European Journal of Plant Pathology **107:** 701–711, 2001.

Characterization of a new begomovirus from Egypt infecting hollyhock (*Althea rosea*)*

L. Bigarré[1,**], M. Chazly[2], M. Salah[2], M. Ibrahim[2], M. Padidam[3], M. Nicole[4], M. Peterschmitt[1], C. Fauquet[3] and J.C. Thouvenel[1,***]
[1]*LPRC, CIRAD-IRD, BP 5035, 34032 Montpellier, Cedex 1, France;* [2]*IRD Laboratory, College of Agriculture, BO 26 GIZA, Code 12211, Cairo, Egypt;* [3]*ILTAB, Danforth Center, University of Missouri Saint-Louis, MO 63121, USA;* [4]*GENETROP, IRD, BP 5045, 34032, Montpellier, Cedex 1, France;*
[**]*Present address: URV, Biotechnology, 3133 AT, Vlaardingen, The Netherlands;* [***]*Author for correspondence (Phone: +33 467615887; Fax: +33 467615505; E-mail: thouvene@mpl.ird.fr)*

Accepted 10 June 2001

Key words: enation, geminivirus, cotton leaf curl, okra leaf curl, hyperplasia

Abstract

A viral isolate from Egypt associated with symptoms of enations and leaf curling on hollyhock (*Althea rosea*) was characterized at the cytopathological and molecular levels. Microscopic observations showed that enations resulted from a reorganization of the vascular tissues, including activation of a cambial activity in the phloem, the development of a palisade parenchyma in place of a spongy one and the differentiation of minor vascular tissues. From this isolate, the full-length DNA-A of a begomovirus (family *Geminiviridae*) was cloned and sequenced. This genome exhibited a genetic organization similar to that of other old-world begomoviruses like *Tomato yellow leaf curl virus* from Israel and *Ageratum yellow vein virus* from Singapore. However, its sequence was significantly distinct (similarity $< 69\%$) from any other geminivirus. This begomovirus was thus considered as representative of a new viral species named *Althea rosea enation virus* (AREV). AREV was agroinfectious on *Nicotiana benthamiana*, on which it induced a severe leaf-curling and vein distortion, but could not re-establish infection on *A. rosea*. To determine if AREV was also associated with a similar disease affecting okra in Upper-Egypt, the partial sequence of the coat protein gene of an isolate was determined. It exhibited 90% nt identity with the hollyhock isolate (97% amino acid), suggesting a genetic heterogeneity in the begomovirus population associated with the enation diseases.

Introduction

Geminiviruses transmitted by the whitefly *Bemisia tabaci* (genus begomovirus) are the cause of important yield losses on many primary crops all over the world. Begomoviruses are all composed of at least one molecule of circular single-stranded DNA (DNA-A) of approximately 2.8 kb, encapsidated in a typical double-geminate particle and encoding the coat protein (*AV1* gene) as well as a replication-associated protein (*Rep*). Apart from a few species (Navot et al., 1991; Dry et al., 1993), begomoviruses possess a second circular DNA genomic component necessary for the systemic spread of the A component and thus indispensable to reproduce the disease. In most species, the second component, named DNA-B, has a similar size to DNA-A and shares a high nucleotide identity in the intergenic region while differing by the encoded products. Recently, another type of component of about 1.4 kb, named DNA-β, was found associated with the DNA-A of *Ageratum yellow vein virus* (AYVV) from Singapore. DNA-β shares no homology with any DNA-A or DNA-B, apart from a highly-conserved nonanucleotide sequence, and is indispensable for

*The GenBank accession numbers of the nucleotide sequences reported in this paper are AF014881 and AF155064

reproducing the disease in association with DNA-A (Saunders et al., 2000).

Begomoviruses are associated with several whitefly-transmitted diseases affecting malvaceae plants in the old world. Geminate particles were purified from okra leaf curl disease (*Abelmoschus esculentus*) in Ivory Coast (Fauquet and Thouvenel, 1987) while serological studies made on different isolates from cotton (*Gossypium* sp.) and okra in Africa confirmed the presence of geminiviruses in the symptomatic plants (Swanson and Harrison, 1993; Konate et al., 1995). In addition, a typical begomovirus DNA-A isolated from the cotton leaf curl disease in Sudan was recently cloned and sequenced (Idris et al., 2000). Although the infectivity of the molecule remains to be assessed, this finding confirmed the likely implication of begomoviruses in the disease in Africa. Finally, extensive surveys with DNA probes in Pakistan revealed that a complex of numerous begomovirus genotypes infected in a non-specific way a range of malvaceae plants such as cotton, okra, hollyhock (*Althea rosea*) as well as non-malvaceae plants.

Despite the accumulation of serological and molecular data, the epidemiological situation is still very confused. First, despite the availability of numerous DNA-A clones from Asia, the etiology of the various diseases cited above was never completely elucidated because of difficulties of verifying Koch's postulate and because of the important genetic heterogeneity between isolates. Second, the cotton leaf curl in Africa refers to isolates with various biological properties and associated with various combinations of symptoms on plants in the fields such as vein mosaic, leaf-shape enations, vein thickening on the main or the secondary network and stunting (Giha and Nour, 1969; Bink, 1975; reviewed in Brown, 1992). More data is greatly needed to elucidate the epidemiological situation of begomoviruses affecting malvaceae plants in Africa.

In Egypt, malvaceous plants are widely cultivated, including okra, several *Hibiscus* sp., and noticeably cotton which represents a major economic resource for the country. To date, no data have been collected on begomoviruses affecting malvaceae plants in the country although typical geminivirus-associated symptoms of leaf enations and curling were observed on different species, rarely on cotton, occasionally on okra and cheeseweed (*Malva parviflora*), and very frequently on hollyhock (*Althea rosea*). The symptoms from hollyhock are transmitted by whiteflies and are indeed very similar to those of the 'small vein thickening syndrome' characteristic of the cotton leaf curl previously reported for cotton and okra in Sudan (Bink, 1975). We thus hypothesized that a begomovirus population may infect hollyhock in Egypt. The sequence of a begomovirus DNA-A genome was determined and it was confirmed that hollyhock is infected by a new viral species that has been named *Althea rosea enation virus* (AREV).

Materials and methods

Source of virus isolates

Leaves of *A. rosea* displaying severe enations on the secondary vein network were collected in a garden in Giza. Transmission of the disease was achieved as follows: a population of *B. tabaci* originating from Cairo was given a 48 h acquisition access period on these leaves, and was transferred on *A. rosea* seedlings cv. 'grande double' (20 insects/plant) for 48 h. The viral isolate was provisionally called AREV. A second isolate was collected from okra leaves exhibiting enations in a field in Aswan (Upper-Egypt) and conserved as calcium chloride-desiccated leaves at room temperature for two months.

Microscopy

Tissue fragments with enations (about 2.0 mm^3) were fixed for 4 h in 2.5% glutaraldehyde buffered with 0.1 M cacodylate (pH 7.2), rinsed in the same buffer and postfixed for 1 h in 1% osmium tetroxide. Fixed tissues were then dehydrated in a graded series of ethanol followed by propylene oxide (one hour), and embedded in Epon 812 (TAAB, England). Semi-thin (1.5 μm) sections were mounted in phosphate-buffered glycerol and examined with a light microscope (Leitz Diaplan) after staining with toluidine blue.

Plant DNA extraction and cloning of AREV

Plants at the 5–10 leaf stage were used to extract total DNA using a CTAB method (Gawel and Jarret, 1991). Viral DNA was initially detected by PCR amplification with two degenerate primers specific for begomoviruses, MP16 and MP82, situated in the intergenic region and the coat protein respectively (Umaharan et al., 1998). This procedure gave a faint DNA fragment of the expected size (530 bp) that could be cloned, sequenced, and used to produce a specific probe for virus detection (referred hereafter as

probe MP16-82). To clone the full-length DNA, total plant DNA was further purified on a cesium chloride gradient. Gradient fractions of 350 µl, starting from the plant genomic DNA, were recovered and the total DNA extracted and submitted to electrophoresis and Southern-blot hybridization with probe MP16-82. Fractions enriched in viral DNA were revealed by the probe and subsequently tested with various endonucleases for production of linear double-stranded viral DNA. *Bam*HI was chosen and used for cloning total DNA into a pBluescript KS(-) vector (Stratagene) using *Escherichia coli* Epicurian XL1 blue competent cells (Stratagene). Colonies carrying viral DNA (pAREV) were selected by hybridization with probe MP16-82. In some experiments, a *Bgl*II/*Bam*HI 2.1 kbp fragment of AREV, excluding the intergenic region, was used as a DNA-A specific probe.

PCR detection of AREV and related isolates

The PCR product obtained from AREV-infected hollyhock with the primers MP16 and MP82 is faint. Moreover, it is not detectable using other malvaceae plant isolates (not shown). Thus, once the sequence of the virus was obtained, another set of more specific primers was designed, oAREV589 (5′-CCCAWKTACAGRAAGCCCAG-3′) and oAREV1166 (GGRTTDGARGCATGHGTACAG), to amplify a fragment of about 580 bp in the 3′ part of the coat protein. The PCR was performed in a 50 µl reaction mix containing 20 pmol of each primer, 1.5 mM of $MgCl_2$, 0.25 mM for each dNTP, and about 100 ng of plant DNA purified with a commercial kit (DNAeasy, Qiagen). Amplification consisted of 40 cycles of 94 °C for 1 min, 54 °C for 1 min and 72 °C for 40 s, followed by an extension cycle of 5 min at 72 °C. A 20 µl sample of the product was analyzed on an agarose gel.

Sequence analysis

One cloned viral DNA (pAREV) was sequenced on both strands with no ambiguities. The unique PCR product derived from the okra isolate was sequenced directly, on one strand, using the primer oAREV1166. Sequencing was performed with the *Taq* dye terminator cycle sequencing kit (Applied Biosystems) and analyzed on an Applied Biosystems 373A sequencer. All the sequences were aligned and the identities determined with the clustal method of the Megalign application of the DNASTAR package (Madison). Phylogenetic trees were produced using a cladistic parsimony approach (PAUP 3.1, Swofford, 1993, Illinois Natural History Survey, Champaign, IL) and a heuristic search with a bootstrap validation (500 replications). Sequence comparisons were made with the following geminiviruses (abbreviation; GenBank accession number): *Tomato yellow leaf curl-Israel* (TYLCV-IL; X15656), TYLCV-Egypt (TYLCV-EG; unpublished), TYLCV-Sardinia (TYLCV-Sar; X61153), TYLCV-Thailand (TYLCV-TH; M59839), *African cassava mosaic virus-Nigeria* (ACMV-NI; X17095), *ACMV-Kenya* (ACMV-KE; J02057), *East African cassava mosaic virus-Tanzania* (EACMV-TZ; Z83256), *Tomato leaf curl virus-New-Delhi* (TLCV-Nde2; U15016), TLCV-Bangalore (TLCV-Ban2; Z48182), TLCV-Australia (TLCV-AU; S53521), *Indian cassava mosaic virus* (ICMV, Z24758), *Ageratum yellow vein virus* (AYVV; X74516), *Cotton leaf curl virus-Pakistan-802a* (CLCuV-PK802a; AJ002455), CLCuV-PK806b (AJ002449), *Okra yellow vein mosaic virus-Pakistan* (OYVMV-PK201a; AJ002451), *Abutilon mosaic virus* (AbMV; X15983) and *Sida golden mosaic virus-Costa Rica* (SiGMV-CR; X99550).

Agroinoculation of AREV

To test the infectivity of the cloned viral DNA, a full-length dimer was inserted in the binary vector pBin-19 and this construction was introduced into two *Agrobacterium tumefaciens* strains, C58MP90 and C58GV2260, the latter being commonly used for genetic transformation of cotton. Recombinant bacteria were plated on Petri dishes at 28 °C for 48 h; colonies were picked with a needle and immediately inoculated in the stem apex. Plants were maintained at 25 °C with a 17 h light/7 h dark cycle in an insect-proof containment chamber. Agroinoculation was performed on hollyhock, *N. benthamiana* and four cultivars of cotton (*G. hirsutum* Stoneville 1324 cv. B75 and *G. barbadense* cvs. Giza 76, Pima S473 and Tadla 43).

Replication in N. tabacum *protoplasts*

For viral replication assays in protoplasts, a cell culture of *N. tabacum* BY2 was electroporated with a 1.8-mer construct of AREV cloned into pBluescript KS(-), as described elsewhere (Padidam et al., 1995). The partial dimer was produced by removing a *Sst*1-*Sty*I

fragment of pAREV containing one of the two *Bam*HI sites, self-ligating the plasmid and finally introducing a full-length genome in the other *Bam*HI site. About two million protoplasts were electroporated with 2 μg of the construct. Total cell DNA was extracted after 48 h as described elsewhere (Padidam et al., 1995) and 5–10 μg were analyzed by electrophoresis and Southern hybridization using a full-length AREV DNA probe labeled by random priming with [α-^{32}P]dCTP.

Results

AREV-associated histopathology

One to two months after whitefly transmission, *A. rosea* plants displayed a range of symptoms varying from mild symptoms, with thin enations but no leaf-curling and no effect on flowering, to severe symptoms with thickening of the veins (Figure 1A–C), pronounced leaf-curling, stunting, leaf proliferation and no flowering. In the most severe cases, the vein thickening developed to form a small cup-shaped enation. Symptoms were also observed on flowers, with light enations on sepals and some variegation on petals. Semi-thin leaf sections of three samples revealed that enations resulted from important internal tissular modifications. Typically, reorganization of leaf tissues consisted in replacement of the spongy parenchyma by a palisade parenchyma (compare Figure 1D and E). Abnormal cambial activity was observed in the main phloem parenchyma which underwent multiple cell divisions. Moreover, we observed in the new parenchyma the *de novo* formation of secondary vascular elements as indicated by the presence of xylem, i.e. lignified cells with secondary cell walls colored in light blue by toluidine blue, and phloem (Figure 1E–G). Another striking modification was the high accumulation of an intercellular material, that stained purple with toluidine blue and red with rhutenium red suggesting its pectic nature (Figure 1E–G). This material appears composed of fibrillar-like matrix closely associated with the middle lamela as observed by electron microscopy (not shown).

Genome analysis and infectivity of AREV

Viral DNA replicative forms as purified by cesium chloride were in too small quantities for detection on agarose gel by simple ethidium bromide staining. However, by using PCR with degenerated primers specific for begomoviruses, a 530 bp fragment was produced and used as a probe. This probe allowed the detection of viral replicative forms in some fractions of the gradient and the susceptibility of some of them to *Bam*H1 endonuclease, which allowed a cloning of these forms in a pBS-KS vector. By screening a plant DNA library with the specific probe, one clone named pAREV was isolated and completely sequenced on both strands (GenBank AF014881). Its sequence is 2755 nts long and exhibits a genetic organization similar to the DNA-A component of the begomoviruses. The DNA codes for at least six open reading frames (ORFs), typically separated by an intergenic region, and determined by homology with begomoviruses as AV1 (coat protein) and AV2 on the viral strand, and AC1 (Replication associated protein, Rep), AC2, AC3, and AC4 on the complementary strand (Figure 2A). Two additional ORFs, AC5 and AC6, have been detected on the complementary strand, but since they are not well conserved among the begomoviruses, their significance remains to be determined.

Southern-blot analysis of DNA extracted from hollyhock naturally infected with AREV revealed several single-stranded (ss)- and double-stranded (ds)-DNA replicative forms: an open ds circular form (oc), which is digested by *Bam*HI to give a linear form of about 2.8 kbp (lin) and two other forms migrating slightly faster and interpreted as ssDNA since they are digested by mungbean nuclease treatment (not shown) and not affected by *Bam*H1 treatment (lines 4 and 5). These forms appear with different DNA purification methods, either by using a commercial kit or by CsCl gradient (see Materials and methods). The presence of two ssDNAs for geminiviruses is not common but is occasionally found for other geminiviruses (Bendhamane, 1995; J. Stanley personal communication). Possibly, they represent circular (intact) and linear (nicked) forms of the genome. Supercoiled dsDNA (sc) is barely detectable in total plant extracts but is clearly detected in inoculated protoplasts (see below). In addition to these replicative forms, at least five low-molecular weight (<1.1 kbp) DNA species were detected, one of these appearing after *Bam*HI restriction (Figure 3, lanes 4 and 5). These forms were interpreted as defective replicative forms of the virus by reference to other geminiviruses (Stenger, 1995; Stanley et al., 1997; Liu et al, 1998).

The ability of the cloned AREV DNA-A to replicate in *N. tabacum* protoplasts was tested by inoculation of cell suspensions. The electroporation of a plasmid carrying a 1.8 repeat of AREV DNA-A resulted in

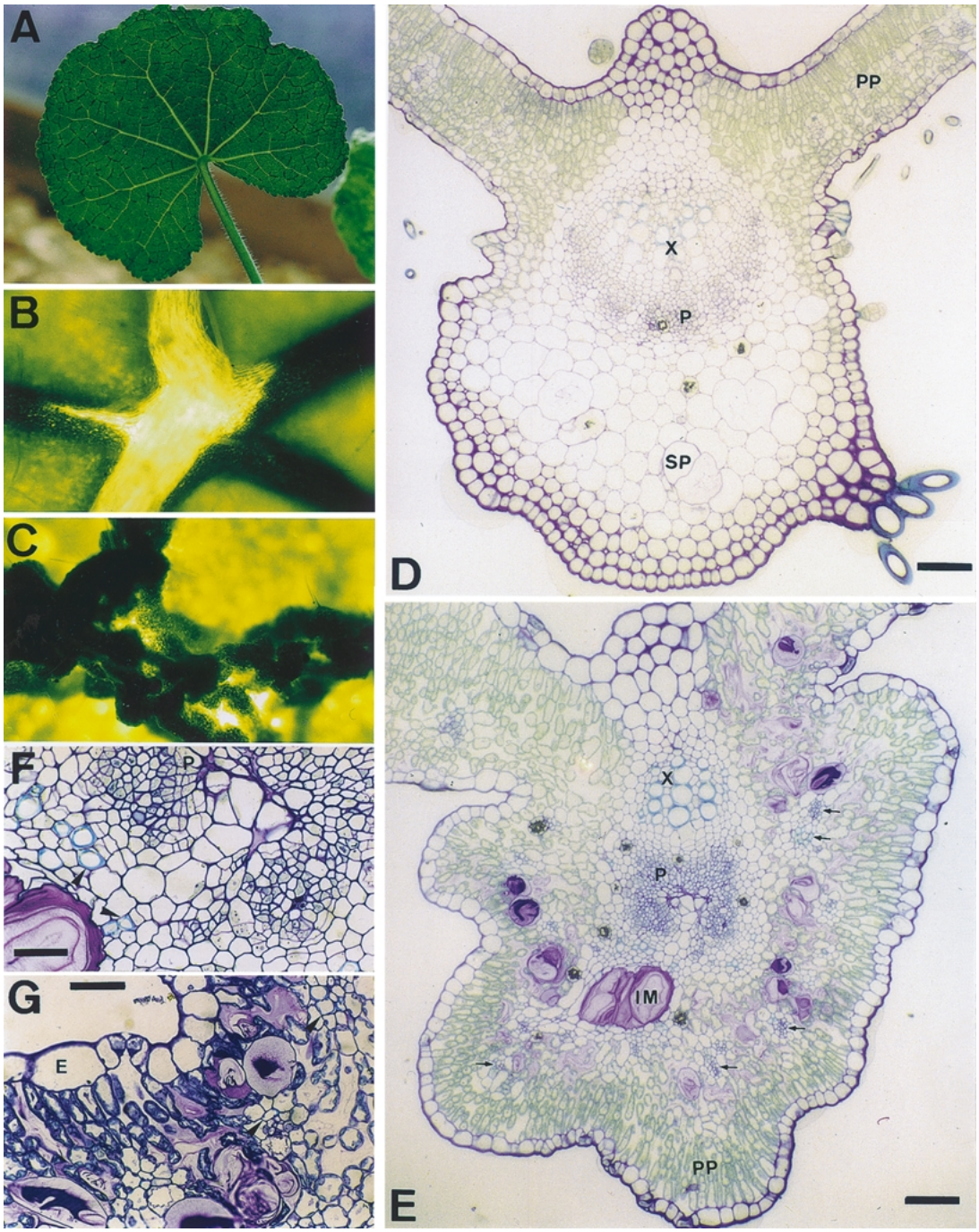
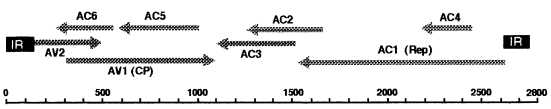


Figure 1. Symptomatology and histopathology in hollyhock plants infected by the *A. rosea* enation virus. (A) Vein darkening on the underside of a leaf. (B) and (C) Details of vein thickening and enations, respectively, on the underside of a leaf. (D) and (E) Light microscopy of leaf vessels of healthy (D) and virus infected (E) plants. Note the reorganization of the vascular system in the infected tissues with the presence of minor veins including thick-wall cells typical of xylem (arrows) and with the accumulation of a purple-colored intercellular material (IM). (F) Magnification of (E) showing the cambial activity of the phloem and the apparition of xylem tissues (arrowheads). (G) Minor veins (arrowheads) in the palisade parenchyma of infected tissues. P, phloem; X, xylem; SP, spongy parenchyma; PP, palisade parenchyma; E, epiderm. Bar D and E: 20 μm; bar F and G: 5 μm.

A



B

	DNA-A	AC1	AC2	AC3	AC4	AV1	AV2
AYVV	**68**	81	50	56	**73**	79	61
CLCuV-PK 802a	60	70	50	54	42	73	54
CLCuV-PK 806b	67	82	52	**60**	64	75	58
OYVMV-PK 201a	65	**83**	50	54	64	73	54
ICMV	61	75	48	51	59	77	60
TLCV-AU	66	82	49	51	**74**	76	51
TLCV-Nde2	61	74	44	54	57	74	57
TLCV-Ban2	62	76	51	**60**	43	68	55
ACMV-KE	64	75	**58**	57	56	75	**67**
EACMV-TZ	65	71	54	58	41	76	55
TYLCV-TH1	62	73	53	56	40	70	61
TYLCV-IL	**68**	77	57	59	46	**83**	61
TYLCV-Sar	65	77	50	56	64	79	53

C

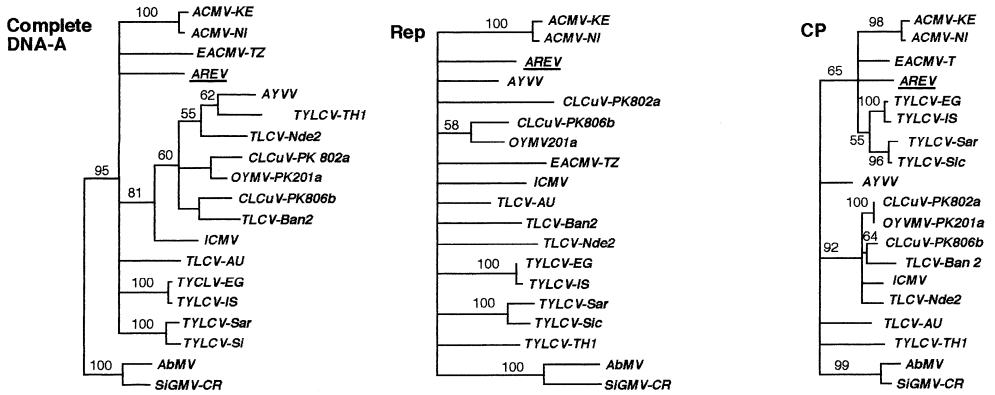


Figure 2. Genetic organization of the AREV genome and its similarities with other begomoviruses. (A) Open reading frames (ORFs) deduced from the sequence of AREV genome; potential encoded proteins are in brackets. (B) Levels of similarities between AREV and other selected begomoviruses. For individual ORFs, the percentages indicate amino acid similarities between putative proteins. For each ORF, the most similar sequence with AREV is indicated in bold. (C) Phylogenetic trees obtained with MEGALIGN and validated with PAUP 3.1. Numbers indicate the percentage support of the bootstrap for each branch. The sequence of an isolate of *Sugarcane Streak Egyptian Virus*, belonging to the mastrevirus genus, was used as an outgroup (not shown).

the release and the replication of a 1-mer copy, as revealed by the detection of replicative forms similar to the forms detected in DNA extracted from infected *A. rosea* leaves (Figure 3). Only the faster-migrating ssDNA form of the two previously described for infected hollyhock DNA appeared in tobacco protoplast DNA. Compared to DNA plant extracts, an additional band below ssDNA is detected which is interpreted as scDNA based on its susceptibility to *Bam*H1 treatment. Two defective forms were weakly detected in protoplast DNA after digestion with *Bam*HI (Figure 3, lane 3).

Infectivity of AREV on whole plants was tested by agroinoculation of *N. benthamiana*, hollyhock and several cotton cultivars. Severe leaf curling and enations on veins appeared on all the leaves of *N. benthamiana* plants 2–3 weeks after agroinoculation (6/15 inoculated) (Figure 4). However, repeated agroinoculation experiments of *A. rosea* and cotton using two different strains of *Agrobacterium* failed to produce any symptoms. Using PCR with primers oAREV589 and oAREV1166, no product was obtained with DNA extracted from the youngest leaf of ten agroinoculated *A. rosea*, although a product of the expected size (578 bp) was detected for the naturally infected plants as well as for the agroinfected *N. benthamiana* (Figure 5). This result indicates that another factor is required for hollyhock infection, possibly a second component. We tried to detect a putative typical DNA-B by using PCR amplification of the full-length genome with two overlapping primers located in the intergenic region, oAREV28 (5′-TCCAAGTCAATCGGTGTGACA-3′) and oAREV29 (5′-ATTCATTTGGTGCCTGCTATC-3′). Since this region is usually very similar between A and B components of bipartite begomoviruses, the two primers should amplify DNA-A as well as a putative DNA-B. Several PCR products were produced by this

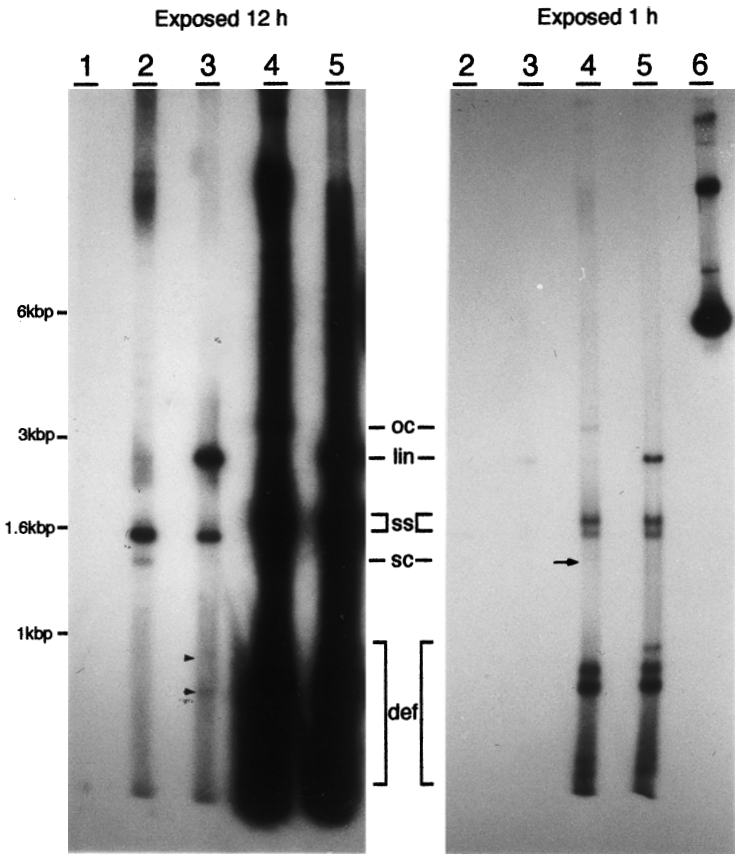


Figure 3. Southern-blot analysis of total DNA extracted from AREV infected tobacco (*N. tabacum)* protoplasts and hollyhock leaves. DNA was run at 30 V overnight in a 1% agarose gel in TAE buffer, without ethidium bromide. 1: Mock inoculated (healthy) tobacco protoplasts; 2: AREV electroporated tobacco protoplasts; 3: Same as 2, restricted with the endonuclease *Bam*HI which cuts once in the viral DNA; 4: Infected hollyhock leaves; 5: same as 4, treated with *Bam*HI; 6: pBluescript containing the 1.8 dimer of AREV (input). Hybridization was carried out with a full-length AREV probe under high stringency conditions. ss: two DNA species not digested by *Bam*HI treatment and interpreted as single-stranded forms of the virus (see text); sc: supercoiled DNA susceptible to restriction by *Bam*HI. Note it is weakly detected in the DNA extracted from leaves as faintly detected here (lane 4, arrow) and under intermediate exposure time (not shown): oc: open circular form; lin: linear form; def: defective forms; weakly detected in protoplasts (lane 3, arrowheads).

method, the expected full-length genome of 2.8 kbp, and a set of at least six products ranging from 1.5 to 0.65 kbp. The composition of the 2.8 kbp product was initially studied by restriction analysis with several endonucleases. It was observed that the product was heterogeneous as it was partially cut by the endonuclease *Eco*RI, for which there is no predicted site in the AREV DNA-A sequence. This result could be due either to the quasi-species (molecular heterogeneity) nature of the virus, or to *Taq* polymerase errors during PCR, or to the presence of a B component possessing an *Eco*RI site. The last hypothesis was tested by analysis of 45 cloned PCR products with primers oAREV589 and oAREV1166. The expected PCR product was obtained with all the clones suggesting that the 2.8 kbp product was mainly, if not exclusively, composed of AREV DNA-A molecules (not shown). Since the shorter PCR products (0.65–1.5 kb) hybridized with a DNA-A specific probe, they were interpreted as amplified defective forms of AREV DNA-A. Finally, this strategy did not allow us to detect a DNA-B component.

Four defective forms of 1.4 kb were cloned and sequenced in order to look for sequences originated from possible recombination events between the DNA-A and a second putative component as described

Figure 4. AREV agroinoculated *N. benthamiana* leaves showing severe curling and pronounced enations on the veins (left), and mild curling and enations (middle), compared to a healthy leaf (right).

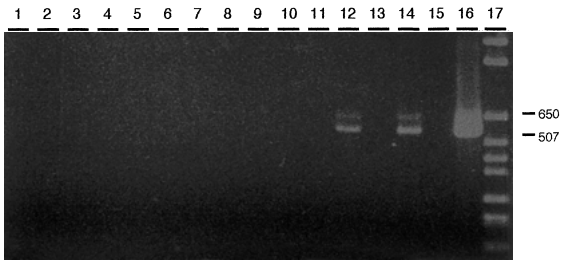


Figure 5. AREV detection in agroinoculated plants based on PCR amplification with primers oAREV589 and oAREV1166. About 100 ng of total plant DNA were submitted to PCR with primers and final products were run on a 1.2% agarose gel. Lanes 1 to 10: Inoculated *A. rosea*; lane 11: Healthy *N. benthamiana*; lane 12: Agroinoculated (symptomatic) *N. benthamiana*; lane 13: Healthy *A. rosea*; lane 14: Positive control: Infected *A. rosea* (whitefly-inoculated); lane 15: PCR with water instead of DNA; lane 16: Positive control: PCR using cloned AREV DNA in a pBluescript; lane 17: Ladder. Apart from the expected 580 bp product, an unexplained additional minor product of about 630 bp was consistently resolved with infected plants, as well as with the plasmid carrying the viral clone, indicating its viral origin.

for the AYVV. Sequence analysis revealed that the four defective forms contained the intergenic region of AREV, parts of the AC1 and AV1 ORFs resulting from deletions and recombination events. Non-viral DNA sequences of 300–500 bp were occasionally inserted in these defective forms. This foreign DNA did not show any homology with any sequence in the databases.

Relationship of AREV with other begomoviruses

Comparison of the complete AREV DNA sequence with other geminiviruses revealed 68% identity with the closest viruses, namely *Tomato yellow leaf curl virus* from Israel and Egypt (TYLCV-IL) and *Ageratum yellow vein virus* from Singapore (AYVV) (Figure 2B). The lowest identity with an old-world begomovirus was observed with CLCuV-PK802a (60% nucleotide identity), which is still higher than the identities with any of the new-world begomoviruses, such as the *Abutilon mosaic virus* (51% nucleotide identity). To confirm that AREV is distinct throughout its entire DNA and does not result from recombination involving other begomoviruses, comparisons were individually carried out in each ORF. The similarity with AC1 (Rep) amino acid sequence ranges between 70 to 83% (Figure 2B). However, no particular relationship was observed between AREV AC1 and the same ORF of any other begomovirus (Figure 2C). Similar conclusions were obtained for other ORFs, AC2, AC3, AC4 and AV2 (Figure 2B), with no clear relationship with any of the other compared viruses; the highest similarity for these four ORFs was obtained with TLCV-AU for the AC4 protein (74%). The phylogenetic tree obtained for the CP protein further confirms that AREV is distinct (68–83% amino acid identity). However, the CP of AREV clustered with the African-Mediterranean geminivirus group which is distinct from the Australasian viruses (Figure 2C). Furthermore, within the African-Mediterranean group, AREV is separated from the TYLCVs subgroup, although the highest similarity level of the AREV CP protein is found with the geographically closest virus, the TYLCV from Egypt and Israel (83–84% amino acid).

Okra is naturally infected by a related, yet distinct, genotype of AREV

Transmission experiments indicated that the hollyhock disease can be whitefly-transmitted to okra (Chazly, unpublished). In order to determine if AREV is also associated with the disease on okra in natural conditions, PCR was carried out with primers oAREV589 and oAREV1166 on DNA from okra leaves exhibiting enations collected in Aswan (Upper-Egypt). A 510 bp region of *AV1* gene was amplified, directly sequenced and showed to be 90% identical to the cognate sequence of AREV (97% amino acid level) (GenBank AF155064). The okra isolate is thus clearly related to AREV although distinct at the nucleic acid sequence level.

Discussion

We report the association of a begomovirus with a whitefly-transmitted disease affecting hollyhock in Egypt. Microscopy showed that the enation phenomenon results from the abnormal development of vascular tissues as already described for the tobacco leaf curl disease (cited in Bawden, 1964). The phloem undergoes abnormal cell division activity, associated with the emergence of additional minor vascular structures. The dark green color of the enations is due to the replacement of the spongy parenchyma by a palisade parenchyma. This parenchyma appears plainly functional since numerous chloroplasts are observed and a network of minor veins is created within the palisade tissue. Comparatively, enations on hollyhock differ from those induced by the *Beet curly top virus* (BCTV), whose effects are limited to hyperplasia and hypertrophia of phloem parenchyma cells (Latham et al., 1997). Moreover, in the case of the *A. rosea* enations, we observed the abundant accumulation of pectic intercellular fibrous material. Although not yet elucidated, the accumulation of this material may reveal a plant reaction to the biotic stress, as pectins have been implicated in some plant–pathogen interactions as a response to limit plant tissue invasion by the pathogen (Rioux et al., 1998). Alternatively, this reaction may be a secondary effect of the abnormal biochemical activity of the palisade parenchyma. Although other enation-inducing begomoviruses have been described in the literature (Tsai et al., 1997; Paximadis et al., 1999), little is known about the onset of the symptoms. The cytological properties of the enations supports the emerging idea that geminiviruses can specifically alter the host cell cycle, maybe to provide a molecular environment favorable for their own replication (reviewed by Hanley-Bowdoin et al., 2000).

AREV DNA-A was unable to induce symptoms on a range of malvaceous plants, including the original host. However, it could systemically invade plants of the permissive host *N. benthamiana* suggesting that the molecule is able to replicate. Initial attempts to find a second genetic component failed. However, Saunders et al. (2000) recently claimed the detection of a DNA-β on a hollyhock isolate from Egypt. Whether this molecule is also present in our isolate remains to be determined, but its association with AREV DNA-A would reinforce the previously noticed similarities between AREV and AYVV: (i) AREV exhibits one of its highest identities with AYVV comparatively to other begomoviruses and (ii) like AYVV, low levels of scDNA and a number of foreign DNA containing-defective forms are found in infected plants, an observation which suggests common features in the replication process for both viruses.

By Southern blot and PCR, we detected defective replicative forms of AREV. Defective DNA forms arising by illegitimate recombination are frequently found with geminiviruses. It was suggested that these short DNAs may have a role in the evolution of the geminiviruses since some defective DNAs carry non-viral sequences of unknown origin (Liu et al., 1998; Stanley et al., 1997; our observations). Furthermore these recombinant molecules can modulate the symptoms (Frischmuth et al., 1997). Interestingly, individual tobacco plants, infected by CLCuV-PK and containing several small defective DNA species, supported the replication of a particular predominant molecule (Liu et al., 1998). This last observation may explain the range of symptoms observed during our transmission assays (results not shown), if we consider that the accumulation of the defective forms differs qualitatively and quantitatively from plant to plant. Alternatively, the range of symptoms may be attributed to differences in the inoculum which may vary during vector transmission and according to the quasi-species composition of the inoculum. Further work is needed to understand the etiology of the enation disease on hollyhock, and in particular, it would be of interest to determine the sequence composition of the defective DNAs and to test their effect on the symptoms of *N. benthamiana* and hollyhock. Presence of non-geminivirus pathogens should also be investigated since ssDNAs related to nanoviruses were also found associated with the cotton leaf curl disease in Pakistan (Mansoor et al., 1999).

When compared to other geminiviruses, AREV exhibits a level of nucleotide identity lower than 90%. If we consider the recommendations of the Second International Workshop on Geminiviruses (Puerto-Rico, 1998) fixing the level of nucleotide identity to distinguish two begomovirus species at 90%, AREV can be considered as a new species within the begomovirus genus. AREV is now recognized as a distinct species in the last report of the International Committee on Taxonomy of Viruses (Rybicki et al., 2000). The okra isolate is highly related to AREV with 97% peptide similarity in the CP. However, both isolates are relatively distinct at the DNA level. Since AV1 is one of the most conserved ORFs in begomoviruses, it is anticipated that

the identity between full-length DNAs found in hollyhock and okra will be less than 90%. For comparison, the amino acid identity between the CPs of CLCuV-PK802a and OYVMV-PK201a is also 97%, whereas identity of their full-length DNA is only 83%. Waiting for the complete sequence of the okra isolate, we propose to name it the okra strain of AREV (AREV-okra).

According to Idris et al. (2000), DNA-A of AREV and an unpublished sequence of a DNA-A cloned from a cotton leaf curl isolate in Sudan share 79%, indicating a genetic diversity in this region of Africa among begomoviruses associated with similar diseases. This data and the difference between AREV and the Egyptian okra isolate suggest that the genetic diversity in the begomovirus population associated with the leaf curl of the malvaceae in Egypt may be in the same range as the one found for the Pakistani population where a complex of distinct variants have invaded a range of malvaceae crops (Zhou et al., 1998; Sanz et al., 2000). The variation of the begomoviruses associated with the so-called 'African cotton leaf curl disease' is probably a major factor to explain the biological diversity of the isolates described in the early literature.

Hollyhock is widely affected by the enation disease in Egypt, and also in Jordan and Israel (our unpublished observations). This plant is clearly an important source of begomoviruses in the region. This situation is favored by the important populations of whiteflies and by the fact that hollyhock is a perennial host that propagates the disease vegetatively. Interestingly, we could transmit our hollyhock isolate on several other malvaceous species, including a weed (*Malva parviflora*) and okra (Chazly, unpublished). AREV may thus be hosted by other plants in natural conditions. The fact that CLCuV-PK related begomoviruses were found on hollyhock in Pakistan make this plant a potential reservoir for strains potentially adapted to cotton in Egypt.

To date, the incidence of the enation disease on Egyptian cotton (mainly *G. barbadense* var.) is very low, possibly as a consequence of the past efforts of breeders in Africa for introducing cotton varieties resistant to the leaf curl (Bink, 1975). This is consistent with the fact that our AREV isolate could be vector-transmitted to a local cotton cv. at a very low efficiency, and with the complete reversion of the symptoms (thin enations) observed within 2 months (unpublished data). Nevertheless, the survey of viral begomovirus populations in Egypt should be intensified to anticipate epidemics on cotton.

Acknowledgements

We thank M. Granier for maintaining the plants, J. Stanley for critically reading the manuscript and T. Bourreau for collecting the okra isolate. This work was supported by IRD, CIRAD and the French government. AREV was maintained in Montpellier with Agreement No. 970323 of the plant protection department.

References

Bawden FC (1964) Symptomatology: changes within infected plants. In: Plant viruses and virus diseases, 4th ed (p. 65) The Ronald press company, NY

Bendahmane M, Schalk HJ and Gronenborn B (1995) Identification and characterization of wheat dwarf virus from France using a rapid method for geminivirus DNA preparation. Phytopathology 85: 1449–1455

Bink FA (1975) Leafcurl and mosaic diseases of cotton in central Africa. Cotton Growing Reviews 52: 233–241

Brown JK (1992) Virus diseases. In: Hillocks RJ (ed) Cotton diseases (pp. 275–320) CAB International

Dry IB, Rigden JE, Krake LR, Mullineaux PM and Rezaian MA (1993) Nucleotide sequence and genome organization of tomato leaf curl geminivirus. Journal of General Virology 74: 147–151

Gawel NJ and Jarret RL (1991) A modified CTAB DNA extraction procedure for Musa and Ipomoea. Plant Molecular Biology Reporter 9: 262–266

Giha OH and Nour MA (1969) Epidemiology of cotton leaf curl virus in the Sudan. Cotton Growing Reviews 46: 105–108

Fauquet C and Thouvenel JC (1987) Okra leaf curl virus. In: Plant viral diseases in the Ivory Coast (pp 96–97) Éditions de l'ORSTOM, Paris

Frishmuth T, Engel M and Jeske H (1997) Beet curly top virus DI DNA-mediated resistance is linked to its size. Molecular Breeding 3: 213–217

Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S and Robertson D (2000) Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. Critical Reviews in chemistry and Molecular Biology 35: 105–140

Idris AM, Lee SH and Brown JK (2000) Identification of a new, monopartite begomovirus associated with leaf curl disease of cotton in Gezira, Sudan. Plant Disease 84: 809

Konate G, Barro N, Fargette D, Swanson MM and Harrison BD (1995) Occurrence of whitefly-transmitted geminiviruses in crops in Burkina Faso, and their serological detection and differentiation. Annals of Applied Biology 126: 121–129

Latham JR, Saunders K, Pinner MS and Stanley J (1997) Induction of plant cell division by beet curly top virus gene C4. Plant Journal 11: 1273–1283

Liu Y, Robinson DJ and Harrison BD (1998) Defective forms of cotton leaf curl virus DNA-A that have different combinations of sequence deletion, duplication, inversion and rearrangement. Journal of General Virology 79: 1501–1508

Mansoor S, Khan SH, Bashir A, Saeed M, Zafar Y, Malik KA, Briddon R, Stanley J and Markham PG (1999) Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology 259: 190–199

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

Padidam M, Beachy RN and Fauquet CM (1995) Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. Journal of General Virology 76: 25–35

Paximadis M, Idris AM, Torres-Jerez I, Villarreal A, Rey MEC and Brown JK (1999) Characterization of tobacco geminiviruses in the Old and New world. Archives of Virology 144: 703–717

Rioux D, Nicole M, Simard M and Ouellette GB (1998) Immunocytochemical evidence that secretion of pectin occurs during gel (gum) and tylosis formation in trees. Phytopathology 88: 494–505

Rybicki EP, Briddon RW, Brown JK, Fauquet C, Maxwell DP, Harrison BD, Markham PG, Bisaro DM, Robinson D and Stanley J (2000) Geminiviridae. In: van Regenmortel SMHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR and Wickner RB (eds) Virus Taxonomy: Classification and Nomenclature of Viruses, the Seventh Report of the International Committee on Taxonomy of Viruses (pp 285–297) Academic Press, San Diego

Sanz AI, Fraile A, Garcia-Arenal F, Zhou X, Robinson J, Khalid S, Butt T and Harrison BD (2000) Multiple infection recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. Journal of General Virology 81: 1839–1849

Saunders K, Bedford ID, Briddon RW, Markham PG, Wong SM and Stanley J (2000) A unique virus complex causes Ageratum yellow vein disease. Proceedings of the National Academy of Sciences USA 97: 6890–6895

Stanley J, Saunders K, Pinner MS and Wong SM (1997) Novel defective interfering DNAs associated with *Ageratum* yellow vein geminivirus infection of *Ageratum conyzoides*. Virology 239: 87–96

Stenger DC (1995) Genotypic variability and the occurrence of less than genome-length viral DNA forms in a field population of beet curly top geminivirus. Phytopathology 85: 1316–1322

Swanson MM and Harrison BD (1993) Serological relationships and epitope profiles of isolates of okra leaf curl geminivirus from Africa and the Middle-East. Biochimie 75: 707–711

Tsai M-C, Liu C-S and Su H-J (1997) Poinsettia leaf curl, a new disease caused by a geminivirus. Journal of Phytopathology 145: 345–350

Umaharan P, Padidam M, Phelps RH, Beachy RN and Fauquet CM (1998) Distribution and diversity of geminiviruses in Trinidad and Tobago. Phytopathology 88: 1262–1268

Zhou X, Liu Y, Robinson DJ and Harrison BD (1998) Four DNA-A variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminiviruses isolates from okra. Journal of General Virology 79: 915–923

Note added in proof

Recently, DNA-β molecules have been isolated from cotton and okra affected by the leaf curl disease in Pakistan (Briddon et al., 2001 Virology; Mansoor et al., 2001 Plant Disease). In both cases, these molecules are necessary in association with a DNA-A to reproduce the symptoms on their natural host.