STUDIES ON PHYLOGEOGRAPHY OF CHILLI-INFECTING BEGOMOVIRUSES AND HOST-VIRUS INTERACTION DURING PATHOGENESIS

Thesis submitted to

Jawaharlal Nehru University, New Delhi

For the award of the degree of

DOCTOR OF PHILOSOPHY

VINOTH KUMAR R

School of Life Sciences
Jawaharlal Nehru University
New Delhi – 110 067

2016

Dedicated to My Parents

SCHOOL OF LIFE SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI -110067

Date: 19, July, 2016

CERTIFICATE

This is to certify that the work embodied in this thesis entitled **"Studies on phylogeography of chilli-infecting begomoviruses and host-virus interaction during pathogenesis"** has been carried out at School of Life Sciences, Jawaharlal Nehru University, New Delhi, India. This work is original and has not been submitted so far in part or in full, for the award of any degree or diploma by any university.

Vinoth Kumar R
Student

Prof. Supriya Chakraborty
Supervisor,
School of Life Sciences, JNU

SUPRIYA CHAKRABORTY, Ph.D.
Professor
School of Life Sciences
Jawaharlal Nehru University
New Delhi - 110067

19.7.16
Dean
School of Life Sciences, JNU

कार्यकारी डीन / Acting Dean
जीवन विज्ञान संस्थान / School of Life Sciences
जवाहरलाल नेहरू विश्वविद्यालय
Jawaharlal Nehru University
नई दिल्ली/New Delhi-110067

ACKNOWLEDGEMENTS

The success of this research depends largely on the encouragement of many others apart from the efforts from my side. I wish to express my sincere gratitude to all the people who have directly or indirectly contributed to my doctoral program.

First and foremost, with a deep sense of gratitude and immense respect, I thank my research supervisor Prof. Supriya Chakraborty, School of Life sciences, Jawaharlal Nehru University, New Delhi for his guidance, understanding and patience. He encouraged me to not only grow as an experimentalist and a biologist but also as an independent thinker. He stood by me in all moments of failure and disappointment and infused me with new zeal and ardor to hang on. The upliftment that he has given to my academic, professional, financial and moral well being cannot be described in words. In every sense, none of this work would have been possible without him.

I am thankful to Dr. B. Chattopadhyay for giving me valuable inputs and guidance during my early days. She encouraged and showed a success path to me whenever I had failed in my results.

I feel highly obliged in expressing my deep sense of gratitude to the Dean, School of Life Sciences, Prof. B. C. Tripathi, conceding me to avail the best of facilities at the elite school of the University. I wish to thank University grant commission (UGC) for providing me Non-NET research fellowship during my Ph. D tenure.

I am deeply obliged to my teacher Dr. S. Sekar (Bharathidasan University, Trichy), for guiding and helping me, to tie down a loose end by providing broad base of knowledge for the research areas. I would like to record my gratitude to Prof. A. K. Nandi, JNU, New Delhi for his right preparatory guidance and cooperation.

I would like to thank people from Jawaharlal Nehru University. They all created a wonderful working environment making my running around worthwhile. My special thanks to Dr. Achuit Kumar Singh (Indian Institute of Vegetable Research) for his valuable words throughout this study. He was most instrumental in evoking my research interests by exposing me to several forefront research issues and for his timely advices to make my research more productive. The liberality and assistance of all my lab mates at JNU, Dr. P. Ranjan, Dr. B. George, Dr. S. Basu, Dr. V. K. Sharma, Dr. N. Kushwaha, Dr. D. Bhattacharya, A. K. Singh, R. Ruhel, D. Raghunathan, Mansi, P. Shekar, N. Sharma, and K. Prabhu has made my goal achievable. My special thanks Pranav Sahu for engaging in scientific discussions and his help during the submission of this thesis.

My heartfelt gratitude and thanks go out especially to, Dr. C. M. Alam, G. Prabu, D. Singh, Kishore, Ved, Tsewang and Manish, for their informal milieu, friendly and lively atmosphere. I would like to extend my thanks to the lab trainees especially, Neha, Priyanka, Raja, Prabakaran, Aruljothi, Rihana, Kavita, Arpita and Krishnendu Sen, for their immense help and co-operation during my experiments. I am thankful to my lab attendants Umesh, Sunny, Satish, Deepak, Rahul and Sujan for arranging lab wares during my stay in the lab and for maintaining the plants in green house.

Besides, I would also like to convey my gratitude to all the respected teachers and administrative staff of SLS for their kind help and support throughout my research period. I would like to thank all the CIF members and the other non-teaching staffs for their indirect help during my experiments.

I thank all my relatives for their cheering words, especially my loving sister and brother-in-law for standing by me through moments of joy and despair. Without the support of my friends it was not possible to bring this work to its logical end. Those who were always a pillar of strength to me includes Gugan, Pres, Alms, Kavin, Swarna, Satish, Chakra, Sen, Prem, Charu, Viky, Vinodh, Viji, Janu, Rajesh and Atri. .

This is an appropriate time to convey my appreciation to the special bond I share with my hostel roommate, Anirban Chatterjee, for giving a lot of memorable stories to share amongst us during my stay in the room. I also thank Kalai and Gopi, for making my stay a memorable one.

My parents have played the most pivotal role to boost my aspiration and have given moral support by patting on my back to see that my confidence is not dented at the time of stress. If we ever had a family motto that would have been – Time is the most Potent and Powerful – a philosophy of life I have been carrying with me every day. The successful completion of my doctoral program is an answer to their prayers and a fruit of their blessings. No words can be enough to thank them, for that, I dedicate my thesis to their feet. However, this is just a small step towards unveiling their dreams.

-

-Vino

LIST OF FIGURES

Figure number	Figure legend	Page number
2.1	The worldwide distribution of chilli cultivation and production.	6
2.2	The distribution of chilli production among various states in India.	6
2.3	Genome organization of the members of the genus included in the *Geminiviridae* family.	10
2.4	Schematic representation of the Rep protein with functional domains.	16
2.5	Schematic representation of the C2 (TrAP) with functional domains.	21
2.6	Model depicting the inhibitory role of C2/L2/AL2 in stress response.	22
2.7	Schematic representation of the various domains of AC3/REn proteins.	25
2.8	Schematic illustration of the various domains of coat protein.	27
2.9	Various domains of BV1/NSP proteins.	31
2.10	Schematic representation of the BC1/MP protein and its various domains.	32
2.11	Processes involved in RNA directed DNA methylation pathway of plants.	37
2.12	Schematic model depicting the self-reinforcing loop between histone and DNA methylation in plants.	37
2.13	Maintenance mechanism of DNA methylation in animals and plants.	38
2.14	Begomovirus-encoded proteins (C2/TrAP and βC1) and inhibition of plant methyl cycle.	43
4.1	The begomovirus-like symptoms noticed on the leaf curl disease affected chilli plants from major chilli growing regions in the country.	64
4.2	Agarose gel electrophoresis showing the digestion pattern of the viral genomic components.	66
4.3	Graphical representation of per cent nucleotide identity plot of begomoviruses using SDT v 1.0 program.	67
4.4	Neighbor-joining phylogenetic tree of ChiLCD-associated begomoviruses isolated here (indicated in bold type) along with the representative begomoviral genomes retrieved from the GenBank.	68
4.5	Neighbor-joining phylogenetic dendrogram of DNA-B components associated with ChiLCD.	72

Figure number	Figure legend	Page number
4.6	Graphical representation of per cent nucleotide identity plot of betasatellites using the SDT v 1.0 program.	75
4.7	Neighbor-joining phylogenetic tree of ChiLCD-associated betasatellites isolated here (indicated in bold type) along with the representative betasatellites retrieved from the GenBank.	76
4.8	Distribution of per cent pairwise identity between the nucleotide sequences of 326 alphasatellites.	78
4.9	Graphical representation of per cent nucleotide identity plot of alphasatellites using SDT v 1.0 program.	79
4.10	Neighbor-joining phylogenetic tree of ChiLCD-associated alphasatellites isolated in this study (indicated in bold type) along with the representative alphasatellites retrieved from the GenBank.	80
4.11	Phylogenetic tree of begomovirus DNA-A encoded AV1 and AC1 ORFs.	85
4.12	Nucleotide diversity of the ORFs encoded by ChiLCD-associated begomoviruses, betasatellites and alphasatellites.	89
4.13	Cluster-based sub-population analysis among the begomoviruses associated with ChiLCD in India.	90
4.14	Distribution of begomoviruses, betasatellites and alphasatellites associated with ChiLCD in India.	95
5.1	Agarose gel electrophoresis showing confirmation of the partial tandem repeat constructs of NgA and Ngβ by digestion.	100
5.2	Agarose gel electrophoresis showing confirmation of the partial tandem repeat constructs of VeA and Veβ by digestion.	100
5.3	Agarose gel electrophoresis showing the confirmation of the partial tandem repeat construct of ChiLCA by restriction digestion.	102
5.4	Infectivity analysis of the viral components cloned from Nagpur and Vellanad regions.	106
5.5	Symptom appearance on *C. annuum* var Sulekha plants inoculated with various begomovirus-betasatellite components.	107

Figure number	Figure legend	Page number
5.6	Relative level of viral DNA in the infected chilli plants.	108
5.7	Symptom appearance on *N. benthamiana* plants inoculated with various combinations containing begomoviruses and betasatellites alongwith alphasatellite components at 21 dpi.	110
5.8	Accumulation of viral DNA in infected *N. benthamiana* plants at 21 dpi.	111
5.9	Transient viral DNA replication assay using *N. benthamiana* leaf disks.	111
6.1	Relative accumulation of virus transcripts in the inoculated *N. benthamiana* plants.	117
6.2	PCR amplification of NbVIM.	117
6.3	Neighbor joining phylogenetic genetic dendrogram of full-length amino acid sequences of VIM proteins.	120
6.4	Agarose gel electrophoresis showing digestion confirmation of viral ORFs cloned into pGADC1 vector.	121
6.5	Interaction between virus-encoded proteins and NbVIM protein through Yeast-two-hybrid assay.	122
6.6	Schematic representation of various domains of NbVIM.	122
6.7	Agarose gel electrophoresis showing the digestion confirmation of NbVIM domains cloned into pGBKT7 vector.	123
6.8	Interaction between V2 and NbVIM proteins through Yeast-two-hybrid assay.	123
6.9	Agarose gel electrophoresis confirming cloning of full-length and domains of NbVIM into BiFC vectors.	124
6.10	Interaction between V2 and NbVIM proteins through BiFC analysis.	126
6.11	Effect of NbVIM on ChiLCV pathogenesis.	127

LIST OF TABLES

Table number	Table caption	Page number
2.1	Composition of the subunits present in RNA Pol II, IV and V of *Arabidopsis thaliana* plant.	35
2.2	Components involved in DNA methylation pathway of animals and plants.	39
3.1	Symptom appearance on *Capsicum* spp. collected from major chilli growing regions in India.	46
3.2	List of primers used to clone the viral genomic components.	52
3.3	Name and accession numbers of the begomoviral genomic components selected from the database to be used in the RDP analysis.	55
3.4	Details of the symptom measurement using 0-5 scale.	59
4.1	Survey, collection and the list of helper begomoviruses identified from the ChiLCD-affected samples from major chilli growing regions in India.	69
4.2	List of betasatellites identified from the ChiLCD-affected samples from major chilli growing regions in India.	74
4.3	Alphasatellites identified from the ChiLCD-affected samples from major chilli growing regions in India.	77
4.4	Recombination analysis of cloned begomoviruses identified from the ChiLCD-affected samples.	81
4.5	Recombination analysis of cloned betasatellites from the ChiLCD-affected samples.	86
4.6	Recombination analysis of cloned alphasatellites from the ChiLCD-affected samples.	88
4.7	Genetic variability among the isolates of begomoviruses, alphasatellites and betasatellites associated with ChiLCD in India.	89

Table number	Table caption	Page number
4.8	Results of neutrality tests and methods to detect positive or negative selected sites in the ORF(s) encoded by ChiLCD associated begomoviruses and betasatellites.	91
4.9	Rate of nucleotide substitution and mutation rate of ChiLCD-associated viral components-encoded ORFs.	92
5.1	Details of the cloning sites used for the generation of partial tandem repeat constructs used in this study.	101
5.2	Infectivity studies of the cloned viral components on *N. benthamiana* and *C. annuum* var. Sulekha plants.	103

ABBREVIATIONS

%	Percent
°C	Degree Celsius
A-rich	Adenine-rich region
BiFC	Bimolecular fluorescent complementation
cDNA	Complementary DNA
ChiLCA	*Chilli leaf curl alphasatellite*
ChiLCD	Chilli leaf curl disease
ChiLCV	*Chilli leaf curl virus*
CroYVMB	*Croton yellow vein mosaic betasatellite*
DAPI	4, 6-diamidino-2-phenylindole
DO	Drop out
DPI	Days-post inoculation
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxy nucleotide triphosphates
dpi	Days post inoculation
dsDNA	Double stranded DNA
EDTA	Ethylene diaminetetra acetic acid
EtBr	Ethidium bromide
g	Gram
GFP	Green fluorescent protein
GDarSLA	*Gossypium darwinii symptomless alphasatellite*
h	Hour
IR	Intergenic region
LiAc	Lithium acetate
kb	Kilo base pairs
LB	Luria-Bertani medium
M	Molar
min	Minutes
ml	Millilitre

mm	Millimeter
mM	Millimolar
ML	Maximum-likelihood
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ng	Nano gram
NJ	Neighbor-joining
OD	Optical density
ORF	Open reading frame
PaLCuV	*Papaya leaf curl virus*
PCR	Polymerase chain reaction
PepLCBV	*Pepper leaf curl Bangladesh virus*
PTR	Partial tandem repeat
RCA	Rolling circle amplification
RDP	Recombination detection program
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
rpm	Revolutions per minute
SCR	Satellite conserved region
SDT	Species demarcation tool
sec	Second
ssDNA	Single stranded DNA
ToLCBDB	*Tomato leaf curl Bangladesh betasatellite*
ToLCJoB	*Tomato leaf curl Joydebpur betasatellite*
ToLCGV	*Tomato leaf curl Gujarat virus*
ToLCNDV	*Tomato leaf curl New Delhi virus*
TRV	*Tobacco rattle virus*
var.	Variety
v/v	Volume by volume
w/v	Weight by volume
µg	Microgram
µl	Microliter
µM	Micromolar

TABLE OF CONTENTS

1 Introduction	1
2 Review of Literature	
2.1 Chilli (*Capsicum* species)	5
2.2 Chilli cultivation	5
2.3 Chilli leaf curl disease (ChiLCD)	7
2.4 The causal organism	
2.4.1 Family: *Geminiviridae*	7
2.4.1.1 Taxonomy	8
2.5 Genome organization of Geminiviruses	
2.5.1 *Becurtovirus*	8
2.5.2 *Begomovirus*	8
2.5.3 *Curtovirus*	9
2.5.4 *Eragrovirus*	9
2.5.5 *Mastrevirus*	10
2.5.6 *Topocuvirus*	11
2.5.7 *Turncurtovirus*	11
2.6 Function of begomovirus components	
2.6.1 DNA-A component	12
2.6.2 DNA-B component	13
2.6.3 Betasatellites	13
2.6.4 Alphasatellites	14
2.7 Functions of genes encoded by DNA-A	
2.7.1 Replication associated protein (Rep)	
2.7.1.1 DNA-A and DNA-B replication	16
2.7.1.2 Satellite DNA replication	17
2.7.1.3 Role in gene expression	18
2.7.1.4 Inhibits TGS and PTGS activities	18
2.7.1.5 Interaction with host factors	
2.7.1.5.1 Plant retinoblastoma-related protein (RBR)	19
2.7.1.5.2 Proliferating cell nuclear antigen (PCNA)	19
2.7.1.5.3 Replication factor C (RFC)	19
2.7.1.5.4 Replication protein A (RPA) and Minichromosome maintenance protein 2 (MCM2)	20
2.7.1.5.5 Histone H3 and RAD51/54	20
2.7.1.5.6 Other cellular factors	20
2.7.2 Transcription activator protein (TrAP)	
2.7.2.1 Gene activation	21
2.7.2.2 Pathogenicity factor	22
2.7.2.3 Suppressor of RNAi silencing	23
2.7.3 Replication enhancer protein (REn)	
2.7.3.1 Virus replication	24
2.7.3.2 Interaction with other cellular proteins	24
2.7.4 AC4 protein	
2.7.4.1 Pathogenicity determinant	25
2.7.4.2 RNA silencing suppression	26
2.7.5 AC5 protein	26
2.7.6 Coat protein (CP)	
2.7.6.1 Nucleo-cytoplasmic shuttling	27

2.7.6.2 Nucleolus localization	28
2.7.6.3 Insect transmission	28
2.7.6.4 Cell-to-cell movement	28
2.7.7 Pre-coat protein	
2.7.7.1 Virus movement	29
2.7.7.2 Pathogenesis	29
2.7.7.3 Suppressor of RNA silencing	29
2.8 Functions of genes encoded by DNA-B	
2.8.1 Nuclear shuttle protein (BV1)	
2.8.1.1 Virus movement	30
2.8.1.2 Pathogenicity determinant	31
2.8.2 Movement protein (BC1)	
2.8.2.1 Inter-cellular trafficking	32
2.8.2.2 Pathogenicity determinant	32
2.9 RNA directed DNA methylation (RdDM)	
2.9.1 Mechanism of RdDM pathway in plants	
2.9.1.1 Step 1: Generation of Pol IV derived siRNAs	33
2.9.1.2 Step 2: Generation of 24nt siRNAs and its binding to their target RNA	34
2.9.1.3 Step 3: Methylation of DNA regions targeted by 24nt siRNAs	36
2.9.2 Maintenance of DNA methylation in plants	
2.9.2.1 VARIANT IN METHYLATION (VIM)	39
2.9.2.2 CG methylation maintenance in animals and plants	40
2.9.2.3 Non-CG methylation maintenance in plants	41
2.10 RdDM as a defense mechanism against geminivirus infection	
2.10.1 Viral DNA methylation and host recovery or symptom remission	42
2.10.2 Viral factors countering TGS pathway	
2.10.2.1 By affecting plant methyl cycle	42
2.10.2.1 By preventing subsequent DNA or histone methylation	43
3 Materials and methods	
3.1 Survey and collection of samples	45
3.2 Genomic DNA isolation from the plant tissues	45
3.3 Rolling circle amplification	47
3.4 Preparation of *E. coli* competent cells	47
3.5 Transformation of *E.coli* cells	47
3.6 Isolation of plasmid DNA	48
3.7 Restriction digestion of plasmid DNA	48
3.8 Dephosphorylation of linearized vectors	49
3.9 Ligation for cloning of DNA	50
3.10 Polymerase chain reaction (PCR)	50
3.11 Gel elution	51
3.12 Cloning of amplified PCR product in the cloning vectors	
3.12.1 Cloning of viral DNA using RCA product	51
3.12.2 Cloning of viral DNA through PCR	52
3.13 Per cent pairwise identity and phylogenetic analysis	53
3.14 Recombination analysis	54
3.15 Selection pressure analysis	54
3.16 Estimation of nucleotide substitution rates	56
3.17 Preparation of *Agrobacterium* competent cells	57

3.18 Agro-transformation	57
3.19 Generation of partial tandem repeat constructs (PTR)	57
3.20 Infectivity analysis	
3.20.1 Agro-inoculation	58
3.20.2 Particle bombardment	58
3.20.3 Detection of viral DNAs	59
3.20.4 Southern hybridization for analyzing viral DNA accumulation	59
3.21 Transcript analysis	
3.21.1 Isolation of total RNAs	60
3.21.2 cDNA synthesis	61
3.21.3 Quantitative real time PCR analysis	61
3.22 Protein-protein interaction studies	
3.22.1 Yeast-two-hybrid assay	61
3.22.2 Bimolecular fluorescence complementation (BiFC) assay	62
3.23 Virus induced gene silencing	62
4 Objective 1	
4.1 Introduction	63
4.2 Results	
4.2.1 Survey and collection of samples	63
4.2.2 Cloning of viral DNAs associated with ChiLCD in India	
4.2.2.1 Association of begomoviral genomic components with Indian ChiLCD	66
4.2.2.2 Association of begomoviral DNA-B component with ChiLCD in India	72
4.2.2.3 Genetic distribution of betasatellites associated with Indian ChiLCD	73
4.2.2.4 Association of alphasatellites with Indian ChiLCD	77
4.2.3 Recombination analysis	
4.2.3.1 Detection of recombination among Indian ChiLCD-associated begomoviruses	78
4.2.3.2 Recombination pattern detected among ChiLCD-associated betasatellites	83
4.2.3.3 Evidence of recombination among the alphasatellites with ChiLCD in India	87
4.2.4 Nucleotide diversity of ChiLCD-associated viral genomic components	87
4.2.5 Rate of nucleotide substitution of cloned viral genomic components	92
4.3 Discussion	93
5 Objective 2	
5.1 Introduction	99
5.2 Results	
5.2.1 Generation of partial tandem repeat (PTR) constructs	
5.2.1.1 NgA and Ngβ	99
5.2.1.2 VeA and Veβs	100
5.2.1.3 ChiLCA (DNA-1)	102
5.2.2. Agro-inoculation of cloned viral DNAs in *N. benthamiana* plants	102
5.2.3 Infectivity of cloned viral molecules on *Capsicum* plants	105
5.2.4 Role of alphasatellite in begomovirus pathogenesis on *N. benthamiana* plants	109
5.3 Discussion	112

6 Objective 3	
6.1 Introduction	115
6.2 Results	
6.2.1 Level of NbVIM transcript upon virus infection	116
6.2.2 Isolation and sequence analysis of the VARIANT IN METHYLATION gene from *N. benthamiana*	117
6.2.2.1 Nucleotide sequence of the clone NbVIM	118
6.2.2.2 Predicted protein product of NbVIM	119
6.2.3 Interaction between geminivirus V2 protein and NbVIM	
6.2.3.1 Cloning of viral ORFs	120
6.2.3.2 Interaction between pre-coat protein and NbVIM through Yeast-two-hybrid assay	121
6.2.3.3 Cloning of various domains of NbVIM into pGBKT7 vector	123
6.2.3.4 Interaction of pre-coat protein with various NbVIM domains through Yeast-two-hybrid assay	
6.2.3.5 Bimolecular fluorescence complementation (BiFC) assay	
6.2.3.5.1 Cloning of the full-length and various domains of NbVIM into BiFC vectors	124
6.2.3.5.2 Protein-protein interaction through BiFC analysis	125
6.2.4 Transient silencing of NbVIM by TRV based-virus induced gene silencing	125
6.3 Discussion	128
7 Summary	131
8 References	133
9 Appendix	156
10 List of publications	163

1. Introduction

Capsicum spp. is an economically important spice cum vegetable crop cultivated worldwide. There are about 5 domesticated and 25 wild species of chilli is available for cultivation worldwide. Among them, *Capsicum annuum* is the most widely cultivated species. The use of chillies in Peru has been dated to pre-historic times; however, the origin of this species is believed to be around 7500 BC in Guatemala and New Mexico (Milind and Sushila, 2012). The extracts from chillies contain various components with very high medicinal values such as anti-cancerous compounds and value added products such as oleoresin. Capsaicin, the active component of chilli having medicinal importance is widely used against the intestinal pathogenic bacteria and to disinfect the oral and gastric mucous membranes (Bosland and Votava, 1999). Capsaicin is being used to treat patients with arthritis, sensory nerve fiber disorders and psoriasis (Chhabra *et al*., 2012). It is also highly rich in minerals such as potassium, iron and magnesium, and vitamins such as vitamin A, C and E, (Simonne *et al.,* 1997). Recently, vitamin P, a derivative from green chillies has been used to protect from secondary irradiation injury (Milind and Shushila, 2012). In addition, chilli is used as coloring agent and in resin making.

The average chilli productivity in our country is reduced mainly due to biotic stress caused by pests and pathogens. Chillies are known to be the host for various bacteria and fungi; however, viruses have emerged to be a serious threat for its cultivation (Varma and Malathi, 2003). Among the viral pathogens, the viruses causing leaf curl disease in chilli are posing a major hurdle to chilli production in our country. Under severe condition, these virus-infected plants remain stunted leading to fewer or no fruit setting. The viruses causing leaf curl disease in chilli belong to the family *Geminiviridae*. It includes a group of twinned icosahedral particles containing circular single stranded genomic component(s). Taxonomically, based on their insect vector, genome organization and host range, the members of this family are categorized into seven genera such as *Becurtovirus, Begomovirus, Curtovirus, Eragrovirus, Mastrevirus, Topocuvirus* and *Turncurtovirus* (Varsani *et al*., 2014). Among them, begomoviruses are geographically distributed throughout the world and are transmitted by the arthropod vector, *Bemisia tabaci* Genn. Begomoviruses can infect wide range of hosts and are considered to be one of the most destructive plant pathogens (Varma and Malathi, 2003).

Based on the requirement of genomic components for successful pathogenesis, they are named either as bipartite (both DNA-A and DNA-B) or monopartite (only DNA-A molecule) (Brown *et al*., 2012). The presence of monopartite and bipartite begomoviruses have mainly documented from the 'Old World' while the 'New World' is considered as the primary source of bipartite begomoviruses with the notable exception of a monopartite begomovirus, *Tomato leaf deformation virus* (ToLDeV) (Rojas *et al*., 2005; Melgarejo *et al*., 2013). Recently, the members of monopartite and few bipartite viruses have been found along with ssDNA satellite molecules named as alphasatellite and betasatellite (Nawaz-ul-Rehman and Fauquet, 2009). Generally, DNA-A or DNA-A like genomic component encodes six ORFs coding for replication associated protein, transcriptional activator protein, replication enhancer protein, silencing suppressor in complementary strand and coat protein, pre-coat protein in virion strand. DNA-B component codes for movement protein and nuclear shuttle protein which aids for intra- and inter-cellular movement. In bipartite begomoviruses, both of these DNA components are needed for systemic infection and for typical symptom induction (Brown *et al*., 2012; Hanley-Bowdoin *et al*., 2013). The majority of 'Old World' begomoviruses are mostly identified along with satellite molecules, named betasatellites of ~1350 nucleotides in length (Saunders *et al*., 2000; Nawaz-ul-Rehman & Fauquet, 2009). These molecules contain a satellite conserved region [SCR], an adenine rich region and a single ORF called βC1 in the complementary sense strand. However, the aspect of trans-replication of these betasatellite molecules by various begomoviruses has been reported to be host specific (Ranjan *et al*., 2014). Besides assisting in virus movement, these betasatellite-encoded βC1s have also been demonstrated to counter RNAi silencing as well as pathogenicity determinants (Saunders *et al*., 2004; Cui *et al*., 2005; Yang *et al*., 2011; Li *et al*., 2014). The satellite-like molecules (alphasatellites) require helper begomoviruses for their intra- and inter-cellular movement, and also have been shown to ameliorate symptoms (Mansoor *et al*., 1999; Nawaz-ul-Rehman *et al*., 2010; Idris *et al*., 2011).

RNA silencing is a highly specific and intricate process employed by host plants as a defence mechanism to combat virus infections. This occurs both at transcriptional and post-transcriptional level (Law and Jacobsen, 2010; Hohn and Vazquez, 2011; Pumplin and Voinnet, 2013). Transcriptional gene silencing (TGS) mainly occurs in the nucleus and it is necessary for the initiation and maintenance of heterochromatic regions (Law

and Jacobsen, 2010). TGS results in DNA methylation and also cross-talk of histone and DNA methylation has been widely reported (Law and Jacobsen, 2010; Du *et al*., 2015). Post-transcriptional gene silencing (PTGS) mechanism that includes miRNA, ta-siRNA and nat-siRNA pathways acts primarily in cytoplasm of the cell (Hohn and Vazquez, 2011). The precursor for both TGS and PTGS is the dsRNAs derived either from the transcription of imperfectly matched hairpins regions, from inverted-repeat loci or by overlapping, bidirectional or convergent transcription (Law and Jacobsen, 2010; Pumplin and Voinnet, 2013). In plants, these gene silencing processes play very essential roles in development, various stresses, maintaining genome integrity and heterochromatin regions, and in defense against numerous pathogens.

Consequently, viruses including the members of *Geminiviridae* also encode various proteins to counter this host mediated gene silencing (defense) mechanisms (Hohn and Vazquez, 2011; Hanley-Bowdoin *et al*., 2013). Several geminiviral proteins such as Rep (Rodriguez-Negrete *et al.,* 2013), TrAP (Dong *et al*., 2003; Vanitharani *et al.,* 2004; Trinks *et al.,* 2005; Wang *et al.,* 2005; Buchmann *et al.,* 2009; Yang *et al.,* 2012; Chung *et al.*, 2014; Castillo-Gonzalez *et al.,* 2015; Kumar *et al*., 2015; Sun *et al.,* 2015; Ye *et al.,* 2015), C4/AC4 encoded protein (Chellappan *et al.,* 2004; Piroux *et al.*, 2007; Dogra *et al*. 2009), C5/AC5 encoded protein (Li *et al.,* 2015), pre-coat protein (Zrachya *et al.,* 2007; Chowda-Reddy *et al.,* 2008; Glick *et al.*, 2008; Fukunaga and Doudna, 2009; Sharma and Ikegami, 2010; Sharma *et al.*, 2010; Zhang *et al.,* 2012; Wang *et al.*, 2014), βC1 encoded protein (Yang *et al.,* 2011; Shen *et al.,* 2012; Li *et al.,* 2014) and alphasatellite-encoded rep proteins (Nawaz-ul-Rehman *et al*., 2010) have been demonstrated to suppress post-transcriptional and/or RNA silencing in plants.

Objectives

Keeping all these informations available in the virus diversity, nature of infection and the counter-defense of viruses against host mechanisms, the following objectives were designed for the present study-

- To asses genetic diversity of begomoviruses associated with chilli leaf curl disease.
- To understand interaction among genomic components in leaf curl disease development.
- To elucidate role of pre-coat protein in host-virus interaction during viral pathogenesis.

2. Review of Literature

2.1. Chilli (*Capsicum* species)

Capsicum spp is an important economically important vegetable and spice crops cultivated throughout the world. Chilli is a self-pollinated plant with a high level of heterogeneity in seed population. Chilli plants are white flower with dark green leaves that grow up to 1.5 m in height. In India, Australia, and New Zealand, heatless species refers to "capsicums" while hot ones are termed as "chilli/chillies". In Mexico, "hot peppers" are called as chile, while ‘pimiento’ name indicates the heatless varieties. The name "pepper" came into existence because of the hotness in the same sense as the condiment black pepper, *Piper nigrum*. Chilli is an important source for the preparation of curries and pickles.

2.2. Chilli cultivation

In corresponds to their medicinal properties, the economic importance of this crops is also significant. Hence, the acreage and production of chilli is gradually increasing globally and placed at the level of 1.5 million hectares and 7 million tonne (NABARD), respectively. Among the countries cultivating chillies, Asian countries account for 50% of the total productivity worldwide and India is the leading producer in terms of area and production which is followed by China and Pakistan. Egypt, Ghana, Nigeria and Tunisia in Africa; Argentina, Brazil, Mexico, Peru and USA in America; Bulgaria, Hungary, Italy, Romania, Spain and Yugoslavia in Europe are the major countries growing chillies in their respective continents. The worldwide chilli consumers include India, China, Germany, Mexico, Thailand, UK and USA (http://www.nabard.org/modelbankprojects/chilli.asp).

In Indian sub-continent, chilli has been cultivated throughout the year in both kharif and rabi seasons. In India, according to National Horticultural Board database for 2013-14 season, the estimated chilli cultivation area is 7.75 lakh ha and chilli production is 14.92 lakh tonne (productivity 1.9 MT/Ha) (http://nhb.gov.in/area-pro/NHB_Database_2015.pdf). The chilli spices account for 16% of the total spice trade in the world and positions next only to black pepper. Among all the spices cultivated in India during 2013-14 season, chilli holds the first position (25.26%) followed closely by

garlic (21.19%) and turmeric (20.14%) (Figure 2.1). The major chilli growing states in the country includes Andhra Pradesh, Karnataka, Madhya Pradesh, Maharashtra, Orissa, Rajasthan, Tamil Nadu and West Bengal. Andhra Pradesh alone contributes 50% of chilli productivity in the country (Figure 2.2).

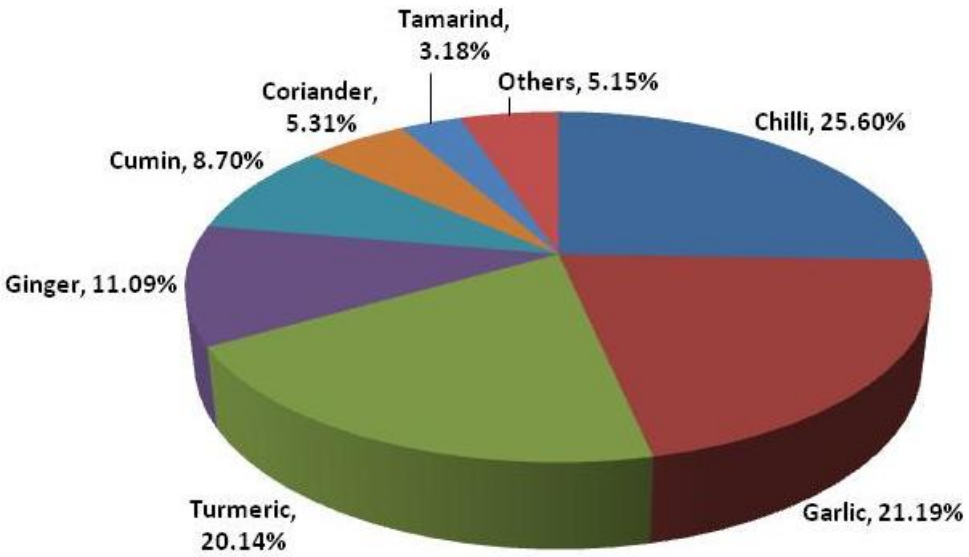


Figure 2.1. Crop cultivation of major spices in India during 2013-14 season (http://nhb.gov.in/area-pro/NHB_Database_2015.pdf).

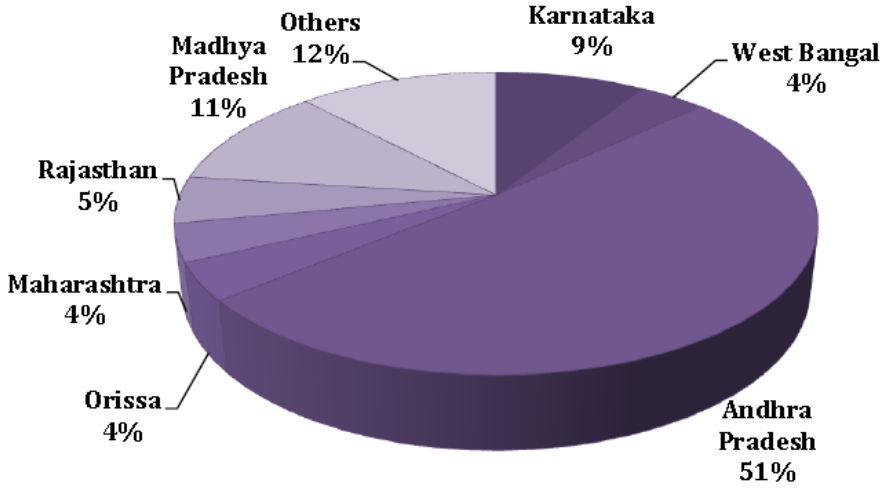


Figure 2.2. Distribution of chilli production among various states in India.

2.3. Chilli leaf curl disease (ChiLCD)

The average chilli productivity in India has been reduced in recent years. The main reason for this decreased productivity is crop loss due to pests and pathogens. Though chilli crops are attacked by several pathogens such as bacteria and fungi, viruses have emerged as a serious threat. Among them, at present, the viruses causing leaf curl disease of chilli (ChiLCD) are posing a major constraint to the cultivation of chillies in our country. Chilli plants affected with ChiLCD bears typical symptoms such as upward or downward leaf curling, leaf rolling, leaf blistering, vein thickening, vein banding, stunting etc. In severe condition, the crop bears no fruits and ultimately, dies. Despite the incidence of ChiLCD was noticed several decades ago (Husain, 1932), the identification of the causal viruses reported a few decades ago (Mishra *et al.*, 1963; Dhanraj and Seth 1968), however the association of begomoviruses with this disease has recently been demonstrated (Chattopadhyay *et al*., 2008).

2.4 The causal organism

2.4.1 Family: *Geminiviridae*

Geminiviridae family includes a group of plant infecting circular ssDNA viruses of ~2.8 kb in size. They infect and cause substantial crop losses in a large number of economically important vegetable and food crops, ornamental plants and fibre crops worldwide (Varma and Malathi, 2003). Based on their nature of twinned icosahedral particles, the term 'Geminivirus' was coined several decades ago (Harrison *et al.,* 1977).

The geminate particles have been reported to be packed in two incomplete icosahedra containing 22 capsomeres of ~18 X 30 nm. Earlier studies suggest that two incomplete icosahedra are attached at a waist and also found twisted by 20^0 (Bottcher *et al.,* 2004). Purified geminate virions do not have any empty cells suggest that the whole structure is found to be stabilized by the interaction of DNA-protein and is concomitantly assembled. Geminivirus genomes consist of one or two circular ssDNA molecules. However, within one twin particle, only one molecule of ssDNA can be encapsidated thus, for a bipartite genome, DNA-A and DNA-B are carries into two twin particles (Jeske, 2009).

2.4.1.1 Taxonomy

At genus level, the International Committee on Taxonomy of Viruses (ICTV) has classified geminiviruses based on insect vector, genome organization, host range and genome-wide pairwise sequence identities into seven genera as *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus* (Varsani *et al.*, 2014).

2.5 Genome organization of geminiviruses

2.5.1 *Becurtovirus*

Beet curly top Iran virus and *Spinach curly top Arizona virus* are the two recognized species in this genus. Unlike other geminiviruses, the members of this genus possess unique nona-nucleotides (TAAGATTCC), genome organization with a spliced replication initiator protein and species demarcation threshold value for this genus has been fixed as 94% (Varsani *et al.*, 2014). These viruses infect dicot plants like *B. vulgaris*, *V. unguiculata*, *S. lycopersicum* and *P. vulgaris* in Iran (Yazdi *et al.*, 2008; Soleimani *et al.*, 2013; Gharouni Kardani *et al.*, 2013) and *S. oleracea* in Arizona (Hernandez-Zepeda *et al.*, 2013).

2.5.2 *Begomovirus*

The type species of this genus is *Bean golden mosaic virus* (BGMV). This is considered as the largest genus in the *Geminiviridae* family. Begomoviruses are classified either as monopartite (contain only DNA-A like molecule) or bipartite (with DNA-A and DNA-B molecules). The insect vector, whiteflies (*Bemisia tabaci* Genn.) are reported to transmit these viruses and they also infect mostly of dicot hosts (Varma and Malathi, 2003; Brown *et al*., 2012). According to genome organization and phylogenetic segregation, these begomoviruses are divided into two regions: ‘Old world’ that includes Africa, Asia, Australia and Europe; ‘New world’ that includes America, Brazil and Mexico (Rybicki, 1994; Nawaz-ul-Rehman and Fauquet, 2009). However, the probable centre of origin for these devastating pathogens is found to be around South East Asia (Nawaz-ul-Rehman and Fauquet, 2009). DNA-A component encodes 6 ORFs, 2 ORFs in the virion strand and 4 in the complementary strand. Virion strand encodes coat protein (AV1) and pre-coat protein (AV2) which helps in virion encapsidation. Complementary strand encoded ORFs

include replication associated protein (C1; Rep), transcription activator protein (C2; TrAP), replication enhancer protein (C3; REn) and a poorly characterized protein (C4) (Brown *et al*., 2012). DNA-B component found only in bipartite viruses encodes a movement protein (BC1; MP) and a nuclear shuttle protein (BV1; NSP) which helps in intra- as well as inter-cellular viral movement. DNA-A component of bipartite begomoviruses from 'New world' is found to lack AV2 ORF and hence, depends entirely on DNA-B encoded proteins for movement (Rybicki, 1994; Varsani *et al.,* 2014). The virion and complementary strand transcription units in DNA-A and DNA-B are separated by a highly conserved non-coding region named common region (CR) (Lazarowitz and Shepherd, 1992). CR contains a nona-nucleotide region (TAATATTAC) forming stem loop structure which act as a cleavage site for Rep to initiate viral replication (Laufs *et al.,* 1995). Virus-specific conserved region (iterons) called Rep-binding sites are located upstream to these stem loop structure (Fontes *et al.,* 1994a, b). Monopartite begomoviruses are often found to be in association with satellite molecules namely, betasatellites and alphasatellites. Both these satellites encode a single ORF in the complementary and virion strand, respectively (Nawaz-ul-Rehman and Fauquet, 2009).

2.5.3 *Curtovirus*

Beet curly top virus is the type species for *Curtovirus* with the size of nearly 3 kb (Stanley *et al.,* 1986). They are transmitted by the insect vector, beet leafhopper (*Circulifer tenellus*) and infect dicotyledonous hosts (Varsani *et al.,* 2014). Like *Mastrevirus*, they are also monopartite viruses but they share less sequence homology between them. Their genome includes 4 ORFs in virion strand, 3 ORFs in complementary strand and short intergenic region. Virion strand encodes a coat protein, ss/ds DNA regulator protein and a putative movement protein whereas complementary strand encodes a REP, Ren and pathogenicity associated protein and a protein involved in cell division (Hormuzdi and Bisaro, 1995).

2.5.4 *Eragrovirus*

These viruses encode a coat protein which resemble to the members of *Mastrevirus*, but the replication proteins are closely related to the viruses belonging to the genus, *Begomovirus. Eragrostis curvula streak virus* is the only member in this genus infecting a monocot species (*E. curvula*) in South Africa (Varsani *et al.,* 2009).

Figure 2.3. Genome organization of members of the genus included in *Geminiviridae* family. The abbreviated names are- MSV-*Maize streak virus*, ECSV-*Eragrostis curvula streak virus*, BCTIV-*Beet curly top Iran virus*, BCTV-*Beet curly top virus*, TPCTV-*Tomato pseudo curly top virus*, TCTV-*Turnip curly top virus*, TYLCV-*Tomato yellow leaf curl virus* and BGMV-*Bean golden mosaic virus*. This figure is adapted from Varsani *et al.* (2014).

They code for four ORFs and the origin of replication sequence is found unusual (TAAGATTCC) (Varsani *et al.*, 2014). Similar to *Becurtovirus*, the species demarcation threshold for this genus has been fixed as 94% (Varsani *et al.*, 2014).

2.5.5 *Mastrevirus*

Type species for this genus is *Maize streak virus*, a monopartite virus and is transmitted by leafhoppers. Mastreviruses are 2.7-2.8 kb in size and are known to infect both

monocot and dicot hosts. These viruses are mostly confined to 'old world' infecting mainly monocots such as maize, sugarcane and wheat (Kraberger *et al.*, 2014), whereas 5 distinct dicot infecting mastreviruses are mostly circulating in Australia (Varsani *et al.*, 2014). Recently, *Chickpea chlorotic dwarf virus*, a dicot infecting mastrevirus has been reported only from Middle East, India and Africa (Kraberger *et al.*, 2014). The genome of this viruses possess a nona-nucleotides (TAATATTAC) which is quite similar to the one noticed among other Geminiviruses (Heyraud *et al.*, 1993). Its genome encodes 2 ORFs each from virion and complementary strand which is separated by long intergenic region (LIR) and short intergenic region (SIR) (Kammann *et al.*, 1991). LIR and SIR contain the origin of replication for the synthesis of virion and complementary strand, respectively.

Virion strand encoded proteins are necessary for viral movement and encapsidation; whereas replication associated proteins are encoded in the complementary strand. Unique to this genus, these viruses are known to regulate their own gene expression through a post-transcriptional splicing event within monocots (Rojas *et al.*, 2005). Recently, an unusual combination of mastrevirus-betasatellite-alphasatellite complex has been identified from wheat in India (Kumar *et al*., 2014).

2.5.6 *Topocuvirus*

A monopartite virus, *Tomato pseudo curly top virus* of ~ 3 kb in size is the type member of this genus. This virus encodes 2 ORFs in the virion strand and 4 ORFs in the complementary strand. It is transmitted to dicot plants in the 'New World' by the vector, treehoppers (*Micrutalis malleifera*). Based on the genome organization, this virus seems to a recombinant between *Mastrevirus* and *Begomovirus* (Briddon *et al.*, 1996).

2.5.7 *Turncurtovirus*

Only member in this genus includes *Turnip curly top virus* reported either from *B. rapa* or *R. sativus* (Briddon *et al.*, 2010). This phylogenetically distinct member is most closely resembles to the members of genus, *Curtovirus*. Its genome encodes 6 rather than 7 proteins. It contains nona-nucelotides (TAATATTAC) similar to *Mastrevirus, Begomovirus, Curtovirus* and *Topocuvirus*. For *Turncurtovirus,* the strain demarcation threshold of 95% has recently been assigned (Varsani *et al.*, 2014).

2.6 Function of begomovirus components

2.6.1 DNA-A component

The proteins encoded by DNA-A component are involved in replication, gene expression, suppression of host's silencing machinery, movement and encapsidation. DNA-A component of Old world begomoviruses encodes 7 ORFs which includes 5 complementary strand and 2 virion strand ORFs, however, New world begomoviruses encodes 6 ORFs with the absence of AV2 ORF. Five complementary strand ORFs includes AC1, AC2, AC3, AC4 and AC5, whereas virion strand ORFs includes AV1 and AV2. AC1 is also called as replication associated protein/Rep which is very much essential for virus replication (Elmer *et al.,* 1988). Though Rep protein does not possess any similarity with known polymerases, it does share some similarity with the replication initiator proteins of bacterial plasmids which undergo replication through rolling circle mechanism (Ilyina and Koonin, 1992). AC2 ORF encodes a transcription activator protein (TrAP), a multi-functional protein which is mainly involved in the transactivation of late viral genes such as AV1 in DNA-A component and BV1 in DNA-B component and several host genes (Sunter and Bisaro, 1992; Haley *et al.,* 1992; Trinks *et al.,* 2005; Rajeswaran *et al.,* 2007).

Various begomoviruses encoded TrAP proteins have been shown to possess transcriptional gene silencing (TGS) and/or post-transcriptional gene silencing (PTGS) suppressor activity (Dong *et al.,* 2003; Yang *et al.,* 2007; 2012). Replication enhancer protein (REn) encoded by AC3 ORF can form homo-oligomers and also binds to Rep protein (Pasumarthy *et al.,* 2010). In addition, REn protein is also known to assist Rep protein in viral replication by interacting with various proteins such as pRBR and PCNA (Settlage *et al.,* 2001; Castillo *et al.,* 2003). AC4 encoded protein has been reported to suppress RNA silencing (Chellappan *et al.,* 2004; Vanitharani *et al.,* 2004; Fondong *et al.,* 2007). It also interacts with a defence associated protein, Shaggy-like kinase (SK) and a cell-cycle regulator protein, related to KPC1 (RKP) (Dogra *et al.,* 2009; Lai *et al*. 2009). Coat protein (AV1) and pre-coat protein (AV2) is encoded in virion strand and mainly involved in intra- as well as inter-cellular movement, viral encapsidation and insect transmission (Boulton *et al.,* 1989; 1993; Liu *et al.,* 1997; 1999).

2.6.2 DNA-B component

The DNA-A and DNA-B sequences have been found to be divergent except a 100-200 nucleotide region (common region; CR). CR is an important determinant for the replication of this genome by DNA-A encoded Rep protein. DNA-B component encodes 2 ORFs, one each in virion and complementary strand which are mostly involved in viral movement (Brown *et al*., 2012). BC1 is a movement protein which is involved in intra- and inter-cellular movement of viral molecules, whereas BC1, a nuclear shuttle protein assists in transport of viral components between nucleus and cytoplasm (Lazarowitz and Shepherd, 1992). BV1 protein is known to be localized to plasma membrane (Pascal *et al.,* 1993; 1994), however, BC1 has been reported to interact with plasmodesmata and thereby increasing the size exclusion limit of plasma membrane (Noueiry *et al.,* 1994). BC1 is also known to be a pathogenicity determinant and with BV1, it plays a vital role in determining host range as well (Pascal *et al.,* 1993; Noueiry *et al.,* 1994).

2.6.3 Betasatellites

Most of the reported begomoviruses has been characterized as bipartite viruses (containing DNA-A and DNA-B molecules), in which both the molecules are needed for successful systemic infection (Brown *et al.,* 2012; Zhou, 2013). Subsequently, begomoviruses from Old world such as *Tomato leaf curl virus* (ToLCV) and TYLCSV have been reported to possess only a DNA-A like component lacking DNA-B component (Kheyr-Pour *et al.,* 1991; Navot *et al.,* 1991). These monopartite viruses have been evolved by associating themselves with satellite molecules (Dry *et al.,* 1997; Saunders *et al.,* 2000). These satellite molecules (called as betasatellites) are half the size (1.3 kb) of the helper components and encode a single ORF in the complementary strand (βC1). These betasatellites requires helper begomoviruses for their replication, encapsidation, insect transmission and systemic spread (Briddon and Stanley, 2006). Recently, a novel DNA motif has been elucidated in contributing to the selective replication of the betasatellite by a helper begomovirus-encoded rep protein (Zhang *et al.,* 2015).

In monopartite begomoviruses, these betaatellites-encoded βC1 proteins have been reported as a pathogenicity determinant and helps in the establishment of disease complex (Saunders *et al.,* 2000; Cui *et al.,* 2005; Briddon and Stanley, 2006). It has been demonstrated to be localized both in nucleus and cytoplasm (Cui *et al.,* 2005) and also in

chloroplast (Bhattacharyya *et al*., 2015). In addition, it has been reported to possess TGS and PTGS silencing suppression activity (Shen *et al.,* 2012; Yang *et al.,* 2012). The βC1 protein is known to bind DNA and interact with SNF1-related kinase (SnRK) and S-adenosyl homocysteine hydrolase (SAHH) (Cui *et al.,* 2005; Shen *et al.,* 2012; Yang *et al.,* 2012). Further, βC1 protein also reported to induce the expression of a host calmodulin-Like protein (rgs-CaM) which leads to the reduced expression of RDR6. This in turn reduces the generation of virus derived secondary siRNA (Li *et al.,* 2014). A recent study demonstrated that a tobacco-encoded RING finger protein (NtRFP1) mediates the degradation of βC1 protein encoded by TYLCCNV-associated betasatellite. This study also reports that plants over-expressing NtRFP1 developed attenuated symptoms upon virus infection (Zhou, 2013; Shen *et al*., 2016).

2.6.4 Alphasatellites

Alphasatellites are circular ss-DNA molecule of 1.4 kb in size which is mainly found in association with monopartite begomovirus-batasatellite complexes in Old world (Mansoor *et al.,* 1999; Briddon *et al.,* 2004; Xie *et al.,* 2010). However, few alphasatellites are also reported to be associated with New world begomoviruses (Paprotka *et al.,* 2010; Romay *et al.,* 2010). Additionally, a new class of satellite DNAs has been found to be associated with a bipartite begomovirus (Fiallo-Olivé *et al.,* 2012). These alphasatellites encode a single ORF in virion strand. Hence, these alphasatellites can replicate independently in the infected cells, however, it depends on the helper viruses for their insect transmission and systemic spread (Saunders and Stanley, 1999). These alphasatellites encoded rolling circle initiator protein does not share any homology with the helper component encoded Rep protein; instead it shares a significant similarity with the Rep protein encoded by nanoviruses (Saunders and Stanley, 1999). The functional significance of this protein is unknown. These alphasatellites has been shown to ameliorate disease in the infected host by decreasing the accumulation of betasatellite molecules (Wu and Zhou, 2005; Idris *et al*., 2011). In addition, rolling circle initiator protein of two cotton associated alphasatellites has been shown to possess RNA silencing activity (Nawaz-ul-Rehman *et al.,* 2010).

2.7. Functions of genes encoded by DNA-A

2.7.1 Replication associated protein (Rep)

Rep protein encoded by geminiviruses is highly conserved in their structure, architecture and function, which is expressed under bidirectional core promoter positioned in the IR region (Hanley-Bowdoin *et al.,* 2013). This protein is indispensible for viral replication and additionally posses various roles such as transcriptional regulation and combating host's anti-viral silencing machinery. Because of these functional roles, this protein relates them with some bacterial plasmids and ssDNA viruses, emphasizing their evolutionary origin with that of prokaryotes (Stenger *et al.,* 1991; Ilyina and Koonin, 1992; Oshima *et al.,* 2001).

Multi-functional Rep protein is 41 kDa in size and possesses various domains such as DNA binding and cleavage domains in the N-terminal region and oligomerization domain in its central region, whereas ATPase domain with Walker A and B motifs resides in the C-terminal region (Orozco *et al.,* 1997) (Figure 2.4). N-terminal region consists of three conserved motifs: Motif I (FLTY) is necessary for specific DNA binding, motif II (HLH) possesses a metal-binding site which is necessary for DNA cleavage and protein conformation (Orozco *et al*., 1998; Gutierrez, 1999). Motif III (YxxKD/E) includes the catalytic site for DNA nicking and binds to cleaved DNA with the help of the conserved Tyr residue (Laufs *et al.,* 1995; Orozco *et al*., 1998). Further, this motif III is also shown to be necessary for DNA ligation by Rep proteins (Orozco *et al.,* 1997). ATPase and helicase activities of the Rep protein have been shown to be around the oligomerization domain in its C-terimal region (Figure 2.4). Further, this helicase activity was found to be associated with either dodecameric or higher oligomeric conformation (Choudhury *et al.,* 2006; Clerot and Bernardi, 2006). Viruses with mutation in either ATP-binding or in Walker motifs do not failed to replicate in yeast or plant cells (Choudhury *et al.,* 2006). The unwinding activity of Rep protein requires oligomerization, ATP-binding and ATP-hydrolysis which further helps in viral DNA replication (Choudhury *et al.,* 2006; Clerot and Bernardi, 2006). In addition, *Mungbean yellow mosaic India virus* (MYMIV) Rep protein has been shown to have topoisomerase activity, however, the role of this activity in virus replication is not known (Pant *et al.,* 2001). Recently, a highly conserved motif

named Geminivirus Rep Sequence (GRS) is identified and shown to be involved in the initiation of the viral replication (Nash *et al*. 2011).

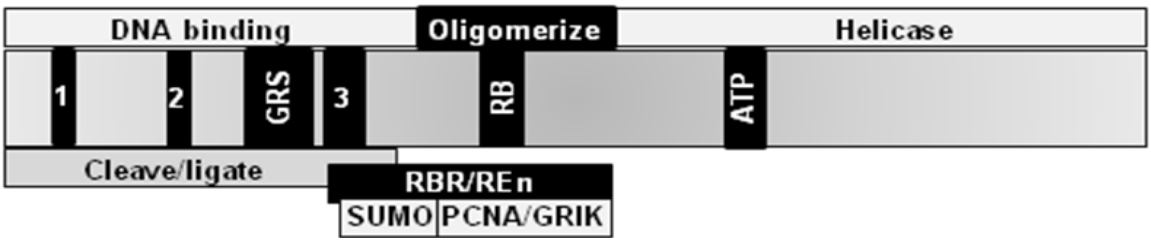


Figure 2.4. Schematic representation of the Rep protein with functional domains. The abbreviated names are: GRS-Geminivirus rep sequence, RBR-retinoblastoma-related protein, REn-replication enhancer, PCNA-proliferating nuclear antigen and GRIK-Geminivirus rep interacting kinase. This figure is adapted from Fondong (2013).

2.7.1.1 DNA-A and DNA-B replication

Geminivirus replication begins with the synthesis of complementary strand using ss-DNA as the template by various unknown host factors. Rep is one of the early viral proteins which recognize the origin of replication (ori) to initiate viral replication (Lazarowitz *et al.,* 1992; Fontes *et al.,* 1994a, b; Laufs *et al.,* 1995; Orozco *et al.,* 1997). The ori includes a large cis-acting region with one (in begomovirus) or more (in mastrevirus) Rep-binding sites (Lazarowitz *et al.,* 1992; Fontes *et al.,* 1992, 1994; Sanz-Burgos and Gutiérrez, 1998; Castellano *et al.,* 1999). These sites are direct repeat sequence which is positioned between TATA box and transcription start site. Prior to this sequence specific binding, Rep proteins form oligomers (Sanz-Burgos and Gutiérrez, 1998; Castellano *et al.,* 1999).

A high affinity artificial zinc finger protein which can bind TYLCV-Rep binding site has been reported to affect TYLCV-Rep binding and subsequently, its replication (Mori *et al.,* 2012). The Rep oligomerization occurs non-specifically i.e., two heterologous Rep proteins can oligomerize without affecting the efficiency of virus replication (Settlage *et al.,* 1996). Rep oligomers are known to be accumulated in the nucleoplasm and involved in the replication of only its cognate and not the heterologous DNA-B component (Lazarowitz *et al.,* 1992; Krenz *et al.,* 2011). Mutation in Rep oligomerization domain hampers virus replication (Orozco *et al.,* 2000).

A unique Rep-binding site (CGGTGT) in MYMIV-ori is found to be located on both sides of the stem loop structure. At initial stage, when the Rep is present at low concentration, the four-iteron sites of MYMIV-IR are occupied and show co-operative binding to four iterons with increasing concentrations of Rep protein (Singh *et al.,* 2008). This induces IR distortion which subsequently necessitates the initiation of the replication. To initiate rolling circle replication (RCR), these DNA bound Rep proteins creates a nick in the phosphodiester bond between the last T and A plus strand within the highly conserved stem-loop structure containing the nona-nucleotide region (TAATATTAC). Rep protein has been demonstrated to possess a conserved NTP-binding motif (Hanson *et al.,* 1995; Rizvi *et al*., 2015). However, ATP does not require for this nicking process, but it can bring conformational change in Rep protein to enhance the cleavage activity (Pant *et al.,* 2001). After nicking, the Rep protein remains occupied to 5′ end of the cleaved DNA strand through a conserved Tyr residue (Arguello-Astorga *et al.,* 1994; Orozco *et al*., 1998; Rizvi *et al*., 2015). After a round of replication, Rep catalyzes nucleotidyl transfer reaction to release a circular ss-DNA genome (Orozco *et al.,* 1997; Steinfeldt *et al.,* 2006). Further, these newly synthesized genomes undergo either addition round of replication or encapsidation for systemic spread.

2.7.1.2 Satellite DNA replication

Tomato leaf curl virus satellite DNA (TLCV sat-DNA), alphasatellites and betasatellites are three types of satellite molecules known to be associated with begomoviruses (Briddon & Stanley, 2006). TLCV sat-DNA does not have any ORFs but known to possess similar nonanucleotides and Rep-binding sites with its cognate DNA-A component, TLCV (Dry *et al.,* 1997). However, these Rep-binding sites are not necessary for either TLCV or TLCV sat-DNA (Lin *et al.,* 2003). Further, the region necessary for their replication has been mapped to 296–335 nucleotides. Begomoviruses-associated alphasatellites can replicate independently, but requires their helper components only for their systemic spread and encapsidation (Saunders and Stanley, 1999).

Unlike alphasatellites, betasatellites depends entirely on their helper begomoviruses for replication, movement and encapsidation. These betasatellites possess a nona-nucleotide region which is quite different from the nona-nucleotides of begomoviruses (Saunders *et al.,* 2000). Unlike the cognate Rep-binding sequence determines the specificity of Rep

proteins, the trans-replication of betasatellites is known to be highly diverse i.e, it can be replicated by diverse begomoviruses including non-cognate ones (Lin *et al.,* 2003; Qing and Xhou, 2009; Patil and Fauquet, 2010; Ranjan *et al*., 2014). However, the exact reason for this promiscuous replication of betasatellites is still obscure. The mechanism behind this less stringent interaction between these satellite ori and Rep is not elucidated. A pyrimidine-rich region and G-box from the promoter of different betasatellites has been shown to regulate the replication of betasatellites (Eini *et al.,* 2009; Zhang *et al.,* 2012). Though the interaction between betasatellite encoded βC1 and Rep protein is known, the functional significance of this interaction in the trans-replication of betasatellites requires further investigation.

2.7.1.3 Role in gene expression

Apart from its role in virus replication, Rep is also involved in transcriptional regulation of self expression through binding to the repeat sequences in CR region (Sunter *et al.,* 1993; Eagle *et al.,* 1994). Hence, binding to these sequences can either initiate replication or repress its own transcription. In addition, a conformational change is also necessary for such transcriptional repression (Orozco *et al.,* 2000). In the Rep protein of *Tomato yellow leaf curl Sardinia virus* (TYLCSV), an RGG sequence has been reported to be involved in this transcriptional repression (Sardo *et al.,* 2011). However, the various host factor(s) involved in this transcriptional repression is not known, but both these processes are reported to occur independent of each other (Eagle *et al.,* 1994). *Wheat dwarf virus* (WDV) RepA has been shown to activate the virion sense gene expression (Collin *et al.,* 1996). Transcription repression of *Tomato golden mosaic virus* (TGMV) Rep protein has been reported to alter the expression of TrAP and REn (Shung and Sunter, 2007).

2.7.1.4 Inhibits TGS and PTGS activities

Rep proteins can suppress TGS activity by reducing the expression of plant DNA methyl transferase (NbMET1) and Chromomethylase (NbCMT3) (Rodriguez-Negrete *et al.,* 2013). Recently, WDV Rep protein has been shown to suppress PTGS by inhibiting both local and systemic silencing (Wang *et al.,* 2014b).

2.7.1.5 Interaction with host factors

Geminivirus Rep proteins are known to interact with various cellular factors to modulate cell cycle, to assist in virus replication and to combat host defense response. The following host proteins are reported to interact with various Rep proteins:

2.7.1.5.1 Plant retinoblastoma-related protein (RBR)

It is known that geminiviruses infect already differentiated cells and hence, it needs to reprogram such type of cells for their replication. Geminivirus encoded Rep proteins are known to interact with RBR protein to reprogram the cell cycle. WDV RepA interacts with RBR through its conserved LXCXE motif, but other geminiviral Rep proteins do not possess this motif (Xie *et al.,* 1995). However, TGMV Rep protein interacts with RBR through RBR-binding domain which resides between the amino acid residues 101 and 180 of the Rep protein. Rep residues K144, A147, L148, I151 and KEE146 of TGMV play a significant role in this TGMV Rep and RBR interaction (Kong *et al.,* 2000; Arguello-Astorga *et al.,* 2004). Similar to RepA, the exogenous application of RBR3 enhance the genes related to host DNA replication including PCNA which will be further utilized by these viruses for their replication (Sabelli *et al.,* 2009).

2.7.1.5.2 Proliferating cell nuclear antigen (PCNA)

Two different Rep proteins have been documented to interact with PCNA (Castillo *et al.,* 2003; Bagewadi *et al.,* 2004). However, ATPase and DNA nicking activities of MYMIV Rep have been shown to be inhibited by PCNA (Bagewadi *et al.,* 2004). During later stage of virus life cycle, this inhibition is believed to be utilized by the virus itself to control the level of its genome in the infected cell.

2.7.1.5.3 Replication factor C (RFC)

WDV Rep protein can interact with the large subunit of wheat RFC1 (Luque *et al.,* 2002). Prior to bind to 3′end of DNA, RFC1 is known to form a complex with PCNA and it also helps in PCNA loading (Mossi & Hubscher, 1998; Gomes *et al.,* 2001). By this way, Rep recruits RFC1 to form a complex with PCNA for their replication.

2.7.1.5.4 Replication protein A (RPA) and minichromosome maintenance protein 2 (MCM2)

Both the subunits RPA32 and RPA70 are known to form trimeric complex with RPA14 and assist in replication by recruiting DNA polymerase α, PCNA and RFC (Loor *et al.,* 1997). However, MYMIV Rep is known to interact only with RPA32 subunit and this interaction cause a negative impact in the nicking and positive impact in the activities on the MYMIV Rep protein *in vitro* (Singh *et al.,* 2008). MYMIV-Rep is also demonstrated to interact with one of the components of MCM complex, MCM2 (Suyal *et al.,* 2013b). MYMIV replication efficiency is impaired in yeast and *Arabidopsis* mcm2 mutants suggesting the positive role of MCM2 in virus replication. Though both MYMIV Rep and MCM proteins are known to possess helicases activity (Bochman and Schwacha, 2008; Suyal *et al.,* 2013b), their combined effect in viral replication is not yet elucidated.

2.7.1.5.5 Histone H3 and RAD51/54

Cabbage leaf curl virus (CbLCV) and TGMV Rep has been reported to interact with histone H3 (Kong and Hanley-Bowdoin, 2002). Though the importance of this interaction is not known, *Abutilon mosaic virus* (AbMV) has been associated with nucleosomes and assembles into minichromosomes (Pilartz and Jeske, 2003). The oligomerization domain of MYMIV Rep also interacts with recombination/repair proteins, RAD51 and RAD54 in its N-terminal region, and emphasized their importance in the virus DNA replication (Kaliappan *et al.,* 2012; Suyal *et al.,* 2013a). RAD54 has been shown to increase the nicking, ATPase and unwinding activities of MYMIV Rep protein *in vitro* (Kaliappan *et al.,* 2012). Further, RAD51D also has been shown to promote *Euphorbia yellow mosaic virus* infection infections (Richter *et al*., 2016).

2.7.1.5.6 Other cellular factors

Tomato yellow leaf curl disease associated Rep protein has been demonstrated to bind to SUMO-conjugating enzyme (SCE-1) and this interaction is necessary for virus replication (Sanchez-Duran *et al.,* 2011). NAC-domain containing proteins such as AtNAC083; geminivirus RepA binding 1 (GRAB1) and GRAB 2 are known to interact with MYMIV and WDV Rep proteins, respectively (Xie *et al.,* 1999; Suyal *et al.,* 2014). TGMV Rep is known to interact with a Ser/Thr kinase named as geminivirus Rep-interacting kinase (GRIK) and a kinesin designated as geminivirus Rep-interacting motor protein (GRIMP)

(Kong and Hanley-Bowdoin, 2002). Interaction of Rep with GRIK might protect these kinases from degradation, whereas the interaction with GRIMP helps Rep to prevent the entry of mitotic phase. Recently it has been demonstrated that Tomato yellow leaf curl disease associated Rep protein binds to β-tubulin 6 and hypothesized the role of microtubules in the movement of the Rep protein (Fondong, 2013). However, the exact mechanistic action and functional importance of these interactions with Rep protein is still not clear.

2.7.2 Transcription activator protein (TrAP)

The mono-directional promoter for the expression of TrAP is located in the coding region of the AC1 gene and TrAP protein is well characterized as a multi-functional protein (Shivaprasad *et al.,* 2005). ORF AC2 of AbMV has been shown to act as a replication brake (Krez *et al.,* 2015). Various functional roles of this protein are noted below:

2.7.2.1 Gene activation

TrAP has been shown to be an ssDNA binding protein with a zinc-binding activation domain and hence, acts as a transcriptional activator (Haley *et al.,* 1992; Sunter and Bisaro, 1992; Hartitz *et al.,* 1999; Shivaprasad *et al.,* 2005) (Figure 2.5). This protein has been shown to trans-activate viral late genes such as coat protein (in DNA-A) and movement protein (in DNA-B) in a non-specific manner. This activity depends on the zinc finger, activation domain and its ability to form oligomers (Sunter *et al.,* 1994; Saunders and Stanley, 1995; Hartitz *et al.,* 1999; Yang *et al.,* 2007; Hanley-Bowdoin *et al.,* 2013). Though TrAP has been reported to be a ssDNA binding protein, it does not bind to dsDNA and so, it relies on the host factors to activate the promoter of target genes (Hartitz *et al.,* 1999; Lacatus and Sunter, 2009).

Figure 2.5. Schematic representation of the AC2 protein (TrAP) with functional domains. The abbreviated names are: NLS-nuclear localization signal and PCD-programmed cell death. This figure is adapted from Fondong (2013).

TrAP of CbLCV and TGMV has been shown to bind to a plant specific DNA binding protein (AtTIFY4B) which is involved in cell cycle arrest (Chung and Sunter, 2014). Through this interaction, TrAP might prevent cell cycle arrest and provides a suitable environment for viral replication. For transcription activation, TrAP protein enters into the nucleus by interacting with the nuclear import protein, karyopherin α (Chandran *et al.*, 2012). Either transient or stable expression of different geminivirus encoded TrAP proteins have been reported to alter several host genes including sugar and jasmonate responsive genes (Trinks *et al.*, 2005; Soitamo *et al.*, 2012). Recently, TrAP protein has been shown to induce expression of the *Arabidopsis* calmodulin-like protein 39 (AtrgsCaM), a known regulator of RNA silencing and binds to it in the nucleus, thereby getting avoided from degradation by AtrgsCaM (Chung *et al.*, 2014). However, in contrast to AC2 of begomoviruses, curtovirus C2 (BCTV) is not a transcription activator and it cannot be called as TrAP (Hormuzdi and Bisaro, 1995).

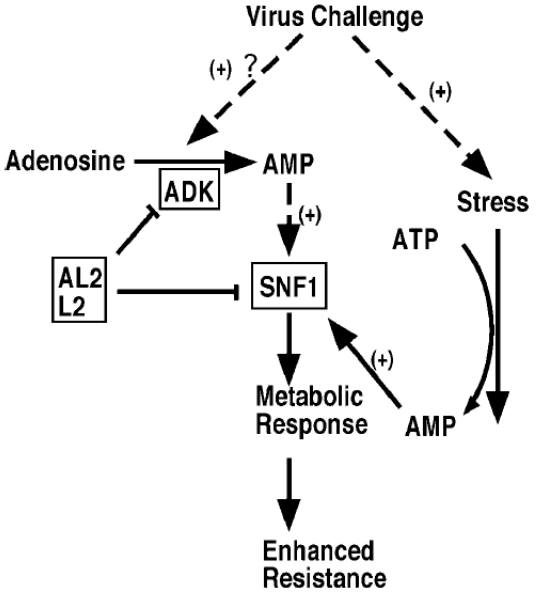


Figure 2.6. Model depicting the inhibitory role of C2/L2/AL2 in stress response. Abbreviated names are: ADK-adenosine kinase and SNF1-sucrose nonfermenting1. This figure is adapted from Hao *et al.* (2003).

2.7.2.2 Pathogenicity factor

As a pathogenicity protein, TrAP of *Tomato leaf curl New Delhi virus* (ToLCNDV) has been shown to counter nuclear shuttle protein-induced hypersensitive reaction (Hussain *et al.*, 2007). TrAP protein interacts with COP9 signalosome (CSN) complex and destabilize

SCF mediated ubiquitination. By this, TrAP regulates several defense related pathways such as jasmonic acid pathway, sulphur assimilation pathway and thereby modulating the host's defense response against these viruses (Lozano-Duran *et al.,* 2011; 2012). Further, Hao *et al.,* (2003) identified that C2/AC2 binds and inhibits sucrose non-fermenting 1 (SNF1) kinase from *Arabidopsis thaliana*, a regulator of enhanced resistance which justifies the enhanced susceptibility upon the transient expression of C2/AC2 (Figure 2.6). In addition, phosphorylation of TrAP by SnRK1 protein has been reported to be responsible for delayed CbLCV infection in *Arabidopsis* (Shen *et al*., 2014). Recently, TrAP of TYLCSV has been elucidated to suppress jasmonate-mediated defence responses (Rosas-Diaz *et al*., 2016).

2.7.2.3 Suppressor of RNAi silencing

Most of the TrAPs encoded by monopartite and bipartite begomoviruses are evidently demonstrated as suppressor of either TGS and/or PTGS (Dong *et al.,* 2003; Vanitharani *et al.,* 2004; Trinks *et al.,* 2005; Wang *et al.,* 2005). TGMV (a begomovirus) and *Beet curly top virus*-BCTV (a curtovirus) have been shown to suppress PTGS by inhibiting the adenosine kinase activity which negatively affects methyl cycle (Wang *et al.,* 2003; 2005). Adenosine kinase is known to be essential for production of s-adenosyl methionine, an important co-factor for methyl transferases (Saze *et al.,* 2003). The AC2-coded protein of MYMIV has been demonstrated to suppress RNA silencing by inhibiting the activities of AtRDR6 and AtAGO1 (Kumar *et al*., 2015). Similarly, to function as a PTGS suppressor, AC2 protein of CbLCV promotes decapping activity of DCP2 which in turn accelerates mRNA turnover rate and also inhibits the siRNA accumulation (Ye *et al.,* 2015).

In addition, C2/AC2 of the members of both begomovirus and curtovirus has been shown to be a suppressor of TGS (Buchmann *et al.,* 2009; Zhang *et al.,* 2011; Yang *et al.,* 2012). Buchmann *et al.,* (2009) demonstrated for the first time that geminivirus C2 and AC2 proteins can suppress TGS and thereby reduce the overall cytosine methylation of the viral genome. *Beet severe curly top virus* (BSCTV) acts as a TGS suppressor by interacting with S-adenosyl-methionine decarboxylase 1 (SAMDC1) and attenuating the degradation of SAMDC1, a key player in the methyl cycle (Zhang *et al.,* 2011). Later, BSCTV C2 also has been shown to affect the generation of virus derived small interfering

RNAs (siRNAs), a precursor for the initiation of RNA directed DNA methylation (RdDM) and thereby reducing the viral DNA methylation (Yang *et al.,* 2012). In this context, demonstration of increase in the virus susceptibility upon transient expression of this protein in tobacco is understandable (Sunter *et al.,* 2001; Hao *et al.,* 2003). Recently, AC2 of *Indian cassava mosaic virus* has been reported to inhibit kryptonite (KYP, a H3K9 methyl transferase) via the activation of transcription repressor RAV2 (RELATED TO ABI3 and VP1) (Sun *et al.,* 2015). However, AC2 protein of TGMV and CbLCV has been shown to interact with the catalytic domain of KYP and further inhibits its methyl transferase activity *in vitro* (Castillo-Gonzalez *et al.,* 2015). Furthermore, using TrAP protein lacking its transcription activation domain, a recent report revealed that this TrAP could reverse TGS in the reproductive plants, independent of ADK inactivation or transcription activation (Jackel *et al*., 2015).

2.7.3 Replication enhancer protein (REn)

Similar to ORF AC2, the expression of ORF AC3 is also driven by the mono-directional promoter located in the AC1 coding region (Shivaprasad *et al.,* 2005). The roles of REn protein in virus life cycles is given below:

2.7.3.1 Virus replication

Though AC3 is not essential for replication, mutant AC3 displays attenuated symptoms and reduced viral DNA accumulation suggesting its accessory role in virus replication (Sunter *et al.,* 1990; Etessami *et al.,* 1991; Morris *et al.,* 1991). Interaction between AC1 and AC3 has been demonstrated in different geminiviruses (Settlage *et al.,* 1996) (Figure 2.7). Etessami *et al*. (1991) using AC3 mutants suggest that AC3 is not essential for synthesis of the coat protein of virus assembly but it can modulate the viral DNA accumulation. The role of REn in virus replication is confirmed by its interaction with AC1 and PCNA (Castillo *et al.,* 2003). Like AC1, without any LXCXE motif, AC3 is also reported to bind to maize RBR suggesting its role in cell cycle regulation (Settlage *et al.,* 2001) (Figure 2.9).

2.7.3.2 Interaction with other cellular proteins

Similar to RepA, this protein also interacts with SlNAC1 which could possibly involved in virus DNA replication (Selth *et al.,* 2005). AC3 of *Tomato leaf curl Kerala virus*

(ToLCKeV) is reported to form oligomers, interacts with Rep protein and further, enhances the ATPase activity of Rep protein (Pasumarthy *et al.,* 2010). Using phage display method, ToLCKeV AC3 has been shown to interact with various cellular proteins, especially, the proteins that are involved in RNA silencing machinery (Pasumarthy *et al.,* 2011).

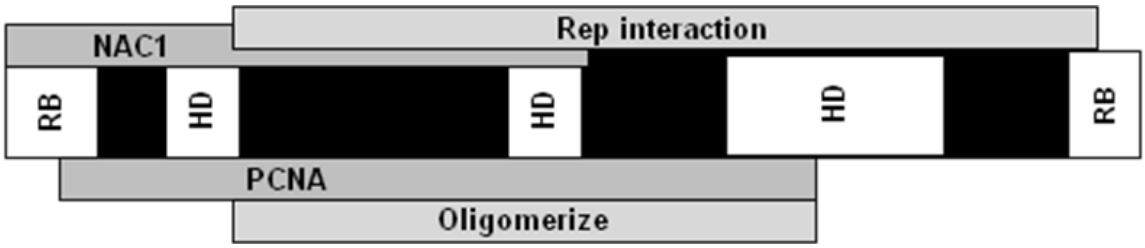


Figure 2.7. Schematic representation of the various domains of AC3/REn proteins. The abbreviated names are: RBR-retinoblastoma related protein, HD-hydrophobic regions and PCNA-proliferating cell nuclear antigen. This picture is adapted from Fondong (2013).

2.7.4 AC4 protein

AC4 ORF is located entirely within the AC1 coding region but in a different reading frame, and this protein is absent in mastreviruses. The viral DNA accumulation was reported to be at low level when test plants were infected with viruses lacking AC4 ORF (Jupin *et al.,* 1994; Teng *et al.,* 2010). N-terminal myristoylation site has been reported to be involved in the delivery of viral DNA to plasmodesmata and also helps in cell-to-cell movement (Rojas *et al.,* 2001). Though AC4 protein of different begomoviruses has been shown to assist in the movement viruses (Pooma and Petty, 1996), however, it is found that this protein is not critical to virus infection (Elmer *et al.,* 1988; Hoogstraten *et al.,* 1996; Bull *et al.,* 2007). The AC4 protein of *Cotton leaf curl Kokhran virus* (CLCuKoV) has been shown to possess ATPase and inorganic pyrophosphatase activities (Guha *et al.,* 2013). Residues such as Arg13 and Cys8 are reported to be the catalytically important residues for these activities. In addition, these enzymatic activities have been shown to get enhanced in the presence of AV2 protein (Guha *et al.,* 2013).

2.7.4.1 Pathogenicity determinant

The AC4 protein has been shown to be responsible for the vein swelling phenotype upon inoculation with BCTV (Stanley *et al.,* 1986; Mills-Lujan and Deom, 2010). In addition,

severe phenotypic abnormalities of transgenic AC4 plants were reported which mimic the symptoms of virus infected plants and impaired cell division (Latham *et al.,* 1997; Piroux *et al.,* 2007; Mills-Lujan and Deom, 2010). Together, it is clear that AC4 is a pathogenicity determinant for various begomoviruses. The various studies have reported that this protein is responsible to induce the gene expression of various cell differentiation regulators such as RKP, ATHB7 and ATHB12 (Lai *et al.,* 2009; Park *et al.,* 2011). This suggests that this protein plays a vital role in cell division and cell development. Recently, C4 protein of BCTV has been reported to induce hyperplasia and also shown to interact with various shaggy-like protein kinases to promote its own phosphorylation and activation, to control cell cycle (Mills-Lujan *et al.,* 2015).

2.7.4.2 RNA silencing suppression

Several monopartite and bipartite begomoviruses encoded AC4 proteins are involved in the suppression of RNA silencing (Chellappan *et al.,* 2004; Fondong *et al.,* 2007). To suppress gene silencing, AC4 protein of cassava mosaic viruses bind either to single stranded form of both microRNAs (miRNAs) and siRNAs (Chellappan *et al.,* 2004). Fondong *et al.,* (2007) suggested that the myristoylation site in the N-terminal region is necessary for this silencing suppression activity and membrane localization. Further, AC4 of ToLCV-Au has been shown to interact with shaggy kinase protein through its C-terminal domain (Piroux *et al.,* 2007; Dogra *et al*. 2009).

2.7.5 AC5

The AC5 ORF is located in the complementary strand, downstream of AC3/C3 ORF and overlaps with the AV1 ORF. This ORF has been reported to be common in MYMIV isolates (Ilyas *et al.,* 2010) and using yeast system, its role in the virus replication was also demonstrated (Raghavan *et al.*, 2004).

The non-essential role of AC5 in virus replication of *Tomato chlorotic mottle virus* (ToCMoV) and *Watermelon chlorotic stunt virus* (WmCSV) has been reported (Kheyr-Pour *et al.,* 2000; Fontenelle *et al.,* 2007). However, role of AC5 in the replication of MYMIV and some of ToLDeV also has been elucidated (Melgarejo *et al.,* 2013). The AC5 protein of MYMIV has shown to be a pathogenicity determinant and also reported to be a suppressor of both PTGS and TGS (Li *et al*., 2015).

2.7.6 Coat protein (CP)

Coat protein, encoded by V1/AV1 ORF, is the only structural protein that forms geminiviral particles (Stanley and Gay, 1983). The viruses with mutated or replaced CP have been reported to produce delayed and attenuated symptoms including reduced ssDNA level with or without altering the level of dsDNA (Briddon *et al.,* 1989; Sunter *et al.,* 1990; Padidam *et al.,* 1996). Besides encapsidation, the role of this protein in insect transmission and intra- as well as inter-cellular movement of viral DNA is described below:

2.7.6.1 Nucleo-cytoplasmic shuttling

The central domain of this protein has been shown to be involved in its multimerization which is also needed for viral capsid assembly and importantly, in insect transmission (Noris *et al.,* 1998; Kheyr-Pour *et al.,* 2000; Hallan and Gafni, 2001; Zhang *et al.,* 2001) (Figure 2.8). This protein can non-specifically bind to ss- and ds-DNA in a co-operative manner, and sequesters ssDNAs form the replication cycle (Saunders *et al.,* 1991; Stenger *et al.,* 1991; Liu *et al.,* 1997). This ssDNA-coat protein binding has been reported to resist nucleolytic cleavage and also maintains stability even at high salt concentration (Palanichelvam *et al.,* 1998).

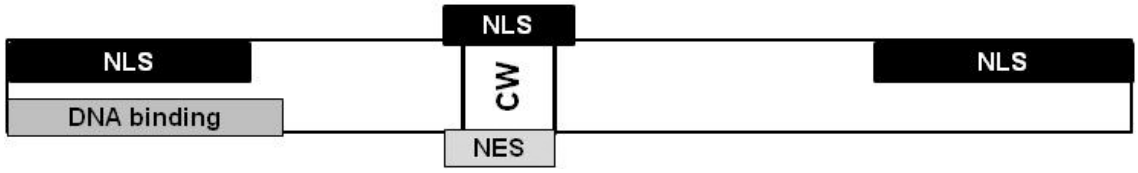


Figure 2.8. Schematic illustration of the various domains of coat protein. The abbreviated names are: NLS-nuclear localization signal, CW-cell wall targeting motif and NES-nuclear export signal. This picture is adapted from Fondong (2013).

CP of CLCuKoV-Dabawali has been shown to possess a conserved C2H2 type zinc finger motif and mutation in this motif affected the strength of CP-DNA binding (Priyadarshini and Savithri, 2009). The ability of CP in nucleo-cytoplasmic shuttling makes it a functional replacement for NSP encoded by DNA-B component for viral movement (Ingham *et al.,* 1995; Qin *et al.,* 1998). However, nucleo-cytoplasmic trafficking mainly depend either on its binding to ssDNA with the help of N-terminal domain or its interaction with importin α/ karyopherin α1, a member of nuclear-pore complex (Kunik *et al.,* 1999; Guerra-Peraza *et al.,* 2005; Pitaksutheepong *et al.,* 2007).

Further, it also possess nuclear localization signal (NLS) in its N-terminal, central and C-terminal regions (Unseld *et al.,* 2001) (Figure 2.8). Besides NLS, a leucine-rich nuclear export signal (NES) has been mapped to its central domain (Ward and Lazarowitz, 1999).

2.7.6.2 Nucleolus localization

CP of both the monopartite and bipartite begomoviruses has also been shown to be accumulated in the nucleolus (Rojas *et al.,* 2001; Fondong, 2013). Several other DNA/RNA viral proteins that are localized to nucleolus can play vital roles in viral infection such as virion formation and systemic spread (Kim *et al.,* 2007; Herranz *et al.,* 2012). In this context, the role of nucleolar localization of CP needs further investigation.

2.7.6.3 Insect transmission

N-terminal domain of CP has been found to interact with GroEL protein of *Arsenophonus*, a bacterial endosymbiont located in the midgut of whitefly (Kunik *et al.,* 1999; Yaakov *et al.,* 2011; Rana *et al.,* 2012). This interaction has been speculated to protect virions in whitefly haemolymph. The CP amino acid residues between 123 and 152 have been mapped to be essential for insect transmission (Noris *et al.,* 1998; Kheyr-Pour *et al.,* 2000). In addition, a change in even one or two amino acid(s) in this region has been shown to have impaired insect transmission (Noris *et al.,* 1998; Kheyr-Pour *et al.,* 2000).

2.7.6.4 Cell-to-cell movement

Coat protein of mastreviruses and monopartite begomoviruses has been reported to be involved in inter-cellular movement of viral molecules (Briddon *et al.,* 1989; Padidam *et al.,* 1996). Further, in monopartite begomoviruses, this function has been presumed through its interaction with another protein encoded by AV2 ORF (Priyadarshini *et al.,* 2011). However, CPs of bipartite begomoviruses are not required for virus infection, it is demonstrated to involve in the cell-to-cell movement and systemic spread of viruses (Fondong, 2013).

2.7.7 Pre-coat protein

The Old World begomoviruses are known to code V2 ORF, but the New World bipartite begomoviruses lack AV2 ORF (Rybicki, 1994; Brown *et al.,* 2012). The viruses with

mutant AV2 ORF replicate inefficiently and accumulate at reduced level of viral DNAs and impaired systemic infection (Padidam *et al.,* 1996; Selth *et al.,* 2004; Mubin *et al.,* 2010).

2.7.7.1 **Virus movement**

V2 of TYLCV and BCTV has been reported to be a pathogenicity determinant, whereas V2 of mastreviruses are involved in intra-cellular movement (Boulton *et al.,* 1989; Stanley *et al.,* 1992; Wartig *et al.,* 1997). This protein has been demonstrated to enhance the nuclear export of TYLCV DNA to cell periphery. Further, AV2 protein has been reported to be localized both in cytoplasm as well as in the nuclear periphery (Rojas *et al.,* 2001; Rothenstein *et al.,* 2007). Priyadarshini *et al* (2009) have demonstrated the interaction of CLCuKoV-Dabawali encoded CP and pre-coat protein and suggest their combined role in systemic movement. However, Hak *et al* (2015) suggest that V2 protein of TYLCV-Israel strain is not needed for the virus movement.

2.7.7.2 Pathogenesis

A putative protein kinase C motif was indentified in *East Africa cassava mosaic Cameroon virus* (EACMCV) and its role in viral pathogenesis has been suggested (Chowda-Reddy *et al.* 2008). Pre-coat proteins of two monopartite begomoviruses have been elucidated to induce hypersensitivity response which was countered by the addition of TrAP protein (Mubin *et al*., 2010). The AV2 of TYLCV has been shown to interact and inhibit the enzymatic reaction of CYP1 protein, a papain-like cysteine protease (Bar-Ziv *et al.,* 2012; Bar-Ziv *et al.,* 2015). Recently, Moshe *et al* (2016) reported the aggregation of V2 protein of TYLCV and this self-aggregation has been subjected to the integrity of cytoskeleton and its binding to the viral DNA.

2.7.7.3 Suppressor of RNA silencing

V2/AV2 ORF of several monopartite and bipartite begomoviruses has been elucidated as a suppressor of RNA silencing, especially PTGS (Zrachya *et al.,* 2007; Chowda-Reddy *et al.,* 2008; Glick *et al.,* 2008; Sharma and Ikegami, 2010; Sharma *et al.,* 2010; Zhang *et al.,* 2012). The interaction of TYLCV V2 with SUPPRESSOR OF GENE SILENCING 3 (SGS3) and cysteine amino acids at 84th and 86th positions are found to be essential for PTGS suppression (Glick *et al.,* 2008). In order to suppress gene silencing, V2 protein

has also been reported to sequester siRNAs (Zhang *et al.,* 2012). Further, it has been suggested that V2 protein competes with SGS3 and possibly reduces the production of secondary siRNAs (Fukunaga and Doudna, 2009). Recently, AV2 of TYLCCNV has been shown to suppress methylation mediated TGS and substitution at amino acid residues at 84 and and 86 did not affect this TGS suppression (Wang *et al.*, 2014a).

2.8 Functions of genes encoded by DNA-B

2.8.1 Nuclear shuttle protein (BV1)

BV1 encoded by the bipartite begomoviruses are also called as nuclear shuttle protein (NSP). The DNA-B component is known to code NSP ORF in the virion-sense strand. In mastreviruses, CP is reported to function as NSP.

2.8.1.1 Virus movement

NSP protein binds non-specifically to viral DNA and facilitates the transport of ssDNA between cytoplasm and nucleus by forming viral DNA-NSP complex (Noueiry *et al.*, 1994; Ward and Lazarowitz, 1999). The NLS of *Squash leaf curl virus* (SqLCV) has been mapped to N-terminal region containing the motif SLEKDLLIDLH which resembles the NLS of other nuclear shuttle proteins (Pascal *et al.,* 1994; Ward and Lazarowitz, 1999) (Figure 2.9). NSP has been reported to be localized in nucleus as well in nucleolus (Rojas *et al.,* 2001; Zhou *et al.,* 2011). For systemic movement, this viral DNA-NSP complex is trapped by movement protein (BC1). Further, this whole complex moves to plasmodesmata for transporting into the adjacent cell. In the adjacent cell, NSP in the complex guides viral DNA into nucleus for further replication (Noueiry *et al.,* 1994; Sanderfoot and Lazarowitz, 1995; 1996; Ward *et al.,* 1997).

The C-terminal region of SqLCV BV1 protein has been reported to be involved in interaction with BC1 protein (Sanderfoot and Lazarowitz, 1995). As both NSP and CP assist in virus movement, they should have common evolutionary origin by sharing redundant functions. For nucleo-cytoplasmic shuttling, the CbLCV-encoded NSP interacts with an *Arabidopsis* nuclear shuttle protein (AtSN1) which acetylates the NSP (McGarry *et al.,* 2003; Carvalho *et al.,* 2006). This led to the segregation of viral DNA for movement and encapsidation. NSPs also reported to be interacting with cytoplasmic

GTPase for the movement of viral DNA-NSP complex from nucleus to cytoplasm (Carvalho *et al.,* 2008). This protein also have been demonstrated to interact with histone H3 forming a complex of viral DNA-NSP-MP-H3 (Zhou *et al.,* 2011). This complex formation helps in the inter- and intra-cellular movement of viral DNA.

Figure 2.9. Various domains of the BV1/NSP proteins. The abbreviated names are: HR-hypersensitive response, AtNSI-*Arabidopsis thaliana* nuclear shuttle protein interactor, NLS-nuclear localization signal, NES-nuclear export signal and MP-movement protein. This picture is adapted from Fondong (2013).

2.8.1.2 Pathogenicity determinant

NSPs of several geminiviruses have been reported to induce hypersensitive response (HR) in various hosts (Garrido-Ramirez *et al.,* 2000). The role of NSP in virus pathogenicity has been documented based on its interaction with a membrane receptor kinase/leucine-rich repeat receptor-like kinase named NSP interacting kinases (NIK) (Fontes *et al.,* 2004; Rocha *et al.,* 2008). In a recent study, NIK1 has been documented to translocate RPL10 into the nucleus, where it interacts with L10-interacting MYB domain-containing protein (LIMYB). This RPL10-LIMYB interaction leads to the activation of the global translation suppression which has been reported further to use as an antiviral strategy in plants (Zorzatto *et al.,* 2015). In addition, the interaction of NSP with a proline-rich extensin-like receptor protein kinase (PERK) has been reported to phosphorylate and regulate the function of NSP (Fontes *et al.,* 2004; Florentino *et al.,* 2006).

2.8.2 Movement protein (BC1)

BC1 encoded by the bipartite begomoviruses are also called as movement protein (MP). It is encoded in the complementary-sense strand of the DNA-B component. However in monopartite begomoviruses, mastreviruses and curtovirus, AV2, V2 and V3 function as movement proteins (Mullineaux *et al.,* 1988; Boulton *et al.,* 1989; Hormuzdi and Bisaro, 1995; Fondong, 2013).

2.8.2.1 Inter-cellular trafficking

MP in co-operation with NSP is required for intercellular and systemic transport of viral DNA molecules (Noueiry *et al.,* 1994; Ward *et al.,* 1997) (Figure 2.10). BC1 protein binds to viral DNA-NSP complexes at the cytoplasmic face of plasma membranes/microsomal vesicles and assists their transfer to the nearby cells (Kleinow *et al.,* 2009). However, MPs of not all the members of geminiviruses reported to bind to DNA for example, MPs of SqLCV (Pascal *et al.,* 1994) and AbMV (Hehnle *et al.,* 2004) cannot bind whereas MPs of BDMV (Rojas *et al.,* 1998) and MYMIV (Radhakrishnan *et al.,* 2008) can bind to both ss- and ds-DNA. The co-ordinated role of CP, MP and C4 was elucidated to be responsible for the delivery of viral DNAs to the cell periphery (Rojas *et al.,* 2001). MPs of SqLCV and CaLCV have been shown to be predominantly associated with ER-derived tubules.

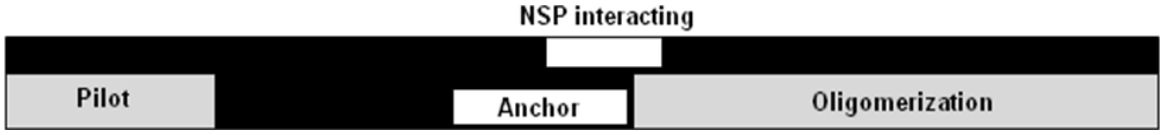


Figure 2.10. Schematic representation of the BC1/MP protein and its various domains. The nuclear shuttle protein is abbreviated as NSP and this is picture is adapted from Fondong (2013).

MPs of *Euphorbia mosaic virus* and *Indian cassava mosaic virus* have been demonstrated to be in co-ordination with microtubules for the viral genome movement (Roberts, 1989; Kim and Lee, 1992). CbLCV MP has been reported to interact with *Arabidopsis* regulator of endocytosis, synaptotagmin to dock the complex containing viral DNAs at plamodesmata for their intercellular locomotion (Lewis and Lazarowitz, 2010). Similarly, AbMV-MP interacts with heat cognate 70-kDa protein to guide the viral DNAs into the neighboring cells (Krenz *et al.,* 2010).

2.8.2.2 Pathogenicity determinant

Transgenic plants expressing the MPs of *Tomato mottle virus* (ToMoV) and BDMV have been reported to display virus infection like symptoms which suggest its role in pathogenesis. The region responsible for pathogenicity has been mapped indirectly to its C-terminus (Duan *et al.,* 1997; Hou *et al.,* 2000; Saunders *et al*., 2001).

2.9 RNA directed DNA methylation (RdDM)

RdDM pathway in plants prevents the transcription of unwanted factor such as repetitive elements, transposons as well as viruses. In this pathway, 24 nt short-interfering RNAs (siRNAs) bind to its complementary DNA loci and recruit various components involved in *de novo* DNA methylation to prevent the expression of target genes (Pumplin and Voinnet, 2013). This mechanism requires dsRNA for initiation by leading to generation of 24 nt siRNAs (Figure 2.11). This mechanism was first elucidated by Wassenegger *et al.*, (1994) using potato spindle tuber viroid RNAs which induces the DNA methylation of its homologous sequences. Unlike in animals, *de novo* methylation in plants includes the addition of methyl group in cytosines in all sequence contexts i.e., CG, CHG and CHH where H can be either A or T or C (Law and Jacobsen, 2010). The plant DNA viruses like geminiviruses, replicate as mini-chromosome in the host cell nuclei, thereby are exposed to RdDM machinery of the host plants (Hanley-Bowdoin *et al*., 2013). Further, the epigenetic marks have been found in the genome of geminiviruses and hence, this RdDM pathway is utilized by the plants to counter these virus infections (Dogar, 2006). Several proteins such as DNA methyl transferase, histone methyl transferase and chromatin remodelers were shown to be involved in histone and DNA methylation pathway.

2.9.1 Mechanism of RdDM pathway in plants

There are three major steps involved in this RdDM pathway culminating in the recruitment of components involved in *de novo* methylation of DNA strands. These steps require participation of RNA polymerases having different subunits with unique functions.

2.9.1.1 Step 1: Generation of Pol IV derived siRNAs

Apart from RNA polymerases I, II and III, *Arabidopsis thaliana* contains two additional plant-specific RNA polymerase subunits such as Pol IV and Pol V. In *Arabidopsis thaliana*, the functionally diversified Pol IV complexes comprise of the second largest subunit, Nuclear RNA polymerase D2a (NRPD2a) along with two distinct largest subunits NRPD1a and NRPD1b (Herr, 2005). The subunit compositions of the RNA polymerases present in *Arabidopsis thaliana* are provided in the Table 2.1. Interestingly, NRPD7A has been found only in Pol IV whereas NRPD7B/NRPE7B has been identified

in Pol IV as well as Pol V (Table 2.1). But NRPD7A/ NRPD7B/NRPE7B were absent in RNA pol II. However, the specific functional significance of the presence of these subunits with the particular RNA polymerases (Pol IV and Pol V) is not quite clear (Saze *et al*., 2012). The RNA polymerases required by RdDM machinery were RNA Pol VI and RNA Pol V (Figure 2.11). The initiation of RdDM involves the generation of long ssRNA transcripts from the transposons, repeats or viruses with the suggested help from RNA Pol IV (Zhang *et al.*, 2007; Pumplin and Voinnet, 2013). This key step also requires the important role of SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SHH1) which has been shown to bind to methylated H3K9 tails (Law *et al.*, 2011; 2013) (Figure 2.11). Further, SHH1 can directly recruit RNA Pol IV for di-methylation of histone H3 at K9 positions, leading to the generation of siRNAs (Du *et al*., 2015). Mutation in the largest subunit of Pol IV, NRPD1 has been reported to reduce the accumulation of 24 nt siRNAs (Huettel *et al.*, 2006). Kanno *et al.*, (2005) reported the precise role of Pol IVa either in generation or in amplification of small RNAs and the involvement of Pol IVb along with DEFECTIVE IN RNA DIRECTED DNA METHYLATION 1 (DRD1) in the downstream *de novo* DNA methylation. Furthermore, a chromatin remodeling factor (CLASSY1) has been shown to be involved in the initial stage of siRNA generation either at the level of Pol IV (step 1) or RDR2 (step 2) activity (Law and Jacobsen, 2010) (Figure 2.11).

2.9.1.2 Step 2: Generation of 24nt siRNAs and its binding to their target RNA

The Pol IV generated transcripts recruit RNA dependent RNA polymerase 2 (RDR2) for the conversion of these ssRNAs into dsRNAs (Figure 2.11). Also the physical association of RDR2 with Pol IV has been shown to be essential for the synthesis of dsRNAs by Pol IV (Law *et al.*, 2011; Haag *et al.*, 2012). However, Pol V and DRD1 are necessary for chromatin methylation and RNA silencing, but it is not essential for siRNA generation (Wierzbicki *et al*., 2008). The dsRNAs generated were subsequently processed by Dicer-like 3 (DCL3) enzyme and other RNAi components into 24 nt siRNAs (Law and Jacobsen, 2010). These siRNAs undergo the addition of a methyl group at their 3′ end which is mediated by HUA ENHANCER 1 (HEN1), a process called as siRNA maturation (Yu *et al*., 2005). The methylation of siRNA also prevents their rapid degradation by the host nucleases. The matured 24 nt siRNAs were then loaded into RNA induced silencing complex (RISC) containing Argonaute (AGO4) protein.

Table 2.1. Composition of the subunits present in RNA Pol II, IV and V of *Arabidopsis thaliana* plant.

RNA polymerase subunits	AGI number	Pol II	Pol IV	Pol V
NRPB1	At4g35800	Yes	No	No
NRPD1	At1g63020	No	Yes	No
NRPE1	At2g40030	No	No	Yes
NRPB2	At4g21710	Yes	No	No
NRPD2/NRPE2	At3g23780	No	Yes	Yes
NRPB3/NRPD3A/NRPE3A	At2g15430	Yes	Yes	Yes
NRPD3B/NRPE3B	At2g15400	No	Yes	Yes
NRPB4	At5g09920	Yes	No	No
NRPD4/NRPE4	At4g15950	No	Yes	Yes
NRPB5/NRPD5A	At3g22320	Yes	Yes	No
NRPD5B/NRPE5B	At2g41340	No	Yes	No
NRPE5	At3g57080	No	No	Yes
NRPB6A/NRPD6A/NRPE6A	At5g51940	Yes	Yes	Yes
NRPB6B/NRPD6B/NRPE6B	At2g04630	Yes	?	?
NRPB7	At5g59180	Yes	No	No
NRPD7A	At3g22900	No	Yes	Yes
NRPD7B/NRPE7B	At4g14660	No	Yes	Yes
NRPB8A/NRPE8A	At1g54250	Yes	?	?
NRPB8B/NRPD8B/NRPE8B	At3g59600	Yes	Yes	?
NRPB9A/NRPD9A/NRPE9A	At3g16980	Yes	Yes	Yes
NRPB9B/NRPD9B/NRPE9B	At4g16265	Yes	Yes	Yes
NRPB10/NRPD10/NRPE10	At1g11475	Yes	Yes	Yes
NRPB11/NRPD11/NRPE11	At3g52090	Yes	Yes	Yes
NRPB12/NRPD12/NRPE12	At5g41010	Yes	Yes	Yes

Yes indicates the 'presence' and no refers to 'absence' of the subunit. This table is adapted from Saze *et al*., 2012 and then modified accordingly.

The loading of siRNAs into AGO4 has been shown to be essential for the activation of RdDM complexes (Figure 2.13). Sometimes AGO6, 8 or 9 also functions similar to AGO4 in this RdDM pathway. This step involves RNA Pol V and AGO4 along with chromatin remodeling factors such as DDR complex which is composed of DRD1,

DEFECTIVE IN MERISTEM SILENCING 3 (DMS3) and RNA-DIRECTED DNA METHYLATION 1 (RDM1) (Figure 2.11). DRD1 has been reported to be involved in the search of DNA sequences targeted by the small RNAs for DNA methylation (Kanno *et al.* 2004). However, the chromatin remodeling activity of DRD1 has not been demonstrated. In addition, 24nt bound-AGO4 along with Pol IVa, RDR2, DCL3 have been shown to be localized to cajal bodies (Li *et al*., 2006) (Figure 2.11). Furthermore, SUPPRESSOR OF TY INSERTION 5-LIKE (SPT5L) protein has been demonstrated to interact with AGO4 (bound with 24 nt siRNAs) and the large subunit of RNA Pol V, Nuclear RNA polymerase E1 (NRPE1) (He *et al*., 2009; Wierzbicki *et al*., 2009) (Figure 2.11). This interaction (SPT5L-AGO4-NRPE1) suggests that SPT5L protein can guide 24nt siRNAs to its Pol V derived target transcripts which is necessary for the recruitment of downstream effector complex involved in *de novo* DNA methylation of target DNA (El-Shami *et al.,* 2007) (Figure 2.11). It has been reported that these Pol V derived transcripts too possess 5′ modification such as triphosphate or a 7meG cap structure without poly-A tail in its 3′ end. A RNA binding chromatin remodeling protein, INVOLVED IN *DE NOVO* 2 (IDN2), has been shown to bind to the RNA:siRNA duplex formed between siRNAs bound to AGO4 and Pol V derived non-coding transcripts (Ausin *et al*., 2009) (Figure 2.11). This binding may also act as a signal for the recruitment of downstream effector complex for DNA methylation. These events seem to be independent of histone methylation, however requires SUVH2 and SUVH9 (functionally inactive histone methyl transferases) which can bind to methylated DNA (Figure 2.11). This binding of SUVHs on the methylated DNA suggests that they have a key role in targeting Pol V (Du *et al*., 2015). Collectively, these steps were found to be required for maintaining a self-reinforcing loop between RdDM pathway and DNA methylation (Figure 2.12).

2.9.1.3 Step 3: Methylation of DNA regions targeted by 24 nt siRNAs

The final step in RdDM pathway includes the recruitment of components needed for DNA methylation (Figure 2.11). This includes the recruitment of *de novo* DNA methyl transferase, DOMAINS REARRANGED METHYL TRANSFERASE 2 (DRM2). It has been reported that 24 nt siRNA and Pol V transcript bound-AGO4 guides DRM2 to the target loci for *de novo* methylation of DNA strand (Zhong *et al*., 2014) (Figure 2.11). However, DRM1, a homolog of DRM2, has been reported to have a negligible role in this

methylation process (Cao and Jacobsen, 2002). In *Arabidopsis thaliana*, four types of DNA methyltransferases actively involved in RdDM pathway are DRM2, CHROMOMETHYLASE 2 (CMT2), CHROMOMETHYLASE 3 (CMT3), DNA METHYL TRANSFERASE 1 (MET1) (Du *et al*., 2015) (Table 2.2). MET1 involves in cytosine methylation at CG position, while CMT2 and CMT3 have been reported to be needed for methylating DNA at CHH and CHG positions, respectively. CMT2, a *de novo* DNA methyltransferase, preferentially to binds unmethylated DNA thereby catalyzing cytosine methylation at CHH and CHG positions (Stroud *et al*., 2014). CMT3 uses hemi-methylated DNA at CHG sites as substrate for methylating cytosines (Du *et al*., 2012). However, the cytosine methylation at CHG and CHG sites (by CMT2 and CMT3) has been reported to be linked with histone methylation such as H3K9me (Figure 2.12).

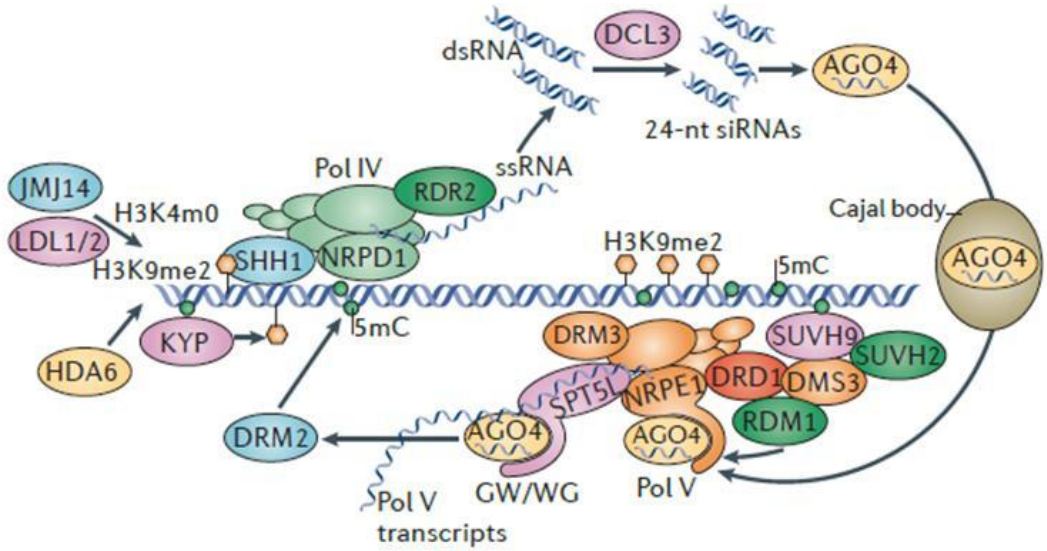


Figure 2.11. Processes involved in the RNA directed DNA methylation pathway of the plants. Various components involved in this mechanism are provided in the picture. The orange colored circles on the DNA strands refer to the methylated cytosine residues. This figure is adapted from Du *et al*., 2015.

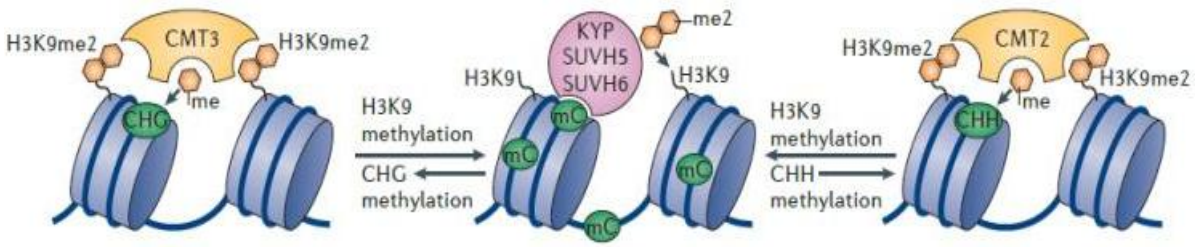


Figure 2.12. Schematic model depicting the self-reinforcing loop between histone and DNA methylation in plants. CMT2, CMT3 and KYP can bind to H3K9 dimethylated histones and, methylates DNA at CHH, CHG and all positions (CG, CHG and CHH), respectively. This figure is adapted from Du *et al*., 2015.

In addition, the role of other histone modifying enzymes such as HISTONE DEACETYLASE 6 (HDA6) and SU(VAR) 3-9 HOMOLOGUE 4 or KRYPTONITE (SUVH4/KYP) in RdDM mechanism has been elucidated (Aufsatz *et al.*, 2002; Jackson *et al.*, 2002; Malagnac *et al.*, 2002; Blevins *et al.*, 2014). The evidences from the recent studies prove that HDA6 functions upstream of Pol IV and generation of siRNAs (Aufsatz *et al.*, 2002; Blevins *et al.*, 2014). Liu *et al.*, (2012) reported the interaction between HDA6 and MET1 suggesting the involvement of histone deacetylation in the RdDM pathway. Another protein, SUVH4/KYP, has been reported to be required for the cytosine methylation especially at CHG position (Jackson *et al.*, 2002; Malagnac *et al.*, 2002). SUVH4/KYP is also identified as the major type of histone methyl transferase (H3K9Me) in *Arabidopsis* plants. The elimination of non-CG methylation (CHH or CHG) resulted in the loss of all H3K9 methylation and further, the elimination of all H3K9 methylation resulted in the loss of all non-CG methylation, which confirms the reinforcing loop model (Jackson *et al*., 2002; 2004) (Figure 2.12). Strikingly, MET1 has been demonstrated to methylate cytosines at CG position without the need of histone methylation (Stroud *et al*., 2013; Du *et al*., 2015).

2.9.2 Maintenance of DNA methylation in plants

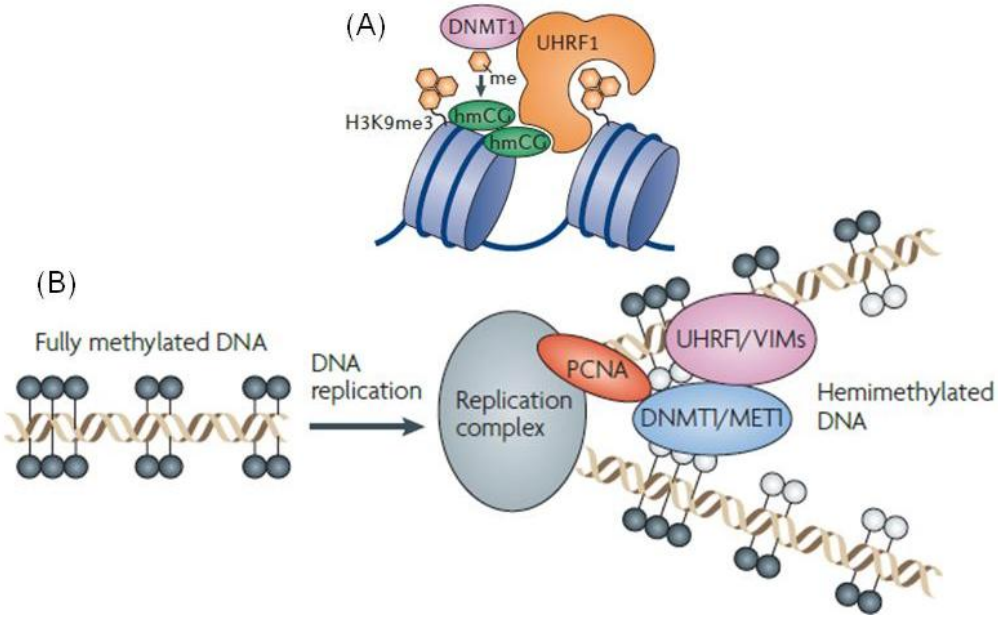


Figure 2.13. Maintenance mechanism of DNA methylation in animals (A) and plants (B). The role of UHRF1/VIM proteins in recruiting *de novo* methyl transferases (DNMT1/MET1) for the maintenance of cytosine methylation. This figure is adapted from Law and Jacobsen (2010).

The various proteins involved in the maintenance of cytosine methylation include hemi-methylated DNA binding proteins, DNA methyl transferase and other chromatin remodeling proteins (Table 2.2).

Table 2.2. Components involved in DNA methylation pathway of animals and plants.

Methylation mechanism	**Animals**	**Plants**
de novo		
CG	DNMT3L (DNMT3A, DNMT3B)	DRM2
CHG	-	CMT3
CHH	-	DRM2, CMT2
All positions	MIWI2 and MIL1	AGO4
Maintenance		
CG	DNMT1, UHRF1, LSH1	MET1, VIM, DDM1
CHG	-	SUVHs, CMT3
CHH	-	SUVHs, CMT2, DRM2
All positions	LSH1	DDM1

2.9.2.1 VARIANT IN METHYLATION (VIM)

The members of the protein family, VARIANT IN METHYLATION (VIM) or ORTHRUS (ORTH) are methyl-cytosine binding proteins that are postulated to be involved in maintenance of DNA methylation in plants (Figure 2.13). In *Arabidopsis thaliana*, six putative VIM proteins have been identified. The amino acid sequence of VIM proteins suggests that it contain a PHD domain (PLANT HOMODOMAIN) and a SRA domain (SET AND RING ASSOCIATED) separating two RING (REALLY INTERESTING GENE) domains. Liu *et al*. (2007) has demonstrated that this protein localizes into the nucleus. The SRA domain has been reported to be involved in the hetero-chromatinization of the centromeric regions (Woo *et al.,* 2007). This protein binds to methylated cytosines (at CG, CHG and CHH sites) through its SRA domain (Woo *et al.,* 2007) (Figure 2.13). In addition, binding of VIM protein with H2B, H3 and H4 has been elucidated (Woo *et al.,* 2007).

VIM protein is localized in the nucleus and involved in maintenance of chromatin structure during cell division. The transgenic plants over-expressing VIM1/ORTH2 has

been shown to cause inhibition of root growth and delayed flowering (Liu *et al*., 2007). Kraft *et al.* (2008) has demonstrated that both the RING domains in VIM protein possess E3 ubiquitin ligase activity *in vitro*. The members of the UBC8 family E2s has been demonstrated to be required for this protein to function as E3 ubiquitin ligase (Kraft *et al.*, 2008). Further, VIM protein has been demonstrated to bind tobacco Su(var), E(z) and Trithorax conserved (SET) domain protein NtSET1. The N-terminus of NtSET1 (SRA domain) or the C-terminus of NtSET1 (SET domain) and the PHD domain of VIM1 has been shown to be needed for their interaction (Liu *et al*., 2007). Recently, *Arabidopsis* VIM proteins have been shown to modulate both DNA methylation and histone methylation in co-operation with MET1, a *de novo* methyl transferase in plants (Kim *et al*., 2014) (Figure 2.13).

2.9.2.2 CG methylation maintenance in animals and plants

The maintenance of CG methylation in animals is performed by DNA METHYL TRANSFERASE 1 (DNMT1), which has been shown to be associated with the replication foci to completely methylate DNA from the hemi-methylated state after each round of replication (Figure 2.13). Its association with replication forks has been achieved through the interaction with PCNA, an essential regulator of DNA replication (Chuang *et al*., 1997). In addition, DNMT1 also has been reported to interact with a hemi-methylated DNA binding protein, UBIQUITIN-LIKE PLANT HOMEODOMAIN AND RING FINGER DOMAIN 1 (UHRF1) (Bostick *et al*., 2007; Sharif *et al*., 2007). Further, a chromatin remodeling factor, LYMPHOID-SPECIFIC HELICASE 1 (LSH1) has also been suggested to be necessary for this mechanism (Dennis *et al*., 2001). These interactions suggest that UHRF1 binds to hemi-methylated DNA and recruits DNMT1 to maintain the methylation status at the CG positions (Law and Jacobsen, 2010).

In plants, VIM proteins are predicted to act similar to UHRF1 and MET1 functions as DNA methyl transferase (Figure 2.13). In addition, a nuclear remodeler protein, DECREASE IN DNA METHYLATION 1 (DDM1), has been shown to behave similar to its mammalian counterparts, LSH1 (Table 2.2). In *Arabidopsis*, it has been elucidated that DNA methyl transferase (MET1) requires DDM1 for accessing histone H1 containing heterochromatin regions (Zemach *et al*., 2013). Furthermore, no significant change in CG methylation level was reported in the mutants of histone methyl transferases, thereby

suggesting that CG methylation by MET1 could be independent of histone methylation in plants (Stroud *et al*., 2013).

2.9.2.3 Non-CG methylation maintenance in plants

The interplay between histone methylation and cytosine methylation has been shown to play a major role in the maintenance of non-CG methylation in plants (Table 2.2) (Bernatavichute *et al*., 2008; Johnson *et al*., 2007). DNA methylation in CMT3 and KYP mutants were reported to be highly impaired, suggesting the involvement of histone H3 methylation with that of cytosine methylation (Jackson *et al*., 2002; 2004). The SRA domain of KYP protein has been shown to bind to methylated DNA at CHG position which revealed that cytosine methylation at CHG position leads to histone methylation (Lindroth *et al*., 2004). Furthermore, CMT3 has been demonstrated to bind to methylated histone H3 through its chromodomain, thereby establishing a cross-talk between histone and DNA methylation (Johnson *et al*., 2007). However, evidence linking interaction between DNA and histone methyl transferases i.e., KYP/SUVH4 and CMT3 is still awaited. Interestingly, a recent binding study underlines that CMT2 binds preferentially to H3K9me2, but CMT3 binds equally to both H3K9me1 and H3K9me2 (Stroud *et al*., 2014). This suggests that recruitment of either CMT2 or CMT3 probably depends on the presence of number of methyl groups on histone H3 tails.

2.10 RdDM as a defense mechanism against geminivirus infection

The ssDNA as genetic material makes geminiviruses prone to TGS mechanism, the major defense mechanism operating in plants to counter viral infection. Several studies using the mutants of various components of RdDM such as ADK, AGO4, DCLs, DDM1, DRB3, DRMs and KYP, reported high level of viral DNA accumulation leading to increased susceptibility to viruses, thereby highlighting the role of TGS as a plant defense against the invading viruses (Raja *et al*., 2008; 2013; Hanley-Bowdoin *et al*., 2013). In comparable to RNA viruses, the increased abundance of 24 nt virus specific siRNAs upon geminivirus infection has been widely reported (Chellappan *et al*., 2004; Zhang *et al*., 2005; Carrillo-Tripp *et al*., 2007; Yadav and Chattopadhyay, 2011). These 24 nt siRNAs were mainly shown to be generated from IR and promoters of Rep and CP (Rodriguez-Negrete *et al*., 2009; Yadav and Chattopadhyay, 2011 Hanley-Bowdoin *et al*., 2013). It has been shown that the heterogeneous viral DNA is methylated in the infected tissues

(Rodriguez-Negrete *et al*., 2009; Hanley-Bowdoin *et al*., 2013). Hence the possibilities of geminiviruses to combat this methylation (defense) process are by 'reviving' unmethylated DNA during viral DNA replication and/or through the action of geminivirus encoded proteins to inhibit this host mediated methylation pathway.

2.10.1 Viral DNA methylation and host recovery or symptom remission

Geminiviruses are not only the target of TGS mechanism, they also induces this host mediated indirect defense mechanism for obvious reasons. In some plants, the role of methylation machinery in the symptom recovery (remission) or resistance from geminivirus infections has been elucidated (Rodriguez-Negrete *et al*., 2009; Yadav and Chattopadhyay, 2011; Hanley-Bowdoin *et al*., 2013). This recovery from virus infection has been found to be associated with reduced accumulation of viral DNA and RNA and the presence of virus-derived 24 nt siRNAs leading to a high level of viral DNA methylation (Rodriguez-Negrete *et al*., 2009; Ceniceros-Ojeda *et al*., 2016). Further, two populations of viral minichromosomes have been reported in the geminivirus-infected plant showing recovery process. From the recovered plants, viral minichromosome was shown to have a compact structure associated with the presence of a chromatin-repressive marker and increased levels of viral DNA methylation (Ceniceros-Ojeda *et al*., 2016).

2.10.2 Viral factors countering TGS pathway

In order to counter this mechanism, several members of the geminiviruses have been reported to encoded proteins that target this TGS pathway at various stages (Hanley-Bowdoin *et al*., 2013). It is apt to say that this inhibition occurs at two stages: blocking the supply of methyl donors for methylation of viral DNAs and/or preventing subsequent DNA/histone methylation.

2.10.2.1 By affecting plant methyl cycle

Plants require universal methyl donor, S-adenosyl methionine (SAM), for successful establishment of DNA methylation (Hanley-Bowdoin *et al*., 2013). This methyl donor is generated through methyl cycle in which the initial precursor is decarboxylated form of SAM (dcSAM) by an enzyme, S-ADENOSYL METHIONINE DECARBOXYLASE 1 (SAMDC1). After regeneration, SAM is then converted via S-adenosyl homocysteine (SAH) which is followed by the conversion of homocysteine (Hcy) by the enzyme, S-

ADENOSYL HOMOCYSTEINE HYDROLASE 1 (SAHH1). Further, Hcy accepts a methyl group to form methionine and then to SAM by ATP. This methyl cycle is necessary for the continuous supply of methyl donor (SAM) for cytosine methylation (Figure 2.14). Two viral-encoded proteins, C2 and βC1, are known to interfere with this methyl cycle (Figure 2.14). C2 of curtovirus has been reported to prevent the proteasomal degradation of SAMDC1 enzyme, thereby preventing the decarboxylation of SAM (Zhang *et al*., 2007). Further, TrAP or C2 proteins have been shown to interact with the methyl cycle enzyme, ADENOSINE KINASE (ADK) to inactivate its enzymatic activity (Wang *et al*., 2005). In addition, βC1 of tomato yellow leaf curl disease associated betasatellite has been demonstrated to affect the host methyl cycle through its interaction with SAHH1 (Yang *et al*., 2011).

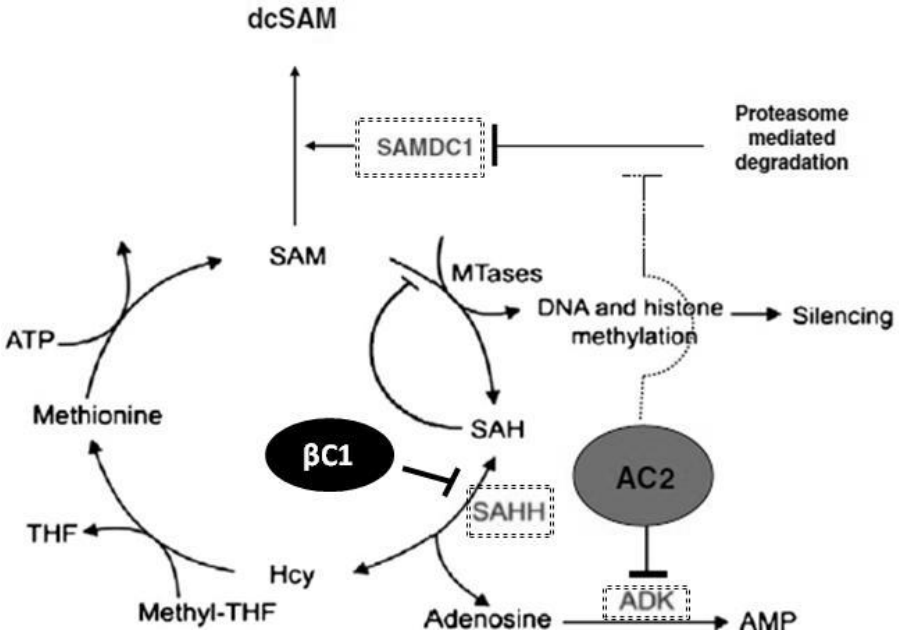


Figure 2.14. Begomovirus-encoded proteins (C2/TrAP and βC1) and inhibition of plant methyl cycle. This figure is modified from Hohn and Vazquez (2011).

2.10.2.2 By preventing subsequent DNA or histone methylation

In addition to their ability to interfere plant methyl cycle, it is eminent for the geminiviruses to block subsequent DNA or histone methylation. It is also widely reported that both cytosine and histone methylation functions as a self-reinforcing loop. In this regard, Rep and C4 proteins have been demonstrated to be down regulating the expression of MET1 and CMT3 genes, thereby, preventing the cytosine and/or histone methylation after every round of viral DNA replication (Rodriguez-Negrete *et al*., 2013).

Earlier study suggests that KYP2 mutant plants were highly susceptible to geminivirus infection, suggesting an important role between histone methylation and virus infection (Raja *et al*., 2008). Recently, TrAP has been shown to interact with the catalytic domain of KYP protein and thereby inhibiting its biological function of histone methyl transferase activity (Castillo-Gonzalez *et al*., 2015).

3. Materials and methods

3.1 Survey and collection of samples

An extensive survey was carried out in all major chilli-growing regions across almost all the states in India. For each collecting region, minimum of nine samples from at least three individual chilli fields were collected. These infected plant samples bearing typical begomovirus-like symptoms such as leaf curling, vein thickening, crinkling of leaves, vein banding, leaf curling and stunting were collected from the field. The details of the places and the GPS co-ordinates from where the plant tissues collected along with the types of symptoms noticed on the infected plants are provided in the Table 3.1. The samples were made available in the laboratory and these samples were used further for the virus characterization.

3.2 Genomic DNA isolation from the plant tissues

Plant genomic DNA was isolated as described by Dellaporta *et al*. (1983). Five hundred µg of leaf sample was ground using micro-pestles and 500 µl extraction buffer was added for homogenization. One hundred µl of 10% SDS was added to the homogenized solution and mixed by inverting the tubes. The microcentrifuge tube was then placed into 65^0C (preset) water bath for 20 min and mixed at regular intervals. The tubes were cooled at room temperature. Then 160 µl of 5M potassium acetate (pH 5.8) was added to the tubes and mixed the contents by inverting the tubes, followed by incubation on ice for 15 min. Then the sample was centrifuged at 13000 rpm for 10 min at 4^0C and the supernatant was collected in fresh microcentrifuge tubes. Equal volume of phenol:chloroform:isoamylalcohol (25:24:1) was added, the tubes were inverted for mixing the contents. The tubes were then centrifuged at 13000 rpm for 10 min at room temperature. Then the upper layer was collected in fresh microcentifuge tubes and 0.8 volume of isopropanol was added and kept at -20^0C from 20 min to 8 h. Microcentrifuge tubes were then spun at 13000 rpm for 20 min at 4^0C. The pellet was washed with 500 µl of 70% ethanol and the air dried for complete removal of ethanol. To the pellet, 50-100 µl of autoclaved distilled water for dissolving the DNA. The concentration of DNA was quantified by spectrophotometer.

Table 3.1. Symptom appearance on *Capsicum* spp collected from major chilli growing regions in India.

State	Place	Year	GPS coordinates	Symptom appearance*
Andhra Pradesh	Guntur	2009	16.30°N 80.45°E	LC, CR, LR, Vt
Assam	Jorehat	2009	26.44°N 94.10°E	LC, CR
Bihar	Chhapra	2010	25.78°N 84.72°E	LC, SL, LB, LR, Vt, St
	Patna	2008	25.36°N 85.7°E	LC, CR, LB, LR
Goa	Keri	2011	26.67°N 73.18°E	LC, CR, Vt
Gujarat	Ahmedabad	2009	23.03°N 72.58°E	LC, CR
Haryana	Pataudi	2007	28.35°N 76.75°E	LC, CR, Vt, St
Himachal Pradesh	Palampur	2009	32.11°N 76.53°E	LC, CR, Vt, St
Jharkhand	Dhanbad	2008	28.00°N 86.45°E	LC, CR
	Ramgarh	2008	30.78°N 76.55°E	LC, CR
Karnataka	Bengaluru	2010	13.81°N 77.29°E	LC, SL, YL
	Tumkur	2008	13.20°N 77.08°E	LC, LR, YL
Kerala	Vellanad	2008	8.34°N 77.30°E	LC, LR. CR
Madhya Pradesh	Jabalpur	2008	23.10°N 79.59°E	LC, SL, LB, LR, St
	Teherka	2008	24.42°N 78.52°E	LC, CR
Maharashtra	Nagpur	2009	21.8°N 79.5°E	LC, SL, LB, LR, Vt, St
	Pune	2009	18.31°N 73.51°E	LC, LR, YL
New Delhi	New Delhi	2010	28.36°N 77.12°E	LC, SL, LB, LR, Vt, YL, St
Orissa	Bhubaneswar	2009	20.14° N 85.50°E	LC, LB, YL
Rajasthan	Bharatpur	2009	27.22°N 77.48°E	LC, LR, YL
	Jodhpur	2009	26.18°N 73.04°E	LC, CR, Vt
Tamil Nadu	Coimbatore	2008	11.0°N 76.58°E	LC, CR, LB, Vt
	Salem	2008	11.44°N 78.4°E	LC, CR, LB, LR, Vt, St
	Villupuram	2008	11.95°N 79.52°E	LC, CR
Uttar Pradesh	Jaunpur	2007	25.44°N 82.41°E	LC, CR, LB, Vt, St
	Ghazipur	2010	25.35°N 83.34°E	LC, SL, LB, LR, St
	Kanpur	2008	26.28°N 80.21°E	LC, SL, LB, LR
	Noida	2007	28.57°N 77.32°E	LC, CR, Vt
	Varanasi	2006	25.20°N 83.0°E	LC, CR, Vt, St, YL
West Bengal	Baruipur	2008	22.20°N 88.26°E	LC, CR, LB
	Kalyani	2006	22.59°N 88.29°E	LC, CR, LB
	Kolkata	2008	22.34°N 88.22°E	LC, CR, LB
	Magrahat	2007	21.41°N 88.4°E	LC, SL, LB, LR

*CR: leaf crinckling, LB: leaf blistering, LC: leaf curling, LR: leaf rolling, SL: small leaves, St: stunting, Vt: vein thickening, YL: yellowing of leaves

3.3 Rolling circle amplification

The rolling circle amplification (RCA) was performed according to the method described in Inoue-Nagata *et al*. (2004). Prior to RCA, total DNA was incubated with RNase A (100 mg/mL) for 30 min at 37^0C to digest RNAs. These RNA free total DNA was extracted using phenol:chloroform mixture and then precipitated using isopropanol. From these purified genomic DNA, 500 ng DNA (<1 µl) was used for RCA using Templifi kit (GE, Chicago, USA) as per manufacturer's instructions. Five µl of sample buffer was then added and mixed gently by micropipette. The contents were incubated at 94^0C for 3 min and snap chilled on ice for 5 min. Further, 5 µl of reaction buffer was added along with 0.2 µl of φ29 DNA polymerase and the contents were mixed gently. Then the tubes were incubated at 30^0C for 4-16 hours, before heat inactivating the enzyme by incubating the tubes at 65^0C for 10 min.

3.4 Preparation of *E. coli* competent cells

E. coli competent cell preparation was carried out by $CaCl_2$ method following Mendel and Higa (1970). A single *E. coli* colony was grown overnight into 3 ml of LB medium at 37˚C (primary culture). The following day, 1 ml of the primary culture was added to 50 ml of LB medium and incubated at 37˚C till OD at 600nm reaches 0.5. Culture was then kept on ice for 15 min and the cells were spun at 6000 rpm for 5 min at 4˚C. The supernatant was discarded and the pellet was re-suspended in 30 ml of ice-cold sterile solution consisting of 0.1M $CaCl_2$ and 24ml of 0.1M $MgCl_2$. This solution was incubated on ice for 30 min and then the cells were pelletized at 6000 rpm for 5 min at 4˚C. Supernatant was discarded and the remaining cells were gently re-suspended in the solution consisting of 225 µl sterile 100% glycerol and 1275 µl of 0.1M $CaCl_2$. The cells was gently mixed for transferring into pre-chilled micro-centrifuge tubes and immediately stored -80˚C.

3.5 Transformation of *E.coli* cells

Transformation of ligated DNAs into *E. coli* competent cells was performed according to Hanahan *et al*. (1991). Prior to the addition of ligation mixture, stored competent cells were thawed completely by keeping on ice for 10-15 min. Then 10 µl of the ligation mixture was added to component cells, mixed by gentle tapping and was incubated on ice for 1 h. Heat shock was given to the cells in the microcentrifuge tubes at 42˚C for 2 min

and the cells were immediately transferred back to ice for 5 min. One ml of LB medium was added into the microcentrifuge tube and incubated at 37˚C for 1 h in an incubator shaker. The cell suspension was centrifuged at 6000 rpm for 5 min at room temperature. Excess of LB was then discarded by simply inverting the microcentrifuge tube in a discard box inside the laminar air flow apparatus. The cells were gently re-suspended in the remaining LB media and the re-suspended cells were then plated onto LA plate supplemented with appropriate antibiotic selection. The plates were then kept at 37˚C for 16 h.

3.6 Isolation of plasmid DNA

Plasmid DNA isolation was carried out by alkaline lysis method as described by Bimboin & Dolly (1979). From overnight grown *E. coli* culture, 3 ml of the culture was pelletized at 13000 rpm for 5 min and re-suspended in 200 µL of alkaline lysis solution I (resuspension buffer: 50mM glucose, 25mM Tris-Cl pH 8, 10mM EDTA pH 8) and kept at room temperature for 5 min. Then, 200 µL of freshly prepared alkaline lysis solution II (lysis buffer: 0.2N NaOH, 1% SDS) was added and mixed gently by inverting the tubes. After incubating at room temperature for 5 min, 200 µL of alkaline lysis solution III (neutralization buffer: 5M Potassium acetate, glacial acetic acid, water) was added and mixed gently which was followed by the incubation on ice for 15 min. Then the tubes were centrifuged at 13000 rpm, 4°C for 15 min. Supernatant was taken in a fresh tube and equal volume of phenol: chloroform: isoamyl alcohol (24:24:1) was added to it, vortexed followed by 5 min incubation at room temperature. This was centrifuged at 13000 rpm, room temperature for 10 min. The upper layer was collected in a fresh tube and 0.8 volume of isopropanol was added. The mixture was kept at -20^{0}C for 20 min and centrifuged at 13000 rpm, 4^{0}C for 25 min. Supernatant was discarded and the pellet was washed with 0.5 ml of 70% ethanol at 13000 rpm, 4°C for 10 min. The pellet dried completely and resuspended with 30 µL of sterile double distilled water. The quality of the plasmid DNA isolated was checked by agarose gel electrophoresis and quantified by spectrophotometer.

3.7 Restriction digestion of plasmid DNA

The reaction mixture for restriction digestion for a particular enzyme(s) was determined as per the manufacturer's instructions (Thermo scientific, Waltham, USA). DNA was

used for restriction digestion either with one enzyme (single digestion) or two enzymes (double digestion). Double digestion was performed in two different ways depending on the enzyme and the buffer compatibility. If the enzymes and buffers were incompatible for the double digestion, the DNA was then first cleaved with one endonuclease, extracted with phenol:chloroform, and precipitated using isopropanol. It was then subjected to restriction digestion with second restriction endonuclease. Finally, the microcentrifuge tubes were incubated at suitable temperature according to manufacturer's instructions for 3-4 hours.

Components of reaction mixture	**Concentrations used**
Template DNA	100ng
10X reaction buffer	2.0 µL
BSA (if required)	0.5 µL (0.1mg/ml)
Restriction enzyme	0.5-1.0 µL (1U/µg)
Sterile distilled water	Final volume adjusted to 20 µL

3.8 Dephosphorylation of linearized vectors

The linearized vector after restriction digestion with an enzyme was heated at recommended temperature based on the enzyme used for 20 min to inactivate the restriction enzyme. Further the linearized vector was extracted by phenol:chloroform and precipitated with isopropanol. This was subsequently used as template for 5' dephophorylation using Antartica phosphatase (New England Biolabs, Ipswich, USA). The tubes were incubated at 37°C for 10-15 min and the phosphatases were heat inactivated at 80° C for 2 min. This purified dephosphorylated plasmid was then eluted using gel elution kit as per the manufacturer's instructions (Thermo scientific, Waltham, USA). The purified dephosphorylated plasmid was treated as vector in ligation mixture set up.

Components of reaction mixture	**Concentrations used**
Template	0.5-1 µg
Reaction buffer	1X
Antartica phsphatase enzyme	0.5-1.0 µL (1U/µg)
Sterile distilled water	Final volume adjusted to 20 µL

3.9 Ligation for cloning of DNA

Prior to ligate the eluted vector and insert DNA, 2.0 µl of both vector and insert DNA were analyzed on 1% agarose gel to estimate their concentration on comparing with 1 kb DNA ladder (Thermo scientific, Waltham, USA). The volume of vector and insert was calculated according to their band intensity. The vector and insert DNA ratio was taken 1:3 and the reaction mixture was made as described below. Ligation mixture was incubated at 16^0C for 16-18 hours. After completion of ligation the ligation mixture was stored at -20° to proceed further for transforming into competent cells.

Components of reaction mixture	Concentrations used
Vector DNA	1.0 µL
Insert DNA	3.0 µL
10X reaction buffer	2.0 µL
ATP (25mM)	1.0 µL
T4 DNA ligase	1.0 µL (10U/µL)
Sterile distilled water	Final volume adjusted to 15 µL

3.10 Polymerase chain reaction (PCR)

From plasmid or genomic DNA, specific DNA fragments were amplified by PCR involving template denaturation, annealing of primers and primer extension using the thermostable *Taq* or *Phusion* DNA polymerases. The components and the concentrations used in the reaction mixture for PCR were described below. The PCR amplified products were electrophoresed on agarose gel electrophoresis and then eluted using gel extraction kit (Thermo scientific, Waltham, USA). PCR condition was carried out standardized based on the specific DNA sequence to be amplified and the primers used.

Components of reaction mixture	Concentrations used
Template DNA (50ng)	1.0 µL
DNA polymerase (3U/µL)	0.5 µL
10X reaction buffer	2.0 µL
25mM $MgCl_2$	1.0 µL
2.5mM dNTP	1.0 µL
Primers (40 nmol/µL)	0.5 µL each
Sterile distilled water	Final volume adjusted to 20 µL

3.11 Gel elution

The desired digested or PCR amplified DNA bands were excised from the agarose gel and the DNAs were extracted using a gel extraction kit as per the manufacturer's instructions (Thermo scientific, Waltham, USA). After gel elution, 2 µl of the DNA products was loading on agarose gel for checking its concentration. The eluted product containing the required DNA fragment of interest was further used for setting up ligation.

3.12 Cloning of amplified PCR product in the cloning vectors

The PCR amplified products were electrophoresed on agarose gel and the PCR amplified products were gently excised using a sterile razor blade and the DNAs from the gel was extracted as described above. A ligation mixture was setup using pTZ57R/T vector supplied within the T/A cloning kit or pJET1.2 vector, if the DNA polymerase used in PCR was *Taq* or *Phusion*, respectively (Thermo scientific, Waltham, USA). The ligation sample was incubated at 16°C for 16 hours before transforming into competent *E. coli* cells. T/A vector allows the blue-white selection on X-gal, IPTG and ampicillin supplemented LB agar plate for easy identification of colonies transformed with the desired clone whereas pJET1.2 vector does not support blue-white selection. The white colonies from the agar plates were inoculated into LB medium supplemented with ampicillin (100mg/L) and cultured in an incubator shaker overnight at 37°C. Plasmid DNA was isolated from these colonies and the clone confirmation was subsequently performed by digestion and PCR following the protocol as mentioned in the earlier sections.

3.12.1 Cloning of viral DNA using RCA product

Concentration of viral DNA isolated by the above mentioned protocol was determined by spectrophotometer. One µg of viral DNA was digested separately in the eppendorf tube with *Hind* III, *Bam* HI, *Pst* I, *Kpn* I and *Xba* I restriction enzymes. Vector such as either *p*UC18 or *p*Bluescript (250ng) was also digested with the same enzymes and was dehosphorylated to avoid re-ligation of plasmid. Enzymes were inactivated by heat from both the vector and digested viral DNA. Ligation was set up for each digested viral DNA with its respective linearized vector using a ratio of 750ng viral DNA and 250ng of vector. Ligation mixture was prepared by above described protocol. This was then incubated at 16°C for overnight and used to transform *E. coli* cells. White blue selection

was applied as discussed in T/A cloning procedure and recombinant plasmids were screened on agarose gel. Uncut recombinant plasmids showing mobility like ~5.2 kb (determined by loading known plasmid construct; other monomeric clones as positive controls in gel) were selected and digested with respective restriction enzymes used for cloning.

3.12.2 Cloning of viral DNA through PCR

In addition to the RCA based method, we also employed PCR based cloning strategy to isolate the complete sequence of begomoviruses associated with ChiLCD in India. PCR was performed with various sets of primers to clone ChiLCD-associated begomoviruses (DNA-A). In addition, a degenerate primer pair was utilized to detect DNA-B component in ChiLCD infected samples (Table 3.2).

Table 3.2. List of primers used to clone the viral genomic components.

Virus component	**Primer sequence (5'-3')**	**Expected Size**	**References**
DNA-A			
ChiLCV	FP: TCAGGATCCATTAGTAAACGAGTTTCC RP: TGCGGATCCCACATTGGAAAGAGTG	2.7 kb	George *et al*., 2014
ToLCNDV	FP: GGATCCATTATTGCACGAATTTCC RP: GGATCCCACATGTTTGTGGATCCAA	2.7 kb	In this study
ToLCGV	FP: GGATCCTTTGTTAAACGAGTTTCC RP: GGATCCCACATAGTACGAACTAC	2.7 kb	In this study
DNA-B			
ToLCNDV	FP: TCTAGAACTSATCKGGTGTCTCTC RP: TCTAGAGAGAGAAACTSCWVCTTC	2.7 kb	George *et al*., 2014
DNA-α			
DNA101/102	FP: CTGCAGATAATGTAGCTTACCAG RP: CTGCAGATCCTCCACGTGTATAG	1.4 kb	Bull *et al*., 2003
UN101/102	FP: AAGCTTGCGACTATTGTATGAAAGAGG RP: AAGCTTCGTCTGTCTTACGAGCTCGCTG	1.4 kb	Bull *et al*., 2003
Alpha	FP: GAATTCCAGAAGACCGCTAGTGAG RP: GAATTCTTCCTCAGCAAGTTTGGCT	1.4 kb	In this study
DNA-β			
Beta	FP: GGTACCACTACGCTACGCAGCAGCC RP: GGTACCTACCCTCCCAGGGGTACAC	1.3 kb	Briddon *et al*., 2003

The universal primers were further employed to clone DNA-α and DNA-β molecules and also used to check presence of these satellites and satellite-like molecules in the infected samples (Briddon *et al*., 2002; Bull *et al*., 2003). The primer combinations used to clone ChiLCD-associated viral genomic components were listed in the Table 3.2. These PCR-amplified products were loaded and eluted from the agarose gel as mentioned in the section 3.11. The eluted DNAs were then used to ligate into either pTZ57R/T or pJET1.2 vector (Thermo scientific, Waltham, USA). All the clones were then subjected to restriction digestion using several restriction enzymes to identify the restriction pattern among the cloned viral components. Among them, one of the representative positive clones was commercially sequenced at the DNA sequencing facility (University of Delhi South Campus), using the primers specific to cloning vectors or the cloned virus molecules.

3.13 Per cent pairwise identity and phylogenetic analysis

Per cent pairwise identity of the cloned viral genomic components was calculated by aligning the full-length viral components along with the representative sequences downloaded from the Genbank as described by Brown *et al*. (2015). The nucleotide identity was estimated using the MUSCLE method available in SDT v1.0 (Muhire *et al*., 2014). The schematic representation of these per cent identity matrix was downloaded, in which the color codes represent the various range of identity scores. The individual per cent pairwise identity values also have been downloaded from SDT v1.0. Similarly, phylogenetic dendrograms were generated by the neighbor-joining method after aligning the viral sequences using MEGA program version 6.0 (Tamura *et al*., 2013). The phylogenetic trees were verified by the bootstrap analysis with 1000 pseudo-replicates as implemented in MEGA 6.0. The separate phylogenetic trees were generated for each of the viral genomic components cloned in this study. The bootstrap values were provided at the nodes of all the phylogenetic trees. All the major clades were color-coded, in which each color represents a particular begomovirus species/alphasatellite group/betasatellite group. The names of all these species/groups were provided in the right side of each of the phylogenetic dendrogram.

3.14 Recombination analysis

The potential recombinant sequences and breakpoint events were detected using recombination detection methods such as RDP, GENECONV, BOOTSCAN, MAXCHI, CHIMAERA, SISCAN and 3SEQ available in RDP3 (Martin *et al*., 2010). The separate RDP dataset for each one of the begomovirus or alphasatellite or betasatellite sequence was prepared including the particular sequence, the sequence with which it possess maximum nucleotide pairwise per cent identity and other sequences from the database to which it share per cent identities between 75% and 90%. The recombination breakpoints for every detected potential recombination event were verified and manually adjusted, if necessary. The identified putative parental sequences may not be the exact or actual ancestors. However, it can be mentioned here that they were the most similar ones to the actual parents among the sequences considered in the dataset. RDP analysis was performed using the Bonferroni corrected *p*-value cut-off of 0.01. The events detected by atleast three recombination detection methods available in RDP3 package were only considered. The name and the accession numbers of the begomoviral genomic components selected from the database to be used in the RDP analysis are provided in Table 3.3.

3.15 Selection pressure analysis

The type of natural selection acting on the ORFs encoded by the viral components were studied using neutrality tests such as Tajima's D, Fu and Li's D* and Fu and Li's F*. These tests were performed using the DnaSP v.5.10 software (Rozas *et al*., 2003). In addition, the detection of sites evolved under diversifying (positive) and purifying (negative) selection pressure was analyzed by various methods. It includes single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), random-effects likelihood (REL), mixed effects model of evolution (MEME) or fast unbiased bayesian approximation (FUBAR) available in the DataMonkey server (www.datamonkey.org) (Pond and Frost, 2005). Bayes factors >50 or posterior probability of $p \geq 0.9$ or *p*-value <0.1 was used as the threshold values for the analyses.

Table 3.3. Name and accession numbers of the begomoviral genomic components selected from the database to be used in the RDP analysis.

Vial components taken from the database for RDP analysis	Accession number
Begomoviruses	
Ageratum enation virus-[Nepal:2001]	AJ437618
Bhendi yellow vein mosaic virus-India [India:Madurai]	AF241479
Chilli leaf curl virus-Khanewal [Pakistan:Khanewal1:2004]	DQ116878
Chilli leaf curl virus-Pakistan [Pakistan:Multan:1998]	AF336806
Croton yellow vein mosaic virus-[India]	AJ507777
Euphorbia leaf curl virus-[China:Guangxi 35:2002]	AJ558121
Mungbean yellow mosaic India virus-[India:New Delhi:Blackgram3:1991]	AF126406
Papaya leaf curl virus-India [India:Lucknow]	Y15934
Papaya leaf curl virus-Pakistan [Pakistan:Cotton:2002]	AJ436992
Pepper leaf curl Bangladesh virus-Bangladesh [Bangladesh:Bogra:1999]	AF314531
Pepper leaf curl Bangladesh virus-Pakistan [Pakistan:Khanewal:2004]	DQ116881
Pepper leaf curl Lahore virus-[Pakistan:Lahore:2004]	AM404179
Pepper leaf curl virus-Malaysia [Malaysia:Klang:1997]	AF414287
Pepper leaf curl virus-Thailand [Thailand]	AF134484
Radish leaf curl virus-[India:Varanasi:2005]	EF175733
Squash leaf curl China virus-India[India:Coimbatore:Pumpkin]	AY184487
Tomato leaf curl Bangalore virus-A [India:Bangalore 1]	Z48182
Tomato leaf curl Gujarat virus-[India:Varanasi:2001]	AY190290
Tomato leaf curl Joydebpur virus-India [India::05]	DQ673859
Tomato leaf curl Joydebpur virus-Bangladesh [Bangladesh]	AJ875159
Tomato leaf curl Karnataka virus-Bangalore [India:Bangalore:1993]	U38239
Tomato leaf curl Kerala virus-[India:Kerala II:2005]	DQ852623
Tomato leaf curl New Delhi virus-India [India:New Delhi:Severe:1992]	U15015
Tomato leaf curl New Delhi virus-India [India:New Delhi:Mild:1992]	U15016
Tomato leaf curl Sri Lanka virus-[Sri Lanka:Bandarawela:1997]	AF274349
Betasatellites	
Ageratum yellow leaf curl betasatellite-[Pakistan:Faisalabad 3:1994]	AJ316027
Bhendi yellow vein betasatellite-[Pakistan:Bahawalpur:1997]	AJ316030
Chilli leaf curl betasatellite-[Pakistan:Gujranwala23:2004]	AM279664
Chilli leaf curl betasatellite-[Pakistan:Sialkot21:2004]	AM279663
Croton yellow vein mosaic betasatellite-[Pakistan:Punjab:2006]	AM410551
Vial components taken from the database for RDP analysis	**Accession number**
Papaya leaf curl betasatellite-[India:Jabalpur:2003]	AY230138
Radish leaf curl betasatellite-[India:Varanasi:2005]	EF175734
Tobacco leaf curl betasatellite-[Pakistan:Rahim Yar Khan:1998]	AJ316033

Tomato leaf curl Ranchi betasatellite-[India:Ranchi:2007]	GQ994096
Tomato leaf curl betasatellite-[IN:Var:06]	AY438559
Tomato leaf curl Bangladesh betasatellite-[Bangladesh:Gazipur:2001]	AJ542489
Tomato leaf curl Bangladesh betasatellite-[India:Lucknow:05]	DQ343289
Tomato leaf curl Bangladesh betasatellite-[India:Rajasthan:2003]	AY438558
Tomato leaf curl Joydebpur betasatellite-[IN:Kenaf:07]	EU43115
Alphasatellites	
Okra leaf curl alphasatellite [PK:NOB3:2002]	AJ512954
Tobacco leaf curl alphasatellite [PK:NOB2:2002]	AJ512955
Bhendi yellow vein mosaic alphasatellite [IN:G2-8:2002]	AJ512957
Tomato yellow leaf curl China alphasatellite [CN:Y89:2003]	AJ579357
Malvastrum yellow mosaic alphasatellite [CN:Hn44:2006]	AM236766
Sida yellow vein Vietnam alphasatellite [VN:Han:2006]	DQ641718
Gossypium darwinii symptomless alphasatellite [PK:Mul:2008]	EU384653
Gossypium mustilinum symptomless alphasatellite [Pk:Mul:2008]	EU384662
Ageratum yellow vein Singapore alphasatellite [OM:Alb43:2005]	FJ956707
Chilli leaf curl Multan alphasatellite [PK:Fai:2009]	FN432360
Ageratum enation alphasatellite [IN:Kan:2009]	FN543100
Bhendi yellow vein mosaic alphasatellite [IN:Har:2007]	FN658716
Tobacco curly shoot alphasatellite [CN:Y216:2010]	FN678901
Ageratum conyzoides associated symptomless alphasatellite [PK:Lah:2006]	FR772085
Cotton leaf curl Burewala alphasatellite [IN:Sri:2010]	GU992937
Cassava mosaic Madagascar alphasatellite [MG:Diana:635A6:2011]	HE984148

3.16 Estimation of nucleotide substitution rates

The NEXUS files of ORF AV1 (for begomoviruses), ORF βC1 (for betasatellites) and ORF α-Rep (for alphasatellites) nucleotide sequences were generated using MEGA 6.0. Only the sequences belonging to single begomovirus species (*Chilli leaf curl virus*), betasatellite group (*Chilli leaf curl betasatellite* and *Tomato leaf curl Bangladesh betasatellite*) and alphasatellite group (*Chilli leaf curl alphasatellite*) were included separately in this analysis. All the inter-species/group recombinant sequences detected by RDP analysis are omitted from the present analysis. The nucleotide substitution rate was calculated using GTR nucleotide substitution model, uncorrelated lognormal relaxed clock model and the BSP coalescence model available in BEAST v.1.7 package (Drummond *et al*., 2012). Bayesian phylogeny using the MCMC method for 40 million steps was used for this analysis. These Bayesian analyses were used as the input file in

the Tracer in BEAST v.1.7 package to estimate the rate of nucleotide substitution and mutation.

3.17 Preparation of *Agrobacterium* competent cells

An *Agrobacterium tumefaciens* strain (either EHA105 or GV3101) colony was used as inoculum for 2 ml LB media and incubated at 28°C at 220 rpm for 16 hours. This was used to inoculate 50 ml of LB media and allowed to shake at 200 rpm, 28°C until OD at 600 nm reaches 0.6-1.0. The culture was kept on ice for 5 min and 30 ml from the culture was taken in a fresh sterile centrifuge tube. The tube was spun at 6500 rpm and the pellet was mixed with 10 ml of ice cold sterile 0.15 M NaCl. The mixture was incubated on ice for 15 min and was centrifuged at 7000 rpm, 4°C for 5 min. This time, the pellet was re-suspended in 1 ml of ice-cold sterile 20 mM $CaCl_2$. These cells were aliquoted into sterile microcentrifuge tubes each containing 100 µL of cells for agro-transformation.

3.18 Agro-transformation

The tube containing 100 µL of competent cells was thawed and 1 µg of plasmid DNA was added by pipette tip and placed on ice for 30 min. Then the tube was frozen in liquid nitrogen for a minute. The tube was thawed in 37°C water bath until the frozen solution becomes liquid. The LB medium of 900 µL was added to it and kept in the shaker at 28°C for 4-5 hours. The cells were centrifuged at 7000 rpm, 4°C for 5 min and the pellet was mixed in 100 µL of LB medium. Then the cells were plated onto the appropriate LA plates and incubated for 48 hours at 28°C.

3.19 Generation of partial tandem repeat constructs (PTR)

The PTR constructs of the cloned begomoviral genomic components were generated in a plant expression vector (pCAMBIA1300/2300) using the viral DNA monomeric clones (DNA-A, DNA-β, DNA-1). In all these cases, two restriction sites were selected around the common region (CR) i.e., one from the left side and the other from the right hand of the CR. Upon digestion with these restriction enzymes, fragment encompassing CR was cloned into the multiple cloning sites (MCS) of pCAMBIA1300/2300 vector. Further, the full-length viral genome component was mobilized into it. The orientation of the tandem repeats were then assessed by restriction digestion using appropriate enzymes used for the cloning.

3.20 Infectivity analysis

3.20.1 Agro-inoculation

The 5-leaf stage *N. benthamiana* plants were used for agro-inoculation of viral PTR constructs. Agro-inoculation was performed on these plants as described by Chattopadhyay *et al*. (2008). *A. tumefaciens* strain EHA105 harboring each PTR construct was inoculated into 2ml LB medium as primary culture. Secondary culture was initiated into 100ml LB media and allowed to grow at 28°C for 36 hours at 220 rpm. The culture was pelletized at 6000 rpm for 15min and then re-suspended in sterile double distilled water. For each inoculum, the optical density (OD 600) was maintained as 0.5 and the equimolar concentration of each construct to be used for agro-inoculation was mixed. The test plants were pricked at the petiole and stem with a fine needle followed by the application of 50µL of the inoculum at the injured site. The symptoms were measured using the scale given in the Table 3.4.

3.20.2 Particle bombardment

Virus inoculation in *C. annuum* var. Sulekha plants was carried out by biolistic delivery as per manufacturer's instructions (Bio Rad, Catalog numbers: 165-2431 and 165-2432). Gold microcarriers of 0.6µm diameter covered with the various viral PTR constructs were used for the particle bombardment. For the coating of viral PTR constructs on gold microcarriers, 25mg of gold microcarriers were mixed with 100µl of 0.05M spermidine and 100µl of viral PTR constructs (3µg/µl, containing equal amount of DNA-A and DNA-β) followed by gentle vortexing and addition of 100µl of 1M $CaCl_2$ during vortexing.

In the next step, viral PTR constructs coated gold microcarriers were loaded into Gold-Coat tubing with the help of 0.1 mg/ml of PVP and gold coated Gold-Coat Tubing was cut into the pieces (cartridges) of 0.5 inches. Finally, the gold microcarriers were bombarded on two upper leaves of *C. annuum* var. *Sulekha* plants from cartridges by Helium gas at 300 psi with the help of Helios Gene Gun (Bio-Rad). The infected plants were maintained in the growth chamber and the appearance of symptoms was monitored using the scale given in the Table 3.4.

Table 3.4. Details of the symptom measurement using 0-5 scale.

Symptom score	Description of symptoms
0	No symptom
1	Mild chlorosis and little deformation in new leaves
2	Leaf curling, chlorosis and/or mosaic
3	Leaf curling, chlorosis, vein clearing and/or yellowing, reduced plant growth
4	Severe leaf curling, shoot apex growth affected, rolling of newly emerging leaves
5	Severe leaf curling, leaf crumpling of whole of the shoot apex, mosaic, stunting

3.20.3 Detection of viral DNAs

From the infected plant samples, the accumulation of viral DNAs was analyzed by Southern blotting. The hybridization and the detection of viral bands were performed as described below. For detecting DNA accumulation of various begomoviruses, the probes were prepared using the complete coat protein gene of the respective begomovirus species. Similarly, βC1 region and α-rep region was used for preparing probe to detect betasatellite molecules and alphasatellite molecules, respectively. The viral DNAs were finally detected using phosphor image analyser (Typhoon, GE, Chicago, USA).

3.20.4 Southern hybridization for analyzing viral DNA accumulation

Southern hybridization was performed for the detecting viral DNA in virus inoculated plants (Southern 1975). Southern hybridization was done according to the procedure as described in Sambrook and Russell (2001). Ten micrograms of genomic DNA was used for analyzing the accumulation of viral DNA in the sample. The genomic DNA was electrophoresed for 4-5 hours in the agarose gel (0.8%) by applying voltage between 50V and 60V for smooth and clear separation of the viral DNA bands. The capillary method was employed for DNA transfer to nylon membrane (MDI, India) as described in Sambrook and Russell, (2001). The agarose gel was then soaked in 0.25M HCl for depurination for 15 min with gentle agitation. The acid depurination enables the partial hydrolysis of the high molecular weight DNA which helps in the proper transfer of DNA

to the nylon membrane. After this depurination, the gel was washed briefly with deionized water followed by the addition of denaturation solution (0.5M NaOH and 1.5M NaCl) for 45 min with gentle rocking. The denatured gel was washed with deionized water and treated with neutrilization solution (0.5M Tris-HCl and 1.5M NaCl pH7.2) twice for 30 min and 15 min.

Using the capillary transfer method, the DNA was then transferred onto the nylon membrane using 10X SSC (3.0 M NaCl and 0.3M sodium citrate). After overnight transfer, the transferred DNAs were cross linked onto the nylon membrane by exposing to the ultraviolet rays. The random probe labelling method was used to prepare radiolabelled probes to detect viral DNA accumulation (Feinberg & Vogelstein, 1983). The hybridization was performed using the radio-labelled (α^{32}P-dCTP) specific probes at 60°C. Then the blots were washed with washing solution (2.0X SSC and 0.1% SDS) for 15 min at room temperature. To remove excess of probes and to rule out nonspecific signals if any, the blots were again washed at 65°C for 5 min. The blots were air dried and kept on the imaging plates for overnight. The viral DNAs were further detected using phosphor image analyser (Typhoon).

3.21 Transcript analysis

3.21.1 Isolation of total RNAs

Hundred microgram of plant tissues were ground using liquid nitrogen and 500 µl of Trizol reagent (Sigma) was added. Subsequently, this was homogenized by adding 500 µl of Trizol reagent again and mixed the content before incubating at room temperature for 5 min. Then 200 µl of chloroform was then mixed. The tubes were kept at room temperature for 15 min. The tubes were then centrifuged for 10 min at 12000 rpm. The supernatant was collected in the sterile tube and to precipitate the total RNAs, 0.8 volume of isopropanol was added and mixed before keeping on ice for 20 min. The tubes were then centrifuged at 12000 rpm for 10 min at 4^0 C. Five hundred microlitre of 75% ethanol was added to the pellet and then centrifuged at 12000 rpm at 4^0 C for 10 min. This step was repeated again and the residual ethanol was removed by pipette tips. The pellets were dried at 37^0 C and sterile DEPC-treated water was added based on the size of the pellet. The extracted total RNAs were dissolved at 65^0 C for 20 min and the tubes were snap

chilled on ice. The dissolved RNAs were store at -80^0 C for further used or the OD was calculated using spectrophotometer.

3.21.2 cDNA synthesis

Five microgram of the extracted total RNAs from leaf tissues were treated with DNase I enzyme. The reaction mixture was incubated at 37^0 C for 30 min. The reaction was stopped by the addition of 5 µl of 25 mM EDTA and by heating at 72^0 C for 10 min before snap chilling the tubes on ice. One microgram of DNase-digested RNA and oligo(dT) primer was used to carry out the first-strand DNA synthesis and using MuMLV reverse transcriptase enzyme (Fermentas). The contents were incubated at 42^0 C for 70 min, followed by heat treatment at 72^0 C for 10 min.

3.21.3 Quantitative real time PCR analysis

The primers used for q-RT-PCR analysis were designed using Primer express v3.0 (Applied Biosystems, California, USA). The cDNAs synthesized from either the mock- or virus-inoculated plants were used for this analysis. SYBR green master mix (Applied Biosystem, California, USA) was used for PCR amplification using Illumina real time PCR machine (Illumina, San Diego, USA). PCR program used for amplification includes initial denaturation of 5 min at 94^{0}C, 40 cycles of denaturation at 94^{0}C for 15 sec, annealing at 55^0 C for 35 sec and extension for 35 sec at 72^0 C. For each sample, three biological and three technical replicates were included. The corresponding actin mRNA amplified from the same RNA samples were used to normalize the target mRNA levels (ΔCt values). Hence, actin mRNAs were served as an internal control and ΔΔCt values were used to plot the graphs.

3.22. Protein-protein interaction studies

3.22.1 Yeast-two-hybrid assay

The DNA encoding full-length and deletion constructs of NbVIM were inserted in frame into the GAL4 DNA binding domain vector pGBKT7. The viral ORFs were cloned into GAL4 activation domain vector pGADT7. The protein-protein interaction studies were performed using yeast strain AH109 as per manufacturer's instructions (Clontech Laboratories, California, USA).

3.22.2 Bimolecular fluorescence complementation (BiFC) assay

The BiFC assay was carried out using epidermal cells of *N. benthamiana* plants. The desired constructs for BiFC cloned in the plant transformation vectors were mobilized into *Agrobacterium* cells. These cells carrying the desired construct were grown in 50 ml LB medium containing appropriated antibiotics. The OD of the culture was adjusted to 0.5 by re-suspending the pellet with infiltration buffer. This mixture was kept in dark for 3 h before infiltrating into the leaves of 3-weeks old *N. benthamiana* plants using the needles syringe at the adaxial surface. Then the epidermal layer from the infiltrated regions was removed 4 days post infiltration and mounted on the slides with 10% glycerol. The nuclei of the epidermal layer were stained with 4, 6-diamidino-2-phenylindole (DAPI) and the epidermal cells were visualized under Andor spinning disk confocal microscope (Andor, Belfast, Ireland). A 340-380 nm excitation filter and a 435-485 nm emission filter were used for the detection of DAPI fluorescence. A 465-495 nm excitation filter and a 515–555 nm emission filter were used for visualizing the GFP fluorescence. The analysis of the captured images was performed using Andor iQ2.7 software.

3.23. Virus induced gene silencing

Virus induced gene silencing was performed using TRV based vectors that includes TRV1 and TRV2. These vectors were mobilized separately into *A. tumefaciens* strain EHA105. *Agrobacterium* cells containing pTRV1 and pTRV2 or pTRV derivatives were cultured overnight at 28°C in LB medium containing with kanamycin (50 μg/ml) and rifampicin (50 μg/ml). The overnight cultures were centrifuged at 6000 rpm for 5 min and re-suspended in the same volume of infiltration buffer containing 10 mM $MgCl_2$, 100 μM acetosyringone and 1 mM MES buffer pH 5.6. Then OD was adjusted to 0.5 and re-suspended cells were kept at room temperature for 3 h. After incubation, the cultures were mixed according to the experimental combinations in 1:1 ratio and used for infiltration at four leaf stage *N. benthaniana* plants using 1 ml needless syringe. After 2 weeks of TRV infiltration, the same plants were either mock-inoculated or with PTR construct of ChiLCV-NgA. Effect of target gene silencing on transcript accumulation was assessed at 20 dpi following virus challenge, by northern blot analysis using radiolabelled probes specific to *NbVIM* transcript. The viral DNA accumulation was analyzed by southern hybridization at 20 dpi using DNA-A and betasatellite specific probes.

Genetic diversity of begomoviruses associated with chilli leaf curl disease in India

4.1 Introduction

Chilli leaf curl disease (ChiLCD) results in severe crop losses especially in the Indian sub-continent and it have been as a major limiting factor for chilli cultivation (Mishra *et al*., 1963; Varma and Malathi, 2003; Senanayake *et al*., 2007; Chattopadhyay *et al*., 2008; Senanayake *et al*., 2013). ChiLCD infected chilli plants exhibit symptoms such as leaf rolling, leaf curling, puckering and reduced size of leaves; in severe cases, the infected plants remain stunted and fail to bear any fruit resulting in complete crop loss. Although the first report of ChiLCD was described almost 5 decades ago (Mishra *et al*., 1963; Dhanraj and Seth, 1968), the association of begomovirus with ChiLCD was confirmed recently (Khan *et al*., 2006; Senanayake *et al*., 2007; Chattopadhyay *et al*., 2008). In the recent years, ChiLCD has spread to major chilli growing regions in India. Monopartite begomoviruses and their betasatellites were found to be associated with ChiLCD (Chattopadhyay *et al*., 2008; Kumar *et al*., 2011). In India till date only two chilli-infecting begomovirus species viz., *Chilli leaf curl virus* (ChiLCV) and *Chilli leaf curl India virus* (ChiLCINV) have been reported but begomoviruses infecting other solanaceous plants such as *Tomato leaf curl New Delhi Virus* (ToLCNDV) and *Tomato leaf curl Joydebpur virus* (ToLCJoV) were found to be associated with this disease (Khan *et al*., 2006; Senanayake *et al*., 2007; Kumar *et al*., 2011; Senanayake *et al*., 2013). Despite the severe losses that are caused by ChiLCD, very little effort was made to understand the diversity of chilli–infecting begomoviruses in India.

4.2 Results

4.2.1 Survey and collection of samples

An extensive survey was carried out in all major chilli-growing regions across almost all the states in India. For each region from where the samples were collected, minimum of nine samples from at least three individual chilli fields were collected. These infected plant samples were taken from the chilli plants bearing typical begomovirus-like symptoms such as vein thickening, leaf curling, crinkling of leaves, vein banding, leaf curling and stunting (Figure 4.1). The details of the places and the GPS co-ordinates

from where the plant tissues collected alongwith the types of symptoms noticed on the infected plants were provided in the Table 4.1.

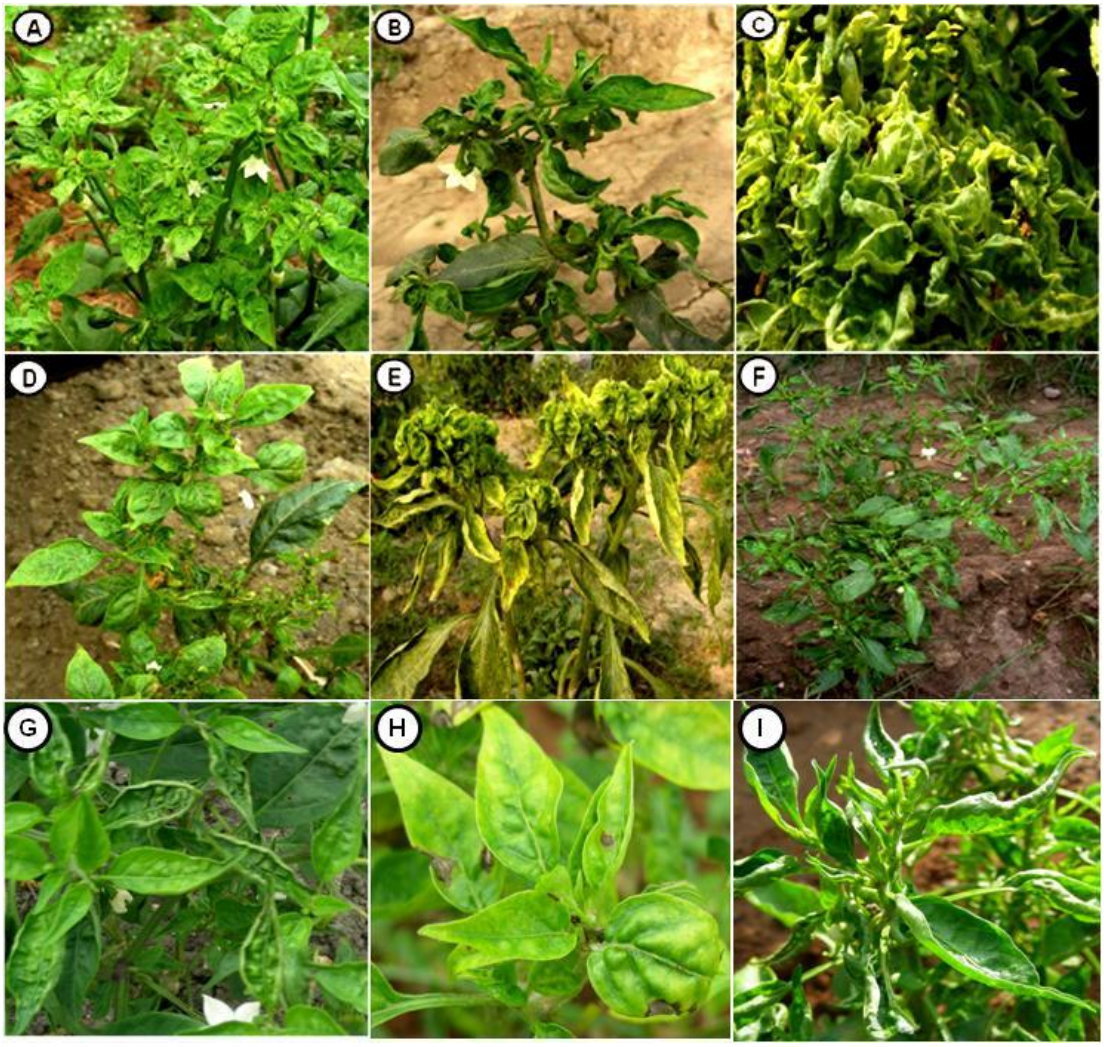
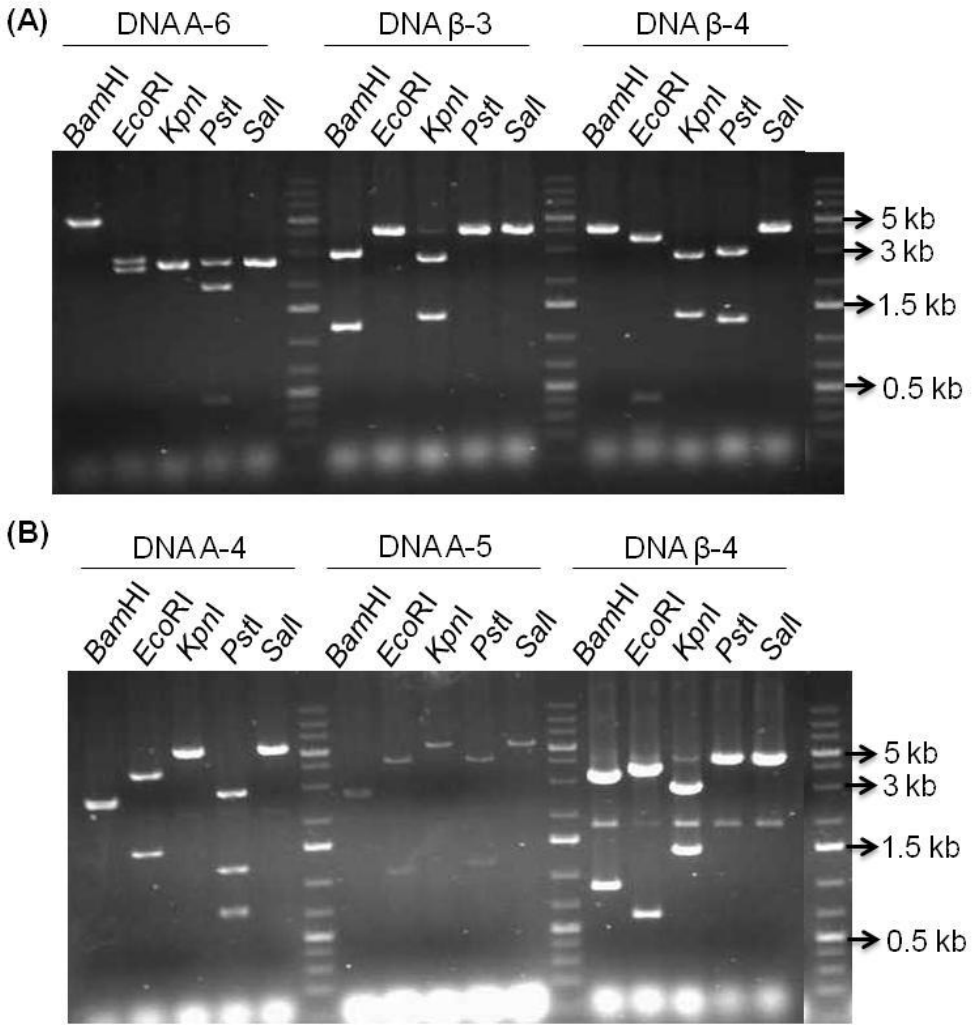


Figure 4.1. The begomovirus-like symptoms noticed on the leaf curl disease affected chilli plants from major chilli growing regions in the country. (A) Coimbatore, (B) Jabalpur, (C) Jodhpur, (D) Salem, (E) Varanasi, (F) Vellanad, (G) Guntur, (H) Tumkur and (I) Bengaluru.

4.2.2 Cloning of viral DNAs associated with ChiLCD in India

Chilli plants exhibiting begomovirus-like symptoms were collected from the major chilli-growing regions across the country. The symptoms observed include leaf curling, leaf crinkling, vein thickening, leaf blistering, leaf rolling, yellow of leaves and stunting (Figure 4.1). The types and severity of the symptoms varied from samples to samples. From these samples, total DNA was isolated and then subjected to either PCR or RCA. The complete viral genomic components were PCR amplified using component specific

primers, and cloned (Table 4.2). Both PCR based amplification for full-length amplification of viral components and RCA using φ29 polymerase, yielded numerous clones of varying sizes (both ~2.8 kb and ~1.3 kb). The viral DNA fragments were then inserted into the cloning vectors. The clones from the respective samples were subjected to restriction digestion using various restriction enzymes to analyze the diversity of the genomic DNA based on restriction fragment length polymorphism (RFLP). Two population of clones (either ~2.8 kb or ~1.3 kb) indicating similar or different restriction patterns were selected for full-length sequencing commercially. The digestion confirmation of the cloned viral genomic DNAs from Nagpur, New Delhi, Salem and Vellanad samples are provided in the Figure 4.2. A total of 87 complete viral sequences (41 of ~2.8 kb and 46 of ~1.3 kb in size) was obtained representing each location of the major chilli-growing regions in India (Tables 4.1-4.3).



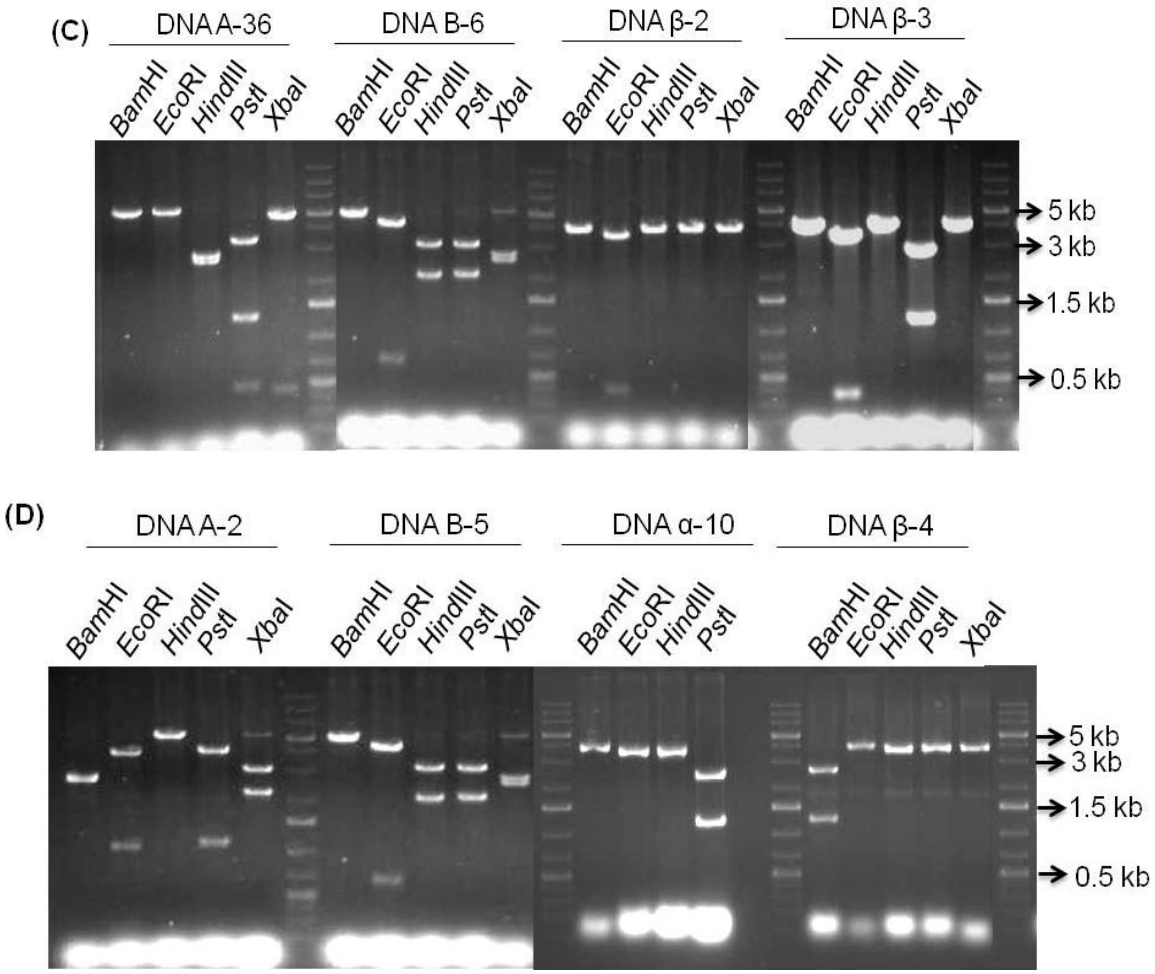


Figure 4.2. Agarose gel electrophoresis showing the digestion pattern of the viral genomic components isolated from Nagpur (A), New Delhi (B), Salem (C) and Vellanad (D) regions. The clone numbers of each of the viral components used for digestion are provided in each panel. Various restriction enzymes used for digesting the cloned genomes are mentioned above each of the gels. The selective bands of DNA ladder are labelled in each of the panels. DNA-A and DNA-B refer to begomovirus components and, DNA-α and DNA-β indicate satellite-like and satellite components.

4.2.2.1 Association of begomoviral genomic components with Indian ChiLCD

All the cloned components were analyzed for the presence of putative viral ORFs using the ORF finder tool. The analysis from the cloned viral components indicate that 35 out of 41 clones of ~2.8 kb size had a typical genome arrangement of DNA-A predominantly resembling with monopartite begomoviruses originating from the OW, whereas 6 clones resembled DNA-B like genomic organization (Table 4.1). The sequences to be half the size of DNA-A, termed as betasatellites were also be found to be associated with the samples irrespective of the nature of helper virus genomic components i.e., monopartite or bipartite (Tables 4.1, 4.2). However, the alphasatellites could not be identified from all

the samples infected with ChiLCD. The genome organization of these satellite molecules suggests the isolation of 33 betasatellite and 13 alphasatellite molecules from these samples (Tables 4.2, 4.3).

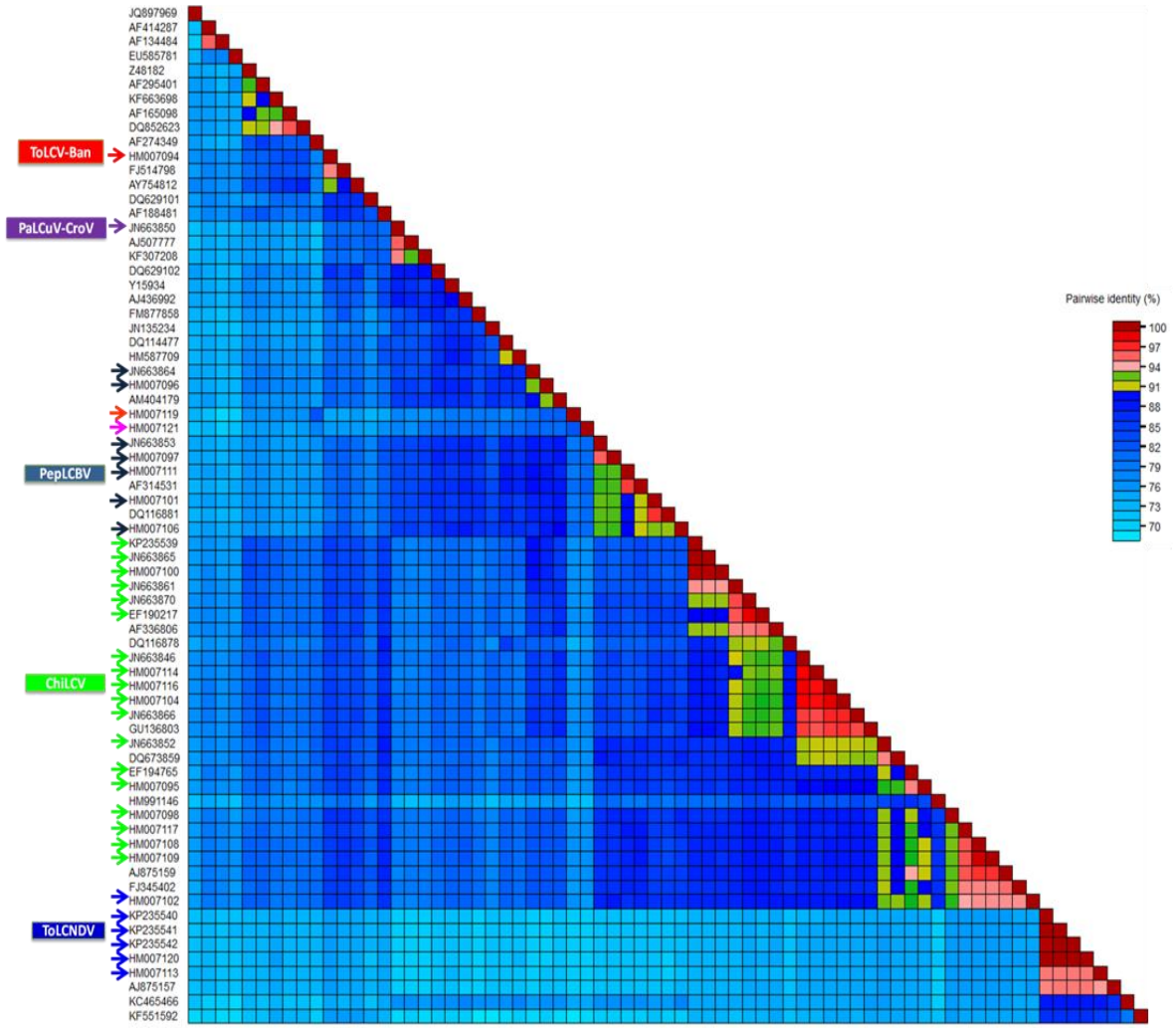


Figure 4.3. Graphical representation of per cent nucleotide identity plot of begomoviruses using SDTv1.0 program. The begomovirus isolates obtained in this study are marked with arrows and the representative begomoviral sequences taken from the database are unmarked.

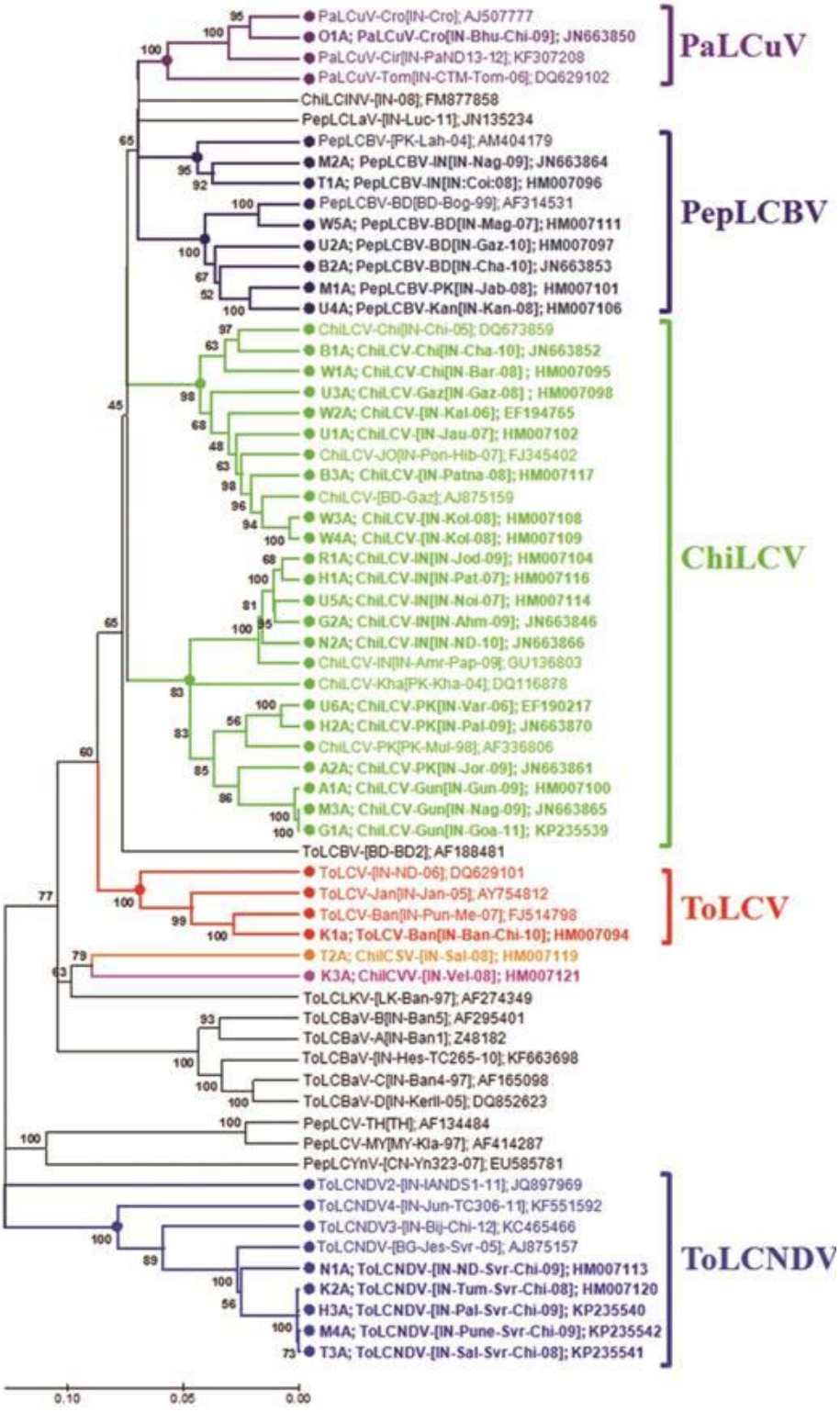


Figure 4.4. Neighbor-joining phylogenetic tree of ChiLCD-associated begomoviruses isolated here (indicated in bold type) along with the representative begomoviral genomes retrieved from the GenBank. CLUSTAL W method was used to align the sequence before the generation of the phylogenetic dendrogram. The bootstrap values of thousand replicates are provided at each node. Major clades in the same colour belong to a particular species which is given at the right side of the dendrogram. The scale bar refers to nucleotide substitution rate per site.

Table 4.1. Survey, collection and the list of helper begomoviruses identified from the ChiLCD-affected samples from major chilli growing regions in India.

Location		Year	DNA A/DNA-A-like[#]				DNA B				Nature of the genome
State	Area		Clone name	Accession number	Cloning method	Isolates*	Clone name	Accession number	Cloning method	Isolates*	
Andhra Pradesh	Guntur	2009	A1A	HM007100	RCA	ChiLCV-Gun	-	-	-	-	Monopartite
Assam	Jorehat	2009	A2A	JN663861	PCR	ChiLCV-PK	-	-	-	-	Monopartite
Bihar	Chhapra	2010	B1A	JN663852	RCA	ToLCJoV-IN[1]	-	-	-	-	Monopartite
			B2A	JN663853	RCA	PepLCBDV-BD[1]	-	-	-	-	
	Patna	2008	B3A	HM007117	RCA	ToLCJoV-BD	-	-	-	-	Monopartite
Goa	Keri	2011	G1A	KP235539	PCR	ChiLCV-Gun	-	-	-	-	Monopartite
Gujarat	Ahmedabad	2009	G2A	JN663846	RCA	ChiLCV-PK	-	-	-	-	Monopartite
Haryana	Pataudi	2007	H1A	HM007116	PCR	ChiLCV-PK	-	-	-	-	Monopartite
Himachal Pradesh	Palampur	2009	H2A	JN663870	PCR	ChiLCV-PK[2]	H2B	JN663871	PCR	ToLCNDV-IN[2]	Monopartite & bipartite
			H3A	KP235540	PCR	ToLCNDV-IN[2]					
Karnataka	Bengaluru	2010	K1A	HM007094	RCA	ToLCKV-Ban	K1B	JN663848	PCR	ToLCNDV-IN	Monopartite
	Tumkur	2008	K2A	HM007120	RCA	ToLCNDV-IN	-	-	-	-	
Kerala	Vellanad	2008	K3A	HM007121	RCA	ChiLCVeV[3]	-	-	-	-	Monopartite
Madhya Pradesh	Jabalpur	2008	M1A	HM007101	RCA	PepLCBDV-PK	-	-	-	-	Monopartite
Maharashtra	Nagpur	2009	M2A	JN663864	PCR	PepLCLV-Ng[4]	-	-	-	-	Monopartite
			M3A	JN663865	PCR	ChiLCV-Gun[4]	-	-	-	-	

	Pune	2009	M4A	KP235542	PCR	ToLCNDV-IN	M4B	KP235543	PCR	ToLCNDV-IN	Bipartite
New Delhi	New Delhi	2010	N1A	HM007113	RCA	ToLCNDV-IN[5]	N2B	JN663867	PCR	ToLCNDV-IN[5]	Monopartite & bipartite
			N2A	JN663866	RCA	ChiLCV-PK[6]	N3B	KP235538	PCR	ToLCGV[5]	
Orissa	Bhubaneswar	2009	O1A	JN663850	RCA	CYVMV	-	-	-	-	Monopartite
Rajasthan	Jodhpur	2009	R1A	HM007104	RCA	ChiLCV-PK	-	-	-	-	Monopartite
Tamil Nadu	Coimbatore	2008	T1A	HM007096	RCA	PepLCLV-Ng[7]	-	-	-	-	Monopartite
	Salem	2008	T2A	HM007119	RCA	ChiLCSmV[8]	T2B	KF471060	PCR	ToLCNDV-IN[8]	Monopartite & bipartite
			T3A	KP235541	PCR	ToLCNDV-IN[8]					
Uttar Pradesh	Jaunpur	2007	U1A	HM007102	PCR	ToLCJoV-BD	-	-	-	-	Monopartite
	Ghazipur	2010	U2A	HM007097	RCA	PepLCBDV-BD[9]	-	-	-	-	Monopartite
			U3A	HM007098	RCA	ToLCJoV-BD[9]	-	-	-	-	
	Kanpur	2008	U4A	HM007106	RCA	PepLCBDV-PK	-	-	-	-	Monopartite
	Noida	2007	U5A	HM007114	PCR	ChiLCV-PK	-	-	-	-	Monopartite
	Varanasi	2006	U6A	EF190217	PCR	ChiLCV-PK	-	-	-	-	Monopartite
West Bengal	Baruipur	2008	W1A	HM007095	PCR	ToLCJoV-BD	-	-	-	-	Monopartite
	Kalyani	2006	W2A	EF194765	PCR	ToLCJoV-BD	-	-	-	-	Monopartite
	Kolkata	2008	W3A	HM007108	RCA	ToLCJoV-BD[10]	-	-	-	-	Monopartite
			W4A	HM007109	RCA	ToLCJoV-BD[10]	-	-	-	-	
	Magrahat	2007	W5A	HM007111	RCA	PepLCBDV-BD	-	-	-	-	Monopartite

*clones followed by the same numbers (in superscript) indicate that they were obtained from the same plant in that particular location.

[#]DNA-A-like and DNA-A refers to monopartite and bipartite begomoviruses, respectively.

The O1A and K1A sequences showed maximum identities of 95.8% and 97.3% with a 'Croton' strain of *Papaya leaf curl virus* (PaLCuV-Cro; AJ507777) and a 'Bangalore' strain of *Tomato leaf curl virus* (ToLCV-Ban; FJ514798), respectively. Therefore, according to the ICTV guidelines, these isolates were thus considered to be the isolates of PaLCuV-Cro and ToLCV-Ban, respectively (Figure 4.3). The complete nucleotide sequences, B2A, M1A, U2A and W5A displayed maximum per cent pairwise identity of 93.1-96.8% with the isolates reported from either Bangladesh or Pakistan strains of *Pepper leaf curl Bangladesh virus* (PepLCBV) (Figure 4.3). These sequences further possessed 92.8–95.8% nucleotide identity among them and hence, positioned themselves in a separate clade along with PepLCBV sequences in the phylogenetic tree (Figure 3).

The sequence of H3A, K2A, M4A, N1A and T3A were found to have the highest sequence identity of 94.2–95.1% with ToLCNDV (AJ875157). Thus, they belong to ToLCNDV species and are grouped as a separate clade in the phylogenetic dendrogram (Figure 4.4). The sequence, U4A, had shown the highest per cent identity of 91.9% with a 'Pakistan' strain of PepLCBV (DQ116881). Similarly, M2A and T1A shared per cent sequence identity between 90.3% and 92.3% with PepLCBV (AM404179). Hence, based on the begomovirus species demarcation threshold value, these U4A and, M2A and T1A should be regarded as the new strains, for which, the name 'PepLCBV-IN' and 'PepLCBV-Kan', respectively were proposed (Figure 4.3).

The G2A, H1A, N2A, R1A and U5A sequences possessed highest sequence identity between 96.3% and 96.9% with an 'Indian' strain of ChiLCV (ChiLCV-IN; GU136803) (Figure 4.3). The sequences of five isolates (B3A, U1A, W2A, W3A and W4A) displayed 96.3–96.9% nucleotide identity with ChiLCV (AJ875159). The sequences of A2A, H2A and U6A were found to share the highest per cent identity of 94.8–95.5% with the 'Pakistan' strain of ChiLCV (ChiLCV-PK; AF336806) (Figure 4.3). Further, A2A, H2A and U6A sequences also possess 95.9-98.6% identity among themselves. However, U3A sequence showed maximum nucleotide identity of 92.8% with ChiLCV (AJ875159) (Figure 4.3). Therefore, on the basis of species demarcation limit for begomoviruses, U3A should be considered to be a new strain of ChiLCV (ChiLCV-Gaz). Similarly, A1A, G1A and M3A sequences had shown the maximum sequence identity of 91.9% with ChiLCV-PK (AF336806) and so they were regarded as the new strains of ChiLCV

(Figure 4.3). In support with their nucleotide per cent identities, all these ChiLCV strains and isolates were found to occupy a separate clade in the phylogenetic tree (Figure 4.4).

Additionally, two sequences shared less than 91% identity with any of the begomoviruses reported so far (Figure 4.3). Therefore, following the species demarcation threshold of begomoviruses indicated above, it is proposed that they belong to two novel begomovirus species. A sequence from the Salem region share maximum nucleotide identity of 82.2% with PepLCLV (AM404179) and so "Chilli leaf curl Salem virus" (ChiLCSmV) name is proposed (Figure 4.3). Similarly, the other sequence, isolated from the Vellanad region, showed maximum nucleotide identity of 81.1% with PepLCBDV (AF314531), and therefore, the name "Chilli leaf curl Vellanad virus" (ChiLCVV) was proposed (Figure 4.3). Hence, ChiLCSmV and ChiLCVV were found separated from other begomovirus species and were uniquely placed in the phylogenetic dendrogram (Figure 4.4).

4.2.2.2 Association of begomoviral DNA-B component with ChiLCD in India

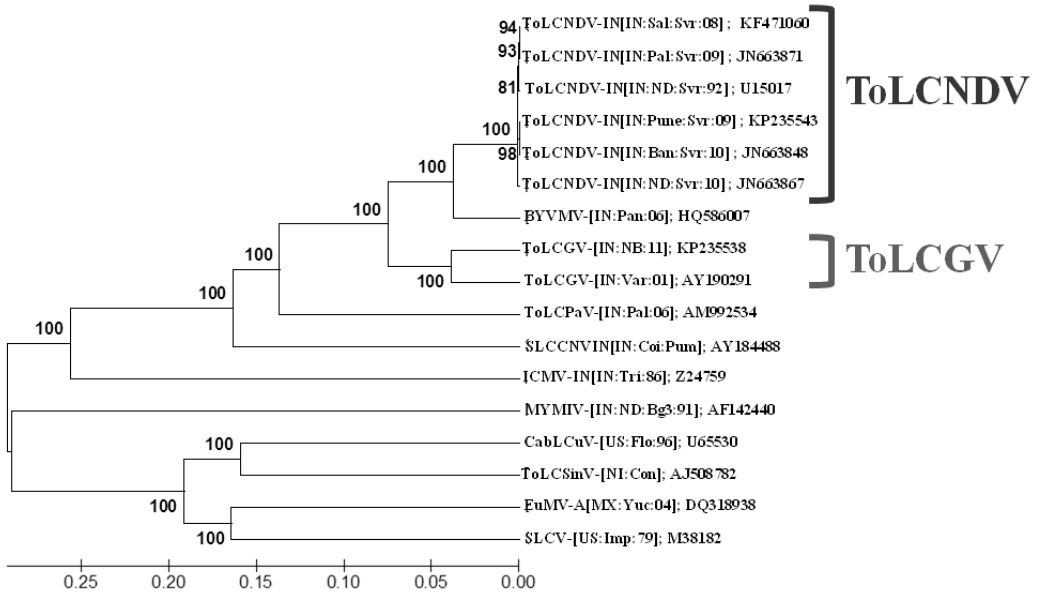


Figure 4.5. Neighbor-joining phylogenetic dendrogram of DNA-B components associated with ChiLCD. Per cent bootstrap values of 1000 replicates are placed at the nodes. The Clustal W algorithm was used to align all the sequences before generating the phylogenetic tree.

Six DNA-B clones (H2B, K1B, M3B, M4B, N2B and T2B) were identified with the ChiLCD-infected field samples from Bengaluru, Palampur, New Delhi, Pune and Salem (Table 5). The size of these clones was found to be of 2696 bp. The sequences, H2B,

K1B, M4B, N2B and T2B shared >99% pairwise sequence identity with the DNA-B component of ToLCNDV-[IN] (U15017). Therefore, these DNA-B sequences should be regarded as the components of the isolates of ToLCNDV (Table 4.1). In addition, a sequence from New Delhi (N3B) shared the highest nucleotide identity with the DNA-B component of Tomato leaf curl Gujarat virus (ToLCGV; AY190291) (Table 4.1). Hence, all these DNA-B components were found to be clustered alongwith their respective species (ToLCGV or ToLCNDV) in the phylogenetic dendrogram (Figure 4.5).

4.2.2.3 Genetic distribution of betasatellites associated with Indian ChiLCD

Majority of the monopartite and bipartite begomoviruses identified from ChiLCD-infected samples were found to be associated with the betasatellites (Tables 4.1, 4.2). The size of the cloned betasatellites was determined to be 1350-1379 nucleotides. On the basis of betasatellite demarcation threshold limitation (78%; Briddon *et al*., 2002) and pairwise nucleotide sequence identity scores, the isolated betasatellites can be broadly classified as Tomato leaf curl Bangladesh betasatellite (ToLCBDB), Radish leaf curl betasatellite (RaLCB), Tomato leaf curl Joydebpur betasatellite (ToLCJoB), Croton yellow vein mosaic betasatellite (CroYVMB) and Tomato leaf curl Ranchi betasatellite (ToLCRnB) (Figure 4.6; Table 4.2). The G1β, K4β and T3β sequences had >99% identity with RaLCB (EF175734) (Figure 4.6). The sequences of M4β, T2β and T4β showed maximum pairwise identities of 90.3-91.5% with CroYVMB (AM410551). Therefore, the betasatellites belonging to CroYVMB and RaLCB were placed separately in the phylogenetic dendrogram (Figure 4.7). Out of 33 betasatellites identified from the ChiLCD-infected samples, 27 were found to be clustered with tomato-associated betasatellite groups such as ToLCJoB, ToLCBDB and ToLCRnB (Figure 4.6; Table 4.2). The sequences of H2β and O1β displayed the highest pairwise identity of 91.5-99.4% with ToLCRnB (GQ994096). The sequences A2β, M2β, U1β, W1β, W2β, W3β and W5β shared the maximum nucleotide identity between 93-94% with ToLCJoB (AJ966244) (Figure 4.6). These ToLCJoBs were found to have 80.6% and 99.6% identity between themselves. The other 18 isolates possessed pairwise nucleotide identities of 88.4-98.2% with ToLCBDB (DQ343289, AY438558). Further, the nucleotide sequence identity among these ToLCBDBs was observed to be 87.7-99.7% (Figure 4.6). Chilli leaf curl betasatellites (ChiLCB) reported from Pakistan was found to separate the ToLCBDBs and ToLCJoBs clusters in the phylogenetic dendrogram (Figure 4.7).

Table 4.2. List of betasatellites identified from the ChiLCD-affected samples from major chilli growing regions in India.

State	Area	Year	GPS co-ordinates	Clone name	Accession number	Betasatellites identified*
Andhra Pradesh	Guntur	2009	16.30°N 80.45°E	A1β	JN663859	ToLCBDB
Assam	Jorehat	2009	26.44°N 94.10°E	A2β	JN663862	ToLCJoB
Bihar	Chhapra	2010	25.78°N 84.72°E	B1β	JN663854	ToLCBDB[1]
				B2β	JN663855	ToLCBDB[1]
	Patna	2008	25.36°N 85.7°E	B3β	HM007118	ToLCBDB
Goa	Keri	2011	26.67°N 73.18°E	G1β	JN663858	RaLCB
Gujarat	Ahmedabad	2009	23.03°N 72.58°E	G2β	JN663847	ToLCBDB
Haryana	Pataudi	2007	28.35°N 76.75°E	H1β	EU582020	ToLCBDB
Himachal Pradesh	Palampur	2009	32.11°N 76.53°E	H2β	JN663872	ToLCRnB
Karnataka	Bengaluru	2010	13.81°N 77.29°E	K1β	JN663849	ToLCBDB
	Tumkur	2008	13.20°N 77.08°E	K2β	JN663875	ToLCBDB
Kerala	Vellanad	2008	8.34°N 77.30°E	K3β	JN663877	RaLCB[2]
				K4β	JN663876	ToLCBDB[2]
Madhya Pradesh	Jabalpur	2008	23.10°N 79.59°E	MIβ	JN663860	ToLCBDB
Maharashtra	Nagpur	2009	21.8°N 79.5°E	M2β	JN663863	ToLCJoB
	Pune	2009	18.31°N 73.51°E	M4β	JN663878	CroYVMB
New Delhi	New Delhi	2010	28.36°N 77.12°E	N1β	JN663868	ToLCBDB[3]
				N2β	JN663869	ToLCBDB[3]
Orissa	Bhubaneswar	2009	20.14° N 85.50°E	O1β	JN663851	ToLCRnB
Rajasthan	Jodhpur	2009	26.18°N 73.04°E	R1β	HM007105	ToLCBDB
Tamil Nadu	Coimbatore	2008	11.0°N 76.58°E	T1β	JN663856	ToLCBDB[4]
				T2β	JN663857	CroYVMB[4]
	Salem	2008	11.44°N 78.4°E	T3β	JN663873	RaLCB[5]
				T4β	JN663874	CroYVMB[5]
Uttar Pradesh	Jaunpur	2007	25.44°N 82.41°E	U1β	HM007103	ToLCJoB
	Ghazipur	2010	25.35°N 83.34°E	U2β	HM007099	ToLCBDB
	Kanpur	2008	26.28°N 80.21°E	U4β	HM007107	ToLCBDB
	Noida	2007	28.57°N 77.32°E	U5β	HM007115	ToLCBDB
	Varanasi	2006	25.20°N 83.0°E	U6β	EF190215	ToLCBDB
	Baruipur	2008	22.20°N 88.26°E	W1β	JN704344	ToLCJoB
West Bengal	Kalyani	2006	22.59°N 88.29°E	W2β	EF190216	ToLCJoB
	Kolkata	2008	22.34°N 88.22°E	W3β	HM007110	ToLCJoB
	Magrahat	2007	21.41°N 88.4°E	W5β	HM007112	ToLCJoB

*clones followed by the same numbers (in superscript) indicate that they were obtained from the same plant in that particular location.

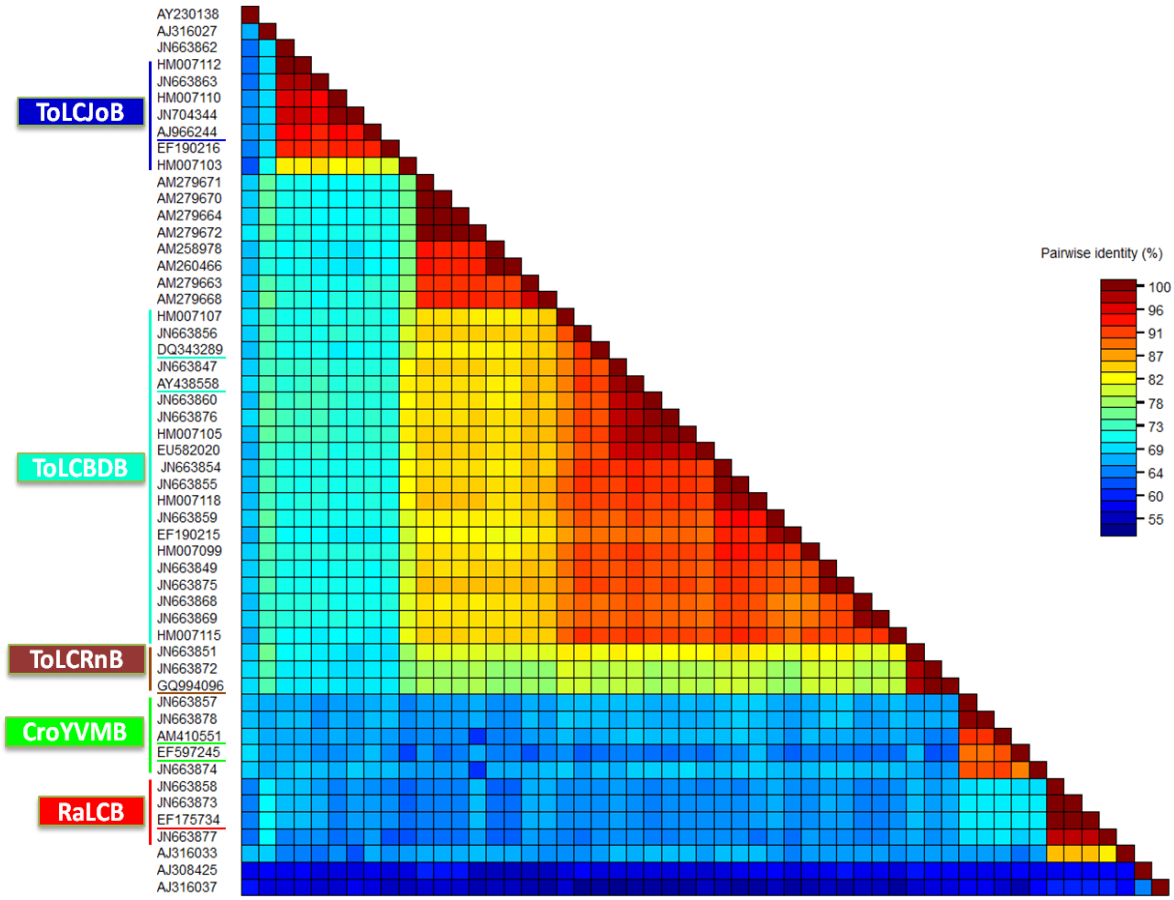


Figure 4.6. Graphical representation of per cent nucleotide identity plot of betasatellites using the SDTv1.0 program. The representative betasatellite sequences for each of the betasatellite group obtained from database are underlined.

The G1β, K4β and T3β sequences shared >99% identity with RaLCB (EF175734) (Figure 4.6). The sequences of M4β, T2β and T4β showed the highest pairwise identities of 90.3-91.5% with CroYVMB (AM410551). Therefore, the betasatellites belonging to CroYVMB and RaLCB were placed themselves separately in the phylogenetic dendrogram (Figure 4.7). Out of 33 betasatellites identified with ChiLCD-infected samples, 27 were found to be grouped with tomato-associated betasatellites such as ToLCJoB, ToLCBDB and ToLCRnB (Figure 4.6; Table 4.2). The sequences of H2β and O1β displayed maximum pairwise identity of 91.5-99.4% with ToLCRnB (GQ994096). The sequences A2β, M2β, U1β, W1β, W2β, W3β and W5β shared the maximum nucleotide identity between 93-94% with ToLCJoB (AJ966244) (Figure 4.6). These ToLCJoBs were found to have 80.6% and 99.6% identity between themselves. The other 18 isolates possessed pairwise nucleotide identity of 88.4-98.2% with ToLCBDB (DQ343289, AY438558). Further, the nucleotide sequence identity among these ToLCBDBs was observed to be 87.7-99.7% (Figure 4.6).

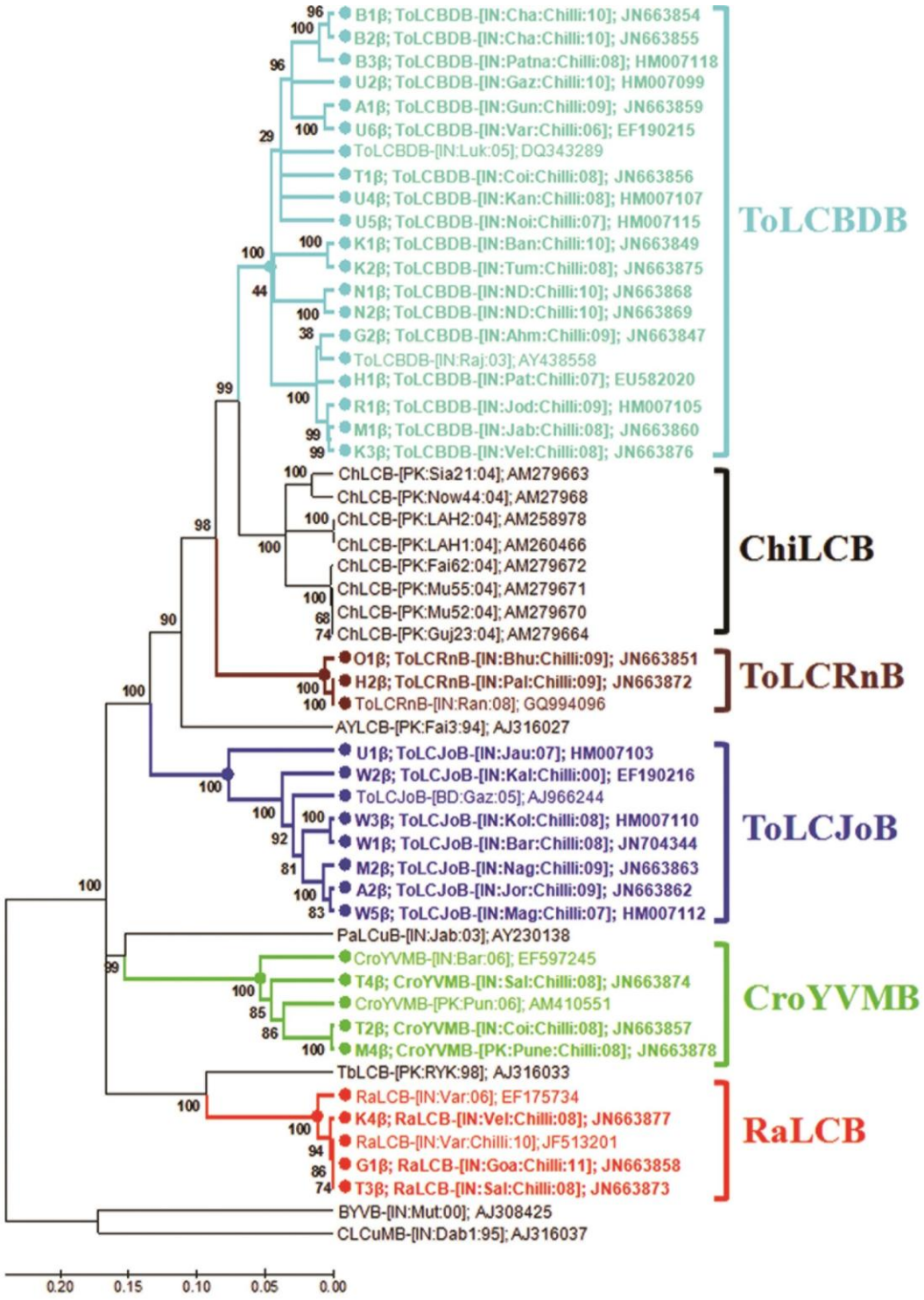


Figure 4.7. Neighbor-joining phylogenetic tree of ChiLCD-associated betasatellites isolated here (indicated in bold type) along with the representative betasatellites retrieved from the GenBank. CLUSTAL W method was employed to align the sequence before the generation of the phylogenetic dendrogram. The bootstrap values of thousand replicates are provided at the nodes. Major clades in the same colour belong to a particular betasatellite group which is given at the right side of the dendrogram. The scale bar refers to nucleotide substitution rate per site.

4.2.2.4 Association of alphasatellites with Indian ChiLCD

Till date, no pairwise sequence identity criterion for demarcation of alphasatellites is available. Therefore, alignment of 328 alphasatellite sequences (including those reported in this study) was carried out using SDT v.1.0 and a pairwise identity matrix was generated. Two major clusters between 59-69% and 76-87% are evident in the pairwise distance comparison of the alphasatellites reported worldwide (Figure 4.8). Based on this pairwise cluster analysis, it is proposed to apply 84% pairwise identity as a criterion for the alphasatellite demarcation. Following this criterion, 13 alphasatellites were categorized into two alphasatellite groups (Figures 4.9, 4.10).

Table 4.3. Alphasatellites identified from the ChiLCD-affected samples from major chilli growing regions in India.

State	Area	Year	GPS co-ordinates	Accession number	Alphasatellites identified
Andhra Pradesh	Guntur	2009	16.30°N 80.45°E	KF471043	ChiLCA
Assam	Jorehat	2009	26.44°N 94.10°E	KF471050	ChiLCA
New Delhi	New Delhi	2010	28.36°N 77.12°E	KF471051	ChiLCA
Himachal Pradesh	Palampur	2009	32.11°N 76.53°E	KF471052	ChiLCA
Jharkhand	Dhanbad	2008	28.00°N 86.45°E	KF471058	GDarSLA
	Ramgarh	2008	30.78°N 76.55°E	KF471049	ChiLCA
Karnataka	Bengaluru	2010	13.81°N 77.29°E	KF471046	ChiLCA
Madhya Pradesh	Jabalpur	2008	23.10°N 79.59°E	KF471044	ChiLCA
	Teherka	2008	24.42°N 78.52°E	KF471039	GDarSLA
Rajasthan	Bharatpur	2009	27.22°N 77.48°E	KF471048	ChiLCA
Tamil Nadu	Villupuram	2008	11.95°N 79.52°E	KF471040	GDarSLA
Uttar Pradesh	Varanasi	2006	25.20°N 83.00°E	KF471037	GDarSLA
West Bengal	Magrahat	2007	21.68°N 88.07°E	KF471045	ChiLCA

With 98-99% nucleotide identity among them, nine alphasatellites (KF471043-46, KF471048-52) isolated from the chilli share the highest sequence identities (81.4-83.4%) with the isolates of Cotton leaf curl Burewala/Multan alphasatellite (CLCuBuA/CLCuMuA) (Figure 4.10). According to the proposed species demarcation threshold (Figure 4.8), these alphasatellites belong to a previously unreported alphasatellite group, Chilli leaf curl alphasatellite (ChiLCA). Furthermore, in the phylogenetic tree, these alphasatellites positioned themselves in a distinct clade (Figure 4.9). The alphasatellites (KF471037- KF471040, KF471054) isolated from a tomato and 3

chilli samples shared maximum per cent nucleotide identity of >87% to already reported alphasatellite group, Gossypium darwinii symptomless alphasatellite (GDarSLA) (Figure 4.10). These isolates possess 97-100% nucleotide identity between each other (Figure 4.10). In consistent with their pairwise sequence identity scores, these GDarSLAs positioned themselves distinctly in the phylogenetic dendrogram (Figure 4.9).

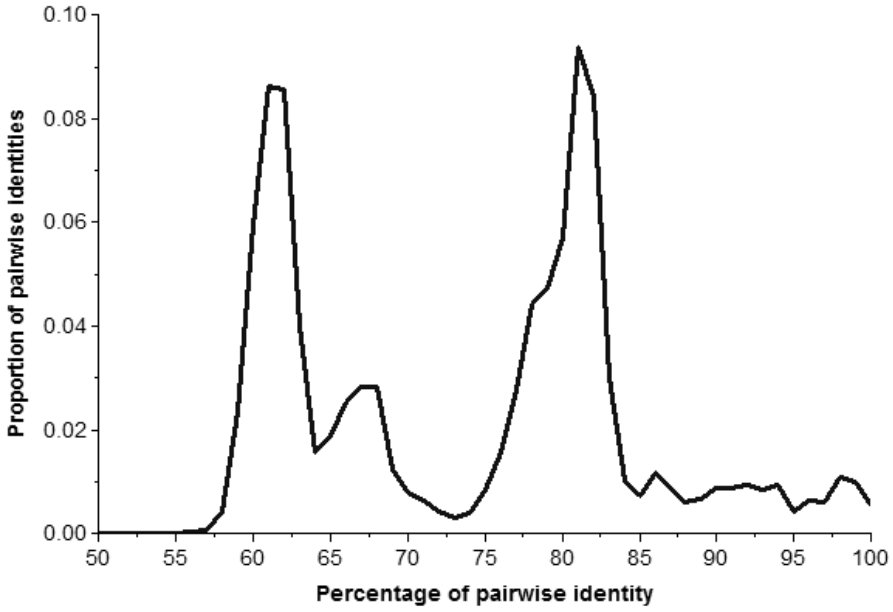


Figure 4.8. Distribution of per cent pairwise identity between the nucleotide sequences of 326 alphasatellites. The pairwise identities were estimated using SDT program after aligning the full-length sequences of 326 alphasatellites including those reported in this study using the MUSCLE method.

4.2.3 Recombination analysis

4.2.3.1 Detection of recombination among Indian ChiLCD associated begomoviruses

In RDP analysis of DNA-A components, significant evidence of recombination in 26 out of 35 isolates were detected (Table 4.4). In ChiLCV-PK [Jorehat, Palampur] isolates, significant breakpoints were located encompassing the AC1 region corresponding to the genome of ChiLCV-[BD-Gaz] and the CR region of Jorehat isolate corresponding to ChiLCV-PK [Mul]. The AV1 fragment in ChiLCV-IN [Ahmedabad, Noida] and [Jodhpur, New Delhi] isolates was probably from ChiLCV-PK[Mul] and PepLCBV, respectively (Table 4.4). However, AV1 of ChiLCV-IN [Pataudi] isolate was probably contributed by ToLCV-Ban whereas the AC1/CR region of this ChiLCV-IN isolate corresponds to various strains of ChiLCVs (Table 8). However, recombination fragments around AC1 and AC2/AC3 in ChiLCV-Gun isolates were found to be contributed by

PepLCBV. Similarly, the recombination fragments detected in ChiLCV-Chi [Baruipur, Chhapra1] was donated by either PepLCBV or ChiLCV strains (Table 4.4). Additionally, in the ChiLCV isolates cloned here, the recombination fragments were mostly contributed by different isolates belonging to the ChiLCV species (Table 4.4). In PepLCBV-BD [Magrahat, Ghazipur1, Chhapra2] isolates, the recombination events was found in the AV1 region, indicating that these fragments were either from ChiLCV-[BD-Gaz] or ChiLCV-Chi (Table 4.4). However, the AC1 region of PepLCBV-BD [Magrahat, Chhapra2] isolates belong to either unknown species or PepLCBV-PK. RDP analysis revealed that all PepLCBV [Jabalpur, Kanpur] isolates identified here possess a single recombination event around AC1 region which was donated by PepLCBV-PK (Table 4.4).

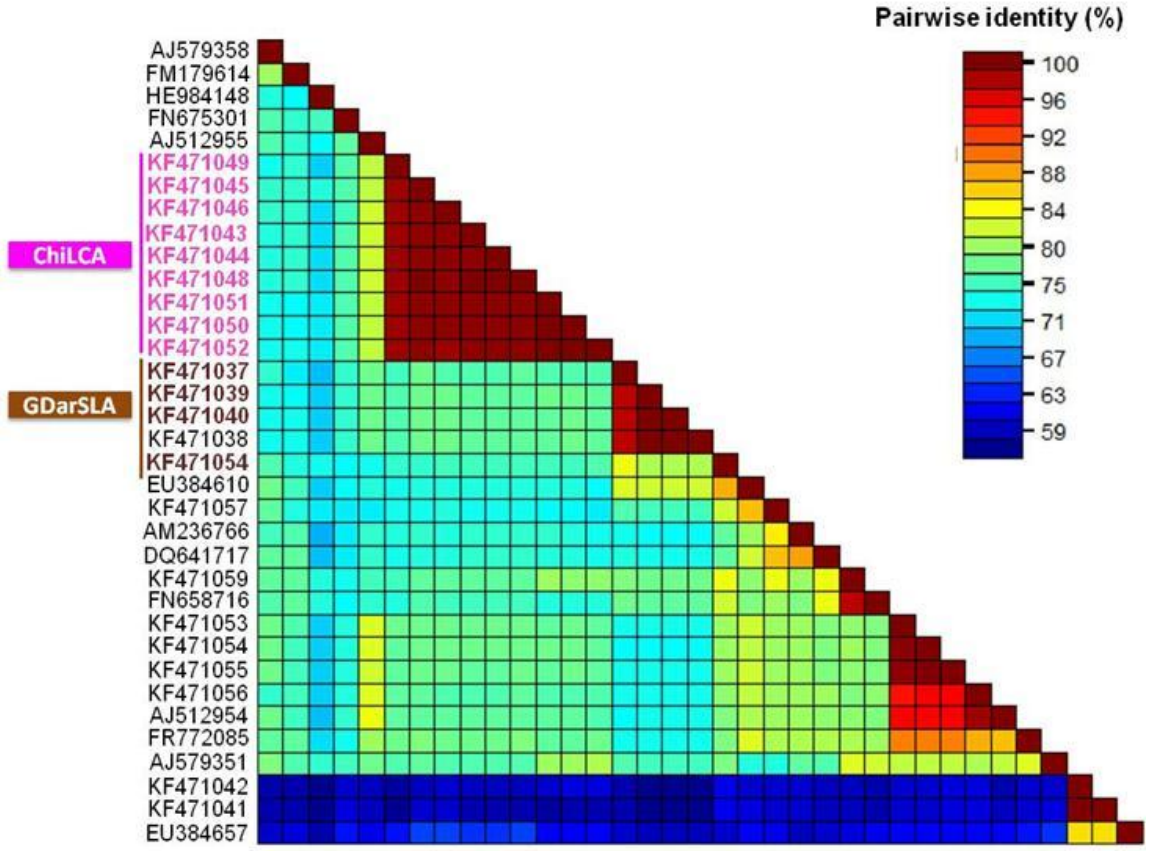


Figure 4.9. Graphical representation of per cent pairwise nucleotide identity plot of alphasatellites using SDTv1.0 program. The alphasatellite sequences identified are provided in bold letters. Different alphasatellite groups are provided in different colour-codes.

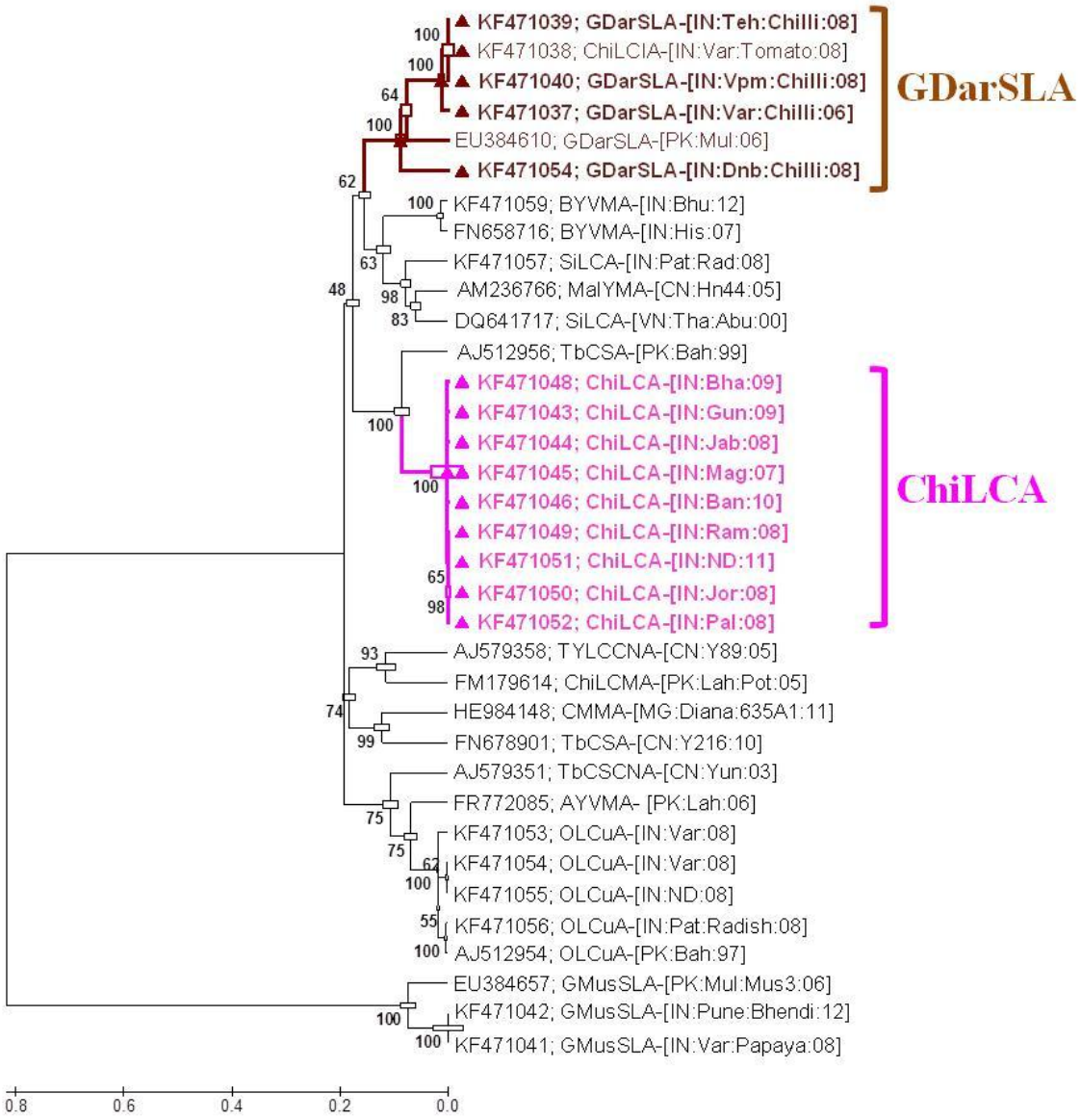


Figure 4.10. Neighbor-joining phylogenetic tree of ChiLCD-associated alphasatellites isolated in this study (indicated in bold type) along with the representative alphasatellites retrieved from the GenBank. The bootstrap values of thousand replicates are provided at the nodes. Major clades in the same colour belong to a particular alphasatellite group. The scale bar refers to nucleotide substitution rate per site.

The CR regions in Coimbatore and Nagpur isolates were probably originated from PepLCBV-BD, PaLCuv-PK was noticed in contributing to the AC1 region of PepLCBV-IN [Coimbatore] isolate and the CR of PepLCBV-IN [Nagpur] isolate. Recombination analysis of ChiLCSV revealed that breakpoints in the AV1 and AC1 regions were possibly derived from PapLCBV and PaLCuV, respectively. Similarly, the AC1 region of ToLCV-Ban [Bengaluru] isolate was contributed by ToLCBaV-D (Table 4.4).

Table 4.4. Recombination analysis of cloned begomoviruses identified from the ChiLCD-affected samples.

Recombinant begomoviruses	Breakpoint positions			Putative minor parent	Putative major parent	Methods detected[@]	Detected *p*-value[#]
	Begin	End	Location*				
ChiLCSV-[IN-Sal-08]	454	1190	AV1	PepLCLV-[PK:Lah:04]	ToLCSLV-[LK:Ban:97]	RMC	7.28 X10^{-11}
	2146	2712	AC1	PaLCuV-PK[PK:Cot:02]	ToLCSLV-[LK:Ban:97]	GBMC	1.94 X10^{-24}
ChiLCV-PK[IN-Jor-09]	1997	2591	AC1	ToLCJoV-BD[BD]	PepLCLV-[PK:Lah:04]	RGBMCS	3.78 X10^{-27}
	2592	2742	CR	ChiLCV-PK[PK:Mul:98]	PepLCLV-[PK:Lah:04]	RGBMC	4.56 X10^{-11}
ChiLCV-PK[IN-Pal-09]	2083	2582	AC1	ToLCJoV-BD[BD]	PepLCLV-[PK:Lah:04]	RGBMCS	1.51 X10^{-26}
ChiLCV-IN[IN-Ahm-09]	512	1084	AV1	ChiLCV-PK[PK:Mul:98]	ToLCJoV-IN[IN::05]	RGBMC	1.34 X10^{-21}
	1230	1790	AC2/AC3	PepLCBDV-BD[BD:Bog:99]	ToLCJoV-IN[IN::05]	RGBMCS	1.76 X10^{-08}
ChiLCV-Gun[IN-Gun-08]	502	1151	AV1	ToLCKVBan-[IN:Ban:93]	ChiLCV-PK[PK:Mul:98]	RBMCS	7.17 X10^{-07}
	2031	2595	AC1	ToLCJoV-BD[BD]	ChiLCV-PK[PK:Mul:98]	RGBMCS	3.65 X10^{-24}
ChiLCV-IN[IN-Jod-09]	323	940	AV1	PepLCLV-[PK:Lah:04]	ChiLCV-PK[PK:Mul:98]	RGBS	8.36 X10^{-33}
	2031	2695	AC1	ToLCJoV-IN[IN::05]	ChiLCV-PK[PK:Mul:98]	RGBMCS	3.06 X10^{-16}
ChiLCV-Gun [IN-Nag1-09] and ChiLCV-Gun[IN-Goa-11]	1090	1359	AC2/AC3	PepLCLV-[PK:Lah:04]	ChiLCV-PK[PK:Mul:98]	RGBMS	4.92 X10^{-12}
	2412	1	CR	PepLCBDV-BD[BD:Bog:99]	ChiLCV-PK[PK:Mul:98]	RGBMS	1.89 X10^{-09}
ChiLCV-IN[IN-ND-10]	7	1438	AV1	PepLCLV-[PK:Lah:04]	ToLCJoV-IN[IN::05]	RGBMCS	2.37 X10^{-19}
	1680	2123	AC1	ChiLCV-PK[PK:Mul:98]	ToLCJoV-IN[IN::05]	RGBMCS	7.95 X10^{-09}
ChiLCV-IN[IN-Noi-07]	47	1086	AV1	ChiLCV-PK[PK:Mul:98]	ToLCJoV-IN[IN::05]	RGBMCS	6.34 X10^{-23}
	1786	7	AC1/CR	ToLCJoV-IN[IN::05]	ChiLCV-PK[PK:Mul:98]	RGBMCS	7.28 X10^{-28}
ChiLCV-IN[IN-	441	1076	AV1	ToLCKVBan-	ToLCJoV-	RGBMC	8.80 X10^{-08}

Pat-07]				[IN:Ban:93]	IN[IN::05]		
	2006	23	AC1/CR	ToLCJoV-IN[IN::05]	ChiLCV-PK[PK:Mul:98]	RGBMCS	1.83 X10^{-27}
PepLCBV-BD[IN-Mag-07]	543	1153	AV1	ToLCJoV-BD[BD]	PepLCBDV-PK[PK:Kha:04]	RGBS	1.72 X10^{-16}
	1483	1642	AC1	Unknown species	PepLCBDV-PK[PK:Kha:04]	GMCS	1.29 X10^{-60}
PepLCBV-BD[IN-Gha1-10]	541	928	AV1	ToLCJoV-BD[BD]	PepLCBDV-PK[PK:Kha:04]	RGBMCS	4.82 X10^{-11}
	963	1165	AV1	Unknown species	PepLCBDV-PK[PK:Kha:04]	GMC	1.77 X10^{-04}
PepLCBV-PK[IN-Jab-07]	2136	2675	AC1	PepLCBDV-PK[PK:Kha:04]	ToLCJoV-IN[IN::05]	RGBMCS	2.20 X10^{-28}
PepLCBV-BD[IN-Chp2-10]	8	691	AV1	ToLCJoV-IN[IN::05]	PepLCBDV-PK[PK:Kha:04]	RGBMCS	2.31 X10^{-14}
	2401	2689	AC1	PepLCBDV-PK[PK:Kha:04]	Unknown species	RGBMCS	5.70 X10^{-22}
PepLCBV-Kan [IN-Kan-08]	1854	2689	AC1	PepLCLV-[PK:Lah:04]	PepLCBDV-PK[PK:Kha:04]	RGBMC	9.09 X10^{-11}
PepLCBV-IN[IN-Coi-08]	2357	2699	AC1	PaLCuV-IN[IN:Luc]	PepLCLV-[PK:Lah:04]	RGBMC	4.11 X10^{-10}
	2724	407	CR	PepLCBDV-BD[BD:Bog:99]	PepLCLV-[PK:Lah:04]	RGBMCS	7.06 X10^{-10}
PepLCBV-IN[IN-Nag1-09]	414	1092	AV1	PaLCuV-PK[PK:Cot:02]	PepLCLV-[PK:Lah:04]	RBMCS	2.44 X10^{-05}
	2619	2702	CR	PepLCBDV-PK[PK:Kha:04]	PepLCLV-[PK:Lah:04]	GBMC	4.71 X10^{-08}
ChiLCV-Chi[IN-Chp1-10]	2713	8	CR	PepLCBDV-BD[BD:Bog:99]	ChiLCV-PK[PK:Mul:98]	RGBMCS	5.45 X10^{-24}
ChiLCV-Chi[IN-Bar-08]	508	883	AV1	ToLCJoV-BD[BD]	ToLCJoV-IN[IN::05]	RGBMCS	1.67 X10^{-16}
	2410	2691	AC1	ToLCJoV-IN[IN::05]	PepLCBDV-PK[PK:Kha:04]	GBMCS	3.58 X10^{-27}
ChiLCV-[IN-Kal-06]	546	962	AV1	ToLCJoV-BD[BD]	Unknown species	RGBMCS	3.07 X10^{-14}
	2115	2588	AC1	ChiLCV-PK[PK:Mul:98]	PepLCBDV-PK[PK:Kha:04]	RGBMCS	1.45 X10^{-24}
ChiLCV[IN-Kol-08]	354	1160	AV1	ToLCJoV-BD[BD]	Unknown species	GMCS	1.62 X10^{-31}
ChiLCV-[IN-	613	1262	AV1	PepLCBDV-	Unknown	RGBMCS	6.77 X10^{-25}

Patna-08]			PK[PK:Kha:04]	species		
	2097	2586	AC1	ChiLCV-PK[PK:Mul:98]	PepLCBDV-PK[PK:Kha:04]	RGBMCS 3.99 X10^{-24}
ChiLCV-[IN-Jau-07]	2564	2721	CR	ToLCJoV-BD[BD]	Unknown species	GBMC 4.17 X10^{-09}
ChiLCV[IN-Gha2-10]	2045	2240	AC1	ChiLCV-PK[PK:Mul:98]	ToLCJoV-BD[BD]	RGBS 3.55 X10^{-08}
ToLCV-Ban[IN-Ban- Chilli-10]	2238	2694	AC1	ToLCKeV-[IN:KerII:05]	ChiLCV-PK[PK:Mul:98]	RGBMCS 4.18 X10^{-15}

*CR, common region of the begomovirus genome.

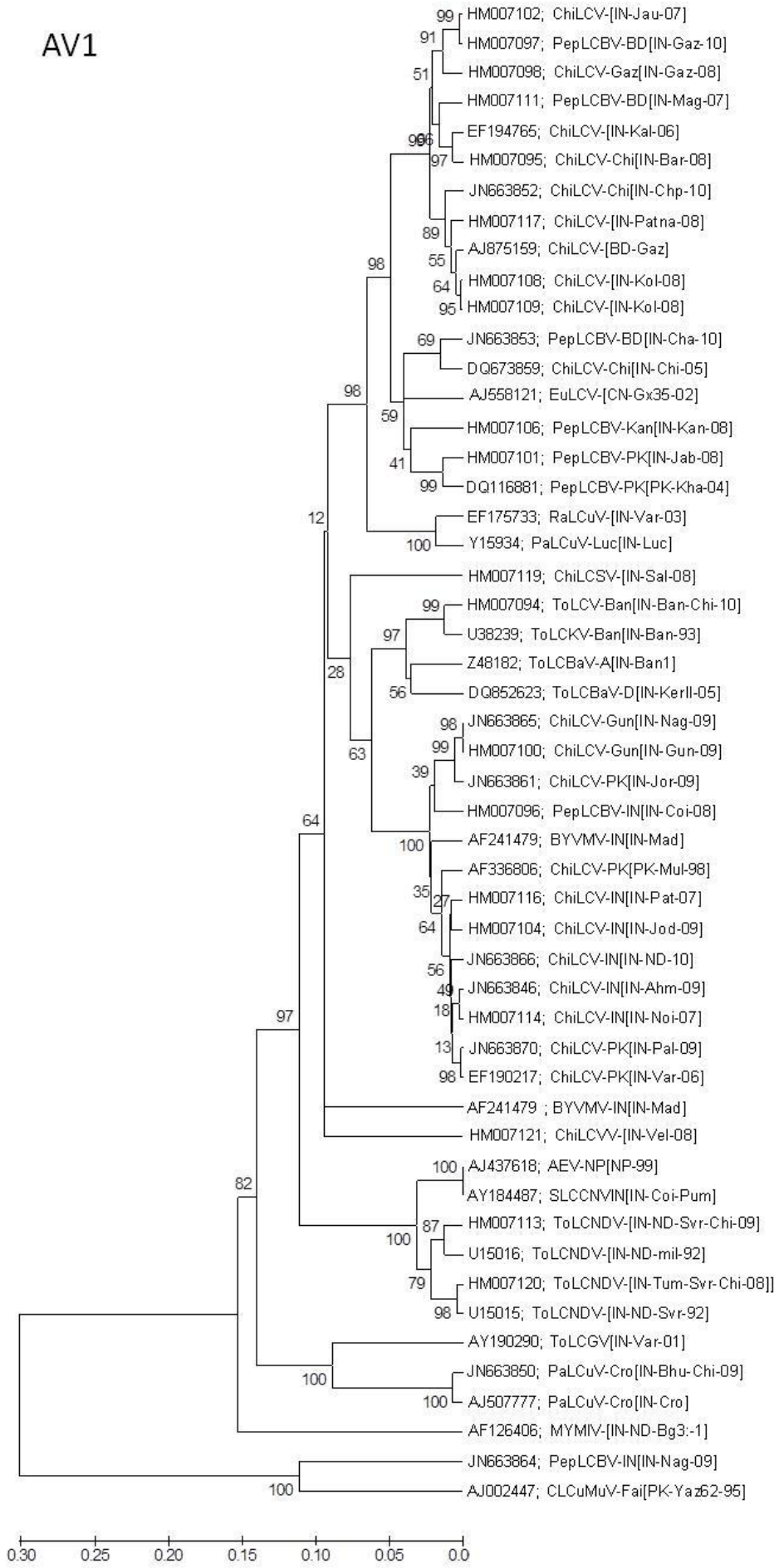
[@]R, RDP; G, GenConv; B, Bootscan; M, MaxChi; C, Chimaera; S, SiScan.

[#]The lowest P-value calculated for the underlined method in the penultimate column.

Interestingly, all the methods available in RDP program have failed to detect any recombination region in the ChiLCVV genome. In the phylogenetic dendrogram of AV1 and AC1 ORFs, the sequences from recombinant molecules were, as anticipated, were segregated alongwith their respective contributors (Figure 4.11). RDP analysis also suggests that the recombination fragments noticed in these ChiLCD-associated begomoviruses were mostly found to be encompassing either AC1 or AV1 regions.

4.2.3.2 Recombination pattern detected among ChiLCD-associated betasatellites

Out of 33 betasatellites identified with ChiLCD-infected samples, 17 betasatellite sequences were found to possess recombination fragments (Table 4.5). As evident from the RDP analysis, SCR and the complete βC1 region of ToLCJoB identified from Jaunpur were shared with ToLCBDB (Table 4.5). A recombination fragment encompassing the βC1 region of ToLCRnB [Bhubaneswar] corresponds to ToLCBDB, whereas the SCR regions of ToLCJoB [Baruipur, Nagpur] and ToLCJoB [Kalyani, Kolkata, Jorehat, Magrahat] were contributed by ToLCBDB and ChLCB, respectively. A recombination event detected around the A-rich region of both CroYVMBs was contributed by Tomato leaf curl betasatellite (Table 4.5). Seven of the ToLCBDB isolates have a recombination fragment encompassing the A-rich region in which either ChLCB or an unknown betasatellite acts as the major parent.


AV1
HM007102; ChiLCV-[IN-Jau-07]
HM007097; PepLCBV-BD[IN-Gaz-10]
HM007098; ChiLCV-Gaz[IN-Gaz-08]
HM007111; PepLCBV-BD[IN-Mag-07]
EF194765; ChiLCV-[IN-Kal-06]
HM007095; ChiLCV-Chi[IN-Bar-08]
JN663852; ChiLCV-Chi[IN-Chp-10]
HM007117; ChiLCV-[IN-Patna-08]
AJ875159; ChiLCV-[BD-Gaz]
HM007108; ChiLCV-[IN-Kol-08]
HM007109; ChiLCV-[IN-Kol-08]
JN663853; PepLCBV-BD[IN-Cha-10]
DQ673859; ChiLCV-Chi[IN-Chi-05]
AJ558121; EuLCV-[CN-Gx35-02]
HM007106; PepLCBV-Kan[IN-Kan-08]
HM007101; PepLCBV-PK[IN-Jab-08]
DQ116881; PepLCBV-PK[PK-Kha-04]
EF175733; RaLCuV-[IN-Var-03]
Y15934; PaLCuV-Luc[IN-Luc]
HM007119; ChiLCSV-[IN-Sal-08]
HM007094; ToLCV-Ban[IN-Ban-Chi-10]
U38239; ToLCKV-Ban[IN-Ban-93]
Z48182; ToLCBaV-A[IN-Ban1]
DQ852623; ToLCBaV-D[IN-KerII-05]
JN663865; ChiLCV-Gun[IN-Nag-09]
HM007100; ChiLCV-Gun[IN-Gun-09]
JN663861; ChiLCV-PK[IN-Jor-09]
HM007096; PepLCBV-IN[IN-Coi-08]
AF241479; BYVMV-IN[IN-Mad]
AF336806; ChiLCV-PK[PK-Mul-98]
HM007116; ChiLCV-IN[IN-Pat-07]
HM007104; ChiLCV-IN[IN-Jod-09]
JN663866; ChiLCV-IN[IN-ND-10]
JN663846; ChiLCV-IN[IN-Ahm-09]
HM007114; ChiLCV-IN[IN-Noi-07]
JN663870; ChiLCV-PK[IN-Pal-09]
EF190217; ChiLCV-PK[IN-Var-06]
AF241479 ; BYVMV-IN[IN-Mad]
HM007121; ChiLCVV-[IN-Vel-08]
AJ437618; AEV-NP[NP-99]
AY184487; SLCCNVIN[IN-Coi-Pum]
HM007113; ToLCNDV-[IN-ND-Svr-Chi-09]
U15016; ToLCNDV-[IN-ND-mil-92]
HM007120; ToLCNDV-[IN-Tum-Svr-Chi-08]]
U15015; ToLCNDV-[IN-ND-Svr-92]
AY190290; ToLCGV[IN-Var-01]
JN663850; PaLCuV-Cro[IN-Bhu-Chi-09]
AJ507777; PaLCuV-Cro[IN-Cro]
AF126406; MYMIV-[IN-ND-Bg3:-1]
JN663864; PepLCBV-IN[IN-Nag-09]
AJ002447; CLCuMuV-Fai[PK-Yaz62-95]
0.30
0.25
0.20
0.15
0.10
0.05
0.0

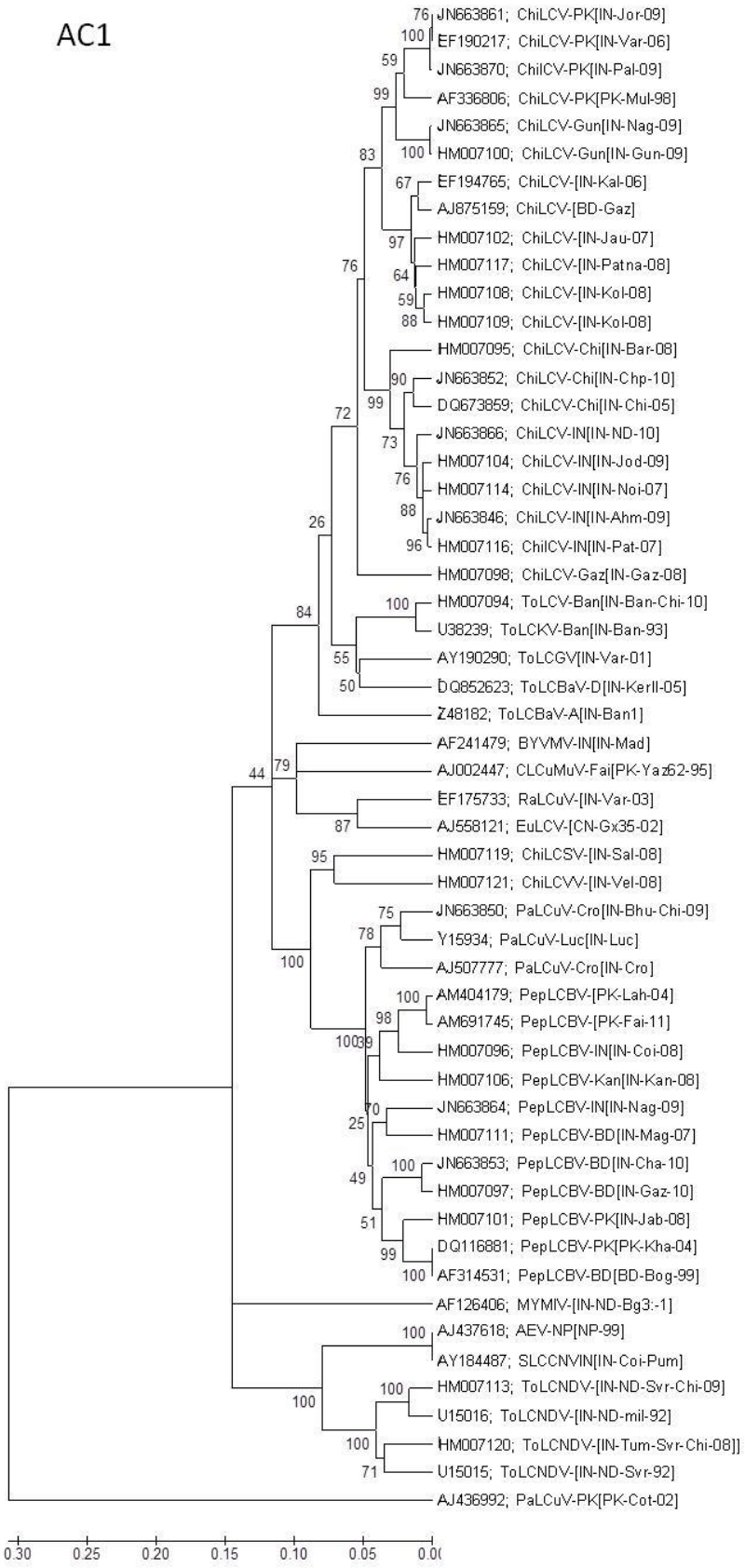


Figure 4.11. Phylogenetic tree of begomovirus DNA-A encoded AV1 and AC1 ORFs. All the sequences were aligned using Clustal W available in MEGA5, prior to tree generation.

Table 4.5. Recombination analysis of cloned betasatellites from the ChiLCD-affected samples.

Recombinant betasatellites	Breakpoint positions			Putative minor parent	Putative major parent	Methods detected[@]	Detected *p*-value[#]
	Begin	End	Location*				
CroYVMB-[IN-Sal-Chilli-08]	779	963	A-rich	ToLCB-[IN:Var:06]	PaLCuB-[IN:Jab:03]	RGBMC	9.49 X10^{-10}
CroYVMB-[IN-Pune-Chilli-09]	832	953	A-rich	ToLCB-[IN:Var:06]	PaLCuB-[IN:Jab:03]	RGBMC	3.25 X10^{-09}
ToLCBDB-[IN-Gha1-Chilli-10]	519	1161	A-rich	ToLCBDB-[BD:Gaz:01]	ChLCB-[PK:Sia21:04]	RGBMCS	3.08 X10^{-06}
ToLCBDB-[IN-Var-Chilli-06]	984	1193	A-rich	ToLCBDB-[IN:Raj:03]	Unknown species	RGBMC	8.36 X10^{-05}
ToLCBDB-[IN-Coi1-Chilli-08]	521	1308	A-rich	ToLCBDB-[BD:Gaz:01]	ChLCB-[PK:Sia21:04]	RGBMCS	5.45 X10^{-07}
ToLCBDB-[IN-Kan-Chilli-08]	472	1366	A-rich	ToLCBDB-[BD:Gaz:01]	ChLCB-[PK:Sia21:04]	RGBMCS	1.17 X10^{-10}
ToLCBDB-[IN-Ahm-Chilli-09]	765	955	A-rich	ToLCBDB-[BD:Gaz:01]	Unknown species	GBMCS	3.78 X10^{-15}
ToLCBDB-[IN-Jod-Chilli-09]	518	1036	A-rich	ToLCBDB-[BD:Gaz:01]	ChLCB-[PK:Sia21:04]	RGBMCS	1.06 X10^{-05}
ToLCBDB-[IN-Vel2-Chilli-08]	716	1077	A-rich	ToLCBDB-[BD:Gaz:01]	ChLCB-[PK:Sia21:04]	RGBMCS	4.83 X10^{-07}
ToLCJoB-[IN-Kal-Chilli-06]	1140	95	SCR	ChLCB-[PK:Guj23:04]	AYLCB-[PK:Fai3:94]	RGBMCS	6.83 X10^{-13}
ToLCJoB-[IN-Kol-Chilli-08]	1123	89	SCR	ChLCB-[PK:Guj23:04]	AYLCB-[PK:Fai3:94]	RGBMCS	5.06 X10^{-12}
ToLCJoB-[IN-Bar-Chilli-08]	1075	1261	A-rich/ SCR	ToLCBDB-[IN:Raj:03]	Unknown species	RGBMCS	1.12 X10^{-14}
ToLCJoB-[IN-Nag-Chilli-09]	1065	1251	A-rich/ SCR	ToLCBDB-[IN:Raj:03]	Unknown species	RGBMCS	3.80 X10^{-15}
ToLCJoB-[IN-Jor-Chilli-09]	1123	92	SCR	ChLCB-[PK:Guj23:04]	AYLCB-[PK:Fai3:94]	RGBMCS	1.99 X10^{-14}
ToLCJoB-[IN-Mag-Chilli-07]	1123	92	SCR	ChLCB-[PK:Guj23:04]	AYLCB-[PK:Fai3:94]	RGBMCS	3.02 X10^{-14}
ToLCJoB-[IN-Jau-Chilli-07]	1276	521	βC1	ToLCBDB-[IN:Raj:03]	ToLCJoB-[IN::Kenaf:07]	RGBMCS	4.54 X10^{-30}
ToLCRnB-[IN-Bhu-Chilli-09]	146	499	βC1	ToLCBDB-[BD:Gaz:01]	Unknown species	GMCS	4.43 X10^{-31}

*SCR, satellite conserved region of betasatellites.

[@]R, RDP; G, GenConv; B, Bootscan; M, MaxChi; C, Chimaera; S, SiScan.

[#]The lowest P-value calculated for the underlined method in the penultimate column.

4.2.3.3 Evidence of recombination among the alphasatellites with ChiLCD in India

The recombination events among these alphasatellites were analyzed using the methods available in the RDP program. As evident from the RDP analysis, all the alphasatellite sequences associated with Indian ChiLCD were found to recombinant in nature by possessing atleast a single recombination event (Table 4.6). The SCR of 4 ChiLCA isolates [Bengaluru, Bharatpur, Guntur, Jabalpur] was contributed by TbCSA. However, ChiLCA [Magrahat] isolate shared SCR with BYVMA. Similarly, the recombination event in the SCR of GDarSLA [Teherka] corresponds to an unidentified donor (Table 4.6). The A-rich region of GDarSLA [Dhanbad] was from TbLCA but for ChiLCA [Palampur], it was donated by MalYMA. In addition, GDarSLA was found to contribute the A-rich region to ChiLCA [Ramgarh] and GDarSLA [Teherka] isolates, whereas for RaLCA (Table 4.6). The unknown alphasatellite(s) has shared a portion of the rep region to ChiLCA [Jorehat, New Delhi] and GDarSLA [Varanasi] isolates. The C-terminal portion of the rep gene in ChiLCA [Villupuram] isolate was contributed by BYVMA (Table 4.6).

4.2.4 Nucleotide diversity of ChiLCD-associated viral genomic components

From an evolutionary perspective, recombination during replication of begomoviruses might not be the sole factor contributing to the genetic variation *de novo*. Mutations and nucleotide substitutions also could contribute to this process of genetic variation. Our results suggest that begomoviruses and betasatellite molecules associated with ChiLCD were noticed to possess higher degree of genetic variability (π>0.08) (Table 4.7). However, Indian ChiLCD-associated alphasatellite found to be the least variable. Our analyses showed ChiLCV and PepLCBV to be the most diverse begomoviruses, and CroYVMB and ToLCBDB to possess much higher genetic variability than the other betasatellites (Table 4.7). However, less genetic variability was observed among both the ChiLCA and GDarSLA isolates.

To analyze the randomness of the observed nucleotide diversity throughout the genome, we analyzed the genetic variability of all the ORFs encoded by ChiLCD-associated components. Among the begomoviruses, the nucleotide diversity was found to be distributed in all the virus-encoded ORFs (Figure 4.12). The nucleotide diversity patterns of the complete genomes of betasatellites and βC1 ORFs, and alphasatellite and αRep were noticed to be highly congruent (Figure 4.12; Table 4.8).

Table 4.6. Recombination analysis of cloned alphasatellites from the ChiLCD-affected samples.

Recombinant alphasatellites	Probable parental sequences		Breakpoint positions		Methods detected[@]	Detected *p*-value range
	Minor parent	Major parent	Begin	End		
ChiLCA-[IN-Ban-10]	TbCSA[CN:Y216:10]	CLCuBurA [IN:Sri:10]	1345	85	RGMC	$4.817X10^{-4}$-$1.823X10^{-7}$
ChiLCA-[IN-Bha-09]	TbCSA[CN:Y216:10]	CLCuBurA [IN:Sri:10]	1350	90	RGMC	$3.444X10^{-4}$-$7.958X10^{-7}$
ChiLCA-[IN-Gun-09]	TbCSA[CN:Y216:10]	CLCuBurA [IN:Sri:10]	1304	70	RBMC	$1.220X10^{-3}$-$8.357X10^{-4}$
ChiLCA-[IN-Jab-08]	TbCSA[CN:Y216:10]	CLCuBurA [IN:Sri:10]	1350	90	RGMC	$3.515X10^{-4}$-$7.939X10^{-7}$
ChiLCA-[IN-Jor-09]	Unknown sequence	SiYVVNA [VN:Han:06]	801	1049	RBMCS	$5.096X10^{-3}$-$3.419X10^{-6}$
ChiLCA-[IN-Mag-07]	BYVMA[IN:G2-8:02]	Unknown sequence	1330	58	RBMS	$6.382X10^{-3}$-$3.299X10^{-12}$
ChiLCA-[IN-ND-11]	Unknown sequence	TbCSA [CN:Y216:10]	135	358	RGBCS	$9.316X10^{-3}$-$6.003X10^{-5}$
ChiLCA-[IN-Pal-09]	TbLCA[PK:NOB2:02]	AYVA [PK:Lah:06]	1105	1142	GMC	$4.836X10^{-3}$-$1.262X10^{-4}$
ChiLCA-[IN-Ram-08]	GDarSLA [PK:Mul:08]	GMusSLA [PK:Mul:08]	1040	1200	GMC3	$1.652X10^{-4}$-$3.892X10^{-9}$
GDarSLA -[IN-Teh-08]	GDarSLA [PK:Mul:08]	GMusSLA [PK:Mul:08]	960	1120	RMCS3	$5.260X10^{-4}$-$1.526X10^{-7}$
	GMusSLA [PK:Mul:08]	MalYMA [CN:Hn44:06]	1310	30	GMS	$2.055X10^{-3}$-$4.540X10^{-6}$
GDarSLA -[IN-Var-08]	Unknown sequence	GDarSLA [PK:Mul:08]	940	1120	RGBMCS3	$2.967X10^{-7}$-$1.027X10^{-22}$
GDarSLA -[IN-Vpm-08]	BYVMA [IN:Har:07]	Unknown sequence	425	990	RGMCS	$8.565X10^{-5}$-$2.694X10^{-32}$
GDarSLA-[IN-Dnd-Chilli-08]	TbLCA [PK:NOB2:02]	TYLCCNA [CN:Y89:03]	1000	1255	GBMCS3	$2.873X10^{-4}$-$6.800X10^{-7}$

[@]R, RDP; G, GenConv; B, Bootscan; M, MaxChi; C, Chimaera; S, SiScan, 3, 3Seq.

Table 4.7. Genetic variability among the isolates of begomoviruses, alphasatellites and betasatellites associated with ChiLCD in India.

Viral components	N	Size	s	η	π	k	θ-w	θ-η
Begomoviruses								
ChiLCV	19	2790	708	849	0.0934	253.5	0.0749	0.0997
PepLCBV	7	2743	576	690	0.0962	263.9	0.0857	0.1160
ToLCNDV	5	2740	134	134	0.0198	53.8	0.0236	0.0243
Betasatellites								
CroYVMB	4	1396	218	228	0.0913	121.2	0.0896	0.1053
ToLCBDB	18	1402	407	500	0.0855	112.4	0.0901	0.1262
ToLCJoB	7	1394	302	340	0.0829	110.0	0.0929	0.1184
Alphasatellites								
ChiLCA	9	1399	37	37	0.0081	10.9	0.0101	0.0101
GDarSLA	4	1378	40	40	0.0146	20.0	0.0158	0.0158

In addition, the selection pressure on these virus-encoded ORFs was analysed using various methods available at Data monkey server and the neutrality tests available in DnaSP v. 5.10. Our results emphasize a large proportion of the sites were found to have evolved under purifying/negative selection (Table 4.8). As ChiLCD-associated begomoviruses possess higher genetic variability, we further checked the presence of any subpopulations within them using a cluster-based method (Figure 4.13).

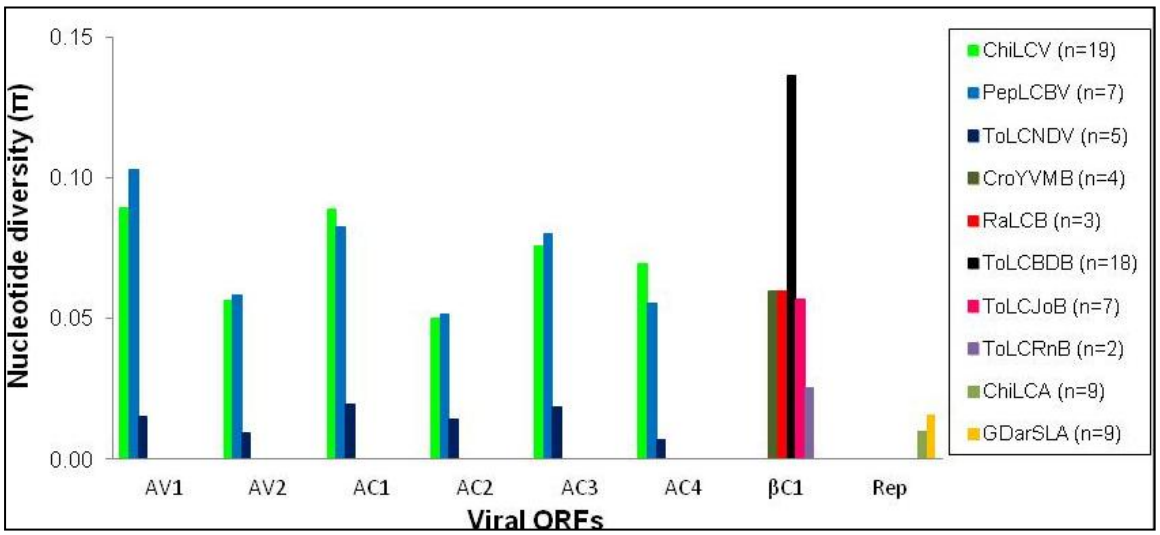


Figure 4.12. Nucleotide diversity of the ORFs encoded by ChiLCD-associated begomoviruses, betasatellites and alphasatellites. The number of sequences of each species or groups used here is also provided in the figure.

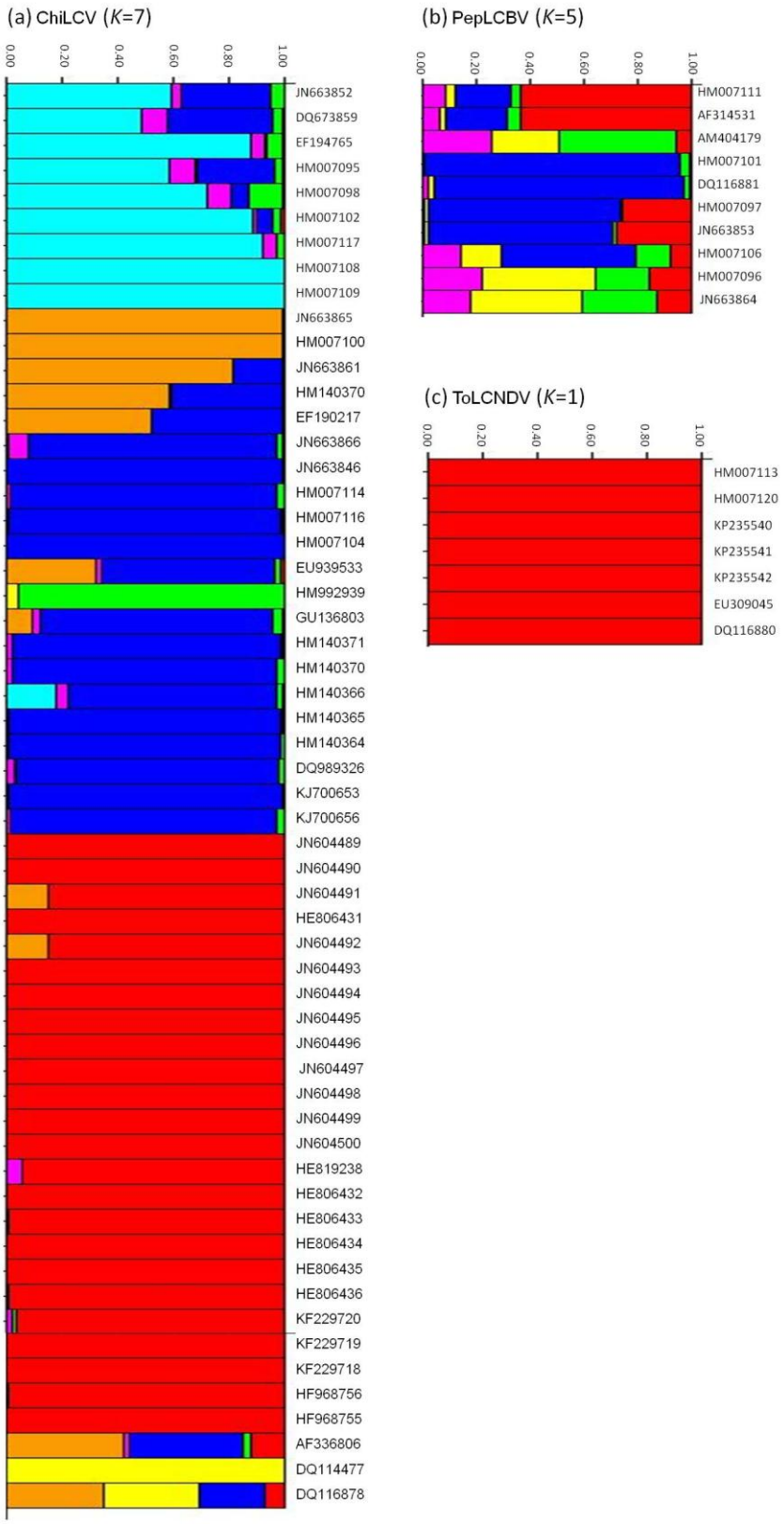


Figure 4.13. Cluster-based sub-population analysis among the begomoviruses associated with ChiLCD in India. Each column numbers in the figure labeled with the accession represents an individual isolates. The *K*-value refers to the number of sub-populations detected in that particular begomovirus species.

Table 4.8. Results of neutrality tests and methods to detect positive or negative selected sites in the ORF(s) encoded by ChiLCD associated begomoviruses and betasatellites.

Viral Components	ORFs	Selection analysis[$]			Neutrality tests		
		SLAC (N/P)	FEL (N/P)	REL (N/P)	Tajima's D	Fu and Li's D*	Fu and Li's F*
Begomoviruses							
ChiLCV (n=19)	AV1	38/0	97/1	131/0	0.8084	0.6171	0.7859
	AV2	7/0	17/0	28/0	0.3749	0.1845	0.2806
	AC1	62/1	102/4	76/5	-0.0062	-0.1930	-0.1597
	AC2	8/0	14/0	5/3	-0.2949	-0.1040	-0.1872
	AC3	7/1	12/2	4/7	0.0630	0.1510	0.1453
	AC4	3/0	9/2	19/0	-0.0161	-0.8468	-0.6987
PepLCBV (n=7)	AV1	37/0	108/1	133/7	-0.0803	0.3345	0.3346
	AV2	4/0	16/0	23/0	-0.6830	-0.5482	-0.6404
	AC1	39/0	85/1	57/7	-0.4608	-0.1153	-0.2141
	AC2	0/0	8/1	16/4	-0.3430	-0.4500	-0.4713
	AC3	8/0	14/1	9/0	-0.4757	-0.3735	-0.4394
	AC4	1/0	4/2	4/2	-0.0955	0.0664	0.0332
ToLCNDV (n=5)	AV1	0/0	0/0	0/0	-1.2452	-1.2452[@]	-1.3413
	AV2	0/0	1/0	0/0	-1.1743	-1.1743	-1.2298
	AC1	1/0	35/1	6/3	-1.1980	-1.1979	1.2977
	AC2	1/0	3/0	0/0	-1.2185	-1.2185	-1.2983
	AC3	2/0	8/0	8/0	-1.2299	-1.2299	-1.3165
	AC4	0/0	0/0	0/0	-1.0485	-1.0485	-1.0519
Betasatellites							
CroYVMB (n=3)	βC1	3/0	15/1	0/0	-0.2014	-0.0537	-0.0853
RaLCB (n=3)	βC1	0/0	0/0	0/0	-0.6124	-0.6124	-0.4787
ToLCBDB n=18)	βC1	9/2	15/1	4/1	-1.9020[@]	-2.4715[@]	-2.6845[@]
ToLCJoB (n=7)	βC1	5/0	17/2	4/0	-0.7660	-0.7958	-0.8737
Alphasatellites							
ChiLCA (n=9)	αRep	0/0	3/0	4/5	-0.099	-1.233	-1.318
GDarSLA (n=3)	αRep	0/0	18/0	0/6	-0.865	-0.865	-0.915

N/P, sites evolved through negative (N) or positive (P) selection;

@ $p<0.05$

[$]SLAC, Single-Likelihood Ancestor Counting; FEL, Fixed-Effects Likelihood and REL, Random-Effects Likelihood

4.2.5 Rate of nucleotide substitution of cloned viral genomic components

The nucleotide substitution rate of AV1 ORF (for begomoviruses), βC1 ORF (for betasatellites) and αRep ORF (for alphasatellite) was estimated using the parameters listed in the methods section using BEAST software. The Bayesian Skyline Plot (BSP) model was used and the datasets were partitioned into three codon positions. The mean nucleotide substitution rate was estimated to be $3.58x10^{-3}$ for AV1 ORF of ChiLCV, $2.57x10^{-4}$ for βC1 ORF of ChiLCB, $5.22x10^{-4}$ for βC1 ORF of ToLCBDB and $2.25x10^{-3}$ for αRep ORF of ChiLCA (Table 4.9). It is possible that the codon degeneracy could play a vital role in the generation of selection pressure leading to genetic variation in the virus evolution. Hence, the mutation rate of all possible nucleotide positions in the codon was estimated using the BSP method (Table 4.9). Further, the mutation rate in the codon position 1 of AV1 ORFs was found to be higher than the mutation rates of the other positions. However, in the βC1 and αRep ORFs, the highest mutation rate was noticed to be in the wobble position of the codon (Table 4.9).

Table 4.9. Rate of nucleotide substitution and mutation rate of ChiLCD-associated viral components-encoded ORFs.

ORFs	Genome names	No. of sequences used	Chain length (in millions)	Mean substitution rate (nucleotide/ site/year)	CP1. mu	CP2. mu	CP3. mu
AV1	ChiLCV	52	20	$3.58X10^{-3}$	1.690	0.366	0.944
	CLCuMuV*	19	40	$4.24X10^{-4}$	1.640	0.449	0.909
	TYLCV[#]	54	20	$4.63X10^{-4}$	-	-	-
	EACMV[@]	71	15	$1.37X10^{-3}$	-	-	-
βC1	ChiLCB	25	40	$2.57X10^{-4}$	0.800	0.622	1.578
	ToLCBDB	31	40	$5.22X10^{-4}$	1.067	0.679	1.256
	CLCuMuB*	39	40	$3.51X10^{-3}$	0.850	0.730	1.430
αRep	ChiLCA	10	40	$2.25X10^{-3}$	0.941	0.860	1.199
	GDarSLA*	63	40	$2.13X10^{-3}$	1.400	0.765	0.831

CP refers to codon position.

Estimated by *Nawaz-ul-Rehman *et al.*, 2012, [#]Duffy and Holmes, 2008, [@]Duffy and Holmes, 2009.

4.3 Discussion

ChiLCD is one of the most severe diseases of chillies in India and cause severe yield reduction in both green and red chillies (Maruthi *et al*., 2007; Chattopadhyay *et al*., 2008). Although, this disease has been known since ages in this country, the damage this diease causes have increased in the recent times; therefore, has become widespread in chilli growing areas in the Indian subcontinent and neighboring countries. Initially, the causal viruses were demarcated based on the type of symptoms (such as enation) induced on host (Mishra *et al*., 1963; Dhanraj and Seth, 1968). Later on, few attempts were also made to identify the causal agent either by partial (Khan *et al*., 2006; Senanayake *et al*., 2007) or complete viral genome sequencing (Kumar *et al*., 2011). Nonetheless, no systematic study was conducted to understand the molecular aetiology of this disease till 2008 (Chattopadhyay *et al*., 2008). The present study provides the first comprehensive survey of begomoviruses associated with the most devastating disease of chilli, ChiLCD in India. The present effort provides the full-length sequences for 84 viral genomic components (35 DNA-A, 6 DNA-B, 10 DNA-1 and 33 betasatellites). In addition to already existing begomoviruses, the presence of unreported monopartite viruses is evident in this study. The identification of novel and existence of previously uncharacterized begomoviruses clearly indicate that the complexity level of diverse pathogens causing ChiLCD is higher than we presume.

The present study identifies ChiLCV as the most prevalent begomovirus species responsible for ChiLCD across India, followed by ToLCNDV and PepLCBV. Nevertheless, the virus species and isolates of ChiLCVs reported here show a greater distinct phylogeographic distribution. Based on prevalence of viral genome, this analysis separates the country into three regions. In the eastern region, states like West Bengal, Bihar and eastern part of Uttar Pradesh, isolates belonging to PepLCBV-BD and ChiLCV were identified with a notable exception of ChiLCV-PK occurring at both Jorehat and Varanasi locations (Figure 4.14). Additionally, mixed infections of PepLCBV and ChiLCV were also found in the eastern Uttar Pradesh (Gazhipur) and Bihar (Chhapra) region. However, ChiLCV was the most abundant species infecting chilli crops in the northern and western region including the states of Haryana, Delhi, Rajasthan, Himachal Pradesh, Gujarat, Maharastra and Uttar Padesh (western part). These data together indicate that ChiLCD in northern and western region of the country is associated with

chilli-infecting begomoviruses identified from Pakistan whereas viruses reported from Bangladesh were predominant in the eastern part of India. Distribution of ChiLCVs is consistent with the spread into India from either of these two neighboring countries. It is relevant to mention here that states adjoining to border region of these countries follow almost similar cropping pattern with that of India which might be responsible for the immigration of the begomoviruses from either Pakistan or Bangladesh through exclusive movement of the insect vector. Additionally, other crop plants as well as weeds grown in abundance also favors the rapid multiplication of vector throughout the year and thereby, enhance movement of these begomoviruses and their satellite components across the country.

Interestingly, the analysis of the isolates from the southern India indicates association of diverse begomoviruses such as PaLCuV-Cro, ToLCV-Ban, ChiLCVV, ToLCNDV, ChiLCV-Gun and PepLCBV-IN (Figure 4.14). Apart from these virus species, the newly identified, ChiLCSmV, also isolated from Salem region in Southern India. These results emphasize that diverse begomoviruses alongwith satellite molecules were noticed to be associated with ChiLCD in the Southern India. A ToLCGV isolate was demonstrated to infect chilli plants under the laboratory condition (Chakraborty *et al*., 2003), it was not surprising to notice the association of several tomato-infecting begomoviruses with ChiLCD in India. Therefore, it can be hypothesized that the breadth of the geographical distribution of any particular begomovirus depends on how efficiently whiteflies transmit that specific virus.

It is well documented that the recombination is a major factor driving the emergence of begomoviruses infecting various crops and weeds in India (Kumari *et al*., 2010; Kumar *et al*., 2011; Singh *et al*., 2012). Chilli-associated begomoviruses contribute to 23 out of 44 recombination breakpoints noticed in ChiLCD-associated begomoviruses and the remaining recombination events (19/44) were generated by either other solanaceous or non-solanaceous crop-infecting begomoviruses. It is noteworthy that ChiLCSmV may have emerged through recombination of three different species such as PaLCuV, PepLCBV and ToLCSLV. As described earlier (Hou & Gilbertson, 1996; Zhou *et al*., 1997; Lefeuvre *et al*., 2007; George *et al*., 2015; Lefeuvre & Moriones, 2015), the current findings highlight the contribution of interspecies recombination in the emergence of chilli begomoviruses. It is also possible that these chilli- and tomato-associated

begomoviruses could be descended from a common ancestor, suggesting a complex history of recombination exists among these begomoviruses in this part of the world. The non-recombinant ChiLCVV might have evolved independently from the earlier recombination event(s). The recombination analyses suggest that the emergence of ChiLCD-associated begomoviruses is mainly due to the significant contribution of solanaceous and non-solanaceous-infecting begomoviruses. Further, the recombination among these viruses could possibly provide higher evolutionary potential for expanding its geographical territory and invading alternate hosts.

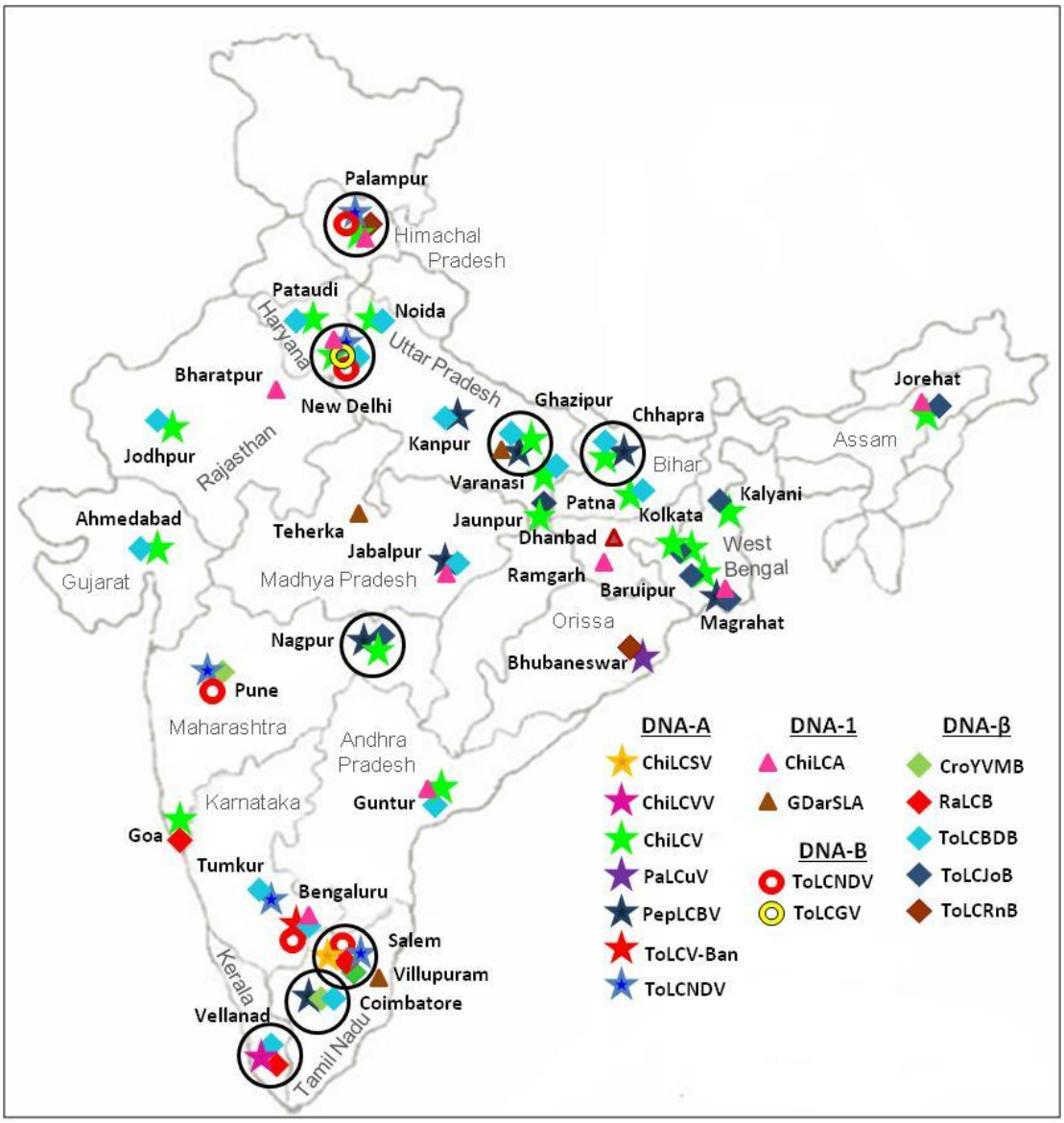


Figure 4.14. Distribution of begomoviruses and associated satellite components with ChiLCD in India. Diverse begomovirus components such as DNA-A, DNA-B, alphasatellites and betasatellites are illustrated in various shapes and colours. The black-coloured circles referred to the mixed infections.

Mixed infection of distinct begomoviruses was noticed in chilli at six locations viz., Chhapra, Gazhipur, New Delhi, Palampur, Salem and Nagpur (Figure 4.14). Noticeably, Nagpur isolate is also a naturally occurring recombinant species originated through inter-specific recombination. Mixed infection often lead to synergistic interaction between participating viruses leading to severe symptom induction coupled with elevated level of DNA replication (Chakraborty *et al*., 2008). Mixed infection, a pre-requisite for recombination, could be responsible and playing significant role in emergence of these ChiLCD-associated begomoviruses. In addition to the presence of monopartite begomoviruses in New Delhi, Salem and Palampur samples, a bipartite begomovirus such as ToLCNDV also identified in these samples (Figure 4.14). The reason for the presence of both monopartite and bipartite begomoviruses together in the same samples and their synergistic interaction, if any, is unclear which needed to be studied in the future. However, it can be speculated that the association of both monopartite and bipartite begomoviruses together could lead to the possible evolution of more severe disease complex. Similar to the helper begomoviruses, the mixed infection was also noticed among the ChiLCD-associated betasatellites identified from the Salem, Coimbatore and Vellanad regions (Figure 4.14). The putative Rep protein binding site of the begomovirus isolates comprises a single core binding site with a potential inverted repeat and the iteron-related domain corresponding to their parental viruses. In spite of sharing maximum of 82.2% with ToLCSLV, the newly identified begomovirus, ChiLCSV, possess similar iteron sequence like that of PepLCBV. This process of origin contribution through the distinct species has been reported to be as “origin donation” which might help the begomoviruses in efficient trans-replication within a suitable host (Chatchawankanphanich & Maxwell, 2002; Saunders *et al*., 2002).

As betasatellites are known to determine pathogenicity for several monopartite begomoviruses, this molecule has been subjected to many scientific investigations (Saunders *et al*., 2004; Cui *et al*., 2005; Yang *et al*., 2011; Li *et al*., 2014). During survey and sample collection, only ToLCBDB and ChiLCB groups were reported with ChiLCD-infected samples in India (Chattopadhyay *et al*., 2008; Kumar *et al*., 2011). The tomato associated betasatellites have been found to be the most prevalent betasatellites associated with ChiLCD in India. Among these, ToLCBDB has been found to co-exist with various begomovirus species in the western part of India. Similarly, ToLCJoB is identified in the eastern and western regions (Figure 13). Though isolates belonging to the begomoviruses

reported from Pakistan (ChiLCV and PepLCBV) with the ChiLCD-infected samples in India was identified, attempts to detect their cognate betasatellites (ChiLCB) in disease samples has not yielded any result. Interestingly, the betasatellites were found along with both bipartite and monopartite begomoviruses from Salem, New Delhi, Pune and Palampur regions. Moreover, a recent report demonstrated the enhanced accumulation of ToLCNDV DNA-A and antagonistic interaction with the ToLCNDV DNA-B component (Jyothsna *et al*., 2013). In this context, the significance of the association of betasatellites with both bipartite and monopartite begomoviruses for the ChiLCD development needs to be studied in future.

The elemental source of genetic variation is mutation, which alongwith recombination determines the genetic makeup of pathogens; finally leading to either compatible or incompatible interaction with the host. Geminiviruses are known to possess high rate of mutation and substitution in consistent with rapidly evolving RNA viruses (Drake, 1993). ChiLCD associated begomoviruses possess high genetic variability than observed in other begomoviruses infecting crops or weeds (Silva *et al*., 2011; Melgarejo *et al*., 2013; Rocha *et al*., 2013). However, the nucleotide diversity noticed in these begomoviruses was not uniformly distributed throughout the genome. Unlike other begomoviruses (Silva *et al*., 2011; Melgarejo *et al*., 2013; Rocha *et al*., 2013), greater diversity in AC1 and AC4 ORFs could provide enhanced fitness for replication or combating of host antiviral silencing activity. High negative values in neutrality tests for βC1 than helper viral ORFs indicates the evolution of large proportion of sites through purifying selection. Alternatively, it can be hypothesized that these betasatellites are undergoing population expansion i.e. they have recently began to expand and therefore, the resultant mutations are unlikely to be lost. These results broadly emphasize that purifying selection is the major evolutionary determinant leading to the emergence of these begomoviruses and betasatellites. Higher mean nucleotide substitution rate estimated for AV1 region in begomoviruses infecting solanaceous crops indicates role of hosts in viral emergence. Nucleotide substitution rate of betasatellites is almost similar to that of helper begomoviruses than betasatellites reported elsewhere. Deamination and lack of repair by host exonucleases are some of the plausible reasons, but no concrete evidences are known till date. Further, the substitution rate of each codon position was observed to be higher in the codon position 1 of AV1, unlike βC1 where codon position 3 (i.e., wobble position) was found to be more amenable for mutation.

ChiLCD poses a serious threat to chilli growing regions and the scenario has worsened in recent past. In the 'New World', chilli is affected by bipartite begomoviruses, *Pepper golden mosaic virus* and *Pepper huasteco yellow vein* virus which are distinct from the Asian ChiLCD-associated begomoviruses (Rentería-Canett *et al*., 2011). Begomovirus association with ChiLCD from other major chilli growing south Asian countries like Pakistan (Tahir *et al*., 2010), Sri Lanka (Senanayake *et al*., 2013), Malaysia (AF414287) and Thailand (AF134484) has been reported. Begomoviruses reported from chilli are also known to naturally infect other hosts like amaranth, bitter gourd, mentha and potato (Mubin *et al*., 2009; Raj *et al*., 2010; George *et al*., 2014; Saeed *et al*., 2014). Greater recombination and mutation rate of these viruses could provide selective advantage in fitness which enables them to adapt efficiently in the newly invaded hosts.

To conclude, the information generated through the comprehensive analysis emphasizes that a large group of monopartite and few bipartite begomoviruses alongwith diverse betasatellites and alphasatellites were associated with ChiLCD in India. Several species of begomoviruses, alphasatellites and betasatellites are identified with the ChiLCD-infected samples. The emergence of these viral components is found to be largely driven by intra- and inter-species recombination and nucleotide substitutions. This is the first systematic study which provides the inclusive details on phylogeographic distribution and the evolutionary dynamics of ChiLCD associated begomoviruses and betasatellites in India. It is clear that ChiLCD-associated begomoviral components are expanding their host range and widening their geographical territory, hence it is necessary to further study the impact of greater genetic variability in these viruses on the epidemiological aspect are the need of hour.

Understanding interaction among cloned chilli-infecting begomoviruses and betasatellites in leaf curl disease

5.1 Introduction

In India, several begomoviruses such as ChiLCV, ChiLCINV, ToLCNDV and ToLCJoV are reported to be associated with ChiLCD (Khan *et al.,* 2006; Senanayake *et al.,* 2007; Kumar *et al.,* 2011). In addition to cause ChiLCD, members of ChiLCVs were also known to be associated with leaf curl disease in other hosts like amaranth, bitter gourd, mentha and potato (Mubin *et al.,* 2009; Raj *et al.,* 2010; George *et al.,* 2014; Saeed *et al.,* 2014). These begomoviruses were found to be associated with various betasatellites and in recent times, this ChiLCD has been reported to spread throughout the major chilli growing regions in India. However, except for the Varanasi strain of ChiLCV (Chattopadhyay *et al*., 2008), these begomoviral genomic components were not proved to be the causal agents of this disease experimentally. Moreover, the infectivity nature and the viral DNA accumulation of these viral genomic components on the natural host, chilli is still lacking. Hence, in this study, the infectivity nature, and the role of satellite molecules in the leaf curl disease development on *N. benthamiana* and *C. annuum* var Sulekha are presented.

5.2 Results

5.2.1 Generation of partial tandem repeat (PTR) constructs

The partial tandem repeat constructs of the various begomoviruses and betasatellite molecules were generated using the strategies as mentioned in the earlier section 3.20 (Table 5.1).

5.2.1.1 NgA and Ngβ

For generating PTR construct of PepLCBV-IN (JN663864; Nagpur isolate), the *Eco*RI (1653)-*Bam*HI (155) digested fragment was ligated into pCAMBIA1300 vector (Table 5.1). This was then followed by mobilization of the monomer of PepLCBV-IN (Nagpur isolate). This PTR construct was designated as NgA. For ToLCJoB (JN663863; Nagpur), the 470 bp region upon digestion using *Kpn*I (1297)-*Bam*HI (411) was cloned in pCAMBIA1300 vector. The full-length *Kpn*I fragment of ToLCJoB (Nagpur) was

mobilized into this, to construct the tandem repeat of ToLCJoB-Nagpur (Figure 5.1). This PTR construct was further referred as Ngβ.

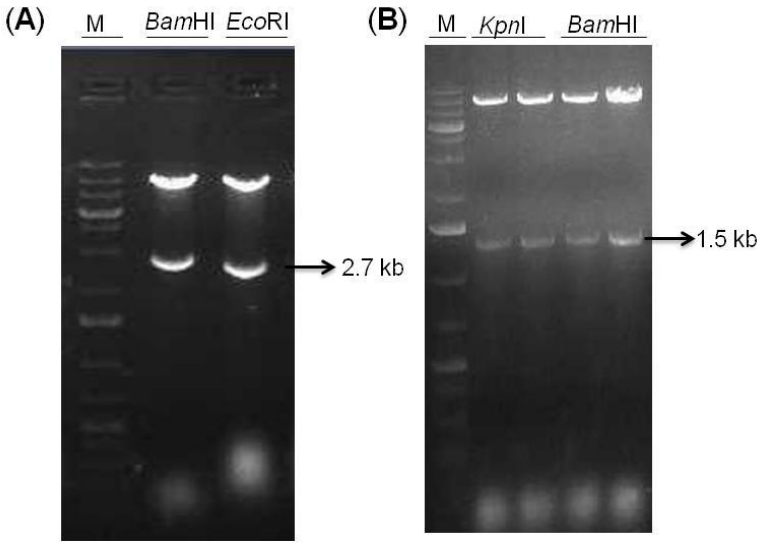


Figure 5.1. Agarose gel electrophoresis showing confirmation of the partial tandem repeat constructs of NgA (A) and Ngβ (B) by digestion. M in lane 1 refers to 1 kb DNA ladder and the restriction enzyme used for digestion is also provided. (A) The DNA used for digestion is clone 1 of NgA (lanes 2 and 3). (B) The DNAs used for digestion are clone 1 of Ngβ (lanes 2 and 4) and clone 2 of Ngβ (lanes 3 and 5).

5.2.1.2 VeA and Veβs

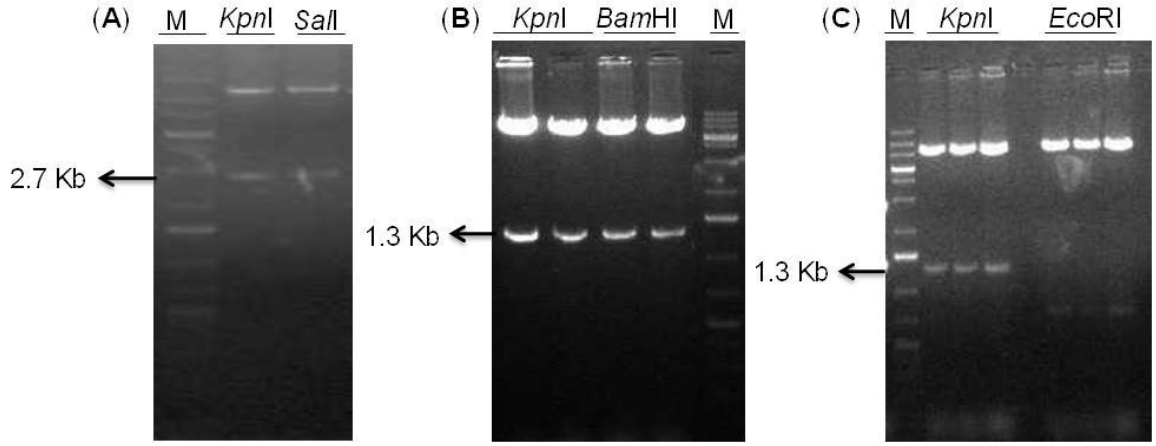


Figure 5.2. Agarose gel electrophoresis showing confirmation of the partial tandem repeat constructs of VeA (A) and Veβ (B) by digestion. M in lane 1 refers to 1 kb DNA ladder and the restriction enzyme used for digestion is also mentioned subsequently. (A) DNA used for digestion is clone 1 of the NgA PTR construct (lanes 2 and 3). (B) DNAs used for digestion are clone 1 of Ngβ (lanes 2 and 4) and clone 2 of Ngβ (lanes 3 and 5) PTR constructs.

Table 5.1. Details of the cloning sites used for generation of partial tandem repeat constructs used in this study.

Genome	Cloned PTR constructs	Partial cloning site	References
DNA-A			
ChA	KR957353; ChiLCV-PK	*Bam*HI (152)-*Pst*I (1449)	Singh *et al.*, 2016
JoA	EF194765; ChiLCV	*Bam*HI (147)-*Sal*I (1150)	Ranjan *et al.*, 2014
KA	KU196750; ToLCNDV	*Kpn*I (1176)-*Xba*I (2250)	Singh *et al.*, 2016
MgA	HM007111; PepLCBV-BD	*Bam*HI (304)-*Kpn*I (1282)	Ranjan *et al.*, 2014
NgA	JN663864; PepLCBV-IN	*Bam*HI (155)-*Eco*RI (1653)	This study
VeA	HM007121; ChiLCVV	*Kpn*I (1620)-*Sal*I (1856)	This study
DNA-B			
KB	KU196751; ToLCGV	*Kpn*I (1543)-*Xba*I (2540)	Singh *et al.*, 2016
DNA-β			
Chβ	KR957354; ToLCBDB	*Kpn*I (1100)-*Bam*H*I* (1149)	Singh *et al.*, 2016
Joβ	EF190216; ToLCJoB	*Eco*RI (511)-*Kpn*I (1310)	Ranjan *et al.*, 2014
Mgβ	HM007112; ToLCJoB	*Eco*RI (511)-*Kpn*I (1308)	Ranjan *et al.*, 2014
Ngβ	JN663863; ToLCJoB	*Bam*HI (411)-*Kpn*I (1297)	This study
Veβ1	JN663876; ToLCBDB	*Bam*HI (1150)-*Kpn*I (1308)	This study
Veβ2	JN663877; RaLCB	*Eco*RI (299)-*Kpn*I (1291)	This study
DNA-1			
DNA-1	KF471051; ChiLCA	*Xba*I (297)-*Pst*I (711)	This study

For generating PTR construct of ChiLCVV (HM007121), a 2.4 kb *SalI* (1856)–*KpnI* (1620) fragment encompassing the CR was ligated into pCAMBIA 2301 vector (Table 5.1). Subsequently, *Kpn*I digested full-length fragment was inserted into it to obtain a PTR construct of ChiLCVV (Figure 5.2). This PTR construct of ChiLCVV was designated as VeA. For ToLCBDB and RaLCB isolated from Vellanad region, 1.2 kb (*Bam*HI and *Kpn*I) and 0.4 kb (*Eco*RI and *Kpn*I) fragments, respectively, were inserted into pCAMBIA 1300 vector (Figure 5.2). This was followed by ligation of the respective

monomers to yield partial tandem repeats of ToLCBDB- and RaLCB-Vellanad (Table 5.1). These PTR constructs was mentioned as Veβ1 and Veβ2, respectively.

5.2.1.3 ChiLCA (DNA-1)

For generating PTR construct of ChiLCA (KF471051; New Delhi isolate), the *Xba*I-*Pst*I digested fragment was cloned into pCAMBIA1300 vector (Table 5.1). This was then followed by the mobilization of the monomer of ChiLCA-New Delhi at *Pst*I site (Figure 5.3). This PTR construct was designated further as DNA-1.

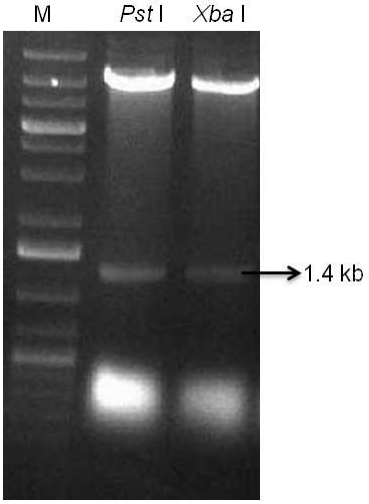


Figure 5.3. Agarose gel electrophoresis showing the confirmation of the partial tandem repeat construct of ChiLCA by digestion. M in lane 1 refers to 1 kb DNA ladder and the restriction enzyme used for digestion is also provided. The DNA used for digestion is clone 1 of ChiLCA PTR construct (lanes 2 and 3).

5.2.2. Agro-inoculation of cloned viral DNAs in *N. benthamiana* plants

The infectivity analysis of the cloned begomoviral molecules such as ChiLCVV (VeA) and its cognate betasatellites (ToLCBDB-Veβ1 and RaLCB-Veβ2), and PepLCBV (NgA) and its cognate betasatellite (ToLCJoB-Ngβ) were studied by infecting *N. benthamiana* plants using these PTR constructs. Five-week old plants were used for the agro-inoculation studies as described earlier. Symptoms were appeared at 8 days post-inoculation (dpi) on either NgA or NgA and Ngβ inoculated *N. benthamiana* plants (Table 5.2). These plants bear symptoms such as leaf rolling, leaf crinkling, leaf curling, and the plants remained stunted (Figure 5.4). However, NgA and Ngβ inoculated *N. benthamiana* plants developed downward leaf curling, a characteristic symptom induced by betasatellites (Figure 5.4).

Table 5.2. Infectivity studies of the cloned viral components on *N. benthamiana* and *C. annuum* var. Sulekha plants.

	Viral constructs	Number of symptomatic plants/ number of inoculated plants	Incubation period (in dpi)[$]	Type of symptoms appeared on the inoculated plants[#]	PCR[@]			
					DNA-A	DNA-β	DNA-B	DNA-1
Nicotiana benthamiana	ChA	2/6	12	MLC	+	-	-	-
	ChA+DNA-1	2/6	13	MLC	+	-	-	+
	ChA+Chβ	6/6	10	DLC	+	+	-	-
	ChA+Chβ+DNA-1	6/6	10	DLC	+	+	-	+
	KA	5/6	11	LC, Vt	+	-	-	-
	KA+DNA-1	5/6	11	LC, Vt	+	-	-	+
	KA+KB	6/6	9	LC, CrL, Vt, YL	+	-	+	-
	KA+KB+DNA-1	6/6	9	LC, CrL, Vt, YL	+	-	+	+
	NgA	10/10	8	LC, LR, TP, St	+	-	-	-
	NgA+DNA-1	10/10	8	LC, LR, TP, St	+	-	-	+
	NgA+Ngβ	10/10	8	DLC, CR, St	+	+	-	-
	NgA+Ngβ+DNA-1	10/10	8	DLC, CrL, TP, St	+	+	-	+
	VeA	10/10	6	LC, St, SL, MY	+	-	-	-
	VeA+Veβ1	10/10	6	DLC, CR, St, SL, MY	+	+	-	-
	VeA+Veβ2	6/6	7	LC, CR, St, SL, MY	+	+	-	-

	Viral constructs	Number of symptomatic plants/ number of inoculated plants	Incubation period (in dpi)[$]	Type of symptoms appeared on the inoculated plants[#]	PCR[@]			
					DNA-A	DNA-β	DNA-B	DNA-1
Capsicum annuum var Sulekha	NgA	10/20	35	MLC	+	-	-	-
	NgA+Ngβ	15/21	25	LC, LR, CrL, LB, Vt, SL, St	+	+	-	-
	VeA	9/20	35	MLC	+	-	-	-
	VeA+Veβ1	14/22	27	LC, LR, CrL, LB, Vt, SL, St	+	+	-	-
	ChA	5/20	55	MLC	+	-	-	-
	ChA+Chβ	11/20	40	LC, CrL, Vt	+	+	-	-
	JoA	7/20	51	MLC	+	-	-	-
	JoA+Joβ	9/19	37	LC, CrL, LR, LB	+	+	-	-
	MgA	5/20	50	MLC	+	-	-	-
	MgA+Mgβ	12/21	38	LC, CrL, LR, LB	+	+	-	-

[$]Incubation period - time taken for the first appearance of symptoms (in dpi).

[#]LC, leaf curling; DLC, downward leaf curling; MLC, mild leaf curling; SL, small leaves; MY, mild yellowing; LB, leaf blistering; LR, leaf rolling; TP, twisting of petiole; St, stunting; LCr, leaf crinkling; Vt, vein thickening.

[@]Primers specific for DNA-A, DNA-B, DNA-β and DNA-1; + and - indicates the presence and absence of viral components, respectively.

Similarly, *N. benthamiana* plants inoculated with VeA developed symptoms such as leaf curling, small leaves and mild yellowing. The symptoms appeared on these plants at 6 dpi (Table 5.2). In addition to other symptoms noticed on VeA inoculated plants, the VeA and Veβ1 inoculated plants developed downward leaf curling (Figure 5.4). However, no downward leaf curling was noticed on the *N. benthamiana* plants inoculated with VeA and Veβ2 (Table 5.2). Further, in comparable to mock-inoculated *N. benthamiana* plants, VeA inoculated plants (either in the presence or absence of Veβ1/Veβ2) remained stunted (Figure 5.4). The presence of these begomoviruses and the betasatellites in the systemic leaves were tested by PCR. All these viral components were readily detectable in the systemic leaves of the infected *N. benthamiana* plants; however the association of betasatellites failed to alter the incubation period (Table 5.2). The accumulation of these viral DNAs by Southern hybridization suggests that the addition of the cognate betasatellites with the begomoviruses did not increase the level of DNA-A component in the inoculated plants at 21 dpi (Figure 5.4).

5.2.3 Infectivity of cloned viral molecules on *Capsicum* plants

In addition to *N. benthamiana* plants, we also studied the infectivity analysis of the cloned begomoviruses and their cognate betasatellites in their natural host, *C. annuum* var. Sulekha. The details of the infectivity analysis including the number of symptomatic plants per total number of plants inoculation, symptom appearance and incubation period are provided in Table 5.2. Apart from the infectious constructs of ChiLCVV-VeA and its cognate betasatellite (ToLCBDB-Veβ1), and PepLCBV-NgA and its cognate betasatellite (ToLCJoB-Ngβ), we have also included the partial tandem repeat constructs of various other begomoviruses and their cognate satellites such as ChiLCV-ChA+ToLCBDB-Chβ, ChiLCV-JoA+ToLCJoB-Joβ and PepLCBV-MgA+ToLCJoB-Mgβ to analyze the infectivity nature of these viral components (Table 5.5). *C. annuum* Var. Sulekha plants infected with begomoviruses alone such as VeA, NgA, ChA, JoA and MgA, developed very mild symptoms (Table 5.2; Figure 5.5). The Southern analysis failed to detect the accumulation of any of these begomoviruses in the systemic leaves of the infected chilli plants, at 90 dpi (Figures 5.6, 5.8). However, presences of all these begomoviruses were readily detected in the systemic leaves of the infected plants by PCR analysis (Table 5.2; Figure 5.6a). This suggests that these begomovirus can replicate and move systemically in the infected chilli plants even in the absence of respective betasatellite. Interestingly, in *C. annuum* var. Sulekha plants co-infected with the begomoviruses and

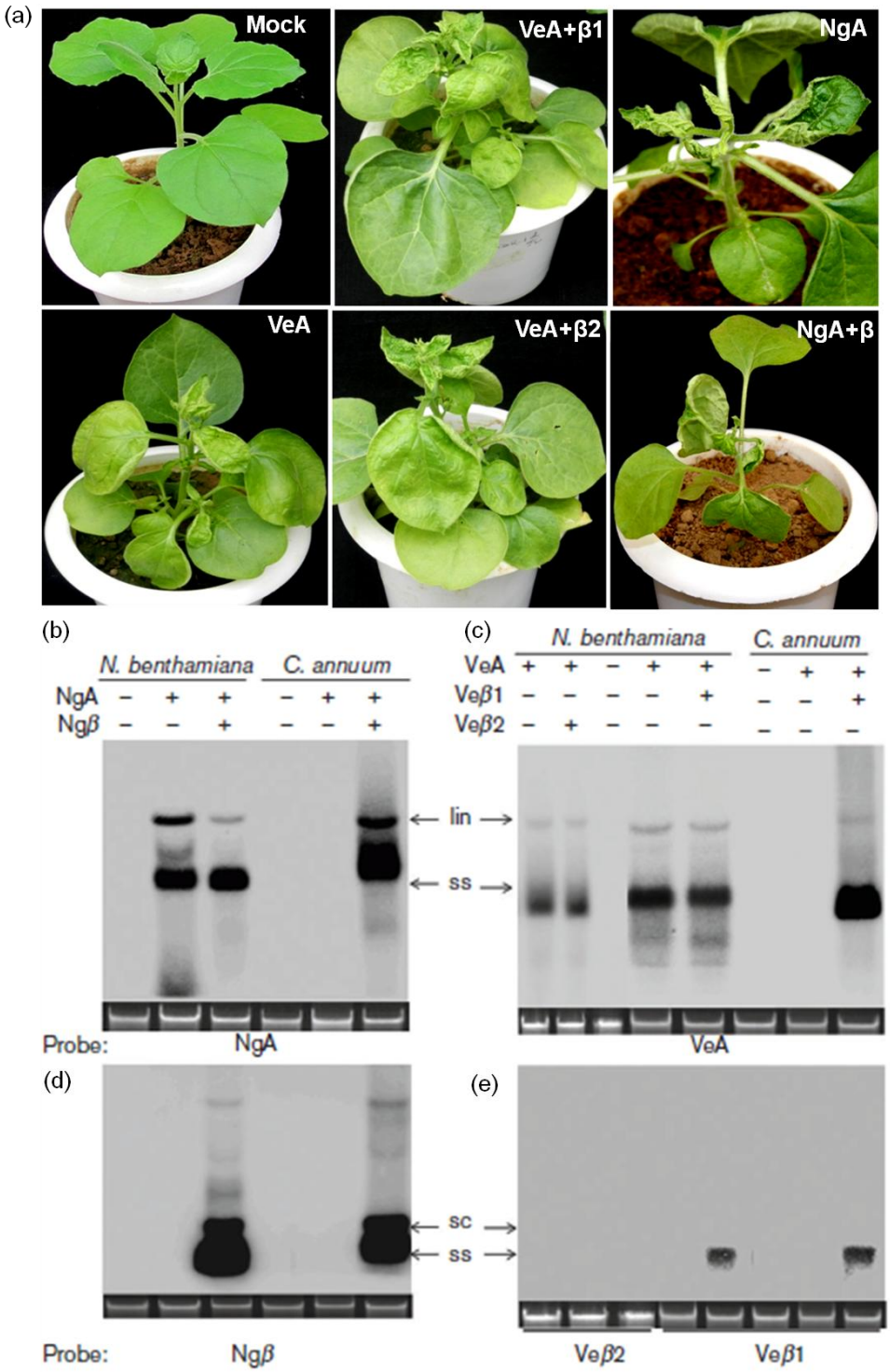


Figure 5.4. Infectivity analysis of the viral components cloned from Nagpur and Vellanad regions. (a) Symptom appearance of *N. benthamiana* plants inoculated with various combinations of viral infectious constructs as mentioned in the figure, at 21 dpi. (b-e) DNA accumulation of the viral genomic components by Southern blotting either at 21 dpi (*N. benthamiana* plants) or at 45 dpi (*C. annuum* var Sulekha plants). The specific probes used are provided at the bottom of each blot. The viral replicative forms are labelled as lin-linear, sc-super coiled and ss-single stranded.

their cognate betasatellites, the symptoms such as leaf crinkling, leaf blistering, leaf curling and vein thickening were noticed (Table 5.2). Further, in comparison to mock and begomoviruses alone-inoculated plants, the co-infected (begomoviruses and betasatellites) chilli plants remained stunted (Figure 5.5).

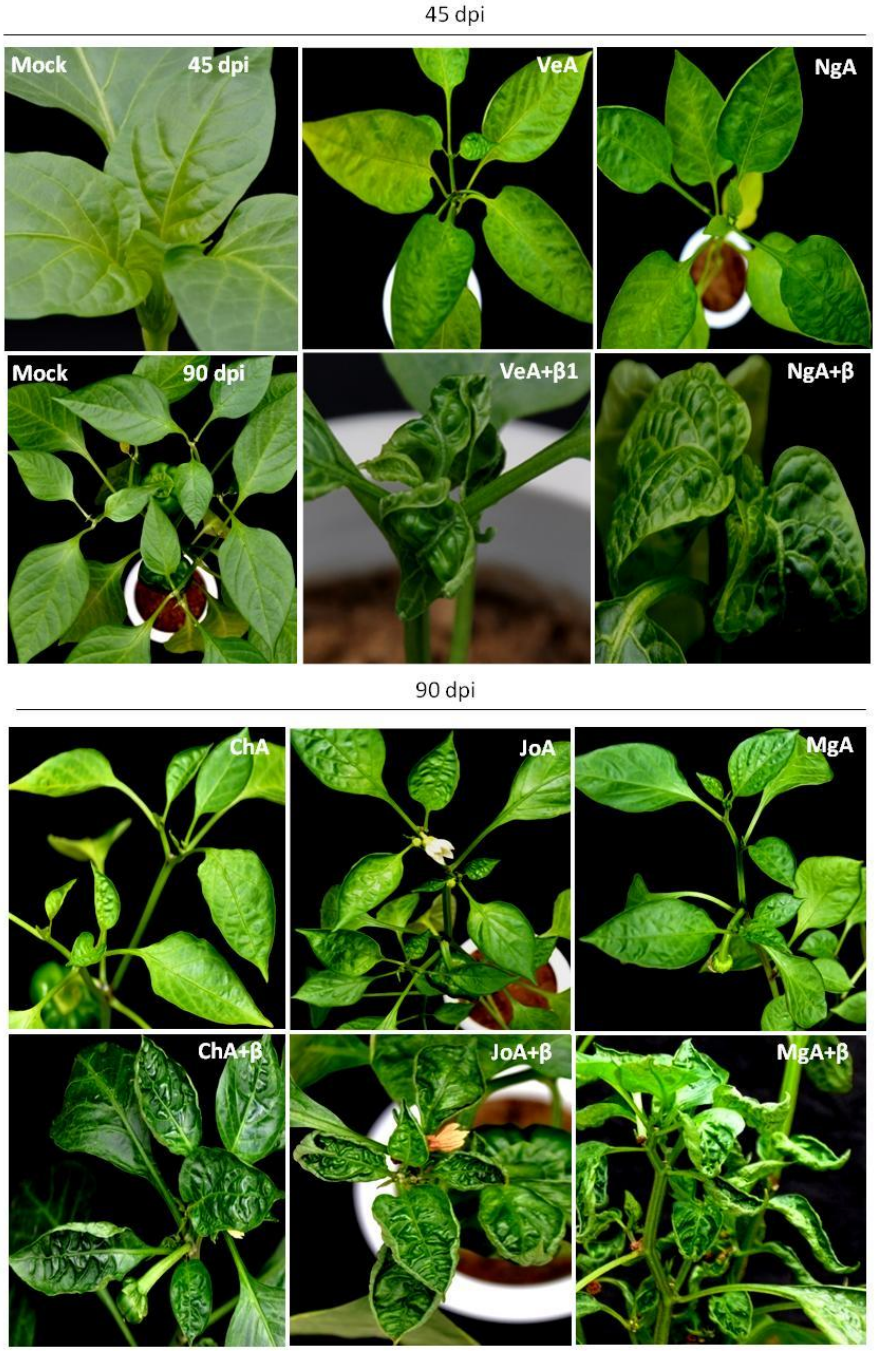


Figure 5.5. Symptom appearance on *C. annuum* var Sulekha plants inoculated with various begomovirus-betasatellite components.

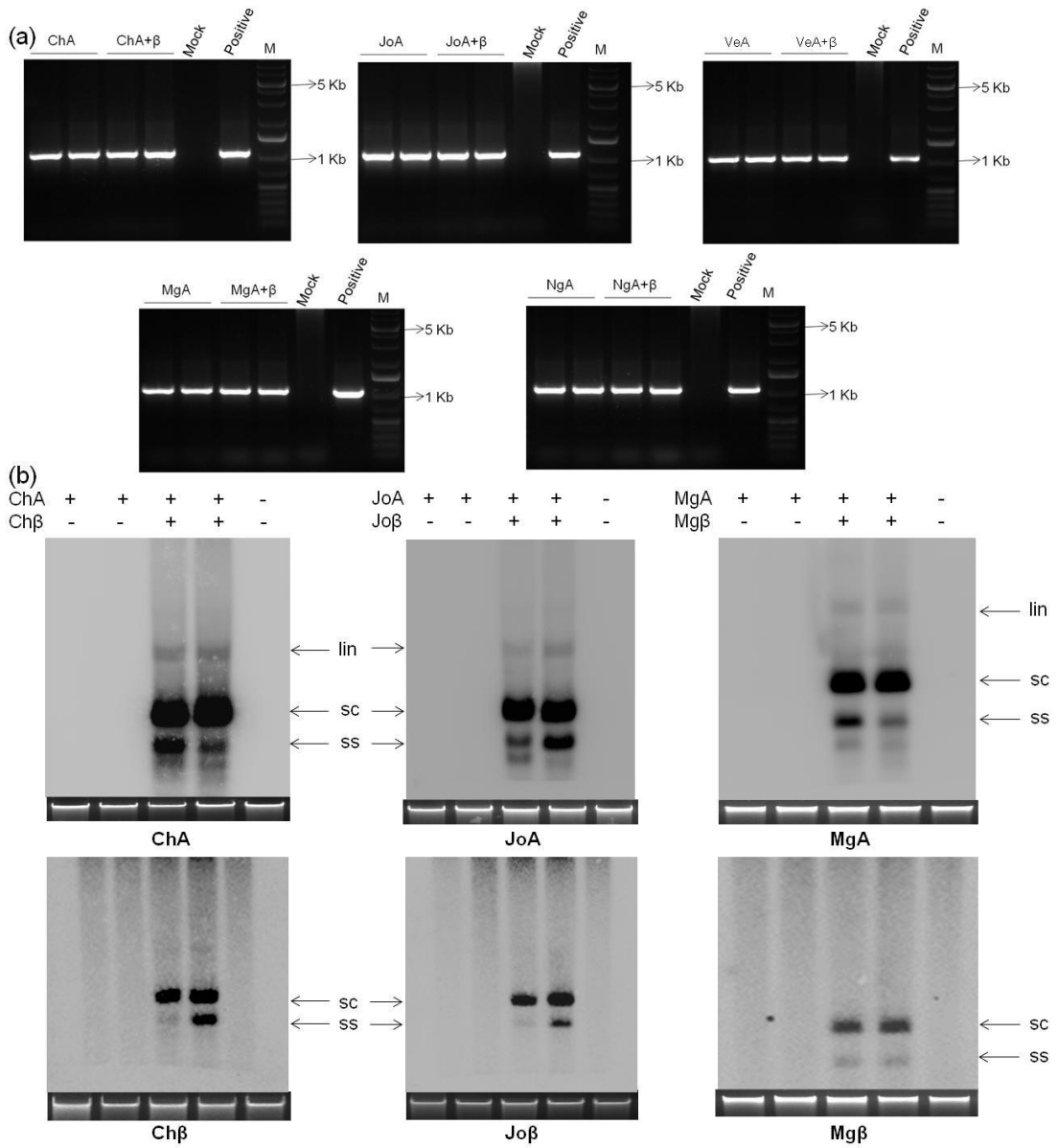


Figure 5.6. Relative level of viral DNA in infected chilli plants. (a) PCR amplification from the DNA extracted from chilli plants infected with begomoviruses alone and its cognate betasatellites, using degenerate primers. Positive indicates DNA isolated from begomovirus-infected plants where as mock denotes plants inoculated with empty vector alone. (b) Accumulation of viral DNAs in the systemic leaves of the *C. annuum* var. Sulekha plants inoculated with various begomovirus-betasatellite components as labelled in the figure, at 90 dpi. The ethidium bromide stained DNA serves as loading control. The viral replicative forms are abbreviated as ln-linear, sc-super coiled and ss-single stranded.

The incubation period in the plants inoculated with both begomoviruses and betasatellites were found to be less than the chilli plants inoculated with begomoviruses alone. Further, the percentage of the number of plants bearing typical begomovirus-like symptoms in the

plants co-inoculated with both the begomoviruses and betasatellites was noticed to be higher than the chilli plants inoculated without the cognate betasatellites (Table 5.2). In the Southern analysis, the higher level of begomovirus DNA accumulation was found to be noticed only in the plants when co-inoculated with begomoviruses and their cognate betasatellites (Figure 5.6). Similarly, the accumulation of betasatellites in these plants was detected in the Southern analysis (Figure 5.6). Altogether, our infectivity analysis on chilli plants suggest that the begomoviruses tested (VeA, NgA, ChA, JoA and MgA) require their cognate betasatellites (Veβ, Ngβ, Chβ, Joβ and Mgβ) for the symptom induction and the overall leaf curl disease development in the natural host, chilli.

5.2.4 Role of alphasatellite in begomovirus pathogenesis on *N. benthamiana* plants

In the previous sections, the effect of betasatellites on disease development has been shown. In addition, to check influence of alphasatellite on the leaf curl disease development on *N. benthamiana*, the PTR construct of DNA-1 (ChiLCA isolate from New Delhi) was generated (Figure 5.6). This ChiLCA-New Delhi PTR construct was co-inoculated with cognate (ChA and KA) and non-cognate (NgA) begomovirus genomic components along with their cognate viral components such as Chβ, KB, and Ngβ, respectively (Table 5.5). *N. benthamiana* plants inoculated with ChA or KA developed mild symptoms and also the presence of DNA-1 did not alter their symptom severity. However, the symptoms were found to enhanced upon co-inoculation with either the betasatellite (Chβ with ChA) or the DNA-B component (KB with KA), respectively (Figure 5.7; Table 5.2).

Similarly, *N. benthamiana* plants inoculated with NgA alone or alongwith Ngβ developed typical begomovirus-like symptoms such as leaf rolling, leaf curling, crinkling of leaves and stunting at 7-8 dpi (Figure 5.7). The presence of DNA-1 with either begomoviruses alone (ChA+DNA-1, NgA+DNA-1 and KA+DNA-1) or begomoviruses along with their cognate genomic components (ChA+Chβ+DNA-1, KA+KB+DNA-1 and NgA+Ngβ+DNA-1) did neither alter the symptom appearance nor the incubation period (Table 5.2). All the viral components were readily detected at 21 dpi by PCR and by Southern analysis (Figure 5.8; Table 5.5). Moreover, in comparison to plants inoculated with begomoviruses (ChA, KA and NgA) alone, the presence of DNA-1 alongwith the begomoviruses failed to influence the accumulation of the DNA level of these helper begomoviruses (Figure 5.8). However, when the alphasatellite was co-inoculated with

begomoviruses and their cognate betasatellites (ChA+Chβ+DNA-1, KA+KB+DNA-1 and NgA+Ngβ+DNA-1), the marginal decrease in the accumulation of betasatellites (Chβ and Ngβ) was noticed (Figure 5.8).

DNA-1 was detected in the systemic leaves alongwith the cognate (ChA, ChA+Chβ, KA and KA+KB) and non-cognate (NgA and NgA+Ngβ) viral genomic components, however the alphasatellite was not found in the systemic leaves of all the inoculated plants. Further, the begomoviruses alone (ChA, KA, NgA) or with the associated viral components (ChA+Chβ, KA+KB and NgA+Ngβ) did not influence the accumulation of the alphasatellite DNA level (Figure 5.8).

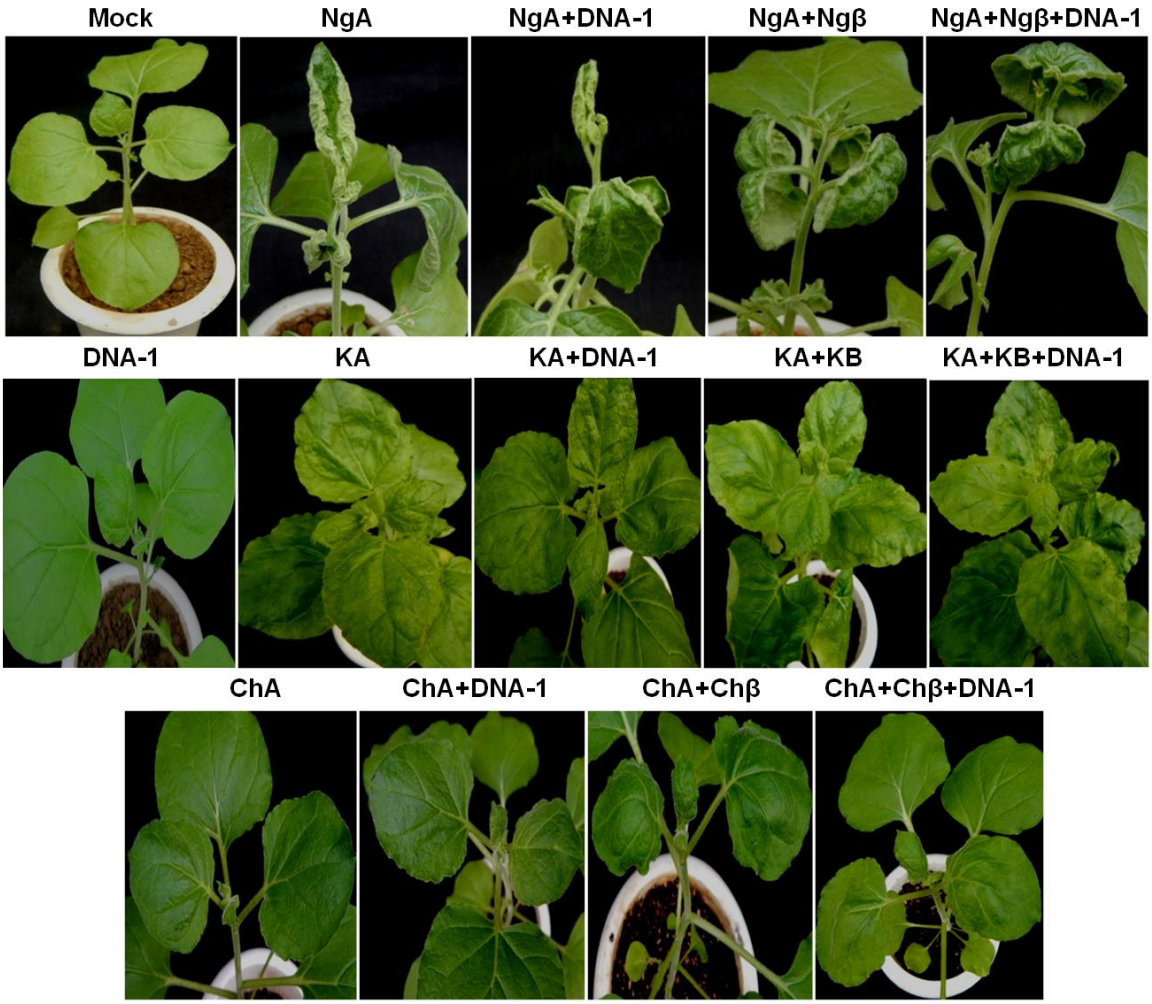


Figure 5.7. Symptom appearance on representative *N. benthamiana* plants inoculated with various combinations containing begomoviruses and betasatellites alongwith alphasatellite components at 21 dpi.

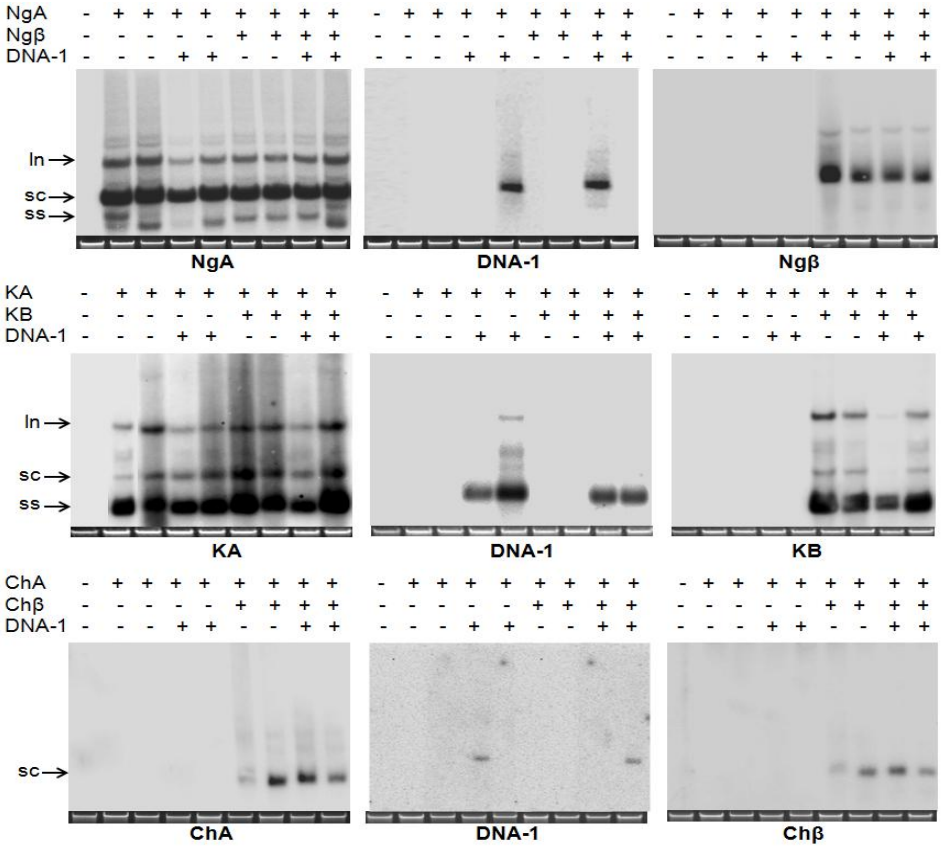


Figure 5.8. Accumulation of viral DNA in infected *N. benthamiana* plants at 21 dpi. Detection of the viral DNAs in the *N. benthamiana* plants inoculated with various begomovirus-betasatellite-alphasatellite components as labeled. The ethidium bromide stained DNA gels serve as loading control. The viral replicative forms are abbreviated as ln-linear, sc-super coiled and ss-single stranded.

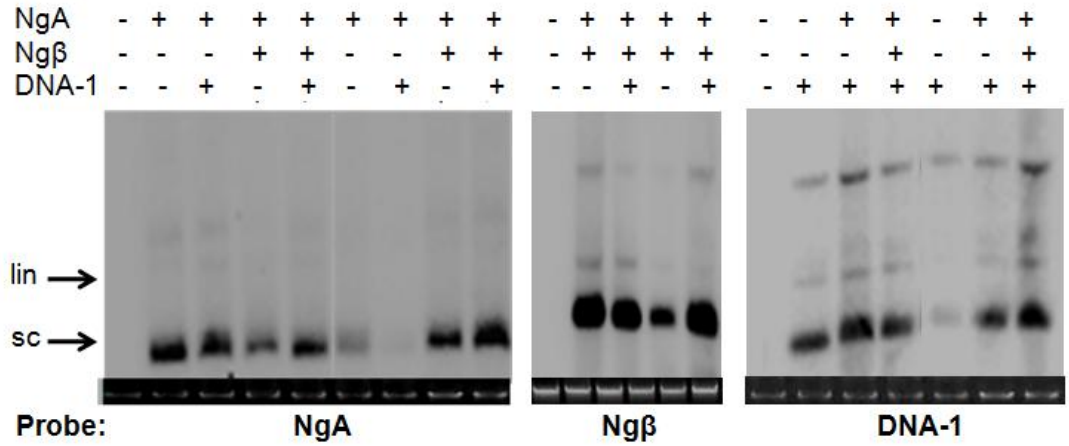


Figure. 5.9. Transient viral DNA replication assay using *N. benthamiana* leaf disks. Level of viral DNA accumulation by Southern blotting in the agro-infected *N. benthamiana* leaf disks at 6 dpi. The viral constructs used for agro-infection are provided above each lane and the probes used are given at the bottom of each blot i.e., NgA, Ngβ and DNA-1.

As the systemic spread of the alphasatellite was found to be sporadic in the inoculated *N. benthamiana* infectivity studies, further attempt was made to study the replicative nature of ChiLCA-New Delhi. This transient replication assay was studied using *N. benthamiana* leaf disks and the NgA and its cognate betasatellite, Ngβ, was used for this study. Total DNA was isolated from the infected *N. benthamiana* leaf disks at 6 dpi. Southern analysis has revealed that DNA-1 alone replicated autonomously without the helper begomovirus or betasatellite (Figure 5.9). Further, this transient assay indicated that the PTR construct of ChiLCA-New Delhi was indeed infectious and also possesses the capability of autonomous replication.

5.3 Discussion

Chilli is a major spice-cum-vegetables used in various food products in India and the cultivation of chillies is mainly hindered by chilli leaf curl disease (ChiLCD). Viruses as the causative agents associated with ChiLCD were reported several decades ago. Later, four distinct begomovirus species such as ChiLCV, ChiLCIV, ToLCJoV and ToLCNDV and two different betasatellites such as ChiLCB and ToLCBDB were found to be associated with ChiLCD in India (Khan *et al.,* 2006; Senanayake *et al.,* 2007; Kumar *et al.,* 2011). The present survey and comprehensive analysis on the diversity of begomoviruses and betasatellites suggest association of various other begomovirus species and betasatellite groups with ChiLCD in India. Begomoviruses reported from chilli are also known to naturally infect other hosts like amaranth, bitter gourd, mentha and potato (Mubin *et al.,* 2009; Raj *et al.,* 2010; George *et al.,* 2014; Saeed *et al.,* 2014). However, detailed studies on role of the reported begomoviral components on leaf curl disease development have not been carried out earlier. Further, the experimental evidence proving that these reported begomoviruses are the causal agents of ChiLCD is lacking. Hence, the present study has investigated the role of the cloned begomoviral genomic components in leaf curl disease development on *N. benthamiana* and *Capsicum* spp. The cloned begomovirus components have been demonstrated to infect *N. benthamiana* as well as *Capsicum* spp. Further, the essentiality of the associated alphasatellites and betasatellites in symptom induction has been elucidated.

The infectivity analysis of NgA and VeA infectious constructs suggest both of these can infect *N. benthamiana* without the help of their cognate betasatellites. It is relevant to

mention here that mixed infection of betasatellites were found in the southern India (Salem, Coimbatore, Vellanad). Both the betasatellites from Vellanad (ToLCBDB and RaLCB) were found to be trans-replicated by their cognate helper begomovirus (ChiLCVeV) in *N. benthamiana*. Similarly, Ngβ (ToLCJoB) was also found to be trans-replicated by its cognate begomovirus (PepLCBV) in *N. benthamiana*. However, RaLCB from Vellanad region could neither influence the symptom severity nor the incubation period. Both NgA and VeA alone could cause systemic infection in *N. benthamiana* plants and the cognate betasatellites could not enhance the accumulation of helper begomoviruses (NgA and VeA) as well. Collectively, these results suggest that these two begomoviruses were obligate monopartite begomoviruses causing ChiLCD in *N. benthamiana*.

In addition, the infectivity nature of ChiLCA with ChiLCD associated begomovirus and betasatellite in *N. benthamiana* was also studied. The results suggest that this alphasatellite did not exert any discernible effect either on the incubation time or on the accumulation of the helper begomovirus genome in infected plants. Similar to other alphasatellites (Kon *et al*., 2009; Idris *et al.*, 2011), this ChiLCA could also partially affect with the accumulation of the betasatellite tested here, ToLCJoB-Nagpur. The collective results from the present study and the previous reports (Briddon *et al.*, 2004; Kon *et al.*, 2009) provide substantial evidence to conclude that ChiLCA possess dispensable role in pathogenesis which furthermore justifies their infrequent association with ChiLCD in India. The sporadic spread of ChiLCA in *N. benthamiana* plants could be the result of its inability to form effective movement complex with the gene products derived from the begomoviruses and/or betasatellites. Altogether, the infectivity studies of ChiLCD-associated begomoviruses, betasatellites and an alphasatellite emphasize that both the alphasatellites and betasatellites were not required for the helper begomoviruses to cause disease in the model host, *N. benthamiana*.

Furthermore, the infectivity nature of these cloned viral genomic components was studied in their natural host, *Capsicum* spp. The mild symptoms were noticed on the chilli plants inoculated with either of the helper begomoviruses such as ChA, JoA, MgA, NgA or VeA. However, the viral replicative forms of ChA, JoA, MgA, NgA or VeA in the systemic leaves could be detected by PCR. These results highlight that these begomovirus molecules (ChA, JoA, MgA, NgA or VeA) can replicate in the natural host. The

infectivity analysis using *N. benthamiana* as well as *Capsicum annuum* plants, further proved that these begomoviruses were indeed the causal agents of ChiLCD. However, in the presence of their cognate betasatellites (ChA+Chβ, NgA+Ngβ, MgA+Mgβ, JoA+Joβ and VeA+Veβ), the inoculated chilli plants displayed severe symptoms till 90 dpi. This is the first report that points out the essentiality of the betasatellite to cause leaf curl disease in chilli and also justifies the ubiquitous presence of betasatellites with the begomoviruses associated with ChiLCD. Additionally, the dependence of ChiLCD-associated monopartite begomoviruses on the betasatellites for symptom development underscores that these begomoviruses require betasatellites for pathogenesis similar to the complexes described earlier (Saunders *et al.,* 2000).

It is noteworthy to mention that the betasatellites have been reported to assist in the systemic movement of monopartite begomoviruses and further, the protein encoded by various betasatellites have been reported to evade host-mediated gene silencing machinery (Shen *et al.,* 2012; Yang *et al.,* 2012; Li *et al.,* 2014). Further, an isolate of ToLCNDV has been reported to trans-replicate a cotton-associated betasatellite and this betasatellite has been shown to complement the defective movement function of this begomovirus (Saeed *et al.,* 2007; Sivalingam and Varma, 2012). In this context, these functions of betasatellites probably have been utilized by ChiLCD-associated begomoviruses to successfully spread and cause leaf curl disease in the natural host, chilli. However, the dispensable nature of the betasatellites in *N. benthamiana* plants could be explained due to its immuno-compromised nature since it lacks functional RNA-dependent RNA polymerase 1 which is the primary antiviral responder in plants (Yang *et al.*, 2004).

In summary, the infectivity of five different monopartite begomovirus-betasatellite complexes was demonstrated on *N. benthamiana* and *C. annuum*. In *N. benthamiana* plants, both the alphasatellites and betasatellites were found to be dispensable for symptom induction. However, this work highlights that the betasatellites are necessary for causing leaf curl disease in the natural host, chilli. Altogether, the present study provides substantial evidence to prove that the monopartite begomoviruses and betasatellites associated with ChiLCD in India are the causal agents of this disease.

Understating the role of pre-coat protein in host-virus interaction

6.1 Introduction

Plants use RNA silencing machinery as a tool of host defense to act against the invading viruses (Law and Jacobsen, 2010; Pumplin and Voinnet, 2013; Du *et al*., 2015). There are two major types of RNAi silencing mechanisms known, transcriptional (TGS) and post-transcriptional (PTGS) gene silencing. In PTGS, siRNAs are generated from the dsRNA precursors and the AGO-bound siRNAs forms RISC leading to target RNA degradation or translation repression. It is widely documented that viruses can be inducers of RNA silencing (Law and Jacobsen, 2010; Pumplin and Voinnet, 2013). However, to overcome this antiviral immune response of host, viruses are evolved to encode proteins termed as viral suppressors of gene silencing (Pumplin and Voinnet, 2013). RNA viruses are known to be the target of PTGS, whereas DNA viruses are targeted at both transcriptional level and post-transcriptional level by the host RNAi surveillance system (Hanley-Bowdoin *et al*., 2013).

Several members of geminiviruses are known to encode proteins to evade this surveillance-cum-defense mechanism (Hanley-Bowdoin *et al*., 2013). Pre-coat protein of begomoviruses has been demonstrated to suppress PTGS by blocking the biological function of SGS3 (Glick *et al.,* 2008), while the βC1 protein represses RDR6 expression through recruitment of calmodulin-like protein (Li *et al*., 2014). With ssDNA as genetic material, geminiviruses are known to form mini-chromosome in the infected plant cells leading to the methylation of viral DNA and its associated histones (Hanley-Bowdoin *et al*., 2013). Further, hyper-susceptibility of geminivirus infections have been elucidated in methylation compromised (both at viral chromatin and histone) mutants (Raja *et al*., 2008; Hanley-Bowdoin *et al*., 2013). These studies provide evidences to the hypothesis that plants employ cytosine and histone methylation as their defense strategy to counter spread of virus infections. It is also widely reported that histone methylation and cytosine methylation are inter-dependent through a self-reinforcing loop (Law and Jacobsen, 2010; Du *et al*., 2015).

Several geminivirus-encoded proteins such as Rep, TrAP, βC1 and pre-coat proteins have been demonstrated to suppress TGS at various levels (Wang *et al*., 2014). For TGS, it is

necessary to generate constant supply of methyl donor i.e., S-adenosyl methionine (SAM). Hence, it is clear that methyl cycle comprising of several key enzymes are very essential for DNA and histone methylation. However, C2 protein (TrAP) has been shown to interfere with the methyl cycle by blocking the proteasomal degradation of S-ADENOSYL METHIONINE DECARBOXYLASE (SAMDC1) and inhibiting the enzymatic activity of ADENOSINE KINASE (ADK) enzyme (Wang *et al*., 2005; Zhang *et al*., 2007). Another begomoviral protein (βC1) has been demonstrated to interact with S-ADENOSYL HOMOCYSTEINE HYDROLASE (SAHH) enzyme and affects its biological function (Yang *et al*., 2011). In addition, C2 protein binds to histone methyl transferase, KRYPTONITE, and inhibiting methyl transferase activity of the target protein (Castillo-Gonzalez *et al*., 2015). Rep protein also functions as TGS suppressor by down-regulating the expression of DNA methylation enzymes such as MET1 and CMT3 (Rodriguez-Negrete *et al*., 2013). Recently, TYLCCNV-V2 protein has been reported to be a suppressor of TGS (Wang *et al.*, 2014), but the exact mechanism through which it acts is not yet understood. In the present study, interaction of V2 with a component of TGS is demonstrated and the biological significance of this association has been discussed.

6.2 Results

6.2.1 Level of NbVIM transcripts upon virus infection

Several host factors have been reported to be differentially regulated upon virus infection, as detected through screening of suppression subtractive hybridization (SSH) library (Kushwaha *et al*., 2015). Further, V2 protein of TYLCCNV has been reported recently to suppress TGS (Wang *et al*., 2014). One of the component of TGS machinery, VARIANT IN METHYLATION (VIM), was found to be differently expressed as found in SSH library screening. Hence, to assess the expression of VIM transcripts upon virus infection, the level of NbVIM was analyzed using qRT-PCR. The level of *NbVIM* transcript upon virus inoculated plants was compared with the mock-inoculated plants. The transcript level was normalized using an internal control (ACTIN). The quantitative RT-PCR results indicate that the accumulation of NbVIM transcript was found to be reduced significantly in the early stage (10 DPI) than the later stage (20 DPI) of ChiLCV infection (Figure 6.1).

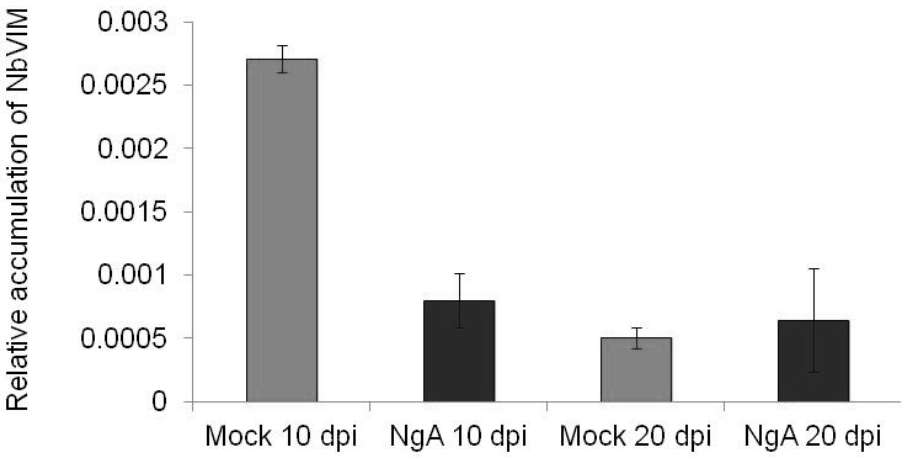


Figure 6.1. Relative accumulation of virus transcripts in the inoculated *N. benthamiana* plan ts. Mock indicates only vector inoculated plants whereas NgA refers to the plants inoculated with infectious construct of ChiLCV-NgA.

6.2.2 Isolation and sequence analysis of the VARIANT IN METHYLATION gene from *N. benthamiana*

To assess the role of VIM on the pathogenesis of ChiLCV, VIM was cloned from the cDNAs of *N. benthamiana*. As the sequence of NbVIM is not available in the public database, GenBank, *A. thaliana* sequence of the VIM gene was used as an input sequence in the Sol genomics network (https://solgenomics.net/) to identity the putative VIM sequence from *N. benthamiana* plants.

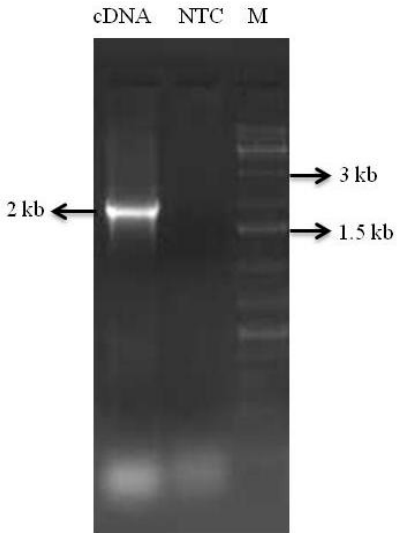


Figure 6.2. PCR amplification of *NbVIM*. Agarose gel electrophoresis indicating the PCR amplified VIM gene from *N. benthamiana* cDNAs. M refers to 1kb DNA ladder and NTC refers to non-template control.

The primers to clone NbVIM were designed based on the putative NbVIM sequence available in the Sol genomics network. *Bam*HI and *Sal*I restriction sites were incorporated into the forward primer and reverse primer, respectively. The complete NbVIM ORF was amplified by PCR using *Phusion* DNA polymerase and the PCR product was eluted from the gel and then cloned into pJET1.2 vector (Figure 6.2). One of the corresponding clones was sequenced completely at the DNA sequencing facility (University of Delhi-South campus, New Delhi). The complete length of NbVIM was found to be 2049 nucleotides and the size of the protein product was predicted to be of 682 amino acids.

6.2.2.1 Nucleotide sequence of the cloned NbVIM

Length: 2049 nucleotides

ATGGCGCACGATAGCGACCTCCCATGTGACGGCGATGGCGTTTGTATGGTGTGCAAAACGAAGCCATCGGAGAAT
GAATCCCTCCTCTGTAAGACTTGTGCCACTCCCTGGCACGTCACCTGCCTCTCTGCTCCTCCCAATTCCATGGCCG
ACGCCGTCAACTGGGAGTGTCCTGATTGCTCCTTGCCTCCCGATGCCCCCCCGGTAGCCGTTAAACAGAATGCTGC
TGCGGCTTCCTCGTCCTCTGGAAGCGATTTGATCGCCGCCATTCGGGTCATAGAATCCGACTCGTCGCTCACCGAA
CAAGAGAAGGCCAAGCAACGACAAGAGCTCTTGAGCGGAAGCGCTAAAGCCAGTAATCTGGAGGGTGTCAACAC
AGACAACGATGTTCTCAACATTCTCGATGGAAGTTTGAACTGCTCTATGTGCATGCAACTGCCGGAGAGACCTGTC
ACGACGCCTTGTGGTCATAACTTTTGCTTGAAGTGTTTCCAAAAGTGGATTGGGCAGCGGAAGAGCACATGTGCTA
ATTGTCGCACCCCTATTCCACCGAAAATGGCTAGCCAGCCTCGAATTAACTCTGCTCTGGTCATTGCCATAAGAAT
GGCAAGAACACTACGATCTGGTGCTTCTGGGGGGACTTCAAAAGTTGCCCATTTCATTCACAACCAAAATAGACCT
GACAAAGCTTTCACCACTGAGCGTGCAAAAAAGTCTGGAAAGGCTAATGCTTGCAGTGGGAAGATTTTTGTCACAA
TTCCTGGAGACCATTTTGGACCAATTTCAGCAGAGAATGATCCAGAGCGCAACTTGGGCGTCTTAGTTGGTGAGAC
TTGGGAGGATCGTTTTGAGTGTCGGCAATGGGGGGCTCATTTCCCTCATGTGGCTGGAATTGCTGGCCAGTCAGATT
ATGGGGCCCAGTCTGTTGCACTTTCTGGTGGATATGAAGATGATGAGGACCATGGTGAATGGTTTCTTTATACTGGA
AGTGGTGGAAGAGACTTGAGTGGTAACAAGCGAACAAACAAGTTACAATCATTTGACCAGAAATTTGATAAGCTT
AATGAAGCTTTACGTGTCAGCTGTCTCAGGGGTTATCCTGTCCGAGTTGTGAGGTCCCATAAGGAGAAGCGTTCTT
CTTATGCACCAGAGAAGGGAGTACGATATGACGGGGTTTACAGGATTGAAAAATGCTGGCGTAAAGCTGGAAAAC
AAGGATTCAAAGTGTGCCGATATCTTTTTATGCGCTGTGATAATGAACCTGCCCCATGGACCAGTGATGAACAGGG
GGACCGTCCTAGGGCCTTGCCTTCCATAAAAGAGCTAAATGGTGCTACAGATATAACTGAAAGAAAGGGATCTCC
TGCTTGGGATTATGATGAGGAGAAGAGTTGTTGGATGTGGAAAAAAGCTCCTCCTAATAGTCGAAAACCTATTCAA
TGTCAAAAAGATGATGATGGAACAAAAGTGAGGAAAGTTAGATCAAAGAAAAATACAATATCTGTGAAGCAAAGG
CTTTTGAAAGAGTTTAGCTGTCTTATTTGCCGGGAAGTGATGAATCTCCCACTTACAACTCCATGTGCTCATAACTT
CTGCAAGGCATGTTTGGAAAGTGCCTTTGCAGGACAGTCTTTCACAAGACAGAGGGGTGAAGGCAGGCGGACGTT
ACGAGTGCAGAAAAATATAATGAAATGCCCGTCATGCCCGACTGATATATCTGATTTCCTCCAGAACCCTCAGATT

AATAGAGAGTTGATGAATGTAATAGAGTCATTGCAGCGCCAAATTGAGGAAGAAAATGCAGGATTAAATGAAGAT
GTTGATGGTGGCTCAGATGGAAAAGATGAGGCTCTAACTGAAGAAGCTGCTGAAACTGATAAACGTGCGGAGGAT
GATGAAGAAGGTGCGAAGCCTATCACTGAGGCTGAAACCAAGCAGACAAACAAGAGGAAGAGGTCCAGTGATAT
TGACAATGTATCACCTAAAAAATCCCTTGACAAGCCTGATGTGGCGCATAATCTTGGGGAAGTTGATCCTGCCCAT
ACCTGA

6.2.2.2 Predicted protein product of NbVIM

Length: 682 amino acids

MAHDSDLPCDGDGVCMVCKTKPSENESLLCKTCATPWHVTCLSAPPNSMADAVNWECPDCSLPPDAPPVAVKQNAA
AASSSSGSDLIAAIRVIESDSSLTEQEKAKQRQELLSGSAKASNLEGVNTDNDVLNILDGSLNCSMCMQLPERPVTTPCGH
NFCLKCFQKWIGQRKSTCANCRTPIPPKMASQPRINSALVIAIRMARTLRSGASGGTSKVAHFIHNQNRPDKAFTTERAK
KSGKANACSGKIFVTIPGDHFGPISAENDPERNLGVLVGETWEDRFECRQWGAHFPHVAGIAGQSDYGAQSVALSGGYE
DDEDHGEWFLYTGSGGRDLSGNKRTNKLQSFDQKFDKLNEALRVSCLRGYPVRVVRSHKEKRSSYAPEKGVRYDGVYR
IEKCWRKAGKQGFKVCRYLFMRCDNEPAPWTSDEQGDRPRALPSIKELNGATDITERKGSPAWDYDEEKSCWMWKK
APPNSRKPIQCQKDDDGTKVRKVRSKKNTISVKQRLLKEFSCLICREVMNLPLTTPCAHNFCKACLESAFAGQSFTRQRG
EGRRTLRVQKNIMKCPSCPTDISDFLQNPQINRELMNVIESLQRQIEEENAGLNEDVDGGSDGKDEALTEEAAETDKRA
EDDEEGAKPITEAETKQTNKRKRSSDIDNVSPKKSLDKPDVAHNLGEVDPAHT

The complete amino acid sequences of various VIM proteins predicted from other plant species were retrieved from the GenBank database. These sequences were aligned along with the isolated NbVIM sequence using the MUSCLE algorithm. The phylogenetic dendrogram analysis suggests that NbVIM possess high similarity with the VIM proteins of other members belonging to *Solanaceae* family and hence, it was positioned along with VIM proteins predicted from other solanaeceous crops (Figure 6.3).

Using this predicted protein sequence, various domains present in NbVIM were predicted using conserved domain database. This analysis suggests the presence of two RING (REALLY INTERESTING GENE) domains (at N-terminal and C-terminal regions) and one domain each of PHD (PLANT HOMEODOMAIN) and SRA (SET- AND RING-ASSOCIATED) in the cloned NbVIM sequence. Among the domains, SRA was found to be large possessing 158 amino acids in length and the two RING domains were not of same size (RING1 is 39 amino acids and RING2 is 57 amino acids).

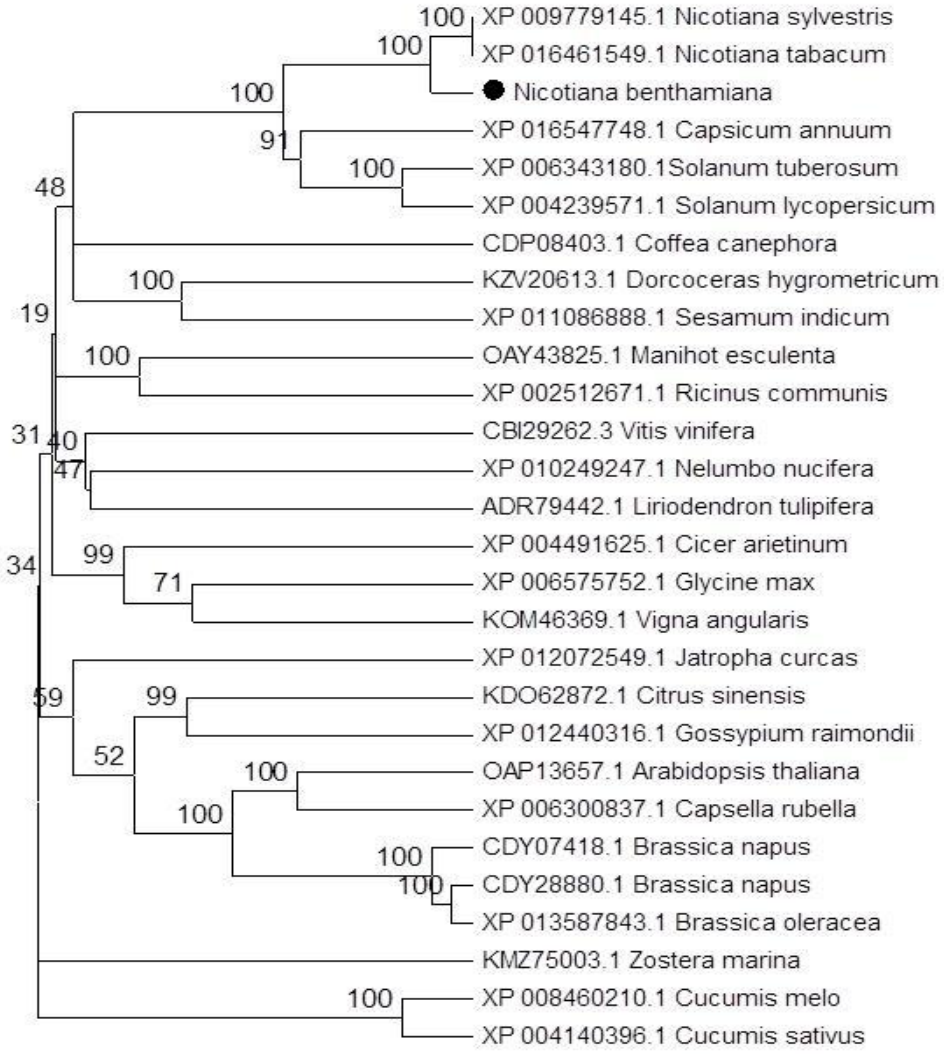


Figure 6.3. Neighbor joining phylogenetic genetic dendrogram of full-length amino acid sequences of VIM proteins. The cloned VIM sequence from *N. benthamiana* and the predicted sequences of other plant species are aligned using MUSCLE algorithm prior to tree generation.

6.2.3 Interaction between geminivirus pre-coat protein and NbVIM

6.2.3.1 Cloning of viral ORFs

For checking interaction between ChiLCV-encoded ORFs and the host factor (NbVIM), viral ORFs were cloned into yeast expression vectors, pGADC1 and pGBKT7. To clone viral ORFs from ChiLCV-New Delhi isolate into pGADC1 vector, *Bam*HI and *Sal*I restriction sites were incorporated into the forward primer and reverse primer, respectively. Clones of pGADT7-C1, pGADT7-C2, pGADT7-C4 and pGADT7-V1 were already available in the laboratory. Therefore, the remaining ORFs of ChiLCV-New Delhi isolate such as, C3 (referred as ChC3), V2 (referred as ChV2) and βC1 (referred as

βC1), were cloned at *Bam*HI and *Sal*I sites in the pGADC1 vector (Figure 6.4A-C). In addition, the V2 ORF of ChiLCV-Nagpur isolate was also cloned into pGADC1 vector at *Bam*HI and *Sal*I restriction sites. This clone was further indicated as NgV2 (Figure 6.4D). Then, NbVIM was cloned at *Bam*HI and *Sal*I sites in pGBKT7 vector to study its interaction with virus-encoded ORFs (Figure 6.4E).

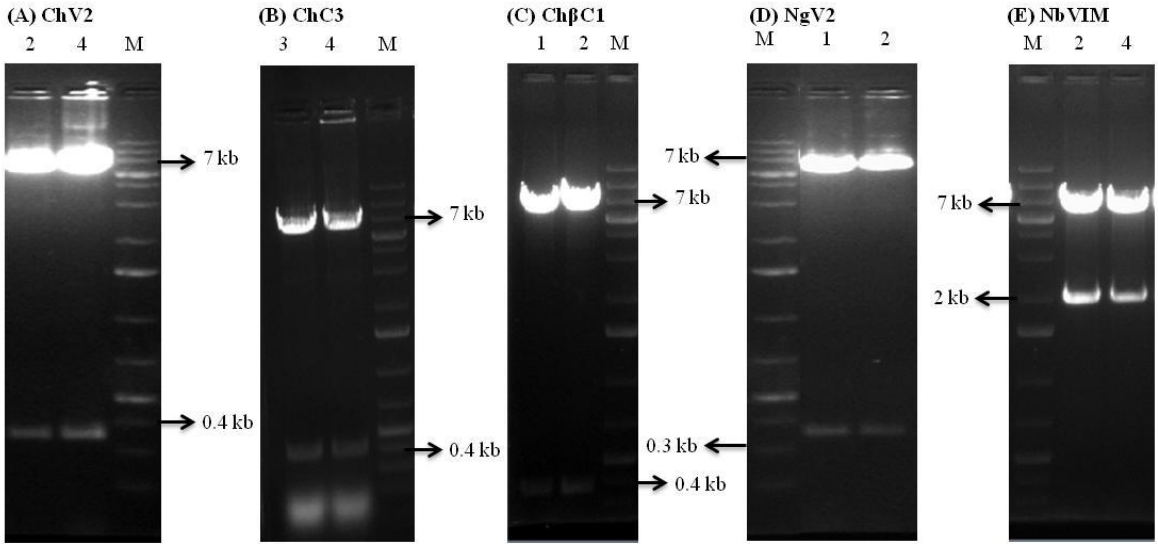


Figure 6.4. Agarose gel electrophoresis showing digestion confirmation of viral ORFs cloned into pGADC1 vector. M refers to 1 kb DNA ladder and the clone numbers used for digestion are provided on top of each gel. The restriction enzymes used for restriction digestion are *Bam*HI and *Sal*I. The DNAs used for digestion are from ChC3 (A), ChV2 (B), ChβC1 (C), NgV2 (D) pGADC1 clones and NbVIM (E) pGBKT7 clones.

6.2.3.2 Interaction between pre-coat protein and NbVIM through Yeast-two-hybrid assay

For studying the physical interaction between virus-encoded ORFs and the cloned NbVIM protein, Yeast-two-hybrid assay was performed. Subsequently, viral ORFs such as C1, C2, C3, C4, V1, V2 and βC1 were cloned downstream to activation domain in yeast expression vector, pGADC1 to generate fusion constructs. Similarly, NbVIM fragments were fused along with DNA binding domain. The yeast cells were co-transformed with appropriate combinations as mentioned earlier and plated onto 2-DO plates. They were subsequently plated onto 3-DO supplemented with 2.5 mM 3-AT. For positive control, yeast cells co-transformed with AD-T-Ag and BD-P53 were used and AD-T-Ag and BD-pLamin C co-transformed yeast cells were used as negative control. The AD-T-Ag and BD-P53 co-transformed yeast cells were grown on the medium,

however, cells co-transformed with AD-T-Ag and BD-pLamin C were failed to grow (Figure 6.5). Further, the growth of yeast cells containing AD-V2 and BD-NbVIM constructs also suggest their physical interaction.

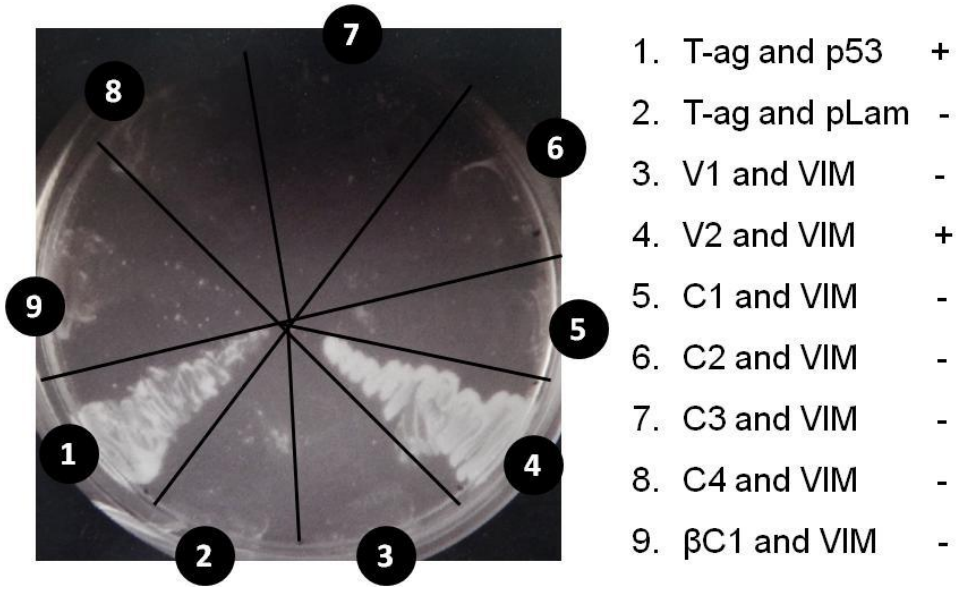


Figure 6.5. Interaction between virus-encoded proteins and NbVIM protein through Yeast-two-hybrid assay. The combinations used to co-transform yeast cells are provided in the right side of the figure. The symbols '+' and '-' indicates positive and negative interactions, respectively.

6.2.3.3 Cloning of various domains of NbVIM into pGBKT7 vector

To identify the regions responsible for interaction between viral ORFs and NbVIM if any, various deletion constructs of NbVIM were designed according to the arrangement of domains present in its sequence. The schematic representation of these constructs is depicted in Figure 6.6.

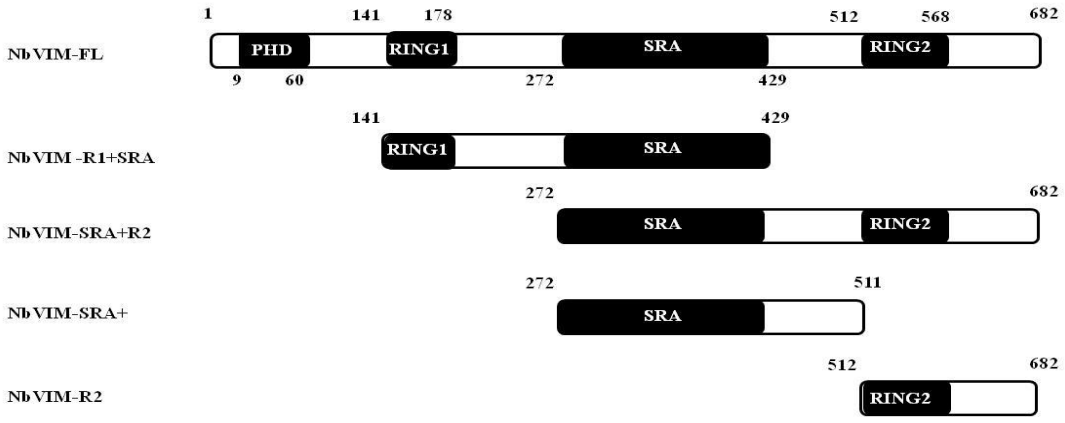


Figure 6.6. Schematic representation of various domains of NbVIM. The names of the construct are provided in the left side and the amino acid co-ordinates are mentioned along with the constructs.

To study the physical interaction of NbVIM with virus-encoded ORFs, the complete sequence and the domain constructs of NbVIM was cloned at *Bam*HI and *Sal*I sites in the pGBKT7 vector (Figure 6.7). All the clones were confirmed through restriction digestion and PCR. Subsequently, these clones were also sequenced.

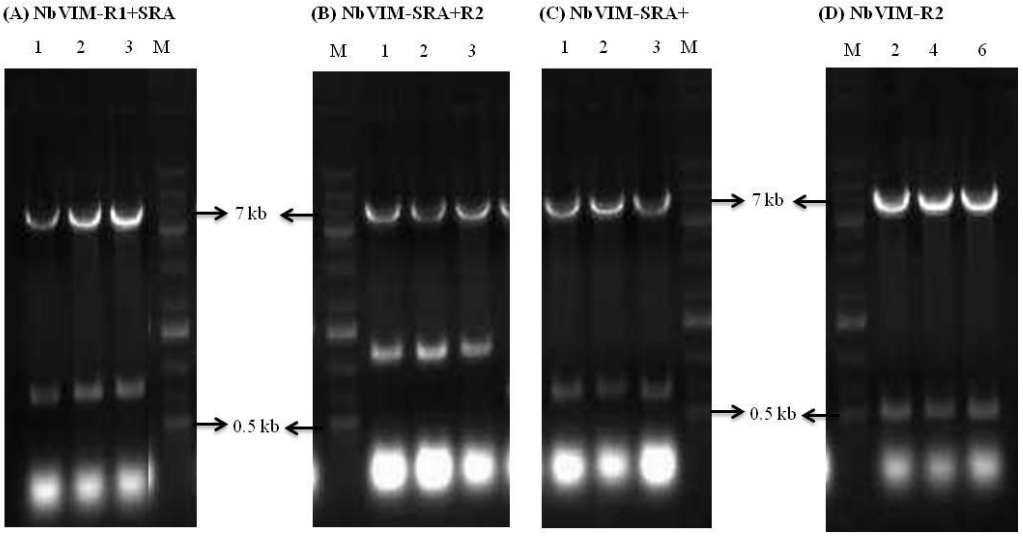


Figure 6.7. Agarose gel electrophoresis showing the digestion confirmation of full-length and domains of NbVIM cloned into pGBKT7 vector. M refers to 1kb DNA ladder and the clone numbers used for digestion are provided with each gel. The restriction enzymes used for restriction digestion are *Bam*HI and *Sal*I. DNAs used for digestion are from NbVIM FL (A), NbVIM R1 (B), NbVIM SRA+R2 (C), NbVIM SRA+ (D) and NbVIM R2 (E) clones.

6.2.3.4 Interaction between pre-coat protein and NbVIM domains through Yeast-two-hybrid assay

Figure 6.8. Interaction between V2 and NbVIM proteins through Yeast-two-hybrid assay. The combinations used to co-transform yeast cells are provided in the right side of the figure. The symbols '+' and '-' indicates positive and negative interactions, respectively.

In order to identify the domain(s) of NBVIM interacting with V2 ORF, several domains of NbVIM were separately co-transformed with the AD-V2 construct. The yeast cells containing AD-V2 and BD-NbVIM SRA+R2 were only able to grow in the 3-DO medium containing 2.5 mM 3-AT (Figure 6.8). This result confirms that NbVIM protein interacts with V2 protein through its SRA and RING2 domain in Yeast-two-hybrid assay.

6.2.3.5 Bimolecular fluorescence complementation (BiFC) assay

6.2.3.5.1 Cloning of the full-length and various domains of NbVIM into BiFC vectors

To further confirm the physical interaction, BiFC analysis was performed, in which the active GFP is reconstituted only when the non-fluorescent N-terminal and C-terminal GFP fragments interact with each other. For BiFC, the V2 ORF was cloned at *Cla*I and *Xho*I sites in pVYNE vector containing 173 amino acids of the N-terminal venus protein. Similarly, the full-length and various domains of NbVIM protein were cloned at *Bam*HI and *Sal*I sites in pVYCE(R) vector having 155 amino acids of C-terminal venus protein. All of these clones were confirmed by digestion with appropriated restriction enzymes (Figure 6.9). Each constructs were then mobilized into *A. tumefaciens* cells (strain GV3101) separately.

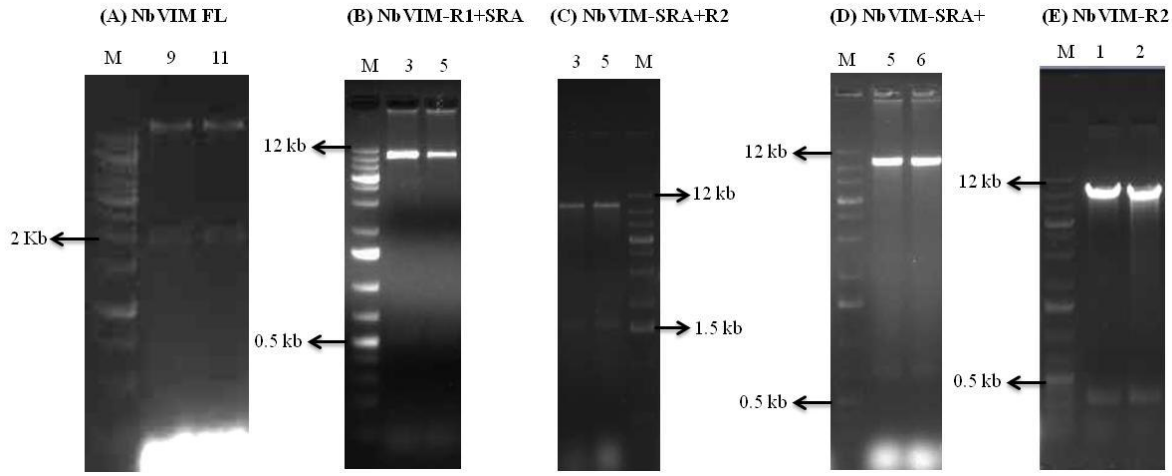


Figure 6.9. Agarose gel electrophoresis confirming cloning of full-length and domains of NbVIM into BiFC vectors. M refers to 1kb DNA ladder and the clone numbers used for digestion are provided with each gel. The restriction enzymes used for restriction digestion are *Bam*HI and *Sal*I. The DNAs used for digestion are from NbVIM FL (A), NbVIM R1 (B), NbVIM SRA+R2 (C), NbVIM SRA+ (D) and NbVIM R2 (E) clones.

6.2.3.5.2 Protein-protein interaction through BiFC analysis

In addition to yeast-two-hybrid assay, BiFC analysis was also carried out to assess the interaction between pre-coat protein and NbVIM proteins *in vivo*. When nVenus-V2 was co-infiltrated along with NbVIM-cVenus, GFP fluorescence was detected in the cell periphery. However, no GFP signal was detected when plants were co-infiltrated with nVenus and NbVIM-cVenus. This result suggests that pre-coat protein and NbVIM protein interacts with each other *in planta* (Figure 6.10). To further identify the domains of NbVIM involved in this interaction, various NbVIM domain constructs were co-infiltrated in the lower epidermis of *N. benthamiana* plants with or without V2 construct. GFP fluorescence was found in the epidermal cells co-infiltrated with nVenus-V2 and NbVIM SRA+R2-cVenus constructs. However, GFP signal was not found in any other combinations used (Figure 6.10). Further, in all combinations, GFP fluorescence was not merged with DAPI signal, suggesting that this interaction occurred in the cytoplasm (Figure 6.10). Collectively, the results from BiFC analysis further validate the interaction between NbVIM and V2 proteins through its SRA and RING2 domains, and this interaction was found to be detected in cell periphery.

6.2.4 Transient silencing of NbVIM by TRV based-virus induced gene silencing

Virus induced gene silencing (VIGS) was performed to study the possible role of NbVIM in viral pathogenesis. VIGS experiment was done using *Tobacco rattle virus* (TRV) based vectors (TRV1 and TRV2) on the *N. benthamiana* plants. To silence NbVIM, 345 nucleotides of C-terminal region of NbVIM were cloned at *Xba*I and *Xho*I sites in TRV2 vector. The restriction digestion and DNA sequencing were used to confirm clones (Figure 6.11A). Silencing of the *NbVIM* transcript was performed by co-infiltrating pTRV1 and pTRV2-C-terminal region of NbVIM into *N. benthamiana* leaves. Accumulation of the NbVIM transcript was evaluated by RNA gel blot analysis after 30 days of infiltration. The transcript levels were compared between the wild type *N. benthamiana* plants and the plants infiltrated with silencing TRV constructs. In wild type *N. benthamiana* plants, NbVIM transcripts were readily detected, however a drastic reduction of NbVIM transcripts was observed in the plants infiltrated with silencing TRV constructs (Figure 6.11B). This result emphasizes that TRV-mediated VIGS assay has efficiently silenced *NbVIM* transcripts in the silenced *N. benthamiana* plants. In addition, in comparable to wild type plants, NbVIM-silenced *N. benthamiana* plants developed severe developmental aberrations (Figure 6.11C).

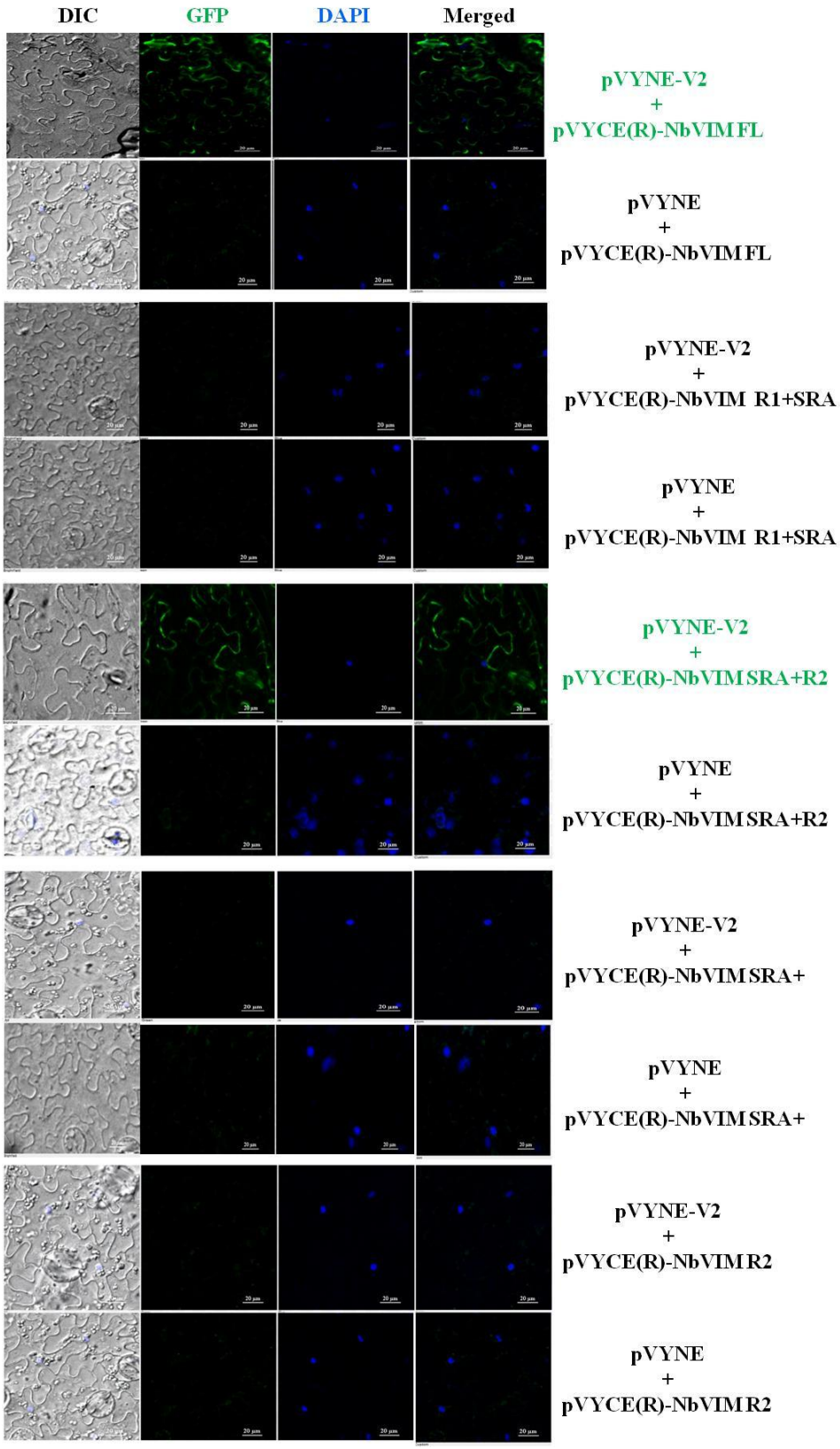


Figure 6.10. Interaction between V2 and NbVIM proteins through BiFC analysis. The combinations used to infiltrate *N. benthamiana* epidermal cells are provided in the right side of the figure. The scale bar measures 20 µM. The nuclei were stained with DAPI dye. Positive interactions are indicated in green text.

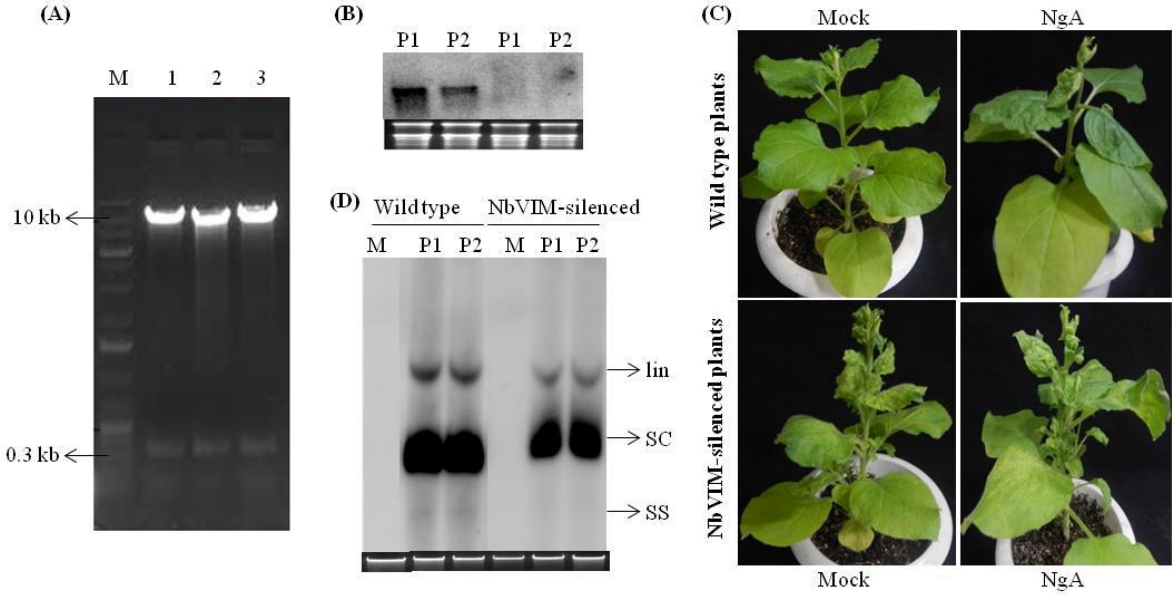


Figure 6.11. Effect of NbVIM on ChiLCV pathogenesis. (A) Agarose gel electrophoresis showing the digestion confirmation of NbVIM C-terminal region cloned into TRV2 vector. M refers to 1 kb DNA ladder and the clone numbers are also provided. The restriction enzymes used for restriction digestion are *Xba*I and *Xho*I. (B) Level of *NbVIM* transcripts in wild type *N. benthamiana* plants (Lanes 1 & 2) and *NbVIM*-silenced *N. benthamiana* plants (Lanes 3 & 4). (C) Symptom appearance on the wild type *N. benthamiana* and *NbVIM*-silenced *N. benthamiana* plants upon mock- and NgA-inoculation, as mentioned in the figure. (D) Accumulation of NgA-specific DNA in wild-type *N. benthamiana* and *NbVIM*-silenced *N. benthamiana* plants upon mock-inoculation (M) and NgA-inoculation (P1 & P2). P1 and P2: DNA/RNA isolated from representative two plants inoculated with vector or ChiLCV-NgA.

Further, infectivity analysis was performed to study the effect of silencing of *NbVIM* on viral pathogenesis. The wild type and *NbVIM*-silenced *N. benthamiana* plants were either mock-inoculated or with NgA infectious construct. Disease development on these inoculated plants was monitored till 20 DPI. Symptoms were initiated at the same time (8 DPI) on both wild type and *NbVIM*-silenced *N. benthamiana* plants. Viral DNA accumulation in these plants was analyzed by Southern blotting at 20 DPI. In comparable to the wild type *N. benthamiana* plants, reduced level of NgA-specific viral DNA in the NbVIM-silenced *N. benthamiana* plants was observed (Figure 6.11D). Results from the Southern analysis suggest that silencing of *NbVIM* affects the accumulation of NgA-specific viral DNA.

6.3 Discussion

RNA silencing plays essential roles in regulation of various developmental processes of plants by altering gene expression either at the transcriptional level through DNA / histone methylation or at the post-transcriptional level through the degradation of target mRNA (Law and Jacobsen, 2010; Hohn and Vazquez, 2011; Saze *et al*., 2012). Plant viruses can either suppress or hijack host defense pathway including the components of antiviral silencing for their survival and subsequent spread within a host (Hohn and Vazquez, 2011). These viruses are widely known to be the inducers and targets of the host-mediated gene silencing pathway (Pumplin and Voinnet, 2013). Geminiviruses also encode several proteins suppressing these antiviral silencing pathways (Hanley-Bowdoin *et al*., 2013). Begomovirus-encoded pre-coat protein (encoded by ORF V2) has been demonstrated to be a pathogenicity determinant and is also involved in systemic spread (Boulton *et al.,* 1989; Priyadarshini *et al*., 2009; Sharma and Ikegami, 2010; Sharma *et al.,* 2010; Stanley *et al.,* 1992; Wartig *et al.,* 1997). These proteins function as suppressor of PTGS by interacting with SGS3 and two cysteine amino acid residues has been demonstrated to be indispensible for PTGS suppression (Glick *et al.,* 2008). Further, the pre-coat protein of TYLCCNV has been shown to counter TGS by reduction cytosine methylation (Wang *et al.,* 2014). However, the mechanistic details of pre-coat protein mediated suppression of TGS are not yet elucidated. In the present study, interaction between the V2 and one of the components of TGS pathway (NbVIM) has been demonstrated. Also, a possible role of NbVIM in ChiLCV pathogenesis has been elucidated.

Pre-coat protein of begomoviruses functions as PTGS suppressor by binding to SGS3, thereby inhibits secondary siRNA generation (Glick *et al.,* 2008; Fukunaga and Doudna, 2009). Further, as begomoviruses contain ssDNA as genetic material, the report of pre-coat protein's TGS suppression activity is quite interesting. The TGS mechanism involves DNA methylation and/or histone methylation and except begomoviruses-encoded TrAP and βC1 proteins, the mechanistic details of TGS suppression activities are yet to be elucidated (Hanley-Bowdoin *et al*., 2013). In this context, the interaction between pre-coat protein and NbVIM, a component of the TGS pathway, possesses functional significance. VIM/UHRF proteins have been known to act as a scaffold protein and are possibly responsible to recruit several players necessary in DNA replication (PCNA) and

gene silencing (MET1) in plants (Chuang *et al*., 1997; Bostick *et al*., 2007; Law and Jacobsen, 2010; Kim *et al*., 2014). In addition, N-terminal region of VIM protein has been shown to be interacting with NtSET, a member of H3K9 methylation (Liu *et al*., 2007). The available information highlights that VIM protein is necessary to modulate gene expression by regulating chromatin and/or histone methylation either directly or indirectly. Hence, the interaction of the pre-coat protein with NbVIM suggests that it can regulate the components of antiviral silencing pathway against geminiviruses, thereby countering this host-mediated defense strategy and enabling them to evade antiviral mechanism.

It is well documented that each domains of VIM proteins posses several essential biological functions. The PHD domain located at the N-terminal region of VIM interacts with NtSET, SRA domain binds to hemi-methylated DNA for maintenance of DNA methylation and both RING domains have been reported to function as ubiquitin E3 ligase (Kraft *et al.,* 2008; Liu *et al*., 2007; Woo *et al.,* 2007). The protein-protein interaction studies emphasize that pre-coat protein interacts with NbVIM protein through SRA and RING (at C-terminal region) domains. Two possible hypotheses can be put forward based on their interaction. One possibility is that interaction of the pre-coat protein with NbVIM could probably inhibit the binding of SRA domain to the hemi-methylated DNA strand and affect the process of maintenance of methylation. This might lead to reduced level of methylation of the viral promoter which ultimately, could increase viral DNA titer in infected plants. However, the level of virus-specific DNA was found to be reduced in the NbVIM-silenced plants than the wild-type *N. benthamiana* plants in VIGS experiment. Collectively, from the present analyses, it can be speculated that inhibition of the SRA domain of NbVIM protein on binding of hemi-methylated DNA by pre-coat protein is not certain. However, more detailed experimental evidences are required to support this assumption. Another hypothesis is that the interaction of pre-coat protein with RING domain at C-terminal of NbVIM might prevent ubiquitylation of the target proteins. However, additional experimentation is required to determine the targets of these VIM proteins to ubiquitylate them. In this scenario, it is tempting to speculate that pre-coat protein could interfere with any of these said biological functions of NbVIM to enable the virus to evade host-mediated defense machinery and favors viral multiplications in a permissive host.

In summary, plant viruses require a living host cell for the infection and spread of the disease, hence they have to suppress or evade host-mediated antiviral strategies. In the present work, interaction of pre-coat protein and a component of TGS pathway (NbVIM) have been demonstrated. This interaction is through SRA and RING domains of NbVIM protein. Silencing of NbVIM in *N. benthamiana* plants reduce the level of virus-specific DNA accumulation and the possible role of NbVIM in viral pathogenesis also has been discussed.

7. Summary

Geminiviruses comprise of ssDNA viruses that can infect monocot as well as dicot plants. The genus, *Begomovirus*, includes viruses one among which cause chilli leaf curl disease (ChiLCD). ChiLCD has been reported to be the major limiting factor for chilli cultivation in the Indian Sub-continent. However, the complexity of the viral components associated with ChiLCD in the country and its infectivity nature is lacking. This study reports the association of begomoviruses and their satellite components with ChiLCD in India and their evolutionary characteristics. In addition, nature of infectivity and the role of pre-coat protein in the pathogenesis have been elucidated. The major findings of this study are summarized below-

1. Survey of major chilli growing regions was carried out across 17 different states in India. Twenty-eight samples from these states were collected from ChiLCD-affected plants.
2. Multiple begomoviruses (35 DNA-A & 6 DNA-B) with betasatellite (33-DNA β) and alphasatellite (13) molecules were isolated and characterized from these samples.
3. A total of 7 distinct begomovirus species, 5 different betasatellite groups and 2 groups of alphasatellites were found to be associated with these ChiLCD samples.
4. One previously unreported begomovirus (*Chilli leaf curl Vellanad virus*) and alphasatellite (Chilli leaf curl alphasatellite) was found to be associated with ChiLCD in India.
5. Emergence of these begomoviruses and associated satellites are largely driven by intra- and/or inter-species recombination.
6. Higher genetic variability was noticed in the ChiLCD-associated begomoviral genomic components.
7. Nucleotide substitution rates (ChiLCV-V1: 2.60×10^{-3} nucleotides/site/year, ChiLCB-βC1: 2.57×10^{-4} nucleotides/site/year, ChiLCA-Rep: 2.25×10^{-3} nucleotides/site/year) were found to be in concurrent with other ssDNA viruses.
8. *Nicotiana benthamiana* and *Capsicum annuum* plants were used for agro-inoculation studies. Both betasatellites and alphasatellite were found to be dispensable in *N. benthamiana* plants.

9. In natural host (*C annuum*), the betasatellites were found to be indispensable for the leaf curl disease development.
10. Three already known begomoviruses and their cognate betasatellites have been demonstrated to infect chilli plants.
11. The present study, for the first time demonstrated that *Chilli leaf curl virus, Chilli leaf curl Vellanad virus* and *Pepper leaf curl Bangladesh virus* along with their cognate betasatellite cause leaf curl disease in chilli, thus fulfilling koch's postulates.
12. The role of V2 encoded protein (pre-coat protein) of *Chilli leaf curl virus*-Nagpur was found to be involved in viral pathogenesis.
13. Reduced accumulation of NbVIM transcript has been detected upon virus infection.
14. ChiLCV-encoded pre-coat protein interacted with *N. benthamiana* VARIANT IN METHYLATION (NbVIM) protein based on yeast-2-hybrid and BiFC assays.
15. Interaction between NbVIM protein and the pre-coat protein of ChiLCV occurs through SRA (SET- AND RING-ASSOCIATED) and RING2 (REALLY INTERESTING GENE 2) domains of the NbVIM protein.
16. TRV based VIGS vector could effectively silence expression of *NbVIM* in *N. benthamiana*.
17. Viral DNA accumulation was found to be lowered in *NbVIM* silenced *N. benthamiana* plants, thereby suggesting possible role of NbVIM in modulating viral pathogenesis.

8. References

Arguello-Astorga G, Lopez-Ochoa L, Kong LJ, Orozco BM, Settlage SB & Hanley-Bowdoin L (2004) A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. J Virol 78:4817–4826.

Aufsatz W, Mette MF, van der Winden J, Matzke M & Matzke AJ (2002) HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. EMBO J 21:6832–6841.

Ausin I, Mockler TC, Chory J & Jacobsen SE (2009) IDN1 and IDN2 are required for de novo DNA methylation in *Arabidopsis thaliana*. Nat Struct Mol Biol 16:1325–1327.

Bagewadi B, Chen S, Lal SK, Choudhury NR & Mukherjee SK (2004) PCNA interacts with *Indian Mungbean yellow mosaic virus* Rep and downregulates Rep activity. J Virol 78:11890–11903.

Bar-Ziv A, Levy Y, Citovsky V & Gafni Y (2015) The *Tomato yellow leaf curl virus* (TYLCV) V2 protein inhibits enzymatic activity of the host papain-like cysteine protease CYP1. Biochem Biophys Res Commun 460:525–529.

Bar-Ziv A, Levy Y, Hak H, Mett A, Belausov E, Citovsky V & Gafni Y (2012) The *Tomato yellow leaf curl virus* (TYLCV) V2 protein interacts with the host papain-like cysteine protease CYP1. Plant Signal & Behav 7:983-989.

Bernatavichute YV, Zhang X, Cokus S, Pellegrini M & Jacobsen SE (2008) Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. PLoS ONE 3:e3156.

Bhattacharyya D, Prabu G, Kumar RK, Kushwaha NK, Sharma VK, Yusuf MdA & Chakraborty S (2015) A geminivirus betasatellite damages structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. J Exp Bot 66(19):5881-5895.

Bimboin HC & Dolly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513-1523.

Blevins T, Pontvianne F, Cocklin R, Podicheti R, Chandrasekhara C, Yerneni S, Braun C, Lee B, Rusch D, Mockaitis K, Tang H & Pikaard CS (2014) A two-step process for epigenetic inheritance in *Arabidopsis*. Mol Cell 54:30–42.

Bochman ML, Schwacha A (2008) The Mcm2-7 complex has in vitro helicase activity. Mol Cell 31:287–293

Bosland PW & Votava EJ (1999) Peppers. In vegetable and spice Capsicums. pp. 177. CABI publishing, New York.

Bostick M, Kim JK, Estève PO, Clark A, Pradhan S & Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 317:1760–1764.

Bottcher B, Unseld S, Ceulemans H, Russell RB & Jeske H (2004) Geminate structures of *African cassava mosaic virus*. J Virol 78:6758-6765.

Boulton MI, Pallaghy CK, Chatani M, MacFarlane S & Davies JW (1993) Replication of *Maize streak virus* mutants in maize protoplasts: evidence for a movement protein. Virology 192:85–93.

Boulton MI, Steinkellner H, Donson J, Markham PG, King DI & Davies JW (1989) Mutational analysis of the virion-sense genes of *Maize streak virus*. J Gen Virol 70:2309–2323.

Briddon RW, Bull SE, Amin I, Mansoor S, Bedford ID, Rishi N, Siwatch SS, Zafar Y, Abdel-Salam AM & Markham PG (2004) Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus-DNA beta complexes. Virology 324:462-474.

Briddon RW & Stanley J (2006) Subviral agents associated with plant single stranded DNA viruses. Virology 344:198–210.

Briddon RW, Bedford ID, Tsai JH & Markham PG (1996) Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, *Tomato pseudo-curly top virus*, suggests a recombinant origin. Virology 219:387-394.

Briddon RW, Bull SE, Mansoor S, Amin I & Markham PG (2002) Universal primers for the PCR-mediated amplification of DNAβ: a molecule associated with some monopartite begomoviruses. Mol Biotechnol 20:315-318.

Briddon RW, Heydarnejad J, Khosrowfar F, Massumi H, Martin DP & Varsani A (2010) *Turnip curly top virus*, a highly divergent geminivirus infecting turnip in Iran. Virus Res 152:169–175.

Briddon RW, Watts J, Markham PG & Stanley J (1989) The coat protein of *Beet curly top virus* is essential for infectivity. Virology 172:628–633.

Brown, J.K., Fauquet, C.M., Briddon, R.W., Zerbini, F.M., Moriones, E., Navas-Castillo, J., 2012. Family *Geminiviridae*. In Virus Taxonomy 9th report of the *International Committee on Taxonomy of Viruses*, pp. 351-373. Edited by King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., London, UK: Elsevier Academic Press.

Brown JK, Zerbini, FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, Fiallo-Olive E, Briddon RW, Hernandez-Zepeda C, Idris A, Malathi VG, Martin DP, Rivera-Bustamante R, Ueda S & Varsani A (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Arch Virol 160:1593-1619.

Buchmann RC, Asad S, Wolf JN, Mohannath G & Bisaro DM (2009) Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. J Virol 83:5005-5013.

Bull SE, Briddon RW & Markham PG (2003) Universal primers for the PCR-mediated amplification of DNA 1: a satellite-like molecule associated with begomovirus-DNAβ complexes. Mol Biotechnol 23:83-86.

Bull SE, Briddon RW, Serubombwe WS, Ngugi K, Markham PG & Stanley J (2007) Infectivity, pseudo-recombination and mutagenesis of Kenyan cassava mosaic begomoviruses. J Gen Virol 88:1624–1633.

Cao X & Jacobsen SE (2002) Role of the *Arabidopsis* DRM methyl transferases in *de novo* DNA methylation and gene silencing. Curr Biol 12:1138–1144.

Ceniceros-Ojeda EA, Rodríguez-Negrete EA & Rivera-Bustamante RF (2016) Two populations of viral mini-chromosomes are present in a geminivirus-infected plant showing symptom remission (Recovery). J Virol 90:3828-3838.

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 97:51–57.

Carvalho CM, Fontenelle MR, Florentino LH, Santos AA, Zerbini FM & Fontes EP (2008) A novel nucleocytoplasmic traffic GTPase identified as a functional target of the bipartite geminivirus nuclear shuttle protein. Plant J 55:869-880.

Carvalho MF, Turgeon R & Lazarowitz SG (2006) The geminivirus nuclear shuttle protein NSP inhibits the activity of AtNSI, a vascular-expressed *Arabidopsis* acetyl transferase regulated with the sink-to-source transition. Plant Physiol 140:1317-1330.

Castellano MM, Sanz-Burgos AP & Gutiérrez C (1999) Initiation of DNA replication in a eukaryotic rolling-circle replicon: identification of multiple DNA–protein complexes at the geminivirus origin. J Mol Biol 290:639–652.

Castillo AG, Collinet D, Deret S, Kashoggi A & Bejarano ER (2003) Dual interaction of plant PCNA with geminivirus replication accessory protein (REn) and viral replication protein (Rep). Virology 312:381–394.

Castillo-Gonzalez C, Liu X, Huang C, Zhao C, Ma Z, Hu T, Sun F, Zhou Y, Zhou X, Wang X-J & Zhang X (2015) Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense. eLife 4:e06671.

Chakraborty S, Pandey PK, Banerjee MK , Kalloo G & Fauquet CM (2003) *Tomato leaf curl Gujarat virus*, a new begomovirus species causing a severe leaf curl disease of tomato in Varanasi, India. Phytopathology 93:1485–1495.

Chakraborty S, Vanitharani R, Chattopadhyay B & Fauquet CM (2008) More virulent recombination and asymmetric synergism between two distinct species of causing tomato leaf curl disease in India. J Gen Virol 89:818–828.

Chandran SA, Levy Y, Mett A, Belausov E, Ramakrishnan U & Gafni Y (2012) Mapping of functional region conferring nuclear localization and karyopherin α-binding activity of the C2 protein of *Bhendi yellow vein mosaic virus*. J Gen Virol 93:1367–1374.

Chatchawankanphanich O & Maxwell DP (2002) *Tomato leaf curl Karnataka virus* from Bangalore, India, appears to be a recombinant begomovirus. Phytopathology 92:637-645.

Chattopadhyay B Singh AK, Yadav T, Fauquet CM, Sarin NB & Chakraborty S (2008) Infectivity of the cloned components of a begomovirus: DNA-β complex causing chilli leaf curl disease in India. Arch Virol 153:533–539.

Chellappan P, Vanitharani R & Fauquet CM (2004) Short interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts and gene silencing targets specific viral sequences. J Virol 78:7465–7477.

Chhabra N, Aseri M L, Goyal V & Sankhla S (2012) Capsaicin: A promising therapy - A critical reappraisal. Int J Nutr Pharmacol Neurol Dis 2:8-15.

Choudhury NR, Malik PS, Singh DK, Islam MN, Kaliappan K & Mukherjee SK (2006) The oligomeric Rep protein of *Mungbean yellow mosaic India virus* (MYMIV) is a likely replicative helicase. Nucleic Acids Res 34:6362–6377.

Chowda-Reddy RV, Achenjang F, Felton C, Etarock MT, Anangfac MT, Nugent P & Fondong VN (2008) Role of a geminivirus AV2 protein putative protein kinase C motif on subcellular localization and pathogenicity. Virus Res 135:115-124.

Chuang LS, Ian HI, Koh TW, Ng HH, Xu G & Li BF (1997) Human DNA-(cytosine-5) methyl transferase–PCNA complex as a target for p21WAF1. Science 277:1996–2000.

Chung HY & Sunter G (2014) Interaction between the transcription factor AtTIFY4B and begomovirus AL2 protein impacts pathogenicity. Plant Mol Biol 86(1-2):185-200.

Chung HY, Lacatus G & Sunter G (2014) Geminivirus AL2 protein induces expression of, and interacts with, a calmodulin-like gene, an endogenous regulator of gene silencing. Virology 460-461:108–118.

Clerot D & Bernardi F (2006) DNA helicase activity is associated with the replication initiator protein rep of tomato yellow leaf curl geminivirus. J Virol 80:11322-11330.

Collin S, Fernandez-Lobato M, Gooding PS, Mullineaux PM & Fenoll C (1996) The two nonstructural proteins from *Wheat dwarf virus* involved in viral gene expression and replication are retinoblastoma-binding proteins. Virology 219:324–329.

Cui X, Li G, Wang D, Hu D & Zhou X (2005) A Begomovirus DNA-β encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. J Virol 79:10764–10775.

Dellaporta SL, Wood J & Hicks JB (1983) A plant DNA mini-preparation: Version II. Plant Mol Biol Rep 1:19-21.

Dennis K, Fan T, Geiman T, Yan Q & Muegge K (2001) Lsh, a member of the SNF2 family, is required for genome-wide methylation. Genes Dev 15:2940–2944.

Dhanraj KS & Seth ML (1968) A strain of tobacco leaf curl virus causing enation chilli (*Capsicum annuum* L.). Indian J Hort 25:70-71.

Dong X, van Wezel R, Stanley J & Hong Y (2003) Functional characterization of the nuclear localization signal for a suppressor of posttranscriptional gene silencing. J Virol 77:7026-7033.

Dogar AM (2006) RNAi dependent epigenetic marks on a geminivirus promoter. Virol J 3:5.

Dogra SC, Eini O, Rezaian MA & Randles JW (2009) A novel shaggy-like kinase interacts with the *Tomato leaf curl virus* pathogenicity determinant C4 protein. Plant Mol Biol 71;25–38.

Drake JW (1993) Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci USA 90:4171–4175.

Drummond AJ, Suchard MA, Xie D & Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973.

Dry IB, Krake LR, Rigden JE & Rezaian MA (1997) A novel subviral agent associated with a geminivirus: The first report of a DNA satellite. Proc Natl Acad Sci USA 94:7088–7093.

Du J, Johnson LM, Jacobsen SE & Patel DJ (2015). DNA methylation pathways and their crosstalk with histone methylation. Nat Rev Mol Cell Biol 16:519-532.

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, Hetzel J, Wohlschlegel JA, Pradhan S, Patel DJ & Jacobsen SE (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell 151:167–180.

Duan YP, Powell CA, Purcifull DE, Broglio P & Hiebert E (1997) Phenotypic variation in transgenic tobacco expressing mutated geminivirus movement/pathogenicity (BC1) proteins. Mol Plant Microbe Interact 10:1065-1074.

Duffy S & Holmes (2009) Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J Gen Virol 90:1539–1547.

Duffy S & Holmes EC (2008) Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. J Virol 82:957–965.

Eagle PA, Orozco BM & Hanley-Bowdoin L (1994) A DNA sequence required for geminivirus replication also mediates transcriptional regulation. Plant Cell 6:1157–1170.

Eini O, Behjatnia SAA, Dogra S, Dry IB, Randles JW & Rezaian MA (2009) Identification of sequence elements regulating promoter activity and replication of a monopartite begomovirus associated DNA b satellite. J Gen Virol 90:253–260.

Elmer JS, Brand L, Sunter G, Gardiner WE, Bisaro DM & Rogers SG (1988) Genetic analysis of the *tomato golden mosaic virus*. II. Requirement for the product of the highly conserved ALI coding sequence for replication. Nucleic Acids Res 16:7043–7060.

El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R & Lagrange T (2007) Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev 21:2539–2544.

Etessami, P, Saunders K, Watts J & Stanley J (1991) Mutational analysis of complementary-sense genes of *African cassava mosaic virus* DNA-A. J Gen Virol 72:1005–1012.

Feinberg AP & Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Ann Biochem 132:6-13.

Fiallo-Olivé E, Martínez-Zubiaur Y, Moriones E & Navas-Castillo J (2012) A novel class of DNA satellites associated with New World begomoviruses. Virology 426:1–6.

Florentino LH, Santos AA, Fontenelle MR, Pinheiro GL, Zerbini FM, Baracat-Pereira MC & Fontes EP (2006) A PERK-like receptor kinase interacts with the geminivirus nuclear shuttle protein and potentiates viral infection. J Virol 80:6648–6656.

Fondong VN (2013) Geminivirus protein structure and function. Mol plant pathol 14:635–649.

Fondong VN, Reddy RV, Lu C, Hankoua B, Felton C, Czymmek K & Achenjang F (2007) The consensus N-myristoylation motif of a geminivirus AC4 protein is required for membrane binding and pathogenicity. Mol Plant Microbe Interact 20:380–391.

Fontenelle MR, Luz DF, Gomes APS, Florentino LH, Zerbini FM & Fontes EP (2007) Functional analysis of the naturally recombinant DNA-A of the bipartite begomovirus *Tomato chlorotic mottle virus*. Virus Res 126:262–267.

Fontes EP, Santos AA, Luz DF, Waclawovsky AJ & Chory J (2004) The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. Genes Dev 18:2545–2556.

Fontes EPB, Eagle PA, Sipe PS, Luckow VA & Hanley-Bowdoin L (1994a) Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. J Biol Chem 269:8459–8465.

Fontes EPB, Gladfelter HJ, Schaffer RL, Petty TD & Hanley-Bowdoin L (1994b) Geminivirus replication origins have a modular organization. Plant Cell 6:405–416.

Fontes EPB, Luckow VA & Hanley-Bowdoin L (1992) A geminivirus replication protein is a sequence-specific DNA binding protein. Plant Cell 4:597–608.

Fukunaga R & Doudna JA (2009) dsRNA with 5' overhangs contributes to endogenous and antiviral RNA silencing pathways in plants. EMBO J 28:545-555.

Garrido-Ramirez ER, Sudarshana MR, Lucas WJ & Gilbertson RL (2000) *Bean dwarf mosaic virus* BV1 protein is a determinant of the hypersensitive response and avirulence in *Phaseolus vulgaris*. Mol. Plant Microbe Interact 13:1184–1194.

George B, Alam CM, Kumar RV, Gnanasekaran P & Chakraborty S (2015) Potential linkage between compound microsatellites and recombination in geminiviruses: evidence from comparative analysis. Virology 482:41–50.

George B, Kumar RV & Chakraborty S (2014) Molecular characterization of *Chilli leaf curl virus* and satellite molecules associated with leaf curl disease of *Amaranthus* spp. Virus Genes 48:397-401.

Gharouni Kardani S, Heydarnejad J, Zakiaghl M, Mehrvar M, Kraberger S & Varsani A (2013) Diversity of *Beet curly top Iran virus* isolated from different hosts in Iran. Virus Genes 46:571–575.

Glick E, Zrachya A, Levy Y, Mett A, Gidoni D, Belausov E, Citovsky V & Gafni Y (2008) Interaction with host SGS3 is required for suppression of RNA silencing by *Tomato yellow leaf curl virus* V2 protein. Proc Natl Acad Sci USA 105:157–161.

Gomes XV, Schmidt SL & Burgers PM (2001) ATP utilization by yeast replication factor C. II. Multiple stepwise ATP binding events are required to load proliferating cell nuclear antigen onto primed DNA. J Biol Chem 276:34776–34783.

Guerra-Peraza O, Kirk D, Seltzer V, Veluthambi K, Schmit AC, Hohn T & Herzog E (2005) Coat proteins of *Rice tungro bacilliform virus* and *Mungbean yellow mosaic virus* contain multiple nuclear-localization signals and interact with importin alpha. J Gen Virol 86:1815–1826.

Guha D, Priyadarshini CGP, Purakayastha A, Thippeswamy R, Lakshmikanth M & Savithri HS (2013) Biochemical characterization of C4 protein of *Cotton leaf curl Kokhran Virus*-Dabawali. Biochim Biophys Acta 1830:3734-3744.

Gutierrez C (1999) Geminivirus DNA replication. Cell Mol Life Sci 56:313-329.

Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, & Pikaard CS (2012) *In vitro* transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Mol Cell 48:811–818.

Hak H, Levy Y, Chandran SA, Belausov E, Loyter A, Lapidot M & Gafni Y (2016) TYLCV-is movement *in planta* does not require V2 protein. Virology 477:56–60.

Haley A, Zhan X, Richardson K, Head K & Morris B (1992) Regulation of the activities of *African cassava mosaic virus* promoters by the AC1, AC2, and AC3 gene products. Virology 188:905–909.

Hallan V & Gafni Y (2001) *Tomato yellow leaf curl virus* (TYLCV) capsid protein (CP) subunit interactions: implications for viral assembly. Arch Virol 146:1765-1773.

Hanahan D, Jessee J & Bloom FR (1991) Plasmid transformation of *Escherichia coli* and other bacteria. Methods Enzymol 204:63-113.

Hanson SF, Hoogstraten RA, Ahlquist P, Gilbertson RL, Russell DR & Maxwell DP (1995) Mutational analysis of a putative NTP binding domain in the replication-associated protein (AC1) of Bean golden mosaic geminivirus. Virology 211:1–9.

Hao L, Wang H, Sunter G & Bisaro DM (2003) Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. Plant Cell 15:1034–1048.

Harrison BD, Barker H, Bock KR, Guthrie EJ, Meredith G & Atkinson M (1977) Plant viruses with circular single-stranded DNA. Nature 270:760-762.

Hartitz MD, Sunter G & Bisaro DM (1999) The *Tomato golden mosaic virus* transactivator (TrAP) is a single-stranded DNA and zinc-binding phosphoprotein with an acidic activation domain. Virology 263:1–14.

Hehnle S, Wege C & Jeske H (2004) Interaction of DNA with the movement proteins of geminiviruses revisited. J Virol 78:7698–7706.

Hanley-Bowdoin L, Bejarano ER, Robertson D & Mansoor S (2013) Geminiviruses: masters at redirecting and reprogramming plant processes. Nat Rev Microbiol 11, 777-788.

Hernandez-Zepeda C, Varsani A & Brown JK (2013) Intergeneric recombination between a new, spinach-infecting curtovirus and a new geminivirus belonging to the genus Becurtovirus: first New World exemplar. Arch Virol 158:2245–2254.

He XJ, Hsu YF, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, Liu HL, Wang CS, Jin H & Zhu JK (2009) An effector of RNA-directed DNA methylation in Arabidopsis is an ARGONAUTE 4- and RNA-binding protein. Cell 137:498–508.

Herr AJ, Jensen MB, Dalmay T & Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. Science 308:118–120.

Herranz MC, Pallas V & Aparicio F (2012) Multifunctional roles for the N-terminal basic motif of *Alfalfa mosaic virus* coat protein: nucleolar/cytoplasmic shuttling, modulation of RNA-binding activity, and virion formation. Mol Plant Microbe Interact 25:1093–1103.

Heyraud F, Matzeit V, Schaefer S, Schell J & Gronenborn B (1993) The conserved nonanucleotide motif of the geminivirus stem-loop sequence promotes replicational release of virus molecules from redundant copies. Biochimie 75:605–615.

Hohn T & Vazquez F (2011) RNA silencing pathways of plants: Silencing and its suppression by plant DNA viruses. Biochimica et Biophysica Acta 1809:588–600.

Hoogstraten RA, Hanson SF & Maxwell DP (1996) Mutational analysis of the putative nicking motif in the replication-associated protein (AC1) of bean golden mosaic geminivirus. Mol. Plant Microbe Interact 9:594–599.

Hormuzdi SG & Bisaro DM (1995) Genetic analysis of *Beet curly top virus*: examination of the roles of L2 and L3 genes in viral pathogenesis. Virology 206:1044–1054.

Hou YM & Gilbertson RL (1996) Increased pathogenicity in a pseudo-recombinant bipartite geminivirus correlates with intermolecular recombination. J Virol 70:5430–5436.

Hou YM, Sanders R, Ursin VM & Gilbertson RL (2000) Transgenic plants expressing geminivirus movement proteins: abnormal phenotypes and delayed infection by *Tomato mottle virus* in transgenic tomatoes expressing the *Bean dwarf mosaic virus* BV1 or BC1 proteins. Mol Plant Microbe Interact 13:297-308.

Huettel B, Kanno T, Daxinger L, Aufsatz W, Matzke AJ & Matzke M (2006) Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. EMBO J 25:2828–2836.

Husain MA (1932) Leaf-curl in cotton and other plants. Nature 129:797.

Hussain M, Mansoor S, Iram S, Zafar Y & Briddon RW (2007) The hypersensitive response to *Tomato leaf curl New Delhi virus* nuclear shuttle protein is inhibited by transcriptional activator protein. Mol Plant Microbe Interact 20:1581–1588.

Idris AM, Shahid MS, Briddon RW, Khan AJ, Zhu JK & Brown JK (2011) An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. J Gen Virol 92:706–717.

Ilyas M, Qazi J, Mansoor S & Briddon RW (2010) Genetic diversity and phylogeography of begomoviruses infecting legumes in Pakistan. J Gen Virol 91:2091–2101.

Ilyina TV & Koonin EV (1992) Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Res 20:3279-3285.

Ingham DJ, Pascal E & Lazarowitz SG (1995) Both bipartite geminivirus movement proteins define viral host range, but only BL1 determines viral pathogenicity. Virology 207:191–204.

Inoue-Nagata AK, Albuquerque LC, Rocha WB & Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. J Virol Methods 116:209–211.

Jackel JN, Buchmann RC, Singhal U & Bisaro DM (2015) Analysis of geminivirus AL2 and L2 proteins reveals a novel AL2 silencing suppressor activity. J Virol 89(6):3176-3187.

Jackson JP, Lindroth AM, Cao X & Jacobsen SE (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyl transferase. Nature 416:556–560.

Jackson JP, Johnson L, Jasencakova Z, Zhang X, PerezBurgos L, Singh PB, Cheng X, Schubert I, Jenuwein T & Jacobsen SE (2004) Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana*. Chromosoma 112:308–315.

Jeske H (2009) Geminiviruses. Curr Top Microbiol Immunol 331:185-226.

Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J & Jacobsen SE (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr Biol 17:379–384.

Jupin I, De Kouchkovsky F, Jouanneau F & Gronenborn B (1994) Movement of tomato yellow leaf curl geminivirus (TYLCV): involvement of the protein encoded by ORF C4. Virology 204;82–90.

Jyothsna P, Haq QMI, Singh P, Sumiya KV, Praveen S, Rawat R, Briddon RW & Malathi VG (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. Appl Microbiol Biotechnol 97:5457–5471.

Kaliappan K, Choudhury NR, Suyal G & Mukherjee SK (2012) A novel role for RAD54: this host protein modulates Geminiviral DNA replication. FASEB J 26:1142–1160.

Kammann M, Schalk HJ, Matzeit V, Schaefer S, Schell J & Gronenborn B (1991) DNA replication of wheat dwarf virus, a geminivirus, requires two cis-acting signals. Virology 184:786–790.

Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M & Matzke AJ (2004) Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Curr Biol 14:801–805.

Khan MS, Raj SK & Singh R (2006) First report of *Tomato leaf curl New Delhi virus* infecting chilli in India. Plant Pathol 55:289.

Kheyr-Pour A, Bananej K, Dafalla GA, Caciagli P, Noris E, Ahoonmanesh A, Lecoq H & Gronenborn B (2000) *Watermelon chlorotic stunt virus* from the Sudan and Iran: sequence comparisons and identification of a whitefly transmission determinant. Phytopathology 90:629–635.

Kim J, Kim JH, Richards EJ, Chung KM & Woo HY (2014) *Arabidopsis* VIM proteins regulate epigenetic silencing by modulating DNA methylation and histone modification in cooperation with MET1. Mol Plant 7:1470-1485.

Kim KS & Lee KW (1992) Geminivirus-induced macrotubules and their suggested role in cell-to-cell movement. Phytopathology 82:664–669.

Kim SH, Macfarlane S, Kalinina NO, Rakitina DV, Ryabov EV, Gillespie T, Haupt S, Brown JW & Taliansky M (2007) Interaction of a plant virus encoded protein with the major nucleolar protein fibrillarin is required for systemic virus infection. Proc Natl Acad Sci USA 104:11115–11120.

Kleinow T, Tanwir F, Kocher C, Krenz B, Wege C & Jeske H (2009) Expression dynamics and ultrastructural localization of epitope-tagged *Abutilon mosaic virus* nuclear shuttle and movement proteins in *Nicotiana benthamiana* cells. Virology 391:212–220.

Kon T, Rojas MR, Abdourhamane IK & Gilbertson RL (2009) Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. J Gen Virol 90:1001-1013.

Kong LJ & Hanley-Bowdoin L (2002) A geminivirus replication protein interacts with a protein kinase and a motor protein that display different expression patterns during plant development and infection. Plant Cell 14:1817–1832.

Kong LJ, Orozco BM, Roe JL, Nagar S, Ou S, Feiler HS, Durfee T, Miller AB, Gruissem W, Robertson D & Hanley-Bowdoin L (2000) A geminivirus replication protein interacts with the retinoblastoma protein through a novel domain to determine symptoms and tissue specificity of infection in plants. EMBO J 19:3485–3495.

Kraberger S, Harkins G W, Kumari S G, Thomas J E, Schwinghamer M W, Sharman M, Collings DA, Briddon RW, Martin DP & Varsani A (2014) Evidence that dicot-infecting mastreviruses are particularly prone to inter-species recombination and have likely been circulating in Australia for longer than in Africa and the Middle East. Virology 444:282-291.

Kraft E, Bostick M, Jacobsen SE & Callis J (2008) ORTH/VIM proteins that regulate DNA methylation are functional ubiquitin E3 ligases. Plant J 56:704–715.

Krenz B, Deuschle K, Deigner T, Unseld S, Kepp G, Wege C, Kleinow T & Jeske H (2015) Early function of the *abutilon mosaic virus* AC2 gene as a replication brake. J Virol 89:3683–3699.

Krenz B, Neugart F, Kleinow T & Jeske H (2011) Self-interaction of *Abutilon mosaic virus* replication initiator protein (Rep) in plant cell nuclei. Virus Res 161:194–197.

Krenz B, Wege C & Jeske H (2010) Cell-free construction of disarmed Abutilon mosaic virus-based gene silencing vectors. J Virol Methods 169:129–137.

Kumar J, Kumar J, Singh SP & Tuli R (2014) Association of satellites with a Mastrevirus in natural infection: Complexity of Wheat dwarf India virus disease. J Virol 88:7093-7104.

Kumar V, Mishra SK, Rahman J, Taneja J, Sundaresan G, Mishra NS & Mukherjee SK (2015) *Mungbean yellow mosaic Indian virus* encoded AC2 protein suppresses RNA silencing by inhibiting Arabidopsis RDR6 and AGO1 activities. Virology 486:158–172.

Kumar Y, Hallan V & Zaidi A (2011) *Chilli leaf curl Palampur virus* is a distinct begomovirus species associated with a betasatellite. Plant Pathol 60:1040–1047.

Kumari P, Singh AK, Chattopadhyay B & Chakraborty S (2010) Molecular characterization of a new species of begomovirus and betasatellite causing leaf curl disease of tomato in India. Virus Res 152:19–29.

Kunik T, Mizrachy L, Citovsky V & Gafni Y (1999) Characterization of a tomato karyopherin alpha that interacts with the *Tomato yellow leaf curl virus* (TYLCV) capsid protein. J Exp Bot 50:731–732.

Kushwaha N, Sahu PP, Prasad M & Chakraborty S (2015) Chilli leaf curl infection highlights the differential expression of genes involved in protein homeostasis and defense in resistant chilli plants. Appl Microbiol Biotechnol 99(11):4757-70.

Lacatus & Sunter (2009) The Arabidopsis PEAPOD2 transcription factor interacts with geminivirus AL2 protein and the coat protein promoter. Virology 392(2):196-202.

Lai J, Chen H, Teng K, Zhao Q, Zhang Z, Li Y, Liang L, Xia R, Wu Y, Guo H & Xie Q (2009) RKP, a RING finger E3 ligase induced by BSCTV C4 protein, affects geminivirus infection by regulation of the plant cell cycle. Plant J 57:905–917.

Latham JR, Saunders K, Pinner MS & Stanley J (1997) Induction of plant cell division by *Beet curly top virus* gene C4. Plant J 11:1273–1283.

Laufs J, Jupin I, David C, Schumacher S, Heyraud-Nitschke F & Gronenborn B (1995) Geminivirus replication: genetic and biochemical characterization of Rep protein function, a review. Biochimie 77:765-773.

Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AM, Strahl BD, Patel DJ & Jacobsen SE (2013) Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature 498:385–389.

Law JA & Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11:204-220.

Law JA, Vashisht AA, Wohlschlegel JA & Jacobsen SE (2011) SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genet 7:e1002195.

Lazarowitz SG & Shepherd RJ (1992) Geminiviruses: Genome structure and gene function. Critical Reviews in Plant Sciences 11:327-349.

Lazarowitz SG, Wu LC, Rogers SG & Elmer JS (1992) Sequence specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. Plant Cell 4(7):799-809.

Lefeuvre P & Moriones E (2015) Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. Curr Opin Virol 10:14–19.

Lefeuvre P, Lett JM, Reynaud B, Martin DP (2007) Avoidance of protein fold disruption in natural virus recombinants. PLoS Pathog 3:e181.

Lewis JD & Lazarowitz SG (2010) *Arabidopsis* synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. Proc Natl Acad Sci USA 107:2491–2496.

Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS & Jacobsen SE (2006) An ARGONAUTE4-containing nuclear processing center colocalized with cajal bodies in *Arabidopsis thaliana*. Cell 126:93–106.

Li F, Huang C, Li Z & Zhou X (2014) Suppression of RNA silencing by a plant DNA virus satellite requires a host calmodulin-like protein to repress RDR6 expression. PLoS Pathog 10:e1003921.

Li F, Xu X, Huang C Gu Z Cao L Hu T, Ding M Li Z & Zhou X (2015) The AC5 protein encoded by *Mungbean yellow mosaic India virus* is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. New Phytologist 208:555-69.

Lin B, Behjatnia SAA, Dry IB, Randles JW & Rezaian MA (2003) High-affinity rep-binding is not required for the replication of a geminivirus DNA and its satellite. Virology 305:353–363.

Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I, Jenuwein T, Khorasanizadeh S & Jacobsen SE (2004) Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. EMBO J 23:4286–4296.

Liu H, Boulton MI & Davies JW (1997) *Maize streak virus* coat protein binds single- and double-stranded DNA *in vitro*. J Gen Virol 78:1265–1270.

Liu L, Saunders K, Thomas CL, Davies JW & Stanley J (1999) *Bean yellow dwarf virus* RepA, but not Rep, binds to maize retinoblastoma protein, and the virus tolerates mutations in the consensus binding motif. Virology 256:270–279.

Liu L, van Tonder T, Pietersen G, Davies JW & Stanley J (1997) Molecular characterization of a subgroup I geminivirus from a legume in South Africa. J Gen Virol 78:2113-2117.

Liu S, YuY, Ruan Y, Meyer D, Wolff M, Xu L, Wang N, Steinmetz A & Shen W-H (2007) Plant SET- and RING-associated domain proteins in hetero-chromatinization. Plant J 52:914–926.

Liu X, Yu CW, Duan J, Luo M, Wang K, Tian G, Cui Y & Wu K (2012) HDA6 directly interacts with DNA methyl transferase MET1 and maintains transposable element silencing in *Arabidopsis*. Plant Physiol 158:119–129.

Loor G, Zhang S, Zhang P, Toomey NL & Lee MYW (1997) Identification of DNA replication and cell cycle proteins that interact with PCNA. Nucleic Acids Res 25:5041–5046.

Lozano-Durán R, García I, Huguet S, Balzergue S, Romero LC & Bejarano ER (2012) Geminivirus C2 protein represses genes involved in sulphur assimilation and this effect can be counteracted by jasmonate treatment. Eur J Plant Pathol 134:49-59.

Lozano-Duran R, Rosas-Diaz T, Gusmaroli G, Luna AP, Taconnat L, Deng XW & Bejarano ER (2011) Geminiviruses subvert ubiquitination by altering CSN-mediated

derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in *Arabidopsis thaliana.* Plant Cell 23:1014–1032.

Luque A, Sanz-Burgos AP, Ramirez-Parra E, Castellano MM & Gutierrez C (2002) Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. Virology 302:83–94.

Malagnac F, Bartee L & Bender J (2002) An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. EMBO J 21:6842–6852.

Mendel M & Higa A (1970) Calcium-dependent bacteriophage DNA infection. J Mol Biol 53:159-162.

Mansoor S, Khan SH, Bashir B, Saeed M, Zafar Y, Malik KA, Briddon RW, Stanley J & Markham PG (1999) Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology 259:190-199.

Martin DP, Lemey P, Lott M, Moulton V, Posada D & Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462-2463.

Maruthi MN, Rekha AR, Mirza SH, Alam SN & Colvin J (2007) PCR-based detection and partial genome sequencing indicate high genetic diversity in Bangladeshi begomoviruses and their whitefly vector, *Bemisia tabaci*. Virus Genes 34:373-385.

McGarry RC, Barron YD, Carvalho MF, Hill JE, Gold D, Cheung E, Kraus WL & Lazarowitz SG (2003) A novel *Arabidopsis* acetyl transferase interacts with the geminivirus movement protein NSP. Plant Cell 15:1605–1618.

Melgarejo TA, Kon T, Rojas MR, Carrasco LP, Zerbini FM & Gilbertson RL (2013) Characterization of a New World monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. J Virol 87:5397-5413.

Milind P & Sushila K (2012) A hot way leading to healthy stay. Int Res J Pharmacy 3:21-25.

Mills-Lujan K & Deom CM (2010) Geminivirus C4 protein alters *Arabidopsis* development. Protoplasma 239;95–110.

Mills-Lujan K, Andrews DL, Chou C-W & Deom CM (2015) The roles of phosphorylation and SHAGGY-Like protein kinases in geminivirus C4 protein induced hyperplasia. PLoS ONE 10:e0122356.

Mishra MD, Raychaudhari SP & Jha A (1963) Virus causing leaf curl of chilli (*Capsicum annuum* L.). Indian J Microbiol 3:73-76.

Mori T, Takenaka K, Domoto F, Aoyama Y & Sera T (2012) Inhibition of binding of *Tomato yellow leaf curl virus* Rep to its replication origin by artificial zinc-finger protein. Mol Biotechnol 54:198-203.

Moshe A, Belausov E, Niehl A, Heinlein M, Czosnek H & Gorovits R (2016) The *Tomato yellow leaf curl virus* V2 protein forms aggregates depending on the cytoskeleton integrity and binds viral genomic DNA. Sci Rep 5:9967.

Mossi R & Hubscher U (1998) Clamping down on clamps and clamp loaders-the eukaryotic replication factor. C Eur J Biochem 254:209–216.

Mubin M, Amin I, Amrao I, Briddon RW & Mansoor S (2010) The hypersensitive response induced by the V2 protein of a monopartite begomovirus is countered by the C2 protein. Mol Plant Pathol 11(2):245–254.

Mubin M, Briddon RW & Mansoor S (2009) Complete nucleotide sequence of *chilli leaf curl virus* and its associated satellites naturally infecting potato in Pakistan. Arch Virol 154:365-368.

Muhire BM, Varsani A & Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS ONE 9:e108277.

Mullineaux PM, Boulton MI, Bowyer P, van Der Vlugt R, Marks M, Donson J & Davies JW (1988) Detection of a non-structural protein of MW11000 encoded by the virion DNA of *Maize streak virus*. Plant Mol Biol 11:57–66.

Nash TE, Dallas MB, Reyes MI, Buhrman GK, Ascencio-Ibanez JT & Hanley-Bowdoin L (2011) Functional analysis of a novel motif conserved across geminivirus rep proteins. J Virol 85:1182-1192.

Navot N, Pichersky E, Zeidan M, Zamir D & Czosnek H (1991) *Tomato yellow leaf curl virus*: a whitefly-transmitted geminivirus with a single genomic component. Virology 185:151-161.

Nawaz-ul-Rehman MS & Fauquet CM (2009) Evolution of geminiviruses and their satellites. FEBS Lett 583:1825-1832.

Nawaz-ul-Rehman MS, Nahid N, Mansoor S, Briddon RW & Fauquet CM (2010) Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. Virology 405:300-308.

Noris E, Vaira AM, Caciagli P, Masenga V, Gronenborn B & Accotto GP (1998) Amino acids in the capsid protein of *Tomato yellow leaf curl virus* that are crucial for systemic infection, particle formation, and insect transmission. J Virol 72:10050–10057.

Noueiry AO, Lucas WJ & Gilbertson RL (1994) Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. Cell 76:925-932.

Oshima K, Kakizawa S, Nishigawa H, Kuboyama T, Miyata S, Ugaki M & Namba S (2001) A plasmid of phytoplasma encodes a unique replication protein having both plasmid- and virus-like domains: clue to viral ancestry or result of virus/plasmid recombination? Virology 285:270–277.

Orozco BM, Gladfelter HJ, Settlage SB, Eagle PA, Gentry RN & Hanley-Bowdoin L (1998) Multiple cis-elements contribute to geminivirus origin function. Virology 242:346–356.

Orozco BM, Kong LJ, Batts LA, Elledge S & Hanley-Bowdoin L (2000) The multifunctional character of a geminivirus replication protein is reflected by its complex oligomerization properties. J Biol Chem 275:6114–6122.

Orozco BM, Miller AB, Settlage SB & Hanley-Bowdoin L (1997) Functional domains of a geminivirus replication protein. J Biol Chem 272:9840–9846.

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 224:390–404.

Palanichelvam K, Kunik T, Citovsky V & Gafni Y (1998) The capsid protein of *Tomato yellow leaf curl virus* binds cooperatively to single-stranded DNA. J Gen virol 79:2829-2833.

Pant V, Gupta D, Choudhury NR, Malathi VG, Varma A & Mukherjee SK (2001) Molecular characterization of the Rep protein of the blackgram isolate of Indian *Mungbean yellow mosaic virus*. J Gen Virol 82:2559–2567.

Paprotka T, Metzler V & Jeske H (2010) The first DNA 1-like a satellites in association with New World begomoviruses in natural infections. Virology 404:148–157.

Park J, Lee HJ, Cheon CI, Kim SH, Hur YS, Auh CK, Im KH, Yun DJ, Lee S & Davis KR (2011) The *Arabidopsis thaliana* homeobox gene ATHB12 is involved in symptom development caused by geminivirus infection. PLoS ONE 6:e20054.

Pascal E, Goodlove PE, Wu LC & Lazarowitz SG (1993) Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. Plant Cell 5:795-807.

Pascal E, Sanderfoot AA, Ward BM, Medville R, Turgeon R & Lazarowitz SG (1994) The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the cell nucleus. Plant Cell 6:995-1006.

Pasumarthy KK, Choudhury NR & Mukherjee SK (2010) *Tomato leaf curl Kerala virus* (ToLCKeV) AC3 protein forms a higher order oligomer and enhances ATPase activity of replication initiator protein (Rep/AC1). Virol J 7:128.

Pasumarthy KK, Mukherjee SK & Choudhury NR (2011) The presence of *tomato leaf curl Kerala virus* AC3 protein enhances viral DNA replication and modulates virus induced gene-silencing mechanism in tomato plants. Virol J 8:178.

Patil BL & Fauquet CM (2010) Differential interaction between cassava mosaic geminiviruses and geminivirus satellites. J Gen Virol 91:1871–1882.

Pilartz M & Jeske H (2003) Mapping of Abutilon mosaic geminivirus mini-chromosomes. J Virol 77:10808–10818.

Piroux N, Saunders K, Page A & Stanley J (2007) Geminivirus pathogenicity protein C4 interacts with *Arabidopsis thaliana* shaggy-related protein kinase AtSKeta, a component of the brassinosteroid signalling pathway. Virology 362:428-440.

Pitaksutheepong C, Vimoltat A, Nathwong B & Attathom S (2007) The N-terminal 62 amino acid residues of the coat protein of *Tomato yellow leaf curl Thailand virus* are responsible for DNA binding. J Gen Plant Pathol 73:72–75.

Pond SLK & Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22:1208-1222.

Pooma W & Petty IT (1996) *Tomato golden mosaic virus* open reading frame AL4 is genetically distinct from its C4 analogue in monopartite geminiviruses. J Gen Virol 77:1947–1951.

Priyadarshini CGP & Savithri HS (2009) Kinetics of interaction of *Cotton leaf curl Kokhran virus*-Dabawali (CLCuKV-Dab) coat protein and its mutants with ssDNA. Virology 386:427–437.

Priyadarshini CGP, Ambika MV, 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(11):e26929.

Pumplin N & Voinnet O (2013) RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat Rev Microbiol 11:745-760.

Qin S, Ward BM & Lazarowitz SG (1998) The bipartite geminivirus coat protein aids BR1 function in viral movement by affecting the accumulation of viral single stranded DNA. J Virol 72:9247–9256.

Qing L & Zhou X (2009) Trans-replication of, and competition between, DNA beta satellites in plants inoculated with *tomato yellow leaf curl china virus* and *tobacco curly shoot virus*. Phytopathology 99:716–720

Radhakrishnan GK, Splitter GA & Usha R (2008) DNA recognition properties of the cell-to-cell movement protein (MP) of soybean isolate of *Mungbean yellow mosaic India virus* (MYMIV-Sb). Virus Res 131:152–159.

Raghavan V, Malik PS, Choudhury NR & Mukherjee SK (2004) The DNA-A component of a plant geminivirus (*Mungbean yellow mosaic India virus*) replicates in budding yeast cells. J Virol 78:2405–2413.

Raj SK, Snehi SK, Khan MS, Tiwari AK & Rao GP (2010) First report of *Pepper leaf curl Bangladesh virus* strain associated with bitter gourd (*Momordica charantia* L.) yellow mosaic disease in India. Australasian Plant Dis Notes 5:14–16.

Raja P, Sanville BC, Buchmann RC & Bisaro DM (2008) Viral genome methylation as an epigenetic defense against geminiviruses. J Virol 82:8997–9007.

Raja P, Jackel JN, Li S, Heard IM & Bisaro DM (2013) Arabidopsis double-stranded RNA binding protein DRB3 participates in methylation-mediated defense against geminiviruses. J Virol 88(5):2611-22.

Rajeswaran R, Sunitha S, Shivaprasad PV, Pooggin MM, Hohn T & Veluthambi K (2007) The mungbean yellow mosaic begomovirus transcriptional activator protein transactivates the viral promoter-driven transgene and causes toxicity in transgenic tobacco plants. Mol Plant Microbe Interact 20:1545-1554.

Rana VS, Singh ST, Priya NG, Kumar J & Rajagopal R (2012) Arsenophonus GroEL interacts with CLCuV and is localized in midgut and salivary gland of whitefly B. tabaci. PLoS ONE 7:e42168.

Ranjan P, Singh AK, Kumar RV, Basu S & Chakraborty S (2014) Host specific adaptation of diverse betasatellites associated with distinct Indian tomato-infecting begomoviruses. Virus Genes 48:334-342.

Rentería-Canett I, Xoconostle-Cázares B, Ruiz-Medrano R & Rivera-Bustamante RF (2011) Geminivirus mixed infection on pepper plants: Synergistic interaction between PHYVV and PepGMV. Virology J 8:104.

Richter KS, Serra H, White CI & Jeske H (2016) The recombination mediator RAD51D promotes geminiviral infection. Virology 493:113-27.

Rizvi I, Choudhury NR & Tuteja N (2015) Insights into the functional characteristics of geminivirus rolling-circle replication initiator protein and its interaction with host factors affecting viral DNA replication. Arch Virol 160(2):375-387.

Roberts IM (1989) *Indian cassava mosaic virus*: Ultrastructure of infected cells. J Gen Virol 70:2729–2739.

Rocha CS, Castillo-Urquiza GP, Lima ATM, Silva FN, Xavier CA, Hora-Junior BT, Beserra-Junior JE, Malta AW, Martin DP Varsani A, Alfenas-Zerbini P, Mizubuti ES, Zerbini FM (2013) Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. J Virol 87:5784–5799.

Rocha CS, Santos AA, Machado JP & Fontes EP (2008) The ribosomal protein L10/QM-like protein is a component of the NIK-mediated antiviral signaling. Virology 38:165–169.

Rodriguez-Negrete E, Lozano-Duran R, Piedra-Aguilera A, Cruzado L, Bejarano ER & Castillo AG (2013) Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytologist 199:464–475.

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. J Virol 83:332–1340.

Rojas MR, Hagen C, Lucas WJ & Gilbertson RL (2005) Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. Ann Rev Phytopathol 43:361-394.

Rojas MR, Jiang H, Salati R, Xoconostle-Cazares B, Sudarshana MR, Lucas WJ & Gilbertson RL (2001) Functional analysis of proteins involved in movement of the monopartite begomovirus, *Tomato yellow leaf curl virus*. Virology 291:110–125.

Rojas MR, Noueiry AO, Lucas WJ & Gilbertson RL (1998) Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. Cell 95:105–113.

Romay G, Chirinos D, Geraud-Pouey F & Desvies C (2010) Association of an atypical alphasatellite with a bipartite New World begomovirus. Arch Virol 155:1843–1847.

Rosas-Díaz T, Macho AP, Beuzón CR, Lozano-Durán R & Bejarano ER (2016) The C2 protein from the Geminivirus *Tomato yellow leaf curl Sardinia virus* decreases sensitivity to jasmonates and suppresses jasmonate-mediated defences. Plants *5*(1):8.

Rothenstein D, Krenz B, Selchow O & Jeske H (2007) Tissue and cell tropism of *Indian cassava mosaic virus* (ICMV) and its AV2 (precoat) gene product. Virology 359:137-145.

Rouhibakhsh A, Haq QMI & Malathi VG (2011) Mutagenesis in ORF AV2 affects viral replication in *Mungbean yellow mosaic India virus*. J Biosci 36:329-340.

Rozas J, Sánchez-DelBarrio JC, Messeguer X & Rozas R (2003) DnaSP: DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496-2497.

Rybicki EP (1994) A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. Arch Virol 139:49-77.

Sabelli PA, Hoerster G, Lizarraga LE, Brown SW, Gordon-Kamm WJ & Larkinsa BA (2009) Positive regulation of minichromosome maintenance gene expression, DNA replication, and cell transformation by a plant retinoblastoma gene. Proc Natl Acad Sci USA 106:4042–4047.

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. J Gen Virol 88:2881-2889.

Saeed ST, Khan A, Kumar B, Ajayakumar PV & Samad A (2014) First report of *Chilli leaf curl India virus* infecting *Mentha spicata* (Neera) in India. Plant Dis 98:164.

Sambrook J & Russell D (2001) Molecular Cloning: a Laboratory Manual, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanchez-Duran MA, Dallas MB, Ascencio-Ibanez JT, Reyes MI, Arroyo-Mateos M, Ruiz-Albert J, Hanley-Bowdoin L & Bejarano ER (2011) Interaction between geminivirus replication protein and the SUMO-conjugating enzyme is required for viral infection. J Virol 85:9789–9800.

Sanderfoot AA & Lazarowitz SG (1995) Cooperation in viral movement: The Geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. Plant Cell 7:1185-1194.

Sanderfoot AA & Lazarowitz SG (1996) Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. Trends Cell Biol 6:353-358.

Sanz-Burgos AP & Gutiérrez C (1998) Organization of the cis-acting element required for wheat dwarf geminivirus DNA replication and visualization of a rep protein–DNA complex. Virology 243:119–129.

Sardo L, Lucioli A, Tavazza M, Masenga V, Tavazza R, Accotto GP & Noris E (2011) An RGG sequence in the replication-associated protein (Rep) of *Tomato yellow leaf curl Sardinia virus* is involved in transcriptional repression and severely impacts resistance in Rep-expressing plants. J Gen Virol 92:204–209.

Saunders K & Stanley J (1999) A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides*: evidence for interfamilial recombination between plant DNA viruses. Virology 264:142–152.

Saunders K & Stanley J (1995) Complementation of *African cassava mosaic virus* AC2 gene function in a mixed bipartite geminivirus infection. J Gen Virol 76:2287–2292.

Saunders K, Bedford ID, Briddon RW, Markham PG, Wong SM & Stanley J (2000) A unique virus complex causes Ageratum yellow vein disease. Proc Natl Acad Sci USA, 97:6890-6895.

Saunders K, Lucy A & Stanley J (1991) DNA forms of the geminivirus *African cassava mosaic virus* consistent with a rolling circle mechanism of replication. Nucleic Acids Res 19:2325–2330.

Saunders K, Norman A, Gucciardo S & Stanley J (2004) The DNA β satellite component associated with ageratum yellow vein disease encodes an essential pathogenicity protein (βC1). Virology 324:37–47.

Saunders K, Salim N, Mali VR, Malathi VG, Briddon RW, Markham PG & Stanley J (2002) Characterization of *Sri Lankan cassava mosaic virus* and *Indian cassava mosaic virus*: evidence for acquisition of a DNA B component by a monopartite begomovirus. Virology 293:63–74.

Saunders K, Wege C, Veluthambi K, Jeske H & Stanley J (2001) The distinct disease phenotypes of the common and yellow vein strains of *Tomato golden mosaic virus* are determined by nucleotide differences in the 3'-terminal region of the gene encoding the movement protein. J Gen virol 82:45-51.

Saze H, Mittelsten Scheid O & Paszkowski J (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat Genet 34:65-69.

Saze H, Tsugane K, Kanno T & Nishimura T (2012) DNA methylation in plants: Relationship to small RNAs and histone modifications, and functions in transposon inactivation. Plant Cell Physiol 53(5):766–784.

Selth LA, Dogra SC, Rasheed MS, Healy H, Randles JW & Rezaian MA (2005) A NAC domain protein interacts with *Tomato leaf curl virus* replication accessory protein and enhances viral replication. Plant Cell 17:311–325.

Senanayake DMJB, Jayasinghe JEARM, Shilpi S, Wasala SK & Mandal B (2013) A new begomovirus-betasatellite complex is associated with chilli leaf curl disease in Sri Lanka. Virus Genes 46:128–139.

Senanayake DMJB, Mandal B, Lodha S & Varma A (2007) First report of *Chilli leaf curl virus* affecting chilli in India. Plant Pathol 56:343.

Settlage SB, Miller AB, Gruissem W & Hanley-Bowdoin L (2001) Dual interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator. Virology 279:570–576.

Settlage SB, Miller AB & Hanley-Bowdoin L (1996) Interactions between geminivirus replication proteins. J Virol 70:6790-6795.

Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M & Koseki H (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912.

Sharma P & Ikegami M (2010) *Tomato leaf curl Java virus* V2 protein is a determinant of virulence, hypersensitive response and suppression of posttranscriptional gene silencing. Virology 396:85-93.

Sharma P, Ikegami M & Kon T (2010) Identification of the virulence factors and suppressors of posttranscriptional gene silencing encoded by *Ageratum yellow vein virus*, a monopartite begomovirus. Virus Res 149:19-27.

Shen Q, Bao M & Zhou X (2012) A plant kinase plays roles in defense response against geminivirus by phosphorylation of a viral pathogenesis protein. Plant Signal Behav 7:888–892.

Shen Q, Hu T, Bao M, Cao L, Zhang H, Song F, Xie Q & Zhou X (2016) Tobacco RING E3 ligase NtRFP1 mediates ubiquitination and proteasomal degradation of a Geminivirus-encoded βC1. Mol Plant 9(6):911-25.

Shen W, Dallas MB, Goshe MB & Hanley-Bowdoin L (2014) SnRK1 Phosphorylation of AL2 delays *Cabbage leaf curl virus* infection in *Arabidopsis*. J Virol 88(18):10598-612.

Shivaprasad PV, Akbergenov R, Trinks D, Rajeswaran R, Veluthambi K, Hohn T & Pooggin MM (2005) Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. J Virol 79:8149–8163.

Shung CY & Sunter G (2007) AL1-dependent repression of transcription enhances expression of Tomato golden mosaic virus AL2 and AL3. Virology 364:112–122

Silva SJC, Castillo-Urquiza GP, Hora-Junior BT, Assuncao IP, Lima GSA, Pio-Ribeiro G, Mizubuti ESG & Zerbini FM (2012) Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. Plant Pathol 61:457–467.

Simonne AH, Simonne EH, Eitenmiller RR, Mills HA & Green NR (1997) Ascorbic acid and provitamin A contents in unusually colored bell peppers (*Capsicum annuum* L.). J. Food Comp Anal 10:299-311.

Singh AK, Chattopadhyay B & Chakraborty S (2012) Biology and interactions of two distinct monopartite begomoviruses and beta-satellites associated with radish leaf curl disease in India. Virol J 9:43.

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. Appl Microbiol Biotechnol 100(9):4035-4049.

Singh DK, Malik PS, Choudhury NR & Mukherjee SK (2008) MYMIV replication initiator protein (Rep): roles at the initiation and elongation steps of MYMIV DNA replication. Virology 380:75–83.

Sivalingam PV & Varma A (2012) Role of betasatellite in the pathogenesis of a bipartite begomovirus affecting tomato in India. Arch Virol 157:1081-1092.

Soitamo AJ, Jada B & Lehto K (2012) Expression of geminiviral AC2 RNA silencing suppressor changes sugar and jasmonate responsive gene expression in transgenic tobacco plants. BMC Plant Biol 12:204.

Soleimani R, Matic S, Taheri H, Behjatnia SAA, Vecchiati M, Izadpanah K & Accotto GP (2013) The unconventional geminivirus *Beet curly top Iran virus*: satisfying Koch's postulates and determining vector and host range. Ann Appl Biol 162:174–181.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503-517.

Stanley J & Gay MR (1983) Nucleotide sequence of Cassava latent virus DNA. Nature 301:260–262.

Stanley J, Markham PG, Callis RJ & Pinner MS (1986) The nucleotide sequence of an infectious clone of the geminivirus *Beet curly top virus*. EMBO J 5:1761–1767.

Steinfeldt T, Finsterbusch T & Mankertz A (2006) Demonstration of nicking/joining activity at the origin of DNA replication associated with the Rep and Rep' proteins of porcine circovirus type 1. J Virol 80:6225–6234.

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. Proc Natl Acad Sci USA 88:8029–8033.

Stroud H, Greenberg MV, Feng S, Bernatavichute YV & Jacobsen SE (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. Cell 152:352–364.

Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ & Jacobsen SE (2014) Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. Nat Struct Mol Biol 21:64–72.

Sun Y-W, Tee C-S, Ma Y-H, Wang G, Yao X-M & Ye J (2015) Attenuation of histone methyltransferase KRYPTONITE mediated transcriptional gene silencing by Geminivirus. Sci Rep 5:16476.

Sunter G & Bisaro DM (1992) Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. Plant Cell 4:1321–1331.

Sunter G, Hartltz MD & Bisaro DM (1993) *Tomato golden mosaic virus* leftward gene expression: autoregulation of geminivirus replication protein. Virology 195:275–280.

Sunter G, Hertitz MD, Hormuzdi SG, Brough CL & Bisaro DM (1990) Genetic analysis of *tomato golden mosaic virus*: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication. Virology 179:69–77.

Sunter G, Stenger DC & Bisaro DM (1994) Heterologous complementation by geminivirus AL2 and AL3 genes. Virology 203:203–210.

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 285:59–70.

Suyal G, Mukherjee SK & Choudhury NR (2013a) The host factor RAD51 is involved in *Mungbean yellow mosaic India virus* (MYMIV) DNA replication. Arch Virol 158:1931–1941.

Suyal G, Mukherjee SK, Srivastava PS & Choudhury NR (2013b) *Arabidopsis thaliana* MCM2 plays role(s) in *Mungbean yellow mosaic India virus* (MYMIV) DNA replication. Arch Virol 158:981–992.

Suyal G, RanaVS, Mukherjee SK, Wajid S & Choudhury NR (2014) *Arabidopsis thaliana* NAC083 protein interacts with *Mungbean yellow mosaic India virus* (MYMIV) Rep protein. Virus Genes 48(3):486-93.

Tahir M, Haider MS & Briddon RW (2010) Chili leaf curl betasatellite is associated with a distinct recombinant begomovirus, *Pepper leaf curl Lahore virus*, in *Capsicum* in Pakistan. Virus Res 149:109–114.

Tamura K, Stecher G, Peterson D, Filipski A & Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725-2729.

Teng K, Chen H, Lai J, Zhang Z, Fang Y, Xia R, Zhou X, Guo H & Xie Q (2010) Involvement of C4 protein of *Beet severe curly top virus* (family *Geminiviridae*) in virus movement. PLoS ONE 5:e11280.

Trinks D, Rajeswaran R, Shivaprasad PV, Akbergenov R, Oakeley EJ, Veluthambi K, Hohn T & Pooggin MM (2005) Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. J Virol 79:2517–2527.

Unseld S, Höhnle M, Ringel M & Frischmuth T (2001) Subcellular targeting of the coat protein of African cassava mosaic geminivirus. Virology 286:373-383.

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. J Virol 78:9487–9498.

Varma A & Malathi VG (2003) Emerging geminivirus problems: A serious threat to crop production. Ann Appl Biol 142:145-164.

Varsani A, Navas-Castillo J, Moriones E, Hernandez-Zepeda C, Idris A, Brown J K, Zerbini F M & Martin D P (2014) Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. Arch Virol 159(8):2193-203.

Varsani A, Shepherd DN, Dent K, Monjane AL, Rybicki EP & Martin DP (2009) A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virol J 6:36.

Wang B, Li f, Huang C, Yang X, Qian Y, Xie Y & Zhou X (2014b) V2 of *Tomato yellow leaf curl virus* can suppress methylation-mediated transcriptional gene silencing in plants. J Gen Virol 95:225–230.

Wang H, Buckley KJ, Yang X, Buchmann RC & Bisaro DM (2003) Adenosine kinase inhibition and suppression of RNA silencing by Geminivirus AL2 and L2 proteins. J Virol 79(12):7410-7418.

Wang H, Hao L, Shung CY, Sunter G & Bisaro DM (2005) Adenosine kinase is inactivated by Geminivirus AL2 and L2 proteins. Plant Cell 15:3020–3032.

Wang Y, Dang M, Hou H, Mei Y, Qian Y & Zhou X (2014a) Identification of an RNA silencing suppressor encoded by a *Mastrevirus*. J Gen Virol 95(9):2082-2088.

Ward BM & Lazarowitz SG (1999) Nuclear export in plants. Use of geminivirus movement proteins for a cell-based export assay. Plant Cell 11:1267-1276.

Ward BM, Medville R, Lazarowitz SG & Turgeon R (1997) The geminivirus BL1 movement protein is associated with endoplasmic reticulum-derived tubules in developing phloem cells. J Virol 71:3726-3733.

Wartig L, Kheyr-Pour A, Noris E, De Kouchkovsky F, Jouanneau F, Gronenborn B & Jupin I (1997) Genetic analysis of the monopartite tomato yellow leaf curl Geminivirus: Roles of V1, V2, and C2 ORFs in viral pathogenesis. Virology 228:132-140.

Wassenegger M, Heimes S, Riedel L & Sanger HL (1994) RNA-directed de novo methylation of genomic sequences in plants. Cell 76:567–576.

Wierzbicki AT, Haag JR & Pikaard CS (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell 135:635–648.

Wierzbicki AT, Ream TS, Haag JR & Pikaard CS (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat Genet 41:630–634.

Woo HR, Pontes O, Pikaard CS & Richards EJ (2007) VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. Gene Dev 21:267–277.

Wu & Zhou XP (2005) Interaction between a tnanovirus-like component and the Tobacco curly shoot virus/satellite complex. Acta Biochimica et Biophysica Sinica 37:25-31.

Xie Q, Sanz-Burgos AP, Guo H, Garcia JA & Gutierrez C (1999) GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. Plant Mol Biol 39:647–656.

Xie Q, Suarez-Lopez P & Gutierrez C (1995) Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant DNA virus: requirement for efficient viral DNA replication. EMBO J 14:4073–4082.

Xie Y, Wu P, Liu P, Gong H & Zhou X (2010) Characterization of alphasatellites associated with monopartite begomovirus/betasatellite complexes in Yunnan, China. Virology J 7:178.

Yaakov N, Levy Y, Belausov E, Gaba V, Lapidot M & Gafni Y (2011) Effect of a single amino acid substitution in the NLS domain of *Tomato yellow leaf curl virus*-Israel (TYLCV-IL) capsid protein (CP) on its activity and on the virus life cycle. Virus Res 158:8–11.

Yadav RK & Chattopadhyay D (2011) Enhanced viral intergenic region-specific short interfering RNA accumulation and DNA methylation correlates with resistance against a geminivirus. Mol Plant Microbe Interact 24:1189–1197.

Yang LP, Fang YY, An CP, Dong L, Zhang ZH, Chen H, Xie Q & Guo HS (2012) C2-mediated decrease in DNA methylation, accumulation of siRNAs, and increase in

expression for genes involved in defense pathways in plants infected with *Beet severe curly top virus*. Plant J 73:910–917.

Yang SJ, Carter SA, Cole AB, Cheng NH & Nelson RS (2004) A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by *Nicotiana benthamiana*. Proc Natl Acad Sci USA 101:6297-6302.

Yang X, Baliji S, Buchmann RC, Wang H, Lindbo JA, Sunter G & Bisaro DM (2007) Functional modulation of the geminivirus AL2 transcription factor and silencing suppressor by self-interaction. J Virol 81:11972–11981.

Yang X, Xie Y, Raja P, Li S, Wolf JN, Shen Q, Bisaro DM & Zhou X (2011) Suppression of methylation-mediated transcriptional gene silencing by βC1-SAHH protein interaction during geminivirus betasatellite infection. PLoS Pathog 7:e1002329.

Yazdi HR, Heydarnejad J, Massumi H (2008) Genome characterization and genetic diversity of *Beet curly top Iran virus*: a geminivirus with a novel nonanucleotide. Virus Genes36:539–545

Ye J, Yang J, Sun Y, Zhao P, Gao S, Jung C, Qu J, Fang R & Chua N-H (2015) Geminivirus activates ASYMMETRIC LEAVES 2 to accelerate cytoplasmic DCP2-mediated mRNA turnover and weakens RNA silencing in *Arabidopsis*. PLoS Pathog 11(10):e1005196.

Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R & Chen X (2005) Methylation as a crucial step in plant microRNA biogenesis. Science 307:932–935.

Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL & Zilberman D (2013) The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell 153:193–205.

Zhang J, Dong J, Xu Y & Wu J (2012) V2 protein encoded by *Tomato yellow leaf curl China virus* is an RNA silencing suppressor. Virus Res 163:51-58.

Zhang J, Zhang X, Wang Y, Hou H & Qian Y (2012) Characterization of sequence elements from *Malvastrum yellow vein betasatellite* regulating promoter activity and DNA replication. Virology 9:1–9.

Zhang P, Vanderschuren H, Futterer J & Gruissem W (2005) Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. Plant Biotechnol J 3:385–397.

Zhang T, Xu X, Huang C, Qian Y, Li Z & Zhou X (2015) A novel DNA motif contributes to selective replication of geminivirus-associated betasatellite by a helper virus-encoded replication-related protein. J Virol 90(4):2077-2089.

Zhang W, Olson NH, Baker TS, Faulkner L, Agbandje-McKenna M, Boulton MI, Davies JW & McKenna R (2001) Structure of the Maize streak virus geminate particle. Virology 279:471–477.

Zhang X, Henderson IR, Lu C, Green PJ & Jacobsen SE (2007) Role of RNA polymerase IV in plant small RNA metabolism. Proc Natl Acad Sci USA 104:4536–4541.

Zhang Z, Chen H, Huang X, Xia R, Zhao Q, Lai J, Teng K, Li Y, Liang L, Du Q, Zhou X, Guo H & Xie Q (2011) BSCTV C2 Attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in *Arabidopsis*. Plant Cell 23(1):273-88.

Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ & Jacobsen SE (2014) Molecular mechanism of action of plant DRM *de* novo DNA methyltransferases. Cell 157:1050–1060.

Zhou X (2013) Advances in understanding begomovirus satellites. Annu Rev Phytopathol 51:357-381.

Zhou X, Liu Y, Calvert L, Munoz D, Otim-Nape GW, Robinson DJ & Harrison BD (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. J Gen Virol 78:2101-2111.

Zhou Y, Rojas MR, Park MR, Seo YS, Lucas WJ & Gilbertson RL (2011) Histone H3 interacts and colocalizes with the nuclear shuttle protein and the movement protein of a geminivirus. J Virol 85:11821–11832.

Zorzatto C, Machado JPB, Lopes KVG, Nascimento KJT, Pereira WA, Brustolini OJB, Reis PAB, Calil IP, Deguchi M, Sachetto-Martins G, Gouveia BC, Loriato VAP, Silva MAC, Silva FF, Santos AA, Chory J & Fontes EPB (2015) NIK1-mediated translation suppression functions as a plant antiviral immunity mechanism. Nature 52: 679-682.

Zrachya A, Glick E, Levy Y, Arzi T, Citovsky V & Gafni Y (2007) Suppressor of RNA silencing encoded by Tomato yellow leaf curl virus-Israel. Virology 358:159-165.

9. Appendix

9.1 Preparation of reagents, buffers and media

Buffer/Reagent/Media	Method of preparation
For Agarose gel electrophoresis	
50X Tris Acetate EDTA (TAE)	242g Tris base, 57.1 ml Glacial acetic acid, 100ml 0.5M EDTA was added and the volume was made up to 800 ml with distilled water. The pH was adjusted to 8.0 and the final volume was then made upto 1 litre using distilled water.
6X DNA loading dye	0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 30% (v/v) glycerol was dissolved in distilled water and stored at 4ºC.
Ethidium bromide (10mg/ml)	1g Ethidium bromide was added to 100 ml of distilled water and stirred using magnetic stirrer for several hours before transferring to a dark bottle and stored at room temperature.
For cloning	
Ampicillin (100mg/ml)	1 g of ampicillin was dissolved in 10 ml of double distilled water, filter sterilized through 0.22µ filter and aliquoted in 1.5ml tubes and stored at -20ºC.
Kanamycin (50mg/ml)	0.5 g of kanamycin was dissolved in 10 ml of double distilled water, filter sterilized through 0.22µ filter and aliquoted in 1.5ml tubes and stored at -20ºC.
Luria broth media	1g of bacto-tryptone, 0.5g of yeast extract and 1g of NaCl was dissolved in 95 ml of water and the pH was adjusted to 7.0 with 5N NaOH. The volume was made up to 100 ml and sterilized by autoclaving.

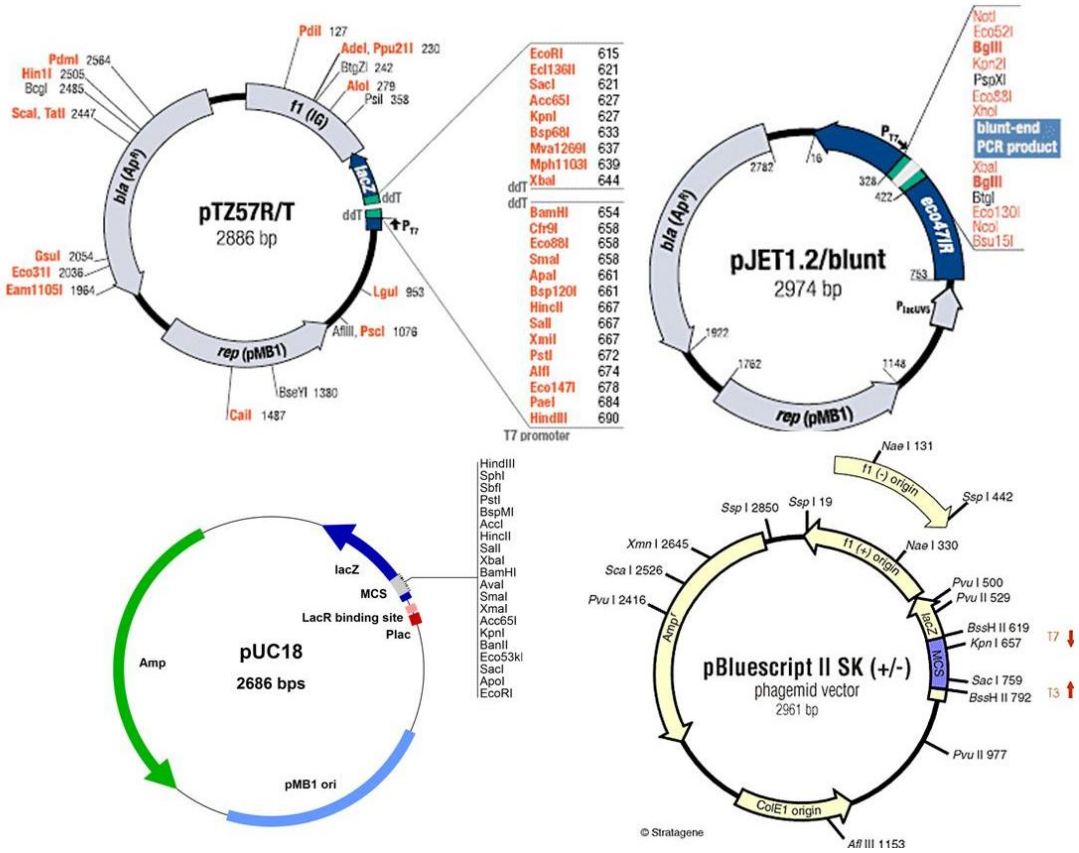
Luria agar media	1g of bacto-tryptone, 0.5g of yeast extract and 1g of NaCl was dissolved in 95 ml of water and the pH was adjusted to 7 with 5N NaOH. The volume was adjusted to 100 ml and 1.5 g of bacto-agar was added. It was then sterilized by autoclaving.
100mM $MgCl_2$	2.03g of $MgCl_2.6H_2O$ was dissolved in distilled water and the volume was adjusted to100ml and sterilized by autoclaving.
100mM $CaCl_2$	1.47g of $CaCl_2.2H_2O$ was dissolved in distilled water, volume adjusted to100ml and sterilized by autoclaving.
For plasmid isolation	
Solution I	It was prepared by adjusting adding: 2.5 ml of 1M Tris-HCl (25mM)-pH 8.0 4.5 ml of 20% Glucose (50mM) 2.0 ml of 0.5M EDTA (10mM) The final volume was made upto 100 ml with sterile distilled water.
Solution II	It was prepared freshly by mixing 2.0 ml of 10N NaOH (0.2N) 10.0 ml of 10% SDS The volume was made upto 100 ml with sterile distilled water.
Solution III (pH 4.8)	It was prepared by adding 5M Potassium acetate - 60ml Glacial acetic acid - 11.5ml Distilled water - 28.5 ml The concentration of potassium acetate found in the final solution was 3M. The solution was autoclaved and stored at 4°C.

For total DNA extraction	
Extraction buffer (100 ml)	The buffer was prepared by the addition of 1o ml of 1 M Tis-Cl pH 8.0, 10 ml of 0.5 M of EDTA pH 8.0, 10 ml of 5 N of NaCl and 500 µl of β-mercaptoethanol. The volume was made up to 100 ml using sterile distilled water.
Agroinfiltration	
Infiltration buffer	The infiltration buffer includes 10 mM MES buffer, 10 mM $MgCl_2$ and 2.25 mM acetosyringone
Southern blotting	
SSC (20X)	It was prepared by adding] 75.3 g NaC!, 88.2 g Sodium citrate (pH adjusted to 7.0 by adding 14 N HCI) in 1 L of distilled water.
Denaturation solution	It was prepared by adjusting final concentration of 1.5 M NaCI and 0.5 M NaOH in distilled water.
Depurination solution	It was prepared by adding 2.8 ml of HCI in 197.2 of sterile distilled water.
Neutralization solution	It was prepared by adding I.SM NaCI, 1.0 M Tris-HCI (pH 7.2).
Denhardt's solution (5OX)	It was prepared by adding 1 % (W/V) ficol, 1 % (W/V) polyvinylpyrrolidone, 1 % (W/V) BSA and was dissolved in distilled water.
Herring sperm DNA	100 mg of DNA was dissolved in 10 ml of sterile distilled H20 and mixed by vortexing, stored at -20°C.
Other commonly used solutions	
$CaCl_2$ (2.5M)	11 g of $CaCl_2.6H_2O$ was dissolved in a final volume of 20ml of distilled water. The solution was sterilized by passing it through a 0.22µ filter. It was then stored in 1ml aliquots at 4°C.

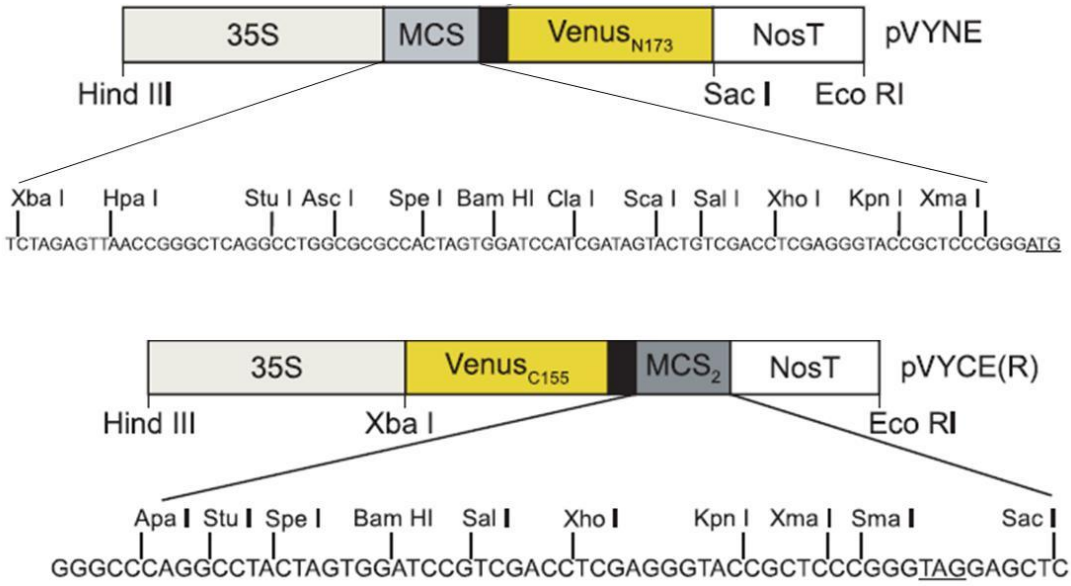
EDTA (0.5M, pH 8.0)	186.1 g of disodium EDTA.2H_2O was added to 800ml of water. It was stirred vigorously on a magnetic stirrer and pH was adjusted to 8.0 with NaOH. It was dispensed into aliquots and sterilized by autoclaving.
Ethanol (70%)	This was made by mixing 70 ml of absolute ethanol with 30 ml of sterile distilled water.
NaOH (10N)	400 g of NaOH pellets were added to 800ml of water with continuous stirring. The volume was adjusted to one litre with distilled water upon the complete dissolution of the pellets in water. The solution was stored at room temperature.
Potassium acetate (5M)	490.15 g of potassium acetate was taken and dissolved in 800 ml of distilled water. The volume was then made upto 1 litre with distilled water.
SDS (10% w/v)	100g of electrophoresis-grade SDS was dissolved in 900ml of H_2O. It was heated to 68°C and stirred with magnetic stirrer to assist dissolution. The volume was adjusted to 1 litre with sterile water and stored at room temperature.
Tris HCl	121.1g of Tris base was dissolved in 800ml of H_2O. The pH was adjusted to 8.0 by adding concentration HCl. The solution was allowed to cool at room temperature and the pH was adjusted. Then, volume of solution was adjusted to 1 litre with H_2O. This was then aliquoted and sterilized by autoclaving.
Chloroform : isoamyl alcohol (24:1)	The mixture was prepared by adding 24 ml of chloroform and 1ml of isoamyl alcohol.

9.2 Vectors used

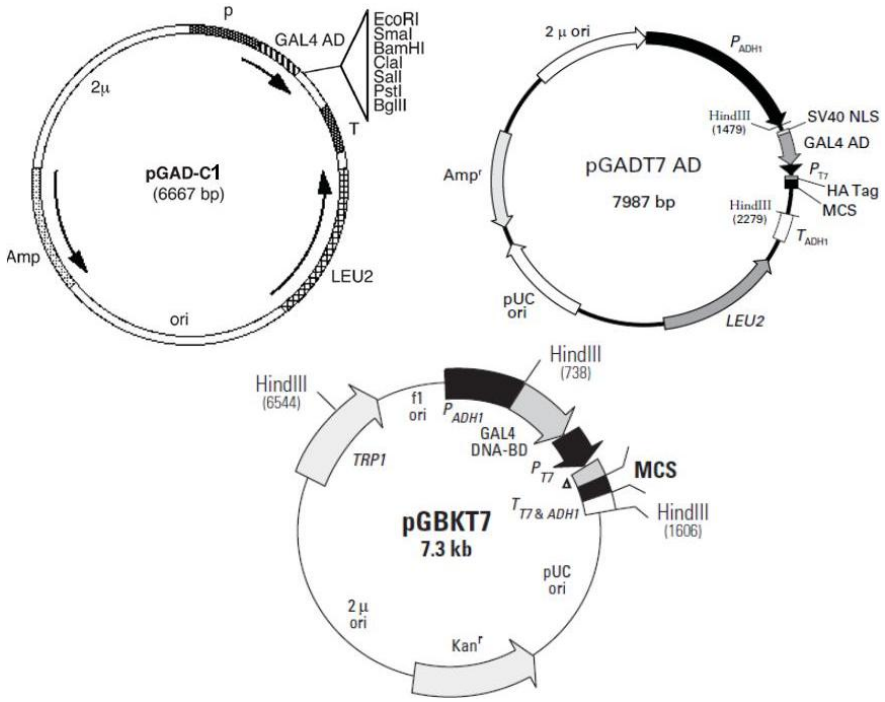
Cloning vectors



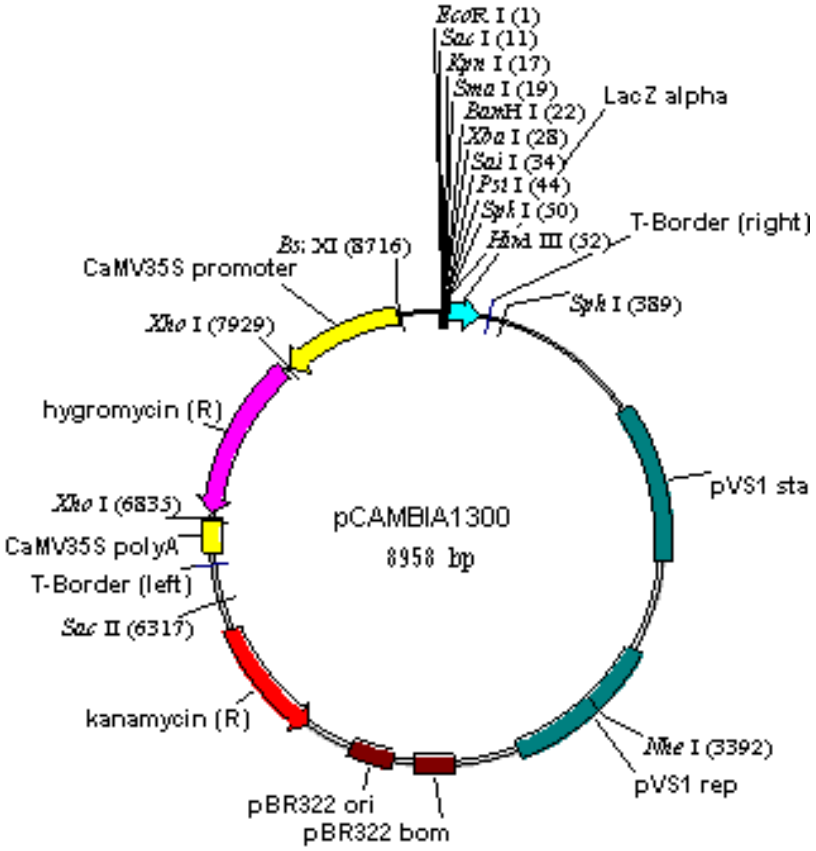
BiFC vectors

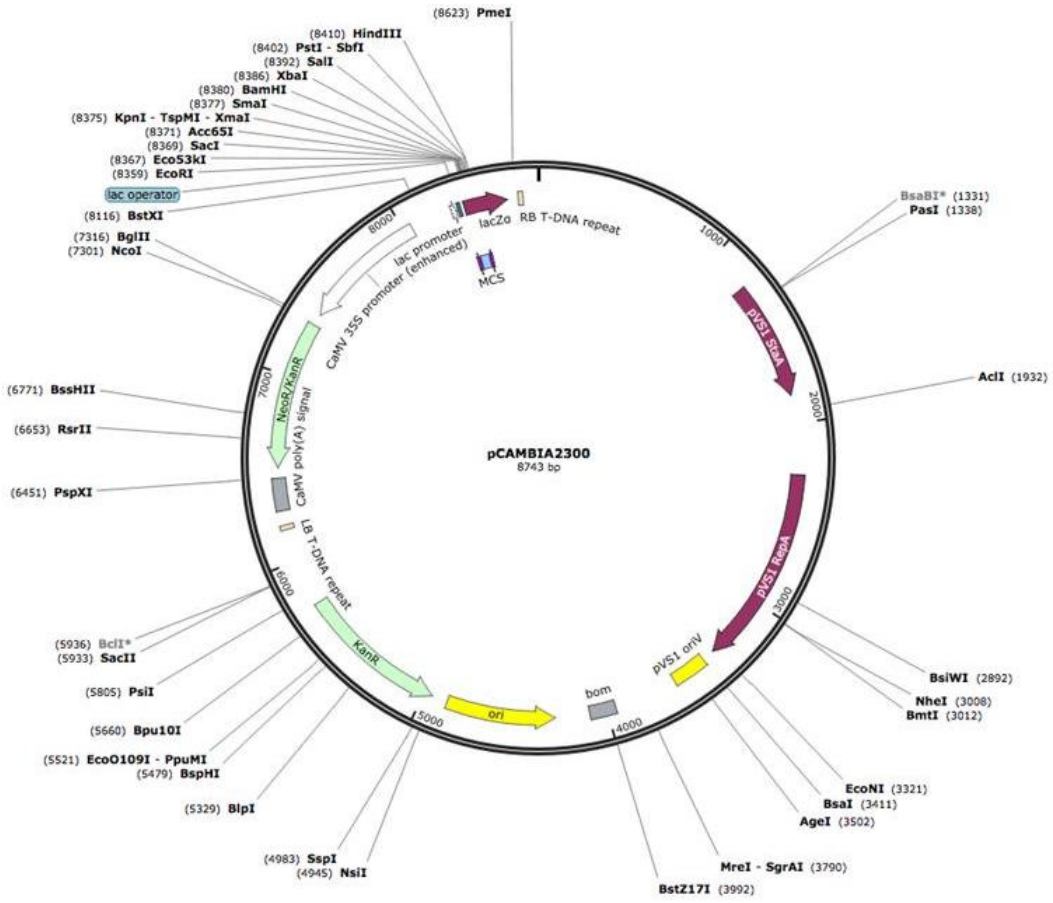


Yeast expression vectors

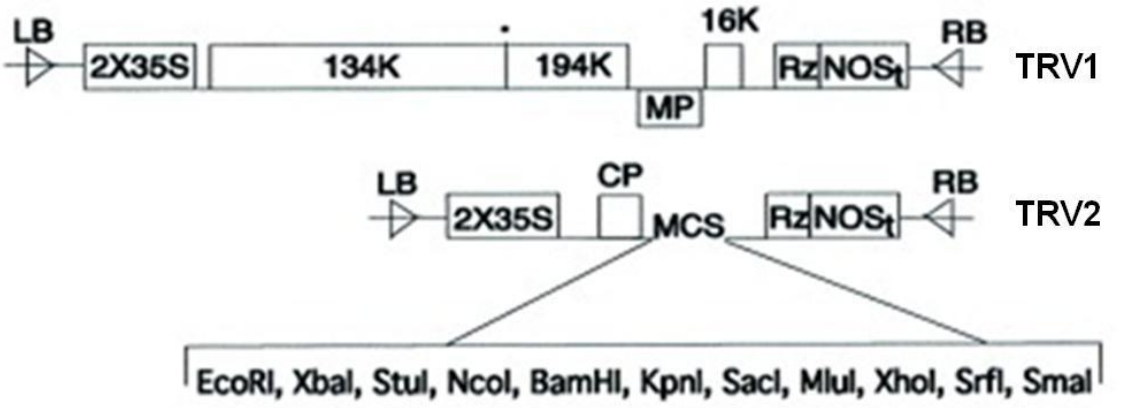


Plant expression vectors





TRV based VIGS vectors



10. List of publications

R Vinoth Kumar, Singh AK, Singh AK, Yadav T, Basu S, Kushwaha N, Chattopadhyay B, Chakraborty S (2015). Complexity of begomovirus and betasatellite populations associated with chilli leaf curl disease in India. *Journal of General Virology* 96: 3157-3172. DOI 10.1099/jgv.0.000254.

George B, Alam CM, R Vinoth Kumar, Prabu G, Chakraborty S (2015). Potential linkage between compound microsatellites and recombination in Geminiviruses: evidence from comparative analysis. *Virology* 482: 41-50. DOI 10.1016/j.virol.2015.03.003.

Ranjan P, Singh AK, R Vinoth Kumar, Basu S, Chakraborty S (2014). Host-specific adaptation of diverse betasatellites associated with distinct Indian tomato-infecting begomoviruses. *Virus Genes* 48: 334-342. DOI 10.1007/s11262-013-1031-y.

George B, R Vinoth Kumar, Chakraborty S (2014). Molecular characterization of *Chilli leaf curl virus* and satellite molecules associated with leaf curl disease of *Amaranthus* spp. *Virus Genes* 48: 397-401. DOI 10.1007/s11262-013-1027-7.

Ranjan P, R Vinoth Kumar, Chakraborty S (2013). Differential pathogenicity among *Tomato leaf curl Gujarat virus* isolates from India. *Virus Genes* 47: 524-531. DOI: 10.1007/s11262-013-0977-0.

R Vinoth Kumar, Sharma VK, Chattopadhyay B, Chakraborty S (2012). An improved plant regeneration and *Agrobacterium*–mediated transformation of red pepper (*Capsicum annuum* L.). *Physiology and Molecular Biology of Plants* 18(4): 357-364. DOI: 10.1007/s12298-012-0132-8.

R Vinoth Kumar, Singh AK, Chakraborty S (2012). A new monopartite begomovirus species, *Chilli leaf curl Vellanad virus*, and associated betasatellites infecting chilli in the Vellanad region of Kerala, India. *New Disease Reports* 25, 20. DOI: 10.5197/j.2044-0588.2012.025.020.

Journal of General Virology (2015), **96**, 3143–3158 **DOI** 10.1099/jgv.0.000254

Complexity of begomovirus and betasatellite populations associated with chilli leaf curl disease in India

R. Vinoth Kumar,[1]† Achuit Kumar Singh,[1]†‡ Ashish Kumar Singh,[1] Tribhuwan Yadav,[1,2] Saumik Basu,[1,3] Nirbhay Kushwaha,[1] Brotati Chattopadhyay[1] and Supriya Chakraborty[1]

Correspondence
Supriya Chakraborty
supriyachakrasls@yahoo.com

[1]Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi-110 067, India

[2]Department of Molecular Biophysics and Biochemistry, School of Medicine, Yale University, USA

[3]Department of Entomology, University of Nebraska, Lincoln, USA

Chilli, which encompasses several species in the genus *Capsicum*, is widely consumed throughout the world. In the Indian subcontinent, production of chilli is constrained due to chilli leaf curl disease (ChiLCD) caused by begomoviruses. Despite the considerable economic consequences of ChiLCD on chilli cultivation in India, there have been scant studies of the genetic diversity and structure of the begomoviruses that cause this disease. Here we report on a comprehensive survey across major chilli-growing regions in India. Analysis of samples collected in the survey indicates that ChiLCD-infected plants are associated with a complex of begomoviruses (including one previously unreported species) with a diverse group of betasatellites found in crops and weeds. The associated betasatellites neither enhanced the accumulation of the begomovirus components nor reduced the incubation period in *Nicotiana benthamiana*. The ChiLCD-associated begomoviruses induced mild symptoms on *Capsicum* spp., but both the level of helper virus that accumulated and the severity of symptoms were increased in the presence of cognate betasatellites. Interestingly, most of the begomoviruses were found to be intra-species recombinants. The betasatellites possess high nucleotide variability, and recombination among them was also evident. The nucleotide substitution rates were determined for the AV1 gene of begomoviruses (2.60×10^{-3} substitutions site^{-1} year^{-1}) and the βC1 gene of betasatellites [chilli leaf curl betasatellite (ChiLCB), 2.57×10^{-4} substitution site^{-1} year^{-1}; tomato leaf curl Bangladesh betasatellite (ToLCBDB), 5.22×10^{-4} substitution site^{-1} year^{-1}]. This study underscores the current understanding of Indian ChiLCD-associated begomoviruses and also demonstrates the crucial role of betasatellites in severe disease development in *Capsicum* spp.

Received 11 December 2014
Accepted 29 July 2015

INTRODUCTION

The family *Geminiviridae* includes a group of twinned icosahedral plant viruses containing circular, single-stranded DNA (ssDNA) and is divided into seven genera based on host range, vector transmission and genome organization: *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus* (Varsani *et al.*, 2014). Begomoviruses are widely distributed geographically, are transmitted through the arthropod vector *Bemisia tabaci*, and are further classified as monopartite with a DNA-A-like component or as bipartite containing DNA-A and DNA-B (Brown *et al.*, 2012). The majority of the monopartite begomoviruses have been reported from the 'Old World' while the begomoviruses of the 'New World' are mostly bipartite, along with a true monopartite begomovirus (Melgarejo *et al.*, 2013; Nawaz-ul-Rehman & Fauquet, 2009). Generally, DNA-A encodes

†These authors contributed equally to this paper.

‡Present address: School of Life Sciences, Central University of Gujarat, Gandhinagar, Gujarat, India.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are EF190215–EF190217, EF194765, EU582020, HM007094–HM007121, JN663846–JN663878, JN704344, KF471060 and KP235538–KP235543.

Seven supplementary figures and six supplementary tables are available with the online Supplementary Material.

000254 © 2015 The Authors *Printed in Great Britain*
Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

six ORFs necessary for virus replication, transcription activation and encapsidation, while the proteins encoded by the DNA-B component facilitate cell–cell and nucleocytoplasmic trafficking of the viral genome. In addition, intracellular movement is controlled by proteins encoded in the DNA-A-like component of monopartite begomoviruses (Hanley-Bowdoin *et al.*, 2013).

Begomoviruses have been reported to be associated with ssDNA satellite molecules and/or satellite-like molecules, namely betasatellites and alphasatellites, respectively (Mansoor *et al.*, 1999; Nawaz-ul-Rehman & Fauquet, 2009; Saunders *et al.*, 2000). The majority of the Old World monopartite begomoviruses co-exist with betasatellites. Betasatellites typically contain a satellite conserved region (SCR), an adenine-rich region and a βC1 ORF. The βC1 ORF has been characterized as a pathogenicity determinant that suppresses host antiviral silencing and is involved in virus movement (Cui *et al.*, 2005; Li *et al.*, 2014; Saunders *et al.*, 2004; Yang *et al.*, 2011). A class of autonomously replicating satellite-like molecules identified as alphasatellites require helper begomoviruses for their intra- and intercellular movement and are reported to ameliorate symptoms in a few cases (Idris *et al.*, 2011; Nawaz-ul-Rehman *et al.*, 2010).

Capsicum spp. are cultivated as vegetables and spice crops throughout the world. Chilli leaf curl disease (ChiLCD) causes severe crop loss in the tropical and subtropical regions of the Indian subcontinent and has emerged as a major factor limiting chilli cultivation in this part of the world (Chattopadhyay *et al.*, 2008; Mishra *et al.*, 1963; Senanayake *et al.*, 2007, 2013; Varma & Malathi, 2003). The symptoms of ChiLCD-infected plants are leaf curling, leaf rolling, puckering and reduced leaf size; in severe cases, the infected plants remain stunted and fail to bear any fruit, resulting in complete crop loss. Although the first incidence of ChiLCD was reported in the 1960s, its association with begomoviruses was confirmed only recently (Chattopadhyay *et al.*, 2008; Khan *et al.*, 2006; Kumar *et al.*, 2012; Senanayake *et al.*, 2007). During the past few years, monopartite begomoviruses and betasatellites associated with ChiLCD have spread to major chilli-growing regions of the Indian subcontinent. To date, in India, chilli leaf curl virus (ChiLCV), chilli leaf curl India virus (ChiLCINV), chilli leaf curl Vellanad virus (ChiLCVV), tomato leaf curl Joydebpur virus and tomato leaf curl New Delhi virus (ToLCNDV) are known to be associated with ChiLCD (Khan *et al.*, 2006; Kumar *et al.*, 2011, 2012; Senanayake *et al.*, 2007; Shih *et al.*, 2007).

Despite the severe losses due to ChiLCD, a systematic study of the begomoviruses associated with ChiLCD in different chilli-growing regions of India is lacking. Here, we have carried out a comprehensive survey to analyse the genetic diversity and the phylogeographical distribution of ChiLCD-associated begomoviruses. This study reveals the association of seven distinct begomovirus species with five different groups of betasatellites with ChiLCD. Most of the ChiLCD-associated begomoviruses and betasatellites were recombinants. The isolated viral components could infect *Nicotiana benthamiana* and *Capsicum annuum*. Furthermore, our study also highlighted the indispensable role of betasatellites in inducing severe leaf curl disease on *Capsicum* spp.

RESULTS

Cloning of ChiLCD-associated viral components across India

Leaf samples from chilli plants exhibiting leaf curling, crinkling and stunting of plants were collected from 28 major chilli-growing regions of India (Fig. S1, Table S1, available in the online Supplementary Material). The symptoms were observed to vary among the samples surveyed (Table S1). Total DNA was extracted from symptomatic leaves, and PCR amplification with abutting primers for DNA-A, DNA-B and betasatellite indicated the association of begomoviruses and betasatellites with the field samples (data not shown). PCR-based amplification of full-length viral components and/or rolling circle amplification (RCA) using φ29 polymerase yielded DNA fragments of either $\sim$2.8 kb or $\sim$1.3 kb, which were cloned (Table 1). Clones from respective samples were digested with several restriction enzymes to identify the diversity based on RFLP and were subsequently partially sequenced. We obtained the complete nucleotide sequences of 41 molecules of $\sim$2.8 kb and 33 molecules of $\sim$1.3 kb in size from samples representing the 28 chilli-growing regions. Cloning method and accession numbers are provided in Table 1.

Diversity and phylogenetic analysis of ChiLCD-associated begomoviruses

Our analysis showed that 35 out of the 41 molecules were $\sim$2.8 kb in size with a DNA-A/DNA-A-like begomovirus genome arrangement, whereas six molecules had a DNA-B-like genomic organization. Betasatellites were present in all the samples irrespective of whether the begomovirus genome detected was either monopartite or bipartite (Table 1).

Based on 91 % nucleotide sequence identity as the species demarcation threshold for DNA-A (Brown *et al.*, 2015) and by pairwise sequence comparison, these DNA-A/DNA-A-like isolates were categorized into seven distinct species including one newly identified species (HM007119). When compared with the genomic sequences available in GenBank, the majority of these sequences had high similarity with previously reported begomoviruses associated with either chilli or tomato (Fig. S2a).

The viral sequences represented in clones O1A and K1A showed maximum identities of 95.8 % and 94.8 % with a 'Croton' strain of papaya leaf curl virus (PaLCuV-Cro;

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Table 1. Details of viral components identified from ChiLCD-infected *Capsicum* spp. in India

Location		Year	GPS coordinates	DNA-A/DNA-A-like*				DNA-β				DNA-B				Nature of genome
State	Area			Clone name	Accession number	Cloning method	Isolate†	Clone name	Accession number	Cloning method	Isolate†	Clone name	Accession number	Cloning method	Isolate†	
Andhra Pradesh	Guntur	2009	16.30° N 80.45° E	A1A	HM007100	RCA	ChiLCV-Gun	A1β	JN663859	PCR	ToLCBDB					Monopartite
Assam	Jorehat	2009	26.44° N 94.10° E	A2A	JN663861	PCR	ChiLCV-PK	A2β	JN663862	PCR	ToLCJoB					Monopartite
Bihar	Chhapra	2010	25.78° N 84.72° E	B1A	JN663852	RCA	ChiLCV-Chi[1]	B1β	JN663854	RCA	ToLCBDB[1]					Monopartite
				B2A	JN663853	RCA	PepLCBV-BD[1]	B2β	JN663855	RCA	ToLCBDB[1]					
	Patna	2008	25.36° N 85.7° E	B3A	HM007117	RCA	ChiLCV	B3β	HM007118	RCA	ToLCBDB					Monopartite
Goa	Keri	2011	26.67° N 73.18° E	G1A	KP235539	PCR	ChiLCV-Gun	G1β	JN663858	PCR	RaLCB					Monopartite
Gujarat	Ahmedabad	2009	23.03° N 72.58° E	G2A	JN663846	RCA	ChiLCV-IN	G2β	JN663847	PCR	ToLCBDB					Monopartite
Haryana	Pataudi	2007	28.35° N 76.75° E	H1A	HM007116	PCR	ChiLCV-IN	H1β	EU582020	RCA	ToLCBDB					Monopartite
Himachal Pradesh	Palampur	2009	32.11° N 76.53° E	H2A	JN663870	PCR	ChiLCV-PK[2]	H2β	JN663872	PCR	ToLCRnB[2]	H2B	JN663871	PCR	ToLCNDV[2]	Monopartite & bipartite
				H3A	KP235540	PCR	ToLCNDV[2]									
Karnataka	Bengaluru	2010	13.81° N 77.29° E	K1A	HM007094	RCA	ToLCV-Ban	K1β	JN663849	PCR	ToLCBDB	K1B	JN663848	PCR	ToLCNDV	Monopartite
	Tumkur	2008	13.20° N 77.08°E	K2A	HM007120	RCA	ToLCNDV	K2β	JN663875	RCA	ToLCBDB					
Kerala	Vellanad	2008	8.34° N 77.30° E	K3A	HM007121	RCA	ChiLCVV[3]	K3β	JN663877	PCR	RaLCB[3]					Monopartite
								K4β	JN663876	PCR	ToLCBDB[3]					
Madhya Pradesh	Jabalpur	2008	23.10° N 79.59° E	M1A	HM007101	RCA	PepLCBV-PK	M1β	JN663860	PCR	ToLCBDB					Monopartite
Maharashtra	Nagpur	2009	21.8° N 79.5° E	M2A	JN663864	PCR	PepLCBV-IN[4]	M2β	JN663863	PCR	ToLCJoB[4]					Monopartite
				M3A	JN663865	PCR	ChiLCV-Gun[4]									
	Pune	2009	18.31° N 73.51° E	M4A	KP235542	PCR	ToLCNDV	M4β	JN663878	PCR	CroYVMB	M4B	KP235543	PCR	ToLCNDV	Bipartite
New Delhi	New Delhi	2010	28.36° N 77.12° E	N1A	HM007113	RCA	ToLCNDV[5]	N1β	JN663868	PCR	ToLCBDB[5]	N2B	JN663867	PCR	ToLCNDV[5]	Monopartite & bipartite
				N2A	JN663866	RCA	ChiLCV-IN[6]	N2β	JN663869	PCR	ToLCBDB[6]	N3B	KP235538	PCR	ToLCGV[5]	
Orissa	Bhubaneswar	2009	20.14° N 85.50° E	O1A	JN663850	RCA	PaLCuV-Cro	O1β	JN663851	PCR	ToLCRnB					Monopartite

Table 1. cont.

Location		Year	GPS coordinates	DNA-A/DNA-A-like*				DNA-β				DNA-B				Nature of genome
State	Area			Clone name	Accession number	Cloning method	Isolate†	Clone name	Accession number	Cloning method	Isolate†	Clone name	Accession number	Cloning method	Isolate†	
Rajasthan	Jodhpur	2009	26.18° N 73.04° E	R1A	HM007104	RCA	ChiLCV-IN	R1β	HM007105	RCA	ToLCBDB					Monopartite
Tamil Nadu	Coimbatore	2008	11.0° N 76.58° E	T1A	HM007096	RCA	PepLCBV-IN[7]	T1β	JN663856	PCR	ToLCBDB[7]					Monopartite
								T2β	JN663857	PCR	CroYVMB[7]					
	Salem	2008	11.44° N 78.4° E	T2A	HM007119	RCA	ChiLCSV[8]	T3β	JN663873	PCR	RaLCB[8]	T2B	KF471060	PCR	ToLCNDV[8]	Monopartite & bipartite
				T3A	KP235541	PCR	ToLCNDV[8]	T4β	JN663874	PCR	CroYVMB[8]					
Uttar Pradesh	Jaunpur	2007	25.44° N 82.41° E	U1A	HM007102	PCR	ChiLCV	U1β	HM007103	RCA	ToLCJoB					Monopartite
	Ghazipur	2010	25.35° N 83.34° E	U2A	HM007097	RCA	PepLCBV-BD[9]	U2β	HM007099	PCR	ToLCBDB[9]					Monopartite
				U3A	HM007098	RCA	ChiLCV-Gaz[9]									
	Kanpur	2008	26.28° N 80.21° E	U4A	HM007106	RCA	PepLCBV-Kan	U4β	HM007107	RCA	ToLCBDB					Monopartite
	Noida	2007	28.57° N 77.32° E	U5A	HM007114	PCR	ChiLCV-IN	U5β	HM007115	RCA	ToLCBDB					Monopartite
	Varanasi	2006	25.20° N 83.0° E	U6A	EF190217	PCR	ChiLCV-PK	U6β	EF190215	PCR	ToLCBDB					Monopartite
West Bengal	Baruipur	2008	22.20° N 88.26° E	W1A	HM007095	PCR	ChiLCV-Chi	W1β	JN704344	PCR	ToLCJoB					Monopartite
	Kalyani	2006	22.59° N 88.29° E	W2A	EF194765	PCR	ChiLCV	W2β	EF190216	RCA	ToLCJoB					Monopartite
	Kolkata	2008	22.34° N 88.22° E	W3A	HM007108	RCA	ChiLCV[10]	W3β	HM007110	RCA	ToLCJoB[10]					Monopartite
				W4A	HM007109	RCA	ChiLCV[10]									
	Magrahat	2007	21.41° N 88.4° E	W5A	HM007111	RCA	PepLCBV-BD	W5β	HM007112	RCA	ToLCJoB					Monopartite

*DNA-A-like and DNA-A refer to monopartite and bipartite begomoviruses, respectively.
†Clones followed by the same numbers (in superscript) indicate that they were obtained from the same plant in that particular location.

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

AJ507777) and a 'Bangalore' strain of tomato leaf curl virus (ToLCV-Ban; FJ514798), respectively, and were thus identified as isolates of the aforesaid respective species (Fig. S2a). With 92.8–95.8 % nucleotide identity among them, B2A, M1A, U2A and W5A were classified as isolates of the species *Pepper leaf curl Bangladesh virus* (PepLCBV), based on their nucleotide identity (93.1–96.8 %) with the isolates from either Bangladesh or Pakistan, and formed a separate clade in the phylogenetic tree (Fig. 1a). Since U4A shared maximum identity of 91.9 % with PepLCBV-PK (DQ116881), and M2A and T1A displayed a maximum identity between 90.3 % and 92.3 % with PepLCBV (AM404179), they were, according to ICTV guidelines, considered to be new strains of the respective species and were named 'PepLCBV-IN' and 'PepLCBV-Kan', respectively (Fig. S2a). The complete nucleotide sequences of H3A, K2A, M4A, N1A and T3A were found to have 94.2–95.1 % sequence identity with ToLCNDV (AJ875157). Thus, they were considered as isolates of ToLCNDV and clustered as a separate clade in the phylogenetic tree (Fig. 1a). The sequences of B3A, U1A, W2A, W3A and W4A shared maximum nucleotide identity between 94 % and 96.9 % with ChiLCV (AJ875159). The sequences of five isolates (G2A, H1A, N2A, R1A and U5A) displayed 96.3–96.9 % nucleotide identity with the 'Indian' strain of ChiLCV (ChiLCV-IN; GU136803)

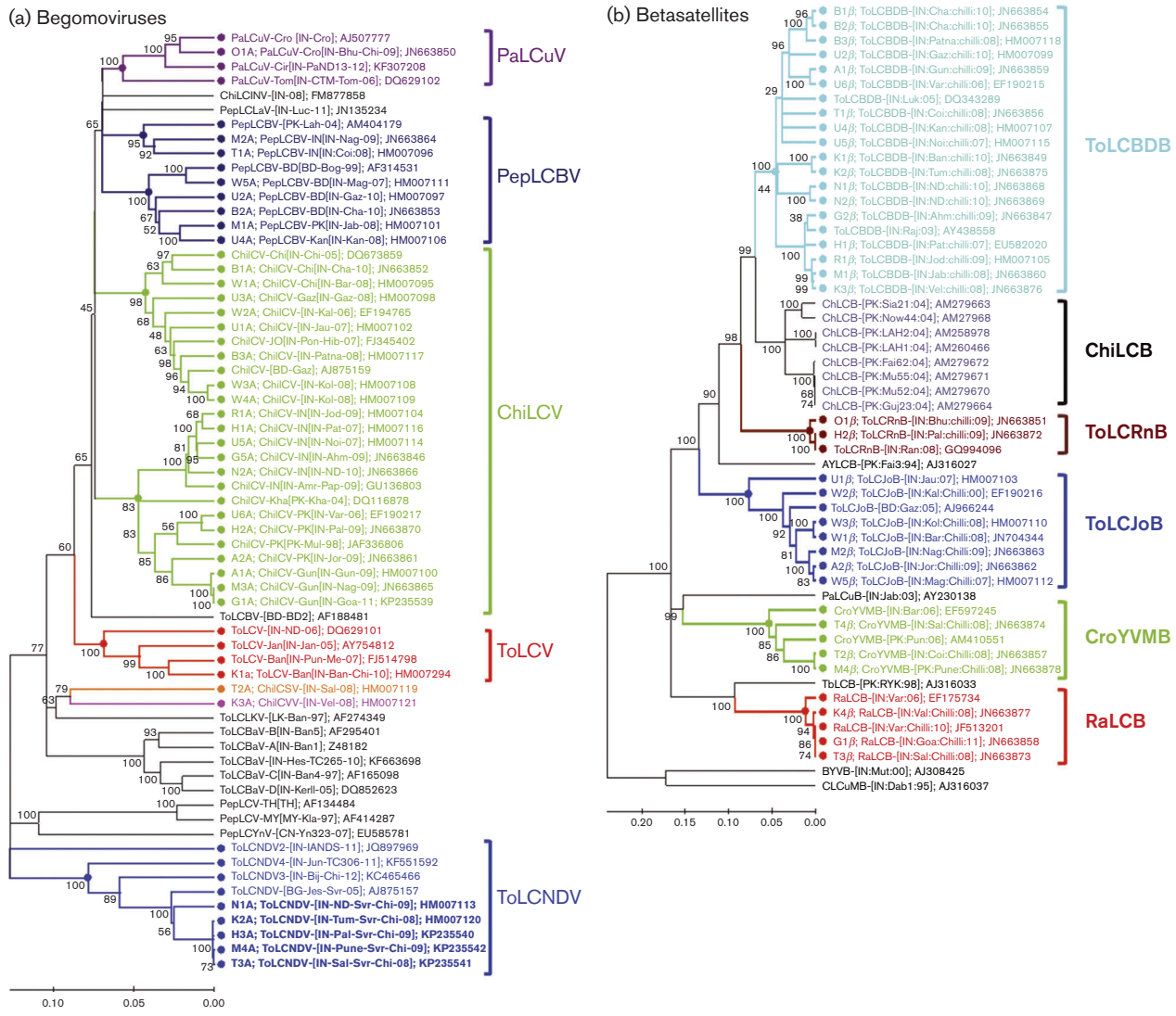


Fig. 1. Neighbour-joining phylogenetic dendrogram of ChiLCD-associated begomoviruses (a) and betasatellites (b) obtained in this study (indicated in bold type) with the most closely related representative genomes available in the database. GenBank accession numbers for all these sequences are shown. Prior to tree generation, all the sequences were aligned using CLUSTAL W. Percentage bootstrap values of 1000 replicates are given at nodes. Major clades in the same colour belong to a particular species, as indicated at the right side of each tree. The scale bars represent the rate of nucleotide substitutions per site.

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

(Fig. S2a). The sequences of A2A, H2A and U6A shared 94.8–95.5 % nucleotide identity with the 'Pakistan' strain of ChiLCV (ChiLCV-PK; AF336806), and 95.9–98.6 % identity among them. Since U3A shared maximum identity (92.8 %) with ChiLCV (AJ875159), according to ICTV guidelines, it was considered as a new strain of ChiLCV (ChiLCV-Gaz). Likewise, A1A, G1A and M3A were found to share a maximum identity of 91.9 % with ChiLCV-PK (AF336806) and hence considered as new strains of ChiLCV (Fig. S2a). Consistent with the nucleotide identities, all the strains/isolates belonging to ChiLCV formed a distinct clade in the phylogenetic dendrogram (Fig. 1a). Results indicated that ChiLCVs were not confined to any single geographical territory of India.

Based on the species demarcation criteria, the isolate from the Vellanad region belongs to chilli leaf curl Vellanad virus (ChiLCVV). A sequence from the Salem region exhibited a nucleotide identity of 82.2 % with tomato leaf curl Sri Lanka virus (AF274349) and we propose the name 'chilli leaf curl Salem virus' (ChiLCSV) for the species. Also note that ChiLCVV and ChiLCSV occupy unique positions in the phylogenetic dendrogram (Fig. 1a).

All five DNA-B sequences identified with these begomoviruses, i.e. H2B, K1B, M4B, N2B and T2B, were 2696 nt in length (Table 1), and shared >99 % nucleotide identity with ToLCNDV DNA-B (U15017). Hence, these DNA-Bs are considered as components of the isolates of ToLCNDV. Similarly, a sequence (N3B) isolated from New Delhi was shown to have maximum sequence identity with tomato leaf curl Gujarat virus (AY190291) DNA-B component. In the phylogenetic analysis, all these isolates grouped with their respective DNA-A (Fig. S3).

Distribution and phylogenetic relationship among betasatellites

Betasatellites 1350–1379 nt in size were isolated from the samples that harboured either monopartite (DNA-A-like) or bipartite (DNA-A and DNA-B) begomoviruses. According to the 78 % identity of species demarcation threshold for betasatellites (Briddon *et al.*, 2008) and pairwise sequence identity, these betasatellites may be broadly grouped as croton yellow vein mosaic betasatellite (CroYVMB), radish leaf curl betasatellite (RaLCB), tomato leaf curl Bangladesh betasatellite (ToLCBDB), tomato leaf curl Joydebpur betasatellite (ToLCJoB) and tomato leaf curl Ranchi betasatellite (ToLCRnB) (Fig. S2b). The sequences of G1β, K4β and T3β showed >99 % identity with RaLCB (EF175734), whereas those of M4β, T2β and T4β shared maximum identities of 90.3–91.5 % with CroYVMB (AM410551). All these weed-associated betasatellites were obtained from samples collected from southern India (Table 1). In the phylogenetic dendrogram, CroYVMB-, RaLCB-, tomato- and chilli-associated betasatellite clades were distinctly placed (Fig. 1b).

Of the 33 betasatellites isolated here, 27 were grouped into three clusters: ToLCJoB, ToLCBDB and ToLCRnB. The ToLCRnB cluster comprised the H2β and O1β sequences, sharing maximum sequence identity of 91.5–99.4 % with ToLCRnB (GQ994096), while the isolates A2β, M2β, U1β, W1β, W2β, W3β and W5β displayed 93–94 % identity to ToLCJoB (AJ966244) (Fig. S2b). The remaining 18 isolates shared 88.4–98.2 % nucleotide sequence identity with ToLCBDB (DQ343289, AY438558). The nucleotide sequence identity among ToLCJoBs ranged between 80.6 % and 99.6 %, whereas it ranged between 87.7 % and 99.7 % for ToLCBDB isolates. In the phylogenetic tree, ToLCBDB and ToLCJoB were grouped as distinct clades separated from chilli leaf curl betasatellite (ChiLCB) reported from Pakistan (Fig. 1b). Betasatellites identified from Assam and West Bengal belonged to ToLCJoB, but ToLCBDBs were distributed throughout the country.

Detection of recombination among ChiLCD-associated begomoviruses and betasatellites

Recombination is known to play a major role in the emergence and evolution of geminiviruses (George *et al.*, 2015; Lefeuvre & Moriones, 2015; Padidam *et al.*, 1999). In recombination analysis of begomoviral components, we detected significant evidence of recombination in the vast majority of isolates (26/35) (Table 2). Among them, nine isolates showed evidence of recombination events in both the AC1 and AV1 regions, whereas in three isolates, the breakpoints were located within the common region (CR) and AV1 (Table 2). In two isolates, recombination breakpoints were detected in the CR and AC1 regions. Furthermore, ChiLCV isolates (Nagpur and Goa) had recombination breakpoints in both the CR and AC2/AC3 regions. Similarly, the Ahmedabad isolate possessed breakpoints in AV1 and AC2/AC3 regions (Table 2). The recombination events located either in AC1 or in AV1 only were identified in five and three isolates, respectively (Table 2). However, no evidence of a recombination event was observed in ChiLCVV. In the phylogenetic tree of AV1 and AC1 ORFs, the sequences from recombinant molecules clustered with their respective contributors (Fig. S4). These results are in line with previous observations of AV1 and AC1 regions as the recombination hotspots (George *et al.*, 2015; Lefeuvre *et al.*, 2007; Lefeuvre & Moriones, 2015). In addition to interspecific recombination events, most of these ChiLCD-associated begomoviruses are obvious intra-species recombinants (Table 2).

Among the 33 betasatellites identified in this study, 17 showed evidence of recombination (Table 2). The recombination breakpoints were located around the A-rich region of CroYVMB (two isolates) and ToLCBDB (seven isolates), the SCR (six ToLCJoB isolates) or the βC1 ORF (sequences from Bhubaneswar and Jaunpur).

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Table 2. Recombination analysis of ChiLCD-associated begomoviruses and betasatellites

Recombinant sequence	Event no.	Breakpoint positions			Putative minor parent	Putative major parent	Methods†	*P* value‡
		Begin	End	Location in the genome*				
ChiLCSV-[IN-Sal-08]	1	454	1190	AV1	PepLCBV-[PK-Lah-04]	ToLCLKV-[LK-Ban-97]	RMC	7.28×10^{-11}
	2	2146	2712	AC1	PaLCuV-PK[PK-Cot-02]	ToLCSKV-[LK-Ban-97]	GBMC	1.94×10^{-24}
ChiLCV-PK[IN-Jor-09]	1	1997	2591	AC1	ChiLCV-[BD-Gaz]	PepLCBV-[PK-Lah-04]	RGBMCS	3.78×10^{-27}
	2	2592	2742	CR	ChiLCV-PK[PK-Mul-98]	PepLCBV-[PK-Lah-04]	RGBMC	4.56×10^{-11}
ChiLCV-PK[IN-Pal-09]	1	2083	2582	AC1	ChiLCV-[BD-Gaz]	PepLCBV-[PK-Lah-04]	RGBMCS	1.51×10^{-26}
ChiLCV-IN[IN-Ahm-09]	1	512	1084	AV1	ChiLCV-PK[PK:Mul:98]	ChiLCV-Chi[IN-Chi-05]	RGBMC	1.34×10^{-21}
	2	1230	1790	AC2/AC3	PepLCBV-BD[BD-Bog-99]	ChiLCV-Chi[IN-Chi-05]	RGBMCS	1.76×10^{-8}
ChiLCV-Gun[IN-Gun-08]	1	502	1151	AV1	ToLCV-Ban[IN-Ban-93]	ChiLCV-PK[PK-Mul-98]	RBMCS	7.17×10^{-7}
	2	2031	2595	AC1	ChiLCV-[BD-Gaz]	ChiLCV-PK[PK-Mul-98]	RGBMCS	3.65×10^{-24}
ChiLCV-IN[IN-Jod-09]	1	323	940	AV1	PepLCBV-[PK-Lah-04]	ChiLCV-PK[PK-Mul-98]	RGBS	8.36×10^{-33}
	2	2031	2695	AC1	ChiLCV-Chi[IN-Chi-05]	ChiLCV-PK[PK-Mul-98]	RGBMCS	3.06×10^{-16}
ChiLCV-Gun[IN-Nag1-09]	1	1090	1359	AC2/AC3	PepLCBV-[PK-Lah-04]	ChiLCV-PK[PK-Mul-98]	RGBMS	4.92×10^{-12}
and ChiLCV-Gun[IN-Goa-11]	2	2412	1	CR	PepLCBV-BD[BD-Bog-99]	ChiLCV-PK[PK-Mul-98]	RGBMS	1.89×10^{-9}
ChiLCV-IN[IN-ND-10]	1	7	1438	AV1	PepLCBV-[PK-Lah-04]	ChiLCV-Chi[IN-Chi-05]	RGBMCS	2.37×10^{-19}
	2	1680	2123	AC1	ChiLCV-PK[PK-Mul-98]	ChiLCV-Chi[IN-Chi-05]	RGBMCS	7.95×10^{-9}
ChiLCV-IN[IN-Noi-07]	1	47	1086	AV1	ChiLCV-PK[PK-Mul-98]	ChiLCV-Chi[IN-Chi-05]	RGBMCS	6.34×10^{-23}
	2	1786	7	AC1/CR	ChiLCV-Chi[IN-Chi-05]	ChiLCV-PK[PK-Mul-98]	RGBMCS	7.28×10^{-28}
ChiLCV-IN[IN-Pat-07]	1	441	1076	AV1	ToLCV-Ban[IN-Ban-93]	ChiLCV-Chi[IN-Chi-05]	RGBMC	8.80×10^{-8}
	2	2006	23	AC1/CR	ChiLCV-Chi[IN-Chi-05]	ChiLCV-PK[PK-Mul-98]	RGBMCS	1.83×10^{-27}
PepLCBV-BD[IN-Mag-07]	1	543	1153	AV1	ChiLCV-[BD-Gaz]	PepLCBV-PK[PK-Kha-04]	RGBS	1.72×10^{-16}
	2	1483	1642	AC1	Unknown species	PepLCBV-PK[PK-Kha-04]	GMCS	1.29×10^{-60}
PepLCBV-BD[IN-Gha1-10]	1	541	928	AV1	ChiLCV-[BD-Gaz]	PepLCBV-PK[PK-Kha-04]	RGBMCS	4.82×10^{-11}
	2	963	1165	AV1	Unknown species	PepLCBV-PK[PK-Kha-04]	GMC	1.77×10^{-4}
PepLCBV-PK[IN-Jab-07]	1	2136	2675	AC1	PepLCBV-PK[PK-Kha-04]	ChiLCV-Chi[IN-Chi-05]	RGBMCS	2.20×10^{-28}
PepLCBV-BD[IN-Chp2-10]	1	8	691	AV1	ChiLCV-Chi[IN-Chi-05]	PepLCBV-PK[PK-Kha-04]	RGBMCS	2.31×10^{-14}
	2	2401	2689	AC1	PepLCBV-PK[PK-Kha-04]	Unknown species	RGBMCS	5.70×10^{-22}
PepLCBV-Kan[IN-Kan-08]	1	1854	2689	AC1	PepLCBV-[PK-Lah-04]	PepLCBV-PK[PK-Kha-04]	RGBMC	9.09×10^{-11}
PepLCBV-IN[IN-Coi-08]	1	2357	2699	AC1	PaLCuV-IN[IN:Luc]	PepLCBV-[PK-Lah-04]	RGBMC	4.11×10^{-10}
	2	2724	407	CR	PepLCBV-BD[BD-Bog-99]	PepLCBV-[PK-Lah-04]	RGBMCS	7.06×10^{-10}
PepLCBV-IN[IN-Nag1-09]	1	414	1092	AV1	PaLCuV-PK[PK-Cot-02]	PepLCBV-[PK-Lah-04]	RBMCS	2.44×10^{-5}
	2	2619	2702	CR	PepLCBV-PK[PK-Kha-04]	PepLCBV-[PK-Lah-04]	GBMC	4.71×10^{-8}
ChiLCV-Chi[IN-Chp1-10]	1	2713	8	CR	PepLCBV-BD[BD-Bog-99]	ChiLCV-PK[PK-Mul-98]	RGBMCS	5.45×10^{-24}
ChiLCV-Chi[IN-Bar-08]	1	508	883	AV1	ChiLCV-[BD-Gaz]	ChiLCV-Chi[IN-Chi-05]	RGBMCS	1.67×10^{-16}
	2	2410	2691	AC1	ChiLCV-Chi[IN-Chi-05]	PepLCBV-PK[PK-Kha-04]	GBMCS	2.60×10^{-27}
ChiLCV-[IN-Kal-06]	1	546	962	AV1	ChiLCV-[BD-Gaz]	Unknown species	RGBMCS	3.07×10^{-14}
	2	2115	2588	AC1	ChiLCV-PK[PK-Mul-98]	PepLCBV-PK[PK-Kha-04]	RGBMCS	1.45×10^{-24}
ChiLCV[IN-Kol-08]	1	354	1160	AV1	ChiLCV-[BD-Gaz]	Unknown species	GMCS	1.62×10^{-31}
ChiLCV-[IN-Patna-08]	1	613	1262	AV1	PepLCBV-PK[PK-Kha-04]	Unknown species	RGBMCS	6.77×10^{-25}
	2	2097	2586	AC1	ChiLCV-PK[PK-Mul-98]	PepLCBV-PK[PK-Kha-04]	RGBMCS	3.99×10^{-24}

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Table 2. cont.

Recombinant sequence	Event no.	Breakpoint positions			Putative minor parent	Putative major parent	Methods†	*P* value‡
		Begin	End	Location in the genome*				
ChiLCV-[IN-Jau-07]	1	2564	2721	CR	ChiLCV-[BD-Gaz]	Unknown species	GBMC	4.17×10^{-9}
ChiLCV[IN-Gha2-10]	1	2045	2240	AC1	ChiLCV-PK[PK-Mul-98]	ChiLCV-[BD-Gaz]	RGBS	3.55×10^{-8}
ToLCV-Ban[IN-Ban- Chilli-10]	1	2238	2694	AC1	ToLCBaV-D[IN-KerII-05]	ChiLCV-PK[PK-Mul-98]	RGBMCS	4.18×10^{-15}
CroYVMB-[IN-Sal-Chilli-08]	1	779	963	A-rich	ToLCB-[IN-Var-06]	PaLCuB-[IN-Jab-03]	RGBMC	9.49×10^{-10}
CroYVMB-[IN-Pune-Chilli-09]	1	832	953	A-rich	ToLCB-[IN-Var-06]	PaLCuB-[IN-Jab-03]	RGBMC	3.25×10^{-9}
ToLCBDB-[IN-Gha1-Chilli-10]	1	519	1161	A-rich	ToLCBDB-[BD-Gaz-01]	ChLCB-[PK-Sia21-04]	RGBMCS	3.08×10^{-6}
ToLCBDB-[IN-Var-Chilli-06]	1	984	1193	A-rich	ToLCBDB-[IN-Raj-03]	Unknown species	RGBMC	8.36×10^{-5}
ToLCBDB-[IN-Coi1-Chilli-08]	1	521	1308	A-rich	ToLCBDB-[BD-Gaz-01]	ChLCB-[PK-Sia21-04]	RGBMCS	5.45×10^{-7}
ToLCBDB-[IN-Kan-Chilli-08]	1	472	1366	A-rich	ToLCBDB-[BD-Gaz-01]	ChLCB-[PK-Sia21-04]	RGBMCS	1.17×10^{-10}
ToLCBDB-[IN-Ahm-Chilli-09]	1	765	955	A-rich	ToLCBDB-[BD-Gaz-01]	Unknown species	GBMCS	3.78×10^{-15}
ToLCBDB-[IN-Jod-Chilli-09]	1	518	1036	A-rich	ToLCBDB-[BD-Gaz-01]	ChLCB-[PK-Sia21-04]	RGBMCS	1.06×10^{-5}
ToLCBDB-[IN-Vel2-Chilli-08]	1	716	1077	A-rich	ToLCBDB-[BD-Gaz-01]	ChLCB-[PK-Sia21-04]	RGBMCS	4.83×10^{-7}
ToLCJoB-[IN-Kal-Chilli-06]	1	1140	95	SCR	ChLCB-[PK-Guj23-04]	AYLCB-[PK-Fai3-94]	RGBMCS	6.83×10^{-13}
ToLCJoB-[IN-Kol-Chilli-08]	1	1123	89	SCR	ChLCB-[PK-Guj23-04]	AYLCB-[PK-Fai3-94]	RGBMCS	5.06×10^{-12}
ToLCJoB-[IN-Bar-Chilli-08]	1	1075	1261	A-rich/SCR	ToLCBDB-[IN-Raj-03]	Unknown species	RGBMCS	1.12×10^{-14}
ToLCJoB-[IN-Nag-Chilli-09]	1	1065	1251	A-rich/SCR	ToLCBDB-[IN-Raj-03]	Unknown species	RGBMCS	3.80×10^{-15}
ToLCJoB-[IN-Jor-Chilli-09]	1	1123	92	SCR	ChLCB-[PK-Guj23-04]	AYLCB-[PK-Fai3-94]	RGBMCS	1.99×10^{-14}
ToLCJoB-[IN-Mag-Chilli-07]	1	1123	92	SCR	ChLCB-[PK-Guj23-04]	AYLCB-[PK-Fai3-94]	RGBMCS	3.02×10^{-14}
ToLCJoB-[IN-Jau-Chilli-07]	1	1276	521	βC1	ToLCBDB-[IN-Raj-03]	ToLCJoB-[IN-Kenaf-07]	RGBMCS	4.54×10^{-30}
ToLCRnB-[IN-Bhu-Chilli-09]	1	146	499	βC1	ToLCBDB-[BD-Gaz-01]	Unknown species	GMCS	4.43×10^{-31}

*CR, common region of a begomovirus genome; SCR, satellite conserved region of betasatellites.
†R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, Chimaera; S, SiScan.
‡The lowest *P* value calculated for the underlined method in the penultimate column.

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Table 3. Nucleotide diversity of ChiLCD-associated Indian begomoviruses and betasatellites

Group	N*	Size	s	η	π	k	θ−w	θ−η
ChiLCV	19	2790	708	849	0.0934	253.5	0.0749	0.0997
PepLCBV	7	2743	576	690	0.0962	263.9	0.0857	0.1160
ToLCNDV	5	2740	134	134	0.0198	53.8	0.0236	0.0243
CroYVMB	4	1396	218	228	0.0913	121.2	0.0896	0.1053
ToLCBDB	18	1402	407	500	0.0855	112.4	0.0901	0.1262
ToLCJoB	7	1394	302	340	0.0829	110.0	0.0929	0.1184

*Details of the abbreviations used are described in Methods.

Genetic structure and substitution rate of begomoviruses and betasatellites

From an evolutionary perspective, recombination during the replication of begomoviruses alone may not lead to the genetic variation *de novo*. Nucleotide substitutions that occur during evolution also contribute to genetic variation. ChiLCD-associated begomoviruses and betasatellites were determined to have a high degree of genetic variability ($\pi > 0.08$). Our analyses showed ChiLCV and PepLCBV to be the most diverse begomoviruses, and CroYVMB and ToLCBDB to possess much higher genetic variability than the other betasatellites (Table 3). To assess the presence of any subpopulations within begomoviruses, a cluster-based method (STRUCTURE) was employed. The results suggest the presence of seven subpopulations among ChiLCV and five subpopulations were noticed among PepLCBV isolates (Fig. S5a, b). However, low nucleotide diversity and the absence of any subpopulations among chilli isolates of ToLCNDV suggest their recent association with this host (Fig. S5c, Table 3). Furthermore, to verify the randomness of nucleotide variability throughout the genome, we analysed the nucleotide diversity of all the ORFs encoded by ChiLCD-associated begomoviruses and betasatellites. Higher nucleotide diversity was observed in all the ORFs encoded by the begomoviruses. The nucleotide diversity patterns of the complete genomes of betasatellites and βC1 ORFs were observed to be highly congruent (Fig. S6). The selection pressure on the ORFs was analysed using single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL) and random-effects likelihood (REL) methods. A large proportion of the sites were found to have evolved under purifying/negative selection (Table S2).

We also estimated the nucleotide substitution rate of AV1 and βC1 ORFs using the parameters listed in Table 4. The mean nucleotide substitution rate was estimated to be 2.60×10^{-3} substitutions site^{-1} year^{-1} for the AV1 gene of ChiLCV, 2.57×10^{-4} substitutions site^{-1} year^{-1} for the βC1 gene of ChiLCB and 5.22×10^{-4} substitutions site^{-1} year^{-1} for the βC1 gene of ToLCBDB (Table 4). Since codon degeneracy could play a vital role in the selection pressure leading to genetic variation, we estimated the mutation rate of all three nucleotide positions in the codon using the BSP method. Strikingly, the mutation rate of the AV1 ORF in codon position 1 was seen to be higher than the mutation rates of the other positions (Table 4). In βC1, however, the mutation rate was found to be highest in the wobble position of the codon; this position of the codon affects the amino acid sequence less often than do the other positions.

Infectivity analysis of newly emerging begomoviruses

Partial tandem repeats of the newly identified begomoviruses NgA (PepLCBV-IN) and VeA (ChiLCVV) infected *N. benthamiana* plants. Notably, the incubation period remained unaltered upon co-inoculation of either NgA or VeA with their cognate betasatellites (Ngβ and Veβs) (Table 5). *N. benthamiana* plants co-inoculated with either NgA + Ngβ or VeA + Veβ1 exhibited symptoms such as downward leaf curling and enations (Fig. 2a). However, agroinoculation of VeA + Veβ2 failed to induce downward leaf curling in *N. benthamiana*, a characteristic symptom induced by betasatellites. In Southern analysis, all these begomoviruses could trans-replicate their cognate betasatellites, but these

Table 4. Mean substitution rate and codon position mutation rate for AV1 and βC1 ORFs encoded by ChiLCD-associated begomoviruses and betasatellites

Parameters used	AV1				βC1		
	ChiLCV	CLCuMuV[a]*	TYLCV[b]	EACMV[c]	ChiLCB	ToLCBDB	CLCuMuB[a]
Number of sequences analysed	52	19	54	71	25	31	39
Chain length (millions)	20	40	20	15	40	40	40
Mean substitution rate (nucleotides site^{-1} year^{-1})	2.60×10^{-3}	4.24×10^{-4}	4.63×10^{-4}	1.37×10^{-3}	2.57×10^{-4}	5.22×10^{-4}	3.51×10^{-3}
Codon position 1 mutation rate	1.690	1.640	–	–	0.800	1.067	0.850
Codon position 2 mutation rate	0.366	0.449	–	–	0.622	0.679	0.730
Codon position 3 mutation rate	0.944	0.909	–	–	1.578	1.256	1.430

*Estimated by: *a*, Nawaz-ul-Rehman *et al.* (2012); *b*, Duffy & Holmes (2008); *c*, Duffy & Holmes (2009).

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Table 5. Infectivity of cloned viral components on *N. benthamiana* and *C. annuum* plants

Host plant	Viral construct	Number of symptomatic plants/number of inoculated plants	Incubation period (days)*	Type of symptoms on inoculated plants†	PCR‡	
					DNA-A	DNA-β
Nicotiana benthamiana	NgA	10/10	8	LC, LR, St	+	−
	NgA + Ngβ	10/10	8	DLC, LCr, St	+	+
	VeA	10/10	6	LC, St, SL, MY	+	−
	VeA + Veβ1	10/10	6	DLC, LCr, St, SL, MY	+	+
	VeA + Veβ2	6/6	7	LC, LCr, St, SL, MY	+	+
Capsicum annuum	NgA	10/20	35	MLC	+	−
var. Sulekha	NgA + Ngβ	15/21	25	LC, LR, LCr, LB, Vt, SL, St	+	+
	VeA	9/20	35	MLC	+	−
	VeA + Veβ1	14/22	27	LC, LR, LCr, LB, Vt, SL, St	+	+
	ChA	5/20	55	MLC	+	−
	ChA + Chβ	11/20	40	LC, LCr, Vt	+	+
	JoA	7/20	51	MLC	+	−
	JoA + Joβ	9/19	37	LC, LCr, LR, LB	+	+
	MgA	5/20	50	MLC	+	−
	MgA + Mgβ	12/21	38	LC, LCr, LR, LB	+	+

*Incubation period is defined as the time taken for the first symptom appearance in days post-inoculation.
†LC, leaf curling; DLC, downward leaf curling; MLC, mild leaf curling; SL, small leaves; MY, mild yellowing; LB, leaf blistering; LR, leaf rolling; St, stunting; LCr, leaf crinkling; Vt, vein thickening.
‡Primers specific for DNA-A and DNA-β; + indicates the presence of viral components.

betasatellites did not influence the accumulation of begomovirus in *N. benthamiana* (Fig. 2b). *C. annuum* plants inoculated with begomoviruses alone (ChA, JoA, MgA, NgA, VeA) developed mild symptoms and the presence of these begomoviruses was ascertained by PCR (Fig. S7; Table 5). However, when these begomoviruses were inoculated along with their cognate betasatellites (ChA + Chβ, JoA + Joβ, MgA + Mgβ, NgA + Ngβ, VeA + Veβ1), symptoms such as stunting, leaf curling, leaf crinkling, blistering and vein thickening were observed on infected *C. annuum* var. Sulekha plants (Fig. 2a). In Southern analysis, no viral DNA accumulation could be detected from the chilli plants inoculated with diverse begomoviruses in the absence of their cognate betasatellites. However, when chilli plants were inoculated with begomoviruses and cognate betasatellites, high levels of viral DNA were detected (Fig. 2b). Collectively, our results suggest that these begomoviruses require betasatellites for development of severe ChiLCD in *Capsicum* spp. (Figs 2b and S7, Table 5).

DISCUSSION

ChiLCD is one of the most devastating diseases of chillies in India and has become a major constraint for production of *Capsicum* (Maruthi *et al.*, 2007). The viruses causing ChiLCD were identified initially based on the type of symptoms induced on the host and later by either partial or complete nucleotide sequencing of the viral genome (Chattopadhyay *et al.*, 2008; Khan *et al.*, 2006; Kumar *et al.*, 2012; Mishra *et al.*, 1963; Senanayake *et al.*, 2007). The present study provides the first comprehensive survey and analysis to our knowledge of begomoviruses associated with ChiLCD in India. Our effort has resulted in the isolation and characterization of 74 genomic components (35 DNA-A/DNA-A-like, 6 DNA-B and 33 betasatellites). The existence of newly identified and previously reported begomoviruses associated with ChiLCD clearly indicates the high degree of species diversity and genetic variability of these pathogens. In addition, the newly emerged begomoviruses have been shown to infect *N. benthamiana* and *Capsicum* spp. Our results demonstrate that the viruses tested are not only associated with but are the causal agents of ChiLCD. Furthermore, the betasatellites are found to be a prerequisite for the induction of severe leaf curl symptoms in *Capsicum* spp.

Our study has identified ChiLCV as the most widely distributed begomovirus associated with ChiLCD across India, followed by PepLCBV and ToLCNDV (Fig. 3). In the eastern region (encompassing states such as West Bengal, Bihar and the eastern part of Uttar Pradesh), ChiLCV and PepLCBV-BD predominate, with a notable exception of ChiLCV-PK in Jorehat and Varanasi. Mixed infections of ChiLCV and PepLCBV were also noticed in eastern Uttar Pradesh (Ghazipur) and the Bihar (Chhapra) region. The isolates from the northern and western regions, including the states of Haryana, Delhi, Rajasthan, Himachal

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

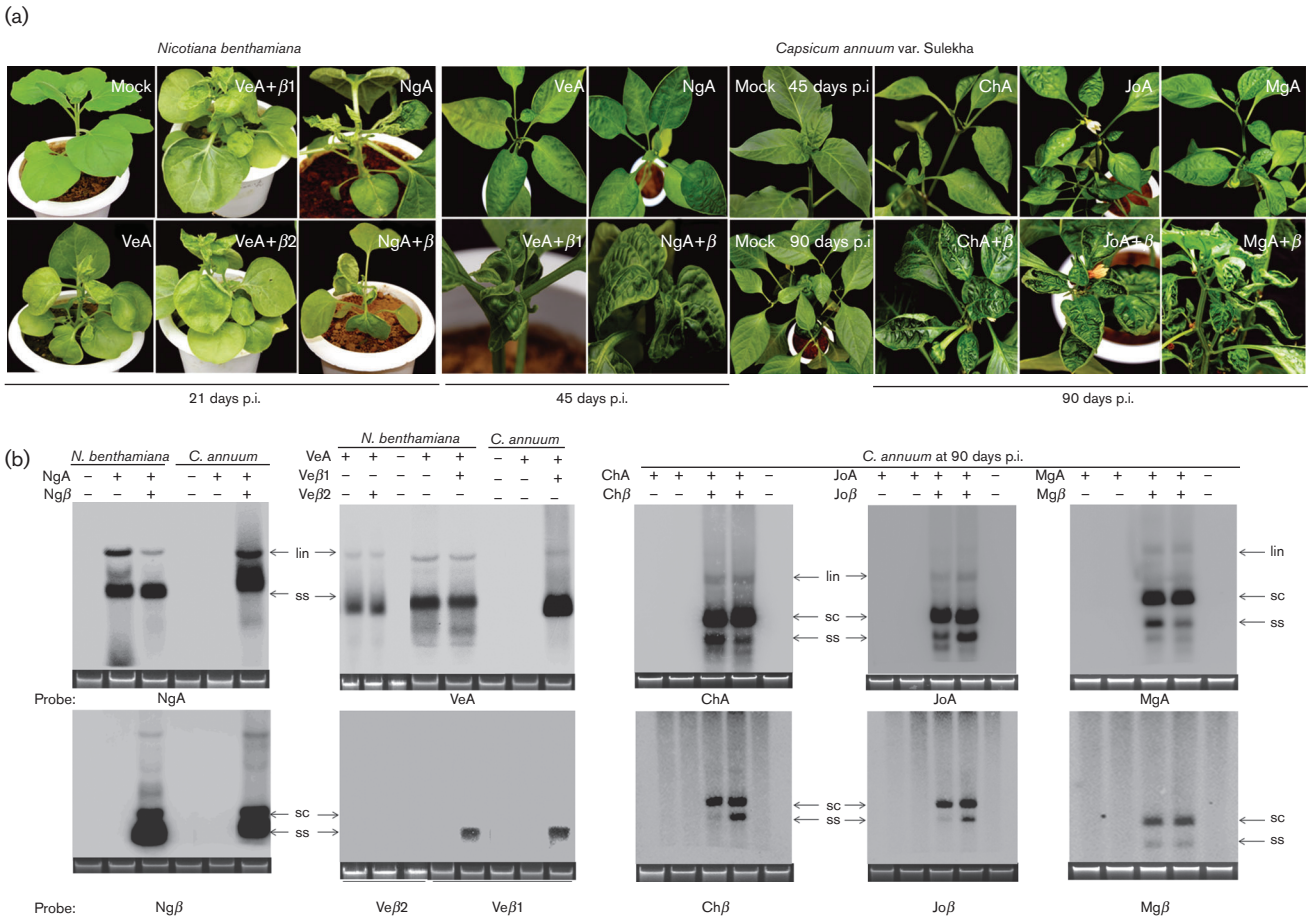


Fig. 2. Infectivity and viral DNA accumulation of cloned components. (a) Appearance of symptoms on *N. benthamiana* and *C. annuum* plants infected with various begomoviruses alone or with their respective cognate betasatellites as indicated. (b) Relative DNA accumulation of various begomoviruses and betasatellites in *N. benthamiana* [at 21 days post-inoculation (p.i.)] and *C. annuum* plants (at 90 days p.i.), as analysed by Southern hybridization for each combination mentioned. Probes used for hybridization are given at the bottom of each blot. The replicative DNA forms are abbreviated as linear (lin), super-coiled (sc) and single-stranded (ss). Ethidium-bromide-stained DNA serves as a loading control. VeA, ChiLCVV; NgA, PepLCBV-IN; ChA, ChiLCV-PK; JoA, ChiLCV; MgA, PepLCBV-BD; Veβ and Chβ, ToLCBDB; Joβ, Mgβ and Ngβ, ToLCJoB.

Pradesh, Goa, Gujarat, Maharashtra and Uttar Pradesh (western part) belong to ChiLCV. Taken together, these data indicate that ChiLCD in these regions is associated with the begomoviruses previously reported from Pakistan, whereas viruses identified in the eastern part of India are predominantly present in Bangladesh (Fig. 3). It is relevant to mention that states adjoining the border region follow similar cropping patterns, which might have facilitated the migration of the viruses through whiteflies.

In southern India, various distinct begomoviruses such as ChiLCV-Gun, ChiLCVV, PaLCuV-Cro, PepLCBV-IN, ToLCV-Ban and ToLCNDV were found to be associated with ChiLCD. In addition, the one previously unreported begomovirus (ChiLCSV) was also identified in this region. Taken altogether, our analysis from southern India indicates that diverse begomoviruses and betasatellites are associated with ChiLCD (Fig. 3). Under laboratory conditions, an isolate of tomato leaf curl Gujarat virus caused systemic infection in chilli (Chakraborty *et al.*, 2003), and hence the association of diverse tomato-infecting begomoviruses with ChiLCD-affected plants is not surprising.

Recombination has been reported to be the major factor in the emergence of begomoviruses infecting crops and weeds in India (Kumar *et al.*, 2011; Kumari *et al.*, 2010; Singh *et al.*, 2012). Begomoviruses associated with chilli and tomato contributed to 32 out of the 44 recombination events detected in ChiLCD-associated begomoviruses, and the remaining recombination events (12/44) were associated with yet undescribed or non-solanaceous-crop-infecting begomoviruses. Our results indicate that the chilli- and tomato-associated begomoviruses have a complex history of recombination in this part of the world, with a relevant contribution of interspecific recombination. It is noteworthy that ChiLCSV may have emerged after multiple recombination events. A non-recombinant

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

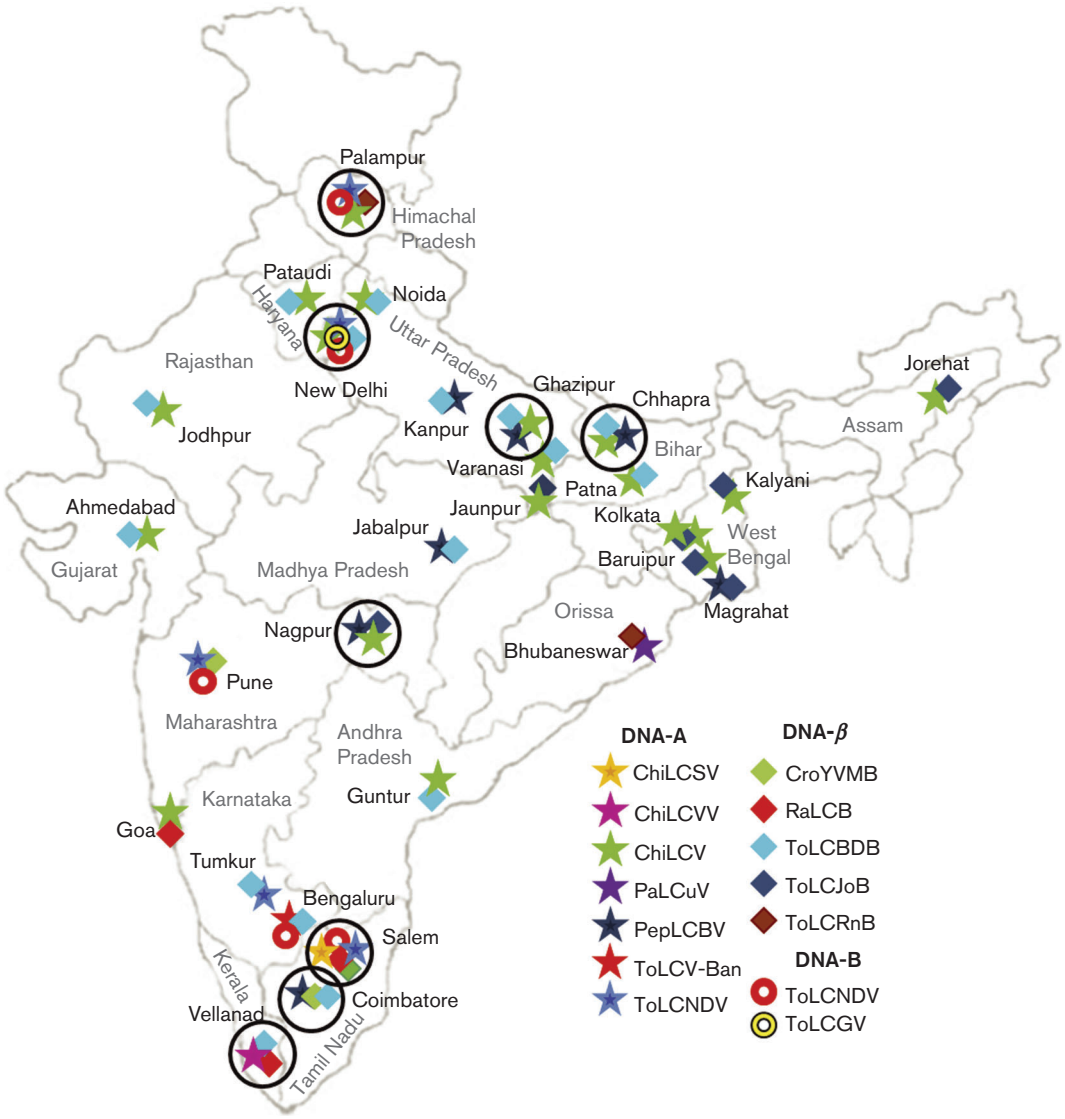


Fig. 3. Distribution of ChiLCD-associated begomoviruses and betasatellites in India identified in this study. Diverse components of DNA-A, DNA-B and betasatellites are depicted in different shapes and colours. Black circles indicate mixed infections.

ChiLCVV could have evolved quite separately after the earlier recombination event(s) and resulted in the present genome organization. These results pinpoint the significant contribution of solanaceous as well as non-solanaceous-infecting begomoviruses in intra- and inter-species recombination leading to the emergence of ChiLCD-associated begomoviruses. Recombination among ChiLCD-associated begomoviruses could provide high evolutionary potential for rapid multiplication, expansion of geographical territory and, more importantly, invasion of new hosts.

Mixed infections of distinct begomoviruses were detected at six locations, viz. Chhapra, Ghazipur, Nagpur, New Delhi, Palampur and Salem (Table 1, Fig. 3). Notably, ChiLCV-PK and PepLCBV-IN have emerged through intra- as well as inter-species recombination. This observation emphasizes that mixed infection, a pre-requisite for recombination, might have facilitated their emergence. Interestingly, New Delhi, Palampur and Salem samples were found to be infected with a monopartite begomovirus (either ChiLCVs or ChiLCSV) and a bipartite begomovirus (ToLCNDV). The presence of a bipartite begomovirus together with a monopartite one could further promote the evolution of severe disease-causing complexes. Possible synergistic interactions, if any, between these monopartite and bipartite begomoviruses in chilli have not been studied so far.

Betasatellites, being important pathogenicity determinants for monopartite begomoviruses, have been the subject of many scientific investigations (Bhattacharyya *et al.*, 2015;

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Cui *et al.*, 2005; Li *et al.*, 2014; Saunders *et al.*, 2004; Yang *et al.*, 2011). At the time of sampling, only two betasatellite groups, ToLCBDB and ChiLCB, were known to be associated with ChiLCD in India (Chattopadhyay *et al.*, 2008; Kumar *et al.*, 2011). In India, the most prevalent betasatellites associated with ChiLCD have been reported from tomato, rather than from weeds. Among these, ToLCBDB has co-existed with diverse begomoviruses predominantly in the western part of India, whereas ToLCJoB is mostly located in the eastern and western regions (Fig. 3). In spite of the prevalence of begomoviruses associated with ChiLCD identified from Pakistan (ChiLCV, PepLCBV), their cognate ChiLCB could not be detected here. It is interesting to note the association of betasatellites with both bipartite and monopartite begomoviruses from New Delhi, Palampur, Pune and Salem. Recently, diverse betasatellites have been shown to enhance the accumulation of ToLCNDV DNA-A and antagonistically interact with the ToLCNDV DNA-B component (Jyothsna *et al.*, 2013). However, the significance of this association for the establishment of ChiLCD needs to be explored.

Mixed infection of betasatellites also persists in the southern region of India (Salem, Coimbatore, Vellanad). Both betasatellites from Vellanad (ToLCBDB and RaLCB) could be trans-replicated by ChiLCVV in *N. benthamiana*. However, they could neither influence symptom severity nor reduce the incubation period (Fig. 2, Table 5). This observation is in agreement with earlier reports suggesting that the betasatellites are dispensable for disease development in *N. benthamiana* (Ranjan *et al.*, 2014; Saunders *et al.*, 2000). Chilli plants inoculated with DNA-A-like sequences alone produced mild symptoms; however, the accumulation of viral DNA as well the severity of the symptoms was increased in the presence of cognate betasatellites (Fig. 2, Table 5). This is the first experimental evidence to our knowledge demonstrating the necessity of betasatellites for development of severe leaf curl disease in chilli, thereby explaining their ubiquitous presence with ChiLCD. Furthermore, their dependence on betasatellites for severe disease development supports the notion that these monopartite begomoviruses require betasatellites for ChiLCD, similar to the complexes reported with ageratum yellow vein and cotton leaf curl diseases (Mansoor *et al.*, 2006; Saunders *et al.*, 2000).

Geminiviruses are known to have high rates of nucleotide substitution, similar to those of RNA viruses (Drake, 1993; Duffy & Holmes, 2008, 2009). ChiLCD-associated begomoviruses possess higher genetic variability than observed in other begomoviruses (Melgarejo *et al.*, 2013; Rocha *et al.*, 2013; Silva *et al.*, 2012). As described by Lima *et al.* (2013), the nucleotide diversity in these begomoviruses was noticed to be non-uniformly distributed throughout the genome. Our results emphasize that purifying selection is the major evolutionary determinant leading to the emergence of these begomoviruses and betasatellites, in agreement with other studies (Lima *et al.*, 2013; Rocha *et al.*, 2013; Silva *et al.*, 2012).

In conclusion, the information generated through our studies demonstrates that ChiLCD is caused by a large group of monopartite begomoviruses and a few bipartite begomoviruses in association with diverse betasatellites. Diverse species of begomoviruses and betasatellites were found at every geographical location in India and their evolution is largely driven by intra- and inter-species recombination and by nucleotide substitution. It is evident from the infectivity analysis that the betasatellites are essential for disease development in the chilli host. As chilli-associated begomoviruses and betasatellites are expanding their host range and geographical distribution (George *et al.*, 2014), further epidemiological studies of the impact of the high genetic variability of these viruses are required.

METHODS

Isolation, cloning and sequencing of full-length genomes. An extensive survey was carried out in all major chilli-growing regions across India. For each region, a total of nine samples (from at least three individual fields) were collected from naturally infected chilli plants exhibiting typical symptoms of ChiLCD (Table S1). Total genomic DNA was extracted from all samples (Dellaporta *et al.*, 1983) and subjected to RCA using φ29 DNA polymerase (Inoue-Nagata *et al.*, 2004). Association of begomovirus with these symptomatic samples was verified by PCR using abutting primers specific for DNA-A and DNA-B components (Rojas *et al.*, 1993; Wyatt & Brown, 1996). RCA products of all the samples were digested with single-cutting restriction endonucleases, viz. *Bam*HI, *Kpn*I, *Hin*dIII or *Pst*I, and cloned into pBluescript II KS (+) vector (Stratagene), previously linearized with the respective restriction enzymes. Simultaneously, PCR with primers specific for DNA-A and DNA-B (George *et al.*, 2014) and betasatellite (Briddon *et al.*, 2002) were used to amplify the full-length genomic components. Details of the primers used to clone viral components are provided in Table S3. PCR products of either ~2.8 kb or ~1.3 kb in size were cloned into the pTZ57R/T vector (Fermentas). All the recombinant clones obtained from each of the samples were digested with various restriction enzymes and were sequenced partially (using M13 primers) to identify the variability among them, if any. On the basis of restriction variability and partial sequencing of the viral genome, one of the representative cloned viral genomes was sequenced commercially at the University of Delhi South Campus.

Phylogenetic and recombination analysis. Using the CLUSTAL W method in SDT v. 1.0 (Muhire *et al.*, 2014) and MEGA5 (Tamura *et al.*, 2011), neighbour-joining phylogenetic dendrograms and percentage pairwise identity of the identified sequences and the representative sequences from the database were generated. For RDP analysis, the begomoviruses were selected according to the criteria followed by George *et al.* (2015) and the sequence accession numbers are listed in Table S4. Recombination breakpoints and their putative parental sequences were detected by the RDP3 program with the Bonferroni-corrected *P* value cut-off of 0.01 as described by George *et al.* (2015) and Martin *et al.* (2010).

Estimation of population structure and nucleotide substitution rates. To assess the genetic variability in the virus populations, parameters in DnaSP v. 5.10 used were total number of segregating sites (s), total number of mutations (η), average number of nucleotide differences between sequences (k), nucleotide diversity (π), Watterson's estimate of the population mutation rate based on the total number of segregating sites ($\theta - w$) and the total number of mutations ($\theta - \eta$). Begomovirus population structure was studied using STRUCTURE v. 2.3.3 as described by Prasanna *et al.* (2010) and Pritchard

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

et al. (2000). The number of subpopulations (*K* value) was estimated by using 100 000 Markov chain Monte Carlo steps after a burn-in period of 100 000 steps. The best-supported *K* values were determined based on ln*P*(*D*) after three independent runs.

Neutrality tests, Tajima's *D*, and Fu and Li's D^* and F^*, were employed to test the hypothesis of what type of selection was occurring in the ORFs encoded by the viral components (Rozas *et al.*, 2003). The detection of sites evolved under positive and negative selection was performed by the methods implemented in the Data-Monkey server (http://www.datamonkey.org): SLAC, FEL and REL (Kosakovsky Pond & Frost, 2005). Bayes factors >50 or *P* value <0.1 were used as thresholds for the analysis. NEXUS files of AV1 and βC1 regions were generated using CLUSTAL W in MEGA5, and nucleotide substitution rate was estimated using the GTR nucleotide substitution model, uncorrelated lognormal relaxed clock model and BSP co-alescence model in the BEAST v. 1.7 package (Drummond *et al.*, 2012; Nawaz-ul-Rehman *et al.*, 2012). The begomovirus and betasatellite sequences used to estimate population structure and nucleotide substitution rate are provided in Table S5.

Generation of infectious clones. For infectivity analysis, newly identified begomoviruses and their associated betasatellites were cloned as partial tandem repeats in plant transformation vectors. For PepLCBV-IN (Nagpur isolate), the *Eco*RI (1653)–*Bam*HI (155) fragment was cloned into pCAMBIA1300, followed by the mobilization of the monomer to generate a partial tandem repeat, designated NgA. For ChiLCVV, a 2.4 kb *Sal*I (1856)–*Kpn*I (1620) fragment containing the CR was cloned into pCAMBIA2301. The full-length monomer linearized with *Kpn*I was then ligated to generate the ChiLCVV tandem repeat construct (designated VeA). For ToLCBDB- and RaLCB-Vellanad, 1.2 kb (*Bam*HI and *Kpn*I) and 0.4 kb (*Eco*RI and *Kpn*I) fragments were cloned into pCAMBIA1300, followed by ligation of the respective monomer to yield partial tandem repeats of ToLCBDB- and RaLCB-Vellanad (referred as Veβ1 and Veβ2, respectively). For ToLCJoB-Nagpur, the 470 bp *Kpn*I (1297)–*Bam*HI (411) fragment was cloned in pCAMBIA1300, followed by insertion of the full-length *Kpn*I fragment to produce the tandem repeat of the ToLCJoB-Nagpur (as Ngβ). Similarly, partial tandem repeat constructs of ChiLCV-PK (ChA-EF190217), PepLCBV-BD (MgA-HM007111), ToLCBDB (Chβ-EF190215), ToLCJoB (Mgβ-HM007112, Joβ-EF190216) and ChiLCV (JoA-EF194765) were also generated and the cloning details are shown in Table S6.

Plant inoculation and Southern blot hybridization. Infectivity of the cloned components was analysed in *N. benthamiana* plants through the *Agrobacterium*-mediated inoculation method (Singh *et al.*, 2012). *C. annuum* var. Sulekha plants were bombarded using a HELIOS Gene Gun apparatus (Bio-Rad) following the manufacturer's instructions. Twenty-five milligrams of gold microcarriers (0.6 μm diameter) were coated with plasmid mixture (3 μg mg^{-1}) of each of these infectious clones and delivered at 300 p.s.i. (pounds per square inch). These plants were maintained in an insect-free growth chamber to monitor the symptom induction. Total genomic DNA (10 μg) from the systemic leaves was resolved on an agarose gel (0.8 %) and transferred onto nylon membrane (Singh *et al.*, 2012). The blots were hybridized separately with the [α^{32}P]dCTP-radiolabelled AV1 region of DNA-A or the monomer of DNA-β specific to the species concerned and the viral DNA was detected using a phosphor image analyser.

ACKNOWLEDGEMENTS

We thank Dr H. C. Prasanna of the Indian Institute of Vegetable Research, Varanasi for his help in STRUCTURE analysis. We also acknowledge the help of Dr Himanshu Pandya and Mr Biju George for providing the infected samples from Ahmedabad and Jodhpur, respectively. This work was supported by a research grant provided by the Department of Biotechnology, Government of India (BT/PR8107/PBD/16/721/2006) and the central instrumentation facility supported by DST-FIST II, DST-PURSE, UGC-RNW and UPOE.

REFERENCES

Bhattacharyya, D., Prabu, G., Kishore Kumar, R., Kushwaha, N. K., Sharma, V. K., Yusuf, Md. A. & Chakraborty, S. (2015). A geminivirus betasatellite damages structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. *J Exp Bot* **66**, 5881–5895.

Briddon, R. W., Bull, S. E., Mansoor, S., Amin, I. & Markham, P. G. (2002). Universal primers for the PCR-mediated amplification of DNA β: a molecule associated with some monopartite begomoviruses. *Mol Biotechnol* **20**, 315–318.

Briddon, R. W., Brown, J. K., Moriones, E., Stanley, J., Zerbini, M., Zhou, X. & Fauquet, C. M. (2008). Recommendations for the classification and nomenclature of the DNA-β satellites of begomoviruses. *Arch Virol* **153**, 763–781.

Brown, J. K., Fauquet, C. M., Briddon, R. W., Zerbini, F. M., Moriones, E. & Navas-Castillo, J. (2012). Family *Geminiviridae*. In *Virus Taxonomy: Classification and Nomenclature of Viruses – Ninth Report of the International Committee on Taxonomy of Viruses*, pp. 351–373. Edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. London: Elsevier.

Brown, J. K., Zerbini, F. M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J. C. F., Fiallo-Olivé, E., Briddon, R. W., Hernández-Zepeda, C. & other authors (2015). Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Arch Virol* **160**, 1593–1619.

Chakraborty, S., Pandey, P. K., Banerjee, M. K., Kalloo, G. & Fauquet, C. M. (2003). *Tomato leaf curl Gujarat virus*, a new begomovirus species causing a severe leaf curl disease of tomato in Varanasi, India. *Phytopathology* **93**, 1485–1495.

Chattopadhyay, B., Singh, A. K., Yadav, T., Fauquet, C. M., Sarin, N. B. & Chakraborty, S. (2008). Infectivity of the cloned components of a begomovirus: DNA β complex causing chilli leaf curl disease in India. *Arch Virol* **153**, 533–539.

Cui, X., Li, G., Wang, D., Hu, D. & Zhou, X. (2005). A *Begomovirus* DNA-β encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. *J Virol* **79**, 10764–10775.

Dellaporta, S. L., Wood, J. & Hicks, J. B. (1983). A plant DNA minipreparation: version II. *Plant Mol Biol Rep* **1**, 19–21.

Drake, J. W. (1993). Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci U S A* **90**, 4171–4175.

Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**, 1969–1973.

Duffy, S. & Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. *J Virol* **82**, 957–965.

Duffy, S. & Holmes, E. C. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *J Gen Virol* **90**, 1539–1547.

George, B., Kumar, R. V. & Chakraborty, S. (2014). Molecular characterization of *Chilli leaf curl virus* and satellite molecules associated with leaf curl disease of *Amaranthus* spp. *Virus Genes* **48**, 397–401.

George, B., Alam, ChM., Kumar, R. V., Gnanasekaran, P. & Chakraborty, S. (2015). Potential linkage between compound microsatellites and recombination in geminiviruses: evidence from comparative analysis. *Virology* **482**, 41–50.

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D. & Mansoor, S. (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nat Rev Microbiol* **11**, 777–788.

Idris, A. M., Shahid, M. S., Briddon, R. W., Khan, A. J., Zhu, J. K. & Brown, J. K. (2011). An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. *J Gen Virol* **92**, 706–717.

Inoue-Nagata, A. K., Albuquerque, L. C., Rocha, W. B. & Nagata, T. (2004). A simple method for cloning the complete begomovirus genome using the bacteriophage ϕ29 DNA polymerase. *J Virol Methods* **116**, 209–211.

Jyothsna, P., Haq, Q. M. I., Singh, P., Sumiya, K. V., Praveen, S., Rawat, R., Briddon, R. W. & Malathi, V. G. (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. *Appl Microbiol Biotechnol* **97**, 5457–5471.

Khan, M. S., Raj, S. K. & Singh, R. (2006). First report of *Tomato leaf curl New Delhi virus* infecting chilli in India. *Plant Pathol* **55**, 289.

Kosakovsky Pond, S. L. & Frost, S. D. W. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* **22**, 1208–1222.

Kumar, Y., Hallan, V. & Zaidi, A. (2011). Chilli leaf curl Palampur virus is a distinct begomovirus species associated with a betasatellite. *Plant Pathol* **60**, 1040–1047.

Kumar, R. V., Singh, A. K. & Chakraborty, S. (2012). A new monopartite begomovirus species, *Chilli leaf curl Vellanad virus*, and associated betasatellites infecting chilli in the Vellanad region of Kerala, India. *New Disease Reports* **25**, 20.

Kumari, P., Singh, A. K., Chattopadhyay, B. & Chakraborty, S. (2010). Molecular characterization of a new species of *Begomovirus* and betasatellite causing leaf curl disease of tomato in India. *Virus Res* **152**, 19–29.

Lefeuvre, P. & Moriones, E. (2015). Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. *Curr Opin Virol* **10**, 14–19.

Lefeuvre, P., Lett, J. M., Reynaud, B. & Martin, D. P. (2007). Avoidance of protein fold disruption in natural virus recombinants. *PLoS Pathog* **3**, e181.

Li, F., Huang, C., Li, Z. & Zhou, X. (2014). Suppression of RNA silencing by a plant DNA virus satellite requires a host calmodulin-like protein to repress RDR6 expression. *PLoS Pathog* **10**, e1003921.

Lima, A. T. M., Sobrinho, R. R., González-Aguilera, J., Rocha, C. S., Silva, S. J. C., Xavier, C. A., Silva, F. N., Duffy, S. & Zerbini, F. M. (2013). Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. *J Gen Virol* **94**, 418–431.

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R., Stanley, J. & Markham, P. G. (1999). Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology* **259**, 190–199.

Mansoor, S., Zafar, Y. & Briddon, R. W. (2006). Geminivirus disease complexes: the threat is spreading. *Trends Plant Sci* **11**, 209–212.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. & Lefeuvre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463.

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

Melgarejo, T. A., Kon, T., Rojas, M. R., Paz-Carrasco, L., Zerbini, F. M. & Gilbertson, R. L. (2013). Characterization of a New World monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. *J Virol* **87**, 5397–5413.

Mishra, M. D., Raychaudhuri, S. P. & Jha, A. (1963). Virus causing leaf curl of chilli (*Capsicum annuum* L.). *Indian J Microbiol* **3**, 73–76.

Muhire, B. M., Varsani, A. & Martin, D. P. (2014). SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* **9**, e108277.

Nawaz-ul-Rehman, M. S. & Fauquet, C. M. (2009). Evolution of geminiviruses and their satellites. *FEBS Lett* **583**, 1825–1832.

Nawaz-ul-Rehman, M. S., Nahid, N., Mansoor, S., Briddon, R. W. & Fauquet, C. M. (2010). Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. *Virology* **405**, 300–308.

Nawaz-ul-Rehman, M. S., Briddon, R. W. & Fauquet, C. M. (2012). A melting pot of Old World begomoviruses and their satellites infecting a collection of *Gossypium* species in Pakistan. *PLoS One* **7**, e40050.

Padidam, M., Sawyer, S. & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218–225.

Prasanna, H. C., Sinha, D. P., Verma, A., Singh, M., Singh, B., Rai, M. & Martin, D. P. (2010). The population genomics of begomoviruses: global scale population structure and gene flow. *Virol J* **7**, 220.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Ranjan, P., Singh, A. K., Kumar, R. V., Basu, S. & Chakraborty, S. (2014). Host-specific adaptation of diverse betasatellites associated with distinct Indian tomato-infecting begomoviruses. *Virus Genes* **48**, 334–342.

Rocha, C. S., Castillo-Urquiza, G. P., Lima, A. T. M., Silva, F. N., Xavier, C. A., Hora-Júnior, B. T., Beserra-Júnior, J. E., Malta, A. W., Martin, D. P. & other authors (2013). Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. *J Virol* **87**, 5784–5799.

Rojas, M. R., Gilbertson, R. L., Russel, D. R. & Maxwell, D. P. (1993). Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Disease* **77**, 340–347.

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497.

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

Saunders, K., Norman, A., Gucciardo, S. & Stanley, J. (2004). The DNA β satellite component associated with ageratum yellow vein disease encodes an essential pathogenicity protein (βC1). *Virology* **324**, 37–47.

Senanayake, D. M. J. B., Mandal, B., Lodha, S. & Varma, A. (2007). First report of *Chilli leaf curl virus* affecting chilli in India. *Plant Pathol* **56**, 343.

Senanayake, D. M. J. B., Jayasinghe, J. E. A. R. M., Shilpi, S., Wasala, S. K. & Mandal, B. (2013). A new begomovirus-betasatellite complex is associated with chilli leaf curl disease in Sri Lanka. *Virus Genes* **46**, 128–139.

Shih, S. L., Tsai, W. S., Green, S. K. & Singh, D. (2007). First report of *Tomato leaf curl Joydebpur virus* infecting chilli in India. *Plant Pathol* **56**, 341.

Silva, S. J. C., Castillo-Urquiza, G. P., Hora-Junior, B. T., Assuncao, I. P., Lima, G. S. A., Pio-Ribeiro, G., Mizubuti, E. S. G. & Zerbini, F. M. (2012). Species diversity, phylogeny and genetic variability of

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

begomovirus populations infecting leguminous weeds in northeastern Brazil. *Plant Pathol* **61**, 457–467.

Singh, A. K., Chattopadhyay, B. & Chakraborty, S. (2012). Biology and interactions of two distinct monopartite begomoviruses and beta-satellites associated with radish leaf curl disease in India. *Virol J* **9**, 43.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

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

Varsani, A., Navas-Castillo, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Brown, J. K., Murilo Zerbini, F. & Martin, D. P. (2014). Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. *Arch Virol* **159**, 2193–2203.

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

Yang, X., Xie, Y., Raja, P., Li, S., Wolf, J. N., Shen, Q., Bisaro, D. M. & Zhou, X. (2011). Suppression of methylation-mediated transcriptional gene silencing by βC1-SAHH protein interaction during geminivirus-betasatellite infection. *PLoS Pathog* **7**, e1002329.

Downloaded from www.microbiologyresearch.org by
IP: 129.93.101.238
On: Mon, 19 Oct 2015 15:38:13

Virology 482 (2015) 41–50

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Potential linkage between compound microsatellites and recombination in geminiviruses: Evidence from comparative analysis

B. George [1], Ch Mashhood Alam [1], R. Vinoth Kumar, Prabu Gnanasekaran, S. Chakraborty *

Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India

ARTICLE INFO

Article history:
Received 14 November 2014
Returned to author for revisions
16 February 2015
Accepted 5 March 2015

Keywords:
IMEx
dMAX
cSSR%
Relative density
Relative abundance
Hotspot
Recombination
Geminivirus

ABSTRACT

The compound microsatellites consist of two or more individual microsatellites, originate from mutation or imperfection in simple repeat sequences. The reports on systematic analysis of the occurrence, size and density of compound microsatellite (cSSR) types are very rare. Our study indicates that cSSRs are clustered at specific regions in the begomovirus genomes. cSSRs were overrepresented in majority of begomovirus genomes indicating that they might have some functional significance. Further, non-random distribution pattern of cSSR in begomovirus genomes was significantly correlated with the recombination breakpoint positions in the genome. The analysis of cSSR regions in the viral genome indicates the presence of stem loop (hairpin) secondary structure. The significance of these findings in biology of geminiviruses is discussed based on our present understanding of recombination and repetitive DNA. To our knowledge, this is the first analysis suggesting the possible association between recombination and microsatellites in any viral genome.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Geminiviridae, the second largest family of plant viruses, is characterized by twinned icosahedral particles with monopartite or bipartite single stranded DNA (ssDNA) genome replicating via dsDNA intermediate in the host nucleus (Hanley-Bowdoin et al., 2013). Depending on the insect vector and genomic characteristics, geminiviruses are grouped into seven genera: *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus* (Adams et al., 2013). Their genomes are characterized by divergent transcriptional units and a 5′ intergenic region containing origin of rolling circle replication (RCR). For bipartite viruses, the two genomic components share a common region containing the origin of replication and regulatory regions for bidirectional transcription. The functions of geminivirus coded genes have been exhaustively studied (Hanley-Bowdoin et al., 2013; Gilbertson et al., 2003; Lazarowitz and Beachy, 1999; Patil and Fauquet, 2009).

Begomoviruses (transmitted by whiteflies, *Bemisia tabaci* Genn.) constitute the largest genus as evident from the number of species available in genome databases and are considered as one of the major constraints of crop production. Based on genome organization, they are further classified as either monopartite (DNA A – like component) or bipartite (DNA A and DNA B components) encoding 5–7 proteins necessary for viral replication, movement and pathogenesis (Fauquet and Stanley, 2003; Fauquet et al., 2008). In addition, monopartite begomoviruses infecting both crops as well as weeds are often associated with satellite and satellite-like molecules named as betasatellites and alphasatellites, respectively (Briddon et al., 2003; Saunders et al., 2008; Chattopadhyay et al., 2008; Kumar et al., 2012; Singh et al., 2012). Recently, association of satellite molecules has also been reported with bipartite begomoviruses (Romay et al., 2010; Sivalingam and Varma, 2012).

In the evolution of ssDNA viruses, the rate of recombination and nucleotide substitution are closer to that in RNA viruses (Drake 1993; Martin et al., 2011). Besides rolling circle replication (RCR), geminiviruses also employ recombination dependent replication (RDR) strategy for their multiplication (Jeske et al., 2001; Preiss and Jeske, 2003). It is known that recombination plays a vital role in the evolution of begomoviruses (Duffy and Holmes, 2008, 2009; Zhou et al., 1997; Monci et al., 2002; García-Andrés et al., 2006). Several studies have also indicated that high frequency of recombination prevails within recombination hotspots in begomoviruses (Ndunguru et al., 2005; Fauquet et al., 2005; García-Andrés et al., 2007).

* Corresponding author. Tel.: +91 11 26704153 (work).
E-mail address: supriyachakrasls@yahoo.com (S. Chakraborty).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.virol.2015.03.003
0042-6822/© 2015 Elsevier Inc. All rights reserved.

The simple sequence repeats (SSRs, also called as microsatellites and minisatellites) are tandem repetitions of relatively short motifs of DNA. The microsatellites may be considered as either simple or compound depending on the constituent sequences. Simple microsatellites are ubiquitously present in both prokaryotic (Mrazek et al., 2007) and eukaryotic (Toth et al., 2000) genomes, but are relatively rare in viruses; especially in those with small genomes such as *human immunodeficiency virus* Type 1, potyviruses and geminiviruses (Chen et al., 2009; Zhao et al., 2011; George et al., 2012). The compound microsatellites (cSSR) are composed of two or more microsatellites located adjacent to each other. Some of these cSSRs are functionally important in regulating gene expression in yeast (Struhl, 1985) and higher metazoans including human (Curi et al., 2005; Borrmann et al., 2003). Importance of SSRs in meiotic recombination (Schultes and Szostak, 1991; Gendrel et al., 2000; Kirkpatrick et al., 1999; Treco and Arnheim, 1986), evolution of species (Bowcock et al., 1994), genome mapping (Hong et al., 2010), differentiation of viral strains (Deback et al., 2009), studying population genetics and linkage association (Rosenberg et al., 2002; Abdurakhmonov et al., 2005) have been documented. Several examples of simple sequence repeats have been found in menovirus (Duke et al., 1990; Hahn and Palemberg, 1995), vesicular stomatitis virus (Barr et al., 1997), hepatitis C virus (Yamada et al., 1996) and human respiratory syncytial virus (Garcia-Barreno et al., 1994). Changes in length of tri- and hexanucleotide repeats at the hemagglutinin cleavage site in avian influenza virus have been associated with increased virulence (Perdue et al., 1997). However, the significance of cSSR is not available for any viruses.

Jeffreys et al. (1998) reported the minisatellite-associated recombination hotspot in humans. Recently, the association of SSRs with hotspots of meiotic recombination in yeast has been reported (Bagshaw et al., 2008). It is known that certain DNA sequence such as dinucleotide repeats [d(CG)n and d(TG)n] can form left-handed Z-DNA with hotspot activity (Stringer, 1985) as recombination enzymes have preference for those sequences (Blaho and Wells, 1987; Kmiec and Holloman, 1984, 1986). Available literature suggests that repeat sequences might have a role in regulating recombination in both prokaryotes and eukaryotes (Schultes and Szostak, 1991; Gendrel et al., 2000; Kirkpatrick et al., 1999; Treco and Arnheim, 1986; Murphy and Stringer, 1986). In order to understand the control of location of recombination in geminiviruses, it would be relevant to study the association of microsatellite repeats with the recombination hotspots. Specifically, the abundance, distribution and variation of cSSRs would give an insight into the mutational changes and evolution of repeat sequences.

Results

Divergence in number, relative density and relative abundance of compound microsatellites in geminivirus genomes

Scanning of geminivirus genomes revealed the presence of 0–4 cSSRs in each (Table S3). Analysis of relative density and relative abundance indicated that 50 out of 181 begomovirus DNA-A/DNA-A like did not possess imperfect microsatellite (cSSR, Fig. 1 and Table S3). On an average, one cSSR was present in each genome, suggesting a lower abundance of cSSRs in geminiviral genomes. Analysis of null distribution of cSSR motifs in begomovirus genomes indicated that their occurrence was significantly higher than expected by chance (Table S3). Maximum percentage of cSSR (27.27%) was observed in *Tomato yellow margin leaf curl virus* followed by *Dicliptera yellow mottle Cuba virus*, *Euphorbia leaf curl Guangxi virus* (EuLGxCV) and *Tomato yellow leaf curl Sardinia virus*, in which, 25% of SSRs could be classified as cSSRs. The maximum relative density of cSSRs (30.2 bp/kb) was observed in EuLGxCV genome. The relative abundance of cSSR varied among all selected genomes from 0.00/kb (in 50 DNA A/DNA-A like which lack cSSR) to 1.46/kb in EuLGxCV (Table S3). The presence of cSSR could be ascertained in all curtovirus species analyzed except in *Pepper yellow dwarf virus* genome, whereas in mastrevirus, 5 out of 11 genomes were devoid of cSSR. The relative density and abundance of cSSR in curtovirus were 8.3 and 0.54 and for mastrevirus it was 5.2 and 0.24, respectively. The relative density and abundance of cSSR in the only available topocuvirus species were 11.5 and 0.69, respectively (Table S3). The range of relative abundance and density of cSSR vis-a-vis various species are shown in Fig. 1.

Effect of dMAX on the occurrence of compound microsatellites

To determine the impact of dMAX, we selected five distinct species of geminiviruses with or without cSSRs. The percentage of individual SSR being part of a cSSR (cSSRs %) was introduced as a measure to check cSSR variability with increase in dMAX (Fig. 2). It is noteworthy that the dMAX value can only be set between 0 and 50 for IMEx. A linear increase of cSSR frequency with dMAX can be expected if SSRs are randomly distributed. Interestingly, we did not observe a linear increase, suggesting a non-random distribution of

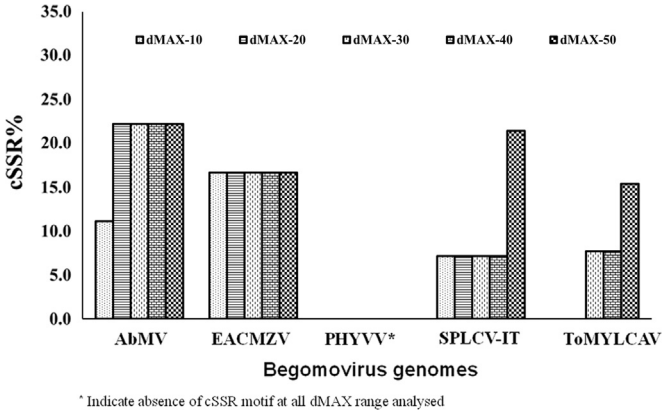


Fig. 2. Effect on cSSR% (percentage of individual microsatellites being part of a compound microsatellite) on varying the dMAX in six different species of begomovirus.

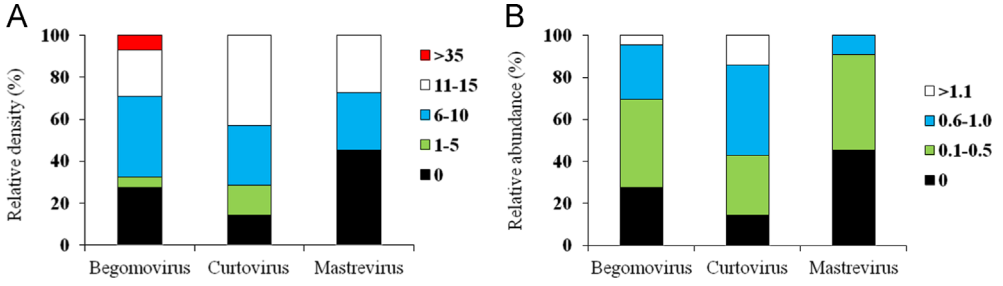


Fig. 1. Relative density (A) and relative abundance density (B) of imperfect microsatellites in geminivirus genomes. Relative density is defined as the total length (bp) contributed by each microsatellite per kb of sequence analyzed whereas relative abundance depicts the number of microsatellites present per kb of the genome.

cSSR in the studied genomes. This could be explained by the fact that all microsatellites being separated by less than the selected dMAX would be accounted as compound microsatellites in the whole genome. In general, genomes lacking cSSR at dMAX 10 also lacked cSSR at higher dMAX tested. These results were similar to that obtained in HIV, potyvirus and *Escherichia coli* genomes (Chen et al., 2011, 2012; Mashhood et al., 2013, Kofler et al., 2008).

Conserved pattern of distribution of cSSR in geminiviral genomes

Out of 131 begomovirus DNA-A/DNA-A like tested, 83% of them possess cSSR either in AC1 or CR regions, whereas, 17% species possess at least one cSSR in the coding region other than AC1 or CR (Fig. 3). In six mastreviruses, except one, all other species contained cSSR in the small intergenic region (Table S3, Fig. 3). In topocuvirus, cSSR was found only in coding regions, whereas in curtoviruses, cSSRs were located in both coding as well as in non-coding regions (Table S3, Fig. 3). Since the above analysis was carried out using one representative from each species, we analyzed distribution patterns of cSSRs in strains or isolates within a particular begomovirus species where more than 5 strains were available in GenBank (Table 1). A total of 456 strains belonging to 38 distinct species of begomovirus DNA-A/DNA-A like were analyzed to identify species specific consensus microsatellite motifs. Among them, except 8 species, others had a conserved consensus species specific motif (Table 1). In 19 species, these consensus motifs were located in the AC1 ORF. Seventy one strains/isolates of 9 distinct begomoviruses shared similar consensus motif (CA)3-x6-(TG)3. Interestingly, *East African cassava mosaic Kenya virus*, *East African cassava mosaic Cameroon virus*, *East African cassava mosaic virus* and *East African cassava mosaic Zanzibar virus* not only shared a conserved consensus motif (TG)*n*-x6-(AT)*y* $n=3$–4, $y=5$–6, but also their location was conserved in AV1 gene. It should be noted that the conserved motif and conserved location of cSSR could be due to the common ancestry of these distinct cassava mosaic geminiviruses.

Polymorphism, complexity and presence of consensus motifs in compound microsatellites

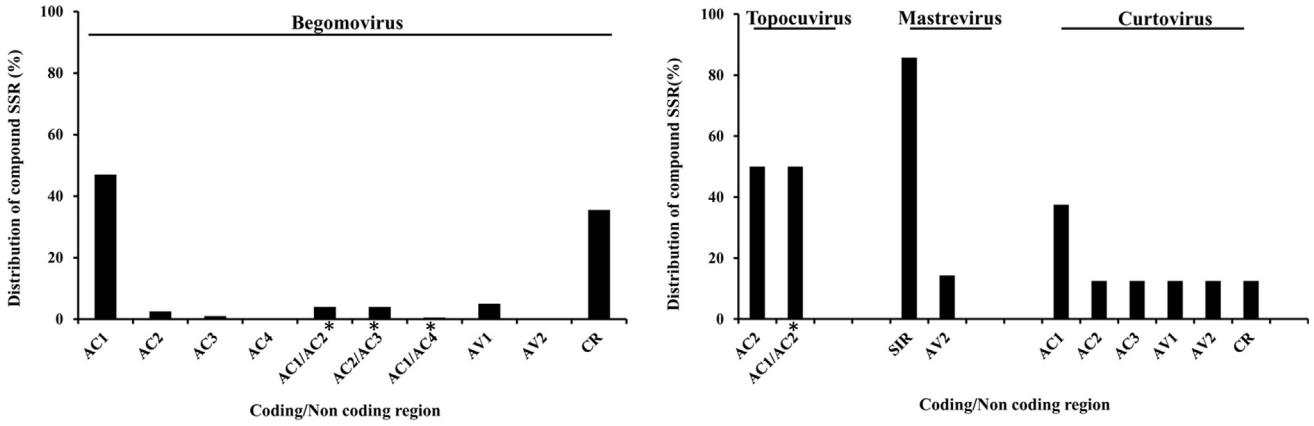
All the extracted motifs were analyzed for the presence of frequently occurring motifs in geminiviral cSSR. Majority of the cSSRs in geminiviruses were composed of very distinct motifs. Among 200 species tested, a self-complimentary motif (CA)3-x6-(TG)3 was the most abundant and present in 28 of begomovirus DNA-A/DNA-A like component (Fig. 4). Motif duplication was observed in cSSRs of 10 begomoviruses (DNA-A/DNA-A like) and in 1 curtovirus genome (Table S3). The most common motifs exhibiting duplication were those consisting of two cSSRs (AT, AG).

Among the analyzed sequences, cSSR with two individual microsatellites (di-SSR) were predominant, while cSSR with three individual microsatellites were rare (Table S3). Notably, cSSR with more than three individual microsatellites were absent. In consistence with prokaryotes, eukaryotes and other viral genomes (Chen et al., 2011, 2012; Kofler et al., 2008), we observed decrease in the number of cSSR with increasing complexity. Secondary structure analysis of region encompassing microsatellite indicated that cSSR motif in the genomes has the potential to form a stem-loop/hairpin structure (Fig. S1). The polymorphic microsatellites, species consensus motifs and their location within geminiviral genome are provided in Table 1. Analysis of cSSR motifs in *East African cassava mosaic virus* indicated that substitution and insertions and deletion events (Indel) resulted in change of motif's length which might lead to change in structure of hairpin loop. Such Indel event in one strain of *East African cassava mosaic virus* resulted in frame shift in AV1 gene (Fig. S1).

Relationship between cSSR motifs and recombination events

To study the relationship, if any, between the presence of microsatellite motif and the recombination breakpoint, we analyzed the recombination pattern in 181 begomovirus sequences. The results indicated that the majority (57.46%) of begomovirus species possessing cSSR motif were also recombinants while 14.36% lacked both motif as well as recombination (dataset 1 in Fig. 5, Table S4). In 17.68% of begomovirus species, motifs were present but no recombination was detected while in 10.16% species, cSSR motif was absent, but the recombination was detected (dataset 1 in Fig. 5, Table S4). In order to find out the correlation between occurrence of cSSR and the type of viral genome, we further categorized them as monopartite (131/181 genomes) and bipartite (50/181 genomes) viruses. It was evident that 64.1% of monopartite and 32% of bipartite begomoviruses were recombinant in nature and possessed cSSR motif(s) (dataset 1 in Fig. 5 and Table S4). Our results suggest that there could be a possible link between presence of cSSR motif and recombination among 181 selected species of begomoviruses.

Further, putative parental sequences of the 181 species of begomoviruses were identified by RDP analysis. In order to validate that whether such a correlation is also present in a separate dataset consisting of the 'putative parental sequences', we analyzed the presence of cSSR motif and recombination among these 'putative parental sequences' (dataset 2 in Fig. 5, Table S4). Results



***Region AC1/AC2, AC2/AC3 or AC1/AC4 are overlapping region between genes**
cSSR motifs located in the overlapping were seperatly counted from those cSSRs which were present in the nonoverlapping regions of begomovirus ORFs

Fig. 3. Differential distribution of cSSRs in coding and non-coding regions of geminivirus genomes.

Table 1
Occurrence of species specific consensus motif in distinct species of begomoviruses.

Begomoviruses (*a*/*b*)*	Consensus motif	Location	*c***
African cassava mosaic virus (11/2)	(CA)3-x6-(TG)3	AC1	9
Ageratum yellow vein virus (9/3)	(CA)3-x6-(TG)3	AC1	4
Ageratum enation virus (20/1)	AGA)3-x1-(A)6	AC1	19
Alternanthera yellow vein virus (4/4)	Absent	–	–
Bean golden yellow mosaic virus (7/5)	Absent	–	–
Chilli leaf curl virus (6/0)	(CA)3-x6-(TG)3	AC1	5
Chino del tomate virus (5/0)	(CTT)3-x5-(AG)3	AC1	4
	(CG)3-x1-(T)6	CR	5
Cotton leaf curl Gezira virus (7/6)	Absent	–	–
Cotton leaf curl Multan virus (16/1)	(CG)3-x1-(T)*n* (*n*=6–8)	CR	15
	(T)6-x8-(GT)3 or (AT)3-x-1-(T)7	AC1	12
Dicliptera yellow mottle virus (7/3)	(TCG)3-x8-(CT)3-x6-(CT)3	AC2/AC3	2
East African cassava mosaic Cameroon virus (6/0)	(TG)*n*-x6-(AT)*y* *n*=3–4 and *y*=5–6	AV1	6
East African cassava mosaic Kenya virus (14/0)	(CA)3-x6-(TG)3 or (CA)3-x6-(TG)3-x5-(AG)3	AC1	13
	(TG)3-x6-(AT)*y* (*y*=5–6)	AV1	5
East African cassava mosaic virus (52/27)	(TG)3-x6-(AT)5	AV1	24
East African cassava mosaic Zanzibar virus (15/0)	(TG)3-x6-(AT)5	AV1	15
Euphorbia yellow mosaic virus (11/0)	Absent	–	–
Honeysuckle yellow vein mosaic virus (6/1)	(CT)3-x-3-(TCT)3 or (CTT)3-x5-(AG)3	AC2 or AC1	4
Honeysuckle yellow vein virus (5/1)	(G)7-x4-(G)6 or (CTT)3-x5-(AG)3	AC1	4
Mungbean yellow mosaic India virus (28/27)	Absent	–	–
Mungbean yellow mosaic virus (10/10)	Absent	–	–
Malvastrum yellow vein Yunnan virus (9/0)	(T)6-x5-(C)7	CR	8
	(CA)3-x6-(TG)3	AC1	9
Macroptilium yellow spot virus(14/0)	(CG)3-x1-(T)6	CR	14
Papaya leaf curl China virus (10/0)	(CA)3-x6-(TG)3	AC1	10
Pepper golden mosaic virus (10/10)	Absent	–	–
Radish leaf curl virus (7/7)	Absent	–	–
Squash leaf curl China virus(8/3)	(CG)3-x1-(T)*n* (*n*=6–8)	CR	5
Sri Lankan cassava mosaic virus (9/1)	(CA)3-x6-(TG)3	AC1	8
Stachytarpheta leaf curl virus (5/0)	(CA)3-x6-(TG)3	AC1	5
Siegesbeckia yellow vein virus (6/0)	(CA)3-x6-(TG)3	AC1	6
Sida micrantha mosaic virus (16/0)	(CG)3-x5–6-(C)7–12	CR	14
Tobacco leaf curl Yunnan virus (5/0)	(CTT)3-x4-(GA)3	AC1	5
	(AT)3-x-1-(T)6	AC1	4
Tomato leaf curl Bangalore virus (7/4)	(CT)3-x0-(AC)3	AC1	3
Tomato leaf curl New Delhi virus (20/6)	(TA)3-x5-(CT)3	AC2/AC3	14
Tomato leaf curl Taiwan virus (10/0)	(CA)3-x6-(TG)3	AC1	2
Tomato severe leaf curl virus (5/3)	(TG)4-x7-(CA)3	CR	13
Tomato yellow leaf curl China virus (17/1)	(CG)3-x1-(T)8-x8-(C)7 or (CG)3-x1-(T)*n* (*n*=6–8)	CR	8
Tomato yellow leaf curl Sardinia virus (8/0)	(AGA)3-x1-(A)6	AC1	8
Tomato yellow leaf curl Thailand virus (8/4)	(AT)3-x6-(CT)3	AC2/AC3	**4**
Tomato yellow leaf curl virus (43/1)	(AT)3-x-1-(T)6	AV1	42

*a/b indicates number of strains per species used/number of strains lacking cSSR; c** denotes number of strains with consensus motif; Absent indicates absence of consensus motif in the species.

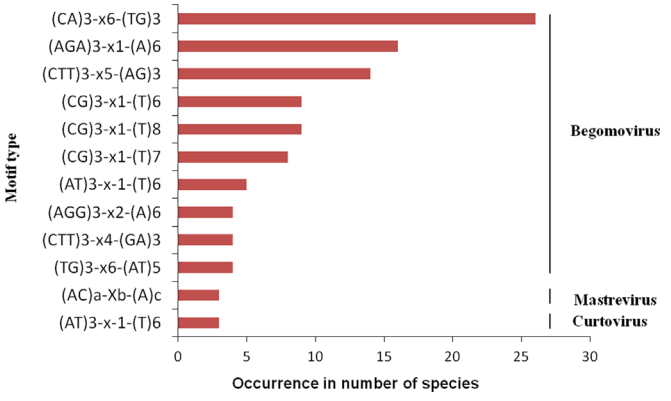


Fig. 4. Frequency of common and abundant motifs among different geminivirus species. A total of 181 species of begomoviruses, 11 species of mastreviruses and 7 species of curtoviruses were analyzed to identify common and abundant motifs.

indicated that 61.24% of parental sequence had both cSSR motif and recombination, whereas 7.75% of these sequences neither possessed cSSR nor recombination. However, 31% of the sequences possessed either cSSR motif or recombination event (Fig. 5). Based on the results obtained at species level, a possible link between cSSR motif and recombination events has been postulated. It is relevant to mention here that the above datasets (1 and 2) represent the association of cSSR and recombination at species level; however, the existence of this correlation between strains of a begomovirus species is yet to be ascertained. Therefore, we considered another dataset (dataset 3 in Table S4) comprising the 132 strains from 7 distinct begomovirus species (both monopartite and bipartite). It is worth noting that all the strains of *Ageratum enation virus*, *East Africa cassava mosaic Kenya virus*, *Tomato yellow leaf curl virus* (TYLCV) and *Sida micrantha mosaic virus* possessed species specific cSSR motif and were recombinant in nature. Similarly, the strains of both *Mungbean yellow mosaic virus* and *Pepper golden mosaic virus* did not possess either cSSR motif or detectable recombination event. Out of 20 strains of *Mungbean yellow mosaic India virus* (MYMIV) analyzed, 11 strains possessed neither cSSR motif nor recombination event. We also observed exception in a few strains of MYMIV where recombination event could be detected in the absence of cSSR motif. Taken together, results from all the 3 datasets indicate a possible linkage between cSSR motifs and recombination in begomoviruses at both species and strain level.

In addition, we have tested the statistical correlation between the occurrence of cSSR and recombination breakpoints in 181

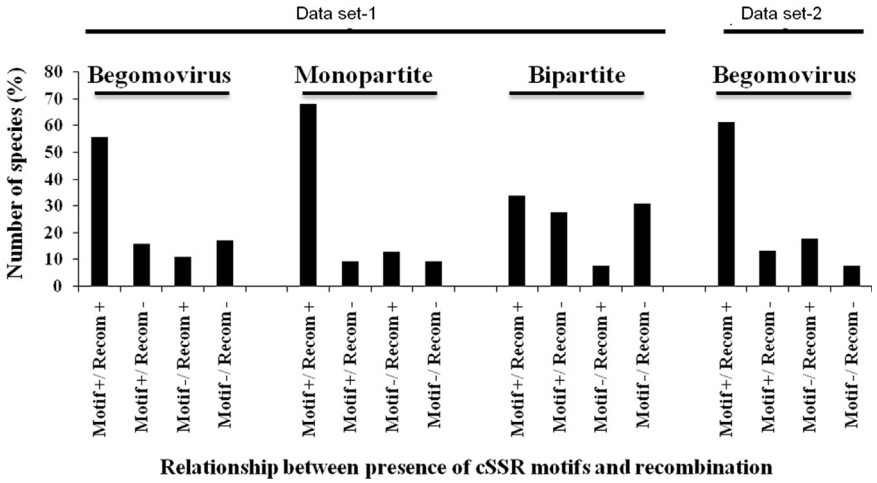


Fig. 5. Correlation between cSSR motif and recombination in DNA A genomes of begomoviruses. Dataset 1 includes 181 DNA-A/DNA-A like genomes of begomoviruses, whereas dataset 2 includes putative parental sequences of 181 begomoviral genomes mentioned in dataset 1.

begomoviruses using Chi-square test. After performing recombination analysis for all the begomovirus genomes taken in our study, parameter '1' was assigned for the presence of motif or recombination and '0' for the absence of motif or recombination. The *P* value for correlation of motif and recombination for DNA A was detected to be $< 10^{-4}$, suggesting a significant correlation between the presence of cSSR motifs and recombination events. The *P*-value for the correlation between the presence of motif and recombination was $< 10^{-4}$ in monopartite DNA-A like genome suggesting significant correlation while that in bipartite genome of DNA A was 0.51 suggesting non-significant correlation. Further, this correlation between cSSR motif and recombination among 181 selected begomovirus species was analyzed quantitatively by one way ANOVA and was found to be highly significant (*P*-value, $< 10^{-4}$). In addition, cSSR motifs as well as recombination breakpoints identified in begomovirus species were mapped onto the linearized genome (Fig. 6a–c). The result indicated that majority of the cSSRs were located in the recombination hotspot regions. In addition, we have also analyzed the physical proximity of the cSSRs and the recombination breakpoints by calculating the number of nucleotides between the end of each cSSR and the closest recombination breakpoint in the sequences which possess both cSSR and recombination events (Fig. 7). The results revealed that majority of the analyzed species (49 out of 99 begomoviruses), cSSRs are located within 400 nt from the recombination break points.

Based on the RDP analysis, AC1, AV1 and CR regions are considered to be the recombination hotspots (Fig. 6c). Therefore, we have searched for a consensus recombination event in each of this recombination hotspot regions to assess relationship between the cSSR and RDP-identified recombination events, if any (Table 2). Analysis of four recombination events indicates a strong relationship between the presence of ≥ 1 cSSR in all of the recombinants associated with each recombination event. However, the presence of a specific cSSR could not be established with any particular recombination event (Tables 2 and S3).

Discussion

We have previously identified the distribution of microsatellite (SSR) in geminiviral genomes (George et al., 2012). In the present study, we further extended our efforts and analyzed the distribution pattern of compound microsatellites (cSSRs) in diverse geminiviral genomes. cSSRs were over-represented in majority of begomovirus genomes indicating that they might have some functional significance. Interestingly, cSSRs are prominently localized in specific coding regions (Rep) and in the intergenic region of the genome. Insertions and deletion events (Indel) in cSSRs within ORFs were not observed except in one case, where it was located at AV1 gene and resulted in frame-shift.

The cSSR% varies between eukaryotes, prokaryotes and viruses. Eukaryotes show higher frequency of cSSR (4–25%, Kofler et al., 2008), which is very low in prokaryotes (1.75–2.85%, Chen et al., 2011). However, in viruses like HIV-1, the frequency of occurrence ranges between 0% and 24% (Chen et al., 2012). Frequency of cSSR in geminiviruses was similar to the range found in HIV (0–25%, Table S3). The reason behind similar frequency of cSSR in genomes of geminivirus and HIV is obscure till date, except that both HIV and geminivirus utilize host replication and recombination machinery. Since replication and recombination are the two processes which regulate the generation and decimation of microsatellite, similar frequency of cSSR in these two eukaryotic viruses could be host-dependent. It will be interesting to study whether viruses sharing the same mode of replication and repair machinery do possess similar frequency of cSSR. All geminiviral species possess similar genome organization, size and base composition; hence, relative density and abundance of cSSR were similar between the genomic components analyzed (Fig. 1).

Around 91.3% of the cSSRs in geminivirus were composed of only two SSR motifs indicating that the individual microsatellites were distributed far from each other and the structure is less complex as compared to other prokaryotic and eukaryotic microsatellites. Eukaryotes, prokaryotes and viruses show varying preferences for microsatellite types and motifs (Chen et al., 2009; Gur-Arie et al., 2000; Karaoglu et al., 2005). This may lead to a bias towards the abundance of one or several motifs in the composition of cSSR. Non-random distribution, identical length of two SSRs forming cSSR and mutually complementary motifs observed in geminiviruses indicated that these cSSRs might originate via recombination. Recombination between homologous SSRs plays an essential role in the generation of self-complementary cSSR (Jakupciak and Wells, 1999). Geminiviruses possess only few cSSR (6.4%) of similar motifs, whereas in eukaryotes this value is higher (90%, Table S3). Our results are in agreement with those observed for HIV (Chen et al., 2012). The rare occurrence of duplication events in viruses could probably be due to high substitution rates in ssDNA viruses compared to those found in RNA viruses (Duffy and Holmes, 2008, 2009; Nawaz-ul-Rehman et al., 2012). Note that majority of strains or isolates of a particular begomovirus species share similar type of motifs and location within the viral genome (Table 1).

DNA viruses being one of the simplest groups of pathogens serve as ideal model system to examine the association of cSSR with recombination events. Notably, position of recombination events has

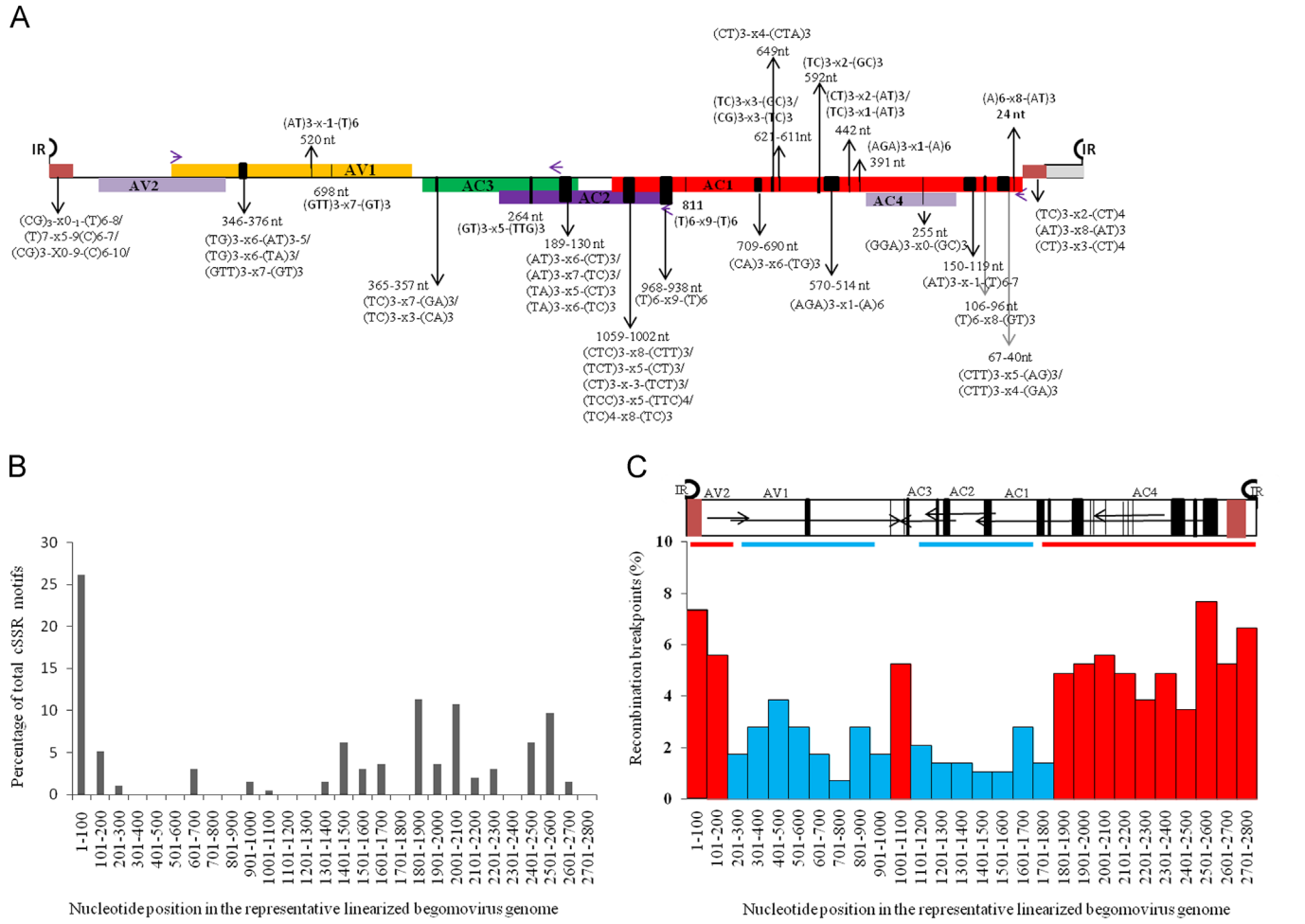


Fig. 6. Genome-wide distribution and location of cSSR motifs and recombination breakpoints in begomoviral genomes. (A) Genome map representing location and sequence of cSSR motifs from 181 begomoviruses analyzed in the study. (B) Distribution of cSSR motifs in the begomoviral genomes. (C) Schematic representation of a linearized begomovirus genome divided into approximately 14 equal segments which represent percentage of total recombination breakpoints analyzed in this study. The red and blue horizontal bars indicate the approximate regions of prominent recombination hotspots and coldspots, respectively. Genomic features: AV2=pre-coat protein; AV1=coat protein gene; AC1=replication initiator protein; AC2=transcriptional activator protein; AC3=replication enhancer protein; AC4=silencing suppressor protein; IR=intergenic region. Black solid boxes represent cSSR location within ORFs. Position of cSSR motifs located in AV1, AC3 corresponds to ATG of AV1, AC3 ORF, respectively. Position of cSSR motifs located in AC1 or AC1/AC2 overlapping region corresponds to ATG of AC1. Position of cSSR motifs in AC2/AC3 overlapping region corresponds to ATG of AC2 ORF. Size of the boxes is to indicate the frequency of occurrence of cSSR at the particular location in begomoviruses in a particular location. Downward pointing arrow represents the conserved cSSR motifs in that particular location. Brown solid boxes indicate cSSR motifs positioned in intergenic (IR) region. Purple arrow represents start codon of ORFs.

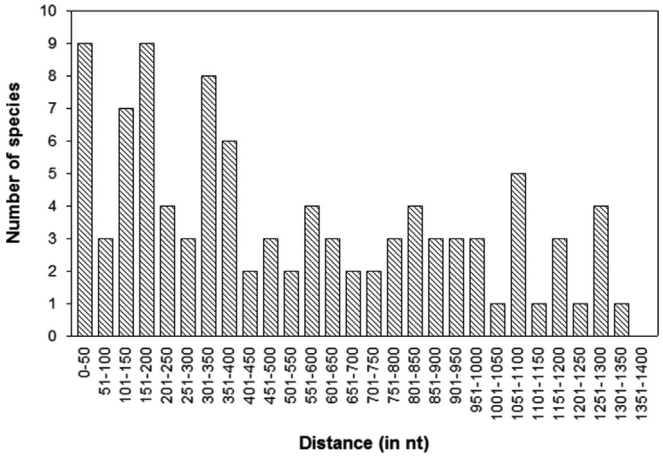


Fig. 7. Frequency of viral genomes that share similar range of distance between the end of cSSR and the nearest recombination break point.

been mapped in the genome of many plant-infecting ssDNA viruses (Monjane et al., 2011; Martin et al., 2011; Prasanna and Rai, 2007; Padidam et al., 1999). The factors that could complicate an association between the cSSR repeats and recombination are likely to be less in geminiviruses as compared to eukaryotes such as yeast and human. For example, the locations of genes and their expression levels are well-characterized in geminiviruses, making it possible to establish link between microsatellites, recombination and transcription.

Our analysis indicates that microsatellite motifs are non-randomly distributed in geminivirus genomes. Approximately 83% of the total cSSR motifs were located in AC1 and/or CR region (Fig. 6, Table S3). A highly significant association (P value, $< 10^{-4}$) was found between the presence of cSSR motifs and recombination in the selected 181 begomoviruses and their putative parental sequences. Further, majority of cSSRs were found to be colocalized with the recombination hotspots. The most obvious factor that possibly contributes to such association is the property of hotspot region in the formation and growth of microsatellite and the regulation of hotspot locations by SSRs. Non-random distribution, strong association of cSSR with recombination events and secondary structure of cSSR provide the evidences for the later mechanism. A large body of evidence shows that a frequent polymerase pause of replication complex occurs when it encounters structures such as hairpins, G-quartets or triplex structures formed by repetitive sequences. Polypurine and polypyrimidine tracts are also known

Table 2
Association between the presence of cSSRs and consensus recombination events in begmoviruses.

Accession no.	Species	Number of cSSR
Recombination breakpoints located within 1934 and 2562 nt		
FJ515747	BYVMV	2
AJ223191	ChaYMV	1
AF139168	DiYMoV	1
AJ558121	EuLCV	1
AF490004	ToCMoV-BA	2
AM236755	MaYMV	2
GU112003	CLCuBV	2
DQ875872	SiYMYuV	2
AM050143	TbLCuCUV	1
AJ608286	ToLCSinV	1
Recombination breakpoints located within 2441 and 93 nt		
HQ189431	HYVV-JR	1
AJ420319	SLCYNV	1
AJ314737	SLCMV-IN	1
AF490004	ToCMoV-BA	3
AF195782	ToLCLV	1
AJ865340	ToLCYTV	1
AJ875157	ToLCNDV	2
AF136222	ToLCPV-A	1
AF084006	ToLCV-Sol	1
X61153	TYLCSV-Sar	4
AY044138	TYLCV-Gez	1
Recombination breakpoints located within 32 and 564 nt		
AJ851005	ALCuV	1
DQ875868	CoYSV	2
AM411424	EuLGxCV	4
AM404179	PepLCLV	1
AF414287	PepLCVMY	2
AB267838	PepLCIV-A	1
AF049336	SiGMV	1
AM238692	SbYVGxV	1
AM183224	SbYVV	1
HM754637	SPLCV-IT	1
AM701756	TbLCZV	1
AJ865339	ToLCMGV-And	1
EU862323	ToLCPKV	2
AF511529	TYLCKaV	2
Recombination breakpoints located within 351 and 1198 nt		
EF165536	CraYVV	2
AF509739	LYMV-VN	1
AJ786711	MaYVYV	2
AJ627904	HgYMV	1
AJ512761	TbLCYnV	1
Z48182	ToLCBV-A	1
AJ558118	ToLCCNV-Bai	1
EU487048	ToLCtV	2
DQ127170	ToLCUV	1
AJ508784	ToSLCV-NI	1
AJ457985	TYLCCNV-Chu	2

to be associated with recombination (Bagshaw et al., 2006). Density of SSRs within genomes positively correlates with regional rates of recombination in *Drosophila* (Schug et al., 1998), yeast (Bagshaw et al., 2008) and *Caenorhabditis elegans* (Denver et al., 2004). Moreover, the role of several repeat types has been experimentally shown to be involved in regulating recombination (Murphy and Stringer, 1986; Napierala et al., 2002, 2004).

We speculate that hairpin structures formed by interrupted repeats in geminivirus genomes possibly act as a recombination signals. These signals might lead to replication pause, double strand breaks and regulation of recombination. Our results thus provide possible link between microsatellites and recombination. Strikingly, occurrence of cSSRs was not found in few of the geminiviral genomes. However, we also observed that out of 27.6% species that lacks cSSR motif, 14.9% also lacks recombination which suggesting a correlation between the two. The occurrence of recombination in 12.7% of species without local microsatellite does not rule out this possible link as regulation of recombination hotspot is heterogeneous in nature (Nishant and Rao, 2006; Kauppi et al., 2004; Petes, 2001).

The association between cSSR motif and recombination events was more prominent in DNA-A like genome of monopartite begomoviruses as compared to DNA A of bipartite begomoviruses (Fig. 4). It may be possible that the recombination frequencies of monopartite and bipartite viruses can be variable due to involvement of several factors that regulate the recombination among viral genomes. For example, both RNA and DNA virus infections are known to induce homologous recombination in plants (Kovalchuk et al., 2003; Richter et al., 2014). It has been suggested that host recombination machinery might be subverted by geminiviruses to promote viral recombination (Richter et al., 2014). Our analysis has established significant correlation between the presence of cSSRs and recombination hotspots predominantly in monopartite begomoviruses. It is possible that the level of involvement of host recombination machinery may influence in either replication or recombination of monopartite and bipartite viruses. However, such speculation needs experimental validation. Few cSSR motifs were also noticed in the recombination coldspot region in the viral genome. Notably, the cSSRs and the recombination events are located within ~400 nucleotides in majority of the begomovirus genomes. Nonetheless, the association of the cSSR motifs (located at a larger distance) with recombination event cannot be ruled out, since the control of hotspot location is shown to be complex and multi-levelled with the involvement of local and distal sequences, transcription factor binding sites and alterations in chromatin structure [reviewed in Nishant and Rao (2006), Kauppi et al. (2004) and Petes (2001)].

The recombination breakpoints that are detectable in naturally occurring ssDNA virus genomes are generally not distributed randomly and occur in clusters within discrete recombination hotspots or within recombination coldspots with lesser frequency (Padidam et al., 1999; Lefeuvre et al., 2009). Mechanistic factors that might influence site-to-site variations in basal recombination rates across these genomes include the locations of replication origins, degrees of sequence similarity between recombining genomes, genomic ssDNA secondary structures, the orientations of genes in relation to directions of rolling circle replication, differential degrees of dsDNA exposure within histone packaged viral mini-chromosomes and ecological as well as epidemiological factors [reviewed by Martin et al. (2011)]. Our results strongly suggest that the pattern of microsatellite distribution within geminiviruses is in accordance with the recombination events. Therefore, we propose that compound microsatellites motif type, length, polymorphism and secondary structure might contribute to the recombination process.

Conclusion

It has been known for long that large number of recombination events is detectable for majority of ssDNA viruses. Nevertheless, the patterns of occurrence of recombination hotspots indicate that interplay between selective and mechanistic processes would determine the general distribution of recombination breakpoints. The comparative analysis of location of microsatellite motif in many geminiviral species emphasizes that in majority of the cases cSSR motifs are located within the recombination hotspot. It may be possible that microsatellites in viral genomes could serve as high affinity binding sites for host proteins which in turn might regulate recombination activity, as has been shown in prokaryotes and eukaryotes. Another possibility is that such repeat sequences might promote crossovers by presenting the recombination apparatus with a specific structure that could be a signal to initiate genetic exchange nearby. Although the present study provides leads to the

possible association between recombination and microsatellites, further experimental evidence will be required to correlate functional role of microsatellite as recombination signals in geminiviruses, as observed in yeast and human.

Materials and methods

Genome sequences of geminiviruses

Sequences from 200 genomes of distinct geminiviruses were selected for analysis of cSSR including 181 DNA-A from distinct species of the genus *Begomovirus*, 11 DNA-A like genome from the genus *Mastrevirus*, seven and one DNA-A like genome from genera *Curtovirus* and *Topocuvirus*, respectively. These sequences were downloaded in FASTA format from the GenBank (http://www.ncbi.nlm.nih.gov) following Fauquet et al. (2008) and the sequences with the accession numbers are provided in Table S1. We relied on the existing annotation ('CDS' features) in differentiating protein-coding and non-coding regions. For identification of species specific consensus microsatellite motifs and microsatellite polymorphism, various strains or isolates within a particular species were considered (Table S2).

Identification of microsatellites in the viral genomes

Imperfect Microsatellite Extractor (IMEx) was used to check the presence of microsatellites in the geminiviral genomes (Mudunuri and Nagarajaram, 2007). Earlier studies on prokaryotic and eukaryotic genomes have focused on microsatellites with lengths of 12 bp (Toth et al., 2000), but due to relatively smaller size of geminivirus genome, cSSR search using these parameters could not yield any microsatellite. Therefore, we detected SSRs and cSSRs using the 'Advance-Mode' of IMEx using the parameters as described in Chen et al. (2012). Other parameters were set as default. The cSSRs were not standardized in order to observe their real composition and evolutionary dynamics. The SSRs were considered as cSSR when the distance separating any two SSRs is less than or equivalent to dMAX (dMAX is the maximum distance between any two adjacent SSR i.e. if dMAX is 10 and two SSRs are present between this distance, then number of cSSR is 1). In order to check the impact of dMAX on identification of cSSR, we varied dMAX value from 10 to 50. The presence of cSSR in each genome was further manually checked and corrected for errors. The microsatellite driven structure was predicted by DNA folding form available within the *M* fold web server (Zuker, 2003).

Estimation of expected number of microsatellites

In order to evaluate whether microsatellites were over- or under-represented in begomovirus genomes, we compared the observed number of microsatellites with the expected number of microsatellites. The expected number of microsatellite composed of M_t (M is motif of the microsatellite with repeat number of t, and its length is L) in a genome of length G was calculated using the formula as given by Dewachter (1981)

$$\text{Exp}\,(M_t) = f(M)^t[1-f(M)][G'(1-f(M))+2L] \quad (1)$$

$$G' = G - tL - 2L + 1 \quad (2)$$

where Exp (M_t) is the expected number of M_t, and $f(M)$ is the probability of M.

Detection of recombination events

Detection of potential recombinant sequences and putative breakpoint events was carried out using several recombination detection methods like RDP, GENECONV, BOOTSCAN, MAXCHI, CHIMAERA, SISCAN and 3SEQ implemented in RDP3 (Martin et al., 2010). The analysis was performed with default setting with a Bonferroni corrected *P*-value cut-off of 0.01. RDP dataset includes the particular sequence, the sequence with which it shared highest percent identity and few sequences to which it possess 75–90% identity. The breakpoint positions and recombinant sequence (s) inferred for every detected potential recombination event were manually checked and adjusted, where necessary using RDP3. Putative parents detected here are not necessarily being the actual parents; however, they could be the most similar ones to the actual parents in the analyzed dataset. The recombination events detected by atleast three independent methods were only considered.

Statistical analysis

Four parameters i.e. the presence or absence of cSSR motif and the presence or absence of recombination events were used to identify co-relation between recombination and cSSR motif. Chi square test was performed using GraphPad Prism version 5 and one way ANOVA was performed using Analyse-it software version 3.76.5.

Acknowledgments

Financial support from the DST-PURSE, UPOE and UGC Resource Networking provided by JNU to SC is gratefully acknowledged. We thankfully acknowledge Prof. S K Goswami, SLS, JNU for critically reading the manuscript and Dr. Manoj Prasad, NIPGR for assistance in statistical analysis (one way ANOVA). We would like to thank the editor and the anonymous reviewers for their constructive comments for the improvement of the manuscript.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2015.03.003.

References

Abdurakhmonov, Y., Abdullaev, A.A., Saha, S., Buriev, Z.T., Arslanov, D., Kuryazov, Z., Rizaeva, S.M., Reddy, U.K., Jenkins, J.N., Abdullaev, A., Abdukarimov, A., 2005. Simple sequence repeat marker associated with a natural leaf defoliation trait in tetraploid cotton. J. Hered. 96, 644–653.

Adams, M.J., King, A.M.Q., Carstens, E.B., 2013. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses. Arch. Virol. 158 (9), 2023–2030.

Bagshaw, A.T.M., Pitt, J.P., Gemmell, N.J., 2006. Association of poly-purine/poly-pyrimidine sequences with meiotic recombination hot spots. BMC Genomics 7, 179.

Bagshaw, A.T.M., Pitt, J.P., Gemmell, N.J., 2008. High frequency of microsatellites in *S. cerevisiae* meiotic recombination hotspots. BMC Genomics 9, 49.

Barr, J.N., Whelan, S.P.J., Wertz, G.W., 1997. cis-acting signals involved in termination of vesicular stomatitis mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. J. Virol. 71, 8718–8725.

Blaho, J.A., Wells, R.D., 1987. Left-handed Z-DNA binding by the recA protein of *Escherichia coli*. J. Biol. Chem. 262, 6082–6088.

Borrmann, L., Schwanbeck, R., Heyduk, T., Seebeck, B., Rogalla, P., Bullerdiek, J., Wisniewski, J.R., 2003. High mobility group A2 protein and its derivatives bind a specific region of the promoter of DNA repair gene ERCC1 and modulate its activity. Nucleic Acids Res. 31, 6841–6851.

Bowcock, A.M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R., Cavalli-Sforza, L.L., 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368, 455–457.

Briddon, R.W., Bull, S.E., Amin, I., Idris, A.M., Mansoor, S., Bedford, I.D., Dhawan, P., Rishi, N., Siwatch, S.S., Abdel-Salam, A.M., Brown, J.K., Zafar, Y., Markham, P.G., 2003. Diversity of DNA beta, a satellite molecule associated with some monopartite begomoviruses. Virology 312, 106–121.

Chattopadhyay, B., Singh, A.K., Yadav, T., Fauquet, C.M., Sarin, N.B., Chakraborty, S., 2008. Infectivity of the cloned components of a begomovirus: DNA beta causing chilli leaf curl disease in India. Arch. Virol. 153 (3), 533–539.

Chen, M., Tan, Z., Jiang, J., Li, M., Chen, H., Shen, G., Yu, R., 2009. Similar distribution of simple sequence repeats in diverse completed human immunodeficiency virus type 1 genomes. FEBS Lett. 583, 2959–2963.

Chen, M., Zeng, G., Tan, Z., Jiang, M., Zhang, J., Zhang, C., Lu, L., Lin, Y., Peng, J., 2011. Compound microsatellites in complete *Escherichia coli* genomes. FEBS Lett. 585, 1072–1076.
Chen, M., Tan, Z., Zeng, G., Zhuotong, Z., 2012. Differential distribution of compound microsatellites in various human immunodeficiency virus type 1 complete genomes. Infect. Genet. Evol. 12, 1452–1457.
Curi, R.A., Oliveira, H.N., Silveira, A.C., Lopes, C.R., 2005. Effects of polymorphic microsatellites in the regulatory region of IGF1 and GHR on growth and carcass traits in beef cattle. Anim. Genet. 36, 58–62.
Deback, C., Boutolleau, D., Depienne, C., Luyt, C.E., Bonnafous, P., Gautheret-Dejean, A., Garrigue, I., Agut, H., 2009. Utilization of microsatellite polymorphism for differentiating herpes simplex virus type 1 strains. J. Clin. Microbiol. 47, 533–540.
Denver, D.R., Morris, K., Kewalramani, A., Harris, K.E., Chow, A., Estes, S., Lynch, M., Thomas, W.K., 2004. Abundance, distribution, and mutation rates of homopolymeric nucleotide runs in the genome of *Caenorhabditis elegans*. J. Mol. Evol. 58, 584–595.
Dewachter, R., 1981. The number of repeats expected in random nucleic acid sequences and found in genes. J. Theor. Biol. 91, 71–98.
Drake, J.W., 1993. Rates of spontaneous mutation among RNA viruses. Proc. Natl. Acad. Sci. USA 90, 4171–4175.
Duffy, S., Holmes, E.C., 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus* (TYLCV). J. Virol. 82, 957–965.
Duffy, S., Holmes, E.C., 2009. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J. Gen. Virol. 90, 1539–1547.
Duke, G.M., Osorio, J.E., Palmenberg, A.C., 1990. Attenuation of Mengo virus through genetic engineering of the 59 noncoding poly(C) tract. Nature 343, 474–476.
Fauquet, C.M., Sawyer, S., Idris, A.M., Brown, J.K., 2005. Sequence analysis and classification of apparent recombinant begomoviruses infecting tomato in the nile and mediterranean basins. Phytopathology 95, 549–555.
Fauquet, C.M., Stanley, J., 2003. Geminivirus classification and nomenclature: progress and problems. Ann. Appl. Biol. 142 (2), 165–189.
Fauquet, C.M., Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M., Zhou, X., 2008. Geminivirus strain demarcation and nomenclature. Arch. Virol. 153, 783–821.
García-Andrés, S., Monci, F., Navas-Castillo, J., Moriones, E., 2006. Begomovirus genetic diversity in the native plant reservoir *Solanum nigrum*: evidence for the presence of a new virus species of recombinant nature. Virology 350 (2), 433–442.
García-Andrés, S., Tomás, D.M., Sánchez-Campos, S., Navas-Castillo, J., Moriones, E., 2007. Frequent occurrence of recombinants in mixed infections of Tomato yellow leaf curl disease-associated begomoviruses. Virology 365, 210–219.
Garcia-Barreno, B., Delgado, T., Melero, J.A., 1994. Oligo(A) sequences of human respiratory syncytial virus G protein gene: assessment of their genetic stability in frameshift mutants. J. Virol. 68, 5460–5468.
Gendrel, C.G., Boulet, A., Dutreix, M., 2000. CA/GT(n) microsatellites affect homologous recombination during yeast meiosis. Genes Dev. 14 (10), 1261–1268.
George, B., Mashhood, A.C., Jain, S.K., Sharfuddin, C., Chakraborty, S., 2012. Differential distribution and occurrence of simple sequence repeats in diverse geminivirus genomes. Virus Genes 45 (3), 556–566.
Gilbertson, R.L., Sudarshana, M., Jiang, H., Rojas, M.R., Lucas, W.J., 2003. Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: insights into DNA trafficking. Plant Cell 15 (11), 2578–2591.
Gur-Arie, R., Cohen, C.J., Eitan, Y., 2000. Simple sequence repeats in *Escherichia coli*: abundance, distribution, composition, and polymorphism. Genome Res. 10, 62–71.
Hahn, H., Palemberg, A.C., 1995. Encephalomyocarditis viruses with short poly (C) tracts are more virulent than their mengovirus counterparts. J. Virol. 69, 2697–2699.
Hanley-Bowdoin, L., Eduardo, R.B., Dominique, R., Shahid, M., 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. Nat. Rev. Microbiol. 11, 777–788.
Hong, Y., Chen, X., Liang, X., Liu., H., Zhou, G., Li, S., Wen, S., Holbrook, C.C., Guo, B., 2010. A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. BMC Plant Biol. 27, 10–17.
Jakupciak, J.P., Wells, R.D., 1999. Genetic instabilities in (CTG.CAG) repeats occur by recombination. J. Biol. Chem. 274, 23468–23479.
Jeffreys, J., Murray, J., Neumann, R., 1998. High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. Mol. Cell 2, 267–273.
Jeske, J., Lütgemeier, M., Preiss, W., 2001. DNA forms indicate rolling circle and recombination-dependent replication of *Abutilon* mosaic virus. EMBO J. 20, 6158–6167.
Karaoglu, H., Lee, C.M., Meyer, W., 2005. Survey of simple sequence repeats in completed fungal genomes. Mol. Biol. Evol. 22, 639–649.
Kauppi, L., Jeffreys, A.J., Keeney, S., 2004. Where the crossovers are: recombination distributions in mammals. Nat. Rev. Genet. 5, 413–424.
Kirkpatrick, D.T., Wang, Y.H., Dominska, M., Griffith, J.D., Petes, T.D., 1999. Control of meiotic recombination and gene expression in yeast by a simple repetitive DNA sequence that excludes nucleosomes. Mol. Cell. Biol. 19, 7661–7671.
Kmiec, E.B., Holloman, W.K., 1984. Synapsis promoted by Ustilago rec1 protein. Cell 36, 593–598.
Kmiec, E.B., Holloman, W.K., 1986. Homologous pairing of DNA molecules by Ustilago rec1 protein is promoted by sequences of Z-DNA. Cell 44, 545–554.
Kofler, R., Schlotterer, C., Luschutzky, E., Lelley, T., 2008. Survey of microsatellite clustering in eight fully sequenced species sheds light on the origin of compound microsatellites. BMC Genomics 9, 612.
Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., Hohn, B., 2003. Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature 423, 760–762.
Kumar, R.V., Singh, A.K., Chakraborty, S., 2012. A new monopartite begomovirus species, Chilli leaf curl Vellanad virus, and associated betasatellites infecting chilli in the Vellanad region of Kerala, India. New Dis. Rep. 25, 20.
Lazarowitz, S.G., Beachy, R.N., 1999. Viral movement proteins as probes for intracellular and intercellular trafficking in plants. Plant Cell 11, 535–548.
Lefeuvre, P., Lett, J.M., Varsani, A., Martin, D.P., 2009. Widely conserved recombination patterns among single-stranded DNA viruses. J. Virol. 83, 2697–2707.
Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26 (19), 2462–2463.
Martin, D.P., Lemey, P., Posada, D., 2011. Analysing recombination in nucleotide sequences. Mol. Ecol. Resour. 11, 943–955.
Mashhood, A.C., George, B., Sharfuddin, C., Jain, S.K., Chakraborty, S., 2013. Occurrence and analysis of imperfect microsatellites in diverse potyvirus genomes. Gene 521 (2), 238–244.
Monjane, A.L., van der Walt, E., Varsani, A., Rybicki, E.P., Martin, D.P., 2011. Recombination hotspots and host susceptibility modulate the adaptive value of recombination during maize streak virus evolution. BMC Evol. Biol. 11, 350.
Monci, F., Sánchez-Campos, S., Navas-Castillo, J., Moriones, E., 2002. A natural recombinant between the geminiviruses, Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303, 317–326.
Mrazek, J., Guo, X., Shah, A., 2007. Simple sequence repeats in prokaryotic genomes. Proc. Natl. Acad. Sci. USA 104, 8472–8477.
Mudunuri, S.B., Nagarajaram, H.A., 2007. IMEx: imperfect microsatellite extractor. Bioinformatics 23, 1181–1187.
Murphy, K.E., Stringer, J.R., 1986. RecA independent recombination of poly[d(GT)-d(CA)] in pBR322. Nucleic Acids Res. 14, 7325–7340.
Napierala, M., Parniewski, P., Pluciennik, A., Wells, R.D., 2002. Long CTG.CAG repeat sequences markedly stimulate intramolecular recombination. J. Biol. Chem. 277, 34087–34100.
Napierala, M., Dere, R., Vetcher, A., Wells, R.D., 2004. Structure-dependent recombination hot spot activity of GAA.TTC sequences from intron 1 of the Friedreich's ataxia gene. J. Biol. Chem. 279, 6444–6454.
Nawaz-ul-Rehman, M.S., Briddon, R.W., Fauquet, C.M., 2012. A melting pot of old world begomoviruses and their satellites infecting a collection of Gossypium species in Pakistan. PLoS One 7 (8), e40050.
Ndunguru, J., Legg, J.P., Aveling, T.A., Thompson, G., Fauquet, C.M., 2005. Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. Virol. J. 22, 2–21.
Nishant, K.T., Rao, M.R., 2006. Molecular features of meiotic recombination hot spots. Bioessays 28, 45–56.
Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265 (2), 218–225.
Patil, B.L., Fauquet, C.M., 2009. Cassava mosaic geminiviruses: actual knowledge and perspectives. Mol. Plant Pathol. 10, 685–701.
Perdue, M.L., Garcia, M., Senne, D., Fraire, M., 1997. Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses. Virus Res. 49, 173–186.
Petes, T.D., 2001. Meiotic recombination hot spots and cold spots. Nat. Rev. Genet. 2, 360–369.
Prasanna, H.C., Rai, M., 2007. Detection and frequency of recombination in tomato-infecting begomoviruses of South and Southeast Asia. Virol. J. 26, 4–111.
Preiss, W., Jeske, H., 2003. Multitasking in replication is common among geminiviruses. J. Virol. 77, 2972–2980.
Richter, K.S., Kleinow, T., Jeske, H., 2014. Somatic homologous recombination in plants is promoted by a geminivirus in a tissue-selective manner. Virology 452, 287–296.
Romay, G., Chirinos, D., Geraud-Pouey, F., Desvies, C., 2010. Association of an atypical alphasatellite with a bipartite New World begomovirus. Arch. Virol. 155, 1843–1847.
Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., Feldman, M.W., 2002. The genetic structure of human populations. Science 298, 2381–2385.
Saunders, K., Briddon, R.W., Stanley, J., 2008. Replication promiscuity of DNA-b satellites associated with monopartite begomoviruses; deletion mutagenesis of the Ageratum yellow vein virus DNA-b satellite localizes sequences involved in replication. J. Gen. Virol. 89, 3165–3172.
Schug, M.D., Hutter, C.M., Noor, M.A., Aquadro, C.F., 1998. Mutation and evolution of microsatellites in *Drosophila melanogaster*. Genetica 102–103 (1–6), 359–367.
Schultes, N.P., Szostak, J.W., 1991. A poly (dA.dT) tract is a component of the recombination initiation site at the ARG4 locus in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 11, 322–328.

Singh, A.K., Chattopadhyay, B., Chakraborty, S., 2012. Biology and interactions of two distinct monopartite begomoviruses and betasatellites associated with radish leaf curl disease in India. Virol. J. 9, 43.
Sivalingam, P.N., Varma, A., 2012. Role of betasatellite in the pathogenesis of a bipartite begomovirus affecting tomato in India. Arch. Virol. 157, 1081–1092.
Stringer, J.R., 1985. Recombination between poly [d (GT).d(CA)] sequences in simian virus 40-infected cultured cells. Mol. Cell. Biol. 5, 1247–1259.
Struhl, K., 1985. Nucleotide sequence and transcriptional mapping of the yeast pet56-his3-ded1 gene region. Nucleic Acids Res. 13, 8587–8601.
Toth, G., Gaspari, Z., Jurka, J., 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 10, 967–981.
Treco, D., Arnheim, N., 1986. The evolutionarily conserved repetitive sequence d(TG. AC)n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. Mol. Cell. Biol. 6, 3934–3947.
Yamada, N., Tanihara, K., Takada, A., Yorihuzi, T., Tsitsumi, M., Shimomura, H., Tsuji, T., Date, T., 1996. Genetic organization and diversity of the 39 noncoding region of the hepatitis C virus genome. Virology 223, 255–261.
Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G.W., Robinson, D.J., Harrison, B.D., 1997. Evidence that DNA-A of a Geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. J. Gen. Virol. 78, 2101–2111.
Zhao, X., Tan, Z., Feng, H., Yang, R., Li, M., Jiang, J., Shen, G., Yu, R., 2011. Microsatellites in different Potyvirus genomes: survey and analysis. Gene 488 (1–2), 52–56.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.

Virus Genes (2014) 48:334–342
DOI 10.1007/s11262-013-1031-y

Host-specific adaptation of diverse betasatellites associated with distinct Indian tomato-infecting begomoviruses

Punam Ranjan · A. K. Singh · R. Vinoth Kumar · Saumik Basu · S. Chakraborty

Received: 10 October 2013 / Accepted: 21 December 2013 / Published online: 3 January 2014
© Springer Science+Business Media New York 2014

Abstract Tomato leaf curl viruses cause major crop loss hindering tomato cultivation worldwide. The 'Old World' begomoviruses are often associated with circular ssDNA satellite molecules called betasatellites. In the present study, replication compatibility of five different betasatellites with three distinct Indian tomato-infecting begomoviruses representing each of a monopartite, a mono-bipartite and a bipartite begomoviruses was studied. All the betasatellites could be trans-replicated by the begomoviruses in *Nicotiana benthamiana* plants, however, not uniformly in tomato. *Tomato leaf curl Joydebpur betasatellite*—Magrahat could not induce symptom with any of these begomoviruses in tomato, whereas only *Tomato leaf curl Gujarat virus* could trans-replicate *Radish leaf curl betasatellite* in this plant species. However, none of the betasatellites were found to complement the movement function of a bipartite begomovirus in tomato. Unlike tomato, the trans-replication/maintenance of betasatellites by these begomoviruses in *N. benthamiana* could be due to its compromised host defence machinery. Co-infection of betasatellites with these viruses did not enhance the helper virus accumulation, but the incubation period was reduced. The possible factors involved in this host-driven adaptability of betasatellites were also discussed.

Keywords *Begomovirus* · Betasatellite · Replication · Host range

Electronic supplementary material The online version of this article (doi:10.1007/s11262-013-1031-y) contains supplementary material, which is available to authorized users.

P. Ranjan · R. V. Kumar · S. Basu · S. Chakraborty (✉)
Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India
e-mail: supriyachakrasls@yahoo.com

Present Address:
A. K. Singh
School of Life Sciences, Central University of Gujarat, Gandhinagar 382 030, Gujarat, India

Introduction

Members of the family, *Geminiviridae*, are the second largest plant-infecting group of viruses encapsidating circular ssDNA genome(s) of ~2.7 kb in size. *Begomovirus*, the largest genus in this family infects wide range of vegetable crops worldwide [1, 2]. Based on the number of encapsidated genomic components, it has been categorised either as monopartite or bipartite. In 'New World' monopartite begomoviruses (such as *Tomato leaf deformation virus*), DNA A component itself was sufficient to induce symptoms systemically [3]. However, the 'Old World' begomoviruses are often found to be associated with novel symptom modulating satellite and satellite-like molecules, named as betasatellites and alphasatellites, respectively [4–7]. In India, tomato leaf curl disease is caused by various mono, mono-bi and/or bipartite begomoviruses along with diverse betasatellites. *Tomato leaf curl Gujarat virus* (ToLCGV), a mono-bipartite virus was reported to be associated with a DNA B molecule in India [8]. However, *Tomato leaf curl New Delhi* (ToLCNDV) is strictly a bipartite species where DNA B is essential for the systemic infection [9]. *Tomato leaf curl Patna virus* (ToLCPaV) is associated with DNA β, however, the betasatellite is not essential for the systemic infection, similar to the case of *Tobacco curly shoot virus* associated betasatellite [10, 11].

Betasatellites are typically half the size of their helper begomoviruses (~1350 bp) encompassing an adenine-rich

region, satellite conserved region (SCR) of a predicted hairpin structure with a loop and a single ORF (named as βC1) in the complementary sense strand. These satellites are highly diverse (<78 % nucleotide sequence identity amongst them) and are associated with the pathogenesis [4–6]. Betasatellite-encoded βC1 has been reported to be a suppressor of post-transcriptional gene silencing (PTGS) and a determinant of pathogenicity [6, 12, 13]. Recent reports have emphasised on the role of A-rich region and SCR in the trans-replication by diverse begomoviruses and various ORFs encoded by begomoviruses for their maintenance [14–16]. In addition, the more relaxed way of origin recognition by diverse begomoviruses-encoded Rep protein has helped betasatellite trans-replication more efficient [16].

Betasatellites and DNA B rely on the interaction between the helper component-encoded Rep and iteron and iteron-like sequences, respectively, for their replication and maintenance [14, 16–18]. A betasatellite reported from cotton has been demonstrated to be trans-replicated by ToLCNDV; however, in turn, βC1 was complementing the defective movement function of the helper virus [19, 20]. In addition, the role of G-box and other promoter elements in the binding of plant nuclear factors regulating the expression of βC1 has been elucidated [21]. In the present study, trans-replication and maintenance of different betasatellites by three distinct tomato-infecting begomoviruses were analysed. Results pinpoint the host adaptation in trans-replication and/or maintenance of betasatellites by diverse begomoviruses. The possible factors involved in determining pathogenicity and functional compatibility of betasatellites with helper viruses and hosts were also discussed.

Materials and methods

Selection of clones and construction of infectious partial tandem dimers

A 450 bp *Bam* HI (304)-*Kpn* I (1282) fragment of *Pepper leaf curl Bangladesh virus*-Magrahat (HM007111) was cloned in pCAMBIA2301 vector. Subsequently, the full-length fragment digested with *Bam*HI was mobilised to obtain a partial tandem repeat construct (indicated as MgA). Likewise, a *Eco*RI (511)-*Kpn*I (1308) fragment of ~0.8 kb containing intergenic region of *Tomato leaf curl Joydebpur betasatellite* from Magrahat (HM007112) region was cloned in pCAMBIA2301 vector followed by mobilisation of full-length betasatellite of 1.3 kb at *Kpn*I site to generate a partial tandem repeat (referred as Mgβ).

Infectious clones of a monopartite (EU862323-ToLCPaV: PA), a mono-bipartite (ToLCGV: AY190290-VA for DNA A, AY190291-VB for DNA B) and a bipartite (ToLCNDV: U15015-NA for DNA A, U15016-NB for DNA B) Indian begomoviruses were selected for the betasatellite trans-replication studies [8–10]. Dhanbad and Ramgarh isolates of *Tomato yellow leaf curl Thailand betasatellite* (EU573713-Dhnβ, GQ994097-Rmβ), *Tomato leaf curl Patna betasatellite* (EU862324-Pβ), *Radish leaf curl betasatellite* (EF175734-Rlβ) and *Tomato leaf curl Joydebpur betasatellite*—Magrahat (HM007112-Mgβ) were the betasatellites used for studying the trans-replication analysis with distinct helper viruses [7, 10, 22].

Agro-inoculation for trans-complementation studies

Partial tandem repeat constructs of MgA and Mgβ were mobilised into *Agrobacterium tumefaciens* EHA105. Other previously reported infectious clones of begomoviruses and betasatellites were also used for the inoculation [7–10, 22]. Three-leaf stage tomato cv. DVRT-1 and five-leaf stage *Nicotiana benthamiana* seedlings were used for inoculation using *Agrobacterium* cells harbouring various partial tandem repeat constructs. Agro-inoculated plants were maintained under controlled insect-free glass house for recording the symptoms till 45 days post agro-inoculation (dpi).

Southern hybridisation for detecting viral DNAs

For assessing the accumulation of viral DNA from the systemic leaves of agro-inoculated plants, the genomic DNA was extracted from the systemic leaves at 25 dpi [23]. For Southern hybridisation, 8 µg of total DNA was fractionated on 0.8 % agarose gel and transferred to nylon membrane (Millipore) according to a standard procedure [24]. To analyse the level of viral DNA accumulation, the blots were hybridised separately with probes specific for DNA A, DNA B and DNA β. For detecting Mgβ, PCR-amplified βC1 fragment was used as probes. Specific probes to detect VA, VB, NA, NB, PA, Pβ, Dhnβ, Rβ and Rmβ were prepared as reported previously [7, 10, 22, 25]. Probes were labelled with [α^{32}P] dCTP using random oligonucleotide-primer synthesis and the bands were detected using Phosphor Imager (Fuji Film).

Analysis of betasatellite sequences

Betasatellites used in this study were aligned using EBI Clustal W tool and the position of conserved sequences and inverted repeats were identified. Per cent identity between the full-length betasatellite sequences, their SCR and βC1 amino acid sequences were derived by EBI sequence alignment tool.

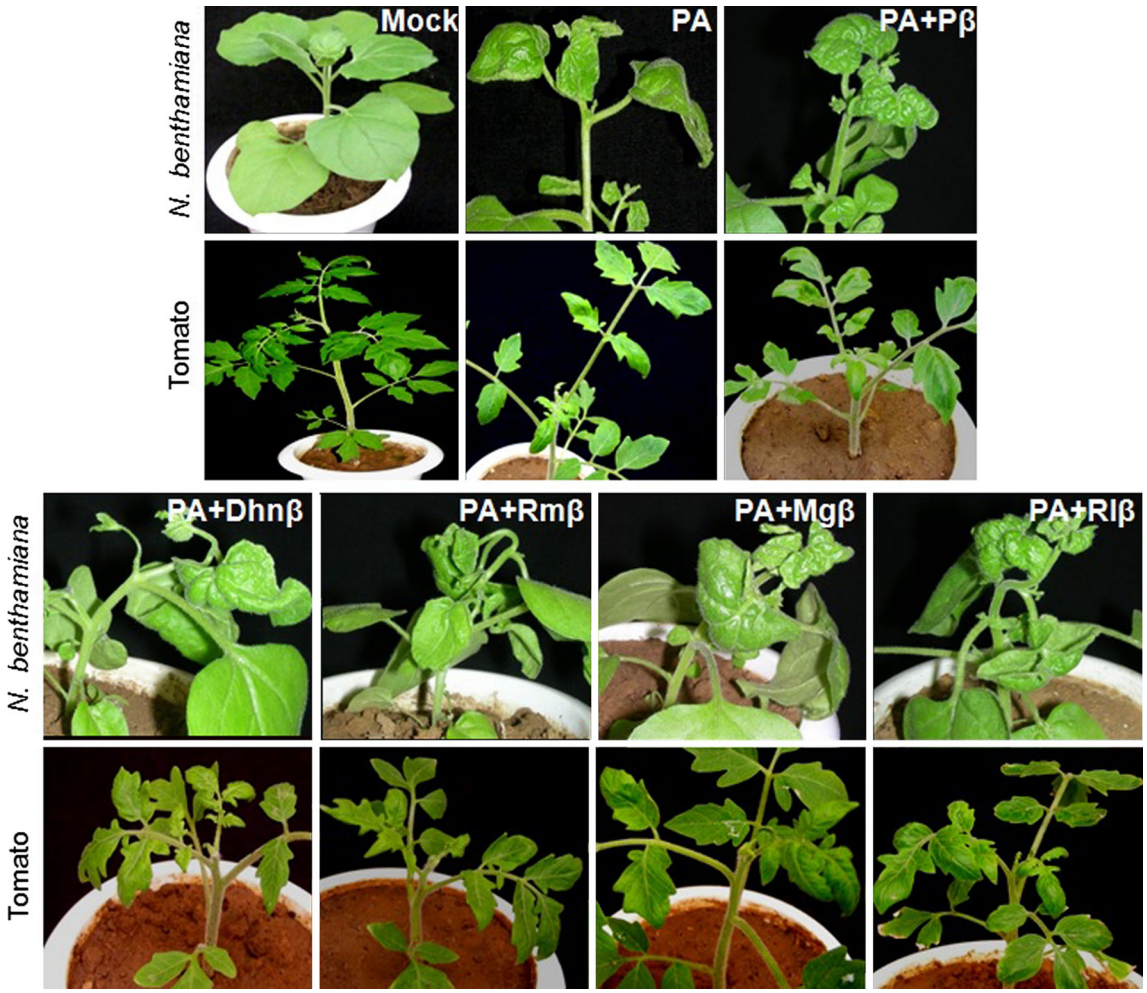


Fig. 1 Symptoms induced by infectious clones, *Tomato leaf curl Patna virus* (PA) alone or with different betasatellites on *N. benthamiana* and tomato cv. DVRT-1 plants at 25 dpi

Results

Effect of betasatellites in the accumulation of helper viruses in *N. benthamiana* and tomato

To understand the trans-replication of betasatellites, Dhnβ, Mgβ, Pβ, Rlβ and Rmβ were co-inoculated with NA, PA or VA on *N. benthamiana* and tomato plants (Figs. 1, 2, 3). All the combinations could infect and produce discernible symptoms on *N. benthamiana*, when inoculated either as DNA A alone or in combination with cognate DNA B/betasatellites. However, variation in symptom appearance and degree of symptoms were observed depending on the combinations used for agro-inoculation. When subjected for PCR using specific primers, the presence of DNA A, DNA B or betasatellite was not detected from all combinations of agro-inoculated plants (Table 1). Tomato plants co-inoculated with begomoviruses (either PA or VA) and betasatellites developed a range of symptoms depending on the combinations used (Table 1; Figs. 1, 2). In general, incubation period on tomato was longer than that observed on *N. benthamiana*, when different combinations were used for inoculation (Table 1). In *N. benthamiana*, unlike NA which induce mild symptom, PA and VA alone inoculated plants have developed distinguishable symptoms (Figs. 1, 2, 3). As expected, the association of DNA B or betasatellite has resulted in increased symptom severity with the reduced incubation period (Table 1). In accordance with symptoms observed on the plants, the viral DNA accumulation was detected in *N. benthamiana* plants inoculated with either NA alone or with betasatellites studied. However, the increased systemic accumulation of viral DNA A (NA) was evident from *N. benthamiana* plants inoculated with NA and different betasatellites, but this accumulation was less in comparable to the level from the plants inoculated with NA and NB (Fig. 4a).

When tomato plants were inoculated either with NA alone or with different betasatellites tested, the helper virus

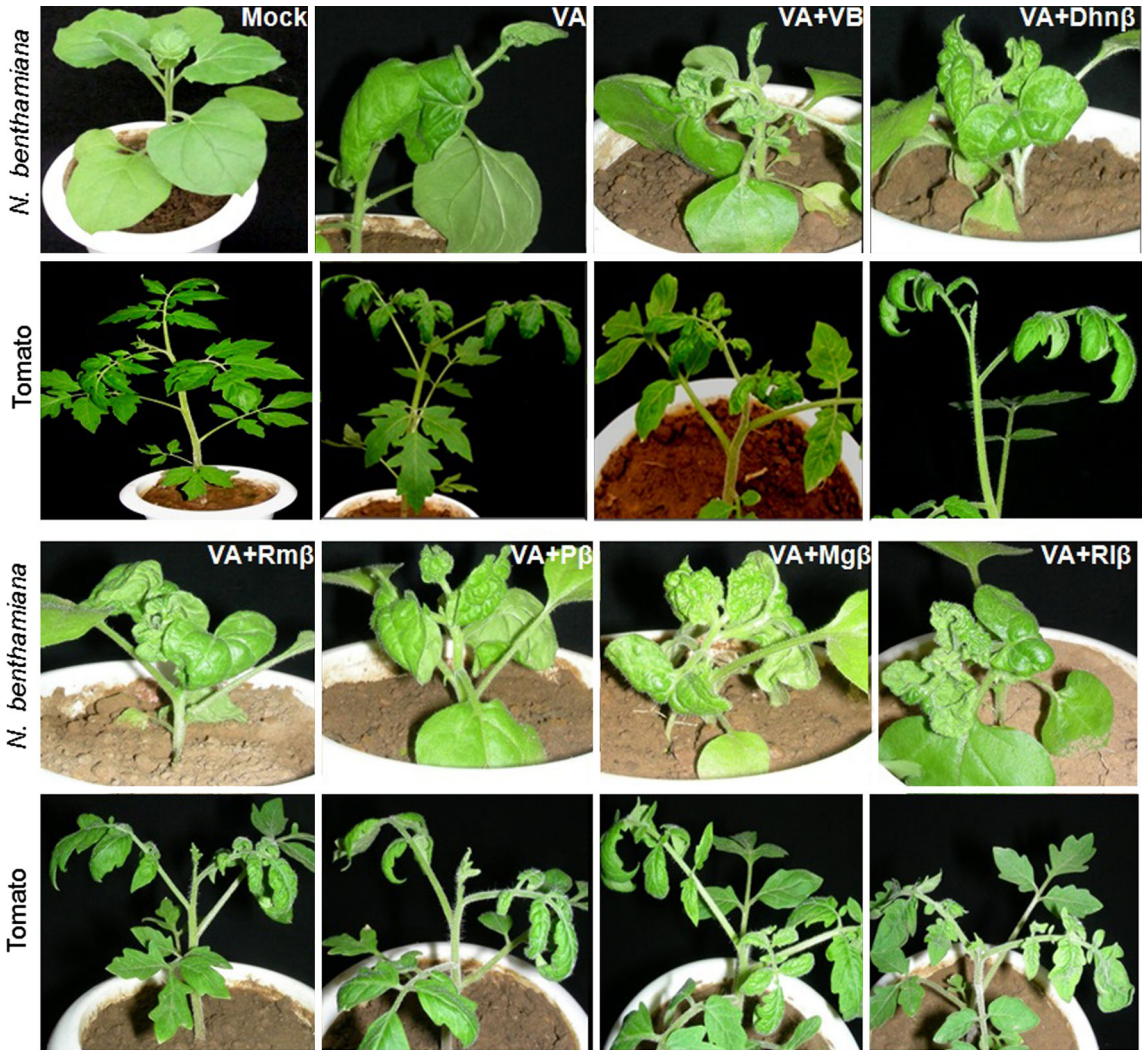


Fig. 2 Infectivity of *Tomato leaf curl Gujarat virus* (VA) alone or with different betasatellites as mentioned in each panel on agro-inoculated *N. benthamiana* and tomato cv. DVRT-1 plants at 25 dpi

could not infect (alone or with the betasatellites) as evidenced by Southern hybridisation (Fig. 4b). Interestingly, when tested through PCR using specific primers, neither the helper virus nor betasatellites could be detected in tomato plants inoculated either with NA alone or with different betasatellites (Table 1). However, the systemic infection and the viral DNA accumulation were observed when the tomato plants co-inoculated with NA and its cognate DNA B (Figs. 3, 4b).

Nicotiana benthamiana and tomato plants were co-inoculated either Rmβ or Dhnβ with PA, NA or VA to study viral DNA accumulation. Despite the presence of viral specific replicative forms in *N. benthamiana*, the accumulation of viral DNA A (PA, NA or VA) did not increase (Fig. 4a). On tomato, in the presence of Dhnβ, VA accumulation was increased during the early period of infection (11–14 dpi), however, at 25 dpi, the accumulation of viral DNA A was found to be similar (data not shown). Similarly, the other betasatellites (Pβ, Mgβ and Rlβ) also failed to exert any influence on the systemic accumulation of DNA A molecules (PA, NA or VA) and thereby not essential for the infectivity of its associated helper viruses (Fig. 4a).

Differential interaction of betasatellites and helper begomoviruses

To determine the comparative efficacy of different betasatellites in influencing the symptoms, the accumulation of betasatellites was analysed. All the betasatellites tested were found replicating with their helper begomoviruses by displaying typical symptoms on *N. benthamiana* plants (Figs. 1, 2, 3). Higher accumulation of Mgβ and Rlβ was observed with NA agro-inoculated plants as compared to

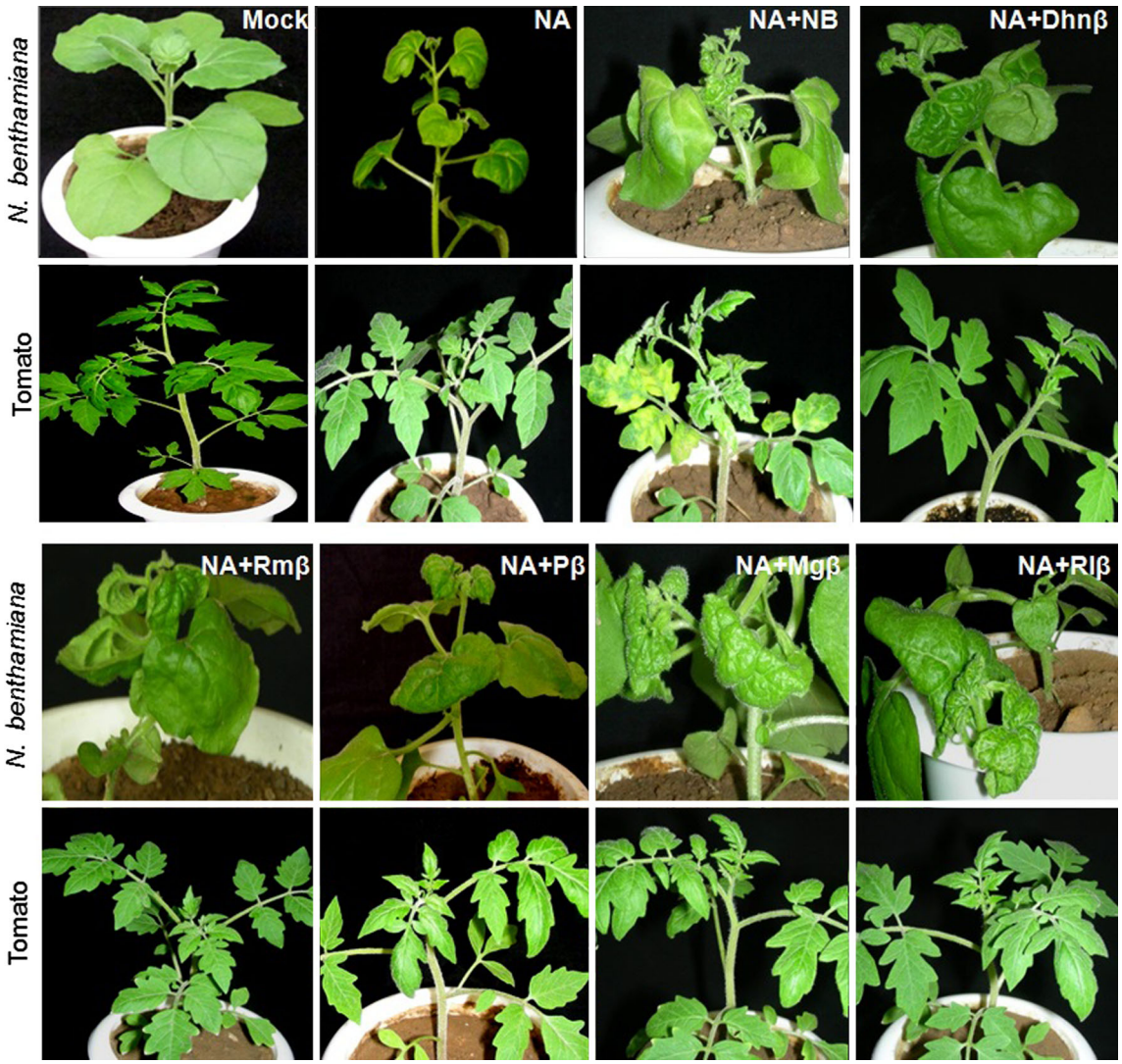


Fig. 3 Symptom induction on *N. benthamiana* and tomato cv. DVRT-1 plants agro-inoculated either with *Tomato leaf curl New Delhi virus* (NA) alone or with different betasatellites at 25 dpi

other betasatellites co-inoculated with NA in *N. benthamiana* plants (Fig. 5a, b). Tomato plants inoculated with different cognate pairs of RlA+Rlβ and MgA+Mgβ remained symptomless till 45 dpi. However, Southern analysis revealed that neither Mgβ nor Rlβ was detected when the tomato plants were co-inoculated with any of the distinct helper begomoviruses studied (Fig. 5a, b). Surprisingly, in tomato, the presence of Rlβ was detected only when co-inoculated with VA (Fig. 5b). Dhnβ and Rmβ were very well detected in the inoculated *N. benthamiana* and tomato plants with VA rather than with NA or PA (Fig. 5c, d). However, all the five betasatellites were detected in *N. benthamiana* plants co-infected with NA and different betasatellites studied, but interestingly, none of the betasatellites were detected from the infected tomato plants (Fig. 5a–e).

Sequence analysis of betasatellites

Dhnβ and Rmβ share 97.9 % nucleotide identity amongst them and possess 72.1 and 71.6 % nucleotide identity with Rlβ sequence, respectively, whereas Pβ shares 76 % nucleotide sequence identity with Mgβ. SCR of Dhnβ and Rmβ was 96.2 % identical, followed by 93.5 % identity between Dhnβ and Rlβ. Per cent identity between SCR of Pβ and Rmβ was found to be 92.8 % (Fig. S1). The position of conserved TATA box and GCTACGC sequence were located at the upstream of βC1 start codon. However, CAAT box in Dhnβ, Rmβ and Rlβ was located comparatively nearer (<40 nt upstream) to the start codon of βC1, whereas it was located farther (418 nt upstream) in the βC1 of Mgβ and Pβ (Fig. S2). Strikingly, the orientation of predicted inverted repeats in Dhnβ, Pβ and Rmβ has been

Table 1 Viral constructs, their infectivity and symptom induction on agro-inoculated plants

Inoculum	Symptomatic plants/ infected plants	Incubation period[a]	Symptoms	PCR	
				DNA A[b]	DNA-B/ DNA-β[c]
Tomato					
PA	9/15	15	LC, SL, MY	+	−
PA+Pβ	10/15	12	LC, CR, SL, YL, St	+	+
PA+Dhnβ	4/6	12	LC, SL, MY, CR	+	+
PA+Rmβ	5/6	12	LC, SL, MY, CR	+	+
PA+Mgβ	5/6	15	LC, SL, MY, CR	+	−
PA+Rlβ	4/6	15	LC, SL, MY	+	−
VA	14/15	9	LC, CR	+	−
VA+VB	15/15	7	LC, CR, LR, YL	+	+
VA+Dhnβ	6/6	9	LC, LR, St	+	+
VA+Rmβ	6/6	8	LC, LR, St	+	+
VA+Pβ	13/15	9	LC, LR, St	+	+
VA+Mgβ	6/6	7	LC, CR, YL	+	+[d]
VA+Rlβ	6/6	9	LC, CR, LR	+	+
NA	0/15	−	−	−	−
NA+NB	15/15	8	LC, CR,YL	+	+
NA+Dhnβ	0/6	−	−	−	−
NA+Rmβ	0/6	−	−	−	−
NA+Pβ	0/15	−	−	−	−
NA+Mgβ	0/6	−	−	−	−
NA+Rlβ	0/6	−	−	−	−
MgA+Mgβ	0/4	−	−	−	−
RlA+Rlβ	0/4	−	−	−	−
Nicotiana benthamiana					
PA	10/12	9	LC, LR, Chls	+	−
PA+Pβ	11/12	7	LC, LR, Bls, VB, IVCL, MO, St, Tw	+	+
PA+Dhnβ	6/6	7	LC, LR, Chls, Tw	+	+
PA+Rmβ	6/6	7	LC, LR, Chls, Tw	+	+
PA+Mgβ	6/6	7	LC, LR, Chls, St, IVCL	+	+
PA+Rlβ	5/6	9	LC, LR, Chls, St, Tw, IVCL	+	+
VA	12/12	8	LC, CR, St	+	−
VA+VB	12/12	5	LC, IVCL, CR, LD, Chls, St, Bls, SL	+	+
VA+Dhnβ	6/6	7	LC, CR, St, Tw	+	+
VA+Rmβ	6/6	6	LC, CR, St, Tw	+	+
VA+Pβ	12/12	8	LC, CR, St, Tw	+	+
VA+Mgβ	6/6	6	LC, CR, St, Tw, IVCL, YL	+	+
VA+Rlβ	6/6	8	LC, CR, St, Tw, IVCL, YL	+	+
NA	8/12	14	LC	+	−
NA+NB	12/12	7	LC, SL,YL, St, MO	+	+
NA+Dhnβ	6/6	10	LC, CR, Tw	+	+
NA+Rmβ	5/6	9	LC, CR, Tw	+	+
NA+Pβ	9/12	12	LC, CR, Tw	+	+
NA+Mgβ	6/6	9	LC, LR, VB, IVCL	+	+
NA+Rlβ	6/6	7	LC, LR, Tw, VB, IVCL	+	+
MgA+Mgβ	4/4	7	LC, LR, Bls, VB, IVCL	+	+
RlA+Rlβ	4/4	8	LC, LR, Tw, VB	+	+

LC leaf curling, *SL* small leaves, *MY* mild yellowing, *LR* leaf rolling, *Chls* chlorosis, *St* stunting, *CR* leaf crinckling, *YL* yellowing, *Bls* blistering, *VB* vein banding, *IVCL* interveinal chlorosis, *MO* mosaic, *Tw* twisting of petiole, *LD* leaf distortion; '+', PCR amplification in all the inoculated plants; '−', no PCR amplification in any of the inoculated plants

[a] Incubation period is defined as the time taken for the appearance of symptoms after inoculation

[b] Specific primer for DNA-A of PA, NA, VA

[c] Specific primer for DNA-B of NB, VB or universal primer for DNA β

[d] PCR amplification in 2 out of 6 plants tested

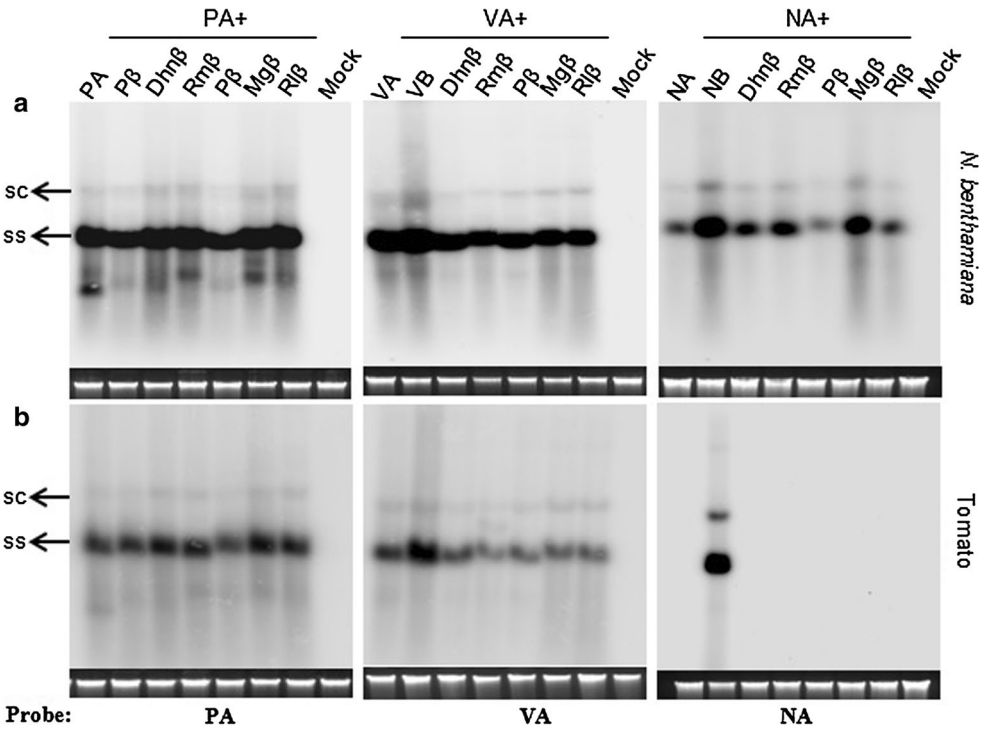


Fig. 4 Comparative level of viral DNA accumulation of *Tomato leaf curl Patna virus* (PA), *Tomato leaf curl Gujarat virus* (VA) and *Tomato leaf curl New Delhi virus* (NA) at 25 dpi. Southern blot analysis showing viral DNA accumulation in *N. benthamiana* (**a**) and tomato plants (**b**) as indicated. The specific probes used for hybridisation are given in the lower end of blots and ethidium bromide-stained DNA is provided below each lane as loading control. The viral forms are abbreviated as *sc* for super coiled and *ss* for single stranded

identified with the potential to form simpler stem-loop structure. However, the inverted repeats in βC1 of Mgβ and Rlβ were suggested to form more than one stem-loop structure (Fig. S3).

Discussion

Tomato is one of the most favoured hosts for begomoviruses causing leaf curl disease leading to complete yield loss in severe condition [1, 8]. Currently, a total of 14 begomoviruses associated with tomato leaf curl disease have been reported from the Indian sub-continent [22]. Amongst them, ToLCNDV is a bipartite virus which depends heavily on DNA B for systemic infection, unlike, DNA A of a mono-bipartite virus, ToLCGV [8, 9]. Similarly, some of the non-tomato infecting bipartite begomoviruses also has been reported to move systemically without DNA B under certain conditions [26, 27]. DNA B encoded proteins have been elucidated in assisting intra- and inter-cellular movement of helper begomoviruses [28]. Further, tomato-infecting monopartite begomoviruses have also been reported in association with diverse betasatellites for typical symptom induction [4, 6]. Betasatellite-mediated amplification of helper virus has been suggested to involve either in systemic movement or conditioning of cells to make them permissive for replication. However, in general, both betasatellites and DNA B serve to augment the accumulation of its helper component in the host cells. Subsequently, it was demonstrated that betasatellites depend on the begomovirus for its replication and its gene product as a pathogenicity determinant [12, 13]. Several factors have been demonstrated for the trans-replication and maintenance of betasatellites, and hence, is not always relaxed [14–16].

In this present study, we have demonstrated that the betasatellites studied could be trans-replicated and maintained by their helper viruses in *N. benthamiana*, but not uniformly in tomato. Rlβ and Mgβ could neither be maintained by their cognate DNA A counterparts nor by the non-cognate helper viruses in tomato, with a notable exception of trans-replication of Rlβ by VA. The absence of systemic infection in tomato plants co-infected with diverse begomoviruses and betasatellites (Rlβ and Mgβ) indicates that there could be the lack of co-ordination between these satellite molecules and/or their gene product with the host factor(s) for inducing the successful infection. Further, the regions demonstrated earlier to be responsible for helper begomoviruses encoded ORFs to trans-replicate/maintain betasatellites were not common amongst the betasatellites studied. A truncated RNA-dependent RNA polymerase in *N. benthamiana* species, an important host-mediated PTGS defence regulator, has been found be associated with increased susceptibility to RNA viruses [29]. As a PTGS suppressor protein, βC1 proteins could easily exploit this defence compromised machinery in *N. benthamiana* plants and thereby efficiently countering the minimal host defence activity to cause systemic infection with NA. Further, βC1 is also known to assist inter-cellular movement of viruses and thus could complement the lack

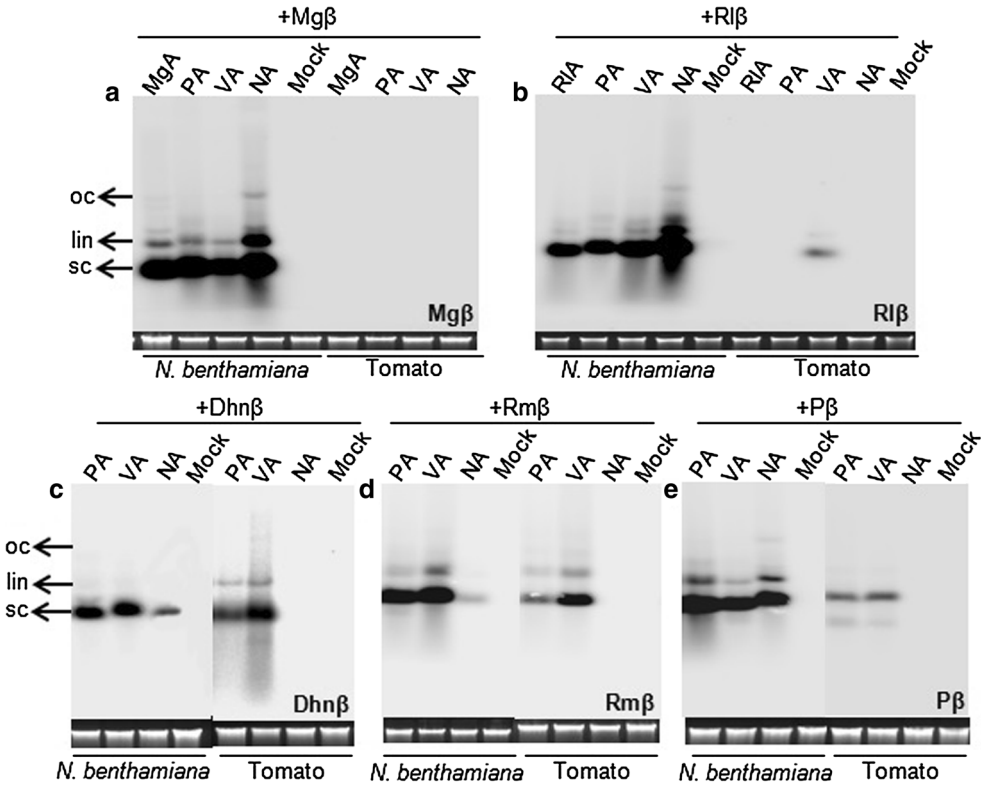


Fig. 5 Relative viral DNA accumulation of *Tomato leaf curl Joydebpur betasatellite* [Mgrβ] (**a**), *Radish leaf curl betasatellite* [Rlβ] (**b**), *Tomato yellow leaf curl Thailand betasatellite*: Dhanbad [Dhnβ] (**c**), *Tomato yellow leaf curl Thailand betasatellite*: Ramgarh [Rmβ] (**d**) and *Tomato leaf curl Patna betasatellite* [Pβ] (**e**) from agro-inoculated plants at 25 dpi. Southern blot analysis showing the level of betasatellite DNA from inoculated *N. benthamiana* and tomato plants, as mentioned above of each blots. The specific probes used for betasatellite detection are indicated within the blots and the viral forms are abbreviated as *oc* for open circular, *lin* for linear and *sc* for super coiled. Ethidium bromide-stained DNA in each blot is shown as a loading control

of movement function in the absence of DNA B. However, diverse βC1s could possess different degrees of silencing suppression activity which can be articulated for the differential accumulation of NA in the presence of various betasatellites. As suggested earlier (15), it could also be possible due to the difference in Rep recognising sequence amongst distinct betasatellites. Moreover, none of the betasatellites studied here could complement the movement function of DNA B in tomato plants; as a result, tomato plants remained symptomless. This further raises a question whether the non-infectious nature of NA with the betasatellites in tomato is due to defect in replication or in inter-cellular movement. From the results, it can be explained that for a successful infection, the interaction between begomoviruses and betasatellites is necessarily be mediated by one or more host factor(s).

TYLCTHB (Dhnβ, Rmβ) and ToLCPaB (Pβ) associated with tomato-infecting begomoviruses in India have been reported to be trans-replicated by VA and PA, respectively, in their natural host, tomato [10, 22]. However, these betasatellites did not significantly enhance the accumulation of helper viruses. Further, both these helper viruses (VA, PA) are itself ascertained as movement competent, a typical characteristic of monopartite viruses. As suggested earlier [11], this type of monopartite begomoviruses could possibly loose associated betasatellite in due course of evolution. It can also be explained that this type of betasatellites probably will offer impaired movement function (loss of function due to lack of evolutionary pressure) and hence, unable to help bipartite type of virus (like NA) to systemically invade its natural host, tomato. Alternatively, it could be possibly due to the lack of efficient interaction between these DNA A molecule and betasatellite product, which is needed to form a functional movement complex or poor interaction of DNA A gene products with betasatellite sequences to form a functional replication complex. Despite βC1 as a pathogenicity determinant, the promoters of βC1 for each betasatellite differ, which is likely to affect its expression and thus could affect infectivity. However, the role of the upstream regulatory elements in the trans-replication/maintenance of betasatellites and/or the expression of βC1 is yet to be studied. Furthermore, it will be interesting to elucidate the role of inverted repeats in the interaction between Rep protein and betasatellite sequence for the trans-replication/maintenance of satellite molecules.

Unlike DNA B component, the trans-replication of diverse betasatellites by begomoviruses-encoded Rep proteins undergoes in a more relaxed way. However, in the present study, none of the betasatellites could complement the movement function for the helper bipartite virus to cause systemic infection. Unlike in *N. benthamiana*, the trans-replication and/or maintenance of betasatellites by diverse tomato-infecting begomoviruses is not uniform in tomato. Our results suggest that the compromised host defence mechanism in *N. benthamiana* species, but not in tomato, could be responsible for the maintenance of studied

betasatellites by diverse Indian begomoviruses. Results also emphasise that trans-replication and/or maintenance of diverse betasatellites by tomato-infecting begomoviruses is a host-driven process.

Acknowledgments PR acknowledges the Council of Scientific and Industrial Research, Government of India for providing the fellowship. This research work was supported by the Department of Science and Technology, Government of India through a grant (SR/SO/PS-87/08) to SC.

References

1. A. Varma, V.G. Malathi, Ann. Appl. Biol. **142**, 145–164 (2003)
2. C.M. Fauquet, R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou, Arch. Virol. **153**, 783–821 (2008)
3. S.S. Campos, A.M. Ayala, B.M. Martín, L.A. Caballero, J.N. Castillo, E. Moriones, Virus Res. **173**, 286–293 (2013)
4. R.W. Briddon, S.E. Bull, I. Amin, S. Mansoor, I.D. Bedford, N. Rishi, S.S. Siwatch, Y. Zafar, A.M. Abdel Salam, P.G. Markham, Virology **324**, 462–474 (2004)
5. B. Chattopadhyay, A.K. Singh, T. Yadav, C.M. Fauquet, N.B. Sarin, S. Chakraborty, Arch. Virol. **153**(3), 533–539 (2008)
6. X. Cui, X. Tao, Y. Xie, C.M. Fauquet, X. Zhou, J. Virol. **78**, 13966–13974 (2004)
7. A.K. Singh, B. Chattopadhyay, S. Chakraborty, Virol. J. **9**, 43 (2012)
8. S. Chakraborty, P.K. Pandey, M.K. Banerjee, G. Kalloo, C.M. Fauquet, Phytopathology **93**, 1485–1495 (2003)
9. M. Padidam, R.N. Beachy, C.M. Fauquet, J. Gen. Virol. **76**, 25–35 (1995)
10. P. Kumari, A.K. Singh, B. Chattopadhyay, S. Chakraborty, Virus Res. **152**(1–2), 19–29 (2010)
11. Z.H. Li, X.P. Zhou, Y. Xie, Phytopathology **95**, 902–908 (2005)
12. P. Gopal, P.P. Kumar, B. Sinilal, J. Jose, A.Y. Kasin, R. Usha, Virus Res. **123**(1), 9–18 (2007)
13. K. Saunders, A. Norman, S. Gucciardo, J. Stanley, Virology **324**(1), 37–47 (2004)
14. K. Saunders, R.W. Briddon, J. Stanley, J. Gen. Virol. **89**, 3165–3172 (2008)
15. M.S. Nawaz-ul-Rehman, S. Mansoor, R.W. Briddon, C.M. Fauquet, J. Virol. **83**, 9347–9355 (2009)
16. Z. Iqbal, M.N. Sattar, A. Kvarnheden, S. Mansoor, R.W. Briddon, Virus Res. **169**, 107–116 (2012)
17. B.L. Patil, C.M. Fauquet, J. Gen. Virol. **91**, 1871–1882 (2010)
18. A. Chatterji, M. Padidam, R.N. Beachy, C.M. Fauquet, J. Virol. **73**, 75481–75489 (1999)
19. M. Saeed, Y. Zafar, J.W. Randles, M.A. Rezaian, J. Gen. Virol. **88**, 2881–2889 (2007)
20. P.V. Sivalingam, A. Varma, Arch. Virol. **157**, 1081–1092 (2012)
21. O. Eini, S.A.A. Behjatnia, S. Dogra, I.A. Dry, J.W. Randles, M.A. Rezaian, J. Gen. Virol. **90**, 253–260 (2009)
22. P. Ranjan, R.V. Kumar, S. Chakraborty, Virus Genes **47**, 524–531 (2013)
23. S.L. Dellaporta, J. Wood, J.B. Hicks, Plant Mol. Biol. Rep. **1**, 19–21 (1983)
24. J. Sambrook, D.W. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd edn. (Cold Spring Harbor Laboratory Press, New York, 2001), pp. 6.39–6.46
25. S. Chakraborty, R. Vanitharani, B. Chattopadhyay, C.M. Fauquet, J. Gen. Virol. **89**, 818–828 (2008)
26. D. Evans, H. Jeske, Virology **194**, 752–757 (1993)
27. F.A. Klinkenberg, J. Stanley, J. Gen. Virol. **71**, 1409–1412 (1990)
28. Y. Gafni, B.L. Epel, Physiol. Mol. Plant Pathol. **60**, 231–241 (2002)
29. S. Yang, S.A. Carter, A.B. Cole, N. Cheng, R.S. Nelson, Proc. Natl. Acad. Sci. USA **101**(16), 6297–6302 (2004)

Virus Genes (2014) 48:397–401
DOI 10.1007/s11262-013-1027-7

Molecular characterization of *Chilli leaf curl virus* and satellite molecules associated with leaf curl disease of *Amaranthus* spp

B. George · R. Vinoth Kumar · S. Chakraborty

Received: 24 October 2013 / Accepted: 13 December 2013 / Published online: 25 December 2013
© Springer Science+Business Media New York 2013

Abstract *Amaranthus*, collectively known as amaranth, is an annual or short-lived perennial plant used as leafy vegetables, cereals and for ornamental purposes in many countries including India. During 2011, leaf samples of *Amaranthus* plants displaying leaf curling, leaf distortion, leaf crinkling and yellow leaf margins were collected from Banswara district, Rajasthan in India. Full-length clones of a monopartite begomovirus, a betasatellite and an alphasatellite were characterized. The complete nucleotide sequence of the isolated begomovirus features as a typical 'Old World' begomovirus with the highest nucleotide per cent identity with *Chilli leaf curl virus* and hence, considered as an isolate of *Chilli leaf curl virus*. The complete nucleotide sequences of betasatellite and alphasatellite possess maximum nucleotide identity with *Tomato yellow leaf curl Thailand betasatellite* and *Chilli leaf curl alphasatellite*, respectively. This is the first report of the association of chilli-infecting begomovirus and satellite molecules infecting a new host, *Amaranthus*, causing leaf curl disease.

Keywords Begomovirus · Betasatellite · Alphasatellite · *Amaranthus* · Leaf curl disease

B. George and R. Vinoth Kumar contributed equally to this study

Electronic supplementary material The online version of this article (doi:10.1007/s11262-013-1027-7) contains supplementary material, which is available to authorized users.

B. George · R. V. Kumar · S. Chakraborty (✉)
Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India
e-mail: supriyachakrasls@yahoo.com

Introduction

The genus, *Amaranthus,* in the family, *Amaranthaceae*, includes approximately 60 species, most of which are cosmopolitan weeds (*A. retroflexus* L., *A. hybridus* L., *A. powellii* L. and *A. spinosus* L.). Amaranth is a versatile warm-seasoned broad leaf plant that can be grown as a grain, ornamental, leafy vegetable or forage crop. Some *Amaranthus* spp. are grown as a grain in India, south-east Asia, Manchuria and North America and is recognized as gluten-free suitable for diets of coeliac disease patients [1, 2]. Like Quinoa, it is a pseudocereal because of its flavour and cooking similarities to grains. It also provides a good source of tocotrienols, proteins, dietary fibre and minerals such as iron, magnesium, phosphorus, copper and manganese [3]. Amaranth is known to prevent hypovitaminosis and anaemia and is also considered as the 'crop of the future' [4, 5].

Geminiviruses are circular single-stranded (ss) DNA viruses comprising either one or two genomic components (DNA-A and DNA-B) of approximately 2.7 kb each and are encapsidated within twined icosahedral virions of 18 × 30 nm in size [6, 7]. The components of bipartite begomoviruses share a highly conserved common region, approximately 200 nucleotides (nt) in length containing repetitive sequence-specific replication initiator protein (Rep) binding motifs called iterons and a nonanucleotide stem-loop structure (TAATATTAC) required for replication of the viral genome [8]. Members of the family, *Geminiviridae*, are the most devastating pathogens infecting food and fibre crops worldwide [8]. In the recent years, these viruses are reported to extend their host range into other commercially important hosts, such as, chilli, cotton and radish [9–11]. Based on genome organization, host range and vector specificity, they are sub-divided into

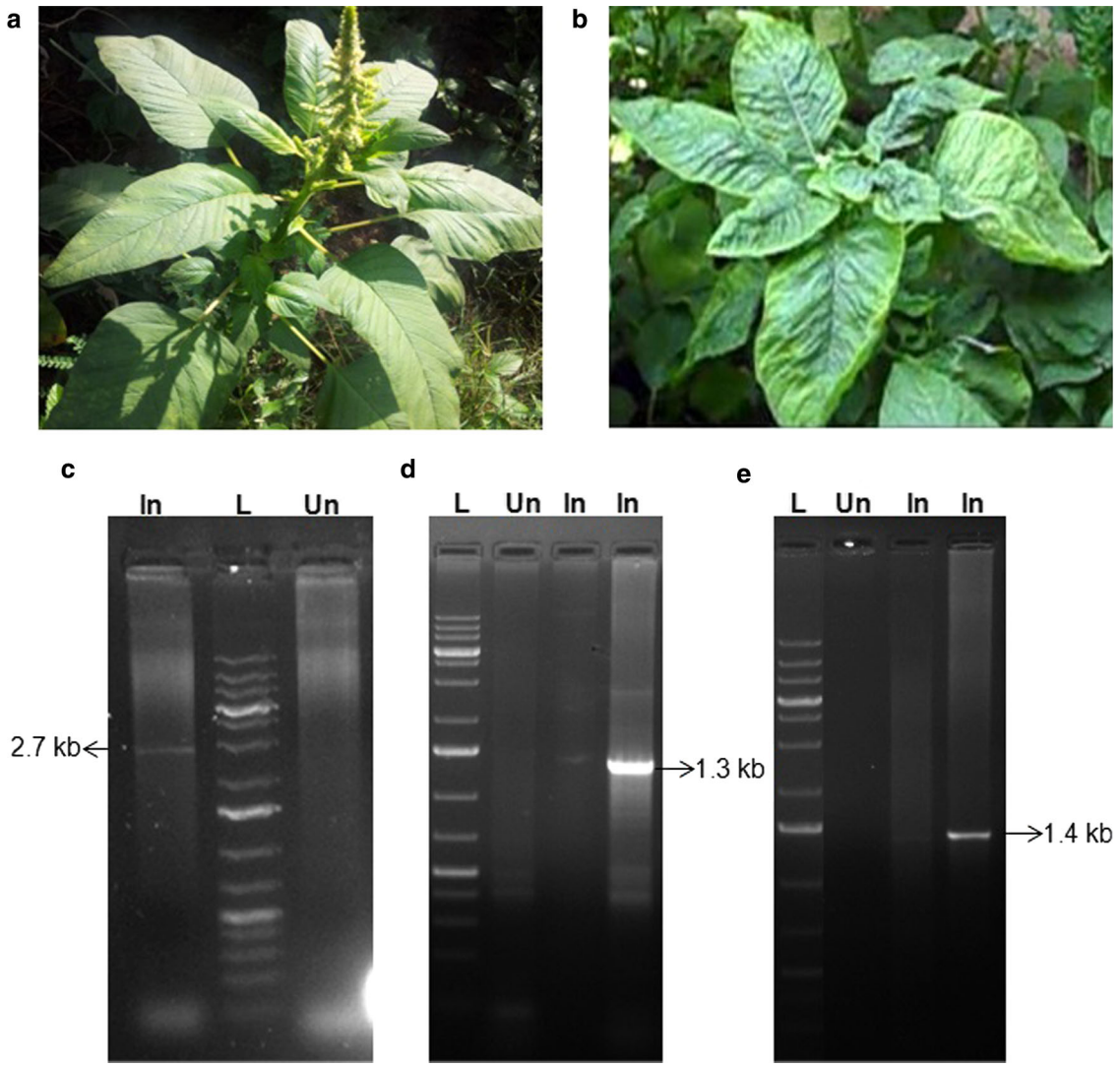


Fig. 1 Photograph showing non-infected (**a**) and infected (**b**) *Amaranthus* plants from the infected field sample. PCR amplification of begomovirus (**c**), betasatellite (**d**) and alphasatellite (**e**) from the infected field samples. The abbreviated labels for L refer to 1 Kb DNA ladder, un to uninfected plant sample and in to virus infected plant sample

seven genera, viz., *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus* [12]. Begomoviruses are transmitted by the insect vector, whitefly (*Bemisia tabaci*), and are further classified either as monopartite or bipartite. Majority of begomoviruses reported from the 'Old World' (OW) are monopartite in nature with the association of circular ssDNA satellites known as betasatellites [13, 14] and/or satellite-like components, alphasatellites [14]. Betasatellites require helper viruses for their replication, encapsidation, insect transmission and movement in plants, whereas alphasatellites are capable of self-replication in host plant, but require helper begomoviruses for movement in plants as well as insect transmission.

During 2011, severe leaf curl disease of *Amaranthus* was observed in Banswara region of Rajasthan, India. In the present study, we have characterized for the first time the association of a monopartite begomovirus alongwith satellite molecules (betasatellite and alphasatellite) causing leaf curl disease in *Amaranthus*.

Methodology

A severe leaf curl disease was observed on *Amaranthus* plants during October, 2011 in Banswara district, Rajasthan, India (Fig. 1a, b). The infected leaves of the plants exhibiting typical geminivirus-like symptoms such as leaf curling, leaf distortion, leaf crinkling and yellow leaf margins were used for total DNA extraction [15] followed by rolling-circle amplification (RCA) using φ29 DNA polymerase. The replicative concatamers yielded through

RCA were subjected to digest with several restriction enzymes such as *Bam*HI, *Hin*dIII, *Kpn*I and *Xba*I and were cloned in pBluescript (SK+) vector. In addition, PCR amplification using degenerate primers specific for DNA-A (FP 5′ TCAGGATCCATTAGTAAACGAGTTTCC 3′ and RP 5′ TGCGGATCCCACATTGGAAAGAGTG 3′) and DNA-B (FP 5′ TCTAGAACTSATCKGGTGTCTCTC 3′ and RP 5′ TCTAGAGAGAGAAACTSCWVCTTC 3′) was used to clone the helper genomic components. Betasatellite and alphasatellite components were PCR amplified using abutting primers as described previously [16, 17]. Subsequently, all the PCR-amplified products of either ~2.7 or ~1.3 kb were then cloned into pTZ57R/T vector (Fermentas). Restriction digestion pattern of various full-length viral clones was analysed by digesting with a range of restriction enzymes and followed by sequencing commercially at DNA sequencing facility, University of Delhi South campus, New Delhi.

Sequences were examined for the presence of begomovirus-specific ORFs (using NCBI ORF Finder) and conserved nonanucleotide sequence. Per cent nucleotide identity of the identified begomovirus and the satellite molecules was calculated by aligning them using Clustal W with the selected begomoviruses and satellites from the database. A phylogenetic dendrogram was also generated by neighbour-joining method with 1,000 bootstrap repetitions using MegAlign programme available in DNASTAR software.

Results and discussion

Amaranth grain is a highly nutritional pseudocereal with a superior amount of proteins and possesses antiallergic and antioxidant activities [18]. Viruses constitute a major limiting factor to vegetable production throughout the world [19]. Only a few reports of viruses (*Amaranthus leaf mottle virus*) infecting *Amaranthus* are available in the literature [20], however, it was not reported to be a host for begomoviruses causing leaf curl disease. Here, we ascertain the association of a begomovirus and the satellite molecules with leaf curl disease from India. Notably, in spite of our repeated attempts, we have failed to obtain full-length genome (DNA-A) of begomoviruses using RCA-based direct cloning procedure, and hence, the full-length clone of a begomovirus was obtained using PCR-based cloning (Fig. 1c–e). In addition, the presence of DNA-B could not be ascertained in the infected *Amaranthus* samples when degenerate primers were used. Several clones obtained through PCR were digested with various restriction enzymes to check the diversity in the cloned begomovirus populations. However, all the clones showed similar restriction pattern, and, therefore, one of the clones was then selected to sequence completely. Complete nucleotide sequence of the begomovirus isolated from *Amaranthus* was found to be of 2,755 nt (KF471061) and sequence analysis showed the genome organization of a typical of OW begomoviruses with six open reading frames (ORFs). Full-length sequence and all the encoded ORFs of the identified begomovirus share maximum per cent identity of ~95 % with *Chilli leaf curl virus* (Table S1). It also segregated with *Chilli leaf curl virus* in the phylogenetic dendrogram (Fig. 2a); hence, it is considered to be an isolate of *Chilli leaf curl virus*. In the tropical and subtropical countries, chilli-infecting begomoviruses have emerged as a major threat by infecting economically important crops such as potato, bitter gourd, papaya, tomato and tobacco leading to complete crop loss [21–26]. Association of chilli-infecting begomovirus with this host plant pinpoints its expanding host range and appears to be the most serious of the newly emerging/re-emerging begomoviruses-associated diseases in Indian sub-continent.

The size of betasatellite isolated was determined to be 1,352 nt (KF471033) encoding a single βC1 protein of 118 aa (coordinates 205–561 nt), adenine rich region (coordinates 706–1,054 nt) and satellite-conserved region (SCR; coordinates 1,240–14 nt). It possesses the highest per cent nucleotide identity (78.6 %) with *Tomato yellow leaf curl Thailand betasatellite*-TYLCTHB (Table S1) and is found to be segregated with them in the phylogenetic tree (Fig. 2b). Based on the proposed species demarcation threshold for betasatellites (78 %), it should be considered as an isolate of TYLCTHB. *Chilli leaf curl virus* is a monopartite begomovirus which depends heavily on betasatellites for the induction of typical symptoms [9]. However, unlike with tomato-infecting begomoviruses [27, 28], the association of TYLCTHB with chilli-infecting begomoviruses has not been reported and hence, the nature of disease development by this newly emerging disease complex causing leaf curl disease is yet to be studied.

The alphasatellite of 1,392 nt (KF471047) has the capacity of encoding 315 aa protein (coordinates 76–1,023 nt) similar to Rep protein of nanoviruses, A-rich region (coordinates 1,010–1,157 nt) and a predicted hairpin structure with the loop sequence of TAGTATTAC. Alphasatellite identified here shares a maximum nucleotide identity of 96 % with unreported *Chilli leaf curl alphasatellite* and thus segregated with them in the phylogenetic analysis (Fig. 2c). On the basis of species demarcation threshold for alphasatellites (83 %), it is proposed to be an isolate of *Chilli leaf curl alphasatellite* (unpublished data). Recently, alphasatellites have been shown to ameliorate the disease development in the infected plant [29, 30]. However, the nature of the association of a monopartite begomovirus with betasatellite and alphasatellite leading to leaf curl disease in *Amaranthus* is unclear which needs further experimentation.

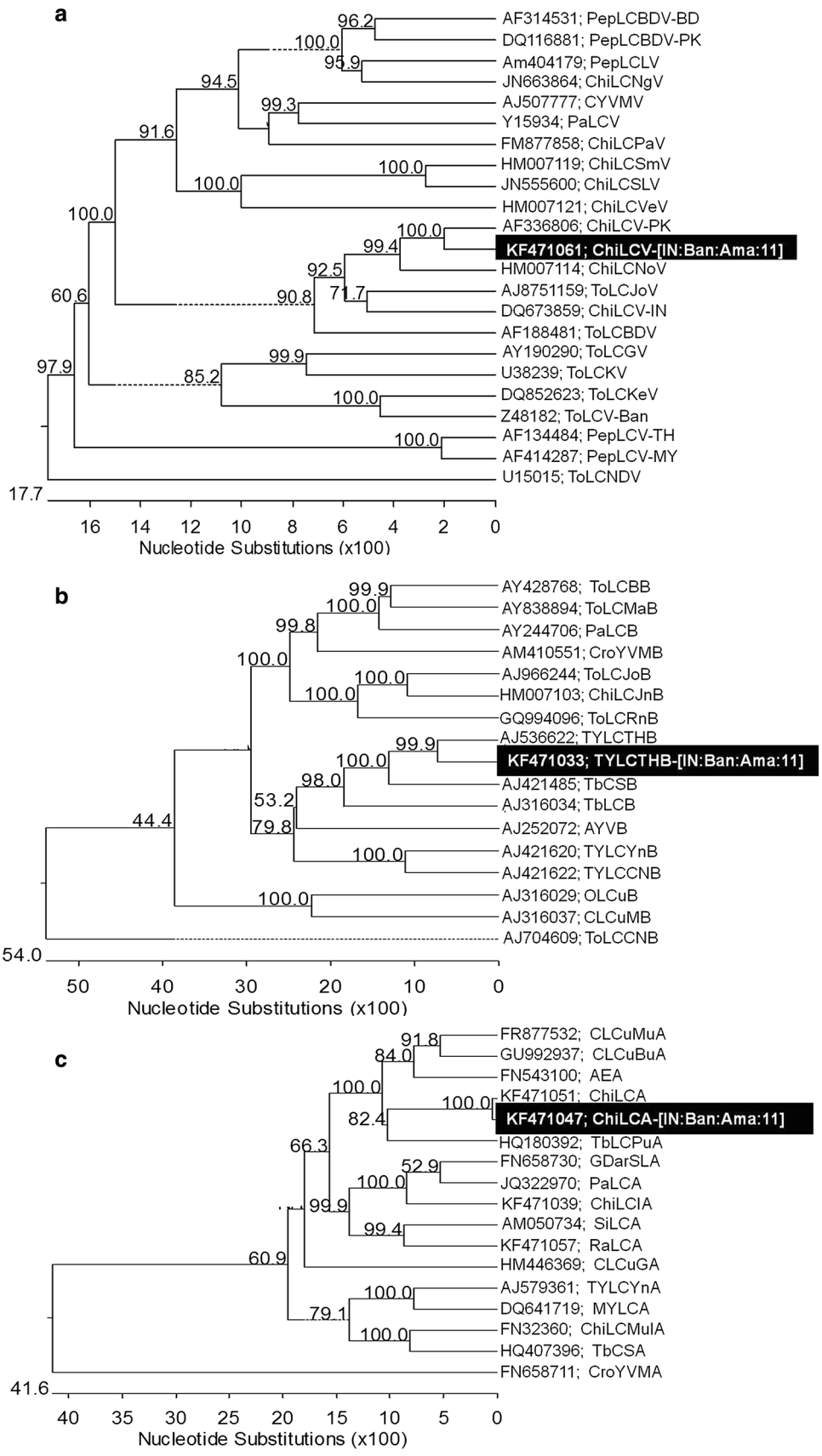


Fig. 2 Phylogenetic dendrogram of begomovirus (**a**), betasatellite (**b**) and alphasatellite (**c**) nucleotide sequences identified in this study with the representative sequences obtained from the database. The sequences were aligned using CLUSTAL W algorithm, and the phylogenetic trees were generated using neighbour-joining method available with the DNAstar package. A bootstrap analysis with 1,000 replicates was performed and provided at the tree nodes

This present study is the first demonstration of a disease complex comprising a monopartite begomovirus, betasatellite and alphasatellite associated with the leaf curl disease in *Amaranthus* in India. Although the extent of the spread of this disease is not known, the results emphasize the greater impact of *Chilli leaf curl virus* by expanding its host range to *Amaranthus*. This disease may emerge as a serious threat and thus, further studies are required to determine the diversity of begomoviruses and associated satellites infecting *Amaranthus* crop.

Acknowledgements Financial support by UGC Resource Net working and DST-PURSE provided by Jawaharlal Nehru University to SC is gratefully acknowledged.

References

1. H. Hauptli, S. Jain, Calif. Agric. **31**, 9 (1977)
2. T. Thompson, J. Am. Diet. Assoc. **101**(5), 586–587 (2001)
3. J. Amaya-Farfan, R. Marcılio, R.C. Spehar, Seg. Alim. Nutr. **12**, 47–56 (2005)
4. B. Reiz, C. Leitzmann, Entwicklung and Landlicher Rawn **19**(4), 8–12 (1985)
5. Marx, Science **198**(4312), 40 (1977)
6. M. Padidam, R.N. Beachy, C.M. Fauquet, J. Gen. Virol. **76**, 249–263 (1995)
7. J. Stanley, D.M. Bisaro, R.W. Briddon, J.K. Brown, C.M. Fauquet, B.D. Harrison, E.P. Rybicki, D.C. Stenger, Geminiviridae, in *Virus Taxonomy*, ed. by C.M. Fauquet, M.A. Mayo, J. Maniloff, U. Desselberger, L.A. Ball (Elsevier/Academic Press, London, 2005), pp. 301–326
8. C.M. Fauquet, D.M. Bisaro, R.W. Briddon, J. Brown, B.D. Harrison, E.P. Rybicki, D.C. Stenger, J. Stanley, Arch. Virol. **148**, 405–421 (2003)
9. B. Chattopadhyay, A.K. Singh, T. Yadav, C.M. Fauquet, N.B. Sarin, S. Chakraborty, Arch. Virol. **153**(3), 533–539 (2008)
10. S. Saeed, Y. Zafar, J.W. Randles, M.A. Rezaian, J. Gen. Virol. **88**(10), 2881–2889 (2007)
11. A.K. Singh, B. Chattopadhyay, S. Chakraborty, Virol. J. **9**, 43 (2012)
12. M.J. Adams, A.M.Q. King, E.B. Carstens, Arch. Virol. (2013). doi:10.1007/s00705-013-1688-5
13. R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou, C.M. Fauquet, Arch. Virol. **153**, 763–781 (2008)
14. R.W. Briddon, J. Stanley, Virology **344**, 198–210 (2006)
15. S.L. Dellaporta, J. Wood, J.B. Hicks, Plant Mol. Biol. Rep. **1**, 19–21 (1983)
16. S.E. Bull, R.W. Briddon, P.G. Markham, Mol. Biotechnol. **23**, 83–86 (2003)
17. R.W. Briddon, S.E. Bull, S. Mansoor, I. Amin, P.G. Markham, Mol. Biotechnol. **20**, 315–318 (2002)
18. V.M. Caselato-Sousa, J. Amaya-Farfan, J. Food Sci. **77**, 4 (2012)
19. T.N. Nyamupingidza, V.Machakaire, 397–406 (2003) In: Plant virology In sub-Saharan Africa, Proc. Conf. Organized by IITA. J. d'A. Hughes and J. Odu, eds. International Institute of Tropical Agriculture, Ibadan, Nigeria
20. M.A. Taiwo, A.T. Owolabi, J. Sci. Res. Dev. **9**, 63–74 (2004)
21. D.M.J.B. Senanayake, B. Mandal, S. Lodha, A. Varma, New Dis. Rep. **13**, 27 (2006)
22. Y. Kumar, V. Hallan, A.A. Zaidi, Plant Pathol. **60**, 1040–1047 (2011)
23. M. Mubin, R.W. Briddon, S. Mansoor, Arch. Virol. **154**(2), 365–368 (2009)
24. S.K. Raj, M.S. Khan, R. Singh, N. Kumari, D. Prakash, Int. J. Food Sci. Nutr. **56**(3), 185–192 (2005)
25. P. Singh-Pant, P. Pant, S.K. Mukherjee, S. Mazumdar-Leighton, Arch. Virol. **157**(7), 1217–1232 (2012)
26. D.M. Senanayake, J.E. Jayasinghe, S. Shilpi, S.K. Wasala, B. Mandal, Virus Genes **46**(1), 128–139 (2012)
27. P. Ranjan, R.V. Kumar, S. Chakraborty, Virus Genes (2013). doi:10.1007/s11262-013-0977-0
28. P. Jyothsna, R. Rawat, V.G. Malathi, Arch. Virol. **158**, 217–224 (2013)
29. A.M. Idris, M.S. Shahid, R.W. Briddon, A.J. Khan, J.K. Zhu, J.K. Brown, Virology **92**, 706–717 (2011)
30. M.S. Nawaz-ul-Rehman, N. Nahid, S. Mansoor, R.W. Briddon, C.M. Fauquet, Virology **405**, 300–308 (2010)

Virus Genes (2013) 47:524–531
DOI 10.1007/s11262-013-0977-0

Differential pathogenicity among *Tomato leaf curl Gujarat virus* isolates from India

Punam Ranjan · R. Vinoth Kumar · S. Chakraborty

Received: 5 July 2013 / Accepted: 26 August 2013 / Published online: 12 September 2013
© Springer Science+Business Media New York 2013

Abstract *Tomato leaf curl Gujarat virus* (ToLCGV) has been identified as one of the most destructive pathogens causing tomato leaf curl disease (ToLCD) in India. In the tomato growing regions of Dhanbad and Ramgarh, plants bearing severe symptoms of ToLCD such as leaf curling, leaf crinkling, yellowing and leaf rolling was observed in the farmer fields. The association of begomovirus in these samples was confirmed by PCR and the causal viruses were identified as the isolates of ToLCGV. However, association of cognate DNA B component could not be ascertained from these samples. Indeed, like other Old World begomoviruses, the present ToLCGV isolates were found to be associated with a particular betasatellite, *Tomato yellow leaf curl Thailand betasatellite* (TYLCTHB). Although DNA A of both ToLCGV isolates could alone infect tomato inducing systemic symptoms, the difference in virulence was observed. Co-inoculation of TYLCTHB reduced the incubation period without influencing the accumulation of helper virus DNA and hence, differential pathogenesis among ToLCGV isolates was governed by the helper component rather than betasatellite. ToLCGV infection with DNA B increases the accumulation of DNA A component of Dhanbad isolate but not of Ramgarh isolate. Results indicated that the begomovirus identified from Ramgarh sample was a mild strain of ToLCGV.

Keywords Begomovirus · Betasatellite · Symptom severity · ToLCGV · Mild isolate

Introduction

Geminiviridae, the second largest family of plant-infecting viruses comprises a group of single-stranded circular DNA with distinctive twinned isometric particle morphology of approximately 20 × 30 nm in size. Based on the host range, genome organisation and vector transmission, it is sub-divided into seven genera as *Becurtovirus, Begomovirus, Curtovirus, Eragrovirus, Mastrevirus, Topocuvirus* and *Turncurtovirus* [1, 2]. Among them, begomoviruses, the largest genus are transmitted by the insect vector whitefly (*Bemisia tabaci* Genn.) and cause devastating disease in majority of the economically important crops worldwide. Further, they were classified as either monopartite (containing a single circular ssDNA of 2.7-kb DNA A) or bipartite (having two similar sized genome of DNA A and DNA B). DNA A encodes six proteins for replication, transactivation and encapsidation of viral components whereas DNA B encodes two proteins which aid the viral molecules to spread systemically from the site of infection [1]. The similarity in iteron region (replication-associated protein-binding sequence) of DNA B and DNA A ensures the binding between the DNA A-encoded Rep and the cognate DNA B for its replication [3]. Monopartite begomoviruses are often associated with satellite molecule(s) named as betasatellite which is about half the size of their helper molecule [4]. Betasatellites contain an adenine-rich region; a highly conserved region termed as satellite conserved region (SCR) and encodes a single ORF, βC1 [4]. These betasatellites depend on their helper component (DNA A)

Electronic supplementary material The online version of this article (doi:10.1007/s11262-013-0977-0) contains supplementary material, which is available to authorized users.

P. Ranjan · R. V. Kumar · S. Chakraborty (✉)
Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: supriyachakrasls@yahoo.com

for their replication. Betasatellite-encoded βC1 protein has been characterized as pathogenicity determinant and a suppressor of post-transcriptional gene silencing [5–7]. Although role of betasatellites has been studied in disease establishment, however, all the betasatellites are not necessarily responsible for disease development [8, 9].

Tomato leaf curl disease (ToLCD) has emerged as the major hindrance for tomato cultivation in India leading to 100 % crop loss in severe condition [10]. Plants affected with this disease bear symptoms such as crinkling, puckering, upward or downward curling of leaves, reduced leaf lamina, stunted growth, vein clearing and yellow mottling [10]. This disease was first reported several decades ago from northern India in 1948 [11]. In India, several closely related begomoviruses have been associated with this disease comprising of 9 monopartite begomoviruses namely, *Tomato leaf curl Bangalore virus*, *Tomato leaf curl Bangladesh virus, Tomato leaf curl Joydebpur virus, Tomato leaf curl Karnataka virus, Tomato leaf curl Kerala virus, Tomato leaf curl Patna virus, Tomato leaf curl Pune virus, Tomato leaf curl Rajasthan virus and Tomato leaf curl Ranchi virus;* 2 bipartite begomoviruses including *Tomato leaf curl New Delhi virus and Tomato leaf curl Palampur virus*; and 1 mono-bipartite virus as *Tomato leaf curl Gujarat virus* (ToLCGV) [10, 12].

ToLCGV, one of the most devastating begomovirus species associated with ToLCD was first reported in Varanasi region in India [12]. Subsequently, the presence of this species was also noticed throughout the major tomato growing regions in India such as Gujarat, Maharashtra and Uttar Pradesh [13]. Further, this virus has been identified from tomato in Nepal and strikingly, also reported to infect a weed, *Xanthium* in Pakistan underlining the hindrance caused by ToLCGV on tomato cultivation and productivity in Indian sub-continent [14]. ToLCGV DNA A component alone is infectious but association of DNA B increases the symptom severity and hence, characterized as a mono-bipartite species [12]. ToLCGV DNA A and its cognate DNA B share a common region (CR) of 155 bp with only 60 % identity between them [12].

In the present study, two different ToLCGV isolates lacking DNA B component have been identified from severely infected tomato plants collected from Dhanbad and Ramgarh district in Jharkhand, India. However, both the helper components DNA A have been found to be in association with a betasatellite, *Tomato yellow leaf curl Thailand betasatellite* (TYLCTHB). The difference in the disease development among these two isolates and the effect of the association of TYLCTHB with these ToLCGV isolates are discussed. Further, the detailed biological and molecular characterization of these two ToLCGV isolates is also reported.

Materials and methods

Cloning and sequence analysis of ToLCGV isolates from Dhanbad and Ramgarh

During 2006–2008, tomato plants exhibiting begomovirus-like symptoms such as leaf curling, leaf crinkling, leaf rolling and mottling were collected from ToLCD-infected fields located in Dhanbad and Ramgarh regions. Genomic DNA was isolated from the collected samples following Dellaporta et al. [15] and subsequently viral replicative forms were isolated by rolling circle amplification using φ29 DNA polymerase (TempliPhi, GE Healthcare, USA) as per manufacturer's instructions [16]. Total DNA was subjected to PCR amplification using abutting primers to check the presence of begomovirus [17–19]. Approximately, 1 μg of concentrated super-coiled DNA was digested with *Kpn*I and then ligated into pUC18 vector previously linearized with the same enzyme. The PCR-amplified betasatellite fragment of ~1.3 kb was then ligated into pTZ57R/T vector (Fermentas, USA). All the clones were confirmed by appropriate restriction digestions. Full-length sequencing of DNA A and DNA β clones of Dhanbad and Ramgarh isolate was carried out commercially by primer walking at DNA sequencing facility, University of Delhi South campus. These sequences were used to derive all the putative ORFs and their amino acid sequences through NCBI ORF finder (http://www.ncbi.nlm.nih.gov/). Percent pairwise sequence identity was generated using the species demarcation tool, SDTv1.0 programme (http://web.cbio.uct.ac.za/SDT) and the phylogenetic tree was constructed using maximum-likelihood algorithm available in MEGA 5 software [20]. The deduced amino acid sequences of ORFs of all three isolates of ToLCGVs were aligned using Clustal W programme (www.ebi.ac.uk).

Generation of infectious tandem repeat constructs

Partial tandem repeats of DNA A (AY190290) and DNA B (AY190291) constructs of ToLCGV Varanasi (denoted as VarA and VarB, respectively) used in this study has been previously described [12]. For preparing a tandem repeat construct of Dhanbad isolate of ToLCGV (EF573714), 1.2-kb fragment containing intergenic region of DNA A of ToLCGV-[Dhn] was generated by *Hin*dIII (84)–*Kpn*I (1,581) digestion and was cloned in pCAMBIA 2,301 vector followed by mobilization of full-length 2.7 kb at *Kpn*I site to yield a partial tandem repeat as pCVADh32 (designated as DhnA). Similarly, a 522 bp, *Bam*HI (455)–*Kpn*I (1,291) fragment of TYLCTHB-[Dhn] (EF573713) was cloned in pCAMBIA 2,301. Then a full-length *Kpn*I fragment was ligated to produce an infectious DNA β construct as pCDhβ2 (hereafter called as Dhnβ).

For producing tandem dimeric construct of Ramgarh isolate of ToLCGV (GQ994098), a *Bam*HI (148)–*Kpn*I

(1,580)-digested 1.3-kb fragment containing intergenic region of DNA A of ToLCGV-[Ram] was first cloned in pCAMBIA 2,301 vector followed by mobilization of full-length 2.7 kb at *Kpn*I site to generate a partial tandem repeat as pCVARm17 (described as RmA). Similarly, *Eco*RI (317)–*Kpn*I (1,291) fragment of 384 bp of TYLCTHB-[Ram] (GQ994097) was cloned in pCAMBIA 2,301; followed by the ligation of a full-length *Kpn*I fragment to produce a partial tandem repeat DNA β construct as pCRmβ5 (called as Rmβ). The orientation of inserts in the generated partial tandem repeats was confirmed by restriction digestion with appropriate enzymes. These infectious constructs were then mobilized into *Agrobacterium tumefaciens* strain EHA 105 for infectivity studies.

Infectivity of partial tandem repeats of cloned DNAs

The generated infectious constructs were introduced into test plants through agro-inoculation method to assess the infectivity. Three-leaf stage tomato *cv*. DVRT-1 plants were used for agro-inoculation either with a single partial tandem repeat construct or a combination of constructs. Transformed *Agrobacterium* strain with an empty pCAMBIA 2,301 vector was used for mock inoculation. All plants were maintained in an insect-proof glasshouse with 16 h of daylight and a temperature of 28 ± 2 °C. The symptom progression was measured till 45 days post inoculation (dpi).

Detection of viral DNA accumulation in the infected tomato plants

For analysing the viral DNA accumulation from the infected tomato plants, the total genomic DNA was isolated at different dpi. For Southern hybridization, approximately 10 μg of genomic DNA was fractionated on 0.9 % agarose gel and then transferred onto nylon membrane (Millipore) following the standard protocol [21, 22]. Approximately 1-kb [*Hin*dIII (84)–*Eco*RI (1,786)] fragment from VarA and a 668-bp [*Xba*I (2,539)–*Eco*RI (1,871)] fragment from VarB were used as probe. For detection of Dhnβ and Rmβ, a TYLCTHB-[Dhn] 836-bp fragment [*Bam*HI (455)–*Kpn*I (1,291)] was used as probe. Viral DNA fragments were labelled with [α^{32}P] dCTP by random oligonucleotide-primer synthesis [23] and the viral DNA accumulation was detected using a PhosphorImager (Fuji Film).

Results

Comparison of nucleotide sequence and ORFs of ToLCGV isolates

From the genomic DNA isolated from the ToLCD-infected samples, PCR amplification using the abutting primers specific for DNA A and betasatellite indicated the presence begomovirus in the samples. Sequence analyses of Dhanbad and Ramgarh clones revealed that viral genome sizes were 2,758 and 2,761 nucleotides, respectively. According to species demarcation threshold value [24], the cloned helper virus component was determined to be the isolates of ToLCGV and share 97–99 % nucleotide identity among them. Therefore, Dhanbad and Varanasi isolates were clustered along with other ToLCGV isolates identified in tomato; however, Ramgarh isolate segregated with a ToLCGV isolate isolated from a weed *Xanthium* in Pakistan (Fig. 1). CR region of the three isolates of ToLCGV share nucleotide identity ranging between 94.6–98.3 %. The iteron sequences were similar among the three isolates as GGTGT-ATT-GGAGT; however, difference in spacer sequence was noticed, for example ATG in the case of Ramgarh isolate whereas ATT in other ToLCGV isolates (Fig. S1). Amino acid differences were also noticed in the viral ORFs among the three ToLCGV isolates (Fig. 2).

The full-length nucleotide sequence length of betasatellites isolated from Dhanbad and Ramgarh was found to be of 1,352 and 1,358 nucleotides having maximum nucleotide per cent identity (99 %) with TYLCTHB. They share <86 % identity with other known betasatellites and possess 98.7 % identity among each other. In accordance with species demarcation threshold for betasatellites [25], the cloned betasatellites were considered to be the isolates of TYLCTHB. Hence, in the phylogenetic dendrogram analysis, they were found segregated with TYLCTHB isolates identified from Pakistan (Fig. 1).

ToLCGV Varanasi isolate is more virulent than other isolates

The infectivity of ToLCGV isolates was studied by infecting tomato plants with either DNA A alone or alongwith cognate betasatellite, TYLCTHB. When a partial tandem repeat infectious VarA clone was inoculated on tomato plants, the initial symptom appeared as early as 9 dpi. However, symptom appearance was delayed till 12 and 13 dpi when agro-inoculated with DhnA and RmA clones, respectively (Table 1). In VarA inoculated plants, symptom severity increases gradually and reached maximum at 21 dpi. At this stage, all the systemic leaves were curled and growth of shoot apex was almost checked, resulting in stunting appearance as compared to those inoculated with either DhnA or RmA (Fig. 3). Interestingly, symptom appearance and disease development were delayed when either DhnA or RmA were inoculated (Table 1). RmA inoculated plants did not exhibit stunting rather the internode length was found to be increased and the plant appeared taller than others. In Southern analysis, the accumulation of viral DNA at 11 dpi remains higher for plants inoculated with VarA as

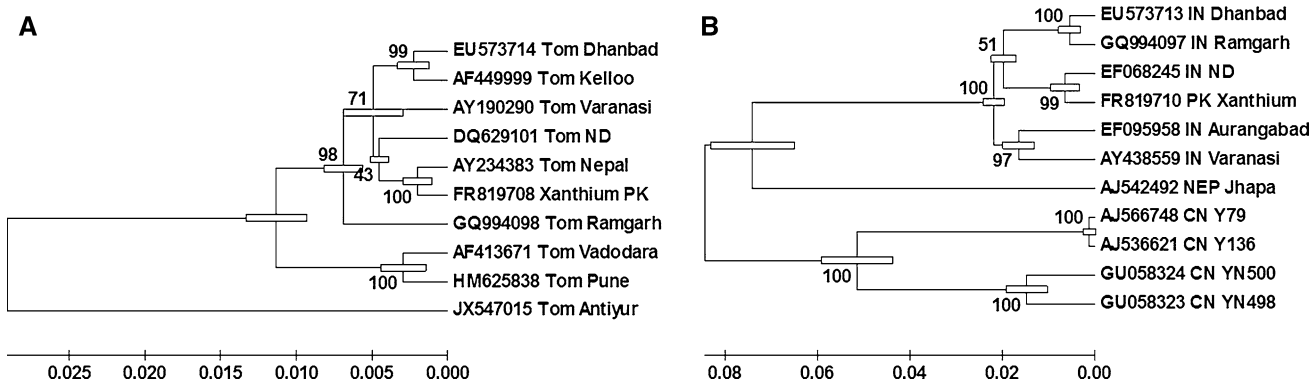


Fig. 1 Phylogenetic dendrogram analysis of ToLCGV DNA A (**a**) and TYLCTHB isolates (**b**). The complete nucleotide sequences reported here were aligned using Clustal W with the other available ToLCGV and TYLCTHB isolates in GenBank. Tree was constructed using maximum-likelihood method within MEGA 5. The percentage bootstrap values of 1,000 replicates are placed at the nodes

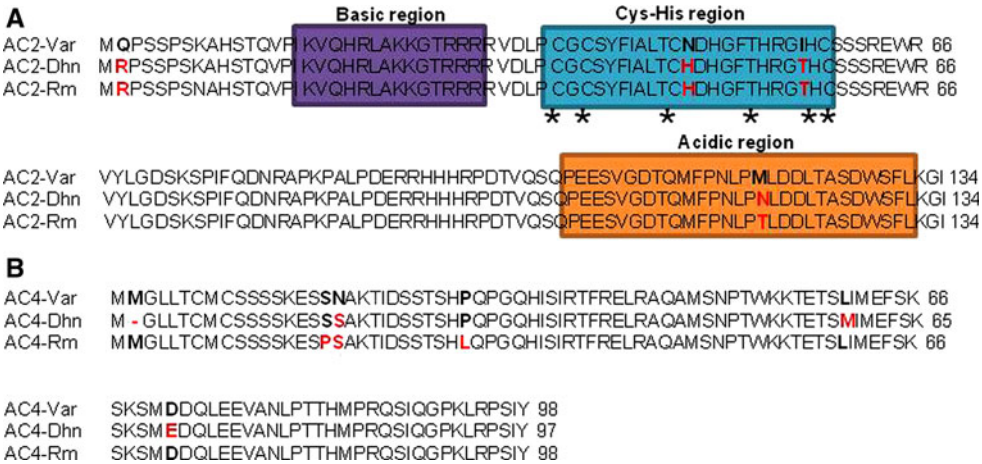


Fig. 2 Multiple alignment of ToLCGV-[Var], ToLCGV-[Dhn] and ToLCGV-[Ram] amino acid sequence of AC2 (**a**) and AC4 (**b**); *red colour font* represents changes in amino acid sequence. The conserved regions were given as *colour boxes* and labelled accordingly (Color figure online)

Table 1 Infectivity of ToLCGV isolates on tomato *cv.* DVRT-1 through agro-inoculation

Inoculum used	No. of plants infected/No. of plants inoculated	Date of symptom appearance (dpi)	Symptom appearance	Symptom severity level at 25 dpi
VarA	14/15	9	LC, CR	++++
DhnA	11/12	12	LC, CR	+++
RmA	10/12	13	LC, CR	+++
VarA + VarB	15/15	7	LC, CR, LR, YL	+++++
DhnA + VarB	12/12	8	LC, CR, LR, YL	++++
RmA + VarB	12/12	12	LC, YL	+++
DhnA + Dhnβ	12/12	9	LC, YL	++++
RmA + Rmβ	11/12	10	LC, YL	++++

LC leaf curling, *CR* leaf crinkling, *YL* yellowing, *LR* leaf rolling

comparable to those inoculated with either DhnA or RmA. Although after 14 dpi, the accumulation of DNA became almost equal in all the three isolates inoculated plants but symptom appearance always remained more severe for plants inoculated with VarA compared to plants inoculated with either DhnA or RmA (Fig. 4a).

Tomato plants co-inoculated with either DhnA + β or RmA + β showed a typical symptom at 9 and 10 dpi, respectively (Table 1). In both the cases, the infected plants developed more severe symptoms than plants inoculated with corresponding DNA A alone (Fig. 3). Interestingly, increased viral DNA A accumulation in tomato plant was

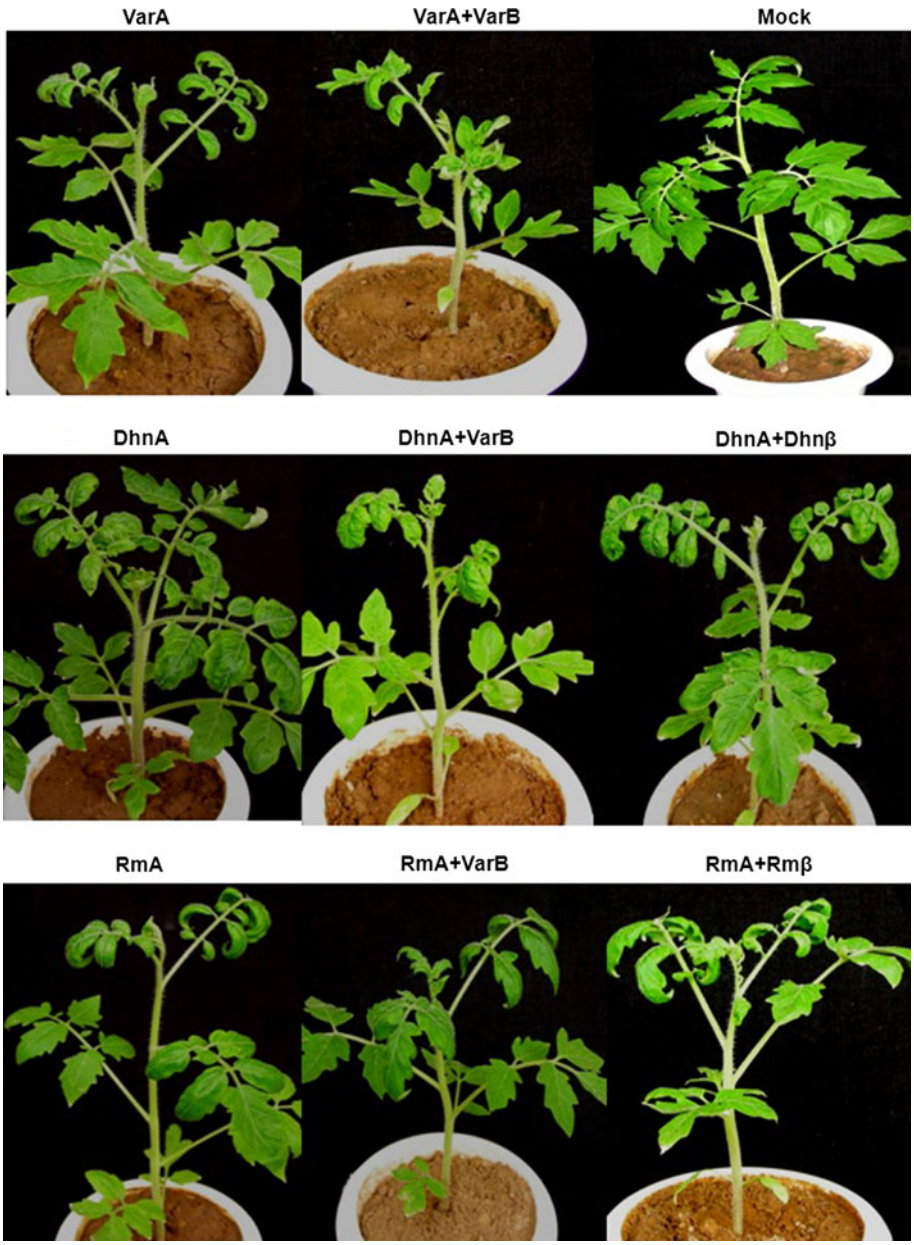


Fig. 3 Symptom severity on tomato *cv.* DVRT-1 plants at 25 dpi. The different combinations used for agro-inoculation were depicted in the figure

observed in the case of DhnA + β than in DhnA alone inoculated plants (Fig. 4a). However, the level of betasatellite DNA accumulation remained unchanged in the plants inoculated with either DhnA + β or RmA + β (Fig. 4b).

VarB increases DNA A accumulation of Varanasi and Dhanbad isolates but not in Ramgarh isolate

The DNA B component of ToLCGV [Var] was responsible for increasing symptom severity coupled with enhanced accumulation of viral DNA as observed in plants co-inoculated with both helper virus components. The more severe symptom was evident from the plants inoculated with VarA + VarB than VarA alone (Fig. 3). When DhnA and VarB were co-inoculated, the earliest symptom appearance was noticed at 8 dpi which was earlier than either DhnA or DhnA + β inoculated plants (Table 1). In DhnA + VarB infected plants, the increased accumulation of viral DNA A was also detected as compared to plants inoculated with either DhnA alone or DhnA + β (Fig. 4a). However, in contrast to VarA and DhnA, the symptom appearance and accumulation of viral DNA remained similar in RmA and RmA + VarB inoculated plants. Strikingly, in DhnA + VarB inoculated plants, DNA B were detected at 11 dpi; however, this was not detected in RmA + VarB inoculated plants till 14 dpi (Fig. 4c).

Symptom severity score was recorded and used to plot a graph depicting disease development in tomato plants

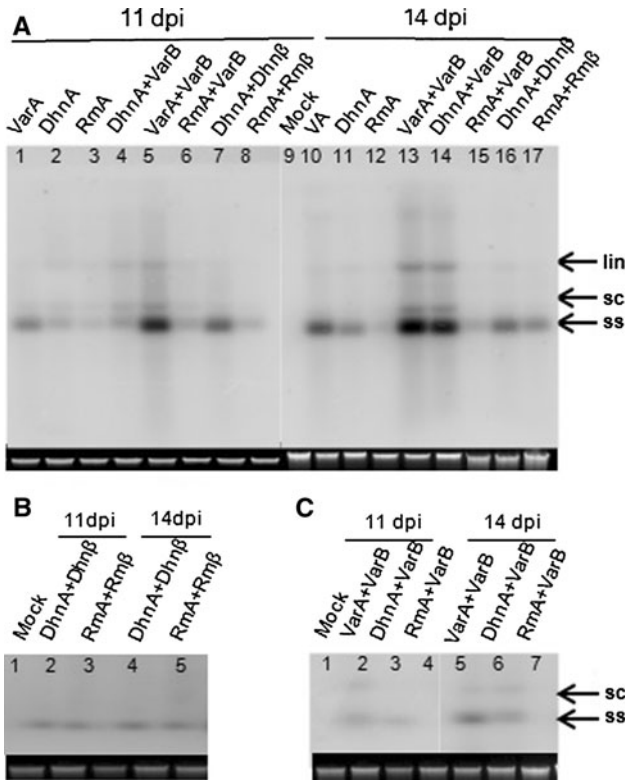


Fig. 4 Comparative level of viral DNA accumulation at 11 dpi and 14 dpi of ToLCGV isolates. Southern blot analysis was carried out from plants agro-inoculated with different combinations as mentioned above each lane. The blots were hybridized with ToLCGV DNA A probe (**a**), TYLCTHB probe (**b**) and ToLCGV DNA B probe (**c**). Ethidium bromide-stained DNA is shown below each lane as loading control. The viral forms abbreviated here are *lin* for linear, *sc* for super-coiled and *ss* for single-stranded

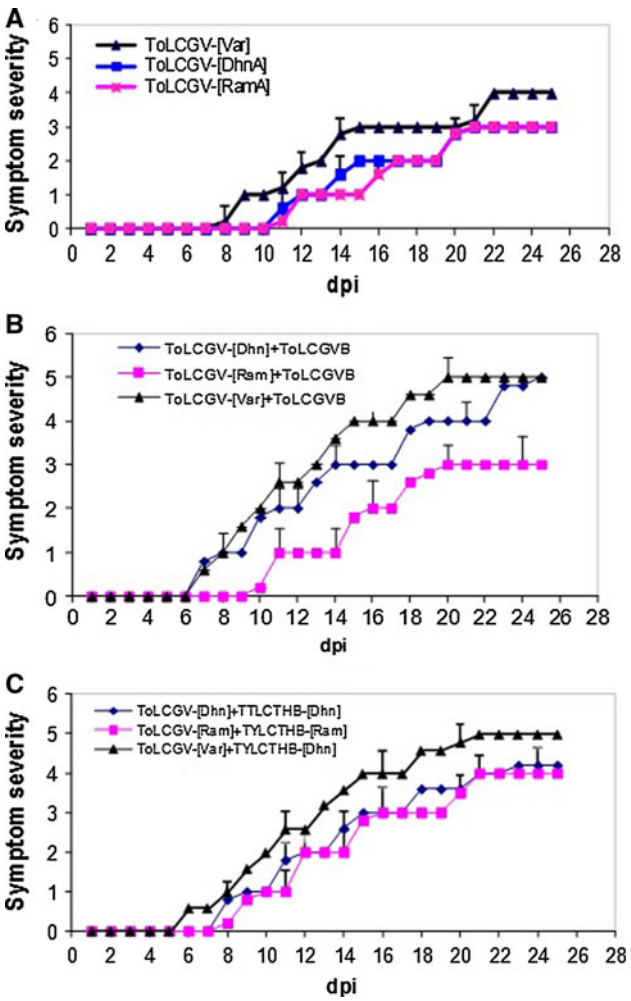


Fig. 5 Graphical representation of symptom severity pattern on tomato agro-inoculated with different constructs combination of three isolates of ToLCGV DNA A alone (**a**), ToLCGV DNA A + TYLCTHB (**b**) and ToLCGV DNA A + DNA B (**c**)

agro-inoculated with different constructs. Graphs revealed that the rate of disease development was slower for RmA inoculated plants than either DhnA or VarA alone inoculated plants (Fig. 5a). The presence of VarB with DhnA enhanced the disease development rate, but unlike VarA + VarB and DhnA + VarB, RmA inoculated plants showed slower disease development in the presence of VarB (Fig. 5b). However, when RmA was co-inoculated with Rmβ, the pattern of disease development was comparatively slower than plants inoculated with either DhnA + β or VarA + Dhnβ (Fig. 5c). Hence, it has been suggested that VarB did not assist faster disease development when inoculated with RmA than DhnA.

Discussion

ToLCGV is one of the most predominant species occurring throughout the India, except extreme southern regions and causes severe ToLCD [9, 14]. It possess a genome organisation typical to majority of Old World monopartite begomoviruses in which DNA A is alone infectious by spreading systemically, but requires its cognate DNA B molecule for typical symptom production [9, 12]. Hence, this has been termed as mono-bipartite in nature. From Pakistan, a ToLCGV isolate from *Xanthium* was characterized to be associated with TYLCTHB [14]. In addition, full-length nucleotide sequences of five ToLCGV isolates are also available in GenBank; however, no information on corresponding molecule of DNA B or betasatellite is known from these samples. This emphasizes the distribution and recent evolution of ToLCGV isolates as a monopartite virus; however, these isolates may have difference in virulence and ability to induce disease severity which has not been demonstrated earlier.

In the present study, two ToLCGV isolates from Dhanbad and Ramgarh were identified possessing a satellite DNA β belonging to the species, *Tomato yellow leaf curl Thailand betasatellite* (TYLCTHB). However, in both

the samples, the attempts failed to identify the association of DNA B. It is interesting to mention here that a recombinant region of *Tomato yellow leaf curl Thailand virus* (TYLCTHV) has been identified in ToLCGV species [26]. So, we speculate that sometime in past, mixed infection of ToLCGV and TYLCTHV has resulted in acquiring its cognate DNA β from TYLCTHV. Despite high degree of sequence identity (>97 %) among ToLCGV isolates, the difference in symptom severity between these isolates was evident from the infectivity studies. However, the DNA A of Varanasi isolate was more virulent in terms of symptom appearance and relatively higher viral DNA accumulation as compared to Dhanbad and Ramgarh isolates.

Analysing the sequences of all three ToLCGV isolates has revealed a difference in the spacer sequence within the iteron region of Ramgarh isolate; however, a characteristic role of the spacer region in determining the Rep protein binding efficiency is yet to be elucidated for any of the begomovirus Rep proteins. The efficiency of viral DNA replication heavily relies on Rep protein which does not allow amino acid difference at its N terminus. Although DNA-binding domain of Rep protein of these isolates was same, we can predict that changes in internal amino acid residues may affect its functional structure, thereby affecting the replication quantitatively. AC2 is known as a transcriptional activator, DNA-binding protein and a suppressor of post-transcriptional gene silencing of diverse begomoviruses [27–29]. The essential *cis* acting conserved amino acid sequences of AC2 in the zinc-finger motif was also identified [28]. Among the ToLCGV isolates, AC2 of Dhanbad and Ramgarh isolates shares 99 % identity between them which is more as compared to 98.3 and 98 % with Varanasi isolate, respectively. However, asparagine and isoleucine amino acid within zinc-finger region has been changed as histidine and threonine in Dhanbad and Ramgarh isolate, respectively. Additionally, a change of the amino acid each in the 2th position (glutamine to arginine) and 119th position (methionine to asparagine and threonine in Dhanbad and Ramgarh isolates, respectively) could possibly affect the interaction between AC2 and the host protein(s), thereby affecting its multi-functional role. Changes in the amino acids (serine, asparagine, and lysine) of AC4 may result in the loss of interaction with the host factors which could be mediated by post-translational modification. Thereby, these amino acid changes in AC2 and AC4 proteins may have negative interference in the host-induced gene silencing suppression activity.

Unlike DNA B containing similar iteron sequence as DNA A, the mechanism by which betasatellites are trans-replicated by helper component is unclear. However, it has been suggested that the betasatellite origin recognition by Rep encoded by helper component is carried out by more relaxed way. In the infectivity studies, co-infection of cognate TYLCTHB with Dhanbad and Ramgarh isolates resulted in reduced incubation period and enhanced symptoms; however, the level of helper component remained unaltered. This trans-replication of cognate TYLCTHB without influencing ToLCGV DNA A accumulation emphasizes that the cognate betasatellite is not responsible for the difference in the virulence among different ToLCGV isolates. Co-inoculation of Dhanbad isolate with DNA B of ToLCGV resulted in the increased symptom severity with enhanced helper component DNA accumulation similar to the plants infected with DNA A and DNA B of Varanasi isolate. However, DNA B component could influence neither symptom severity nor accumulation of DNA A of Ramgarh isolate. This could be explained as suggested [30] that the host adaptation is the cumulative effect of recruitment of host cellular factors to participate in virus multiplication, systemic dissemination and successful suppression of antiviral defence mechanism, which could ultimately determine the overall virulence after the invasion of viral components. On the basis of these results, it can be concluded that the virus from Ramgarh is a mild ToLCGV isolate.

ToLCGV is identified as a mono-bipartite begomovirus causing severe disease in tomato; however, the available data suggest that ToLCGV prefers a betasatellite (TYLCTHB) over its cognate DNA B, leading to a different disease scenario. In this context, the present study is valuable as it indicates the difference in the virulence of various ToLCGV isolates in the natural host, tomato. Further, it will be interesting to understand the molecular basis of ubiquitous association of TYLCTHB with ToLCGV isolates in India.

Acknowledgments PK is grateful to the Council of Scientific and Industrial Research, Government of India for providing fellowship. The research work was supported by The Department of Science and Technology, Government of India through a Grant (SR/SO/PS-87/08) to SC.

References

1. A.M.Q. King, M.J. Adams, E.B. Carstens, E.J. Lefkowitz, *In Virus taxonomy* (Elsevier Academic Press, London, 2011)
2. M.J. Adams, A.M.Q. King, E.B. Carstens, Arch. Virol. (2013). doi:10.1007/s00705-013-688-5
3. A. Chatterji, M. Padidam, R.N. Beachy, C.M. Fauquet, J. Virol. **73**, 75481–75489 (1999)
4. M.S. Nawaz-ul-Rehman, C.M. Fauquet, FEBS Lett. **583**, 1825–1832 (2009)
5. K. Saunders, A. Norman, S. Gucciardo, J. Virol. **324**, 37–47 (2004)
6. X. Cui, G. Li, D. Wang, D. Hu, X. Zhou, J. Virol. **79**, 10764–10775 (2005)
7. Z.H. Li, Y. Xie, X. Zhou, Phytopathology **95**(8), 902–908 (2005)
8. X. Yang, Y. Xie, P. Raja, S. Li, J.N. Wolf, Q. Shen, D.M. Bisaro, X. Zhou, PLoS Pathog. **7**(10), e1002329 (2011)

9. K. Saunders, R.W. Briddon, J. Stanley, J. Gen. Virol. **89**(12), 3165–3172 (2008)
10. S. Chakraborty, Tomato leaf curl viruses from India, *Encyclopedia of virology*, vol. 5 (Elsevier, Amsterdam, 2008), pp. 124–133
11. R.S. Vasudeva, J. Sam Raj, Phytopathology **38**, 364–369 (1948)
12. S. Chakraborty, P.K. Pandey, M.K. Banerjee, G. Kalloo, C.M. Fauquet, Phytopathology. **93**(12), 1485–1495 (2003)
13. R.V. Chowda Reddy, J. Colvin, V. Muniyappa, S.E. Seal, Arch. Virol. **150**, 845–867 (2005)
14. M. Mubin, S. Akhtar, I. Amin, R.W. Briddon, S. Mansoor, Virus Genes **44**(1), 112–119 (2012)
15. S.L. Dellaporta, J. Wood, J.B. Hicks, Plant Mol. Biol. Rep. **1**, 19–21 (1983)
16. D. Haible, S. Kober, H. Jeske, J. Virol. Meth. **135**(1), 9–16 (2006)
17. S.D. Wyatt, J.K. Brown, Phytopathology **86**, 1288–1293 (1996)
18. M.R. Rojas, R.L. Gilbertson, D.R. Russel, D.P. Maxwell, Plant Dis. **77**, 340–347 (1993)
19. R.W. Briddon, S.E. Bull, S. Mansoor, I. Amin, P.G. Markham, Mol. Biotechnol. **20**, 315–318 (2002)
20. K. Tamura, D. Peterson, N. Peterson, G. Stecher et al., Mol. Bio. Evo. **28**, 2731–2739 (2011)
21. J. Sambrook, D.W. Russell, *Molecular cloning: a laboratory manual*, 3rd edn. (Cold Spring Harbor Laboratory Press, New York, 2001), pp. 6.39–6.46
22. P. Kumari, A.K. Singh, V.K. Sharma, B. Chattopadhyay, S. Chakraborty, Arch. Virol. **156**, 769–783 (2011)
23. A.P. Feinberg, B. Vogelstein, Ann. Biochem. **132**, 6–13 (1983)
24. C.M. Fauquet, R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou, Arch. Virol. **153**, 783–821 (2008)
25. R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou, C.M. Fauquet, Arch. Virol. **153**, 763–781 (2008)
26. H.C. Prasanna, M. Rai, Virol. J. **4**, 111 (2007)
27. G. Sunter, D.M. Bisaro, Virology **180**, 416–419 (1991)
28. M.D. Hartitz, G. Sunter, D.M. Bisaro, Virology **263**, 1–14 (1999)
29. X. Yang, S. Baliji, R.C. Buchmann, H. Wang, J.A. Lindbo, G. Sunter, D.M. Bisaro, J. Virol. **81**, 11972–11981 (2007)
30. I.T.D. Petty, S.C. Carter, M.R. Morra, J.L. Jeffrey, H.E. Olivey, Virology **277**, 429–438 (2000)

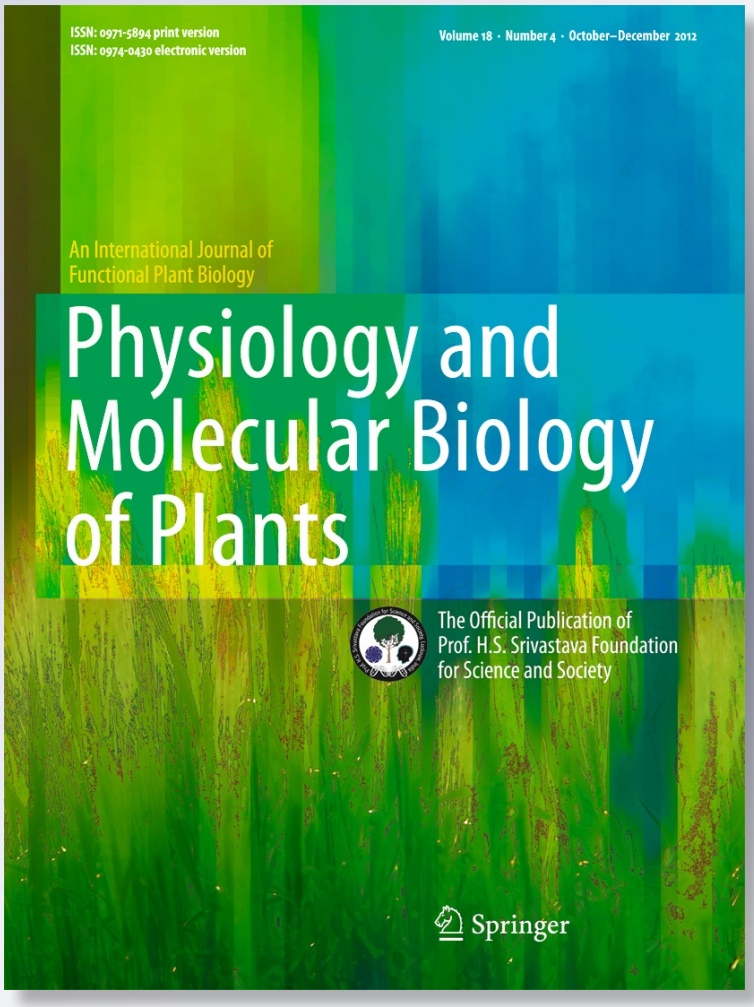
An improved plant regeneration and Agrobacterium *- mediated transformation of red pepper (*Capsicum annuum *L.)*

R. Vinoth Kumar, V. K. Sharma, B. Chattopadhyay & S. Chakraborty

Physiology and Molecular Biology of Plants
An International Journal of Functional Plant Biology

ISSN 0971-5894
Volume 18
Number 4

Physiol Mol Biol Plants (2012)
18:357-364
DOI 10.1007/s12298-012-0132-8



Springer

Your article is protected by copyright and all rights are held exclusively by Prof. H.S. Srivastava Foundation for Science and Society. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Author's personal copy

Physiol Mol Biol Plants (October–December 2012) 18(4):357–364
DOI 10.1007/s12298-012-0132-8

RESEARCH ARTICLE

An improved plant regeneration and *Agrobacterium* - mediated transformation of red pepper (*Capsicum annuum* L.)

R. Vinoth Kumar · V. K. Sharma · B. Chattopadhyay · S. Chakraborty

Published online: 23 August 2012
© Prof. H.S. Srivastava Foundation for Science and Society 2012

Abstract *Capsicum annuum* (red pepper) is an important spice cum vegetable crop in tropical and subtropical countries. Here, we report an effective and reproducible auxin free regeneration method for six different red pepper cultivars (ACA-10, Kashi Anmol, LCA-235, PBC-535, Pusa Jwala and Supper) using hypocotyl explants and an efficient *Agrobacterium*-mediated transformation protocol. The explants (hypocotyls, cotyledonary leaves and leaf discs) collected from axenic seedlings of six red pepper cultivars were cultured on either hormone free MS medium or MS medium supplemented with BAP alone or in combination with IAA. Inclusion of IAA in the regeneration medium resulted in callus formation at the cut ends of explants, formation of rosette leaves and ill defined shoot buds. Regeneration of shoot buds could be achieved from hypocotyls grown in MS medium supplemented with different concentrations of BAP unlike other explants which failed to respond. Incorporation of GA_3 in shoot elongation medium at 0.5 mg/l concentration enhanced the elongation in two cultivars, LCA-235 and Supper, while other cultivars showed no significant response. Chilli cultivar, Pusa Jwala was transformed with βC1 ORF of satellite DNA β molecule associated with Chilli leaf curl Joydebpur virus through *Agrobacterium tumefaciens*. Transgene integration in putative transformants was confirmed by PCR and Southern hybridization analysis.

Keywords *Capsicum annuum* · Plant regeneration · *Agrobacterium tumefaciens* · Genetic transformation · Transgenic plants · Chilli leaf curl virus

R. V. Kumar · V. K. Sharma · B. Chattopadhyay · S. Chakraborty (✉)
School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: supriyachakrasls@yahoo.com

Introduction

Capsicum annuum, a member of the family *Solanaceae,* is an important crop grown worldwide. It is used as a spice, a source of red pigment and as a vegetable in many countries. In addition to its use as a food, its powder disinfects the oral and gastric mucous membranes and destroys the pathogenic bacteria in the intestine (Bosland and Votava 1999). Red pepper is susceptible to many pathogens including viruses, fungi, bacteria and nematodes. The approaches aiming for genetically engineered red pepper plants for making resistant varieties or value addition are protracted due to the lack of highly efficient regeneration method. Because of lack of efficient transformation method for the cultivated species of red pepper, understanding of host virus interaction is limited to the model plant like *Arabidopsis thaliana* or in transient system using leaf disc assay.

Plant regeneration via organogenesis and somatic embryogenesis in *Capsicum* spp. has been reported for diverse explants such as cotyledons (Gunay and Rao 1978; Ebida and Hu 1993; Manoharan et al. 1998; Golegaonkar and Kantharajah 2006), hypocotyls (Gunay and Rao 1978; Ramirez-Malagon and Ochoa-Alejo 1996), leaf discs (Sanatombi and Sharma 2008) and nodal segments (Ahmad et al. 2006). However, these reports are genotype-specific and consequently, the regeneration protocol as well as transformation procedure need to be established for each commercial cultivar for exploiting the potential of genetic engineering. Additionally, achieving elongation of in vitro formed buds was a major problem in red pepper (Valera-Montero and Ochoa-Alejo 1992) and sweet pepper varieties (Ebida and Hu 1993). Ill-defined buds or leafy shoot structures either resist elongation or produce rosettes of distorted leaves which generally do not develop into normal shoots. Cytokinins, in supplement with auxins, have been shown to induce shoot buds proliferation from different types of

Author's personal copy

explants in various plant species. Almost all the protocols of red pepper regeneration using hypocotyls rely on both the cytokinins and auxins (Gunay and Rao 1978; Sanatombi and Sharma 2008). However, *Capsicum* is known to have high endogenous auxin content and therefore, the exogenous supply of auxins induces callus formation instead of shoot buds (Christopher and Rajam 1994).

Recent advances in genetic engineering have opened up new avenues for crop improvement. Genetic engineering of red pepper for the purpose of obtaining useful traits like disease resistance, stress tolerance and studying the functions of host genes is dependent upon the development of a reliable regeneration and transformation protocols. Unlike other members of the *Solanaceae* family, red pepper is considered to be a highly recalcitrant species with respect to in vitro regeneration and genetic transformation. *Agrobacterium*-mediated transformation is widely applied for red pepper transformation but its efficiency relies on the effective shoot regeneration and the genotype. The protocols available for red pepper transformation either could not be reproduced successfully or with very low efficiency in different genotypes (Manoharan et al. 1998; Zhu et al. 1996; Lee et al. 2004; Li et al. 2003).

In this paper, we report a reliable auxin-free (or auxin-independent) regeneration protocol for six different cultivars using hypocotyls with significantly improved regeneration frequency. A highly reproducible transformation protocol for *C. annuum L.* var Pusa Jwala has also been demonstrated.

Materials and methods

Establishment of axenic seedlings

Seeds of all red pepper cultivars were obtained from Indian Institute of Vegetable Research, Varanasi, India. Mature seeds were washed 3 times with sterile distilled water and surface sterilized with 70 % ethanol for 1 min followed by 3 immediate washes with sterile distilled water. Finally, they were submerged with 0.1 % mercuric chloride for 5 min with continuous shaking and then rinsed several times with sterile distilled water. The seeds were blotted dry on sterile filter paper and sown in bottles containing 50 ml of MS medium containing 3 % sucrose (Murashige and Skoog 1962). pH of the medium was adjusted to 5.8 with 1 N KOH before the addition of 0.25 % phytagel (Sigma) and autoclaved at 121 °C for 20 min. Seeds were germinated at temperature of 25±1 °C and 16-h photoperiod (irradiance of 100 μmol m^{-2} s^{-1}). Hypocotyls and cotyledonary leaves were taken from 2 weeks old seedlings while leaf discs were taken from 4 weeks old plants. All the above steps were carried out under sterile conditions.

Culture media and conditions

Shoot regeneration ability of different explants was tested on MS medium alone or supplemented either with different concentrations of BAP (varied from 1.0 mg/L to 5.0 mg/L) or in combination with IAA (0.1 mg/L to 1.0 mg/L). Explants producing shoot buds were transferred to elongation medium comprising of MS medium with or without GA_3. Elongated shoots were excised from the explants and transferred to rooting medium containing either MS basal medium or in supplementation with NAA 1.0 mg/L. In vitro regenerated plants were potted in agropeat with covers to maintain moisture for 2 weeks. After 2 weeks, the edges of the covers were cut and plants were kept for another 1 week under controlled conditions before finally transferring them to green house. The plants were then transferred to pots containing soil in green house. Regeneration response (in per cent) was calculated by dividing the total number of responding explants with total number of explants taken for the study. To optimize minimal inhibitory dosage of kanamycin on shoot induction, varying concentrations of kanamycin on *C. annuum* var Pusa Jwala explants (0, 10, 30, 50 and 100 mg/L) were tested.

Cloning of viral ORF in plant transformation vector

For transformation, βC1 ORF of Chilli leaf curl Joydebpur virus (GenBank accession number: EF190216) was PCR amplified using forward primer (5′-*GGAT CC*ATGTTTCACCTCACAAAAT-3′) and reverse primer (5′-*GTCGAC*TCATATATGTGCATTTATACC-3′). PCR reaction was performed as follows: initial denaturation at 94 °C for 4 min; followed by 10 cycles at 94 °C for 1 min, 62 °C for 45 s, 72 °C for 45 s and the reaction was terminated at 72 °C for 15 min as final extension step. The amplified product was analyzed by 1 % agarose gel electrophoresis and cloned into binary vector pBinAR to generate pBinARβC1. After confirming the clone by sequencing, the construct pBinARβC1 was mobilized into *Agrobacterium tumefaciens* strain EHA 105.

Agrobacterium-mediated transformation

The hypocotyl explants from *C. annuum* var Pusa Jwala were placed on pre-culture medium containing BAP alone for 2 days. A single colony of *A. tumefaciens* strain EHA 105 carrying pBinARβC1 was inoculated into 4 ml LB medium containing kanamycin (50 mg/L) and rifampicin (30 mg/L) and incubated at 28 °C for a day. One ml of this bacterial culture was used as starter culture and inoculated into fresh 50 ml LB medium containing the above mentioned antibiotics. The culture was then allowed to grow at 28 °C for 6–8 h until the OD reaches 0.6–0.8. The cells were

Author's personal copy

then harvested at 5,000 rpm, 28 °C for 10 min and were diluted in 30 ml liquid MS medium to obtain cell density (OD) of 0.05–0.1. The explants were submerged into bacterial culture for 5–10 min on a rotary shaker at low speed (120 rpm), blot dried on sterile filter paper and co-cultivated with the bacterial culture provided with 100 μM of acetosyringone for 2 days on co-cultivation medium. After 2 days, the explants were washed with liquid MS medium containing 500 mg/L cefotaxime and blot dried. Furthermore, they were cultured on selection medium (shoot regeneration medium) provided with the above mentioned antibiotics for shoot regeneration.

Molecular characterization of transgenic plants

Total genomic DNA was extracted from the leaves according to Dellaporta et al. (1983) and polymerase chain reaction (PCR) using gene specific primers was carried out to detect the presence of transgene. The following oligonucleotides were used for PCR analysis: primers specific for viral βC1 ORF of Chilli leaf curl Joydebpur virus (as described earlier) generating a fragment size of 380 bp and primers specific for the gene, *vir*-D1 [*A. tumefaciens*] generating a fragment of ~200 bp using the primer forward sequence as 5′-ATGTCGCAAGGCAGTAAGCCC-3′ and the reverse as 5′-GGAGTCTTTCAGCATGGAGCA-3′. The reaction was performed as follows: initial denaturation at 94 °C for 4 min; followed by 30 cycles at 94 °C for 1 min, 62 °C for 45 s, 72 °C for 5 min and the reaction was terminated at 72 °C for 15 min as final extension step. The amplified products were analyzed by 1 % agarose gel electrophoresis. For Southern hybridization and to check transgene integration, 20 μg of total DNA from each test plant was digested with *Hind*III and pBinARβC1 digested with *Bam*HI – *Sal*I was used to release viral ORF. Digested DNAs were fractioned on 0.8 % agarose gel and blotted by capillary method onto Hybond N^+ membrane (Amersham Pharmacia, UK) using transfer buffer as 20X SSC. Southern blots were hybridized with [α^{32}P] dCTP labelled βC1 ORF using random labelling method (Amersham Biosciences). The hybridization was carried out according to Sambrook and Russell (2001).

Results

Selection of suitable explants for regeneration

In the preliminary experiments, various explants such as 2 weeks-old cotyledons, hypocotyls (0.5–1 cm) and 4 weeks-old leaf discs of all six red pepper cultivars were tested for shoot induction on MS medium supplemented with different concentrations of BAP and IAA. Among the different explants tested, only hypocotyls showed shoot induction while the other explants formed callus at the cut end of the explants or formation of rosette leaves and ill-defined shoot buds. Hence, we have used hypocotyls as explants for further experimentations.

Shoot regeneration using hypocotyls explants

In the first phase of experiment, a suitable auxin independent protocol was established to determine the regeneration potential of hypocotyl explants. The regeneration potential of hypocotyls from six popular Indian red pepper cultivars were tested either on MS medium supplemented without or with BAP (1–4 mg/L). The regeneration potential of each red pepper cultivars was calculated as the percentage of hypocotyls showing initial regeneration response (Fig. 1). Shooting was observed as early as 8 days from the tip of hypocotyls on shoot induction medium with 3 to 4 shoot buds emerging from hypocotyl tip (Fig. 2a). The red pepper cultivars, Kashi Anmol and LCA-235 showed regeneration frequency of 96 % and 80 %, respectively in medium supplemented with BAP 4.0 mg/L. Chilli varieties, Pusa Jwala and PBC-535 showed variable degree of regeneration frequency of 90 % (in BAP 2.0 mg/L) and 93 % (in BAP 3.0 mg/L), respectively. The cultivar Supper showed regeneration frequency of 80 % on MS medium and ACA-10 cultivar showed 60 % regeneration potential on MS medium with BAP 1.0 mg/L. Our results using six different cultivars have shown that no single combination was inclusively effective for all six varieties tested (Fig. 1).

Elongation and rooting of shoots

For elongation of small shoot buds, we have used either MS basal medium or MS medium supplemented with different concentrations of GA_3. We could achieve elongation of desired shoot buds in all the genotypes in MS medium

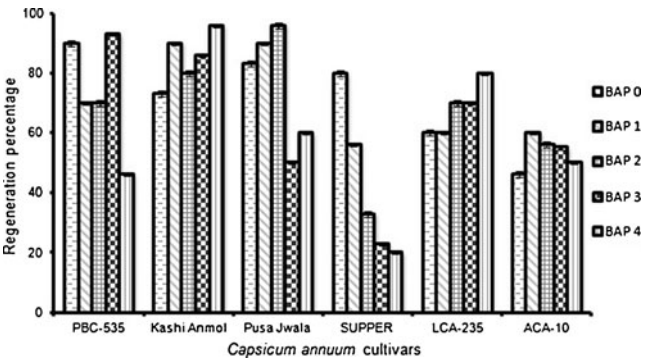


Fig. 1 The regeneration potential (in per cent) of six different cultivars of Indian red pepper cultured on different hormonal combinations. The experiment was repeated thrice and the values were representative of average regeneration response after 14 days on shoot bud induction medium and expressed as means ± SE (*n=10*). The concentration of BAP was used as mg/L

Author's personal copy

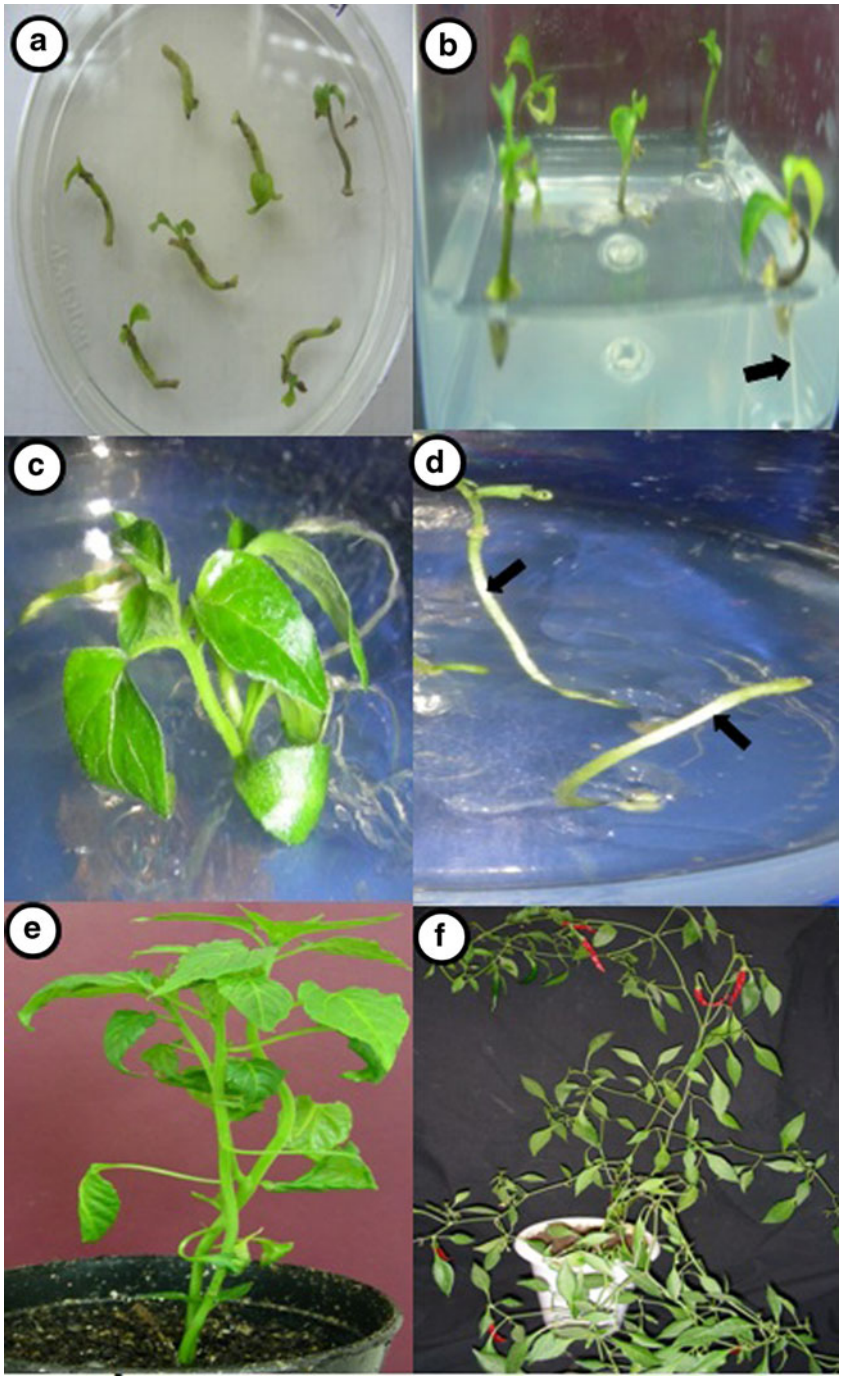


Fig. 2 Different stages of plant regeneration using hypocotyls explants of *Capsicum annuum*; **a** Induction of shoot bud after 1 week culture, **b**- Shoot bud elongation and rooting after 3 weeks culture, **c** Shoot elongation on MS basal medium, **d** Shoot elongation on MS+1 mg/L GA_3 medium, **e** Well elongated plantlet after hardening on agropeat, **f** Fully grown regenerated plantlet in growth chamber

lacking any hormones (Fig. 2b). Our results also showed that MS medium supplied with 0.5 mg/L GA_3 enhanced the rate of elongation of shoot buds but further increase in GA_3 (1.0–2.0 mg/L) concentration did not provide any advantage (Fig. 2c and d). For achieving roots, the well elongated shoots were cultured on MS medium supplemented with NAA. We observed 100 % rooting from the shoots transferred to MS medium provided with 0.5 mg/L NAA. These well rooted plantlets were then acclimatized in growth chamber under controlled growth conditions.

Transformation of red pepper var Pusa Jwala

To establish an efficient method for *Agrobacterium*-mediated transformation, red pepper hypocotyls explants from 2-weeks old in vitro grown seedlings of *C. annuum* var Pusa Jwala and *A. tumefaciens* harbouring pBinARβC1 (Fig. 3) were used. Prior to infection, the hypocotyl explants were pre-cultured on MS medium containing BAP- 2 mg/L. The cell density of *Agrobacterium* was a critical factor for the regeneration of infected explants. After exposure to

Author's personal copy

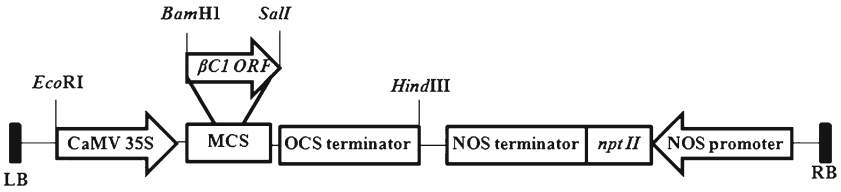


Fig. 3 T-DNA construct of binary vector, pBINAR harboring viral βC1 ORF of Chilli leaf curl Joydebpur virus. Viral gene was cloned into *Bam*H1 and *Sal*I restriction sites. *LB* Left border, *RB* Right border, *CaMV 35S Cauliflower mosaic virus* 35S promoter, *MCS* Multiple cloning site, *OCS Octapine synthase*, *NOS Nopaline synthase*

undiluted *Agrobacterium* culture of OD 0.5–1.0, red pepper explants showed severe necrosis resulting in reduced regeneration potential. Hence, the cell density was diluted and we optimized the culture OD to 0.05–0.1. The same concentration was subsequently used for transformation of hypocotyls.

After infection of the red pepper hypocotyl explants with *A. tumefaciens* strain EHA-105 harbouring pBinARβC1, these explants were co-cultivated on MS medium containing BAP-2 mg/L and 100 μM of acetosyringone for 2 days in dark condition. Following co-cultivation with *Agrobacterium*, the hypocotyls were transferred to selection medium. For selecting transformants, we used kanamycin as a selectable agent in the culture medium along with cefotaxime (500 mg/L). In earlier reports, varying amount of kanamycin was used for selection of transformants of sweet pepper varieties. At 100 mg/L of kanamycin concentration, we noticed that very few explants survived and hence, we further used 100 mg/L kn concentration as minimal inhibitory dosage for *Agrobacterium*-mediated red pepper transformation. The regenerated shoots on selection media were sub-cultured every 2 weeks for a month and the explants were conditioned as described earlier. All those regenerated plants were phenotypically normal like non-transformed lines when grown on soilrite (Fig. 2e) and could bear fruits like non-transgenic ones on soil (Fig. 2f).

Molecular characterization of putative transformants

About 90 hypocotyls were taken for *Agrobacterium*-mediated transformation and 11 transformants were obtained on selection medium. Total genomic DNA was isolated from the leaves of all kanamycin resistant plants and subjected to PCR analysis using Chilli leaf curl Joydebpur βC1 specific primers. The amplification of viral gene in PCR confirmed T_0 plants were shown in Fig. 4a. To exclude the possibility of *Agrobacterium* contamination in transformed plants, total DNAs were analysed with PCR using primers specific for *vir*-D1 region of this bacterium. There was no PCR amplification observed in the samples except from the DNA isolated from *Agrobacterium* (Fig. 4b). PCR analysis was repeated at least thrice which confirmed the presence of transgene in the putative transformed plants on selection medium. Genomic DNA isolated from PCR positive lines was subjected to Southern blot analysis by digesting with *Hind*III that does not cut internally within the transgene. Hence, number of bands generated upon digestion with *Hind*III corresponds to the integrated copy number of the transgene. Southern blot analysis revealed the transgene integration and its number of copies among the progeny of three individual T_0 plants analysed. A single hybridization signal was detected from three independent lines whereas as expected no hybridization signal was detected in non-transformed line (Fig. 4c). These data provided the confirmation of the integration of transgene in the red pepper genome.

Discussion

Plant regeneration remains an important bottleneck for the successful establishment of highly reproducible transformation. Optimization of different parameters (age of explant, type of explant, medium requirement, concentration of phytohormones) is critical for plant regeneration because of their direct effect on genetic transformation efficiency. In this communication, we report an improved method for regeneration of six Indian red pepper cultivars and a reliable *Agrobacterium*-mediated transformation for the cultivar, *Pusa Jwala*.

In earlier reports of red pepper regeneration, various explants viz. cotyledonary leaves, hypocotyls, shoot tips, leaf discs and nodal segments have been studied to test their regeneration potential but the response for different explants varied from one genotype to another. However, in red pepper, hypocotyls and cotyledonary leaves had responded well than other explants reported so far (Gunay and Rao 1978; Ramirez-Malagon and Ochoa-Alejo 1996; Manoharan et al. 1998). In preliminary stage, we have started with all possible explants such as cotyledonary leaves, leaf discs and hypocotyls. However, except hypocotyls, no other explants responded in the auxin/cytokinin medium. The callus induction at the cut end of cotyledonary leaves and leaf discs was observed and thus we preferred hypocotyls for further experiments. In pepper tissue culture, BAP has been highly effective in shoot bud induction than other cytokinins (Gunay and Rao 1978; Joshi and Kothari 2007; Sanatombi

Author's personal copy

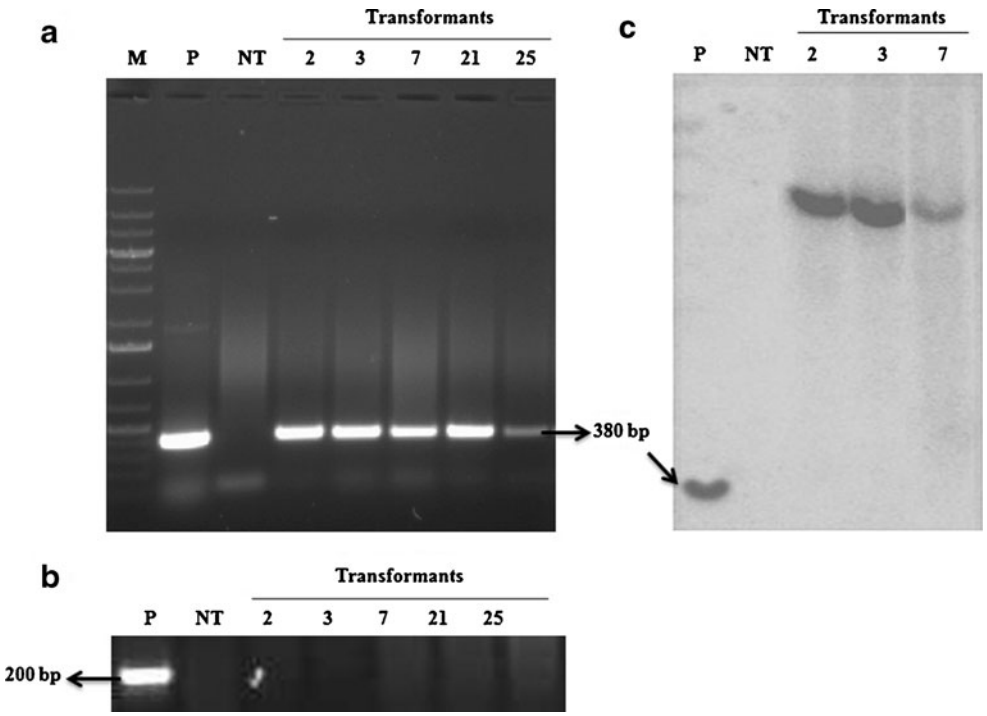


Fig. 4 Molecular analysis of transgenic *C. annuum* var Pusa Jwala by PCR (**a**,**b**) and Southern hybridization (**c**); **a** detection of βC1 ORF of *Chilli leaf curl Joydebpur virus* where pBINARβC1 ORF represents positive control (P); non-transgenic lines as negative control (NT); transgenic lines 2, 3, 7, 21 and 25 (Transformants); marker (M); **b** PCR analysis using vir-D1 gene specific primers with plasmid DNA from *Agrobacterium tumefaciens* as positive control (P); **c** Integration of βC1 fragment in red pepper. Total DNAs from plantlets were digested with Hind III and 380 bp viral ORF was used both as probe and as well as positive control (P)

and Sharma 2008). The earlier reports of red pepper regeneration emphasized on a balanced auxin/cytokinin ratio for multiple shoot buds induction (Steinitz et al. 1999). In contrast, the formation of callus and leafy shoots in the regeneration medium supplied with low/high auxin could be due to the high endogenous auxin content of red pepper (Christopher and Rajam 1994). Hence, we opted for an auxin-free (independent) approach for regeneration of hypocotyls in all six red pepper cultivars. However, in contrast to earlier reports (Christopher and Rajam 1994, 1996), the addition of triiodobenzoic acid (TIBA) has not improved multiple shoot bud proliferation in the present study. On auxin-free regeneration medium, a variation in the regeneration response of hypocotyls was obsereved. In consistent with Ochoa-Alejo and Ireta-Moreno (1990), this variation among the cultivars could be due to the difference in their genotypic nature. This result convincingly emphasizes the role of type and nature of the genotype used in the regeneration of different *Capsicum* species.

The formation of rosette leaves, malformed shoots or blind leaf structures which fails to elongate further has been a major obstacle in red pepper tissue culture practices (Steinitz et al. 1999). Earlier reports cautioned that even high cytokinin level could inhibit the elongation of newly formed shoot (Huetteman and Preece 1993). To minimize the inhibitory effect of cytokinin used in shoot proliferation medium, we have attempted shoot elongation in MS medium devoid of cytokinin. All the explants could achieve elongation well irrespective of their genotypic nature. GA_3 alone or in combination with silver nitrate/copper sulphate was found to be efficient for shoot elongation in different genotypes (Hyde and Phillips 1996; Golegaonkar and Kantharajah 2006; Joshi and Kothari 2007; Kumar et al. 2007). However, unexpectedly in our work at higher GA_3 concentration (>1 mg/L), the elongated shoots of two cultivars (LCA-235 and Supper) appeared thin and weak which could not survive after hardening. GA_3 supplement enhanced the shoot elongation; however it could not result in elongation of small leafy structures or ill defined buds. This emphasized that MS medium lacking phytohormones could be effective for shoot elongation. NAA has been used for in vitro rooting of shoots in *Capsicum* and other species. But strikingly, the observation of rooting was quite evident in 70 % of the shoots on elongation medium lacking any auxins. Rooting in parallel with elongation of shoots, reduces the overall duration taken for generating transgenic plants.

Agrobacterium-mediated transformation is widely used for genetic transformation in plant system. Chilli unlike other members of the *Solanaceae* family is known to be a highly recalcitrant species. The available red pepper transformation protocols are either irreproducible or having less transformation efficiency (Li et al. 2003; Lee et al. 2004), which reflects the necessity for an effective transformation method. For achieving high transformation efficiency, it is necessary to optimize the critical factors affecting the health of the explants after infection. One of them is the *Agrobacterium* cell density used for transformation. Li et al. (2003) suggested use of *Agrobacterium* cell density as OD 0.3–0.5 for red pepper transformation. However, at the same cell density we noticed *Agrobacterium* accumulation at the cut ends of hypocotyls till 30 days after transfer to selection medium. This residual *Agrobacterium* completely colonizes

Author's personal copy

resulting in necrosis and thereby, reducing their regeneration potential. Similarly, at OD 0.5–1.0, grape leaf discs showed severe necrosis whereas this effect was reduced to greater extent by diluting the cell density (Das et al. 2002). Hence, we optimized culture density by diluting to OD <0.1 to prevent bacterial colonization. In addition, removal of the excessive culture from the explants after infection is important to prevent the accumulation of bacterial cells and ultimately death of the explants.

After co-cultivation, the putative transformants were selected on kanamycin selection medium because of its significant role played in recovering transformed shoots. PCR analysis using viral gene specific primers demonstrated that 12.2 % of the selected transformants contained the desired transgene while *Agrobacterium* contamination is ruled out as confirmed by PCR using *vir*-D1 specific primers. This observed response could be attributed to the combined effect of parameters optimized such as bacterial cell density and phytohormone combination used for hypocotyl regeneration. In plants transformed with transgene under constitutive promoter, a negative correlation could be possible between transgene number and its expression due to homology-dependent transcriptional gene silencing (Datla et al. 1993). Hence, Southern blot analysis was performed to determine the number of transgene integrated in the red pepper genome. This present analysis yielded a single band and as a result, 3 out of 11 independent lines proved to be βC1-transgenic with single copy transgene integration in the genome. The integration of transgene in a single position could eliminate the possibility of homology based transgene silencing.

In summary, we have demonstrated a reliable auxin free regeneration protocol and the capability of shoot elongation and rooting using hormone free MS medium for six popular Indian red pepper cultivars. This method shall considerably reduce the time required for regeneration and transformation of red pepper cultivars. The βC1 transgenic plants developed in this study will certainly help in understanding the βC1 mediated pathogenesis in its natural host, a recalcitrant cultivated crop. The genetic transformation of red pepper described here also opens up new possibilities in generating transgenic red pepper plants desired for superior agronomic traits and in studying the gene function of *Capsicum* species.

Acknowledgements The research work was financially supported by the Department of Biotechnology, Government of India, New Delhi. VKS gratefully acknowledge Council of Scientific and Industrial Research for providing fellowship.

References

Ahmad N, Siddique I, Anis M (2006) Improved plant regeneration in *Capsicum annuum*. L from nodal segments. Biol Plant 50(4):701–704. doi:10.1007/s10535-006-0110-5

Bosland PW, Votava EJ (1999) Peppers: vegetable and spice capsicums. CABI publishing, New York, p 177

Christopher T, Rajam MV (1994) In vitro clonal propagation of *Capsicum* species. Plant Cell Tissue Organ Cult 38:25–29. doi:10.1007/BF00034439

Christopher T, Rajam MV (1996) Effect of genotype, explant and medium on in vitro regeneration of red pepper. Plant Cell Tissue Organ Cult 46:245–250. doi:10.1007/BF02307101

Das DK, Reddy MK, Upadhyaya KC, Sopory SK (2002) An efficient leaf-disk culture method for the regeneration via somatic embryogenesis and transformation of Grape (*Vitis vinifera* L.). Plant Cell Rep 20:999–1005. doi:10.1007/s00299-002-0441-4

Datla RS, Bekkaoui F, Hammerlindl JK, Pilate G, Dunstan D, Crosby W (1993) Improved high-level constitutive foreign gene expression in plants using an AMV RNA4 untranslated leader sequence. Plant Sci 94:139–149. doi:10.1016/0168-9452(93)90015-R

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1:19–21. doi:10.1007/BF02712670

Ebida AIA, Hu CY (1993) In vitro morphogenetic responses and plant regeneration from pepper (*Capsicum annuum* cv. Early California Wonder) seedling explants. Plant Cell Rep 13:107–110. doi:10.1007/BF00235301

Golegaonkar PG, Kantharajah AS (2006) High frequency adventitious shoot bud induction and shoot elongation of chile pepper (*Capsicum annuum.* L). In Vitro Cell Dev Biol – Plant 42:341–344. doi:10.1079/IVP2006785

Gunay AL, Rao PS (1978) In vitro regeneration from hypocotyls and cotyledon explants of red pepper. Plant Sci Lett 11:365–372. doi:10.1016/0304-4211(78)90024-X

Huetteman CA, Preece JE (1993) Thidiazuron: a potent cytokinin for woody plant tissue culture. Plant Cell Tissue Organ Cult 33:105–119. doi:10.1007/BF01983223

Hyde C, Phillips GC (1996) Silver nitrate promotes plant regeneration of chile pepper (*Capsicum annuum* L.) via organogenesis. In Vitro Cell Dev Biol – Plant 32:72–80. doi:10.1007/BF02823134

Joshi A, Kothari SL (2007) High copper levels in the medium improve shoot bud differentiation and elongation from the cultured cotyledons of *Capsicum annuum* L. Plant Cell Tissue Organ Cult 88:127–133. doi:10.1007/s11240-006-9171-6

Kumar V, Sharma A, Prasad BCN, Gururaj HB, Giridhar P, Ravishankar GA (2007) Direct shoot bud induction and plant regeneration in *Capsicum frutescens* Mill: influence of polyamines and polarity. Acta Physiol Plant 29:11–18. doi:10.1007/s11738-006-0002-5

Lee YH, Kim HS, Kim JY, Jung M, Park YS, Lee JS, Choi SH, Her NH, Lee JH, Hyung NI, Lee CH, Yang SG, Harn CH (2004) A new selection method for pepper transformation: callus-mediated shoot formation. Plant Cell Rep 23:50–58. doi:10.1007/s00299-004-0791-1

Li D, Zhao K, Xie B, Zhang B, Luo K (2003) Establishment of a highly efficient transformation system for pepper (*Capsicum annuum*). Plant Cell Rep 21:785–788. doi:10.1007/s00299-003-0581-1

Manoharan M, Vidya CSS, Sita GL (1998) *Agrobacterium*-mediated genetic transformation in red pepper (*Capsicum annuum* var. Pusa Jwala). Plant Sci 131:77–83. doi:10.1016/S0168-9452(97)00231-8

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays of tobacco tissue cultures. Physiol Plant 15:473–497. doi:10.1111/j.1399-3054.1962.tb08052.x

Ochoa-Alejo N, Ireta-Moreno L (1990) Cultivar differences in shoot forming capacity of hypocotyl tissue of red pepper pepper

Author's personal copy

(*Capsicum annuum*) cultured in vitro. Sci Horti 42:21–28. doi:10.1016/0304-4238(90)90144-4

Ramirez-Malagon, Ochoa-Alejo N (1996) An improved and reliable chilli pepper (*Capsicum annuum*) plant regeneration methods. Plant Cell Rep 16:226–231. doi:10.1007/BF01890873

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual; 3rd edition. Cold Spring Harbor Laboratory Press, New York, USA, pp 6.39–6.46

Sanatombi K, Sharma GJ (2008) In vitro plant regeneration in six cultivars of *Capsicum* species using different explants. Biol Plant 52(1):141–145. doi:10.1007/s10535-008-0029-0

Steinitz B, Wolf D, Matzevitch-Josef T, Zelcer A (1999) Regeneration in vitro and genetic transformation of pepper (*Capsicum* species)-The current state of art. Capsicum Egg Plant Newsl 18:9–15

Valera-Montero LL, Ochoa-Alejo N (1992) Novel approach for red pepper (*Capsicum annuum*) plant regeneration: shoot induction in rooted hypocotyls. Plant Sci 84:215–219. doi:10.1016/0168-9452(92)90137-B

Zhu YX, Ouyang WJ, Zhang YF, Chen ZL (1996) Transgenic sweet pepper plants from *Agrobacterium*-mediated transformation. Plant Cell Rep 16:71–75. doi:10.1007/BF01275453

New Disease Reports

A new monopartite begomovirus species, *Chilli leaf curl Vellanad virus*, and associated betasatellites infecting chilli in the Vellanad region of Kerala, India

R. Vinoth Kumar, A.K. Singh and S. Chakraborty

Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi -110 0676, India

*E-mail: supriyachakrasls@yahoo.com

Received: *19 Jan 2012.* ***Published:*** *05 Apr 2012.* ***Keywords:*** *Capsicum annuum, leaf curling, sequence identity*

Capsicum annuum (chilli) is one of the most important commercial crops cultivated in the Indian sub-continent for vegetable, spice and industrial purposes. Leaf curl disease of chilli has become a serious problem in the Vellanad district of Kerala, one of the major chilli growing areas of India. Disease incidence up to 100% was recorded during December 2008 in this region. The infected plants exhibited typical upward leaf curling, crinkling, puckering and reduction of leaf area along with stunting of whole plants (Fig. 1).

Total DNA was isolated from chilli plants and applied as template for PCR amplification of begomovirus-specific fragments using degenerate primer pairs (PAL1v1978/PAR1c496) for detection of *Begomovirus* DNA-A (Rojas *et al*., 1993) and Beta01/Beta02 for DNA- (Briddon *et al*., 2002). Specific fragments of 1.5 kb for DNA-A and 1.4 kb for DNA-, respectively, were obtained from samples associated with symptoms. Rolling circle amplification (RCA) using φ29 DNA polymerase was carried out on the total DNA preparation, followed by digestion with *Kpn I*. An amplicon of ~2.8 kb was gel-eluted and cloned into *Kpn* I-linearized pBluescript II KS(+). Restriction digestion of plasmid DNA from the resulting clones indicated the presence of one type of molecule. Betasatellite-specific amplicons of ~1.4 kb were cloned into pTZ57R/T. The corresponding satellite (sat) DNAs could be categorised into two groups based on their restriction pattern. In spite of repeated attempts, we failed to isolate DNA-B from the plants showing appropriate symptoms, indicating that the isolate was a monopartite begomovirus.

Full-length sequencing revealed a viral genome (HM007121) of 2,788 nt with seven predicted ORFs (AV1, AV2, AC1, AC2, AC3, AC4, AC5). Sequence alignment using MegAlign (DNASTAR, Madison, WI) showed the virus to have the highest levels of nucleotide sequence identity (77%) with a 'Pakistan" strain of *Pepper leaf curl Bangladesh virus* (AF314531). With other chilli-infecting begomoviruses in the Indian sub-continent, there was between 64 and 76% sequence identity. The present virus is distinctly placed in a separate clade in the phylogenetic tree (Fig. 2). According to the ICTV guidelines (Fauquet *et al*., 2008), it has to be considered a new begomovirus species for which the name *Chilli leaf curl Vellanad virus* (ChiLCVeV) is proposed. The associated betasatellites (GenBank Accession Nos. JN663876, 1355 nt and JN663877, 1372 nt) each contain a single ORF C1. They share maximum identity of 95.6% and 96.9% with *Radish leaf curl betasatellite* (EF175734) and *Tomato leaf curl Bangladesh betasatellite* (AY438558), respectively. Following ICTV guidelines (Briddon et al., 2008) along with phylogenetic analysis (Fig. 3) and nucleotide sequence identity data, the two molecules are therefore identified as isolates of satDNAs described earlier. Old World begomoviruses have been shown to be widely associated with one or more DNA betasatellites (Briddon *et al*., 2003; Chattopadhyay *et al*., 2008). Together, these findings indicate that a new begomovirus species, ChiLCVeV,along with two distinct betasatellites (*Radish leaf curl betasatellite* and *Tomato leaf curl Bangladesh betasatellite)* are associated with chilli leaf curl disease in Vellanad, Kerala.

Acknowledgements

The authors gratefully acknowledge the Department of Biotechnology, India, for funding the research.

References

Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X, Fauquet CM, 2008. Recommendations for the classification and nomenclature of the DNA-β satellites of begomoviruses. *Archives of Virology* **153**, 763-781. [doi:10.1007/s00705-007-0013-6]

Briddon RW, Bull SE, Amin I, Idris AM, Mansoor S, Bedford ID, Dhawan P, Rishi N, Siwatch SS, Abdei-Salam AM, Brown JK, Zafar Y, Markham PG, 2003. Diversity of DNA β; a satellite moleculeassociatedwith some monopartite begomoviruses. *Virology* **312**, 106-121. [doi:10.1016/j.physletb.2003.10.071]

Briddon RW, Bull SE, Mansoor S, Amin I, Markham PG, 2002. Universal primers for the PCR-mediated amplification of DNAβ: a molecule associated with some monopartite begomoviruses. *Molecular Biotechnology* **20**, 315-318. [doi:10.1385/MB:20:3:315]

Chattopadhyay B, Singh AK, Yadav T, Fauquet CM, Sarin NB, Chakraborty S, 2008. Infectivity of the cloned components of a begomovirus: DNA-β complex causing chilli leaf curl disease in India. *Archives of Virology* **153**, 533-539. [doi:10.1007/s00705-007-0017-2]

Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X, 2008. Geminivirus strain demarcation and nomenclature. *Archives of Virology* **153**, 783-821. [doi:10.1007/s00705-008-0037-6]

Rojas MR, Gilbertson RL, Russel DR, Maxwell DP, 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly transmitted geminiviruses. *Plant Disease* **77**, 340-347. [doi:10.1094/PD-77-0340]

Figure 1

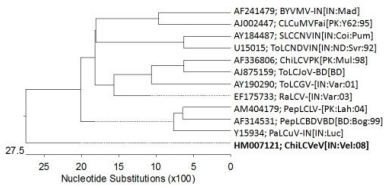


Figure 2

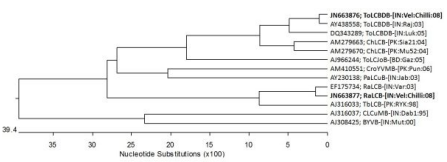


Figure 3

To cite this report: Vinoth Kumar R, Singh AK, Chakraborty S, 2012. A new monopartite begomovirus species, *Chilli leaf curl Vellanad virus*, and associated betasatellites infecting chilli in the Vellanad region of Kerala, India. *New Disease Reports* **25**, 20. [doi:10.5197/j.2044-0588.2012.025.020]

©2012 The Authors

This report was published on-line at www.ndrs.org.uk where high quality versions of the figures can be found.