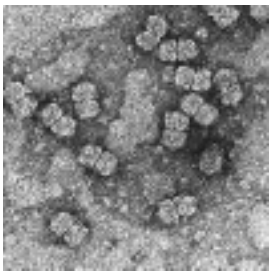
Genome Characterization of Whitefly-transmitted Geminiviruses of Cotton and Development of Virus-resistant Plants through Genetic Engineering and Conventional Breeding



Final Report
1996-2001

JOHN INNES CENTRE

NIBGE

THE UNIVERSITY OF ARIZONA
TUCSON ARIZONA

The report is compiled and edited by following

Dr. Shahid Mansoor
Dr. Rob Briddon
Ms. Shaheen Asad
Dr. Yusuf Zafar
Dr. Judith K. Brown
Dr. Peter G. Markham
Dr. Kauser A. Malik

Foreword

Cotton is an important commodity, both in terms of its contribution to the national economies of a number of developing countries as well as in terms of its many applications. As a natural product, cotton faces several challenges. In addition to competition from man-made fibres, cotton has to overcome a wide range of nature-related problems, such as cotton diseases, which reduce both quality and quantity of cotton produced. Specifically the Cotton Leaf Curl viral disease (CLCuV) and the Cotton Leaf Crumple viral disease (CLCrV) have caused severe losses in production in a number of countries, including Pakistan where the production decline of 40% in the 1992/3 and 1993/4 cotton seasons was mainly attributed to the CLCuV transmitted by whiteflies.

The project *Genome Characterization of Whitefly-transmitted Geminiviruses of Cotton and Development of Virus-resistant Plants through Genetic Engineering and Conventional Breeding* was designed by three collaborating research institutes: the *National Institute for Biotechnology and Genetic Engineering* (NIBGE) in Pakistan, the *John Innes Centre* of the United Kingdom and the *University of Arizona* in the United States of America. The project proposal was submitted by the International Cotton Advisory Committee (ICAC) to the Common Fund for Commodities for financing. The project was implemented during 1996 – 2001 under the overall direction of NIBGE, the Project Executing Agency. Technical supervision and guidance was provided by the International Cotton Advisory Committee.

Findings of this research project have been published in a range of scientific articles and publications, accessible to all who have a direct interest in the subject of whitefly-transmitted diseases.

This Technical Paper is a concise reflection of the results obtained by the project and can be seen as an account of the work that has been undertaken in the framework of the project. The report is published as a CFC Technical Paper in line with the Fund's practice to share the experiences gained in its projects with Common Fund Member Countries and others that have an interest in the subject matter.

It is hoped that this publication will contribute to increasing the knowledge base in the field of engineering cotton to become more virus-resistant, and that it will lead to tangible benefits for cotton producing countries which are affected by virus-based cotton diseases.

Dr. Rolf W. Boehnke
Managing Director
Common Fund for Commodities

Project Summary

1. Title

Genome Characterization of Whitefly-transmitted Geminiviruses of Cotton and Development of Virus-resistant Plants through Genetic Engineering and Conventional Breeding

2. Number

CFC/ICAC-07

3. Project Execution Agency

National Institute of Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan (Dr. Kauser A. Malik)

4. Location

National Institute for Biotechnology and Genetic Engineering (NIBGE), P O Box 577, Jhang Road, Faisalabad, Pakistan (PI Dr. Yusuf Zafar, email; yzafar@nibge.org)

Cotton Research Institute (CRI), Ayub Agricultural Research Institute, Faisalabad, Pakistan (PI Ch. Waheed Sultan)

Department of Disease and Stress Biology, John Innes Centre, Norwich, NR4 7UH, Norfolk, UK (PI Dr. Peter Markham email; rob.briddon@bbsrc.ac.uk)

Department of Plant Sciences, Forbes Bld. Room 303, University of Arizona, Tucson, AZ 85721 USA (PI Dr. Judith K. Brown Email; jbrown@ag.arizona.edu)

5. Starting date

12th January, 1996

6. Completion date

31st December, 2001

7. Financing

Total Project Cost:	US$ 3,926,518
CFC Financing:	US$ 1,549,770
Counterpart contribution	
Pakistan:	US$ 1,324,629
United Kingdom:	US$ 791,226
USA	US$ 245,608
ICAC	US$ 15,285
Total	**US$ 2,376,748**

Preface

Six years tripartite project entitled "Genome characterization of whitefly-transmitted geminiviruses of cotton and development of virus-resistant plants through genetic engineering and conventional breeding" administratively came to an end in December 2001. A detailed report regarding achievements of this project is being presented. An International Symposium on Cotton Biotechnology was planned for November 2001 to share the results of this project with the International scientific community but had to be postponed due to unfortunate happenings since September 11, 2001.

As is evident from this report, the project has been able to achieve all the objectives as outlined in the original proposal. This has been due to the dedication and commitment of all the collaborating partners. I would like to thank all of them. This project was initiated at a time when cotton leaf curl virus (CLCuV) was playing havoc with the cotton crop in Pakistan. The results obtained through this project have gone a long way in combating this menace by undertaking the biology of the virus, its genetic make up, its diversity and mode of infection. Great effort was put in proving the Koch's postulate by inducing the symptom in cotton plant. Moreover, optimization of cotton transformation system and adoption of various strategies for imparting resistance against CLCuV have resulted in various transgenic lines, which are at various stages of evolution. A series of diagnostic tools have also been developed which can expedite screening of new cotton germplasm.

The financial support provided by the Common Fund for Commodities (CFC) under the supervision of the International Cotton Advisory Committee (ICAC) has been crucial in achieving the project objectives. The mid term evaluation of the project by Dr. Norma L. Trolinder and Dr. Johannes F.J.M. van den Heuvel has also been great help in streamlining some of our activities. The financial support provided by the Punjab Government for cotton biotechnology research to NIBGE is gratefully acknowledged.

I am also personally thankful to Mr. Sietse van der Werff, First Project Manager of CFC and Dr. Rafiq Chaudhry of ICAC for their continuous support in both administration and technical matters. Without their help, patience and understanding execution of this project would not have been possible.

Dr. Kauser A. Malik
Project Execution Agency
Member Biosciences
Pakistan Atomic Energy Commission
Islamabad, Pakistan

Symptoms of diseases associated with whitefly-transmitted geminiviruses. Top left; cotton leaf curl disease in Pakistan, Top right; cotton leaf curl disease in Sudan; bottom left; cotton leaf crumple in Arizona, bottom right; cotton leaf crumple from Brazil

Acknowledgements:

The Project Execution Agency and staff at three collaborating groups thank all those who contributed to this project. Special thanks are also due to all those unofficial collaborators who contributed materials and time and without whom this project could not have been completed.

These include:
Prof. Jeffery W. Davies, Department of Disease and Stress Biology, John Innes Centre, Norwich, UK
Dr. John Stanley, Department of Disease and Stress Biology, John Innes Centre, Norwich, UK
Prof. David C. Baulcombe, The Sainsbury Laboratory, John Innes Centre, Norwich, UK.
Prof. Conrad P. Lichtenstein, Queen Mary and Westfield College, London, UK.
Dr. Claude M. Fauquet, ILTAB, Donald Danforth Centre, USA.
Prof. N. Rishi and Dr. P. Dhawan, (Department of Plant Pathology) and Dr. S.S. Siwatch (Department of Plant Breeding) CCS Haryana Agricultural University, Hisar - 1125004, India.
Dr. A. M. Abdel Salam, Plant Pathology Department, Faculty of Agriculture, Cairo University, Giza, Egypt.
Dr. D. Russell, Natural Resources Institute, Pest Management Department, Central Avenue, Chatham Maritime, Kent, ME4 4TB, UK.
Dr. I. Dasgupta, Dept. of Plant Molecular Biology, University of Delhi, South Campus, New Delhi 110021, India.
Mr. Jahangeer Khan Tareen, former Chairman, Agriculture Task Force, Government of Punjab, Pakistan.
Dr. Neil Forrester, DeltaPine, Australia.
Dr. Ali Rezaian, CSIRO, Plant Industry, Adelaide, Australia.
Dr. Zahoor Ahmad, ex Director, Central Cotton Research Institute (CCRI), Multan, Pakistan.
Dr. Zafar Altaf, former Secretary, MINFAL, Islamabad, Pakistan.
Dr. Barkat Ali Soomro, Vice President PCCC, Karachi, Pakistan.
Dr. Zakir Hussain, ex Cotton Commissioner, MINFAL, Islamabad, Pakistan.
Mr. Munir-ud-Din Khan, ex Director Cotton Research Institute, Multan/Faisalabad, Pakistan.
Mr. Akmal Khan, ex Director Cotton Research Institute, Rahim Yar Khan, Pakistan.
Dr. Ahmad Saleem Akhtar, Director, Plant Protection Institute, Faisalabad, Pakistan
Prof. Muhammad Yusuf, Department of Entomology, University of Agriculture, Faisalabad, Pakistan.

Table of Contents:

EXECUTIVE SUMMARY

Whitefly-transmitted geminiviruses are recognized as the most widespread and the most severe pathogens of cotton. Among cotton diseases associated with geminiviruses cotton leaf curl disease in Pakistan is the most serious pathogen that has devastated cotton crop in Pakistan and has spread to all cotton growing areas of Northern India. Cotton diseases associated with begomoviruses are also known to occur in cotton growing areas of Africa where a disease similar to cotton leaf curl disease in Pakistan is found in the field. In Americas cotton leaf crumple and other diseases of cotton are also associated with geminiviruses. Besides these diseases, geminiviruses found on several hosts belonging to the family *Malvacae* are a potential threat to cotton crop as some of these viruses could be transmitted to cotton under experimental conditions.

Realizing the importance of geminiviruses to sustainable cotton production, the project was submitted to the Common Funds for Commodities (CFC) by The International Cotton Advisory Committee in 1993. The Fund, after clarifications on use of genetic engineering for breeding resistance in cotton, approved the project in 1995. NIBGE was accepted as Project Executing Agency (PEA) with components at the John Innes Centre, Norwich, UK and University of Arizona, USA.

The objectives of the project were

- Characterization of cotton geminiviruses found in Pakistan, Africa and Americas.
- Development of *in vitro* tissue culture techniques for cotton transformation
- Development of virus resistant cotton both by conventional breeding and genetic engineering

Major achievements of the project

Besides achieving all the objectives, the project has contributed significantly to our understanding of geminiviruses, development of scientific infrastructure at NIBGE, transfer of technology and most importantly development of a nucleus of dedicated scientists that are ready to accept new challenges in the field. The knowledge and materials generated in the project has a potential for applications in understanding cotton genome and solving related problems in other crops.

The major achievements of the project are outlined below

Characterization of cotton begomoviruses

1. First report of routine purification of geminivirus particles from infected cotton plants
2. Characterization of several distinct geminiviruses associated with the disease
3. Infectious clones of four distinct geminiviruses associated with the disease
4. The first report of a functional DNA satellite-like molecule DNA 1
5. Identification of the satellite molecule DNA ß essential for infectivity

6. Identification of functions of DNA ß
7. Identification of disease complexes similar to cotton leaf curl disease on several plant species in Pakistan
8. Establishment of an efficient system for the transmission of the disease by whiteflies
9. Identification of factors that affect disease transmission by whiteflies
10. Identification of natural hosts of cotton leaf curl disease
11. Development of molecular diagnostic tests for the detection of cotton geminiviruses
12. Production of antisera against *in vitro* expressed coat protein
13. Cloning of genomic components of cotton leaf curl disease from Sudan and Egypt and the infectious β clone
14. Cloning of genomic components of cotton leaf crumple virus from Arizona, USA
15. Establishment of sequences and a database for identification of geminiviruses that infect cotton and other related hosts
16. Establishment of infectivity of sida golden mosaic virus from South America on cotton

Development of *in vitro* transformation of cotton

1. Establishment of a routine system for transformation of regenerable varieties of cotton
2. Establishment of a system for transformation of local elite cultivars of cotton based on transformation of mature embryos

Development of virus resistant genotypes

1. Evaluation and confirmation of natural sources of resistance against cotton leaf curl disease
2. Development of constructs for plant transformation based on expression of sense and antisense RNA
3. Transformation of constructs in tobacco and cotton and development of virus-resistant plants
4. Development of constructs based on the expression of the coat protein to provide broad-spectrum resistance against distinct begomoviruses associated with the disease
5. Transformation of coat protein constructs in tobacco
6. Development of more constructs for resistance based on the findings that DNA ß is essential for infectivity
7. Transformation of constructs based on DNA ß in tobacco
8. Breeding of virus resistant cotton genotypes by conventional breeding and release of several varieties (FH-900, FH-901, MNH 552, MNH 554)
9. Multilocation trials of virus resistant cotton genotypes
10. Approval and distribution of virus resistant genotypes to farmers in Pakistan
11. Availability of sources of resistance to cotton leaf curl disease for their evaluation for resistance to geminivirus diseases in other parts of the world.

FINAL REPORT

Geminiviruses are single-stranded DNA viruses with circular genome that are grouped in family *Geminiviridae*. Geminiviruses have been divided into four genera based on their insect vectors and genome organization. Whitefly-transmitted geminiviruses (genus *Begomovirus*) are the biggest and the most important group of geminiviruses. At the time of writing the project proposal for studying the causative agent of cotton leaf curl disease-Pakistan (CLCuD-Pak) two types of begomoviruses were known. The bipartite begomoviruses (the vast majority of the then characterized begomoviruses) have genomes consisting of two components with genes involved in replication, control of gene expression and insect transmission encoded on DNA A and genes involved in inter and intracellular movement in plants encoded on DNA B. The second group consists of monopartite begomoviruses with a component homologous to the DNA A of the bipartite viruses but no DNA B.

Project Rationale and Objectives

Begomoviruses are a major threat to the productivity and quality of cotton grown in the subtropical and tropical regions of the world. Begomoviruses are recognized as the most abundant and most severe viral pathogens of cotton, worldwide. The most important diseases known to be transmitted by the whitefly vector are the cotton leaf curl viral (CLCuV) disease and the cotton leaf crumple viral (CLCrV) disease. The CLCuV has previously caused heavy losses in cotton producing countries like Sudan, Egypt and Nigeria. Recent years have already shown an upsurge in the recorded outbreaks of the whitefly vector, leading to dramatic losses in cotton production in Pakistan, Chad, India and Mali as a result of the virus epidemics. Other affected countries are the USA, the Dominican Republic, Guatemala, Nicaragua and Togo.

It was envisaged that, with the combined efforts of the collaborating laboratories, the predominant begomovirus pathogens of cotton occurring in both Old and New World countries will be characterized in biological and genetic terms. The information thus developed will be used to investigate epidemiological parameters and the infra-structural basis to develop geminivirus resistant cotton.

The project comprised of the following four components:

(i) biological and genetic characterization of begomoviruses of cotton
(ii) development of *in-vitro* transformation systems of cotton;
(iii) Development of geminivirus-resistant genotypes of cotton by conventional breeding and genetic engineering and to carry out field trials to establish the validity of resistance in different locations.

Project implementation and results achieved

Objective 1

Biological and molecular characterization of whitefly-transmitted geminiviruses of cotton

Summary

The project identified that cotton leaf curl disease is associated with several distinct monopartite begomoviruses showing homology to other begomoviruses from the region. Full-length clones of these viruses were systemically infectious to tobacco and cotton but were unable to develop typical disease symptoms, suggesting that unusual components are required to cause symptoms. The conclusion resulted in concerted efforts to identify additional components that resulted in the discovery of DNA 1 and DNA ß.

This project identified a third group of begomoviruses, namely monopartite begomoviruses that are associated with a satellite molecule (termed DNA ß) to induce symptoms in some hosts. DNA ß was shown to require the begomovirus for replication, spread in plants and insect transmission, behaving as a satellite. The identification of this molecule, as well as the satellite-like molecule DNA 1, has opened up a whole new area of geminivirus research. In fact DNA 1 and DNA ß are the first examples of functional satellites in DNA viruses.

Four distinct begomoviruses cloned from infected cotton were found to be associated with a single species of DNA ß. Co-inoculation of any of distinct DNA A with DNA ß was sufficient to cause disease symptoms. Distinct begomoviruses associated with the disease are very often found in multiple infection but these results showed that multiple infection is not necessary to develop disease symptoms.

Another phenomenon observed in the course of the project was the emergence of new begomoviruses by recombination of existing viruses and expansion in the host range of cotton leaf curl and other begomoviruses. Several natural hosts and other begomoviruses from the region were detected in several new host plants.

The project showed that the geographic range of diseases associated with DNA ß molecules covered all areas where begomoviruses are known, with the exception of the Americas. This suggests that the DNA ß and begomovirus interaction may be an ancient one. These diseases cause losses in a number of important crops including cotton, peppers, okra, tobacco and tomato. Initial evidence would seem to suggest that the diseases caused by begomovirus-DNA ß complexes are displacing diseases caused by the previously prevalent bipartite and monopartite begomoviruses. The reason for this is unclear but may indicate that begomovirus-DNA ß complexes have a selective advantage over the other begomoviruses. Certainly the scale and number of disease caused by begomovirus-DNA ß complexes are on the increase in this region.

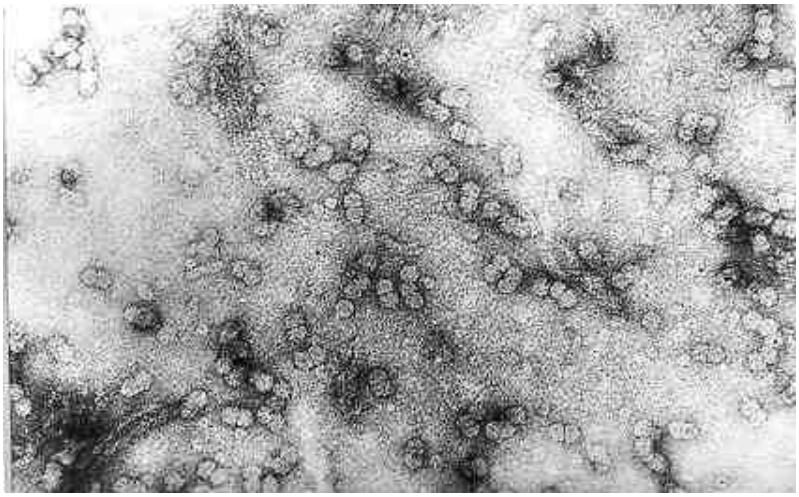
Co-inoculation of distinct begomoviruses and DNA ß revealed interesting data about the replication and interactions between DNA A and DNA ß. Sequence analysis of DNA ß was unable to predict rep binding domains previously identified on DNA A. Distinct DNA ß cloned from Pakistan were replicated in trans by ageratum yellow vein virus (AYVV) from Singapore and developed severe symptoms on *Nicotiana benthamiana*. On the other hand, inoculation of cotton ß with DNA A of AYVV was unable to induce disease symptoms, showing that there are both virus and host specific interactions between DNA A and DNA ß. Similarly, all distinct begomoviruses associated with cotton leaf curl disease trans replicated DNA ß.

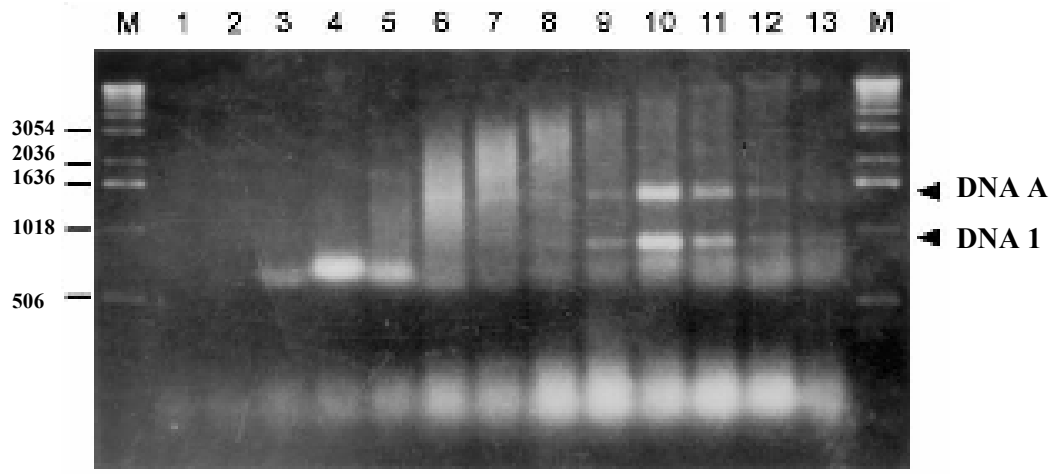
A single gene (C1) was shown to be conserved between all the DNA ß molecules sequenced and mutation of this gene abolished induction of severe symptoms. The promoter analysis revealed a constitutive strong promoter activity. The expression of DNA ß from PVX vector resulted in severe chlorite and necrotic symptoms, suggesting that the gene is a virulence determinant. ß C1 was shown to suppress post-transcriptional gene silencing (PTGS), a plant defense mechanism. ßC1 complemented 25K protein of PVX, suggesting that DNA ß also promote virus movement.

Detailed report on objective 1

1.1 Isolation of geminate particles from cotton

A protocol for routine purification of geminate particles from infected cotton plants was established. Previously all attempts to purify virus particles from infected cotton failed probably because of production of interfering compounds in cotton. Previously the routine protocol was to inoculate laboratory host plants and then purify the geminate particles from alternate/lab host. Several published protocols for geminivirus purification from the infected leaf tissues were tried. Finally we were able to establish protocol for the purification geminate particles from cotton. For isolation of geminate particles fresh tissue collected from the field (100g) was used. Leaves were frozen in liquid nitrogen and homogenized in Warring blender after adding 200 ml of ice-cold virus isolation buffer (100 mM Na-citrate buffer pH 6.0, 18.5 mM Ascorbic acid, 60 mM Na_2SO_3, 1% v/v 2-mercaptoethanol, 5mM EDTA). The homogenate was made 2.5% (v/v) with Triton X-100 stirred overnight, squeezed through four layers of cheesecloth and clarified by 10 min centrifugation at 8,000g. The supernatant was centrifuged for 3 hr at 90,000g in Beckman SW 28 rotor. Pellets were suspended in a CEM buffer (10 mM Na-citrate, 1 mM EDTA & 0.1 % 2-mercaptoethanol) and submitted to a similar low speed centrifugation followed by high-speed centrifugation. Pellets were again suspended in 10% (w/v) sucrose in CEM (1ml of buffer per equivalent of 100g of leaves), homogenized with test tube and clarified by 10 min centrifugation at 4,000g. The supernatant was applied to a 32ml linear, 10-50% (w/v) sucrose gradient in CEM (extracts obtained from 100g of tissue were applied on each gradient). The gradients were fractionated (2ml per fraction) after being subjected to a 14-hour centrifugation at 90,000g in the SW 28 rotor. Virus was concentrated by centrifugation at 45000 rpm for 1 hour in SW 50.1 Ti rotor. The geminate particles were visualized by JEOL 1010 electron microscope in 2% aqueous urinyl acetate.





Purification of geminate particles and identification of DNA forms encapsidated in geminate particles. Top; geminate particles isolated from Characterization of DNA forms isolated from geminate particles revealed presence of small DNA molecules corresponding to DNA 1 and DNA A.

1.2. Isolation of CLCuV DNA from purified geminate particles

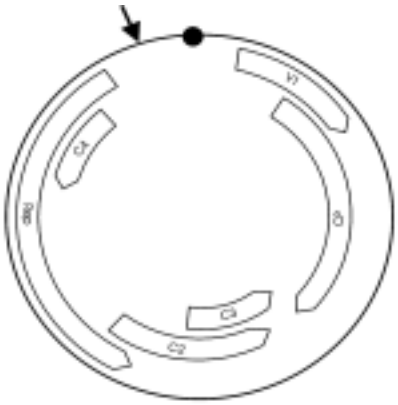
Viral DNA was extracted from sucrose gradient fractions or from concentrated viral preparations by phenol-chloroform method, following 2 hr incubation in 0.2% (w/v) sodium dodecyl sulfate (SDS) and 50μg/ml proteinase K. After ethanol precipitation, DNA was suspended in sterile de-ionized water and stored at -20 ^{0}C. An aliquot of viral DNA was run on 1% agarose gel and DNA bands were visualized on ethidium bromide stained gel. Two types of DNA molecules were found to be associated with virus particles; a usual full-length molecule and an unusual DNA molecule that was approximately the half in size. Characterization of molecules associated with the smaller band resulted in identification of DNA 1 and DNA β.

1.3 Characterization of begomoviruses associated with cotton leaf curl disease in Pakistan

The genomes of the vast majority of begomoviruses consist of two components (Padidam *et al.*, 1995, Rybicki, 1994). The first of these components (DNA A) encodes all viral functions required for replication, control of gene expression and encapsidation. The second, DNA B, encodes two products involved in movement of the virus between and within plant cells (Noueiry *et al.*, 1994).

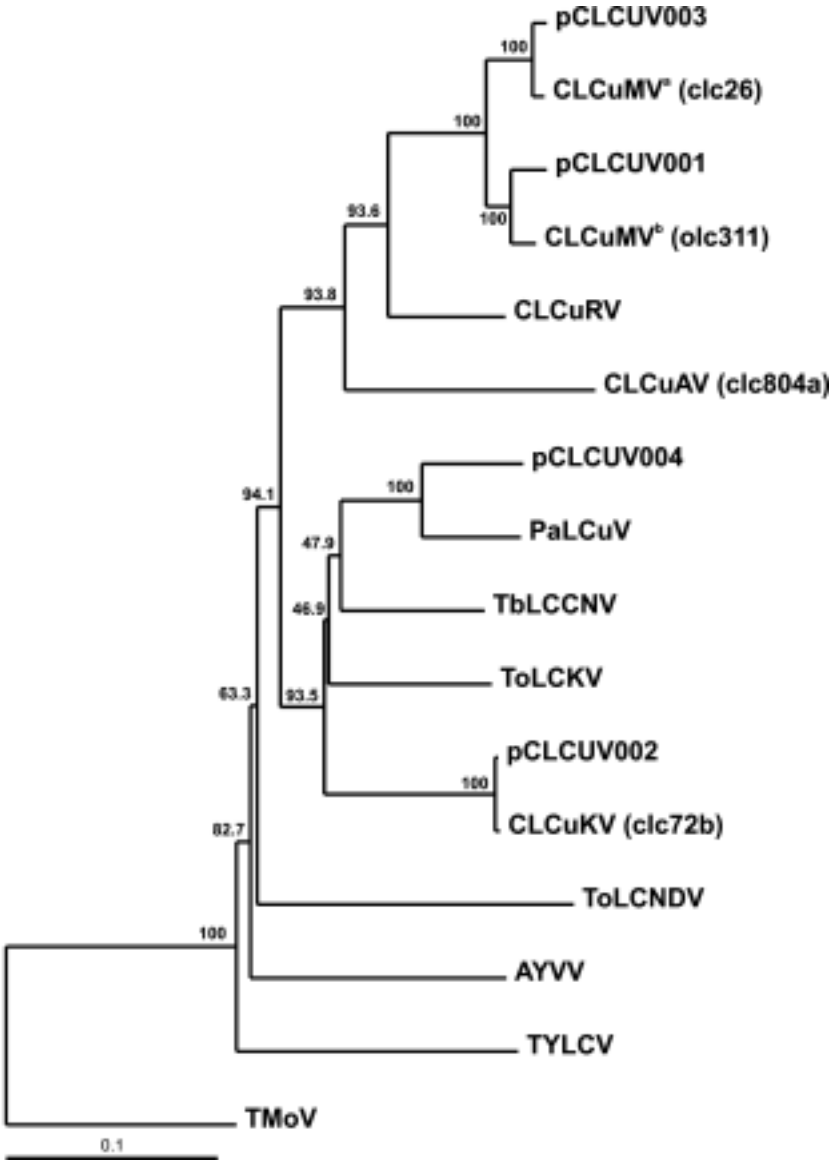
A small number of begomoviruses, mostly isolated from tomato, have monopartite genomes, appearing to be able to dispense with the functions encoded by DNA B. For these viruses various gene products encoded by their genomes (a homologue of the DNA A component of bipartite viruses) appear able to take over the functions of the DNA B products in movement (Rojas *et al.*, 2001). The group of monopartite begomoviruses includes *Tomato yellow leaf curl virus*, a major cause of losses to tomato crops globally.

Cotton leaf curl disease (CLCuD) has been shown to be associated with several viruses of the genus *Begomovirus* (family *Geminiviridae*) that were collectively referred to as cotton leaf curl virus (Briddon *et al.*, 2000, Hameed *et al.*, 1993, Harrison *et al.*, 1997, Mansoor *et al.*, 1993, Nadeem *et al.*, 1997, Zhou *et al.*, 1998). The study of Zhou *et al.* (1998), into the diversity of begomoviruses associated with CLCuD, concluded that there were essentially four variants (now classified as species) of begomoviruses infecting cotton in Pakistan. These species have been named *Cotton leaf curl Multan virus* (CLCuMV), *Cotton leaf curl Alalabad virus* (CLCuAV), *Cotton leaf curl Rajasthan virus* (CLCuRV) and *Cotton leaf curl Khokhran virus* (CLCuKV) represented by clones clc26, clc804a, olc311 and clc72b respectively. None of these studies, however, showed a causative relationship between the viruses and the disease. Only one of the viruses, CLCuMV, was represented by an infectious clone which was not capable of inducing characteristic leaf curl symptoms in host plants (Briddon *et al.*, 2000).




```
                                                        TATA
pCLCUV003       2623- TACT.C.TG. GGAATTGGAG ..ACTGGATA CA.ATTTATA G.TGT..CTC CAAATGGCAT -2673
CLCuMV (clc26)  2624- TACT.C.TG. GGAATTGGAG ..ACTGGATA CA.ATTTATA G.TGT..CTC CAAATGGCAT -2674
pCLCUV001       2602- TTCT.C.TG. GCAATTGGTG TAACGGGGTG CA.ATATATA GGTGTA.CCC CAAATGGA.T -2659
CLCuMV (clc311) 2602- CTCT.C.TG. GCAATCGGTG TAATGGGGTG CA.ATATATA GGTGTA.CCC CAAATGGCAT -2650
pCLCUV002       2625- TCTTGCAAAC TCTCCCCGGT ATATCGGTGA TCAATATATA G.TG.ATCAC CAAATGGCAT -2682
CLCuKV (clc72b) 2624- TCTTGCAA.C TCTCCCCGGT ATATCGGTGA TCAATATATA G.TG.ATCAC CAAATGGCAT -2680
pCLCUV004       2636- TCTGAGGA.. ...ATGGGGG ACTCCGGGGA CGCATTTATA T.GGCGTCCC CCAATGGCTT -2689
PaLCuV          2631- TCTGAGGA.. ...ATGGGGG ACTCTGGGGA CGCATTTATA T.GGTGTCCC CAAATGGCTT -2684
CLCuAV (clc804a) 2614- AACT.CATA. TCAATTGGGG ACACTGGGGA CTCATTTATA C.TGCGTCCC CAAGTGGCAT -2670
CLCuRV          2630- TCTTGCAAC. TCTCCCCGGT ATATCGGTGA TCAATATATA G.TGA.TCAC CAAATGGCAT -2686
```

Characterization of begomoviruses associated with cotton leaf curl disease in Pakistan identified distinct begomoviruses. Phylogenetic relationship revealed that these viruses are related with other begomoviruses from the Indian subcontinent. For a description of these viruses please see text.



Characterization of begomoviruses associated with cotton leaf curl disease in Pakistan identified distinct begomoviruses. Phylogenetic relationship revealed that these viruses are related with other begomoviruses from the Indian subcontinent. For a description of these viruses please see text.

Infectious clones of cotton leaf curl virus were obtained from cotton isolates maintained at John Innes Centre. Cotton plants showing typical symptoms of cotton leaf curl disease were collected from a field in the vicinity of Multan, Pakistan, in 1992. The isolate was maintained by grafting infected scion to healthy, susceptible cotton (var. S-12) plants. Further isolates were obtained from Faisalabad, Pakistan and were maintained by whitefly transmission and grafting to healthy cotton (var. S12) and tobacco (*Nicotiana tabacum* cv. Samsun). Clones of CLCuV (PK3) and CLCuKV (pCLCV) were obtained by cloning from restriction endonuclease (*Hind*III and *Xho*I respectively) digested supercoiled (sc) DNA as previously described for *Beet curly top virus* (Briddon *et al.*, 1998). The linearized molecules were ligated into Bluescript vector.

A single clone (pCLCUV004) was produced by polymerase chain reaction (PCR) mediated amplification, from *N. tabacum* graft inoculated with a CLCuD isolate originating from Faisalabad, using primers designed around a unique *Cla*I restriction endonuclease site identified in the published sequence of CLCuAV. (Zhou *et al.*, 1998; EMBL accession no. AJ436992; virion-sense primer [Pak 4F] 5'-TTTATCGATCTGAAATGTGCCCCAGTC-3', complementary-sense primer [Pak 4R] 5'-TTTATCGATGGAAGATCTGCAAGAGGA-3'. The PCR product was ligated into the pGEM-T (Promega) vector. The complete nucleotide sequences of three full-length clones were determined by dideoxy nucleotide chain termination sequencing using the PCR-based "BIG DYE" kit (Perkin Elmer Cetus) and specific internal primers (Genosys).

The complete nucleotide sequences of clones pCLCV002, pCLCV003 and pCLCV004 were determined. The clones were 2750nt, 2750nt and 2745nt in length. All three clones had the typical genome arrangement of the monopartite begomoviruses (equivalent to the DNA A components of bipartite begomoviruses) consisting of 6 positionally and sequence conserved genes; two in the virion-sense (V2 and the coat protein) and four in the complementary-sense (the replication associated protein (Rep), the transcriptional activator protein (TrAP), the replication enhancer protein (Ren) and C4). By both sequence alignments and phylogenetic comparisons (Figure 2) clone pCLCV002 is most similar to CLCuKV (99.6% nucleotide sequence similarity to clone CLCuKV-72b of Zhou *et al.* (1998)). Clone pCLCV003 is most similar to CLCuMV (98.8% nucleotide sequence similarity to clone CLCuMV-26 of Zhou *et al.* (1998)). An alignment of the intergenic regions of selected begomoviruses is shown in Figure. This shows the predicted Rep binding sites (iterons) of clones pCLCV002 and pCLCV003 to be the same as those of clones CLCuKV-72b and CLCuMV-26 respectively. From these findings it is clear that clones pCLCV002 and pCLCV003 represent isolates of the begomovirus species CLCuKV and CLCuMV respectively.

In contrast, clone pCLCV004 shows the highest levels of nucleotide sequence similarity (90.6%) to *Papaya leaf curl virus* (PaLCuV) originating from India (Saxena *et al.*, 1998) and The tree supports the species demarcation and using the 90% identity threshold segregates with this virus in the phylogenetic comparisons. The predicted iteron sequences of clone pCLCUV004 are identical to those of PaLCuV showing this clone to represent a distinct isolate of PaLCuV-Cot.

In addition, the sequence alignments and phylogenetic analyses include the cotton isolated clone pCLCV001 (Briddon *et al.*, 2000). This clone shows the highest levels of sequence

similarity to CLCuMV (97% nucleotide sequence similarity to clone OLC-311 of Zhou *et al*.(1998)). The predicted iteron sequences of pCLCV001 are identical to those of CLCuMV showing this clone to be an isolate of CLCuMV.

The description of species of cotton leaf curl begomoviruses

Cotton leaf curl Alabad virus (Cotton leaf curl virus - Pakistan3; CLCuV-Pk3)	(CLCuAV)	
Cotton leaf curl Alabad virus - [802a]	AJ002455	(CLCuAV-[802a])
Cotton leaf curl Alabad virus - [804a]	AJ002452	(CLCuAV-[804a])
Cotton leaf curl Rajasthan virus		(CLCuRV)
Cotton leaf curl Rhajasthan virus	AF363011	(CLCuRV)
Cotton leaf curl Kokhran virus (Cotton leaf curl virus - Pakistan2; CLCuV-Pk2) (Pakistani cotton leaf curl virus)		(CLCuKV)
Cotton leaf curl Kokhran virus - [72b]	AJ002448	(CLCuKV-[72b])
Cotton leaf curl Kokhran virus - [806b]	AJ002449	(CLCuKV-[806b])
Cotton leaf curl Kokhran virus - [Faisalabad1] (Cotton leaf curl virus - Pakistan2 [Faisalabad1]; CLCuV-PK2[Fai1])	Aftab	(CLCuKV-[Fai1])
Cotton leaf curl Multan virus (Cotton leaf curl virus - Pakistan1; CLCuV-Pk1)		(CLCuMV)
Cotton leaf curl Multan virus – [26]	AJ002458	(CLCuMV-[26])
Cotton leaf curl Multan virus - [62]	AJ002447	(CLCuMV-[62])
Cotton leaf curl Multan virus – [Faisalabad1] (Cotton leaf curl virus - Pakistan1 [Faisalabad1]; CLCuV-PK1[Fai1])	X98995	(CLCuMV-[Fai1])
Cotton leaf curl Multan virus – [Faisalabad2] (Cotton leaf curl virus - Pakistan1 [Faisalabad2]; CLCuV-PK1[Fai2])	Aftab	(CLCuMV-[Fai2])
Cotton leaf curl Multan virus – [Faisalabad3]	AJ132430 (CLCuMV-[Ok])	(CLCuMV-[Fai2]) (CLCuMV-[Mul])

1.4 Begomoviruses associated with cotton leaf curl disease are monopartite but induce symptoms atypical of cotton leaf curl disease

The majority of begomoviruses comprises two genomic components called DNA A and DNA B. A number of strategies were employed in an effort to identify a second genomic component (DNA B). Two pairs of primers and a pair designed to DNA B components of Old World begomoviruses, which have successfully amplified the DNA B components of ACMV and *Asystasia* golden mosaic designed to conserved regions of the DNA B genomic components of begomoviruses amplified products from diseased cotton, when used with low annealing temperatures during PCR. Sequence analysis however did not show the presence of products corresponding to a DNA B component. A pair of primers designed to an area within the non-coding, intergenic region of the genome of CLCuV (equivalent to the common region of bipartite begomoviruses) and which would be expected to amplify both genomic components (if present), amplified only the DNA A component as judged by restriction mapping of the PCR products. Similarly a single primer, designed to the repeated sequence in the stem of the conserved (between geminiviruses) stem loop structure only amplified the DNA A component. Analysis of viral replicative form DNA purified from cotton and digested with various restriction enzymes revealed only a single species of molecules with a length of approximately 2800 bp.

As well as the full-length genomic component of CLCuV, PCR amplification with the intergenic region primers produced a band of approximately half genomic size. Cloning and sequencing of molecules from this band revealed it to consist of population of so called "defective DNA". These molecules consist of varying amounts of CLCuV DNA (always maintaining the major part of the intergenic region, particularly the region with the nonanucleotide motif) with deletions, rearrangements, inversions and duplications, as well as insertions of sequences of unknown origin. Liu *et al*. for CLCuV, following isolation from tobacco, have previously described these molecules. Several molecules were isolated and found to be similar to those presented earlier. One molecule was isolated which had no counterpart in the work of Liu *et al*. The sequence of this molecule is available in the EMBL, DDJB and GenBank databases under accession number AJ242974. This molecule consisted of 1346bp and contained the intergenic region of CLCuV and a large insertion (approximately 1000bp) of indeterminate origin. The techniques used to identify a potential DNA B genomic component included digestion of purified replicative form DNA with various restriction endonucleases to identify two populations of DNA molecules (DNA A and a putative DNA B), each with an approximately length of (2700bp) and PCR based methods. The PCR assays were conducted with primers designed to a conserved region of begomovirus

DNA B C1 gene (virion strand primer GGAGATCTAGATTAGCATCCTTATGCGG; complementary strand primer GCCGAATTCAAGGGCACGTAAACTGAAAGTCTA GTC), the DNA B primers, designed to the intergenic region of CLCuV corresponding to the "common region" of bipartite begomoviruses (virion strand primer CCGTAACGCGTAATTTGAGAAATCATTTCAAAATCC; complementary strand primer TTTAAACGCGTATGCCATTTGGGGTACACCTATATA) and a single primer designed to the repeated sequence (for CLCuV) of the stem of a possible stem loop structure forming the origin of replication for geminiviruses (TATTACCGGATGGCCG). On the basis of these results we concluded that begomoviruses associated with cotton leaf curl disease are monopartite.

The clones of CLCuV were systemically infectious to both *Nicotiana benthamiana* and cotton. Infected plants did not exhibit symptoms characteristic of cotton leaf curl disease, producing mild leaf curling, yellowing and some stunting. These findings suggest that the begomovirus, CLCuV, is not or not the sole cause of cotton leaf curl disease.

Constructs for *Agrobacterium*-mediated inoculation were produced by cloning an approximately 1300bp *Hind* III, *Bam* HI fragment of clone pCLCUV001 into the binary vector pBin19. The full-length *Hin*d III insert of pCLCUV001 was then ligated into the unique *Hin*d III restriction site of the pBin19 construct, to yield a partial repeat of the viral genome. This construct was transferred to *Agrobacterium tumefaciens* strain $C58^{nal}$ and used to inoculate plants. Symptoms of infection typically took up to six weeks to appear for *N. benthamiana* and were mild yellowing and downward curling of the upper leaves. Inoculated cotton plants did not show any symptoms of infection, other than possibly some stunting, although systemic infection could be detected by hybridization on Southern blots. Levels of viral DNA products in *N. benthamiana* infected with the cloned CLCuV were higher than those present in *N. benthamiana* plants infected with cotton leaf curl disease by grafting which showed full leaf curl disease symptoms. For cotton plants, however, viral nucleic acid levels were higher in cotton leaf curl diseased plants than in plants infected with the cloned virus.

1.5 Distinct begomoviruses associated with cotton leaf curl disease are found in multiple infections

Primers specific for four begomovirus species associated with CLCuD were used to assess the prevalence of multiple infections in field isolated cotton. A specific primer pair for the

PCR-mediated detection of CLCuV (Pak-1-PK3) was designed to the sequence at the overlap of the Rep and TrAP genes (virion sense primer [C1Pak1F]; 5'-GCTTTCCATGGGTTGTAGTTGAACTGGAT-3') and the 5' end of the Rep gene (complementary-sense primer [C1Pak1R];5'-AAGTACCATGGCCTCCCCCCAAACGTT TTAA-3'). These primers amplify an approximately 1300bp fragment of the virus. Primers for the detection of CLCuAV were designed to the intergenic region (CLCuV Pak-4R, complementary-sense; 5'TTAAGCTTTGAGTCTGCATCGTCGTGCG3') and the 5' end of the C3 gene (VBTB, virion-sense; 5'-TGATGAGTTCCCTGTGCGTGAATCCATGGTTGT-3'). The design of a multiplex PCR procedure for the specific detection of CLCuV (Pak-1-PK3) and CLCuKV (Pak3-pCLCV) was reported earlier (Mansoor *et al.*, 1999). Briefly, sequences conserved between these two species were identified in the *rep* gene to which a primer was designed (CLCuV-F, virion-sense; 5'-CTGTCTAGATTTGCATTTAAA TTATGAAATTG-3'). Primers (PCL2 and CLCuVPk1-C442) specific for CLCuKV (Pak3-pCLCV) and CLCuV (Pak-1-PK3) were designed to the 5' end of the rep gene (complementary-sense; 5'-CATGCCTCCAAAGCGGAACGGTATTTATT-3') and the 5' end of the C4 gene (complementary-sense; 5'-CGACCATGGGAGCCCTCATCTCCATGTGC-3'). These primers amplify fragments of approximately 510bp and 360bp respectively. A CLCuMV (Pak2-H65) specific primer was designed to the intergenic region 5' of the hairpin in the complementary sense (CR Pak2R, complementary-sense; 5'-TGGAGCGTGA GGATTTTGAAAT GATTTCTCA-3') and was used in combination with CLCuV-F.

Samples of cotton showing the typical symptoms of CLCuD were collected from selected sites across all cotton growing regions of Pakistan and subjected to PCR analysis using primers specific for four begomovirus species. Of 19 cotton samples analyzed, 15 showed the presence of more than one begomovirus species of which two plants showed the presence of all four species tested for. The most commonly detected species was CLCuMV (detected in 17 samples) while CLCuAV was detected in the fewest samples. The primers used to detect CLCuAV would also amplify PaLCuV. Prior to this analysis PaLCuV had not previously been detected in cotton and this cross-reactivity was not considered.

Table 1. PCR-mediated detection of four begomoviruses in CLCuD-affected cotton plants collected from cotton-growing areas of Pakistan

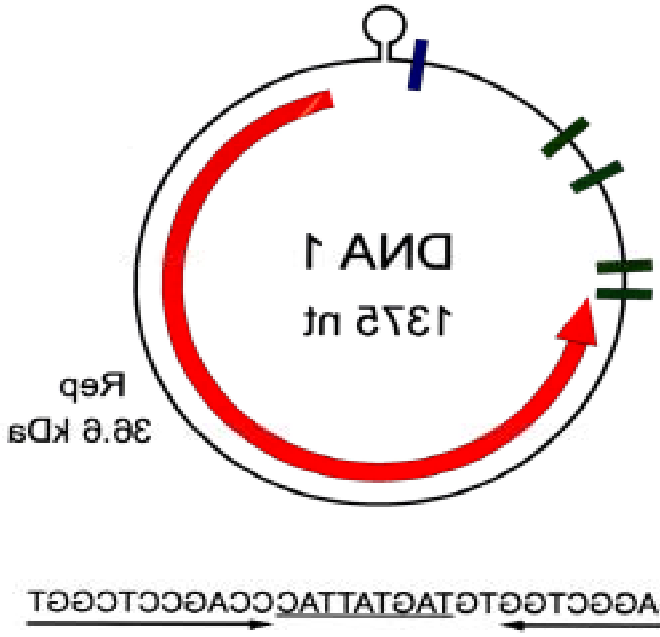
		Species detected			
Location	**Year**	**CLCuMV (Pak-1-PK3)**	**CLCuMV Pak2-H65**	**CLCuKV Pak3-pCLCV**	**CLCuAV Pak4-PK4**
Multan	**1992**	+	+	-	-
Faisalabad	**1993**	+	+	+	-
Chichawatni	**1996**	-	-	+	-
Mianchannu	**1996**	+	-	-	-
Rahim Yar Khan	**1996**	+	+	-	-
Bahawalpur	**1996**	+	+	-	-
Faisalabad	**1996**	-	+	-	-
Dera Ghazi Khan	**1997**	+	+	+	+
Dera Ghazi Khan	**1997**	+	+	+	-
Rahim Yar Khan	**1997**	+	+	-	-
Rahim Yar Khan	**1997**	+	+	-	-
Multan	**1997**	+	+	+	-
Multan	**1997**	+	-	+	+
Dera Ghazi Khan	**1997**	+	+	+	+
Multan	**1997**	+	+	-	+
Faisalabad	**1997**	+	-	-	-
Multan	**1997**	+	+	+	-
Multan	**1997**	+	+	+	-
Ghotki	**1997**	+	-	-	+
Sukkur	**1997**	+	-	-	+
Rahim Yar Khan	**1998**	-	+	-	-
Faisalabad	**1998**	+	+	+	-
Multan	**1998**	+	+	+	+

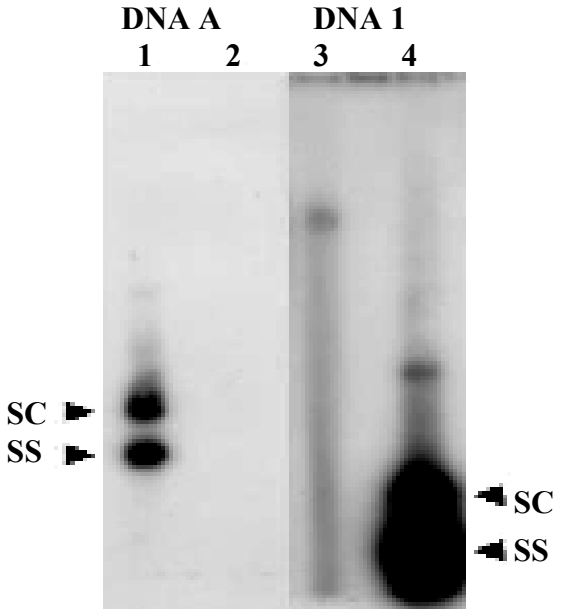
1.6 Identification of a novel circular single stranded DNA associated with cotton leaf curl disease in Pakistan

The fact that symptom severity of the disease correlated with the level of CLCuV DNA in susceptible cotton genotype suggested that a whitefly-transmitted geminivirus is the causative agent of the disease. However, although cloned copies of the genomic component cause a systemic infection when reintroduced into *Nicotiana benthamiana* and cotton, symptoms are delayed, extremely mild, and not typical of cotton leaf curl disease. All attempts to isolate a genomic component equivalent to DNA B were unsuccessful. Attempts were made to identify additional DNA components that were involved in symptom development. Virus particles were purified by sucrose gradient centrifugation from cotton plants growing in the field and that were showing typical leaf curl disease symptoms. Fractions were analyzed for the presence of characteristic geminivirus particles, and DNA was isolated from each fraction and analyzed by agarose gel electrophoresis. The DNA associated with fractions containing virus contained two DNA species that were present at similar levels. Of these, only the slower migrating DNA hybridized efficiently with a full-length CLCuV DNA probe, whereas a weak signal was associated with the faster migrating DNA that was referred to as DNA 1.

To characterize DNA 1, viral supercoiled (sc) DNA was isolated from symptomatic cotton plants and purified by CsCl density gradient centrifugation. The DNA was cloned after digestion with BamHI. Clones were screened for viral DNA using probes made from either scDNA or full-length CLCuV DNA. Clones that were detected using the scDNA probe, but not the CLCuV DNA probe, were selected for further analysis. Sequence analysis of one such clone, pBS-CLCV1, showed that its 1.4-kb insert shared no homology with CLCuV DNA or to DNA B sequences of other begomoviruses, indicating that it was not derived from a defective genomic component.

To verify that DNA 1 was associated with geminivirus particles, as suggested from the sucrose gradient centrifugation data, virus was isolated from infected *N. benthamiana* extracts by immunotrapping using poly-clonal antiserum raised against African cassava mosaic virus (ACMV) coat protein. The coat protein is the most highly conserved begomovirus gene product, and ACMV coat protein polyclonal antiserum is known to cross-react with other begomoviruses. The ACMV antiserum used in this experiment specifically detects CLCuV coat protein on Western blots of diseased cotton extracts. To avoid false positive results arising from the amplification of small quantities of DNA adsorbed onto the surface of the plastic tubes, Southern blot analysis rather than PCR amplification of viral DNA fragments was used to analyze immunotrapped DNA. Such an analysis showed the presence of DNA 1 associated with immunotrapped particles. The DNA migrated to the approximate position of ssDNA extracted from a symptomatic cotton plant. This confirmed that a single-stranded form of DNA 1 was immunotrapped by the ACMV coat protein antiserum. The sequences of the pGEM-CLCV1 and pBS-CLCV1inserts were established (GenBank Accession numbersAJ132344 and AJ132345) and shown to be 1375 and 1376nucleotides in length, respectively, and closely related to Comparison with database nucleotide sequences indicated that DNA 1 has between 43 and 53% identity to a genomic component of members of the *Nanoviridae* family of plant DNA viruses and is most closely related to milk vetch dwarf virus (MVDV). The most highly conserved nucleotide sequences are those encompassing a virion-sense open reading frame (ORF) with the capacity to encode a protein of 36.6 kDa. This putative protein is referred to as the replication associated protein





Identification of a nanovirus like component DNA 1, the first report of a functional DNA satellite found associated with cotton leaf curl disease in Pakistan. Top; the genome organization of DNA 1 shows a single gene for the rep protein. Bottom; DNA 1 replicate independently, DNA1 was inoculated either independently on with DNA A of cotton leaf curl disease. The lanes 1 and 2 were probed with DNA A while lanes 3 and 4 were probed with DNA 1. The detection of replicative DNA forms in lane 4 where only DNA1 was inoculated showed that DNA 1 replicates independent of DNA A.

There are 57 nucleotide differences between the Rep ORFs of pGEM-CLCV1 and pBS-CLCV1, resulting in 11 amino acid to as the replication associated protein (Rep) because it shows strong homology (34±47% identity, 57±64% similarity) to nanovirus changes. The (Rep) as it shows strong homology (34-47 % identity, 57-64 % similarity) to nanovirus reps of which MVDV Rep is the most closely related. DNA 1 Rep contains consensus motifs that are characteristic of proteins involved in rolling circle replication, including a tyrosine residue at position 90 that may be involved in binding to nascent DNA during the initiation of rolling circle replication and a consensus NTP-binding motif, GGEGKS (amino acids 188±193), that occurs in nanovirus and geminivirus Reps. Other features of DNA 1 include the sequence TAGTATTAC located within a potential stem-loop structure (nucleotides 1357±1376/1±13; numbering according to the pBS-CLCV1 insert). This sequence is identical to that found within stem loop structures of many nanovirus DNA components (Sano *et al*., 1998 and references therein), and closely resembles the TAATATTAC sequence that is found within similar structures in geminivirus DNA components. The virion-sense DNA strand is nicked within this sequence during the initiation of rolling circle replication of both nanoviruses and geminiviruses. Of note, the sequence CCTCGGTTCCCTC (nucleotides 7±19) partially overlaps the stem-loop sequence and is repeated immediately downstream (nucleotides 21±33; the insert of pGEM-CLCV1 contains two mismatches in this downstream sequence). Such a reiterated sequence may participate in replication or the control of Rep expression. A consensus TATA box occurs upstream of the Rep ORF although, interestingly, it is located on the opposite side of the stem-loop sequence. Several consensus transcript polyadenylation signals (AATAAA) occur downstream of the ORF. However, an ATTGTA motif (nucleotides 996±1000) that is frequently associated with functional polyadenylation signals, including that of the nanovirus coconut foliar decay virus (CFDV), is suitably positioned only upstream of the first motif, suggesting that one of the first two adjacent motifs (nucleotides 1020±1025 and 1032±1037) participates in transcript polyadenylation. Also, the proposed noncoding region contains two noticeable A-rich sequences [nucleotides 1117±1156 (73%) and 1195±1226 (81%)] flanking imperfect inverted repeat sequences (nucleotides 1162±1178 and 1180±1194). An increase in the size of the noncoding region between the 39 terminus of the Rep ORF and the stem-loop sequence largely accounts for the difference in size of DNA 1 and the majority of nanovirus components that are between 1000 and 1100 nucleotides in length. DNA 1 can replicate autonomously. A partial repeat of CLCuV DNA (pCLCV-Pak1; manuscript in preparation) or a dimer of DNA 1 (pGEM-CLCV1d) were introduced biolistically into tobacco leaf explants together with an expression cassette containing the uidA gene. Full-length circular copies of the genomic DNAs are produced from the cloned DNA either by homologous recombination between the repeat sequences or by a replicative mechanism if two copies of the origin of replication are present. Leaf explants harvested 24 h after inoculation showed blue spots when stained for uidA activity (data not shown), confirming the delivery of the inoculum into viable cells. Southern blot analysis showed that ssDNA and scDNA forms of both CLCuV DNA and DNA 1 are produced de novo in the inoculated tissue, indicating that both DNAs can replicate autonomously.

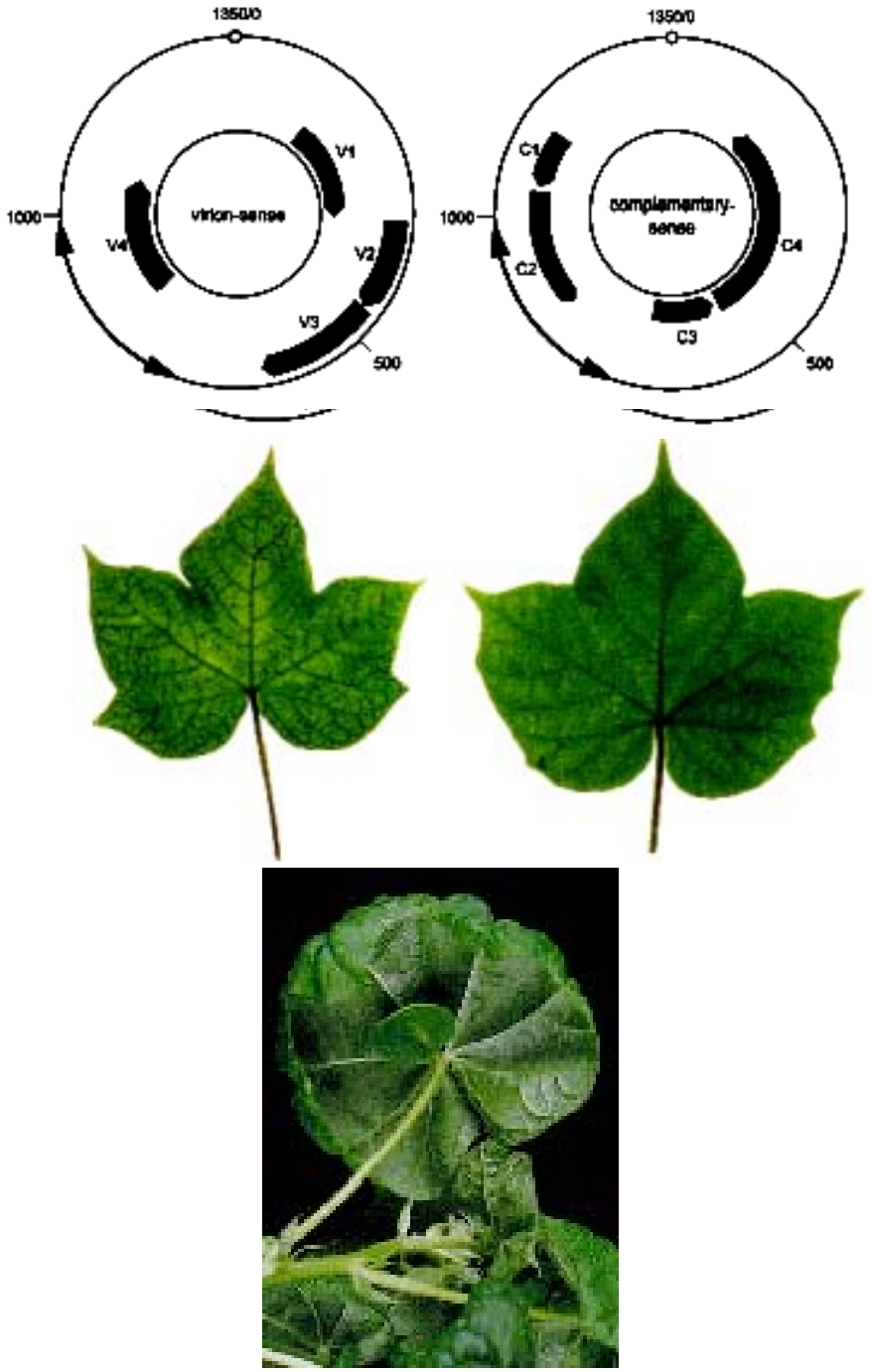
DNA 1 (1375 and 1376 nucleotides in pGEM-CLCV1 and pBS-CLCV1, respectively) is almost exactly half the size of CLCuV DNA (;2750 nucleotides; Zhou *et al*., 1998) but shows no significant homology to either the CLCuV DNA or its defective forms, which are also approximately half the size of the genomic component (Liu *et al*., 1998). The defective

DNAs are produced by deletion and rearrangement of the genomic DNA, but they all retain intergenic region sequences containing cis elements necessary for replication in trans by the Rep protein encoded by the full-length component. They are produced de novo during chronic infection of laboratory hosts such as tobacco and tomato (Liu *et al*., 1998), and occur in cotton plants infected in the field.

DNA 1 is unrelated to other geminivirus DNAs and to the satellite DNA associated with TLCV (Dry *et al*., 1997), which is significantly smaller (682 nucleotides) and depends on the geminivirus for its replication. However, it shows a close relationship to some genomic components of members of the *Nanoviridae*, specifically those that encode the Rep protein. It has the capacity to encode a 36.6-kDa Rep protein that is highly homologous to counterparts encoded by nanoviruses. Consistent with this observation, we have demonstrated that DNA 1 can replicate autonomously in tobacco to produce both single- and double-stranded DNA forms. DNA 1 differs from nanovirus components in two main respects. Firstly, it is whitefly-transmitted, whereas all known nanoviruses are transmitted by either aphids or planthoppers (Magee,1940; Grylls and Butler, 1959; Julia, 1982). Secondly, it is significantly larger than the multiple components of banana bunchy top virus (BBTV), faba bean necrotic yellows virus (FBNYV), subterranean clover stunt virus (SCSV), and MVDV, which are all between 1000 and 1100 nucleotides in length (Boevink *et al*., 1995; Burns *et al*.,1995; Katul *et al*., 1997, 1998; Sano *et al*., 1998). The exception is the genomic component of CFDV that is only slightly smaller at 1291 nucleotides (Rohde *et al*., 1990). Interestingly, CFDV is planthopper-transmitted while the other nanoviruses are aphid-transmitted, suggesting a correlation between genome size and encapsidation constraints resulting from adaptation to different insect vectors. This, coupled with the fact that only a single CFDV genomic component has been isolated (Rohde *et al*., 1990), raises the interesting possibility that it, too, has an associated geminivirus component. Defective DNAs associated with begomoviruses are maintained at approximately half the size of the genomic components, presumably to facilitate encapsidation and whitefly transmission. Hence DNA 1 may have adapted its size to allow encapsidation within If this proves to be the case, our results support the hypothesis that the size of the DNA rather than its sequence is important for encapsidation. Whitefly-transmitted diseases similar to cotton leaf curl disease were reported for tobacco, hibiscus, and other plants long before they were associated with cotton (Pal and Tandon, 1937; Pruthi and Samuel, 1942). This suggests that the nanovirus-like component is indigenous to Pakistan rather than being a recent import, but the disease has only recently had a major impact due to intensive cultivation of susceptible cotton cultivars coupled with the emergence of insecticide-resistant whiteflies resulting from extensive spraying regimes to control cotton leaf curl disease.

1.7 Molecular characterization of DNA β, a DNA satellite essentially required for causing cotton leaf curl disease symptoms

Further search for DNA components required for induction of cotton leaf curl disease resulted in identification of a single-stranded DNA molecule approximately 1350 nucleotides in length which, when co-inoculated with the begomovirus to cotton, induces symptoms typical of CLCuD including vein swelling, vein darkening, leaf curling and enations. This molecule (termed DNA β) requires the begomovirus for replication and encapsidation. The CLCuV/DNA 1/DNA β complex, together with a similar complex previously identified in



DNA ß is essential for the development of typical symptoms of cotton leaf curl disease. Top; the genome organization of CLCuV ß DNA reveals small open reading frames arranged bidirectionally. Centre; co-inoculation of DNA A and DNA ß results in enations, typical of cotton leaf curl disease. Bottom; leaf like enations developed on plants inoculated with DNA A and DNA ß on cotton.

Ageratum conyzoides, represent members of an entirely new type of infectious, disease causing agents.

The discovery of DNA β was made by designing oligonucleotide primers based on DNA sequence of a 1346 nucleotide recombinant molecule (EMBL accession number AJ242974; which is referred to as cotton leaf curl recombinant 01 (CLCR01)), consisting of the intergenic region of CLCuV with the remaining sequence of indeterminate origin, in diseased cotton (Briddon *et al*., 2000). A pair of non-overlapping, abutting primers (V4570/V4571) were designed to the sequence of unknown origin of CLCR01. These primers span a unique *Sal*I restriction endonuclease site present in the sequence of CLCR01. PCR amplification with nucleic acids extracted from CLCuD infected cotton originating from Pakistan utilizing primer pair V4570/V4571 amplified an approximately 1350bp fragment. The product of this PCR was cloned into pGEM T-Easy vector yielding approximately 10 clones. Restriction endonuclease analysis of the 10 clones showed all to have the same restriction pattern and to be different from CLCR01, possibly indicating that CLCR01 is a minor component of the DNA molecules in the plants analyzed. A single clone (CLCβ01) was chosen at random for further analysis. An additional primer pair (Beta01/Beta02), designed to sequence of CLCβ01 which is not present in CLCR01, also amplified an approximately 1350bp fragment. The amplified DNA was cloned into pGEM T-Easy vector yielding 10 clones and a single clone (CLCβ02) was selected at random for further analysis.

The complete nucleotide sequences of clones CLCβ01 and CLCβ02 were determined in both orientations and consist of 1351 and 1349 nucleotides, respectively. These sequences are available in the EMBL, DDJB and GenBank databases under accession numbers AJ292769 and AJ298903, respectively. The component represented by clones CLCβ01 and CLCβ02 will henceforth be referred to as CLCuD DNA β, in line with the nomenclature used for a similar molecule isolated from *Ageratum conyzoides*. Nucleotide numbering for CLCuD 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 DNA replication. This sequence differs from that of DNA 1 and the majority of nanovirus components, from which DNA 1 is proposed to have evolved (Mansoor *et al*., 1999) which have a TAGTATTAC loop sequence. More recently some DNA 1 molecules have been identified with the geminivirus-like (TAATATTAC) loop sequence**.** The sequences of CLCuV DNA β contain an A-rich region between nucleotides 766 and 984 and show 96% overall nucleotide sequence similarity with the majority of the nucleotide changes occurring just downstream of the nonanucleotide motif in a putative non-coding region.

Analysis of the sequences of clones CLCβ01 and CLCβ02 identified four open reading frames (ORFs) with a predicted coding capacity (beginning from an in-frame methionine codon) above 4 kDa in each strand. These ORFs are denoted as being encoded on the virion (V) or complementary (C) strands. Of these ORFs, only two (V3 and C4) are predicted to be functional coding regions by the program TESTCODE. This program predicts functional ORFs based on codon usage.

A search of the EMBL nucleotide sequence databases identified the presence of a fragment of DNA β amongst the sequences of a number of defective CLCuV molecules isolated by Liu *et al.* (1998a) from *Nicotiana tabacum* experimentally infected with CLCuD. The recombinant molecule (EMBL accession number AJ222705; which is referred to as CLCR02) consists of mainly CLCuV DNA, but has a stretch of approximately 211 nucleotides of DNA β (nucleotides 555 to 354, corresponding to nucleotides 804 to 1004 of DNA β clone CLCβ01). This finding demonstrates that DNA β is not unique to the CLCuD isolate examined as part of this study. The sequence contained in the database entry AJ222705 is in the complementary-sense and lacks the nanonucleotide motif due to the design of the primers. Comparisons of the sequence of CLCuD DNA β to the DNA β molecule associated with *Ageratum* yellow vein disease (AYVD) shows relatively low sequence identity (53%) at the nucleotide level. A dot plot comparison shows the sequences to have the highest levels of similarity over a stretch of approximately 86 nucleotides upstream of the predicted hairpin structure. Alignments (conducted on the basis of position) of the predicted amino acid sequences encoded by the ORFs of CLCuD DNA β and AYVD DNA β identifies only limited amino acid sequence similarity. The highest levels of similarity (57%) were detected between ORF C1 and the positional equivalent in AYVD DNA β. The relatively high levels of sequence conservation identified here may be due to the fact that these ORFs overlap the A-rich region and are consequently lysine rich, probably indicating that they are not translated. CLCuD DNA β ORF C4 is positionally analogous to the ORF predicted by TESTCODE to be functional for AYVD DNA β (Saunders *et al*., 2000).

Table 2. Features of CLCuD DNA β open reading frames and comparison to those of AYVD DNA β

ORF	No. of amino acids (CLCβ01/CLCβ02) [% amino acid similarity*]	Predicted molecular weight (kDa) (CLCβ01/CLCβ02)	Similarity to AYVD DNA β ORFs (%)
V1	61/81 [98.4]	7.5/9.8	-
V2	39/40 [100]	4.3/4.4	56/56
V3	49/49 [93.8]	5.7/5.7	43/39
V4	65/65 [93.8]	7.5/7.5	25/23
C1	61/41 [92.7]	7.6/5.1	57/52
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	46/46

*Percentage amino acid similarity (Schwartz and Dayhoff, 1978) between predicted products of clones CLCβ01 and CLCβ02.

Both CLCβ01 and CLCβ02 were found to be infectious to cotton following co-inoculation with CLCuV. Subsequently all inoculations were conducted using CLCβ02. Following biolistic inoculation with CLCuV and DNA β, symptoms in cotton appeared within 18 to 21 days post inoculation. The efficiency of infection to cotton following biolistic inoculation was 16% (10 plants infected of 60 inoculated, the result of 4 independent experiments). Inoculation of *Nicotiana benthamiana* with CLCuV and DNA β was less efficient (10%, 5 plants infected of 50 inoculated, the result of 5 independent experiments) with symptoms appearing within 10 to 15 days of inoculation. This compares to 6 weeks for symptoms to appear for inoculation of *N. benthamiana* with just CLCuV (Briddon *et al*., 2000).

Symptoms induced in cotton by co-infection with CLCuV and CLCuD DNA β are illustrated in Figure 4. The first leaf to develop symptoms exhibited deep downward cupping 2 days prior to appearance of full foliar symptoms. The symptoms exhibited were downward curling of the leaf margins, vein swelling, vein darkening and enations which eventually developed into leaf like structures on the main veins on the undersides of symptomatic leaves. On a small number of plants upward curling of the leaf margins was evident. These symptoms are typical of the symptoms of CLCuD seen in the field and contrast with the symptoms of infection of cotton with just CLCuV which consist of mild stunting and yellowing. *N. benthamiana* plants co-infected with CLCuV and DNA β exhibit downward leaf curling, leaf crumpling and stunting indistinguishable from those of infection of *N. benthamiana* by AYVV and AYVD DNA β. Although infected *N. benthamiana* plants flowered they set no seed. *N. benthamiana* plants infected with just CLCuV exhibited mild stunting and yellowing (Briddon *et al.*, 2000).

Analysis of both cotton and *N. benthamiana* plants inoculated with only CLCuD DNA β, by Southern blot hybridization, was unable to show the spread of this molecule in plants, nor its replication in inoculated leaves (20 plants analyzed for each species). Southern blot analysis of cotton plants infected with both CLCuV and DNA β, compared to plants infected with a wild isolate of CLCuD was carried out. Typical DNA forms indicative of replication are evident for both CLCuV and DNA β, indicating efficient *trans*-replication and/or *trans*-movement of DNA β by CLCuV in cotton. For both probes some DNA forms (marked as sub-genomic; sg), migrating faster than the unit length molecules, are detected. These bands most probably represent defective forms of CLCuV and DNA β produced by errors in replication or by recombination. Comparison of cotton plants co-infected with CLCuV and CLCuD DNA β to cotton plants infected with only CLCuV show an amplification of CLCuV in the presence of DNA β. In co-infected plants CLCuV shows considerably higher levels of virus-specific sub-genomic DNA molecules.

Southern blot analysis of cotton plants co-inoculated with CLCuV and DNA β, which did not become symptomatic, identified a small number of plants in which CLCuV replication was evident without DNA β (3 out of 26 plants analyzed). The ability of CLCuV to infect cotton, but inducing only very mild symptoms, has been noted previously (Briddon *et al*., 2000). PCR amplification with primer pair V4570/4571 from nucleic acid extracts produced from the inoculated leaves of both *N. benthamiana* and cotton, which had been inoculated with CLCβ02, did not produce a DNA β specific product indicating that CLCuD DNA β is not capable of autonomous replication (results not shown) but instead relies on CLCuV for replication functions. Similar experiments with leaves inoculated with both CLCuV and CLCβ02 yielded the expected full-length DNA β product.

The presence in both CLCuD DNA β and AYVD DNA β of an A rich region, which may have originated by sequence duplication to satisfy size requirement for encapsidation and/or virus

movement (Saunders *et al*., 2000), may indicate that the progenitor molecule, from which DNA β has evolved, was less than its present half geminivirus genomic component size (approx 1350 nucleotides). For the DNA 1 molecules identified for CLCuD and AYVD, a similar size increase from that of nanovirus components (from 1000 nucleotides to approx. 1350 nucleotides) has been proposed (Mansoor *et al*., 1999; Saunders and Stanley, 1999). It is possible that DNA β molecules originated with nanoviruses, although no molecules with similarity to DNA β have been identified in nanovirus infected plants and the range of symptoms induced by nanoviruses do not include those exhibited by CLCuD and AYVD.

The two molecules, which make up the genome of bipartite begomoviruses, such as *African cassava mosaic virus* (Stanley and Gay, 1983), share a so called a common region of approximately 200 nucleotides. This region contains *cis*-acting elements for replication and gene expression including the nonanucleotide motif (TAATATTAC), which contains the nick site for initiation of rolling circle replication (Laufs *et al*., 1995), and the Rep binding site (Fontes *et al*., 1994; Argüello-Astorga *et al*., 1994). CLCuV and CLCuD DNA β do not share a common region, although DNA β must have a functional Rep binding site to allow *trans*-replication from CLCuV. It is likely that the conserved region of DNA β molecules is functionally analogous to the common region of begomoviruses. This region contains a nanonucleotide motif and may also contain a cryptic Rep binding site. The high levels of sequence conservation within the common region of such otherwise dissimilar DNA β components may be due to the need to maintain the elements required for *trans*-replication and gene expression, possibly being an adaptation for interaction with disparate begomoviruses.

Recombination between components has been shown to play a major part in the evolution of geminiviruses. Evidence of intermolecular recombination occurring between CLCuV and CLCuD DNA β (resulting in molecules CLCR01 and CLCR02) illustrates the possibility for further adaptation and evolution of the CLCuD complex. It is interesting to note that CLCR01 consists essentially of the sequence of CLCuD DNA β with a geminivirus origin of replication (intergenic region). We have thus far not been able to show any biological activity of CLCR01, other than the ability to be *trans*-replicated by CLCuV (results not shown), but this molecule is not far removed from a begomovirus DNA B-like component with a true common region.

1.8 Relaxed specificity of interaction between DNA A and DNA β

The replication of geminiviruses occurs by a rolling circle mechanism with the origin of replication consisting of a conserved stem loop structure, with the loop sequence TAATATTAC, which is nicked by the virus-encoded replication associated protein (Rep) during the initiation of replication, and a number of upstream repeat sequences ("iterons") which are the sequence specific binding sites for Rep (Fontes *et al*., 1994; Chatterji *et al*., 1999). For bipartite begomoviruses, *trans*-replication of the DNA B component is achieved by virtue of a shared sequence, the "common region", which is present on both components and contains the stem loop structure and iterons. The Rep iteron interaction is very sequence specific (Fontes *et al* 1994; Chatterji *et al*., 1999) such that interspecific replication can occur only if iterons are conserved or very similar (Saunders *et al*., 2002). In contrast to DNA B components, the DNA β components do not contain the iterons of their respective helper begomoviruses suggesting that some other motif(s), cryptic Rep binding sites, act to facility *trans*-replication of these molecules (Briddon *et al*., 2001).

DNA β requires the helper begomovirus for replication, movement in plants and insect transmission, presumably by *trans*-encapsidation in the begomovirus coat protein. The

specificity of interaction between DNA β and distinct begomoviruses associated with the disease was investigated by using three further species of monopartite begomoviruses, isolated from cotton, with distinct Rep binding specificities.

All three species were capable of *trans*-replicating the satellite molecule and inducing CLCuD symptoms in cotton, indicating that the interaction between begomovirus and DNA β is relaxed in comparison to the interaction between DNA A and DNA B components. Only a single DNA β "species" has been detected in CLCuD affected cotton suggesting that this has the capacity to recruit unrelated begomoviruses. A single DNA β component has been identified in CLCuD isolates originating from geographically widely dispersed sites and is capable of interacting with all four begomovirus species to induce leaf curl symptoms in cotton. This is the first demonstration of the association of a single DNA β molecule with multiple begomoviruses and has implications on our understanding of the evolution of such complexes.

Cotton samples showing typical of CLCuD symptoms were collected from various locations in Pakistan. Total DNA was resolved in duplicate agarose gels using TNE buffer and were blotted to Hybond NX (Amersham) membranes. DNA β molecule cloned from cotton and a DNA β molecule originating from leaf curl affected okra originating from Pakistan were radioactively-labeled with ^{32}PdCTP (NEN) by oligo-labeling. Duplicate blots were hybridized with these probes and were washed at high stringency (0.1xSSC, 65^{o}C) for the CLCuD DNA β and okra leaf curl DNA β probes. The cotton β probe detected the presence of DNA β in all samples while okra β probe did not hybridized to any sample. These results showed that only a single species of DNA β is associated with the disease.

Young *N. benthamiana* and cotton (var. S-12) plants at the 3-4 leaf stage were inoculated with cloned viruses and DNA. by a biolistic procedure. Gold particles were coated with excised inserts of full length clones of viruses and DNA β. Following inoculation, cotton plants were maintained in a growth cabinet at 35^{o}C with a 16 hour photoperiod and *N. benthamiana* in an insect proof glasshouse at 25^{o}C with supplementary lighting to give a 16 hour photoperiod. Plants were checked daily for the appearance of symptoms. Cotton seedlings, infected with the cloned components of CLCuD by biolistic inoculation, were placed in perspex cages and exposed to approximately 1000 *Bemisia tabaci*. Following a 48 hour acquisition access period 4 healthy cotton seedlings were placed in the cage and the source plant was removed. After approximately 10 days the plants were removed from the cage, sprayed with insecticide and maintained in an insect proof glasshouse at 25^{o}C with supplementary lighting to give a 16 hour photoperiod.

All three clones were infectious to both the experimental host *N. benthamiana* and to cotton. Symptoms of infection for *N benthamiana* typically appeared within 12 days of inoculation. For all three clones the symptoms were downward leaf curling, swelling of the veins on the undersides of leaves and vein yellowing with some swelling of the veins on the upper leaf surface. The yellowing of veins on the upper leaf surface was more pronounced for infections with pCLCV003 than the other clones. Initially infected *N. benthamiana* plants showed little chlorosis but a bright yellow mottling developed within a week of the first symptoms. These symptoms are very similar to the symptoms reported for infection of *N. benthamiana* with AYVV and AYVD DNA β with the exception that, for the cotton clones, the yellow vein symptoms were more pronounced.

Of the three clones analyzed only pCLCUV002 was transmitted from cotton plants, infected by biolistic inoculation, to healthy test seedlings. Two of four test cotton seedlings showed typical

symptoms of CLCuD approximately three weeks after exposure to viruliferous insects. The symptoms exhibited by these plants were poor with very little leaf curling and vein darkening.

These findings make it clear that the epidemic of CLCuD occurring across Pakistan and western India has a complex genetic structure, involving numerous distinct begomovirus species. The involvement of more than two begomoviruses in a disease epidemic is unique for geminiviruses. Although a number of symptomatically indistinguishable diseases are known to be caused by genetically distinct begomoviruses, these viruses usually have different geographical host ranges. For CLCuD, despite the frequency of dual and multiple infection of cotton plants in the field, no apparent synergism is evident. The severity of symptoms appears more related to the cultivar of cotton infected, the level of whitefly infestation and the age at which plants were infected than to the species/strain of virus associated with the infection. The common denominator of CLCuD across the affected areas of Pakistan and western India appears to be CLCuD DNA β. It is therefore likely that a particular "species" of DNA β is driving the epidemic. This satellite molecule is well adapted to the "exotic" cotton varieties (particular the variety S12 in which the epidemic is believed to have started) under cultivation on the Indian subcontinent and is able to recruit diverse begomoviruses, resulting in the present situation of multiple virus species supporting, apparently, a single satellite "species". The recruitment of other begomoviruses is likely to occur through co-infection of alternate hosts, most probably within the *Malvaceae*, with other begomovirus/DNA β complexes and subsequent transmission back to cotton by the vector whitefly *Bemisia tabaci*. Several malvaceous plant species harbor begomovirus/DNA β complexes, including hollyhock, okra and hibiscus, that could have acted as a source of viruses for the diversity that we now encounter in cotton.

1.9 Universal primers for amplification of DNA β from diverse plants in the Old World

Although the evolutionary origin of DNA β remains uncertain, sequence comparisons of the two DNA β molecules available showed them to have only very limited overall nucleotide sequence identity (53%). However, just upstream of a predicted hairpin structure (with the loop sequence TAATATTAC, a motif that DNA β shares with geminiviruses and, for geminiviruses, shown to be the origin of virion strand DNA replication the DNA β molecules have a conserved region of approx 80 nucleotides. This "conserved sequence" has been suggested to be important in *trans*-replication of DNA β by the begomovirus Rep, possibly containing cryptic Rep binding sites. A pair of abutting oligonucleotide primers (Beta01/Beta02) was designed to the "conserved sequence" of the DNA β molecules. Primers Beta01 and Beta02 are 25 nucleotides in length and non-degenerate, annealing to a region just upstream of the conserved hairpin loop structure. This region is highly conserved between CLCuD DNA β and AYVD DNA β and for all DNA β molecules, which have thus far been cloned and sequenced. The primers introduce a *Kpn*I restriction endonuclease recognition site, which allows recovery of the complete amplified molecule, intact without additions, following cloning into T-vectors. Thus far the introduced *Kpn*I site has been found to be unique for all DNA β molecules cloned. Total nucleic acid extracts were produced by standard, well established methods; typically the CTAB method for cotton and *Ageratum* and the method of Covey and Hull for tobacco and other plant species. Amplification conditions utilized were typically 35 cycles of melting at 94^{o}C for 1 min, annealing at 50^{o}C for 1 min and extension for 1.5 min at 72^{o}C. For some samples a lower annealing temperature (45^{o}C) proved advantageous, although this usually lead to a dramatic increase in non-specific, background

amplification yielding a smear on ethidium bromide stained agarose gels. The results of PCR amplifications with Beta01 and Beta02 from extracts produced from field-infected plants is shown in. PCR with these primers typically produces a major and at 600– 700 bp and a minor band at approx 1350 bp. Amplification with nucleic acid extracts produced from healthy plants produce no products although some nonspecific bands can be seen at lower (below 50^{o}C) annealing temperatures. When used with nucleic acid templates extracted from plants experimentally infected with clones of the helper geminivirus and DNA β, only a single product is detected . The smaller products amplified from field collected material is believed to emanate from defective DNA β molecules (deletion mutants) and molecules which result from recombination between DNA β, DNA 1 and the helper geminivirus. These less than full length molecules, although not detected in young recently experimentally inoculated plants, rapidly build up as the infection progresses and the plant ages. For several diseases including CLCuD and AYVD originating from Pakistan, clones of the approx 1350 bp PCR product, produced with primers Beta01 and Beta 02, have proven infectious to host plants yielding typical disease symptoms when inoculated with clones of the respective helper begomovirus.

1.11 Diversity of DNA β; a satellite molecule associated with some monopartite begomoviruses

DNA β isolated from cotton and ageratum are the first examples of functional satellite in DNA viruses. Like ToLCVsat, these are sequence unrelated to their helper viruses and require their helpers for replication, movement in plants and insect transmission. Unlike the ToLCVsat, DNA β affect the replication of their helper begomoviruses and alter the symptoms induced in some host plants. Satellites are a common feature of a number of RNA viruses. These molecules are classically defined as viruses or nucleic acids that depend on a helper virus for replication, are dispensable for the replication of the helper virus and lack sequence homology to the helper virus genome. Satellites often interfere with the replication of their helper viruses and attenuate symptoms. A small number of satellites, however, exacerbate symptoms and can produce symptoms not associated with the helper virus. There are several diseases in the Old World that are known to be associated with begomoviruses. Characterization of begomoviruses associated with these diseases has identified only a single component related to DNA A but all efforts to clone a DNA component equivalent of DNA B remain unsuccessful. The discovery of DNA β suggested that these diseases might be associated with begomoviruses that essentially required DNA β.

In this study we have assessed the geographical and sequence diversity of the DNA β associated with begomoviruses. We show that diseases associated with DNA β molecules are widespread in the Old World but apparently absent from the New World. The sequences of DNA β molecules analyzed show both a host and geographic segregation. Comparisons of sequences identify conserved features of DNA β molecules, which must play an important role in their function.

The geographic origins of plant materials infected with DNA complexes are summarized in Table 3. For the vast majority of samples total nucleic acids were extracted from field collected leaf samples. The origins of 4 other DNA β molecules are detailed in Table 3 which were not produced as part of this study but which are included in the analysis; two clones isolated from cotton originating from Pakistan (CLCuV-Pak and CLCuV-Pak [Briddon *et al.*, 2001]), a clone isolated from okra (BYVV-Ind, Jose and Usha, unpublished) and a clone isolated from *Ageratum conyzoides* originating from Singapore (Saunders *et al.*, 2000). Total nucleic acids

were extracted from all leaf samples by the CTAB method, with the exception of those from Egypt which were extracted using the Phytopure kit (Amersham), as described by the manufacturer. A pair of "universal primers" (beta01/beta02) for the polymerase chain reaction(PCR)-mediated amplification of DNA β molecules were used. For some isolates (HYVMD and OYVD-Egypt), amplifications with primer pair beta01/beta02 produced products of less than 1300 nucleotides. In these cases the small PCR products were cloned and partially sequenced to allow the design of specific primers for amplification of the whole DNA β component. Clones OYV01-Egy (SB36-1 virion-sense primer GGTACCCACCGGG AAATGATAGTTTCACCG, complementary-sense primer GGTACCGGCCTCAATTCA CGACACGCGCGG) and HYVM01 (u96-2 virion-sense primer TCTAGATAAGTT TTGGTTTTATTGCACTTTG, complementary-sense primer TCTAGAAGGTGG GGCCCACCTTTCCCCCG) were produced in this fashion. Clones OLC02-Egy, OLC03-Egy, OLC02-Egy and HLCr02-Egy were produced by using a primer pair derived from beta01 and beta 02 (prSat1284V [GGTACCACTACGCTACGCAG] and prSat1270C [GGTACCTACCCTCCCAGGGG]. PCR amplifications were conducted as described previously with annealing temperatures ranging from 45 to 50^{o}C. All PCR products were cloned into either pGem-T or pGem-T Easy vectors (Promega).

The infectivity of cloned DNA β molecules to plants was determined by co-inoculation with their respective begomovirus (for CLCuD and related DNA. molecules as detailed in Table 3) or by co-inoculation with *Ageratum* yellow vein virus (AYVV) (Tan *et al.*, 1995). Inoculations were conducted by biolistic inoculation with linearized clones coated on to 1.3 μM gold particles and a hand held biolistics gun. Inoculated *N. benthamiana* plants were maintained in insect proof glasshouses at 25oC with a 16 hour daylight regime. Cotton plants inoculated with cloned components were maintained in a growth cabinet at a constant temperature of 35^{o}C with a 16 hour day length. Plants were monitored daily for the appearance of symptoms.

The complete nucleotide sequences of 26 DNA β molecules were determined. These sequences are available in the DDJB, EMBL and GenBank nucleotide sequence databases under the accession numbers detailed in Table 4. In addition to these, the sequences of four further DNA β molecules were included in this analysis; two originating from cotton leaf curl disease infected cotton from Pakistan (CLC01-Pak and CLC02-Pak; Briddon *et al.*, 2001), one isolated from yellow vein disease affected *Ageratum conyzoides* originating from Singapore (AYV01-Sin; Saunders *et al.*, 2000) and one isolated from yellow vein disease infected okra (BYV01-Ind; Jose and Usha, unpublished [EMBL accession no. AJ308425]).

Of the 30 sequences analyzed, 24 represent potentially full-length DNA β. molecules and have been shown to be infectious to plants, when inoculated with their corresponding begomovirus component, inducing typical disease symptoms in the host from which they were isolated (CLCu01-Pak, CLCu02-Pak; [Briddon, 2001 #1239]. Whereas 3 (AYV01, OYV01-Egy and HLCr01-Egy) are clearly deletion mutants with a size <1000 nucleotides. The nature of the deletions suggests that they are naturally occurring mutants rather than being PCR derived.

Comparisons of the sequences of all DNA β molecules, including both the apparently full-length molecules and those presumed to be deletion mutants, identified 3 absolutely conserved features.

1. A stem loop structure with the loop sequence TAATATTAC. For geminiviruses a similar stem loop structure, also with the loop sequence TAATATTAC, has been shown to be the origin of virion strand DNA replication. A single DNA. molecule

(OLC02-Pak NIB5-2) was identified with the loop sequence TAGTATTAC, common to genomic components of the members of the genus *Nanovirus* and the nanovirus-like DNA 1 molecules shown to be associated with the two begomovirus/DNA. complexes thus far characterized.

2. A region of high sequence homology which was first identified by Briddon *et al*. (Briddon, 2001) as the "conserved region". The primer pair beta01/beta02, with which the majority of the DNA. molecules were amplified, were designed to this region (Briddon, *et al*., 2002). The conserved region is approximately 118 bases in length, includes the nanoanucleotide containing hairpin structure and is typically conserved between all the DNA. molecules analyzed.
3. An A rich region approximately 370 to 420 nucleotides upstream of the "conserved region". The A rich region is typically between 160 and 280 bases in length and has between 57% and 65% A content (the overall A content of DNA β molecules is between 28% and 38%).

Analysis of all presumed full-length DNA. molecules (those of over 1200 bases in length) identified a single, positionally conserved, open reading frame (ORF). This ORF (identified as ORF C1 by Saunders *et al*. (2000) for AYVD DNA. and ORF C4 for CLCuD DNA β by Briddon *et al*. (2001) is encoded on the complementary strand. We shall henceforth refer to this ORF as gene C1. The conserved region is 75 to 123 nucleotides in length and encompasses the nonanucleotide containing hairpin structure.

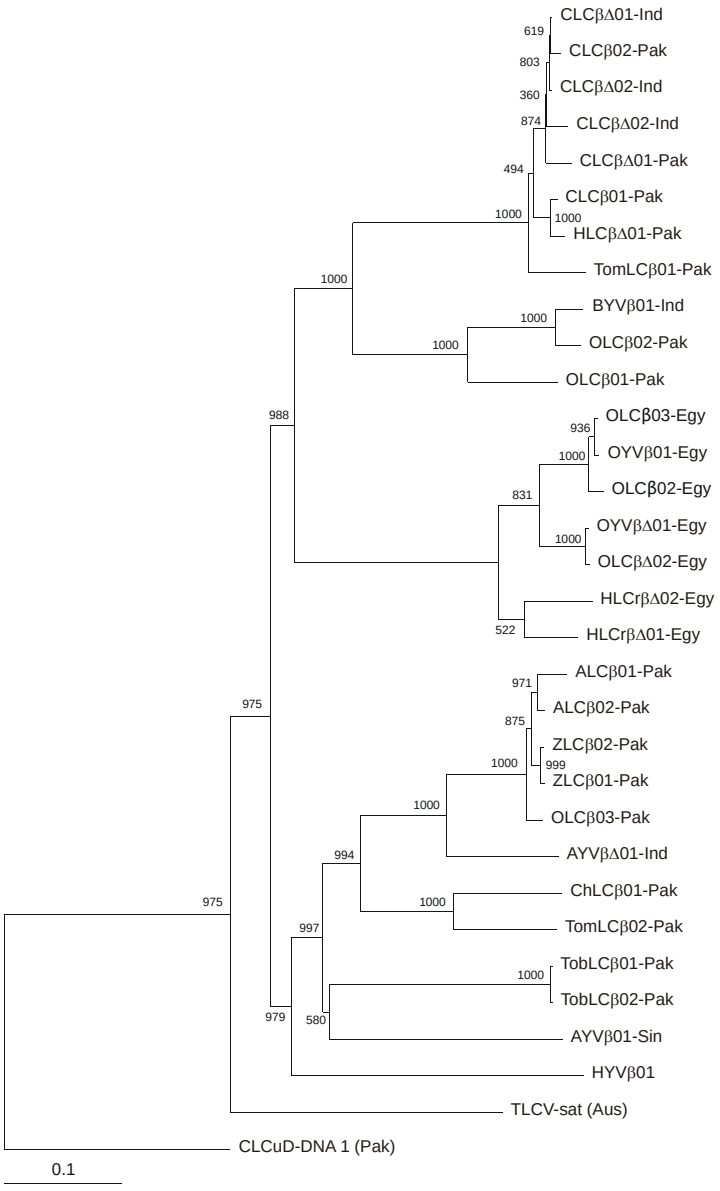
The features of the product predicted to be encoded by gene C1 of the potentially full- length DNA. molecules are summarized in Table 6. An alignment of the predicted amino acid sequences of all gene C1 products is shown in the Figure. The typical length of the product encoded by gene C1 appears to be 118 amino acids. Some gene C1 products have an extended amino terminal leader. However, a conserved methionine suggests that this is start of the core active sequence of the protein. This is substantiated by the fact that u774, with a N-terminal deletion of gene C1 does not induce the typical symptoms of CLCuD in cotton. Despite the low levels of nucleotide sequence conservation between the diverse DNA molecules the alignment highlights clearly conserved amino acid sequence blocks in the gene C1 product.

The Figure details the extent of the sequence deletions for the mutant DNA. molecules which were examined. This analysis show all mutants encountered as part of this study to have sequence deletions downstream of the hairpin structure and encompassing all or part of the gene C1 coding region. For all mutants the hairpin, "conserved region" and A-rich region is maintained. This is strong evidence suggesting that these sequences are essential for *trans*-replication of the DNA. molecule by its begomovirus helper.

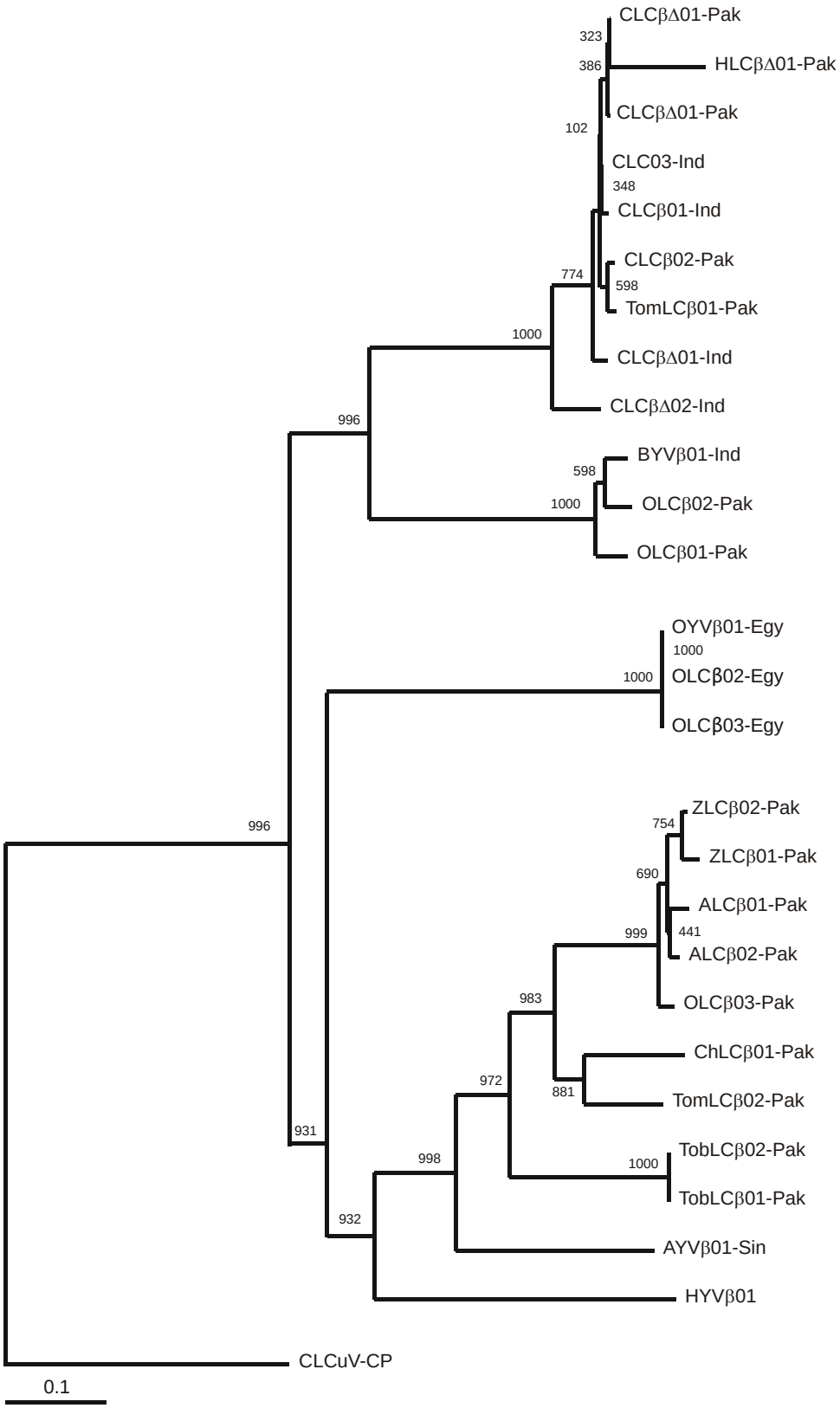
The infectivity of the DNA β molecules was investigated either by co-inoculation with CLCuV to *N. benthamiana* and cotton (*G. hirsutum* cv. S12; for the DNA β molecules isolated from cotton originating on the Indian subcontinent) or by co-inoculation with AYVV to *N. benthamiana* and, for DNA β molecules isolated from *Ageratum conyzoides*. It has previously been shown that AYVV can support the trans-replication of the unrelated DNA β of CLCuD in *N. benthamiana*. The presence of CLCuD DNA. induces a shift in the symptoms of the infection from upward leaf curling (when AYVV infects *N. benthamiana* on its own) to downward leaf curling. A similar change in symptoms is induced in *N. benthamiana* in the presence of AYVV DNA β. Of the several DNA β molecules isolated from cotton, all were *trans*-replicated by CLCuV and induced symptoms typical of CLCuD in cotton and the typical downward leaf curling symptoms of CLCuD in *N. bentamiana.* A single DNA β molecule

isolated from *A. conyzoides* originating from Pakistan was able to induce typical yellow vein symptoms of AYVD when co-inoculated to *A. conyzoides* with AYVV.

Phylogenetic dendrograms resulting from alignments of the complete nucleotide sequences of DNA β molecules (30 sequences of both presumed full-length and mutant, deleted molecules analyzed) and the predicted amino acid sequences of their C1 genes (24 sequences analyzed) are shown in Figures.



Phylogenetic relationship of distinct DNA β molecules cloned from Pakistan and other parts of the world. The data revealed that highly diverse DNA β molecules are associated with monopartite begomoviruses in the old world.



Phylogenetic relationship of diverse DNA β associated with begomoviruses based on βC1 gene, the conserved gene present in all known β molecules.

The tree derived from an alignment of the complete nucleotides sequences divides the DNA β molecules into two "super groups". The first includes molecules isolated from plant species within the family *Malvaceae*. Within this group, DNA β molecules originating from Egypt, consisting of molecules isolated from okra and hollyhock, are distinct from those originating from the Indian sub-continent. The *Malvaceae* isolated DNA βs originating from the Indian sub-continent fall into two groups, those isolated from okra (also called bhendi) and those isolated from either cotton or hibiscus. This latter group also includes a single DNA β molecule isolated from tomato (TomLCV01-Pak).

The second super group includes DNA β molecules originating from diverse plant species. The DNA β molecules originating from the Indian sub-continent isolated from Ageratum and Zinnia (both species within the family *Compositae*) form a single cluster which also includes a molecule isolated from okra (OLC03-Pak) but not the Ageratum isolated DNA β originating from Singapore. Although the DNA β molecules isolated from chilies and tomato co-segregate in this analysis they show only a distant relationship to the DNA β molecules isolated from a third Solanaceous host tobacco. Instead the tobacco isolated molecules show a more distant relationship to the satellite molecule, isolated from tomato, originating from Australia, now believed to represent a mutant, deleted DNA β.

The tree derived from the predicted amino acid sequences of the C1 gene gives the same overall clustering consisting of two major groups; the Indian sub-continent, Malvaceae derived molecules and a more diverse group. The major difference to the nucleotide tree is in the relationship of the molecules originating from Egypt, all of which originate from okra. These sequences are very distinct and basal to the diverse group rather than the other *Malvaceae* derived molecules. On this tree also the C1 genes of the DNA β molecules isolated from solanaceous hosts show a somewhat closer relationship.

We have assessed isolates of a number of diseases for the presence of DNA β satellites. The majority of the diseases examined have previously been assumed to be caused by begomoviruses, due to their transmissibility by the whitefly vector of begomoviruses (*Bemisia tabaci*) and include *Ageratum* leaf curl disease, *Zinnia* leaf curl disease, chili (also know as pepper) leaf curl disease and tomato leaf curl disease. Further diseases examined have previously been shown to be associated with begomoviruses including honeysuckle yellow vein mosaic disease, *Hibiscus* leaf curl disease, okra yellow vein disease and okra leaf curl disease. Hollyhock leaf crumple disease was shown by Abdel-Salam *et al*. (Abdel-Salam, 1998) to be associated with a whitefly- transmitted geminivirus which they named hollyhock leaf crumple virus (HLCrV). An infectious clone of this virus has recently been published but was, confusingly, named *Althea rosea* enation virus (Bigarre, 2001). All these virus diseases were shown to be associated with DNA β molecules in the present study.

The screening of disease isolates for the presence of DNA β did not detect its presence in any of the material collected in the New World. Isolates investigated included *Abutilon* mosaic, sida golden mosaic (originating from Puerto Rico and Columbia) and cotton leaf crumple disease (originating from Arizona). These three diseases are known to be caused by bipartite begomoviruses (Frischmuth *et al*., 1991;

Höfer *et al*., 1997; Nadeem *et al*., 1997). A number of Old World diseases also were shown not to be associated with DNA β including cassava mosaic disease (isolates originating from Kenya, Uganda, Nigeria and India), watermelon chlorotic stunt disease (originating from Yemen) and tomato yellow leaf curl disease (originating from Spain and Sicily). It is possibly not surprising to find that these diseases are not associated with DNA βs infectious clones of begomoviruses have been demonstrated to cause the full range of symptoms. With the exception of tomato yellow leaf curl these disease are caused by bipartite begomivuses.

The analysis conducted shows DNA β only to be associated with begomoviruses originating from the Old World. The only begomovirus-DNA β complex known to occur in the New World is that involving honeysuckle yellow vein mosaic virus (HYVMV) infecting variegated honeysuckle (*Lonicera japonica* Thunb. var. aureo-reticulata). Variegated honeysuckle is a perennial climbing vine which is grown as an ornamental for its striking yellow vein net symptoms. It is native to Korea, China and Japan and is thought to have been introduced to both Europe and North America in the early part of the 19th Century. The begomovirus associated with variegated honeysuckle has characteristics associated with the Old World begomoviruses and is thus not a native of New World. It is also interesting to note that thus far DNA. has only been found associated with monopartite begomoviruses and no native monopartite begomoviruses have been identified in the New World, although *Tomato yellow leaf curl virus* (originating in the Middle East) has recently been introduced to the Carribean, Central America and the southeastern USA. These finding thus strongly suggest that the begomovirus-DNA β interaction evolved after the divergence of the Old World and New World begomoviruses; the divergence is estimated to have occurred some 130 million years ago.

The diseases associated with begomovirus-DNA β complexes which have thus far been identified fall into two basic types; those associated with vein yellowing and those associated vein swelling and the production of enations. On the whole the vein yellowing diseases produce little if any leaf deformation of their natural host plants with vein yellowing the only cytopathological effect. The vein swelling/enation group of diseases are associated, in most cases with severe cytopathological changes including vein swelling, vein darkening and in severe diseases, such as CLCuD and HLCrD, the production of enations and leaf-like enations on the veins on the undersides of leaves. For HLCrD a study of the histological changes induced in hollyhock included replacement of spongy parenchyma by palisade parenchyma in leaves and abnormal cambial activity in phloem parenchyma leading to the formation of secondary vascular elements (Bigarré *et al.*, 2001). These changes are in most respects similar to the changes induced by CLCuD in cotton (Ashraf *et al*., 1999; P.G. Markham, unpublished results). For one disease, however, mixed symptoms were seen. The *Ageratum* isolate from which DNA β clone ALC01-Pak was isolated exhibited both yellow vein and enation type symptoms with severe leaf distortion. It is as yet unclear whether this isolate was a mixture of two DNA β requiring virus isolates/species or whether one begomovirus-DNA β combination is capable of inducing both disease phenotypes in a single host. Inoculation of ALC01-Pak with AYVV-Sin to *A. conyzoides* produced only yellow vein symptoms. This result suggest that either this isolate was a mixture or that the leaf curl symptoms are a feature of the virus associated with this isolate and the yellow vein symptoms are

determined by the DNA β. Resolution of this question will have to await the production of full-length begomovirus clones from this isolate to inoculate with ALC01-Pak.

The present cotton leaf curl disease epidemic affecting all the cotton growing areas of Pakistan and is now sweeping in to western India, originated near the city of Multan in the mid 1980s. The nucleotide sequence similarity (98.9%) of CLCuD-associated DNA β molecules originating from western India to those originating from Pakistan confirms that the causative agent of the disease is the same in both regions and is presumably being carried into India by wind-blown whitefly. Several other instances of leaf curl symptoms on cotton have been reported from southern India and New Delhi, well away from the advancing front of the epidemic Nateshan *et al.*, 1996). It remains to be shown whether these diseases also are caused by a begomovirus/DNA complex. A leaf curl disease of cotton has been known in Sudan since the 1900s and was recently shown to be associated with a monopartite begomovirus (Idris *et al.*, 2000). It will be interesting to see whether this virus too requires a satellite molecule.

Our results demonstrate that tomato leaf curl disease (TLCD) originating from Pakistan is associated with at least two distinct DNA β molecules. The first, TomLC01-Pak, falls in the *Malcaceae* super group of DNA β molecules, albeit basal to representatives isolated from cotton. TomLC02-Pak, in contrast, falls in the second super group and is most similar to a DNA β isolated from a second solanaceous host (ChLC01-Pak). It would appear therefore that TLCD in Pakistan could be caused by two distinct DNA β molecules. TLCD on the India sub-continent is caused by either a bipartite begomovirus (Padidam *et al.*, 1995; Mansoor *et al.*, 1997) or a monopartite begomovirus (accession no. U38239). Clones of the monopartite virus (Tomato leaf curl Karnataka virus) are infectious to tomato, producing authentic TLCD symptoms, in the absence of either a DNA B or a DNA β and this virus is thus unlikely to be the one associated with the DNA β molecule identified in tomato.

All the DNA β molecules characterized so far have a putative stem-loop structure with the loop sequence TAATATTAC (the so-called nonanucleotide motif); a motif they share with geminiviruses. A single DNA β molecule, OLC02-Pak (NIB5-2), contains the nonanucleotide sequence TAGTATTAC, common to viruses of the genus *Nanovirus* as well as the satellite-like molecules (DNA 1) associated with begomovirus/DNA β complexes. For both geminiviruses and nanoviruses, the nonanucletide is the origin of virion strand DNA replication. This sequence is nicked (between bases 7 and 8) by the virus-encoded replication associated protein (Rep), yielding an origin for extension of the 3' end of the nicked strand by a host-encoded DNA polymerase using the complementary strand as a template. Since clone OLC02-Pak NIB5-2 is infectious in the presence of AYVV, it is probable that nonanucleotide motif TAGTATTAC is recognized by a geminivirus Rep. However, the mechanism of interaction between begomovirus Rep and DNA remains unclear. The begomovirus origin of replication comprises repeated sequences (iterons) that are the binding sites for Rep, in addition to the predicted stem-loop structure containing the nonanucleotide loop sequence (reviewed by Hanley-Bowdoin *et al.*, 1999).

Both theoretical and experimental studies of the specificity of Rep binding have shown that iterons have a 5bp core sequence ($GGN_1N_2N_3$) and a variable number of

additional nucleotides which are species (sometimes isolate) specific (Argüello-Astorga *et al*., 2001; Chatterji *et al*., 1999). The conserved region of DNA molecules contains a number of highly conserved GG dinucleotides. It is intriguing to speculate that one or more of these GG dinucleotides could be the basis of the DNA β Rep binding motif; a motif which breaks the usual sequence specificity for Rep/iteron interaction.

The satellite molecule (ToLCVsat) identified by Dry *et al.*, associated with *Tomato leaf curl virus* (ToLCV) in Australia, is clearly derived from a DNA β progenitor. It maintains the A rich region (51% A content, approximately 157 nucleotides in length) but lacks the coding region for the C1 gene. In this respect TLCVsat is very similar to the small defective DNA β molecules isolated as part of this study, such as HLCr01-Egy. However, the region of ToLCVsat equivalent to the conserved region of DNA β molecules contains extensive sequence deletions, insertions and rearrangements. Also the sequence of the predicted hairpin structure of ToLCVsat, with the exception of the nonanucleotide motif, is entirely different to that of the DNA β molecules. It is apparent, therefore, that the ToLCVsat has diverged significantly from DNA β. One of the insertions in ToLCVsat is mirrored by an insertion at the same position, albeit somewhat larger (26 nucleotides compared to 14 nucleotides), in OYV01-Egy. This insertion is not found in the full-length parent molecule (OYV01-Egy), which was isolated from the same plant. These insertions within the conserved regions may thus be a feature of the degeneration of DNA β molecules following the deletion of the C1 gene. It is possible that these insertions and mutations within the defective DNA β molecules are tolerated due to their small size. Similar insertions in a full length DNA β might be expected to interfere with replication of the molecule and consequently be out competed by the intact DNA β.

Table 3. Origins of disease isolates

Clone	Origin (country/town[state]/year)	Plant Species	Symptoms
CLCβ1-Pak (u77-4)	Pakistan/Faisalabad/1994	*Gossypium hirsutum* [cotton]	LC,VD,E,LE
CLCβ01-Pak[1] (u77-5)	Pakistan/Faisalabad/1994	*Gossypium hirsutum* [cotton]	LC,VD,E,LE
CLCβ02-Pak[1] (u89-2)	Pakistan/Faisalabad/1994	*Gossypium hirsutum* [cotton]	LC,VD,E,LE
ALCβ01-Pak (f24-9)	Pakistan/Faisalabad/2000	*Ageratum conyzoides*	LC,VY,E,
ALCβ02-Pak (nib8-2)	Pakistan/Faisalabad/1994	*Ageratum conyzoides*	LC,VD
ZLCβ01-Pak (nib9-1)	Pakistan/Faisalabad/1994	*Zinnia elegans*	LC,VD
ZLCβ02-Pak (nib18-1)	Pakistan/Faisalabad/1994	*Zinnia elegans*	LC,VS
HLCβ01-Pak (u81-1)	Pakistan/Faisalabad/1994	*Hibiscus rosa-sinensis*	LC,VD,E,LE
OLCβ01-Pak (nib1-4)	Pakistan/Gojra/1997	*Hibiscus esculentis* [okra]	LC,YM
OLCβ02-Pak (nib5-2)	Pakistan/Bahawalpur/1997	*Hibiscus esculentis* [okra]	VY

Table 3 (*continued*)

OLCβ03-Pak (nib10-3)	Pakistan/Bahawalpur/1997	*Hibiscus esculentis* [okra]	VY
ChLCβ01-Pak (nib16-1)	Pakistan/Mian Channu/1997	chillies	LC,VS,
TobLCβ01-Pak (nib12-1)	Pakistan/Rahim Yar Khan/1998	*Nicotiana tabacum*	LC
TobLCβ02-Pak (nib2-3)	Pakistan/Bahawalpur/1999	*Nicotiana tabacum*	LC
TomLCβ01-Pak (nib23-1)	Pakistan/Okara/1999	tomato	LC,E
TomLCβ02-Pak (nib14-1)	Pakistan/Rahim Yar Khan/1997	tomato	LC,YM
CLCβ01-Ind (u80-1)	India/Dabwali[Rajasthan]/1995	*Gossypium hirsutum* [cotton]	LC,VD,E,LE
CLCβ02-Ind (u79-1)	India/Dabwali[Rajasthan]/1995	*Gossypium hirsutum* [cotton]	LC,VD,E,LE
CLCβ01-Ind (u79-4)	India/Dabwali[Rajasthan]/1995	*Gossypium hirsutum* [cotton]	LC,VD,E,LE
AYVβ01-Ind (sb8-4)	India/[Punjab]/1997	*Ageratum conyzoides*	VY
BYVβ01-Ind[2]	India/Madurai[Tamil Nadu]/?	*Hibiscus esculentis* [okra]	VY
OYVβ01-Egy (sb36-1)	Egypt/[Fayoum]/1995	*Hibiscus esculentis* [okra]	VY
OLCβ02-Egy(JKBsat3)	Egypt/Cairo/2000	*Hibiscus esculentis* [okra]	
OLCβ03-Egy(JKBsat10)	Egypt/Cairo/2000	*Hibiscus esculentis* [okra]	
OYVβ01-Egy(sb18-1)	Egypt/[Fayoum]/1995	*Hibiscus esculentis* [okra]	VY,YM
OLCβ02-Egy(JKBsat27)	Egypt/Cairo/2000	*Hibiscus esculentis* [okra]	
HLCrβ01-Egy (f28-3)	Egypt/Cairo/1995	*Althea rosea* [hollyhock]	LC,VD,E,LE
HLCrβ02-Egy (JKB14)	Egypt/Cairo/2000	*Althea rosea* [hollyhock	LC,VD,E,LE
AYVβ01-Sin[3]	Singapore/?	*Ageratum conyzoides*	VY
HYVMβ01 (u96-2)	United Kingdom/Norwich/1996	*Lonicera japonica* [honeysuckle]	VY

1. Briddon *et al.* (2001).
2. EMBL accession no. AJ308425
3. Saunders *et al.* (2000).
4. LC - leaf curl; VD - vein darkening; VY - vein yellowing; VS - vein swelling; E - enations; LE - leaf-like enations; YM - yellow mottling of leaves.

Table 4. Features of DNA β molecules

Clone	Size (bp)	EMBL Accession No.	Nucleotide sequence similarity to CLCβ02-Pak/AYVβ01-Sin (%)	Infectivity[1] (trans-replication/ symptoms)	C1 gene product (no. of amino acids [predicted molecular weight kDa])	Percentage amino acid sequence similarity of C1 gene product to that of CLCβ02-Pak/AYVβ01-Sin
CLCβ)01-Pak (u77-4)	1247	AJ299443	95/55	C/-(def.)	99	98/47
CLCβ01-Pak (u77-5)	1351	AJ292769	96/53	C/-	118	97/47
CLCβ02-Pak (u89-2)	1349	AJ298903	100/53	C/-	118	100/46
ALCβ01-Pak (f24-9)	1351	AJ316026	51/57	A/-	138	42/68
ALCβ02-Pak (nib8-2)	1352	AJ316027	53/59		138	43/69
ZLCβ01-Pak (nib9-1)	1349	AJ316041	52/60	def	74	39/33
ZLCβ02-Pak (nib18-1)	1350	AJ316028	53/61		138	43/70
HLCβ01-Pak (u81-1)	1347	AJ297908	95/53	def (?)	105	86/44
OLCβ01-Pak (nib1-4)	1361	AJ316029	63/53		146	66/45
OLCβ02-Pak (nib5-2)	1362	AJ316030	60/54	C in Nb	99	64/43
OLCβ03-Pak (nib10-3)	1369	AJ316031	54/61		138	44/69
ChLCβ01-Pak (nib16-1)	1387	AJ316032	53/60		149	42/67
TobLCβ01-Pak (nib12-1)	1356	AJ316033	55/62		127	40/68
TobLCβ02-Pak (nib2-3)	1356	AJ316034	55/62		127	40/68
TomLCβ01-Pak (nib23-1)	1370	AJ316035	92/56		118	99/46
TomLCβ02-Pak (nib14-1)	1374	AJ316036	52/59		118	43/67
CLCβ)01-Ind (u80-1)	1248	AJ291601	96/56	no (def.)	118	97/44
CLCβ)02-Ind (u79-1)	1353	AJ316037	99/54	no (def.)	77	84/40
CLCβ02-Ind (u79-4)	1351	AJ316038	99/54	C/-	118	97/45

Table 4 (*continued*)

CLCR03-Ind (u79-2)	1348	AJ298903			118	
AYVβ)01-Ind (sb8-4)	902	AJ316042	53/55	def	-	-
BYVβ01-Ind	1353	AJ308425	60/54	?	140	66/44
OYVβ01-Egy (sb36-1)	1307	AJ316039	55/52	C/-	117	44/48
OLCβ02-Egy(JKBsat3)	1350	AF397217	56/50		117	44/48
OLCβ03-Egy(JKBsat10	1305	AF397215	56/52		117	44/48
OYVβ)01-Egy(sb18-1)	785	AJ316043	57/55	def	-	-
OLCβ)02-Egy(JKBsat27)	774	AF397216	57/54	def	-	
HLCrβ)01-Egy (f28-3)	660	AJ316044	57/58	def	-	-
HLCrβ)02-Egy(JKB14)	741	AF397214	55/54	def	-	-
AYVβ01-Sin	1347	AYE252072	53/100	A/-	118	46/100
HYVMβ01 (u96-2)	1344	AJ316040	50/53	A/∃	116	43/52

1. Infectivity is given as ability to be transreplicated by either AYVV (A) or CLCuV(C). The host in which this was achieved is shown superscript. Symptoms are only given for DNA β molecules inoculated to *N. benthamiana* with AYVV. The symptoms are designated as either like AYVV in the absence of AYVD DNA β (A; upward leaf curling) or like an AYVV with AYVD DNA β infection (β; downward leaf curling). For infections of *A. conyzoides* with AYVV and a DNA β or cotton with CLCuV and a DNA β the symptoms were as described for AYVV with AYVD DNA β (vein yellowing) or CLCuV with CLCuD DNA β (vein swelling, vein darkening and the production of enations) respectively.

1. Infectious at low efficiency to both N. benthamiana and cotton.u774
2. Infectious to both *N. benthamiana* and *A. conyzoides*.
3. Infectious to *N. benthamiana*.
4. Recombinant DNA β molecule described by Briddon *et al*. (2001).

1.12 Understanding the role of DNA β in cotton leaf curl disease

In this study mutagenesis and functional analysis of a DNA β molecule isolated from cotton was carried out to define its role in the process. Interactions between unrelated plant viruses are common, and some viruses depend on such interactions for their survival. Frequently, a virus lacks some essential molecular function that another provides. A comparative analysis of the sequences of DNA β isolated from various crops and locations suggests that DNA β molecules encode a single conserved open reading frame (ORF) on the complementary-sense (Briddon *et al*., 2001) and suggests an essential role for this ORF.

Analysis of putative open reading frames (ORFs) encoded by DNA βs cloned from various host plants has shown the presence of a single conserved open reading frame in the complementary sense termed as βC1. Mutagenesis analysis showed that βC1 is essential for the development of typical disease symptoms in cotton. The ORF is transcribed at a high level in infected plants. The promoter region of βC1 showed strong constitutive expression when fused to the reporter gene GUS. The expression of βC1 from a PVX vector developed severe chlorotic and necrotic symptoms on *Nicotiana benthamiana* plants, and accumulated higher levels of viral RNA, suggesting that βC1 is a virulence determinant. PVX vector expressing βC1 was unable to systemically infect *Nicotiana tabacum* var. Samsun, suggesting that βC1 became a target of host defense responses in this host. However, no visible hypersensitive response (HR) was found on the inoculated leaf when expressed either from PVX vector or by agroinfiltration of a binary vector expressing βC1 from 35S promoter. The ability of βC1 to overcome antiviral host defense responses mediated either by gene silencing or R gene was tested. βC1 suppressed post-transcriptional gene silencing of a transgene (GFP) that was initiated either by ectopic expression of the transgene or by another virus carrying GFP sequences. βC1 suppressed production of short RNAs that are associated with PTGS. Similarly, the inoculation of PVX expressing DNA βC1 prevented the hypersensitive response (HR), and accumulated detectable levels of viral RNA on inoculated leaves of transgenic plants expressing Rx, an extreme resistance gene against PVX. Our results indicate that DNA β encodes an essential virulence determinant that suppresses host defense response mediated either by gene silencing or by a resistance gene. This is the first example of a virulence determinant encoded by a DNA satellite, which is essential for induction of disease symptoms.

DNA β with mutated βC1 is replicated by helper virus but is unable to develop typical disease symptoms in cotton. We identified two populations of DNA β cloned from cotton samples originating from India and Pakistan. One of these was having a small deletion at N terminal but was still infectious. In order to avoid potential that alternative start codon might be involved, mutation was created in such a way that would stop βC1 expression from both these clones. The mutant was replicated by CLCuV DNA A and moved systemically in *Nicotiana benthamiana* plants. However, symptoms exhibited were those shown by inoculation of DNA A alone. Inoculation of the mutant DNA β with DNA A failed to show symptoms of cotton leaf curl disease in cotton. The data confirmed that effects of DNA β are not due to replication enhancement as replication of DNA A was unaffected in *N. benthamiana.*

The transcription of the βC1 was investigated by Northern analysis using a cloned probe encompassing the ORF. The viral transcripts were readily detected in infected plants, showing that the ORF is transcriptionally active. Similar results were found in cotton (data not shown). The symptoms induced by CLCuV suggest that the virus is phloem restricted. Since we used RNA extracted from whole leaf suggest that the ORF is transcribed at high levels. In order to assess whether the expression is regulated by a DNA A encoded protein, the promoter region of β C1 was used to drive the expression of a reporter gene GUS. The construct was cloned in a binary vector. The clone was used to express the GUS transiently in tobacco and *N. benthamiana*. High level expression of GUS was directed by β C1 promoter in the absence of DNA A that was visually similar to those derived by 35S promoter. These results confirmed that β C1 is expressed from a strong constitutive promoter that results in readily detectable levels of transcripts.

The data so far established that βC1 is an essentially required protein for the development of disease symptoms. The ORF was cloned in PVX vector to assess phenotype that result when βC1 is expressed in a heterologous system. The expression of the protein from PVX vector resulted in severe necrosis, chlorosis and leaf curling on *N. benthamiana* plants. The clone with a mutation in βC1 was unable to show these symptoms and further confirmed that these effects are mediated by the expression of the protein. The severity of symptoms and accumulation of higher level of viral RNA expressing βC1 showed that the protein is a virulence determinant. PVX is infectious on *N. tabacum* and produces symptoms in this host. However, inoculation of PVX expressing βC1 failed to show systemic symptoms. In few plants a leaf above the inoculated leaf showed necrosis in the veins but was unable to spread further. These resulted suggested that βC1 might have become a target of host defense responses in this host. However, there was no hypersensitive response on the inoculated leaf. On the other hand, inoculation of PVX with βC1 mutant resulted in systemic infection. Northern analysis showed that detectable levels of the virus are present in the inoculated leaf but no detectable levels were present in systemic leaves. We further tested whether high level expression would result in HR on infiltrated leaf. Transient expression of βC1 35S promoter was unable to show detectable HR response as has been found for other avirulence proteins. Plants inoculated with PVX AC2, another geminivirus protein with a suppressor activity were systemically infectious and high level of viral RNA were detected in systemic leaves. βC1 was systemically infectious on *N. glutinosa*, ruling out possibility that the vector was mutated.

Severe necrotic and chlorotic symptoms developed by PVX expressing βC1 were typical of suppressors of gene silencing encoded by DNA and RNA viruses (Voinnet *et al*., 1999). These suppressors are often viral pathogenecity determinants with the ability to suppress PTGS. PTGS of a transgene can be initiated by a replicating virus containing part of coding sequences (VIGS) or by gene silencing signal initiated by ectopic expression of sequences homologous to the transgene. GFP transgenic plants where GFP expression was blocked by PTGS were inoculation with PVX expressing βC1 protein. The expression of βC1 was able to suppress silencing of GFP initiated either by signaling or by a replicating virus, showing that βC1 is a pathogenecity determinant with a suppressor of gene silencing activity.

Post-transcriptional gene silencing is associated with the production of short RNA species of 20-25 nucleotides with sequence homology to targeted transcripts. Suppressors of gene silencing may block the production of these RNA species. The ability of βC1 to block production of short RNA species was assessed by a transient expression system. Agroinfiltration of GFP on non-silenced GFP transgenic plants resulted in silencing of GFP expression while the introduction of GFP and βC1 on non-silenced GFP transgenic plants blocked gene silencing as evidenced by a bright green fluorescence even after 10 days post-infiltration. An assay for short RNA species confirmed the absence of short RNA species at all time points. There appears to be two species of short RNAs and both of these are absent in plants where βC1 was agroinfiltrated with GFP. The data confirmed that βC1 is a suppressor that completely blocked the production of short RNA species.

There is another type of defense response mediated by resistance genes. The inability of the PVX expressing βC1 to infect systemically in the absence of HR and the presence of detectable levels of virus in the inoculated leaf was intriguing, suggesting that βC1 probably prevented HR but showed restricted movement out of the patch. In order to further understand the effect we used an established system of extreme resistance (Rx) against PVX. Mechanical inoculation of PVX on these transgenic plants results in no detectable levels without HR. On the other hand when the virus is expressed from 35S promoter delivered by agro-infiltration, it causes HR on inoculated leaf and no virus is detected in systemic leaves. The inoculation of PVX expressing βC1 prevented detectable HR and accumulated detectable level of viral RNA while mutant clone readily developed HR. Similarly, plants inoculated with PVXAC2 developed HR. Northern analysis confirmed the presence of detectable levels only in the inoculated leaf while no viral DNA was detected in systemic leaf and no symptoms were found on systemic leaves. These results confirm that βC1 can prevent HR both in Rx transgenic *N. benthamiana* plants.

Distinct DNA β molecules associated with begomovirus diseases in the Old World suggested that DNA β encode essentially required functions in these complexes. We have shown that mutations in the conserved ORF called βC1 abolish these affects. Our further analysis has shown that βC1 is a virulence determinant that interferes with resistance pathways mediated either by disease resistance genes or by gene silencing.

Mutation analysis has clearly established that the protein rather than coding sequences or RNA mediates DNA β associated effects. Viroids and satellites associated with some RNA viruses do not code for proteins and thus RNA itself is able to modulate symptom development. Viroids are shown to mediate their effect by gene silencing. The data suggesting that the protein encoded by the βC1 ORF mediate these effects is supported by the fact that defective interfering molecules associated with AYVV or CLCuV that lack either promoter or the gene are unable to develop disease symptoms (Stanley *et al*., 1997; Briddon *et al*., 2001). On the other hand a recombinant molecule containing intact ORF was infectious on ageratum and developed symptoms typical of AYVV (Saunders *et al*., 2001). The virion-sense promoter of begomoviruses is activated *in trans* by AC2 for expression of BV1 and the coat protein. The promoter region of βC1 showed strong constitutive promoter activity when fused to a reporter gene GUS, suggesting that DNA A encoded proteins are not required for the expression of this gene.

The expression of βC1 from PVX vector developed severe chlorotic and necrotic symptoms and accumulated higher level of viral RNA. Interestingly, enations or vein thickening symptoms were not developed on inoculated plants, suggesting that either another gene or combination of viral gene induces formation of enations. The symptoms from PVX vector are relevant to the fact that PVX encodes a weak suppressor (25K) that block gene silencing signal but is unable to suppress maintenance of gene silencing (Voinnet *et al*., 2000). Plants inoculated with PVX expressing βC1 showed enhanced susceptibility and accumulated higher level of viral RNA even in upper leaves, suggesting that βC1 suppress host defense responses in this host. Pathogenecity determinants encoded by plant pathogens are often target of host defense responses mediated by resistance genes. The restricted levels of viral RNA in *Nicotiana tabacum* suggest that βC1 became a target of host defense responses. It would be interesting to find out whether the same protein is the avirulence factor for natural resistance in virus-resistant cotton cultivars.

Programmed cell death (PCD) also called hypersensitive response (HR) in plants is part of disease and stress response in plants. There is an emerging hypothesis that stress response is a part of disease resistance pathways in plants. Several animal viruses have been found to suppress this defense response by preventing PCD. Rx is an extreme resistance gene that prevents accumulation of viral nucleic acids when activated by PVX coat protein (Bendahmane *et al*., 1995). The absence of HR on the inoculated leaf suggested the βC1 was able to prevent pathway leading to HR and is supported by the fact that detectable virus levels accumulated in the inoculated leaf. We tested some other suppressors cloned in PVX vector to find out whether this is a general feature of suppressors. Whether the prevention of HR is a general property of pathogenecity determinants of begomoviruses is not known. However, AC2 encoded by another geminivirus was unable to prevent HR. Similarly other suppressors encoded by RNA viruses were also unable to prevent HR. These finding suggest that βC1 interfere with host stress responses that result in HR.

βC1 was found to be a suppressor of gene silencing. Another viral protein encoded by geminiviruses called AC2 also has gene silencing activity. Although we don't know whether suppression of gene silencing is a general property of this protein encoded by geminiviruses belonging to different genera. These suppressors may either act at different steps in gene silencing pathway or may have cumulative effects. Different suppressors of gene silencing act at different steps in gene silencing pathway. There also evidence that co-infiltration of several suppressors result in enhanced suppression of PTGS. Geminiviruses are known to have synergistic effects when co-inoculated may result from additive effects. The fact that DNA A of CLCuV is infectious on it's own support the hypothesis that these suppressors might have additive effect. Identification of targets of two suppressors encoded by geminiviruses would help to understand how these suppressors are complemented.

Our results show that DNA β encodes an essentially required virulence determinant that interferes with disease resistance pathways mediated either by a disease resistance gene or by gene silencing. Interestingly, while begomoviruses associated with geminivirus complexes involving DNA β show divergence based on geographical distribution but DNA β show relationships based on their hosts. The

conserved nature of this protein suggests that distinct DNA βββ associated with different hosts might have evolved to overcome host defense responses in distinct hosts.

Epigenetic gene silencing result either from the inhibition of transcription called transcriptional gene silencing (TGS) or from sequence specific degradation of target transcripts and is termed post-transcriptional gene silencing (PTGS). PTGS has been shown to be a part of defense against foreign nucleic acids including transgenes, viruses and transposons. RNA silencing can thus be envisioned as a form of immune system that operates at nucleic acid level where specificity of the response is determined by the sequence homology. Since the majority of plant viruses have RNA genome, this RNA based system forms a potent antiviral defense. Transcripts of plant DNA viruses and transgenes that are localized in the nucleus may be perceived as foreign and thus are also target of PTGS. A key factor in PTGS is the involvement of double stranded RNA that triggers sequence specific degradation and generation of short RNA that are incorporated into an RNAase complex. A remarkable aspect of gene silencing is that once initiated it can be transported by gene silencing signal. PTGS is often associated with methylation of coding sequences. The role of DNA methylation in PTGS is not clear but it may be related to aberrant RNA production. The mechanism of PTGS appears to be highly conserved across kingdoms as evidenced by recent genetic and biochemical studies conducted in plants, animals and fungi. These studies have identified a host encoded RNA dependant RNA polymerase (RdRP) that converts single-stranded RNA into double-stranded RNA that may be a trigger for RNA silencing. Recent studies show that in Drosophila RNAase III related protein, Dicer, process dsRNA into 21-23nt RNA *in vitro*. These short RNAs are incorporated into RISC-mediated ssRNA degradation.

Plant viruses have evolved mechanisms to overcome this defense system by coding proteins that suppress gene silencing. The idea of suppression of gene silencing was prompted by the findings that unrelated plant viruses often show synergistic effects. These proteins are able to overcome antiviral host response and can simultaneously overcome transgene induced or virus-induced gene silencing. The majority of these proteins were previously recognized as pathogenicity determinants. These suppressors are highly diverse in sequence and structure, indicative of an evolutionary convergence. The highly diverse nature of these proteins suggests that silencing suppressors have evolved as an additional feature of multifunctional viral proteins. These viral proteins may have evolved due to divergence of host proteins that these suppressors interact in diverse plant species. Another feature of these suppressors is that the same protein cloned from related viruses may differ highly in their silencing activity. The suppressors of gene silencing studied so far suggest that these proteins differ in their ability to suppress DNA methylation or suppression of two species of short RNAs species associated with gene silencing, suggesting that they interfere at different steps in gene silencing pathway. For example 25K protein of PVX was shown to suppress gene silencing signal. The other suppressors were found to affect maintenance of gene silencing. Nevertheless these suppressors are valuable tools to understand gene silencing pathway.

Geminiviruses are among diverse plant viruses that were found to suppress gene silencing. The geminivirus protein with suppressor activity called AC2 is found among all geminiviruses belonging to the genus begomovirus while a functional

homolog exits in geminiviruses belonging to the genus curtovirus or topocuviruses. The other suppressor was found to be encoded by a DNA satellite associated with a group of monopaqrtite begomoviruses that require this satellite for systemic infection of their natural host plant. The diverse nature of this protein suggests that this suppressor has evolved to function in highly diverse host plants. Intriguingly all suppressors except AC2 known so far have a role in virus movement. This suggests that another geminivirus protein may also suppress gene silencing that has a role in the movement of the virus.

Here we have studied how βC1 suppress gene silencing. We show here that βC1 suppress production of both species of short RNAs associated with gene silencing. We found that βC1 suppress gene silencing signal and complements 25K movement protein of PVX and has also a role in the movement of cotton leaf curl virus in cotton.

1.13 New hosts of cotton leaf curl virus and other begomovirus diseases in Pakistan

1.13.1 Widespread Occurrence of *Cotton leaf curl virus* on Radish in Pakistan

The current epidemic of cotton leaf curl disease (CLCuD) in Pakistan started in 1988 with the natural host range limited to a few plant species in the family Malvaceae. However, we have observed expansion in the host range of the virus, and several non-Malvaceous plants were found to be infected with the virus. Characteristic symptoms of CLCuD such as leaf curl and enations have been observed on radish plants, primarily in kitchen gardens. However, in 1999, levels of infection of 10 to 90% were observed both in commercial fields and kitchen gardens in the Punjab province of Pakistan. Both symptomatic and nonsymptomatic samples were collected from five different locations. Total DNA was isolated, dot-blotted on nylon membrane, and a full-length clone corresponding to DNA A of cotton leaf curl virus was labeled with (α32)P dCTP and used as a probe for the detection of a begomovirus. Strong signals were observed in symptomatic plants while no signals were observed in nonsymptomatic plants. Infection with a begomovirus was further confirmed by polymerase chain reaction (PCR) using degenerate primers for DNA A. Primers specific for the two distinct begomoviruses associated with CLCuD were also used in PCR reactions, and products of the expected size were obtained from all symptomatic samples, confirming infection with begomoviruses similar to those associated with CLCuD. A full-length probe of a nanovirus-like molecule associated with cotton leaf disease, called DNA 1 was labeled with (α32)P dCTP and detected the virus only in symptomatic plants. Similarly, primers specific for DNA 1 amplified a product of expected size when used in PCR. On the basis of symptomatology and the detection of specific viral components associated with the disease, we confirmed that radish plants are infected with *Cotton leaf curl virus* (CLCuV). Since radish is a short duration crop, infection of CLCuV in radish may not serve as a direct source of infection for the next cotton crop. However, it is a potential threat to tomato crops which overlap with radish in the Punjab province. The detection of CLCuD in radish is another example of the mobilization of begomoviruses to previously unknown hosts.

1.13.2 Evidence for the association of a bipartite geminivirus with tomato leaf curl disease in Pakistan

Tomato leaf curl disease is the most important constraint on tomato production in Pakistan, where it is found throughout the country. The disease, which occurs in high incidence in Punjab and Sindh provinces, causes 30 to 40% yield losses in the spring crop and uneconomically high losses when grown as an autumn crop. The symptoms of the disease include upward or downward leaf curling, vein thickening, and stunting of the plant. The disease is transmitted by *Bemisia tabaci* whiteflies (non-B, biotype K) and is suspected to be caused by a geminivirus. For the detection of geminivirus, total DNA was extracted from infected plants, fractionated in an agarose gel, transferred to a nylon membrane, and Southern blotted. A full-length clone of DNA-A of cotton leaf curl virus from Pakistan was labeled with [α32P]dCTP by the oligo-labeling method and hybridized at medium stringency. Geminivirus DNA forms that are normally found in infected plants were detected in plants with tomato leaf curl disease but not in healthy plants. To further confirm the presence of a whitefly-transmitted geminivirus, universal primers for dicot-infecting geminiviruses were used in polymerase chain reaction (PCR) and a product of expected size (approximately 2.7 kb) was detected. The 2.7-kb PCR-amplified DNA from diseased tomato plants was labeled with [α32P]dCTP and used as probe in Southern hybridization. This probe also detected geminivirus DNA forms at medium stringency. Both monopartite and bipartite geminiviruses transmitted by whiteflies have been reported to cause leaf curl symptoms on tomato from the Eastern hemisphere. Degenerate primers (PBLv2040 and PCRc1), which amplify B component DNA, were used to determine if tomato leaf curl was monopartite or bipartite. A product of expected size (0.65 kb) was amplified, suggesting this virus to be bipartite. DNA-B PCR product obtained from diseased tomato plants was hybridized as described above and detected geminivirus DNA forms at medium stringency. Samples of diseased tomato plants were collected from tomato fields throughout Punjab. DNA-A was detected in all 20 samples whereas DNA B was detected in 17 samples when hybridized by dot blot method at medium stringency. Our data show that tomato leaf curl virus from Pakistan is a bipartite geminivirus. This is the first evidence for a bipartite geminivirus in tomato plants from Pakistan.

1.13.3 Evidence that watermelon leaf curl disease in Pakistan is associated with *Tomato leaf curl virus*-India, a bipartite begomovirus

Whitefly-transmitted geminiviruses (begomoviruses) have emerged as major constraints on food and fiber crops worldwide, and there are several examples of begomovirus mobilization in previously unknown host plants. Here we report on evidence that leaf curl disease of watermelon in Pakistan is caused by *Tomato leaf curl virus*-India (TLCV-India). Leaf curl disease of watermelon, characterized by leaf curling and mottling and stunted plant growth, was observed at several locations in the Punjab Province of Pakistan. Symptomatic and asymptomatic leaf samples were collected from three locations, and total DNA was isolated by the cetyl trimethyl ammonium bromide method and resolved in agarose gel. A full-length clone of *Cotton leaf curl virus* DNA A was labeled with (α [32P])dCTP and used as a general probe in Southern hybridization. The probe detected characteristic geminivirus DNA forms in infected watermelon plants, whereas no signal was detected in asymptomatic plants. The association of a begomovirus was confirmed further by polymerase chain

reaction (PCR) amplification with degenerate primers PAL1V and pAR1c. Samples were screened for infection by TLCV-India, because of symptom similarity. A full-length clone of DNA B of TLCV-India was labeled with (α [32P])dCTP by random priming and was used as a specific probe in Southern hybridization. The probe detected geminivirus DNA forms, showing that the disease is associated with TLCV-India. Primers TLCV1 (GAGGTACCAAAACTTGTCGTTTTGATTCGG), in the virion-sense, and TLCV2 (GCCCATGGTTCTTTGCTCGGAGAACAAGAA), in the complementary-sense, were designed based on the sequence of DNA A of TLCV-India. These primers were used in PCR and amplified a product of the expected size from infected plants. Similarly, primers TLCVBC1 (GCGGATCCTTATTCCGTAATTATATCTGCA), in the virion-sense, and TLCV BC2 (CACCATGGCAATAGGAAATGATGGTATGGG), in the complementary-sense, were designed based on the sequence of DNA B of TLCV-India. These primers amplified a product of expected size when used in PCR. The results show that watermelon leaf curl disease in Pakistan is associated with TLCV-India. This the first report of detection of a begomovirus in watermelon in Pakistan and the first report of detection of TLCV-India on a plant other than tomato from Southeast Asia.

1.13.4 Association of a monopartite begomovirus producing subgenomic DNA and a distinct DNA β on *Croton bonplandianus* showing yellow Vein symptoms in Pakistan

The recent discovery that monopartite begomoviruses on ageratum and cotton essentially require a DNA satellite called DNA β is leading to identification of several other hosts that have similar disease complexes. A weed species (*Croton bonplandianus*) belonging to the family *Euphorbiaceae* is one such example. *C. bonplandianus* is widely distributed on wastelands throughout the Punjab Province in Pakistan. It very often shows yellow vein symptoms indicating infection by a begomovirus. To detect a begomovirus, both symptomatic and asymptomatic plants were collected from several widely separated locations in the Punjab Province. Total DNA was isolated from these samples by the cetyl trimethyl ammonium bromide (CTAB) method, resolved in an agarose gel, and blotted on a nylon membrane. A full-length clone of DNA A of *Cotton leaf curl virus* (CLCuV) labeled with (α32)PdCTP was used as a probe in Southern hybridization. The probe detected hybridizing bands only in symptomatic plants, confirming the presence of a begomovirus. In addition to hybridizing bands of the expected sizes, smaller bands were also detected, suggesting the presence of subgenomic molecules derived from DNA A. Universal polymerase chain reaction (PCR) primers for dicot-infecting geminiviruses were used in PCR for amplification of DNA A of the begomovirus associated with the disease. The use of these primers in PCR was expected to result in amplification of full-length DNA A. In addition to a product of the expected size (2.7 to 2.8 kb), another product of approximately 1.4 kb was amplified. The presence of subgenomic DNAs that are derived from DNA A is an indicator of the monopartite nature of begomoviruses, because in bipartite begomoviruses subgenomic DNAs are derived solely from DNA B. The presence of a DNA β, a DNA satellite associated with certain monopartite begomoviruses, was suspected because of symptoms and the possible monopartite nature of the virus. Universal primers for amplification of DNA β were used in PCR for amplification of a putative DNA β. The PCR reaction yielded a product of expected size (approximately 1.4 kb). A probe from the amplified product was made by the oligo-labeling method. The probe detected hybridizing bands in all

symptomatic samples collected from three locations, confirming the association of a DNA β with the disease. A duplicate blot when hybridized with a DNA β associated with ageratum yellow vein disease did not hybridize to these samples. These results confirm that yellow vein disease on this weed is associated with a monopartite begomovirus and a distinct DNA β.

1.13.5 Association of a begomovirus and nanovirus-like molecule with Ageratum yellow vein disease in Pakistan

Whitefly-transmitted geminiviruses (begomoviruses) cause heavy losses to many food and fiber crops in Pakistan. Many weeds also show symptoms typical of begomoviruses. Ageratum (*Ageratum conyzoides*) is a common perennial weed in Pakistan, growing along irrigation canals, that often shows symptoms, such as yellow vein and mosaic, suggesting infection by a begomovirus. To confirm this, symptomatic and asymptomatic ageratum plants were collected from three locations in the Punjab Province of Pakistan, and total DNA was isolated, subjected to agarose gel electrophoresis, transferred to a nylon membrane, and Southern blotted. Total DNA isolated from cotton infected with *Cotton leaf curl virus* (CLCuV), tomato infected with *Tomato leaf curl virus* from Pakistan (TLCV-Pak), tobacco infected with *African cassava mosaic virus* (ACMV) from Nigeria, and healthy tobacco were included as controls. A full-length clone of CLCuV DNA A was labeled with (α [32P])dCTP by oligo-labeling and hybridized at medium stringency. The probe detected characteristic geminivirus DNA forms in symptomatic ageratum and plants infected with CLCuV, TLCV-Pak, and ACMV, while no signal was detected in asymptomatic ageratum from the field or healthy tobacco. To confirm infection by a begomovirus, degenerate primers WTGF (GATTGTACGCGTCCDCCTTTAATTT GAAYBGG), designed in the rep gene of begomoviruses, and WTGR (TANACGCGTGGCTTCKRTACATGGCCTDT), designed in the coat protein gene of DNA A of begomoviruses, were used in polymerase chain reaction (PCR). Degenerate primers (PBLv2040 and PCRc1) also were used in PCR. A product of expected size (approximately 1.4 kb) was obtained with DNA A primers from symptomatic ageratum, while no product was obtained with DNA B primers in the same sample. Previously we were unable to detect a DNA component equivalent to begomovirus DNA B in cotton showing symptoms of cotton leaf curl disease. We recently reported a novel circular DNA molecule that was approximately half as long as the full-length DNA A (CLCuV DNA-1) associated with CLCuV that share homology to plant nanoviruses. The supercoiled replicative form of viral DNA isolated from infected ageratum plants indicated the presence of smaller molecules, as was found in cotton leaf curl disease, suggesting that a nanovirus-like molecule might be associated with ageratum yellow vein disease. A duplicate blot of samples used in Southern hybridization with the DNA A probe was prepared, and a probe of the full-length clone of the nanovirus-like molecule (CLCuV DNA-1) was prepared as described for DNA A. The probe detected characteristic nanovirus DNA forms in ageratum with yellow vein symptoms and cotton infected with CLCuV, while no signal was detected in plants infected with TLCV-Pak or ACMV, healthy tobacco, or asymptomatic ageratum. Abutting primers PB2-F and PB2R, designed based on the CLCuV DNA-1 sequence, were unable to amplify a PCR product from ageratum with yellow vein symptoms, suggesting the nanovirus-like molecule associated with ageratum yellow vein disease is distinct from CLCuV DNA-1. Our results show that yellow vein disease of ageratum in Pakistan is associated with a begomovirus

infection and single-stranded circular DNA molecule with similarity to CLCuV DNA1.

1.13.6 Association of a disease complex involving a begomovirus, DNA 1 and a distinct DNA β with leaf curl disease of okra in Pakistan

Okra leaf curl disease (OLCD), characterized by either upward or downward leaf curl and stunted plant growth, is one of the major diseases of okra (*Hibiscus esculentis* L.) in Pakistan. OLCD is transmitted by the whitefly *Bemisia tabaci* and is suspected of being associated with a whitefly-transmitted geminivirus (Genus *Begomovirus*). Total DNAs isolated from both symptomatic and healthy okra plants collected from several locations in Pakistan were resolved on agarose gels and blotted to nylon membranes. A full-length DNA A clone of *Cotton leaf curl virus* (CLCuV) from Pakistan was labeled with 32PdCTP and used as a probe at medium stringency. The probe detected the presence of characteristic geminivirus DNA forms in infected plants, while no hybridization was observed to healthy plant extracts, confirming the association of a begomovirus with OLCD. Degenerate oligonucleotide primers based on conserved sequences of DNA B components of begomoviruses were used in PCR for the detection of a potential DNA B. No amplification was observed with these primers from okra plants, while amplification of a product of expected size was obtained from plants infected with *African cassava mosaic virus*, suggesting the lack of a genomic component equivalent to DNA B. We have reported previously that monopartite begomoviruses on cotton and *Ageratum conyzoides* in Pakistan are associated with a disease complex involving a DNA component termed DNA 1, which shows homology to components of nanoviruses that encode the replication-associated protein. Recently, another molecule, DNA β, has been identified, associated with Ageratum yellow vein disease from Singapore and with cotton leaf curl disease (CLCuD) from Pakistan. These molecules are DNAs satellite and are essential for the development of typical disease symptoms in their respective hosts. Duplicate blots were probed for the presence of DNAs homologous to DNA 1 and DNA β (using full-length clones of these molecules isolated from CLCuD originating from Pakistan and washed at medium stringency. The probes detected bands hybridizing to DNA 1 in extracts from infected okra plants but not DNA β. No hybridizing bands were detected for either probe in extracts from healthy okra. A pair of primers, designed to conserved sequences in DNA β molecules, were used in PCR for the amplification of DNA β from symptomatic plants. The use of these primers amplified a product of the expected size (approximately 1.35 kb) from extracts of infected okra plants. The amplified DNA was cloned in TA cloning vector and labeled with α32PdCTP. The use of this as a probe detected the presence of a hybridizing band in infected okra plants, while no signal was observed in extracts from cotton plants showing symptoms of CLCuD. These results show that OLCD in Pakistan is associated with a DNA β molecule that is distinct from that reported on cotton and *Ageratum*. In particular, the DNA β of CLCuD and OLCD originating from Pakistan are sufficiently diverse not to cross-hybridize under the conditions used here, and are most likely different disease complexes. To our knowledge this is the first report of the association of a whitefly-transmitted begomovirus/DNA 1/DNA β complex with okra leaf curl disease.

1.14 Diagnostics for cotton leaf curl virus: production of antisera against the coat protein and the rep protein expressed in *E. coli*

The efficient control of plant viral diseases heavily relies on easy and reliable detection of the casual pathogen. Out of three main options available for the diagnosis of plant viruses; serological techniques employing polyclonal antibodies are easy and straight forward. The polyclonal antibodies are easier to obtain and have good specificity when made with highly purified antigen, therefore are usually better choice for the field-testing of *geminivirus* infection.

Based on the DNA sequence data, the PCR cloning strategy was used for the in frame cloning of coat protein and rep associated protein genes in pET-32a vector system (Novagen, USA). Positive clones were transformed into the *E. coli* strains *BL21 (DE3)* and *AD494 (DE3)* and transformants were selected through induction by the addition of IPTG (isopropyl-β thiogalactopyranoside) to the growing cultures. To determine the maximum protein expression, a time course study was performed on SDS poly acrylamide gel electrophoresis (SDS-PAGE). At the end of 3 hrs induction, the cells were collected and disrupted by sonication and the expressed protein was purified through affinity chromatography (utilizing histidine tag incorporated on the N-terminal of recombinant protein), under non-denaturing conditions (Novagen, USA). The column purified protein sample was dialyzed twice against 50 mM Tris-HCl overnight in dialysis tubing with a molecular cut off size of 6 kDa.

A pair of 3-4 months old female albino rabbits was injected subcutaneously with approximately 500 µg of each of the purified protein, mixed thoroughly with an equal volume of Freund's incomplete adjuvant (Sigma) to make a milky emulsion. Three injections were given with a gap of one week. The booster dose was prepared by mixing equal volume of Fruend's complete adjuvant (Sigma) with protein sample. Blood was collected one week after the booster dose from the immunized rabbits for determining the antibody titer. A 5 ml aliquot of blood was taken by the heart puncture and transferred into the sterile screw caped test tube. The antibodies were isolated from serum. An amount of 400-600 µl clean serum was obtained per 5 ml of the blood. Dilutions of this stock antiserum were made and equal volume of the diluted serum was mixed with equal volume of antigen in a microtiter plate. After sealing with parafilm the tray was incubated at 37 ^{0}C overnight. The positive agglutination indicated by granular clump formation leaving clear supernatant, was read and recorded under the microscope. Negative agglutination was indicated by settling of cells in the bottom of well and turbid supernatant.

The activity of the antisera was tested using three approaches:

1. Cross-reacting the antiserum with cell sap extracted from healthy and infected cotton plants. A 60-µl fraction of cellular extract was mixed with an equal volume of antiserum dilution and placed at 37 ^{o}C and observed for positive and negative reaction.
2. Western blotting against total cellular proteins of *E. coli* host cells expressing AV1 and AC1 proteins.
3. Immunosorption electronmicroscopy.

Coat protein (AV1) and replication-associated (AC1) proteins were expressed as 48 kDa and 58 kDa fusion proteins, respectively as indicated by SDS PAGE analysis of

the total cellular proteins during the time course study. Maximum expression for these proteins was observed, 3 hrs after the addition of IPTG to the growing cultures. The purified fusion protein when cut with the enterokinase yielded about 30 kDa coat protein and 40 kDa rep protein, which are in lines with predicted molecular weights. All the cellular proteins were efficiently removed in the process of purifying of these proteins through affinity chromatography. Bradford assay was used to determine the protein concentration of these proteins in the purified extract. The protein concentration of the coat protein in purified extract ranged from 92-108 µg/ml. The final protein concentration for the rep protein was in the range of 40-55 µg/ml.

The antisera titers were determined against the purified proteins. Anti coat protein and anti rep serum showed a titer of 1:12800 and 1:6400 respectively, when reacted with purified antigen. Anti coat protein serum raised against the protein detected protein expressed in *E. coli* during Western blot analysis.

The antibodies produced against the coat protein and the rep protein were tested with infected and healthy plants to find out their practical applications. Fresh emerging leaves of both infected and healthy plants were extracted with 1X PBS. A strong reaction with high titer (1:6400, almost the same as with the purified protein) was observed for antisera raised against the coat protein. No false reaction was observed, as all the control wells were free of any precipitation. The antisera against rep showed positive reaction only against the purified protein at a titer of 1:6400, which indicate a very mild agglutination. Immunosorbant electron microscopy by using the anti coat protein serum detected the geminate particles in the infected cell sap, which were not visualized without antibody coating.

1.15 Diversity and distribution of begomoviruses associated with Malvaceous spp., Including cotton

PCR and sequencing for core Cp and CP to achieve provisional identification

Polymerase chain reaction (PCR) was applied to detect and establish provisional identity of begomoviruses through amplification of a ~575 bp fragment of the begomoviral coat protein gene (*CP*), referred to as the 'core' region of the CP gene (core *CP*). The core *CP* fragment contains conserved and unique regions, and was hypothesized to constitute a sequence useful for begomovirus classification. Virus relationships were predicted by distance and parsimony analyses using the A component (bipartite viruses) or full genome (monopartite viruses), *CP* gene, core *CP*, or the 200 5'-nucleotides (nt) of the *CP*. Reconstructed trees and sequence divergence estimates yielded very similar conclusions for all sequence sets, while the *CP* 5'-200 nt was the best strain discriminator. Alignment of the core *CP* region for 52 field isolates with reference begomovirus sequences permitted provisional virus identification based on tree position and extent of sequence divergence. Geographic origin of field isolates was predictable based on phylogenetic separation of field isolates examined here. A 'closest match' or genus-level identification could be obtained for previously undescribed begomoviruses using the BLAST program to search a reference core *CP* database located at our website and/or in GenBank. Here, we describe an informative molecular marker that permits provisional begomovirus

identification and classification using a begomoviral sequence that is smaller than the presently accepted, but less accessible *CP* sequence (Brown *et al*., 2001). We have examined several hundred samples for the project and have obtained the core CP and/or CP for all for which amplicons could be obtained. In the final report, a phylogenetic tree and GenBank Accession numbers will be provided so sequences may be accessed and diversity and distribution of malvaceous-infecting begomoviruses may be documented in detail.

1.16 Infectious clones and molecular characterization of new and emergent begomoviruses of cotton

Begomoviruses of Malvaceous hosts: hollyhock and okra and associated satellite DNAs in Egypt

The apparent full-length, monopartite genome was cloned and the DNA sequence was determined for Hollyhock leaf crumple-associated virus (HLCrV) from hollyhock (2755 nt) and for Okra leaf curl-associated virus from okra (OkLCV) (2764 nt) exhibiting leaf crumpling and curling symptoms, respectively. The HLCrV and OkLCV genomes contained six predicted open reading frames (>10 kDa) as is characteristic of other monopartite begomoviruses. There was no evidence for a DNA B component in symptomatic hollyhock or okra plants. Phylogenetic analysis for HLCrV and OkLCV genomes with well-studied begomoviruses indicated that they are members of an Old World clade, also containing Cotton leaf curl virus from Sudan (CLCuV-SD), with which they shared 84% and 95% nt identity, respectively. In addition, HLCrV was 98% identical to a begomoviral genome cloned independently from *Althaea rosea* in Egypt {AF014881}, provisionally, Althaea rosea enation-associated virus (AREV). And, OkLCV shared 99% identity with a begomoviral DNA cloned from okra exhibiting enation symptoms, also from Egypt, provisionally, Okra enation virus (OkEV) {AF155064}. Interestingly, plants with which AREV and OkEV were associated exhibited foliar enations, and not leaf crumpling and curling, as was observed for HLCrV and OkLCV, respectively. Based on the 90% species rule, CLCuV-SD and OkEV may be considered distinct begomoviruses, despite an unusually high shared nt identity, if the atypically the lower (88%) shared *CP* nt identity and the basis for host range differences are taken into account. On the other hand, OkLCV is 99% identical to OkEV, suggesting that they are isolates of the same begomoviral species, even though distinct symptoms occurred in okra source plants. Also, satellite-like DNAs (sat DNA) ranging in size from 741-1350 nt were cloned from hollyhock-HLCrV: sat 65 (741 nt), and from okra-OkLCV: sat 3, 10, and 27 (744-1350 nt). Phylogenetic analysis indicated that OkLCV-sat 3 and 10 shared 94% identity, and that OkLCV-sat 10 and 27 were 88% identical to HLCrV-sat 65. Collectively, these satellite-like DNAs were less than 38%, 41%, and 23% identical to the sat DNA associated with *Ageratum yellow vein virus* {AJ252072}, CLCuV-PK sat DNA {AJ298903}, and *Tomato leaf curl virus* AU {U74627}, respectively. It is hypothesized that different individual or suites of sat DNAs, together with an OkLCV/OkEV or HLCrV/AREV genome, may be responsible for distinct symptoms observed for the same host species. Further, unique sat DNA(s) associated with HLCrV/AREV from hollyhock may further distinguish these diseases from leaf curl of cotton, with which their closest relative (~95% nt identity), CLCuV-SD, has been associated.

Cotton leaf curl virus-Sudan

Genomic-size begomoviral DNAs (2761 bp) were cloned and the sequences obtained for natural population of isolates exhibiting leaf curl and vein thickening symptoms in cotton, okra, and *Sida alba* from Gezira, and in okra from Shambat. Attempts to detect a B component were unsuccessful. Comparison of the four apparent full-length begomoviral DNAs revealed 99.3-99.5% shared nt identity, indicating that they are the same viral species, hereafter, referred to as Cotton leaf curl virus-Sudan (CLCuV-SD). Host range studies revealed that the wild type okra isolate of CLCuV-SD was transmissible from okra by the whitefly vector to okra, M. parviflora, and hollyhock, but not to cotton, whereas, the cotton isolate of CLCuV-SD infected cotton and hollyhock, but not okra. The genome organization of CLCuV-SD resembled that of other monopartite begomoviruses and encoded six (predicted) open reading frames (ORFs). The intergenic region (IR) sequence for CLCuV-SD DNAs was identical and contained a putative Rep-associated protein (Rep) binding element, GGTACTCA. Comparison of the Rep binding element with other begomoviruses indicated 100% identity with Okra leaf curl virus from Egypt (OLCV-EG), with which it shared its highest nt identity at 95.5%, and Tomato yellow leaf curl viruses from Israel, Japan, and Portugal, all begomoviruses from the Eastern Hemisphere. Recombination analysis revealed one major and one minor intermolecular sequence exchange in CLCuV-SD and OLCV-EG genomes, in which the majority of the genomes are apparently the same (99% shared identity), whereas, the coat protein (CP) ORF and a small region in the IR were replaced with sequences for either one and possibly two different parents. The CP ORFs of CLCuV-SD and OLCV-EG share 87.9% nt identity, suggesting that they are separate species based on the <90% ICTV rule. BLAST analysis for CLCuV-SD and OLCV-EG CP ORFs identified no prospective extant parent of either CP ORFs, and constraint analysis of the CP ORF for CLCuV-SD isolates indicated the ORF exhibited high constraints to divergence. Collectively, results suggest that replacement of the CP ORF through recombination could have involved by at least two different mechanisms. In the first, a simultaneous (or nearly so, in evolutionary time) substitution of CP and IR sequences originating from two distinct genotypes occurred, thereby immediately or ultimately, eliminating the original parent virus. In the second model, a single CP/IR sequence replacement occurred, followed by differential divergence of the CP ORFs, and since then, divergence of CP ORFs of CuCLV-SD and possibly OLCV-EG has been highly constrained. Irrespective of evolutionary histories, these viruses appear highly host adapted to the Malvaceae and may, therefore, be predicted to be highly recalcitrant to change.

Infectious clones for four *Cotton Leaf Curl Virus*-Sudan associated DNAs and four satellite DNAs from three Malvaceous hosts

Full-length clones of *Cotton leaf curl virus*-Sudan (CLCuSDV) (2.8 kb) were obtained from symptomatic cotton, okra-SH, okra-GZ, and *Sida* spp. and their DNA sequences were determined. The four viral genomes were found to be less than 3% divergent, indicating the begomoviral component associated with symptomatic source plants is of the same virus (Idris and Brown, *in press)*.

Prospective satellite DNAs (1.3 kb) from cotton, okra and *Sida* spp. were amplified by PCR, cloned, and the sequences were determined. Phylogenetic analysis and sequence comparisons indicated the satellite DNAs were distinct from one another. In January, 2001, viral genomic and satellite DNA clones were sent to Dr. Rob Briddon at JI to carry out infectivity assays using biolistic inoculation of cotton (*Gossypium hirsutum*) and *N. benthamiana*. The *Sida* isolate of CLCuSDV was infectious to cotton and tobacco and typical leaf curl disease symptoms were observed in cotton co-inoculated with satellite DNA isolated from *Sida* spp.

When a full-length clone for *Cotton leaf curl virus*–Pakistan (CLCuPKV) and CLCuV-SD associated Sida spp. satDNA were co-inoculated to cotton, leaf curl disease symptoms were observed in cotton 10 days post-inoculation. This result indicates that both CLCuV-PK and CLCuSDV-Sida clones are infectious in cotton when present together with CLCuV-SD associated satellite DNA from *Sida*. Second, results provide preliminary proof of Koch's Postulates for CLCuV from Sudan. Experiments are underway to complete infectivity assays for all possible combinations of viral genomic components and satellites in cotton, okra, and Sida spp.

Monopartite Begomoviral DNAs Cloned and Sequenced

I.	Cotton SD 1	cotton Gezira
II.	Cotton SD 3	cotton GZ
III.	Cotton SD 4	cotton GZ
IV.	Cotton SD 10	cotton GZ
V.	Okra SD25	okra GZ
VI.	Okra SD28	okra GZ
VII.	Sida SD83	Sida GZ
VIII.	Sida SD84	Sida GZ

DNA Satellites Cloned and Sequenced

I. Cotton 41
I. Cotton 46
II. Cotton 53
III. Cotton 55
IV. Okra 43
V. Okra 44
VI. Okra 45
VII. Sida 29
VIII. Sida 31
IX. Sida 32
X. Sida 33

Molecular analysis of infectious clones for *cotton leaf crumple virus* (Arizona, California, USA; Sonora, Mexico) reveals it is a double recombinant between *Sida golden mosaic* and *bean calico mosaic viruses*

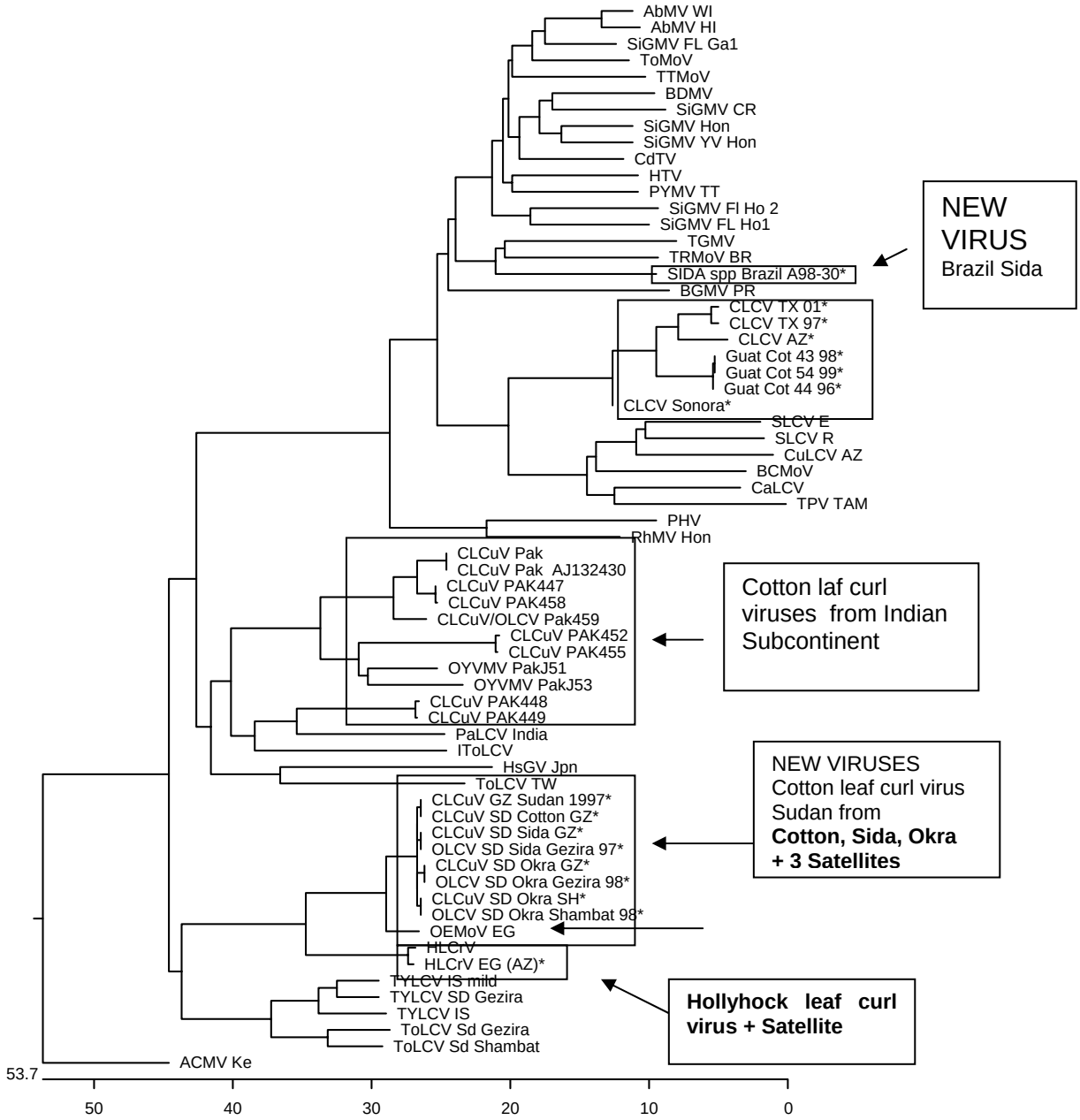
Cotton leaf crumple virus (CLCrV), a begomovirus occurring in Arizona and California, USA and in Sonora, Mexico, was shown to be a distinct bipartite begomovirus species. CLCrV-Sonora (CLCrV-Son) shares >97% nt sequence identity with the previously characterized isolate of CLCrV from Arizona (Brown et. al., 2001). Experimental and natural host range studies indicated that CLCrV has a relatively narrow host range within the *Malvaceae* and *Fabaceae* families. The genome of Sonora isolate, designated CLCrV-Son, was cloned and completely sequenced. Cloned CLCrV-son DNA A and B components were infectious by biolistic inoculation to cotton and bean, and progeny virus was transmissible by the whitefly vector, *Bemisia tabaci*, thereby completing Koch's postulates for the first time. CLCrV-AZ DNA A shared highest nucleotide (nt) sequence identity with *Bean calico mosaic virus* (BCMoV), *Cabbage leaf curl virus* (CaLCV) and, *Squash leaf curl virus* (SLCV-R), at 76%, 76%, and 75%, respectively. The CLCrV DNA B component shared highest nt sequence identity with *Potato yellow mosaic virus* (PYMV), *Tomato mottle virus* (ToMoV), and *Abutilon mosaic virus* at 67%, 66%, and 66%, respectively. Collectively, these results provide intriguing evidence that CLCrV is a double recombinant, having sequences from two distinct species, and forming a new prototype group among the New World begomoviruses. The *cis*-acting begomovirus replication specificity element, GGAGTCTGGAGC, in the CLCrV origin of replication is similar to that of *Pepper golden mosaic virus* (PepGMV), suggesting that reassortants among components of CLCrV and PepGMV may be possible. These results demonstrate that the limited hosts of CLCrV constituted a barrier for natural co-replication in the same host with other begomovirus which resulted in begomovirus isolates bearing distinct biological and molecular properties (Idris and Brown, *submitted*).

GEMINIDETECTIVE Website: URL: http:gemini@biosci.arizona.edu.

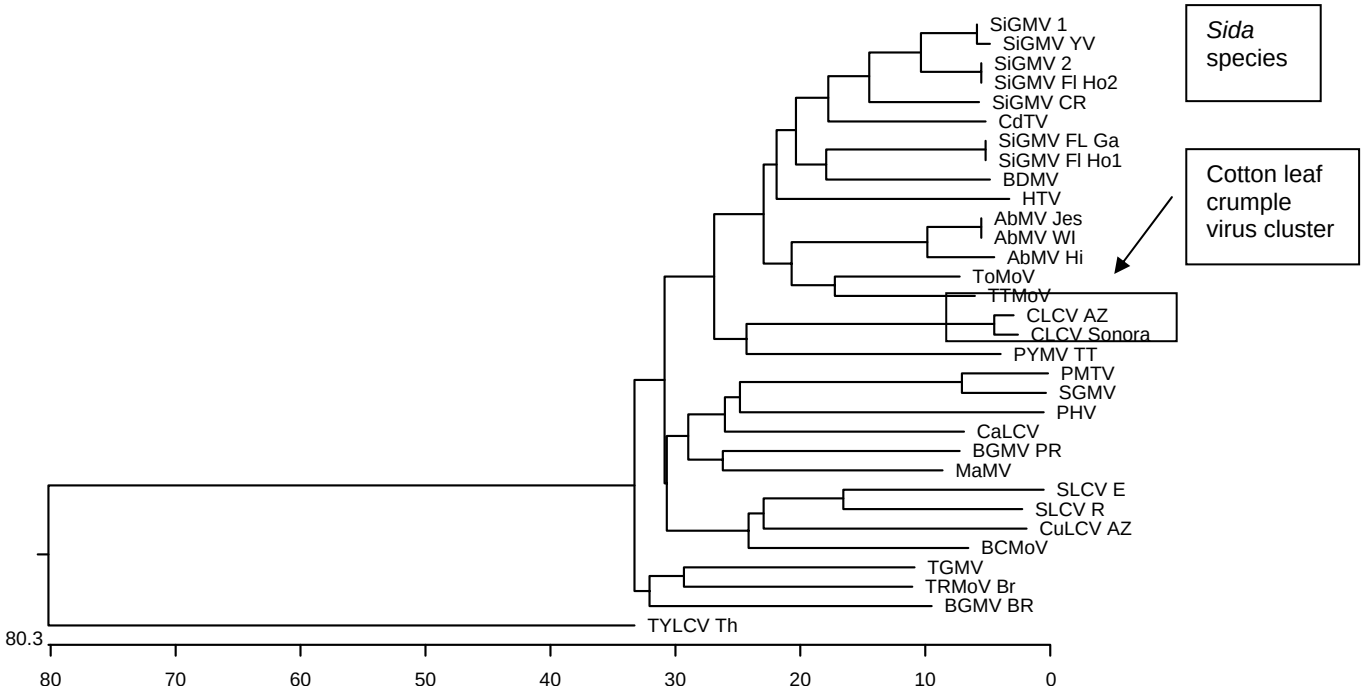
Enables user to establish provisional identification of whitefly-transmitted viruses by comparative analysis using the database for the core coat protein gene provided at the site; site also contains PCR primers and parameters for amplifying the core Cp sequence together with information regarding geographical distribution, disease symptoms, host range, and relevant references.

ARIZONA Field Trial 2000- Resistance to cotton leaf crumple virus

The varieties tested are listed by rank of lint yield. Different degrees of tolerance to foliar symptom development and yield loss were observed for cotton lines and varieties tested in Arizona in 2000. We observed significant variation for *Cotton leaf crumple virus* disease symptoms among the 12 lines for which data are available. The 'adapted' controls (susceptible Arizona varieties) produced higher lint yields and seed cotton yields than the introductions. Fiber quality of the introductions tended to be superior compared to the 'adapted' control cotton plants, but the value (i.e. how much the lint is worth per acre) of the control varieties was significantly greater than the introductions. One new observation was that there was no relationship between symptom severity rating and yield, even among the introductions.



Full-length A components or monopartite genomes for select cotton and malvaceous-infecting begomoviruses for which full-length clones and/or clones and their respective satellites (when they occurred) have been obtained.



Phylogenetic relationships of bipartite cotton and malvaceous-infecting begomoviruses based on the B component sequence. Cotton leaf crumple virus B component is more closely related to PYMV than to members of the SLCV cluster, as is seen for the A component of these viruses/strains.

Objective 2

Development of an *in vitro* system for cotton transformation

Summary

At the time of writing this project only few cotton genotypes amenable to regeneration were available. Seeds of those varieties were obtained to establish their tissue culture at NIBGE. Since transformation of cotton takes longer, construct for engineering virus resistance were first evaluated in tobacco. A large number of cotton genotypes grown in Pakistan were evaluated for regeneration response but none of those were found to respond to regeneration despite use of different conditions and hormone levels. A system of meristem culture was optimized for biolistic gun transformation system. Another approach adopted was to establish a system of *Agrobactrium*-mediated transformation of mature embryos of cotton. The transformation system although of low efficiency as applicable to all cotton genotypes. The system was used for transformation of available constructs in local elite varieties of cotton.

2.1 *In vitro* studies of cotton for callus induction, regeneration and transformation

In vitro, regeneration is the prerequisite for the genetic manipulation of plant. Regeneration is highly genotypic dependent process. In cotton few cultivars are responsive to regeneration, which included mostly exotic cultivars like Coker, Siokra and others. The latest gene transfer technology in plants that biolistic gun, requires callus, suspension cells, leaves, meristem tip or any other regenerable explant. Lack of regeneration capability of local elite cotton cultivars makes the meristem tip the most suitable alternate procedure for genetic transformation of cotton. Relatively with simple procedure apical tip produces shoot directly without any hazard of somaclonal variations and chromosomal abnormalities.

Four basal tissue culture media viz. MS, B5, N6 and LS were tested for callus induction response. The results indicated that MS medium was superior to all other three media combinations for tissue survival and callus induction. MSO medium without any growth regulator initiated callus on the cut surfaces of cotyledons and hypocotyls tissues but when subcultured on the same medium for further growth. The medium did not support further growth and maintenance of the callus died. This medium was further modified by the replacement of MS vitamins with B5 vitamins. This medium produced better callus as compared to the previous one with MS vitamins.

The four different types of explants were selected for callus induction, hypocotyls, cotyledons, radical and petiole. Cotton varieties NIAB-78, S-12, and Coker 312 were used in these studies. Twenty explants were cultured for each treatment. The results indicated that significant differences were present for callus induction from cotyledons, hypocotyls, roots and petiole tissues. Small callus of very low quality was produced on root tissues, which was discarded in the initial phases of the experiments.

Maximum callus was obtained from hypocotyls tissues followed by cotyledon and petiole tissues. Morphology of callus was also different from different sources. Hypocotyls tissues gave better quality callus as compared to cotyledon and petiole tissues respectively. Media formulation showed that on both media, tissues produced callus but in some cases cotyledonary tissues gave better callus on NAA and Kinetin medium. Generally 2,4-D and Kinetin media was found superior to NAA and Kinetin. In the initial stage of callus induction, non-significant differences were observed among genotypes but true embryogenic callus was observed only from Coker- 312 variety. The other two varieties did not produce embryogenic callus on these two mediums.

Twenty-two cotton lines were selected: *Gossypium hirsutum* L. NIAB-78, S-12, MNH-93, Gohar-87, FH87. FH-682, SLS-1, NIAb-26, BH-36, CIM-70, CIM-109, CIM-240, RH-1, B-557, A-1-85, A-18/87, AEM-52, Coker-312, Siokra 1-3, Siokra-324 and *Gossypium arboreum* L. Ravi In the first experiment hypocotyls, cotyledons and root sections of four cotton varieties viz, NIAB-78, S-12, Gohar-87 plants raised in culture tubes were excised aseptically and placed on MS medium supplemented with B5 vitamins and with the following hormone combinations (mg/l) 0.1, 0.5,1.0 and 2.0 2,4-D, 0.1 2,4D plus0.1 kinetin, 0.1 2,4-D plus 0.5 kinetin and 0.5 NAA plus 0.1 kinetin In the second experiment, explants from 22 cotton varieties were cultured on two defined media(mg/l of 0.1 2,4-D+0.5 kinetin and 0.1 2,4-D + 1.0 2iP). Callus produced was sub- cultured for maintenance and regeneration on 18 different media containing 2,4-D, Kinetin, NAA, BAP, 2iP and Zeatin. In the first experiment, where only four different cotton varieties were cultured on seven different media containing different concentration of 2,4-D hormone alone, no callus was observed from root tissues. At 0.1 and 0.5 mg/l 2,4-D media, hypocotyls tissues of all the varieties gave callus but at 2.0mg/l 2,4-D, no callus was observed from all the four varieties. When combinations of 2,4-D and kinetin were used for callus induction, steady callus proliferation was observed from all varieties and tissues with exception of Gohar-87 root tissues. Microscopic studies on transverse section of explants revealed actively dividing sub-epidermal cells as early as 6 days after the beginning of culture. These results showed that combination of 2,4-D and kinetin was better for callus growth in cotton. A variable response to callus induction was observed on NAA and kinetin media where callus was obtained from cotyledon tissues of NIAB-78, S-12 and MNH-93 varieties and root growth was observed from cotyledon and hypocotyls tissues of NIAB-78 and S-12 varieties.

In the second experiment, 22 cotton varieties were cultured on two defined media. These two media were selected from the preliminary experiments on callus induction and regeneration. Results indicated that all the varieties induced callus with variable responses. Highest callus induction was observed from Coker- 312 followed by coker-304, Siokra1-3 and Siokra-324 respectively. In the beginning 3-4 weeks, response to callus induction was equal in all the varieties and light green/ cream color callus appeared. In the later stages of growth, the differences appeared and varied widely from genotype to genotype. The cotyledon tissues produced single type of callus. This callus was found lower in quality as compared to hypocotyls callus.

The callus produced on both the above media were sub-cultured on 40 different combinations of MSO, MSk (MS+KNO_3 1.9 g/l), BAP, IAA, NAA, kinetin and 2,4-D and zeatin hormones. No callus growth was observed on IAA and BAP combinations.

Medium containing zeatin, callus remained green, dividing but no embryo was detected. In combination of Zeatin, IAA and Kinetin, callus became brown and dead.

Response to embryogenesis was restricted to Coker and Siokra lines. Ninety percent seeds of coker-312 showed embryogenic callus. . No response was observed from all other varieties.. In some varieties, like CIM-240 and NIAB-78, proembryo like structures were observed but did not mature to complete embryos. When the tissues were placed for longer time on the same medium or sub cultured to another medium for growth, the callus started browning, stopped growing and eventually died. Callus of Coker and Siokra lines was sub cultured on MSk medium on monthly intervals and alive for a period of 6-8 months. This callus gave embryos on solid and liquid medium.

Screening of 22 genotypes substantiates the genotype specificity of the embryogenic response in cotton. Coker-312 exhibits a higher index of embryogenesis under the same set of treatments than any other cultivars tested. The data from these experiments suggested that genetic components rather than culture components are the most critical factor in obtaining efficient regeneration in cotton.

2.1.1 Comparison of cotton varieties for *Agrobacterium* Transformation

Cotton varieties Coker-312, HS-200 and S-12 were subjected to the advanced procedures of *Agrobacterium* mediated transformation. Two strains LBA4404 and AGL1 harboring p35S GUS INT plasmid. Cotton hypocotyls inoculated with *Agrobacterium tumefaciens* developed callus after 3 weeks on kanamycin medium. The first morphological change was formation, which was observed after 6-7 days of inoculation. Within two weeks, enough callus was visible growing directly out of hypocotyl tissues. The number of callus produced after 3 weeks varied from 3-5 per hypocotyl. Tissues transformed with Ti- vector showed rapid growth similar to that control (without kanamycin). While all cultivars were susceptible to *Agrobacterium* strains, there was a difference in the percentage of explants developing transgenic callus depending upon the explants and cultivar source and on the transformation conditions. AGL1 strain showed higher rate of transformation than LBA4404. The transformed callus was subcultured on kanamycin and GUS expression was studied. The results indicated the transformed callus turned blue. The embroils and plantlets were obtained from this transformed callus, which also showed GUS expression (Fig: 4-9)

2.1.2 Response of different cotton cultivars for meristem tip culture

For meristem tip culture, seedlings of AEM- 52, N-26, NIAB-78, S-12, FH-87, FH-634, FH- 682, MNH- 93, A-185, A-18/87, Coker –304, GOHAR-87, CIM-70, CIM-240, CIM-109, B-557, SL-41, BH-36, RH-1and *G. arboreum* L. RAVI were grown aseptically for 2-10 days to get the meristem tip. Under the dissecting microscope, meristem tip was removed and placed on shooting medium (1/2MS, MS, MS+0.46mM Kinetin, MS+0.93 mom Kinetin, MS+ 0.45mM 2,4-D+ 2.32 mom Kinetin and MS+ 0.45 mM2, 4-D+ 2.46 mom 2iPand subsequently on rooting medium following combinations; 1/2MS, MS, MS+ IBA (0.98mM, 2.46mM, 4.92 mom, 9.84mM and 24.60 mom and Ms + 2.68 mom NAA + 0.46 mom Kinetin). Rooting plants were transferred to sterile sand and to the soil. Unlike many other crop

plants, which can be grown easily in tissues, cotton has proved to be very difficult. With the addition of sucrose and agar in media, tissues cells readily became brown/black, watery and ultimately lose the capacity of division. The problem was mitigated by the replacement of sucrose and agar media with media containing glucose (30g/l), philately (2.5g/l) and MgCl2 (0.75g/l). Basal portion of tips did not turn brown and grew normally. In the first experiment, three cultivars S-12, NIAB-78 and MNH-93 were tested on 28 different medium and growth regulators. As the media formulation was very diverse, response was also quite different. Size of meristem/ shoot tip also contributed significantly to the rate of plant formation. For the first two-week tips grew normal on all the 28 media. In later stages, some proved detrimental and caused death of tips tissues. On1/2 MS, MS, MS+0.46 mom Kinetin, MS +0.93mM kinetin and MS+2.32mM kinetin, shoots grew but callus and root formation was not observed. On media containing 2,4-D along with kinetin and 2iP, callus developed at the base of the meristem tip, tip growth was normal in the first two weeks and then retarded, leaves shed and tip degenerated. Media containing higher concentration of kinetin did not show good response. Tip cultured on media containing IAA (11.41 mom) with kinetin (2.32 mom) initiated roots from S-12, NIAB-78, and MNH-93 but callus was also present at the base and tip growth was also present at the base and tip growth was not good as on simple kinetin (0.46 mM and 0.93 mM) media. In media containing NAA (10.74 mm) and Kinetin (0.46 mM and 2.32 mM) rooting response was higher than shoot development. Callus was also present at the base. Very little shoot growth was observed but not rooting on the media containing NAA (5.37 mom) and 2iP(2.46). Highest rooting was observed when (2.68 mom) with kinetin (0.46) was used. All three varieties cultured on this concentration developed roots. In the light of these results, all other 16 cultivars were cultured on 6 defined media. MS simple medium, MS with 0.46 mom and 93mM kinetin were found highly suitable for shoot tip growth in cotton Kinetin enhanced the division of tip cells and favor the formation of leaves. Media containing 2,4-D and 2iP enhanced the callus formation but no root formation respectively. Shoots from 10 cultivars, which were developed on kinetin media, were transferred to rooting medium. All the varieties developed roots only on media containing 2.68 mom NAA and 0.46 mom kinetin. Shoot tips developed roots more readily than hypocotyls segments, which indicates the response of leaf, and tip material for root development. The results of meristem shoot tip indicated that the size of meristem/shoot tip contributed significantly to the rate of plant formation. Mortality rate was 50% when meristem of less than 0.5 mm size. The results of these experiments revealed that media containing 0.46 mom kinetin was better for shoot development whereas rooting was observed on media containing 2.68 mM nephthalene acetic acid (NAA) and 0.46 mM kinetin. No intervarietal variability was observed. We concluded that the methodology is simple and replaces the existing protocols for meristem tip culture of cotton (Figure 2.10).

2.1.3 Biolistic transformation of cotton

The biolistc process represents a completely new approach to the problem of how to deliver DNA into intact cells and tissues. A wide variation of biological and physical parameters like quality and quantity of DNA, quality and age of spermidine, quality of helium, quality of ethanol, quality of NDA precipitation, particle size, gap width, bombardment pressure, age of explant, target distance, osmoticum and cultural condition were optimized for particle bombardment. Cotton varieties DP 50, Acala-

15/17, CIM-240, NIAB-78, S-12 and Coker 312 seeds were used for these experiments. Meristems and mature embryos were used as intact explants. Results of the experiments with meristems showed that the week GUS expression was observed on these tissues but the main thing is that these tissues were fully covered with small hairs which are present on the cotton plants and on the meristem area, the number of hairs are much higher. It was found that this explant is not suitable for transformation studies. Some of bombarded mature embryos were stained for GUS activity. Some of the tips were found completely blue while in other cases it was counted that only 20 spots were enough to complete the blue tip of embryo. The maximum GUS expression was obtained when bombarded at 9 cm target distance.

In other experiments, split meristems were used in the bombardment studies. Explants of cotton varieties Acala 15/17 and CIM-240 were splited and shot with plasmids pCAMBIA2301 (GUS+NPT11) and pBI221 (GUS). The shot tissues were put on kanamycin for a period of three months with monthly sub- culturing. Small leaf from each plant derived from shot meristem tips, which survived and grew on kanamycin medium were checked for GUS activity. The results showed the GUS expression in the leaf tissues. Furthermore cross sections of leaf and roots also revealed the Gus expression. (See Figure)

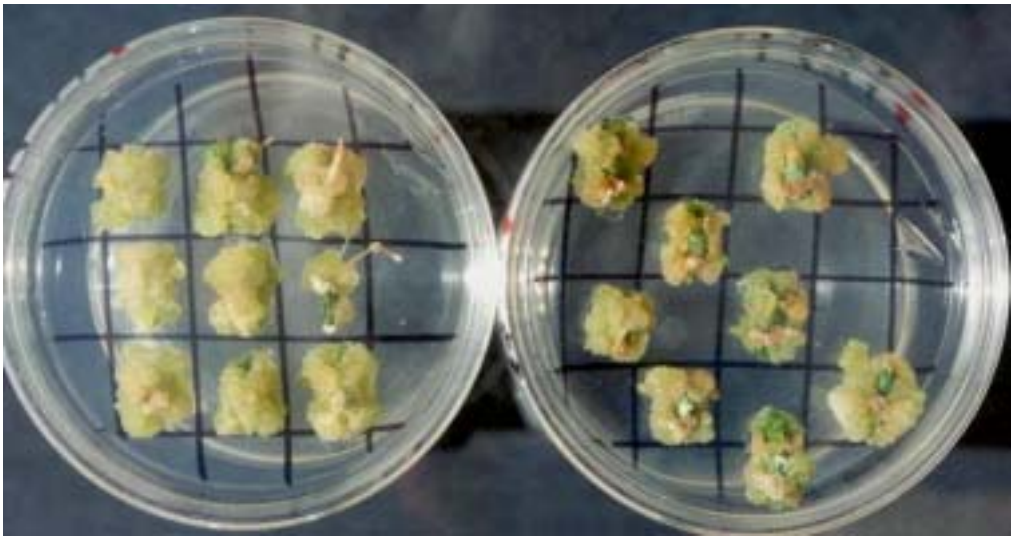


Figure 2.1 Hypocotyl sections of cotton varieties Coker- 312(left) and S-12 cultured on MS medium containing 0.1mg/l 2,4-D and 0.5mg/l kinetin

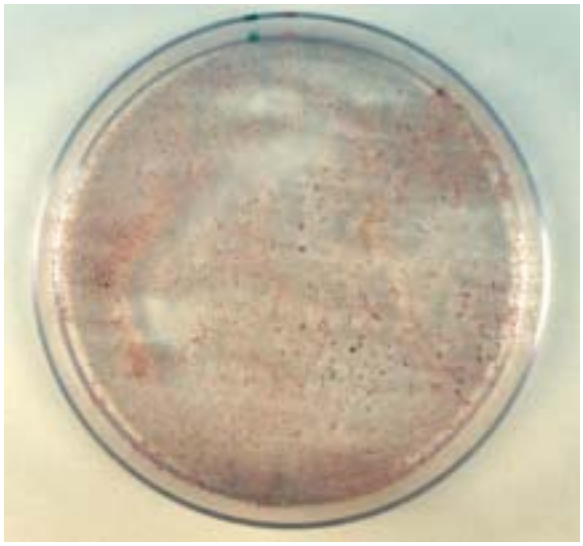


Figure 2.2 Dense plating of suspension culture of cotton variety Coker-312 on embryo maturation medium

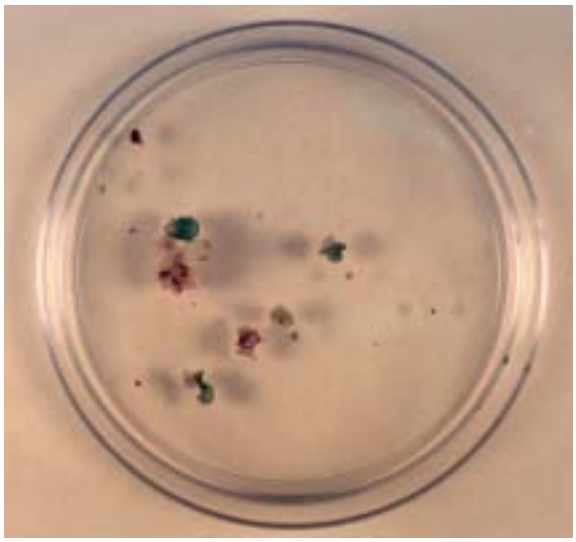


Figure 2.3 Thin plating of suspension culture of cotton- 312 on embryo maturation medium

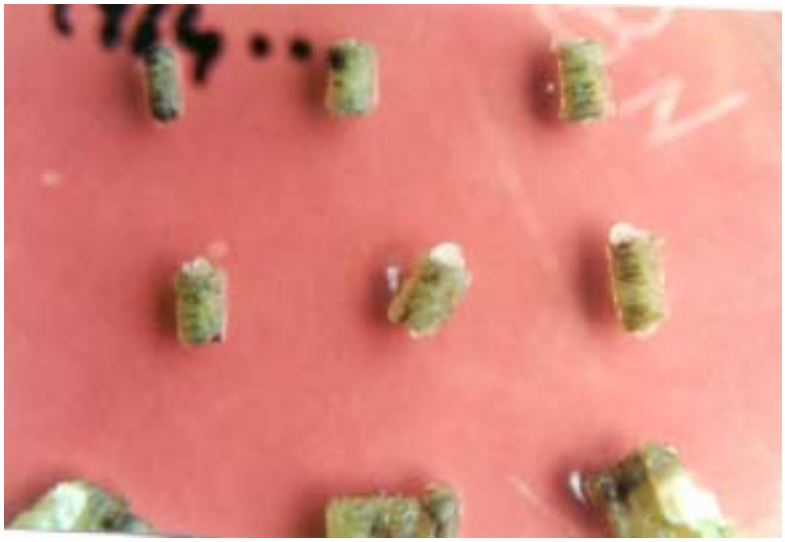


Figure 2. 4 Transgenic cotton coker-312 appearing from hypocotyls section on selection medium

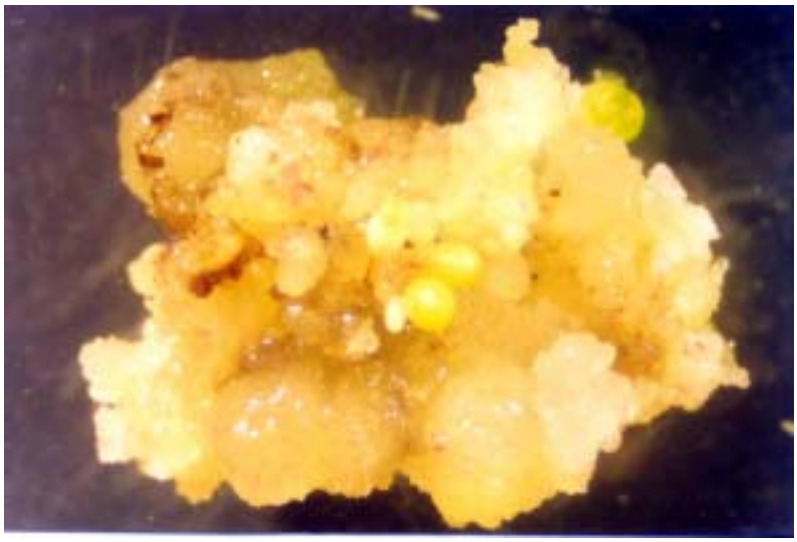


Figure 2.5 Transgenic callus of cotton Coker-312 with large number of somatic embryos and plantlets

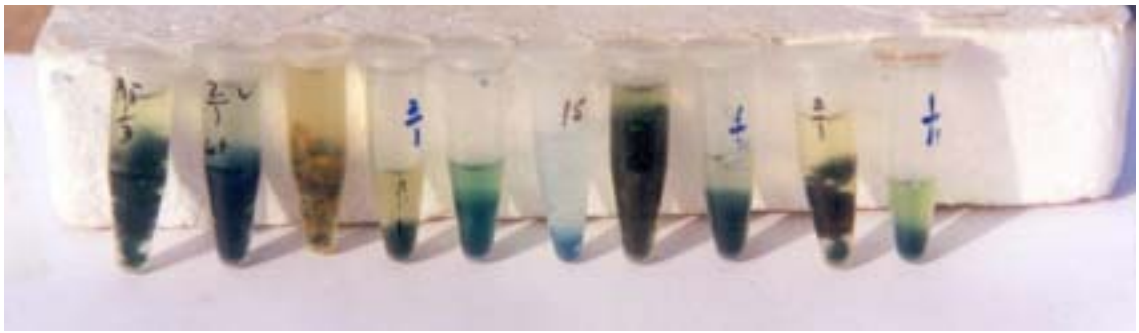


Figure 2.6 Callus and embryogenic tissues of independent transformants of Coker-312 with GUS expression

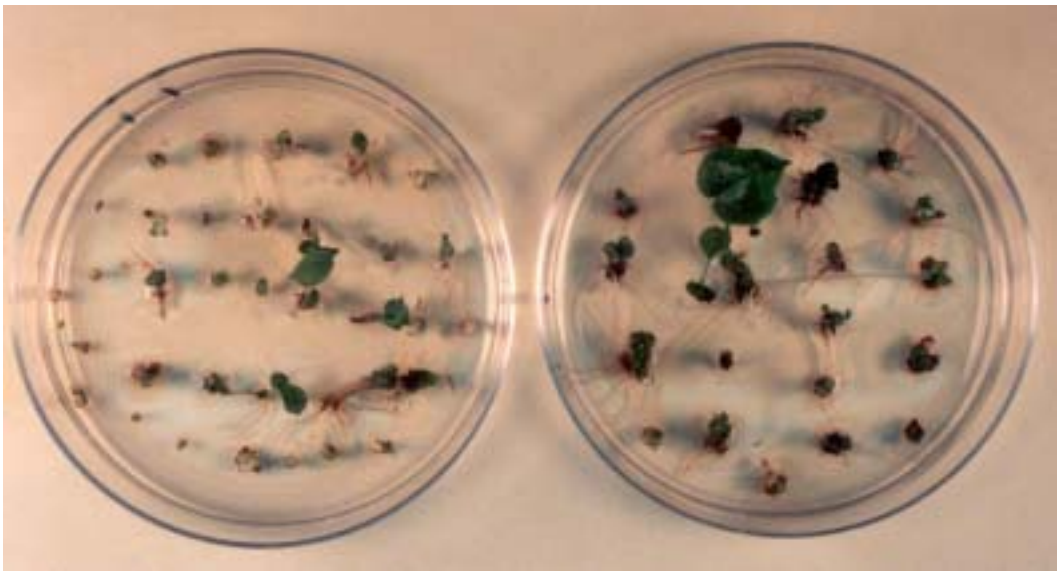


Figure 2.7 Somatic embryos of cotton variety Coker –312 cultured on germination medium showing the presence of shoots and roots on selection medium

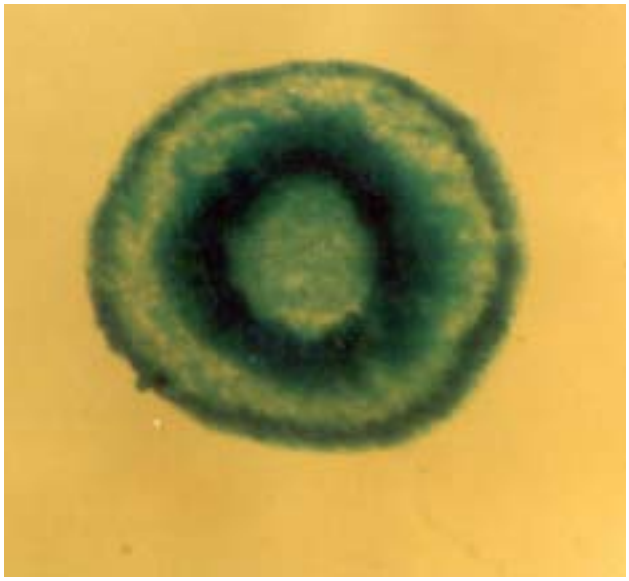


Figure 2. 8 Cross section of stem tissues of transgenic Coker –312 showing GUS expression

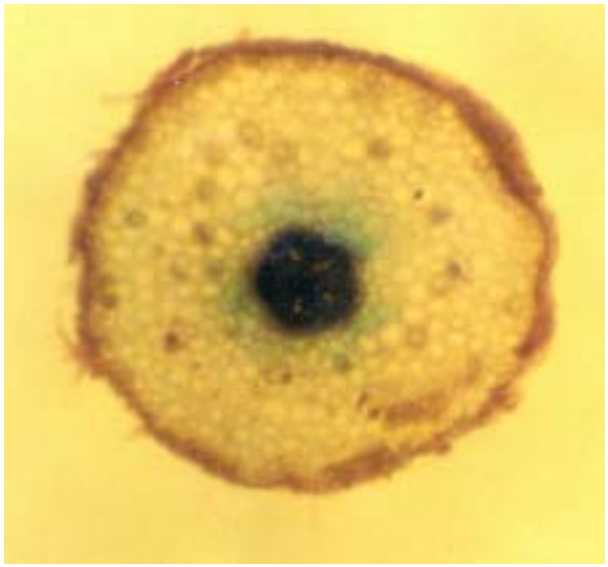


Figure 2. 9 Cross section of root of transgenic Coker 312 showing the expression of GUS

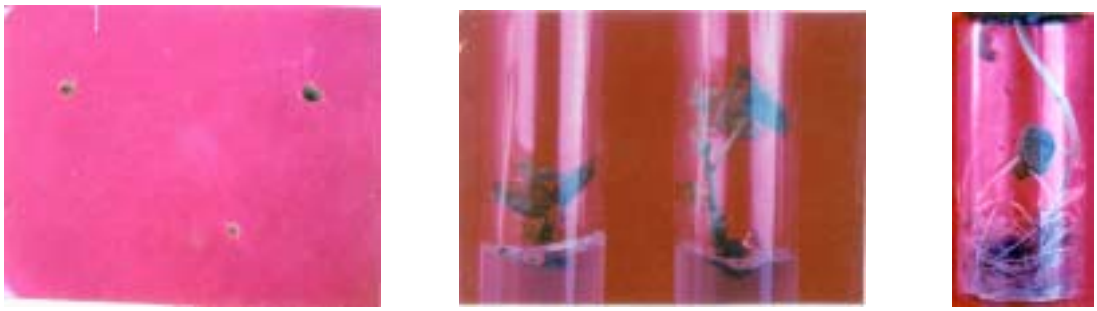


Figure 2.10 Meristem shoot tips (left), different stages of shoot growth (middle) and plantlet root formation (right) of cotton elite cultivars NIAB-78.

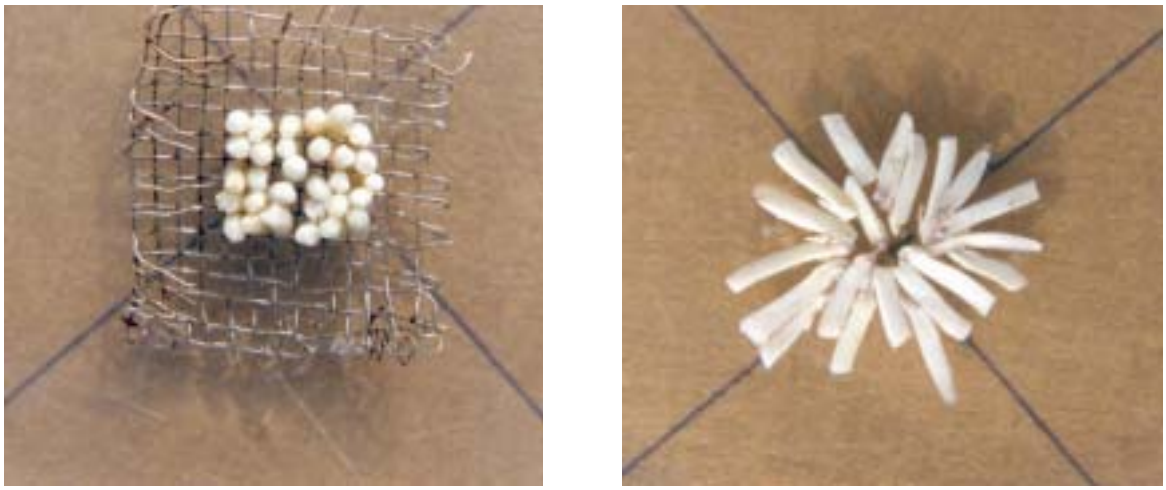


Figure 2.11 Arrangement of mature /split embryos for biolistic bombardment

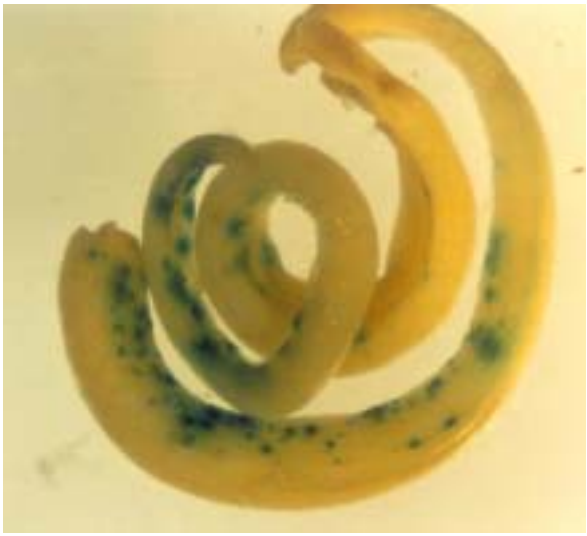


Figure 2.12 Split embryos showing GUS spots after Bombardment

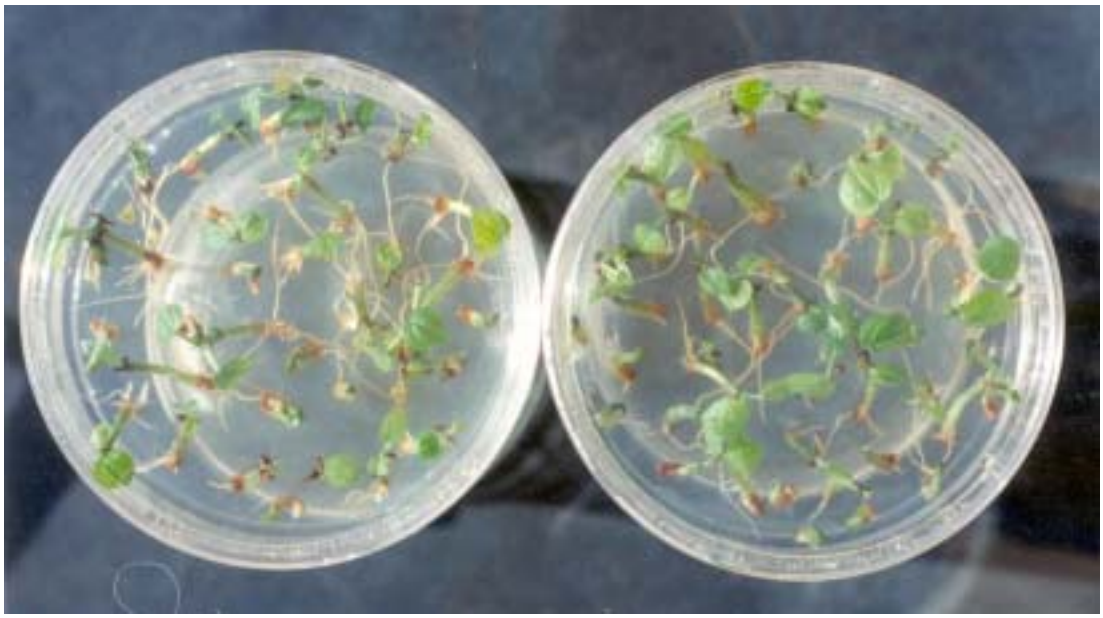


Figure 2. 13 Shoots and roots appearing from mature embryos after bombardments

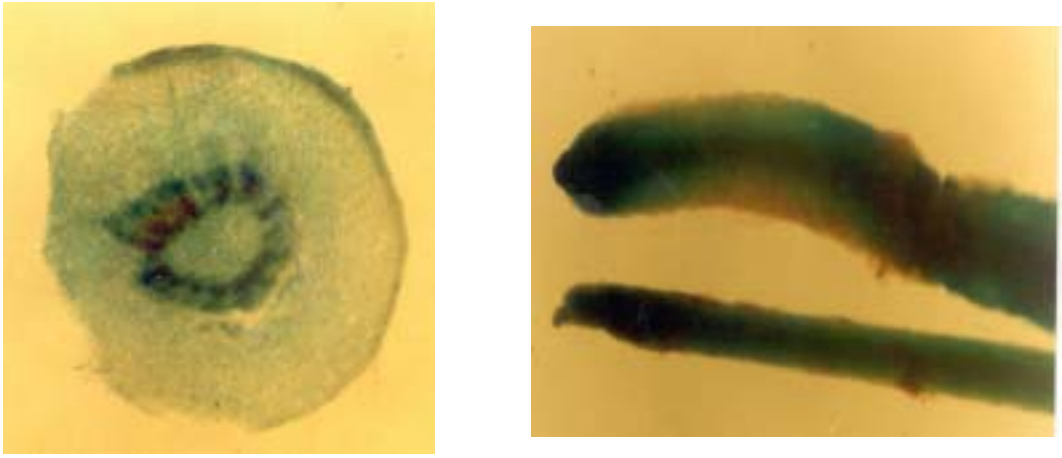


Figure 2.14. Stem section (left) and root section (right) of transgenic plants derived from splitted embryos of CIM-240 after bombardment

Objective 3

Development of virus resistant cotton genotypes

Summary

Both conventional breeding and genetically engineered resistance approaches were used for the development of virus resistant elite cotton genotypes. Identification of sources of natural resistance and their confirmation by available tools was the first step in selection of sources of resistance. This was an important task because previously only field evaluation was used to confirm resistance. It was confirmed that LRA5166 and CP15/2 are good sources where no detectable levels of viral DNA are present when tested by dot-blot hybridization.

At the time of writing of project two types of begomoviruses were known; bipartite begomoviruses having DNA A and DNA Band monopartite begomoviruses homologous to DNAA of bipartite begomoviruses. The approaches for engineering resistance at that time were based on inhibition of replication by antisense RNA or interference with the use of defective interfering DNA. The project identified that several distinct monopartite begomoviruses are associated with the disease that requires a single species of DNA beta to develop symptoms. Based on our understanding of cotton leaf curl disease new approaches were adopted to create genetically engineered resistance. These new constructs were evaluated in tobacco to assess their ability to suppress the disease.

3.1 Genetic engineering of tobacco/cotton for CLCuV resistance

A variety of strategies have been employed to engineer virus-resistant transgenic plants. One exploits the natural phenomenon of cross-protection but unlike with RNA viruses, has had limited success with DNA viruses. For, begomoviruses, expression of truncated defective transdominant viral coat protein, replicase and movement proteins has proved more promising. Another approach is to express antisense transgenes complementary to a target mRNA. The original rationale of antisense RNA technology was that by pairing with a complementary target mRNA, antisense would inhibit expression of homologous genes by preventing translation or promoting degradation of the target mRNA. Indeed this technology has been successfully applied to engineering resistance to geminiviruses. However antisense is actually part of complex natural pathways for gene regulation by homology sensing mechanisms where sense transcripts are also able to silence gene expression.

A 1.5 kb Sst I/Xho I fragment containing the enhanced CaMV 35S promoter and poly A tail, flanking a Sma I site was subcloned from plasmid pJIT60 (gift from Dr. P. Mullineaux, John Innes Centre, UK) into the pBluescript II KS+ to yield pSQW1 as an plant expression cassette. CLCuV DNA A genes were amplified by PCR [cycle 95ºC, 5 min, then 40 cycles of 95ºC, 1 min, 50ºC, 1 min, 72ºC, 1 min, then 72ºC, 10 min] of CLCuV Pak2 using the primer pairs: [i] for D1/4, 5'-AGTCAACATGCCTCCAAAGC-3', 5'-AGCTAGTTCCTTAATGACTC-3' [ii] For D1, 5'-ACTAGCTCCTAAAGAT TTTG-3', 5'-AAGATCGCATTCTTTACTCG-3'and [iii] For D1/d2/d3, 5'-TGCAATCTTCATCAGCCTC3' 5'-AAGATGATTGGTCTACAAATAC-3', end filled with T4 DNA polymerase and

subsequently cloned in sense and antisense orientations into the Sma I site of pSQW1. The orientation was determine by restriction sequences analysis The 6 expression cassettes, sense and antisense of D1/4, D1 and D1/d2/d3, were then each individually sub- cloned as Sst I/EcoR V fragments into the Sst I/Hpa I sites of pGA482. The plasmid recombinants were transformed into *Agrobacterium tumefaciens* strain LBA4404 by electroporation.

Nicotiana tobacum (Samsun) was transformed using leaf disc as explants. About 15-20 leaf discs were used for each construct in each experiment. Experiments were carried in three- four different batches. The putative transgenic plants were obtained on MS medium supplemented with kanamycin (100 ug/l) and cefotaxime (250mg/l) and subsequently transferred to soil for getting the progeny. T_0 lines were self-pollinated and T_1 seeds germinated on MS medium containing 500ug/ml of kanamycin and T_1 seedling were transplanted into soil a month after germination. T_2 and T_3 seedlings were similarly produced. *Gossypium hirsutum* L, cotton, cultivar Coker 312, transformation and regeneration were carried out using hypocotyl as explants. All of the six constructs were used to transform the cotton Coker -312 in four to five different batches. Transformants carrying the D1/4, D1, and D1/d2/d3 antisense transgenes were transferred to soil and grown as T0 lines. Fertile transgenic lines were self-fertilized to produce T_1, T_2 and T_3 progeny as above while the regenerants are being collected from others constructs.

Local elite cultivars of cotton S–12 and NIAB –78 were selected on the basis of their susceptibility to CLCuD in the initial experiments while in later experiments, attention was focused to transform S–12. Unfortunately no local variety is responsive to regeneration in tissue culture due to their genotype specificity. For transforming the genes in these cultivars, mature embryos were inoculated with *Agrobacterium* LBA4404 harboring the specific transgenes. For each construct 40-50 thousands embryos were agro- inoculated. In one of the initial experiments, biolistic transformation of D1 antisense construct was carried out using mature embryos (20,000) as explants. Transformants were obtained on kanamycin (50 mg/l) and cefotaxime and transferred to soil for getting the selfed T0 and subsequently T1, T2 and T3 generations.

System was optimizes for rearing viruliferous whiteflies on CLCuD-infected cotton plants under containment. Seedlings were exposed to viruliferious whiteflies at the four to five leaf stages for 120 days at 28-30 ^{0}C for tobacco and 32-35ºC for cotton. The presence or absence of symptoms was observed on weekly basis.

The presence of transgenes was analyzed by PCR using specific primers for CLCuV DNA A transgenes and nptII on the isolated total genomic DNA. Southern blotting was performed using 15µg of Hind III digested genomic DNA probing with PCR fragments (labeled with the Radi-primed ^{32}PdCTP-labelling kit, GIBCO-BR). Hybridization, washing and detection were carried using standard procedure. Total RNA from tobacco and cotton leaves was extracted using Trizol reagent (GIBCO BRL) and $LiCl_2$/ guanidine isothoicyanate procedure of cotton RNA. Northern blotting was performed using 20µg of total RNA in formaldehyde agarose gel. Multiplex PCR was used to amplify CLCuV DNA A variants Pak 1 and Pak2 using the primers:

I) 5'-ATGTCGAAGCGACTCCGATATCGTCATTTCTACG-3',
II) 5'-TGATGAGTTCCCCTGTGCGTGAATCCATGGTTGT-3',
III) 5'-GAATAAATACCGTTCCGCTTTGGAGGCATGTTG-3'

Southern blotting, as above, of 5µg of total undigested plant DNA was used to detect CLCuD DNA β, using a full length DNA β as a probe.

The rate of transcription or expression of two tobacco resistant sense lines of each construct *in vivo* was analyzed by nuclear run- on assay. For each construct, nuclei from six young leaves (20-30mm) of two resistant sense lines were extracted in the extraction buffer.

Total RNA from two resistant unexposed sense lines for each construct was extracted using TRIzol reagent (GIBCO BRL) for the detection of double-stranded RNA. Northern blot analysis was performed to detect respective genes. About 100 ug of total RNA was treated with RNase one (0.1-0.2 units/ug of RNA) to degrade single-stranded RNA. Small RNAs was extracted from the same lines (described above) in the extraction buffer. *In vitro* transcripts of internally labeled of respective genes were synthesized in both sense and antisense orientations using SP6 or T7 polymerase. The triple antibody sandwich ELISA (TAS) and western blotting were performed on the extracted protein of the same plants described above.

We constructed *Agrobacterium* binary vectors with expression cassettes driving, in the sense and antisense orientations, three different regions of the AC1-4 genes [i] D1/4, the first half of AC1 and all of the AC4 ORF; [ii] D1, the last two thirds of AC1 excluding overlap with AC2 and AC4; and [iii] D1/d2/d3, the last 97 bases of the AC1, all except the initiation codon of AC2 and the 3' 23 bases of AC3. About 20-25 independent transgenic plant lines were produced in three independent experiments for each of the six constructs in tobacco [cv. Samsun cotton Coker-312 and variable number of lines (about 55) of S-12 and NIAB-78. Twenty selected T0 lines for each construct were selfed to produce T1 plants. A subset of these T0 lines, 12 lines/construct, all phenotypic ally normal, were verified by PCR to contain both transgene and the nptII gene encoding kanamycin resistance [the selectable marker] and similarly selfed to make T2 and T3 lines. Southern blots of selected transgenic lines show transgenes present in one to three copies per genome (Fig: 1) Because of the greater difficulty in obtaining cotton transformants, with less available embryogenic transgenic callus material, some T0 plants were derived from a common original transgenic culture and thus perhaps from the same original transformed cell.

Transgenic tobacco and cotton lines expressing viral RNA are resistant to CLCuD and inhibit viral DNA A and DNA β. replication: We performed four independent inoculation experiments with viruliferous whiteflies carrying CLCuD components with a mixture of two different but closely related DNA A components Pak 1and Pak 2 prevailing at NIBGE, Faisalabad, Pakistan.

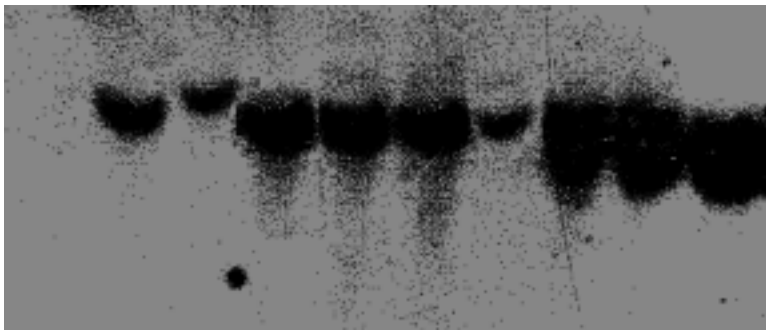


Figure 3.1. Genomic southern of trangenic cotton Coker-312 showed the presence of 1-2copies of transgene/genome comparable to control (Left to right).

.

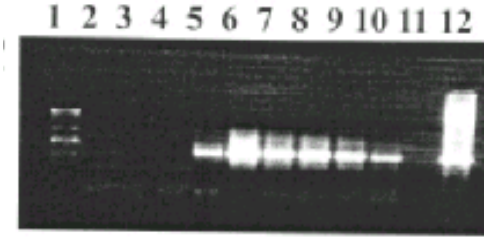


Figure 3.2. Multiplex PCR showing the absence /presence of replication of cotton leaf curl virus DNA A of Pak1 and Pak2 strains in the /resistant/ susceptible plants.

We tested for virus resistance in untransformed controls and kanamycin resistant T1 plants carrying the transgenes, the lines were scored as resistant if greater than 70% of plants showed no symptoms; accumulation of viral DNAs were analyzed by multiplex PCR to discriminate between Pak1 and Pak2 and by Southern blot analysis to detect the DNA β component (Figure 3.2 and 3.3).

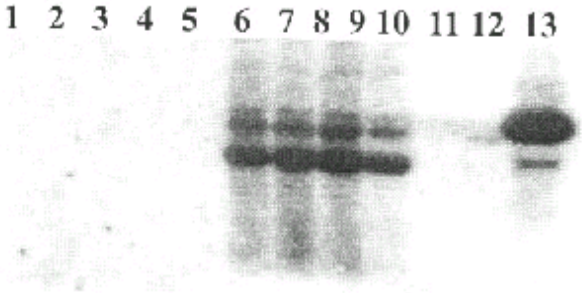


Figure 3.3 Genomic Southern of total genomic DNA showing the absence/presence of CLCu DNA β in the resistant/ susceptible cotton coker-312 plants.

A total of 72 lines of tobacco tested, many were still fully susceptible to infection and some showed mild and delayed symptoms No recovery phenomenon was observed in any of these susceptible transgenic lines. But six T1 antisense transgenic tobacco lines (1of D1/4; 3 of D1; 2 of D1/d2/d3), 12 tobacco sense lines(four lines of each construct D1/4; D1; D1/d2/d3), and two cotton Coker D1 antisense lines out of 12 fertile lines of antisense D1, 15 lines of S-12 sense/antisense D1/4, D1, D1/d2/d3 neither developed symptoms nor contained detectable amounts of DNAs of either CLCuD components, so indicating inhibition of viral DNA replication. In contrast, control plants, like susceptible transgenic plants, all showed the typical symptoms of CLCuD –vein thickening, enation and leaf curling; multiplex PCR revealed the presence of both Pak 1and Pak 2 in amplifying 1.2 and 1.5 kb diagnostic fragments respectively (Figure 3.2, 3.3, 3.4 and 3.5).

Figure 3.4. Transgenic tobacco expressing antisense D1 is resistant to CLCuD. The T2 and T3 lines derived from all resistant lines of tobacco and cotton Coker-312 were challenged to viruliferous whiteflies under containment and found 90-100% resistance in all the lines (Figure 3.5).

Figure 3.5. Transgenic cotton Coker-312 expressing/not expressing antisense D1 gene are susceptible/resistant to CLCuD.

Cotton lines derived from bombarded embryos with antisense D1 did not showed resistance to CLCuD. Cotton S-12 T2 lines sense/antisense of D1/4, D1 and D1/d2/d3 showed 80-100% resistance under containment. The limited seed number of these lines is being planted in the NIBGE field for checking their stability. While one line of each of S-12 and NIAB-78 of D1 antisense construct showed 60-77% resistance in the NIBGE /Rahim Yar khan fields which indicated the stable Mendelian inheritance in these now homozygous lines.
Northern blot analysis, performed using tissue from the same plants chosen for the Southern analysis, yielded transcripts of the following sizes in lines transformed with the constructs: D1/4, 466nts; D1, 543nts; D1/d2/d3 526nts. No transcripts were detected by northern analysis of selected resistant lines after infection in contrast to susceptible lines and controls (Figure 3.6).

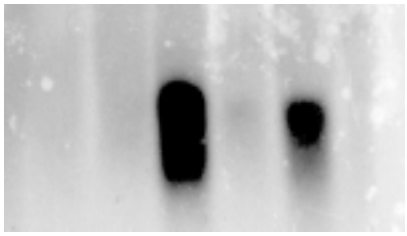


Figure 3.6. Resistant transgenic cotton Coker-312 lines expressing the D1 antisense (543bp).

Analysis of resistance mechanism in tobacco sense lines: The sense transgenes had originally been designed as negative controls to the antisense lines but, as they also showed resistance to CLCuD, we investigated whether resistance involved homology sensing mechanisms leading to post-transcriptional gene silencing, PTGS. Here mRNA is degraded by 21-23 nt long so-called guide RNAs or small interfering RNAs, siRNAs which prime synthesis of dsRNA from an mRNA template by an RNA-dependent RNA polymerase; the dsRNA is then processed by a dsRNA-specific RNase, Dicer, to more siRNAs to repeat the cycle of dsRNA synthesis and degradation.

The total RNA from two resistant, but uninfected sense lines for each construct was treated with RNase I to remove ssRNA and saw discrete size classes of dsRNA at 400nt, 280nt, 155nt and smaller. (Fig 3.5). We also were able to detect faintly, in low molecular weight RNA preparations, approx. 22nt siRNAs hybridizing with strand-specific RNA probes.

For the analysis of resistance mechanism in tobacco sense lines, two possible mechanisms were studied: (i) Resistance involved homology sensing mechanism leading to post-transcriptional gene silencing (PTGS) where mRNA is degraded in a PCR cycle by 21-23 nt long so called guide RNAs or small interfering RNAs, siRNAs which prime synthesis of dsRNA from an mRNA template by an RNA- dependent RNA polymerase; the dsRNA is then processed by a dsRNA-specific RNase, Dicer, to more siRNAs to repeat the cycle of dsRNA synthesis and degradation. (ii) Resistance is via production of trans-dominant defective geminivirus proteins encoded by the transgenes.

Results of low MW RNA preparations showed, approx, 22nt siRNAs very faint bands hybridizing with strand-specific RNA probes. We concluded as geminivirus and transgene RNA is not detected in resistant lines following infection, and resistant tobacco sense lines revealed dsRNA and small interfering RNAs, the most likely mechanism is via post –transcriptional gene silencing. siRNAs can also promote transcriptional gene silencing, TGS, by promoter methylation. As uninfected resistant lines express transgene RNA, TGS cannot be operating prior to infection; but we cannot rule out any such epigenetic gene silencing by methylation of geminivirus and transgene promoters following infection.

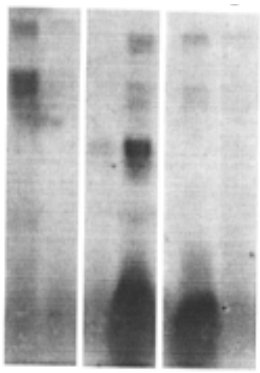


Figure 3.7. Northern blots of two uninfected sense tobacco lines of each construct indicating the presence of double-stranded RNA of discrete sizes.

The antibodies provided by ADGEN Company and Dr Marie (USA) did not cross-react with the transgenes of cotton leaf curl virus (DNA A AC1-4). However, we cannot rule out this possibility as proper antibodies were not available and an additional resistance mechanism via production of trans dominant defective geminivirus proteins encoded by transgenes may act prior to PTGS that removes all geminivirus RNA in infected plants.

3.2 Development of virus resistant plants through conventional breeding

Efforts to identify sources of natural genetic the disease and their utilization for the development of new commercial cultivars have been extensive (Ali, 1997). Upland cotton which is an allotetraploid (AADD), is cultivated in more than 95 % of the total cotton-areas in Pakistan. Sources of genetic resistance identified so far falls into two main categories. In the first are two diploid cotton species of Asiatic origin, namely *G. arboreum L.* (AA) and *G. herbaceum L.* (AA) but so far these have not been utilized in the development of new commercial varieties because of their diploid nature. The second category includes two genotypes belonging to *G. hirsutum*, namely LRA5166 and CP15/2. These sources of resistance have been used for the development of new virus-resistant genotypes such as CIM 1100, CIM 448, CIM 443, and other advanced breeding lines. In virus-susceptible cotton genotypes there is considerable variation in disease incidence and symptom severity. A disease scale

based on the symptom severity of CLCuV in infected plants was used. The resistance of plant genotypes might or might not correlate with the level of virus in the plant. Such information has important implications concerning the mechanism and type of resistance. For example, it has been proposed that inhibition of virus accumulation and short/long distance movement corresponds to constitutively/ expressed dominant alleles (Atabekov Y Dorokhov, 1984).

Development of virus resistant varieties through conventional breeding was another important objective of the project. This component was exclusively taken up by Cotton Research Institute (CRI), Faisalabad which is the largest monocrop institute of the Punjab Government with five Cotton Research Stations (CRS) at Multan, Sahiwal, Vehari, Bahawalpur, and Rahim Yar Khan. Prior to the epidemic of cotton leaf curl virus, (1992-3) the varieties of CRI covered nearly 60% area of cotton. However, CLCuD drastically reduced the share of commercial varieties on the farmer field. Breeders extensively searched for some source of CLCV resistance in the local exotic germplasm. The funding by CFC/ICAC accelerated the pace of conventional breeding programs of CRI, Faisalabad.

3.2.1 Evaluation of natural sources of resistance against CLCuD

Cotton genotypes representing highly susceptible, tolerant and resistant germplasm were collected from primary and secondary gene pools. Plants were grown in an experimental plot at NIBGE for the evaluation of natural resistance. Natural disease incidence and symptom severity were recorded for each genotype. A visual system of scoring for disease intensity was devised by the CLCuV project co-ordination unit (Karim, 1996), based on the following scale: 0=No symptoms, 1= Thickening of small veins, scattered, 2= Thickening of all veins but no curling, 3= Thickening of veins and curling of leaves at the top (light effect), 5= Thickening of veins and curling of leaves on entire plant and dwarfing of plant (severe effect). Pooled representative samples of young leaves from ten individual plants of each cultivar were collected. Total DNA was isolated from 1-2 g of pooled sample by a CTAB method (chapter2). Total DNA isolated from cotton plants (var. S-12) grown in sterile culture was used as the control. The quality and quantity of DNA was measured spectrophotometrically.

Total DNA isolated from plants was denatured using NaOH/EDTA (0.4M NaOH, 0.2 MEDTA) for 10 minutes followed by heat denaturation in a boiling water bath for 5 minutes. Samples were immediately chilled on ice and approximately 5 μg DNA (10 μ1) of each genotype was potted onto a nylon membrane using a vacuum manifold. The membrane was placed on filter paper soaked in neutralization solution (1M Tris HC1pH 7.5, 1.5 M NaC1) and was then briefly dipped in 2xSSC and air dried. DNA was fixed to the membrane using an automatic UV cross linker (Stratagene, USA). A full-length clone of DNA A (CLCuV-Pak) was used as a probe. The DNA was labeled with [^{32}P]dCTP by the random-primer labeling system from Amersham (Ready Prime, Amersham, UK). The probe was hybridized with the membrane-bound DNA samples overnight in a hybridization oven at 65 ^{0}C. Excess probe was washed from the membrane at 65 ^{0}C for 30 minutes using 2X SSC and 1 % SDS. This was followed by another wash of 1XSSC and 0.1% SDS for 30 minutes. After washing, the membrane was exposed to an X-ray film and developed after overnight exposure at – 70 ^{0}C.

Two-fold serial dilutions of 1:1, 1:2, 1:4, 1:8, 1:16, 1:32 and 1:64 were made starting with 5 µg DNA. In the last row of dot-blots, 5 µg of total DNA isolated from the cotton plant (var.S-12) grown in sterile culture was used as a control. The final volume of serial dilutions was 10 µ1 and were denatured before spotting on a membrane as described above. The dilutions of some highly susceptible, less susceptible and resistant lines were made for comparison.

A series of two-fold dilution of CLCuV infected DNA was used to determine relative level of viral DNA (Figure 3.8). Serial dilution suggest that CIM 70, Nazelli and *G. barbadense* appear to be the most susceptible genotypes as they accumulated higher levels of viral DNA. No signal was found in LRA 5166, CP15/2, CIM 448 and CIM 435, confirming that these are resistant genotypes. NIAB Karishma developed a signal at the first two dilutions.

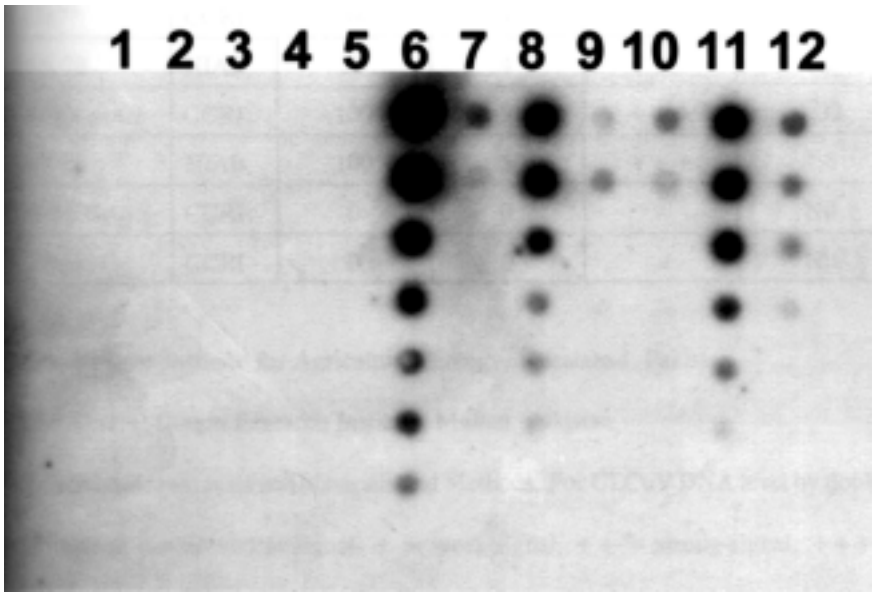


Figure 3.8. The level of begomovirus DNA in cotton genotypes. The level of viral DNA as judged by signal strength on slot-blot (top panel). The signal of a highly susceptible cotton (CIM 70) has been shown by arrow (top panel). In the bottom panel, the relative level of begomovirus DNA has been estimated by serial-dilutions. The serial dilutions starts from 1:1 (first row) to 1:64 (seventh row in lane 6) Lane 1, LRA 5166; lane 2, CP15/2; lane 3, CIM 448; lane 4 CIM 435; lane 5 CIM 443; lane 6 CIM 70; lane 7, NIAB Karishma; lane 8, Nazelli; lane 9, H 1787; lane 10, NIAB078; lane 11, *G. barbadense*; lane 12, NIAB 86.

3.2.2 Collection of germplasm resistant to leaf curl virus

At Cotton Research Institute (CRI), Faisalabad, 414 accessions including 292 exotic and 122 local accessions of *Gossypium hirsutum*. L are available. The available cotton germplasm has been evaluated for leaf curl virus resistance/ susceptibility. At present various sources of resistance to leaf curl virus are available at CRI, Faisalabad. (Table 3.1).

Table 3.1 Sources of resistance against cotton leaf curl disease

Primary Sources of Resistance	**First generation approved resistant varieties**	**Second generation approved resistant varieties**
1. CP-15/2	1. FH-634	1. FH-900
2. LRA-5166	2. CIM-1100	2. FH-901
3. KIVI-1021	3. CIM-448	3. CIM-482
	4. CIM-443	4. BH-118
	5. FVH-53	5. MNH-554
		6. MNH-552

3.2.3 Development of virus resistant varieties

Extensive crossing with exotic source of resistance, evaluation of segregated generations and multi location trials resulted in development of four commercial varieties. In Pakistan, varietal approval system is executed by an independent body, Punjab Seed Council (PSC) in the province while Federal Seed Certification & Registration Department (FSC&RD) is the supreme body for approval and maintenance of seed quality in the country (Figs 3.9 & 3.10).
The four varieties FH-900, FH 901, MNH 552, MNH 554 were released for general cultivation in year 2002. Some salient characteristics of these varieties are given in table: 3b.2

Table 3.2 List of CLCV resistance varieties registered/ released by CRI, Faisalabad

SrNo	***Variety***	***Year***	***Yield Potential) Kg/ha***	***Ginning Outturn (%)***	***Staple Length (mm)***	***Salient Features***
1	FH-900	2000	38.00	37.5	29.8	Virus resistant/heat tolerant
2	FH-901	2000	37.00	38.8	26.7	Virus resistant
3	MNH-554	2000	37.00	42.5	29.0	Virus and drought resistant
4	MNH-552	2000	38.00	40.0	27.1	Virus and heat resistant

Figure 3.9. Characteristics of virus resistant variety FH-900 developed by the project. The figure show plant shape and shape of flower, leaf and bolls of this variety.

These varieties now covered nearly 15% of the total area (2.1 million ha) under cotton in Pakistan. This achievement was made possible due to focused research program

and strict evaluation at each stage for CLCuV resistance. During the report period some other virus resistant potential candidate varieties have also been developed. These have been registered with national seed registration authorities (FSC&RD) by CRI, Faisalabad (Table 3.3).

Table 3.3 Candidate varieties released by CRI, Faisalabad, Pakistan

SrNo	***Name of Strain***	***Yield Potential (Kg/ha)***	***Ginning Outturn (%)***	***Staple Length (mm)***	***Fineness (Mic.) Ug/sq.inch***	***Remarks***
1	FH-925	3800	38.5	28.5	4.6	Short stature, big bolls
2	FH-930	3700	38.0	28.5	4.6	---
3	FH-945	3700	40.0	29.0	4.5	Heat tolerant (HT)
4	FH-1000	3900	39.0	28.5	4.5	Heat tolerant
5	MNH-633	4700	43.0	27.0	4.7	HT, high GOT
6	MNH-642	4500	47.5	27.5	4.3	V. high GOT,HT,EM

Thus the varieties developed by the CRI, Faisalabad are commercially utilized by the farmers and related industries. The project made vital contributions in strengthening the breeding program for cotton leaf curl virus resistance. The research expertise and facilities will also useful for breeding program in cotton.

Cotton Research Institute, Faisalabad, carried out extensive breeding work for evolution of new virus resistant high yielding cotton varieties. After multilocation trials three new varieties namely FH 634, FH 900 and FH 9001 were approved for cotton cultivation in Pakistan. Another virus resistant variety with the commercial name NIBGE1 has been tested extensively at NIBGE and now is undergoing varietal trials at multiple locations.

There is a potential that virus-resistant cotton genotypes developed in the project may confer broad-spectrum resistance to cotton begomoviruses in other parts of the world. Tests carried out so far at University of Arizona and the John Innes Centre is expected to provide interesting results. Genotypes susceptible to cotton leaf curl disease were found to be susceptible to CLCV from Arizona, CLCuV from Sudan and SiGMV from South America. Virus-resistant cotton genotypes are being evaluated to confer resistance to CLCV in Arizona, and to SiGMV at John Innes Centre.

Figure 3.10. Characteristics of virus resistant variety FH-900 developed by the project. The figure show plant shape and shape of flower, leaf and bolls of this variety.

References cited

Argüello-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.

Azzam, O., Frazer, J., de la Rosa, D., Beaver, J. S., Ahlquist, P., and Maxwell, D. P. (1994). Whitefly transmission and efficient ssDNA accumulation of bean golden mosaic geminivirus requires functional coat protein. *Virology* **204,** 289-296.

Briddon, R. W., Bedford, I. D., Tsai, J. H., and Markham, P. G. (1996). Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. *Virology* **219,** 387-394.

Briddon, R. W., 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.

Briddon, 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.

Briddon, R. W., and Markham, P. G. (1995). Geminiviridae. *In* Virus Taxonomy; Sixth Report of the International Committee on Taxonomy of Viruses@ (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 158-165. Springer, Wien.

Briddon, R. W., and Markham, P. G. (2001). Cotton leaf curl disease. *Virus Res*. **71**, 151-159.

Briddon, R. W., Watts, J., Markham, P. G., and Stanley, J. (1989). The coat protein of beet curly top virus is essential for infectivity. *Virology* **172,** 628-633.

Czosnek, H., and Laterrot, H. (1997). A worldwide survey of tomato yellow leaf curl viruses. *Arch. Virol*. **142**, 1391-1406.

Devereaux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387-395.

Fontes, E. P. B., Gladfelter, 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.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., and Robertson, D. (1999). Geminiruses: models for plant DNA replication, transcription, and cell cycle regulation. *Crit. Rev. Plant Sci.* **18,** 71-106.

Hussain, T., and Mahmood, T. (1988). A note on leaf curl disease of cotton. *The Pakistan Cottons* **32,** 248-251.

Hussain, T., Tahir, M., and Mahmood, T. (1991). Cotton leaf curl virus. *Pakistan J. Phytopath.* **3,** 57-61.

Idris, A. M., and Brown, J. K. (2000). Identification of a new, monopartite begomovirus associated with leaf curl disease of cotton in Gezira, Sudan. *Plant Dis.* **84,** 809.

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S. G., Schell, J., and Gronenborn, B. (1995). *In vitro* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Natl. Acad. Sci., USA* **92,** 3879-3883.

Liu, Y., Robinson, D. J., and Harrison, B. D. (1998a). 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.

Liu, S., Briddon, R. W., Bedford, I. D., Pinner, M. S., and Markham, P. G. (1998b). Identification of genes directly and indirectly involved in the insect transmission of African cassava mosaic geminivirus by *Bemisia tabaci*. *Virus Genes* **18,** 5-11.

Mansoor, S., Bedford, I., Pinner, M. S., Stanley, J., and Markham, P. G. (1993). A whitefly-transmitted geminivirus associated with cotton leaf curl disease in Pakistan. *Pak. J. Bot.* **25,** 105-107.

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R. W., Stanley, J., and Markham, P. G. (1999). Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology* **259,** 190-199.

Mansoor, S., Khan, S. H., Hussain, M., Zafar, Y., Pinner, M. S., Briddon, R. W., Stanley, J., and Markham, P. G. (2000). Association of a begomovirus and nanovirus-like molecule with *Ageratum* yellow vein disease in Pakistan. *Plant Dis.* **84,** 101.

McGlashan, D., Polston, J. E., and Bois, D. (1994). Tomato yellow leaf curl virus in Jamaica. *Plant Dis.* **78,** 1219.

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

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

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

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

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

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

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* **97,** 6890-6895.

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

Schwartz, R. M. and Dayhoff, M. O. (1978). *In* Atlas of protein sequence and structure (M. O. Dayhoff, Ed.), volume 5, pp. 353-362 National Biomedical Research Foundation, Washington DC.

Papers published (NIBGE and JIC components)

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

Briddon, 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.

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

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

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

Briddon, RW, Bull SE, Amin I, Idris, AM, Mansoor S, Bedford ID, Dhawan P, Rishi N, Siwatch SS, Abdel-Salam AM, Brown, JK, Zafar, Y and Markham PG 2003. Diversity of DNAβ; a satellite molecule associated with some monopartite begomoviruses. *Virology* (accepted).

Khan, S.A., Hussain, D., Askari,E., Stewart, J.M.D., Malik, K.A. and Zafar, Y. (2000). Molecular phylogeny of Gossypium species by DNA fingerprinting. *Theoretical and Applied Genetics,* 101,931-938.

Mansoor, S., Khan, S. H., Saeed, M., Bashir, A., Zafar. Y. and. Malik. K. A 1997. Evidence for the association of a bipartite geminivirus with tomato leaf curl disease in Pakistan. *Plant Disease* 81, 958.

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R.W., Stanley, J., and Markham, P. G. (1999). Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology* 259, 190-199.

Mansoor, S., Bashir, A, Saeed, M., Khan, S. H., Hussain, M., Zafar Y., and Malik, K. A., (1999). Rapid multiplex PCR for the specific detection of two whitefly-transmitted geminivirus species associated with cotton leaf curl disease in Pakistan. *Pakistan Journal of Botany,* 31, 115-123.

Mansoor, S., Khan, S. H., Hussain, M., Zafar, Y., Pinner, M. S., Briddon, R. W., Stanley, J., and Markham, P. G. (2000). Association of a begomovirus and nanovirus-like molecule with Ageratum yellow vein disease in Pakistan. *Plant Disease* 84, 101.

Mansoor, S., Mukhtar, S., Hussain, M., Amin, I., Bashir, A., Zafar Y., and Malik K. A., (2000). Wide-spread occurrence of cotton leaf curl virus on radish in Pakistan. *Plant disease* 84, 101.

Mansoor, S., S. H. Khan, M. Hussain, N. Mushtaq, Y. Zafar, and K. A. Malik (2000). Evidence that watermelon leaf curl disease in Pakistan is associated with tomato leaf curl virus-India. *Plant Disease* 84, 102.

Mansoor, S., Amin, I., Hussain, M., Zafar, Y., Bull, S. E., Briddon, R. W., and Markham, P. G. (2001). Association of a disease complex involving a begomovirus, DNA 1 and a distinct DNA ß with leaf curl disease of okra in Pakistan. *Plant Disease* 85, 922.

Mansoor, S., Saqib, M., Hussain, M., Briddon R W., Malik K. A., and Zafar, Y. Development of gene expression and gene silencing vectors based on cotton leaf curl virus for functional genomics in cotton. The Third ICGI Workshop. *Cotton Science* 14, 45.

Mansoor, S., Briddon, R.W., Zafar, Y and Stanley, J. (2003). Geminivirus disease complexes: an emerging threat. *Trends in Plant Sciences* (in press).

Rahman M, Hussain, D., and Zafar, Y. Estimation of genetic divergence among elite cotton cultivars-genotypes by DNA fingerprinting technology. *Crop Science*, 42, 2137-2144.

Papers submitted:

Asad, S., Haris, W.A.A., Bashir, A., Zafar, Y., Malik, K. A., Malik, N.N., and Lichtenstein, C.P. (2003). Transgenic tobacco expressing geminiviral RNAs are resistant to the serious viral pathogen causing cotton leaf curl disease. *Archives of Virology* (Submitted).

Mansoor, S., Briddon, R.W., Bull, S. E., Bedford I. D., Bashir, A., Hussain, M., Saeed M., Zafar, Y., Malik K. A., Fauquet, C., and Markham, P. G. (2003). Relaxed specificity for interaction of cotton begomoviruses with DNA beta. *Archives of Virolgy* (submitted).

Rahman M, Hussain, D., Malik, T.A., and Zafar, Y (2003). Inheritance of cotton leaf curl virus in *Gossypium hirsutum* L. (Submitted to *Phytopathology*).

Saeed, M., Bashir, A., Faisal, A., Ahmad, N., Qamar, Z., Malik, K. A. and Zafar, Y. Development of antibodies against in-vitro expressed coat Protein and replication associated protein of cotton leaf curl virus. (Submitted to Microbiology & Immunology, Japan)

Manuscripts in preparation

Asad, S., Haris, W.A.A., Zafar, Y., Lichtenstein,C.P., and Malik, K. A. (2003). Expression of viral RNAs in cotton develops different levels of resistance against cotton leaf curl virus.(In preparation for Molecular Breeding).

Asad, S., Haris, W.A.A., Zafar, Y., Lichtenstein, C.P., and Malik, K. A. (2002). Cosuppresssion of cotton leaf curl virus (CLCuV) in *Nicotiana benthamiana*. (In preparation for Plant Microbe Interaction)

Briddon, R.W., Bull, S.E., Bedford, I.D., Rishi, N., Siwatch, S.S., Saunders, K., Stanley, J. and Markham, P.G. (2003). Pathogenic recombinant between cotton leaf curl virus and cotton leaf curl disease DNA: implications for the future evolution of the virus complex.

Briddon, R.W., Bull, S.E., Bedford, I.D., Saunders, K., Stanley, J. and Markham, P.G. (2003). Maintenance of the satellite-like molecule, DNA 1, by geminiviruses requires the coat protein.

Briddon, R.W., Bull, S.E., Mansoor, S., Bedford, I.D., Rishi, N., Siwatch, S.S., Abdel-Salam, A.M. and Markham, P.G. (2003). Diversity of DNA 1; a satellite-like molecule associated with monopartite begomoviruses.

Briddon, R. W., and Markham, P. G. (2003). Episomal replication of *Nicotiana tabacum* integrated geminivirus sequences following infection by the cotton leaf curl disease complex.

Mansoor, S., Briddon, R.W., Bull, S.E., Bedford, I.D., Bashir, A., Hussain, M., Saeed, M., Zafar, Y., Malik, K. A., and Markham, P.G. (2003). Characterization of begomoviruses associated with cotton leaf curl disease.

Papers presented at conferences/symposia:

Asad, S., Bashir, A., Haris, W.A.A., Zafar, Y., Lichtenstein,C.P., and Malik, K.A. (2001). Expression of antisense RNA of viral genes confers resistance against cotton leaf curl disease in tobacco. Paper abstract in Third international geminivirus symposium held at John Innes Centre, UK on 24-28 July 2001.

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. Paper presented in third International Geminivirus Symposium, 24-28 July 2001, John Innes Centre, Norwich, UK.

Zafar, Y. (2000). Plant Biotechnology-Providing Biotech. Solutions to Agricultural Problems. Paper presented in OIC-COMSTECH meeting on Biotechnology for Development, Islamabad, Pakistan, 24-26 January, 2000.

Asad, S., Bashir, A., Haris, W.A.A., Zafar, Y., Lichtenstein, C.P. and Malik., K.A. (1999). Genetic engineering for leaf curl virus resistance in cotton. Paper abstract in 2nd National Symposium on Plant Tissue Culture and Genetic Engineering. June 1-3, Islamabad, Pakistan.

Asad, S., Bashir, A., Haris, W.A.A., Zafar, Y., Lichtenstein, C.P., and Malik, K.A. (1999). Genetic engineering for leaf curl virus resistance in cotton. Paper abstract in

2nd National Symposium on Plant tissue culture and genetic engineering. June 1-3, Islamabad, Pakistan.

Hussain, M., Mansoor, S., Khan, S.H., Bashir, A., Zafar, Y., and Malik, K.A. (1999). Broad-spectrum resistance in cotton against geminiviruses. Paper abstract in National Conference on Tissue Culture and Genetic Engineering NARC, Islamabad, June 1-3 1999,

Mansoor, S., Hussain, M., Khan, S.H., Bashir, A., Zafar, Y., Malik, K.A., Briddon, R. W., and Markham, P. G. (1999). Analysis of the distribution of cotton leaf curl disease associated begomoviruses in Pakistan. (Abstract) VII International Plant Virus Epidemiological Symposium, Almeria, Spain. April 11-16, 1999.

Mansoor, S., Khan, S. H. Hussain, M., Bashir, A. Saeed, M., Zafar, Y., Stanley, J., Briddon, R. W., Markham, P. G., and Malik, K. A. (1999). DNA variants among Pakistan isolates of cotton leaf curl virus. Proceedings of ICAC-CCRI Regional Consultation on Insecticide Resistance Management in Cotton. June 28 to July 1,1999. Multan, Pakistan. pp 261-276.

Mansoor, S., Nosheen, M., Hussain, M., Khan, S.H., Bashir, A., Zafar, Y., and Malik, K.A. (1999). Evidence for trans-ovarial transmission of cotton leaf curl virus and tomato leaf curl virus by whitefly Bemisia tabaci. Paper abstract in second National Conference of Plant Pathology. September,27-29, Faisalabad, Pakistan.

Mansoor, S., Nosheen, M., Hussain, M., Khan, S.H., Bashir, A., Zafar, Y., and Malik, K. A. (1999). Cotton leaf curl disease in Pakistan: A novel combination of Monopartite Begomoviruses and a novel whitefly-transmitted DNA component showing homology to plant Nanoviruses. Paper abstract in second National conference of Plant Pathology. September 27-29, Faisalabad, Pakistan.

Asad, S., Bashir, A., Haris, W.A.A., Zafar, Y., Lichtenstein,C.P., and Malik, K.A. (1998). Pathogen based resistance in cotton to suppress cotton leaf curl virus (CLCuV) Paper abstract in 6th National Conference of Plant Sciences, October 20-22, Peshawer, Pakistan.

Asad, S., Bashir. A., Harris, W.A.A., Zafar, Y., Lichtenstein, C. P., and Malik, K. A. (1998). Development of resistance in cotton against cotton leaf curl Virus (CLCuV) using viral genes. Paper Abstract in /World Cotton Research Conference. September, 6-12, Athenes, Greece (Abstract Book, pp. 272).

Asad, S., Bashir, A., Haris, W.A.A., Saeed, N.A., Zafar, Y., Lichtenstein, C.P., and Malik, K.A. (1998). Development of resistance in cotton against cotton leaf curl virus (CLCuV) using viral genes. Paper abstract in World Cotton Research Conference, September 6-12, Athens, Greece. (Abstract Book-pp.272)

Asad, S., Bashir, A., Harris, W.A.A., Zafar, Y., Lichtenstein, C.P. and Malik, K. A. (1998) Pathogen based resistance in cotton to suppress cotton leaf curl virus. Paper Abstract in 6th National Conference of Plant Sciences, October 20-22, Peshawar, Pakistan.

Hussain, M., Mansoor, S., Khan, S. H., Bashir, A., Zafar, Y., and Malik, K. A. (1998). New constructs for genetically engineered resistance against cotton leaf curl virus. Paper abstract in 6th National Conference of Plant Sciences, October, 20-22, Peshawar, Pakistan.

Khan, S. H., Mansoor, S., Hussain, M., Bashir, A., Zafar. Y., and Malik, K. A. (1998). Molecular detection of whitefly-transmitted geminiviruses in cotton-growing areas of Pakistan by dot-blot hybridization and polymerase chain reaction. Paper abstract in 6th National Conference of Plant Sciences, October, 20-22, Peshawar, Pakistan.

Mansoor, S., Khan, S.H., Hussain, M., Bashir, A., Zafar, Y., and Malik, K. A. (1998). Coinfection and distribution of variable geminiviruses associated with cotton leaf curl disease in Pakistan. Paper abstract in 6th National Conference of Plant Sciences, October, 20-22, Peshawar, Pakistan.

Tariq, M., Mansoor, S., Khan, S. H., Hussain, M., Asad, S., Mukhtar, Z., Zafar, Y., and Malik, K. A. (1998). Molecular identification of tomato leaf curl virus (ToLCV) and development of genetically engineered resistance in tomato. Paper abstract in 6th National Conference of Plant Scientists, October, 20-22, Peshawar, Pakistan.

Bashir, A., Saeed, M., Mansoor, S., Shabnam, S., Zafar, Y, Malik, K. A., Beachy, R. N., and Fauquet, C. M, 1997. Evidence for the Presence of Two New Geminivirus Species Infecting Cotton in Pakistan. 5th International Congress of Plant Molecular Biology, September 21-27, 1997, Singapore.

Asad, S., Bashir, A., Haris, W.A.A., Saeed, N.A., Zafar, Y., Lichtenstein,C.P., and Malik, K.A. (1997). Use of antisense RNA technology to suppress cotton leaf curl virus (CLCuV) in a model plant system. Poster abstract in International Conference on Plant Moleculr Biology, March 25-28, Singapore.

Asad, S., Bashir, A., Haris, W.A.A., Saeed, N.A., Zafar, Y., Lichtenstein, C.P., and Malik, K.A. (1997). Use of antisense RNA technology to suppress cotton leaf curl virus (CLCuV) in cotton. Poster abstract in Scientific Exhibition by Pakistan Science Foundation, October 28, Islamabad, Pak.

Zafar, Y., Bashir, A., Mansoor, S., Saeed, M., Asad, S., Saeed, N. A., Briddon, R. W., Markham, P. G., Fauquet, C. M., and Malik, K. A. (1997). Cotton leaf curl virus epidemic in Pakistan: virus characterization, diagnosis and development of virus-resistant cotton through genetic engineering. 56th Plenary Meeting of the International Cotton Advisory Committee October 1997 Asuncion, Paraguay.

Bashir, A., Shabnam, S., Aftab, S., Saeed, M., Saeed, N. A., Mansoor, S., Zafar Y., and Malik. K. A., 1996. Isolation, Identification and Molecular Characterization of Cotton Leaf Curl Virus in Pakistan. Proceeding of the Bellagio Conference "Whiteflies and Viruses: Menace to World Agriculture" August 12-16,1996. Bellagio Italy.

Mansoor, S., S. H. Khan, M. Hussain, A. Bashir, Y. Zafar and K. A. Malik (1998). Co-infection and distribution of variable geminiviruses associated with cotton leaf

curl disease in Pakistan. Paper abstract in 6th National Conference of Plant Sciences, October, 20-22, Peshawar, Pakistan.

Sultan. H. Khan, S. Mansoor, M. Hussain, A. Bashir, Y. Zafar and K. A. Malik (1998). Molecular detection of whitefly-transmitted geminiviruses in cotton-growing areas of Pakistan by dot-blot hybridisation and polymerase chain reaction. Paper abstract in 6th National Conference of Plant Sciences, October, 20-22, Peshawar, Pakistan.

Mansoor, S. S. H. Khan, M. Hussain, A. Bashir, M. Saeed, Y. Zafar, J. Stanley, R. Briddon, P. G. Markham and K. A. Malik (1999). Cotton leaf curl disease in Pakistan: A novel combination of monopartite begomoviruses and a novel whitefly-transmitted DNA component showing homology to plant nanoviruses. 2nd National Conference of Plant Pathology Sept 27-29, 1999, Faisalabad, Pakistan.

Mansoor, S., N. Mushtaq, M. Hussain, S. H. Khan, A. Bashir, Y. Zafar, and K. A. Malik (1999). Evidence for transovarial transmission of cotton leaf curl virus and tomato leaf curl virus by whitefly *Bemisia tabaci*. 2nd National Conference of Plant Pathology September, 27-29, 1999, Faisalabad, Pakistan.

Mansoor, S., S. H. Khan, A. Bashir, M. Saeed, Y. Zafar, Rob W. Briddon, J. Stanley, P. G. Markham and K. A. Malik (1999). Cotton leaf curl disease in Pakistan: Molecular characterization, and genetically engineered virus resistance. In 2nd National Symposium on Plant Tissue Culture and Genetic Engineering at ABI, National Agriculture Research Centre. June 1-3, 1999.

Hussain, M, Mansoor, S, Iram, S; Zafar, Y, and Malik, K. A. (2000). Genetically-engineered broad-spectrum resistance against whitefly-transmitted geminiviruses through mutated viral movement protein gene (BC1). 7th National Conference of Plant Scientists. November 14-16, Lahore, Pakistan.

Husnain, G, Mansoor, S., Hussain, M, Iram, S., Haider, S., Mukhtar, Z., Khan, N.A, Swatti, Z.A and Zafar, Y. (2000). Promoter studies of a novel nanovirus-like component associated with cotton leaf curl disease. 7th National Conference of Plant Scientists. November 14-16, Lahore, Pakistan.

Musawar, Y, Mansoor, S, Hussain, M, Amin, I, Haider, S, and Zafar, Y. (2000). Genetically-engineered broad-spectrum resistance through viral coat and pre-coat protein genes. 7th National Conference of Plant Scientists. November 14-16, Lahore, Pakistan.

Khan, M.A, S. Mansoor, Hussain, M; Iram, S and Zafar, Y. (2000). Transgenic suppression of CLCuV DNA-1 to elucidate its role in the disease development. 7th National Conference of Plant Scientists. November 14-16, Lahore, Pakistan.

Bull, S., R.W. Briddon., I. Amin, S. Mansoor, I. D Bedford., A. A. Abdul-Salam, S. S. Siwatch., N. Rishi., P. Dhawan., Y. Zafar., K. A., Malik, and P.G. Markham 2001. Diversity of DNA β; a satellite molecule essential for the symptom induction by a number of monopartite begomoviruses. 3rd International Geminivirus Symposium, John Innes Center, Norwich Norfolk, UK 24- 28 July 2001.

Mansoor, S., R. W. Briddon, M. Saeed, A. Bashir, I. Amin, M. Tariq, M. Hussain, S, Iram, Y, Zafar, K. A. Malik and P. G. Markham 2001. Genomic component of begomoviruses associated with tomato leaf curl disease in Pakistan. 3rd International Geminivirus Symposium, John Innes Center, Norwich Norfolk, UK 24- 28 July 2001.

Briddon, R.W., S. E. Bull, I. D Bedford., S. Mansoor, Y. Zafar., K. A., Malik, and P.G. Markham 2001. Analysis of DNA beta; an essential component of cotton leaf disease. 3rd International Geminivirus Symposium, John Innes Center, Norwich Norfolk, UK 24- 28 July 2001.

Mansoor, S., M. Hussain, Y. Musawar, M. A. Khan, S. Iram, Y. Zafar, R. W. Briddon, P. G. Markham and K. A. Malik. 2001. Strategies for broad-spectrum resistance against cotton leaf curl and tomato leaf curl disease in Pakistan. 3rd International Geminivirus Symposium, John Innes Center, Norwich Norfolk, UK 24-28 July 2001.

Mansoor, S., R. W. Briddon, O. Voinnet, P. G. Markham and D. Baulcombe. 2001. Understanding the role of DNA beta in cotton leaf curl disease in Pakistan. 3rd International Geminivirus Symposium, John Innes Center, Norwich Norfolk, UK 24-28 July 2001.

Thesis completed

Ph.D

Mansoor, S., 1999. Cotton leaf curl disease in Pakistan; Molecular characterisation, diagnostics and genetically-engineered virus-resistance. Ph.D. thesis. University of East Anglia, Norwich, U.K.

M.Phil

Muhammad Afzal Dogar (2002). Engineering resistance against tomato leaf curl virus. Quaid -i-Azam Univ./NIBGE, Faisalabad.

Faiza Noreen (2002). Pathogenecity determinant study of AC2, AC3 and AC4 genes of cotton leaf curl virus. Quaid -i-Azam Univ./NIBGE, Faisalabad.

Haniea Jamal (2001). Cosuppression of cotton leaf curl virus in a model plant. M.Phil Thesis, Department of Botany, University of Agriculture, Faisalabad.

Habib-ur-Rehman (2001). Engineering Resistance against tomato leaf curl geminivirus. Quaid-i-Azam Univ./NIBGE, Faisalabad.

Yasir-ul-Musawar (2001). Broad-spectrum resistance based on coat protein gene. Quaid -i-Azam Univ./NIBGE, Faisalabad.

Amir Khan (2001). Transgenic suppression of nanovirus component associated with cotton leaf curl disease. Quaid -i-Azam Univ./NIBGE, Faisalabad.

Imran Amin (2001). Diversity of DNA β associated with whitefly-transmitted geminiviruses in Pakistan. Department of Biochemistry, University of Agriculture, Faisalabad.

Amir Faisal (2000). Utilization of cotton leaf curl virus coat (AV1) and replication associated (AC1) protein for polyclonal antibody production. Quaid-i-Azam Univ./NIBGE, Faisalabad.

Muhammad Tariq (1999). Molecular identification of tomato leaf curl virus and development of transgenic resistance in a model system. Quaid -i-Azam Univ./NIBGE, Faisalabad.

Sher Afzal Khan (1999). Molecular Phylogeny of *Gossypium* spp. based on DNA finger printing Quaid-i-Azam Univ./NIBGE, Faisalabad.

Shazia Jilani (1998). A comparative study of three variety independent transformation methods for cotton (*G. hirsutum* L.) Quaid-i-Azam Univ./NIBGE, Faisalabad.

Saima Hashmi (1998). Construction of partial dimmers for cotton leaf curl disease related DNA components and evaluation for their infectivity to cotton and tobacco through particle bombardment. Quaid -i-Azam Univ./NIBGE, Faisalabad.

Nazish Ahmed (1998). Development of polyclonal antibodies against pre-coat/AC1 protein of cotton leaf curl disease. Quaid-i-Azam Univ./NIBGE, Faisalabad.

Mohsin A. Khan (1996). Studies on genetic transformation of cotton by using Agrobacterium and biolistic methods. Quaid-i-Azam Univ./NIBGE, Faisalabad.

Naveed Aziz (1996). Evaluation of genetic variability among cotton varieties by random amplified DNA (RAPD) markers. Quaid -i-Azam Univ./NIBGE, Faisalabad.

M.Sc.

Muhammad Mubeen Raja (2000). Introduction of a foreign gene in local cultivars of cotton by *Agrobacterium tumefaciens*. Department of Biochemistry, University of Agriculture, Faisalabad.

Mamoona Naz Molecular Analyses of transgenic tobacco plants. M.Sc Thesis, Department of Chemistry, University of Agriculture, Faisalabad.

Ghulam Husnain (2000). Promoter studies of cotton leaf curl virus DNA 1 component. Department of Plant Breeding and Genetics. NWFP Agriculture University, Peshawar.

Nosheen Mushtaq (1999). Molecular studies on virus-vector relationship in cotton leaf curl disease. Department of Chemistry, University of Agriculture, Faisalabad.

Thesis (in progress)

Ph. D

Sajjad Haidar. Disease resistance gene cloning and characterization from *Gossypium arboreum.* Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad.

Mujaddad-ur-Rehman. Virus resistant cotton plants through genetic engineering. Department of Biological Sciences, QAU/NIBGE, Faisalabad.

Ghulam Sarwar. Mechanism of resistance and diversity in source of natural resistance against cotton leaf curl disease. Department of PBG, University of Agriculture, Faisalabad.

Shaheen Asad. Use of antisense RNA technology to suppress CLCuV infection in a model plant.

Muhammad Siddique. Expression of truncated replicase genes for developing CLCuV resistance. Department of Botany University of Punjab, Lahore.

Chapters in Books/Proceedings

Malik, T.A. and M. Rahman. (1998). A chapter "Genetically Engineering Plants for deserts". (Book in press).

Rahman M, D Hussain and Y Zafar. 2001. "DNA Markers: A review". Proceedings, First SESCOB Biotechnology Conference, January 11, 1996, Univ. of Agric. Faisalabad.

Zafar,Y .2001 Pakistan in;Agricultural Biotechnology: Country case studies-A Decade of Development,Edited by GabrielleJ,Persley and L.Reginald Maclntyre.CABI International Publisher,UK

Zafar Y, S Mansoor, M. Rahman, S Asad, M. Saeed, M Hussain, W. A. Haris and K A Malik. 2001. Cotton Biotechnology: Contribution of NIBGE. In: New Horizons in DNA Technology, Centre for Agricultural Biochemistry & Biotechnonlogy (CABB), Univ. of Agric. Faisalabad. (in press).

Publications (University of Arizona component)

Brown, J.K. 2000. Molecular markers for the identification and global tracking of whitefly vector-begomovirus complexes. *Virus Research* 71, 233-260.

Costa, H.S., D.M. Wescot, D.E. Ullman, R.C. Rosell, J.K Brown, and M.W. Johnson. 1996. Virus-like particles in the mycetocytes of the sweetpotato whitefly, *Bemisia tabaci* (Homoptera: Aleyrodidae). *Journal of Invertibrate Pathology* 67,183-186.

Torres-Pacheco, I., J. A. Garzon-Tiznado, J. K. Brown, A. Becerra-Flora, and R. Rivera-Bustamante. 1996. Detection and distribution of geminiviruses in Mexico and the southern United States. *Phytopathology* 86, 1186-1192.

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

Idris, A. M. and J.K. Brown. 1998. Sinaloa tomato leaf curl geminivirus: biological and molecular evidence for a new subgroup III virus. *Phytopathology* 88, 648-657.

Rosell, R.C., I. Torres-Jerez, and J.K. Brown. 1999. Temporal pathway of geminivirus in whitefly extracts, saliva, hemolymph and honeydew. *Phytopathology* 89, 239-246.

Brown, J. K., A. M. Idris, I. Torres-Jerez, G. K. Banks, and S. D. Wyatt. 2001. The core region of the coat protein gene is highly useful for establishing the provisional identification and classification of begomoviruses. *Archives of Virology*, 146, 1-18.

Short Communications

Idris, A. M., J. Bird, and J.K. Brown. 1999. First report of a bean-infecting begomovirus from *Macroptilium lathyroides* in Puerto Rico that is distinct from bean golden mosaic virus. *Plant Disease* 83, 1071.

Idris, A.M., and J. K. Brown. 2000. Identification of new, monopartite begomovirus associated with leaf curl disease of cotton in Gezira, Sudan. *Plant Disease* 84, 809.

Published Proceedings

Brown, J.K. 1996. Distribution and genetic variability of whitefly-transmitted geminiviruses of cotton. Pages 275-279 in: Proceeding of 1996. Beltwide Cotton Conference, Nashville TN, January, 1996.

Brown, J.K., J. Bird, D.R. Frohlich, R.C. Rosell, I.D. Bedford, and P.G. Markham. 1995. The relevance of variability within the *Bemisia tabaci* species complex to epidemics caused by Subgroup III geminiviruses. Pages 77-92 in:*Bemisia '95: Taxonomy, Biology, Damage, Control and Management* (D. Gerling and R. T. Mayer, eds). Intercept Publications, Wimborne, UK.

Frohlich, D.R., J.K. Brown, I.D. Bedford, and P.G. Markham. 1996. Mitochondrial 16S ribosomal subunit as a molecular marker in Bemisia, and implications for population variability. Pages 143-145 in: *Bemisia '95: Taxonomy, Biology, Damage, Control and Management* (D. Gerling and R.T. Mayer, eds). Intercept Public., Wimborne, UK.

Rosell, R.C., I.D. Bedford, P.G. Markham, D.R. Frohlich, and J.K. Brown. 1996. Morphological variation in Bemisia populations. Pages 147-149 in: *Bemisia '95: Taxonomy, Biology, Damage, Control and Management*. Intercept Publications, Wimborne, UK (D. Gerling and R. T. Mayer, eds).

Brown, J.K. 1997. Global diversity and distribution of cotton-infecting geminiviruses: an esssential requisite to developing sustainable disease resistance. Pages 39-48 in Proceedings of: Technical Seminar at the 56th Plenary Meeting of the International Cotton Advisory Committee. Asuncion, Paraguay, October 27-31, 1997.

Brown, J.K. 1998. Global diversity and distribution of whitefly-transmitted geminiviruses of cotton. Pages 155-161 in: *Proceeding of 1998 Beltwide Cotton Conference*, San Diego, CA, January, 5-9 1998.

Brown, J.K. and S.D. Wyatt. 1997. The core region of the coat protein gene permits virus identification and predicts an evolutionary history of whitefly-transmitted geminiviruses. American Phytopathological Society Proc. Caribbean Division Meeting, San Jose, Costa Rica, Nov 10-12, 1997. Phytopathology S88:120.

Brown, J.K. 2000. Molecular markers for the identification and global tracking of whitefly vector-begomovirus complexes. *Virus Research* 71, 233-260.

Brown, J.K. 2001. The Molecular Epidemiology of Begomoviruses. Pages 279-316 in: *Trends in Plant Virology* (J. A. Khan and J. Dykstra), The Haworth Press, Inc., NY. 537pp.

Brown, J.K. and H. Czosnek. 2001.Whitefly transmitted viruses. Pages 65-100 in: *Advances in Botanical Research*, Academic Press, NY.

Brown, J.K. 2001. Plant Virus and Mycoplasma Diseases of Cotton. Pages **XX** In: *Cotton Compendium*. American Phytopathological Society, St. Paul, MN.

Brown, J.K., A.M. Idris, I. Torres-Jerez, G.K. Banks, and S.D. Wyatt. 2001. The core region of the coat protein gene is highly useful for establishing the provisional identification and classification of begomoviruses. *Archives of Virology*, 146, 1-18.

Idris, A.M., and J.K. Brown. 2002. Molecular analysis *of Cotton leaf curl virus*-Sudan reveals an evolutionary history of recombination. *Virus Genes* (*in press*).

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

Manuscripts submitted

Brown, J.K. and Idris, A.M. 200x. Diversification patterns of New World begomoviruses of malvaceous hosts reveal species and strains that have diverged from common ancestors. *Archives of Virology* .

Idris, A.M. and J.K. Brown. 200X. Molecular analysis of infectious *Cotton leaf crumple virus* reveals a new begomovirus clade, and a species that has undergone recombination with *Sida golden mosaic* and *Bean calico mosaic* viruses. *Phytopathology* (submitted).

Idris, A.M., Briddon, R.W., and Brown J.K. Closely related DNA satellites are essential for cotton leaf curl disease symptom induction in malvaceous species from the Upper Nile Basin. *Virus Research* (submitted).

Idris, A.M., C.M. Fauquet, and J.K. Brown. 200x. Three new monopartite begomovirus genomes associated with Tomato leaf curl disease from Sudan and evidence for interspecies recombination. *Phytopathology* (submitted).

Manuscripts In Preparation

Idris, A.M. and J.K. Brown. 200x. Phylogenetic relationships for cloned *Cotton leaf crumple virus* reveals it is the sole member of a previously unknown New World begomovirus clade and a double recombinant.

Idris, A.M., S. Bull, R. Briddon, and J.K. Brown. 200x. The *Cotton leaf curl virus*-Sudan genome cloned from three malvaceous hosts is identical, while the associated satellite DNAs are phylogenetically diverse and necessary for symptom development.

Abstracts

Brown, J.K. 1999. New and emergent cotton-infecting viruses transmitted by Homopteran vectors: new problems of international importance. Second Congress of Cotton, Ribeirao Prieto, Brazil, September 6-11, 1999.

Brown, J.K., A.M. Idris, I. Torres-Jerez, and S.D. Wyatt. 1999. The core region of the coat protein gene of the geminiviridae is phylogenetically informative: introducing an interactive website for begomovirus identification using the core cp sequence. Page 80 In: Proceedings of VIIth International Plant Virus Epidemiology Symposium, Almeria, Spain, April 11-16, 1999.

Brown, J.K. and A.M. Idris. 2000. Molecular identification and phylogenetic relationships between cotton-infecting begomo-viruses from the Americas, and their relationship to Old and New World begomoviruses of Malvaceous hosts. *Phytopathology* 90: S10.

Idris, A.M., and J.K. Brown. 2000. Molecular characterization of full-length begomovirus DNAs from cotton, hollyhock and okra in Sudan in relation to the epidemiology of cotton leaf curl disease-Sudan. *Phytopathology* 90:S37.

Brown, J.K., A.M. Idris, L.V. Resende, and G.E. de França. 2000. Four new begomovirus species of vegetable and weed hosts in Pernambuco and Sao Paulo, Brazil. Proceedings in The Ninth Latin American and Caribbean Conference on Whiteflies and Geminiviruses. Panama City, Panama, November 21-24, 2000.

Brown, J.K., M. Palmieri, M.M. Doyle, A.M. Idris, and R.D. Martyn. 2000. Diversity and distribution of begomoviruses from vegetable and weed hosts in Guatemala and Honduras by comparative analyses of viral coat protein and common region nucleotide sequences. Proceedings in the Ninth Latin American and Caribbean Conference on Whiteflies and Geminiviruses. Panama City, Panama, November 21-24, 2000.

Brown, J.K., M. Palmieri, M.M.Doyle, A.M. Idris, and R.D.Martyn. 2001. Diversity of begomoviruses from vegetables, cotton, and weeds in Guatemala and Honduras by viral coat protein sequence analysis. Page 72 in: Fourth International Seminar of Plant Health and 41st Annual Meeting of the Caribbean Division of the American Phytopathological Society, Cuba, June 10-15, 2001. APS On-line: http://www.apsnet.org/meetings/div/cr)1abs.asp.

Idris, A.M., and J.K. Brown. 2001. Satellite DNAs associated with monopartite begomoviruses of malvaceous hosts in Sudan. *Phytopathology* 91: S41.

Brown, J.K., A.M. Idris, L.V. Resende, G.E. de França. 2000. Vegetable and weed-infecting begomoviruses in the states of Pernambuco and Sao Paulo, Brazil. *Phytopathology* 90:S10.

Brown, J.K. and A.M. Idris. 2000. Molecular identification and phylogenetic relationships between cotton-infecting begomo-viruses from the America's, and their relationship to Old and New World begomoviruses of Malvaceous hosts. *Phytopathology* 90: S10.

Brown, J.K., A.M. Idris, L.V. Resende, and G.E. de França. 2000. Four new begomovirus species of vegetable and weed hosts in Pernambuco and Sao Paulo, Brazil. Proceedings in The Ninth Latin American and Caribbean Conference on Whiteflies and Geminiviruses. Panama City, Panama, November 21-24, 2000.

Brown, J.K., M. Palmieri, M.M. Doyle, A.M. Idris, and R.D. Martyn. 2000. Diversity and distribution of begomoviruses from vegetable and weed hosts in Guatemala and Honduras by comparative analyses of viral coat protein and common region nucleotide sequences. Proceedings in the Ninth Latin American and Caribbean Conference on Whiteflies and Geminiviruses. Panama City, Panama, November 21-24, 2000.

Idris, A.M., and J.K. Brown. 2000. Molecular characterization of full-length begomovirus DNAs from cotton, hollyhock and okra in Sudan in relation to the epidemiology of cotton leaf curl disease-Sudan. *Phytopathology* 90:S37.

Brown, J.K., M. Palmieri, M.M. Doyle, A.M. Idris, and R.D.Martyn. 2001. Diversity of begomoviruses from vegetables, cotton, and weeds in Guatemala and Honduras by viral coat protein sequence analysis. Page 72 in: Fourth International Seminar of Plant Health and 41st Annual Meeting of the Caribbean Division of the American Phytopathological Society, Cuba, June 10-15, 2001. *Phytopathology* S91: XX.

Idris, A.M., and J.K. Brown. 2001. Satellite DNAs associated with monopartite begomoviruses of malvaceous hosts in Sudan. *Phytopathology* 91: S41.

Popular Articles

Brown, JK. 2001. Global Diversity and Distribution of Whitefly-transmitted geminiviruses of cotton and Malvaceous hosts. California, Arizona, and Texas Cotton Magazine. Western Agriculture Publishing Company.

Websites

Gemini-Detect*ive* Website. Brown, J.K. and S.D. Wyatt, copyright 1997. Enables user to establish provisional identification of whitefly-transmitted viruses by comparative analysis using the core coat protein database provided at the site; site also provides PCR primers and parameters for amplification of the begomoviral core Cp sequence, together with information on geographical distribution, disease symptoms, and host range, plus relevant references. URL: http://gemini.biosci.arizona.edu

Compendia Contribution

Brown, J.K. 2002. Plant Virus and Mycoplasma Diseases of Cotton. Cotton Compendium. American Phytopathological Society, St. Paul, MN.

Invited Seminars, Symposia, & Meeting Presentations

Seminar: Whitefly-transmitted geminiviruses of cotton. Cotton Growers Organization, Caborca, Sonora, Mexico. August 21, 1997. Invitation from Delta Pine International.

Seminar: Phylogenetically informative sequences for begomoviruses and an interactive website/database for virus identification. Department of Plant Sciences, The University of Arizona. November 24, 1998.

Presentation: New and Emergent Cotton-Infecting Viruses Transmitted by Homopteran Vectors: New Problems of International Importance. 11th Brazilian Congress of Cotton, Ribeirao Preto, Brazil; September 6-14, 1999.