Arch Virol (2001) 146: 71–86

Archives of Virology

Additional Rep-encoding DNAs associated with banana bunchy top virus

C. L. Horser[*], **M. Karan**[*,**], **R. M. Harding,** and **J. L. Dale**

Centre for Molecular Biotechnology, Queensland University of Technology, Brisbane, Queensland, Australia

Accepted July 5, 2000

Summary. Banana bunchy top nanovirus (BBTV) has a multicomponent circular single-stranded DNA (cssDNA) genome consisting of at least six components. We have cloned, sequenced and analysed two additional cssDNA components, designated BBTV-S1 and S2, associated with a Taiwanese BBTV isolate. The sequences of BBTV-S1 and S2 comprised 1109 and 1095 nucleotides (nt), respectively, and like BBTV DNA-1, potentially encoded replication initiation proteins (Reps). However, the genome organisation of BBTV-S1 and S2 differed from that of BBTV DNA-1 in that (i) the stem sequence of the CR-SL was not conserved, (ii) the internal gene was absent and (iii) the probable TATA boxes were located 5′ of the stem-loop. Further, sequence and phylogenetic analysis of the Rep genes indicated that BBTV DNA-S1 and S2 were distinct from BBTV DNA-1. When different geographical isolates of BBTV were tested for the presence of BBTV-S1/S2, these components were detected in various isolates from Vietnam, Taiwan, the Philippines, Tonga and Samoa but were not detected in isolates from Australia, Egypt, Fiji, and India. Based on these results, BBTV-S1 and S2 do not appear to be integral components of the BBTV genome and represent additional Rep-encoding DNAs associated with BBTV.

Introduction

Banana bunchy top virus (BBTV) is an isometric virus with a circular single-stranded DNA (cssDNA) genome consisting of at least six components (BBTV

[*]These authors contributed equally to the work described here.

[**]Current address: Laboratoire de Physiologie Cellulaire Vegetale, URA 576 CEA-CNRS-Universite Joseph Fourier Grenoble, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France.

DNA-1 to -6) [4, 7]. Each of these components is approximately 1 kb and shares a common genome organisation: all have a stem-loop, a potential TATA box 3′ of the stem-loop, at least one major gene in the virion sense and a polyadenylation signal associated with each gene [2–4]. Beetham et al. [2, 3] demonstrated that the major genes of DNA-1 to 6, as well as the small gene within the major gene of DNA-1, were transcriptionally active. Two regions of homology have been identified between all six BBTV components: (i) a 69 nucleotide (nt) stem-loop common region (CR-SL) which comprises the potential stem-loop and has at least 62% homology between components and (ii) a major common region (CR-M), a 66-92 nt region located 5′ of the CR-SL with at least 76% homology between components [4]. The major gene of DNA-1 encodes a replication initiation protein (Rep) [6] while DNA-3 encodes the viral coat protein [18]. The presence of an LXCXE motif in the predicted gene product of BBTV DNA-5 suggests it may function as a retinoblastoma binding-like protein [19] while BBTV DNA-4 and 6 appear to encode proteins involved in cell-to-cell movement and nuclear shuttling, respectively [19]. No putative function has been determined for BBTV DNA-2.

BBTV DNA-1 has been identified in all BBTV isolates tested [8]. Further, Karan et al. [8, 10] and Wanitchakorn et al. [20] demonstrated that there are two groups of BBTV isolates based on sequence comparisons of DNA-1, 3 and 6. The South Pacific group comprises isolates from Australia, Burundi, Egypt, Fiji, India, Tonga and Western Samoa while the Asian group comprises isolates from the Philippines, Taiwan and Vietnam. Importantly, BBTV DNA-1 to 6 have been confirmed in all isolates tested from both the South Pacific and Asian groups, implying that all six DNAs are integral components of the BBTV genome [10].

Three additional cssDNAs have been associated with BBTV in Taiwan [21, 22]. These components, designated here as BBTV DNA-Y [22], BBTV-W1 and -W2 [21], are all approximately 1 kb and contain a putative stem-loop structure with a nonanucleotide loop sequence almost identical to that found in BBTV DNA-1 to 6. Unlike BBTV DNA-1 to 6, however, the CR-SL and CR-M of BBTV DNA-Y, W1 and W2 are not conserved, while the TATA box is located 5′ of the stem-loop [4, 7, 21, 22]. Two components, BBTV-W1 and W2, each contain a major open reading frame (ORF) which appears to encode a Rep based on the presence of a GXGKT/S motif, while BBTV DNA-Y contains two small ORFs of unknown function [21, 22]. Interestingly, the potential Reps of BBTV-W1 and W2 share 50.3% and 52.9% sequence similarity to BBTV DNA-1 at the nucleotide level [21], respectively, indicating that they are not variants of the Rep encoded by BBTV DNA-1 but represent distinct Reps. This is further supported by the absence, in BBTV-W1 and W2, of the small internal ORF present in BBTV DNA-1 [2, 22].

In this paper, we report the isolation and characterisation of two new putative Rep-encoding components associated with Taiwanese BBTV isolates and the detection of these components in some, but not all, geographical isolates of BBTV.

Materials and methods

Virus isolates

BBTV-infected tissue was obtained from Australia, Egypt, Fiji, Hawaii, India, Samoa, Tonga, the Philippines, Taiwan and Vietnam. Tissue was used immediately or stored at −80 °C.

Total nucleic acid extraction

For PCR, banana midrib tissue was ground to a fine powder in liquid nitrogen. Approximately 0.2 g of powder was suspended in 500 µl 1% SDS and extracted with phenol/chloroform (1:1). The aqueous phase was collected and nucleic acid precipitated with 100% ethanol, washed with 70% ethanol and resuspended in 50 µl of water. For Southern hybridisation, total nucleic acid was extracted from up to 5 g of banana tissue by grinding to a fine powder in liquid nitrogen and emulsifying in 10 ml of extraction buffer (100 mM Tris-HCl, pH 8.0, 25 mM EDTA, 1% SDS, 0.2% 2-mercaptoethanol) and 10 ml of phenol:chloroform (1:1). After centrifugation, the aqueous phase was retained and extracted with an equal volume of chloroform. Nucleic acids were precipitated, washed in 70% ethanol and resuspended in 600 µl of TE buffer, pH 8.0.

DNA amplification and cloning

PCR primers were derived from the sequence of BBTV DNA-Y [22] and from the sequence of clones characterised in this study. PCR mixes comprised 20 pmoles of each primer, 10 mM Tris-HCl pH 8.3, 5 mM KCl, 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.5 to 1.25 U Taq polymerase and 1 µl of total nucleic acid extract (diluted 1/100 to 1/1000 in water). PCR mixes were denatured at 94 °C for 4 min followed by 30 cycles of 94 °C for 1 min, 42–52 °C for 1 min and 72 °C for 2 min followed by 1 cycle of 72 °C for 10 min. PCR amplified products were cloned into pGEM-T (Promega) as previously described [12].

Sequence analysis

Clones were sequenced with an Applied Biosystems 373A DNA Sequencer. DNA templates were prepared as recommended by the manufacturer's protocol and primers used for sequencing were either universal sequencing primers (US Biochemical) or primers designed from cloned DNA. Sequences used for nucleotide and amino acid comparisons were obtained from GenBank through the Australian National Genomic Information Service (ANGIS). Programs within the English Genetic Computer Group (EGCG) included MAP, TRANSLATE for nucleotide sequence translation and DISTANCES to determine nucleotide and amino acid pairwise percentage similarities. Programs with the Phylogeny Inference Package (PHYLIP) were used for phylogenetic inference using amino acid alignments created in PILEUP (EGCG). ESEQBOOT was used to create one hundred bootstraps and the robustness of the tree was assessed using EPROTPARS and the consensus tree drawn using ECONSENSE and PHYLIP2TREE in Evolutionary Analysis programs of Web ANGIS GCG.

Southern blotting and hybridisation

Nucleic acids were electrophoresed in 1.5% agarose gels in TAE buffer, pH 7.8 [14] and stained with ethidium bromide. Prior to transfer, gels were denatured for 30 min in 0.5 M NaOH, 1.5 M NaCl followed by neutralisation for 30 min in 1 M Tris-HCl, pH 8.0, 1.5 M NaCl. Nucleic acids were transferred to positively charged nylon membranes (Roche) using capillary transfer [16]. Membranes were baked at 80 °C for 2 h then washed in 2×SSC. Hybridisation probes to detect BBTV DNA-1, 3 and S1 were amplified using primers designed from sequences of the complete ORF of BBTV DNA-1, 3 (Australia) and S1 (Taiwan),

and were labelled using DIG-11-dUTP (9:1) (Roche). Membranes were prehybridised, hybridised, washed under high stringency conditions and developed as recommended by the manufacturer (Roche).

Nucleotide sequence data

Nucleotide sequences for BBTV-S1 and S2 have been deposited with GenBank, accession numbers: AF216221 and AF216222. The GenBank accession numbers of sequences used for phylogenetic analyses were AR010225 (BBTV-1A), AR010233 (BBTV-1T), U02312 (BBTV DNA-Y), L32166 (BBTV-W1), L32167 (BBTV-W2); X80879, Y11405, AJ005964 and AJ005966 for FBNYV DNA-1, 2, 3 and 9, respectively; AB000920, AB000921, AB000922 and AB009047 for MDV DNA-1, 2, 3 and 10, respectively; U16731 and U16736 for SCSV DNA-2 and 6, respectively, and M29963 (CFDV).

Results

Amplification of additional Rep-encoding DNAs from BBTV isolates

Two primers, SF1 (5′ GATGATCGAAGCATCATCTGGG 3′, nt 554-575) and SR1 (5′ GGGTTCCTCGTCAATTGCCT 3′, nt 553-534), were designed from the sequence of BBTV DNA-Y [22] such that the primers were immediately adjacent and reversed in orientation and would amplify a full-length copy of BBTV DNA-Y. These primers were used in PCR with nucleic acid extracts from BBTV-infected bananas collected from Taiwan, Australia, Fiji, Tonga, Burundi, Western Samoa, India, the Philippines and Vietnam. A product was only amplified from the Taiwanese extract. This product of approximately 1 kbp was cloned and two clones (BBTV S1-a and BBTV S1-b) were fully sequenced. The clones contained inserts of 1109 nt and 1108 nt, respectively, and there were 8 nt differences between the two.

Two additional primers, SF3 (5′ GAAGACTTTCTCGCTCT 3′, nt 117-133) and SR2 (5′ TC GCTCAGCTGGGGAGG 3′, nt 116-100), were designed from the sequences of BBTV S1-a and S1-b. These primers were also adjacent and reversed in orientation such that amplification would result in full-length dsDNA copies of BBTV DNA-Y, BBTV S1-a and S1-b. Using these primers with the Taiwanese extract as template, a 1 kbp product was amplified and cloned. The sequence of one clone, BBTV S1-c, was very similar to BBTV S1-a and S1-b with only 13 nt differences between the three clones. A consensus sequence (1109 nt) was derived from the three clones and designated BBTV-S1 (Fig. 1a). At the nucleotide level, BBTV-S1 shared only 77% sequence homology with BBTV DNA-Y [22] and 40% sequence homology with BBTV DNA-1 from Australia [7]. However, sequence analysis of a second clone, BBTV S2-a, revealed approximately 64% similarity with BBTV-S1, 59% similarity to BBTV DNA-Y and 39% similarity to BBTV DNA-1 (Australia) over the entire nucleotide sequence. To verify the presence and sequence of BBTV S2-a, two degenerate primers, TF1 (5′ TTCSMRKGCTCACTGGGAGA 3′, nt 266-285) and TR1 (5′ TACYTCTTCTTCAATC3′, nt 262-246), were designed from sequences conserved between BBTV-S1 and S2-a. These primers were 3 nt apart and reversed in their orientation such that they would prime the amplification of near full-length

dsDNA copies of both BBTV-S1 and S2-a. Using the Taiwanese extract as template, a 1 kbp product was amplified and cloned. Two clones, designated BBTV S2-b1 and S2-b2, were sequenced. When the sequences were compared to S2-a, there were only 13 nt differences; a consensus sequence (1095 nt) was derived from BBTV S2-a, S2-b1 and S2-b2 and was designated BBTV-S2 (Fig. 1b). This sequence had 64% similarity to BBTV-S1, 59% similarity to BBTV DNA-Y and 39% similarity to BBTV DNA-1 (Australia) over the entire nucleotide sequence.

Comparison of BBTV DNA-1, W1, W2, S1, S2 and DNA-Y

When the sequences of BBTV-S1 and S2 were analysed, each contained one large ORF which had appropriately located possible TATA boxes and polyadenylation signals. These ORFs both potentially encoded 33 kDa proteins which appeared to be Reps based on the presence of dNTP binding motifs, GNEGKS (BBTV-S1) and GGEGKS (BBTV-S2) (Fig. 1a, 1b) [5]. The published sequence of BBTV DNA-Y contained two small ORFs (nt 55-360 and nt 392-604) [22] situated within the same region as the large ORFs of BBTV-S1 and S2. Manipulation of this sequence at two sites, a deletion at nt 238 (G) and a substitution at position 606 (G-C), created an ORF (nt 55-909) which also potentially encoded a 33 kDa protein with a dNTP binding motif (GNEGKS). This modified sequence was designated BBTV-Y1.

The genome organisation of BBTV-S1 and S2 resembled that of BBTV-Y1, BBTV-W1 and W2 and BBTV DNA-1 (Australia, Taiwan) in that (i) each was a circular DNA molecule of approximately 1.1 kb, (ii) each potentially encoded a putative Rep of approximately 33 kDa with appropriately located possible TATA boxes and polyadenylation signals, (iii) the potential polyadenylation signal of BBTV DNA-1, S1, S2, Y1, W1, and W2 (AATAAA) was fully conserved and located 22 nt 5′ of the end of the major ORF (Fig. 1a, 1b) and (iv) each contained a sequence 5′ of the ORF which was capable of forming a stable stem-loop structure (Fig. 1a, 1b). Further, the loop sequences were very similar to the highly conserved nonanucleotide loop sequence of BBTV DNA-1 to 6 (Fig. 2) [4].

However, while there were a number of similarities between BBTV DNA-1 (Australia, Taiwan) and BBTV-S1, S2, Y1, W1 and W2, there were also very marked differences (Fig. 3). Apart from the loop sequence, the CR-SL (of BBTV DNA-1 to 6) was not conserved in BBTV-S1, S2, Y1, W1 or W2. Further, although the CR-M of the Taiwanese BBTV DNA-1 (DNA-1T) had an equivalent in BBTV-S1 (95% sequence similarity), no equivalent CR-M was present in BBTV-S2, Y1, W1 or W2. The internal ORF identified in BBTV DNA-1 (Australia, Taiwan) [2, 4, 9] was also not present in the putative Rep-encoding sequences of BBTV-S1, S2, Y1, W1 or W2. Finally, the probable TATA boxes of BBTV DNA-1 to 6 (TATAA) were all located 3′ of the stem-loop sequence [4, 7] whereas the probable TATA boxes of BBTV-S1, S2, Y1, W1 and W2 were all located 5′ of the stem-loop.

When the nucleotide and predicted amino acid sequences of the Rep genes of BBTV DNA-1 (Australia and Taiwan) were compared with the major ORF

(a)

```
                CGAGGAAGTGGCCTAGTATTACCCACTTCCTCGCCCTTCTTCCTCGCCCCTACGTCATCA
   1                                                                        60

                GTATGTCATCTTTTAAATGGTGCTTCACTCTGAACTACTCCTCCGCAGCTGAGCGAGAAG
  61                                                                       120
                  M  S  S  F  K  W  C  F  T  L  N  Y  S  S  A  A  E  R  E  D

                ACTTTCTCGCTCTTCTGAAGGAGGAAGAGCTTAACTACGCAGTTGTCGGCGACGAAGTCG
 121                                                                       180
                  F  L  A  L  L  K  E  E  E  L  N  Y  A  V  V  G  D  E  V  A

                CTCCGAGCTCCGGCCAGAAGCACCTACAGGGATATCTATCCCTGAAGAAATCTATAAAAC
 181                                                                       240
                  P  S  S  G  Q  K  H  L  Q  G  Y  L  S  L  K  K  S  I  K  L

                TTGGAGGATTGAAGAAGAAGTATTCTTCGAGAGCTCACTGGGAGAGGGCGAGAGGATCTG
 241                                                                       300
                  G  G  L  K  K  K  Y  S  S  R  A  H  W  E  R  A  R  G  S  D

                ACGAAGATAATGCTAAGTATTGTTCTAAAGAAACCCTAATTCTTGAATTGGGGTTTCCCG
 301                                                                       360
                  E  D  N  A  K  Y  C  S  K  E  T  L  I  L  E  L  G  F  P  A

                CCTCTCAAGGCTCTAATAGACGGAAACTATCGGAGATGGTTTCTAGATCTCCAGAACGCA
 361                                                                       420
                  S  Q  G  S  N  R  R  K  L  S  E  M  V  S  R  S  P  E  R  M

                TGAGAATTGAACAGCCTGAGATATATCACAGATACACATCTGTGAAGAAGTTAAAAAAAT
 421                                                                       480
                  R  I  E  Q  P  E  I  Y  H  R  Y  T  S  V  K  K  L  K  K  F

                TCAAGGAGGAATTCGTTCATCCTTGCCTCGATAGACCATGGCAGATTCAATTGACGGAGG
 481                                                                       540
                  K  E  E  F  V  H  P  C  L  D  R  P  W  Q  I  Q  L  T  E  A

                CAATTGACGAGGAACCCGATGATCGAAGCATCATCTGGGTCTATGGTCCGAATGGTAATG
 541                                                                       600
                  I  D  E  E  P  D  D  R  S  I  I  W  V  Y  G  P  N  G  N  E

                AGGGGAAATCAACATATGCGAAGTCATTAATGAAGAAGGACTGGTTCTACACCAGAGGTG
 601                                                                       660
                  G  K  S  T  Y  A  K  S  L  M  K  K  D  W  F  Y  T  R  G  G

                GGAAGAAGGAGAACATACTGTTCTCTTACGTGGACGAAGGATCTGAGAAGCATATTGTAT
 661                                                                       720
                  K  K  E  N  I  L  F  S  Y  V  D  E  G  S  E  K  H  I  V  F

                TTGATATTCCTCGCTGTAATCAGGATTATTTAAATTATGATGTTATAGAGGCATTAAAGG
 721                                                                       780
                  D  I  P  R  C  N  Q  D  Y  L  N  Y  D  V  I  E  A  L  K  D

                ATAGGGTGATAGAGAGTACTAAATATAAACCTATTAAGTTAGTTGAATTGATTAATATAC
 781                                                                       840
                  R  V  I  E  S  T  K  Y  K  P  I  K  L  V  E  L  I  N  I  H

                ATGTAATTGTCATGGCTAATTTCATGCCAGAATTCTGTAAAATCTCCGAAGATAGAATAA
 841                                                                       900
                  V  I  V  M  A  N  F  M  P  E  F  C  K  I  S  E  D  R  I  K

                AGATTATTTATTGTTAAAAAGGAAATTATATTATGCACTATGACAATCGTACGCTATGAC
 901                                                                       960
                  I  I  Y  C  *
```

```
     AAAAGGGGACCACAAAGACTCGGGGGTTGATTGCGACATCCTAACGATTAAGGGCCGCAG
 961                                                              1020

     GCCCGTCAAGATGGAATGAACGGTCAGATTTGATTGCTTAGCCACGAAGGAACAACTTTA
1021                                                              1080

     AGTTCCTCGCCCC**TATATA**TAGTTTTGCC
1081                              1109
```

(b)

```
      AGGAGGAGCGCCTAGTATTACCGCTCCTCCTCGCCTTTCCTCCTCGCCCCTGACGTCATC
    1                                                              60

      ATTATGTCCTCTTTTAAATGGTGCTTCACTCTGAATTATTCCTCCGCAGCGGAGCGAGAA
   61                                                             120
         M  S  S  F  K  W  C  F  T  L  N  Y  S  S  A  A  E  R  E

      GACTTTCTCGCTCTTCTGAAGGAGGAAGAGTTAAATTACGCTGTCGTCGGCGACGAAGTC
  121                                                             180
      D  F  L  A  L  L  K  E  E  E  L  N  Y  A  V  V  G  D  E  V

      GCTCCGAGCACCGGTCGGAAGCACCTCCAGGGATATCTATCCCTGAAGAAATCTATTAAG
  181                                                             240
      A  P  S  T  G  R  K  H  L  Q  G  Y  L  S  L  K  K  S  I  K

      CTTGGTGGATTGAAGAAGAGGTACTCTTCGAAGGCTCACTGGGAGAGGGCGAGAGGAACT
  241                                                             300
      L  G  G  L  K  K  R  Y  S  S  K  A  H  W  E  R  A  R  G  T

      GATGAACAGAATCGCGGATACTGTTCGAAGGAAACCCTAGTTCTTGAACTGGGTACTCCG
  301                                                             360
      D  E  Q  N  R  G  Y  C  S  K  E  T  L  V  L  E  L  G  T  P

      GTGGTCCCTGGTTCGAAGAAGCGCAAGCTTCTCGAGAGATTCAGAGAGAGCCCTGAAGAA
  361                                                             420
      V  V  P  G  S  K  K  R  K  L  L  E  R  F  R  E  S  P  E  E

      TTGAAGATGGAGGATCCATCCAAGTATCGCAGATGCTTGGCAGTGGAATCGTTGAACAAT
  421                                                             480
      L  K  M  E  D  P  S  K  Y  R  R  C  L  A  V  E  S  L  N  N

      GCTAGGAAAAATTCTGAATGGGTTCATGAACTAAGAGAATGGCAAAATAAATTAATTCAA
  481                                                             540
      A  R  K  N  S  E  W  V  H  E  L  R  E  W  Q  N  K  L  I  Q

      CACATCGAAGGTGTTCCTGATGATCGAAGTATCATCTGGGTATACGGTCCCAACGGAGGC
  541                                                             600
      H  I  E  G  V  P  D  D  R  S  I  I  W  V  Y  G  P  N  G  G

      GAAGGAAAGTCAACCTTCGCAAGATATCTATCATTAAAACCTGGATGGGGATATATCAAC
  601                                                             660
      E  G  K  S  T  F  A  R  Y  L  S  L  K  P  G  W  G  Y  I  N

      GGTGGAAAGACGTCGGATATGATGCACATCATAACGATGGATCCTGATAATCATTGGATT
  661                                                             720
      G  G  K  T  S  D  M  M  H  I  I  T  M  D  P  D  N  H  W  I
```

Fig. 1 (*continued*)

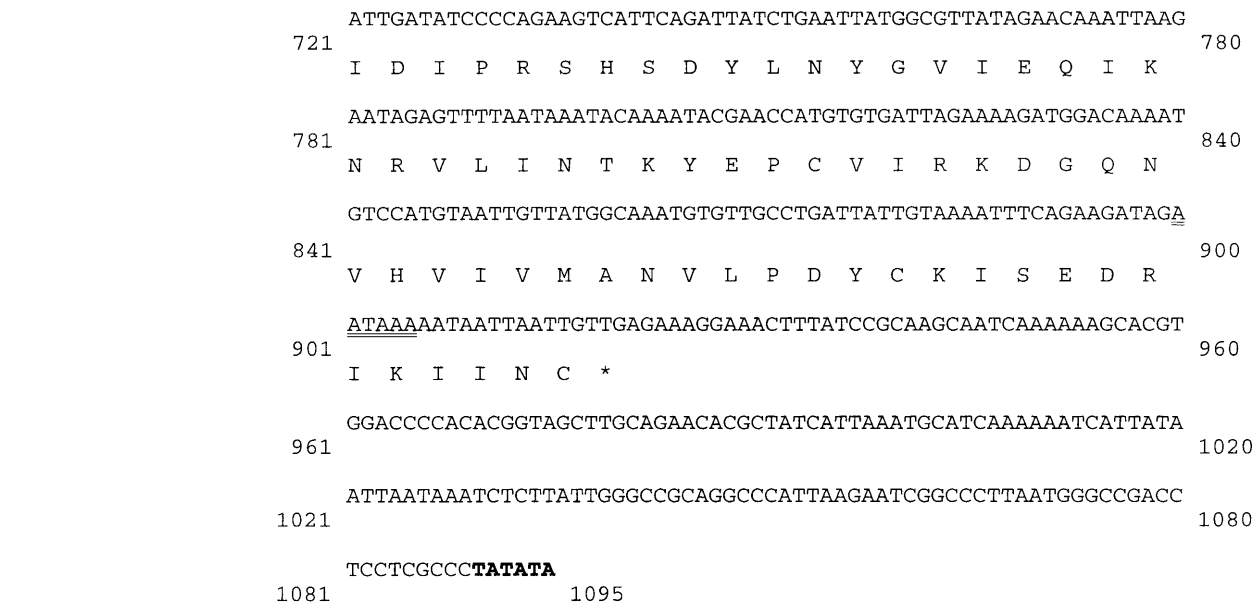

Fig. 1. Nucleotide and amino acid sequence of BBTV-S1 (**a**) and BBTV-S2 (**b**). Potential stem-loop is underlined, dNTP binding motif is bold and underlined, potential polyadenylation signal is double underlined and putative TATA box is in bold

sequences of BBTV-S1, S2, Y1, W1 and W2 (Tables 1 and 2), less than 50% and 36% homology was observed between BBTV DNA-1 (Australia and Taiwan) and BBTV-S1, S2, Y1, W1 and W2 at the nucleotide and amino acid levels, respectively. However, between BBTV-S1, S2, Y1, W1 and W2, there was between 60–97% nucleotide sequence homology and between 61–99% amino acid identity. These comparisons suggested that BBTV-Y and W1, and BBTV-S2 and W2, were variants of one another and that BBTV-S1 was a separate molecule.

Phylogenetic analysis of BBTV-S1 and S2

The relationships between the Reps associated with BBTV and those of other nanoviruses, including subterranean clover stunt virus (SCSV) [1], faba bean necrotic yellows virus (FBNYV) [11] and milk vetch dwarf virus (MDV) [15], were determined by phylogenetic analysis of the Rep sequences (Fig. 4). This analysis revealed that (i) BBTV DNA-1 from Australia (BBTV-1A) and Taiwan (BBTV-1T) grouped together and these were closely grouped with FBNYV-2, (ii) BBTV-Y1/W1 and S2/W2 grouped together, with BBTV-S1 grouped separately although more closely to BBTV-Y/W1 and (iii) MDV-10 and FBNYV-7, SCSV-6 and MDV-3, and SCSV-2 and MDV-2 grouped closely. Further, MDV-1, FBNYV-9, FBNYV-1 and CFDV grouped separately, with the CFDV Rep being the most dissimilar of all the Reps analysed. These relationships were confirmed by amino acid sequence comparisons (Table 3) which showed that BBTV-S1, S2/W2 and Y1/W1 had higher amino acid identity with the Reps encoded by FB-

Sequence	Stem	Loop	Stem
BBTV-1	AGCGCTGGGG	CTTATTATTAC	CCCCAGCGCT
BBTV-2	GGCGCTGGGG	CTTATTATTAC	CCCCAGCGCC
BBTV-3	AGCGCTGGGG	ACTATTATTAC	CCCCAGCGCT
BBTV-4	AGCGCTGGGG	CTTATTATTAC	CCCCAGCGCT
BBTV-5	AGCGCTGGGG	CTTATTATTAC	CCCCAGCGCT
BBTV-6	AGCACGGGGG	ACTATTATTAC	CCCCCGTGCT
BBTV 1-6*	AGCACGGGGG	CTTATTATTAC	CCCCAGCGCT
BBTV-S1	CGAGGAAGTGG	CCTAGTATTAC	CCACTTCCTCG
BBTV-S2	AGGA-G-CG	CCTAGTATTAC	CG-C-TCCT
BBTV-Y1	CCGAGG---TGG	CTTAGTATTAC	CCACCTCGG
BBTV-W1	CCGAGG---TGG	CTTAGTATTAC	CCACCTCGG
BBTV-W2	AGGAGG-AG-CG	GCTAGTATTACC	CGCTCCTCCT
Loop invariant sequence		TA-TATTAC	

BBTV 1-6* stem-loop consensus

Fig. 2. Stem and loop sequences of BBTV-associated components. Underlined letters represent conserved nucleotides

NYV (excluding FBNYV-2), SCSV, MDV and CFDV (39.1–55.96% homology) than with BBTV DNA-1 (<36% homology).

Geographical distribution of BBTV-S1 and S2

To determine whether components equivalent to BBTV-S1 or S2 were present in other BBTV infections, nucleic acid extracts from BBTV-infected bananas from both the South Pacific (Australia, Egypt, Fiji, India, Tonga and Samoa) and Asian (Philippines, Taiwan and Vietnam) groups of isolates were Southern blotted and hybridised with probes generated from BBTV-S1, BBTV DNA-1 (Australia) or BBTV DNA-3 (Fig. 5). The sequence similarity between BBTV-S1, S2 and Y1 indicated that they would probably cross-hybridise and this was verified for BBTV-S1 and S2 (results not shown). BBTV DNA-1 and BBTV-S1 did not cross-hybridise, while we have previously shown that probes generated from DNA-1 and 3 of South Pacific BBTV isolates hybridise with their respective components in the Asian group of isolates [9].

In duplicate Southern blots, BBTV infection was confirmed in all isolates by hybridisation with probes generated from BBTV DNA-1 and 3 (Fig. 5). Using the BBTV-S1/S2 probe, hybridisation was only observed with isolates from Tonga

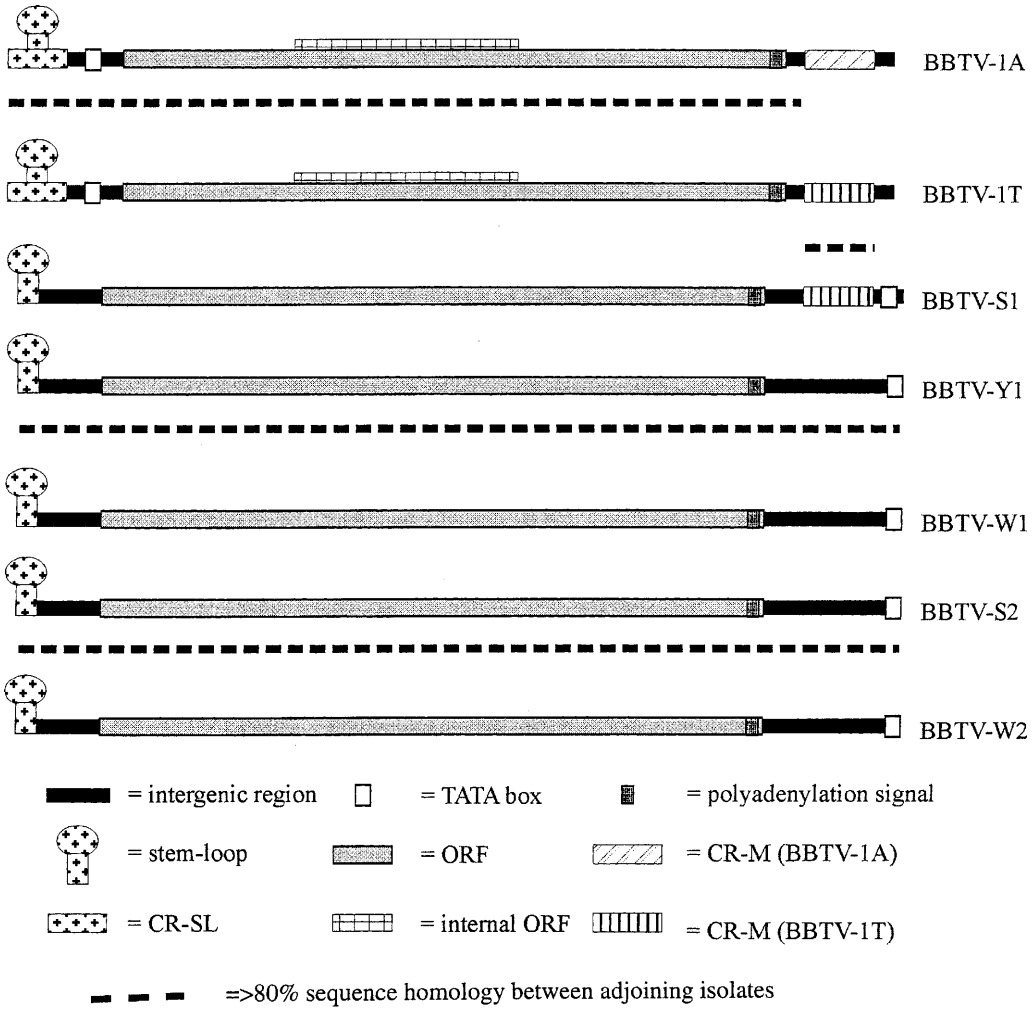

Fig. 3. Genome organisation of BBTV-associated Reps. BBTV-1A and 1T represent BBTV DNA-1 from Australia and Taiwan, respectively

Table 1. Nucleotide sequence similarity of potential Reps encoded by BBTV-1A, 1T, BBTV-S1, S2, Y1, W1 and W2

	BBTV-1A	BBTV-1T	BBTV-S2	BBTV-W2	BBTV-W1	BBTV-Y1	BBTV-S1
BBTV-1A	100.0	90.3	42.9	43.3	47.4	47.7	46.4
BBTV-1T		100.0	43.4	43.6	47.6	48.2	47.0
BBTV-S2			100.0	94.0	60.7	60.6	61.7
BBTV-W2				100.0	70.0	62.8	60.0
BBTV-W1					100.0	96.9	79.1
BBTV-Y1						100.0	79.3
BBTV-S1							100.0

Table 2. Amino acid sequence similarity of potential Reps encoded by BBTV-1A, 1T, BBTV-S1, S2, Y1, W1 and W2

	BBTV-1A	BBTV-1T	BBTV-S2	BBTV-W2	BBTV-W1	BBTV-Y1	BBTV-S1
BBTV-1A	100.0	94.4	34.1	34.8	35.4	35.2	35.3
BBTV-1T		100.0	34.1	34.1	35.4	35.2	36.3
BBTV-S2			100.0	92.7	61.4	61.2	64.2
BBTV-W2				100.0	62.8	62.3	61.4
BBTV-W1					100.0	99.0	87.7
BBTV-Y1						100.0	87.3
BBTV-S1							100.0

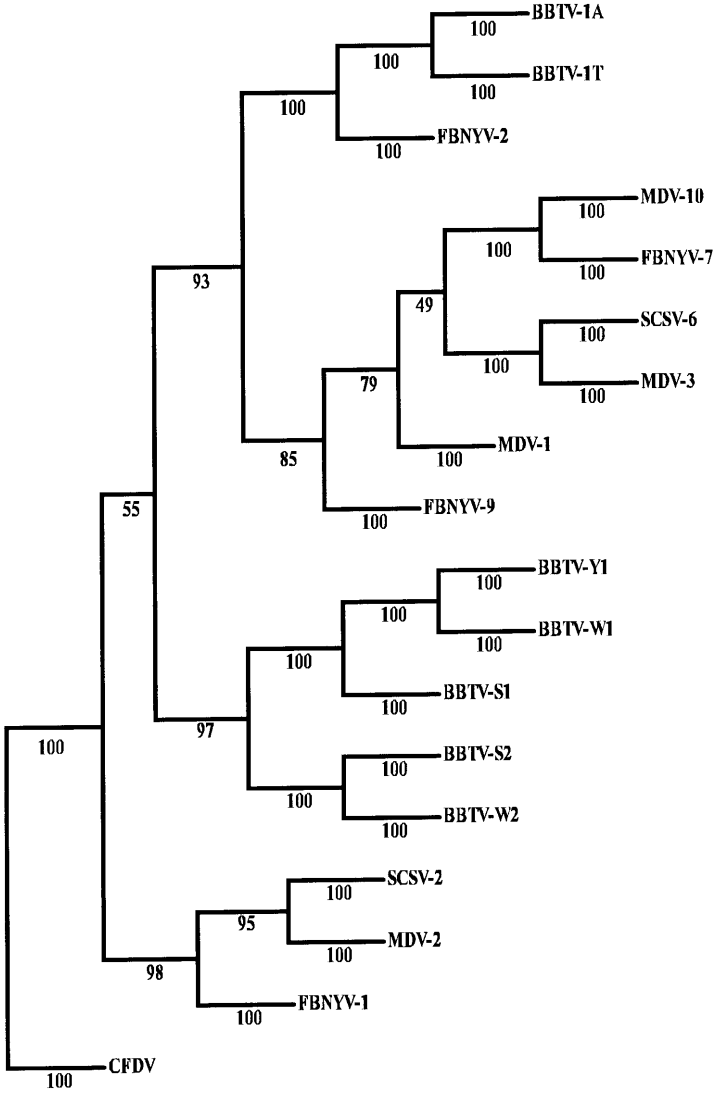

Fig. 4. Phylogenetic inference of Rep encoding proteins associated with BBTV, FBNYV, MDV, SCSV and CFDV. Numbers at nodes represent percent bootstrap scores

Table 3. Amino acid sequence similarity of potential Reps encoded by BBTV, FBNYV, MDV, SCSV and CFDV

	BT-1A	BT-1T	FB-2	FB-7	MD-10	MD-3	SC-6	MD-1	FB-9	BT-S2	BT-W2	BT-W1	BT-Y1	BT-S1	MD-2	SC-2	FB-1	CFDV
BT-1A	100.0	94.4	54.6	36.2	35.8	36.5	36.4	37.8	32.6	34.1	34.8	35.4	35.2	35.3	37.6	35.6	34.0	33.0
BT-1T		100.0	55.6	37.6	36.9	38.0	37.5	37.8	34.4	34.1	34.1	35.4	35.2	35.3	38.0	36.0	35.1	33.7
FB-2			100.0	37.0	38.5	38.9	38.8	38.7	33.5	33.5	32.7	36.3	36.1	34.3	35.9	36.1	36.7	29.0
FB-7				100.0	89.8	59.9	58.0	55.7	47.3	51.7	52.4	45.4	45.6	44.2	47.4	45.3	48.9	39.3
MD-10					100.0	58.5	56.9	57.1	47.3	51.7	51.7	45.0	45.2	44.2	46.2	45.7	48.5	40.7
MD-3						100.0	73.3	50.4	50.0	51.5	51.1	41.5	41.7	41.1	43.8	42.9	44.1	37.3
SC-6							100.0	47.7	52.3	51.3	50.6	39.6	39.7	39.1	45.2	43.8	47.7	37.6
MD-1								100.0	48.0	43.9	44.2	45.9	45.7	44.8	42.3	41.3	47.5	35.9
FB-9									100.0	41.4	40.7	40.4	41.3	39.3	40.7	39.3	50.4	37.3
BT-S2										100.0	92.7	61.4	61.2	64.2	54.9	51.3	43.1	47.2
BT-W2											100.0	62.8	62.3	61.4	55.6	52.0	42.4	47.2
BT-W1												100.0	99.0	87.7	56.0	53.5	41.3	42.1
BT-Y1													100.0	87.3	55.4	53.3	41.4	41.8
BT-S1														100.0	54.9	52.7	41.3	42.0
MD-2															100.0	81.5	64.4	45.5
SC-2																100.0	61.2	46.2
FB-1																	100.0	38.2
CFDV																		100.0

BT BBTV; *FB* FVNYV; *MD* MDV; *SC* SCSV

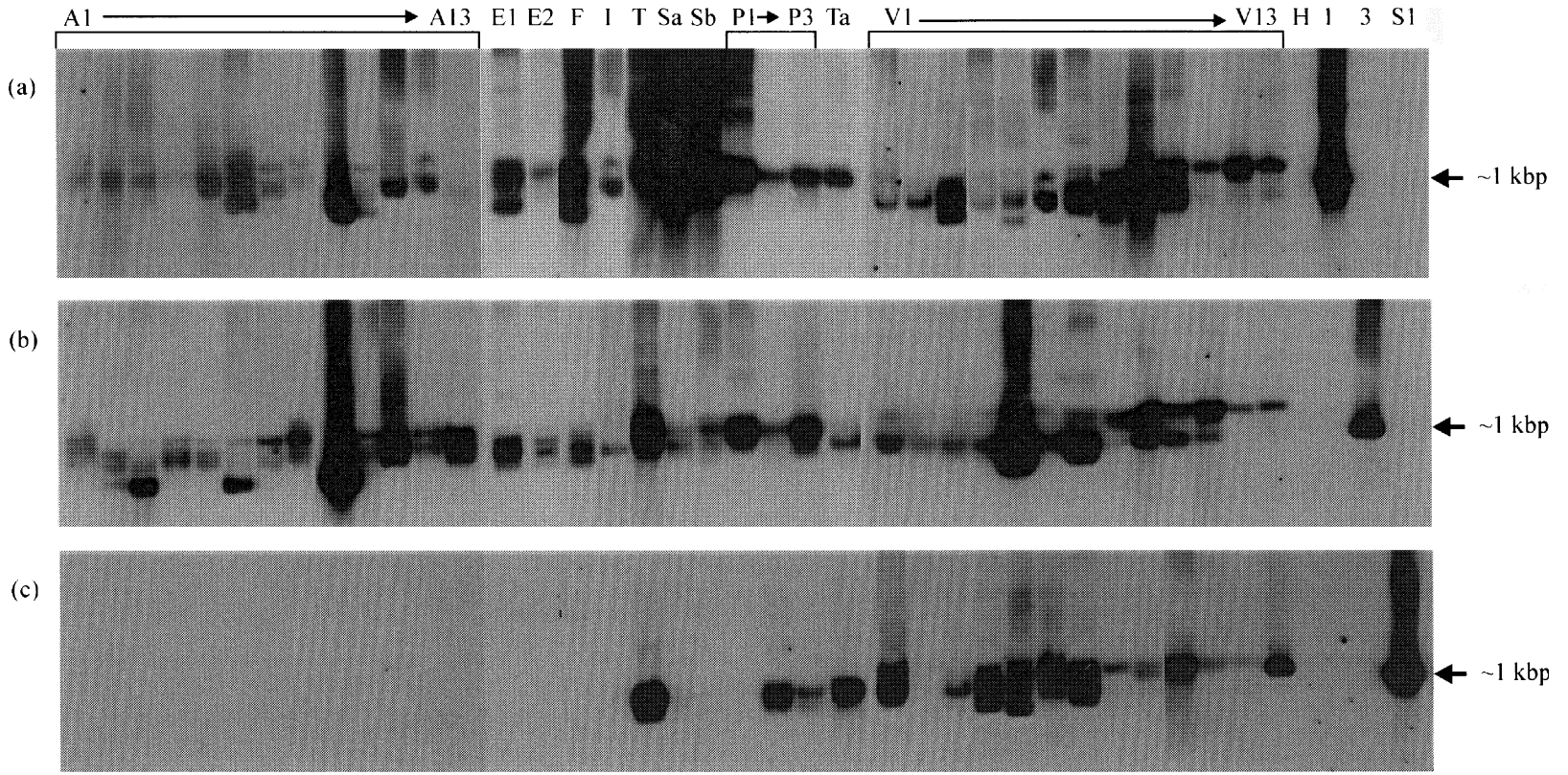

Fig. 5. Geographical distribution of BBTV-S1 – like sequences. *A1–A13* Australia; *E1, E2* Egypt; *F* Fiji; *I* India; *T* Tonga; *Sa, Sb* Samoa; *P1–P3* Philippines; *Ta* Taiwan; *V1–V13* Vietnam; *H* healthy. 1, 3 and S1 are component controls for BBTV DNA-1, 3 and S1, respectively. All samples were probed with BBTV DNA-1 (**a**), 3 (**b**) and S1 (**c**)

(1/1), Samoa (1/2), the Philippines (2/3), Taiwan (1/1) and Vietnam (13/13) although the hybridisation levels with some isolates (Samoa, Vietnam) were sometimes weak. The probe did not hybridise with nucleic acid extracted from any of the Australian, Egyptian, Fijian and Indian isolates.

Discussion

Six cssDNA components (BBTV DNA-1 to 6) have been consistently associated with BBTV worldwide in all geographical isolates tested [8, 10] although it has yet to be confirmed whether these components represent the complete infectious unit of BBTV. A further three DNA components, DNA-Y, W1 and W2, have also been associated with Taiwanese BBTV infections. We have now isolated two additional DNA components associated with BBTV infections, namely BBTV-S1 and S2. Like BBTV DNA-1, these additional components, with the exception of BBTV-Y, all potentially encoded Reps. However, it is possible that DNA-Y represents a variant Rep-encoding sequence since simple manipulation of the sequence creates a large ORF which potentially encodes a Rep (designated BBTV-Y1). Based on phylogenetic analysis, BBTV-S2 and W2, and BBTV-Y1 and W1 appeared to be variants with BBTV-S1 representing a distinct Rep-encoding sequence.

The results presented in this paper provide evidence that these additional DNA components are not integral to the BBTV genome. Firstly, the additional components are not present in all BBTV infections; although highly prevalent in the Asian group of BBTV isolates, they were only detected in two isolates from the South Pacific group, Tonga and Samoa. Although it is possible that these additional components, or related sequences, are more widespread in South Pacific isolates but not detected with the probes used in this study, we believe this is unlikely as the two groups of isolates differ by approximately 10% at the nucleotide level (based on BBTV DNA-1 and 6 sequences) [8, 10] and there is no reason to expect significantly more variability within the BBTV-S1, S2 and Y1 equivalent sequences between the two groups of isolates. Secondly, integral components of the BBTV genome would all be expected to share conserved features including the CR-M, the CR-SL and the location of the probable TATA boxes. While BBTV DNA-1 to 6 all share these conserved features, BBTV-S1 and S2 and Y1 do not, with the exception of BBTV-S1 which contained the CR-M of the Taiwanese BBTV DNA-1. Finally, there is no obvious explanation for BBTV to encode more than one Rep, since members of the *Geminiviridae*, the only other group of ssDNA plant viruses, only encode a single Rep.

There are now three other nanoviruses for which significant sequence data is available; subterranean clover stunt virus (SCSV) [1], faba bean necrotic yellows virus (FBNYV) [11] and milk vetch dwarf virus (MDV) [15]. For each of these viruses, several Rep-encoding components have also been identified. Four potential Rep-encoding components have been associated with FBNYV with FBNYV DNA-2 considered to encode the "master" Rep because (i) it is the only FBNYV Rep with the conserved CR-SL, (ii) it is consistently identified in numerous samples tested and (iii) it possesses the ability to initiate self-replication and

initiate the replication of the other six non-Rep components [17]. This suggestion is supported by phylogenetic analysis which showed that FBNYV-2 groups more closely with BBTV DNA-1, which we consider encodes the "master" Rep of BBTV, than to the other FBNYV Rep-encoding components. Two Rep-encoding components, SCSV DNA-2 and 6, have been associated with SCSV but we do not consider these to be the "master" Rep since (i) apart from DNA-2 and 6, all other components of SCSV share an intergenic region of 95% homology over 107 nt [1], possibly including the equivalent of the BBTV CR-M and (ii) these components are not found in all infections [1]. The genomes of MDV and CFDV, a possible nanovirus, also contain three and one Rep-encoding sequence, respectively, none of which we believe to encode the "master" Rep [1, 13, 15]. Recently, an additional Rep-encoding component, thought to represent the "master" Rep, has been isolated from SCSV and MDV infections [17], but the sequences of these components are not yet available for analysis. Interestingly, a comparison of the amino acid sequences of the nanovirus Reps showed that BBTV-S1, S2 and Y1 are more closely related to the Reps encoded by CFDV, MDV and FBNYV (except FBNYV DNA-2) than they are to the "master" Rep encoded by BBTV DNA-1.

Based on their limited geographical distribution and different genome organisation, we propose that BBTV-S1 and S2 (and Y1, W1 and W2) are non-essential Rep-encoding components of the BBTV genome. Although it is possible that these components are associated with another nanovirus, considering that similar DNA components have been isolated from three other nanoviruses (FBNYV, MDV and SCSV) would suggest that these additional Rep-encoding components are a characteristic of nanovirus genomes.

Acknowledgements

The authors wish to thank H. J. Su and A. K. Allam for providing banana samples from Taiwan and Egypt, respectively, C.H. was supported by an Australian Postgraduate Research Award.

References

1. Boevink P, Chu PWG, Keese P (1995) Sequence of subterranean clover stunt virus DNA: affinities with the geminiviruses. Virology 207: 354–361
2. Beetham PR, Hafner GJ, Harding RM, Dale JL (1997) Two mRNAs are transcribed from banana bunchy top virus DNA-1. J Gen Virol 78: 229–236
3. Beetham PR, Harding RM, Dale JL (1999) Banana bunchy top virus DNA-2 to 6 are monocistronic. Arch Virol 144: 89–105
4. Burns TM, Harding RM, Dale JL (1995) The genome organisation of banana bunchy top virus: analysis of six ssDNA components. J Gen Virol 76: 1471–1482
5. Gorbalenya AE, Koonin EV, Wolf YI (1990) A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS Lett 262: 145–148
6. Hafner GJ, Stafford MR, Wolter LC, Harding RM, Dale JL (1997) Nicking and joining activity of banana bunchy top virus replication protein in vitro. J Gen Virol 78: 1795–1799

7. Harding RM, Burns TM, Hafner GJ, Dietzgen RG, Dale JL (1993) Nucleotide sequence of one component of the banana bunchy top virus genome contains a putative replicase gene. J Gen Virol 74: 323–328
8. Karan M, Harding RM, Dale JL (1994) Evidence for two groups of banana bunchy top virus isolates. J Gen Virol 75: 3541–3546
9. Karan M (1995) Sequence diversity of DNA components associated with banana bunchy top virus. PhD thesis, Queensland University of Technology
10. Karan M, Harding RM, Dale JL (1997) Association of banana bunchy top virus DNA components 2 to 6 with banana bunchy top disease. Molecular Plant Pathology On-line [http://wwwbspporguk/mppol/]
11. Katul L, Timchenko T, Gronenborn B, Vetten HJ (1998) Ten distinct circular ssDNA components, four of which encode putative replication-associated proteins, are associated with the faba bean necrotic yellows virus genome. J Gen Virol 79: 3101–3109
12. Marchuk D, Drumm M, Saulino A, Collins FS (1991) Construction of T-vector, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res 21: 1154
13. Rohde W, Randles JW, Langridge P, Hanold D (1990) Nucleotide sequence of a circular single-stranded DNA associated with coconut foliar decay virus. Virology 176: 648–651
14. Sambrook J, Maniatis T, Fritsch EF (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
15. Sano Y, Wada M, Hashimoto Y, Matsumoto T, Kojima M (1998) Sequences of ten circular ssDNA components associated with the milk vetch dwarf virus genome. J Gen Virol 79: 3111–3118
16. Southern E (1985) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517
17. Timchenko T, Katul L, Sano Y, Gronenborn B, Vetten HJ (1999) Identification of the master rep protein encoding components of three nanoviruses. In "XIth International Congress of Virology Sydney, Australia, 1999" abstract No. VW56.05 page 106
18. Wanitchakorn R, Harding RM, Dale JL (1997) Banana bunchy top virus DNA-3 encodes the viral coat protein. Arch Virol 142: 1673–1680
19. Wanitchakorn R, Hafner GJ, Harding RM, Dale JL (2000a) Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6. J Gen Virol 81: 299–306
20. Wanitchakorn R, Harding RM, Dale JL (2000b) Sequence variability in the coat protein gene of two groups of banana bunchy top isolates. Arch Virol 145: 1–10
21. Wu RY, You LR, Soong TS (1994) Nucleotide sequence of two circular single-stranded DNA's associated with banana bunchy top virus. Phytopathology 84: 952–958
22. Yeh H, Su HJ, Chao Y (1994) Genome characterisation and identification of viral-associated dsDNA component of banana bunchy top virus. Virology 198: 645–652

Authors' address: Dr. J. L. Dale, Centre for Molecular Biotechnology, Queensland University of Technology, GPO Box 2434, Brisbane, Queensland 4001, Australia.

Received January 14, 2000