Development of Antibodies, Infectious clones and Infectivity studies of *Cotton leaf curl Burewala virus* (CLCuBuV)

MALIK NAWAZ SHUJA

(2009-NUST-DirPhD-V&I-45)

ATTA-UR-RAHMAN SCHOOL OF APPLIED BIOSCIENCES,

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD, PAKISTAN

September, 2017

A thesis submitted in partial fulfillment of the requirement for the degree of

Doctor of Philosophy (PhD) in **Applied Biosciences**

Development of antibodies, infectious clones and infectivity studies of *Cotton leaf curl Burewala virus* (CLCuBuV)

Malik Nawaz Shuja

(2009-NUST-DirPhD-V&I-45)

SUPERVISOR: **Dr. Muhammad Tahir**
(Assistant Professor, Head of Department,
Plant Biotechnology, ASAB, NUST, Islamabad)

OVERSEAS STUDY HOST: **Dr. Stephan Winter**
(Plant Virus Head,
Leibniz Institute DSMZ-Braunschweig, Germany)

PLACES OF STUDY: Atta-ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad

Leibniz Institute DSMZ-Deutsche Sammlung von Mikroorghanismen und Zellkulturen, Messeweg, Braunschweig, Germany

ABSTRACT

Cotton leaf curl disease (CLCuD) is a major hindrance to cotton production across Pakistan. The disease is caused by a complex of begomoviruses with a single known betasatellite, Cotton leaf curl Multan betasatellite. Leaf samples from cotton plants, both symptomatic (leaf curling, vein thickening, enation) and non-symptomatic, were collected from major areas in Punjab, Pakistan during 2010-2013. Total DNA was isolated from these samples using a CTAB-method, followed by PCR. From the symptomatic samples, 30 full-length begomovirus components (10 DNA-A, 10 betasatellite and 10 alphasatellite) were cloned and completely sequenced. The isolates C28 and C49 with their associated cognates alpha-and beta-satellites were used to produce partial tendem repeat constructs for agroinoculation. Cotton leaf curl Burewala virus (recently re-named as Cotton leaf curl Kokhran virus-Burewala) with Cotton leaf curl Multan alphasatellite and betasatellite was the most frequent virus identified in most of the cotton belt of Punjab. One new strain, a recombinant alphasatellite was identified. For this new strain, the name "Cotton leaf curl Burewala virus-Layyah" was proposed. The partial tandem repeat constructs were infectious to *Nicotiana benthamiana, N. tabacum, Solanum lycopersicum, Cucurbita maxima* and cucumber. For immunological studies, polyclonal antisera were raised to the coat protein of CLCuKoV-Bu. Antiserum was tested as serologically positive in arrays of assays like DAS-ELISA and western blot for detection of viral coat protein of the begomovirus. The present study is significant for the development of antibodies for rapid detection of cotton infecting begomoviruses on a large scale and production of infectious clones for screening material for resistance studies.

ACKNOWLEDGEMENTS

"Give thanks to Allah, and whoever gives thanks, he gives thanks for (the good of) his ownself. And whoever is unthankful, then verily, Allah is All-Rich (free of all needs), Worthy of All-Praise"

(Al-Quran 31:12)

All thanks are for "ALLAH Almighty" whose blessings enabled me to seek knowledge and invigorate me for this task. Words can't express my feeling of thankfulness for "ALLAH Almighty". Who have helped, gave me courage and ability to complete this task. Without the Graces of ALLAH Almighty, it would not have been accomplished.

Thank you Allah for everything, the Big things and Small! Thank you for blessing me much more than I deserve. Without You, I am Nothing.

I offer my salutation to The Holy Muhammad (peace be upon Him), the source of guidance as a whole forever. There is a saying of the Holy Prophet Muhammad (peace be upon Him)

"Seeking knowledge is an obligation upon every Muslim"

Although any work can be attributed to the skills of brains, strength of hands and the fortune's favor but this study also represent the dedicated efforts of numerous individuals, many of whom deserve special mention.

I am also highly appreciative to our former Director/Principal Dr. Ishtiaque Qadri (who opened the door of NCVI for welcome) and Dr. Muhammad Ashraf, and present Principal Dr. Peter John of Atta-ur-Rahman School of Applied Biosciences for their giant support and guidance whenever I needed.

"A Muslim shall never leave another Muslim helpless in the time of need"

I feel utmost pride and highly honored to express my sincere gratitude to respected supervisor Dr. Muhammad Tahir (HoD/ Assistant Professor). I am highly obliged and have heartfelt gratefulness to him for giving me the opportunity to work in his laboratory and under his kind and dynamic supervision. His generous, untiring guidance and remarkable suggestion from the initial to the final level has enabled me to develop an understanding of the subject. His tremendous skilled experienced and encouraging attitude granted me the confidence to face the challenges of PhD.

PhD is a long journey, many helping hands co-operated at different stages. It will be injustice, if I don't mention their names on this special occasion. I would like to pay my best gratitude to my Guidance and Examination Committee (GEC) members, including Dr. Najam-us-Saher Sadaf Zaidi, Dr. Rabia Amir and Dr.M. Zeeshan Hyder (External, COMSAT Institute of Information Technology, Islamabad). Their unlimited assistance, suitable advices and encouraging remarks always helped to overcome any hurdle that came in my way when doing PhD. These advisable guidelines will be always appreciated and remembered. I am also highly thankful to all faculty members at ASAB for their moral and spiritual support.

I feel pleasure in acknowledging my mates for their encouraging behavior and supportive statements, whenever I got down in my research. I can't ignored the time when they took me to café and released my tension with spicy foods and hot tea to change my mindset from research point of view and refresh it for the next task. I would like load of thanks to my PhD friends; Dr. Baber Aslam, Dr. Kashif, Dr. Rehan Zafar Paracha,Dr. Fahad, Dr. Fazal Jalil,

Dr. Hayat Khan, Dr. Mohsin Jamal, Dr. Ammar Raheel, Dr. Altaf Baloch,Dr. Zia-ur-Rahman Farooqi, M. Sajid, Imran Raza, Bilal Zulfiqar, Sadia Malik, Shabina Ishtiaq, Alia Sadiq, Sumaira Salahuddin, Fakiha Azal, Amna Shah, Hira, Saba Munawar and many more.

Last but not the least, my lab fellows with whom I spend most of my time. It will be completely unfairness with them if I missed their names on the stage of success. More credit goes to them for their uncountable prayers, scientific and moral support, advices and encouragement and their marvelous discussion and suggestions. I am so happy to give tribute to my dearest lab friends Dr.Muhammad Ali, Sidra Anwer, Dr. Sehrish Akbar, Muhammad Arshad, Atifa Mehmood, Aqsa Ikram, Irfan Fareed Bhatti, Kehkashan Naeem, Maria Khalid, Shaheen Bibi, Fizza Akhtar, Anam Mustujab, Azkia Khan, Javaria Khan, Yaqoot Imtiaz, Ayesha Zaheer, Amal Khan, Zeeshan Yousaf, Amir Lal, Ayesha Liaqat, Ayesha Ishfaq, Maryum Khan, Humda Qayyum and Syeda Wajiha Gilani. There are many other hands, without their support and cooperation, success can't be achieved. Miss Fozia (Lab Technician, who taught me how to work in the lab), Fida-ur-Rahman and Kamran Ahmed (providing chemicals).

With deep sense of acknowledgment, I would like to express my extreme and profound appreciation and sincere thanks to the worthy overseas supervisor Dr. Stephan Winter (Head Plant Virus, DSMZ-Germany) and his staff members. For him, there is no any match. Dr. Wulf Menzel, Dr. Monika Gotz (thanks for taking me with your family on tours), Dr. Dennis Knierim, Dr. M.R. Beena (I never feel boring and home sickness in her company). For superb technical assistance: Nanni, there seems to be nothing you cannot fix (trained me how to work in the lab, she is so sweet and polite), Beate (trained me how to work in green house), Agnes (so sweet and nice in her talks, trained me in protein expression and purification), Lara and Vera (trained me how to potting for new plants), Sabine (I have never seen such a cool minded, patient and polite lady in my whole life like Sabine) and my lovely and jolly friends Pirasteh (for hosting with Iranian foods), Hoang, Naheed (and her husband Hadi) and Samar Sheat (your own baking cake was so yummy and unique).I am indebted all of them for their help and their care about my work. I am grateful to all of them for not only giving me opportunity to work in very competitive environment but their kind nature and freedom of research also gave me confidence to learn a lot of skills and able to complete my project in short period of time. Their way of conducting research, novel ideas and remarkable suggestions are the source of inspiration for me. Their affection and hospitality made my first foreign visit memorable. I am also thankful for them for inviting me on the dinner parties and tours where I came to know about how friendly and compassionate German people are.

I would extend my thanks to the Ministry of Science and Technology (MoST), Higher Education Commission (HEC) Pakistan and ICARDA for providing funds and honoring me the scholarship to complete my degree.

I would take this opportunity to show my deep gratitude to my parents, my brothers and sisters, and my beloved wife, who experienced all the ups and downs of my research and they constantly prayed for me throughout my academic career.

Finally, thank you ALLAH Almighty, once again, for having all this for the start of my new journey in science….

Malik Nawaz Shuja

LIST OF ABBREVIATIONS

aa	amino acid
AAP	acquisition access period
APS	ammonium persulphate
BLAST	basic local alignment search tool
bp	base pair
$CaCl_2$	calcium chloride
ccc	covalently closed circular
CLCuD	cotton leaf curl disease
CP	coat protein
CR	common region
CTAB	cetyl trimethyl ammonium bromide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	double stranded DNA
dsRNA	double stranded RNA
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid disodium salt
g	gram
HR	hypersensitive response
hr	hour
ICTV	International Committee on Taxonomy of Viruses
IPTG	isopropyl-beta-D-1-thiogalactopyranoside

IR	intergenic region
IRD	iteron related domain
kb	kilo bases (nucleotides)
KCl	potassium chloride
kDa	kilo Dalton
LB	Luria-Bertani Broth
LIR	large intergenic region
M	molar
MCS	multiple cloning site
mg	milligram
$MgSO_4$	magnesiun sulphate
$MgSO_4.7H_2O$	magnesiun sulphate heptahydrate
min	min
ml	milliliter
mM	millimolar
MP	movement protein
MW	molecular weight
NaCl	sodium chloride
NaOH	sodium hydroxide
NSP	nuclear shuttle protein
nt	nucleotide
NW	New World
OD	optical density
ORF	open reading frame

OW	Old World
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PDR	pathogen derived resistance
pH	paviour of hydrogen
PVP	polyvinyl pyrrolidone
RCA	rolling cycle amplification
RCR	rolling cycle replication
RDP	recombination detection program
RDR	recombination-dependent replication
RdRP	RNA dependent RNA polymerase
REn	replication enhancer protein
Rep	replication associated protein
RF	Replicative form
RNA	ribonucleic acid
rpm	revolutions per min
SCR	satellite conserved region
SDS	sodium dodecyle sulphate
sec	second
SIR	small intergenic region
ssDNA	single stranded DNA
ssRNA	single stranded RNA
STD	Species Demarcation Tool
TAE	tris-acetate EDTA

Taq	*Thermus-aquaticus*
TGS	transcriptional gene silencing
TrAP	transcriptional activator protein
UV	ultra violet
VIGS	virus induced gene silencing
v/v	volume/volume
w/v	weight/volume
X-Gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside
µg	microgram
µl	microliter
µm	micrometer
µM	micromole

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. History

Plant viral disease can historically be divided into three phases, discussed in the coming lines. A descriptive phase during 1883-1951, is called the "Classical Discovery Period". The second phase during 1952-1983 is called "Early Molecular Era" in which new techniques were developed and advanced work at molecular level was described. Currently, in the third phase more techniques are available to evaluate virus genomes, gene functions, and plant transformation for resistance to virus diseases (Zaitlin and Palukaitis, 2000).

Diseases caused by plant viruses were initially recognized by their symptoms rather than identification of their source, the viruses. The first known example of plant viral disease, the yellow vein net disease of *Eupatrium mankinon*, is in an eighth century Janpanese poem (Bock, 1982, Hull, 2002)

Perhaps it does frost

In this village morn by O morn

For the grass, I saw in the field of summertime

Has already turned yellow

This poem rather acknowledges and appreciates the novel colouration associated with the "disease" condition. The appearance of such symptoms on leaves as curling, yellow veins during the summer time on plants was exhibited by geminivirus and its associated satellite DNA (Saunders *et al.*, 2003). Hence, infection by the geminivirus *Abutilon mosaic virus* and its associated betasatellite establishes this horticulturally desirable trait (Saunders *et al.*, 2003; Jeske *et al.*, 1977).

1.2. Classification of viruses

Viral classification is a continuous and constant work in progress done by International Committee on Taxonomy of Viruses (ICTV) with identification and addition of new viral species. ICTV organizes viruses into families (-viridae), genera (-virus), and species; viruses are also grouped based on the genomic nature (RNA or DNA), single or double stranded (ss /ds) genome and positive or negative (+/-) sense nucleic acid. Approved virus species (~2000) are divided into six orders, 87 families and 349 genera (Brown *et al.*, 2012). A minimum of 10-100 viruses for each host species of micro-organisms, plants and animals is suggested (ICTV 9th report). Upto ~47% of causal agents of all emerging plant diseases are due to viruses (Anderson *et al.,* 2004). Recently, approximately 800 plant viruses belonging to 90 genera of 20 families infect plants worldwide (Brown *et al.*, 2012).

1.2.1. RNA Viruses

RNA viruses are sub-grouped into three categories; positive sense single-stranded RNA (ss(+)RNA), negative sense single-stranded RNA (ss(-)RNA) and double stranded RNA (dsRNA) viruses. RNA viruses with ds genome include two families; *Partitiviridae* having isomertric particles either 30 or 40 nm in diameter and that conatin two or more dsRNA segments, family *Reoviridae* are classified with ten or more segmented isometric particles.

Viruses with ss(+)RNA genome includes nine families; *Bromoviridae, Closteroviridae, Comoviridae, Flexiviridae Luteoviridae, Potyviridae, Sequiviridae, Tombusviridae* and *Tymoviridae*. Family *Luteoviridae* possess icosahedral particles of 20-30nm in diameter with 32 capsomeres per nucleocapsid, family *Potyviridae* has filamentous rod shape particles with ~10kb genome, which is translated into a single polyprotein of ~70kDa.

Two families *Rhabdoviridae* and *Bunyaviridae* possess a negative sense (-ss) single stranded RNA genomes. Family *Rhabdoviridae* comprises viruses that form bullet shaped particles and exist as monopartite, while the family *Bunyaviridae* covers plant viruses with

membrane bound particles and contains tripartite ssRNA genomes made up of ambisense components.

1.2.2. DNA Viruses

Similarly, DNA viruses are classified on the basis of their genome organization either (ss) single stranded or (ds) double stranded DNA as described below.

1.2.2.1. Family *Caulimoviridae*

This family of viruses are termed pararetroviruses (replicate by reverse transcription). There are 53 species in the family *Caulimoviridae,* which are divided into eight genera such as *Rosadnavirus* (*Rose yellow vein virus*), *Cavemovirus* (*Cassava vein mosaic virus*), *Petuvirus* (*Petunia vein clearing virus*), *Caulimovirus* (*Cauliflower mosaic virus*), *Soymovirus* (*Soybean chlorotic mottle virus*), *Badnavirus* (*Commelina yellow mottle virus*), *Solendovirus* (*Sweet potato vein clearing virus*) and *Tungrovirus* (*Rice tungro bacilliform virus*). Viruses in the family *Caulimoviridae* are mostly monopartite, circular dsDNA genome that range from 6000bp to 8000bp in size, transmitted either mechanically or through aphids. Depending on the virus, DNA can contain one ORF (as in petuviruses), up to eight ORFs (as in Soymoviruses). Virions are non-capsid which may be in bacilliform or isometric structure.

1.2.2.2. Family *Nanoviridae*

This family of plants viruses are composed of ten species, which are divided into two genera, *Nanovirus* and *Babuvirus.* The genome of each genus is composed of multiple segments, each of which is approximately 1kb in size with a single open reading frame (ORF). These viruses are 20 nm in diameter with icosahedral particles that encapsidate small circular ssDNA. The type species of the genus *Babuvirus* is *Banana bunchy top virus,* with six segments that are transcribed unidirectionally. The other genus is *Nanovirus,* which infects dicotyledonous plants and includes species *Faba bean necrotic yellow vein virus, Milk vetch*

dwarf virus, and *Subterranean clover stunt virus*, resemble animal infecting circoviruses in genome size and genome architecture but differs from them in virion size and morphology. Viruses in this family are transmitted by aphids.

1.3. Family *Geminiviridae*

Geminiviridae constitutes the second largest family of plant viruses with ssDNA genome, which infect both monocots and dicots, throughout the world, in the tropical and sub-tropical regions. Recently, they have spread into temperate zones because of alteration in agricultural practices and variation in environmental conditions as well as globalization in the trade of agricultural products led to the dissemination of insect vectors more rapidly (Fauquet *et al*., 2003). Viruses in this family produce some specific symptoms like yellow mosaic, yellow mottle, leaf curling, stunting, and streaks. Since 19th century, such symptoms have been observed in plants (Wege *et al*., 2000). It was in the 1970s when ssDNA viruses were found with these symptoms (Harrison *et al.,* 1977). This group of viruses was coined as geminivirus by Goodman in 1977 (Goodman, 1977a, 1997b), later in 1980s they were classified into a new plant virus family *Geminiviridae* (Rybicki, 1994). Virions in this family are twinned (geminate) quasi-icosahedral capsids. The capsid contains nucleic acid as either one or two small circular ssDNA components, each of which is 2.5 - 3.0 kb (approximately 30 x 20 nm in size) (Krupovic *et al*., 2009; Bottcher *et al.,* 2004; Zhang *et al*., 2001; Harrison, 1985) (Figure. 1.1). Geminiviruses replicate in the nuclei of the infected host plant cells through an intermediate replicative form (RF) dsDNA molecule. The RF is used as a template for transcription and replication, both of which depend upon the host DNA transcription and replication machinery (Jeske, 2007). Transmission of these viruses by insect vectors occurs in a persistent, circulative manner (Arguello-Artorga *et al.*, 1994; Sunter *et al.*, 1994; Harrison and Robinson, 1999, Varma and Malathi, 2003).

1.4. Geminiviruses

Among plant infecting virus families, the family *Geminiviridae* causes the most economic losses. These viruses infect a broad range of plants worldwide including both, monocots and dicots (Gutierrez, 2000). The economic losses in Pakistan from 1992-1997 due to cotton geminiviruses are estimated to be US $5.0 billion (Briddon and Markham, 2001). Correspondingly, in the last two decades, losses due to cassava geminiviruses in Africa were estimated to be US $ 1.2-2.3 billion (Thresh and Cooter, 2005). Also in India, geminiviruses caused US $ 300 million losses in grain legumes (Varma and Malathi, 2003) and US $ 140 million in tomato crop alone in Florida (Moffat, 1999).

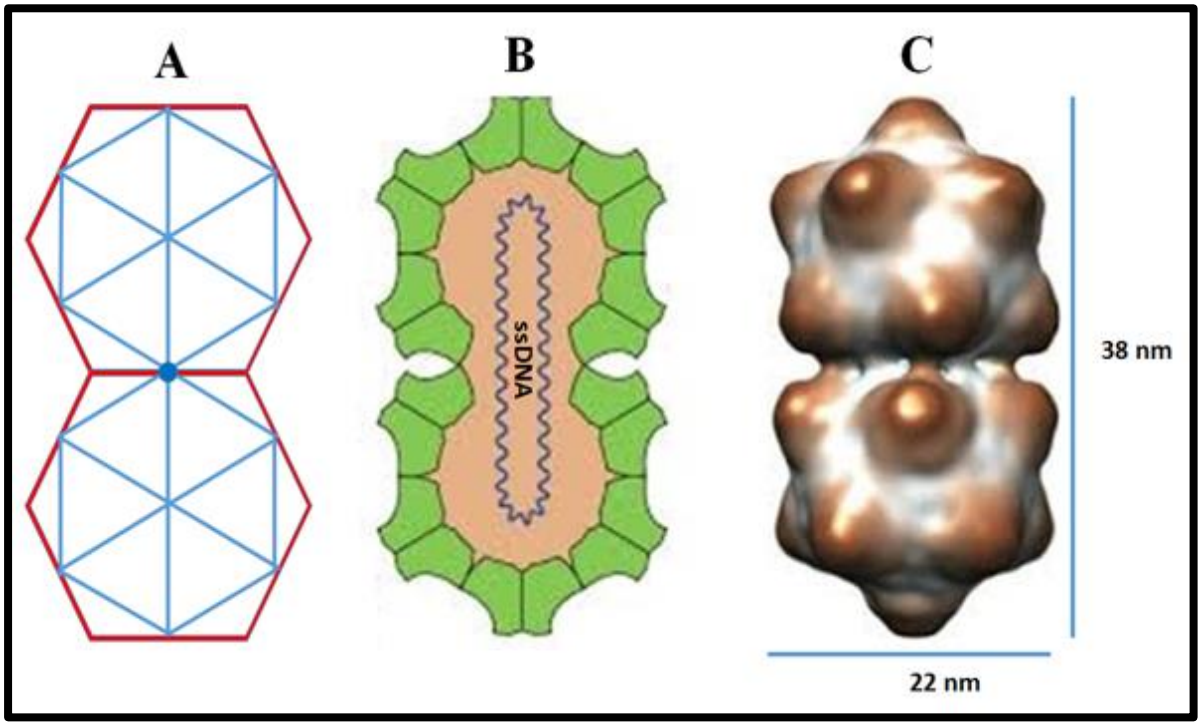


Figure 1.1. Structure of typical Geminivirus. (A) Geometrical representation of icosahedral structure; (B) Graphical representation of geminate structure; (C) Three-dimensional reconstruction of a geminate particles from cryo-electron microscopy data. Image C was reproduced from Shepherd *et al*. (2010).

1.4.1. Geminivirus Taxonomy and Nomenclature

Certain rules that have been established, followed to name/classify new viruses into species, stains and variants (Van Regenmortal, 2006). The number of geminiviruses is growing

rapidly with every passing day. *Geminiviridae* is the second largest family among plant viruses. This complicates taxonomy and nomenclature of geminiviruses. Although, the taxonomic list is regularly and efficiently updated (Brown *et al*., 2012; Fauquet and Stanley, 2005; Fauquet *et al*., 2003), there is a need for a rationale and all-inclusive way to name and classify the newly identified geminiviruses. Therefore, to make the taxonomic standards clearer and nomenclature guidelines more transparent, the ICTV has proposed a comprehensive demarcation criteria (Brown *et al.,* 2012) for classifting and naming geminiviruses. Rapid increases in deposition of new viruses to the family *Geminiviridae* (specifically genus begomovirus), due to the recent innovation in cloning and low-cost sequencing, various constrains were elevated in the previously recommended scenario of species demarcation. To encompass these matters, previous guidelines were revised for genus begomoviruses and established a new revised one: pairwise sequence comparison was done using Sequence Demarcation Tool (SDT) (Muhire *et al.*, 2014) with MUSCLE (Edgar, 2004) alignment option and phylogenetic tree was predicted using PHYML3.0 method implemented in MEGA 5.2 (Tamura *et al.,* 2011). The following guidelines were setup for new virus:

- An isolate having <91% sequence identity (with available full genome or DNA A component in the databases) it would be considered as a new species
- An isolate having ≤94% sequence identity to isolates of two or more in the databases, it would be considered as a new strain of the same species
- An isolate having ≥94%, sequence identity to one or few isolates from a particular strain; it would be a variant of the same species (Brown *et al*., 2012)

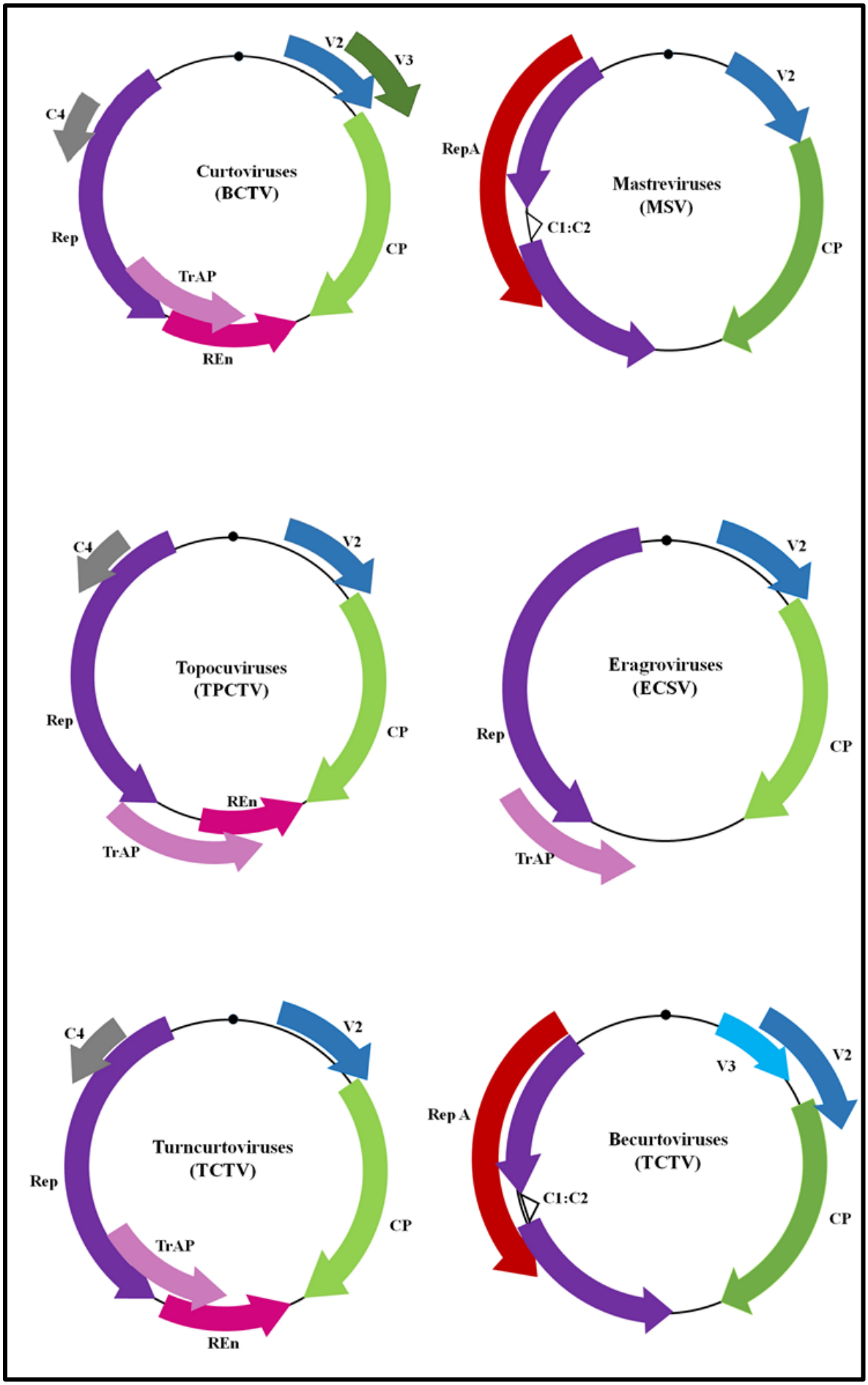
To name a new isolate the following recommended nomenclature structure is usually followed:

Species name, strain descriptor (symptoms, host, location, and/ or a letter A, B, C etc)

[Variant descriptor (Country- location-[host]-year)**]**

To date, more than 300 official geminivirus species belonging to the genera: *Mastrevirus, Curtovirus, Topocuvirus, Begomovirus, Becurtovirus, Turncurtovirus, Eragrovirus, Grablovirus* and *Capulavirus* have been reported (Varsani *et al*., 2017; Zerbini *et al*., 2017; Brown *et al.*, 2015; Roumagnac *et al*., 2015; Varsani *et al*., 2014; Adams *et al*., 2013; Haydarnejad *et al.,* 2013; Brown *et al*., 2012; Briddon *et al*., 2010). The genus *Begomovirus* is comprised of at least 283 species, representing the largest and most important plant infecting viruses (Brown *et al*., 2015; 2012). Those that have a monopartite genome component infecting both monocotyledonous and/or dicotyledonous plants are members of genus *Mastrevirus,* transmitted by leafhoppers, of which *Maize streak virus* (MSV) is the type species (Willment *et al.*, 2007). Like Mastreviruses, the genus *Curtovirus* also have a monopartite genome and are transmitted by leafhoppers to dicotyledonous plants only and *Beet curly top virus* (BCTV) is the type species (Baliji *et al*., 2004). *Tomato pseudo curly top virus* (ToPCTV) with a monopartite genome is the only member of genus *Topocuvirus* and is transmitted by a treehopper to dicotyledonous plants (Briddon *et al*., 1996). Genus *Begomovirus* are whitefly (*Bemisia tabaci*) transmitted geminiviruses that infect dicotyledonous plants (Stanley *et al*., 2005). Begomoviruses have either a monoparite genome (DNA A) or a bipartite genome (DNA A and DNA B) of approximately equal size. *Bean golden mosaic virus* (BGMV) is the type species of Begomoviruses (Brown *et al*., 2012; Seal *et al*., 2006). Genus *Bercurtovirus* have a monopartite genome and are transmitted by leafhopper (*Circulifer haematoceps*) to dicotyledonous plants and *Beet curl top Iranian virus* (BCTIV) is the type species. Genera *Eragrovirus* and *Turncurtovirus* have a monopartite genome and have single species: *Eragrostis curvula streak virus* (ECSV) and *Turnip curly top virus* (TCTV), and are transmitted by leafhoppers to monocots and dicots plants in case of *Turnip curly top virus,* respectively (Varsani *et al*., 2014). The last two genera of the family *Geminiviridae* are *Capulavirus* and

Grablovirus, characterized by a single component molecule and are transmitted by aphid vectors (*Aphis craccivora*) and alfalfa treehoppers (*Spissistilus festinus* Say); the type species are *Alfalfa leaf curl virus* (ALCV) and *Grapevine red blotch virus* (GRBV), respectively (Varsani *et al.*, 2017; Brown *et al.*, 2015; Roumagnac *et al*., 2015) (Figure 1.2).


V2
V3
C4
Curtoviruses
(BCTV)
Rep
CP
TrAP
REn
RepA
V2
Mastreviruses
(MSV)
C1:C2
CP
C4
V2
Topocuviruses
(TPCTV)
Rep
CP
REn
TrAP
V2
Eragroviruses
(ECSV)
Rep
CP
TrAP
C4
V2
Turncurtoviruses
(TCTV)
Rep
CP
TrAP
REn
Rep A
V3
V2
Becurtoviruses
(TCTV)
C1:C2
CP

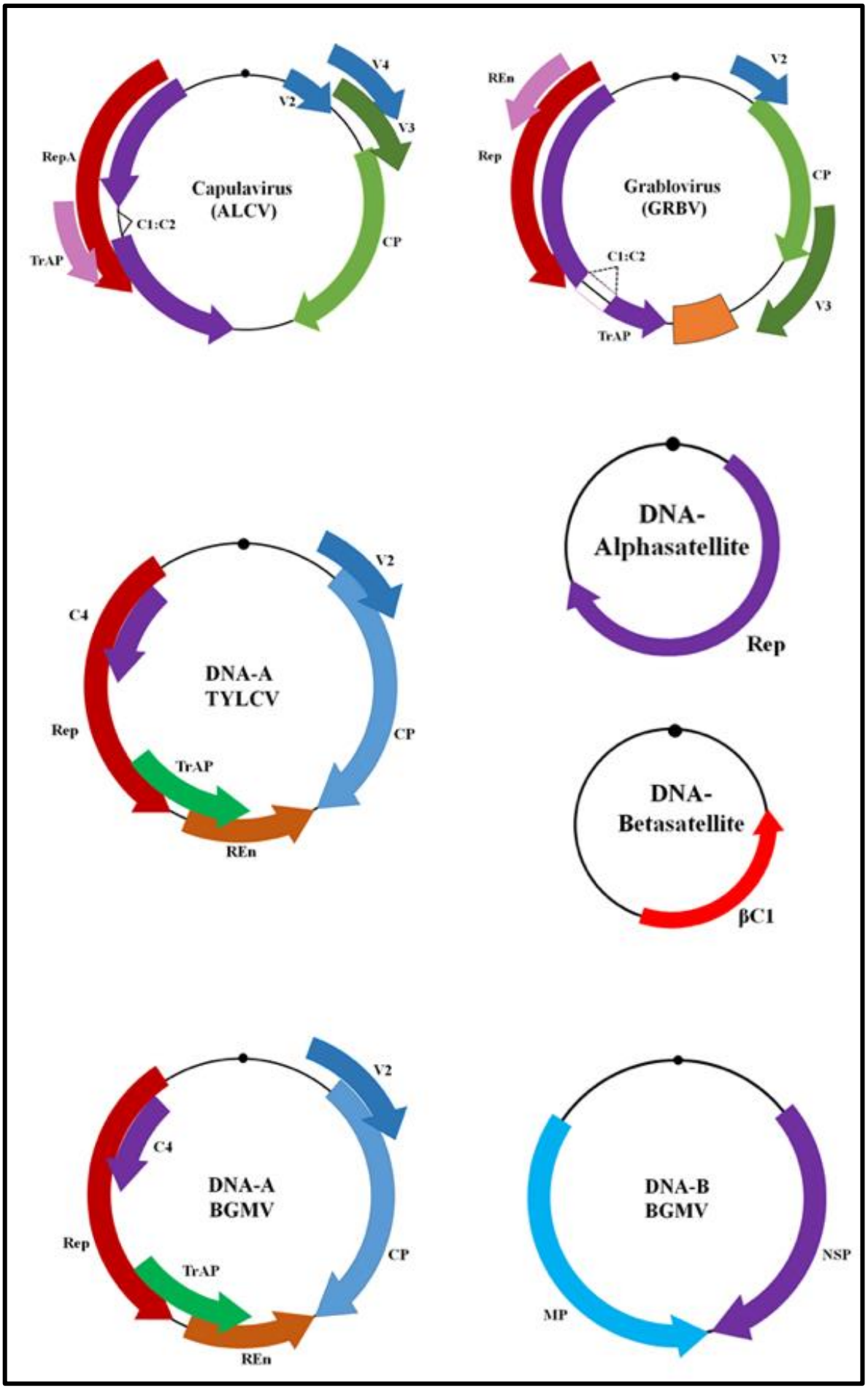

V4
V2
V3
RepA
Capulavirus
(ALCV)
C1:C2
CP
TrAP
REn
V2
Rep
CP
Grablovirus
(GRBV)
C1:C2
V3
TrAP
DNA-
Alphasatellite
Rep
V2
C4
DNA-A
TYLCV
CP
Rep
TrAP
REn
DNA-
Betasatellite
βC1
V2
C4
DNA-A
BGMV
CP
Rep
TrAP
REn
DNA-B
BGMV
NSP
MP

Figure 1.2. Genome organization of the nine genera of the family *Geminiviridae*. Each genus is presented by their type species: *Mastrevirus* (MSV), *Curtovirus* (BCTV), *Topocuvirus* (TPCTV), *Begomovirus* ([monopartite] TYLCV and [bipartite] BGMV), *Becurtovirus* (BCTIV), *Turncurtovirus* (TCTV), *Eragrovirus* (ECSV), *Grablovirus* (GRBV) and *Capulavirus* (ALCV). *Mastrevirus, Curtovirus, Topocuvirus, Becurtovirus, Turncurtovirus, Eragrovirus, Grablovirus* and *Capulavirus* have monopartite genomes, whereas majority of begomoviruses originating from old world (OW) have monopartite genomes and few have bipartite genomes. All begomoviruses originating from the new world (NW) possess bipartite genomes. The bipartite genomes of begomoviruses consist of two genomic components designated as DNA A and DNA B. Bipartite begomoviruses in the NW lack the V2 ORF on the virion-sense strand of DNA A. The stem-loop structure containing the nona-nucleotide sequence is represented as a black circle. The genes are: coat protein (CP), pre-coat protein (V2), replication associated protein (Rep), replication enhancer protein (REn), transactivator protein (TrAP), movement (MP), and nuclear shuttle protein (NSP).

1.4.2. Replication of *Geminivirus* Genome

Geminiviruses replicate their ssDNA genomes via dsDNA Replicative Form (RF) intermediates by rolling-circle mechanism and/or a recombination-dependent mechanism in the nucleus of the host cells in which the viral protein Rep plays a vital role (Gutierrez, 1999; Hanley-Bowdoin *et al*., 1999; Jeske *et al.,* 2001; Preiss and Jeske., 2003). Rolling circle replication is a type of nucleic acid replication to synthesize multiple copies of circular nucleic acid molecules, rapidly (Jeske *et al.,* 2001).

The infection cycle starts with insect vector feeds on an infected host and transfers viral particles to another plant where the virus uncoats and the ssDNA is transported to the nucleus where transcription and replication occur. Transcription is a bidirectional process from promotor sequences in the intergenic region (IR). This is a complex phenomenon giving rise to several overlapping mRNAs. These transcripts overlap in several nucleotides at their 3' ends (Hanley-Bowdoin *et al*., 1989; Hanley-Bowdoin *et al.,* 1999; Shivaprasad *et al*., 2005; Sunter and Bisaro, 1989). Two conserved proteins on complementary strands of the Rep and REn, along with some known cellular host factors such as RF-C, PCNA, RPA, RAD54, SCE1 and DNA polymerases *etc* (Yadava *et al.,* 2010) are implicated in virus replication. Rep is a

multitasking protein that has role in initiation of rolling circle replication (RCR) (Elmer *et al.*, 1988; Hanley-Bowdoin *et al.,* 1990). Once inside the host nucleus, a short RNA primer attaches to the genomic ssDNA to initiate the synthesis of complementary sense DNA. This intern produces dsDNA RF intermediates (Donson *et al*., 1984; Saunders *et al.,* 1992). The precise initiation and termination depend upon the multitasking Rep protein. The dsDNA RF is used as a template for RCR (Figure 1.3). A similar mechanism of RCR is also exhibited by some ssDNA encoding bacteriophages and other prokaryotes (Bisaro, 1996; Noris *et al*., 1996; Saunders *et al.,* 1991; Stenger *et al*., 1991; Timmermans *et al.,* 1994).

The iterons are repeated sequences located in the IR of geminivirus, used as recognition sites for binding of Rep-protein. Any alteration in the sequence of iterons due to mutation, affect the Rep binding ability (Arguello-Astorga and Ruiz-Medrano, 2001). Binding of Rep protein to the iteron sites is followed by the action of Rep protein, to nick the plus strand of the dsDNA at a specific site called double strand origin (DSO), usually in the loop sequence between position 7 and 8 of the nona-nucleotide motif (TAATATT↓AC) of virion sense strand, attachment of Rep to the exposed 5'-terminus and elongation of the 3'-terminus using the complementary-sense template (Laufs *et al.,* 1995; Kong *et al.,* 2000; Luque *et al*., 2002; Selth *et al*., 2005).

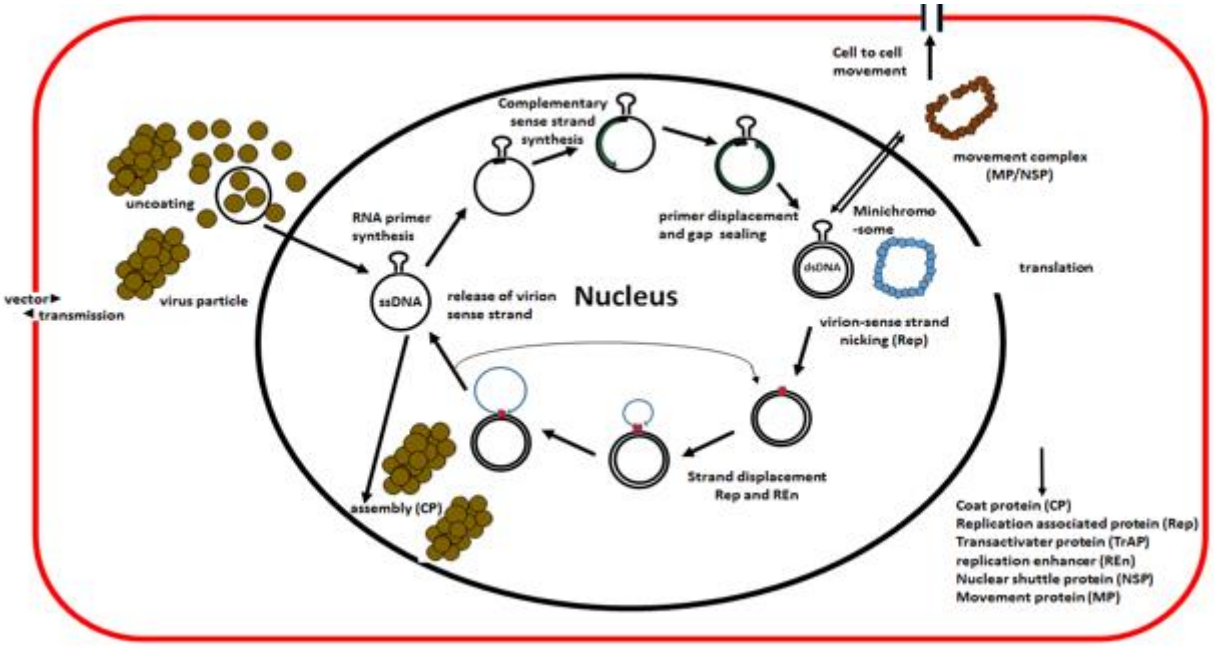


Figure 1.3. Schematic representation of replication strategy of geminiviruses. Virus particles are introduced into the host cell during insect feeding. Uncoated ssDNA is converted into dsDNA intermediate-known as replicative form (RF) by using host factors in the nucleus. ssDNA is synthesized from the dsDNA template by a rolling circle replication (RCR) involving Rep and REn viral protein in association with host factors, and is assembled into virus particles that accumulate in the nucleus. dsDNA also serves as transcriptionally active minichromosomal form (shown here). The nascent viral ssDNA may re-enter the replication cycle or become encapsidated. Intracellular and cell-to-cell transportation of both dsDNA and ssDNA is mediated by V2 and CP in the case of monopartite begomoviruses and by NSP and MP for bipartite begomoviruses (Bull, 2004; Stanley *et al*., 2005)

The second and last stage of the RCR is the synthesis of ssDNA from the dsDNA RF. The ssDNA is then assembled into virus particles that accumulate in the nucleus of the host cell. This stage starts after a nicking in the nona-nucleotide which serves as the origin of replication. Multifunctional Rep protein attaches to the 5'-terminus of the nicked DNA, whereas, during the replication, the 3' end of the cleaved strand serves as a primer. DNA polymerase moves around the template and synthesizes the virion strand. The replicated unit length virion strand is then separated from the replication machinery and ligates as a circular ssDNA also mediated by Rep protein. The resulting newly synthesized ssDNA can either re-

enter into the DNA replication pool to serve as template for replication or can associate with the CP, or can be transported outside of the nucleus and to the neighboring cell, through plasmodesmata (Jeske *et al.,* 2001). In case of monopartite begomovirus, the cellular and systemic movement is entirely dependent on the encoded CP and V2 proteins (Sharma and Ikegami, 2009), while NSP and MP of DNA B in case of bipartite begomovirus performed the transportation activity (Sanderfoot and Lazarowitz, 1996; Sanderfoot *et al.,* 1996; Lazarowitz, 1999; Lazarowitz and Beachy, 1999; Gafni and Epel, 2002).

Recombination-dependent replication (RDR), in addition to RCR (Jeske *et al*., 2001; Alberter *et al.,* 2005) is alternate strategy where the host factors in combination with the Rep protein or even alone are necessary to produce DNA intermediates for replication. Like begomovirus (AbMV) and some well-known viruses also produce viral DNA intermediates similar to RDR, as analogous to T4 phages (Mosing *et al*., 2001; Mosig, 1998; Kornberg and Baker, 1992). The presence of both RCR and RDR replicative forms suggest this multitasking activities different replication strategies to overwhelmed host defenses (Jeske *et al.,* 2001; Preiss and Jeske, 2003). The newly synthesized ssDNAs can then face three fates. Either they re-enter the DNA replication pool for further copies, encapsidate (coating) to make virions, or be transported outside the nucleus to the neighboring cell, through plasmodesmata, utilizing viral MPs (Stenger *et al*., 1991; Heyraud *et al*., 1993; Stanley, 1995).

1.5. Genus *Begomovirus*

Begomovirus is the largest and most economical important genus of the family *Geminiviridae.* The name begomovirus is derived from Bean golden mosaic virus (the virus now reclassified as *Bean golden mosaic yellow virus*) (van Regenmortel *et al.,* 1997). Begomoviruses have much bigger vistas, engulfing most economically important crops in the tropical and sub-tropical regions of the world. Virus species in this genus are solely transmitted by a species complex of whitefly *Bemisia tabaci* (Gennadius) in a persistent and circulative

manner. Due to high level of host species for whitefly, begomoviruses become major destructive to the agricultural output of many warmer areas of the world including Pakistan (Brown *et al*., 1995). Previous studies highlighting that begomoviruses infect dicotyledonous plants only (Navas-Castillo *et al.*, 2011) including cassava mosaic, cotton and tomato leaf curl, and bean golden mosaic. According to estimates losses estimated up to several billion dollars per year (Harrison and Robinson, 2002; Varma and Malathi, 2003). From the last three decades, begomoviruses are emerged as major threat in food, fiber, and ornamental crops worldwide, including bean, cassava, cotton, cucurbits, squash, tobacco, tomato and many more (Brown,1994; Polson and Anderson, 1997; Rybicki and Pieterson. 1999; Briddon and Markham, 2000; Moriones and Navas-Castillo, 2000; Monsalve-Fonnegra *et al*., 2001; Otim-Nape *et al*., 2001, Morales and Anderson, 2001; Ribeiro *et al*., 2003; Varma and Malathi, 20003, Fargette *et al*, 20006, Mansoor *et al*., 2006).

1.5.1. Genome Organization

Genus *Begomovirus* of the family *Geminiviridae* is the largest group of plant infecting viruses, which can be classified into two groups, based on the phylogenetic analysis. The Old World (OW) viruses (those originating from Africa, Asia, Australia and Europe) and the New World (NW) viruses (those originating from the Americas) (Padidam *et al.*, 1999; Paximadis *et al.*, 1999; Raybicki, 1994). Furthermore, they are divided into two categories (monopartite or bipartite), based on their genomic architecture consist of either one or two circular ssDNA molcules (Figure 1.4).

1.5.1.1. Bipartite Begomoviruses

NW begomoviruses typically have genomes consisting of two separately encapsidated genomic components, referred to DNA-A and DNA-B, covalently closed, circular ssDNA molecules of about 2.5 ~3.0 kb in size (Brown *et al.*, 2012; Stanley *et al.*, 2005). These

components share no sequence identity with each other except ~200 nucleotides fragment of intergenic region (IR), known as common region (CR; Hanley-Bowdoin *et al.*, 1999). The common region of both molecules consists of a hairpin stem loop structure with a highly conserved sequence called nona-nucleotide (TAATATT/AC), serves as origin of replication.

DNA-A component of bipartite begomovirus contains 5-6 overlapping open reading frames (ORFs) on virion-sense and complementary-sense strand. Two ORFs encoding for V2 protein (V2) and coat protein (CP) is on virion-sense strand, while four ORFs encoding for Replication associated protein (Rep), replication enhancer protein (REn), a transcriptional activator protein (TrAP) and C4 protein (C4) are located on the complementary-sense strand. The DNA-B component contains two OFRs, encoding for nuclear shuttle protein (NSP) and movement protein (MP) on sense and complementary strand. Bipartite begomoviruses reported from the New World (NW) can be distinguish from Old World (OW) viruses, lack V2 gene on the virion-sense strand of DNA-A component (Brown *et al*., 2012).

DNA-A component expresses viral proteins required for DNA replication, encapsidation and other factors to control both plant and viral gene expression in the host cell (Townsend *et al.*, 1986; Etessami, *et al*., 1989) while DNA-B encodes proteins required for transportation of viral proteins in an inter- and intra-cellular movement and symptom development (Von Arnim and Stanley, 1992; Sanderfoot and Lazarowitz, 1996; Bisaro, 1996). In several OW bipartite begomoviruses, DNA alone is efficient and sufficient for systemic infection and movement, wheras, in NW begomoviruses, both genomic components are essential for systemic infection (Rojas *et al.*, 2005). The known function of begomovirus encoded proteins are summarized in Table 1.1.

1.5.1.2. Monopartite Begomoviruses

Majority of begomoviruses reported from the OW are monopartite, lack a DNA-B component, except a few are bipartite. Monopartite begomovirus has a single genomic component homologous to the DNA-A (DNA-A like) molecule of bipartite begomoviruses. Like bipartite begomovirus DNA-A molecule, monopartite genome encodes all required proteins essential for viral DNA replication, transcription and encapsidation (Navot *et al.*, 1991; Rojas *et al.*, 1998). Pre-coat protein (V2) and coat protein (CP) of the monopartite begomoviruses performed the same transportation activity, reported in DNA-B of bipartite begomoviruses. In addition, the genomes of some known monopartite begomoviruses alone have the potential to induce wild type disease symptoms. Best examples of them are *Tomato leaf curl virus* (TLCV), *Tomato yellow leaf curl virus* (TYLCV) and *Tomato yellow leaf curl Sardina virus* (TYLCSV) (Navot *et al*., 1991; Dry *et al*., 1993; Bananej *et al.*, 2004; Scholthof *et al.*, 2011).

1.5.1.2.1. Begomovirus-DNA satellite complex

Satellite molecules/ components are small nucleic acid components which may be classically define as "virus or nucleic acid that depend on their helper virus for replication, are dispensable for the replication of the helper virus, and lack sequence similarity to the helper virus genome" (Murant and May, 1982). The concept of satellite molecule associate is always highlighting RNA viruses, as a common feature of them. However, the circle of satellites is not strictly bound to RNA viruses only, DNA viruses also found with its association. Recently, majority of monopartite begomoviruses have been found to also be associated with DNA satellites, known as beta-satellite and alpha-satellite (Briddon and Stanley, 2006). Most of these satellites interfere in helper virus replication and attenuate symptoms. On the contrary, a small number of satellite molecules, are known that exacerbate symptomology or produce novel

disease symptoms which otherwise are not attributed to the helper virus (Collmer and Howell, 1992; Roossinck *et al*., 1992). The complex of virus- satellite mostly prevails in the OW including Africa and Asia, but recently association of alphasatellite, in the absence of betasatellite, with bipartite begomovirus from the NW also reported (Paprotka *et al*., 2010; Romay *et al*., 2010)

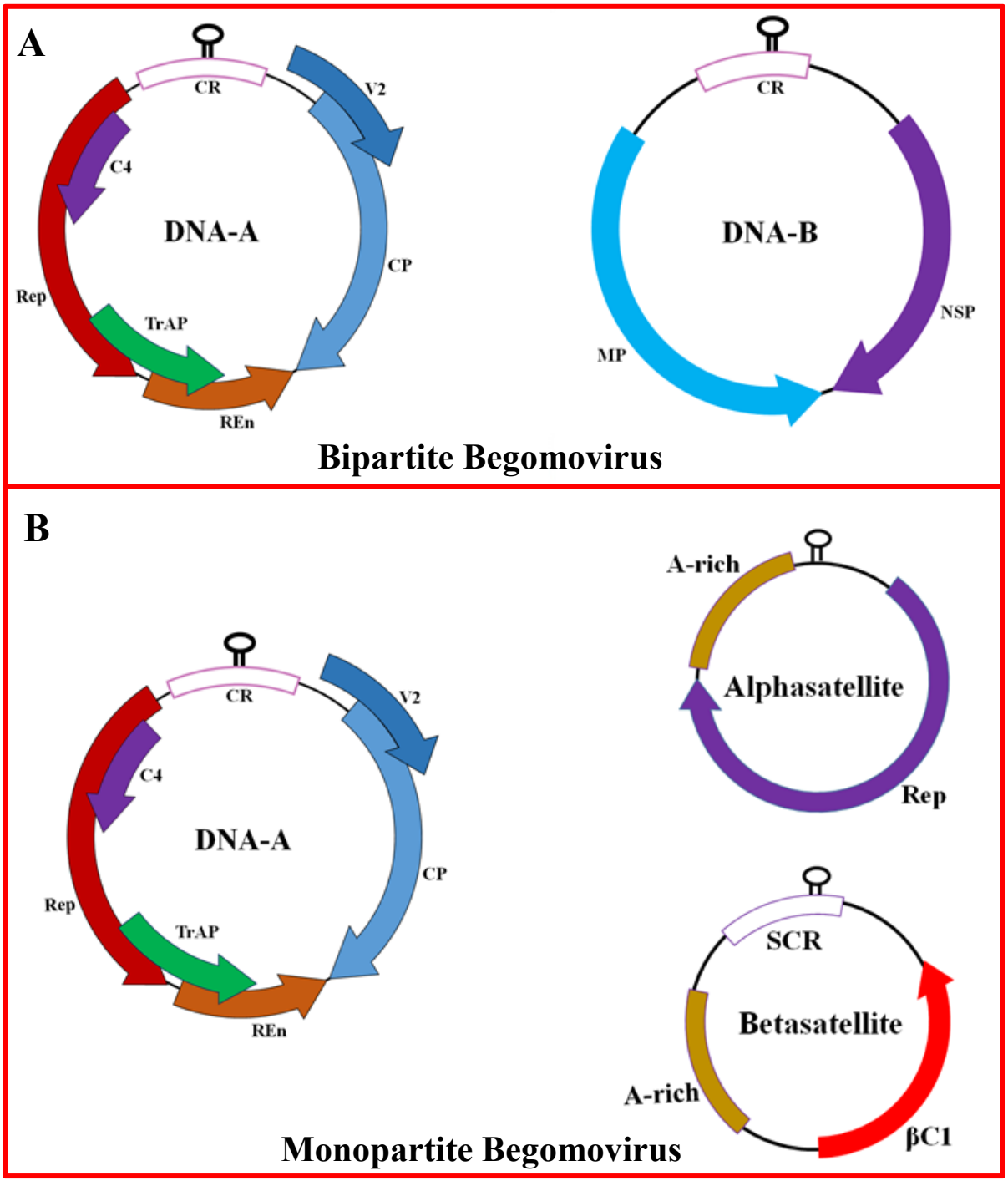


Figure 1.4. Genome organizations of bipartite and monopartite begomoviruses and their associated DNA alpha- and beta-satellite.

1.5.1.2.1.1. DNA Betasatellite

Cotton plants were agro-inoculated with Cotton leaf curl viruses (CLCV) to verify Koch's postulates. Plant did not reproduce symptoms which were observed in the field, like severe curling of leaf, dark and thick veins, and in some case a small leaflet-known as enation on the underside of leaves. This indicates missing of some genomic component which serve as a causal agent during the whole process of inoculation (Briddon *et al.,* 2001). Similar, reports were also documented in case of *Ageratum conyzoides* by re-inoculation using *Ageratum yellow vein virus* (AYVV) to reproduced yellow vein symptoms, suggested the lack of the other factor in association for symptoms development (Saunders *et al*., 2000). Several molecules were characterized from the infected plants, containing similarity in the origin of replication with begomoviruses. Initially, all these were categorized as defective molecules but when co-inoculated with begomoviruses, typical symptoms like thick veins, leaf curling and enations were appeared on the plants (Saunders *et al.,* 2000; Briddon *et al.,* 2001; Jose and Usha, 2003; Zhou *et al*., 2003).

Initially, this novel molecule was named as DNA-β and later termed as betasatellite. The first ever begomovirus associated DNA satellites to be discovered, was Tomato leaf curl satellite (referred as ToLCV-sat) in a tomato plant associated with monopartite begomovirus in Australia (Dry *et al*., 1997). This small satellite circular molecule was about 682 nucleotides in length with no open reading frames, but, had some sequence identity to its helper virus in the stem-loop structure containing the conserved geminivirus TAATATTAC motif and Rep binding motif (iteron) (Behjatnia *et al.,* 1998). ToLCV-sat showed no apparent effect on the symptoms induction by the helper virus but it was found dependent on the helper begomovirus, atleast for its own replication, and encapsidation.

Betasatellites are ssDNA molecules of approximately half the size of their helper components (~1350 bp) with lack of sequence similarity among the two, except the presence of putative stem-loop structure containing the nona-nucleotide motif. All betasatellite contains some basic features; an adenine rich region (A-rich), satellite conserved region (~100 nt), and a single ORF (*βC1* gene) in the complementary orientation of the molecule (Briddon *et al*., 2003; Jose and Usha 2003; Zhou *et al*., 2003; Bull *et al*., 2004; Cui *et al.,* 2004). The βC1 gene is encode for multi-function protein of approximately 118AA, involved in determination of the host range, movement and accumulation of both satellite and helper virus (Zhou *et al*., 2003). Protein encoded by betasatellite has multiple activities: pathogenicity enhancement (Zhou, 2013), suppressor of post transcriptional gene silencing (PTGS) to overwhelmed host defense mechanism, potential involvement in virus drive in *planta* (Saeed *et al*., 2007), upregulation of viral DNA levels *in planta* (Briddon *et al*., 2001;Saunders *et al*., 2000), capability of binding of RNA/DNA (Cui *et al*., 2005) as well as interacting with a range of host derived factors (Eini *et al.,* 2009). In addition to the multiple functions, betasatellite encoded protein (βC1) also engaged in a variety of complexes (Cheng *et al*., 2011).

Along with this, some small satellites were also found which contain SCR and A-rich regions but lack either all or part of βC1 gene which is essential for symptom determination, were presumed as defective betasatellites (Briddon *et al.,* 2003, Zhou *et al*., 2003, Saunders *et al*., 2004; Cui *et al*., 2005).

1.5.1.2.1.2. Satellite-like alphasatellite

Like betasatellites, alphasatellites are circular ssDNA molecules usually in association with monopartite begomoviruses (Mansoor *et al.,* 1999, Briddon *et al.,* 2004, Stanley, 2004) (Figure. 1.4). Alphasatellites, formerly called DNA 1 are of approximately 1400 bp in length, were assumed to be evolved from Nano viruses-another ssDNA plant infecting viruses, which possess some basic feature like; A-rich region, a stem loop structure containing a nona-

nucleotide sequence (TAGTATTAC) and a single gene known as 'Alpha-Rep' on the sense strand encoding for replication initiator protein of about 315 amino acid (Briddon *et al*., 2004; Kumar *et al*., 2014; Shahid *et al*., 2014). These molecules are competent of self-replication but may require a helper for insect transmission and movement (Fiallo-Olivé *et al*., 2012; Rehman *et al*., 2012). Previously, alphasatellites were found associated with OW begomoviruses. But recently few alphasatellites have also been found associated with the NW begomovirus (Rojas *et al*., 2005, Paprotka *et al.,* 2010).

Table 1.1. Begomovirus encoded proteins and their known functions.

Gene	Protein	Functions	References
(A)V2	V2	Symptoms determination, TGS suppression, movement and involve in CP expression.	Bull *et al.*, 2007; Padidam *et al.*, 1996; Rigden *et al*., 1993; Rojas *et al.*, 2001; Wang *et al.*, 2013
(A)V1/CP	CP	White-fly mediated transmission, encapsidation, accumulation of viral ssDNA, nuclear localization.	Guerra-Peraza *et al.*, 2005; Rojas *et al*., 2001; Kunik *et al*., 1998;Boulton *et al*., 1991; Briddon *et al*., 1990
(A)C1/Rep	Rep	Essential for replication, DNA-binding, nicking-ligation, oligomerization, initiation of virion-strand DNA replication, control of cell cycle, involve in ATPase activity, has site specific topoisomerase activity, induces the expression of a host DNA synthesis protein, proliferating cell nuclear antigen (PCNA), in non-dividing plant cells, binds with retinoblastoma related protein (pRBR).	Arguello-Astorga *et al*., 2007; Campos-Olivas *et al*., 2002; Castillo *et al*., 2003; Choudhury *et al*., 2006; Collin *et al*., 1996; Desbiez *et al.*, 1995; Egelkrout *et al*., 2002; Heyraud-Nitschke *et al*., 1995; Kong and Hanley-Bowdoin, 2002; Kong *et al*., 2000; Luque *et al*., 2002;Orozco and Hanley-Bowdoin, 1998; Orozco *et al.*, 1996; Raghavan *et al*., 2004.
(A)C2/TrAP	TrAP	Transcriptional activator for virion sense genes, overcomes other host defense mechanism including programmed cell death (PCD), suppressors of RNA silencing, interact with and inactivates SNF1 kinase and adenosine kinase (ADK).	Hao *et al*., 2003; Baulcombe, 2004; Ding *et al*., 2004; Roth *et al*., 2004; Voinnet, 2005; Vanitharani *et al*., 2004; van Vezel *et al*., 2002a; Wang *et al*., 2003, 2005; Buchmann *et al.*, 2009; Moffatt *et al*.,

			2002, Yang *et al*., 2007; Hussain *et al*., 2007; Mubin *et al*., 2010.
(A)C3/REn	REn	Enhance viral DNA replication	Castillo *et al*., 2003; Settlage *et al*., 1996, 2001, 2005; Sunter *et al*., 1990; Wang *et al*., 2005.
(A)C4	C4	Involved in symptoms development, movement, suppressors of PTGS.	Rigden *et al*., 1994; Jupin *et al*., 1994; Rojas *et al*., 2001;2005; Etessami *et al*., 1989; Elmer *et al*., 1988, Vanitharani *et al*., 2004
(B)V1/NSP	NSP	Transport of viral DNA between the nucleus and cytoplasm, host range determination and virus transmission, interacts with host factors, pathogenicity determinant.	Pascal *et al*., 1993; 1994; Sanderfoot *et al*., 1996; Brough *et al*., 1988; Etessami *et al*., 1988; Mariano *et al*., 2004; McGarry *et al*., 2003, Garrido-Ramirez *et al*., 2000, Hussain *et al*., 2007.
(B)C1/MP	MP	Involved the cell-to cell movement and pathogenicity determinant.	Noueiry *et al*., 1994; Rojas *et al*., 1998; Sudarshana *et al*., 1998; Gilbertson *et al*., 2003; Hussain *et al*., 2007, Frischmuth *et al*., 2007; Jeske, 2009.

1.6. Importance of cotton

Cotton has long been known to as a major source of food for humans, feed for animals/livestock and fiber for industry across the world. The use of cotton fibers to make fabrics has long been known for at least 7000 years (Cantrell, 2005; Sunilkumar *et al*., 2006). Throughout the world, ~32.6 million hectares (mha) area is covered by cotton with a production estimated at 27.6 million tons for the year 2011/2012 (Anonymous, 2011). The genus *Gossypium* has ~51 species. Of the cultivated species, *Gossypium arboreum* L. and *Gossypium herbaceum* L. have diploid, whilst *Gossypium hirsutum* L. and *Gossypium barbadense* L. have allotetraploid genomes (Wendel and Cronn, 2003). *G. hirsutum* (also known as upland cotton)

and *G. barbadense* (also known as Pima/Egyptian cotton), have their centre of origin in Central America and northern Peru, respectively. *G. hirsutum* is the most widely cultivated species likely due to its quality fiber, with a thick secondary wall, giving it a strength. *G. arboreum* and *G. herbaceum*, on the otherhand, are native to the OW. The fiber they produce however is not much valuable. Contrary to the upland cotton, *G. arboreum* is adapted to hot dry environments and is immune to CLCuD, and is thus still considered as important commercial species.

Cotton is an important cash crop for Pakistan known as "white gold". Pakistan is the fifth largest producer of cotton, fourth largest consumer of cotton, third largest exporter of raw cotton and leading exporter of yam in the world. Cotton is planting on major cultivated portion of cultivated land in Pakistan, covering 15 % of cultivable area which engulf around 3 million hectares with production estimated at ~12.3 million bales for 2013/14. Cotton production supports Pakistan's industrial sector, facilitating 400 textile mills and hundreds of ginning factories, heavily depend on cotton, which engages around 1.3 million farmers. Cotton and cotton products account for 8.2 % of the value added in agriculture and contribute about 10 % to GDP and 55 % to foreign exchange. Along with this, it also adds over US$ 2.5 billion to the national economy. In the current scenario, 30 % to 40 % of cotton ends up as domestic consumption, while the remaining is exported as a raw cotton, yarn, cloth, and garments. Due to increasing industries in the country, Pakistan annually imports around 1.5-2.0 million bales to meet the local demand. Pakistan consecutively missing its target for the last seven years due to disease called Cotton leaf curl disease (CLCuD).

1.7. Cotton leaf curl disease in Pakistan

Major hindrance to cotton production is the CLCuD which is caused by a complex of begomoviruses. This first evidence of CLCuD was noted earlier near the city of Multan in 1967

(Aslam and Gilani, 2000; Hussain and Ali, 1975), but, an epidemic appeared in 1985. During this year, the disease rapidly spread across most growing areas of cotton in Pakistan and into northwestern portion of India. This major outbreak of CLCuD was referred to as the "1st epidemic" in Pakistan, in which several monopartite begomoviruses were involved specifically *Cotton leaf curl Multan virus* (CLCuMuV), *Cotton leaf curl Kokhran virus* (CLCuKoV), *Cotton leaf curl Alabad virus* (CLCuAV), *Cotton leaf curl Rajasthan virus* (CLCuRaV), *Papaya leaf curl virus* (PaLCuV) and *Cotton leaf curl Bangalore virus* (CLCuBaV) with alpha- and beta-satellite components (Mansoor *et al.*, 2003b; Briddon *et al*., 2003).

With the passage of time, CLCuD was spread from Pakistan to India and from India to China (infecting cotton, okra and hibiscus plants). Current reports show the presence of CLCuMuV in the Philippines (infecting hibiscus plants). More interestingly, in all these locations of Asian part, only Faisalabad strain of CLCuMuV was found present, while Pakistan is considered the centre of diversity.

The financial losses due to the CLCuD epidemic were estimated at US $5 billion during 1992-97 (Briddon and Markham, 2001). Cotton production in Pakistan was recovered to the pre-epidemic levels in the late 1990s by the introduction and adoption of resistant cotton cultivars, developed through conventional breeding programs (Rehman *et al.*, 2000). Unluckily, from 2001 onwards, the previously found resistant varieties succumbed to display CLCuD (Mahmood *et al.*, 2003; Mansoor *et al*., 2003b), and then no commercial cotton cultivar show resitance to the disease, other than a low level of tolerance. This resistance breaking was attributed to a distinct begomovirus, *Cotton leaf curl Burewala virus* [now identifed as Cotton leaf curl Kokhran virus-Burewala (CLCuKoV-Bu) after revision by Brown *et al.*, (2015)], in association with a recombinant form of CLCuMB (Amin *et al*., 2006; Amrao *et al*., 2010). This epidemic on cotton was noticed on all previously resistant cultivars in the vicinity of Burewala district, referred to as the "2nd epidemic". From 2001 onwards, the Burewala strain was

disseminated with other regions of Pakistan and northwestern India with un-expected losses in cotton as well as in other commercial crops.

1.8. Revision of cotton infecting begomoviruses taxonomy

Brown *et al* (2015) revised the list of all previously documented begomoviruses in the databases. Begomovirus, the largest genus in the family *Geminiviridae*, is comprised of at least 283 species, which encompass more than 3500 full-length begomoviruses. Large number of sequences in the databases emphasizing its prime importance as well as significant degree of genetic diversity which is widespread in the tropical and sub-tropical regions

Previously classification of begomoviruses was based on pairwise sequence distance was more stable and easier for all. This method of classification was also adopted by ICTV concerned with the *Anelloviridae* and the *Circoviridae.* Due to increasing number of begomoviruses, certain constraints were also encountered with their nomenclature. To come over this problem, a new strategy for its classification was delineated, STD (species demarcation tool). The cutoff value for new isolate of begomovirus was <91 %. (details are given in section 1.4.1).

Based on new/current revision of begomoviruses by Brown *et al.*, (2015), the taxonomy of begomoviruses has been revised. The number of cotton infecting begomovirus species has been reduced to five, previously from 9-10 documented species. From the total list, cotton infecting begomoviruses are given in the Table 1.2. Based on this revision, new names of isolates will be used in this study.

Table 1.2. Revision of Cotton infecting begomovirus taxonomy by ICTV. Species names are shown in bold italics, and isolate names are given in regular font. For species that have known

strains, one isolate from each strain is shown, and the type isolate is the first one listed. Sequence accession numbers are also listed.

Specie	Strains	Accession No.
Cotton leaf curl Alabad virus		
Cotton leaf curl Alabad virus	Cotton leaf curl Alabad virus-Alabad	AJ002452
Cotton leaf curl Alabad virus	Cotton leaf curl Alabad virus–Haryana	GU112081
Cotton leaf curl Alabad virus	Cotton leaf curl Alabad virus–Karnal	GU112004
Gossypium punctatum mild leaf curl virus	Cotton leaf curl Alabad virus-Lobatum	FJ210467
Gossypium punctatum mild leaf curl virus	Cotton leaf curl Alabad virus-Multan	EU384575
Cotton leaf curl Banglore virus		
Cotton leaf curl Banglore virus	Cotton leaf curl Banglore virus	AY705380
Cotton leaf curl Gezira virus		
Cotton leaf curl Gezira virus	Cotton leaf curl Gezira virus–Sudan	AF260241
Cotton leaf curl Gezira virus	Cotton leaf curl Gezira virus-Burkina Faso	FN554540
Hollyhock leaf crumple virus	Cotton leaf curl Gezira virus–Cairo	AJ542539
Cotton leaf curl Gezira virus	Cotton leaf curl Gezira virus–Cameroon	FM210276
Cotton leaf curl Gezira virus	Cotton leaf curl Gezira virus–Egypt	AF155064
Hollyhock leaf crumple virus	Cotton leaf curl Gezira virus–Hollyhock	AF014881
Okra leaf curl virus-Cameroon	Cotton leaf curl Gezira virus–Lysoka	FM164726
Bean leaf curl Madagascar virus	Cotton leaf curl Gezira virus–Madagascar	AM701757
Okra leaf curl Mali virus	Cotton leaf curl Gezira virus–Mali	EU024120
Cotton leaf curl Gezira virus	Cotton leaf curl Gezira virus–Niger	EU432373
Cotton leaf curl Gezira virus	Cotton leaf curl Gezira virus-Okra	FJ868828
Cotton leaf curl Kokhran virus		
Cotton leaf curl Kokhran virus	Cotton leaf curl Kokhran virus–Kokhran	AJ496286
Cotton leaf curl Burewala virus	Cotton leaf curl Kokhran virus–Burewala	AM421522
Cotton leaf curl Burewala virus–Layyah	Cotton leaf curl Kokhran virus–Layyah	HF549182
Cotton leaf curl Shadadpur virus	Cotton leaf curl Kokhran virus-Shadadpur	FN552001
Cyamopsis tetragonoloba leaf curl virus	Cotton leaf curl Kokhran virus–Lucknow	GU385879
Cotton leaf curl Multan virus		
Gossypium darwinii symptomless virus	Cotton leaf curl Multan virus–Darwinii	EU365613
Cotton leaf curl Multan virus	Cotton leaf curl Multan virus–Faisalabad	AJ002447
Cotton leaf curl Multan virus	Cotton leaf curl Multan virus-Hibiscus	JN807763
Cotton leaf curl Multan virus	Cotton leaf curl Multan virus–Hisar	AJ132430
Cotton leaf curl Multan virus	Cotton leaf curl Multan virus–Pakistan	EU365616
Cotton leaf curl Rajasthan virus-	Cotton leaf curl Multan virus- Rajasthan	AF363011

1.9. Importance of Begomoviruses in Pakistan

As globalization of human activities, free transport of plant hosts and vectors by global trade and modern pattern of agricultural practices increase coupled with climate change, the

viruses and their strains are evolving rapidly. Geminiviruses specifically begomoviruses make the largest group and are quite diverse in nature.

Geographically, in terms of severity and genetic variability, Pakistan lies on the hot plate of geminiviruses/begomoviruses. Like African region, this region was also account for epidemic of begomoviruses for the three decades. Many important crops were engulfed and substantial losses were accounted.

Although, many factors are contributing towards in the emergence of new begomoviruses, including, the over cropping, increased recombination between geminiviruses specifically begomoviruses and environmental conditions for vectors. The phenomenon of recombination is very frequent, occurs at every taxonomic level and is a significant contributor in evoluting new geminiviruses (Fondong *et al.,* 2000; Padidam *et al*., 1999; Stanley *et al*., 1985).

Pakistan has faced heavy losses in their national economy, because of CLCuD during 1992-97, accounts for 5 billion US $ in cotton only (Briddon and Markham, 2000). The disease on muskmelon, *Cucumis melo,* apparently caused 100 percent yield loss in some areas of Punjab, Pakistan (Malik *et al*., 2006). Whereas, the begomoviruses infection in Pakistan has increased its vista to greater extent during last decade to plants like *Hibiscus syriacus* (Riaz *et al*., 2016), *Duranta repens* (Mustujab *et al.,* 2015), *Hibiscus rosa-sinensis* (Akhtar *et al.,* 2014), *Luffa cylindrical* (Zia-ur-Rahman *et al.,* 2013), *Capsicum spp*. (Tahir *et al.,* 2010; Shafiq *et al.,* 2010), *Codiaeum variegatum* (Anwar *et al.,* 2012), *Ricinus communis* (Fareed *et al.,* 2012), Legumes (Ilyas *et al*., 2010), *Sonchus arvensis* (Mubin *et al.,* 2010*), Momordica charantia* (Tahir *et al.,*2010), *Rhynchosia minima* (Ilyas *et al.,* 2009), Potato (Mubin *et al.,* 2009), *Zinnia elegans, Solanum nigrum* and *Ageratum conyzoides* (Haider *et al.,* 2007b) *Vigna aconitifolia* (Qazi *et al*., 2006), *Vinca minor* (Haider *et al.,* 2007a), Bell pepper (Tahir and Haider, 2006),

Duranta erecta (Iram *et al.*, 2004), Tomato, Chilli and pepper (Hussain *et al.*, 2004; Shih *et al*., 2003; Siddiqui *et al.*, 1999).

1.10. Objectives of this study

The study has been designed with the following objectives:

1) Identification of begomovirus isolates infecting cotton.

2) Development of infectious clones of CLCuBuV and infectivity studies.

3) Development of polyclonal antibodies against coat protein gene of CLCuBuV.

2. REVIEW OF LITERATURE

Begomoviruses are economically critical to vegetable, ornamental plants and fiber crops worldwide. Mutation and recombination are a diversifying force to the evolution of new begomoviruses. A lot of work has been done on crop-infecting begomoviruses. Pakistan is an agricultural country and begomoviruses are causing devastating yield losses in important cash and food crops. There is a need to regularly identify these viruses. A summary of work done in Pakistan and elsewhere on new species, isolates/variants of begomoviruses and epidemics caused by these viruses are summarized below.

Kothandaraman *et al.* (2016) investigated infected black gram (*Vigna mungo* L. Hepper) showing the yellow discoloration of pods and seeds. They were suggesting the presence of *Mung bean yellow mosaic virus* (MYMV) in the seed. Different seed portions were processed by using polymerase chain reaction (PCR), Southern blot analysis and sequencing. Sequence analysis showed the presence of MYMV in seed coat, cotyledon and embryonic axes. The presence of virus was further confirmed by double antibody sandwich- enzyme linked immunosorbent assay (DAS-ELISA) and immunosorbent electron microscopy (ISEM). Seedlings from infected seeds did not shows any begomoviral symptoms, indicating as MYMV is a seed born virus.

Lozano *et al*. (2016) collected plant samples *Ipomoea batatas* and *Ipomoea indica* from Spain and *Merremia dissecta* from Venezuela. Samples were investigated for the presence of begomoviruses and its associated satellites. They obtained 18 clones, which were non-coding sub viral molecules of ~700 nts in size. Analysis of these clones showed maximum structural identity with ToLCV-sat, comprising A-rich region, satellite conserved region (SCR), a stem-loop structure containing a nona-nucleotide and a second predicted stem-loop. These features deviate them from betasatellite group and a distinct class was proposed named deltasatellite.

Saleem *et al*. (2016) investigated the phenomena of recombination for some old begomoviruses which were usually involved in evolution of new viruses through recombination. Remarkable example of this was the Cotton leaf curl Burewala virus-caused epidemic in cotton across Pakistan and India, was recombinant between CLCuMuV and CLCuKoV. Recombination detection programs (RDP) and phylogenetic analysis revealed that CLCuKoV and CLCuMuV were highly recombinant virus. Furthermore, they found that in majority cases, CP was usually donated by CLCuKoV while Rep was donated by CLCuMuV with a rate 4.96×10^{-4} and 2.706×10^{-4}. Beside this, they found the propagation of CLCuMuV at an alarming rate in the Asian countries, specifically Pakistan, India and China, which may be a threat in near future for economical crops like cotton.

Snehi *et al*. (2016) observed severe mosaic like symptoms on ornamental species of Jatropha (*J. podagrica, J. multifida and J. integerrima*) in the garden at Lucknow. Samples were collected from infected jatropha plants and processed for the presence of begomovirus using specific primers in a diagnostic PCR. All species of jatropha were detected positive for amplicons of ~1.2 kb. Clones were sequenced and analyzed, shows highest nucleotide identities and close phylogeic relationships *Jatropha mosaic India virus* in *J. podagrica* (HQ848382); *Tomato leaf curl Patna* virus in *J. multifida* (HQ848381) and Papaya leaf curl virus in *J integerrima* (JQ043440). This was a new report of multiple begomoviruses infecting ornamental jatropha species from India.

Sohrab *et al.* (2016) developed transgenic cotton plants using βC1 gene in antisense orientation using Cauliflower mosaic virus-35S promoter and nos-terminater for expression. Transgenic and healthy plants were challenged through agrobacterium mediated inoculation. Transgenic plants were observed symptomless and healthy plants were noted with begomoviral symptoms appearance, indicating the resistance of plants.

Srivatava *et al.* (2016) identified an ornamental plant Hibiscus (*Hibiscus rosa-sinensis*) with symptoms like severe leaf curling, vein thickening and enation on most of the plants in the garden in Lucknow during 2011. DNA was extracted from infected and healthy plants and subjected to diagnostic PCR, using universal primers for coat protein. All suspected plants were found positive. Full-length virus from two samples was amplified using rolling cycle amplification (RCA) method, while associated betasatellite was amplified using universal primers. Analysis of the sequences had 98% nucleotide identity with CLCuMuV and CLCuMB. This was the first report of CLCuMuV-CLCuMB complex infecting Hibiscus in India.

Zaidi *et al.* (2016) collected infected cotton sample across major cotton belt in Pakistan to investigate the status of CLCuD, usually caused by CLCuKoV-Bu/CLCuMB complex. 31 samples were investigated using diagnostic PCR, showed the presence of CLCuKoV-Bur in all samples with CLCuMB while bipartite begomovirus ToLCNDV was also found in 20 samples only. They also studied that co-infection of ToLCNDV with CLCuKoV-Bu/CLCuMB complex lead to enhance the level of CLCuMB, using quantitative PCR, which is symptom determinant.

Baltes *et al.* (2015) introduced another novel strategy to combat geminiviruses, which is responsible for heavy economical losses in the world. They targeted *Bean yellow dwarf virus* (BeYDV) using CRISPR-Cas system. They found in transient assays using BeYDV-based replicons that the CRISPR-Cas regents reduced the copy number of viral genomes by introducing mutation in the viral genome. Transgenic plants producing CRISPR-Cas regents were inoculated with BeYDV, showed resistance, indicating a novel strategy for engineering resistance against geminiviruses, specifically begomoviruses.

Hak *et al.* (2015) identified some tomato infecting begomoviruses with mutant V2. They found that such virus has symptomless effect on plants. To study the role of V2 gene, two mutant versions of the virus, one impaired in V2 silencing-suppression activity, and other

carrying a non-translatable V2. They observed that both mutant viruses spread in the infected plants with newly emerged leaves with the same rate as wild type but the DNA accumulation was ten-time lower in comparison with the wild type. They concluded that this is due to lack of silencing suppression activity.

Tahir *et al.* (2015) isolated four begomovirus and three associated betasatellites from a wild weed *Ageratum conyzoides*, collected from Pakistan and Nepal. Virus showed highest level of nucleotide sequence identity with the available *Ageratum Enation Virus* (AEV) in the databases, while its counterpart reveals maximum identity with Ageratum yellow leaf curl betasatellite (AYLCB). Phylogenetic analysis of the virus sequences discrete those into two different clades, showing two different species, for which, Nepal and Indian strains were proposed. Agrobacterium mediated inoculation into *Nicotiana benthamiana, N. tabacum, Solanum lycopersicon* and *A. conyzoides* produced leaf curl symptoms except in *A. conyzoides.* They suggested that the AEV infecting plants with wide range from weeds to vegetables.

Rishishwar *et al.* (2015) collected samples of okra from ten different locations to investigate the causal agent of Bhendi yellow vein mosaic disease (BYVMD) in India. Coat protein was amplified from suspected samples with diagnostic primer sets and analyzed, revealed that seven of them belong to *Bhendi yellow vein mosaic virus* (BYVMV) and the rest were *Mesta yellow vein mosaic virus* (MeYVMV). Full-length viral DNA of begomoviruses were also amplified with specific primers and analyzed using RDP. They found that the full-length viruses were recombinant, resulting from MeYVMV and *Malvastrum yellow vein Yunnan virus* (MYVYNV) as parents. MYVYNV was first time identified on Indian region. Betasatellites and alphasatellites sequences were also amplified and studied. They found that these sequences were variants of previously reported satellites from this region infecting Okra.

Roumagnac *et al.* (2015) reported the new genus of family *Geminiviridae* known as *Capulavirus.* They demonstrated that the classification of geminiviruses is based on genome

organization, nucleotide sequence similarities and biological properties. The type of species include in the new genus Capulavirus is Alfa-alfa curl virus which is transmitted by aphid. This was the first report that aphid as an insect vector to transmitted geminivirus.

Ashraf *et al.* (2014) have compared the expression activity of bidirectional promoter of begomovirus specifically *Cotton leaf Burewala virus* (CLCuBuV) with CaMV 35S promoter. Expressional level of both promoter was compared by using an agro-infiltrated construct in the *Nicotiana tabacum* and *Gossypium hirsutum* and its expression activity was monitored by a GUS reporter gene. They found that begomovirus (CLCuBuV) Rep gene promoter showed a strong, consistent transient expression of the reporter gene (GUS) in *N. tabacum* and *G. hirsutum* leaves and demonstrated GUS activity two- to three-fold higher than the CaMV 35S promoter.

Shahid *et al*. (2014) collected samples of four different tomato plants with Tomato leaf curl disease. Two different virus isolates were found infecting tomato plants from two distinct locations. Sequence analysis showed, tomato plants of Central Komae infected with TYLCV whereas plants from South West of Komae were found associated with AYVV. Also, Tomato leaf curl Java betasatellite (ToLCJaB) was isolated from plants harboring AYVV. Tomato plants infected with TYLCV, AYVV and ToLCJaB were reported with two alphasatellite clones. These alphasatellites shared 99% identity and showed maximum identity with Sida yellow vein China alphasatellite.

Shahid *et al.* (2014) collected tomato plant samples in 2011, exhibiting leaf curl symptoms near Komae, Japan. Amplification through PCR, cloning and then sequencing was performed for all begomoviral components. Two samples were taken; one showed to be infected by TYLCV and other to be infected with AYVV. Plants infected with AYVV were seen to be associated with a beta satellite molecule of Tomato leaf curl Java betasatellite (ToLCJaB). This betasatellite has never been seen before, to be present in Japan. No

betasatellite was seen to be associated with TYLCV. Both the viruses contained an additional alpha satellite known to be Sida yellow vein China alphasatellite. This alphasatellite was only reported previously from Yunnan and Nepal. This suggests begomoviruses along with their associated satellites are now introducing in Japan.

Marval *et al.* (2014) provided an in-silico recombination analysis of Ageratum enation virus (AEV) isolated from an ornamental plant Marigold. Using Recombination Determination Program (RDP) AEV showed six recombination events based on the similarity of the sequence alignment proving AEV as a recombinant virus. First recombination was shown to have sequences from PedLCV and CrYVMV having starting and ending nucleotide coordinates at 2665-287. Second recombination event was encounter at nucleotide positions 388 to 935 encompassing sequences from TLCuKV and CrYVMV. Third recombination was shown to occur at 938 to 1042 comprising sequences from AEV and TLCuRV. At 1042 to 1200 nucleotide position fourth recombination event was confirmed via RDP having sequences from AEV and TLCuRV.Fifth recombination event was shown to begin at 1237 nucleotide position to 1366 nucleotide, major parent being TbCuSV and minor being CrYVMV. Final recombination was encountered at 2074 to 2210 nucleotide position having *Euphorbia leaf curl virus* and *Bhendi yellow vein Bhubhaneswar virus*.

Srivastava *et al*. (2014) reported the association of *Pedilanthus leaf curl virus* (PedLCV) with *Tabernaemontana coronaria* and *Cestrum nocturnum* in Lucknow, India. Diseased plants of Crape jasmine (*T. coronaria*, family *Apocynaceae*) displayed yellowing of leaves, mottling and dwarfing whereas night blooming jasmine (*C. nocturnum*, family *Solanaceae*) had leaf curling symptoms. For molecular characterization of begomoviral infection a DNeasy plant mini kit was used for total DNA extraction and amplified using Rolling circle amplification. Sequence analysis showed that the virus has highest nucleotide identity with *Pedilanthus leaf curl virus* (PedLCV).

George *et al*. (2014) collected leaf samples ofAmaranthus plants having severe leaf curl disease. They characterized for the first-time full-length clones of a monopartite begomovirus having both alphasatellite and betasatellite molecules. The isolated full-length begomovirus of 2755 nts shared highest sequence identity with *Chilli leaf curl virus*. The isolated alphasatellite and betasatellite molecules shared maximum sequence identity with Chilli leaf curl alphasatellite and Tomato yellow leaf curl Thailand betasatellite respectively.

Manzoor *et al*. (2014) collected cotton sample showed begomoviral symptoms from Punjab province, Pakistan. Full-length virus clones (~2.7 kb) were amplified and sequenced, showed maximum identity with CLCuBuV as a dominant virus in this region from 2001 onward. Based on species demarcation threshold, viruses identified were isolates of CLCuBuVs. Along with this, a small number of leafhopper-transmitted (Genus *Mastrevirus*) viruses were also identified from cotton plants. Four chickpea chlorotic dwarf virus (CpCDV) were sequenced, showed the highest levels of nucleotide sequence identity to isolates of CpCDV strains C and D previously reported from Pakistan, but has not been identified in cotton. This was the first report of isolation CpCDV from cotton.

Fiallo-Olivé *et al*. (2014) reported a new bipartite begomovirus species infecting *Jacquemontiasp,* with the common name morning glory from the family *Convolvulaceae*, being grown in Venezuela. It showed yellow mosaic symptoms in infected plants and so was named as *Jacquemontia yellow mosaic virus (*JaYMV). For JaYMV DNA A, highest nucleotide similarity of 78.4% and 77.0% were obtained with *Merremia mosaic virus* (MerMV) and *Jacquemontia mosaic Yucatan virus* (JacMYuV), respectively; while DNA-B showed 74.6% nucleotide sequence identity with JacMYuV. Recombination analysis showed JacMYuV and MerMV as major and minor parents of JaYMV, respectively. They also noted that MerMV is also infecting tomato crops in Venezuela, providing a viral evolutionary linkage

between *Convolvulacae* species and tomato crops i.e. spreading from wild plant species to economically important food crops.

Ullah *et al.* (2014) identified Eggplants (*Solanum melongena* L.) with severe leaf mosaic and mottling symptoms. Full-length virus was amplified through RCA and its counterpart with betasatellite-specific primers. Clones were sequenced and analyzed, showed 99% identity with CLCuBuV and 97% nucleotide sequence identity with Cotton leaf curl Multan betasatellite ($CLCuMB^{Mul}$). Based on species demarcation threshold, both were identified as variant of CLCuBuV and CLCuMB. This was the first report of eggplant as an alternative host of cotton infecting begomovirus in Pakistan.

Kumar *et al.* (2014) verified the function of bC1 gene of betasatellite as pathogenicity determinant. Cotton leaf curl Multan betasatellite (CLCuMB) was used as a VIGS by replacing its bC1 gene with a fragment of host gene (Su) or a reported transgene (uidA). The modified CLCuMB was further co-agroinoculated with CLCuMuV, CLCuKoV and AEV separately into *Nicotiana tabacum, N. benthamiana, Solanum lycopersicum, Arabidopsis thaliana* and *Gossypium hirsutum* plants. The plants were observed healthy, showed efficient silencing of the targeted genes in the helper virus.

Wang *et al*. (2014) reported *Indian cassava mosaic virus* isolate from Singapore (ICMV-SG). Its complete nucleotide sequence was determined. Infection caused by ICMV-SG was more severe in *Nicotiana benthamiana*, compared to other isolates *of Indian cassava mosaic virus.* Infectious clones of DNA-A of ICMV-SG, alone caused severe damage to *N. benthamiana* while less damage was seen in *Jatropha curcus*. It was proved after agro-infection assays that for severe damage in *Jatropha curcus,* both DNA-A and DNA-B were required.

De Almeida *et al*. (2013) collected five samples from infected cotton plants exhibiting chlorotic spots, leaf distortion. These samples were collected from Paraiba state. DNA-A and

DNA-B genomic components were sequenced and cloned in a pBlue-script after rolling circle amplification. Digestion was performed by using restriction enzyme *Xba*l. Cloned DNA-A and DNA-B genomes were of 2670 and 2650 nucleotides. DNA-A showed a highest level of nucleotide sequence identity (77.8 percent) to tomato common mosaic virus. Thus, it was named as cotton chlorotic spot virus (CCSV). While DNA-B showed highest levels of nucleotide identity (67.8 percent) to *Tomato yellow vein streak virus*.

Harimalala *et al*. (2013) isolated begomovirus and two alphasatellite from cassava plant showing mosaic disease in Madagascar. They analysed the sequences. Virus was an isolate of *East African cassava mosaic Kenya virus* (EACMKV), while both satellite molecules had maximum relatedness to Cotton leaf curl Gezira alphasatellite (CLCuGeA) with 80% nucleotide sequence identity. Alphasatellite contained the basic feature as 315AA Rep protein on virion strand, an adenine rich region and satellite conserved region (SCR) which containing a stem loop structure called nona-nucleotide (TAG/ATATTAC). Based on cut-off value for the alphasatellite as 83% nucleotide identity, the molecule identified here were proposed as new species, named as Cassava mosaic alphasatellite.

Hussain *et al.* (2013) isolated nucleotide sequences of alpha satellites being associated with monopartite begomoviruses from the plant *Xanthim strumarium* L. Samples were collected from Pakistan. Highest levels of nucleotide sequence identity (96.7%) for alpha satellite was seen with a Gossypium darwinii isolate. The second highest nucleotide sequence identity (85.8%) was seen with Papaya leaf curl alphasatellite isolate. Therefore, the name given was *Gossypium darwanii* symptomless alphasatellite isolate Xanthium, as reported a new isolate of *Gossypium darwani* symptomless alphasatellite species.

Ilyas *et al.* (2013) isolated two begomovirus clones (KN4 and KN6) from Madagascar periwinkle (*Catharanthus roseus*) which is a common ornamental plant in this region. They found that KN4 showed maximum nucleotide sequence identity (86.5 %) with chili leaf curl

India virus, followed by PaLCV with 84.4 % maximum identities. Based on species demarcation threshold level, KN4 was a new species, for which a name "Catharanthus yellow mosaic virus" was proposed, infecting Catharanthus. On the other hand, KN6 showed highest level of nucleotide sequence identity (95.9 % to 99 %) to Papaya leaf crumple virus, which was currently reported virus from India. Previous literature showed that KN4 and KN6 hybrid from the recombination of Pedilanthus leaf curl virus and Croton yellow vein mosaic virus.

Lee *et al.* (2013) described another strategy for resistance against geminiviruses rather than targeting viral protein or RNAs. They produced transgenic *Nicotiana tabacum*, expressing 3D8 single chain variable antibody (scFv), which has DNase activities. They found that this protein was produced in the cytosol of protoplast in the *N. tabacum* plants where it targets the viral DNA directly, rather than its products. They observed that the transgenic plants were completed resistant to Beet curly top virusand the Beet severe curly top virus, without viral accumulation or symptoms development. Furthermore, this strategy was much more effective to targets ssDNA viruses without adverse effects on the transgenic plants growth and production.

Leke *et al.* (2013) demonstrated recombinant virus infecting Okra plants in West Africa. The virus isolated from Okra plants was found to be a recombinant of CLCuGeV and *Okra yellow crinkle virus* (OYCrV). The virus shared 87.8% nucleotide identity with CLCuGeV. Therefore, it can be regarded as a new species of begomoviruses. New strains of betasatellites and alphasatellites were also isolated from the infected plants. Betasatellite from CLCuGeV had 93.3% identity to CLCuGeV and was named as CLCuGeB whereas; alphasatellite shared 97.3% identity and was named as CLCuGeA. Another alphasatellite had 95.2% Okra leaf curl Burkina Faso alphasatellite so it was named as OLCuBFA. The results reflected diversity of begomoviruses and satellite complexes in infected okra plants in Africa.

Marwal *et al*. (2013) collected infectious leaf samples from an ornamental plant (*Lantana camara*; Common name: Spanish Flag) showing yellow vein disease. The presence of begomovirus infection was confirmed through PCR by using degenerate primers specific for Coat protein (CP). Further confirmation was done by southern blot hybridization. Phylogenetic analysis was also performed, which showed begomovirus infecting *L.camara* to cluster with an isolate of *Tomato leaf curl Bathinda virus*. This demonstrated the first report showing begomovirus to be infecting *L. Camara*.

Geetanjali *et al.* (2013) observed an infected wild morning glory plant (*Ipomoea purpurea*) with whitefly in New Delhi, India during 2009. Viruliferous Whitefly were transferred to healthy wild morning glory plant (Ipomoea purpurea). They observed two distinct symptoms were appeared on the plant. DNA was extracted from the plant and subjected to RCA. Multiple clones were sequenced. Virus was 91.8 to 95.3% identical to Sweet potato leaf curl virus (SPLCV) but phylogenetically discrete from previously reported from Brazil, China, Japan and USA. They found that the virus associated with two distinct betasatellites - Croton yellow vein mosaic betasatellite and Papaya leaf curl betasatellite. This was the first report that SPLCV infecting new species of morning glory (*I. Purpurea*) in India. Furthermore, they also reported of SPLCV with two distinct betasatellite infecting *I. purpurea*.

Kamaal *et al*. (2013) observed leaf curl symptoms in French bean (*Phaseolus vulgaris*). The complete DNA A sequence showed 80 % identity with *Cotton leaf curl Bangalore virus* and comprised 2741 nucleotides with six ORFs. According to the species demarcation criteria it was found as a new begomovirus species named as *French bean leaf curl virus*. Similarly, an associated betasatellite molecule showed 80 % identity with Papaya leafcurl betasatellite and comprised of 1,379 nucleotides with a single ORF bC1. No recombination events were seen in either DNA-A or associated betasatellite of FbLCV.

Kumar *et al.* (2013) isolated two new species of begomovirus infecting okra in Bihar, India. Both isolate showed maximum nucleotide sequence identity with *Radish leaf curl virus* (RaLCV), for which RaLCV-Hajipur1 and RaLCV-Hajipur2 name was proposed. Associated satellites were identified as Cotton leaf curl Burewala alphasatellite and Tomato leaf curl Bangladesh betasatellite (ToLCBDB). *N. tabacum* were agroinoculated alone or with both isolates in association with their cognates.symptoms were appeared after 27 days of post inoculation. Isolates were amplified from infected tabacum and sequenced. Sequence analysis revealed the presence of new isolate, RaLCV-Hajipur3, resulting from recombination of RaLCV-Hajipur1 and RaLCV-Hajipur2 as major and minor parents.

Pramesh *et al*. (2013) reported a weed infecting monopartite begomovirus, Croton yellow vein mosaic virus that causes bright yellow vein symptoms in croton but leaf curl symptoms in other host species including tobacco species, vegetables and ornamental plants. In this study, it was revealed that CYVMV DNA-A along with a betasatellite (CroYVMB) is responsible for yellow vein mosaic disease of croton.

Xie *et al.* (2013) characterized a full-length sequence of begomovirus from *Malvastrum coromandelianum* and tomato. Sequence analysis across the whole genome showed highest levels of sequence identities with TYLCCV, except for the region of C4 on the complementary sense strand. Based on demarcation threshold level, a new name Tomato leaf curl Yunnan virus (TLCYnV) was proposed. Virus was agro-inoculated to various plants, showed highly infectious except *M. coromandelianum.* They were also investigated that TLCYnV was alone infecting tomato in the field. Transgenic expression of C4 revealed that TLCYnV could produce more abnormalities in comparison of C4 protein of TYLCCNV.

Zia-ur-Rehman *et al.* (2013) collected *Luffa cylindrica* (Gia tori) in the region of Burewala during 2011. Plants exhibiting symptoms like leaf curling and stunting growth, usually caused by Begomovirus. Total DNA was isolated from five different symptomatic

leaves using CTAB method. Plant extract was analyzed using 1.1kb fragment of CLCuBuV probe in the southern blot hybridization. They found all sample positive with the probe. Rolling circle amplification (RCA) was performed from the total DNA. RCA amplified products were digested with *EcoR*I and cloned into pGEM-3Zf+ (Promega, Madison, WI) plasmid vector. Associated alpha and betasatellite were amplified from T.DNA with using abutting primer for alpha and beta molecule. Clones were sequenced and analyzed. Virus was determined to be 2753 nts and showed maximum nucleotide identity with CLCuBuV. Beta and alpha molecules were also sequenced and closely related to CLCuMB and *Gossypium darwinii* symptomless alphasatellite (GDaSA). This constitutes the first report that a complex like CLCuBuV-CLCuMB-GDaSA infecting cucurbitaceous species and *L.cylindrica* as an alternative host of CLCuD.

Singh *et al.* (2012) characterized two new monopartite begomoviruses infecting radish plants in India. First virus was *Radish leaf curl Varanasi virus* was a new recombinant species sharing 87.7% of identity with *Tomato leaf curl Bangladesh virus*. The other virus isolated from radish was variant of *Croton yellow vein mosaic virus*-India having 95.8% nucleotide identity. Furthermore, RDP analysis showed RaLCV as a probable recombinant having sequences from *Euphorbia leaf curl virus* and PaLCuV. When these viruses were inoculated in plants, they induced mild symptoms. However, when they were coinfected with their betasatellites, resulted in symptom severity and increased in viral titers.

Anwar *et al.* (2012) identified a new variant of ClYMV from infected croton plants which is grown as an evergreen shrub in gardens. Samples showing mild leaf curl and yellowing symptoms were collected from Lahore in 2012. Full-length virus was amplified and sequenced showing 99.1% identity to ClYMV. Attempts to amplify DNA-B or β satellite component failed. It was a first report of ClYMV infecting croton and hence proving croton as a host plant of begomoviruses.

Fareed *et al.* (2012) observed leaf curling and vein thickening on caster bean, grown as ornamental plant in this region for oil production. Suspected symptoms were looking as caused by begomovirus. DNA was extracted from infected plant. Degenerative primers were used for detection initially. Full-length virus was amplified by using abutting primers and associated betasatellite was cloned with its universal primers. They determined the sequence to be 2759bp with maximum identity level (98%) with CLCuKoV-Bu and betasatellite (1373bp) as CLCuMB with 96% nucleotide identity. Based on the species demarcation criteria, both were the variant of CLCuKoV-Bu and CLCuMB. This was the first report that CLCuBuV-CLCuMB complex infecting caster bean as an alternative host.

Hina *et al.* (2012) screened different kinds of cotton plants to check for begomoviral infection. Viruses were amplified through PCR using specific primers, then cloned and sequenced. The study revealed CLCuBuV alone to be responsible for destroying the most important economic crops. Pairwise sequence comparison of isolates from this study showed 98.9 percent to 99.55 percent homogeneity with CLCuBuV. Among all, one of the virus isolates had lacked Rep region, which is always found in begomoviruses.

Ilyas *et al.* (2012) from Pakistan first reported infection of begomoviruses in *Catharanthus roseus*. They isolated two viruses from the infected plant. Sequence analysis of both the viruses isolated proved them to be a recombinant of *Pedilanthus leaf curl virus* (PedLCV) and *Croton yellow vein mosaic virus* (CrYVMV). First virus isolated KN4 was a new species having 86.5% identity with unpublished chili leaf curl India virus (ChiLCIV) and 84.4% identity to PaLCV. On the other side second virus KN6 was shown to have 99% identity with *Papaya leaf crumple virus* (PaLCrV), reported from India.

Rocha *et al.* (2012) collected leaf samples of weeds and tomatoes from tomato growing areasduring 2008 and 2010. Tomato mottle leaf curl virus isolates were obtained from the samples. Molecular characterization of DNA-A was performed. Sequence analysis showed

Tomato mottle leaf curl virus to be a new world begomovirus with highest levels of nucleotide sequence identity to begomoviruses from Brazil and Central America. Brazilian begomoviruses in most cases were of recombinant nature, the same pattern was followed by Tomato mottle leaf curl virus. This was said to be the first report for Tomato mottle leaf curl virus as a new species of begomovirus.

Tiendrébéogo *et al.* (2012) collected leaf samples from cassava showing mosaic disease symptoms. These leaf samples were collected from the areas of Burkino Faso. DNA-A component from four samples were cloned and sequenced. Sequence analysis showed 99.9 % sequence identity to African cassava mosaic virus (ACMC) while less than 89 % sequence identity to previously discovered begomovirus. DNA-B component was also cloned and sequenced. As all the genomic components of begomoviruses had an old world begomovirus organization, it was a distinct species, which was named as African cassava mosaic Burkino Faso virus (ACMBFV). After recombination analysis, it was observed that inter species recombination had occurred. Major part belonging to the ACMV isolates of West Africa and minor part belonging to Tomato leaf curl Cameroon virus and Cotton leaf curl Gezira virus. This was the first report of recombination event between a monopartite African begomovirus and a bipartite begomovirus.

Zulfiqar *et al*. (2012) characterized virus from *Vernonia cinerea* showing yellow vein. The sequence was determined to be 2739 nts and arranged with typical architecture of begomovirus. They analyzed the sequenced, showed highest (78.9%) nucleotides sequence identity with the recent *Vernonia yellow vein virus* (VeYVV) from India. They also cloned alpha- and beta-satellite associated with virus. They found highest nucleotide identity 51.2% for betasatellite with Vernonia yellow vein virus betasatellite (VeYVVB), and alphasatellite showed maximum relatedness 70.7 % with *Gossypium mustelinium* symptomless alphasatellite (GMusSLA). Based on species demarcation threshold, virus associated with cognate alpha and

betasatellite was a new species for which the name Vernonia yellow vein Fujian virus (VeYVFJV) was proposed.

Cuong *et al.* (2011) collected leaf samples from tomato and papaya exhibiting leaf curl symptoms in the areas of northern Vietnam. These samples were amplified, cloned and then sequenced. The isolates were named as tomato-89, tomato-100 and papaya-31. Their sequences were submitted to GenBank. After molecular characterization of these isolates, it was confirmed that these viruses belonged to Tomato leaf curl Hainan virus. This virus was detected in Vietnam for the first time. It was found in Papaya which was a new host of the virus, while the virus isolated from tomato was a distinct species which was named as Tomato leaf curl Hanoi virus. This was a recombination of Papaya leaf curl China virus and Ageratum leaf curl virus.

Gaur *et al.* (2011) collected leaf samples from stunted *Mimosa pudica* L. showing small yellow leaves. The presence of begomoviruses was doubted due to the presence of Bemesia tabaci on the leaves of the plants. Primers specific for coat protein of begomovirus were used to detect the presence of begomoviruses. Infected samples were cloned and sequenced. High levels of nucleotide sequence identity, 97 percent were seen with Ageratum yellow vein virus-Guangxi (AYVV-Gx) and with AYVV. It was the first report of begomovirus infecting Mimosa pudica, making it as an alternative host.

Kumari *et al.* (2011) characterized a new species of tomato infecting begomovirus having 88.3% if identity with Tobacco curly shoot virus (TbCSV) and was thus named as Tomato leaf curl Ranchi virus (ToLCRnV). The betasatellite isolated showed 74.5% identity with Tomato leaf curl Bangladesh beta was therefore named as Tomato leaf curl Ranchi beta. Phylogenetic analysis showed ToLCRnV in close relationship with Tomato leaf curl Bangladesh virus, Tobacco curly shoot virus and Tomato leaf curl Gujarat virus. It was also observed that β satellite was efficiently replicated by DNA A whereas, DNA B showed less

accumulation. Leaf disk assay was also done which showed that DNA A can trans-replicate homologous DNA B and β as compared to heterologous ones.

Malik *et al.* (2011) constructed the partial dimer of *Tomato leaf curl Palampur virus* (ToLCPMV) after isolation from muskmelon, exhibiting severe yellow vein disease in association with *Zucchini yellow mosaic virus* (ZYMV). ToLCPMV has been reported from India and Iran but its partial dimer was not yet constructed.They did the analysis of isolates, showed maximum relatedness with the reported ToLCPMV-bipartite begomovirus. DNA A have six ORFs while DNA B with two ORFs, having truncated NSP in comparison with the reported NSP of ToLCPMV. They produced the partial tandem repeats of ToLCPMV and agroinoculated on *N. benthamiana*. The dimer did not produce any symptom with the truncated NSP. They used component B of ToLCNDV with DNA A of ToLCPMV, lead to a hypersensitive response along the veins in the *N. benthamiana*. The agrobacterium mediated infiltration of muskmelon with a construct expressing NSP of a ToLCNDV under the control of Cauliflower mosaic virus 35S promoter induced hypersensitive response in the veins. They suggested that NSP gene was responsible for the hypersensitive response. The infection in the muskmelon with defective NSP may be due to the synergistic interaction with ZYMV of potyvirus.

Pratap *et al.* (2011) isolated a new variant of bipartite Tomato leaf curl New Delhi virus from eggplant in India. Virus was isolated and cloned. Further analysis of the sequences proved it to be a variant with virus having 97.6% nucleotide identity with Tomato leaf curl New Delhi virus [India:NewDelhi:Papaya:2005]. DNA B showed 94.1% identity with Tomato leaf curl New Delhi virus [India-Nagpur-Eggplant-2009]. Variants pathogenicity was confirmed through agroinfiltration showing yellow mosaic symptoms in eggplant and leaf curling symptoms in tomato plants.

P**riyadarshini *et al.* (2011)** performed a series of experiments to study the function of V2 and Coat protein (CP) of the Cotton leaf curl Kokhran virus-Dabawali. Function of both gene were analyzed *in vitro* and *in vivo* by over-expression in *E. coli*, insect cells and transient expression in the plant system. They found that purified recombinant protein of V2 and CP interact each other using ELISA and surface plasmon resonance. During confocal microscopic studies of Sf21 cells expressing V2 and CP protein, they found V2 protein localized to the cell periphery while CP protein restricted to the nucleus of the cell. In one of the experiment, deletion of the N terminal localization signal of CP limited it distribution to the cytoplasm of the cell. In plant system, GFP-V2 was observed to cell periphery and YFP-CP mostly in the nucleus by agro infiltration into the *N. benthamiana* plants. When both were co-infiltrated, CP was found both in the cytoplasm as well as in the nucleus along with V2 protein. From these experiment, they concluded that V2 and CP interact each other in the intra and inter cell movement.

Zulfiqar *et al.* (2011) collected leaf samples of Morning glory exhibiting yellow mosaic and vein yellowing symptoms from different provinces of China. The genomic component, DNA-A was amplified from both the samples, one to be 2827 nucleotides while other to be 2801 nucleotides. They were named as JS-1 and Y338. JS-1 showed highest levels of nucleotide sequence identity, 97 percent with Sweet potato leaf curl virus (SPLCV) while DNA-A of Y338 showed highest levels of nucleotide sequence identity, 97.8 percent with SPLCV isolate from China. After phylogenetic analysis, JS-1 was crowded with SPLCV-US and Y338 being crowded with SPLCV-CN, confirming them to be new strains of Sweet potato leaf curl virus.

Amrao *et al.* (2010a) demonstrated that the first epidemic in cotton was caused by a complex of begomoviruses with association of a single betasatellite called Cotton leaf curl Multan betasatellite (CLCuMB in 1980s. In combat of CLCuD, new varieties of cotton were

introduced in the late 1990s. With a gap of some years, begomoviral symptoms were re-appeared on the resistant cultivars in the late 2000s. From the symptomatic cotton leaves, full-length virus was cloned along its counterpart. Virus was analyzed, showed to be recombinant of two previously identified viruses in cotton. Associated betasatellite molecule was also recombinant containing a fragment of approx. 100 nts of ToLCB. Based on species demarcation, a new name for virus was proposed as Cotton leaf curl Burewala virus. They found the lack of an intact TrAP protein in all burewala isolates and suggested that the causal agent for the second epidemic in cotton was "Burewala" strain only.

Amrao *et al.* (2010b) cloned begomoviruses and their associated components from six samples in the areas of Sindh. Out of the six samples, one was Cotton leaf curl Kokhran virus isolate. Sequence analysis of the other five clones showed nucleotide sequence identity to be 90 percent with begomoviruses seen to be involved in CLCuD. These isolates were grouped under species *Cotton leaf curl Shahadpur virus* (CLCuShV). After reviewing CLCuShV in detail, it was seen to have arisen due to recombination events. Two samples showed the association of an alphasatellite, making them a newly discovered species. All six samples were seen to be associated to betasatellites. Sequence analysis of two betasatellites showed them to be CLCuMB, both were recombinants. They suggested that begomovirus causing CLCuD in Sindh was entirely different from the begomovirus responsible for causing CLCuD in Punjab.

Graham *et al.* (2010) amplified, cloned and sequenced two DNA As and one DNA B Malvastrum americanum plants from Jamaica exhibiting yellow mosaic symptoms. They did the analysis of the genome components, representing typical bipartite begomovirus having six ORFs in DNA A and two ORFs in DNA B reported from the Western Hemisphere in Jamaica. One DNA A was a new species, named as Malvastrum yellow mosaic Helshire virus (MaYMHV), while other DNA A and DNA B form a cognate fair and represent a new species identified as Malvastrum yellow mosaic Jamaica virus (MaYMJV). DNA B of the MaYMJV

showed most relatedness to Sida golden mosaic virus-[USA:Florida] (SiGMV-[US:Flo]) and Sida golden mosaic Costa Rica virus-[Costa Rica] (SiGMCRV-[CR]). To verify Koch's postulate, they infected *M. americanum* with both components of MaYMJV, which produced yellow mosaic symptoms, as observed in plants during collection. They found 84% nucleotide sequence identities between both DNA-As MaYMJV and MaYMHV. Their phylogenetic analysis showed both cluster in the same clade with begomoviruses which infecting malvaceous weeds of Cuba and Florida. Recombination Detection Program (RDP) results showed that the genomes of MaYMJV and MaYMHV were evident of of inter-species recombination.

Herna´ndez-Zepeda *et al.* (2010) collected leaf samples from Rhynchosia minima exhibiting bright golden mosaic symptoms typical of begomoviruses from Yuctan. Both genomic components, DNA-A and DNA-B were cloned and sequenced. Nucleotide comparison of the two components showed highest levels of sequence identity, (87 percent for DNA-A and 71 percent for DNA-B) with cabbage leaf curl virus, thus distinct species and named as Rhynchosia yellow mosaic Yuctan virus (RhYMYuV). Phylogenetic analysis showed Squash leaf curl virus to be the closest relative of RhYMYuV. After looking at recombination analysis, it was seen that DNA-A had arisen due to inter molecular recombination.

Mubin *et al.* (2010) studied *Sonchus arvensis* which is a common weed in Pakistan. They isolated begomovirus, two betasatellites and two alphasatellites from S. arvensis showing yellow vein disease. Upon analysis, virus showed 95-99% nucleotide sequence identity with Alternanthera yellow vein virus (AlYVV) reported from China, Vietnam and India. One betasatellite shared nucleotide sequence identity (91.4-95.3%) with AYLCB, other showed (78.2 to 99.9%) identity with isolates of CLCuMB. Along with betasatellite, two alphasatellites were also predicted. The one was similar to PotLCuA and other was an isolate of HLCuA.They

concluded that monopartite begomovirus may be associated with multiple satellites which are predominant in this area.

Paprotka *et al.* (2010) collected Brazilian weeds showing symptoms like begomovirus. Virus along with satellites were amplified and sequenced. Alphasatellite was analyzed, showed typical genome architecture of begomoviral satellite molecule possessing a single ORF coding potentially for a replication-associated protein (Rep), a conserved hairpin loop containing Nona-nucleotides, and an adenine-rich region. They also sequenced two full-length begomoviruses (Euphorbia mosaic virus, EuMV and Cleome leaf crumple virus, ClLCrV). Alphasatellite was co-inoculated with their helper virus (EuMV and ClLCrV) into the Arabidopsis thaliana plants. Infected plant produced mild symptoms. Further they also investigated that ClLCrV trans-replicated the satellite more efficiently than the other help virus.

Tahir *et al.* (2010) cloned full-length begomovirus with its associated betasatellite from capsicum species. Begomovirus isolated from capsicum was a new species, thus given the name Pepper leaf curl Lahore virus (PepLCLV). It was thought to be a recombinant with the parents - PaLCuV and *Chili leaf curl virus*. Associated betasatellite was confirmed to be Chili leaf curl betasatellite (ChLCB). It was said to be a cognate of chili leaf curl betasatellite, which was first seen to be infecting potato. PepLCLV is among the increasing monopartite begomoviruses, which are seen to be accompanying a betasatellite, and among the numerous species infecting capsicum. In consideration of only been recognized in Pakistan, PepLCLC and ChLCB, both signify a geographically new begomovirus betasatellite complex.

Tahir *et al.* (2010) collected leaf samples with yellow mosaic symptoms from *Cucurbita pepo* plants from Lahore. Both genomic components; DNA-A and DNA-B were cloned and then sequenced. DNA-A showed 98.4% nucleotide identity, while DNA-B showed 89.6% nucleotide identity to an Indian Squash Leaf Curl China Virus (SLCCNV) strain.

Tahir *et al.* (2010) collected *M. charantia* leaf samples containing yellow vein symptoms, possibly begomoviral infection from areas near Lahore in 2004. Full-length clones of begomoviruses were obtained from infected leaves and sequenced. DNA A nucleotide sequence, showed 86.9 % nucleotide sequence identity to one of the isolates from Tomato Leaf curl New Delhi Virus (ToLCNDV), a distinct species, thus given the name Bitter gourd Yellow Vein Virus (BYVV). It was shown to have resulted from interspecific recombination between ToLCNDV and Tomato leaf curl Bangladesh virus (ToLCBDV). Nucleotide sequence identity of DNA B was 97.2 % to an Indian strain of Squash leaf curl China Virus.

He *et al.* (2009) isolated full-length begomovirus from allamanda plants from Guangdong, China. Virus was sequenced, comprised of 2755 nts with typical arrangement of begomovirus with an overlapping six ORFs. Sequence analysis was done, compared with the available reported sequences in the database. They found the maximum sequence identity (81.2%) to Tomato leaf curl Guangdong virus (ToLCGuV) isolate G2. Based on the threshold level for demarcation of species in the genus Begomovirus, virus isolated from allamanda plant was a distinct species, for which for which the name Allamanda leaf curl virus (AlLCV) was proposed.

Mubin *et al*. (2009) found multiple begomoviral components related to yellow vein disease of a common weed, *Digera arvensis.* A full-length clone of 2752 nucleotides was isolated that showed 98% sequence identity with *Cotton leaf curl Rajasthan virus.* Another clone of 1386 nucleotides that had the arrangement typical of alphasatellites was found to be correlated to Potato leaf curl alphasatellite (PotLCA). In addition to these, two betasatellite species (Ageratum yellow leaf curl betasatellite, and a recombinant of Tobacco leaf curl betasatellite and CLCuMB) were also isolated from the same sample. Such multiple begomoviral infection highlighted the importance of weeds as a reservoir for economically important begomoviruses.

Ohnesorge *et al.* (2009) reported interaction of CP protein with small HSP of transmission vector whitefly, *Bemisia tabaci*. To search for interacting proteins, CP was used to screen *Bemisia tabaci* cDNA library using the yeast two-hybrid assay. A member of the small HSP family (termed as BtHSP16) was identified. The interaction was further verified by an *in vitro* pull-down assay. Furthrmore, binding domain was mapped at the variable N-terminal portion of the CP, whereas full-length BtHSP16 is required for the interaction.

Samad *et al.* (2009) reported Tomato leaf curl Pakistan virus from Mentha samples. Total DNA was extracted from symptomatic and non-symptomatic mentha samples from different locations. The presence of a begomovirus was detected by PCR using begomovirus coat protein gene-specific primers. A product 771 bp in size was amplified successfully only from samples with symptoms. Amplicons were cloned into pCR-TOPOT/A. The selected clones were sequenced in both orientations. Sequence analysis showed the highest levels of sequence identity (93%) with the begomovirus Tomato leaf curl Pakistan virus [Pakistan-Rahim Yar Khan 1-2004] DQ116884.

Tahir *et al*. (2009a) reported a recombinant begomovirus infecting capsicum *spp.* in Pakistan and its association with Chili leaf curl betasatellite (ChLCB). It caused leaf curl disease in infected plants and so was named as Pepper leaf curl Lahore virus (PepLCLV). Recombination analysis showed presence of recombination event between *Chili leaf curl virus* (ChLCV) and PaLCuV, which ultimately gave birth to PepLCLV. Sequence analysis showed that the first 2160 nucleotides have maximum similarity of 92.8% with ChLCV-Pakistan and the remaining 586 nucleotides to have maximum similarity of 88.7% with PaLCuV-India. Its betasatellite showed 86.5% to 94.5% sequence identity with ChLCB, which is more than the species demarcation threshold of 78% for betasatellites, suggesting it to be an isolate of ChLCB. This type of disease caused by begomovirus-ChLCB pair in capsicum is first time identified in Pakistan.

Tahir *et al*. (2009b) reported for the first time the cognate partner of Tobacco leaf curl betasatellite which is a monopartite virus infecting *Pedilanthus tithymaloides*, in Pakistan. It caused leaf curl disease in infected plants and so was named as *Pedilanthus leaf curl virus* (PedLCV). Sequence analysis using Clustal X and MegAlign program and dendrogram was analyzed using Phylip and Treeview. Sequence similarity of 90.3% was obtained with virus infecting tomato crops in Pakistan, indicating them to be distinct strains of the same virus PedLCV, while it showed 86.3% identity with *Radish leaf curl virus* (RaLCV), designating it to be distinct species. Primer specific for B component, did not show any amplification, characterizing it to be a monopartite virus.

Duffy *et al.* (2008) estimated the evolution of single stranded TYLCV by calculating mean genomic substitution rate. Sequences submitted in the GenBank during year 1988-2006 were analyzed for evolution studies. Mean genomic substitution rate was found to be 2.88 X 10^{-4} nucleotide substitution per site per year (subs/site/year). Calculating the substitution rate of CP gene, a conserved region, gave a result of 4.63 X 10^{-4} subs/site/year. Intergenic region which shows more frequent substitution gave a value of ~1.56 X 10^{-3} subs/site/year.

Fauquet *et al.* (2008) determined and proposed geminiviral taxonomy. The distribution of pairwise sequence comparisons identified two peaks - one at 85-94% nt identity (corresponding to "strain" level, and one at 92-100% identity (corresponding to "variant" level) - determined for pairs of sequences below the species taxonomic level. They used the Clustal V algorithm (DNAStar MegAlign software) for the sequence alignment. Furthermore, they proposed guidelines for descriptors for each of these levels, to standardize nomenclature below the species level.

Haider *et al.* (2008) surveyed the vicinity of Punjab University-Pakistan for the presence of begomovirus in 2008. Some ornamental plants like *Vinca minor* L. exhibiting severe symptoms similarly produced by begomovirus. DNA from an infected and healthy *V.*

minor plant was isolated and diagnostic PCR was performed using coat protein amplification primers sets for the presence of virus. Amplified product (approx. 0.7 kb) was cloned, sequenced and analyzed. The nucleotide sequence shared 93% identity with previously reported *Pedilanthus leaf curl virus*. Altogether, this is the first report of PedLCV- a begomovirus infecting ornamental plant in Pakistan.

McLaughlin *et al*. (2008) identified virus from infected weeds and crops in Belize. Total 171 plants of tomato, squash, pepper, red kidney bean, string bean and weeds with leaf curling, yellowing, mottling, mosaic and stunted growth symptoms were collected. A modified version of the Dellaporta extraction protocol was used for total DNA extraction and samples were screened via nucleic acid hybridization using geminiviral probes. According to their research, *Pepper golden mosaic virus* was isolated from tomato and pepper while *Tomato mottle virus*-[Flo] was determined in tomato and sweet pepper only. Weed species *Euphorbia heterophylla* along with some samples of hot and sweet pepper were found infected by *Merremia mosaic virus.*

Ogawa *et al.* (2008) collected honeysuckle *(Lonicera japonica*) plants exhibiting yellow vein mosaic like begomoviral symptoms from Nara Prefecture and Ibaraki Prefecture, Japan. They isolated two viruses (Nara Virus-1 and Nara Virus-2), a betasatellite molecule (DNAbeta-Nara) and defective DNAs from samples of Nara Prefecture, and one begomovirus (Ibaraki virus) and a betasatellite DNA (DNAbeta-Ibaraki) were cloned from Ibaraki Prefecture. Nara Virus-1 shared highest nucleotide identity (94%) with Nara Virus-2 and (90%) with Ibaraki virus. They found that Nara virus-1 and Nara virus-2 had a hybrid genome of a HYVMV and TbLCJV, for which HYVMV-Nara1 and HYVMV-Nara2 names were proposed. Furthermore, Ibaraki virus was determined as a strain of Tobacco leaf curl Japan virus (TbLCJV), determined hereafter as TbLCJV-Hs[Iba]. They observed that TbLCJV-Hs[Iba] or HYVMV-Nara2 partial dimer could produce yellowing, leaf crinkling and stunting symptoms

when tomato plant was agro-inoculated but in presence of DNA beta more severe symptoms were appeared on tomato plants. Also, further they reported that the satellite associated with virus possessing iterons of their helper virus which were not previously reported from Japan, and this complex was previously responsible of *Tomato yellow dwarf virus* while honeysuckle serve as reservoir for such viruses' recombination.

Haider *et al.* (2007) collected leaf samples from three different plants (*Zinnia elegans, Solanum nigrum* and *Ageratum conyzoides*) showing possible begomoviral infection. The samples were subjected to PCR using degenerate primers specific for coat protein. The coat protein gene was amplified, cloned and sequenced. MSA and a phylogenetic tree was constructed. The isolates of coat protein were named as Zinnia leaf curl virus (ZLCV) and Ageratum yellow vein virus- Pakistan (AYVV-PK). These viruses were seen to be close relatives of Tomato leaf curl virus-India, Indian cassava mosaic virus and CLCV. One isolate named as Solanum yellow leaf curl virus was mechanically infectious strain of Tomato leaf curl virus-India, while ZLCV was seen to be a mechanically infectious strain of AYVV-PK.

García-Andrés *et al.* (2007) studied recombination events that occur frequently in begomoviruses. For this, tomato was chosen as a model plant and it was infected with Tomato yellow leaf curl disease, TYLCSV and TYLCV. Recombination events were assessed after numerous post inoculations. Genetic exchange was observed in a substantial viral population. After recombination analysis, it was seen that recombination occurred in the intergenic region (IR) and REn ORF, while a second region identified somewhere between the 3'- terminus of Rep ORF also showed genetic exchange. The identification of a second recombination site would not only be helpful in intermittent DNA replication, but the replication complex would keep on swapping between homologous regions of DNA template.

Kon *et al.* (2007) reported a recombinant of Tomato leaf curl Java virus and Ageratum yellow vein virus-Java isolated from *Ageratum conyzoides* L. plants showing yellow vein

symptoms. DNA A of 2747 and betasatellite of 1360 base pairs was amplified, cloned and sequenced. Sequence analysis confirmed that virus is 91% similar to Tomato leaf curl Java virus. Whereas, ORFs C1 and C4 shared maximum identity of 91% and 95% with Ageratum yellow vein virus-Java respectively. Therefore, it was concluded that virus is a recombinant having sequence of Ageratum yellow vein virus- Java from 2389–2692. Infection with only ToLCJV-Ageratum induced no symptoms, however, when coinfected with the betasatellite produced yellow vein symptoms. *Nicotiana benthamiana* plants were infected with the three viruses ToLCJV, AYVV-[Java] and ToLCJV-Ageratum showed similar symptoms which suggested exchange of C4 fragment which functions as pathogenicity determinant.

Saeed *et al.* (2007) showed association of monopartite-associated DNA betasatellite, implicated in CLCuD, with a bipartite species *Tomato leaf curl New Delhi virus* (ToLCNDV). According to research, DNA betasatellite has similar functions as those provided by DNA B component; therefore, it can replace DNA B component. Like DNA B encoded proteins, betasatellite encoded βC1 can provide movement functions. Increased accumulation of DNA A along with systemic infection was observed by co-infection of DNA A and DNA betasatellite. No symptoms were observed when only DNA A was inoculated. To confirm the role of βC1, co-inoculation of DNA A and βC1 constructs was carried out; this showed systemic infection and DNA A accumulation; however, no symptoms were observed when co-inoculated with mutant βC1 construct. This provides evidence that begomoviruses, whether mono- or bipartite, can get associated with betasatellites, resulting in emergence of new infections.

Singh *et al.* (2007) were the first ones to observe radish leaf curl disease in India. Plants showing typical begomovirus symptoms were collected. Begomovirus twinned icosahedral particles were observed under electron microscope and it was further confirmed by PCR using degenerate primers typical for DNA-A and primers for DNA-beta was also used. Sequence

analysis showed DNA-A to be of 2756 nucleotides, a maximum identity (87.7 percent) with Tomato leaf curl Bangladesh virus and DNA-beta was seen to be of 1358 nucleotides, a maximum identity (84.9 percent) with Tobacco leaf curl virus. Based on identity less than 89 percent, this virus was said to be a novel begomovirus species given the name Radish leaf curl virus.

Vadivukarasi *et al.* (2007) compared replication initiator protein (Rep) with rolling circle replication initiators. It was seen that Rep protein was conserved among each group of organisms when compared. Through recombination analysis, genetic exchange between Rep gene and the intergenic region was discovered. Consequences of these recombinations and an increasingly great number of host range was discussed.

Amin *et al.* (2006) indicated that the satellite associated with CLCuD was recombinant. They investigated the satellite from both archival (pre-2001) and cotton samples recently collected near Vehari (Pakistan). They found the satellite associated with resistant breaking was recombinant-containing major portion of CLCuD betasatellite with minor fragment of ToLCB. This recombinant satellite molecule was isolated in the cotton samples collected recently, not from samples collected before 2001. They concluded that the DNA satellite was recently mobilized into the cotton, likely from tomato and that recombination plays a driving force towards the evolution of these satellites.

Bull *et al*. (2006) studied the genetic variability and distribution of *Cassava mosaic viruses* in Kenya. Cassava mosaic disease is the most common cause of economic loss in Africa. These cassava infecting begomoviruses have greater potential to undergo genetic diversity and hence can cause great epidemics. Molecular analysis showed a wide range of variability among *Cassava mosaic viruses*, characterizing them into distinct groups. These groups were further analyzed through sequence analysis. This resulted in the identification of *East African Cassava mosaic virus* and *East African Cassava mosaic Zanzibar virus* isolates

and a new species whose name *East African Cassava mosaic Kenya* virus was suggested. These isolates showed specific geographic distribution, indicating restricted movement of these viruses from their native locations and the significance of these viruses from surrounding countries which may play pivotal role in aggravating their diversity.

Ogbe *et al.* (2006) conducted a survey, in which they visited 418 farms. 48 percent of the cassava from these farms was severely and semi severely infected, while 52 percent cassava showed mild symptoms. Of the total leaf samples, 1397 that were collected, symptoms were seen in 1106 samples. When the infected samples were subjected to PCR, 74.1 percent samples were infected with *African cassava mosaic virus* (ACMV), infection to EACMV was seen in 0.3 percent samples while 24.4 percent samples were seen to be containing mixed infection. ACMV was found majorly in single and mixed infection samples. Two new variants of ACMV were identified from fields. They also discovered ACMV and EACMV in *Senna occidentalis* (L.) and in a weed (*Combertum confertum* Lams). Altogether in this report, the two plants were identified as two new natural hosts of these viruses.

Tahir *et al*. (2006) observed mild leaf symptoms on *Duranta repens* (pigeon berry) grown as an ornamental plant. Total DNA was extracted from both healthy and infected leaf samples. A product of approx. 1.5 kb was obtained using a pair of primers designed by Mansoor *et al*., 1998. Sequence analysis of this segment shared 91% identities with *Croton yellow vein mosaic virus*. DNA B of approx. 2.8kb was also amplified which had 94% nucleotide sequence identity with *Tomato leaf curl New Delhi virus* segment B. This study revealed the presence of a bipartite begomovirus inducing leaf curl disease in *D. repens.*

Cui *et al.* (2005) demonstrated the binding activity and expression analysis of βC1 protein. Using electrophoretic mobility shift assays and UV- crosslinking experiments showed that βC1 is capable 0f binding both single stranded and double stranded DNA irrespective of any specificity of sequence or size. Nicotiana benthamiana plants were co-inoculated with GFP

transgene along with TYLCCNV-Y10 and β. Patch agroinfiltration assay showed suppression of GFP transgene silencing as GFP mRNA accumulation was observed in leaf patches of N. benthamiana plants. Immunogold labeling provided evidence of nuclear localization of βC1 protein in infected *N. benthamiana* plants. Mutations in nuclear localization sequence of β, showed failure to induce symptoms, suppression of RNA silencing and nuclear accumulation. It was concluded that βC1 is required for symptom induction and RNA silencing suppression.

Haider *et al.* (2005) detected begomoviruses from 5 different host plants, characterize them and built a relationship among them. The studies revealed that only SYLCV was mechanically transmissible to *Nicotiana benthamiana* while all viruses were graft transmissible. Serological studies showed their relevance to *African cassava mosaic virus* (ACMV). Indirect ELISA results clearly differentiated three viruses Ageratum yellow vein virus-Pakistan (AgYVV-PK), Solanum yellow leaf curl virus (SYLCV) and Eclipta prostrata yellow vein virus (EPYVV). The other two were confirmed as a strain. It was concluded that AgYVV and SYLCV were closely related to CLCuV than ACMV or EPYVV.

Khan *et al*. (2005) observed yellow vein net disease of *Calendula* onseveral ornamental plants in India. Whiteflies were used to experimentally transform naturally infected *Candula* disease to healthy seedlings. Total DNA was extracted from both symptomatic and non symptomatic leaf samples. PCR amplification using Coat protein region of *Tomato leaf curl New Delhi virus* (ToLCNDV; Hallan, 1998) was followed by cloning and sequence analysis.To confirm the authenticity of PCR results, cross-hybridization was performed using ToLCNDV DNA-A probe. In this study, the isolated virus showed 95, 94 and 93% of nucleotide sequence identities with *Tobacco curly shoot virus* (AF240675), *Ageratum enation virus* (AJ437618) and *Tomato leaf curl Bangladesh virus* (AF188481), respectively. This was the first report of *Calendula officinalis* infected by a begomovirus.

Tahir and Haider (2005) collected infected leaf samples showing yellow blotch symptoms in Bitter gourd from the areas of Lahore. The presence of begomovirus was confirmed through PCR amplification by using primers specific for coat protein sequence. The amplified product was cloned and then sequenced. Nucleotide sequence identity of 95 percent was seen with Tomato leaf curl New Delhi virus (ToLCNDV) revealing it to be a new strain ToLCNDV found for the first time in bitter gourd.

Briddon *et al.* (2004) observed the diversity of DNA-β in east and south-East Asia. CLCV and OLCV from Pakistan and AYVV from Singapore were isolated. Seventeen (17) betasatellites were isolated, cloned and sequenced for diversity studies. Presence of DNA β was confirmed in almost all the virus complxes except for just two begomoviruses from Far East Asia. It was then concluded that β satellites are conserved in both their sequence and architecture.

Gutierrez *et al.* (2004) demonstrated geminivirus proteins interacting with host plant cell cycle proteins and DNA replication factors. To determine these interactions yeast two hybrid assays were utilized. The study showed interaction of wheat dwarf virus RepA protein interacting with host plant retinoblastoma related (RBR) protein. Moreover, it was confirmed that viral Rep protein is not involved in RBR interaction, but deletion of C-terminal of the Rep protein imparts the same function as that of RepA protein suggesting interaction of this domain with RBR binding motif (LXCXE), which is also confirmed through secondary structure analysis.

Hussain *et al.* (2004) studied alternative hosts of CLCuD. For this, samples of chili were collected from Vehari and Khanewal. Total DNA was isolated from symptomatic leaves. A sample showing DNA betasatellite containing CLCuD was also included. Begomovirus infection was confirmed through PCR. Further confirmation was done by Southern blot hybridization. CLCuD DNA beta probes were used in southern blots. Strong signal was yielded

which showed the presence of CLCuD in Chili. This makes chili as an alternative host of CLCuD.

Wang *et al*. (2004) compared six papaya samples all showing downward leaf curl symptoms. Comparative study revealed the presence of two different species, *Papaya leaf curl Guangdong virus* (PaLCuGDV) infecting group II isolates and *Papaya leaf curl China virus* (PaLCuCNV) infecting group I isolates. According to amino acid sequence identity of Coat Protein (CP), G2 and GD2 showed 97.7% and 94.2% identity with *Pepper leaf curl virus* (PepLCV) respectively. This amino acid comparison analysis revealed that these viruses have common ancestor for coat protein sequence.

Briddon *et al.* (2003) verified DNA-beta as a symptom inducing component, which is in complex with monopartite begomoviruses. These satellite molecules were determined to be inducing symptoms in infected AYVV from Singapore and CLCuMuV from Pakistan. 26 samples of betasatellites were isolated, cloned and sequenced. Sequence analysis showed high levels of conservation for a single open reading frame, for an adenine rich region and for a satellite conserved region (SCR). This was the region containing the hairpin loop sequence. Phylogenetic analysis revealed two groups of satellite molecules, one that had originated from Malvaceae, and the other group of plants that had originated from Solanaceae and Compositae. DNA molecules showed close connections between both the host and geographic origin. This study showed the adaptation of begomoviruses with their associated satellite molecules.

Stonor *et al*. (2003) studied geminiviruses which are transmitted by whiteflies. DNA probes to detect new species of geminiviruses were used. A single species, Tomato leaf curl virus was acquired by using these probes which was found to occur in Australia. This area was found to be colonized by Australasian biotype of whitefly species, *Bemesia tabaci*. A new biotype in coastal regions of Queensland was found to be 150 kilo meters far from TLCV-infected area. After sample collection from the infected area and their sequence analysis, a

variation of 14 percent in the total sequence was seen in this region plus new viral iterons were also observed.

Mansoor *et al.* (2003) stated DNA B as an essential component of bipartite begomoviruses. Its replication depends on the Rep protein encoded by DNA A. the two DNA molecules interact by short motifs called iterons which are Rep binding sites. Further it was revealed that DNA β isolated from CLCuMV, a monopartite virus, lacks iterons but is trans-replicated by DNA A of the virus. Moreover, their interaction was confirmed by grafting a monopartite CLCuD sample in healthy plant, susceptible plant and via whitefly in healthy plant. All three plants showed symptoms as that of bipartite viruses, confirming interaction of β molecules with viruses.

Jose and Usha. (2003) demonstrated the interaction of DNA betasatellite with *Bhendi yellow vein mosaic virus* (BYVMV), causing a major threat to okra plant in India. Their research study comprised of full-length DNA amplification of DNA A and betasatellite molecule followed by cloning, sequence analysis partial repeat constructs of BYVMV genome and DNA betasatellite and insertion into binary vector, for agro-inoculation. They found that agro-inoculation of BYVMV alone, produces mild leaf curl symptoms whereas yellow vein mosaic symptoms typical of bhendi yellow vein mosaic disease (BYVMD) were produced when BYVMV was inoculated along with DNA beta.

Briddon *et al*. (2002) found an extremely conserved region located upstream of the hairpin loop. They intended a set of abutting primers from this conserved region that amplify full-length DNA beta component of approx. 1350 nts from infected plants. This single stranded DNA molecule needs a helper begomovirus for insect transmission and replication in host plant cells.

Idris *et al.* (2002) worked on samples of hollyhock and okra with leaf crumpling and curling symptoms respectively. Full-length viruses with monopartite characteristics were cloned from both the samples named as HLCrV and OkLCV showing identity to old world begomoviruses such as CLCuV-SD having 84% and 95% identities. Individual alignments showed HLCrV to be 98% identical to AREV and OkLCV being 99% identical to OkEV. Alignments showed identity to viruses having enation symptoms and this can be explained by the amplification of satellite DNAs from both the viruses ranging from (741-1350 nucleotides).

Monci *et al*. (2002) studied a recombinant genome of TYLCSV and TYLCV. Both viruses are distinct species and found in southern Spain. An infectious clone of this recombinant was made and characterized. Recombination analysis showed recombination to have occurred in the IR and at the 3' end of the replication enhancer protein (Ren). Recombinant was seen to be much more dangerous as it had a wide host range, it could efficiently be transmitted by whitefly Bemesia tabaci and it was becoming prevalent in nature.

Xie *et al*. (2002) isolated new begomovirus species, named *Tobacco curly shoot virus* (TCSV) which was causing shoot curling and retarded growth in Tobacco plants in Yunnan. To confirm the characteristics typical of begomovirus, basic diagnostic tests such as whitefly transmission test, virion morphology using Immunosorbent Electron Microscopy (ISEM) and serological tests using Triple Antibody Sandwich-ELISA (TAS-ELISA) were performed. Total DNA extraction was followed by amplification, cloning and sequence analysis. Sequence analysis of complete DNA A component showed 85% nucleotide sequence identity with both *Papaya leaf curl virus* (PapLCV-IN) and *Tomato leaf curl virus* (ToLCV-IN) from India. In contrast, sequence analysis based on coat protein showed 98% amino acid sequence identity with 72b isolate of CLCV from Pakistan. TCSV showed close relationship with Asian begomoviruses, with different parts of genome similar to different viruses, showing different ancestors for different parts.

Rojas *et al.* (2001) analyzed proteins involved in the movement of TYLCV and suggested that CP and C4 are homologs of βC1. Monopartite begomoviruses have CP and C4 proteins which mediate movement of virus through mesophyll or epidermal plasmodesmata. Unlike bipartite begomoviruses movement protein βC1 which mediates cell to cell movement, CP and C4 have narrow capacity of doing that. They showed it via microinjection of Escherichia coli expressed proteins along with GFP fusion proteins expressed transiently. Expression of the proteins were examined at different growth stages of tomato plants resulting in the observation that only phloem cells of shoot apical, stem, leaves and floral tissue were infected.

Saunders *et al.* (2001) demonstrated the importance of β satellite in inducing symptoms (a recombinant, having sequences from both DNA-A and β) by examining infection of AYVV in *Ageratum conyzoides* imparting yellow vein symptoms. They observed that DNA-A alone cannot induce symptoms and that satellite DNA is responsible for imparting symptoms during viral infection. For that they made a chimera similar to recDNA-Aβ17 and inoculate it into Nicotiana glutinosa. Inoculation resulted in mild symptoms but similar to that seen in AYVV infection. Based on above mentioned observations they concluded that complex of recombinant DNA satellite molecules with viruses are responsible for symptom modulation in plants.

Briddon *et al.* (2001) demonstrated that β satellite is necessary for inducing symptoms typical to that of begomoviruses. An infectious clone of CLCuV was developed and inoculated in cotton plants. There was an absence of typical vein swelling, vein darkening, leaf curling and enation symptoms from infected plants. They also described that complex of virus/ DNA-A/ β resulting in recombinant satellite having fragments from both DNA-A and β is totally a different molecule having the same function in symptom modulation.

Xie *et al.* (2001) characterized a new species of begomoviruses, Tobacco curly shoot virus. Virus was isolated from tobacco plants showing curly shoot symptoms from Yunnan. DNA A of the virus was amplified cloned and sequenced. The sequencing showed virus to be

of 2746 nucleotides having 6 ORFs. Similarity analysis showed 85% sequence identity with Tomato leaf curl virus from India. Coat protein showed 98% sequence identity with CLCV from Pakistan.

Pant *et al*. (2001) characterized a bipartite *Mungbean yellow mosaic virus* (MYMV-Bg) from blackgram varieties in India. The complete nucleotide sequence was compared with other geminiviruses. In this study a unique genome organization, sequence divergence between Common Region (CR) of both genomic components was observed. To study viral DNA replication mechanism, Rep protein was over expressed in bacterial host system. Their study demonstrates that ATP is directly related with the enhanced cleavage activity of Rep. As the concentration of ATP increases the rate of cleavage increases.

Briddon and Markham (2000) documented the reported of CLCuD, infecting Cotton in Pakistan's. CLCuD was initially reported in Punjab in 1967, later in 1985. By 1990, CLCuD became a major epidemic in Pakistan. It has now spread to the surrounding regions such as West and South of Pakistan. Reports of CLCuD were also found from some regions of India. Symptoms observed of CLCuD were vein darkening, swelling, a leaf like out-growth at the underside of the leaf. It was seen that CLCuD was spread due to a single species of whitefly known as *Bemesia tabaci*. The virus was ascribed to be begomovirus from the family geminivirus. But later in 1999, after studying the etiology of the disease it was seen to be much more complex than presumed.

Briddon *et al*. (2000) isolated CLCV and found it equal to DNA A of a bipartite begomovirus. When the clones were infected to cotton and Nicotiana benthamiana, leaf curling and yellowing was observed. Their findings suggest that CLCuV may contribute in the transmission of the disease.

Lotrakul *et al*. (2000) molecularly characterized a distinct bipartite begomovirus from an infected ornamental plant *Dicliptera sexangularis* (sixangle foldwing), found in Lee County, Florida. The begomoviral infection was confirmed by whiteflies, mechanical and graft inoculations followed by PCR amplification and southern blot hybridization. Both genomic components, DNA A and DNA B were successfully extracted, cloned and transformed into DH5α cells. The virus was named *Dicliptera yellow mottle virus* (DiYMV). Phylogenetic analysis based on coat protein, replication initiation protein and nuclear shuttle protein indicated a close phylogenatic relationship of *Dicliptera yellow mottle virus* (DiYMV) with the new world begomoviruses.In the sequence comparison studies, different percentage identities were obtained for different ORFs of DiYMV. On the basis of Rep and NSP ORFs, DiYMV shows close relationship with BGMV type II isolates; however, based on of derived AC4 amino acids and common region, it shows close relationship with PYMV-VE. This may be, in part, due to the possible occurrence of recombination event among new world begomoviruses, specifically belonging to Caribbean Basin. The other possibility can be a sudden shift of divergence away from PYMV-like virus, in the evolutionary process of DiYMV.

Sanz *et al.* (2000) suggested that multiple infections occur in cotton and other plants of malvaceous species along with recombination. It was identified that cotton plants having begomovirus infection in case of CLCuD encompass sequences from two or more begomoviruses. It was also observed that cotton plants contained sequences from non-malvaceous species and some non-malvaceous plant contained sequence of malvaceous species. Chimeric sequences were also seen, sequences having elements from malvaceous/non-malvaceous plants. A recombination site was also present at origin of replication of viruses. Hence it was concluded that recombination and multiple infection plays an important role in diversification of begomoviruses.

3. MATERIALS AND METHODS

3.1. Sample collection

During a survey of cotton farms, plants exhibiting begomoviral symptoms, including leaf curling, dark green vein thickening, enation on lower side of the leaf, were observed in Punjab province of Pakistan in 2010-13. Young leaves from suspected plants were collected in different locations in Punjab province. These young leaves were kept in a zip bag on ice during transportation. Symptomatic plants were also photographed with camera.

3.2. Total DNA Extraction

Total DNA was extracted from leaf tissue with the CTAB method (Doyle and Doyle 1990). Approximately, 1gm of leaf tissue was crushed in pestle and mortar to fine powder in the presence of liquid nitrogen. To the fine powder, 25ml extraction buffer [100 mM Tris, 1.4 M NaCl, 20 mM EDTA (pH 8), 2% CTAB, 0.2% β-mercaptoethanol] was added in the mortar and homogenized. The slurry was then shifted to a 50ml sterilized falcon tube and mixed by inverted several times before incubation at 62°C for 30-45 min in a water bath shaker. The mixture was allowed to cool in the fume hood for short time up to 3-5 min. Equal volume (25 ml) of chloroform: Iso-amyl solution was added and mixed by inverting gently. The slurry was divided into two equal volumes in new 50ml falcon tubes and centrifuged (Eppendorf Centrifuge 5804R) at 9000 rpm for 10 min at room temperature to separate the phases. The upper aqueous phase (up to 20 ml) was transferred to a new sterile 50 ml falcon tube. An equal volume of chilled isopropanol was added to the supernatant and mixed by inverting gently. The tubes were incubated over night at room temperature and then

centrifuged at 9000 rpm for 10 min. The supernatant was discarded and the pellet washed with 1ml cold wash buffer [76% ethanol and 10 mM ammonium acetate] and incubated for 20 min at room temperature. The tube was centrifuged at 9000 rmp for 5 min and the supernatant decanted and the pellet was dried at 37°C for 30-60 min. Finally, the pellet was re- suspended in a 1ml TE buffer [10 mM Tris- HCl and 1 mM EDTA] and the DNA stored at -20°C.

3.3. Quantification of DNA

DNA concentration was analyzed by visualizing on 1% [w/v] agarose gel in TAE buffer and by a spectrophotometer. Five microliters of total DNA was loaded on 1% agarose gel along with a high range DNA marker (GeneRuler^{TM}, 10Kb DNA Ladder, Fermentas) and quantified visually using UV-light while comparing with the marker bands.

The concentration of dsDNA was calculated on a spectrophotometr with absorbance at 260 nm wavelength. The TE buffer was used as blank. DNA concentration was determiend by following formulae:

DNA concentration (ug/ml) = E × OD_{260} × dilution factor

(Whereas, E is an extinction co-efficient = 50 for dsDNA)

3.4. Amplification of DNA

3.4.1. PCR amplification

For amplification of DNA by PCR, a final volume 50 µl reaction mixture containing 5µL 1X *Taq* polymerase buffer (Thermo Fisher Scientific), 1.5-2 mM $MgCl_2$, 100-200 µM dNTPs, 2.5 µM of each primer (Table 3.1), 2.5U of *Taq*

polymerase (Thermo Fisher Scientific) and approximately 100ng of DNA template were prepared in a 0.25 ml or 0.5 ml PCR tube. The reaction mixture was incubated in a thermal cycler (Applied Biosystems, ProFlex). The thermal program was as a preheat treatment of 94 °C for 1 min followed by 30 cycles of 94 °C for 1 min, 50-58 °C for 1 min, 72 °C for varying times (dependent upon the length of the fragment to be amplified; typically, 1 min per 1000 nucleotides to be amplified), followed by a final incubation of 10 min at 72°C. The amplicons were then visualized onto 1% agarose gel in 1x TAE buffer.

Table 3.1. Names, sequences and brief description of primers used in this study

Primer Name	Sequence (5'……3')	Used for
WTG-F **WTG-R**	5'GATTGTACGCGTCTAATTTGAAYBGG 5'TANACGCGTGGCTTACATGGGCCTDT	Amplification of DNA-A of all Whitefly transmitted viruses
CP-F **CP-R**	5'-ATGHSVAAGCGWMCMGSMGATAT-3' 5'-TTAATTBVHDAYHSHRTCATARAARTA-3'	Amplification of DNA-A Coat Protein
CLCV1 **CLCV2**	5'-CCGTGCTGCTGCCCCCATTGTCCGCGTCAC-3' 5'-CTGCCACAACCATGGATTCACGCACAGGG-3'	Amplification of DNA-A
BurX-F **BurX-R**	5'-*CTCGAG*AGTGTCCCCGTCCTTGTCG-3' 5'-*CTCGAG*TGGGGAGAGTTTCAGATCG-3'	Amplification of DNA-A
BurN-F **BurN-R**	5'-*CCATGG*TTGTGGCAGTTGATTGACAGATAC-3' 5'-*CCATGG*ATTCACGCACAGGGGAACCC-3'	Amplification of DNA-A
KTB-F **KTB-R**	5'-*CTGCAG*AGGTCACCTTGTCATTTCCTTC-3' 5'-*CTGCAG*CATCATTTGTGAGCGCATATTC-3'	Amplification of DNA-B
MP-F **MP-R**	5'-CACCATGGCAATAGGAAATGATGGTATGGG-3' 5'-AAGGATCCTCTTATTTTTTGAATAAATTTGG-3'	Amplification of DNA-B Movement Protein

Beta-01 **Beta-02**	5'-*GGTACC*ACTACGCTACGCAGCAGCC-3' 5'-*GGTACC*TACCCTCCCAGGGGTACAC-3'	Amplification of betasatellite
Alpha-F **Alpha-R**	5'-*AAGCTT*AGAGGAAACTAGGGTTTC-3' 5'-*AAGCTT*TTCATACARTARTCNCRDG-3'	Amplification of alphasatellite
CP27NF **CPXR**	5'-TT*CCATGGC*TAAGCGACCAGCAGACATC-3' 5'-TG*CTCGAG*ATTTGTCACGGAATCAT-3'	Cloning of CLCuKoV-Bu Coat Protein into pET Vector

3.4.2. Rolling- circle amplification

For amplification of circular DNA molecules, rolling-circle amplification was used. A reaction mixture of 20 µL containing 100 to 200 ng of total DNA (extracted from infected plant samples) used as a template, 50 µM random hexamer primers, 2 µL 10Xφ29 DNA polymerase reaction buffer (330 mM Tris-acetate [pH 7.9 at 37 °C], 100mM magnesium acetate, 660 mM potassium acetate, 1% [v/v] Tween 20, 10mM DTT) was prepared and incubated at 94 °C for 3 min for denaturation of dsDNA in PCR machine. The mixture was cooled at room temperature and mixed using 1mM dNTPs and 5-7 units of φ29 DNA polymerase and incubated at 30°C for 18-20 hours. Finally, the φ29 DNA polymerase in RCA was inactivated at 65 °C for 10 min. From RCA 5µL was then analyzed on 1% agarose gel.

3.4.3. DNA isolation from agarose gels

Amplified viral DNA was extracted using the Silica Bead DNA Gel Extraction Kit (Thermo Fisher Scientific) as per manufacturer's instructions. Specific band of amplified DNA fragment of required size from 1% agarose gel was cut and slice placed in an Eppendorf tube, Gel slice was incubated with binding buffer solution (6 M sodium iodide solution) at 3 times the weight of the gel slice at 55°C for 5-10 min or until the

gel was completely dissolved. Seven microliters of silica suspension was added to the DNA binding mixture and vortexed for a short time before incubation at 55°C for 5 min. The mixture was centrifuged at 12000 rpm for 30 sec. and the supernatant decanted. The pellet was washed three time with 500µl of 5% cold wash buffer. Finally, after the pellet got dried was dissolved in 25-30 µl of TE buffer.

3.5. Cloning of Amplified DNA

3.5.1. Cloning of PCR Product

PCR amplified DNA fragments purified from agarose gel were quantified and ligated into the vector pTZ57R/T (InsT/Aclone^{TM} PCR Cloning Kit, Thermo Fisher Scientific) according to manufacturer's instruction or into the binary vector pGreenII / pGreen 0029 for the partial dimers.

A ligation reaction mixture of 30 µl containing ~400 ng DNA fragment, 6µl 5X ligation buffer, 3 µl pTZ57R/T vector and 5 units of T4 DNA Ligase, was prepared in a 1.5 ml microcentrifuge tube and incubated at 4 °C overnight. On the next day the incubated ligation mixture was transferred into competent *Escherichia coli* cells by heat-shock method or electric method. Transformed cells were spread on solid LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl and 1.5% agar) plates containing 100 µg/mL ampicillin, spread with 20 µl X-Gal (50 mg/ mL) and 100 µl IPTG (24 mg/ ml) and incubated at 37 °C for 16 hours. Next day, the single white colonies were picked and cultured in 5-10 ml aliquots of LB liquid media in an autoclaved test tubes or small size flasks and grown at 37 °C for 16 hours.

3.5.2. Small-scale plasmid preparation

Alkaline lysis method was used for plasmid isolation with modifications made by Birnboim and Doly (1979) and Ish-Horowicz and Burke (1981). *E. coli* colonies

were picked from agar plates using a sterile loop or tip. A single bacterial colony transformant containing the recombinant plasmid was inoculated into 5-10 ml of LB medium (1% tryptone, 0.5% yeast extract and 1% NaCl) containing 100 µg/ml of ampicillin or 50 µg/ml of kanamycin in a 100 ml flask. The culture was incubated at 37 °C for 16 hours with shaking. On the following day, the overnight culture was centrifuged at 14,000 rpm (Eppendorf centrifuge, 5415R) for 1 min and the supernatant discarded. The pellet was resuspended in 100 µl of ice-cold Solution-I (50 mM glucose, 25 mM Tris-Cl [pH 8.0] and 10 mM EDTA [pH 8.0]) (Heidolph, Model D-91126) and 200 µl of freshly prepared Solution-II (0.2 N NaOH and 1% SDS) was added. The mixture was mixed by inverting 5-6 times and 150 µl of ice cold Solution-III (3M potassium acetate [pH 8.0]) added. The mixture was inverted 5-6 times. The tube was incubated for 5 min on ice and centrifuged at maximum speed for 10 min. The supernatant was transferred to a new 1.5 ml Eppendorf tube. An equal volume of Phenol: Chloroform (1:1) was added and mixed by vortexing or pipetting. After centrifugation at 12,000 rpm for 2 min at 4 °C, the supernatant was transferred to a new tube to which 2 volumes of 100 % ethanol were added to precipitate the nucleic acids. The tube was left for two min at room temperature. The pellet was collected by centrifugation at 12,000 rpm for 5 min at 4 °C. The supernatant was discarded and pellet was rinsed with 500-1000 µl 70% ethanol and centrifuged for short time. The ethanol was discarded and pellet was air dried for 10 min. The DNA was dissolved in 40-60 µl of TE buffer containing RNase and stored at -20 °C.

3.5.3. Intermediate-scale plasmid preparation

Recombinant plasmids were isolated from 5-10 ml overnight culture using QIA prep spin miniprep kit (QIAGEN, Hilden, Germany) or Gene JET Plasmid Miniprep Kit (Thermo Fisher Scientific). A single colony was inoculated to 5 ml LB medium

containing 100 µg/ml ampicillin and grown at 37 °C for 14-16 hours on an orbital shaker. The overnight culture was centrifuged at 12,000 rpm for 1 min. This step was repeated 2-3 times to recover enough pellet for plasmid isolation according to manufacturer's instructions. The pellet was re-suspended in 250 µl Buffer P1, followed by addition of 250 µl Buffer P2 and then mixed by inverting 5-6 times until the solution became viscous and less turbid. This was followed by addition of 350 µl Buffer N3 and mixing by inverting 5-6 times. The mixture was centrifuged at 13,000 rpm for 10 min at room temperature. The supernatant was transferred to QIA prep spin column followed by centrifugation for 1 min and the flow through was discarded. To remove nuclease activity, 0.5 ml of Buffer PB was added to the column followed by addition of 0.75 ml of Buffer PE and centrifugation for 1 min at room temperature. The flow through was discarded and the column centrifuged for 1 min to remove the residual of wash buffer. To elute DNA, the column was placed in a 1.5 ml sterilized Eppendorf tube and 50 µl of Buffer EB was added to the column and incubated at room temperature for 1 min and centrifuged at 13,000 rpm for 1 min to collect DNA. The quantity and quality of purified DNA was measured by agarose gel or using a spectrophotometer (Nanodrop 1000, Thermo Fisher Scientific).

3.5.4. Large-scale plasmid preparation

To obtained a large amount of purified DNA, 100 ml of LB medium containing the appropriate antibiotic 100 µg/ml ampicillin or 50 µg/ml kanamycin was inoculated with 2-3 beads of glycerol stock and allowed to grow overnight at 37 °C in shaker . On the following day culture was transferred to 50 ml falcon tubes and centrifuged at 10,000 rpm for 5 min at 4 °C to harvest the bacterial cells. Pellet was re-suspended completely into 10 ml of Buffer P1. 10 ml of Buffer P2 was added and the cape on falcon tube was closed tightly. The tube was inverted 5-6 times before incubated at

room temperature for 5 min. Followed by addition of 10 ml of chilled Buffer 3, mixed thoroughly by vigorously inverting 4-6 times and incubated on ice for 20 min. The solution was filtered and collected in a 50 ml falcon tube. QIAGEN-tip 500 was equilibrated by 10 ml Buffer QBT and allowed to the column to empty by gravity flow. Supernatant was applied to the column and flow through was discarded. The QIAtip column was washed twice by adding 30 ml Buffer QC. To elute DNA, column was placed in a sterilized tube and 15 ml of Buffer QF was added and the flow through was collected. DNA was precipitated by adding 10.5 ml isopropanol at room temperature and mixed gently. The mixture was aliquoted into 2 ml or 1.5 ml Eppendorf tubes and centrifuged at 14,000 rpm for 30 min at 4 °C. Pellet was washed with 70% ethanol and collected in one tube. The tube was centrifuged at 14,000 rpm for 10 min and pellet was air dried, re-suspended into TE Buffer. The quantity and quality of purified DNA was determined by agarose gel or using a spectrophotometer (Nanodrop 1000, Thermo Fisher Scientific).

3.6. Microbiology techniques

3.6.1. Preparation of Competent of *Escherichia coli* ($DH_5\alpha$) cells

The procedure of Cohen *et al.* (1972) was followed to prepare competent cells. A single colony from a freshly grown plate of *E.coli* cells was transferred into 5 ml LB low medium (1% tryptone , 0.5% yeast extract, 0.5% NaCl) and incubated overnight (14-16 hours) at 37 °C with shaking. 100 ml LB medium was refreshed with 2 ml of the overnight culture and the flask was incubated at 37 °C for 2 hours in a shaking incubator. The culture was transferred to 50 ml falcon tube and incubated on ice until chilled. Centrifugation at 4,000 rpm for 10 min at 4 °C was used to harvest the cells. The supernatant was discared and the pellet was resuspended in 10 ml of ice-cold 50 mM $CaCl_2$ and left on ice for 40 min. The cells were again pelleted as before and the

supernatant was discarded again. Finally, the pellet was re-suspended in 2 ml of ice-cold 50 mM $CaCl_2$ and incubated on ice until used or mixed with filter-sterilized cold glycerol (in approx. 3:1 ratio). Aliquots of 100 µl were made, shaken gently and stored at -80 °C.

3.6.2. Transformation of competent *E. coli* cells ($DH_5\alpha$)

Competent *E. coli* ($DH_5\alpha$) cells were transformed using heat-shock method. Initially the competent *E. coli* cells were prepared with the standard protocol described by Sambrook and Russel (2001). Approx. 100 µl of competent cells were added to a 30 µl ligation mixture, mixed gently by pipetting and incubated on ice for 40-60 min. The cells were heat shocked at 42 °C for 2 min in a dry bath exactly and placed on ice for further 2 min. 1 ml of LB medium was added to each tube and incubated at 37 °C for 60-90 min in a shaker incubator. Transformed cells were spread onto LB medium agar plates, supplemented with appropriate antibiotics. Plates were incubated inverted at 37 °C overnight.

3.6.3. Preparation of electro-competent *E. coli* ($DH_5\alpha$) cells

A single colony *E. coli* strain-DH5α was inoculated into 5 ml LB low medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) and incubated overnight in shaking incubator. 2 ml of the overnight culture was diluted into 100 ml of LB low medium in a 250 ml flask. The flask was incubated at 37 °C for 2 hours. The culture was chilled and pellet down by centrifugation at 5000 rpm (10 min / 4 °C). The supernatant was discarded and pellet was re-suspended with 10 ml of sterile, ice cold 10% glycerol by vortexing. The volume was adjusted to 40 ml with chilled 10% glycerol. The tubes were centrifuged at 5000 rpm (10 min / 4 °C). This washing step was repeated twice with

chilled 10% glycerol. Finally, the pellet was re-suspended in 0.5 ml ice chilled 10 % glycerol and aliquoted into 1.5 ml sterile Eppendorf tubes and stored at -80 °C.

3.6.4. Transformation of electro-competent cells of *E. coli* (DH$_5\alpha$)

Plasmid or ligation mixture approx.2 µl was added to the electro-competent cell (40 µl) and mixed gently. The mixture was transferred to chilled electroporation cuvette on ice and placed in the chamber in MicroPulser (Bio-Red) chamber until the cuvette was seated between the contacts in the base of the chamber and the charge button on the MicroPulser was pressed to charge the cuvette until a beep like sound was produced. The cuvette was removed and 1 ml SOC medium (Bacto tryptone, Bacto Yeast Extract, 5M NaCl, 1M KCl, 1M $MgCl_2$, 1M $MgSO_4$, 1M glucose) was added and the cells were transferred to an Eppendorf tube. The cells were incubated at 37 °C for 60 min, whilst shaking. The cells were spread on solid LB plate containing the appropriate antibiotics. The plates were incubated at 37 °C overnight in an incubator.

3.6.5. Preparation of competent *Agrobacterium tumefaciens* (GV3101) cells

A single colony was picked from a plate of *Agrobacterium tumefaciens* (strain GV3101) and inoculated into 10 ml of LB medium containing 50 µg of rifampicin in a 50 ml autoclaved flask and placed in a shaker at 160 rpm for 48 hours at 28 °C. 2ml of the culture was transferred to 1 L flask containing 250 ml Low- LB medium with 50 µg of rifampicin and incubated at 28 °C for atleast 3-4 hours at 225 rpm to an OD_{600} of at least 0.5-0.8. The culture was transferred to an autoclaved tube and incubated on ice for 10-15 min. Cells were harvested by centrifugation at 4000 rpm for 10 min at 4°C. Supernatant decanted and pellet was re-suspended in 15 ml of cold 20 mM $CaCl_2$ and incubated on ice for 40 min. Once again, the cells were harvested and the supernatant

was discarded. The pellet was then re-suspended in 5 ml of cold 20 mM $CaCl_2$ and 100 µl of aliquots were prepared and stored at -80 °C for future use.

3.6.6. Transformation of *A. tumefaciens* (GV3101)

Cells of agrobacterium were transformed following a standard protocol. 1µg of plasmid was added to 100 µl of competent cell and mixed gently. The mixture was dipped in liquid nitrogen. The transformation mixture was thawed for 20-30 min at 37 °C. 1 ml of LB broth medium was added and mixed by pipetting once or twice but gently. The mixture was incubated at 28 °C for 3-4 hours in shaker incubator. A total of 300 µl of the transformed cells were spread onto the LB medium agar plates supplemented with 50 µg each of rifampicin for *A. tumefaciens* and kanamycin of pGreen. Plates were incubated at 28 °C for 48 hours.

3.6.7. Preparation of electro-competent *Agrobacterium tumefaciens* (C58C1 or LB4404) cells

Single colony was used to inoculate 10 ml LB broth added with 50 µg of rifampicin and incubated at 28 °C at 180 rpm for 2 days. 5 ml of the culture was added to 250 ml low LB medium containing 50 µg rifampicin in 1L flask and placed at 28 °C for 3-4 hours in shaker incubator at 225 rpm to an OD_{600} at least 0.5-0.8. Culture was transferred to autoclaved tubes and incubated on ice for 10-20 min. Cells were harvested by centrifugation for 15 min at 4000 rpm at 4°C. Supernatant was discarded and pellet was dissolved in 10 ml cold SDW (sterile distilled water) and finally the volume was made upto 40 ml with addition of cold SDW. The cells were pelleted by centrifugation at 4000 rpm for 10 min at 4 °C. This step was repeated twice. The pellet was re-suspended in 10 ml of sterile cold 10% [v/v] glycerol water and centrifuged. This step of washing was also repeated. Finally, the pellet cells were re-suspended in

3-4 ml of ice-cold 10% [v/v] glycerol, aliquoted in 1.5 ml Eppendorf tubes and stored at -80 °C.

3.6.8. Transformation of *A. tumefaciens* (C58C1 or LB 4404)

Plasmid DNA (approx. 1-2 µl) was added to the thawed electro-competent (40 µl) cell and mixed by pipetting but gently. The mixture was transferred to a sterilized cold electroporation cuvette and the cuvette was set in the chamber slide and pushed the slide into the MicroPulser (Bio-Rad) chamber in such a way that the cuvette was seated between the contacts in the base of the electric shock chamber. After proper setting of the cuvette in the chamber, pressed the button set for "Agr" and waited to charge. After a beep, cuvette was removed immediately from the chamber and 1ml SOC medium (10 µl 0f 2MMg^{++}+10µl 2M Glucose + 980 µl SOB medium) was added to cuvette and mixed by pipetting. The transformed mixture was transferred back to 1.5 ml Eppendorf tube and incubated at 28 °C for 3-4 hours in shaker incubator. 100-200 µl of electroporated cells were plated on LB agar plates containing rifampicin for *A. tumefaciens* and kanamycin for pGreen. The plates were kept in a 28° C incubator for 48 hours.

3.6.9. Preparation of electro-competent *E. coli* (BL21-CodonPlus) cells

A single colony of *E. coli* (BL21-CodonPlus) cells was picked from a fresh LB agar plate and inoculated into 10 ml Low LB medium. The culture was incubated at 37 °C in the shaker for overnight. 2-3 ml of the overnight culture was added to 100 ml low LB medium containing in a 1L flask and placed at 37 °C for 3-4 hours in shaker incubator until reached to an OD_{600} at least 0.4-0.8. Culture was transferred to autoclaved tubes and incubated on ice for 20-30 min. Cells were harvested by centrifugation for 15 min at 4000 rpm at 4°C. Supernatant was discarded and pellet was

dissolved into 10 ml cold SDW and finally the volume was made up to 40 ml with addition of cold SDW. The cells were pelleted by centrifugation at 4000 rpm for 10 min at 4 °C. This step was repeated again. The pellet was re-suspended in 10 ml of sterile cold 10% [v/v] glycerol water and centrifuged. This washing step was also repeated. The pellet cells were finally re-suspended in 3-4 ml of sterilized ice-cold 10% [v/v] glycerol, aliquoted in 1.5 ml Eppendorf tubes and snap freeze tubes in liquid nitrogen and stored the frozen tubes at -80 °C.

3.6.10. Transformation of *E. coli* (BL21-CodonPlus) cells

Plasmid DNA (approx. 1-2 µl) was added to the thawed electro-competent (40 µl) cell and mixed by gentle pipetting. The mixture was transferred to a sterilized cold electroporation cuvette and the cuvette was set in the chamber slide and pushed the slide into the MicroPulser (Bio-Rad) chamber. After proper setting of the cuvette in the chamber, pressed the button and waited to charge. After a beep, cuvette was removed immediately from the chamber and 1ml SOC medium (10 µl 0f 2MMg^{++}+10µl 2M Glucose + 980 µl SOB medium)/ Low LB broth was added to cuvette and mixed by pipetting. The transformed mixture was transferred back to 1.5 ml Eppendorf tube and incubated at 37 °C for 30-60 min in shaker incubator. 100-200 µl of electroporated cells were spread on LB agar plates containing kanamycin. The plates were kept in a 37° C incubator for overnight.

3.7. Plasmid Digestion

Digestion of purified plasmid was performed using restriction endonucleases following manufacturer's instructions (Thermo Fisher Scientific or Biolab). A total reaction of 30µl or 10 µl was used when screening plasmid for the expected insert, 10-

30 µl of digestions kept for 2 hours when Fast digest enzymes (Biolabs) were used or overnight in case of conventional enzymes (Thermo Fisher Scientific).

3.8. Sequencing and sequence analysis

Plasmid colonies were selected and purified using a GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) and sent to Macrogen (South Korea) for sequencing with universal M13F [-20] and M13R [-20] primers. To extend the sequence, specific primers were designed (Primer walking). The sequences data (contigs) were assembled and analysed using Lasergene package of sequence analysis software (DNAStar Inc., Madison WI, USA). Sequence similarity searches (NCBI BLASTn) were performed by comparing the sequences to the other begomovirus/alpha-betasatellite sequences in the database (http://www.ncbi.nlm.nih.gov/ BLAST/). For the location of open reading frames in the genome, an online ORF Finder Program was used (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and with GeneQuest (Lasergene).Sequences were finally submitted to EMBL database (http://www.ebi.ac.uk/emb).

For full-length genome sequence analysis Geneious (version 7.1.7) software package or MEGA6.0 (Tamura *et al*., 2013) was used. Multiple sequences were aligned using CLUSTAL X (Thompson *et al.,* 1997) or MUSCLE (Edgar, 2004).Un-weighted Pair Group Method (UPGMA), Neighbor-Joining (NJ) and/or Maximum Likelihood (ML) methods with bootstrapping of 1000 replicates was used to infer phylogenetic relationships. Phylogenetic trees were constructed using Neighbor Joining algorithm of Clustal X and viewed, manipulated and printed using Tree view (Page, 1996). The figures were further refined using CorelDraw X5.

In some cases, recombination analysis were also performed using Plotcon, which is part of EMBOSS (Rice *et al.,* 2000) and analysis were run on-line

(http://emboss.bioinformatics.nl/), and Recombination Detection Program (RDP4) freely available at (http://darwin.uvigo.es/rdp/rdp.htm1) (Martin*et al.*,2015), with automated default settings (x-over option) of RDP, Chimera, MaxChi, Siscan, Geneconv and 3Seq was used.

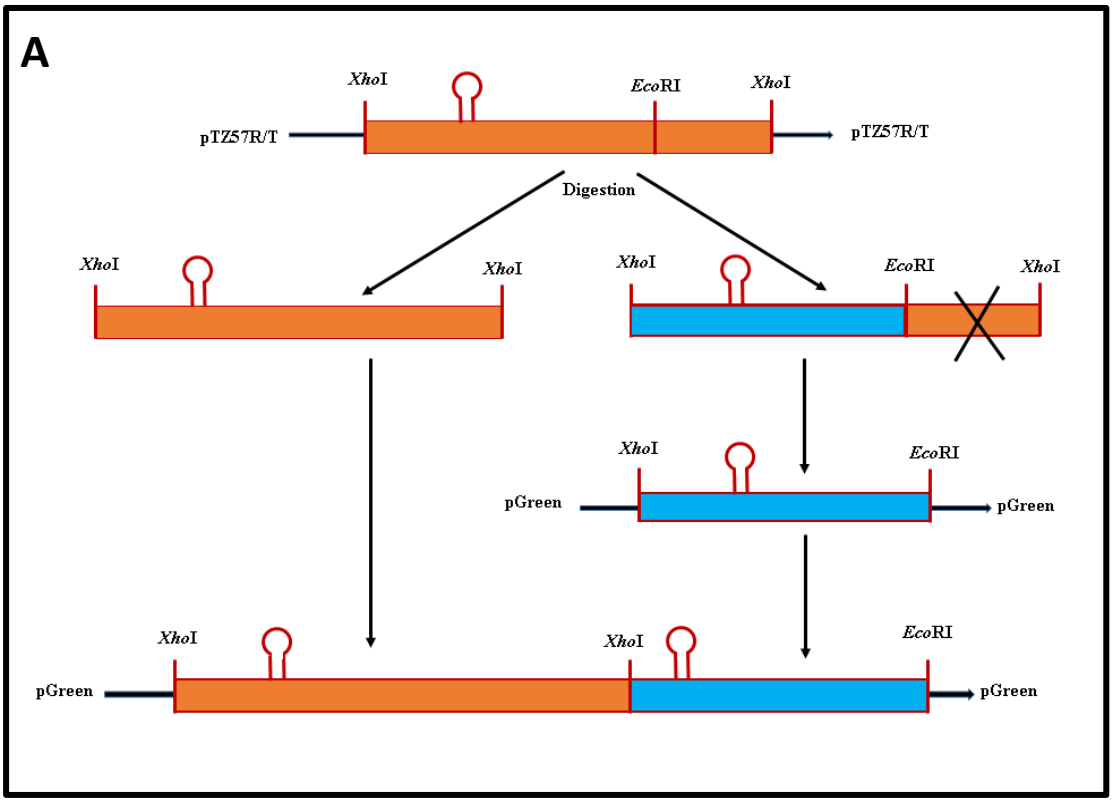
3.9. Construction of partial tandem repeats

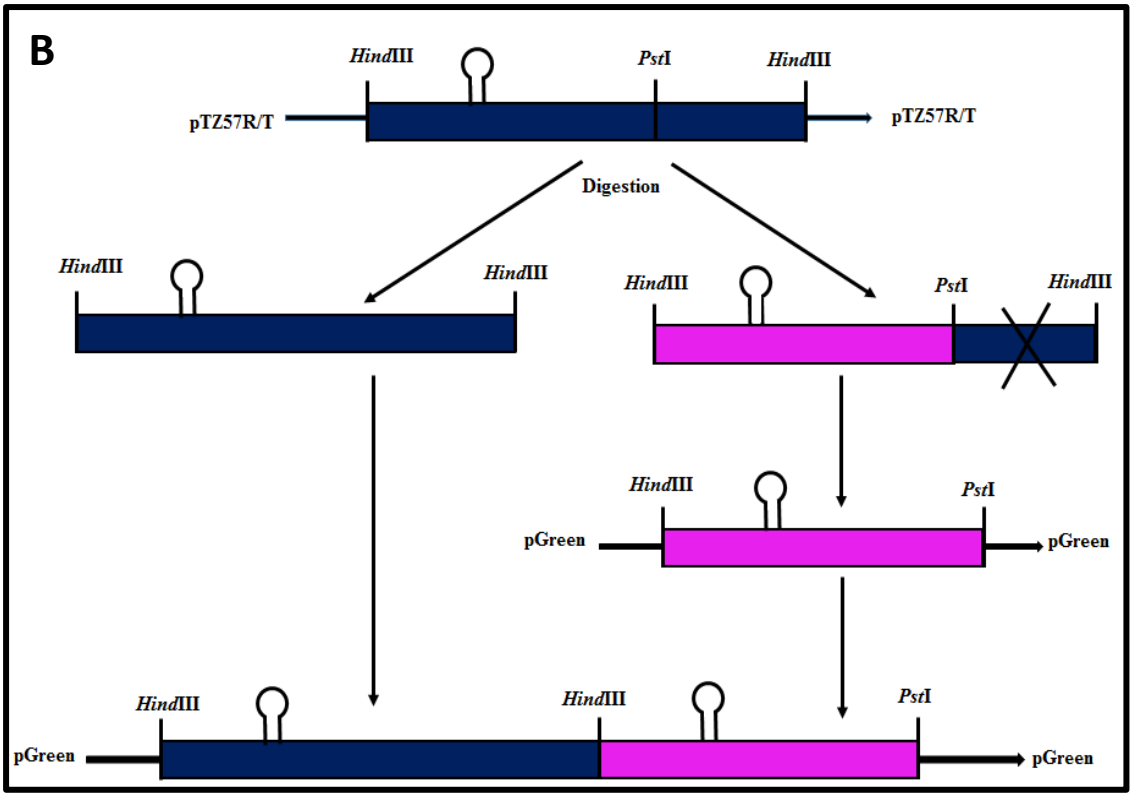
Full-length genome of begomoviruses was amplified, cloned into pTZ57R/T (Thermo Fisher Scientific) using abutting primers BurXF/BurXR to produce the clone pTZ-A-FL. Sequenced clones were used for construction of infectious clones. A fragment containing the entire intergenic region was released by digestion with *Xho*I - *EcoR*I and cloned with the same sites in the binary vector pGreen0029 to produce pGreen-A-2.1. Then full-length clone from the plasmid was released with *Xho*I and inserted into pGreen-A-2.1 at *Xho*I site to yield a clone pGreen-A-4.9 (containing a 2.1-mer partial tandem repeat [Figure 3.1A]).

The plasmid containing full-length genome of alphasatellite was digested with *Hind*III and *Kpn*I to release a fragment containing the entire intergenic region. The digested fragment was inserted in the same sites in the binary vector pGreen0029 to produce a clone of pGreen-α-0.9. Full-length alphasatellite fragment was released from the plasmid by digestion with *Hind*III and the fragment was sub-cloned into pGreen-α-0.9 in the same site to yield a clone pGreen-α-2.3 containing a 0.9-mer partial tandem repeat (Figure 3.1 B).

Similarly, the plasmid containing full-length genome of betasatellite was digested with *Kpn*I and *Sal*I to release a fragment containing the entire intergenic region. The digested fragment was inserted with the same site in the binary vector pGreen0029 to produce a clone of pGreen-β-0.55. Full-length betasatellite fragment

was released from the plasmid by digestion with *Kpn*I and the fragment was sub-cloned into pGreen-β-0.55 with same site to yield a clone pGreen-β-1.9 containing a 0.55-mer partial tandem repeat (Figure 3.1 C)





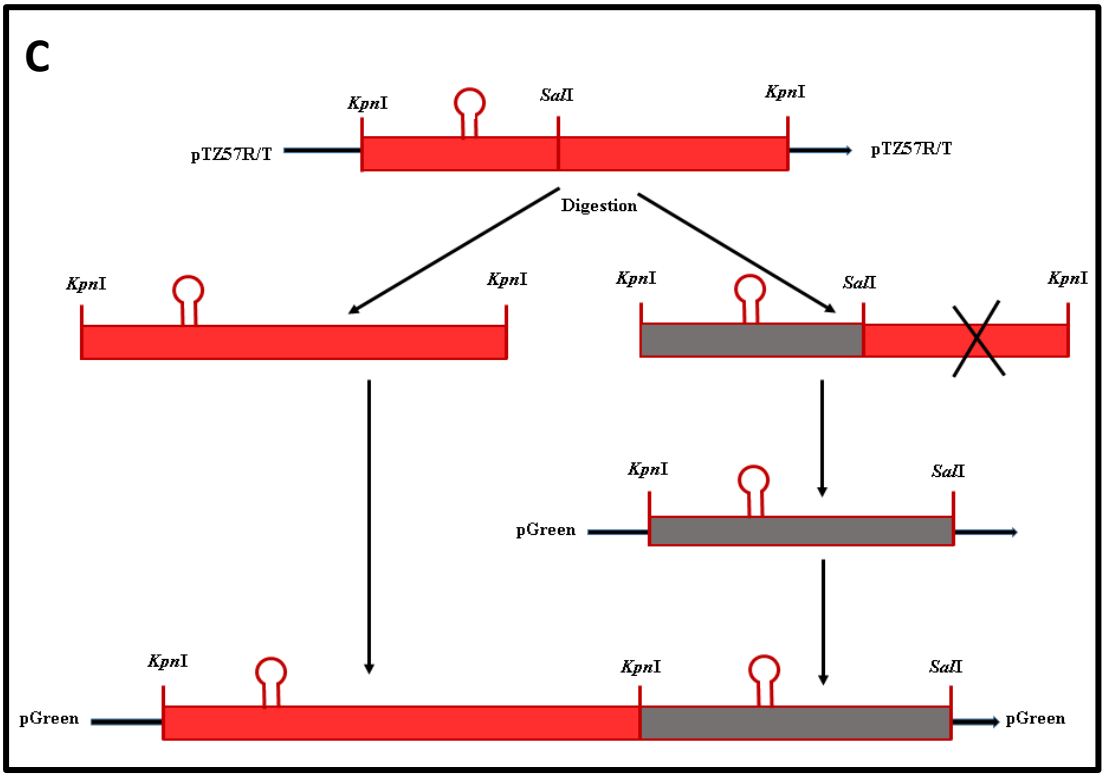


Figure 3.1. Construction of partial tandem repeat consructs of Cotton leaf curl Kokhran virus-Burewala with its associated cognates Cotton leaf curl Multan alphasatellite and Cotton leaf curl Multan betasatellite.

3.10. Agrobacterium-mediated inoculation

Partial dimer clones pGreen-A-4.9 and pGreen-β-1.9 and pGreen-A-3.4 were electroporated to *Agrobacterium tumefaciens* strain C58C1 or LB4404. For agro-inoculation, glycerol stocks were streaked on solid LB plates containing 50mg/ml each of the two antibiotics (kanamycin and rifampicin) and incubated at 28 °C for 48 hours. A single colony of bacterial cell was inoculated to 5 ml LB broth containing kanamycin and rifampicin and placed at 28 °C in shaker incubator for 48 hours and 1ml culture was diluted in 100 ml LB medium containing required amount antibiotics and incubated at 28 °C with continuous shaking. Cells were harvested by centrifugation at 5000 rpm at 20 °C for 15 min. For infiltration, 10 mM $CaCl_2$ and 100µM acetosyringone were added to the cells and incubated at 28 °C in shaker incubator till OD_{600} reached 1. The activated agrobacterium inoculum was used either to infiltrate 2-3 fully extended leaves per plant at the 5 to 6 leaf stage using 5ml needless syringe or to inoculate at the nodes of the young leaves with very fine needle. After agrobacterium mediated inoculation, plants were kept in an insect free glasshouse at 25 °C temperature with supplementary lighting to give a 16-hour photoperiod.

3.11. Development of coat protein construct

3.11.1. Construction of recombinant plasmid

Based on the available sequence of CLCuKoV-Bu (accession number HF549184), sequenced specific set of primers were designed for amplification of coat protein gene. The nucleotide sequence of the primers and primer restriction site (italic and underline) are given in Table 3.1. The PCR amplified product was purified by using agarose gel electrophoresis and cloned into pTZ57R/T plasmid. Coat protein clones were digested with *Nco*I and *Xho*I enzymes, and sub-cloned into expression vector

(pET28a) with same restriction sites. The conditions used for PCR were: 95°C for 3 min, 35 cycles of [95°C for 1 min, 52°C for 1 min, and 72°C for 2 min.], and a final extension was carried out at 72°C for 10 min. The amplicon thus obtained was cloned in pTZ57R/T plasmid using an InsTAclone PCR Cloning Kit (Thermo Fisher Scientific) and sequenced in both orientations (Macrogen, Seoul, South Korea).

3.11.2. Ligation and screening of recombinant clones

The purified PCR full-length coat protein gene was ligated in pTZ57R/T vector. Ligation mixture was prepared in 1.5 ml sterilized Eppendorf tube by adding PCR products, 6 µl of 5x ligation buffer (200 mM Tris-HCl, 50 mM $MgCl_2$, 50 mM DTT, and 5 mM ATP pH 7.8 at 25 °C), 3 µl of pTZ57R/T vector, 1 µl of T4 DNA ligase (5U) and volume was made to 30 µl with NF water (according to manufacturer's instructions). The tubes were vortex, spin down for 4-5 seconds and incubated at 4 °C overnight (for maximum ligation). The plasmid was named as pTZ-CP.

The ligation was transformed in *E.coli* strain $DH_5\alpha$ as in section 3.6.2. (Heat-shock method) or 3.6.4. (Electroporation). A recombinant colony was inoculated in 5 ml of LB medium containing 5ul of 100 µg/ml ampicillin. Plasmid extraction was done by using kits as describe in section 3.5.2 and run on 1% agarose gel electrophoresis.

3.11.3. Cloning of coat protein gene in pET expression system

The pET system is widely used for cloning and genes expression in *E. coli* host. The desired/ target gene(s) in this system is under the control the strong bacteriophage T7 transcription signal. The T7 RNA polymerase promoter is one of the strong promotors, when it is fully induced it uses most of the host resources to synthesize the target protein. On the other hand, the expression level can be monitored using inducer.

The target protein expression can be initiated by transferring a recombination plasmid into the bacterial system.

For the construction of recombinant plasmids pET28a-CPBur; pTZ-CPBur plasmids were digested with *Nco*I-*Xho*I restriction enzymes to release CP, which was then inserted into a similarly digested pET-28a (+) vector by ligase-mediated cloning. The resulting recombinant plasmids (pET28a-CPBur) was initially transformed into $DH_5\alpha$ cells. [this referred as hereby a "Lab Construct"; kindly provided by Dr. Muhammad Tahir (supervisor) for research work]. Presence of an insert in transformants were confirmed later by colony PCR, restriction digestion and nucleotide sequencing from Macrogen, Seoul, South Korea.The positive purified plasmid was then used to transform into *E. coli* BL21 CodonPlus expression host and the transformants were selected on LB-agar plates containing kanamycin.

3.11.4. Expression of the recombinant coat protein (rCP)

E.coli BL21 (CodonPlus) harboring the positive recombinant plasmids (pET28a-CPBur) were streak on LB agar plate containing 50 µg/ml kanamycin and allowed to grow overnight at 37°C. For expression studies, a single colony from pET28a-CPBurCodon Plus was picked with sterile loop from LB medium was inoculated into 10 ml Luria Bertani (LB) broth containing kanamycin (50 µg/ml) and grown overnight at 37°C with shaking incubator. Next morning about 1-2 ml of overnight culture was added into 50 ml of fresh LB broth containing kanamycin and incubated at 37 °C until the optical density (OD_{600}) reached 0.6 to 0.8. At this stage, the cells were induced by addition of Isopropyl-D-thiogalactoside (IPTG) at a final concentration of 0.5 mM and the fermentation was continued for another 6 hours at 37 °C in shaker incubator. Samples fractions were collected at 2 hours intervals during the

fermentation and subjected to 12% SDS PAGE. Finally, cells were harvested by centrifugation (6,500 rpm at 4°C for 15 min) and stored at -20°C.

3.11.5. Optimization of expression conditions

For the optimization of expression conditions of coat protein of begomovirus, fermentation of the transformed colonies was done in the 100 ml flask containing 10 ml LB broth (kanamycin 50 µg/ml). From the overnight culture, 1 ml culture was used to inoculate fresh 50 ml LB broth in a 500 ml sterile flask and incubated at 37 °C in shaker incubator until the culture reaches an OD ~0.5-0.8 at 600 nm (Donovan *et al.*, 1996).

Different optimization studies were done like, time course study after induction with IPTG, post induction 1ml samples were collected from the culture, at the interval of 2 hours up till 10 hours and OD of the culture was measured at 600 nm. Samples were harvested at 1,000 g for 2-3 min and pellets were analyzed on 12 % SDA-PAGE for total cell protein analysis and to evaluate the maximum level of expression. Un-induced pET28a-CPBur was used a control to test the effectiveness of expression conditions.

For the optimization of inducer concentration, transformed *E. coli* cells harboring the recombinant plasmids were induced with different IPTG concentrations from 0.2mM to 1.0 mM. All the culture was continued up till maximum level of expression and 1 ml samples from each culture was collected and cells were down by centrifugation at 12,000 g for 2-3 min and pellets were resuspended and analyzed on 12 % SDS-PAGE.

Similarly, for the optimization of cultivation temperature, the cultures were incubated at various ranges of temperature from 20 °C to 37 °C, for maximum

expression level at post IPTG induction. The cells were collected and further processed. Sonication was done to find the characteristic of induced protein i.e. soluble or insoluble at low to high temperature. Each time, samples were taken and analyzed on 12 % SDS-PAGE and measured the %age of expression by using documentation system, Dolphin software.

3.11.6. Sub-cellular localization of expressed protein

To check sub-cellular location of coat protein, 100 ml of LB broth containing 50mg/ml kanamycin was inoculated with 1 % of overnight inoculum/ culture and induced with 0.5 mM IPTG. The culture was grown at 37°C in shaking incubator for 6 hours. Cells were harvested by centrifugation at 6500 rpm for 15 min. Supernatant was discarded and the cell pellet was resuspended in 50 mM Tris-Cl (pH 8.0) to final $OD_{600}$of 10. Cells were lysed by sonication for 30-60 min at 60 amplitude with lysis cycle of 30 seconds. Cell lysate was centrifuged at 6500 rpm for 15 min at 4°C to separate the soluble and insoluble (inclusion bodies) protein. To the post induction, soluble and insoluble samples 100 µl of 2 x SDS loading dye [100 mM Tris-HCl (pH 6.8), 4 % (w/v) SDS, 20 % (v/v) glycerol, 200 mM dithiothrietol (DTT) and 0.2 % bromophenol blue] was added and vortex for short time. Samples were denatured/boiled for 5 min at 95 °C properly mixed by passing from 3cc syringe to reduce the viscosity. Samples were short spin and loaded on 12 % SDS gel against 5µl Phage Ruler (Thermo Fisher Scientific 26619) for protein weight determination. SDS-PAGE was stained for 1 hour in staining buffer [1.25 gm of Coomassie Brilliant Blue R-250, 450 mL methanol and water 450 ml and 100 ml glacial acetic acid] and de-stained overnight in the detaining buffer [200 ml methanol, 100 ml glacial acetic acid and water 700 ml for 1L]. SDS-PAGE analysis of fractions showed that the protein was expressed in inclusion bodies (Wang *et al*., 2005).

3.11.7. Solubilization of rCP from inclusion bodies

After the expression and localization of coat protein was confirmed, the harvested cells after sonication were washed thrice with IBs wash buffer [50 mM Tris-Cl (pH 8.0), 10 mM EDTA, 100 mM NaCl, 1% Triton X-100] and final washing was done two times with 50mM Tris-HCl (pH 8.0) and each time centrifuged at 6500 rpm for 15 min. Purified IBs containing the rCP were then solubilized with IB solubilization buffer pH 11.0 [50mM Tris-HCl, 6M Guanidine-HCl, 10 mM β-mercaptoethanol, 250 mM NaCl]. The suspensions were solubilized for 2-4 hrs in shaking incubator at room temperature. The solubilized IBs were then centrifuged at 6500 rpm for 20 min to obtain a clear supernatant for subsequent refolding. The supernatant and pellet obtained after solubilization were also loaded on 12% SDS-PAGE to check solubilization.

3.11.8. Purification of His-Tag expressed protein

Optimizing expression and solubilization protocol, the IBs from 1L of culture were solubilized in 50 mM Tris-HCl [pH 8.0]. Finally, the recombinant poly-histidine-tagged expressed protein was purified using glass column (Bio-Rad) containing Ni-IDA resin under native conditions. 3 ml of fresh Ni-IDA resin was loaded in the column and washed with 10 volumes of distilled water. The resin was equilibrated 10 volumes of Ni-binding buffer [50 mM Tris-Cl pH 8.0, 6 M guanidine-HCl, 0.5 M NaCl, 10 mM imidazole]. Flow rate of 0.5 ml/min was maintained throughout the process. 5 ml of sample was loaded onto column. The column was then washed with 10 column volumes of Ni-wash buffer-I [50 mM Tris-Cl pH 8.0, 6 M guanidine-HCl, 0.5 M NaCl and 50 mM imidazole] followed by washing with 5 column volumes of Ni-wash buffer-II [50 mM Tris-Cl pH 8.0, 6 M guanidine-HCl, 0.5 M NaCl and 100 mM imidazole]. Elution of bound proteins was done with 5 ml of Ni-elution buffer-I [50 mM Tris-Cl pH 8.0, 6 M guanidine-HCl, 0.5 M NaCl and 250 mM imidazole] and Ni-elution buffer II [50

mM Tris-Cl (pH 8.0), 6 M guanidine-HCl, 0.5 M NaCl and 500 mM imidazole]. Eluted protein was collected in 15 ml falcon tube and stored at 4 °C. All the fractions were check by SDS-PAGE.

Along with this, coat protein was also eluted by denaturing the polyacrylamide gels. After gel electrophoresis, gel was stained with Coomassie brilliant blue R250 for 60 to 90 min on shaking slide. Staining solution was decanted and gel was washed 2x times with distilled water. After washing, gel was incubated in the de-staining solution on shaker at room temperature. When the gel portion was properly destained, the washing steps was repeated that no traces left. With surgical blade, the gel piece was sliced to cubes of minimum possible size and placed in a 15 ml falcon tube.Gel pieces were washed with dH_2O. The gel pieces were meshed with pestle after repeated freeze thaw process. 1.5 ml of denaturing-polyacrylamide gel (D-PAG) elution buffer [50 mM Tris-Cl pH 8.0, 150 mM NaCl, 2m M EDTA and 0.1% SDS] was added to the gel andincubated in shaking at 37°C overnight. Next day the gel slurry was centrifuged at 7500 x g for 20 min at room temperature and supernatant containing protein was separated.

Protein sample of about 1 ml was taken in 15 ml falcon tube and 4 volumes of ice-cold acetone was added. Sample was briefly vortexed to mix and placed at -20°C for 60 min. Precipitated proteins were collected by centrifugation at 7500 x g for 15 min at 4 °C. Acetone was decanted and pellet was left to air dry for about 30 min. The pellet was dissolved in guanidine-HCl containing solubilization buffer [20 mM Tris-Cl (pH 8.0), 6 M guanidine-HCl, 2 mM β-mercaptoethanol]. Solubilized protein was further processed for refolding.

3.11.9. Refolding

The technique applied for refolding of recombinant coat protein (rCP) was by dilution method. Ice cold refolding tank or sink (50 mM Tris-HCl pH 11, 5M guanidine-HCl, 1mM PMSF, 1 mM cystine and 5 mM cysteine) was made and kept at 4 °C up till the end of refolding process. The solubilized protein was poured in the dialyses tube and was shaking by magnetic stirrer in the tank constantly at 4 °C to carry out air oxidation. Buffer in the tank was continuously changed until 0 M guanidine-HCl was used. The pH was then reduced from 11 to 9.5 gradually.

3.11.10. Quantification of purified protein

Quantification of purified protein of pET-CP was achieved by Bradford method (1987), by taking absorbance at 280 nm on UV Vis-spectrophotometer (AE-S70-2U) Bovine serum albumin (BSA) as a standard.

3.11.11. Production of antisera against the recombinant CP (rabbit anti-rCP Ab)

Antisera were raised in two rabbits of New Zealand White breed as shown in Figure 3.2. Each rabbit received a total amount of ~1000 µg of histidine tagged rCP in five injections over a period of four weeks interval per immunization. Initially, 200 µg/ml of purified rCP were emulsified with 1ml of Freund's complete adjuvant (FCA) [1:1, vol/vol] and was injected to each rabbit at 3-4 different sites subcutaneously. Followed by Freund's incomplete adjuvant (FIA) as booster dose with four weeks interval. Blood was collected from the marginal vein of the ear of each rabbit, after every four weeks interval. First bleed was drawn before the administration of the first injection (Freund's complete adjuvant) and it said as control or zero bleed (B_0). In total span of the production of antisera, blood was collected five times from each rabbit at 10 days after each booster injection and centrifuged at 6,000 rpm for 10 min. Serum

was separated, collected and stored at -20 °C for further titration and selection of the bleed.

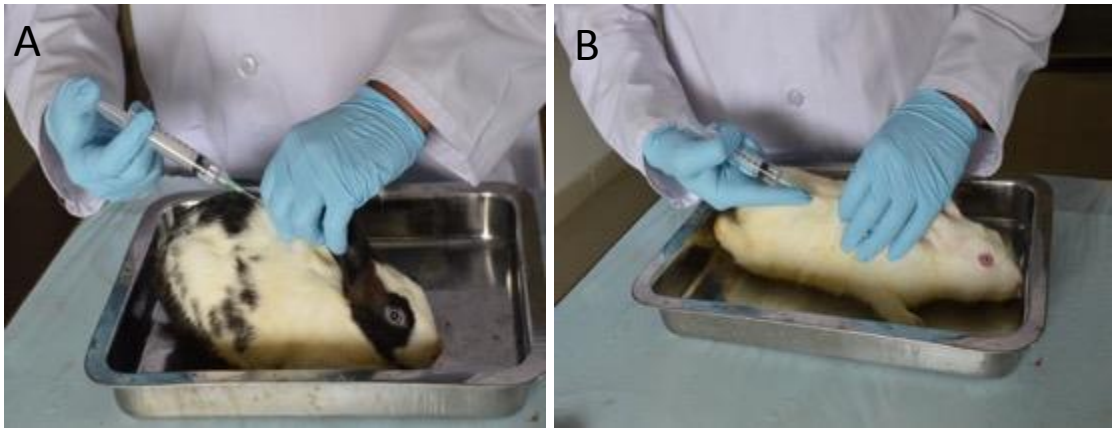


Figure 3.2. A and B. Production of polyclonal anti-rCP antibodies (Ab) in rabbits.

3.12. Double antibody sandwich ELISA (DAS-ELISA)

DAS-ELISA was performed essentially as described by Clark and Adams (1977) and Givord *et al*., (1994) for plant viruses. Serum was diluted in the coating buffer [15 mM Na_2CO_3, 35 mM $NaHCO_3$, 3mM NaN_3, (pH 9.8)]. From dilution, 200 µl was added to each well of the microtiter plate. Plate was tightly covered with plastic wrap and incubated for 3-4 hours at 37 °C. The plate was washed three times with PBS-Tween [137 mM NaCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, 2.68 mM KCl, 3.07 mM NaN_3, 0.05 % Tween (pH 7.4)] carefully. After each washing, plate was blotted by tapping on tissue paper to remove the PBS-Tween. 200 µl of extract sample was added to duplicate or triplicate wells of the plate. Plate was again covered with plastic wrap and incubated overnight at 4 °C. Next day, plate was washed with PBS-Tween. Washing step was repeated three times, followed by tapping on tissue paper. 200 µl of enzyme conjugate IgG-AP (kindly provided by Dr. Stephan Winter, DSMZ-Germany) was added to each well. Plate was covered with plastic wrap and incubated at 37 °C for

3-4 hours. After completion of incubation, plate was washed with PBS-Tween as described previously. 200 µl of freshly prepared substrate [1 mg/ml para-nitrophenyl-phosphate in substrate buffer] was added to each well. Plate was covered with plastic wrap and incubated at 37 °C for 3-4 hour, or if necessary to obtained clear reactions. Plate was visualized by observation due to yellowish colour appearance in the positive wells. The results were also assessed by ELISA reader taking absorbance at 405 nm.

3.13. Western blot analysis

Western blot was done based on the method described by Blake *et al*., (1984), and Mahmood and Yang (2012). After SDS-PAGE using Mini-PROTEAN tetra cell (Bio-Rad), gel was transferred to the tray containing transfer buffer [25 mM Tris, 190 mM glycine, 20% methanol (pH 8.3)] and incubated at room temperature for 15 min on a shaking bench.Nitrocellulose membrance was wet in the deionized water for 3 min. Sponge and filter paper were also soak in the transfer buffer for 3-5 min. Accessories for western blot cassette were assembled as shown in the figure 3.3. Roller was used to remove air bubbles from the sandwich, before the clamp were tighten.

Finally, cassette was set in the electrophoresis chamber containing transfer buffer. Apparatus was set to ensure that the blotting sandwich/cassette was covered with buffer and electrode in the proper direction. Cooling unit was also placed in the chamber to keep the temperature low. Protein from SDS-PAGE gel were transferred to nitrocellulose membrane, by using the method of Towbin *et al.,* (1979), by running apparatus for 150 min at 20V. The membrane was immediately transfer to tray containing blocking buffer [3% bovine serum albumin (BSA) in TBST Buffer {20mM Tris (pH 7.5), 150 mM NaCl, 0.1% Tween}]. It was placed on shaking surface for 15 min at room temperature. Blocking solution was poured off and primary antibody was

added to 10 ml of 5% BSA solution which was added to membrane. Membrane was incubated for 10-20 min at room temperature or overnight at 4 °C on shaker. Primary antibody was pour off and membrane was washed with wash buffer [20mM Tris (pH 7.5), 150 mM NaCl, 0.1% Tween] immediately. Second and third wash was done after incubation for 5 min at room temperature on shaker. At the end of the third wash, secondary antibody was added to the membrane and incubated to 10-20 min at room temperature on a shaking surface. Secondary antibody was poured off and membrane was washed 3x times with wash buffer as described previously. Finally, substrate was added to the membrane and incubated at room temperature on shaker until the development of color.

Figure 3.3. Schematic representation of western blot cassette.

4. RESULTS

4.1. Sample collection

Cotton leaf symptomatic and asymptomatic samples were collected from diverse cotton growing areas of Punjab, Pakistan during the year 2010-2013. Map of Pakistan displaying the major cotton growing areas are highlighted (Figure 4.1). Cotton samples displaying begomoviral like disease symptoms; thick and dark veins, upward or downward leaves, a small leaflet on lower side of leaves called enation and shunted growth were collected from the major cotton growing areas in Punjab as shown in the Figure 4.2. The complete information regarding each sample collection, date of collection and displayed disease symptoms by the plants is summarized in the Table 4.1. Samples were well labeled in plastic zippers and stored at -80 ^{o}C for further processing.

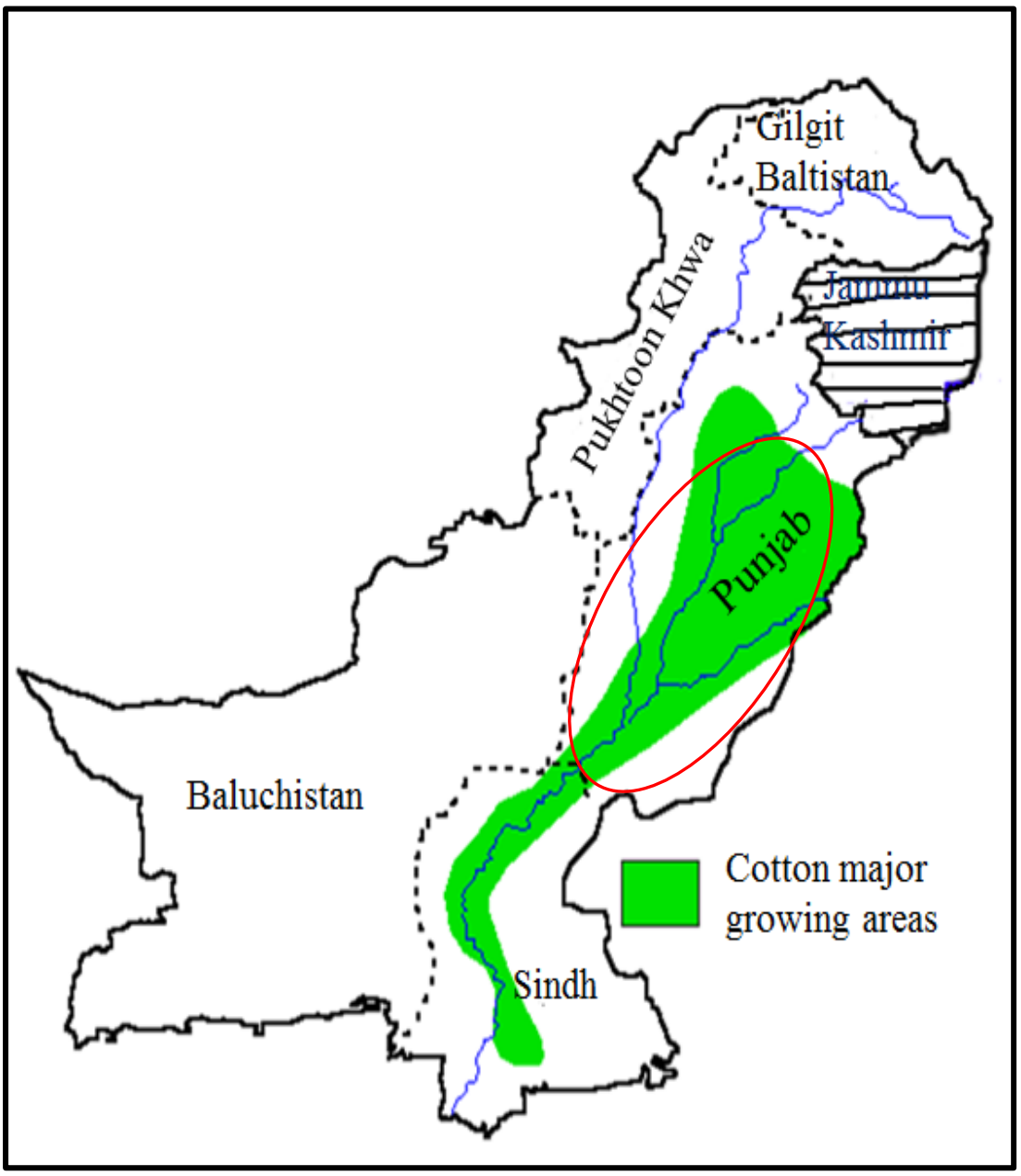


Figure 4.1. Map of Pakistan. Green portion of the map representing the major cotton growing areas in Punjab and Sindh Province.

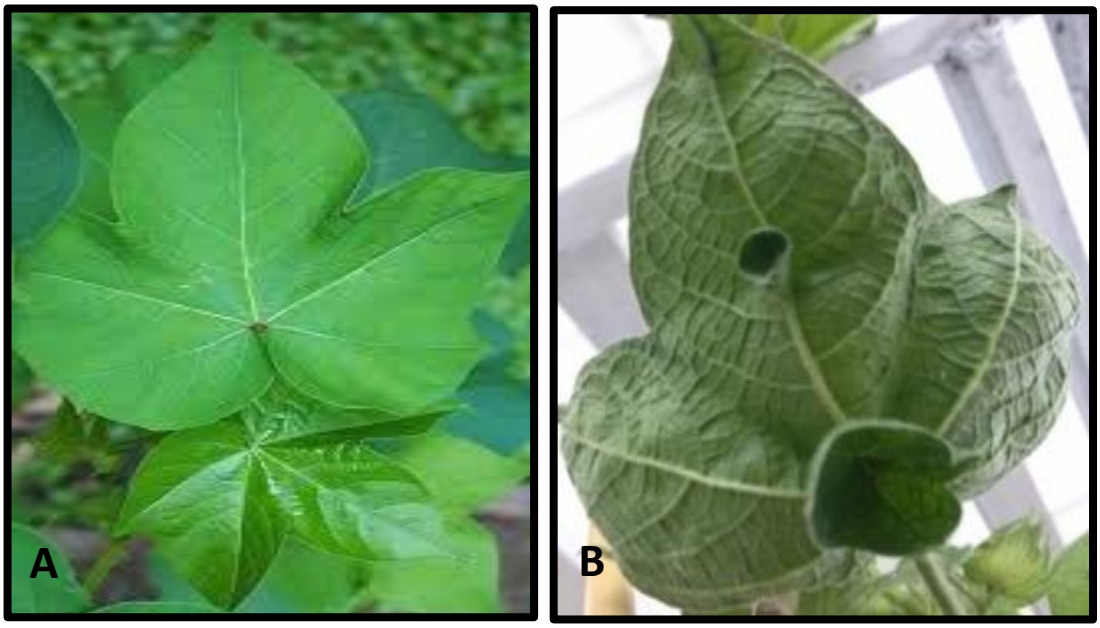


Figure 4.2. Healthy and infected cotton leaf samples were collected from areas around Faisalabad for this study. A) Healthy cotton leaf and B) infected cotton leaf showing leaf curling, vein thickening, a small leaflet on lower side called enation symptoms.

Table 4.1. Origins of samples, symptoms exhibited and year of collection.

Plant Species	Common Name	Isolate Code	Symptoms*	Location	Date of Collection
Gossypium hirsutum	Cotton	C-1	LC, VT, En	Islamabad	Oct, 2010
		C-2	LC, VT	Islamabad	Oct, 2010
		C-3	LC, VT	Islamabad	Oct, 2010
		C-4	LC, VT, En	Islamabad	Oct, 2010
		C-5	LC, VT	Islamabad	Oct, 2010
		C-6	LC, VT, En	Islamabad	Nov, 2012
		C-7	LC, VT	Islamabad	Nov, 2012
		C-8	LC, SL, SG,En	Mian wali	Jul, 2011
		C-9	LC, VT, En	Mian wali	Jul, 2011
		C-10	LC, VT, SG	Mian wali	Jul, 2011
		C-11	LC, En	Mian wali	Jul, 2011
		C-12	VT, En, UC	TobaTek Singh	Aug, 2011
		C-13	LC, En	Toba Tek Singh	Aug, 2011
		C-14	SG, SL, LC	Vehari	Oct, 2011
		C-15	LC, VT, En	Vehari	Oct, 2011
		C-16	LC, VT	Vehari	Oct, 2011
		C-17	LC, SG, En	Vehari	Oct, 2011
		C-18	VT, En, UC	Burewala	Oct, 2011
		C-19	LC, En	Burewala	Oct, 2011
		C-20	SG, SL, LC,En	Burewala	Oct, 2011
		C-21	LC, VT	Bahawalpur	Oct, 2011
		C-22	LC, VT	Bahawalpur	Oct, 2011
		C-23	LC, VT	Bahawalpur	Oct, 2011
		C-24	VT, En, UC	Bahawalpur	Oct, 2011
		C-25	VT, En, UC	Bahawalpur	Oct, 2011
		C-26	LC, En, SG	Faisalabad	Dec, 2010
		C-27	VT, En, UC	Faisalabad	Dec, 2010
		C-28	VT, En, UC	Faisalabad	Dec, 2010
		C-29	LC, VT, En	Faisalabad	Dec, 2010
		C-30	LC, VT	Faisalabad	Dec, 2010

		C-31	LC, SL, En	Faisalabad	Dec, 2010
		C-32	VT, En, UC	Faisalabad	Dec, 2010
		C-33	SG, SL, VT	Jhang	Sept, 2011
		C-34	LC, VT, SG	Jhang	Sept, 2011
		C-35	LC, En, SG	Jhang	Sept, 2011
		C-36	VT, En, UC	Jhang	Sept, 2011
		C-37	VT, LC, En	Jhang	Sept, 2011
		C-38	SG, VT, En	Muzaffargarh	Dec, 2011
		C-39	LC, VT	Muzaffargarh	Dec, 2011
		C-40	VT, En, UC	Muzaffargarh	Dec, 2011
		C-41	VT, En, SL	Muzaffargarh	Dec, 2011
		C-42	LC, En, SG	Muzaffargarh	Dec, 2011
		C-43	VT, En, UC	Muzaffargarh	Dec, 2011
		C-44	VT, En, UC	Dhok Ali Khan	Jul, 2011
		C-45	VT, En, UC	Dhok Ali Khan	Jul, 2011
		C-46	VT, En, UC	Dhok Ali Khan	Jul, 2011
		C-47	VT, En, UC	Dhok Ali Khan	Jul, 2011
		C-48	LC, VT, En	Khushab	Jul, 2011
		C-49	VT, UC	Layyah	Oct, 2010
		C-50	SG, VT, En	Bhakkar	Oct. 2011
		C-51	VT, En, LC	Bhakkar	Oct. 2011
		C-52	SG, SC, En	Multan	Sept, 2011
		C-53	LC, VT, En	Multan	Sept, 2011
		C-54	LC, En	Multan	Sept, 2011
		C-55	VT, En, LC	Bahawal Nagar	Sept, 2011
		C-56	LC, En	Bahawal Nagar	Sept, 2011
		C-58	LC, VT, En	Okara	Oct. 2011
		C-59	VT, En, LC	Okara	Oct. 2011
		C-60	SG, SC, LC, En	Okara	Oct. 2011
		C-61	SG, SC, En	Khanewal	Oct, 2011
		C-62	VT, LC, En	Khanewal	Oct, 2011
		C-63	SG, VT, En	Rahim Yar Khan	Oct, 2013
		C-64	LC, VT, En	Rahim Yar Khan	Oct, 2013

*Symptoms are denoted as Vein thickening (VT), Enations (En), Upward Leaf Curling (UC), Shunted Growth (SG), Severe Leaf Curling (SC) and Leaf curling (LC)

4.2. Begomoviral Diagnostic PCR

The following samples were collected from healthy and symptomatic cotton farms from different districts in Punjab. To start with, the presence of a begomovirus in the symptomatic samples was confirmed using two diagnostic primer sets (Table 3.1) were used in the thermocycler with total nucleic acid samples extracted from the plants. The first primer set used was WTGF/WTGR (Mansoor *et al.,* 2000a), degenerate primers. This set of primer pair was reported to detect all types of whitefly transmitted geminiviruses and could theoretically amplify 1500bp product, covering the IR and most of the *Rep* gene. Another set of primer CLCV1/CLCV2 was used which was more specific to cotton infecting begomoviruses, covering a product of 1100bp CP in virion sense strand and the TrAP region (Figure 4.3A). Most of the cotton infecting begomoviruses are monopartite and associated with a small satellite components. The presence of betasatellite in the samples was identified and amplified employing universal primer pair Beta01/ Beta02 (Briddon *et al.,* 2002). The primer set amplify a fragment of approximately 1350bp corresponding to full length betasatellites (Figure 4.3C). Along with betasatellite, the occurrence of alphasatellite (previously known as DNA 1) was also detected in the samples using AlphaF/ AlphaR primer set (Zia-ur-Rahman *et al.,* 2013) (Figure 4.3D).

For the presence of DNA B (present in the bipartite begomovirus) in the samples, previously known primer sets were used as diagnostic in the PCR.MPF/MPR (Husain *et al*., 2004) were designed to amplify approximately 850bp product of the DNA B, covering the *MP* gene (Figure 4.3E). Along with this, an abutting primer sets KTBF/KTBR (Tahir *et al.,* 2010) was also used to amplify full-length DNA-B, if diagnosed in the samples (Figure 4.3F). The sequences of all the primers used are in Table 3.1.

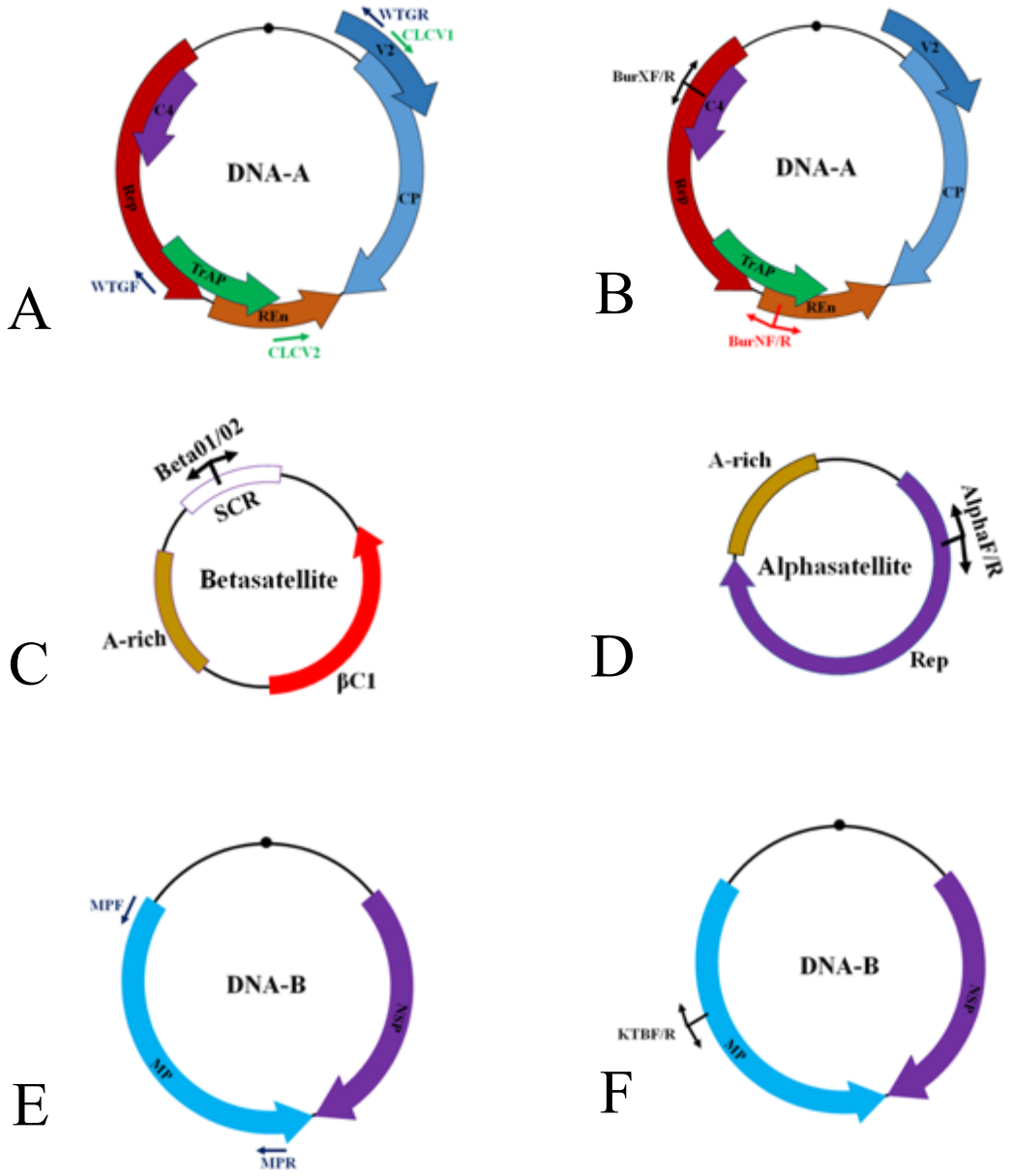


Figure 4.3. Position of the diagnostic and specific primers used to detect the presence of the genomes (or DNA A genomic components) and DNA B components of begomoviruses. (A) Diagnostic primer pairs (WTGF/WTGR) (Mansoor *et al.*, 2000a) for amplification most the *Rep* gene covering the intergenic region of all whitefly transmitted geminiviruses and specific primer pair (CLCV1/CLCV2) annealing in the coat protein gene and TrAP gene, designed to sequences conserved among begomoviruses from the Indian subcontinent (Hussain *et al.*, 2003) and (E) diagnostic primer pairs (MPF/MPR) for DNA-B of bipartite begomoviruses. Primer sites and orientation are marked with arrows, (B) position of abutting primers (BurXF/ BurXR, and BurNF/ BurNR) for the amplification of full-length DNA-A is marked with arrows (Amrao *et al.*, 2010); Fareed *et al.*, 2012), (C) universal primer pair (Beta01/Beta02) for amplification of begomoviruses associated DNA-betasatellite (Briddon *et al.*, 2002), (D) abutting primer pair (AlphaF/AlphaR) for amplification of alphasatellite (Zia-ur-Rahman *et al.*, 2013) and (F) primer pair (KTBF/KTBR) for amplification of full-length DNA B genomic component. The sequences of primers are given in the Table 3.1.

Table 4.2. Results of diagnostic PCR to detect the occurrence of a begomovirus genomic component (or DNA A), DNA B, alphasatellite and associated betasatellite.

Sample code	Diagnostic Primer pairs				
	WTGF/ WTGR	CLCV1/ CLCV2	Beta01/ Beta02	AlphaF/ AlphaR	MPF/ MPR
C-1	-	-	-	-	-
C-2	-	-	-	-	-
C-4	-	+	+	-	-
C-5		+	-	-	-
C-6	+		-	-	
C-7		+	+		-
C-8		+			-
C-9		+	+		-
C-10		+	-		-
C-12	+	+			-
C-13	+				-
C-15		+			-
C-16		+			-
C-17	+				
C-18		+			-
C-19		+			-
C-20		+			-
C-21	+				-
C-22	+		+		-
C-23		+			-
C-24		+			-
C-25		+			-
C-26		+			
C-27	+				-
C-28*	+	+	+	+	-
C-29		+			-
C-30		+			-
C-31		+			
C-32*	+	+	+	-	-
C-33	+				-
C-34	+	+			-
C-35		+			-
C-36		+			-
C-37		+	-	-	-
C-38	+				-

C-39	+				**-**
C-40		+			**-**
C-42	+	+			**-**
C-43		+			**-**
C-44			+	-	**-**
C-45		+		+	**-**
C-46	+		+		**-**
C-47*		+	+	+	**-**
C-48		+			**-**
C-49*		+	+	+	**-**
C-50*		+	+	+	**-**
C-51	+			-	**-**
C-52*		+	+	+	**-**
C-53		+	-	-	**-**
C-54	+		+		**-**
C-55*		+	+	+	**-**
C-56	+				**-**
C-57	+		+		**-**
C-58*	+	+	+	+	**-**
C-59		+	+		**-**
C-60	+		+	-	**-**
C-61		+	-	-	**-**
C-62*		+	+	+	**-**
C-63	+		+		**-**
C-64*		+	+	+	**-**

Positive (+) sign specifies the presence of a product while minus (-) sign is for the absence of a product in diagnostic PCR reactions, () no sign or empty specifies that sample was not processed for PCR. Sample with (*) was further processed for amplification of full-length genomic components

The diagnostic PCR showed the presence of monopartite begomoviruses associated with betasatellites and alphasatellites shown in Table 4.2. None of the plants showed the presence of a bipartite begomovirus. The diagnostic primers gave the expected size products of 1500bp, and 1100bp with both WTGF/WTGR and CLCV1/CLCV2 primer sets. Partial sequences with CLCV1/CLCV2 primer sets, were sequenced and analyzed using NCBI and Lasergene (DNAStar Inc., Madison, WI,

USA) software. Pairwise sequence comparisons against all begomovirus genome (or DNA-A component) sequences available in the databases were run on-line using pairwise sequence comparison (PASC) software (http://www.ncbi.nlm.nih.gov/sutils/pasc/viridty.cgi) and sequence distance table (Table 4.3) was produced. All sequences showed more than 98% nucleotide sequence identity with each other and above 97 % with the available sequences belong to Cotton leaf curl Kokhran virus-Burewala (previously known as Cotton leaf curl Burewala virus) in the databases. Based on these partial sequences, these results elucidating the widespread of a single strain in this region, infecting cotton. Due to the presence of the same strain of begomovirus in all analyzed samples, sequencing with partial clones was not further proceeded.

The primers gave the expected size products of ~2750 bp with both specific abutting primer pairs BurNF/BurNR and/or BurXF/BurXR, respectively. Universal primer sets Beta01/Beta02 and AlphaF/AlphaR also yield the expected size of product of ~1300 bp and ~1400 bp, corresponding to betasatellite and alphasatellite, respectively. Contrary to this, the primer pair MPF/MPR for amplification of the movement protein or specific abutting primer pair for amplification of the full-length clone corresponding to DNA B genomic component of bipartite begomoviruses was unable to yield any product in all PCR reactions that were previously found positive with CLCV1/CLCV2 and WTGR/WTGR. The amplified products were cloned and sequenced.

Table 4.3. Percentage nucleotide sequence identity among partial sequences obtained with CLCV1/CLCV2 primer sets from isolatesC5,C7, C9, C12,C15, C18, C20, C25,C28, C34, C32, C47, C48, C50, C52, C55, C58, C59, C62 and C64 with selected begomoviruses in the databases. Sequences were selected of the most similar species from the Blast comparison.

Isolates	ToLCNDV (2)	PaLCV (2)	CLCuGeV-EG (3)	CLCuBaV (2)	CLCuAlV-Al (2)	CLCuAlV-Mul (2)	CLCuMV-Fai (4)	CLCuKoV-Sha	CLCuKoV-Ko (4)	CLCuKoV-Bu (10)	CLCV (24)
CLCV	75.8-76.0	88.9-94.3	72.2-73.8	79.7-81.1	77.2-78.7	79.3-80.6	79.2-80.4	91.3-93.1	94.5-96.9	98.0-100	**98.2-100**
CLCuKoV-Bu	75.7-75.9	90.1-93.3	72.6-73.9	80.0-81.1	78.5-78.7	77.5-78.3	80.3-80.4	92.8-93.1	96.1-96.8	**99.8-100**	
CLCuKoV-Ko	75.9-76.1	92.7-96.1	73.5-74.7	83.7-84.2	76.6-77.6	76.9-77.7	77.5-78.3	89.6-90.4	**99.2-100**		
CLCuKoV-Sha	76.0-76.5	84.8-88.0	71.8-73.0	78.9-79.9	72.7-83.0	85.9-86.0	86.4-86.7	**99.3-100**			
CLCuMV-Fai	74.5-74.8	75.8-77.7	70.3-70.8	81.9-82.6	89.2-89.2	89.0-89.2	**99.8-100**				
CLCuAlV-Mul	75.9-76.2	75.5-77.4	71.6-71.7	81.4-82.5	91.3	**100**					
CLCuAlV-Al	74.2-74.8	75.4-77.4	70.6-70.9	83.1-85.5	**99.7-100**						
CLCuBaV	77.5-78.0	76.8-80.0	71.2-71.7	**97-100**							
CLCuGeV-EG	70.7-71.2	71.8-74.0	**98.4-100**								
PaLCuV-Pk	75.3-77.2	**95-100**									
ToLCNDV	**98.6-100**										

From overall amplification and sequencing, thirty (30) full-length sequence data is presented in this study, of which ten (10) belongs to monopartite begomoviruses (DNA A component), ten (10) clones are corresponding to DNA betasatellites and ten (10) were monopartite begomovirus alphasatellites associated with them. The diagnostic primers did not yield products in PCR reactions containing nucleic acid from the healthy cotton plants. The sequences obtained exhibited more than 98% of nucleotide similarity with the available sequences in the databases.

4.3. Nucleotide Sequence Analysis

4.3.1. Analysis of the cloned DNA A components of begomoviruses

After confirmation with diagnostic PCR, the presence of begomovirus in the nucleic acid of all samples collected from different locations were processed for amplification of full-length clones by using specific abutting primer pairs BurXF/BurXR (Amrao *et al*., 2010) and BurNF/BurNR (Freed *et al.*, 2012) in the PCR reactions. All amplicons were cloned and sequenced.

4.3.1.1.Analysis of the sequence of the DNA A components of begomovirus isolates C28A, C32A, C47A,C49A, C50A, C52A, C55A, C58A, C62A and C64A

The nucleotide sequences of ten (10) full-length begomoviruses were determined through DNA sequencing. Genome size of all isolates identified here was 2,759 nucleotides in length, except the isolate C49 with 2751bp in length.

An initial search using the public sequence databases under BLASTn directed that nine DNA-A components are sharing >95% nucleotide sequence identities among them and with the available CLCuKoV-Bu sequences in the databases, while isolate from C49A is exhibiting 92% nucleotide sequence identities with all isolates identified

in this study and with all previously submitted sequences in the databases. Closer inspection of the isolates identified containing six ORFs with a typical arrangement of the DNA A components of the Old World begomoviruses (Figure 1.4B), consisting of two ORFs (CP and V2) on the virion-sense strand and four ORFs (Rep, TrAP, REn, C4) on the complementary sense-strand. The features of these components are summarized in Table 4.4.

Table 4.4. Characteristic features of begomovirus isolates obtained from *Gossypium hirsutum.*

Start Codon-Stop Codon Predicted size (no. of amino acid)/molecular weight (kDa)							
Plant Sample	**V2**	**CP**	**Rep**	**REn**	**TrAP**	**C4**	**C5**
C-28	132-488 118/13.7	292-1062 256/29.7	2596-1505 363/40.7	1463-1059 134/15.5	1608-1501 35/4.3	2682-2137 181/20.3	807-283 174/19.8
C-32	132-488 118/13.7	292-1062 256/29.6	2596-1505 363/40.8	1463-1059 134/15.6	1608-1501 35/4.3	2682-2242 146/16.3	807-283 174/19.9
C-47	132-488 118/13.6	292-1062 256/29.7	2596-1505 363/40.7	1463-1059 134/15.7	1608-1501 35/4.3	2682-2242 146/16.3	807-283 174/19.9
C-49	118-474 118/13.7	278-1048 256/29.6	2588-1497 363/40.8	1455-1051 134/15.6	1600-1481 39/4.6	2674-2129 181/20.5	
C-50	132-488 118/13.7	292-1062 256/29.6	2596-1505 363/40.7	1463-1059 134/15.4	1608-1501 35/4.3	2682-2242 146/16.3	807-283 174/19.8
C-52	132-488 118/13.6	292-1062 256/29.7	2596-1505 363/40.7	1463-1059 134/15.3	1608-1501 35/4.4	2682-2137 181/20.4	
C-55	132-488 118/13.6	292-1062 256/29.7	2596-1505 363/40.6	1463-1059 134/15.4	1608-1501 35/4.3	2682-2242 146/16.3	807-283 174/19.7
C-58	132-488 118/13.7	292-1062 256/29.6	2596-1505 363/40.7	1463-1059 134/15.4	1608-1501 35/4.3	2682-2242 146/16.3	807-283 174/19.8
C-62	132-488 118/13.5	292-1062 256/29.7	2596-1505 363/40.6	1463-1059 134/15.3	1608-1501 35/4.2	2682-2242 146/16.2	
C-64	132-488 118/13.6	292-1062 256/29.6	2596-1505 363/40.7	1463-1059 134/15.4	1608-1501 35/4.3	2682-2242 146/16.3	807-283 174/19.8

The predicted coding capacity of the C4 gene of C28, C49 and C52 was higher (181 amino acid) in comparison with the rest of isolates (146 amino acids). On the other hand, C5 gene, encoding for 174 amino acid, was found missing in three isolates as shown in the Table 4.4. Similarly, the table reveals that the coding capacity of the TrAP in the isolate C49 DNA A was also higher (39 amino acid) than those of the other isolates (35 amino acid). The coding capacity for the ORF TrAP in Burewala-strain has also a premature termination due to two stop codons at different positions. The isolate C49A TrAP gene is slightly encoding for larger protein due to a point mutation in the sequence in comparison with all available sequences in the databases or/and obtained here.

The sequence of the intergenic region from all isolates in the present study were aligned using Clustal X **(**Figure 4.4). This region contains a series of *cis*-acting DNA elements (involved in begomovirus replication and transcriptional regulation of *Rep* gene) (Arguello-Astorga *et al.,* 1994; 2001). All such elements are present in the sequences of all isolates including; (i) Rep-binding site (iterons), (ii) TATA boxes, iii) GC boxes, (iv) core iteron [GGAG]N conserved in all CLCV isolates, and (v) hairpin stem loop structure that contains the conserved nona-nucleotide (TAATATTAC) sequences nicked by the Rep protein to initiate DNA replication. Although, there was no variation found in the sequences of iterons among all isolates, the stem-loop and nona-nucleotide sequences are highly conserved.

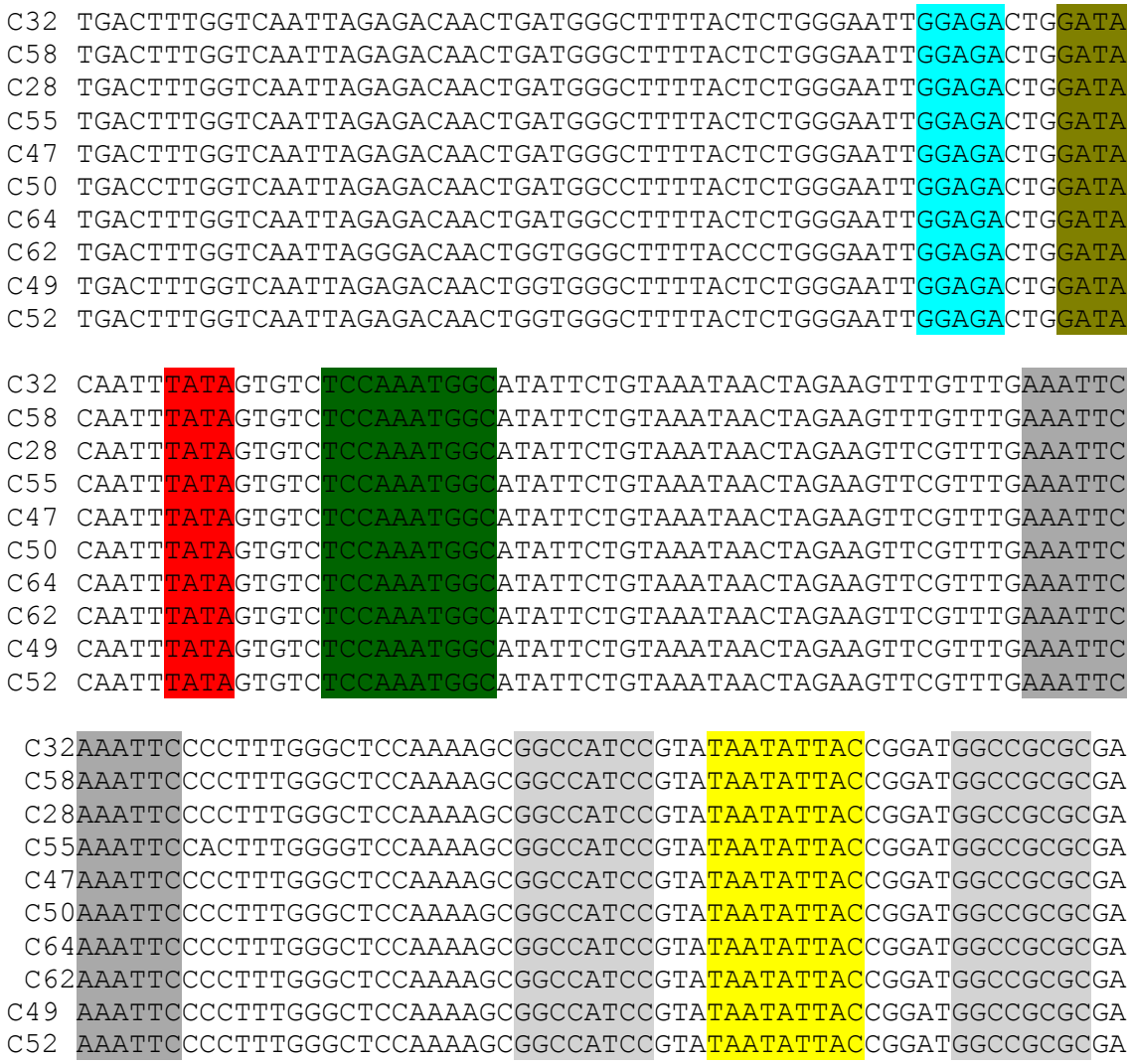


Figure 4.4. Multiple alignment of IR sequences of begomovirus isolates conserved nonanucleotide sequences [TAATATTAC] (highlighted in yellow), the conserved stem-loop (highlighted in light gray), repeated iterons [AAATTC] (highlighted in dark gray), core iteron [GGAG]N (light sky blue), GATA motif (highlighted in olive green).

4.3.1.2. Comparison of the DNA A components to other begomoviruses

An initial trawl of the public sequences databases using BLASTn, the virus isolates of the present study (belong to DNA A) indicates a high levels of nucleotide sequence identity to isolates of Burewala strains. A table (Table4.5) is constructed to show the percentage nucleotide sequence identities among these isolates with selected

viruses, based upon most similar sequences (one isolate for each species), available in the database.

In the table, all isolates, except C49, showed a 97% to 99% nucleotide sequence identity among them and 96 % to 98 % with the available sequences representing CLCuKoV-Bu, followed by CLCuMuV-Fai with 89% to 90% and CLCuKoV-Ko with 88 % to 89%, respectively. Minimum nucleotide sequence identity for C28, C32, C47, C50, C52, C55, C58, C62 and C64 was observed with CLCuGeV-EG (73%). On the other hand, DNA A from isolate C49 showed a maximum nucleotide identity of 87.7 % to 92.1% with the isolates identified here, followed by 92% with CLCuKoV-Bu reported from India, 91% by CLCuKoV-Sha and 89% by CLCuMuV-Fai, repectively. Minimum nucleotide sequence identity for C49 was observed with CLCuGeV-EG (73%).

Based on the reported scenario for the genus begomovirus description, nucleotide sequence identity is more than 94%, indicating an isolate to be a variant of a same species, while nucleotide sequence identity is less than 94% but more than 91% specify an isolate to be a new strain of a species and finally, nucleotide sequence identities of an isolate if less than 91% with reported sequences in the databases may predict a new species (Brown *et al.*, 2012). The DNA A sequences of isolates C28, C32, C47, C50, C52, C55, C58, C62 and C64 showed high levels of nucleotide sequence identities (96.8 to 98.2 %) among them and available sequences in the Genbank. Based on the currently accepted species demarcation threshold (more than 95%), all isolates are variants of single species, except C49.

Table 4.5. Percentage nucleotide sequence identity between the complete sequences of the DNA As of isolatesC28, C32, C47, C49, C50, C52, C55, C58, C62 and C64 with selected begomoviruses in the databases. Sequences were selected of the most similar species from the Blast comparison.

Isolates DNA A Components	**ToLCV-[IN-Ban-Chi-08]**	**PaLCuV-PK[PK-Cot-02]**	**CLCuGeV-EG[EG-Cai-Okr]**	**CLCuKoV-Sha[PK-TanA-05]**	**CLCuBaV-[IN-Ban-04]**	**CLCuAlV-Al[PK-Koh802a-96]**	**CLCuKoV-Ko[PK-Man806b-96]**	**CLCuMuV-Fai[CN-Fz1-Hib-12]**	**CLCuKoV-Bu[IN-SriGang-05**	**C28A, C32A, C47A, C50A, C52A, C55A, C58A, C62A, C64A**	**C49A**
C49A	79.8	80.2	73.2	91.7	85.2	79.1	84.8	89.9	**92.1**	87.7-92.1	**100**
C28A, C32A, C47A, C50A, C52A, C55A, C58A, C62A,C64A	80.1-81.7	81.6-83.1	73.2-73.8	89.4-90.6	83.8-84.4	75.6-76.6	88.2-89.4	**89.3-90.6**	**96.4-98.2**	**100**	
CLCuKoV-Bu[IN-SriGang-05]	81.3	82.4	73.7	89.6	83.7	75.8	87.5	90.4	**100**		
CLCuMuV-Fai[CN-Fz1-Hib-12]	76.8	75.6	73.7	85.5	85.2	83.1	79.1	**100**			
CLCuKoV-Ko[PK-Man806b-96]	85.2	86.2	75.9	91.4	81.2	74.6	**100**				
CLCuAlV-Al[PK-Koh802a-96]	75.1	79.2	70.4	77.9	79.2	**100**					
CLCuBaV-[IN-Ban-04]	79.5	76.7	74.3	85.4	**100**						
CLCuKoV-Sha[PK-TanA-05]	86.6	80.6	74.7	**100**							
CLCuGeV-EG[EG-Cai-Okr]	72.9	71.6	**100**								
PaLCuV-PK[PK-Cot-02]	83.4	**100**									
ToLCV-[IN-Ban-Chi-08]	**100**										

Percentage valves more than 91% (species demarcation threshold for begomovirus DNA A, Brown *et al.*, 2012) are in bold

Table 4.6. The predicted amino acid percentage identities of the DNA A gene products of the virus isolates characterized in this study with sequences available in the databases.

Isolate	CP	V2	Rep	TrAP	REn	C4	C5
C28	CLCuKoV-Ko [PK-806b-98] 99	CLCuKoV-Sha [PK-anJ-05] 99	CLCuKoV-Ko [IN-IVRI-45-13] 99	CLCuKoV-Bu [PK-Neo18-12] 97	CLCuKoV-Bu [PK-Kha5-04]99	CLCuKoV-Bu [PK-KoB-04] 99	CLCuKoV-Bu [IN-Dab1-10] 99
C32	CLCuKoV-Bu [PK-RYK7-04] 99	CLCuKoV-Ko [PK-806b-97] 99	CLCuKoV-Bu [IN-Han-10] 99	CLCuKoV-Bu [PK-RYK18-12] 97	CLCuKoV-Bu [PK-Pun-06] 99	CLCuKoV-Bu [PK-Kha5-04] 99	CLCuKoV-Bu [IN-Pun-07] 99
C47	CLCuKoV-Ko [Pk-806b-97] 99	CLCuKoV-Ko [PK-Fai1] 99	CLCuKoV-Bu [IN-Han-10] 99	CLCuKoV-Bu [IN-Sri-05] 97	CLCuKoV-Bu [IN-Sri1-10] 99	CLCuKoV-Bu [PK-Fai13-04] 99	CLCuKoV-Bu [PK-Veh2-04] 99
C49	CLCuKoV-Bu [PK-RYK7-04] 99	CLCuKoV-Sha [PK-TanJ-05]99	CLCuMuV-Fai [IN-Sri-CLR11-10] 98	CLCuKoV-Sha [PK-Sha-LS4-04] 97	CLCuMuV-IN [IN:Abo:03]99	CLCuMuV-IN [IN-Abo-03]99	RLCV-[IN-gp1-05] 97
C50	CLCuKoV-Sha [PK:TanJ:05] 99	CLCuKoV-Ko [PK-806b-96]100	CLCuKoV-Bu [In:Han:10] 99	CLCuKoV-Bu [IN-Sri-05] 100	CLCuKoV-Bu [PK-Veh2-04] 99	CLCuKoV-Bu [PK-Fai13-04] 100	CLCuKoV-Bu [IN-Faz-04] 99
C52	CLCuKoV-Bu [PK-RYK7-04]99	CLCuKoV-Sha [PK -TanJ-05]99	CLCuKoV-Bu [PK-RM31-11]99	CLCuKoV-Bu [PK-RYK18-12]97	CLCuKoV-Bu [PK-Veh2-04]99	CLCuKoV-Bu [PK-Fai11-12]91	CLCuKoV-Bu [PK-Fai11-12]91
C55	CLCuKoV-Bu [PK-Veh-MV2A-09] 96	CLCuKoV-Bu [PK-RM31-11]98	CLCuKoV-Bu [IN-Han-10]99	CLCuKoV-Bu [PK-RYK18-12] 97	CLCuKoV-Bu [PK-Kha5-04] 99	CLCuKoV-Bu [PK-Pun-06]99	CLCuKoV-Bu [IN:Pun:07]99
C58	CLCuKoV-Bu [PK-RYK7-04]99	CLCuKoV-Bu [IN-IARI30-12] 99	CLCuKoV-Bu [IN-Sri-05] 99	CLCuKoV-Bu [IN-Sri-05] 97	CLCuKoV-Bu [IN-Veh-06] 99	CLCuKoV-Bu [IN-Veh-06] 99	CLCuKoV-Bu [IN-Veh-06] 99
C62	CLCuKoV-Bu [PK-RYK7-04] 99	CLCuKoV-Ko[PK-Fai1] 99	CLCuKoV-Bu[IN-Sri-05]98	CLCuKoV-Bu [IN-Luc-09]95	CLCuKoV-Bu [IN-Dab3-10] 99	CLCuKoV-Bu [PK-Kha5-04]99	CLCuKoV-Bu [IN-Har-10]97
C64	CLCuKoV-Sha [PK-Sha-LS6-04]99	CLCuKoV-Sha [PK-Sha-LS4-04]100	CLCuKoV-Bu [PK-Ric-12]99	CLCuKoV-Bu [PK-RYK18-12] 97	CLCuKoV-Bu [PK-Lay20-04]99	CLCuKoV-Bu [PK-Kha5-04]99	CLCuKoV-Bu [PK-RYK18-12]98

CLCuKoV-Bu is featured by the lack of an intact TrAP encoding gene (Amrao *et al*., 2010a; Rajagopalan *et al*., 2012; Zaffalon *et al*., 2011). The absence of a complete TrAP gene in our isolate led to the prediction that this gene product may have been the avirulence determinant recognized by the resistance in cotton and that CLCuKoV-Bu overcame resistance by not expressing the complete protein (Amrao *et al*., 2010; Briddon *et al.*, 2014). Isolate C49A also lacks an intact TrAP gene, due to two stop codons in the sequence at different position. Instead of encoding a predicted 35-amino-acid residual TrAP, as was predicted for earlier CLCuKoV-Ko isolates (Amrao *et al*., 2010a; Rajagopalan *et al*., 2012), the isolate C49A encodes a predicted 39-amino-acid TrAP (Table 4.4). Isolate C49A shared high level of amino acid similarity with available amino acid sequences of the begomoviruses in the databases. All gene products of each isolate were summarized in the Table 4.6.

The available DNA A sequences of cotton infecting begomoviruses in the databases has been revised. CLCuKoV has five distinct strains; Burewala [-Bu], Kokhran [-Ko], Lucknow [Lu], Shadadpur strains [-Sha] and one has recently added, identified here as Layyah strain [-La] (Brown *et al.*, 2015). The comparison of these strains and isolates of CLCuKoV-Bu, originating from Pakistan, India, and China with their percentage nucleotide identities are given in the Table 4.5. The DNA A sequence of isolate C49 showed high levels of nucleotide sequences identities of 92.1 % with CLCuKoV-Bu, followed by CLCuKo-Sha with 91.7 %, respectively (Table 4.5). Similarly, high level of nucleotide sequence identities among other isolates was determined 96.8-98.2 %. As the maximum nucleotide identity was more than 91 % but less than 94%, so based on the current accepted species demarcation criteria (Brown *et al.*, 2012), C28, C32, C47, C50, C52, C55, C58, C62 and C64 are variants of CLCuKoV-Bu species, while isolate C49 DNA A component is apparently a new strain.

4.3.1.3. Phylogenetic analysis

A phylogenetic dendrogram was produced using MUSCLE embedded in MEGA6.0 software and the percentage of replicates tree in which the accompanying taxa clustered together in bootstrap iterations (1000 replicates).

A phylogenetic tree of ten begomovirus DNA-A components (C28A, C32A, C47A, C49A, C50A, C52A, C55A, C58A, C62A and C64A) obtained here was compared with the previously reported forty-one begomovirus full-length sequences corresponding to DNA-A component shown to be associated with CLCuD and selected other sequences were based upon highest nucleotide similarity sequences from the GenBank.

A phylogenetic dendrogram was produced based upon the alignments of selected begomoviruses with an out-group and the sequences of the isolates of present study is shown in Figure 4.5. It is evident from the tree that the isolates C28A, C32A, C47A, C50A, C52A, C55A, C58A, C62A and C64A cluster in a group with CLCuKoV-Bu, all reported to infect cotton and other malvaceous plants in this region and northwestern India, with a high bootstrap values, and more distantly associated with the other CLCuD-associated begomoviruses except for ToLCV and PaLCuV. Each group in the dendrogram was supported with high bootstrapping values. All isolates except isolate C49A in the group of CLCuKoV-Bu are shown to be very closely related with very short branches lengths. A phylogenetic analysis of isolate C49A showed it to segregate with a low bootstrap value (45%) from CLCuKoV-Bu cluster and displayed to be the most distinct isolate and lies at the basal to the rest of isolates, consistent with idea that this represent a new strain belonging to the same species.

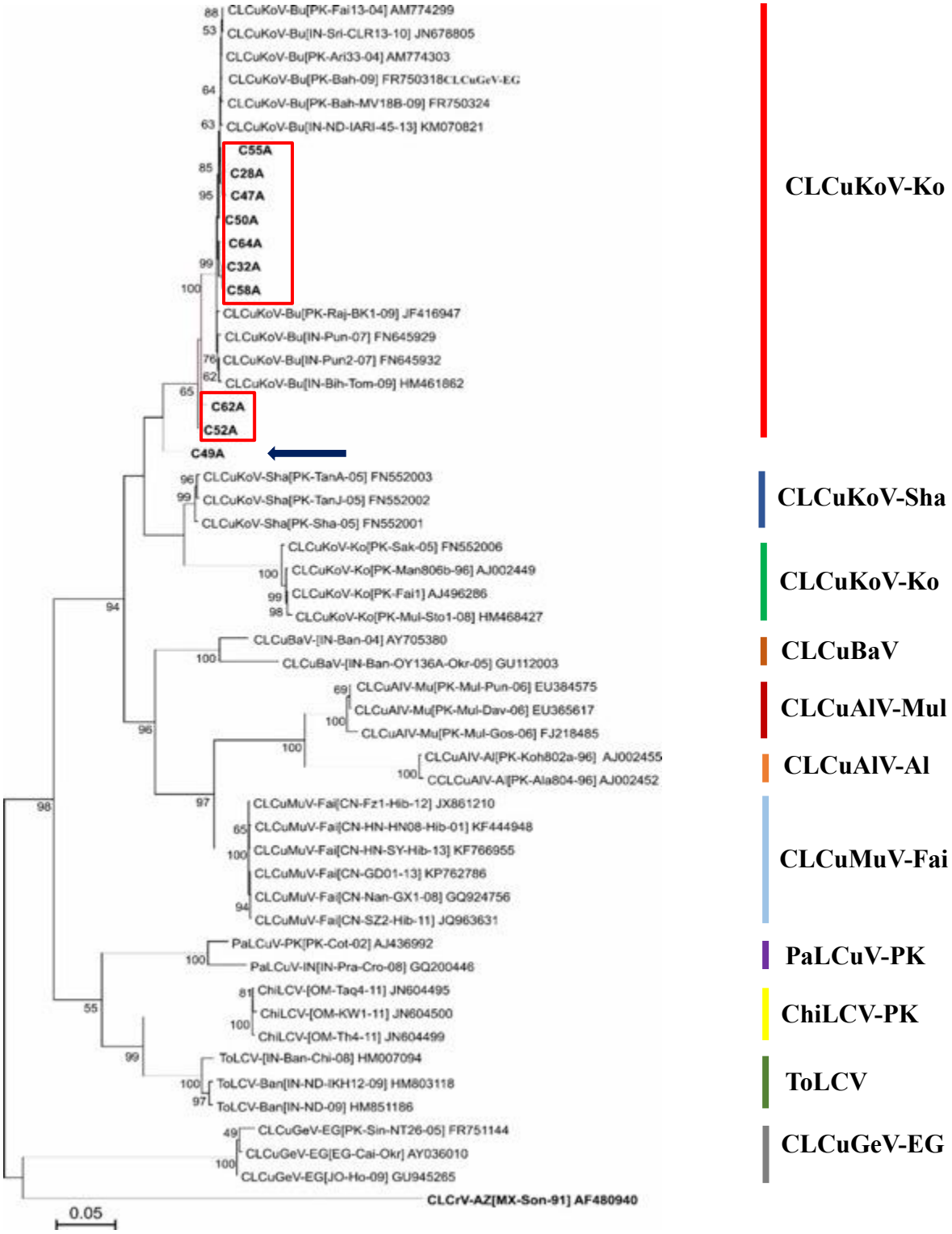


Figure 4.5. Phylogenetic tree based upon an alignment of the complete sequences of the genomes (DNA A genomic components) of fifty-two selected begomoviruses from databases, using MUSCLE method in MEGA 6.0 software Isolates characterized here are shown in bold with red boxes. Each color represents a group of single species. The tree was rooted on the complete DNA A sequence of *Cotton leaf crumple virus* (CLCrV-AZ[MX-Son-91]), a distantly related begomovirus originating from the New World. The numbers at the nodes indicates the bootstrap confidence valves (1000 replicates)

4.3.1.4. Analysis of the sequence of the DNA A component of isolate C49

The complete nucleotide sequence of isolate C49 was determined to be 2751bp in length. Initial blast of the sequence showed maximum nucleotide identity with the available sequences corresponding to CLCuKoV-Bu in the databases. Structural analysis of DNA A component of isolate C49 revealed it to have features typical of the genomes (or DNA A components) of begomoviruses originating from the OW. The sequence contain five overlapping predicted genes, two in the virion-sense strand (encoding coat protein and V2 protein) and three genes on the complementary-sense strand (encode Rep, REn, and the C4 proteins). Moreover, the virion-sense strand and complementary-sense strand are separated by a large intergenic region (LIR), was also present. The length of intergenic region of the isolate C49, however, was 441 nucleotides. The rest of isolates obtained in this study contained 455 nucleotides long LIR (Figure 4.4). The features of DNA-A component of isolate C49 is summarized in Table 4.4.

The closer inspection of the intergenic region of isolate C49 revealed a series of *cis*-acting DNA elements (involved in begomovirus replication and transcriptional regulation of *Rep* gene) (Arguello-Astorga *et al.,* 1994; 2001). All such elements are present in the sequences of all isolates including; (i) Rep-binding site (iterons), (ii) TATA boxes, iii) GC boxes, (iv) core iteron [GGAG]N conserved in all CLCV isolates, and (v) hairpin stem loop structure that contains the conserved nona-nucleotide (TAATATTAC) sequences nicked by the Rep protein to initiate DNA replication (Figure 4.4).

A table (Table 4.5) is constructed to show the percentage nucleotide sequence identity among these isolates with selected viruses, based upon most similar sequences (one isolate for each species), available in the database. In the table, isolate

C49, showed an 87.7 % to 92.1 % nucleotide sequence identity with the isolates identified in this study, followed by 92.1 % with CLCuKoV-Bu reported from India, 91% with CLCuKoV-Sha and 89% with CLCuMuV-Fai isolates, respectively. Minimum nucleotide sequence identity for isolate C49 was observed CLCuGeV-EG (73 %).

Nucleotide sequence distance and phylogenetic analysis of isolate C49 showed its distinct position. The sequence of isolate C49 showed a closed relation with the CLCuKoV-Bu strains in the databases. So, further analysis was done to characterize this isolate. Sequence alignment of all CLCuKoV-Bu isolates available in the databases (a total of 66 sequences) showed pairwise identities ranging from 92.8 to 100 %, with a distinct peak at ~98.5 % (Figure 4.6). Furthermore, percentage identity values below 93 % identified for the three pairs (JF502353-JF510461, 92.5%; JF502353-JF510458, 92.8% and JF502353-FHM461867, 92.8 %). The result of pairwise identity for isolate JF502353 indicates that this is an unusual CLCuKoV-Bu isolate, but the fact for the remaining 63 pairwise comparisons the identity values were found >94% (the demarcation threshold for distinguishing begomovirus variant from strain) (Brown *et al.*, 2015).

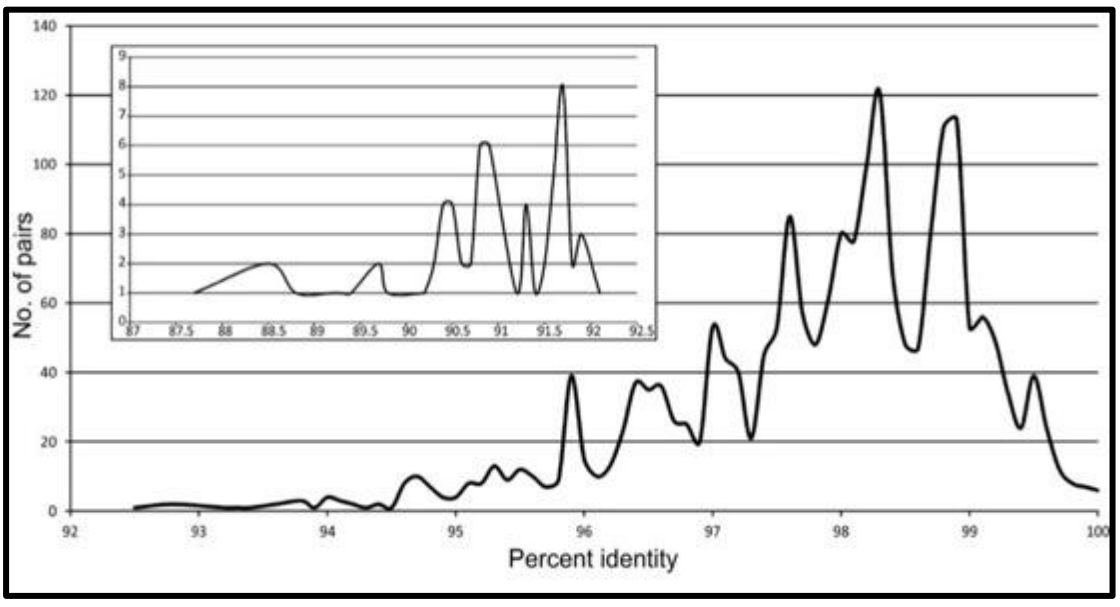


Figure 4.6. Sequence relatedness of isolate C49 DNA A to CLCuKoV-Bu. Graphical illustration of pairwise nucleotide sequence identities of 66 CLCuKoV-Bu isolates and (inset) pairwise comparison of the sequence of C49 DNA A to 66 CLCuKoV-Bu isolates.

4.3.1.5. Recombination analysis of isolate C49 DNA A component

Blast and phylogenetic analysis isolate C49A showed a distinct position in the table and phylogenetic tree. Plotcon and recombination analysis was performed for the characterization of this distinct isolate. Previous reports indicating that CLCuKoV-Bu is a recombinant, comprising of sequences derived from two old world begomoviruses, *Cotton leaf curl Kokhran virus*- strain Kokhran (CLCuKoV-Ko; covering the virion sense genes) and *Cotton leaf curl Multan virus* (CLCuMuV; spanning the complementary sense genes) (Amrao *et al.*, 2010). Closer analysis of isolate C49A also showed a chimeric nature (involvement of CLCuMuV and CLCuKoV-Ko), using Plotcon (Rice, 2000) and RDP4 (Martin *et al.*, 2015).

A comparison of individual sequence alignments between CLCuKoV-Bu and CLCuMuV or CLCuKoV-Ko using Plotcon is shown in Figure 4.7. This elucidate the high level of nucleotide sequence similarity between CLCuKoV-Bu and CLCuKoV-Ko in the virion sense sequences but greater similarity between CLCuKoV-Bu and CLCuMuV for the complementary-sense sequences. Isolate C49A was individually aligned and compared with CLCuKoV-Ko, CLCuMuV and CLCuKoV-Bu, indicating recombinant nature from the same viruses from which CLCuKoV-Bu was evolved but looking distinct in breakpoints locations graphically.

Chimeric nature of the isolate C49A was further explored with RDP. RDP results provided strong evidence in favor of possible recombination in the isolate C49A genome at different places, suggested a more origin involving possible back recombination with CLCuMuV (Figure 4.8). A recombinant event was foreseen in a fragment of 1, 238 nucleotides in size, positioned spanning nucleotides 61 and 1, 300

(breakpoints). The event was indicated and supported by all the given parameters used in RDP4 with identified P-valve as p = 1.365 x 10^{-97}. The analysis highlighted the involvement of CLCuMuV as major parent (MP), while *Cotton leaf curl Kokhran virus-Kokhran* (CLCuKoV-Ko) was determined as minor parent.

The second possible recombination event was predicted in a fragment of 400 nucleotides, exist between nucleotides 901 and 1301 (breakpoints). The event was indicated and supported by all possible parameters used in RDP4. The identified P-valve was determined as p = 3.01 x $10^{-45.}$ Some minor traces were also predicted in the genome of C49A nucleotides existed between 1553 and 1615 nucleotides with determined P-valve, p =1.695 x 10^{-5}. However, the fact that the 5' recombination breakpoints of the two events suggests that rather than originating from back recombination of CLCuKoV-Bu with CLCuMuV, isolate C49A (CLCuKoV-La) resulted ratherthan from a distinct recombination probably occurred between CLCuMuV and CLCuKoV-Ko. The possible map of recombinant molecule of isolate C49 DNA A is shown in the Figure 4.9.

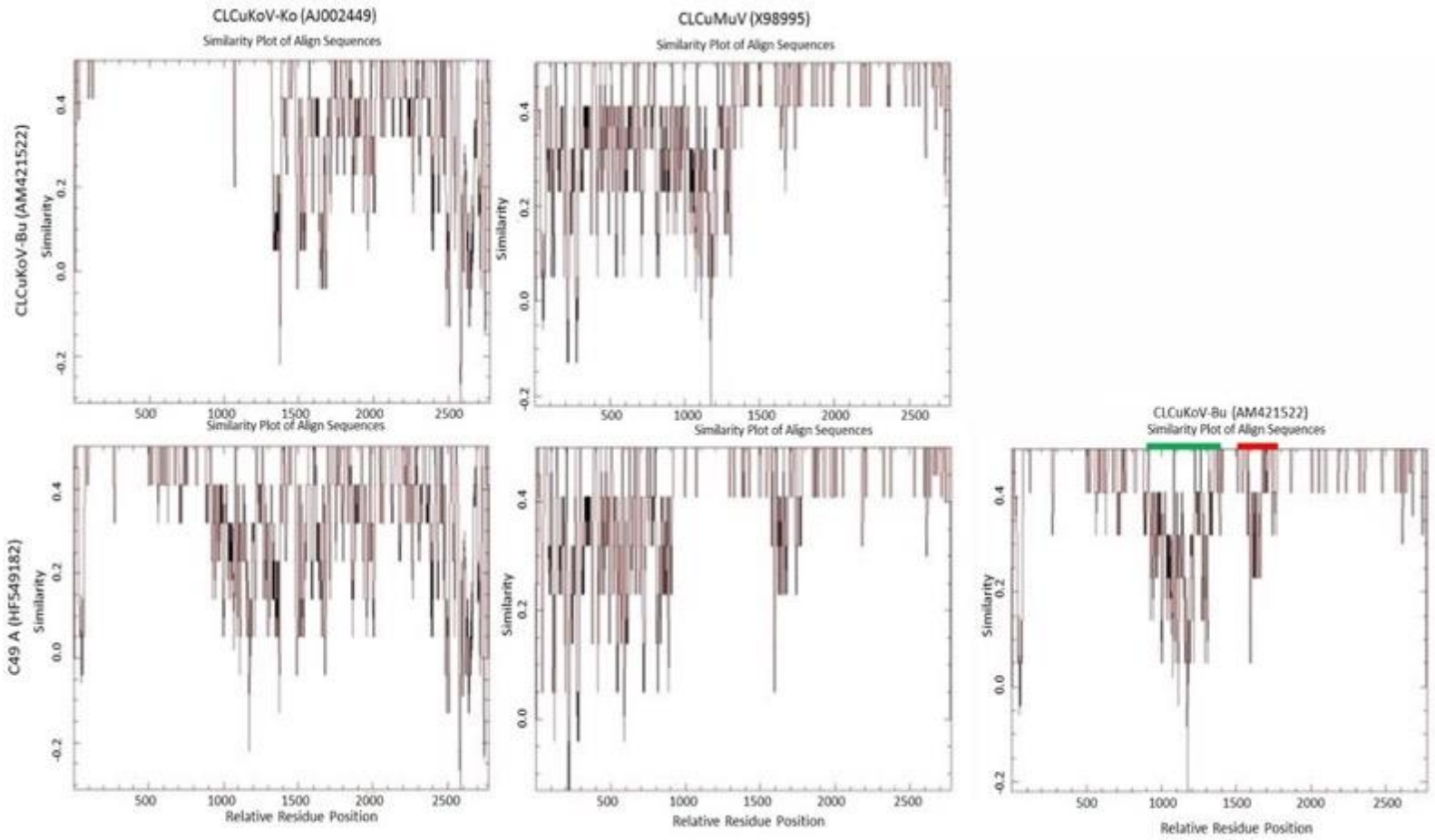


Figure 4.7. Pairwise Plotcon comparisons of the sequence relatedness of isolate C49A, CLCuMuV, CLCuKoV-Ko and CLCuKoV-Bu. The database accession number of each isolate used in the comparison is given. For the C49A comparison with CLCuKoV-Bu the green bar highlights the sequences differing between the two isolates that suggest the recombinant fragment originating from CLCuMuV is smaller for C49A than for CLCuKoV-Bu. The red bar indicates the possible second recombinant fragment which RDP4 suggests originates from CLCuMuV.

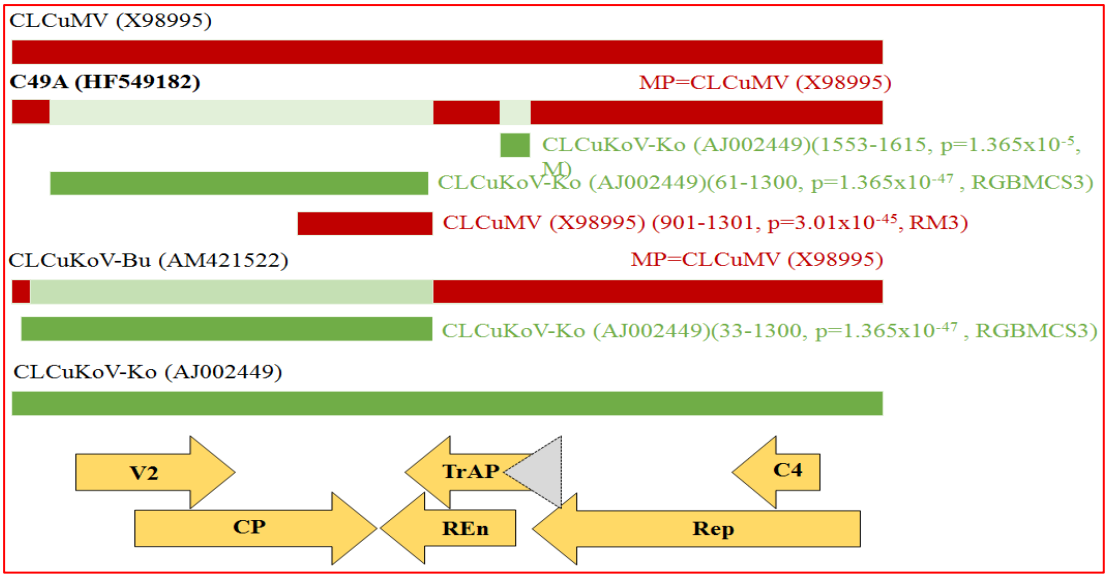


Figure 4.8. RDP analysis of C49A and CLCuKoV-Bu. For each isolate the database accession number is given. The major parent (MP) for each recombinant sequence is indicated. For each recombinant fragment the origin (minor parent), position (nucleotide coordinates of the 5' and 3' breakpoints), highest p-value and methods supporting the event are indicated. The methods are RDP (R), GENECONV (G), Bootscan (B), Max Chi (M), Chimera (C), SiScan (S) and 3Seq (3)

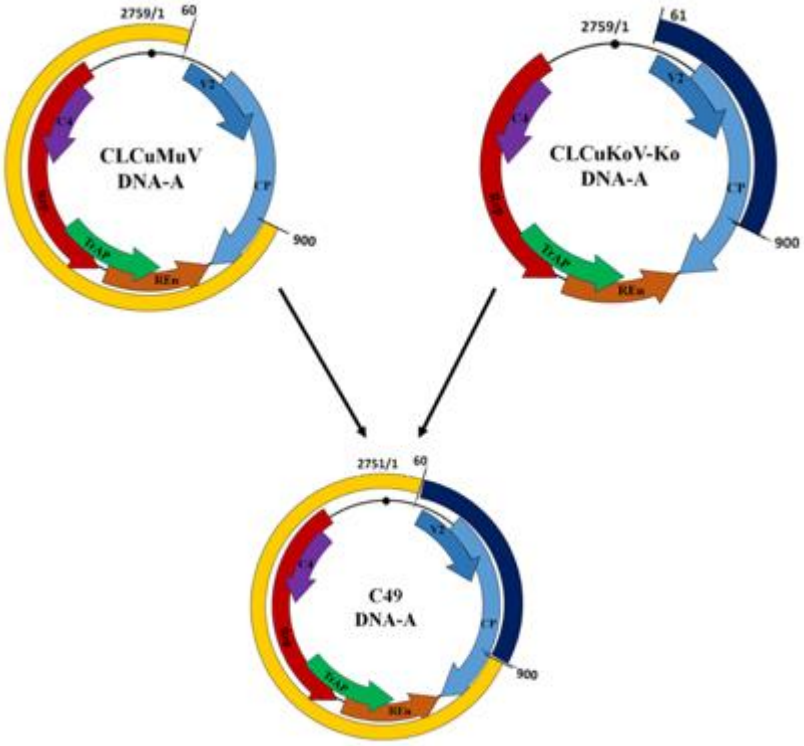


Figure 4.9. Map of an isolate C49 DNA A showing the sequences originating from *Cotton leaf curl Multan virus* (CLCuMuV) and *Cotton leaf curl Kokhran virus* - strain Kokhran (CLCuKoV-Ko). The fragments are colored as Yellow, originating from CLCuMuV, and Dark Blue, originating from CLCuKoA-Ko.

4.3.1.6. Closing remarks for virus

The virus characterization described here has concludes on one new strain and nine newly characterized isolates. Brown *et al*., 2012 proposed new guidelines for the updated nomenclature of geminiviruses. Table 4.7 list the isolate descriptors proposed to be used that describe the virus isolates characterized in this study.

Table 4.7. Begomovirus sequences with the respective acronym and GenBank accession number reported from CLCuD infected cotton

Plant sample code	Virus name	Acronym	GenBank Accession No.
C-28	Cotton leaf curl Kokhran virus-Burewala [Pakistan-Faisalabad-C28-2010]	CLCuKoV-Bu[PK-Fai-C28-10]	HF549180
C-32	Cotton leaf curl Kokhran virus – Burewala [Pakistan-Faisalabad-GMT2-2010]	CLCuKoV-Bu[PK-Fai-GMT2-10]	HF549181
C-47	Cotton leaf curl Kokhran virus - Burewala [Pakistan-Dhok Ali Khan-GMT63-2011]	CLCuKoV-Bu[PK-DAK-GMT63-11]	KR816002
C-49	Cotton leaf curl Kokhran virus-Layyah [Pakistan-Layyah-2011]	CLCuKoV-La[PK-Lay-11]	HF549182
C-50	Cotton leaf curl Kokhran virus-Burewala [Pakistan:Bhakkar-GMT-16:2011]	CLCuKoV-Bu[PK:Bha-GMT-16-11	HF549183
C-52	Cotton leaf curl Kokhran virus-Burewala[Pakistan-Multan-GMT66-2011]	CLCuKoV-Bu[PK-Mul-GMT66-11]	KR815998
C-55	Cotton leaf curl Kokhran virus-Burewala[Pakistan- Bahawalnagar-GMT69-2011]	CLCuKoV-Bu[PK-BWN-GMT69-11]	KR815999
C-58	Cotton leaf curl Kokhran virus - Burewala [Pakistan-Okara-GMT-18-2011	CLCuKoV-Bu[PK-Oka-GMT-18-11]	HF549184
C-62	Cotton leaf curl Kokhran virus-Burewala[Pakistan-Khanewal-GMT72-2011]	CLCuKoV-Bu[PK-KNW-GMT72-11]	KR816000
C-64	Cotton leaf curl Kokhran virus-Burewala[Pakistan-Rahim Yar Khan-GMT75-2013]	CLCuKoV-Bu[PK-RYK-GMT75-13]	KR816001

4.3.2. Analysis of the cloned betasatellite components of begomoviruses

Cotton hosts a complex of begomoviruses. Diagnostic PCR confirmed the presence of begomovirus in the expected samples which are usually associated with satellite molecule(s). For the presence of betasatellite in nucleic acid of suspected sample, universal primers set Beta01/Beta02 (Briddon *et al.*, 2002) was used in the thermocycler. Full-length betasatellite sequences of ten isolates (C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β and C64β) were amplified and sequenced in both orientations.

4.3.2.1. Analysis of the sequence of the betasatellite components of isolates C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β and C64β

The complete nucleotide sequences of the satellite C32β and C52β were determined to be 1,349 bp in length while the sequences obtained for C47β, C49β and C64β were 1350 bp long. Similarly, C28β, C50β and C55β were determined as 1,351 bp, C58β as 1,371 bp, while the largest satellite was C62β with 1372bp.

These betasatellites contains all the three basic feature of a satellite (Figure 1.4C, middle); SCR which were approximately 118 nts in length and contains the hairpin structure TAATATTAC at the 3' end. A region of sequence rich in Adenine residues (A) was located between coordinates 664 to 983 with an overall A contents of 49 to 54%. The βC1 genes of all the ten betasatellites encode 118 amino acids on the complementary sense strand with a predicted molecular weight of ~13.7 kDa (Table 4.8).

The betasatellites isolated from symptomatic cotton samples collected from different locations in Punjab province shared 87.9-99.0 % nucleotide sequence identities among them which shows that these are isolates from the same species. The sequences of these betasatellites shared identitie CLCuKoV-Bu [IN-Veh-06] 99s from 89.6 to 92.2 % with reported Cotton leaf Multan betasatellites, followed by MaYVB (60.1-64.2 %). Based on the present accepted species demarcation threshold of 78 % nucleotide sequence identities criteria for classifying betasatellites (Briddon *et al.*, 2008), these betasatellites identified in this study are isolate of the same species. Amino acid identity of βC1 gene for all isolates were determined which were 99 % with already reported betasatellite molecules identified in cotton. The percentage nucleotide sequence identities of these betasatellites with reported betasatellites available in the databases are shown in Table 4.9.

There are two types/versions of CLCuMB associated with CLCuD affecting cotton in Pakistan and India. SCR of CLCuMB having approximately 100 nucleotides fragment derived from other betasatellite associated with Tomato leaf curl disease (ToLCB), refer this betasatellite as the "Burewala" strain, of CLCuMB (CLCuMB^{Bur}) and the non-recombinant CLCuD-associated betasatellite as the "Multan" strain of CLCuMB (CLCuMB^{Mul}) reported during 1st epidemic. The SCR region of the all betasatellite isolates, identified here, falls in the category that have a recombinant fragment originating from Tomato leaf curl betasatellite as the characteristic of the CLCuMB associated with resistance breaking, known as the "Burewala" strain of CLCuMB (CLCuMB^{Bur}) (Amin *et al.*, 2006; Azar *et al*., 2012).

Table 4.8. Genomic features of CLCuD associated betasatellite component

Isolate	ORFs	Start codon	Stop codon	Predicted size (no. of amino acid)	Predicted molecular weight (kDa)	Predicted Amino acid Sequence identity
C28	βC1	550	194	118	13.6	CLCuMB[IN-His3-04] 99
C32	βC1	551	195	118	13.6	CLCuMB[IN-Sri-08] 99
C47	βC1	551	195	118	13.7	CLCuMB[PK-RYK5-12] 99
C49	βC1	551	195	118	13.7	CLCuMB[PK-AS6-10] 99
C50	βC1	551	195	118	13.7	CLCuMB[PK-AS6-10] 100
C52	βC1	551	195	118	13.6	CLCuMB[PK-AS47-10] 99
C55	βC1	551	195	118	13.7	CLCuMB[PK-AS21-10] 99
C58	βC1	551	195	118	13.6	CLCuMB[IN-Sri-08] 99
C62	βC1	551	195	118	13.6	CLCuMB[PK-Fai-BarbB-11] 99
C64	βC1	551	195	118	13.7	CLCuMB[PK-AS47-10] 99

Table 4.9. Percentage nucleotide sequence identity between the complete sequences of the betasatellite of isolates C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β and C64β with selected betasatellites in the databases. Sequences were selected of the most similar species from the Blast comparison.

Isolates of betasatellite	PaLCuB -[IN-Coi-BG-CBE-12]	ChiLCB -[IN-Joh-04]	CrYVMB -[IN-Bhub-OYBHU-06]	AYLCB -[IN-WSFBI-10]	MaYVB -[CN-Yun-Y217-05]	ToYLCThB -[IN-Ramg-07]	CLCuMB -[IN-bt18]	C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β, C64β
C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β, C64β	39.3-40.6	39.0-40.5	37.8-42.7	36.7-38.8	60.1-64.2	39.4-42.3	**89.6-92.2**	**87.6-99.0**
CLCuMB-[IN-bt18]	38.3	38.9	41.1	37	60.7	39.3	**100**	
ToYLCThB-[IN-Tamg-07]	45.4	47.4	51.6	47.9	38.4	**100**		
MaYVB-[CN-Yun-Y217-05]	36.1	38.7	40.1	36.3	**100**			
AYLCB-[IN-WSFBI-10]	45.7	55.7	47.9	**100**				
CrYVMB-[IN-Bhub-OYBHU-06]	52.7	45.6	**100**					
ChiLCB-[IN-Joh-04]	53.1	**100**						
PaLCuB-[IN-Coi-BG-CBE-12]	**100**							

Percentage valves more than 78% (species demarcation threshold for betasatellite, Briddon *et al.*, 2008) are in bold.

4.3.2.2. Phylogenetic analysis of betasatellites

Phylogenetic analysis of the identified isolates corresponding to Cotton leaf curl Multan betasatellite (C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β and C64β) was made by comparing them with thirty-seven (37), already reported, betasatellites randomly selected from the databases.

Phylogenetic tree was constructed based upon an alignment of the complete nucleotide sequences of isolates identified in a begomoviral complex from Punjab region with selected betasatellites available in the database is shown in Figure 4.10.

Phylogenetic dendrogram shows that the betasatellite C28β, C32β, C47β, C49β, C50β, C52β, C55β, C58β, C62β and C64β clustered with the betasatellite group corresponding to CLCuMB with a high bootstrap values. There are two species/version of CLCuMB; $CLCuMB^{Mul}$ (non-recombinant) and $CLCuMB^{Bur}$ (recombinant) (previously described in section 4.3.2). It is evident from the tree that all identified betasatellite components clustered with the recombinant version of $CLCuMB^{Bur}$ in a close cluster, previously reported from Pakistan and India infecting plants within the family Malvaceae.

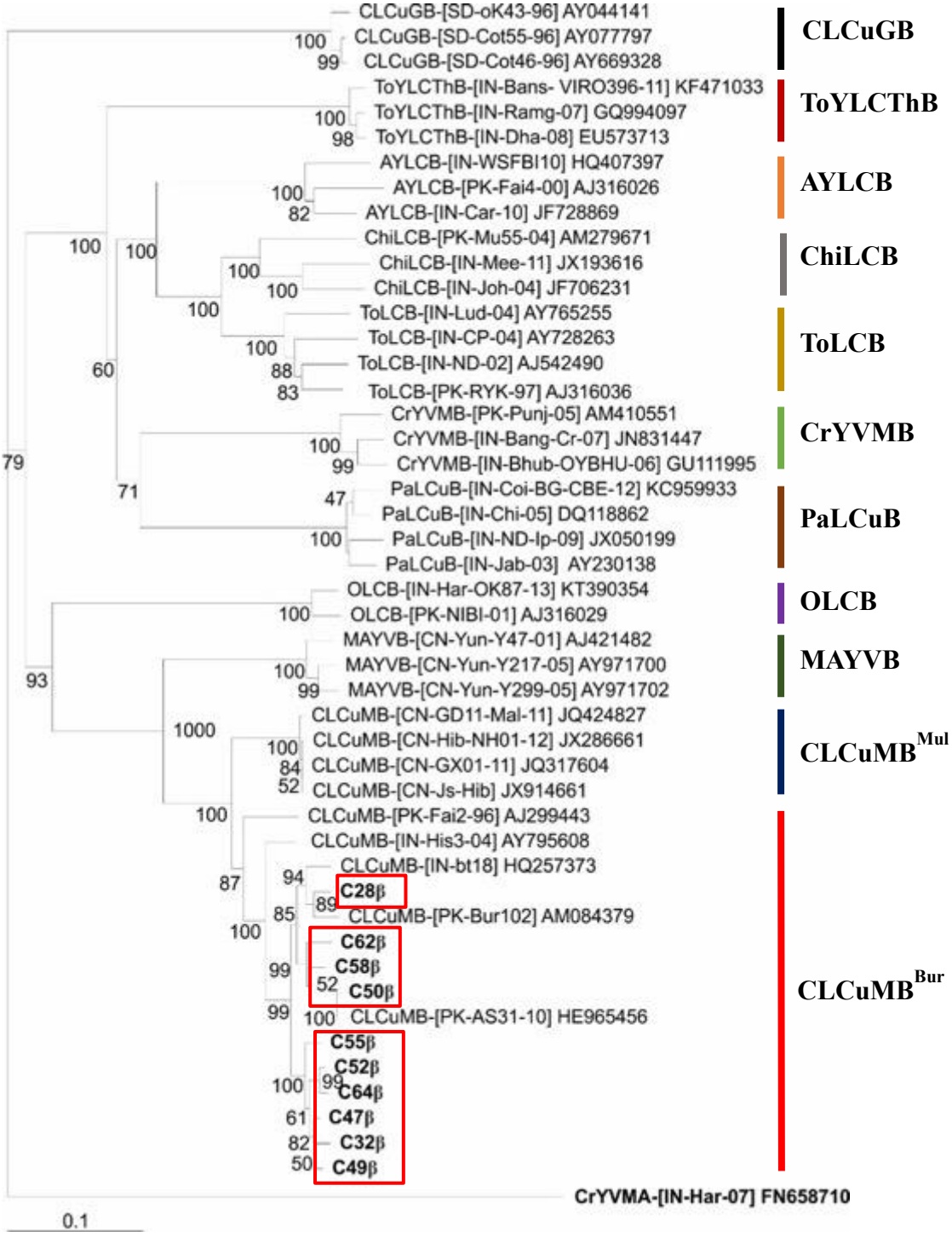


Figure 4.10. Phylogenetic analysis of betasatellites. Neighbor joining phylogenetic dendrogram was constructed, based upon alignment of the complete nucleotide sequences of the forty-seven betasatellites clones available in the database. In each case, the isolate descriptor and accession number are given. Betasatellites cloned and characterized here are shown in bold with red boxes. Each color represents a group of single species. The numbers at the nodes indicates the bootstrap confidence valves (1000 replicates). The alignment was arbitrary rooted on an out-group the complete alphasatellite sequence of Croton yellow vein mosaic alphasatellite (CrYVMA-[IN-Har-07]), a distantly related begomovirus associated satellite originating from the Old World

4.3.2.3. Closing remarks for betasatellite

The betasatellites described here concluded with ten newly characterized isolates identified in this study. Briddon *et al*., 2008 proposed guidelines for the updated nomenclature of betasatellites. Table 4.10 list the isolate descriptors proposed for the betasatellites characterized in this study.

Table 4.10. Begomovirus associated DNA betasatellites with their respective acronym and GenBank accession number reported from CLCuD infected cotton

Betasatellite	Acronym	Accession
Cotton leaf curl Multan betasatellite-[Pakistan-Faisalabad-C28-2010]	CLCuMB-[PK-Fai-C28-10]	HF549185
Cotton leaf curl Multan betasatellite-[Pakistan-Faisalabad-C32-2010]	CLCuMB-[PK-Fai-C32-2010]	HF549186
Cotton leaf curl Multan betasatellite-[Pakistan-Dhok Ali Khan- GMT62- 2011]	CLCuMB-[PK-DAK-GMT62-11]	KR816003
Cotton leaf curl Multan betasatellite-[Pakistan-Layah-C49-2010]	CLCuMB-[PK-Lay-C49-10]	HG000665
Cotton leaf curl Multan betasatellite[PK-Bakkar-C50-2011]	CLCuMB-[PK-Bak-C50-11]	HF549187
Cotton leaf curl Multan betasatellite-[Pakistan-Multan-GMT65-2011]	CLCuMB-[PK-Mul-GMT65-11]	KR816004
Cotton leaf curl Multan betasatellite-[Pakistan-Bahawalnagar-GMT68- 2011]	CLCuMB-[PK-BWN-GMT68-11]	KR816005
Cotton leaf curl Multan betasatellite-[Pakistan-Okara-C58-2011]	CLCuMB-[PK-Oka-C58-11]	HF549188
Cotton leaf curl Multan betasatellite-[Pakistan-Khanewal-GMT71-2011]	CLCuMB-[PK-KNW-GNT71-11]	KR816006
Cotton leaf curl Multan betasatellite-[Pakistan-Rahim Yar Khan-GMT74-2013]	CLCuMB-[PK-RYK-GMT74-13]	KR816007

4.3.3. Analysis of the cloned alphasatellite - components of isolates C28α, C47α, C49α, C50α1, C50α2, C52α, C55α C58α, C62α and C64α

After confirmation with diagnostic PCR for DNA A, the presence of associated alphasatellite in the nucleic acid of all extracted samples were also processed for amplification of full-length by using specific abutting primer pairs AlphaF/AplhaR (Zia-ur-Rahman *et al.,* 2013). In this study, ten full-length clones (C28α, C47α, C49α, C50α1, C50α2, C52α, C55α C58α, C62α and C64α) were selected for sequencing in both orientations.

4.3.3.1. Analysis of the sequence of alphasatellite components of isolates C28α, 47α, C49α, C50α1, C50α2, C52α, C55α C58α, C62α and C64α

Sequence analysis of the obtained isolates revealed that the total genome lengths of begomovirus associated alphasatellites were 1315-1374 nts.

The sequences of all isolates obtained here, except isolate C50α2, signifies the typical characteristics of alphasatellites (previously known as DNA1), consisting of an Adenine rich region (52-58%, coordinates 1100-1260), a predicted hairpin-loop structure with the nona-nucleotide TAGTATTAC forming part of the loop and a single conserved gene encoding a Rep in the virion-orientation to govern self- replication (coordinates 77-1024) with the capacity to code a 315 amino-acid protein (Table 4.11).

An initial comparison of the ten sequences obtained with sequences in the databases using BAST indicated them to be most similar to begomoviruses associated alphasatellite previously identified in cotton and other hosts of begomovirus in the old world. An alignment was then produced with all isolates identified here along with selected alphasatellites from other hosts to determine pairwise similarities (Table 4.12)

This showed that the isolates have the highest nucleotide sequence identities of 87.5-99 .% among them, excluding isolate C50α2, followed by CLCuMA (83.9-88.9%)

and lowest nucleotide sequence identity was observed for MalYVMA (67.6-70.7%). On the contrary, the isolate C50α2 showed the highest level of nucleotide sequence similarities.to isolates (C28α, C47α, C49α, C50α1, C52α, C55α C58α, C62α and C64α) (56.4-66.5%), with CLCuMA (55.9 %), followed by CLCuShA (48.8%), respectively.

To compare isolate C50α2 with other reported alphasatellites in the databases, lowest level of nucleotide similarity was noted for SiYVCNA (40.5 %). Based on the currently assumed species demarcation threshold of 83% nucleotide sequence identity, these findings indicate that alphasatellites identified in the resistant cotton varieties corresponding to the isolate of the same species except isolate C50α2.

Table 4.11. Characteristic features of begomovirus associated alphasatellite isolates obtained from *Gossypium hirsutum.*

Isolate	ORFs	Start Codon	Stop codon	Size (no. of amino acid)	Molecular weight (kDa)
C28α	Rep	77	1024	315	36.6
C47α	Rep	77	1024	315	36.6
C49α	Rep	77	1024	315	36.7
C50α1	Rep	77	1024	315	36.7
C50α2	Alpha-Rep	77	475	132*	15.5
	Virus- Rep	679	446	77*	9.0
	Virus- C4	765	514	83*	9.5
C52α	Rep	77	1024	315	36.5
C55α	Rep	77	1024	315	36.6
C58α	Rep	77	1024	315	36.5
C62α	Rep	77	1024	315	36.7
C64α	Rep	77	1030	317	36.8

Table 4.12. Percentage nucleotide sequence identity between the complete sequences of alphasatellite isolates C28α, C47α, C49α, C50α1,C502α, C52α, C55α, C58α, C62α and C64α with selected alphasatellites in the database. Sequences were selected of the most similar species from the blast comparison.

Isolates	ToLCA-[IN-LN-11]	SiYVCNA-[CN-Y340-10]	MalYVMA-[CN-Hn39-06]	GDaSA-[PK-Tom-2-06]	MeYMA-[IN-10b-RCA-al-F-07]	CLCShA-[PK-DSC-07]	CLCuMA-[PK-AS1-01]	C28α, C47α, C49α, C50α1, C52α, C55α, C58α, C62α, C64α	C50α2
C50α2	42.5	40.5	43.0	41.0	47.7	48.8	**55.9**	**56.4-66.5**	**100**
C28α,C47α,C49α,C50α1, C52α, C55α, C58α, C62α, C64α	69.7-72.1	67.2-70.8	67.6-70.7	67.1-71.0	74.0-78.1	75.1-79.0	**83.9-88.9**	**87.5-100**	
CLCuMA-[]PK-AQ16-11	71.3	70.4	69.8	69.9	76.6	77.7	**100**		
CLCuShA-[PK-DSC-07]	65.0	76.2	73.9	77.9	80.9	**100**			
MeYMA-[IN-10b0RCA-al-F-07]	69.3	70.2	73.3	72.6	**100**				
GDaSA-[PK-Tom-2-06]	64.9	73.8	68.5	**100**					
MalYVMA-[CN-Hn39-06]	65.6	74.6	**100**						
SIYVCNA-[CN-Y340-10]	76.2	**100**							
ToLCA-[IN-LN-11]	**100**								

Percentage values more than 83% (species demarcation threshold for alphasatellite, Mubin *et al.*, 2009 are in bold

4.3.3.2. Phylogenetic analysis of alphasatellite isolates

A phylogenetic dendrogram was produced using the Neighbor-Joining (NJ) algorithm in Clustal X2 software and the percentage of replicates tree in which the accompanying taxa clustered together in bootstrap iterations (1000 replicates) were shown next to the nodes.

Ten begomovirus/betasatellite complex associated alphasatellites (C28α,C32α, C47α, C49α, C50α1,C50α2, C52α, C55α C58α, C62α) identified here were compared with the thirty-seven already reported alphasatellite complete sequences selected based on their high nucleotide similarity sequences identified from the GenBank-NCBI.

A phylogenetic analysis based on full-length genome sequences of alphasatellites components revealed that C28α, C32α, C47α, C49α, C50α1, C50α2, C52α, C55α C58α, C62α grouped in a close cluster with CLCuMA reported to infect cotton and other malvaceous plants in this region and northwestern India, with a high bootstrap values, and more distantly associated with the other groups of begomoviruses-associated alphasatellites (Figure 4.11). Each group in the dendrogram indicated a well support of bootstrap values. All alphasatellites except component C50α2 in the group of CLCuMB showed to be very closely related with very short branches lengths. A phylogenetic analysis of component C50α2 showed it to segregate to a distinct position in the dendrogram, without clustering in any group, indicating as a divergent entity from all previously reported alphasatellites in the database.

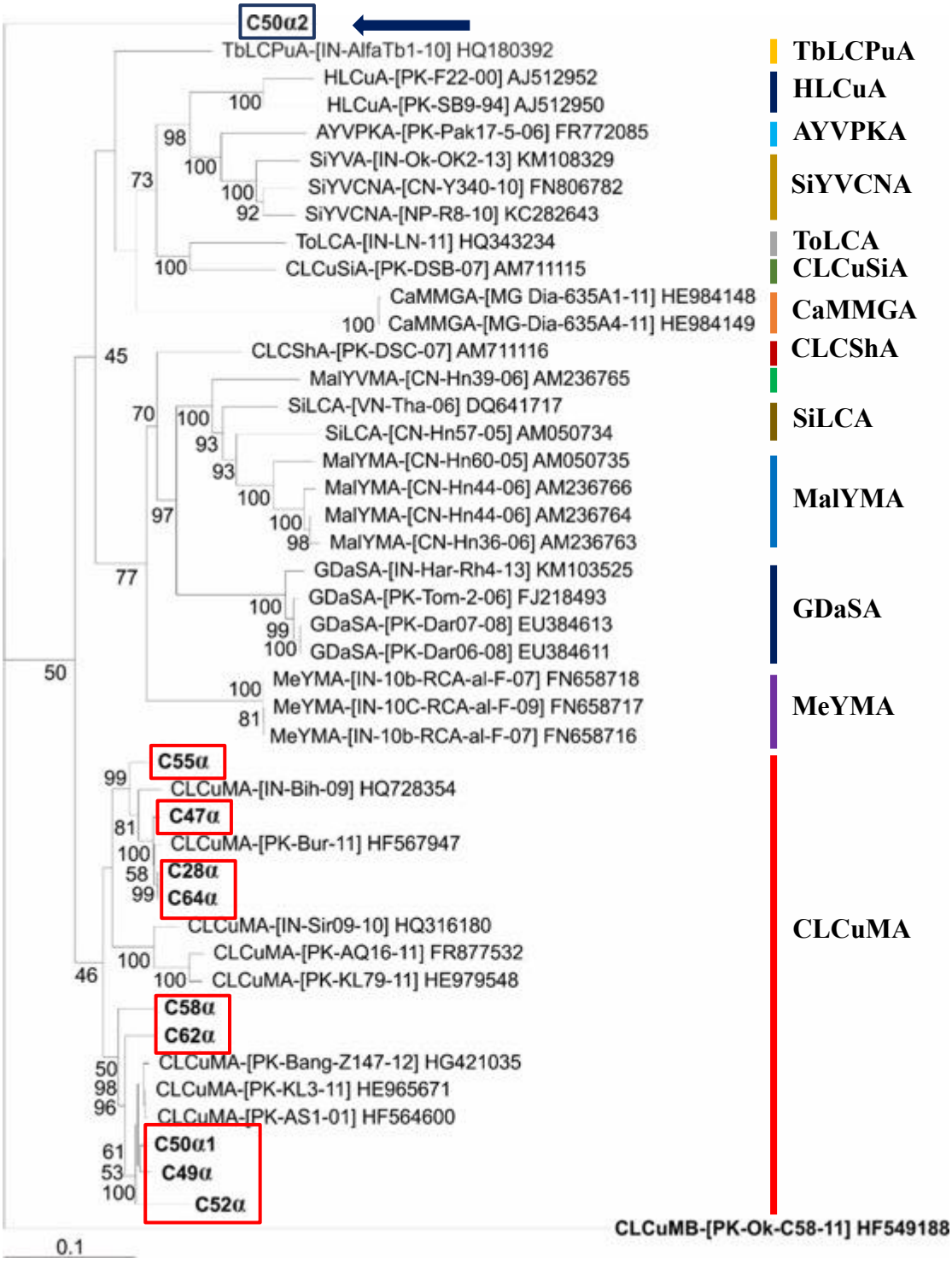


Figure 4.11. Phylogenetic analysis of alphasatellites. Neighbor joining phylogenetic dendrogram was constructed, based upon alignment of the complete nucleotide sequences of the forty-five alphasatellites clones available in the database. In each case, the isolate descriptor and accession number are given. Alphasatellites cloned and characterized here are shown in bold with red boxes. Each color represents a group of single species. The numbers at the nodes indicates the bootstrap confidence values (1000 replicates). The alignment was arbitrary rooted on an out-group the complete betasatellite sequence of Cotton leaf curl Multan Betasatellite (CLCuMB-[PK-OK-C58-11]), a distantly related betasatellite originating from the OW

4.3.3.3. Analysis of the sequence of the alphasatellites components of isolates C50α1 and C50α2

The expected size products of both DNA-A and its associated betasatellite were amplified from the nucleic acid extracted from sample C50. An attempt for the presence of a cloning of third component, usually associated with begomovirus-betasatellite complex, a set of abutting primers AlphaF/AlphaR (Table 3.1), were used to amplify and cloned.

Two full-length clones C50α1 and C50α2 were cloned and sequenced in both orientations, sequences were determined to be 1365 and 1315 nucleotides in length. The sequence of isolate C50α1 has an architecture typical of an alphasatellite, comprising of an A-rich region (53% A; coordinates 1107-1260), a predicted hairpin-loop structure with the nonanucleotide TAGTATTAC forming part of the loop and a single conserved gene encoding a Rep in the virion-sense orientation (coordinates 77-1024) with the capacity to encode a 315 amino-acid protein (Briddon *et al*. 2004) (Table 4.11). Comparison of the sequence of C50α1 with full-length alphasatellite sequences available in the databases showed that it has high levels of nucleotide sequence identity to isolates of CLCuMA with the highest identity (98.9%) to CLCuMA-[PK-AS1-01] (HF564600))(Table 4.12). Phylogenetic analysis of C50α1 showed this to segregate with isolates of CLCuMA, confirming the identification of this as an isolate of CLCuMA (Figure 4.11).

The second molecule cloned from the same sample using the alphasatellite primers, component C50α2, shareed only 79% nucleotide sequence identities with C50α1. Based on the sequence comparison of all components obtained for this study, component C50α2 was noted distinct among these components. This distinctness was further confirmed by phylogenetic analysis, when it segregates with a separate group

with a low boot strap values. Closer inspection of the component was done, revealed a chimeric nature of the sequence. Between coordinates 954 and 412 (fragment B in Figure 4.12) the sequences of C50α2 and C50α1 are nearly identical with just a single nucleotide change at position 340. Most of the remaining C50α2 sequence (fragment A in Figure 4.12) instead derives from CLCuKoV-Bu; the sequence between coordinates 523 and 953 showing 99.5% identity with an isolate of the "Layyah" strain of CLCuKoV (CLCuKoV-Lay; HF549182) but only 96.1% identity to C50A. The sequence of isolate C50α2 between coordinates 413 and 522 (fragment C in Figure 4.12) showed 92% nucleotide sequence identity to an isolate of Acalypha yellow vein mosaic alphasatellite originating from India (FN658711) but inserted in reverse orientation. C50α2 thus contains two origins of replication, that typical of an alphasatellite with the hairpin structure containing the TAGTATTAC nonanucleotide sequence (arbitrarily shown at position zero in Figure 4.12), and a begomovirus *ori* with a TAATATTAC nonanucleotide sequence (coordinates 836 to 844). The molecule also contains two truncated begomovirus genes (3' truncated Rep and C4 genes in the complementary-sense) and a 3' truncated alphasatellite Rep gene (Table 4.11, Figure 4.12).

Sequence comparison of the component C50α2 with full-length alphasatellite sequences available in the databases showed that it has high levels of nucleotide sequence identity to the identified components (C28α, C47α, C49α, C50α1, C52α, C55α C58α, C62α) with 56.4 -66.4 %, and CLCuMA with the highest identity (59.5%) to CLCuMA-[PK-AS1-01] (HF564600). Based on the currently applicable species demarcation threshold of 83% nucleotide sequence identity, these findings indicate that component C50α1 of alphasatellites identified in the resistant cotton varieties corresponding to the isolate of the same species, while component C50α2 is apparently

a recombinant molecule. The results were further confirmed by phylogenetic analysis of component C50α2 showed this to segregate with separate group, confirming the identification of this as a new recombinant component from isolate C50 (Figure 4.11).

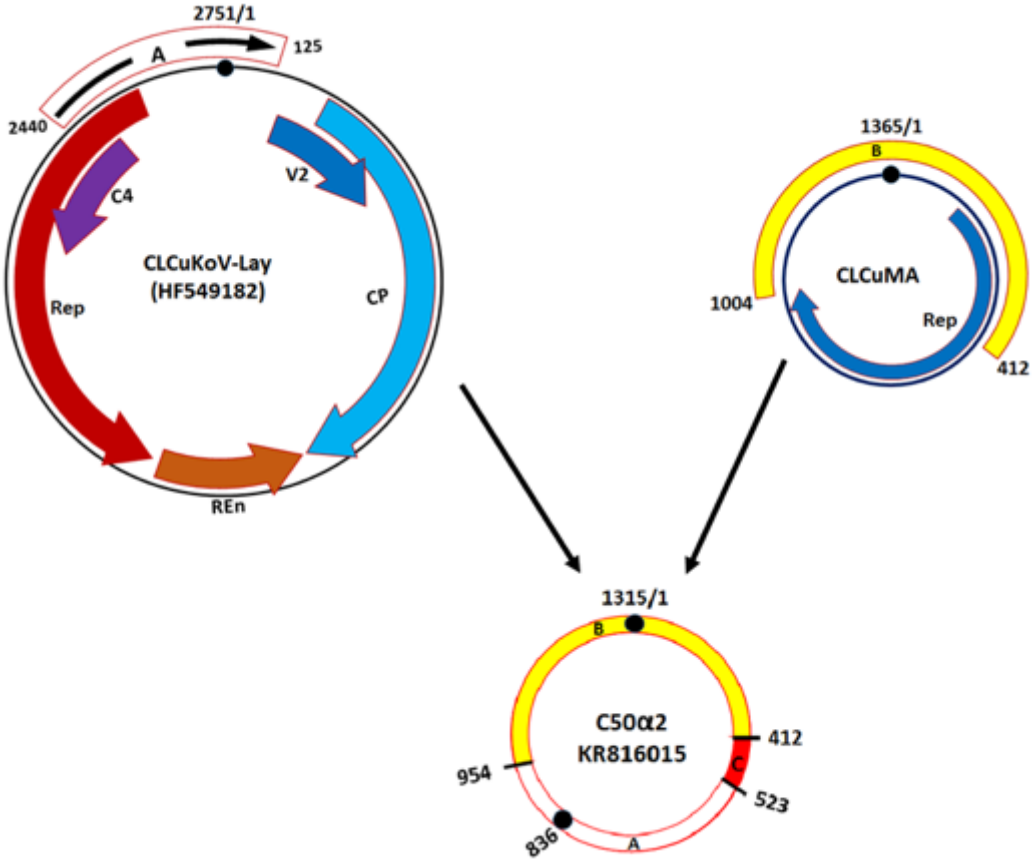


Figure 4.12. Map of an isolate C50α2 showing the sequences originating from *Cotton leaf curl Kokhran virus* - strain Layyah (CLCuKoV-Lay) and Cotton leaf curl Multan alphasatellite (CLCuMA). The fragments are labeled as A, originating from CLCuKoV-Lay, and B, originating from CLCuMA. The sequence marked C originates from *Acalypha* yellow vein mosaic alphasatellite (AcYVMA; accession number FN658711) but is inserted in the reverse orientation.

4.3.1.4. Closing remarks for alphasatellites

The characterization of alphasatellites described here is concluded with identification of a one new recombinant molecule and nine newly characterized isolates. Mubin *et al*., 2009 proposed new guidelines to update nomenclature of alphasatellites. Table 4.13 lists the isolate descriptors proposed to describe the alphasatellites characterized in this study.

Table 4.13. DNA alphasatellite sequences with the respective acronym and GenBank accession number reported from CLCuD infected cotton

Alphasatellite	Acronym	GenBank Accession No.
Cotton leaf curl Multan alphasatellite-[Pakistan-Faisalabad-GMT56- 2010]	CLCuMB-[PK-Fai-GMT56-10]	KR816016
Cotton leaf curl Multan alphasatellite [Pakistan-Dhok Ali Khan- GMT61-2011]	CLCuMA-[PK-DAK-GMT61-11]	KR816017
Cotton leaf curl Multan alphasatellite-[Pakistan-Layah- GMT57-2010]	CLCuBuA-[PK-Lay-GMT57-10]	KR816008
Cotton leaf curl Multan alphasatellite-[Pakistan-Bakkar-GMT58-2011]	CLCuMB-[PK-Bak-GMT58-11]	KR816009
Cotton leaf curl Multan alphasatellite-[Pakistan-Bakkar-GMT59-2011]	CLCuMB-[PK-Bak-GMT59-11]	KR816015
Cotton leaf curl Multan Alphasatellite-[Pakistan-Multan-GMT64-2011]	CLCuBuA-[PK-Mul-GMT64-11]	KR816010
Cotton leaf curl Multan alphasatellite-[Pakistan-Bahawalnagar- GMT67- 2011]	CLCuBuA-[PK-BWN-GMT67-11]	KR816011
Cotton leaf curl Multan alphasatellite-[Pakistan-Okara-GMT60-2011]	CLCuMB-[PK-Oka-GMT60-11]	KR816012
Cotton leaf curl Multan alphasatellite-[Pakistan- Khanewal- GMT70- 2011]	CLCuMA-[PK-KNW-GMT70-11]	KR816013
Cotton leaf curl Multan alphasatellite-[Pakistan-Rahim Yar Khan-GMT73-2013]	CLCuMA-[PK-RYK-GMT73-13]	KR816014

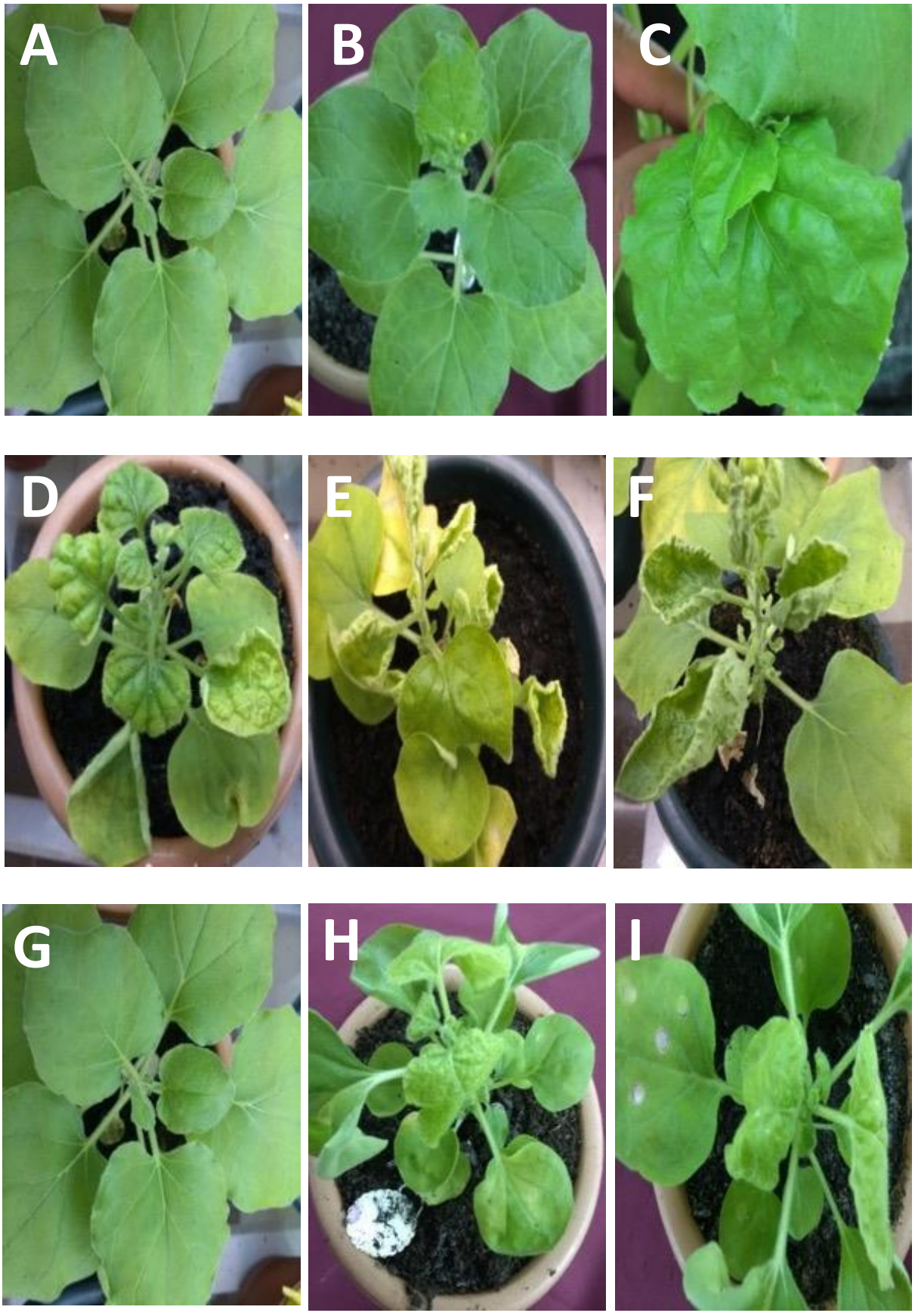
4.4. Infectivity analysis

As it has been described in the previous reports from Asia and Africa that leaf curl diseases are caused by complex of begomoviruses associated with satellite molecule(s). In Pakistan. CLCuD is caused by multiple geminiviruses, majority of them are begomoviruses and their strains, including CLCuKoV, CLCuMuV, CLCuBaV, CLCuGeV, CLCuRaV, CLCuAlV, GPMLCuV, TYLCBaV, PaLCV, ACMV and CpCDV. A detailed study was planned to carry out the potential interactions of the virus with the associated DNA satellites to access their compatibility. Partial repeat constructs of genomic components and DNA satellites were produced (section 3.9). *Nicotiana benthamiana* plants were grown up to 4-5 leaves stage in the glass house, under 24°C temperature and 16/8 hours light and dark period. The plants were agro-inoculated at the axils of the leaves using *Agrobacterium tumefaciens*

In *N. benthamiana* seedlings inoculated with Cotton leaf curl Kokhran virus-Burewala (CLCuKoV-Bu; isolate C28A) with its cognates CLCuMB (C28β) and CLCuMA (C28α) symptoms were apparent on all new emerging leaves. Initially, light depression (dimples) on the upper surface of the leaves were noted but with the passage of time, curling of leaves from the margin were also appeared after 15-19days post inoculation. With the passing days, the symptoms become more severe and prominent, like upward leaf curling, vein thickening and reduced leaflet size, as well as shunted growth was observed all over. In *N. benthamiana* inoculated with isolate C49A with its cognates (C49α and C49β), the symptoms were appeared but not severe as observed in inoculation with C28A with its cognates (Figure 4.13).

These differences in symptoms in *N.benthamiana* by these isolates with their cognates force to have a close analysis of these helper viruses. This might be the recombination in the isolate C49A (previously discussed in section 4.3.1.5). Full-length

sequences of isolate C28A and C49A were aligned using ClustalW for plotcon analysis, reveal that isolate C49A contain recombinant fragments in *CP* and *Trap* genes, derived from CLCuMuV. Previous reports show that *CP* gene is involved in nuclear localization. This might be the possible reason for isolate C49A with poor nuclear localization, further poor silencing efficiency by expressing 39AA TrAP protein rather than 35AA *TrAP* protein efficiency as in isolate C28A. Furthermore, it is evident from the phylogenetic tree that isolate C28A cluster with CLCuKoV-Bu group but isolate C49A produce a distinct branch, specifying it distinctiveness.


A
B
C
D
E
F
G
H
I

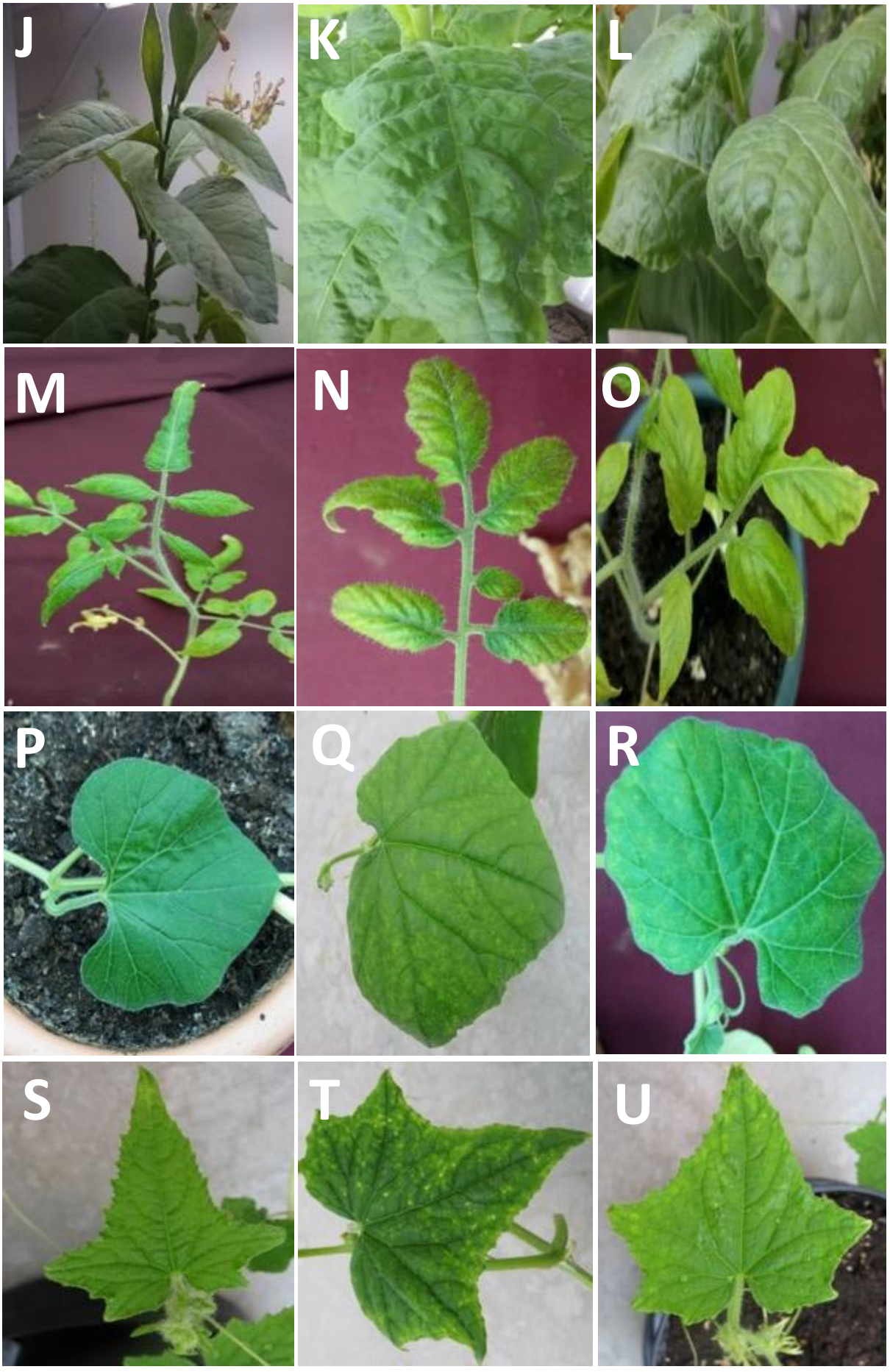

J
K
L
M
N
O
P
Q
R
S
T
U

Figure 4.13. Infectivity analysis, *Nicotiana benthamiana, Nicotiana tabacum,* tomato, pumpkin and cucurbits plant were agro-inoculated with C-28A (C, J, M, P and S), C28A and C28β (D, E, H, K, N, Q, and T), C28A, C28β and C28α (F, I, L, O, R and U), C49A (G),C49A, and C49β (H) C49A,C49β and C49α (I), mock plant (B) and control healthy plant (A).

4.4.1. Agro-inoculation of plants with component C28A and its associated cognates alphasatellite and betasatellite

For infectivity studies, multiple plants were inoculated, including sixty (60) *N. benthamiana,* twenty (20) *N. tabacum,* twenty (20) *Solanum lycopersicum,* and forty (40) cucurbits (cucumber and pumpkin) plants were agro-inoculated with all possible combination of CLCuKoV-Bu (isolate C28A) with its associated cognate DNA satellites (Table 4.14). In each experiment, 2-4 plants were used as a healthy controls. Symptoms seemed within 15-25 days post inoculation in agro inoculated plants but there was no symptom seen in mock and healthy control plants. Experimental plants/seedlings inoculated with Cotton leaf curl Kokhran virus-Burewala (CLCuKoV-Bu; isolate C28A) produced no or mild visual symptoms (Figure 4.13 C, J, M and P). CLCuKoV-Bu (Isolate C-28A) in the presence of its cognate CLCuMB (C28β) started to show symptoms with 15-17 days post inoculation, with increased upward leaf curling, a greater reduction in leaf size, followed by severe stunting (Figure 4.13 D, E, H, K, N, Q, and T). However, CLCuKoV-Bu (C28A)-CLCuMB (C28β) complex in the presence of CLCuMA (C28α), the symptoms appeared but delayed as well as not severe as in the absence of alphasatellite. Initially, light depression (dimples) on the upper surface of the leaves were noted but with the passage of time, curling of leaves from the margin were also appeared after 17-19 days post inoculation. As the days pass on, the symptoms become more severe and prominent, like upward leaf curling, vein thickening and reduced leaflet size, as well as stunted growth was observed (Figure

4.13F, I, L, O, R and U). The results were confirmed by PCR using diagnostic primers. All symptomatic plants gave positive PCR with expected size of PCR product for genomic component, alpha-and betasatellite molecules, but there was no amplification in healthy control and mock inoculated plants.

4.4.2. Agro-inoculation of plants with isolate C49A and its associated cognates alphasatellite and betasatellite

Twenty (20) *N. benthamiana* plants were agro-inoculated with Cotton leaf curl Kokhran virus-Lay (component C49A) and its cognate satellites (Table 4.14). Initially, dimples like symptoms were seen but later leaves become narrow, pointed and twisted leading to leaf curl like symptoms, appeared after 18-25 days post inoculation. Plants agro-inoculated with isolate C49A only, produced no or very mild symptoms (Figure 4.13G), but addition of C49β, enhance the symptoms (Figure 4.13H). However, in the presence of alphasatellite, symptoms appeared to be delayed (Figure 4.13I). Stunted growth was observed in all agro-inoculated plants while healthy and mock plants were noticed as symptomless (Figure 4.13A and B). Results were confirmed by PCR for genomic component, alpha and beta-satellite. All infected plants gave expected size products but there was no amplification in healthy control plants.

Table 4.14. Plant species agro-inoculated with isolate C28 and C49, and its cognates satellites (CLCuMB) and (CLCuMA).

Isolate	Host	No. of plants infected / inoculated	No. of healthy plants	Days post inoculation	PCR			Symptoms on inoculated plants
					DNA-A	Alpha satellite	Beta satellite	
C-28	*N. benthamiana*	60/60	4	15-25	+	+	+	D, LC
	N. tabacum	20/20	4	15-25	+	+	+	D, LC
	S. lycopersicum	15/15	3	15-25	+	+	+	D, YV, LC
	Cucumber	20/20	2	15-25	+	+	+	D, YV, LC
	Pumpkin	20/20	2	15-25	+	+	+	D, YV, LC
C-49	*N. benthamiana*	20/20	4	18-25	+	+	+	D, LC

D, Dimples on upper leaf surface, LC, leaf curling, YV, yellow veins

4.5. Bacterial expression of coat protein gene

For the expression, full-length CP gene of Burewala strain [CLCuKoV-Bu] was amplified using specific primer sets (Table 3.1). PCR fragments were initially cloned in pTZ57R/T vector and *E. coli* strain DH5α was transformed with it. Positive clones were initially confirmed with restriction digestion and colony PCR. Required fragment was released from pTZ57R/T vector and sub-cloned in a pET expression vector to produce pET28a-CPBur. This is referred as hereby "Lab Construct (LC)" [kindly provided by Dr. Muhammad Tahir (PhD Supervisor) for my research work].The map of the construct is shown in the Figure 4.14.

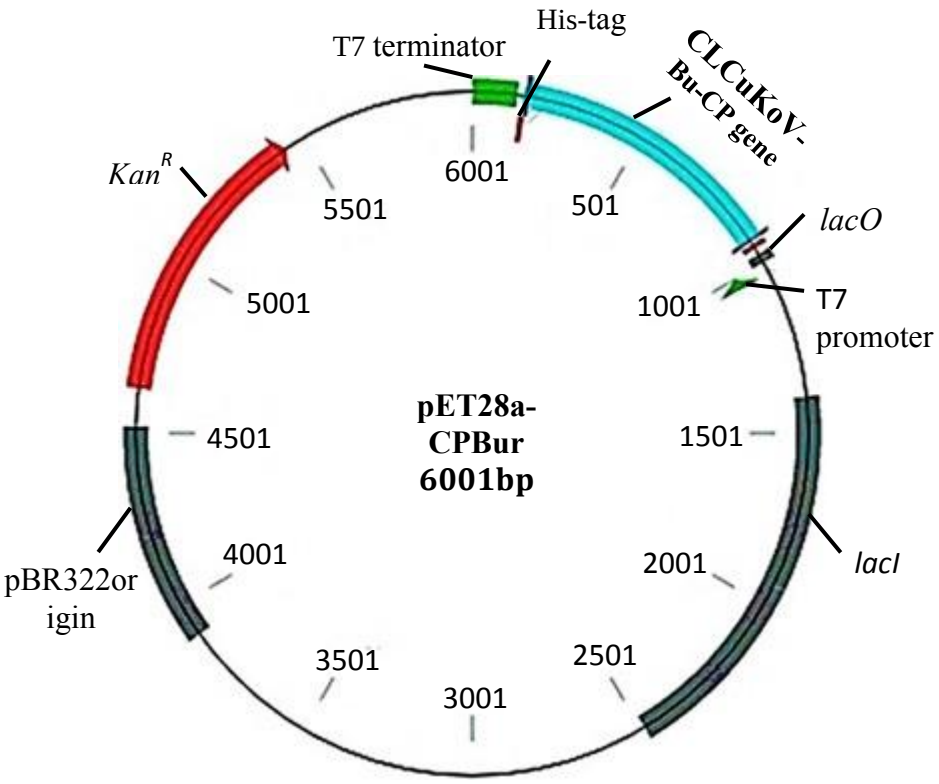


Figure 4.14. Schematic representation of lab construct (pET28a-CPBur).

4.5.1. Confirmation of pET28a-CPBur

Plasmid of the LC was initially confirmed subjected to PCR, using CP specific primer sets that produced an expected size of amplicon ~780 bp as shown in the Figure 4.15A. Recombinant expression plasmids (pET28a-CPBur) was also confirmed by restriction digestion using unique enzymes (*Nco*I/*Xho*I). To observe the restriction analysis, digested plasmid was run on 1% agarose gel and two bands was observed corresponding to pET28a vector and inserted gene i.e. at ~5.3kb and ~0.78kb respectively, as shown in Figure 4.15B. Positive recombinant plasmid was finally transformed into *E. coli* strain of BL21 (CodonPlus) for the expression of the CP gene under T7 promotor. Positive colonies were scrutinized by colony PCR as in Figure 4.15 C.

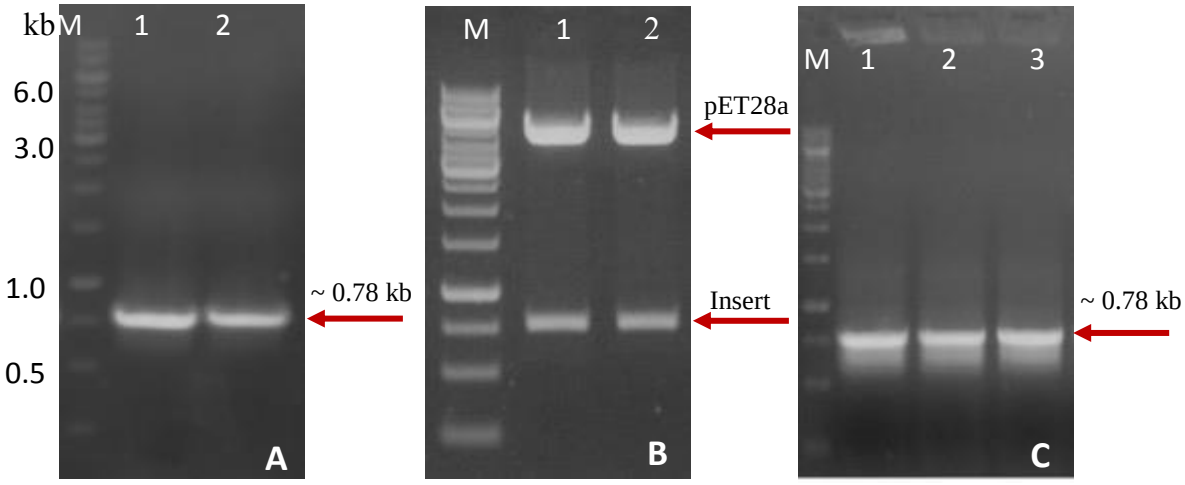


Figure 4.15. Schematic representation of recombinant expression construct in pET28a, restriction analysis and screening of positive transformants by colony-PCR. A: Schematic diagram of expression vector, pET28a-CPBur under T7 promoter. A: Lane 1 and 2 are the PCR product on 1 % agarose gel, B: Lane 1 and 2 are the restriction digestion of clones pET28a-CPBur (*Nco*I/*Xho*I), M; 1 kb DNA Ladder. C: Colony PCR Lane 1 -3, M; 1 kb DNA Ladder (Thermo Fisher Scientific).

4.5.2. Optimization of Expression Conditions

The plasmid pET28a-CPBur was optimized for expression at different conditions including the effect of time, temperature of incubation and concentation using IPTG as an inducer (details are in section 3.11.15).

4.5.2.1. Effect of time

Effect of time on expression of pET28a-CPBur with IPTG was studied using 50 ml LB broth (containing kanamycin) was inoculated with 1% of ON culture at 37 C. IPTG of 0.5 mM concentration was used to induce the cells at 0.6-0.8 OD_{600}. 1 ml each of the un-induced sample (as a control) and post-induction samples (taken after every 2 hrs interval) was taken from culture to observe the maximum level of expression up to 12 hrs. Cells were harvested, lysed with lysis buffer and expression was analyzed on 12% SDS PAGE. Enhanced expression was seen at 4-6 hrs of induction and was maintained in further IPTG induction experiments.

4.5.2.2. Effect of IPTG concentration

The effect of IPTG concentration on pET28a-CPBur expression was studied. Different concentration of IPTG (02, 0.4, 0.6, 0.8 and 1.0 mM) were induced at 0.6-0.8 OD 600. At 6 hrs of post-induction, 1 ml of sample was collected from each culture and the cells were spin down, lysed by lysis buffer and observed on 12 % SDS-PAGE. The expression was seen as maximum at 0.4-0.6 mM IPTG concentration. So, 0.5 mM concentration of IPTG was maintained in all experiments.

4.5.2.3. Effect of temperature

To study the effect of temperature, different conditions of temperature (18 °C-37 °C) were used. 1 % of overnight incubated culture was re-freshed in 50 ml of LB broth (containing kanamycin). IPTG of 0.5 mM concentration was used to induce at

0.6-0.8 OD_{600}. Pre- and post-induction samples (1 ml) were collected after 6 hrs of incubation. Sonication was done for post induction samples to find the characteristic of protein whether soluble or insoluble (inclusion bodies) at low to high temperature. Sample were run on 12 % SDS-PAGE to observe the level of expression. Low temperature has no effect on the nature of protein. The expression was seen maximum at 37 °C and it was maintained for next all experiments with IPTG.

4.5.3. Expression of recombinant plasmid (pET28a-CPBur) in *E. coli* strains

The expression of positive transformants was evaluated by inoculating 10 ml LB medium (50 µg/ml kanamycin) and kept at 37 °C in shaking incubator for an overnight growth. Next morning, 50 ml LB broth was refreshed with 1% of overnight culture containing kanamycin (50 µg/ml) and incubated at 37 °C in shaker. Upon 0.6-0.8 optical density (OD_{600}), 0.5 mM IPTG was induce and pre-induction sample was taken as control. After 6 hours of post induction, OD_{600} was again measured i.e. reached to near 1.6-2.0 and the cells were harvested. Cells were treated with 50 mM Tris-HCl buffer, followed by sonication to find out either the protein is in soluble or in the insoluble portion (inclusion bodies) (details are in material and methods section 3.11.5). To analyze the level of expression, the sample was measured by running on 12 % SDS-PAGE. Approximately 30-40% of total cell protein was observed for recombinant construct (Figure 4.16).

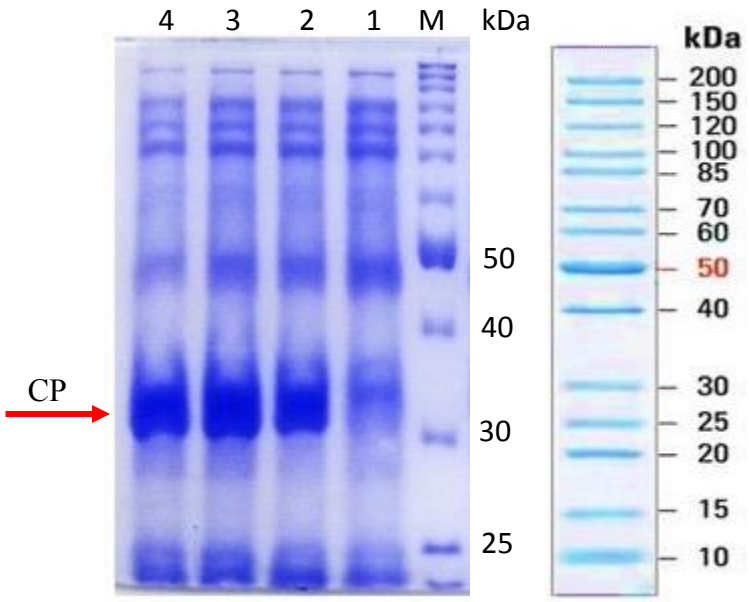


Figure 4.16. SDS-PAGE (12 %) showing expression of recombinant protein at 32 kDa in *E. coli* BL21CodonPlus transformed with pET28a-CPBur construct. M, Page Ruler protein ladder (Thermo Fisher Scientific), Lane 1: uninduced; Lane 2, 3 and 4 post-induced samples at 2, 4 and 6 hrs, respectively.

4.5.4. Sub-cellular localization of expressed protein

To check sub-cellular location of the expressed protein, in an optimized condition, 100 ml of LB broth containing kanamycin was inoculated with 1% overnight inoculum and induced with 0.5 mM IPTG. The culture was grown at 37 °C in shaking incubator for 6 hours. Cells were pelleted by centrifugation and followed by sonication (details are in section 3.11.6). Cell lysate was centrifuged and run on SDS-PAGE for visualization of total cell protein (TCP), soluble protein (Sol) and insoluble protein (inclusion bodies, IBs). Analysis of fractions showed that the protein was expressed in inclusion bodies as shown in Figure 4.17.

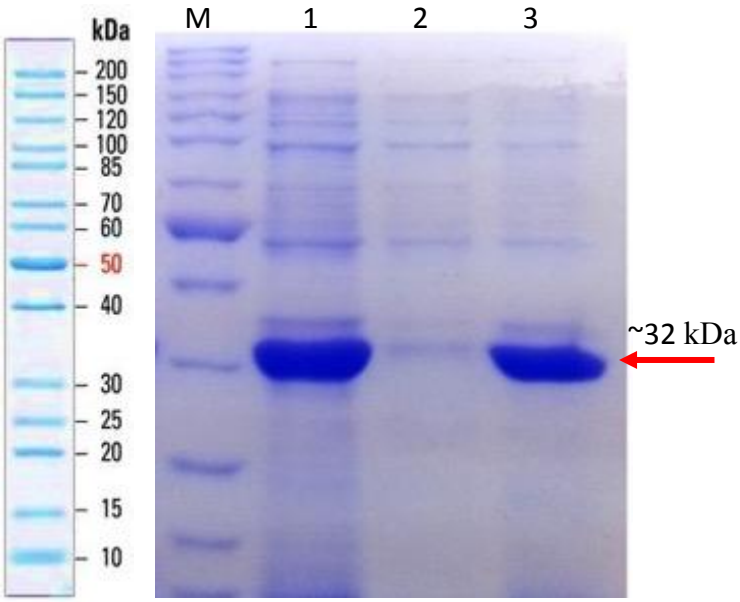


Figure 4.17. SDS-PAGE of sub-cellular fractions of CP in constructpET28a-CPBur.M: Page Ruler protein ladder (Thermo Fisher Scientific). Lane 1: TCP, Lanes 2: Soluble protein, and Lanes 3: insoluble inclusion bodies (IBs).

4.5.5. Production of coat protein from recombinant construct pET28a-CPBur

After optimization of all conditions for highest level of expression,1 liter of fresh LB broth containing kanamycin (200 ml LB broth/flask) was inoculated with 1 % of overnight inoculum and set at 37 °C (details are in section 3.11.7) with 0.5 mM IPTG induction. The cells were pelleted and washed with washing buffer and then resuspended into Tris-HCl buffer and subjected to sonication. Sonication was made to separate the soluble and insoluble (inclusion bodies) potrions (details are in section 3.11.6). All the samples were prepared for 12 % SDS-PAGE to observe level of expression as shown in Figure 4.16. The insoluble protein (IBs) was proceeded for further solubilization, purificationand refolding.

4.5.6. Purification and refolding of coat protein (CP)

4.5.6.1. Washing of Inclusion bodies and refolding of coat protein

Insoluble inclusion bodies were washed thrice with wash buffer to remove cell debris and associated host protein (details are section 3.11.7). 10 ml of wash buffer was used per gram of cells. Final washing was done with 50 mM Tris-HCl. Purified IBs were then progressed for solublization. Purified IBs were solubilized in IB solubilization buffer at the rate of 10 OD (detail is given in section 3.11.7). The solubilized inclusion bodies were taken to refolding process.

Refolding of recombinant protein was done by the process of dilution method. Ice cold refolding tank of 50 ml was made and kept at 4 °C for refolding process to complete (details are in materials and method section 3.11.8). The solubilized protein was poured in the dialyses tube and was shaken by magnetic stirrer in the tank constantly at 4 °C to carry out air oxidation. Buffer in the tank was continuously changed until 0 M guanidine-HCl was used. The pH was then reduced from 11 to 9.5 gradually and was run on 12 % SDS-PAGE as shown in the Figure 4.18.

4.5.6.2. Purification of coat protein

Purification of coat protein was done using glass column (Bio-Rad) containing 3 ml Ni-IDA. The resin was washed with 10 volumes of distilled water, followed by equilibration with 10 volumes of Ni-binding buffer through Ni-Chromatography column. 4 ml solubilized protein was loaded to column containing 6M guanidine-HCl. Flow rate of 0.5 ml/min was maintained throughout the process. 5ml of sample was loaded onto column. The column was then washed with 10 column volumes of Ni-

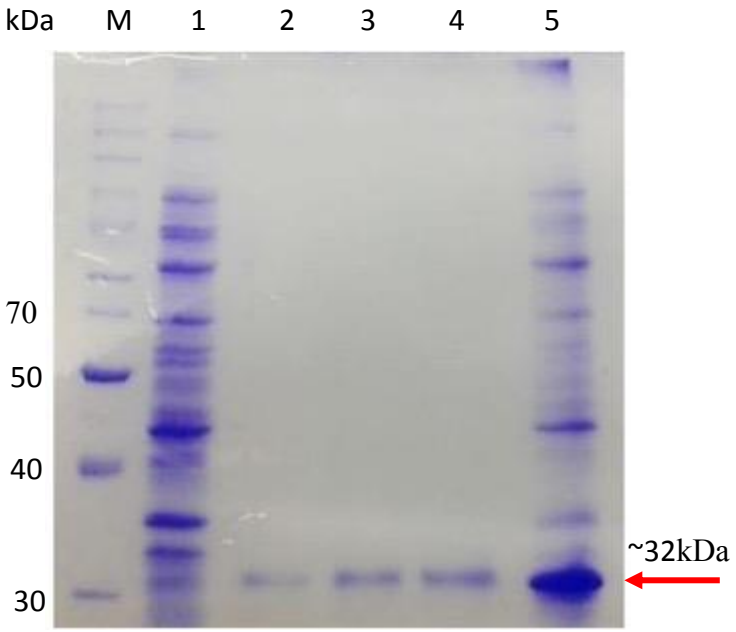


Figure 4.18. SDS-PAGE (12 %) analysis of refolded protein after removal of guanidine-HCl through dialysis. M: protein marker, Lane 1: unbounded protein, Lane 2, 3 and 4: refolded protein, and Lane 5: TCP.

wash buffer I followed by washing with 5 column volumes of Ni-wash buffer II. Elution of bound proteins was done with 2.5 ml of Ni-elution buffer I. Samples of all fractions were prepared for SDS-PAGE by removing guanidine-HCl.

Purification of coat protein was also done by elution from denaturing polyacrylamide gel. Solubilized protein was mixed with loading sample to run on 12 % SDS-PAGE. Gels were stained with Coomassie brilliant blue R250. When the gel portion was properly destained, it was placed against the remaining unstained gel and gel slice was cut corresponding to the required protein band. With surgical blade the gel piece was sliced to cubes of minimum possible size and placed in a 15 ml falcon tube. Gel pieces were washed with dH2O. The gel pieces were meshed with pestle after repeated freeze thaw process.

1.5 ml of denaturing-polyacrylamide gel (D-PAG) elution buffer was added to the gel and incubated in shaking chamber at 37°C overnight. Next day the gel slurry was centrifuged at 7500 x g for 20 min at room temperature and supernatant containing

protein was separated.1 ml of protein sample was taken in 15 ml falcon tube and 4 volumes of ice-cold acetone were added. Sample was briefly vortexed to mix and placed at -20°C for 60 min. Precipitated proteins were collected by centrifugation at 7500 x g for 15 min at 4°C. Acetone was decanted and pellet was left to air dry for about 30 min. The pellet was dissolved in guanidine-HCl containing solubilization buffer. 5ul of checked on 12% SDS-PAGE gel as shown in the Figure.4.19.

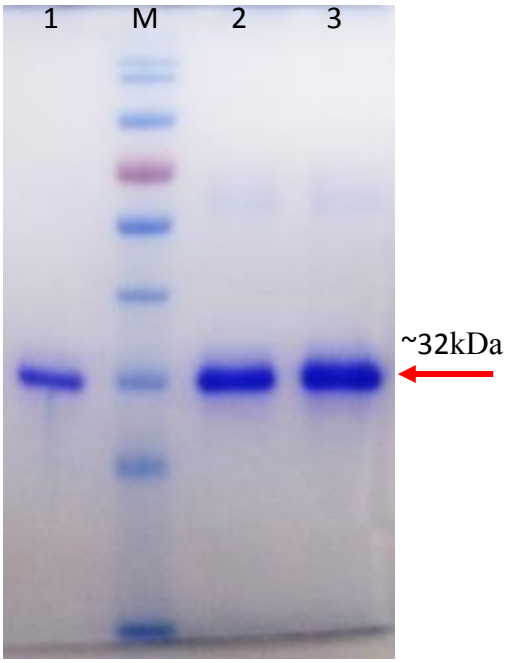


Figure 4.19. SDS-PAGE (12 %) analysis of gel based purified protein after removal of guanidine-HCl through ethanol precipitation. M: protein marker, Lane 1: 5 µl, Lane 2: 10 µl and Lane 3: 15, respectively.

4.6. Production of polyclonal antisera (rabbit anti-rCP27 Ab)

Two rabbits (New Zealand breed i.e. rabbit 1 and 2) were immunized at one month interval with 200 µg of rCP27 (recombinant coat protein expressed from LC) through subcutaneous application for five months. Details are mentioned in material and methods, section 3.11.10. Blood was drawn from the marginal veins at interval of

10 days, after each booster dose, starting from the second injection. Blood was centrifuged and serum was collected and stored at -20 °C for further use.

4.7. Testing of antisera

Polyclonal antisera were tested with enzyme linked immunosorbent assay (ELISA) and the western blot.

4.7.1. DAS-ELISA

N. benthamiana plants earlier infected with partial tandem repeats in the green house were collected for the detection of begomovirus, against which polyclonal antisera have been raised. Different dilution of the serum was made, like 1:10, 1:100, 1:500, 1:1000, 1: 10,000 and 1:100,000.

Antiserum was evaluated by double antibody sandwich ELISA (DAS-ELISA), which was done earlier described by Clark and Adams (1977) and Givord *et al*., (1994) for plant viruses. Different dilution of antiserum was used to coat the microtiter plate. Leaf samples (as shown in the Figure 4.13) from both plants (healthy and infected *N. benthamiana* plant) at the rate of 1: 20 (w:v) in extraction buffer was used, from which 200 µl was loaded per well of the microtiter plate as earlier describe by Clark and Adams (1977). Goat-anti-rabbit-IgG-AP conjugate was used at 1:1000 dilution as mentioned on the delivery note of DSMZ, Germany, as a secondary antibody.

Finally, addition of substrate produced yellowish color in all those wells, which were coated with leaf extract of infected plants, earlier confirmed by PCR also, as shown in the Figure 4.20. While, on the other hand, no color or very mild color were noted in all those wells where negative control, healthy control or no serum was used. The result of DAS-ELISA was further confirmed by ELISA reader, as shown in the Table. 4.15.

As were using different dilution of serum from both rabbits in DAS-ELISA, best results were noted with dilution of 1:500 from both rabbits. The results of crude serum from both rabbits were compared with commercially available kit for DAS-ELISA of ToLCV (DSMZ-Germany), were used as reference in a dilution of 1:1000, which gave position reaction for all three test samples with OD_{405} values in the range of 1.87-1.96, almost double in efficiency when comparison with antibodies raised here, shown in the Table 4.15.

Serum from the blood of both rabbits ($1B_0$ and $2B_0$) which were drawn before immunizaion, detect no virus in the infected samples, after absorbance at 405 nm, and gave a range of absorbance from 0.02 to 0.23, which was almost similar with the absorbance obtained with negative control (0.024-0.046), healthy control (0.03-0.07) and with no serum (0.028-0.043).

Absorbance with infected samples which were used in triplicate form in two columns of the plate to avoid contamination. Reactivity of serum from rabbit 1 (1B) was noted with 1.39-1.57, in the wells 7B, 7C, 7D, 8B, 8C, and 8D. Similarly, reactivity of serum from rabbit-2 (2B) was also measured at 405nm, showed 1.45-157. Both rabbits produce almost same level of serum containing polyclonal antibodies for the detection of CLCuKoV-Bu, as shown in the Table 4.15.

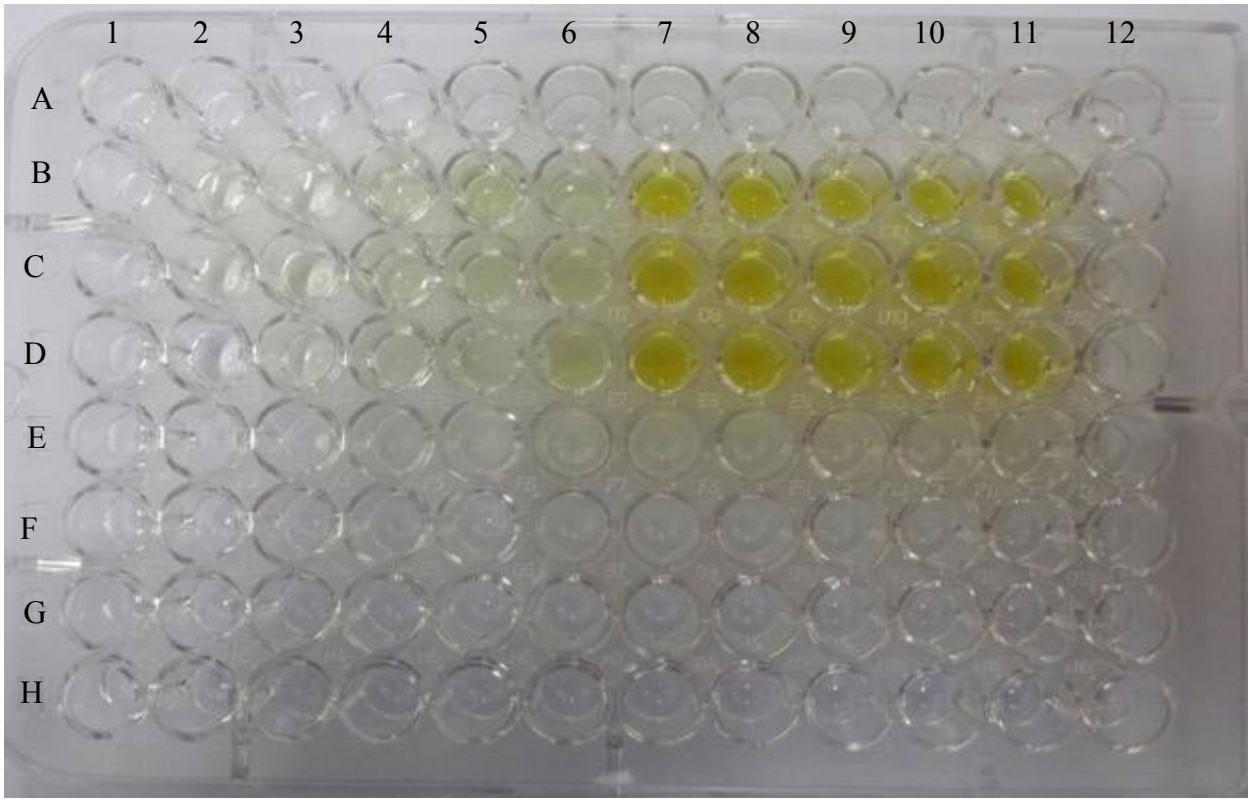


Figure 4.20. DAS-ELISA for *N. benthamiana* plants infected with CLCuKoV-Bu and associated cognates CLCuMA and CLCuMB.

Table 4.15. DAS-ELISA detection of *Cotton leaf curl Kokhran virus-Burewala* in leaf extracts of *N. benthamiana*.

Well No.		Absorbance (405nm)	ELISA reaction (+ or -)
Negative Control	2B	0.024	-
	2C	0.022	-
	2D	0.046	-
No Serum	3B	0.030	-
	3C	0.028	-
	3D	0.043	-
Healthy Control	4B	0.050	-
	4C	0.034	-
	4D	0.079	-
IB_o	5B	0.026	-
	5C	0.040	-
	5D	0.060	-
$2B_o$	6B	0.233	-
	6C	0.220	-
	6D	0.187	-
$1B_1$	7B	1.393	+
	7C	1.454	+
	7D	1.388	+
	8B	1.421	+
	8C	1.457	+
	8D	1.574	+
$2B_1$	9B	1.453	+
	9C	1.466	+
	9D	1.527	+
	10B	1.451	+
	10C	1.457	+
	10D	1.487	+
Positive Control	11B	1.875	+
	11C	1.921	+
	11D	1.963	+

4.7.2. Western Blotting

The polyclonal antisera developed in two rabbits [Rabbit-1 (R_1) and Rabbit-2 (R_2)] against the expressed coat protein of cotton infecting Burewala stain was evaluated in western bloting for their specific reaction with purified protein.

In western blotting, crude antisera collected from the blood of both rabbits (R_1 and R_2) were used for their ability to detect 32 kDa protein. However, the post-boost antisera of both rabbits (after immunization) was negative with curde protein from healthy *N. benthamiana* plants (Lane 1 and 3), while positive for crude protein from infected *N. benthamiana* plants (Lane 2 and 4); earlier confirmed with diagnostic PCR. The rabbit anti-rCP antibodies was also able to show a positive serological reaction with proteins of less than 24 kDa present in the infected *N. benthamiana* plants (Lane 2 and 4). It reacted with whole bacterially expressed protein (32 kDa) from pET28-CPBur construct (Lane 5 and 6), while using post-boost antisera (after immunization), showed a strong positive reaction with 32 kDa recombinant protein of the CLCuKoV-Bu (Lane 7 and 8) as shown in the Figure 4.21.

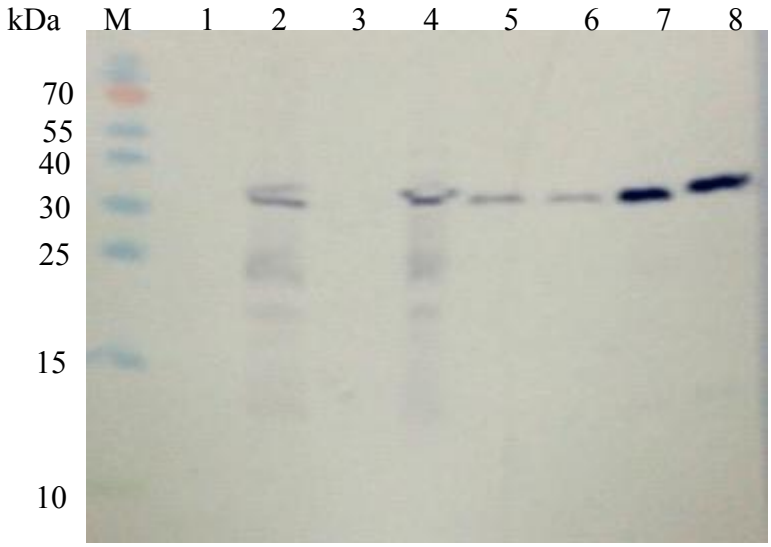


Figure 4.21. Western blot analysis of the recombinant protein of CLCuKoV-Bu with post boost antiserum. Lane 1 and 3 (crude protein from healthy *N. benthamiana* plants), Lane 2 and 4 (crude protein from infected *N. benthamiana* plants), Lane 5 and 6 (total cell protein of bacterially expressed in pET28-CPBur construct), Lane 7 and 8 (purified protein from pET28-CPBur construct) and M: Page Ruler.

5. DISCUSSION

Phytopathogenic viruses are responsible for the major limitations to agricultural productivity throughout the World. However, the claimed losses cause the most hardships in the tropical and sub-tropical regions in developing countries. There are several factors that account for this and they comprise the fact that the diversity and prevalence of plant-infecting viruses of agricultural significance are higher. Furthermore, the environmental conditions are more favourable for the vectors of the viruses, and that the farmers cannot manage/afford the expensive chemical control agents with their low budgets as well as, in many cases, not having the educational standard and training to use them efficiently.

It is unfortunate fact that Pakistan geographical position, in common with all other countries of southern Asia, serve as a hub for harboring all taxonomic groups of plant infecting viruses infecting cultivated, non-cultivated, introduced and endemic plants (Mali and Ragegore 1979; Naidu *et al*., 1989; Nadeen *et al*., 1999; Raikhy *et al*., 2003; Ali *et al*., 2004; Haider *et al*., 2004; Mandal *et al*., 2004; Verma *et al*., 2004; Iram *et al*., 2005; Lauren *et al*., 2006; Amin *et al*., 2007; Tahir *et al*., 2010) which, at least in part, elucidates the low agricultural productivity in the country. Amongst these phyto-pathogens, prime viruses are the whitefly-transmitted geminiviruses (begomoviruses) (Monga *et al*., 2011, Brown *et al.*, 2015).

Begomoviruses have become the most destructive group of plant viruses. Like other viruses in the tropical and sub-tropical regions of the world, begomoviruses play a vital role to limit crop production, distribution of insect vectors and global trade of plant materials (Seal *et al*., 2006). Another concern has been noticed in begomoviral

infections that the re-emergence of begomviral diseases is an interplay of a complex of begomoviruses and satellite DNA components (Saunders *et al.,* 2001; Varma and Malathi, 2003; Bull *et al.,* 2004a). However, the abundant number of begomovirus isolates (Brown *et al.,* 2012, 2015), the continued reports of new species as well as the high degree of genetic diversity within species (García-Andrés *et al.,* 2006; Patil *et al.,* 2005; Sanz *et al., 1*999; Ooi *et al*., 1997; Stenger, 1994) suggest that begomoviruses have a high mutation rate (Duffy and Holmes, 2008; Padidam *et al.,* 1999). This may be the reason which lead to the emergence of new species in a short possible time with a high diversity among begomoviruses, infecting plants with a wide host range.

Pakistan serves as a hot spot region, where cotton is severely affected by begomoviruses, includes herbs, shrubs, vegetables, field crops and fruits crops. Begomoviruses infecting dicots including *Ageratum conyzoides, Solanum nigrum* and *Zinnia elegans* (Haider *et al.,* 2007), Capsicum spp. (Tahir *et al.,* 2010; Shafiq *et al.,* 2010), chili, pepper and tomato (Hussain *et al*., 2004; Shih *et al*., 2003; Mansoor *et al.,* 1997), *Eclipta prostrata* (Haider *et al.,* 2006), *Duranta erecta* (Iram *et al*., 2005; Mustujab *et al.,* 2015), *Duranta repens* (Tahir *et al.,* 2006), C*apsicum annuum* (Tahir and Haider, 2005), Mungbean (Bashir *et al.,* 2006), *Croton bonplandianus* (Amin *et al*., 2002), *Sonchus arvensis* (Mubin *et al.,* 2010), Legumes (Ilyas *et al*.,2010), *Vinca minor* (Haider *et al.,* 2007, 2008), okra, radish and watermelon (Mansoor *et al*., 2001, 2000b, 2000c), *Momordica charantia* (Tahir *et al*., 2010), *Rhynchosia minima* (Ilyas *et al.,* 2009), papaya (Nadeem *et al.,* 1997), *Vigna aconitifolia* (Qazi *et al*., 2006), *Ricinus communis* (Fareed *et al.,* 2012), *Luffa cylindrical* (Zia-ur-Rahman *et al.,* 2013), *Urtica dioica* (Iqbal *et al.,* 2016).

Cotton is among the most important crops grown for its food, feed and fiber sources from the ancient times. It covers ~3 million hectares land during growing

season and accounts for 60% of the export products. Initially, farmers were used to grow their local cultivars, which were resistant to many biotic factors but at the cost of low yield and reduced fiber quality. For high production and yield, new cultivars were introduced from the United States. The production was increased many folds in comparison with local cultivars. The introduced cultivars were also noticed to be susceptible to a variety of pathogens, among which CLCV was dominant (Sattar *et al.*, 2013; Farooq *et al.,* 2011, Saeed *et al.,* 2004; Sanz *et al.,* 2000; Khan and Khan, 2000; Zhou *et al.*, 1998; Harrison *et al*., 1997; Mansoor *et al*., 2011, 2006, 2003, 1999, 1998, 1993). CLCuD is recorded as one of the most disparaging disease of cotton. CLCuD can be caused by a single virus or in form of complex of begomoviruses with a disease specific satellite-CLCuMB (Mansoor *et al*., 2003b; Briddon *et al.,* 2003; Amrao *et al*., 2010a).

Pakistan has been affected by two epidemics in cotton. First was noted in the era of 1980s in the vicinity of Multan, in which several monopartite begomoviruses were involved specifically *Cotton leaf curl Multan virus* (CLCuMuV), *Cotton leaf curl Kokhran virus* (CLCuKoV), *Cotton leaf curl Alabad virus* (CLCuAV), *Cotton leaf curl Rajasthan virus* (CLCuRaV), and *Papaya leaf curl virus* (PaLCuV), *Cotton leaf curl Bangalore virus* (CLCuBaV) associated with alphasatellite and betasatellite molecules (Mansoor *et al*., 2003b; Briddon *et al*., 2003), referred to as "1st epidemic" in Pakistan. The financial losses due to the CLCuD epidemic were estimated at US $5 billion during 1992-97 (Briddon and Markham, 2001). Unluckily, from 2001 onwards, previously resistant varieties succumbed to CLCuD (Mahmood *et al.,* 2003; Mansoor *et al*., 2003b), and now no commercially grown cotton varieties show any more than tolerance to the disease. This resistance breaking was shown to be associated with a distinct begomovirus, Cotton leaf curl Burewala virus [now called as Cotton leaf curl Kokhran

virus-Burewala (CLCuKoV-Bu) after revision by Brown *et al.*, (2015)], which was unknown before the resistance breaking strain associated with a recombinant form of CLCuMB (Amin *et al*., 2006; Amrao *et al*., 2010). This epidemic on cotton was noticed/appeared on all previously resistant cultivars in the vicinity of Burewala district, referred to "2nd epidemic". From 2001 onwards, burewala strain was disseminated to the other regions of Pakistan and northwestern India with un-expected losses in cotton as well as in other commercial crops. Therefore, the main objective of this study was molecular characterization of selected virus isolate(s), if present, from the cotton crop. In this study, I have investigated the cotton fields in the major growing districts of Punjab (serve as a hub of cotton) were screened. Symptoms on cotton leaves were noticed from severe to mild leaf curling, enation and vein thickening and whitefly load were seen from maximum to minimum in the field. Adjacent fields were also inspected, where seasonal vegetables were grown. Begomoviral symptoms and whitefly were also seen on these vegetables and some known weeds, indicated as an alternate hosts of begomoviruses. This study identified nine virus variants and a single new strain. The aim of this work was to provide baseline information for evaluating the potential impact of cotton infecting begomovirus(es) to understand their evolution and thus design more robust and broad-spectrum control strategies to combat.

Availability of new and advance techniques in molecular biology has made the job of cloning and reduced the cost and efficiency of sequencing. This is why there has been a rapid increase in the number of available begomovirus genome sequences in the global databases. To this date, more than 3000 of full-length genomes corresponding to DNA A of begomoviruses are available in the public databases, among which approximately 200 isolates have been submitted from Pakistan (Brown *et al.*, 2015). More than 180 isolates are from Punjab region only (Ilyas *et al.*, 2010; Shih *et al.*, 2003)

which indicating a clearer genetic diversity and distribution of begomoviruses in this region in Pakistan. More interestingly, all isolates reported in this study belongs to the most robust strain of CLCuKoV-Bu currently. This strain is more dominant among all previously reported cotton infecting begomoviruses and widely present in cotton. Our reports are in accordance with previous reports from same region of Pakistan, indicating the dominancy of this strain in this region as well as in adjacent border areas of India (Hina *et al*., 2012; Rajagopalan *et al*., 2012; Zaffalon *et al*., 2012; Amrao *et al.,* 2010).

Begomovirus species/strain is not restricted to the isolation host/species. Interestingly, cotton is the host of several begomovirus species and their strains (as previously discussed). Similarly, host is also not limited to specific species/strain of begomovirus in a specific geographical area. CLCuV is an important phytopathogen, the host range of which is bounded to infect cotton not only in Pakistan, but also reported from India, China and African. Along with cotton, CLCuV is also reported from other plant hosts including Brinjal, Cucurbits (*Tinda, Kali tori*), Hibiscus, *Convolvulus arvensis*, *Rumex dentatus,* Water Melon, Cow Pea and Lilly plants, Okra, *Malvaviscus penduliflorus*, papaya, tomato, *Ricinus communis*, *Luffa cylindrica*, China rose (Anonymous, 1993;Fareed *et al.,* 2012; Mao *et al.,* 2008;Zia-ur-Rahman *et al.*, 2013; Akhtar *et al.,* 2014).

Geminiviruses might have evolved by several different mechanisms, leading from nucleotide substitution to homologous recombination and component capture. Contrary to other expectation, recombination/component exchange (known as pseudo-recombination) is considered as a major driving force in evolution of new geminiviruses (begomoviruses) (Padidam *et al*., 1999; Roye *et al*., 2000; Pita *et al*., 2001; Harrison and Robinson 2005; Rothenstein *et al.*, 2006; Seal *et al.*, 2006). Previous studies in this regard highlighting the basic species of begomoviruses infecting cotton in this region

were not more than seven in the early 1990s. Later, the number of virus species/ strains were increased, due to recombination between the basic species. A prime example is the ongoing epidemic of CLCuD that originate from Punjab region and spread to Sindh region in Pakistan (previously virus free zone) and continues to spread northwestern India and China. The severe epidemic is attributed to a recombinant strain of CLCuBV, resulted from the probable recombination between CLCuMuV and CLCuKoV (Amrao *et al.*, 2010a). Similarly, *Cotton leaf curl Shadadpur virus* [now a strain of Cotton leaf curl Kokhran virus-Shadadpur (CLCuKoV-Sha) after revision by Brown *et al*. (2015)] is also evolved from the recombination between CLCuMuV and CLCuKoV, in which the major and minor parents remain the same but the breaking point are different from Burewala virus (Amrao *et al.*, 2010b). There are other numerous examples of recombination/pseudo-recombination in literature. Probably the most striking is begomoviruses associated with the pandemic of cassava mosaic disease (CMD) that originated in northern Uganda, spread across eastern Africa and continues through central and western Africa. This pandemic was attributed to a recombinant strain of *East African cassava mosaicvirus*-Uganda (EACMV-UG) (Zhou *et al.*, 1997) which derived a small stretch of the CP gene of *African cassava mosaic virus* (ACMV).

According to the latest established taxonomy for geminiviruses, cotton can be infected by eleven (11) different species on the Indian subcontinent, elucidating the status of CLCuD here is more complex than in Africa, caused by monophyletic virus (CLCuGeV). Those begomoviruses, which are mostly associated with CLCuD, are categorized into five different species engulfing CLCuMuV, CLCuKoV, CLCuBaV, CLCuAlV and CLCuGeV (Brown et al., 2015; Muhire, 2014), while the other six species including *African cassava mosaic virus* (ACMV), *Chickpea chlorotic dwarf virus* (CpCDV), *Okra enation leaf curl virus* (OEnLCV), *Papaya leaf curl virus*

(PaLCV), *Tomato leaf curl Bangalore virus* (ToLCBV) and *Tomato leaf curl New Delhi virus* (ToLCNDV) have also been identified in cotton (Chowda-Reddy *et al*., 2005; Hameed *et al.,* 2014; Kirti *et al.,* 2004; Kumar *et al.,* 2010; Manzoor *et al*., 2014; Mubin *et al*.,2012; Nawaz-ur-Rahman *et al*., 2012; Sinha *et al*., 2013; Zhou *et al*., 1998), belong to the genus begomovirus, except one species CpCDV, which is Mastrevirus, reported to infect cotton in this region. More interestingly, of these all species, Koch's Postulates was only satisfied in case of CLCuMuV, CLCuKoV and PaLCV (Briddon *et al*., 2001; Mansoor *et al.,* 2003). In the present investigation, CLCuKoV-Bu was unable to develop any or very mild (almost looking healthy plant) symptoms in the model plants when inoculated alone without a betasatellite. This indicates the need of the satellite for symptoms enhancement, which was achieved when virus was in association of its cognate- CLCuMB(Figure 4.12). This is also the best example of association in cotton, like synergism between EACMV-UG and ACMV in case of cassava, provide a distinct suppressor of PTGS that is required to establish and maintain the severe disease in cassava and cotton; either virus alone does not have this ability (Briddon *et al*., 2002; Vanitharani *et al*. 2004).

The phenomenon of recombination is not restricted to a single species, common among animal and plant RNA viruses (Simon and Bujarski, 1994) and DNA viruses (Roossinck, 1997). Recombination widely contributed in the emergence of new population and has been related to the advent of some serious disease, as discussed earlier (Zhou *et al*., 1997; Moffat, 1999; Padidam *et al*., 1999; Cui *et al*., 2004; Li *et al*., 2005; Martin *et al*., 2011). The pre-requisite for recombination to occur is co-infection of the same cell in a single host plant (Mansoor *et al*., 2005). Isolate C49A investigated from Layyah has overall high levels of sequence identity to CLCuKoV-Bu (acc.no. JF509747) with missing of some region, showing a chimeric nature, while

previous report from this region indicates the dominancy of CLCuKoV-Bu in cotton farms. Most common features in the isolate C49A inspected were derived from the same viruses from which Burewala and Shadadpur strains of Kokhran virus and Rajasthan virus was documented (Amrao *et al.*, 2010a, 2010b; Kumar *et al.*, 2010). Details of isolate C49A shows maximum nucleotide sequence identity in the virion sense with CLCuKoV-Ko, while in the complementary sense indicated high level of identities with CLCuMuV, elucidating that this component/molecule is likely an interspecific recombinant.

To verify Koch's Postulates, partial tandem repeats of isolates C28 (including begomovirus, alpha- and beta-satellite) and isolates C49 (including begomovirus, alpha- and beta-satellite) in a binary vector (discussed section 3.9) were produced. Agro-inoculation of various plants including cotton was performed. Unfortunately, cotton plants produced no symptoms indicative of older/mature plants selection at the time of inoculation. Host range studies for isolate C28 was also determined in *Nicotiana benthamiana, Nicotiana tabacum,* Solanaceae (tomato), and Cucurbitaceae.

All begomoviruses are transmitted with a sole vector whitefly (*Bemisia tabaci*) in a persistently circulating manner. Movement of viruliferous whitefly to the straddle regions without any physical hindrance in the way as well as by short distance are already known. It may be the sole cause with no time in spreading of identified virus isolates across border infecting economical crops (Rajagopalan *et al*., 2012). Begomoviruses on the Indian sub-continent is distinct from begomoviruses in the Middle East. This is most probably due to the geographical barrier to spreading/dispersal posed by the Sulaiman mountain range, which runs down mostly to the east of Balochistan, separating Pakistan from the Iranian plateau. Likewise, the begomoviruses of China and the sub-continent are also discrete, most likely due to the

obstacle posed by the Himalayan mountain range. Geographically, there is a natural hindrance for vector in movement from Pakistan to China, indicating that CLCuMuV was introduced from Pakistan to India and from there to China, not from here directly to China, as exist of a natural obstacle of Himalayan mountain range. Natural barriers do not completely, but mostly, limited the range of vector, but not virus. Unusually, such situation was occurred with TYLCV that originate from Middle East, when spread globally including the Caribbean's island, New World, North America and in some part of Asia, but does not report either in India or Pakistan, indicating the nature of virus is distinct (Lefeuvre *et al*., 2010; Moffat, 1999). Based on geographical location, the question thus stands up for how come CLCuGeV present in southern Pakistan (Tahir *et al.,* 2011). Possibility might be the alternative host of CLCuGeV which came here through trade, presence of ancient routes for trade between Africa and the subcontinent both overland and by sea, rather than whitefly migration.

Basically, monopartite begomoviruses are group/or categorized into three types, based on the association of DNA betasatellite molecule. The first are the true monopartite begomoviruses which do not require a betasatellite and, in the field, are not found associated with betasatellites. The viruses represent this group includes such as TYLCSV *and* TYLCV (Kheyr-Pour *et al.,* 1991; Dry *et al.,* 1993). The second group has a more facultative relationship with betasatellites. These monopartite begomoviruses do not require the betasatellite to infect plants but, in the field, are often (but not always) associated with betasatellites (Li *et al*., 2005). The last group comprises of viruses which have an absolute prerequisite for the betasatellite to infect the plant species from which they were isolated, such as the begomoviruses causing CLCuD and TYLCCNV (Briddon *et al*., 2001; Cui *et al*., 2004).

A wealth of information on the existence and variability of DNA satellite molecules associated with monopartite begomoviruses has been mostly from OW, specifically from Asian region (Bull *et al.,* 2004; Nawaz-ul-Rahman and Fauquet, 2009; Sivalingam *et al.,* 2010). In this study, begomovirus-betasatellite complex was identified that represents the third group of monopartite begomoviruses, which need its associated betasatellite for symptom determination. In-depth analysis of betasatellites, identified in this study, showed a recombinant version of Cotton leaf curl Multan betasatellite ($CLCuMB^{Bur}$), consisting of, for the most part, sequences derived from the Multan strain Cotton leaf curl Multan betasatellite ($CLCuMB^{Mul}$) but with a small amount of fragment/ region of sequence derived from another well known betasatellite called Tomato leaf curl betasatellite (Briddon *et al*., 2003; Amin *et al.,* 2006). The significance of this recombination between a cotton and tomato betasatellite remain unclear, in a large part due to the fact that we have yet to conclude the precise function of the SCR of betasatellites. Their sequence analysis showed all the basic features of betasatellites (Briddon *et al.,* 2003), as previously identified from this region.

For experimental studies, partial tandem repeats of CLCuMB (C28β and C49β) were agroinoculated alone to *Nicotiana benthamiana, Nicotiana tabacum, Solanaceae* (tomato), and *Cucurbitaceae,* produced no symptom, indicating its dependency on the helper virus for trans-replication. Similarly, agroinoculation of virus alone was unable to produce symptoms in a model plants *N. benthamiana, N. tabacum*, tomato and cucurbits. But, severe symptoms were noted when both partial tandem repeats were agroinculated to *N. benthamiana, N. tabacum, Solanaceae* (tomato), and *Cucurbitaceae*. These results are in agreement with the previous experiments performed by Amrao *et al*., (2010).

Previous studies indicating that the interaction of CLCuMB with cotton infecting begomoviruses (CLCuMuV, CLCuKoV-Ko, CLCuKoV-Bu, CLCuKoV-Sha) is not specific, to trans-replicate, shared only a small amount of region that produced step loop structure, indicating a relaxed relationship. Deletion analysis and naturally occurring mutants suggest that it is the sequence between the A-rich region and the satellite common region (SCR) which may be involved in Rep binding (Nawaz-ul-Rahman *et al*., 2009; Saunders *et al.,* 2008). This region is highly variable, contains sequences which resemble helper virus iterons and is possibly an adaptation all owing betasatellites to rapidly adapt to distinct begomoviruses (distinct Rep recognition sequences). Like other species of begomovirus, ToLCV have also the ability of trans-replication of CLCuMB (Saeed *et al.,* 2008).

The third component of begomovirus-betasatellite complexes are the alphasatellites. Aphasatellites are not a true satellite like betasatellite, as they are capable of autonomous replication, and thus best described as satellite-like molecule/ component. Literature reveals that aplasatellites are considered as opponent for symptoms enhancement during a complex establishment. It interfere in viral replication and deceased the symptoms. Unlike betasatellite, alphasatellites are not pathogenicity determinant (Mansoor *et al.,* 1999; Saunders and Stanley, 1999) except, possibly acting as a minor "dampener" of severity of the disease with which they are associated.

Alphasatellites were reported with the complex during the first epidemic of cotton in Pakistan (Mansoor *et al*., 1999). Surprisingly, only one type of alphasatellite, CLCuMA, exist with the complex. Amrao *et al*., (2010) could not find any alphasatellite associated with the second epidemic, with CLCuKoV-Bu infected cotton crop in Pakistan.

Suspected cotton samples collected on wide range with the aim to know the diversity of begomovirus in this vicinity as a continuation of previous efforts (Mansoor *et al*., 1999; Briddon and Markham, 2000; Briddon *et al*., 2001;Briddon *et al.*, 2002; Briddon *et al*., 2003; Mansoor *et al*., 2003a; Mansoor *et al.*, 2003b; Briddon *et al*., 2004, Amin *et al*., 2006, Mansoor *et al*., 2006; Saeed *et al.*, 2007; Nawaz-ul-Rahman *et al.*, 2009, Amrao *et al.*, 2010; Amrao *et al.*, 2010b; Farooq *et al*., 2011; Azhar *et al*., 2012; Sattar *et al.*, 2013; Briddon *et al*., 2014). In this study, I found the association of CLCuMA with the CLCuKoV-Bu-CLCuMB complex in all suspected samples, except one sample C32 from Faisalabad. This result indicating the recent mobilization of alphasatellite to the complex, involved in the second epidemic of CLCuD in Pakistan (Amrao *et al.*, 2010a). Interestingly, there was no sample found in which alphasatellite was present only with the virus, without presence of betasatellite.

Like betasatellite, the relationship of alphasatellite to the complex is also relaxed. Zia-ur-Rahman *et al.*, (2013) found *Gosyypium darwinii* symptomless alphasatellite (GDaSA) with CLCuKoV-Bu-CLCuMB complex, infecting *Luffa cylindrica* (a new alternate host), highlighting the diversity of cotton infecting begomovirus with other diverse combination and fecund relationship. Previous reports suggested association of alphasatellites with monopartite begomoviruses only. But, Paprotka *et al*., (2010) and Romay *et al*., (2010) isolated alphasatellite associated with bipartite begomoviruses infecting *Cleome affinis*, sida plants (weeds) and watermelon in the NW. This draw a diverse picture of begomoviruses with associated satellites in the world (Nawaz-ul-Rehamn and Faquet, 2009).

The exact function/role of an alphasatellite is still unclear, except autonomous replication (having Rep gene) and reduced virus and betasatellite DNA levels in *planta* (Mansoor *et al.*, 1999; Saunders and Stanley, 1999). While findings of Shahid (thesis)

showed a conflict with earlier reports that the alphasatellite associated with tomato disease in Oman states ameliorates the symptoms. Possible explanation of symptom amelioration in tomato plants may be the failure of interaction of alphasatellite with the complex. But, the results suggest that there may be more to the story which warrants more detailed investigation, for future.

Recombination between begomoviruses and their associated betasatellites appears to be common and it was the identification of such molecules that led directly to the discovery of betasatellites (Saunders *et al.,* 2000). In contrast, recombination between betasatellites and alphasatellites, as well as between begomoviruses and alphasatellites, appears to be less common. This may suggest that there is a mechanistic constraint to recombination between these molecules, or that the progeny of such recombination events is less well maintained. Recently Huang *et al*. (2013) reported a three-parent-recombinant molecule containing sequences originating from two distinct betasatellites and an alphasatellite. Unusually, this molecule retained the origins of replication of two of the parent molecules – of the alphasatellite and of the betasatellite.

In our study, isolate C50-alpha2 showed the same architecture, but, with the helper virus, rather than betasatellite, having two origins of replication of the parent molecules-of the alphasatellite and of the new identified strain of CLCuKoV-lay, which is less encountered. These finding divert attention to investigate in detail such molecule with two origin of replications, like C50-alpha2. Here is a pressing need to further investigate their effect via agro-inoculation.

The last part of this study was the development of polyclonal antibodies, against coat protein ofCLCuKoV-Bu, which is widely spread in this region, infecting wide host ranges of crops. Literature reveals that coat protein gene is highly conserved among all cotton infecting begomoviruses, was cloned in the bacterial expression system, globally

used for expression of recombinant protein (Rosano and Ceccarelli, 2014; Swartz, 2001; Baneyx, 1999). The gene was expressed under strong promotor of pET vector. The draw back of the bacterial expression system is the expression of gene in the aggregated form known to be inclusion bodies (IBs).

Next step was to yield purified recombinant protein isolation and solubilization of the IBs as illustrated by Fahnert *et al*., 2004, Palmer and Wingfield, 2004 and Rudolph *et al.,* 1996. Various techniques were used for disruption of IBs like French press or Sonication. Mechanical disruption of bacterial cells (sonication) was preferred followed by low speed centrifugation. The isolated IBs were washed with wash buffer containing Tritron-X 100, to remove the cell protein and other nonspecific cell materials which might hinder and reduce the refolding process as explained by Maachupalli-Reddy *et al*., 1997 and by Lilie *et al*., 1998. Isolated IBs were solubilized to obtained soluble protein. Normally, solubilization is done by using high concentration of Guanidine hydrochloride (Gnd-HCl), SDS, Urea etc. Here we used 6M Gnd-HCl with β-mercaptoethanol in Tris-HCl buffer, in which protein remained in its native like secondary structure (Kurucz *et al*., 1995).

Finally to convert protein into its native and active form, to refold to 3D structure was needed for the normal activities. There are several methods and techniques used by which recombinant proteins are refolded like, direct dilution, by dialysis, size exclusion chromatography matrix-assisted protein refolding membrane controlled denaturant removal and lastly through hydrophobic interaction chromatography (Jungbauer and Kaar, 2007; Ventura and Villaverde, 2006; Markus *et al*., 2005; Lefebvre *et al*., 2004; Tsumoto *et al.*, 2004; Vallejo and Rines, 2004).

Dilution method was applied as it is a simple refolding procedure and the solubilized proteins were not directly diluted at once in a refolding sink. Rather, it was added in the refolding sink in a pulsatile manner at a rate of 0.1 ml/min in refolding sink that contained renaturation solution. This phenomenon was done to achieve the productive refolding of the recombinant protein and avoid gaining unproductive aggregates following the procedure of Katoh and Katoh in 2000 and De Bernardez Clark *et al.,* 1999. More commonly arginine is used during the refolding process to avoid the aggregates formation (Tsumoto *et al*., 2003, 2004). But in our study, we used glycine instead of arginine as a cost effective substitute. Two amino acid, cysteine and cysteine, were also added in the refolding sink to provide the redox potential for disulphide bonds rearrangements. The refolding sink was placed at 4 °C with constant stirring to avoid the formation of aggregates during refolding process. Previous studies identified that low temperature reduces the folding rate of the protein (Vallejo and Rinas, 2004). Thus, to handle the situation, the time duration was increased up till 24 hours for renaturation with constant stirring.

Finally, the purified protein was subcutaneously injected to two rabbits of New Zealand white breed with Freund's complete adjuvant and Freund's incomplete adjuvant for a period of five months with one month interval. Blood was drawn after ten days for each booster dose and stored at -20 °C for future use.

ELISA is universally used for rapid detection of viruses infecting plants, animals, humans and birds (Khatabi *et al.,* 2012; Zhang *et al.,*2006; Gelfi*et al*., 1999; Ksiazek*et al.,* 1999; Mowat and Dawson, 1987; Katz *et al*., 1986; Clark and Adams, 1977). Antisera against CP of Burewala strain raised in rabbits were used to perform ELISA of all those infected *N. benthamiana* plants in the green house which were experimentally infected with Burewala strain of Kokhran virus. Antisera from both

rabbits showed a serologically positive response with the leaf extracts of infected plants. Our results are in agreement with the results obtained by Iracheta-Caardenas *et al.*, (2008) using polyclonal antibodies for PVX, Khatabi *et al*., (2012) for alfa mosaic virus and Gulati-Sakhuja *et al*., (2006) for *Pelargonium zonate spot virus.*

To check the effectiveness, antisera was also used to perform western blot and positive results were obtained with the post boost antisera (after immunization). The results of DAS-ELISA and western blot were in agreement with PCR results from the same infected *N. benthamiana* plants. These antibodies were initially utilized against infected plants of *N. benthamiana* grown in green house and will be used for viral detection in other naturally infected plants in future.

Analysis of CP gene sequence of CLCuKoV-Bu showed a high level of sequence identity with other strains of cotton infecting begomoviruses, like, CLCuKoV-Ko, CLCuMuV, CLCuKoV-Sha, CLCuKoV-La (Brown *et al.,* 2015). Burewala strain isolates from other plants also shared high levels of nucleotide and amino acid sequence identity when compared with strains representing phylogenetically distinct group of the same virus (Saleem *et al*., 2016; Sattar *et al.*, 2013). These observations suggest that the CP of CLCVs is highly conserved, possibly due to multiple crucial functions that it plays in its life cycles of the virus. Regardless, due to the presence of high levels of CP sequence identities among CLCVs, it is likely that polyclonal antibodies generated to recombinant CP of Burewala strain will cross react with wide range of CLCV isolates infecting cotton and other economical plants species in this region.

Conclusions

We conclude that cotton crops in Pakistan, specifically in the province of Punjab (precisely southern and central part of Punjab), and Sindh are severely affected by monopartite begomovirus CLCuKoV-Bu, the only strain which showed a widespeard distribution in this region of post-resistance in cotton. In this study, nine of ten districts showed widespread of the highly pathogenic strain of CLCuD infecting cotton and other malvaceous species.

Only a single new strain of CLCuD was identified, from Layyah region, indicating a recently emerged virus from parental strains. Associated cognates with cotton infecting begomoviruses were CLCuMBs and CLCuMAs, except one alphasatellite was chimeric, derived major sequence from CLCuMA and minor sequences from CLCuKoV-La and AcaYVMA.

Infectivity studies of CLCuKoV-Bu and CLCuKoV-La with its associated CLCuMB and CLCuMA satellites were encountered in *N. bethamiana, N. tabacum, Solanum lycoperscium* and *cucurbitacae* family, produced symptoms, including vein thickening, leaf curling, small/reduced leaf size and overall shunted plant growth.

Polyclonal antisera were raised against the coat protein of CLCuKoV-Bu isolate of CLCuD, expressed in *E. coli* using the pET expression system. The expressed protein was purified and injected into rabbits subcutaneously. Blood was drawn after each booster dose. Polyclonal antiserum was serologically positive in arrays of assays like DAS-ELISA and western blot for detection of viral coat protein of begomovirus.

REFERENCES

Adams, M. J., King, A. M. and Carstens, E. B. (2013). Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses. *Archives of Virology,* 158(9): 2023-2030.

Akhtar, K. P., Ullah, R., Saeed, M., Sarwar, N. and. Mansoor, S. (2014). China rose (*Hibiscus rosa-sinensis*): A new natural host of Cotton leaf curl Burewala virus in Pakistan. *Journal of Plant Pathology,* 96 (2): 385-389.

Alberter, B., Rezaian, M. A. and Jeske, H. (2005). Replicative intermediates of Tomato leaf curl virus and its satellite DNAs. *Virology,* 331: 441-448.

Ali, A., Natsuaki, T. and Okuda, S. (2004). Identification and molecular characterization of viruses infecting cucurbits in Pakistan. *Journal of Phytopathology,* 152: 677-682.

Amin, I., Mansoor, S., Amrao, L., Hussain, M., Irum, S., Zafar, Y., Bull, S. E. and Briddon, R. W. (2006). Mobilization into cotton and spread of a recombinant Cotton leaf curl disease satellite. *Archives of Virology,* 151:2055-2065.

Amin, I., Mansoor, S., Iram, S., Khan, M. A., Hussain, M., Zafar, Y., Bull, S. E., Briddon, R. W. and Markham, P. G. (2002). Association of a monopartite begomovirus producing subgenomic DNA and a distinct DNA betasatellite on *Croton bonplandianus* showing yellow vein symptoms in Pakistan. *Plant Disease,* 86: 444.

Amrao, L., Akhter, S., Tahir, M. N., Amin, I., Briddon, R. W., Mansoor, S. (2010b). Cotton leaf curl disease in Sindh province of Pakistan is associated with recombinant begomovirus components. *Virus Research,* 153: 161–165.

Amrao, L., Amin, I., Shahid, M. S., Briddon, R. W. and Mansoor, S. (2010a). Cotton leaf curl disease in resistant cotton is associated with a single begomovirus that lack an intact transcriptional activator protein. *Virus Research,* 152: 153-163.

Anonymous. (1993). A Research compendium of Cotton leaf curl viral disease and its management. PARC, Islamabad. p 62

Anwer, S., Tahir, M., Zaidi, N. S. S. Briddon R. W. (2012).First report of *Clerodendron yellow mosaic virus* infecting croton. *Journal of Plant Pathology,* 94.

Arguello-Astorga, G. R. and Ruiz-Medrano, R. (2001). An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. *Archives of Virology,* 146: 1465-1485.

Argüello-Astorga, G. R., Guevara-González, R. G., Herrera-Estrella, L. R. and Rivera-Bustamante R. F. (1994). Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology*, 203(1): 90-100.

Arguello-Astorga, G., Ascencio-Ibanez, J. T., Dallas, M. B., Orozco, B. M. and Hanley-Bowdoin, L. (2007). High frequency reversion of geminivirus replication protein mutants during infection. *Journal of Virology,* 81(20): 11005-11015.

Ashraf, M. A., Shahid, A. A., Rao, A. Q., Bajwa, K. S. and Husnain, T. (2014). Functional Characterization of a Bidirectional Plant Promoter *from Cotton leaf curl Burewala virus* using an *Agrobacterium*-Mediated Transient Assay.*Viruses.* 6: 223-242.

Aslam, M. and Gilani, A. (2000). Resistance of different cotton varieties to Cotton leaf curl virus under field conditions. *Journal of Research Sciences*,11:42–45.

Azar, M. T., Akhtar, S. and Mansoor, S. (2012). Cotton leaf curl Multan betasatellite strains cloned from *Gossypium barbadense* further supports selection due to host resistance. *Virus Genes*, 45:402–405.

Balaji, S., Black, M. C., French, R., Stenger, D. C. and Sunter, G. (2004). Spinach curly top virus: a newly described Curto species from southwest Texas with incongruent gene phylogenies. *Phytopathology*, 94: 772-779.

Baltes, N. J., Hummel, A. W., Konecna, E., Cegan, R., Bruns, A. N., Bisaro, D. M. and Voytas, D. F. (2015). Conferring resistance to geminiviruses with theCRISPR–Cas prokaryotic immune system. *Nature Plants*, DOI: 10.1038.

Bananej, K., Kheyr-Pour, A., Hosseini Salekdeh, G. and Ahoonmanesh, A. (2004). Complete nucleotide sequence of Iranian Tomato yellow leaf curl virus isolate: further evidence for natural recombination amongst begomoviruses. *Archives of Virology*, 149: 1435-1443.

Baneyx. F. (1999). Recombinant protein expression in Escherichia coli. *Current Opinions in Biotechnology*, 10: 411-421.

Bashir, M., Jamali, R. A. and Ahmad, Z. (2006). Genetic resistance in mungbean and mash bean germplasm against Mungbean yellow mosaic begomovirus. *Mycopath*, 4(2): 1-4.

Baulcombe, D. (2004). RNA silencing in plants. *Nature*, 431: 356-363.

Bisaro, D. (1996). Geminivirus DNA replication. In DNA replication in eukaryotic cells, Cold Spring Harbor Laboratory press. Pp: 833—854.

Blake, M. S., Johnston, K. H., Russell-Jones, G. J. and Gotschlich, E. C. (1984). A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. *Analytical Biochemistry*, 136(1):175-179.

Bottcher, B., Unseld, S., Ceuleamns, H., Russell, R. B. and Jeski, H. (2004). Geminate structures of *African cassava mosaic virus*. *Journal of Virology*, 78: 6758-6765.

Boulton, M. I., King, D. I., Markham, P. G., Pinner, M. S. and Davies, J. W. (1991). Host range and symptoms are determined by specific domains of the maize streak virus genome. *Virology*, 181: 312-318.

Bradford, M. M.(1976). Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical. Biochemistry*,72: 248–254.

Briddon, R. W. and Markham, P. G. (2000). Cotton leaf curl disease. *Virus Research*, 71(1-2): 151-159.

Briddon, R. W., Akbar, F., Iqbal, Z., Amrao, L., Amin, I., Saeed, M., and Mansoor, S. (2014). Effects of genetic changes to the begomovirus/betasatellite complex causing Cotton leaf curl disease in South Asia post-resistance breaking. *Virus Research*,186:114-119.

Briddon, R. W., Bedford, I. D., Tsai, J. H. and Markham, P. G. (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, R. W., Bull, S. E., Amin, I., Idris, A. M., Mansoor, S., Bedford I. D., Dhawan, P., Rishi, N., Siwatch, S. S., Abdel-Salam, A. M., Brown, J. K., Zafar, Y. and Markham, P. G. (2003). Diversity of DNA β: a satellite molecule associated with some monopartite begomoviruses. *Virology*, 312: 106–121.

Briddon, R. W., Bull, S. E., Amin, I., Mansoor, S., Bedford, I. D., Rishi, N., Siwatch, S. S., Zafar, Y., Abdel-Salam, A. M. and Markham, P. G. (2004). Diversity of DNA 1: a satellite-like molecule associated with monopartitebegomovirus – DNA β complexes. *Virology*, 324: 462 – 474.

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

Briddon, R. W., Pinner, M. S., Stanley, J. and Markham, P. G. (1990). Geminivirus coat protein replacement alters insect specificity. *Virology*, 177: 85-94.

Briddon, R.W. and Stanley, J. (2006). Sub-viral agents associated with plant single-stranded DNA viruses. *Virology,* 344(1): 198-210.

Briddon, R.W. Bull, S. E., Mansoor, S., Amin, I and Markham, P. G. (2002). Universal primers for the PCR-mediated amplification of DNA b a molecule associated with some monopartite begomoviruses. *Molecular Biotechnology,* 20: 315 – 318.

Brough, C. L., Hayes, R. J., Morgan, A. J., Coutts, R. H. A., and Buck, K. W. (1988). Effects of mutagenesis in vitro on the ability of cloned tomato golden mosaic virus DNA to infect *Nicotiana benthamiana* plants. *Journal of General Virology*, 69: 503-514.

Brown, J. (1994). Current status of *Bemisia tabaci* as a plant pest and virus vector in agroecosystems worldwide. *FAO Plant Protection Bulletin*, 42: 3-32.

Brown, J. K., Zerbini, F. M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J. C., Fiallo-Olivé, E., Briddon, R. W., Hernández-Zepeda, C., Idris, A, Malathi, V. G., Martin, D. P., Rivera-Bustamante, R., Ueda, S.and Varsani, A. (2015). Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of Virology*, 160(6):1593-1619.

Brown, J. K., Coats, S. A., Bedford, I. D., Markham, P. G., Bird, J. and Frohlich, D. R. (1995). Characterization and distribution of esterase electromorphs in the

whitefly, *Bemisia tabaci* (Genn.)(Homoptera:*Aleyrodidae*). *Biochemical Genetics*, 33: 205-214.

Brown, J. K., Fauquet, C. M., Briddon, R. W., Zerbini, F. M., Moriones, E. and Navas-Castillo, J. (2012). *Geminiviridae*: In: King A. M. Q., Adams M. J., Carstens, E. B., Lefkowits E. J.,(eds) Virus Taxonomy-Ninth Report of the International Committee on Taxonomy of Viruses. Associated Press, Elsevier Inc. London, Waltham, San Diego pp. 351-373.

Buchmann, R. C., Asad, S., Wolf, J. N., Mohannath, G. and Bisaro, D. M. (2009). Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. *Journal of Virology,* 83: 5005–5013.

Bull, S. E. (2004). Diversity of cassava infecting geminiviruses in Kenya. Thesis, Department of Disease and Stress Biology, John Innes Centre, United Kingdom.

Bull, S. E.,Briddon, R. W.,Sserubombwe., W. S,Ngugi, K.,Markham, P. G. and Stanley, J. (2006). Genetic diversity and phylogeography of cassava mosaic viruses in Kenya.*Journal of General Virology,* 87(10):3053-65.

Bull, S. E., Tsai, W. S., Briddon, R. W., Markham, P. G., Stanley, J. and Green, S. K. (2004). Diversity of begomovirus DNA β satellites of non-malvaceous plants in east and south-east Asia. *Archives of Virology,* 149: 1193–1200.

Campos-Olivas, R., Louis, J. M., Clérot, D., Gronenborn, B. and Gronenborn, A. M. (2002). 1H, 13C, and 15N assignment of the N-terminal, catalytic domain of the replication initiation protein from the geminivirus TYLCV. *Journal of Biomolecular* NMR 24, 73-74.

Castillo, A. G., Collinet, D., Deret, S., Kashoggi, A. and Bejarano, E. R. (2003). Dual interaction of plant PCNA with geminivirus replication accessory protein (REn) and viral replication protein (Rep). *Virology*, 312: 381-394.

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

Cheng, X., Wang, X., Wu, J., Briddon, R. W. and Zhou, X. (2011). βC1 encoded by Tomato yellow leaf curl China betasatellite forms multimeric complexes in vitro and in vivo. *Virology*, 409(2): 156-162.

Choudhury, N. R., Malik, P. S., Singh, D. K., Islam, M. N., Kaliappan, K. and Mukherjee, S. K. (2006). The oligomeric Rep protein of *Mungbean yellow mosaic India virus* (MYMIV) is a likely replicative helicase. *Nucleic Acids Research*, 34: 6362–6377.

Chowda Reddy, R V., Muniyappa, V., Colvin, J., Seal, S. (2005). A new begomovirus isolated from *Gossypium barbadense* in southern India. *Plant Pathology*, 54: 570–570.

Cohen, S. N., Chang, A. C., Boyer, H. W. and Helling, R. B. (1973). Construction of biologically functional bacterial plasmid in vitro. *National Academy of Sciences of the Unites States of America*, 70(11): 3240-3244.

Collin, S., Fernández-Lobato, M., Gooding, P. S., Mullineaux, P. M. and 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.

Collmer, C. W. and Howell, S. H. (1992). Role of satellite RNA in the expression of symptoms caused by plant viruses. *Annual Review of Phytopathology,* 30: 419-442.

Cui, X. F., Tao, X. R., Xie, Y., Fauquet, C. M. and Zhou, X. P. (2004). ADNAbeta associated with Tomato yellow leaf curl China virus isrequired for symptom induction. *Journal of Virology,* 78: 13966–13974.

Cui, X., Li, G., Wang, D., Hu, D. and Zhou, X. (2005). A begomovirus DNAβ-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. *Journal of Virology,* 79: 10764-10775.

Cuong, H. V., Hai, L. V., Tiep, T. N. and Hao, N. B. (2011). Molecular characterization of Tomato leaf curl Hainan virus and Tomato leaf curl Hanoi virus in Vietnam. *International Society of Southeast Asian Agricultural Sciences*, 2: 70-82.

De Almeida, M. M. S., Jain, S., Barroso, P. A. V., Hoffmann, L. V., de Lucena, M. G., Resende, R. D. O.and Inoue-Nagata, A. K. (2013). Complete sequence of a new bipartite begomovirus infecting cotton plants in Brazil. *Genome Announcements*. 1(6):doi:10.1128/genomeA.00661-13.

De BernardezClark, E., Schwartz, E. and Rudolph, R. (1999). Inhibition of aggregation side reactions during in vitroprotein folding. *Methods in Enzymology,* 309:217–236.

Desbiez, C., David, C., Mettouchi, A., Laufs, J., and Gronenborn, B. (1995). Rep protein of Tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proceedings of the National Academy of Sciences and the United States of America,* 92: 5640– 5644.

Ding, S. W., Li, H., Lu, R., Li, F. and Li, W. X. (2004). RNA silencing: a conserved antiviral immunity of plants and animals. *Virus Research,* 102: 109– 115.

Donson, J., Morris-Krsinich, B. A. M., Mullineaux, P. M., Boulton, M. I. and Davies, J. W. (1984). A putative primer for second-strand synthesis of *Maize streak virus* is virion-associated. *EMBO Journal,* 3: 3069– 3073.

Dry, I. D., Rigam, J. E., Krake, L. R., Mullineaux, P. M. and Rezaian, M. A. (1993). Nucleotide sequence and genome organization of Tomato leaf curl geminivirus. *Journal of General Virology,* 74: 147– 151.

Duffy, S. and Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus* (TYLCV). *Journal of Virology,* 82(2): 957– 965.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics,* 5: 1– 19.

Egelkrout, E. M., Mariconti, L., Settlage, S. B., Cella, R., Robertson, D. and Hanley-Bowdoin, L. (2002). Two E2F elements regulate the proliferating cell nuclear antigen promoter differently during leaf development. *Plant Cell,* 14: 3225-3236.

Eini, O., Dogra, S., Selth, L. A., Dry, I. B., Randles, J. W. and Rezaian, M. A. (2009). Interaction with a host ubiquitin-conjugating enzyme is required for the pathogenicity of a geminiviral DNA beta satellite. *Molecular Plant and Microbial Interaction,* 22(6): 737– 746.

Elmer, J. E., Brand, L., Sunter, G., Gardiner, W. E., Bisaro, D. M. and Rogers S. G. (1988). Genetic analysis of the tomato mosaic golden virus II. The product of the ALI coding sequence is required for replication. *Nucleic. Acids Research,* 16: 7043– 7060.

Etessami, P., Callis, R., Ellwood, S. and Stanley, J. (1988). Delimitation of essential genes of cassava latent virus DNA 2. *Nucleic Acids Research,* 16: 4811– 4829.

Etessami, P., Watts, J. and Stanley, J. (1989). Size reversion of African cassava mosaic virus coat protein gene deletion mutants during infection of *Nicotiana benthamiana.Journal of General Virology,* 70: 277– 289.

Fahnert, B., Lilie, H. and Neubauer, P. (2004). Inclusion bodies: Formation and utilization. *Advance Biochemical Engineering/Biotechnology*, 89: 93– 142.

Fareed, I., Tahir, M. and Briddon, R. W. (2012). First report of Cotton leaf curl Burewala virus infecting *Ricinus communis*. *Journal of Plant Pathology,* 94 (S85).

Fargette, D., Konate, G., Fauquet, C., Muller, E., Peterschmitt, M. and Thresh, J. M. (2006). Molecular ecology and emergence of tropical plant viruses. *Annual Review of Phytopathology*, 44: 235– 260.

Farooq, A., Farooq, J., Mahmood, A., Shakeel, A., Rehman, A., Batool, A., Riaz, M., Shahid, M. T. H. and Mehboob, S. (2011). An overview of Cotton leaf curl virus disease (CLCuD) a serious threat to cotton productivity. *Australian Journal of Crop Science,* 5: 1823–1831

Fauquet, C. M., Bisaro, D. M., Briddon, R. W., Brown, J. K., Harrison, B. D., Rybicki, E. P., Stenger, D. C. and Stanley, J. (2003). Revision of taxonomic criteria for species demarcation in the family *Geminiviridae,* and an updated list of begomovirus species. *Archives of virology,* 148: 405– 421.

Fiallo-Olivé, E., Chirinos, D. T., Geraud-Pouey, F., Moriones, E. and Navas-Castillo, J. (2014). Complete genome sequence of *Jacquemontia yellow mosaic virus*, a novel begomovirus from Venezuela related to other New World bipartite begomoviruses infecting *Convolvulaceae*. *Archives of Virology,* 159(7):1857-60.

Fondong,V. N., Pita, J. S., Rey, M. E., de Kochko, A., Beachy, R. N. and Fauquet, C. M. (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *Journal of General Virology,* 81: 287– 297.

Frischmuth, S., Wege, C., Hülser, D. and Jeske, H. (2007). The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast. *Protoplasma* 230: 117–123.

Futami, J., Tsushima, Y., Tada, H., Seno, M. and Yamada, H. (2000). Convenient and efficient in vitro folding of disulfide-containing globular protein from crude bacterial inclusion bodies. *Journal of Biochemistry,* 127: 435–441.

Gafni, Y. and Epel, B.L. (2002). The role of host and viral proteins in intra and inter-cellular trafficking of geminiviruses. *Physiological and Molecular Plant Pathology,* 60: 231– 241.

García-Andrés, S., Monci, F., Navas-Castillo, J. and 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: 433 – 442.

García-Andrés, S., Tomás, D. M., Sánchez-Campos, S., Navas-Castillo, J. and Moriones, E. (2007). Frequent occurrence of recombinants in mixed infections of Tomato yellow leaf curl disease-associated begomoviruses. *Virology,*365(1): 210– 219.

Garrido-Ramirez, E. R., Sudarshana, M. R. and Gilbertson, R. L. (2000). Bean golden yellow mosaic virus from Chiapas, Mexico: characterization,

pseudorecombination with other bean-infecting geminiviruses and germ plasm screening. *Phytopathology,* 90: 1224-1232.

Gaur, R. K., Prajapat, R.., Marwal, A., Sahu, A. and Rathore, M. S. (2011). First report of a begomovirus infecting*Mimosa pudica* in India. *Journal of Plant Pathology,* 93(4): Supplementary pp. S4.80

Geetanjali, A.S., Shilpi, S. and Mandal, B. (2013). Natural association of two different betasatellites with *Sweet potato leaf curl virus* in wild morning glory (*Ipomoea purpurea*) in India. *Virus Genes,* 47(1): 184–188.

Gelfi, J., Chantal, J., Phong, T. T., Py, R. and Boucraut-Baralon, C. (1999). Development of an ELISA for detection of myxoma virus-specific rabbit antibodies: test evaluation for diagnostic applications on vaccinated and wild rabbit sera. *Journal of Veterinary Diagnostic Investigation,* 11(3):240-455.

George, B., Kumar, R. V, and 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

Gilbertson, R. L., Sudarshana, M., Jiang, H., Rojas, M. R. and Lucas, W. J. (2003a). Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: insights into DNA trafficking. *Plant Cell,* 15: 2578-2591.

Givord, L., Fargeite, D., Kounounguissa, B., Thowenel, J. C. and Van Regenmortel, M. H. V. (1994). Diagnostic d'infection à geminivirus à l'ade d'anticorps monoclonaux dirigés contre le virus de la mosaïque africaine du manioc. *Agricultural Research Bullietin Gembloux,* 29(2):161-168.

Graham, A. P., Martin, D. P.and Roye, M. E. (2010). Molecular characterization and phylogeny of two begomoviruses infecting *Malvastrum americanum* in

Jamaica: evidence of the contribution of inter-species recombination to the evolution of malvaceous weed-associated begomoviruses from the northern Caribbean. *Virus Genes*, 40(2): 256– 266.

Guerra-Peraza, O., Kirk, D., Seltzer, V., Veluthambi, K., Schmit, A. C., Hohn, T. and Herzog, E. (2005). Coat proteins of Rice tungro bacilliform virus and Mungbean yellow mosaic virus contain multiple nuclear-localization signals and interact with importin α. *Journal of General Virology,* 86: 1815–1826.

Gulati-Sakhuja, A., Sears, J. L., Nunez, A., and Liu, H.-Y. (2009). Production of polyclonal antibodies against Pelargonium zonate spot virus coat protein expressed in Escherichia coli and application for immunodiagnosis. *Journal of Virology Methods*, 160:29-37.

Gutierrez, C. (1999). Geminivirus DNA replication. *Cell and Molecular Life Sciences,* 56: 313-329.

Gutierrez, C. (2000). Geminiviruses and the plant cell cycle.*Plant Molecular Biology*,43(5-6):763-72.

Gutierrez, R. A., Green, P. J., Keegstra, K. and Ohlrogge, J. B. (2004). Phylogenetic profiling of the Arabidopsis thaliana proteome: what proteins distinguish plants from other organisms? *Genome Biology,* 5: 53.

Haider M. S., Tahir, M., Evants, A. A. F.and Markham, P. G. (2007b). Coat protein gene sequence analysis of three begomovirus isolates from Pakistan and their affinities with other begomoviruses. *Pakistan Journal of Zoology,* 39: 165-170.

Haider, M. S., Liu, S., Ahmed, W., Evans, A. A. F. and Markham, P. G. (2005). Detection and relationship among begomoviruses from five different host plants, based on ELISA and Western Blot analysis. *Pakistan Journal of Zoology,* 37(1): 69-74.

Haider, M.S., Tahir, M., Latif, S. and Briddon, R. W. (2006). First report of Tomato leaf curl New Delhi virus infecting *Eclipta prostrata* in Pakistan. *Plant Pathology*, 55: 285.

Haider, M.S., Tahir, M., Saeed, A., Ahmed, S., Parveen, R. and Rashid, N. (2008). First report of begomovirus infecting the ornamental plant *Vinca minor* L. *Australasian plant Disease Note*, 3: 150– 151.

Haider, S., Tahir, M., Saeed, M., Shah, *A. H.*, Rashid, N., Javid, M. A. and Iqbal, J. (2007a). *Vinca minor*; another host of a tomato infecting begomovirus in Pakistan. *African Crop Science Conference Proceedings*, 8: 905– 907.

Hak, H., Levy, Y., Chandran, S. A., Belausov, E., Loyter, A., Lapidot, M. and Gafni, Y. (2015). TYLCV-Is movement in planta does not require V2 protein. *Virology*, 477:56–60.

Hameed, U., Zia-ur-Rehman, M., Herrmann, H. W., Haider, M. S. and Brown, J. K. (2014). First Report of *Okra enation leaf curl virus* and Associated Cotton leaf curl Multan betasatellite and Cotton leaf curl Multan alphasatellite Infecting Cotton in Pakistan: A new member of the Cotton leaf curl disease Complex. *Plant Disease*, 98: 1447–1447.

Hanley-Bowdoin, L., Elmer, J. S. and Rogers, S. G. (1989). Functional expression of the leftward open reading frames of the component of tomato golden mosaic virus in transgenic tobacco plant. *Plant Cell*, 1: 1057– 1068.

Hanley-Bowdoin, L., Elmer, J. S. and Rogers, S. G. (1990). Expression of functional replication protein from *Tomato golden mosaic virus* in transgenic tobacco plants. *Proceedings of the National Academy of Sciences USA*, 87: 1446– 1450.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S. and Robertson, D. (1999). Geminiviruses: models for plant DNA replication, transportation, and cell cycle regulation. *Critical Review of Plant Sciences,* 18: 71– 106.

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

Harimalala, M.,De Bruyn, A.,Hoareau, M.,Andrianjaka, A.,Ranomenjanahary, S.,Reynaud, B.,Lefeuvre, P.andLett, J. M. (2013). Molecular characterization of a new alphasatellite associated with a cassava mosaic geminivirus in Madagascar. *Archives of Virology,* 158(8):1829-1832.

Harrison, B. D. (1985). Advances in geminivirus research. *Annual Review of Phytopathology,* 23: 55– 82.

Harrison, B. D. and Robinson, D. J. (1999). Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). *Annual Review of Phytopathology,* 37: 369– 398.

Harrison, B. D., and Robinson, D. J. (2005). Another quarter century of great progress in understanding the biological properties of plant viruses. *Annuals of Applied Biology,* 146: 15-37.

Harrison, B. D., Liu, Y. L., Khalid, S., Hameed, S., Otim-Nape, G. W. and Robinson, D. J. (1997a). Detection and relationships of Cotton leaf curl virus and allied whitefly-transmitted geminiviruses occurring in Pakistan. *Annals of Applied Biology,* 130: 61-75.

Harrison, B. D., Swanson, M. M. and Fargette, D. (2002). Begomovirus coat protein: serology, variation and functions. *Physiological and Molecular Plant Pathology,* 60: 257–271.

He, Z. F., Mao, M. J., Yu, H., Li, H. P.and Chen, X. (2009). Molecular characterization of a distinct begomovirus infecting *Allamanda cathartica* in Guangdong, China. *Archives of Virology*, 154(8): 1199– 1202.

Hernández-Zepeda, C., Brown, J. K., Moreno-Valenzuela, O. A., Argüello-Astorga. G., Idris, A. M., Carnevali, G. and Rivera-Bustamante, R. F. (2010). Characterization of Rhynchosia yellow mosaic Yucatan virus, a new recombinant begomovirus associated with two fabaceous weeds in Yucatan, Mexico. *Archives of Virology*, 155(10): 1571– 1579.

Heyraud, F., Matzeit, V., Kammann, M., Schaefer, S., Schell, J. and Gronenborn, B. (1993). Identification of the initiation sequence for viral-strand DNA synthesis of wheat dwarf virus. *EMBO Journal*, 12: 4445– 4452.

Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J. and Gronenborn, B. (1995). Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic Acids Research*, 23: 910– 916.

Hina, S., Javed, M. A., Haider, S. and Saleem, M. (2012). Isolation and sequence analysis of cotton infecting begomoviruses. *Pakistan Journal of Botany*, 44: 223– 230.

Huang, C., Xie, Y., Zhao, L., Ren, H., Li. Z. (2013). A naturally occurring defective DNA satellite associated with a monopartite begomovirus: Evidence for recombination between alphasatellite and betasatellite. *Viruses*, 5: 2116– 2128.

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

Hussain, M., Mansoor, S., Iram, S., Zafar, Y. and Briddon, R.W. (2007). The hypersensitive response to Tomato leaf curl New Delhi virus nuclear shuttle

protein is inhibited by transcriptional activator protein. *Molecular Plant Microbe Interaction,* 20(12): 1581– 1588.

Hussain, M., Qazi, J., Mansoor, S., Iram, S., Bashir, M., and Zafar, Y. (2004). First Report of *Mungbean yellow mosaic India virus* on mungbean in Pakistan. *New Disease Report*. 13: 9.

Hussain. K., Mehmood, M. A., Naheed, N., Khan, A. H., Akram, A., Rahman, M., Azeem, F. and Shabnum, S. (2013). Molecular characterization of begomovirus associated alphasatellite from an asymptomatic weed plant; *Xanthium strumarium* L. *Pakistan Journal of Life and Social Sciences,*11(3): 233– 237.

Hussain, T. and Ali, M. (1975). A review of cotton diseases of Pakistan.*Pakistan Cotton,* 19:77–78.

Idris, A. M., Hussein, M. H., Abdel-Salam, A. M. and Brown, J. K. (2002). Genetic variability of satellite DNA associated with monopartite begomoviruses of okra and hollyhock exhibiting vein-thickening symptoms. *Arab Journal of Biotechnology,* 5:67-82.

Ilyas, M.,Nawaz, K.,Shafiq, M.,Haider, M. S.andShahid, A. A. (2013). Complete nucleotide sequences of two begomoviruses infecting Madagascar periwinkle (*Catharanthus roseus*) from Pakistan. *Archives of Virology*,158(2):505-510.

Ilyas, M., Qazi, J., Mansoor, S. and Briddon, R. W. (2009). Molecular characterization and infectivity of a "Legumovirus" (genus *Begomovirus*: family *Geminiviridae*) infecting the leguminous weed *Rhynchosia minima* in Pakistan. *Virus Research*, 145(2): 279– 284.

Ilyas, M., Qazi, J., Mansoor, S. and Briddon, R. W. (2010). Genetic diversity and phylogeography of begomoviruses infecting legumes in Pakistan. *Journal of General Virology*, 91(8):2091– 2101.

Iqbal, M. J., Hussain, W., Zia-Ur-Rehman, M., U. Hameed, U. and Haider, M. S. (2016). First report of *Chilli leaf curl virus* and associated alpha- and beta-satellite DNAs infecting Nettle Weed (*Urtica dioica*) in Pakistan. *Plant Disease*, 100(4): 870– 870.

Iram, S., Amrao, L., Mansoor, M. S., Malik, A. H., Briddon, R. W. and Zafar, Y. (2004). First report of a begomovirus associated with leaf curl disease of *Duranta erecta* in Pakistan. *New Disease Reports*, 10: 8.

Iram, S., Amrao, L., Mansoor, M. S., Malik, A. H., Briddon, R. W. and Zafar, Y. (2005). First report of a begomovirus associated with leaf curl disease of *Duranta erecta* in Pakistan. *Plant pathology*, 54(2): 260.

Jeske H. (2009). Geminiviruses. *Current Topics in Microbiology and Immunology*. 331: 185 –226.

Jeske, H. (2007). Replication of geminiviruses and use of rolling circle amplification for their diagnosis. In: *Czosnek, H. (Ed), Tomato leaf Curl Virus Disease pp.* 141– 156 Springer, Dordrecht.

Jeske, H., Lutgemeier, M. and Preiss, W. (2001). DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *EMBO Journal*, 20: 6158– 6167.

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

Jungbauer, A. and Kaar, W. (2007). Current status of technical protein refolding.*Journal of Biotechnology*, 128: 587 – 596.

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

Kamaal, N., Akram, M., Pratap, A. and Yadav, P.(2013). Characterization of a new begomovirus and a beta satellite associated with the leaf curl disease of French bean in northern India. *Virus Genes*, 46(1):120– 127.

Katoh, S. and Katoh, Y. (2000). Continuous refolding of lysozyme with fed-batch addition of denatured protein solution. *Process Biochemistry***,** 35(10): 1119–1124.

Katz, D., Hilliard, J. K., Eberle, R. and Lipper, S. L. (1986). ELISA for detection of group-common and virus-specific antibodies in human and simian sera induced by herpes simplex and related simian viruses. *Journal of Virology Methods*, 14(2):99-109.

Khan, A. A., Naqvi, Q. A., Khan, M. S., Singh, R.and Raj, S. K. (2005). First report of a begomovirus infecting *Calendula* in India. *Plant Pathology*, 54(4): 569.

Khatabi, B., He, B., and Hajimorad, M. R. (2012). Diagnostic potential of polyclonal antibodies against bacterially expressed recombinant coat protein of Alfalfa mosaic virus. *Plant Disease,* 96:1352-1357.

Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G. P., Crespi, S. and Gronenborn, B. (1991). Tomato yellow leaf curl virus fromSardinia is a whitefly-transmitted monopartite geminivirus. *Nucleic Acids Research,* 19: 6763–6769.

Kirthi, N., Priyadarshini, C. G. P., Sharma, P., Maiya, S. P., Hemalatha, V., Sivaraman, P., Dhawan, P., Rishi, N., Savithri, H. S. (2004). Genetic variability of begomoviruses associated with Cotton leaf curl disease originating from India. *Archives of Virology*, 149: 2047–2057.

Kon, T., Kuwabara, K., Hidayat, S. H. and Ikegami M. (2007). A begomovirus associated with Ageratum yellow vein disease in Indonesia: evidence for natural

recombination between Tomato leaf curl Java virus and Ageratum yellow vein virus-[Java]. *Archives of Virology*, 152(6):1147–1157.

Kong, L. and 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, L. J., Orozco, B, M., Roe, J. L., Nagar, S, S., Ou, S., Feiler, H. S., Durfee, T., Miller, A. B., Gruissem, W., Robertson, D. and 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 Journal, 19: 3485–3495.

Kornberg, A. and Baker, T. A. (1992). *DNA replication*. 2nded. New York: W. H. Freeman and Co.

Kothandaraman, S. V., Devadason, A. and Ganesan, M. V. (2016).Seed-borne nature of a begomovirus, Mung bean yellow mosaic virus in black gram.*Applied Microbiology and Biotechnology*, 100(4): 1925–33.

Krupovic, M., Ravantt, J. J. and Bamford, D. H. (2009). Geminiviruses: a tale of plasmid becoming a virus. *BMC Evolutionary Biology*, 9: 112.

Ksiazek, T.G., West, C. P., Rollin, P. E., Jahrling, P. B. and peters, C. J.(1999).ELISA for the Detection of Antibodies to Ebola Viruses. *The Journal of Infectious Diseases*, 79.

Kumar, A., Kumar, J. and Khan, J. A. (2010). Sequence characterization of Cotton leaf curl virus from Rajasthan: phylogenetic relationship with other members of geminiviruses and detection of recombination. *Virus Genes,* 40: 282– 289.

Kumar, J., Gunapati, S., Kumar, J., Kumari, A., Kumar, A., Tuli, R. and Singh, S. P. (2014).Virus-induced gene silencing using a modified betasatellite: a potential

candidate for functional genomics of crops. *Archives of Virology*, 159(8): 2109–2113.

Kumar, J., Singh, S. P., Kumar, A., Khan, J. A and Tuli, R. (2013). Recombination study using Radish leaf curl virus isolates.*African Journal of Microbiology Research*, 7(27): 3542–3546.

Kumari, P., Singh, A. K., Sharma, V. K., Chattopadhyay, B.and Chakraborty, S. (2011). A novel recombinant tomato-infecting begomovirus capable of trans complementing heterologous DNA-B components. *Archives of Virology*, 156(5): 769–783.

Kunik, T., Palanichelvam, K., Czosnek, H., Citovsky, V. and Gafni, Y. (1998). Nuclear import of the capsid protein of *Tomato yellow leaf curl virus* (TYLCV) in plant and insect cells. *Plant Journal*, 13: 393–399.

Kurucz, I., Titus, J. A., Jost, C. A. and Segal, D. M. (1995). Correct disulphide pairing and efficient refolding of detergent-solubilized single-chain Fv proteins from bacterial inclusion bodies. *Molecular Immunology*, 12: 1443–1452.

Laufs, J., Schumacher, S., Geisler, N., Jupin, I. and Gronenborn, B. (1995b). Identification of the nicking tyrosine of geminivirus Rep protein. *FEBS Letters*, 377: 258–262.

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S.G., Schell, J. and Gronenborn, B. (1995a). *In vitro* cleavage and joining at the origin of replication by the replication initiator protein of Tomato yellow leaf curl virus. *Proceedings of the national Academy of Sciences USA*, 92: 3879–3883.

Lauren, J., Crawford, Toba, A. M., Osman, Frank, P. B., Robert, H. A. C., Clive, M. B. and Kenneth, W. B. (2006). Molecular characterization of a partitivirus from Ophiostoma Himal-ulmi. *Virus Genes*, 33: 33–39.

Lazarowitz, S. G. (1999). Probing plant cell structure and function with viral movement proteins. *Current Opinion in. Plant Biology,* 2: 332–338.

Lazarowitz, S. G. and Beachy, R. N. (1999). Viral movement proteins as probes for intracellular and intercellular trafficking in plants. *Plant cell,* 11: 535–548.

Lee, G., Shim, H. K., Kwon, M. and Lee, S. (2013).A nucleic acid hydrolyzing recombinant antibody confers resistance to curtovirus infection in tobacco. *Plant Cell Tissue and Organ Culture*, 115:179-187.

Lefebvre, B. G., Gage, M. J. and Robinson, A. S. (2004). Maximizing recovery of native protein from aggregates by optimizing pressure treatment. *Biotechnology Program,* 20: 623 – 629.

Lefeuvre, P., Martin, D. P., Harkins, G., Lemey, P., Gray, A. J., Meredith, S., Lakay, F., Monjane, A., Lett, J. M., Varsani, A.and Heydarnejad, J. (2010). The spread of *Tomato yellow leaf curl virus* from the Middle East to the World. *PLoS Pathogens,* 6: e1001164.

Leke, W. N., Sattar, M. N., Ngane, E. B., Ngeve, J. M., Kvarnheden, A. and Judith, K. Brown, J. B. (2013). Molecular characterization of begomoviruses and DNA satellites associated with okra leaf curl disease in Cameroon.*Virus Research,* 174: 116– 125.

Li, Z. H., Xie, Y. and Zhou, X. P. (2005). Tobacco curly shoot virus DNAbis not necessary for infection but intensifies symptoms in a host-dependent manner. *Phytopathology,* 95: 902–908.

Lilie, H., Schwarz, E. and Rudolph, R. (1998). Advances in refolding of proteins produced in E. coli.*Current Opinion in Biotechnology,* 9(5):497-501.

Lotrakul, P., Valverde, R. A. and Landry, A. D. (2000). Biological and Molecular Properties of a Begomovirus from Dicliptera sexangularis. *Phytopathology,* 90(7):723-729.

Lozano, G., Trenado, H. P., Fiallo- Olivé, E., Chirinos, D., Geraud-Pouey, F., Briddon, R. W. and Navas-Castillo, J. (2016). Characterization of Non- coding DNA Satellites Associated with Sweepoviruses (Genus *Begomovirus, Geminiviridae*)–Definition of a Distinct Class of Begomovirus- Associated Satellites. *Frontiers in Microbiology*, 7:162.

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

Hussain, M., S. Mansoor, S., I. Amin, I., S. Iram, S., Y. Zafar, Y., Malik, K. A. and Briddon, R. W. (2003). First report of Cotton leaf curl disease affecting chili peppers. *Plant Pathology,* 52: 809.

Maachupalli-Reddy, J., Kelley, B. D. and De Bernardez Clark, E. (1997). Effect of inclusion body contaminants on the oxidative renaturation of hen egg white lysozyme. *Biotechnology Program*, 13(2):144-150.

Mahmood, T. and Yang, P. (2012). Western Blot: Technique, Theory, and Trouble Shooting. *North American Journal of Medical Sciences,* 4(9): 429–434.

Mali, V. R. and Rajegore, S. B., (1979). Occurrence of cucumber mosaic virus on banana in India. *Plant Disease Reporter,* 63: 138-142.

Malik, A. H., Briddon, R. W. and Mansoor, S. (2011). Infectious clones of *Tomato leaf curl Palampur virus* with a defective DNA B and their pseudo-recombination with *Tomato leaf curl New Delhi virus*. *Virology Journal*, 8: 173.

Malik, A. H., Mansoor, S., Iram, S., Briddon, R.W. and Zafar, Y. (2006).A severe outbreak of Melon yellow mosaic disease caused by Zucchini yellow mosaic virus in the Punjab province of Pakistan. *Plant Pathology,* 55: 285.

Malygin, A., Parakhnevitch, N. and Karpova, G. (2005). Human ribosomal protein S13: cloning, expression, refolding, and structural stability.*Biochimica et Biophysica Acta,* 1747(1): 93–97.

Mandal, B., Mandal, S., Pun, K. B., and Varma, A. (2004). First report of the association of a nanovirus with foorkey disease of large cardamom in India. *Plant Disease,* 88: 428.

Mansoor, S., Amin, I. and Briddon, R. W. (2011). Geminiviral diseases of cotton. In Stress Physiology in Cotton. Edited by D. M. Oosterhuis. Cordova, Tennessee, USA: The Cotton Foundation.

Mansoor, S., Amin, I., Iram, S., Hussain, M., Zafar, Y., Malik, K. A. and Briddon, R. W. (2003a). Breakdown of resistance in cotton to Cotton leaf curl disease in Pakistan. *Plant Pathology,* 52: 784.

Mansoor, S., Amrao, L., Amin, I., Briddon, R. W., Malik, K. A. and Zafar, Y. (2006). First report of Cotton leaf curl disease in central and southern Sindh province in Pakistan. *Plant Disease,* 90: 826.

Mansoor, S., Bedford, I., Pinner, M. S., Stanley, J. and Markham, P. G. (1993). A whitefly-transmitted geminivirus associated with Cotton leaf curl disease in Pakistan. *Pakistan Journal of Botany,* 25: 105–107.

Mansoor, S., Briddon, R. W., Bull, S. E., Bedford, I. D., Bashir, A., Hussain, M., Saeed, M., Zafar, Y., Malik, K. A., Fauquet, C. M. and Markham, P. G. (2003b). Cotton leaf curl disease is associated with multiple monopartite begomoviruses supported by single DNA β. *Archives of Virology,* 148: 1969–1986.

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R., Stanley, J., and 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., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R. W., Stanley, J. and 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., Khan, S. K., Hussain, M., Zafar, Y., Pinner, M. S., Briddon, R. W., Stanley, J. and Markham, P. G., (2000a). Association of a Begomovirus and Nanovirus-like Molecule with Ageratum Yellow Vein Disease in Pakistan.*TheAmerican Phytopathological Society*. 84(1). 101.

Mansoor, S., Zafar, Y., and Briddon, R. W. (2006). Geminivirus disease complexes: the threat is spreading. *Trends in Plant Science*, 11: 209–212.

Manzoor, M. T., Ilyas, M., Shafiq, M., Haider, M. S., Shahid, A. A. and Briddon, R.W. (2014). A distinct strain of chickpea chlorotic dwarf virus (genus Mastrevirus, family *Geminiviridae*) identified in cotton plants affected by leaf curl disease. *Archives of Virology*, 159(5):1217–1222.

Mao, M. J., He, Z. F., Yu, H. and Li, H. P. (2008). Molecular characterization of Cotton leaf Curl Multan virus and its satellite DNA that infects Hibiscus rosa-sinensis. *Chinses Journal of Virology*, 24 (1): 64–68.

Mariano, A. C., Andrade, M. O., Santos, A. A., Carolino, S. M. B., Oliveira, M. L., Baracat-Pereira, M. C., Brommonshenkel, S. H. and Fontes, E. P. B. (2004). Identification of a novel receptor-like protein kinase that interacts with a geminivirus nuclear shuttle protein. *Virology*, 318: 24–31.

Markus, K., Markus, F., Sandra, B. and Pleiss, J. (2005). How to find soluble proteins: a comprehensive analysis of a / b hydrolases for recombinant expression in *E. coli*. *BMS Genomics,* 6: 49 – 58.

Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P. and Varsani, A. (2011). Recombination in eukaryotic single stranded DNA viruses. *Viruses,* 3: 1699–1738.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, 1(1): vev003. Doi: 10.1093/ve/vev003

Marwal, A., Prajapat, R. and Gaur, R. K. (2014). In Silico Recombination Analysis of DNA-A sequence from Begomovirus reported in India: This identified recombinant is the evolution from other viruses prevailing at different geographical region of Pakistan and China. *International Journal of Current Microbiology and Applied Sciences,* 3: 489–497.

Marwal, A., Sahu, A. K. and Gaur, R. K. (2013). Molecular Characterization of Begomoviruses and DNA Satellites Associated with a New Host Spanish Flag (*Lantana camara*) in India. *ISRN Virology*, Article ID915703

McGarry, R. C., Barron, Y. D., Carvalho, M. F., Hill, J. E., Gold, D., Cheung, E., Kraus, W. L. and Lazarowitz, S. G. (2003). A novel Arabidopsis acetyltransferase interacts with the geminivirus movement protein NSP. *Plant Cell,* 15: 1605–1618.

McLaughlin, P. D., McLaughlin, W. A., Maxwell, D. P. and Roye, M. E. (2008). Identification of Begomoviruses Infecting Crops and Weeds in Belize. *Plant Viruses,* 2(1):58-63.

Moffat, A. S. (1999). Geminiviruses emerge as serious crop threat. *Science,* 286: 1835.

Moffatt, B. A., Stevens, Y. Y., Allen, M. S., Snider, J. D., Periera, L. A., Todorova, M. I., Summers, P. S., Weretilnyk, E. A., Martin-Caffrey, L. and Wagner, C. (2002). Adenosine kinase deficiency is associated with developmental abnormalities and reduced transmethylation. *Plant Physiology,* 128: 812– 821.

Monci, F., Sánchez-Campos, S., Navas-Castillo J. and 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(2):317-326.

Monga, D., Chakrabarty, P. K., Kranthi, R. (2011). Cotton leaf Curl Disease in India-recent status and management strategies. Presented in 5th meeting of Asian Cotton Research and Development Network Held in Lahore in Feb 23–25.

Monsalve-Fonnegra, Z., Argüello-Astorga, G. and Rivera-Bustamante, R. (2001). Geminivirus replication and gene expression. In Plant Viruses as Molecular Pathogens, pp. 257-277. Edited by Khan, J. A. and J. Dijstra, J. New York: The Haworth Press.

Morales, F. J. and Anderson, P. K. (2001). The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. *Archives of Virology*, 146: 415–441.

Moriones, E. and Navas-Castillo, J. (2000). *Tomato yellow leaf curl virus*, and emerging virus complex causing epidemics worldwide. *Virus Research*, 71: 123–134.

Mosig, G. (1998). Recombination and recombination-dependent DNA replication in bacteriophages T4. *Annual Review of Genetics,* 32: 379–413.

Mosig, G., Gewin, J., Luder, A., Colowick, N. and Vo, D. (2001). Two recombination-dependent DNA replication pathways of bacteriophages T4, and their roles in mutagenesis and horizontal gene transfer. *PNAS,* 98: 8306–8311.

Mowat, W. P. and Dawson, S. (1987). Detection and identification of plant viruses by ELISA using crude sap extracts and unfractionated antisera. *Journal of Virology Methods,* 15(3):233-247.

Mubin, M., Amin, I., Amrao, L., Briddon, R. W. and Mansoor, S. (2010a). The hypersensitive response induced by the V2 protein of a monopartite begomovirus is countered by the C2 protein. *Molecular Plant Pathology,* 11(2): 245–254

Mubin, M., Briddon, R. W. and Mansoor, S (2009). Diverse and recombinant DNA betasatellites are associated with a begomovirus disease complex of *Digera arvensis*, a weed host.*Virus Research,* 142(1-2): 208–212.

Mubin, M., Briddon, R. W. and Mansoor, S. (2009). Complete nucleotide sequence of chili leaf curl virus and its associated satellites naturally infecting potato in Pakistan. *Archives of Virology,* 154(2): 365–368.

Mubin, M., Mansoor, S. and Briddon R. W. (2012). Mastrevirus sequences in a begomovirus-infected plant. *Virus Genes,* 44: 536–538.

Mubin, M., Shahid, M. S., Tahir, M. N., Briddon, R. W. and Mansoor, S. (2010). Characterization of begomovirus components from a weed suggests that begomoviruses may associate with multiple distinct DNA satellites.*Virus Genes,* 40(3): 452–457.

Muhire, B. M., Varsani, A. and Martin, D. P. (2014). SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLOS One*: e108277.

Mustujab, A., Briddon, R. W. and Tahir, M. (2015). First report of *Catharanthus yellow mosaic virus* infecting *duranta* repens in Pakistan. *Journal of Plant Pathology*, 97:1

Nadeem, A., Mehmood, T., Tahir, M., Khalid, S. and Z. Xiong, Z (1997). First Report of Papaya Leaf Curl Disease in Pakistan. *The American Pathological Society*, 81(11): 1333.

Naidu, R. A., Manohar, S. K., Reddy, D. V. R., Reddy, A. S. (1989). A plant rhabdovirus associated with peanut veinal chlorosis disease in India. *Plant Pathology*, 38: 623–626.

Navas-Castillo, J., Fiallo-Olivé, E., and Sánchez-Campos, S. (2011). Emerging virus diseases transmitted by whiteflies. *Annual Review of Phytopathology*, 49: 219–248.

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

Nawaz-ul-Rehman, M. S. and Fauquet, C. M. (2009). Evolution of geminiviruses and their satellites. *FEBS Letter.* 583: 1825–1832.

Nawaz-ul-Rehman, M. S., Briddon R. W. and 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.

Noris, E., Jupin, I., Accotto, G. P. and Gronenborn, B, (1996). DNA- binding activity of the C2 protein of Tomato yellow leaf curl geminivirus. *Virology*, 217: 607–612.

Noueiry, A. O., Lucas, W. J. and Gilbertson, R. L. (1994). Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell*, 76: 925–932.

Ogawa, T., Sharma, P. and Ikegami, M. (2008). First report of a strain of*Tobacco leaf curl Japan virus* associated with a satellite DNA in honeysuckle in Japan. *Plant Pathology,* 57 (2): 391.

Ogbe, F.O., Dixon, A.G.O., Hughes, J.D., Alabi, O.J. and Okechukwu, R. (2006). Status of cassava begomoviruses and their new natural hosts in Nigeria.http://agris.fao.org/agris-search/search.do?recordID=US201301073978

Ohnesorge, S. and Bejarano, E. R. (2009). Begomovirus coat protein interacts with a small heat-shock protein of its transmission vector (Bemisia tabaci). *Insect Molecular Biology,* 18(6): 693–703.

Ooi, K., Ohshita, S., Ishii, I., Yahara, T., 1997. Molecular phylogeny of geminivirus infecting wild plants in Japan. *Journal of Plant Research*, 110: 247–257.

Orozco, B. M. and Hanley-Bowdoin, L. (1996). A DNA structure is required for geminivirus replication origin function. *Journal of Virology,* 7: 148–158.

Orozco, B. M. and Hanley-Bowdoin, L. (1998). Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. *Journal of Biological Chemistry,* 273: 24448–24456.

Otim-Nape, G. W., Alicai, T. and Tresh, J. M. (2001), changes in the incidence and severity of cassava mosaic virus disease, varietal diversity and cassava production in Uganda. *Annual of Applied Biology,* 138: 313–327.

Padidam, M., Beachy, R. N. and Fauquet, C. M. (1996). The role of AV2 ("precoat") and coat protein in viral replication and movement in Tomato leaf curl geminivirus. *Virology*, 224: 390–404.

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

Palmer, I. and wingfield, P. T. (2004). Preparation and Extraction of Insoluble (Inclusion-Body) Protein from *Escherichia coli. Current Protocols in Protein Science*, CHAPTER: Unit–6.3. doi:10.1002/0471140864.ps0603s38.

Pant, V.,Gupta, D.,Choudhury, N. R.,Malathi, V. G.,Varma, A. and Mukherjee, S. K. (2001).Molecular characterization of the Rep protein of the blackgram isolate of Indian mungbean yellow mosaic virus. *Journal of General Virology,* 82(10): 2559-2567.

Paprotka, T., Metzler, V. and Jeske, H. (2010). The first DNA 1-like satellites in association with New World begomoviruses in natural infections. *Virology*, 404(2): 148–157.

Pascal, E., Goodlove, P. E., Wu, L. C., and Lazarowitz, S. G. (1993). Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. *Plant Cell*, 5: 795–807.

Pascal, E., Sanderfoot, A. A., Ward, B. M., Medville, R., Turgeon, R., and Lazarowitz, S. G. (1994). The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the cell nucleus. *Plant Cell*, 6: 995–1006.

Paximadis, M., Idris, A. M., Torres-Jerez, I., Villarreal, A., Rey, M. E. C. and Bron, J. K. (1999). Characterization of tobacco geminiviruses in the Old and New World. *Archives of Virology*, 144: 703–717.

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

Polson, J. E. and Anderson, P. (1997). The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. *Plant Disease,* 81: 1358–1369.

Pramesh, D., Mandal, B., Phaneendra, C. and Muniyappa, V. (2013). Host range and genetic diversity of Croton yellow vein mosaic virus, a weed-infecting monopartite begomovirus causing leaf curl disease in tomato. *Archives of Virology*, 158(3):531–542.

Pratap, D., Kashikar, A. R. and Mukherjee, S. K. (2011). Molecular characterization and infectivity of a Tomato leaf curl New Delhi virus variant associated with newly emerging yellow mosaic disease of eggplant in India. *Virology Journal*, 8:305

Preiss, W. and Jeske, H. (2003). Multitasking in replication is common among geminiviruses. *Journal of Virology*, 77: 2972–2980.

Priyadarshini, C. G. P., Ambika, M. V., Tippeswamy, R. and Savithri, H. S. (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.

Qazi, J., Mansoor, S., Amin, I., Awan, A. Y., Briddon, R. W. and Zafar, Y. (2006). First report of Mungbean yellow mosaic India virus on mothbean in Pakistan. *Plant Pathology,* 55: 818.

Raghavan, V., Malik, P. S., Choudhury, N. R. and Mukherjee, S. K. (2004). The DNA-A component of a plant geminivirus (Indian mungbean yellow mosaic virus) replicates in budding yeast cells. *Journal of Virology,* 78: 2405–2413.

Raikhy, G., Hallan, V., Kulshrestha, S., Sharma, M. L., Ram, R. and Zaidi, A. A. (2003). Molecular characterization of an Indian isolate of Carnation etched ring virus. *Acta Virology,* 43: 105–111.

Rajagopalan, P. A., Naik, A., Katturi, P., Kurulekar, M., KankanalluI, R. S., Anandalakshmi, R. (2012). Dominance of resistance-breaking Cotton leaf curl Burewala virus (CLCuBuV) in northwestern India. *Archives of Virology*, 157:855–868.

Riaz, H., M. Ashfaq, M., Mukhtar, T. and Riaz. T. (2016). First report of *Euphorbia yellow leaf curl virus* infecting*Hibiscus syriacus. Journal of Plant Pathology*, Doi: 10.4454/JPP.

Ribeiro, S. G., Ambrozevicius, B., Avil, A. C., Bezerra, I. C., Calegario, R. F., Fernandes, J. H., Lima, M. F., de Mello, R. N., Rocha, H, and Zerbini, F. M. (2003). Distribution and genetic diversity of tomato infecting begomoviruses in Brazil. *Archives of Virology,* 148: 281–295.

Rice, P., Longden, I. and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite (2000). *Trends in Genetics*, 16: 276–277.

Rigden, J. E., Dry, I. B., Mullineaux, P. M. and Rezaian, M. A. (1993). Mutagenesis of the virion-sense open reading frames of Tomato leaf curl virus. *Virology,* 193: 1001–1005.

Rigden, J. E., Krake, L. R., Rezaian, M. A. and Dry, I. B. (1994). ORF C4 of Tomato leaf curl virus is a determinant of symptom severity. *Virology,* 204: 847–850.

Rishishwar, R., Mazumdar, B. and Dasgupta, I. (2015). Diverse and recombinant begomoviruses and various satellites are associated with Bhendi yellow vein mosaic disease of okra in India. *Journal of Plant Biochemistry and Biotechnology,* 24: 470–475.

Rob W. Briddon, R. W., Bull,S. E., Amin, I., Mansoor, S., Bedford, I. D., Rishi, N., Siwatch, S. S., Zafar, Y., Abdel-Salam, A. M.and Markham P. G. (2004). Diversity of DNA 1: a satellite-like molecule associated with monopartitebegomovirus – DNA β complexes. *Virology*, 324: 462 – 474.

Rocha, C. S., Xavier, C. A. D., Lima, A. T. M., Silva, F. N. and Zerbini, F. M. (2012). Molecular analysis of the DNA-A of the begomovirus Tomato Mottle leaf curl virus (ToMoLCV). *Virus Reviews and Research*, DOI: 10.17525/vrrjournal.v16i1-2.67.

Rojas, M. R., Hagen, C., Lucas, W. J. and Gilbertson, R. L. (2005). Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. *Annual Review of Phytopathology*, 43: 361–394.

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

Rojas, M. R., Noueiry, A. M., Lucas, W. J. and Gilbertson, R. L. (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. and Desbiez, C. (2010). Association of an atypical alphasatellite with a bipartite New World begomovirus. *Archives of Virology*, 155(11): 1843–1847.

Roossinck, M. J., Sleat, D. and Palukaitis, P. (1992). Satellite RNAs of plant viruses: structures and biological effects. *Microbiology Revision,* 56: 265–279.

Rosano and Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5:172 Doi: 10.3389/fmicb.2014.00172

Roth, B. M., Pruss, G. J. and Vance, V. B. (2004). Plant viral suppressors of RNA silencing. *Virus Research,*102: 97–108.

Rothenstein, D., Haible, D., Dasgupta, I., Dutt, N., Patil, B. L and Jeske, H. (2006). Biodiversity and recombination of cassava-infecting begomoviruses from southern India. *Archives of Virology,* 151: 55–69.

Roumagnac, P., Granier, M., Bernardo, P., Deshoux, M., Ferdinand, R., Galzi, S., Fernandez, E., Julian, C., Abt, I., Filloux, D., Mesléard, F., Varsani, A., Blanc, S., Martin, D. P. and Peterschmitt, M. (2015). Alfalfa Leaf Curl Virus: an Aphid-Transmitted Geminivirus. *Journal of Virology,* 89(18): 9683–9688.

Roye, M. E., Mclaughlin, W. A. and Maxwell, D. P. (2000). The evolution of new virus genes: Interspecies recombination among two geminiviruses from Jamaica. *Jamaican Journal of Scientific Technology*. 11: 42-46.

Rudolph, R. and Lilie, H. (1996). In vitro refolding of inclusion body proteins.*Faseb Journal*, 10:49–56.

Rybicki, E. P. (1994). A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Archives of Virology*, 139 (1-2): 49–77.

Rybicki, E. P. and Pietersen, G. (1999). Plant virus disease problems in the developing world. *Advances in Virus Research* 53: 127–175.

Saeed, M., Behjatnia, S. A. A., Mansoor, S., Zafar, Y., Hasnain S., Rezaian, M. A. (2004). A single complementary-sense transcript of a geminiviral DNA β satellite is determinant of pathogenicity. *Molecular Plant Microbe Interaction*, 18: 7–14.

Saeed, S., Y. Zafar, Y., Randles, J. W. and Rezaian, M. A. (2007). A monopartite begomovirus-associated DNA satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection. *Journal of General Virology*, 88(10): 2881–2889.

Saeed. M. (2008) Limitations observed in the use of agroinoculation for geminivirus research. *Virus Genes*, 37: 434–435.

Samad, A., Gupta, M. K., Shasany, A. K., Ajayakumar, P. V. and Alam, M. (2009). Begomovirus related to *Tomato leaf curl Pakistan virus* newly reported in *Mentha* crops in India. *Plant Pathology*, 58(2): 404.

Sambrook, J., and Russel, D. W. (2001). Molecular cloning: A Laboratory Manual, 3rd edn. N.Y: Cold Spring Harbor Laboratory.

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

Sanderfoot, A. A., Ingham, D. J. and Lazarowitz, S. G. (1996). A viral movement protein as a nuclear shuttle. *Plant Physiology*, 110: 23–23.

Sanz, A. I, Fraile, A, García-Arenal, F., Zhou, X., Robinson, D. J., Khalid, S., Butt, T. and Harrison, B. D.(2000). Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. *Journal of General Virology*, 81: 1839–1849.

Sanz, A. I., Fraile, A., Gallego, J. M., Malpica, J. M. and Garcia-Arenal, F. (1999). Genetic variability of natural populations of Cotton leaf curl geminivirus, a single-stranded DNA virus.*Journal of Molecular Evolution,* 49(5): 672–81.

Sattar, M. N., Kvarnheden, A., Saeed, M., Briddon, R. W. (2013). Cotton leaf curl disease—an emerging threat to cotton production worldwide. *Journal of General Virology,* 94:695–710.

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

Saunders, K., Bedford, I. D. and Stanley, J. (2001). Pathogenicity of a natural recombinant associated with Ageratum yellow vein disease: implications for geminivirus evolution and disease etiology. *Virology,* 282: 38–47.

Saunders, K., Bedford, I. D., Briddon, R. W., Markham, P. G., Wong, S. M. and Stanley, J. (2000). A unique virus complex causes Ageratum yellow vein disease. *Proceedings of the National Academy of Sciences USA*, 97(12): 6890–6895.

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

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

Saunders, K., Lucy, A, and Stanley, J. (1992). RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. *Nucleic Acids Research,* 20: 6311–6315.

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

Scholthof, K. B. G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C. and Foster, G.D. (2011). Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology,* 12(9): 938–954.

Seal, S. E., van den Bosch, F. and Jeger, M. J. (2006). Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. *Critical Reviews in Plant Sciences,* 25: 23–46.

Selth, L. A., Dogra, S. C., Rasheed, M. S., Healy, H., Randles, J. W. and Rezaian, M. A. 2005. A NAC domain protein interacts with Tomato leaf curl virus replication accessory protein and enhances viral replication. *Plant Cell*, 17: 311–325.

Selth, L. A., Randles, J. W. and Rezaian, M. A. 2004. Host responses to transient expression of individual genes encoded by *Tomato leaf curl virus*. *Molecular Plant-Microbe Interaction,* 17: 27–33.

Sereikaite, J., Statkute, A., Morkunas, M., Radzevicius, K., Borromeo, V., Secchi, C. and Bumelis. V. A. (2007). Production of recombinant milk growth hormone in *E. coli*. *Applied Microbiology and Biotechnology*, 74(2): 316–23.

Settlage, S. B., Miller, A. B., Gruissem, W. and 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, S. B., Miller, B. and Hanley-Bowdoin, L. (1996). Interactions between geminivirus replication proteins. *Journal of Virology,* 70: 6790–6795.

Settlage, S. B., See, R. G. and Hanley-Bowdoin, L. (2005). Geminivirus C3 protein: replication enhancement and protein interactions. *Journal of Virology,* 79: 9885–9895.

Shafiq, M., Asad, S., Zafar, Y., Briddon, R. W. and Mansoor, S. (2010). Pepper leaf curl Lahore virus requires the DNA B component of *Tomato leaf curl New Delhi virus* to cause leaf curl symptoms. *Virology Journal,* 7: 367.

Shahid, M. S., Ikegami, M., Waheed, A., Briddon, R. W. and Natsuaki, K. T. (2014). Association of an alphasatellite with Tomato yellow leaf curl virus and Ageratum yellow vein virus in Japan Is Suggestive of a Recent Introduction.*Viruses,* 6: 189–200.

Sharma, P. and Ikegami, M. (2008). The begomoviruses Honeysuckle yellow vein mosaic virus and Tobacco leaf curl Japan virus with DNA betasatellites cause yellow dwarf disease of tomato. *Virus Research*, 137(2): 235–244.

Sharma, P. and Ikegami, M. (2009). Characterization of signals that dictate nuclear/nucleolar and cytoplasmic shuttling of the capsid protein of Tomato leaf curl Java virus associated with DNAβ satellite. *Virus Research,* 144: 145–153.

Shih, S. L., Tsai, W. S., Green, S. K., Khalid, S., Ahmad, I., Rezaian, M. A. and Smith, J. (2003). Molecular Characterization of Tomato and Chili Leaf Curl Begomoviruses from Pakistan. *Disease Note*, 87(2): 200.

Shivaprasad, P. V., Akergenov, T., Trinks, D., Raheswaran, R., Veluthambi, K., Hohn, T. and Pooggin, M. (2005). Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. *Journal of Virology,* 79: 8149–8163.

Siddique, S. A., Khalid, S., Ilyas, M. B. and Aulakh, M. A. (1999). Evaluation op peppers lines against Tomato leaf curl virus. *Pakistan Journal of Botanical Sciences*, 2: 1494–1496.

Singh, A.K., Chattopadhyay, B. and Chakraborty, S. (2012). Biology and interactions of two distinct monopartite begomoviruses and betasatellites associated with radish leaf curl disease in India. *Virology Journal*, 9: 43.

Singh, A. K., Chattopadhyay, B., Pandey, P. K., Singh, A. K. and Chattopadhyay, S. (2007). A new begomovirus species causing leaf curl disease of Radish in India. *Disease Notes*, 91(8): 1053.

Sinha, V., Kumar, A.and Bhatnagar, D. J. A. K. (2013). Association of Cotton leaf curl Multan virus and its satellite molecules with leaf curl disease of papaya in India. *New Disease Reports*, 27: 9.

Sivalingam, P. N., Malathi, V. G. and Varma, A. (2010). Molecular diversity of the DNA-β satellites associated with Tomato leaf curl disease in India. *Archives of Virology*, 155: 757–764.

Snehi, S. K., Raj, S. K., Prasad, V. and Singh V. (2016). Molecular Detection and Identification of Begomovirus Isolates Associated with Mosaic Disease of Ornamental Jatropha species from India. *International Journal of Bacterio-Virology*, 1: 004.

Sohrab, S. S., Kamal, M. A., Abdul Ilah., Azamal Husen., Bhattacharya, P. S. and Rana, D. (2014). Development of Cotton leaf curl virus resistanttransgenic cotton using antisense ßC1 gene. *Saudi Journal of Biological Sciences*, 23: 358 –362.

Srivastava, A., Kumar, S. and Raj, S. K. (2014). Association of *Pedilanthus leaf curl virus* with yellow mottling and leaf curl symptoms in two jasmine species grown in India. *Journal of General Plant Pathology*, 80: 370–373.

Srivastava, A., Kumar, S., Jaidi, M. and Raj S. K. (2016). Association of Cotton leaf curl Multan virus and its associated betasatellite with leaf curl disease of Hibiscus rosa-sinensis in India. *New Disease Reports*, 33: 4.

Stanley, J. (1995). Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. *Virology*, 206: 707–712.

Stanley, J., Bisaro, D. M., Briddon, R. W., Brown, J. K., Fauquet, C. M., Harrison, B. D., Rybicki, E. P. and Stenger, D. C. (2005). *Geminiviridae* In: Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., Ball, L. A. (Eds), VIII Report of the International Committee on Taxonomy of Viruses. Virus Taxonomy. Elsevier/ Academic Press, London, pp. 1163–1169.

Stenger, D. C. (1994a). Complete nucleotide sequence of the hypervirulent CFH strain of Beet curly top virus. *Molecular Plant-Microbe Interaction*, 7: 154–157.

Stenger, D. C. (1994b). Strain-specific mobilization and amplification of a transgenic defective-interfering DNA of the geminivirus Beet curly top virus. *Virology*, 203: 397–402.

Stenger, D. C., Revington, G. N., Stevenson, M. C. and Bisaro, D. M. (1991). Replication release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. *Proceedings of National Academy of Sciences USA*, 88: 8029–8033.

Stonor, J., Hart, P., Gunther, m., Debarro, P. and Rezaian, M. A. (2003). Tomato leaf curl geminivirus in Australia: occurrence, detection, sequence diversity and host range. *Plant Pathology*, 52(3): 379-388.

Sudarshana, M. R., Wang, H. L., Lucas, W. J. and Gilbertson, R. L. (1998). Dynamics of bean dwarf mosaic geminivirus cell-to-cell and long-distance movement in Phaseolus vulgaris revealed, using the green fluorescent protein. *Molecular Plant-Microbe Interactions*, 11: 277–291.

Sunter, G. and Bisaro, D. M. (1989). Transcription map of the B genome component of Tomato golden mosaic virus and comparison with A component transcripts. *Virology*, 173: 647–655.

Sunter, G., Hartitz, M. D., Hormuzdi, S. G., Brough, C. L. and Bisaro, D. M. (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, D. C. and Bisaro, D. M. (1994). Heterologous complementation by geminivirus AL2 and AL3 genes. *Virology*, 203: 203–210.

Swartz, J.R. (2001). Advances in *Escherichia coli* production of therapeutic proteins. *Current Opinions in Biotechnology*, 12: 195–201.

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

Tahir, M. N., Amin, I., Briddon, R. W. and Mansoor, S. (2011). The merging of two dynasties–identification of an African Cotton leaf curl disease-associated begomovirus with cotton in Pakistan. *PLoS ONE*, 6: e20366.

Tahir, M., Amin, I., Haide, M. S., Mansoor, S. and Briddon, R. W. (2015). Ageratum enation virus—A Begomovirus of Weeds with the Potential to Infect Crops. *Viruses*, 7: 647–665

Tahir, M., Haider M. S.and Briddon, R. W. (2010). Chili leaf curl betasatellite is associated with a distinct recombinant begomovirus, Pepper leaf curl Lahore virus, in Capsicum in Pakistan. *Virus Research*, 149:109–114.

Tahir, M., Haider, M. S. and Briddon, R. W. (2010). Complete nucleotide sequence of Bitter gourd yellow vein virus and its associated betasatellite infecting *Momordica charantia*. *Archives of Virology*, 155: 1901–1905.

Tahir, M., Haider, M. S. and Briddon, R. W. (2010). First report of Squash leaf curl China virus in Pakistan. *Australasian Plant Disease Note,* 5(1): 21–24.

Tahir, M., Haider, M. S., Shah, A.H., Rashid, N. and Saleem, F. (2006). First report of bipartite begomovirus associated with leaf curl disease of *Duranta repens* in Pakistan. *Journal of Plant Pathology*, 88: 337.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MeEGA5: molecular evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution,* 28: 2731–39.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30(12): 2725–2529.

Thresh, J.M. and Cooter, R. J. (2005). Strategies for controlling cassava mosaic virus disease in Africa. *Plant Pathology*, 54: 587-614.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Harimalala,M. A., De Bruyn, A., Villemot, J., Traoré, V. S. E., Konaté,G., Traoré, A. S., Barro, N., Reynaud, B., Traoré, O. and Lett, J. M. (2012). Evolution of *African cassava mosaic virus*by recombination between bipartite and monopartite begomoviruses. *Virology Journal*, 9: 67.

Timmermans, M. C. P., Das, O. P. and Messing, J. (1994). Geminiviruses and their uses as extrachromosomal replicons. *Annual Review of Plant Physiology and Plant Molecular Biology,* 45: 79–112.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences, USA,* 76(9): 4350–4354.

Townsend, R., Watts, J. and Stanley, J. (1986). Synthesis of viral DNA forms in *Nicotiana plumbaginifolia* protoplasts inoculated with cassava latent virus (CLV); evidence for the independent replication of one component of the CLV genome. *Nucleic Acids Research* 14: 1253–1265.

Tsumoto, K., Ejima, D., Kumagai, I. and Arakawac, T. (2003). Practical considerations in refolding proteins frominclusion bodies. *Protein Expression and Purification,* 28:1–8.

Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S.and Arakawa. T. (2004). Role of Arginine in ProteinRefolding, Solubilization, and Purification. *Biotechnology Programme,* 20: 1301–1308.

Ullah, R., Akhtar, K.P., Hassan, I., Saeed, M., Sarwar, N. and Mansoor, S. (2014). Evidence of *Cotton leaf curl Burewala virus* variant and its associate betasatellite causing yellow mosaic of Eggplant (*Solanum melongena*) in Pakistan. *Journal of. Phytopathology*, 163(3): 233–237.

Vadivukarasi, T., Girish, K. R. and Usha, R. (2007). Sequence and recombination analysis of the geminivirus replication initiator protein. *Journal of Bioscience,* 32(1):17–29.

Vallejo, L. F and Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microbial Cell Factories,* 3: 11.

van Wezel, R., Dong, X., Blake, P., Stanley, J. and Hong, Y. (2002a). Differential roles of geminivirus Rep and AC4 (C4) in the induction of necrosis in *Nicotiana benthamiana*. *Molecular Plant Pathology,* 3: 461–471.

Vanitharani, R., Chellappan, P., Pita, J. S. and Fauquet, C. M. (2004). Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and

suppression of post-transcriptional gene silencing. *Journal of Virology*, 78: 9487–9498.

Varma, A. and Malathi, V. G. (2003). Emerging geminivirus problems: A serious threat to crop production. *Annuls of Applied Biology*, 142: 145–164.

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

Varsani, A., Roumagnac, P., Fuchs, M., Navas-Castillo, J., Moriones, E., dris, A., Briddon, Bustamante, R. R., Zerbini, F. M. and Martin, D. P. (2017). Capulavirus and Grablovirus: two new genera in the family Geminiviridae. *Archives of Virology*, doi: 10.1007/s00705-017-3268-6.

Ventura, S. and Villaverde, A. (2006). Protein quality in bacterial inclusionbodies. *Trends in Biotechnology*, 4: 179 – 185.

Verma, R., Prakash, S., Tomer, S. P. S. (2004). First report of Zucchini yellow mosaic virus in cucumber (*Cucumis sativus*) in India. *Plant Disease*, 88: 906.

Voinnet, O. (2005a). Induction and suppression of RNA silencing: insights from viral infections. *Nature Genetics*, 6: 206–221.

Voinnet, O. (2005b). Induction and suppression of RNA silencing: insights from viral infections. *Nature Review of Genetics*, 6: 206-220.

Voinnet, O. (2005c). Non-cell autonomous RNA silencing. FEBS Letters 579, 5858.

Von Arnim, A. and Stanley, J. (1992). Determinants of tomato golden mosaic virus symptom development located on DNA B. *Virology*, 186: 286–293.

von Regenmortel, M. H. V., Bishop, D. H. L., Fauquet, C. M., Mayo, M. A., Maniloff, J. and Calisher, C.H. (1997). Guidelines for the demarcation of virus species. *Archive of virology,* 142: 1505–1518.

Wang, B. L., Xu, Y., Wu, C. Q., Xu Y. M. and Wang, H. H. (2005). Cloning, expression, and refolding of a secretory protein ESAT-6 of Mycobacterium tuberculosis. *Protein Expression and Purification,* 39: 184–188.

Wang, B., Li, F., Huang, C., Yang, X., Qian, Y., Xie, Y. and Zhou, X. (2013). V2 of Tomato yellow leaf curl virus can suppress methylation-mediated transcriptional gene silencing in plant. *Journal of General Virology*.doi:10.1099/vir.0.055798-0

Wang, G., Sun, Y. W., Xu, R. R., Qu, J., Tee, C. S., Jiang, X. and Ye, X. (2014). DNA-A of a highly pathogenic Indian cassava mosaic virus isolatedfrom *Jatropha curcas* causes symptoms in *Nicotiana benthamiana*. *Virus Genes*, 48:402–405.

Wang, H., Buckley, K. J., Yang, X., Buchmann, R. C. and Bisaro, D. M. (2005a). Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. *Journal of Virology,* 79: 7410–7418.

Wang, H., Hao, L., Shung, C. Y., Sunter, G. and Bisaro, D. M. (2003a). Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. *Plant Cell,* 15: 3020–3032.

Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W. and Chen, X. Y. (2005b). Control of root cap formation by MicroRNA-targeted auxin response factors in Arabidopsis. *Plant Cell,* 17: 2204–2216.

Wang, X., Xie, Y. and Zhou, X. (2004). Molecular characterization of two distinct Begomoviruses from papaya in China. *Virus Genes*, 29(3): 303–309.

Wege, C., Gotthardt, R. D., Frischmuth, T. and Jeske, H. (2000). Fulfilling Koch"s postulates for *Abutilon mosaic virus*. *Archives of Virology,* 145: 2217–2225.

Willment, J. A., Martin, D. P., Palmer, K. E., Schnippenkoetter, W. H., Shepherd, D.N. and Rybicki, E. P. (2007). Identification of long intergenic region sequences involved in maize streak virus replication. *Journal of General Virology*, 88: 1831–1841.

Xie, Y., Zhao, L., Jiao, X., Jiang, T., Gong, H., Wang, B., Briddon, R.W. and Zhou, X. (2013). A recombinant begomovirus resulting fromexchange of the C4 gene.*Journal of General Virology*, 94: 1896–1907.

Xie, Y., Zhou, X., Zhang, Z. andQi, Y. (2002). *Tobacco curly shoot virus* isolated in Yunnan is a distinct species of *Begomovirus.Chinese Science Bulletin,* 47(3): 199–201.

Yadava, P., Suyal, G. and Mukherjee, S. K. (2010). Begomovirus DNA replication and pathogenicity. *Current Science*. 98: 360–368.

Yang, X., Baliji, S., Buchmann, R. C., Wang, H., Lindbo, J. A., Sunter, G. and Bisaro, D. M. (2007). Functional modulation of the Geminivirus AL2 transcription Factor and silencing suppressor by self-interaction. *Journal of Virology*, 81: 11972–11981.

Zaffalon, V., Mukherjee, S., Reddy, V., Thompson, J. and Tepfer, M. (2012). A survey of geminiviruses and associated satellite DNAs in the cotton-growing areas of northwestern India. *Archives of Virology,* 157:483–495.

Zaidi, S. S., Shafiq, M., Amin, I., Scheffler, B. E., Scheffler, J. A., Briddon, R. W. and Mansoor, S. (2016). Frequent occurrence of *Tomato leaf curl New Delhi virus* in Cotton leaf curl disease affected cotton in Pakistan. *PLoS ONE,* 11(5): e0155520. doi:10.1371/journal.

Zaitlin, M. and Palukaitis, P. (2000). Advances in understanding plant viruses and virus disease. *Annual Review of Phytopathology*, 38: 117-143.

Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, P., Varsani, A. and ICTV Report Consortium.(2017).ICTV Virus Taxonomy Profile: *Geminiviridae. Journal of General Virology*, 98:131–133.

Zhang, A., Jin, M., Liu, F., Guo, X., Hu, Q., Han, L., Tan, Y. and Chen, H. (2006). Development and evaluation of a DAS-ELISA for rapid detection of avian influenza viruses. *Avian Disease*.50(3):325-30.

Zhang, W., Olson, N. H., Baker, T. S., Faulkner, L., Agbandje-McKeen, M., Boulton, M. I., Davies, J. W. and McKenna, R. (2001). Structure of the Maize streak virus geminate particle. *Virology*, 279: 471–477.

Zhou, X. (2013). Advances in understanding begomovirus satellites. *Annual Review of Phytopathology*, 51: 357–381.

Zhou, X., Liu, Y., Calvert, L., Munoz, C. and Otim-Nape, G. W. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology*, 78: 2101–2111.

Zhou, X., Liu, Y., Robinson, D. J. and Harrison, B. D. (1998). Four DNA-A variants among Pakistani isolates of Cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. *Journal of General Virology*, 79: 915–923.

Zhou, X., Xie, Y., Peng, Y. and Zhang, Z. (2003b). Malvastrum yellow vein virus, a new Begomovirus species associated with satellite DNA molecule. *Chinese Science Bulletin*, 48: 2205–2209.

Zhou, X., Xie, Y., Tao, X., Zhang, Z., Li, Z. and Fauquet, C. M. (2003a). Characterization of DNA β associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. *Journal of General Virology*, 84: 237–247.

Zia-ur-Rehman, M., Herrmann, H. W., Hameed, U., Haider, M. S. and Brown, J. K. (2013). First detection of *Cotton leaf curl Burewala virus* and Cognate Cotton leaf curl Multan betasatellite and Gossypium darwinii symptomless alphasatellite in Symptomatic *Luffa cylindrical* in Pakistan. *Disease Notes,* 97(8).pp. 1122.

Zulfiqar, A., Ren, H., Zhang, J., Li, J., Cui, X. and Xie. Y. (2011). Molecular characterization of two isolates of *Sweet Potato leaf curl virus* infecting *Ipomoea indica* in China. *African Journal of Biotechnology,* 10: 45.

Zulfiqar, A., Zhang, J., Cui, X., Qian, Y., Zhou, X. and Xie, Y. (2012). A new begomovirus associated with alpha- and betasatellite molecules isolated from *Vernonia cinerea* in China. *Archives of Virology*, 157(1): 189–191.

APPENDICES

Appendix I (Sequences derived with Primer sets CLCV1/CLCV2 in diagnostic PCR)

>C50

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGCGATGGGTTCCCCTGTGCGTGAAT

>C28

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGGGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAAAACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTAAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTTATGGGTTCCCCTGTGCGTGAAT

>C32

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT

ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C49

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGCCCATGTAAGGTTCA
GTCCTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTCATGTGTGTCAGTGATGTTACTCGTGG
AACTGGGCTGACCCATAGAGTGGGTAAGAGATTTTGTGTCAAGTCTGTGTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACCAAGAATCACACGAATAGTGTGATGTTTTTTTTAGTTAGAGACCGTAG
ACCAGTTGATAAACCTCAGGATTTTGGAGAGGTGTTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAGGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCGCTTGTGAAGAAGTTTATTAGGGTTAATAATTATGTTGTTTACAATCA
ACAGGAAGCAGGAAAATACGAGAATCATACGGAAAATGCGTTAATGCTTTATATGGCTTGTACCCACGC
TAGCAACCCTGTTTATGCTACGTTGAAGATTAGGATATATTTTTATGACTCTGTAACGAATTGATATTA
ATAAAGTTTGAATTTTATTTCTGAATATTGATCTACATACATAGTTTGTTGGATTACATTGTACAATAC
ATGTTCTACAGCTTTAATAACTAAATTAATTGAAATTACACCGAGATTGTTCAGATATTTGAGGACTTG
GGTTTTGAATACCCTTAAGAAAAGACCAGTCTGAGGCTGTAAGGTCGTCCAGATTCGGAAGGTTAGAAA
ACACTTGTGCAGTCCCAGAGCTTTCCGAGTGTTGTAGTTGAACTGGATCTTGATCGTGAGTATGTCCAT
ATTCGTCGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAAGGGGGATTTGGAACCTCCCAGAT
ATATGCGCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C58

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C64

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTGCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT

CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C47

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTAAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATCACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTTAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTTATGGGTTCCCCTGTGCGTGAAT

>C62

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGCCCATGTAAGGTTCA
GTCCTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTCATGTGTGTCAGTGATGTTACTCGTGG
AACTGGGCTGACCCATAGAGTGGGTAAGAGATTTTGTGTCAAGTCTGTGTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACCAAGAATCACACGAATAGTGTGATGTTTTTTTTAGTTAGAGACCGTAG
ACCAGTTGATAAACCTCAGGATTTTGGAGAGGTGTTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAGGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCTTAAGTTATTTAAATATTTTAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAAAGGGGATTTGGAACCTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C52

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGCCCATGTAAGGTTCA
GTCCTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTCATGTGTGTCAGTGATGTTACTCGTGG
AACTGGGCTGACCCATAGAGTGGGTAAGAGATTTTGTGTCAAGTCTGTGTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACCAAGAATCACACGAATAGTGTGATGTTTTTCTTAGTTAGAGACCGTAG
ACCAGTTGATAAACCTCAGGATTTTGGAGAGGTGTTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAGGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATGTCCATATTCGT

CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAAAGGGGATTTGGAACCTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C40

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGGCGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C5

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTTATGGGTTCCCCTGTGCGTGAAT

>C15

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCATGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C34

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTTAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

>C12

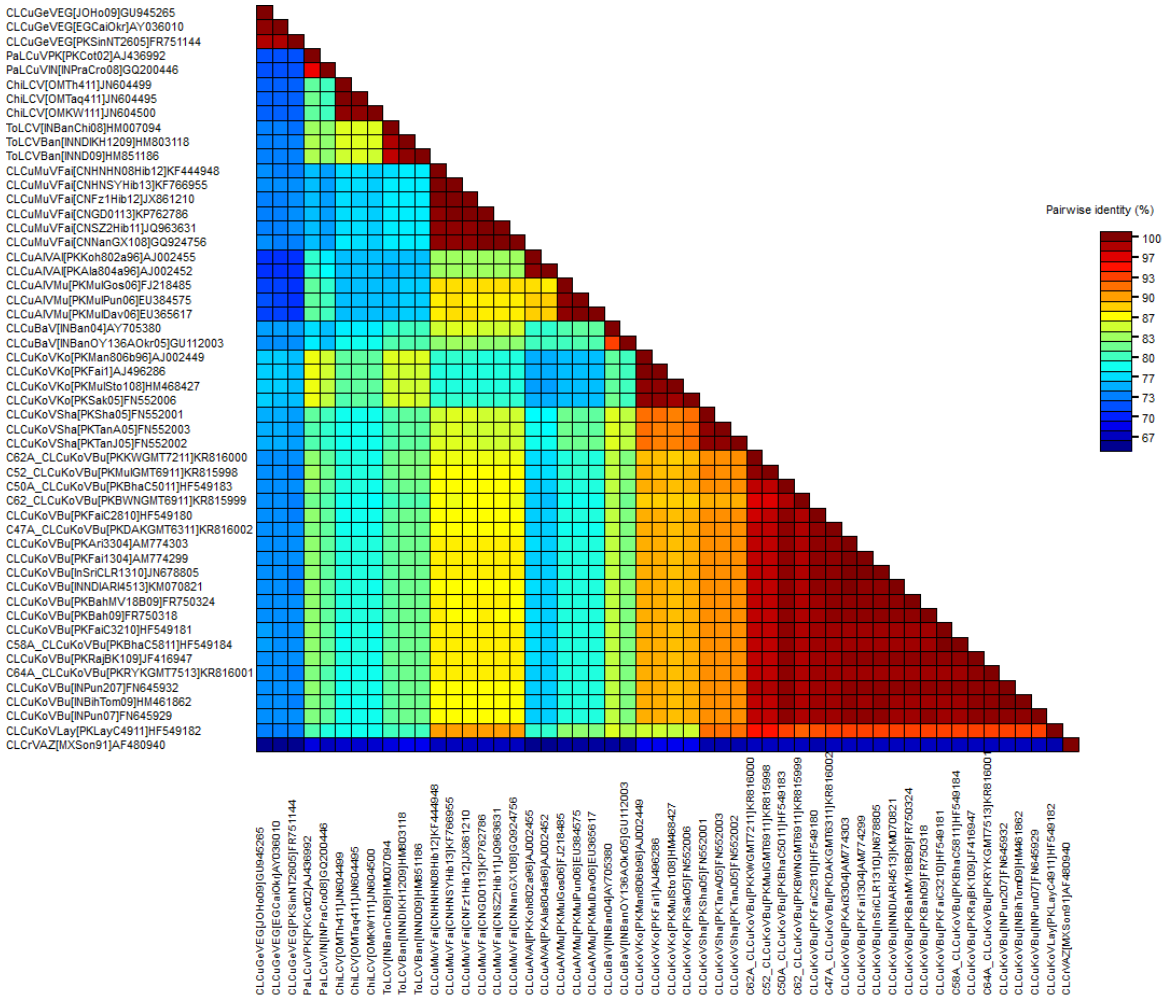
CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTTAAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTTATGGGTTCCCCTGTGCGTGAAT

>C7

CCGTGCTGCTGCCCCCATTGTCCGCGTCACCAAAGCAAAAGCATGGGCGAACAGGCCCATGAACAGAAA
GCCCAGGATGTACAGGATGTACAGAAGTCCAGATGTTCCTAGAGGATGTGAAGGTCCATGTAAGGTTCA
GTCGTTTGAGTCCAGACATGATATTCAGCATATAGGTAAAGTAATGTGTGTTAGTGATGTTACTCGTGG
TACTGGGCTGACCCATAGAGTTGGTAAGAGATTTTGTGTTAAGTCTGTTTATGTGTTGGGTAAGATCTG
GATGGATGAGAACATTAAGACGAAGAATCACACGAATAGTGTGATGTTTTTCTTGGTTAGAGATCGTAG
ACCTGTTGATAAACCTCAAGATTTTGGAGAGGTATTTAATATGTTTGATAATGAGCCCAGTACGGCGAC
TGTGAAGAATGTTCATCGTGATAGGTATCAAGTTCTGCGCAAATGGTATGCAACTGTCACCGGTGGACA
ATACGCTTCAAAGGAACAAGCTCTCGTGAAGAAATTTATTAGAGTTAATAATTATGTTGTGTATAACCA
GCAGGAAGCTGGCAAGTATGAGAATCATTCTGAGAATGCTTTAATGTTGTATATGGCGTGTACTCACGC
CTCTAACCCAGTGTATGCTACCTTGAAGATACGGATCTACTTCTATGATTCCGTGACAAATTAATAAAT
ATTGAATTTTATTGAAGATGATTGGTCTACAAATACAACATGTTGTAATACATTCCATAATACATGATC
AACTGCTCTAACTACATTATTAATACTGACAATTCCTAAGTTATTTAAATATTTTAGCACTTGAGTCCT
AAAGACCCTTAAGAAACGACCAGTCGGAGGCTGTGAGGTCATCCAGATTCGGAAAGCTATGAAACATTT
GTGTATCCCCAACGCTTTCCTCAGGTTGTGATTGAACTGGATCCTGATCGTGAGTATATCCATATTCGT
CGTGAATGGACGGTTGACGTGGCTGATGATCTTGAAATAGAGGGGATTTGGAACTTCCCAGATATATGC
GCCATTCCCTGCTTGAGCTGCAGTGATGGGTTCCCCTGTGCGTGAAT

Appendix II (Partial Genome based Pairwise Sequence Comparisons by STD)

Appendix III (Full Genome based Pairwise Sequence Comparisons by STD)



Appendix IV (Solutions for Plasmid Isolation)

Solution 1 (500 ml, pH 8.0, Filter, Store at 4 °C)

Glucose (50 mM)	4.5 g
Tris-HCl (1M)	12.5 ml
EDTA (0.5 M)	10 ml
Sterilze water	400 ml

Solution 2 (0.2 N NaOH; 0.5% SDS) (100 ml, Store at Room Temperature)

NaOH (1N)	20ml
10% SDS	10 ml
Sterile water	70 ml

Solution 3 (500 ml, pH 4.8-5.2, filter sterilized, Store at Room Temperature)

Glacial Acetic Acid	57.5 ml
Potassium Acetate (5 M Stock)	300 ml
Sterile distilled water	142.5 ml.

Wash Buffer (100 ml, Store at -20 °C)

Trizma base (100 mM)	1.21 g
Ammonium Acetate (8 M stock)	125 µl
Ethanol (95 %)	79 ml

CTAB Isolation Buffer (1L, pH 8.0, Autoclave)

CTAB (2%)	20 g
Trizma base (100 mM)	12.1g
EDTA (20 mM)	7.44 g
NaCl (1.4 M)	81.82g
β-mercaptoethanol	2% (add after autoclave or before use)

TAE Electrophoresis Buffer (50X) (1L, Autoclave, Store at Room Temperature)

Trisma-base	242 g
$Na_2EDTA.1H2O$	37.2 g
Glacial Acetic Acid	57.1 ml
Distilled water	930 ml

8 M Ammonium Acetate Solution (1L, Filter, Store at Room Temperature)

Ammonium acetate	616.64 g

YEP medium (1L, pH 7.5 with NaOH, Autoclave)

Meat extracts	3g
Bacto tryptone	5g
Bacto-yeast extract	1g
Sucrose	5g
$MgSO_4$ (1M)	2ml

SOB medium (1L, pH 7.0 with NaOH, Autoclave)

Bacto tryptone	20g
Bacto-yeast extract	5g
NaCl	0.58g
KCl	0.2g

2 M Mg^{++} stock solution (100ml, Autoclave)

$MgCl_2.6H_20$	20.3g
$MgCl_2.7H_20$	24.6g

SOC medium

SOB	980µl
2 M Mg^{++}	10µl
2 M Glucose	10µl

Appendix V (Solutions for Protein Purification, Refolding and Dialysis)

Ni-Binding Buffer (100 ml)

50mM Tris-HCl	5 ml (1 M stock)
6M Guanidine-HCl	11.4 g
NaCl (0.5 stock	25 ml
10 mM Imidozole	68 mg

Ni-wash buffer-I (200 ml)

50 mM Tris-HCl	10 ml (1 M stock)
6M Guanidine-HCl	11.4 g

0.5 M NaCl	5.85 g
50 mM Imidozole	0.68 g

Ni-wash buffer-II (200 ml)

50 mM Tris-HCl	10 ml (1 M stock)
6M Guanidine-HCl	11.4 g
0.5 M NaCl	5.85 g
50 mM Imidozole	1.36 g

Ni-elution buffer-1 (50 ml)

50 mM Tris-HCl	2.5 ml (1 M stock)
6M Guanidine-HCl	2.85 g
0.5 M NaCl	5.85 g
250 mM Imidozole	0.85 g

Ni-elution buffer-II (50 ml)

50 mM Tris-HCl	2.5 ml (1 M stock)
6M Guanidine-HCl	2.85 g
0.5 M NaCl	5.85 g
500 mM Imidozole	1.7 g

Denaturing-polyacrylamide gel (D-PAG) elution buffer (100 ml, pH. 8.0, filter sterilize)

50 mM Tris-Cl	5 ml (1M stock)
150 mM NaCl,	3 ml (5M stock)
2 mM EDTA	0.4 ml (0.5 M stock)
0.1 % SDS	1 ml (10 % stock)

Refolding buffer (20 ml, pH 11.0, filter sterilized)

50 mM Tris-HCl	0.4 ml (1M stock)
5 M guanidine-HCl	9.5g
1mM PMSF	200 µl (100 mM stock)
1 mM cystine	1ml (20 mM stock)
5 mM cysteine	1ml (100 mM stock)

Appendix VI (Publications)

Based on the study conducted for the present thesis, the following manuscripts have been published:

1. **Malik Nawaz Shuja**, Rob. W. Briddon and Muhammad Tahir (**2014**). Identification of a distinct strain of Cotton leaf curl Burewala virus. *Archives of Virology,* **159**:2787–2790.

2. **Malik Nawaz Shuja**, Muhammad Tahir and Rob W. Briddon (**2017**). Occurrence of a recombinant molecule carrying sequences derived from an alphasatellite and the helper virus in Cotton leaf curl disease affected cotton. *Tropical Plant Pathology,* DOI. 10.1007/s40858-017-0161-5.