Molecular characterization of a subgroup I geminivirus from a legume in South Africa

Li Liu,1 Tertia van Tonder,2 Gerhard Pietersen,2 Jeffrey W. Davies1 and John Stanley1

1 Department of Virus Research, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
2 Plant Protection Research Institute, Private Bag X134, Pretoria 0001, Republic of South Africa

A South African geminivirus for which we propose the name bean yellow dwarf virus (BeYDV) has been isolated from French bean (*Phaseolus vulgaris* cv. Bonus) showing stunting, chlorosis and leaf curl symptoms. A full-length cloned copy of the viral genome produced characteristic symptoms of the disease when reintroduced into French bean by agroinoculation, and was systemically infectious in *Nicotiana benthamiana*, *N. tabacum*, *Lycopersicon esculentum*, *Datura stramonium* and *Arabidopsis thaliana*. BeYDV resembles subgroup I geminiviruses which infect monocotyledonous plants in having a single DNA component, two non-overlapping virion-sense (V1 and V2) and two overlapping complementary-sense (C1 and C2) coding regions, and an intron within the complementary-sense coding regions that is excised to produce a C1C2 fusion protein. It is most closely related to tobacco yellow dwarf virus from Australia, the only subgroup I geminivirus previously known to infect dicotyledonous plants, although it is sufficiently dissimilar (65% nucleotide sequence identity) to be considered a distinct virus.

A disease occurring on French beans (*Phaseolus vulgaris*) has recently been responsible for severe yield losses in South Africa. The highest incidence of the disease occurred in the Northern Province and Mpumalanga districts, where it has been estimated to cause 85–92% reduction in bean yield in French bean (cv. Bonus). Symptoms of the disease are brittle and leathery primary leaves, thickened and shortened internodes and downward curling of young leaves. Preliminary identification by negative staining of virus-like particles with a characteristic twinned morphology associated with the disease suggested geminivirus aetiology (T. van Tonder & G. Pietersen, unpublished).

Geminiviruses are small single-stranded DNA plant viruses that are conveniently subdivided on the basis of genome organization, host range and insect vector (Briddon & Markham, 1995). All subgroup I members have single genomic components, are transmitted by a variety of leafhoppers and, with only one exception, infect monocotyledonous plants. In contrast, most subgroup III members have two genomic components (DNA A and B) that are organizationally distinct from their subgroup I counterparts, and are transmitted by the whitefly *Bemisi tabaci* (Gennadius) solely to dicotyledonous hosts. Subgroup II members infect dicotyledonous plants and their single genomic component in many respects closely resembles DNA A of subgroup III members, although they are transmitted by leafhoppers or treehoppers.

Geminiviruses that have previously been isolated from beans include bean golden mosaic virus (Howarth et al., 1985) and bean dwarf mosaic virus (Hidayat et al., 1993), both of which are subgroup III members originating from the New World. Bean summer death virus (BSDV), a leafhopper-transmitted geminivirus originating from Australia, was considered to be a strain of tobacco yellow dwarf virus (TYDV; Thomas & Bowyer, 1984) that is closely related to subgroup I geminiviruses but adapted to dicotyledonous plants (Morris et al., 1992).

To investigate the aetiology of the disease in South Africa, symptomatic plants were collected from Malelane in the Mpumalanga district. Immunosorbent electron microscopy (Roberts, 1986) using antisera against TYDV and BSDV (provided by J. E. Thomas, Department of Primary Industries, Australia), beet curly top virus (BCTV; provided by G. I. Mink, Washington State University, USA) and chickpea chlorotic dwarf virus (CCDV; provided by D. V. R. Reddy, ICRISAT, India) revealed weak serological relationships between the new virus and TYDV, BSDV and CCDV. No serological relationship was observed with BCTV, although all antisera readily detected the homologous virus in positive controls (data not shown).
Fig. 1. Organization of the BeYDV genome. The arrangement of virion-sense (V) and complementary-sense (C) ORFs with the potential to encode proteins in excess of 10 kDa molecular mass are shown relative to the large and small intergenic regions (LIR and SIR, respectively). With the exception of ORFs C1 and C2, ORFs start at the first in-frame ATG codon. ORF C1 starts at the second in-frame ATG and all of ORF C2 is shown (for reasons discussed in the text). Shaded ORFs are those that are conserved between subgroup I members. Removal of the intron (nucleotides 1722–1807, indicated by lines) serves to fuse ORFs C1 and C2 (ORF C1C2, from which Rep is produced), and ORFs C4 and C1 (ORF C4C1, found only in BeYDV and TYDV). Predicted molecular masses of proteins encoded by each ORF (coordinates of start and stop codons are shown in parentheses) are as follows: V1 (129–404), 10.2 kDa; V2 (420–1154), 27.1 kDa; C1 (2398–1523), 33.4 kDa; C1C2 (2398–1311), 39.4 kDa; C3 (2358–1978), 14.4 kDa; and C4C1 (1956–1523), 12.2 kDa. Nucleotide numbering starts at the adenine residue located immediately downstream of the Rep nick site (J) within the TAATATT+AC motif that is a feature of all geminiviruses analysed to date.

Total nucleic acids were extracted from individual symptomatic plants (Covey & Hull, 1981) and analysed by blot hybridization using probes derived from representatives of geminivirus subgroups I (TYDV), II (BCTV) and III (African cassava mosaic virus (ACMV) DNA A). Each probe readily detected the virus from which it was derived, but only the TYDV probe additionally hybridized weakly to viral DNA in extracts from symptomatic plants (data not shown). Using primers based on the TYDV sequence, a DNA fragment of approximately 930 bp was PCR-amplified specifically from infected plants. Sequence analysis indicated that the amplified DNA was related (approximately 66% identity) to the analogus region of the TYDV genome. On the basis of this relationship to TYDV and disease symptoms, we propose that the virus be called bean yellow dwarf virus (BeYDV), and refer to it by that name in this paper.

Because attempts to transmit the virus by mechanical inoculation to a number of dicotyledonous species were unsuccessful, the viral DNA was introduced into plants by agroinoculation. Digestion of pSK-BYD081 with Clal and PstI produced viral DNA fragments of approximately 1600 and 1000 bp (Fig. 1) that were cloned into pBluescript II SK+ to give pSK-BYD0.6 and pSK-BYD0.4, respectively. The Clal insert from pSK-BYD081 was cloned into these constructs to give pSK-BYD1.6 and pSK-BYD1.4, having two copies of the small intergenic region (SIR) and two copies of the large intergenic region (LIR), respectively. The partial repeats were subcloned into pBin19 (Bevan, 1984) using flanking Kpn1 and Xhol sites to give pBinBYD1.6 and pBinBYD1.4. Agrobacterium tumefaciens tumefaciens containing the Ti plasmid pGV3850 (Zambrzycki et al., 1983) was transformed with pBinBYD1.6 and pBinBYD1.4, and viral DNAs were introduced into plants by agroinoculation as described by Tan et al. (1995). Both clones were systemically infectious in Nicotiana benthamiana and N. tabacum (cv. Samsun), and clone pBinBYD1.4 was subsequently shown to infect Datura stramonium, tomato (Lycopersicon esculentum), French bean (cv. Bonus and Top Crop) and Arabidopsis hialiana ecotype Ler (clone pBinBYD1.6 was not tested in these hosts). Plants became stunted, and leaves developed interveinal chlorosis and severe downward curling symptoms. Symptoms observed in French bean (cv. Bonus) were similar to those that occur in naturally infected plants. The presence of BeYDV DNA in symptomatic tissues of each of these plant species was confirmed by blot hybridization, and typical geminate particles, isolated from N. benthamiana essentially as described by Dollet et al. (1986), were detected by electron microscopy after negative staining (data not shown). The clones did not produce systemic symptoms when agroinoculated into sugarbeet (Beta vulgaris cv. Giant Western), pea (Pisum sativum cv. Vedette), chickpea (Cicer arietinum), maize (Zea mays L. cv. Golden Bantam x Yellow Hybrid) and wheat (Triticum aestivum cv. Maris Huntsman), and viral DNA was not detectable in upper asymptomatic leaves, indicating that none of these plant species is a host for BeYDV.

The sequence of both DNA strands of the pSK-BYD081 insert was established using an automated model 373A DNA sequencing system (Applied Biosystems). The cloned BeYDV DNA is 2561 nucleotides in length, and encodes two virion-sense open reading frames (ORFs) (V1 and V2) and two complementary-sense ORFs (C1 and C2) that are conserved between all subgroup I geminiviruses, separated by two intergenic regions, the LIR and SIR (Fig. 1). The LIR contains the ubiquitous nonanucleotide sequence TAATATT+AC, located at the apex of a potential stem–loop structure, that is...
nicked at a specific position (↓) by the replication-associated protein (Rep) during the initiation of virion-sense DNA replication (Laufs et al., 1995). It also contains a tandem repeat of the sequence TGGAGCCA (nucleotides 2448–2463), located next to the complementary-sense consensus TATA box, that is reminiscent of the Rep-binding site within the tomato golden mosaic virus intergenic region (Fontes et al., 1994). Interestingly, a tandem repeat of the sequence GTGAGCCA (five nucleotides in common with the BeYDV sequence) occurs in the same relative position in TYDV DNA. However, the BeYDV and TYDV LIR sequences exhibit relatively low overall identity (43%), supporting the idea that they are distinct viruses.

Conservation of putative gene products between subgroup I members (Table 1) suggests possible functions for the BeYDV genes. The product of maize streak virus (MSV) ORF V1 is required for virus movement (Boulton et al., 1993). Of note, the hydrophobic domain identified in MSV protein V1 and conserved in BeYDV. MSV ORF V2 encodes the coat protein (Morris-Krsinich et al., 1985). Geminivirus coat proteins are probably a major determinant of insect-vector specificity (Bridgon et al., 1990). Although the BeYDV vector remains to be identified, the BeYDV coat protein shows greatest similarity to that of TYDV (Table 1), which may reflect a closer relationship of the BeYDV vector to Orosius argentinatus, the TYDV vector (Thomas & Bowyer, 1984), than to the vectors of subgroup I members that infect monocotyledonous plants. Wheat dwarf virus (WDV) ORF C1 encodes a regulatory protein involved in the control of virion-sense gene expression (Collin et al., 1996).

Although the first in-frame ATG of BeYDV ORF C1 occurs at nucleotide 2470, it is the second in-frame ATG, at nucleotide 2398, that corresponds to the putative initiation codons found in other subgroup I geminiviruses. This, together with the fact that the first in-frame ATG is located upstream of the complementary-sense consensus TATA box (nucleotides 2440–2436), suggests that it is the downstream ATG that is functional.

ORF C2 does not contain an appropriately placed ATG, suggesting that it is expressed by post-transcriptional processing as has been demonstrated for digitaria streak virus (DSV), WDV and TYDV (Accoto et al., 1989; Schalk et al., 1989; Morris et al., 1992). To investigate splicing in BeYDV, total nucleic acids were extracted from infected D. stramonium plants agroinoculated with cloned viral DNA, and high-molecular-weight RNAs were isolated by salt precipitation and treated with DNase I (Pharmacia) as described by Covey & Hull (1981). A cDNA to complementary-sense transcripts was synthesized by reverse transcription using a primer corresponding to nucleotides 1435–1451 within ORF C2. DNA fragments were PCR-amplified from the cDNA using primers corresponding to nucleotides 1533–1562 and nucleotides 2227–2256 (complementary-sense) that flank the predicted intron. Size markers were PCR-amplified from clones pSK-BYD081 (wild-type DNA) and pSK-BYDΔintron (derived from pSK-BYD081 by deletion of the putative intron; to be described elsewhere) using the same primers. In addition to the anticipated 724 bp fragment produced from unspliced transcript, a smaller less abundant fragment occurred with an electrophoretic mobility comparable to that of the fragment produced from the deletion mutant pSK-BYDΔintron (data not shown). Sequence analysis of clones derived from these fragments confirmed the presence of an intron encompassing nucleotides 1722–1807 (Fig. 1). Transcript splicing serves to fuse ORFs C1 and C2, from which a Rep protein homologous to its counterparts in other subgroup I members may be expressed (Table 1).

In addition to the four conserved ORFs, BeYDV encodes complementary-sense ORF C3 with the capacity to encode a 14·4 kDa protein, and ORFs C4 and C1 are fused as a result of splicing to produce ORF C4C1 with a coding capacity of 12·2 kDa (Fig. 1). Analogous ORFs are present in TYDV.
Fig. 2. Phylogenetic tree based on complete nucleotide sequences of subgroup I members, rooted using DNA A of the African subgroup III member ACMV as the outgroup. Numbers above the branches indicate the number of times the branch was recovered. The BeYDV sequence was compared with those of TYDV (GenBank accession number M81103), WDV (X02869), MSV (X01633), panicum streak virus (PSV; X60168), DSV (M23022), sugarcane streak virus (SSV; M82918), chloris striate mosaic virus (CSMV; M20021), miscanthus streak virus (MSV; D01030), and ACMV (K02057).

(Morris et al., 1992), but not in subgroup I members that infect monocotyledonous plants (Table 1).

The complete nucleotide sequences of subgroup I members were aligned with the PileUp program of GCG software (Devereux et al., 1984), and a phylogenetic tree was established using the parsimony method after resampling using the SEQBOOT program of PHYLIP version 3.5 (Felsenstein, 1989). This revealed that BeYDV is more closely related to TYDV (65% identity) than to other subgroup I members (ranging from 46–49% identity) (Fig. 2). Comparison of amino acid sequences predicted for the conserved ORFs V1, V2, C1 and C1C2 shows a similar trend (Table 1), which is consistent with host-range adaptation being a major contributory factor to evolutionary divergence. This is not surprising in view of the limited genetic information encoded by the viral genome, implying a high dependence on host factors to maintain a productive infection of the host plant. For example, the regulation of virion-sense gene expression by protein C1 (Collin et al., 1996) may require host factors (Fenoll et al., 1990). Also, both Rep and C1 protein have been shown to interact with human retinoblastoma (Rb) protein (Xie et al., 1995; Collin et al., 1996), and an Rb homologue has recently been isolated from maize (Xie et al., 1996). Thus, the viral proteins may modulate the action of cell cycle regulatory proteins in order to create a cellular environment conducive to virus replication. The consensus Rb-binding motif, LeuXCyxGlu, found in the majority of subgroup I members is maintained at the analogous position in BeYDV ORF C1 (amino acids 205–209).

There have been no previous reports to indicate that adaptation of subgroup I viruses to dicotyledonous hosts is a widespread phenomenon. However, BeYDV and TYDV have been isolated from the widely separated geographical locations of South Africa and Australia, respectively. Furthermore, a leafhopper-transmitted geminivirus from India, CCDV, appears to be unrelated to subgroup II and III members but shows a weak serological relationship to TYDV and BSDV (Horn et al., 1993) as well as to BeYDV (this study; T. van Tonder & G. Pietersen, unpublished). Thus, CCDV may represent a third example of a subgroup I member that is adapted to dicotyledonous hosts. BeYDV shares several hosts in common with TYDV (this article; Morris et al., 1992), and symptoms induced by TYDV and BSDV (Thomas & Bowyer, 1984) in many respects resemble those described here for BeYDV, although the level of homology that exists between these viruses clearly demonstrates that BeYDV is a distinct virus.

Genetic analysis of BeYDV is currently in progress to investigate whether the predicted BeYDV genes perform similar functions to their counterparts in subgroup I members that infect monocotyledonous plants, and to establish if the novel ORFs C3 and C4C1 are functional. It is anticipated that this work will provide a basis for the development of non-conventional resistance strategies, not only against BeYDV, but also against important pathogens such as MSV that infect hosts less amenable to genetic transformation. Finally, a comparative study of gene function and regulatory elements of subgroup I members that infect monocotyledonous and dicotyledonous plants will provide a unique opportunity to investigate virus factors that influence host range.

We thank M. S. Pinner (Norwich) and G. G. K. Kasdorf, A. E. C. Jooste, R. H. Olivier and S. M. Staples (Pretoria) for their technical assistance. Thanks also to J. E. Thomas, G. I. Mink and D. V. R. Reddy for kindly providing antisera for this work. L.L. was supported by a studentship from the John Innes Foundation. This work was funded in part by a grant-in-aid to the John Innes Centre from the BBSRC.

References


Received 29 January 1997; Accepted 23 April 1997