MOLECULAR CHARACTERIZATION OF GEMINIVIRUS CAUSING YELLOW VEIN MOSAIC IN PUMPKIN (*Cucurbita moschata* Duch. Ex Poir.)

By

SAHNA HAMSA N.H

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680 656

KERALA, INDIA

2011

DECLARATION

I, hereby declare that this thesis entitled **"Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex Poir.)"** is a bonafide record of research work done by me during the course of research and that it has not been previously formed the basis for the award to me of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara
09-1-12

Sahna Hamsa N.H
(2009-11-116)

CERTIFICATE

Certified that this thesis entitled **"Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex Poir.)"** is a bonafide record of research work done independently by **Ms. Sahna Hamsa N.H** under my guidance and supervision and that it has not formed the basis for the award of any degree, diploma, fellowship or associateship to her.

Vellanikkara
09-1-12

Dr. D. Girija
(Major Advisor, Advisory Committee)
Professor and Head
Department of Agricultural Microbiology
College of Horticulture
Vellanikkara, Thrissur

CERTIFICATE

We, the undersigned members of the advisory committee of **Ms. Sahna Hamsa N. H** a candidate for the degree of **Master of Science in Agriculture**, with major field in **Plant Biotechnology**, agree that the thesis entitled **"Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex Poir.)"** may be submitted by **Ms. Sahna Hamsa N.H** in partial fulfillment of the requirement for the degree.

Dr. D. Girija
(Major Advisor, Advisory Committee)
Professor and Head
Department of Agricultural Microbiology
College of Horticulture
Vellanikkara, Thrissur

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

Dr. R. Keshavachandran
(Member, Advisory committee)
Professor
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture
Vellanikkara, Thrissur.

Dr. Sally K. Mathew 9/1/12
(Member, Advisory committee)
Professor and Head
Dept. of Plant Pathology
College of Horticulture
Vellanikkara, Thrissur.

EXTERNAL EXAMINER: Dr. E.V. Soniya
Scientist
Rajiv Gandhi Centre for Biotechnology
Poojappura, Thiruvananthapuram

ACKNOWLEDGEMENT

In connection with this venture, I sincerely place on record my profound sense of gratitude and indebtedness to numerous people who have helped me in this project. And so comes the time to look back on the path traversed during the endeavour and to remember the faces behind the action with a sense of gratitude. Nothing of significance can be accomplished without the acts of assistance, words of encouragement and gestures of helpfulness from others.

First and foremost I bow my head to the God who blessed me with the power and courage to complete this endeavour successfully in spite of the most difficult times face by me during the period of my study.

With deep respect I express my whole hearted gratitude and unforgettable indebtedness to Dr. D. Girija, Professor, CPBMB and the chairperson of my advisory committee for her inimitable guidance, valuable suggestions, constructive criticism, unceasing interest, constant encouragement and above all the extreme patience, understanding and whole hearted co-operation rendered throughout the course of the study. I consider myself being fortunate in having the privilege of being guided by her.

I express my sincere gratitude to Dr. P.A. Nazeem, Professor and Head, CPBMB and member of my advisory committee for her valuable suggestions and guidance rendered me for the completion of the research programme and preparation of the thesis.

I am grateful to Dr. R. Keshavachandran, Professor, CPBMB and member of my advisory committee for his constructive suggestions and ever willing help rendered throughout this venture.

I am deeply obliged to Dr. Sally K. Mathew, Professor and Head, Department of Plant pathology and member of my advisory committee for his unfailing support and enthusiasm, relevant suggestions and whole hearted cooperation throughout the period of investigation.

I also avail this opportunity to pay my sincere obligations and heartfelt thanks to Dr. P.A. Valsala, Dr. R. Sujatha and Dr. Abitha of CPBMB for their encouragement and kind help offered at different stages of the study.

I have no words to express my deep sense of gratitude to all my senior friends Sivaji, Sudhamayee, Sunil and Gajanan for their valuable assistance and guidance extended throughout the investigation.

Words cannot really express the true friendship that I relished from my friends **Lubina P.A** *and* **Somya P.P** *and my batchmates* **Sarveswar Sah, Manjunath Mogalai, Chavan Pradeep Uttam Rao, Rakesh kumar Choudhary and Achinthyakumar Dolui.** *for the heartfelt help, timely suggestions and back-up which gave me enough mental strength to get through all mind-numbing circumstances.*

I have no words to express my deep sense of gratitude to my junior friends **Renju, Resmi, Meera, Fahitha, Asok and Amar** *for their moral support and encouragement.*

I extend myriad of gratitude to my friends and research workers of CPBMB **Deepa, Firoz, Sunil, Anju and Nithin Roy and Nitha** *for their kind co-operation and guidance during entire period of my study.*

I wish to express my sincere thanks to all the non-teaching staff members and labourers of CPBMB and Department pf Agricultural microbiology for their whole hearted cooperation and timely assistance.

I am deeply indebted to my **Parents** *without their moral support, blessings and affection this would not have been a success. I, once again express my heartful thanks to all those who helped me in completing this venture.*

The award of fellowship by JNU is greatfully acknowledged.

Sahna Hamsa

TABLE OF CONTENTS

LIST OF TABLES

LIST OF PLATES

LIST OF APPENDICES

ABBREVIATION

A	Adenine
bp	Base pair
BLAST	Basic local alignment search tool
C	Cytosine
°C	degree Celsius
cm	Centimeter
CPBMB	Centre for Plant Biotechnology and Molecular Biology
DMSO	Dimethyl sulphoxide
DNA	Deoxyribo Nucleic Acid
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetra acetic acid
G	Guanine
g	Gram
IPTG	Isopropyl thio galactoside
KAU	Kerala Agricultural University
kb	kilo base
LB	Luria Bertani
M	Mole
mg	Milligram
min	minute
ml	Millilitre
mM	Millimole
µg	Microgram
µl	Microlitre
NCBI	National Centre for Biotechnology Information
ng	Nanogram

OD	Optical Density
ORF	Open reading frame
PCR	Polymerase chain reaction
pH	Hydrogen ion concentration
%	Percentage
RNA	Ribo Nucleic Acid
RNase	Ribonuclease
rpm	Rotations per minute
sec	Second
T	Thymine
TAE	Tris Acetate EDTA
TE	Tris EDTA
U	Unit
V	Volts
v/v	Volume by volume
w/v	Weight by volume
X-gal	5- bromo 4 chloro 3- indolyl β-D galactosidase

Introduction

1. INTRODUCTION

Pumpkin (*Cucurbita moschata*) is an important vegetable grown extensively in India throughout the year for its mature and immature fruits. The word pumpkin originated from the word 'pepon' which in Greek means large melon. Initially adapted from the French word 'pompon' and changed to pumpion by the British and later American colonists changed the word to pumpkin. The crop is believed to have originated in North America. However, evidences from history reveal that pumpkin seeds were found in Mexico from 7000 to 5500 BC. It belongs to the family *Cucurbitaceae* with chromosome number 42. The plant is monoecious and honey bees are the major pollinating agents.

According to FAO, total production of pumpkin is 15.6 million metric tons per year. India commands the largest acreage worldwide (29.9 per cent) in this crop, followed by China (18 per cent) and Ukraine (4.3 per cent). China has recorded maximum production and highest productivity followed by USA and South Africa. A variety of pumpkins are grown all over India and it is a popular vegetable of Kerala. It is grown mainly for its mature and immature fruits which are rich in carotene.

The crop was found to be infected with several virus diseases like pumpkin mosaic (PM), pumpkin yellow vein mosaic (PYVM), squash leaf curl mosaic, pumpkin enation mosaic and pumpkin mild mosaic (Ghosh and Mukhopadhyay, 1979). All these diseases were reported from time to time primarily on the basis of symptoms on the hosts, physical properties, transmission and host range. Among these diseases only a few have been extensively studied.

The cultivation of pumpkin suffered a serious setback during the recent years due to severe outbreak of mosaic diseases, particularly yellow vein mosaic and pumpkin mosaic (Latha, 1992). PYVM was reported in northern India during early 1940s (Vasudeva and Lal, 1943). Later the disease was reported throughout India and

is caused by a geminivirus, belonging to genus *Begomovirus*. The virus is transmitted by whitefly vector, *Bemisia tabaci* and not sap transmissible (Capoor and Ahmed, 1975; Jayashree *et al.,* 1999). The incidence of PYVMD and the associated yield loss was reported maximum during February to May, when the vector population is at its peak (Muniyappa *et al.*, 2003) and even up to 90 per cent yield loss was recorded in pumpkin (Singh *et al.,* 2009). Two species of geminiviruses causing PYVMD have been reported in India, one from northern India- Tomato leaf curl New Delhi virus-India (Maruthi *et al.,* 2007) and the other from southern India- Squash leaf curl China virus-India, both are bipartite. Recently, bipartite *Squash leaf curl Palampur virus* has also been reported to be associated with PYVMVD (Jaiswal *et al.,* 2011). Although attempts have been made to characterize the causal agent based on its biological characteristics, information on molecular biology of the causal organism is scanty and there were no reports from Kerala.

The symptom appears as faint clearing of tertiary veins which gradually spreads to secondary and primary veins. In the later stages the entire leaves may turn yellow, making it difficult to distinguish from other diseases, pest attack caused by sucking insects or nutrient deficiency symptoms. The virus also takes nearly one week to express the characteristic symptom (Muniyappa *et al.,* 2003) or might remain in the latent stage without producing any external symptoms. Altering planting dates and timely chemical protection against insect vectors has been the principle means of disease control. Also, other plants which serve as reservoir hosts have to be eradicated. Hence, early detection of the disease is important.

Detection of plant viruses is becoming more critical as globalization of trade, particularly in horticultural commodities is increasing. The potential effects of climate change have further aggravated the spread of viruses and their vectors. The epidemics caused by newly emerging or re-emerging begomoviruses are becoming frequent and appearing even in new regions, previously free from such diseases. Techniques for

geminivirus detection include biological indexing, electron microscopy, antibody-based detection, including ELISA, nucleic acid hybridizations like dot blot, squash blot and Southern blot techniques, polymerase chain reaction and other DNA polymerase-mediated assays and microarray detection.

Methods based on PCR are sensitive, fairly inexpensive and require minimal skill to perform. PCR based techniques are most sensitive than ELIZA tests and can detect pathogens present in lower concentration also. In addition to detection of virus, an additional advantage of the method is that the amplicon can be sequenced to provide further data about strain types (Webster *et al.,* 2004). In case of pumpkin, mixed infections are common; PCR could be used to detect pathogen in those conditions also.

Cultivation of resistant varieties is the most efficient method of control of disease, particularly those caused by viruses. None of the available pumpkin cultivars or varieties so far available or reported in the country possesses resistance to PYVMV. CM-214, a line from Nigeria was reported resistant to PYVMV and PM (Peter, 2007). The variety is highly erratic in bearing and the fruits are warty. Crop improvement for resistant varieties using conventional breeding methods is difficult. Hence, other approaches like pathogen derived resistance mechanisms like coat protein and movement protein mediated resistance have to be applied for developing stable resistant varieties for which knowledge about the viral genes are essential. Therefore the present study was undertaken on molecular characterization of geminivirus causing yellow vein mosaic in pumpkin and developing a PCR based diagnostic kit to detect the disease.

Review of literature

2. REVIEW OF LITERATURE

Pumpkin is an important vegetable grown extensively in India throughout the year for its mature and immature fruits. The crop is infected by several virus diseases like yellow vein mosaic, cucumber mosaic, bottle gourd mosaic and pumpkin mosaic. The cultivation of the crop suffered a serious set back during the last few years due to the serious outbreak of mosaic diseases particularly yellow vein mosaic and pumpkin mosaic (Latha, 1992). In recent years the yellow vein mosaic has become a major threat to the crop. It is caused by a geminivirus belonging to *Begomovirus* genus, which is characterised by its dicotyledonous hosts and whitefly (*Bemisia tabaci*) vector. Maruthi *et al.* (2007) reported pumpkin was infected with a monopartite begomovirus in Bangladesh, whereas bipartite *Squash leaf curl China virus* (SLCCNV), *Tomat leaf New Delhi virus* (ToLCNDV), and *Tomato leaf curl Palampur virus* (ToLCPV) was reported from India (Jaiswal *et al.*, 2011).

Geminiviruses form the second largest family of plant viruses. These viruses have caused devastating effects on crops. Until last decade, these epidemics were relatively localized. Discussions in the Third International Geminivirus Symposium revealed that geminivirus diseases have spread throughout the world (Boulton, 2003). Viruses belonging to Begomoviruses were identified to be the main culprits that infect a wide variety of crops including pulses, solanaceous crops, bhindi, bitter gourd and several weed plants. Pumpkin yellow vein mosaic (PYVM) is also caused by geminivirus belonging to the genus *Begomovirus*. Their propensity for genetic recombination and the presence of species nonspecific satellite molecules provide unlimited evolutionary possibilities for the emergence of new disease problems. This, coupled with the widespread occurrence of the polyphagous B biotype of the whitefly vector and the dissemination of infected planting material, has resulted in an expansion of virus host ranges in new geographical locations (Polston, 1999).

The study deals with molecular characterization of geminivirus causing yellow vein mosaic in pumpkin and developing a PCR based molecular diagnostic kit to detect PYVMV. PCR is a sensitive and reliable tool for the detection of plant viruses. Geminiviruses are small, single stranded DNA viruses and thus are well suited for detection by PCR. Several degenerate primers based on conserved sequences of whitefly transmitted geminiviruses have been designed for detection of these viruses (Rojas *et al.,* 1993; Briddon and Markham, 1993; Wyatt and Brown, 1996). These primers can be used to amplify previously un-characterised geminiviruses. On the other hand, PCR primers designed on the basis of non-conserved sequences can be used to specifically detect a particular virus strain of the same virus (Rybicki *et al.,* 1990, McGovern *et al.,* 1994).

PYVM causes more than 90 per cent loss depending upon the stage of the infection. Some control of the disease may be obtained by controlling the vectors, but it is time consuming, costly and environmentally undesirable. Thus, recent efforts have been directed at understanding the interaction between the vectors and viruses with the ultimate aim of preventing virus acquisition or transmission. Czosneck *et al.* (2002) reported the presence of receptors in the midgut of whitefly vector of *Tomato leaf curl virus*. Once these are identified, it is possible to block the virus transmission by developing appropriate transgenic plants. Reports on molecular characterization of geminivirus causing PYVM in pumpkin are very few. Therefore, the relevant literature on geminiviruses infecting plants other than pumpkin is also reviewed here.

2.1 Symptomatology of PYVMV in pumpkin

Capoor and Ahmed (1975) described a mosaic in pumpkin (*C. moschata*) and designated as pumpkin yellow vein mosaic (PYVM). It also affected *C. pepo* and *C. maxima.* It was neither seed borne nor sap transmissible but transmitted in nature by whitefly in a semi- persistent manner. Bhargava and Bhargava (1977) confirmed the

whitefly transmission of PYVMV and recorded 60 per cent incidence in *C. moschata*. The symptom of PYVMV was described by Jayashree in 1984. Initial symptoms were faint yellowing of finer veins, finally developing into characteristic vein yellowing. In advanced stages, large chlorotic areas appeared on leaves making the yellow network of veins inconspicuous. Affected plants were stunted and the leaf size was reduced. There were very few female flowers and the flowers dropped prematurely.

Muniyappa *et al.* (2003) reported that the faint vein clearing symptom developed first in the tertiary veins of younger leaves. The symptoms extended gradually to secondary and primary veins as prominent vein yellowing and later, coalesced to a yellow mosaic. As the disease symptoms progressed, vein yellowing developed on older leaves causing early senescence and the plant growth was retarded. PYVMV infected summer squash (*C. pepo*) and winter squash (*C. maxima)* showed leaf distortion and mosaic symptoms. Infection of bottle gourd (*L. siceraria*) was associated with a mild mosaic but there was no vein yellowing. Flowers produced on infected seedlings senesced prematurely. Fruit set was reduced and the fruits developed were shriveled, under-sized and unfit for marketing (Maruthi *et al.,* 2007; Singh *et al.,* 2009).

2.2 Geminiviruses and their origin

The term 'Geminivirus', from the latin word, 'gemini' meaning twins was proposed by Harrison *et al.* (1977) for the small viruses with paired particles having single stranded genomes. The unique twin particle morphology of the virions (Bock *et al.,* 1974) and their single stranded DNA genome (Goodman, 1977) created interest in this group. Depending upon the genome organization, biological properties, vector and host range, geminiviruses are currently divided into *Mastrevirus, Curtovirus, Topocuvirus* and *Begomovirus* (Ribicki *et al.,* 2000; Fauquet and Stanely, 2003). Most

of the *geminiviruses* are transmitted by whiteflies and belong to the genus *Begomovirus*. These have great economic importance and enormous diversity resulting from their widespread geographic distribution and host adaptation (Varma and Malathi, 2003).

Scientists have put forward several hypotheses regarding possible origin of geminiviruses. According to Koonin and Ilyina (1992), geminiviruses descend from bacterial plasmids. In the late 1990s sequences with relatively high similarity to replication associated proteins of geminivirus were found in some phytoplasmal plasmids. On the basis of these similarities among replication associated proteins and comparative homology based structural modeling of viral capsid proteins, Krupovic *et al.* (2009) proposed that geminiviruses have evolved from phytoplasmal extra chromosomal DNAs (EcDNA) by acquiring a capsid protein coding gene from a co-invading plant RNA virus. But recently, Saccardo *et al.* (2011) suggested that the EcDNA encoded Rep is not of phytoplasmal origin but has been acquired by phytoplasmas through horizontal transfer from a geminivirus or its ancestor and the evolution of geminivirus capsid protein in land plants showed missing links, while the sequence analysis using metagenomic data showed significant similarity to geminivirus.

Even though the evolutionary history of geminiviruses still remains unresolved, they are evolving successfully to invade new hosts, break host resistance and move virus particles within and between the plants. Humans have played an important role in the last century by way of carrying infected plants and planting material to new areas and thereby enhancing their evolution. The greatest molecular diversity of geminiviruses and their satellites resides in Southeast Asia revealing a possible center of origin (Nawaz-ul-Rehman and Fauquet, 2009).

2.2.1 Genetic makeup of Geminivirus

Geminiviruses possess single stranded, circular DNA genomes packaged within geminate particles. They may be monopartite with single DNA molecules (DNA A) or bipartite with two DNA (DNA A and DNA B) molecules. The monopartite viruses are found to be usually associated with a satellite DNA molecule. Based on the genomic organization and biological properties, the family *Geminiviridae* has been classified into four genera: *Mastrevirus, Topocuvirus, Curtovirus* and *Begomovirus* (Fauquet and Stanley, 2003).

Members of the genus *Mastrevirus* have a single genomic component encoding four proteins, infect mainly monocotyledonous plants and are transmitted by leafhoppers. *Topocuvirus* members have a single genomic component encoding six proteins, infect dicotyledonous plants and are transmitted by treehoppers. The *Curtovirus* members have a single genomic component encoding seven proteins, infect dicot plants and are transmitted by leafhoppers. Members of the genus *Begomovirus* have either one genomic component (monopartite) or two genomic components (bipartite) (Plate 1a and b). Six proteins are encoded by the single genome (DNA A) of monopartite viruses whereas; in bipartite viruses five-six proteins are encoded by DNA A and two proteins by DNA B. These viruses infect dicot plants and are transmitted by whiteflies (Fauquet and Stanley, 2003).

Most of the viruses causing yellow vein mosaic in crop pants belong to genus *Begomovirus.* They have characteristic geminate particles (18 nm×30 nm) encapsidating a single stranded circular DNA (Harrison and Robinson, 1999; Rojas *et al.*, 2005). DNA- A (2.5 kb) encodes proteins required for encapsidation and viral replication and DNA B (2.6 kb) for movement functions. DNA-A generally possesses 6 ORFs with 2 ORFs [AV1 (CP) and AV2] in the virion sense DNA and 4 ORFs (AC1

and AC4) in complementary sense DNA (Lazarowitz, 1992). DNA A encodes replication-associated protein (Rep), replication enhanced protein (REn), coat protein (CP), and transcriptional activator protein (TrAp). DNA-B encodes nuclear shuttle protein (NSP) and movement protein (MP), which are involved in virus movement within the host plant.

The DNA β, associated with monopartite viruses is circular, single stranded and approximately half the size of the genomic DNA-A (1300 bp). It has little sequence similarity to either DNA-A or DNA-B molecule of begomoviruses, except for a conserved hairpin structure and a TAATATTAC loop sequence (Xie *et al.*, 2002). It is essential for inducing typical symptoms in the hosts as it contains the functional gene, βC1 (Briddon *et al.*, 2003; Briddon and Stanley, 2006). This gene was found responsible for inducing yellow vein disease in *Ageratum, Eupatoriuim* and bhendi, leaf curl in cotton, yellow leaf curl disease in tomato and many other crops. (Briddon *et al.*, 2001; Jose and Usha, 2003; Kon *et al.*, 2006; 2007; Saunders *et al.*, 2003; 2004). Kon *et al.* (2007) reported βC1 gene was responsible for post transcriptional gene silencing (PTGS) suppression in the host, *N. benthamiana* when co-inoculated with *Tomato leaf curl virus.*

2.3 Functions of coat protein (CP) and movement protein (MP)

2.3.1 Coat protein

Coat protein represents the building block of viral particle and forms the shell of the geminate particles that are typical for begomoviruses. Geminivirus virions, because of their size, cannot enter by free diffusion through the nuclear pore complex. CP shuttles the viral genome into the host nucleus. This is based mainly on the

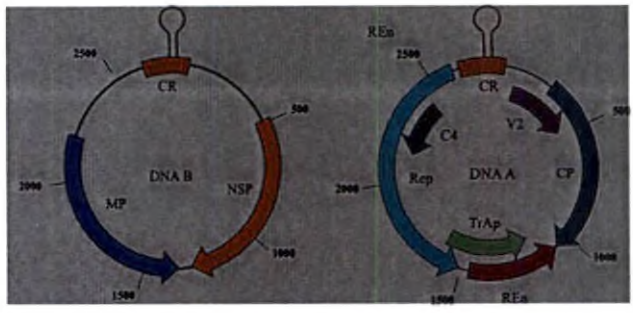
karyophilic nature of the CP (Kunik *et al.*, 1998) as well as its ssDNA-binding activity (Palanichelvam *et al.*, 1998).

Gafni and Epel (2002) reported that coat protein (CP) gene is absolutely essential for viral infection and movement. This is true for all monopartite but not for bipartite geminiviruses as they possess separate movement protein (MP) gene for performing movement functions within the plant.

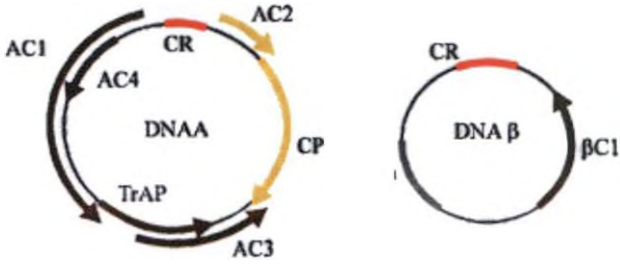
Coat protein plays the basis of serological detection and identification methods. CP shows a geography-related pattern of variation which was revealed by comparing epitope profiles, reactions with monoclonal antibodies and deduced amino acid sequences. Its properties include binding to and protecting viral ssDNA, self-binding, nuclear targeting, nuclear export and with monopartite-genome begomoviruses, mediating systemic movement in plants. CP is needed for transmission by the whitefly vector, *Bemisia tabaci*. It controls virus transport through the whitefly gut wall into the haemocoele, where it binds to GroEL analogue, produced by a bacterial endosymbiont, in a way that may protect virus particles from degradation. It is strongly conserved among different begomoviruses and serves as an outstanding example for multifunctional viral protein (Harrison, 2002).

2.3.2 Movement protein

Movement proteins are essential for systemic infection of the hosts. Begomoviruses are mostly phloem limited and encode two proteins required for virus movement, BR1 and BL1 (Ingham *et al.*, 1995). BR1 is ssDNA binding protein that localizes to nuclei of infected cells and functions as a nuclear shuttle protein (Sanderfoot *et al.*, 1996a). BL1 is directly responsible for viral pathogenic properties and transient expression assays in tobacco protoplasts revealed that BL1 provides



a. **Bipartite genome organization**



b. **Monopartite genome organization**

c. Bemisia tabaci

Plate. 1 Organisation of begomovirus genome

directionality to viral movement by cooperatively interacting with BR1 to redirect it from the nucleus to the cell periphery (Sanderfoot *et al.*, 1996b).

Qin *et al.* (1998) reported that the movement of bipartite geminiviruses such as *Squash leaf curl virus* (SLCV) requires the cooperative interaction of two essential virus-encoded movement proteins, BR1 and BL1 for systemic infection.

2.4 Host range of Geminivirus

Begomovirus contains more than 230 species (Fauket *et al.*, 2008). As a group they have wide host range, but infect only dicotyledonous plants. Worldwide they are responsible for large economic loss to many economically important crops including bean, cassava, cotton, pepper, potato, cucurbits, tobacco, tomato, etc. Their immense ability for genetic recombination and thereby development of new strains helped in extending the host range. Also, global trade in agricultural products encouraged the dissemination of their insect vectors. In recent years geminiviruses have been reported in many plants including ornamentals in which there were no previous reports of infection. (Fauket *et al.*, 2003). Affected plants exhibited a wide range of symptoms such as vein clearing, leaf curling, stunded growth and poor yield (Briddon, 2003; Shih *et al.*, 2009; Kumar *et al.*, 2010).

Cucurbits are highly susceptible to geminivirus attack. They are mainly affected by *Squash leaf curl virus* (SLCV), *Squash leaf curl China virus* (SLCCV), *Watermelon curly mottle virus* (WCMoV) and *Tomato leaf curl New Delhi virus* (ToLCNV). The original strain of the *Squash leaf curl virus* (SLCV-CA) was reported in California (USA) in *Cucurbita pepo* (marrows). WCMoV reported in Arizona (USA) has a wider host range, attacking cucumbers (*Cucumis sativus*), melons (*Cucumis melo*) and watermelons (*Citrullus lanatus*). Other hosts include *Cucurbita*

foetidissima, *C. maxima*, *C. moschata* and wild *Cucurbita* spp. *Phaseolus vulgaris* is the only non-cucurbitaceous host (McCreight and Kishaba, 1991). Liao *et al.* (2006) reported *Squash leaf curl virus* infecting cucurbits in Vietnam.

ToLCNDV has been reported to infect many cucurbits such as cantaloupe (Samretwanich *et al.* 2000a), cucumber (Samretwanich *et al.* 2000b), wax gourd, muskmelon (Samretwanich *et al.* 2000c), watermelon (Mansoor *et al.* 2000) and bitter gourd (Tahir and Haider, 2005). ToLCNDV was also reported in tomato (Padidam *et al.*, 1995), chilli pepper (Hussain *et al.* 2004), potato (Usharani *et al.* 2004). Ito *et al.* (2008) reported bipartite TLCNV-[Cuc:Thai] causing yellow leaf curl disease in cucumber, bottle gourd and muskmelon.

Muniyappa *et al.* (2005) reported that Pumpkin yellow vein mosaic virus (PYVMV) transmission to non-cucurbitaceous crops was limited. Among forty three plants belonging to 16 families tested for their susceptibility to PYVMV, only four of 11 cucurbit species were found to be virus hosts. All inoculated pumpkin, winter squash, summer squash and all but one bottle gourd plants developed disease symptoms.

Whitefly transmitted geminiviruses (WTGs) was reported in several weeds which are found to be reservoir hosts of these viruses. Konate *et al.* (1995) reported *Sida acuta* as a wild host of okra yellow leaf curl virus. Colvin *et al.* (2003) conducted study on sixty four weed species that occured commonly in and around ToLCV infected fields during three seasons. Among 39 weed species infested with *B. tabaci* during different seasons, the weed, *Euphorbia geniculata,* was found to be one of the important reservoir hosts.

2.4.1 Vector Transmission

Geminiviruses depend on their insect vectors for their transmission but there are some viruses which are reported to be vegetatively transmitted through infected stocks, such as geminiviruses associated with cassava mosaic disease. One reason for the appearance of epidemics is the spread of the vectors of geminiviruses (Rybicki and Pietersen, 1999). Geminiviruses have evolved a highly dependent relationship between their host plant and their vector. There are many species of sucking insects like whiteflies, which can acquire viruses but a few are actually capable to transmit to the next host (Liu *et al.*, 1997). There are a few species of leafhoppers (12-18) and one species of treehopper which can transmit geminiviruses. It is actually the coat protein which determines the specificity of geminivirus transmission by an insect vector.

Begomoviruses are transmitted by the whitefly, *Bemisia tabaci* (Gennadius) (Plate 1c). *B. tabaci* was first described in 1889 (Gennadius, 1889). All begomoviruses are transmitted by *B. tabaci* and the transmission efficiency is different for different viruses. Over 100 begomoviruses are transmitted by at least two biotypes of *B. tabaci* to more than 20 cultivated species of socioeconomic importance. Begomovirus transmission by *B. tabaci* is circulative and non-propagative (Markham *et al.*, 1994; Gray and Gildow, 2003). As the feeding time increases, the amount of virus acquired by the whitefly increases.

Hunter *et al.* (1998) proposed a model for the movement of begomoviruses in the whitefly vector. Virus particles are ingested along with plant fluids into the whitefly oesophagus and foregut. As food enters the filter chamber, excess water is shunted to the ileum of the hindgut. Thus nutrients and begomoviruses are concentrated in the filter chamber. Begomovirus particles adsorb to specific sites on the alimentary membrane or to sites along the anterior region of the midgut. Begomovirus particles move out of these tissues into the hemolymph, eventually

invading the salivary glands. There is no evidence for viral replication in the insect vectors (Czosnek, 2001).

Muniyappa and Reddy (1976) reported that whiteflies required a minimum of 30 minutes acquisition access period (AAP) to transmit Horse gram yellow vein mosaic virus. Ioannou (1985) reported that minimum acquisition and inoculation feeding periods for the vector, *Bemisia tabaci* to transmit yellow leaf curl virus was 20 and 30 minutes, respectively in tomato. The latent period was 21-24h and the vector was infective for about 12 days, but there was no transovarial transmission. After acquisition by whiteflies, begomoviruses are persistent and are retained for periods ranging from a few weeks to life (Duffus, 1987).

An acquisition access period of 10 minutes and minimum inoculation access period of 20-30 minutes by *B. tabaci* was sufficient for successful transmission of Indian Cassava mosaic virus, ToLCV and ToLCV-Ban4 (Mathew and Muniyappa, 1991; Mansour and Al-musa, 1992; Ramappa, 1993; Muniyappa *et al.*, 2000).

In case of pumpkin yellow vein mosaic virus, Muniyappa *et al.* (2005) reported that a minimum period of 30 min for whiteflies to acquire PYVMV for per cent transmission and an AAP of atleast 6 h was required for 100 per cent transmission efficiency. IAP of 10 min and 6 h resulted in 1.0 per cent and 100 per cent transmission respectively. It also required a minimum latent period of 6 h for transmitting the virus.

2.5 Detection of geminiviruses

Detection deals with establishing the presence of a particular target organism in a sample, with special emphasis on symptomless individuals. Recombinant DNA

technology has significantly contributed to our understanding of plant viruses. Several techniques are available for plant virus diagnosis and detection, each with its own advantages and disadvantages (Webster *et al.,* 2004).

It was necessary to perform time consuming indexing for virus detection and this process was obviously not suited to routine analysis of a large number of samples. Till recently, the detection and diagnostic techniques available for viruses were biological indexing, electrophoresis, electron microscopy and serological techniques based on precipitation. These methods for detection are either not sensitive enough (immunological tests), too slow (biological tests on indicator plants), or difficult to standardize (immunofluorescence). This prompted for the search of alternative diagnostic techniques (Lopez *et al.,* 2003). Compared to other methods, PCR based diagnosis permits the detection of viruses even at low levels (Fenby *et al.,* 1995).

2.5.1 Serological detection

Serological approach was an important turning point in virology because the pathogen could be identified and detected much more rapidly than biological indexing and culturing methods (Miller and Martin, 1987). Serological techniques based on precipitation are still in use for detection and diagnosis of pathogens.

2.5.1.1 Enzyme Linked Immunosorbent Assay (ELISA)

Detection using ELISA based on antigen antibody reactions are widely used method for practical plant virus detection. Many commercial plant virus detection kits have been developed based on this. It requires labeling of every antibody to be used which is time consuming and expensive. The method also requires polyclonal or monoclonal specific antisera for each virus of interest that do not cross react with plant

proteins. ELISA has being used for detection of begomoviruses like SLCV, *Cassava latent virus*, PYVMV etc.

Sequeira and Harrison (1982) used ELISA to detect *Cassava latent virus* (CLV) in *Nicotiana benthamiana* leaves and naturally infected cassava from Angola. Mathew (1989) successfully used ELISA to detect ICMV in crude sap of infected cassava, eara rubber, *N. benthamiana, N. tobaccum* cv. Jayasri, using ICMV-H antiserum.

Muniyappa *et al.* (1991) used different monoclonal antibodies of ACMV and ICMV to detect and characterization of tomato leaf curl geminivirus. *Cotton leaf curl virus* was detected in cotton cultivars and weed hosts by using monoclonal ELISA (Nateshan, 1992).

Swanson *et al.* (1998) used a panel of 25 monoclonal antibodies (MAbs) raised against particles of two heterologous whitefly-transmitted geminiviruses (begomoviruses) in triple antibody-sandwich ELISA (TAS-ELISA) to determine the detectability and epitope profiles of 26 Indian isolates of tobacco leaf curl virus (TLCV) and 13 of *Croton yellow vein mosaic virus* (CYVMV). Virus isolates serologically indistinguishable from TLCV were detected in symptom-bearing weeds (*Acanthospermum hispidum, Ageratum conyzoides, Euphorbia geniculata* and *Parthenium hysterophorus*) found in leaf curl-affected tobacco fields. They found that TAS-ELISA with MAb SCR 18 was a more sensitive test for detecting Indian TLCV isolates than with double antibody-sandwich ELISA with polyclonal antibodies.

Begomovirus was detected by Abouzid *et al.* (2002) using polyclonal rabbit antisera to the coat protein of bean golden mosaic virus Brazil isolate (BGMV), *Cabbage leaf curl virus* (CabLCV), *Tomato yellow leaf curl virus* (TYLCV), and *Tomato mottle virus* (ToMoV).

PYVMV was detected in pumpkin by indirect ELISA using ICMV polyclonal antibody (Jayashree, 1996). Muniyappa *et al.* (2003) serologically detected PYVMV in diseased pumpkin plants and viruliferous *B. tabaci* using polyclonal antibodies in a double antibody sandwich ELISA. Reactions with monoclonal antibodies in triple antibody ELISA showed that PYVMV has an epitope profile distinct from those of other begomoviruses from the Indian subcontinent.

Serological tests using DAC-ELISA with 10^{-1} dilution of Bitter gourd distortion mosaic virus (BGDMV) showed that the virus was closely related to *Squash leaf curl China virus.* Positive reaction was also observed with monoclonal antibody of TLCV (Zachariya, 2007).

Bitter gourd yellow vein mosaic virus (BGYMV) has been detected by ACMV polyclonal antibody and ICMV monoclonal antibody. Optimum antigen and ICMV monoclonal antibody dilution to detect BGYMV was 1:10 and 1:1 respectively (Rajinimala *et al.,* 2009).

2.5.2 Electron microscopic (EM) detection:

Most of the viruses, because of their extremely small size are well below the resolving power of light microscope. Therefore electron microscopy is the only way to visualize the virus particles. Immunosorbent electron microscopy (ISEM) is widely used for this purpose and results accumulated have made it possible to classify viruses based on their distinct morphology and symmetry.

Immunosorbent electron microscopy (ISEM) using antibody coated carbon filmed grids, was employed to detect several whitefly transmitted geminiviruses in the purified preparations as well as in the crude sap. Sequeira and Harrison (1982) observed Cassava latent virus (CLV) particles trapped in CLV antibody coated grids.

Roberts *et al.* (1984) used homologous and heterologous antibody coated carbon filmed grids in ISEM and found that strong relationship existed among five whitefly transmitted gemini viruses viz., ACMV, BGMV, EVMV (Euphorbia mosiac virus), SLCV and TomGMV.

Mathew (1989) used ISEM successfully to detect ICMV particles in the leaf extracts of *N. benthamiana* and cassava infected with ICMV-H and ICMV-14 using ICMV-H antibody coated grids. He also found that ICMV-14 antiserum could also detected HyMV in infected french bean and ToBLCV in infected *N. tobaccum.*

Mandal (1989) successfully detected CYVMV particles in the leaf extracts of croton infected with CYVMV using Indian cassava mosaic virus antibody coated grids. Harrison *et al.* (1991) also used ISEM technique to detect seven different whitefly transmitted geminiviruses from India.

Tomato leaf curl virus was detected using ISEM technique by Muniayppa *et al.* (1991). Typical geminate particles were detected in the leaf extract of cotton leaf curl infected cotton by using squash leaf curl virus antiserum (Nateshan, 1992).

Potato plants showing apical leaf curl and pinkish pigmentation were examined for the presence of geminivirus with electron microscope. Clarified virus concentrate showed presence of a high concentration of geminate virus particles ca. 28×17 nm (Garg *et al.,* 2001).

Electron microscopic observation of purified preparation of BGDMV showed geminate particles of 18-20 x 30 nm along with flexuous rod shaped particles of 750 nm resembling poty virus,indicating mixed infection of gemini and poty viruses (Zachariya, 2007).

2.5.3 Molecular detection and Characterisation of Geminivirus

Technologies for molecular detection of plant pathogens have undergone two major breakthroughs. The first was the advent of antibody based detection, in particular monoclonal antibodies and ELISA. Then came DNA based technologies, such as polymerase chain reaction (PCR) which made rapid detection possible even the pathogen copy number is low as it can amplify the DNA several million fold (Mullis and Faloona, 1987). PCR technology eventually replaced the immunological assays. PCR based cloning and sequencing has made molecular characterization an easy task.

2.5.3.1 Isolation of DNA

Isolation of good quality DNA is a pre-requisite for PCR amplification. Depending upon the plants, appropriate DNA isolation protocol has to be developed. Care should be taken to remove all the contaminating agents like proteins, polyphenols and RNA. These components can hinder the PCR reaction by inhibiting the Taq polymerase activity. Scientists have developed different protocols to isolate DNA from infected plants. CTAB method is the most commonly used method for nucleic acid extraction from geminivirus infected samples.

Total DNA was isolated by Mansoor *et al* (1998) from young symptomatic leaves of cotton infected with cotton leaf curl disease by modified CTAB method described by Doyle and Doyle (1987).

Jose and Usha (2000) described a protocol for the extraction of geminiviral DNA from *Bhendi yellow vein mosaic virus* infected okra which has high amounts of mucilage and other phenolic compounds. They extracted DNA with a buffer

containing sodium citrate at pH 6 followed by PEG precipitation of the virus and alkali lysis. The extraction buffer eliminated the mucilage and other polyphenols, PEG precipitated the viral particles and DNA and the alkali lysis enriches the replicative forms of the viral DNA. This method was superior to three other common plant DNA extraction protocols the DNA was used for PCR amplification of coat protein gene.

Li *et al.* (2004) extracted total nucleic acid from symptomatic tomato leaves showing leaf curl following the method of Xie *et al.* (2002). They used these DNA samples to carry out PCR reactions using degenerate primers for amplifying portions of DNA A and DNA β.

Ito *et al.* (2008) extracted total nucleic acids from the leaf tissues of infected cucumber plants, as described by Kon *et al.* (2003).

Rouhibakhsh *et al.* (2008) extracted DNA from symptomatic legumes with modified CTAB method. The buffer contained mercaptoethanol (5%) and sodium chloride (1.4 to 2.0 M). The protocol is useful for the detection and identification of begomoviruses infecting grain legumes. CTAB method was also reported in bitter gourd yellow vein mosaic virus (BGYMV) infected bitter gourd (Rajinimala *et al.*, 2009; Tariq *et al.*, 2010)

Total DNA was extracted from young pumpkin plants showing typical symptoms of vein yellowing, mosaic and crumpling symptoms on leaves, mosaic and distortion and reduction of size of fruits, using the method of Pich and Schubert with some modifications (Jaiswal *et al.*, 2010). Watcharawongpaiboon and Chungwongse (2007) extracted genomic DNA from pumpkin using the micro preparation procedure described by Fulton *et al.* (1995)

2.5.3.2 PCR based detection

PCR based tools play a vital role in diagnosis, detection and identification of viruses in plants. Traditional virus diagnosis required bioassay, determination of host range, symptomatology, virus particle morphology and vector relations. PCR made this whole process simple, rapid and highly reproducible. Several universal primers are available for geminivirus detection. For specific detection, mainly primers based on conserved regions of coat protein and movement proteins genes (Bela-ong and Bajet, 2005).

Navot *et al.* (1992) amplified the genomic DNA molecule of an Israel; isolate of TYLCV from DNA extracts of TYLCV infected tomato M-82 by polymerase chain reaction (PCR) using synthetic oligonucleotides complimentary to different regions of the viral genome as primers.

Nakhla *et al.* (1993) employed PCR technique for the identification and characterization of four tomato yellow leaf curl virus isolates of Egypt using specific primers of TYLCV-ICR (Israel isolate).

Rojas *et al.* (1993) used PCR for detection and differentiation of whitefly transmitted geminivirus using degenerate primers. DNA-A and DNA-B fragments of 15 previously uncharacterized geminiviruses were amplified. Deng *et al.* (1994) designed primers to amplify approximately 500 bp fragment of DNA-A of five well characterized whitefly transmitted geminiviruses. Deng's degenerate primer was used to amplify 560 bp fragment from BGYMV infected bitter gourd (Rajinimala *et al.*, 2009).

Several degenerate and specific primers were designed to detect TYLCV. McGlashan *et al.* (1994) amplified viral degenerate primers specific for genome A

(PASIV 1978 and PARIC 496) and genome B (PBCIv 2040 and PCRC1). Nakhla and Maxwell (1994) amplified the full length 2.8 kb viral DNA from yellow leaf curl infected tomato plant using primers PTYC IV 2182 and PTYC 1C 2140. Ramos *et al.* (1996) carried out PCR amplification with tomato plants infected with TYLCV and full nucleotide sequence of the virus was determined.

Khan (2000) detected ToLCV both in its host (tomato) and vector *B. tabaci* by employing geminivirus specific degenerated primers using PCR. Muniyappa *et al.* (2000) used polymerase chain reaction for detection of ToLCV using degenerate primers both in infected tomato plants and viruliferous *B. tabaci*. Morris *et al.* (2002) reported that PCR was found to be a more sensitive method than triple-antibody sandwich enzyme-linked immunosorbent assay (TAS ELISA) for the detection of TYLCV isolates in all hosts.

Muniyappa *et al.* (2003) detected PYVMV using CP and MP specific primers. CP primers amplified from the nucleotide positions 300 (the 5' end of CP) to 1100 bases (~70 bases downstream of the 3' end of CP). MP primers amplified a ~570bp PCR product within the movement protein on DNA-B. Singh *et al.* (2009) also detected geminivirus in PYVM infected pumpkin leaf samples. Primers, PYABFP and PYABRP were used to amplify full length DNA A and primers PYBHFP and PYBHRP for amplifying full length DNA B.

Geminivirus, causing yellow leaf curl disease of wax gourd plants, was detected using polymerase chain reaction with geminivirus-specific primers. (Sawangjith, 2009).

Maruthi *et al.* (2007) used universal and species-specific primers to detect begomoviruses in seven crops (chilli, okra, papaya, pumpkin, sponge gourd, tomato and yardlong bean) and two common weeds (*Ageratum conyzoides* and *Croton*

bonplandianum). At least five distinct species of tomato leaf curl viruses infected tomato and other host-plants. Manickam *et al.* (2001) detected the gemini virus present in weed species like *Croton sparsiflorus, Malvastrum coromendalianum* by PCR.

Sinha *et al.* (2011) detected *Pepper leaf curl virus* (PepLCV) in chilli by PCR with coat protein gene specific primer pairs. Amplification product of size 1kb specific to coat protein gene was amplified.

2.6 Cloning and sequencing

The term cloning refers to isolation and amplification of an individual gene sequence by insertion of that sequence into a suitable vector (e.g. plasmid, phage, cosmid) where it can be replicated. The precise order of the nucleotide sequence of the insert can be determined by DNA sequencing. Usually, geminivirus genomes are cloned by either directly restricting the replicative form which is isolated electrophoretically from the total infected nucleic acid (Hamilton *et al.*, 1982) or by amplifying the full or partial genomes with geminivirus-specific primers (Wu *et al.*, 1996).

The complete genome of a Thailand isolate of the geminivirus, *Tomato yellow leaf curl virus* (TYLCV-Th) was cloned and sequenced. The results indicated that genome consists of two DNAs each slightly greater than 2700 nucleotides in length and designated as A-DNA and B-DNA. (Rochester *et al.*, 1994).

Harrison (1995) cloned and sequenced the coat protein (CP) genes of an isolate of tomato leaf curl virus (ITmLCV; 2749 nt) obtained from southern India, Nigeria and two regions of Saudi Arabia.

Padidam *et al.* (1995) sequenced the genomes of two isolates of tomato leaf curl gemini virus from India (ToLCV-India). It contained DNA A and B components. Muniyappa *et al.* (2000) used degenerate DNA-A-specific primers to amplify and clone the genome of a ToLCV isolate (named ToLCV-Ban4) from field-infected tomato plants growing in Bangalore. The full-length of 2759 nucleotide long DNA-A like viral genome was sequenced (Muniyappa *et al.*, 2000).

The genomic DNA of tobacco leaf curl geminivirus (TLCV) from tomato plants with leaf curl disease in Japan has been cloned and sequenced. The single circular DNA molecule comprises 2,761 nucleotides (Shimizu and Ikegami, 1999).

Yin *et al.* (2001) sequenced the complete DNA sequence (2734 nucleotides) of the monopartite genome of *Tomato yellow leaf curl China virus* (TYLCCNV). Sequence comparisons with other geminiviruses showed that TYLCCNV belongs to Begomovirus from the Old World.

The complete nucleotide sequence DNA A and B of squash leaf curl disease in the Philippines was determined. DNA-A and DNA-B comprise 2739 and 2705 nucleotides, respectively; the common region is 174 bases in length. Five ORFs were found in DNA-A and two in DNA-B (Kon *et al.*, 2003).

Praveen *et al.* (2004) cloned and sequenced replicase gene of three isolates of whitefly transmitted geminivirus causing leaf curl disease of tomato in India. Nucleotide sequence and the derived amino acid sequence for the replicase gene of these isolates was determined and analyzed with the published sequences.

Liao *et al.* (2007) cloned PCR products from wax gourd plant showing yellowing and enation DNA sequences of the correct clones were confirmed by an automatic DNA sequencing system (ABI 3730 XL DNA Analyzer, Applied

Biosystems, CA). The CP gene was further cloned into the vector pET21d(+), which was then used to express proteins in *Escherichia coli* BL21(DE3) and to produce specific antiserum for future diagnosis purposes. PCR products of geminivirus-specific primers from yellow leaf curl infected wax gourd were cloned using pJet 1.2/blunt cloning vector (Fermentas). The ligation mixture was used to transform *E.coli* DH5-competent cells by heat shock transformation. The complete genomic sequence was determined by sequencing overlapping DNA fragments.

PCR products of size ~530 bp obtained from the intergenic region and capsid protein of DNA-A from tomato, chilli, dolichos, *Ageratum*, *Croton*, okra, papaya, pumpkin and yardlong bean using the Deng A/B primers were cloned into the pGEMT-Easy vector (Promega, UK). Two independent clones were sequenced (Maruthi *et al.*, 2007).

2.7 Homology and Phylogenetic analysis

During recent years there has been a tremendous accumulation of nucleic acid and protein data of viruses that allow for establishing more accurate relationship among individual members. These sequences can be readily used for establishing homologies and constructing phylogenetic trees. Homology and phylogenetic analysis of coat protein, DNA A and B and their intergenic region are widely used in classification and characterization of geminiviruses.

Kon *et al.* (2003) reported that the causative agent of squash leaf curl disease of pumpkin in the Philippines is an isolate *Squash leaf curl China virus* based on the sequence homology of DNA-A (88 per cent) and deduced coat protein amino (98 per cent). Muniyappa *et al.* (2003) reported that, on sequence and phylogenetic analysis of the virus causing pumpkin yellow vein mosaic showed similarity to the bipartite *Tomato leaf curl New Delhi virus.*

Malathi *et al.* (2007) reported PYVMD in North and South India were caused by two distinct begomoviruses and shared only approximately 88 per cent DNA-A nucleotide identity. The South Indian isolate was most closely related to *Squash leaf curl China virus* between 91 and 96 per cent identities, and the two North Indian isolates to *Tomato leaf curl New Delhi virus* between 94 and 96 per cent identities.

Comparison of DNA-A and DNA-B components along with the intergenic region of Squash leaf curl China virus from Varanasi, India (SLCCNVIN[IN:Var:Pum]) showed highest sequence identity with the SLCCNV-IN[IN:Luc:Pum] and SLCCNV-IN [IN:Coi:Pum]. Phylogenetic analysis also shows close relationship to other cucurbit infecting geminiviruses (Singh *et al.*, 2009).

Ito *et al.* (2008) reported *Tomato leaf curl New Delhi virus* infecting cucumber, bottle gourd and muskmelon in Thailand DNA A showed 97 per cent nucleotide sequence identity with ToLCNDV-[Luffa] and 95 per cent with ToLCNDV-Mild.

Phylogenetic relationship clearly indicated two subsets for ToLCV, one belonging to tomato leaf curl virus (ToLCV) having bipartite genome and the other having the monopartite genome (Rojas *et al.* 2005). .

Guo *et al.* (2008) reported that *Tomato yellow leaf curl Thailand virus,* TYLCTHV-[Y72] is a monopartite begomovirus with DNAβ which may represent an evolutionary intermediate between the begomoviruses requiring DNAβ and begomoviruses dispensable of DNAβ.

Kon *et al.* (2009) reported the sequence of the okra leaf curl disease associated betasatellite was most identical (87–99 per cent) to isolates of cotton leaf curl Gezira

betasatellite (CLCuGB) and had near identity (99.2 per cent) with betasatellite from tomato in Mali (CLCuGB-[Mali:Tomato:2005].

2.8 Geminiviruses in weed plants

Weeds have been reported as important reservoirs for some geminiviruses (Roye *et al.* 1997; Rojas *et al* 2000; Sanz *et al.* 2000). Begomoviruses have been reported to infect a variety of weeds and most of them were monopartite in nature (Brown and Bird 1992; Frischmuth *et al.* 1997; Roye *et al.* 1997; Umaharan *et al.* 1998).

Tomato mottle virus-Flo (ToMoV-[Flo]) was found to infect the tropical soda apple (*Solanum varium*) in Florida (McGovern *et al.* 1994). *Potato yellow mosaic Trinidad virus*-[Trinidad and Tobago] (PYMTV-TT) was found in five weed hosts (Umaharan *et al.* 1998). *Sida golden mosaic Honduras virus* (SiGMHV) was found in tomato growing areas of Nicaragua infecting tomato for which *Sida* spp. serve as natural hosts (Rojas *et al.,* 2000).

Saunders *et al.* (2000) reported monopartite begomoviruses such as *Ageratum yellow vein virus* (AYVV), *cotton leaf curl Multan virus* (CLCuMV), and *Malvastrum yellow vein virus* (MYVV) were associated with satellite DNA molecules in docot weeds. Zhou *et al.* (2003) detected begomovirus in weed plant, *Malvastrum coromandelianum* showing yellow vein symptom degenerate primer pair PA and PB amplified partial intergenic region and CP gene of DNA-A.

Most of the viruses reported from weeds are monopartite and associated with DNA β which is essential for inducing the characteristic disease symptoms as it encodes an essential pathogenicity protein. Ageratum yellow vein virus (AYVV) from Singapore, Eupatorium yellow vein virus (EpYVV), tomato yellow leaf curl Java virus

(ToLCJAV), and Ageratum yellow vein virus from Indonesia (AYVV-[Java]) were associated with DNA β (Tao *et al.,* 2004).

Guo and Zhou (2006) reported two virus isolates from *Sida cordifolia* showing mild upward leaf-curling symptoms in Hainan province of China. Complete sequence comparison with other reported begomoviruses revealed highest sequence identity (71.0 per cent) with that *of Tobacco leaf curl Yunnan virus.* The isolate was named as *Sida leaf curl virus* (SiLCV). Two additional circular single-stranded satellite DNA molecules corresponding to DNAβ and DNA A were found associated with SiLCV isolates which showed similarity to *Ageratum leaf curl virus* (AJ316027).

2.9 Developing virus resistant varieties

Diseases caused by plant viruses are not easy to manage and their control mainly involves the use of insecticides to destroy insect vectors, the use of virus-free propagating materials and selection of plants with appropriate resistance genes. Eventhough insecticides can control vectors, often the virus has already been transmitted to the plant before the insect vector is killed. The use of resistant cultivars is the most efficient means of control, however plant virus resistance genes are commonly unavailable and their introgression into some crops is not straight forward (Sharma *et al.,* 2000).

Genetic engineering has emerged as a "million dollar technology" to fight against viruses in comparison to conventional breeding. Various strategies have been engaged to develop genetically engineered resistance against plant viruses (Prins *et al.,* 2008). Among them, there are pathogen derived resistance strategies and non-pathogen derived resistance strategies. Both of these approaches can be used as a promising remedy against a number of viruses (Sanford and Johnson, 1985).

In pathogen derived resistance, certain key gene products of a pathogen present in the plant in a non-functional form, in excess, or at an inappropriate stage during viral replication cycle disrupt the infection. Pathogen-derived resistance is mediated either by the protein encoded by a transgene (protein-mediated) or by the transcript produced from the transgene (RNA-mediated) (Smith *et al.*, 1994).

2.9.1 Protein mediated pathogen derived resistance

This type of pathogen-derived resistance employs the use of viral genes through the expression of their protein product. Coat protein gene (Beachy, 1990), sequences from the viral replicase gene (Nguyen *et al.*, 1996), Rep protein gene (Noris *et al.*, 1996; Brunetti *et al.*, 1997) viral movement protein gene (Baulcombe, 1996) have been used as a source of pathogen derived resistance in various plant species.

Coat protein mediated resistance

Among the different strategies for controlling viruses by genetic engineering, the coat protein mediated strategy is the most promising. Classical cross protection studies suggested that the expression of the viral coat protein in the host would ultimately interfere with the uncoating of the virus in such a way that the virus genome would be less available for initiation of virus replication (Hamilton, 1980). Data supporting this assumption was obtained in protoplasts of tobacco transformed with TMV coat protein (Osbourn et al., 1989). Initially, coat protein mediated- resistance was thought to be effective not only against the virus from which the coat protein sequence is derived but also against distantly related viruses. Furthermore, it was suggested that the resistance would be characterized by a positive correlation between the level of protein expression and resistance. Resistant plants would exhibit high

levels of expression of the transgene protein (Beachy, 1994). There are many examples of resistance responses associated with transformation with coat protein genes. These responses can be grouped into two categories depending upon the characteristics of the construct (Stam *et al.*, 1997). With translatable coat protein transgenes, resistance is often expressed as a reduction in symptom severity, a reduction of the rate of systemic disease and a reduction of virus accumulation (Beachy, 1994; Loesch-Fries *et al.*, 1987; Van Dun *et al.*, 1988; Powell-Abel *et al.*, 1986). For non-translatable coat protein genes, the resistance can be characterized as immunity to infection, even in the presence of high doses of inoculum (Lindbo and Dougherty, 1992; Smith *et al.*, 1994).

Kunik *et al.* (1998) reported coat protein mediated resistance in transgenic tomato plants showing over expression of *nv1* gene producing high levels of the TYLCV CP. Mutations in the CP gene of Tomato yellow leaf curl Sardinia virus (TYLCSV) have been shown to interfere with systemic infection, as well as particle formation and insect transmission (Noris *et al.*, 1998). Hallan and Gafni (2001) suggested that these mutations prevent homotypic interaction of the TYLCV CP, and therefore viral particles fail to form and insect transmission is prevented. Two possible risks associated with viral coat-protein expressing plants are heterologous encapsidation and recombination. Recombination between mutant viral genome and coat protein transgene can result in a permanent change in genetic constitution of the virus progeny resulting in infectious recombinant virus.

Movement proteins have been used to confer resistance to geminivirus infection. Tobacco plants were transformed with the sense and antisense sequences of the movement genes (*BV1* and *BC1*) of ToMoV (Duan *et al.*, 1997). Transgenic plants that expressed a spontaneously mutated form of the BC1 gene showed resistance to ToMoV and cabbage leaf curl virus. The resistant phenotypes observed included a

delay in symptom development, recovery from mild early symptoms and complete absence of symptoms and virus replication.

Brunetti *et al.* (1997) showed that tomato plants (*Lycopersicon esculentum* cv. Moneymaker) transformed with a truncated version of the replication-associated protein gene (*AC1*) of TYLCV were resistant to TYLCV infection. Plants transformed with the transgene in the sense direction accumulated the truncated Rep protein and were resistant to TYLCV. Accumulation of the Rep protein was required to confer resistance to TYLCV. Tobacco plants transformed with the *Rep* gene also have shown resistance to geminivirus infection. Noris *et al.* (1996) demonstrated that the expression in tobacco of a Rep protein of TYLCV-Sr lacking the carboxy-terminus conferred resistance to TYLCV-Sr infection.

2.9.2 RNA silencing

RNA silencing is a homology-dependent mechanism of RNA degradation. It is induced by double stranded RNA (dsRNA) and leads to sequence specific gene silencing at the post-transcriptional level (Hannon, 2002). Recent reports indicate that geminivirus infected plants can recover or show reduction of symptoms (Akbergenov, 2006; Cao *et al.,* 2002). Such mechanisms are correlated with RNA silencing involving different fundamental processes of gene regulation, denovo histone and DNA methylation, establishment of heterochromatin and control of transposon mobility (Lipman and Martienssen, 2004; Voinet, 2005). RNA silencing in plants provides an adaptive immune system for recognizing and inactivating viruses

Plant virus-derived proteins which act as suppressors and counter the RNA silencing was reported for *Tomato yellow leaf curl China virus* (Dong *et al.*, 2003). Small RNAs of 21, 22, and 24 nt have been reported by Akbergenov (2006) and

Vanderschuren (2007) in *Nicotiana benthamiana* and cassava plants infected with African cassava mosaic virus (ACMV) and in *A. thaliana* infected with Cabbage leaf curl virus (CaLCuV) respectively.

Zrachiya *et al.* (2006) reported small interfering double stranded RNAs (siRNAs), derived from an intron-hairpin RNA (ihpRNA) construct and targeting the v1 gene product, on CP accumulation in tomato. siRNA targeted against the TYLCV CP gene, did not show disease symptoms 7 weeks post-inoculation with the virus, while non-transgenic control plants were infected within 2 weeks post inoculation.

Pepper golden mosaic virus (PepGMV)-infected pepper plants showed a recovery phenotype, which has been associated with the presence of virus-derived small RNAs. PepGMV was targeted by both posttranscriptional and transcriptional gene silencing mechanisms involving siRNAs of 21 to 22 nucleotides related to the coding regions (Rep, TrAP, REn, and movement protein genes) and a 24-nt population primarily associated to the intergenic regions (Rodríguez-Negrete, 2009).

Materials and methods

3. MATERIALS AND METHODS

The study entitled "Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex. Poir)" was carried out at the Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara during the period 2009-2011. The materials used and the methodologies adopted in this study are described below.

3.1. Materials

3.1.1 Chemicals, glassware and plastic ware

The chemicals used for the study were of good quality (AR grade) from the firms MERCK, SRL and HIMEDIA and SISCO. Molecular Biology Grade enzymes and buffers were supplied by Bangalore Genei Ltd. All the plastic wares used were obtained from Axygen and Tarson India Ltd.

3.1.2 Equipment and machinery

The equipment items available at Centre for Plant Biotechnology and Molecular Biology were used for the study. Cultures were incubated in Incubator Shaker, DK- S1010 (Dai Ki Scientific Co. Korea). Centrifugation was carried out in KUBOTA centrifuge (Kubota, Japan). DNA amplification was carried out in Eppendorf Master Cycler, Gradient (Eppendorf, Germany) and horizontal gel electrophoresis system (BIO-RAD) was used for electrophoresis. Imaging system UVP GelDoc-IT^{TM} was used for imaging the gel.

3.2 Isolation of genomic DNA

3.2.1 Collection of leaf samples

Yellow vein mosaic infected pumpkin leaf samples (Plate.3) were collected from the research field of Department of Olericulture, College of Horticulture, Vellanikkara, Thrissur. Tender leaves showing yellow vein symptoms were collected early in the morning, stored on ice and brought to lab for DNA isolation.

3.2.3 Genomic DNA isolation

Isolation of good quality DNA is one of the prerequisites for doing PCR. The total genomic DNA was extracted by the procedure reported by Rogers and Benedich (1994), with slight modification.

3.2.3.1 DNA isolation (Rogers and Benedich, 1994)

The following reagents were used in the procedure

Reagents

1. 2X CTAB extraction buffer
2. 10% CTAB solution
3. TE Buffer
4. Isopropanol
5. Chloroform: isoamylalcohol (24:1, v/v)
6. Ethanol 70% and 100%

The details of preparation of the reagents are provided in Appendix 1.

Procedure

- Leaf sample (1g) was weighed accurately and ground using liquid nitrogen in a pre-chilled mortar and pestle in the presence 10mM (50µl β-mercaptoethanol) and a pinch of sodium metabisulphate
- 3 ml of extraction buffer was added to the ground sample
- The ground tissue was transferred into a 50ml oakridge tube containing 4ml pre-warmed 2X CTAB extraction buffer
- The contents were mixed and incubated at 65°C for 20-30 minutes with occasional mixing and gentle inversion.
- Equal volume of chloroform: isoamyl alcohol (24:1) was added, mixed by gentle inversion and centrifuged at 10000 rpm for 15 min at 4 °C.
- The aqueous phase was removed with a wide bore pipette, transferred to a fresh oakridge tube and added 1/10 volume of 10% CTAB solution. The contents were mixed by gentle inversion and centrifuged at 10000 rpm for 15 minutes at 4 °C.
- The above two steps were repeated
- The aqueous phase was transferred to a fresh tube with a wide bore pipette and 1/6 volume of isopropanol was added and kept at -20 °C for 1 hour for complete precipitation of the DNA.
- The tube was centrifuged at 10000 rpm for 15 minutes at 4°C. The supernatant was discarded.
- The pellet was washed with 70% ethanol by spinning for 30 minutes at 1000 rpm.
- The pellet was air dried and dissolved in 250µl TE buffer and stored at -20 °C.

3.2.4 Purification of DNA

The DNA isolated from pumpkin leaf tissue contained RNA as contaminant and was purified by phenol precipitation and RNase treatment (Sambrook *et al.,* 1989).

Reagents

1. Phenol: chloroform mixture (1:1, v/v)
2. Chilled isopropanol
3. 70% ethanol
4. TE buffer
5. Chloroform: isoamyl alcohol (24:1, v/v)
6. 1% RNase

The RNase A from Sigma, USA was used to prepare RNase. One per cent solution was prepared by dissolving RNase A in TE buffer at 100°C for 15 minutes. The solution was cooled to room temperature, dispensed into aliquots and stored at -20°C.

3.2.4.1 Procedure for RNase treatment

1. 2µl of the RNase solution was added to100µl DNA sample and incubated at 37°C in a dry bath (Genei, Thermocon) for 1 hour.
2. The volume was made up to 250µl with distilled water and equal volume of phenol: chloroform mixture was added.
3. Centrifuged at 12000rpm for 10 minutes at 4 °C. The aqueous phase was collected in a fresh eppendorf tube and added equal volume of chloroform: isoamyl alcohol (24:1).
4. Again centrifuged at 12000rpm for 10 minutes at 4 °C.

5. The above two steps were repeated and added 0.6 volume of chilled isopropanol.
6. The mixture was incubated at -20 °C for 30 minutes
7. Centrifuged at 10000rpm for 15 minutes at 4 °C.
8. The obtained pellet was washed with 70% ethanol and air dried.
9. The pellet was dissolved in 250µl TE buffer

3.2.5 Checking the quality of DNA

3.2.5.1 Agarose gel electrophoresis

Agarose gel electrophoresis was performed based on the method described by Sambrook *et al.* (1989) to check the quality of DNA and also to separate the amplified products.

Materials required for agarose gel electrophoresis were:

1. Agarose (Bangalore Genei, Low EEO)
2. 50X TAE buffer (pH 8.0)
3. Electrophoresis unit (Biorad, USA), power pack, casting tray and comb
4. 6X Loading/ Tracking dye (Bangalore Genei, Bangalore)
5. Ethidium bromide solution (stock 10mg/ml; working concentration 0.5µg/ml (5µl/100ml gel))
7. Gel documentation and analysis system (UVP GelDoc-IT^{TM} imaging system)

(Chemical composition of the buffer and dyes are given in Appendix III)

3.2.5.2 Electrophoresis

1. 1X TAE buffer was prepared from 50 X TAE stock solution.

2. Gel tray was prepared by sealing the ends with cellophane tape. The comb was placed in the gel tray about 1 inch from one end of the tray and positioned such that the teeth were about 1-2mm above the surface of the tray.
3. Agarose (0.8 per cent (w/v) for genomic DNA and 1.0 per cent (w/v) for PCR) was weighed and added to 1X TAE. It was boiled till the agarose get melted and dissolved completely. The solution was allowed to cool to lukewarm temperature (42-45°C).
4. Two µl of ethidium bromide (0.5µg/ml) was added as an intercalating dye of DNA and mixed well.
5. The warm solution was poured into the tray to a depth of about 5mm. After the gel was completely set (15 to 20 minutes at room temperature), the comb and cellophane tape were carefully removed.
6. The gel was placed in the electrophoresis tank with the wells near the cathode and submerged in 1X TAE buffer to a depth of 1cm.
7. A piece of cellophane tape was pressed on a solid surface and 1µl 6X loading dye was dispensed in small quantity on the tape. A quantity of 3 to 5µl of DNA was added to each slot, mixed well by pipetting in and out for 2 to 3 times. Then the mixture was loaded in the wells, with the help of micropipette. λDNA/*Eco*R1+ *Hind*III double digest (Bangalore Genei) was used as the molecular weight marker.
8. The cathode and anode were connected to the power pack and the gel was run at a constant voltage of 100 volts.
9. The power was turned off when the tracking dye reached at about 3cm from the anode end.

3.2.5.3 Gel documentation

The DNA bands separated by electrophoresis were viewed and photographed using Vision Works LS software and UVP GelDoc-IT^{TM} imaging system.

3.2.6 Purity of DNA

The purity of DNA was further analysed by using NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA). The absorbance of nucleic acid sample was measured at a wavelength of 260 nm and 280 nm. OD 260/280 ratio was used to assess the purity of nucleic acids. A ratio of 1.8 to 2 indicates that the DNA is of good quality and free from protein

3.3 PCR amplifcation of coat protein and movement protein genes

3.3.1 Primer designing for coat protein and movement protein genes

Nucleotide sequence information about coat protein and movement protein genes of *geminiviruses* infecting vegetables were collected from NCBI nucleotide database.

The following steps were followed to design the primers:

1. The complete nucleotide sequences of geminiviruses infecting vegetables like cucurbits and tomato, reported by various research workers were downloaded from NCBI nucleotide database (http://www.ncbi.nlm.nih.gov) and copied in FASTA format in a notepad.

2. To assess the homology of the downloaded nucleotide sequences, multiple sequence alignment was done using Clustal W (www. ebi.ac.uk/clustalw/) (Thompson *et al.*, 1994).
3. Specific primers were designed based on conserved sequences and considering the following factors:
    a. The conserved boxes selected should have GC content not less than 50 per cent
    b. Melting temperature (Tm = 4 GC + 2 AT) ranged between 55°C and 65°C.
    c. It is preferable to have G or C at 3' end.
    d. There should not be any complementarily between forward and reverse primers.
    e. Stretches of single base should not appear within the primer sequence.
    f. Each primer should be 18 to 24bp long
    g. For designing forward primer the sequence was used and for the reverse primer, the reverse complementary sequence was taken.
    h. Based on above conditions forward and reverse primers were designed to amplify 700bp within the movement protein gene and full length ORF of coat protein gene. The primers for movement protein gene were named MFP and MRP and coat protein gene primers as CFP and CRP. The details of the sequences used are given in Tables 1 and 2.

3.3.2. PCR amplification

The total DNA extracted from symptomatic leaves of pumpkin was used as template for PCR amplification. The viral DNA regions were amplified using primers specific to coat protein and movement protein genes of geminiviruses. Polymerase

Table 1. Details of nucleotide sequences used to design primers for coat protein gene

SL No.	NCBI Accession	Description
1	AY184487.3	*Squash leaf curl China virus* - [Pumpkin :Coimbatore] segment DNA-A, complete sequence
2	AY686500.1	*Pumpkin yellow vein mosaic virus* coat protein gene, complete cds
3	AM286433.1	*Tomato leaf curl New Delhi virus* - [Potato] isolate Happur segment DNA A, complete sequence
4	DQ026296.2	*Squash leaf curl China virus* - [Pumpkin: Lucknow] segment DNA A, complete sequence

Table 2. Details of nucleotide sequences used to design primers for movement protein gene

SL No.	NCBI Accession	Description
1	AY184488.1	*Squash leaf curl China virus* - [Pumpkin :Coimbatore] segment DNA-B, complete sequence
2	FJ859881.1	*Squash leaf curl China virus* [India:Varanasi:Pumpkin] segment DNA-B, complete sequence
3	EF043233.1	*Tomato leaf curl New Delhi virus* - [Potato] isolate Happur segment DNA B, complete sequence
4	EU479711.1	*Squash leaf curl virus* isolate YL segment DNA B, complete sequence

chain reaction was carried out in Eppendorf *Master Cycler, Gradient* (Eppendorf, Germany).

3.3.2.1 Composition of the reaction mixture for PCR

1. Genomic DNA (100ng)	-	1.0µl
2. 10X Taq assay buffer	-	2.5µl
3. dNTP mix (10mM)	-	1.0µl
4. Forward primer (10pM)	-	1.0µl
5. Reverse primer (10pM)	-	1.0µl
6. Taq DNA polymerase (0.3 U)	-	2.0µl
7. Autoclaved distilled water	-	16.5µl
		25.0µl

The reaction was set in a 200µl microfuge tube chilled over ice flakes. A momentary spin was given to mix the reagents and set in thermal cycler for amplification under suitable programme. A negative control was maintained without the template DNA.

3.3.2.2 Thermal cycler programme

Steps followed in PCR, temperature profile and time duration of each step for amplification of coat protein gene are given in Table 3.

3.3.2.3 Standardisation of the annealing temperature by gradient PCR

Gradient PCR was carried to determine the annealing temperature of the primers. A temperature range of 55±10°C was given for annealing. The range of annealing temperature tested is indicated in Table 4.

Table 3. Temperature profile for amplification of coat protein and movement genes by PCR

Sl. No	Step	Temperature (°C)	Time (min)
1	Initial denaturation	94	2.00
2	Denaturation	94	0.45
3	Annealing	Varied with the primers	1.00
4	Primer extension	72	1.00
5	Step 2 to 4	29 cycles	
6	Final extension	72	10.00

Table 4. Details of annealing temperature in Gradient PCR

Lane No.	Temperature (°C)
1	45.0
2	45.3
3	46.4
4	48.0
5	50.4
6	53.0
7	55.8
8	58.5
9	61.0
10	63.1
11	64.7
12	65.6

3.4 Cloning and sequencing of the PCR amplified fragments

3.4.1 Gel elution of PCR amplified fragments

Products obtained from PCR reactions were loaded separately on 1.0 per cent (w/v) agarose gel and the desired band in each case was eluted using DNA Gel Extraction Kit (Axygen Biosciences, USA) following the procedure as per the manufacturer's guidelines).

1. The DNA fragment of interest was excised from the agarose gel with a clean sharp scalpel under ultraviolet illumination. The excised gel slice was briefly placed on absorbent towel to remove the residual buffer.
2. Gel slice was weighed in a 2ml micro centrifuge tube. The gel slice was then crushed with a pipette tip and spun at 12000g for 30 sec.
3. Three volumes of buffer DE-A (w/v) was added to one volume of the gel.
4. The tube was incubated at 75°C until the gel was completely dissolved (6-8 min).
5. Buffer DE-B was added at half the volume of Buffer DE-A .
6. A miniprep column was placed in a microfuge tube. The solubilised agarose gel was transferred to the column and centrifuged at 12000 rpm for 1 min.
7. The filtrate from the 2ml microfuge tube was discarded.
8. The miniprep column was returned to the 2 ml microfuge tube and 500 µl of Buffer W1 was added. This was centrifuged at 12000 rpm for 30 sec.
9. The filtrate was again discarded from the 2ml microfuge tube and the miniprep column was placed into the microfuge tube. Then 700µl of Buffer W2 was added and centrifuged at 12000 rpm for 1 min.

10. The filtrate from the 2ml microfuge tube was discarded and the miniprep column was placed in the 2ml microfuge tube. It was centrifuged at 12000 rpm for 1 min.
11. The miniprep column was transferred into a clean 1.5 ml microfuge tube and 25-30µl of the eluent was added to the centre of the membrane to elute the DNA. It was allowed to stand for 1 min at room temperature and then centrifuged at 12000 rpm for 1 min.

3.4.2 Cloning

3.4.2.1 Preparation of competent cells

Competent cells for plasmid transformation were prepared using Genei Competent cell preparation Kit (B) from Bangalore Genei, following the manufacturer's guideline. Reagents used for competent cell preparation are given in Appendix IV.

Procedure:

Day 1

1. *Escherichia coli* JM 109 was streaked on LB agar plate from the stab and incubated at 37°C for 16 to 18 hours.

Day 2

1. 100ml LB broth in 1 litre conical flask was inoculated with 10 to 12 moderately sized colonies from LB agar plates.
2. Overnight incubation was given at 37°C in a shaker at 200 rpm. When the OD_{600} reached 0.3 (3 to 3.5 h only), growth was arrested by chilling. The flask was chilled for 20 minutes.

3. The entire culture was transferred into a 50ml centrifuge tube and centrifuged at 3500 rpm for 15 minutes at 4°C.
4. The supernatant was discarded. Keeping the tubes on ice, the bacterial pellet was resuspended in 33.3ml ice cold solution A.
5. The tubes were kept on ice for 20 minutes and then centrifuged at 3500 rpm for 15 minutes at 4°C.
6. Supernatant was discarded and pellet chilled on ice. The pellet was resuspended in 5 to 6ml of ice cold solution A.
7. The suspension was kept on ice for 10 minutes and aliquots of 100µl were dispensed in chilled 1.5ml Eppendorf tubes.
8. The tubes were frozen on ice for a few minutes before storing at -80°C.

3.4.2.2 Checking the competence of the competent cells

The competence of the cells prepared was confirmed by transformation using pUC18. The cells were plated on LB agar containing 50mg/l ampicillin, layered with X-gal and IPTG.

Procedure

1. The vial with competent cells was thawed on ice.
2. pUC18 (1 µl) was added to the vial, contents mixed gently and kept on ice for 40 minutes.
3. Then it was rapidly taken from ice; heat shock was given at 42°C exactly for 90 seconds without shaking and placed back on ice for 5 minutes.
4. 250µl of LB broth was added to the vial under sterile conditions and the tube was inverted twice to mix the contents.
5. The tube was incubated at 37 °C for 1 hour

6. 100µl and 250µl aliquots of the competent cells were plated on LB/ampicillin (50mg/l) plates layered with IPTG (6µl) and X-gal (60µl). (Stock: Ampicillin-5mg/ml in water, IPTG-200mg/ml in water, X-gal-20 mg/ml in DMSO.

3.4.2.2.1 Calculation of transformation efficiency

The transformation efficiency was calculated using the formula (Tu *et al.*, 2005) given below.

$$\text{Transformation efficiency} = \frac{\text{total number of colonies obtained}}{\text{Amount of DNA plated (µg/ml)}}$$

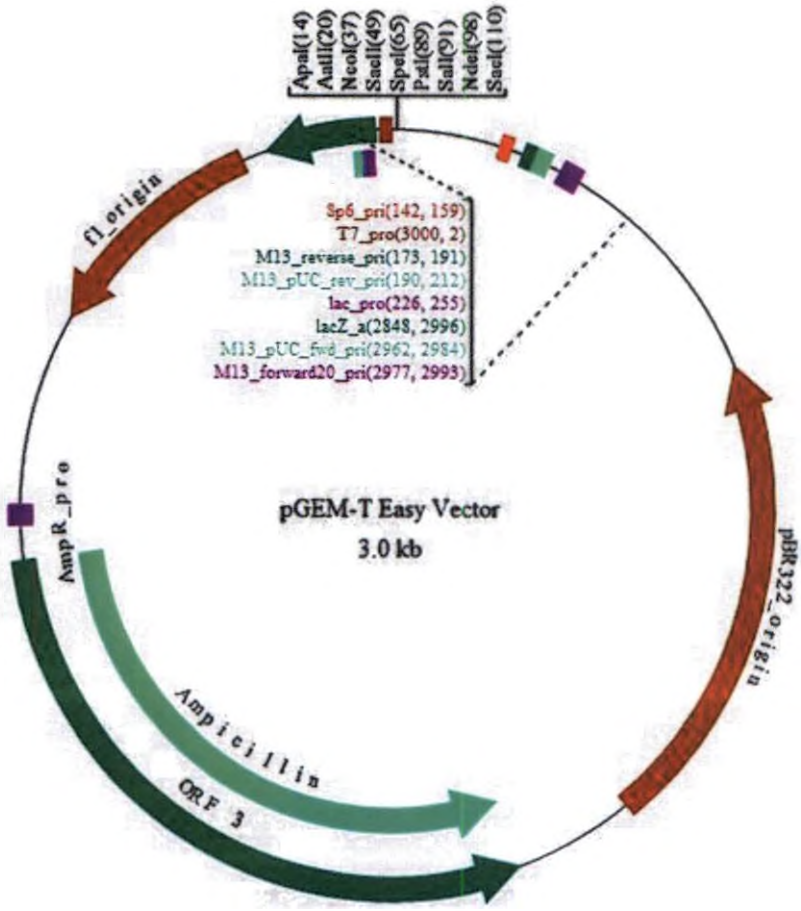
3.4.2.3 Ligation

The eluted product was ligated in pGEMT vector (Plate 2) using pGEMT Easy Vector System (Promega Corporation, USA), following the manufacturer's protocol.

1. Reaction mixture was prepared as described below:

2X rapid ligation buffer	-	5.0µl
pGEMT Easy Vector (50ng)	-	1.0µl
PCR product	-	1.0µl
T4 DNA ligase (3 units/µl)	-	1.0µl
Deionised water	-	2.0µl
		10.0µl

2. The reaction mixture was incubated for one hour at room temperature. Then it was kept at 4^0C overnight. Next day it was used for transformation in competent cells of *E. coli.*



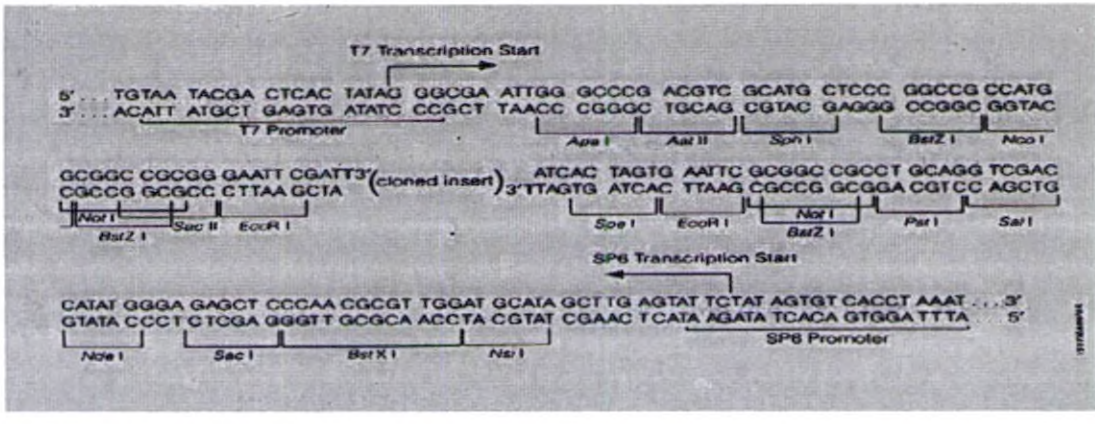


Plate. 2 Structure and sequence of pGEM-T Easy vector

3.4.2.4 Transformation of competent cells with ligated DNA

1. The vial containing competent cells was thawed on ice.
2. The ligated product was added to the competent cells, contents mixed gently and kept on ice for 40 minutes.
3. The tube was taken from ice; heat shock was given at 42°C exactly for 90 seconds without shaking and placed back on ice for 5 minutes.
4. 250µl of LB broth was added and the tube was inverted twice
5. The tube was incubated at 37°C for one hour with shaking.
6. 100µl and 250µl aliquots of the transformed cells were plated on LB agar/ampicillin layared with Xgal and IPTG as in the case of checking competence (3.4.2.2) and incubated overnight at 37°C.
7. The recombinant clones were selected based on blue-white screening.

3.4.3 Confirmation of the presence of DNA insert

The putative transformants which appeared as white colonies were picked from the plate and streaked on grided LB agar plates containing ampicillin, X-gal and IPTG. The plate was incubated overnight at 37°C and blue colonies were also streaked as a control. This was done to check whether the white colonies retained their colour and also to get more colonies of each transformant. Then it was stored at 4°C for further use.

3.4.3.1 Colony PCR

The putative transformants which appeared as white colonies were picked from the plate, sub cultured on LB agar plates containing ampicillin, X-gal and IPTG and incubated overnight at 37 °C to get more number of colonies of each transformant.

Colony PCR was carried out by using the specific primers (T7 and SP6) to confirm the presence of insert. Recombinant bacterial colony was taken by inoculation loop, mixed with 20µl sterile water and kept at 94°C for 2 minutes for denaturation. After a brief centrifugation to sediment the bacterial cell constituents, 2µl of supernatant was taken and used as a template DNA for amplification of specific DNA insert. A blue colony was taken from each plate as a negative control. The products were run on one per cent agarose gel. The gel was examined for the presence of bands with the expected amplicon size. Details of the steps followed, temperature profile and time duration of each step of colony PCR is given in Table 5.

3.4.3.2 Composition of the reaction mixture for colony PCR

a) Template DNA	-	2.0µl
b) 10X Taq assay buffer	-	2.5µl
c) d NTP mix (10mM)	-	1.0µl
d) Forward primer	-	1.0µl
e) Reverse primer	-	1.0µl
f) Taq DNA polymerase (0.3 U)	-	2.0µl
g) Autoclaved distilled water	-	10.5µl
		20.0µl

Table 5. Temperature profile for colony PCR

Sl. No	Step	Temperature (°C)	Time (min)
1	Initial denaturation	94	2.00
2	Denaturation	94	0.45
3	Annealing	55	1.00
4	Primer extension	72	1.00
5	Step 2 to 4	29 cycles	
6	Final extension	72	10.00

3.4.4 Maintenance of the clones

3.4.4.1 Preparation of pure culture of recombinant bacteria

Single white colonies from transformed plate were streaked on LB agar containing ampicillin with a flame sterilized loop. The plate was incubated overnight at 37 °C for the bacterial multiplication. Then it was stored at 4 °C for further use.

3.4.4.2 Preparation of stabs

LB agar medium was melted, added antibiotic ampicillin (50mg/ml) and poured into sterile storage vial, aseptically and allowed to solidify. In the laminar air flow (LAF), single white colonies were taken with a sterile bacterial loop and plunged into the LB agar media in the vials and sealed with parafilm. The cultures were incubated overnight at 37°C and stored at 4 °C for further use.

3.4.5 Sequencing of DNA clones

After confirmation of the presence of the expected insert in the clone, the recombinant clones with nearly 900bp and 700bp inserts were sequenced at Chromous Biotech PVT. LTD. Bangalore using T7 primer to obtain 5'-3' sequence information of the insert from the forward region, using automated sequencer (ABI–31100 Genetic Analyzer, USA). SP6 primer was used to obtain 3'-5' sequence information.

3.5 *In silico* sequence analysis

3.5.1 Vector screening

The sequences obtained were subjected to vector screening to find the vector contamination in the nucleic acid sequences. It was done by VecScreen tool available at the NCBI website (www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html). The located vector sequences were removed using the tool, Bioedit.

3.5.2 Nucleic acid sequence analysis

3.5.2.1 Homology Search

The Blastn programme (http://www.ncbi.nm.nih.gov/blast/) was used to find out the homology of nucleotide sequences (Altschul *et al.*, 1997) with the NCBI database.

3.5.2.2 Detection of open reading frame (ORF)

The programme ORF finder (www.ncbi.nlm.gov/orf) of NCBI was used to find the open reading frames in the nucleotide sequences. The ORFs were noted and saved. Blastp was used to find out the homology the of deduced amino acid sequences. The results obtained were saved for further interpretations.

3.5.2.3 Nucleotide composition

The nucleotide composition of the sequences was determined by nucleotide statistics (NASTATS) tool offered by Biology Workbench (http://seqtool.sdsc.edu/).

3.5.2.4 Restriction analysis

The restriction sites available in the DNA sequences for various restriction enzymes were detected by restriction analysis tool offered by NEB cutter V2.0 (http;//tools.neb.com/NEBcutter2/). The sites for all Type II and commercially available Type III restriction enzymes that cut the sequence just once were detected.

3.5.3 Amino acid analysis

Physical and chemical properties of the deduced amino acid sequences were determined using amino acid statistics tools (AASTATS) offered by Biology Workbench (http://seqtool.sdsc.edu/). Other analyses carried out were Kyte and Doolyte (Kyte and Doolyte, 1987) hydropathy plot using hydropathy tool of molecular tool kit (www. vivo. colostate. Edu/molkit/index.html) and secondary structure prediction using Sopma programme of Expasy tool (www.expasy.org/tools).

3.5.4 Phylogenetic analysis

Phylogenic analysis was done using the software MEGA 5.05 version. Dendrogram was constructed by Neighbor-Joining Tree method with thousand boot strapping. The details of the accessions used for phylogenetic analysis are given in the Table 6 and 7.

Table 6. Accessions used for phylogenetic analysis of coat protein gene

Accession	Description
AY184487.3	*Squash leaf curl China virus* - [Pumpkin :Coimbatore] segment DNA-A, complete sequence
DQ026296.2	*Squash leaf curl China virus* - [Pumpkin: Lucknow] segment DNA A, complete sequence
AM286794.1	*Squash leaf curl China virus* - [Cucurbita pepo: Lahore] AV2 gene, CP gene, AC1 gene, TrAP gene, ReN gene, AC4 gene and AC5 gene, clone CPoAL4(3)
EU573715.1	*Squash leaf curl China virus* - [Pumpkin:Varanasi] segment DNA-A, complete sequence
GU967381.1	Squash leaf curl China virus-[Varanasi:Pumpkin:08] segment DNA-A, complete sequence
AY396151.2	*Squash leaf curl China virus* - [Pumpkin: Lucknow] coat protein (AV1) gene, complete cds
AM286434.1	*Tomato leaf curl New Delhi virus*-[Pumpkin:New Delhi] segment DNA-A, complete sequence, isolate 2
AM286433.1	*Tomato leaf curl New Delhi virus*-[Pumpkin:New Delhi] segment DNA-A, complete sequence, isolate 1
EF043230.1	*Tomato leaf curl New Delhi virus* - [Potato] isolate Happur segment DNA A, complete sequence
AM850115.1	*Tomato leaf curl New Delhi virus* DNA-A complete genome, isolate Himachal
NC010307.1	*Emilia yellow vein virus*-[Fz1], complete genome
FN645923.1	*Bhendi yellow vein mosaic virus* segment A, complete sequence, clone 10c-RCA-AII-F
FN645917.1	*Bhendi yellow vein mosaic virus* segment A, complete sequence, clone 10a1-RCA-AI-F
GQ288400.1	*Okra Yellow Vein Mosaic Virus* isolate Aurangabad coat protein (AV1) gene, complete cds
FJ561298.1	*Bhendi yellow vein Haryana virus* [2003:Karnal] isolate OY76 segment DNA-A, complete sequence
EF531603.1	*Ageratum yellow vein virus* isolate F4 AV2 protein (AV2) gene, complete cds; and coat protein (AV1) gene, partial cds
EF531602.1	*Ageratum yellow vein virus* isolate F3 AV2 protein (AV2) gene
EF531601.1	*Ageratum yellow vein virus* isolate F1 AV2 protein (AV2) gene
EF527823.1	*Ageratum yellow vein virus* isolate F2, complete sequence

Table 7. Accessions used for phylogenetic analysis of movement protein gene

NCBI Accession No	Description
HM566113.1	*Squash leaf curl China virus DNA B*, complete sequence
GU967382.1	*Squash leaf curl China virus* [Varanasi: pumpkin: 08], segment B, complete sequence
AF509742.1	*Squash leaf curl China virus* [B] segment DNA B, complete sequence
AM260208.1	*Squash leaf curl China virus* DNA B genome, isolate Hn6
AM260208.2	*Squash leaf curl China virus* DNA B genome, isolate G25
AB085794.1	*Squash leaf curl Philippine virus* DNA B genome, complete sequence
AM286435.1	*Tomato leaf curl New Delhi virus* [Pumpkin: New Delhi}, segment DNA B complete sequence, isolate 2
DQ873412.1	*Tomato leaf curl New Delhi virus* [chilli: Bahrain], movement protein gene complete CDS
AY184488.1	*Squash leaf curl China virus-* [Pumpkin: Coimbatore], segment DNA-B, complete sequence
AM778959.1	*Squash leaf curl China virus* segment DNAB, complete sequence
AM709505.1	*Squash leaf curl China virus* DNA B, complete sequence
FJ85988101	*Squash leaf curl China virus* [India: Varanasi: Pumpkin] segment DNA-B, complete sequence

3.6 Validation of the technique

For validation of the technique, samples of yellow vein mosaic infected pumpkin leaves were collected from the fields of Olericulture Deparment, College of Horticulture, Vellanikkara and from farmers' fields in Palakkad and Malappuaram districts. Five infected samples were taken from each location. Two healthy pumpkin leaf samples from Olericulture Department field and one each from the other two locations were also collected.

Apparently healthy leaf samples from infected and healthy plants were collected from Olericulture Department field and farmers' field in Pattambi (Plate 5) and total genomic DNA was extracted. These four DNA samples were used for PCR amplifications with the coat protein and movement protein gene specific primers.

Then PCR amplifications were carried out with the DNA of the fifteen infected samples and four healthy samples.

3.7 Primer designing for full length movement protein gene

A nucleotide sequence of ~700bp was obtained by sequencing the PCR amplicon obtained with the first set of primers specific to movement protein gene. However, this contained only partial ORF and hence an attempt was made to amplify the full length ORF by designing another set of primers. These primers were designed based on the conserved boxes in five accessions of *Squash leaf curl China virus*, the details of which are given in the Table 8.

3.8 Detection of PYVMV in plants other than pumpkin

3.8.1 Collection of leaf samples

Leaf samples of weed plants showing yellow vein symptom were collected from the pumpkin field and nearby boundaries. Leaf samples of *Ageratum conyzoides*, *Emilia sonchifolia*, *Synedrella nodiflora* and *Hibiscus surattensis* (Plate. 4) which were suspected to be the possible collateral hosts of the virus and exhibited prominent yellow vein symptoms were taken for DNA extraction. Among the cultivated crops, leaf samples of yellow vein mosaic infected okra, ash gourd and bitter gourd leaves infected with distortion mosaic were also collected (Plate 3).

3.8.2 Isolation of genomic DNA

DNA was extracted from all the plants except okra by the CTAB method described earlier. The method was slightly modified to isolate DNA from okra as the leaves contain high amount of mucilage and other phenolic compounds that could affect the quality and quantity of the extracted DNA.

3.8.2.1 DNA isolation from Okra

The procedure reported by Rogers and Benedich (1994) was modified by increasing the concentration of NaCl from 1.4M to 2M in 2X extraction buffer and from 0.7M to 1M in CTAB solution (Appendix II).

Reagents

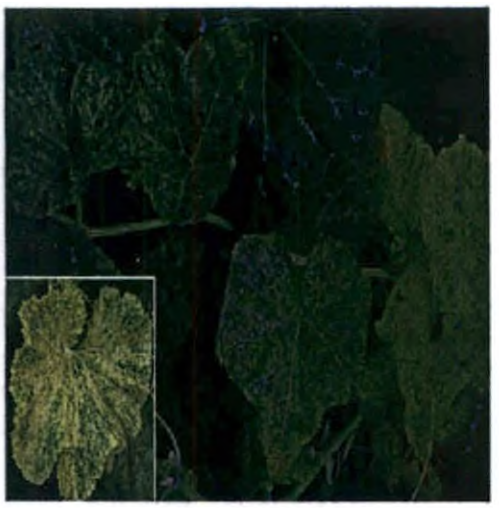
1. 2X CTAB extraction buffer
2. 10% CTAB solution

3. TE Buffer
4. Isopropanol
5. Chloroform: isoamylalcohol (24:1, v/v)
6. Ethanol 70% and 100%

The details of preparation of the reagents are provided in Appendix II.

Procedure

- Leaf sample (1g) was weighed accurately and ground using liquid Nitrogen in a pre-chilled mortar and pestle in the presence 10mM (50µl β-mercaptoethanol) and a pinch of Sodium metabisulphate.
- 4ml of extraction buffer was added to the ground sample
- The ground tissue was transferred into a 50ml oakridge tube containing 6ml pre-warmed 2X CTAB extraction buffer and mixed well.
- The contents were incubated at 65°C for 20-30 minutes with occasional mixing and gentle inversion.
- Equal volume of chloroform isoamyl alcohol mixture was added, mixed by gentle inversion and centrifuged at 10000 rpm for 15 min at 4 °C.
- The aqueous phase was removed with a wide bore pipette, transferred to a fresh oakridge tube and added 1/10 volume of 10% CTAB solution. The contents were mixed by gentle inversion and centrifuged at 10000 rpm for 15 minutes at 4 °C.
- The above two steps were repeated
- Equal volume of chloroform isoamyl alcohol mixture was added, mixed by gentle inversion and centrifuged at 10000 rpm for 15 min at 4 °C. The above aqueous phase was removed to a new centrifuge tube.
- The step was repeated.
- The aqueous phase was transferred to a fresh tube with a wide bore pipette and 1/6 volume of isopropanol was added and kept at -20 °C for 1 hour for complete precipitation of the DNA.



a. Pumpkin

b. Okra

c. Ash gourd

d. Bitter gourd

a, b and c: Yellow vein mosaic

d: Distortion mosaic

Plate. 3 Mosaic symptoms in infected leaf samples

a. *Synedrella nodiflora (Venalpacha)*

b. *Emilia sonchifolia (Muyalchevian)*

c. *Ageratum conyzoides (Appa)*

d. *Hibiscus surattensis*

Plate. 4 Leaf samples of weed plants with yellow vein mosaic symptoms

a. Sample 1

b. Sample 2

c. Sample 3

d. Sample 4

a, b and c: Collected from Olericulture Department field

d: Collected from Pattambi

Plate. 5 Apparently healthy pumpkin leaf samples used for validation of PCR

- The tube was centrifuged at 10000 rpm for 15 minutes at 4°C. The supernatant was discarded.
- The pellet was washed with 70% ethanol by spinning for 30 minutes at 1000 rpm.
- The pellet was air dried and dissolved in 250 µl TE buffer and stored at -20°C.

3.8.3 PCR amplification in weed plants and other crops

PCR amplifications with the primers designed specific to coat protein and movement protein genes were carried out using the DNA extracted from the above mentioned plant species using the thermal cycles mentioned in section 3.3.2.2 and the products were separated by agarose gel electrophoresis.

Table 8. Accessions used for designing primers for full length movement protein gene

Sl No.	NCBI Accession no.	Description
1	AY184488.1	*Squash leaf curl China virus* - [Pumpkin :Coimbatore] segment DNA-B, complete sequence
2	FJ859881.1	*Squash leaf curl China virus* [India:Varanasi:Pumpkin] segment DNA-B, complete sequence
3	AM260207.1	*Squash leaf curl China virus*, complete genome B, isolate Hn61
4	HM566113.1	*Squash leaf curl China virus* segment DNA-B, complete sequence
5	AM778959.1	*Squash leaf curl China virus* segment DNA-B, complete sequence

Results

4. RESULTS

The results of the study on "Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex Poir)" undertaken during the period 2009-2011 at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Vellanikkara are presented in this chapter.

4.1 Total DNA isolation from pumpkin infected with yellow vein mosaic virus (PYVMV)

Total DNA extracted from infected pumpkin leaves showed intact single band of DNA on one per cent agarose gel (Plate 6a). However, there was RNA contamination and RNase treatment successfully removed RNA (Plate 6b).

4.1.1 Quantification of DNA

The quantity of DNA was assessed by spectrophotometry using Nanodrop and the sample contained 4193 ng/µl of DNA. The OD_{260}/OD_{280} value was 1.87 indicating that there was no protein contamination.

4.2 PCR amplification of coat protein and movement protein genes

4.2.1 Primer designing

Primers specific for coat protein (CP) and movement protein (MP) genes were designed based on the conserved regions obtained from multiple sequence alignment. In MSA nucleotide sequences of geminiviruses retrieved from NCBI database was aligned and the accessions used are presented in Tables 1 and 2. The forward and

reverse primers of coat protein gene were named as CFP and CRP respectively and the MP gene primers were described as MFP and MRP. Details of primers are given in Table 9.

Table 9. Details of the primers

Sl No.	Primer	Nucleotide Sequence	Length (bp)	Tm (°C)	Expected amplicon size (bp)
1	CFP	5'GCAAACAACTATGGCGAAGC3'	20	60.0	896
	CRP	5'TGTTGGGGGTGATTGGTATT3'	20	58.0	
2	MFP	5'GATGCATGAAAATTGATCACG3'	21	58.0	678
	MRP	5'GGA/GGCGTTTTCATTA/C/TGATTTC3	21	58.0	

4.2.2 Standardization of annealing temperature

The annealing temperatures of the primers were determined by gradient PCR. A temperature range of 55±10°C was given for the PCR amplification. Twelve samples with different annealing temperatures were kept and all the other components of the reaction were kept constant.

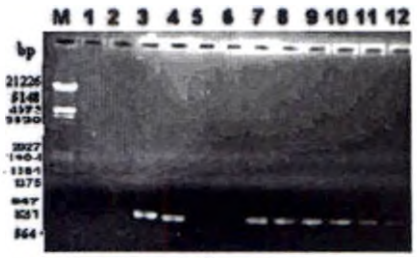
In case of coat protein gene, amplification was obtained in eight samples at annealing temperatures of 46.4, 48.0, 55.8, 58.5, 61.0, 63.1, 64.7 and 65.6°C (Table 4). When electrophoresed in one per cent agarose gel, the PCR product obtained at 55.5°C was found to be the best giving intact distinct band (Plate 6c).

For movement protein gene, amplification was obtained in seven samples. The amplicon obtained at 55.8°C produced the best band on electrophoresis (Plate 6d). Amplification was obtained at 46.4, 55.8, 58.5, 61.0 and 63.1°C. The band obtained at

a. Before RNase treatment

b. After RNase treatment

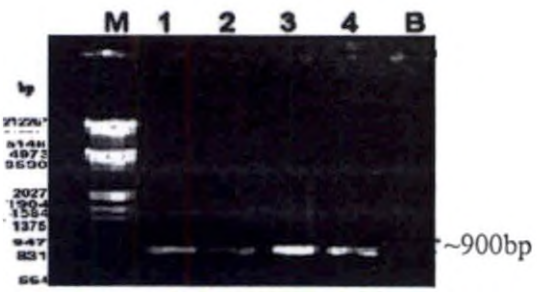
M: λDNA *Eco*R1/*Hind*Ill double digest
Lane1: Infected pumpkin DNA sample



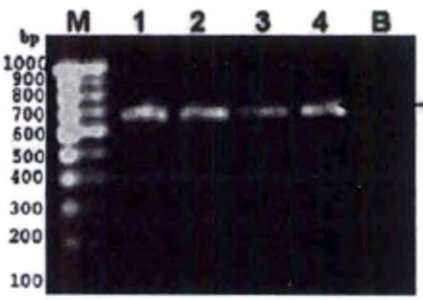
c.Coat protein gene

d. Movement protein gene

M: λDNA *Eco*R1/*Hind*Ill double digest
Lane1-12: Infected pumpkin, different annealing temperatures
Lane 1- 45.0 °C, 2- 45.3 °C, 3- 46.4 °C, 4- 48.0 °C, 5- 50.4 °C, 6- 53.0 °C, 7- 55.8 °C, 8- 58.5 °C, 9- 61.0 °C, 10- 63.1 °C, 11- 64.7 °C, 12- 65.6 °C



e.Coat protein gene
M: λDNA *Eco*R1/*Hind*Ill double digest
Lane1-4: Infected pumpkin, B: Blank



f. Movement protein gene
M: 100bp ladder
Lane 1-4: Infected pumpkin
B: Blank

Plate 6. CP and MP genes of PYVMV amplified by PCR

58.5 °C was sharp and intact. Hence for further PCR amplifications the annealing temperatures 55.5 and 58.5 °C were selected for CP and MP genes respectively (Plate 6d).

4.2.4 PCR amplification

PCR amplification of CP and MP genes was carried out separately with the specific primers designed, using 100ng of the template DNA. The PCR products were electrophoresed in one per cent agarose gel. For both CP and MP primers, a single band of expected size (~900bp and ~700bp respectively) was obtained (Plate 6e and 6f). There were no bands in the negative control suggesting that there was no primer dimer formation.

4.3 Cloning and sequencing of the PCR products

4.3.1 Gel elution of PCR products

The amplicons of coat protein (~900bp) and movement protein (~700bp) genes were eluted from the gel and purified using the gel extraction kit. The eluted products (3µl) was checked on one per cent agarose gel and intact bands of corresponding molecular weight was obtained suggesting that the quality of the product was good (Plate 7a and 8a). The quantity as assessed by nanodrop spectrophotometry for CP and MP genes was found to be 46 and 49 ng/µl respectively. The OD_{260}/OD_{280} value for CP gene was 1.82 while 1.86 for MP gene. This product was used for cloning and sequencing studies.

4.3.2 Transformation efficiency of competent cells

Competent cells of *E. coli* JM 109 strain when tested with pUC18 plasmid exhibited high degree of transformation efficiency. Numerous blue colonies were obtained when plated on LBA containing ampicillin layered with X-gal and IPTG (Plate 7b).

The transformation efficiency was found to be $1x10^7$ cfu/μg DNA, as shown below:

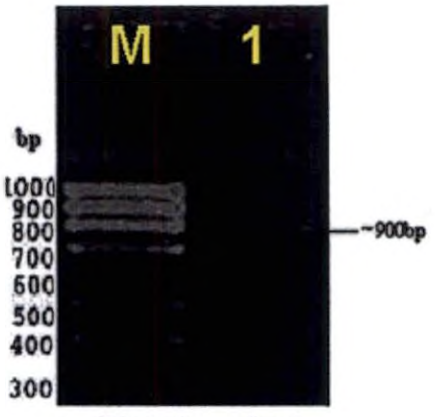
$$\text{Transformation efficiency} = \frac{\text{total number of colonies obtained}}{\text{Amount of DNA plated (μg/ml)}}$$

$$= \frac{200x\ 10^6}{20} = 1x10^7$$

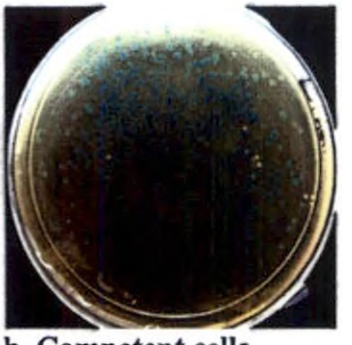
4.3.3 Transformation and confirmation of transformation

After confirmation of competence, the ligated product was transferred into competent *E. coli* JM109 cells using the heat shock method. When the transformed *E. coli* cells were grown in LBA ampicillin plates overlaid with X-gal and IPTG, a combination of blue and white colonies were obtained after overnight incubation confirming successful transformation (Plate7c and 8b).

White colonies were plated on gridded LBA ampicillin plates overlaid with X-gal and IPTG and blue colonies were also taken as a control. Most of the white colonies retained their colour, even though some got changed to blue colonies indicating the absence of the insert (Plate 7d and 8c).

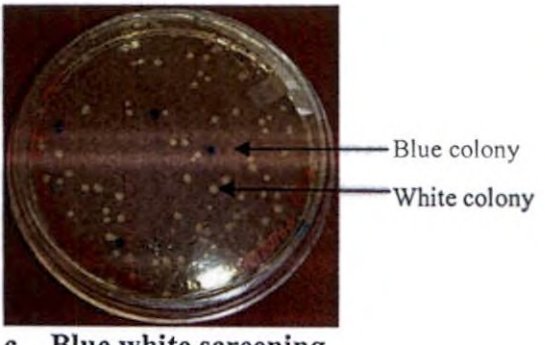


a. Eluted product



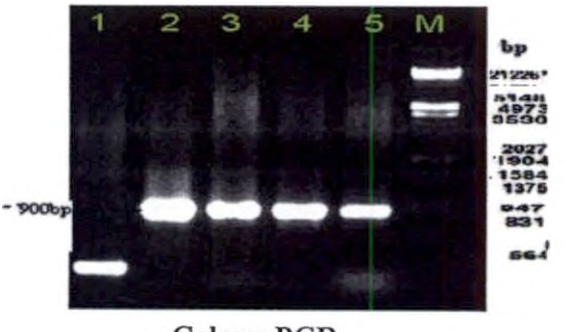
b. Competent cells transformed with pUC18

M: λDNA *Eco*R1/*Hind*III double digest
1: Eluted product



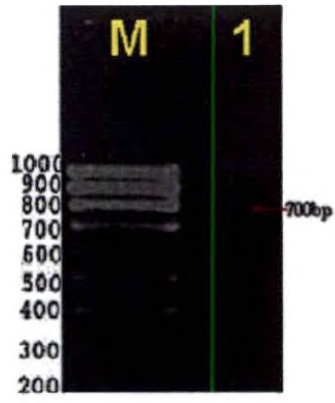
c. Blue white screening

d. Recombinants on LB with X-gal

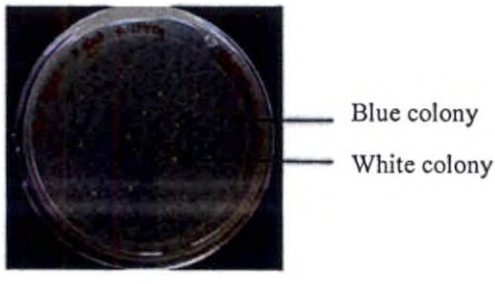


Colony PCR
Lane M: λDNA *Eco*R1/*Hind*III double digest
Lane 1: Blue colony, Lane 2-5: White colonies

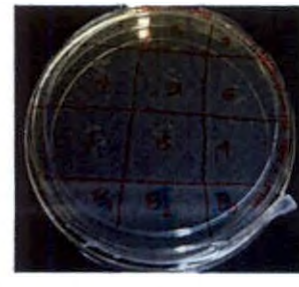
Plate. 7 Cloning and confirmation of CP gene from infected pumpkin sample



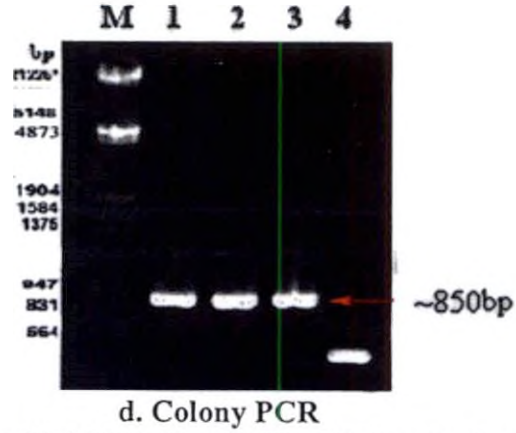
a. Eluted Product
Lane M: λDNA *Eco*R1/*Hind*III double digest
Lane 1: Eluted product



b. Blue white screening



c. Plated on LB layered with X-gal and IPTG



d. Colony PCR
Lane M: λDNA *Eco*R1/*Hind*III double digest
Lane 1-3: White colonies, Lane 4: Blue colony

Plate. 8 Cloning and confirmation of MP gene from infected pumpkin sample

4.3.4 Confirmation of presence of the insert

The presence of insert was further confirmed by performing colony PCR using universal primers T7 and SP6 as forward and reverse primers respectively. The white colonies yielded amplicons of expected size for both CP (900bp) and MP (700bp) genes. This confirmed the presence of insert DNA and the size of amplicons obtained for blue colonies was about 200bp which resulted from the amplification of the short region between the T7 and SP6 region of the vector (Plate 7e and 8d).

4.3.5 Sequencing of CP and MP amplicons

Cloned products of ~900bp from CP gene and ~700bp from MP gene were sequenced at Chromous Biotech using ABI 31100 Genetic Analyser that uses fluorescent labeled dye terminators and primers. The universal primers T7 and SP6 were used for forward and reverse sequencing respectively. The graphical outputs and nucleotide sequences of CP and MP genes are presented in Plate 10 and 13 respectively.

4.4 Full length amplification of movement protein gene

Sequencing results revealed that PYVMV from Kerala is closely related to SLCCV. Therefore, those sequences that share maximum homology with the partial MP gene sequence obtained were used for designing primers. The primers were designed to get the full length ORF of MP gene. Amplification was obtained at 46.4, 48.0, 50.4, 53.0, 58.5, 61.0, 63.1, 64.7 and 65.6 °C (Plate 9a). The annealing temperature was standardized as 58.5°C by gradient PCR. Amplification with the primers MFP1 and MRP1 yielded amplicons of ~1300bp (Plate 9b).

The PCR sample was directly sequenced without cloning using T7 and SP6 primers and a nucleotide sequence of 1317bp was obtained. The sequence was named as Pumpkin yellow vein mosaic virus movement protein (PYVMV MP) gene complete cds [Pumpkin: Kerala]. Sequence analysis was carried out using standard bioinformatics tools.

Table 10. Details of primers for amplification of full length movement protein gene

Primer	Sequence	Length (bp)	Tm (°C)	Expected amplicon size (kb)
Forward (MFP1)	CAATGTAATGC/TTAAATTACATTG	23	59	1.3
Reverse (MRP1)	CCATACCCCAATATACCATAG	21	57	

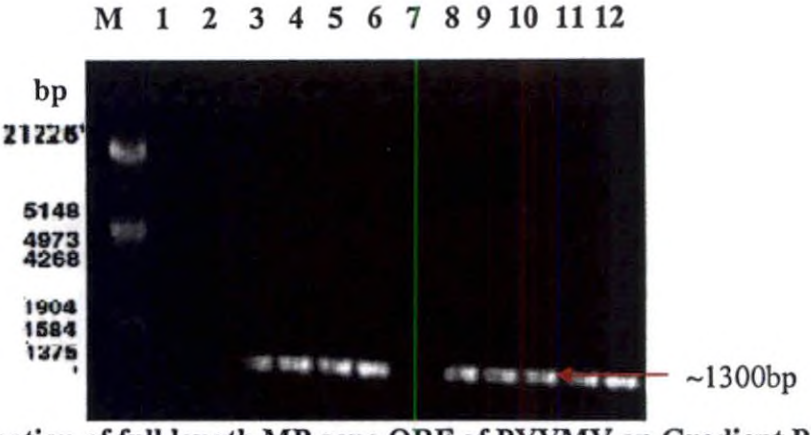
4.5 *In silico* sequence analysis

Different computer algorithms were used to analyze the nucleotide sequences and the deduced amino acid sequences obtained for coat protein and movement protein genes.

4.5.1 Nucleotide sequence analysis of coat protein gene

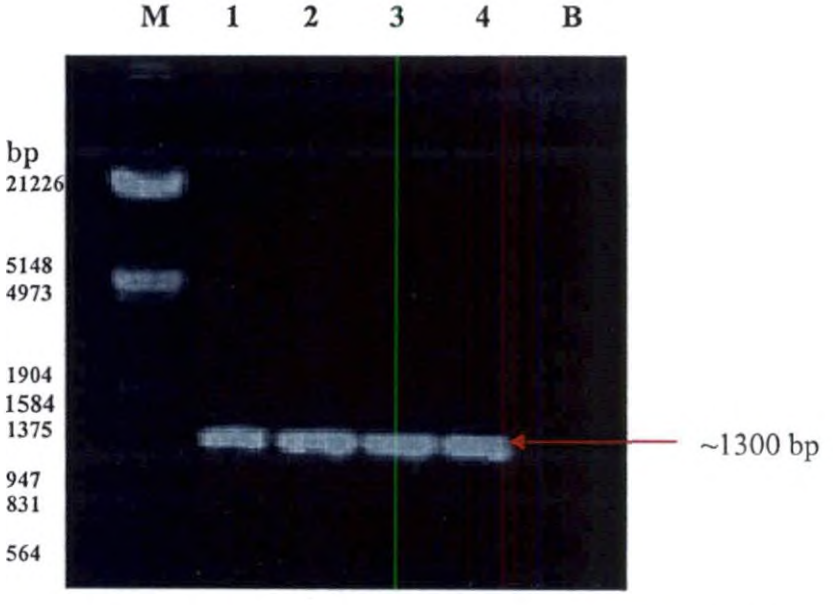
4.5.1.2. Vecscreen

There was no vector contamination in the forward sequence and the reverse sequence of coat protein showed a similarity of 53bp to the vector sequence. The



a. Amplification of full length MP gene ORF of PYVMV on Gradient PCR

Lane M: λDNA *Eco*R1/*Hind*lll double digest
Lane 1-12: infected pumpkin, different annealing temperatures
Lane 1- 45.0 ^{0}C, 2- 45.3 ^{0}C, 3- 46.4 ^{0}C, 4- 48.0 ^{0}C, 5- 50.4 ^{0}C, 6- 53.0 ^{0}C, 7- 55.8 ^{0}C, 8- 58.5 ^{0}C, 9- 61.0 ^{0}C, 10- 63.1 ^{0}C, 11- 64.7 ^{0}C, 12-65.6 ^{0}C



b. Amplification of full length MP gene ORF of PYVMV on PCR

Lane M: λDNA *Eco*R1/*Hind*lll double digest
Lane 1-4: infected pumpkin, Lane B: Blank

Plate. 9 PCR amplification of movement protein gene (full length ORF) of PYVMV

vector sequences were removed; forward and reverse sequences were merged by Emboss merger.

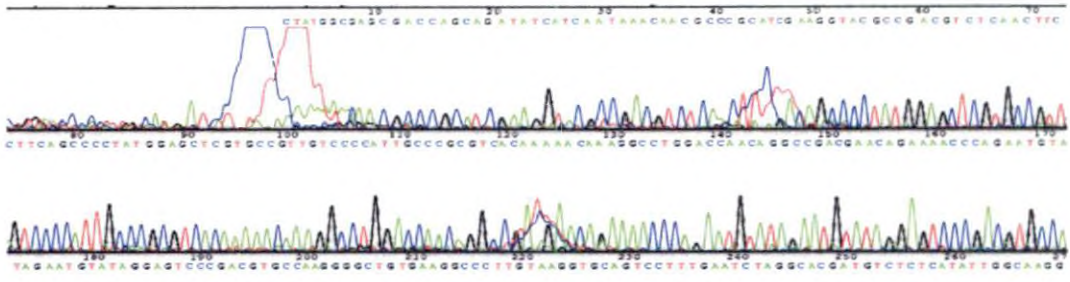
The CP gene sequence named as 'Pumpkin yellow vein mosaic virus coat protein gene [Pumpkin: Kerala]' and contained 891bp. The nucleotide sequence obtained is given below:

>Pumpkin yellow vein mosaic virus coat protein gene (891 bp)

CTATGGCGAAGCGACCAGCAGATATCATCATTTCAACGCCCCGCATCGAAGGTACGCCGACGTCTCAACTTCG
ACAGCCCCTATGGAGCTCGTGCCGTTGTCCCCATTGCCCGCGTCACAAAAACAAAGGCCTGGACCAACAGG
CCGATGAACAGAAAACCCAGAATGTATAGAATGTATAGGAGTCCCGACGTGCCAAGGGGCTGTGAAGGCCC
TTGTAAGGTGCAGTCCTTTGAATCTAGGCACGATGTCTCTCATATTGGCAAGGTCATGTGTGTTAGTGATGT
TACACGAGGAACCGGACTCACACATCGCGTAGGGAAGCGATTCTGTGTGAAATCTGTCTATGTGCTGGGGA
AGATATGGATGGATGAAAATATCAAGACTAAAAACCATACTAACAGTGTCATGTTTTTTTTGGTTCGTGATCG
TCGTCCTACAGGATCTCCCCAAGATTTTGGGGAAGTCTTTAACATGTTTGATAATGAACCGAGCACAGCAAC
GGTGAAGAATATGCATCGTGATCGTTATCAAGTCCTACGGAAGTGGCATGCGACTGTGACGGGAGGAACAT
ATGCTTCTAGAGAGCAAGCATTAGTTAGGAAATTTGTTAGGGTCAATAATTATGTTGTCTACAACCAACAAG
AAGCCGGCAAGTATGAGAATCTACTGAAAATGCATTAATGTTGTATATGGCTTGTACTCACGCATCGAACCC
TGTATACGCGACTTTGAAAATCCGACTATTTTTATGATTCGGTAACAAATTAATATAAGTTCAATTTCACATC
ATATGTTGTCCATACATCAATAAATAAA

4.5.1.3 Nucleotide Blast

Homology search of the nucleotide sequences obtained were carried out with blastn. Coat protein gene sequence showed maximum similarity of 98 per cent to *Squash leaf curl China virus* [Pumpkin: Coimbatore] DNA A (AY184487.1) with 99 per cent query coverage. It also showed 94 per cent similarity to *Tomato leaf curl New Delhi virus.* Details of accessions showing homology with the amplicons under the study are provided in Table 11.

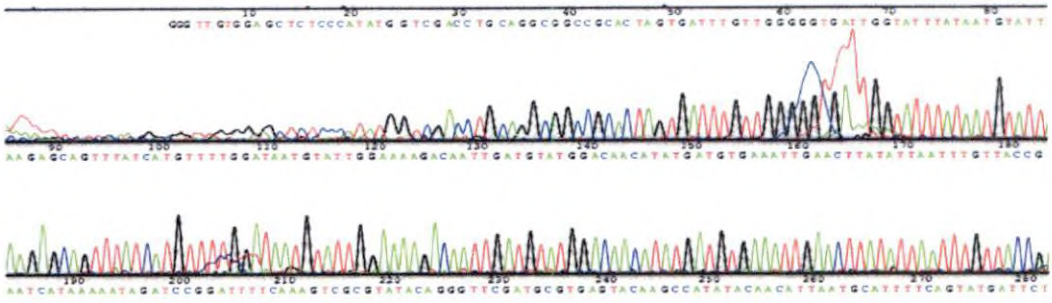


a. Graphical output

>PMC-1_T7
AATAAACAACGCCCGCATCGAAGGTACGCCGACGTCTCAACTTCCTTCAGCCCCTATGGA
GCTCGTGCCGTTGTCCCCATTGCCCGCGTCACAAAAACAAAGGCCTGGACCAACAGGCCG
ACGAACAGAAAACCCAGAATGTATAGAATGTATAGGAGTCCCGACGTGCCAAGGGGCTGT
GAAGGCCCTTGTAAGGTGCAGTCCTTTGAATCTAGGCACGATGTCTCTCATATTGGCAAG
GTCATGTGTCTTAATGATGTTACACGAGGAACCGGACTCACACATCGCGTAGGGAAGCGA
TTCTGTGTGAAATCTGTCTATGTGCTGGGGAAGATATGGATGGATGAAAATATCAAGACT
AATAACCATACTAACAGTGTCATGTTTTTTTTTGGTTCGTGATCGTCGTCCTACAGGATCT
CCCCAAGATTTTGGGGAAGTCTTTAACATGTTTGATAATGAACCGAGCACAGCAACGGTG
AAGAATATGCATCGTGATCGTTATCAAGTCCTACGGAAGTGGCATGCGACTGTGACGGGA
GGAACATATGCTTCTAGAGAGCAAGCATTAATTAGGAAATTTGTTAGGGTCAATAATTAT
GTTGTCTACAACCAACAAGAAGCCGGCAAGTATGAGAATCATACTGAAAAT

b. Nucleotide sequence

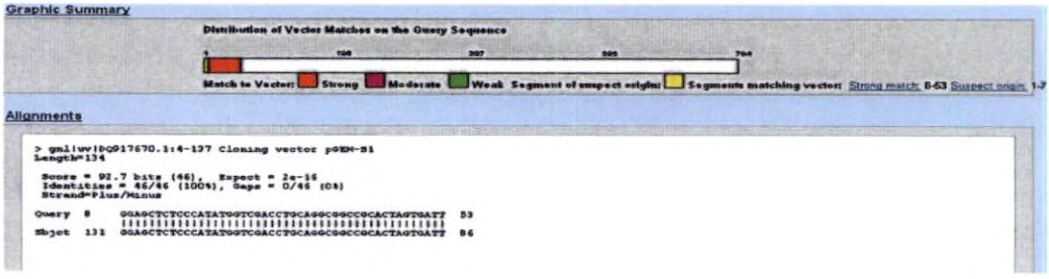
Plate. 10a Details of coat protein gene with forward primer (T7)



a. Graphical output

>PMC-1_SP6
GGGTTGTGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTGTTGGG
GGTGATTGGTATTTATAATGTATTAAGAGCAGTTTATCATGTTTTGGATAATGTATTGGA
AAAGACAATTGATGTATGGACAACATATGATGTGAAATTGAACTTATATTAATTTGTTAC
CGAATCATAAAAATAGATCCGGATTTTCAAAGTCGCGTATACAGGGTTCGATGCGTGAGT
ACAAGCCATATACAACATTAATGCATTTTCAGTATGATTCTCATACTTGCCGGCTTCTTG
TTGGTTGTAGACAACATAATTATTGACCCTAACAAATTTCCTAACTAATGCTTGCTCTCT
AGAAGCATATGTTCCTCCCGTCACAGTCGCATGCCACTTCCGTAGGACTTGATAACGATC
ACGATGCATATTCTTCACCGTTGCTGTGCTCGGTTCATTATCAAACATGTTAAAGACTTC
CCCAAAATCTTGGGGAGATCCTGTAGGACGACGATCACGAACCAAAAAAAAACATGACACT
GTTAGTATGGTTTTTAGTCTTGATATTTTCATCCATCCATATCTTCCCCAGCACATAGAC
AGATTTCACACAGAATCGCTTCCCTACGCGATGTGTGAGTCCGGTTCCTCGTGTAACATC
ACTAAGACACATGACCTTGCCAATATGAGAGACATCGTGCCTAGATTCAAAGGACTGCAC
CTTTACAAGGGCCTTCACAGCCCCTTGGCACGTCGGGACTCCTATACATTTCTATACATT
TCTGGGTTTTCTGT

b. Nucleotide sequence



c. Vecscreen output

Plate. 10b Details of coat protein gene with reverse primer (SP6)

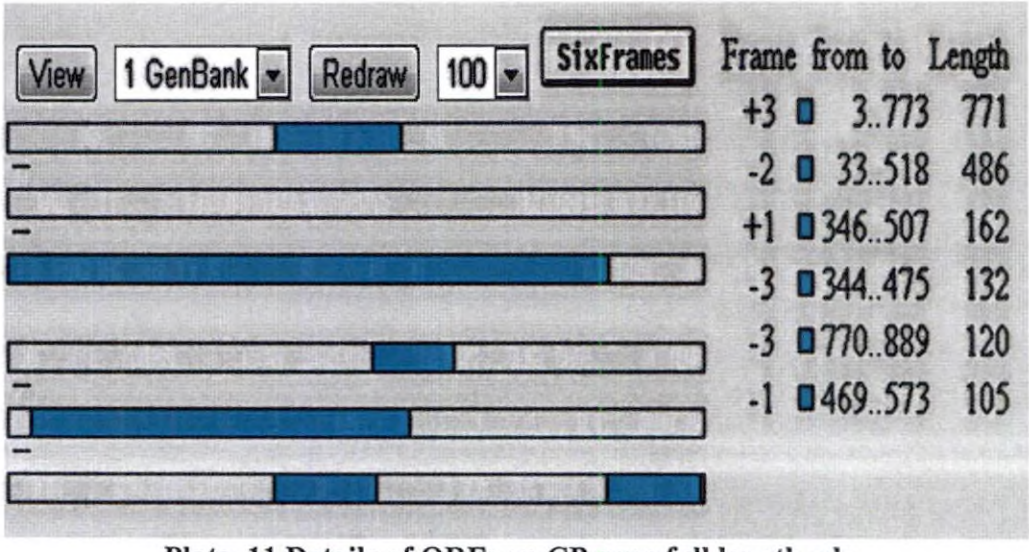


Plate. 11 Details of ORFs on CP gene full length cds

Table 11. Details of Blastn output of PYVMV CP gene complete cds

Accession	Description	Maximum score	Query coverage	E value	Maximum identity
AY184487.3	*Squash leaf curl China virus*-[Pumpkin Coimbatore] segment A	1585	99%	0.0	98%
DQ026296.2	*Squash leaf curl China virus*-[Pumpkin :Lucknow] segment A	1498	99%	0.0	97%
AM286794.1	*Squash leaf curl China virus*-[Cucurbita pepo: Lahore]AV2	1485	99%	0.0	96%
EU573715.1	*Squash leaf curl China virus*-[Pumpkin :Varanasi]segment A	1469	99%	0.0	96%
GU967381.1	*Squash leaf curl China virus*-[Varanasi: Pumpkin:08]segment A	1404	99%	0.0	95%
AM286434.1	*Tomato leaf curl New Delhi virus*- [Pumpkin: New Delhi] segment A	1327	97%	0.0	94%
AY396151.2	*Squash leaf curl China virus*-[Pumpkin :Lucknow] coat protein	1325	86%	0.0	97%
AY939926.1	*Tomato leaf curl New Delhi virus* segment A, complete sequence	1319	99%	0.0	93%
AM286433.1	*Tomato leaf curl New Delhi virus*-[Pumpkin: New Delhi]segment A	1315	97%	0.0	93%
U15015.2	*Tomato leaf curl New Delhi virus*-Severe segment A, complete sequence	1314	99%	0.0	93%

4.5.1.4 Open reading frame

The sequences were translated in all six opening reading frames (http://www.ncbi.nlm.nih.gov.ORF finder). There were six open reading frames (Plate. 12) in Pumpkin yellow vein mosaic virus coat protein gene [Pumpkin: Kerala], with the longest one located on +3 strand and starting from base 3 to 773, having a length of 771 bases. Details of the ORFs are given in Table 12 and 20.

4.5.1.5 Composition of nucleotide

Nucleotide composition of full length coat protein and movement protein genes were determined using Biology Work Bench (http://seqtool.sds.edu/). The A+T and G+C base pair composition of coat protein gene was 54.9 per cent and 45.1 per cent respectively. The details of nucleotide composition of sequences are given in the Table 21.

4.5.1.6 Restriction analysis

The restriction sites in the DNA sequence was analysed by NEBcutter V2.0 and found the large, non-overlapping open reading frames using the *E.coli* genetic code. The sites for all Type II and commercially available Type III restriction enzymes that cut the sequence just once were detected. Details of the restriction sites for 10 common restriction endonucleases in the sequence are provided in Table 13. Coat protein gene contained unique restriction sites for 64 endonucleases belonging to Type II and III (Plate 12).

Table 12. Details of ORFs on CP gene full length cds

No	Nucleotide sequence	Deduced aminoacid sequence
1	5'ATGGCGAAGCGACCAGCAGATATCATCATTTCAACGCCCGCATCGAAGGTA CGCCGACGTCTCAACTTCGACAGCCCCTATGGAGGTGCCGTTGTCCCCATTGC CCGCGTCACAAAAACAAAGGCCTGACAACAGGCCGATGAACAGAAAACCCAG AATGTATAGAATGTATGGAGTCCCGACGTGCCAAGGGGCTGTGAAGGCCCTT GTAAGGTGCAGTCCTTTGAATCTAGGCACGATGTCTCTCATATTGGCAAGGTC ATGTGTGTTAGTGATGTTACACGAGGAACCGGACTCACACATCGCGTAGGGA AGCGATTCTGTGTGAAATCTGTCTATGTGCTGGGGAAGATATGGATGGATGAA AATATCAAGACTAAAAACCATACTAACAGTGTCATGTTTTTTTTGGTTCGTGAT CGTCGTCCTACAGGATCTCCCCAAGATTTTGGGGAAGTCTTTAACATGTTTGAT AATGAACCGAGCACAGCAACGGTGAAGAATATGCATCGTGATCGTTATCAAG TCCTACGGAAGTGGCATGCGACTGTGACGGGAGGAACATATGCTTCTAGAGA GCAAGCATTAGTTAGGAAATTTGTTAGGGTCAATAATTATGTTGTCTACAACC AACAAGAAGCCGGCAAGTATGAGAATCATACTGAAAATGCATTAATGTTGTA TATGGCTTGTACTCACGCATCGAACCCTGTATACGCGACTTTGAAAATCCGGA TCTATTTTTATGATTCGGTAACAAATTAA 3'	MAKRPADIIISTPASKVRRRLNF DSPYGARAVVPIARVTKTKAW TNRPMNRKPRMYRMYRSPDVP RGCEGPCKVQSFESRHDVSHIG KVMCVSDVTRGTGLTHRVGKR FCVKSVYVLGKIWMDENIKTK NHTNSVMFFLVRDRRPTGSPQ DFGEVFNMFDNEPSTATVKNM HRDRYQVLRKWHATVTGGTY ASREQALVRKFVRVNNYVVYN QQEAGKYENHTENALMLYMA CTHASNPVYATLKIRIYFYDSV TN*
2	5'ATGCATATTCTTCACCGTTGCTGTGCTCGGTTCATTATCAAACATGTTAAAG ACTTCCCCAAAATCTTGGGGAGATCCTGTAGGACGACGATCACGAACAAAAA ACATGACACTGTTAGTATGGTTTTTAGTCTTGATATTTTCATCCATCCATATCTT CCCCAGCACATAGACAGATTTCACACAGAATCGCTTCCCTACGCGATGTGTGA GTCCGGTTCCTCGTGTAACATCACTAACACACATGACCTGCCAATATGAGAGA CATCGTGCCTAGATTCAAAGGACTGCACCTTACAAGGGCCTTCACAGCCCCTT GGCACGTCGGGACTCCTATACATTCTATACATTCTGGGTTTTCTGTTCATCGGC CTGTTGGTCCAGGCCTTTGTTTTTGTGACGCGGGCAATGGGGACAACGGCACG AGCTCCATAGGGGCTGTCGAAGTTGAGACGTCGGCGTACCTTCGATGCGGGCG TTGA 3'	MHILHRCCARFIIKHVKDFPKIL GRSCRTTITNQKKHDTVSMVFS LDIFIHPYLPQHIDRFHTESLPY AMCESGSSCNITNTHDLANMR DIVPRFKGLHLTRAFTAPWHVG TPIHSIHSGFSVHRPVGPGLCFC DAGNGDNGTSSIGAVEVETSA YLRCGR*
3	5'ATGTTAAAGACTTCCCCAAAATCTTGGGGAGATCCTGTAGGACGACGATCA CGAACCAAAAAAAACATGACACTGTTAGTATGGTTTTTAGTCTTGATATTTTC ATCCATCCATATCTTCCCCAGCACATAG 3'	MCWGRYGWMKISRLKTILTVS CFFWFVIVVLQDLPKILGKSLT CLIMNRAQQR*
4	5'ATGTGCTGGGGAAGATATGGATGGATGAAAATATCAAGACTAAAAACCATA CTAACAGTGTCATGTTTTTTTTGGTTCGTGATCGTCGTCCTACAGGATCTCCCC AAGATTTTGGGGAAGTCTTTAACATGTTTGATAATGAACCGAGCACAGCAACG GTGA3'	MLKTSPKSWGDPVGRRSRTKK NMTLLVWFLVLIFSSIHIFPST*
5	5'ATGTTGGGGGTGATTGGTATTTATAATGTATTAAGAGCAGTTTATCATGTTTT GGATAATGTATTGGAAAAGACAATTGATGTATGGACAACATATGATGTGAAA TTGAACTTATATTAA 3'	MLGVIGIYNVLRAVYHVLDNV LEKTIDVWTTYDVKLNLY*
6	5'ATGTTCCTCCCGTCACAGTCGCATGCCACTTCCGTAGGACTTGATAACGATC ACGATGCATATTCTTCACCGTTGCTGTGCTCGGTTCATTATCAAACATGTTAA 3'	

*Stop codon

Table 13. Details of restriction analysis of PYVMV CP gene complete cds

Restriction enzyme	Recognition sequence	Number of cuts	Position of restriction sites (Nucleotide)
*Alu*I	AG/CT	Nil	Nil
*Bam*HI	G/GATCC	Nil	Nil
*Btg*I	C/CRYGG	Nil	Nil
*Dpn*I	GA/TC	4	173, 189, 269, 490
*Hae*III	GG/CC	Nil	Nil
*Hinf*I	G/ANTC	3	47, 71, 402
*Hind*III	A/AGCTT	Nil	Nil
*Hpa*I	GTT/AAC	Nil	Nil
*Mbo*I	/GATC	4	171, 187, 267, 488
*Eco*RI	G/AATTC	Nil	Nil

R- A or G, Y- C or T, N- Any nucleotide

4.5. 1.7 Amino acid sequence analysis of coat protein

The nucleotide sequence was converted to amino acid sequence by using Expasy translate tool. The amino acid sequence was subjected to various computer algorithms in order to characterize them.

The proportion of each amino acid in CP was calculated by using 'AASTATS' tool (http://seqtool.sdsc.edu/). Coat protein was rich in the amino acid valine (molar per cent of 10.55) and next to this was argenine and threonine, (10.16 and 7.42 per cent respectively). The details of amino acid composition of the CP gene are presented in Table 27.

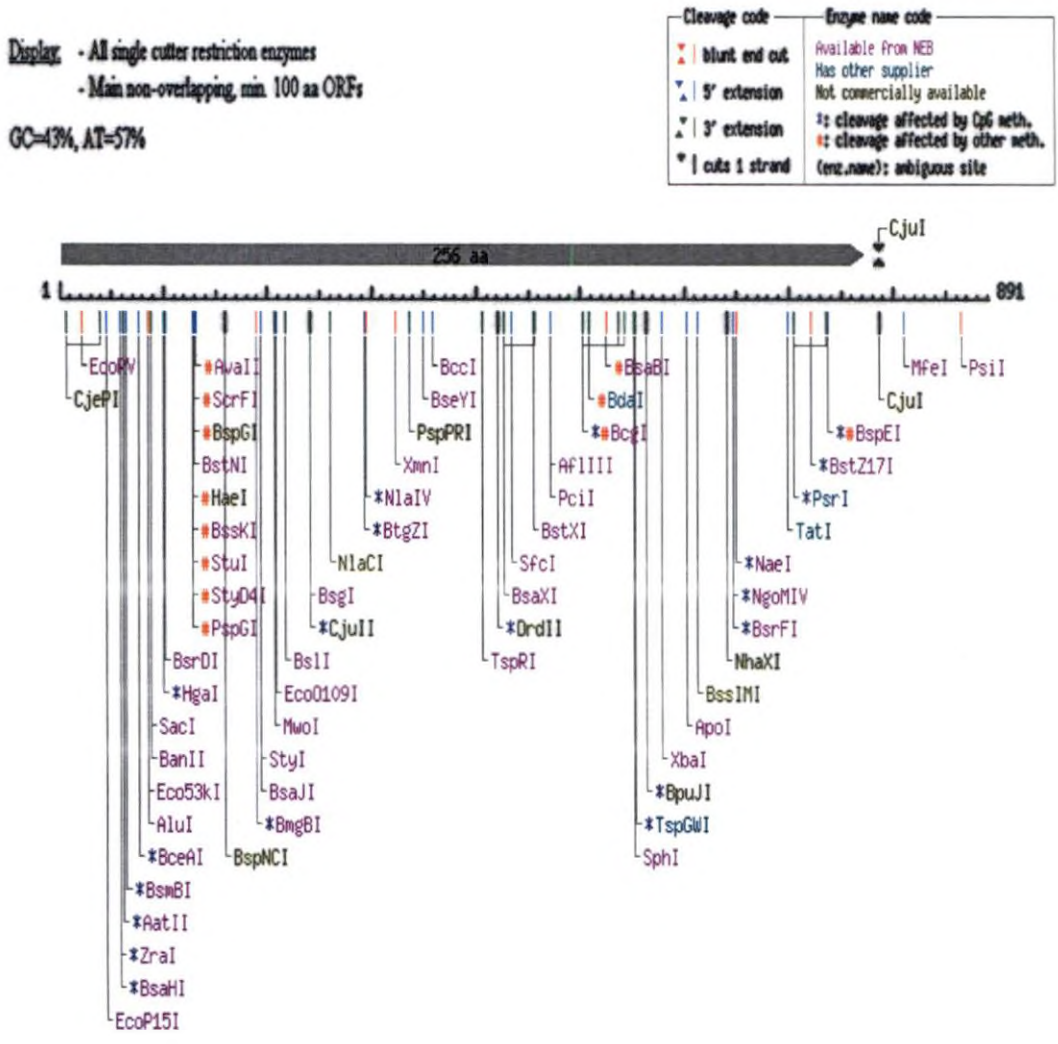


Plate 12. Diagramatic representation of restriction analysis of coat protein gene

Blastp analysis of all the ORFs was carried out with the deduced amino acid sequences. The longest ORF, +3(3-773) showed maximum similarity to coat protein of *Squash leaf curl China virus* - [Pumpkin:Varanasi] and *Tomato leaf curl New Delhi virus* (Table 14). Conserved domain of geminivirus coat protein/BR1 was detected. The ORFs of –1 (33-518) and -3 (770-889) frames showed maximum similarity to AC5 and AC3/replication enhancer proteins of SLCCV respectively (Table 15 and 16). Conserved domains of geminivirus AC4/5 and AL3 super families were detected in the ORFs –1 (33-518) and -3 (770-889) respectively. The ORF +1 (346-507) showed similarity only to V3 protein of *Soybean crinkle leaf virus* (NP543114) and AV1 protein of *East African cassava mosaic Kenya virus* (CAJ78317.1) with scores, 44.3 and 38.9 respectively. The ORF -3(344-475) showed similarity to three accessions- AC3 protein [*Tomato leaf curl New Delhi virus*- Luffa]] (AAD14627) with score 78.6, AC5 protein [*Chenopodium leaf curl virus* VEM]] (AEF12626.1) and AC5 protein [*Chenopodium leaf curl virus*] (AEF12619.1) with a score of 43.5. On Blastp analysis of the ORF -3 (469-573), no sequences with significant similarity was obtained.

Table 14. Sequences producing significant alignments with the ORF +3 (3-773) of CP

Accession	Description	Maximum score	Query coverage	E value
AAA92817.1	AV1[*Tomato leaf curl New Delhi virus*- Mild]	536	100%	0.0
ACE79034.1	Coat protein [*Squash leaf curl China virus*- Pumpkin: Varanasi]	536	100%	0.0
AEF65940.1	Coat protein [*Tomato leaf curl New Delhi virus*]	535	100%	0.0
CAH17805.1	Coat protein [*Tomato leaf curl New Delhi virus*]	535	100%	0.0
CA144681.1	Coat protein [*Tomato leaf curl New Delhi virus*- Severe [Jessore]]	535	100%	0.0
ACA24132.1	Coat protein [*Tomato leaf curl New Delhi virus*- Gorakhpur]	535	100%	0.0
CAL15144.1	Coat protein [*Tomato leaf curl New Delhi virus* [Pumpkin: New Delhi]]	535	100%	0.0
AAV98398.1	Coat protein [*Tomato leaf curl virus*]	535	100%	0.0
AAL60213	AVI protein [*Tomato leaf curl New Delhi virus* [PkT5/6]]	535	100%	0.0

Table 15. Sequences producing significant alignments with the ORF -2(33-518) of CP

Accession	Description	Maximum score	Query coverage	E value
AAY63957.3	AC5 [*Squash leaf curl China virus* - [Pumpkin: Lucknow]]	320	100%	2e-111
AAY63957.3	AC5 protein [*Squash leaf curl China virus* - [Cucurbita pepo: Lahore]]	316	100%	6e-110
ACE79039.1	AC5 protein [*Squash leaf curl China virus* - [Pumpkin:Varanasi]]	315	100%	1e-109
ADM36027.1	AC5 [*Tomato leaf curl New Delhi virus*-Severe]	306	100%	3e-106
ABB52030.1	C5 [*Tomato leaf curl New Delhi virus* - chili pepper]	304	100%	2e-105
NP_803221.1	AC5 protein [*Tomato leaf curl New Delhi virus*]	304	100%	3e-105
ADF28564.1	AC5 protein [*Squash leaf curl China virus*- [Varanasi:Pumpkin:08]]	304	99%	3e-105
ADN65589.1	Unknown [*Tomato leaf curl New Delhi virus*]	303	100%	4e-105
ABA00494.1	AC5 protein [*Tomato leaf curl New Delhi virus*]	303	100%	5e-105
ABJ97336.1	AC5 [*Tomato leaf curl New Delhi virus*-Papaya [India:New Delhi:Papaya: 2005	301	100%	3e-104

Table 16. Sequences producing significant alignments with the ORF 770-889 (-3) of CP

Accession	Description	Maximum score	Query coverage	E value
AAW55551.1	Replication enhancer protein[*Squash leaf curl China virus*]	77.0	100%	5e.18
CAJ97385.1	AC3 protein[*Squash leaf curl China virus*]	75.5	100%	2e.17
AAP47126.1	Replication enhancer protein[*Squash leaf curl China virus*-[K]]	75.5	100%	2e.17
ACB30401.1	Replication enhancer protein[*Squash leaf curl China virus*-[Wax Gourd:Nakhon Pathom]]	75.5	100%	2e.17
BAF69026.1	C3[*Squash leaf curl China virus*-[Thailand]]	74.7	100%	5e.17
BAA78101.1	Rep protein 3[*Squash leaf curl China virus*-[B]]	74.3	100%	6e.17
ACA61026.1	C3 protein[*Squash leaf curl China virus*- Pumpkin: Varanasi]	73.6	100%	1e.16
CAJ97379.1	AC3 protein[*Squash leaf curl China virus*]	73.2	100%	2e.16
ADK78875.1	AC3 protein[*Squash leaf curl China virus*]	73.2	100%	2e.16
ABM92354.1	AC3 protein[Squash leaf curl Philippines virus-[Taiwan]]	73.2	100%	2e.16

Secondary structure of CP was predicted by SOPMA tool offered by Expasy. CP contained mostly random coils (51.95 per cent) followed by extended strands, alpha helix and beta turn. Details of the output are given in Plate 13.

Kyte and Doolittle hydropathy plot was used to characterize the hydrophobic character, which is useful in predicting transmembrane regions and potential antigenic sites that are likely exposed on the protein's surface. The details of the plots are given in the Plate 14. There were no regions with score above 1.6 which indicated the absence of putative transmembrane regions. The regions above the midpoint represent hydrophobic and below are hydrophilic. The hydrophilic regions may be exposed to the surface of proteins and generally serve as antigenic sites. For predicting antigenic sites window size of seven was selected. Both hydrophilic and hydrophobic regions were present in the coat protein.

4.5.2 Nucleotide sequence analysis of movement protein gene

4.5.2.1 Vecscreen

There was no vector contamination in the forward sequence. The reverse sequence showed a similarity of 22bp to the vector sequence (Plate 15). The vector sequences were removed; forward and reverse sequences were merged by Emboss merger.

The MP gene sequence was given the name 'Pumpkin yellow vein mosaic virus movement protein gene partial cds [Pumpkin: Kerala]'. The nucleotide sequence obtained is given below:

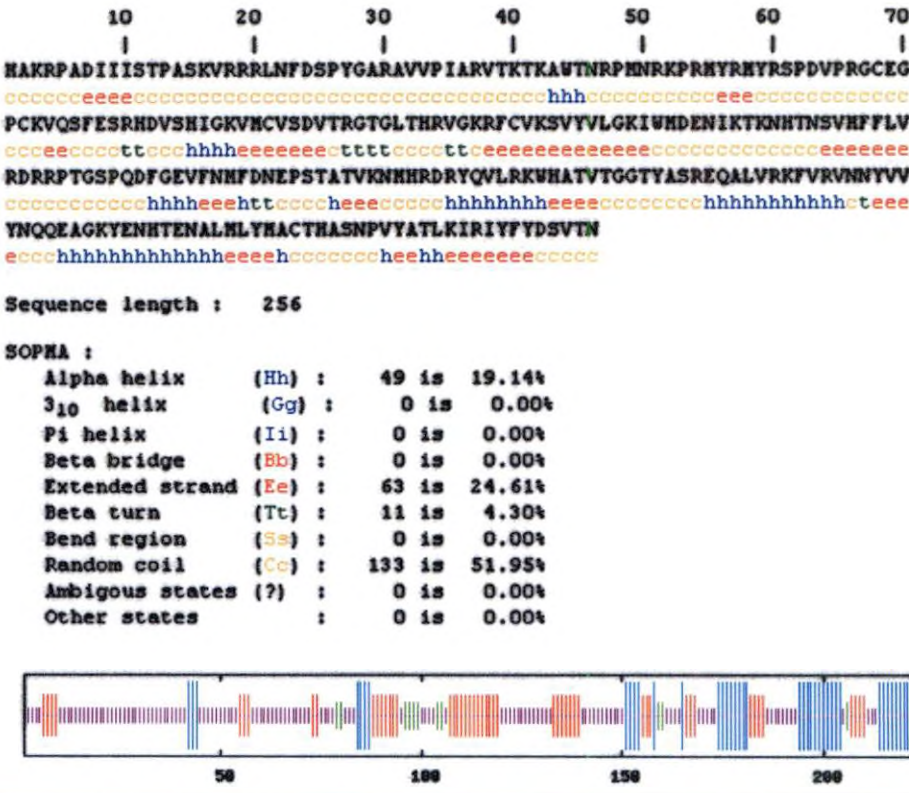
>Pumpkin yellow vein mosaic virus movement protein gene (702 bp)

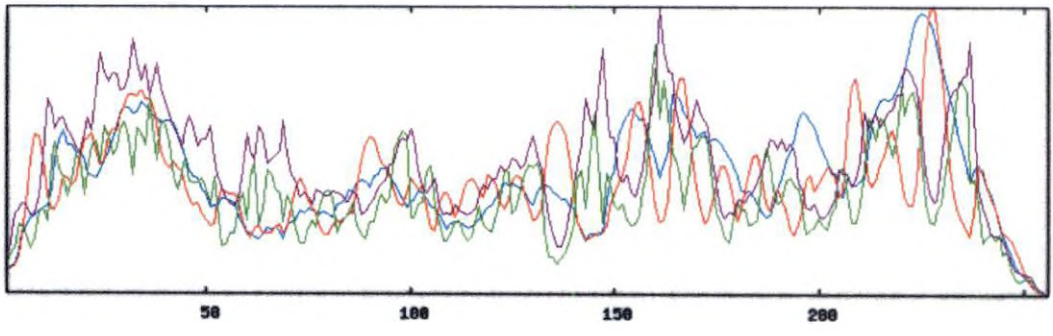
CATGGCCGCGGGATTGATGCATGAAAATTGATCACGTTCTACTTGAATATCGCAGTCAAGTACCAATTAATG
CAGTCGGGCATGTGGTTATTGAGATTCACGATATGAGATTGACAGAAGGAGACACGAAACAAGCAGAATTC
ACAATTCCGATAAAATGCAACTGCAATCTGCACTACTACTCATCAACATACTTCTCTGTTAAGGATAAAAATC
CTTGGAGAGTTGAATACAGGGTGGAGAACACGAACGTAGTAAATGGAGTACATTTTTGTAAGATGCTCGGTA
AATTAAAACTATCCTCGGCCAAACATTCGACTGACGTCGAATTCCGAGCACCAAGAATAGAGATACAGAGCA
AAGAATTCACCGTTAACGACATAGATTTCTGGTCAGTGGGATCCAAACCACAAACAAGACGGCTTGTAGATG
GATCAAGACTAATGGGTCACAGTTCAAGATCGTTACGCGTCCCACATCTTGCGATTGGTCCGAACGAATCGT
GGGCTAGTAGATCGGAGATTGGGCTGGGTTCTGTTACCAGCAGACCATATAAGAACCTCAGCGGATTAGAT
GAATCCGCATAGATCCTGGTCCATCGGCATCACAAGTGGGAAGTATTACGAGGGACGAGATAACTGAAATAA
TAACTAAAACAGTAGAACAATGTATGAAATCAAATGAAAACGCCCCATCACTA

The sequence containing full length ORF of MP gene was named as 'Pumpkin yellow vein mosaic virus movement protein gene complete cds [Pumpkin: Kerala]'. The nucleotide sequence obtained is given below:

>MP gene complete cds (1317 bp)

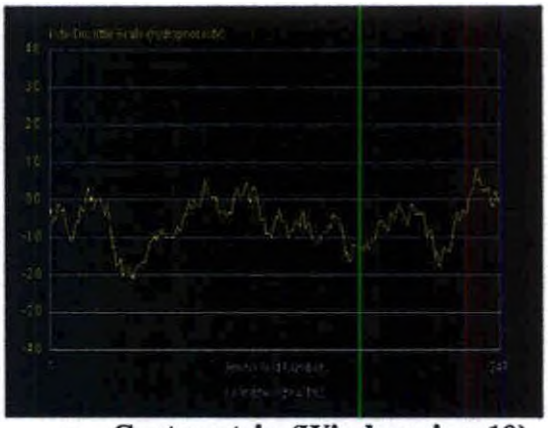
GAGAGAGATGTGCGACTTCTCTCTAACTTTCCATATAAATAAACTTCAAACGGTGTCGTTATTTGCATGGATT
GCTGGAAAGCTATTTGATGGGCCATTTTTCAATAGCCCAACAGATTATATACTCACAGCCCAATTTTAACAAT
TAGGCCTTTTGCTAAAGATCCAATTTAAAAGTAATTTAATTTTGATCTCAGCCGTTCAAAATTAGCTAGGTAT
GATAGATCATCAGTCACTGATGCAATTGTTTTGCTCCGTGGGAAATCTGTTATAAATAGATGTTAGATGTTCC
ATAGAAACACAATTGTCATGTGTTTCATCTCCAAAATATCCTTTCTTCGTTTCAAACCCTTCTTTAGCATACG
TTTCATAGTTTATTCACCTGCGGAGCCATTTCTGCTGCGCAGCCAAAATGTCCATAAGAAATGATAACATGG
GTCTTGGTGTTGGAGGGTACATAGAATCGGATCGCGTAGAATACGCTCTCACAAACGACGCTGCGGAGGTC
ACTCTGACGTTCCCTTCCATGTTCGAACAGAAAATCAGTCAATTGAGAAACAGATGCATGAAAATTGATCAC
GTTCTACTTGAATATCGCAGTCAAGTACCAATTAATGCAGTCGGGCATGTGGTTATTGAGATTCACGATATG
AGATTGACAGAAGGAGACACGAAACAAGCAGAATTCACAATTCCGATAAAATGCAACTGCAATCTGCACTAC
TACTCATCAACATACTTCTCTGTTAAGGATAAAAATCCTTGGAGAGTTGAATACAGGGTGGAGAACACGAAC
GTAGTAAATGGAGTACATTTTTGTAAGATGCTCGGTAAATTAAAACTATCCTCGGCCAAACATTCGACTGAC
GTCGAATTCCGAGCACCAAGAATAGAGATACAGAGCAAAGAATTCACCGTTAACGACATAGATTTCTGGTCA
GTGGGATCCAAACCACAAACAAGACGGCTTGTAGATGGATCAAGACTAATGGGTCACAGTTCAAGATCGTTA
CGCGTCCCACATCTTGCGATTGGTCCGAACGAATCGTGGGCTAGTAGATCGGAAATTGGGCTGGGTTCTGTT
ACCAGCAGACCATATAAGAACCTCAGCGGATTAGATGAATCCGCAATAGATCCTGGTCCATCGGCATCACAA
GTGGGAAGTATTACGAGGGACGAGATAACTGAAATAATAACTAAAACAGTAGAACAATGTATGAAATCAAAT
GTAAACGCTCCACTAACTAAAGGCGTATGATAAATTTATCGCTAAATTTATTTATAAAAGTTATTTTAAGCAT
CATGAATGCATTAACC



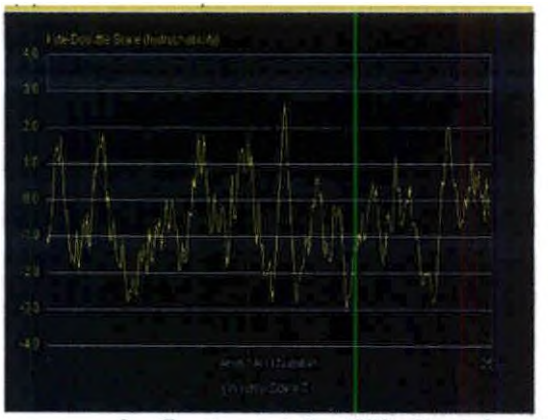


SOPMA output of coat protein

Plate. 13 Secondary structure prediction of deduced coat protein by SOPMA

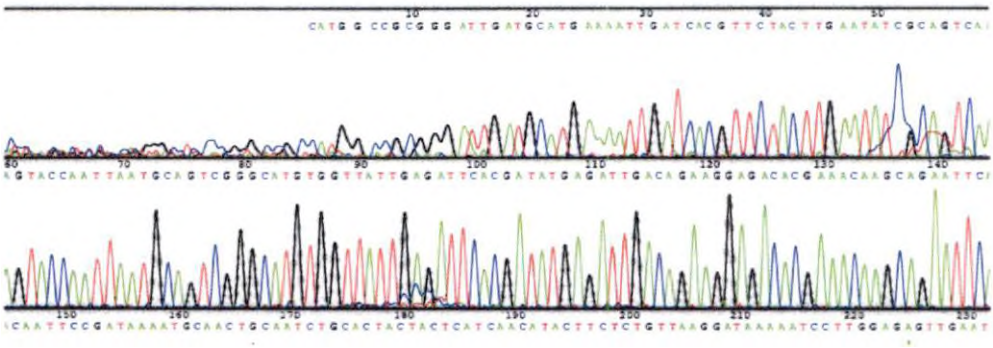


a. Coat protein (Window size=19)



b. Coat protein (Window size=7)

Plate. 14 Kyte-Doolittle hydropathicity plots of CP. The average hydrophobicity of each amino acid residue was calculated using the algorithm over window size of 19 and 7 and plotted as a function of window position



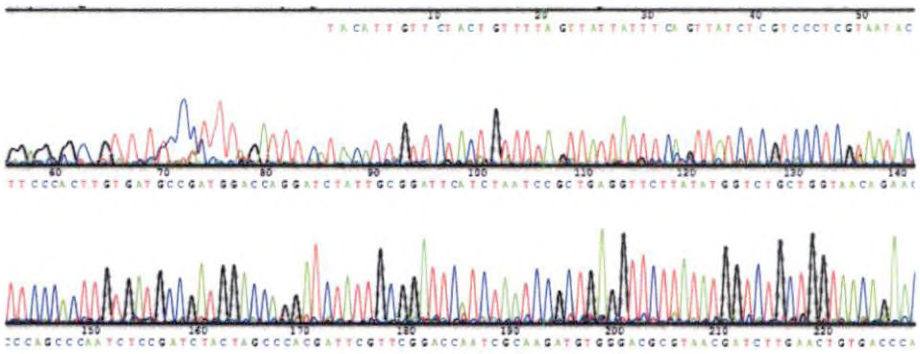
a. Graphical output

>MP-T7

CATGGCCGCGGGATTGATGCATGAAAATTGATCACGTTCTACTTGAATATCGCAGTCAAGTACCAA
TTAATGCAGTCGGGCATGTGGTTATTGAGATTCACGATATGAGATTGACAGAAGGAGACACGAAAC
AAGCAGAATTCACAATTCCGATAAAATGCAACTGCAATCTGCACTACTACTCATCAACATACTTCTC
TGTTAAGGATAAAAATCCTTGGAGAGTTGAATACAGGGTGGAGAACACGAACGTAGTAAATGGAGT
ACATTTTTGTAAGATGCTCGGTAAATTAAAACTATCCTCGGCCAAACATTCGACTGACGTCGAATTC
CGAGCACCAAGAATAGAGATACAGAGCAAAGAATTCACCGTTAACGACATAGATTTCTGGTCAGTG
GGATCCAAACCACAAACAAGACGGCTTGTAGATGGATCAAGACTAATGGGTCACAGTTCAAGATCG
TTACGCGTCCCACATCTTGCGATTGGTCCGAACGAATCGTGGGCTAGTAGATCGGAGATTGGGCTG
GGTTCTGTTACCAGCAGACCATATAAGAACCTCAGCGGATTAGATGAATCCGCATAGATCCTGGTC
CATCGGCATCACAAGTGGGAAGTATTACGAGGGACGAGATAACTGAAATAATAACTAAAACAGTAG
AACAATGTATGAAATCAAATGAAAACGCCCCATCACTA

b. Nucleotide sequence

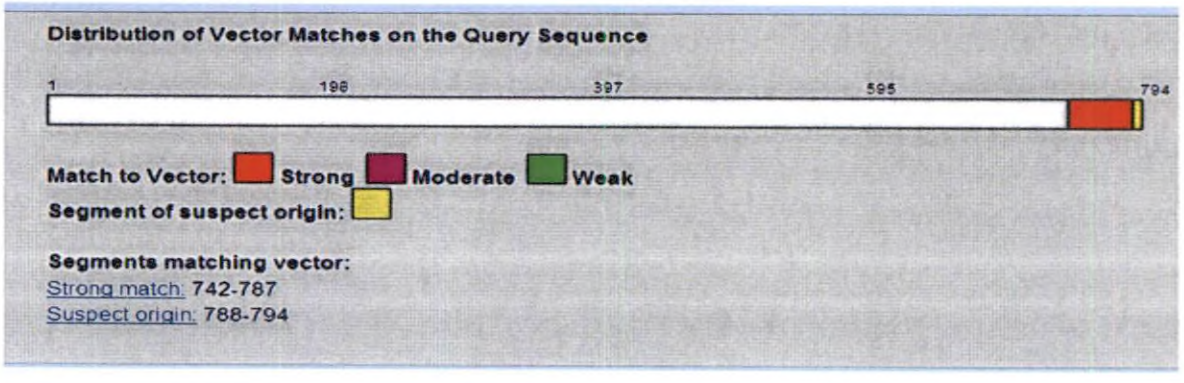
Plate. 15a Details of movement protein gene with forward primer (T7)



a. Graphical output

TTGTTGGGGGTGATTGGTATTTATAATGTATTAAGAGCAGTTTATCATGTTTTGGATAATGTATTGGAAAAGACAA
TTGATGTATGGACAACATATGATGTGAAATTGAACTTATATTAATTTGTTACCGAATCATAAAAATAGATCCGGAT
TTTCAAAGTCGCGTATACAGGGTTCGATGCGTGAGTACAAGCCATATACAACATTAATGCATTTTCAGTATGATTC
TCATACTTGCCGGCTTCTTGTTGGTTGTAGACAACATAATTATTGACCCTAACAAATTTCCTAACTAATGCTTGCTC
TCTAGAAGCATATGTTCCTCCCGTCACAGTCGCATGCCACTTCCGTAGGACTTGATAACGATCACGATGCATATTC
TTCACCGTTGCTGTGCTCGGTTCATTATCAAACATGTTAAAGACTTCCCCAAATCTTGGGGAGATCCTGTAGGACG
ACGATCACGAACCAAAAAAAACATGACACTGTTAGTATGGTTTTTAGTCTTGATATTTTCATCCATCCATATCTTC
CCCAGCACATAGACAGATTTCACACAGAATCGCTTCCCTACGCGATGTGTGAGTCCGGTTCCTCGTGTAACATCAC
TAAGACACATGACCTTGCCAATATGAGAGACATCGTGCCTAGATTCAAAGGACTGCACCTTTACAAGGGCCTTCA
CAGCCCCTTGGCACGTCGGGACTCCTATACATTTCTATACATTTCTGGGTTTTCTGT

b. Nucleotide sequence



c. Vecscreen output

Plate. 15b Details of movement protein gene with reverse primer (SP6)

4.5.2.2 Nucleotide Blast

Homology search of the nucleotide sequences obtained were carried out with blastn. Movement protein gene sequences showed maximum similarity of 98 per cent to *Squash leaf curl China virus* [Pumpkin: Coimbatore] DNA B (AY184488.1)

Details of accessions showing homology with the amplicons under the study are provided in Table 17 and 18.

4.5.2.3 Open reading frame

The sequences were translated in all six opening reading frames (http://www.ncbi.nlm.nih.gov.ORF finder). There were two open reading frames in the sequence, Pumpkin yellow vein mosaic virus coat protein gene [Pumpkin: Kerala]. The longest ORF was of 567 bases long (Plate 17).

Pumpkin yellow vein mosaic virus movement protein gene complete cds [Pumpkin: Kerala] contained five open reading frames with the longest one of 846 bases length. The details of the ORFs obtained in each sequences are given in Table 19 and 20.

4.5.2.4 Nucleotide composition of Movement protein gene

Nucleotide composition movement protein gene sequences were determined using nucleotide statistics tool offered by Biology Work Bench (http://seqtool.sds.edu/). The A+T and G+C base pair composition of movement protein was 57.7 per cent and G+C was 43.3 per cent respectively. The details of nucleotide composition of sequences are given in the Table 21.

Table 17. Sequences producing significant alignments with the MP gene partial cds

Accession	Description	Maximum score	Query coverage	E value	Maximum identity
AY184488.1	*Squash leaf curl China virus*-[Pumpkin Coimbatore] segment DNA B	1185	97%	0.0	98%
FJ859831.1	*Squash leaf curl China virus*-[India: Varanasi:Pumpkin] segment DNA B	1085	97%	0.0	94%
AM250207.1	*Squash leaf curl China virus*-complete DNA B genome, isolate Hn61	1029	97%	0.0	93%
HM566113.1	*Squash leaf curl China virus*-segment DNA B, complete sequence	1024	97%	0.0	93%
AM709505.1	*Squash leaf curl China virus* segment DNA B complete sequence	1024	97%	0.0	93%
GU967382.1	*Squash leaf curl China virus*-[Varanasi:Pumpkin: 08] segment DNA B	1018	97%	0.0	93%
AF509742.1	*Squash leaf curl China virus*-[B] segment DNA B, complete sequence	1007	97%	0.0	93%
AM260208.1	*Squash leaf curl China virus*-complete DNA B genome, isolate G25	979	97%	0.0	92%
ABO85794.1	*Squash leaf curl Philippines virus* DNA B, complete sequence	929	97%	0.0	91%

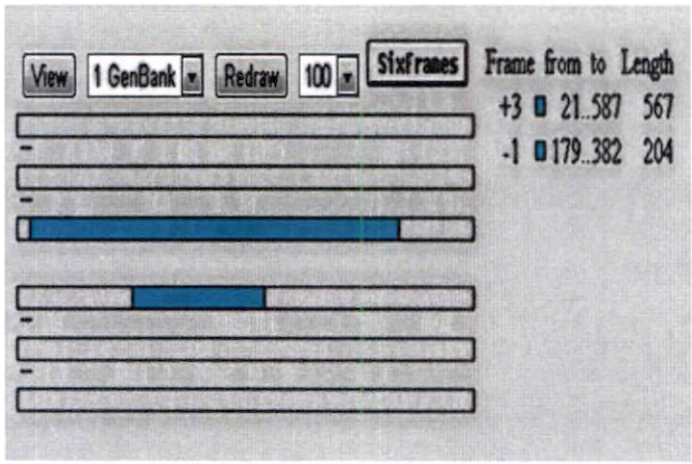
Table 18. Sequences producing significant alignments with MP gene complete cds

Accession	Description	Maximum score	Query coverage	E value	Maximum identity
AY184488.1	*Squash leaf curl China virus*-[Pumpkin Coimbatore] segment DNA-E	2198	100%	0.0	96%
GU967382.1	*Squash leaf curl China virus*-[Varanasi :Pumpkin :08] segment DNA-B	1618	100%	0.0	88%
FJ859881.1	*Squash leaf curl China virus* [India :Varanasi :Pumpkin]segment DNA-B, complete sequence	1613	100%	0.0	88%
AM709505.1	*Squash leaf curl China virus* DNA-B, complete sequence	1607	100%	0.0	88%
AM778959.1	*Squash leaf curl China virus* segment DNA-B, complete sequence	1602	100%	0.0	88%
HM566113.1	*Squash leaf curl China virus* segment DNA-B, complete sequence	1500	100%	0.0	87%
AM260207.1	*Squash leaf curl China virus* complete DNA-B genome, isolate Hn61	1463	97%	0.0	87%
AF509742.1	*Squash leaf curl China virus*-[B] segment DNA-B, complete sequence	1454	100%	0.0	86%
AM260208.1	*Squash leaf curl China virus* complete DNA genome, isolate G25	1293	79%	0.0	89%
AB085794.1	*Squash leaf curl Philippines virus* DNA-B, complete sequence	1240	72%	0.0	90%

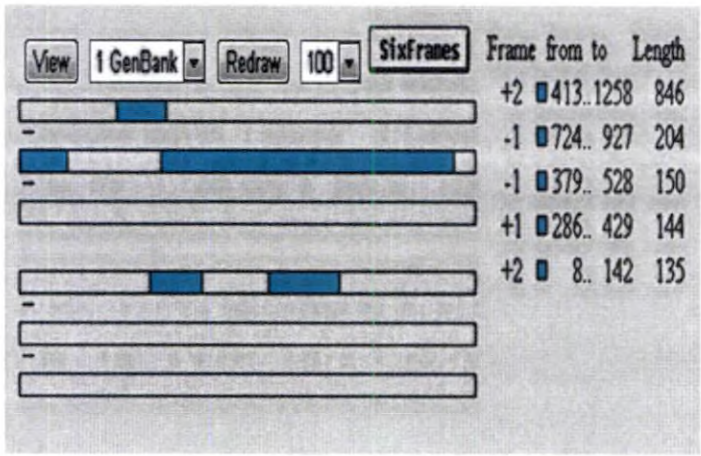
Table. 19 Details of ORFs on MP gene partial and complete cds

No	Nucleotide sequence	Deduced amino acid sequence
1	ATGAAAATTGATCACGTTCTACTTGAATATCGCAGTCAAGTACCAAT TAATGCAGTCGGGCATGTGGTTATTGAGATTCACGATATGAGATTGA CAGAAGGAGACACGAAACAAGCAGAATTCACAATTCCGATAAAATG CAACTGCAATCTGCACTACTACTCATCAACATACTTCTCTGTTAAGG ATAAAAATCCTTGGAGAGTTGAATACAGGGTGGAGAACACGAACGT AGTAAATGGAGTACATTTTTGTAAGATGCTCGGTAAATTAAAACTAT CCTCGGCCAAACATTCGACTGACGTCGAATTCCGAGCACCAAGAAT AGAGATACAGAGCAAAGAATTCACCGTTAACGACATAGATTTCTGG TCAGTGGGATCCAAACCACAAACAAGACGGCTTGTAGATGGATCAA GACTAATGGGTCACAGTTCAAGATCGTTACGCGTCCCACATCTTGCG ATTGGTCCGAACGAATCGTGGGCTAGTAGATCGGAGATTGGGCTGG GTTCTGTTACCAGCAGACCATATAAGAACCTCAGCGGATTAGATGA ATCCGCATAG	MKIDHVLLEYRSQVPINAVGH VVIEIHDMRLTEGDTKQAEFTI PIKCNCNLHYYSSTYFSVKDKN PWRVEYRVENTNVVNGVHFC KMLGKLKLSSAKHSTDVEFRA PRIEIQSKEFTVNDIDFWSVGSK PQTRRLVDGSRLMGHSSRSLR VPHLAIGPNESWASRSEIGLGS VTSRPYKNLSGLDESA*
2	ATGTCGTTAACGGTGAATTCTTTGCTCTGTATCTCTATTCTTGGTGCT CGGAATTCGACGTCAGTCGAATGTTTGGCCGAGGATAGTTTTAATTT ACCGAGCATCTTACAAAAATGTACTCCATTTACTACGTTCGTGTTCT CCACCCTGTATTCAACTCTCCAAGGATTTTTATCCTTAACAGAGAAG TATGTTGATGAGTAG	MSLTVNSLLCISILGARNSTSVE CLAEDSFNLPSILQKCTPFTTFV FSTLYSTLQGFLSLTEKYVDE*
3	ATGGAAGGGAACGTCAGAGTGACCTCCGCAGCGTCGTTTGTGAGAG CGTATTCTACGCGATCCGATTCTATGTACCCTCCAACACCAAGACCC ATGTTATCATTTCTTATGGACATTTTGGCTGCGCAGCAGAAATGGCT CCGCAGGTGA	MEGNVRVTSAASFVRAYSTRS DSMYPPTPRPMLSFLMDILAAQ QKWLRR*
4	ATGTTCCATAGAAACACAATTGTCATGTGTTTCATCTCCAAAATATC CTTTCTTCGTTTCAAACCCTTCTTTAGCATACGTTTCATAGTTTATTC ACCTGCGGAGCCATTTCTGCTGCGCAGCCAAAATGTCCATAAGAAA TGA	MFHRNTIVMCFISKISFLRFKPF FSIRFIVYSPAEPFLLRSQNVHK K*
5	ATGTGCGACTTCTCTCTAACTTTCCATATAAATAAACTTCAAACGGT GTCGTTATTTGCATGGATTGCTGGAAAGCTATTTGATGGGCCATTTT TCAATAGCCCAACAGATTATATACTCACAGCCCAATTTTAA	MCDFSLTFHINKLQTVSLFAWI AGKLFDGPFFNSPTDYILTAQF*

* Stop codon



a. ORFs in MP gene partial cds



b. ORFs in MP gene full length sequence

Plate. 16 Details of ORFs on MP gene sequences

Table 20. Open reading frames of coat protein and movement protein genes

Genes	No. of ORFs	Location	Length	Frame	Conserved domain
CP	6	3-773	771	+3	Geminivirus coat protein/BR1
		33-518	486	-2	Geminivirus AC4/5
		346-507	162	+1	Nil
		344-475	132	-3	Nil
		770-889	120	-3	Geminivirus AL3 protein
		469-573	105	-1	Nil
MP (partial)	2	21-587	567	+3	Geminivirus BL1
		179-382	204	-1	Nil
MP (Full length)	5	413-1258	846	+2	Geminivirus BL1
		724-927	204	-1	Nil
		379-528	150	-1	Nil
		286-429	144	+1	Nil
		8-142	135	+2	Nil

Table 21. Nucleotide statistics of full length CP and MP genes

Sl. No.	Gene sequence	Nitrogen base percentage (%)					
		A	T	G	C	A and T	G and C
1	Coat protein	29.3	25.6	24.4	20.8	54.9	45.1
2	Movement protein	30.3	30.9	14.9	23.9	61.2	38.8

4.5.2.5 Restriction analysis

The restriction sites in the DNA sequence was analysed by NEBcutter V2.0 and found the large, non-overlapping open reading frames using the *E.coli* genetic code. The sites for all Type II and commercially available Type III restriction enzymes that cut the sequence just once were detected. Details of the restriction sites for 10 common restriction endonucleases in the sequence are provided in Table 22. Movement protein gene contained unique restriction sites for 72 restriction endonucleases belonging to Type II and III (Plate. 17).

Table 22. Details of restriction analysis of PYVMV MP gene full length sequence

Restriction enzyme	Recognition sequence	Number of cuts	Position of restriction sites (Nucleotide)
*Alu*I	AG/CT	1	82
*Bam*HI	G/GATCC	1	944
*Btg*I	C/CRYGG	4	255, 1004, 1132, 1134
*Dpn*I	GA/TC	8	164, 191, 225, 468, 576, 946, 979, 1006
*Hae*III	GG/CC	3	94, 150, 851
*Hinf*I	G/ANTC	3	640, 1043, 1021
*Hind*III	A/AGCTT	Nil	Nil
*Hpa*I	GTT/AAC	1	919
*Mbo*I	/GATC	8	162, 189, 223, 574, 944, 977, 1058, 1060
*Eco*RI	G/AATTC	3	6

R- A or G, Y- C or T, N- Any nucleotide

4.5.2.6 Amino acid sequence analysis of movement protein

The nucleotide sequences were converted to their respective amino acid sequences using Expasy translate tool. These amino acid sequences were subjected to various computer algorithms in order to characterize them.

The proportion of each amino acid in MP was calculated by using 'AASTATS' tool (http://seqtool.sdsc.edu/). Movement protein was rich in serine (10.32 per cent) and valine (8.54 per cent). The details of amino acid composition are presented in the Table 23.

Blastp analysis of the longest ORFs on the MP sequences showed maximum similarity to movement protein of *Squash leaf curl China virus*. The details are provided in Table 24. Conserved domain of geminivirus movement protein (BL1 super family) was detected on the ORF. Blastp analysis of the remaining ORFs revealed that, there were no sequences showing significant similarity with these sequences.

Secondary structure of MP was predicted by SOPMA tool offered by Expasy. MP contained mostly random coils (51.95 and 50.89 per cent respectively) followed by extended strand, alpha helix and beta turn. Details of the output are given in Plate. 18.

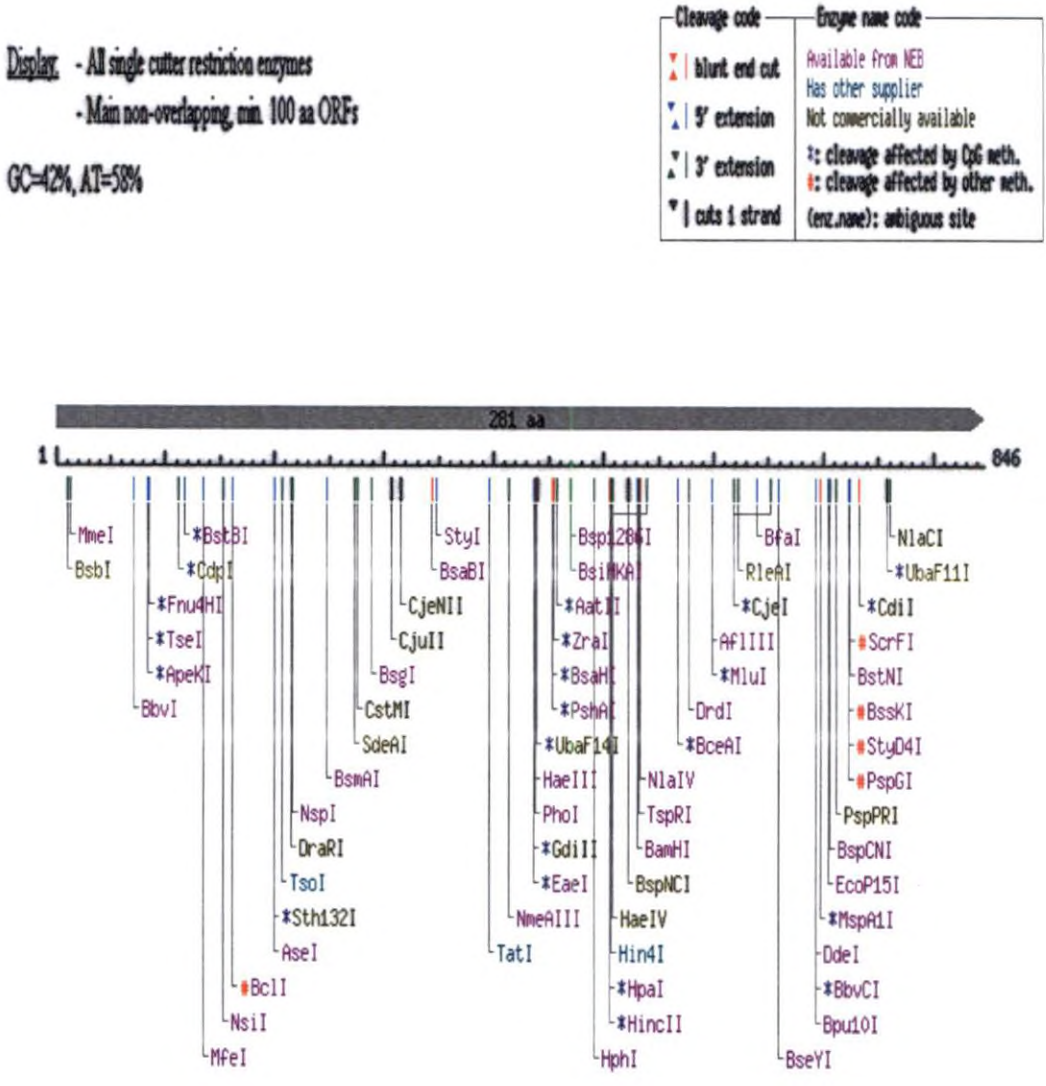


Plate. 17 Diagramatic representation of restriction analysis of MP gene

Table 23. Deduced amino acid composition (%) of CP and MP

Polarity		Amino acid (%)	Coat protein	Movement protein
Non-polar		Gly	5.08	6.41
		Ala	6.25	4.27
		Val	10.55	8.54
		Leu	3.52	6.76
		Ile	3.52	6.76
		Met	4.30	2.85
		Pro	5.08	4.27
		Phe	3.91	2.85
		Trp	1.17	1.07
Polar	Uncharged	Ser	5.86	10.32
		Thr	7.42	6.05
		Cys	1.95	1.78
		Tyr	5.08	2.85
		Asn	6.25	5.69
		Gln	2.34	2.85
	Basic	Lys	6.64	5.69
		Arg	10.16	6.41
		His	3.12	2.85
	Acidic	Asp	4.30	4.98
		Glu	3.52	6.76

Table 24. Sequences producing significant alignments with the ORF +2(413-1258) on blastp analysis of MP

Accession	Description	Maximum score	Query coverage	E value
AAQ91817.1	Movement protein[*Squash leaf curl China virus*-[Pumpkin Lucknow]]	396	100%	2e.139
ACX33151.1	Movement protein[*Squash leaf curl China virus*-India]	395	100%	3e.139
ADK78879.1	BC1 protein[*Squash leaf curl China virus*]	393	100%	3e.138
CAM94794.1	Movement protein[*Squash leaf curl China virus*]	391	100%	2e.137
ADF28570.1	Movement protein[*Squash leaf curl China virus*-Varanasi: Pumpkin: 08]]	390	100%	3e.137
YP 006445.1	Movement protein[*Squash leaf curl Philippines virus*]]	389	100%	1e.136
CAJ97389.1	BC1 protein[*Squash leaf curl China virus*]	389	100%	1e.136
YP 293687.1	BC1[*Squash leaf curl China virus*]	389	100%	1e.136
CAJ97391.1	BC1 protein[*Squash leaf curl China virus*]	389	100%	1e.136
CAQ86174.1	Movement protein[*Squash leaf curl China virus*]	387	100%	1e.135

Kyte and Doolittle hydropathy plot was used to characterize the hydrophobic character for predicting the membrane-spanning domains, potential antigenic sites and regions that are likely exposed on the protein's surface. The details of the plots are given in the Plate 19. The regions above the midpoint represent hydrophobic and below are hydrophilic. The hydrophilic regions may be exposed to the surface of proteins and generally serve as antigenic sites. There were no putative transmembrane regions present in the MP gene.

4.5.3 Phylogenetic analysis

Phylogenetic analysis of coat protein and movement protein genes was carried out by mega 5.05 software (Plate. 20 and 21). Phylogenetic trees were constructed by neighbor joining tree method giving thousand bootstrapping. Phylogenetic tree of CP and MP genes showed that PYVMV from Kerala is closely related to SLCCV: Coimbatore with maximum bootstrap score of 99. Coat protein gene of *Ageratum, Emilia* and okra formed into different groups in the phylogenetic tree indicating that they are not much related. Strains of *Tomato leaf curl New Delhi virus* also formed a different clad in both coat and movement protein analysis but with comparatively higher scores.

4.6 Validation of the technique

Validation was carried out with DNA of leaf samples from 15 pumpkin plants with yellow vein symptom and 4 healthy pumpkin samples collected from three different locations. DNA was isolated by CTAB method (Plate. 22a and b) and the quality of the DNA isolated was good. Quantity was assessed by Nanodrop spectrophotometry (Table 25). PCR amplification was carried out with all the 19 DNA samples using coat protein and movement protein gene specific primers.

```
         10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MSIRNDNNGLGVGGYIESDRVEYALTNDAAEVTLTFPSNFEQKISQLRNRCNKIDHVLLEYRSQVPINAV
eeccccceeeeetceecttthheeeecttcccceeeecccchhhhhhhhhhtccccccheeeehttcccccc
GHVVIEIHDNRLTEGDTKQAEFTIPIKCNCNLHYYSSTYFSVKDKNPWRVEYRVENTNVVNGVHFCKNLG
ceeeeeeccccccccccchheeeeeecccccceeeeeeeeeecccccceeeeeeccttccttcehhhhhh
KLKLSSAKHSTDVEFRAPRIEIQSKEFTVNDIDFWSVGSKPQTRRLVDGSRLNGHSSRSLRVPNLAIGPN
heeecccccccceecccccceeeecccccttteeeeeecccccchheeccccccccccccccccceecctt
ESWASRSEIGLGSVTSRPYKNLSGLDESAIDPGPSASQVGSITRDEITEIITKTVEQCNKSNVNAPLTKG
ccccccccccccccccchhhhhhccccccccccchhccccchhhhhhhhhhhhhhhhhttcccccccc
V
c

Sequence length :   281

SOPMA :
   Alpha helix      (Hh) :    49 is  17.44%
   3(10) helix      (Gg) :     0 is   0.00%
   Pi helix         (Ii) :     0 is   0.00%
   Beta bridge      (Bb) :     0 is   0.00%
   Extended strand  (Ee) :    70 is  24.91%
   Beta turn        (Tt) :    19 is   6.76%
   Bend region      (Ss) :     0 is   0.00%
   Random coil      (Cc) :   143 is  50.89%
   Ambigous states  (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

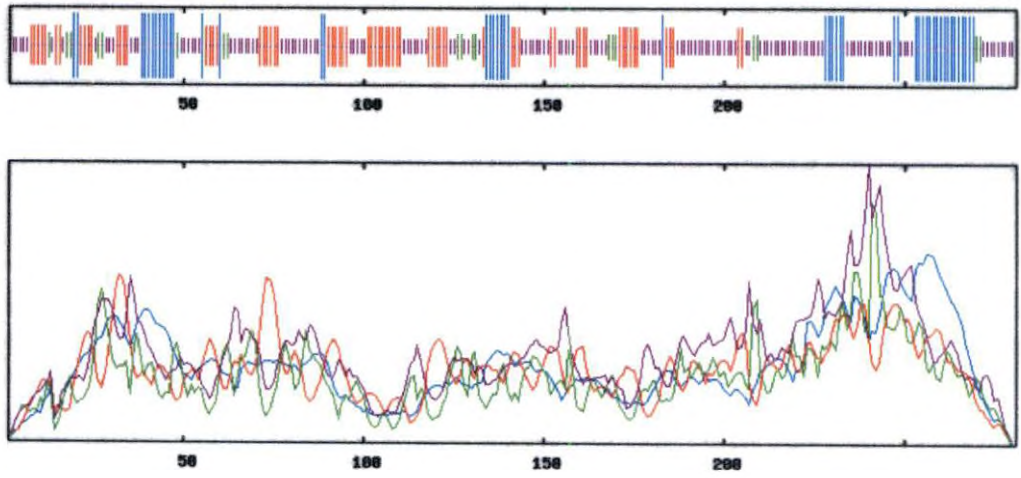


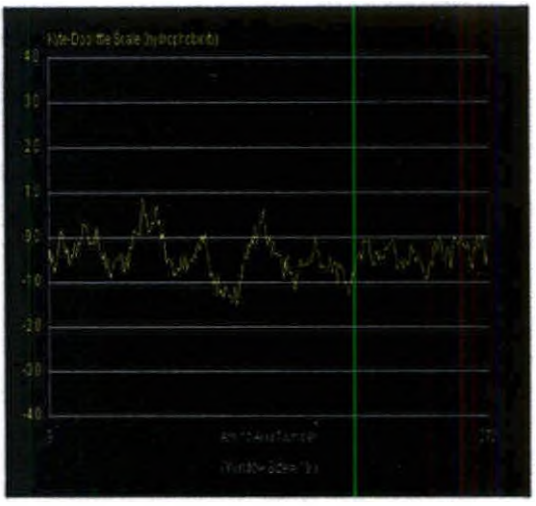
Plate. 18 Secondary structure prediction of deduced movement protein by SOPMA

Amplicons of expected size were obtained in all fifteen infected samples and no amplification with healthy samples (Plate 22 c and d). This indicated that these primers could be used in a diagnostic kit for detection of PYVMV.

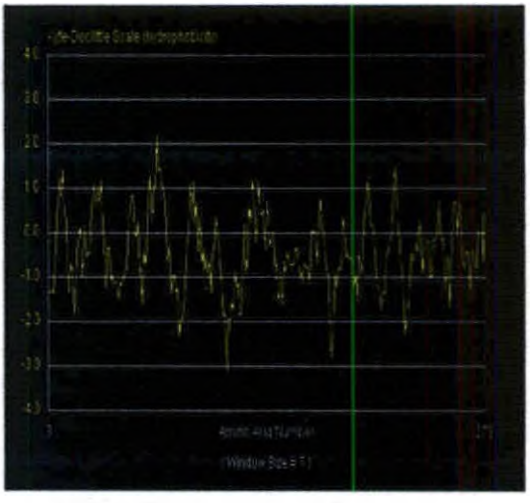
PCR amplification was carried out with apparently healthy leaf samples taken from infected field. Amplification was observed in symptomless leaves of symptom carrying plant (Plate 23b and c). This indicated that these primers could also be used to detect latent infection where no external symptoms are produced.

4.7 Detection of PYVMV in weeds and crop plants with yellow vein symptom

DNA was isolated from weed plants like *Emilia sonchifolia, Ageratum conyzoides, Synedrella nodiflora* and *Hibiscus surattensis* (Plate 24a) and also from vegetable crops like okra, ash gourd and bitter gourd. The quality of DNA isolated was good as all the samples gave intact band of DNA on agarose gel electrophoresis. The quantity was assessed by spectrophotometry (Table 26). PCR amplification was carried out with MP and CP specific primers using these samples and DNA from infected pumpkin leaves were kept as control (Plate 24b and c). No amplification was obtained in any of these samples except the control which indicated that none of these plants are acting as collateral hosts of the virus.

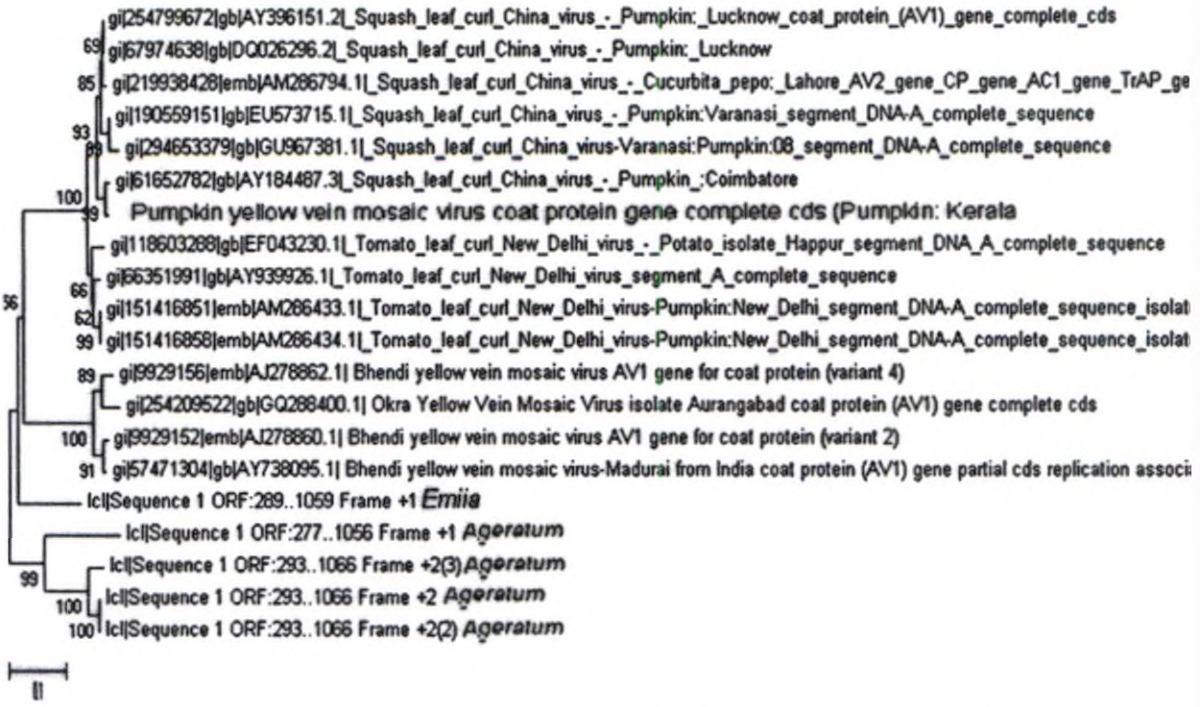


a. Movement protein (Window size=19)

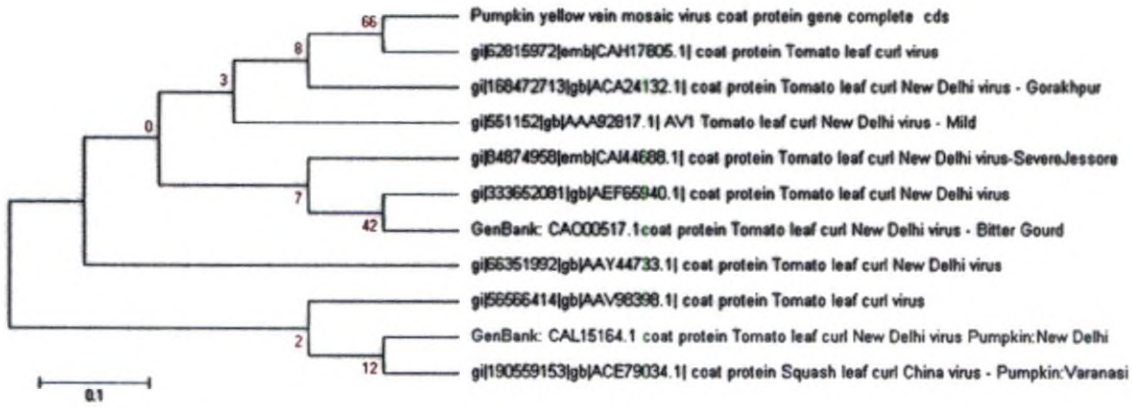


b. Movement protein (Window size=7)

Plate. 19 Kyte-Doolittle hydropathicity plots of MP. The average hydrophobicity of each amino acid residue was calculated using the algorithm over window size of 19 and 7 and plotted as a function of window position

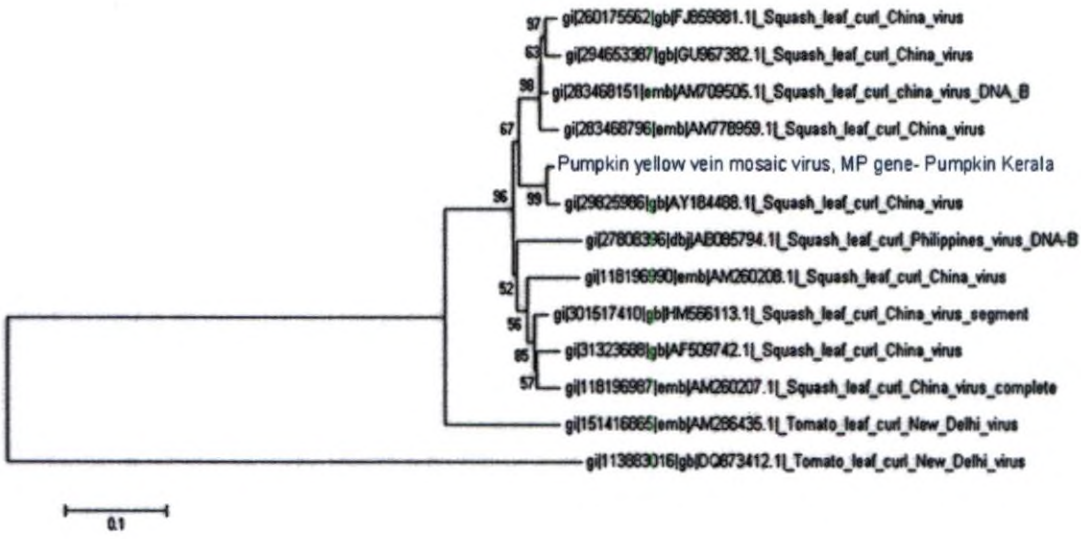


a. The phylogenetic tree of coat protein gene inferred using the Neighbor-Joining method.

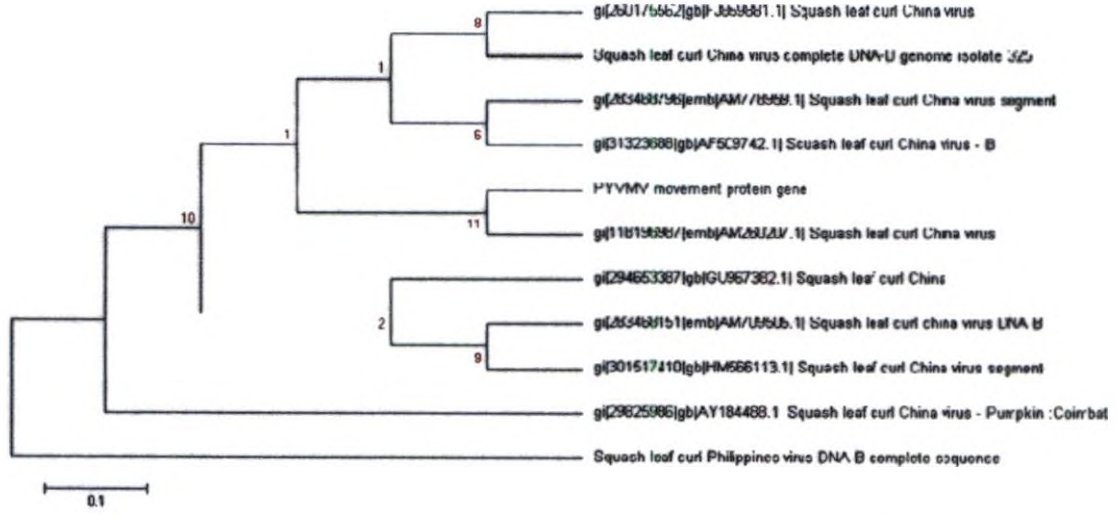


b. The phylogenetic tree of coat protein (deduced amino acid sequences) inferred using the Neighbor-Joining method.

Plate. 20 Phylogenetic analysis of coat protein gene



a. The phylogenetic tree of movement protein gene inferred using the Neighbor-Joining method.



b. The phylogenetic tree of movement protein (deduced amino acid sequences) inferred using the Neighbor-Joining method.

Plate. 21 Phylogenetic analysis of movement protein gene

Table 25. Quantity of total DNA extracted for validation of the technique

Sl. No	Samples	Optical density ratio (260/280)	Quantity of ds DNA(ng/µl)
1	PS1	1.92	2461.7
2	PS2	1.95	2234.2
3	PS3	1.99	2029.3
4	PS4	2.03	1838.3
5	PS5	1.93	1034.9
6	PS6	1.90	2391.0
7	PS7	1.95	1947.9
8	PS8	1.97	1914.6
9	PS9	1.89	1943.6
10	PS10	1.83	1450.6
11	PS11	1.91	1356.3
12	PS12	2.00	1974.9
13	PS13	1.88	1129.3
14	PS14	1.86	1883.2
15	PS15	1.85	1730.6
16	PH1	1.84	2491.5
17	PH2	1.95	1929.4
18	PH3	1.96	2255.8
19	PH4	1.94	2560.1

PS: Pumpkin leaves with yellow vein mosaic symptom

PH: Healthy pumpkin leaves

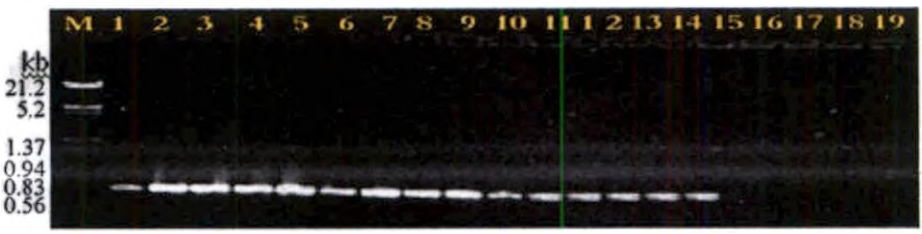
a. Genomic DNA isolated from PYVM infected samples

Lane M: λDNA *Eco*R1/*Hind*lll double digest; 1-15: Infected pumpkin

Lane M: λDNA *Eco*R1/*Hind*lll double digest

Lane 1- 4: Healthy pumpkin

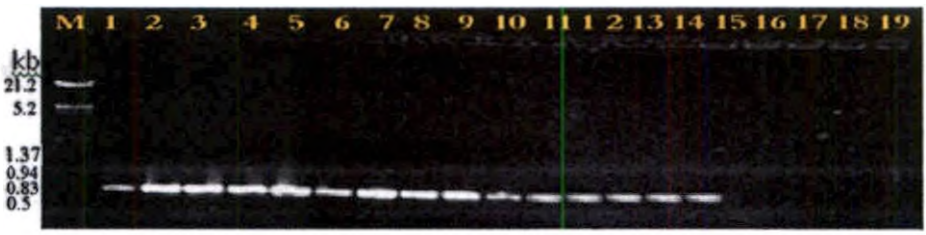
b. Genomic DNA isolated from healthy pumpkin



c. PYVMV detected by PCR amplification of MP gene

Lane M: λDNA *Eco*R1/*Hind*lll double digest; 1-15: Infected pumpkin;

16-19: Healthy pumpkin



d. PYVMV detected by PCR amplification of CP gene

Lane M: λDNA *Eco*R1/*Hind*lll double digest; 1-15: Infected pumpkin;

16-19: Healthy pumpkin

Plate. 22 Validation of the MP and CP gene specific primers using **infected and healthy pumpkin samples**

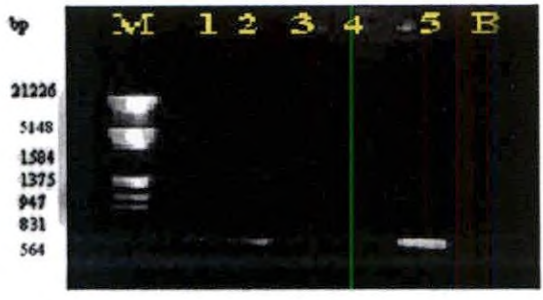
Table 26. Quantity of total DNA extracted from weeds and vegetable crops showing yellow vein symptom

Sl. No	Samples	Optical density ratio (260/280)	Quantity of ds DNA(ng/µl)
1	*Ageratumconyzoides*	1.82	2161.3
2	*Emilia sonchifolia*	1.91	2344.2
3	*Synedrella nodiflora*	1.89	2232.9
4	*Hibiscus surattensis*	1.92	1630.5
5	Ash gourd	1.93	2341.9
6	Bitter gourd	1.96	2566.0
7	Okra	1.95	947.9

a. Genomic DNA

Lane M: λ DNA *Eco*RI/*Hind*lll double digest

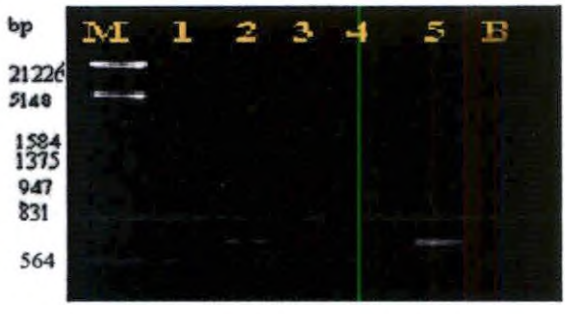
Lane 1- 4: Apparently healthy pumpkin



b. PCR amplification of CP gene

M: λ DNA *Eco*R1/*Hind*lll double digest

1-4: Apparently healthy samples, 5: Infected pumpkin sample, B: Blank

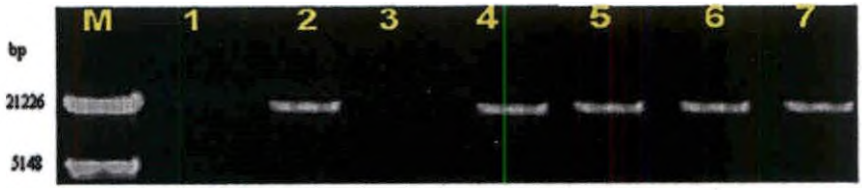


c. PCR amplification of MP gene

M: λ DNA *Eco*R1/*Hind*lll double digest

Lane 1-4: Apparently healthy samples, 5: Infected pumpkin sample, B: Blank

Plate. 23 PCR amplification of PYVMV in apparently healthy pumpkin

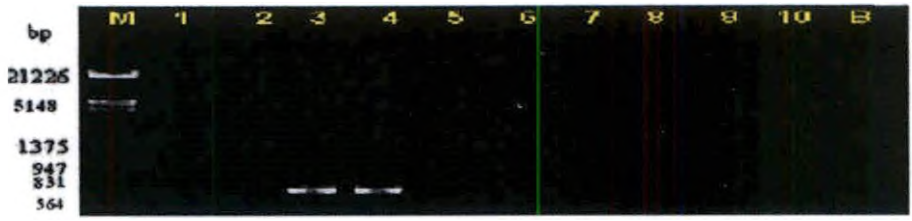


a. Genomic DNA

Lane M: λDNA *Eco*RI/*Hind*III double digest

Lane 1- *Emilia*; 2- *Synedrella*; 3- *Ageratum*; 4-*Hibiscus*

5- Okra; 6- Ash gourd; 7- Bitter gourd



b. PCR amplification of CP gene

Lane M: λDNA *Eco*RI/*Hind*III double digest; 1- *Emilia*; 2- *Synedrella*

3 & 4- Infected pumpkin; 5- *Ageratum*; 6- *Hibiscus*; 7- Okra; 8- Ash gourd;

9- Bitter gourd; 10- Healthy pumpkin; B- Blank

b. PCR amplification of MP gene

Lane M: λDNA *Eco*RI/*Hind*III double digest; 1- Infected pumpkin; 2-*Emilia*;

3- *Synedrella*; 4- *Ageratum*; 5- *Hibiscus*, 6- Okra; 7- Ash gourd; 8- Bitter gourd;

9-Healthy pumpkin; B- Blank

Plate. 24 PCR amplification of PYVMV in weeds and vegetable crops showing yellow vein symptoms

Discussion

5. DISCUSSION

Pumpkin (*Cucurbita moschata*) belongs to Cucurbitaceae family and is considered as native to Central America. The term pumpkin also refers to the species *Cucurbita pepo*, *Cucurbita mixta* and *Cucurbita maxima*. The biggest producers of pumpkins include the United States, Mexico, India and China. It is a warm weather crop and cultivated throughout the year. The plant is monoecious and almost all parts are edible, including the fleshy shell, the seeds, the leaves and even the flowers.

The crop is infected by a number of virus diseases mainly pumpkin mosaic (PM) and pumpkin yellow vein mosaic disease (PYVMD). Among these PYVMD has emerged as a major threat during the recent years and has been reported throughout India. The disease is caused by a geminivirus belonging to genus begomovirus and transmitted by the whitefly vector but not by aphids or by sap inoculation (Capoor and Ahmed, 1975; Jayashree *et al.*, 1999). The incidence of PYVMD and the associated yield loss is found maximum during February to May, when the vector population is at its peak. Infected plants exhibited yellowing of veins on younger leaves which developed into mosaic patches in the later stages of infection. Affected plants were stunted and flowers dropped prematurely. Although attempts have been made to characterize the causal agent based on its biological characteristics, information on molecular biology of the causal organism is scanty. Two species of geminiviruses causing PYVMD have been reported in India, one from northern India and the other from southern India, both are bipartite (Muniappa *et al.*, 2003 and Maruthi *et al.*, 2007). In the past decade both the prevalence and distribution of the whitefly transmitted viruses have increased the disease incidence and the impact has been devastating.

Accurate disease diagnosis and precise identification of any pathogen involved is an essential prerequisite for understanding plant disease and controlling them

effectively. The lack of rapid, accurate and reliable means by which plant pathogens can be detected and identified has been one of the main limitations in integrated disease management. For instance, the major control measure for viral diseases is proper vector control and the vectors usually vary according to virus species. In pumpkin, only PYVMV is transmitted by whitefly whereas, the commonly occurring PM and cucumber mosaic are transmitted by aphids. Hence timely and rapid detection can ensure better control by managing the vector population using suitable chemicals. In the advanced stages, it is difficult to identify the disease as the entire leaf turns yellow. The confusing symptoms may even lead to a wrong diagnosis of the disease.

Traditional methods of identifying plant pathogens can be slow and inconclusive and this has prompted the search of alternative diagnostic techniques. Traditionally, characterization of the pathogen by cultural and biochemical test has been employed to identify potential pathogens. This is a relatively slow process, often requires a skilled taxonomist to reliably identify the pathogen and is not possible for virus as it is not culturable.

A great variety of methods have been developed that permit differentiation of viral pathogens. These methods, initially based solely on identifying the distinct biological characteristics of different viruses, were soon supplemented with methods based on light or electron microscopy and serology. Subsequently Enzyme Linked Immunosorbant Assay (ELISA) was also used for detection of plant viruses. Many commercial plant virus detection kits have been developed based on ELISA. However, detection of geminiviruses through ELISA has certain limitations. Although it is a sensitive method for virus detection, it has a low specificity. Development of specific antibodies and low antigenicity of virus increase the limitation boundary of ELISA method for geminivirus detection. Direct ELISA requires the labeling of every antibody to be used, which can be a time consuming and expensive proposition. Moreover, one major problem is the cross-reaction of the antibodies produced against

one geminivirus coat protein with other geminiviruses. This cross-reaction is due to high level of homology between different geminiviruses which result in the presence of many shared epitopes. Therefore, it has become difficult to distinguish between two closely related strains or species. It has also been observed that ELISA is effective in detecting virus in purified preparations but not in crude extract of infected plants and whitefly transmitted geminivirus (WTG) purification is difficult due to the peculiar physical and chemical properties of the virion particle. Also, the virion is poorly immunogenic and the capsid protein of WTG are not antigen wise distinct with available polyclonal and monoclonal antisera preparations (Wyat and Brown, 1996). Therefore, modifications were introduced in ELISA to enhance the sensitivity of detection over a period of time. Even though serological methods could detect the virus, the analysis cannot provide conclusive data about the nature of the virus. Hence serological methods were eventually replaced by nucleic acid based methods which are more sensitive and specific. Certain nucleic acid based techniques have the potential to detect single nucleotide differences and may be useful in distinguishing different strains of a virus (Miller and Martin, 1988).

Nucleic acid hybridization method fulfilled some of the requirements of geminivirus detection method. Different kinds of hybridization methods have been developed like squash blot, dot blot and southern blot for the geminivirus detection. All of them are more sensitive and specific than ELISA techniques and both qualitative and quantitative assays can be performed with single technique. Several geminiviruses which were not detected by ELISA were identified and characterized by nucleic acid hybridization methods. Apart from showing good sensitivity and specificity nucleic acid hybridization techniques also have some limitations. Since many geminivirus genomes share common conserved nucleotide sequences; during hybridization non specific binding may occur and also hybridization techniques is not sensitive for very low level (10 ng) viral genome present in plant samples (Kushwaha *et al.*, 2010).

In order to detect such an ultra low level viral genome, PCR came in to the trend which revolutionised the geminivirus detection methods. PCR based technology shows great promise to solve this problem by accurate and rapid diagnosis. PCR and its direct and indirect use in identifying by amplification of gene products are vital to both research and applied fields (Crowther, 1998). The major advantage of PCR as a detection method is its rapidity, sensitivity and robustness. It has high potential to detect very small number of copies of a gene or pathogen. It offers great analytical sensitivity much higher than conventional probes. Since many geminiviruses have conserved regions in the genome, universal primers designed from these regions could be used to detect any geminivirus infection in the sample. Primers designed from conserved regions have been used for identification and characterization of several geminiviruses. Different kinds of PCR like multiplex PCR, nested PCR, RT- PCR, rolling circle amplification (RA) have also been used for detection and characterization of geminiviruses.

Keeping all these facts in view, the present study was taken up with the objectives of molecular characterization of geminivirus causing yellow vein mosaic in pumpkin and developing a molecular diagnostic tool based on PCR for detection of the virus.

5.1 DNA isolation

Isolation of good quality DNA is a prerequisite for performing PCR. Most research workers has reported CTAB method for extraction of the total plant DNA from infected samples (Mansoor *et al.,* 1998; Rouhibaksh *et al.,* 2008; Rajinimala *et al.,* 2009; Tariq *et al.,* 2010). In the present study the total plant DNA from PYVMD infected samples were isolated by the method described by Rogers and Benedich (1994). DNA was extracted from tender leaves collected early in the morning so as to

minimize the interference of polyphenols. Babu (2000) reported that the quality and the quantity of DNA isolated was best when tender leaves were used as compared to matured and half matured leaves. The use of tender leaves for DNA isolation in pumpkin leaves has been reported by Singh *et al.* (2009), Jaiswal *et al.* (2010) and Narayana *et al.* (2010). The amount of leaf tissue used was 1g, as too much sample might increase the viscosity of DNA precipitated and lead to shearing of genomic DNA.

Excess liquid nitrogen was used for homogenization of the leaf tissue, and this could be the reason for obtaining intact DNA. Liquid nitrogen freezes the tissues and helps to prevent degradation of nucleic acid. Similar observations have been made by Sharma *et al.* (2002) in sorghum, chickpea, wheat and soybean, Lodhi *et al.* (1994) in grapevine cultivars and Padmalatha and Prasad (2006) in medicinal and aromatic plants.

The presence of high amounts of polyphenols was a problem during DNA extraction from pumpkin leaves and could reduce the yield and purity of DNA by binding covalently with the extracted DNA thus making it useless for most of the research applications. The problem was overcome by the addition of sodium metabisulphate, β-mercaptoethanol and Poly Vinyl Pyrrolidine (PVP) along with the extraction buffer. β-mercaptoethanol disrupts the protein disulphide bonds and is thereby capable of initiating protein degradation. PVP helps in the removal of polyphenolic compounds by binding with hydrogen bonds and forming a complex. Nesbit *et al.* (1995) and Padmalatha and Prasad (2006) have also reported the use of β-mercaptoethanol and PVP for overcoming phenolic contamination in plant samples.

CTAB (cetyl trimethyl ammonium bromide), a detergent used in the extraction buffer in Rogers and Benedict protocol has dual functions. It helps in the disruption of the cell membrane thereby releasing nucleic acids into the extraction buffer; while on

the other hand, it prevents co-precipitation of polysaccharides with nucleic acid by acting as a selective precipitant of nucleic acids. By these actions CTAB must have helped in the recovery of relatively pure DNA in the present study. Sharma *et al.* (2002) reported that nucleic acids form tight complexes with polysaccharides creating a gelatinous pellet that contains embedded DNA, and polysaccharides also co-precipitate with DNA after alcohol addition during DNA isolation leading to viscous solutions.

Extraction buffer contains EDTA, which could effectively chelate Mg^{2+} ions and mediate aggregation of nucleic acid. Isopropanol (0.6 volume) was used for initial precipitation of DNA at low temperature (-20°C) and two volumes of ethanol was used for final precipitation. The pellet was dissolved in TE buffer for long term storage. EDTA present in TE buffer could chelate and remove Mg^{2+} ions, which was required for nuclease activity. The extracted DNA samples contained RNA as contaminant. Large amount of RNA in the DNA sample can chelate Mg^{2+} ions and reduce the yield of the PCR. Hence, RNAse treatment was performed to remove RNA contamination.

DNA isolation from okra was carried out following the same protocol with minor modification. Okra leaves contain high amount of mucilage and polyphenolic compounds which make the DNA isolation difficult. The major obstacle for DNA extraction from mucilaginous plants is the presence of gelling polysaccharides and these polysaccharides co-precipitate with DNA. Hence, it is necessary to avoid as far as possible by keeping them in solution (Dellaporta *et al.,* 1983). Polysaccharides prevent access to enzymes for PCR and some workers advocate use of hydrolytic enzyme (Rether *et al.,* 1993) or ion exchange resins (Marchal and Drouard, 1995) for removal of polysaccharides from nucleic acids. Barnwell *et al.* (1998) reported that the step wise increase in the concentration of CTAB could remove polysaccharides and polyphenols effectively. For extraction of genomic DNA from okra, the concentration of NaCl was made to 3 per cent in the extraction buffer and 1 per cent in

10% CTAB solution. The quantity of buffer used was increased as it diluted the mucilage and helped in the removal.

The extracted DNA was analyzed by agarose gel electrophoresis using 0.8 percent agarose gel. Good quality DNA was obtained as single sharp bands and the RNA contamination was high which was removed by RNAse treatment. The quantity of the DNA obtained was estimated by spectrophotometry. The quantity of the DNA isolated ranged from 790 to 4917 ng/µl. The ratio of absorbance at 260nm to absorbance at 280nm ranged from 1.8 to 2 which indicated that the DNA was of good quality. The quantity of DNA recovered for okra was less compared to pumpkin and other plant samples which might be due to the entrapment of the DNA in the mucilage. Since one PCR reaction requires only 25-100ng template, the DNA was found to be sufficient for further PCR reactions.

5.2 PCR based detection

The first step towards minimizing the damages caused by the virus is timely detection. Biological virus indexing and serological methods like enzyme-linked immunosorbent assay are performed for detection. But, identification of whitefly-transmitted geminiviruses by traditional methods has been difficult and thus reliable detection of PYVMV remains to be a challenge. Serology alone is not suitable, since high viral titre is difficult to prepare (Brown *et al.*, 2001).

Other approaches such as polymerase chain reaction (PCR) amplification of specific genomic regions and DNA sequencing offer attractive opportunities for detection, identification and characterization of begomoviruses. Using PCR the time required for detection is less than 24 h whereas conventional methods require several days to confirm the pathogen (Lopez *et al.,* 2003). In addition, geminiviral genome is

more suitable for PCR amplification because it replicates via a double-stranded circular DNA intermediate (replicate form), which can serve as template for DNA amplification. In recent years, PCR-based rapid detection methods have been developed for several viruses (Khan *et al.*, 2006; Chattopadhyay *et al.*, 2008).

5.3 Primer designing

Proper amplification relies on the set of two short synthetic oligonucleotide primers. Good test performance requires that the primers bind only to the corresponding nucleotide sequence of the viral genome. Thus the specificity of the assay and accuracy of the results depend upon the design of PCR primers.

Geminiviruses are one group of plant viruses for which a significant database with large number of nucleotide and amino acid sequences is being developed, especially for coat protein and movement protein gene sequences. They have highly conserved regions in many parts of their genome. The coat protein gene is more conserved than the remaining part of the genome (Padidam *et al.,* 1995). However, a short N-terminal region (60-70 amino acids) of the CP is more variable than the rest of the CP sequence and is a close representation of the genome. PCR primers based on conserved sequences can be used to clone and sequence the N-terminal sequences of the CP of the geminiviruses; this sequence can be used to classify a particular virus isolate. Movement protein gene also posseses conserved regions but was found to be less compared to coat protein gene.

From the previous reports it was found that the yellow vein mosaic in pumpkin is caused by a bipartite geminivirus in India. The virus obtained from South India and North India showed maximum similarity to Tomato leaf curl New Delhi Virus (ToLCNDV) and Squash leaf curl china virus (SLCCV) respectively (Muniyappa *et*

al., 2003; 2005; Maruthi *et al.*, 2005; Singh *et al.*, 2009). Only a few accessions of viruses causing yellow vein mosaic in pumpkin were available in the NCBI data base. Hence, for primer designing, sequences of geminiviruses reported in other vegetable crops like tomato were also used, considering the fact that they might have particular regions in their genome which remain conserved.

The conserved regions in the sequences were located by performing multiple sequence alignment (MSA). It is a tool that determines the levels of homology, and hence relatedness between a number of globally related sequences. When designing primers, the conservation of the region, the degeneracy of the genetic code and parameters of PCR reaction must be considered. In MSA, sequences are aligned optimally by bringing together the greatest number of similar characters into the same column of alignment. For primer designing, conserved regions in the nucleotide level are considered but when sufficient nucleotide data is not available, amino acid sequences can be considered, accounting the degeneracy of genetic code. In this study, as plenty of nucleotide sequences are available. Hence MSA at amino level was not considered for primer designing.

For designing coat protein gene specific primers, sequences of DNA A were used, as it encodes the protein. The DNA A was found to be highly conserved, especially within and nearby the coat protein gene ORF. Therefore, the primer pair was designed without degeneracy.

MP is encoded by DNA B and the primers were designed to amplify ~700bp within the MP gene. Movement protein gene was comparatively less conserved so that degenerate primers were designed. Degenerate primers have a number of options at several positions in the sequences so as to allow annealing and amplification of the

gene in related organisms. The forward and reverse primers had degeneracy at one and four nucleotide positions respectively.

While designing the primers, care was taken to avoid complimentarity between the forward and reverse primers. Regions having high G+C content were selected. The choice of primer annealing temperature is the most critical factor in designing specific primers with high amount of specificity. Primers were designed in such a way that the annealing temperature comes between 55-60°C. The length of the primer sequences for coat protein and movement protein genes were 20bp and 21bp respectively for both the FP and RP, the annealing temperature should be similar for getting amplification. If the temperature is too high no annealing occurs and if too low, non specific annealing will increase dramatically. Therefore, the annealing temperature must be determined and optimized empirically. Optimal annealing temperature is often approximately 3-5 °C higher than the T_m of the primers. The primers were synthesized at Sigma Aldrich Chemicals Pvt. Ltd., Bangalore.

The primers were designed in such a way to amplify the full length CP gene (~900bp) and approximately 700bp within in the movement protein gene. The expected amplicon size for the primers CPF and CPR was 896bp and 678bp for the primers MPF and MPR.

5.4 Amplification of CP and MP genes

5.4.1 Standardization of PCR conditions

PCR method has high potential for detecting minute quantities of virus in the plant tissue (Spiegel *et al.*, 1993). In the present study, Initial optimization of the PCR conditions was carried out, since the PCR parameters are very crucial for specific

amplification (Park and Kohel, 1992). Annealing temperature is the important factor governing amplification in PCR. If it is very low, nonspecific amplification will take place. Spurious product may get amplified, due to nonspecific amplification. If annealing temperature is very high annealing will not take place or too little product will be produced. Therefore twelve different annealing temperatures was tried using gradient PCR by giving a temperature range of 55±10°C. CP primers gave amplifications at 46.4, 55.8, 58.5, 61.0, 63.1, 64.7 and 65.0 °C. The MP primers produced amplifications at 46.4, 55.8, 58.5, 61.0 and 63.1°C. CP and MP primers gave the best band with good concentration and expected size at 55.8 and 58.5°C respectively. The quality and quantity of the amplicons were less in all other temperatures. Hence for further experiments these temperatures were used as the annealing temperature.

Primer and template concentrations are also important for specific amplification. If the primer concentration is low the quantity of the amplified product may be less and higher primer concentration may promote misspriming and accumulation of non specific product. There are also chances of primer dimer formation especially when complimentarity exists between the primers. During the PCR cycle, the template gets multiplied in the range of 2^n , where n is number of cycles. The primer and template dilutions of 1/10 and 1/100 dilutions respectively were tried which gave good amplifications. Similar observations were reported by Singh *et al.* (2009). Hence the same concentrations were used for further PCR amplifications also.

5.4.2 PCR amplification

PCR amplification with CPF and CPR primers yielded amplicons of ~900bp from the infected samples. The concentration of the amplified products was high

indicating higher concentration of the viral nucleic acids in the test samples. MPF and MPR primers yielded amplicons of ~700bp. Both the primer sets yielded single intact bands (Plate 6e and f). No primer dimer formation was observed in any of the reactions. As the amplicons are of small size, it can be easy amplified. Hence these primers could be used in developing molecular diagnostic kit for rapid detection of the pathogen.

5.5 Transformation and cloning

5.5.1 Preparation and screening of competent cells

Competent cells were prepare from *E. coli* JM109 strain as per the procedure described on section 3.4.2.1 were highly competent and showed high degree of transformation efficiency when transformed with the plasmid pUC18, containing ampicillin resistance marker. *E. coli* cells alone cannot grow in ampicillin containing media, since they had no resistance encoding sequence. But all the competent cells cells harbouring the plasmids can grow in the media. The transformation efficiency was calculated as 1 x 10^7 cfu/ µl. The survival rate was high and thus the competent cells prepared were found to be efficient for other cloning works.

5.5.2 Transformation of DNA

The cloning vector used was pGEM-T Easy with a size of 3kb. It contains T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase (Plate 2). Insertional inactivation of the α- peptide allows the recombinant clones to be directly identified by the colour screening of indicator plates. The multiple cloning region of the vector includes

restriction sites conveniently arranged for use. The sites allow release of the insert by digestion with a single restriction enzyme.

The eluted DNA was inserted in the vector and the ligated product was transferred to *E. coli* JM109 cells using the heat shock at 42°C. The cells containing the ligated product could later be picked up from the media containing 5-bromo 4-chloro 3-indolyl β-D galactoside (X-gal) and isopropyl thiogalactoside (IPTG) based on blue white screening. The pGEMT vector contained polycloning sites inside the β-galactosidase gene encoding region. Insertion of the new sequence disrupts the reading frame of the gene. The bacterial cell and the vector together provided the complete protein as a result of α-complementation (Ullmann *et al.,* 1967). The colonies which have not taken up the plasmid can further utilize the substrate and appear as blue colonies on X-gal chromogenic substrate (Horwitz *et al.,* 1964). All the transformed colonies harbouring the recombinant plasmid appeared in white colour due to disruption of α-complementation.

5.5.3 Confirmation of recombination

Some times the cells will appear as white colonies if the X-gal and IPTG are not evenly spread or due to the delay in the development of the blue colour. Therefore, the white colonies obtained were again plated on LB layered with X-gal and IPTG. In the second plate, some of the colonies appeared as white in the first plate turned to blue.

The presence of the insert in the cells was confirmed by performing colony PCR. For this DNA from white colonies were used as template. The universal primers T7 and SP6 were used as the forward and reverse primers respectively. The CP and MP specific primers yielded amplicons of ~900 and ~700 bp respectively. The blue

colonies which were used as control yielded amplicons of small size compared to white colonies. This might be produced by the annealing of the primers at the T7 and SP6 region present in the vector.

5.6 Amplification of full length movement protein gene

The ORF of the movement protein obtained from the first set of primers were partial. From the NCBI data base it was found that the full length ORF of movement proteins of all the related sequences were 846bp. It is better to take full length MP gene to get significant results from sequence analysis. So primers MP1F and MP1R were designed to get the full length MP gene. The expected amplicon size was 1363bp. The annealing temperature was standardized at 55.8°C by gradient PCR. The amplicon obtained was ~1300bp. The PCR product was directly sequenced by ABI prism sequence analyser. The sequence analysis revealed the presence of full length ORF of MP gene.

5.7 *In silico* sequence analysis

5.7.1 Nucleotide sequence analysis of coat protein gene

When the sequences were subjected to BLAST homology search, the CP gene sequence showed maximum identity of 98 per cent with *Squash leaf curl China virus* [Pumpkim: Coimbatore] and then with *Squash leaf curl China virus* [Pumpkin: Lucknow] (97 per cent). It also showed similarity to *Tomato leaf curl New Delhi virus* (93 per cent). The high level of identity of the sequences with the sequences in the database might be due to the conserved regions in these genes.

6 ORFs were detected in the CP gene sequence corresponding to six possible translation frames (+1, +2, +3 and -1, -2,-3). The largest ORF was 771 bp long in the +3 reading frame and this correspond to the full length CP gene. On Blastp analysis it showed maximum similarity to coat protein of *Squash leaf curl China virus* [Pumpkim: Varanasi] and Tomato leaf curl New Delhi virus with the highest score of 536. The ORFs of -2 (33-518) and -3 (770-889) showed maximum similarity to AC5 (score: 320) and replication enhancer proteins (77) respectively. The three conserved domains located were geminivirus coat protein/BR1, AL4/5 and AL3 family.

The nitrogen base composition was typical to prokaryotic genes with high A+T and low G+C content for CP gene. The A+T content of CP gene was 54.9 per cent.

Restriction analysis of the sequence revealed the cleavage sites for different enzymes. There were 64 restriction sites for various restriction enzymes in CP gene.

The evolutionary relationship of the CP with the other selected sequences was drawn by constructing phylogenetic tree. Phylogram is the branching diagram assumed to be an estimate of phylogeny, branch lengths are proportional to the amount of inferred evolutionary change. PYVMV:Kerala and SLCCV: Coimbatore fell into the same clusture (Plate 21). In CP gene analysis, sequences of geminiviruses causing yellow vein mosaic disease in okra, *Emilia* and *Ageratum* were also included to find the extent of their relationship between these sequences. They formed clusters in the tree according to their geographical origin of isolation with distinct branches for monopartite and bipartite viruses indicating the sequence difference between them. Padidam *et al.,* (1995) also reported the same pattern of clustering and they suggested that, any new virus isolate having more than 90 per cent sequence identity to a previously characterized virus genome should be called as a strain of that particular virus species. Hence, PYVMV from Kerala can be considered as a strain of SLCCV.

The analysis of the amino composition indicates that the CP is rich in non polar amino acid valine (10.55 per cent) followed by basic amino acid arginine (10.16 per cent).

Secondary structure prediction showed the proportion of different structures viz., alpha helix, beta sheet and random coils. The CP gene contains 19.14 per cent alpha helix, 4.3 per cent beta turn and 51.95 per cent random coils. Random coils were most abundant and it usually represents parts of protein that lacks secondary structure. The α helices and β sheets formed rest of the structural backbone. Hence the 3D structure of the CP would be helical in shape.

Sequences were subjected to Kyte and Doolittle (1982) hydropathy plot analysis (Plate 19). No putative transmembrane region could be detected. Positive hydropathy scores indicate increased hydrophobicity and negative values show an increase in hydrophilic amino acids and the putative antigenic sites.

5.7.2 Nucleotide sequence analysis of movement protein gene

When the sequences were subjected to BLAST homology search, the MP gene sequences also showed maximum similarity to *Squash leaf curl China virus* [Pumpkim: Coimbatore] (98 per cent) and *Tomato leaf curl New Delhi virus* (83 per cent). The high level of identity of the sequences with the sequences in the database might be due to the conserved regions in these genes.

There were two open reading frames in the sequence, Pumpkin yellow vein mosaic virus movement protein gene [Pumpkin: Kerala]. The longest ORF was of 567 bases long. On Blastp analysis, it showed maximum similarity to *Squash leaf curl China virus* [Pumpkim: Varanasi] and *Tomato leaf curl New Delhi virus* with the

highest score of 536. *Pumpkin yellow vein mosaic virus* complete cds had five open reading frames with the longest being 846 bases length and on blastp analysis with maximum similarity to SLCCV- Pumpkin: Lucknow with the highest score of 396.

The nitrogen base composition was typical to prokaryotic genes with high A+T and low G+C content. The A+T and G+C content of MP gene were 57.7 per cent and 42.3 per cent respectively.

Restriction analysis of the full length MP gene sequence revealed the cleavage site for different enzymes. There were 72 restriction sites for various restriction enzymes within the sequence.

The evolutionary relationship and MP sequences with other selected sequences were drawn by constructing phylogenetic tree. Phylogram is the branching diagram assumed to be an estimate of phylogeny, branch lengths are proportional to the amount of inferred evolutionary change. PYVMV:Kerala and SLCCV: Coimbatore fell into the same clusture.

The analysis of the amino composition revealed that MP is rich in polar amino acid Serine (10.32 per cent) followed by non-polar amino acid Valine (8.54 per cent). Tryptophan (1.07 per cent) is the least abundant amino acid present in MP.

The secondary structure prediction showed the proportion of different structures viz., alpha helix, beta sheet and random coil. The MP gene contains 17.44 per cent alpha helix, 6.76 per cent beta turn and 50.89 per cent random coils. Random coils were most abundant and it usually represents parts of protein that lacks secondary structure. The α helices and β sheets form rest of the structural backbone. Hence the 3D structure of the MP would be helical in shape.

Sequences were subjected to Kyte and Doolittle (1982) hydropathy plot analysis (Plate 20). No putative transmembrane region could be detected as all the aminocaids had score below 1.6. Positive hydropathy scores indicate increases hydrophobicity and negative values show an increase in hydrophilic amino acids and the putative antigenic sites.

The sequence analysis shows that it is possible to classify geminiviruses based on the sequence comparisons, and that a short region of the genome is sufficient to classify an isolate. Highly conserved PCR primers can be used to clone and sequence this short region and to classify a new virus isolate.

5.8 Validation of the technique

Validation of the PCR detection of pumpkin yellow vein mosaic virus was carried out using 15 infected and 4 healthy pumpkin samples. PCR amplification with the CP and MP gene specific primers yielded amplicons of expected size in all the 15 infected samples and no amplication was obtained in the healthy sample indicating the absence of viral nucleic acid. This also shows that the amplification is reproducible and the technique can be used for the easy and rapid detection of the virus.

In PCR amplification with apparently healthy leaves from infected field, one of the four samples produced amplification. This leaf sample was collected from an infected plant which also contained leaves with the characteristic symptoms. The presence of amplicons indicated that the primer pairs can detect the virus even without or before the development of the symptom. Hence this kit can be used to detect latent infection where no external symptoms will be produced and also to identify the symptomless carriers.

5.9 PCR amplification in other plants

Numerous common weeds were found infected with geminiviruses (Brown and bird, 1992), but little is known of their relationship with crop infecting geminiviruses. These plants may act as collateral hosts for the virus and they must be identified and eradicated from the premises of the field. In an attempt to identify the collateral host of PYVMV, PCR amplification was also performed in weed plants, *Ageratum conyzoides, Emilia sonchifolia, Synedrella nodiflora,* and *Hibiscus surattensis* showing typical yellow vein symptom. Okra and ash gourd with yellow vein symptom and bitter gourd infected with distortion mosaic were also used for PCR amplification. None of these samples produced amplification. Reports on *Ageratum* and okra revealed that the yellow vein mosaic in these plants are caused by monopartite begomovirus with only DNA A and a satellite DNA molecule while, PYVM is caused by bipartite virus (Muniyappa *et al.,* 2003; 2005; Xiong *et al.,* 2006; Swanson *et al.,* 2007; Singh *et al.,* 2009; Shih *et al.,* 2009). The absence of amplicons in okra, bittergourd and ash gourd could be due to the fact that the virus could be a different strain. This also indicated that none of these plant species are acting as collateral host for the virus. However, there may be some plants which may be acting as source of the virus inocula without producing any external symptom or the symptom may not be expressed as yellow vein as in the case of PYVMV infected bottle gourd exhibited crinkling of leaves (Muniyappa *et al.,* 2003). Hence elimination of such hosts are important for the disease management.

In conclusion, geminivirus causing yellow vein mosaic of pumpkin in Kerala has been characterized. The primers designed were used for amplification of CP and MP genes which were then cloned and sequenced. Sequence analysis of both coat protein and movement protein genes revealed that the virus is closely related to bipartite *Squash leaf curl China virus.* Further investigation based on electron

microscopic study is required to confirm the bipartite nature of the virus. Light microscopy cannot be used to visualize virus particles as they are minute (nano scale). The primers designed in the study could be used for molecular detection of PYVMV.

The sequence information of coat protein and movement protein genes obtained in the study could be further exploited in crop improvement programmes for developing PYVMV resistant varieties. The sequence information of different genes could be used to serve the purpose. Multiple approaches for engineering resistance to geminiviruses have been employed in plants, in most cases in the model plant system. Such technologies can also be applied to crop plants like pumpkin and they would bring significant benefit to crop breeding and production. Geminivirus genes that could confer pathogen derived resistance include CPs, replicases and MPs. Defective interfering RNAs, DNAs and non translated RNAs also play a major role in disease resistance.

Significant advances in understanding the mechanism of resistance and the nature of genes involved will lead increased and sustainable resistance (Beachy, 1997). CP and MP mediated resistance can confer resistance to a broader range of viruses. CP-MR is the first reported and most widely studied mechanism. It has been used to confer resistance to a number of viruses in a variety of plant species. The CP produced by the transgene interferes with the disassembly of the virus thus inhibiting the infection. Furthermore specific mutations to the known sequences can confer much greater levels of resistance than the wild type. Similarly movement protein mediated resistance inhibits the virus by interfering with the virus movement within the plant. Hence, the knowledge in structure of the virus gene could aid in design of mutant genes that possess increased efficacy and breadth in protection.

Summary

6. SUMMARY

Plant virus diseases appear to be proliferating at ever increasing rates. Pumpkin is also known to be affected by many virus diseases. Pumpkin yellow vein mosaic (PYVM) is the major disease that is responsible for reduction in the yield in many pumpkin growing tracts all over the world. Considerable efforts are needed to control this disease including accurate and rapid diagnosis using suitable technologies. Polymerase chain reaction in different formats is the most widely used molecular technique for detection of bacteria and viruses. The application of PCR has now made it possible to amplify low copy number of viral DNA molecules and their subsequent detection and characterization. Hence the study was undertaken for molecular characterization and detection of the virus. The work undertaken and the salient findings obtained in the study "Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex. Poir)" are given below.

1. Leaf samples with typical yellow vein symptom were collected from the research field of Department of Olericulture, College of Horticulture, Vellanikkara, Thrissur.
2. DNA isolated from these samples yielded intact bands indicating the DNA was of good quality and Quantity.
3. Primers were designed for coat protein and movement protein genes based on the homology within the conserved regions of nucleotide sequences.
4. PCR amplification with the CP and MP gene specific primers yielded amplicons of expected size ie., 891bp and 702bp respectively.
5. The amplified fragments were eluted, cloned in plasmid vector and competent *E. coli* cells were transformed with the ligated product. A combination of blue and white colonies were obtained after overnight incubation confirming successful transformation.

6. Presence of the insert was confirmed by colony PCR in which amplicons of the expected size was obtained from the white colonies.
7. Full length coat protein and movement protein genes were cloned and sequenced.
8. Nucleotide sequence analysis of CP and MP genes:
   a. Sequence analysis with blastn revealed maximum identity to *Squash leaf curl China virus*.

   b. Phylogenetic tree constructed for the sequences revealed high level evolutionary relationship with SLCCV.
9. Amino acid sequence analysis of CP and MP genes:

   a. Sequence analysis with blastp revealed maximum identity to SLCCV and *Tomato leaf curl New Delhi virus*.

   b.Phylogenetic tree constructed using the sequences revealed high level evolutionary relationship with SLCCV and ToLCNDV.
10. Secondary structure of CP and MP predicted using SOPMA showed that random coils are the major structural components. Alpha helices and beta sheets are more or less equal in their proportion.
11. The technique was validated using 15 infected and 4 healthy pumpkin samples. Amplification was obtained in all the infected samples and no amplification was observed in the healthy sample.
12. PCR amplification was obtained in apparently healthy leaf samples from infected field indicating that the primers could be used to detect latent infection.
13. PCR amplification was not obtained with yellow vein mosaic infected okra, ash gourd, *Ageratum, Emilia, Hibiscus* and *Synedrella* showing yellow vein symptoms and bitter gourd with distortion mosaic which indicated that none of these plants are acting as the collateral host of the virus.

Sequence and phylogenetic analysis of coat protein and movement protein gene sequences showed maximum similarity to bipartite *Squash leaf curl china virus* (SLCCV) from Coimbatore. Hence Pumpkin yellow vein mosaic virus from Kerala can be considered as a strain of SLCCV.

References

REFERENCES

Abouzid, A. M., Freitas-Astua, J., Purcifull, D.E., Polston, J.E., Beckham, K. A., Crawford, W. E., Petersen, M. A., Peyser, B., Patte, C., and Hiebert, E., 2002. Serological studies using polyclonal antisera prepared against the viral coat protein of four begomoviruses expressed in *Escherichia coil. Plant Dis. 86*: 1109-14.

Akbergenov, R., A. Si-Ammour, T. Blevins, I. Amin, C. Kutter, H. Vanderschuren, P. Zhang, W. Gruissem, F. Meins, Jr., T. Hohn, and Poogin, M. M. 2006. Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Res. 34:462-471.

Babu, T. P. H. 2000. RAPD analysis to assess the genetic stability in tissue culture derived black pepper (*Piper nigrum* L.) plants. MSc (Ag) thesis, Kerala Agricultural University, Thrissur, 83p.

Baulcombe, D. 1996. Mechanisms of pathogen-derived resistance to viruses in transgenic plants. *Plant Cell.* 8:18333-1844.

Barnwell, P., Blanchard, A. N., Bryant, J. A., Smirnoff, N., and Weir, A. F. 1998. Isolation of DNA from the highly mucilaginous succulent plant shape *Sedum telephium. Plant Mol. Biol. Reptr.* 16 (2): 133-138.

Beachy, R. N., Loesch-Fries, S., and Tumer, N. E. 1990. Coat protein resistance against virus infection. *Annu. Rev. Phytopathol.* 28: 451-474.

Beachy, R. N. 1994. Introduction: transgenic resistance to plant viruses. *Semin. Virol.* 4: 327-328.

Beachy, R. N. 1997. Mechanism and application of pathogen-derived resistance in transgenic plants. *Curr. Opin. Biotechnol.* 8: 215-220.

Bela-ongl, B. B. and Bajet, N. B. Molecular Detection of Whitefly-Transmissible Geminiviruses (Family *Geminiviridae,* Genus *Begomovirus)* in the Philippines. *Philippines J. Sci.* 136 (2): 87-101.

Bhargava, B. and Bhargava, K. S. 1977. Cucurbit Mosaic Viruses in Gorakhpur. *Indian J. Agric. Sci.* 47:1-5.

Bock, K. R., Guthrie E. J., and Woods, R. D. 1974. Purification of maize streak virus and its relationship to viruses associated with streak disease of sugarcane and *Panicum maximum. Ann. Appl. Biol.* 77: 289-296.

Boulton, M. I. 2003. Geminiviruses: major threats to world agriculture. *Ann. Appl. Biol.* 142: 143.

Briddon, R. W. and Markham, P. G. 1994. Universal primers for the PCR amplification of dicot infecting geminivirus. *Mol. Biotechnol.* 1: 202-205.

Briddon, R. W., Mansoor, S., Bedford, I. D., Pinner, M. S., Saunders, K., Stanley, J., Zafar, Y., Malik, K. A., and Markham, P.G. 2001. Identification of DNA components required for induction of cotton leaf curl disease. *Virology.* 285: 234-243.

Briddon, R. W., 2003. Cotton leaf curl disease, a multicomponent begomovirus complex. *Mol. Plant Pathol.* 4: 427-434.

Briddon, R. W., Bull, S. E., Amin, I., Idris, A. M., Mansoor, S., Bedford, I. D., Dhawan, P., Rishi, N., Siwatch, S. S., Abdel-Salam, A. M., Brown, J. K., Zafar,

Y, and Markham, P. G. 2003. Diversity of DNA beta; a satellite molecule associated with some monopartite begomoviruses. *Virology.* 32: 223- 234.

Briddon, R. W., and Stanley, J. 2006. Sub-viral agents associated with plant infecting single stranded DNA viruses. *Virology.* 344: 198-210.

Brown, J. K. and Bird, J.1992. whitefly transmitted geminiviruses in the Americas and the Caribbean Basin: past and present. *Plant Dis.* 76: 220-225.

Brunnetti, A., Tavazza, M., Norris, E., Tavazza, R., Cacigali, P., Ancora, G., Crespi, S., and Accoto, G. P. 1997. High expression of truncated viral rep protein confers resistance to tomato leaf curl virus in transgenic tomato. *Mol. Plant. Microbe. Inter.* 10: 571-579.

Buck, K. R., Guthrie, E. J., and Woods, R. D. 1974. Purification of maize streak virus and its relationship to viruses associated with streak diseases of sugarcane and *Panicum maximum. Ann. Appl. Biol.* 77: 289-296.

Capoor, S. P., and Ahmed, R. U. 1975. Yellow Vein Mosaic disease of field pumpkin and its relationship with the vector, *Bemisia tabaci. Indian Phytopathol.* 28: 241-246.

Chatterji, A., Beachy, R. N. and Fauquet, C. M. 2001. Expression of the oligomerization domain of the replication-associated protein (Rep) of tomato leaf curl New Delhi virus interferes with DNA accumulation of heterologous geminiviruses.. *J. Biol. Chem.* 276: 25631-25638.

Crowther, J. R. 1998. Detection of viruses in livestock. *Parasitology* 177: 29-40.

Czosnek, H., Ghanim, M., Morin, S., Rubinstein, G., Fridman, V. and Zeidan, M. 2001. Whiteflies: vectors, and victims, of geminiviruses. *Adv. Virus. Res.* 57: 291-321.

Czosneek, H., Ghanim, M., and Ghanim, M. 2002. The circulative pathway of begomoviruses in the whitefly vector *Bemisia tabaci*- insights from studies with *Tomato yellow leaf curl virus. Ann. Biol.* 140: 215-231.

Dellaporta, S. L., Wood, J., and Hicks, J. B. 1993. A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* 1: 19-21

Deng, A., McGrath, P. F., Robinson, D. J., and Harrison, B. D. 1994. Detection and differentiation of whitefly transmitted gemini viruses in plants and vector insects by the polymerase chain reaction with degenerate primers. *Ann. Biol.* 125: 327-336.

Dong. X., Van Wezel, R. Stanley, J and Hong, Y. 2003. Functional characterization of the nuclear localization signal for a suppressor of post- transcriptional gene silencing *J. Virol.* 77: 7026-7033.

Doyle, J. J. and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue, *Phytochem Bull.* 19-11-15.

Duan, Y. P. Powell, C.A. Pureifull, D.E. Broglio. P and Hiebert, E. 1997. Phenotypic variation in transgenic plants expressing mutated geminivirus movement/ pathogenicity (BCI) proteins. *Plant Microbe interact.* 10:1065-1074.

Duffus, J. E.1987. Whitefly transmission in plant viruses. In Current Topics in Vector of squash leaf curl virus (SqLCV) Research. Vol. 4 K.F. Harris (Ed.) *Springer* Verlag New York. 73-91.

Fauquet, C. M. and Stanley, J. 2003. Geminivirus classification and nomenclature; progress and problems. *Ann Appl. Biol.* 25: 165-198.

Fauquet, C. M. and Stanley, J. 2005 Revising the way we conceive and name viruses below the species level: a review of geminivirus taxonomy calls for new standardized isolate descriptors *Arch virol.* 150:2151-2179.

Fauquet, C. M. Briddon R. W. Brown, J. K. Moriones, E. Stanley, J. Zerpini, M., and Zhou, X. 2008. Geminivirus strain demarcation and nomenclature *Arch Virol.* 153:783-821.

Fenby, N. S. Scott, N. W. Slater, A., and Elliott, M. C. 1995. PCR and non – isotopic labeling techniques for plant virus detection. *Cell Mol. Biol.* 41(5):639-652.

Frischmuth, T., Engel, M., Lauster S., and Jeske, H. 1997. Nucletoide sequence evidence for the occurrence of three distinct whitefly – transmitted,Sida-infecting bipartite bipartite geminiviruses in Central America. *J Gen. Virol.* 78 (10): 2675-2682.

Fulton, T. M. Chunwonge, J., and Tanksley, S. D. 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *plant Mol. Biol. Rep.* 13:207-209.

Gafni, Y. and Epel. B. L. 2002. The role of host and viral proteins in intra and inter cellular trafficking of geminiviruses. *Physiol Mol. Plant path.* 60:231-241.

Garg, I. D., Khurana, S. M., Kamar S. and Lakra B. S. 2001. Association of a geminivirus with potato apical leaf curl in India and its immuno- electron microscopic detection. *Virology.* 28: 227-32.

Gennadius P. 1989. Diseases of tobacco plantation in the Trikonia. The aleurodid of tobacco. The aleurodid of tobacco. *Ellenike Georgia.* 5:1-3.

Ghosh, S. K. and Mukhopadhyay, S. 1979. Viruses of Pumpkin (*Cucurbita moschata*) in West Bengal. *J. Phytopathol.* 92(2): 172:184.

Goodman, R. M. 1977. Single stranded – DNA genome in a whitefly – transmitted plaint virus. *Virology.* 83:171-179.

Guo, X and Zhou, X 2006. Molecular characterization of a new begomovirus infecting *Sida cordifolia* and its associated satellite DNA molecules. *Virus genes.* 33(3): 279 -285.

Hallan V., Gafni Y. 2001. *Tomato yellow leaf' Curl virus* (TYLCV) capsid protein (CP) submit interactions: implications for viral assembly. interactions: implications for viral assembly. *Arch. Virology.* 146:1765-1773.

Hamilton, R. I. 1980. Defences triggered by previous invaders: viruses. In Horsfall, J.G. and Cowling, E.B. (eds.), Plant Disease. Academic Press, Vol.5, pp. 279-303.

Hamilton, W. D. O., Bisaro, D. M., and Buck, K.W. 1982. Identification of novel DNA forms in tomato golden mosaic virus infected tissue: Evidence for a two component genome. *Nucleic Acids Res.* 4901-4912.

Hannon, G.J. 2002. RNA interference. *Nature.* 418:244-251.

Harrison, B. D., Barker H, Bock K. R., Guthrie E. J,. Meredith G., and Atkinson M. 1977 Plant viruses with Circular single-stranded DNA. *Nature.* 270:760-762.

Harrison B. D., Muniyappa V., Swanson M. M. Roberts I. M., Robinson D. J 1991. Recognition and differentiation of seven whitefly –transmitted geminiviruses from India and their relationships to African cassava mosaic and Thailand mung bean yellow mosaic viruses. *Ann. Appl. Biol.* 118:299-308.

Harrison B. D. and Robinson D. J 1999. Natural genomic and antigenic variation in whitefly– transmitted geminiviruses (Begomoviruses). *Annu.Rev. Phytopathol.* 37: 369-398.

Harrison B. D. and Swanson M. M. and Fargette D. 2002. Begomovirus coat protein: serology, variation and functions. *Physiol.Mol.Plant Path.* 60(5): 257-271.

Horwitz J. P., Chua J., Curby, R. J., Tomson, A. J., Daroog, M. A. Fisher, B. E., Mauricio, J., and Klundt, I. 1964. Sustrates for cytochemical demonstration of enzyme activity- I. Some substituted 3- Indoyl - β-D – glycopyranosides, *J.Med. Che.* 7: 574-575.

Hussain, M., Mansoor, S., Iram, S., Zafar Y., and Briddon, R. W. 2004. First report of Tomato leaf curl New Delhi virus affection chilli pepper Pakistan, *Plant, Pathol.* 53(6):794.

Ingham, D. J., Pascal, E., and Lazarowitz, S. G. 1995. Both geminivirus movement proteins define viral host range, but only BL 1 determines viral pathogenicity. *Virology.* 207:191-204.

Ioannou, N. 1985. Yellow leaf curl and other diseases of tomato in Cyprus. *Plant Pathol.* 345:428-434.

Ito T., Sharma, P., Kittipakorn, K., and Ikegami, M. 2008. Complete nucleotide sequence of a new isolate of tomato leaf curl New Delhi virus inflecting cucumber, bottle gourd and muskmelon in Thailand. *Arch. Virol.* 153:611-613.

Jaiswal, N., Saritha R. K., Datta, D., Sing, M., Dubey, R. S., Rai, A. B., and Rai, M. 2011. Molecular Characterization of Tomato leaf curl Palampur virus and Pepper leaf curl betasatellite Naturally Infecting Pumpkin (*Cucurbita moschata*) in India. *Indian J. Virol.* 21(2): 128-132.

*Jayashree, K. 1996: studies on Yellow vein mosaic disease of pumpkin M.Sc. (Ag) thesis, Tamil Nadu Agricultural University, Coimbatore India. 80p.

Jayashree, K, Pun K.B. and Doraiswamy, S. 1999, Virus –vector relationship of yellow vein mosaic virus and whitefly (Bemisia tabaci) in pumpkin. *India Phytopathol.* 52:3-10.

Jose, J. and Usha R. 2000. Extraction of geminiviral DNA from a highly mucilaginous plant (*Abelmoschus esculentus). Plant. Mol. Rep.* 18:349-355.

Jose, J. and Usha, R. 2003. Bhendi yellow vein mosaic disease in India is caused by association of a DNA beta statellite with a begomovirus. *Virology.* 305: 310-317.

Khan, J. A. 2000. Detection of tomato yellow leaf curl geminivirus in its vector *Bemisia tabaci. Indian J. Exp. Biol.* 38:512-515.

Kon, T., Dolores L.M., Bajet N.B., Hase S., Takahashi H and Ikegami M. 2003. Molecular Characterization of a Strain of squash Leaf Curl China Virus from the Philippines. *J. Phytopathol.* 151 (10): 535-539.

Kon, T., Kuwabara, K., Hidayat, S. H. and Ikegami, M. 2006. A begomovirus associated with ageratum yellow vein disease in Indonesia : evidence for natural recombination between tomato leaf curl Java virus and Ageratum yellow vein virus (Java) *Arch. Virol.* 152:1147-1157.

Kon, T., Sharma, P., and Ikegami, M. 2007. Suppressor of RNA silencing encoded by the monopartite tomato leaf curl Java begomovirus. *Arch. Virol.* 152: 1273-1282.

Kon, T., Rojas, M.R., Abdourhamane, I.K. & Gilbertson, R.L. 2009. Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *J. Gen Virol.* 90: 1001-1013.

Konate, G., Barro, N., Fargette, D., Swanson, M. M., and Harrison, B.D. 1995. Occurrence of whitefly- transmitted geminiviruses in crops in Burkina Faso, and their serological detection and differentiation. *Ann. Appl. Biol.* 126 (1): 121-129.

Koonin, E. V., and Ilyina, T. V. 1992. Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA initiation proteins. *J.Gen. Virol.* 73: 2763-2766.

Krupovic, M., Ravantti, J. J., and Bamford, D. H. 2009. Geminiviruses: a tale of a plasmid becoming a virus. *BMC Evol. Biol.* 9: 112-123.

Kumar, Y., Bhardwaj, P., Hallan, V., and Zaidi, A. A. 2010. Detection and characterization of Ageratum enation virus and a nanovirus –like satellite

DNA1 from zinnia causing leaf curl symptoms in India. *J. Plant. Pathol.* 76:395-398.

Kunik, T., Karuppaiah, P., Czoszcnek, H., Citovsky, V., and Gafni, Y. 1998. Nuclear import of tomato yellow leaf curl virus (TYLCV) in plant and insect cells. *The plant J.* 13: 393-399.

Kushwaha, N., Singh, A. K., Chattopadhyay, B., and Chakraborty, S. 2010. Recent advances is geminivirus detection and future perspectives. *J. Plant Prot. Sci.* 2(1): 1-18.

Kyte, J. and Doolittle, R. F. 1982. A simple method of for displaying hydropathic character of a protein.*J. Mol. Biol.* 157 (1): 32-105.

Latha, P. 1992. Selection of mosaic resistance resistance in pumpkin. M.Sc. (Hort.) thesis. Kerala Agricultural University. 10p.

Lazarowitz, S.G. 1992. Geminiviruses genome structure and gene function. *Crit. Rev. Plant Sci.* 11: 327-349.

Liao, Y., Hu, C., Lin, T., Chang, C., and Deng, D. 2007. Identification of Squash leaf curl Philippines virus on Benincasa hispida in Taiwan. *Plant Pathol. Bull.* 16:11-18.

Lindbo, J. A. and Dougherty, W. G. 1992. Untranslatable transcripts of the tobacco etch virus coat protein gene sequence can interfere with tobacco etch virus replication in transgenic plants and protoplasts. *Virology.* 189: 725-733.

Lippman, Z. and Martienssen, R. 2004. The role of RNA interference in heterochromatic silencing. *Nature.* 431 (7006): 364-370.

Liu, S., Bedford, I. D., Briddon, R.W., and P.G. Markham. 1997. Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. *J. Gen. Virol.* 78 (7): 1791-1794.

Lodhi, M. A., Ye, G. N., Weeden, N. F., and Reisch, B.I . 1994. A simple and efficient method for DNA extraction from grapevine cultivars, Vitis species and Ampelopsis. *Plant Mol. Biol. Rep.* 12: 6-13.

Loesch-Fries, L. S., Merlo, D., Zinnzn, T., Burhop, L., Hill, K., Krahn, K., Jarvis, N., Nelson, S. and Halk, E. 1987. Expression of alfalfa mosaic virus RNA 4 in transgenic plants confers virus resistance. *EMBO J.*, 6, 1845-1851.

Lopez, M. M., Bazzi, C., and Gieder, K. 2003. Innovative tools for detection of plant pathogenic viruses and bacteria. Int. Microbiol., 6: 233-243.

Mandal, B. 1989, Studies on croton yellow vein mosaic virus. M.Sc. (Ag) thesis. University of Agricultural Science, Bangalore. 92p.

Manickam, K., Rabindran, R., Ganapathy T., Doraiswamy, S. 2001. Detection of Gemini viruses in dicot weeds using polymerase chain reaction. *Madras Agric. J.* 88: 534-535.

Mansoor, S., Hussain, M., Khan, S. H. Bashir., A. Leghari., A. B. Panwar, G.A., Siddiqui W. A., Zafar Y., and Malik K. A. 1998. Polymerase Chain Reaction based detection of cotton leaf curl and other whitefly-transmitted geminiviruses from Sindh. Pakistan J. Biol. Sci. 1 (1): 39-43.

Mansoor, S., Mukhtar S., Hussain M., Amin I., Zafar Y., and Malik K. A. 2000. Widespread occurrence of cotton leaf curl virus on radish in Pakistan. *Plant. Dis.* 84: 809.

Marechal- Drouard, L. and Guillemaut, P. 1995. A powerful but simple technique to prepare polysaccharide- free DNA quickly and without phenol extraction. *Plant Mol. Biol. Reptr.* 13: 26-30.

Markam, P. G., Bedford S. L., and Pinner M. S. 1994. The transmission of geminivirus by Bemisia tabaci. *Pestic. Sci.* 1994, 42: 123-128.

Maruthi, M. N., Colvin J., Briddon R. W., Bull S.E., and Muniyappa, V. 2003. Pumpkin yellow vein mosaic virus: a novel begomovirus infecting cucurbits. *J. Plant Pathol.* 85: 63-5.

Maruthi, M. N., Rekha, A. R., Mirza, S. H., Alam S. N., and Colvin, J. 2007. PCR- based detection and partial genome sequencing indicate high genetic diversity in Bangladeshi begomoviruses and their whitefly vector, Bemisia tabaci. *Virus Genes.* 34: 373-385.

Maruthi, M. N., Rekha, A. R., and Muniyappa, V. 2007. Pumpkin yellow vein mosaic disease is caused by two district begomoviruses: Complete viral sequences and comparative transmission by an indigenous Bemisia tabaci and the introduced B- biotype. *EPPO Bull.* 37: 412-419.

Mathew A.V., Host range of Indian cassava mosaiv virus. 1989. *Indian Phytopathol.* 46: 16-23.

Mathew, A. V. and Muniyappa, V. 1991. Transmission of Indian Cassava mosaic virus by Bemisia tabaci .*Fitologia Brasileira.* 16: 122-125.

McGlashan, D., Poston, J. E., and Bois, D, 1994. Tomato yellow leaf curl geminivirus in Jamaica. *Plant Dis.* 78: 1219.

McGovern, R. J., Polston, J. E., Danyluk, G. M., Hiebert, E., Abouzid A. M., and Stansly, P. A. 1994. Identification of a natural weed host of tomato mottle geminivirus in Florida. *Plant Dis.* 78: 1102-1106.

Miller, S. A. and Martin R. R. 1988. Molecular diagnosis of plant disease. *Annu. Rev. Phytopathol.* 46: 409-432.

Morris, J., Harju V., Cambridge J., Boonham, N., and Henry, C. 2002, Virus transmission studies and diagnostic detection of four begomo virus isolates in selected crop hosts and in Bemisia tabaci, biotype- B. *EPPO Bull.* 32: 41-45.

Mullis K. B. and Faloona F. A. 1987. Specific synthesis of DNA in vitro via a polymerasecatalyzed chain reaction. *Methods Enzymol.* 155: 335-350.

Muniyappa V. and Reddy H. R. 1976, Studies on the yellow mosaic disease of horsegram (Dolichos biflorus Linn.). Virus- vector relationship. *Mysore J. Agric. Sci.* 10: 605-610.

Muniyappa, V., Swanson, M. M., Duncan, G. H., and Harrison B. D. 1991. Particle purification, properties and epitope variability of Indian tomato leaf curl geminivirus. *Ann. Appl. Biol.* 118: 595-604.

Muniyappa, V., Venkatesh, H. M., Ramappa, H. K., Kulkarni, R. S., Zeidan, M., Tarba, C. Y., Ghanim M., and Czosnek, H. 2000. Tomato leaf curl virus from Bangalore (ToLCV-Ban 4): Sequence comparison with Indian ToLCV isolates, detection in plants and insects, and vector relationships. *Arch. Virol.* 145: 1583-1598.

Muniyappa, V., Maruthi, M. N., Babitha, C. R, Colvin J., Briddon R. W. and Rangaswamy K. T, 2003. Characterisation of pumpkin yellow vein mosaic virus from India. *Ann. Appl. Biol.* 142: 323-31.

Nakhla, M. K., Mazyad ,H. M., and Maxwell, D. P., 1993. Molecular characterization of four Tomato yellow leaf curl virus isolates from Egypt and development of diagnostic methods. *Phytopathol. Mediterr.* 32: 163-173.

Nakhla, M. K. and Maxwell, D. P. 1994. Widespread occurrence of the eastern Mediterranean strain of tomato yellow leaf curl geminivirus in tomatoes in the Dominican republic. *Plant. Dis.* 78: 926.

*Nateshan, H. M., 1992. Biological properties of cotton leaf curl virus. M.Sc. (Ag) thesis. University of Agricultural Science, Bangalore, India. 123 p.

Navot, N., Zieden, M., Pichersky, E., Zamir, D., and Czoznek, H. 1992. Use of Polymerase Chain Reaction to Amplify Tomato Yellow Leaf Curl Virus DNA from infected plants and viruliferous whiteflies. *Phytopathology.* 82 (10): 1192-1202.

Nawaz-ul-Rehman, M. S. and Fauquet, C. M. 2009. Evolution of geminiviruses ad their satellites. *FEBS Lett.* 18(12) : 1825-32.

Nesbit, K. A., Potts, B. M., Vaillancourt, R. E., Wests, A. K., and Reid, J. B. 1995. Partitioning and distribution of RAPD variation in a forest tree species, Eucalyptus globules (Myrtaceae). *J. Heredity.* 74: 628-637.

Nguyen, L., Lucas, W. J., Ding, B., and Zaitlin, M. 1996. Viral RNA trafficking is inhibited in replicase- mediated resistant transgenic tobacco plants. Proc. Natl Acad. Sci. USA. 93: 12643-12647.

Noris, E., Accotto G. P., Tavazza, R., Brunetti, A., Crespi, S. and Tavazz, M. 1996. Resistance to Tomato yellow leaf curl geminivirus in Nicotiana benthamiana plants transformed with a truncated viral C1 gene. *Virology.* 224:130-138.

Osbourn, J. K., Watts, J. W., Beachy R. N., and Wilson T. M. A. 1989. Evidence that nucleocapside disassembly and a later step in virus replication are inhibited in transgenic tobacco protoplasts expressing TMV coat protein. *Virology.* 172: 370-373.

Padidam, M., Beachy, R. N., and Fauquet C. M. 1995. Classification and identification of geminiviruses using sequence comparisons. *J. Gen. Virol.* 7: 249-263.

Padmalatha, K. and Prasad M.N. V. 2006. Optimisation of DNA isolation and PCR protocol for RAPD analysis of selected medicinal and aromatic plants of conservation concern from peninsular India. *African J. Biotech.* 5(3): 230-234.

Palanichelvam, K., Kunik, T., Citovsky, V., and Gafni, Y. 1998. The capsid protein of alone. Tomato yellow Leaf Curl Virus binds co-operatively to single–stranded DNA. *J. Gen. Virol.* 78: 2829-2833.

Peter, K. V. 2007. Underutilized and underexploited horticultural crops. Vol. 1. New India Publishing. 238p.

Polston, J. E., McGovern, R. J., Brown L.G. 1999. Introduction of tomato yellow leaf curl virus in Florida and implications for the spread of this and other geminiviruses of tomato. *Plant Dis.* 83: 984-988.

Powell, A. P, Nelson, R. S., De, B., Hoffmann, N., Rogers, S. G., Fraley R. T., and Beachy, R. N.1986. Delay in disease development in transgenic plants that express the tobacco mosaic coat protein gene. *Science* 232:738-743.

Praveen, S., Dasgupta, A., and Varma, A. 2004. Phylogenetic analysis and homologies of the replicase of tomato leaf curl geminiviruses: implications for obtaining pathogen derived resistance. *Virus Genes.* 28(2) 195-9.

Prins, M., Laimer, M., Noris, E., Schubert, J., Wassenegger, M. and Tepfer, M. 2008. Strategies for antiviral resistance in transgenic plants. *Mol. Plant. Pathol* 9(1): 73-83.

Qin, S., Ward, B. M., and Lazarowitz, S.1998. The bipartite geminivirus coat protein aids BRI function in viral movements by affection the accumulation of viral single –stranded DNA J. *Virol* 72:9247-9256.

Rajinimala, N., Rabindran, R., and Ramaiah, M. 2009. Detection of Bittergourd Yellow Mosaic virus (BGYMV) in bitter gourd (*Momordica charantia*) thorough Triple Antibody Sandwitch (TAS) ELISA and Polymerase Chain Reaction (PCR) *Arch. Phytopathol. and Plant Protec.* 42(6): 559-565.

*Ramappa H. K. 1993. Detection, host resistances and integrated management of tomato leaf curl geminivirus disease. Ph.D. Thesis. University of Agricultural Science, Bangalore.

Ramos, P. L., Guerra, O., Dorestes, V., Ramirez, N., Riverabustamante P., and Oramas, P. 1996. Detection of TYLCV in Cuba. *Plant Dis.* 80(10): 1208.

Reddy, B. A., Patil M. S. and Ram, R. T. 2010. Effect of plant extracts on the incidence of tomato leaf curl virus and whitefly population under filed conditions, Karnataka J. *Agric. Sci.* 23 (5): 807-808.

Rether, B., Delmas, G., and Laouedj, A. 1993. Isolation of polysaccharide –free DNA from plants. *Plants Mol. Biol. Rep.* 11:333-337.

Roberts, I. N. Robinson, D. J., and Harrison, B. D. 1984. Serological relationships and genome homologies among geminiviruses. *Ann.Appl.Biol.* 105: 483-493.

Rochester, D. E., Fauquet, C. M., Dapaulo, J. J, and Beachy, R. N. 1994. Complete nucleotide sequence of the geminivirus tomato yellow leaf curl virus (Thailand isolate). *J.Gen. Virol* 75:477-485.

Rodriguez –Negrete, E. A. Carrillo- Tripp J., and Rivera –Bustamante, R. F. 2009. RNA silencing against Geminivirus: Complementary Action of Posttranscriptional Gene silencing and Transriptional Gene silencing in Host Recovery *J. Virol.* 83(3): 1332-1340.

Rogers, S. O., and Benedich, A. J., 1994. Extraction of total cellular DNA from plants algae and fungi. *Plant Mol. Biol. Manual.* 91:1-18.

Rojas, A., Kvarnheden, A., and Valkonen, J. P. T. 2000. Geminiviruses infecting tomato crops in Nicaragua. *Plant Dis.* 84: 843-846.

Rojas, M. R., Hagen C., Lucas J. W., and Gilbertson, L. R. 2005. Expoiting Chinks in the Plant's Armor: Evolution and Emergence of Geminiviruses. *Annu. Rev. Phytopathol.* 43: 361-394.

Rouhibakhsh, A., Priya, J., Periasamy, M., Haq, Q. M. I., Malathi, V. G. 2008. An improved DNA isolation method and PCR protocol for efficient detection of multicomponents of begomovirus in legumes. *J. Virol. Methods.* 147:37-42.

Roye, M. E., Mc Laughlin W. A., Nakhla, M. R. and Maxwell D. P. 1997. Genetic diversity among geminiviruses associated with three weed species, *Sida spp. Macriptilium lathyroides* and *Wissadula amplissima in Jamaica. Plant Dis* 81:1251-7.

Rybicki, E. K. and Hughes F. L. 1990. Detection and typing of maize streak virus and other distantly related geminiviruses of grasses by polumerase chain reaction amplification of a conserved viral sequence. *J. Gen Virol.* 71 (11): 1317-2519.

Rybicki, E. P. and Pietersen, G., 1999. Plant virus disease problems in the developing world. *Virus Res.* 53: 127-175.

Rybicki, E., Briddon, R., Brown, J., Fauquet, C., Maxwell, D., Stanley, J., Harrison B. D., Markham, P. G. Bisaro D., and Robinson D., 2000. Family Geminiviridae. In "Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses" (M. H.V. Van Regenmortel, C.M. Fauquet, D.H. L. Bishop E., Carstens M.K., Estes S.M., Lemon J, Maniloff M.A., Mayo D.J., Mc Geoch, C.R. Pringle and R.B. Wickner, Eds), pp. 285-297. Academic Press, New York/ San Diego.

Saccardo, F., Cettull, E., Palmano, S., Noris, E., and Firro, G. 2011. On the alleged origin of geminiviruses from extrachromosomal DNA's of phytoplasmas. *BMC Ecol. Biol.* 11:185-197.

*Sambrook, J., Fritsh E.F., and Maniatis T., 1989. Molecular Cloning, Cold Spring harbor Laboratory Press.

Samretwanich, K., Chiemsombat, P., Kittipakorn, K. and Ikegami, M. 2000a. A new geminivirus associated with a yellow leaf curl disease of pepper in Thailand. *Plant .Dis.* 84(9): 1047-1047.

Samretwanich, K., Chiemsombat, P., Kittipakorn, K., and Ikegami, M. 2000b. Tomato leaf curl geminivirus associated with cucumber yellow leaf disease in Thailand. *J. phytopathol.* 148: 615-617.

Samretwanich, K., Chiemsombat, P. Kittipakorn, K., and Ikegami, M., 2000c. Yellow leaf disease of Muskmelon from Thailand caused by Tomato leaf curl virus. *Plant Dis.* 84: 707.

Sanford, J. C. and Johnson, S. A. 1985. The Concept of parasite- derived resistance –Deriving resistance genes from the parasite's own genome. *J. Theor. Biol.* 113: 395-405

Sanderfoot, A. A. Ingham, D. L., and Lazarowitz, S. G. 1996. A Viral movement protein as a nuclear shuttle. The geminivirus BRI movement protein contains domains essential for interaction with BLI and nuclear localization. *Plant Physiol.* 23-33.

Sanderfoot, A. A. and Lazarowitz, S.G. 1996. Getting it together in plant virus movement cooperative interactions between bipartite geminivirus movement proteins. *Trends Cell Biol.* 6: 353-358.

Sanz, A.I., Fraile A., Garcia- Arenal F., Zhou X., Robinson D. J., Khalid S., Butt T., Harrison B. D. 2000. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. *J. Gen. Virol.* 81. 1839-49.

Saunders, K., Bedford, I. D., Briddon, R. W., Markham, P. G., Wong, S. M., and Stanley, J. 2000. A unique virus complex causes Ageratum yellow vein disease. *Proc. Natl. Acad. Sci. U.S.A.* 97: 6890-6895.

Saunders, K., Bedford, I. D., Yahara, T., and Stanley, J. 2003. Aetiology: The earliest recorded plant virus disease *Nature.* 422: 831.

Saunders, K., Norman, A., Gucciardo, S., and Stanley, J. 2004. The DNA β satellite component associated with ageratum yellow vein disease encodes an essential pathogenicity protein (βCI) *Virology.* 324: 37-47.

Sequeira, J. C. and Harrison B. D. 1982. Serological studies on cassava latent virus. *Ann. Appl. Biol.* 101:33-42.

Sharma, A., Murayama, A., Osaki, T., and Ikegami, M. 1998. Characterization of virus –specific DNA forms from Tomato tissue infected by tobacco leaf curl virus: evidence for a single genomic component producing defective DNA molecules. *Plant Pathol.* 47:787-793.

Shimizu, S. and Ikegami, M. 1999. Complete nucleotide sequence and the genome organization of tobacco leaf curl geminivirus from Japan. *Microbiol. Immun*, 43:989-992.

Smith, D. R. and Maxwell, D. P. 1994. Requirement of the Common region of DNA- B and the BLI open reading frame of bean golden mosaic

geminivirus for infection of *Phaseolus vulgaris*. *Phytopatholgy*. 84133-138.

Shih, S. L., Kumar, S., Tsai, W. S., Lee, L. M., and Green, S. K. 2009. Complete nucleotide sequences of okra isolates of Cotton leaf curl Gezira virus and their associated DNA and from Niger, *Arch. Virol.* 154:369-372.

Singh, A. K. Mishra, K. K. Chattopadhyay, B., and Chakraborty, S. 2009. Biological and molecular characterization of a begomovirus associated with yellow vein mosaic disease of pumpkin from Northern India. *Virus Genes*. 39: 359-370.

Spiegel, S., Martin, R. R., Legget, P. Ter- Borg, M., and Postman, J. 1993. Recent developments in therapy and virus- detection procedures for international movement of clonal germplasm. *Plant Dis.* 70:722-726.

Stam, M. Mol. J. N. M., and Kooter, J. M. 1997. The silencing of genes in transgenic plants. *Ann Bot.* 79: 3-12.

Swangjith, S. 2009. The complete nucleotide sequash leaf curl China virus- [Wax gourd] and its phylogenetic relationship to other geminiviruses. *Sci. Asia* 35:131-136.

Swanson, M. M., Valand, G. B., Muniyappa, V., and Harrison, B. D. 1998. Serological detection and antigenic variation of two whitefly- transmitted geminiviruses: tobacco leaf curl and croton yellow vein mosaic viruses. *Ann. Appl. Biol.* 132 (3): 427-435.

Tahir, M. and Haider, M. S. 2005 First report of a begomovirus associated with leaf curl disease of bell pepper in Pakistan. *New Dis. Rep.* 12:25.

Tahir, M., Haider, M. S. Briddon R. W. 2010. Complete nucleotide sequences of a distinct bipartite begomovirus, bitter gourd yellow vein virus, infecting Momordica charantia. , bitter gourd yellow vein virus, infecting *Momordica charantia. Arch. Virol.* 155:1901-1905.

Tao, Y., Rao P. K. Bhattacharjee, S., and Gelvin, S. B. 2004 Expression of plant protein phosphatase 2C interferes with nuclear import of the *Agrobacterium* T- complex protein Vir D2. *Proc. Natl Acad Sei. USA 101:5164-5169.*

Tu, Z., He, G., and Li, K. X. 2005. An improved system for competent cell preparation and high efficiency plasmid transformation using different E. coli strains. *Electron. J. Biotechnol.*.8(1):113-120.

Ullmann, A., Jacob, F., and Monod, J. 1967. Characterization by in vitro complementation of a peptide corresponding to an operator proximal segment of the B- Galactosidase structural gene of *Escherichia coli. J. Mol. Biol.* 24:339-343.

Umaharan, P. Padidam, M Phelps R.H. Beachy, R.N. and Fauquet, c.M. 1998, Distribution and diversity of geminiviruses in Trinidad and Tobago. *Phytopathol.* 88: 1262-1268.

Usharani, K. S. Surendranath, B., Paul – Khurans, S. M. Garg, I. D. and Malathi, V. G. 2004. Potato leaf curl- a new disease of potato in northern India caused by strain of Tomato leaf curl New Delhi virus. *Plant Pathol.* 53: 235.

Vanderschuren, H., Akbergenov, R., Pooggin, M. M., Hohn, T., Grussem, W., and Zhang, P. 2007. Transgenic cassava resistance to African cassava

mosaic virus in enhanced by viral DNA- A bidirectional promoter Derived siRNAs. *Plant Mol. Biol.* 64:549-557.

Van Dun, C. M. P. Overduin, B. Van Vloten., Doting, L. and Bol. J. F. 1988. Transgenic tobacco plants expressing tobacco streak on mutated alfgalfa mosaic virus coat protein does not cross- protect against alfalfa mosaic virus infection. *Virology* . 164: 383-389.

Varma, A. and Malathi, V. G. 2003. Emerging geminivirus problems: a serious threat to crop production *Ann. Appl.* Biol. 142:145-164.

Vasudeva, R. S.and Lal, T. B. 1943 Mosaic diseases of bottle gourd. *Indian J. Agric. Sci.* 13: 182-191.

Voinnet, O. 2005. Induction and suppression of RNA silencing : insights from viral infections. *Nat. Rev. Genet.* 6:206-220.

Watcharawongpaiboon, N. and Chunwongse, J. 2007. Development of microsatellite markers from and enriched genomic library of pumpkin (*Cucurbita moschata* L.) *Songlkanakarin J Sci. Technol.* 29. (5): 1217-1223.

Webster, C.G. Wylie. S.J. and Jones M.G.K. 2004. Diagnosis of plant viral pathogens. *Curr. Sci.* 86:1604-1607.

Wu, Z. C. Hu, J. S. Polston, J. E. Ullman, D. E., and Hiebert, E. 1996. Complete nucleotide sequence of a nonvector –transmissible strain of Abutilon mosaic geminivirus in Hawaii. *Phytopathology.* 86:608-613.

Wyat, S. D. and Brown, J. K. 1996. Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. *Phytopathology.* 86 (12) 1288 -1293.

Xie, Y., Zhou, X. P. Zhang, Z. K. and Qi, Y. J. 2002. Tabacco curly shoot virus isolated in Yunnan is a distinct species of Begomovirus. *Chin. Sci. Bull.* 47:197-200.

Xiong, Q., Fan, S., Wu, J., and Zhou. 2006. Ageratum yellow vein China virus is Distinct Begomovirus Species. Associated with a DNAβ Molecule. *Virology.* 97(4): 405-411.

Yin, Q. Y., Yang, H. Y., Gong, Q. H., Wang, H., Liu, Y. L., Hong, Y. G., and Tien, P. 2001. Tomato yellow leaf curl China virus: monopartite genome organization and agroinfection of plaint. *Virus Res.* 81:69-76.

Zachariya, R. M. 2007. Characterization and management of distortion mosaic virus. MSc. (Ag.). Thesis. Kerala Agricultural university. 60p.

Zhou, X. P., Liu, Y., and Calvert, L. 1997. Evidence that DNA- A of a geminivirus associated with severe mosaic disease in Uganda has arisen by interspecific recombination. *J. Gen. Virol.* 78:2101-2111.

Zhou, X. P., Xie, Y., Peng, Y., and Zhang, Z. 2003. *Malvastrum yellow vein virus*, a new *Begomovirus* species associated with satellite DNA molecule. *Chin. Sci. Bull.* 48(20): 2206- 2210.

Zrachya, A., Kumar, P. P. Ramakrishna, U., Levy, Y., Loyter, A. Arazi, T., Lapidot M., and Gafni, Y. 2007. Production of si RNA targeted against TYLCV coat protein transcripts leads to silencing of its expression and resistance to the virus. *Transgenic Res.* 16:385 – 398.

* Originals not seen

Appendices

APPENDIX-1

LIST OF LABORATORY EQUIPMENTS USED FOR THE STUDY

Name of the instrument	Make
Spectrophotometer	Spectronic Genesys-5, USA
Nanodrop Spectrophotometer	NanoDrop Technologies Inc., U.S.A
Referigerated centrifuge	Kubota, Japan
Thermal cycler	Eppendorf, Germany
Gel documentation system	Biorad, USA
Horizontal electrophoresis system	Biorad, USA

APPENDIX-II

COMPOSITION OF REAGENTS USED FOR DNA ISOLATION

1. Rogers and Bendich (CTAB) method

2X CTAB extraction buffer
CTAB (2%,v/v)
100mM Tris buffer (pH 8)
20mM EDTA (pH 8)
1.4M NaCl
1 % PVP

10% CTAB solution
10% CTAB (w/v)
0.7M NaCl

TE buffer
10mM Tris (pH 8)
10mM EDTA (pH 8)

2. Modified CTAB method

2X CTAB extraction buffer
CTAB (2%,v/v)
100mM Tris buffer (pH 8)
20mM EDTA (pH 8)
2.0 M NaCl
1 % PVP

10% CTAB solution
10% CTAB (w/v)
1.0 M NaCl

TE buffer
10mM Tris (pH 8)
10mM EDTA (pH 8)

APPENDIX –III

COMPOSITON OF BUFFERS AND DYES USED FOR GEL ELECTROPHORESIS

1. TAE buffer 50X (for 1 litre)

242g Tris base

57.1 ml glacial acetic acid

100 ml 0.5M EDTA

2. TBE buffer 10X (for 1 litre)

54g Tris base

27.5g Boric acid

20ml 0.5M EDTA (pH 8.0)

3. Loading Dye (6 X)

0.25% bromophenol blue

0.25% xylene cyanol

30 % glycerol in water

4. Formamide dye

Formamide-10ml

Xylene cyanol-10mg

Bromophenol blue-10mg

0.5M EDTA (pH-8.0)-200µl

APPENDIX IV

COMPOSITION OF DIFFERENT MEDIA USED IN THE STUDY

1. Luria Bertani (LB) broth

Tryptone	-	10 g
Yeast Extract	-	5 g
NaCl	-	5 g
pH adjusted to	-	7± 0.2
Distilled water	-	to make up to 1000ml

2. Luria Bertani agar medium

Tryptone	-	10 g
Yeast Extract	-	5 g
NaCl	-	5 g
Agar	-	20g
pH adjusted to	-	7± 0.2
Distilled water	-	make up to 1000ml

II. Reagent used for competent cell preparation

Solution A

Ice - cold 100mM $CaCl_2$

Abstract

MOLECULAR CHARACTERIZATION OF GEMINIVIRUS CAUSING YELLOW VEIN MOSAIC IN PUMPKIN

(*Cucurbita moschata* Duch. Ex Poir.)

By

Sahna Hamsa N. H

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the

requirement for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680 656

KERALA, INDIA

2011

size of 1363bp were designed to get full length ORF (846bp) of movement protein gene.

The technique was validated with DNA from 15 PYVM infected and 4 healthy pumpkin leaf samples collected from Palakkad, Thrissur and Malappuaram districts. The virus was detected only in the diseased samples. Hence these primers could be used in developing a molecular diagnostic tool to detect the virus. PCR amplifications were carried out in weed plants like *Emilia sonchifolia, Ageratum conyzoides, Hibiscus surattensis* and *Synedrella nodiflora* and crop plants like okra and ash gourd with yellow vein symptom to check whether these plants serve as the collateral hosts of the virus. PCR amplification was also performed in bitter gourd with distortion mosaic symptom. No amplifications were obtained in plants other than pumpkin. Using the primers PYVMV was detected in apparently healthy mature leaves of infected pumpkin. Hence, these primers could be used to detect latent infection.

Sequence and phylogenetic analysis of coat protein and movement protein gene sequences showed maximum similarity to bipartite Squash leaf curl china virus (SLCCV) from Coimbatore. Hence Pumpkin yellow vein mosaic virus (PYVMV) from Kerala can be considered as a strain of SLCCV.

MOLECULAR CHARACTERIZATION OF GEMINIVIRUS CAUSING YELLOW VEIN MOSAIC IN PUMPKIN (*Cucurbita moschata* Duch. Ex Poir.)

By

Sahna Hamsa N. H

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the

requirement for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680 656

KERALA, INDIA

2011

ABSTRACT

The study entitled "Molecular characterization of geminivirus causing yellow vein mosaic in pumpkin (*Cucurbita moschata* Duch. Ex poir.)" was carried out in the Molecular Biology Laboratory of Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara during the period 2009-2011. The objectives of the study included the molecular characterization of geminivirus causing yellow vein mosaic disease in pumpkin and developing a PCR based diagnostic kit.

Yellow vein mosaic infected pumpkin leaf sample was collected from the field of Olericulture Department, College of Horticulture, Vellanikkara and total genomic DNA was extracted by CTAB method. Specific primers for coat protein and movement protein genes were designed based on the sequences of geminiviruses infecting vegetables, obtained from the NCBI database. PCR amplification was carried out using these primers. Amplicons of size ~900bp and ~700bp were obtained for coat protein and movement protein genes respectively.

The purified PCR products were ligated in pGEMT plasmid vector and cloned. The recombinant *E. coli* cells were selected based on blue white screening on LB agar containing ampicillin layered with X-gal and IPTG. After confirmation of the inserts by colony PCR, the clones were sequenced. Sequences of 891bp and 702 bp were obtained for coat protein and movement protein genes respectively.

Sequence analysis was carried out with standard bioinformatics tools. On blastn analysis both coat protein and movement protein sequences showed maximum similarity to Squash leaf curl China virus (SLCCV) from Coimbatore. For coat protein gene, full length ORF of 771bp was obtained and the ORF of movement protein was partial. Primers MP1F and MP1R with expected amplicon

size of 1363bp were designed to get full length ORF (846bp) of movement protein gene.

The technique was validated with DNA from 15 PYVM infected and 4 healthy pumpkin leaf samples collected from Palakkad, Thrissur and Malappuaram districts. The virus was detected only in the diseased samples. Hence these primers could be used in developing a molecular diagnostic tool to detect the virus. PCR amplifications were carried out in weed plants like *Emilia sonchifolia, Ageratum conyzoides, Hibiscus surattensis* and *Synedrella nodiflora* and crop plants like okra and ash gourd with yellow vein symptom to check whether these plants serve as the collateral hosts of the virus. PCR amplification was also performed in bitter gourd with distortion mosaic symptom. No amplifications were obtained in plants other than pumpkin. Using the primers PYVMV was detected in apparently healthy mature leaves of infected pumpkin. Hence, these primers could be used to detect latent infection.

Sequence and phylogenetic analysis of coat protein and movement protein gene sequences showed maximum similarity to bipartite Squash leaf curl china virus (SLCCV) from Coimbatore. Hence Pumpkin yellow vein mosaic virus (PYVMV) from Kerala can be considered as a strain of SLCCV.