ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

VIROLOGY

Virology 312 (2003) 106–121

www.elsevier.com/locate/yviro

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

Rob W. Briddon,[a,*] Simon E. Bull,[a] Imran Amin,[b] Ali M. Idris,[c] Shahid Mansoor,[b] Ian D. Bedford,[a] Poonam Dhawan,[d] Narayan Rishi,[d] Surender S. Siwatch,[e] Aly M. Abdel-Salam,[f] Judith K. Brown,[c] Yusuf Zafar,[b] and Peter G. Markham[a]

[a] *Department of Disease and Stress Biology, John Innes Centre, Colney Lane, Norwich, NR4 7UH, UK*
[b] *National Institute of Biotechnology and Genetic Engineering, P.O. Box 577, Jhang Road, Faisalabad, Pakistan*
[c] *Department of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA*
[d] *Department of Plant Pathology, CCS Haryana Agricultural University, Hisar-125 004, India*
[e] *Department of Plant Breeding, CCS Haryana Agricultural University, Hisar-125 004, India*
[f] *Plant Pathology Department, Faculty of Agriculture, Cairo University, Giza, Egypt*

Received 18 November 2002; returned to author for revision 20 December 2002; accepted 24 February 2003

Abstract

DNA β molecules are symptom-modulating, single-stranded DNA satellites associated with monopartite begomoviruses (family *Geminiviridae*). Such molecules have thus far been shown to be associated with *Ageratum yellow vein virus* from Singapore and *Cotton leaf curl Multan virus* from Pakistan. Here, 26 additional DNA β molecules, associated with diverse plant species obtained from different geographical locations, were cloned and sequenced. These molecules were shown to be widespread in the Old World, where monopartite begomoviruses are known to occur. Analysis of the sequences revealed a highly conserved organization for DNA β molecules consisting of a single conserved open reading frame, an adenine-rich region, and a region of high sequence conservation [the satellite conserved region (SCR)]. The SCR contains a potential hairpin structure with the loop sequence TAA/GTATTAC; similar to the origins of replication of geminiviruses and nanoviruses. Two major groups of DNA β satellites were resolved by phylogenetic analyses. One group originated from hosts within the *Malvaceae* and the second from a more diverse group of plants within the *Solanaceae* and *Compositae*. Within the two clusters, DNA β molecules showed relatedness based both on host and geographic origin. These findings strongly support coadaptation of DNA β molecules with their respective helper begomoviruses.

Keywords: geminivirus; begomovirus; satellite; DNA β; evolution; diversity

Introduction

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 (Murant and Mayo, 1982). The majority of satellites interfere with the replication of their helper viruses and attenuate symptoms. A small number of satellites, however, are known that exacerbate symptoms or produce novel symptoms not otherwise associated with the helper virus (reviewed by Collmer and Howell, 1992).

Begomoviruses (family Geminiviridae) are single-stranded DNA viruses that infect dicotyledonous plants, are transmitted by a single species of whitefly [*Bemisia tabaci* (Gennadius)], and typically have genomes consisting of two components (Rybicki et al., 2000). The DNA A component encodes all virus functions required for DNA replication, control of gene expression and insect transmission. The DNA B component encodes two genes involved in virus

* Corresponding author. Fax: +1603-450045.
E-mail address: rob.briddon@bbsrc.ac.uk (R.W. Briddon).

0042-6822/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0042-6822(03)00200-9

movement within plants (Stanley, 1983; Noueiry et al., 1994) and their products are symptom determinants for the bipartite begomoviruses (Klinkenberg and Stanley, 1990; von Arnim and Stanley, 1992). However, a small number of monopartite begomoviruses have been identified that lack the second component. For these viruses, all viral products required for replication, gene expression, whitefly transmission, and systemic infection are encoded on a single component (a homolog of the DNA A component of the bipartite begomoviruses; Navot et al., 1991; Rojas et al., 2001).

Previously, a single begomovirus, *Tomato leaf curl virus* (ToLCV), originating from Australia, was shown to be associated with a single-stranded DNA satellite molecule. This satellite had no discernable effects on either viral replication or on the symptoms caused by ToLCV (Dry et al., 1997). The ToLCV satellite (ToLCV-sat) is approximately 682 nt in length and sequence unrelated to ToLCV and requires ToLCV for replication, spread in plants, and insect transmission. Recently satellite molecules related to the ToLCV-sat were identified for several other monopartite begomoviruses. These molecules, called DNA β, are approximately 1350 nt in length (approximately half that of the genomes of their helper viruses) and, like the ToLCV-sat, are unrelated in sequence to their helper viruses as well as requiring their helper viruses for replication and movement in plants and insect transmission. Unlike the ToLCV-sat, DNA β satellites affect the replication of their helper begomoviruses and alter the symptoms induced in some host plants (Saunders et al., 2000; Briddon et al., 2001).

Both *Ageratum yellow vein virus* (AYVV) and *Cotton leaf curl Multan virus* (CLCuMV) are monopartite begomoviruses for which cloned genomes have been shown to be infectious to the experimental host *Nicotiana benthamiana* but were either not infectious or poorly infectious to the hosts from which they were isolated, *Ageratum conyzoides* and cotton respectively (Tan et al., 1995; Briddon et al., 2000). Additionally, clones of CLCuMV were found to induce symptoms in cotton that were not typical of cotton leaf curl disease (CLCuD). When inoculated with their respective DNA β components, both viruses induced typical disease symptoms in their natural hosts (Saunders et al., 2000; Briddon et al., 2001). For *Ageratum* yellow vein disease (AYVD) of *A. conyzoides* the typical symptoms are a bright yellow chlorosis along the veins. Symptoms for CLCuD of cotton are typically leaf curling, vein swelling, vein darkening, and the formation of enations on veins on the undersides of leaves that, in some varieties, develop into cup-shaped leaf-like structures (Briddon and Markham, 2000).

As well as DNA β, CLCuMV and AYVV are also associated with a second circular, single-stranded DNA molecule. The DNA 1 satellite-like molecules are approximately half the size of begomovirus components ($\approx$ 1350 nts) and encode a rolling-circle replication initiator protein with similarity to those of nanoviruses (Mansoor et al., 1999; Saunders and Stanley, 1999). Consequently, DNA 1 molecules are capable of self-replication in host plants but require the helper begomovirus for movement in plants as well as insect transmission.

In this study we have assessed the geographical and sequence diversity of the DNA β satellites 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 nucleotide sequences of DNA β molecules examined cluster as two major groups based on the plant host from which they were isolated. Within these two groups the DNA β molecules segregate both by host and geographic origin. Sequences comparisons identified conserved features of DNA β molecules that presumably play an important role in their function.

Results

Geographic distribution of DNA β-associated diseases

The geographic origins of diseases for which an association with a DNA β satellite was demonstrated are listed in Table 1. In addition to these, the use of either hybridization with heterologous DNA β probes or PCR-mediated amplification with universal DNA β primers (Briddon et al., 2002) showed the presence of a DNA β in three further disease isolates from *A. conyzoides*. Two, originating from Kenya and Uganda, showed a typical yellow vein net chlorosis very similar to that of AYVD originating from Pakistan (Fig. 1D). The third, *Ageratum* enation disease originating from Nepal, had symptoms consisting of a very mild vein darkening and minute enations on the veins on the undersides of leaves with no chlorosis.

A small number of begomovirus isolates originating from the New World were screened for the presence of DNA β and were shown not to contain the satellite. Isolates investigated included *Abutilon* mosaic disease, *Sida* golden mosaic disease (originating from Puerto Rico and Columbia), and cotton leaf crumple disease (originating from Arizona). These three disease have been shown to be caused by or associated with bipartite begomoviruses (Frischmuth et al., 1991; Höfer et al., 1997; A.M. Idris, manuscript in preparation). Nor were a number of plant diseases originating from the Old World associated with DNA β. Those investigated included 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). For these diseases clones of their associated begomoviruses have been demonstrated to cause the full range of symptoms in experimentally inoculated plants similar to those exhibited by the plants from which they were isolated. With the exception of tomato yellow leaf curl, these diseases are caused by bipartite begomoviruses (Navot

Table 1
Origins of disease isolates

Disease	Clone[a]	Origin (country/town[state]/year)	Plant species	Helper begomovirus	Symptoms[b]
Cotton leaf curl disease	CLCDβΔ01-Pak◇	Pakistan/Faisalabad/1994	*Gossypium hirsutum* [cotton]	cotton leaf curl viruses[c]	LC, VD, E, LE
Cotton leaf curl disease	CLCDβ01-Pak[d]◇	Pakistan/Faisalabad/1994	*Gossypium hirsutum* [cotton]	cotton leaf curl viruses[c]	LC, VD, E, LE
Cotton leaf curl disease	CLCDβ02-Pak[d]◇	Pakistan/Faisalabad/1994	*Gossypium hirsutum* [cotton]	cotton leaf curl viruses[c]	LC, VD, E, LE
Ageratum yellow leaf curl disease	AYLCDβ01-Pak	Pakistan/Faisalabad/2000	*Ageratum conyzoides*	nd	LC, VY, E
Ageratum yellow leaf curl disease	AYLCDβ02-Pak	Pakistan/Faisalabad/1994	*Ageratum conyzoides*	nd	LC, VY, VD
Zinnia leaf curl disease	ZLCDβ01-Pak	Pakistan/Faisalabad/1994	*Zinnia elegans*	nd	LC, VD
Zinnia leaf curl disease	ZLCDβ02-Pak	Pakistan/Faisalabad/1994	*Zinnia elegans*	nd	LC, VS
Hibiscus leaf curl disease	HLCDβΔ01-Pak	Pakistan/Faisalabad/1994	*Hibiscus rosa-sinensis*	nd	LC, VD, E, LE
Okra leaf curl disease	OLCDβ01-Pak	Pakistan/Gojra/1997	*Hibiscus esculentis* [okra]	nd	LC, YM
Okra leaf curl disease	OLCDβ02-Pak	Pakistan/Bahawalpur/1997	*Hibiscus esculentis* [okra]	nd	VY
Okra leaf curl disease	OLCDβ03-Pak	Pakistan/Bahawalpur/1997	*Hibiscus esculentis* [okra]	nd	VY
Chilli (pepper) leaf curl disease	ChLCDβ01-Pak	Pakistan/Mian Channu/1997	*Capsicum annuum* [chillies]	nd	LC, VS
Tobacco leaf curl disease	TobLCDβ01-Pak	Pakistan/Rahim Yar Khan/1998	*Nicotiana tabacum* [tobacco]	nd	LC
Tobacco leaf curl disease	TobLCDβ02-Pak	Pakistan/Bahawalpur/1999	*Nicotiana tabacum* [tobacco]	nd	LC
Tomato leaf curl disease	TomLCDβ01-Pak	Pakistan/Okara/1999	*Lycopersicon esculentum* [tomato]	nd	LC, E
Tomato leaf curl disease	TomLCDβ02-Pak	Pakistan/Rahim Yar Khan/1997	*Lycopersicon esculentum* [tomato]	nd	LC, YM
Cotton leaf curl disease	CLCDβΔ01-Ind□	India/Dabwali[Rajasthan]/1995	*Gossypium hirsutum* [cotton]	cotton leaf curl viruses[c]	LC, VD, E, LE
Cotton leaf curl disease	CLCDβΔ02-Ind□	India/Dabwali[Rajasthan]/1995	*Gossypium hirsutum* [cotton]	cotton leaf curl viruses[c]	LC, VD, E, LE
Cotton leaf curl disease	CLCDβ01-Ind□	India/Dabwali[Rajasthan]/1995	*Gossypium hirsutum* [cotton]	cotton leaf curl viruses[c]	LC, VD, E, LE
Ageratum yellow vein disease	AYVDβΔ01-Ind	India/[Punjab]/1997	*Ageratum conyzoides*	nd	VY
Okra yellow vein disease	BYVDβ01-Ind[e]	India/Madurai[Tamil Nadu]/?	*Hibiscus esculentis* [okra]	*Bhendi yellow vein mosaic virus*	VY
Okra yellow vein disease	OYVDβ01-Egy	Egypt/[Fayoum]/1995	*Hibiscus esculentis* [okra]	nd	VY
Okra leaf curl disease	OLCDβ01-Egy	Egypt/Cairo/2000	*Hibiscus esculentis* [okra]	*Cotton leaf curl Gezira virus*	LC
Okra leaf curl disease	OLCDβ02-Egy	Egypt/Cairo/2000	*Hibiscus esculentis* [okra]	*Cotton leaf curl Gezira virus*	LC
Okra yellow vein disease	OYVDβΔ01-Egy	Egypt/[Fayourn]/1995	*Hibiscus esculentis* [okra]	nd	VY, YM
Okra leaf curl disease	OLCDβΔ01-Egy	Egypt/Cairo/2000	*Hibiscus esculentis* [okra]	*Cotton leaf curl Gezira virus*	LC
Hollyhock leaf crumple disease	HLCrDβΔ01-Egy	Egypt/Cairo/1995	*Althea rosea* [hollyhock]	*Hollyhock leaf crumple virus*	LC, VD, E, LE
Hollyhock leaf crumple disease	HLCrDβΔ02-Egy	Egypt/Cairo/2000	*Althea rosea* [hollyhock]	*Hollyhock leaf crumple virus*	LC, VD, E, LE
Ageratum yellow vein disease	AYVDβ01-Sin[f]	Singapore/?	*Ageratum conyzoides*	*Ageratum yellow vein virus*	VY
Honeysuckle yellow vein mosaic disease	HYVMDβ01	United Kingdom/Norwich/1996	*Lonicera japonica* [honeysuckle]	*Honeysuckle yellow vein mosaic virus*	VY

Note. nd, not determined.

[a] Clones with the same symbol (◇ or □) were obtained from the same plant.

[b] LC, leaf curl; VD, vein darkening; VY, vein yellowing; VS, vein swelling; E, enations; LE, leaflike enations; YM, yellow mottling of leaves.

[c] At least four begomovirus species (previously called cotton leaf curl virus) are associated with CLCuD on the Indian sub-Continent (Zhou et al. 1998).

[d] Briddon et al. (2001).

[e] EMBL accession no. AJ308425

[f] Saunders et al. (2000).

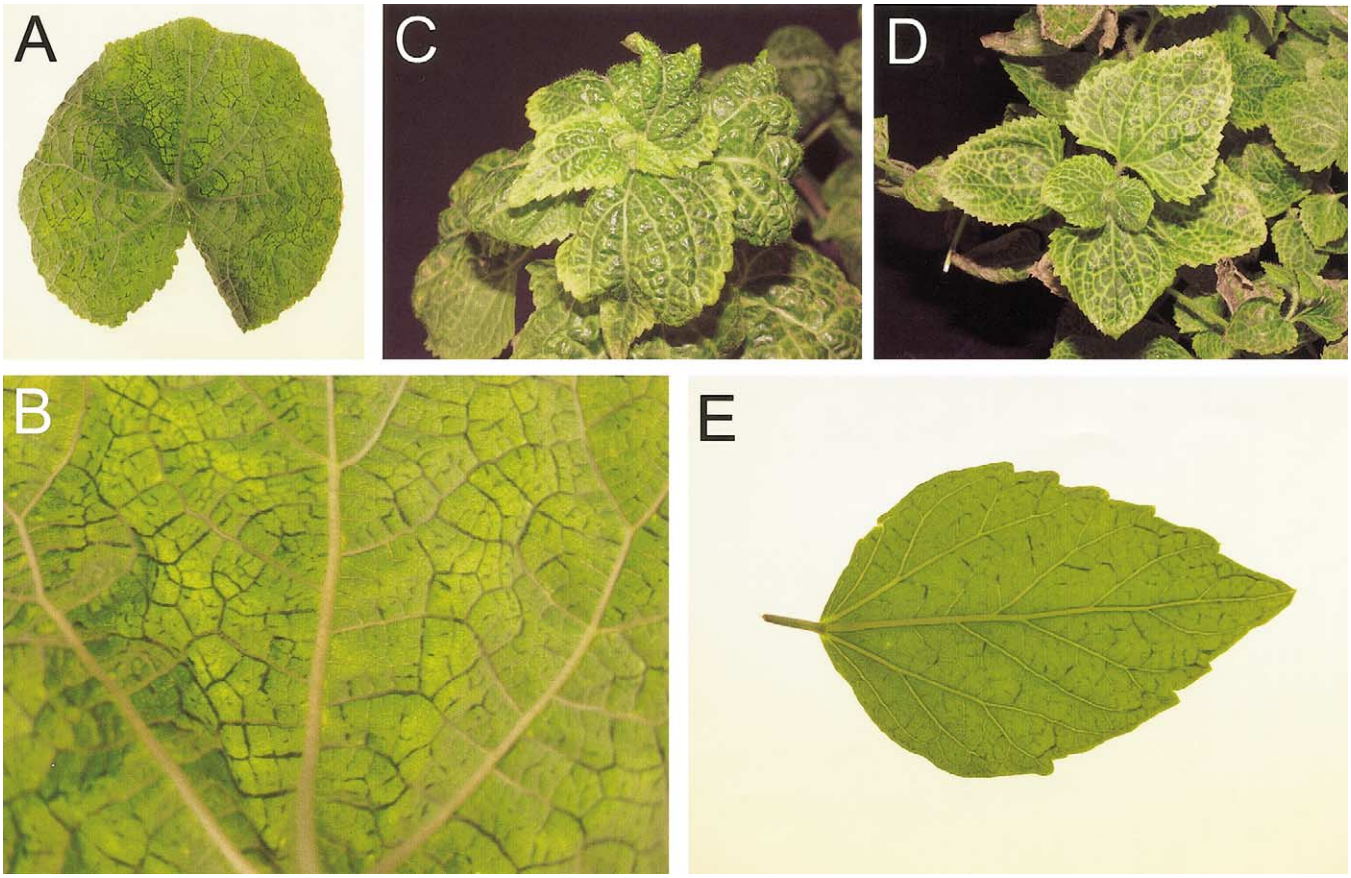


Fig. 1. Foliar symptoms of hollyhock leaf crumple disease (HLCrD) of *Althea rosea* originating from Egypt (A and B), *Ageratum* yellow leaf curl disease (AYLCD) of *A. conyzoides* from Pakistan (C), *Ageratum* yellow vein disease (AYVD) of *A. conyzoides* originating from India (D), and *Hibiscus* leaf curl disease (HLCD) of *Hibiscus rosa-sinensis* originating from Pakistan (E).

et al., 1991; Kheyr-Pour et al., 1992, 2000; Briddon et al., 1998).

Cloning and sequence determination

Clones of DNA β molecules were obtained by PCR-mediated amplification from 10 different plant species, originating from four countries, as indicated in Table 1. This table also indicates the helper virus, where known, with which each DNA β is associated. The foliar symptoms of four of the diseases for which clones of DNA β clones were obtained are shown in Fig. 1. 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 2. Hereafter DNA β satellites will be named after the disease with which the are associated: cotton leaf curl disease (CLCD), *Ageratum* yellow leaf curl disease (AYLCD), *Ageratum* yellow vein disease (AYVD), *Zinnia* leaf curl disease (ZLCD), *Hibiscus* leaf curl disease (HLCD), okra leaf curl disease (OLCD), okra yellow vein disease (OYVD), chilli leaf curl disease (ChLCD), tobacco leaf curl disease (TobLCD), tomato leaf curl disease (TomLCD), bhendi yellow vein disease (BYVD), hollyhock leaf crumple disease (HLCrD), and honeysuckle yellow vein mosaic disease (HYVMD).

In addition, the sequences of four further DNA β molecules were included in the analysis presented here; two originating from CLCuD affected cotton from Pakistan (CLCDβ01-Pak and CLCDβ02-Pak; Briddon et al., 2001), one isolated from yellow vein disease affected *A. conyzoides* originating from Singapore (AYVDβ01-Sin; Saunders et al., 2000), and one isolated from yellow vein disease affected okra from India [BYVDβ01-Ind (EMBL accession no. AJ308425)].

Of the 30 sequences analyzed, 25 represent potentially full-length DNA β molecules. Three of these were previously shown to be infectious to plants, when coinoculated with their corresponding begomovirus components, and induced typical disease symptoms (AYVDβ01-Sin, CLCDβ01-Pak, CLCDβ02-Pak; Saunders et al., 2000; Briddon et al., 2001). Of the remaining molecules, 5 (AYVD$\beta\Delta$01-Ind, OYVD$\beta\Delta$01-Egy, OLCD$\beta\Delta$01-Egy, HLCrD$\beta\Delta$01-Egy, and HLCrD$\beta\Delta$02-Egy) are believed to represent deletion mutants, because each has a size below 1000 nts.

Table 2
Features of DNA β molecules

Clone	Size (bp)	EMBL accession no.	C1 gene product [no. of amino acids (predicted molecular weight, kDa)]	Infectivity		
				N. benthamiana[a]	cotton[b]	*A. conyzoides*[c]
CLCDβΔ01-Pak	1247	AJ299443	99 (11.6)	β	Y	nd
CLCDβ01-Pak	1351	AJ292769	118 (13.7)	β	Y	nd
CLCDβ02-Pak	1349	AJ298903	118 (13.7)	β	Y	nd
AYLCDβ01-Pak	1351	AJ316026	138 (16.0)	β	nd	Y
AYLCDβ02-Pak	1352	AJ316027	138 (16.1)	β	nd	nd
ZLCDβ01-Pak	1349	AJ316041	74 (8.5)	A	N	nd
ZLCDβ02-Pak	1350	AJ316028	138 (16.1)	A	nd	nd
HLCDβ01-Pak	1347	AJ297908	105 (12.6)	A	nd	nd
OLCDβ01-Pak	1361	AJ316029	146 (17.2)	A	nd	nd
OLCDβ02-Pak	1362	AJ316030	99 (11.5)	A	nd	nd
OLCDβ03-Pak	1369	AJ316031	138 (16.1)	β	nd	nd
ChLCDβ01-Pak	1387	AJ316032	149 (17.5)	A	nd	nd
TobLCDβ01-Pak	1356	AJ316033	127 (14.6)	β	nd	nd
TobLCDβ02-Pak	1356	AJ316034	127 (14.6)	β	nd	nd
TomLCDβ01-Pak	1370	AJ316035	118 (13.6)	β	Y	nd
TomLCDβ02-Pak	1374	AJ316036	118 (13.7)	β	N	nd
CLCDβΔ01-Ind	1248	AJ291601	118 (13.7)	A	N	nd
CLCDβΔ02-Ind[d]	1353	AJ316037	77 (8.7)	A	N	nd
CLCDβ01-Ind	1351	AJ316038	118 (13.7)	β	Y	nd
AYVDβΔ01-Ind	896	AJ316042	—	nd	nd	nd
BYVDβ01-Ind	1353	AJ308425	140 (16.5)	nd	nd	nd
OYVDβ01-Egy	1307	AJ316039	117 (13.7)	β	Y	nd
OLCDβ01-Egy	1350	AF397217	117 (13.7)	β	nd	nd
OLCDβ02-Egy	1305	AF397215	117 (13.7)	β	nd	nd
OYVDβΔ01-Egy	759	AJ316043	—	nd	nd	nd
OLCDβΔ01-Egy	774	AF397216	—	nd	nd	nd
HLCrDβΔ01-Egy	660	AJ316044	—	A	nd	nd
HLCrDβΔ02-Egy	741	AF397214	—	nd	nd	nd
AYVDβ01-Sin	1347	AJ252072	118 (13.9)	β	nd	Y
HYVMDβ01	1344	AJ316040	116 (13.6)	β	nd	nd

[a] Infectivity of DNA β molecules determined by co-inoculation with AYVV to *N. benthamiana.* Symptoms induced are denoted as either typical of an AYVV infection of *N. benthamiana* in the absence of AYVD DNA β (producing upward leaf curling; A) or typical of an AYVV infection of *N. benthamiana* in the presence of AYVD DNA β (producing downward leaf curling; β).

[b] Infectivity determined by co-inoculation with CLCuMV to cotton. The infectivity is denoted as positive (Y) in which a proportion of inoculated plants developed symptoms and in which both CLCuMV and the DNA β molecule could be detected (as determined by Southern blot hybridization) or negative (N) in which no inoculated plants developed symptoms. In no plants was either only CLCuMV or only the DNA β molecule detected. Infected plants showed typical CLCuD symptoms. nd, not determined.

[c] Infectivity determined by co-inoculation with AYVV to *A. conyzoides.* The infectivity is denoted as positive (Y) in which a proportion of inoculated plants developed symptoms and in which both AYVV and the DNA β molecule could be detected (as determined by Southern blot hybridization). In no plants was either only AYVV or only the DNA β molecule detected. Infected plants showed typical AYVD symptoms. nd, not determined.

[d] Defective molecule with frameshift mutation in the C1 gene.

Sequence comparisons

Table 3 summarizes the relative nucleotide sequence similarities between DNA β satellites. These show between 49 and 99% similarity. For molecules that are unrelated (from different diseases and/or different geographic areas) the maximum nucleotide sequence similarity is around 71%.

Comparisons of the nucleotide sequences of all DNA β molecules, including both the apparently full-length molecules (those of over 1200 bases in length) and those presumed to be deletion mutants, identified three absolutely conserved features: (1) a predicted stem loop structure with the loop sequence TAA/GTATTAC [the nonanucleotide sequence (NS)], (2) a region of high sequence similarity (first identified by Briddon et al., (2001) that will henceforth be referred to as the "satellite conserved region" (SCR), and (3) an adenine- (A) rich region approximately 370 to 420 nts upstream of the SCR.

The total number of open reading frames (ORFs) identified differed for each of the DNA β molecules. However, the presumed full-length DNA β molecules contained only a single positionally and sequence conserved ORF. A consensus genome map for DNA β, showing the relative positions of each of the conserved features identified, is presented in Fig. 2.

Table 3
Pairwise sequence comparisons of the complete nucleotide sequences and predicted amino acid sequences of the putative gene C1 products of DNA β satellites

Clone	CLCDβΔ01-Pak	CLCDβ01-Pak	CLCDβ02-Pak	AYLCDβ01-Pak	AYLCDβ02-Pak	ZLCDβ01-Pak	ZLCDβ02-Pak	HLCDβ01-Pak	OLCDβ01-Pak	OLCDβ02-Pak	OLCDβ03-Pak	ChLCDβ01-Pak	TobLCDβ01-Pak	TobLCDβ02-Pak	TomLCDβ01-Pak	TomLCDβ02-Pak	CLCDβΔ01-Ind	CLCDβΔ02-Ind	CLCDβ01-Ind	AYVDβΔ01-Ind	BYVDβ01-Ind	OYVDβ01-Egy	OLCDβ01-Egy	OLCDβ02-Egy	OYVDβΔ01-Egy	OLCDβΔ01-Egy	HLCrDβΔ01-Egy	HLCrDβΔ02-Egy	AYVDβ01-Sin	HYVMDβ01
CLCDβΔ01-Pak	**100**	100	98.0	41.4	42.5	40.5	41.4	90.8	63.6	61.6	42.4	40.4	36.3	36.4	97.9	40.4	99.0	83.1	99.0	—	63.6	44.9	44.9	44.9	—	—	—	—	43.4	37.4
CLCDβ01-Pak	95.0	**100**	97.5	42.4	43.2	40.5	43.2	90.6	64.4	61.6	44.1	41.5	39.8	39.8	98.3	43.2	97.5	83.1	98.3	—	64.7	46.7	45.7	45.7	—	—	—	—	46.6	43.1
CLCDβ02-Pak	95.5	96.1	**100**	42.4	43.2	39.2	43.2	88.7	66.1	63.6	44.1	41.5	39.8	39.8	99.2	43.2	96.6	84.4	98.3	—	66.1	44.0	44.0	44.0	—	—	—	—	45.8	43.1
AYLCDβ01-Pak	52.6	50.5	50.8	**100**	97.8	94.6	97.1	38.7	43.1	43.4	96.4	78.2	72.4	72.4	42.4	84.7	46.7	37.6	41.5	—	43.8	47.9	47.9	47.9	—	—	—	—	67.8	53.5
AYLCDβ02-Pak	53.7	52.0	53.1	96.8	**100**	95.9	97.8	39.6	43.5	44.4	97.1	79.0	73.2	73.2	43.2	85.6	41.5	39.0	42.4	—	43.8	48.7	48.7	48.7	—	—	—	—	68.6	54.3
ZLCDβ01-Pak	54.2	50.8	52.0	95.5	97.8	**100**	98.6	35.5	36.5	37.8	93.2	87.3	73.2	78.4	39.2	86.5	39.2	32.4	39.2	—	37.8	45.9	46.0	45.9	—	—	—	—	77.0	54.1
ZLCDβ02-Pak	54.0	51.5	52.7	96.0	98.2	99.4	**100**	40.6	43.8	43.4	96.4	80.4	78.4	74.8	43.2	87.3	41.5	37.7	42.4	—	44.5	49.6	49.6	49.6	—	—	—	—	70.3	56.0
HLCDβ01-Pak	93.7	98.3	95.5	49.9	51.5	51.5	51.9	**100**	61.3	56.3	40.6	39.6	74.8	37.7	88.7	41.5	87.7	67.7	88.7	—	61.3	40.0	40.0	40.0	—	—	—	—	43.4	42.9
OLCDβ01-Pak	62.9	62.1	63.0	51.8	53.5	53.1	53.2	62.5	**100**	96.0	44.5	44.9	42.5	42.5	66.1	46.6	63.6	57.1	65.3	—	99.3	46.2	46.2	46.2	—	—	—	—	44.9	44.8
OLCDβ02-Pak	59.1	59.2	60.0	50.9	53.1	52.4	52.1	59.7	79.4	**100**	44.4	45.5	42.4	42.4	63.6	47.5	62.6	62.3	62.6	—	97.0	43.4	43.4	43.4	—	—	—	—	43.4	41.4
OLCDβ03-Pak	55.2	52.3	53.7	95.3	97.4	96.6	97.0	50.8	53.4	51.8	**100**	79.0	74.0	74.0	44.1	85.6	42.4	39.0	43.2	—	45.3	49.6	49.6	49.6	—	—	—	—	69.5	55.2
ChLCDβ01-Pak	54.6	52.2	53.3	68.6	71.6	70.5	71.2	53.0	53.5	51.3	71.5	**100**	74.0	74.0	41.5	87.3	39.8	37.7	40.7	—	46.3	47.9	47.9	47.9	—	—	—	—	66.9	52.6
TobLCDβ01-Pak	55.3	55.6	54.8	61.0	63.4	63.0	63.3	56.3	55.3	53.0	62.6	61.9	**100**	100	39.8	80.5	37.3	37.7	39.0	—	41.7	47.0	47.0	47.0	—	—	—	—	67.8	54.3
TobLCDβ02-Pak	55.2	54.9	54.7	61.8	63.4	63.3	63.1	56.2	55.3	52.9	61.6	61.7	99.8	**100**	43.2	80.5	40.7	39.0	42.4	—	48.3	47.0	47.0	47.0	—	—	—	—	67.8	56.0
TomLCDβ01-Pak	90.8	93.0	92.4	53.4	55.0	54.2	54.7	92.3	57.9	57.0	54.2	51.6	58.0	57.6	**100**	39.8	97.5	84.4	98.3	—	66.1	44.0	44.0	44.0	—	—	—	—	45.8	43.1
TomLCDβ02-Pak	54.0	51.9	52.2	68.6	70.2	70.3	69.7	51.0	52.0	50.3	71.2	79.9	61.9	61.9	54.1	**100**	37.3	37.7	39.0	—	41.7	43.0	43.0	47.0	—	—	—	—	67.8	54.3
CLCDβΔ01-Ind	95.6	95.1	96.5	53.2	56.0	55.1	56.2	93.9	62.9	60.0	55.3	53.0	53.0	53.6	90.5	54.5	**100**	84.4	97.5	—	63.6	43.0	43.0	43.1	—	—	—	—	44.1	40.5
CLCDβΔ02-Ind	97.4	96.4	98.7	50.9	52.5	52.2	51.6	95.8	63.5	60.3	51.4	53.0	55.7	54.9	92.6	52.5	97.6	**100**	84.4	—	63.6	43.4	43.4	43.4	—	—	—	—	40.3	37.6
CLCDβ01-Ind	97.3	96.7	99.0	51.4	53.6	52.4	53.2	96.0	63.5	60.8	52.0	52.8	55.0	54.2	92.8	52.3	97.3	99.6	**100**	—	65.2	44.8	44.8	44.8	—	—	—	—	44.9	42.2
AYVDβΔ01-Ind	50.2	52.6	52.2	74.7	77.5	78.1	78.7	52.2	53.3	51.5	79.1	59.2	57.7	57.7	56.9	59.2	49.0	52.9	53.1	**100**	—	—	—	—	—	—	—	—	—	—
BYVDβ01-Ind	60.9	57.8	60.2	50.0	52.3	53.0	51.6	60.8	79.8	99.8	50.5	51.7	53.6	53.9	54.9	53.1	60.8	61.1	61.3	50.7	**100**	47.0	47.0	47.0	—	—	—	—	44.1	44.8
OYVDβ01-Egy	50.8	55.7	55.3	50.2	51.7	53.0	54.0	56.3	51.5	53.8	52.4	51.1	54.3	54.5	55.2	52.1	53.4	55.6	55.9	49.5	53.4	**100**	100	100	—	—	—	—	47.9	46.6
OLCDβ01-Egy	53.3	49.5	50.6	48.4	49.9	50.2	50.9	50.7	49.6	49.1	50.6	50.4	51.5	51.5	50.4	50.5	53.2	49.6	49.5	50.2	49.4	96.9	**100**	100	—	—	—	—	47.9	46.6
OLCDβ02-Egy	51.8	56.2	56.4	50.8	53.1	53.2	54.8	57.2	51.6	54.4	53.1	52.8	54.5	54.7	56.1	54.5	53.8	56.0	55.8	49.2	53.1	98.5	98.3	**100**	—	—	—	—	47.9	46.6
OYVDβΔ01-Egy	51.9	58.4	57.2	48.9	51.7	51.7	53.5	58.3	56.6	51.6	51.9	52.7	55.5	55.3	56.9	53.2	53.5	57.5	57.5	53.2	50.7	89.9	88.1	91.0	**100**	—	—	—	—	—
OLCDβΔ01-Egy	50.9	57.2	56.9	48.5	51.6	51.3	52.2	58.9	54.6	54.7	48.8	52.2	53.8	53.6	58.0	53.9	53.8	57.2	57.1	52.5	53.0	89.4	87.9	90.8	99.5	**100**	—	—	—	—
HLCrDβΔ01-Egy	54.1	58.6	57.0	50.4	54.7	56.8	58.0	57.4	60.1	58.5	57.4	55.2	54.7	59.1	53.4	56.7	53.2	57.4	54.8	55.2	56.6	74.2	75.5	75.8	89.5	85.4	**100**	—	—	—
HLCrDβΔ02-Egy	50.1	55.2	55.3	46.0	50.6	49.8	50.4	57.5	52.7	53.2	51.0	53.2	55.7	54.8	56.3	50.8	49.7	57.7	57.0	52.3	54.1	82.3	81.6	83.1	91.2	92.4	90.6	**100**	—	—
AYVDβ01-Sin	55.4	53.2	53.0	56.5	59.2	59.8	60.6	53.3	52.6	53.6	60.9	59.7	61.9	61.9	56.3	58.9	56.4	54.2	54.0	55.1	53.7	52.0	50.0	52.2	56.1	53.8	58.0	54.4	**100**	52.6
HYVMDβ01	48.6	49.9	50.0	55.2	55.6	54.5	55.3	51.2	48.2	47.0	55.7	52.5	54.0	53.7	51.3	51.1	50.0	50.5	50.5	52.3	47.9	50.2	49.2	50.2	48.0	51.2	53.1	52.0	52.9	**100**

Note. Percentage nucleotide sequence similarities between the complete nucleotide sequences are shown below the diagonal. Percentage amino acid sequence similarities of the putative gene C1 products are shown above the diagonal.

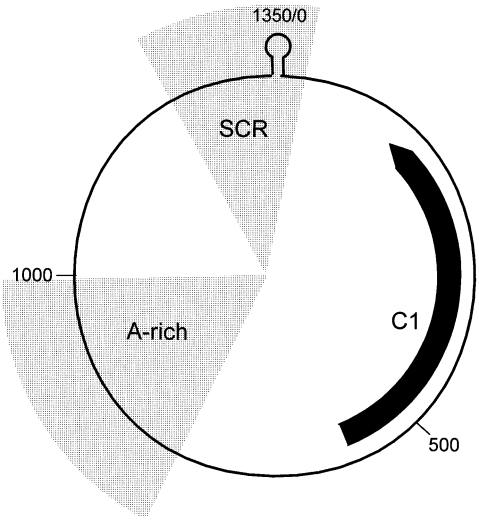


Fig. 2. Consensus genome map for DNA β molecules. The position and orientation of the conserved C1 gene is shown as an arrow within the circle. The relative position of the satellite conserved region (SCR) and A-rich region are shown as shaded slices.

Analysis of the "satellite conserved region" (SCR)

The SCR is approximately 200 nts in length and encompasses, at the 3′ end, a predicted hairpin structure with similarity to the origin of virion-strand DNA replication for the geminiviruses and nanoviruses. The majority of the DNA β molecules analyzed contain the sequence TAATATTAC within the loop, typical of the geminiviruses. Only a single DNA β molecule (OLCDβ02-Pak) was identified with the loop sequence TAGTATTAC, typical of the nanoviruses. The sequences of the stem of the hairpin structure of all DNA β molecules show near identity (Fig. 3). This is in contrast to both the geminiviruses and the nanoviruses, as well as the begomovirus-associated DNA 1 molecules, where each "species" has a distinct stem sequence.

The SCR of the DNA β molecules examined were greater than 65% conserved compared to a typical level of 50% overall nucleotide sequence conservation between distinct DNA β molecules. Within the SCR were three blocks of conserved sequence (as indicated in Fig. 3) interspersed with sequence conserved only between related DNA β molecules (being the Indian subcontinent *Malvaceae*-isolated DNA β molecules, the Egyptian *Malvaceae*-isolated molecules, and the non-malvaceous DNA β molecules). The largest of the conserved sequence blocks (indicated as Block A in Fig. 3) is up to 105 nts in length and encompasses the predicted hairpin structure with the NS. The sequence of Block A was greater than 75% conserved. Of the DNA β molecules, the SCRs of clones AYLCDβ01-Pak and HYVMDβ01 were the most diverse. A further well-conserved sequence block (Fig. 3D) occurs outside the SCR, just downstream of the hairpin structure.

Although clearly related to DNA β, the region of the ToLCV-sat equivalent to the SCR has less than 58% nucleotide sequence similar to the sequences of DNA β. The most prominent difference was around the predicted hairpin structure (illustrated by a 40-base insertion immediately upstream of the hairpin; Fig. 3). Consequently the sequence of the stem of the hairpin structure of ToLCV-sat is distinct from that of the DNA β molecules.

Analysis of A-rich region

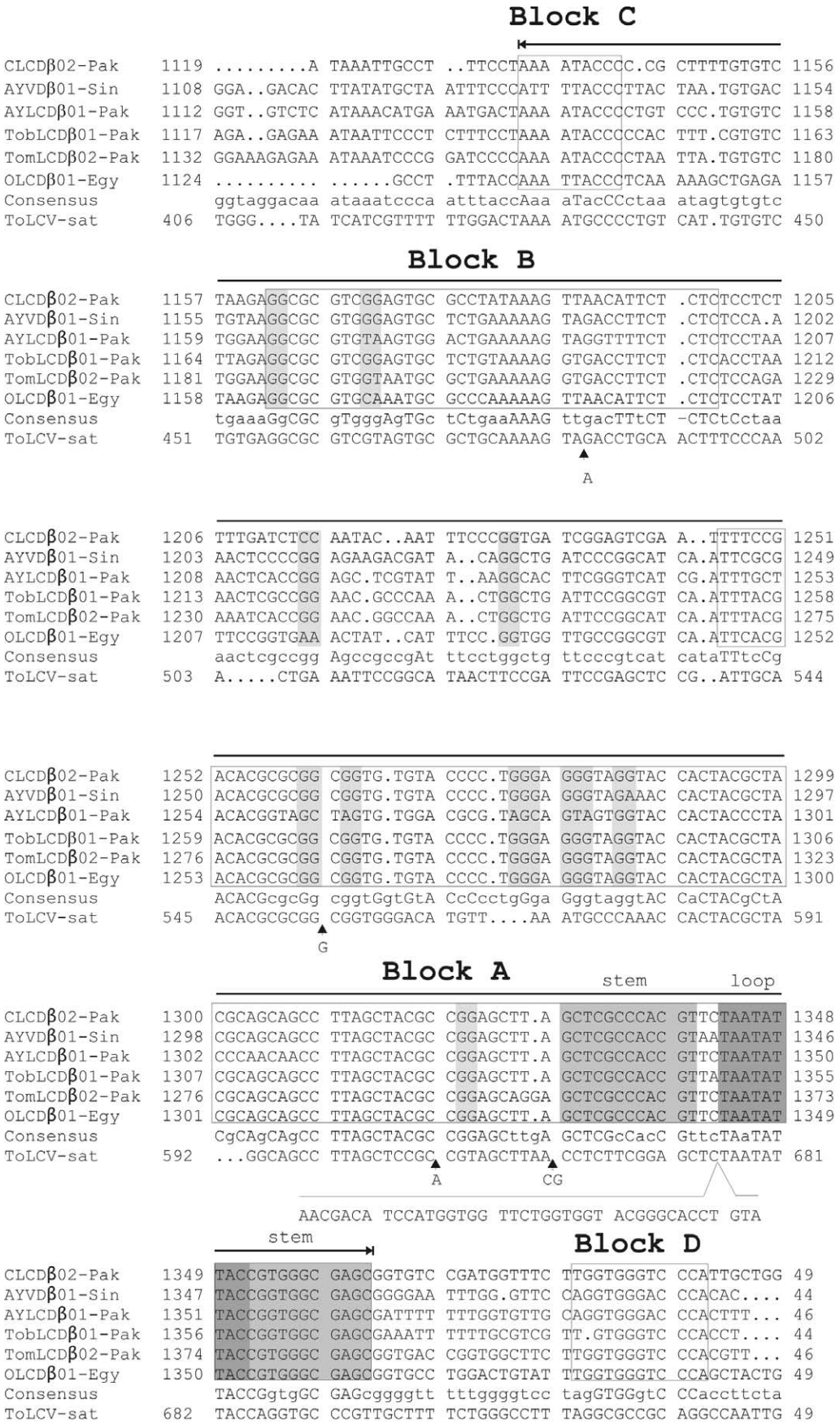
The A-rich region is typically between 160 and 280 bases in length and has between 57 and 65% A content (the overall A content of DNA β molecules is between 28 and 38%). The A-rich region is maintained for all DNA β molecules as well as the ToLCV-sat. This region has repeated blocks of up to 11 consecutive A residues. The sequences of the A-rich regions of the DNA β molecules segregate with and form the same groups as the SCR sequences in phylogenetic analyses (as discussed later).

Analysis of the potential coding region

The single conserved ORF identified for all full-length DNA β molecules was previously identified as ORF C1 by Saunders et al., (2000) for AYVD DNA β and ORF C4 for CLCuD DNA β by Briddon et al., (2001) and is encoded on the complementary strand (Fig. 2). We shall henceforth refer to this ORF as complementary-sense gene 1 (gene C1). The features of the product predicted to be encoded by gene C1 of the full-length DNA β molecules are summarized in Table 2. An alignment of the predicted amino acid sequences of all gene C1 products is shown in Fig. 4.

The typical length of the product predicted to be encoded by gene C1 is 118 amino acids with a predicted molecular weight of 13.7 kDa. Some putative C1 proteins, such as that of AYLCDβ01-Pak, have an extended amino terminal leader (Fig. 4). However, a conserved downstream methionine (indicated with an arrow in Fig. 4) provides a potential initiation site for translation that would produce an approx-

Fig. 3. Alignment of the nucleotide sequences of the satellite conserved regions of selected DNA β molecules to the homologous region of ToLCV-sat. Gaps (.) were introduced to optimize the alignment. Insertions in the sequence of ToLCV-sat are shown below the alignment. A consensus sequence (for all DNA β molecules but excluding the ToLCV-sat) is shown with absolutely conserved nucleotides shown in uppercase. The extent of the satellite conserved region is indicated by the arrow above the alignment. The conserved hairpin structure is indicated by the dark shading (for the nonanucleotide sequence which forms part of the loop) and the two lighter shaded boxes (for the stem). Well-conserved GG dinucleotide sequences within the conserved region are indicated by light shading. The highly conserved sequence blocks (A, B, and C) within the SCR are indicated by boxes as well as a conserved block of sequence just downstream of the hairpin structure (D).


Block C
CLCDβ02-Pak 1119 .....A TAAATTGCCT ..TTCCTAAA ATACCCC.CG CTTTTGTGTC 1156
AYVDβ01-Sin 1108 GGA..GACAC TTATATGCTA ATTTCCCATT TTACCCTTAC TAA.TGTGAC 1154
AYLCDβ01-Pak 1112 GGT..GTCTC ATAAACATGA AATGACTAAA ATACCCCTGT CCC.TGTGTC 1158
TobLCDβ01-Pak 1117 AGA..GAGAA ATAATTCCCT CTTTCCTAAA ATACCCCCAC TTT.CGTGTC 1163
TomLCDβ02-Pak 1132 GGAAAGAGAA ATAAATCCCG GATCCCCAAA ATACCCCTAA TTA.TGTGTC 1180
OLCDβ01-Egy 1124 ....GCCT .TTTACCAAA TTACCCTCAA AAAGCTGAGA 1157
Consensus ggtaggacaa ataaatccca atttaccAaa aTacCCctaa atagtgtgtc
ToLCV-sat 406 TGGG....TA TCATCGTTTT TTGGACTAAA ATGCCCCTGT CAT.TGTGTC 450
Block B
CLCDβ02-Pak 1157 TAAGAGGCGC GTCGGAGTGC GCCTATAAAG TTAACATTCT .CTCTCCTCT 1205
AYVDβ01-Sin 1155 TGTAAGGCGC GTGGGAGTGC TCTGAAAAAG TAGACCTTCT .CTCTCCA.A 1202
AYLCDβ01-Pak 1159 TGGAAGGCGC GTGTAAGTGG ACTGAAAAAG TAGGTTTTCT .CTCTCCTAA 1207
TobLCDβ01-Pak 1164 TTAGAGGCGC GTCGGAGTGC TCTGTAAAAG GTGACCTTCT .CTCACCTAA 1212
TomLCDβ02-Pak 1181 TGGAAGGCGC GTGGTAATGC GCTGAAAAAG GTGACCTTCT .CTCTCCAGA 1229
OLCDβ01-Egy 1158 TAAGAGGCGC GTGCAAATGC GCCCAAAAAG TTAACATTCT .CTCTCCTAT 1206
Consensus tgaaaGgCGC gTgggAgTGc tCtgaaAAAG ttgacTTtCT -CTCtCctaa
ToLCV-sat 451 TGTGAGGCGC GTCGTAGTGC GCTGCAAAAG TAGACCTGCA ACTTTCCCAA 502
A
CLCDβ02-Pak 1206 TTTGATCTCC AATAC..AAT TTCCCGGTGA TCGGAGTCGA A..TTTTCCG 1251
AYVDβ01-Sin 1203 AACTCCCCGG AGAAGACGAT A..CAGGCTG ATCCCGGCAT CA.ATTCGCG 1249
AYLCDβ01-Pak 1208 AACTCACCGG AGC.TCGTAT T..AAGGCAC TTCGGGTCAT CG.ATTTGCT 1253
TobLCDβ01-Pak 1213 AACTCGCCGG AAC.GCCCAA A..CTGGCTG ATTCCGGCGT CA.ATTTACG 1258
TomLCDβ02-Pak 1230 AAATCACCGG AAC.GGCCAA A..CTGGCTG ATTCCGGCAT CA.ATTTACG 1275
OLCDβ01-Egy 1207 TTCCGGTGAA ACTAT..CAT TTCC.GGTGG TTGCCGGCGT CA.ATTCACG 1252
Consensus aactcgccgg AgccgccgAt ttcctggctg ttcccgtcat cataTTtcCg
ToLCV-sat 503 A.....CTGA AATTCCGGCA TAACTTCCGA TTCCGAGCTC CG..ATTGCA 544
CLCDβ02-Pak 1252 ACACGCGCGG CGGTG.TGTA CCCC.TGGGA GGGTAGGTAC CACTACGCTA 1299
AYVDβ01-Sin 1250 ACACGCGCGG CGGTG.TGTA CCCC.TGGGA GGGTAGAAAC CACTACGCTA 1297
AYLCDβ01-Pak 1254 ACACGGTAGC TAGTG.TGGA CGCG.TAGCA GTAGTGGTAC CACTACCCTA 1301
TobLCDβ01-Pak 1259 ACACGCGCGG CGGTG.TGTA CCCC.TGGGA GGGTAGGTAC CACTACGCTA 1306
TomLCDβ02-Pak 1276 ACACGCGCGG CGGTG.TGTA CCCC.TGGGA GGGTAGGTAC CACTACGCTA 1323
OLCDβ01-Egy 1253 ACACGCGCGG CGGTG.TGTA CCCC.TGGGA GGGTAGGTAC CACTACGCTA 1300
Consensus ACACGcgcGg cggtGgtGtA CcCcctgGga GggtaggtAC CaCTACgCtA
ToLCV-sat 545 ACACGCGCGG CGGTGGGACA TGTT....AA ATGCCCAAAC CACTACGCTA 591
G
Block A
stem
loop
CLCDβ02-Pak 1300 CGCAGCAGCC TTAGCTACGC CGGAGCTT.A GCTCGCCCAC GTTCTAATAT 1348
AYVDβ01-Sin 1298 CGCAGCAGCC TTAGCTACGC CGGAGCTT.A GCTCGCCACC GTAATAATAT 1346
AYLCDβ01-Pak 1302 CCCAACAACC TTAGCTACGC CGGAGCTT.A GCTCGCCACC GTTCTAATAT 1350
TobLCDβ01-Pak 1307 CGCAGCAGCC TTAGCTACGC CGGAGCTT.A GCTCGCCACC GTTATAATAT 1355
TomLCDβ02-Pak 1276 CGCAGCAGCC TTAGCTACGC CGGAGCAGGA GCTCGCCCAC GTTCTAATAT 1373
OLCDβ01-Egy 1301 CGCAGCAGCC TTAGCTACGC CGGAGCTT.A GCTCGCCCAC GTTCTAATAT 1349
Consensus CgCAgCAgCC TTAGCTACGC CGGAGCttgA GCTCGcCacC GttcTAaTAT
ToLCV-sat 592 ...GGCAGCC TTAGCTCCGC CGTAGCTTAA CCTCTTCGGA GCTCTAATAT 681
A
CG
AACGACA TCCATGGTGG TTCTGGTGGT ACGGGCACCT GTA
stem
Block D
CLCDβ02-Pak 1349 TACCGTGGGC GAGCGGTGTC CGATGGTTTC TTGGTGGGTC CCATTGCTGG 49
AYVDβ01-Sin 1347 TACCGGTGGC GAGCGGGGAA TTTGG.GTTC CAGGTGGGAC CCACAC.... 44
AYLCDβ01-Pak 1351 TACCGGTGGC GAGCGATTTT TTTGGTGTTG CAGGTGGGAC CCACTTT... 46
TobLCDβ01-Pak 1356 TACCGGTGGC GAGCGAAATT TTTTGCGTCG TT.GTGGGTC CCACCT.... 44
TomLCDβ02-Pak 1374 TACCGTGGGC GAGCGGTGAC CGGTGGCTTC TTGGTGGGTC CCACGTT... 46
OLCDβ01-Egy 1350 TACCGTGGGC GAGCGGTGCC TGGACTGTAT TTGGTGGGTC CCAGCTACTG 49
Consensus TACCGgtgGC GAGcggggtt tttggggtcc tagGTGGgtC CCaccttcta
ToLCV-sat 682 TACCAGGTGC CCGTTGCTTT TCTGGGCCTT TAGGCGCCGC AGGCCAATTG 49

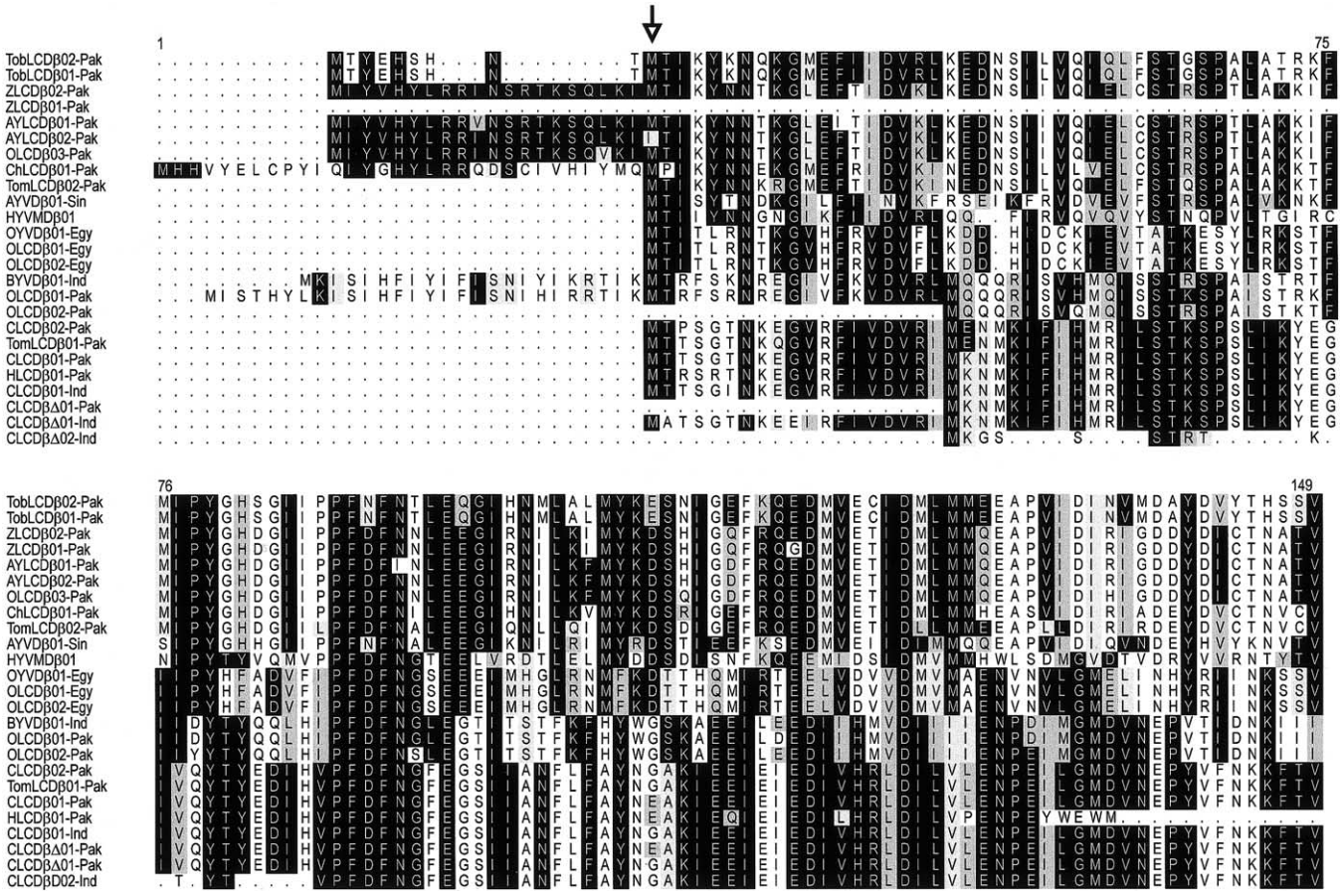


Fig. 4. Alignment of the amino acid sequences of the predicted products of gene C1 for DNA β molecules. Gaps (.) were introduced to optimize the alignment. Amino acids are highlighted, within a column, in black to indicate identity and shaded gray to indicate functional similarity. The conserved possible alternative initiation methionine codon is indicated by an arrow.

imately 118-amino-acid product for these molecules. The cotton-isolated DNA β clone CLCD$\beta\Delta$01-Pak encodes an N-terminally truncated form of the putative C1 protein (99 amino acids). This clone was infectious to cotton in the presence of CLCuMV, albeit with reduced efficiency (results not shown), producing typical CLCuD symptoms, and suggests that the N-terminal sequences of the C1 protein are dispensable for its function. A somewhat larger N-terminal truncation (due to a frameshift) in clone CLCD$\beta\Delta$02-Ind (truncating the predicted C1 protein to 77 amino acids) appears not to be tolerated. This clone is not infectious to cotton in the presence of CLCuMV. In contrast, clone HLCDβ01-Pak (isolated from hibiscus) has a C-terminal truncation of the C1 protein. This clone was also not infectious to cotton in the presence of CLCuMV, possibly indicating that these sequence are essential for C1 function and even small deletions here are not tolerated.

A pairwise comparison of the percentage similarity of the amino acid sequences of the putative gene C1 products is shown in Table 3. Despite the low levels of amino acid sequence conservation between the diverse DNA β molecules (between 35 and 80% sequence similarity for unrelated molecules) the alignment highlights clearly conserved amino acid sequence blocks in the gene C1 product. Database searches revealed no sequence relationship for any of the gene C1 products and indicated no presence of any conserved sequence motifs, which might provide some indication of the function, if any, of this putative protein.

Analysis of deletion mutants

The extent of sequence deletions for mutant DNA β molecules are detailed in Fig. 5. For all DNA β deletion mutants the SCR and A-rich region were maintained. The majority of deletions encountered were downstream of the hairpin structure, encompassing all or part of the gene C1 coding region. The nature of these deletions, being distal to the cloning site (or region where the primers used to amplify them are situated), suggests that they represent naturally occurring mutants rather than being PCR-derived mutations or cloning artifacts. The deleted molecules remained capable of *trans*-replication by either CLCuMV or AYVV (Table 2), indicating that an intact C1 gene is not required for this process. With the exception of clone CLCD$\beta\Delta$01-Pak, the deletion mutants did not induce the DNA β associated symptoms in cotton or *A. conyzoides*, suggesting that an

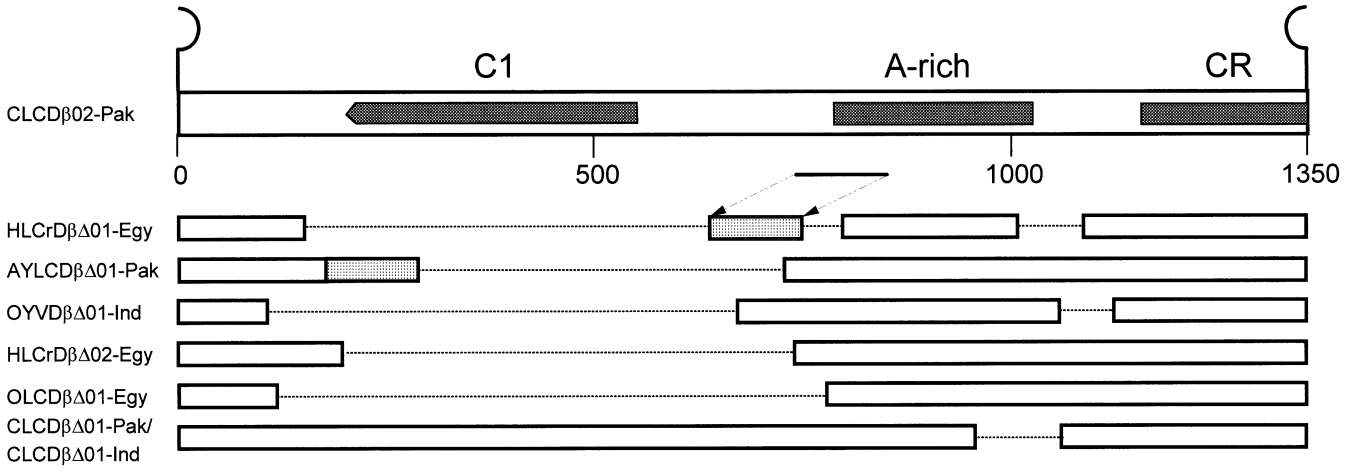


Fig. 5. Graphical representation of the deleted sequences of the less than full-length DNA β molecules. Deleted sequences are indicated with a dotted line. The deleted molecules are compared to the full-length molecule CLCDβ02-Pak. The shaded box for HLCrDβΔ01-Egy is a partially repeated sequence originating from the area indicated by the line above. The shaded box for AYLCDβΔ01-Pak indicates an inverted region.

intact C1 gene is required for this process. Clone CLCDβΔ01-Pak has a deletion of approximately 100 nts at the 3′-end of the A-rich region. This clone has low infectivity to cotton in the presence of CLCuMV, but induces typical CLCuD, indicating that some changes in this area of the genome can be tolerated.

Recombinant DNA β molecules

Analysis of DNA β sequences, using the Recombination Detection Program (RDP; Martin and Rybicki, 2000), detected a large number of potential recombination events between the molecules examined. These recombinations were predicted, in all cases, to have occurred between geographically related DNA β molecules. The most prominent example of recombination is provided by TomLCDβ01-Pak. This molecule shows the highest levels of similarity to the cotton DNA β molecules (92% nucleotide sequence identity to CLCDβ02-Pak) and groups with these in phylogenetic analyses (as discussed below). However, the SCR of TomLCDβ01-Pak shows higher levels of similarity to the second tomato isolated DNA β (TomLCDβ02-Pak; 90% nucleotide sequence identity) than it does to the cotton isolated DNA β molecules (<77%). The nucleotide sequence of the C1 gene of TomLCDβ01-Pak shows 99% identity to that of CLCDβ02-Pak but only 49% to that of TomLCDβ02-Pak. Similarly, the A-rich region of this satellite has 98% identity with that of CLCDβ02-Pak but only 52% similarity to that of TomLCDβ02-Pak. Thus TomLCDβ01-Pak is likely a recombinant DNA β molecule with an SCR derived from a tomato DNA β and the remainder of the molecule originating from a cotton-like DNA β.

Recombination is also apparent for okra-isolated clone OLCDβ03-Pak. This molecule has a C1 gene and A-rich region most similar to an *Ageratum* DNA β originating from Pakistan (97% nucleotide sequence identity to clone AYLCDβ01-Pak). The SCR of this molecule has only 85% similarity to AYLCβ01-Pak and 70% similarity to other DNA β molecules originating from okra. The highest levels of similarity of the SCR of this molecule are to a molecule originating from *Ageratum* from India (95% identity to AYVβΔ01-Ind).

Infectivity of DNA β molecules

The infectivity of DNA β molecules was investigated by coinoculation with an infectious clone of the helper virus AYVV to *N. benthamiana* (Table 2). All the DNA β molecules that were tested, including those with deletions, were capable of *trans*-replication and maintenance in *N. benthamiana* by AYVV, as judged by Southern blot hybridization (results not shown). The DNA β components fall into two classes based on the symptoms induced in *N. benthamiana* in the presence of AYVV. The first group consists of all molecules with a size below 1200 nts as well as ZLCDβ02-Pak, OLCDβ01-Pak, ChLCDβ01-Pak, CLCDβΔ01-Ind, CLCDβΔ02-Ind, and HLCDβ01-Pak, which had no apparent effects on the symptoms induced by AYVV in *N. benthamiana*. The symptoms in this case were upward leaf curling, vein swelling, and the formation of small enations on the veins on the undersides of leaves. These symptoms were indistinguishable from the symptoms induced by AYVV in the absence of AYVD DNA β in this host. All other DNA β molecules, in the presence of AYVV, induced symptoms typical of AYVV with its own DNA β in *N. benthamiana*, consisting of a downward leaf curl and mottling. The symptom phenotype correlated with the presence of an intact C1 gene (Table 2; as discussed above). However, some DNA β molecules with apparently intact C1 genes were not capable of inducing the change in symptoms (ZLCDβ02-Pak, OLCDβ01-Pak, and ChLCDβ01-Pak). The reason for this is unclear but may indicate that there are

other mutations within the clones or that AYVV is not capable of interacting with all DNA β molecules to effect the change of symptom phenotype.

Several additional inoculations were conducted to investigate the infectivity of the DNA β clones. Clone AYLCDβ01-Pak, isolated from *A. conyzoides* originating from Pakistan, was co-inoculated with AYVV to *A. conyzoides*. The foliar symptoms induced were typical of AYVD, consisting of a deep yellow chlorosis along the veins, but were milder than either those induced by AYVV with its own DNA β or of the field isolate from which clone AYLCDβ01-Pak was obtained. As well as the yellow veinal chlorosis, the field isolate from which clone AYLCDβ01-Pak was obtained exhibited a severe downward leaf curling (Fig. 1C) that was not evident in *A. conyzoides* following co-infection with AYVV.

A number of DNA β molecules were also co-inoculated with CLCuMV to cotton. In addition to CLCDβ01-Pak and CLCDβ02-Pak (for which infectivity has been demonstrated previously; Briddon et al., 2001), clones CLCDβΔ01-Pak, TomLCDβ01-Pak, CLCDβ01-Ind, and OYVDβ01-Egy were infectious to cotton, in the presence of CLCuMV, and induced typical CLCuD symptoms, including downward leaf curling, vein swelling, vein darkening, and the formation of small enations on the veins on the underside of leaves.

Phylogenetic comparisons

The putative phylogenetic relationships of DNA β molecules deduced from alignments of their complete nucleotide sequences, the predicted amino acid sequences of their C1 genes (24 sequences analyzed), as well as the nucleotide sequences of their conserved regions and A-rich regions are shown in Fig. 6. The dedrograms based on the complete nucleotide sequences revealed two major groups of DNA β molecules. The first consisted of molecules isolated from species within the family *Malvaceae* as well as a single molecule isolated from tomato (TomLCDβ01-Pak). Within the *Malvaceae* group, three clusters of molecules were evident: (1) DNA β molecules isolated from cotton and *Hibiscus*, as well as the single molecule isolated from tomato, all originating from the Indian sub-continent (denoted as Mal^A); (2) molecules isolated from okra originating from the Indian subcontinent (Mal^B); (3) molecules isolated from okra and hollyhock originating from Egypt (Mal^C).

The second major group of DNA β molecules (denoted as "Diverse") consisted of molecules originating from a diverse range of host plants. The molecules isolated from *Ageratum* and *Zinnia* (both in the family *Compositae*) originating from the Indian subcontinent formed a single cluster together with a single molecule isolated from okra originating from Pakistan (OLCDβ03-Pak). Although the DNA β molecules originating from Pakistan isolated from chilies and tomato cosegregated in this analysis, they showed only a distant relationship to the DNA β molecules isolated from a third solanaceous host, tobacco. Instead, the tobacco isolated molecules showed a closer relationship to the satellite molecule isolated from *A. conyzoides*, originating from Singapore (Saunders et al., 2000). The single molecule isolated from honeysuckle (HYVMDβ01) grouped with and was basal to the Diverse group, whereas the ToLCV-sat (originating from Australia) showed no particular affinity for either of the two major groups of DNA β molecules and was basal to all the DNA β molecules.

The dendrograms derived from alignments of the predicted amino acid sequences of the C1 protein (Fig. 6B) and the complete nucleotide sequences (Fig. 6A) were generally congruent, predicting two distinct groups with generally similar evolutionary relationships. The major difference between the two trees was in the relationship of the DNA β molecules originating from Egypt (Mal^C). Although the Mal^C DNA β molecules segregated with the other *Malvaceae* DNA β molecules based on nucleotide sequences, they were predicted to be more distantly related to the other DNA β molecules based on the C1 protein dendrogram, being basal to the Diverse group. This relationship of the Mal^C sequences was supported by a high bootstrap confidence value.

A dendrogram deduced for an alignment of the nucleotide sequence of the SCR of all the DNA β molecules (Fig. 6C) gave similar groupings to those predicted for the complete nucleotide sequences. In this tree the tomato isolated DNA β molecule TomLCDβ01-Pak, which grouped with the Mal^A molecules based on complete nucleotide sequences and the amino acid sequences of the C1 protein, segregated with and was basal to the other *Solanaceae* DNA β molecules within the Diverse group. This suggests that TomLCDβ01-Pak is a recombinant molecule, as concluded above.

Dendrograms derived from alignments of the nucleotide sequences of the A-rich region of DNA β molecules gave groupings identical to that obtained for the SCR (Fig. 6D). In this analysis the tomato isolated DNA β, TomLCDβ01-Pak, grouped with the Mal^A molecules and represents the only molecule for which the segregation based on the A-rich region does not correspond to that based on the SCR.

Within both the Malvaceae and Diverse groups, DNA β molecules clustered according to the host from which they were isolated as well as by recent geographic origin for molecules originating from the Indian subcontinent and Egypt.

Discussion

The first symptom modulating satellite molecules for DNA viruses were recently shown to be associated with the monopartite begomoviruses AYVV and CLCuMV (Saunders et al., 2000; Briddon et al., 2001). These two full-length DNA β molecules are just below half the length of their helper begomoviruses (49.2% for the AYVD-associated

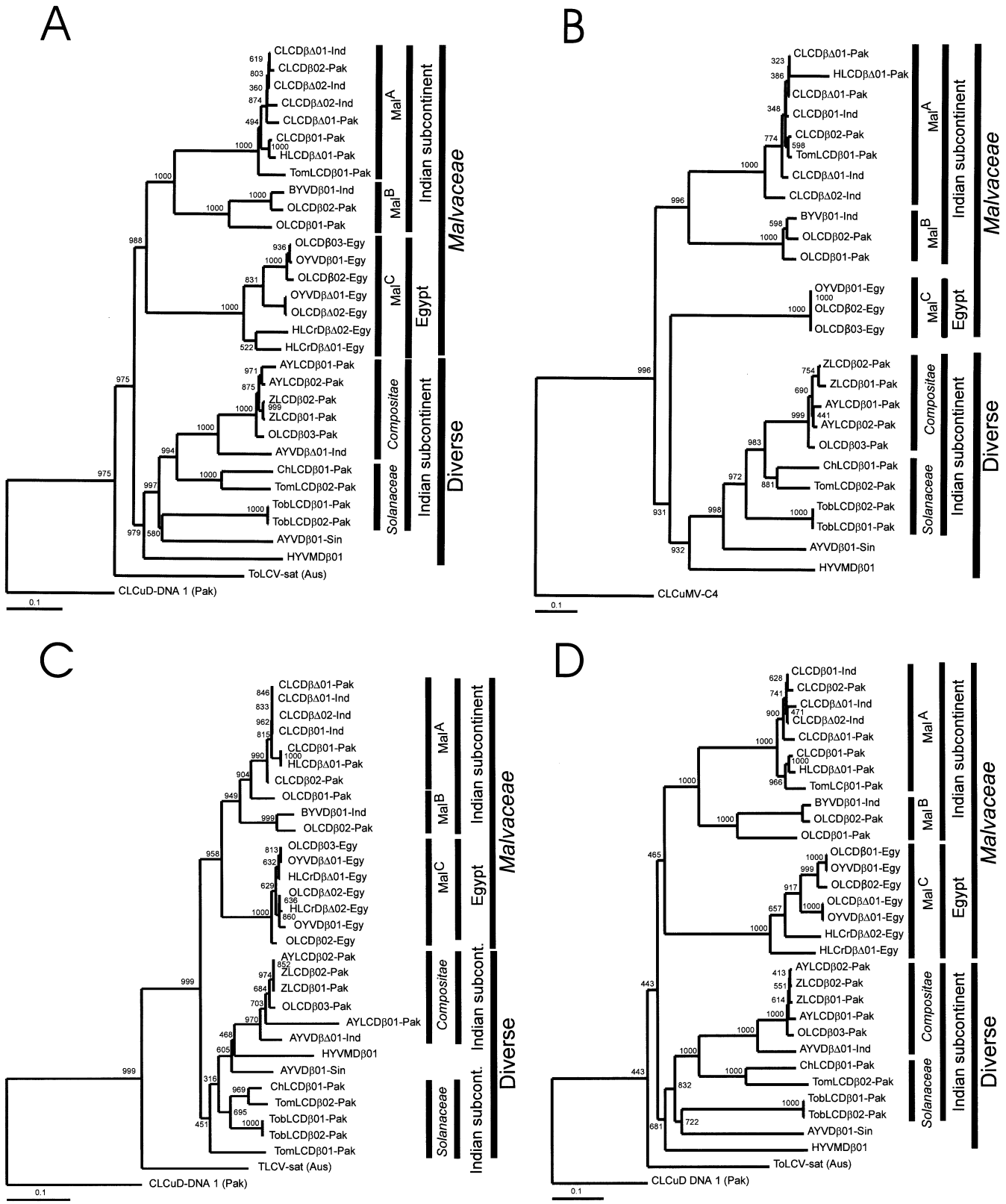


Fig. 6. Neighbor-joining phylogenetic dendrograms based on alignments of the complete nucleotide sequences of DNA β components (A), predicted amino acid sequences of the putative C1 genes (B), the nucleotide sequences of the SCR (C), and A-rich regions (D). Vertical distances are arbitrary; horizontal distances are proportional to calculated mutation distances. Numbers at nodes indicate percentage bootstrap scores (1000 replicates). The trees were rooted on the sequence of CLCuD DNA 1 (a molecule distantly related to DNA β but of a similar size), the predicted amino acid sequence of the C4 protein of CLCuMV (an unrelated protein of similar size), the positionally equivalent sequence of CLCuD DNA 1 (surrounding and immediately upstream of the conserved hairpin structure), and the A-rich region of CLCuD DNA 1 respectively.

molecule originating from Singapore and 49.5% for the CLCuD-associated molecules originating from Pakistan). Our analysis of 26 further satellites has identified biologically active DNA β molecules ranging in size from 1247 (CLCuDβΔ01-Pak) to 1374 nts (TomLCDβ02-Pak) in length. The sequences of these presumed full-length DNA β molecules have three conserved features: an A-rich region, a conserved region (SCR), and a single open reading frame (the putative coding region of gene C1).

The SCR has at its 3′ extremity a putative stem-loop structure with a highly conserved stem sequence and the NS TAATATTAC (typical of the geminiviruses) or, for a single molecule (OLCDβ02-Pak), TAGTATTAC (typical of the nanoviruses). For both the geminiviruses and the nanoviruses the hairpin structure acts as the origin of virion-strand replication (Heyraud-Nitschke et al., 1995; Hafner et al., 1997). To initiate replication the NS is nicked by the virus-encoded replication-associated protein (Rep). Since clone OLCDβ02-Pak is competent for *trans*-replication by AYVV, it is probable that the NS motif TAGTATTAC is recognized by a geminivirus Rep. However, the mechanism of interaction between begomovirus Rep and DNA β remains unclear and the NS of DNA β molecules has yet to be shown to function as an origin.

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 NS (reviewed by Hanley-Bowdoin et al., 1999). The two components of the bipartite begomoviruses have a region of high sequence similarity [the "common region" (CR)], encompassing the origin of replication, that allows the replication of both components by the Rep encoded on one of the components. Rep recognition of the iterons is extremely sequence-specific (Fontes et al., 1994) and is usually confined to a particular begomovirus species, although interspecific replication can occur if the iterons are conserved (Saunders et al., 2002). DNA β molecules do not contain the conserved iterons of their helper begomoviruses, at least for those molecules for which a sequence of the corresponding helper begomovirus is available (Briddon et al., 2001; Saunders et al., 2000), suggesting that DNA β satellites contain alternative motif(s) for Rep binding.

For DNA β molecules the SCR is positionally analogous to the CR of bipartite begomoviruses (just upstream of the NS containing hairpin). This would seem the most likely place to find cryptic Rep binding sites for *trans* replication of DNA β molecules by begomoviruses. Both theoretical and experimental studies of the specificity of Rep binding have shown that iterons have a 5-bp core sequence ($GGN_1N_2N_3$) and a variable number of additional nucleotides which are species (sometimes isolate) specific (Argüello-Astorga and Ruiz-Medrano, 2001; Chatterji et al., 1999, 2000). The SCR contains a number of highly conserved GG dinucleotides (Fig. 3). It is possible that one or more of these GG dinucleotides, in conjunction with the highly conserved sequence blocks (within the SCR), could be the basis of the DNA β Rep binding motif, a motif that breaks the usual sequence specificity for Rep/iteron interaction and acts, possibly, as a generic recognition sequence for Reps.

The A-rich region, just upstream of the SCR, is a feature of all the DNA β molecules examined and a feature in common with the DNA 1 molecules (Mansoor et al., 1999; Saunders and Stanley, 1999). For DNA 1 it has been suggested that the A-rich sequences may be a "stuffer" (a region of arbitrary sequence) required to increase the size of the molecule from that of a nanovirus component (~1000 nts) to that required for encapsidation by a begomovirus coat protein (half the size of a begomovirus DNA component, ~1400 nts; Mansoor et al., 1999; Saunders and Stanley, 1999). For all the DNA β molecules that contained deletions, both the A-rich region and the SCR were maintained. Rather the C1 coding sequence was partially or wholly deleted in these molecules. This is strong circumstantial evidence to suggest that both the A-rich region and the SCR are required for, or at least play a part in, maintenance of DNA β by a begomovirus. Possibly the A-rich region has a function in complementary-strand DNA replication.

Only a single conserved (in both sequence and position) ORF was identified for the presumed full-length DNA β molecules. The correlation between an ability to induce the change in symptoms in *N. benthamiana* (when co-infected with AYVV) and an intact C1 gene is strong circumstantial evidence suggesting that this gene plays an important role in the function of this satellite and is involved in symptom modulation. Conversely it also demonstrates that this gene is not required for *trans* replication by the helper begomovirus. This does not rule out the possibility that other coding sequences, possibly being host or virus adapted, are present on DNA β. Mutational analysis will be required to identify the full complement of genes encoded by these molecules.

The phylogenetic comparisons show DNA β molecules to group according to the host from which they were isolated (*Malvaceae* and non-*Malvaceae*). This most likely represents the host adaptation of the associated helper begomovirus and coadaptation of its DNA β rather than host adaptation of the satellite. This assumption is supported by OLCβ03-Pak (a nonmalvaceous satellite isolated from okra) and TomLCDβ01-Pak (a recombinant molecule, with a C1 gene originating from a malvaceous DNA β, isolated from a solanaceous host). Also we have shown that CLCuD can be experimentally transmitted to tomato and clones of CLCuMV and CLCuD DNA β are infectious to tobacco (R.W. Briddon, unpublished results). However, determination of the precise contribution made by each of the components to the host range of the complex will have to await further studies.

The ToLCV-sat identified in Australia (Dry et al., 1997) is clearly related to the DNA β satellites. It maintains the A-rich region (51% A content, ~157 nts in length) but lacks the coding region for the C1 gene. In this respect

the ToLCV-sat is very similar to the small defective DNA β molecules isolated as part of this study, such as HLCrDβΔ01-Egy. However, the region of ToLCV-sat equivalent to the SCR of DNA β molecules is very different from that of DNA βs. It has some residual SCR sequences but an entirely different hairpin structure and adjacent sequences. This may be an indication that there is second class of DNA β molecules that remain to be identified. Alternatively ToLCV-sat may have gained a distinct origin of replication as a result of recombination with another virus or viral component. However, the donor is unlikely to have been ToLCV as this virus shares no significant sequences with the ToLCV-sat (Dry et al., 1997).

The evolution of DNA β satellites has clearly followed two paths leading to the two major clusters (*Malvaceae* and non-*Malvaceae*) of molecules we see in the present. The North African DNA β satellites (Mal^{C}) appear to have arisen by a recombination event between members of these two clusters resulting in malvaceous satellites with a C1 gene from a nonmalvaceous satellite. In view of the limited numbers of molecules originating from southeast Asia and the Far East that were available for this study, it is impossible to draw any firm conclusions about the evolutionary history and genetic complexity of DNA β in these areas. Since the two DNA β molecules originating from these regions (AYVDβ01-Sin and HYVMDβ01 respectively) were distinct in this analysis, it seems likely that additional geographically and genetically distinct clusters of these molecules remain to be identified.

Our analysis has identified only DNA β 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. aureoreticulata); a perennial climbing vine which is grown as an ornamental. 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 (R.W. Briddon, unpublished results) and is thus not a native of the New World. It is also interesting to note that thus far DNA β has been found associated only with monopartite begomoviruses and no native monopartite begomoviruses have been identified in the New World, presumably because they evolved after the separation of the Americas from the Gondwana landmass (Rybicki, 1994). However, the global trade in agricultural products has led to the introduction of the monopartite begomovirus *Tomato yellow leaf curl virus* (which originated in the Middle East) into southern Europe, the Carribean, Central America, and the southeastern USA (McGlashan et al., 1994; Polston et al., 1999). These findings provide some evidence to suggest that the begomovirus–DNA β interaction evolved after the divergence of the Old World and New World begomoviruses; this divergence is estimated to have occurred some 130 million years ago (Rybicki, 1994).

Efforts are now underway to better understand the nature of DNA β molecules in two major areas. The first is to determine the nature of the interaction of begomovirus-encoded Rep with DNA β, a molecule which lacks the conventional Rep binding sequences. This will provide information on the specificity, or lack thereof, in the adaptation between begomoviruses and their satellites. The second area seeks to elucidate the mechanism of action of DNA β in its interaction(s) with host plants. Work here is centred on understanding the function and mechanism of action of the C1 gene product in the infection process.

Methods and materials

Origins of virus infected plant materials

The geographic origins of plant materials infected with begomovirus–DNA β complexes are summarized in Table 1. For the vast majority of samples total nucleic acids were extracted from field collected leaf samples. For a small number of samples [all of the *A. conyzoides* isolates (originating from India and Pakistan), *Hibiscus rosa-sinensis* originating from Pakistan, hollyhock (*Althea rosea*) originating from Egypt and the cotton isolates (originating from India and Pakistan)] live plants were maintained by vegetative propagation (*Ageratum* and *Hibiscus*) or graft transmission (cotton originating from Pakistan) in an insect-proof glasshouse maintained at 25°C with a 16-h lighting regime. The origins of four other DNA β molecules that were not produced as part of this study but are included in the analysis are detailed in Table 1; two clones isolated from cotton originating from Pakistan (CLCDβ01-Pak and CLCDβ02-Pak; Briddon et al., 2001), a clone isolated from okra (BYVDβ01-Ind, Jose and Usha, unpublished) and a clone isolated from *A. conyzoides* originating from Singapore (AYVDβ01-Sin; Saunders et al., 2000).

PCR-mediated amplification and cloning of DNA β molecules

Total nucleic acids were extracted from all leaf samples by the CTAB method, with the exception of some of the samples originating 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 have been described previously (Briddon et al., 2002). These primers were designed to allow the amplification of the entire DNA β molecule and have been shown previously to produce products which are, in many cases, infectious to plants (Briddon et al., 2001). The primers introduce a *Kpnl* restriction endonuclease site

for recovery of the full-length insert. Clones OLCDβ02-Egy, OLCDβ03-Egy, OLCDβΔ02-Egy, and HLCrDβΔ02-Egy were produced with truncated versions of primers β01/β02 (virion strand primer 5′-GGTACCACTACG-CTACGCAG-3′; complementary strand primer 5′-GGT-ACCTACCCTCCCAGGGG-3′).

For some isolates (HYVMD and OYVD-Egypt), amplifications with primer pair beta01/beta02 produced products of less than 1300 nts. 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 OYVDβ01-Egy (virion-sense primer 5′-GGTACCCACCGGGAAATGATAGTTTCACCG-3′, complementary-sense primer 5′-GGTACCGGCCTCAAT-TCACGACACGCGCGCGG-3′) and HYVMDβ01 (virion-sense primer 5′-TCTAGATAAGTTTTGGTTTTATTG-CACTTTG-3′, complementary-sense primer 5′-TCTAGA-AGGTGGGGCCCACCTTTCCCCCG-3′) were produced using specific primers.

PCR amplifications were conducted as described by Briddon et al. (2002) with annealing temperatures ranging from 45° to 50°C. All PCR products were cloned into either pGem-T or pGem-T Easy vectors (Promega).

Determination of infectivity of DNA β clones

The infectivity of cloned DNA β molecules to plants was determined by co-inoculation with their respective begomovirus (for CLCuD and related DNA β molecules as detailed in Table 1) or by co-inoculation with *Ageratum* yellow vein virus (AYVV; Tan et al., 1995). The use of AYVV to investigate infectivity was based on the observation that this virus is capable of the *trans*-replication of unrelated DNA β satellites in *N. benthamiana* (Briddon et al., manuscript in preparation). The symptoms of an AYVV infection of *N. benthamiana* differ depending on whether AYVD DNA β is present (Saunders et al., 2000); from upward leaf curl in the absence of AYVD DNA β to downward leaf curl in the presence of AYVD DNA β. A similar change in symptom phenotype was noted when AYVV was inoculated with CLCuD DNA β (Briddon et al., manuscript in preparation). Thus a productive interaction was indicated by the change of symptoms of an AYVV infection of *N. benthamiana* in the presence of a DNA β.

Inoculations were conducted by biolistic inoculation with excised clones coated on to 0.95 or 1.3 μM gold particles and a hand-held biolistics gun, as described previously (Briddon et al., 1998, 2001). Inoculated *N. benthamiana* and *A. conyzoides* plants were maintained in insect-proof glasshouses at 25°C with a 16-h lighting regime. Inoculated cotton plants were maintained in a growth cabinet at a constant temperature of 35°C with a 16-h lighting regime. Plants were monitored daily for the appearance of symptoms.

Sequence determination and sequence analysis

The complete nucleotide sequences of 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 3700 automated sequencer. Sequence information was stored, assembled, and analyzed using version 7 of the program library of the Genetics Computer Group (Devereaux, Haeberli, and Smithies, 1984).

Phylogenetic analyses were conducted on matrices of aligned sequences using the neighbor-joining and bootstrap options of Phylip version 3.5c (J. Felsenstein, Department of Genetics, University of Seattle, WA, USA) running on an IBM compatible personal computer. Sequence alignments were produced using CLUSTAL W (Thompson, Higgins, and Gibson, 1994) running on a Unix mainframe computer. Phylogenetic dendrograms were viewed, manipulated, and printed using Treeview (Page, 1996). Potential recombination events between DNA β molecules were identified using RDP (Martin and Rybicki, 2000).

Acknowledgments

The authors thank J. Stanley for providing the AYVV clone. The first author is particularly grateful to Darren Martin for taking the time to explain the intricacies of RDP. This work was supported by the Common Fund for Commodities (overseen by the International Cotton Advisory Committee under Project No. ICAC/07) and the Department for Environment, Food and Rural Affairs (DEFRA) under Project No. PH01 47T. The CFC/ICAC project was a collaboration between the University of Arizona (Tucson, AZ, USA), the National Institute of Biotechnology and Genetic Engineering (Faisalabad, Pakistan), the Cotton Research Institute (Faisalabad, Pakistan), and the John Innes Centre (JIC, Norwich, UK). The University of Arizona laboratory also recognizes the invaluable support from Cotton Incorporated, USA. The authors at the JIC acknowledge the support of the Biotechnology and Biological Sciences Research Council. In the UK, viruses were held and manipulated with the authorization of DEFRA under Licence No. PHL 11D/3825(7/2001). In the USA, viral DNA extracts were received and manipulated under USDA-APHIS Permit 34631.

References

Argüello-Astorga, G.R., Ruiz-Medrano, R., 2001. An iteron-related domain is associated to motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. Arch. Virol. 146, 1465–1485.

Briddon, R.W., Bull, S.E., Mansoor, S., Amin, I., Markham, P.G., 2002. Universal primers for the PCR-mediated amplification of DNA β: a

molecule associated with some monopartite begomoviruses. Mol. Biotechnol. 20, 315–318.

Briddon, R.W., Markham, P.G., 2000. Cotton leaf curl virus disease. Virus Res. 71, 151–159.

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

Briddon, R.W., Liu, S., Pinner, M.S., Markham, P.G., 1998. Infectivity of African cassava mosaic virus clones to cassava by biolistic inoculation. Arch. Virol. 143, 2487–2492.

Chatterji, A., Padidam, M., Beachy, R.N., Fauquet, C.M., 1999. Identification of replication specificity determinants in two strains of tomato leaf curl virus from New Delhi. J. Virol. 73, 5481–5489.

Chatterji, A., Chatterji, U., Beachy, R.N., Fauquet, C.M., 2000. Sequence parameters that determine specificity of binding of the replication-associated protein to its cognate site in two strains of Tomato leaf curl virus-New Delhi. Virology 273, 341–350.

Collmer, C.W., Howell, S.H., 1992. Role of satellite RNA in the expression of symptoms caused by plant viruses. Ann. Rev. Phytopathol. 30, 419–442.

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

Dry, I., Krake, L.R., Rigden, J.E., Rezaian, M.A., 1997. A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. Proc. Natl. Acad. Sci., USA 94, 7088–7093.

Fontes, E.P.B., Eagle, P.A., Sipe, P.S., Luckow, V.A., Hanley-Bowdoin, L., 1994. Interaction between a geminivirus replication protein and origin DNA is essential for viral repliation. J. Biol. Chem. 269, 8459–8465.

Frischmuth, T., Zimmat, G., Jeske, H., 1991. The nucleotide sequence of abutilon mosaic virus reveals prokaryotic as well as eukaryotic features. Virology 185, 596–604.

Hafner, G.J., Stafford, M.R., Wolter, L.C., Harding, R.M., Dale, J.L., 1997. Nicking and joining activity of banana bunchy top virus replication protein in vitro. J. Gen. Virol. 78, 1795–1799.

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

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

Höfer, P., Engel, M., Jeske, H., Frischmuth, T., 1997. Nucleotide sequence of a new bipartite geminivirus isolated from the common weed *Sida rhombifolia* in Costa Rica. J. Gen. Virol. 78, 1785–1790.

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

Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G.P., Crespi, S., Gronenborn, B., 1992. Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. Nucleic Acids Res. 19, 6763–6769.

Klinkenberg, F.A., Stanley, J., 1990. Encapsidation and spread of African cassava mosaic virus DNA A in the absence of DNA B when agroinoculated to *Nicotiana benthamiana*. J. Gen. Virol. 71, 1409–1412.

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

Martin, D.P., Rybicki, E.P., 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16, 562–563.

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

Murant, A.F., Mayo, M.A., 1982. Satellites of plant viruses. Ann. Rev. Phytopathol. 20, 49–70.

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

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

Page, R.D.M., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358.

Polston, J.E., McGovern, R.J., 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.

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

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

Rybicki, E.P., Briddon, R.W., Brown, J.E., Fauquet, C.M., Maxwell, D.P., Harrison, B.D., Markham, P.G., Bisaro, D.M., Robinson, D., Stanley, J., 2000. Geminiviridae, in: van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wicker, R.B. (Eds.), Virus Taxonomy: Seventh Report on the International Committee on Taxonomy of Viruses, Academic Press, San Diego, pp. 285–297.

Saunders, K., Bedford, I.D., Briddon, R.W., Markham, P.G., Wong, S.M., Stanley, J., 2000. A novel virus complex causes *Ageratum* yellow vein disease. Proc. Natl. Acad. Sci. USA 97, 6890–6895.

Saunders, K., Salim, N., Mali, V.R., Malathi, V.G., Briddon, R.W., Markham, P.G., Stanley, J., 2002. Characterisation of *Sri Lankan cassava mosaic virus* and *Indian cassava mosaic virus*: evidence for acquisition of a DNA B component by a monopartite begomovirus. Virology 293, 63–74.

Saunders, K., 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 disease plant DNA viruses. Virology 264, 142–152.

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

Tan, P.H.N., Wong, S.M., Wu, M., Bedford, I.D., Saunders, K., Stanley, J., 1995. Genome organization of Ageratum yellow vein virus, a monopartite whitefly-transmitted geminivirus isolated from a common weed. J. Gen. Virol. 76, 2915–2922.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

von Arnim, A., Stanley, J., 1992. Determinants of tomato golden mosaic virus symptom development located on DNA B. Virology 186, 286–293.

Zhou, X., Liu, Y., Robinson, D.J., Harrison, B.D., 1998. Four variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. J. Gen. Virol. 79, 915–923.