Virology **285,** 234–243 (2001)
doi:10.1006/viro.2001.0949, available online at http://www.idealibrary.com on IDEAL®

Identification of DNA Components Required for Induction of Cotton Leaf Curl Disease

R. W. Briddon,*[1] S. Mansoor,† I. D. Bedford,* M. S. Pinner,* K. Saunders,* J. Stanley,* Y. Zafar,† K. A. Malik,† and P. G. Markham*

*Department of Virus Research, John Innes Centre, Norwich Research Park, Colney Lane, Norwich NR4 7UH, United Kingdom; and †National Institute of Biotechnology and Genetic Engineering, P.O. Box 577, Jhang Road, Faisalabad, Pakistan

Received January 7, 2001; returned to author for revision March 7, 2001; accepted April 10, 2001

Cotton leaf curl disease (CLCuD) is a major constraint to cotton production in Pakistan. Infectious clones of the monopartite begomovirus cotton leaf curl virus (CLCuV), associated with diseased cotton, are unable to induce typical symptoms in host plants. We have identified and isolated a single-stranded DNA molecule approximately 1350 nucleotides in length which, when coinoculated 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 *Ageratum conyzoides*, represent members of an entirely new type of infectious, disease-causing agents. The implications of this finding to our understanding of the evolution of new disease-causing agents are discussed.

Key Words: geminivirus; whitefly; cotton; DNA β; recombination; evolution.

INTRODUCTION

The classification of the *Geminiviridae* divides the family into four genera. The basis of this division rests upon genome arrangement and the taxonomy of the insect vector, either whitefly (*Begomovirus*), leafhopper (*Mastrevirus* and *Curtovirus*), or treehopper (*Topocuvirus*). The majority of the economically important geminiviruses fall into the genus *Begomovirus*, which presently encompasses close to 100 distinct species which are all transmitted by the whitefly *B. tabaci* and infect only dicotyledonous plants (Briddon and Markham, 1995). Prominent examples of begomoviruses that cause significant agricultural losses include the cassava mosaic viruses (the factor most limiting to cassava production across south and central Africa (Thottappilly, 1992)) and *Tomato yellow leaf curl virus* (TYLCV), a severe problem to tomato cultivation across the Mediterranean region (Czosnek and Laterrot, 1997), which has now spread to the Caribbean (McGlashan *et al.,* 1994), Florida (Polston *et al.,* 1999), and numerous other countries.

The vast majority of begomoviruses have genomes consisting of two circular single-stranded DNA molecules of approximately equal size (Stanley and Gay, 1983; Rybicki, 1994). The genome component designated DNA A encodes viral functions required for replication (Stanley, 1983; Hanley-Bowdoin *et al.,* 1999), as well as the coat protein, which plays an essential role in insect transmission (Briddon *et al.,* 1989; Azzam *et al.,* 1994). The second component encodes two products involved in movement within and between cells in host tissues (Noueiry *et al.,* 1994). A small number of begomoviruses have genomes which consist of just a single component, equivalent to DNA A of the bipartite viruses. The most notable and economically most significant example of a monopartite begomovirus is TYLCV (Navot *et al.,* 1991).

During the last decade the output of cotton from Pakistan has been seriously compromised due to an epidemic of cotton leaf curl disease (CLCuD), with losses between 1992 and 1997 estimated at US$5 billion. The disease has very characteristic symptoms, including leaf curling, vein swelling, vein darkening, and enations on the veins on the underside of leaves, which frequently develop into leaf-like structures (Briddon and Markham, 2001). Although first described as early as the 1960s, the present epidemic originated in 1986 in the vicinity of the city of Multan (Hussain and Mahmood, 1988; Hussain *et al.,* 1991). Spreading mainly northeast under the influence of the prevailing winds, the disease crossed into the state of Punjab, India in 1995.

CLCuD was initially assumed to be caused by a begomovirus, due to its transmission by the whitefly *B. tabaci.* PCR-mediated amplification showed the presence of a begomovirus, named *Cotton leaf curl virus* (CLCuV), in CLCuD-infected cotton (Mansoor *et al.,* 1993). Attempts to identify a second genomic component (DNA B) were not successful, suggesting that CLCuV is a monopartite begomovirus. However, full-length clones of CLCuV, although infectious, were unable to induce the symptoms typical of CLCuD in cotton and other host plants. This indicated either that CLCuV plays no part in

[1] To whom correspondence and reprint requests should be addressed. Fax. +1603 450045. E-mail: rob.briddon@bbsrc.ac.uk.

0042-6822/01 $35.00

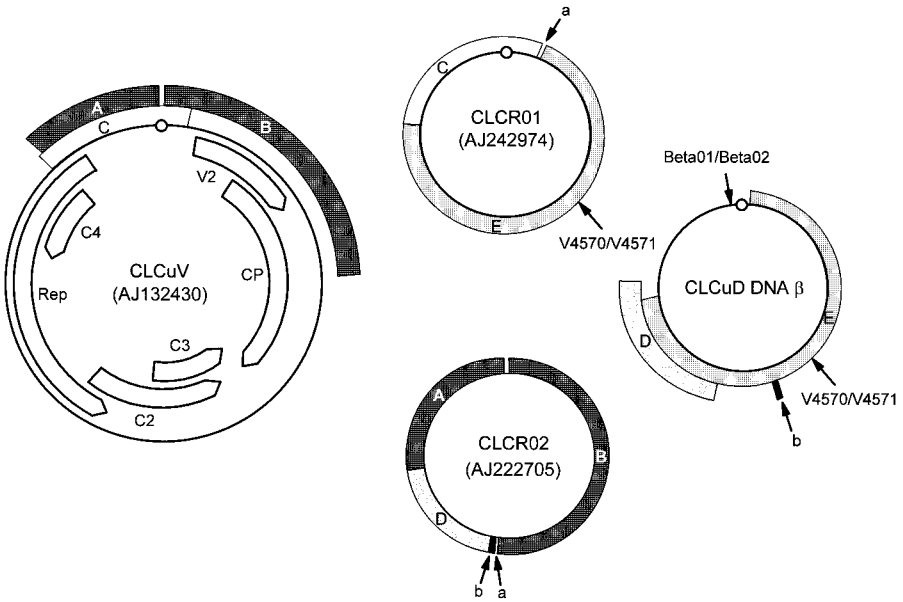


FIG. 1. Genome maps of CLCuV and and CLCuD DNA β showing the origins of sequences found in molecules CLCR01 and CLCR02. The fragments are labeled as A–C (originating from CLCuV) and D and E (originating from CLCuD DNA β). The positions of abutting primers, designed to the sequence of CLCR01 (V4570/V4571) and sequence unique to DNA β (primers Beta01/Beta02), used to amplify CLCuV DNA β are shown. The region of sequence on CLCR02 (marked as b) originates from CLCuD DNA β but is inserted in the reverse orientation. The origin of two small segments of DNA on CLCR01 and CLCR02 (marked as a) remains unclear.

causing CLCuD or that it requires some other factor(s) to induce symptoms of the disease in plants (Briddon *et al.,* 2000).

Subsequently a smaller (approx. 1350 nucleotides) single-stranded DNA molecule (termed DNA 1) was shown to be associated with CLCuD (Mansoor *et al.,* 1999). This molecule, capable of self-replication in plant cells but requiring CLCuV for spread in plants as well as insect transmission, was shown to play no part in inducing symptoms. DNA 1 shows high levels of similarity to nanovirus components which encode the replication associated protein (Rep). Nanoviruses are a genus of plant-infecting, multicomponent, single-stranded DNA viruses associated with diseases such as banana bunchy top and faba bean necrotic yellows (Timchenko *et al.,* 2000).

In this report we characterise an additional single-stranded DNA molecule isolated from CLCuD-affected tissue and demonstrate that the causative agent of the disease is a complex between this novel molecule and CLCuV.

RESULTS

A unique circular DNA molecule associated with cotton leaf curl disease

We have previously shown the presence of a 1346 nucleotide recombinant molecule (EMBL Accession No. AJ242974; which we shall henceforth refer to as cotton leaf curl recombinant 01 (CLCR01)), consisting of the intergenic region of CLCuV (Fig. 1) with the remaining sequence of indeterminate origin, in diseased cotton (Briddon *et al.,* 2000). A pair of nonoverlapping, 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 utilising primer pair V4570/V4571 amplified an approximately 1350-bp fragment. The product of this PCR was cloned into pGEM T-Easy vector (Promega), yielding approximately 10 clones. Restriction endonuclease analysis of the 10 clones shown all to have the same restriction pattern and to be different from CLCR01 (results not shown), 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 the sequence of CLCβ01 which is not present in CLCR01, also amplified an approximately 1350-bp 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.

Sequence 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 Nos. AJ292769 and AJ298903, respectively. The component

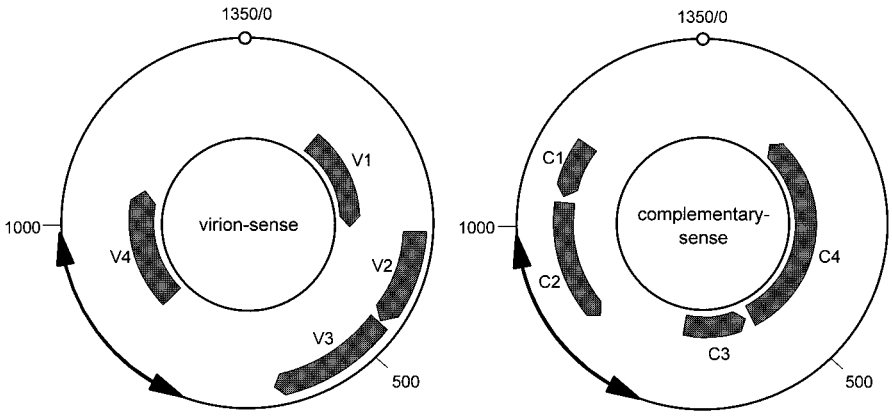


FIG. 2. Genome map of CLCuV DNA β showing the positions of predicted open reading frames (shown by arrows within the circle) encoded in both the virion-sense (left) and the complementary-sense (right) beginning in each case from an in-frame methionine codon. Only ORFs with a coding capacity above 4 kDa are shown. The position of the A-rich region is shown as a double-headed arrow.

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* (Saunders *et al.,* 2000). 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 (Laufs *et al.,* 1995; Stanley, 1995). 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 (J. Stanley and K. Saunders, unpublished results). The sequences of CLCuV DNA β contain an A-rich region between nucleotides 766 and 984 (58% A; Fig. 2) and show 96% overall nucleotide sequence similarity with the majority of the nucleotide changes occurring just downstream of the nonanucleotide motif in a putative noncoding 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 (Fig. 2). Of these ORFs, only two (V3 and C4) are predicted to be functional coding regions by the program TESTCODE (results not shown). This program predicts functional ORFs based on codon usage. The coding capacities of these ORFs are listed in Table 1.

A search of the EMBL nucleotide sequence databases identified the presence of a fragment of DNA β among 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 No. AJ222705; which we shall henceforth refer 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; Fig. 1). This finding demonstrates that DNA β is not unique to the CLCuD isolate examined as part of this

TABLE 1

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][a]	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

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

study. The sequence contained in the database entry AJ222705 is in the complementary sense and lacks the nonanucleotide motif due to the design of the primers.

Comparison to DNA β associated with Ageratum yellow vein disease

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 (Fig. 3A) shows the sequences to have the highest levels of similarity over a stretch of approximately 86 nucleotides upstream of the predicted hairpin structure. An alignment of this "conserved region" between the two DNA β molecules and its association to the TAATATTAC motif of the hairpin structure is shown (Fig. 3B).

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 (Table 1). 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).

Infectivity and symptoms

Both CLCβ01 and CLCβ02 were found to be infectious to cotton following coinoculation 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 postinoculation. The efficiency of infection to cotton following biolistic inoculation was 16% (10 plants infected of 60 inoculated, the result of four independent experiments). Inoculation of *Nicotiana benthamiana* with CLCuV and DNA β was less efficient (10%, 5 plants infected of 50 inoculated, the result of five 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 coinfection with CLCuV and CLCuD DNA β are illustrated in Fig. 4. The first leaf to develop symptoms exhibited deep downward cupping 2 days prior to appearance of full foliar symptoms (Fig. 4D). The symptoms exhibited were downward curling of the leaf margins, vein swelling, vein darkening, and enations (Figs. 4A and 4B), which eventually developed into leaf-like structures on the main veins on the undersides of symptomatic leaves (Fig. 4E). On a small number of plants upward curling of the leaf margins was evident (results not shown). 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 (Briddon *et al.,* 2000). *N. benthamiana* plants coinfected 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 (results not shown). 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).

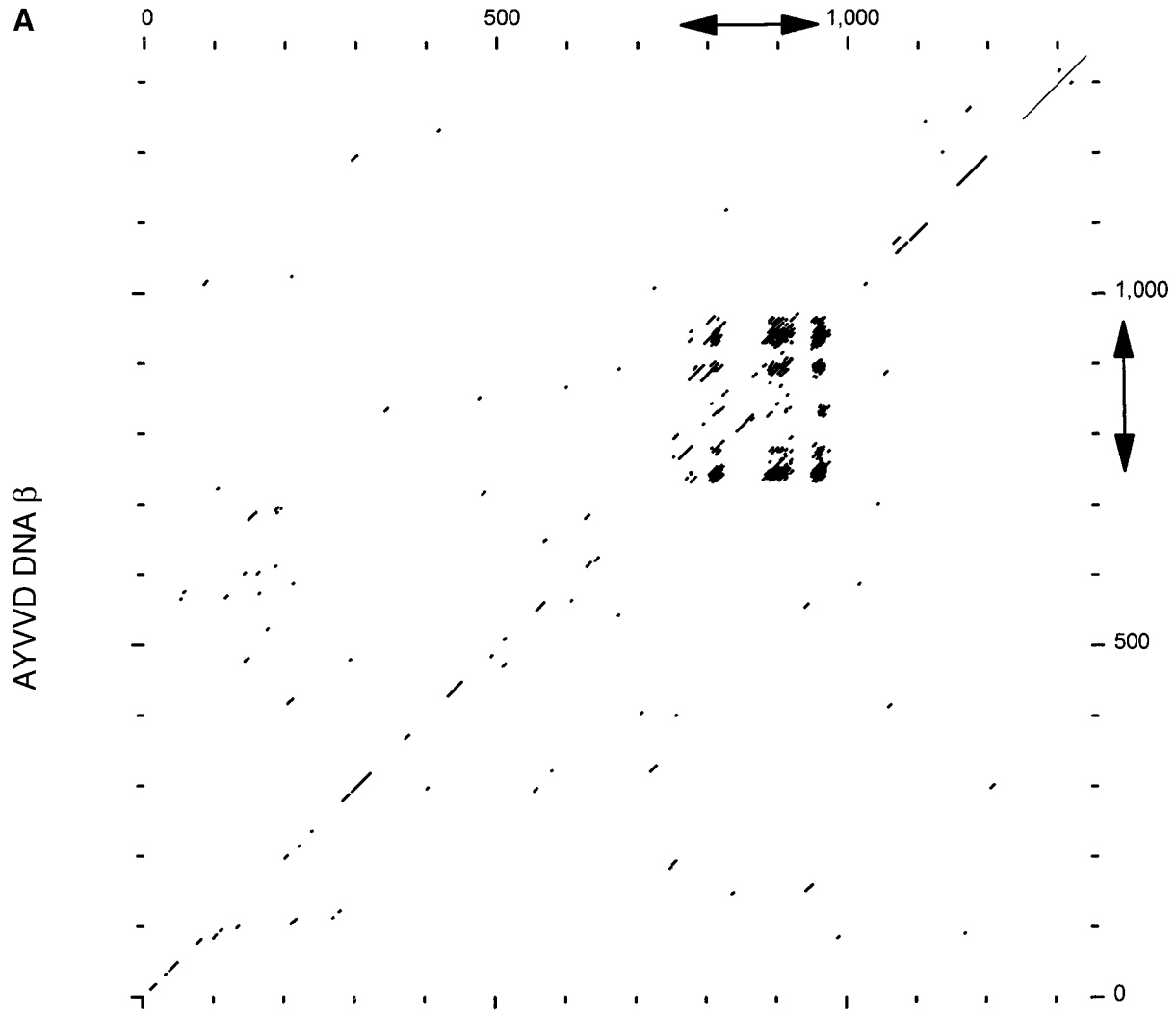
Analysis of CLCuV and DNA β inoculated plants

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 analysed for each species; results not shown). Southern blot analysis of cotton plants infected with both CLCuV and DNA β, compared to plants infected with a wild isolate of CLCuD, is shown in Fig. 5. 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 subgenomic; 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 coinfected with CLCuV and CLCuD DNA β to cotton plants infected with only CLCuV show an amplification of CLCuV in the presence of DNA β (Fig. 6). In coinfected plants CLCuV shows considerably higher levels of virus-specific subgenomic DNA molecules.

Southern blot analysis of cotton plants coinoculated 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 analysed; Fig. 6). 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.

Attempts at plant-to-plant transmission of progeny virions resulting from coinfection of cotton with CLCuV and



B

```
CLCuD-   1241 GTCGAATTTTACGACACGCGCGGCGGTGTGTACCCCTGGGAGGGTAGAAA 1290
AYVD-    1237 GCATCAATTCGCGACACGCGCGGCGGTGTGTACCCCTGGGAGGGTAGAAA 1286

CLCuD-   1291 CCACTACGCTACGCAGCAGCCTTAGCTACGCCGGAGCTTAGCTCGCCCAC 1340
AYVD-    1287 CCACTACGCTACGCAGCAGCCTTAGCTACGCCGGAGCTTAGCTCGCCACC 1336

                        *
CLCuD-   1341 GCTT      CGTGGGCGAGCGGTGCCTCCGGGGACGCAGTGGGTCC 49
AYVD-    1337 GTAA      CGGTGGCGAGCGGGGAATTTGGGTTCCAGGTGGGACC 49
                        *
```

FIG. 3. DOTPLOT comparison of the sequences of CLCuD DNA β (pCLCβ02) and AYVD DNA β with a window size of 21 and a stringency of 13 (A). The position of the A-rich region is shown for each sequence by a double-headed arrow. Alignment of the sequences of CLCuD DNA β and AYVD DNA β surrounding the predicted hairpin structure (B). The position of the nonanucleotide sequence is shown within a gray box and nucleotide number 1 is indicated by an asterisk. The "conserved region" between the two sequences is shown underlined.

CLCuD DNA β, using *B. tabaci*, were not successful (three independent experiments utilizing separate source plants, three healthy test seedlings, and approximately 1000 insects per experiment). Southern blot analysis of nucleic acids extracted from groups of 10 *B. tabaci* which were given a 48-h acquisition access pe-

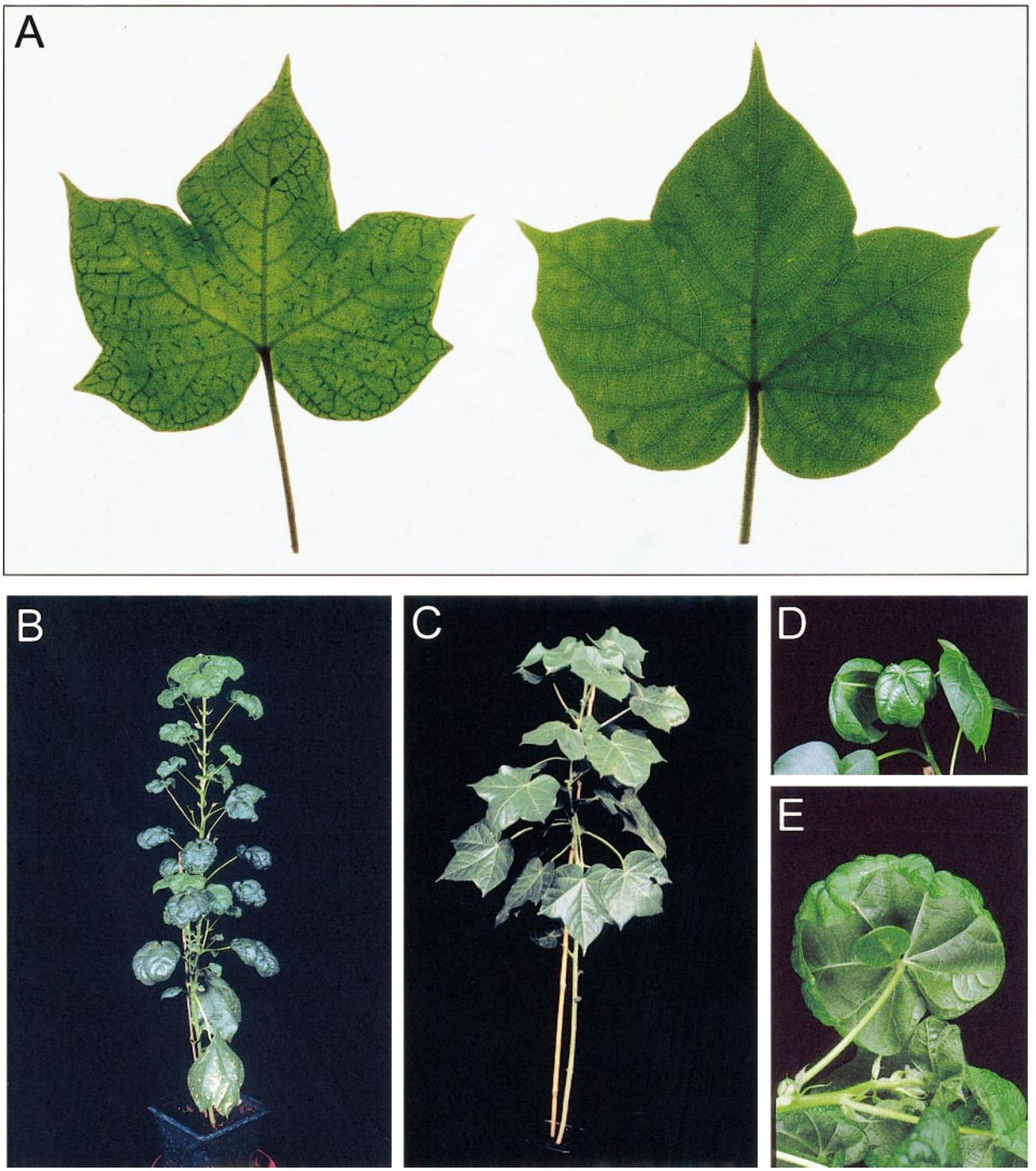


FIG. 4. Symptoms induced by infection of cotton with clones of CLCuV and CLCuD DNA β by biolistic inoculation. (A) Comparison of a leaf from an infected plant (left) with a leaf from a healthy plant (right) photographed with transmitted light. An infected (B) and a healthy plant (C) photographed approximately 2 months after inoculation. (D) Initial symptoms on the upper leaves of a cotton plant occurring 10 to 15 days after inoculation. (E) Leaf-like enations which develop on the main veins on the undersides of symptomatic leaves.

riod on cotton coinfected with CLCuV and CLCuD DNA β showed that both these components could be acquired (CLCuV and CLCuD DNA β ssDNA detected in four of eight groups of 10 insect extracts examined), consistent with DNA β being *trans*-encapsidated by the coat protein of CLCuV.

DISCUSSION

Cotton leaf curl disease is the major constraint to cotton production in Pakistan and is of increasing importance to the cotton industry of India. By analysing the sequences of indeterminate origin contained by recombinant, subgenomic DNA molecules associated with

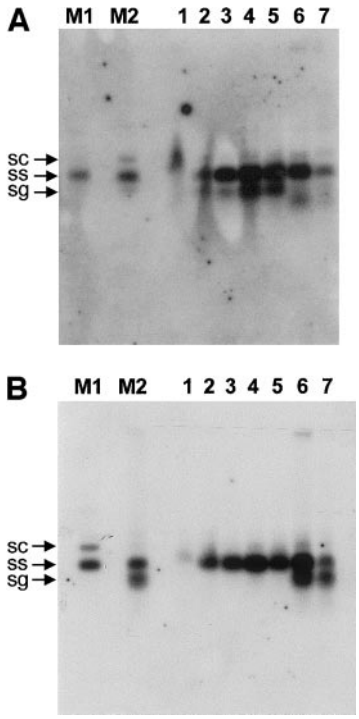


FIG. 5. Southern blot analysis of total nucleic acids extracted from cotton plants showing CLCuD symptoms following coinoculation with clones of CLCuV and CLCuD DNA β (lanes 1–7). Lanes marked as M1 and M2 were loaded with total nucleic acid extracts of an *N. benthamiana* plant (infected with CLCuD by graft inoculation) and cotton plant (field collected plant) showing symptoms of infection by CLCuD, respectively. Approximately equal amounts of nucleic acids (10 μg) were loaded in each lane. Duplicate blots were probed with radioactively labeled, full-length clones of CLCuV (A) and CLCuD DNA β (B). The positions of single-stranded (ss), supercoiled (sc), linear (lin), and subgenomic (sg) DNA forms are shown.

CLCuD, we have identified a novel molecule essential for symptom induction. CLCuD DNA β, in conjunction with CLCuV, represents the infectious unit responsible for CLCuD and is the first agriculturally destructive disease conclusively attributable to a begomovirus/DNA β complex.

CLCuD DNA β is approximately half the length of the begomovirus component. Unlike DNA 1, a half-genomic length molecule previously shown to be associated with CLCuD which shows high levels of sequence similarity to Rep encoding components of nanoviruses (Mansoor *et al.,* 1999; Saunders and Stanley, 1999), DNA β shows no significant similarity to sequences in databases, with the exception of AYVD DNA β. CLCuD DNA β shares little sequence homology with CLCuV, although it requires the virus for *trans*-replication and encapsidation as evidenced by the ability of *B. tabaci*, the vector of begomoviruses, to acquire DNA β from plants coinfected with CLCuV and DNA β. By analogy to the function of DNA B of bipartite begomoviruses, which encodes two gene products involved in virus movement (Noueiry *et al.,*1994), the product(s) encoded by DNA β may be involved in cell-to-cell movement. Alternatively it may act as a replication enhancer which manipulates the cellular environment to make it favorable for virus replication, possibly by overcoming host defense mechanisms such as gene silencing. The results of the analysis of plants infected with CLCuV alone and in conjunction with CLCuD DNA β, showing the presence of CLCuD DNA β leading to an increase in the levels of viral nucleic acids, is consistent with both these hypotheses.

Satellite RNAs of many plant RNA viruses have been described (reviewed by Roossinck *et al.,* 1992). By definition satellites have no sequence homology to their helper virus and are completely dependent upon their helper viruses for replication but may encode their own coat protein or rely on their helper virus for encapsidation. Satellites may also affect the titer of the helper virus in plants as well as alter the symptoms induced by the helper virus. Thus DNA β molecules should be seen, taxonomically, as DNA satellites of begomoviruses, rather than true virus components.

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 mol-

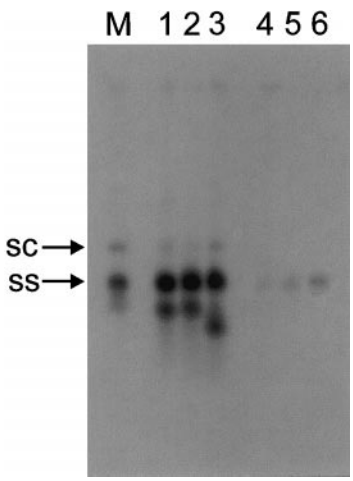


FIG. 6. Southern blot analysis of total nucleic acids extracted from cotton plants coinfected with CLCuV and CLCuD DNA β, showing typical CLCuD symptoms (lanes 1–3), and plants infected only with CLCuV (lanes 4–6), showing very mild symptoms, probed with radioactively labeled CLCuV. The sample in lane M was extracted from a field-collected cotton plant infected by CLCuD. The positions of single-stranded (ss) and supercoiled (sc) viral DNA forms are shown.

ecules 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.

Although clearly related and of common ancestry, the precise origin of the two DNA β molecules thus far identified and characterised remains unclear. *A. conyzoides* exhibiting yellow vein symptoms not dissimilar to those reported in Singapore occur throughout the Indian subcontinent and have been shown to be associated with a DNA 1 molecule (Mansoor *et al.,* 2000). It is likely therefore that *Ageratum* yellow vein disease on the Indian subcontinent is caused by a begomovirus–DNA β complex and is a possible source for the complex now affecting cotton crops. There are however a number of different plants species in Pakistan, India, and Nepal which exhibit symptoms comparable to CLCuD, including okra, *Hibiscus* spp., and hollyhock. Although shown to be associated with begomoviruses, identification and analysis of their possible DNA β components will be required to determine their relationship to the complex causing disease in cotton.

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 "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 nonanucleotide 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. Further analysis of the ability of other begomoviruses, particularly of those also associated with DNA β molecules, to *trans*-replicate DNA β molecules may provide an answer to this question.

The reason for the failure to insect transmit the progeny virions from plants coinfected with CLCuV and CLCuD DNA β remains unclear. The fact that both components can be acquired by *B. tabaci* shows that the coat protein of CLCuV can encapsidate both the begomovirus and the DNA β component and that these are able to interact with receptors in whiteflies predicted to be involved in virus acquisition (Liu *et al.,* 1998b). It is possible that the coat protein of the CLCuV clone used here has defective signal(s) for passage of the virus into the salivary glands of insects, thus preventing transmission. Further clones of CLCuV are being tested to investigate this possibility. The role that DNA 1 may play in insect transmission is also under investigation, although the results obtained with AYVD would suggest that this component is not required for insect transmission (Saunders *et al.,* 2000).

Although the origin of the DNA β molecules associated with CLCuD and AYVD remains unclear, the origin of their DNA 1 molecules rests, almost certainly, with the nanoviruses (Mansoor *et al.,* 1999; Saunders and Stanley, 1999). CLCuD complex is thus the product of a reassortment of virus genomic components, in this case a begomovirus (CLCuV), a component originating from nanoviruses (CLCuD DNA 1), and a new component of unknown origin (CLCuD DNA β), leading to a new disease. This finding illustrates a mechanism of evolution, component exchange, for single-stranded DNA viruses (Saunders and Stanley, 1999). Recently Timchenko *et al.* (2000) have suggested that reassortment of genomic components between different nanoviruses may be important in the evolution of new viruses and virus variants. The findings with CLCuD and AYVD, of the presumed capture of DNA 1 from nanoviruses, demonstrate that even reassortment and exchanges between viruses of different families are a factor in the evolution of these viruses (Saunders and Stanley, 1999). The possibility that DNA β originates with an, as yet, unidentified group of viruses cannot be ruled out.

Recombination between components has been shown to play a major part in the evolution of geminiviruses (Briddon *et al.,* 1996; Padidam *et al.,* 1999). Evidence presented here 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.

Further examples of this type of begomovirus-DNA β complex are under analysis, including examples isolated from honeysuckle and *Eupatorium* species, which should provide further information on the function and origin of the DNA β component. Recently Idris and Brown (2000) reported a monopartite begomovirus associated with cotton leaf curl disease originating from Sudan. It will be interesting to see whether this also requires a DNA β component to produce full disease symptoms in cotton.

The evidence presented here demonstrates that the

begomovirus CLCuV requires a unique single-stranded molecule, which we have named CLCuD DNA β, to induce symptoms of CLCuD in cotton. Thus CLCuD is a member of new type of infectious agents. This new group of pathogens consists typically of a monopartite begomovirus, a DNA 1 component originating from nanoviruses and a single-stranded satellite-like molecule (DNA β), which is approximately half the length of the begomovirus component and which has no sequence similarity to either the helper begomovirus or the DNA 1. CLCuD is the first crop disease attributable to this new group of pathogens and also includes *Ageratum* yellow vein disease (Saunders *et al.,* 2000), a disease of weed species.

MATERIALS AND METHODS

Virus collection and maintenance

The virus isolate (CLCV-C) was collected in cotton from the district of Multan, Pakistan and was maintained in cotton (*Gossypium hirsutum* cv. S12) both by insect transmission, using the whitefly *B. tabaci* (Gennadius), and by grafting infected scions to healthy cotton seedlings. Infected plants were maintained in insect-proof glasshouses at 25°C with supplementary lighting to give a 16-h photoperiod. Periodically the plants were transferred to a growth cabinet (Fisons) maintained at 35°C with a 16-h photoperiod to maintain symptoms.

PCR amplification, cloning, and sequence analysis

The isolation and sequencing of clone pCLCUV001, a full-length infectious clone of the begomovirus CLCuV, from virus isolate CLCuV-C has been described previously (Briddon *et al.,* 2000).

Full-length clones of the DNA β molecule were obtained by PCR amplification with primer pair V4570/V4571 (virion strand primer 5′-GTCGACTATGAACCTGACTCCCTCCTTG-3′ and complementary strand primer 5′-GTCGACGTTCGCATCATGAAGAATATGAAG-3′; these primers are designed around a naturally occurring *Sal*I restriction endonuclease site (underlined) occurring in both DNA β and CLCR01) and cloning of products into pGEM T-Easy vector (Promega) to yield clone CLCβ01. A further clone (CLCβ02) was obtained with primer pair Beta01/Beta02 (virion strand primer 5′-GGTACCACTACGCTACGCAGCAGCC-3′; complementary strand primer 5′-GGTACCTACCCTCCCAGGGGTACAC-3′; these primers introduce a unique *Kpn*I restriction endonuclease site (underlined)), designed to sequence which CLCβ01 does not share with CLCR01.

The complete nucleotide sequences of both DNA β clones were determined by dideoxynucleotide chain-termination sequencing using the PCR-based BIG DYE kit (Perkin–Elmer Cetus) and specific internal primers (Genosys). Reaction products were resolved on an ABI 377 automated sequencer. Sequence information was stored, assembled, and analysed using Version 7 of the program library of the Genetics Computer Group (Devereaux *et al.,* 1984).

Inoculation of plants

The inserts of clones pCLCUV001, CLCβ01, and CLCβ02 were excised from their cloning vectors using *Hin*dIII, *Sal*I, and *Kpn*I, respectively. DNA A and DNA β inserts were mixed in equal proportions, coated onto gold particles, and inoculated to *N. benthamiana* and cotton (*G. hirsutum* cv. S12) seedlings essentially as described previously (Briddon *et al.,* 1998, 2000). Cotton was inoculated into the growing tip and youngest leaf upon the expansion of the first true leaf. Following inoculation, cotton seedlings were maintained in a growth cabinet at a constant temperature of 35°C with a 16-h daylength. *N. benthamiana* seedlings were maintained in an insect-proof glasshouse at 25°C with supplementary lighting to give a 16-h daylength.

Detection of CLCuV and CLCuD DNA β components

Nucleic acids were isolated from plants and insects (in groups of 10) as previously described (Liu *et al.,* 1998) and electrophoresed in 1% agarose gels in TNE buffer (40 mM Tris acetate [pH 7.5], 20 mM sodium acetate, 2 mM EDTA). DNA was transferred to Hybond-NX (Amersham Pharmacia) and hybridised to [α^{32}P]dCTP (New England Nuclear)-labeled probes of full-length CLCuV and CLCuD DNA β clones.

ACKNOWLEDGMENTS

This work was funded by the Common Fund for Commodities under Contract ICAC/07. The project is a collaboration between the University of Arizona (Tucson), the National Institute of Biotechnology and Genetic Engineering (Faisalabad, Pakistan), the Cotton Research Institute (Faisalabad, Pakistan), and the JIC (Norwich, UK). Their inputs to this work are gratefully acknowledged. The authors also gratefully acknowledge the support of the Biotechnology and Biological Sciences Research Council. Viruses were held and manipulated with the authority of the Ministry of Agriculture, Fisheries, and Food under the Plant Health (Great Britain) Order 1993 (SI No.1993/1320); license numbers PHF 1419/819/113 and PHF 49/123(103).

REFERENCES

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-Verlag, Vienna.

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. *Pak. Cottons* **32,** 248–251.

Hussain, T., Tahir, M., and Mahmood, T. (1991). Cotton leaf curl virus. *Pak. 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.), Vol. 5, pp. 353–362. National Biomedical Research Foundation, Washington, D.C.

Stanley, J. (1983). Infectivity of the cloned geminivirus genome requires sequences from both DNAs. *Nature* **305,** 643–645.

Stanley, J. (1995). Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. *Virology* **206,** 707–712.

Stanley, J., and Gay, M. G. (1983). Nucleotide sequence of cassava latent virus DNA. *Nature* **301,** 260–262.

Thottappilly, G. (1992). Plant virus diseases of importance to African agriculture. *J. Phytopath.* **134,** 265–288.

Timchenko, T., Katul, L., Sano, Y., de Kouchkovsky, F., Vetten, H. J., and Gronenborn, B. (2000). The master Rep concept in nanovirus replication: Identification of missing genome components and potential for natural genetic reassortment. *Virology* **274,** 189–195.