

BIOLOGICAL AND MOLECULAR CHARACTERIZATION OF COTTON LEAF CURL GEMINIVIRUS FROM INDIA

G.RADHAKRISHNAN

T7125
IARI



T-7125



Division of Plant Pathology
Indian Agricultural Research Institute
New Delhi – 110 012

2002

BIOLOGICAL AND MOLECULAR CHARACTERIZATION OF COTTON LEAF CURL GEMINIVIRUS FROM INDIA

A Thesis

by

G.RADHAKRISHNAN

submitted to the Faculty of the Post Graduate School,
Indian Agricultural Research Institute, New Delhi,
in partial fulfillment of the requirements
for the award of the degree of

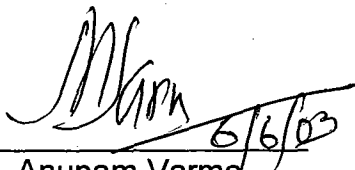
Doctor of Philosophy

in

**PLANT PATHOLOGY
2002**

Approved by:

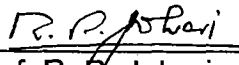
Chairman:


Prof. Anupam Varma

Members:


Dr. V. G. Malathi


Prof. K. R. Koundal


Prof. R. P. Johari

Prof. Anupam Varma
National Professor

Division of Plant Pathology
Indian Agricultural Research Institute
New Delhi - 110 012 (India)

CERTIFICATE

This is to certify that the thesis entitled, "**Biological and Molecular Characterization of Cotton Leaf Curl Geminivirus from India**", submitted to the Faculty of the Post-Graduate School, Indian Agricultural Research Institute, New Delhi, in partial fulfillment of the requirements for the award of the degree of **Doctor of Philosophy in Plant Pathology**, is a faithful record of *bona fide* research work carried out by **Mr. G.Radhakrishnan** under my guidance and supervision and that no part of the thesis has been submitted for any other degree or diploma.

The assistance and help received during the course of these investigations, have been duly acknowledged by him.

Place: New Delhi

Date : Dec 12, 2002



(Anupam Varma)
Chairman,
Advisory Committee

*Affectionately dedicated
to
my wife Renuka*

ACKNOWLEDGMENTS

I express my deepest sense of gratitude and sincere thanks to my chairman Prof. Anupam Varma, for suggesting the problem, valuable guidance, incessant encouragement, constructive criticism, expedient advice and affection during the course of investigation and preparation of this manuscript. His adeptness, creative thoughts, perspicuous mind and words shall for ever remain in my mind.

I place my sincere and heartfelt gratitude to my co-chairperson Dr.V.G.Malathi, for her inspiring lectures, guidance, fruitful advice, valuable suggestions, affectionate encouragement and untiring efforts to develop an independent thinking on me, especially in molecular biology work. I feel highly indebted for her maternal affection and blessings.

I owe a special debt of gratitude to the members of my advisory committee Prof. K.R.Koundal and Prof. R.P.Johari for their kindness, valuable suggestions and cooperation throughout the progress of this dissertation. I am highly thankful to Dr.P.Bahadur, Professor and Dr.D.V.Singh, Head, Division of Plant Pathology for their kindness and providing necessary facilities to undertake this study.

I am grateful to Dr.Y.S.Ahlawat, Dr.F.R.Niazi, Dr.R.K.Jain, Dr.Ishwar Bhat, Dr.Padma Ramachandran, Dr.Shelly Praveen, Dr.Bikash Mandal, Dr. V.K.Baranwal, Dr. P. Sinha, Dr.Jayashree Jayaraman, Dr.Reshmi Aggarwal, Dr.K.D.Srivastava, Dr.Jasvir Singh, Dr. R.P.Pant, Dr.P.Anandkumar, Dr.T.Mohapatra of IARI, Dr.S.K.Mukherjee (ICGEB, Delhi) and Dr.I.Dasgupta (Delhi Univ.) for their help and encouragement throughout the course of study.

I tender my profound thanks to Dr.D.Monga (CICR, Sirsa), Dr.P.Jeyakumar, (CICR, Sirsa), Dr.B.D.Ajmera (RAU, Sriganaganagar), Dr.M.S.Chauhan (HAU, Hisar) Dr.P.S.Sekhon (PAU, Ludhiana), Dr.(Mrs.)J.Amudha (CICR, Nagpur), Dr.V.P.Singh (IARI, New Delhi) and Dr.R.P.Singh (IARI, New Delhi) for providing their help in collection of samples and cotton seeds. I am immensely thankful to Dr.K.V.Prabhu, Arun Kumar, Sudhir and other staff of National Phytotron Facility, IARI for their whole hearted cooperation and timely help throughout the course of investigation.

Diction is not enough to express my thanks to Usha, Suren, Natarajan uncle, Dhandapani, Selvakumar, Periasamy, Anil, Nagma and Sivalingam for their help, encouragement and cooperation which made my research work enjoyable and worthy.

I am grateful to acknowledge the cheerful cooperation of my classmates Santosh, Suruli, Dutta, my labmates Savarni, Siva, Anirban, Al Shami, Behram, Hoa, Sohrab, Vikas, Meenakshi, Shally, Jyoti, Shalini, Renu, Petikam, Srinivasan, Thirumalai, K.K.Singh, my beloved friends

Siva (nates) Vivek, Maity, Basu, Sureja, Panda, Ramesh (Ento), Ramesh (Gen), Jayashree, Sekhar, Anandha, Gautam, Sivanandha, Bhoopathi and Balki and my beloved seniors Gandhi, Mohan, Dhinakaran, Sivakumar, Makesh, Narayana, Viswanathan and Suresh.

Perhaps words are not enough to express my indebtedness to my beloved parents, wonderful brothers and caring relatives whose unboundful love, affection, sacrifice and constant encouragement inspired me all the way of my research and study.

My eternal gratitude goes to Renuka for her love, affection, inspiration, encouragement and moral support without which my long cherished aspirations would have vanished.

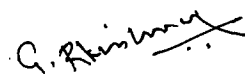
My special thanks are also due to Gandhiji, Paswanji, Rampreetji, Shibuji, Mehra madam, Baljeet madam, Tyagiji, Bishamberji, Amarjeet and Pallavi for their help whenever needed.

I am grateful to Mr. and Mrs. Sekhar for meticulous typing of this manuscript and Purushottamji of Trivikram Photo Studio for photographic work.

The Senior Research Fellowship provided by CSIR during the tenure of my Ph.D. programme is thankfully acknowledged.

Place : New Delhi

Date : 12.12.2002


(G. RADHAKRISHNAN)

CONTENTS

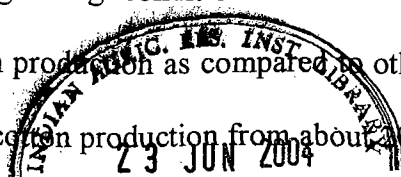
CHAPTER	PAGE
1. INTRODUCTION	1 – 5
2. REVIEW OF LITERATURE	6 – 50
3. MATERIALS AND METHODS	51
4. RESULTS AND DISCUSSION	52 – 115
5. SUMMARY	116 – 119
BIBLIOGRAPHY	i – xvii
APPENDICES	I – LI

1. INTRODUCTION

Cotton, one of the most ancient and important commercial-crops next only to food grains, is cultivated in India from sub-Himalayan region of Punjab in the North to Kerala in South and from dry regions of Kutch in the West to high rainfall areas of Manipur in East. Among the cotton growing countries, India occupies the foremost position in cotton acreage covering nearly 8.76 mha, but ranks third in production, after China and USA, with 11.64 mbales of lint, with a productivity of 226 kg/ha (Anonymous, 2001).

Cotton is extensively cultivated in three distinct agro-ecological zones viz., northern, central and southern regions. Four different species of *Gossypium* and their F1 hybrids are grown. The species composition has shifted from the predominance of diploid (*G. arboreum* and *G. herbaceum*) till the early sixties to that of *G. hirsutum* and its tetraploid hybrids and *G. barbadense*. Even now *G. arboreum* is grown widely in India covering around 29% of total cotton area, mostly in the cotton belt of Rajasthan. *G. herbaceum* with around 21% area under it, is confined mainly to the western coast extending from Gujrat to Tamil Nadu, whereas, *G. barbadense* is grown in less than 1% area mainly in southern India. Maximum area (>50%) is under *G. hirsutum*. The productivity ranges from 177 to 538 kg/ha in different states as the crop is grown predominantly as a rainfed crop, which constitutes about 64 per cent of the total crop area (Anonymous, 2001).

The northern cotton zone comprising of Punjab, Rajasthan, Haryana and western U.P. has witnessed a high growth of cotton production. As this zone is endowed with irrigation and relatively uniform growing conditions and cultural practices it has high potential for further increase in production as compared to other zones. This zone contributes about 27% to the total cotton production from about 20%



of total area with a productivity of 450 kg/ha and is considered as the 'cotton basket' of the country. Two species of cotton namely, *G. hirsutum* and *G. arboreum* cover an area of 85 and 15 per cent, respectively, of the total cotton area in this zone. The central zone comprising Maharashtra, Gujarat and Madhya Pradesh covers 56% of the whole cotton area and contributes 47% production with an average productivity of 270kg/ha. The southern cotton zone comprising Andhra Pradesh, Tamil Nadu and Karnataka contributes 26% to the national production with 23% of the area and an average productivity of 367kg/ha (Anonymous, 2001).

Insect pests and diseases have always been the major constraint in the cultivation of cotton in these zones. The severity of this problem can be gauged from the fact that, cotton is the single largest crop accounting for nearly 50% of the total plant protection chemicals used in the country (Gupta, 2001). Cotton is attacked by a number of insect pests and diseases. The major insect pests are aphids, jassids, whiteflies, thrips, leaf roller, red cotton bug, mealy bug, stem weevil, ash weevil, tobacco cutworm, bollworms and gram caterpillar. The important diseases are root rot, anthracnose, leaf blight, leaf curl, bacterial blight, wilt, *Alternaria* leaf spot, *Cercospora* leaf spot, *Helminthosporium* leaf spot and grey mildew.

In the recent years, leaf curl disease of cotton (CLCuD) caused by a whitefly (*Bemisia tabaci* Genn.) transmitted geminivirus (WTG) is posing a major threat to the cotton cultivation in northern India (Varma, 1994) with the yield losses varying from 68% to 79% depending on the cultivars used. The worst affected states being Rajasthan, Haryana and Punjab. CLCuD, was first reported from Nigeria in 1912 on the native cotton species *G. peruvianum* and *G. vitifolia* (Farquharson, 1912), has caused serious economic losses in Sudan (Nour and Nour, 1964). The disease was not known in the Indian sub-continent until late 1960's when it was first observed in a small area at Multan in Pakistan, and by 1987 this CLCuD reached an epidemic

proportion in major cotton-growing areas of Pakistan (Hameed *et al.*, 1994; Briddon and Markham, 2000). Since then, it has spread to about 1 million hectares of cotton (*G. hirsutum*), mainly in Punjab province and the northern edge of Sindh in Pakistan (Zhou *et al.*, 1998). The losses caused by this disease to the Pakistan economy are calculated to be about US \$ 5 billion between 1992 and 1997 (Briddon and Markham, 2000).

In India, CLCuD was first observed in the experimental fields of IARI, New Delhi in 1989 in a few accessions of *G. barbadense* (Anonymous, 1990; Varma *et al.*, 1993) and by 1993 it assumed epidemic proportion in *G. hirsutum* crops adjoining international border with Pakistan near Sri Ganganagar in Rajasthan and Punjab (Varma, 1994; Narula *et al.*, 1999).

Around the same time, this disease was also found in the kitchen gardens in Bangalore (Nateshan *et al.*, 1996), although even the highly susceptible cultivars of *G. barbadense* were free from this disease in Coimbatore in 1995 and are free even now (Varma, A: personal communication). The occurrence of the disease, however, remained unnoticed in other areas and by 1994 the virus had devastated more than 1,00,000 ha. Now this disease has spread to all the major cotton growing areas of Punjab, Haryana and Rajasthan and has posed an imminent threat to all the cotton farmers of the country (Varma, 1995).

CLCuD is a whitefly transmitted viral disease and the symptoms manifest primarily in the form of downward or upward curling of leaves. The very characteristic symptoms include leaf curling, thickening of veins and enations that frequently develop into cup shaped, leaf-like structures on the undersides of leaves. The finding that cotton leaf curl disease (CLCuD) is transmitted specifically by the whitefly *B. tabaci* led researchers to believe that a begomivirus might cause the disease. The aetiology of the virus was established in 1989 from India when Enzyme

Linked Immuno Sorbant Assay (ELISA) and Immuno Sorbant Electron Microscopy (ISEM) tests conducted using polyclonal antisera to India cassava mosaic virus (ICMV) showed the causative agent to be a geminivirus (Varma, *et al.*, 1993). Later, Mansoor *et al.*, (1993) showed the presence of a begomovirus in CLCuD infected plants by PCR using primers designed to allow the amplification of all dicot-infecting begomoviruses (Briddon and Markham, 1994). The begomovirus associated with CLCuD is named cotton leaf curl virus (CLCuV).

The genome of CLCuV consists of a single component, equivalent to the DNA-A component of the bipartite begomoviruses (Briddon *et al.*, 2000). The cloned genome of CLCuV DNA-A was infectious to cotton and *Nicotiana benthamiana*, but did not induce symptoms typical of CLCuD. This was conclusive evidence that CLCuV DNA-A is not, or not the sole, causative agent of CLCuD. Attempts to identify a second genomic component (DNA-B) by several workers were not successful, suggesting that CLCuV is a monopartite begomovirus. However, Pappu *et al.*, (1997) cloned the DNA-B component of an isolate of the virus from India and also sequenced partially, but so far information on the complete sequence is lacking. An advance was made in the aetiology of CLCuV with the report of a novel single-stranded, circular DNA molecule that was associated with CLCuD (Mansoor *et al.*, 1999). This DNA molecule is 1376 bases long (termed as DNA 1) and appears to encode a single protein with similarity to the replication-associated protein (Rep) of nanoviruses (Mansoor *et al.*, 1999). This molecule is capable of self-replication in plant cells but require CLCuV for spread in plants as well as insect transmission (Mansoor, *et al.*, 1999). Infectivity studies using cloned DNA components have demonstrate that DNA 1 moves systemically throughout cotton when co-inoculated with CLCuV DNA-A but is unable to confer typical leaf curl symptoms, suggesting that one or more additional viral components are associated with the disease.

Recently, Briddon *et al.*, (2001) have identified and isolated a single-stranded DNA molecule approximately 1350 nucleotides in length from infected cotton plants which, when co-inoculated with the CLCuV DNA-A, induces symptoms typical of CLCuD, including vein swelling, vein darkening, leaf curling and enations. This molecule (termed as DNA- β) requires the CLCuV DNA-A for replication and encapsidation. The CLCuV DNA-A/DNA1/DNA β complex, together with a similar complex previously identified in *Ageratum conyzoides* causing Ageratum yellow vein disease (Saunders *et al.*, 2000), represent members of an entirely new type of infectious, disease causing agents.

In India, most of the work has been on the detection of CLCuV by electron microscopy and serology (Varma *et al.*, 1993), occurrence of the disease and its spread under field conditions in Rajasthan, Punjab and Haryana (Varma, 1995) and identification of resistant cultivars (Narula, *et al.*, 1999). Studies on the host range, vector relations, epitope profile and serological relationships of CLCuV isolate from southern India have also been carried out (Nateshan *et al.*, 1996).

CLCuD is emerging as a major constraint in the north-western cotton growing areas. There is an urgent need to contain this disease to improve productivity of cotton. For developing management strategies a good understanding of the causal virus and its biology is essential. For achieving this goal, the present study was undertaken with the following specific objectives:

1. Biological characterization of an isolate of CLCuV from Rajasthan.
2. Cloning of the selected isolate and its molecular characterization.
3. Detection of additional genomic components.
4. Molecular detection of CLCuV genomic components.
5. Determining variability among CLCuV isolates in North-West India.

2. REVIEW OF LITERATURE

The recent outbreak of leaf curl disease caused by a whitefly (*Bemisa tabaci* Genn) transmitted geminivirus (WTG), in epidemic proportion in Northern India has posed serious threat to the cultivation of cotton (Varma, 1994). The disease was first reported from Nigeria in 1912 on *Gossypium peruvianum* and *G. vitifolia* (Farquharson, 1912). After causing serious economic losses in Sudan (Nour and Nour, 1964), Egypt (Smith, 1982) and Pakistan (Hameed *et al.*, 1994), the disease affected cotton in Northern India during 1993 in pockets of Rajasthan (Varma, 1994). In India, this disease was first observed in I.A.R.I., New Delhi in 1989 in few isolated plants of *G. barbadense* (Anonymous, 1990; Varma *et al.*, 1993). The infected plants develop very characteristic symptoms which include leaf curling, darkened veins, vein swelling and enations that frequently develop into cup shaped, leaf-like structures on the undersides of leaves. Studies on the aetiology of cotton leaf curl disease (CLCuD) have shown the disease causing agent to be a member of an entirely new type of infectious agents. This new group of pathogens consists typically of a monopartite *Begomovirus* (CLCuV), a DNA1 component originating from nanoviruses and a single-stranded DNA molecule (DNA- β), which is approximately half the length of the *Begomovirus* component (DNA-A) and which has no sequence similarity to either the helper *Begomovirus* or the DNA1 (Briddon *et al.*, 2001). CLCuD is the first crop disease attributable to this new group of pathogens and also includes Ageratum yellow vein disease (Saunders *et al.*, 2000), a disease of weed species. Here the current information on the disease is briefly reviewed.

VIRAL DISEASES OF COTTON

Virus diseases of cotton have only been of sporadic importance to global cotton production. Recent devastating epidemics in Pakistan and other areas have

brought new awareness to the potential for disaster of a pathogen once considered to be of minor importance. However, under changing conditions the pathogen emerges as a serious problem.

Brown (1992) extensively reviewed viral diseases of cotton. A brief summary of suspect or putative viral diseases of cotton, the vectors and geographical distribution is provided in Table 2.1 & 2.2 (Brown, 1992).

Table 2.1 Summary of partially characterized virus and virus-like diseases cotton

Disease	Vector/ Transmission	Geographical distribution	References
African cotton mosaic	<i>Bemisia tabaci</i>	Africa (central, western)	Bink, 1973
Cotton anthocyanosis ^a	<i>Aphis gossypii</i>	Brazil	Costa, 1957
Cotton blue disease	<i>A. gossypii</i>	Africa, South America	Cauquil & Follin, 1983
Cotton leaf crumple ^a	<i>B. tabaci</i>	USA (South-Western)	Dickson <i>et al.</i> , 1954
Cotton leaf curl (Cotton leaf crinkle)	<i>B. tabaci</i>	Africa (Chad, central, Nigeria, Sudan, Togo, Western)	Farquharson 1912, Golding, 1930 Kirkpatrick, 1931
Cotton yellow vein ^a (Texas vein-clearing virus)	<i>Scaphytopius</i> <i>albifrons</i>	USA (Texas)	Rosberg, 1957

^a Diseases for which virus-like particles have been observed. (Adapted from Brown, 1992)

Table 2.2 Summary of suspect virus-like disorders of cotton

Disease	Vector/transmission	Geographical distribution	Original references
Acromania (crazytop)	Unknown	USA (Arizona), China, Haiti, Sudan	Bedford, 1938 Cook, 1920, 1923
Cotton (common) yellow mosaic	<i>Bemisia tabaci</i>	Brazil	Costa and Forster, 1938; Costa, 1955
Cotton mosaic	<i>B. tabaci</i> Unknown	Central America USA (Louisiana)	Quant, 1977 Neal, 1946
Cotton leaf curl	<i>B. tabaci</i>	Philippines	Reyes <i>et al.</i> , 1959; Nour, 1960
Cotton leaf mottle	Grafting	Sudan	Nour, 1959
Cotton leaf roll	<i>Aphis gossypii</i> <i>A. gossypii</i>	Thailand Paraguay	Delalande, 1970
Cotton leaf roll/leaf curliness	<i>A. gossypii</i>	USSR (former)	Moskovetz, 1941
Infectious variegation	Grafting, <i>B. tabaci</i>	India	Mali, 1978
Indian leaf crumple	Grafting, <i>B. tabaci</i>	India	Mali, 1977
Murcha vermelhao/ purple wilt	<i>A. gossypii</i>	South America	Costa and Sauer, 1954
Terminal stunt	Unknown	USA (Texas)	Sleeth <i>et al.</i> , 1963
Tobacco streak	Possibly thrips	Brazil	Costa, 1961
Viral wilt	Grafting	India	Mali, 1979

(Adapted from Brown, 1992)

While over 20 virus diseases of cotton have been described in the APS cotton disease compendium, only a few have actually been shown to be of virus etiology.

The main ones that have been proven to actually of virus etiology, include two

geminiviruses (Brown and Nelson, 1984; Mansoor *et al.*, 1993 and Nadeem *et al.*, 1997), and tobacco streak virus (Cauquil and Folin, 1983).

The geminiviruses are cotton leaf crumple virus (CLCrV) and cotton leaf curl virus (CLCuV). Leaf crumple is found in the South western USA and Central and South America, while leaf curl has been described in Africa and Asia (Nelson *et al.*, 1998). These viruses are quite different when the DNA sequences are compared (Nadeem *et al.*, 1997) and produce different symptoms in the field. These two cotton geminiviruses have been sporadically important over the years. Leaf crumple, first described in the 1950's in the USA, is occasionally a problem when cotton is grown as a perennial, or if it is planted late or near spring vegetables supporting high populations of whiteflies (Nelson *et al.*, 1998).

Cotton mosaic disease, caused by tobacco streak virus (TSV) has recently been found to be widespread in Pakistan (Nelson *et al.*, 1998). Like leaf curl, it too has a long history in Africa and elsewhere. Tobacco streak virus is readily transmitted in sap, in contrast to the geminiviruses and has been shown experimentally to suppress cotton plant productivity when infections occur early (Cauquil and Folin, 1983). The virus was widespread in the Punjab area of Pakistan in 1997 (Nelson *et al.*, 1998). Attempts to experimentally transmit TSV by thrips, as is the case in some other related viruses, has not yet been successful (Cauquil and Folin, 1983; Nelson *et al.*, 1998).

OCCURRENCE AND GEOGRAPHICAL DISTRIBUTION OF COTTON LEAF CURL DISEASE

Cotton leaf curl was first reported by Farquharson in 1912 in native cottons grown in Nigeria. According to Massey and Andrews (1932), the first case of leaf curl in the Sudan was probably in 1923. The disease was also recorded in Tanzania (Jones and Mason, 1926; Kirkpatrick, 1930), Egypt (Smith, 1982) and in Africa North of

equator (El Nur and Abu Salih, 1970; Cauquil and Follin, 1983; Brown, 1994). The leaf curl was first observed in Multan, Pakistan in 1967 (Hussain and Ali, 1975). In India, this disease was first observed in IARI, New Delhi in 1989 in few isolated plants of *G. barbadense* (Anonymous, 1990) and later on *G. hirsutum* near Sri Ganganagar, Rajasthan in 1993 (Ajmera, 1994). Around the same time, this disease was also found in the kitchen gardens in Bangalore, Karnataka (Nateshan *et al.*, 1996).

INCIDENCE AND LOSSES CAUSED DUE TO CLCuD

First outbreak of this disease occurred in Nigeria in 1924 (Jones and Mason, 1926). With the increased cultivation of cotton, the disease became widespread, and heavy losses were reported in 1928-29 (Kirkpatrick, 1931; Massey and Andrews, 1932; El Nur and Abu Salih, 1970) and since then was found to have caused serious economic losses in Sudan (Nour and Nour, 1964). In Pakistan, this disease reached an epidemic proportion in major cotton-growing areas during the year 1987 (Hameed *et al.*, 1994) and since then it has spread out from a small area to effect about 1 million hectares of cotton (*G. hirsutum*), mainly in Punjab province and the northern edge of Sindh (Zhou *et al.*, 1998). The losses to the Pakistan economy are calculated to have been 5 billion US \$ in the 5 years between 1992 and 1997 (Briddon and Markham, 2000).

In India, since the first record of CLCuD in IARI, New Delhi in 1989 (Anonymous, 1990) and later on near Sri Ganganagar, Rajasthan in 1993 (Ajmera, 1994), the disease continued to spread and in subsequent years 1994 and 1995 it occupied around 300 and 500 hectares area, respectively in Sri Ganganagar, Rajasthan (Narula *et al.*, 1999). In 1996 its outbreak took place and about 10946 ha of area was found affected with this disease. In Punjab also, the disease was noticed at Ferozepur, Bhatinda and Faridkot districts in 1994 in an area of 1200 ha but in 1995 it spread to

Abohar and Muktsar district. The disease further spread and covered about 1500 ha in 1996 crop season. The situation in Haryana was better and up to 1996 the disease was limited to about 20ha. However, the year 1997 saw a sudden flare up of this disease in all the three states and as per rough estimates an area of about 2.19 lakh ha was infested due to this disease (Table 2.3) (Narula *et al.*, 1999).

Table 2.3 Area affected by leaf curl of cotton in North India (Narula *et al.*, 1999)

State	Area affected (ha)				
	1993	1994	1995	1996	1997
Rajasthan	Patches	480	343	10946	100000
Punjab	Nil	1200	40	1500	80000
Haryana	Nil	20	1.0	17	38610
North Zone	Patches	1700	384	12463	218610

CLCuV occurred widely in part of North Western India on *G. hirsutum* during 1994, its incidence varying from 1% to 97% on cvs. F846, Pakistani NIAB-72, F505, LH1134, LH900, PL 104, RST9, Somnath and Ganganagar Ageti (Singh *et al.*, 1994).

YIELD LOSS DUE TO THE DISEASE

Lewin (1927) reported reduction in the size of bolls due to a reduction in the number of seeds per boll. Massey (1934) made a quantitative examination and showed reduction in yield up to 50 percent on account of CLCuV infection. Andrews (1936) conducted an experiment on estimation of losses and found that this disease reduced weight of opened bolls by 48.6 per cent, number of green bolls by 48.6 percent and number of bud and flowers by 4.3 percent. Tarr (1957) stated that the leaf curl disease outbreak in Gezira, Sudan caused 30-40 percent yield loss. The cultivar Sakel (*G. barbadense*) was affected and suffered severe losses (Tarr, 1964).

Shafique *et al.*, (1994) in Pakistan noted that cotton yields of different popular varieties S12, Niab 78, MNH 93, CIM109 and CIM70 during 1992-93 were 553, 676, 670, 733 and 556 kg/ha, respectively, showing there by that in the highly affected districts average yields of different cotton varieties decreased to 50 percent. The lint yield per hectare dropped down from 768 kg/ha in 1991-92 to 543 kg/ha in 1992-93 and 488 kg/ha in 1993-94 reducing the production from 12.8 m bales in 1991-92 to 9.05 m bales in 1992-93 and 8.04 m bales in 1993-94.

In India, extensive field experiments have been conducted to estimate the yield loss due to CLCuD on popular varieties F846, RST9, HS6 and Pakistani Nerma (NIAB-72) mainly in Rajasthan (Ajmera, 1996), Punjab (Singh *et al.*, 1994) and Haryana (Monga *et al.*, 1998). (Table 2.4).

Table 2.4 Yield loss due to CLCuD

Cultivar	Reduction in seed cotton (%)	Reduction in opened bolls (%)
F846	10.5-92.2	14.9-87.4
RST9	32.9	22.8
HS6	81.4	88.4
Pakistani Nerma (NIAB 72)	39.9-79.7	30.7-68.6

(Singh *et al.*, 1994; Ajmera, 1996 and Monga *et al.*, 1998)

Ten promising germplasm lines were tested to study the effect of CLCuV on yield and quality and parameters (Monga *et al.*, 1999).. There was a drastic reduction in seed cotton yield in diseased plants (68.1-79.2%) of all the tested lines. Reduction in boll weight ranged from 1.6-37.1 per cent whereas boll number reduction varied from 54.3-73.2 per cent. The reduction in boll number due to CLCuV was move as compared to boll weight (Monga *et al.*, 1999).

The disease also affects the quality of cotton fibre. Ajmera (1997) reported reduction in 2.5 per cent span length by 2.5 percent in RST9 and 3.5 percent in F846 whereas in bundle strength, it was 3.4 percent in RST9 and 6.1 per cent in F846.

SYMPTOMS OF THE DISEASE

External symptoms

The initiation of disease is characterized by small vein thickening (SVT) type symptoms on young upper leaves of plants. Dark green bead like thickening of small veins can be clearly seen by holding the leaf against sunlight and observing it from lower side. These irregular thickenings gradually extend and coalesce to form continuous reticulation of small veins. The disease is further characterized by upward curling of leaves which occurs because of the uneven growth of veinal tissues on the abaxial side of the leaves. Veins of the leaves become thickened which are more pronounced on the underside. Affected leaves become thick, leathery, brittle and more green than healthy leaves. Similar symptoms are also observed on bracts. The disease also causes vein swelling and enations of the veins which frequently develop into cup shaped, leaf-like outgrowths on the undersides of leaves. A spiral twisting of the petiole, peduncle, and the youngest parts of the stem is also common; while in severe cases buds are either totally suppressed or die shortly after formation, causing partial or complete sterility of the plant. In plants affected at early age, reduction of internodal length leading to stunting and reduced flowering/fruiting is observed (Massey, 1932; Nour and Nour, 1964; Narula *et al.*, 1999).

Symptoms may first appear in the leaves near any growing point e.g., that of a monopodium or sympodium – but in by far the greater number of cases the first signs of disease are seen in the leaves near the main growing point. After the first symptoms have appeared, secondary growth from nodes below that of the infected leaf usually shows the disease in increased severity.

Internal symptoms

Tissues, which have suffered infection during active growth show striking changes. The principal effect is the stimulation of normal non-meristematic tissue to assume meristematic activity, leading to the production of various forms of tissue. Severe hypertrophy of the cells of the leaves, bract and stem occurs, resulting in the malformations described above. In the stem irregular bundles develop in the cortex with a formation of central xylem elements surrounded by a cambium and phloem. These bundles are mainly confined to the ribbed portion of the stem, but there are definite signs of meristematic activity in the whole of the cortical layer. In the infected petiole new vascular bundles may appear in the medulla.

An infected leaf is of a much thicker texture than a healthy one, and it is in the leaf that the formation of new growth is most evident. In the infected leaf, there is no well-defined palisade layer, rectangular cells with large intercellular spaces occurring on both sides of the leaf, which take the place of normal palisade tissue. Formation of similar cells is found also below the larger vascular bundles on the under surface of the leaf, giving rise to an abnormal amount of spongy tissue. Numerous conducting elements are formed in this spongy tissue, and generally there is in the whole of the leaf both externally and internally growth without control (Massey, 1932).

Ashraf and Zafar (1999) studied the physiological characteristics in CLCuV infected cotton cultivars. They observed that leaf nitrogen content increased considerably due to disease whereas leaf potassium and calcium contents were lower in the diseased leaves of *G. hirsutum* cv. S12 than in healthy ones. A marked reduction in chlorophyll b content was observed in the diseased leaves of *G. hirsutum* cv. S12. They also observed decrease in leaf water potential in *G. hirsutum* cv. S12 and NIAB-Krishma due to disease.

Ashraf *et al.*, (1999) observed that in the midrib of the diseased leaf of S12, the xylem and phloem elements were partially or completely collapsed, the vascular elements of the veins were severely affected and thickening of midribs and veins in diseased leaves seemed to be mainly due to proliferation of parenchymatous tissues in these leaf parts.

Gurdeep *et al.*, (1998) studied the influence of CLCuV infection on the polyphenol metabolism in the leaves of resistant and susceptible cotton plants. They observed that peroxidase activity increased in moderately infected leaves compared with in healthy leaves, but decreased in the severely infected leaves. Catechol oxidase increased in all leaves with increasing disease severity.

Incubation period of the virus within the cotton plants

The incubation period of the virus in the cotton plant varied from 8 to 34 days (Cowland, 1934; Tarr, 1951; Couteaux, *et al.*, 1968; Yassin and El Nur, 1970). Considerable variation in the length of the incubation period of the virus within the cotton plant is encountered. It was found to range from 1 – 5 weeks or even larger, with an average of 11-19 days. (El Nur and Abu Salih, 1970). Tarr (1951) reported that plant age had no effect on the length of incubation period, and that higher temperature tended to be associated with shorter incubation periods. However, Abu Salih (1965) reported an increase in the minimum incubation period with plant age. Similar results were reported by El Nur and Abu Salih (1966). They reported that when the plant age (in weeks) was 2, 4, 6, 8, 10, 12, 14, and 16, the mean minimum incubation period (in days) were 12, 15, 24, 34, 41, 29, 35 and 55, respectively.

THE CAUSAL VIRUS

The taxonomy of the *Geminiviridae* divides the viruses into four genera. The division is on the basis of insect vector [either whitefly (*Begomovirus*), leafhopper (*Mastrevirus* and *Curtovirus*), or treehopper (*Topocuvirus*)] and genome arrangement.

Most of the economically important geminiviruses are in the genus *Begomovirus*, which consists of more than 70 species that are all transmitted by the whitefly *B. tabaci* and infect only dicotyledonous plants. Examples of economically destructive begomoviruses include cassava mosaic viruses (the most limiting factor to cassava production across South and Central Africa; Thottapilly, 1992), tomato yellow leaf curl viruses (TYLCVs), a severe problem to tomato cultivation across the Mediterranean region (Czosnek and Laterrot, 1997; Moriones and Navas-Castillo, 2000) and mungbean yellow mosaic virus which infects large number of leguminous crops in the Indian subcontinent, results in most severe economic losses in pulse production in this region (Chenulu and Varma, 1988; Varma *et al.*, 1992). Although the majority of begomoviruses have bipartite genomes, an increasing number are being identified that have only single component equivalent to the DNA-A component of the bipartite viruses. TYLCV is the most notable and economically most significant example of a monopartite *Begomovirus* (Navot *et al.*, 1991).

The finding that CLCuD is transmitted specifically by the whitefly *B. tabaci* led researchers to believe that a *Begomovirus* might cause the disease. The aetiology of the virus was first established in 1989 from India when ELISA and ISEM tests conducted using polyclonal antisera to Indian cassava mosaic virus (ICMV) proved the causative agent to be a geminivirus (Varma *et al.*, 1993). Later, Mansoor *et al.*, (1993) showed the presence of a *Begomovirus* in CLCuD-infected plants by PCR using primers designed to allow the amplification of all dicot-infecting geminiviruses (Bridson and Markham, 1994). The *Begomovirus* associated with CLCuD was named cotton leaf curl virus (CLCuV). Subsequent PCR-based studies by various groups established the diversity of CLCuV in Pakistan, showing that there were more than four distinct, yet closely related, *Begomovirus* species associated with CLCuD (Nadeem *et al.*, 1997; Zhou *et al.*, 1998). However, none of these investigations

attempted infectivity studies with their cloned components, or even tried to show the presence of a second (DNA-B) component. Nevertheless, in several reports, it was concluded that CLCuV DNA-A was the causative agent of CLCuD (Hameed *et al.*, 1993; Liu *et al.*, 1998; Zhou *et al.*, 1998). The findings of Briddon *et al.* (2000) showed that the genome of CLCuV consists of a single component (equivalent to the DNA-A component of the bipartite begomoviruses). The cloned genome of CLCuV DNA-A was infectious to cotton and *Nicotiana benthamiana*, but did not induce symptoms typical of CLCuD. Plants infected with the cloned component (DNA-A) of CLCuV showed very mild symptoms, with none of the typical characteristics of CLCuD such as vein darkening, enations and leaf-like growths on veins on the undersides of leaves (Briddon *et al.*, 2000). This was conclusive evidence that CLCuV DNA-A is not, or not the sole, causative agent of CLCuD. Attempts to identify a second genomic component (DNA-B) by several workers were not successful, suggesting that CLCuV is a monopartite *Begomovirus*. However, Pappu *et al.*, (1997) cloned the DNA-B component of an isolate of the virus from India and also sequenced partially, but so far information on the complete sequence is lacking.

An advance was made in studies of the aetiology of CLCuD with the report of a novel single-stranded, circular DNA molecule that was associated with CLCuD (Mansoor *et al.*, 1999). This DNA molecule is 1376 bases long (termed as DNA1), is approximately half the size of CLCuV DNA-A but shares no sequence homology indicating that it is not a defective geminivirus component and appears to encode a single protein with similarity to the replication-associated protein (Rep) of nanoviruses (Mansoor *et al.*, 1999). This molecule is capable of self-replication in plant cells but require CLCuV for spread in plants as well as insect transmission (Mansoor *et al.*, 1999).

Nanoviruses are a genus of plant infecting multicomponent, single-stranded DNA viruses associated with diseases such as banana bunchy top and faba bean necrotic yellows (Timchenko *et al.*, 2000).

A similar nanovirus-like component (NLC) has been detected subsequently in *Ageratum conyzoides* plants originating from Singapore and exhibiting yellow vein symptoms (Saunders and Stanley, 1999). As with CLCuD, ageratum yellow vein disease was assumed to be caused by a monopartite *Begomovirus* Ageratum yellow vein virus (AYVV) (Tan *et al.*, 1995). Clones of AYVV DNA-A were not infectious to *A. conyzoides*. Subsequently, a nanovirus-like component, closely related to that detected in CLCuD was identified (Saunders and Stanley, 1999) in yellow vein symptomatic *A. conyzoides*.

Infectivity studies using DNA components have demonstrated that DNA1 moves systemically throughout cotton when co-inoculated with CLCuV DNA-A but is unable to confer typical leaf curl symptoms, suggesting that one or more additional viral components are associated with the disease (Mansoor *et al.*, 1999).

Similar results were also seen in Ageratum yellow vein disease (AYVD) with AYVV DNA1 and AYVV DNA-A (Saunders and Stanley, 1999). In AYVD, AYVV DNA1 is dependent on AYVV DNA-A for systemic infection of *A. conyzoides* and *Nicotiana benthamiana*. *A. conyzoides* plants coinfecting with AYVV DNA-A and DNA1 remain asymptomatic, indicating that additional factors are required to elicit yellow vein disease (Saunders and Stanley, 1999).

Finally, Saunders *et al.*, (2000) demonstrated that yellow vein disease of *A. conyzoides* results from co-infection of AYVV DNA-A (2741 nt) and a novel circular DNA that is approximately half its size (1347 nt), (designated as DNA- β). Apart from the sequence TAATATTAC, common to all geminiviruses and containing the initiation site of rolling circle replication, DNA- β shows negligible sequence

homology either to AYVV DNA-A or to DNA-B associated with bipartite begomoviruses (Saunders *et al.*, 2000). DNA- β depends on DNA-A for replication and is encapsulated by DNA-A encoded coat protein and so has characteristics of a DNA satellite (Saunders *et al.*, 2000).

Recently, Briddon *et al.*, (2001) have identified and isolated a single-stranded DNA molecule approximately 1350 nucleotides in length from infected cotton plants which, when co-inoculated with the CLCuV DNA-A, induces symptoms typical of CLCuD, including vein swelling, vein darkening, leaf curling and enations. This molecule (also termed as DNA- β) requires the CLCuV DNA-A for replication and encapsidation.

The CLCuV DNA-A/DNA1/DNA- β complex (Mansoor *et al.*, 1999; Briddon *et al.*, 2000; Briddon *et al.*, 2001), together with a similar complex previously identified in *Ageratum conyzoides* causing AYVD (Tan *et al.*, 1995; Saunders and Stanley, 1999; Saunders *et al.*, 2000), represent members of an entirely new type of infectious, disease causing agents.

GENOME ORGANIZATION OF GEMINIVIRUSES

The geminiviruses are the single stranded DNA (ssDNA) viruses infecting a large number of important crop plants. They are characterized by a circular single stranded DNA genome of 2.5-3.0kb, encapsidated in unique geminate quasi-isometric virus particles of size 15-18 x 30 nm (Goodman, 1977; Harrison *et al.*, 1977; Stanley, 1985; Davies and Stanley, 1989; Lazarowitz, 1992; Hanley-Bowdoin *et al.*, 1999; Harrison and Robinson, 1999).

According to the latest taxonomy of the family *Geminiviridae*, the following four genera have been recognized viz., *Mastrevirus*, *Curtovirus*, *Begomovirus* and *Topocuvirus* (Rybicki *et al.*, 2000). Their genome organization and their vectors distinguish these four genera from one another.

1. Genus *Mastrevirus* [type species: Maize streak virus (MSV)]

This genus contains the economically important MSV. All members have narrow host ranges and, with the exception of tobacco yellow dwarf virus (TYDV) and bean yellow dwarf virus (BeYDV) which infect dicotyledonous species, their host ranges are limited to species in the Poaceae. Mastreviruses are transmitted by leafhoppers in a circulative, non-propagative manner (Harrison, 1985). They have genomes comprising a single component of ssDNA. The nucleotide sequence of the MSV genome (Nigerian strain) revealed a single DNA component of 2687 nucleotides that had no detectable overall sequence similarity with genomes of geminiviruses with two DNA components (Begomoviruses) (Mullineaux *et al.*, 1984). Analysis of ORFs with potential for polypeptides of 10 kDa or larger revealed four potential coding regions arranged as shown in Figure 2.1. There are two intergenic regions, the long intergenic region (LIR) and the short intergenic region (SIR). Most, if not all, of the transcription starts in the LIR and terminates in the SIR. There are two ORFs on the viral strand (given the designation V) and two are on the complementary strand (given the designation C). ORF V1 encodes the cell-to-cell movement protein, ORF V2 the virion coat protein, and ORF C1 and C2 encode the two replication-associated proteins.

2. Genus *Curtovirus* [type species: Beet curly top virus(BCTV)]

The type member of this genus used to be an economically important virus of sugarbeet and has a wide host range. Curtoviruses are transmitted by leafhoppers in the circulative non-propagative manner (Harrison, 1985). They have a monopartite ssDNA genome. The sequence of 2993 nucleotides in an infectious clone of BCTV has been established (Stanley *et al.*, 1986). The organization of the single-component genome (Fig. 2.1) reveals that *Curtovirus* DNAs produce six or seven proteins. The product of ORF V1 is the virion coat protein, that of ORF V2 is involved in regulation

of the genomic ss and ds DNA levels, and that of ORF V3 facilitates the cell-to-cell movement of the virus. The four proteins encoded by the complementary-sense ORFs are involved in virus replication. ORF C1 encodes the replication initiator (Rep) protein, C3 a protein analogous to the replication enhancer protein of begomoviruses, and C4 a protein that can initiate cell division; the function of the product of ORF C2 is unknown.

3. Genus *Topocuvirus* [type species : Tomato pseudo-curly top virus (TPCTV)]

This is a very recently designated genus (Pringle, 1999) that has been split off from the *Curtovirus* genus. It has a similar genome organization to the curtoviruses but it is transmitted by the treehopper *Micrutalis malleifera*. The single component genome of TPCTV has properties of both mastreviruses and begomoviruses (Briddon *et al.*, 1996) (Fig. 2.1). The coat protein has features more akin to the leafhopper-transmitted mastreviruses than to those transmitted by whiteflies whereas the organization of the complementary-sense genes is similar to that of the single-component (monopartite) begomoviruses.

4. Genus *Begomovirus* [type species: Bean golden mosaic virus (BGMV)]

This is the largest geminivirus family, members of which have narrow host ranges among dicotyledonous species. They are transmitted by whiteflies (*B. tabaci*). Most members of this genus have their genomes divided between two DNA components (bipartite), though genomes of some members are monopartite. Bean golden mosaic virus (BGMV), African cassava mosaic virus (ACMV), tomato golden mosaic virus (TGMV), mungbean yellow mosaic virus (MYMV) are the example of bipartite *Begomovirus* group while tomato yellow leaf curl virus (TYLCV) from Israel, Thailand and Sardinia have a monopartite genome (Rochester *et al.*, 1990; Navot *et al.*, 1991; Kheyr-Pour *et al.*, 1991).

Padidam *et al.*, (1995) had subdivided whitefly transmitted geminiviruses (WTGs) into two group i.e., WTGs from the New World e.g. Brazilian isolate of bean golden mosaic virus (BGMV), Californian isolate of squash leaf curl virus (SqLCV) and WTGs from the Old World e.g. Indian cassava mosaic virus (ICMV), Thailand isolate of MYMV (MYMV-T) on the basis of sequence analysis of N-terminal region (60-70 amino acid) of the coat protein gene and biological properties.

The bipartite begomoviruses

Among the bipartite begomoviruses, ACMV was the first virus to be cloned and sequenced (Stanley and Gay, 1983). The genome of several other begomoviruses have now been cloned and sequenced. Comparison of genome of all these viruses revealed that the genome of bipartite begomoviruses consists of two circular DNA molecules designated as DNA-A and DNA-B encapsidated by a geminate particle (Stanley, 1985; Morinaga *et al.*, 1987; Elmer *et al.*, 1988). Both the genomic components are essential for virus infectivity (Hamilton *et al.*, 1983; Stanley, 1983). In DNA-A there are at least five consensus open reading frames (ORFs) among which one (AV1) is present in viral sense and the other four (AC1, AC2, AC3, AC4) are present in complimentary sense. In DNA-B there are two functional ORFs, one in each sense (BV1 and BC1) and all these ORFs are highly conserved with respect to location, orientation and composition (Howarth *et al.*, 1985; Stanley, 1985) (Fig. 2.1). DNA-A component encodes all the information necessary for viral replication and encapsidation (Rogers *et al.*, 1980; Sunter *et al.*, 1990; Bisaro, 1996; Hanley-Bowdoin *et al.*, 1999). In contrast the DNA-B component cannot replicate in the absence of DNA-A, but is required for systemic movement and symptom production (Brough *et al.*, 1988; Etessami *et al.*, 1988; Von Arnim *et al.*, 1993; Sanderfoot *et al.*, 1995; Sung and Coutts, 1995; Gilbertson *et al.*, 1996). Transient replication studies in leaf disc or

protoplast also showed that DNA-A component can replicate autonomously while DNA-B cannot (Townsend *et al.*, 1996; Lazarowitz *et al.*, 1991).

Within both DNA-A and DNA-B there is a non-coding intergenic region (IR) (~200bp) which has a conserved nonamer sequence (5'-TAATATTAC-3') potentially capable of forming a hairpin secondary structure which is an essential *cis*-element required for DNA replication (Revington *et al.*, 1989) and which has been genetically identified as part of the geminivirus plus-strand origin of DNA replication (Stenger *et al.*, 1991; Lazarowitz *et al.*, 1992; Bisaro, 1996). This IR, which is similar within two DNA populations of a virus but different between the viruses, is characteristic of bipartite begomoviruses. Among the ORFs of DNA-A component, AC1 codes for a 39 kDa replication initiator protein (Rep) which is a prerequisite for the replication of both DNA-A and DNA-B (Elmer *et al.*, 1988). Rep has no polymerase activity but is an origin recognition protein (Fontes *et al.*, 1992) and has a site specific endonuclease activity which cleaves at TAATATTAC and ligation activity for plus-strand viral DNA (Laufs *et al.*, 1995; Orozco *et al.*, 1998).

AC1 negatively regulates its own transcription (Eagle *et al.*, 1994), and also has ATP/GTPase activity (Desbiez *et al.*, 1995). Another, very interesting function of Rep involves its proven ability, in mastreviruses and begomoviruses, to bind plant homologues of the mammalian retinoblastoma (Rb) protein. This is a cell cycle regulatory protein, and binding to Rep may switch cells from a quiescent state to pre-S phase, in which DNA replication machinery is produced (Gutierrez, 2000). Lazarowitz *et al.*, (1992) demonstrated sequence specific recognition of the replication origin by the AC1 encoded protein, thereby localizing the origin to an ~90 nucleotide segment, proximal side of the IR that includes the conserved geminiviral stem-loop structure and ~60 nucleotides of the 5' upstream sequence. AC2 and AC3 are the regulatory genes. The AC2 gene product (transcription activator protein or

TrAP) transactivates transcription of both coat protein gene (AV1) and movement gene BV1 (Sunter and Bisaro, 1991, 1992). AC3 gene product (replication enhancer protein or Ren) is necessary for normal accumulation of viral DNA in infected plants (Sunter *et al.*, 1990; Sung and Coutts, 1995; Bisaro, 1996). Disruption of AC3 ORF was shown to delay and ameliorate disease symptoms, which is correlated with reduced replication of DNA-B (Morris *et al.*, 1991). ORF AC4 overlaps with highly conserved ORF AC1 and the function of its gene product is not yet known.

In viral sense of DNA-A, AV1 codes for a 29kDa coat protein which has 85-90 per cent sequence homology within all whitefly transmitted geminiviruses except at the N-terminal region. N-terminal region seems to play a major role in insect transmission (Noris *et al.*, 1998). AV1 protein product also interacts with BV1 protein product and thus has a role in movement of the virus (Ingham *et al.*, 1995).

In DNA-B, two movement proteins BV1 and BC1 help in both short distance (cell to cell) and long distance movement (Noueiry *et al.*, 1994; Pascal *et al.*, 1994; Ingham *et al.*, 1995). BV1 is a nuclear shuttle protein that binds viral ssDNA and BC1 protein product drags BV1-ssDNA complex to cell periphery and helps in entering the viral nucleic acid into the adjacent uninfected cells and also enters sieve elements (Sanderfoot and Lazarowitz, 1995). Besides this, BC1 is also known to encode protein product responsible for disease symptoms (Pascal *et al.*, 1993) and BV1 to determine host range properties (Ingham and Lazarowitz., 1993). Ingham *et al.*, (1995) demonstrated that both BC1 and BV1 contribute to viral host range properties.

The monopartite begomoviruses

The single-component nature and the gene organization of the monopartite begomoviruses were discovered only recently, hence aspects concerning gene organization and function in these viruses have been studied only recently and new information is forthcoming. The gene organization of the monopartite WTGs has been

shown to be somewhat similar to that found in the bipartite WTGs except that all ORFs are arranged within a single component of slightly larger size. These viruses appear to have several genes analogous in size and function to those in the bipartite WTGs and several unique ORFs for which function(s) are not yet completely understood.

The monopartite WTGs have at least six predicted ORFs encoding polypeptides of greater than 9.5 to 10 kDa (Navot *et al.*, 1991; Dry *et al.*, 1993; Kheyr-Pour *et al.*, 1991) that encode products termed V1 and V2 (CP), transcribed in the rightward, clockwise direction (V=viral strand), whereas C1 (Rep), C2, C3, and C4 are transcribed in the leftward or counter clockwise direction (C = complementary strand). V2 (CP), C1 (Rep), C2 and C3 appear to be functionally analogous to AV1 (CP), AC1 (Rep), AC2, and AC3, respectively, in the bipartite counterparts (Bisaro 1996). In these viruses, a unique ORF, the pre-coat, designated V1 because it is the first rightward ORF, making the coat protein gene the second, and hence the designation V2 instead of AV1 as in the bipartite members (Fig. 2.1). The C4 and V1 gene products of the monopartite viruses appear to be unique and may not have analogous functional homologues in the bipartite members; instead, these two gene products are thought to be involved in movement, in contrast to the functional analogs, BC1 and BV1 in bipartite members (Brown, 1996).

In the monopartite WTGs, the CP is also thought essential for viral movement, whereas the CP of the bipartite viruses enhances viral movement but is not absolutely required for systemic spread of the virus. One prospective intermediate virus, the tomato yellow leaf curl virus from Thailand (TYLCV-TH) has two genomic components with a DNA-A component having a gene arrangement like the monopartite members and a DNA- B component arranged like those of bipartite members. The monopartite-like DNA-A genome alone can incite a viable systemic

infection of tomato in the absence of the DNA-B component (Rochester *et al.*, 1990). When the DNA-A and DNA-B components are co-inoculated to tomato, symptoms develop slightly more rapidly and initially appear somewhat more severe, however, within several days, symptoms in infected tomato are indistinguishable from those caused by DNA-A alone (Rochester *et al.*, 1990). Whether the DNA-B component enhances viral movement but is not absolutely required for systemic infection by TYLCV-TH because the virus has functional V1 and C4 genes on the DNA-A component as in other monopartite viruses has not been addressed experimentally (Brown, 1996).

GENOME ORGANIZATION OF COTTON LEAF CURL VIRUS

Recently, Briddon *et al.*, (2001) have identified the DNA components required for induction of cotton leaf curl disease. They have demonstrated that the causative agent of the disease is a complex between a novel single-stranded DNA molecule approximately 1350 nucleotides in length (termed DNA- β) and CLCuV DNA-A. Sequencing and molecular characterization have provided detailed information on the genome organization of CLCuV.

Briddon *et al.* (2000) obtained full-length clones of CLCuV DNA-A by cloning PCR amplification products produced using specific primers. The sequence consisted of 2722 nucleotides and is available in the EMBL, DDJB and GenBank nucleotide sequence databases under accession number AJ132430. Analysis of the sequence in both orientations identified 6 conserved (between Old World begomoviruses) ORFs with the capacity to encode proteins of predicted molecular mass of 9 kDa or more (Table 2.5). The positions and orientation of these ORFs are shown in Figure 2.2. The genome of CLCuV was homologous to the DNA-A genomic component of bipartite begomoviruses and maintained the typical genome

organization of begomoviruses originating from the Old World with six conserved ORFs (Bridson *et al.*, 2000).

Table 2.5. Predicted open reading frames of CLCuV (Bridson *et al.*, 2001)

ORF	Start	Stop	No. of amino acids	Predicted molecular weight (kDa)
V1(CP)	277	144	256	29.6
V2	117	443	109	12.7
C1	2581	1501	361	40.9
C2	1599	1150	150	17.3
C3	1454	1052	134	15.6
C4	2430	2149	94	10.7

Bridson *et al.*, (2001) have identified a novel molecule essential for symptom induction. CLCuV DNA- β , in conjunction with CLCuV DNA-A, represents the infectious unit responsible for cotton leaf curl disease and is the first agriculturally destructive disease conclusively attributable to a *Begomovirus* / DNA- β complex.

CLCuV DNA- β is approximately 1350 nucleotides, half the length of the *Begomovirus* component which was cloned from PCR amplification with nucleic acids extracted from CLCuV infected cotton originating from Pakistan utilizing primer pair V4570/V4571 (Bridson *et al.*, 2001). The complete nucleotide sequences of clones CLC β 01 (1351 nucleotides) and CLC β 02 (1349 nucleotides) were determined. These sequences are available in the EMBL, DDJB and GenBank databases under accession numbers AJ292769 and AJ298903, respectively. Nucleotide numbering for CLCuV DNA- β , as for geminivirus genomic components, proceeds from the 3'A in the nonanucleotide sequence TAATATTAC. This motif, which DNA- β shares with geminiviruses, forms the loop of a predicted stem loop structure which, for geminiviruses, contains the nick site for initiation of virion-sense

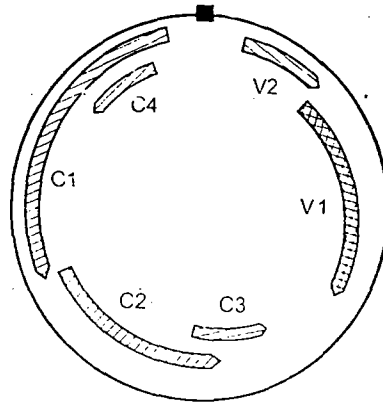


Figure 2.2. Genome map of CLCuV DNA-A showing the positions of predicted open reading frames (shown by arrows within the circle) encoded in both the virion-sense (left) and the complementary-sense (right) strands. (Adapted from Briddon *et al.*, 2000)

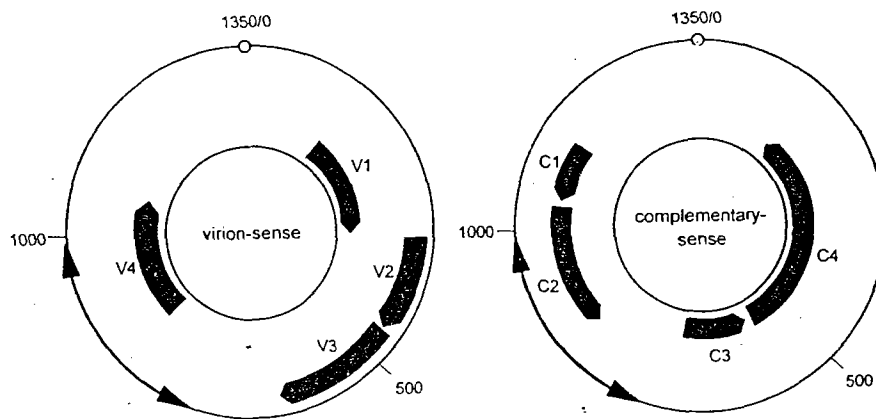


Figure 2.3. Genome map of CLCuV DNA- β showing the positions of predicted open reading frames (shown by arrows within the circle) encoded in both the virion-sense (left) and the complementary-sense (right) beginning in each case from an in-frame methionine codon. Only ORFs with a coding capacity above 4 kDa are shown. (Adapted from Briddon *et al.*, 2001)

DNA replication (Laufs *et al.*, 1995; Stanley, 1995). Analysis of the sequences of clones CLC β 01 and CLC β 02 identified four ORFs with a predicted coding capacity above 4 kDa in each strand (Fig.2.3; Table 2.6) (Briddon *et al.*, 2001).

Table 2.6 Features of CLCuV DNA- β ORFs (Briddon *et al.*, 2001)

ORF	No.of amino acids (CLC β 01/CLC β 02) (% amino acid similarity)*	Predicted molecular weight (kDa) (CLC β 01/CLC β 02)
V1	61/81 (98.4)	7.5/9.8
V2	39/40 (100)	4.3/4.4
V3	49/49 (93.8)	5.7/5.7
V4	65/65 (93.8)	7.5/7.5
C1	61/41 (92.7)	7.6/5.1
C2	49/49 (95.9)	5.8/5.8
C3	53/53 (94.3)	6.3/6.3
C4	118/118 (97.5)	13.7/13.7

* Percentage amino acid similarity between predicted products of clones CLC β 01 and CLC β 02

The sequences of CLCuV DNA- β contain an A-rich region between nucleotides 766 and 984 (58%A; Fig.2.3). CLCuV DNA- β shows no significant similarity to sequences in databases, with the exception of AYVV DNA- β (Briddon *et al.*, 2001). CLCuV DNA- β shares very little sequence homology with CLCuV DNA-A, although it requires the CLCuV DNA-A for trans-replication and encapsidation as evidenced by the ability of *B. tabaci* to acquire DNA- β from plants co-infected with CLCuV DNA-A and DNA- β . Thus, DNA- β molecules should be seen, taxonomically, as DNA satellites of begomoviruses rather than true virus components (Saunders, *et al.*, 2000; Briddon *et al.*, 2001). The results of the analysis of plants infected with CLCuV DNA-A alone and in conjunction with CLCuV DNA- β , showing the presence of CLCuV DNA- β leading to an increase in the levels of viral nucleic acids shows

that the product(s) encoded by DNA- β may be involved in cell-to-cell movement and act as a replication enhancer (Briddon *et al.*, 2001). Mansoor *et al.*, (2001) has demonstrated by mutagenesis experiments that β C1 gene is essential for symptom induction in cotton. The expression of β C1 from PVX vector developed strong chlorotic and necrotic symptoms on *Nicotiana benthamiana*, suggesting that β C1 encodes a pathogeneticity determinant. However, the PVX vector expressing β C1 was unable to develop systemic infection in *Nicotiana tabacum* var. Samsun, suggesting that β C1 is a target of host defense responses (Mansoor *et al.*, 2001). The PVX vector expressing β C1 was assayed for the suppression of post-transcriptional gene silencing (PTGS) of a transgene-green fluorescent protein (GFP) initiated either by ectopic expression of the transgene or the expression of part of the GFP sequence from PVX vector. β C1 suppressed PTGS of GFP initiated either by the transgene or by a gene silencing signal. These results showed that the pathogenicity determinant encoded by β C1 is a suppressor of PTGS (Mansoor *et al.*, 2001). It has been shown recently that the 25K protein of PVX, which is essential for the movement of PVX, blocks the gene silencing signal. The deletion of 25K gene inhibits virus (PVX) movement. The 25K gene was deleted from the PVX vector expressing β C1. The mutant was systemically infectious to *N. benthamiana* and developed symptoms of PVX expressing β C1. These results showed that β C1 complements the function of 25K gene which itself is a suppressor of gene silencing (Mansoor *et al.*, 2001).

PHYLOGENETIC RELATIONSHIP, VARIABILITY AND EVOLUTION

Phylogenetic relationships between different begomoviruses have been studied by comparing either nucleotide or amino acid sequence. The analysis showed that the relationship is closest among viruses from the same geographical area (Rybacki, 1994; Hong and Harrison, 1995; Padidam *et al.*, 1995; Padidam *et al.*, 1999). Begomoviruses of a specific geographical region shares more sequence homology

amongst them, than with begomoviruses of different geographical region, infecting the same host. Region specific selection pressure are operating more efficiently than host range determinants. Based on amino acid comparison of CP, the phylogenetic tree of begomoviruses can be divided into two major branches, containing old world and new world viruses, respectively. Within them there are clusters depending on the region where they occur (Harrison and Robinson, 1999).

The genome of CLCuV DNA-A has features typical of the old world begomoviruses, including a V2 gene, the function of which remains obscure. Phylogenetic tree derived from complete nucleotide sequences of the genomes (or DNA-A genomic components) of cotton leaf curl viruses and other selected begomoviruses shows a clear geographic relationship; viruses originating from the Indian subcontinent being more closely related to each other than to viruses from other regions (Fig. 2.4). It would seem probable, therefore, that CLCuV originated in this region and was not introduced from elsewhere (Briddon *et al.*, 2000; Briddon and Markham, 2000).

Investigations on begomoviruses infecting agriculturally important crops, all over the world, revealed that rich diversity exist among the *Begomovirus* populations (Harrison and Robinson, 1999). These variation could be due to mutations which are amplified by various secondary genome alterations brought about by acquisition of additional DNA molecules, recombination or pseudorecombination. These amplifying mechanisms are greatly enhanced by multiple infections of various viruses/strains in the same host and assortment of genomic components inside the vectors. (Harrison and Robinson, 1999). Recombination between two distinct begomoviruses i.e., interspecific recombination like okra yellow vein mosaic virus (OYVMV) and cotton leaf curl virus (CLCuV) is observed more commonly in Pakistan (Zhou *et al.*, 1998). The results of Zhou *et al.*, (1998) revealed the existence of four distinct forms of

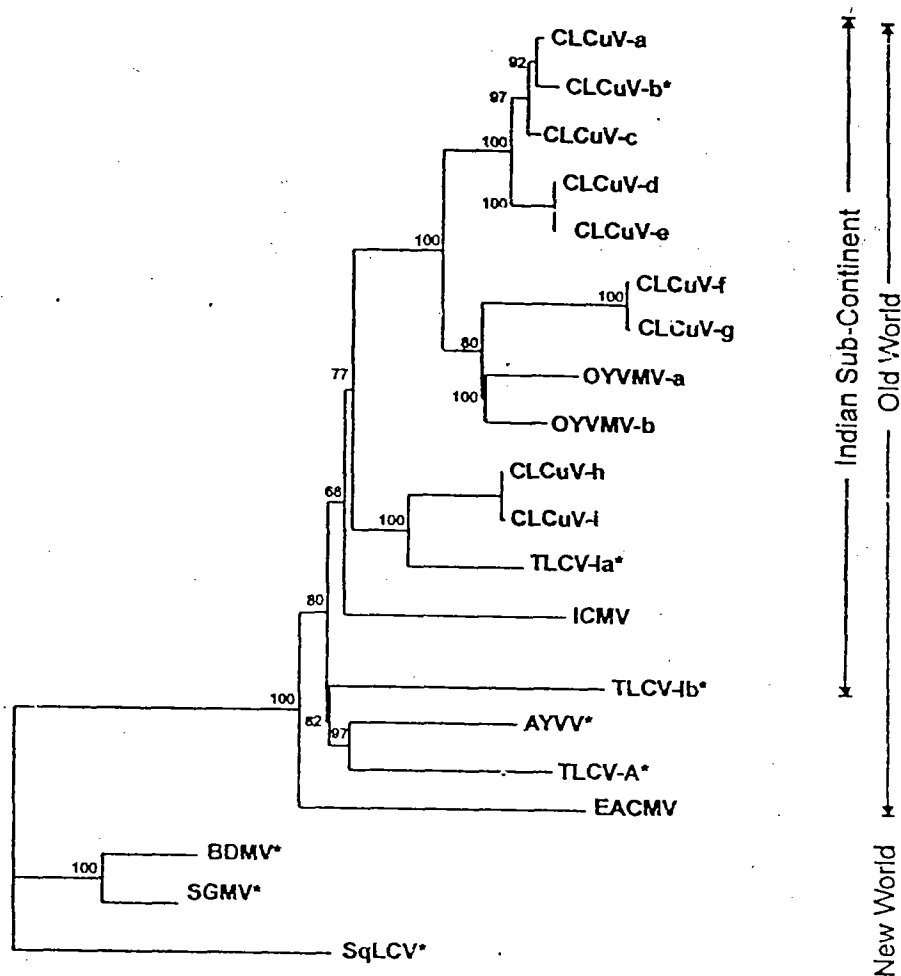


Figure 2.4. Phylogenetic dendrogram based on alignments of the complete nucleotide sequences of the genomes (or DNA-A genomic components) of *Cotton leaf curl virus* (CLCuV) and other selected begomoviruses. The geographic origins of the viruses are indicated on the right. The viruses are CLCuV, *Okra yellow vein mosaic virus* (OYVMV), *Tomato leaf curl virus* (TLCV), *Indian cassava mosaic virus* (ICMV), *Ageratum yellow vein virus* (AYVV), *East African cassava mosaic virus* (EACMV), *Bean dwarf mosaic virus* (BDMV), *Sida golden mosaic virus* (SGMV) and *Squash leaf curl virus* (SqLCV). Suffix (a-i) denotes different isolates of viruses. Vertical distances are arbitrary, horizontal distances are proportional to calculated mutation distances. Numbers at nodes indicate percentage bootstrap scores (100 replicates). (Adapted from Briddan and Markham, 2000)

CLCuV-PK, and have provided evidence that recombination involving virus isolates associated with okra yellow vein mosaic disease has played a role in the genesis of two of the CLCuV-PK variants (CLCuV-PK types clc 804a and clc 72b). The chance evolution of virus variants with increased biological fitness may lead to sudden outbreak of epiphytotics, as documented in the evolution of Ugandan isolate of cassava mosaic virus [the Uganda variant (UgV)] (Zhou *et al.*, 1997). This virulent Ugandan isolate of cassava mosaic virus (UgV) has arisen by recombination of DNA-A of East African cassava mosaic virus (EACMV) and DNA-A of African cassava mosaic virus (ACMV) (Zhou *et al.*, 1997). Recombination seems to have been the key to the development of the cotton leaf curl epidemic, as it has too in the cassava mosaic epidemic, spreading through Uganda and into neighbouring countries (Zhou *et al.*, 1998; Zhou *et al.*, 1997). Such epiphytotics in turn generate additional recombinants and other kinds of variants, especially where the vector has a range of hosts in which a variety of mixed infections by different begomoviruses can occur. The host genotypes, environmental conditions and agricultural practices constantly influence the interplay of host-virus-vector relationships and putting constant selection pressure on the genome of virus, which is acting as the driving force behind the rapid evolution of begomoviruses.

AGROINOCULATION

The Ti plasmid of *Agrobacterium tumefaciens* delimited by border sequences of 25 bp long is transferred to the plant cell where it gets integrated with the plant nuclear DNA (Chilton *et al.*, 1980). Any DNA between border sequences in *Agrobacterium* is also transferred to plants. The process of introduction of viral genome into plants mediated by the Ti plasmid of *Agrobacterium* was termed 'agroinfection' by Grimsley *et al.*, (1986). Elmer *et al.* (1988) suggested that agroinoculation rather than agroinfection was more accurate term for *Agrobacterium*

mediated delivery of virus genome into plant cell since the symptoms observed were mediated by the virus and not by *Agrobacterium* which acted merely as a delivery agent. Agroinoculation has been found to be an efficient and sensitive method of introduction of several viruses in plants (Table 2.7). For successful infection of plants by virus through agroinoculation require more than one copy of viral DNA in tandem dimeric orientation (Mandal *et al.*, 1997) but in the case of cassava latent virus (CLV) both monomeric and dimeric constructs were found to be infective (Morris *et al.*, 1988).

Agroinoculated plants produce systemic characteristic symptoms in 10-20 days with CLV (Morris *et al.*, 1988), 7 days TGMV (Elmer *et al.*, 1988), 14-21 days with TYLCV (Kheyr-Pour *et al.*, 1991), 13-18 days with MYMV (Mandal *et al.*, 1997), 21 days with PYMV (Buragohain *et al.*, 1994), 21 days with ToLCV (Tripathi, 2000), 10-14 days with AbMV (Wege *et al.*, 2000), 10-18 days with AYVV (Saunders *et al.*, 2000) and 18 to 21 days with CLCuV (Briddon *et al.*, 2001).

The percentage of infection of agroinoculated plants varies from host to host and virus to virus. The mechanism by which viral genome infects plants by agroinoculation is not well understood. For the geminiviruses, two possible mechanisms have been suggested (Grimsley *et al.*, 1987; Elmer *et al.*, 1988).

(a) homologous recombination between tandem repeated genome produces a circular unit length double stranded viral DNA molecule, and (b) replication between the single stranded origin to release a single stranded circular molecule which can be converted to double stranded replicative forms.

The natural life cycle, replication, spread and symptom development of cloned viral DNAs of several begomoviruses have been well demonstrated through agroinoculation. Recently other alternative methods of introduction of viral genome into plant cells have also been demonstrated: bombardment of tungsten or gold

particles coated with the viral genomic DNA components by biolistic method e.g., for pepper haustico geminivirus (PHV) (Garzon-Tiznado *et al.*, 1993), ToLCV (Padidam *et al.*, 1995) and CLCuV (Briddon *et al.*, 2001).

Table 2.7 Viral infectivity system in begomoviruses

Virus	Test host	Reference
A) Agroinoculation		
Bean golden mosaic virus (BGMV)	<i>Phaseolus vulgaris</i> var. Top crop	Morinaga <i>et al.</i> , (1988)
Cassava latent virus (CLV)	<i>N. clevelandii</i> <i>N. benthamiana</i> <i>Lycopersicon esculentum</i>	Morris <i>et al.</i> , (1988)
Mungbean yellow mosaic virus (MYMV)	<i>Vigna mungo</i> <i>Vigna radiata</i> <i>Glycine max</i> <i>Cajanus cajan</i>	Mandal (1994) Mandal <i>et al.</i> , (1997) Chakraborty (1996)
Potato yellow mosaic virus (PYMV)	<i>N. benthamiana</i> <i>L. esculentum</i> <i>N. benthamiana</i>	Buragohain <i>et al.</i> , (1994) Sung and Coutts (1995)
African cassava mosaic virus (ACMV)	<i>N. benthamiana</i> <i>N. benthamiana</i>	Saunders and Stanley (1995) Morris <i>et al.</i> , (1991)
Squash leaf curl virus (SLCV)	<i>N. benthamiana</i>	Lazarowitz and Lazdins (1991)
Sida golden mosaic virus (SGMV)	<i>Phaseolus vulgaris</i> <i>L. esculentum</i>	Hofer <i>et al.</i> , (1997)
Tomato golden mosaic virus (TGMV)	<i>N. benthamiana</i>	Elmer <i>et al.</i> , (1988) Stenger <i>et al.</i> , (1992)
Tomato yellow leaf curl (TYLCV)	<i>N. benthamiana</i> <i>L. esculentum</i>	Kheyr-Pour <i>et al.</i> , (1991) Rochester <i>et al.</i> , (1990) Czosnek <i>et al.</i> , (1993) Noris <i>et al.</i> , (1994) Navot <i>et al.</i> , (1991)

Tomato leaf curl virus (ToLCV)	<i>N. benthamiana</i> <i>L. esculentum</i>	Dry <i>et al.</i> , (1993) Tripathi, (2000)
Abutilon mosaic virus (AbMV)	<i>Abutilon</i> <i>sellovianum</i>	Wege <i>et al.</i> , (2000)
South African Cassava mosaic virus (SACMV)	Cassava (<i>Manihot</i> <i>esculenta</i>)	Berrie <i>et al.</i> , (2001)
Ageratum yellow rein virus (AYVV)	<i>Ageratum</i> <i>conyzoides</i> , <i>N. benthamiana</i>	Saunders <i>et al.</i> , (2000)
Maize streak virus (MSV)	<i>Zea mays</i> cv. Jubilee	Schnippenkoetter <i>et al.</i> , (2001)
Panicum streak virus (Pan SV)	<i>Zea mays</i> cv. Jubilee	Schnippenkoetter <i>et al.</i> , (2001)
B) Biolistic method		
Tomato leaf curl virus (ToLCV)	<i>L. esculentum</i>	Padidam <i>et al.</i> , (1995)
African cassava mosaic virus (ACMV)	Cassava (<i>Manihot</i> <i>esculenta</i>)	Briddon <i>et al.</i> , (1998)
Cotton leaf curl virus (CLCuV)	Cotton (<i>Gossypium</i> <i>hirsutum</i> cv. S12), <i>N. benthamiana</i>	Briddon <i>et al.</i> , (2001)

HOST RANGE OF CLCuV

Range is defined here as those plant species that support virus infection, thus, the term is applied to plant species that are naturally occurring hosts, and to those that can be experimentally infected by various means of inoculation. The host ranges of WTGs are generally thought to be rather narrow in that they are often confined to one or several plant families. However, as additional viruses are discovered and better characterized, it is clear that geminivirus host range phenotypes span a range from narrow to quite broad.

Perusal of literature shows that host range studies has been well carried out for three cotton leaf curl virus isolates from different geographical regions viz., CLCuV-S (Sudan isolate) (Tarr, 1957; Nour and Nour, 1964; El Nur and Abu Salih, 1970), CLCuV-K (Karnataka isolate) (Nateshan *et al.*, 1996) and CLCuV-PK (Pakistan isolate) (Harrison *et al.*, 1997).

Experimental transmission trials have been conducted with the leaf curl agent by the cotton whitefly *B. tabaci* from infected Sakel cotton (*Gossypium barbadense*) in Sudan. (Tarr, 1957; Nour and Nour, 1964; El Nur and Abu Salih, 1970). Although leaf curl from cotton affects a large number of species within the Malvaceae, it nevertheless has a relatively narrow host range. Host species which have been identified include *Abutilon theophrasti*, *Althaea rosea*, *A. ficifolia*, *A. kurdica*, *A. nudiflora*, *A. pontica*, *A. sulphurea*, *Gossypium barbadense*, *G. hirsutum*, *Hibiscus cannabinus*, *H. esculentus*, *H. ficulneus*, *H. huegelii*, *H. trionum*, *H. sabdariffa*, *Lavatera cretica*, *Malva alcea*, *M. silvestris*, *M. moschata*, *Malvaviscus arboreus*, *Pavonia hastata*, *Sida acuta*, *Sida alba*, *S. cordifolia*, *S. rhombifolia*, *Abelmoschus esculentus* in Malvaceae and *Corchorus fascicularis* in Tiliaceae (Tarr, 1957; Nour and Nour, 1964; El Nur and Abu Salih, 1970). All these species showed varying degrees of vein thickening and often curling and reduction of the leaves. Canquil and Follin, 1983 reported *G. arboreum* and *G. herbaceum* as resistant to leaf curl. The first report of this disease was from Nigeria in 1912 on the native cotton species *G. peruvianum* and *G. vitifolia* (Farquharson, 1912).

Nateshan *et al.*, (1996) determined the host range of CLCuV isolate (CLCuV-K) from Karnataka, South India by whiteflies (*B. tabaci*) fed on infected *G. barbadense* plants. CLCuV-K was transmitted by *B. tabaci* to 24 species in six plant families which include Asteraceae (*Acanthospermum hispidum*, *Ageratum conyzoides*, *Cosmos bipinnata*, *Dahlia sp.*, *Sonchus brachyotis*, *Synedrella nodiflora* and *Zinnia*

elegans), Euphorbiaceae (*Croton bonplandianum* and *Phyllanthus niruri*), Fabaceae (*Phaseolus vulgaris*), Malvaceae (*Althaea rosea*, *Gossypium barbadense* and *G. hirsutum*), Oxalidaceae (*Oxalis latifolia*), Solanaceae (*Capsicum annuum*, *Datura stramonium*, *Lycopersicon esculentum*, *Nicandra physalodes*, *Nicotiana benthamiana*, *N. glutinosa*, *N. occidentalis*, *N. rustica*, *N. sylvestris* and *N. tabacum*). All infected species developed leaf curl, leaf cupping or leaf rolling and several produced enations or thickened veins. In their study, the following species remained symptom free – *Abelmoschus esculentus*, *Arachis hypogaea*, *Beta vulgaris*, *Cajanus cajan*, *Crotolaria juncea*, *Cucumis sativus*, *Daucus carota*, *Eleusine coracana*, *Helianthus annuus*, *Lablab purpureus*, *Macrotyloma uniflorum*, *Manihot esculenta*, *Paspalum scrobiculatum*, *Phaseolus lunatus*, *Raphanus sativus*, *Solanum melongena*, *Sorghum bicolor*, *Spinacia oleracea*, *Vigna mungo*, *V. radiata* and *V. unguiculata* (crop plants); *Petunia hybrida*, *Salvia officinalis*, *Tagetes erecta* (ornamental plants); *Amaranthus viridis*, *Chenopodium amaranticolor*, *C. quinoa*, *Datura metel*, *Euphorbia geniculata*, *Galinsoga parviflora*, *Malvastrum coromandelianum*, *Parthenium hysterophorus* and *Physalis floridana* (weeds).

CLCuV-K infected all cultivars tested of *G. barbadense* and one of six cultivars of *G. hirsutum* but none of *G. arboreum* or *G. herbaceum* (Nateshan *et al.*, 1996). Based on the host range studies, CLCuV-K (Nateshan *et al.*, 1996) appears to have a host range wider than that reported for CLCuV-S (Tarr, 1957; Nour and Nour, 1964; El Nur and Abu Salih, 1970).

Similar host range studies were conducted for the Pakistan isolate of CLCuV (CLCuV-PK) by whiteflies fed on infected *G. hirsutum* plants (Harrison *et al.*, 1997). CLCuV-PK was transmitted by whiteflies (*B. tabaci*) to seven plant species, in three plant families, which include Fabaceae (*Phaseolus vulgaris*), Malvaceae (*Abelmoschus esculentus* and *Hibiscus cannabinus*) and Solanaceae (*Datura*

stramonium, *Lycopersicon esculentum*, *Nicotiana benthamiana* and *N. tabacum*) (Harrison *et al.*, 1997). In all species, symptoms consisted of vein thickening, and in some instances vein greening, together with slight or moderate leaf curling. Based on host range studies, CLCuV-PK resembled more like CLCuV -K (South Indian isolate) in infecting plant species outside the family Malvaceae, including some in the Solanaceae and Fabaceae (Harrison *et al.*, 1997).

In order to find out whether plant species other than cotton are naturally infected with CLCuV-PK, Harrison *et al.*, 1997 carried out serological studies based on epitope profiles of CLCuV-PK on a range of plants with virus like symptoms. They found that okra (*Abelmoschus esculentus*) with yellow vein symptoms, okra with leaf curl symptoms, *Hibiscus tiliaceus* with leaf curl symptoms, radish (*Raphanus sativus*) with leaf curl symptoms, sunflower (*Helianthus annuus*) with vein thickening symptoms, viruses found in China rose (*Hibiscus rosa-sinensis*) and hollyhock (*Althaea rosea*) had epitope profiles indistinguishable from CLCuV-PK, suggesting that these plants are putative natural hosts of CLCuV-PK. Mansoor *et al.*, (2000) confirmed by PCR and dot-blot hybridization that radish plants are infected with CLCuV naturally.

TRANSMISSION

The cotton leaf curl causal agent is transmissible naturally only by its whitefly vector *B. tabaci* and is not passaged by transovarial means (Golding, 1930; Kirkpatrick, 1930, 1931; Massey, 1932; Nour and Nour, 1964; El Nur and Abu Salih, 1970). It is not transmitted by aphids, flea beetles, thrips, jassids or mites (Kirkpatrick, 1931; Mustafa, 1995). The agent is transmitted by grafting, but not by seed or by mechanical inoculation with sap or by soil (Kirkpatrick, 1931; Tarr, 1951; Massey, 1932; Nour and Nour, 1964). Singh *et al.*, (1994) also demonstrated that none of the plants grown from the seeds of the CLCuV infected plants of LH900.

F846 and Pakistani Niab72 exhibited disease symptoms which indicated that CLCuV is not transmitted through the seed. Single *B. tabaci* are able to transmit the leaf curl agent (Kirkpatrick, 1931; Yassin and El Nur, 1970; Briddon and Markham, 2000), but greater transmission efficiency is observed when additional *B. tabaci* are present (Kirkpatrick, 1931; Cauquil and Follin, 1983). Kirkpatrick (1931) reported that immobile larval stages of *B. tabaci* were able to acquire the virus by feeding on infected leaves and then to transmit it as adults, with no further recourse to infected tissue.

In experimental studies, a 3hrs 20min acquisition access feeding (AAF) and 30 min inoculation access feeding (IAF) were required for transmission and the transmission threshold is reported as 6hrs 30min for CLCuV-S (Sudan isolate) (Kirkpatrick, 1930, 1931), whereas Nateshan *et al.*, (1996) reported that a minimum of 1h acquisition access period (AAP) and a minimum of 5 min inoculation access period (IAP) were required for transmission and the transmission threshold was 8hrs for CLCuV-K (Karnataka isolate). The whole process of acquiring the virus and infecting a healthy plant was accomplished in 6hrs 30min to 8hrs, suggesting that the virus is circulative and requires a relatively short latent period in the vector (Kirkpatrick, 1931; Nateshan *et al.*, 1996). If the AAP and IAP are increased up to 24hrs each, efficiency of virus transmission increases which is shown by the increased proportion of *G. barbadense* seedlings that became infected from 20% to 87% and 13% to 73%, respectively (Nateshan *et al.*, 1996).

Whiteflies that acquire the virus (CLCuV-S) remain infective for 12 days and probably for life, therefore a persistent, circulative relationship is postulated (Kirkpatrick, 1931), whereas Nateshan *et al.*, (1996) reported that CLCuV-K was transmitted for at least 9 days after *B. tabaci* left the virus source plants.

Nateshan *et al.*, (1996) also studied the transmission efficiency of male and female *B. tabaci*. They reported that female *B. tabaci* were more efficient in transmitting CLCuV-K, infecting 19/25 (76%) of the test plants than male *B. tabaci* in which 12/25 (48%) of the test plants were infected.

VECTOR ECOLOGY AND DISEASE EPIDEMIOLOGY

Until quite recently, WTGs were known primarily to infect non-cultivated plants in dry subtropical and tropical climates, with the few exceptions being diseases of cassava, tobacco, and tomato described primarily in the Old World. From the 1970's to the present, whitefly transmitted viruses were suspected to be involved in increasingly more prevalent diseases of crops, including cotton and vegetables that were grown in large monoculture systems in both the Old and New Worlds. The increasingly wider distribution and higher population levels of whitefly vector populations on a global scale can be directly implicated in the development of epidemics and incidental infections in crops throughout the world.

The whitefly, *Bemisia tabaci* (Genn.) is a member of the family Aleyrodidae (order: Hemiptera: sub-order: Homoptera). Members of the *Bemisia tabaci* (Genn.) complex, among which are the tobacco, cotton or sweet potato whitefly, and *B. argentifolli*, or the B-biotype of *B. tabaci*, are the only known whitefly vectors of begomoviruses (Brown, 1996). *B. tabaci* is readily capable of establishing to extreme population levels, particularly in crops grown under irrigated, arid conditions in both field and greenhouse systems. In addition, this whitefly has the potential to colonize a wide range of dicotyledonous species, among which are primarily vegetable and fiber species of great importance to worldwide agricultural production efforts.

The vector *B. tabaci* is one of the most formidable enemies in tropical agriculture. Although this cosmopolitan insect has a wide host range, besides cotton, tobacco and sweet potato, a large number of plants such as bhindi, cassava, tomato,

chillies, mungbean, pigeonpea, soybean and urdbean have been shown to be hosts in experimental studies. In nature, it has been observed on 142 plant species belonging to 23 different families (Saikia and Muniyappa, 1986) and 99 plant species belonging to 20 different families (Pimpale and Summanwar, 1984). From Pakistan, Ali *et al.*, (1995) reported 128 species as hosts of whitefly. Among vegetables brinjal was the most important host for carry over of whiteflies. Lantana and sunflower among ornamentals, *Achyranthus*, *Citrullus* and *Xanthium* among weeds, Mulberry among fruit and forest plants were most important host which help in maintaining whitefly population during off-cotton season in Pakistan.

From India, some of the weed plants recorded as hosts of whitefly are *Acanthospermum hispidum*, *Blainvillea rhomboiden*, *Euphorbia geniculata*, *Flovaria aurtralasica*, *Malvastrum caromandalianum*, *Perstophe hicalyculata*, *Physalis minima*, *Solanum nigrum*, *Trianthema portuloreastrum* and forty five plants served either as hosts or as food plants (Mariappan and Narayanasamy, 1986). Singh *et al.*, (1994) observed that whitefly remained active throughout the year in Punjab. Cotton varieties intercropped in citrus orchards exhibited more severe disease symptoms and higher incidence of its vector, *B. tabaci*. Potato, sunflower, cucurbits, tomato, chillies, rapeseed, okra, beans, mentha and brinjal were the important crops on which whitefly population developed during different months. Some weed plants like *Achyranthes aspera* var *porphyristachya*, *Xanthium strumarium*, *Mentha longifolia*, *Abutilon indicum*, *Sida spp.* *Amaranthus spinosus*, *Elusine indica*, *Euphorbia hirta*, *Datura metel*, *Solanum nigrum*, *Digera muricata*, *Convolvulus arvensis*, *Corchorus capsularis*, *Malva parviflora*, growing within or around cotton field also harboured adult population of whiteflies (Singh *et al.*, 1994). Host preference of whiteflies varies from region to region. In West Africa *B. tabaci* colonise on cowpea commonly but in

North India they have not been found to colonise in this host as other most preferable hosts are available (Varma, 1986).

The vector *B. tabaci* thrive best under humid and hot weather conditions. The insects survive in small number on bushy perennial plants like *Lantana camara* in subtropical areas during unfavourable winter and dry summer months. In tropical areas having no distinct winter, *B. tabaci* survives throughout the year with fluctuation of population depending on the prevailing weather conditions (Varma, 1986). Under semi-tropic, North Indian conditions minimum temperature and minimum relative humidity are the most important factors (Subrahmanyam, 1979; Varma and Subrahmanyam, 1986; Pimpale and Summanwar, 1986; Krishnareddy, 1989). Whereas under South Indian conditions (tropical) low rainfall, low humidity and high temperature are favourable (Murugesan *et al.*, 1977). The life span of the vector varies from season to season and host to host. Varma *et al.*, (1992) reported that about 11 days during summer and up to 83 days in winter are required to complete the life cycle of the vector and about 15 generations have been found in one year. Varma and Srivastava (1996) also reported that the life cycle of the vector was completed in 24-30 days in the months of January to March in West Bengal conditions, although the life span on different host varied. In West Bengal conditions whitefly could be found on 17 wild hosts belonging to widely different plant families and of which *Anisomeles ovata* has been found to be the most preferred host (Varma and Srivastava, 1996).

Recognition of the factors involved in the epidemiology of a certain virus may lead the way to efficient control of the disease by attacking the cycle at its weakest point. Lewin (1927) reported that the amount of damage caused by infection depended largely on the age of the plant when it was first infected. If young, it became stunted and bore little fruit. Further El Nur and Abu Salih (1966) and El Nur

(1967) reported that cotton plants were most susceptible at 2, 4 and 6 week old stages. Based on field observations (Ali *et al.*, 1995) it was noticed that if the disease appeared at late stage when the plants were aged and probably mature enough to tolerate the attack, the crop gave good yield even in the presence of leaf curling at the top. However, when most of the crop was attacked at young stage it suffered severely resulting in almost complete loss of the crop.

Date of sowing is another factor which plays a decisive role in disease development and its spread. The incidence of CLCuV was studied in susceptible cotton variety S12 planted on different dates i.e., mid April to end of June at 15 days interval (Ali *et al.*, 1995). The maximum incidence of disease was recorded as 99.7 per cent in 30th April planting. In this sowing date, an incidence of 94.8 per cent had been observed on 2nd of August after which the incidence increased at a slow rate and reached 99.2 per cent in the end of August and 99.7 per cent in the beginning of October. Maximum incidence percentage reached was 97.0 per cent in 15th April, 98.0% in 15th May, 90.0% in 30th May, 63.0% in 15th June and 37.9% in 30th June sowing. This study reveals that the incidence is more in early sown crop as compared to late sown crop. Observations on the progress of disease have shown that in general, the percentage of incidence increase slowly during the month of June, jumps up to 80% during July and August and after that the increase becomes slow. The rate of increase of the disease in the crop is also related to its degree of susceptibility (Idris, 1990). Monga *et al.*, (1998) however, after studying the incidence of CLCuV on ten promising germplasm lines noted that the progress of disease in general was maximum during the month of August as compared to July and September / October.

Further observations showed that rainy periods promote lush growth of weeds and crops on which whitefly population multiplies and which also serve as reservoirs of the virus for the infection of the young cotton crop (Bink, 1975). Early

multiplication of the vector, which reaches high population levels when the cotton is still young and highly susceptible, leads to early appearance of the disease. Further it was noted that richly fertilized crop provided environment congenial for the development of virus attack. Another observation has revealed that the disease was more near the orchards more often in one (wind) direction of the orchard and not in the other. These observations however need further confirmation with systematic experimentation (Narula *et al.*, 1999). Giha (1960) showed that spread of the CLCuD was always in the direction of the wind.

Epidemiology of vector transmitting the disease is equally important in disease development. A temperature range of 27-30°C is required for the whiteflies to complete their development in minimum time (Gerling *et al.*, 1986). Oviposition is impaired by rain and low humidity (under 60%) and temperature extremes. There are two main ways by which *B. tabaci* move (a) short active movement over distances measured in meters (Melamed – Madjar *et al.*, 1979; Fauquet *et al.*, 1986), (b) long distance passive movement controlled mainly by wind (Youngman *et al.*, 1986). The longest flight distance measured was 7 km. The main flight activity took place during morning hours but sometimes a second short peak was observed in the late afternoon (Fauquet *et al.*, 1986).

In Pakistan, whiteflies are usually a problem in the mid-to late season (August, September, October), when CLCuD is already well established. There is no data to indicate that there is significant secondary spread within cotton crops when whitefly numbers are high, late in the growing season (Briddon and Markham, 2000). Since CLCuV is not seed-borne, both the virus and insect vector must survive on alternative hosts. Interestingly, the native or 'Desi' cotton (*G. arboreum*), appears not to be susceptible to the disease (Briddon and Markham, 2000). There are no data to indicate the initial source of the infection in cotton each year. Sources may be weeds or the

surviving ratoon cotton infected in the previous season. Attempts by farmers to influence the annual onset of the disease by intensive weed control or removing (roguing) diseased cotton had no impact on infection rates. Most of the cotton grown in Pakistan is *G. hirsutum*, with approximately 3% *G. arboreum* and some *G. barbadense*.

The cotton varieties being used in Pakistan have played a major role in the epidemic. More recent cotton variety selections have increased tolerance and some appear immune (including CIM 1100, CIM448 and CIM443) although the use of immune varieties incurs some yield penalty (Briddon and Markham, 2000). A key issue during the early years of the epidemic was variety identification and seed quality, which was resolved by RFLP mapping and certification (Iqbal *et al.*, 1997).

G. barbadense was widely grown in the Sudan between 1930 and 1960 when it was susceptible to CLCuD (Tarr, 1951; Nour and Nour, 1964). The problem in Sudan was solved by selecting cotton varieties with a high level of tolerance to the disease (Nour and Nour, 1964).

MANAGEMENT OF COTTON LEAF CURL DISEASE

A variety of approaches have been tried to control the whitefly transmitted viral diseases (Varma, 1993; 1997). The management strategies mainly focused on methods to prevent initial infection and on the use of genetic resistance. For effective management of cotton leaf curl virus disease, development of resistant varieties, control of vector whiteflies and weeds harbouring the virus seem to be the most relevant approaches.

In order to study occurrence of cotton leaf curl virus disease in India, a survey was initially carried out by Sirsa Regional Station of Central Institute of Cotton Research, Nagpur, in cotton growing areas of Rajasthan and Punjab in July, 1994 (Anonymous, 1994). The disease was noticed in Sri Ganganagar, Abohar and Fazilka

areas. The varieties affected were F846, RST9, LH900, F1054, Niab 78, Ganganagar Ageti and Zhurar. Variety F846 appeared to be highly susceptible followed by RST9 and LH900 varieties. A committee was then constituted by Indian Council of Agricultural Research (ICAR) to study disease affected areas and make recommendations for containing the spread of leaf curl disease of cotton (Anonymous, 1994). The committee visited the affected areas in August, 1994 and made the following recommendations:

1. All the varieties of *G. hirsutum* cotton grown in the area (F846, RST9, Pakistani Narma, LH900, Ganganagar Ageti, F1054, F505, HS6 and Zhurar) were found to be affected by leaf curl disease although the incidence varied from 5 to 95 percent from place to place.
2. A distinct gradient in the incidence of the disease was observed. Maximum incidence and severity was observed in areas close to the international border while less and less disease was recorded as one moved away from the international border.
3. Based on the incidence of leaf curl disease in different locations it was observed that the varieties Pakistani Narma, F846, LH900, RST9, Zhurar, F505 and F1054 are highly susceptible to cotton leaf curl geminivirus.
4. Desi cotton (*G. arboreum*) varieties remained completely free from the disease.

Considering the severity of the disease in the major cotton belt of North-West India and the virulence of the pathogen which causes this disease the following recommendations were made by the committee:

A. Measures requiring immediate action

1. Chemical control of whitefly.
2. Identification of resistant and susceptible genotypes.
3. Rogueing of affected plants and weeds.

B. Long term measures to manage the leaf curl disease

As a long term measure, it has been suggested by ICAR Committee on Cotton Leaf Curl Disease, 1994, that the entire cotton growing belt in North-West Zone adjoining international border up to 20 km should grow only Desi cotton (*G. arboreum*) which will form buffer zone I, as Desi cotton varieties have been found free from leaf curl disease. After the buffer zone I of the *G. arboreum*, resistant varieties of *G. hirsutum* may also be allowed to grow in next 20 km as buffer zone II. This will substantially reduce the spread of leaf curl virus coming across from the North-West international border.

Monitoring of whitefly population in crops being grown in buffer zone I and II should be done and prevention of infestation by whitefly should be taken up as soon as it occurs so as to avoid migration of whitefly and spread of CLCuV to cotton growing areas.

Development of resistant varieties

A greater effort has been on developing plant varieties resistant to CLCuV. Most of the resistant varieties have been identified on the basis of field resistance. In Pakistan, out of 1200 world collections, resistance to CLCuV was noted on three varieties. These were LRA 5166 and CP-15/2 introduced from India and Cedix from USA (Ali *et al.*, 1995). Similarly, some the strains like CIM 435, CIM 443, CIM 448 and CIM 1100 developed by Central Cotton Research Institute in Pakistan are reported resistant and show no symptoms of virus attack (Ali *et al.*, 1995).

In India, programme of screening of breeding material and *G. hirsutum* germplasm has been undertaken at Agricultural Research Station, Sri Ganganagar, Punjab Agricultural University, Ludhiana, Chowdhury Charan Singh Haryana Agricultural University, Hisar and Central Institute of Cotton Research, Regional Station, Sirsa. Resistant breeding programme involving resistant varieties/strains and

local cultivars is also being undertaken in these institutions and is at various stages of development (Narula *et al.*, 1999). It was noted that variety RS875 developed at Agricultural Research Station, Sri Ganganagar and varieties LRA 5166 and LRK 516 developed at Central Institute for Cotton Research, Regional Station, Coimbatore did not show any symptoms of virus attack. Similarly hybrid LHH 144 developed at Punjab Agricultural University, Ludhiana also remained free from leaf curl disease at hot spots and also under artificial screening conditions (Narula *et al.*, 1999).

Singh *et al.*, (1997) screened 213 cotton entries (210 of *G. hirsutum* and 3 of *G. arboreum*) against CLCuV using its vector *B. tabaci* in Punjab. *G. arboreum* varieties, LD327, RG8 and RG18 were tolerant. *G. hirsutum* entries exhibited differential reaction towards disease as 84 entries remained free from disease symptoms.

Asad *et al.*, (2001) showed that expression of antisense RNA of viral genes confers resistance against cotton leaf curl disease in tobacco. Tobacco plants expressing antisense RNA of Rep, TrAP and Ren genes under the control of the 35S CaMV promoter, obtained by *Agrobacterium* mediated transformation showed a consistent resistance to CLCuD when screened for disease resistance by exposure to viruliferous whiteflies carrying CLCuV. They also reported that resistance was due to suppression of viral DNA replication and the resistance mediated by transgenes was inherited in a Mendelian fashion throughout a three-generation study.

Genetics of resistance

There are very limited studies to determine the genetics to WTGs in plants. Most of the cultivars resistant to viruses are genetically simple with resistance and susceptibility segregating in a typical Mendelian pattern involving a single locus. (Khetarpal *et al.*, 1990). The mechanism of resistance varies not only from virus to virus but also from host to host for the same virus (Varma, 1993). In the case of grains

legumes, conflicting evidences are available indicating that the resistance genes are dominant in some cases and recessive in others. The number of genes involved in resistance to MYMV appears to be one or two (Varma *et al.*, 1992). Similar observations have also been recorded in other crops. In most of the cases resistance to WTGs appears to be recessive (Varma, 1993). Dominance of susceptibility over resistance to WTGs is also reflected by the emergence of disease like cowpea golden mosaic, okra leaf curl and cotton leaf curl in India (Varma *et al.*, 1993; Varma *et al.*, 1997). The cotton leaf was first noticed in 1989 in susceptible *G. barbadense* cotton introduced from Sudan and okra leaf curl in 1990 in the germplasm introduced from West Africa (Varma *et al.*, 1997).

Nateshan *et al.*, (1996) studies the reactions of *Gossypium* species and genotypes against CLCuV-K. All five genotypes of *G. barbadense* and four out of five of *G. hirsutum* x *G. barbadense* developed symptoms. In contrast, only one (cv. Sharada) out of six cultivars *G. hirsutum*, and none of *G. arboreum* or *G. hubaceam* was infected. These results show that many sources of resistance to CLCuV-K occur in *Gossypium* germplasm, indicate that *G. barbadense* is the main source of susceptibility and suggest that susceptibility is dominant to resistance (Nateshan *et al.*, 1996).

Vector management

The timely application of chemicals, particularly granular systemic insecticides, have been found very effective not only in reducing vector population but also in checking the spread of WTGs. The success of this control measures depends on efficient monitoring of the vector and timely application of chemicals (Varma, 1976; Varma *et al.*, 1992). Usually insecticidal control is better for persistent and semi-persistent viruses such as WTGs, which require longer acquisition and inoculation feeding periods by their vectors (Varma, 1983). These insecticides are

shown to reduce the population of whiteflies in different crops from 50 to 99% and incidence of disease up to 70%, resulting in manifold increase of yield of crops like mungbean which are severely affected by MYMV (Varma *et al.*, 1992). Similar effect of insecticidal treatment has also been found in other crops. The increase in yield in various crops by treatment with insecticides may not be entirely due to the reduction in disease incidences but also through protection provided against other insect pests (Varma *et al.*, 1992). Some of the chemicals like aldicarb also appear to have hormonal effect in increasing yield (Varma *et al.*, 1997). Singh *et al.*, (1994) reported that Ethion 50 EC at 800ml and Triazophos 40 EC at 600 ml/acre were both effective against the vector.

Paecilomyces farinosus parasitizes *B. tabaci* (Nene, 1973). This fungus has been shown to reduce the population of vector on treated plants in field conditions (Varma *et al.*, 1992). Rashid and Khan (2000) evaluated the action of antagonistic microbes and bio-insecticides against CLCuV. They reported that *Paecilomyces lilacinus* decreased CLCuV disease severity on NIAB-78 and CIM-240 significantly, whereas *Verticillium chlamydosporium* did so on NIAB-78 only. Larvo BTK (a bio-insecticide) reduced CLCuV significantly on NIAB-78 and CIM-109 while Nimbokil 60 EC (a neem product) did so on MNH 93 and CIM-109. Khan *et al.* (2000) reported that foliar application of neem leaf extract and 1 % neem oil completely inhibited CLCuV transmission (100%) by whiteflies. Neem-seed extract, neem flavonoids (glycosides and aglycone, both at 3% concentration in water or ethanol) and 0.5% neem oil resulted in 80% reduction of CLCuV transmission.

Management through cultural practices

The manipulation of planting date to avoid peak vector population effectively reduces chances of infections. The incidence of CLCuV was studied in susceptible cotton variety S12 planted on different dates i.e., mid April to end of June at 15 days

interval in Pakistan (Ali *et al.*, 1995). They reported that incidence was more in early sown crop as compared to late sown crop.

Since, cotton is grown for only part of the year (and without ratooning), alternate weeds and cultivated hosts probably serve as virus reservoirs. Elimination of volunteer perennial cotton and suspected or known alternate hosts (weed reservoirs) is considered partially effective in reducing disease incidence (Tarr, 1964; Bink, 1975). The eradication of malvaceous weeds that may carry the virus is hardly worth attempting (El-Nur and Abu Salih, 1970). Similarly, attempts by farmers in Pakistan to influence the annual onset of the disease by intensive weed control or removing (roguing) diseased cotton had no impact on infection rates (Bridson and Markham, 2000). Quarantine legislation forbids the unauthorized importation of malvaceous planting material, and the growing of okra before 15th September in Sudan is strictly prohibited (El-Nur and Abu Salih, 1970). As okra is a staple food, this prohibition is not strictly followed and off-season okra is found in many localities (El-Nur and Abu Salih, 1970). The presence of wild okra in the Gezira, Sudan complicates the issue still further (El Nur and Abu Salih, 1970).

3. MATERIALS AND METHODS

Materials and methods for various experiments are given in chapter 4 along with the results for easy reference. Details of various materials and methods, not covered in chapter 4, have been included in Appendix 1 and 2.

4. RESULTS AND DISCUSSION

The results and discussion are presented in the form of research papers for easy reference.

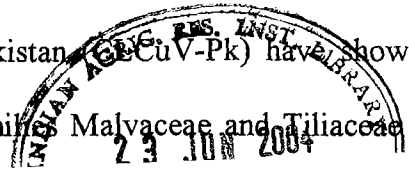
	CONTENTS	PAGE No.
4.1	Biological characterization of an isolate of cotton leaf curl virus from northern India and identification of sources of resistance.	53 – 59
4.2	Molecular characterization of cotton leaf curl virus from Rajasthan, India.	60 – 81
4.3	Association of novel DNA molecules with cotton leaf curl virus.	82 – 90
4.4	Molecular detection of DNA- β associated with whitefly transmitted geminiviruses in infected plants and viruliferous whiteflies.	91 – 101
4.5	Molecular variability among cotton leaf curl virus isolates in North India.	102 - 115

4.1 BIOLOGICAL CHARACTERIZATION OF AN ISOLATE OF COTTON LEAF CURL VIRUS FROM NORTHERN INDIA AND IDENTIFICATION OF SOURCES OF RESISTANCE

Leaf curl disease of cotton (CLCuD) was first reported from Nigeria on *Gossypium peruvianum* and *G. vitifolia* in 1912 (Farquharson, 1912) and since then it has caused serious economic losses in Sudan (Nour and Nour, 1964), Egypt (Smith, 1982) and Pakistan (Zhou *et al.*, 1998). In Pakistan, CLCuD was first observed in 1967 (Hussain and Ali, 1975) and it reached an epidemic proportion by 1987 (Hameed *et al.*, 1994). Since then the epidemic has covered about 1 million hectares of major cotton growing areas of Punjab and northern parts of Sindh in Pakistan (Zhou *et al.*, 1998).

In India, CLCuD was first observed in the experimental fields of Indian Agricultural Research Institute, New Delhi in 1989 in a few isolated plants of *G. barbadense* (Anonymous, 1990; Varma *et al.*, 1993). Later, in 1993, CLCuD appeared in Sri Ganganagar district of Rajasthan affecting *G. hirsutum* crops (Ajmera, 1994), and posed a major threat to cotton cultivation in northern India (Varma, 1994, 1995). The disease was shown to be caused by a whitefly transmitted geminivirus on the basis of positive reaction with polyclonal antiserum to Indian cassava mosaic virus (ICMV) in ELISA and immuno-sorbent electron microscopy (Varma *et al.*, 1993). Around the same time, this disease was also found in the kitchen gardens in Bangalore, Karnataka (Nateshan *et al.*, 1996). In 1997, the disease flared up in all the major cotton growing areas of Punjab, Haryana and Rajasthan in northern India infesting an area of about 0.2 million ha (Narula *et al.*, 1999).

Earlier studies on the biological properties of CLCuV isolates from Sudan (CLCuV-S), Karnataka, India (CLCuV-KI) and Pakistan (CLCuV-Pk) have shown that the host range of CLCuV-S is restricted to families Malvaceae and Tiliaceae (



Tarr, 1957; Nour and Nour, 1964; El-Nur and Abu Salih, 1970), whereas CLCuV-KI has a wider host range as it could also infect plants belonging to families Asteraceae, Euphorbiaceae, Fabaceae, Oxalidaceae and Solanaceae (Nateshan *et al.*, 1996), and CLCuV-Pk has intermediate host range as it could infect plants belonging to families Malvaceae, Solanaceae and Fabaceae (Harrison *et al.*, 1997). *G.arboreum* and *G.herbaceum* are not infected by any of these isolates (Cauquil and Follin, 1983; Nateshan *et al.*, 1996; Briddon and Markham, 2000).

In the present study whitefly transmission characteristics, symptomatology, host range and resistance in different genotypes of *Gossypium* spp. to isolate CLCuV-GI from Sri Ganganagar, India is reported.

MATERIALS AND METHODS

Virus isolate

CLCuV was isolated from a *G. hirsutum* plant showing severe disease symptoms at a farm in Sri Ganganagar (Rajasthan), and the isolate was designated as CLCuV-GI. It was maintained on *G. hirsutum* cv. LH 900 by whitefly (*Bemisia tabaci*) transmission in an insect-proof glass house.

Whitefly culture and transmission

A virus free colony of *B. tabaci* was maintained on *Nicotiana tabacum* cvs Xanthi and White Burley in a separate insect-proof glass house. For transmission, *B. tabaci* were given different acquisition access period (AAP) on CLCuV infected plants and transferred to cotton cv. LH 900 seedlings @ five whitefly per test seedling for the desired inoculation access period (IAP). After IAP the whiteflies were killed by spraying Decis 2.8 EC @ 0.5% and the test plants maintained in an insect proof glasshouse and observed for six weeks for symptom development. In all transmission experiments cotton seedlings at two fully expanded leaves stage were used.

Host range studies

For determining the host range, test plants were raised from seeds in an insect-proof glass house. Plants were inoculated at seedling stage using 10 whiteflies per seedling with AAP and IAP of 48h each. The inoculated plants were maintained in an insect-proof glass house and observed for symptom development for up to six weeks.

The test plants, which developed symptoms, were tested by back inoculation to cotton cv. LH 900 using 20 whiteflies per seedling with AAP and IAP of 48h each as described above, for confirming their infection by CLCuV and determining the ability of *B. tabaci* to acquire the virus from different hosts.

Identification of resistance in *Gossypium* spp. to CLCuV

126 genotypes of *G. hirsutum*, 19 of *G. barbadense*, 14 of *G. arboreum* and two of *G. herbaceum* were screened for resistance to CLCuV by whitefly inoculation. Ten seedlings of each line were inoculated using 10 whiteflies per seedling given AAP and IAP of 48h each and then the plants were kept in an insect-proof glass house for symptom development. The seeds of various genotypes of *Gossypium* spp. were obtained from IARI, New Delhi, Punjab Agricultural University, Ludhiana and Central Institute of Cotton Research, Nagpur. The evaluation of the reactions of these genotypes was made on the basis of per cent plant infection and graded as follows: no infection: highly resistant; up to 20% infection: resistant; 21 to 50% infection: moderately resistant; 51 to 70% infection: susceptible; 71 to 100% infection: highly susceptible.

RESULTS AND DISCUSSION

Whitefly transmission characteristics

CLCuV-GI was easily transmitted by whitefly (*B. tabaci*) to cotton cv. LH 900 producing typical symptoms of vein thickening, leaf curling and enations on the veins on the abaxial side of leaves, which frequently develop into large leaf-like out

growths. In severely infected plants all the leaves are curled, further growth is retarded and boll-bearing capacity is drastically reduced (Fig.4.1.1).

The virus could be transmitted by whitefly given a minimum AAP of 30 minutes and IAP of 10 minutes. The efficiency of transmission, however, increased with increase in AAP or IAP (Table 4.1.1). Best transmission was obtained when the whitefly were given AAP and IAP of 48 hrs.

The minimum acquisition and inoculation access periods required for CLCuV-GI differ slightly from that of CLCuV-KI which required a minimum AAP and IAP of 1h and 5 min, respectively (Nateshan *et al.*, 1996), and CLCuV-S, which required a minimum 3h 20 min AAP and 30 min IAP (Kirkpatrick, 1931). In transmission experiments reported here *G. hirsutum* seedlings were used whereas earlier workers (Nateshan *et al.*, 1996; Kirkpatrick, 1931) used *G. barbadense* seedlings. As *G. barbadense* is known to be more susceptible to CLCuV than *G. hirsutum*, more efficient transmission by whitefly is expected. But, on the contrary transmission was more efficient in the present experiments compared to the transmission of CLCuV-S (Kirkpatrick, 1931). These differences in transmission characteristics could be due to (a) greater virulence of CLCuV-GI, (b) differences in transmission efficiency of *B. tabaci* cultures, and (c) differences in the susceptibility of the test plants and their growing conditions.

Host range and symptomatology

CLCuV-GI was readily transmitted by *B. tabaci* to 9 species in 3 plant families (Malvaceae, Solanaceae and Fabaceae). The plant species which developed typical leaf-curl symptoms were *Abelmoschus esculentus*, *Althaea rosea*, *G. barbadense*, *G. hirsutum*, *Lycopersicon esculentum*, *Nicotiana benthamiana*, *N. tabacum*, *Phaseolus vulgaris* and *Physalis floridana* (Table 4.1.2). The host range of various isolates of CLCuV resembles each other. Like the present isolate other

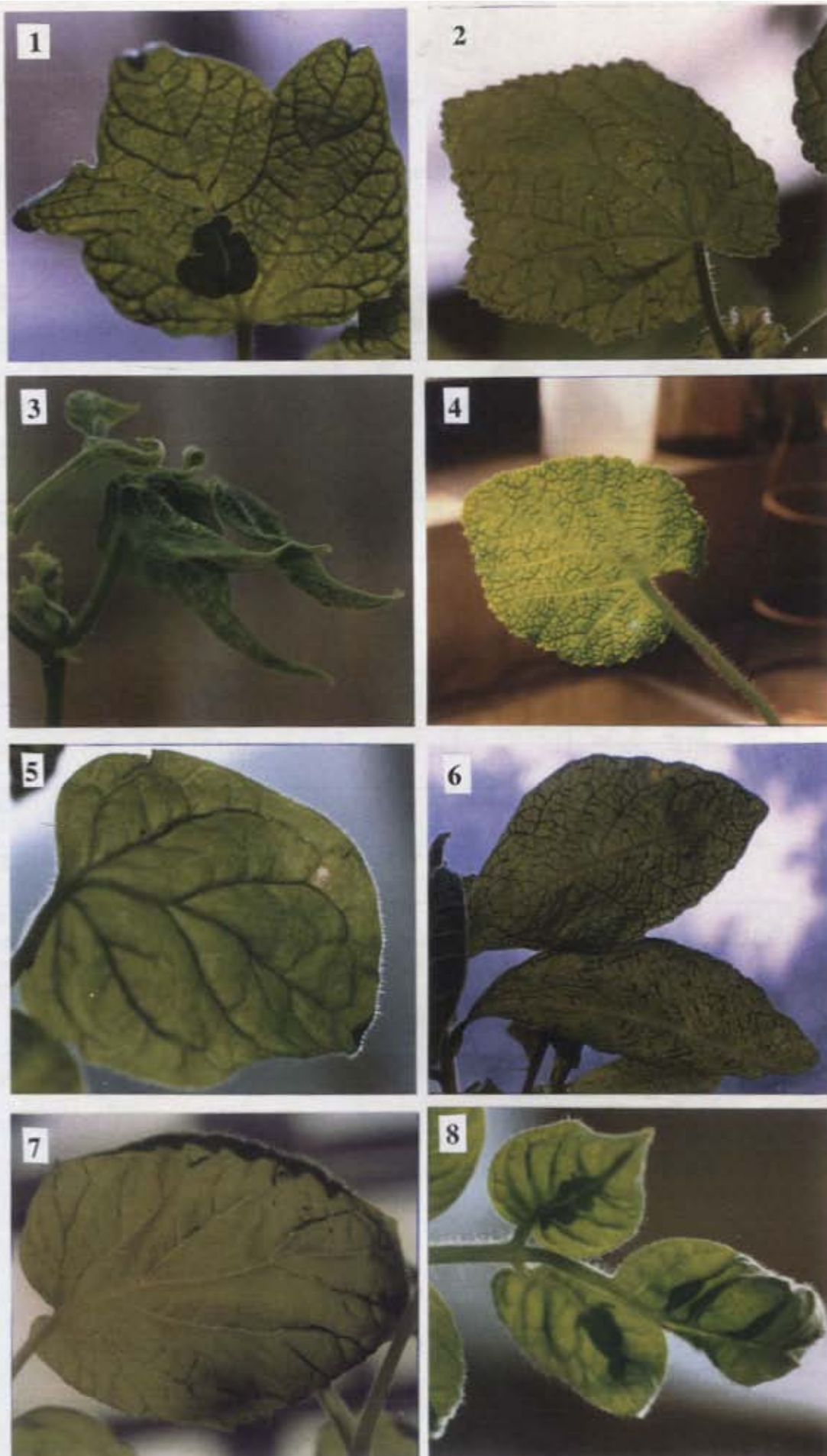


Figure 4.1.1 Symptoms produced in (1) *Gossypium hirsutum*, (2) *Abelmoschus esculentus*, (3) *Phaseolus vulgaris*, (4) *Althaea rosea*, (5) *Nicotiana benthamiana*, (6) *N. tabacum*, (7) *Physalis floridana* and (8) *Lycopersicon esculentum* through whitefly inoculation of CLCuV-GI

Table 4.1.1 Effect of acquisition and inoculation access periods on transmission of CLCuV-GI by *B. tabaci*

Experiment 1		Experiment 2	
AAP*	Transmission ⁺	IAP**	Transmission ⁺
5 min	0/10	5 min	0/10
10 min	0/10	10 min	1/10
15 min	0/10	15 min	3/10
30 min	1/10	30 min	3/10
1h	2/10	1h	4/10
3h	4/10	3h	5/10
6h	6/10	6h	5/10
12h	6/10	12h	7/10
24h	7/10	24h	8/10
48h	10/10	48h	10/10

* Inoculation access period was 48h

+ No. of plants infected / No. of plants inoculated. For inoculation @ 5 whiteflies per cotton cv. LH900 seedling were used.

** Acquisition access period was 48h

Table 4.1.2 Host range of cotton leaf curl virus (CLCuV-GI)

Host Plant	*No. of plants infected/No. of plants inoculated	Symptoms induced
1. Chenopodiaceae		
a. <i>Chenopodium amaranticolor</i>	0/10	----
2. Cucurbitaceae		
a. <i>Cucumis sativus</i> cv. Poinsette	0/5	----
b. <i>Lagenaria siceraria</i> cv. Pusa Naveen	0/5	----
c. <i>Luffa cylindrica</i> cv. Pusa Chikni	0/5	----
d. <i>Momordica charantia</i> cv. Pusa Do Mousami	0/5	----
3. Fabaceae		
a. <i>Cajanus cajan</i> cv. Pusa Upkar	0/5	----
b. <i>Glycine max</i> cv. Pusa 16	0/5	----
c. <i>Phaseolus vulgaris</i> cv. Sel. 9	1/5	Mild leaf curling and vein thickening
d. <i>Vigna mungo</i> cv. Local	0/5	----
e. <i>V. radiata</i> cv. Pusa Baisakhi	0/5	----
f. <i>V. unguiculata</i> cv. Pusa Komal	0/5	----

4. Malvaceae		
a. <i>Abelmoschus esculentus</i> cv. Parbhani Kranti	4/10	Mild leaf curling and vein thickening
b. <i>Althaea rosea</i> cv. Pusa Deepika	5/5	Mild leaf curling and vein thickening in all
cv. Pusa Kristina	5/5	
cv. Pusa Sweta	5/5	
c. <i>Gossypium arboreum</i> (line AKA-01-9)	0/10	----
d. <i>G. barbadense</i> (line GS 14-193)	10/10	Severe leaf curling, vein thickening and enations
e. <i>G. herbaceum</i> cv. Russian-5	0/10	----
f. <i>G. hirsutum</i> cv. LH 900	10/10	Leaf curling, vein thickening and enations
5. Solanaceae		
a. <i>Capsicum annuum</i> cv. Pusa Jwala	0/10	----
b. <i>Datura metel</i>	0/10	----
c. <i>Lycopersicon esculentum</i> cv. Pusa Ruby	10/20	Mild leaf curling, vein thickening and enations
d. <i>Nicotiana benthamiana</i>	8/10	Leaf curling and vein thickening
e. <i>N. tabacum</i> cv. White Burley	10/10	Leaf curling, vein thickening and enations
f. <i>Physalis floridana</i>	8/10	Mild leaf curling and vein thickening
g. <i>Solanum melongena</i> cv. Pusa Kranti	0/10	----

* Seedlings were inoculated with CLCuV-GI, using 10 whiteflies per seedling with AAP and IAP of 48h each.

isolates of CLCuV also did not infect *G. arboreum* and *G. herbaceum* (Cauquil and Follin, 1983; Nateshan *et al.*, 1996; Briddon and Markham, 2000). The reaction of the present isolate in various species of *Gossypium*, *A. esculentus* and *A. rosea* resembles that of CLCuV-S (Kirkpatrick, 1931; Tarr, 1951). Information on reaction of other hosts, found susceptible in the present investigation, to CLCuV-S is not available. The isolate from Pakistan (CLCuV-Pk) induced symptoms similar to those of the present isolate on *A. esculentus*, *Gossypium* spp., *L. esculentum*, *N. benthamiana* and *N. tabacum*. However the symptoms produced by CLCuV-Pk on *P. vulgaris* were milder compared to those produced by the present isolate. The host range of the present isolate differs from that of CLCuV-KI in its ability to infect *A. esculentus* and *P. floridana*, producing typical symptoms, whereas both these species were not susceptible to CLCuV-KI (Nateshan *et al.*, 1996). The present isolate also differs from CLCuV-KI, as it could not infect *Capsicum annuum* whereas CLCuV-KI causes vein clearing and leaf curl in this species. The characteristic symptoms produced by CLCuV-GI in tomato also differ from that of CLCuV-KI. CLCuV-GI induced typical enations which developed into large leaf-like out-growths on the abaxial side of leaves of tomato, whereas CLCuV-KI causes only leaf curl resembling the symptoms caused by tomato leaf curl virus (ToLCV), and no enation is developed. The symptoms produced by CLCuV-GI on *A. rosea*, *N. benthamiana*, *N. tabacum* and *P. vulgaris* were, however, similar to that produced by CLCuV-KI (Nateshan *et al.*, 1996). Various other species which were not infected by present isolate like *Cajanus cajan*, *Chenopodium amaranticolor*, *Cucumis sativus*, *Datura metel*, *Solanum melongena*, *Vigna mungo*, *Vigna radiata* and *Vigna unguiculata* were also not infected by CLCuV-KI. These hosts have not been tested with the other isolates (CLCuV-S and CLCuV-Pk).

The differences observed in host range and symptomatology of various isolates could be due to differences in genotypes and growing conditions used for different isolates. However, based on the limited information available on host range and symptomatology, CLCuV-GI appears distinct from CLCuV-KI and resembles CLCuV-Pk.

Efficiency of transmission of CLCuV from experimental hosts

CLCuV-GI was transmitted by whiteflies from various host species back to cotton cv. LH 900 in which typical vein thickening, enations and leaf curling symptoms developed (Table 4.1.3). The efficiency of transmission from different species, however, differed. Most efficient transmission was obtained from *N. tabacum*, followed by *L. esculentum*, *N. benthamiana* and *A. esculentus*. In North-West Indian cropping system, *A. esculentus* is grown in fields adjoining cotton fields, therefore, *A. esculentus* could serve as a potential source of inoculum of CLCuV. Similarly, *L. esculentum* could also be an important source of inoculum. *P. floridana* is a common weed in North-West India. Although transmission of CLCuV-GI from this host was not very efficient, it could be an important reservoir of the virus.

Identification of sources of resistance in *Gossypium* spp. to CLCuV

Of the 126 genotypes of *G. hirsutum* screened against CLCuV, 62 genotypes were found to be highly resistant, 21 resistant, 31 moderately resistant, 7 susceptible and 5 highly susceptible whereas all the 19 genotypes of *G. barbadense* screened were found to be highly susceptible. All the 14 genotypes of *G. arboreum* and two genotypes of *G. herbaceum* tested were highly resistant (Table 4.1.4).

All the genotypes of *G. barbadense* developed typical leaf curling, thickening of veins, formation of veinal-net and enations on the under surface of leaves; some plants also developed twisting of the petiole, peduncle and the youngest parts of the

Table 4.1.3 Back inoculation on cotton cv. LH 900

S.No.	Host	*No. of plants infected / No. of plants inoculated
1.	<i>A. esculentus</i>	2/3
2.	<i>A. rosea</i>	1/5
3.	<i>L. esculentum</i>	2/5
4.	<i>N. benthamiana</i>	2/3
5.	<i>N. tabacum</i>	4/5
6.	<i>P. vulgaris</i>	1/3
7.	<i>P. floridana</i>	1/5

* Acquisition access period of 48h on infected host plant and inoculation access period of 48h on cotton cv. LH 900. For inoculation @ 20 whiteflies per cotton cv. LH 900 seedling were used.

Table 4.1.4 Reaction of *Gossypium* species and genotypes to CLCuV-GI*

Grade	Species	No.	Genotypes
Highly Resistant	<i>G. arboreum</i>	14	AKA-01-1, AKA-01-2, AKA-01-3, AKA-01-4, AKA-01-5, AKA-01-6, AKA-01-7, AKA-01-8, AKA-01-9, AKA-01-10, AKH-8828, Desi-10NLLY/NLLY, RG 8, 30858 NLL ₄
	<i>G. barbadense</i>	Nil	----
	<i>G. herbaceum</i>	2	Baluchistan-7, Russian-5
	<i>G. hirsutum</i>	62	GH. 12-17-37 (SS), GH. 15-41-25 (SS), GH. 16-38 (SS), GH. 73-7-39-4 (S), RS 810, RS 875, RS 2013, 3x21-1-5, 8-6-11, 8-6-31-7-5, 8-6-31-9-2, 13-221, 13-237, 13-250, 14-218, 14-242, 14-257, 15BC1-1-2P ₁ (HS), 15 BC ₁ -1-2P ₁ , 15F ₂ -1-1-1-1-P ₁ (HS), 17-P ₅₂ -P ₁₀ , 31-38-105-3-2-P ₁ (SS), 31-38-105-5P ₁ (SS), 59, 63, 69, 95-2, 95-7, 95-9, 95-15, 95-16, 95-23, 95-27, 102-37 (S), 179-10-14-1, 180-1-1, 218, 230, 343, 350, 394, 411, 414, 422, 443, 444, 452, 463, 477, 533, 609, 616, 643, 658, 676, 702, 707, 802, 811, 814, 934, 4515
Resistant	<i>G. arboreum</i>	Nil	----
	<i>G. barbadense</i>	Nil	----
	<i>G. herbaceum</i>	Nil	----
	<i>G. hirsutum</i>	21	GH. 12-17-24 (S), GH. 12-17-47 (S), GH. 12-17-20 (SS), GH. 21-1-24 (SS), GH. 73-7-39-4 (SS), GH. 73-8 (SS), 13x2-1-1, 31-38-105-3P ₇ (SS), 31-38-105-5-P ₃ (SS), 31-38-107-3-1 (SS), 68, 95-10, 95-18, 95-22, 180-2-1, 180-2-2, -132-3-6 (SS), 342, 353, 388, 1118 (1118-37-1)
Moderately Resistant	<i>G. arboreum</i>	Nil	----
	<i>G. barbadense</i>	Nil	----
	<i>G. herbaceum</i>	Nil	----

Susceptible	<i>G.hirsutum</i>	31	GH. 12-17-37 SPs late, GH. 15-41-30 (SS), GH.73-E8 (SS), GH. 73-8 (S), GH.254 (SS), H1098, LH1556, Pusa 8-6, 13x2-1, 3x21-1-6, 14-237, 56, 57, 58, 61, 70, 95-1, 95-24, 95-29, -132-3-6 (S), 141, 316, 343 D, 693, 811, 816, 833, 840, 919, 935
	<i>G. arboreum</i>	Nil	----
	<i>G.barbadense</i>	Nil	----
	<i>G.herbaceum</i>	Nil	---- .
Highly Susceptible	<i>G.hirsutum</i>	7	F846, GH. 12-17-60 (SS), 31-38-1-P ₁ (SS), 95-28, 179-10-2-3, 480, 942
	<i>G. arboreum</i>	Nil	----
	<i>G.barbadense</i>	19	GB. 13-218, GB. 13-219(OP), GB. 13-223, GB. 13-225, GB. 13-226, GB. 14-193, GB. 14-199, GB.14-200, GB. 68-E.20 (SS), GB. 72-7-3-1 (SS), GB. 68 Br (SS), GB.68 Br ₂ (SS), GB. 68-E ₁ -P ₃ (SS), 13-217 (SS), 13-218 (SS), 13-219 (SS), 14-199 (SS), 14-200 (SS), 16-193 (SS)
	<i>G.herbaceum</i>	Nil	----
	<i>G.hirsutum</i>	5	GH. 73E. 8P ₂₃ (S), LH900, 468 (Br) (S), (321-1x11-152P ₁) P ₂₄ , 31-38-105-3-1-1 (SS)

*10 seedlings of each genotype were inoculated with CLCuV-GI, using 10 whiteflies per seedling with AAP and IAP of 48h each.

stem. The symptoms in *G. barbadense* were more severe than in susceptible *G. hirsutum* genotypes.

Nateshan *et al.*, (1996), also examined the reaction of all the *Gossypium* spp. against CLCuV, and found that all five genotypes of *G. barbadense*, and four out of five of *G. hirsutum* x *G. barbadense* interspecific hybrids were susceptible. Similarly, Singh *et al.* (1996) reported that all the five *G. hirsutum* x *G. barbadense* interspecific Israeli hybrids were highly susceptible to CLCuV and produced only shriveled bolls. In contrast, only one out of six cultivars of *G. hirsutum*, and none of *G. arboreum* (five cvs) or *G. herbaceum* (three cvs) or *G. arboreum* x *G. herbaceum* interspecific hybrids (three) were infected (Nateshan *et al.*, 1996).

The present findings show that good sources of resistance are available in *G. arboreum* and *G. herbaceum* as no infection was obtained in any of the genotypes tested. These species are also not susceptible to CLCuV-S, CLCuV-KI and CLCuV-Pk (Cauquil and Follin, 1983; Nateshan *et al.*, 1996; Briddon and Markham, 2000). Good degree of resistance has also been observed in *G. hirsutum*. Some of the genotypes like RS 875 of *G. hirsutum* and RG 8 of *G. arboreum* which have been found highly resistant to CLCuV in the present investigation were also reported to be resistant to CLCuV by earlier workers (Singh *et al.*, 1997; Narula *et al.*, 1999).

These results clearly demonstrate that good sources of resistance to CLCuV occur in *Gossypium* germplasm, and also indicate that *G. barbadense* is the main source of susceptibility, although, Fadlalla and Ibrahim (1999) reported *G. barbadense* cv. Barakat as resistant to CLCuV infection in Sudan. These resistances can be transferred to agronomically suitable varieties of cotton by conventional breeding. Further studies are also required to identify and isolate the resistance gene(s) from the identified resistant genotypes, so that resistance is transferred to agronomically suitable genotypes with greater ease.

4.2 MOLECULAR CHARACTERIZATION OF COTTON LEAF CURL VIRUS FROM RAJASTHAN, INDIA

Cotton leaf curl disease (CLCuD) has emerged as a serious constraint in the production of cotton in the north-western cotton belt of India (Varma, 1995). The disease is caused by whitefly (*Bemisia tabaci*) transmitted *Cotton leaf curl virus* (CLCuV) – a Begomovirus (Varma *et al.*, 1993; Mansoor *et al.*, 1993). The genus *Begomovirus* of the family *Geminiviridae*, comprise of whitefly transmitted viruses having typical geminate particles, encapsidating a single stranded circular DNA genome of approximate length of 2.5 to 2.7 kb (Harrison and Robinson, 1999). Most of the begomoviruses have bipartite genome, referred to as DNA-A and DNA-B (Stanley, 1983; Hanley-Bowdoin *et al.*, 1999). The DNA-A component has two open reading frames (ORFs) on viral sense, ORF AV1 (coat protein) and AV2, and four in the complementary sense referred as AC1 (replication initiation protein, Rep), AC2 (transcription activator protein, TrAP), AC3 (replication enhancer protein, REn) and AC4. Functions of ORFs AV2 and AC4 are not yet understood (Hanley-Bowdoin *et al.*, 1999; Harrison and Robinson, 1999). The second component, DNA-B has one ORF in viral sense and one in complementary sense. These ORFs code for proteins involved in movement within and between cells in host tissues (Noueiry *et al.*, 1994). Between the two genomic components of begomoviruses a region of ~180 to 250 nucleotides is common and highly specific for a virus. This common region has origin of replication (*Ori*) and regulatory sequences for both rightward and leftward open reading frames (Hanley-Bowdoin *et al.*, 1999; Harrison and Robinson, 1999).

Some begomoviruses, like *Tomato yellow leaf curl virus* (TYLCV; Navot *et al.*, 1991) have a single genomic component, equivalent to DNA-A of the bipartite viruses, and some other monopartite begomoviruses like the *Ageratum yellow vein virus* (AYVV; Saunders *et al.*, 2000), also require another component designated as

DNA- β , which is about a half of the size of DNA-A, to cause infection. DNA- β depends on DNA-A for replication and encapsidation. Like AYVV, CLCuV isolates, characterized so far, are also shown to be monopartite (Briddon *et al.*, 2000) and with one of the isolate (CLCuV-Pk; GenBank Acc.No.AJ132430) association of DNA- β has also been demonstrated (Briddon *et al.*, 2001). CLCuV DNA-A alone is infectious to cotton and *Nicotiana benthamiana* but does not induce typical leaf curl symptoms (Briddon *et al.*, 2000), unless the plants are coinoculated with CLCuV associated DNA- β (Briddon *et al.*, 2001). CLCuV complex (DNA-A and DNA- β), however, could not be transmitted by whiteflies from plants coinoculated by biolistic method (Briddon *et al.*, 2001).

Various isolates of CLCuV from Pakistan show considerable diversity in the sequences of DNA-A based on which three distinct species were proposed by Fauquet *et al.*(2000); two of these species were recognized earlier by the International Committee on Taxonomy of Viruses (ICTV; Rybicki *et al.*, 2000). One more species has been found in Sudan (GenBank Acc.No.AF155064) and Egypt (GenBank Acc.No.AY036010). Nomenclature of the geminivirus is still evolving. To avoid confusion the nomenclature used by ICTV will be followed in the present study which reports cloning, sequencing, agroinfection and vector transmission of agroprogeny of an isolate of CLCuV from Sri Ganganagar, Rajasthan, India, and genomic diversity among CLCuV.

MATERIALS AND METHODS

Virus isolate

CLCuV isolated from a *G. hirsutum* plant showing severe disease symptoms at a farm in Sri Ganganagar (Rajasthan, India), was maintained on *G. hirsutum* cv. LH 900 by whitefly (*Bemisia tabaci*) transmission in an insect-proof glass house. The isolate was designated as CLCuV-GI.

PCR cloning

a) Primer designing

A pair of abutting primers was designed, to amplify full-length DNA-A (~2.7kb) based on conserved sequences of CLCuV in Pakistan (GenBank Acc. No.- AJ002447, AJ002448, AJ002449, AJ002452, AJ002455 and AJ132430). These primers were designed to the C' terminal of AV1 gene (CPR primer) and C' terminal of AC3 gene (FV-AC primer).

The two primers used were:

1. CPR : 5'-TAATATCAATTCGTTACAGAG-3' (21 nucleotides)
2. FV-AC : 5'-ATAAAGTTTGAATTTTAT TTC-3'. (21 nucleotides)

In addition the following primers were designed to amplify the DNA-A fragment from upstream of nonanucleotide sequence (CRC-AC primer) to C' terminal of AC3 gene (FV-AC primer).

1. CRC-AC.: 5'-TATTAAACGGATGGCCGCTTT-3' (21 nucleotides)
2. FV-AC : 5'-ATAAAGTTTGAATTTTAT TTC-3' (21 nucleotides)

CPF and CPR primers were designed on the basis of published nucleotide sequence of DNA-A of CLCuV, to amplify complete coat protein gene (~750 bp):

1. CPF: 5'-AATTATGTCGAAGCGAGCTGC-3' (21 nucleotides)
2. CPR: 5'-TAATATCAATTCGTTACAGAG-3' (21 nucleotides)

Similarly, the following set of abutting primers having a *KpnI* site at its 5' end were also designed to amplify the second genomic component of CLCuV i.e., CLCuV DNA-β, based on the published nucleotide sequence of CLCuV DNA-β of CLCuV-Pk (GenBank Acc. No.-AJ298903; Briddon *et al.*, 2001).

The two primers used were:

1. Beta 1: 5'-GGTACCACTACGCTACGCAGCAGCC-3' (25 nucleotides)
2. Beta 2: 5'-GGTACCTACCCTCCCAGGGGTACAC-3' (25 nucleotides)

All these primers were synthesized by the GENSET Singapore Biotech. Pvt. Ltd.

b) PCR amplification

For PCR amplification of CLCuV DNA-A and DNA- β components from infected cotton plants, total DNA was isolated from 100mg of leaf tissue using DNeasy Plant Mini Kit (Qiagen Inc., USA) according to manufacturer's instructions and 200ng of total DNA was taken for PCR amplification.

The PCR reaction mixture consisted of 2.5 U of *Taq* DNA polymerase (MBI Fermentas, USA), 5 μ l of 10x PCR buffer (without MgCl₂), 3 μ l of 25 mM MgCl₂, 1 μ l of 10 mM dNTP and 1 μ l of 100ng/ μ l primers (DNA-A primers) or (DNA- β primers). Total reaction volume was made up to 50 μ l using sterile distilled water.

PCR amplification protocol was standardized employing different temperatures of denaturation, annealing and extension based on T_m values of the primers. The successful PCR amplification of the CLCuV DNA-A fragments and CLCuV DNA- β was obtained in a PHC-3 (Techne) thermal cycler.

Amplification of full-length DNA-A (~2.7kb) and DNA- β (~1.4kb) was obtained with the following programme: first denaturation at 94°C for 2min, then 30 cycles of denaturation (94°C for 1min), annealing (55°C for 2min) and extension (72°C for 4min for DNA-A and 3min for DNA- β). At the end of the cycles, a final extension step at 72°C for 5 min and then the temperature was decreased to 4°C until the reaction mixtures were removed. PCR conditions and amplification programme for CRC-AC+FV-AC primers and CPF+CPR primers were followed as described for amplifying DNA- β .

PCR products were analyzed by electrophoresis in 1% agarose gel, stained with ethidium bromide (0.5 μ g/ml), compared with the marker 1kb DNA ladder (MBI Fermentas, USA) and photographed on a UV transilluminator.

c) Cloning

Individual amplicons from the cotton sample were eluted from the gel using Gel Extraction Kit (Qiagen Inc., USA) according to manufacturer's instructions. Cloning of CLCuV DNA-A and DNA- β components was done according to Sambrook *et al.* (1989). The gel eluted amplified DNA-A and DNA- β products were cloned into pUC18 and pGEM-T easy vector, respectively. *Escherichia coli* strain DH5 α was transformed with the recombinant plasmid by CaCl₂ method (Mandel and Higa, 1970).

Transformants containing recombinant plasmids were selected on Luria agar plates containing ampicillin (50 μ g/ml), X-gal (2%) and IPTG (100 mM) (Vieira and Messing, 1982). Presence of the inserts was confirmed by releasing the insert from the plasmid by using *Hind*III + *Sac*I restriction enzyme (double digestion) in the case of DNA-A clones and *Kpn*I restriction enzyme in the case of DNA- β clones. The restricted products were analyzed by electrophoresis in 1% agarose gel as described earlier.

Construction of tandem dimers of CLCuV-GI DNA-A and DNA- β

Tandem dimeric constructs of CLCuV-GI DNA-A and DNA- β were first cloned in pUC19 and pGEM-T easy vectors, respectively. Two copies of PCR amplified full-length DNA-A (2.7kb) and one PCR amplified 1.7kb DNA-A fragment (0.6mer) were ligated in tandem to give an insert size of ~7.1kb and cloned in pUC19 vector to give a 2.6mer clone.

For generating DNA- β dimeric constructs, DNA- β clone in pGEM-T easy vector was partially digested with *Kpn*I releasing DNA- β and DNA- β +vector fragments. These were gel eluted and religated to obtain dimeric construct in pGEM-

T easy vector having ~2.8kb insert. 2.6mer construct of DNA-A and dimeric construct of DNA- β were used for transforming *E.coli* DH5 α to get the required clones.

The tandem 2.6mer DNA-A insert (~7.1kb) was released from pUC19 vector using *Hind*III + *Sac*I restriction enzymes (double digestion) and from pGEM-T easy vector DNA- β tandem dimeric insert (2.0mer) was released by *Eco*RI restriction. These tandem dimeric inserts were then recloned into the binary vector pBin19; the DNA-A tandem 2.6mer insert was then directionally cloned at *Sac*I and *Hind*III site, whereas, DNA- β tandem dimeric insert was cloned at *Eco*RI site. *Escherichia coli* strain DH5 α was transformed with the recombinant plasmid by CaCl₂ method (Mandel and Higa, 1970). The transformants containing recombinant plasmids were selected on Luria agar plates containing kanamycin (50 μ g/ml), X-gal (2%) and IPTG (100 mM). Presence of the insert was confirmed by releasing the insert from the plasmid by using *Hind*III + *Sac*I restriction enzyme (double digestion) in the case of DNA-A clones and *Eco*RI restriction enzyme in the case of DNA- β clones. The restricted products were analyzed by electrophoresis in 1% agarose gel as described earlier.

Agroinoculation

For agroinoculation of plants, CLCuV-GI DNA-A and DNA- β tandem dimers in the binary vector pBin19 were introduced into EHA105 strain of *Agrobacterium tumefaciens* by triparental mating; pRK2013 was the helper plasmid. *Agrobacterium* containing CLCuV constructs were selected against kanamycin (50 μ g/ml) and rifampicin (50 μ g/ml). Cotton (*Gossypium hirsutum*) cv. LH900 and Frenchbean cv. Sel9 plants were inoculated by pricking the tender hypocotyl region near the cotyledon with a fine entomological needle at the sprouted seed stage by 48hrs old *Agrobacterium tumefaciens* culture containing the CLCuV DNA-A and DNA- β

constructs in equal proportion. The inoculated sprouted seeds were soaked in the *Agrobacterium tumefaciens* culture over-night before planting. The plants were maintained in a Conviron growth cabinet of the National Phytotron Facility, IARI, at 28-35°C, 80-90 RH and 22,000 lux, and the biosafety regulations of the Department of Biotechnology, Government of India were followed throughout the experiments.

Detection of CLCuV in agroinoculated plants

CLCuV DNA-A was detected in agroinoculated cotton and Frenchbean plants by PCR using CPF and CPR primers and DNA- β was detected using Beta1 and Beta2 primers. DNA extraction from plants, PCR conditions and amplification programme were followed as described earlier for amplifying DNA- β .

Transmission of CLCuV from agroinfected plants by whiteflies

Transmission of CLCuV agroprogeny from agroinoculated plants was carried out by whitefly in an insect-proof glasshouse of the containment facilities. A virus free colony of *B. tabaci*, maintained on *Nicotiana tabacum* cvs Xanthi and White Burley in a separate insect-proof glasshouse, was used in this experiment. The agroinoculated test plants, which developed symptoms, were tested by back inoculation to cotton (*Gossypium hirsutum* cv. LH900 and *G. barbadense* line GB14-193) using 20 whiteflies per seedling with acquisition access period (AAP) and inoculation access period (IAP) of 72hrs each. Cotton seedlings at two fully expanded leaves stage were used for whitefly inoculation. After IAP, the whiteflies were killed by spraying Decis 2.8 EC @ 0.5%. The inoculated plants were maintained in the containment facilities and observed for symptom development.

CLCuV DNA-A and DNA- β was detected from back inoculated cotton plants by PCR using coat protein and DNA- β primers, as described earlier.

Sequencing and sequence analysis

Complete DNA-A and DNA- β clones were sequenced by primer walking method using automated DNA sequencer (ABI Prism, Perkin Elmer, University of Delhi, South Campus, New Delhi). The sequence data were assembled and analysed using the software programme Bioedit version 5.0.9. Multiple nucleotide and putative amino acid sequence alignments were done in CLUSTALW(1.7) programme (<http://www.ebi.ac.uk/clustalw>). Comparisons were made with sequences obtained from GenBank nucleotide sequence database. Relationship dendrograms were constructed from the aligned sequences using the neighbour-joining method and bootstrap option of Treecon version 1.3b (500 x bootstrap replicates). Accession numbers of the geminivirus sequences reported and compared in the present investigation, geographical origin and other details are given in table 4.2.1.

RESULTS

Cloning of CLCuV DNA-A and DNA- β

Using abutting primers, CPR+FV-AC and Beta1 + Beta2, full length DNA-A (~2.7 kb) and DNA- β (~ 1.4 kb) were amplified consistently, from CLCuV infected cotton plants (Fig.4.2.1). The PCR amplicons were gel eluted and then, DNA-A (~2.7 kb) fragment was cloned into *Sma*I cut pUC18 and DNA- β (~1.4 kb) fragment was cloned in pGEM-T easy vector. In all, 17 clones of DNA-A and 36 clones of DNA- β were obtained. DNA-A and DNA- β clones were designated as pCLCA and pCLC β , respectively. Of these clones only one of DNA-A and one of DNA- β was selected at random for further studies. The restriction analysis pCLCA and pCLC β clones showed the presence of full-length DNA-A (~2.7 kb) and DNA-B (~1.4kb) inserts, respectively (Fig.4.2.2).

Table 4.2.1.Geminivirus sequences used in this study

Sl. No.	Acronym	Virus name	Geographic origin	GenBank Acc. No.*
1.	AbMV-Ger	Abutilon mosaic virus	Germany	XI5983
2.	ACMV-Cam	African cassava mosaic virus	Cameroon	AF112352
3.	AEV-Nep	Ageratum enation virus	Nepal	AJ437618
4.	AREV-Egy	Althaea rosea enation virus	Egypt	AF014881
5.	AYVV-Sg	Ageratum yellow vein virus	Singapore	X74516 AJ252072 (DNA-β)
6.	AYVV-Tw	Ageratum yellow vein virus	Taiwan	AF307861
7.	BCTV-Cal	Beet curly top virus	California	M24597
8.	BDMV-Col	Bean dwarf mosaic virus	Colombia	M88179
9.	BGMV-PR	Bean golden mosaic virus	Puerto Rico	M10070
10.	BYVMV-In	Bhendi yellow vein mosaic virus	Madurai, India	AF241479 AJ308425 (DNA-β)
11.	CLCuV-311Pk	Cotton leaf curl virus	Multan, Pakistan	AJ002459
12.	CLCuV-72bPk	Cotton leaf curl virus	Faisalabad, Pakistan	AJ002448
13.	CLCuV-802a Pk	Cotton leaf curl virus	Lahore, Pakistan	AJ002455
14.	CLCuV-[804a]	Cotton leaf curl virus	Alabad, Pakistan	AJ002452
15.	CLCuV-[Gez]	Cotton leaf curl virus	Gezira, Sudan	AF155064
16.	CLCuV-[Cot]	Cotton leaf curl virus	Gezira, Sudan	AF260241
17.	CLCuV-[Okr/EG]	Cotton leaf curl virus	Egypt	AY036010
18.	CLCuV-[Okr/Gez]	Cotton leaf curl virus	Gezira, Sudan	AY036006
19.	CLCuV-[Sida]	Cotton leaf curl virus	Gezira, Sudan	AY036007
20.	CLCuV-[Fai1]	Cotton leaf curl virus	Faisalabad, Pakistan	AJ496286
21.	CLCuV-[26]	Cotton leaf curl virus	Multan,	AJ002458

			Pakistan	
22.	CLCuV-[62]	Cotton leaf curl virus	Multan, Pakistan	AJ002447
23.	CLCuV-[Fai2]	Cotton leaf curl virus	Faisalabad, Pakistan	AJ496287
24.	CLCuV-[Mul]	Cotton leaf curl virus	Multan, Pakistan	AJ496461
25.	CLCuV-806b Pk	Cotton leaf curl virus	Faisalabad, Pakistan	AJ002449
26.	CLCuV-DI	Cotton leaf curl virus	Delhi, India	AY146962 (AC1) AF465618 (AV1) AY146960 (AV2) AY146956 (IR) AF534185 (DNA-β)
27.	CLCuV-GI	Cotton leaf curl virus	Sri Gangana- gar, India	AF363011 AY083590 (DNA-β)
28.	CLCuV-HI	Cotton leaf curl virus	Hisar, India	AY146964 (AC1) AF465620 (AV1) AY146959 (AV2) AY146955 (IR) AF534186 (DNA-β)
29.	CLCuV-LI	Cotton leaf curl virus	Ludhiana, India	AY146963 (AC1) AF465619 (AV1) AY146958 (AV2) AY146954 (IR) AF534187 (DNA-β)
30.	CLCuV-Pk	Cotton leaf curl virus	Pakistan	AJ132430
31.	CLCuV-Pk1	Cotton leaf curl virus	Pakistan	AJ292769 (DNA-β)
32.	CLCuV-Pk2	Cotton leaf curl virus	Pakistan	AJ298903 (DNA-β)
33.	CLCuV-SI	Cotton leaf curl virus	Sirsa, India	AY146961 (AC1) AF465621 (AV1) AY146957 (AV2) AY146953 (IR) AF534188 (DNA-β)
34.	ICMV	Indian cassava mosaic virus	India	NC001932

35.	MYMIV-Bg	Mungbean yellow mosaic virus	New Delhi, India	AF126406
36.	OEV-Egy	Okra enation virus	Egypt	NC002510
37.	OLCV-Sd	Okra leaf curl virus	Sudan	AF397217 (DNA-β)
38.	OYVMV-201 Pk	Okra yellow vein mosaic virus	Faisalabad, Pakistan	AJ002451
39.	OYVMV-301 Pk	Okra yellow vein mosaic virus	Multan, Pakistan	AJ002453
40.	PaLCV-In	Papaya leaf curl virus	Lucknow, India	Y15934
41.	SiGMV-Hon	Sida golden mosaic virus	Honduras	Y11099
42.	SLCV-Ch	Squash leaf curl virus	China	AB027465
43.	TGMV-Bra	Tomato golden mosaic virus	Brazil	K02029
44.	TLCV-Ch	Tobacco leaf curl virus	China	AF240674
45.	ToLCV-Ban 4	Tomato leaf curl virus	Bangalore, India	AF165098
46.	ToLCV-Ban II	Tomato leaf curl virus	Bangalore, India	U38239
47.	ToLCV-NDe	Tomato leaf curl virus	New Delhi, India	U15016
48.	TYLCV-Is	Tomato yellow leaf curl virus	Israel	NC001996

* <http://www.ncbi.nlm.nih.gov>

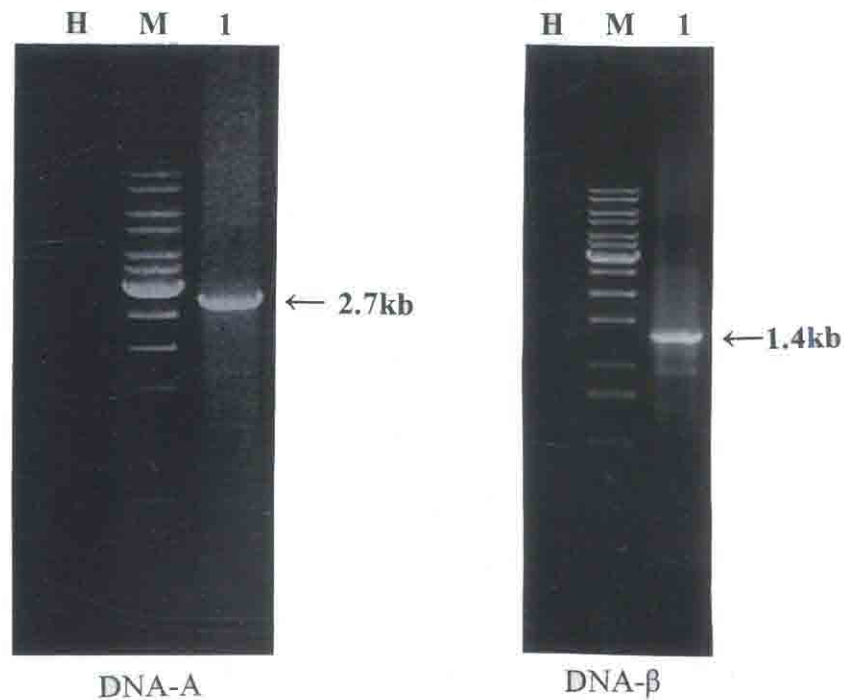


Figure 4.2.1. PCR amplification of CLCuV-GI DNA-A (~2.7kb) and DNA-β (~1.4kb) from total DNA of infected cotton cv. LH900 plants.
Lane 1: Infected cotton, 2: Healthy cotton control, M: 1kb DNA ladder

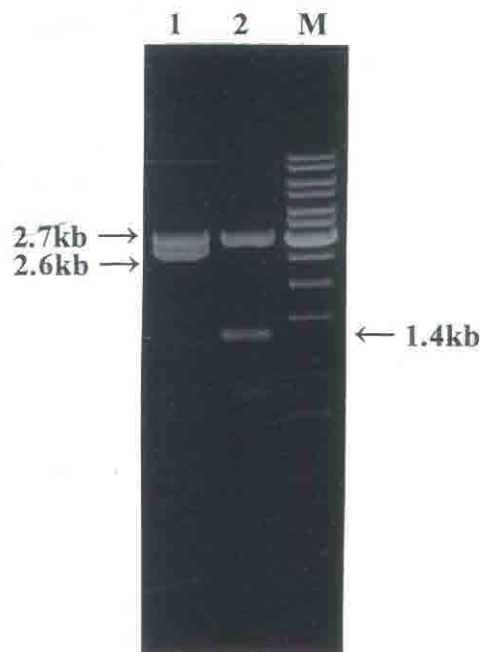


Figure 4.2.2. Agarose gel electrophoresis of CLCuV-GI DNA-A (~2.7kb) cloned in pUC18 restricted with *Hind*III+*Sac*I (lane 1) and DNA-β (~1.4kb) cloned in pGEM-T easy vector restricted with *Kpn*I (lane 2).
M: 1kb DNA ladder

Agroinfection

To prove the infectious nature of the cloned viral genomic components, agroinoculation was done on to cotton and Frenchbean plants. Tandem 2.6mer construct of DNA-A was released from pUC19 by digestion with *Hind*III and *Sac*I and recloned in pBin19 (Fig.4.2.3a). Tandem orientation of the full-length DNA-A in 2.6 mer clone was confirmed by restriction with *Kpn*I (Fig.4.2.4).

Similarly, tandem dimeric construct of DNA- β was released from pGEM-T by *Eco*RI and recloned in pBin19 (Fig.4.2.3b) and tandem orientation confirmed by restriction with *Ssp*I (Fig.4.2.5).

In all, 8 tandem 2.6mer clones of DNA-A and 22 tandem dimeric clones of DNA- β were obtained in pBin19. Of these, one clone each of DNA-A and DNA- β (designated as pBinCLCA and pBinCLC β) were selected at random. Both the clones were mobilized into *Agrobacterium tumefaciens* strain EHA105 by triparental mating using pRK2013 as helper plasmid, and used for infectivity tests.

Cotton (*Gossypium hirsutum*) cv. LH900 and Frenchbean cv. Sel.9 were inoculated with *A. tumefaciens* strain EHA105 containing tandem dimeric constructs of DNA-A (pBinCLCA) and DNA- β (pBinCLC β). Five plants of cotton cv. LH900 and 15 plants of Frenchbean cv. Sel.9 were inoculated. Four out of five cotton plants developed typical symptoms 24 to 27 days after inoculation (Table 4.2.2; Fig.4.2.6a,b) similar to those developed in plants by whitefly inoculation (Fig.4.2.6c). The agroinfected plants remained stunted and the leaves developed mild upward curling, vein swelling, vein darkening and enations (Fig.4.2.6a,b). All the agroinoculated plants of Frenchbean cv Sel.9 developed symptoms 11 to 12 days after inoculation (Table 4.2.2; Fig.4.2.7a.). The symptom first appeared as very mild mosaic on the first trifoliolate leaves; later severe mosaic, mottling and severe downward curling of all the leaves developed (Fig.4.2.7b.). Some leaves also

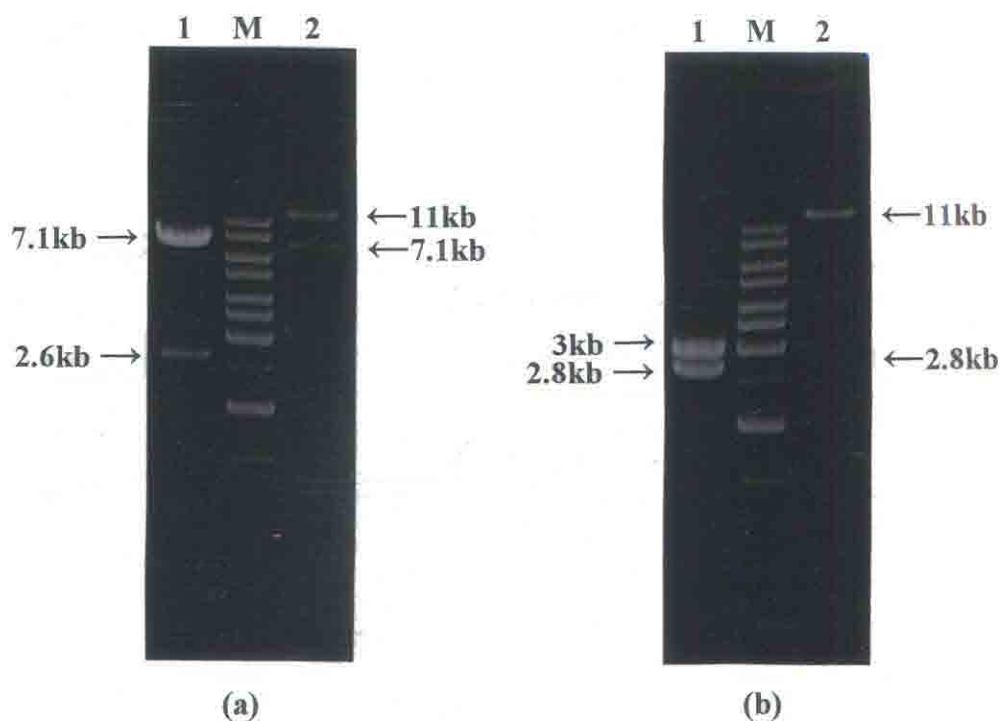
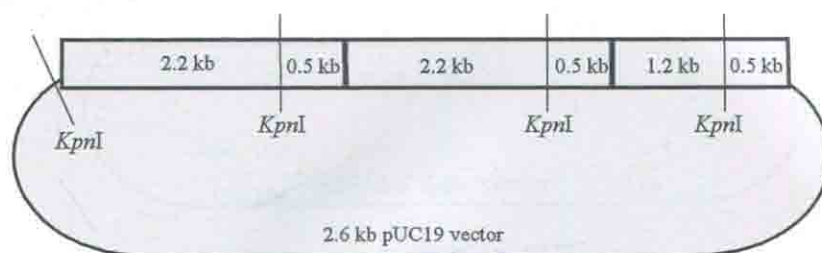
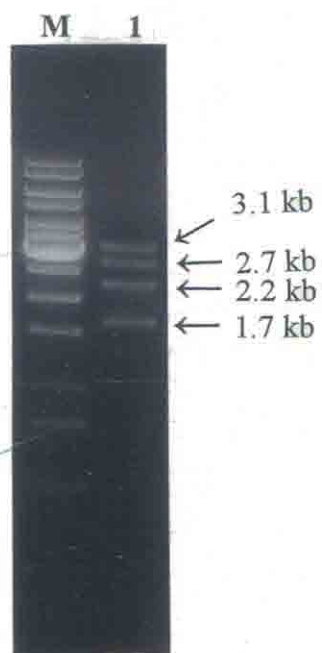


Figure 4.2.3. Agarose gel electrophoresis of tandem dimeric constructs of CLCuV-GI DNA-A (a) and DNA- β (b)

- (a) Tandem dimeric constructs of CLCuV-GI DNA-A (~7.1kb) in pUC19 (lane 1) and pBin19 (lane 2) released by restriction with *Hind*III+*Sac*I
 - (b) Tandem dimeric constructs of CLCuV-GI DNA- β (~2.8kb) in pGEM-T easy vector (lane 1) and pBin19 (lane 2) released by restriction with *Eco*RI
- M: 1kb DNA ladder

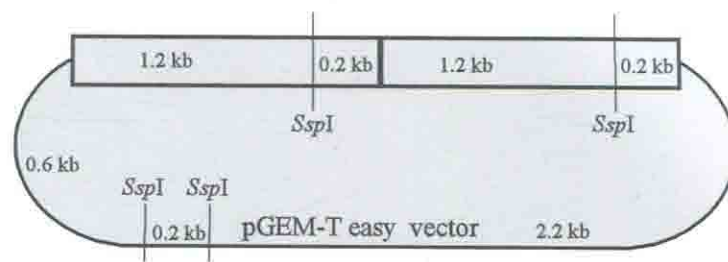


(a)

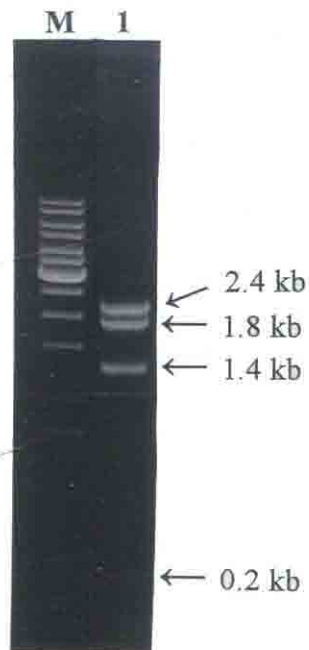


(b)

Figure 4.2.4. Diagrammatic representation of tandem 2.6mer DNA-A construct (box) in pUC19 vector (line) (a) and agarose gel electrophoresis of tandem 2.6mer DNA-A construct in pUC19 vector restricted with *Kpn*I (lane1) (b).
M: 1kb DNA ladder



(a)



(b)

Figure 4.2.5. Diagrammatic representation of tandem dimeric DNA- β construct (box) in pGEM-T easy vector (line) (a) and agarose gel electrophoresis of dimeric DNA- β construct in pGEM-T easy vector restricted with *Ssp*I (lane1) (b).
M: 1kb DNA ladder

(a)



(b)



(c)



Figure 4.2.6. Cotton (*G. hirsutum* cv. LH900) plants showing typical vein thickening (a) and enation symptoms (b) after agroinoculation with tandem dimeric constructs of CLCuV-GI DNA-A and DNA- β . For comparison, a leaf from CLCuV infected cotton (*G. hirsutum* cv. LH900) plant inoculated by whiteflies (c).

(a)



(b)



(c)



Figure 4.2.7. Frenchbean cv. Sel.9 plants showing severe downward leaf curl (a), mosaic and mottling (b) and vein thickening (c) symptoms after agroinoculation with tandem dimeric constructs of CLCuV-GI DNA- α and DNA- β .

Table 4.2.2. Infection of cotton and Frenchbean with the cloned DNA components (DNA-A + DNA- β) of CLCuV-GI using *A. tumefaciens* strain EHA105

S. No.	Host	No. of plants infected / No. of plants inoculated	Incubation period (days)	Symptoms* observed	PCR test DNA-A DNA- β (CP)
1.	<i>G. hirsutum</i> cv. LH900	4/5	24	MLC, VT, E	+ +
2.	Frenchbean cv. Sel.9	15/15	11	SMM, SDLC	+ +

* MLC : Mild leaf curl ; VT : Vein thickening ; E : Enation ; SMM : Severe mosaic and mottling ; SDLC : Severe downward leaf curl

Table 4.2.3. Whitefly transmission of CLCuV agropropy from agroinfected plants to healthy cotton plants.

Agroinfected plants used for acquisition	Test plants	No. of plants infected / No. of plants Inoculated*	Incubation period (days)	Symptoms observed ⁺	PCR test DNA-A DNA- β (CP)
<i>G. hirsutum</i> cv. LH900	<i>G. hirsutum</i> cv. LH900	4/10	24	MLC, VT, E	+ +
	<i>G. barbadense</i> line GB14-193	6/6	22	SLC, VT, E	+ +
Frenchbean cv. Sel.9	<i>G. hirsutum</i> cv. LH900	0/10	—	—	— —
	<i>G. barbadense</i> line GB14-193	0/6	—	—	— —

* Inoculation was done using 20 whiteflies per seedling with acquisition access period and inoculation access period of 72hrs each.

+ MLC : Mild leaf curl ; VT : Vein thickening ; E : Enation ; SLC : Severe leaf curl

developed irregular thickening of the veins (Fig.4.2.7c). Equal number of plants were also inoculated with 2.6mer clone of DNA-A or dimeric clone of DNA- β . None of these plants developed any symptoms for up to 60 days.

Agroinfection was confirmed by PCR detection of CLCuV DNA-A and DNA- β using CPF, CPR, Beta 1 and Beta 2 primers. CLCuV coat protein gene fragment (~ 750 bp) and full length DNA- β (~ 1.4 kb) was amplified from all the symptomatic plants of cotton cv LH900 and Frenchbean cv Sel.9 tested (Table 4.2.2; Fig.4.2.8), confirming infectious nature of the tandem dimeric constructs of DNA-A (pBinCLCA) and DNA- β (pBinCLC β). Viral DNA could not be detected by PCR in the agroinoculated asymptomatic cotton cv. LH900 and in the plants, which were agroinoculated using the DNA-A and DNA- β constructs individually.

Transmission of CLCuV from agroinfected plants

Whiteflies (*Bemisia tabaci*) could acquire CLCuV agroprogeny from agroinoculated symptomatic cotton cv. LH900 plants after 72hr of AAP and transmit the virus to healthy 7 days old cotton seedlings (*G. hirsutum* cv. LH900 and *G. barbadense* line GB14-193) with IAP of 72hr. The agroprogeny was more efficiently transmitted to *G. barbadense* than to *G. hirsutum* (Table 4.2.3). Typical leaf curl symptoms with vein swelling, vein darkening (Fig.4.2.9) and enations, which were indistinguishable from those associated with naturally infected plants, developed in both the species; the leaf curling, in *G. barbadense* were more severe (Fig.4.2.9b) than in *G. hirsutum*. The time taken for symptom development in the two species was almost similar (Table 4.2.3). CLCuV DNA-A and DNA- β were detected from back inoculated cotton plants by PCR using coat protein and DNA- β primers (Fig.4.2.10), as described earlier, confirming infection by CLCuV and demonstrating the ability of *B. tabaci* to acquire the virus agroprogeny from agroinfected plants.

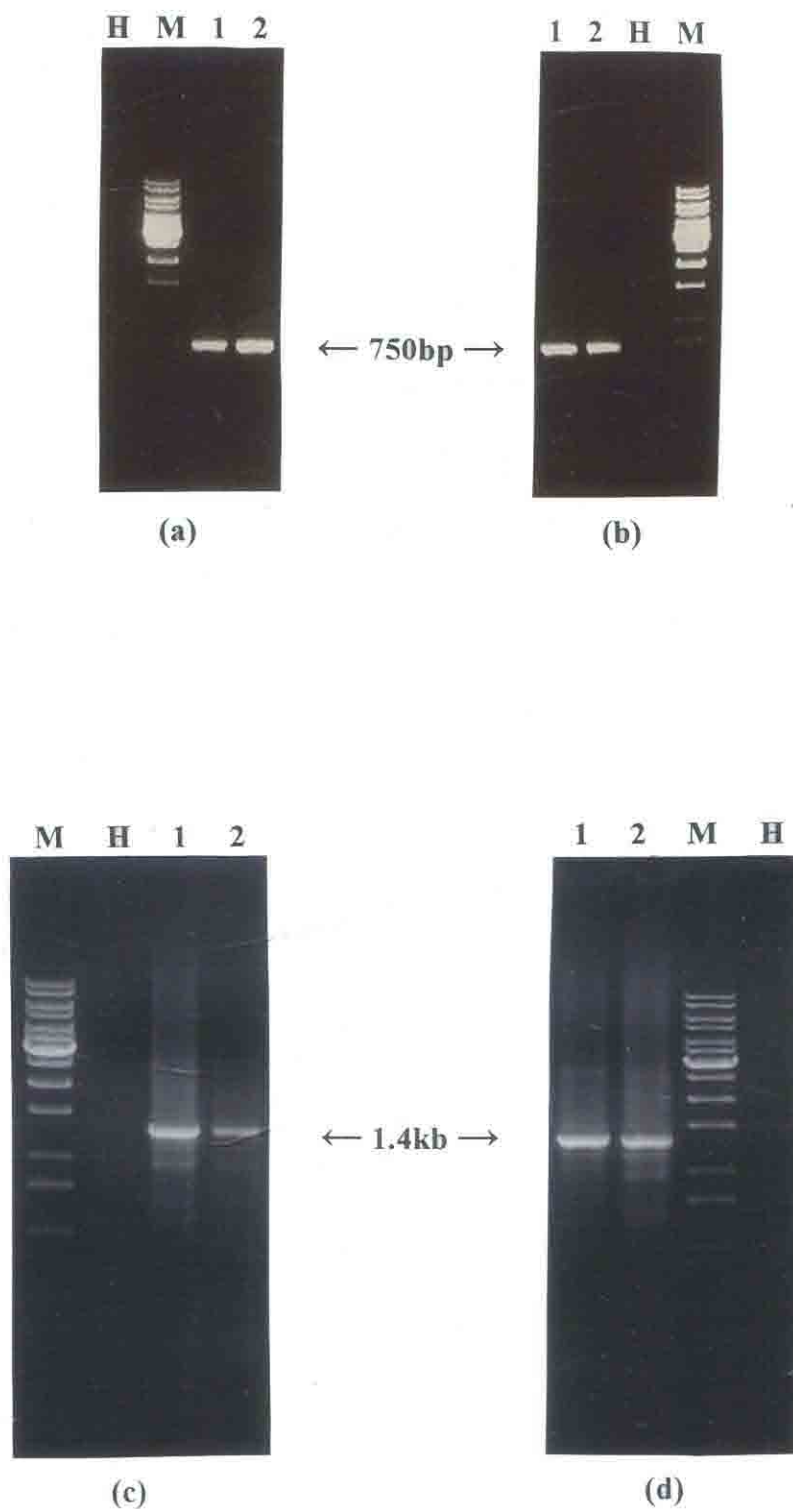


Figure 4.2.8. PCR detection of CLCuV-GI DNA components [DNA-A (coat protein gene, ~750bp) and DNA-β (~1.4kb)] in agroinoculated plants using coat protein primers (a, b) and DNA-β primers (c, d).

(a, c) Lane 1: Agroinoculated cotton cv. LH900 plants, 2: Positive control (cotton cv. LH900 plants inoculated by whiteflies), H: Healthy cotton control, M: 1kb DNA ladder
 (b, d) Lane 1: Agroinoculated Frenchbean cv. Sel.9 plants, 2: Positive control (cotton cv. LH900 plants inoculated by whiteflies), H: Healthy cotton control, M: 1kb DNA ladder



(a)



(b)

Figure 4.2.9. Symptoms induced by infection of CLCuV agroprogeny from agroinoculated cotton (*G.hirsutum* cv. LH900) plants by whitefly inoculation. A leaf from *G.hirsutum* cv. LH900 plant showing typical vein thickening symptoms (a) and *G.barbadense* line GB14-193 plant showing typical vein thickening, enation and severe leaf curl symptoms (b).

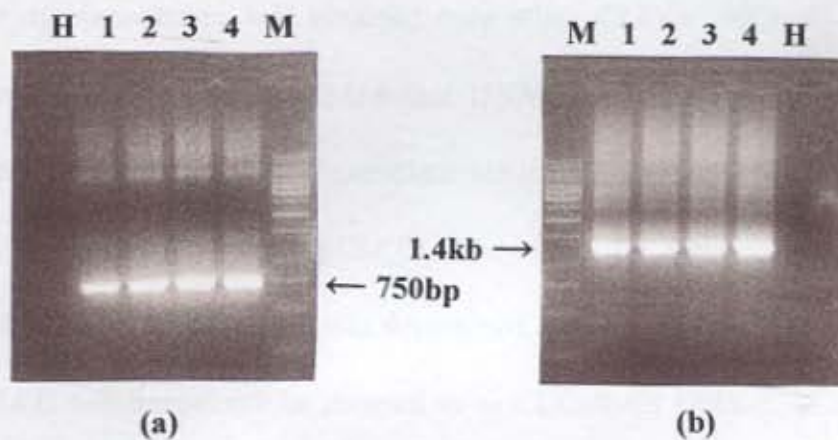


Figure 4.2.10. PCR detection of CLCuV-GI DNA components [DNA-A (coat protein gene, ~750bp) and DNA-β (~1.4kb)] in cotton plants after whitefly transmission from agroinoculated cotton (*G.hirsutum* cv. LH900) plants using coat protein primers (a) and DNA-β primers (b)

Lane 1: Agroinoculated *G.hirsutum* cv. LH900 plants used for acquisition, 2: *G.hirsutum* cv. LH900 plants inoculated with CLCuV agroprogeny by whiteflies, 3: *G.barbadense* line GB14-193 plants inoculated with CLCuV agroprogeny by whiteflies, 4: Positive control (cotton cv. LH900 plants inoculated by whiteflies), M: 1kb DNA ladder

H. Healthy Cotton

Attempts to back transmit CLCuV from agroinfected Frenchbean cv. Sel.9 to cotton (*G. hirsutum* cv. LH 900 and *G. barbadense* line GB14-193) plants, by whitefly given AAP and IAP of 72hr, were, however, not successful (Table 4.2.3).

Analysis of DNA-A and DNA- β sequences

As agroinoculation was obtained only when CLCuV DNA-A and DNA- β were coinoculated, it was concluded that DNA-A and DNA- β are the genomic components of CLCuV-GI. The complete nucleotide sequence of DNA-A clone (pCLCA) and DNA- β clone (pCLC β), shown to be infectious and whitefly transmissible in this investigation was determined. The component represented by the clone pCLC β will henceforth be referred to as CLCuV-GI DNA- β , in line with the nomenclature used for a similar molecule isolated from *Ageratum conyzoides* (Saunders *et al.*, 2000) and cotton (Briddon *et al.*, 2001). The full-length sequence of CLCuV-GI DNA-A was 2754 bp (GenBank Acc. No.-AF363011) and of DNA- β was 1350 bp (GenBank Acc. No.-AY083590) (Appendix 3). These sequences are available in the GenBank, DDJB and EMBL nucleotide sequence databases. Nucleotide numbering proceeds from the 3' A in the conserved nonanucleotide sequence 5'-TAATATTAC-3'. This motif, which DNA- β shares with geminiviruses, forms the loop of a predicted stem loop structure, which, for geminiviruses, contains the nick site of the Rep protein for initiation of virion-strand DNA replication (Laufs *et al.*, 1995; Stanley *et al.*, 1995).

Genome organization of DNA-A

The genome organization of CLCuV-GI DNA-A is similar to the DNA-A genomic component of bipartite begomoviruses and maintained the typical genome organization of begomoviruses originating from the Old World (Hanley-Bowdoin *et al.*, 1999; Harrison and Robinson, 1999), with six conserved open reading frames ORFs (Fig.4.2.11), including an AV2 gene, the function of which remains obscure.

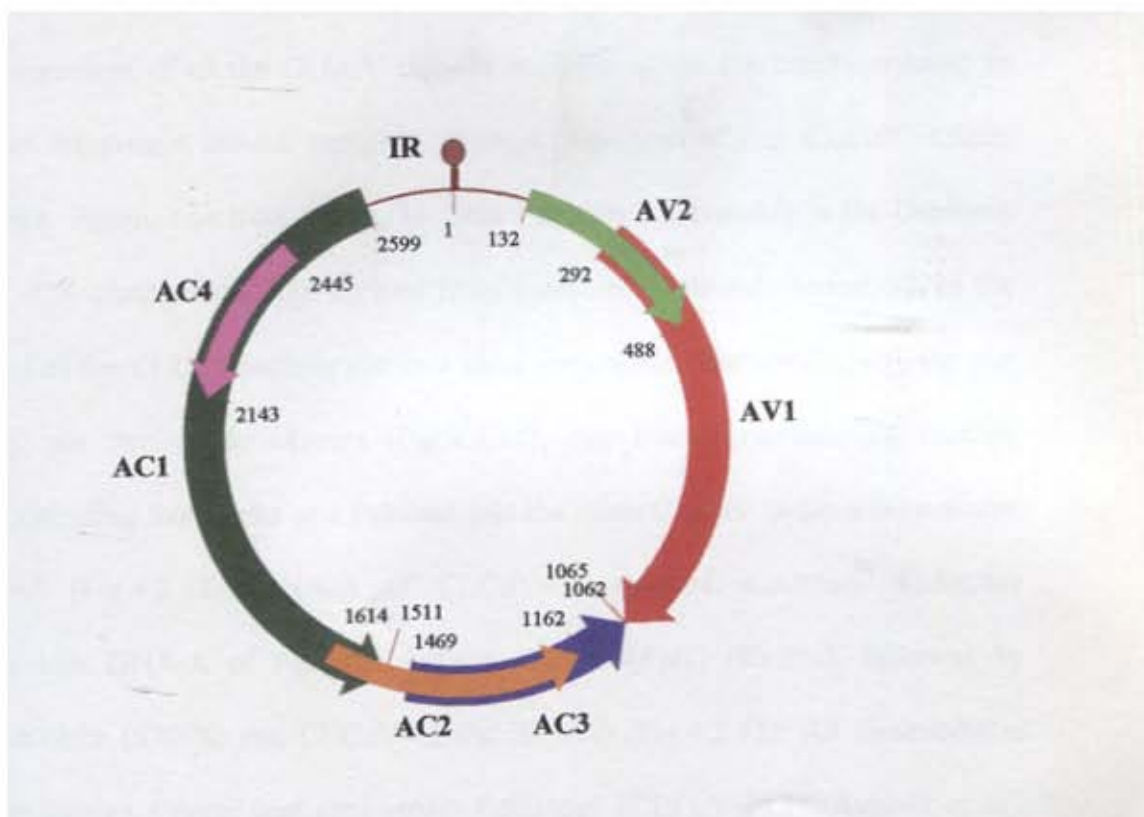


Figure 4.2.11. The genome organization of CLCuV-GI DNA-A (2754 bp, GenBank Acc. No.-AF363011). Arrows represent ORFs in both orientations. ORFs are indicated as being either on the viral sense (V) or complementary sense (C) strands.

Two of the six ORFs are on the viral sense strand and four are on the complementary sense strand (Fig.4.2.11) coding for predicted proteins of > 9 kDa (Table 4.2.4).

Sequence comparison of CLCuV-GI DNA-A with the other CLCuV isolates

Nucleotide sequence of CLCuV-GI DNA-A was compared to the complete DNA-A sequences of all the CLCuV isolates available in the GenBank database. In addition to the present isolate, complete DNA-A sequences of four CLCuV isolates from Gezira, Sudan, one from Egypt, 11 from Pakistan are available in the GenBank database. The phylogenetic tree derived from complete nucleotide sequences of the DNA-A of all the CLCuV isolates shows a clear geographic relationship, with the tree branching into two major clusters (Fig.4.2.12); one branch representing CLCuV isolates originating from India and Pakistan and the other CLCuV isolates from Sudan and Egypt (Fig.4.2.12). DNA-A of CLCuV-GI showed maximum nucleotide identities with DNA-A of Pakistan isolates CLCuV-[Fai1] (88.2%), followed by CLCuV-806bPk (87.8%) and CLCuV-72bPk (87.6%) (Fig.4.2.12). All these isolates belong to species *Cotton leaf curl virus- Pakistan2* (CLCuV-Pk2) (Rybicki *et al.*, 2000) and form one separate cluster. CLCuV-GI clustered with CLCuV-Pk2 but branched separately within this cluster, occupying a distinct place in the tree (Fig.4.2.12). All the other Pakistan CLCuV isolates had less than 87% nucleotide sequence identity with CLCuV-GI. Among Pakistan isolates, CLCuV-[804a] and CLCuV-802aPk showed maximum diversity of 23% with CLCuV-GI, whereas all the Sudan isolates were most distantly related, having more than 30% diversity with CLCuV-GI. These differences clearly indicate that CLCuV-GI is a new species of CLCuV.

Comparison of CLCuV-GI DNA-A with other begomoviruses

Nucleotide sequence comparisons of CLCuV-GI DNA-A to the sequences of selected begomoviruses described from India and Pakistan, showed maximum

Table 4.2.4. Features of CLCuV-GI DNA-A (2754 bp, GenBank Acc. No.- AF363011) open reading frames

ORF	Start	Stop	No. of amino Acids	Predicted molecular weight (kDa)
AV1 (CP)	292	1062	256	29.7
AV2	132	488	118	13.8
AC1 (Rep)	2599	1511	362	40.7
AC2 (TrAP)	1614	1162	150	17.3
AC3 (REn)	1469	1065	134	15.6
AC4	2445	2143	100	11.1

Table 4.2.5. Comparison of CLCuV-GI DNA-A identity with DNA-A and various ORFs of selected begomoviruses from India and Pakistan

Virus	Total DNA-A	IR	AV1	AV2	AC1	AC2	AC3	AC4
CLCuV-Pk	86.8	56.0	94.9	88.1	83.7	97.3	97.0	59.0
CLCuV-72b Pk	87.6	91.9	96.4	75.4	94.7	58.2	67.9	98.0
CLCuv-806b Pk	87.8	91.9	96.4	75.4	95.5	60.2	68.6	98.0
CLCuV-802a Pk	77.5	55.6	97.2	92.5	70.2	79.3	86.5	39.0
CLCuV-311 Pk	86.6	57.3	95.7	94.2	85.0	96.6	98.5	68.0
OYVMV-201 Pk	83.9	53.6	97.2	92.5	86.7	82.0	85.0	78.0
OYVMV-301 Pk	82.0	71.6	95.7	90.9	84.8	78.6	86.5	61.0
BYVMV-In	82.0	77.7	96.4	92.5	80.9	76.6	85.8	64.7
ICMV	70.7	34.8	90.6	63.5	73.5	58.9	70.1	57.8
PaLCV-In	74.2	46.4	94.9	72.0	68.8	61.5	67.9	38.0
ToLCV-NDe	71.1	46.5	91.7	60.1	75.4	51.3	61.7	38.0
ToLCV-BanII	76.7	51.4	94.5	72.0	80.1	60.2	70.1	49.0
ToLCV-Ban4	74.4	47.3	79.6	66.1	79.6	55.6	65.6	44.0
MYMIV-Bg	60.8	32.9	64.9	28.8	69.1	44.3	41.4	38.2

All the ORFs compared using amino acid sequences.

nucleotide identities (> 82%) to OYVMV and BYVMV (Table 4.2.5). All other *Begomovirus* sequences had less than 82% identity.

The phylogenetic tree derived from complete nucleotide sequences of the DNA-A genomic components of cotton leaf curl viruses and other selected begomoviruses shows a clear geographic relationship; viruses originating from the Indian sub-continent except MYMIV-Bg were more closely related to each other than to viruses from other regions (Fig.4.2.13). These sequences highlight the genetic difference between the New World (TGMV, BGMV, BDMV, AbMV and SiGMV) and Old World begomoviruses.

Organization of intergenic region and its comparison with other begomoviruses

The intergenic region (IR) of 286bp between the start codons of AV2 and AC1 genes represents the most characteristic feature of the begomoviruses. Except for the conserved nonanucleotide sequence (TAATATTAC), which has the DNA nick site *Ori* required for the initiation of replication of viral sense strand, this region is most variable. The IR of CLCuV-GI has the repeat sequence or iterons or Rep binding sequence TGATCA like some other isolates of CLCuV (Table 4.2.6). These iterons are present as three copies in tandem orientation in the IR (Table 4.2.6). The first copy of these sequences is at 136nt coordinate upstream of *Ori* of replication in the nonanucleotide sequence which is near the 5' end of the IR, the second copy is located just to the 5' side or immediate upstream of the TATA motif, which is located at 87nt coordinate upstream of *Ori* of replication in the nonanucleotide sequence and the third copy is located on the 3' side or immediate downstream of the TATA motif. The iteron sequence and arrangement of CLCuV-GI are similar to CLCuV-72bPk and CLCuV-806bPk (Table 4.2.6), as described earlier by Zhou *et al.* (1998) and Sanz *et al.*, (2000).

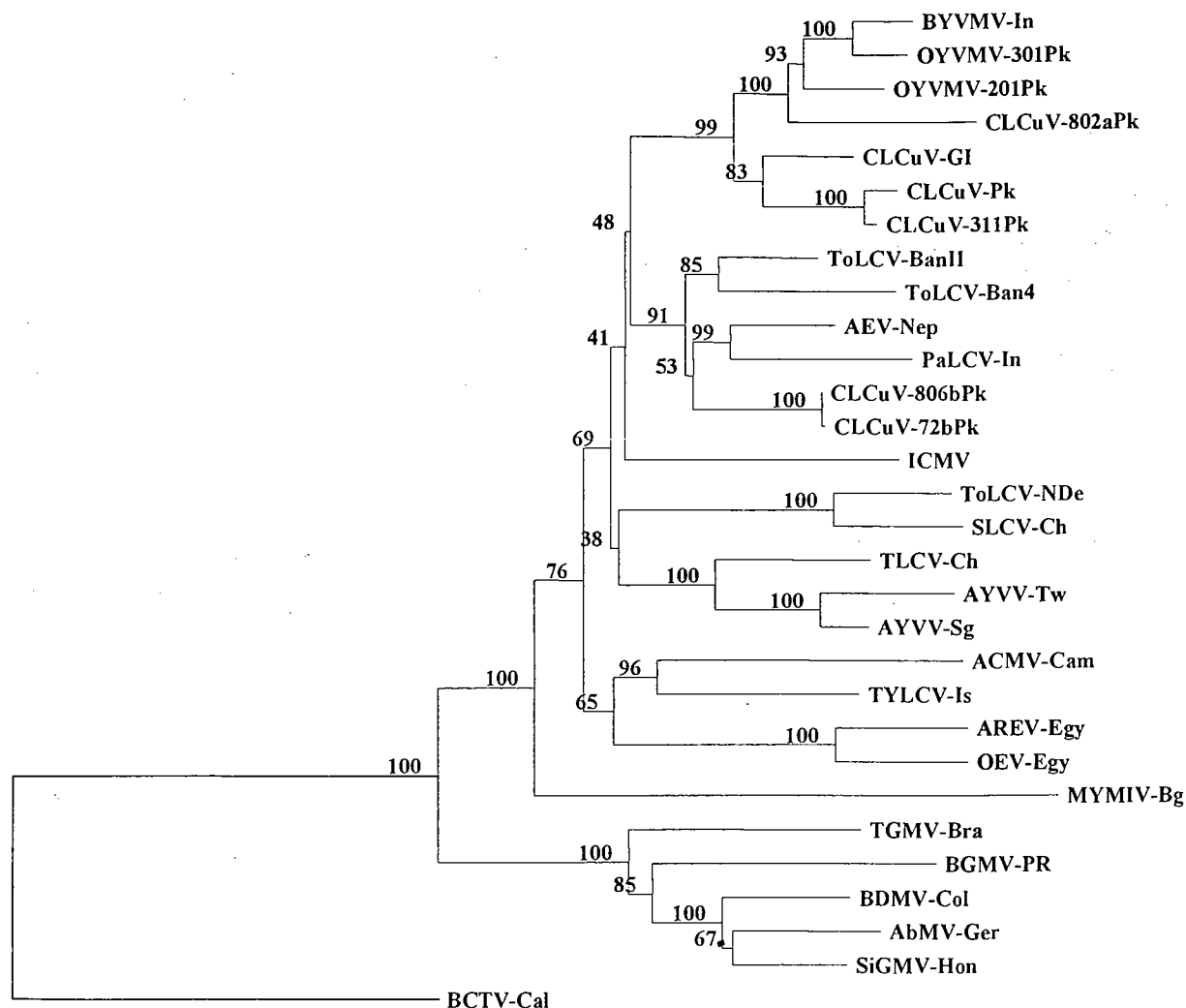


Figure 4.2.13. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the complete DNA-A. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the BCTV-Cal sequence.

Table 4.2.6. Iteron sequence and their arrangement in the intergenic region of CLCuV and some other related Begomoviruses

S.No	Iteron sequence	Copy No.	Virus	Iteron copies and arrangement in relation to 'TATA' boxes*
1	TGATCA	3	CLCuV-GI, CLCuV-72b Pk, CLCuV-806bPk	→ →TATA→
2	GGTGTA	3	CLCuV-Pk, CLCuV-311 Pk	→ → TATA→
3	GGGGAACT	2	CLCuV-802aPk	→ → TATA
4	CGGT(A/G)CCC	2	OYVMV-201Pk	→ →TATA
5	GTGAGA	2	OYVMV-301Pk	→ TATA→
6	TGAGACCC	2	BYVMV-In	→ TATA→

* Note: Arrow (→) indicates iteron sequence

Comparison of nucleotide sequence identity of intergenic region of CLCuV-GI DNA-A with the other Begomoviruses showed >90% sequence identity with CLCuV-72bPk and CLCuV-806bPk and less than 80% identity with other viruses (Table 4.2.5). It was interesting that the sequence identity of IR is greater with BYVMV-In (77.7%) and OYVMV-301Pk (71.6%) than with the other CLCuV isolates, which have < 60% identity.

Phylogenetic tree derived from nucleotide sequences of the intergenic region of cotton leaf curl viruses and other selected begomoviruses shows that begomoviruses from cotton and bhendi form a separate cluster (Fig.4.2.14). Within this cluster, there are two sub clusters, CLCuV-GI clustering with bhendi/okra begomoviruses and more closely with CLCuV-72bPk and CLCuV-806bPk.

Multiple alignment of the IR nucleotide sequences of 8 begomoviruses isolated from malvaceous hosts (3 from bhendi/okra and 5 from cotton) was done to locate the variable and identical regions (Fig.4.2.15). The lengths of the IR of these isolates range from 257 to 286 nucleotides. The sequences show a highly conserved central region of 46 nucleotides ie., 23nt upstream and 23nt downstream of *Ori* of replication in the nonanucleotide sequence TAATATTAC among these begomoviruses (Fig.4.2.15). At 3' upstream region of *Ori* of replication, between ~ 24nt coordinate to ~132nt coordinate and at 5' downstream region of *Ori* of replication, between ~ 24nt coordinate to ~89nt coordinate, there is considerable sequence variability. Another substantial block of conserved sequences is also present in the 3' side of the IR sequence, whereas, the 5' side of the IR has relatively greater proportion of variable blocks.

Comparison of virus encoded proteins

In contrast to the variability of the IR, the most conserved region among *Begomovirus* genomes is the AV1 or coat protein (CP) gene. This was also reflected

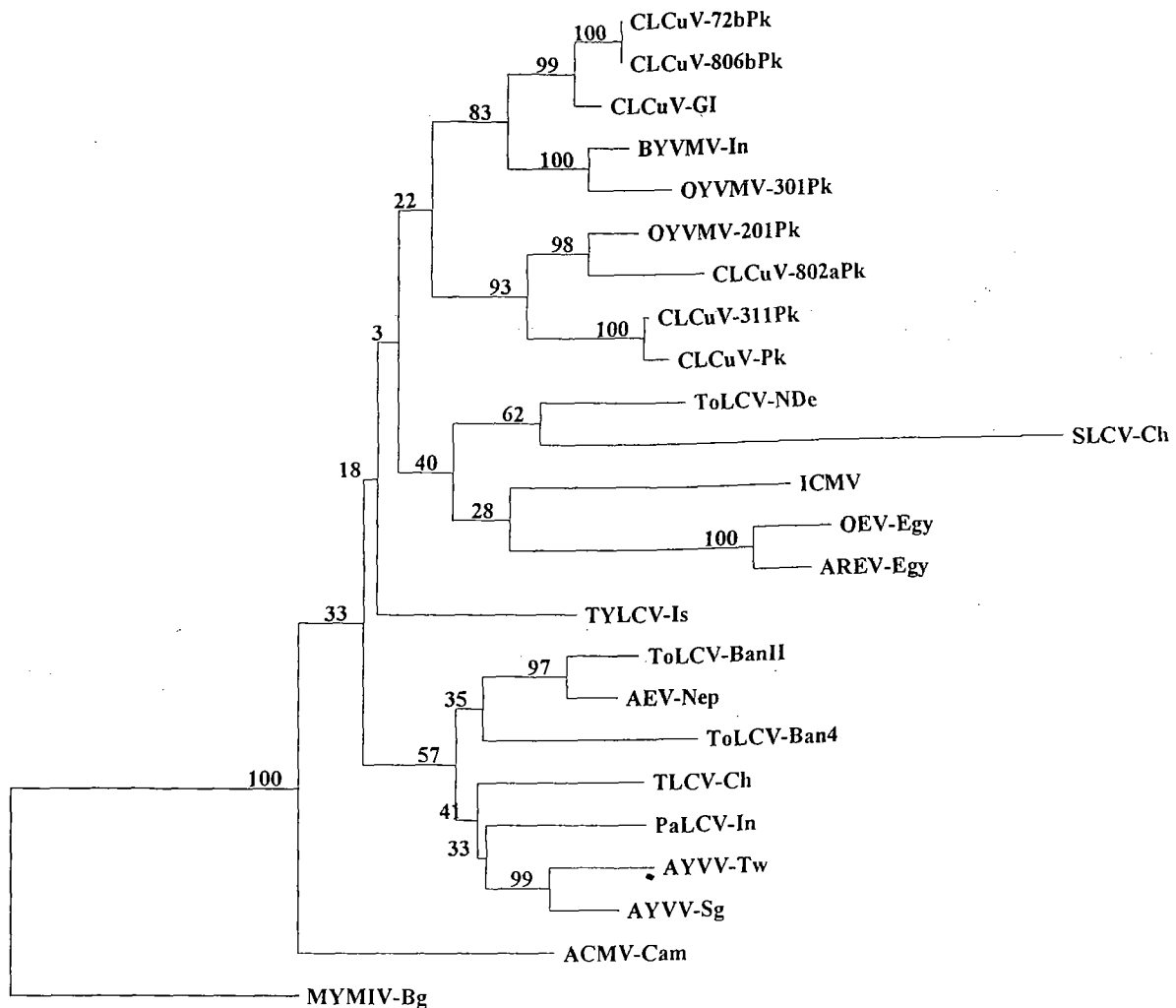


Figure 4.2.14. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the intergenic region. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

CLCuV-GI	1	---GTTGACTAAAATTGATCAAC--GATTGACCGCTCTTGCAAOTCTCCCCGGTATATCGG	56
CLCuV-Pk	1	---TTTGGCAA--TCGGTGTACAC-TCTAATTCTCTGGCAATTG---GTG-TAA-CGG	47
CLCuV-806bPk	1	---GTTGACTAAAATTGATCAAC--GATTGACCGCTCTTGCAAOTCTCCCCGGTATATCGG	56
CLCuV-72bPk	1	---GTTGACTAAAATTGATCAAC--GATTGACCGCTCTTGCAAOTCTCCCCGGTATATCGG	56
CLCuV-802aPk	1	---TTTGTCTAT--CGGGGACAC--TTAAAACTCATATCAATTG---GGG-ACACTGG	47
CLCuV-311Pk	1	---TTTGGCAA--TCGGTGTACAC-TCTAATTCTCTGGCAATCG---GTG-TAA-TGG	47
OYVMV-301Pk	1	---GTTGACCAAGTGAGAGGACCCGATTGACCGCTCTTGCAAOTCTCCCCTGTATATTGG	57
OYVMV-201Pk	1	TGACTTTGGTCAA--TCGGTACCCAGATCTAATCCTATGTCAATTG---GTG-AA--CGG	51
BYVMV-In	1	---GTTGACTATTTTITGAGACCC--GATTGACCGCTCTTACAACCTCTCCCCAGTATATCGG	56
CLCuV-GI	57	TGATCAATATATAGTGATCAG--CAAATGGCA-----TAATCGTAATAAAAAAACT	105
CLCuV-Pk	48	GGTGCAATATATAGGTGTACCCCAAATGGCA-----TATC-GTAATGGGAGAAAT	95
CLCuV-806bPk	57	TGATCAATATATAGTGATCAG--CAAATGGCA-----TAATCGTAATAAAAAAACT	105
CLCuV-72bPk	57	TGATCAATATATAGTGATCAG--CAAATGGCA-----TAATCGTAATAAAAAAACT	105
CLCuV-802aPk	48	GGACTCATTTTATACTGCGTCCCTAAATGGCATAAATGTAAATATTITGCTTTATGTTTG-	106
CLCuV-311Pk	48	GGTGCAATATATAGGTGTACCCCAAATGGCAT-----TATC-GCAATTTGAGAAAT	97
OYVMV-301Pk	58	GTCTCAATATATAGTGAGACC--CAAATGGCA-----TTATCGTAATTTGGGACAT	106
OYVMV-201Pk	52	TGCCCTATATATAGTGGGTAC--TGAATGGCAT-----TATTGTGAATTACAAAAGG	101
BYVMV-In	57	GTCCCTATATATAGTGAGACC--CAAATGGCA-----TAATGTGAATAAAAACAACT	105
CLCuV-GI	106	TTAATTTGAAATTCAAACCAAAGGCTAAAGCGGCCATCCGTTTAATATTACCGGATGGC	165
CLCuV-Pk	96	C-AITTTGAGATCTCAGCTCCAA--AA-GCGGCCATCCGTATAATATTACCGGATGGC	151
CLCuV-806bPk	106	TTAATTTGAAATTCAAACCAAAGGCTAAAGCGGCCATCCGTTTAATATTACCGGATGGC	165
CLCuV-72bPk	106	TTAATTTGAAATTCAAACCAAAGGCTAAAGCGGCCATCCGTTTAATATTACCGGATGGC	165
CLCuV-802aPk	107	A-AITTTCAAATTCCAAACGCTCCAA--AAAGCGGCCATCCGTATAATATTACCGGATGGC	163
CLCuV-311Pk	98	C-AITTTCAAATCTCAGCTCCAA--AAAGCGGCCATCCGTATAATATTACCGGATGGC	154
OYVMV-301Pk	107	A-GATTCAAATCTCAGCTCCAA--AAAGCGGCCATCCGACTAATATTACCGGATGGC	163
OYVMV-201Pk	102	A-AATTCAAATCTACAGCTCCAA--AAAGCGGCCATCCGTATAATATTACCGGATGGC	158
BYVMV-In	106	TTAATTTGAAATTCAAACGAAAAGGCTAAAGCGGCCATCCGTATAATATTACCGGATGGC	165
CLCuV-GI	166	CGCGCGATTTTTTTTTGG-TGGGCCCTTACCATTAAACACTTGTTCGGCCAAATCATATGACTCC	224
CLCuV-Pk	152	CGCGCGATTTTTTTTAT--G-----TGGGCCCCCGATGTATGAGATTGCTC---CC	195
CLCuV-806bPk	166	CGCGCGATTTTTTTTTGG-TGGGCCCTTACCATTAACTCTTGTTCGGCCAAATCATATGACGCG	223
CLCuV-72bPk	166	CGCGCGATTTTTTTTTGG-TGGGCCCTTACCATTAACTCTTGTTCGGCCAAATCATATGACGCG	223
CLCuV-802aPk	164	CGCGCGATTTTTTTTTAGTGG-----TGGGTCCAGAACGCACGACGATGCAG---AC	211
CLCuV-311Pk	155	CGCGCGATTTTTTTTAT--G-----TGGGCCCCCGATGTATGAGATTGCTC---CC	198
OYVMV-301Pk	164	CGCGCGATTTTTTTTAAGTGG-----TGGGTCCAGAACGCACGACGATGCAG---AC	211
OYVMV-201Pk	159	CGCGCGATTTTTTTTTAGTGG-----TGGGTCCAGAACGCACGACGATGCAG---AC	206
BYVMV-In	166	CGCGCGATTTTTTTTAAG-TGG-----TGGGTCCAGAACGCACGACGATGCAG---AC	212
CLCuV-GI	225	TCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATTCA	283
CLCuV-Pk	196	TCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATTCA	254
CLCuV-806bPk	224	TCAAAGCTTAAATAATTCTCCCGCTTATTATAAGTACTTCGTTGCTAAGTATGCGTTTGA	283
CLCuV-72bPk	224	TCAAAGCTTAAATAATTCTCCCGCTTATTATAAGTACTTCGTTGCTAAGTATGCGTTTGA	283
CLCuV-802aPk	212	TCAAAGCTTAGATAACGCTCCTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATTCA	270
CLCuV-311Pk	199	TCAAAGCT-AAATAACGCTCCCGCCACTATAAGTACTTGCGCACTAAGTTTCAAATTCA	257
OYVMV-301Pk	212	TCAAAGCTTAGATAACGCTCCTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATTGA	270
OYVMV-201Pk	207	TCAAAGCTTAGATAACGCTTCTTTG-GCTATAAGTACTTGCGCACTAAGTTTCAAATTCA	265
BYVMV-In	213	TCAAAGCTTAGATAACGCTCCTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATTCA	271
CLCuV-GI	284	AAC	286
CLCuV-Pk	255	AAC	257
CLCuV-806bPk	284	AAA	286
CLCuV-72bPk	284	AAA	286
CLCuV-802aPk	271	AAC	273
CLCuV-311Pk	258	AAC	260
OYVMV-301Pk	271	AAC	273
OYVMV-201Pk	266	AAC	268
BYVMV-In	272	AAA	274

Figure 4.2.15. Multiple alignment of the nucleotide sequence of the intergenic region of DNA-A of 9 malvaceous Begomovirus isolates from India and Pakistan. Blocks of conserved regions are shown in reverse contrast.

by >90% sequence identity with 12 of the 14 viruses compared. Maximum amino acid sequence identity (97.2%) of the coat protein gene of CLCuV-GI was to CLCuV-802aPk and OYVMV-201Pk (Table 4.2.5). Comparisons of coat protein amino acid sequences lead to a conclusion that relationships are close among viruses from the same geographical area.

The phylogenetic tree representing ORF AV1 of the selected begomoviruses shows two major branches, containing Old World and New World viruses, respectively (Fig.4.2.16). Within the Old World branch, there are three sub clusters: one contains the viruses from Africa and the Mediterranean region and a second comprised exclusively of viruses from Indian sub-continent and China, whereas the third cluster contains AYVV-Sg and AYVV-Tw, from Singapore and Taiwan, respectively, which seem relatively distant from the other Old World viruses.

The AV2 gene occurs only in the Old World begomoviruses and partly overlaps with AV1. Nevertheless, the AV2 protein was more variable than the AV1. More than 90% amino acid sequence identity of AV2 gene of CLCuV-GI was found with two CLCuV and three okra/bhendi isolates (Table 4.2.5), whereas, it was < 88% with all the other isolates/viruses compared. Phylogenetic tree showed that all the malvaceous begomoviruses from Indian sub-continent cluster together with the exception of CLCuV-806bPk and CLCuV-72bPk, which cluster separately (Fig.4.2.17).

Since the AC1 gene product (replication initiation protein or Rep) binds in a sequence-specific way to the iterons in the IR (Fontes *et al.*, 1994; Lazarowitz *et al.*, 1992), variability of the IR coincides with variability of the Rep sequence, particularly in its N' terminal half, which contains the domain that interacts with the IR. The highest amino acid sequence identity of CLCuV-GI Rep was to CLCuV-806bPk

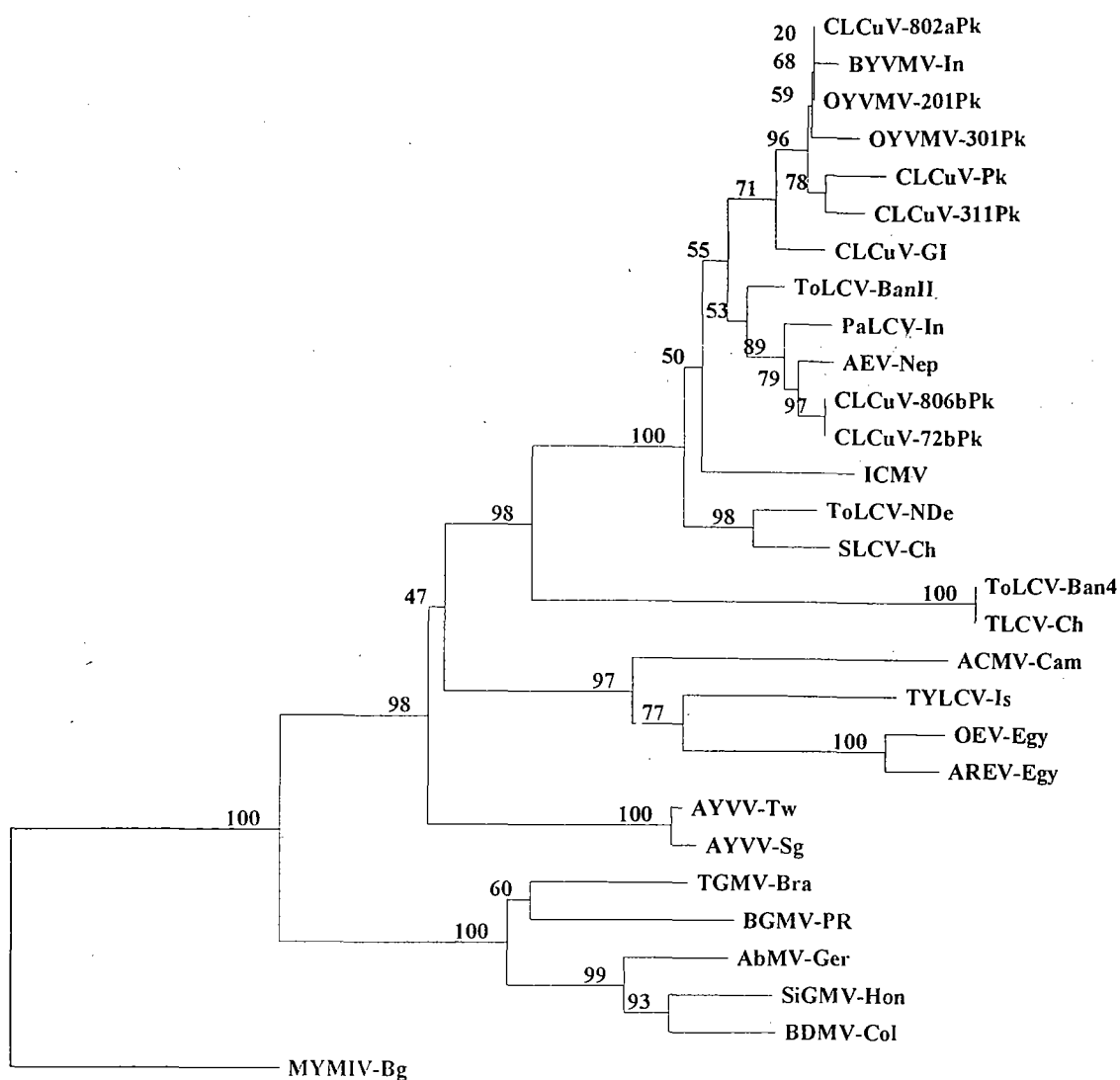


Figure 4.2.16. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AV1. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

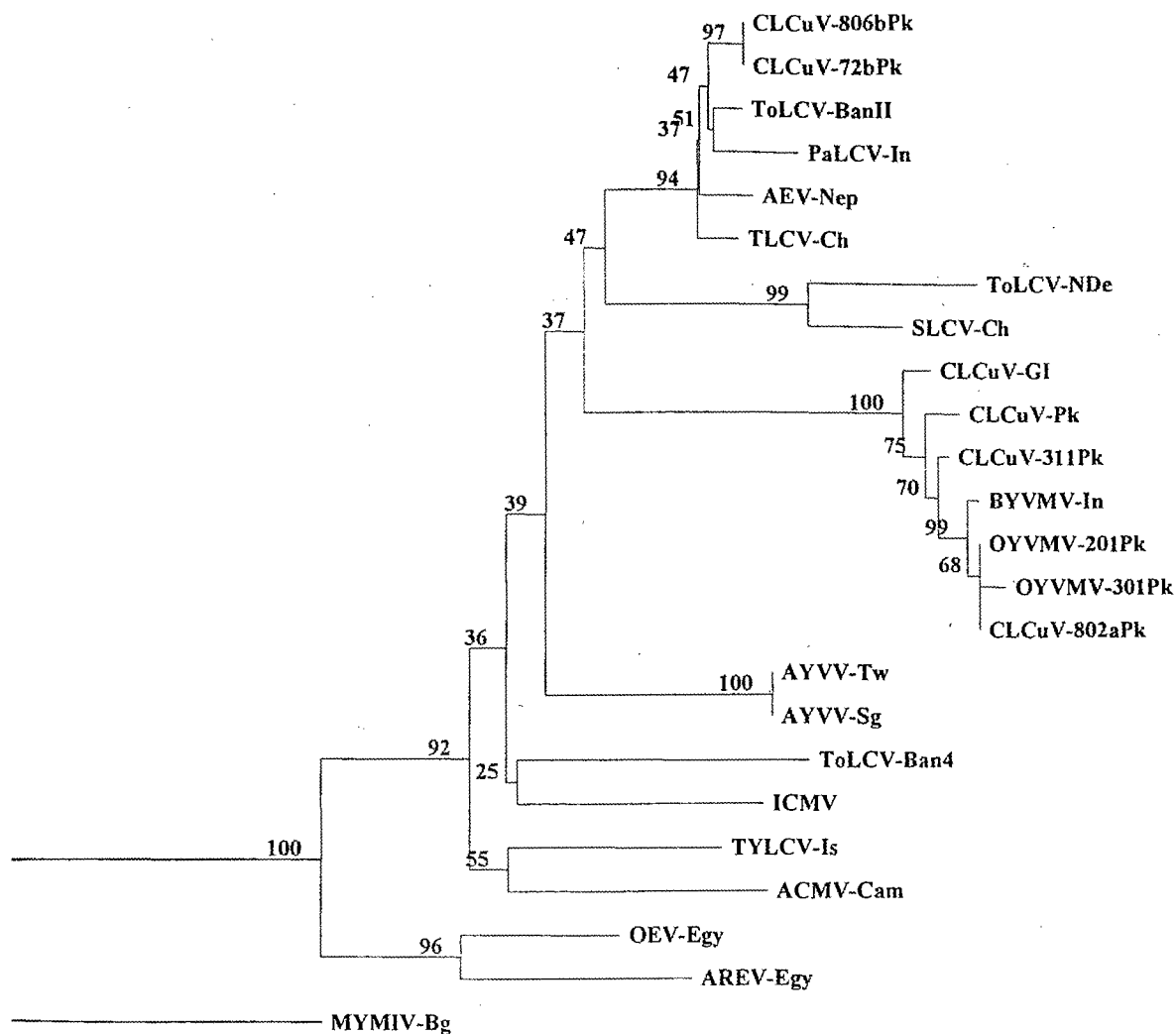


Figure 4.2.17. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AV2. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

(95.5%) followed by CLCuV-72bPk (94.7%), and it was < 87% to the other viruses (Table 4.2.5).

Comparison of Rep amino acid sequences in a phylogenetic tree (Fig.4.2.18) indicates more complex relationships in that sub clustering of viruses occasionally obscures some of the vector-related and /or geographically distinct relationships. Nevertheless, there is clear division between the ORF AC1 of Old World and New World begomoviruses. Within Old World begomoviruses branch, CLCuV-GI clusters with CLCuV-806bPk and CLCuV-72bPk (Fig. 4.2.18).

Both the AC2 and AC3 proteins have a degree of variability intermediate between those of CP and Rep, and show the same division between Old World and New World viruses as the other DNA-A genes. The highest amino acid sequence identity of AC2 gene (transcription activator protein or TrAP) of CLCuV-GI was to CLCuV-Pk (97.3%) followed by CLCuV-311Pk (96.6%) and rest others being less than 85% (Table 4.2.5). The identity ranged from 58.2% to 97.3% with CLCuV isolates from Pakistan. Phylogenetic tree showed that all the malvaceous begomoviruses from Indian sub-continent cluster together with the exception of CLCuV-806bPk and CLCuV-72bPk, which cluster separately (Fig. 4.2.19).

Comparison of AC3 gene (replication enhancer protein or REn) sequences indicates that CLCuV-GI AC3 gene amino acid sequence has the highest identity with CLCuV-311Pk (98.5%) followed by CLCuV-Pk (97.0%) (Table 4.2.5). The identity pattern was more or less similar to the identity pattern with AC2 gene. Phylogenetic tree derived from AC3 gene amino acid sequence from selected begomoviruses showed many small clusters within the Old World Begomoviruses (Fig.4.2.20). Here again clustering of CLCuV-GI was similar to that of TrAP.

AC4 gene comparison revealed that CLCuV-GI has the highest amino acid sequence identity (98.0%) with CLCuV-806bPk and CLCuV-72bPk (Table 4.2.5).

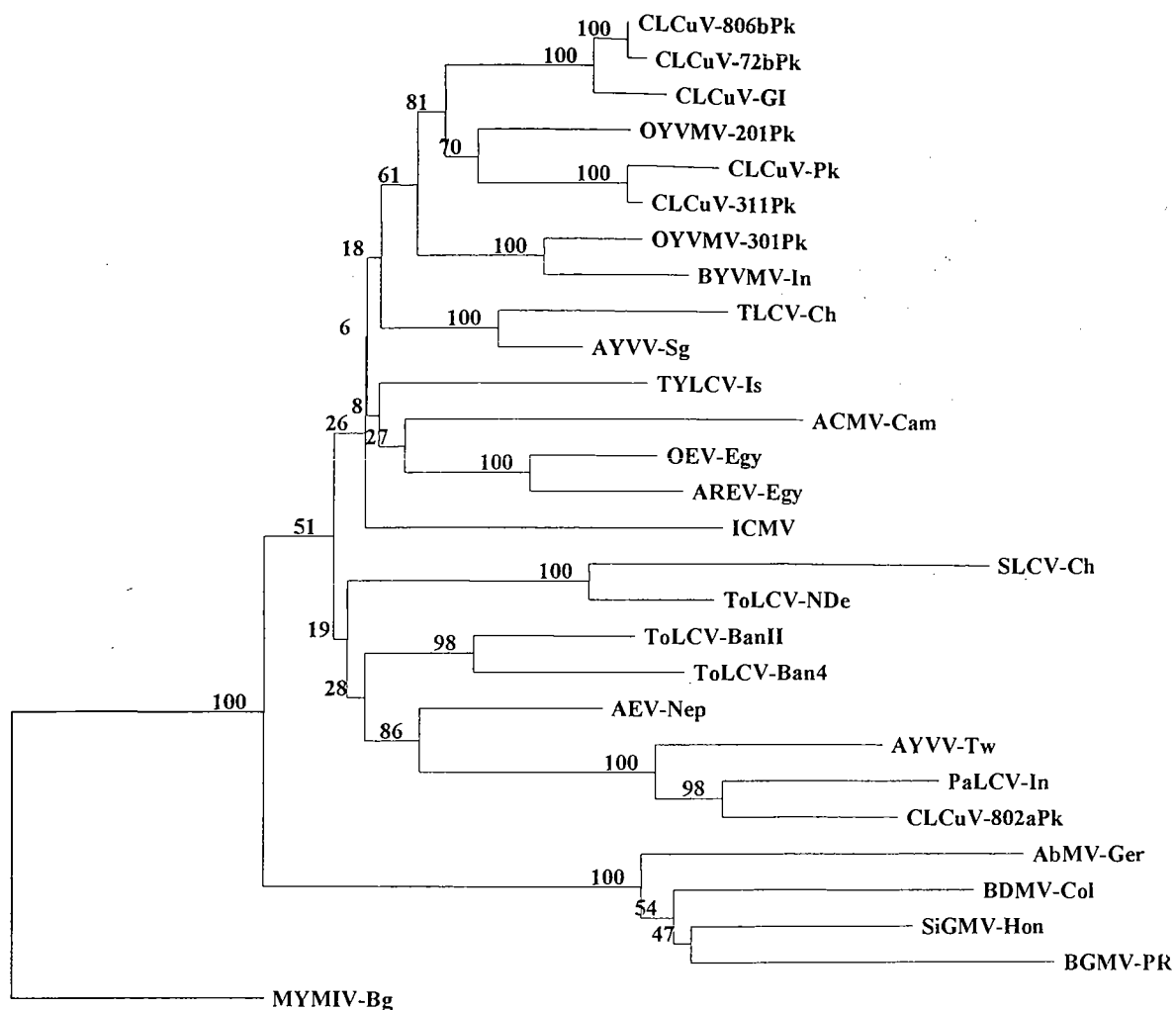


Figure 4.2.18. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AC1. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

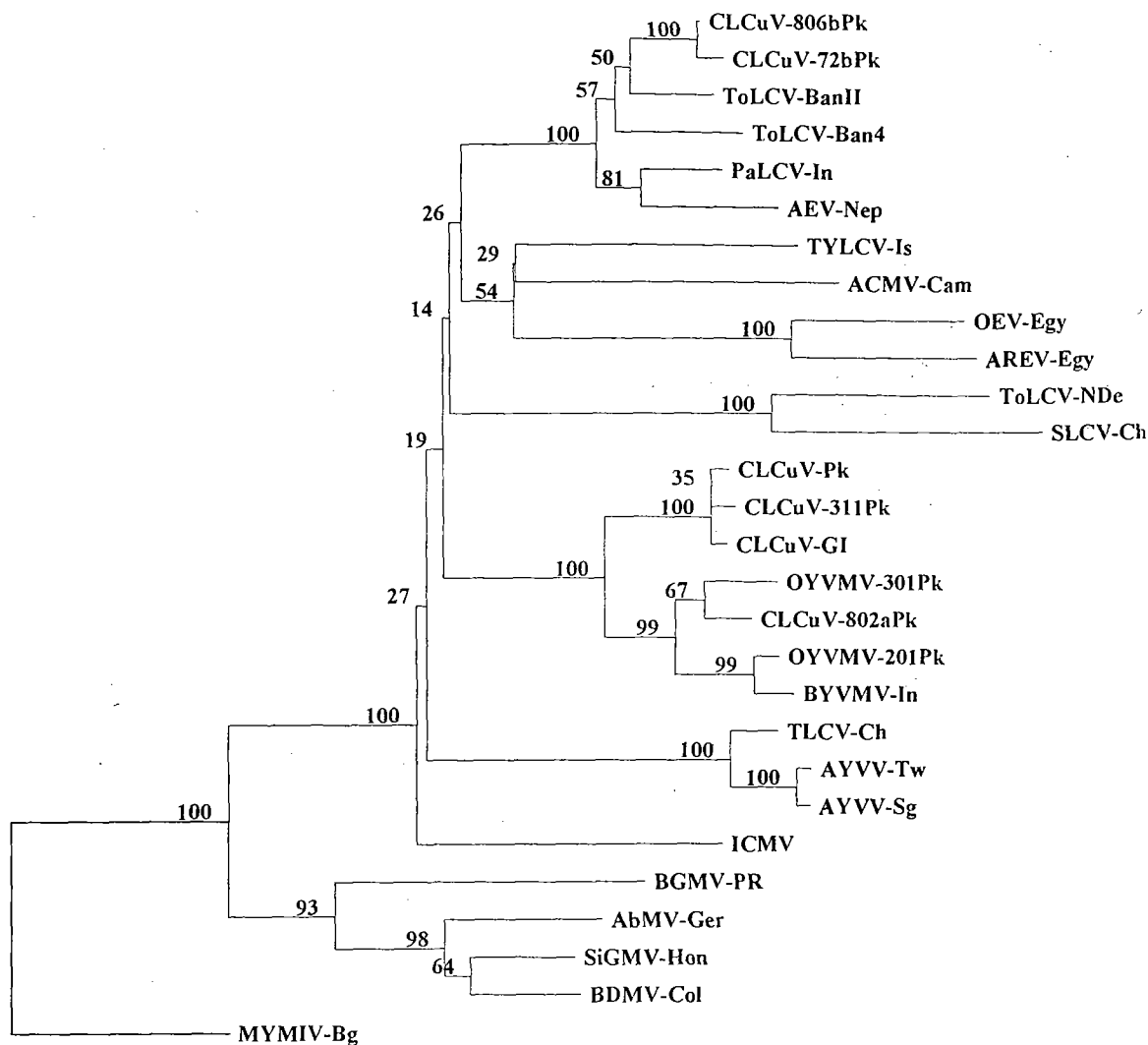


Figure 4.2.19. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AC2. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

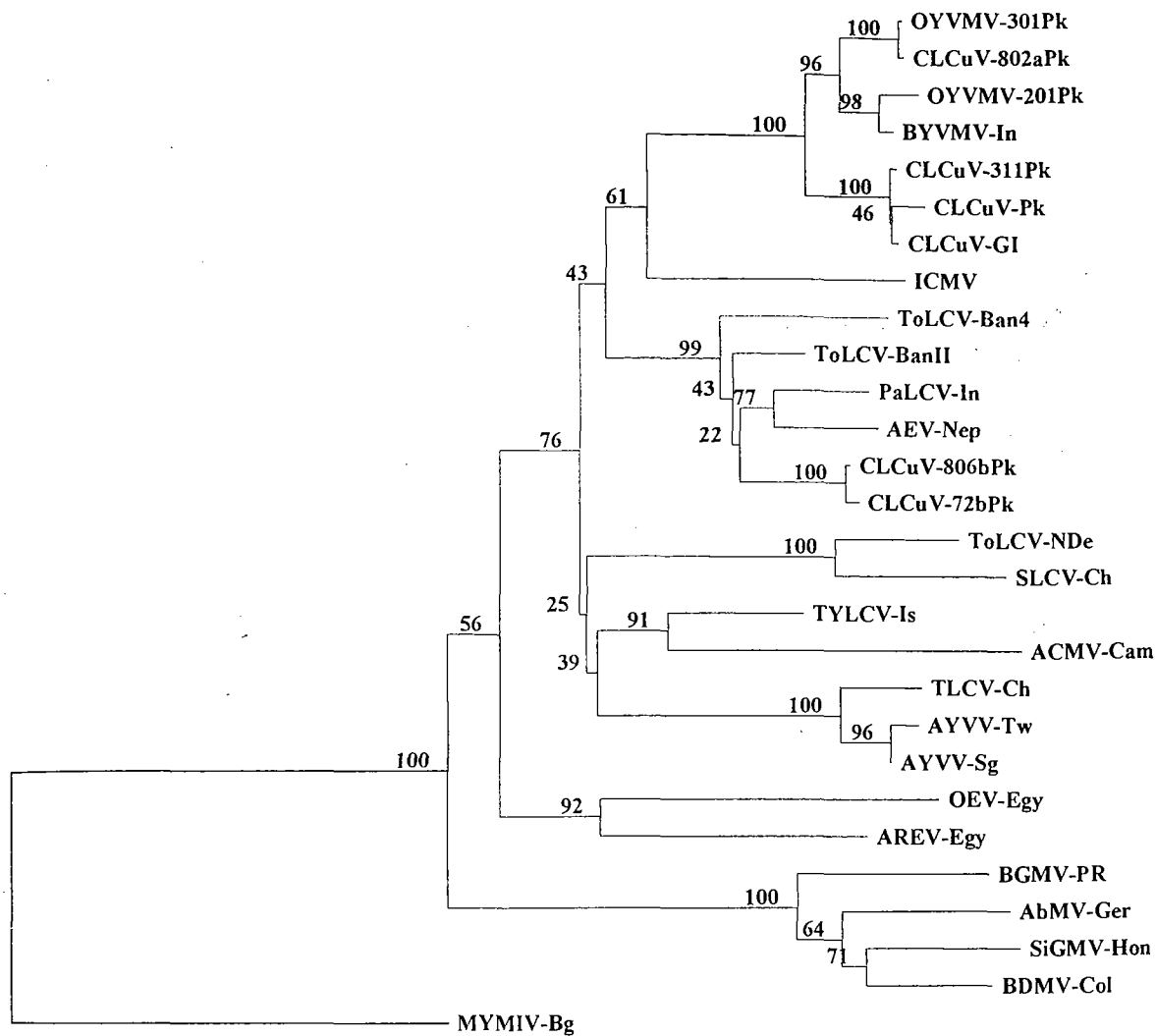


Figure 4.2.20. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AC3. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

Identity was less than 80% with other viruses selected for the study. It was interesting to note that among all the DNA-A genes, CLCuV-GI AC4 gene had maximum identity with CLCuV-806bPk and CLCuV-72bPk, and least identity with other selected begomoviruses. Hence, it can be said that among the CLCuV-GI DNA-A genes, AC4 gene is the most variable gene. Since this AC4 gene totally spans the N' terminal half of the Rep gene (AC1), which contains the Rep domain that interacts with the iterons in the IR, variability of the IR coincides with the variability of AC4 gene sequence. Phylogenetic tree derived from AC4 gene sequence does not give a clear geographical virus clusters (Fig.4.2.21), although CLCuV-GI clustered with CLCuV-806bPk and CLCuV-72bPk, as expected from the amino acid sequence identity.

Genome organization of DNA- β

Analysis of the sequence in both orientations identified nine ORFs (beginning from an in-frame methionine codon) with a predicted coding capacity of proteins of >4 kDa (Table 4.2.7). These ORFs are denoted as being encoded on the virion (β V) or complementary (β C) strands and the positions and orientations of these ORFs are shown (Fig.4.2.22). Of all the ORFs, ORF β C4 of CLCuV-GI DNA- β has the highest predicted coding capacity of 13.7 kDa (Table 4.2.7). ORF β C4 of CLCuV-GI DNA- β is the positional analogue of β C1 ORF in DNA- β of other begomoviruses reported till date.

Sequence comparison of DNA- β

Comparison of the complete nucleotide sequence of CLCuV-GI DNA- β to the DNA- β molecules associated with other begomoviruses shows relatively low sequence identity of 57.3% with AYVV-Sg DNA- β , 54.4% with BYVMV-In DNA- β and 49.7% with OLCV-Sd DNA- β , whereas, the identity was 93.7% with the

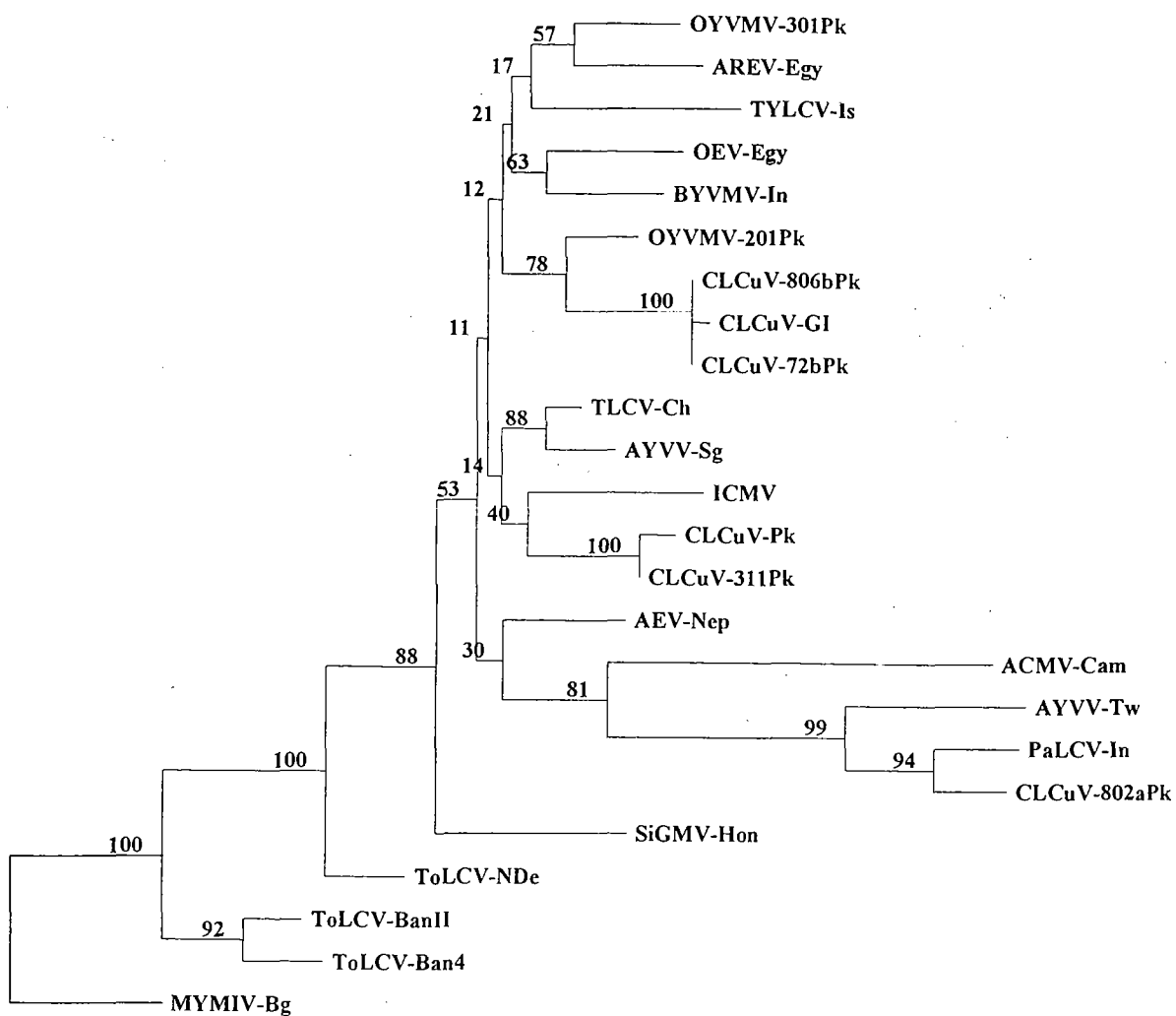


Figure 4.2.21. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AC4. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

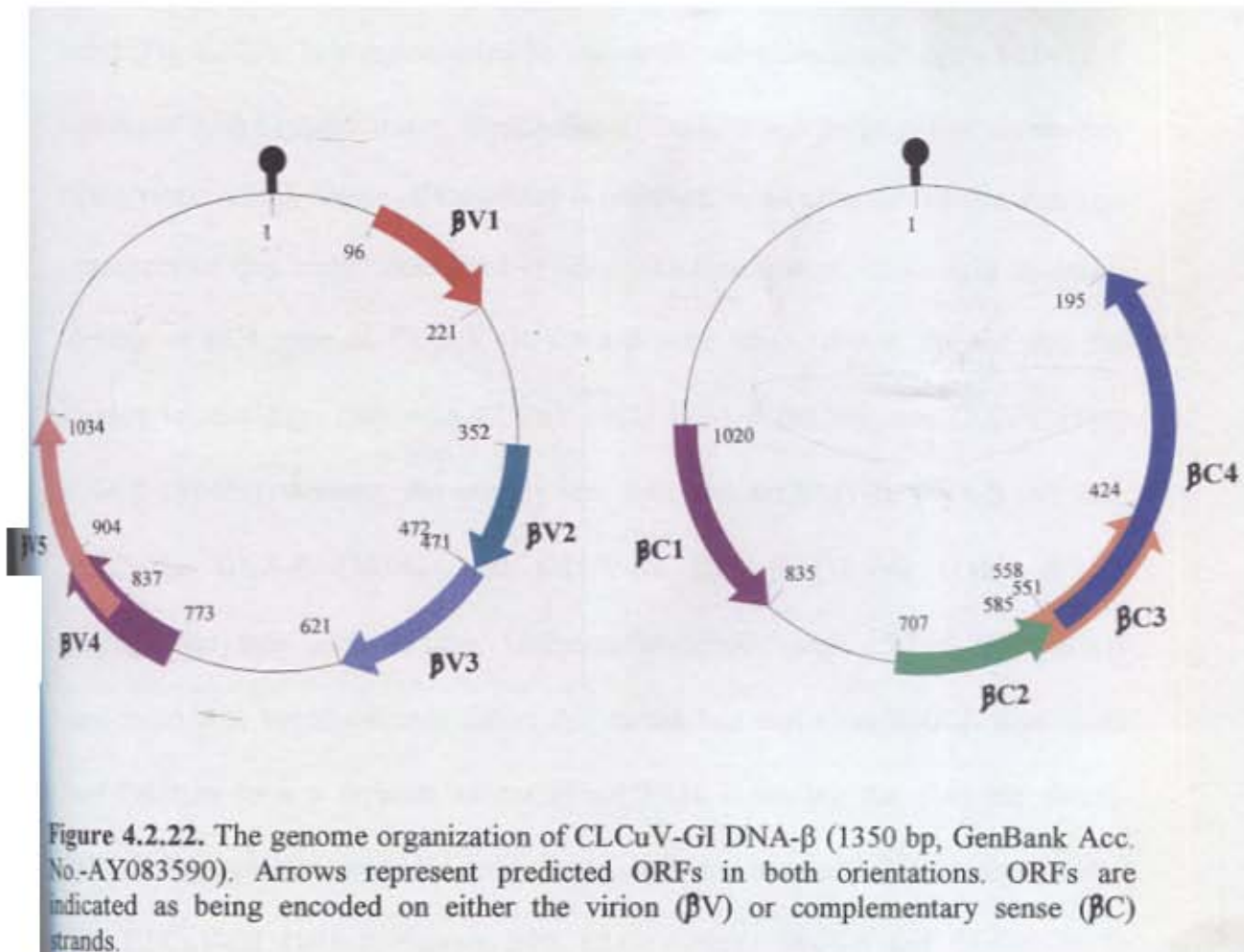


Table 4.2.7. Features of CLCuV-GI DNA- β (1350 bp, GenBank Acc. No.-AY083590) open reading frames

ORF	Start	Stop	No. of amino Acids	Predicted molecular weight (kDa)
β V1	96	221	41	4.7
β V2	352	471	39	4.4
β V3	472	621	49	5.8
β V4	773	904	43	5.2
β V5	837	1034	65	7.6
β C1	835	1020	61	7.6
β C2	558	707	49	5.9
β C3	424	585	53	6.3
β C4	195	551	118	13.7

Table 4.2.8. Comparison of CLCuV-GI DNA- β identity with other begomoviruses associated DNA- β

Virus	Total DNA- β	β C1 gene
CLCuV-Pk(1)	93.5	97.4
CLCuV-Pk(2)	93.7	98.3
BYVMV-In	54.4	42.1
OLCV-Sd	49.7	31.3
AYVV-Sg	57.3	33.0

CLCuV-Pk(2) DNA- β and 93.5% with CLCuV-Pk(1) DNA- β (Table 4.2.8). DNA- β sequences have a highly conserved region over 100 nucleotides upstream and 15 nucleotides downstream of the *Ori* of replication in the nonanucleotide TAATATTAC motif (Fig.4.2.23). This region could be said as the most conserved region in DNA- β associated with begomoviruses, irrespective of the host and geographical distribution of the virus. A high degree of variability is observed in the upstream and downstream sequences of this highly conserved region. Comparison of amino acid sequence identity of β C4 gene of CLCuV-GI DNA- β with other DNA- β showed that the identity is very high only with CLCuV-Pk(2) DNA- β (98.3%) and CLCuV-Pk(1) DNA- β (97.4%), whereas, the identity was low with BYVMV-In DNA- β (42.1%), AYVV-Sg DNA- β (33.0%) and OLCV-Sd DNA- β (31.3%) (Table 4.2.8). Phylogenetic tree derived from complete nucleotide sequence of the DNA- β associated with begomoviruses shows that cotton leaf curl virus DNA- β from India and Pakistan form a separate cluster (Fig.4.2.24), indicating that they are closely related. Phylogenetic tree based on the amino acid sequence of β C4/ β C1 also showed that CLCuV-GI DNA- β clusters with CLCuV-Pk(1) DNA- β and CLCuV-Pk(2) DNA- β (Fig.4.2.25), indicating that they are closely related.

DISCUSSION

The results presented here show that CLCuV has two genomic components, represented by DNA-A and DNA- β , which are required to cause cotton leaf curl disease (CLCuD), as has been shown by agroinfection of cotton plants. The virus could also be back transmitted to cotton by whitefly (*G. tabaci*) from agroinfected plants. Earlier, Briddon *et al.*(2001) also showed using biolistic inoculation that CLCuD is caused by the co-infection of CLCuV-Pk1 DNA-A and DNA- β , although they could not back transmit the progeny virus by whitefly. In the present

CLCuV-G1 1 ACCGTGGGCGAGCGGTC CGATGGCTT TGC TGGG CUC - C TGGT TTTAC TGT TTTGACT AT TGGGCGAT T T T T S - G G T T A A - 96
CLCuV-Pk1 1 ACCGTGGGCGAGCGGTC TC-CGGGGA SCA TGGGTCCTA C-GTGTGGT TTTAC TGT TTTGACT AT TGGGCGAT T T T T S - G G T T A A - 95
CLCuV-Pk2 1 ACCGTGGGCGAGCGGTC CGATGGT TTT TGC TGGG CUC - C TGGT TTTAC TGT TTTGACT AT TGGGCGAT T T T T S - G G T T A A - 96
BYVMV-In 1 ACCGTGGGCGAGCGGAT T T T -AGT GTT TGGAA C C C - C T A T A C A - - - - - T G A G T T A - - - - - T G G T G A T T C T C A - T A T C A G S G - 80
OLCV-Sd 1 ACCGTGGGCGAGCGGTC TGGACTGTATTTG TGGT TCCCGGCA TGGCAT TAA GAGCTCTTGGATT TGGG TGA C C C A C C C A A T T S T A T - 100
AYVV-Sg 1 ACCGTGGGCGAGCGGTC ATT-TGGGTT TAC TGGGTCCTA CAC TGGACCA C C S G A G A A C C G - - - - - T G G G C C A A - A T S - G G C C - 94

CLCuV-G1 97 TGT TGT G C C T T G G T A G S G G C - - - - - T T T T G G A T T S G G C T C T G A T T T G T T G G - - - - - T A T T G T G T T A T A T S - A T T G C T G G T T T G T G - 192
CLCuV-Pk1 97 TGT TGT G C C T T G G T A A A G A C - - - - - T T T T T G G A T T S G G C T C T G A T T T G T T G G - - - - - T A T T G T G T T A T A T S - A T T G C T G G T T T G T G - 192
CLCuV-Pk2 96 TGT TGT G C C T T T G G A G S G G C - - - - - T T T T T G G A T T S G G C T C T G A T T T G T T G G - - - - - T A T T G T G T T A T A T S - A T T G C T G G T T T G T G - 192
BYVMV-In 81 A A T G T T G A T G A A A G G A A T C A A G T T T G C A A T T T T A T T A T A C A G T C C A T A A A A T T A T T A T C A A A - - - - - C A A A T A C A A C A - 177
OLCV-Sd 101 T A T T T A A T T A T G T T T A A T A - - - - - A A T T G C C G T T A A T G A T A T A T A A A A A A A A T A T G T A A - A A T T A T T A T A - T A T A A A G A T A A - 193
AYVV-Sg 94 - A T S G C T T T T A T T C A A A C A - - - - - C A A A T T T C G A A G C C C A T A T A T T A T C A T T - - - - - S A A A A - T A C T T A A A A T A T G C C A A A A T A A C - 189

CLCuV-G1 193 A T T A A C G G T G A C T T T A T C A T A A A C C T T G T T T A A A C A A T A T T C T C T G T T A G A G T C A T G T A A A A G T C T G A T T - 292
CLCuV-Pk1 193 A T T A A C G G T G A C T T T A T C A T A C A C C T T G T T T A C C C A T C A T A T C T C T G T T T C A G T C A G T A T A T C A T C C A G A G T - 292
CLCuV-Pk2 193 A T T A A C C G G T G A C T T T A T C A T A C A C C T T G T T T A C C C A T C A T A T C T C T G T T T C A G T C A G T A T A T C A T C C A G A G T - 292
BYVMV-In 178 A A T A A T T A T T C T C T A T A C A T A G T T A C T C T G T T T A C C C A T C A T A T C T C T G T T T C A G T C A G T A T A T C A T C C A G A G T - 277
OLCV-Sd 194 C A T T A C G G T A T A T T A T A T C T G C C A A T T A T T A T T A T T C C A T T C C A T T C C A G T C C C T C C A T C C A G C C A G T A T - 293
AYVV-Sg 189 - T T A A C G G T T A C T T T A T C A T C A T A T T C C T G T T A T T G A T A T T T T G G A C T T C C T T G C T C T G A T A T A T T A T T A C A C - 288

CLCuV-G1 293 T A T C T C T A T C T C G A T C T C T C T A T C T T G C C C G T T T G G C C A T A G G A A T C G C T A G A T C C C T T C A A A C C G T T A G T G A A T S A A T G - 392
CLCuV-Pk1 293 T A T C T C T A T C T C G A T C T C T C T A T C T T G C C C T G T T G T G C C A T A G G A A T C G C T A G A T C C C T C T C A A A C C G T T A G T G A A T S A A T G - 392
CLCuV-Pk2 293 T A T C T C T A T C T C G A T C T C T C T A T C T T G C C C G T T G T T G T G C C A T A G G A A T C G C T A G A T C C C T C T C A A A C C G T T A G T G A A T S A A T G - 392
BYVMV-In 278 T A T C T C T C T C A G A T T T C T C T C T C T T T A A C C C A A T A T G A C T T G T T A G T C C C T C T C A A A C C G T T A G T G A A T S A A T G - 377
OLCV-Sd 294 A T T C T C C T C C T T A C A T C S G T G A G T T G T T T T G T C A T A T T C C T T A C C A T T C C T T A C C A T T C C T C A G C C C G T T A A G T C G A A T G G A T G - 392
AYVV-Sg 289 T A T C T C T A T T T G A T C T C T C A T T T G T A A T T C A T C A T C C T T A T A T T C T T G A T A C C T T T C T A G A T A T A T T A A T T A A T G G T G A T - 388

CLCuV-G1 393 T A T C T C T A T T G A C T T A T C C A T C C T C A T A C T T G A T T A G A T G G T G A C T T G T T A T A G T A T T C T A T C T G A T T A G A G T C T C A T A T C C C - 492
CLCuV-Pk1 393 T A T C T C T A T T G A C T T A T C C A T C C T C A T A C T T G A T T A G A T G G T G A C T T G T T A T A G T A T T C T A T S T A G T A G T C A T A T C A T T C T C - 492
CLCuV-Pk2 393 T A T C T C T A T T G A C T T A T C C A T C C T C A T A C T T G A T T A G A T G G T G A C T T G T T A T A G T A T T C T A T S T A G T A G T C A T A T C A T T C C C - 492
BYVMV-In 378 T A T T T T T A T T G A T T A T C C A T C C A T C C T T C T T T G A T T G A T T G A T T G A T T G T T A G T A T T G T T A G T A T T C T A T T C C T A T T C C - 477
OLCV-Sd 394 C A T A C C G C C S T G A T A T G T A T A T A A A C G T T G A T T C C T G A T A T G T T C T T T G T T C C G T T A C T C A T C T - - - - - T G C A C T A T A T A G C C G T - 490
AYVV-Sg 389 C T A C C T A T S A T S C T A T G S A T S G A A A T T G T T C T G A C T A G T T G G T G A C C T G T T A A A C C T T C T C G C A C T C T G A T T A T T C T G A T - 488

CLCuV-G1 493 A T G T G C G A C C G C A C T A T G A C C T S C C C C T C T T G T T T C C C C T G C T T C A T T T C G C T A T T G A G S A T T T A - - - - - T T S - G T G C A G C T T - 590
CLCuV-Pk1 493 A T G T G C G A C C G C A C T A T G A A C C T S A C C C C C T C T T G T T T C C C C T G C T T C A T T T T C G C T A T T G A T A G S A T T T A - - - - - T T S - G T G C A G C T T - 590
CLCuV-Pk2 493 A T G T G C G A C C G C A C T A T G A A C C T S A C C C C C T C T T G T T T C C C C T G C T T C A T T T T C G C T A T T G A T A G S A T T T A - - - - - T T S - G T G C A G C T T - 590
BYVMV-In 478 A T G S C C C T A C A T C T A C T T A T G A C A C A T C C C C C C G T T T T T C T G A A C C - - - - - T C A T C T T G A T T T S C S T A T T G A A - - - - - A T G T T G A A A G A T A - 574
OLCV-Sd 491 T T G S G A G A C A C C G A C T C G A S T C S C C C C T G G G T T T T A A T G T A A T C A T A T T T A A A - - - - - T T C T G T T G A C A T A C A G G A A - A T A - 589
AYVV-Sg 489 C T G A C T T G A C T T T A T A T A G A A C A A T T T C C T A T C T S T A T G A T A T A T C A T - - - - - T T T T T T A T G S A A T A T A T C T T G A T G C T A T - 584

CLCuV-G1 591 T A T A S C T T A T G A S A - - - - - A T A T T T S T G A C T T G T G T T G T G A S T A T T A T T A T A T G A T T S C C A T T A A S G G A A A S T G A C C T S G A G A C G T - 690
CLCuV-Pk1 591 T A T A S C T T A T G A S A - - - - - A T A T T T S T G A C T T G T G T T G T G A S T A T T A T T A T A T G A T T S C C A T T A A S G G A A A S T G A C C T S G A G A C G T - 690
CLCuV-Pk2 591 T A T A S C T T A T G A S A - - - - - A T A T T T S T G A C T T G T G T T G T G A S T A T T A T T A T A T G A T T S C C A T T A A S G G A A A S T G A C C T S G A G A C G T - 690
BYVMV-In 575 T A T A S C T G A A T G A A T A T A T T T T C A T A T A A T G T G G A T G A A - - - - - A A T T A T A G A T A T T T C A T A T - - - - - A T G T G C G C T - - - - - A - 652
OLCV-Sd 590 T A T A T G A T C A T A - - - - - A A A T A T A T C T A A G A - - - - - T A T T G A T T A T A T A A A C C G T A T A T C C T G A T - - - - - T C C A C A C T - - - - - T T C - 669
AYVV-Sg 585 T A T A T T S G A T C - - - - - T G T C T T T A T G T G C T T C - - - - - A T T A A T A T C T G C A A T T A T C T C T C - - - - - A - - - - - T A T - - - - - T - 655

CLCuV-G1 691 A T T A C C G T T T G T A T S T T G C - - - - - T S A A T C T T A T A C T G G G T T G A C C C G T A T T C C T T A T A T A C T S T S A A C C G G T S A A S G A A A C - 789
CLCuV-Pk1 691 A T T A C C G T T T G T A T S T T G C - - - - - T S A A T C T T A T A C T G G G T T G A C C C G T A T T C C T T A T A T A C T S A A A C C G G T S A A S G A A A C - 789
CLCuV-Pk2 691 A T T A C C G T T T G T A T S T T G C - - - - - T S A A T C T T A T A C T G G G T T G C C G G T - - - - - T C C T T A T A C T S A A A C C G G T S A A S G A A A C - 788
BYVMV-In 653 A T T T C T A A A A C T C A T C C T A T C C A A A A A G G T C A T G - - - - - G A A T T T A T T A G C T C A G A A T C G G A C G A C T A A G G T A C C C A C A C T T C - 752
OLCV-Sd 670 G T A T C A T A C A T A T A C A C S C - - - - - C A T A C G A T A A S C A T C T A A G A G A T A A A T T A T A T A A G S A A A A A T A A T A A T A A T T A A T - 766
AYVV-Sg 656 T T S T T G T G T A C T C A T T G A T G T T A T G T T T G S A T T - - - - - F A A T S A G A A A A A G A A T T C C T T T C T T T A T A G A A A G A A S A A A A A S A A - 755

CLCuV-G1 790 G G A A C - T G A A - - - - - G G A A A A A A A A A A A A A A C - - - - - A A G C A T A T T A T T A T A G A A A A A T G S G - - - - - S C S A T C G A A A G C G - - - - - A A - 880
CLCuV-Pk1 790 G G A A C - T G A A - - - - - G G A A A A A A A A A A A A A A C - - - - - A A G C A T A T T A T T A T A G A A A A A T G S G - - - - - S C G A C A T C G A A C C G A - - - - - A A C - 880
CLCuV-Pk2 789 G G A A C - T G A A - - - - - G G A A A A A A A A A A A A A A C - - - - - A A G C A T A T T A T T A T A G A A A A A T G S G - - - - - S C G A C A T C G A A C C G A A C - - - - - A A - 879
BYVMV-In 753 C A T C T T C A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 848
OLCV-Sd 767 T A C A A G A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 858
AYVV-Sg 756 A C T A T A T A C A T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 847

CLCuV-G1 881 G S A A A - - - - - A G A A A - - - - - A A T A A A A A A - - - - - T A A A - - - - - S A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 969
CLCuV-Pk1 881 G S A A A - - - - - A G A A A - - - - - A A T A A A A A A - - - - - T A A A - - - - - S A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 969
CLCuV-Pk2 880 G S A A A - - - - - A G A A A - - - - - A A T T A A A A A - - - - - T A A A - - - - - S A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 968
BYVMV-In 849 A S A A A A G G G A C C G C C G - - - - - A A A A A A A A A T C C G G A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 948
OLCV-Sd 859 G A G G S - - - - - C T T T G T T A A A A A A A - - - - - A C - - - - - T T T A T T G C C G C A C T - - - - - C A T A T G T T - - - - - T T G A A T C C G G T C T C A A G C T A T T T T - 951
AYVV-Sg 848 A S A A A T C C T A T A T A T A T A T S A G A A T T G A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 947

CLCuV-G1 970 A A A - - - - - A T A T C C G A A A C T T T T T T T T T G A A S T G A C T G T S G G T T T - - - - - A C A T T A C T A T T T G G T A A A S G T A G T - - - - - A T T G A C A T A A - 1056
CLCuV-Pk1 970 A A A - - - - - A T A T C C G A A A C T T T T T T T T T G A A S T G A C T G T S G G T T T - - - - - A C A T T A C T A T T T G G T A A A S G T A G T - - - - - A T T G A C A T A A - 1056
CLCuV-Pk2 969 A A A - - - - - A T A T C C G A A A C T T T T T T T T T G A A S T G T C T G T S G G T T T - - - - - A C A T T A C T A T T T G G T A A A S G T A G T - - - - - A T T G A C A T A A - 1055
BYVMV-In 949 G G T A A A A C G A A G C C G T T T T G S G S A T A G G A A A A S G A A C G G T C A A T T A C A T T T A C A T T A G C S C G A A A T C G A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 1048
OLCV-Sd 952 C C G G T A G T - - - - - T G A A T T T A T G C A A - A T C C G T C C T A T T T T T T T T A T T T T A T T C C G T T A T T C - - - - - C C C A T C C C T A T T A T A C C C T A T A T C - 1045
AYVV-Sg 948 A A A T A A A A T G A T A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A - 1043

CLCuV-G1 1057 T T A A A - A A T T T G T C A C A T A G G T A A A T T G A T A C C A T A - - - - - T A T C C G G T A T C A A T T G G T G C T C A T A T T G G T G - - - - - T C C T A A A A A C - 1142
CLCuV-Pk1 1057 A A A G T A A T T T G T C A C A T A G G T A A A T T G A T A C C A T A - - - - - T A T C C G G T A T C A A T T G G G A C T A T A A T T G C C - - - - - T T C T A A A A A C - 1143
CLCuV-Pk2 1056 T T A A A - A A T T G G A G G C C A T A G G T A A A T T G T A C C A T A - - - - - T A T T G S G T C A A T T G G G A C T A T A A T T G C C T - - - - - T T C T A A A A A C - 1141
BYVMV-In 1049 G T A A A T T T G A C A G A C A C C A T A G G T A A A T T G A G A C C A T A - - - - - T A T C G G T T C T A A T T G G T A C T T G G G S C T T A T A A A G T G C T T G T T A C C - - - - - A T T T S C - 1146
OLCV-Sd 1046 G T G C C - T C T T A T A T A G T A G A T - - - - - C C A A T A G A - G T A C A T A G A T A T C C G G T C C A T T C A T T S T A C T A T T A S T C T A A T G C A A G C C T T A C C A A T T A C - 1141
AYVV-Sg 1044 G G T A A A T T A T T G T C T A T A G G T A A A T T G T - - - - - T C A T A - - - - - T A T C G G A C A T T T T G A G A T T G G C A C A C T A T A - T G C T A A T C C T - - - - - T T T T G - 1139

CLCuV-G1 1143 C C T A T T C T S T C T C - - - - - T G A A S G C G C G T S A G T G C G C T C A A A A G C S T T C T C T C T C A C T A A A A C C A C C G G A C G C C A A A C T G G T S T T G C G G C - 1241
CLCuV-Pk1 1144 C C C C G T T T S T C T C - - - - - T A A G A G C G C G C T S A G T G C G C C T A A A G T A A C A T C T C T C C C T T T T G A T C T C A T A C - - - - - A T T T C C S T S A T G G A - 1240
CLCuV-Pk2 1142 C C C C G T T T S T C T C - - - - - T A A G A G C G C G C T S A G T G C G C C T A A A G T A A C A T C T C T C C C T T T T G A T C T C A T A C - - - - - A T T T C C S T S A T G G A - 1238
BYVMV-In 1147 C C T A C T T T A A T - - - - - T A T A T C C C G C G T A A G T G C G C T T A A A G T T S T C T C T C T C T C A G T T C C G A T A A A C C T - - - - - A T T T C C S A T A T G G A - 1242
OLCV-Sd 1142 C C T A A A A A G C T G A G A A A C G G C G C T S C A T T G C G C C A A A A A G T A C A T C T C T C C C T A T T T C C G T A A T C C G G T S T G C G G C - 1241
AYVV-Sg 1140 C C T A C T A A S T G A C - - - - - T S T A A G C G C G T S A G T G T T C A A A A G T A G C C T C T C T C T C A A A A C T C C C G G A G A G C C T A C A G C T A C C G - 1238

CLCuV-G1 1242 A T C A A - - - - - T T A T G A C A C G C G C G C G T G T G T A C C C T G G G A G G T A S T A C C A C T A C G C T A C G A G A G C C T A G C T A C C G C G A G C T T A C C T C G C C A C - 1339
CLCuV-Pk1 1241 G T C C A A T T A T G A C A C G C G C G C G T G T G T A C C C T G G G A G G T A S T A C C A C T A C G C T A C G A G A G C C T A G C T A C C G C G A G C T T A C C T C G C C A C - 1340
CLCuV-Pk2 1239 G T C C A A T T A T G A C A C G C G C G C G T G T G T A C C C T G G G A G G T A S T A C C A C T A C G C T A C G A G A G C C T A G C T A C C G C G A G C T T A C C T C G C C A C - 1338
BYVMV-In 1243 G T C C A A T T A T G A C A C G C G C G C G T G T G T A C C C T G G G A G G T A S T A C C A C T A C G C T A C G A G A G C C T A G C T A C C G C G A G C T T A C C T C G C C A C - 1342
OLCV-Sd 1242 G T C A A - - - - - T T A C A C G C G C G C G T G T G T A C C C T G G G A G G T A S T A C C A C T A C G C T A C G A G A G C C T A G C T A C C G C G A G C T T A C C T C G C C A C - 1339
AYVV-Sg 1239 A T C A A - - - - - T T C G A C A C G C G C G C G T G T G T A C C C T G G G A G G T A S T A C C A C T A C G C T A C G A G A G C C T A G C T A C C G C G A G C T T A C C T C G C C A C - 1336

CLCuV-G1 1340 C T T A A T A T T 1350
CLCuV-Pk1 1341 C T T A A T A T T 1351
CLCuV-Pk2 1339 C T T A A T A T T 1349
BYVMV-In 1343 C T T A A T A T T 1353
OLCV-Sd 1340 C T T A A T A T T 1350
AYVV-Sg 1337 C T A A T A A T T 1347

Figure 4.2.23. Multiple alignment of the nucleotide sequence of DNA-β of 6 Begomovirus isolates. Blocks of conserved regions are shown in reverse contrast.

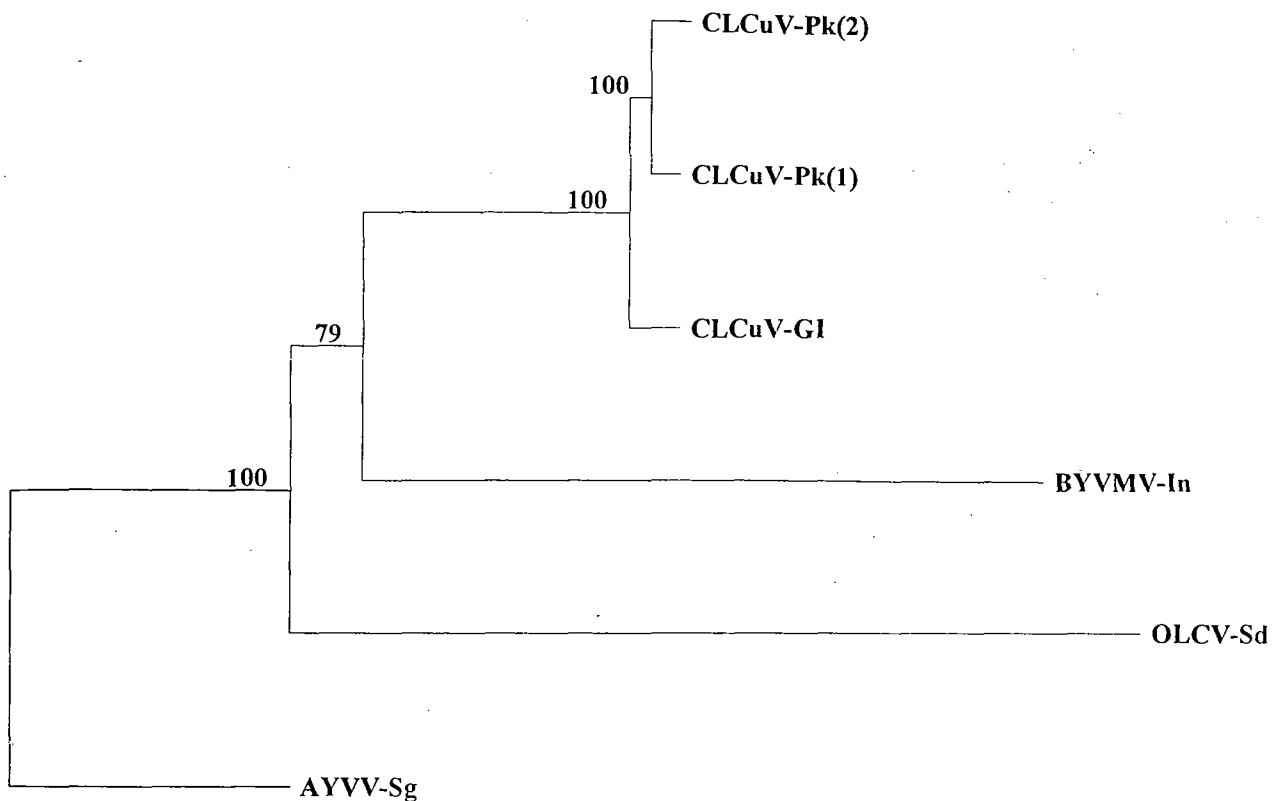


Figure 4.2.24. Phylogenetic relationship based on complete DNA- β sequence of various CLCuV isolates and other begomoviruses. The dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the AYVV-Sg sequence.

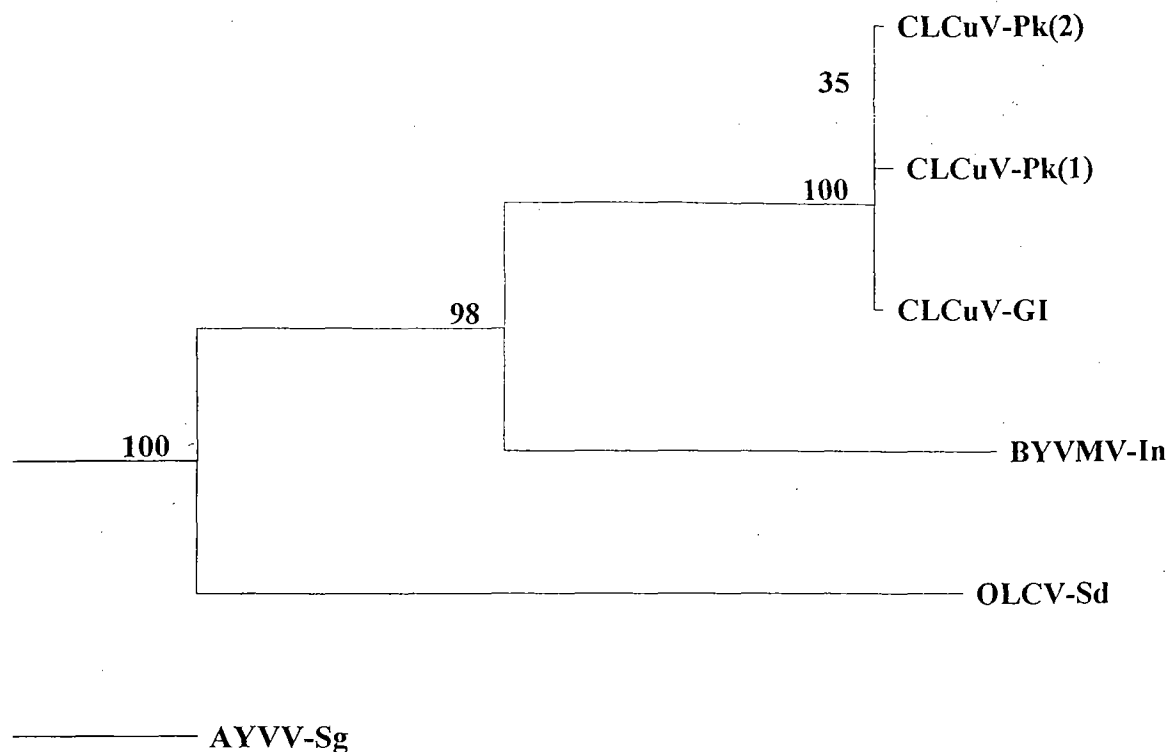


Figure 4.2.25. Phylogenetic relationship based on ORF β C1 of DNA- β between various CLCuV isolates and other begomoviruses. The dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the AYVV-Sg sequence.

investigation, 80% plants of cotton could be agroinfected, whereas, Briddon *et al.*(2001) could infect only 16% plants. This difference in the efficiency of infection could be due to the differences in the method of inoculation and the type of constructs used.

Briddon *et al.*(2001) used monomer of DNA-A and DNA- β coated onto gold particles for inoculating cotton plants by biolistic method, whereas, in the present investigation full length tandem dimeric constructs of DNA-A and DNA- β in the binary vector pBin19 introduced into EHA105 strain of *Agrobacterium tumefaciens* were used for inoculating cotton and Frenchbean plants. The efficiency of agroinoculation of Frenchbean plants was even better, although the virus could not be back transmitted to cotton by whitefly. It is difficult to explain, as whiteflies have been shown to transmit CLCuV to cotton earlier (Chapter 4.1). Agroinfection of cotton and back transmission of the progeny virus, however, confirms Koch's postulates that CLCuD is caused by coinfection with CLCuV DNA-A and DNA- β for the first time. Similarly, Saunders *et al.* (2000) have proved that Ageratum yellow vein disease (AYVD) is caused by the co-infection of AYVV DNA-A and DNA- β by agroinoculation and also back transmitted the virus agroprogeny through whiteflies. BYVMV-In is the only other virus with which DNA- β has been found associated (Jose and Usha, 2002), but information about its role in causing the disease is not known.

The relationship between CLCuV-GI and other CLCuV isolates was the most essential aspect of the study as it was expected to give a clue to the origin of the disease epidemics in India. The present isolate of CLCuV differs from all earlier reported CLCuV isolates in sequence identity of DNA-A by more than 11%, clearly indicating that it is a new species of begomoviruses as per the ICTV (Rybecki *et al.*, 2000) criteria of defining species of begomoviruses. Thus there are five distinct

species of begomoviruses causing cotton leaf curl disease, represented by various isolates studied so far. (Table 4.2.9).

Although none of the earlier isolates of different species of CLCuV have more than 90% homology in DNA-A, comparison of different ORFs, indicated that AV1 is the most conserved gene of CLCuV as it had >95% homology with various isolates of CLCuV and other begomoviruses infecting malvaceous hoots (OYVMV and BYVMV). This gene also had >90% sequence identity with ICMV, PaLCV and two isolates of ToLCVs.

All the other ORFs have >90% sequence identity only with CLCuV isolates indicating similarity among various CLCuVs. Most diverse region was found to be IR which had >90% homology only with the isolates of CLCuV-Pk2 (Table 4.2.5), whereas, sequence identity with other CLCuV isolates was <58%. The iteron sequence and their arrangement also resembled the isolates of CLCuV-Pk2 (Table 4.2.6). Similar affinity is also observed in ORFs AC1 and AC4. These ORFs had <87% homology with all the other isolates of CLCuV as well as other begomoviruses. In contrast to IR and AC1, ORFs AC2 and AC3 of the present isolate had >97% homology with the isolates of CLCuV-Pk1 and <87% identity with the other isolates. These affinities indicated that the present isolates may be a recombinant of CLCuV-Pk1 and CLCuV-Pk2, with CLCuV-Pk1 contributing ORFs AC2 and AC3 and CLCuV-Pk2 contributing IR and ORFs AC1 and AC4 (Fig. 4.2.26).

The ORFs AV1 could have originated from any of the CLCuVs or malvaceous begomoviruses, but as it has >97% sequence identity with CLCuV-Pk3, this species may have been the main source of ORF AV1. Judging from maximum sequence identity, it appears that CLCuV-Pk1 might have contributed ORF AV2.

The significant feature of sharing identity at 5' half of the intergenic region and the sequence at the N' terminal sequence of Rep, that controls binding of the Rep

Table 4.2.9. Comparison of CLCuV-GI DNA-A identity with all the CLCuV isolates

Begomovirus*	Isolate	% identity of DNA-A sequence with the present isolate
Cotton leaf curl virus- Pakistan 1 (CLCuV-Pk1)	26	85.6
	62	85.7
	Fai 2	85.6
	Mul	85.9
	Pk	86.8
	311Pk	86.6
Cotton leaf curl virus- Pakistan 2 (CLCuV-Pk2)	Fai 1	88.2
	806bPk	87.8
	72bPk	87.6
Cotton leaf curl virus- Pakistan 3 (CLCuV-Pk3)	804a	77.6
	802aPk	77.5
Cotton leaf curl virus-Sudan 1 (CLCuV-Sd1)	Gez	69.7
	Cot	69.2
	Sida	68.9
	Okr/Gez	68.9
	Okr/EG	69.5
Cotton leaf curl virus-India 1 (CLCuV-In1)	Sri Ganganagar, Rajasthan	100.0

* Names given as per the procedure followed by ICTV seventh report (Rybicki *et al.*, 2000)

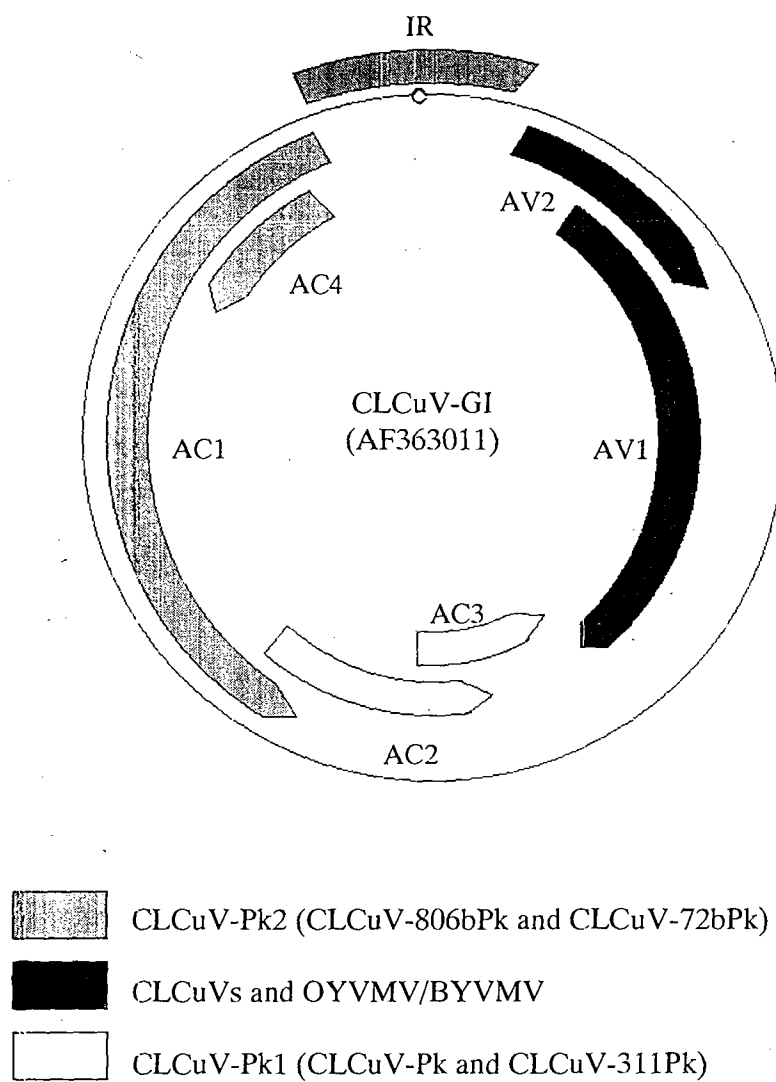


Figure 4.2.26. Probable origin of various ORFs in CLCuV-GI DNA-A

protein to the iteron across the isolates is also seen in CLCuV-GI. The cluster CLCuV-GI, CLCuV-806bPk, CLCuV-72bPk, show high identity both at N' terminal Rep and 5' half of the intergenic region indicating their relatedness. It should be noted that CLCuV-GI IR nucleotide sequence identity was the highest (91.9% each) with these isolates i.e., CLCuV-806bPk and CLCuV-72bPk, which clearly shows that the identity of the IR coincides with identity of the Rep sequence.

DNA- β molecule associated with the present isolate has >93% sequence identity with two DNA- β molecules associated with CLCuV-Pk1. However, all the DNA- β molecules associated with CLCuVs are distantly related to DNA- β associated with BYVMV, OLCV and AYVV suggesting considerable diversity in DNA- β . Actual role of DNA- β has not been determined so far. In the present investigation neither DNA-A nor DNA- β could infect plants by individual agroinoculation. Briddon *et al.*, (2000), however, reported infection of cotton as well as *N. benthamiana* by inoculation with DNA-A, which developed atypical symptoms. The typical symptoms developed only when the plants were coinfecting with DNA-A and DNA- β . Therefore, DNA- β seems to be playing important role in symptom development in plants.

CLCuV seems to be still evolving and developing recombinants by swapping fragments. The present isolate appears to have evolved as a recombinant of the CLCuVs which had been spreading in Pakistan causing serious economic losses. The recombinant must have moved across the border from Pakistan and spread in the cotton fields near the international border in Sri Ganganagar. CLCuV was not known to occur in the Indian sub-continent until late 1960s when it was first spotted in some pockets in Pakistan. However, bhendi (okra) yellow vein mosaic disease has been known in the region for the long time (Kulkarni, 1924; Varma, 1963). AV1 gene of

CLCuV has >97% sequence identity with OYVMV. It is possible that CLCuV might have evolved from OYVMV through gradual genomic changes.

Comparison with all other begomoviruses from Indian sub-continent, Africa, Asia and Mediterranean region established the geography based affinities. Phylogenetic analyses, based upon alignments of the sequences of the DNA-A and DNA- β of CLCuV-GI with the DNA-A and DNA- β genomic component of other selected begomoviruses shows that CLCuV-GI is most closely related to CLCuV isolates from Pakistan. Therefore it seems probable, that CLCuV originated in this region (Indian sub-continent) and was not introduced from elsewhere.

4.3 ASSOCIATION OF NOVEL DNA MOLECULES WITH COTTON LEAF CURL VIRUS

Cotton leaf curl virus (CLCuV) is emerging as a complex *Begomovirus*. DNA-A of CLCuV resembles the genome of monopartite begomoviruses (Briddon *et al.*, 2000), but unlike other monopartite begomoviruses, plants infected with CLCuV DNA-A do not develop symptoms (Briddon *et al.*, 2000; Chapter 4.2), indicating requirement of additional components for disease induction. Association of different types of apparent genomic components of CLCuV has been detected by various workers. Mansoor *et al.* (1999) found association of a single stranded, circular DNA molecule, approximately a half the size of CLCuV DNA-A and designated it as DNA-1 component. It shares no sequence homology with DNA-A, indicating that it is not a defective geminivirus component and appears to encode a single protein similar to the replication-associated protein (Rep) of nanoviruses. DNA-1 is capable of self-replication in plant cells but requires CLCuV DNA-A for spread in plants as well as for vector transmission. Briddon *et al.* (2001) and in the present investigation (Chapter 4.2) association of another circular single stranded DNA molecule, which is also approximately a half the size of CLCuV DNA-A and resembles DNA- β component of *Ageratum* yellow vein virus (AYVV; Saunders *et al.*, 2000) has been found. Apart from the nonanucleotide sequence TAATATTAC, which is common to all geminiviruses and contains the initiation site of rolling circle replication, DNA- β of CLCuV shows negligible sequence homology to CLCuV DNA-A or to DNA-1 or to DNA-B associated with bipartite begomoviruses (Briddon *et al.*, 2001; Chapter 4.2). DNA-A and DNA- β have also been shown to form recombinant molecules (Briddon *et al.*, 2001). Cotton plant develop typical cotton leaf curl disease (CLCuD) when inoculated with CLCuV DNA-A + DNA- β but not when DNA- β is replaced by DNA-1 indicating involvement of DNA- β in CLCuD (Briddon *et al.*, 2001).

Small sub-genomic DNA species are also often associated with geminivirus infection (Stanley and Townsend, 1985; Chatterji, 1994). These DNAs are derived by partial deletion of the wild type viral genome and, as such, show a high degree of sequence homology to the helper virus. Liu *et al.*, (1998) have reported the presence of several circular ds DNA molecules, also called as defective DNA molecules, in the CLCuV infected plants that were mostly about a half the size of CLCuV DNA-A. These molecules appeared to be derived from CLCuV DNA-A by sequence deletion, duplication, inversion and rearrangement and, in a few instances, insertion of sequences of unknown origin. All such defective DNA molecules had the intergenic region of CLCuV DNA-A and 5' part of the AC1 (Rep) gene (Liu *et al.*, 1998). Similar defective DNA forms have also been shown to be present in the case of Ageratum yellow vein disease (Stanley *et al.*, 1997) and in the tobacco plants infected with tobacco leaf curl virus (TLCV; Zhou *et al.*, 2001).

Here association of novel DNA molecules (NDMs), which differ from DNA-A, DNA- β and DNA-1, in addition to sub-genomic DNAs, with CLCuD is reported. NDMs are not derived from the helper virus genome as these do not have even the intergenic region of CLCuV DNA-A.

MATERIALS AND METHODS

Virus isolate

CLCuV was isolated from a *G. hirsutum* plant showing severe disease symptoms at a farm in Sri Ganganagar (Rajasthan), and the isolate was designated as CLCuV-GI. It was maintained on *G. hirsutum* cv. LH900 by whitefly transmission in an insect-proof glasshouse.

PCR cloning

a) Primer designing

A set of abutting primers were designed to amplify full length genomic components of CLCuV based on the published nucleotide sequence of CLCuV DNA-A (GenBank Acc. No.- AF363011, AJ002447, AJ002448, AJ002449, AJ002452, AJ002455 and AJ132430). The following primers were designed to amplify the intergenic region of CLCuV DNA-A flanking stem-loop sequence having a part of nonanucleotide sequence at their 5' end.

1. CRV-AV : 5'-TTACCGGATGGCCGCGGATTT-3' (22 nucleotides)
2. CRC-AC : 5'-TATTAAACGGATGGCCGCTTT-3' (21 nucleotides)

b) PCR amplification

For PCR amplification of CLCuV DNA components from infected cotton plants total DNA was isolated from 100mg of leaf tissue using DNeasy Plant Mini Kit (Qiagen Inc., USA). 200ng of total DNA was taken for PCR amplification. Similarly isolated DNA from healthy cotton plants served as negative control.

For the amplification of CLCuV DNA component from viruliferous whiteflies, a batch of 10 non-viruliferous adult whiteflies were given an acquisition access period (AAP) of 24h on CLCuV infected cotton (cv. LH900) plants. After AAP, total DNA from these viruliferous whiteflies was isolated using DNeasy Tissue Kit (Qiagen Inc., USA). 250 ng of total DNA was taken for PCR amplification. Similarly DNA was extracted from a batch of 10 adult whiteflies which were fed on healthy cotton plants for use as negative control.

The PCR reaction mixture consisted of 2.5U of *Pfu* DNA polymerase (MBI Fermentas), 5µl of 10x PCR buffer (without MgCl₂), 3µl of 25mM MgCl₂, 1µl of 10mM dNTP and 1µl of 100ng/µl CRV-AV and CRC-AC primers each for the amplification of CLCuV DNA components in the infected cotton plants and

viruliferous whiteflies. Total volume was made upto 50µl using sterile double distilled water.

PCR amplification was done with the following programmes in a PHC-3 (Techne) thermal cycler. First denaturation at 94°C for 2mins, then 30 cycles of denaturation (94°C for 1min), annealing (55°C for 2min), extension (72°C for 4min). At the end of the cycles, a final extension step at 72°C for 5 min and then the temperature was decreased to 4°C until the reaction mixtures were removed. PCR products were analysed by electrophoresis in 1% agarose gel, stained with ethidium bromide (0.5µg/ml) and photographed on a UV transilluminator.

c) Cloning

Individual amplicons from the infected cotton plants were eluted from the gel using Gel Extraction Kit (Qiagen Inc., USA), and cloned according to Sambrook *et al.* (1989) into *Sma*I restricted pUC18 vector. *Escherichia coli* strain DH5α was transformed with the recombinant plasmid by CaCl₂ method (Mandel and Higa, 1970).

Transformants containing recombinant plasmids were selected on Luria agar plates containing ampicillin (50 µg/ml), X-gal (2%) and IPTG (100 mM) (Vieira and Messing, 1982). Presence of the insert was confirmed by PCR, as described above. The amplified products were analyzed by electrophoresis in 1% agarose gel as described earlier and the amplicons were compared with the marker 1kb DNA ladder (MBI Fermentas, USA).

Extraction of DNA from plants for Southern hybridization

To detect viral DNA components in cotton and various experimental hosts (*Abelmoschus esculentus*, *Althaea rosea*, *Lycopersicon esculentum*, *Nicotiana benthamiana*, *Phaseolus vulgaris* and *Physalis floridana*), which were inoculated by CLCuV-GI using whiteflies in the glasshouse, total nucleic acid from 1g of the leaf

tissue of inoculated experimental hosts was extracted (Dellaporta *et al.*, 1983). The DNA pellets were re-suspended in 50µl sterile double distilled water and used for Southern hybridization.

Southern hybridization

The DNAs extracted from infected cotton and experimental hosts were electrophoresed in 1% agarose gel and transferred onto nitrocellulose membrane (NCM; Amersham, Inc., USA). Similarly the PCR amplified products from infected cotton and viruliferous whiteflies were also electrophoresed in 1% agarose gel and transferred onto NCM. Standard procedure of Southern hybridization was followed (Sambrook *et al.*, 1989). To confirm the viral nature of the clones, the inserts were released from the plasmid vector clones pEDC2 and pEDC1.5 and used for probe preparation. The CLCuV specific DNA A probe was prepared from a full length CLCuV-GI DNA-A (GenBank Acc. No. AF363011) clone pCLCA after incorporating ($\alpha^{32}\text{P}$) dCTP by random priming method (Feinberg and Vogelstein, 1983), using Random Primer Labelling Kit (Genei Pvt Ltd., Bangalore). Radiolabelled probe was added @ 0.5×10^6 dpm/ml of prehybridization solution. Prehybridization and hybridization was done at 65°C for 4h and 18h, respectively. Hybridized blots were washed in 2x SSC (0.3M Sodium Chloride and 0.03M Sodium Citrate, pH 7.4); 0.1% SDS thrice at 65°C for 15min each time. Autoradiographs were prepared by exposing Kodak X-ray film for 24-48h at 70°C.

Sequencing

The clones were sequenced using M13 Forward and M13 Reverse primers. DNA sequencing was carried out using Taq polymerase (Sequencing Kit, MBI Fermentas) by dideoxy chain termination method (Sanger *et al.*, 1977). The products were run in Pharmacia LKB macrophore system. Each sequence was generated thrice for a single reaction. The sequence data were assembled and analysed using the

software programme Bioedit version 5.0.9. Multiple nucleotide sequence alignments were done in CLUSTALW(1.7) programme (<http://www.ebi.ac.uk/clustalw>). Comparisons were made with sequences obtained from GenBank nucleotide sequence database.

RESULTS

Cloning and sequence analysis of extra DNA components

PCR amplification of DNA from CLCuV infected plants and viruliferous whiteflies, with abutting primers (CRV-AV+CRC-AC), resulted in amplicons varying from 0.6 kb to 2.7 kb in infected cotton and viruliferous whiteflies (Fig.4.3.1). No such amplicon was observed in DNA from healthy cotton and non-viruliferous whiteflies. Various amplicons from infected cotton plants were cloned in *Sma*I restricted pUC18 vector and the recombinant DNAs analysed by PCR (Fig.4.3.2). Recombinant clones having inserts of 0.6 kb, 1.1 kb, 1.3 kb, 1.5 kb, 2.0 kb and 2.7 kb were obtained (Fig.4.3.2).

Partial nucleotide sequences were generated for these clones and analysed by BLAST search (www.ncbi.nlm.nih.gov/BLAST). The clone with 2.7 kb insert was found to be DNA-A and 1.3 kb clone as sub-genomic DNA-A as the sequences were identical to the CLCuV-GI DNA-A intergenic region (GenBank Acc No. AF363011). The remaining four clones having 0.6, 1.1, 1.5 and 2.0kb inserts, however had sequences of unknown origin except for nonanucleotide sequence aligning with begomoviruses. Their sequences (Fig.4.3.3) were compared by software programme Bioedit version 5.0.9. The clone pNDM0.6 was found to be sub-genomic of clone pNDM2 showing 98.8% sequence identity (Table 4.3.1). The other three clones had <50% sequence identities (Table 4.3.1 and 4.3.2). For further studies two clones, with inserts of 2kb (pNDM2) and 1.5kb (pNDM1.5) were selected.

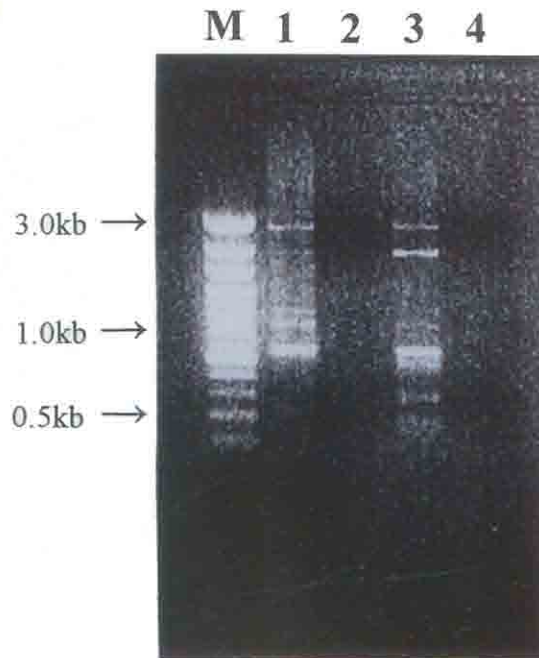


Figure 4.3.1. PCR amplification of novel DNA molecules from total DNA of infected cotton and viruliferous whiteflies using CRV-AV+CRC-AC primers
Lane 1: Infected cotton, 2: Healthy cotton, 3: Viruliferous whiteflies, 4: Non-viruliferous whiteflies, M: 100bp DNA ladder.

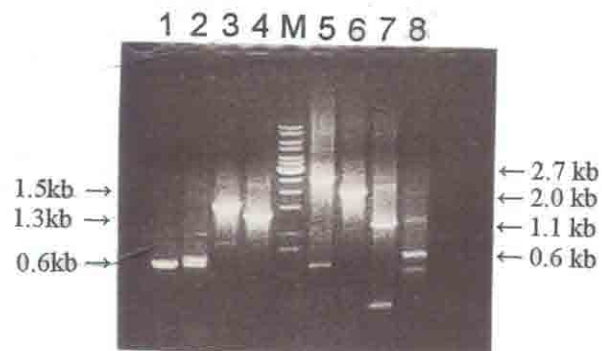


Figure 4.3.2. PCR amplification of novel DNA molecules from different clones in pUC18 using CRV-AV+ CRC-AC primers.
Lane 1: ~0.6 kb, lane 2: ~0.6 kb, lane 3: ~1.5 kb, lane 4: ~1.3 kb, lane 5: ~2.7 kb, lane 6: ~2.0 kb, lane 7: ~1.1 kb, lane 8: ~0.6 kb, M: 1kb DNA ladder.

2kb NDM

TCAGCTATCGTGA CTATACAGATGCTATGATGACTTGTGCTAGCGGTGTTATGACTTGCGAGA
CTGCTCTATATAATGTAAAATAAGTGTACAGTGTTCTTTAAACTTGACGAGCGACTGTTCTTGT
GAGTTTCGAGATGCACCAGTTTGC GCGTGATGTTGCCAAAGTCGTTCGAAGAGTTACTCGTTA
CAAGTAAAGGGTGGTACACAACAAGTTTGACATGCGTTAAACGATATCGACTCGGTTTCGCAG
TTCTAACTCCAATGGTCTAAAAATATAGCGTCCTTTGACAGTACCATTGCAAAGCGGCCATCC
GTTTAATATTACCGGATGGCGCGCGGATTTTTTAGTAGCAGGGGGTAGGGGCGTGAGCGTAG
GGTAGGAAAGCAGTTTTTGTGGATTGTGAGTCGGCCTTTACGATCTCAGGGAAGTATGACT
CTCATTTCTGTTACGCTTAGGTAAAGACTAACCTAAGAACATAAAATACCGAAGTGTATTAGC
TGTTCCATTATTTCTAATGTCAA

1.5kb NDM

ATCAGAGACTTAGGACAGAACAGTCAGCGACAAACGCAAGGAAATTCTTTCTCTCCTTCCTTT
CTTTCTGATTGTTTCTTCGTTCCGGTAAAACTCACAAGTTTGC GTAAAGGAGTCGAGGGACA
CATCGCATCGTGACAGGTTCTGTCCTCTGTCCATCTTGTGTAATTTAAAGTAAATGTAGAAGA
AAACTGCCGTGGTAAGGAGTAATGCCTATGAATTTTCCAGAGTTGCCAAATTTCCCTTGATAA
AACATGTATTTTGCACAACATTTATGCGTATATTTCTTGAAATTTT CAGATATTTTAGATTAA
ATTGCGTAGAAAATTGTCCGAAAATTTTGGAAAATTATATTCACGATTTTCCCAGTAAATTCG
GTTTTTATCGAAGGAACTTGGCAAACCTCTGAAGGCCCATACGGCGTTCTTCCTTAGCACGGC
AGAAAAGGCGCAGAAAGATTCTTTCATCCGTACATACTGTTTTATCTCATTCTTTATTTCCGTA
AGCTCTCCCGTTTCCAACCTCATTCTGTTTGTTTAACTATTTAAAGCAGCCATCCGTTTAATA
TTACCGGATGGCCGCGCGGATTTGAAAGTGGACGAAAAACTCATGTGAGATAGGAAACATGCT
ATAGTCAAGAACATGCCACGTTGAAATCTTAAATTTTCTGTTTTGCTTCGGACAAGACGCTG
ATAGCAACATCATGAGTTATAATGCGGTACCCCAAGTAGCAATGACCTTTTAAACATTTTTC
AAAAGCTCTCAAAAAGATGTTAAAATGTTCTGATTAGGAAACCCTTTTTGTATAATTTCTACA
AGTAAAATTCAAGGGAAGAGTGTTAAGCAAAACTCTAAAAATGCGAGTATCAATGTAGATTT
TAACCTTTTTTTCAAACATTTTTAGTGTTGACAAATACAGATATTTTTTACCATAACATTTTATT
CCTCGTAAATTTTACTTAAAAAATAGCCTTAAAAAGAGTTTCTTAAAACTGACATTGTAACTTTT
TTTTATAAAGTTGTCTTTTTGATATTTCAAGAGCTTTTTTAAAGTTTCTCTTCAAAGTCGTTCTT
ACTAGGGAGACATTTTTCTAACAACGACTCTATTGCACGTGAATTTTTTCGA

1.1kb NDM

TTACCGGATGGCCGCGCGGATTTGCAATGGTACTGTCAGAGACGCTATATTTTTAGACCATTGG
AGTTAGAACTGCGAAACCGACTGGATATCGTTTAAACGCATGTCAAACCTTGTGTGTACCACCC
TTATCTTGTAACGAGCTAATTCCGAACGAACCTGGCACATCACGCGCACT

0.6kb NDM

TTGCGCGTGATGTTGCCAAAGTCGTCGAAGAGTTAGCTCGTTACAAGTAAAGGGTGGTACACA
ACAAGTTTGACATGCGTTAAACGATATCGACTCGGTTTCGCAGTTCTAACTCCAATGGTCTAA
AAATATAGCGTCCTTTGACAGTACCATTGCAAAGCGGCCATCCGTTTAATA

Figure 4.3.3. Nucleotide sequence of the viral sense strand of CLCuV novel DNA molecules (NDMs). Nonanucleotide sequence are in bold and primer sequences are underlined.

Table 4.3.1. Comparison of nucleotide sequence identity in the upstream region of nonanucleotide among the NDMs.

NDM	0.6kb NDM	1.5kb NDM	2.0kb NDM
0.6kb NDM	100		
1.5kb NDM	48.0	100	
2.0kb NDM	98.8	48.0	100

Table 4.3.2. Comparison of nucleotide sequence identity in the downstream region of nonanucleotide among the NDMs.

NDM	1.1kb NDM	1.5kb NDM	2.0kb NDM
1.1kb NDM	100		
1.5kb NDM	47.0	100	
2.0kb NDM	36.7	39.5	100

Viral nature of the NDMs

Viral nature of the two NDMs was shown by Southern hybridization. Radiolabelled probes to full length CLCuV-GI DNA A (pCLCA), pNDM2 and pNDM1.5 were used to differentiate DNA population in the infected plants and viruliferous whiteflies.

Southern analysis of DNA extracted from infected cotton and experimental hosts were done with all the above probes (Fig.4.3.4). Typical viral replicative forms were seen with pCLCA probe. Double stranded, open circular forms were in higher concentration than super coiled and single stranded DNA (Fig.4.3.4a). With pNDM2 probe, strong hybridization was obtained with bands of ~ 0.6 kb and moderate hybridization with ~ 2.0 kb band (Fig.4.3.4b). These bands were distinct from those, which hybridized with pCLCA probe. With pNDM1.5 probe, strong hybridization was obtained only with the ~ 1.5 kb band (Fig.4.3.4c).

To examine whether these NDMs were encapsidated or not, Southern analysis was done with the PCR amplified products of DNA extracted from viruliferous whiteflies and infected cotton plants (Fig.4.3.5). With DNA-A (pCLCA) probe, hybridization with bands at viral genome length (2.7kb) and sub-genomic DNA-A components was obtained (Fig.4.3.5a), whereas, pNDM2 probe hybridized with bands at ~ 2 kb and ~ 0.6 kb, indicating that these NDMs are associated not only with CLCuV affected plants but also with viruliferous whiteflies suggesting their encapsidation (Fig.4.3.5b). The infected cotton plants seem to have larger amounts of 0.6kb NDM compared to 2kb NDM. In the whiteflies, however, the amount 2kb NDM appeared to be relatively more than that of 0.6kb NDM. No hybridization was obtained with DNA from healthy cotton plant and non-viruliferous whiteflies with the two probes.

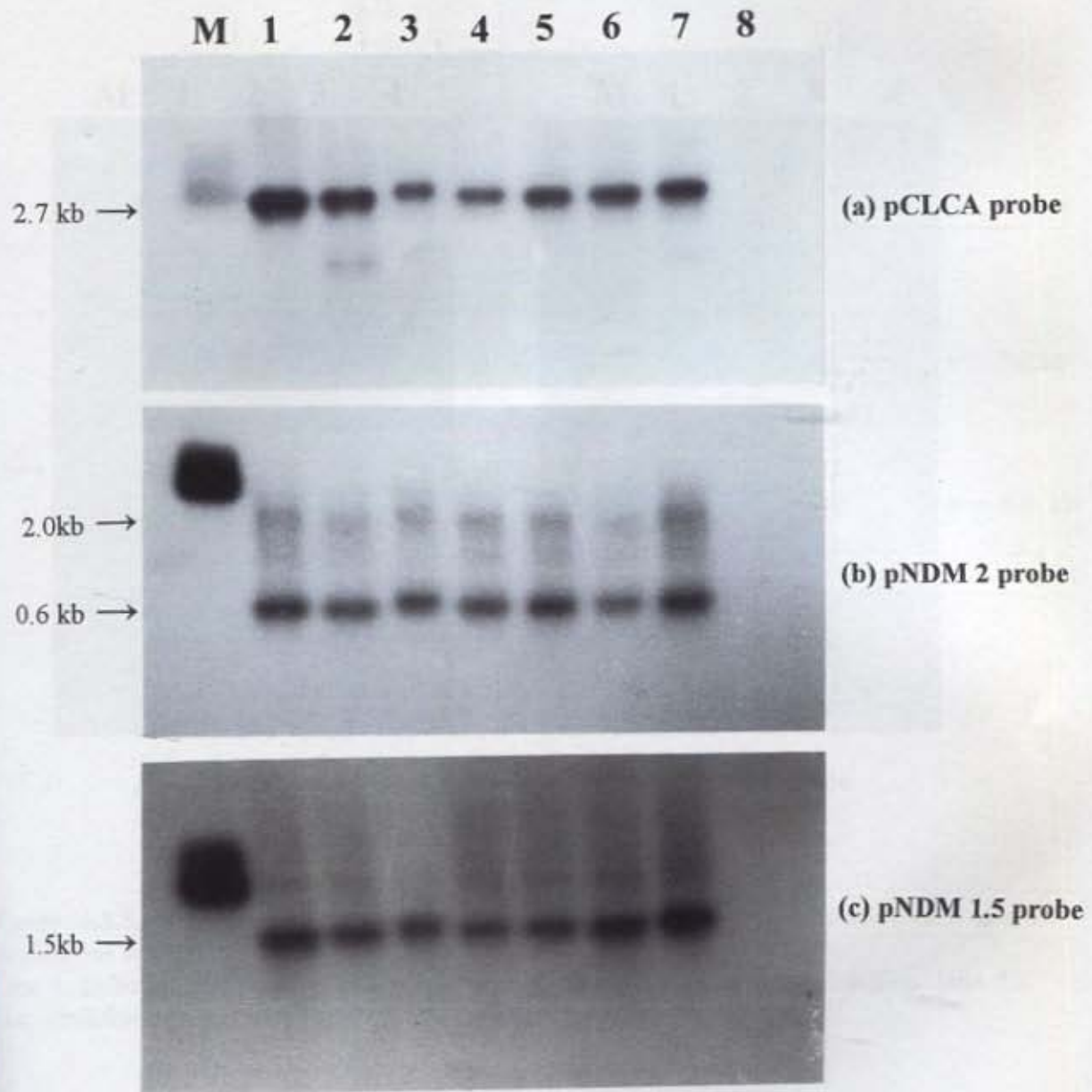


Figure 4.3.4. Southern analysis of infected cotton plants and test hosts.
 Lane 1: Cotton, lane 2: Frenchbean, lane 3: Hollylock, lane 4: Bhendi, lane 5: *Physalis floridana*, lane 6: *N. benthamina*, lane 7: Tomato, lane 8: Healthy cotton,
 Lane M: 1 kb DNA ladder.

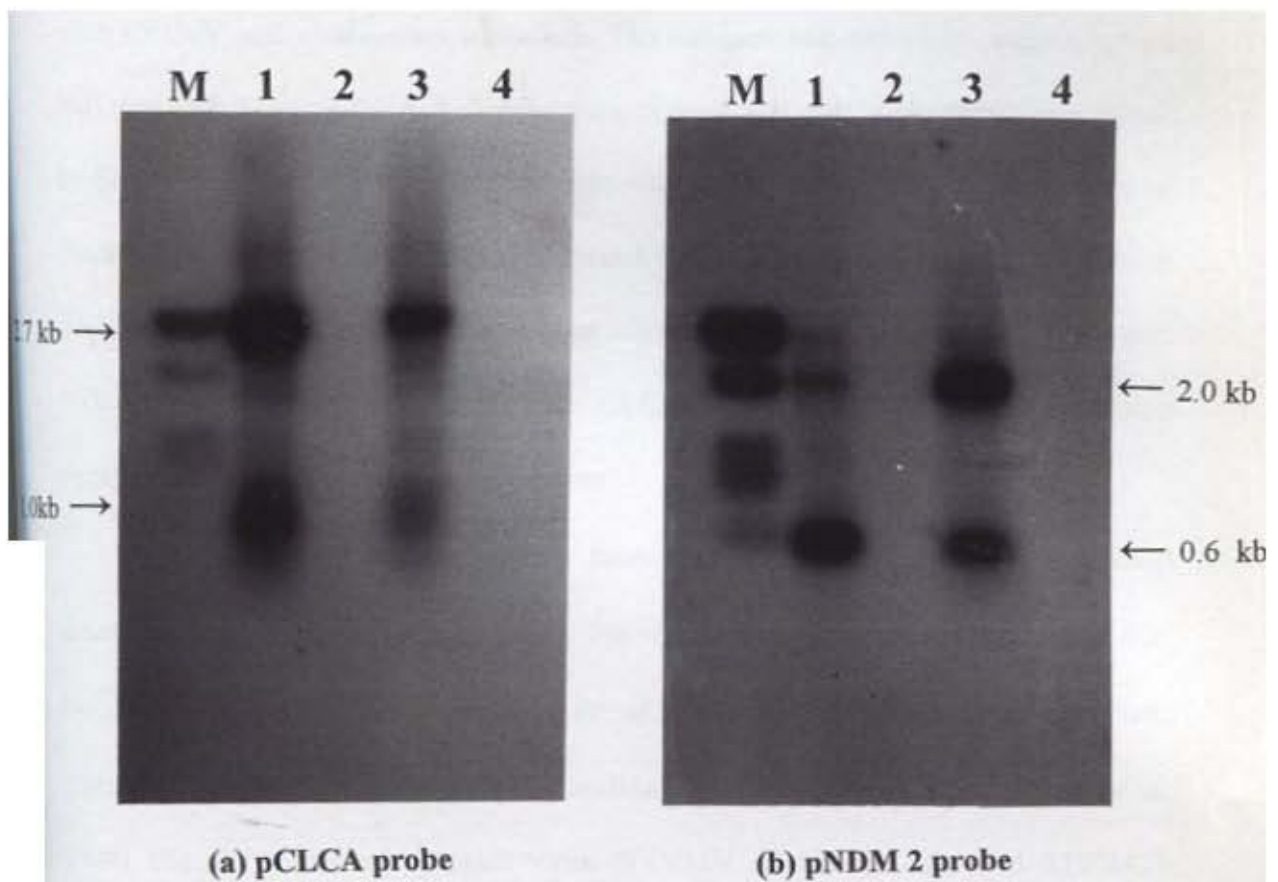


Figure 4.3.5. Southern analysis of PCR amplified products of infected cotton and viruliferous whiteflies.

Lane 1: Infected cotton, lane 2: Healthy cotton, lane 3: Viruliferous whiteflies, lane 4: Non-viruliferous whiteflies, lane M: 100 bp DNA ladder.

DISCUSSION

Both, sub-genomics and NDMs were found associated with plants infected with CLCuV and viruliferous whiteflies. The sub-genomic molecules were nearly a half size of full-length DNA-A. Viral nature of the 1.5kb and 2.0kb NDMs was shown by Southern hybridization, as hybridization was obtained only with the DNA samples from CLCuV infected plants and viruliferous whiteflies only and not with DNA from healthy cotton plants and non-viruliferous whiteflies. These findings also indicate that NDMs are not only associated with CLCuV in infected plants but are also encapsidated and transmitted by whiteflies.

NDMs do not have sequence homology with the sequences of DNA- β associated with CLCuV (GenBank Acc. No. AY083590, Chapter 4.2; GenBank Acc. No. AJ292769; AJ298903, Briddon *et al.*, 2001;) or DNA- β associated with Ageratum yellow vein virus, AYVV (GenBank Acc. No. AJ252072, Saunders *et al.*, 2000), bhendi yellow vein mosaic virus, BYVMV (GenBank Acc. No. AJ308425, Jose and Usha, 2002) and okra leaf curl virus, OLCV (GenBank Acc. No. AF397217). The sequences of NDMs also do not have any sequence homology with CLCuV DNA-1 component (Mansoor *et al.*, 1999), with the defective DNA molecules (Liu *et al.*, 1998) or with the DNA-A/DNA- β recombinant molecules (Briddon *et al.*, 2001). All the defective DNA molecules and DNA-A/DNA- β recombinant molecules contain the intergenic region of CLCuV DNA-A (Liu *et al.*, 1998; Briddon *et al.*, 2001). In contrast, NDMs reported here do not have the CLCuV DNA-A intergenic region except for the primer sequence including nonanucleotide sequence TAATATTAC, which is common to all geminiviruses and contains the nick site/initiation site of rolling circle replication. These NDMs, also do not have the replication-associated protein (Rep) binding sequences (iterons) of CLCuV-GI DNA-A, indicating that there might be some other sequences in the upstream of

nonanucleotide sequence which is recognized by the Rep protein, for the replication of NDMs in infected plants. NDMs reported here display characteristics analogous to the RNA satellites (Roossnick *et al.*, 1992). Such molecules have not been found to be associated with any other geminivirus. The role of these NDMs in the CLCuD is not yet understood.

4.4 MOLECULAR DETECTION OF DNA- β ASSOCIATED WITH WHITEFLY TRANSMITTED GEMINIVIRUSES IN INFECTED PLANTS AND VIRULIFEROUS WHITEFLIES

Whitefly transmitted geminiviruses (WTGs; genus *Begomovirus*, family *Geminiviridae*) have emerged as serious threat to the cultivation of crop plants in various parts of the world (Varma *et al.*, 1997). WTGs have unique twin particle morphology of their virions and mostly have bipartite genome designated as DNA-A and DNA-B. Some of the WTGs, however, have monopartite genome and a few of these have also been shown to depend on a satellite-like ssDNA component designated DNA- β (Saunders *et al.*, 2000). The cotton leaf curl disease (CLCuD), first noticed in a few plants of *G. barbadense* in the experimental fields of IARI in 1989 (Anonymous, 1990; Varma *et al.*, 1993) and later on *G. hirsutum* near Sri Ganganagar in Rajasthan in 1993 (Ajmera, 1994), has spread to all the major cotton growing areas of Punjab, Haryana and Rajasthan and is a threat to cotton cultivation in the region (Varma, 1995). CLCuD is caused by a DNA- β associated monopartite geminivirus - *cotton leaf curl virus* (CLCuV; Mansoor *et al.*, 1993; Briddon and Markham 2000; Radhakrishnan *et al.*, 2001). DNA- β has also been shown to be associated with yellow vein disease of *Ageratum conyzoides* (Saunders *et al.*, 2000), which also occurs in India (Varma, 1963) but its association with DNA- β has not been shown.

Here we report association of DNA-A and DNA- β , not only with plants experimentally infected with CLCuV but also in viruliferous whiteflies and various crop, ornamental and weed plants naturally infected with WTGs under field conditions. We also report occurrence of sub-genomic DNA- β molecules in plants infected with CLCuV.

MATERIALS AND METHODS

1. Detection by nucleic acid spot hybridization (NASH)

a) Detection of CLCuV in experimentally inoculated test plants

Abelmoschus esculentus cv. Parbhani Kranti (bhendi), *Althaea rosea* cv. Pusa Deepika (Hollyhock), *Lycopersicon esculentum* cv. Pusa Ruby (tomato), *Nicotiana benthamiana*, *N. tabacum* cv. White Burley, *Phaseolus vulgaris* cv. Sel. 9 (Frenchbean) and *Physalis floridana*, were experimentally infected with CLCuV-GI (Radhakrishnan *et al.*, 2001) at seedling stage using 10 whiteflies (*Bemisia tabaci*) per seedling with acquisition access period (AAP) and inoculation access period (IAP) of 48h each. The inoculated plants were maintained in an insect-proof glass house and observed for symptom development for up to six weeks. Total nucleic acid from 1g leaf tissue of individual test plants was extracted according to Dellaporta *et al.* (1983). The DNA pellet was suspended in sterile double distilled water and used for NASH tests. Three different dilutions of total DNA i.e., 1:1, 1:10 and 1:100 were dotted on the nitrocellulose membrane (NCM; Amersham Inc., USA). Total DNA extracted from infected and healthy cotton plants were used as positive and negative controls, respectively. After dotting, the membranes were dried and baked at 80°C for 2h in vacuum.

b) Detection of CLCuV in viruliferous whiteflies

For the detection of CLCuV in viruliferous whiteflies, a batch of 25 non-viruliferous adult whiteflies were given an AAP of 24h on CLCuV infected cotton (cv. LH 900) plants. After AAP, the whiteflies were pooled and crushed in 500 µl of TE buffer (100 mM Tris-HCl, pH8.0; 10 mM EDTA, pH 8.0) with the help of a Como-mini drill. The DNA from the crushed whiteflies was extracted by phenol/chloroform and chloroform/isoamyl alcohol treatment and concentrated by isopropanol precipitation (Sambrook *et al.*, 1989), the concentrated DNA was

suspended in sterile double distilled water and used for NASH. Three different dilutions of total DNA i.e., 1:1, 1:10 and 1:100 were dotted on the nitrocellulose membrane as described for plant samples. Total DNA extracted from a batch of 25 adult whiteflies which fed on healthy cotton plants served as negative control.

c) Detection of geminivirus infection in naturally infected plants

(i) Weed plants

Twelve dicot weeds growing in and around severely leaf curl affected cotton fields were collected from Sri Ganganagar (Rajasthan) and Sirsa (Haryana). Total nucleic acid from 1g of the leaf tissue of these weeds was extracted and used as described for plants above except only two different dilutions of total DNA (1:1 and 1:10) was dotted on the nitrocellulose membrane.

(ii) Crop and ornamental plants

Naturally infected crop/ornamental plants showing apparent infection by WTGs were collected from the experimental fields of Indian Agricultural Research Institute, New Delhi. Total DNA was isolated from 1g of leaf tissue as described earlier. Samples were taken from plants affected with the following diseases: Ageratum leaf crumple, Ageratum yellow vein, bhendi leaf curl, bhendi yellow vein mosaic, chilli leaf curl, cowpea golden mosaic, hollyhock enation, hollyhock yellow vein, Luffa leaf distortion mosaic, mungbean yellow mosaic and tomato leaf curl. In addition cassava mosaic caused by Indian cassava mosaic virus (ICMV) maintained in an insect proof glass house at IARI was also tested.

d) Preparation of radiolabelled probe and hybridization

The CLCuV specific DNA probe was prepared from a full-length CLCuV-GI DNA- α (GenBank Acc. No.AF363011) clone (pCLCA) and full-length CLCuV-GI DNA- β (GenBank Acc. No.AY083590) clone (pCLC β) of Sri Ganganagar isolate (CLCuV-GI) after incorporating (α - ^{32} P) dCTP by random priming method (Feinberg

and Vogelstein, 1983), using Random Primer labelling kit (Genei Pvt. Ltd., Bangalore). Radiolabelled probe was added @ 0.5×10^6 dpm/ml of prehybridization solution. Prehybridization and hybridization was done at 65°C for 4h and 18h, respectively. Hybridized blots were washed in 2x SSC (0.3 M Sodium Chloride and 0.03 M Sodium Citrate, pH7.4); 0.1% SDS, thrice at 65°C for 15 min each time. Autoradiographs were prepared by exposing Kodak X-ray film for 24-48h at -70°C.

2. Detection by polymerase chain reaction (PCR)

Considering the sensitivity and precision required, a polymerase chain reaction (PCR) based detection technique was developed for the rapid detection of CLCuV in infected plants and viruliferous whiteflies.

a) Primer designing

Based on the published nucleotide sequence of DNA-A of Pakistan isolates of CLCuV (GenBank Acc. No. AJ002449, AJ002447, AJ132430, AJ002455, AJ002448 and AJ002452), a set of primers were designed to amplify complete coat protein gene (~750 bp) of CLCuV. The two primers used were:

1. CPF: 5'-AATTATGTCGAAGCGAGCTGC-3' (21 nucleotides)
2. CPR: 5'-TAATATCAATTCGTTACAGAG-3' (21 nucleotides)

Similarly, the following set of abutting primers having a *Kpn*I site at its 5' end were also designed to amplify the second genomic component of CLCuV i.e., CLCuV DNA-β, based on the published nucleotide sequence of CLCuV DNA-β of CLCuV-Pk (GenBank Acc. No.-AJ298903; Bridson *et al.*, 2001):

1. Beta 1: 5'-GGTACCACTACGCTACGCAGCAGCC-3' (25 nucleotides)
2. Beta 2: 5'-GGTACCTACCCTCCCAGGGGTACAC-3' (25 nucleotides)

These primers were synthesized by the GENSET Singapore Biotech. Pvt. Ltd.

b) Extraction of DNA from plants and whiteflies

For the detection of CLCuV from infected plants by PCR, total DNA was isolated from 100 mg of leaf tissue using DNeasy Plant Mini Kit (Qiagen Inc., USA) according to manufacturer's instructions. 200ng of total DNA was taken for PCR amplification. Total DNA from CLCuV infected cotton and healthy cotton served as positive and negative control, respectively. For the detection of CLCuV in viruliferous whiteflies, a batch of 10 non-viruliferous adult whiteflies were given an AAP of 24 h on CLCuV infected cotton cv. LH900. After AAP, total DNA from these viruliferous whiteflies was isolated using DNeasy Tissue Kit (Qiagen Inc., USA) according to manufacturer's instructions. 250ng of total DNA was taken for PCR amplification. Total DNA extracted from a batch of 10 adult whiteflies, which were given feeding period of 24 h on healthy cotton plants was used as negative control.

c) PCR conditions

The reaction mixture consisted of 2.5 U of *Taq* DNA polymerase (MBI Fermentas, USA), 5µl of 10x PCR buffer (without MgCl₂), 3µl of 25 mM MgCl₂, 1µl of 10 mM dNTP and 1µl of 100ng/µl primers (CPF/CPR) or (Beta1/Beta2). Total reaction volume was made up to 50µl using sterile distilled water.

PCR amplification protocol was standardized employing different temperatures of denaturation, annealing and extension based on T_m values of the primers. The successful PCR amplification of the CLCuV coat protein gene and CLCuV DNA-β was obtained with the following programmes in a PHC-3 (Techne) thermal cycler. First denaturation at 94°C for 2 mins, then 30 cycles of denaturation (94°C for 1 min), annealing (55°C for 2 min) and extension (72°C for 3 min). At the end of the cycles, a final extension step at 72°C for 5 min was given and then the temperature was decreased to 4°C until the reaction mixtures were removed. PCR

products were analyzed by electrophoresis in 1% agarose gel, stained with ethidium bromide (0.5 µg/ml) and photographed on a UV transilluminator.

RESULTS

Detection of CLCuV DNA-A and DNA-β by NASH and PCR in experimentally infected plants and viruliferous whiteflies

CLCuV DNA-A and DNA-β were detected in all the experimentally infected plants (bhendi, cotton, Frenchbean, hollyhock, *N. benthamiana*, *N. tabacum*, *P. floridana* and tomato) by NASH using radiolabelled pCLCA (Fig.4.4.1a) and pCLCβ (Fig.4.4.1b) probes. In all the host plants, viral DNA could be detected up to 1:100 dilutions, although hybridization at this dilution was weaker as compared to that obtained at 1:1 and 1:10 dilutions. The level of viral DNA in infected cotton, Hollyhock, bhendi, *N. tabacum* and tomato was found to be relatively higher than in Frenchbean, *P. floridana* and *N. benthamiana* (Fig.1a). The overall concentration of CLCuV DNA-β was less than that of DNA-A in all the plants tested as DNA-β could be detected only in 1:1 dilution, except in samples from cotton and Frenchbean in which weak hybridization was also obtained at 1:10 dilution (Fig. 4.4.1b).

CLCuV DNA-A and DNA-β were also detected in viruliferous whiteflies by NASH using radiolabelled pCLCA (Fig.4.4.2a) and pCLCβ (Fig.4.4.2b) probes. In viruliferous whiteflies, however, viral DNA was detected only up to 1:10 dilution of total DNA sample from viruliferous whiteflies and below this dilution i.e., at 1:100 dilution level, no hybridization was obtained indicating small amount of viral DNA in viruliferous whiteflies.

The coat protein gene fragment of CLCuV of ~ 750bp was amplified from all the infected test hosts and viruliferous whiteflies by PCR using CPF and CPR primers. No size difference was observed between the amplicons (Fig.4.4.3a). There was no amplification in samples from healthy cotton and non-viruliferous whiteflies

Similarly, full length CLCuV DNA- β (~1.4kb) was amplified using abutting primers (Beta 1 and Beta 2 primers) from all the infected test hosts and from viruliferous whiteflies. There was no amplification in samples from healthy cotton and non-viruliferous whiteflies (Fig.4.4.3b). Unlike the uniform size of amplicons obtained by coat protein gene primers, considerable heterogeneity was observed in the amplicons obtained by DNA- β primers. In addition to the expected amplicons at ~1.4 kb, which is the full length CLCuV DNA- β , many smaller amplicons which could be sub-genomic DNA- β were also obtained. The sub-genomic DNA- β forms were maximum in the case of bhendi and *P.floridana* (four each), followed by *N.tabacum*, cotton and tomato (two each), and Hollyhock and *N.benthamiana* (one each). The sub-genomic DNA- β form in hollyhock (~0.6 kb) was amplified to a greater extent than the full length DNA- β (~ 1.4 kb). Sub-genomic DNA- β were also amplified from viruliferous whiteflies in addition to the full length DNA- β (~ 1.4 kb) amplicon (Fig.4.4.3b).

Detection of infection by WTGs in field infected crop, ornamental and weed plants by NASH using probes to CLCuV DNA-A and DNA- β

a) Detection of infection by WTG in crop and ornamental plants

CLCuV CP gene probe pCLCA probe could detect WTG infection in all the plants tested by NASH (Fig.4.4.4a). The hybridization obtained with bhendi leaf curl, bhendi yellow vein mosaic, hollyhock enation, hollyhock yellow vein, cassava mosaic and tomato leaf curl was on par with that obtained with cotton leaf curl. The hybridization, however, was relatively less strong with Luffa leaf distortion mosaic, chilli leaf curl, cowpea golden mosaic and mungbean yellow mosaic. This difference in hybridization could be due to variation in the concentration of WTGs in the naturally infected plants or differences in nucleotide sequence homology.

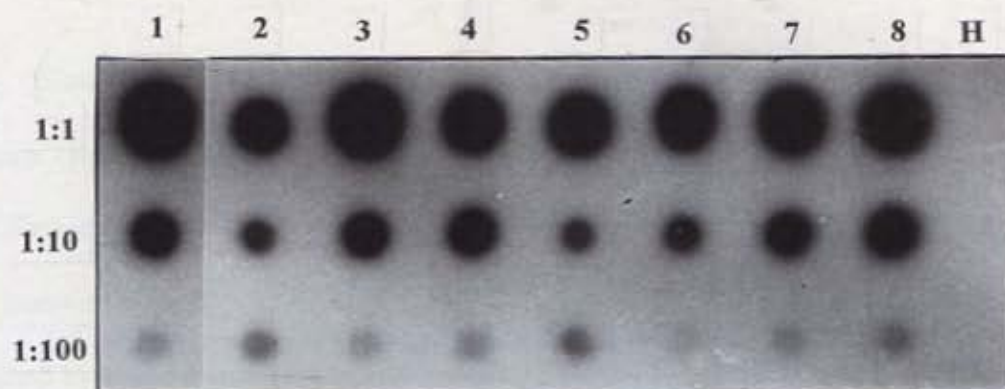


(a)

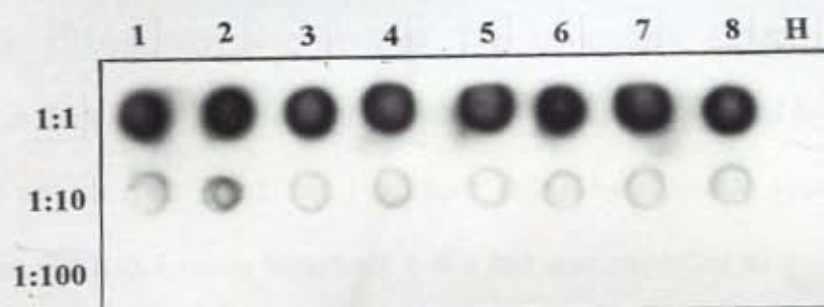


(b)

Figure 4.4.3 Detection of CLCuV in experimental hosts and viruliferous whiteflies by PCR amplification using coat protein (a) and DNA- β (b) primers
 1. *N. tabacum*, 2. Cotton, 3. Frenchbean, 4. Hollyhock, 5. Bhendi, 6. *P. floridana*,
 7. *N. benthamiana*, 8. Tomato, 9. Viruliferous whiteflies,
 H - Healthy Cotton, NW - Non-viruliferous whiteflies, M – 1kb DNA ladder



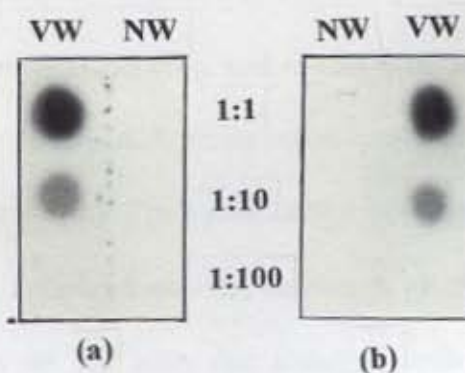
(a)



(b)

Figure 4.4.1 Detection of CLCuV in experimental hosts using probes to CLCuV DNA-A (a) and DNA-β (b)

1. Cotton, 2. Frenchbean, 3. Hollyhock, 4. Bhendi, 5. *P. floridana*, 6. *N. benthamiana*, 7. *N. tabacum*, 8. Tomato, H. Healthy Cotton



(a)

(b)

Figure 4.4.2 Detection of CLCuV in viruliferous whiteflies by probes to CLCuV DNA-A (a) and DNA-β (b)

VW – Viruliferous whiteflies, NW – Non-viruliferous whiteflies

Like the pCLCA probe, hybridization was also obtained with pCLC β probe (Fig.4.4.4b). But, the only difference was that weak hybridization was obtained with chilli leaf curl, cowpea golden mosaic and mungbean yellow mosaic.

b) Detection of infection in weeds

In order to find out the weeds, which may be alternate hosts of CLCuV under field conditions, dicot weeds harbouring adult population of whiteflies growing within and around leaf curl affected cotton fields were collected from Sri Ganganagar (Rajasthan) and Sirsa (Haryana). Twelve dicot weeds from each location were collected at random. None of these weeds showed any symptom of viral infection except for *Corchorus acutangularis*, which showed severe leaf curl-like symptoms (Fig.4.4.5), no other weed showed any symptom of viral infection. In addition, *Ageratum conyzoides* showing typical leaf crumple and yellow vein symptoms on the experimental farms in Delhi were tested.

Radiolabelled pCLCA (CLCuV DNA-A) probe detected virus infection in *A. conyzoides*, *Corchorus acutangularis*, *Melilotus indica* and *Parthenium hysterophorus* (Figs.4.4.4a & 4.4.6a). Hybridization was stronger with DNA from *A. conyzoides* (Fig.4.4.4a), *C. acutangularis* and *P. hysterophorus* than with DNA from *M. indica* (Fig.4.4.6a). Infection in *A. conyzoides* (Fig.4.4.4b), *C. acutangularis* and *M. indica* (Fig. 4.4.6b) was also detected with pCLC β probe whereas no hybridization was obtained with *P. hysterophorus*. Hybridization was stronger with DNA from *A. conyzoides* and *C. acutangularis* than *M. indica*.

No hybridization was detected with any of the following other weeds collected from Sri Ganganagar (Rajasthan): *Abutilon indica*, *Amaranthus viridis*, *Convolvulus arvensis*, *Digera arvensis*, *Euphorbia hirta*, *Heliotropium indicum*, *Phyllanthus niruri*, *Trianthema monogyna* and *Tribulus terrestris*, and from Sirsa (Haryana): *Achyranthes aspera*, *Amaranthus spinosus*, *A. viridis*, *Chenopodium album* C



Figure 4.4.5 Severe leaf curl-like symptoms on *Corchorus acutangularis*

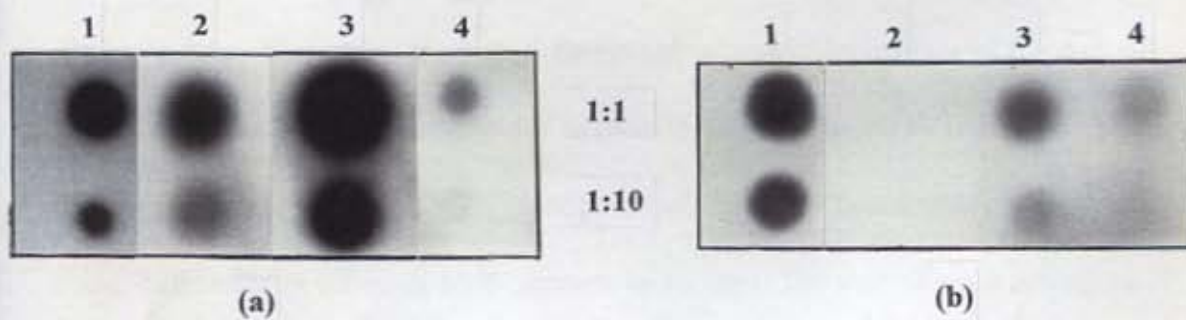


Figure 4.4.6 Detection of WTGs associated with weeds by NASH using CLCuV DNA-A (a) and DNA-β (b) probes
1. Infected cotton (control), 2. *Parthenium hysterophorus*, 3. *Corchorus acutangularis*, 4. *Melilotus indica*

amaranticolor, *Eclipta alba*, *Malva parviflora*, *Sisymbrium irio*, *Solanum nigrum*, *Trianthema monogyna*, *Tribulus terrestris* and *Vernonia cinerea*.

DISCUSSION

The probe to CLCuV CP gene could detect infection in plants by CLCuV and a range of other WTGs. This is expected as CP (or AV1) gene of WTGs is known to be highly conserved (Harrison and Robinson, 1999). Novel DNA- β components have been shown to be associated with CLCuV (Briddon *et al.*, 2001), Ageratum yellow vein virus (AYVV; Saunders *et al.*, 2000) and bhendi yellow vein mosaic virus (BYVMV; Jose and Usha, 2002). In NASH much difference was not observed in the amount of CLCuV DNA-A in different experimental hosts, but cotton and Frenchbean seem to favour multiplication of DNA- β more than the other hosts. In PCR amplification also no difference was observed in the amplicons from CLCuV CP gene primers, but the DNA- β primers not only amplified the expected ~1.4 kb fragments differentially but also detected the presence of smaller fragments in different plant hosts and viruliferous whiteflies. In hollyhock, ~0.6kb amplicon was more predominant than the full length DNA- β (~1.4 kb) amplicon. In bhendi and *P. floridana*, besides a predominant amplicon of ~1.4 kb, several other smaller fragments were also seen. They may represent different sub-genomic derivatives of DNA- β . The new DNA molecules may develop due to recombination between DNA-A and DNA- β as has been shown in AYVV (Saunders *et al.* 2000). Whether such recombinants occur differently in different hosts remains to be seen. The role of such molecules in pathogenicity also needs elucidation. Presence of both DNA-A and DNA- β of CLCuV in all the experimental hosts shows that both the components are efficiently encapsidated and transmitted by whiteflies. This was also confirmed by the presence of subgenomic DNA- β forms in whiteflies. The role of sub-genomic DNA- β forms associated with the cotton leaf curl disease remains to be understood. On the basis of

variations in sub-genomic DNA- β forms, different hosts of CLCuV, seem to fall in five groups; (i) Frenchbean, (ii) *N. benthamiana*, (iii) hollyhock, (iv) *N. tabacum*, cotton and tomato, and (v) bhendi and *P. floridana*. These differences indicate involvement of some host factors inducing production of the sub-genomic forms.

All the DNA- β characterized so far have a conserved domain of ~100 nucleotides upstream of the nonanucleotide sequence (Bridson *et al.*, 2001). The overall sequence of CLCuV-GI DNA- β (GenBank Acc. No. AY083590) shares 54.4% identity with DNA- β of BYVMV, the only other WTG from India found to have DNA- β (Jose and Usha, 2002). The present investigations have demonstrated association of DNA- β with several WTGs causing diseases in a variety of crop, ornamental and weed plants indicating association of DNA- β component a common feature of the WTGs occurring in northern India. The epidemiological as well as etiological significance of DNA- β , particularly those associated with WTGs like Indian cassava mosaic virus, which are known to have bipartite genome is puzzling.

Detection of infection in *C. acutangularis* and *M. indica* by both pCLCA and pCLC β probes suggests that these weeds could be the alternate hosts of CLCuV as these were collected from cotton fields severely affected by CLCuD. Whereas, *P. hysterophorus*, must be harbouring some other WTG as it did not hybridize with pCLC β probe. The samples from plants of *A. conyzoides*, affected with leaf crumple and yellow vein diseases, also hybridized with both pCLCA and pCLC β probes. *A. conyzoides* could also be a putative host of CLCuV.

Earlier workers (Harrison *et al.*, 1997; Mansoor *et al.*, 2000) have suggested that weeds may be important reservoir of CLCuV based on the detection of infection by WTG by serology and PCR. In order to find out whether plant species other than cotton are naturally infected with CLCuV-PK, Harrison *et al.*, 1997 carried out

serological studies based on epitope profiles of CLCuV-PK on a range of plants with virus like symptoms. They found that okra (*Abelmoschus esculentus*) with yellow vein symptoms, okra with leaf curl symptoms, *Hibiscus tiliaceus* with leaf curl symptoms, radish (*Raphanus sativus*) with leaf curl symptoms, sunflower (*Helianthus annuus*) with vein thickening symptoms, viruses found in China rose (*Hibiscus rosa-sinensis*) and hollyhock (*Althaea rosea*) had epitope profiles indistinguishable from CLCuV-PK, suggesting that these plants are putative natural hosts of CLCuV-PK. Mansoor *et al.*, (2000) confirmed by PCR and dot-blot hybridization that radish plants are infected with CLCuV naturally.

Further studies are required to transmit WTGs from weed hosts to cotton and other crop plants for identifying the main sources of infection of economically important WTGs.

4.5 MOLECULAR VARIABILITY AMONG COTTON LEAF CURL VIRUS ISOLATES IN NORTH INDIA

Leaf curl disease has emerged as a serious threat to the cultivation of cotton in the north-western cotton belt of Indian sub-continent (Varma and Malathi, 2003). The disease is caused by cotton leaf curl virus (CLCuV) a whitefly transmitted geminivirus (WTG) belonging to the genus *Begomovirus* (Varma *et al.*, 1993). The genome organization of CLCuV DNA-A resembles that of monopartite WTGs, but it requires another molecule designated DNA- β for disease induction (Briddon *et al.*, 2001). DNA-A sequences of a large number of CLCuV isolates from Pakistan and one isolate from India are available in the GenBank database. Sequences of DNA- β molecules associated with two CLCuV isolates from Pakistan (Briddon *et al.*, 2001) and one from India (Radhakrishnan *et al.*, 2001) are also available. Considerable molecular variability has been observed in DNA-A of CLCuV and various isolates, sharing >90% sequence identity, can be placed into five distinct species.

The leaf curl disease of cotton (CLCuD) is fast spreading in India, but information on the variability of the causal virus is not available. In the present investigation molecular variability in DNA-A and DNA- β of four CLCuV isolates from Delhi, Haryana and Punjab is studied and their sequence identity is compared with the known CLCuV sequences from Indian sub-continent.

MATERIALS AND METHODS

Virus isolates

Young fully expanded leaves of cotton, expressing severe symptoms, were collected from four different locations in north-western India, and CLCuV isolated by

transmission through whitefly (*Bemisia tabaci*) to cotton (*Gossypium hirsutum*) cv. LH900. All the isolates were maintained in the same host in an insect proof glass-house. One isolate each from Delhi (CLCuV-DI), Hisar, Haryana (CLCuV-HI), Ludhiana, Punjab (CLCuV-LI) and Sirsa, Haryana (CLCuV-SI) were selected for further studies.

PCR cloning

a) Primer designing

Based on the published nucleotide sequences of DNA-A of CLCuV (GenBank Acc. No. –AJ002447, AJ002448, AJ002449, AJ002452, AJ002455 and AJ132430) two pairs of primers were designed to amplify complete DNA-A as two fragments. One set of primer was designed to amplify the region downstream of nonanucleotide sequence (CRV-AV primer) to C' terminal of AV1 gene (CPR primer). The second set of primers were designed to amplify the rest of the DNA-A fragment from upstream of nonanucleotide sequence (CRC-AC primer) to C' terminal of AC3 gene (FV-AC primer). The two sets of primers used for amplifying DNA-A were:

1. CRV-AV : 5'-TTACCGGATGGCCGCGGATTT-3' (22 nucleotides)
2. CPR : 5'-TAATATCAATTCGTTACAGAG-3' (21 nucleotides)
3. CRC-AC : 5'-TATTAAACGGATGGCCGCTTT-3' (21 nucleotides)
4. FV-AC : 5'-ATAAAGTTTGAATTTTAT TTC-3' (21 nucleotides)

Similarly, the following set of abutting primers having a *KpnI* site at its 5' end were also designed to amplify the second genomic component of CLCuV i.e., CLCuV DNA-β, based on the published nucleotide sequence of CLCuV DNA-β of CLCuV-Pk (GenBank Acc. No.-AJ298903; Briddon *et al.*, 2001). The two primers used were:

1. Beta 1: 5'-GGTACCACTACGCTACGCAGCAGCC-3' (25 nucleotides)
2. Beta 2: 5'-GGTACCTACCCTCCCAGGGGTACAC-3' (25 nucleotides)

These primers were synthesized by the GENSET Singapore Biotech. Pvt. Ltd.

b) PCR amplification

For PCR amplification of CLCuV DNA-A and DNA- β components from infected cotton plants, total DNA was isolated from 100mg of leaf tissue using DNeasy Plant Mini Kit (Qiagen Inc., USA) according to manufacturer's instructions. 200ng of total DNA was taken for PCR amplification.

The PCR reaction mixture consisted of 2.5 U of *Taq* DNA polymerase (MBI Fermentas, USA), 5 μ l of 10x PCR buffer (without MgCl₂), 3 μ l of 25 mM MgCl₂, 1 μ l of 10 mM dNTP and 1 μ l of 100ng/ μ l primers (DNA-A primers) or (DNA- β primers). Total reaction volume was made up to 50 μ l using sterile distilled water.

PCR amplification protocol was standardized employing different temperatures of denaturation, annealing and extension based on T_m values of the primers. The successful PCR amplification of the CLCuV DNA-A fragments and CLCuV DNA- β was obtained with the following programmes in a PHC-3 (Techne) thermal cycler. First denaturation at 94°C for 2 min, then 30 cycles of denaturation (94°C for 1 min), annealing (55°C for 2 min) and extension (72°C for 3 min). At the end of the cycles, a final extension step at 72°C for 5 min was given and then the temperature was decreased to 4°C until the reaction mixtures were removed. PCR products were analyzed by electrophoresis in 1% agarose gel, stained with ethidium bromide (0.5 μ g/ml) and photographed on a UV transilluminator.

c) Cloning

Individual amplicons from the cotton sample were eluted from the gel using Gel Extraction Kit (Qiagen Inc., USA) according to manufacturer's instructions. Cloning of CLCuV DNA components was done according to Sambrook *et al.*, (1989). The gel eluted

amplified DNA product was cloned into pGEM-T easy vector (Promega, USA). *Escherichia coli* strain DH5 α was transformed with the recombinant plasmid by CaCl₂ method (Mandel and Higa, 1970).

Transformants containing recombinant plasmids were selected on Luria agar plates containing ampicillin (50 μ g/ml), X-gal (2%) and IPTG (100 mM) (Vieira and Messing, 1982). Presence of the insert was confirmed by releasing the insert from the plasmid by using *NotI* restriction enzyme in the case of DNA-A clones and *KpnI* restriction enzyme in the case of DNA- β clones. The restricted products were analyzed by electrophoresis in 1% agarose gel as described earlier and were compared with the marker 1kb DNA ladder (MBI Fermentas, USA).

Sequence determination and analysis

The clones were sequenced using automated DNA sequencer (ABI Prism, Perkin Elmer, University of Delhi, South Campus, New Delhi). The sequences were analysed using the software programme Bioedit version 5.0.9. Multiple nucleotide and putative amino acid sequence alignments were done in CLUSTALW(1.7) programme (<http://www.ebi.ac.uk/clustalw>). Comparisons were made with sequences obtained from GenBank nucleotide sequence database. Relationship dendrograms were constructed from the aligned sequences using the neighbour-joining method and bootstrap option of Treecon version 1.3b (500 x bootstrap replicates). Accession numbers of the geminivirus sequences reported and compared in the present investigation, geographical origin and other details are given in table 4.5.1. In DNA-A, N' terminal AC1, complete intergenic region, complete AV2 and AV1 gene were sequenced. In DNA- β , ~ 450 nucleotide upstream and ~ 450 nucleotide downstream of the nonancucleotide sequence were sequenced.

Table 4.5.1. Geminivirus sequences used in this study

Sl. no.	Acronym	Virus name	Geographic origin	GenBank Acc. No.*
1.	AYVV-Sg	Ageratum yellow vein virus	Singapore	AJ252072 (DNA- β)
2.	BYVMV-In	Bhendi yellow vein mosaic virus	Madurai, India	AF241479 AJ308425 (DNA- β)
3.	CLCuV-311Pk	Cotton leaf curl virus	Multan, Pakistan	AJ002459
4.	CLCuV-72bPk	Cotton leaf curl virus	Faisalabad, Pakistan	AJ002448
5.	CLCuV-802a Pk	Cotton leaf curl virus	Lahore, Pakistan	AJ002455
6.	CLCuV-806b Pk	Cotton leaf curl virus	Faisalabad, Pakistan	AJ002449
7.	CLCuV-DI	Cotton leaf curl virus	Delhi, India	AY146962 (AC1) AF465618 (AV1) AY146960 (AV2) AY146956 (IR) AF534185 (DNA- β)
8.	CLCuV-GI	Cotton leaf curl virus	Sri Gangana-gar, India	AF363011 AY083590 (DNA- β)
9.	CLCuV-HI	Cotton leaf curl virus	Hisar, India	AY146964 (AC1) AF465620 (AV1) AY146959 (AV2) AY146955 (IR) AF534186 (DNA- β)
10.	CLCuV-LI	Cotton leaf curl virus	Ludhiana, India	AY146963 (AC1) AF465619 (AV1) AY146958 (AV2) AY146954 (IR) AF534187 (DNA- β)
11.	CLCuV-Pk	Cotton leaf curl virus	Pakistan	AJ132430
12.	CLCuV-Pk(1)	Cotton leaf curl virus	Pakistan	AJ292769 (DNA- β)
13.	CLCuV-Pk(2)	Cotton leaf curl virus	Pakistan	AJ298903 (DNA- β)

14.	CLCuV-SI	Cotton leaf curl virus	Sirsa, India	AY146961 (AC1) AF465621 (AV1) AY146957 (AV2) AY146953 (IR) AF534188 (DNA-β)
15.	ICMV	Indian cassava mosaic virus	India	NC001932
16.	MYMIV-Bg	Mungbean yellow mosaic virus	New Delhi, India	AF126406
17.	OLCV-Sd	Okra leaf curl virus	Sudan	AF397217 (DNA-β)
18.	OYVMV-201 Pk	Okra yellow vein mosaic virus	Faisalabad, Pakistan	AJ002451
19.	OYVMV-301 Pk	Okra yellow vein mosaic virus	Multan, Pakistan	AJ002453
20.	PaLCV-In	Papaya leaf curl virus	Lucknow, India	Y15934
21.	ToLCV-Ban 4	Tomato leaf curl virus	Bangalore, India	AF165098
22.	ToLCV-Ban II	Tomato leaf curl virus	Bangalore, India	U38239
23.	ToLCV-NDe	Tomato leaf curl virus	New Delhi, India	U15016

* <http://www.ncbi.nlm.nih.gov>

RESULTS

Complete DNA-A of the four CLCuV isolates from north-western India were PCR amplified as two fragments, of ~1.7 kb (using FV-AC + CRC-AC primers) and ~1.0 kb (using CPR + CRV-AV primers) (Fig.4.5.1a). Similarly, full-length DNA- β of the four isolates were also PCR amplified, as ~ 1.4kb fragment (using Beta1 + Beta2 primers) (Fig. 4.5.2a). The PCR products were then cloned into pGEM-T easy vector (Fig.4.5.1b; Fig.4.5.2b) and sequenced (Appendix 3). These sequences are available in the GenBank, DDJB and EMBL nucleotide sequence databases.

SEQUENCE ANALYSIS

Intergenic region sequences

To consider the intergenic region (IR) of CLCuV isolates, sequences in between the start codon of ORF AC1 to start codon of ORF AV2 were aligned and percentage identity determined. The non-coding intergenic region is highly specific for a WTG species, and give a good indication of variability amongst different isolates.

Comparison of nucleotide sequences of intergenic region clearly showed that CLCuV isolates from India, including the one from Sri Ganganagar (CLCuV-GI) reported earlier (Radhakrishnan *et al.*, 2001), show up to 45% divergence (Table 4.5.2). CLCuV-HI and CLCuV-LI are the most closely related isolates as they have 98.4% identity, followed by CLCuV-GI and CLCuV-SI which share 90.5% identity. The remaining combinations of the Indian isolates share <84% identity, and the Delhi isolate is most distinctive (Table 4.5.2). The Indian isolates share 53.5% to 96.1% sequence identity with the Pakistan isolates. CLCuV-HI, CLCuV-LI and CLCuV-GI seem to be more closely related to Pakistan isolates than to the other Indian isolates, and CLCuV-DI shows greater diversity also with the Pakistan isolates (Table 4.5.2). The sequence

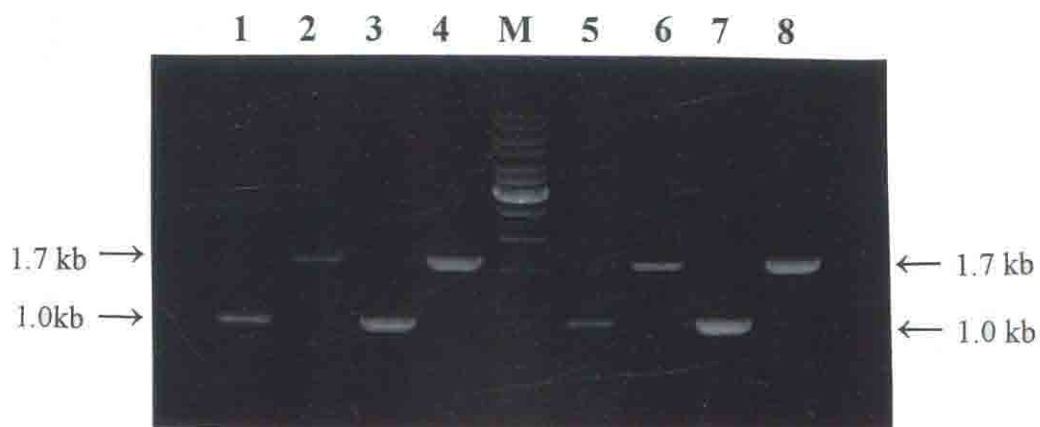


Figure 4.5.1a. PCR amplification of two fragments (1.7kb and 1.0kb) of CLCuV DNA-A of different isolates.

Lane 1&2: Delhi isolate, 3&4: Hisar isolate, 5&6: Ludhiana isolate, 7&8: Sirsa isolate
M: 1kb DNA ladder



Figure 4.5.1b. Agarose gel electrophoresis of CLCuV DNA-A clones of North Indian isolates restricted with *NotI* restriction enzyme.

Lane 1&5: Delhi isolate, 2&6: Hisar isolate, 3&7: Ludhiana isolate, 4&8: Sirsa isolate
M: 1kb DNA ladder



Figure 4.5.2a. PCR amplification of CLCuV DNA- β .

Lane 1: Delhi isolate, 2: Hisar isolate, 3: Ludhiana isolate, 4: Sirsa isolate

M: 1kb DNA ladder

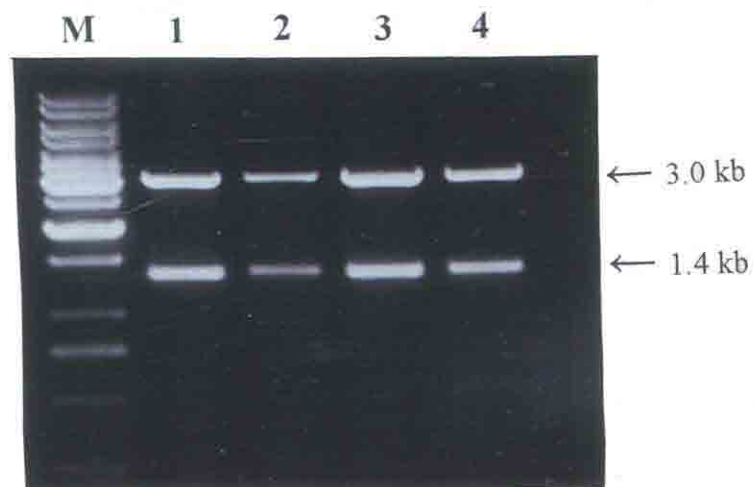


Figure 4.5.2b. Agarose gel electrophoresis of CLCuV DNA- β clones restricted with *Kpn*I.

Lane 1: Delhi isolate, 2: Hisar isolate, 3: Ludhiana isolate, 4: Sirsa isolate

M: 1kb DNA ladder

Table 4.5.2. Comparison of nucleotide identity in the intergenic region among some of the begomoviruses from the Indian sub-continent

Virus	CLCuV-DI	CLCuV-GI	CLCuV-HI	CLCuV-LI	CLCuV-SI
CLCuV-DI	100.0				
CLCuV-GI	61.2	100.0			
CLCuV-HI	83.4	58.4	100.0		
CLCuV-LI	83.4	58.4	98.4	100.0	
CLCuV-SI	55.0	90.5	64.3	64.3	100.0
CLCuV-Pk	80.0	56.0	94.2	94.2	60.9
CLCuV-72b Pk	55.6	91.9	53.5	53.5	83.9
CLCuV-806b Pk	55.6	91.9	53.5	53.5	83.9
CLCuV-802a Pk	59.2	55.6	67.8	66.7	55.0
CLCuV-311 Pk	81.7	57.3	96.1	96.1	63.0
OYVMV-201 Pk	64.6	53.6	72.6	73.0	56.3
OYVMV-301 Pk	47.8	71.6	53.5	53.5	74.1
BYVMV-In	49.5	77.7	55.5	55.8	81.2
ICMV	33.9	34.8	31.6	31.9	32.2
PaLCV-In	48.0	46.4	47.4	46.4	43.9
ToLCV-NDe	53.3	46.54	60.2	60.2	48.3
ToLCV-BanII	51.1	51.4	50.4	51.1	51.4
ToLCV-Ban4	50.7	47.3	49.3	50.0	49.3
MYMIV-Bg	32.1	32.9	33.7	34.4	32.3

identity of IR of the Indian isolates of CLCuV to other begomoviruses is less than 82% and in a majority of comparisons it is less than 50% indicating variability in the IR region of begomoviruses. Phylogenetic tree based on nucleotide sequences of IR (Fig.4.5.3), also shows similar relationships. The CLCuV-GI and CLCuV-SI cluster with CLCuV-72bPk and CLCuV-806bPk, and CLCuV-DI, CLCuV-LI, CLCuV-HI, CLCuV-311Pk and CLCuV-Pk cluster together. These groupings indicate two distinct lineages.

Multiple alignment of the IR nucleotide sequences was done to locate the variable and identical regions (Fig.4.5.4). A core central region (of 46 nucleotides ie., 23nt upstream and 23nt downstream of *Ori* of replication) adjacent to the nonanucleotide TAATATTAC, is highly conserved in all the Indian isolates except for one nucleotide change in the 11th nucleotide upstream from the *Ori* of replication in CLCuV-DI. This block of sequence also shows high degree of conservation in all the Begomoviruses affecting malvaceous plants. However, at 3' upstream region of *Ori* of replication, between ~ 24nt coordinate to ~132nt coordinate and at 5' downstream region of *Ori* of replication, between ~ 24nt coordinate to ~ 89nt coordinate, there is considerable sequence variability.

All the IR contain repeat sequences or the Rep binding iteron sequences of the kind described by Zhou *et al.*(1998) and Sanz *et al.* (2000). The Rep binding iteron sequence for the Indian CLCuV isolates falls into two groups, either TGATCA or GGTGTA (Table 4.5.3). These iterons are present as three copies in tandem orientation in the IR. The first copy of these sequences occur near the 5' end of the IR, the second copy occur just to the 5' side or immediate upstream of the TATA motif, which is located at ~ 87nt coordinate upstream of *Ori* of replication. The third copy of the iteron occurs on the 3' side or immediate downstream of the TATA motif. The arrangement of these iterons

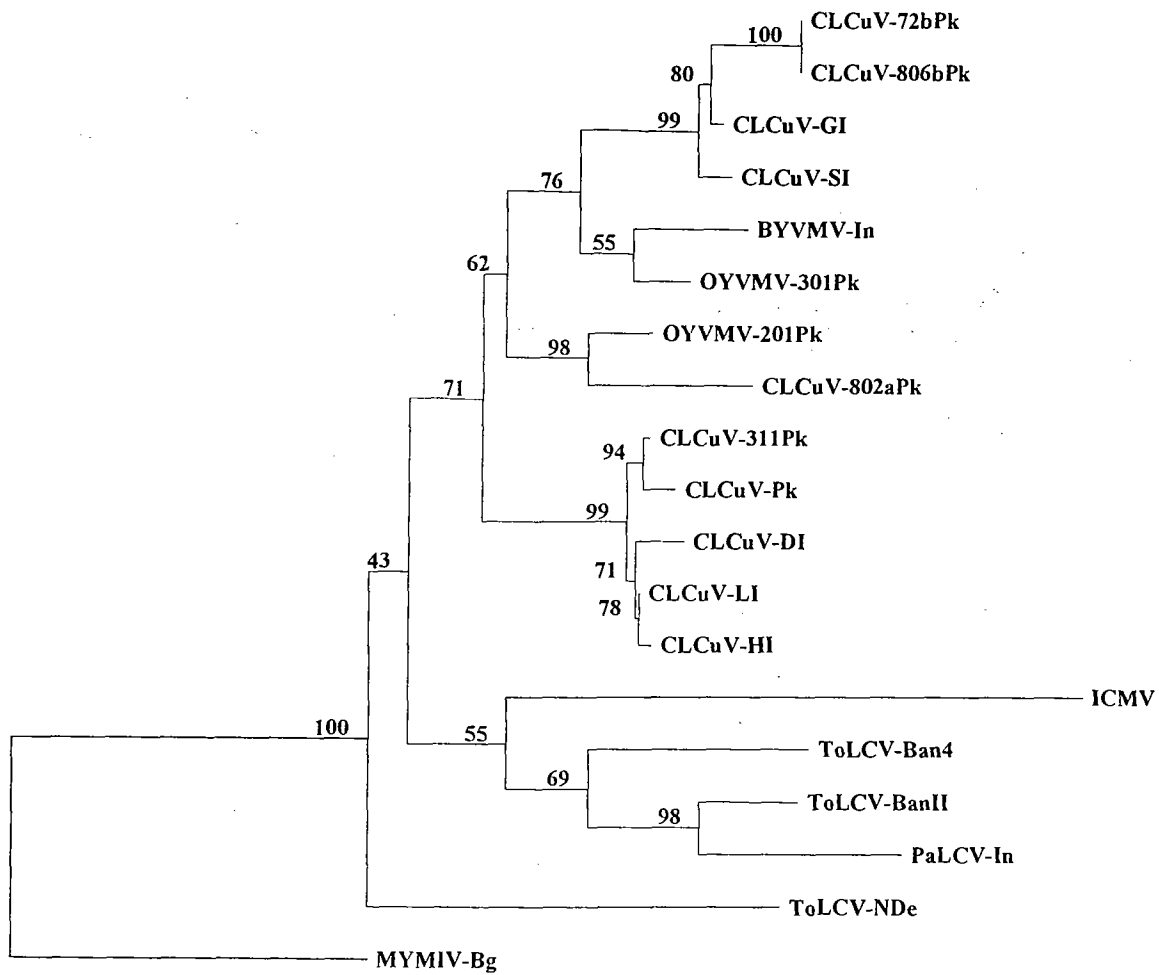


Figure 4.5.3. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the intergenic region. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence. •

```

CLCuV-DI 1 ---TTTGGCAATCGGTGT-ACA-C-TCTAATTCTATGGCAATC-----GGGTAA-CGGGGTCCAAATATATAGGTGTA 66
CLCuV-GI 1 ---CTTGACTAAAATTGATCACC-GATTGACCGCTCTTGCAACTCTCCCCGGTATAT-CGGTGATCAATATATAGTGATC 75
CLCuV-HI 1 ---TTTGGCAATCGGTGT-ACA-C-TCTAATTCTATGGCAATC-----GGGTAA-CGGGGTCCAAATATATAGGTGTA 66
CLCuV-LI 1 ---TTTGGCAATCGGTGT-ACA-C-TCTAATTCTATGGCAATC-----GGGTAA-CGGGGTCCAAATATATAGGTGTA 66
CLCuV-SI 1 ---CTTGACTAAAATTGATCACC-GATTGACCGCTCTTGCAACTCTCCCCGGTATAT-CGGTGATCAATATATAGTGATC 75
CLCuV-Pk 1 ---TTTGGCAATCGGTGT-ACA-C-TCTAATTCTATGGCAATC-----GGGTAA-CGGGGTCCAAATATATAGGTGTA 66
CLCuV-806bPk 1 ---CTTGACTAAAATTGATCACC-GATTGACCGCTCTTGCAACTCTCCCCGGTATAT-CGGTGATCAATATATAGTGATC 75
CLCuV-72bPk 1 ---CTTGACTAAAATTGATCACC-GATTGACCGCTCTTGCAACTCTCCCCGGTATAT-CGGTGATCAATATATAGTGATC 75
CLCuV-802aPk 1 ---TTTGTCTATCGGGGAC-AC---TTAAACTCATATCAATT-----GGSGACACTGGGGACTCATTTATATCTGCGT 66
CLCuV-311Pk 1 ---TTTGGCAATCGGTGT-ACA-C-TCTAATTCTATGGCAATC-----GGGTAA-TGGGGTCCAAATATATAGGTGTA 66
OYVMV-301Pk 1 ---CTTGACCAAGTGAGAGGACCCGATTGACCGCTCTTGCAACTCTCCCCGTATAT-TGGGTCTCAATATATAGTGAGA 76
OYVMV-201Pk 1 TGACTTGGTCAATCGGTAC-CCA-GATCTAATCTATGTCAATT-----GGGTAA-CGGTCCCTATATATAGTGGGT 70
BYVMV-In 1 ---CTTGACTATTTTGGAG-ACCCGATTGACCGCTCTTACAACCTCTCCCCAGTATAT-CGGTCCCTATATATAGTGAGA 75

CLCuV-DI 67 CCCCAAATGGCAT-----TATC-GTAATTGAGAAAT--CATTTCAAAATCCTCAGCTCCAA--AAAGCGGCCAT 132
CLCuV-GI 76 AC-CAAATGGCA-----TAATGGTAATA-AAAAACTTTAATTGAAATTCAAACCAAAGGCTAAAGCGGCCAT 143
CLCuV-HI 67 CCCCAAATGGCAT-----TATC-GTAATTGAGAAAT--CATTTCAAAATCCTCAGCTCCAA--AAAGCGGCCAT 132
CLCuV-LI 67 CCCCAAATGGCAT-----TATC-GTAATTGAGAAAT--CATTTCAAAATCCTCAGCTCCAA--AAAGCGGCCAT 132
CLCuV-SI 76 AC-CAAATGGCA-----TAATGGTAATA-AAAAACTTTAATTGAAATTCAAACCAAAGGCTAAAGCGGCCAT 143
CLCuV-Pk 67 CCCCAAATGGA-----TATC-GTAATGGGAGAAAT--CATTTGAGATCCTCAGCTCCAA--AA-CCGGCCAT 129
CLCuV-806bPk 76 AC-CAAATGGCA-----TAATGGTAATA-AAAAACTTTAATTGAAATTCAAACCAAAGGCTAAAGCGGCCAT 143
CLCuV-72bPk 76 AC-CAAATGGCA-----TAATGGTAATA-AAAAACTTTAATTGAAATTCAAACCAAAGGCTAAAGCGGCCAT 143
CLCuV-802aPk 67 CCTTAAATGGCATAAATGTAATTTTGCCTTTATGTTG---AATTTCAAAATTCAAACGCTCCAA--AAAGCGGCCAT 141
CLCuV-311Pk 67 CCCCAAATGGCAT-----TATC-GCAATTTGAGAAAT--CATTTCAAAATCCTCAGCTCCAA--AAAGCGGCCAT 132
OYVMV-301Pk 77 CC-CAAATGGCA-----TATCGTAATTGGACAA--TAGATTCAAAATCCTCAGCTCCAA--AAAGCGGCCAT 141
OYVMV-201Pk 71 AC-TGAATGGCAT-----TATTTGTAATTACAAAGG--AAATTCAAAACTACAGCTCCAA--AAAGCGGCCAT 136
BYVMV-In 76 CC-CAAATGGCA-----TAATTGTAATA-AAACAACCTTTAATTGAAATTCAAACGAAAGGCTAAAGCGGCCAT 143

CLCuV-DI 133 CCGTTTAATATTACCGGATGGCGCGGATTTTTTACC GGATGGCTGCGCGATTTT TGTGGGCTTACCATTAACTT 212
CLCuV-GI 144 CCGTTTAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 202
CLCuV-HI 133 CCGTTTAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 178
CLCuV-LI 133 CCGTTTAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 178
CLCuV-SI 144 CCGTTTAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 189
CLCuV-Pk 130 CCGTATAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 175
CLCuV-806bPk 144 CCGTTTAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 201
CLCuV-72bPk 144 CCGTTTAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 201
CLCuV-802aPk 142 CCGTATAATATTACCGGATGGCGCGGATTTTTTTAG-----TGTGGGCTTACCATTAACTT 191
CLCuV-311Pk 133 CCGTATAATATTACCGGATGGCGCGGATTTTTT-----TGTGGGCTTACCATTAACTT 178
OYVMV-301Pk 142 CCGACTAATATTACCGGATGGCGCGGATTTTTTTAAG-----TGTGGGCTTACCATTAACTT 192
OYVMV-201Pk 137 CCGTATAATATTACCGGATGGCGCGGATTTTTTTAG-----TGTGGGCTTACCATTAACTT 186
BYVMV-In 144 CCGTATAATATTACCGGATGGCGCGGATTTTTT-AAG-----TGTGGGCTTACCATTAACTT 193

CLCuV-DI 213 GTCGGCCAATCATATGACTCCCTCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATT 291
CLCuV-GI 203 GTCGGCCAATCATATGACTCCCTCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATT 281
CLCuV-HI 178 --CGATTTATGAGATTGCTCCCTCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATT 255
CLCuV-LI 178 --CGATTTATGAGATTGCTCCCTCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATT 255
CLCuV-SI 189 --CGATTTATGAGATTGCTCCCTCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATT 266
CLCuV-Pk 175 --CGATTTATGAGATTGCTCCCTCAAAGCT-AAATAACGCTCCCGCACACTATAAGTACTTGCGCACTAAGTTTCAAATT 252
CLCuV-806bPk 202 GTCGGCCAATCATATGACGCGCTCAAAGCTTAATAAATCTCCCGCTTATTATAAGTACTTCTGTGCTAAGTATCGGTTT 281
CLCuV-72bPk 202 GTCGGCCAATCATATGACGCGCTCAAAGCTTAATAAATCTCCCGCTTATTATAAGTACTTCTGTGCTAAGTATCGGTTT 281
CLCuV-802aPk 191 --GAACGACGACGATGCAGACTCAAAGCTTAGATAACGCTCCTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATT 268
CLCuV-311Pk 178 --CGATTTATGAGATTGCTCCCTCAAAGCT-AAATAACGCTCCCGCCACTATAAGTACTTGCGCACTAAGTTTCAAATT 255
OYVMV-301Pk 192 ---AACGACGACGATGCAGACTCAAAGCTTAGATAACGCTCCTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATT 268
OYVMV-201Pk 186 --GAACGACGACGATGCAGACTCAAAGCTTAGATAACGCTTCTTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATT 263
BYVMV-In 193 ---AACGACGACGATGCAGACTCAAAGCTTAGATAACGCTCCTTCG-GCTATAAGTACTTGCGCACTAAGTTTCAAATT 269

CLCuV-DI 292 CAAAC 296
CLCuV-GI 282 CAAAC 286
CLCuV-HI 256 CAAAC 260
CLCuV-LI 256 CAAAC 260
CLCuV-SI 267 CAAAC 271
CLCuV-Pk 253 CAAAC 257
CLCuV-806bPk 282 GAAAA 286
CLCuV-72bPk 282 GAAAA 286
CLCuV-802aPk 269 GAAAC 273
CLCuV-311Pk 256 CAAAC 260
OYVMV-301Pk 269 GAAAC 273
OYVMV-201Pk 264 AAAAC 268
BYVMV-In 270 AAAAA 274

```

Figure 4.5.4. Multiple alignment of the nucleotide sequence of the intergenic region of DNA-A of 13 malvaceous Begomovirus isolates from India and Pakistan. Blocks of conserved regions are shown in reverse contrast.

Table 4.5.3. Iteron sequences and their arrangement in the intergenic region of CLCuV and related begomoviruses

S.No	Iteron	Copy No.	Virus	Iteron arrangement
1	TGATCA	3	CLCuV-GI, CLCuV-SI, CLCuV-72bPK, CLCuV-806bPk	→ →TATA→
2	GGTGTA	3	CLCuV-DI, CLCuV-HI, CLCuV-LI, CLCuV-Pk, CLCuV-311Pk	→ → TATA→
3	GGGGACACT	2	CLCuV-802aPk	→ → TATA
4	CGGT(A/G)CCC	2	OYVMV-201Pk	→ →TATA
5	GTGAGA	2	OYVMV-301Pk	→ TATA→
6	TGAGACCC	2	BYVMV-In	→ TATA→

* Note: Arrow (→) indicates iteron sequence

arrangement of these iterons and their sequences are also similar with the CLCuV isolates from Pakistan (Table 4.5.3). The CLCuV-GI and CLCuV-SI shares the same iteron sequence and arrangement with the CLCuV-72bPk and CLCuV-806bPk, whereas, CLCuV-DI, CLCuV-HI and CLCuV-LI are similar to CLCuV-Pk and CLCuV-311Pk, also indicating two distinct lineages identified by overall sequence comparison of IR.

AV1 gene

All the isolates of CLCuV and begomoviruses from other malvaceous hosts compared show >95% amino acid sequence identity in the CP gene (Table 4.5.4). A notable feature is the near identical (99.2%) sequence identity of CLCuV-LI with CLCuV-802aPk and OYVMV-301Pk. All the Indian isolates also show high degree of sequence identity (>90%) with ICMV, PaLCV-In, ToLCV-NDe and ToLCV-BanII, but ToLCV-Ban4 and MYMIV-Bg show greater diversity in CP gene (Table 4.5.4).

As determined by sequence analysis CLCuV-LI and CLCuV-SI were closer to CLCuV-802aPk and OYVMV-301Pk, whereas, the other three Indian isolates ie., CLCuV-GI, CLCuV-HI and CLCuV-DI isolates grouped together along with CLCuV-311Pk and CLCuV-Pk by phylogenetic analysis (Fig. 4.5.5).

ORF AV2

The predicted amino acid sequence of the AV2 gene shows very close affinity (>94%) between all the CLCuV isolates from India (Table 4.5.5); the sequences of CLCuV-LI and CLCuV-SI are identical and show only 5% divergence from the other Indian isolates; similarly CLCuV-DI and CLCuV-HI are identical and CLCuV-DI is near identical (99.1% identity) to CLCuV-GI. All the five Indian isolates show more than 90% identity in the ORF AV2 with CLCuV-802aPk and CLCuV-311Pk and OYVMV-201Pk, OYVMV-301Pk and BYVMV-In. Whereas there was considerable divergence (30%)

Table 4.5.4. Comparison of amino acid identity in the coat protein gene among begomoviruses from the India sub-continent

Virus	CLCuV-DI	CLCuV-GI	CLCuV-HI	CLCuV-LI	CLCuV-SI
CLCuV-DI	100.0				
CLCuV-GI	99.6	100.0			
CLCuV-HI	99.2	99.6	100.0		
CLCuV-LI	96.0	96.4	96.0	100.0	
CLCuV-SI	95.3	95.7	95.3	97.2	100.0
CLCuV-Pk	94.5	94.9	94.5	96.4	96.4
CLCuV-72b Pk	96.0	96.4	96.0	93.7	92.1
CLCuV-806b Pk	96.0	96.4	96.0	93.7	92.1
CLCuV-802a Pk	96.8	97.2	96.8	99.2	98.0
CLCuV-311 Pk	95.3	95.7	95.3	97.2	97.6
OYVMV-201 Pk	96.8	97.2	96.8	99.2	98.0
OYVMV-301 Pk	95.3	95.7	95.3	99.2	96.4
BYVMV-In	96.0	96.4	96.0	98.4	97.2
ICMV	90.2	90.6	90.2	91.7	90.6
PaLCV-In	94.5	94.9	94.5	93.3	92.5
ToLCV-NDe	91.4	91.7	91.4	91.7	90.2
ToLCV-BanII	94.1	94.5	94.1	94.9	93.7
ToLCV-Ban4	79.6	79.6	79.2	79.2	77.3
MYMIV-Bg	64.9	69.9	64.9	64.9	63.8

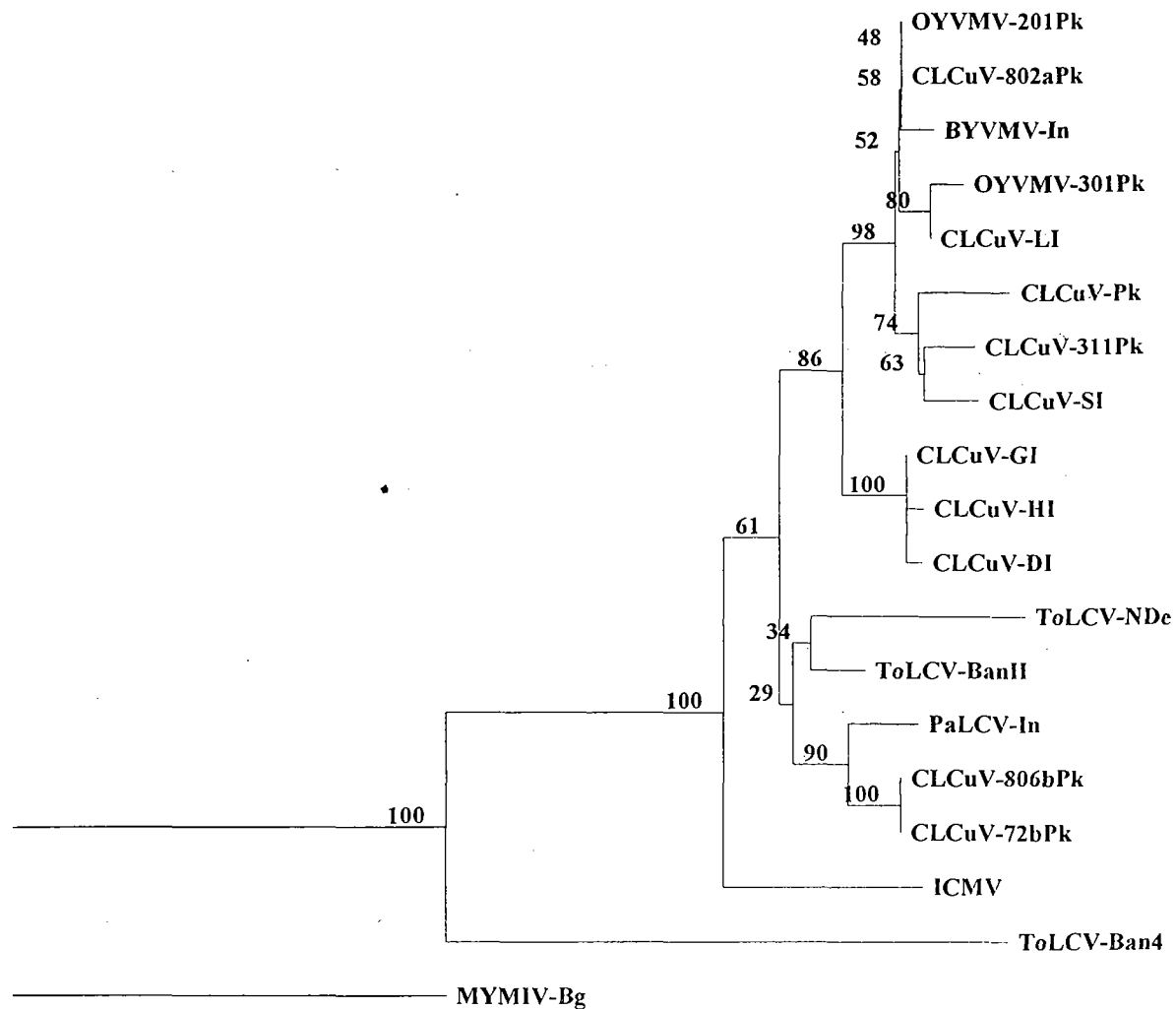


Figure 4.5.5. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the ORF AV1. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

Table 4.5.5. Comparison of amino acid identity in the AV2 gene among begomoviruses from India and Pakistan

Virus	CLCuV-DI	CLCuV-GI	CLCuV-HI	CLCuV-LI	CLCuV-SI
CLCuV-DI	100.0				
CLCuV-GI	99.1	100.0			
CLCuV-HI	100.0	99.1	100.0		
CLCuV-LI	95.0	94.2	95.0	100.0	
CLCuV-SI	95.0	94.2	95.0	100.0	100.0
CLCuV-Pk	88.9	88.1	88.9	87.6	87.6
CLCuV-72b Pk	76.2	75.4	76.2	71.9	71.9
CLCuV-806b Pk	76.2	75.4	76.2	71.9	71.9
CLCuV-802a Pk	93.3	92.5	93.3	96.6	96.6
CLCuV-311 Pk	95.0	94.2	95.0	100.0	100.0
OYVMV-201 Pk	93.3	92.5	93.3	96.6	96.6
OYVMV-301 Pk	91.7	90.9	91.7	95.0	95.0
BYVMV-In	93.3	92.5	93.3	96.6	96.6
ICMV	64.4	63.5	64.4	63.6	63.6
PaLCV-In	72.8	72.0	72.8	68.5	68.5
ToLCV-NDe	61.0	60.1	61.0	59.5	59.5
ToLCV-BanII	72.8	72.0	72.8	68.5	68.5
ToLCV-Ban4	65.2	66.1	65.2	62.8	62.8
MYMIV-Bg	29.6	28.8	29.6	28.9	28.9

when compared with the other isolates from Pakistan viz., CLCuV-72bPk, CLCuV-806bPk and CLCuV-Pk. It is interesting to see that CLCuV-LI and CLCuV-SI are identical to CLCuV-311Pk. Identities with other begomoviruses from India were negligible.

Phylogenetic analysis clearly showed the relatedness amongst CLCuV isolates based on amino acid sequence of AV2 gene (Fig.4.5.6). Three North Indian isolates i.e., CLCuV-DI, CLCuV-GI and CLCuV-HI clustered together, whereas, the other two isolates i.e., CLCuV-LI and CLCuV-SI are more closer to CLCuV-311Pk and branch separately.

Comparison of N' terminal AC1 sequence

Comparison of predicted amino acid sequence of the N' terminal region of the replication initiation protein (Rep) gave unexpected results. There was distinct separation between CLCuV isolates from India (Table 4.5.6; Fig.4.5.7). CLCuV-GI and CLCuV-SI isolates have 98% identity between them and show only ~73% to 74% identity with the other three isolates (CLCuV-DI, CLCuV-LI and CLCuV-HI). Their relationship with Pakistan isolates also reveals their separate identity. CLCuV-GI and CLCuV-SI isolates have higher affinity to Pakistan isolates CLCuV-72bPk, CLCuV-806bPk (>97%), 70-80% with CLCuV-Pk, CLCuV-311Pk, OYVMV-201Pk, OYVMV-301Pk and BYVMV-In, but considerably less (<55%) with CLCuV-802aPk. The isolates CLCuV-LI, CLCuV-DI, and CLCuV-HI showed contrasting relationship. Their identity with CLCuV-Pk and CLCuV-311Pk were 94-97%, less than 75% with CLCuV-72bPk and CLCuV-806bPk and only 50% with CLCuV-802aPk. In contrast, all the Indian isolates of CLCuV showed 75 to 85% identity with OYVMV-201Pk, OYVMV-301Pk and BYVMV-In. Among the Indian isolates, CLCuV-DI was identical to CLCuV-HI, and similarly, CLCuV-SI was

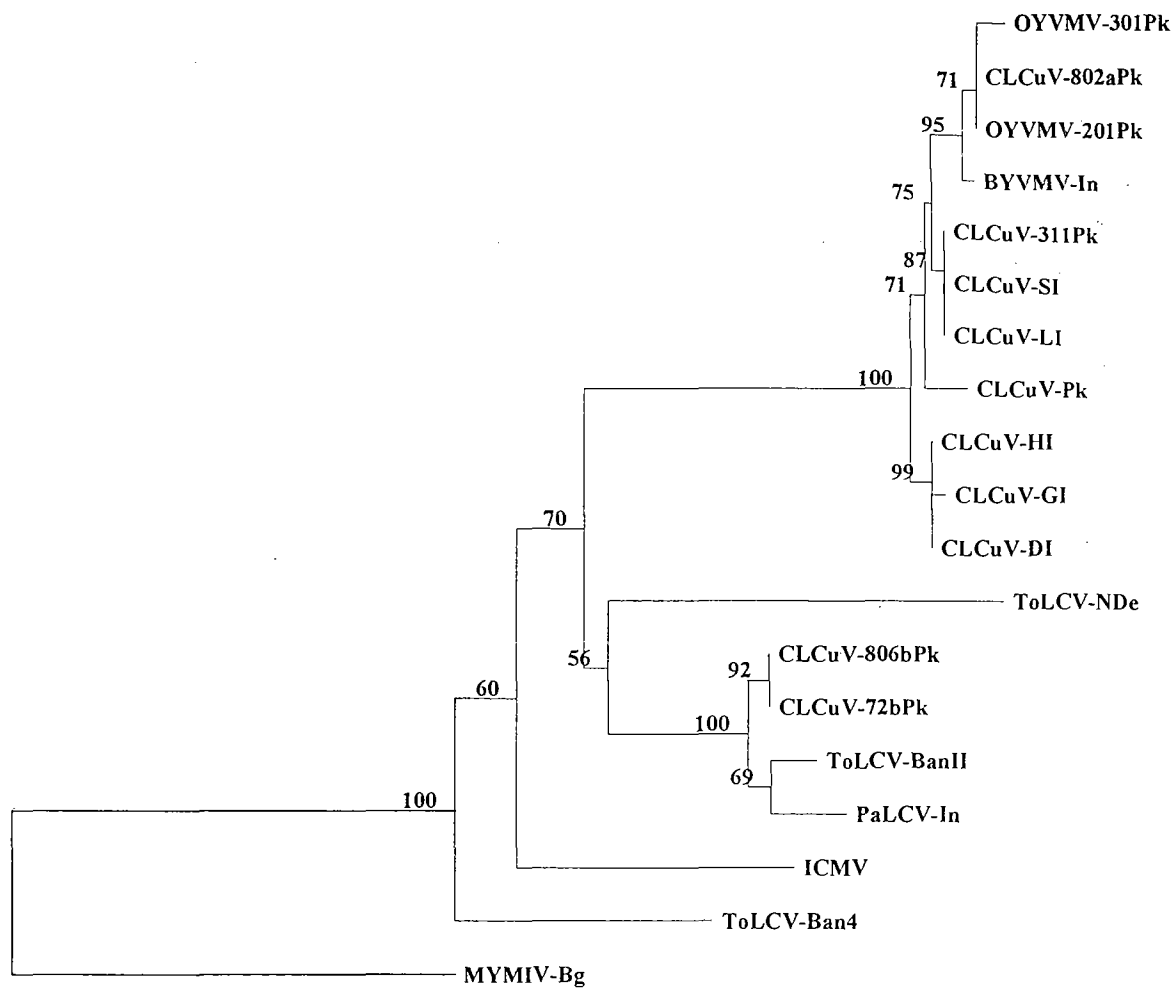


Table 4.5.6. Comparison of amino acid identity in the N' terminal of AC1 gene among begomoviruses from India and Pakistan

Virus	CLCuV-DI	CLCuV-GI	CLCuV-HI	CLCuV-LI	CLCuV-SI
CLCuV-DI	100.0				
CLCuV-GI	72.7	100.0			
CLCuV-HI	100.0	72.7	100.0		
CLCuV-LI	99.0	71.8	99.0	100.0	
CLCuV-SI	74.5	98.1	74.5	73.6	100.0
CLCuV-Pk	95.4	71.8	95.4	94.4	73.6
CLCuV-72b Pk	74.5	97.2	74.5	73.6	99.0
CLCuV-806b Pk	74.5	98.1	74.5	73.6	100.0
CLCuV-802a Pk	50.4	51.8	50.4	49.5	53.6
CLCuV-311 Pk	97.2	71.8	97.2	96.3	73.6
OYVMV-201 Pk	76.5	78.1	76.5	75.6	80.0
OYVMV-301 Pk	80.0	83.4	80.0	79.0	84.4
BYVMV-In	78.1	79.8	78.1	77.2	81.6
ICMV	78.3	78.1	78.3	77.4	80.0
PaLCV-In	47.7	48.1	47.7	46.8	50.0
ToLCV-NDe	76.5	75.4	76.5	75.6	77.2
ToLCV-BanII	76.5	80.0	76.5	75.6	80.9
ToLCV-Ban4	72.9	78.1	72.9	72.0	79.0
MYMIV-Bg	66.6	66.3	66.6	65.7	67.2

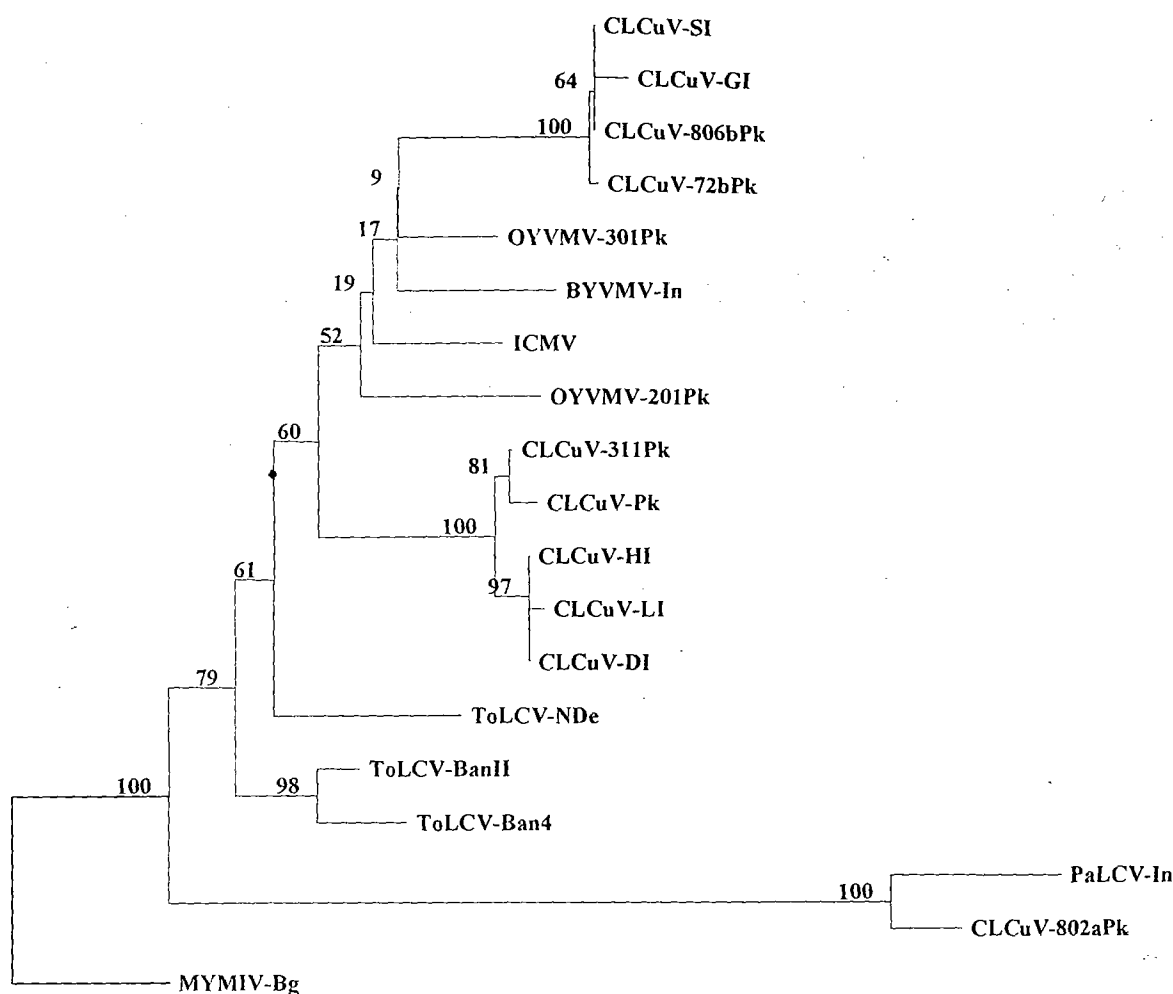


Figure 4.5.7. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the N' terminal ORF AC1. Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the MYMIV-Bg sequence.

identical to CLCuV-806bPk and near identical to CLCuV-72bPk in the N' terminal AC1. All the isolates in fact showed less identity with CLCuV-802aPk. With other begomoviruses like ICMV, ToLCV-BanII, ToLCV-Ban4 and ToLCV-NDe, it was between 72-80%.

Phylogenetic tree derived from the amino acid sequence of the N' terminal AC1 gene (Fig.4.5.7) was analogous to the phylogenetic tree obtained for the IR (Fig.4.5.3). Since N' terminal half of the Rep gene (AC1) contains the Rep domain that interacts with the iterons in the IR, variability of the IR coincides with the variability of N' terminal AC1 gene.

Comparison of DNA- β sequence

Partial nucleotide sequence data were generated for DNA- β associated with the four CLCuV isolates. For the isolate CLCuV-GI, full length DNA- β sequence have been reported earlier (GenBank Acc. No.- AY083590).

DNA- β nucleotide sequence data are available only for two Pakistan isolates of CLCuV, one Sudanese isolate of OLCV, one Indian isolate of BYVMV and one Singapore isolate of AYVV. Analysis of all the available sequence data revealed the diversity amongst DNA- β (Table 4.5.7). DNA- β of CLCuV-GI and CLCuV-DI are identical (100%), but all the other Indian isolates are considerably diverse (Table 4.5.7). CLCuV-GI and CLCuV-DI have closer sequence identity with the two isolates [CLCuV-Pk(1) and CLCuV-Pk(2)] from Pakistan than the other Indian isolates (Table 4.5.7). The sequence identity with DNA- β associated with OLCV-Sd, BYVMV-In and AYVV-Sg is even more remote (Table 4.5.7). The DNA- β of CLCuV-LI isolate is most distinct, showing 38-51% diversity with the known DNA- β . Absence of closer relationship was also seen with CLCuV-HI.

Table 4.5.7. Comparison of nucleotide identity in the CLCuV DNA- β of North Indian isolates with other begomoviruses associated DNA- β

Virus	CLCuV-DI	CLCuV-GI	CLCuV-HI	CLCuV-LI	CLCuV-SI
CLCuV-DI	100.0				
CLCuV-GI	100.0	100.0			
CLCuV-HI	70.8	70.8	100.0		
CLCuV-LI	56.1	56.1	48.7	100.0	
CLCuV-SI	83.9	83.9	63.0	52.6	100.0
CLCuV-Pk1	90.7	90.7	66.6	51.7	86.8
CLCuV-Pk2	91.3	91.3	63.7	53.4	86.2
BYVMV-In	58.5	58.5	52.8	49.6	58.2
OLCV-Sd	50.8	50.8	45.7	48.4	49.4
AYVV-Sg	55.5	55.5	48.7	62.1	53.3

All the DNA- β sequences have the highest level of identity over a stretch of 100 nucleotides upstream and 18 nucleotides downstream of the *Ori* of replication in the nonanucleotide TAATATTAC motif (Fig. 4.5.8). This region could be said as the highly conserved region present in the DNA- β associated with begomoviruses, irrespective of the host and geographical distribution of the virus. High degree of variability is observed in the upstream and downstream sequence of this highly conserved region.

Phylogenetic tree derived from the nucleotide sequences of the DNA- β associated with begomoviruses (Fig.4.5.9) shows that CLCuV-SI is closely related to CLCuV-Pk(1) and CLCuV-Pk(2) which cluster together, whereas, CLCuV-GI and CLCuV-DI are more closely related to each other and branch separately. In contrast, DNA- β of CLCuV-LI is very distinct and branches separately from the rest of the DNA- β compared.

DISCUSSION

In the present investigation the nucleotide sequencing has been done for the specific region encompassing N' terminal region of ORF AC1 (330 nucleotides), complete intergenic region, AV2 gene and AV1 gene. This region has been selected, as variability in begomoviruses is more seen in the intergenic region near the *Ori* of replication, as at this region most recombination occurs. Interestingly, no evidence was found for recombination in the IR or in the N' terminal AC1 gene. The association of the 5' half of the IR and its cognate AC1 gene was therefore maintained. This is probably required for viability because, for the genomic DNA to be replicated, a specific amino acid motif in the N' terminal portion of AC1 encoded Rep protein must recognize the iterons in the 5' half of the IR (Lazarowitz *et al*, 1992; Fontes *et al.*, 1994).

The comparison data clearly revealed that all the isolates share the core conserved region (23nt upstream and 23nt downstream of *Ori* of replication in the nonanucleotide

Figure 4.5.8. Multiple alignment of the nucleotide sequence of DNA- β of 10

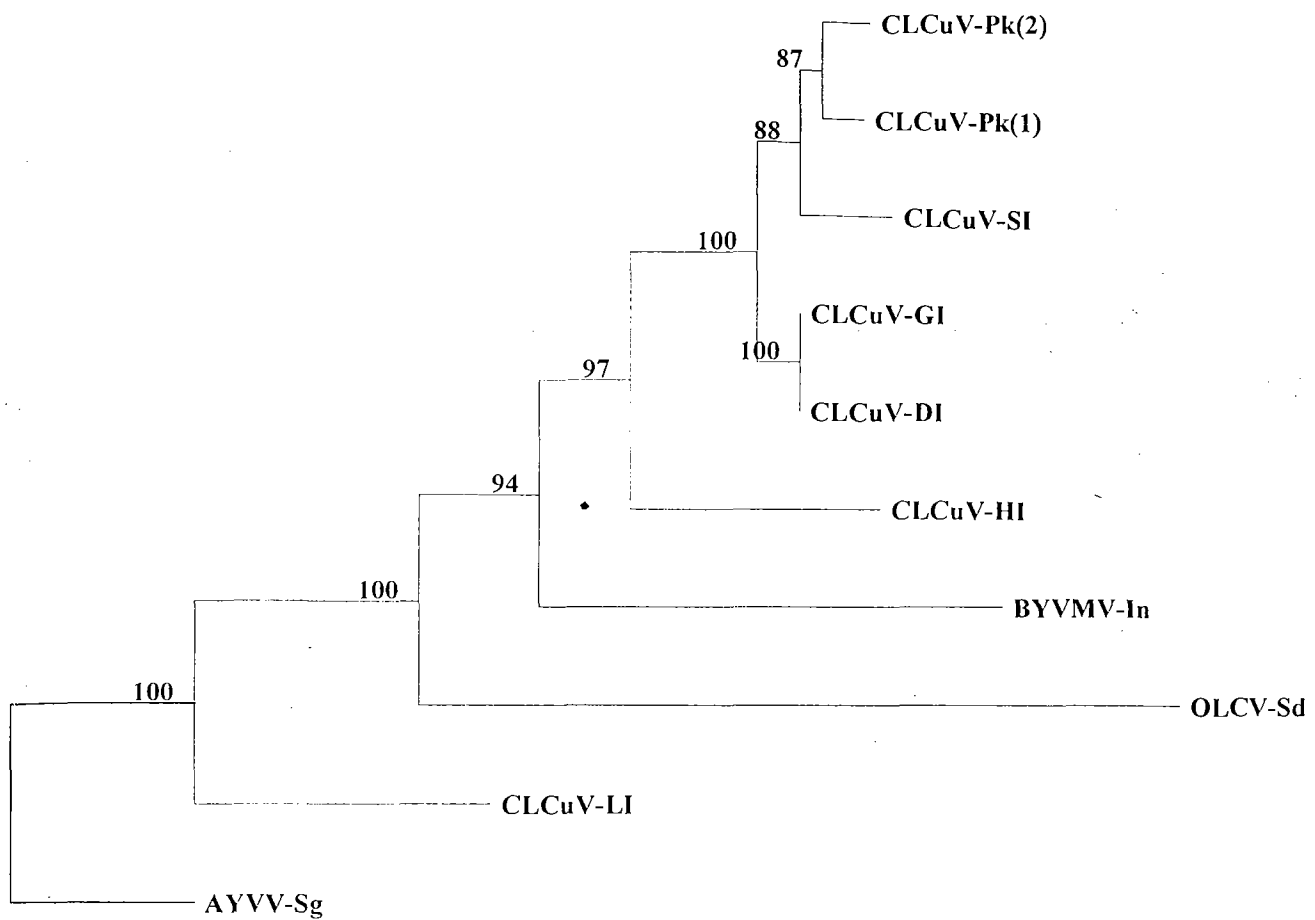


Figure 4.5.9. Phylogenetic relationship between various CLCuV isolates and other begomoviruses in the DNA- β . Dendrogram was constructed using the neighbour-joining method with bootstrapping (500 replicates) in Treecon for Windows version 1.3b on sequences aligned using CLUSTALW 1.7 version. The numbers at nodes refer to number of times (in percentages) in which the branching was supported. The tree was rooted on the AYVV-Sg sequence.

sequence) but shows variability at 5' intergenic region. After the nonanucleotides, at the 3' region again there is strong core that is highly conserved.

Intergenic region is the most specific part of DNA-A for a given *Begomovirus* (Harrison and Robinson, 1999). Comparison of IR nucleotide sequences, the Rep binding iterons and their arrangements give a good indication of relationship of different begomoviruses (Arguello-Astorga *et al.*, 1994; Zhou *et al.*, 1998; Sanz *et al.*, 2000). As the iteron sequences interact with replication initiation protein (Rep protein) of ORF AC1, the isolates showing high identity in IR are expected to share high percentage identity in N' terminal ORF AC1. This was also found in the CLCuV isolates. Three of the Indian isolates viz., CLCuV-DI, CLCuV-HI and CLCuV-LI showed high sequence identity at IR and N' terminal ORF AC1, and shared the same iteron sequences and their arrangements. But isolates CLCuV-GI and CLCuV-SI have different, iteron (TGATCA) like the isolates CLCuV-72bPk and CLCuV-806bPk from Pakistan.

On the contrary when comparison is done for the coat protein gene (AV1), close (>99% sequence identity) relationship was found among isolates CLCuV-DI, CLCuV-HI and CLCuV-GI and all begomoviruses from malvaceous hosts had nearly 95% sequence identity. A notable feature is the high level of per cent identity (>90%) with non-malvaceous begomoviruses too, like ICMV, PaLCV-In, ToLCV-NDe and ToLCV-BanII but not with ToLCV-Ban4 and MYMIV-Bg which share <80% sequence identity. It shows diversity even in AV1 gene which is the most conserved gene of begomoviruses (Harrison and Robinson, 1999).

ORF AV2 was found to be most distinctive, as it showed high (>94%) sequence identity among the Indian isolates and CLCuV-311Pk. Sequence identity of the Indian isolates was also >92% with CLCuV-802aPk and >90% with begomoviruses from

bhendi/okra, but <89% with CLCuV-Pk and <76% with CLCuV-72bPk and CLCuV-806bPk. ORF AV2, therefore, can be useful in distinguishing the begomoviruses occurring in the old world. The affinities in AV2 observed here show that precursor of some of the isolates could be BYVMV and other isolates might have originated from non-malvaceous hosts.

ORFs AV1 and AV2 of CLCuV-LI and CLCuV-SI greatly resemble (>98% identity in AV1 and >95% identity in AV2) that of OYVMV/BYVMV isolates, which is even greater than most of the CLCuV isolates compared, whereas, the rest of the region studied i.e., IR and N' terminal AC1 do not show high affinity with OYVMV/BYVMV isolates but resemble CLCuV isolates. In this region, i.e. in IR and N' terminal AC1, OYVMV/VYVMV sequences have a lower percentage identity than that of all the other CLCuV isolates, which suggests that some evolutionary drift has occurred since CLCuV-LI and CLCuV-SI evolved, or that the recombination involved a close relative of OYVMV/BYVMV instead of this virus itself. Although it is not clear in either of these examples involving OYVMV/BYVMV, whether the CLCuV isolate or the OYVMV/BYVMV isolate is the parental one, OYVMV/BYVMV seems to be the more likely alternative because okra/bhendi yellow vein mosaic disease is known to occur in the Indian sub-continent for a long time (Kulkarni, 1924; Varma, 1963), whereas CLCuD appeared only recently (Anonymous, 1990; Varma *et al.*, 1993).

Sequence analysis and comparison of North Indian isolates of CLCuV with the malvaceous Begomoviruses from Indian sub-continent has provided a clear picture of the extent of variation among them. This study clearly showed that there are at least four distinct forms of CLCuV DNA-A component in North India. One form comprising CLCuV-DI and CLCuV-HI isolates, the second very distinct and separate represented by

CLCuV-LI, the characteristic third form is that of CLCuV-GI and fourth form CLCuV-SI, which shows relationship with all the isolates. All the four forms of DNA-A showed varying level of affinity with other CLCuV and BYVMV/OYVMV isolates in different regions of their genomes, indicating possible recombinant events in nature. Because of the many potential sites available for homologous recombination between isolates of the same virus, intraspecies recombination might be expected to be the most frequent kind (Harrison and Robinson, 1999).

When all the genomic regions are taken together different isolates from India fall into two distinct groups. The first group consists of CLCuV-DI, CLCuV-HI and CLCuV-LI and the second group consists of CLCuV-GI, and CLCuV-SI. CLCuV-DI, CLCuV-HI and CLCuV-LI resemble each other closely in iteron, AV1, AV2 and N-terminal AC1, however, IR of CLCuV-DI is different having less than 84% sequence identity with the other two isolates of the group. As IR is considered to be most specific for a particular virus (Harrison and Robinson, 1999), CLCuV-DI could also be taken as distinct from CLCuV-HI and CLCuV-LI inspite of the similarities in other regions of their genome. Members of the second group i.e., CLCuV-GI and CLCuV-SI, however, resemble each other in all the regions compared.

The overall sequence identity of 1530 bases compared also shows that CLCuV-DI shares 94.3% identity with CLCuV-HI and CLCuV-HI shares 92.7% identity with CLCuV-LI (Table 4.5.8). These three isolates, therefore, appear to belong to one viral species of *Begomovirus*. The overall sequence identity of these isolates with the other two isolates i.e., CLCuV-GI and CLCuV-SI was much greater. Whether these two groups form two different viral species can be ascertained only when full-length sequence of

Table 4.5.8. Comparison of nucleotide sequence identity in the N'terminal AC1, IR, AV2 and AV1 regions (1530 bp) among CLCuV isolates from North India.

Virus	CLCuV-DI	CLCuV-GI	CLCuV-HI	CLCuV-LI	CLCuV-SI
CLCuV-DI	100				
CLCuV-GI	89.5	100			
CLCuV-HI	94.3	87.9	100		
CLCuV-LI	87.8	81.4	92.7	100	
CLCuV-SI	81.5	91.1	84.3	84.8	100

DNA-A of these isolates is available, however, there are indications of existence of two different species of begomoviruses causing leaf curl disease in India.

Variability observed in DNA- β was unexpected, as in a small geographical region, the DNA- β associated with DNA-A component was expected to show greater identity. DNA- β of CLCuV-GI and CLCuV-DI were near identical and had high sequence identity (>90%) with CLCuV-Pk(1) and CLCuV-Pk(2) followed by CLCuV-SI (83.9% identity). Highest diversity (>43%) was observed in DNA- β of CLCuV-LI amongst the CLCuV isolates, but it resembled more (62% identity) with DNA- β of AYVV-Sg. This study clearly showed that there are at least four distinct CLCuV DNA- β components in North India. One type is associated with CLCuV-DI and CLCuV-GI, next to is CLCuV-SI DNA- β which shares 52-84% similarity with the first type, the third type is of CLCuV-HI DNA- β with homology of 49-71% and the fourth type is represented by CLCuV-LI DNA- β having 49-56% homology.

An interesting finding is that all the DNA- β sequences have a highly conserved region (which is near identical) 100nt upstream and 18nt downstream of the nonanucleotide TAATATTAC motif, irrespective of the host and geographical distribution of the virus. The significance of this conserved region in all the known DNA- β shows that this region must be important for association with begomoviruses. Actual function of DNA- β , however, is yet to be understood.

5. SUMMARY

Cotton leaf curl disease (CLCuD) caused by whitefly transmitted *Cotton leaf curl virus* (CLCuV), (genus *Begomovirus*, family Geminiviridae) is a major constraint in cotton cultivation in North-West India. Association of geminivirus-like particles with CLCuV affected plants was demonstrated by ELISA and ISEM earlier, but no information was available on the biology, molecular characteristics and variability of the virus in India. Therefore, present studies were undertaken with the following objectives: (1) biological characterization of an isolate of CLCuV from Rajasthan, (2) cloning of the selected isolate and its molecular characterization, (3) detection of additional genomic components, (4) molecular detection of CLCuV genomic components and (5) determining variability among CLCuV isolates in North-West India.

An isolate of CLCuV from Sri Ganganagar (CLCuV-GI), the major cotton growing area in North-West India, was characterized. The virus could be transmitted by whitefly (*Bemisia tabaci*) when given a minimum acquisition and inoculation access period of 30 minutes and 10 minutes, respectively. *B. tabaci* could transmit the virus efficiently from *Gossypium hirsutum* to *G. hirsutum* and *G. barbadense* and seven other species belonging to plant families Fabaceae, Malvaceae and Solanaceae. All the infected plants developed typical leaf vein thickening and curling characteristic of CLCuV infection. The virus could be back inoculated from these hosts to *G. hirsutum* cv. LH 900. To identify sources of resistance to CLCuV, 126 genotypes of *G. hirsutum*, 19 of *G. barbadense*, 14 of *G. arboreum* and two of *G. herbaceum* were tested against CLCuV by whitefly inoculation. All the genotypes of *G. arboreum* and *G. herbaceum* were found to be highly resistant (no infection) whereas all the genotypes of *G. barbadense* were highly susceptible (100% infection);

in *G. hirsutum*, 62 genotypes were highly resistant, 5 genotypes highly susceptible and the remaining genotypes showed varying degree of resistance.

Full length CLCuV-GI DNA-A and DNA-2 of Sri Ganganagar isolate were amplified from infected cotton plants by PCR and cloned. Tandem dimeric constructs of full length CLCuV-GI DNA-A and DNA- β were made in pBin19 vector. Typical symptoms were obtained in cotton (*Gossypium hirsutum*) cv. LH900 when these genomic components were delivered through *Agrobacterium tumefaciens* strain EHA105. From such agroinoculated infected cotton plants, CLCuV was back transmitted by whiteflies to cotton (*G. hirsutum*) cv. LH 900 and *G. barbadense* line GB14-193, producing typical cotton leaf curl symptoms. The confirmation of Koch's postulates through agroinoculation and whitefly transmission of CLCuV from agroinoculated plants has been demonstrated for the first time in the world. The infectious clones were sequenced completely (DNA-A, 2754bp, GenBank Acc. No.AF363011; DNA-2, 1350bp, GenBank Acc. No.AY083590). The genome organization of CLCuV-GI DNA-A showed all the features of DNA-A of bipartite begomoviruses. CLCuV-GI differs from all earlier reported CLCuV isolates in sequence identity by more than 11%, clearly indicating that it is a new species of *Begomovirus* as per the ICTV criteria of defining species of begomoviruses. *In silico* analysis of CLCuV-GI DNA- β nucleotide sequence revealed that there are nine ORFs, which code for proteins of > 4 kDa. Phylogenetic analysis showed that CLCuV-GI DNA- β is related to CLCuV DNA- β of an isolate from Pakistan.

Both, sub-genomics and novel DNA molecules (NDMs) were found associated with plants infected with CLCuV and viruliferous whiteflies. The sub-genomic molecules were nearly a half size of full-length DNA-A. NDMs are not only associated with CLCuV in infected plants but are also encapsidated and transmitted by whiteflies. These NDMs do not have sequence homology with CLCuV DNA-A,

DNA- β or DNA-1 and their sequences were of unknown origin except for nonanucleotide sequence.

For rapid and sensitive diagnosis of CLCuV nucleo-diagnostic methods have been developed. Nucleic acid spot hybridization (NASH), using radiolabelled probes to CLCuV DNA-A and CLCuV DNA- β , was found efficient in detecting CLCuV both in plants and viruliferous whiteflies. These probes also detected infection by other whitefly transmitted geminiviruses (WTGs) in crop, ornamental and weed plants. Hybridization of DNA- β probe with plant samples naturally infected with other WTGs shows that the association of DNA- β molecules could be a common feature of WTGs in northern India. CLCuV was also efficiently detected by PCR using primers to amplify coat protein gene and full length DNA- β molecules from infected plants and viruliferous whiteflies. By PCR, using DNA- β primers, many sub-genomic DNA- β forms were also amplified from CLCuV infected test hosts and viruliferous whiteflies. The number of sub-genomic forms of DNA- β varied in different hosts suggesting involvement of some host factors in their development.

Sequence analysis and comparison of North Indian isolates of CLCuV DNA-A with other malvaceous begomoviruses from Indian sub-continent has shown considerable variation among them. Based on sequence identity, the five isolates from North-West India have been shown to be of four distinct types, i.e., CLCuV-DI and CLCuV-HI, CLCuV-LI, CLCuV-GI and CLCuV-SI. All the four types of DNA-A showed varying level of affinity with other CLCuV and BYVMV/OYVMV isolates in different regions of their genomes, indicating possible recombinant events in nature.

Similarly, sequence analysis and comparison of CLCuV DNA- β clearly showed that there are at least four distinct CLCuV DNA- β components in North India. One type is associated with CLCuV-DI and CLCuV-GI, next to is CLCuV-SI

DNA- β which shares 52-84% similarity with the first type, the third type is of CLCuV-HI DNA- β with homology of 49-71% and the fourth type is represented by CLCuV-LI DNA- β having 49-56% homology.

7-7125

BIBLIOGRAPHY

- Abu Salih, H.S. (1965). Rep. Res. Div. Minist. Agric. Sudan: 53-54.
- Ajmera, B.D. (1994). Occurrence of leaf curl virus on American Cotton (*G.hirsutum*) in North Rajasthan. *Poster presentation, National Seminar on Cotton Production Challenges in 21st Century*, April 18-20, Hisar, India.
- Ajmera, B.D. (1996). Status and strategies of leaf curl of cotton in north India. *Paper presented at National Seminar on century of cotton in India*, December 21st Surat, India.
- Ajmera, B.D. (1997). Status of leaf curl of *hirsutum* cotton in north western Rajasthan. *Poster presentation, International conference on Integrated plant disease management for sustainable agriculture*, November 10-15, New Delhi, India.
- Ali, M., Ahmed, Z., Tanveer, M. and Mehmood, T. (1995). Cotton leaf curl virus (CLCV) in the Punjab. *Current situation and review of Central Cotton Research Institute*, Multan, Pakistan. 117 p.
- Andrews, F.W. (1936). The effect of leaf curl disease on the yield of cotton plant. *Cotton Growers Review*. XIII : 287-293.
- Anonymous (1990). *Cotton leaf curl virus is a whitefly transmitted geminivirus*. Annual Report, IARI.
- Anonymous (1994). Leaf curl disease of cotton in north west India. *Report of the ICAR committee*. 15p.
- Anonymous (2001). Agriculture statistics at a glance. Published by Ministry of Agriculture, Govt. of India.
- Argeüllo-Astorga, G.R., Guevara-González, L.R., 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: 90-100.
- Asad, S., Haris, W.A.A., Bashir, A., Zafar, Y., Malik, K.A., Malik, N.N. and Lichtenstein, C.P. (2001). Expression of antisense RNA of viral genes confers resistance against cotton leaf curl disease in tobacco. In: *3rd International Geminivirus Symposium*. July 24-28, 2001, John Innes Centre, UK, p.8.
- Ashraf, M. and Zafar, Z.U. (1999). Some physiological characteristics in resistant and susceptible cotton cultivars infected with cotton leaf curl virus. *Biologia Plantarum*, 42(4): 615.

- Ashraf, M., Zafar, Z.U., McNeilly, T. and Veltkamp, C.J. (1999). Some morpho-anatomical characteristics of cotton (*Gossypium hirsutum* L.) in relation to resistance to cotton leaf curl virus (CLCuV). *Angewandte – Botanik*. **73**(3-4): 76-82.
- Bedford, H.W. (1938) Entomological Section, Agricultural Research Service, Report 1936-37. *Report, Agricultural Research Service, Anglo Egyptian Sudan, 1937*, 50-65.
- Berrie, L.C., Rybicki, E.P., and Rey, M.E.C. (2001). Complete nucleotide sequence and host range of South African cassava mosaic virus: Further evidence for recombination amongst begomoviruses. *J. Gen Virol*. **82**: 53-58.
- Bink, F.A. (1973). A new contribution to the study of cotton mosaic in Chad. I. Symptoms, transmissions by *Bemisia tabaci* Genn. II. Observations on *B. tabaci*. III. Other various diseases on cotton and related plants. *Coton et Fibres Tropicales* **28**: 365-378.
- Bink, F.A. (1975). Leaf curl and mosaic diseases of cotton in Central Africa. *Empire Cotton Growing Review*. **52**: 133-141.
- Bisaro, D.M. (1996). Geminivirus DNA replication. In: *DNA Replication in Eukaryotic Cells* (Ed., DePamphilis, M.L.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 833-854.
- Bridson R.W. and Markham, P.G. (2000). Cotton leaf curl virus disease. *Virus Res.*, **71**: 151-159.
- Bridson, R.W. and Markham, P.G. (1994). Universal primers for the PCR amplification of dicot-infecting geminiviruses. *Molecular Biotechnology*, **1**: 202-205.
- Bridson, R.W. and Markham, P.G. (2000). Cotton leaf curl virus disease. *Virus Research*, **71**:151-159.
- Bridson, 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.
- Bridson, R.W., Liu, S., Pinner, M.S., and Markham, P.G. (1998). Infectivity of African cassava mosaic virus clones to cassava by biolistic inoculation. *Arch. Virol*. **143**: 2487-2492.
- Bridson, R.W., Mansoor, S., Bedford, I.D., Pinner, M.S., and Markham, P.G. (2000). Clones of cotton leaf curl geminivirus induce symptoms atypical of cotton leaf curl disease. *Virus Genes*. **20**: 17-24.
- Bridson, 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.

- Brown, J.K. (1992). Virus diseases of cotton. In: Cotton Diseases (Ed. R.J. Hillocks). Commonwealth Agricultural Bureaux International, Oxon, UK, 415 pp.
- Brown, J.K. (1994). Current status of *Bemisia tabaci* as a plant pest and virus vector in agro eco systems world wide. In: *FAO Bulletin* **42/1-2**: 86p.
- Brown, J.K. (1996). Molecular biology and epidemiology of subgroup III, *Geminiviridae*. Plant-Microbe Interactions Review Series, (Eds. G. Stacey and N. Keen). Chapman and Hall, pp.125-195.
- Brown, J.K. and Nelson, M.R. (1984). Geminata particles associated with cotton leaf crumple disease in Arizona. *Phytopathology*, **74**: 987-990.
- Buragohain A.K., Sung, Y.K., Coffin, R.S. and Coutts, R.H.J.A. (1994). The infectivity of dimeric potato yellow mosaic geminivirus clones in different hosts. *J. Gen. Virol.*, **75**: 2857-2861.
- Cauquil, J. and Follin, J.C. (1983). Presumed virus and mycoplasma-like organism diseases in subsaharan Africa and the rest of the world. *Coton et Fibers Tropicales* . **38**: 293-317.
- Chakraborty, S. (1996). Biological and molecular characterization of mungbean yellow mosaic geminivirus. Ph.D. Thesis, IARI, New Delhi, 150pp.
- Chatterji, U. (1994). Characterization of tobacco leaf curl geminivirus (TLCV). Ph.D. thesis, IARI, New Delhi p.101.
- Chenelu, V.V. and Varma, A. (1988). Virus and virus like diseases of pulse crops commonly grown in India. pp 338-370. In: *Pulse Crops* (Eds., Baldev, B., Ramanujam, S., and Jain, H.K.), Oxford and IBH, New Delhi.
- Chilton, M.D., Saiki, R.K., Yadav, N., Gordon, M.P. and Quetier, F. (1980). tDNA from *Agrobacterium* Ti plasmid is the nuclear DNA fraction of crown gall tumour cells. *Proc. Nat. Acad. Sci., USA*, **77**: 4060-4064.
- Cook, O.F. (1920). A disorder of cotton plants in China: club-leaf or crytosis. *Journal of Heredity*. **11**: 99-110.
- Cook, O.F. (1923). Malformation of cotton plants in Haiti: a new disease named smalling or stenosis, causing abnormal growth and sterility. *Journal of Heredity*. **14**: 323-325.
- Costa, A.S. (1955). Studies on *Abutilon* mosaic in Brazil. *Phytopathologische Zeitschrift*. **24**: 97-112.
- Costa, A.S. (1957). Anthocyanosis, a virus disease in cotton in Brazil. *Phytopathologische Zeitschrift*. **28**: 167-186.
- Costa, A.S. (1961). Studies on Brazilian tobacco streak. *Phytopathologische Zeitschrift*. **42**: 113-138.

- Costa, A.S. and Forster, R. (1938). Nota preliminary sobre une nova molestia de virus do algodoeiro: mosaico das nervura. *Compinas Institute Agronomico Boletim Technico*, 51.
- Costa, A.S. and Sauer, H.F.G. (1954). Vermelhao do algodoeiro. *Bragantia* 13: 237-246.
- Couteaux, L., Lefort, P.L. and Kaukuvu, E. (1968). Some observations on cotton leaf curl in *G. barbadense* at the Anie Mono Station, *Cotton Fibr. Trop.* 23: 506-507.
- Cowland, J.W. (1934). Gezira entomological section, G.A.R.S. *Final report on experimental work. 1932-33. A Rept. Gezira Agric. Res.*, pp. 107-125.
- Czosnek, H. and Laterrot, H. (1997). A world wide survey of tomato yellow leaf curl viruses. *Arch. Virol.* 142: 1391-1406.
- Czosnek, H., Kheyr-Pour, A., Gronenborn, B., Remetz, E., Zeidan, M., Altman, A., Rabinowitch, H.D., Vidavsky, S., Kedar, N., Gafni, Y. and Zamir, D. (1993). Replication of tomato yellow leaf curl virus (TYCLV) DNA in agroinoculated leaf tissue from selected tomato genotypes. *Pl. Mol. Biol.*, 22: 95-1005.
- Davies, W.J. and Stanley, J. (1989). Geminivirus genes and vectors. *Trends in Genetics*, 5: 77-81.
- Delalande, P. (1970). *Quelques observations sur la maladie bleue des contonniers. Campagne 1969-70, rapport non public.* Station centrale de Bambari, RCA and IRCT, Paris.
- Dellaporta, S.L., Wood, J. and Hicks, J.D. (1983). A plant DNA miniprep: Version II. *Plant Mol. Biol. Repr.*, 4: 19-21.
- Desbiez, C., David, C., Mettouchi, A., Lauffs, J. and Gronenborn, B. (1995). Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proc. Natl. Acad. Sci., USA.* 92: 5640-5644.
- Dickson, R.C., Johnson, M.McD. and Laird, E.F. (1954). Leaf crumple, a virus disease of cotton. *Phytopathology*, 44: 479-480.
- Ditta, G., Stanfield, S., Corbin, D. and Helinski, D.R. (1980). Broad host range DNA cloning system for gram negative bacteria. Construction of gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA.* 77: 7347-7351.
- Dry, I.B., Rigden, J.E., Krake, L.R., Mullineaux, P.M. and Rezaian, M.A. (1993). Nucleotide sequence and genome organization of tomato leaf curl geminivirus. *J. Gen. Virol.* 74: 147-151.
- Eagle, P.A., Orozco, B.M. and Hanley-Bowdoin, L. (1994). A DNA sequence required for geminivirus replication also mediates transcriptional regulation. *Plant cell* 6: 1157-1170.

- Elmer, J.S., Brand, I., Sunter, G., Gardiner, W.E., Bisaro, D.M. and Rogers, S.G. (1988). Genetic analysis of the tomato golden mosaic virus II: The product of the AL1 coding sequence is required for replication. *Nucl. Acids Res.* **16**: 7043-7060.
- El-Nur, E. (1967). Annual report. Agric. Res. Div. Sudan 1966-67. pp.129-31.
- El-Nur, E. and Abu Salih, H.S. (1970). Cotton leaf curl virus disease. *PANS*, **16**: 121-131.
- El-Nur, E. and Abu-Salih, H.S. (1966). Annual Report of Research Division, Ministry of Agriculture Sudan 1965-66 pp 118-120.
- Etessami, P., Callis, R., Ellwood, S. and Stanje, J. (1988). Delimitation of essential genes of cassava latent virus DNA-2. *Nucl. Acids Res.* **16**: 4811-4829.
- Fadlala, A.S. and Ibrahim, G. (1999). Release of "Barakat S" a new better quality extra long cultivar of Sudan Sakel cotton, resistant to Fusarium wilt. *Sudan J. of Agricultural Res.* **2**: 95.
- Farquharson, C.O. (1912). *Report of a mycologist*. Annual Report. Agriculture Department, Nigeria.
- Fauquet, C., Fargette, D. and Thouvenel, J.C. (1986). Development of African cassava mosaic virus at a Regional level in Ivory Coast. *Proceeding of the Workshop on Epidemiology of plant virus diseases* (Orlando, F.L.) pp.VII-22-24.
- Fauquet, C.M. Maxwell, D.P., Gronenborn, B. and Stanley, J. (2000). Revised proposal for naming geminiviruses. *Arch. Virol.* **145**(8): 1743-1761.
- Feinberg, A.P. and Vogelstein, B.(1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**:6-13.
- Fontes, E.P.B., Gladefelter, H.J., Schaffer, R.L., Petty, I.T.D., and Hanley-Bowdoin, L. (1994). Geminivirus replication origins have a modular organization *Plant Cell.* **6**: 405-416.
- Fontes, E.P.B., Lucknow, V.A. and Hanley-Bowdoin, L. (1992). A geminivirus replication proteins is a sequence specific DNA binding protein. *Plant Cell* **4** : 597-608.
- Garzon-Tiznado, J.A., Torres-Pacheco, I., Ascencio-Ibanez, J.T., Herrera-Estrella, L. and Rivera-Bustamante, R.F. (1993). Inoculation of peppers with infectious clones of a new geminivirus by biolistic procedure. *Phytopathology.*, **83**: 514-521.
- Gerling, D., (1986). Natural enemies of *Bemisia tabaci*, biological characteristics and potential as biocontrol agents. A review. *Agric. Ecosys. Environ.* **17**: 99-110.

- Gerling, D., Horowitz, A.R. and Paumgartney 1986. Autecology of *Bemisia tabaci* *Agriculture Eco systems and Environment*, **17**: 5-19.
- Giha, O.H. and Nour, M.A. (1969). Epidemiology of cotton leaf curl in the Sudan. *Cotton Growing Review*. **46**: 105-118.
- Gilbertson, R.L. and Lucas, W.J. (1996). How do viruses traffic on the 'vascular highway'? *Trends Plant Sci.* **1**: 260-268.
- Golding, F.D. (1930). A vector of leaf curl of cotton in Southern Nigeria. *Emp. Cotton Growers Review* **VII**: 120-126.
- Goodman, R.M. (1977). Single-stranded DNA genome in a whitefly-transmitted plant virus. *Virology* **83**: 171-179.
- Grimsley, N., Hohn, B., Hohn, T. and Walden, R. (1986). Agroinfection, an alternative route for viral infection of plants by using the Ti plasmid. *Proc. Nat. Acad. Sci., USA*, **83**: 3282-3286.
- Grimsley, N., Hohn, T., Davis, J.W. and Hohn, B. (1987). *Agrobacterium* mediated delivery of infectious maize streak virus into maize plants. *Nature*, **325**: 177-179.
- Gupta, G.P. (2001). Alternate strategies for insect pest management in cotton. In: *Alternate Strategies for Insect Pest Management in Major Crops* (Ed. V.K. Sehgal). Division of Entomology, IARI, New Delhi, 232-236 p.
- Gurdeep, K., Sohal, B.S., Singh, J., Bajaj, K.L., Kaur, G. and Singh J. (1998). Influence of cotton leaf curl virus on the polyphenol metabolism of resistant and susceptible cotton leaves. *Plant Disease Research*, **13**(1): 23-27.
- Gutierrez, C. (2000). DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J.* **19**: 792-799.
- Hameed, S., Khalid, S., Ehsan-ul-Haq and Hashrni, A.A. (1993). Cotton leaf curl disease in Pakistan caused by a whitefly transmitted geminivirus. *Plant Dis.* **78**: 529.
- Hameed, S., Khalid, S., Ehsan-Ul-Haq and Hashrni, A.A. (1994). Cotton leaf curl disease in Pakistan is caused by a whitefly transmitted geminivirus. *Plant Disease*. **78**:529.
- Hamilton, W.D.O., Bisaro, D.M., Coutts, R.H.A. and Buck, K.W. (1983). Demonstration of the bipartite nature of the genome of single stranded DNA virus by infection with the cloned DNA components. *Nucl. Acids Res.* **11** : 7387-7396.
- Hanley-Bowdoin, L., Settloge, S.B., Orozco, B.M., Nagar, S. and Robertson, D. (1999). Geminiviruses: Models of plant DNA replication, transcription and cell cycle regulation. *Crit. Rev. Plant Sci.* **18**: 71-106.

- Harrison, B. D. and Robinson, D. J. (1999). Natural genomic and antigenic variation in whitefly-transmitted geminiviruses. *Annu. Rev. Phytopath.* **37**: 369-398.
- Harrison, B.D. (1985). Advances in geminivirus research. *Annu. Rev. Phytopathol.* **23**: 55-82.
- Harrison, B.D., Barker, H., Bock, K.R., Guthrie, E.J., Meredith, G. and Atkinson, M. (1977). Plant viruses with circular single stranded DNA. *Nature* **270**: 760-762.
- Harrison, B.D., Liu, Y.L. Khalid, S., Hameed, S., Otim-Nape, G.W. and Robinson, D.J. (1997). Detection and relationships of cotton leaf curl virus and allied whitefly transmitted geminiviruses occurring in Pakistan. *Ann. Appl. Biol.* **130**: 61-75.
- Hoefer, P., Engel, M., Jeske, H., Frischmut, T. (1997). Nucleotide sequence of a new bipartite geminivirus isolated from the common weed *Sida rhombifolia* in Costa Rica. *J. Gen. Virol.*, **78**: 1785-1790
- Hong, V.G. and Harrison, B.D. (1995). Nucleotide sequences from tomato leaf curl viruses from different countries: evidence for three geographically separate branches in evolution of the coat protein of whitefly-transmitted geminiviruses. *J. Gen. Virol.* **76** :2043-49.
- Howarth, A.J., Caton, J., Bossert, M., and Goodman, R.M. (1985). Nucleotide sequence of bean golden mosaic virus and a model for gene regulation in geminiviruses. *Proc. Natl. Acad. Sci. USA.* **82** : 3972-3976.
- Hussain , T. and Ali, M. (1975). A review of cotton diseases of Pakistan. *The Pakistan Cotton*, **19**: 71-86.
- Idris, A.M. (1990). Cotton leaf curl virus disease in Sudan, *Med. Fac. Lanbow, Rijksunir, Gent*, **55**(29)L: 1990-2029.
- Ingham, D.J., and Lazarowitz, S.G. (1993). A single missense mutation in the BR1 movement protein alters the host range of squash leaf curl virus. *Virology* **196** : 694-770.
- Ingham, D.J., Pascal, E. and Lazarowitz, S.G. (1995). Both bipartite geminivirus movement proteins define viral host range, but only BL1 determines viral pathogenicity. *Virology* **207** : 191-204.
- Iqbal, M.J., Azia, N., Saeed, N.A., Zafar, Y. and Malik K.A. (1997). Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. *Theor. Appl. Genet.* **94**: 139-144.
- Jones, G.H. and Mason, I.G. (1926). Studies on two obscure diseases of cotton. *Ann. Bot.*, **1**, **160**: 759-72.
- Jose, J. and Usha, R. (2002). Molecular characterization of bhendi yellow vein mosaic virus. In: *XIV Annual Conference and National Seminar, Indian Virological Society*, Bangalore, Jan, 18-20, 2002. (Abstract) p.18.

- Khan, N.U., Abro, H.K., Gul, H., Kumbhar, M.B. and Abdur, R. (2000). Leaf curl virus in American upland cotton in NWFP. *Pakistan J. Phytopathology*. **12**(2): 79-86.
- Khetarpal, R.K., Maury, Y., Cousin, R., Burghofer, A. and Varma, A. (1990). Studies on resistance on pea seed borne mosaic virus and new pathotypes. *Ann. Appl. Biol.* **116**: 297-304.
- Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G.P., Crespi, S. and Gronenborn, B. (1991). Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. *Nucleic Acid Res.* **19**: 6763-6769.
- Kirkpatrick, T.W. (1930). Leaf curl in cotton. *Nature*. **125**: 672.
- Kirkpatrick, T.W. (1931). Further studies on leaf-curl of cotton in the Sudan. *Bulletin of Entomological Research*. **22**: 323-363.
- Krishnareddy, M. (1989). Studies on yellow mosaic and leaf crinkle disease of blackgram. Ph.D. Thesis, IARI, New Delhi, 263 pp.
- Kulkarni, G.S. (1924). Mosaic and other related diseases of crops in the Bombay Presidency. *Poona Agricultural College Magazine*. **6**: 6-12.
- Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S.G., Schell, J. and Gronenborn, B. (1995). *In vivo* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Nat. Acad. Sci., USA*, **92**: 3879-3883.
- Lazarowitz, S. (1992). Geminiviruses : Genome structure and gene function. *Crit. Rev. Plant Sci.* **11** : 327-349.
- Lazarowitz, S.G. and Lazdins, I.B. (1991). Infectivity and complete nucleotide sequence of the cloned genomic components of a bipartite squash leaf curl geminivirus with a broad host range phenotype. *Virology* **180** : 58-69.
- Lazarowitz, S.G., Wu, L.C. Rogers, S.G. and Elmer, J.S. (1992). Sequence specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. *Plant Cell* **4** : 800-809.
- Lewin, C.J. 1927. Incidence of leaf curl of cotton in Southern Nigeria. *Sixth Ann. Bull. Deptt. Agric. Nigeria*. **6**: 70-77.
- Liu, Y., Robinson, D.J., and Harrison, B.D. (1998). Defective forms of cotton leaf curl virus DNA-A that have different combinations of sequence deletion, duplication, inversion and rearrangement. *J. Gen. Virol.* **79**: 1501-1508.
- Mali, V.R. (1977). Cotton leaf crumple virus disease – a new record for India. *Indian Phytopathology*. **30**: 326-329.
- Mali, V.R. (1978) Infectious variegation – a new virus disease of cotton in India. *Current Science*. **47**: 304-305.

- Mali, V.R. (1979). Viral wilt – a new disease hitherto unrecorded on cotton. *Current Science*. **48**: 687-688.
- Mandal, B. (1994). Molecular characterization of mungbean yellow mosaic virus. Ph.D. Thesis, IARI, New Delhi.
- Mandal, B., Varma, A. and Malathi, V.G. (1997). Systemic infection of *Vigna mungo* using the cloned DNAs of the blackgram isolate of mungbean yellow mosaic geminivirus through agroinoculation and transmission of the progeny virus by whiteflies *J. Phytopath.* **145**: 503-510.
- Mansoor, S., Bedford, I.D., Pinner, M.S., Stanley, J. and Markham, P.G. (1993). Whitefly transmitted geminivirus associated with cotton leaf curl disease in Pakistan. *Pakistan Journal of Botany*. **25**: 105-107.
- Mansoor, S., Briddon, R.W., Vionnett, O., Markham, P.G. and Baulcombe, D. (2001). Understanding the role of DNA- β in cotton leaf curl disease in Pakistan. In: *3rd International Geminivirus Symposium*, July 24-29, 2001, John Innes Centre, U.K. p.58.
- Mansoor, S., Hussain, M., Yasir-ul-Musawar, Khan, M.A., Iram, S., Zafar, Y., Briddon, R.W., Markham, P.G. and Malik, K.A. (2001). Strategies for broad-spectrum resistance against cotton leaf curl and tomato leaf curl disease in Pakistan. In: *3rd International Geminivirus Symposium*, July 24-29, 2001, John Innes Centre, U.K. p.59.
- 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., Mukhtar, S., Hussain, M., Amin, I., Zafar, Y. and Malik, K.A. (2000). Widespread occurrence of cotton leaf curl virus on radish in Pakistan. *Plant Disease*, **84**:809.
- Mariappan, V. and Narayanasamy, P. (1986). Epidemiology of tomato leaf curl disease. In: *National Seminar on Whitefly Transmitted Plant Virus Diseases*. June. 25-27. IARI. New Delhi. 28p.
- Massay, R.E. and Andrews, F.W. (1932). The leaf curl disease of cotton in the Sudan. *Empire Cotton Growers Review*. **9**: 32-45.
- Massey, R.E. 1934. Report of the Annual Agric. Res. Gezira, Sudan, 1933, pp. 126-146.
- Melamad-Madjar, V.S., Cohen, M., Chen, and Tam, S. (1978). Determination of whitefly population in cotton. *Hassadeh* **58**: 1913-1915.
- Melamed-Madjar, V., Cohen, S., Chen, M., Tam, S. and Rasillioi, D. (1979). Observations on populations of *Bemisia tabaci* Gennadius (Homoptera: Aleyrodidae) on cotton adjacent to sunflower and potato in Israel. *Israel Journal of Entomology*. **13**: 71-78.

- Nadeem, A., Weng, Z., Nelson, M.R. and Xiong, Z. (1997). Cotton leaf crumple virus and cotton leaf curl virus are two distantly related geminiviruses. *Molecular Plant Pathology On-line*. <http://www.bspp.org.uk/mppol/1997/0612nadeem>
- Narula, A.M., Monga, D., Chauhan, M.S. and Sheo Raj. (1999). Cotton leaf curl virus disease in India: The challenge ahead. *J. Cotton Res. Dev.* **13**(2):129-138.
- Nateshan, M.M., Muniappa, V., Siwanson, M.M. and Harrison, B.D. (1996). Host range, vector relations and serological relationships of cotton leaf curl virus from Southern India. *Ann. Appl. Biol.* **128**:233-244.
- Navot, N., Pichersky, E., Zeidan, M., Zamir, D. and Czonsnek, H. (1991). Tomato yellow leaf curl virus : A whitefly transmitted geminivirus with a single genomic component. *Virology* **185**: 151-161.
- Neal, D.C. (1946). A possible mosaic disease of cotton observed in Louisiana in 1946. *Phytopathology*, **37**: 434-435.
- Nelson, M.R., Nadeem, A., Ahmed, W. and Orum, T.V. (1998). Global assessment of cotton virus diseases. In: *Proceedings of Beltwide cotton conference*, Jan 5-9, San Diego, CA, USA. p. 161-162.
- Nene, Y.L. (1973). Virus diseases of some warm weather pulse crops in India. *Plant Dis. Repr.* **57** : 463-467.
- Noris, E., Hidalgo, E., Accotto, G.P. and Moriones, E. (1994). High similarity among the tomato yellow leaf curl virus isolates from the West Mediterranean basin – the nucleotide sequence of an infectious clone from Spain. *Arch. Virol.* **135**: 165-170.
- Noris, E., Vaira, A.M., Caciagli, P., Masenga, V., Gronenberg, B. and Accotto, G.P. (1998). Amino acids in the capsid protein of tomato yellow leaf curl virus that are crucial for systemic infection, particle formation, and insect transmission. *J. Virol.* **72** : 10050-57.
- Noueiry, A.O., Lucas, W.J. and Gilbertson, R.L. 1994. Two proteins of a plant DNA virus coordinate nuclear and plasmodesmatal transport. *Cell* **76** : 925-932.
- Nour, M and Nour, J.J. (1964). Identification, transmission and host range of cotton leaf curl virus infection of cotton in Sudan. *Empirical Cotton Growers Review*. **36**:32-34.
- Nour, M.A. (1959). Cotton leaf mottle: a new virus disease of cotton. *Empire Cotton Growers Review*. **36**: 32-34.
- Nour, M.A. (1960). On leaf curl of cotton in the Philippines. *FAO Plant Protection Bulletin*, **8**: 55-56.
- Orozco, B.M., Gladfelter, H.J., Settlege, S.B., Eagle, P.A., Gentry, R. and Hanley-Bowdoin, L. (1998). Multiple cis elements contribute to geminivirus origin function. *Virology* **242** : 346-356.

- Padidam, M., Beachy, R.N. and Fauquet, M. (1995). Classification and identification of geminivirus using sequence comparisons. *J. Gen. Virol.* **76** : 249-263.
- Padidam, M., Sawyer, S. and Fanquet, C.M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology* **265** : 218-25.
- Pappu, S.S., Chatterjee, A., Pappu, H.R. and Varma, A. (1997). Characterization of cotton leaf curl geminivirus from India: nucleotide sequence of DNA-B. *Phytopathology*. **87**: 575.
- Pascal, E., Good Lone, P.E., Wu, L.C. and Lazarowitz, S.G. (1993). Transgenic tobacco expressing the geminivirus BL1 protein exhibits 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 BL1 exhibit symptoms of viral disease. *Plant Cell* **5** : 795-807.
- Pimpale, T.D. and Summanwar, A.S. (1984). Studies on different stages in the life cycle and influence of season on the duration of different generations of the whitefly *Bemisia tabaci* Genn. *Pestology* **8**: 15-19.
- Pimpale, T.R. and Summanwar, A.S. (1986). Some observation on biology, behaviour and bionomics of whiteflies (*Bemisia tabaci* Genn.) – the vector of some plant viruses. In: *National Seminar on Whitefly Transmitted Plant Virus Diseases*, IARI, New Delhi, 18pp.
- Pringle, C.R. (1999). Virus taxonomy at the XIth International Congress of Virology, Sydney, Australia, 1999. *Arch. Virol.* **144**: 2065-2070.
- Pringle, C.R. (1997). Virus taxonomy 1997. *Arch. Virol.* **142**: 1727-1733.
- Quant, G.L. (1977). Virosis del algodonero en Guatemala. *El Algodonero*, **28**: 4-7.
- Radhakrishnan, G., Malathi, V.G. and Varma, A. (2001). Novel features of cotton leaf curl disease in India. In: *3rd International Geminivirus Symposium*, July 24-28, 2001, John Innes Centre, Norwich, Norfolk, UK, p.53.
- Rashid, A. and Khan, M.A. (2000). Evaluation of antagonistic microbes and bio-insecticides against leaf crul virus and bacterial blight of cotton. *Pakistan J. Phytopathology*. **12**(2): 137-141.
- Revington, G.N., Sunter, G. and Bisaro, D.M. (1989). DNA sequences essential for replication of the B genome component of tomato golden mosaic virus. *Plant cell* **1** : 985-992.
- Reyes, G.M., Martinez, A.L. and Chinte, P.T. (1959). Three virus diseases of plants new to the Philippines. *FAO Plant Protection Bulletin*, **7**: 141-143.
- Rochester, D.E., Kositratana, W. and Beachy, R.N. (1990). Systemic movement and symptom production following agroinoculation with a single DNA of tomato yellow leaf curl geminivirus (Thailand). *Virology* **178**: 520-26.

- Rogers, S.G. and Weiss, B. (1980). Exonuclease III of *Escherichia coli* K-12 an AP endonuclease. *Methods Enzymol.*, **65**: 201-211.
- Roossnick, M.J., Sleat, D. and Palukaitis, P. (1992). Satellite RNAs of plant viruses: Structures and biological effects. *Microbiol. Rev.*, **56**:265-279.
- Rybicki, E.B., Briddon, R.W., Brown, J.K., Fauquet, C.M., Maxwell, D.P., Harrison, B.D., Markham, P.G., Bisaro, D.M., Robinson, D., and Stanley, J. (2000). *Geminiviridae*. In: "*Virus Taxonomy. Seventh Report of the international Committee on Taxonomy of Viruses*" (M.H.V. Regenmortel, C.M. Fauquet, D.H.L. Bishop, E.B. Carstens, M.K. Estes, S.M. Lemon, J. Maniloff, M.A. Mayo, D.J. McGeoch, C.R. Pringle and R.M. Wickner, Eds.), pp.285-297. Academic Press, San Diego.
- Rybicki, E.P. (1994). A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Arch. Virol.* **139** : 49-77.
- Saikia, A.K. and Muniyappa, V. (1986). Epidemiology and control of tomato leaf curl virus. In: *National Seminar on Whitefly Transmitted Plant Virus Diseases*. June, 25-27, IARI, New Delhi, p.30-31.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning -A laboratory manual. Cold Spring Harbour Laboratory, New York, USA.
- 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** (8) : 1185-1194.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74** : 5463-5467.
- Sanz, A.I., Fraile, A., Garcia-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. *J. Gen. Virol.* **81**: 1839-1849.
- Saunders, K. and Stanley, J. (1995). Complementation of African cassava mosaic virus AC2 gene function in a mixed bipartite geminivirus infection. *J. Gen. Virol.*, **76**: 2287-2292.
- Saunders, K., and Stanley, J. (1999). A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides*: Evidence for interfamilial recombination between plant DNA viruses. *Virology*. **264**: 707-712.
- 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 aetiology. *Virology*. **282**: 38-47.
- Saunders, K., Bedford, I.D., Briddon, R.W., Markham, P.G., Wong, S.M. and Stanley, J. (2000). A novel virus complex causes *Ageratum* yellow vein disease. *Proc. Natl. Acad. Sci. USA*. **94**: 7088-7093.

- Schnippenkoetter, W.H., Martin, D.P., Hughes, F.L., Fyvie, M., Willment, J.A. James, D., von Wechmar, M.B. and Rybicki, E.P. (2001). The relative infectivities and genomic characterization of three distinct mastreviruses from South Africa. *Arch. Virol.* **146**: 1075-1088.
- Shafique, M., Sharif, M. and Farooq, V. (1994). Cotton leaf curl virus in the cotton and mixed zone of the irrigated Punjab. Agriculture Economics Research Institute (PARC), AART, Faisalabad, 1933.
- Singh, J., Gill, R.S., Brar, A.P.S., Sandhu, B.S. and Singh, J. (1996). Evaluation of some exotic cotton hybrids against insect pests. *Crop Improvement*, **23**(2): 275-277.
- Singh, J., Sohi, A.S., Mann, H.S. and Singh, J. (1997). Screening of cotton germplasm against cotton leaf curl viral disease using its vector *Bemisia tabaci* (Genn.). *J. of Res.*, **34**(3): 294-298.
- Singh, J.A.S., Sahi Mann, H.S. and Kapur, S.P. (1994). Studies on whitefly *Bemisia tabaci* (Genn) transmitted cotton leaf curl virus disease in Punjab. *J. Insect Sci.* **7**: 194-198.
- Sleeth, B., Lambe, R.C. and Hubbard, J.L. (1963). Terminal stunt of cotton in South Texas. *Plant Disease Reporter*. **47**: 587-589.
- Smith, K.M.S. (1982). *Plant viruses*. 6 Ed. Chaman and Hall, London.
- Stanley, J. (1983). Infectivity of the cloned geminivirus genome requires sequences from both DNA. *Nature* **305** : 643-645.
- Stanley, J. (1985). The molecular biology of geminiviruses. *Adv. Virus Res.* **30** : 139-177.
- 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. and Gay, M.R. (1983). Nucleotide sequence of cassava latent virus DNA. *Nature* **301** : 260-262.
- Stanley, J. and Townsend, R. (1985). Characterization of DNA forms associated with cassava latent virus infection. *Nucl. Acids Res.*, **13**: 2189-2206.
- Stanley, J., Markham, P.G., Callis, R.J. and Pinner, (1986). The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus. *EMBO J.*, **5**: 1761-1767.
- Stanley, J., Saunders, K., Pinner, M.P., and Wong, S.M. (1997). Novel defective interfering DNAs associated with ageratum yellow vein geminivirus infection of *Ageratum conyzoides*. *Virology*. **239**: 87-96.

- Stenger, D.C., Davies, H.R. and Bisaro, D.M. (1992). Limited replication of tomato golden mosaic virus DNA in explants of non host species. *Mol. Pl. Microbe Interaction*, **5**: 525-527.
- Stenger, D.C., Revington, G.N., Stevenson, M.C. and Bisaro, D.M. (1991). Replicational release of geminivirus genomes from tandemly repeated copies : Evidence for rolling – circle replication of a plant viral DNA. *Proc. Natl. Acad. Sci. USA*. **88** : 8029-8033.
- Subrahmanyam, K. (1979). Studies on yellow mosaic of blackgram. Ph.D. Thesis, IARI, New Delhi.
- Sung, Y.K. and Coutts, R.H.A. (1995). Pseudorecombination and complementation between potato yellow mosaic geminivirus and tomato golden mosaic geminivirus. *J. Gen. Virol.* **76** : 2809-2815.
- Sunter, G. and Bisaro, D.M. (1991). Transactivation in a geminivirus : AL2 gene product is needed for coat protein expression. *Virology* **180** : 416-419.
- Sunter, G. and Bisaro, D.M. (1992). Transactivation of geminivirus AR1 and geminivirus BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. *Plant Cell*. **4**: 1321-1331.
- Sunter, G., Harititz, 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.
- Tan, P.H.N., Wong, S.M., Wu, M., Bedford, I.D., Saunders, K., and Stanley, J. (1995). Genome organization of ageratum yellow vein virus, a monopartite whitefly-transmitted geminivirus isolated from a common weed. *J. Gen. Virol.* **76**: 2915-2922.
- Tarr, S.A.J. (1951). Leaf curl disease of cotton. Commonwealth Mycological Institute, Kew.
- Tarr, S.A.J. (1957). The host range of cotton leaf curl virus in Sudan. *Empire Cotton Growers Review*, **34**: 258-262.
- Tarr, S.A.J. (1964). Virus Diseases of Cotton. Miscellaneous Publication No.18. Commonwealth Mycological Institute, Kew, Surrey, England, 23pp.
- Thottappilly, G. (1992). Plant virus diseases of importance to African agriculture. *J. Phytopath.* **134**: 265-288.
- Timchenko, T., Katul, L., Sano, Y., de Kouchkovsky, F., Vetten, H.J., and Gronenborn, B. (2000). The master Rep concept in nanovirus replication: Identification of missing genome components and potential for natural genetic reassortment. *Virology*. **274**: 189-195.
- 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. *Nucl. Acids. Res.* **14**: 1253-1265.
- Tripathi, S. (2000). Characterization and management of tomato leaf curl geminivirus. Ph.D. thesis, IARI, New Delhi, p.87.
- Varma, A. (1976). Recent trends in control of plant virus disease. *Proc. Natl. Acad. Sci. India.* **46**: 193-206.
- Varma, A. (1983). Control of plant virus vectors an integral part of pest management. In: *Principles and Concepts of Integrated Pest Management* (Eds. R.A. Agrawal, G.P. Gupta, Prem Kishore and Dinesh Chandra). Division of Entomology, IARI, New Delhi, 48-53 p.
- Varma, A. (1986). Ecology of whitefly transmitted virus. In: *National Seminar on Whitefly Transmitted Plant Virus Diseases*, IARI, New Delhi, pp. 41.
- Varma, A. (1993). Integrated management of plant virus diseases. In: *Crop Protection and Sustainable Agriculture*. Wiley Chichester. (Ciba Foundation Symposium 177), pp. 140-157.
- Varma, A. (1994). *Leaf curl disease of cotton in North West India*. Report of the ICAR Committee. September, 1994.
- Varma, A. and Srivastava, K.P. (1996). Advances in the management of insect vectors of viral diseases of plants. In: *Recent advances in Indian Entomology*, (Ed. O.P. Lal) APC Publications, New Delhi. p 43-62.
- Varma, A. and Subrahmanyam, K. (1986). Epidemiology of mungbean yellow mosaic virus in *Vigna mungo*. In: *National Seminar on Whitefly Transmitted Plant Virus Diseases*, IARI, New Delhi, p.46..
- Varma, A. (1995). *Leaf curl disease of cotton in North West India*. Report of the ICAR committee. September, 1995.
- Varma, A., Dhar, A.K. and Mandal, B. (1992). MYMV transmission and control in India. In: *Mungbean Yellow Mosaic Disease* (Eds., Green, S.K. and Kim, D.), Asian Vegetable Research and Development Centre, Taipei, p. 8-27.
- Varma, A., Malathi, V.G. and Tripathi, S. (1997). Whitefly transmitted viral diseases of plants in India. In: *Plant Pathology: Fifty years of Research in India*. (Eds. Varma, J.P. and Varma, A). Indian Council of Agricultural Research, New Delhi p. 1-27.
- Varma, A and Malathi, V.G. (2003). Emerging geminivirus problems: A serious threat to crop production. *Ann. Appl. Biol.* **142**: 145-164.
- Varma, A., Malathi, V.G., Handa, A., Ailon, M., Harrison, B.D., Verma, J.P., Singh, R.P., Singh, M., Srivastava, M. and Singh, J. (1993). Occurrence of leaf curl of cotton and okra in Northern India. In: *6th International Congress of Plant Pathology*, July 28-August 6, Montreal, Canada, 17.5.14.

- Varma, Anupam (1997). Application of biotechnology for plant pest management: current status and future prospects. In: *Proceedings of Regional Expert Consultation on Application of Biotechnology in Plant Pest Management*. February 25-28. New Delhi, India, p 21-66.
- Varma, P.M. (1963). Transmission of plant viruses by whiteflies. *Bull. Natl. Inst. Sci., India*, **24**:11-33.
- Vieira, J. and Messing, J. (1982). The pUC plasmids and M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primer. *Gene*, **19**: 259-268.
- Von Arnim, A., Frischmuth, T. and Stanley, J. (1993). Detection and possible functions of African Cassava mosaic virus DNA-B gene products. *Virology* **192** : 264-272.
- Wege, C., Gotthardt, R.D., Frischmuth, T. and Jeske, H. (2000). Fulfilling Koch's Postulates for Abutilon mosaic virus. *Arch. Virol.* **145**: 2217-2225.
- Yasin, A.M. and El-Nur, E. (1970). Transmission of cotton leaf curl virus by a single insect of *Bemisia tabaci*. *Pl. Dis. Report.* **54**: 528-531.
- Youngman, R.R., Toscano, N.C., Jones, V.P.K. and Natwick, E.T. (1986). Correlations of seasonal trap counts of *Bemisia tabaci* (Homoptera: Aleyrodidae) in south eastern California. *Econ. Ento.* **79**: 67-70.
- 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 Oka. *J. Gen. Virol.* **79**: 915-923.
- Zhou, X.P., Xie, Y., Zhang, Z.K., Qi, Y.J. and Wu, J.J. (2001). Molecular characterization of a novel defective DNA isolated from tobacco tissues infected with tobacco leaf curl virus. *Acta Virol.* **45**(1): 45-50.

APPENDIX 1

MATERIALS AND METHODS

Quantitative Estimation of DNA

- 5 µl plasmid or plant DNA was electrophoresed in 1% agarose gel along with 1 kb DNA ladder. Serial dilution of 200 ng, 500 ng, 750 ng and 1 µg of the marker in 20 µl of loading mix was electrophoresed along with the samples. Intensity of sample DNA was compared with the serial concentration of molecular weight marker's bands and an approximate estimate was obtained.
- 5 µl of plasmid/plant DNA was diluted to 500 µl with 1 X TE. Absorbance was read at 260 nm using 1 X TE as reference in a UV spectrophotometer having 1 cm path length. Concentration of the DNA in the sample was calculated using the formula $50 \mu\text{g of DNA/ml} = 1 \text{ OD}$.

Agarose Gel Electrophoresis

Viral replicative DNA forms or total nucleic acid from plants or restricted plasmid DNA were mixed with 4 µl of 6 X loading dye and sterilized water to make up the volume to 20 µl and were electrophoresed in 1 per cent agarose gel prepared in 1 X tris-acetate-EDTA (TAE) buffer containing Et.Br. (0.5 µg/ml). Electrophoresis was carried out at 60V for 1-2h. An aliquot (500 ng) of 1kb DNA ladder was similarly mixed with the dye and electrophoresed to serve as molecular weight marker.

Southern Hybridization

Southern hybridization was carried out following the procedure given by Southern (1975) as described below.

Southern transfer

- After visualizing agarose gel on the UV transilluminator (UVP gel documentation system) it was trimmed to retain only relevant portion and gel size measured accurately.
- The left hand top corner of the gel was cut to fix the correct orientation of the gel.
- The DNA in the gel was alkali-denatured by soaking the gel in denaturation solution for 45 min. with gentle agitation.
- After denaturation, the gel was neutralized by soaking it in neutralization solution for 30 min. with gentle agitation.
- A piece of Whatman filter paper (3 mm) was soaked in 10 X SSC and kept over a glassplate which in turn was placed over a tray containing 10 X SSC (transfer buffer).
- The gel was then kept inverted down on the glass plate over the soaked filter paper.
- A piece of nitrocellulose membrane (NCM) was cut into same size of the gel and soaked in 10 X SSC. A cut mark same as the orientation of gel was given in the NCM and was placed over the gel. Enough care was taken to avoid air bubble.
- Over the NCM, 3 small pieces (same as the size of gel) of Whatman filter paper (3 mm) were kept and a bunch of ordinary blotting papers was kept over it which helps in capillary transfer of DNA from gel to membrane. A weight of ~1kg was kept over the blotting paper to facilitate efficient capillary action.
- The capillary transfer of DNA was allowed to proceed for 18 h.
- After transfer, the membrane was washed in 6 X SSC and air dried.
- The membrane was baked at 80⁰C for 24 h in vacuum.

Prehybridization

- The baked NCM blots were kept in hybridization cylinders.
- Prehybridization solution was added at the rate of 0.2 ml/sq cm.
- The cylinders were then incubated at 65⁰C for 4h in hybridization oven with gentle rotation.

Probe production by random primer labelling method

- DNA fragments obtained either after restriction or PCR amplification were separated in low melting point (LMP) agarose gel (0.7 per cent) electrophoresis at 40V in 1 X TAE buffer.
- The LMP-agarose block containing the DNA fragment was cut from the gel and kept in boiling water bath for 5 min. 20 µl of gel mix was taken into a sterile eppendorf tube and 1µl of random hexanucleotide primer was added and boiled for 1 min. at 100⁰C.
- Reaction mixture for probe was prepared by mixing the following reagents using Random Primer Labelling Kit (Bangalore Genei Pvt. Ltd.) and radiolabelling was done with (α -³²P) dCTP by random priming method (Feinberg and Vogelstein, 1984).

Reaction mixture:

Denatured DNA (~30 ng)	20.0 µl
Primers (random hexa nucleotide 100 µg/µl)	1.0 µl
10 X labelling buffer	5.0 µl
dNTPs (10mM) (excluding dCTP)	2.0 µl
Dithiothreitol (DTT) (20 mM)	2.5 µl
Nuclease free BSA	2.0 µl
(α - ³² P) dCTP (4000 Ci m/mol)	1.0 µl
Klenow (DNA Pol I) (3 U/µl)	1.0 µl
Sterile double distilled water	<u>15.5 µl</u>
	<u>50.0 µl</u>

- The reaction mixture was then incubated at 37°C for 2h.

Hybridization

- The double stranded ³²P-labelled DNA probe was denatured in a boiling water bath for 5 min.
- The denatured probe was added to prehybridization solution at 0.5×10^6 cpm/ml concentration.
- It was allowed to hybridize for 18h at 65°C in hybridization oven with gentle rotation.

Authoradiography

- The hybridization solution was discarded and the membrane was washed 3 times with 2 X SSC, 0.1 per cent SDS at 65°C, each time with duration of 15 min.
- The washed membrane was dried on a paper towel, then placed within folds of cling film, placed in a lead cassette and exposed to X-ray film (INDU) for 18h at -70°C.
- Autoradiograph were then developed as per X-ray films manufacture's instructions.

Cloning of Viral Genome

The PCR amplified full-length DNA- α and novel DNA molecules were cloned as blunt end molecules in *Sma*I cut dephosphorylated pUC18/19 vector (MBI, Fermentas) (Fig.A1). PCR amplified DNA- α fragments and full-length DNA- β molecules were cloned in pGEM-T easy vector (Promega Inc., USA) (Fig.A2).

Blunt ending T₄ DNA polymerase strategy

T₄ DNA polymerase (MBI, Fermentas) removes the 3' and 5' overhang of dsDNA which can then be directly ligated with any blunt end

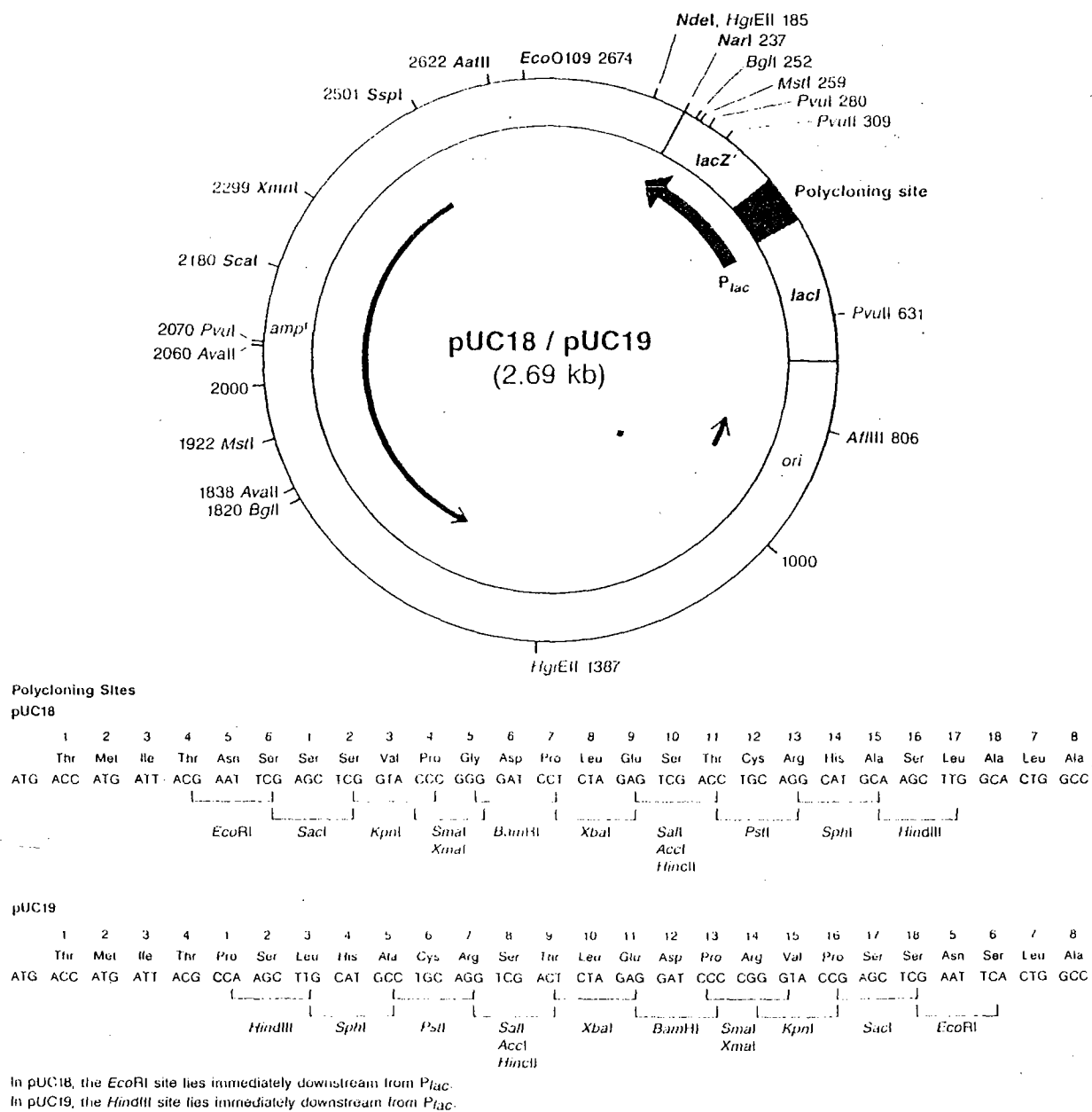


Figure A.1. The genome map of the vector pUC18/pUC19

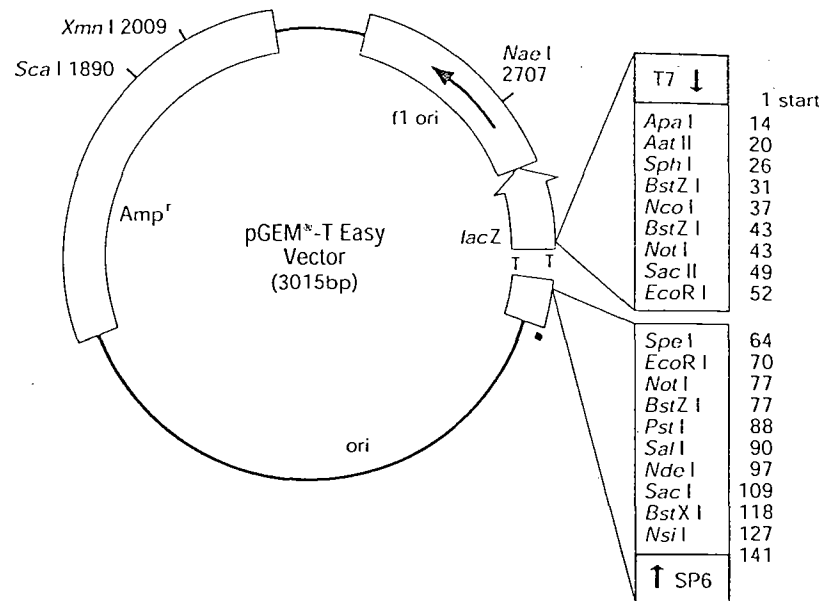


Figure A.2. The genome map of the vector pGEM-T easy

vector. This strategy was used to clone viral RF DNA assuming that there may be considerable linear forms. The reaction conditions are as given below:

Viral DNA (~200ng)	10 μ l
5 X reaction buffer	4 μ l
2 mM dNTP	1 μ l
T ₄ DNA Pol. (10U/ μ l)	1 μ l
Sterile double distilled water	<u>4 μl</u>
	<u>20 μl</u>

After mixing, a pulse spin was given in table top microfuge, and the reaction mixture was incubated at 11⁰C for 20 min. To inactivate the enzyme the reaction mixture was incubated at 75⁰C for 10 min. The reaction mixture was then ready for ligation.

Ligation

Both the T₄ DNA Pol. treated PCR amplified full-length DNA- α and novel DNA molecules and PCR amplified DNA- α fragments and full-length DNA- β molecules were separately ligated with *Sma*I cut dephosphorylated pUC18/19 vector and PGEM-T easy vector, respectively. The ligation mixture was as follows:

Viral DNA	20 μ l
Linearised dephosphorylated <i>Sma</i> I cut pUC18/19 vector (150 ng) or pGEM-T easy vector (50 ng)	1 μ l
10 X ligase buffer	3 μ l
10 mM ATP	3 μ l
T ₄ DNA ligase (2 U/ μ l)	1 μ l
Sterile double distilled water	<u>2 μl</u>
	<u>30 μl</u>

The ligation mixtures were incubated at 15⁰C for 18h. Both the ligation mixtures were then ready for transformation.

Preparation of competent cells

The competent cells were prepared by CaCl_2 method as described by Mendel and Higa (1970).

- 50 ml Luria Broth (LB) was inoculated with overnight grown culture of DH5 α strain of *Escherichia coli* (strategene) and incubated at 37°C for 1h 15 min. with constant shaking at 200 rpm in a shaker incubator till the bacterial growth as measured by optical density reached 0.3 OD at 600 nm.
- The culture was then aseptically transferred to 40 ml sterile screw capped tubes and kept on ice for 10 min.
- The cells were centrifuged at 5000 rpm for 10 min. at 4°C in a Sorvall SS-34 rotor to pelletize the cells.
- The cells were resuspended gently in 10 ml ice cold 0.1 M MgCl_2 solution and centrifuged at 5000 rpm for 10 min. at 4°C.
- The pellet was resuspended in 10 ml ice cold 0.1 M CaCl_2 solution and kept on ice for 1h.
- The cells were recovered by centrifuging at 5000 rpm for 10 min. at 4°C and the pellets were resuspended in 1 ml of chilled 0.1 M CaCl_2 and used for transformation after keeping on ice for 1h.

Transformation of competent cells

- 200 μl competent cells were added to 20 μl of the ligation mixture in a sterile microfuge tube, gently mixed and kept on ice for 1h.
- The bacterial cell-DNA mixture was given a heat shock at 42°C for 2 min. 1 ml of LB medium was then added and the transformants were allowed to grow at 37°C for 1h in shaker incubator at 200 rpm.
- 200 μl of cell suspension were aseptically plated either as such or after concentration on Luria Agar (LA) plate containing ampicillin, X-gal and

IPTG (100 μ l of 1 per cent ampicillin, 200 μ l of 2 per cent X-gal, and 20 μ l of 0.1 M IPTG for 100 ml LA).

- The plates were incubated overnight at 37°C.

Selection of transformants

The transformants were selected on the basis of blue/white colony colour. The white colonies were selected and subsequently plated on IXA (IPTG-X-gal-Ampicillin) plates. This plate having individual transformants in grid served as master plate.

Rapid Screening for the Recombinant Clones having Insert

White recombinant bacterial colonies may be formed either due to insertion of viral genome or due to even a small deletion in the *lac Z* operon of vector itself. To know which white colonies have the insert a rapid screening method was used which was as follows:

- A small amount of overnight grown bacterial colonies were picked from master grid individually with the help of sterile tooth pick and mixed with 50 μ l of 10 mM EDTA (pH 8.0) in sterile microfuge tubes. One or two blue colonies of pUC18 transformed bacterial cells were also taken for control.
- 50 μ l of fresh lysis solution (2 N NaOH, 0.5 per cent SDS, 20 per cent sucrose) was added in each tube and vortexing was done for 30 seconds which caused rapid disruption of bacterial cell wall.
- The mixture was then incubated at 70°C for 5 min. and was cooled down to room temperature.
- 1.5 μ l of 4 M KCl and 0.5 μ l of 0.4 per cent of bromophenol blue were added to each tube and vortexing was done for 30 seconds.
- The mixture was then incubated for 5 min. on ice and centrifugation was done at 10,000 rpm for 30 min. at 4°C in table top centrifuge.

- Bacterial cell debris was removed and 30 µl of supernatant from each tube along with the control were electrophoresed in 0.7 per cent agarose gel.
- The lanes which show higher plasmid band than the control was considered as recombinant plasmid and the respective colonies were taken for isolation of recombinant plasmid DNA.

Isolation of Recombinant Plasmid DNA by Miniprep Method

Isolation of recombinant plasmid DNA was done by modified alkaline lysis method (Brinboim and Doly, 1979).

- Selected white colonies which were presumed to contain viral insert in rapid screening were individually inoculated in 2 ml of LB medium containing ampicillin (50 µg/ml) in sterile capped culture tubes.
- Tubes were then incubated overnight at 37⁰C at 200 rpm in a shaker incubator.
- The overnight grown bacterial cells were then transferred to 1.5 ml sterile eppendorf tube and cells were harvested by centrifuging in a table top centrifuge for 1 min. Care was taken to remove the medium adhering to the cell pellet.
- The pellet was resuspended in 100 µl of solution I and mixed vigorously by vortexing.
- Then 200 µl of freshly prepared lysis solution i.e. solution II was added and mixed gently.
- 150 µl of ice cold solution III was then added and mixed gently with lysed cell suspension and the mixture was kept on ice for 15 min.
- The chromosomal DNA and the bacterial cell debris were removed by centrifuging at 15,000 rpm for 15 min. at 4⁰C in a table top centrifuge.
- The supernatant was again centrifuged for another 15 min. at 15,000 rpm at 4⁰C to pelletize any unwanted bacterial debris.

- The supernatant was taken and equal volume of phenol : chloroform : isoamyl alcohol (25 : 24 : 1) was added. It was vortexed well, centrifuged in a table top centrifuge for 15 min. at room temperature.
- The clear aqueous phase was transferred to fresh eppendorf tube.
- The DNA in aqueous phase was precipitated by adding 0.8 volume of isopropanol, and kept on ice for 10 min.
- This mixture was then centrifuged at 15,000 rpm for 20 min. at 4°C.
- To the pellet 200 µl of 70% ethanol was added. The tube was rotated well so that the pellet from the wall gets suspended in 70% alcohol. This ensures removal of adhering salts by 70% alcohol. DNA was then pelletized by centrifuging at 15,000 rpm for 5 min.
- The pellet was finally suspended in 30 µl sterile double distilled water.

Large Scale Plasmid DNA Preparation

Large scale plasmid DNA preparation of selected DNA-A and DNA-B clones were done following QIAGEN Midi plasmid purification protocol (QIAGEN Plasmid Midi Kit). The procedure is described below:

- A single colony from a freshly streaked plates of each clone was picked and was inoculated in 50 ml LB containing ampicillin (50 µg/ml) in a screw capped conical flask.
- The culture was grown for 18 h in shaker incubator at 37°C with 200 rpm rotation.
- The bacterial cells were then harvested by centrifugation at 5000 rpm for 10 min. at 4°C in RC-5C Sorvall centrifuge with SS-34 rotor.
- The pellet was resuspended in 4 ml of buffer P1 and vortexed.
- 4 ml of buffer P2 was added, mixed gently and incubated at room temperature for 5 min.

- 4 ml of chilled P3 buffer was then added to the lysed cell suspension and mixed immediately but gently and incubated on ice for 15 min.
- After incubation the bacterial chromosome and debris were removed by centrifugation at 15,000 rpm for 30 min. at 4°C.
- After removing the supernatant promptly, the same step was carried out once again.
- The supernatant was loaded onto QIAGEN-tip 100 [previously equilibrated with 4 ml of buffer QBT and allowed to enter the resin by gravity flow.
- QIAGEN-tip 100 was then washed twice with 10 ml of buffer QC.
- The recombinant plasmid DNA was eluted with 5 ml of buffer QF.
- DNA was then precipitated by adding 3.5 ml of isopropanol and centrifugation was done at 15,000 rpm at 4°C for 30 min.
- DNA pellet was washed with 2 ml of 70% alcohol and centrifugation was done at 15,000 rpm at 4°C for 10 min.
- After air-drying, the pellet was dissolved in 100 µl of sterile double distilled water.

After estimating the DNA concentration, appropriate aliquots were used for various experiments.

Restriction Analysis

- Selected recombinant plasmids were subjected to digestion with different restriction enzymes and incubated at 37°C for 4h. (except *Sma*I, which was incubated at 30°C for 4h). Restriction was done to know the size of the insert and to develop a physical map.

The reaction conditions were as follows:

Plasmid DNA (500-1000 ng)	1.0 μ l
10 X reaction buffer	2.5 μ l
Restriction enzyme (10 U/ μ l)	0.5 μ l
Sterile double distilled water	<u>21.0 μl</u>
	<u>25.0 μl</u>

After restriction the DNA samples were electrophoresed in 1 per cent agarose gel. Fragment sizes were assessed in comparison with 1 kb DNA markers.

Nucleic Acid Spot Hybridization Test (NASH Test)

- 1g leaf tissue of all the plants being tested were ground in liquid nitrogen.
- Total nucleic acid was extracted in 2 per cent CTAB buffer as given earlier. Total nucleic acid isolated from healthy and whitefly inoculated cowpea plants served as negative and positive control respectively.
- Appropriate size of NCM (Amersham, USA) was cut, 1 cm square grids were made with soft pencil and soaked in water, then in 10 X SSC for 5 min.
- Excess SSC was drained out and the membrane was kept on filter paper for dotting.
- The samples (DNA) prepared were heat denatured by keeping them in boiling water bath for 5 min. and then dotted on NCM. About 2.5 μ g of total DNA was dotted for all plants.
- After dotting, the membranes were dried and baked at 80⁰C for 2h in vacuum.
- Prehybridization and hybridization of these blots with radiolabelled (∞ ³²P) probe were done separately.

DNA Sequencing by Didioxy Chain Termination Method

DNA sequencing was carried out using *Taq* polymerase (Sequencing kit, MBI Fermentas) by didioxy chain termination method (Sanger *et al.*, 1977) following were the steps involved:

(1) Alkaline denaturation of plasmid DNA

- 8 μ l (3-4 μ g) of large-prep plasmid DNA was taken in a sterile eppendorf. 2 μ l of 2 M NaOH was added and was mixed gently by pipetting the contents up and down. The mixture was incubated in room temperature for 10 min.
- 3 μ l of 3 M sodium acetate solution (pH 4.8) and 7 μ l of deionised water was added and was mixed gently but thoroughly.
- 60 μ l of ethanol was added and mixed thoroughly and was incubated for 15 min. at -20°C .
- To harvest the single stranded DNA pellet, centrifugation was done at 4°C for 30 min. at 10,000 rpm.
- 200 μ l of cold 70 per cent ethanol was added to the pellet for washing and again centrifugation was done at 15,000 rpm for 30 min. at 4°C .
- Pellet was dried under vacuum and was dissolved in 7 μ l of deionised water.

(2) Primer annealing

- Following components were mixed in a sterile microfuge tube.

Denatured template DNA	7 μ l
Primer (1 pmol/ μ l)	1 μ l
10 X sequencing buffer	2 μ l
Deionised water	4 μ l
	14 μ l

- The mixture was then incubated at 70°C for 3 min. and then 45°C for 10 min.

For sequencing two primers were used, M13 forward primer and M13 reverse primer.

(3) Labelling reaction

- Following reagents were added to the eppendorf tube which contain 14 µl of annealed primer/template mixture:

Labelling mix	3 µl
(α - ³⁵ S) dATP (10 µCi)	1 µl
Reader TM <i>Taq</i> DNA Pol.	1 µl
Deionised water up to	20 µl

- The mixture was then incubated at 45°C for 30 sec.

(4) Termination reaction

- Four termination mixtures were prepared by dispensing 4 µl each of A, C, G, T in 4 separately labelled eppendorf.
- 4 µl of labelling mixture was then taken to each of the four marked (A, C, G, T) termination mixture aliquoted before and mixed gently.
- The mixtures were incubated at 70°C for 5 min.
- 4 µl of STOP solution was added in each tube, centrifuged for few seconds for proper mixing.

(5) Denaturing polyacrylamide gel electrophoresis

Polyacrylamide gels were prepared by copolymerisation of acrylamide and a cross-linker, N'N'-methylene-bis-acrylamide. Common ratio of these two is 19 : 1. To keep the DNA denatured sequencing gels were prepared with a high concentration of urea (7-8 M). 6% DNA sequencing gels were 0.2 mm thick. Components of the sequencing gel used in the present investigation were as follows:

Acrylamide	5.7 g
Bis-acrylamide	0.3 g
Urea	50.0 g
10 X TBE	10 ml
Sterile double distilled water	40 ml

After mixing all the contents the mixture was filtered and volume was made upto 100 ml with double distilled water and 25 μ l of TEMED and 500 μ l freshly prepared of 10 per cent ammonium persulphate solution were added to it. After that the gel mix was carefully poured between sequencing and casting glass plates. Before pouring the glass plate on which gel would bind was cleaned with alcohol and treated with bind silane many times while the double glass plate which helps in maintaining temperature was cleaned with alcohol and treated with repel silane many times. After pouring the gel these two glass plates were clamped and was allowed to set for 2h. After setting the gel apparatus in sequencing system (Pharmacia LKB macrophore system) comb was fixed very carefully. Samples were boiled for 5 min. and kept on ice. 2.5 μ l of samples were then loaded in each well in A, C, G, T sequence. The gel was allowed to run at 2000 V, 55⁰C for 2h (in the case of short run) or 5h (in the case of long run). In the case of short run, the run was stopped when the blue dye (bromophenol blue) touches the bottom of the gel while in the case of long run, the run was stopped after 45 min. of release of green dye (xylene cyanol) from the bottom of the gel. Sequence generated were confirmed by repeating it minimum three times for a single reaction.

After cooling, the glass plate which contained gel was removed and immersed in 10 per cent glacial acetic acid solution for 15 min. to fix the gel properly. The gel (along with glass plate) was then gently washed in tap water, dried in oven at 80⁰C for overnight. After that Konica X-ray film was kept on it for developing autoradiogram.

Agroinoculation (triparental mating)

The tandem dimeric constructs of CLCuV DNA-A and DNA- β in the binary vector pBin19 (Fig.A3) were mobilized from *E. coli* DH5 α to *Agrobacterium tumefaciens* strain LBA4404 using pRK2013 as helper plasmid in a triparental mating system (Ditta *et al.*, 1980). *Agrobacterium tumefaciens*

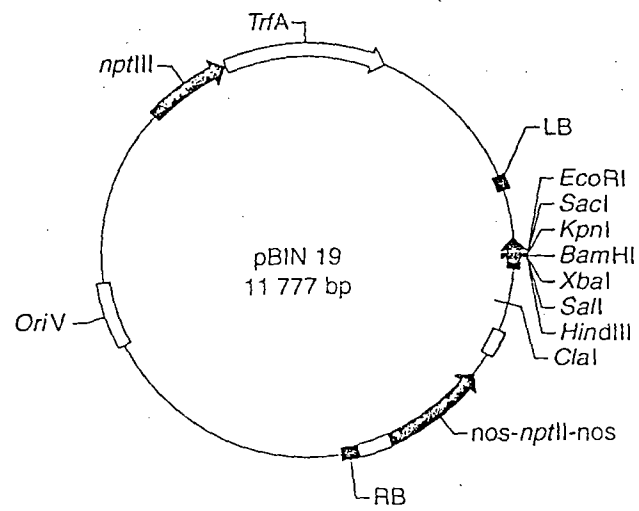


Figure A.3. The genome map of the vector pBin19

LBA4404 was grown in LB medium containing kanamycin (50 µg/ml) at 28°C for 24h with continuous shaking at 200 rpm in shaker incubator. The *E. coli* DH5α cells containing dimeric constructs of CLCuV DNA-A and DNA-β and *E. coli* JM 83 cells containing helper plasmid pRK2013 were also grown separately in LB medium containing kanamycin (50 µg/ml) for 24h at 37°C with continuous shaking at 200 rpm. Two ml of *A. tumefaciens* LBA4404, one ml of helper *E. coli* JM 83 and one ml of *E. coli* DH5α cells (containing dimeric constructs of CLCuV DNA-A and DNA-β) were transferred into a culture tube and mixed well. An aliquot of 50-100 µl of this mixture was layered on 1 sq. cm. sterile NCM placed on LA plates. The plates were air dried and sealed with parafilm and incubated at 28°C for 24h. The bacterial cells which grew on NCM were diluted in 5 ml of sterile water and mixed with gentle shaking. These were further diluted ten times in sterile water and from that 10 µl suspension spread on LA plates, containing kanamycin (50 µg/ml) and then the plates allowed for incubation for 36h at 28°C. The bacterial colonies from each construct were inoculated in 5 ml of LB containing kanamycin (50 µg/ml) and rifampicin (50µg/ml) and incubated at 28°C for 24h. For agroinoculation of plants, both the constructs i.e. DNA-A and DNA-β were inoculated separately as well as in mixture of equal proportion.

APPENDIX 2

COMMON REAGENTS, BUFFERS AND MEDIA USED

I. ANTIBIOTICS

Ampicillin Stock solution (50 mg/ml) of the antibiotic was made in double distilled water, filter sterilized (through 0.22 micron filter) and distributed into 200 μ l aliquots and stored at -20°C . It was used at a concentration of 50 $\mu\text{g/ml}$.

II. ELECTROPHORESIS REAGENTS

50X TAE	Tris base	242.0 g
	Glacial acetic acid	57.1 ml
	0.5 M EDTA (pH 8.0)	100 ml
	Distilled water to 1 Litre	
Loading dye	1% Bromophenol blue	200 μ l
	Glycerol	200 μ l
	10% SDS	60 μ l
	0.5 M EDTA	50 μ l
	10X TAE	60 μ l
	Distilled water	30 μ l

III. MEDIA

Luria Agar medium	Bacto-tryptone	10.0 g
	Bacto-yeast extract	5.0 g
	NaCl	10.0 g
	Agar	15.0 g
	Deionized water	950 ml

pH was adjusted to 7.0 with 5N NaOH and volume made upto 1 litre with deionized water. It was dispensed in 100 ml aliquots in 250 ml flasks and was sterilized by autoclaving for 20 min. at 15 p.s.i.

Luria Broth medium	Tryptone	10.0 g
	Yeast extract	5.0 g
	NaCl	5.0 g
	Deionized water	950 ml

pH was adjusted to 7.0 with 5N NaOH and volume made upto 1 litre with deionized water. The medium was aliquoted into 50 ml in 250 ml flasks and sterilized by autoclaving for 20 min at 15 p.s.i.

IV. PLASMID ISOLATION BUFFERS

Solution I (resuspension buffer)	25 mM Tris HCL (pH 8.0)
	50 mM Glucose
	10 mM EDTA
Solution II (lysis buffer)	0.2 N NaOH
	1% SDS
Solution III (neutralization buffer)	3 M Sodium acetate pH 4.8

V. SOUTHERN HYBRIDIZATION SOLUTIONS

Denaturing solution	1.5 M NaCl
	0.5 M NaOH
Neutralizing solution	1.5 M NaCl
	1.0 M Tris. HCl (pH 7.4)
Denhardt's reagent (50X)	1g Polyvinylpyrrolidone sp. (Sigma)
	1g Ficoll (Type 400, Sigma)
	1 g Bovine serum albumin (Fraction V. Sigma)
	100 ml Double distilled water

This stock solution was stored at -20°C . It was used at working concentration of 5X Denhardt's reagent.

VI. PREHYBRIDIZATION SOLUTION

5X Denhardt's reagent

6X SSC

0.3% SDS

$100\ \mu\text{g ml}^{-1}$ yeast RNA

VII. OTHER BUFFERS

DNA Extraction buffer	0.1 M Tris. HCL, pH 8.0
	1.5 M NaCl
	0.01 M EDTA
	2% CTAB (Cetyl trimethyl ammonium bromide)

Membrane washing buffer	0.1% SDS
	2X SSC

10X CIAP	10 mM ZnCl_2
Dephosphorylation buffer	10 mM MgCl_2
	100 mM Tris. HCl (pH 8.3)

10X Ligation buffer	0.5 M Tris. HCl (pH 7.6)
	0.5 M MgCl_2
	0.1 M Dithiothreitol
	$500\ \mu\text{g ml}^{-1}$ Bovine serum albumin
	(Fraction V, Sigma Chemical Co., A 8022)

VIII. DNA MOLECULAR WEIGHT MARKER

1 kilobase (1kb) DNA ladder of MBI Fermentas was used as marker. The ladder is formed by fourteen DNA fragments of 10 kb, 8 kb, 6 kb, 5 kb, 4 kb, 3.5 kb, 2.5 kb, 2 kb, 1.5 kb, 1 kb, 0.75 kb, 0.5 kb and 0.25 kb.

**IX. COMPOSITION OF BUFFERS USED IN QIAGEN PLASMID
MIDI KIT**

Buffer	Composition
Buffer P1 (Resuspension Buffer)	50 mM Tris. Cl, pH 8.0; 10 mM EDTA; 100 µg/ml RNase A
Buffer P2 (Lysis Buffer)	200 mM NaOH, 1% SDS
Buffer P3 (Neutralization Buffer)	3.0 M potassium acetate, pH 5.5
Buffer QBT (Equilibration Buffer)	750 mM NaCl; 50 mM MOPS, pH 7.0; 15% isopropanol; 0.15% Triton® X-100
Buffer QC (Wash Buffer)	1.0 M NaCl; 50 mM MOPS, pH 7.0; 15% isopropanol
Buffer QF (Elution Buffer)	1.25 M NaCl; 50 mM Tris. Cl, pH 8.5; 15% isopropanol

X. PREPARATION OF COMMONLY USED STOCK SOLUTION

Solution	Method of preparation
0.1 M Adenosine triphosphate (ATP)	60.0 mg of ATP was dissolved in 0.8 ml of distilled water. The pH was adjusted to 7.0 with 0.1 N NaOH and volume made upto 1 ml with distilled water. The solution was dispensed into small aliquots and stored at -70°C.
1M CaCl₂	54.0 g of CaCl ₂ .2H ₂ O was dissolved in 200 ml of pure water (Mill-Q or equivalent). The solution was sterilized by passing through a 0.22 micron filter and stored in 1 ml aliquots at 4°C.
0.5 M EDTA (pH 8.0)	186.1 g of ethylenediamine tetra acetic acid disodium salt.2H ₂ O was added to 800 ml of distilled water, stirred vigorously on a magnetic stirrer, pH was adjusted to 8.0 with NaOH (20.0 g of NaOH pellets). Volume made upto 1 L with distilled water, dispensed into aliquots and sterilized by autoclaving.
Ethidium bromide (10 mg ml⁻¹)	1.0 g of ethidium bromide was added to 100 ml of distilled water and stirred on a magnetic stirrer for several hours to ensure that the dye has dissolved. The solution was transferred to a dark bottle and stored at room temperature.

Phenol : Chloroform : isoamyl alcohol

Buffer saturated phenol, chloroform and isoamyl alcohol were mixed in the ratio of 25 : 24 : 1. The equilibrated mixture was stored under a layer of 0.01 M Tris-HCl (pH 7.6) at 4°C in dark glass bottle.

Chloroform : isoamyl alcohol

Buffer saturated chloroform and isoamyl alcohol were mixed in the ratio of 24 : 1.

IPTG (Isopropyl-β-D-thiogalactopyranoside)

A solution of IPTG was made by dissolving 2.0 g of IPTG in 8 ml of distilled water. Volume was made upto 10 ml with distilled water and sterilized by filtration through a 0.22 micron disposable filter. The solution was dispensed into 1 ml of aliquots and stored at -20°C.

1 M MgCl₂

203.3 g of MgCl₂.6H₂O was dissolved in 800 ml of distilled water. The volume was made upto 1L, dispensed into aliquots and sterilized by autoclaving.

3M Sodium acetate (pH 4.8)

408.1 g of NaOAc.3H₂O was dissolved in 800 ml of distilled water. The pH was adjusted to 4.8 with glacial acetic acid. Volume made upto 1L with distilled water, dispensed into aliquots and sterilized by autoclaving.

5 M NaCl

233.8 g of NaCl was dissolved in 800 ml of distilled water, volume made upto 1L with

distilled water, dispensed into aliquots and sterilized by autoclaving.

10% Sodium dodecyl sulphate (SDS)

100.0 g of electrophoresis grade SDS was dissolved in 900 ml of distilled water, heated at 68°C to assist dissolution and pH was adjusted to 7.2 by adding few drops of concentrated HCl. The volume was made up to 1L with distilled water, dispensed in aliquots.

10(N) NaOH

Dissolve 400 g of NaOH in 800 ml of distilled water and make up the volume to 1L with distilled water.

20X SSC

175.3 g of NaCl and 88.2 g of sodium citrate were dissolved in 800 ml of distilled water. The pH was adjusted to 7.4 with 10N solution of NaOH. Volume was adjusted to 1L with distilled water, dispensed into aliquots and sterilized by autoclaving.

1M Tris. HCl

121.1 g of Tris. base was dissolved in 800 ml of distilled water. pH was adjusted to the desired value by adding concentrated HCl (for pH 7.4, HCl 70 ml; for pH 8.0, HCl 42 ml). The solution was allowed to cool down to room temperature before making final adjustment to the pH. The volume was made upto 1L with distilled water, dispensed into aliquots and sterilized by autoclaving.

X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside)

The stock solution was made by dissolving X-gal in dimethyl formamide to make a 20 mg/ml⁻¹ solution and stored at -20°C.

**XI. MATERIALS, SOURCE, CATALOGUE NUMBER OF
CHEMICALS AND REAGENTS**

Material	Source	Catalogue No.
Vector		
pUC18 and pUC19 (<i>Sma</i> I cut, dephosphorylated) and pGEM-T easy vector	Stratagene, USA Promega, USA	212205
Cells		
<i>Escherichia coli</i> DH 5 α	Stratagene, USA	200233
Nucleic acid marker		
1 kb DNA ladder	MBI Fermentas	SM 0311
Enzymes		
Restriction endonucleases	Bangalore Genei Pvt. Ltd., MBI Fermentas	MBE 3, GA 2151
Nuclease		
RNase A	Sigma Chemical Cio.	R5000
Other enzymes		
Calf intestinal alkaline phosphatase (CIAP)	Stratagene, USA	600015
T ₄ DNA ligase	Amersham, USA	E2050Y
Radioisotope		
(α - ³² P) dCTP Sp. activity 4000 Ci m/Mol	Board of Radiation and Isotope Technology (BRIT), Bombay	LCP 102
(α - ³⁵ S) dATP (10 μ Ci)	Board of Radiation and Isotope Technology (BRIT), Bombay	LCS -35

Membrane

Nitrocellulose membrane (NCM)	Amersham, USA	RPN 2020E
-------------------------------	---------------	-----------

Other reagents

5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal)	USB	12385
--	-----	-------

Isopropyl- β -D-thiogalactopyranoside (IPTG)	USB	17886
--	-----	-------

PCR Reagents and <i>Taq</i> and <i>Pfu</i> DNA Pol.	MBI, Fermentas	
---	----------------	--

Antibiotics

Ampicillin (Sodium salt)	Sigma Chemical Co.	A 9518
--------------------------	--------------------	--------

Kits

Random primer labelling kit	Bangalore Genei Pvt. Ltd.	KT-04
-----------------------------	---------------------------	-------

Plasmid midi/mini kit, DNeasy Plant Mini kit and DNeasy Tissue Kit	QIAGEN, USA	1012507
--	-------------	---------

Reader TM DNA sequencing kit	MBI, Fermentas	K#1811
---	----------------	--------

APPENDIX 3

GenBank data sheet of CLCuV sequences reported in the present study



PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM

Search **Nucleotide** for

Limits Index History Clipboard
 Display Default View as **HTML** Save Add to Clipboard

☐ 1: AF363011 **Cotton leaf curl virus segment A, complete sequence** Protein, Related Sequences, Taxonomy

LOCUS AF363011 2754 bp DNA circular VRL 20-JUN-2001
 DEFINITION Cotton leaf curl virus segment A, complete sequence.
 ACCESSION AF363011 AF329671 AF337533 AF337534 AF339845 AF363010
 VERSION AF363011.1 GI:14495285
 KEYWORDS .
 SOURCE cotton leaf curl virus.
 ORGANISM cotton leaf curl virus
 Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
 REFERENCE 1 (bases 1511 to 2599)
 AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.
 TITLE Nucleotide sequence of replication initiation protein of Indian isolate of cotton leaf curl virus (CLCuV-G)
 JOURNAL Unpublished
 REFERENCE 2 (bases 292 to 1062)
 AUTHORS Radhakrishnan,G., Malathi,V.G., Usharani,K.S. and Varma,A.
 TITLE Nucleotide sequence of cotton leaf curl virus (CLCuV-G) coat protein gene from India
 JOURNAL Unpublished
 REFERENCE 3 (bases 2143 to 2445)
 AUTHORS Radhakrishnan,G., Malathi,V.G., Usharani,K.S. and Varma,A.
 TITLE Nucleotide sequence of cotton leaf curl virus (CLCuV-G) AC4 gene from India
 JOURNAL Unpublished
 REFERENCE 4 (bases 1162 to 1614)
 AUTHORS Radhakrishnan,G., Malathi,V.G., Usharani,K.S. and Varma,A.
 TITLE Nucleotide sequence of cotton leaf curl virus (CLCuV-G) AC2 gene from India
 JOURNAL Unpublished
 REFERENCE 5 (bases 1065 to 1669)
 AUTHORS Radhakrishnan,G., Malathi,V.G., Usharani,K.S. and Varma,A.
 TITLE Nucleotide sequence of cotton leaf curl virus (CLCuV-G) AC3 gene from India
 JOURNAL Unpublished
 REFERENCE 6 (bases 1 to 2754)
 AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.
 TITLE Complete Nucleotide Sequence of DNA A of Cotton Leaf Curl Virus (CLCuV-G) from India
 JOURNAL Unpublished
 REFERENCE 7 (bases 1 to 2754)
 AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-2001) Advanced Center For Plant Virology, Division of Plant Pathology, Indian Agricultural Research Institute, Pusa Road, New Delhi, Delhi 110012, India
 COMMENT On Jun 20, 2001 this sequence version replaced gi:13625990
 gi:12239618 gi:13123718 gi:13123720 gi:13123722.
 FEATURES Location/Qualifiers
 source 1..2754
 /organism="cotton leaf curl virus"
 /isolate="CLCuV-G"
 /db_xref="taxon:53010"
 /country="India: Rajasthan, Sriganaganagar"

```

gene      /translation="MMLPSKPLKSNGTCMSSYVYWTIPSYLISDGDVDSILM"
          complement(424..585)
          /gene="beta C3"
CDS       complement(424..585)
          /gene="beta C3"
          /codon_start=1
          /product="beta C3 protein"
          /protein_id="AAL89478.1"
          /db_xref="GI:19568133"
          /translation="MHTKQFHQISRNDNERNKQGGSQVHSRRSHHGEYEDLHSHEDTI
          HKVTIANQV"
gene      472..621
          /gene="beta V3"
CDS       472..621
          /gene="beta V3"
          /codon_start=1
          /product="beta V3 protein"
          /protein_id="AAL89473.1"
          /db_xref="GI:19568128"
          /translation="MKIFIFSMRTSTMNLTPSLFVPLVVISAYLMELFSVHAFIDLM
          NRYFV"
gene      complement(558..707)
          /gene="beta C2"
CDS       complement(558..707)
          /gene="beta C2"
          /codon_start=1
          /product="beta C2 protein"
          /protein_id="AAL89477.1"
          /db_xref="GI:19568132"
          /translation="MTTRVIRLHRHFIPMDNHIIINHSQPHNYTKYLFIKSINACTL
          NNSIK"
gene      773..904
          /gene="beta V4"
CDS       773..904
          /gene="beta V4"
          /codon_start=1
          /product="beta V4 protein"
          /protein_id="AAL89474.1"
          /db_xref="GI:19568129"
          /translation="MRKGKTELKRKNKEKKQGYIIYERNGSAANRNKKTQGRKNKK"
gene      complement(835..1020)
          /gene="beta C1"
CDS       complement(835..1020)
          /gene="beta C1"
          /codon_start=1
          /product="beta C1 protein"
          /protein_id="AAL89476.1"
          /db_xref="GI:19568131"
          /translation="MVKPHSHLQTTTFSDMFLFFLFFYYSQMTFSNLYFFSLYFLFF
          LFPWVFLFRFAALPFLS"
gene      837..1034
          /gene="beta V5"
CDS       837..1034
          /gene="beta V5"
          /codon_start=1
          /product="beta V5 protein"
          /protein_id="AAL89475.1"
          /db_xref="GI:19568130"
          /translation="MKEMGAQRIETGKPKEREKIKSKEKKNINSKTSFESNKKKEKT
          KTYPKTLSFEGDCVVLFFTVM"

```

```

BASE COUNT      395 a      225 c      315 g      415 t
ORIGIN

```

```

1 accgtgggagc agcgggtgcc gatggcttct tgggtgggtcc cattgctggt attgacttga
61 tttgacttat attgggccaa tttaatgggt tgaaaatggt tgggcctttg gaagagggtc
121 ttttatggaa tagggctctg tatattgttg gtgatatgtg tgttaaatat gcattgctgg
181 tttgtgtttg gaatttaaag ggtgaacttt ttattgaata agtacgggtc gtttacatcc
241 attcccaata tctctgggtt ttcaagtaca agtatatcaa gtctgtgaac tatatcttct
301 atctcgatct cttctatctt tgccccgttg tatgcgaata ggaaattcgc tatgatgctc
361 ccttcaaaqc ctttqaagtc gaatggaacg tgtatgtctt cgtacgtgta ctggacgata

```

NCBI Sequence Viewer

/db_xref="GI:14495288"
 /translation="MGLLTCMFSSSSKGISSARICDSSTWSPQAGQHISIRTYRELN
 APTSNPTPRRTGTFSTGGNFRSTEGQQEEGNRQPMTLTPQHLTREVSRRLLLESLRN"

BASE COUNT	734 a	556 c	630 g	834 t			
ORIGIN	1	accggatggc	cgcgcgattt	ttttgggtggg	ccttaccatt	aacacttgtc	ggccaatcat
	61	atgactccct	caaagctaaa	taacgctccc	gcacactata	agtacttgcg	cactaagttt
	121	caaattcaaa	catgtgggat	ccactattaa	acgaattccc	tgatacgggt	cacgggtttc
	181	gggtgatgct	ttctgtgaaa	tatttgcaac	ttttgtcgca	ggattattca	ccggatacgc
	241	ttgggtacga	gttaatacgg	gattttaattt	gtattttacg	ctcccgtagt	tatgtcgaag
	301	cgagctgccg	atatcgatcat	ttctacgccc	gcgtcgaaaag	tacgcccggc	tctgaacttc
	361	ggcagcccat	acaccagccg	tgctgctgcc	cccattgtcc	gcgtcacaaa	acaacaggca
	421	tggacaaaca	ggcctatgaa	caggaagccc	aggatgtaca	ggatgtacag	aagtccagat
	481	gttcctagag	gatgtgaagg	tccatgtaag	gttcagtcgt	ttgagtcag	acatgatatt
	541	cagcatatag	gtaaagtaat	gtgtgttagt	gatgttactc	gtggtagctg	gctgacccat
	601	agagttggta	agagattttg	tgtcaagtct	gtttatgtgt	tgggtaagat	atggatggat
	661	gagaacatta	agacgaagaa	tcacacgaat	agtgtgatgt	ttttcttggt	tagagatcgt
	721	agacctgttg	ataaacctca	agatttttga	gaggtattta	atatgtttga	taatgagccc
	781	agtacggcga	ctgtgaagaa	tgttcatcgt	gataggtatc	aagttctgcg	caaatggat
	841	gcaactgtca	ccggtggaca	atacgtttca	aaggaacagg	ctttgggtcaa	gaagtttgct
	901	agagttaaca	attatgttgt	ttacaatcaa	caggaagcag	gaaaatacga	gaatcatacg
	961	gaaaatgcgt	taatgcttta	tatggcttgt	actcacgcta	gcaaccctgt	ttatgctacg
	1021	ttgaagatta	ggatatattt	ttatgactct	gtaacgaatt	gatattaata	aagtttaaat
	1081	tttatttctg	aatattgatc	tacatacata	gtttgttggg	ttacattgta	caatacatgt
	1141	tctacagctt	taataactaa	attaattgaa	attacaccga	gattgttcag	atatttgagg
	1201	acttggtttt	tgaataccct	taagaaaaga	ccagtctgag	gctgtaaggt	cgtccagatt
	1261	cggaaggtta	gaaaacactt	gtgcagtcct	agagctttcc	gcgtgttgta	gttgaactgg
	1321	atcctgatcg	tgagtatgtc	catattcgct	gtgaatggac	ggttgacgtg	gctgatgatc
	1381	ttgaaataaa	ggggatttgg	aacctcccag	atatatgcgc	cattccctgc	ttgagctgca
	1441	gtgatgggtt	cccctgtgcg	tgaatccatg	gttggtggcag	ttgattgaca	gataataaga
	1501	acaccgcgat	tcaagatcta	ctctcctcct	cctgttgccg	ctcttcgctt	ccctgtgctg
	1561	tactttgatt	ggtacctgag	tacatgggtc	tatcaagtgt	gatgaagatc	gcattcttta
	1621	aagcccaatt	ttttagtgtc	gaattcttct	cttcatccaa	aaactcttta	tagctagaat
	1681	tgggtccctg	attgcagagg	aagatgggtg	gaattccgcc	tttaatttga	actggcttcc
	1741	cgtattttgt	atttgattgc	cagtcctttt	gggcccccat	gaactcctta	aagtgcctta
	1801	ggtaatgtgg	gttgacgtca	tcaatgacgt	taaaccaggc	gtcattactg	tatacccttg
	1861	ggctcagatc	tagatgtcca	cacagataat	tatgtggacc	taatgatctg	gccacatcgc
	1921	tcttccccgt	cctactgtca	ccctcaatca	ctacacttat	tggtctattg	gcccgcgcag
	1981	cggcactgac	gacgttctcg	gcagcccaca	cttcaagttc	ttctggaact	tgatcgaaaag
	2041	aagaagagga	aaaaggagaa	acataaggag	ctgggtggctc	ctgaaagatt	ctgtctagat
	2101	ttgcatttaa	attatgaaat	tgcagtacaa	aatccttagg	agctagttcc	ttaatgactc
	2161	taagagcctc	cgacttactt	cccgcgttaa	gtgctgcggc	gtaagcgtca	ttggctgtct
	2221	gttgccctcc	tcttgctgac	cttccgtcga	tctgaaattg	ccccagtcg	agaatgtccc
	2281	cgtccttctc	ggtgtaggat	ttgacgtcgg	agctggattt	agctccctgt	atgttcggat
	2341	ggaaatgtgc	tgacctgctt	ggggagacca	agtcgaagaa	tcgcatattc	tggcacttga
	2401	aattcccttc	gaactggatg	agaacatgca	agtgaggagt	cccatcttcg	tgaagctctc
	2461	tgagatttct	aatatatttt	tttgaagttg	gggtttgtat	atttaataat	tgggaaagtg
	2521	cttccctctt	ggtgagagaa	catttgggat	aagtgatgaa	atagtttttg	gaataaatac
	2581	cgttccgctt	tggaggcatg	ttgactaaaa	ttgatcaccg	attgaccgct	cttgcaactc
	2641	tccccggtat	atcgggtgatc	aatatatagt	gatcaccaaa	tggcataatg	gtaataaaaa
	2701	aactttaatt	tgaattcaa	accaaaggcc	taaagcggcc	atccgtttaa	tatt



PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

Search

Nucleotide

for

Limits

Preview/Index

History

Clipboard

De

Cubby: not logged in

Display

default

Save

Reset

Go to Clipboard

☐ 1: AY083590. Cotton leaf curl ...[gi:19568125]

Protein, Taxonomy

LOCUS AY083590 1350 bp DNA linear VRL 20-MAR-2002
 DEFINITION Cotton leaf curl virus CLCuV-G segment DNA beta, complete sequence.
 ACCESSION AY083590
 VERSION AY083590.1 GI:19568125
 KEYWORDS .
 SOURCE cotton leaf curl virus.
 ORGANISM cotton leaf curl virus
 Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
 REFERENCE 1 (bases 1 to 1350)
 AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.
 TITLE Complete nucleotide sequence of DNA beta of cotton leaf curl virus (CLCuV-G) from Sri Ganganagar, India
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1350)
 AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAR-2002) Advanced Centre for Plant Virology, Division of Plant Pathology, Indian Agricultural Research Institute, Pusa Road, New Delhi, Delhi 110012, India
 FEATURES Location/Qualifiers
 source 1..1350
 /organism="cotton leaf curl virus"
 /isolate="CLCuV-G"
 /db_xref="taxon:53010"
 /note="segment: DNA beta"
 gene 96..221
 /gene="beta V1"
 CDS 96..221
 /gene="beta V1"
 /codon_start=1
 /product="beta V1 protein"
 /protein_id="AAL89471.1"
 /db_xref="GI:19568126"
 /translation="MFGPLEEGLLWNRALYIVGDMCVKYALLVCVWNLNGELFIE"
 gene complement(195..551)
 /gene="beta C4"
 CDS complement(195..551)
 /gene="beta C4"
 /codon_start=1
 /product="beta C4 protein"
 /protein_id="AAL89472.1"
 /db_xref="GI:19568134"
 /translation="MTTSGTNKEGVRFIVDVRIMENMKIFIHMRILSTKSPSLIKYEGIVQYTYEDIHVPFDENGFEGSIIANFLFAYNGAKIEEIEIEDIVHRLDILVLENPEILGMDVNEPYLFNKKFTV"
 gene 352..471
 /gene="beta V2"
 CDS 352..471
 /gene="beta V2"
 /codon_start=1
 /product="beta V2 protein"
 /protein_id="AAL89473.1"
 /db_xref="GI:19568137"

gene /translation="MMLPSKPLKSNMTCSSYVYWTIPSYLISDGDVDSILM"
complement(424..585)
/gene="beta C3"
CDS complement(424..585)
/gene="beta C3"
/codon_start=1
/product="beta C3 protein"
/protein_id="AAL89478.1"
/db_xref="GI:19568133"
/translation="MHTKQFHQISRNDNERNKQGGSQVHSRRSHHGEYEDLHSHEDTI
HKVTIANQV"
gene 472..621
/gene="beta V3"
CDS 472..621
/gene="beta V3"
/codon_start=1
/product="beta V3 protein"
/protein_id="AAL89473.1"
/db_xref="GI:19568128"
/translation="MKIFIFSMRTSTMNLTPSLFVPLVVISAYLMELFSVHAFIDLM
NRYFV"
gene complement(558..707)
/gene="beta C2"
CDS complement(558..707)
/gene="beta C2"
/codon_start=1
/product="beta C2 protein"
/protein_id="AAL89477.1"
/db_xref="GI:19568132"
/translation="MTTRVIRLHRHFIPLMDNHIIINHSQPHNYTKYLFIKSINACTL
NNSIK"
gene 773..904
/gene="beta V4"
CDS 773..904
/gene="beta V4"
/codon_start=1
/product="beta V4 protein"
/protein_id="AAL89474.1"
/db_xref="GI:19568129"
/translation="MRKGKTELKRKNKEKKQGYIIYERNGSAANRNRKTQGKRKNKK"
gene complement(835..1020)
/gene="beta C1"
CDS complement(835..1020)
/gene="beta C1"
/codon_start=1
/product="beta C1 protein"
/protein_id="AAL89476.1"
/db_xref="GI:19568131"
/translation="MVKPHSHLQTTTFSDMFLFFLFFYYYSQTMTFSNLYFFSLYFLFF
LFPWVFLFRFAALPFLS"
gene 837..1034
/gene="beta V5"
CDS 837..1034
/gene="beta V5"
/codon_start=1
/product="beta V5 protein"
/protein_id="AAL89475.1"
/db_xref="GI:19568130"
/translation="MKEMGAQRIETGKPKEREKIKSKEKKNINSKTSSFESNKKKEKT
KTYPKTILSFEGDCVLPFTVW"

BASE COUNT 395 a 225 c 315 g 415 t
ORIGIN

```
1 accgtgggag agcgggtgcc gatggcttct tgggtgggtcc cattgctggt attgacttga
61 tttgacttat attgggccaa tttaatgggt tgaaaatggt tgggcctttg gaagagggtc
121 ttttatggaa tagggctctg tatattgttg gtgatatgtg tgttaaatat gcattgctgg
181 tttgtgtttg gaatttaaac ggtgaacttt ttattgaata agtacgggtc gtttacatcc
241 attcccaata tctctgggtt ttcaagtaca agtatatcaa gtctgtgaac tatatcttct
301 atctcgatct cttctatctt tgccccgttg tatgcgaata ggaaattcgc tatgatgctc
361 ctttcaagag ctttcaagag ctttcaagag ttttcaagag ctttcaagag ctttcaagag
```

```

421 ccttcatact tgattagcga tgggtgacttt gtggatagta tcctcatgtg aatgaagatc
481 ttcataattct ccatgatgcg aacgtcgact atgaacctga ctccctcctt gtttggtccg
541 ctcgttgtca tttctgctta tttgatggaa ttgtttagtg tgcattgcatt tatagactta
601 atgaacagat attttgtgta gttgtgtggt tgtgagtgat ttattattat gtgattgtcc
661 attaaagggg taaagtgcg atggagacgt attacacgtg ttgtcatggt ggcttgaaat
721 ctttatacat gggtttgtac cgtatatacc tatatatacg gatagaaaac ggatgagaaa
781 aggaaaaacg gaactgaaga ggaaaaacaa agaaaagaaa caaggatata ttatttatga
841 aagaaatggg agcgcagcga atcgaaacag gaaaacccaa ggaaagagaa aaaataaaaa
901 gtaaaagagaa aaaaaatata aattcgaaaa cgtcatcggt tgagagtaat aaaaaaaaaag
961 aaaaaacaaa aacatatccg aaaacgttgt cgtttgaagg tgactgtgtg gttttaccat
1021 ttactgtgtg gtaaattgta agtgattgaa gaataattaa aaaatttgct accaataggt
1081 aaattgatca ccgatatac ggtgatcaat tgggtgcctca taattggtgt tgcctaaaat
1141 acccctattt ctgtgtcttg aaggcgcgtg gaagtgcgct gaaaaaggtg ttcttctctc
1201 acctaaaacc caccggaacg ccaaaactgg ctgttgccgg catcaattta cgacacgcgc
1261 ggcgggtgtg acccctggga gggtaggtac cactacgcta cgcagcagcc ttagctacgc
1321 cggagcttag ctgcccacg ctttaatat

```

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH



PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

Search for

Limits

Preview/Index

History

Clipboard

Cubby: not logged in

 ☐ 1: AF465618. Cotton leaf curl ...[gi:18449102]

Related Sequences, Protein, Taxonomy

LOCUS AF465618 771 bp DNA linear VRL 31-JAN-200:
DEFINITION Cotton leaf curl virus CLCuV-D coat protein (AV1) gene, complete
cds.

ACCESSION AF465618

VERSION AF465618.1 GI:18449102

KEYWORDS .

SOURCE cotton leaf curl virus.

ORGANISM cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 771)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-D)
coat protein gene from Delhi, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 771)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (06-JAN-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source

1..771

/organism="cotton leaf curl virus"

/isolate="CLCuV-D"

/db_xref="taxon:53010"

/country="India: Delhi"

gene

1..771

/gene="AV1"

CDS

1..771

/gene="AV1"

/codon_start=1

/product="coat protein"

/protein_id="AAL69989.1"

/db_xref="GI:18449103"

/translation="MSKRAADIVISTPASKVRRRLNFGSPYTSRAAIPVRVTQQQI
TNRPMNRKPRMYRSPDVPRGCEGPCKVQSFESRHDIQHLGKVMCVSDVTRGTGI
HRVGKRFCKSVYVLGKIWMDENIKTKNHTNSVMFFLVRDRRPVDPKQDFGEVFNMI
NEPSTATVKNVHRDRYQVLRKQYATVTGGQYASKEQALVKKFVRVNNYVVYNQQEA
YENHTENALMLYMACTHASNPVYATLKIRIYFYDSVTN"

BASE COUNT 223 a 137 c 195 g 216 t

ORIGIN

```
1 atgtcgaagc gagctgcoga tatcgtcatt tctacgccc cgctcgaagt acgccggcgt
61 ctgaacttcg gcagcccata caccagccgt gctgctgcc ccattgtccg cgtcacaaaa
121 caacaggcat ggacaaacag gcctatgaac aggaagccca ggatgtacag gatgtacaga
181 agtccagatg ttcttagagg atgtgaaggt ccatgtaagg ttcagtcgtt tgagtcacaga
241 catgatattc agcatttagg taaagtaatg tgtgttagtg atgttactcg tggtagctggg
301 ctgacccata gagttggtta gagattttgt gtcaagtctg tttatgtgtt gggaagata
361 tggatggatg agaacattaa gacgaagaat cacacgaata gtgtgatgtt tttcttggtt
421 agagatcgta gacctgttga taaacctcaa gattttggag aggtatttaa tatgtttgat
481 aatgagccca gtacggcgac tgtgaagaat gttcatcgtg ataggtatca agttctgcgt
541 aaatggatg caactgtcac cgggtggaca taccgttcaa aggaacaggc tttggtcaag
601 aagtttgtca gagttaacaa ttatgttggt tacaatcaac aggaagcagg aaaatacag
661 aatcatacgg aaaatgcgtt aatgctttat atggcttgta ctacgcctag caaccctgtt
721 tatgctacgt tgaagattag gatataattt tatgactctg taacgaattg a
```

//



Entries Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

Search Nucleotide for

Limits

Preview/Index

History

Clipboard

Detail

Cubby: not logged in

Display default Save Text Add to Clipboard

1: AF465620. Cotton leaf curl ...[gi:18449106]

Related Sequences, Protein, Taxonomy

LOCUS AF465620 771 bp DNA linear VRL 31-JAN-2002
DEFINITION Cotton leaf curl virus CLCuV-H coat protein (AV1) gene, complete cds.

ACCESSION AF465620
VERSION AF465620.1 GI:18449106

KEYWORDS

SOURCE cotton leaf curl virus.

ORGANISM cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 771)

AUTHORS Radhakrishnan, G., Malathi, V.G. and Varma, A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-H) coat protein gene from Hisar, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 771)

AUTHORS Radhakrishnan, G., Malathi, V.G. and Varma, A.

TITLE Direct Submission

JOURNAL Submitted (06-JAN-2002) Advanced Centre for Plant Virology, Division of Plant Pathology, Indian Agricultural Research Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source

1..771

/organism="cotton leaf curl virus"

/isolate="CLCuV-H"

/db_xref="taxon:53010"

/country="India: Hisar"

gene

1..771

/gene="AV1"

CDS

1..771

/gene="AV1"

/codon_start=1

/product="coat protein"

/protein_id="AAL69991.1"

/db_xref="GI:18449107"

/translation="MSKRAADIVISTPASKVRRRLNFGSPYTSRAAAPIVRVTKQQAW
TNRPMNRKPRMYRMYRSPDVPRGCEGPKVQSFESRHDIQHIGKVMCVSDVTRGTGLT
LRVGKRFCKVSVYVLGKIWMDENIKTKNHTNSVMFFLVDRRPPVDKPDQDFGEVFNMF
NEPSTATVKNVHRDRYQVLRKQYATVTGGQYASKEQALVKKFVRVNNYVVYNQQEAGK
YENHTENALMLYMACTHASNPVYATLKIRIYFYDSVTN"

BASE COUNT 223 a 138 c 195 g 215 t

ORIGIN

```

1 atgtcgaagc gagctgccga tatcgtcatt tctacgcccg cgtcgaaagt acgcggcgct
61 ctgaacttcg gcagcccata caccagccgt gctgctgcc ccattgtccg cgtcacaaaa
121 caacaggcat ggacaaacag gcctatgaac aggaagccca ggatgtacag gatgtacaga
181 agtccagatg ttcttagagg atgtgaaggt ccatgtaagg ttcagtctgt tgagtccaga
241 catgatattc agcatatagg taaagtaatg tgtgttagtg atgttactcg tgggtactggg
301 ctgaccctta gagtttgtaa gagattttgt gtcaagtctg tttatgtgtt gggtaagata
361 tggatggatg agaacattaa gacgaagaat cacacgaata gtgtgatgtt tttcttggtt
421 agagatcgta gacctgttga taaacctcaa gattttggag aggtatttaa tatgtttgat
481 aatgagccca gtacggcgac tgtgaagaat gttcatcggt ataggtatca agttctgocg
541 aaatggtatg caactgtcac cgggtggaca tacgcttcaa aggaacaggc tttggtcaag
601 aagtttgtca gagttaacaa ttatgttgtt tacaatcaac aggaagcagg aaaatacagag
661 aatcatacgg aaaatgcgtt aatgctttat atggcttgta ctacgcgtag caacctgtt
721 tatgctacgt tgaagattag gatataattt tatgactctg taacgaattg a

```

//



PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

Search for

Limits

Preview/Index

History

Clipboard

De

Cubby: not logged in

 ☐ 1: AF465619. Cotton leaf curl ...[gi:18449104]

Related Sequences, Protein, Taxonomy

LOCUS AF465619 771 bp DNA linear VRL 31-JAN-2002
 DEFINITION Cotton leaf curl virus CLCuV-L coat protein (AV1) gene, complete cds.

ACCESSION AF465619
 VERSION AF465619.1 GI:18449104

KEYWORDS

SOURCE cotton leaf curl virus.

ORGANISM cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 771)

AUTHORS Radhakrishnan, G., Malathi, V.G. and Varma, A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-L) coat protein gene from Ludhiana, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 771)

AUTHORS Radhakrishnan, G., Malathi, V.G. and Varma, A.

TITLE Direct Submission

JOURNAL Submitted (06-JAN-2002) Advanced Centre for Plant Virology, Division of Plant Pathology, Indian Agricultural Research Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES

source

Location/Qualifiers

1..771

/organism="cotton leaf curl virus"

/isolate="CLCuV-L"

/db_xref="taxon:53010"

/country="India: Ludhiana"

gene

1..771

/gene="AV1"

CDS

1..771

/gene="AV1"

/codon_start=1

/product="coat protein"

/protein_id="AAL69990.1"

/db_xref="GI:18449105"

/translation="MSKRAADIVISTPASKVRRRLNFGSPYTSRAAAPIVRVTKQQAW
 TNRPMMNRKPRMYRMYRSPDVPRGCEGPCKVQS FESRHDVVHIGKVMCISDVTRGVGLT
 HRIGKRFCKSVYVLGKIWM DENIKTKNHTNSVMFFLVRDRRPTDKPQDFGEVFNMF
 NEPSTATVKNMHRDRYQVLRKWHATVTGGQYASKEQALVKKFIRVNNYVVYNQOEAGK
 YENHTENALMLYMACTHASNPVYATLKIRIIFYDSVTN"

BASE COUNT 215 a 155 c 195 g 206 t

ORIGIN

```

1 atgtcgaagc gagctgccga tatcgtcatt tctacgcccg cgtcgaaagt acgccggcgt
61 ctgaacttcg gcagccctta caccagccgt gctgctgcc ccattgtccg cgtcacaaaa
121 caacaggcat ggacaaacag gcctatgaac aggaaaccca gaatgtaccg gatgtacaga
181 agtccggatg ttccaagggg atgtgagggt ccctgtaagg tacagtcgtt tgaatctcga
241 cagcatgttg tccatattgg taaggtaatg tgtatttctg atgttacgcg tggagtcggt
301 ttgacctatc gtataggtaa gcgtttttgt gtcaagtcag tttatgtttt aggtaagata
361 tggatggacg agaacatcaa gaccaagaac catacgaatt cggatgatgtt tttccttggt
421 cgtgatcgac gaccgacaga taaaccacaa gattttggtg aagtatttaa tatgtttgat
481 aacgagccca gtacggccac cgtaaagaac atgcataggg atcggtacca ggtgttgagg
541 aaatggcatg caaccgttac tgggtgtcaa tatgcatcga aggaacaggc gttggttaag
601 aagtttatca gggttaacaa ttatgttgtt tacaaccagc aggaggcagg aaaatacagag
661 aatcacaccg agaatgcatt gatgctttat atggcttgta ccgatgctag taaccagtt
721 tatgctacgc ttaagattcg gatatatatt tatgactctg taacgaattg a

```

//



PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM

Search for

Limits

Preview/Index

History

Clipboard

De

☐ 1: AF465621. Cotton leaf curl ...[gi:18449108]

Related Sequences, Protein, Taxonomy

LOCUS AF465621 771 bp DNA linear VRL 31-JAN-2002
DEFINITION Cotton leaf curl virus CLCuV-S coat protein (AV1) gene, complete cds.

ACCESSION AF465621
VERSION AF465621.1 GI:18449108

KEYWORDS

SOURCE cotton leaf curl virus.

ORGANISM cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 771)

AUTHORS Radhakrishnan, G., Malathi, V.G. and Varma, A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-S) coat protein gene from Sirsa, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 771)

AUTHORS Radhakrishnan, G., Malathi, V.G. and Varma, A.

TITLE Direct Submission

JOURNAL Submitted (06-JAN-2002) Advanced Centre for Plant Virology, Division of Plant Pathology, Indian Agricultural Research Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source

1..771

/organism="cotton leaf curl virus"

/isolate="CLCuV-S"

/db_xref="taxon:53010"

/country="India: Sirsa"

gene

1..771

/gene="AV1"

CDS

1..771

/gene="AV1"

/codon_start=1

/product="coat protein"

/protein_id="AAL69992.1"

/db_xref="GI:18449109"

/translation="MSKRAADIVISTPASKVRRRLNFGSPYTTRAAAPIVRVTKQQA
TNKPMNRKPRMYRSPDVPNGCGGPKVQSFSRHDVVHIGKVMCISDVTRGVGL
HRIGKRFVKSVYVLGKIWMDENIKTKNHTNSVMFFLVRDRRPVDKPDQFGEVFNMF
NEPSTATVKNSHRDRYQVLRKWHATVTGGQYASKEQALVKKFVRVNNYVVYNQQEAG
YENHTENALMLYMACTHASNPVYATLKIRIYFYDSVTN"

BASE COUNT 220 a 146 c 188 g 217 t

ORIGIN

```
1 atgtcgaagc gagctgccga tctcgtcatt tctacgccg cgctcgaagt acgccggcgt
61 ctgaacttcg gcagcccata caccaccgt gctgctgcc ccattgtccg cgtcacaaaa
121 caacaggcat ggacaacaaa gcctatgaac aggaagccca gaatgtatcg gatgtacaga
181 agtcgggatg ttccaaatgg ttgtggaggc ccatgtaagg tacaatcttt tgagtcgaga
241 catgatgtcg ttcatattgg taaggtaatg tgtatttctg atgttactcg tgggtgcggt
301 ttgacccatc gtatttgtaa acgtttttgt gtcaagtcag tttatgtttt aggtaagata
361 tggatggatg aaaatataaa gaccaagaat cacacgaatt cggtgatgtt ctttttagtc
421 cgcgatcgac gtctgttgta caaacctcag gattttgggt aggtattcaa tatgtttgac
481 aacgaacctg gtacagcaac tgtgaagaat agtcatagg accgttacca ggtgttgagg
541 aaatggcatg caaccgttac ggggtggtaa tatgctagta aggaacaagc tttggtcaag
601 aagtttgtca gagttaacaa ttatgttggt tacaatcaac aggaagcagg gaaatacgag
661 aatcatacgg aaaatgcgtt aatgctttat atggcttgta ctacacgtag caaccctggt
721 tatgctacgt tgaagattag gatataatgt tatgactctg taacgaattg a
```

LOCUS AY146960 357 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-D AV2 protein (AV2) gene, complete
cds.

ACCESSION AY146960

VERSION AY146960

KEYWORDS .

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 357)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-D)
AV2 protein gene from Delhi, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 357)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..357

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-D"

/country="India: Delhi"

gene 1..357

/gene="AV2"

CDS 1..357

/gene="AV2"

/note="movement/pre-coat protein"

/codon_start=1

/product="AV2 protein"

/protein_id="AAN61155"

/translation="MWDPLLNEFPD TVHGFR CMLSVKYLQLLSQDYSPDTLGYELIRD

LICILRSRNYVEASCRYRHFYARVESTPASELRQPIHQPCCPHCPRHKT TGMDKQAY
EQEAQDVQDVQKSRCS"

BASE COUNT 88 a 88 c 87 g 94 t

ORIGIN

1 atgtgggatac cactattaaa cgaattccct gatacgggtc acggggttcg gtgtatgctt
61 tctgtgaaat atttgcaact ttgtgcgag gattattcac cggatacgct tgggtacgag
121 ttaatacggg atttaatttg tattttgcgc tcccgttaatt atgtcgaagc gagctgccga
181 tatcgtcatt tctacgcccc cgtcgaaagt acgccggcgt ctgaacttcg gcagcccata
241 caccagccgt gctgctgccc ccattgtccg cgtcacaaaa caacaggcat ggacaaacac
301 gcctatgaac aggaagceca ggatgtacag gatgtacaga agtccagatg ttcttag

LOC JS AY146959 357 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-H AV2 protein (AV2) gene, complete
cds.

ACCESSION AY146959

VERSION AY146959

KEYWORDS .

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 357)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-H)
AV2 protein gene from Hisar, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 357)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..357

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-H"

/country="India: Hisar"

gene 1..357

/gene="AV2"

CDS 1..357

/gene="AV2"

/note="movement/pre-coat protein"

/codon_start=1

/product="AV2 protein"

/protein_id="AAN61154"

/translation="MWDPLLNEFPD TVHGFR CMLSVKYLQLLSQDYSPDTLGYELIRD

LICILRSRNYVEASCRYRHFYARVESTPASELRQPIHQPCCPHCPRHKTTGMDKQAY
EQEAQDVQDVQKSRCS"

BASE COUNT 89 a 88 c 86 g 94 t

ORIGIN

1 atgtgggatac cactattaaa cgaattccct gatacgggtc acggggttcg gtgtatgctt
61 tctgtgaaat atttgcaact ttgtcgcag gattattcac cggatacgct tgggtacgag
121 ttaatacggg atttaatttg tattttacgc tcccgttaatt atgtcgaagc gagctgccga
181 tatcgtcatt tctacgcccg cgtcgaaagt acgccggcgt ctgaacttcg gcagcccata
241 caccagccgt gctgctgccc ccattgtccg cgtcacaaaa caacaggcat ggacaaacac
301 gcctatgaac aggaagccca ggatgtacag gatgtacaga agtccagatg ttcttag

//

LOCUS AY146958 366 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-L AV2 protein (AV2) gene, complete
cds.

ACCESSION AY146958

VERSION AY146958

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 366)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-L)
AV2 protein gene from Ludhiana, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 366)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..366

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-L"

/country="India: Ludhiana"

gene 1..366

/gene="AV2"

CDS 1..366

/gene="AV2"

/note="movement/pre-coat protein"

/codon_start=1

/product="AV2 protein"

/protein_id="AAN61153"

/translation="MWDPLLNEFPD TVHGFR CMLSVKYLQLLSQDYSPDTLGYELIRD

LICILRSRNYVEASCRYRHFYARVESTPASELRQPIHQPCCPHCPRHKT TGMDKQAY
EQEAQNVSDVQKSGCSKGL"

BASE COUNT 91 a 88 c 90 g 97 t

ORIGIN

1 atgtgggatac cactattaaa cgaattccct gatacgggtc acgggtttcg gtgtatgctt
61 tctgtgaaat atttgcaact ttgtcgcag gattattcac cggatacgct tgggtacgag
121 ttaatacggg atttaatttg tattttacgc tcccgaatt atgtcgaagc gagctgccga
181 tategtcatt tctacgcccg cgtcgaaagt acgccggcgt ctgaacttcg gcagcccata
241 caccagccgt gctgctgcc ccatgttcg cgtcacaaaa ccaaggcat ggacaaacag
301 gcctatgaac aggaagccca gaatgtatcg gatgtacaga agtccggatg ttccaaaggg
361 ttgtga

//

LOCUS AY146957 366 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-S AV2 protein (AV2) gene, complete
cds.

ACCESSION AY146957

VERSION AY146957

KEYWORDS .

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 366)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Complete nucleotide sequence of cotton leaf curl virus (CLCuV-S)
AV2 protein gene from Sirsa, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 366)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..366

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-S"

/country="India: Sirsa"

gene 1..366

/gene="AV2"

CDS 1..366

/gene="AV2"

/note="movement/pre-coat protein"

/codon_start=1

/product="AV2 protein"

/protein_id="AAN61152"

/translation="MWDPLLNEFPD TVHGFR CMLSVKYLQLLSQDYSPDTLGYELIR

LICILRSRNYVEASCRYRHFYARVESTPASELRQPIHQPCCPHCPRHKTTGMDKQAY
EQEAQNVSDVQKSGCSKGL"

BASE COUNT 91 a 88 c 90 g 97 t

ORIGIN

1 atgtgggatac cactattaaa cgaattccct gatacgggtc acgggtttcg gtgtatgctt
61 tctgtgaaat atttgcaact ttgtcgag gattattcac cggatacgt tgggtacgag
121 ttaatacggg atttaattg tattttacgc tccgtaatt atgtcgaagc gagctgccga
181 tategtcatt tctacgcccg cgtcgaaagt acgcccggcg ctgaacttcg gcagcccata
241 caccagccgt gctgtgccc ccattgtcgc cgtcacaaaa caacaggcat ggacaaacag
301 gcctatgaac aggaagccca gaatgtatcg gatgtacaga agtcggatg ttccaaaggg
361 ttgtga

//

LOCUS AY146956 296 bp DNA linear VRL 30-OCT-2002

DEFINITION Cotton leaf curl virus CLCuV-D intergenic region, complete sequence.

ACCESSION AY146956

VERSION AY146956

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 296)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Nucleotide sequence of intergenic region of cotton leaf curl virus (CLCuV-D) from Delhi, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 296)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..296

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-D"

/country="India: Delhi"

misc_feature 1..296

/note="intergenic region"

BASE COUNT 82 a 71 c 57 g 86 t

ORIGIN

1 ttggcaatc ggtgtacct ctaattctat ggcaatcggg gtaacgggg gcaatatata
61 ggtgtacccc aatggcatt atcgtaattt gagaaatcat ttcaaaatcc tcacgtcca
121 aaaagcggcc atccgtttaa tattaccgga tggctgcgcg atttttacc ggatggctgc
181 gcgattttt gtgggcctta ccattaacac ttgtcggcca atcatatgac tccctcaaag
241 ctaaataacg ctccgcaca ctataagtac ttgcgcacta agtttcaaat tcaaac

LOCUS AY146955 260 bp DNA linear VRL 30-OCT-2002

DEFINITION Cotton leaf curl virus CLCuV-H intergenic region, complete sequence.

ACCESSION AY146955

VERSION AY146955

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 260)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Nucleotide sequence of intergenic region of cotton leaf curl virus (CLCuV-H) from Hisar, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 260)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,

Division of Plant Pathology, Indian Agricultural Research

Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..260

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-H"

/country="India: Hisar"

misc_feature 1..260

/note="intergenic region"

BASE COUNT 72 a 62 c 51 g 75 t

ORIGIN

```
1 ttggcaatc ggtgtacact ctaattctat ggcaatcggg gtaacggggg gcaatttata
51 ggtgtacccc aaatggcatt atcgtaattt gagaaatcat ttcaaaatcc tcacgctcca
121 aaaagcggcc atccgtttta tattaccgga tggccgcgcg atttttttg tgggcccccg
131 atttatgaga ttgctccctc aaagctaaat aacgctcccg cacactataa gtacttgcg
241 actaagtffc aaattcaaac
```

//

LOCUS AY146954 260 bp DNA linear VRL 30-OCT-2002

DEFINITION Cotton leaf curl virus CLCuV-L intergenic region, complete sequence.

ACCESSION AY146954

VERSION AY146954

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 260)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Nucleotide sequence of intergenic region of cotton leaf curl virus (CLCuV-L) from Ludhiana, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 260)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology, Division of Plant Pathology, Indian Agricultural Research Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..260

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-L"

/country="India: Ludhiana"

misc_feature 1..260

/note="intergenic region"

BASE COUNT 72 a 61 c 51 g 76 t

ORIGIN

```
ttggcaatc ggtgtacact ctaattctat ggcaatcggg gtaacggggg gcgatatata
11 ggtgtacccc aaatggcatt atcgtaattt gagaaatcat ttcaaaatcc tcacgttca
121 aaaagcggcc atccgtttaa tattaccgga tggccgcgcg atttttttg tgggcccccg
181 atttatgaga ttgtccctc aaagctaaat aacgtcccg cactataa gtacttgcgc
241 actaagtttc aaattcaaac
```

//

LOCUS AY146953 271 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-S intergenic region, complete
sequence.

ACCESSION AY146953

VERSION AY146953

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 271)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Nucleotide sequence of intergenic region of cotton leaf curl virus
(CLCuV-S) from Sirsa, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 271)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..271

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-S"

/country="India: Sirsa"

misc_feature 1..271

/note="intergenic region"

BASE COUNT 84 a 64 c 47 g 76 t

ORIGIN

1 gttgactaaa attgatcacc gattgaccgc tcttgcaact ctccccggta tatcggtgat
61 caatatatag tgatcaccaa atggcataat ggtaataaaa aacctttaat ttgaaattca
121 aaccaaagg ctaaagcggc catccgttta atattaccgg atggcgcgcg cgtttttt
181 gtgggcccc gatttatgag attgctccct caaagctaaa taacgctccc gcacactata
241 agtacttgcg cactaagttt caaattcaaa c

//

LOCUS AY146962 349 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-D replication initiation protein (AC1)
gene, partial cds.

ACCESSION AY146962

VERSION AY146962

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 349)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Partial nucleotide sequence of cotton leaf curl virus (CLCuV-D)
replication initiation protein (AC1) gene from Delhi, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 349)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..349

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-D"

/country="India: Delhi"

gene 1..>349

/gene="AC1"

CDS 1..>349

/gene="AC1"

/codon_start=1

/product="replication initiation protein"

/protein_id="AAN61157"

/translation="MPRDLKTNAKNYFLTYPKCSLTKEETLSQLLNLHTPVNKKYIKI

CRELHQNGEPHLHVLIQFEGKYQCTNKRFFDLVSPTRSAHFHPNIQGAKESSSDVKAYI
DKDGDITLEWGEFQI"

BASE COUNT 122 a 94 c 61 g 72 t

ORIGIN

1 atgccacgtg atttaaaac aaacgcaaa aactacttcc tcacttatcc caaatgctct
61 ctaactaaag aagaaccct ctccaatta ctaaactac atacaccagt aaacaaaaa
121 tacatcaaga tctgcagaga gcttaccac aatggggaac ctcatttca cgtgctcatc
181 cagttcgaag ggaagtacca atgcacgaat aaaagattct tcgacctggt atcccaacc
241 aggtcagcac atttccatcc gaacattcag ggagctaaat caagtcaga tgtcaaggcc
301 tacatcgaca aggacgggga cactctcgag tggggagagt ttcagatcg

//

LOCUS AY146964 388 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-H replication initiation protein (AC1)
gene, partial cds.

ACCESSION AY146964

VERSION AY146964

KEYWORDS .

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 388)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Partial nucleotide sequence of cotton leaf curl virus (CLCuV-H)
replication initiation protein (AC1) gene from Hisar, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 388)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..388

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-H"

/country="India: Hisar"

gene 1..>388

/gene="AC1"

CDS 1..>388

/gene="AC1"

/codon_start=1

/product="replication initiation protein"

/protein_id="AAN61159"

/translation="MPRDLKTNAKNYFLTPKCSLTKEETLSQLLNLHTPVNKKYIKI

CRELHQNGEPHLHVLIQFEGKYQCTNKRFFDLVSPTRSAHFHPNIQGAKSSSDVKAYI
DKDGDITLEWGQFQIDGRSARGGQQTAN"

BASE COUNT 138 a 103 c 73 g 74 t

ORIGIN

1 atgccacgtg atttaaaaac aaacgccaaa aactacttcc tcacttatcc caaatgctct
61 ctaactaaag aagaaacct ctccaatta ctaaacctac atacaccagt aaacaaaaa
121 tacatcaaga tctgcagaga gttcaccaa aatggggaac ctcatttca cgtgctcatc
181 cagttcgagg ggaagtacca atgcacgaat aaaagattct tcgacctggt atcccaacc
241 aggtcagcac atttcatcc gaacattcag ggagctaaat caagtcaga tgtcaaggcc
301 tacatcgaca aggacgggga cactctcgag tggggacagt ttcatatcga tggaagatca
361 gcaagaggag gacaacagac agccaacg

LOCUS AY146963 396 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-L replication initiation protein (AC1)
gene, partial cds.

ACCESSION AY146963

VERSION AY146963

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 396)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Partial nucleotide sequence of cotton leaf curl virus (CLCuV-L)
replication initiation protein (AC1) gene from Ludhiana, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 396)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..396

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-L"

/country="India: Ludhiana"

gene 1..>396

/gene="AC1"

CDS 1..>396

/gene="AC1"

/codon_start=1

/product="replication initiation protein"

/protein_id="AAN61158"

/translation="MPRDLKTNAKNYFLTYPKCSLTKEETLSQLLN LH TPVNKKYIKI

CRELHQNGEPHLHVLIQFEGKYQCTNKRFFDLVSPTRTAHFHPNIQGAKSSSDVKAYI
DKDGDTLGWGEFQIDGRSARGGRQTANDAY"

BASE COUNT 139 a 105 c 77 g 75 t

ORIGIN

1 atgccacgtg atttaaaac aaacgcaaa aactacttc tcacttatcc caaatgctct

61 ctaactaaag aagaaacct ctccaatta ctaaacctac atacaccagt aaacaaaaa

121 tacatcaaga tctgcagaga gttcaccaa aatggggaac ctcatcttca cgtgctcatc

181 cagttcgagg ggaagtacca atgcacgaat aaaagattct tcgacctggt atccccaaac

241 aggacagcac atttcatcc gaacattcag ggagctaaat caagctcaga tgtaaggcc

301 tacatcgaca aggacgggga cactctcggg tggggagagt ttcagatcga tggaagatca

361 gcaagaggag gacgacagac agccaacgac gcttac

//

LOCUS AY146961 328 bp DNA linear VRL 30-OCT-2002
DEFINITION Cotton leaf curl virus CLCuV-S replication initiation protein (AC1)

gene, partial cds.

ACCESSION AY146961

VERSION AY146961

KEYWORDS

SOURCE Cotton leaf curl virus.

ORGANISM Cotton leaf curl virus

Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

REFERENCE 1 (bases 1 to 328)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Partial nucleotide sequence of cotton leaf curl virus (CLCuV-S)
replication initiation protein (AC1) gene from Sirsa, India

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 328)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..328

/organism="Cotton leaf curl virus"

/db_xref="taxon:53010"

/clone="CLCuV-S"

/country="India: Sirsa"

gene 1..>328

/gene="AC1"

CDS 1..>328

/gene="AC1"

/codon_start=1

/product="replication initiation protein"

/protein_id="AAN61156"

/translation="MPPKRNGIYSKNYFITYPKCSLTKEEALSQLLN!QTPTS KKYIR

ICRELHEDGTPHLHVLIQFEGKFKCQNMRFDDLVSRSRSHFHPNIQGAKSSSDVKSY
IEKDGD I"

BASE COUNT 107 a 86 c 56 g 79 t

ORIGIN

1 atgcctccaa agcggaacgg tatttatcc aaaaactatt tcactactta tccaaatgc

61 tctctacca aagaggaaagc acttcccaa ttattaaata taaaacccc aactcaaaa

121 aaatatatta gaatctgcag agagcttcac gaagatggga ctctcactt gcatgttctc

181 atccagttcg aagggaatt caagtgccag aatatgcgat tcttcgactt ggtctcccca

241 agcaggtcag cacattcca tccgaacata caggagacta aatccagctc cgacgtcaaa

301 tctacatcg agaaggacgg ggacattc

//

LOCUS AF534185 902 bp DNA linear VRL 28-AUG-2002
DEFINITION Cotton leaf curl virus-associated DNA beta isolate CLcuV-D gene
sequence.

ACCESSION AF534185

VERSION AF534185

KEYWORDS

SOURCE Cotton leaf curl virus-associated DNA beta.

ORGANISM Cotton leaf curl virus-associated DNA beta

Viruses; Satellites; Satellite Nucleic Acids;

Begomovirus-associated DNA beta-like sequences.

REFERENCE 1 (bases 1 to 902)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (06-AUG-2002) Advanced Centre for Plant Virology,

Division of Plant Pathology, Indian Agricultural Research

Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..902

/organism="Cotton leaf curl virus-associated DNA beta"

/isolate="CLcuV-D"

/db_xref="taxon:134386"

/country="India: Delhi"

BASE COUNT 266 a 161 c 210 g 265 t

ORIGIN

```
1 caaggatata ttattatga aagaaatggg agcgcagcga atcgaaacag gaaaacccaa
61 ggaaagagaa aaaataaaaa gtaaagagaa aaaaaatata aatcgaaaa cgatcatggt
121 tgagagtaat aaaaaaaaaa aaaaaacaaa aacatatccg aaaacgttgt cgttgaagg
181 tgactgtgtg gttttacat ttactgtgtg gtaaattgga agtgattgaa gaataattaa
241 aaaatttgc accaatagg aaattgatca ccatatatc ggtgatcaat tgggcctca
301 taattggtg tgcctaaaat accctattt ctgtgtctgg aaggcgcgtg gaagtgcgt
361 gaaaaagggtg ttcttcttc acctaaaacc caccggaacg caaaactgg ctgtgccgg
421 catcaattta cgacacgcgc ggcggtgtgt accctggga gggtaggtac cactacgcta
481 cgcagcagcc ttagctacgc cggagcttag ctgcccacg cttaatat accgtgggcg
541 agcgggtgcc gatggcttct tgggtgggtc cattgctggt atgacttga ttgacttat
601 attgggcca ttaattgggt tgaaaatgt tgggccttg gaagagggtc tttatggaa
661 tagggctctg tatattgtg gtgatgtg tgtaaatat gcattgctgg ttgtgttg
721 gaattaaac ggtgaacttt ttattgaata agtacgggtc gtttacatc attccaata
781 tctctgggtt ttcaagtaca agtatatcaa gtctgtgaac tatacttct atctgatct
841 ctctatctt tgccccgtg tatgcgaata ggaaattcgc tatgatgct cttcaaagc
901 cg
```

//
LOCUS AF534186 1026 bp DNA linear VRL 28-AUG-2002
DEFINITION Cotton leaf curl virus-associated DNA beta isolate CLCuV-H gene
sequence.

ACCESSION AF534186

VERSION AF534186

KEYWORDS .

SOURCE Cotton leaf curl virus-associated DNA beta.

ORGANISM Cotton leaf curl virus-associated DNA beta

Viruses; Satellites; Satellite Nucleic Acids;

Begomovirus-associated DNA beta-like sequences.

REFERENCE 1 (bases 1 to 1026)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (06-AUG-2002) Advanced Centre for Plant Virology,

Division of Plant Pathology, Indian Agricultural Research

Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..1026

/organism="Cotton leaf curl virus-associated DNA beta"

/isolate="CLCuV-H"

/db_xref="taxon:134386"

/country="India: Hisar"

BASE COUNT 323 a 191 c 225 g 287 t

ORIGIN

1 gaaatgggag cgcagcgaat cgaaacagga aaacccaagg aaagagaaaa aataaaaagt
61 aaagagaaaa aaaatataaa ttcgaaaacg tcatcgtttg agagtaataa aagaaaagca
121 ataacaaatc aaacatgagc agagagaaat aaaaacaaga aatatctgtg gtccccactg
181 aaaaaataaa aagaaaaataa aaaataaaat acaactgttt ttactgagtc gtaaatcagt
241 agttcaataa ttactgaagt taatgggtaa atatagaaag gtaaatttgt caccaatagg
301 taaattgatac accgatatat cggtgatcaa ttgggtgcctt ataattgggtg ttgcctaaaa
361 tacccttatt tctgtgtctg gaaggcgcgt ggaagtgcgc taaaaaagggt gttcttctct
421 cacctaaaat cgcgcggaac gccaaaactg gctgttgccg gcatcaattt acgacacgcg
481 cggcgggtgtg taccctggg agggtaggta ccactacgct acgcgcagcc tttagctacgc
541 cggagcttag ctgcccacg cttaatatatt accgtgggcg agcgggtgcct cgggggacgc
60 agtgggtccc acatgttaga tctgcagaat taatgggcgc ttaaactgta ttgggcctat
661 ttgtattggg cttttatcat atatttaaat tgggcttcgt atattagaga atgtttggaa
721 tttaaacggt gaacttttta ttgaatacgt acggttcgtt tacatccatt cccaatatct
781 ctgggttttc aagtacaagt atatcaagtc tgtgaactat atcttctgtc tcgatctctt
841 ctatctttgc ctggttag gcgaaatagga aatcgctat gatgctccct tcaaagccgt
901 tgaagtcgaa tggaacgtgt atgtcttcgt acgtgtactg gccgatccct tcatactga
961 ttacgatgg tgactttgtg gatagtatcc tcattgtgaat gaagatcttc atatttcca
1021 tgatgc

LOCUS AF534187 1013 bp DNA linear VRL 28-AUG-2002
DEFINITION Cotton leaf curl virus-associated DNA beta isolate CLCuV-L gene
sequence.

ACCESSION AF534187

VERSION AF534187

KEYWORDS

SOURCE Cotton leaf curl virus-associated DNA beta.

ORGANISM Cotton leaf curl virus-associated DNA beta
Viruses; Satellites; Satellite Nucleic Acids;
Begomovirus-associated DNA beta-like sequences.

REFERENCE 1 (bases 1 to 1013)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (06-AUG-2002) Advanced Centre for Plant Virology,
Division of Plant Pathology, Indian Agricultural Research
Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..1013

/organism="Cotton leaf curl virus-associated DNA beta"

/isolate="CLCuV-L"

/db_xref="taxon:134386"

/country="India: Ludhiana"

BASE COUNT 361 a 202 c 209 g 241 t

ORIGIN

```
1 gataaaaaaa aggaaattaa aaaatgaatt caaaagaatt tagaagacat aataagaaag
61 aagaaaaata taaacagaaa atatgaaagg aaaacataaa catnaatctc ttgagaaaat
121 ttgggagcgc agcgaaatag agccaaacag aaataaataa ccgagaacat gagcagagaa
181 aatggaaaaa acaaacgaaa ctaatagggt cctctatgca aaaataagaa aatccgtaca
241 cagtaattaa tcattaatta ctgcgcagta aatgtgaata aaattaacca acagggttaat
301 ttggagact ccaataggta aataagacct caatatattg gggctcfaat tggggactgg
361 aaagagaaat aaatccccga tccccaaaat acccctaatt gtgtgtcttg aaggcgctg
421 gtaatgcgct gaaaaagggt acctctctc tctagaact caccggaacc gccaaactgg
481 ctgattccgg catcaattta cgacacgcgc ggcgggtgtg acccctggga gggtaggtac
541 cggtagcact acgtacgca gcagccttag ctacgccgga gcttagctcg ccacgttct
601 aatattaccg tgggcgagcg gtgaccgatg gcttcttggg gggteccatg ttgaaggaga
661 aagataaatg ggctagtta atgcaattgg gcctgaattg aaatgggctt ggaccagtag
721 attcgggact gggccaatac aataagataa gaaatgggtt cgtaaattaa aacaaaggat
781 ttattcattt caaatgcatt acgcactcac acacacacat tcgtacgtac atcgtattca
841 tctctatac gtatatcaac taatggagct tcttccatca tgagtaaate aattgtctcc
901 accatacct cttgtcgaac ctctctata acagaatcct tgtacatgat ttgcaacaga
961 tttgtatac ctctcttag agcgttgaag tcgaacggag gtatgatccc atc
```

LOCUS AF534188 1026 bp DNA linear VRL 28-AUG-2002
DEFINITION Cotton leaf curl virus-associated DNA beta isolate CLCuV-S gene
sequence.

ACCESSION AF534188

VERSION AF534188

KEYWORDS

SOURCE Cotton leaf curl virus-associated DNA beta.

ORGANISM Cotton leaf curl virus-associated DNA beta

Viruses; Satellites; Satellite Nucleic Acids;

Begomovirus-associated DNA beta-like sequences.

REFERENCE 1 (bases 1 to 1026)

AUTHORS Radhakrishnan,G., Malathi,V.G. and Varma,A.

TITLE Direct Submission

JOURNAL Submitted (06-AUG-2002) Advanced Centre for Plant Virology,

Division of Plant Pathology, Indian Agricultural Research

Institute, Pusa Road, New Delhi, Delhi 110012, India

FEATURES Location/Qualifiers

source 1..1026

/organism="Cotton leaf curl virus-associated DNA beta"

/isolate="CLCuV-S"

/db_xref="taxon:134386"

/country="India: Sirsa"

BASE COUNT 306 a 181 c 248 g 291 t

ORIGIN

1 aaccaaagaa aagaacaag gatattat ttatgaaaga aatgggagcg cagcgaatcg
61 aaacaggaaa acccaaggaa agagaaaaaa taaaaggaa agagagaaaa taaaagtaa
121 agagaaaaaa aatataaatt cgaaaacgtc atcgttgag aagagggaaa aaaaagaaa
181 aaacaaaatt agaaccgaga acgttgccgt ttgaagggtga ctgggtggtt ttaccattta
241 ctgggtggta aatggtaagt gattgaagaa ggggttaaaa tagagactct gataggtaaa
301 ttgatcaccg atatacggg gatcaattgg tgaccaataa attaccttc ctaaaatacc
361 cccgcttttg tgtctaatag gcgcgtcgga gtgcgctgat aaagttaaca ttctctctc
421 tacaatgaga cctaatacaa tttcccggtg atcggagtcg aattttccga cagcgcgagg
481 ggtgtgtacc cctgggaggg taggtaccac tacgtacgc agcagcctta gctacgccgg
541 agcttagctc gccacgctt taatattacc gtgggcgagc ggtgcccggt ggccgcgcag
601 tgggtccac tgctgtctt gactggattt gacttcta atgtccaattt attgggcttg
661 gaagtgttg ggcctttgaa gatggcctt tatggattat gggtctgtt gtgttggtg
721 atatgtgtg taaatatgca ttgctggtt gtggttgaa ttaaacggt gaactctta
781 ttgaatatgt atggttcgat tacatccatt ccaatatct ctgggtttc aagtacaagt
841 atatcaagtc tgtgaactat atctctatc tcatctctt ctatcttgc cccgttgat
901 gcgaatagga aattcgtat gacactccct tcaaagccgt tgaagtcgaa tggaacgtg
961 atgtctctg acgtgtactg gatgatccct tcacttga ttacgatgg tgactttgt
1021 gatagt

//

7-7/25

