A STUDY ON siRNA MEDIATED RESISTANCE AGAINST BEGOMOVIRUS(ES) CAUSING PAPAYA LEAF CURL DISEASE

Thesis submitted to
Department of Biotechnology
School of Biosciences & Biotechnology
Babasaheb Bhimrao Ambedkar University,
Lucknow-226025
India

For the Degree of

Doctor of Philosophy

in

BIOTECHNOLOGY

Submitted by

Saurabh Verma

Enrolment No. : 004/14

Supervisor
Prof. Sangeeta Saxena
Department of Biotechnology
Babasaheb Bhimrao Ambedkar University
Lucknow-226025

Co-Supervisor
Dr. Pradhyumna Kumar Singh
Principal Scientist
Plant Molecular Biology & Genetic Engineering
CSIR-National Botanical Research Institute
Lucknow-226001

2018

Dedicated to

Almighty God

&

my family

UNDERTAKING FROM THE CANDIDATE

I, **Saurabh Verma**, have presented this thesis to enlighten the topic with whatever little I known.

This thesis mainly deals with the topic "**A STUDY ON siRNA MEDIATED RESISTANCE AGAINST BEGOMOVIRUS(ES) CAUSING PAPAYA LEAF CURL DISEASE**" . I have mainly focused on the development of a generic siRNA strategy against a group of begomoviruses causing papaya leaf curl disease. Proper care has been taken on my behalf to ensure that no mistake has been committed in this thesis.

(Saurabh Verma)

Declaration

I, **Saurabh Verma**, hereby declare that the research work embodied in the thesis entitled "**A STUDY ON siRNA MEDIATED RESISTANCE AGAINST BEGOMOVIRUS(ES) CAUSING PAPAYA LEAF CURL DISEASE**" submitted by me for the award of degree of philosophy is my own original research work carried out by me under the guidance of Dr. Sangeeta Saxena and Dr. Pradyumna Kumar Singh.

I declare that I have faithfully acknowledged, given credit to and referred to the research workers wherever their works have been cited in the text and the body of the thesis. I further certify that I have not willfully lifted up some other's work, para, text, data, results, etc., or available at web-sites and included them in this Ph.D. thesis and cited as my own work. I declare that the content in this thesis is free from plagiarism.

Place: (Saurabh Verma)

Date: Enrollment no. 004/14

CERTIFICATE

This is to certify that the thesis titled "**A STUDY ON siRNA MEDIATED RESISTANCE AGAINST BEGOMOVIRUS(ES) CAUSING PAPAYA LEAF CURL DISEASE**" submitted by Mr. **Saurabh Verma** is an original research work and has not been previously submitted in part or full for the award of any other degree or diploma to this or any other University.

The thesis submitted to Babasaheb Bhimrao Ambedkar University, Lucknow satisfies all the requirements as stipulated in the Doctor of Philosophy (Ph.D.) regulations- 1999 as amended in 2008/2010/2013 and it is fit for submission and evaluation for the award of the degree of Doctor of Philosophy of the University.

Date:

Supervisor

Prof. Sangeeta Saxena
Department of Biotechnology
Babasaheb Bhimrao Ambedkar University
Lucknow-226025

Co-Supervisor

Dr. Pradhyumna Kumar Singh
CSIR-NBRI
Lucknow

Head of the Department

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY
LUCKNOW
ESTABLISHED 1996

बाबासाहेब भीमराव अम्बेडकर विश्वविद्यालय

(केन्द्रीय विश्वविद्यालय)

विद्या विहार, रायबरेली रोड, लखनऊ–226025

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(A Central University)

Vidya Vihar, Raebareli Road, Lucknow-226025

Letter No.-159**/COE/BBAU/2015** **Dated:** 16/06/15

Ph.D. Course Work Certificate

This is to certify that **Mr. Saurabh Verma**, Enrollment No. 004/14 Ph.D. Research Scholar, Department of Biotechnology of this University has successfully completed his Ph.D. Course work in the examination held during December, 2014.

(A.K. Maurya)
Deputy Registrar (Exam.)

Acknowledgements

I *thank the Almighty God for enriching my life with beautiful things in the world and providing me the spiritual and physical strength to face the quandaries of life.*

This thesis would not have been possible without her support, I truly acknowledge the guidance rendered upon me by my supervisor Dr. Sangeeta Saxena, an eminent Professor and a researcher in the field of Plant Virology. I thank her for the able guidance, constant supervision, keen interest and moral support, which she rendered me from time to time. It was her constructive criticism, fruitful discussion and critical analysis, which helped me to develop as a researcher and a scholar in the area of plant biotechnology and bioinformatics. She always encouraged me to work on the ideas igniting from my naïve mind and helped me to carry them till completion. I also appreciate her banters and her life experiences, which helped me to develop a better attitude towards life.

I would like to thank my co-supervisor, Dr. Pradhyumna Kumar Singh, an eminent scientist, a teacher and a dynamic personality who always strives for excellence in all his endeavors. I have learned a lot from his constructive criticisms, inquisitive questioning and the novel approach towards solving scientific problems. I also sincerely thank him for providing workspace and resources required for accomplishment of my research objectives in this thesis.

I express my deep sense of gratitude and appreciation to Late Dr. Shirish Anand Ranade, an eminent teacher, researcher and comrade who helped me to realize the importance of being human. I had long discussions with him that helped me in satisfying my scientific and philosophical inquisitiveness and helped me become better every day. I still feel his presence around me whenever I feel despondent, his ever refreshing smile brings happiness on my face. He unconditionally rendered his lab facilities and scientific suggestions to me whenever I needed.

I feel privileged in expressing my sincere regards and gratitude to the Vice Chancellor, Babasaheb Bhimrao Ambedkar University for allowing me to pursue my research goals and for providing necessary infrastructure for research.

I want to express my gratitude to the Head of Department, Biotechnology, BBAU, for admitting me to the Ph.D. program and providing all the necessary support and facilities during my course work and research tenure.

I would like to thank the Director, CSIR-National Botanical Research Institute, for allowing me to use their workspace and facilities crucial in accomplishing the objectives in this thesis.

I would like to thank Dr. Prabodh Kumar Trivedi, an outstanding scientist and a generous human being. He was always helpful and generous enough to spare some time despite his busy schedule. He has greatly influenced the scientific planning and execution of my experiments. He has also facilitated his laboratory and other resources, which immensely helped in achieving the goals of my thesis.

I sincerely express my gratitude to Dr. Samir V. Sawant, an eminent scientist, a great teacher and wise counsellor. I thank him for believing in my abilities and supporting me when I needed the most. I found him as an astounding teacher who is always ready to impart drops from his sea of

knowledge. I thank him for providing me early opportunities starting from my summer training until now.

I would like to sincerely thank Dr. Sunil G. Babu, Assistant Professor, Department of Biotechnology, BBAU, an amicable human being and a teacher for providing cordial atmosphere and helping me out when doubts creep into my mind. I will always relish his friendly company and a cheerful smile.

I would like to thank Dr. Hemant Kumar Yadav, for providing his laboratory space and facilities at CSIR-NBRI.

I would like to thank my colleagues of Lab 208 in BBAU, SAR lab (presently lab 418), Plant Genomics Lab and Proteomics Lab for providing a friendly environment, healthy scientific discussions and help rendered to me throughout my stay at BBAU and CSIR-NBRI.

I want to profoundly thank my friendsDr. Abu-Farhan, Dr. Prashant Kumar, Dr. Saurabh Pandey and Mr. Anshu Alok for providing constant support and encouragement to me.

I want to thank Mr. Moti Lal, Mr. Ram Kumar, Mr. Ravindra, Mr. Jaibhan, Mr. Krishna Kant, Mr. Ramu and Mr. Chaitram for being a great help in executing my experiments.

I would like to thank the Department of Biotechnology office staff Mr. Deep, Mr. Sahu, and Mr. Noor for their cooperation in the day to day proceedings. Especially, I am grateful to Mr. Deep for the pain he took in processing our requests and application from department to administration at University.

I would like to thank Dr. Smita Kumar, whose endless efforts, time, advice and support allowed me to wade through the most difficult phase of my life. She has been the pillar of strength for me and stood by me in the most arduous of situations. I fall short of words while acknowledging her role in completion of this thesis. All I can mention about her is that she is a loving, caring and an ideal human being who became my partner for life and changed it.

I feel to be really fortunate that I am writing this thesis as it would not have been possible without the untiring efforts of Dr. Sumit Kumar. He is one person who ran from pillar to post for me. He always stood beside me as a brother, a friend, a guide and spiritual healer. I have learned a lot and still learning many things from him. I owe my thesis and accomplishments to him.

Last but not the least, I would like to thank my families for the constant moral, affection and support,which enabled me to achieve my goals in life.

Words are sometimes hard to find when one tries to say thank you for something so priceless as loving criticism, considerate helpfulness and valuable guidance. Though facts must be evidently acknowledged and honest thankfulness must be unequivocally stated. This is what I have humbly attempted to do here.

(Saurabh Verma)

Table of Contents

List of Tables

List of Figures

Introduction

1. Introduction

Carica papaya L. also known as Papaya is a sole member of family *Caricaceae*. It is a fruit bearing tree with large deeply palmate multi-lobed leaves. The flowers mature into large, soft, and yellowish to red colored fleshy fruits. Papaya is native to Mexico and other subtropical areas such as Brazil and Colombia. Papaya growth requires high temperature range and evenly distributed precipitation throughout the year (Mishra et al. 2007). Papaya fruit is a popular table food with high nutraceutical and broad range of industrial applications in food processing industry. It is a major commercial tropical horticulture crop and cultivated on large scale in India, Brazil, South-East Asian sub-continent and Central American region (FAOSTAT 2013). Large-scale cultivation of papaya is hampered due to major disease outbreaks caused due to fungi, bacteria, phytopthora and viruses. The major pathogens of papaya include microorganisms causing powdery mildew disease, leaf blight, papaya mosaic, papaya ring spot, papaya leaf curl, foot rot and other such infections (De la Cruz Medina et al. 2003). The papaya leaf curl disease (PLCD) is the major threat to papaya cultivation causing severe damage in northern regions of India (Saxena et al. 2016). Papaya leaf curl virus (PaLCuV), which is a member of begomovirus group of plant infecting Geminiviruses causes PLCD. This group of virus has caused large-scale economic losses to the food grain and horticulture crop production in the world. PaLCuV has been identified to be transmitted through *Bemisia tabaci* Gennadius, also known as whitefly (Guo et al. 2015). Whiteflies along-with its various biotypes, spread begomovirus species across the globe in tropical and sub-tropical regions. The begomovirus-whitefly interaction complex has been regarded as a global threat to the world crop production. Whiteflies are potentially harmful to the crops such as tomato, papaya, cassava, tobacco, cotton, potato, chickpea, sugarcane, wheat, maize and other commercial crops providing the bulk of the global food platter (Luo et al. 2017). They cause damage to crops in two ways i.e. through direct feeding on the phloem sap and by transmission of Begomoviruses (Byrne and Bellows 1991).

PLCD not only affects leaf development but also alters development of whole papaya plant including premature fall of fruits. The plant shows stunted growth and low probability of fruit development and completion of its normal life cycle. The symptoms consequently result in drastic decrease in total papaya production. The production in 2012-13 was approximately 5381.7 metric ton, which reduced to less than 4681 metric ton in 2014-15 in India with upto 8% culmination of cultivation area (Horticulture Statistics Division, Department of Agriculture, Cooperation and Farmers Welfare, Ministry of Agriculture). Many physical, chemical and biological control strategies are in current practice as Farm Management techniques for control of PLCD. However, these strategies are unable to control the large-scale wipe out of papaya cultivation in north India. Therefore, the future control strategies are required to devise tools and techniques to identify, isolate and provide resistance to PLCD for healthy papaya cultivation in India.

The control strategies against begomovirus could be devised by understanding the molecular basis of their initiation, propagation and elimination inside the host cell i.e. papaya, and disease associated crops. Therefore, it is a prerequisite to dissect the PLCD genome and identify the genes and proteins responsible in host defense and virus infection. It is important to identify the plant defense pathways and understand the mechanisms by which the PLCD components i.e. DNA-A and betasatellites, suppress the plant defense system and invade the plant genetic organization (Pooggin and Hohn 2004). It would be beneficial in designing the antiviral strategy to produce PLCD resistant transgenic papaya. The first successful virus resistant transgenic papaya was produced against *Papaya Ring Spot Virus* (PRSV), utilizing the strategy based on RNA interference (RNAi) technology (Gonsalves 1998). RNAi utilizes the knowledge about molecular basis of virus infection and introducing molecular components such as anti-sense transgenes or targeted entry of designed small interfering RNAs (siRNAs) containing plant binary vectors (McIntyre and Fanning 2006; Velten et al. 2012). These are designed against the viral factors that lead to post-transcriptional suppression of the viral genes in plants. Therefore, elaborate study of PLCD in papaya holds a

promise to develop a successful technology to prevent the damage caused by PLCD infection in Indian sub-continent.

Objectives

1. **Collection of healthy and leaf curl infected Papaya samples.**
2. **Isolation of DNA from infected and healthy plant samples for molecular analysis.**
3. **PCR screening of infected leaf samples for presence of begomoviral genome.**
4. **Identification and selection of genomic region(s) from viral isolate(s) for siRNA based strategy.**
5. **Introduction of above region(s) in appropriate vector for siRNA based resistance against begomovirus(es) causing papaya leaf curl disease.**

Review of Literature

2. Review of Literature

2.1 Leaf curl disease of papaya and weeds form a complex

The *C. papaya* L. also called Papaya is a member of family *Caricaceae*, widely used for its nutraceutical and commercial value. It is native to southern Mexico and Central American region but now cultivated all over the world. India and Brazil are among the largest producing nations of the world. The fruit has very high nutritional value and has a daily-recommended allowance for Vitamin A, C, folate, riboflavin, niacin, thiamine, iron, potassium, calcium and fibers in it. The stems, leaves, fruit and roots of papaya plants are used in production of commercially important enzyme *papain*, which is widely used in food processing industry and is used as a medicinal supplement in case of digestive problems (Australian Government 2008). Due to above properties, various varieties of papaya are now grown all through globe from kitchen gardens to big farms in tropical and sub-tropical regions of the world. Such a magnanimous scale of cultivation attracts many infectious agents such as bacteria, phytopthora, fungi, insects and viruses.

Among all symptoms listed, the leaf curl, crinkle, crumple and distortion has common occurrence in India (Borah and Dasgupta 2012). The leaf crinkle symptom was first reported to occur in papaya (Thomas and Krishnaswamy 1939). The disease was identified to be caused by a geminivirus infecting tobacco (Nariani, 1956) and later *Papaya leaf curl virus* (PaLCuV) in Pakistan was reported in papaya (Nadeem et al. 1997). Symptoms such as stunted growth, crippled, darkened leaves, curled lamina, deformed leaf and stem shape, vein swelling and enation characterized this disease. The genome organization of PaLCuV was first reported to be composed of bipartite begomovirus DNA components (Saxena et al. 1998a,b,c). Thus, PaLCuV was identified as a bipartite begomovirus, a class of whitefly-transmitted geminiviruses (WTGs) (Deng et al. 1994, Brown et al. 2015). Later, studies have revealed the monopartite nature of PaLCuV infecting papaya in different regions of China (Zhang et al. 2010), National Capital Region (NCR), India (Singh-Pant et al. 2012) also its association with weeds was presented in northern regions of India (Srivastava et al. 2014).

Therefore, this disease is now considered to be caused by a monopartite begomovirus group, which is found to be associated with smaller satellite molecules.

A molecular study was conducted related to spatial distribution of leaf curl disease of papaya in NCR, India. Typical leaf curl samples were collected from infected and healthy papaya plants. Sequence based investigations revealed that the leaf curl disease of papaya was actually not a single virus disease but was caused by various other begomoviruses like *Chili leaf curl virus* (ChLCuV), *Tomato leaf curl virus* (ToLCV), *Papaya leaf crumple virus* (PapayaLCV), *Pedilanthus leaf curl virus* (PedLCV) and *Ageratum leaf curl virus* (AgLCV) (Singh-Pant et al. 2012). On the other hand, PaLCuV is known to cause leaf curl disease in tomato, cucurbits, jatropha and other weeds (Verma and Saxena 2017a). Generally, disease caused by a group of similar or mixed type of viruses is considered a "disease complex". This term is justified in this case, due to the combined effect of the interaction of various papaya leaf curl disease (PLCD) causing viruses inside their hosts and the level of molecular diversity arising due to their genetic mixing. Therefore, we prefer referring this disease a PLCD complex rather than PLCD.

2.2 The genetic basis of Begomovirus causing PLCD complex

The group of plant infecting viruses called, Geminiviruses, cause various symptoms in plants leading to reduced productivity and developmental anomalies. It is a group of plant infecting single stranded (ss) DNA viruses in which the genome is carried inside a geminate shaped capsid, hence the name derived as Geminivirus (Gemini= geminate). Earlier, the Geminivirus family was sub-divided into genus *Mastrevirus, Curtovirus, Topocuvirus* and *Begomovirus* on the basis of number of genomic components i.e., Monopartite containing a single component genome e.g. *Maize streak virus* and Bipartite containing a two component genome e.g. *Tomato Golden Mosaic Virus* (King et al. 2012). Later, five new genera were added to geminivirus viz. *Becurtovirus, Eragrovirus, Turncurtovirus, Capulavirus* and *Grablovirus* due to their distinct genomes, host and transmission vector (Varsani et

al. 2014, 2017). There are few viruses which have a half-size defective components i.e., ss-DNA satellites encapsidated into the isometric virions e.g. *Tomato Leaf Curl New Delhi virus*.

Begomoviruses, a subgroup III type of Geminivirus, have been reported to cause leaf curl symptoms in papaya. These plant viruses are characterized by single stranded genome; geminate (twinned) particles made up of 110 protein subunits and appears as pentameric capsomeres. Begomovirus replicate inside the host nucleus via formation of a double stranded DNA (ds-DNA) intermediate through rolling circle mechanism. The intermediate ds-DNA is present in open and supercoiled circular forms and complex with histone proteins to form viral minichromosomes (Gutierrez et al. 2004). The single stranded-DNA (ss-DNA) ranges from 2.5 kb to 2.9 kb in genome size and present as monopartite (DNA-A) or bipartite (both DNA-A and DNA-B) (Stanley et al. 2005). These viruses are capable of replication in both cases due to the presence of multifunctional protein encoding open reading frames (ORFs), which form a variety of interactions with host cellular machinery for their own multiplication, movement and defense related functions (Fondong 2013). Begomovirus' helper DNA i.e. DNA-A, contains 4-6 ORF's named as AC and AV when present on complementary sense and viral sense strand respectively, in case of bipartite begomoviruses. In monopartite begomoviruses, the ORF's are named as C and V when present on complementary and viral sense strands respectively. In case of a disease complex, where both mono- and bipartite begomoviruses are involved, the AC and AV terms shall be used to denote these multifunctional ORF's. A 200 nucleotide long intergenic region (LIR) with hairpin-like loop containing origin of replication, nonanucleotide conserved motif and iterons serving as promoter and enhancer elements (Borah et al. 2016).

A monopartite begomovirus causing a severe leaf curl symptom, comprises of DNA-A component (2.7kb in size) and a betasatellite (DNA-β) corresponding to genome size of 1.3 kb. Both components share approximately 200bp sequences within the intergenic region. The DNA-A virion sense strand encode coat protein (CP) gene AV1 and AV2 and the complementary-sense strand

encodes replication associated protein (Rep) gene AC1 and transcription activator protein (*TrAP*) gene AC2, a replication enhancer protein (*REn*) gene AC3 and a host silencing suppressor protein called AC4. DNA-β contains a single coding region termed DNA βC1 protein gene, which suppresses the host RNA silencing machinery. The β-satellite does not have much sequence similarity to DNA-A except that it encodes a recognizable begomoviral sequence from the stem-loop and TAATATTAC sequence. Therefore, DNA-β is dependent upon the DNA-A portion of begomovirus genome for its replication and encapsidation, which is an essential condition for the maintenance of disease in the papaya fields. The genome of almost all the betasatellites contains a ~120-nts long highly conserved region, called satellite conserved region (SCR) and an Adenine rich (A-rich) region (~57-65% A content), typically ~160 – 280-nts long. The A-rich region was earlier reported to be important in maintaining the structural integrity of betasatellites. Later on, the mutations based studies proved it to be important in functioning of βC1 ORF, helping it in betasatellite genome accumulation and its expression regulating symptom severity (Zhou 2013). Almost all begomoviruses associated with PLCD complex were found to be associated with betasatellite molecules (Singh-Pant et al. 2012; Verma and Saxena 2017a, 2018). Therefore, betasatellites are an important 'infection partner' of DNA-A in case of PLCD complex.

2.3 Molecular functions of begomovirus proteins causing PLCD

The replication and transcription are two most important biological functions required for multiplication and survival of begomovirus inside their host cell. The replication of begomovirus ss-DNA occurs inside the host nucleus via Rep dependent rolling circle replication (RepRCR) mechanism involving open, linear double stranded (ds) DNA replicative intermediates. The RepRCR involves formation of a complementary anti-sense DNA synthesis on the virion sense strand to produce ds-DNA utilizing the host DNA replication factors. The virion ss-DNA synthesis is initiated immediately downstream of 3' thymidine residue present in the conserved TAATATT/AC sequence region within intergenic region by the *replicase* (Rep) protein which is a virus encoded replication

associated protein. The viral DNA replication model involves activation of Rep protein, which initiates release of ss-DNA template from the virion into the nucleus. RNA primer and host DNA polymerase synthesize the complementary strand. The ds-DNA gets associated with the host nucleososmes and form minichromosomes. The host RNA polymerase II (Pol II) then transcribes Rep mRNA and other proteins assisting in viral DNA replication and transcription. Rep protein with the help from PCNA complex binds *v-ori*, the starting point of replication present in IR. Thus, producing multiple viral DNA particles inside host nucleus by 'nick and join' mechanism (Stenger et al. 1991; Pilartz and Jeske 2003; Gutierrez et al. 2004). Thus, the RepRCR mechanism is an important step in the introduction of a begomovirus systemic infection to the plant and its' propagation inside plant cells. Therefore, Rep is the only essential viral protein required for its replication inside the host nucleus.

The transcription of viral mRNA requires promoter and enhancer elements, which are present in IR of viral DNA. IR is a mosaic region containing transcriptionally active elements (TAEs) and iterons responsible for initiation of bidirectional transcription of viral mRNAs (Shivaprasad et al. 2005; Borah et al. 2016). Transcription activator protein (*TrAP*) expression is controlled by a promoter present in upstream Rep coding region of DNA-A. Its' role has been reported in activation of transcription of CP and other movement related proteins by suppressing host viral silencers (Bisaro 2006). It is reported to interact *in vitro* with Arabidopsis *H3K9me2 histone methyltransferase* and inhibits its activity, thus reducing methylation marks and suppressing transcriptional gene silencing (TGS) (Castillo-González et al. 2015). *TrAP* also regulates the CP promoter expression by suppressing host suppressor proteins i.e. those participating in defense pathways (Sunter and Bisaro 1997; Berger and Sunter 2013). It performs transactivation in a manner similar to Adenovirus E1A protein that does not require it to bind specifically with related ss- or ds-DNA (Ferrari et al. 2008). Thus, *TrAP* has an important role in gene activation, virus pathogenicity and host RNAi machinery suppression. AC3 (C3) also known as Replication Enhancer protein (*REn*) encodes the symptom

severity factor. It enhances viral DNA replication by interacting with Rep and PCNA, thus assist in RepRCR (Jeske et al. 2001; Settlage et al. 2005).

AC4, the least conserved protein of begomoviruses has divergent functions such as symptom development, virus movement and host RNAi silencing. Its' role in vein swelling phenotype was reported to be caused by abnormal cellular development and differentiation (Stanley and Latham 1992; van Wezel et al. 2002; Mills-Lujan and Deom 2010). It causes vein swelling due to its interaction with various cell cycle regulator e.g. RKP in tobacco and ATHB7 and ATHB12 in Arabidopsis (Lai et al. 2009; Park et al. 2010). The RNAi suppression is an important function of AC4 protein found in *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV)-like viruses. AC4 inhibits RNAi by binding with miRNAs and siRNAs (Vanitharani et al. 2004; Chellappan et al. 2005) and by interactions in post translationally modified N- and C-terminus region (Fondong et al. 2007; Hipp et al. 2016; Roshan et al. 2018).

CP is the only structural protein encoded by a begomovirus ORF, its primary function is to provide encapsidation to the viral ss-DNA synthesized during DNA replication (Klinkenberg et al. 1989). CP is known to be a determinant of insect transmission as the mutations in this region resulted into reduced transmission rate via whiteflies (Briddon et al. 1990; Ghanim 2014). Apart from structural function, CP helps in viral DNA movement in and out of nucleus in absence of other movement related proteins e.g. *Squash leaf curl virus* (SqLCV) can replace function of DNA-B encoded MP by CP gene product (Rojas et al. 2001; Priyadarshini et al. 2011). Overall, CP gene function has been conserved for structural function, whereas the multifunctional nature of this protein is economically utilized by begomoviruses from time to time. AV2 protein is also known as pre-coat protein as its start codon precedes AV1 ORF, however, it lays more into overlapping region. This protein is normally not essential for any particular function yet mutations might result in decrease in viral DNA accumulation inside host and loss of mutants resulting into survival of wild type variety only (Padidam et al. 1996; Mubin et al. 2007). Its role in RNA silencing suppression in

case of EACMV and *Ageratum yellow vein virus* (AYVV) has been reported to play an important part in host RNAi suppression (Chowda-Reddy et al. 2008; Roshan et al. 2018).

βC1, the only protein encoding ORF of betasatellite is dependent upon DNA-A for its replication and transcription (Briddon and Stanley 2006). There have been evidences of its role in symptom development e.g. *Tomato yellow leaf curl China betasatellite* (TYLCCNB) by inducing abnormal cell division (Cui et al. 2005; Guan and Zhou 2006) and by multimerization (Cheng et al. 2011). Symptom development in case of *Radish leaf curl betasatellite* (RaLCB) was found to be through downregulation of chlorophyll synthesis related genes rather than structural genes of chloroplast (Bhattacharyya et al. 2015). The pathogenicity severity was distorted in βC1 mutated viral transgenes thus, the symptom development could be attributed to the formation of complexes with host defense related transcription factors in order to regulate viral replication (Yang et al. 2008; Eini et al. 2009). Betasatellite C1 protein has been heavily linked to host PTGS suppression due to its DNA and RNA binding activity in TYLCCNB, Cotton leaf curl Multan betasatellite (CLCuMuB), *Bhendi yellow vein mosaic betasatellite*, etc. (Gopal et al. 2007; Kon et al. 2007; Qazi et al. 2007). Studies have supported the role of βC1 in suppression of methylation mediated transcriptional gene silencing (TGS) (Raja et al. 2010; Yang et al. 2011a,b; Rogans et al. 2016). The reversal of cytosine methylation susceptibility in case of TYLCCNV and *Beet curly top virus* L2- mutants when co-inoculated with betasatellites proves its role in reversal of TGS (Sunter et al. 2001; Zhang et al. 2011). βC1 has a nuclear localization signal called 49KKK51, which interacts with *S-adenosyl homocysteine hydrolase* (SAHH) to reduce SAHH activity. SAHH is responsible for maintenance of methyl production via *S-adenosyl methionine* enzyme. Therefore, decrease in availability of SAM would eventually result in suppression of epigenetic mechanism of regulation of TGS (Yang et al. 2011b; Pooggin 2013). In an experiment in which ToLCNDV DNA-A when co-inoculates with a defective CLCuMuB particles produced hindrance in the DNA-A molecules movement. Later, it was reported that βC1 localizes around nucleus and co-localizes with endoplasmic reticulum network

around cellular margins (Saeed et al. 2007; Patil and Fauquet 2010; Saeed 2010). Thus, βC1 protein of betasatellites has evolved to be multifunctional in nature and adapts to the host and molecular machinery of its helper DNA i.e. DNA-A of begomovirus. This property of betasatellites has allowed the PLCD complex begomoviruses to infect tobacco, tomato, chili and cucurbits in natural condition and *Arabidopsis* when infected in controlled conditions; use of weeds as their reservoir host when the replicative primary host is absent. Therefore, the host range continuously expands as they spread from one geographic region to another.

2.4 Origin of host diversity in PLCD complex

Early leaf curl infection in papaya was limited to PaLCuV only, but later on reports of it's occurrence in different parts by various other monopartite begomoviruses opened the theory of "one disease - too many viruses", thus forming a disease complex. This broad host and begomovirus range probably resulted out of various interspecific recombinations. The recombination and mutational dynamics is the main reason behind most of the molecular events leading to emergence of new species causing severe symptoms. The recombination events arising due to the RepRCR tend to occur at high frequency in the viral origin of replication (*v-ori*). This region acts as a recombination *hot-spot* due to the natural tendency of double strand repair phenomenon by nick and join mechanism by Rep protein. Also, the monomeric single strand units of begomoviruses are released from large sized concatamers in this region. However, the above phenomenon was absent in betasatellite molecules indicating the significance of DNA-A in enhancing the genetic diversity and emergence of newer begomoviral species. In an experiment carried over 120 days, *Tomato yellow leaf curl virus* and *Tomato leaf curl Comoros virus* were co-infiltrated and the pattern of recombination was studied. It was observed that large number of recombination events occurred in genetically similar regions, in regions leading to strengthening of a particular protein-protein, protein-DNA or long range interactions and in events which lead to natural selection of the phenotype/function helping begomovirus to overcome external obstacles in its survival (Martin et al.

2011). Although, the exact pattern of physical recombination events i.e. recognition of physical recombination boundaries are not explored in depth, it should be explored in future.

2.5 RNA interference: PTGS and TGS

2.5.1 Posttranscriptional Gene Silencing (PTGS)

RNA interference (RNAi) is a mechanism involving inhibition of mRNA production through production of small homologous RNA molecules ranging from 21 – 25-nts in size called small RNAs (sRNAs) . It is highly specific, effective and thermodynamically favorable mechanism. It was first reported by Baulcombe and Hamilton (1999) as a phenomenon of production of 25-nts long dsRNA molecules from their dsRNA precursor regions. This phenomenon was later termed as Posttranscriptional gene-silencing (PTGS), which provides innate immunity to plants against external organisms, transgenes, transposable elements and viral mRNAs. RNA dependent RNA polymerase (RDR), DICERs or DICER-LIKEs (DCL) and ARGONAUTE (AGO) proteins are important enzymes having a key role in plant sRNA biogenesis and function. There are six RDRs in *A. thaliana*, which function in synthesis of dsRNA precursor (Willmann et al. 2011). DCLs are endonucleases, which catalyze dsRNA precursors to release short dsRNA duplexes (Xie 2005; Margis et al. 2006; Liu et al. 2009). AGO then engages this short dsRNA duplex and releases one of the RNA strand and results in mRNA-RISC-RNA adduct formation in highly specific manner to its target mRNA sequence (Fang and Qi 2016). PTGS involves processing of precursor dsRNA by DCLs to produce sRNA duplexes. It is followed by capture of guide sRNA strand by ARGONAUTE (AGO) proteins to form RNA induced silencing complexes (RISC) (Rand et al. 2005; Leuschner et al. 2006; Ameres et al. 2007). The above process results in recognition and binding of RISC to its homologous perfectly complementary sequence in mRNA region resulting into cleavage of mRNA into 21 – 25-nts molecules called micro RNAs (miRNAs) and small interfering RNAs (siRNAs) belonging to primary sRNAs (Lu et al. 2008). However, when the complementarity is not perfect,

then the preferred mechanism is through RISC mediated translational inhibition of target mRNA transcription (Brodersen and Voinnet 2009).

These small non-coding RNAs i.e. miRNAs and siRNAs lie at the pinnacle of gene expression regulation and have been reported to be the master regulators of plant growth and development and in plant responses against various abiotic and biotic stresses. The mechanism of regulation by miRNA is mostly through downregulation of mRNAs involved in a particular function e.g. miR394 targets LCR functioning in regulation of shoot meristem (Knauer et al. 2013); miR156 targets SPL family mRNAs promoting flowering and tillering (Aukerman 2003; Wang and Wang 2015); miR165/166 in maintaining procambial activity and lateral root growth (Sakaguchi and Watanabe 2012; Jia et al. 2015). However, siRNAs mainly take part in conferring innate immunity to plants against various pathogenic infection including diseases caused by plant viruses (Axtell et al. 2006; Aregger et al. 2012; Creasey et al. 2014; Borges and Martienssen 2015). The endogenous siRNAs are produced by action of RDR1 – RDR2 in concert with DCL2 – DCL4 on long non-coding RNAs, primarily producing primary siRNA and secondary siRNAs. Primary siRNAs include heterogeneous siRNA (hetsiRNA), a 24-nts long primary siRNA processed by DCL3. The secondary siRNA include phased or trans-acting siRNA (tasiRNAs) and natural anti-sense small interfering RNAs (NAT-siRNAs), which are 21 and 22-nts long secondary siRNAs processed by DCL4 and 2 (Axtell 2013). The sRNAs thus produced are processed and modified in plants at the 3'-end by mainly *2'-O*-methylation, which is important for their stability and protection from degradation pathways involving 3'-uridylation mechanism (Ibrahim et al. 2010; Scheer et al. 2016).

2.5.2 Transcriptional Gene Silencing (TGS)

Plants take part in RNA silencing through PTGS (*refer* previous section) and also transcriptional gene silencing involving RNA dependent DNA methylation (RdDM) causing CG, CHG and CHH methylation marks on foreign transgenes to direct them for transcriptional repression termed as TGS. The TGS mechanism involves 24-nts long hetsiRNA contributing to identification and suppression

of methylation imprints at 5-methyl cytosine residues, made by DNA METHYLTRANSFERASE (DNMT) class of enzymes. This methylation pattern is then maintained synergistically by DEFICIENT IN DNA METHYLATION 1 (DDM1) and RdDM in plants. The heterochromatin is formed only after sustenance of DNA methylation, through activation of H3K9 METHYLTRANSFERASE enzyme (Zvereva and Pooggin 2012; Pooggin 2013).

The above canonical pathway of RdDM could not properly explain the trans-generational inheritance of DNA methylation patterns. Therefore, several non-canonical mechanisms have been proposed to explain TGS in plants (Raja et al. 2010). The non-canonical pathways involving miRNA-directed DNA methylation (miRNA-dDM) i.e. pol II derived inverted repeat (IR) sequences gets cleaved into a 24-nts siRNAs and produces RdDM in *trans* manner and epigenetic transcriptional repression of the rest of gene family e.g. Mutator family TE in maize. miRNA-dDM can occur in both cis- and trans- direction to cause TGS. A RDR6 RdDM based pathway involves TAS genes, which facilitate miRNA mediated processing of dsRNA by RDR6 enzyme to produce 21 – 22-nts siRNA, to repress transposable elements (TE). Apart from above two mechanisms non-canonical pathways take part in TGS (Cuerda-Gil and Slotkin 2016). Therefore, both canonical and non-canonical pathways fill in the voids in biogenesis of siRNA involved in TGS, which is now identified as an important mechanism involved in plant resistance against a geminivirus infection.

2.6. Suppression of RNAi by begomovirus is an important mechanism for its survival

Both RNA and DNA plant viruses have evolved several mechanisms to override host defense pathways for their own survival, transmission and symptom development (Burgyán and Havelda 2011). RNA virus i.e. *Potyvirus* was identified to inhibit established plant RNA silencing through HC-Pro protein, which was found to inhibit host immunity by interacting with rgsCaM and inhibiting dsRNA processing resulting into failure to produce effective siRNA population. Similarly,

geminivirus transcripts from DNA-A and betasatellite molecules are known to take part in suppression of both PTGS and TGS mode of RNA interference.

DNA-A component of begomovirus genome has been widely related to suppression of PTGS and TGS in host plants. Mainly, the suppression phenomenon occurs through mechanisms involving AC2 (*TrAP*), AC4 alone or AC4/AC2 partnership. Local RNAi suppression by *Mungbean yellow mosaic virus* (MYMV) *TrAP* is induced through transactivation of viral promoters and activation of host RNA silencing suppressor proteins e.g. WEL1 (Werner Exonuclease-Like 1) (Trinks et al. 2005). In another experiment, *Tobacco golden mosaic virus* (TGMV) AC2 protein inhibited TGS mechanism by interfering with proteins responsible for maintenance of DNA methylation. Here, the AC2 protein interacts with adenosine kinase (ADK) protein functioning in methyl group transfer during RdDM. From microscopic evidences, AC2 subcellular localization is known to be influenced by phosphorylation induced by ADK. Therefore, an ADK-deficient host plant shows failure in TGS mediated silencing. Other proteins of TGS associated with AC2 mediated suppression are HEN1, SAHH, homology dependent gene silencing 1 (HOG1) and SNF1 (by inactivation of ADK). AC2 protein of *Mungbean yellow mosaic Indian virus* (MYMIV) interacts with RNAi components i.e. RDR6 and AGO1, thus, inhibiting synthesis of tasiRNAs mediated silencing (Lózsa et al. 2008; Kumar et al. 2015). The synergistic interaction of *African cassava mosaic virus* (ACMV), a recovery type virus with AC4 suppressor and *East African cassava mosaic virus* (EACMV) (non-recovery type) with strong AC2 suppressor activity were able to cause severe infection of cassava plantations. Another study revealed a newly emergent EACMV-Cameroon detected in samples showing mixed symptoms, proposed to form a similar synergism as EACMV-AC2 component (Fondong et al. 2000; Vanitharani et al. 2004). A study has shown that AC4/Rep synergism in ACMV and *Tomato leaf curl China virus* (TYLCCNV) was responsible for appearance of necrotic symptoms, which otherwise was absent when only AC4 component was expressed under constitutive promoter expression (van Wezel et al. 2002). Further studies are required to elucidate mechanistic aspects of

the synergistic interaction of AC4 with Rep and AC2 to get in-depth insight into the RNA silencing suppression mechanism.

βC1, a nuclear localized protein from TYLCCNV, is required for TGS reversal and DNA methylation (Zhong et al. 2017). It acts through inactivation of SAHH, resulting in reduction of 3'-cytosine methylation. It is a functional substitute to mutated C2 protein of TYLCCNV (Yang et al. 2011b). βC1 of *Cotton leaf curl Multan betasatellite* (CLCuMuB) binds with DNA and RNA, reported to interfere with binding interaction of dsRNA, sRNAs, thus inhibiting PTGS (Qazi et al. 2007; Amin et al. 2011; Tiwari et al. 2013). Thus, βC1 is an important begomoviral component important for symptom severity through suppression of TGS and PTGS (Zhou 2013).

AV2, functionally least characterized begomovirus protein of *Velvet bean severe mosaic virus* is a strong repressor of RNA silencing when compared to *TrAP* and CP (Jangid et al. 2017). The Rep protein is a known cell cycle modulating component of begomoviral machinery. Rep downregulates MET1, CMT3 and ROS1; thus triggering host DNA reduplication and resulting into reduced cytosine methylation levels in cell. Therefore, Rep mediated rescue of begomovirus replication from methylation-sensitive repression is an important augmentation of viral RNA silencing suppressor machinery (Rodríguez-Negrete et al. 2009, 2013). There is still insufficient information related to role of various begomoviral proteins taking part in host PTGS and TGS suppression. But, it is an important mechanistic aspect of begomovirus machinery which provides it an edge to virus in the molecular arms race of host plant against viral infections (Sahu et al. 2014).

2.7 RNAi strategies: Past, present and tomorrow

Initially, discovered as a 'co-suppression' phenomenon in Petunia plants overexpressing *chalcone synthase* gene. Co-suppression resulted into depigmentation of flowers as a consequence of reduction in transgene and its mRNA expression levels (van der Krol et al. 1988; Napoli 1990). Later, various studies identified that coat protein mRNA of plant-infecting viruses have potential to provide partial

to complete resistance when introduced in anti-sense manner (Beachy et al. 1990; Gonsalves 1998; Krubphachaya et al. 2007). Transgenic papaya carrying coat protein mediated resistance against *Papaya ringspot virus* (PRSV) is a classic example of the level of resistance provided by anti-sense RNA technology (Gonsalves 2006). However, the success of this technology could not be guaranteed against geminivirus. An approach involving simultaneous expression of viral segments in sense and anti-sense manner separated by an intron was found to produce transgenics showing potency against geminivirus (Varsha Wesley et al. 2001; Hirai and Kodama 2008). The principle behind above approach was to induce the RNAi machinery of plants to synthesize 21 – 24-nts siRNA against a specific region by mimicking it as a hairpin containing dsRNA precursor fragment. Thus, the natural RNAi machinery of plants, which generally produce decoys of primary and secondary virus siRNAs (vsRNAs) from putative hot-spot regions of viral genome was diverted to produce specific and efficient targeted siRNAs to produce geminivirus resistant transgenics (Younis et al. 2014). As per current scenario, several cases in India and around the world, have been reported, where the siRNA strategy has found success in conferring resistance. The siRNA harboring transgenic plants against *Bean Golden Mosaic Virus* (BGMV) (Bonfim et al. 2007), *Cotton Leaf Curl Virus* (CLCuV) (Hashmi et al. 2011), *Ageratum yellow vein virus* (AYVV) and *Melon yellow spot virus* (MYSV) (Yang et al. 2014) were developed and evaluated successfully. The transgenic plants against *Mungbean yellow mosaic virus* (MYMV) and *African cassava mosaic virus* (Vanitharani et al. 2003; Galvez et al. 2014) have found success in field trials. The insect resistance transgenic cotton plants were developed against *B. tabaci*, which indirectly inhibits transmission of geminiviruses (Thakur et al. 2014, Shukla et al. 2016). Interestingly, some of the transgenic horticulture crops produced using siRNA strategy were approved for human consumption e.g. *Bean golden mosaic virus* resistant Pinto bean (approved by Brazil, 2011), *Papaya ring spot virus* resistant papaya (approved by USA, 1996; Canada, 2003; Japan, 2011), high oleic acid soybean line MON87705 (approved by New Zealand, 2010).

Transgenic plants containing short and long-intronic hairpin containing viral transgenes and fragments became a popular approach from 2002 - 2016 as the levels of siRNAs produced were high and the resistance developed was sustainable to some extent (Pooggin 2017). However, occurrence of frequent recombination events resulted into breaking anti-viral resistance in transgenic plants (Pita et al. 2001; van der Walt et al. 2009). Recently, field trials of a TYLCV resistant tomato transgenic plant harboring Rep containing intron-hairpin siRNA construct was conducted till F4 – F6 transgenic generations (Fuentes et al. 2006). This transgenic tomato plant showed successful resistance against TYLCV but due to intervention by a previously unknown virus identified as *Tomato latent virus* (TLV), the resistance was suppressed (Fuentes et al. 2016). TLV was a recombinant virus inheriting DNA-A complementary sense strand ORFs from a distant begomovirus and viral sense strand ORFs from TYLCV. Various strategies have been employed against different variants of ToLCNDV, but all of them failed due to frequent emergence of a new recombinant or a new variant possessing mutation in viral genomes conferring resistance against RNAi constructs (Jyothsna et al. 2013; Sahu et al. 2014; Kumar and Naqvi 2016; Zaidi et al. 2016). RNAi initiative for developing transgenic cassava resistant against ACMV and associated cassava disease complex in Africa is underway (De Bruyn et al. 2016; Fondong 2017). Similarly, research is ongoing to develop transgenic cotton plants resistant against CLCuMV and associated cotton leaf curl disease complex viruses in Pakistan and India (Saeed 2010; Akmal et al. 2017; Zubair et al. 2017a; Shweta et al. 2018). In an interesting siRNA based strategy, transgenics with generic resistance were developed against *Pepper golden mosaic virus* (PepGMV) and *Tomato chino La Paz virus* (ToChLPV) using 700-800-nts upper half fragment of DNA-A carrying more than 85% sequence conservation (Medina-Hernández et al. 2013). This experiment provided a support to the generic approach involving long hpRNAi containing constructs, which when designed on the basis of sequence conservation of a particular region, have potential to provide a sustainable resistance against a genera of begomovirus like papaya, chili, tomato and cassava disease complexes.

Apart from RNAi based strategies, new tools have been discovered which act in a sequence specific manner and result in whole genome editing. This technique is termed *Clustered regularly interspaced palindromic repeats* (CRISPR) and forms a molecular editing system when adjunct with *Cas9* protein (Li et al. 2014). It is in exploratory stages at the moment, which when successful has potential to provide resistance through targeted mutagenesis and geminiviral genome editing (Cong et al. 2013; Zaidi et al. 2016; Abdullah et al. 2017; Gebre et al. 2018).

Therefore, begomovirus infection in India plays a crucial role in increasing suffering of farmers apart from natural calamities, which has severely affected food productivity. This situation has caused mass migration of farmers from agriculture to other profitable occupation and then we might face a national food crisis. Therefore, the immense loss caused by the begomovirus class of Geminiviruses needs special attention and this problem could be solved by restricting their national and global presence using siRNA strategy.

Materials & Methods

3. Materials and Methods

3.1 Chemicals and experimental material

Table 3.1 List of Laboratory chemicals and enzymes used in this study

Material	Source	Purpose
Plant Material		
Papaya (*Carica papaya* cv Pusa delicious)	National Seeds Corporation Ltd., New Delhi	Experimental plant
Tobacoo (*Nicotiana benthamiana*)	Plant Genomics Lab. CSIR-NBRI, Lucknow	Experimental plant
Bacterial Culture		
Escherichia coli (DH5α strain)	Invitrogen Life Technologies, Carlsbad, USA	Bacterial transformation
Agrobacterium tumefaciens (LBA4404 strain)	Invitrogen Life Technologies, Carlsbad, USA	Plant transformation
Vectors		
pTZ57R/T	Thermo Scientific	Cloning
pUC19	GE Lifecare	Cloning
pCAMBIA1301	Plant Genomics Lab. CSIR-NBRI, Lucknow	Cloning
pBI121	Plant Genomics Lab. CSIR-NBRI, Lucknow	Cloning
DNA Ladders and Markers		
Lambda HindIII/EcoR1 digested	Fermentas	Agarose Gel Electrophoresis
1kb plus Ladder	Invitrogen	Agarose gel Electrophoresis
Restriction Enzymes and DNA modifying enzymes		
Restriction endonucleases T4 DNA Ligase Calf Intestinal Alkaline Phosphatase (CIAP) Taq DNA polymerase	Fermentas	DNA modification and cloning
TempliPhi Amplification kit (100rx)	GE Lifecare	Viral DNA amplification
Real time reagents	ABI biosystems (Invitrogen)	
General Chemicals		
Electrophoresis grade chemicals	HiMedia Laboratories Sigma Chemicals	Genei Chemicals India
High purity HPLC purified chemicals and solvents	Sigma Chemicals Rankem chemicals	
Tissue culture medium, salts and hormones	Sigma Chemicals HiMedia Laboratories GIBCO Invitrogen Corporation	

3.2 Experimental conditions for various types of culture

3.2.1 Bacterial culture

E.coli cells were cultured in Luria Bertani (LB) broth (1% NaCl; 1% Tryptone; 0.5% Yeast extract dissolved in water, pH 7.0±0.2). The cells were incubated at 37℃ with 200-250 rpm shaking for 12-18 h. The bacterial colony forming unit (c.f.u.) were sub-cultured on LB-Agar (LA) plates containing 1.5% Agar with appropriate antibiotics added before pouring.

3.2.2 Agrobacterium culture

Agrobacterium strain was cultured on YEM (1% D-Mannitol, 0.01% NaCl, 0.04% yeast extract, 0.02% $MgSO_4$, 0.05% K_2HPO_4) broth while incubation was done at 28℃±2℃ for 2 days. The *Agrobacterium* c.f.u. were maintained on YEM–Agar (1.5%) medium containing petri plates with suitable antibiotics added just before pouring.

3.2.3 Plant growth conditions

The plants were maintained in controlled glasshouse conditions dedicated for optimal growth of tobacco and papaya plants. The light conditioning was set for 14 h day and 10 h night for tobacco and 16 h day and 8 h night for papaya. The temperature was maintained at 22-24℃ for tobacco and 28℃±3℃ for papaya with relative humidity range of 50-60% during day and night cycles. The tobacco plants were grown in plastic pots filled with Agro-peat/vermiculite mixture in 1:1 ratio. The papaya plants were grown in medium sized earthen pots containing clay: peat: vermiculite mixture in 1:1:1 ratio.

3.3 Database management and *in-silico* analysis

3.3.1 Viral genome retrieval and database management

The viral genome sequences were retrieved from National Centre for Biotechnology (NCBI) website (https://www.ncbi.nlm.nih.gov/) Nucleotide database (https://www.ncbi.nlm.nih.gov/ nuccore/). Search terms such as “leaf curl”, “begomovirus”, “DNA-A”, “DNA-B” and “betasatellite” were used

to retrieve the nucleotide sequences of interest. Annotation of the sequences in FASTA format was done for further analysis.

3.3.2 Multiple Sequence Alignment (MSA)

The procedure of multiple sequence alignment was performed using various algorithms available in MEGA version 6.0, a freely available software suite (Tamura et al. 2013). The large dataset of viral DNA sequences retrieved from NCBI was aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm (Edgar 2004) integrated into MEGA 6. The default parameters were selected to obtain alignment with highest accuracy and refined position information (Hall 2013). Smaller sized datasets of viral DNA sequences were aligned using CLUSTALW (Thompson et al. 1994) integrated into MEGA 6 software suite using default parameters.

3.3.3 Construction of Phylogenetic tree

The phylogenetic analysis was performed to infer evolutionary information of the DNA sequences retrieved for the terms as described in section 3.3.1. All DNA sequences were aligned using methodology as described in section 3.3.2. The input file saved in FASTA format was uploaded into TOPALi v2.5 software (www.topali.org/), an integrated platform to perform high-throughput, parallel computing on a distant server (Milne et al. 2009). This platform was used to perform evolutionary study of DNA-A sequences of retrieved begomovirus related to PLCD in Indian sub-continent. Bayesian algorithm and Maximum Likelihood (ML) algorithms are popular choices when analyzing distantly related, large number of population (Felsenstein 1981; Ronquist et al. 2012). Both of the above studies can be done using tools integrated into TOPALi v2.5 software suite i.e. MrBayes v.3.1.1 for Bayesian calculations and PhyML-aLRT v2.4.5 (www.atgc-montpellier.fr/phyml/) for ML based evolutionary study (Guindon et al. 2005, 2010). The parameter selection for this study was decided by performing a model test using Model Selection Tool available in TOPALi v2.5. The suitable model was set for both Bayesian and ML based study to infer phylogenetic tree depicting evolutionary distances between input DNA sequences with appropriate

statistical significance set by performing bootstrapping steps (Freedman 1981). The inferred trees from above studies were prepared for illustration using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

The ML algorithm available in MEGA 6 was used to infer evolutionary relationship between DNA-A and betasatellite sequences retrieved from NCBI website. The default parameters were employed to infer a statistically relevant phylogenetic association of DNA sequences of begomoviruses as required for the study of evolutionary distances and relationships.

3.3.4 Sequence Demarcation Tool analysis

Sequence Demarcation Tool (SDT) v1.2 (http://web.cbio.uct.ac.za/~brejnev/) (Muhire et al. 2014) was used for the taxonomic classification of DNA viruses infecting plants as it was recommended by International Committee on Taxonomy of Viruses (ICTV). This is highly automatable tool to classify a smaller dataset e.g. DNA-A and betasatellite components of PLCD complex, based on pairwise genetic identity of DNA sequences of new and unknown viral isolates. The viral DNA-A sequences retrieved in section 3.3.1 were used in FASTA format for the analysis in SDT v1.2. The output is illustrated in the form of a color-coded distance matrices to aid in the classification of viral sequences according to the ICTV approved demarcation criterion (Brown et al. 2015). The DNA-A sequences showing close percentage proximity with each other were grouped under one representative begomovirus causing PLCD. Similar analysis was carried out for analysis of Sultanpur isolate DNA-A component causing papaya leaf curl disease.

3.3.5 Recombination analysis

The recombination analysis was conducted to detect and analyze recombination breakpoints in a viral sequence using Recombination Detection Program ver.4 (RDP4) (http://web.cbio.uct.ac.za/ ~darren/ rdp.html). RDP4 program simultaneously uses various recombination detection methods to detect and characterize the recombination boundaries and hence an event, which arises due to a

possible recombination between two or more virus genomes (Martin et al. 2015). The input file containing all putative parental and recombinant viral DNA sequences was prepared as mentioned in section 3.3.1 and 3.3.2. The input file in *.fasta* format is used to feed the program with aligned DNA sequences. RDP4 has an OPTION toolbar, which contains various parameters set at default values. The parameters are optimized for most of the basic recombination analysis, but require careful selection for investigation of complex sequences such as large DNA sequence dataset containing DNA-A and betasatellite sequences. The parameters such as Methods and the basic operation settings were modified according to the nature of investigation. General analysis provided significant results in few cases, therefore, the guidelines for parameter selection was followed to perform the recombination analysis using RDP4 software. Various recombination detection programs available in RDP4 are RDP (Martin and Rybicki 2000), BOOTSCAN (Salminen et al. 1995; Martin et al. 2005), CHIMAERA (Posada and Crandall 2001), MAXCHI (Smith 1992), GENECONV (Padidam et al. 1999), SISCAN (Gibbs et al.s 2000), 3SEQ (Boni et al. 2007), VisRD (Lemey et al. 2009) and BURT (Martin et al. 2015). All recombination analysis were done for circular viral DNA, with highest acceptable p-values set at default for each program, breakpoint polishing selected with 'require topological evidence' setting ON to generate phylogenetically meaningful results. The output file was saved in default format i.e. .rdp and recombinant screening was done manually as described in the RDP4 software suite instruction manual (http://web.cbio.uct.ac.za/~darren/RDP4Manual.pdf).

3.3.6 *in silico* prediction of siRNA molecules

3.3.6.1 Selection of conserved regions

The DNA sequences retrieved and aligned in section 3.3.1-3.3.2 were investigated for the percentage conservation among them. Generally, the sequence conservation percentage is maintained at 100% to design a specific and effective siRNA. However, in case of generic approach, the conservation was kept between 70-85%.

3.3.6.2 Prediction of siRNAs

The conserved consensus regions were analyzed with pssRNAit software (http://plantgrn.noble.org/pssRNAit/), a platform for plant siRNA prediction. pssRNAit incorporates various algorithms for efficient and specific siRNA prediction with efficiency score based on the features such as homology, mRNA-siRNA binding energy, siRNA accessibility, secondary structure and off-target screening against query plant genome. It also provides score for the algorithms and rules used to predict each siRNA. The parameters were selected to design efficient and specific siRNAs against highly conserved consensus sequences and off-target filtering was performed against *Carica papaya* genome. The results were compiled into MS Excel file to carry out further investigations.

3.4 Sample collection

The *Carica papaya* L. leaf samples from plants showing upward curl, downward curl, vein swelling, leaf mottling, crumpling with darkened green coloration and stem swiveling were collected in zip lock bags containing silica pouches. The leaves were wrapped in aluminium foil and stored at -20℃ until further use. The samples were collected from several parts of Lucknow (26.8467° N, 80.9462° E), Sultanpur (26.2648° N, 82.0727° E), Bareilly (28.3670° N, 79.4304° E) and Malihabad (26.9168° N, 80.7076° E).

3.5 Routine laboratory protocols

3.5.1 Isolation of genomic DNA from plant tissue

Isolation of genomic DNA from fresh leaf tissue was done by modified CTAB procedure (Doyle and Doyle 1987). Preheated CTAB isolation buffer [2% hexadecyltrimethylammonium bromide (Bangalore Genei), 1.4 M NaCl, 20 mM Na-EDTA, 100 mM Tris-HCl, pH 8.0 and freshly add 0.2% 2-mercaptoethanol] was added to 0.2-0.5 g leaf tissue pulverized with a mortar and pestle in liquid nitrogen till a fine powder was obtained. The powder in CTAB buffer was incubated for 30 min in water bath at 65℃ with occasional gentle swirling. Gently, equal volume of chloroform-isoamyl alcohol (24:1) was added, mixed and centrifuged at 10,000 rpm for 5 min at room temperature (RT).

The aqueous layer solution was collected while rest was discarded. Total DNA and impurities were precipitated with ice-cold isopropanol (2/3 volumes) after centrifugation at above condition. The DNA pellet was washed twice with 75% alcohol and air-dried; dissolved in 1ml TE buffer [10 mM Tris-HCl+1 mM Na-EDTA; pH 7.5]. The DNA thus obtained was used for diagnostic purpose i.e. PCR based virus screening. Further purification procedure (see section 3.5.2) was employed to prepare DNA for sophisticated techniques.

3.5.2 Procedure for viral DNA enrichment

The genomic DNA isolated in section 3.5.1 was diluted with sterile deionized water to make final volume up to 0.5 ml. To this, RNase (20 µg/ml) was added and incubated for 30 min in water bath at 37℃ and then 0.2 ml Solution II [0.2 N NaOH+ 1% w/V sodium dodecyl sulphate (SDS, Sigma)] was added and incubated for 5 min with gentle mixing. Immediately, 0.3ml ice cold Solution III [60 ml 5 M K-acetate; 11.5 ml glacial CH_3COOH and water to final volume 100 ml] was added and incubated on ice for 20 min. The microfuge tubes were centrifuged at 12,500 rpm for 20 min at 4℃ in a fixed angle bench top centrifuge. The supernatant was collected while the white precipitate was discarded. The resulting solution was extracted sequentially with phenol:chloroform:isoamyl alcohol (25:24:1), chloroform:isoamyl alcohol (24:1) and then DNA was precipitated with 0.7 volumes isopropyl alcohol. The DNA pellet was washed thrice with ice cold 75% ethanol and air-dried in sterile environment. The dried DNA pellet was dissolved in 50 µl TE solution and stored at 4℃.

3.5.3 Rolling circle amplification (RCA) of viral DNA from genomic DNA

Rolling circle amplification (Richardson et al. 2003) of viral DNA was performed using phi DNA polymerase enzyme kit i.e. Templiphi™ 100 DNA amplification kit (GE Healthcare). 50 ng of enriched viral DNA was added to 5 µl Sample buffer (SB) and denatured at 95℃ for 3 min and immediately cooled on ice for 5-10 min. Reaction solution was prepared by adding 5 µl Reaction buffer (RB) and 0.2 µl enzyme mix. The final reaction was prepared by adding 5.2 µl reaction

solution and incubated at 30℃ for 16-18h in a thermal cycler. The reaction product was removed on completion and stored at -20℃ until further use.

3.5.4 Restriction enzyme mediated digestion of DNA and analysis

Plasmid/genomic DNA/RCA product (0.1-2 µg) was used for restriction enzyme mediated DNA digestion. A 50 µl reaction mixture was prepared containing 1unit/µg DNA of restriction enzyme and reaction components as per manufacturer's instruction. The mixture was incubated in water bath set at 37℃ for 30 min to 12 h as per requirement. The digested plasmid DNA was treated with Calf intestinal alkaline phosphatase (CIAP) enzyme, purified (see section 3.5.6) and stored at -20℃. For restriction analysis, 1 µg of RCA product was digested as mentioned above and resolved on 0.8% agarose gel. For confirmation of cloned inserts, 200 ng of total DNA was used for restriction analysis at above-mentioned conditions.

3.5.5 Polymerase chain reaction (PCR) amplification of nucleic acids

Polymerase chain reaction (Mullis and Faloona 1987) was used to amplify various viral DNA segments using forward and reverse primers. The PCR reaction mixture contains PCR buffer, primers, dNTPs, Taq/other commercial DNA polymerase enzymes, DNA template, $MgCl_2$ and

Table: 3.2 List of primers and PCR conditions used in this study

Primer name	Primer sequence (5'-3' direction)	Amplicon details	PCR conditions
UN1V 1979	ACCTACGCGGCCGGCATCTGCAGGCC CACATYGTCTTYCCNGT	Upper half	94℃=30s, 59℃=2m, 72℃=3m, 35cycles, 72℃=15m
UN1C715	ACCTACGCGGCCGGATTTCTGCAGTT DATRTTYTCRTCCATCCA	Upper half	Product size= 1.5-1.6 kb
NBL1V 2040	ACCTACGCGGCCGGCCTCTGCAGCAR TGRTCKATCTTCATACA	Coat protein	94℃=30s, 60℃=2m, 72℃=1m, 35cycles, 72℃=5m
NCRC1	ACCTACGCGGCCGCTAGCTGCAGCAT ATTTACRARWATGCCA	Coat protein	
NPCR154	ACCTACGCGGCCGGGTAATATTATAH CGGATGG	Coat protein	Product size= 520-600 bp
N1C1960	ACCTACGCGGCCGTGGACTGCAGACN GGNAARACNATGTGGGC	Bottom half	94℃=30s, 59℃=2m, 72℃=3m, 35cycles, 72℃=15m
NR1V722	ACCTACGCGGCCGATATCTGCAGGGN AARATHTGGATGGA	Bottom half	Product size= 1.2-1.4 kb

KBet01 (Kβ01)	ACCTAGGTACCACTACGCTACGCAGC AGCC	Betasatellite	94℃=30s, 55℃=1m, 72℃=3m, 35cycles, 72℃=15m Product size= 1.2-1.4 kb
KBet02 (Kβ02)	ACCTAGGTACCTACCCTCCCAGGGGT ACAC	Betasatellite	
M13UF	CCCAGTCACGACGTTGTAAAACG	lacZ(pUC)	94℃=30s, 52℃=1m, 72℃=1.5m, 35cycles, 72℃=15m Product size= depends upon insert size
M13UR	AGCGGATAACAATTTCACACAGG	lacZ(pUC)	
KGF1335	GGATATACAGAAGCCCAGATGTTCC	Full length DNA-A (AC1-AC4 region)	94℃=30s, 56.2℃=50s, 72℃=3m, 35cycles, 72℃=15m Product size= 2.74 kb
KGR1336	TGTACATCTTGGGCTTTCGG	Full length DNA-A (AV2-AV5-AV1 region)	
SA1825EF	GGATGGAATTCGCTATACACCTTTGG GCTCAG	siRNA (upper half) forward	94℃=30s, 58.8℃=1m, 72℃=3m, 35cycles, 72℃=15m Product size= 1.4 kb
SA472SR	TCTGAGCTCGCACGTCGGGACTTCTA TACATTC	siRNA (upper half) reverse	
SA472BR	ATCGGGATCCGCACGTCGGGACTTCT ATACATTC	siRNA (upper half insertion of BamH1 site) reverse	
RTMF	CGATCGAATTCGTAAGTCTGATTTTTG ACTCTTC	Intron (hairpin)	94℃=30s, 54.9℃=50s, 72℃=1m, 35cycles, 72℃=15m Product size= 120bp
RTMR	ATGCGCTAATGAGAAGTCACCACCAG GAGG	Intron (hairpin)	
SRTB2	TAATGAGAAGTCACCACCAGGAGG	Bridged primer at intron for asymmetric PCR	94℃=30s, 56.2℃=1.5m, 72℃=4m, 35cycles, 72℃=15m Used in combination with SA472R primer Product size= 2.8 kb/1.5kb
FB1F	GGCACTCACAAACGTCTATTTC	qRT-PCR F-box gene	94℃=30s, 56.2℃=30s, 72℃=50s, 35cycles, 72℃=3m
FB2R	ACCTGGGAGGCATCCTGCTTAT	qRT-PCR F-box gene	
CPR1	CAGCAACGGTGAAGAACATGGC	qRT-PCR CP gene	94℃=30s, 52.2℃=30s, 72℃=30s, 35cycles, 72℃=3m
CPR2	GCTCCCTAGATGCATATGTTCCTCC	qRT-PCR CP gene	
C1R1	CTCTACCATGTCCTCCTGTC	qRT-PCR c1 gene	94℃=30s, 51.5℃=30s, 72℃=30s, 35cycles, 72℃=3m
C1R2	CAACAATTTGGAGGAAGGGG	qRT-PCR c1 gene	

sterile deionized water. PCR consists of 3 different steps i.e. denaturation at 94℃, annealing at 50-65℃ and extension at 65-72℃ for appropriate time with 25-35 cycles. The PCR was employed in diagnosis of begomovirus in plant leaves using degenerate primers and amplification of full length viral isolates with condition specific for each reaction (Table 3.2). The PCR amplified fragments were resolved on 0.8-1.2% agarose gel through electrophoresis.

3.5.6 DNA extraction and purification from agarose gel

The PCR amplified fragments resolved on agarose gel were sliced off under ultra-violet lamp illumination and transferred to Buffer SET provided with Nucleopore® SureExtract Spin PCR clean up/Gel extraction kit (Genetix Biotech Asia Pvt. Ltd.) and incubated at 50℃ for 5-10 min. The gel when melted completely was transferred to spin columns and centrifuged for 1 min at 11,000xg at RT. The flow through was discarded and column was washed twice with 700 µl Buffer SET3 and spin-dried. The DNA in columns was eluted with either TE buffer or sterile deionized water to required concentration.

3.5.7 Preparation of plasmid DNA using alkaline lysis method

Overnight bacterial culture (5ml) were harvested by centrifugation at 8,000 rpm for 1 min. The pellet was washed with 0.1X phosphate buffer and resuspended in 0.1 ml of chilled Solution I [50mM Tris-HCL, pH8.0; 10mM Na-EDTA, pH8.0; 50mM Glucose (filter sterilized)]. Freshly prepared Solution II (2 ml) was added and gently mixed by inverting 5 times. Upon lysis completion, 0.15ml of chilled Solution III was added to neutralize the solution and incubated on ice for 15-20 min. The supernatant was collected by centrifugation at 13,000 rpm at 4℃ for 20 min. The supernatant was processed as described in section 3.5.2.

3.5.8 Preparation of *E.coli* competent cells and transformation

E.coli DH5α cells were grown in 100 ml LB medium (sub-cultured on LA plates and 5 ml LB previously) at 37℃ until OD_{600} reached 0.4-0.6. The cells were harvested at 3000 rpm at 4℃ in SS34 tubes. The loosely bound pellet was washed with 20 ml of 100 mM $CaCl_2$ for 5 min on ice and centrifuged to collect clean pellet. Again, the pellet was suspended in 20 ml of 100 mM $CaCl_2$ and incubated for 10-15 min on ice. The pellet was collected by centrifugation and dissolved in 4 ml glycerol resuspension buffer [100mM$CaCl_2$+ 50% sterile glycerol] with incubation on ice for 1-3 h.

Subsequently, the cells were transferred to chilled 1.5 ml microfuge tubes in 0.1 ml aliquots and stored at -80℃ until further use.

The gel purified PCR/restriction digested fragments (both containing 3'A overhangs) were ligated with appropriate plasmid (preparation in section 3.5.7) i.e. pTZ57R/T using T4 DNA ligase enzyme (1unit/reaction) according to manufacturer's instructions. The ligated product was used directly for transformation. DH5α cells were transferred on ice from -80℃ and allowed to thaw for 20-30 min. The ligated product was added and allowed to incubate for 10-15 min. The heat shock treatment was carried out at 42℃ for 90 s and immediately chilled on ice for 5 min. Approximately, 1 ml of LB was added to the cells and allowed to grow at 37℃ for 1 h with slow shaking. These cells were plated on LA plates with appropriate antibiotics and allowed to grow overnight or for 16-18 h at 37℃. The colony forming units thus obtained were screened with colony PCR to obtain positive clones.

3.5.9 Preparation of *Agrobacterium* competent cells and transformation

Agrobacterium tumefaciens strain LBA4404 harboring helper plasmid only was cultured in YEM medium [0.4g/L Yeast extract, 0.1g/L NaCl, 0.2 g/L $MgSO_4$, 10g/L Mannitol, 0.5g/L K_2HPO_4, pH 7.0] supplemented with streptomycin (25 mg/L) and rifampicin (25 mg/L) till OD_{600} reached ~0.5 absorbance. The culture was pre-chilled at 4℃ for 30-45 min before harvesting by centrifugation at 3000 rpm for 20 min at 4℃. The pellet was washed with 0.1X phosphate buffer and suspended in chilled 10 mM $CaCl_2$ solution. After incubation of 2-3 h on ice, the cells were harvested at 25,000 rpm and resuspended in 10 mM $CaCl_2$ solution containing equal volume of >80% glycerol. The cells were transferred to 1.5 ml microfuge tubes as 0.1-0.2 ml aliquots and stored at -80℃.

Approximately, 1 µg of a plant transformation vector-containing insert was added to the LBA4404 competent cells and incubated on ice for 20-30 min. The cells were pulverized in liquid nitrogen and immediately treated with a heat shock in a water bath set at 37℃ for 5 min. The cells

were incubated on ice for 5-10 min and then incubated in incubator shaker at 28℃ for 3-5 h at 200 rpm after addition of YEM medium (800-900 µl). After the appropriate culture growth, the cells were plated on YEM agar plates containing appropriate antibiotic. The plates were incubated at 28℃ for 48 h until satisfactory colony formation was obtained. These colonies were further processed for screening of positive clones.

3.6 Preparation of infectious clones containing dimeric forms of viral genome

Upper half of the begomovirus was amplified and sequenced as mentioned in the section 3.5.5. The full-length genome of begomovirus was amplified by designing abutting primers (with appropriate restriction sites) against the region within upper half of the sense and anti-sense strands. The amplified fragments were restriction digested and partial ligation was carried out to obtain dimeric forms of the begomoviral DNA. This dimer was ligated into a restriction digested pCAMBIA1301 plasmid at appropriate restriction site in multiple cloning sequence region. This plasmid was propagated into suitable *Agrobacterium* strain. The clones thus obtained were screened with colony PCR using suitable primers. Positive clones were further sequenced and introduced into the tobacco plants.

3.7 Agroinfiltration procedure for mobilization of infectious clones in tobacco

3.7.1 Preparation of media and stock solutions

Murashige and Skoog (MS) medium (Murashige and Skoog 1962) was used for all the tissue culture experiments and was prepared in distilled water, sterilized by autoclaving. The MS medium constituents are listed in Table 3.3. The phytohormones used in the study were prepared as 1 mg/ml stocks.

Table 3.3 Composition of Murushige and Skoog medium (MS) used for *in-vitro* culture

Constituents	Concentration (mg/L)
Macronutrient components	
$MgSO_4.7H_2O$	370
KNO_3	1900

NH_4NO_3	1650
$CaCl_2.2H_2O$	440
KH_2PO_4	170
Micronutrient components	
FeSo4.7H2O	27.80
Na2.EDTA	37.30
MnSO4.4H2O	22.30
H3BO3	6.2
ZnSO4.4H2O	8.6
Na2MoO4.2H2O	0.25
KI	0.83
CuSO4.5H2O	0.025
CoCl2.6H2O	0.025
Supplementary components	
Sucrose	30,000
Inositol	100
Glycine	2.0
Nicotinic acid	0.5
Pyridoxine-HCl	0.5
Thiamine-HCl	0.1

Double distilled water used to make up volume to 1 litre.
pH of the medium was adjusted to 5.8 before autoclaving

3.7.2 Agro-inoculation/infiltration procedure in *Nicotiana* plants

10 ml YEM medium supplemented with 1 µl acetosyringone (200 mM), 100 µl MES buffer (2-(N-morpholino)-ethane sulfonic acid, 1 M) and the appropriate antibiotics were prepared. 20 µl glycerol stock of the *Agrobacterium* strain harboring viral DNA components was pipetted into the above medium and allowed to grow for 2 days at 28°C and 200 rpm to an OD_{600} of approximately 1.0. The cells were harvested by centrifugation at 3,000 rpm for 10 min. The pellet was dissolved in freshly prepared MMA medium (20 g sucrose, 5 g MS salts without vitamins, 1.95 g MES, 2 ml (1 M) NaOH, 1 ml acetosyringone (200 mM); finally adjust pH to 5.6 and make volume with ddH_2O to 1 L] to obtain an OD_{600} of 0.3. The cells were gently vortexed and resuspended. The cultures were kept at RT for 2-5 h before infiltration.

The *Agrobacterium* transformed cells in equimolar i.e. 1:1 molar ratio concentration were inoculated on *N. benthamiana* plants (3-4 week old). For inoculating *N. benthamiana* plants, junction of 4^{th} and 5^{th} leaf were pricked and one drop of inoculum was added. For agroinfiltration, 1 ml syringe was used to infiltrate the viral cultures on the dorsal side with

constant negative pressure on ventral side with fingers to create a vacuum for easy infiltration of culture. All inoculated plants along with the plant inoculated with empty construct (mock) were allowed to grow at 28°C in glass house chamber maintaining 16 h day and 8 h night period.

3.8 Whitefly mediated viral acquisition and transmission

3.8.1 Whitefly colony maintenance

The non-viruliferous whiteflies i.e. *Bemisia tabaci* Gennadius were maintained on cotton plants grown in glasshouses at 28°C in cages maintaining 14 h photoperiod and 30-50% relative humidity. Cages were built with wooden planks and insect-proof net (50 x 50 mesh size). Several other types of cages were built from pet bottles for transmission experiments.

3.8.2 Acquisition stage

The newly emergent (1-3 days post emergence) non-viruliferous adult whiteflies were collected from cotton plants and approximately 100-200 non-viruliferous whiteflies were transferred per cage containing 6-8 symptomatic *Nicotiana benthamiana* plants (6-7 weeks old). They were left for 36-48 h for acquisition of replicating viral genome. The viruliferous whiteflies were then collected for the transmission procedure.

3.8.3 Infection/Transmission stage

A fresh insect-proof cage was prepared and 3 papaya plants were kept per cage. Papaya plants (more than 10 weeks old) were grown under controlled glass house conditions. Approximately, 100-200 viruliferous whiteflies were released into these cages and allowed to transmit virus through feeding for 24 h. The whitefly colonies were terminated by using Pegasus® (Syngenta, India) insecticide used as per manufacturer's instructions. The papaya plants were maintained in insect-proof cages until completion of experiments.

3.8.4 Study of symptoms

All virus transmitted papaya plants were studied daily for symptom development and compared with mock to record the symptom appearance. Score was given for symptoms as per the criteria mentioned in Table 3.4 (Respective score was given if one or more of the symptoms from that category appeared).

Table 3.4 Score criterion set for evaluation of symptom severity in infected plants

Label	Score	Symptom description
Symptom-0	0	No symptoms
Symptom-1	1	Mild chlorosis, slight deformation in new leaves, systemic leaf has little crumpling
Symptom-2	2	Curling of leaf, chlorosis/mosaic
Symptom-3	3	Curling of leaf, chlorosis/mosaic, vein clearing, reduced plant growth
Symptom-4	4	Curling/rolling of newly emerging leaves and shoot apex growth reduced,
Symptom-5	5	Severe mosaic, curling/crumpling of whole shoot apex, stunted growth, no new leaves

3.9 Construction of siRNA binary vector for viral RNA silencing

3.9.1 Preparation of a binary vector for siRNA insertion

pBI121 (Addgene), plant binary vector was used for the construction of siRNA containing expression vector. pBI121 plasmid is a plant expression vector, 14,758 bp in size, contains CaMV35s promoter (346 bp), β-glucuronidase (GUS) reporter gene (1812 bp), kanamycin selection marker within T-DNA borders. It contains several restriction sites, which could be cleaved to insert the gene of interest. For this reason, BamH1 (5821 bp position) and Sac1 (7715 bp position) restriction sites were cleaved and dephosphorylated as described in previous sections. The cleaved plasmid pBI121-(BamH1-Sac1) was resolved on 0.8% agarose gel and recovered using gel extraction procedure mentioned in previous sections. The prepared vector pBI121-(BamH1-Sac1) was stored at -20°C until further step.

3.9.2 Preparation of siRNA cassette

siRNA expression cassette consists of a siRNA producing region derived from viral source in sense and antisense orientation. The two regions are separated by an intron from unrelated plant host,

which acts as a hairpin-forming loop for stability of siRNAs. *Arabidopsis thaliana* (Col-0) RTM1 gene was used to amplify a 120 bp intron region using primers RTMF and RTMR (Thakur et al. 2014). Here, TLCV-Sultanpur isolate DNA-A upper half region was used to derive siRNA producing region. The primers were designed from the nucleotide at position 1825 corresponding to region near AC1 tail (SA1825EF) until the end of AV2 tail region at position 472 (SA472SR) as mentioned in Table 3.2. Both amplicons were digested with EcoR1 restriction enzyme (primers containing respective sites) and ligation was performed. The resulting ligated product was ~1.5 kb in size and was used as a template for asymmetric PCR. The asymmetric PCR was performed by designing bridged primer SRTB2 from intron end and RTMR reverse primer (AV2 tail region with SacI restriction site). The resulting product was ~3 kb in size and carried ~1.4 kb upper half region in sense and anti-sense orientation. A PCR Primer SA472BR was designed to incorporate a BamH1 restriction site at anti-sense end of the siRNA cassette.

3.9.3 Insertion into binary vector pBI121-(BamH1-Sac1)

The final siRNA cassette was double digested with BamH1-SacI restriction enzymes and ligated with pBI-(BamH1-Sac1) plasmid. The resulting plasmid pBS-Sul was transformed in *Agrobacterium* LBA4404 and positive clones were screened and selected on kanamycin YEM-Agar plates. The clones were sequenced using specific primers (Table 3.2) and confirmed for the presence of full siRNA cassette, correct orientation and in-frame insertion.

3.10 Isolation of total RNA from infected and control plants

Total RNA (100-150 mg) was isolated from the control and infected leaf samples using PureLink™ RNA mini kit (Invitrogen™). The tissue was pulverized in liquid nitrogen and made in to a fine powder using autoclaved mortar and pestle. The finely ground powder was dissolved in 500 µl Lysis buffer (freshly added 2-mercaptoethanol) and processed as given in manufacturer's instruction

manual (MAN0000406-2012). The isolated RNA was quantified using Nanodrop spectrophotometer for concentration and purity using A260nm/280nm.

3.11 Viral load quantification using quantitative Real Time PCR technique

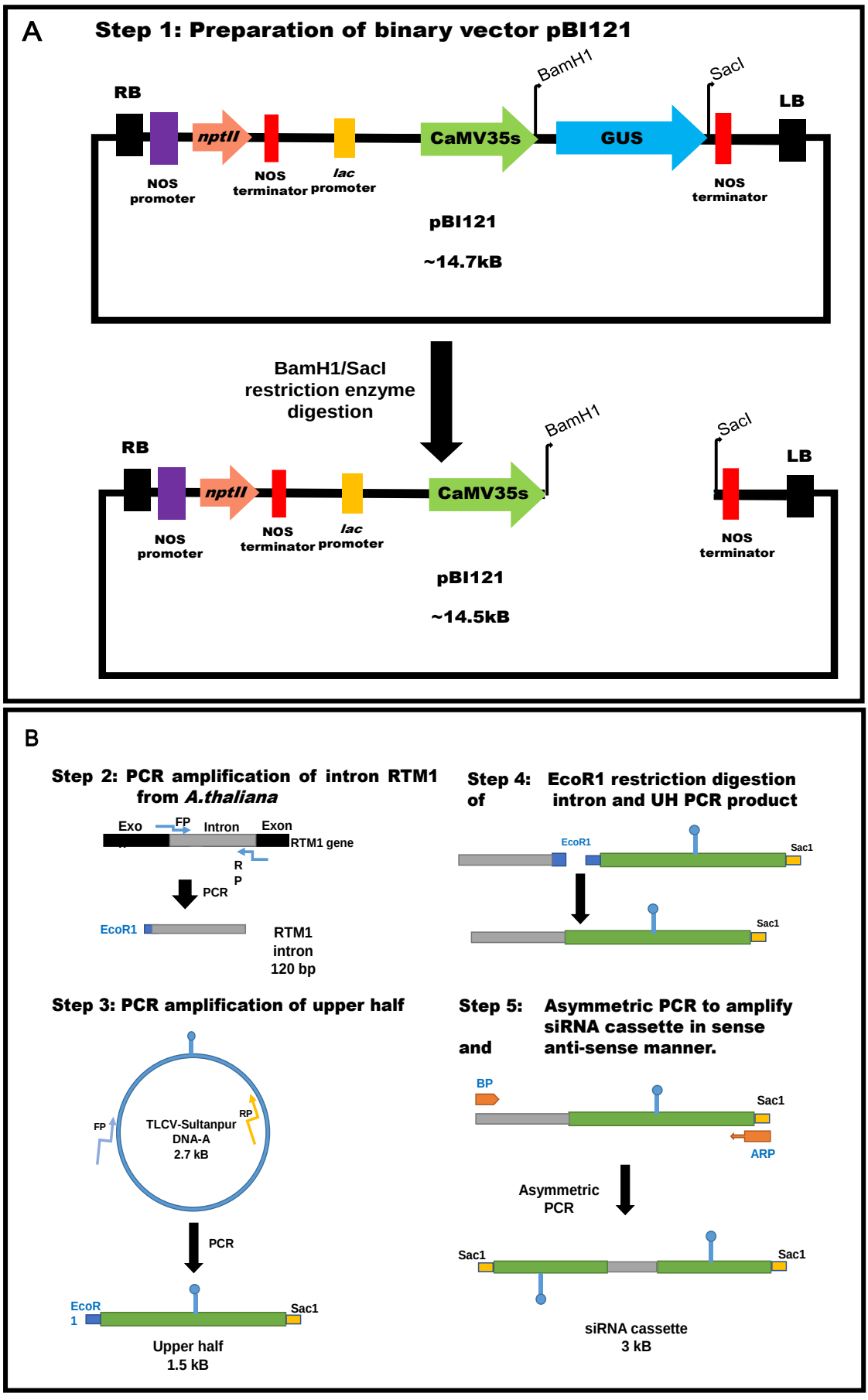
3.11.1 Preparation of cDNA

Total RNA (5 µg) treated with RNase free DNase was used for first strand cDNA synthesis. SuperScript™ IV First-Strand Synthesis System (Invitrogen) was used to synthesize cDNA first strand.

Cocktail-I [1 µl oligo-dT primer (50 µM), 10µl (~5 µg) total RNA, 13 µl sterile deionized water) was prepared in a 100 µl PCR tube and contents were annealed at 65°C for 5 min and immediately cooled down on ice. **Cocktail II** [4 µl 5x SSIV buffer, 1 µl of 10 mM dNTP mix, 1 µl of 0.1 M DTT and 1 µl Superscript IV™ RT (200 U/µl)] was prepared and mixed to the annealed RNA mixture. The contents of the tube were incubated at 50°C for 10-15 min and inactivated at 80°C for 10 min. The contents were stored at -20°C for short-term storage and at -80°C for long-term storage in a volatile buffer.

3.11.2 Quantitative Real Time PCR (qRT-PCR) analysis

The cDNA prepared was used for the quantification of viral load after confirming viral DNA through rolling circle amplification based detection. Viral titer was quantified through amplification of short 200 bp coat protein and C1 transcripts with the help of *Power* SYBR™ Green PCR Master Mix and (Applied Biosystems™). The qRT-PCR was performed according to the manufacturer's instruction (4367218 Rev.E) in Applied Biosystems 7500/7500 Fast real time PCR system. The results were analyzed with Applied Biosystems Software v2.0. *Nicotiana benthamiana F-box* (F-box) gene was used as a reference gene in the analysis. The viral quantification was done by calculating final fold change and comparing ΔΔCt values of reference and infected samples (Varkonyi-Gasic et al. 2007; Noris and Miozzi 2015). Appropriate biological and technical replicates were maintained throughout all the experiments.


A
Step 1: Preparation of binary vector pBI121
RB
nptII
NOS promoter
NOS terminator
lac promoter
CaMV35s
BamH1
GUS
SacI
NOS terminator
LB
pBI121
~14.7kB
BamH1/SacI restriction enzyme digestion
RB
nptII
NOS promoter
NOS terminator
lac promoter
CaMV35s
BamH1
SacI
NOS terminator
LB
pBI121
~14.5kB
B
Step 2: PCR amplification of intron RTM1 from A.thaliana
Exo
FP
Intron
Exon
RTM1 gene
R
P
PCR
EcoR1
RTM1 intron 120 bp
Step 4: EcoR1 restriction digestion of intron and UH PCR product
EcoR1
Sac1
Sac1
Step 3: PCR amplification of upper half
FP
TLCV-Sultanpur DNA-A 2.7 kB
RP
PCR
EcoR 1
Sac1
Upper half 1.5 kB
Step 5: Asymmetric PCR to amplify siRNA cassette in sense and anti-sense manner.
BP
Sac1
ARP
Asymmetric PCR
Sac1
Sac1
siRNA cassette 3 kB

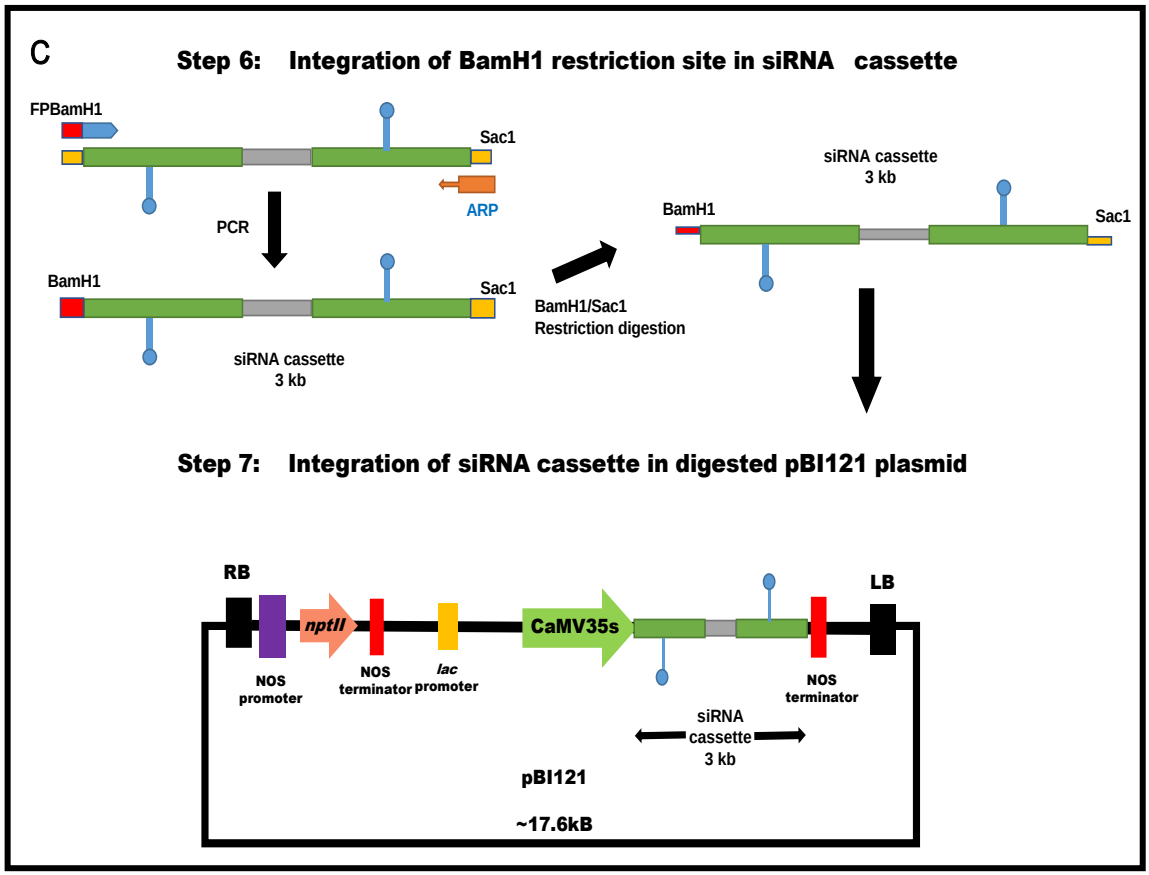


Figure 3.1 Construction of siRNA cassette and siRNA expression vector: (A) STEP 1: Preparation of binary vector pBI121-(BamH1-SacI) by restriction digestion of pBI121with BamH1 and SacI enzymes. (B) STEP 2: Intron fragment amplification resulted in synthesis of a 120bp hairpin template. STEP 3: Amplification of a 1.5kb fragment from DNA-A containing upper half of *Tomato leaf curl virus [India:Sultanpur: Papaya:2016]*. STEP 4: Ligation (EcoR1 site) of intron template and upper half fragment resulted into a 1.5kb fragment template for next step i.e. asymmetric PCR. STEP 5: Asymmetric PCR using bridged primer (BP) and reverse primer of upper half fragment (containing a Sac1 site) resulted into amplification of 2.95kb i.e. ~3kb, major product with SacI sites on both end, the product contained upper half in sense (viral hairpin upwards) and anti-sense (viral hairpin downwards). (C) STEP 6: Integration of BamH1 restriction site in siRNA cassette by incorporating BamH1 site towards antisense terminal instead of SacI site. Step 7: The resulting cassette from step 6 was incorporated into pBI-(BamH1-SacI) plasmid producing siRNA expression plasmid named pBS-Sul.

3.12 Experimental design for evaluation of siRNA efficacy in *N. benthamiana* plants

siRNA based resistance strategies are homology dependent post-transcriptional mode of gene silencing. Therefore, a plant binary vector based hairpin construct containing a genomic fragment of DNA-A was constructed. The aim for incorporation of this fragment was effective suppression of the transcription of genes lying within this fragment. The upper half PCR amplified fragment of DNA-A

component of *Tomato leaf curl virus [India:Sultanpur:Papaya:2016]* (MH105055) contains AV2 and partial CP gene fragment on viral sense strand and AC4, partial AC1 and AC5 gene fragments on viral complementary strand. Therefore, the siRNA decoy produced using siRNA construct pBS-Sul (Figure 3.1) could suppress any or individual mRNA production for the above-mentioned gene fragments.

For siRNA efficacy assessment: WT, Mock (pBI121 only), SA1301-5.4/SB1301-2.6 (1:1) only (S1301-5.4/2.6), pBS-Sul+S1301-5.42.6 (siRNA/Vir), S1301-5.4/2.6+pBS-Sul (Vir/siRNA) plasmids inside *Agrobacterium* LBA4404 cultures were agro-infiltrated. In 3rd type of agroinfiltration, siRNA construct pBS-Sul was infiltrated 7 days prior to introduction of equimolar cultures containing SA1301-5.4/B1301-2.6 (1:1). In 4th type of agroinfiltration, equimolar cultures containing SA1301-5.4/B1301-2.6 (1:1) plasmids was inoculated 7 days prior introducing siRNA construct containing culture. This condition was simulated to assess the silencing potential of siRNA construct on virus challenged field crops. The saplings were transferred at their 75 days stage i.e. when the immune system was developed.

N. benthamiana plants at 3-4 leaf stage were used to evaluate the efficacy of siRNA against genomic components of PLCD complex i.e. DNA-A of *Tomato leaf curl virus [India:Sultanpur:Papaya:2016]* and *Tomato leaf curl betasatellite - Naj 2 [India:New Delhi: Papaya:2009]*. Agroinfiltration of infectious clones containing genomic dimers i.e. 5.4 kb DNA-A fragment and 2.6 kb betasatellite fragments, initiated the virus challenge in tobacco plants. The infectious clones propagated in *Agrobacterium* cell line LBA4404 were agroinoculated in tobacco plants. Suitable control plant lines i.e. WT and Mock plants containing pBS-Sul plasmid only were used for the comparative evaluation of siRNA efficacy. The siRNA construct was introduced into the tobacco plants in two ways: first, at 0 day and in another at 7th day of experiment. S1301-5.4/2.6 (equimolar mixture) plasmid was positive control for the virus challenge. The leaves were collected from plants grown in same conditions in glass house, at suitable time intervals i.e. 7, 15, 36, 50th day

and processed separately for DNA and RNA based assays i.e., RCA and qRT-PCR. Real time analysis was performed for comparison of relative expression of CP (DNA-A) probe and c1 (betasatellite) probe in plants inoculated with infectious clones for evaluation of siRNA efficacy of pBS-Sul hairpin construct.

3.13 Statistical analysis

The statistical analysis for evaluation of significant changes observed in experiments e.g., infectious clones efficacy, viral and siRNA treatment efficacy and suitable time-period for observed changes, was performed with the help of embedded tools in Microsoft Excel package and IBM SPSS 20.0 software package, IBM Analytics Inc. (https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software). Kruskal Wallis H test was performed for the estimation of significance of symptom development (Ruxton and Beauchamp 2008) at different time-periods (alpha factor=0.0005). Post-hoc analysis of above data was performed using Mann-Whitney U test, a non-parametric test of null hypothesis i.e. the two variables are equal and there is no significant difference and the alternate hypothesis that the two variables have significant different response to a treatment (Kasuya 2001), followed by Friedman and Wilcoxon Signed Ranks Test (Laerd Statistics 2013).

Results & Discussion

4. Results and Discussion

4.1. *in-silico* investigation into Papaya leaf curl disease (PLCD) group

4.1.1. DNA-A begomoviral components of PLCD complex

The PLCD complex consists of 13 begomovirus species associated with β satellites (Figure 4.1). This begomovirus complex component is present in Central Asia i.e. China, India, Pakistan and parts of Bangladesh. *Papaya leaf curl China virus* (PaLCuCNV), *Papaya leaf curl Guandong virus* (PaLCuGNV) *Tomato leaf curl Hainan virus* (ToLCuHaV) were mainly reported in China infecting papaya, tomato and members of family *Euphorbiaceae*. The Chinese components are associated with *Papaya leaf curl China betasatellites*. There are 18 different virus isolates comprising the Chinese PLCD complex components. The Pakistan PLCD complex consists of *Papaya leaf curl Faisalabad virus*, an unclassified virus that has been excluded from begomovirus. The Indian PLCD complex consists of 28 isolates and has diverse nature in terms of species reported to cause leaf curl symptom in papaya. Nine begomovirus species i.e. *Chili leaf curl virus* (ChiLCV), *Papaya leaf curl virus* (PaLCuV), *Pedilanthus leaf curl virus* (PeLCuV), *Papaya leaf crumple virus* (PapayaLCV), *Tomato leaf curl virus* (ToLCV), *Tomato leaf curl New Delhi virus* (ToLCNDV*), Papaya yellow leaf curl virus* (PaYLCuV) etc. have been known to infect papaya. Overall, the PLCD complex consists of 47 different isolates identified to cause leaf curl disease in papaya and associated weeds and crops.

4.1.2. High molecular diversity of DNA-A in PLCD complex

The phylogenetic analysis of thirty-five isolates belonging to Indian PLCD complex (Annexure 1) was performed. The CLUSTALW output was used for Bayesian inference with MrBayes v.3.1.1 using General time reversible with gamma (GTR-G) substitution model (Larget and Simon 1999) with two runs for 2,50,000 generations, sample frequency at 10 and burn-in ratio at 42% i.e. first 1,05,000 generations were excluded from the inference. The same set of PLCD complex DNA-A sequences were analyzed for maximum likelihood analysis using PhyML-aLRT v2.4.5 and Hasegawa, Kishino and Yano - Gamma (HKY- G) as substitution model (Hasegawa et al. 1985).

The above phylogenetic analysis produced likelihood values that were very similar to each other i.e. 26743.85 and 26732.18 for MrBayes and PhyML respectively (Figure 4.2 and 4.3). The close lying likelihood scores indicate that the results and interpretation for MrBayes tree analysis will be same as PhyML-aLRT tree analysis.

The genetic complexity of Indian PLCD group was determined employing posterior distribution of model GTR-G parameters through applying Markov Chain Monte Carlo (MCMC) methods available in MrBayes analysis output (Larget and Simon 1999; Ronquist and Huelsenbeck 2003; Gamerman and Lopes 2006) (Figure 4.2). Potential scale reduction factor (PSRF) converged to value 1.0 for all parameters providing acceptable inferences in above analysis. The convergence in the range of 1.0-1.2 indicate that the clade credibility tree has branches lying in the region of acceptable posterior probability i.e 95% or more.

Therefore, branch length in the cladogram, represents the measured expected substitution per site, which signifies a phylogenetic relatedness among all clades and sub-clades. Further, MrBayes and PhyML tree results were mid-rooted tree (Figure 4.2 and 4.3), divided into two major clades and an out-group i.e. PLCD19 (*Tomato leaf curl New Delhi virus*). Other out-groups lying in the two major clades include PLCD13, PLCD14, PLCD34, PLCD35 and PLCD31 representing CLCuMuV, PaLCuV [India:New Delhi:tomato:2005], PaYLCuV [DP2], ToLCuV [C1], PapayaLCV [A-87] respectively.

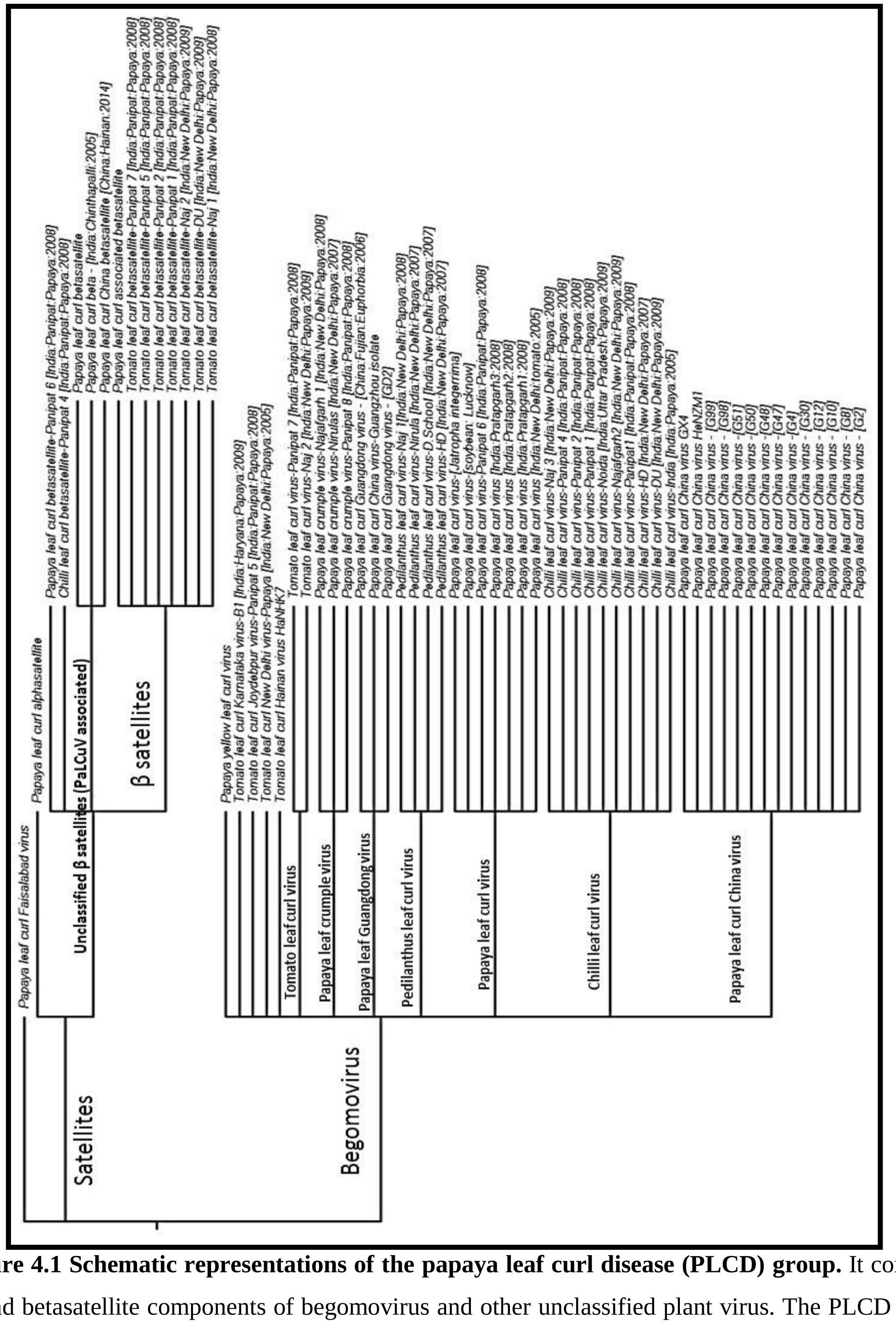


Figure 4.1 Schematic representations of the papaya leaf curl disease (PLCD) group. It consists of DNA-A and betasatellite components of begomovirus and other unclassified plant virus. The PLCD group contains diversity in strains causing leaf crumpling and curl like symptoms in papaya and associated crops and weeds.

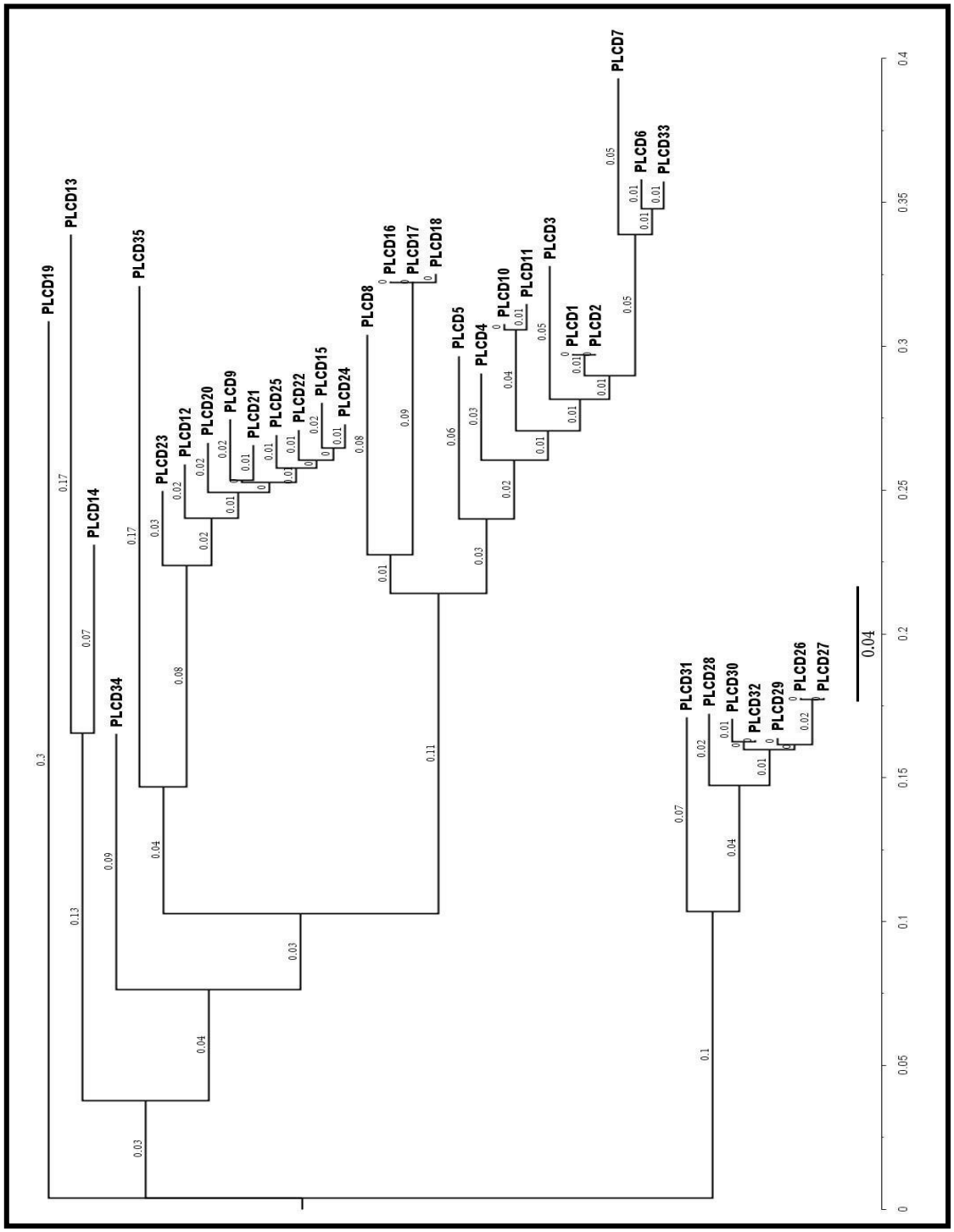


Figure 4.2 Phylogenetic tree based on the complete DNA-A sequences of the begomoviruses detected in *C. papaya* and additional sequences from associated crops and weeds infecting begomoviruses causing leaf curl disease. The tree was constructed by Bayesian inference using the GTR-G nucleotide substitution model available in MrBayes v3.1.1 method in TOPALi v2.5. Numbers at the branches indicate estimated clade credibility score inferred from statistical analysis of posterior probabilities assessed for the substitution model.

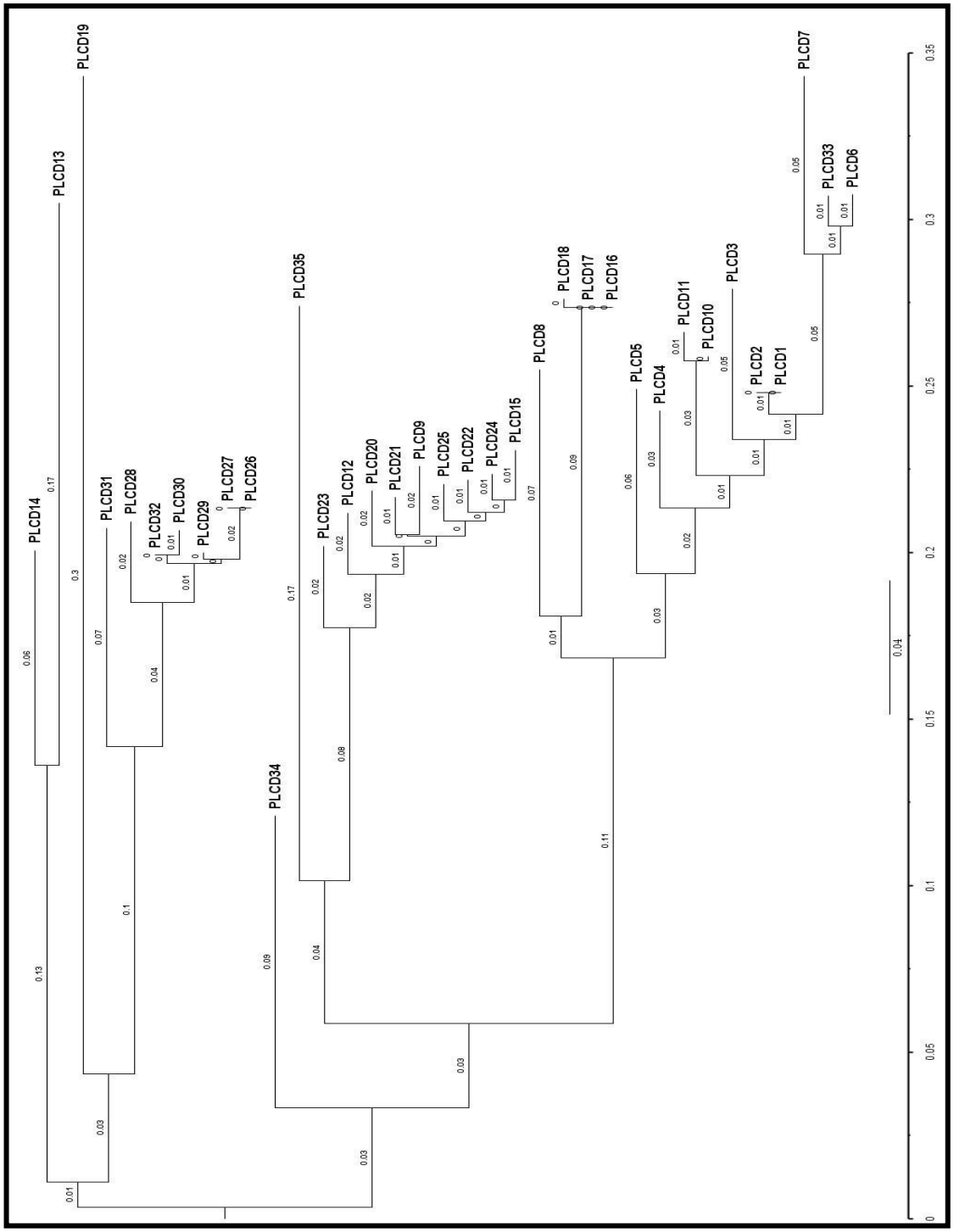


Figure 4.3 A maximum likelihood tree based upon complete DNA-A sequences of the begomoviruses detected in *C. papaya* and additional sequences from associated crops and weeds infecting begomoviruses causing leaf curl disease. The tree was constructed by using PhyML-aLRT v2.4.5 tool in TOPALi v2.5 and the HKY-G nucleotide substitution model. Numbers at the branches indicate estimated substitution rate per nucleotide position.

The first major clade consists of Papaya leaf crumple virus isolates within branch lengths 0.1 to 0.17 signifying low expected substitution rates per site (Figure 4.3). The narrow branch length, is the reason for narrow host range i.e. virus infecting only *A. paniculata*, as its reservoir host, which is

an annual weed widely found in south India whereas tomato and papaya as a primary host in north India. This indicates that this new virus was introduced in northern regions via anthropogenic activities. During its course for survival, the *Papaya leaf crumple virus* isolates have adapted tomato and papaya as its primary host and evolved in parallel with *Papaya leaf curl virus* group in the northern regions. The second major clade was further sub-divided into two minor sub-clades i.e. ChiLCV and PaLCuV. The ChiLCV clade consists of ChiLCV and CLCuBuV [LK_2N] whereas PaLCuV clade clustered the papaya and associated crop infecting PaLCuV isolates according to their hosts. The ChiLCV clade has a narrow branch length range i.e. 0.24 - 0.29 signifying low substitution rate per site. Therefore, whole ChiLCV clade shows close genetic relatedness and host singularity as evident from *C. papaya* being the only host crop infected by members of this clade. The pattern suggests that these clade viruses have evolved to infect papaya crops lying nearby chili fields in north Indian region and they are under very low mutational pressure to evolve and expand their host range. The PaLCuV clade on the other hand, has a very broad range of branch length i.e. 0.23 – 0.39 signifying high substitution rate per site; showing expansive genetic relation and diverse host range. Analysis also suggests that they might have evolved independently on diverse hosts in the geographically different regions. The out-group species were un-related to the clade forming viral species. These outgroup viral species i.e. *Tomato leaf curl isolate C1* and *Papaya yellow leaf curl virus isolate DP2* have acquired the ability to infect papaya and associated host crops and weeds.

4.1.3. Species diversity identification on the basis of DNA-A by pairwise sequence comparison

The International Council for Taxonomy of Viruses (ICTV) has decided to solve the species nomenclature issues by using pairwise sequence identity criterion as a standard methodology for species demarcation. Earlier cutoff i.e., >89% sequence similarity was elevated to >91% and use of SDT tool was made mandatory for naming of new viral species. Therefore, this tool was employed to study the taxonomic status of species reported to be associated with Indian PLCD complex.

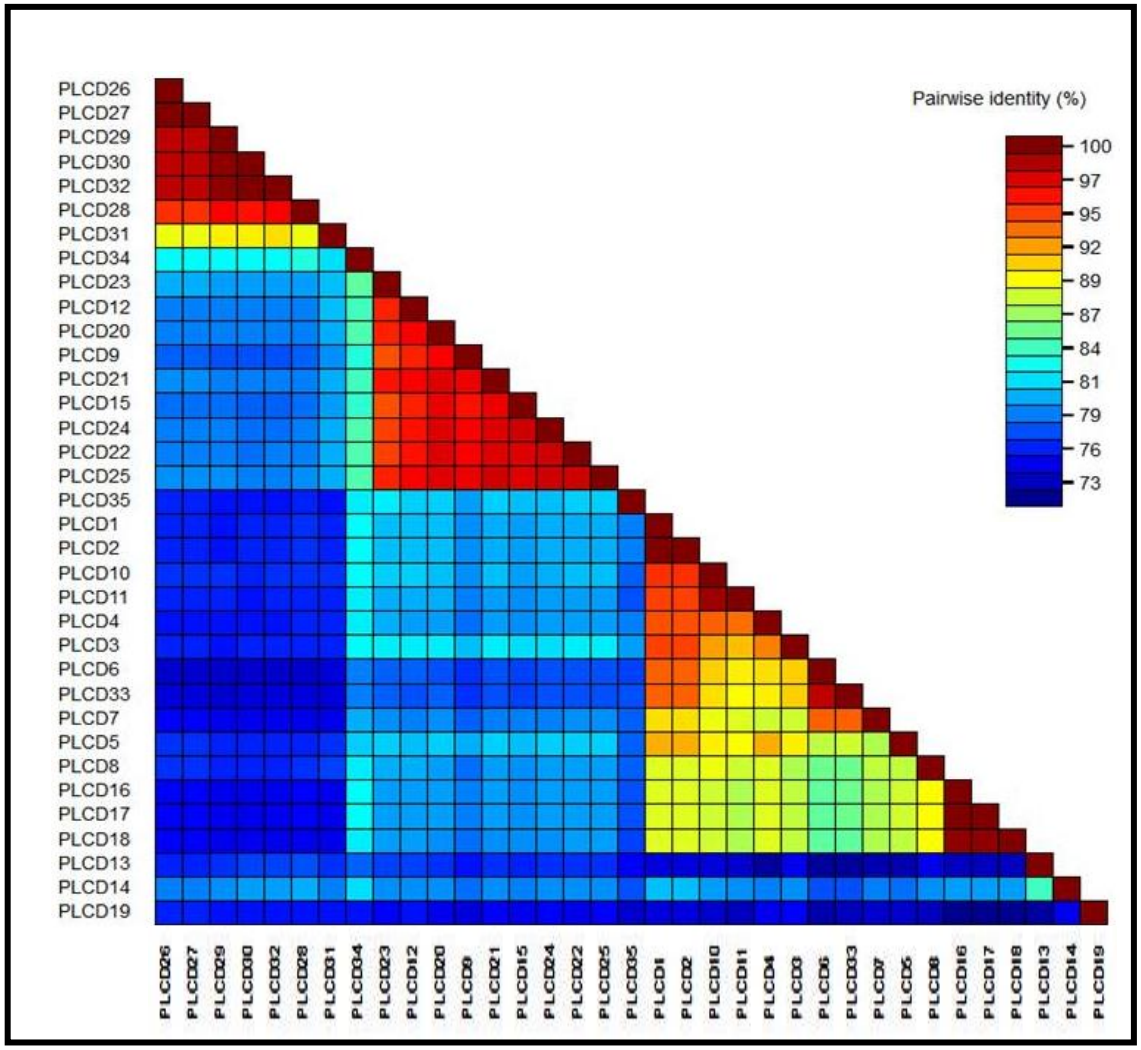


Figure 4.4 Pairwise identity estimation of sequence similarity using Sequence Demarcation Tool v1.2. The matrix output indicates the diversity of begomoviruses infecting papaya and associated crops and weeds. The color key indicates highest and least similarity with brown and blue color respectively.

The pairwise matrix (Figure 4.4) obtained by using SDT v1.4 shows a similarity matrix characteristic of a complex association as the minimum and maximum similarity obtained was >75% and 100%. The species with highest similarity group i.e. >90% to 100% were grouped together using inbuilt function of the software. The PLCD complex was distributed into 11 different groups (Table 4.1). The groups were labeled as *Papaya leaf curl virus* group, *Chili leaf curl virus* group, *Papaya leaf crumple virus* group and *Papaya leaf curl virus (Pratapgarh)*, these groups consist of 10, 9, 6 and 3 isolates respectively. Among all 35 PLCD complex isolates, 28 belonged to above mentioned groups and rest of the DNA-A sequences were phylogenetically isolated as individual species

associated with Indian PLCD complex. Further, recombination analysis can prove their categorization into above mentioned groups.

Table.4.1 SDT analysis of PLCD DNA-A components. Result of pairwise sequence demarcation using SDT v1.2 program for inferring pairwise sequence similarity among begomovirus components in PLCD complex (Figure 4.10) The groups have been distributed according to the >90 to 100% sequence similarity with each other.

Group No.	No. of species	PLCD complex	Virus Taxon included in group
1	6	PLCD26; PLCD27; PLCD28; PLCD29; PLCD30; PLCD32	**Papaya leaf crumple virus**
2	1	PLCD31	**Papaya leaf crumple virus (A-87)**
3	1	PLCD34	**Papaya yellow leaf curl virus (DP2)**
4	9	PLCD9; PLCD12; PLCD21 PLCD22; PLCD23;PLCD24; PLCD25	**Papaya leaf curl virus (New Delhi/2016)** **Chili leaf curl virus (Amritsar/2009)**
5	1	PLCD35	**Tomato leaf curl virus (C1)**
6	10	PLCD1; PLCD2; PLCD3; PLCD4; PLCD5; PLCD6; PLCD7; PLCD10; PLCD11; PLCD33	**Papaya leaf curl virus**
7	1	PLCD8	**Papaya leaf curl virus (WB2&5)**
8	3	PLCD16; PLCD17; PLCD18	**Papaya leaf curl virus (Pratapgarh)**
9	1	PLCD13	**Cotton leaf curl Multan virus**
10	1	PLCD14	**Papaya leaf curl virus (New Delhi/2005)**
11	1	PLCD19	**Tomato leaf curl New Delhi virus**

4.2. Recombination pattern study of DNA-A in Indian PLCD complex

Recombination is an important molecular event that leads to translocation, transversion, addition, deletion or frame-shift mutations in a genome (Lima et al. 2017). Due to the rolling circle replication mechanism in begomoviruses, these organisms have evolved to survive inside their hosts through acquisition of useful regions from other viruses or the host themselves, so that they can successfully evade host mediated RNA interference based immune response (van der Walt *et al.*, 2009). In the above analysis, it was observed that the Indian PLCD complex is well defined with major contribution made by PaLCuV, ChiLCV, PapayaLCV and discreet reports of other species infecting

papaya and associated crops and weeds. To study if recombination is responsible for occurrence of such complex virus-host interaction; the DNA-A sequences of Indian PLCD complex were subjected to recombination analysis using RDP v4.95 using available recombination analysis tools (Figure 4.5) and default parameters as recommended by RDP4 manual. The analysis ended up predicting 66 recombination events, which further reduced to 12 events after careful manual screening as per the RDP4 manual.

According to the manually screened recombination events results (Table 4.2), event no. 10 i.e. PLCD9 and PLCD31 (inferred closest to the unknown minor parent) seems to be the major contributor of fragments giving a chance for origin of a new recombinant species e.g., PLCD8 and thirteen others i.e. PLCD1, PLCD2, PLCD3, PLCD4, PLCD5, PLCD6, PLCD7, PLCD10, PLCD16, PLCD17, PLCD18, PLCD23 and PLCD33. Thus, a *Chili leaf curl virus* and *Papaya leaf crumple virus*, infecting papaya and *A. paniculata* contributed significantly to the Papaya leaf curl virus sub-clade, which partially explains the clades affinity to infect weeds and associated crops. The event no. 5, PLCD2 and PLCD28 (inferred closest to the unknown minor parent) gave rise to PLCD33, PLCD6 and PLCD7 (already contains fragments from event no.10), resulting into *Papaya leaf curl virus* isolates capable of infecting soybean, croton and tomato. Similarly, PLCD5 became a weed infecting recombinant after acquiring fragments through event no.1 i.e. *Papaya leaf curl virus* infecting *S. lycopersicum* and *Chili leaf curl virus* infecting papaya. Overall, the Indian PLCD complex appears to be a repertoire of recombinant betasatellite components as evident from the recombination analysis. As the level of complexity in recombination increases, the host range of recombinant virus also seems to expand accordingly.

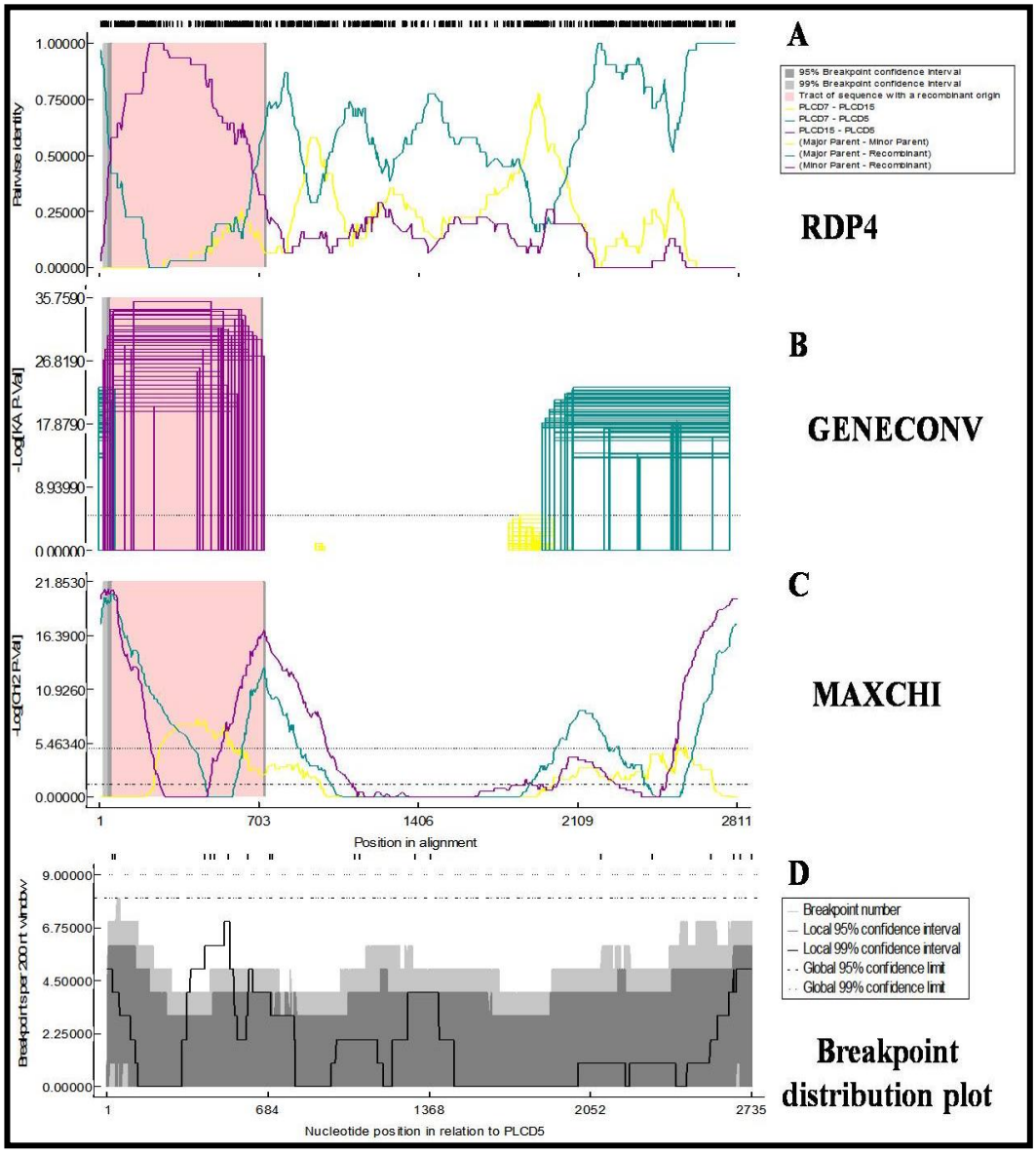


Figure 4.5 Recombination analysis showing results of event number 1 using Recombination Detection Program v4.95 (RDP v4.95). The pink region indicates the occurrence of the recombination event with intersection points in the case of RDP and overlapping peaks in case of MAXCHI on the plot as the recombination starting and ending breakpoints. The grey region at the ends represents 99% and 95% confidence interval of prediction of breakpoints as mentioned in the key. A. RDP output; B. GENECONV output; C. MAXCHI output and D. Recombination breakpoint distribution plot: The light and dark grey regions in the plot represent the 99% and 95% confidence limit respectively for a 200nt scanning window. The breakpoint distribution was assessed with PLCD5 DNA-A sequence as a reference sequence just for qualitative purpose.

Table. 4.2 Recombination analysis of Indian PLCD complex using RDPv4.95 program. The analysis was performed against DNA-A sequences of 35 begomoviruses using default parameters (Figure 4.6.1) and following algorithms.

Event No.	Recombinant	Major parent	Minor Parent	Frequency of recombin-ation event	Detection Method						
					R	G	B	M	C	S	T
1	PLCD5	PLCD7	PLCD15	1	+	+	+	+	+	+	+
2	PLCD14	PLCD13	PLCD8	1	+	+	+	+	+	+	+
3	PLCD35	PLCD8*	PLCD3	1	+	+	+	+	+	+	+
4	PLCD7	PLCD10*	PLCD6	1	+	+	+	+	+	-	+
5	PLCD33	PLCD2	PLCD28*	3	+	+	+	+	+	+	+
6	PLCD13	PLCD2*	PLCD28	2	+	+	+	+	+	+	+
7	PLCD19	PLCD29	PLCD8	1	+	+	+	+	+	+	+
8	PLCD26	PLCD2	PLCD7	2	+	+	-	+	+	-	+
9	PLCD3	PLCD4	PLCD28*	1	+	+	+	+	+	+	+
10	PLCD8	PLCD9	PLCD31*	14	+	+	+	+	+	+	+
11	PLCD23	PLCD1	PLCD28	1	+	+	+	+	+	+	+
12	PLCD34	PLCD28	PLCD12	1	+	+	+	+	+	+	-
Total = 12			Total frequency = 29								

R=RDP; G=GENCONV; B=BOOTSCAN; M=MAXCHI; C=CHIMAERA; S=SISCAN; T=3SEQ

* Unknown parent detected in recombination analysis but the specified parent was the closest inferred sequence in this analysis.

4.3. Betasatellite enhances the molecular complexity prevalent in PLCD complex

Betasatellite molecules are the truncated version of a larger begomoviral genomic component associated with leaf curl disease. Many monopartite begomoviruses are incapable to produce viral symptoms and spread in usual manner due to lack of movement factors. These factors are important for cell-cell movement and nuclear trafficking of begomoviral DNA-A protein/nucleic acid components. The betasatellite molecules have been shown to enhance symptoms of begomovirus infection in plants and help in its systemic spread across host plants (Zhou 2013). Therefore, they become a prime candidate for investigation into their potential for siRNA based strategy. A

molecular investigation was carried out to get insight into the molecular complexity prevalent in betasatellite molecules associated with PLCD complex. Then, recombination based investigation was carried out to determine the potential contribution of recombination in providing betasatellite molecules the evident genomic complexity.

4.3.1. Betasatellite associated with DNA-A of begomoviral components of PLCD complex

The PLCD complex is mainly composed of monopartite begomoviruses, which are associated with either betasatellite or alphasatellite or both. Since, this complex is found to be closely related to tomato, chili and weeds infecting monopartite begomoviruses, therefore, a comprehensive list of betasatellites associated with DNA-A of the above begomoviruses was compiled (Annexure 2). The betasatellite molecules of *Ageratum leaf curl virus, Chili leaf curl virus, Cotton leaf curl Burewala virus, Papaya leaf curl virus, Tomato leaf curl virus, Tomato yellow leaf curl virus and Tomato leaf curl* Among 51 betasatellite sequences, 39 were identified in India, 10 in Pakistan and 2 from Chinese region. Total papaya, tomato, capsicum, guar and other crop infecting betasatellite molecules were 10, 11, 9, 5 and 12 out of 51 in the list. The betasatellite molecules associated with other crops and weeds along-with papaya were included in this study to evaluate their diversity and recombination potential. This helps in minimizing interference with resistance strategy at later stages of this study.

4.3.2. Betasatellite molecules are highly diverse and have potential to enhance PLCD complexity

All 51 betasatellite DNA sequences were retrieved from NCBI and saved as FASTA format and annotated as TBS1-TBS51 according to their taxonomic placements. Total genome alignment was performed with MUSCLE algorithm in MEGA6 software suite. The aligned DNA sequences of betasatellite molecules were included in phylogenetic analysis and a Maximum Likelihood tree was inferred (Figure 4.6).

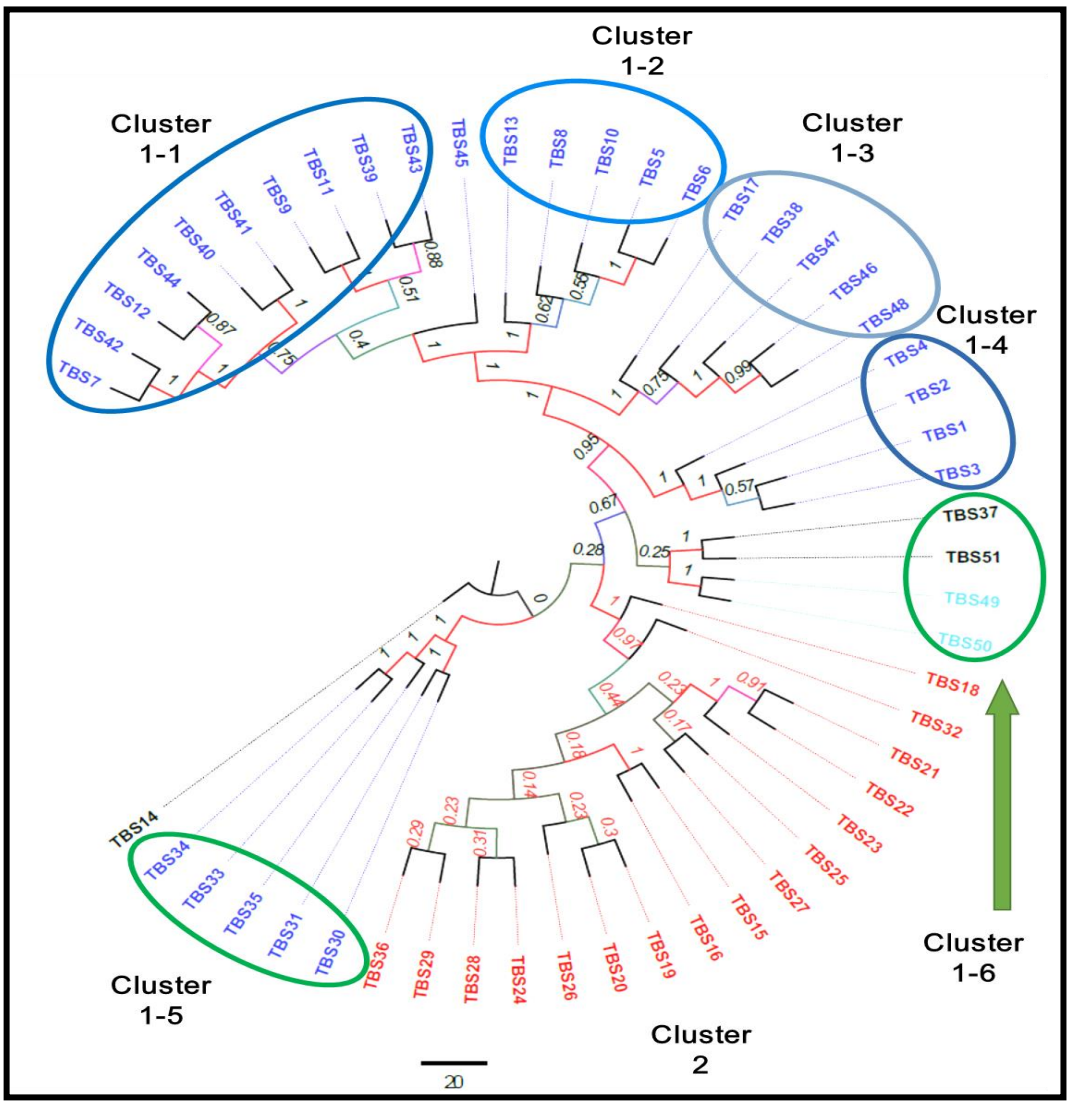


Figure 4.6

Phylogenetic analysis of betasatellite component of PLCD complex and associated leaf curl disease group. The tree was divided into two clusters i.e. Cluster-1 and Cluster 2 on the basis of likelihood predicted by bootstrap step estimation. The Cluster-1 was further subdivided into five sub-clusters on the basis of host, virus and region. Cluster-2 is sparsely related, therefore, all genomic components in this group were analyzed separately. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-13451.7843) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.5473)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 9.4207% sites). The analysis involved 51 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 878 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.*Joydebpur virus* were included in the list of 51 betasatellite molecules.

The tree was clearly demarcated into two types of clusters viz. genetically related CLUSTER-1 (blue colored group) and genetically distant CLUSTER-2 (Red to magenta colored groups). Chinese taxon is illustrated separately in light blue color and outgroups in black colored labels. CLUSTER-1 was

further grouped into Group1-1 to Group1-6 (Table 4.3). Each sub-clade in CLUSTER-2 was analyzed separately.

CLUSTER 1-1 consists of betasatellites infecting primarily papaya plant in New Delhi, Panipat (Haryana, India) and Punjab (Pakistan). These were found to be genetically very close i.e. phylogenetic likelihood predicted in 80-100% bootstrap steps. TBS7, 39, 40, 41, 42, 43, 44 and 45 were papaya infecting betasatellites belonging to Chili and Tomato leaf curl disease causing DNA-A component. CLUSTER 1-2 consists of Chili leaf curl associated betasatellites infecting capsicum and tuber plants in Punjab (Pakistan), Palampur (India), Rajasthan (India) and Punjab (India) regions. TBS 13 is an outlier in this group belonging to betasatellite associated with Cotton leaf curl Burewala disease from Punjab (India) region. CLUSTER 1-3 consists of tomato and guar plants in New Delhi, Haldwani in India and Bhakkar in Pakistan.

Table 4.3 Details of clusters obtained through phylogenetic analysis of betasatellite in PLCD complex

Cluster	Betasatellite	Feature
1-1	TBS7, TBS9, TBS11, TBS12, TBS39, TBS40, TBS41, TBS42, TBS43, TBS44	**Chili and Tomato leaf curl betasatellites infecting papaya plants.**
1-2	TBS5, TBS6, TBS8, TBS10, TBS13*	**Chili leaf curl betasatellite group**
1-3	TBS17*, TBS38, TBS46, TBS47, TBS48	**Tomato leaf curl betasatellite group**
1-4	TBS1, TBS2, TBS3, TBS4	**Ageratum leaf curl betasatellite group infecting weeds and crops other than papaya.**
1-5	TBS30, TBS31, TBS33, TBS34, TBS35	**Papaya leaf curl betasatellite group infecting crops other than papaya.**
1-6	TBS37, TBS49, TBS50, TBS51	**Tomato leaf curl betasatellite infecting tomato and chili in China, India and Bangladesh**

* Outgroup species not related to other members in host or species but occur together due to predicted nucleotide sequence similarity.

These betasatellites were found to be associated with primarily Guar and Tomato leaf curl disease. CLUSTER 1-4 is a group of weed infecting betasatellite associated with Ageratum leaf curl disease

reported mainly in Central parts of Uttar Pradesh and Rajasthan, India. CLUSTER 1-5 was found to be associated with Papaya leaf curl disease causing DNA-A component infecting guar and capsicum plants in western provinces of India and adjoining areas in Pakistan. CLUSTER 1-6, a group of *Tomato leaf curl betasatellite* molecules infecting tomato and capsicum in China, India and Bangladesh.

CLUSTER 2 appeared to be a complex clade consisting of papaya leaf curl associated betasatellite components reported from various parts of India. The betasatellite components in this group are genetically diverse and phylogenetically distant. The betasatellite components infect various crops such as papaya, tomato, cucurbits, weeds and other herbaceous plants growing in tropical regions of India and Pakistan. The papaya infecting group consist of TBS15 and 16 whereas tomato being infected by TBS24, 25, 32 and 36. A unique group of TBS21, 22 and 23 has been reported from Coimbatore and Pudukottai districts (Tamil Nadu, India), which infect various gram (Cereal) varieties in that region. TBS18, 19, 20 and 28 were infecting cucurbits and tuber varieties. The weed infecting group consists of TBS27 and 29 infecting *Parthenium* specifically. The wide genetic variation in this group provides the basis of genetic complexity in this study.

Although, not all betasatellite molecules selected for phylogenetic analysis were associated with PLCD complex, yet, their probability of becoming one in future cannot be obviated. Therefore, this analysis holds importance when a designing a siRNA based strategy against betasatellite molecules. The unrelated betasatellite species lying in same CLADE could be included into the possible target genomes groups selected for siRNAs based generic approach. In addition, evaluation of recombination in above results of betasatellite molecules could be done.

4.4. Recombination analysis of Betasatellite with DNA-A component of PLCD complex

A recombination analysis was conducted to investigate the recombination boundary distribution which provides an evidence of recombination pattern prevalent between the DNA-A and betasatellite

components of PLCD complex. DNA-A dataset of PLCD complex (Annexure 1) was aligned with betasatellite dataset (Annexure 2) to prepare input for RDP4 software based recombination analysis. The recombination analysis of betasatellite component did not predict any recombination boundary, which includes genomic fragments from DNA-A components of PLCD. The betasatellite dataset was analyzed for recombination boundaries and breakpoint distribution plot were obtained (Figure 4.7). Eleven recombination events were detected with statistical relevance and qualifying the criteria set initially (Table 4.4). In total, 50 recombination signals were predicted under 11 separate events. Event no. 4 was detected in 14 different viral sequences namely, TBS29,15, 16, 19, 20, 21, 22, 23, 24, 27, 28, 32, 36 and partially evident in TBS25. Similarly, event no. 10 was detected in 11 different recombinant betasatellite sequences namely, TBS 27, 15, 16, 19, 20, 24, 26, 28, 29, 36 and partially evident in TBS 25 and 32. Such a prominent recombination event prediction signifies that the major and minor parents *Papaya leaf curl virus betasatellite isolate In:Var:Pum:08:1* and *Tomato leaf curl betasatellite isolate Cluster bean, clone NGS-D1* respectively, were responsible for evolution of 15 putative recombinant betasatellite components reported to be associated with *Papaya leaf curl virus* reported from Indian tropical regions. Another pair of major and minor parents i.e. *Papaya leaf curl virus betasatellite clone BG-CBE and Cotton leaf curl Burewala betasatellite, clone L2-RCA-b1-F* respectively, have significant contribution in evolution of betasatellite components of above mentioned components of PLCD complex.

The betasatellite components associated with leaf curl disease of papaya and tomato have undergone significant evolution through recombination. Recombination has enabled these betasatellite components to evolve and infect different crops in varied tropical regions of Indian sub-continent.

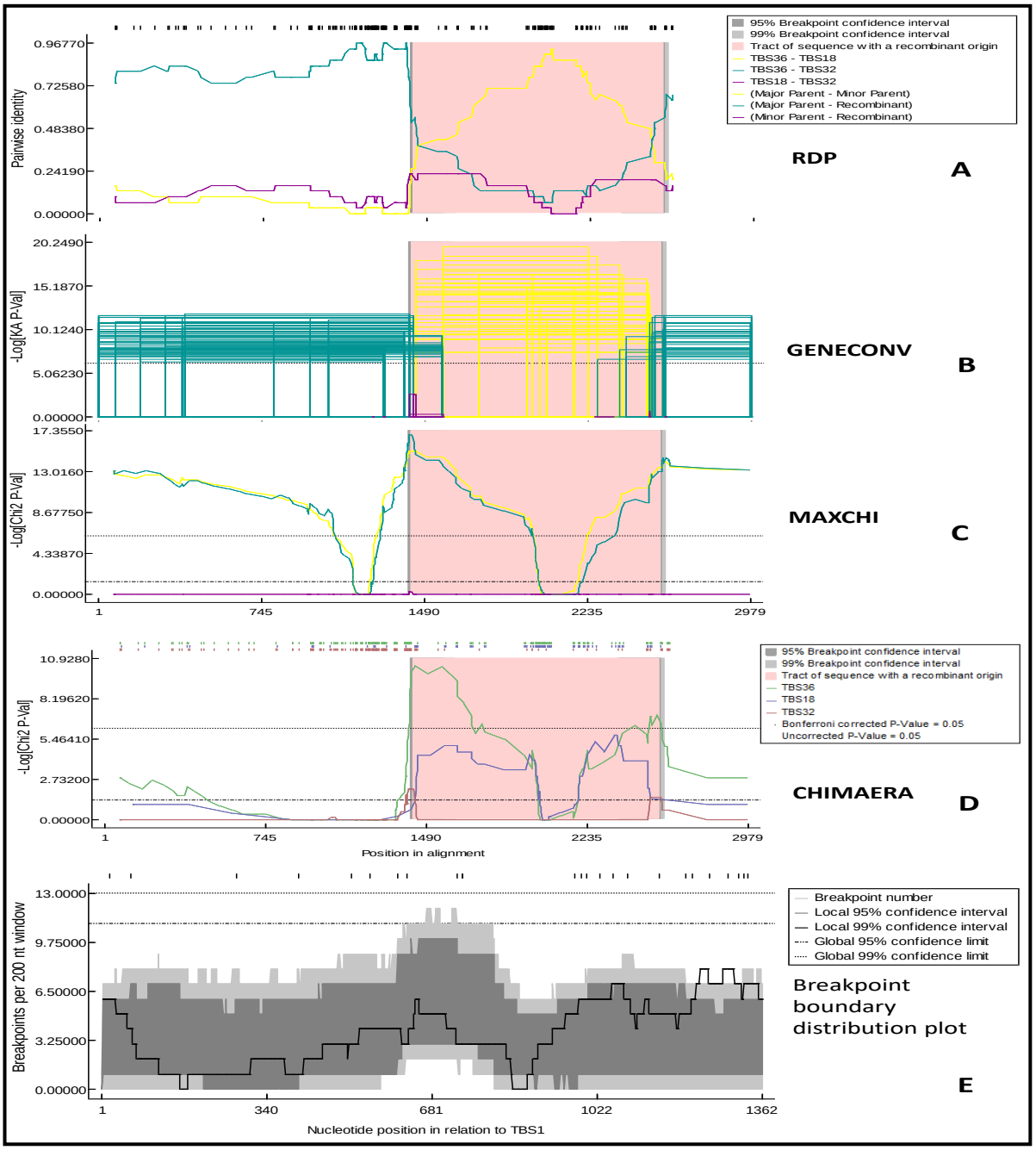


Figure 4.7 Recombination analysis showing results of event number 1 using Recombination Detection Program v4.95 (RDP v4.95). The pink region indicates the occurrence of the recombination event with intersection points in the case of RDP and overlapping peaks in case of MAXCHI on the plot as the recombination starting and ending breakpoints. The grey region at the ends represents 99% and 95% confidence interval of prediction of breakpoints as mentioned in the key. A. RDP output; B. GENECONV output; C. MAXCHI output; D. CHIMAERA output; E. Recombination breakpoint distribution plot: The light and dark grey regions in the plot represent the 99% and 95% confidence limit respectively for a 200nt scanning window. The breakpoint distribution was assessed with TBS1 sequence as a reference sequence just for qualitative purpose.

Apart from PLCD complex, the above mentioned parental pair has contributed in evolution of tomato leaf curl disease associated betasatellite components. Other predicted recombination events

i.e. event no. 11 and 57 also gave rise to recombinant viral sequences associated with tomato leaf curl disease complex (TLCD) (Moriones et al. 2017).In addition, the broad host spectrum of PLCD and TLCD could also be attributed to recombination. Owing to replication dependent recombination, the resistance strategies against leaf curl diseases in particular, became potentially prone to failure.

This study also explains the occurrence of interspecies infections i.e. association of Papaya betasatellite with TLCD and Tomato betasatellite with PLCD. These two disease complexes along-with ChLCD make up for the majority of leaf curl disease in Indian sub-continent, which are caused by a monopartite begomovirus. Therefore, betasatellite recombination hot-spots should be identified and future siRNA based strategies should be based upon these regions along-with the DNA-A component associated with respective monopartite begomoviruses.

Table. 4.4 Recombination analysis of Betasatellite genomes reported in Indian sub-continent using RDPv4.95 program. The analysis was performed against Betasatellite and DNA-A sequences of 51 begomoviruses using default parameters and following algorithms

Event No.	Recombinant	Major parent	Minor Parent	Frequency of recombination event	Detection Method						
					R	G	B	M	C	S	T
1	TBS32	TBS3	TBS18	2	+	+	+	+	+	+	+
2	TBS8	TBS10	TBS11	1	+	+	+	+	+	+	+
3	TBS4	TBS1	TBS34	1	+	+	+	+	+	+	+
4	TBS29	TBS18	TBS47	15	+	+	+	+	+	+	+
5	TBS27	TBS32	TBS15*	1	+	+	+	+	-	+	+
10	TBS27	TBS21	TBS14	13	+	+	-	+	-	+	+
11	TBS38	TBS39*	TBS6	5	+	+	-	+	+	+	+
12	TBS3	TBS2	TBS1	1	-	+	+	+	+	+	+
13	TBS35	TBS2	TBS1	4	+	+	+	+	+	+	-
15	TBS16	TBS26	TBS27*	2	+	+	+	+	+	-	-
57	TBS36	TBS27	TBS26	5	+	+	+	+	+	+	+
Total = 11			Total frequency = 50								

R=RDP; G=GENCONV; B=BOOTSCAN; M=MAXCHI; C=CHIMAERA; S=SISCAN; T=3SEQ

* Unknown parent detected in recombination analysis but the specified parent was the closest inferred sequence in this analysis.

4.5. Collection of papaya leaf curl infected leaf samples of *Carica papaya* L.

Leaf samples were collected from various sites located in Lucknow, Bareilly, Sultanpur and Malihabad regions of Uttar Pradesh, India. Total 19 sites were investigated and the leaf samples showing symptoms of leaf curl, vein swelling and other associated anomalies (Figure 4.8 A-F) were collected along-with non-symptomatic leaves of *Carica papaya* L. plants. The details of the

Table 4.5 Leaf curl infected *C. papaya* L. samples collected from various sites in U.P., India

S.No.	Sample Name	Place of collection/Date	Symptoms
1	BR1	Kitchen Garden, Shastri Nagar, Bareily- 12/07/15	Yellowish mosaic pattern+vein swelling + inward curling.
2	BR2	Kitchen Garden, Triveni nagar, Bareily- 12/07/15	Yellowish mosaic pattern + vein swelling + inward curling.
3	BR3	Vacant plot, Triveni nagar, Bareily- 12/07/15	Yellowing (old leaves) + vein swelling + Inward curling.
4	BR4	Kitchen Garden, Karamcharinagar, Bareily- 12/07/15	Vein swelling (Apical leaves) + Inward curling.
5	BR5	Kitchen Garden, Lajpatnagar 1- Bareily- 12/07/15	Yellowish mosaic pattern + vein swelling + inward curling + leaf mottling.
6	PC1	Highway road (Lucknow to Sultanpur)- 30/07/15	Vein swelling + average inward curl
7	PC2	Road side field, Aliganj, Sultanpur- 30/07/15	Vein swelling + severe inward curl, no fruits
8	PC3	Road side field, Aliganj, Sultanpur- 30/07/15	Vein swelling + severe inward curl, no fruits
9	BR6	Kitchen Garden, Lajpatnagar 1- Bareily- 12/07/15	Vein swelling + inward curling.
10	BR7	Kitchen Garden, Lajpatnagar 2- Bareily- 12/07/15	Vein swelling + inward curling.
11	BR8	Kitchen Garden, Lajpatnagar 2- Bareily- 12/07/15	Healthy
12	PC4	Department of Biotechnology, BBAU, Lucknow- 31/07/15	Vein swelling + average inward curling
13	PC5	Department of Biotechnology, BBAU, Lucknow- 31/07/15	Vein swelling + severe inward curling
14	PC6	Department of Biotechnology, BBAU, Lucknow- 31/07/15	Vein swelling + mild inward curling
15	KG1	Hostel garden, KGMU, Lucknow- 02/08/15	Vein Swelling + upward curling
16	KG2	Hostel garden, KGMU, Lucknow- 02/08/15	Vein Swelling + inward curling
17	CISH1	Nethouse, CISH, Rahmankheda, Malihabad- 03/11/14	Vein swelling + severe inward curling
18	CISH2	Nethouse, CISH, Rahmankheda, Malihabad- 03/11/14	Vein swelling + inward curling + mottling + dark green papery leaves
19	CISH3	Nethouse, CISH, Rahmankheda, Malihabad- 03/11/14	Yellowish mosaic + vein swelling + severe inward curling

collected samples is provided in the Table 4.5. The symptoms mainly comprise of yellowing of leaves producing mosaic pattern along-with vein swelling and inward (downward) curling as evident from Bareilly (BR1-5) and CISH3 (Figure 4.9 A); vein swelling, upward curling of leaves along-with mosaic pattern all over the leaves collected from KGMU, Lucknow (KG1) (Figure 4.9 B) and severe vein swelling and inward curling of leaves found in Sultanpur (PC1-3), Bareilly (BR6-8), Lucknow (PC4-6; KG2; CISH1-2) (Figure 4.9 C&D).

4.6. Isolation of DNA from healthy and symptomatic papaya leaves

The DNA extraction is an important step in the efficient PCR mediated detection of any Begomovirus, therefore, viral DNA was checked both qualitatively and quantitatively along with plant genomic DNA (G-DNA). Various methods were used to extract plant G-DNA. A viral DNA enrichment procedure was performed to detect the Begomovirus genome through PCR. A good quality viral DNA was isolated using enrichment procedure and yield was between 50ng/µl – 150ng/µl; samples were resolved on 0.8% agarose gel electrophoresis (Figure 4.10). Lanes 1-19 represent total enriched DNA samples isolated from papaya leaves.

4.7. Polymerase chain reaction mediated detection of viral components in healthy and symptomatic leaves of papaya plant

The PCR mediated screening for DNA-A component of begomoviruses was performed with degenerate primer sets targeting DNA-A genome. Different primer sets were used to amplify various fragments from regions of a 2.4 - 2.7 kb size ss-DNA-A and ~ 1.2 - 1.4 kb DNA-β monopartite genomes of leaf curl causing begomovirus (Table 3.2). The enriched genomic DNA of 19 leaf curl samples of papaya plants were subjected to PCR mediated screening using upper half region (UH region) of DNA-A, the RCA and the betasatellite specific primer set. Out of 19 samples of leaf curl infected leaves, six were identified to be positive for UH region PCR with degenerate primer set amplifying 1.4-1.6 kb region of the DNA-A component (Figure 4.11).

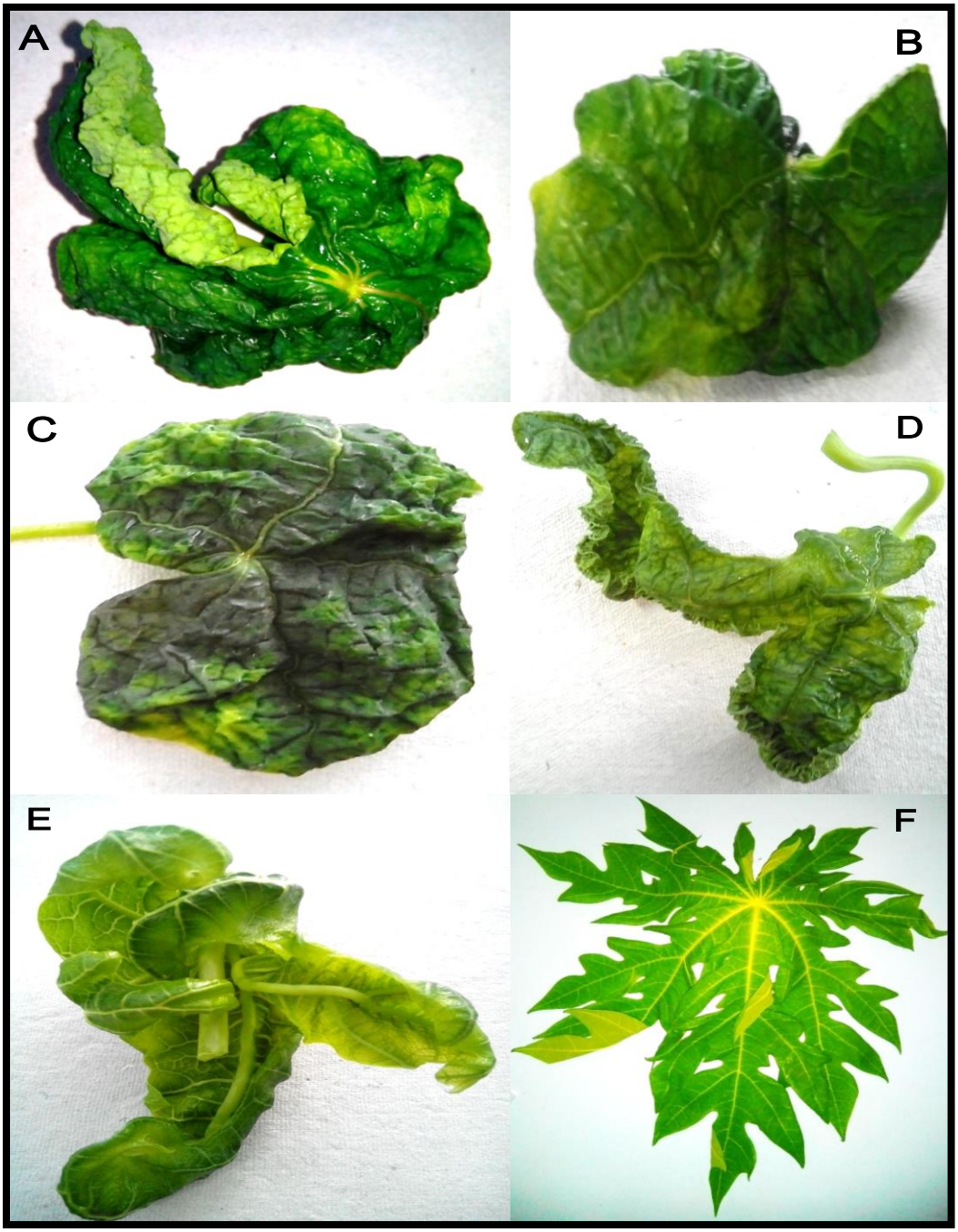


Figure 4.8 Different leaf curl samples collected from various sites in Uttar Pradesh, India. (A) typical lamina curled and twisted, dark green color, vein swelled curled leaf, (B) crumpled, down--ward curling with mosaic patches, (C) leaf crumpled, mosaic patches, dark green color, (D) up-ward curled, lamina crippled, (E) downward curl, vein swelling, slight discoloration, (F) healthy papaya leaf sample.

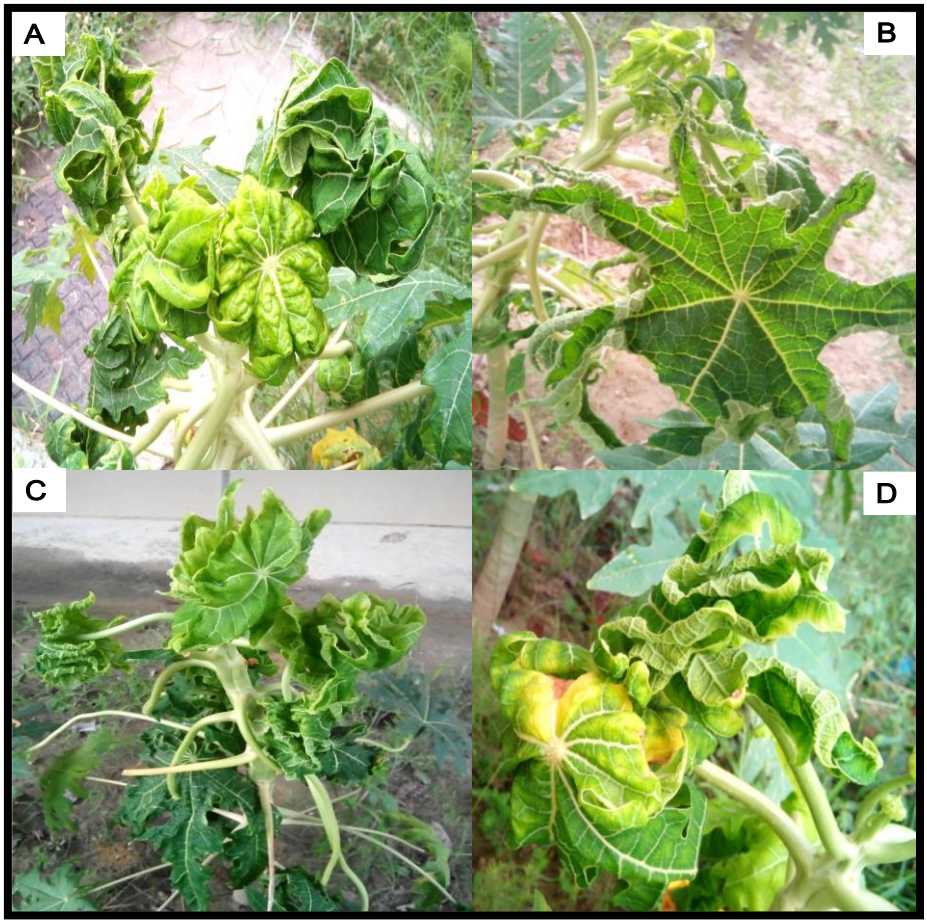


Figure 4.9 Field symptoms observed in papaya leaf curl infected papaya plants form various sites in Uttar Pradesh, India. (A) reduced growth, downward curling, sporadic dark and light green coloration, vein swelling; (B) reduced growth, upward curling, shoestring type leaves, vein swelling; (C) reduced plant growth with prominent mixed leaf curling, vein swelling, dark green color, sporadic chlorosis; (D) reduced plant growth, inward curling, vein swelling, vein clearing, mosaic chlorosis.

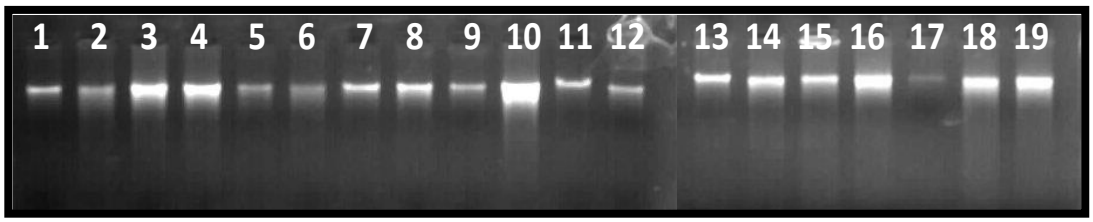


Figure 4.10 Isolation of total DNA from healthy and symptomatic papaya leaves. Total DNA was isolated from papaya leaves using modified CTAB procedure and later enriched using method mentioned in section 3.5.2. DNA samples were resolved on 0.8% Agarose gel electrophoresis to assess quality of viral DNA. Lanes 1-19 represent following samples 1= BR1, 2= BR2, 3= BR3, 4= BR4, 5= BR5, 6= PC1, 7= PC2, 8= PC3, 9= BR6, 10= BR7, 11= BR8, 12= PC4, 13= PC5, 14= PC6, 15= KG1, 16= KG2, 17= CISH1, 18= CISH2, 19= CISH3.

These UH positive samples were subjected to RCA, which produced positive concatamers in all samples. The high molecular weight (HMW) concatamers did not pass through wells i.e. >50kb whereas, low molecular weight (LMW) concatamers resolved on 0.8% agarose gel above 21 kb region (Figure 4.12). Therefore, RCA reaction resulted into concatamers in the range of 7 – 16 times the genome size of DNA-A, DNA-B, DNA-β components (Knierim and Maiss 2007; Wyant 2011). The resulting RCA product was subjected to the screening for the presence of a putative betasatellite component of monopartite begomoviruses. Six RCA positive samples were diagnosed with degenerate primer sets designed to amplify 1.2 – 1.4 kb fragment of betasatellite genome (Table 3.2). Out of the above six samples, four samples from site no. 2, 4, 5 and 8 produced amplicon fragment in the range of ~1.2 – 1.4kb corresponding to betasatellite component genome (Figure 4.13). The two samples i.e. site 14 and 15 (faint band ~1kb) did not amplify DNA-β component, therefore, they might be bipartite begomovirus or might be DNA-A with defective interfering fragments (Patil and Dasgupta 2006).

4.8 Cloning and Sequencing of full-length fragment of DNA-A and betasatellite components of leaf curl sample from Sultanpur

The leaf curl positive UH fragments from samples 2, 4, 5, 8 were cloned into pTZ57R/T vector provided with TA Cloning Kit (Invitrogen). Sample no. 8 was selected for further analysis due to evidence of its association with a betasatellite molecule and characteristic severe leaf curl symptoms of PLCD i.e. vein swelling and severe inward leaf curling. The 1.6 kb fragment was amplified from sample no.8 using primers containing Not1 restriction sites (Table 3.2). The 1.6 kb UH fragment was resolved on 0.8% agarose gel and eluted with gel elution kit. The resulting fragments were ligated at 16℃ for overnight to pTZ57R/T vector plasmid containing A overhangs providing sticky ends for ligation. The resulting plasmid-insert product was transformed in *E.coli* DH5α competent cells. The resulting clones were screened for the presence of inserted fragment using M13U forward and reverse primer sets in colony PCR. A positive fragment of ~1.6 kb confirms the presence of UH

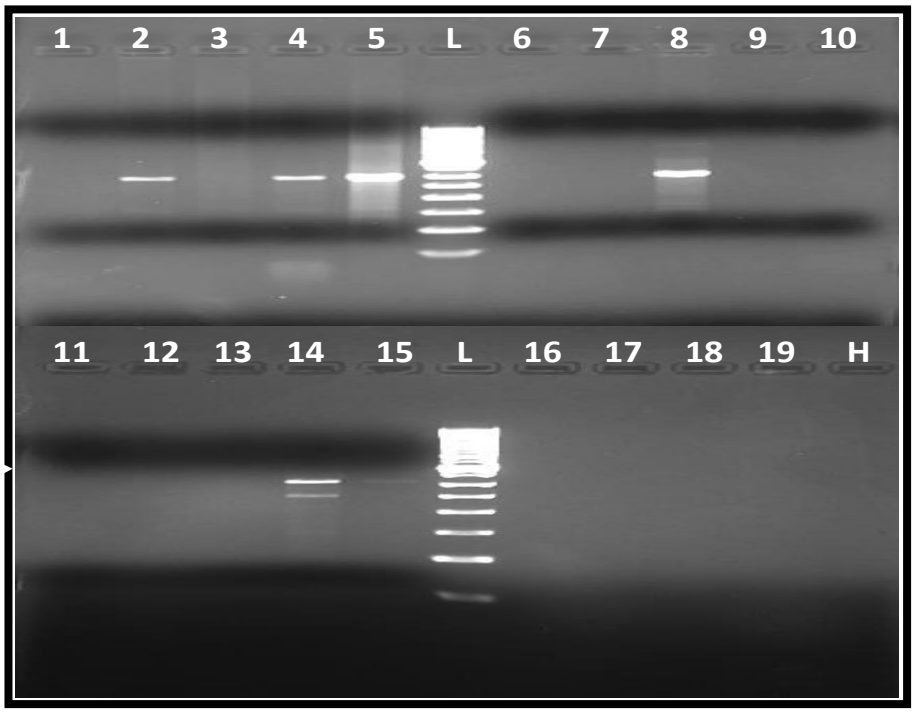


Figure 4.11 Polymerase chain reaction mediated detection of viral components in healthy and symptomatic leaves of papaya plant using UH primer set. PCR mediated screening of total DNA isolated from papaya leaves. Upper-half degenerate set of primers (Table 3.2) were used to amplify ~1.4-1.6kb fragment (indicated with white arrows) from DNA-A component of leaf curl infected and healthy samples. 1= BR1, 2= BR2, 3= BR3, 4= BR4, 5= BR5, 6= PC1, 7= PC2, 8= PC3, 9= BR6, 10= BR7, 11= BR8, 12= PC4, 13= PC5, 14= PC6, 15= KG1, 16= KG2, 17= CISH1, 18= CISH2, 19= CISH3, L= Lambda HindIII/EcorI digested DNA ladder, H=healthy plant DNA.

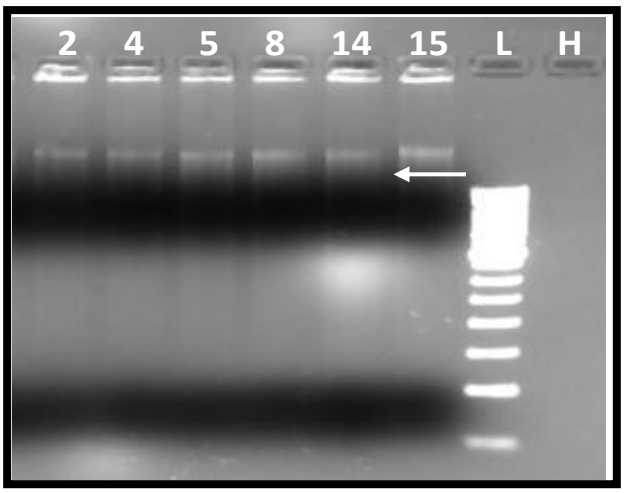


Figure 4.12 Rolling circle assay based detection of UH positive viral components in healthy and symptomatic leaves of papaya plant using betasatellite PCR. PCR mediated screening of total DNA isolated from papaya leaves. High molecular weight (HMW) product in wells, low molecular weight (LMW) product (indicated with white arrow) resolved in 0.8% agarose gel electrophoresis; 2= BR2, 4= BR4, 5= BR5, 8= PC3, 14= PC6, 15= KG1, L= 250bp step up DNA ladder, H=healthy DNA (negative control).

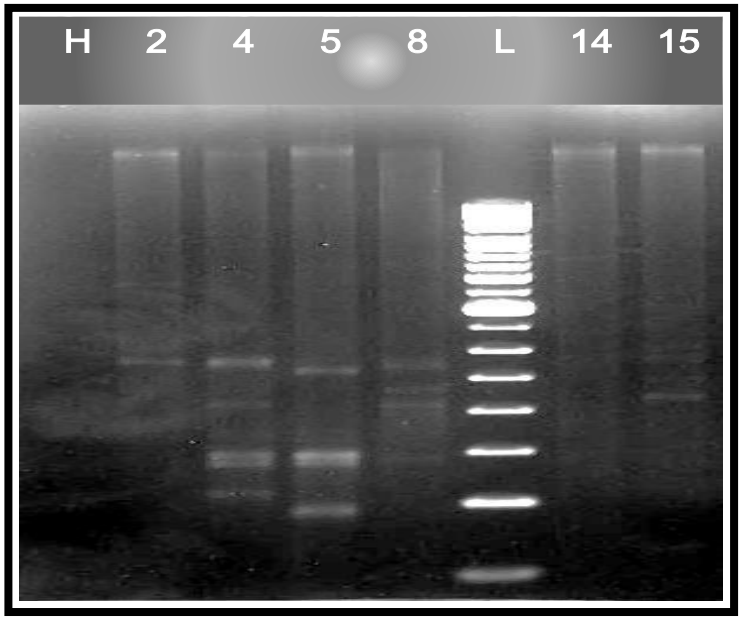


Figure 4.13 Polymerase chain reaction mediated detection of viral components in healthy and symptomatic leaves of papaya plant using betasatellite specific PCR. PCR mediated screening of RCA positive DNA samples isolated from papaya leaves. Betasatellite degenerate set of primers (Table 3.2) were used to amplify ~1.2-1.4kb fragment (indicated with white arrow) from DNA-β component of leaf curl infected and healthy samples. 2= BR2, 4= BR4, 5= BR5, 8= PC3, 14= PC6, 15= KG1, L= 250bp step up DNA ladder, H=healthy DNA (negative control).

region. Total plasmid DNA was isolated from positive clones and Sanger sequencing at Eurofins Inc. was performed. Similarly, betasatellite fragment was amplified from sample no. 8 using above procedure to clone and sequence ~1.3 kb size fragments of betasatellite components.

The sequencing results provided information about the UH and betasatellite regions of the viral component in positive sample no. 8. Abutting primers (Table 3.2) was designed against a common AV2-AV1 encoding open reading frame on viral sense strand to amplify whole viral genome i.e., 2.7 kb fragment through PCR. The resulting 2.7 kb fragment amplified using abutting primers was purified using PCR purification kit. It was later, cloned into pTZ57R/T plasmid and positive clones were sequenced using Sanger sequencing (Annexure 3 and 4).

4.9. Infection clones of DNA-A and betasatellite components are causative agents of leaf curl disease of *Carica papaya L.*

It is important to prove Koch's postulate so that the causative factor for plant disease could be established for a pathogen isolated from infected plant sample. Viral DNA-A and betasatellite

components isolated and identified from papaya leaves showing infection symptoms were introduced into an expression vector in dimeric forms to function as infectious clones. These infectious clones could then be introduced into a host plant to prove Koch's postulate.

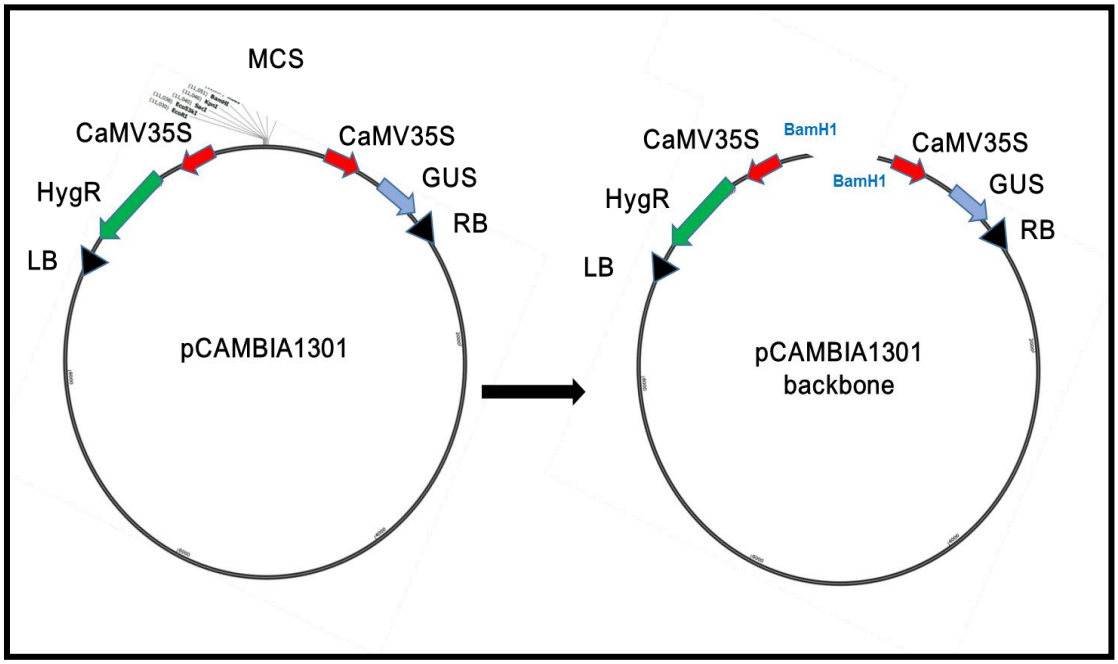


Figure 4.14 Preparation of infectious clones of DNA-A and betasatellite components: Preparation of plant binary vector for insertion of dimeric viral fragments.

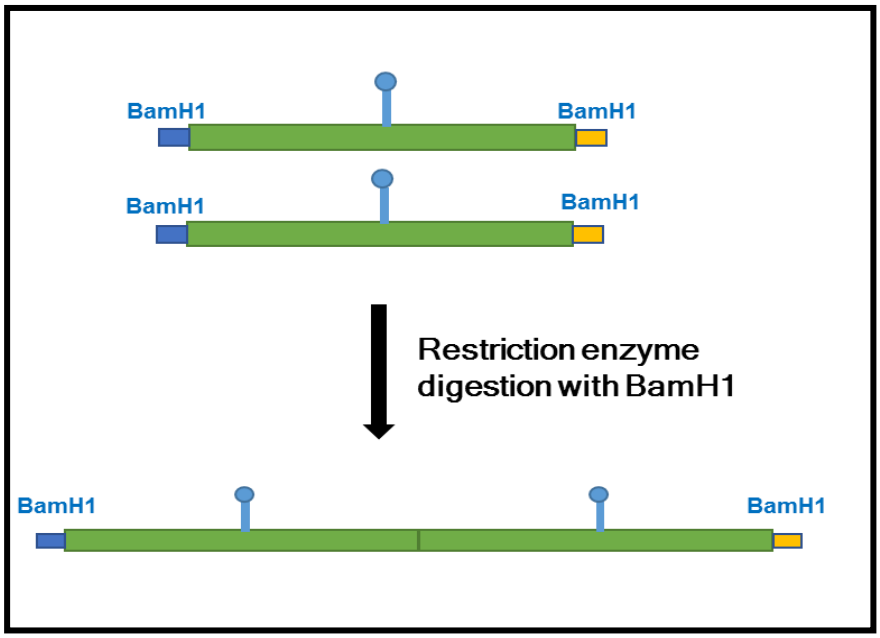


Figure 4.15 Preparation of infectious clones of DNA-A and betasatellite components: Preparation of dimeric fragment (5.4 kb for DNA-A and 2.6 kb for betasatellite) for integration into plant binary vector

The vector backbone for infectious clone was prepared by restriction digestion of pCAMBIA1301 plasmid with BamH1 enzyme to produce 14.7 kb nucleotide product, which was gel purified and dephosphorylated using CIAP enzyme at 37℃ for 15 min in water bath (Figure 4.14). The insertion fragments were prepared using previously PCR amplified products of DNA-A and betasatellite component i.e. 2.7 kB and 1.3 kB fragments. Betasatellite monomeric fragments were subjected to restriction digestion with BamH1 restriction enzyme and ligation was performed for short duration i.e. 30℃ for 30-60 min, to produce a 2.6 kb dimer. The full length PCR amplified 2.7 kB fragment of DNA-A genome was similarly subjected to restriction digestion and ligation to produce a dimer (partial ligation product) i.e. 5.4 kb ligation product (Figure 4.15). The ligated dimeric fragments of 5.4 kb and 2.6 kb were purified using QIAEX II Gel Extraction Kit (Qiagen, GmBH, Germany). Purified dimers were ligated using T4 DNA ligase (Thermo Scientific, Waltham, USA) into a BamH1 restriction enzyme digested pCAMBIA1301 binary vector. The resulting plasmids were named as SA1301-5.4 and SB1301-2.6 for DNA-A and beta satellite respectively (Figure 4.16). Above steps have been illustrated in Figure 4.17, where A-C represent molecular modifications done to obtain infectious clones SA1301-5.4 and SB1301-2.6.

Similar procedure was used to produce infectious clones for *Pepper golden mosaic virus*, a bipartite begomovirus naturally infecting solanaceous plants in Mexico and Central America (Carrillo-Tripp et al. 2007). The plants infected with these infectious clones prepared for investigation related to symptom appearance, resistance and remission. *N. benthamiana* plants infected with the infectious clones were studied upto 35dpi and the role of PTGS was studied. Therefore, dimeric clones were employed for conducting infectivity and siRNA based assays.

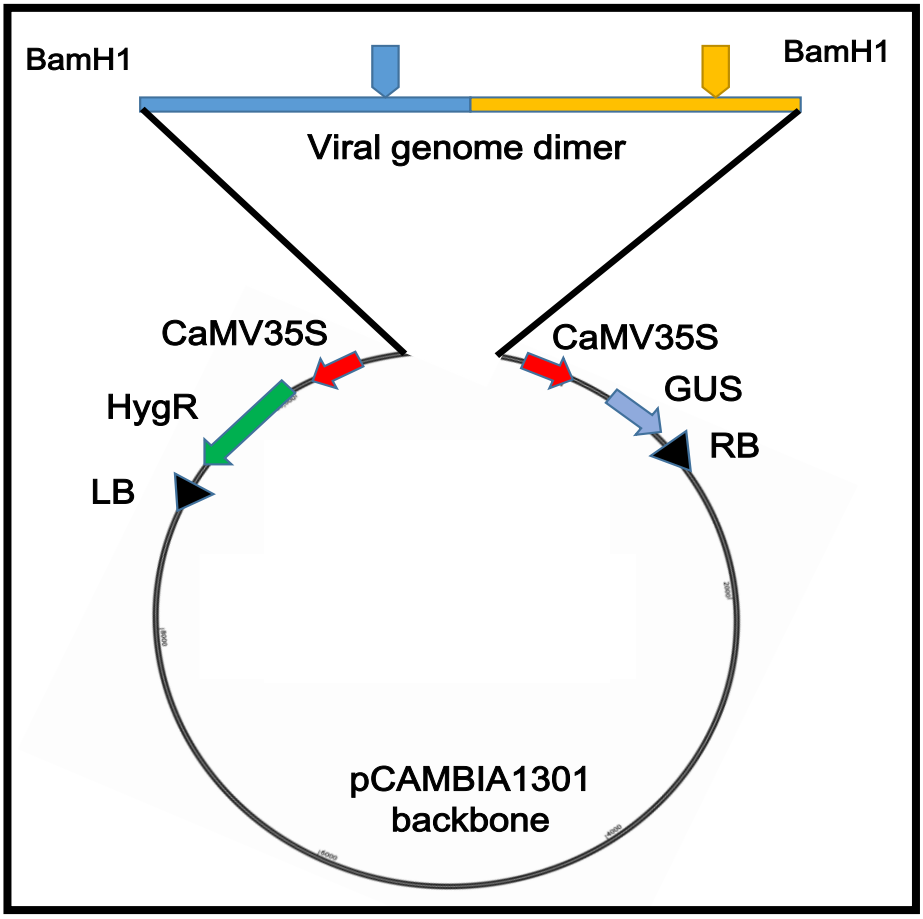


Figure 4.16 Illustration of infectious clones of DNA-A and betasatellite components after ligation of their respective dimeric fragments Full length fragment of DNA-A was amplified using abutting primer set KGF1335 and KGR1336 (Table 3.2). pCAMBIA1301 binary expression vector was linearized and dephosphorylated at BamH1 restriction site. The amplified viral genomes (2.7/1.3kb) cloned in pTZ57R/T were retrieved after cleavage with BamH1 restriction enzyme. Dimerization was performed by performing partial ligation for 30 minutes. The resulting dimer was purified from agarose gel and integrated into previously digested pCAMBIA1301 vector. The resulting plasmids were designated as SA1301-5.4 for DNA-A and SB1301-2.6 for betasatellite component.

4.10. Qualitative and Quantitative estimation of begomovirus DNA-A and betasatellite infectivity in agroinoculated plants.

The plasmids SA1301-5.4 and SB1301.2.6 were transformed into *Agrobacterium* strain LBA4404 competent cells and transferred into 4-5 weeks old *Nicotiana benthamiana* plants using agroinfiltration procedure (Du et al. 2014). Negative control tobacco plants were mock inoculated with pCAMBIA 1301. The plants were maintained in insect proof cages at 22℃ with 16h light and 8h dark cycles of growth conditions until appearance of symptoms. Whitefly mediated acquisition and transmittance procedure was performed as mentioned in section 3.9. The pots were removed

after 24h and kept in separate insect proof cages under controlled conditions as mentioned earlier. The symptoms started developing after 30 day post inoculation (dpi). The samples were collected and tested for presence of viral genome using PCR based diagnostic technique as mentioned in following sections (Verma and Saxena 2017b).

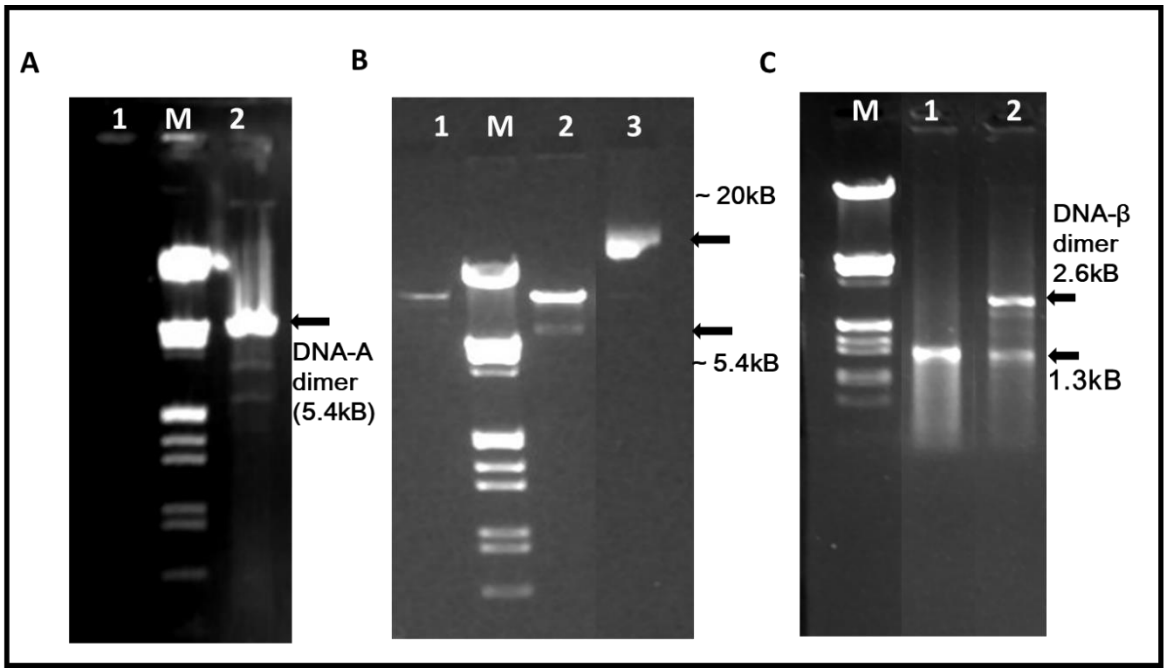


Figure 4.17 Construction of infectious clones of PLCD complex. (A) Ligation of DNA-A dimer: 1=NTP (no template); M= Lambda H/E DNA ladder; 2= DNA-A dimer (30 minutes). (B) Integration of DNA-A dimer in pCAMBIA1301: 1= Linearized pCAMBIA1301, M= Lambda H/E DNA ladder, 2=digested SA1301-5.4 plasmid, 3=SA1301-5.4 plasmid. (C) Dimerization of betasatellite molecule (~1.3kb): M=Lambda H/E DNA ladder, 1=PCR product betasatellite (0 min), 2= ligated dimer product (30 min) and remaining monomer product (30 min).

4.10.1. Infectivity assay in *Nicotiana benthamiana* plant

The plasmids SA1301-5.4 and SB1301.2.6 were mobilized into *Agrobacterium* cells and transferred into *Nicotiana benthamiana* plants using agroinfiltration technique (Du et al. 2014). Both plasmids were introduced individually as well as in *equimolar ratio* in tobacco plants. Negative control tobacco plants were subjected to mock inoculation with pCAMBIA1301plasmid. The transformed plants were grown in insect proof cages at 22℃ with 16h light and 8h dark cycles of growth conditions until appearance of symptoms. The plants were observed for 6, 15 and 36 days post inoculation (dpi) for symptoms and qualitative assay for infectivity of SA1301-5.4 and SB1301-2.6

infectious plasmids in *N. benthamiana*. The symptoms scores assessed were statistically validated using non-parametric tests to derive significance values (p value) due to absence of normal distribution data in the experiments. *F-box* gene was used as a reference gene for qRT-PCR based study for relative quantification of viral replication. Coat protein gene, a component of DNA-A and c1, a component of betasatellite genome were used to estimate the viral load in tobacco plants. The results of symptom based observation supplemented with RCA based virus genome detection were used to calculate **% infectivity = no. of symptomatic plants\total no. of RCA positive plants *100** (Figure 4.18).

Table 4.6 Infectivity assay in *N. benthamiana* to assess potential of infectious clones SA1301-5.4 and SB1301-2.6 to cause disease symptoms. Control plants were WT tobacco plants, Mock were infiltrated in 6 tobacco plants with pCAMBIA1301 vector only, SA1301-5.4 and SB1301-2.6 were infiltrated separately in 16 plants each while equimolar mixture (1:1) S1301-5.4/2.6 was infiltrated in another 16 tobacco plants. The scores were recorded for each plant and mean score obtained was listed in table.

	N	**Non-Symptomatic**	**Symptomatic**	**Mean Score**	**SD (+/-)**	**RCA**	**% Infectivity**
Control	6	6	0	-	-	-	-
Mock	6	6	0	-	-	-	-
SA1301-5.4	16	10	6	1.5	0.5	11	**54.5**
SB1301-2.6	16	15	1	0.5	0.15	6	**16.7**
S1301-5.4/2.6 (1:1)	16	1	15	2.6	0.47	15	**100**

The values for % infectivity was calculated to be 54.5% for SA1301-5.4 infectious clone with 1.5 ±0.5 mean symptom score. The mean score for SB1301-2.6 was 0.5 ±0.15 with 1 incidence of symptomatic plant. The tobacco plants co-infiltrated with both genomes were calculated to show 100% infectivity with mean score of 2.6 ±0.47 (Table 4.6).Thus, the equimolar mixture of infectious clones i.e. S1301-5.4/2.6 was the most efficient to cause leaf curl infection in tobacco plants (Figure 4.19). This result is cohort with disease causing frequency of genomic components in monopartite

begomoviruses causing leaf curl disease (Saeed et al. 2007; Akhtar et al. 2011). The above observation were used to derive statistical significance of infectivity assays in tobacco plants. According to the Kruskal Wallis test and post hoc analysis by Mann Whitney U test, the symptom differences at different time interval was most significant ($p < 0.005$) when comparison of SA1301-5.4 was done against S1301-5.4/2.6. The Wilcoxon Signed Ranks test was conducted to determine the most efficient treatment time interval. The period between 15-36 days was found most significant for the treatment in all cases ($p < 0.5$). Therefore, the viral treatment S1301-5.4/2.6 was studied at time points 15 to 36 days for determination of infection and resistance based parameters. The confidence of data beyond this time interval was inconclusive due to initiation of senescence phenotypic changes and decreased innate immunity of plants (Figure 4.20).

The symptoms produced in *N. benthamiana* by dimeric clones of DNA-A and betasatellite components was studied. The equimolar mixture of both components produced mild symptoms of leaf curl disease with prominent leaf crumpling and appearance of yellow mosaic patterns at later stages of plant development. This signifies the fact that tobacco is not the most suitable host for this virus isolate causing PLCD in central U.P. region of India. The above observations were found to be in correlation with various studies, where infectious clones were introduced in plants and the symptoms were studied for different combinations of begomoviral components i.e. containing DNA-A, DNA-B, betasatellite or any combination of these components (Rochester et al. 1990; Brown et al. 2005; Bhattacharyya et al. 2015). Similarly, infectious clones have been widely employed to study recombination, pseudorecombination and defective genomic component formation in case of mixed begomovirus infections (Stanley 1995, 2004; Fondong et al. 2000; Briddon et al. 2001; Pita et al. 2001; Malik et al. 2011). The symptoms studied through inoculation of infectious clones of these begomoviruses have resulted in confirmation of the infectious clones to be functional and can cause mild symptoms of PLCD in *N. benthamiana* plants. This is an important result, as the RNAi based experiments will be performed in this model plant and the observations will be used to interpret the

efficacy of resistance through introduction of a siRNA producing cassette. In above assays, the CP gene based experiment set-up comprises of WT, Mock, SA1301-5.4 plasmid and S1301-5.4/2.6 equimolar mixture agroinoculated into tobacco plants (3-4 weeks old). Leaf samples were collected from 3 different plants for each sample. Therefore, 3 biological and 2 technical replicates were collected and RNA samples were isolated from 100mg leaf samples.

The DNase-A treated RNA samples were subjected to reverse transcriptase-mediated ds-cDNA production. The final product was diluted to 5ng/µl and 2 µl was added to 25µl qRT-PCR reaction assay mixture. The real time estimation of viral load in infected and control plants showed significant fold change in viral mRNA expression i.e. CP and c1 transcripts. The CP gene primers could probe replication and expression of DNA-A component of PLCD complex in tobacco plant. The c1 gene was used to probe the viral load due to betasatellite component in tobacco plants. Therefore, CP and c1 gene expression was studied for estimation of DNA-A and betasatellite load in tobacco plant respectively.

Figure 4.18 Infectivity assay of infectious clones in *N. benthamiana* (A) Inoculation stage (4-5 weeks) (B) 15 dpi stage SA1301-5.4 only(C) 15 dpi stage SB1301-2.6 only (D) 15 dpi stage S1301-5.4/2.6 (1:1) equimolar mixture.

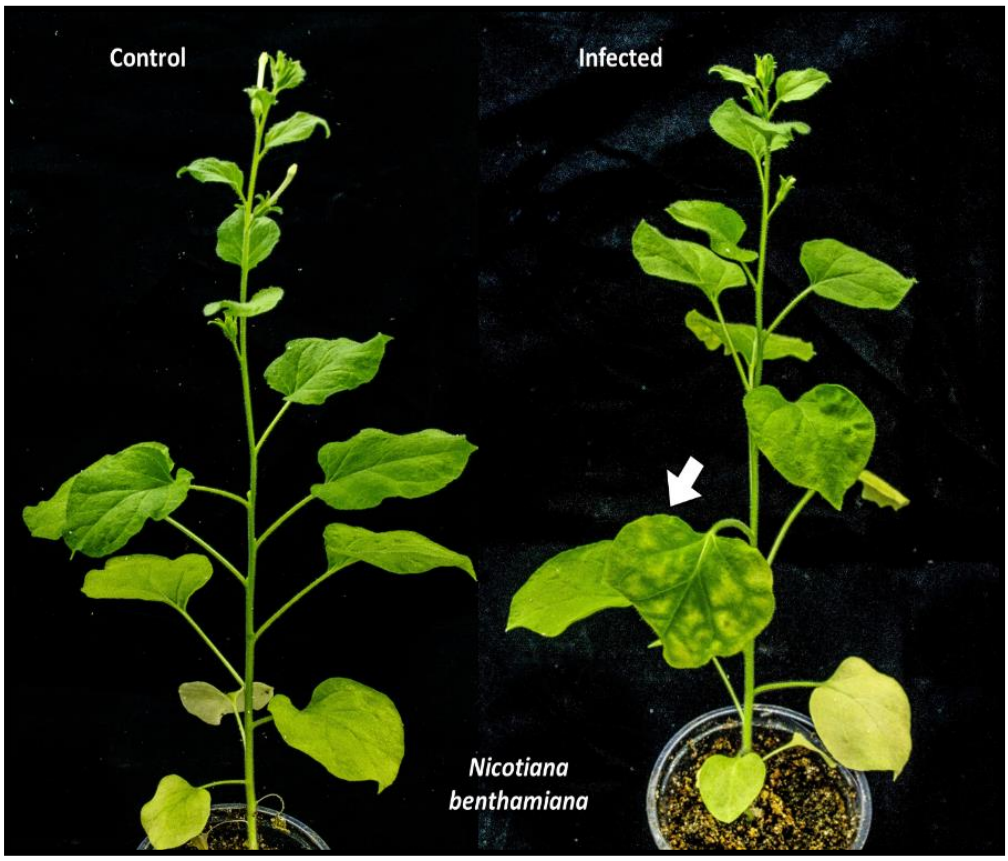


Figure 4.19 Comparison of WT and S1301-5.4/2.6 (1:1) plants at 36 dpi stage. Control (WT) tobacco plants was compared to plants harboring equimolar ratio of S1301-5.4/2.6 virus expression vector at 36th day post inoculation. A reduced plant growth, crumpling, upward curling of lamina and mosaic pattern was characteristic symptom of infection using viral expression vector in *N. benthamiana*.

Figure 4.20 Comparison of S1301-5.4/2.6 (1:1) plants at various post inoculation stages. S1301-5.4/2.6 infected tobacco plants were compared at different days post inoculation. (A) 15 dpi stage (score 2, (B) 36 dpi stage (score 3), (C) 36 dpi stage (score 2) and (D) 50 dpi stage (score 4). All stages show reduced plant growth, crumpling, upward curling of lamina and mosaic pattern. The plants were studied until 50 days, after this stage the plants complete their life cycle.

In present experiment, the CP gene expression was found to be upregulated in infectious component containing tobacco plants. At 15 dpi, the CP expression in plants with DNA-A component only i.e. SA1301-5.4 and equimolar (1:1) betasatellite component i.e. SB1301-5.4/2.6 was found to be upregulated by more than 5 and 15 folds respectively (Figure 4.21 A). This expression pattern was observed until completion of experiments i.e. till 36 dpi stage. The present results were found to be similar to 15 and 36 dpi CP gene expression pattern in symptomatic leaves

upon introduction of PepGMV DNA-A and DNA-B components in chili (Carrillo-Tripp et al. 2007). The temporal changes in CP gene expression in S1301-5.4/2.6 plants were found to be consistently

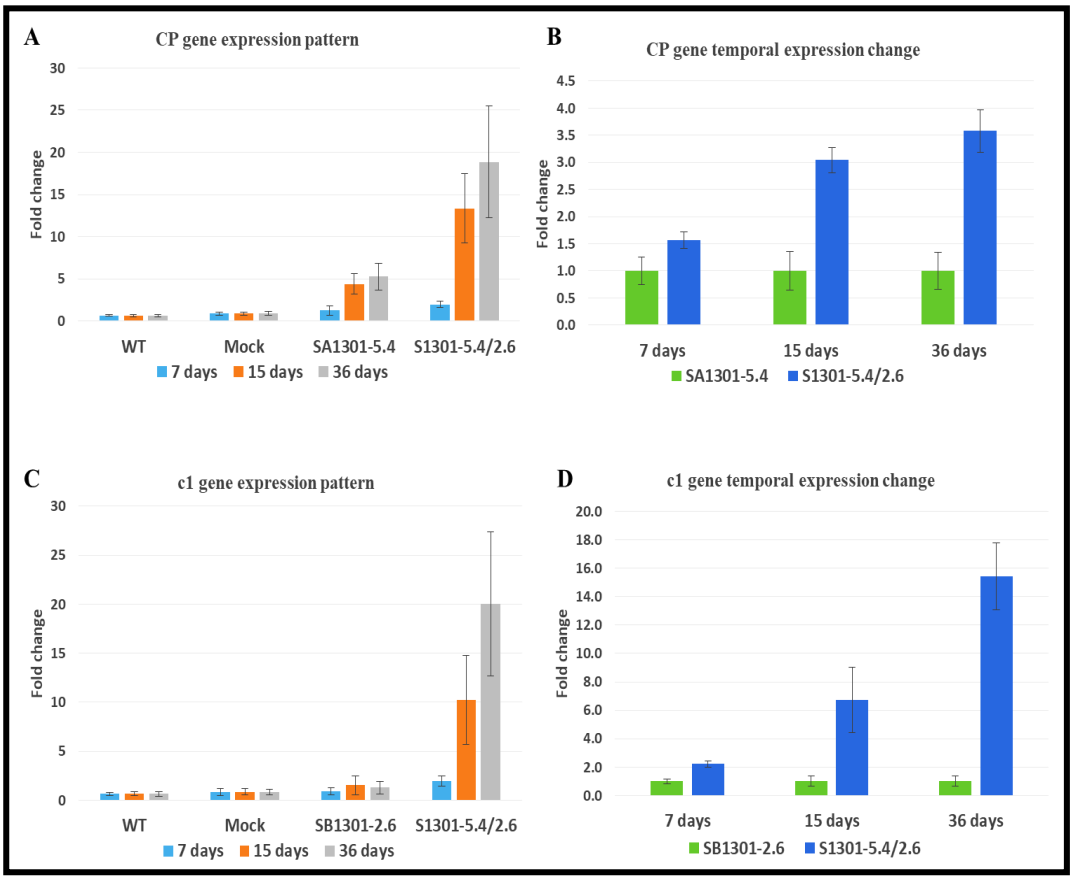


Figure 4.21 Real time analysis of infectivity assay in *N. benthamiana* plants. This figure illustrates comparative expression of viral genes at 7 days, 15 days and 36 days after different treatments in WT, mock and virus genome infiltrated plants. (A) Relative expression of CP gene after different days post infection (dpi), the increase in CP gene expression indicates increasing viral load (DNA-A) in plant samples. (B) Comparative expression of CP gene after infiltration of SA1301-5.4 and S1301-5.4/2.6 at different temporal points, expression of CP in SA1301-5.4 samples was taken as reference. (C) Relative expression of c1 gene after different days post infection (dpi), the increase in c1 gene expression indicates increasing viral load (betasatellite component) in plant samples. (D) Comparative expression of c1 gene after infiltration of SB1301-2.6 and S1301-5.4/2.6 at different temporal points, expression of c1 in SB1301-2.6 samples was taken as reference Relative expression levels of betasatellite increases after 15 days although DNA-A component does not show significant increase.

upregulated by 1.5, 3.0 and 3.5 folds at 7, 15 and 36 dpi when compared against SA1301-5.4 plants (Figure 4.21 B). Similar expression pattern was obtained for c1 gene when relative expression was compared for infectious clones against control plants. SB-1301-2.6 accumulated at lower levels in absence of helper DNA component i.e. SA1301-5.4 (Figure 4.21 C). This result confirms that betasatellite component depends upon DNA-A Rep gene interaction for initiation of its replication

and transcription (Eini and Behjatnia 2016; Zubair et al. 2017b). Therefore, c1 accumulation could be attributed to synergistic interaction with DNA-A component SA1301-5.4 in equimolar inoculation plants, which resulted in 9 and 20 folds increase in expression between 15-36 dpi stages (Figure 4.21 C). The temporal changes in c1 gene expression was found to be upregulated by 2, 6 and 15 folds at 7, 15 and 36 dpi stages respectively (Figure 4.21 D). The high level accumulation of c1 gene transcripts could be attributed to its' functional significance in both long and short range movement of viral particles inside cells (Bhattacharyya et al. 2015; Singh et al. 2016). Thus, resulting in severity of leaf curl, vein swelling and leaf deformation symptoms.

In above experiments, the CP and c1 gene expression at various stages of plant development were studied. The expression levels were consistently high for the equimolar mixture containing both viral components i.e. DNA-A and betasatellite molecules as compared to single genomic component containing plants. This signifies that the symptoms developed using both infectious components most possibly represent the real infection pattern, if not the real symptoms as observed in the papaya leaf samples from Sultanpur region.

4.10.2. Whitefly mediated transmission of PLCD complex components in *C. papaya* L. plants

The procedure adopted for this experiment is similar to that in section 4.10. The only difference in this experiment was that papaya plants were infected instead of tobacco plants during inoculation phase. The samples were collected and tested for presence of viral genome using isothermal PCR based diagnostic technique i.e. RCA at 7, 14, 36 and 50 dpi.

Table 4.7 Transmissivity assay in *C. papaya* L. using s1301-5.4/2.6 equimolar mixture of infectious clones. Control plants were WT tobacco plants, Mock were exposed to non-viruliferous whitefly (pCAMBIA1301 only) challenge in 2 tobacco plants, while 6 plants were exposed to S1301-5.4/2.6 carrying viruliferous whiteflies.

	N	Non-Symptomatic	Symptomatic	Mean Score	SD (+/-)	RCA	% Transmissivity
Control	2	2	0	-	-	0	0
Mock	2	2	0	-	-	0	0
S1301-5.4/2.6	6	0	6	4.3	0.47	6	100

RCA based screening of both symptomatic and non-symptomatic *C. papaya* plants was performed to detect viral genome components. The successful detection of a viral DNA component from symptomatic leaf samples confirmed the completion of whitefly mediated transmittance in WT papaya plants. The results of RCA screening were recorded and % transmissivity was calculated (Table 4.7).

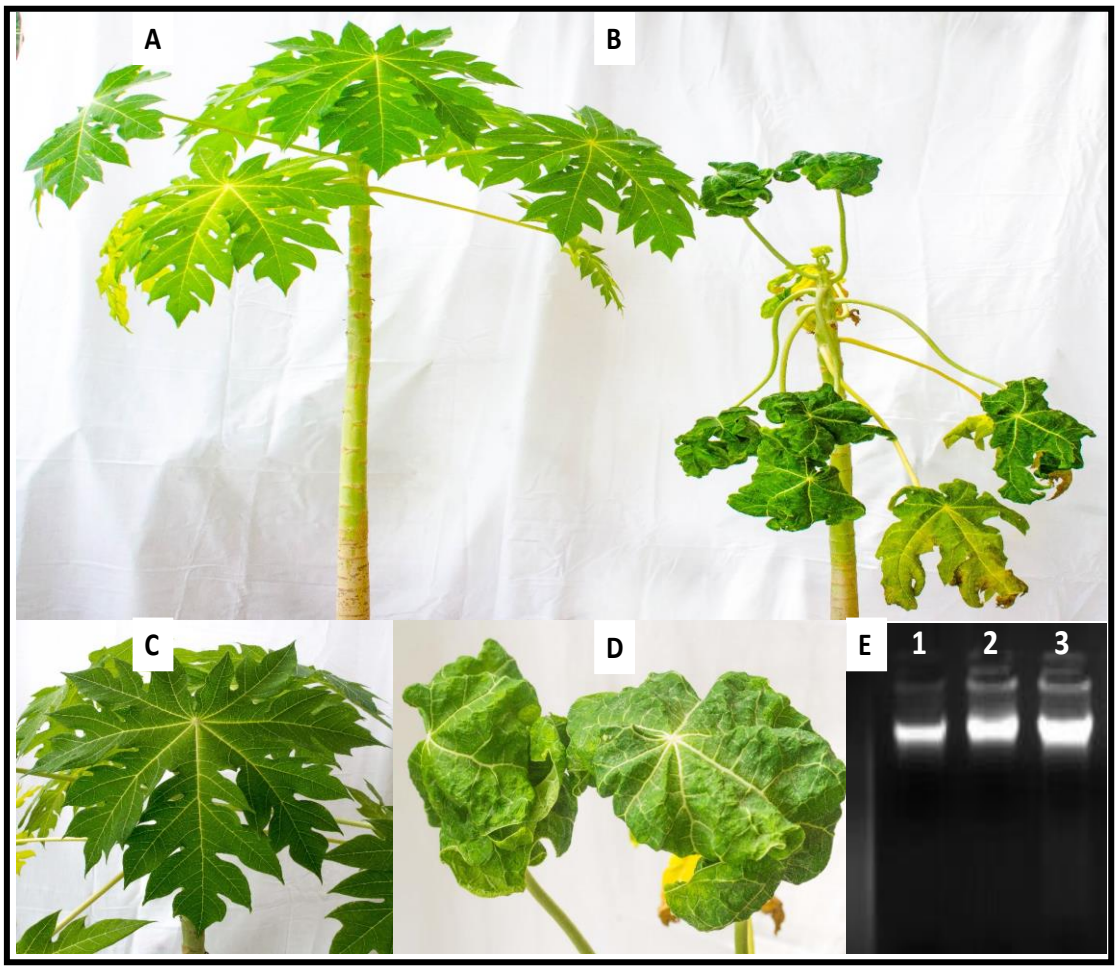


Figure 4.22 Whitefly mediated transmission of DNA-A and betasatellite components into *C. papaya* L. Figure illustrates symptoms developed due to transmission of viral genome in papaya plants after 50 dpi. (A) Control papaya plant (B) infected papaya plant, 50 days after transmission (C) Healthy leaf of control papaya plant (D) infected symptoms show typical crumpled, stunted growth, dark green coloration and inward curling of leaves with curling prominent in fresh leaves (E) RCA detection of positive infection (Lanes 2 and 3; Lane 1=positive control).

All parameters were tabulated for symptomatic and non-symptomatic plants and real time analysis was performed with CP, c1 and F-box gene probes. Transmission efficiency was found to be 100% for plants challenged with viruliferous whiteflies carrying equimolar infectious particles i.e.

S1301-5.4/2.6 plasmid. The symptoms were scored 5 out of scale of 5 with mean score of 4.3 ± 0.47 in all 6 infected papaya plants after 50 days of transmission using viruliferous whiteflies (Figure 4.22 A-E).

The real time analysis was conducted at 7, 15, 36 and 50 dpi plant samples from WT and S1301-5.4/2.6 papaya plants exhibiting scores of zero and 5 respectively (Figure 4.22 A-D). CP and c1 gene expression was found to be upregulated in all samples after 7 dpi also RCA amplicons were detected in all S1301-5.4/2.6 papaya samples (Figure 4.22) The relative expression of CP gene was upregulated by 15 folds at 15 dpi. The same expression reached around 26 folds after 36 dpi and was maintained at same level till 50 dpi. Contrastingly, c1 gene relative expression was upregulated by 10.5 fold at 15 dpi, 25 fold at 36 dpi and 30 fold at 50 dpi (Figure 4.23 A and C). The CP gene temporal expression at 15, 36 and 50 dpi samples was compared to that at 7 dpi. The temporal expression increased by 7 and 9 fold for 15 and 36 dpi, after which it did not show any increase (Figure 4.23 B). The relative expression was found to attain replication plateau as evident by the maximum symptom i.e. score 5, exhibited by papaya plants. The temporal expression of c1 gene was also upregulated by ~6, 13 and 15 fold at 15-50 dpi stage papaya leaf samples (Figure 4.23 D).

Thus, the relative fold change in gene expression of both CP and c1 gene correlates the results obtained for % transmissivity experiment and prove that both genomic components i.e. DNA-A and betasatellite are responsible for successful propagation of this PLCD complex component in papaya through their vectors i.e. whiteflies.

The efficiency to acquire, retain and transmit a begomoviral DNA component depends upon the biotype of whitefly involved in infection. It has been reported that all biotypes are able to acquire and transmit viral DNA component, but transmission efficiency depends upon the biotype of whitefly. The ideal acquisition access period for the biotype Q was found to be between 18-24 h, which resulted in 100% whiteflies with acquired viral DNA. The retention of viral DNA was found

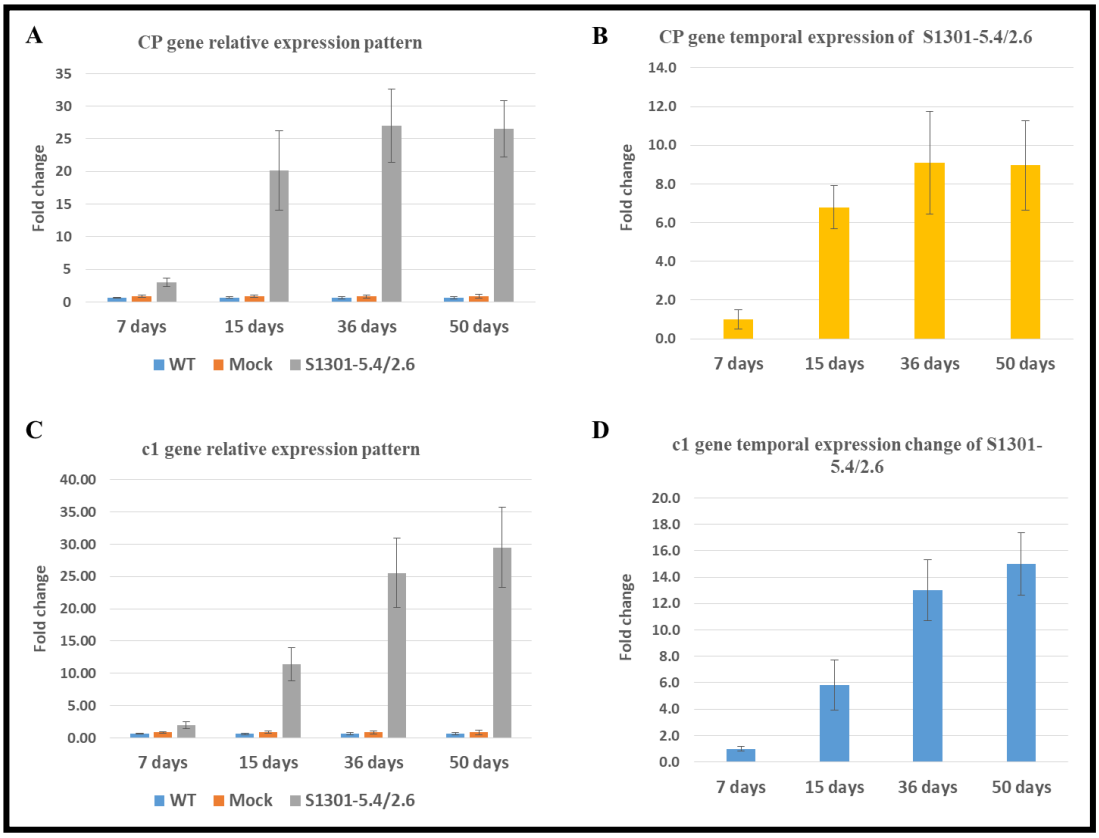


Figure 4.23 Real time analysis of infectivity assay in *C. papaya* L. plants. This figure illustrates comparative expression of viral genes at 7 days, 15 days, 36 days and 50 after different treatments in WT, mock and virus genome infiltrated plants. (A) Relative expression of CP gene after different days post infection (dpi), the increase in CP gene expression indicates increasing viral load (DNA-A) in plant samples. (B) Comparative expression of CP gene after infiltration of SA1301-5.4 and S1301-5.4/2.6 at different temporal points, expression of CP in S1301-5.4/2.6 plant samples when 7 days expression was taken as reference. (C) Relative expression of c1 gene after different days post infection (dpi), the increase in c1 gene expression indicates increasing viral load (betasatellite component) in plant samples. (D) Comparative expression of c1 gene after infiltration of S1301-5.4/2.6 at different temporal points, expression of c1 in S1301-5.4/2.6 samples when 7 days expression value was reference. Relative expression levels of betasatellite increases drastically after 15 days although DNA-A component does not show such a difference at various time points.

to be under ideal range within first 20 h, therefore, the whiteflies were used for transmission phase within first 4 h after acquisition phase (Guo et al. 2015, 2018, Rosen et al. 2015). The transmission rate was found to be >100% when more than 20 viruliferous whiteflies were employed for transmission, hence the exposure of 100 whiteflies in our experiment is in compliance with 100% transmission efficiency of whitefly biotype Q. The 100% transmission rate was observed in above experiment which also signifies that the whitefly biotype Q is actually the transmitting agent for this PLCD complex component.

4.11. BLAST analysis of DNA-A and betasatellite clones

The sequences of full length DNA-A clone and betasatellite components were subjected to BLAST tool based analysis. The DNA-A component had 99% nucleotide identity with *Tomato leaf curl virus* isolate Raebareli DNA-A segment (JX232220.1). Therefore, the present DNA-A component isolated from Sultanpur site in India was named as *Tomato leaf curl virus [India:Sultanpur:Papaya:2016]* [MH105055] abbreviated as *TLCV [India:Sultanpur:Papaya:2016]*, according to the new convention.

The betasatellite component showed 100% nucleotide identity with a *Tomato leaf curl betasatellite-Naj 2 [India:New Delhi:Papaya:2009]* (HM143911.1) associated with *Tomato leaf curl virus* isolate found infecting papaya at Najafgarh site in New Delhi, India.

4.11.1. Molecular investigation of DNA-A component isolated from Sultanpur region

The genomic sequence of DNA-A component of *TLCV [India:Sultanpur:Papaya:2016]* (Annexure 5) was subjected to ORF analysis using SnapGene viewer (http://www.snapgene.com/products /snapgene_ viewer/) software. The software provided following open reading frames corresponding to: AV2 (120-458=339bp) and AV1 (280-1050=771bp) present on viral strand; AC1 (1496-2581=1086bp); AC4 (2248-2424=177bp); AC2 (1174-1593=420bp); AC3 (1047-1454=408bp) and AC5 (310-672=363bp) present on complementary strand of DNA-A genome. The viral strand ORFs encoded coat protein (AC1) 29.5kDa and AV2 protein 12.9kDa in size whereas the complementary strand ORFs encoded Rep (AC1) 40.8kDa, TrAP (AC2) 15.8kDa, Ren (AC3) 16.1kDa, putative silencing suppressor protein (AC4) 6.6kDa and AC5 encoded putative protein 13.1kDa in size. Thus, the circular single stranded DNA-A component contains all the necessary ORFs required for infectivity, survival and transmittance across host compartment.

4.11.2. Molecular investigation of betasatellite component isolated from Sultanpur region

The betasatellite component *Tomato leaf curl betasatellite-Naj2 [India:New Delhi:Papaya:2009]* (HM143911.1) was found to be 1370bp in size, circular, single stranded DNA molecule containing betasatellite associated ORF called C1 (201-563=363bp) encoding a 13.7kDa protein, A rich region (759-971bp=213bp) a stem loop containing conserved region (common to all Begomovirus genomic components) (Annexure 6).

4.12. Identification of siRNA target regions

A phylogenetic analysis of DNA-A component of PLCD complex i.e. *TLCV [India:Sultanpur:Papaya:2016]* was performed using ML algorithm and it was found that it was closest to *Tomato leaf curl New Delhi virus-Papaya [India:New Delhi:Papaya:2005]* in 100% of trees obtained in bootstrap analysis (Figure 4.24). It was found 86% closer to *PapayaLCV* and *ChLCV* group. Therefore, whole DNA-A genome of *TLCV [India:Sultanpur:Papaya:2016]* was investigated for siRNA based targeting using methods as mentioned in section 3.3.6.2, for designing siRNAs. The results were tabulated (Annexure 7 and 8) for total 7 siRNA predictions fulfilling criterion for effective and non-toxic RNA interference. The top siRNA candidate score was 10.42 with one off-target, it was targeting AC3 (1453-1473) coding regions of DNA-A (Annexure 7). While the lowest ranked candidate score was 9.07, targeting AC1/AC4 (2360-2380) protein encoding regions. Maximum siRNA candidates were predicted to target AV1, AV1/AV2, AC1, AC/AC4 protein encoding regions (Annexure 8) i.e. five out of seven, therefore, this region was analyzed further for a generic resistance strategy against PLCD complex.

It is important to mention that the predicted siRNA candidates are not the only population being synthesized by plants machinery. The plants produce a decoy of siRNA which enables effective, broad range targeting of viral genomes. Therefore, the upper half fragment containing viral genome from nucleotide position 1825 – 472 was used as fragment to generate in-vivo siRNA

candidates via hairpin siRNA containing construct. The upper half was also used for this strategy as it showed nucleotide conservation in the range of 73-79%, which is very near to criterion of >85% nucleotide conservation (Figure 4.25). One advantage offered by choice of this region is that it contains a viral hairpin, non-coding region containing various promoter elements and origin of replication of DNA-A genome. Therefore, this fragment was chosen in favor of bottom half region of DNA-A. Also, the siRNA candidates predicted for upper half fragments were less toxic to *N. benthamiana* due to weak or moderate off-target binding as evident from high expect score and higher off-target accessibility score i.e. lower the score better will be off-target-siRNA binding. This fragment has been a choice in many siRNA based strategies where a resistance was developed against more than one type of begomovirus (Bucher et al. 2006; Saxena et al. 2011, 2013; Medina-Hernández et al. 2013). Various non-coding regions have been reported to cause transcriptional gene silencing (TGS), methylation dependent mechanism and other non-canonical pathways have been reported to carry out such a silencing phenomenon (Blevins et al. 2011; Keller et al. 2013; Pooggin 2017). Earlier, it was reported that this region is important in virus replication and provides promoter elements for efficient transcription of viral ORFs into viral mRNAs (Pooggin et al. 2003;

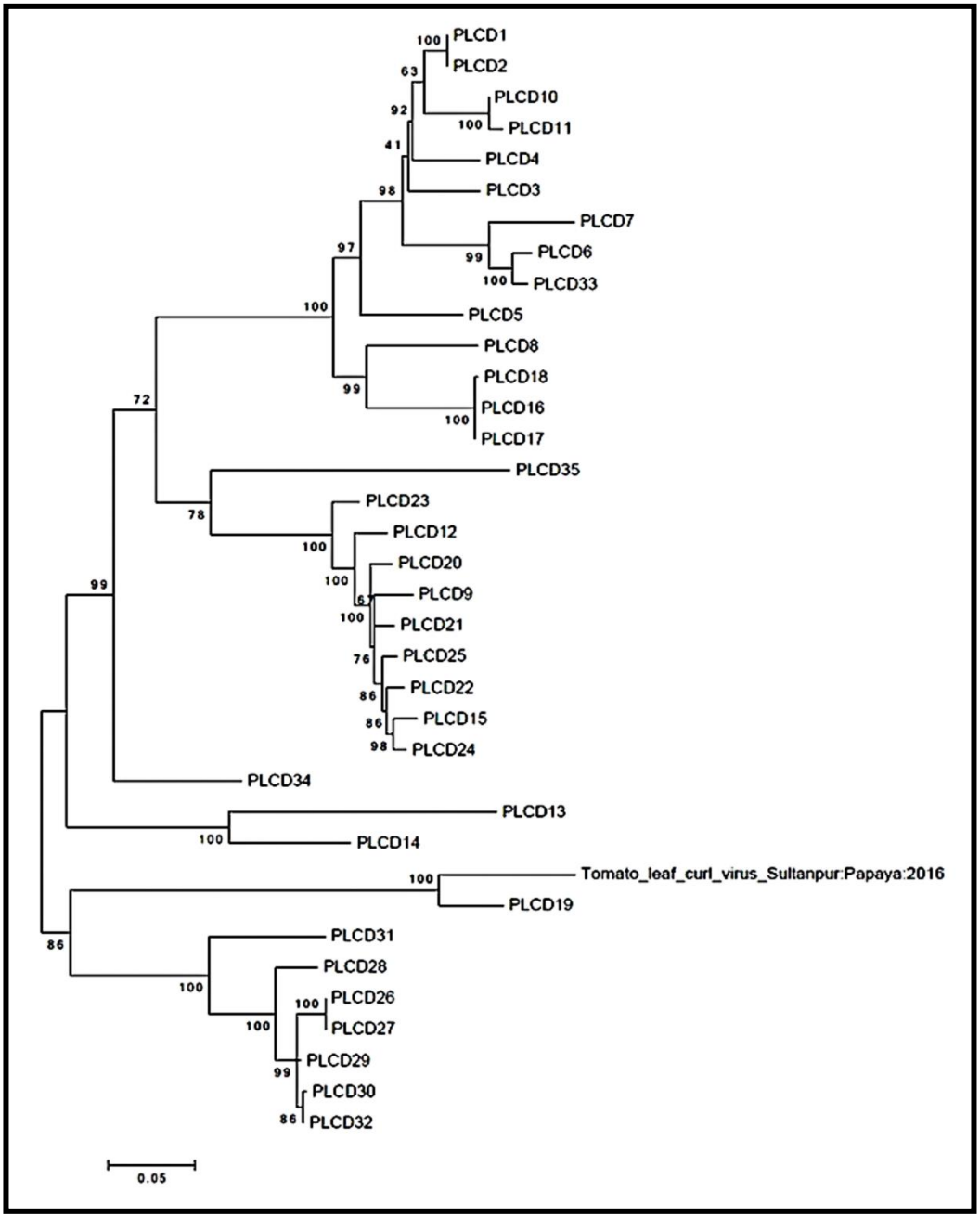


Figure 4.24 Phylogenetic status of DNA-A component in PLCD complex. The evolutionary history was inferred by using the ML method based on the Tamura-Nei model. The tree with the highest log likelihood (-24854.0143) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+*G*, parameter = 0.7466)). The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 10.1469% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 36 nucleotide sequences. All positions containing gaps and missing data were eliminated.

Shivaprasad et al. 2005). Therefore, upper half becomes a natural choice for a generic RNAi based strategy against PLCD complex.

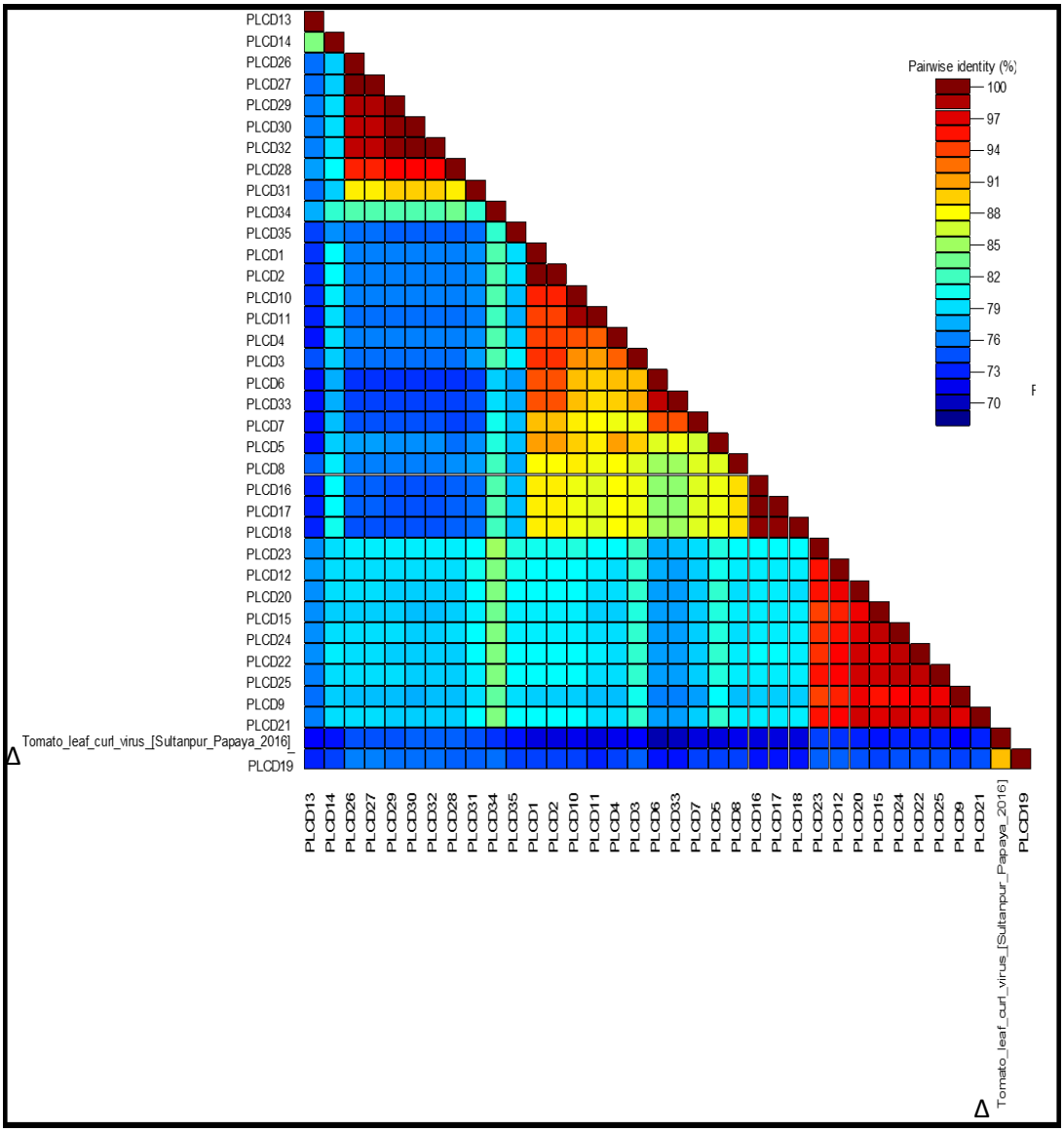


Figure 4.25 DNA-A component of PLCD complex Pairwise identity estimation of sequence similarity using Sequence Demarcation Tool v1.2. *Tomato leaf curl virus [Sultanpur:Papaya:2016]* (indicated by Δ symbol) shows sequence similarity in the range of 73-80% when compared against other DNA-A components of PLCD complex. The matrix output indicates the diversity of begomoviruses infecting papaya and associated crops and weeds. The color key indicates highest and least similarity with brown and blue color respectively.

4.13 Evaluation of siRNA construct pBS-Sul efficacy against PLCD complex components in *N. benthamiana* plants

The plant binary vector containing siRNA cassette (section 3.10) was introduced into *N. benthamiana* plants through agroinoculation and its efficacy was evaluated at different time intervals

after virus challenge. Two control plants (WT) were inoculated with pBI121 only, while 2 were used as WT controls in the siRNA challenge experiment. Six tobacco plants were inoculated with equimolar mixture i.e. S1301-5.4/2.6, which was labelled as virus only and shall be ussed henceforth. The siRNA efficacy was checked in two ways to simulate the *in-vivo* conditions i.e. one plant was inoculated with equimolar mixture containing both viral infectious components and then siRNA construct pBS-Sul was introduced. This plant is labelled as Vir/siRNA and shall be used to describe this particular case. In another siRNA experiment, siRNA construct pBS-Sul was introduced before challenging the plants with viral components. The plants were labelled as siRNA/Vir in this case. Six number of plants were inoculated (see section 3.8) in both cases and studied for symptom appearance, resistance and recovery for 50 days. The leaf samples were recovered for each timeline i.e. 7 days, 15days, 36 days and 50 days keeping 2 biological and 3 technical replicates.

4.13.1 Phenotype and RCA based qualitative evaluation of siRNA efficacy

The plants were observed for symptomatic changes at different time intervals i.e. at 7, 15, 36 and 50 days stage, total DNA was isolated from them for PCR based screening. The phenotypes related to disease symptom development such as curling, vein swelling, chlorosis and vein clearing were observed for each sample and tabulated (Table 4.8).

Table 4.8 RNA silencing efficacy of siRNA construct pBS-Sul evaluated in *N. benthamiana* (36 days). The table represents mean symptom score (with SD values) recorded for 6 biological replicates across time period of 7 days to 50 days in virus, siRNA/Vir and Vir/siRNA treated *N. benthamiana* plants. The increase in mean score of siRNA/Vir treatment at 15-50 days with relative to those of other two treatments owes to the advancement in ageing process of tobacco plants.

	N	Mean Score (Scale 5) (±SD)			
		7 days	15 days	36 days	50 days
Virus only	6	0.33±0.39	1.3±0.47	3.83±0.37	5±0
siRNA/Vir	6	0	0.33±0.42	2.33±0.74	4.83±0.4
Vir/siRNA	6	0.33±0.41	1.33±0.43	3.33±0.45	5±0

The symptom was scored at the scale of 5 (when compared to Virus only positive control) to assess the efficacy of siRNA in reducing symptoms in infected tobacco plants (Figure 4.26). The siRNA/Vir samples were inoculated with siRNA construct 7 days before introduction of virus, therefore, these plants show late and decreased symptom initially but later on i.e. after 30 days stage, the symptoms increased 50% more than virus only plant, owing to ageing and reduced innate immunity. However, further experiments need to be performed to assess the changes occurring at molecular and protein level.

Figure 4.26 Phenotype of virus challenged *N. benthamiana* plants at 15 dpi stage (A) siRNA/Vir tobacco plants in which siRNA construct pBS-Sul was infiltrated 7 days prior to infiltration of S1301-5.4/2.6. (B) Vir/siRNA plants in which pBS-Sul was infiltrated 7 days after virus infiltration. (C) Virus only (S1301-5.4/2.6) infiltrated tobacco plants

The total DNA isolated from leaves was subjected to RCA analysis using pUC19 plasmid and DNA from WT tobacco plant as positive and negative control respectively. After 36 dpi, all 3 WT and Mock plants remained non-symptomatic (Figure 4.27) and correlated well with absence of RCA amplification in their DNA samples (Table 4.9). The tobacco plants with pBS-Sul did produce a detectable amount of RCA amplification product due to presence of binary plasmid in it. But, it cannot be considered significant due to absence of any type of exposure to viral genome in these plants. All S1301-5.4/2.6 plants.

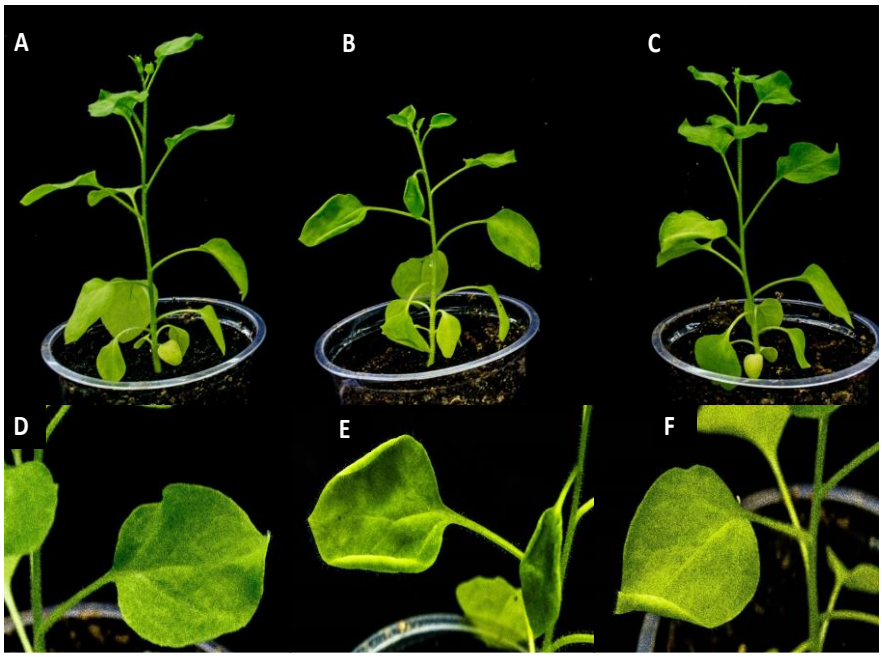


Figure 4.27 Symptoms in N. benthamiana plants at 15 dpi stage (A) & (D) siRNA/Vir plants showing very mild leaf curl symptoms. (B) & (E) Virus only plants showing typical lamina curled with stunted growth and chlorosis. (C) & (F) Vir/siRNA plants showing laminar curling along-with chlorosis and slightly reduced growth.

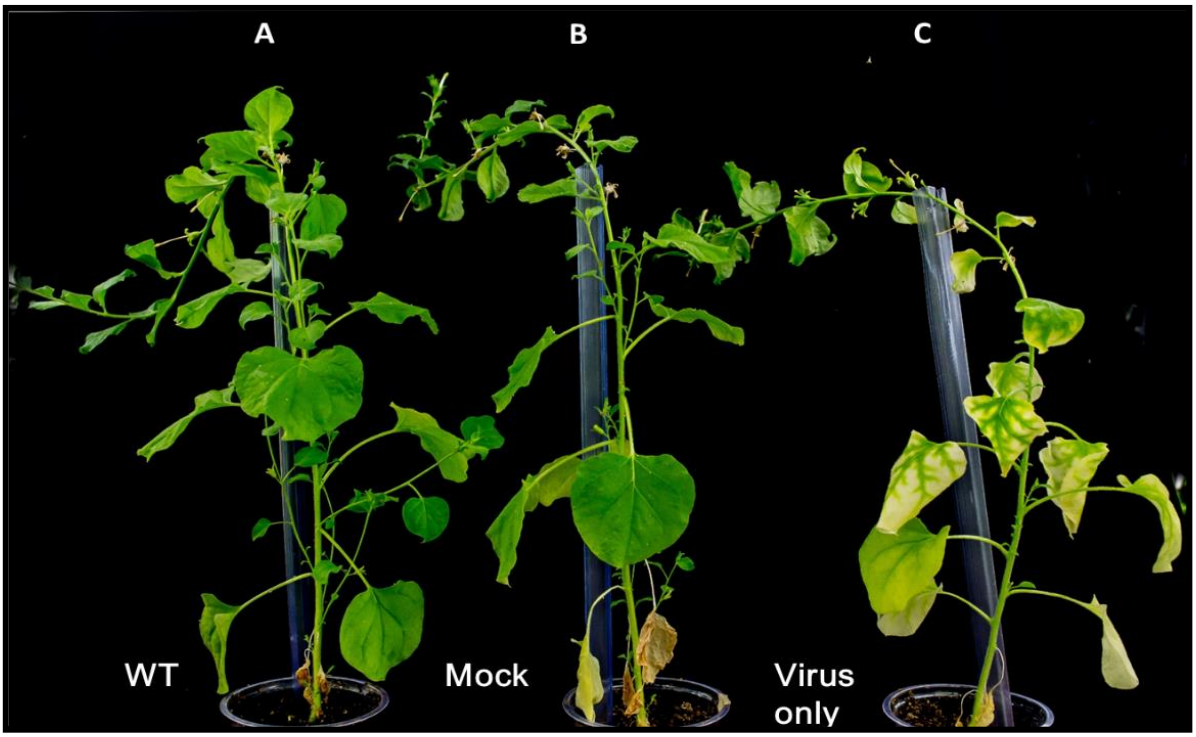


Figure 4.28 Comparative symptoms of *N. benthamiana* plants at 50 days stage without siRNA treatment (~35 dpi) (A) WT tobacco plant. (B) pBS-Sul infiltrated mock tobacco plant. (C) S1301-5.4/2.6 (1:1) plant showing severe chlorosis, leaf curl, reduced plant growth.

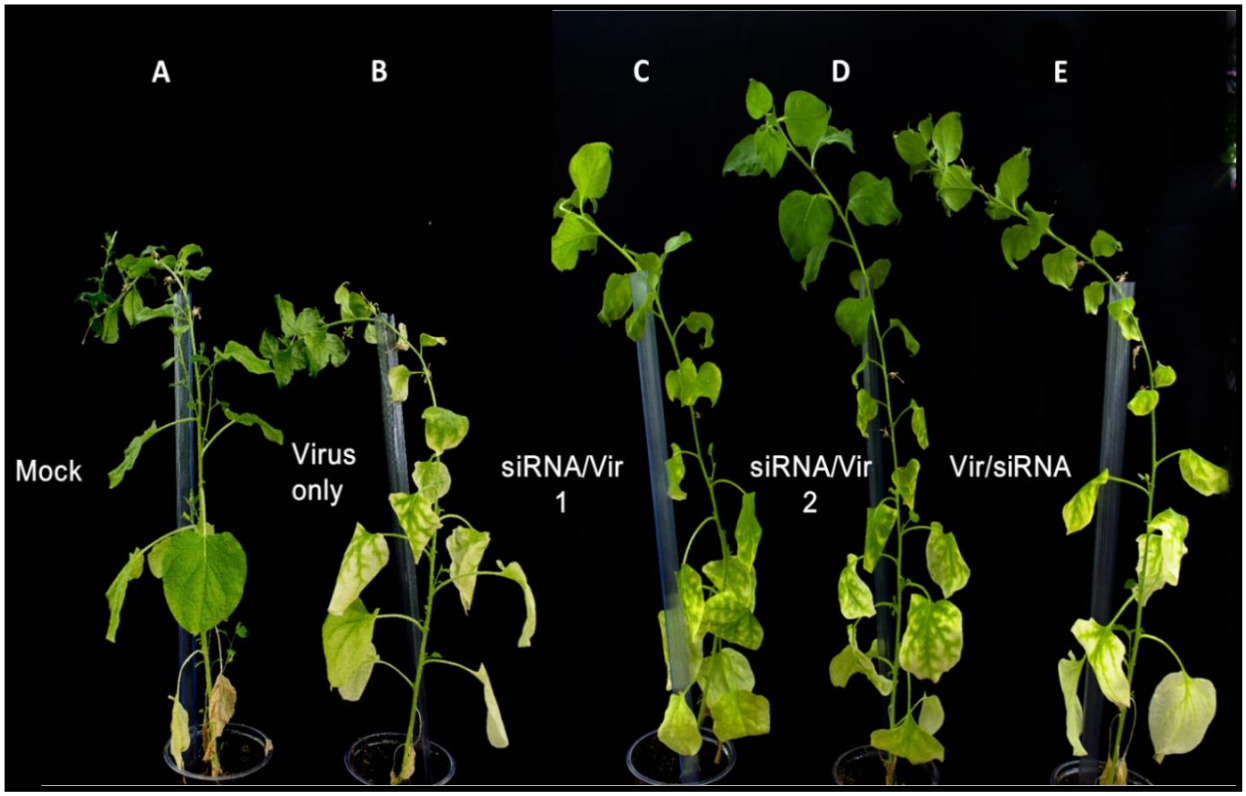


Figure 4.29 Comparative symptoms of *N. benthamiana* plants at 50 days stage (~35 dpi) (A) Mock tobacco plant. (B) Virus only, severe symptoms. (C) & (D) siRNA/Vir plants showing slightly reduced symptoms with late recovery. (E) Vir/siRNA plant with severe symptoms, apparent plant senescence with slight late recovery.

were symptomatic with a score in the range of 4-5 at 15 to 36 days stage i.e. curling of leaf, vein clearing, chlorosis/mosaic appearance with reduced plant growth (Figure 4.27 and 4.29). These plants were RCA positive for all DNA samples (Table 4.9) and also positive for PCR amplification of CP gene using degenerate primer set. When symptoms were compared for siRNA efficacy in siRNA treated plant samples, the 35 dpi stage plants in siRNA/Vir and Vir/siRNA were slightly less symptomatic with reduced chlorosis and showing better upper branch growth. This result indicate slight recovery of plant's innate immunity against viral infection (Figure 4.29).

The symptom attenuation is considered to be an indication of RNAi activity in plants (Zhang et al. 2005; van der Walt et al. 2008). It has been observed that imperfect, chimeric and mutated virus genomes cannot produce efficient and severe infection symptoms. This reduction in infection severity is termed as 'attenuation'. Hence, symptom attenuation is an indication that the virus is not producing the desired symptoms resulting in to attenuated infection.

Table 4.9 RNA silencing efficacy of siRNA construct pBS-Sul evaluated in *N. benthamiana* (36 days). The control plants are WT tobacco plants, mock plants were infiltrated with pBS-Sul (siRNA vector only). Equimolar mixture S1301-5.4/2.6 was used as positive control for virus infection. The siRNA treatment was simulated in two way to check its efficacy against begomoviral components. First, siRNA was infiltrated before virus challenge at 7 days stage while the virus was infiltrated at 0 days stage i.e. 7 days prior to siRNA infiltration.

	N	Non-Symptomatic	Symptomatic	RCA			% Infectivity		
				7 days	15 days	36 days	7 days	15 days	36 days
Control	3	3	0	0	0	0	-	-	-
Mock	3	3	0	0	0	0	-	-	-
Virus only	6	0	6	3	6	6	50	100	100
siRNA/Vir	6	0	6	0	2	6	0	33.3	100
Vir/siRNA	6	0	6	3	6	6	50	100	100

The phenotypic observations in above experiment i.e. the reduction in mean score of 2.33 indicate that the symptom attenuation occurred in siRNA/Vir plants as compared to Virus only and Vir/siRNA plants having mean score value of 3.83 and 3.33 respectively (Figure 4.29). Although, virus was inoculated at later stages of plant development in siRNA/Vir plants, the innate immunity of all tobacco plants was supposed to be same at same developmental stage. Therefore, attenuation in symptom severity at 36 days stage in siRNA/Vir plants could be attributed to prior introduction of pBS-Sul containing siRNA cassette under constitutive expression of CaMV35S promoter. It was observed that the virus only plants did not show even a slight symptom attenuation at same stage as the siRNA treated tobacco plants as evident from increase in symptom mean score at all stages (Figure 4.29 B). Therefore, at this stage the apparent attenuation of symptoms can be seen as induced via introduction of siRNA construct pBS-Sul. However, the level of accumulation of viral genome particles and transcripts would be a fair indicator of the efficacy of siRNAs in producing PTGS in tobacco plants against PLCD component DNA-A and betasatellite molecules.

4.13.2 Quantitative estimation of siRNA efficacy using qRT-PCR based assay

The quantitative estimation of siRNA efficacy to suppress viral transcripts was performed with the help of CP, c1 and F-box primer probes in qRT-PCR. The decrease in relative expression of viral transcripts indicate effect of PTGS mediated suppression of viral transcripts in tobacco plants. RNA isolation was performed for samples at each time point i.e. 7, 15, 36 and 50 dpi and the cDNA samples were prepared and stored at -20℃ until required. A qRT-PCR was performed for CP gene and c1 gene probes of DNA-A and betasatellite components of PLCD complex. Real time analysis was performed against 2 biological and 3 technical replicates for each sample at different time points.

The non-symptomatic WT and Mock tobacco plants did not show any significant up or downregulation in CP nor c1 gene expression at any time point (Figure 4.30.A&C). Since, the CP and c1 relative gene expression show more than 15 fold change in gene expression when compared with the F-box gene background, therefore, a relative change in gene expression was studied. The gene expression was compared against virus only CP and c1 gene backgrounds as represented by blue and orange bars (Figure 4.30.B&D) whereas a comparison was done in background of Vir/siRNA treatment as shown by grey bars, at different time intervals i.e. 15-36 days stage. The siRNA/Vir treatment was shown to exhibit nominal reduction in CP and c1 gene expression i.e. 0.24 and 0.15 fold at 15 day stage but gradually the expression the level equivalent to the background until 50 days stage i.e. between 1-1.1 fold of virus only plant samples. The Vir/siRNA plant samples showed equivalent levels of CP and c1 gene expression when compared to background samples. The c1 gene was found to be slightly more upregulated in case of 15 day samples, which gradually got nullified due to enhanced c1 levels in virus only plants at later stages (Figure 4.30.D).

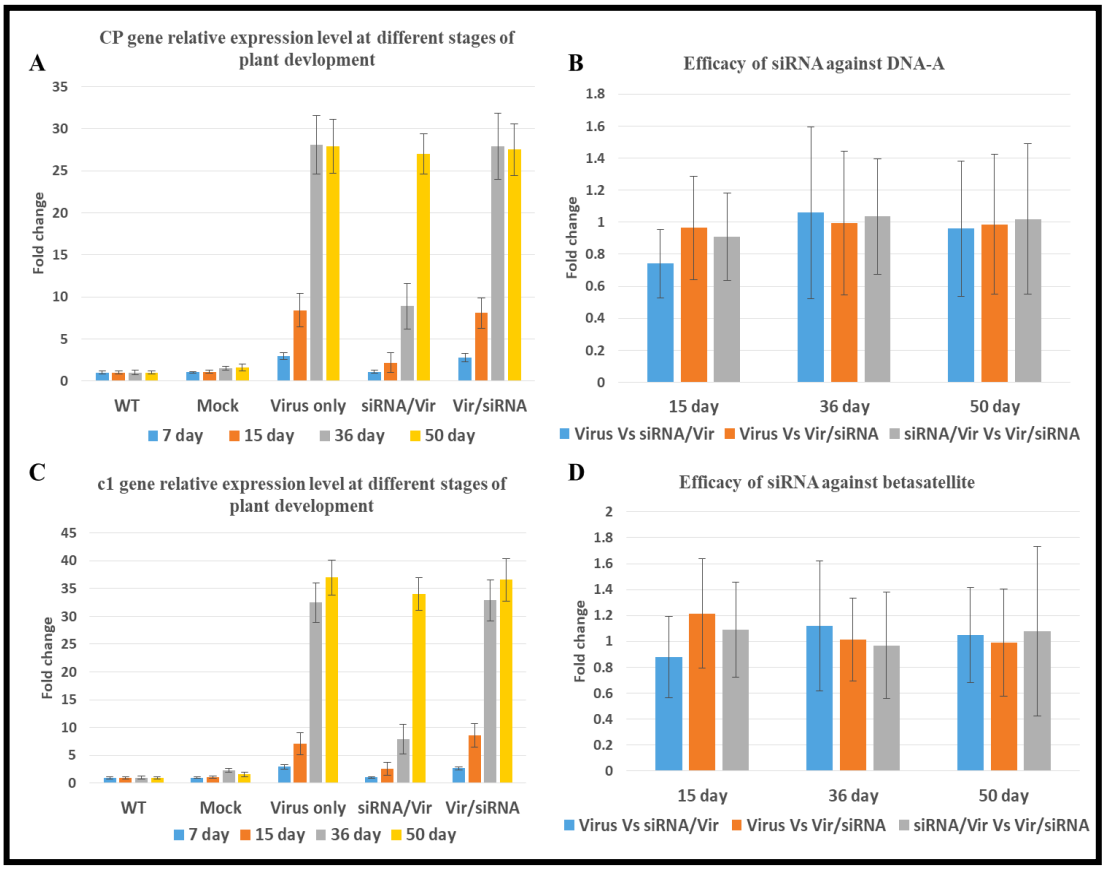


Figure 4.30 Quantitative estimation of siRNA efficacy using qRT-PCR analysis (A) Expression of CP gene in siRNA treated, virus infected, mock treated and control WT *N. benthamiana* plants. (B) Efficacy of siRNA was analyzed by comparing the increase or decrease in expression of CP gene in relation to virus only plants. The grey bars present comparison of siRNA treatment relative to siRNA/Vir treated plant. The comparative difference in CP gene expression indicates reduced/increases DNA-A accumulation. (C) Relative expression of c1 gene in siRNA treated, virus infected, mock treated and control WT *N. benthamiana* plants. (D) Efficacy of siRNA was analyzed by comparing the increase or decrease in expression of c1 gene in relation to virus only plants. The grey bars present comparison of siRNA treatment relative to siRNA/Vir treated plant. The comparative difference in CP gene expression indicates reduced/increases DNA-A accumulation.

The grey bars represent comparative change in gene expression under various siRNA treatments with Vir/siRNA taken as background sample. The comparison indicate slightly reduced CP gene expression at 15 day stage, which gradually levels to 1 at later stages. The change in c1 expression did not follow any trend and show random nature i.e. increased levels decreased at 36 day stage but reached to level 1 at later stage.

The statistical analysis was conducted for above siRNA treatments i.e. two-way ANOVA to assess the significance of CP and c1 gene expression at various time intervals and also among themselves. The p value obtained for above test was more than 0.05 i.e. the expression changes were

not found to be significant. Hence, the treatment cannot be considered effective in providing resistance against begomovirus component of PLCD complex.

It was observed that CP gene transcript level has no correlation with that of c1 gene, as the replication of betasatellites is dependent entirely on the Rep gene transcript levels in host cell (Zhang et al. 2016). Therefore, the c1 gene seems to express more at later stages (after 30 dpi) of viral infection and gradually increase their population even when the helper genome i.e. DNA-A, expression reaches its plateau. It has been shown that DNA-A of *Cotton leaf curl Multan virus* and *Cotton leaf curl Burewala virus* were unable to suppress RNA silencing mechanism of cotton plants, however, the suppression was successful in presence of an associated betasatellite component (Amin et al. 2011). In another study, *Cotton leaf curl Burewala betasatellite* associated with an old world begomovirus was able to transreplicate efficiently in presence of a new world bipartite begomovirus i.e. *Cabbage leaf curl virus*, after few generations due to introduction of few mutations in betasatellite component (Nawaz-ul-Rehman et al. 2009). Therefore, an elevated level of c1 gene expression depicts a natural mechanism exhibited by betasatellite components i.e. they adapt to helper DNA component and utilize it for their own replication. Thus, mutations and host RNAi suppression seems to be a possible mechanism by which betasatellite components interfere with siRNA based resistance strategy. Therefore, this seems to be a possible reason for symptoms severity at later stages of virus infection in PLCD complex.

The results of siRNA efficacy indicate that the siRNA can be efficient in controlling viral population only when introduced in host plants at-least 7 days before viral entry (Figure 4.12.5.B). The expectation that siRNA decoy production in response to introduction of siRNA hairpin cassette corresponding to AC1, AC4, AC5, AV2 and AV1 ORFs of PLCD complex DNA-A component could suppress virus accumulation inside host cell was fulfilled at early stages i.e. 7-15 dpi, as evident from decrease in fold change as compared to virus only plants. But, abrupt increase in CP

and c1 transcripts level by more than ~9 fold (Figure 4.30 C) indicate that the induced silencing mechanism due to pBS-Sul construct had failed to control viral accumulation.

This failure could be attributed to large amount of betasatellite accumulation in mid-to later stages of plant development and virus infection due to their role in host RNAi suppression (Amin et al. 2011). It was also observed that RNAi hairpin constructs were able to inhibit systemic movement when the plants were infected with only DNA-A component of *Cotton leaf curl Multan virus*. But, the same constructs failed to produce mildest of resistance when the plants were infected with both DNA-A and *Cotton leaf curl Multan betasatellite* molecules (Mubin et al. 2011). Therefore, betasatellite have ability to compromise the host resistance machinery by suppression of RNAi components and enhancing accumulation of V-sRNAs (Yang et al. 2011a). Growing evidences of betasatellite adapting rapidly by causing multiple mutations might result in to an efficient co-operation with helper DNA. This causes the betasatellite to transreplicate and adapt to the new environment and helper DNA (Nawaz-ul-Rehman et al. 2009).

Plant RNAi machinery is known to act in two ways i.e. via PTGS and TGS pathways. TGS pathway is mediated by the methylation mechanism in which epigenetic modifications on viral transgenes act as signals for the plant RdRP's to promote siRNA production and thus result into TGS. A study has shown that c1 gene of *Tomato yellow leaf curl China betasatellite* (TYLCCNB) can complement the AL2 gene function of *Tomato leaf curl China virus* (TYLCCNV) by interacting with S-adenosyl homocysteine hydrolase (SAHH) enzyme by inhibiting it's in vitro activity (Yang et al. 2011b). Other mechanisms, which might be involved in host RNAi suppression, have been discussed in review of literature section of this thesis.

In this study, the mechanism and components involved in symptom attenuation could not be predicted at this stage due to limited knowledge about replication, transcription, translation and RNAi suppression components of these PLCD complex components. However, it was observed that

the plant DCLs' have preference for the type of siRNAs produced in response to a particular virus as evident from a comparative study on TYLCV and *Tomato latent virus* (TLV). Here, the tomato transgenic plants carrying intronic hairpin constructs were found to be symptomless in case of TLV infection whereas TYLCV caused severe infection (Fuentes et al. 2006a, 2016). Further, betasatellite c1 gene is known to interact with ASYMMETRIC LEAVES 2 (AS2) (Yang et al. 2008) and ubiquitin conjugating enzyme (SlUBC3) (Eini et al. 2009) and suppress plant defense pathways to increase its accumulation. A major mechanism responsible for such an interaction is the ability of betasatellites to forms multimeric units, which help in formation of complexes with above mentioned host molecular factors (Cheng et al. 2011). Therefore, the observed attenuation of symptoms could be a possible case of betasatellite-mediated suppression of host RNAi machinery.

Conclusion

5. CONCLUSION

This study is significant as this is the first time that the leaf curl diseases in papaya was shown to be forming a PLCD complex. In this study, begomovirus causing leaf curl symptoms collected from Sultanpur region was identified, isolated, cloned and sequenced. The sequenced full length begomovirus fragment was found to be associated with a betasatellite causing PLCD. The infectious clones of both DNA-A and betasatellite components of PLCD complex developed exactly the same symptoms in experimental papaya plants as observed in the infected papaya plants (from field) used to isolate both components. Later, through BLAST analysis, the DNA-A was found to have 99% sequence similarity with a tomato leaf curl virus isolate from Raebareli hence, DNA-A component was renamed as *Tomato leaf curl virus [India:Sultanpur:Papaya:2016]* [MH105055] (a Sultanpur isolate) and betasatellite component was found to be 100% similar to *Tomato leaf curl betasatelllite-Naj2 [India:New Delhi:Papaya:2009]* [HM143911.1] hence the name was used as such.

Due to 75-80% sequence similarity of the Sultanpur isolate DNA-A component with other components of PLCD complex, this isolate was chosen for a generic resistance strategy based on siRNA dependent PTGS. The siRNA based generic strategy was shown to slow down begomoviral infection thus providing resistance against PLCD complex. Based upon this study, following can be concluded:

1. PLCD group of begomoviruses constitute a disease complex in Indian sub-continent.
2. Different symptomatic leaf curl disease infected papaya samples were collected from central Uttar Pradesh region in India.
3. 6 samples were found to be positive in PCR for begomovirus.
4. The Sultanpur isolate was chosen for cloning and to prove Koch's postulate as it shows the most severe and prominent infection in papaya plants.

5. The viral isolate was genetically similar to a leaf curl causing begomovirus, which was later named as *Tomato leaf curl virus [India:Sultanpur:Papaya:2016]* [MH105055] and *Tomato leaf curl betasatelllite-Naj2 [India:New Delhi:Papaya:2009]* [HM143911.1] due to their nucleotide similarity score.
6. Infectious clones containing DNA-A and betasatellite in dimeric form was constructed and named as p1301-5.4 and p1301-2.6 respectively.
7. The begomoviral components were successfully transmitted via Whitefly from infected tobacco plants to healthy papaya plants..
8. PLCD is caused due to DNA-A and betasatellite components of a monopartite begomovirus responsible for leaf curl disease in India.
9. siRNA expression cassette containing binary vector pBS-Sul was constructed and introduced into tobacco plants using agro-infiltration.
10. siRNA/Vir plants showed better resistance against PLCD complex genomic components as compared to Vir/siRNA plants, which were previously inoculated with viral mixture.
11. siRNA cassette against upper half (UH) region could effectively suppress virus replication.
12. The systemic movement of virus was inhibited in presence of pBS-Sul as evident from healthy, non-symptomatic leaves in upper portion of tobacco plants.
13. pBS-Sul might require further fine tuning by trimming the large sequence into smaller overlapping fragments, which are capable of producing siRNAs.

Summary

6. SUMMARY

This thesis work involves study on development of a resistance strategy against begomoviruses causing papaya leaf curl disease (PLCD). Papaya occupies status of an important cash crop of India and one of the important source of foreign revenue through export across world. In India, leaf curl disease severely affects papaya production. PLCD is caused by begomoviruses such as PaLCuV, PapLCV, ToLCV, ChLCV and other associated viruses. The proposed study employed both molecular and *in-silico* approaches to dissect molecular complexity of PLCD and disease resistance potential of the siRNA strategy. In past, many approaches based on siRNA based tools were proposed like artificial miRNA, hp-siRNA and anti-sense RNA, but most of them enjoyed partial success in eliciting resistance against evolving genomes of begomoviruses. Therefore, generic resistance based approach was evaluated for potential to produce efficient siRNA based silencing in plants.

In silico analysis of the DNA-A of begomoviruses, causing PLCD was conducted to evaluate their genetic complexity and recombination patterns within the group. Within the group of begomoviruses of PLCD, there were 11 sub-groups, where PaLCuV, PapLCV and ChLCV/PaLCuV were the major groups forming the PLCD complex. Further, 12 recombination events were observed in 29 predicted viral recombinant sequences signifying contribution of this phenomenon in enhancement of genetic complications of PLCD complex. Therefore, PLCD shall be considered a disease complex i.e. disease caused by complex interactions of more than two types of begomoviruses. Similarly, betasatellite genomic components were investigated for the prevailing complexity. Phylogenetically, betasatellite components in Indian sub-continent were divided into six clusters represented by betasatellite molecules associated to particular host plants or to a particular DNA-A associated diseases pattern i.e. begomovirus, host and symptoms. Among all, PaLCuB infecting papaya and ChLCB/ToLCB infecting papaya crops were closely associated with each other

as compared to other clusters. Recombination study on betasatellite components revealed much more intricate interactions giving rise to genetic variation in betasatellite genomes i.e. 11 recombination events occurring among 50 betasatellite sequences. All presented experimental observations provide an evidence of molecular complexity prevailing in DNA-A of PLCD. Therefore, PLCD is definitely a disease complex consisting of cross-interaction between DNA-A and betasatellite components of various leaf curl causing begomoviruses.

Leaf samples were collected from various symptomatic papaya plants from Bareily, Sultanpur, Lucknow and Malihabad regions of Uttar Pradesh in India. After PCR based diagnostics, cloning and sequencing of begomovirus components, DNA-A and betasatellite components causing leaf curl disease symptoms in papaya were isolated. Infectious clones were prepared by arranging together genomic dimers of DNA-A and betasatellite components and introducing them in a plant binary vector, pCAMBIA1301. DNA-A infectious clone was named as pB1301-5.4 and betasatellite was named p1301-2.6, according to the size of dimers i.e. 5.4 kb and 2.6 kb. These infectious clones were found to cause crumpled and mosaic symptoms in *Nicotiana benthamiana* plants, when co-infiltrated together in 1:1 molar mixture. The infectious clones were found to be functional and cause typical mild leaf curl symptom in tobacco plants. Severe leaf curl symptoms developed in papaya plants when above begomoviral infectious particles were introduced via viruliferous whiteflies previously fed upon infected tobacco plants. Therefore, Koch's postulate was proved and the above two begomovirus components are considered to be a part of the PLCD complex in India.

Through BLAST and SDT analysis, the DNA-A and betasatellite components were later identified to be a new DNA-A isolate named *Tomato leaf curl virus [India:Sultanpur:Papaya:2016]* [MH105055] and *Tomato leaf curl betasatelllite-Naj2 [India:New Delhi:Papaya:2009]* [HM143911.1]. These were the begomoviral components identified to be the causative agents of PLCD in samples collected from Sultanpur region in central Uttar Pradesh.

Till now, the RNAi based strategies focused more on specific and single gene siRNA/artificial miRNA based disease control. Few reported that overlapping regions from DNA-A genome were successful in eliciting resistance (Ye et al. 2014). Therefore, this strategy based upon overlapping regions fulfills our criterion for a generic strategy against PLCD complex components. siRNA producing potential of DNA-A was evaluated. The Upper half region (UH) was chosen for siRNA based strategy as it contains AC1/AC4, IR and AV1/AV2/AC5 encoding overlapping ORFs. These regions were found to be functionally important for replication, coat protein synthesis, movement and host RNAi suppression in host plants. Therefore, UH region was introduced into a plant binary vector pBI121 as a siRNA expression cassette named pBS-Sul. Theoretically, similar types of cassettes are capable of producing a decoy of targeted siRNA specifically induced under effect of CaMV35s promoter, resulting in constitutive expression of protective ds-RNA molecules ranging from 21 – 24-nts size (Blevins et al. 2011; Aregger et al. 2012; Pooggin 2017). RNAi experiment was designed to simulate field type conditions i.e. plants in which pBS-Sul was introduced either before or after infection. The plants were co-infected with p1301-5.4/2.6 (1:1) equimolar mixture to produce maximum possible disease severity index. As compared to Virus only plants, the siRNA/Vir plants showed better resistance from 14^{th} until 36^{th} day, as the pBS-Sul expressed siRNAs before introduction of viral genomic components. The Vir/siRNA plants did not show any significant resistance for above period due to previous infection of viral components, which might have suppressed their production. After 50^{th} day, both siRNA treated plants show similar levels of CP gene expression concluding that viral DNA reached their replication plateau, however, betasatellite molecules show unprecedented rise in expression from 36^{th} days onwards. This signifies that the siRNA construct might have controlled the late expression of DNA-A encoded proteins and its efficacy against betasatellite molecules requires further investigations. Therefore, the siRNA strategy proposed in this thesis has potential to become a basis for further siRNA based research to get a better insight into the UH mediated resistance through expression of siRNA decoys.

Bibliography

BIBLIOGRAPHY

Abdullah E, Idris A, Saparon A (2017). Establishing RNA virus resistance in plants by harnessing CRISPR immune system. ARPN Journal of Engineering and Applied Sciences. 12:279-281. doi: 10.1111/ijlh.12426

Akhtar S, Briddon RW, Mansoor S (2011). Reactions of Nicotiana species to inoculation with monopartite and bipartite begomoviruses. Virology Journal. 8:279-281. doi: 10. 1186/1743-422X-8-475

Akmal M, Baig MS, Khan JA (2017). Suppression of cotton leaf curl disease symptoms in Gossypium hirsutum through over expression of host-encoded miRNAs. Journal of Biotechnology. 263:279-281. doi: 10.1016/j.jbiotec.2017.10.003

Ameres SL, Martinez J, Schroeder R (2007). Molecular Basis for Target RNA Recognition and Cleavage by Human RISC. Cell. 130:279-281. doi: 10.1016/j.cell.2007.04.037

Amin I, Hussain K, Akbergenov R, et al (2011). Suppressors of RNA Silencing Encoded by the Components of the Cotton Leaf Curl Begomovirus-BetaSatellite Complex. Molecular Plant-Microbe Interactions. 24:279-281. doi: 10.1094/MPMI-01-11-0001

Aregger M, Borah BK, Seguin J, et al (2012). Primary and Secondary siRNAs in Geminivirus-induced Gene Silencing. PLoS Pathogens. 8:279-281. doi: 10.1371/journal.ppat.1002941

Aukerman MJ (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2-Like Target Genes. The Plant Cell Online. 15:279-281. doi: 10.1105/tpc.016238

Australian Government (2008). The Biology of Carica papaya L (papaya , papaw , paw paw). Australian Government.

Axtell MJ (2013). Classification and Comparison of Small RNAs from Plants. Annual Review of Plant Biology. 64:279-281. doi: 10.1146/annurev-arplant-050312-120043

Axtell MJ, Jan C, Rajagopalan R, Bartel DP (2006). A Two-Hit Trigger for siRNA Biogenesis in Plants. Cell. 127:279-281. doi: 10.1016/j.cell.2006.09.032

Beachy RN, Loesch-Fries S, Tumer NE (1990). Coat Protein-Mediated Resistance Against Virus Infection. Annual Review of Phytopathology. 28:279-281. doi: 10.1146/annurev.py.28.090190.002315

Berger MR, Sunter G (2013). Identification of sequences required for AL2-mediated activation of the tomato golden mosaic virus-yellow vein BR1 promoter. Journal of General Virology. 94:279-281. doi: 10.1099/vir.0.050161-0

Bhattacharyya D, Gnanasekaran P, Kumar RK, et al (2015). A geminivirus betasatellite damages the structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. Journal of Experimental Botany. 66:279-281. doi: 10.1093/jxb/erv299

Bisaro DM (2006). Silencing suppression by geminivirus proteins. Virology 344:158–168

Blevins T, Rajeswaran R, Aregger M, et al (2011). Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. Nucleic Acids Research. 39:279-281. doi: 10.1093/nar/gkr119

Bonfim K, Faria JC, Nogueira EOPL, et al (2007). RNAi-Mediated Resistance to Bean golden mosaic virus in Genetically Engineered Common Bean (Phaseolus vulgaris). Molecular Plant-Microbe Interactions. 20:279-281. doi: 10.1094/MPMI-20-6-0717

Boni MF, Posada D, Feldman MW (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics. 176:279-281. doi: 10.1534/genetics.106.068874

Borah BK, Dasgupta I (2012). Begomovirus research in India: A critical appraisal and the way ahead. Journal of Biosciences. 37:279-281. doi: 10.1007/s12038-012-9238-y

Borah BK, Zarreen F, Baruah G, Dasgupta I (2016). Insights into the control of geminiviral promoters. Virology. 495:279-281. doi: 10.1016/j.virol.2016.04.033

Borges F, Martienssen RA (2015). The expanding world of small RNAs in plants. Nature Reviews Molecular Cell Biology. 16:279-281. doi: 10.1038/nrm4085

Briddon RW, Mansoor S, Bedford ID, et al (2001). Identification of DNA components required for induction of cotton leaf curl disease. Virology. 285:279-281. doi: 10.1006/viro.2001.0949

Briddon RW, Pinner MS, Stanley J, Markham PG (1990). Geminivirus coat protein gene replacement alters insect specificity. Virology. 177:279-281. doi: 10.1016/0042-6822(90)90462-Z

Briddon RW, Stanley J (2006). Subviral agents associated with plant single-stranded DNA viruses. Virology. 344:279-281. doi: 10.1016/j.virol.2005.09.042

Brodersen P, Voinnet O (2006). The diversity of RNA silencing pathways in plants. Trends Genet. 22:268–280

Brown JK, Idris AM, Ostrow KM, et al (2005). Genetic and Phenotypic Variation of the *Pepper golden mosaic virus* Complex. Phytopathology. 95:279-281. doi: 10.1094/ PHYTO-95-1217

Brown JK, Zerbini FM, Navas-Castillo J, et al (2015). Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Archives of virology. 160:279-281. doi: 10.1007/s00705-015-2398-y

Bucher E, Lohuis D, van Poppel PMJA, et al (2006). Multiple virus resistance at a high frequency using a single transgene construct. Journal of General Virology. 87:279-281. doi: 10.1099/vir.0.82276-0

Burgyán J, Havelda Z (2011). Viral suppressors of RNA silencing. Trends in Plant Science. 16:279-281. doi: 10.1016/j.tplants.2011.02.010

Byrne DN, Bellows TS (1991). Whitefly Biology. Annual Review of Entomology. 36:279-281. doi: 10.1146/annurev.en.36.010191.002243

Carrillo-Tripp J, Lozoya-Gloria E, Rivera-Bustamante RF (2007). Symptom Remission and Specific Resistance of Pepper Plants After Infection by *Pepper golden mosaic virus*. Phytopathology. 97:279-281. doi: 10.1094/PHYTO-97-0051

Castillo-González C, Liu X, Huang C, et al (2015). Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense. eLife. 4:279-281. doi: 10.7554/eLife.06671

Chellappan P, Vanitharani R, Fauquet CM (2005). MicroRNA-binding viral protein interferes with Arabidopsis development. Proceedings of the National Academy of Sciences. 102:279-281. doi: 10.1073/pnas.0504439102

Cheng X, Wang X, Wu J, et al (2011). βC1 encoded by tomato yellow leaf curl China betasatellite forms multimeric complexes in vitro and in vivo. Virology. 409:279-281. doi: 10.1016/j.virol.2010.10.007

Chowda-Reddy R V., Achenjang F, Felton C, et al (2008). Role of a geminivirus AV2 protein putative protein kinase C motif on subcellular localization and pathogenicity. Virus Research. 135:279-281. doi: 10.1016/j.virusres.2008.02.014

Cong L, Ran FA, Cox D, et al (2013). Multiplex genome engineering using CRISPR/Cas systems. Science. 339:279-281. doi: 10.1126/science.1231143

Creasey KM, Zhai J, Borges F, et al (2014). MiRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. Nature. 508:279-281. doi: 10.1038/nature13069

Cuerda-Gil D, Slotkin RK (2016). Non-canonical RNA-directed DNA methylation. Nature Plants. 2:279-281. doi: 10.1038/nplants.2016.163

Cui X, Li G, Wang D, et al (2005). A Begomovirus DNA β -Encoded Protein Binds DNA , Functions as a Suppressor of RNA Silencing , and Targets the Cell Nucleus. Journal of Virology. 79:279-281. doi: 10.1128/JVI.79.16.10764

De Bruyn A, Harimalala M, Zinga I, et al (2016). Divergent evolutionary and epidemiological dynamics of cassava mosaic geminiviruses in Madagascar. BMC Evolutionary Biology. 16:279-281. doi: 10.1186/s12862-016-0749-2

De la Cruz Medina J, Vela Gutiérrez G, García H. (2003). PAPAYA Post-harvest Operations. INPho, Post-harvest Compendium.

Deng D, McGrath PF, Robinson DJ, & Harrison BD (1994). Detection and Differentiation of whitefly- transmitted Gemini- viruses in plants and vector insects by the Polymerase chain reaction with degenerate primers. Annals of Applied Biology. 125:279-281.

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin 19:11-15.

Du J, Rietman H, Vleeshouwers VGAA (2014). Agroinfiltration and PVX Agroinfection in Potato and Nicotiana benthamiana. Journal of Visualized Experiments. 83. e50971. doi: 10.3791/50971

Eini O, Behjatnia SAA (2016). The minimal sequence essential for replication and movement of Cotton leaf curl Multan betasatellite DNA by a helper virus in plant cells. Virus Genes. 52:279-281. doi: 10.1007/s11262-016-1354-6

Eini O, Behjatnia SAA, Dogra S, et al (2009). Identification of sequence elements regulating promoter activity and replication of a monopartite begomovirus-associated DNA β satellite. Journal of General Virology. 90:279-281. doi: 10.1099/vir.0.002980-0

Edgar RC (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 32:279-281. doi: 10.1093/nar/gkh340

Fang X, Qi Y (2016). RNAi in Plants: An Argonaute-Centered View. The Plant Cell. 28:279-281. doi: 10.1105/tpc.15.00920

FAOSTAT. FAO, Rome. 2013.; http://faostat.fao.org

Felsenstein J (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. Journal of Molecular Evolution. 17:279-281. doi: 10.1007/BF01734359

Ferrari R, Pellegrini M, Horwitz GA, et al (2008). Epigenetic reprogramming by adenovirus e1a. Science. 321:279-281. doi: 10.1126/science.1155546

Fondong VN (2013). Geminivirus protein structure and function. Molecular Plant Pathology. 14:279-281. doi: 10.1111/mpp.12032

Fondong VN (2017). The Search for Resistance to Cassava Mosaic Geminiviruses: How Much We Have Accomplished, and What Lies Ahead. Frontiers in Plant Science. 8:279-281. doi: 10.3389/fpls.2017.00408

Fondong VN, Pita JS, Rey MEC, et al (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. Journal of General Virology. 81:279-281. doi: 10.1099/0022-1317-81-1-287

Fondong VN, Reddy RVC, Lu C, et al (2007). The Consensus N -Myristoylation Motif of a Geminivirus AC4 Protein Is Required for Membrane Binding and Pathogenicity. Molecular Plant-Microbe Interactions. 20:279-281. doi: 10.1094/MPMI-20-4-0380

Freedman DA (1981). Bootstrapping Regression Models. The Annals of Statistics. 9:279-281. doi: 10.1214/aos/1176345638

Fuentes A, Carlos N, Ruiz Y, et al (2016). Field Trial and Molecular Characterization of RNAi-Transgenic Tomato Plants That Exhibit Resistance to Tomato Yellow Leaf Curl Geminivirus. Molecular Plant-Microbe Interactions. 29:279-281. doi: 10.1094/MPMI-08-15-0181-R

Fuentes A, Ramos PL, Fiallo E, et al (2006). Intron – hairpin RNA derived from replication associated protein C1 gene confers immunity to Tomato Yellow Leaf Curl Virus infection in transgenic tomato plants. 15:279-281. doi: 10.1007/s11248-005-5238-0

Galvez LC, Banerjee J, Pinar H, Mitra A (2014). Engineered plant virus resistance. Plant Science. 228:279-281. doi: 10.1016/j.plantsci.2014.07.006

Gamerman D, Lopes HF (2006). Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference. CRC Press. doi: 10.1002/15213773(20010316)40:6<9823::AID - ANIE9823>3.3.CO;2-C

Gebre M, Nomburg JL, Gewurz BE (2018). Crispr–Cas9 genetic analysis of virus–Host interactions. Viruses. 10:279-281. doi: 10.3390/v10020055

Ghanim M (2014). A review of the mechanisms and components that determine the transmission efficiency of Tomato yellow leaf curl virus (Geminiviridae; Begomovirus) by its whitefly vector. Virus Research. 186:279-281. doi: 10.1016/j.virusres.2014.01.022

Gibbs MJ, Armstrong JS, Gibbs AJ (2000). Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics. 16:279-281. doi: 10.1093/bioinformatics/16.7.573

Gonsalves D (1998). Control Of Papaya Ringspot Virus In Papaya: A Case Study. Annual Review of Phytopathology. 36:279-281. doi: 10.1146/annurev.phyto.36.1.415

Gonsalves D (2006). Transgenic Papaya: Development, Release, Impact and Challenges. Adv. Virus Res. 67:317–354

Gopal P, Pravin Kumar P, Sinilal B, et al (2007). Differential roles of C4 and βC1 in mediating suppression of post-transcriptional gene silencing: Evidence for transactivation by the C2 of Bhendi yellow vein mosaic virus, a monopartite begomovirus. Virus Research. 123:279-281. doi: 10.1016/j.virusres.2006.07.014

Guan C, Zhou X (2006). Phloem specific promoter from a satellite associated with a DNA virus. Virus Research. 115:279-281. doi: 10.1016/j.virusres.2005.08.002

Guindon S, Dufayard JF, Lefort V, et al (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Systematic Biology. 59:279-281. doi: 10.1093/sysbio/syq010

Guindon S, Lethiec F, Duroux P, Gascuel O (2005). PHYML Online - A web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Research. 33:279-281. doi: 10.1093/nar/gki352

Guo T, Guo Q, Cui XY, et al (2015). Comparison of transmission of Papaya leaf curl China virus among four cryptic species of the whitefly Bemisia tabaci complex. Scientific Reports. 5:279-281. doi: 10.1038/srep15432

Guo T, Zhao J, Pan LL, Geng L, Lei T, Wang XW, et al. (2018). The level of midgut penetration of two begomoviruses affects their acquisition and transmission by two species of Bemisia tabaci. Virology. 515:66–73. doi:10.1016/j.virol.2017.12.004.

Gutierrez C, Ramirez-Parra E, Mar Castellano M, et al (2004). Geminivirus DNA replication and cell cycle interactions. Veterinary microbiology. 98:279-281.

Hall BG (2013). Building phylogenetic trees from molecular data with MEGA. Molecular Biology and Evolution. 30:279-281. doi: 10.1093/molbev/mst012

Hasegawa M, Kishino H, Yano T (1985). Dating the human-ape split by a molecular clock of mitochondrial DNA. Evolution. 22:279-281. doi: 10.1007/bf02101694

Hashmi JA, Zafar Y, Arshad M, et al (2011). Engineering cotton (*Gossypium hirsutum* L.) for resistance to cotton leaf curl disease using viral truncated AC1 DNA sequences. Virus Genes. 42:279-281. doi: 10.1007/s11262-011-0569-9

Hipp K, Rau P, Schäfer B, et al (2016). Translation, modification and cellular distribution of two AC4 variants of African cassava mosaic virus in yeast and their pathogenic potential in plants. Virology. 498:279-281. doi: 10.1016/j.virol.2016.07.011

Hirai S, Kodama H (2008). RNAi Vectors for Manipulation of Gene Expression in Higher Plants. The Open Plant Science Journal. 2:279-281. doi: 10.2174/1874294700802010021

Horticulture Statistics Division, Department of Agriculture, Cooperation and Farmers Welfare, Ministry of Agriculture. http://agricoop.nic.in/statistics/horticulture

Ibrahim F, Rymarquis LA, Kim E-J, et al (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. Proceedings of the National Academy of Sciences. 107:279-281. doi: 10.1073/pnas.0912632107

Jangid A, Chandel V, Kumar Y, Rishi N (2017). RNA silencing suppressor activity of Velvet bean severe mosaic begomovirus DNA-A encoded proteins. Archives of Phytopathology and Plant Protection. 50:279-281. doi: 10.1080/03235408.2017.1316028

Jeske H, Lütgemeier M, Preiß W (2001). DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. EMBO Journal. 20:279-281. doi: 10.1093/emboj/20.21.6158

Jia X, Ding N, Fan W, et al (2015). Functional plasticity of miR165/166 in plant development revealed by small tandem target mimic. Plant Science. 233:279-281. doi: 10.1016/j.plantsci.2014.12.020

Jyothsna P, Haq QMI, Singh P, et al (2013). Infection of tomato leaf curl New Delhi virus (ToLCNDV), a bipartite begomovirus with betasatellites, results in enhanced level of helper virus components and antagonistic interaction between DNA B and betasatellites. Applied Microbiology and Biotechnology. 97:279-281. doi: 10.1007/s00253-012-4685-9

Kasuya E (2001). Mann-Whitney U test when variances are unequal. Anim. Behav. 61:1247–1249

Keller C, Kulasegaran-Shylini R, Shimada Y, et al (2013). Noncoding RNAs prevent spreading of a repressive histone mark. Nature Structural and Molecular Biology. 20:279-281. doi: 10.1038/nsmb.2619

King AMQ, Adams MJ, Carsten EB, Lefkowitz EJ (2012). Virus Taxonomy: Classification and Nomenclature of Viruses. Ninth Report of the International Committee on Taxonomy of Viruses.

Klinkenberg FA, Ellwood S, Stanley J (1989). Fate of African Cassava Mosaic Virus Coat Protein Deletion Mutants after Agroinoculation. Journal of General Virology. 70:279-281. doi: 10.1099/0022-1317-70-7-1837

Knauer S, Holt AL, Rubio-Somoza I, et al (2013). A Protodermal miR394 Signal Defines a Region of Stem Cell Competence in the Arabidopsis Shoot Meristem. Developmental Cell. 24:279-281. doi: 10.1016/j.devcel.2012.12.009

Knierim D, Maiss E (2007). Application of Phi29 DNA polymerase in identification and full-length clone inoculation of tomato yellow leaf curl Thailand virus and tobacco leaf curl Thailand virus. Archives of Virology. 152:279-281. doi: 10.1007/s00705-006-0914-9

Kon T, Sharma P, Ikegami M (2007). Suppressor of RNA silencing encoded by the monopartite tomato leaf curl Java begomovirus. Archives of Virology. 152:279-281. doi: 10.1007/s00705-007-0957-6

Krubphachaya P, Juricek M, Kertbundit S (2007). Induction of RNA-mediated Resistance to Papaya Ringspot Virus Type W. BMB Reports. 40:279-281. doi: 10.5483/BMBRep.2007.40.3.404

Kumar V, Mishra SK, Rahman J, et al (2015). Mungbean yellow mosaic Indian virus encoded AC2 protein suppresses RNA silencing by inhibiting Arabidopsis RDR6 and AGO1 activities. Virology. 486:279-281. doi: 10.1016/j.virol.2015.08.015

Kumar V, Naqvi AR (2016). Tomato leaf curl New Delhi virus (ToLCNDV) encoded AC2 associates with host mirnas by directly interacting with AGO1. Journal of RNAi and Gene Silencing. 12:279-281.

Laerd Statistics (2013). Friedman Test in SPSS Statistics - How to run the procedure, understand the output using a relevant example | Laerd Statistics. In: laerd Stat. https://statistics.laerd.com/spss-tutorials/friedman-test-using-spss-statistics.php

Lai J, Chen H, Teng K, et al (2009). RKP, a RING finger E3 ligase induced by BSCTV C4 protein, affects geminivirus infection by regulation of the plant cell cycle. Plant Journal. 57:279-281. doi: 10.1111/j.1365-313X.2008.03737.x

Larget B, Simon DL (1999). Markov Chain Monte Carlo Algorithms for the Bayesian Analysis of Phylogenetic Trees. Mol Biol Evol. 16:279-281.

Lemey P, Lott M, Martin DP, Moulton V (2009). Identifying recombinants in human and primate immunodeficiency virus sequence alignments using quartet scanning. BMC Bioinformatics. 10:279-281. doi: 10.1186/1471-2105-10-126

Leuschner PJF, Ameres SL, Kueng S, Martinez J (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO Reports. 7:279-281. doi: 10.1038/sj.embor.7400637

Li JF, Zhang D, Sheen J (2014). Cas9-based genome editing in arabidopsis and tobacco, 1st edn. Elsevier Inc.

Lima ATM, Silva JCF, Silva FN, et al (2017). The diversification of begomovirus populations is predominantly driven by mutational dynamics. Virus Evolution. 3:279-281. doi: 10.1093/ve/vex005

Liu Q, Feng Y, Zhu Z (2009). Dicer-like (DCL) proteins in plants. Funct. Integr. Genomics 9:277–286

Lózsa R, Csorba T, Lakatos L, Burgyán J (2008). Inhibition of 3′ modification of small RNAs in virus-infected plants require spatial and temporal co-expression of small RNAs and viral silencing-suppressor proteins. Nucleic Acids Research. 36:279-281. doi: 10.1093/nar/gkn365

Lu Y Du, Gan QH, Chi XY, Qin S (2008). Roles of microRNA in plant defense and virus offense interaction. Plant Cell Reports. 27:279-281. doi: 10.1007/s00299-008-0584-z

Luo Y, Chen Q, Luan J, et al (2017). Towards an understanding of the molecular basis of effective RNAi against a global insect pest, the whitefly Bemisia tabaci. Insect Biochemistry and Molecular Biology. 88:279-281. doi: 10.1016/j.ibmb.2017.07.005

Malik AH, Briddon RW, Mansoor S (2011). Infectious clones of Tomato leaf curl Palampur virus with a defective DNA B and their pseudo-recombination with Tomato leaf curl New Delhi virus. Virology Journal. 8:279-281. doi: 10.1186/1743-422X-8-173

Margis R, Fusaro AF, Smith NA, et al (2006). The evolution and diversification of Dicers in plants. FEBS Letters. 580:279-281. doi: 10.1016/j.febslet.2006.03.072

Martin D, Rybicki E (2000). RDP: detection of recombination amongst aligned sequences. Bioinformatics. 16:279-281. doi: 10.1093/bioinformatics/16.6.562

Martin DP, Lefeuvre P, Varsani A, et al (2011). Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. PLoS Pathogens. 7:279-281. doi: 10.1371/journal.ppat.1002203

Martin DP, Murrell B, Golden M, et al (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution. 1:279-281. doi: 10.1093/ve/vev003

Martin DP, Posada D, Crandall KA, Williamson C (2005). A Modified Bootscan Algorithm for Automated Identification of Recombinant Sequences and Recombination Breakpoints. AIDS Research and Human Retroviruses. 21:279-281. doi: 10.1089/aid.2005.21.98

McIntyre GJ, Fanning GC (2006). Design and cloning strategies for constructing shRNA expression vectors. BMC Biotechnology. 6:279-281. doi: 10.1186/1472-6750-6-1

Medina-Hernández D, Rivera-Bustamante RF, Tenllado F, Holguín-Peña RJ (2013). Effects and effectiveness of two RNAi constructs for resistance to pepper golden mosaic virus in Nicotiana benthamiana plants. Viruses. 5:279-281. doi: 10.3390/v5122931

Mills-Lujan K, Deom CM (2010). Geminivirus C4 protein alters Arabidopsis development. Protoplasma. 239:279-281. doi: 10.1007/s00709-009-0086-z

Milne I, Lindner D, Bayer M, et al (2009). TOPALi v2: A rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics. 25:279-281. doi: 10.1093/bioinformatics/btn575

Mishra M, Chandra R, Saxena S (2007) Papaya. In: Kole C (Ed.) Genome Mapping and Molecular Breeding in Plants-Fruits and Nuts. USA. Springer. p.p. 343-350.

Moriones E, Praveen, S, and Chakraborty S (2017). Tomato leaf curl new delhi virus: An emerging virus complex threatening vegetable and fiber crops. Viruses 9:264. doi:10.3390/v9100264.

Mubin M, Hussain M, Briddon RW, Mansoor S (2011). Selection of target sequences as well as sequence identity determine the outcome of RNAi approach for resistance against cotton leaf curl geminivirus complex. Virology Journal. 8:279-281. doi: 10.1186/1743-422X-8-122

Mubin M, Mansoor S, Hussain M, Zafar Y (2007). Silencing of the AV2 gene by antisense RNA protects transgenic plants against a bipartite begomovirus. Virology Journal. 4:279-281. doi: 10.1186/1743-422X-4-10

Muhire BM, Varsani A, Martin DP (2014). SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. PLoS ONE. 9:279-281. doi: 10.1371/journal.pone.0108277

Mullis, K. B., & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. Methods in Enzymology, 155:335–350. https://doi.org/10.1016/0076-6879(87)55023-6

Murashige T, Skoog F (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiologia Plantarum. 15:279-281. doi: 10.1111/j.1399-3054.1962.tb08 052.x

Nadeem A, Mehmood T, Tahir M (1997). First report of papaya leaf curl disease in Pakistan. Plant Disease. 81:279-281. doi: 10.1094/PDIS.1997.81.11.1333B

Napoli C (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. The Plant Cell Online. 2:279-281. doi: 10.1105/tpc.2.4.279

Nariani TK (1956). Leaf curl of papaya. Indian Phytopathology. 9: 151-157.

Nawaz-ul-Rehman MS, Fauquet CM (2009). Evolution of geminiviruses and their satellites. FEBS Letters. 583:279-281. doi: 10.1016/j.febslet.2009.05.045

Noris E, Miozzi L (2015). Chapter 6: Real-Time PCR protocols for the quantification of the begomovirus *Tomato Yellow Leaf Curl Sardinia Virus* in tomato plants and in its insect vector. Methods in Molecular Biology. 1236:279-281. doi: 10.1007/978-1-4939-1743-3_6

Padidam M, Beachy RN, Fauquet CM (1996). The role of AV2 (’precoat’) and coat protein in viral replication and movement in tomato leaf curl geminivirus. Virology. 224:279-281. doi: 10.1006/viro.1996.0546

Padidam M, Sawyer S, Fauquet CM (1999). Possible Emergence of New Geminiviruses by Frequent Recombination. Virology. 265:279-281. doi: 10.1006/viro.1999.0056

Park J, Hwang HS, Buckley KJ, et al (2010). C4 protein of Beet severe curly top virus is a pathomorphogenetic factor in Arabidopsis. Plant Cell Reports. 29:279-281. doi: 10.1007/s00299-010-0923-8

Patil BL, Dasgupta I (2006). Defective interfering DNAs of plant viruses. CRC. Crit. Rev. Plant Sci. 25:47–64

Patil BL, Fauquet CM (2010). Differential interaction between cassava mosaic geminiviruses and geminivirus satellites. Journal of General Virology. 91:279-281. doi: 10.1099/vir.0.019513-0

Pilartz M, Jeske H (2003). Mapping of Abutilon Mosaic Geminivirus Minichromosomes. Journal of Virology. 77:279-281. doi: 10.1128/JVI.77.20.10808-10818.2003

Pita JS, Fondong VN, Sangaré A, et al (2001). Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. Journal of General Virology. 82:279-281. doi: 10.1099/0022-1317-82-3-655

Pooggin MM, Hohn T (2004). Fighting geminiviruses by RNAi and vice versa. Plant Molecular Biology. 55:279-281. doi: 10.1007/s11103-004-0116-3

Pooggin M, Shivaprasad P V, Veluthambi K, Hohn T (2003). RNAi targeting of DNA virus in plants. Nature Biotechnology. 21:279-281. doi: 10.1038/nbt0203-131b

Pooggin MM (2013). How can plant DNA viruses evade siRNA-directed DNA methylation and silencing? International Journal of Molecular Sciences. 14:279-281. doi: 10.3390/ijms140815233

Pooggin MM (2017). RNAi-mediated resistance to viruses: a critical assessment of methodologies. Current Opinion in Virology. 26:279-281. doi: 10.1016/j.coviro.2017.07.010

Posada D, and Crandall, KA (2001). Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. Proceedings of the.National Academy of Science. 98:13757–13762. doi:10.1073/pnas.241370698.

Priyadarshini CGP, Ambika M V., Tippeswamy R, Savithri HS (2011). Functional characterization of coat protein and V2 involved in cell to cell movement of cotton leaf curl Kokhran Virus-Dabawali. PLoS ONE. 6:279-281. doi: 10.1371/journal.pone.0026929

Qazi J, Amin I, Mansoor S, et al (2007). Contribution of the satellite encoded gene βC1 to cotton leaf curl disease symptoms. Virus Research. 128:279-281. doi: 10.1016/j.virusres.2007.04.002

Raja P, Wolf JN, Bisaro DM (2010). RNA silencing directed against geminiviruses: Post-transcriptional and epigenetic components. Biochim. Biophys. Acta - Gene Regul. Mech. 1799:337–351

Rand TA, Petersen S, Du F, Wang X (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell. 123:279-281. doi: 10.1016/j.cell.2005.10.020

Richardson, P. M., Detter, C., Schweitzer, B., & Predki, P. F. (2003). Practical applications of rolling circle amplification of DNA templates. Genetic Engineering, 25:51–63. https://doi.org/10.1007/978-1-4615-0073-5_3

Rochester DE, Kositratana W, Beachy RN (1990). Systemic movement and symptom production following agroinoculation with a single DNA of tomato yellow leaf curl geminivirus (Thailand). Virology. 178:279-281. doi: 10.1016/0042-6822(90)90349-V

Rodríguez-Negrete E, Lozano-Durán R, Piedra-Aguilera A, et al (2013). Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytologist. 199:279-281. doi: 10.1111/nph.12286

Rodriguez-Negrete EA, Carrillo-Tripp J, Rivera-Bustamante RF (2009). RNA Silencing against Geminivirus: Complementary Action of Posttranscriptional Gene Silencing and Transcriptional Gene Silencing in Host Recovery. Journal of Virology. 83:279-281. doi: 10.1128/JVI.01474-08

Rogans SJ, Allie F, Tirant JE, Rey MEC (2016). Small RNA and methylation responses in susceptible and tolerant landraces of cassava infected with South African cassava mosaic virus. Virus Research. 225:279-281. doi: 10.1016/j.virusres.2016.08.011

Rojas MR, Jiang H, Salati R, et al (2001). Functional analysis of proteins involved in movement of the monopartite begomovirus, Tomato yellow leaf curl virus. Virology. 291:279-281. doi: 10.1006/viro.2001.1194

Ronquist F, Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:279-281. doi: 10.1093/bioinformatics/btg180

Ronquist F, Teslenko M, Van Der Mark P, et al (2012). Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Systematic Biology. 61:279-281. doi: 10.1093/sysbio/sys029

Rosen R, Kanakala S, Kliot A, et al (2015). Persistent, circulative transmission of begomoviruses by whitefly vectors. Current Opinion in Virology. 15:279-281. doi: 10.1016/j.coviro.2015.06.008

Roshan P, Kulshreshtha A, Kumar S, et al (2018). AV2 protein of tomato leaf curl Palampur virus promotes systemic necrosis in Nicotiana benthamiana and interacts with host Catalase2. Scientific Reports. 8:279-281. doi: 10.1038/s41598-018-19292-3

Ruxton GD, Beauchamp G (2008). Some suggestions about appropriate use of the Kruskal-Wallis test. Animal Behaviour. 76:279-281. doi: 10.1016/j.anbehav.2008.04.011

Saeed M (2010). Tomato leaf curl virus and Cotton leaf curl Multan betasatellite can cause mild transient symptoms in cotton. Australasian Plant Disease Notes. 5:279-281. doi: 10.1071/DN10021

Saeed M, Zafar Y, Randles JW, Rezaian MA (2007). A monopartite begomovirus-associated DNA β satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection. Journal of General Virology. 88:279-281. doi: 10.1099/vir.0.83049-0

Sahu PP, Sharma N, Puranik S, Prasad M (2014). Post-transcriptional and Epigenetic Arms of RNA Silencing: A Defense Machinery of Naturally Tolerant Tomato Plant Against Tomato Leaf Curl New Delhi Virus. Plant Molecular Biology Reporter. 32:279-281. doi: 10.1007/s11105-014-0708-2

Sakaguchi J, Watanabe Y (2012). miR165/166 and the development of land plants. Development, Growth & Differentiation. 54:279-281. doi: 10.1111/j.1440-169X.2011.01318.x

Salminen MO, Carr JK, Burke DS, McCutchan FE (1995). Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS research and human retroviruses. 11:279-281. doi: 10.1089/aid.1995.11.1423

Saxena S, Hallan V, Singh BP and Sane PV (1998a). Leaf curl disease of Carica papaya from India may be caused by a bipartite geminivirus. Plant Disease.82(1): 126.

Saxena S, Hallan V, Singh BP and Sane PV (1998b). Evidence from nucleic acid hybridization test for a geminivirus infection causing leaf curl disease of papaya in India. Indian Journal Experimental Biology. 36: 229-232.

Saxena S, Hallan V, Singh BP and Sane PV (1998c). Nucleotide sequence and intergeminiviral homologies of the DNA-A of papaya leaf curl geminivirus from India. Biochem. Mol. Biol. Int.,45(1): 101-113.[now renamed as IUBMB Life, International Union of Biochemistry and Molecular Biology, ISSN; 1521-6543]

Saxena S, Kesharwani RK, Singh V, Singh S (2013). Designing of putative siRNA against geminiviral suppressors of RNAi to develop geminivirus-resistant papaya crop. International Journal of Bioinformatics Research and Applications. 9:279-281. doi: 10.1504/IJBRA.2013.050655

Saxena S, Singh N, Ranade SA, Babu SG (2011). Strategy for a generic resistance to geminiviruses infecting tomato and papaya through in silico siRNA search. Virus Genes. 43:279-281. doi: 10.1007/s11262-011-0649-x

Scheer H, Zuber H, De Almeida C, Gagliardi D (2016). Uridylation Earmarks mRNAs for Degradation… and More. Trends Genet. 32:607–619

Settlage SB, See RG, Hanley-Bowdoin L (2005). Geminivirus C3 protein: replication enhancement and protein interactions. Journal of virology. 79:279-281. doi: 10.1128/JVI.79.15.9885-9895.2005

Shivaprasad P V, Akbergenov R, Trinks D, et al (2005). Promoters , Transcripts , and Regulatory Proteins of Mungbean Yellow Mosaic Geminivirus Promoters , Transcripts , and Regulatory Proteins of Mungbean Yellow Mosaic Geminivirus †. 79:279-281. doi: 10.1128/JVI.79.13.8149

Shukla AK, Upadhyay SK, Mishra M, et al (2016). Expression of an insecticidal fern protein in cotton protects against whitefly. Nature Biotechnology. 34:279-281. doi: 10.1038/nbt.3665

Shweta, Akhter Y, Khan JA (2018). Genome wide identification of cotton (Gossypium hirsutum)-encoded microRNA targets against Cotton leaf curl Burewala virus. Gene. 638:279-281. doi: 10.1016/j.gene.2017.09.061

Singh AK, Kushwaha N, Chakraborty S (2016). Synergistic interaction among begomoviruses leads to the suppression of host defense-related gene expression and breakdown of resistance in chilli. Applied Microbiology and Biotechnology. 100:279-281. doi: 10.1007/s00253-015-7279-5

Singh-Pant P, Pant P, Mukherjee SK, Mazumdar-Leighton S (2012). Spatial and temporal diversity of begomoviral complexes in papayas with leaf curl disease. Archives of Virology. 157:279-281. doi: 10.1007/s00705-012-1287-x

Smith JM (1992). Analyzing the mosaic structure of genes. Journal of Molecular Evolution. 34:279-281. doi: 10.1007/BF00182389

Srivastava A, Jaidi M, Kumar S, et al (2014). Association of Papaya leaf curl virus with the leaf curl disease of grain amaranth (Amaranthus cruentus L.) in India. Phytoparasitica. 43:279-281. doi: 10.1007/s12600-014-0424-3

Stanley J (1995). Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs withinthe conserved nonanucleotide motif during the initiation of rolling circle DNA replication. Virology. 206:279-281. doi: 10.1016/S0042-6822(95)80093-X

Stanley J (2004). Subviral DNAs associated with geminivirus disease complexes. In: Veterinary Microbiology. pp 121–129

Stanley J, Boulton MI, Davies JW (2001). *Geminiviridae*. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK

Stanley J, Latham JR (1992). A symptom variant of beet curly top geminivirus produced by mutation of open reading frame C4. Virology. 190:506-509. Doi: 10.1016/0042-6822(92)91243-N

Stenger DC, Revington GN, Stevenson MC, Bisaro DM (1991). Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. Proceedings of the National Academy of Sciences. 88:279-281. doi: 10.1073/pnas.88.18.8029

Sunter G, Bisaro DM (1997). Regulation of a geminivirus coat protein promoter by AL2 protein (TrAP): Evidence for activation and derepression mechanisms. Virology. 232:279-281. doi: 10.1006/viro.1997.8549

Sunter G, Sunter JL, Bisaro DM (2001). Plants expressing tomato golden mosaic virus AL2 or beet curly top virus L2 transgenes show enhanced susceptibility to infection by DNA and RNA viruses. Virology. 285:279-281. doi: 10.1006/viro.2001.0950

Tamura K, Stecher G, Peterson D, et al (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution. 30:279-281. doi: 10.1093/molbev/mst197

Thakur N, Upadhyay SK, Verma PC, et al (2014). Enhanced whitefly resistance in transgenic tobacco plants expressing double stranded RNA of v-ATPase a gene. PLoS ONE. 9:279-281. doi: 10.1371/journal.pone.0087235

Thomas KM and Krishnaswamy CS (1939). Leaf crinkle- A transmissible disease of papaya. Current Science. 8: 316.

Thompson JD, Higgins DG, Gibson TJ (1994). ClustalW: improving the sensitivity of progressive multiple sequence aligment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Research Acids Res. 22:279-281. doi: 10.1093/nar/22.22.4673

Tiwari N, Sharma PK, Malathi VG (2013). Functional characterization of βc1 gene of Cotton leaf curl Multan betasatellite. Virus Genes. 46:279-281. doi: 10.1007/s11262-012-0828-4

Trinks D, Rajeswaran R, Shivaprasad P V., et al (2005). Suppression of RNA Silencing by a Geminivirus Nuclear Protein, AC2, Correlates with Transactivation of Host Genes. Journal of Virology. 79:279-281. doi: 10.1128/JVI.79.4.2517-2527.2005

Van der Krol AR, Lenting PE, Veenstra J, et al (1988). An anti-sense chalcone synthase gene in transgenic plants inhibits flower pigmentation. Nature. 333:279-281. doi: 10.1038/333866a0

van der Walt E, Palmer KE, Martin DP, Rybicki EP (2008). Viable chimaeric viruses confirm the biological importance of sequence specific maize streak virus movement protein and coat protein interactions. Virology Journal. 5:279-281. doi: 10.1186/1743-422X-5-61

van der Walt E, Rybicki EP, Varsani A, et al (2009). Rapid host adaptation by extensive recombination. Journal of General Virology. 90:279-281. doi: 10.1099/vir.0.007724-0

van Wezel R, Dong X, Blake P, et al (2002). Differential roles of geminivirus Rep and AC4 (C4) in the induction of necrosis in Nicotiana benthamiana. Molecular Plant Pathology. 3:279-281. doi: 10.1046/j.1364-3703.2002.00141.x

Vanitharani R, Chellappan P, Fauquet CM (2003). Short interfering RNA-mediated interference of gene expression and viral DNA accumulation in cultured plant cells. Proceedings of the National Academy of Sciences. 100:279-281. doi: 10.1073/pnas.1733874100

Vanitharani R, Chellappan P, Pita JS, Fauquet CM (2004). Differential Roles of AC2 and AC4 of Cassava Geminiviruses in Mediating Synergism and Suppression of Posttranscriptional Gene Silencing. Journal of Virology. 78:279-281. doi: 10.1128/JVI.78.17.9487-9498.2004

Varkonyi-Gasic E, Wu R, Wood M, et al (2007). Protocol: A highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Methods. 3:279-281. doi: 10.1186/1746-4811-3-12

Varsani A, Navas-Castillo J, Moriones E, et al (2014). Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. Archives of Virology. 159:279-281. doi: 10.1007/s00705-014-2050-2

Varsani A, Roumagnac P, Fuchs M, et al (2017). Capulavirus and Grablovirus: two new genera in the family Geminiviridae. Archives of Virology. 162:279-281. doi: 10.1007/s00705-017-3268-6

Varsha Wesley S, Helliwell CA, Smith NA, et al (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. Plant Journal. 27:279-281. doi: 10.1046/j.1365-313X.2001.01105.x

Velten J, Cakir C, Youn E, et al (2012). Transgene silencing and transgene-derived siRNA production in tobacco plants homozygous for an introduced AtMYB90 construct. PLoS ONE. 7:279-281. doi: 10.1371/journal.pone.0030141

Verma S, Saxena S (2017). Dissecting Papaya Leaf Curl Disease (PLCD) Complex and Assessing its Potential for siRNA Based Targeting. International Journal of Plant and Environment. 3:279-281. doi: 10.18811/ijpen.v3i02.10433

Verma S, Saxena S (2018). In-silico Investigation of Betasatellite Complexity in Papaya Leaf Curl Disease Complex. International Journal of Plant & Environment. 4: (Unpublished).

Wang H, Wang H (2015). The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. Mol. Plant 8:677–688

Willmann MR, Endres MW, Cook RT, Gregory BD (2011). The Functions of RNA-Dependent RNA Polymerases in Arabidopsis. The Arabidopsis Book. 9:279-281. doi: 10.1199/tab.0146

Wyant PS (2011). The use of rolling circle amplification (RCA) for diagnosis and characterization of geminiviruses

Xie Z (2005). Expression of Arabidopsis MIRNA Genes. Plant Physiology. 138:279-281. doi: 10.1104/pp.105.062943

Yang CF, Chen KC, Cheng YH, et al (2014). Generation of marker-free transgenic plants concurrently resistant to a DNA geminivirus and a RNA tospovirus. Scientific Reports. 4:279-281. doi: 10.1038/srep05717

Yang JY, Iwasaki M, Machida C, et al (2008). βCl, the pathogenicity factor of TYLCCNV, interacts with AS1 to alter leaf development and suppress selective jasmonic acid responses. Genes and Development. 22:279-281. doi: 10.1101/gad.1682208

Yang X, Wang Y, Guo W, et al (2011a). Characterization of small interfering RNAs derived from the Geminivirus/Betasatellite complex using deep sequencing. PLoS ONE. 6:279-281. doi: 10.1371/journal.pone.0016928

Yang X, Xie Y, Raja P, et al (2011b). Suppression of methylation-mediated transcriptional gene silencing by βc1-sahh protein interaction during geminivirus-betasatellite infection. PLoS Pathogens. 7:279-281. doi: 10.1371/journal.ppat.1002329

Ye J, Qu J, Mao HZ, et al (2014). Engineering geminivirus resistance in Jatropha curcus. Biotechnology for Biofuels. 7:279-281. doi: 10.1186/s13068-014-0149-z

Younis A, Siddique MI, Kim CK, Lim KB (2014). RNA interference (RNAi) induced gene silencing: A promising approach of hi-tech plant breeding. International Journal of Biological Sciences. 10:279-281. doi: 10.7150/ijbs.10452

Zaidi SSEA, Mansoor S, Ali Z, et al (2016). Engineering Plants for Geminivirus Resistance with CRISPR/Cas9 System. Trends in Plant Science. 21:279-281. doi: 10.1016/j.tplants.2016.01.023

Zhang H, Ma X, Qian Y, Zhou X (2010). Molecular characterization and infectivity of Papaya leaf curl China virus infecting tomato in China. Journal of Zhejiang University SCIENCE B. 11:279-281. doi: 10.1631/jzus.B0900176

Zhang P, Vanderschuren H, Fütterer J, Gruissem W (2005). Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. Plant Biotechnology Journal. 3:279-281. doi: 10.1111/j.1467-7652.2005.001 32.x

Zhang T, Xu X, Huang C, et al (2016). A Novel DNA Motif Contributes to Selective Replication of a Geminivirus-Associated Betasatellite by a Helper Virus-Encoded Replication-Related Protein. Journal of Virology. 90:279-281. doi: 10.1128/JVI.02290-15

Zhang Z, Chen H, Huang X, et al (2011). BSCTV C2 Attenuates the Degradation of SAMDC1 to Suppress DNA Methylation-Mediated Gene Silencing in *Arabidopsis*. The Plant Cell. 23:279-281. doi: 10.1105/tpc.110.081695

Zhong X, Wang ZQ, Xiao R, et al (2017). iTRAQ analysis of the tobacco leaf proteome reveals that RNA-directed DNA methylation (RdDM) has important roles in defense against geminivirus-betasatellite infection. Journal of Proteomics. 152:279-281. doi: 10.1016/j.jprot.2016.10.015

Zhou X (2013). Advances in Understanding Begomovirus Satellites. Annual Review of Phytopathology. 51:279-281. doi: 10.1146/annurev-phyto-082712-102234

Zubair M, Zaidi S, Shakir S, et al (2017a). An Insight into Cotton Leaf Curl Multan Betasatellite, the Most Important Component of Cotton Leaf Curl Disease Complex. Viruses. 9:279-281. doi: 10.3390/v9100280

Zubair M, Zaidi SSEA, Shakir S, et al (2017b). Multiple begomoviruses found associated with cotton leaf curl disease in Pakistan in early 1990 are back in cultivated cotton. Scientific Reports. 7:279-281. doi: 10.1038/s41598-017-00727-2

Zvereva AS, Pooggin MM (2012). Silencing and innate immunity in plant defense against viral and non-viral pathogens. Viruses. 4:279-281. doi: 10.33

Annexure

ANNEXURE 1

Table A1 List of DNA-A sequences of Indian PLCD complex begomoviruses causing leaf curl disease in papaya and associated crops and weeds.

[AV1=Coat protein; AV2=variable function; AC1=replication initiation protein; AC2=transcription activator protein; AC3=replication enhancer protein; AC4=host RNA machinery suppressor; AC5, AC6 and AV3=uncharacterized open reading frames of begomoviruses of PLCD complex.]

Accession ID	Virus	Annotation	Host	Size	AV2	AV1	AC3	AC2	AC1	AC4	AC5	AC6	AV3
Y15934.1	**Papaya leaf curl virus**	PLCD-1	*Carica papaya*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457	301-618		
NC_004147.1	**Papaya leaf curl virus**	PLCD-2	*Carica papaya*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457	301-618		
HM143914.1	**Papaya leaf curl virus segment DNA A**	PLCD-3	*Nicotiana glutinosa*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
JN135233.1	**Papaya leaf curl virus**	PLCD-4	*Amaranthus cruentus L.*	2746	150-506	310-1083	1077-1481	1222-1626	1529-2614	2200-2457			
JQ954859.1	**Papaya leaf curl virus**	PLCD-5	*Aster alpinus L.*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KM525657.1	**Papaya leaf curl virus**	PLCD-6	*Croton bonplandianus L.*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KU376493.1	**Papaya leaf curl virus isolate CN2**	PLCD-7	*Solanum lycopersicum*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KX302713.1	**Papaya leaf curl virus isolate Wellington clone WB2&5**	PLCD-8	*Carica papaya*	2763	149-511	315-1085	1082-1270	1267-1623	1535-2674	2400-2621	729-980	1227-1487	
KY800906.1	**Papaya leaf curl virus isolate India/New Delhi/Papaya/2016**	PLCD-9	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1529-2611	2161-2460			
KY026597.1	**Papaya leaf curl virus clone Rad38**	PLCD-10	*Raphanus sativus*	2745	143-499	303-1073	1070-1474	1215-1619	1522-2607	2193-2450			
KY026598.1	**Papaya leaf curl virus clone Rad07**	PLCD-11	*Raphanus sativus*	2746	143-499	303-1073	1070-1474	1215-1619	1522-2607	2193-2450			
GU136803.1	**Chilli leaf curl virus Isolate India:Amritsar: Papaya: 2009**	PLCD-12	*Carica papaya*	2763	147-512	307-1077	1074-1478	1219-1623	1526-2460	2161-2460			

JN558352.1	**Cotton leaf curl Multan virus isolate**	PLCD-13	*Carica papaya*	2725	117-482	277-1047	1050-1454	1147-1599	1496-2581	2128-2430			
DQ629103.1	**Papaya leaf curl virus [India:New Delhi: tomato:2005]**	PLCD-14	*Carica papaya*	2765	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2450			
KX302707.1	**Cotton leaf curl Burewala virus isolate Guntur clone LK_2N**	PLCD-15	*Carica papaya*	2758	131-487	291-1061	1058-1462	1294-1503	1504-2595	2241-2681			
GQ200446.1	**Papaya leaf curl virus [India:Pratapgarh1:2008] clone SHLD-NIFL-Pra-01**	PLCD-16	*Crotalaria juncea*	2738	146-502	306-1076	1060-1464	1205-1609	1512-2597	2183-2500			
GQ200447.1	**Papaya leaf curl virus [India:Pratapgarh2:2008] clone SHLD-NIFL-Pra-02**	PLCD-17	*Crotalaria juncea*	2738	146-502	306-1076	1060-1464	1205-1609	1512-2597	2183-2500			
GQ200448.1	**Papaya leaf curl virus [India:Pratapgarh2:2008] clone SHLD-NIFL-Pra-03**	PLCD-18	*Crotalaria juncea*	2738	146-502	306-1076	1060-1464	1205-1609	1512-2597	2183-2500			
DQ989325.1	**Tomato leaf curl New Delhi virus-Papaya [India:New Delhi:Papaya:2005]**	PLCD-19	*Carica papaya*	2735	127-465	287-1057	1199-1603	1054-1464	1506-2597	2264-2440	317-802		48-434
DQ989326.1	**Chilli leaf curl virus-India [India:Papaya:2005]**	PLCD-20	*Carica papaya*	2764	148-504	308-1078	1075-1479	1220-1624	1527-2612	2162-2455	338-823		
HM140364.1	**Chilli leaf curl virus-DU [India:New Delhi:Papaya: 2009]**	PLCD-21	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2460			
HM140365.1	**Chilli leaf curl virus-HD [India:New Delhi:Papaya: 2007]**	PLCD-22	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2460			
HM140366.1	**Chilli leaf curl virus-Panipat1 [India:Panipat:Papaya:2008]**	PLCD-23	*Carica papaya*	2761	145-510	305-1075	1072-1476	1217-1621	1524-2609	2159-2458			
HM140370.1	**Chilli leaf curl virus-Najafgarh2 [India:New Delhi:Papaya: 2009]**	PLCD-24	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2460			
HM140371.1	**Chilli leaf curl virus-Noida**	PLCD-25	*Carica*	2762	146-	306-	1073-	1218-	1525-	2160-			

	[India:Uttar Pradesh:Papaya :2009]		*papaya*		502	1076	1477	1622	2610	2459			
HM140367.1	**Papaya leaf crumple virus-Panipat 8 [India:Panipat: Papaya:2008]**	PLCD-26	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
NC_014707.1	**Papaya leaf crumple virus-Panipat 8 [India: Panipat:Papaya:2008]**	PLCD-27	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
HM140368.1	**Papaya leaf crumple virus-Nirulas [India:New Delhi :Papaya:2007]**	PLCD-28	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
HM140369.1	**Papaya leaf crumple virus-Najafgarh 1 [India:New Delhi:Papaya:2008]**	PLCD-29	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
KJ028210.1	**Papaya leaf crumple virus clone Moh7**	PLCD-30	*Solanum nigrum*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
KM359408.1	**Papaya leaf crumple virus isolate A-87**	PLCD-31	*Andrographis paniculata*	2737	121-459	281-1051	1048-1452	1193-1600	1524-2588	2129-2431			
KR052159.1	**Papaya leaf crumple virus isolate Mohali**	PLCD-32	*Carica papaya*	2736	120-458	280-1050	1047-1451	1192-1599	1523-2587	2128-430			
JN807765.2	**Papaya leaf curl virus-[soybean: Lucknow]**	PLCD-33	*Glycine max*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KX353622.2	**Papaya yellow leaf curl virus isolate DP2**	PLCD-34	*Carica papaya*	2759	145-501	305-1075	1072-1476	1217-1621	1524-2606	2159-2452			
KP725055.1	**Tomato leaf curl virus isolate C1**	PLCD-35	*Carica papaya*	2757	144-491	304-1074	1071-1475	1216-1620	1532-2608	2158-2451			

ANNEXURE 2

Table A2. List of Betasatellite virus genomes used in this study. Betasatellite sequences causing leaf curl disease in papaya and associated crops and weeds. c1= betasatellite ORF

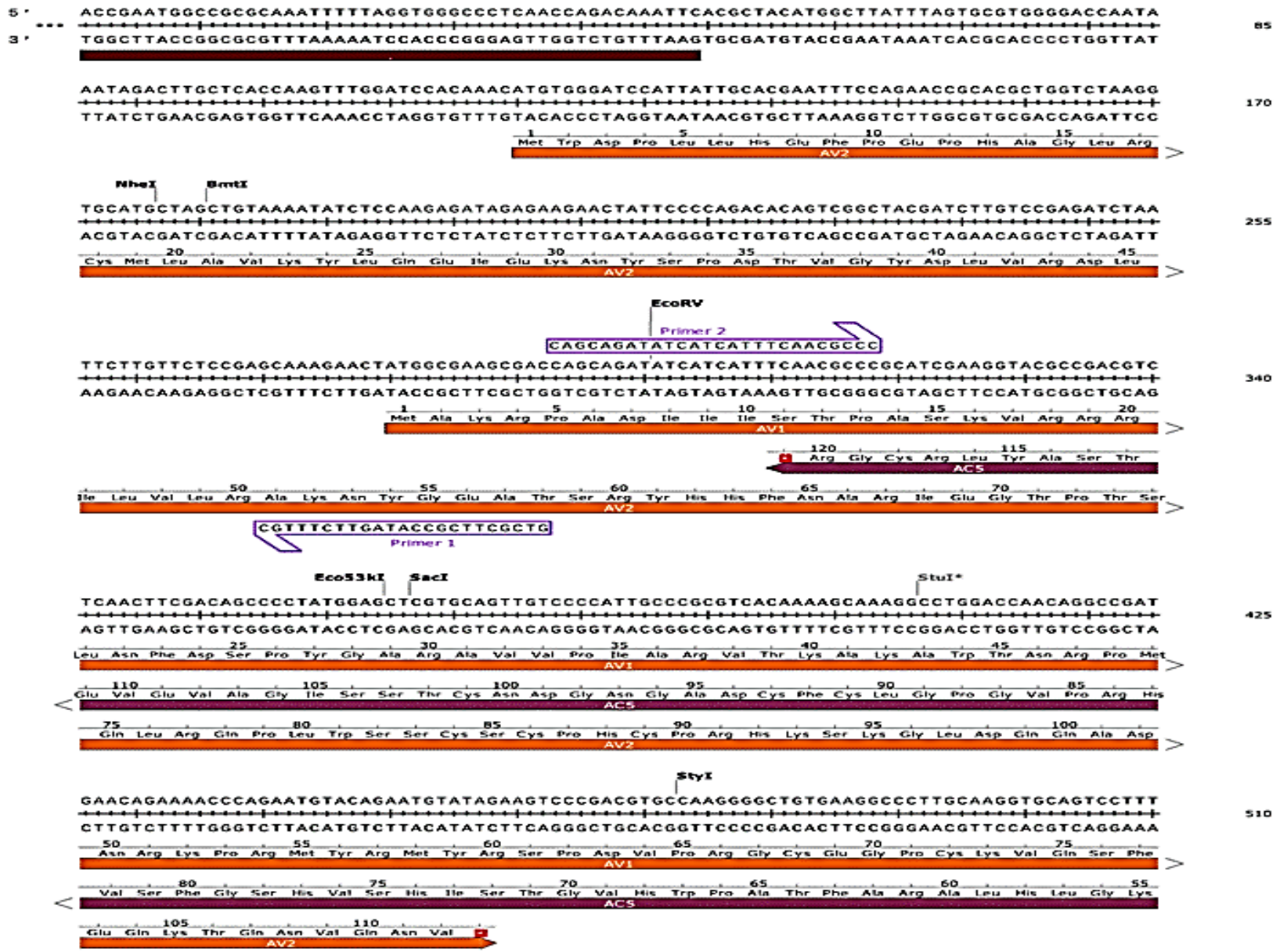
Annotation	Satellite	Accession No.	Host	Geography	Length	c1 gene
TBS 1	Ageratum leaf curl betasatellite clone beta	JX512904.2	*Amaranthus hypochondriacus L.*	Lucknow, India	1362	180-596
TBS 2	Ageratum leaf curl betasatellite isolate Jaipur	KY089034.1	*Helianthus sp.*	Jaipur, India	1366	180-596
TBS 3	Ageratum leaf curl betasatellite isolate NBRI-B1	KR922821.1	*Calendula officinalis*	Lucknow, India	1363	180-596
TBS 4	Ageratum leaf curl betasatellite isolate Sikar M 2	KC589700.1	*Tagetes patula*	Lakshmangarh, Rajasthan, India	1335	180-608
TBS 5	Chilli leaf curl betasatellite - [Pakistan:Potato:2008]	FM179615.1	*Solanum tuberosum*	Punjab, Pakistan	1399	201-563
TBS 6	Chilli leaf curl betasatellite-[India:Palampur:2008]	FM877803.2	*Capsicum sp.*	Palampur, India	1376	200-562
TBS 7	Chilli leaf curl betasatellite-Panipat 4 [India:Panipat: Papaya:2008]	HM143904.1	*Carica papaya*	Panipat, Haryana, India	1369	201-557
TBS 8	Chilli leaf curl virus satellite DNA beta	EU582020.1	*Capsicum sp.*	Pataudi, India	1380	200-562
TBS 9	Chilli leaf curl virus satellite DNA beta C1 gene, isolated from Capsicum annuum in Pakistan	FN179279.1	*Capsicum annuum*	Pakistan	1387	201-563
TBS 10	Chilli leaf curl betasatellite isolate ToLCBDB-[IN:Nar:Chil:04]	JF706231.1	*Capsicum sp.*	Jodhpur, India	1380	200-562
TBS 11	Chilli leaf curl betasatellite isolate Meerut	JX193616.1	*Capsicum sp.*	Meerut, India	1390	221-670
TBS 12	Chilli leaf curl betasatellite isolate India:Punjab:TC241:2009	KJ605111.1	*Solanum lycopersicum*	Muskabad, Punjab, India	1373	201-563
TBS 13	Cotton leaf curl Burewala betasatellite	NC_013802.1	*Gossypium hirsutum*	Punjab, India	1354	198-554
TBS 14	Cotton leaf curl Burewala betasatellite, clone L2-RCA-bl-F	FN658722.1	*Gossypium hirsutum*	Punjab, India	1354	198-554
TBS 15	Papaya leaf curl virus-associated DNA beta	AY244706.1	*Carica papaya*	New Delhi, India	1372	201-557
TBS 16	Papaya leaf curl virus-associated DNA beta	NC_004706.1	*Carica papaya*	New Delhi, India	1372	201-557
TBS 17	Papaya leaf curl virus betasatellite isolate PRM	GU370715.1	*Solanum lycopersicum*	New Delhi, India	1377	200-556
TBS 18	Papaya leaf curl virus betasatellite isolate In:Var:Pum:08:1	HM101173.1	Pumpkin	Varanasi, India	1370	201-557
TBS 19	Papaya leaf curl virus betasatellite isolate PaLCuB-IYV:Del	JX050199.1	*Ipomoea purpurea*	New Delhi, India	1367	201-557

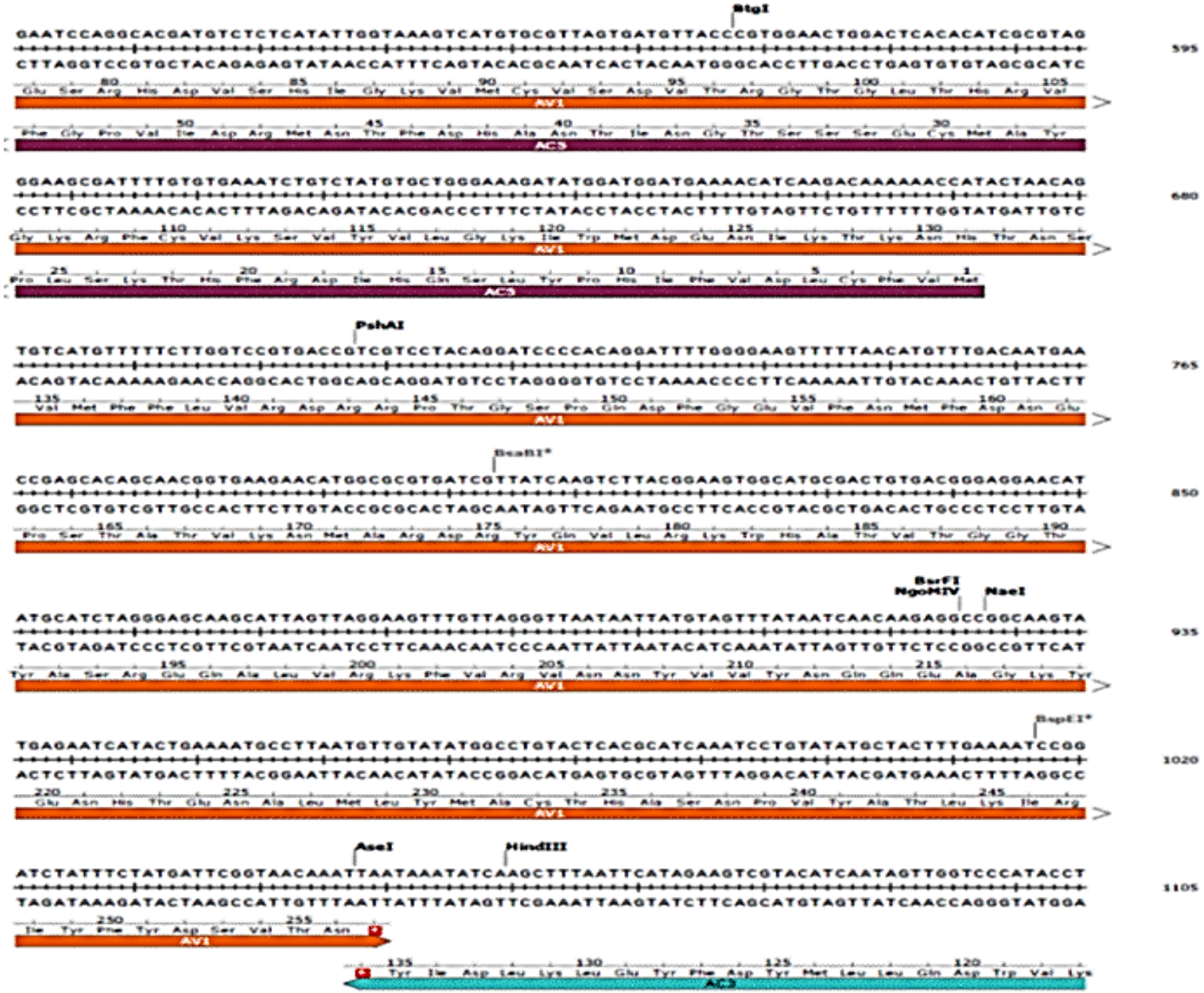
TBS 20	Papaya leaf curl virus betasatellite isolate PaLCuB-Pumpkin:IARI	JX040472.1	Pumpkin	New Delhi	1367	201-557
TBS 21	Papaya leaf curl virus betasatellite clone BG-CBE beta	KC959933.1	Black gram	Coimbatore, India	1351	188-556
TBS 22	Papaya leaf curl virus betasatellite clone GG-CBE	KC959934.1	Green gram	Coimbatore, India	1358	188-556
TBS 23	Papaya leaf curl virus beta-satellite clone BG-VBN	KC959935.1	Black gram	Pudukottai, India	1359	188-556
TBS 24	Papaya leaf curl virus beta- satellite isolate India:Pune: TC255:2010	KJ605112.1	*Solanum lycopersicum*	Pune, India	1367	200-556
TBS 25	Papaya leaf curl virus beta-satellite isolate India:Bangalore: TC281:2010	KJ605113.1	*Solanum lycopersicum*	Bangalore, India	1367	201-557
TBS 26	Papaya leaf curl betasatellite complete sequence, isolate Palampur	LN831955.1	*Valeriana jatamansi*	Palampur, India	1367	201-557
TBS 27	Papaya leaf curl betasatellite isolate NBRI	JX987089.2	*Parthenium hysterophorus L.*	Lucknow, India	1367	189-557
TBS 28	Papaya leaf curl betasatellite isolate GMT32	KT948074.1	*Cucurbita pepo*	Rawalpindi, Pakistan	1368	195-551
TBS 29	Papaya leaf curl betasatellite, clone Par-B1	LN906595.1	*Parthenium hysterophorus L.*	Lahore, Pakistan	1362	195-551
TBS 30	Papaya leaf curl betasatellite isolate India-Valsad-Cluster bean-2015	KT253636.1	*Cyamopsis tetragonoloba*	Valsad, Gujrat, India	1367	219-575
TBS 31	Papaya leaf curl betasatellite isolate India-Bhavnagar-Cluster bean-2015	KT253637.1	*Cyamopsis tetragonoloba*	Bhavnagar, Gujrat, India	1367	219-575
TBS 32	Papaya leaf curl betasatellite isolate DPB1	KX353621.1	*Solanum lycopersicum*	Sikar, Rajasthan, India	1367	189-557
TBS 33	Papaya leaf curl betasatellite isolate 2491	KY825245.1	*Capsicum* sp.	Islamabad, Pakistan	1331	221-577
TBS 34	Papaya leaf curl betasatellite isolate 2481	KY825246.1	*Capsicum* sp.	Islamabad, Pakistan	1369	221-577
TBS 35	Papaya leaf curl betasatellite isolate 49sb	KY825247.1	*Capsicum* sp.	Islamabad, Pakistan	1368	220-576
TBS 36	Tomato leaf curl virus-associated DNA beta	NC_004715.1	*Solanum lycopersicum*	Jabalpur, Madhya Pradesh, India	1424	195-221
TBS 37	Tomato leaf curl betasatellite isolate ToLCuB[IN:Bah]	KM201278.1	*Solanum lycopersicum*	India	1353	191-571
TBS 38	Tomato leaf curl betasatellite - [India:Halwadni:2007]	EU847239.1	*Solanum lycopersicum*	Haldwani, Uttarakhand, India	1370	200-556
TBS 39	Tomato leaf curl betasatellite-Naj 1 [India:New Delhi:Papaya:2008]	HM143909.1	*Carica papaya*	Najafgarh, New Delhi, India	1369	199-555

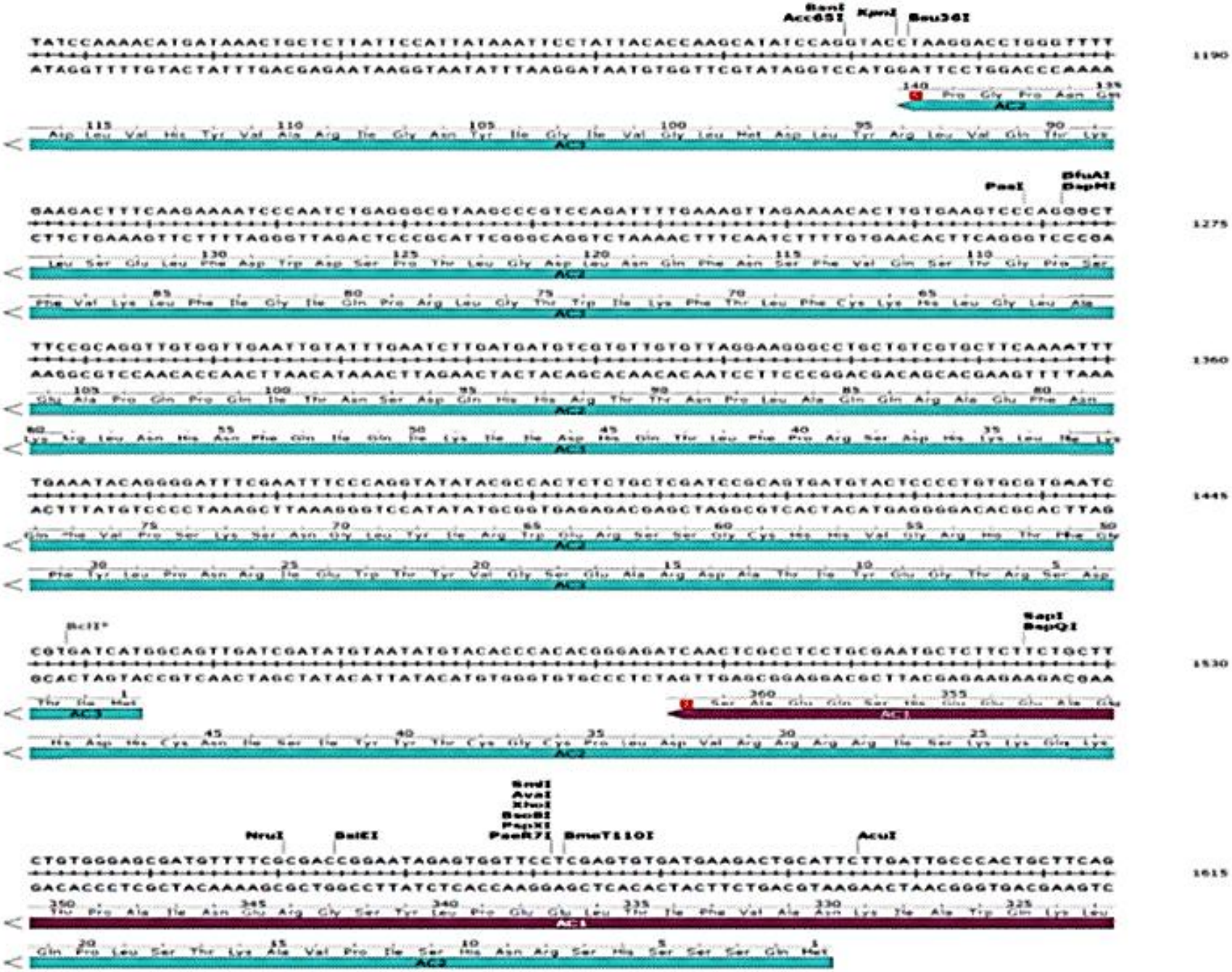
TBS 40	Tomato leaf curl betasatellite-DU [India:New Delhi:Papaya:2009]	HM143910.1	*Carica papaya*	New Delhi, India	1370	201-557
TBS 41	Tomato leaf curl betasatellite-Naj 2 [India:New Delhi:Papaya:2009]	HM143911.1	*Carica papaya*	Najafgarh, New Delhi, India	1370	201-557
TBS 42	Tomato leaf curl betasatellite-Panipat 1 [India:Panipat:Papaya:2008]	HM143901.1	*Carica papaya*	Panipat, Haryana, India	1369	201-557
TBS 43	Tomato leaf curl betasatellite-Panipat 2 [India:Panipat:Papaya: 2008]	HM143902.1	*Carica papaya*	Panipat, Haryana, India	1370	200-556
TBS 44	Tomato leaf curl betasatellite-Panipat 5 [India:Panipat:Papaya: 2008	HM143905.1	*Carica papaya*	Panipat, Haryana, India	1373	201-557
TBS 45	Tomato leaf curl betasatellite-Panipat 7 [India:Panipat:Papaya :2008	HM143907.1	*Carica papaya*	Panipat, Haryana, India	1375	201-557
TBS 46	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D17	LT009401.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 47	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D1	LT009403.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 48	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D22	LT009406.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 49	Tomato yellow leaf curl China virus-associated DNA beta complete genome, isolate Y87	AJ457818.1	*Solanum lycopersicum*	Yunnan, China	1344	209-565
TBS 50	Tomato yellow leaf curl China virus-associated DNA beta complete genome, isolate Y38	AJ420315.2	*Solanum lycopersicum*	Yunnan, China	1338	209-565
TBS 51	Tomato leaf curl Joydebpur beta virus	NC_010236.1	*Capsicum* sp.	Kalyani, West Bengal, India	1370	193-573

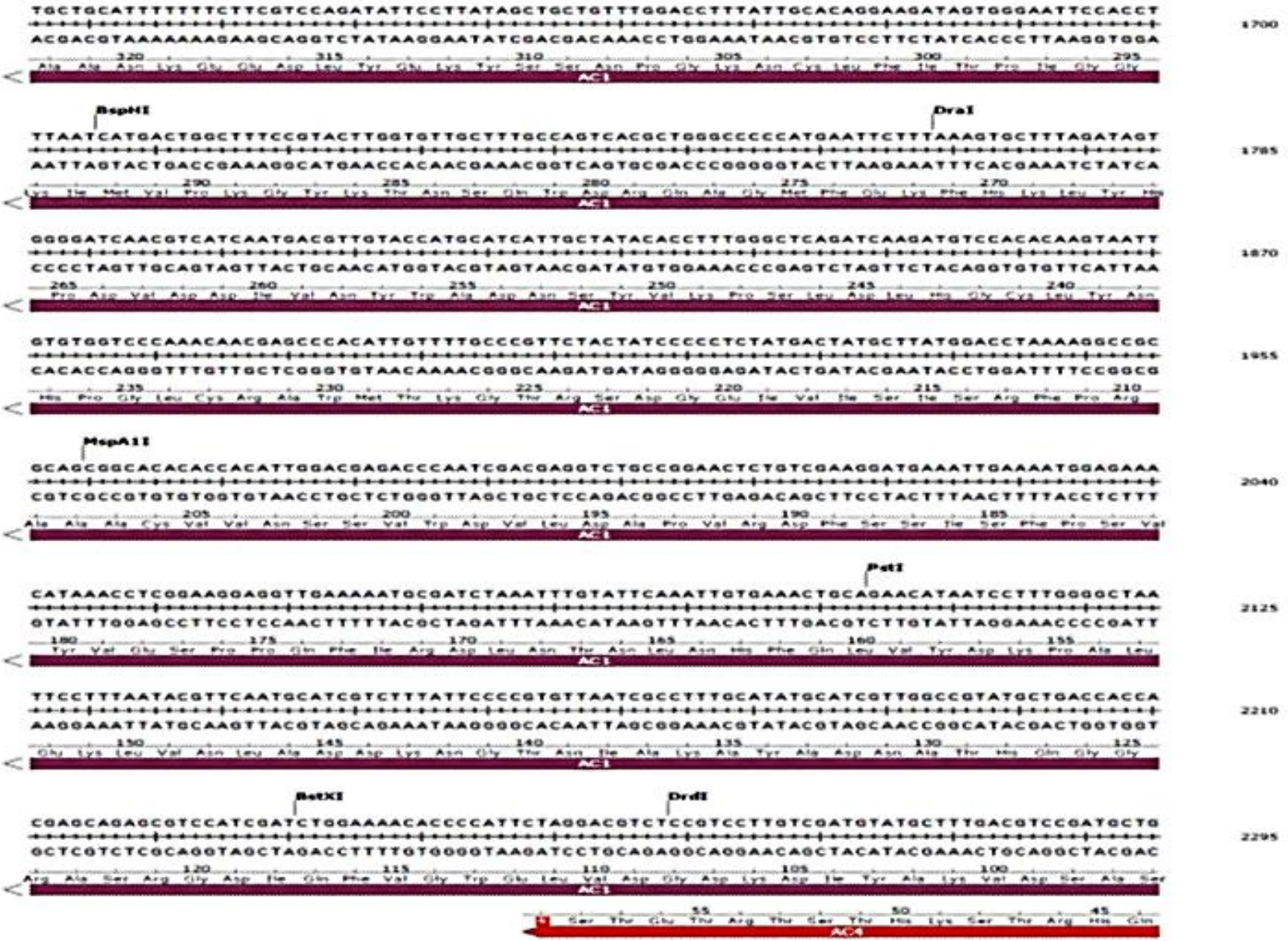
ANNEXURE 3

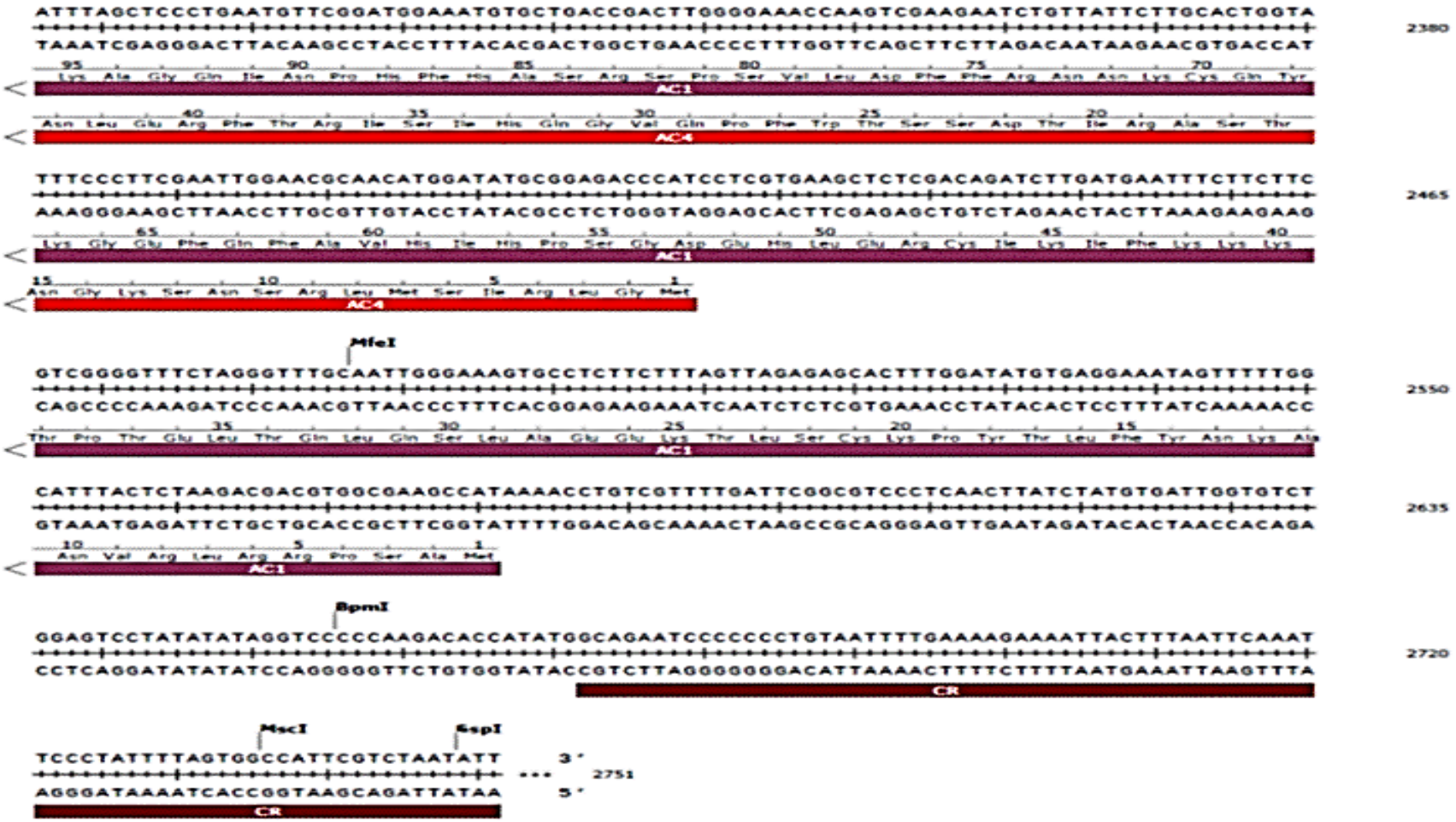
Figure A1. Annotated nucleotide sequence of ***Tomato leaf curl virus [India:Sultanpur:Papaya:2016]*** **(MH105055)**




BtgI
595
AV1
AC3
680
PshAI
765
BseBI*
850
BsrFI
NgoMIV
NaeI
935
BspEI*
1020
AseI
HindIII
1105

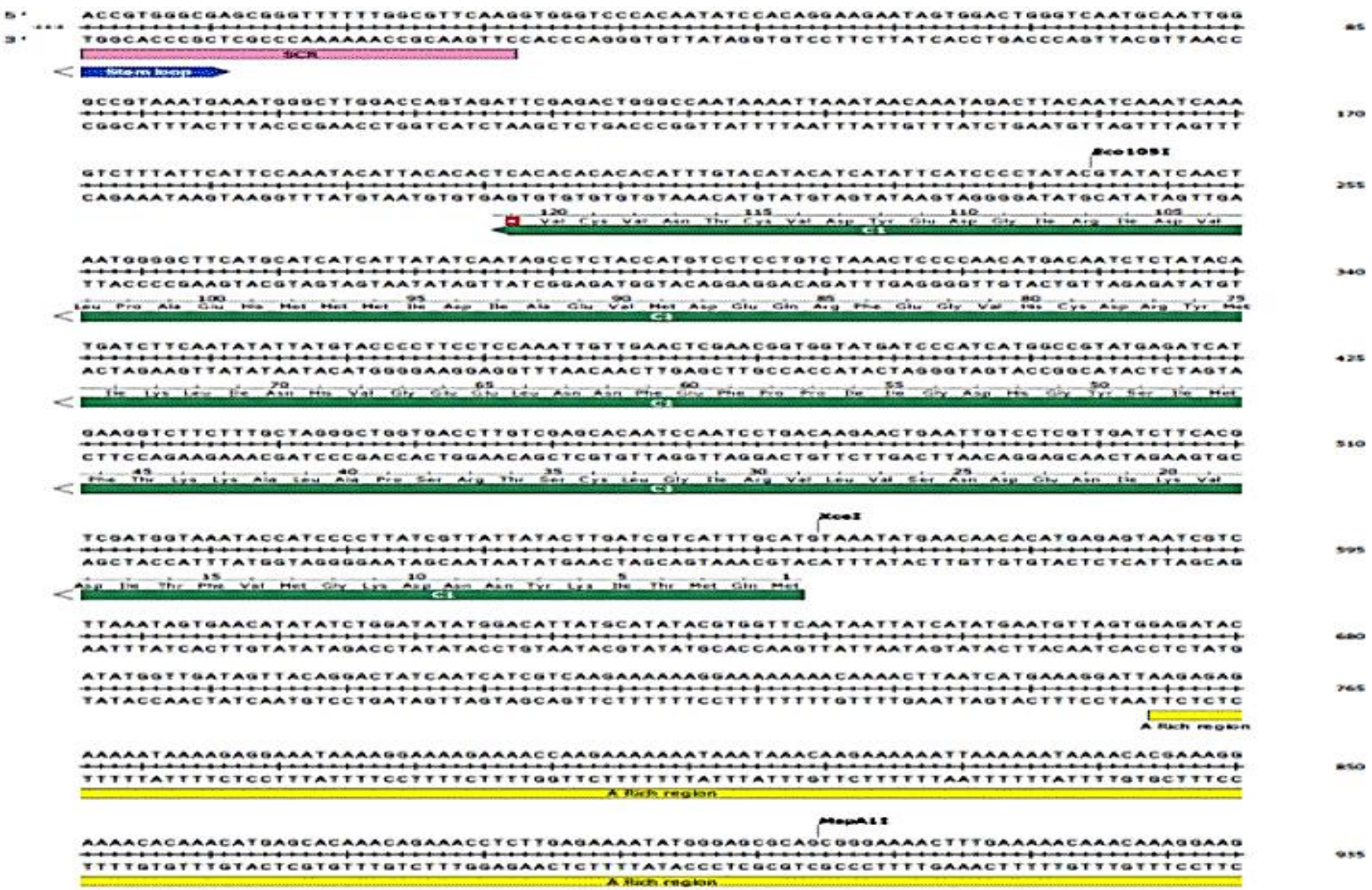

BsmI
Acc65I
KpnI
Bsu36I
PaeI
BfuAI
DapMI
BclI*
SapI
BspQI
NruI
BsiEI
SmaI
AvaI
XhoI
BsoBI
PagXI
PaeR7I
BmeT110I
AcuI


1700
1785
1870
1955
2040
2125
2210
2295
BspHI
DraI
MspA1I
PstI
BstXI
DrdI
ACI
ACI
ACI
ACI
ACI
ACI
ACI
ACI


ATTTAGCTCCCTGAATGTTCGGATGGAAATGTGCTGACCGACTTGGGGAAACCAAGTCGAAGAATCTGTTATTCTTGCACTGGTA
TAAATCGAGGGACTTACAAGCCTACCTTTACACGACTGGCTGAACCCCTTTGGTTCAGCTTCTTAGACAATAAGAACGTGACCAT
2380
AC1
AC4
TTTCCCTTCGAATTGGAACGCAACATGGATATGCGGAGACCCATCCTCGTGAAGCTCTCGACAGATCTTGATGAATTTCTTCTTC
AAAGGGAAGCTTAACCTTGCGTTGTACCTATACGCCTCTGGGTAGGAGCACTTCGAGAGCTGTCTAGAACTACTTAAAGAAGAAG
2465
AC1
AC4
MfeI
GTCGGGGTTTCTAGGGTTTGCAATTGGGAAAGTGCCTCTTCTTTAGTTAGAGAGCACTTTGGATATGTGAGGAAATAGTTTTTGG
CAGCCCCAAAGATCCCAAACGTTAACCCTTTCACGGAGAAGAAATCAATCTCTCGTGAAACCTATACACTCCTTTATCAAAAACC
2550
AC1
CATTTACTCTAAGACGACGTGGCGAAGCCATAAAACCTGTCGTTTTGATTCGGCGTCCCTCAACTTATCTATGTGATTGGTGTCT
GTAAATGAGATTCTGCTGCACCGCTTCGGTATTTTGGACAGCAAAACTAAGCCGCAGGGAGTTGAATAGATACACTAACCACAGA
2635
AC1
BpmI
GGAGTCCTATATATAGGTCCCCCAAGACACCATATGGCAGAATCCCCCCCCTGTAATTTTGAAAAGAAAATTACTTTAATTCAAAT
CCTCAGGATATATATCCAGGGGGTTCTGTGGTATACCGTCTTAGGGGGGGACATTAAAACTTTTCTTTTAATGAAATTAAGTTTA
2720
CR
MscI
SspI
TCCCTATTTTAGTGGCCATTCGTCTAATATT 3'
2751
AGGGATAAAATCACCGGTAAGCAGATTATAA 5'
CR

ANNEXURE 4

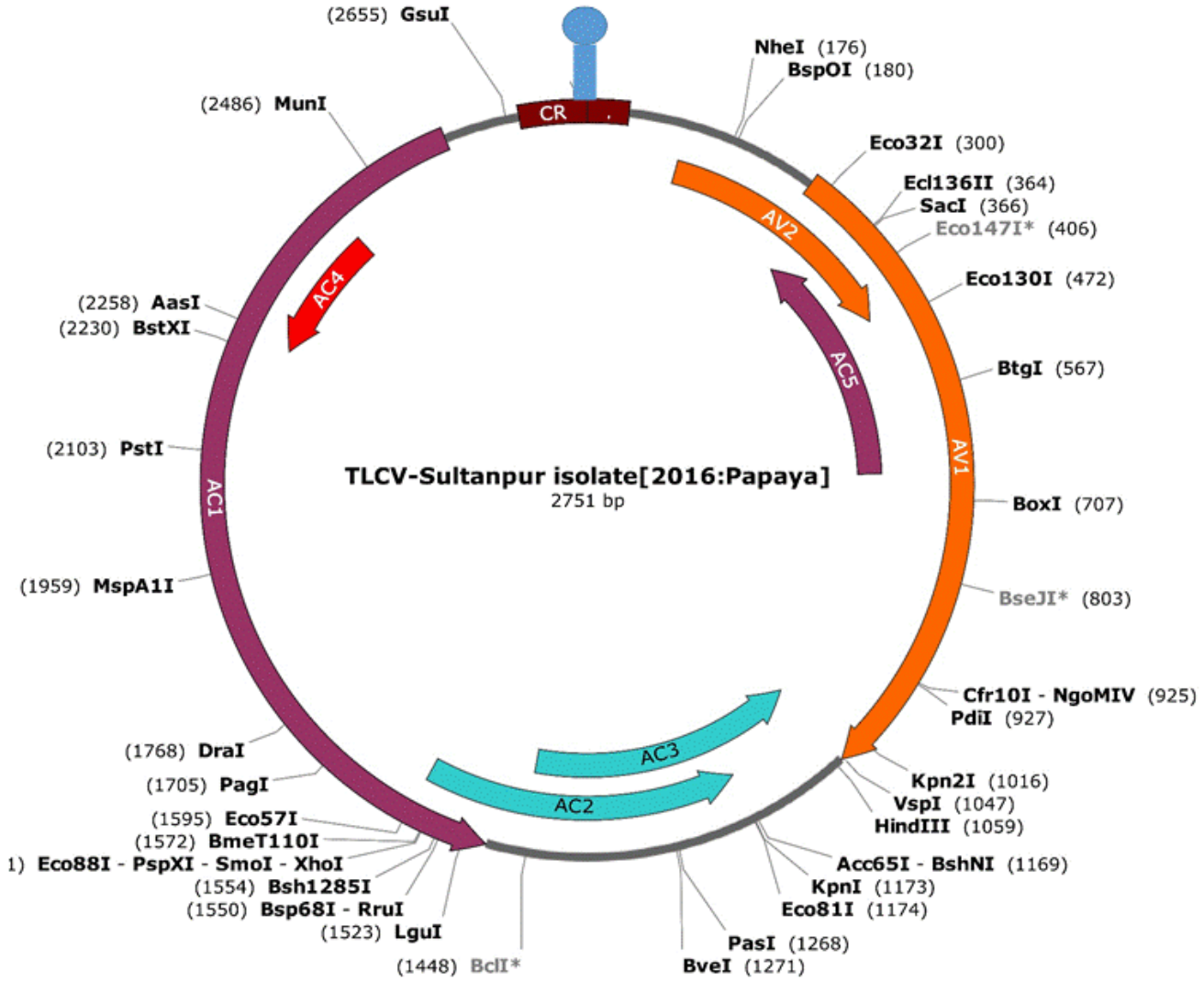
Figure A2. Annotated nucleotide sequence of ***Tomato leaf curl betasatellite-Naj2 [India:New Delhi:Papaya:2009]*** **(HM143911.1)**




ACACATGAGCAAAAAAAGGAAATACAAAACTAAACTACGAGGGCTTTACATGAAATTTTAGAAAGTCCGTACACAGTAATTAATC
1020
TGTGTACTCGTTTTTTTCCTTTATGTTTTGATTTGATGCTCCCGAAATGTACTTTAAAATCTTTCAGGCATGTGTCATTAATTAG
A Rich region
Eco130I
ATTAATTACTGCGTAGTAAATGTGAATAAAATTAACCCAAGGGGTTAATTTTGGAGTCTCCAATAGGTAAATGAGTCCCCAATTG
1105
TAATTAATGACGCATCATTTACACTTATTTTAATTGGGTTCCCCAATTAAAACCTCAGAGGTTATCCATTTACTCAGGGGTTAAC
Acc65I
KpnI
BamHI
AGGCACCGATATATCGGTACCTCGGAGAGAGAAATAAATCCCGGATCCCGAAAATACCCCCAAGTCTGTGTCTGGAAGGCGCGTG
1190
TCCGTGGCTATATAGCCATGGAGCCTCTCTCTTTATTTAGGGCCTAGGGCTTTTATGGGGGTTCAGACACAGACCTTCCGCGCAC
SCR
GTAATGCGCTGAAAAAGGCGACCGTCTCTCTCCTAAAACTCACCGGAACCTCCAAACTGGCTGATGCCGGTGTCAATTTTCGACA
1275
CATTACGCGACTTTTTCCGCTGGCAGAGAGAGGATTTTGAGTGGCCTTGGAGGTTTGACCGACTACGGCCACAGTTAAAAGCTGT
SCR
SacI
CGCGCGGCGGTGTGTACCCCTGGGAGGGTAGAAACCTCTACGCTACGCAGCAGCCTTAGCTACGCCGGAGCTGAGCTCGCCCACG
1360
GCGCGCCGCCACACATGGGGACCCTCCCATCTTTGGAGATGCGATGCGTCGTCGGAATCGATGCGGCCTCGACTCGAGCGGGTGC
SCR
TTCTAATATT 3'
1370
AAGATTATAA 5'
SCR
Stem loop

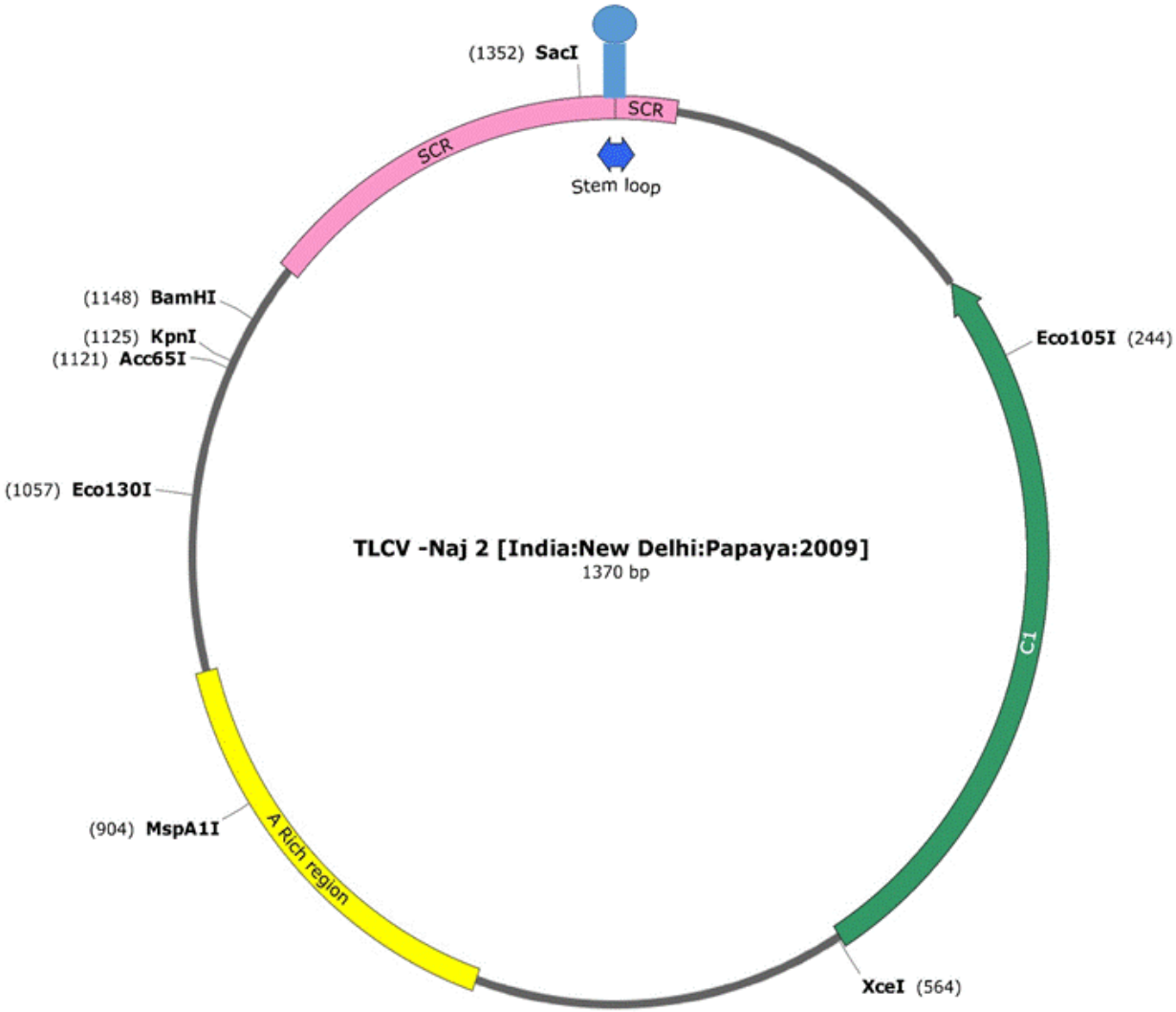
ANNEXURE 5

DNA-A: *Tomato leaf curl virus [India:Sultanpur:Papaya:2016] [MH105055]*



ANNEXURE 6

Betasatellite: *Tomato leaf curl betasatellite-Naj 2 [India:New Delhi:Papaya: 2009]*



ANNEXURE 7

Table A3. Identification of target region in DNA-A component of PLCD complex. Target sequences with their predicted siRNA parameters using online pssRNAit (http://plantgrn.noble.org/pssRNAit/server. (AS = antisense strand; S = sense strand)

siRNA rank	siRNA		Region	Efficiency	RISC Binder		Target accessibility	Off-targets
	AS	S			AS	S		
1	UACAUAUCGAUCAACUGCCAU	GGCAGUUGAUCGAUAUGUAAU	1453-1473	10.42	0.34	-2.21	18.754	1
2	UAACGCACAUGACUUUACCAA	GGUAAAGUCAUGUGCGUUAGU	536-556	9.45	1.01	-0.62	18.219	0
3	AACCACUCUAUUCCGGUCGCG	CGACCGGAAUAGAGUGGUUCC	1549-1569	9.36	0.65	-0.45	24.875	3
4	UAGCCGACUGUGUCUGGGGAA	CCCCAGACACAGUCGGCUACG	215-235	9.15	1.95	-2.15	19.896	1
5	UAACUAAUGCUUGCUCCCUAG	AGGGAGCAAGCAUUAGUUAGG	857-877	9.14	0.81	0.11	11.122	2
6	UACCAGUGCAAGAAUAACAGA	UGUUAUUCUUGCACUGGUAUU	2360-2380	9.07	0.32	-0.41	23.519	2
7	UCCUGUGCAAUAAAGGUCCAA	GGACCUUUAUUGCACAGGAAG	1660-1680	9.07	0.15	-1.32	19.809	1

ANNEXURE 8

Table A4. Off-target prediction analysis. Expect score and off-target accessibility were used to infer strength of off-targets.

siRNA rank	Region	Target gene	Predicted Off-target	Off-targets	Expect	Strength
1	1453-1473	AC3	1	AT4G03153.1 Kinase interacting (KIP1-like) family protein	3.5	Moderate
2	536-556	AV1/AC5	0			-
3	1549-1569	AC2	3	AT5G43415.1 transposable element gene AT2G21760.1 pre-tRNA AT3G59910.1 Ankyrin repeat family protein	3.5 3.5 3.5	Weak Weak Weak
4	215-235	AV2	1	AT1G13030.1 sphere organelles protein-related	3.5	Moderate
5	857-877	AV1	2	AT5G49490.1 AGAMOUS-like 83 AT2G06900.1 transposable element gene	2.5 3.5	Moderate Weak
6	1660-1680	AC1	1	AT2G20380.1 Galactose oxidase/kelch repeat superfamily	2.5	Weak
7	2360-2380	AC1/AC4	2	AT1G35530.2 DEAD/DEAH box RNA helicase family protein AT1G02230.1 NAC domain containing protein 4	3.5 3.5	Moderate Weak

List of Publications

1. Verma, S. and Saxena, S. (2017). Dissection of Papaya Leaf Curl Disease (PLCD) Complex and Assessing it's Potential for siRNA Based Targeting. International Journal of Plant & Environment 3:17-29

2. Verma, S. and Saxena, S. (2018). *in-silico* identification of betasatellite complexity in Papaya leaf curl disease complex. International Journal of Plant & Environment (Accepted for publication in Vol. 4 No.2)

3. Saxena, S., Singh, V.K., and Verma, S. (2016). PCR mediated detection of sex and PaLCuV infection in papaya - A review. Journal of Applied Horticulture 18:80-84

4. Saxena, S. and Verma, S. (2016) . Chapter 20. Harnessing the Variability in Plant-Virus-Vector complex interaction in Begomovirus family to prevent viral disease outbreak. In: Sobti, R.C., Jaiswal, K. and Misra, S, Recent advances in Applied Biosciences. Narendra Publishing House, New Delhi.

5. Verma, S. and Saxena, S. (2017). Chapter 3: Recent Advancement in Diagnosis of Begomoviruses. In Saxena, S. and Tiwari, A.K. (Ed). Begomoviruses: Occurrence and Management in Asia and Africa, Springer Nature, Singapore.

6. Shukla, N.*, Verma, S*, Babu, S.G. and Saxena, S. (2017). Chapter 8: Strategy for Generic Resistance Against Begomoviruses Through RNAi. In Saxena, S. and Tiwari, A.K. (Ed). Begomoviruses: Occurrence and Management in Asia and Africa, Springer Nature, Singapore.

List of Conferences/Workshops/Symposiums

- **Poster** presentation titled "Identification of siRNA hotspots in Begomoviruses causing leaf curl disease." at **8th International Geminivirus Symposium & The 6th International ssDNA Comparative Virology Workshop**, New Delhi, 7th Nov-11th November 2017
- **Oral** presentation titled **"In-silico Approach to dissect the molecular diversity of New World Begomovirus"** at **3rd Lucknow Science Congress (3rd LUSCON) and National Conference on Science for Society: An Interdisciplinary Approach** organized by BBAU, Lucknow from **31stOctober-2nd November 2015**.
- **Oral** presentation titled **"Old World Begomovirus: A Molecular Study"** at **26th All India Congress of Zoology (26th AICZ) and International Symposium on Innovation in Animal Sciences for Food Security, Health Security and Livelihood-2015** organized by BBAU, Lucknow & the Zoological Society of India from **29th-31st October 2015**.
- Presented a poster titled **"Engineering Resistance against Begomovirus infection in plants by targeting multiple viral genes"** at the **Fifth International Conference on Plant & Environmental Pollution (5th ICPEP)**, organized by The International Society of Environmental Botanists (ISEB) & CSIR-National Botanical Research Institute (NBRI) held at NBRI, Lucknow from **24th-27th February, 2015**.
- Participated in the **International Symposium on "Genetics & Genomics in Modern Clinical Medicine"** organized by The Indo UK Genetic Education Forum, Wales Gene Park, Cardiff University, UK and King George Medical University, Lucknow; held at K.G.M.U., Lucknow on **9th and 10th February 2015**. Uttar Pradesh Medical Council has awarded **6 accreditation hours** for the course.
- Participation as a **Volunteer** in **15th Indo-US Flow Cytometry Workshop: Application of Flow Cytometry in Biomedical Research**, held at Department of Biotechnology, SB & BT, Babasaheb Bhimrao Ambedkar University, Lucknow from **29th – 31st October 2014**.
- Hands-On-Training on SEM, FTIR, FPLC and Ion Chromatography held at University Science and Instrumentation Center, BBAU, Lucknow during 18th-20th February 2015.
- Training in Good Lab Practice (GLP) organized by cCAMP, Bangalore at NCBS-TIFR Bangalore campus.

International Journal of Plant and Environment
Vol. 3, No. 2 (2017), 17-29

Dissecting Papaya Leaf Curl Disease (PLCD) Complex and Assessing its Potential for siRNA Based Targeting

Saurabh Verma and Sangeeta Saxena*

Department of Biotechnology, Babasaheb Bhimrao Ambedkar University, Lucknow-226024, INDIA

Publication Info

***Article history*:**
Received : 02.05.2017
Accepted : 24.07.2017
DOI :

***Key words*:**
Begomovirus
Genetic variability
Papaya leaf curl disease complex
Phylogenetic analysis
siRNA

Corresponding author:
Dr. Sangeeta Saxena
Tel.: +91- 9450645342
Email:
dr_sangeeta_saxena@yahoo.com

Abstract

Carica papaya L. (papaya) production is severely affected by leaf curl and related disease symptoms caused by various species of begomoviruses making it a complex disease. Papaya leaf curl disease (PLCD) complex has broad host range and high genetic variability, therefore, the study was conducted to dissect the genetic complexity of PLCD. Begomoviruses infect various host crops and weeds associated with cultivation of papaya. Intermixing among their genomes and mutational forces drive molecular variability in the disease complex, which enables them to expand their host range. This molecular variability is important determinant when RNA interference based strategy is applied against these viruses. In this study, we have observed that the broad host range of PLCD complex viral components is due to high frequency of recombination events that took place at different time scales. The rise of recombinants led to host variability resulting in similar symptomatic infection over new crops and weeds in nearby fields. Therefore, we have taken into account molecular variability and recombination regions of the viral genome as the targets for providing resistance against these viruses, thus, preventing serious implications of similar diseases spreading over to new hosts.

1. Introduction

Papaya fruit is used for various medicinal and commercial purposes apart from being an edible delicacy and an important source of daily nutrition like minerals, vitamins and papain. Due to its commercial importance, it is cultivated at large scale in almost all parts of India. However, in last decade, papaya farming has suffered a huge setback due to various pathogens infecting papaya crops in every season, especially the rainy season. Fungus, anthracnose, leaf curl, crumpled growth and early fruit fall are some of the problems related to rainy season (Roden and Ingle, 2009; Singh *et al.,* 2012; Varun *et al.,* 2017). Among all problems, the leaf curl disease of papaya has been a major adversity faced by papaya plantations causing a deep dip in papaya productivity (Saxena and Verma, 2016). Upward or downward leaf curl, dark green coloration, vein swelling, leaf mottling, crumpling and crippled plant development (Saxena *et al.,* 2016) characterize the leaf curl disease symptoms in papaya caused by a plant virus infection. This disease is endemic and causes huge losses to papaya farmers in India, especially in the northern and central regions of Uttar Pradesh, Haryana, Delhi and Rajasthan. Apart from papaya leaf curl, crops in India such as cotton, tomato, okra, chili and squash are highly susceptible to leaf curl disease caused by various other begomoviruses (Varma and Malathi, 2003; Borah and Dasgupta, 2012).

Begomoviruses are single stranded DNA viruses, characterized by the geminate shaped structure consisting of a capsid and viral genome. Their genome reported to be containing DNA-A, DNA-B, beta (DNA-β) and alpha satellite molecules. The genetic composition of these viruses depends upon their host plants and vector biotypes i.e. *Bemisia tabaci* (Gennadius) also known as whitefly (Brown *et al.,* 1995; Gutierrez, 2000; David, 2003; King *et al.,* 2011). Monopartite begomoviruses consist of DNA-A and associated satellites, whereas, bipartite consist of DNA-A and DNA-B. The common genetic component, DNA-A, of begomoviruses harbor structural and multi-functional protein encoding open reading frames (ORFs). All genetic components have been found to be associated with a 200bp long, highly conserved region composed of repetitive elements, promoter region, origin of replication and a hairpin loop (Harrison, 1985). Satellite molecules are incomplete defective genetic components which may or may not be functional inside the host crops. In addition, DNA-β has been reported to be associated with disease severity in several cases (Zhou, 2013). DNA-B harbors ORFs responsible for the movement of viral DNA from cytoplasm to nucleus, nucleus to cytoplasm and from one cell to another cell, hence also known as movement component of bipartite begomoviruses. Together, these ORFs encode proteins that take part in structure formation, replication initiation, transcriptional activation, replication

enhancement, host immune suppression and viral movement (Fondong, 2013). Hence, begomoviruses are complete parasites of plants, forming a very tight virus-vector-host complex (Hanley-Bowdoin *et al.,* 2013; Saxena and Verma, 2016).

The PLCD (Nadeem *et al.,* 1997; Saxena *et al.,* 1998a,b,c) is reported to be a complex system due to different begomoviral species involved in causing this symptom. The begomovirus isolates belonging to Papaya leaf curl virus (PaLCuV), Papaya leaf curl China virus (PaLCuCNV), Papaya leaf curl Guangdong virus (PaLCuGdV), Chili leaf curl virus (ChiLCV), Tomato leaf curl virus (ToLCV), Pedilanthus leaf curl virus (PeLCuV), Papaya leaf crumple virus (PapayaLCV) and beta (β) satellites associated with other monopartite begomoviruses causes severe leaf deformation symptoms in plants (Fig. 1). Among all PLCD causing begomoviruses, the Indian species form the components of largest disease complex infecting papaya and associated crops and weeds (Hallan *et al.,* 1998a,b; Raj *et al.,* 2008) (Table 1). There is an urgent need to develop assessment parameters to propose an effective and sustainable resistance strategy against PLCD complex as it is caused by more than one type of begomovirus species (Singh-Pant *et al.,* 2012).

In this study, several parameters such as conservation, sequence diversity and the occurrence of recombination events have been studied to assess the level of complexity of PLCD. The complexity will decide if this disease complex has potential for a broad spectrum strategy that could target most of the components of PLCD complex. Therefore, this study on PLCD complex is necessary for assessing the potential for siRNA based targeting (Saxena *et al.,* 2011, 2013; Ghosal *et al.,* 2012) of PLCD complex to develop PLCD resistant papaya crops.

2. Materials and Methods

2.1. Taxonomic analysis and sequence retrieval of PLCD complex

The taxonomic diversity of the leaf curl disease complex in papaya was investigated using Taxonomy browser at NCBI (https://www.ncbi.nlm.nih.gov/taxonomy). The virus species were manually screened for their presence in India and infection host as *Carica papaya* and associated weeds and crop plants. The

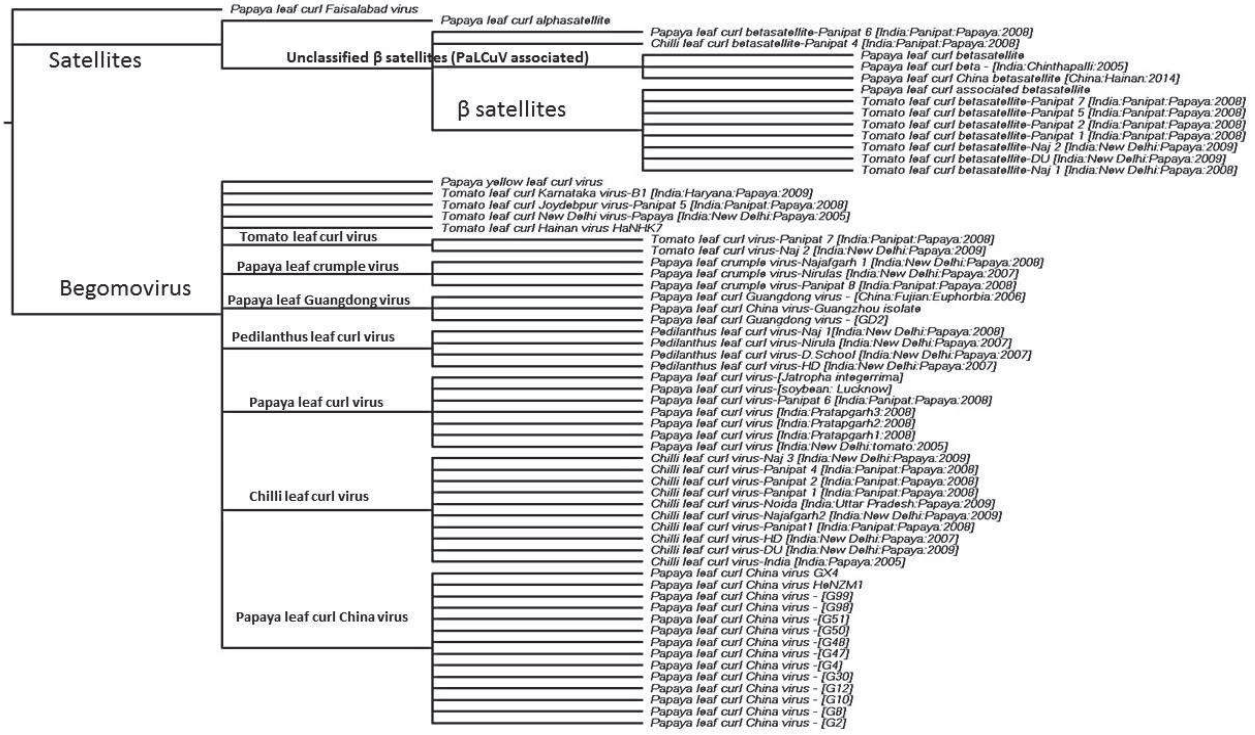


Fig. 1: Schematic representation of the Papaya leaf curl disease (PLCD) complex. The PLCD complex consists of Begomoviruses, incomplete defective satellite DNA molecules and unclassified viruses and satellite DNA molecules. The PLCD complex contains diverse strains exhibiting leaf curl and crumpling symptoms in papaya and associated crops and weeds

Table 1: List of DNA-A sequences of Indian PLCD complex begomoviruses causing leaf curl disease in papaya and associated crops and weeds

Accession ID	Virus	Annotation	Host	Size	AV2	AV1	AC3	AC2	AC1	AC4	AC5	AC6	AV3
Y15934.1	Papaya leaf curl virus	PLCD-1	*Carica papaya*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457	301-618		
NC_004147.1	Papaya leaf curl virus	PLCD-2	*Carica papaya*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457	301-618		
HM143914.1	Papaya leaf curl virus segment DNA A	PLCD-3	*Nicotiana glutinosa*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
JN135233.1	Papaya leaf curl virus	PLCD-4	*Amaranthus cruentus* L.	2746	150-506	310-1083	1077-1481	1222-1626	1529-2614	2200-2457			
JQ954859.1	Papaya leaf curl virus	PLCD-5	*Aster alpinus* L.	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KM525657.1	Papaya leaf curl virus	PLCD-6	*Croton bonplandianus* L.	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KU376493.1	Papaya leaf curl virus isolate CN2	PLCD-7	*Solanum lycopersicum*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KX302713.1	Papaya leaf curl virus isolate Wellington clone WB2&5	PLCD-8	*Carica papaya*	2763	149-511	315-1085	1082-1270	1267-1623	1535-2674	2400-2621	729-980	1227-1487	
KY800906.1	Papaya leaf curl virus isolate India/New Delhi/ Papaya/ 2016	PLCD-9	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1529-2611	2161-2460			
KY026597.1	Papaya leaf curl virus clone Rad38	PLCD-10	*Raphanus sativus*	2745	143-499	303-1073	1070-1474	1215-1619	1522-2607	2193-2450			
KY026598.1	Papaya leaf curl virus clone Rad07	PLCD-11	*Raphanus sativus*	2746	143-499	303-1073	1070-1474	1215-1619	1522-2607	2193-2450			
GU136803.1	Chilli leaf curl virus isolate India: Amritsar: Papaya: 2009	PLCD-12	*Carica papaya*	2763	147-512	307-1077	1074-1478	1219-1623	1526-2460	2161-2460			
JN558352.1	Cotton leaf curl Multan virus isolate	PLCD-13	*Carica papaya*	2725	117-482	277-1047	1050-1454	1147-1599	1496-2581	2128-2430			
DQ629103.1	Papaya leaf curl virus [India:New Delhi: tomato: 2005]	PLCD-14	*Carica papaya*	2765	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2450			
KX302707.1	Cotton leaf curl Burewala virus isolate Guntur clone LK_2N	PLCD-15	*Carica papaya*	2758	131-487	291-1061	1058-1462	1294-1503	1504-2595	2241-2681			
GQ200446.1	Papaya leaf curl virus [India:Pratapgarh1:2008] clone SHLD-NIFL-Pra-01	PLCD-16	*Crotalaria juncea*	2738	146-502	306-1076	1060-1464	1205-1609	1512-2597	2183-2500			

Accession ID	Virus	Annot-ation	Host	Size	AV2	AV1	AC3	AC2	AC1	AC4	AC5	AC6	AV3
GQ200447.1	Papaya leaf curl virus [India:Pratapgarh2:2008] clone SHLD-NIFL-Pra-02	PLCD-17	*Crotalaria juncea*	2738	146-502	306-1076	1060-1464	1205-1609	1512-2597	2183-2500			
GQ200448.1	Papaya leaf curl virus [India:Pratapgarh2:2008] clone SHLD-NIFL-Pra-03	PLCD-18	*Crotalaria juncea*	2738	146-502	306-1076	1060-1464	1205-1609	1512-2597	2183-2500			
DQ989325.1	Tomato leaf curl New Delhi virus-Papaya [India:New Delhi: Papaya: 2005]	PLCD-19	*Carica papaya*	2735	127-465	287-1057	1199-1603	1054-1464	1506-2597	2264-2440	317-802		48-434
DQ989326.1	Chilli leaf curl virus-India [India: Papaya: 2005]	PLCD-20	*Carica papaya*	2764	148-504	308-1078	1075-1479	1220-1624	1527-2612	2162-2455	338-823		
HM140364.1	Chilli leaf curl virus-DU [India: New Delhi : Papaya: 2009]	PLCD-21	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2460			
HM140365.1	Chilli leaf curl virus-HD [India: New Delhi: Papaya: 2007]	PLCD-22	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2460			
HM140366.1	Chilli leaf curl virus-Panipat1 [India: Panipat: Papaya: 2008]	PLCD-23	*Carica papaya*	2761	145-510	305-1075	1072-1476	1217-1621	1524-2609	2159-2458			
HM140370.1	Chilli leaf curl virus-Najafgarh2 [India:New Delhi: Papaya: 2009]	PLCD-24	*Carica papaya*	2763	147-503	307-1077	1074-1478	1219-1623	1526-2611	2161-2460			
HM140371.1	Chilli leaf curl virus-Noida [India: Uttar Pradesh: Papaya :2009]	PLCD-25	*Carica papaya*	2762	146-502	306-1076	1073-1477	1218-1622	1525-2610	2160-2459			
HM140367.1	Papaya leaf crumple virus-Panipat 8 [India: Panipat: Papaya: 2008]	PLCD-26	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
NC_014707.1	Papaya leaf crumple virus-Panipat 8 [India: Panipat: Papaya: 2008]	PLCD-27	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
HM140368.1	Papaya leaf crumple virus-Nirulas [India:New Delhi: Papaya: 2007]	PLCD-28	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		

Accession ID	Virus	Annot-ation	Host	Size	AV2	AV1	AC3	AC2	AC1	AC4	AC5	AC6	AV3
HM140369.1	Papaya leaf crumple virus-Najafgarh 1 [India:New Delhi: Papaya: 2008]	PLCD-29	*Carica papaya*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
KJ028210.1	Papaya leaf crumple virus clone Moh7	PLCD-30	*Solanum nigrum*	2736	120-485	280-1050	1047-1451	1192-1599	1523-2587	2128-2430	258-548		
KM359408.1	Papaya leaf crumple virus isolate A-87	PLCD-31	*Andrographis paniculata*	2737	121-459	281-1051	1048-1452	1193-1600	1524-2588	2129-2431			
KR052159.1	Papaya leaf crumple virus isolate Mohali	PLCD-32	*Carica papaya*	2736	120-458	280-1050	1047-1451	1192-1599	1523-2587	2128-430			
JN807765.2	Papaya leaf curl virus-[soybean: Lucknow]	PLCD-33	*Glycine max*	2746	150-506	310-1080	1077-1481	1222-1626	1529-2614	2200-2457			
KX353622.2	Papaya yellow leaf curl virus isolate DP2	PLCD-34	*Carica papaya*	2759	145-501	305-1075	1072-1476	1217-1621	1524-2606	2159-2452			
KP725055.1	Tomato leaf curl virus isolate C1	PLCD-35	*Carica papaya*	2757	144-491	304-1074	1071-1475	1216-1620	1532-2608	2158-2451			

AV1=Coat protein; AV2=variable function; AC1=replication initiation protein; AC2=transcription activator protein; AC3=replication enhancer protein; AC4=host RNA machinery suppressor; AC5, AC6 and AV3=uncharacterized open reading frames of begomoviruses of PLCD complex.

details were transformed into a schematic representation using the available tree tool. The manually screened virus species were used to retrieve nucleotide sequence information of complete DNA-A genome using Nucleotide browser at NCBI (https://www.ncbi.nlm.nih.gov/nuccore). All sequences were transferred as a FASTA format into a text file. The ORF information was retrieved using Sequence information tool for every nucleotide sequence.

2.2. Phylogenetic analysis of PLCD complex

Phylogenetic analysis was performed using two algorithms i.e. Bayesian and Maximum likelihood. Bayesian analysis was performed using TOPALi v2.5 software (www.topali.org/). The input file for bayesian phylogenetic analysis was prepared by performing a multiple sequence alignment using CLUSTALW software (www.ebi.ac.uk/Tools/msa/clustalw2/). The parameters selection was based upon prediction of suitable model by Model Selection tool and final analysis was run using MrBayes version 3.1.1 tool (http://mrbayes.sourceforge.net/) available in TOPALi v2.5. The maximum likelihood algorithm was used using the same input file as used previously. PhyML-aLRT version 2.4.5 tool (www.atgc-montpellier.fr/phyml/) available in TOPALi v2.5 software was used for inferring phylogenetic association of PLCD complex viral DNA-A components. The inferred trees were prepared for illustration using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/ software/figtree/).

2.3. Pairwise similarity and analysis

To construct a similarity matrix based on pairwise alignment of DNA-A, the pairwise distance estimation of DNA-A sequences of PLCD complex was carried out using Sequence demarcation tool (SDT) version 1.2 (web.cbio.uct.ac.za/SDT). The aligned viral components showing 90-100% sequence similarity were grouped together and assumed to be isolates of same species.

2.4. Recombination analysis

The CLUSTALW aligned file was used as input file for recombination analysis using Recombination detection program (RDT) v4.94 software (http://web.cbio.uct.ac.za/~darren/rdp.html). The recombination events and breakpoint distribution analysis was conducted using various tools available in Options toolbar of RDT software with default parameters as stated in RDP manual. The recombination events were screened manually according to the procedure listed in RDP manual. The recombination events were selected and analyzed manually for breakpoint boundary and parental contribution to assess recombinant viral component.

3. Results and Discussion

3.1. Components of PLCD complex

The PLCD complex consists of 13 begomovirus species associated with β satellites (Fig. 1). The begomovirus complex components are distributed mainly in Central Asia i.e. China, India, Pakistan and parts of Bangladesh. PaLCuCNV, PaLCuGnV and Tomato leaf curl Hainan virus (ToLCuHaV) are mainly reported in China infecting papaya, tomato and euphorbia. The Chinese components are associated with Papaya leaf curl China betasatellites. There are 18 different virus isolates comprising the Chinese PLCD complex components. The Pakistan PLCD complex consists of Papaya leaf curl Faisalabad virus, an unclassified virus that has been excluded from begomovirus. The Indian PLCD complex shows the most diverse nature in terms of number of species reported to cause leaf curl symptom in papaya. Nine begomovirus species i.e. ChiLCV, PaLCuV, PeLCuV, PapayaLCV, ToLCV, ToLCNDV, PaYLCuV etc. have been observed to infect papaya with broad host range. The Indian PLCD complex consists of 28 isolates belonging to the above mentioned begomoviruses. Overall, the PLCD complex consists of 47 different isolates identified to cause leaf curl disease in papaya and associated weeds and crops.

3.2. High molecular diversity of PLCD complex

The phylogenetic analysis was performed for manually screened isolates of Indian PLCD complex. Total thirty five isolates were chosen for phylogenetic analysis. The CLUSTALW output was used for Bayesian inference with MrBayes v.3.1.1 (Huelsenbeck and Ronquist, 2001) using General time reversible with gamma (GTR-G) substitution model (Larget and Simon, 1999) with two runs for 2,50,000 generations, sample frequency at 10 and burn-in ratio at 42% i.e. first 1,05,000 generations were excluded from the inference. The same set of PLCD complex DNA-A sequences were analyzed for maximum likelihood analysis using PhyML-aLRT v2.4.5 (Guindon and Gascuel, 2003) and HKY-G (Hasegawa, Kishino and Yano, 1985-Gamma) as a substitution model (Hasegawa *et al.,* 1985). The above phylogenetic analysis produced likelihood values that were very similar to each other i.e. 26743.85 and 26732.18 for MrBayes and PhyML respectively. The close lying likelihood scores means that the two trees can be interpreted in similar manner without much divergence. Therefore, the results and interpretation for MrBayes tree will be same as PhyML-aLRT tree in all future discussions.

The complexity of Indian PLCD complex is evident from the posterior distribution of model GTR-G parameters obtained using Markov Chain Monte Carlo (MCMC) methods in MrBayes inference (Larget and Simon, 1999; Ronquist and Huelsenbeck, 2003; Gamerman and Lopes, 2006) (Fig. 2). The use of GTR-G substitution model resulted in convergence of potential scale reduction factor (PSRF) to 1.0 for all the parameters, therefore, the analysis carried out using above mentioned models and default parameters was true and the bayesian inference could be accepted for the Indian PLCD complex DNA-A dataset. The PSRF convergence with range from 1.0 to 1.2 also suggest that the cladogram represents a clade credibility tree, whose branches lie in the region of 95% or more accumulated posterior probability. Therefore, the branch lengths represent the measured expected substitution per site that signifies the phylogenetic relatedness among all clades and sub-clades of the tree.

The MrBayes and PhyML tree resulted in a mid-rooted tree (Fig. 2 and 3) which was divided into two major clades and an out-group i.e. PLCD19 (Tomato leaf curl New Delhi virus). Other out-groups lying in the two major clades include PLCD13, PLCD14, PLCD34, PLCD35 and PLCD31 representing CLCuMuV, PaLCuV [India:New Delhi:tomato:2005], PaYLCuV [DP2], ToLCuV [C1], PapayaLCV [A-87] respectively. The first major clade consists of Papaya leaf crumple virus isolates within branch lengths 0.1 to 0.17 signifying low expected substitution rates per site. Despite having narrow branch length, notably, this virus species used only *A. paniculata* as its reservoir host, which is an annual weed widely found in south India, whereas tomato and papaya as a primary host in north India. This indicates that this is a new virus that has been introduced in northern regions via anthropogenic activities. During its course for survival, the Papaya leaf crumple virus isolates have adapted tomato and papaya as its primary host and undergone parallel evolution with Papaya leaf curl virus group in the northern regions. Its recombination analysis can provide much more details about its origin and propagation pattern.

The second major clade is further subdivided into two minor sub-clades i.e. ChiLCV and PaLCuV. The ChiLCV clade consists of ChiLCV and CLCuBuV [LK_2N] whereas PaLCuV clade clustered the papaya and associated crop infecting PaLCuV isolates according to their hosts. The ChiLCV clade has a narrow branch length range i.e. 0.24 - 0.29 signifying low substitution rate per site. Therefore, whole ChiLCV clade shows close genetic relatedness and host singularity as evident from *C. papaya* being the only host crop infected by members of this clade. The pattern suggests that these clade viruses have evolved to infect papaya crops lying nearby

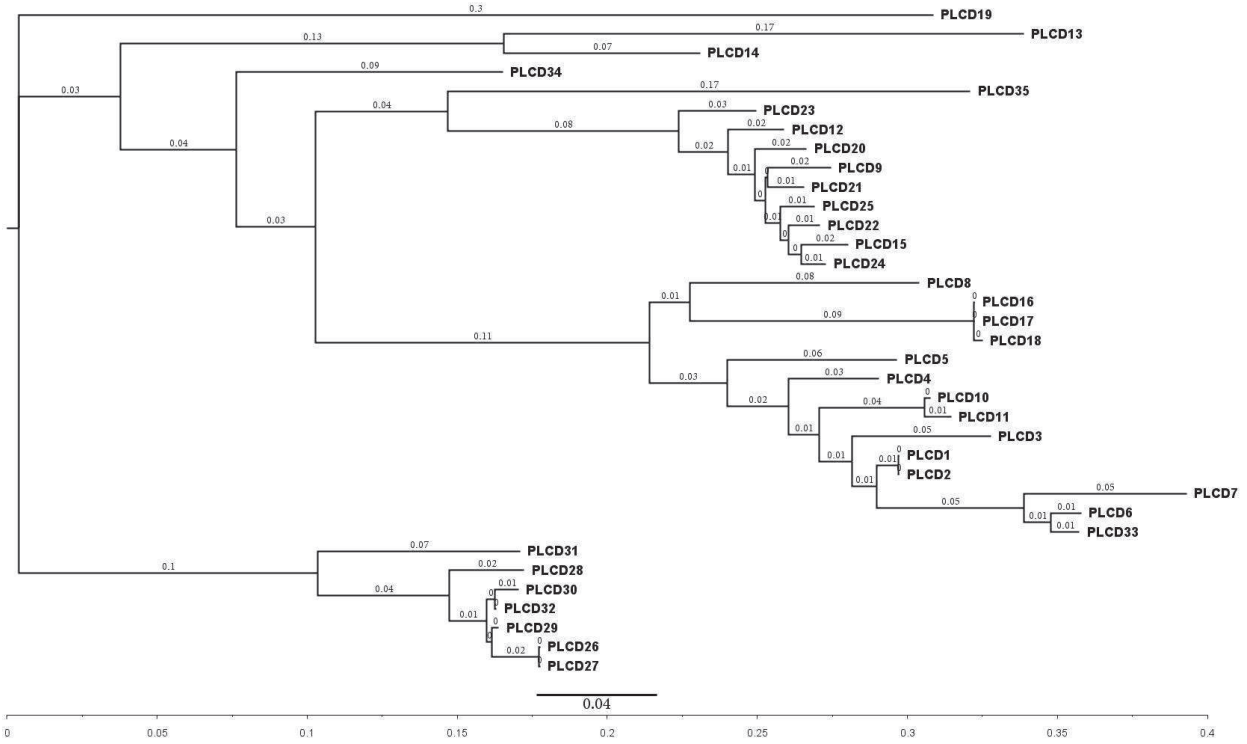


Fig. 2: MrBayes tree analysis. Phylogenetic tree based on the complete DNA-A sequences of the begomoviruses detected in *C. papaya* and additional sequences from associated crops and weeds infecting begomoviruses causing leaf curl disease. The tree was constructed by Bayesian inference using the GTR-G nucleotide substitution model available in MrBayes v3.1.1 method in TOPALi v2.5. Numbers at the branches indicate estimated clade credibility score inferred from statistical analysis of posterior probabilities assessed for the substitution model

chili fields in north Indian region and they are under very low mutational pressure to evolve and expand their host range. The PaLCuV clade on the other hand, has a very broad range of branch length i.e. 0.23-0.39 signifying high substitution rate per site; showing expansive genetic relation and diverse host range. Analysis also suggests that they might have evolved independently on diverse hosts in the geographically different regions. The out-group species are sparingly related to the clade species and have acquired the ability to infect papaya and associated host crops and weeds.

3.3. Species diversity identification by pairwise sequence comparison

The International Council for Taxonomy of Viruses (ICTV) has decided to solve the species nomenclature issues by using pairwise sequence identity criterion as a standard methodology for species demarcation. Earlier cutoff i.e., >89% sequence similarity was elevated to >91% and use of SDT tool was made mandatory for naming of new viral species (Brown *et al.,* 2015). Therefore, this tool was employed to study if the taxonomic species associated with Indian PLCD complex are really as diverse as indicated by the phylogenetic analysis.

The pairwise matrix (Fig. 4) obtained by using SDT v1.4 shows a similarity matrix characteristic of a complex association as the minimum and maximum similarity obtained was >75% and 100%. The species with highest similarity group i.e. >90% to 100% were grouped together using inbuilt function of the software. The PLCD complex thus got distributed into 11 different groups (Table 2). The groups labeled as Papaya leaf curl virus group, Chili leaf curl virus group, Papaya leaf crumple virus group and Papaya leaf curl virus (Pratapgarh) consist of 10, 9, 6 and 3 isolates respectively. Among all 35 PLCD complex isolates, 28 belonged to above mentioned groups and rest of the DNA-A sequences were phylogenetically isolated as individual species associated with Indian PLCD complex. Though, this grouping based on pairwise similarity is useful in inferring a good correlation with phylogenetic analysis done using MrBayes and PhyML-aLRT algorithms, still, it would be too premature to declare them similar or different species due to a significant overall similarity score in the range of >70%

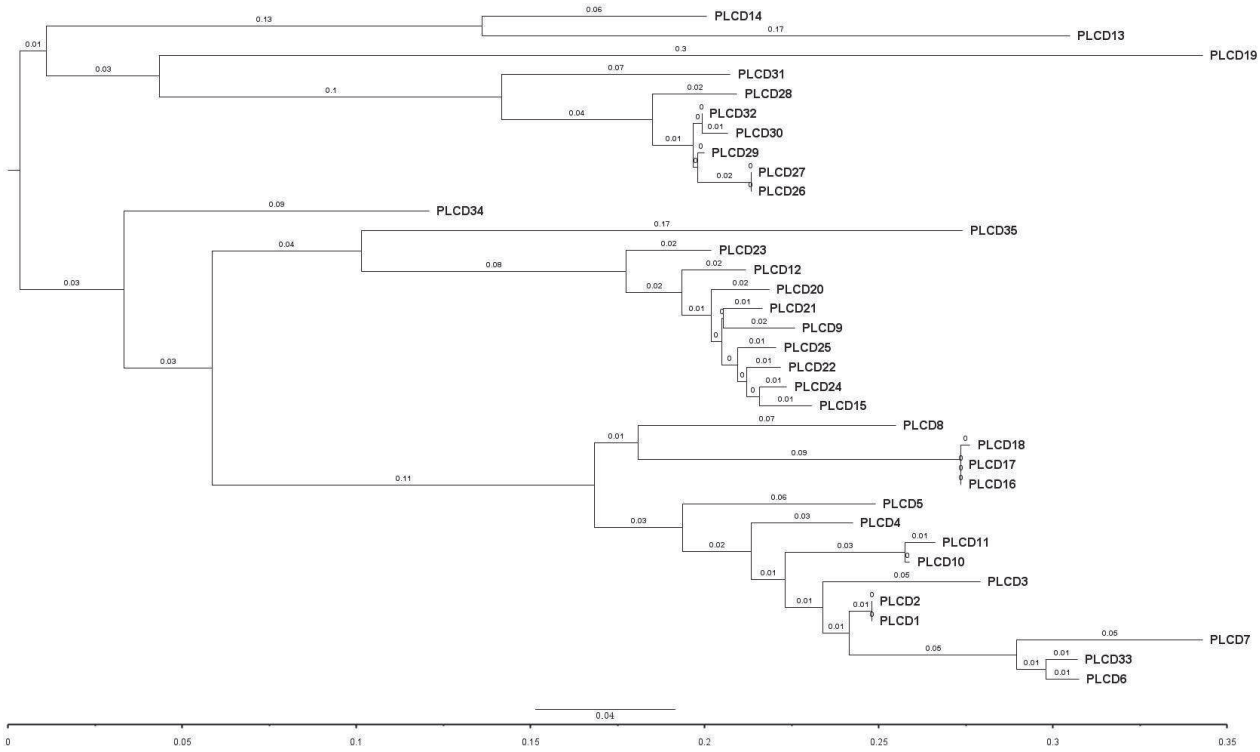


Fig. 3: PhyML tree analysis. A maximum likelihood tree based upon complete DNA-A sequences of the begomoviruses detected in *C. papaya* and additional sequences from associated crops and weeds infecting begomoviruses causing leaf curl disease. The tree was constructed by using PhyML-aLRT v2.4.5 tool in TOPALi v2.5 and the HKY-G nucleotide substitution model. Numbers at the branches indicate estimated substitution rate per nucleotide position

to 80%. Therefore, it was important to study the recombination pattern of the DNA-A sequences involved in the Indian PLCD complex to solve the complex nature of PLCD group of begomoviruses.

Pairwise sequence demarcation was carried out using SDT v1.2 program. The groups have been distributed according to the >90 to 100% sequence similarity with each other.

3.4. Recombination pattern of Indian PLCD complex

Recombination is an important molecular event that leads to translocation, transversion, addition, deletion or frame-shift mutations in a genome (Lima *et al.,* 2017). Due to the rolling circle replication mechanism in begomoviruses, these organisms have evolved to survive inside their hosts through acquisition of useful regions from other viruses or the host themselves, so that they can successfully evade host mediated RNA interference based immune response (van der Walt *et al.,* 2009). In the above analysis, it was observed that the Indian PLCD complex is well defined with major contribution made by PaLCuV, ChiLCV, PapayaLCV and discreet reports of other species infecting papaya and associated crops and weeds. To study if recombination is responsible for occurrence of such complex virus-host interaction; the DNA-A sequences of Indian PLCD complex were subjected to recombination analysis using RDP v4.95 (Martin and Rybicki, 2000) using available recombination analysis tools (Fig. 5) and default parameters as recommended by RDP4 manual (Smith, 1992; Padidam *et al.,* 1999; Gibbs *et al.,* 2000; Posada and Carandall, 2001; Martin *et al.,* 2005; Boni *et al.,* 2007). The analysis ended up predicting 66 recombination events, which further reduced to 12 events after careful manual screening as per the RDP4 manual (Martin *et al.,* 2015) (Table 3).

According to the manually screened recombination events results, event no. 10 i.e. PLCD9 and PLCD31 (inferred closest to the unknown minor parent) seems to be the major contributor of fragments giving a chance for origin of a new recombinant species e.g., PLCD8 and thirteen others i.e. PLCD1, PLCD2, PLCD3, PLCD4, PLCD5, PLCD6, PLCD7, PLCD10,

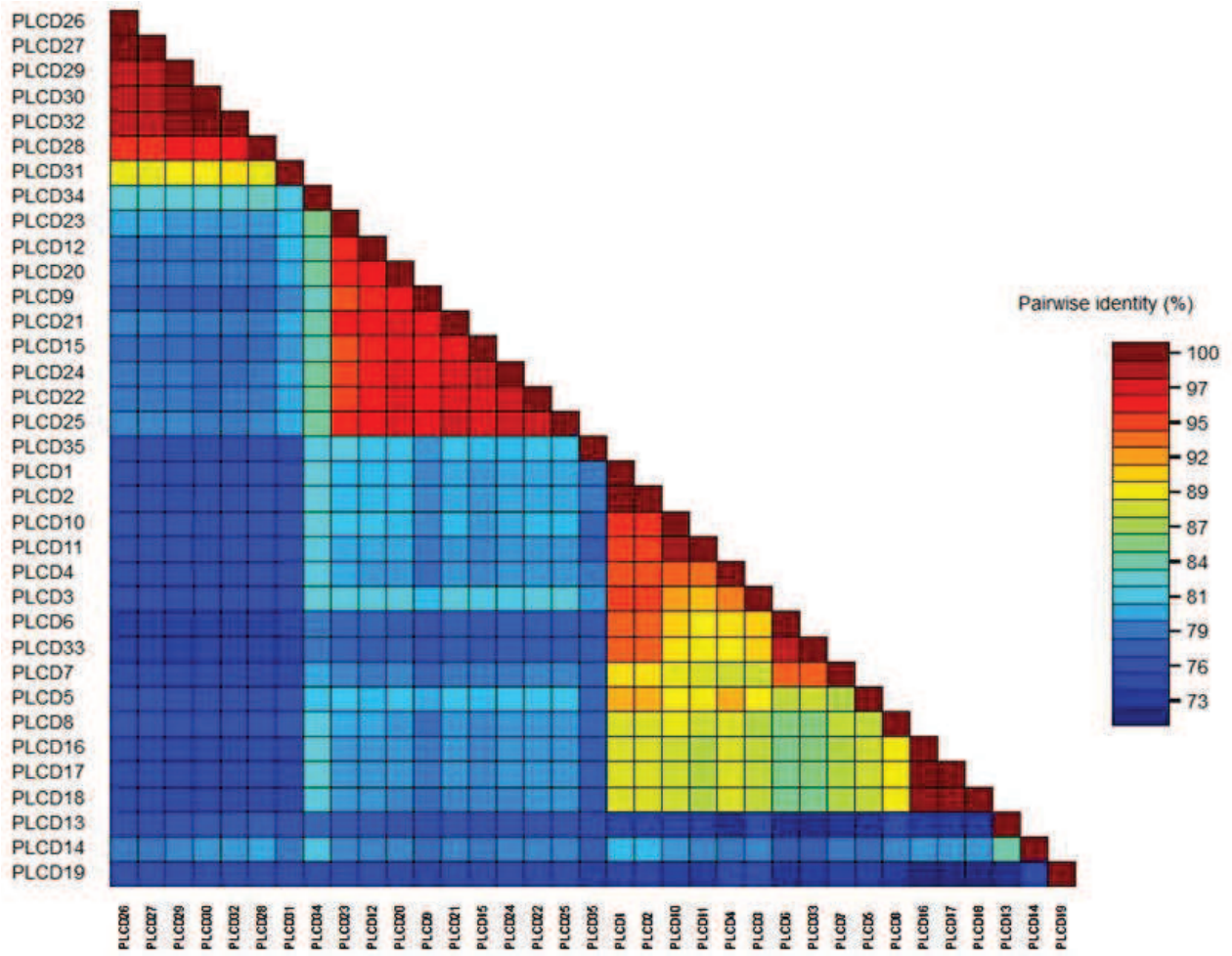


Fig. 4: Estimation of pairwise sequence similarity using Sequence Demarcation Tool v1.2. The matrix output indicates the diversity of begomoviruses infecting papaya and associated crops and weeds. The color key indicates highest and least similarity with brown and blue color respectively

Table 2: Pair wise sequence similarity among begomovirus components in PLCD complex

Group No.	No. of species	PLCD complex	Virus Taxons included in group
1	6	PLCD26; PLCD27; PLCD28; PLCD29; PLCD30; PLCD32	Papaya leaf crumple virus
2	1	PLCD31	Papaya leaf crumple virus (A-87)
3	1	PLCD34	Papaya yellow leaf curl virus (DP2)
4	9	PLCD9; PLCD12; PLCD21 PLCD22; PLCD23;PLCD24; PLCD25	Papaya leaf curl virus (New Delhi/2016) Chili leaf curl virus (Amritsar/2009)
5	1	PLCD35	Tomato leaf curl virus (C1)
6	10	PLCD1; PLCD2; PLCD3; PLCD4; PLCD5; PLCD6; PLCD7; PLCD10; PLCD11; PLCD33	Papaya leaf curl virus
7	1	PLCD8	Papaya leaf curl virus (WB2&5)
8	3	PLCD16; PLCD17; PLCD18	Papaya leaf curl virus (Pratapgarh)
9	1	PLCD13	Cotton leaf curl Multan virus
10	1	PLCD14	Papaya leaf curl virus (New Delhi/2005)
11	1	PLCD19	Tomato leaf curl New Delhi virus

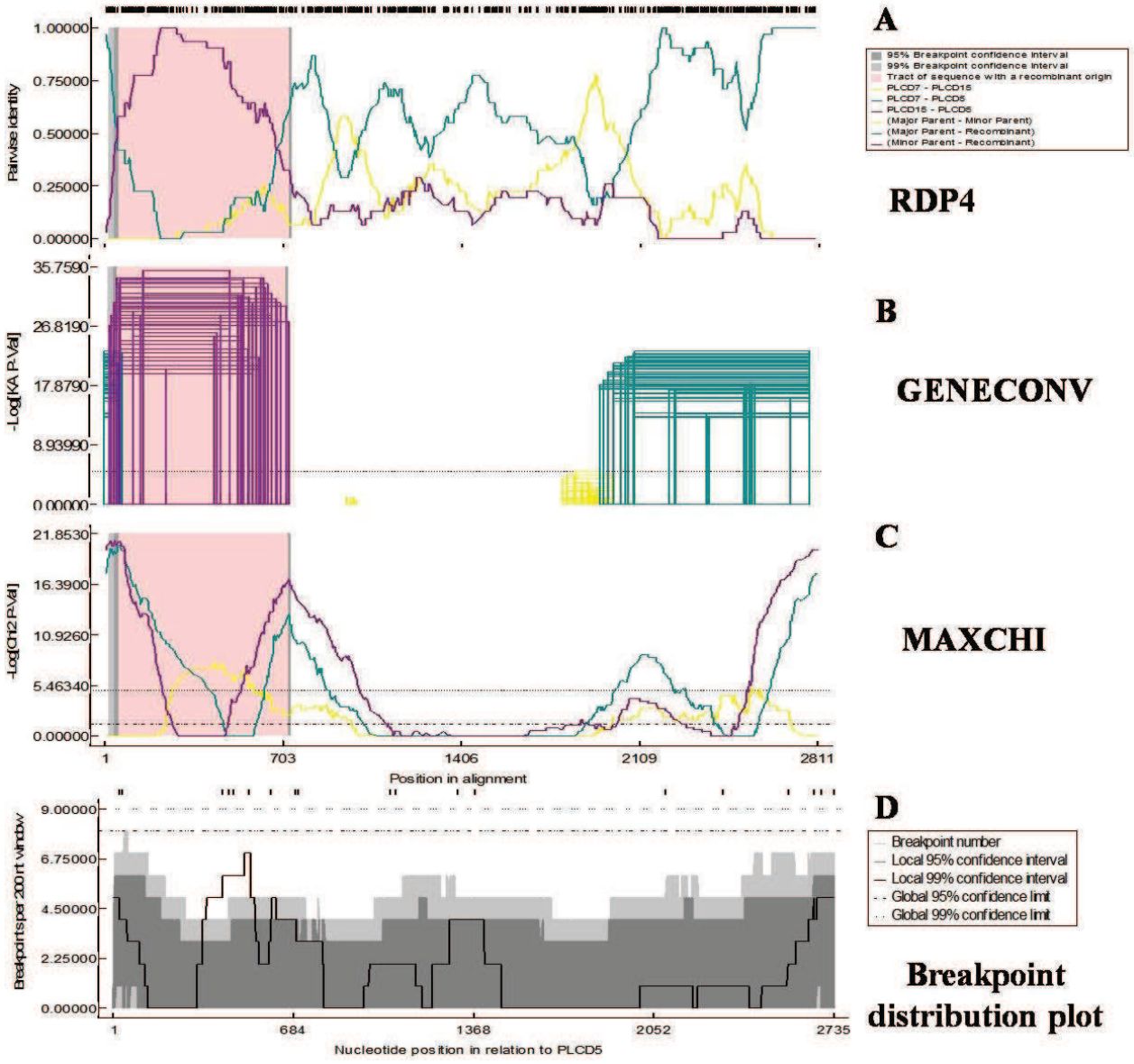


Fig. 5: Recombination analysis showing results of event number 1 using Recombination Detection Program v4.95 (RDP v4.95). The pink region indicates the occurrence of the recombination event with intersection points in the case of RDP and overlapping peaks in case of MAXCHI on the plot as the recombination starting and ending breakpoints. The grey region at the ends represents 99% and 95% confidence interval of prediction of breakpoints as mentioned in the key. A. RDP output; B. GENECONV output; C. MAXCHI output and D. Recombination breakpoint distribution plot: The light and dark grey regions in the plot represent the 99% and 95% confidence limit respectively for a 200nt scanning window. The breakpoint distribution was assessed with PLCD5 DNA-A sequence as a reference sequence just for qualitative purpose

PLCD16, PLCD17, PLCD18, PLCD23 and PLCD33. Thus, a Chili leaf curl virus and Papaya leaf crumple virus infecting papaya and *A. paniculata* contributed significantly to the Papaya leaf curl virus sub-clade, which partially explains the clades affinity to infect weeds and associated crops. The event no. 5, PLCD2 and PLCD28 (inferred closest to the unknown minor parent) give rise to PLCD33, PLCD6 and PLCD7 (already contains fragments from event no. 10) resulting into a *G. max*, Croton and *S. lycopersicum* infecting Papaya leaf curl virus isolates. Similarly, PLCD5 became a weed infecting recombinant after acquiring fragments

Table 3: Recombination analysis of Indian PLCD complex using RDPv4.95 program

Event No.	Recombinant	Major parent	Minor Parent	Frequency of recombination event	Detection Method						
					R	G	B	M	C	S	T
1	PLCD5	PLCD7	PLCD15	1	+	+	+	+	+	+	+
2	PLCD14	PLCD13	PLCD8	1	+	+	+	+	+	+	+
3	PLCD35	PLCD8*	PLCD3	1	+	+	+	+	+	+	+
4	PLCD7	PLCD10*	PLCD6	1	+	+	+	+	+	-	+
5	PLCD33	PLCD2	PLCD28*	3	+	+	+	+	+	+	+
6	PLCD13	PLCD2*	PLCD28	2	+	+	+	+	+	+	+
7	PLCD19	PLCD29	PLCD8	1	+	+	+	+	+	+	+
8	PLCD26	PLCD2	PLCD7	2	+	+	-	+	+	-	+
9	PLCD3	PLCD4	PLCD28*	1	+	+	+	+	+	+	+
10	PLCD8	PLCD9	PLCD31*	14	+	+	+	+	+	+	+
11	PLCD23	PLCD1	PLCD28	1	+	+	+	+	+	+	+
12	PLCD34	PLCD28	PLCD12	1	+	+	+	+	+	+	-
Total=12	Total frequency = 29										

*Unknown parent detected in recombination analysis but the specified parent was the closest inferred sequence in this analysis.
The analysis was performed against DNA-A sequences of 35 begomoviruses using default parameters and following algorithms. R=RDP; G=GENCONV; B=BOOTSCAN; M=MAXCHI; C=CHIMAERA; S=SISCAN; T=3SEQ

through event no. 1 i.e. Papaya leaf curl virus infecting *S. lycopersicum* and Chili leaf curl virus infecting papaya. Overall, the Indian PLCD complex is a repertoire of recombinants as evident from the recombination analysis and as the level of recombination complexity increases, the host range of recombinant virus also expands.

4. Conclusion

The PLCD complex (Fig. 1) comprises of genetically diverse group of begomoviruses as observed from their distinct phylogenetic patterns, pairwise sequence similarity and recombination analysis in this study. The diverse range of hosts mentioned in the study could be well attributed to putative recombination events between begomovirus species and strains infecting different host crops or weeds. Hence, in case of high genetic variability, the symptoms should only be used as a parameter for estimation of severity index while studying disease impact. The overall, >75% sequence similarity among begomoviruses of PLCD complex (Fig. 4) in this study signifies evidence of some sequence conservation that could be used to design siRNA based strategy.

The high frequency recombination fragments have prospects to be targeted for siRNA based strategy until and unless the target regions lie within the borders of a recombination hotspots i.e. a region most amenable to undergo complete transposition under normal recombination circumstances. Hence, the recombination hotspots despite being genetically mobile fragments can be employed as regions contributing to present and future siRNA strategies. Similar study can be employed to assess disease complexes such as Cassava mosaic disease (Pita *et al.,* 2001), Cotton leaf curl disease (Farooq *et al*., 2011) and Tomato leaf curl disease (Moriones *et al.,* 2017) complex widespread in central and south-east Asia, Africa and American continents. Therefore, the recombination patterns and sequence conservation trends of similar disease complexes need careful investigation to develop an efficient and sustainable strategy against begomovirus disease complexes in future.

Acknowledgment

SS and SV acknowledge Babasaheb Bhimrao Ambedkar University, Lucknow for providing the necessary infrastructure required to carry out this research. SV acknowledges Babasaheb Bhimrao Ambedkar University, Lucknow for fellowship.

References

Boni, M.F., Posada, D. and Feldman, M.W. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **176**:1035-1047.

Borah, B.K. and Dasgupta, I. 2012. Begomovirus research in India: A critical appraisal and the way ahead. *Journal of Biosciences* **37**:791-806.

Brown, J.K., Frohlich, D.R. and Rosell, R.C. 1995. The sweetpotato or silverleaf whiteflies: Biotypes of *Bemisia tabaci* or a species complex? *Annual Reviews of Entomology* **40**:511-534.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C.F. and Fiallo-Olive, E. 2015. Revision of begomovirus taxonomy based on pairwise sequence comparisons. *Archives of Virology* **160**:1593-1619.

David, R.J. 2003. Plant viruses transmitted by whiteflies. *European Journal of Plant Pathology* **109**:195-219.

Farooq, A., Farooq, J., Mahmood, A., Shakeel, A., Rehman, A., Batool, A., et al. 2011. An overview of cotton leaf curl virus disease (CLCuD) a serious threat to cotton productivity. *Australian Journal of Crop Science* **5**:1823-1831.

Fondong, V.N. 2013. Geminivirus protein structure and function. *Molecular Plant Pathology*. **14**:635-649.

Gamerman, D. and H. F. Lopes. 2006. MarkovChain Monte Carlo: Stochastic Simulation for Bayesian Inference. 2nd ed. Chapman and Hall, London, pp. 342.

Ghosal, S., Das, S., and Chakrabarti, J. 2012. Chapter 11: Computational approaches for designing efficient and specific siRNAs. In: Pérez-Sánchez, H. (Ed.) *Bioinformatics*, InTech open, Croatia-EUROPEAN UNION, pp. 261-276.

Gibbs, M.J., Armstrong, J.S. and Gibbs, A.J. 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* **16**:573-582.

Gutierrez, C. 2000. Geminiviruses and the plant cell cycle. *Plant Molecular Biology* **43**:763-772.

Guindon, S. and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**:696-704.

Harrison, B.D. 1985. Advances in geminivirus research. *Annual Reviews of Phytopathology* **23**:55-82.

Hallan, V., Saxena, S. and Singh, B.P. 1998a. Ageratum, Croton and Malvastrum harbor geminiviruses: evidence through PCR amplification. *World Journal of Microbiology and Biotechnology* **14**:931-932.

Hallan, V., Saxena, S. and Singh, B.P. 1998b. Yellow net of *Triumfetta* is caused by a Geminivirus: A first report. *Plant Disease* **82**:127-127.

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D. and Mansoor, S. 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Reviews in Microbiology* **11**:777-788.

Hasegawa, M., Kishino, H. and Yano, T. 1985. Dating of human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**:160-174.

Huelsenbeck, J.P., Ronquist, F. 2001. MrBayes: Bayesian inference of phylogeny. *Biometrics* **17**:754-755.

King, A.M.Q., Adams, M., Eric, B.C. and Lefkowitz, E.J. 2011. Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier. San Diego, CA, pp. 1339.

Larget, B. and Simon, D. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution* **16**:750-759.

Lima, A.T.M., Silva, J.C.F., Silva, F.N., Castillo-Urquiza, G.P., Silva, F.F. and Seah, Y.M. 2017. The diversification of begomovirus populations is predominantly driven by mutational dynamics. *Virus Evolution* **3**:1-14.

Martin, D. and Rybicki, E. 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**:562-563.

Martin, D.P., Posada, D., Crandall, K.A. and Williamson, C. 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Research and Human Retroviruses* **21**:98-102.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**:1-5.

Moriones, E., Praveen, S. and Chakraborty, S. 2017. Tomato Leaf Curl New Delhi Virus: An emerging virus complex threatening vegetable and fiber crops. *Viruses* **9**:e264.

Nadeem, A., Mehmood, T., Tahir, M., Khalid, S. and Xion, Z.1997. First report of Papaya Leaf Curl Disease in Pakistan. *Plant Disease* **81**:1333-1333.

Padidam, M., Sawyer, S. and Fauquet, C.M. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**:218-225.

Pita, J.S., Fondong, V.N., Sangare, A., Ogwal, S. and Fauquet, C.M. 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology* **82**:655-665.

Posada, D. and Crandall, K.A. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proceedings of National Academy of Science USA* **98**:13757-13762.

Raj, S.K., Snehi, S.K., Khan, M.S., Singh, R. and Khan, A.A. 2008. Molecular evidence for association of Tomato leaf curl New Delhi virus with leaf curl disease of papaya (*Carica papaya L.*) in India. *Australasian Plant Disease Notes* **3**:152-155.

Roden, L.C. and Ingle, R.A. 2009. Lights, Rhythms, Infection: The role of light and the circadian clock in determining the outcome of plant-pathogen interactions. *Plant Cell* **21**:2546-2552.

Ronquist, F. and Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**:1572-1574.

Saxena, S., Hallan, V., Singh, B.P. and Sane, P.V. 1998a. Leaf Curl Disease of *Carica papaya* from India may be caused by a bipartite geminivirus. *Plant Disease* **82**:126-126.

Saxena, S., Hallan, V., Singh, B.P. and Sane, P.V. 1998b. Nucleotide sequence and inter-geminiviral homologies of the DNA-A of papaya leaf curl Geminivirus from India. *Biochemistry and Molecular Biology International* **45**:101-113.

Saxena, S., Hallan, V., Singh, B.P. and Sane, P.V. 1998c. Evidence from nucleic acid hybridization test for a Geminivirus infection causing leaf curl disease of papaya in India.

Indian Journal of Experimental Biology **36**:229-232.

Saxena, S., Kesarwani, R.K., Singh, V. and Singh, S. 2013. Designing of putative siRNA against geminiviral suppressors of RNAi to develop Geminivirus resistant Papaya crop. *International Journal of Bioinformatics Research and Applications* **9**:3-12.

Saxena, S., Singh, N., Ranade, S.A. and Babu, G.S. 2011. Strategy for a generic resistance to Geminivirus infecting papaya and tomato through *in-silico* siRNA search. *Virus Genes* **43**:409-434.

Saxena, S., Singh, V.K. and Verma, S. 2016. PCR mediated detection of sex and PaLCuV infection in papaya: A review. *Journal of Applied Horticulture* **18**:80-84.

Saxena, S. and Verma, S. 2016. Chapter 20: Harnessing the genetic variability in Plant-Virus-Vector complex interaction in Begomovirus family to prevent viral diseases. In: Sobti, R.C., Mishra, S. and Jaiswal, K. (Eds.) *Recent Advances in Applied Biosciences*, Narendra Publishing House, New Delhi, pp. 173-183.

Singh, P., Mishra, A.K. and Tripathi, N.N. 2012. Assessment of mycoflora associated with postharvest losses of papaya fruits. *Journal of Agricultural Technology* **8**:961-968.

Singh-Pant, P., Pant, P. and Mazumdar-Leighton, S. 2012. Spatial and temporal diversity of begomoviral complexes in papayas with leaf curl disease. *Archives of Virology* **157**:1217-1232.

Smith, J.M. 1992. Analyzing the mosaic structure of genes. *Journal of Molecular Evolution* **34**:126-129.

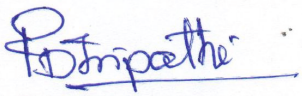
Varma, A. and Malathi, V.G. 2003. Emerging geminivirus problems: A serious threat to crop production. *Annals of Applied Biology* **142**:145-64.

Varun, P., Ranade, S.A. and Saxena, S. 2017. A molecular insight into papaya leaf curl—a severe viral disease. *Protoplasma* doi: 10.1007/s00709-017-1126-8.

van der Walt, E., Rybicki, E.P., Varsani, A., Polston, J.E., Billharz, R., Donaldson, L., et al. 2009. Rapid host adaptation by extensive recombination. *Journal of General Virology* **90**:734-746.

Zhou, X. 2013. Advances in understanding Begomovirus Satellites. *Annual Reviews of Phytopathology* **51**:357-381.

Sub.: International Journal of Plant and Environment (IJPE): Decision letter for IJPE/2018/74

Name of the Journal	International Journal of Plant and Environment
ISSN No.	ISSN No.: 2454-1117 (Print) ISSN No.: 2455-202X (Online)
UGC Journal List No.	43696
Manuscript No.	**IJPE/2018/74**
Title	***In-Silico* Investigation of Betasatellite Complexity in Papaya Leaf Curl Disease Complex**
Authors	Saurabh Verma and Sangeeta Saxena
Corresponding Authors	**Mr. Saurabh Verma and Dr. Sangeeta Saxena**
Date of acceptance	09 March, 2018
Volume, Issue, Year	Volume: 4; Issue: 2; Month & Year: July, 2018
From:	ijpeeditors@gmail.com
To:	saddybiot@gmail.com , dr_sangeeta_saxena@yahoo.com
Subject	International Journal of Plant and Environment (IJPE): Decision letter for MS No. **IJPE/2018/74**
	Dear Mr. Saurabh Verma and Dr. Sangeeta Saxena, We are pleased to accept your manuscript entitled "***In-Silico* Investigation of Betasatellite Complexity in Papaya Leaf Curl Disease Complex**" for publication in International Journal of Plant and Environment (IJPE). Thank you for choosing to publish your work with us. We will send you the copy right / author agreement form within a week time. To view the contents of previous volume of IJPE please click: **https://www.myresearchjournals.com/index.php/IJPEn/issue/view/990** For visit our website (ISEB, CSIR-NBRI, Lucknow) please click: **http://www.isebindia.com/** Thanking you and best regards,  R.D. Tripathi, FNASc Emeritus Scientist (CSIR) & Emeritus Professor (AcSIR) Chief Editor, International Journal of Plant and Environment Plant Ecology & Environmental Science Division CSIR-National Botanical Research Institute Lucknow-226001, India Tel: 91-522-2297825 Ext: 825 (O); 0522-2721568 (R) Mobile: ++91-7376679779 Email: ijpeeditors@gmail.com; tripathird@gmail.com

In-silico investigation of betasatellite complexity in Papaya leaf curl disease complex

Saurabh Verma* and Sangeeta Saxena

Department of Biotechnology, Babasaheb Bhimrao Ambedkar University, Lucknow-226024, INDIA

*Author for correspondence: saddybiot@gmail.com

Mobile No.: +91-9565427565

Running title: Betasatellite molecules and PLCD complex

Abstract

Papaya leaf curl disease complex (PLCD) impose heavy losses to papaya farmers across India. It is mainly reported to be caused by *Papaya leaf curl virus* and associated begomoviruses carrying a DNA-A and betasatellite molecules. Both components are required for infliction of severe symptoms. The investigation into molecular diversity is necessary to devise intervention techniques against these begomoviruses. Based on the outcome of diversity, preferred resistance strategy against a disease complex should be generic in nature, which provides an advantage of a broad range resistance against a variety of plant virus causing same disease symptoms. For this purpose, we studied the molecular diversity through investigation of genetic complexity of the betasatellite genome using various computational methods and probed genetic complexity of betasatellite component. Multiple sequence alignment and recombination analysis were performed to understand the evolutionary relationship of betasatellite components with a DNA-B occurring as additional genome in case of bipartite begomoviruses. The overall results indicate that the betasatellite molecules have evolved independently of DNA-A component and their

association is a result of co-infection and due to inter- and intraspecific interaction with various forms of virus infection in plants. The information generated through this study has potential application in designing intervention strategies against PLCD, which is widespread in Indian sub-continent region that includes countries like India, Pakistan, Nepal and Bangladesh.

Keywords: *Begomovirus, Betasatellite, generic resistance, papaya leaf curl disease complex, Recombination*

1. Introduction

Papaya, is grown mainly in tropical and sub-tropical regions of Indian sub-continent spanning from Pakistan in west to Bangladesh in east and adjoining north east and south east nations such as Thailand, Malaysia, Maldives and Indonesian group of islands (Mishra *et al.*, 2007). Farmers in these areas grow papaya due to favorable environmental conditions, large scale consumption and high net income from total productivity. But, papaya plant is also amenable to tough climate, exposing it to abiotic conditions such as cold, drought, flooding and osmotic stress (Jeyakumar *et al.*, 2007). Biotic stress upon papaya is one of the major concerns nowadays, especially disease caused by viruses have played an important destructive interference in total productivity in Indian sub-continent. Papaya is infected by several viruses such as begomoviruses which also infect crops such as chili, tomato, papaya, potato, cucurbits, etc. Recently, papaya crops in Brazil and Mexico were reported to be infected with, a sticky disease symptom named as Papaya Meleira virus, a new type of double stranded (ds)-RNA virus capable of causing complete loss in production (Antunes *et al.,* 2016). Papaya leaf curl disease complex has resulted in a complete wipe-out of papaya cultivation in northern regions of Indian sub-continent. It is reported to be caused by a group of single strand (ss)-DNA viruses, which are monopartite in nature and named

as *Papaya leaf curl virus* (PaLCuV) and other associated viruses as well (Borah and Dasgupta, 2012; Singh-Pant *et al.*, 2012). Recently, an *in-silico* study of papaya leaf curl disease genomic components has shown that this disease group is actually a complex, where each member virus has mutated and adapted to suit its host plant and geographical region (Verma and Saxena, 2017). Therefore, this disease is now called 'Papaya leaf curl disease complex' (PLCD) to accommodate the complexity in member composition and their genetic variability. Apart from DNA-A component of a monopartite begomovirus, the members of PLCD complex have been shown to be associated with betasatellite, which is a defective component containing partial sequence of DNA-A. This component contains 200bp long intergenic region, regions containing an open reading frame (ORF) on complementary-sense strand named as c1 ORF (βC1), iterons, promoter elements and a Adenine (A) nucleotide rich sequence (Zhou, 2013). This c1 ORF has been shown to be a pathogenicity determinant and responsible for physical deformation and cellular damage of infected leaves (Tahir and Mansoor, 2011; Bhattacharyya *et al.*, 2015). Therefore, due to this reason they are required for efficient infection in case of many monopartite begomovirus. PLCD complex has emerged as one such complex where 99% cases of symptoms and positive infection has been associated with the presence of these betasatellite molecules.

Betasatellite molecules genomic size is nearly half of DNA-A i.e. ~1.2-1.5kb. They are single stranded DNA particles and encapsidated within geminate particles (similar to other geminiviruses). Betasatellite molecules are reported to utilize Rep protein (DNA-A dependent synthesis) to initiate replication and transcription of βC1. Apart from being pathogenicity determinant, βC1 has also been shown to be helping in viral DNA accumulation, host RNAi suppression and helping in viral movement (Yang *et al.*, 2011; Jia *et al.*, 2016; Haxim *et al.*, 2017). Because of the above reasons, the role of bipartite begomoviruses have been slowly taken

over by monopartite begomoviruses populations in old world region. This phenomenon accounted for more than 250 betasatellite sequences being deposited in last 5-8 years. Such a large number of betasatellite components associated with DNA-A have given rise to several disease complexes in old world like Cotton leaf curl disease complex in cotton producing regions of Punjab area, spanning both India and Pakistan (Zubair, *et al.*, 2017a,b). Such complexes have given rise to many epidemics in past few years leading to large scale economic losses (Pita *et al.*, 2001; da Silva *et al.*, 2011; Moriones *et al.*, 2017; Sattar *et al.*, 2017). Therefore, this study was carried out to understand the genomic complexity of these betasatellites with respect to PLCD complex in Indian sub-continent. The genetic complexity arises due to genomic recombination and strand exchange phenomenon occurring due to interactions of two or more different types of viruses in same plant. This results in genetic exchange resulting in formation of a new viral particle capable of causing much more diversely severe symptoms in same or in new hosts (Lefeuvre *et al.*, 2009; Martin *et al.*, 2011; Pearson *et al.*, 2016).

2. Materials and Methods

2.1. Retrieval of genomic components associated with papaya leaf curl disease in Indian sub-continent from GenBank

The viral genome belonging to DNA-A component of PLCD complex were retrieved and annotated (Verma and Saxena, 2017). Betasatellite components were retrieved from National Centre for Biotechnology (NCBI) website (https://www.ncbi.nlm.nih.gov/) Nucleotide database (https://www.ncbi.nlm.nih.gov/ nuccore/). "Betasatellite" term was used to retrieve nucleotide sequences of interest. Annotation of the sequences in FASTA format was compiled for further analysis (Table 1).

2.2. Phylogenetic analysis of betasatellite components associated with leaf curl disease

MUSCLE algorithm integrated into MEGA 6.0 software package was used to perform sequence alignment of annotated betasatellite components associated with PLCD complex (Edgar, 2004). The default parameters were selected to obtain alignment with highest accuracy and refined position information (Hall, 2013). Phylogenetic analysis for elucidation of evolutionary information related to betasatellite sequences was done using Maximum Likelihood (ML) algorithm available in MEGA 6.0 software package. ML method based on the Tamura-Nei model was used to infer evolutionary history.

2.4. Recombination analysis

A recombination-based investigation was conducted to get insight into inter-sequence interaction of DNA-A and betasatellite components causing leaf curl disease of papaya and associated weeds and crops. Recombination Detection Program ver.4 (RDP4) (http://web.cbio.uct.ac.za/ ~darren/ rdp.html) was employed for such analysis. Various recombination detection modules are available in this package which independently analyze aligned set of sequences for recombination breakpoints, recombination boundaries and conclude with a p-value for each event, which predicts the probability of occurrence of that event (Martin *et al.*, 2015). The default parameters were used for analysis of 51 betasatellite sequences aligned against 35 DNA-A sequences (Verma and Saxena, 2017). Events predicted by more than 4 tools were accepted and rest were rejected. All predicted events were manually screened and accepted after careful investigation of phylogenetic and recombination data inferred by various tools in RDP4.

3. Results and Discussion

3.1. Betasatellite components of PLCD complex and other associated crops and weeds

The PLCD complex is mainly composed of monopartite begomoviruses, which are associated with either betasatellite or alphasatellite or both. Since, this complex is found to be closely associated with tomato, chili and weeds infecting monopartite begomoviruses, therefore, a comprehensive list of betasatellites associated with DNA-A of the above begomoviruses was compiled (Table 1). The betasatellite molecules associated with weeds i.e. *Ageratum leaf curl virus,* tomato i.e. *Tomato leaf curl virus, Tomato yellow leaf curl virus* and *Tomato leaf curl Joydebpur virus,* chili i.e. *Chili leaf curl virus,* cotton i.e. *Cotton leaf curl Burewala virus,* papaya i.e. *Papaya leaf curl virus* were included in the list of 51 betasatellite molecules. Among all the above considered in our study, 39 were from in India, 10 from Pakistan and 2 were obtained from a Chinese region. Overall, the group total of betasatellites infecting papaya, tomato, capsicum, guar and other crop were 10, 11, 9, 5 and 12 respectively, out of 51 mentioned in the list (Table 1). The betasatellite molecules associated with crops and weeds other than papaya were considered in this study to evaluate their diversity and recombination potential. This helps in minimizing interference with resistance strategy to be employed in future.

3.2. Genetic variability of betasatellite in PLCD complex

All 51 betasatellite molecules DNA sequences were retrieved from NCBI in FASTA format and annotated as TBS1-TBS51 according to their taxonomic placements. Total genome alignment was performed with MUSCLE algorithm in MEGA6 software suite. The aligned DNA sequences of betasatellite molecules were subjected to phylogenetic analysis and a Maximum Likelihood tree was obtained (Fig 1). The tree was clearly demarcated into two types of clusters viz. genetically related CLUSTER-1 (blue colored group) and genetically distant CLUSTER-2 (Red to magenta colored groups). Chinese taxon is illustrated separately in light blue color and

outgroups in black colored labels. CLUSTER-1 was further grouped into Group1-1 to Group1-5. The CLUSTER-2 was not sub-divided into sub-clades, thus, analyzed separately.

CLUSTER 1-1 comprises of betasatellite components infecting papaya in New Delhi and Panipat regions (India) and Punjab (Pakistan). CLUSTER 1-1 was found to be genetically very close i.e. phylogenetic likelihood predicted in 80-100% bootstrap steps. TBS7, 39, 40, 41, 42, 43, 44 and 45 were papaya-infecting betasatellites associated with Chili leaf curl and Tomato leaf curl disease causing DNA-A component. CLUSTER 1-2 consists of Chili leaf curl associated betasatellites infecting capsicum and tuber plants in Punjab (Pakistan), Palampur (India), Rajasthan (India) and Punjab (India) regions. CLUSTER 1-3 consists of tomato and guar plants in New Delhi, Haldwani in India and Bhakkar in Pakistan. These betasatellites were found to be associated with primarily Guar and Tomato leaf curl disease. CLUSTER 1-4 is a group of weed infecting betasatellite infecting weeds with Ageratum leaf curl disease reported mainly in Central parts of Uttar Pradesh and Rajasthan, India. CLUSTER 1-5 was found to be linked to Papaya leaf curl disease causing DNA-A components infecting guar and capsicum plants in western provinces of India and adjoining areas in Pakistan. TBS 13 is an outlier in this group belonging to betasatellite associated with Cotton leaf curl Burewala disease from Punjab (India) region.

CLUSTER 2 was observed to have complex clade pattern consisting of papaya leaf curl associated betasatellite components reported from various parts of India. The betasatellite components in this group were cladistically distant relatives as evident from low bootstrap predictions. These betasatellite components infect various crops such as papaya, tomato, cucurbits, weeds and other herbaceous plants growing in tropical regions of India and Pakistan. The papaya-infecting group consist of TBS15 and 16 while tomato was found to be infected by TBS24, 25, 32 and 36. A unique group of TBS21, 22 and 23 isolates reported from Coimbatore

and Pudukottai districts (Tamil Nadu, India), infect various gram (Cereal) varieties in that region. TBS18, 19, 20 and 28 were found to be infecting cucurbits and tuber varieties. The weed infecting group comprises of TBS27 and 29 infecting *Parthenium* specifically.

From above phylogeny of betasatellite molecules, it is clear that these components have evolved in a mixed environment, interacting with more than one type viral components. Due to this reason, these defective partial viral components i.e. betasatellites, have evolved to infect a variety of host plants. This is evident from the above cladistic analysis that hosts from different families have played vital part in evolution related to different betasatellite molecules, expanding their host range and helping in their persistent presence in the region throughout year and across seasons. The reason for their omnipresent nature needs deeper investigation using recombination analysis with DNA-A components.

3.4. Recombination pattern in betasatellite component of Asian PLCD complex

A recombination analysis was carried out to investigate the genomic breakpoints in begomoviral sequences, which provide an evidence of recombination pattern prevalent between the DNA-A and betasatellite components of PLCD and other leaf curl causing complex. DNA-A dataset of PLCD complex used to investigate recombination pattern within DNA-A components was aligned with betasatellite dataset to prepare input for RDP4 software based recombination analysis. The output from various algorithm was analyzed for authenticity of analysis for breakpoint distribution and recombinants using RDP, GENCONV, MAXCHI, CHIMAERA and analysis of breakpoint boundary distribution with reference to TBS1 genome sequence (Fig. 2).

The recombination analysis of betasatellite component did not predict any recombination boundary, which includes genomic fragments from DNA-A components of PLCD. However,

eleven recombination events were detected with statistical relevance and qualifying the criteria set initially (Table 2). In total, 50 recombination signals were predicted under 11 separate events. Event no. 4 was detected in 14 different recombinants namely, TBS29, 15, 16, 19, 20, 21, 22, 23, 24, 27, 28, 32, 36 and partially evident in TBS25. Similarly, event no. 10 was detected in 11 different recombinants namely, TBS 27, 15, 16, 19, 20, 24, 26, 28, 29, 36 and partially evident in TBS 25 and 32 (Table 2). Such a prominent recombination event prediction signifies that the major and minor parents *Papaya leaf curl virus betasatellite isolate [In:Var:Pum:08:1]* and *Tomato leaf curl betasatellite* isolate *Cluster bean*, clone *NGS-D1* respectively, were responsible for evolution of 15 putative recombinant betasatellite components reported to be associated with *Papaya leaf curl virus* DNA-A isolated from Indian tropical regions. Another pair of major and minor parents i.e. *Papaya leaf curl virus betasatellite clone BG-CBE* and *Cotton leaf curl Burewala betasatellite*, clone *L2-RCA-b1-F* respectively, have significant contribution in evolution of betasatellite components of above mentioned components of PLCD complex. Apart from PLCD complex, the above-mentioned parental pair has contributed in evolution of tomato leaf curl disease associated betasatellite components. Other predicted recombination events i.e. event no. 11 and 57 also gave rise to recombinants associated with tomato leaf curl disease complex (TLCD) (Moriones *et al.*, 2017).

Recombination events have been reported in many cases where more than one type of betasatellite was isolated and were found to cause mixed type of symptoms in sunn hemp (Kumar *et al.*, 2010) and tomato (Kumar *et al.*, 2013). Another case was reported where a bipartite begomovirus i.e. *Tomato leaf curl Gujarat virus* was found to be associated with betasatellite rather than DNA-B component in *Xanthium strumarium* (a weed), which acts as a reservoir for interspecific mixing of DNA-A and associated genomic components (Mubin *et al.*,

2012). These natural recombinants are proof of natural selection of factors that increase fitness of newer recombinant progenies as in betasatellite genomes. Experimental investigations have provided valuable insights into the selection of recombinant viruses which exhibit recovery of wild-type portions with enhanced fitness when compared with synthetic chimaeras, in case of mixed infections (van der Walt *et al.*, 2009). Recent studies have indicated that mutations have played a pivotal role in evolution of viruses, but recombination has important role in enhancing synergistic interactions in case of mixed infections (Pita *et al.*, 2001; Lima *et al.*, 2017). Such synergistic adaptations often are transferred in more than one progeny, which further propagates into several other recombinant betasatellite genomes, thus, giving rise to recombination 'hot-spots' (Lefeuvre *et al.*, 2007; Paul *et al.*, 2016).

4. Conclusion

In this study, a significant finding that the betasatellite components associated with leaf curl disease complex of papaya and tomato have actually evolved through recombination phenomenon sheds light on the much more complex interaction among its members. This recombination has enabled these betasatellite components to evolve and infect different crops in varied tropical regions of Indian sub-continent. The broad host spectrum of PLCD and TLCD could be attributed to this phenomenon, therefore, the resistance strategies against these diseases have failed in past and are potentially prone to failure in near future. This study also explains the occurrence of interspecies infections i.e. association of Papaya betasatellite with TLCD and Tomato betasatellite with PLCD. The two disease complex make up for the large proportion of leaf curl disease in Indian sub-continent, caused by a monopartite ss-DNA viruses. Therefore, betasatellite recombination hot-spots should be identified and future siRNA based strategies

should be based upon these regions along-with the DNA-A component associated with respective monopartite begomoviruses.

Acknowledgment

SS and SV acknowledges Babasaheb Bhimrao Ambedkar University, Lucknow for providing the necessary infrastructure to carry out this research. SV acknowledges Babasaheb Bhimrao Ambedkar University, Lucknow for fellowship.

References

Antunes, T.F.S, Amaral, R.J.V., Ventura, J.A. *et al.* 2016. The dsRNA virus papaya meleira virus and an ssRNA virus are associated with papaya sticky disease. *PLoS ONE* **11**, e0155240.

Bhattacharyya, D., Gnanasekaran, P., Kumar, R.K., *et al.* 2015. A geminivirus betasatellite damages the structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. *Journal of Experimental Botany* **66**:279-281.

Borah, B.K. and Dasgupta, I. 2012. Begomovirus research in India: A critical appraisal and the way ahead. *Journal of Biosciences* **37**:791-806.

da Silva, S.J.C., Castillo-Urquiza, G.P., Hora Júnior, B.T., *et al.* 2011. High genetic variability and recombination in a begomovirus population infecting the ubiquitous weed Cleome affinis in northeastern Brazil. *Archives of Virology* **156**:279-281.

Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:279-281.

Hall, B.G. 2013. Building phylogenetic trees from molecular data with MEGA. *Molecular Biology and Evolution* **30**:279-281.

Haxim, Y., Ismayil, A., Jia, Q., *et al.* 2017. Autophagy functions as an antiviral mechanism against geminiviruses in plants. *eLife* **6**:279-281.

Jeyakumar, P., Kavino, M., Kumar, N. and Soorianathasundaram, K. 2007. Physiological performance of papaya cultivars under abiotic stress conditions. *Acta Horticulturae* 209–216.

Jia, Q., Liu, N., Xie, K., Dai, Y., Han, S., Zhao, X., et al. 2016. CLCuMuB βC1 Subverts Ubiquitination by Interacting with NbSKP1s to Enhance Geminivirus Infection in Nicotiana benthamiana. *PLoS Pathogen.* **12**:1–30.

Kumar, A., J. Kumar, Z. A. Khan, N. Yadav, V. Sinha, D. Bhatnagar, and J. A. Khan. 2010. Study of Betasatellite Molecule from Leaf Curl Disease of Sunn Hemp (*Crotalaria juncea*) in India. *Virus Genes* **41**:432–40.

Kumar, J., Singh, S. P., Kumar, A., Khan, J. A., and Tuli, R. 2013. Detection and characterization of a new betasatellite: Variation in disease symptoms of tomato leaf curl Pakistan virus-India due to associated betasatellite. *Archives of Virology* **158**:257–261.

Lefeuvre, P., Lett, J.M., Varsani, A. and Martin, D.P. 2009. Widely Conserved Recombination Patterns among Single-Stranded DNA Viruses. *Journal of Virology* **83**:279-281.

Lefeuvre, P., Martin, D. P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., et al. 2007. Begomovirus "melting pot" in the south-west Indian Ocean islands: Molecular diversity and evolution through recombination. *Journal of General Virology* **88**:3458–3468.

Lima, A. T. M., Silva, J. C. F., Silva, F. N., Castillo-Urquiza, G. P., Silva, F. F., Seah, Y. M., et al. 2017. The diversification of begomovirus populations is predominantly driven by mutational dynamics. *Virus Evolution* **3**.

Martin, D.P., Lefeuvre, P., Varsani, A., *et al.* 2011. Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. *PLoS Pathogens* **7**:279-281.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**:1-5.

Mishra, M., Chandra, R. and Saxena, S. 2007. Papaya. In: Kole C, ed. *Fruits and Nuts*. Berlin, Heidelberg: Springer Berlin Heidelberg, 343–351.

Moriones, E., Praveen, S. and Chakraborty, S. 2017. Tomato Leaf Curl New Delhi Virus: An emerging virus complex threatening vegetable and fiber crops. *Viruses* **9**:e264.

Mubin, M., Akhtar, S., Amin, I., Briddon, R. W., and Mansoor, S. 2012. Xanthium strumarium: A weed host of components of begomovirus-betasatellite complexes affecting crops. *Virus Genes* **44**:112–119.

Paul, P., Nag, D., and Chakraborty, S. 2016. Recombination hotspots: Models and tools for detection. *DNA Repair (Amst).* **40**:47–56.

Pearson, V.M., Caudle, S.B., Rokyta, D.R. 2016. Viral recombination blurs taxonomic lines: examination of single-stranded DNA viruses in a wastewater treatment plant. *Peer Journal* **4**:279-281.

Pita, J.S., Fondong, V.N., Sangare, A., Ogwal, S., Fauquet, C.M. 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology* **82**:655–665.

Sattar, M.N., Iqbal, Z., Tahir, M.N. and Ullah, S. 2017. The prediction of a new CLCuD epidemic in the old world. *Frontiers in Microbiology* **8**:279-281.

Singh-Pant, P., Pant, P., Mazumdar-Leighton, S. 2012. Spatial and temporal diversity of begomoviral complexes in papayas with leaf curl disease. *Archives of Virology* **157**:1217-1232.

Tahir, M.N. and Mansoor, S. 2011. C1 of chili leaf curl betasatellite is a pathogenicity determinant. *Virology Journal* **8**:279-281.

van der Walt, E., Rybicki, E. P., Varsani, A., Polston, J. E., Billharz, R., Donaldson, L., et al. 2009. Rapid host adaptation by extensive recombination. *Journal of General Virology* **90**, 734–746.

Verma, S. and Saxena, S. 2017. Dissecting Papaya Leaf Curl Disease (PLCD) Complex and Assessing its Potential for siRNA Based Targeting. *International Journal of Plant and Environment* **3**:279-281.

Yang, X., Xie, Y., Raja, P., *et al.* 2011. Suppression of methylation-mediated transcriptional gene silencing by βc1-sahh protein interaction during geminivirus-betasatellite infection. *PLoS Pathogens* **7**:279-281.

Zhou, X. 2013. Advances in understanding Begomovirus Satellites. *Annual Reviews of Phytopathology* **51**:357-381.

Zubair, M., Zaidi, S., Shakir, S., *et al.* 2017a. An Insight into Cotton Leaf Curl Multan Betasatellite, the Most Important Component of Cotton Leaf Curl Disease Complex. *Viruses* **9**:279-281.

Zubair, M., Zaidi, S.S.E.A., Shakir, S., *et al.* 2017b. Multiple begomoviruses found associated with cotton leaf curl disease in Pakistan in early 1990 are back in cultivated cotton. *Scientific Reports* **7**:279-281.

Table 1. List of Betasatellite virus genomes used in this study causing leaf curl disease in papaya and associated crops and weeds.

Annotation	Satellite	Accession No.	Host	Geography	Length	c1 gene
TBS 1	Ageratum leaf curl betasatellite clone beta	JX512904.2	*Amaranthus hypochondriacus L.*	Lucknow, India	1362	180-596
TBS 2	Ageratum leaf curl betasatellite isolate Jaipur	KY089034.1	*Helianthus sp.*	Jaipur, India	1366	180-596
TBS 3	Ageratum leaf curl betasatellite isolate NBRI-B1	KR922821.1	*Calendula officinalis*	Lucknow, India	1363	180-596
TBS 4	Ageratum leaf curl betasatellite isolate Sikar M 2	KC589700.1	*Tagetes patula*	Lakshmangarh, Rajasthan, India	1335	180-608
TBS 5	Chilli leaf curl betasatellite - [Pakistan:Potato:2008]	FM179615.1	*Solanum tuberosum*	Punjab, Pakistan	1399	201-563
TBS 6	Chilli leaf curl betasatellite-[India:Palampur:2008]	FM877803.2	*Capsicum sp.*	Palampur, India	1376	200-562
TBS 7	Chilli leaf curl betasatellite-Panipat 4 [India:Panipat: Papaya:2008]	HM143904.1	*Carica papaya*	Panipat, Haryana, India	1369	201-557
TBS 8	Chilli leaf curl virus satellite DNA beta	EU582020.1	*Capsicum sp.*	Pataudi, India	1380	200-562
TBS 9	Chilli leaf curl virus satellite DNA beta C1 gene, isolated from Capsicum annuum in Pakistan	FN179279.1	*Capsicum annuum*	Pakistan	1387	201-563
TBS 10	Chilli leaf curl betasatellite isolate ToLCBDB-[IN:Nar:Chil:04]	JF706231.1	*Capsicum sp.*	Jodhpur, India	1380	200-562
TBS 11	Chilli leaf curl betasatellite isolate Meerut	JX193616.1	*Capsicum sp.*	Meerut, India	1390	221-670
TBS 12	Chilli leaf curl betasatellite isolate India:Punjab:TC241:2009	KJ605111.1	*Solanum lycopersicum*	Muskabad, Punjab, India	1373	201-563
TBS 13	Cotton leaf curl Burewala betasatellite	NC_013802.1	*Gossypium hirsutum*	Punjab, India	1354	198-554
TBS 14	Cotton leaf curl Burewala betasatellite, clone L2-RCA-bl-F	FN658722.1	*Gossypium hirsutum*	Punjab, India	1354	198-554
TBS 15	Papaya leaf curl virus-associated DNA beta	AY244706.1	*Carica papaya*	New Delhi, India	1372	201-557
TBS 16	Papaya leaf curl virus-associated DNA beta	NC_004706.1	*Carica papaya*	New Delhi, India	1372	201-557
TBS 17	Papaya leaf curl virus betasatellite isolate PRM	GU370715.1	*Solanum lycopersicum*	New Delhi, India	1377	200-556

TBS 18	Papaya leaf curl virus betasatellite isolate In:Var:Pum:08:1	HM101173.1	Pumpkin	Varanasi, India	1370	201-557
TBS 19	Papaya leaf curl virus betasatellite isolate PaLCuB-IYV:Del	JX050199.1	*Ipomoea purpurea*	New Delhi, India	1367	201-557
TBS 20	Papaya leaf curl virus betasatellite isolate PaLCuB-Pumpkin:IARI	JX040472.1	Pumpkin	New Delhi	1367	201-557
TBS 21	Papaya leaf curl virus betasatellite clone BG-CBE beta	KC959933.1	Black gram	Coimbatore, India	1351	188-556
TBS 22	Papaya leaf curl virus betasatellite clone GG-CBE	KC959934.1	Green gram	Coimbatore, India	1358	188-556
TBS 23	Papaya leaf curl virus beta-satellite clone BG-VBN	KC959935.1	Black gram	Pudukottai, India	1359	188-556
TBS 24	Papaya leaf curl virus beta- satellite isolate India:Pune: TC255:2010	KJ605112.1	*Solanum lycopersicum*	Pune, India	1367	200-556
TBS 25	Papaya leaf curl virus beta-satellite isolate India:Bangalore: TC281:2010	KJ605113.1	*Solanum lycopersicum*	Bangalore, India	1367	201-557
TBS 26	Papaya leaf curl betasatellite complete sequence, isolate Palampur	LN831955.1	*Valeriana jatamansi*	Palampur, India	1367	201-557
TBS 27	Papaya leaf curl betasatellite isolate NBRI	JX987089.2	*Parthenium hysterophorus L.*	Lucknow, India	1367	189-557
TBS 28	Papaya leaf curl betasatellite isolate GMT32	KT948074.1	*Cucurbita pepo*	Rawalpindi, Pakistan	1368	195-551
TBS 29	Papaya leaf curl betasatellite, clone Par-B1	LN906595.1	*Parthenium hysterophorus L.*	Lahore, Pakistan	1362	195-551
TBS 30	Papaya leaf curl betasatellite isolate India-Valsad-Cluster bean-2015	KT253636.1	*Cyamopsis tetragonoloba*	Valsad, Gujrat, India	1367	219-575
TBS 31	Papaya leaf curl betasatellite isolate India-Bhavnagar-Cluster bean-2015	KT253637.1	*Cyamopsis tetragonoloba*	Bhavnagar, Gujrat, India	1367	219-575
TBS 32	Papaya leaf curl betasatellite isolate DPB1	KX353621.1	*Solanum lycopersicum*	Sikar, Rajasthan, India	1367	189-557
TBS 33	Papaya leaf curl betasatellite isolate 2491	KY825245.1	*Capsicum* sp.	Islamabad, Pakistan	1331	221-577
TBS 34	Papaya leaf curl betasatellite isolate 2481	KY825246.1	*Capsicum* sp.	Islamabad, Pakistan	1369	221-577
TBS 35	Papaya leaf curl betasatellite isolate 49sb	KY825247.1	*Capsicum* sp.	Islamabad, Pakistan	1368	220-576
TBS 36	Tomato leaf curl virus-associated DNA beta	NC_004715.1	*Solanum lycopersicum*	Jabalpur, Madhya Pradesh, India	1424	195-221
TBS 37	Tomato leaf curl betasatellite isolate	KM201278.1	*Solanum*	India	1353	191-571

	ToLCuB[IN:Bah]		*lycopersicum*			
TBS 38	Tomato leaf curl betasatellite - [India:Halwadni:2007]	EU847239.1	*Solanum lycopersicum*	Haldwani, Uttarakhand, India	1370	200-556
TBS 39	Tomato leaf curl betasatellite-Naj 1 [India:New Delhi:Papaya:2008]	HM143909.1	*Carica papaya*	Najafgarh, New Delhi, India	1369	199-555
TBS 40	Tomato leaf curl betasatellite-DU [India:New Delhi:Papaya:2009]	HM143910.1	*Carica papaya*	New Delhi, India	1370	201-557
TBS 41	Tomato leaf curl betasatellite-Naj 2 [India:New Delhi:Papaya:2009]	HM143911.1	*Carica papaya*	Najafgarh, New Delhi, India	1370	201-557
TBS 42	Tomato leaf curl betasatellite-Panipat 1 [India:Panipat:Papaya:2008]	HM143901.1	*Carica papaya*	Panipat, Haryana, India	1369	201-557
TBS 43	Tomato leaf curl betasatellite-Panipat 2 [India:Panipat:Papaya: 2008]	HM143902.1	*Carica papaya*	Panipat, Haryana, India	1370	200-556
TBS 44	Tomato leaf curl betasatellite-Panipat 5 [India:Panipat:Papaya: 2008	HM143905.1	*Carica papaya*	Panipat, Haryana, India	1373	201-557
TBS 45	Tomato leaf curl betasatellite-Panipat 7 [India:Panipat:Papaya :2008	HM143907.1	*Carica papaya*	Panipat, Haryana, India	1375	201-557
TBS 46	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D17	LT009401.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 47	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D1	LT009403.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 48	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D22	LT009406.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 49	Tomato yellow leaf curl China virus-associated DNA beta complete genome, isolate Y87	AJ457818.1	*Solanum lycopersicum*	Yunnan, China	1344	209-565
TBS 50	Tomato yellow leaf curl China virus-associated DNA beta complete genome, isolate Y38	AJ420315.2	*Solanum lycopersicum*	Yunnan, China	1338	209-565
TBS 51	Tomato leaf curl Joydebpur beta virus	NC_010236.1	*Capsicum* sp.	Kalyani, West Bengal, India	1370	193-573

c1= betasatellite ORF

Table 2 Recombination analysis of Betasatellite genomes reported in Indian sub-continent using RDPv4.95 program.

Event No.	Recombinant	Major parent	Minor Parent	Frequency of recombination event	Detection Method						
					R	G	B	M	C	S	T
1	TBS32	TBS3	TBS18	2	+	+	+	+	+	+	+
2	TBS8	TBS10	TBS11	1	+	+	+	+	+	+	+
3	TBS4	TBS1	TBS34	1	+	+	+	+	+	+	+
4	TBS29	TBS18	TBS47	15	+	+	+	+	+	+	+
5	TBS27	TBS32	TBS15*	1	+	+	+	+	-	+	+
10	TBS27	TBS21	TBS14	13	+	+	-	+	-	+	+
11	TBS38	TBS39*	TBS6	5	+	+	-	+	+	+	+
12	TBS3	TBS2	TBS1	1	-	+	+	+	+	+	+
13	TBS35	TBS2	TBS1	4	+	+	+	+	+	+	-
15	TBS16	TBS26	TBS27*	2	+	+	+	+	+	-	-
57	TBS36	TBS27	TBS26	5	+	+	+	+	+	+	+
Total = 11			**Total frequency = 50**								

* Unknown parent detected in recombination analysis but the specified parent was the closest inferred sequence in this analysis

The analysis was performed against Betasatellite and DNA-A sequences of 51 begomoviruses using default parameters and following algorithms: R=RDP; G=GENCONV; B=BOOTSCAN; M=MAXCHI; C=CHIMAERA; S=SISCAN; T=3SEQ

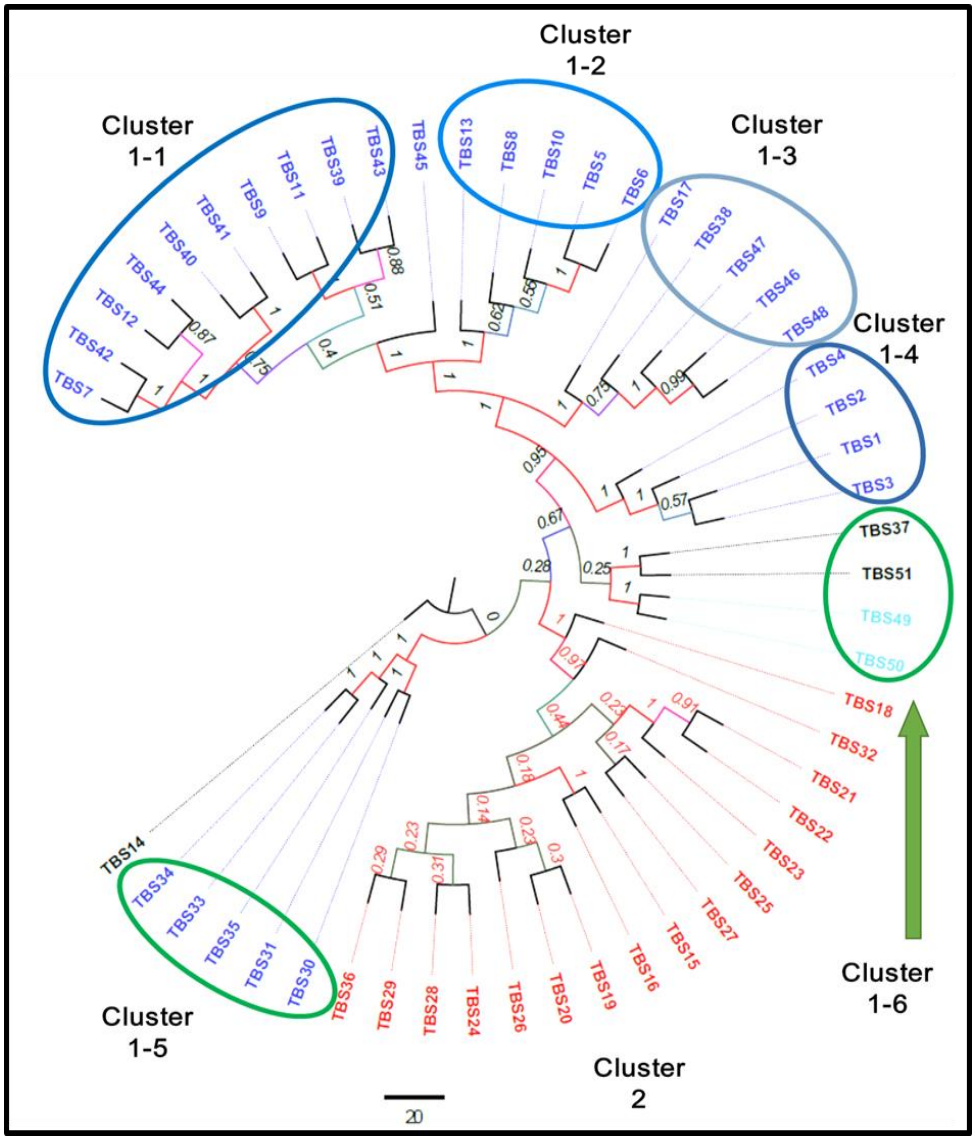


Fig. 1: Phylogenetic analysis of betasatellite component of PLCD complex and associated leaf curl disease group. The tree was divided into two clusters i.e. Cluster-1 and Cluster 2 on the basis of likelihood predicted by bootstrap step estimation. The Cluster-1 was further subdivided into five sub-clusters on the basis of host, virus and region. Cluster-2 is sparsely related, therefore, all genomic components in this group were analyzed separately. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-13451.7843) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.5473)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 9.4207% sites). The analysis involved 51 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 878 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.

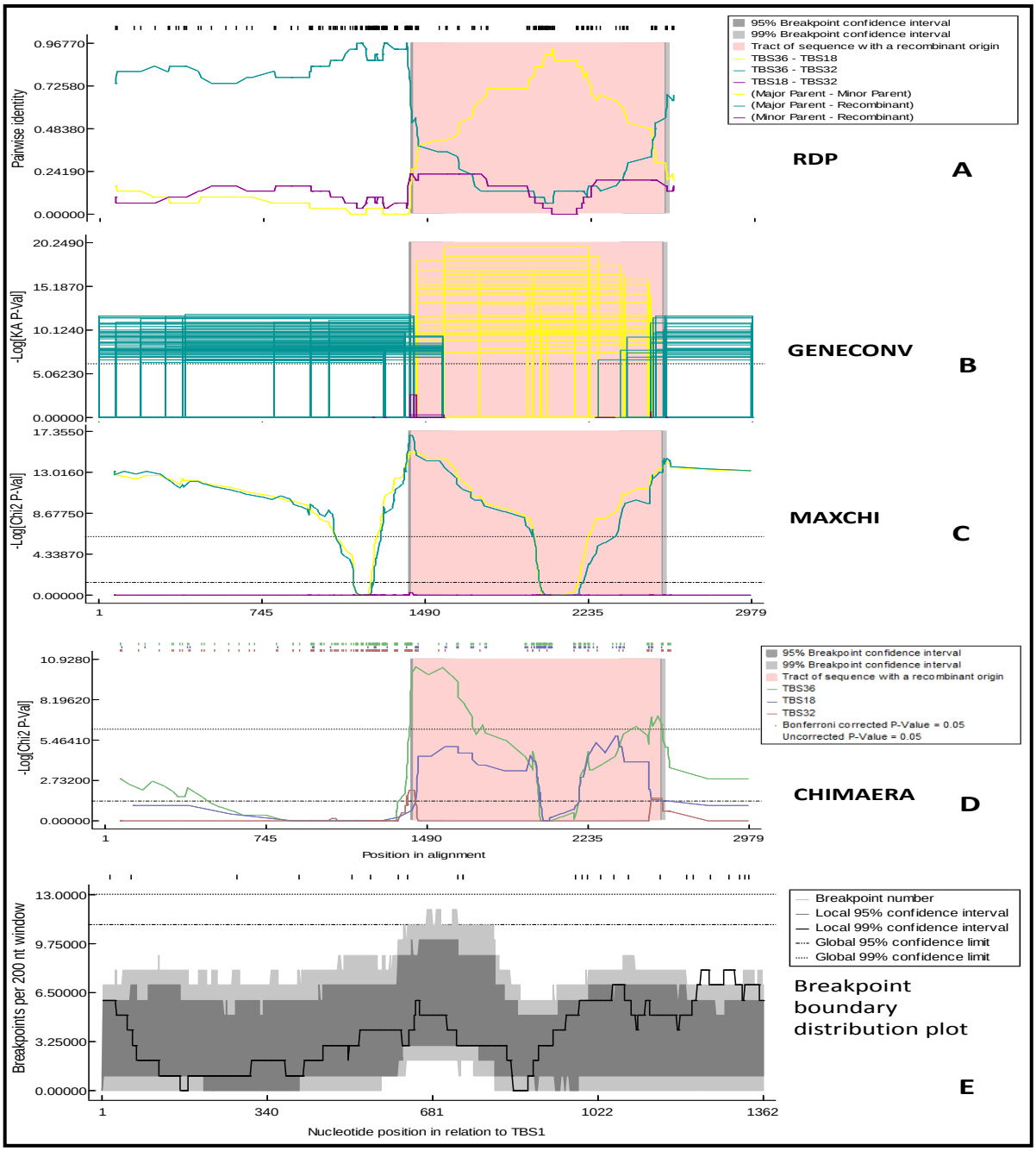


Fig. 2: Recombination analysis showing results of event number 1 using Recombination Detection Program v4.95 (RDP v4.95). The pink region indicates the occurrence of the recombination event with intersection points in the case of RDP and overlapping peaks in case of MAXCHI on the plot as the recombination starting and ending breakpoints. The grey region at the ends represents 99% and 95% confidence interval of prediction of breakpoints as mentioned in the key. A. RDP output; B. GENECONV output; C. MAXCHI output; D. CHIMAERA output; E. Recombination breakpoint distribution plot: The light and dark grey regions in the plot represent the 99% and 95% confidence limit respectively for a 200nt scanning window. The breakpoint distribution was assessed with TBS1 sequence as a reference sequence just for qualitative purpose.

CURRICULUM-VITAE

Saurabh Verma

DOB: 16th October 1982

Babasaheb Bhimrao Ambedkar University
Department of Biotechnology
School of Biosciences & Biotechnology
Vidya Vihar, Raebareli road
Lucknow-2260025. INDIA.
Contact (M0bile):+919565427565
email:saurabhv1982@gmail.com

Education

- **PONDICHERRY UNIVERSITY** — **PUDUCHERI. INDIA**
  Master of Science in Biotechnology — **1st Division (CGPA 7.25/10) - 2007**
- **LUCKNOW UNIVERSITY** — **LUCKNOW, INDIA**
  Bachelor of Science (Zoology, Chemistry, Botany) — **1st Division (61.3%) - 2005**
- **CLASS 12th** — **ISC, New Delhi, INDIA**
  Biology — **1st Division (60.6%) – 2000**
- **CLASS 10th** — **ICSE, New Delhi, INDIA**
  1st Division (75%)

Honors and Awards

- **Junior Research Fellowship** awarded from **Council for Scientific and Industrial Research, India (2008)**
- **Junior Research fellowship from Indian Council for Medical Research (2008)**
- **Project assistantship** at **National Botanical Research Institute, Lucknow (January 2008 till September 2008)**
- **Qualified GATE 2007** exam conducted by **Indian Institute of Technology, India.**
- **Post Graduate scholarship** from **Department of Biotechnology, MHRD, Govt. of India (2005 - 2007)**

- Research Publications in UGC recognized journals - 4