene biosynthesis enzyme leads to andromonoecy in two Cucumis species. *PLoS One* 4:e6144

Bowman, J. L., Smyth, D. R., and Meyerowitz, E. M. 1991. Genetic interactions among floral homeotic genes of Arabidopsis. *Development* 112: 1-20.

Brownie, J., Shawcross, S., Theaker, J., Whitcombe, D., Ferrie, R., Newton, C., and Little, S. 1997. The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Res.* 25(16): 3235-3241.

Carpenter, R. and Coen, E. S. 1990. Floral homeotic mutations produced by transposon-mutagenesis in *Antirrhinum majus*. *Genes Dev.* 4: 1483-149.

Chakravarthy, H. L. 1990. *Momordica charantia* L. In: Fascicles of flora of India. Botanical Survey of India, Calcutta. pp. 87–95.

Charlesworth, D. and Mank, J. E. 2010. The birds and the bees and the fowers and the trees: lessons from genetic mapping of sex determination in plants and animals. *Genetics* 186: 9-31.

Chen, H., Sun, J., Li, S., Cui, Q., Zhang, H., Xin, F., Wang, H., Lin, T., Gao, D., Wang, S., and Li, X. 2016. An ACC oxidase gene essential for cucumber carpel development. *Mol. Plant* 9(9): 1315-1327.

Coen, E. S. and Meyerowitz, E. M. 1991. The war of the whorls: Genetic interactions controlling flower development. *Nature* 353: 31-37.

Collard, B.C., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica 142*(1): 169-196.

Cruden, R. W., and Lloyd, R. M. 1995. Embryophytes have equivalent sexual phenotypes and breeding systems —why not a common terminology to describe them. *Amer. J. Bot.* 82:816–825.

Cui, J., Luo, S., Niu, Y., Huang, R., Wen, Q., Su, J., Miao, N., He, W., Dong, Z., Cheng, J., and Hu, K. 2018. A RAD-based genetic map for anchoring scaffold sequences and identifying QTLs in bitter gourd *(Momordica charantia)*. *Front. Plant Sci*. 9: 477.

Darwin, C., 1876. *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom: By Charles Darwin*. John Murray, Albemarle Street.

Dellaporta, S. L. and Calderon-Urrea, A. 1994. The sex determination process in maize. *Science 266*(5190):1501-1505.

Desai UT and Musmade AM. 1998. Pumpkins, squashes and gourds. In: Salunkhe DK, Kadam SS (eds) *Handbook of vegetable science and technology: production, composition, storage and processing* pp. 273–298

Destefano-Beltran, L. J. C., Caeneghem, W. V., Gielen, J., Richard, L., Montegu, M. V., and Straeten, D. V. D. 1995. Characterization of three members of the ACC synthase gene family in *Solanum tuberosum* L. *Mol. Gen. Genet*. 246: 496-508.

Dey, S. S., Behera, T. K., Munshi, A. D., and Pal, A. 2010. Gynoecious inbred with better combining ability improves yield and earliness in bitter gourd (*Momordica charantia* L.). *Euphytica* 173(1): 37-47.

Dey, S.S., Behera, T., Munshi, A., Rakshit, S., and Bhatia, R. 2012. Utility of gynoecious sex form in heterosis breeding of bitter gourd and genetics of associated vegetative and flowering traits. *Indian J. Hort.* 69:523-529.

Dutta, S., Kumawat, G., Singh, B. P., Gupta, D. K., Singh, S., Dogra, V., Gaikwad, K., Sharma, T. R., Raje, R. S., Bandhopadhya, T. K. and Datta, S. 2011. Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [*Cajanus cajan* (L.) Millspaugh]. *BMC plant biology 11*(1):1-13.

Ecker, J.R., 1995. *The ethylene signal transduction pathway in plants*. *Science*. pp. 667-675.

El Shawaf, I. I. S. and Baker, L.R. 1981. Inheritance of parthenocarpic yield in gynoecious pickling cucumber for onec-over mechanical harvest by diallel analysis of six gynoecious lines. *J. Amer. Soc. Hortic. Sci.* 106(3):359–364.

Esteras, C., Nuez, F., Picó, B., YiHong, W., Behera, T. K., and Kole, C. 2011. Genetic diversity studies in Cucurbits using molecular tools. *Genetics, Genomics and Breeding of Cucurbits 2*(1): 140-198.

Fan, J., Li, W., Dong, X., Guo, W., and Shu, H. 2007. Ectopic expression of a hyacinth AGL6 homolog caused earlier flowering and homeotic conversion in Arabidopsis. *Sci. China C. Life Sci.* 50: 676-689.

Fazio, G. 2002. Comparative study of marker-assisted and phenotypic selection and genetic analysis of yield components in cucumber. *Theor. Appl. Genet.* 119(4):621-34.

Fazio, G., Staub, J. E., and Stevens, M.R. 2003. Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. *Theor. Appl. Genet.* 107(5): 864-874.

Gaikwad, A. B., Saxena, S., Behera T. K., Archak. S., and Meshram, S. 2014. Molecular marker to identify gynoecious line in bitter gourd. *Indian J. Hortic.* 71(1): 142-144.

Gallego, F.J. and Martinez, I., 1996. Molecular typing of rose cultivars using RAPDs. *Journal of Horticultural Science*, *71*(6), pp.901-908.

Galun, E. 1962. Study of the inheritance of sex expression in the cucumber. The interaction of major genes with modifying genetic and non-genetic factors. *Genetica 32*(1):134-163.

Giovannoni, J.J., Wing, R.A., Ganal, M.W. and Tanksley, S.D.1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic acids res. 19*(23): 6553-6568.

Goraichuk, I. 2018. Re: Cause of failed DNA sequencing using reverse primer?. Retrieved from:https://www.researchgate.net/post/Cause_of_failed_DNA_sequencing_using_reverse_primer/5b928e504921ee478c3c16f6/citation/download.

Goujon, M., Mc William, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. 2010. A new bioinformatics analysis tools framework at EMBL– EBI. *Nucleic Acids Res. 38*(2): 695-W699.

Gramzow, L. and Theissen, G. 2010. A hitchhiker's guide to the MADS world of plants. *Genome Biol.* 11(6): 1-11.

Grant, S., Houben, A., Vyskot, B., Siroky, J., Pan, W.H., Macas, J., and Saedler, H. 1994. Genetics of sex determination in flowering plants. *Dev. Genet. 15*(3):214-230.

Grimplet, J., Martinez-Zapater, J. M., and Carmona, M. J. 2016. Structural and functional annotation of the MADS-box transcription factor family in grapevine. *BMC Genomics* 17(1): 1-23.

Gunnaiah, R., Vinod, M.S., Prasad, K. and Elangovan, M. 2014. Identification of candidate genes, governing gynoecy in bitter gourd (*Momordica Charantia* L.) by *in-silico* gene expression analysis. *Int J Computer Appl*. *2*: 5-9.

Guo, S., Zheng, Y., Joung, J. G., Liu, S., Zhang, Z., Crasta, O. R., Sobral, B. W., Xu, Y., Huang, S., and Fei, Z. 2010. Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics* 11(1): 1-13.

Gupta. P. K. and Roy, J. K. 2002. Molecular markers in crop improvement: present status and future needs in India. *Plant Cell Tiss. Org. Cult.* 70(3):229–234.

Gupta. P. K., Balyan, H. S., Sharma, P. C., and Ramesh, B. 1996. Microsatellites in plants: a new class of molecular markers. *Curr. Sci.* 70:45–54.

Hagen, G., and Guilfoyle, T. 2002. Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Mol Biol*. 49: 373-385.

Heiser, C. B. 1979. *The gourd book*. University of Oklahoma Press.

Honma, T. and Goto, K. 2001. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* 409: 525-529.

Hseih and Yin-yin Do. 2019. Genome-wide Identification and analysis of MADS-box gene family in bitter gourd. Conference. doi: 10.13140/RG.2.2.29416.29445.

Hsu, H. F., Huang, C. H., Chou, L. T., Yang, C. H. 2003. Ectopic expression of an orchid (*Oncidium Gower Ramsey*) AGL6-like gene promotes flowering by activating flowering time genes in *Arabidopsis thaliana*. *Plant Cell Physiol.* 44: 783-794.

Huang, P. L., Parks, J. E., Rottmann, W. H., and Theologis, A. 1991. Two genes encoding l-aminocyclopropane-l-carboxylate synthase in zucchini (*Cucurbita pepo)* are clustered and similar but differentially regulate. *Proc. Natl. Acad. Sci. USA* 88:7021-7025.

Huang, X. and Madan, A. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9(9): 868-877.

Ishiguro, S., and Nakamura, K. 1994. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5′ upstream regions of genes coding for sporamin and β-amylase from sweet potato. *Mol. Gen. Genet. 244*(6):563-571.

Iwamoto, E. and Ishida, T. 2006. Development of gynoecious inbred line in balsam pear (*Momordica charantia* L.). *Hortic. Res.*

Jadhav, B. P., Patil M. B., Patil S. Y., Halakude, I. S., and Rajput, J. C. 2018. Development and maintenance of novel source of gynoecy in bittergourd (*Momordica charantia* L.). *Int. J. Curr. Res. 10*(1):63802-63805.

Jat, G. S., Munshi, A. D., Behera, T. K., Choudhary, H., Dash, P., Ravindran, A., and Kumari, S., 2019. Genetics and molecular mapping of gynoecious (F) locus in cucumber (*Cucumis sativus* L.). *J. Hortic. Sci. Biotechnol.* 94(1): 24-32.

John, J. K. and Antony, V. T. 2008. Ethnobotanical investigations in the genus Momordica L. in the Southern Western Ghats of India. *Genet. Resour. Crop Evol.* 55: 713–721.

Johnson, P. R. and Ecker, J. R. 1998. The ethylene gas signal transduction pathway: a molecular perspective. *Annu. Rev. Genet. 32*(1):227-254.

Kahana, A., Silberstein, L., Kessler, N., Goldstein, R.S. and Perl-Treves, R. 1999. Expression of ACC oxidase genes differs among sex genotypes and sex phases in cucumber. *Plant molecular biology*. *41*(4):517-528.

Kamachi, S., Sekimoto, H., Kondo, N., and Sakai, S. 1997. Cloning of a cDNA for a 1 aminocyclopropane-1-carboxylate synthase that is expressed during development of female flowers at the apices of *Cucumis sativus* L. *Plant Cell Physiol.* 38:1197–1206.

Kang, H. G., Noh, Y. S., Chung, Y. Y., Costa, M. A., An, K., and An, G. 1995. Phenotypic alterations of petal and sepal by ectopic expression of a rice MADS box gene in tobacco. *Plant Mol. Biol.* 29: 1-10.

Kater, M. M., Colombo, L., Franken, J., Busscher, M., Masiero, S., Campagne, M. M. V. L., and Angenent, G.C. 1998. Multiple AGAMOUS homologs from cucumber and petunia differ in their ability to induce reproductive organ fate. *The Plant Cell* 10(2): 171-182.

Kater, M. M., Franken, J., Carney, K. J., Colombo, L., and Angenent, G. C. 2001. Sex determination in the monoecious species cucumber is confined to specific floral whorls. *Plant Cell 13*(3):481-493.

Kaufmann, K., Melzer, R., and Theissen, G. 2005. MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347(2): 183-198.

Kempin, S. A., Mandel, M. A., and Yanofsky, M. F. 1993. Conversion of perianth into reproductive organs by ectopic expression of the tobacco floral homeotic gene NAG1. *Plant Physiol.* 103: 1041-1046.

Kende, H. 1993. Ethylene biosynthesis. *Annu. Rev. Plant Biol*. *44*(1):283-307.

Kim, H.S., Abbasi, N. and Choi, S.B. 2013. Bruno-like proteins modulate flowering time via 3' UTR dependent decay of SOC1 mRNA. *New Phytol. 198*(3):747-756

Klee, H.J. and Giovannoni, J.J. 2011. Genetics and control of tomato fruit ripening and quality attributes. *Annu. Rev. Genet.* 45: 41-59.

Knopf, R. R. and Trebitsh, T. 2006. The female specific *Cs-ACS*1G gene of cucumber. A case of gene duplication and recombination between the nonsex specific 1-aminocyclopropane-1-carboxylate synthase gene and a branched-chain amino acid transaminase gene. *Plant Cell Physiol.* 47: 1217-1228.

Kole, C., Olukolu, B. A., Kole, P., Rao, V. K., Bajpai, A., Backiyarani, S., *et al*. 2012). The first genetic map and positions of major fruit trait loci of bitter melon (*Momordica charantia*). *J. Plant Sci. Mol. Breed.* 1:1.

Korbie, D., and Mattick, J. 2008. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* **3:**1452–1456.

Kubicki, B., 1969. Investigations on sex determination in cucumber (*Cucumis sativus* L.). III. Variability of sex expression in the monoecious and gynoecious lines. *Genet Pol*.

Kubicki, B., 1969b. Investigations on sex determination in cucumber (*Cucumis sativus* L.). V. Genes controlling intensity of femaleness. *Genet. Pol. 10*:69-85.

Kubicki, B., 1969. Investigations on sex determination in cucumber (*Cucumis sativus* L.). VII. andromonoecism and hermaphroditism. *Genet Pol*. *10*:101-121.

Kubicki, B., 1972. The mechanism of sex determination in flowering plants.

Kumari, M. 2015. Physiological and molecular characterisation of gynoecious line and genetics of fruit traits in bitter gourd *(Momordica charantia L.)*. Doctoral dissertation. Division of vegetable science Indian Agricultural Research Institute, New Delhi.

Lei, Y., Sun, Y., Wang, B., Yu, S., Dai, H., Li, H., Zhang, Z., and Zhang, J. 2020. Woodland strawberry WRKY71 acts as a promoter of flowering via a transcriptional regulatory cascade. *Hortic. Res. 7*(1):1-14.

Li, H., Liang, W., Jia, R., Yin, C., Zong, J., Kong, H., and Zhang, D. 2010. The AGL6-like gene OsMADS6 regulates floral organ and meristem identities in rice. *Cell Res.* 20(3): 299-313.

Li, S. B., Xie, Z. Z., Hu, C. G., and Zhang, J.Z., 2016. A review of auxin response factors (ARFs) in plants. *Front. Plant Sci. 7:*47.

Li, T., Jiang, Z., Zhang, L., Tan, D., Wei, Y., Yuan, H., Li, T., and Wang, A. 2016. Apple (*Malus domestica*) MdERF2 negatively affects ethylene biosynthesis during fruit ripening by suppressing MdACS1 transcription. *Plant J.* 88: 735-748.

Li, Z., Han, Y., Niu, H., Wang, Y., Jiang, B. and Weng, Y. 2020. Gynoecy instability in cucumber (*Cucumis sativus* L.) is due to unequal crossover at the copy number variation-dependent Femaleness (F) locus. *Hortic. Res. 7*(1):1-15.

Li, Z., Liu, S., and Pan, J. 2009. Molecular isolation of the M gene suggests that a conserved -residue conversion induces the formation of bisexual flowers in cucumber plants. *Genetics* 182: 1381-1385.

Li, Z., Wang, S., Tao, Q., Pan, J., Si, L., Gong, Z., and Cai, R. 2012. A putative positive feedback regulation mechanism in *CsACS2* expression suggests a modified model for sex determination in cucumber (*Cucumis sativus* L.). *J. Exp. Bot.* 63: 4475-4484.

Liang, X., Abel, S., Keller, J. A., Shen, N. F., and Theologis, A. 1992. The 1-aminocyclopropane-l-carboxylate synthase gene family of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 89:11046-11050.

Lieberman, N. 1979. Biosynthesis and action of ethylene. *Annu. Rev. Plant Physiol.* 30: 533-591.

Lin, Z., Zhong, S., and Grierson, D. 2009. Recent advances in ethylene research. *J. Exp. Botany* 60(12): 3311-3336.

Liu, G., Li, C., Yu, H., Tao, P., Yuan, L., Ye, J., Chen, W., Wang, Y., Ge, P., Zhang, J. and Zhou, G. 2020. GREEN STRIPE, encoding methylated TOMATO AGAMOUS-LIKE 1, regulates chloroplast development and Chl synthesis in fruit. *New Phytol. 228*(1): 302-317.

Liu, G., Li, C., Yu, H., Tao, P., Yuan, L., Ye, J., Chen, W., Wang, Y., Ge, P., Zhang, J. and Zhou, G. 2020. Green stripe, encoding methylated tomato agamous-like 1, regulates chloroplast development and Chl synthesis in fruit. *New Phytol. 228*(1):302-317.

Liu, K., Yuan, C., Li, H., Lin, W., Yang, Y., Shen, C., and Zheng, X. 2015. Genome-wide identification and characterization of auxin response factor (ARF) family genes related to flower and fruit development in papaya (*Carica papaya* L.). *BMC Genomics*, *16*(1):1-12.

Liu, M., Gomes, B. L., Mila, I., Purgatto, E., Peres, L. E., Frasse, P., Maza, E., Zouine, M., Roustan, J. P., Bouzayen, M., and Pirrello, J. 2016. Comprehensive profiling of ethylene response factor expression identifies ripening-associated ERF genes and

their link to key regulators of fruit ripening in tomato. *Plant Physiol.* 170(3): 1732-1744.

Luo, X., Sun, X., Liu, B., Zhu, D., Bai, X., Cai, H., Ji, W., Cao, L., Wu, J., Wang, M. and Ding, X. 2013. Ectopic expression of a WRKY homolog from Glycine soja alters flowering time in Arabidopsis. *PLoS One*, *8*(8), p.e73295

Ma, H., Yanofsky, M. F., and Meyerowitz, E. M. 1991. AGL1– AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev.* 5: 484-495.

Malepszy, S. and Niemirowicz-Szczytt, K. 1991. Sex determination in cucumber (*Cucumis sativus*) as a model system for molecular biology. *Plant science*, *80*(1-2):39-47.

Mammadov, J., Aggarwal, R., Buyyarapu, R., and Kumpatla, S. 2012. SNP markers and their impact on plant breeding. *Int. J. Plant Genomics*. 1–11.

Mandel, M. A., Bowman, J. L., Kempin, S. A., Ma, H., Meyerowitz, E. M., and Yanofsky, M. F. 1992. Manipulation of floral structure in transgenic tobacco. *Cell* 71: 133-143.

Manye, Y., Maojun, Z., Ying, X., Lin, T., Jie, B., and Fan, C. 2004. Cloning and expression of a MADS box gene from bitter melon (*Momordica charantia* L.). *J. Beijing For. Univ.* 26(4): 30-34.

Manzano, S., Aguado, E., Martínez, C., Megías, Z., García, A. and Jamilena, M. 2016. The ethylene biosynthesis gene CitACS4 regulates monoecy/andromonoecy in watermelon *(Citrullus lanatus*). *PloS one 11*(5).

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fermandes, R., Morin, H., Pitrat, M., Dogimont, C., and Bendahmane, A. 2009. A transposon-induced epigenetic change leads to sex determination in melon. *Nature* 461: 1135-1139.

Masiero, S., Colombo, L., Grini, P. E., Schnittger, A., and Kater, M. M. 2011. The emerging importance of type I MADS box transcription factors for plant reproduction. *Plant Cell* 23: 865-872.

Matasyoh, L.G., Wachira, F.N., Kinyua, M.G., Muigai, A.W.T. and Mukiama, T.K., 2008. Leaf storage conditions and genomic DNA isolation efficiency in *Ocimum gratissimum* L. from Kenya. *Afr. J. Biotechnol. 7*(5).

Matsumura, H., Hsiao, M.C., Lin, Y.P., Toyoda, A., Taniai, N., Tarora, K., Urasaki, N., Anand, S.S., Dhillon, N.P., Schafleitner, R. and Lee, C.R. 2020. Long-read bitter gourd *(Momordica charantia*) genome and the genomic architecture of nonclassic domestication. *Proceedings of the National Academy of Sciences*.*117*(25): 14543-14551.

Matsumura, H., Miyagi, N., Tanaiai, N., Fukushima, M., Tarora, K., Shudo, A., and Urasaki, N. 2014. Mapping of the gynoecy in bitter gourd (*Momordica charantia*) using RAD-seq analysis. *PLoS ONE* 9 (1): e87138.

Mattoo, A. K. and Suttle, J. C. 1991. The plant hormone ethylene. CRC press. Boca Raton, Florida, pp. 337.

Meagher, T. R. 2007. Linking the evolution of gender variation to floral development. *Ann. Bot.* 100(2):165– 176.

Meena, P. 2014. Molecular characterization and DNA fingerprinting of selected cashew (*Anacardium occidentale* L.) varieties of KAU. M.Sc. thesis, Kerala Agricultural University, Thrissur, 120 p.

Miao, H., Zhang, S., Wang, X., Zhang, Z., Li, M., Mu, S., Cheng, Z., Zhang, R., Huang, S., Xie, B., and Fang, Z. 2011. A linkage map of cultivated cucumber (*Cucumis sativus* L.) with 248 microsatellite marker loci and seven genes for horticulturally important traits. *Euphytica.* 182(2): 167-176.

Mibus, H. and Tatlioglu, T. 2004. Molecular characterization and isolation of the F/f gene for femaleness in cucumber (*Cucumis sativus* L.). *Theor. Appl. Genet.* 109: 1669-1676.

Miniraj N, Prasanna KP, Peter KV. 1993. Bitter gourd (*Momordica* spp.). In: Kalloo G, Bergh BO (eds) Genetic improvement of vegetable crops. Pergamon Press, Oxford. pp. 239–246

Minkov, A. S., Levi, A., Wolf, S., and Trebtish, T. 2008. ACC Synthase genes are polymorphic in watermelon (*Citrullus spp.*) and differentially expressed in flowers and in response to auxin and gibberellins. *Plant Cell Physiol.* 49(5):740–750.

Minnu Ann, J. 2019. Breeding for gynoecy in bitter gourd (*Momordica charantia* L.). M.Sc. (Ag.) Thesis. Kerala Agricultural University, Thrissur.

Mishra, S. 2012. Genetics of economic traits and sex modification in bitter gourd *(Momordica charantia L.).* PhD Thesis, Division of Vegetable Science, IARI.

Mishra, S., Behera, T. K, Munshi, A.D., Gaikwad, K., and Mohapatra, T. 2014. Identification of RAPD markers associated with gynoecious trait ('gy-1' gene) in bitter gourd (*Momordica charantia* L.). *Aust. J. Crop Sci*. 8(5): 706-710.

Mishra, S., Behera, T. K., Munshi, A. D., Bharadwaj, C., and Rao, A. R. 2015. Inheritance of gynoecism and genetics of yield and yield contributing traits through generation mean analysis in bitter gourd. *Indian J. Hortic.* 72(2): 218-222.

Mohanty, J. N. and Joshi, R. K. 2018. Molecular cloning, characterization and expression analysis of MADS-box genes associated with reproductive development in *Momordica dioica* Roxb. *Biotech* 8(3): 1-13.

Morgante, M., and Olivieri, A. 1992. PCR-amplified microsatellites as markers in plant genetics. *Plant J.* 3:175–182.

Nakagawa, N., Mori, H., Yamazaki, K., and Imaseki, H. 1991. Cloning of a complementary DNA for auxin-induced 1-aminocycIopropane-l-carboxylate synthase and

differential expression of the gene by auxin and wounding. *Plant Cell Physiol*. 32:1153-1163.

Neal, P. R. and Anderson, G. J. 2005. Are 'mating systems' 'breeding systems' of inconsistent and confusing terminology in plant reproductive biology? or is it the other way around?. *Plant Sys. Evol.* 250(3-4):173–185

Olson, D. C., White, J. A., Edlman, L., Harkins, R. N., and Kende, H. 1991. Differential expression of two genes for 1-aminocyclopropane-l-carboxylate synthase in tomato fruits. *Proc. Natl. Acad. Sci. USA* 88:5340-5344.

Pawelkowicz, M., Pryszcz, L., Skarzynska, A., Woycicki, R. K., Posyniak, K., Rymuszka, J., Przybecki, Z., and Plader, W. 2019. Comparative transcriptome analysis reveals new molecular pathways for cucumber genes related to sex determination. *Plant Reprod.* 1-24.

Peakall, R., Gilmore, S., Keys, W., Morgante, M., and Rafalski, A. 1998. Cross-species amplification of Soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for the transferability of SSRs in plants. *Mol. Biol. Evol.* 15:1275–1287.

Peng, J., Carol, P., Richards, D. E., King, K. E., Cowling, R. J., Murphy, G. P., and Harberd, N. P. 1997. The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev.* 11:3194– 3205.

Peng, S. M., Luo, T., Zhou, J. Y., Niu, B., Lei, N. F., Tang, L., and Chen, F. 2008. Cloning and quantification of expression levels of two MADS-box genes from *Momordica charantia*. *Biol. Plant.* 52(2): 222-230.

Perl-Treves R and Rajagopalan PA. 2006. Close, yet separate: patterns of male and female floral development in monecious species. In: Ainsworth C (ed) Flowerdevelopment and manipulation. Blackwell, Oxford.

Perl-Treves, R. 1999. Male to female conversion along the cucumber shoot: approaches to study sex genes and floral development in *Cucumis sativus.* In: Ainsworth, C. C. (ed). Sex determination in plants. Bios Scientific Publication, Oxford, UK. pp. 189-215.

Perl-Treves, R. and Rajagopalan, P. A., 2018. Close, yet separate: patterns of male and female floral development in monoecious species. *Annual Plant Reviews online* pp.117-146.

Perl-Treves, R., Kahana, A., Rosenman, N., Xiang, Y., and Silberstein, L. 1998. Expression of multiple AGAMOUS-like genes in male and female flowers of cucumber (*Cucumis sativus* L.). *Plant Cell Physiol.* 39(7): 701-710.

Pitrat, M. 2008. Melon. In: Prohens, J. and Nuez, F. (eds) Handbook of plant breeding: vegetable: Vol. 1. Springer Science, New York, pp. 283-316.

Pnueli, L., Hareven, D., Rounsley, S. D., and Yanofsky, M. F. 1994. Isolation of the tomato AGAMOUS gene TAG1 and analysis of its homeotic role in transgenic plants. *Plant Cell* 6: 163-173.

Poole, C. F. and Grimball, P. C. 1939. Inheritance of new sex forms in *Cucumis melo* L. *J. Hered.* 30: 21-25.

Poritz, Mark A., and Kirk M. Ririe. 2014. Getting things backwards to prevent primer dimers. *J. Mol. Diagn.* 16(2): 159-162.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. 1996 The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for gemplasm analysis. *Mol. Breed.* 2:225-238.

Purugganan, M. D., Rounsley, S. D., Schmidt, R. J., and Yanofsky, M.F. 1995. Molecular evolution of flower development: Diversification of the plant MADS-box regulatory gene family. *Genetics.* 140: 345-356.

Pysh, L. D., Wysocka-Diller, J. W., Camilleri, C., Bouchez, D., Benfey, P. N. 1999. The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. *Plant J.* 18:111– 119.

Ram, D., 2002. Occurrence, identification and preliminary characterization of gynoecism in bitter gourd (*Momordica charantia*). *Indian J. Agr. Sci.* 72: 348-349.

Ram, D., Kumar, S., Banerjee, M. K., Singh, B., and Singh, S. 2002. Developing bitter gourd (*Momordica charantia* L.) populations with a very high proportion of pistillate flowers. *REPORT-CUCURBIT GENETICS COOPERATIVE* 25: 65-66.

Ram, D., Kumar, S., Singh, M., Rai, M., and Kalloo, G. 2006. Inheritance of gynoecism in bitter gourd (*Momordica charantia* L.). *J. Hered.* 97(3): 294-295.

Rao, G., Behera, T. K., Gaikwad, A. B., Munshi, A. D., Jat, G. S., and Boopalakrishnan, G. 2018. Mapping and QTL analysis of gynoecy and earliness in bitter gourd (*Momordica charantia* L.) using Genotyping-by-Sequencing (GBS) technology. *Front. Plant Sci.* 9: 1555.

RAPDs. *J. Hort. Sci.* 71(6): 901-908.

Reshmika, P. K., Pradeepkumar, T., Krishnan, S., and Sureshkumar, P. 2019. Evaluation of bitter gourd Hybrids. *Electronic. J. Plant Breed. 10*(4):1617-1623.

Robinson, R. W. and Decker-Walters, D. S., 1997. *Cucurbits*. Cab international.

Robinson, R. W., and Decker-Walters, D.S. 1999. *Cucurbits*. CAB International, Wallingford, Oxford, UK, p. 226.

Robinson, R. W., RW, R., HM, M., TW, W., and GW, B. 1976. Genes of the Cucurbitaceae.

Rogers, S.O. and Bendlich, A.J. 1994. Extraction of total cellular DNA from plants, algae and fungi. In: Gelvin, S.B. and Schilperoort, R.A. (eds.) Plant Molecular Biology Manual, 2nd Edition, Kluwer Academic Publishers, Dordrecht, pp. 1-8.

Rottmann, W. H., Peter, G. F., Oeller, P. W., Keller, J. A., Shen, N. F., Nagy. B. P., Taylor, L. P., Campbell, A. D., and Theologis, A. 1991. 1- Aminocyclopropane-1-carboxylate synthase in tomato is encoded by a multigene family whose transcription is induced during fruit and floral senescence. *J. Mol. Biol.* 222:937-961.

Rudich, J., Baker, L. R., Scott, J. W., and Sell, H. M. 1976. Phenotypic stability and ethylene evolution in androecious cucumber. *Amer. Soc. Hort. Sci.* 101: 48-51.

Rudich, J., Halevy, A. H., and Kedar, N. 1972a. Ethylene evolution from cucumber plants as related to sex expression. *Plant Physiol.* 49(6): 998-999.

Rudich, J., Halevy, A. H., and Kedar, N. 1972b. The level of phytohormones in monoecious and gynoecious cucumbers as affected by photoperiod and ethephon. *Plant Physiol.* 50(5): 585-590.

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. 2010. WRKY transcription factors. *Trends Plant Sci. 15*(5):247-258.

Saedler, H., and Huijser, P. 1994. Molecular biology of flower development in *Antirrhinum majus* (snapdragon). *Gene* 135: 239-243.

Saha, K., Bhoi, S. and Mishra, P.K.S. 2018. Molecular breeding for gynoecy trait in Bitter gourd (*Momordica charantia* L.). *J. Pharmacogn. Phytochem. 7*(5): 86-89.

Saito, S., Fujii, N., Miyazawa, Y., Yamasaki, S., Matsuura, S., Mizusawa, H., Fujita, Y. and Takahashi, H. 2007. Correlation between development of female flower buds and expression of the Cs-ACS2 gene in cucumber plants. *J. Exp. Bot*. 58:2897–2907.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. *Molecular cloning: a laboratory manual* (Ed. 2). Cold spring harbor laboratory press.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain- terminating inhibitors. *Proceedings of the national academy of sciences*, 74(12): 5463-5467.

Schaefer, H and Renner, S.S. (2010). A three-genome phylogeny of *Momordica* (Cucurbitaceae) suggests seven returns from dioecy to monoecy and recent long-distance dispersal to Asia. *Mol Phylogenet. Evol*. 54:553–560.

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. *Proc. R.l Soc. B: Biol Sci. 276*(1658):843-851.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., and Sommer, H. (1990). Genetic control of flower development: Homeotic genes in Antirrhinum majus. Science 250, 931–936.

Sharma, B. and Kramer, E. 2013. Sub-and neo-functionalization of APETALA3 paralogs have contributed to the evolution of novel floral organ identity in Aquilegia (columbine, Ranunculaceae). *New Phytol.* 197(3): 949-957.

Shen, C., Yue, R., Yang, Y., Zhang, L., Sun, T., Xu, L., Tie, S., and Wang, H. 2014. Genome-wide identification and expression profiling analysis of the Aux/IAA gene family in *Medicago truncatula* during the early phase of *Sinorhizobium meliloti* infection. *PLoS One*, *9*(9):e107495.

Shifriss, O. 1961. Sex control in cucumbers. *J. Hered*. 52(1):5–12.

Shukla, A., Singh, V. K., Bharadwaj, D. R., Kumar, R., Rai, A., Rai, A. K., Mugasimangalam, R., Parameshwaran, S., Singh, M., and Naik, P. S. 2015. *De novo* assembly of bitter gourd transcriptomes: Gene expression and sequence variations in gynoecious and monoecious lines. *PLoS ONE* 10(6): e0128331.

Silverstone, A. L., Ciampaglio, C. N., and Sun, T. P. 1998. The Arabidopsis RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. *Plant Cell 10*(2):155-169.

Singh, A. K., Chaurasia, S., Kumar, S., Singh, R., Kumari, J., Yadav, M. C., Singh, N., Gaba, S., and Jacob, S.R. 2018. Identification, analysis and development of salt

responsive candidate gene based SSR markers in wheat. *BMC Plant Biol.* 18(1): 1-15.

Singh, K., Foley, R. C., and Oñate-Sánchez, L. 2002. Transcription factors in plant defense and stress responses. *Curr Opin Plant Biol.* 5(5): 430-436.

Smaczniak, C., Immink, R. G., Angenent, G. C., and Kaufmann, K. 2012. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development* 139(17): 3081-3098.

Staub, J. E., Robbins, M. D., and Wehner, T. C. 2008. Cucumber. In: Prohens, J. and Nuez, F., (eds) Handbook of plant breeding; Vegetables I: asteraceae, brassicaceae, chenopodiaceae, and cucurbitaceae. Springer, NY, pp. 241–282

Stokes, T. L., Kunkel, B. N., and Richards, E. J. 2002. Epigenetic variation in Arabidopsis disease resistance. *Genes Dev.* 16(2):171–182.

Su, Y., Liu, J., Liang, W., Dou, Y., Fu, R., Li, W., Feng, C., Gao, C., Zhang, D., Kang, Z., and Li, H. 2019. Wheat AGAMOUS LIKE 6 transcription factors function in stamen development by regulating the expression of Ta APETALA3. *Development* 146(20).

Takahashi, H. and Jaffe, M. J. 1984. Further studies of auxin and ACC induced feminization in the cucumber plant using ethylene inhibitors. *Phyton* 44(1): 81-86.

Takahashi, H. and Suge, H. 1980. Sex expression in cucumber plants as affected by mechanical stress. *Plant Cell Physiol.* 21(2): 303-310.

Tang X, Wang H, Brandit A. S. and Woodson W. R. 1993. Organization and structure of the 1-aminocyclopropane-1-carboxylate oxidase gene family *Petunia hybrid. Plant Mol. Bio.* 23: 1151-1164.

Tanksley, S.D. 1983. Gene mapping. In: *Developments in Plant Genet. Breed.* Elsevier. 1: 109-138.

Tanurdzic, M. and Banks, J. A., 2004. Sex-determining mechanisms in land plants. *Plant Cell 16*(1):S61-S71.

Tao, Q., Niu, H., Wang, Z., Zhang, W., Wang, H., Wang, S., Zhang, X., and Li, Z., 2018. Ethylene responsive factor ERF110 mediates ethylene-regulated transcription of a sex determination-related orthologous gene in two Cucumis species. *J. Exp. Bot.* 69(12): 2953-2965.

Thakur, P.P., Mathew, D., Nazeem, P.A., Abida, P.S., Indira, P., Girija, D., Shylaja, M.R. and Valsala, P.A. 2014. Identification of allele specific AFLP markers linked with bacterial wilt [*Ralstonia solanacearum* (Smith) Yabuuchi *et al*.] resistance in hot peppers (*Capsicum annuum* L.). *Physiol.Mol. plant Pathol*.*87*:19-24.

Theissen, G. 2001. Development of floral organ identity: stories from the MADS house. *Curr. Opin. Plant Biol.* 4(1): 75-85.

Theissen, G. and Saedler, H. 2001. Floral quartets. *Nature* 409(6819): 469-471.

Tiwari, S. B., Hagen, G., and Guilfoyle, T. 2003. The roles of auxin response factor domains in auxin-responsive transcription. *Plant Cell.* 15: 533-543.

Trebitsh, T., Rudich, J., and Riov, J. 1987. Auxin, biosynthesis of ethylene and sex expression in cucumber *(Cucumis sativus). Plant Growth Regul.* 5: 105-113.

Trebitsh, T., Staub, J. E., and O'Neill, S. D. 1997. Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the female (F) locus that enhances female sex expression in cucumber. *Plant Physiol. 113*(3):987-995.

Tsuchimoto, S., Van der Krol, A. R., and Chua, N.-H. 1993. Ectopic expression of pMADS3 in transgenic petunia phenocopies the petunia blind mutant. *Plant Cell* 5: 843-853.

Ulmasov, T., Hagen, G., Guilfoyle, T. J., 1997. ARF1, a transcription factor that binds to auxin response elements. *Science*. 276: 1865-1868.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S.G. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40(15): e115-e115.

Urasaki, N., Takagi, H., Natsume, S., Uemura, A., Taniai, N., Miyagi, N., Fukushima, M., Suzuki, S., Tarora, K., Tamaki, M. and Sakamoto, M. 2017. Draft genome sequence of bitter gourd (*Momordica charantia*), a vegetable and medicinal plant in tropical and subtropical regions. *Dna Res. 24*(1): 51-58.

Varalakshmi, B., Pitchaimuthu, M., Rao, E. S., Krishnamurthy, D., Suchitha, Y., and Manjunath, K.S.S. 2014. Identification preliminary characterization and maintenance of gynoecious plants, IIHRBTGy-491 and IIHRBTGy-492 in bitter gourd. In *Int Bitter Gourd Conf (BiG2014). Hyderabad, March, AVRDC at ICRISAT* (Vol. 36).

Wang QM and GW Zeng. 1997. Morphological and histochemical study on sex differentiation on Momordica charantia. *Journal of Zhejiang Agriculture University*. 23: 149-153.

Wang, D. H., Li, F., Duan, Q. H., Han, T., Xu, Z. H., and Bai, S. N. 2010. Ethylene perception is involved in female cucumber flower development. *Plant J.* 61:862–872.

Wangkumhang, P., Chaichoompu, K., Ngamphiw, C., Ruangrit, U., Chanprasert, J., Assawamakin, A. and Tongsima, S. 2007. WASP: a Web-based Allele-Specific PCR assay designing tool for detecting SNPs and mutations. *BMC genomics 8*(1):1-9.

Weigel, D. and Meyerowitz, E. M. 1994. The ABCs of floral homeotic genes. *Cell* 78: 203-209.

Weising, K., Atkinson, R.G. and Gardner, R.C., 1995. Genomic fingerprinting by microsatellite-primed PCR: a critical evaluation. *Genome Res. 4*(5):249-255.

Wettasinghe, R. and Peffley, E.B., 1998. A rapid and efficient extraction method for onion DNA. *Plant breed. 117*(6):588-589.

Win, K.T., Zhang, C., Song, K., Lee, J.H. and Lee, S. 2015. Development and characterization of a co-dominant molecular marker via sequence analysis of a genomic region containing the Female (F) locus in cucumber (*Cucumis sativus* L.). *Mol. Breed. 35*(12): 229.

Wu, T., Qin, Z., Zhou, X., Feng, Z. and Du, Y., 2010. Transcriptome profile analysis of floral sex determination in cucumber. *J. plant physiol*. *167*(11): 905-913.

Xiao, Y. Y., Chen, J. Y., Kuang, J. F., Shan, W., Xie, H., Jiang, Y. M., and Lu, W. J. 2013. Banana ethylene response factors are involved in fruit ripening through their interactions with ethylene biosynthesis genes. *J. Exp. Bot*. 64: 2499-2510.

Yamasaki, S., Fujii, N. and Takahashi, H. 2003. Photoperiodic regulation of CsACS2, CS-ACS4 and CS-ERS gene expression contributes to the femaleness of cucumber flowers through diurnal ethylene production under short-day conditions. *Plant Cell Environ*. 26:537–546.

Yamasaki, S., Fujii, N., and Takahashi, H. 2000. The ethylene regulated expression of CS-ETR2 and CS-ERS genes in cucumber plants and their possible involvement with sex expression in flowers. *Plant Cell Physiol.* 41:608– 616.

Yamasaki, S., Fujii, N., Matsuura, S., Mizusawa, H., and Takahashi, H. 2001. The M locus and ethylene-controlled sex determination in andromonoecious cucumber plants. *Plant Cell Physiol.* 42: 608-619.

Yamazaki, Y., Fujii, N., and Takahashi, H. 2003. Characterization of ethylene effects on sex determination in cucumber plants. *Sex. Plant Reprod.* 16:103–111.

Yang, S. F. and Hoffman, N. E. 1984. Ethylene biosynthesis and its regulation in higher plants. *Annu. Rev. Plant Physiol.* 35:155–189.

Yang, F.S., 1985. Biosynthesis and action of ethylene. *Hort Science*, *20*(1): 41-45.

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A., and Meyerowitz, E.M. 1990. The protein encoded by the Arabidopsis homeotic gene agamous resembles transcriptional factors. *Nature* 346: 35-39.

Yin, T. and Quinn, J. A., 1995. Tests of a mechanistic model of one hormone regulating both sexes in *Cucumis sativus* (Cucurbitaceae). *Am. J. Bot.* 82(12): 1537-1546.

Yin, W., Yu, X., Chen, G., Tang, B., Wang, Y., Liao, C., Zhang, Y., and Hu, Z. 2018. Suppression of SlMBP15 inhibits plant vegetative growth and delays fruit ripening in tomato. *Front. Plant Sci.* 9: 938.

Yuan, J., Njiti, V. N., Meksem, K., Iqbal, M. J., Triwitayakorn, K., Kassem, M. A. *et al.* 2002. Quantitative trait loci in two soybean recombinant inbred line populations segregating for yield and disease resistance. *Crop Sci*. 42: 271–277.

Zarembinski, T. I. and Theologis, A. 1993. Anaerobiosis and plant growth hormones induce two genes encoding 1-aminocyclopropane-1-carboxylate synthase in rice (*Oryza sativa* L.). *Molecular Biology of the Cell*, *4*(4):363-373.

Zhang, L., Chen, L., and Yu, D. 2018. Transcription factor WRKY75 interacts with DELLA proteins to affect flowering. *Plant Physiol. 176*(1):790-803.

Zhang, Z., Zhang, H., Quan, R., Wang, X. C., and Huang, R. 2009. Transcriptional regulation of the ethylene response factor LeERF2 in the expression of ethylene biosynthesis genes controls ethylene production in tomato and tobacco. *Plant Physiol.* 150(1): 365-377.

Zhou, W. B., Luo, S. B., and Luo, J.N. 1998. *An early maturing and high yielding bitter gourd F1 hybrid-'Cuilu No. 1*.

Zouine, M., Fu, Y., Chateigner-Boutin, A.L., Mila, I., Frasse, P., Wang, H., Audran, C., Roustan, J.P., and Bouzayen, M. 2014. Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. *PloS one*, *9*(1):e84203.

Annexure

ANNEXURE I

List of laboratory equipments used for the study

Refrigerated centrifuge	:	Kubota 6500, Japan , Eppendorf
Horizontal Refrigerated centrifuge	:	BIO-RAD, USA
Thermal cycler	:	Sure Cycler 8800 (Agilent Technology), Veriti thermal cycler (Applied Biosystems, USA) and ProFlex PCR system (Applied Biosystems, USA)
Gel documentation system	:	Gel Doc^{TM} XR^{+} (BIORAD)
Micropipettes	:	Eppendorf
Spectrophotometer	:	NanoDrop^{R} ND-1000

ANNEXURE II

Reagents required for DNA isolation

1. 2X CTAB extraction buffer (100ml)

CTAB	:	2 g
Tris base	:	1.21 g
EDTA	:	0.745 g
NaCl	:	8.18 g
PVP	:	1.0 g

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

2. 10 per cent CTAB solution

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

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

4. Chilled isopropanol

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

5. Ethanol (70 %)

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

6. TE buffer (pH 8, 100ml)

Tris HCl (10 mM) : 0.1576 g
EDTA (1 mM) : 0.0372g

(The solution was prepared, autoclaved and stored at room temperature)

ANNEXURE III

Composition of buffers and dyes used for gel electrophoresis

a. TAE buffer (50X) (for 1 L)

Tris Acetic Acid	242 g
Glacial Acetic Acid	57.1 mL
0.5 M EDTA (pH 8)	100 mL

(final volume made up to 1 L)

b. Loading dye (6X)

0.25 per cent bromophenol blue

0.25 per cent xylene cyanol 30 per cent glycerol in water

c. Ethidium Bromide

The dye was prepared as a stock solution of 10 mg/ml in water and was stored at room temperature in dark bottle.

ANNEXURE IV

Reagents in the gel elution and purification kit (QIAquick)

- Buffer QG (solubilization buffer) with pH indicator (yellow colour)
- Buffer PE (96-100% ethanol added)
- Elution Buffer

(All the contents of the kit needs to be stored in 15-25°C)

MOLECULAR CHARACTERIZATION OF GYNOECY IN BITTER GOURD (*Momordica charantia* L.)

by

NIVETHITHA B.

(2018-11-007)

ABSTRACT

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Agriculture

(Plant Biotechnology)

Faculty of Agriculture

Kerala Agricultural University

DEPARTMENT OF PLANT BIOTECHNOLOGY

COLLEGE OF AGRICULTURE

VELLANIKKARA, THRISSUR - 680656

KERALA, INDIA

2021

ABSTRACT

Bitter gourd *(Momordica charantia* L.) is an important vegetable of tropics and sub-tropics of Asia belonging to Cucurbitaceae family. The immature fruits of bitter gourd are valued for its culinary and medicinal importance. Having highest amount of ascorbic acid and iron content, it is considered to be the most nutritive among all the cucurbitaceous vegetables. Heterosis is well exploited in bitter gourd for early maturity, increased yield and other agronomic traits through development of hybrids. However, the production of hybrids is labour-intensive involving manual bagging and hand pollination, thereby increasing the cost of seed production.

The predominant sex form of bitter gourd is monoecious which bears separate male and female flowers on the same plant. However, gynoecious type bearing only female flowers were also reported in few locations of India, Japan and China. These gynoecious bitter gourd lines can be exploited as female parent to make hybrid seed production economical and easier, as it eliminates the need for emasculation and assisted pollination. It also aids in maintaining the genetic purity of hybrids and helps in harnessing the benefit of hybrid vigour including early maturity and high yield.

Usually, the sex expression in cucurbits is highly influenced by environment and hormones, which makes the early phenotypic identification of gynoecy challenging. Identification of molecular markers tightly linked to gynoecy trait would ease the identification of gynoecious line in breeding programmes.

In this study, 20 putative candidate genes governing sex expression were selected from different literatures and the gene sequences were retrieved from bitter gourd and cucumber genome assembly in NCBI GenBank database. Four ethylene biosynthesis genes (*ACO1, ACS2, ACS3. ACS4* and *ACS7*), three MADS-box transcription factor encoding genes (*AG6, MADS-boxTF23* and *McAG2*), one auxin related gene (*CsARF5*) and 12 WRKY transcription factor encoding genes comprise the list. A total of 20 gene-specific primer sets were designed from the selected regionof each of the 20 genes. Genomic DNA was isolated from monoecious genotype, MC-136 and gynoecious, KAUMCGy-101. Efforts were made to amplify all the 20

genes, however, only seven gene-specific primer sets designed from bitter gourd produced PCR amplification. The amplicons of expected product size from both the samples were sequenced.

Sequence variation analysis was done by comparing the monoecious and gynoecious sequences generated to the reference sequence (monoecious) available in NCBI database for bitter gourd. Six valid variations including three SNPs and three In/Dels were found in *AG6* and *McAG2*. All the variations except SNP1 of *AG6* were present in non-coding regions.

In/Del of 48 bp ($[TC]_{24}$) in *AG6* gene caused a significant difference in the number of 'TC' repeats in the sequences which was used for the development of SSR marker. The marker showed clear polymorphism between monoecious and gynoecious genotypes used. It was further validated in three other monoecious lines namely, Priya, Priyanka and the wild type *M. charantia* var. *muricata.* Expected banding pattern showing polymorphism for gynoecious line was obtained for all the three monoecious lines screened.

Thus, the study identified two potential candidate genes*, AG6* and *McAG2* for sex expression in bitter gourd. The SSR marker developed needs to be validated in large number of population and more number of varieties to confirm its use as a reliable polymorphic marker in marker-assisted selection of gynoecious lines of bitter gourd.