Evidence for two groups of banana bunchy top virus isolates

Mirko Karan, Robert M. Harding and James L. Dale*

Centre for Molecular Biotechnology, School of Life Science, Queensland University of Technology, Brisbane, Queensland, Australia

Banana bunchy top virus (BBTV) DNA component 1 from isolates from 10 different countries was cloned and sequenced and the sequences were aligned and compared. This analysis indicated two groups: the South Pacific group (isolates from Australia, Burundi, Egypt, Fiji, India, Tonga and Western Samoa) and the Asian group (isolates from the Philippines, Taiwan and Vietnam). The mean sequence difference within each group was 1.9 to 3.0% and between isolates from the two groups was approximately 10%, but some parts of the sequences differed more than others. However, the protein encoded by the major open reading frame, which is probably a replicase, differed by approximately 5%. The region from the beginning of the stem-loop sequence to the potential TATA box was identical in all isolates except for a two nucleotide change in the Western Samoan isolate and a single change in that of the NSW isolate. These results, together with other evidence, suggest that BBTV has spread to bananas after the initial movement of bananas from the Asian Pacific regions to Africa and the Americas.

Introduction

Banana bunchy top disease is the most serious of the viral diseases of bananas and is caused by banana bunchy top virus (BBTV), which is transmitted by aphids in a persistent manner. The virus has 18 to 20 nm isometric virions which contain circular ssDNA molecules of about 1 kb (Harding et al., 1991) and a single species of coat protein of Mr 20·1K. Harding et al. (1993) reported the sequence of one component of the BBTV genome. This component, BBTV DNA-1, was 1·111 kb and contained (i) a potential stem-loop structure with a sequence that was virtually identical to that of the invariant loop sequence of the geminiviruses and (ii) one long open reading frame (ORF) which would have encoded a protein of Mr 33·6K, likely to be a replicase as it contains the nucleotide-binding motif, GGEGKT. A further five ssDNA components of the BBTV genome have been identified (Burns et al. (1994). Each of these components had the same stem-loop sequence and four had a single large ORF in the DNA found in the virion, but not in its complement. Further, a second region of about 90 nucleotides, the Common Region-Major, was found to be closely similar in all components (Fig. 1). Yeh et al. (1993) have recently reported the sequence of an additional ssDNA component associated with BBTV. This component has little sequence similarity with BBTV DNA-1 or the other five BBTV ssDNA components (Burns, 1994).

Banana bunchy top disease is currently widespread in Asia and the South Pacific and is present in some parts of Africa but is not found in all banana-growing regions. It has not been recorded in Central or South America or in the Caribbean. Harding et al. (1991) and Thomas & Dietzgen (1991) isolated BBTV from bananas in Australia. Further, Dietzgen & Thomas (1991) demonstrated that BBTV was associated with the disease in Hawaii, Indonesia and Tonga, and Wu & Su (1990) demonstrated that the virus was present in Taiwan. However, the association of BBTV with the disease has not been confirmed for most countries where the disease occurs.

In this paper, we demonstrate that BBTV DNA-1 is present in all isolates of BBTV tested from 10 different countries and from the sequence data generated we have determined that there are probably two groups of BBTV isolates.

Methods

Collection of plant material. Banana tissue infected with BBTV was obtained from Africa (Burundi and Egypt), Asia (Taiwan, Vietnam, Philippines and India) and the South Pacific (Fiji, Tonga, Western Samoa and Australia (New South Wales (NSW), Queensland)) (Table 1).

Total nucleic acid extraction. Approximately 1 g of leaf midrib tissue was ground to a powder in liquid nitrogen. From this, 0·2 g of powdered tissue was mixed with 500 μl 1% SDS followed by extraction with phenol/chloroform (1:1) and then chloroform. The aqueous phase was collected and the nucleic acids were precipitated with ethanol, washed with 70% ethanol and resuspended in 50 μl water.

Amplification, cloning and sequencing. Oligonucleotide primers for PCR were designed using the published sequence of BBTV component...
Table 1. Geographical origins of isolates of BBTV

<table>
<thead>
<tr>
<th>Isolate origin</th>
<th>Abbreviation</th>
<th>Cultivar</th>
<th>Supplied by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taiwan</td>
<td>T</td>
<td>?</td>
<td>H. J. Su</td>
</tr>
<tr>
<td>Burundi</td>
<td>B</td>
<td>?</td>
<td>M. Iskra</td>
</tr>
<tr>
<td>Egypt</td>
<td>E</td>
<td>Cavendish</td>
<td>A. K. Allam</td>
</tr>
<tr>
<td>Fiji</td>
<td>F</td>
<td>Viemama</td>
<td>Collected by authors</td>
</tr>
<tr>
<td>Tonga</td>
<td>To</td>
<td>Cavendish</td>
<td>Collected by authors</td>
</tr>
<tr>
<td>W. Samoa</td>
<td>W</td>
<td>Cavendish</td>
<td>Collected by authors</td>
</tr>
<tr>
<td>Vietnam</td>
<td>V</td>
<td>Cavendish</td>
<td>Collected by authors</td>
</tr>
<tr>
<td>Philippines</td>
<td>P</td>
<td>Cavendish</td>
<td>Collected by authors</td>
</tr>
<tr>
<td>Australia</td>
<td>Q, N</td>
<td>Cavendish</td>
<td>Collected by authors</td>
</tr>
<tr>
<td>India</td>
<td>I</td>
<td>Robusta</td>
<td>J. E. Thomas</td>
</tr>
</tbody>
</table>

1 (Harding et al., 1993; GenBank accession no. S56276). The primer combinations used were; F3 5' GGAAGAGCCTCCTCATCTGCT-TCAGAGCC 3' (nt 287 to 258, Fig. 2) and FPCR4 5' TTCCCAG- GCGCACACCTTGAGAAAAAG 3' (nt 284 to 313) (Burundi, Egypt, Fiji, India, Qld, NSW, Taiwan, Tonga, Western Samoa), RV1 5' TACAGGATAATGCCTGAA 3' (nt 426 to 442) and RPCR1 5' TGACATCAAATAAATTA 3' (nt 425 to 409) (Philippines), F3 and FPCR3 5' CAGGCGCAGACCTTGGAGAAACGAAAG 3' (nt 288 to 317) (Vietnam). The primer pairs F3/FPCR4 had a four nucleotide overlap at their 5' ends. The primer pairs RV1/RPCR1 and F3/FPCR3 were adjacent. All primer pairs were reversed in their orientation such that they would prime the amplification of full-length dsDNA copies of BBTV DNA-1. Different primer pairs were used as not all primer pairs hybridized to the DNA of all isolates. The reaction mix (50 μl) contained 20 pmol each primer, 200 μM each dNTP, 5 mM-MgCl₂, 1.5 mM-Tris–HCl pH 8.3, 1 μl total nucleic acid extract (diluted 1:100 or 1:1000 in water) and 0.5 to 1.25 U Taq polymerase (Cetus). The mix was heated to 94 °C for 1 min; then 30 cycles of 94 °C for 1 min, 40 to 65 °C for 1 min, and 72 °C for 2 min; and finally 1 cycle of 72 °C for 10 min using a thermal cycler (Perkin Elmer Cetus).

The PCR products were cloned directly into either plasmid pCR 2000 (TA Cloning Kit; Invitrogen), pGEM-T (Promega) or Smal-digested, T-tailed pBluescript II KS plasmid (Stratagene) using the method of Marchuk et al. (1991).

Cloned DNA was sequenced using [35S]dATP and a Sequenase kit (US Biochemical) as recommended by the manufacturer. Reaction products were analysed by electrophoresis in a 6% (w/v) polyacrylamide gel containing 7 M-urea. Gels were fixed, dried and exposed to Agfa Curix RP1 film. The primers used for sequencing were either universal sequencing primers (US Biochemical) or 17 to 30 nucleotide primers complementary to appropriate regions of the cloned DNA.

Computer analysis. The computer programs used for sequence analysis were accessed from the Australian National Genomic Information Service (ANGIS), Sydney University. CLUSTAL V (Higgins et al., 1992) was used to align DNA and amino acid sequences and to construct Neighbor-Joining trees. The Wisconsin Genetics Computer Group (GCG) package of programs version 7.3 (Devereux et al., 1984) was used to translate DNA sequences (TRANSLATE) and to create similarity matrices (DISTANCES). Trees were drawn using the DRAWTREE program (PHYLIP package version 3.5c; Felsenstein, 1993).

Results

BBTV DNA-1 was amplified from 10 isolates of BBTV using two adjacent, outwardly extending primers designed from the original sequence of this component (Harding et al., 1993). The PCR products, which represented full-length dsDNA copies of this component, were cloned and sequenced. It was possible to amplify BBTV DNA-1 from all these isolates.

The sequences from the isolates were aligned and compared with each other and the original sequence of BBTV DNA-1 (Harding et al., 1993). However, we used a different numbering convention such that nucleotide 1 of BBTV DNA-1 is the first nucleotide of the putative stem–loop structure. The sequences of three PCR clones derived from a single Queensland (Nambour, Australia) BBTV isolate and pBT338 (Harding et al., 1993) (Currumbin) were compared; there were only five nucleotide differences between them, three of which were in the major ORF and two of these would have resulted in amino acid changes. When compared with a New South Wales (Australia) BBTV isolate, again there were only five nucleotide differences (Fig. 2), three within the ORF, one of which would have resulted in an amino acid change. These results suggested that there was little sequence heterogeneity among Australian isolates of BBTV.

However, when all the isolates were aligned and compared (Fig. 2), they formed two very distinct groups (Fig. 3a). Isolates from the South Pacific, Australia, Africa and India formed one group, the South Pacific group. There was a maximum of only 3.8% nucleotide sequence difference (mean of 1.9%) over the full sequence when each isolate was compared with all others in the group. The second group, the Asian group, included isolates from Taiwan, the Philippines and Vietnam. These isolates were a maximum of only 4.2% different (mean of 3.0%) at the nucleotide level when compared with each other. However, they were a minimum of 8.6% different (mean of 9.6%) from all isolates of the South Pacific group (Table 2). To estimate PCR error rate an
Two groups of BBTV isolates

Fig. 2. Alignment of nucleotide sequences of BBTV-1 from 11 isolates. The ORF is in bold; the potential TATA box is double underlined; the Common Region-Stem Loop is underlined; the potential poly(A) signal is in bold type and underlined; and the Common Region-Major is in bold type and double underlined.
additional eleven PCR clones derived from the various isolates were partially or fully sequenced. The sequence difference between clones of the same isolate varied from 0% to 1.2% and the average variability from the additional 7.66 kb of sequence was 0.37%.

BBTV DNA-1 has two regions. One is the untranslated region which includes the Common Region-Stem–Loop (CR-SL), the Common Region-Major (CR-M) and a potential TATA box. The other has a single major ORF which includes a dNTP-binding motif and a polyadenylation signal. CR-SL includes the potential stem–loop structure, the sequence of which was the same in all isolates, except for two nucleotide differences in the Western Samoan isolate (Fig. 2). The potential TATA box sequence, TATAAATA, was the same in all isolates (Fig. 2). Further, the 20 nucleotide sequence between the 3' end of the stem–loop structure and the 5' end of the TATA box was also the same in all isolates other than one nucleotide change (G to A) in one isolate (NSW) as well as the 23 nucleotides 5' of the stem–loop structure (Fig. 2). This would suggest that this nucleotide sequence has an important function for instance as the promoter for the major ORF.

The most interesting sequence in the untranslated region was the CR-M. It has previously been found that this sequence of approximately 90 nucleotides was similar in six BBTV DNA components of one isolate (Queensland) (Burns, 1994). In this study, we found that CR-M is also closely similar in all isolates of each isolate group. For instance, within the South Pacific group, there was only a maximum of 10% sequence difference between the isolates (mean 3.5%) (Table 2). Of these isolates, that from Tonga was the most different and the other isolates differed at most by 3%. The Asian group differed, at most, by only 3% between the isolates and had a mean difference of only 2%. In contrast, there was minimum sequence difference of 30% (mean 32%) between the CR-M of the South Pacific group and the Asian group (Table 2). A taxonomy based on CR-M similarity (Fig. 3b) has the same topology as, but greater differences than, one derived from the full DNA-1 nucleotide sequences.

The major ORF of BBTV DNA-1 was first identified in the Queensland isolate and consisted of 858 nucleotides encoding a protein of 286 amino acids with an $M_r$ of 33.6K. This protein was believed to be a replicase as it
had a dNTP binding motif (Harding et al., 1993). This ORF was present in all 11 isolates sequenced in this study (Fig. 2). Further, the amino acid sequence of this dNTP-binding motif (GGEGKT) was fully the same in all isolates. The potential polyadenylation signal, AATAAA, located 5' of the stop codon (Fig. 2) was also the same in all isolates except that there was one nucleotide change in one isolate (Egypt: AATAAA to AATATA). When the nucleotide sequence of the ORF of each isolate and its encoded amino acid sequence were compared, the two groups, South Pacific and Asian, were again evident (Figure 3c and d). The ORFs were less variable than the complete sequences and the amino acid sequences were less variable than the nucleotides encoding them (Table 2). This reflected the fact that 103 of the 858 nucleotides within the ORF were variable, but only 21 were in first codon position, 12 in the second position and 70 in the third position (Fig. 2).

Discussion

We have compared the sequences of BBTV DNA-1 from 11 isolates all from different geographical regions. The major outcome of this comparison is that there appear to be two groups of BBTV isolates infecting bananas: the South Pacific group which includes isolates from Fiji, Western Samoa, Tonga, Australia, India, Burundi and Egypt, and the Asian group which includes isolates from Taiwan, the Philippines and Vietnam. This was based on the comparisons of the full sequences of BBTV DNA-1 from the isolates as well as the comparisons of selected regions within the component (Fig. 2; Table 2). However, some of the regions of BBTV DNA-1 were closely similar in all isolates irrespective of group. These regions probably have sequence-specific functions and included the stem–loop structure, the potential TATA box, the intervening sequence between the stem–loop structure and the potential TATA box (possibly a promoter region), and the dNTP-binding motif and the polyadenylation signal within the major ORF. The major ORF was present in all isolates but there were some nucleotide and amino acid sequence differences, particularly between isolates from the two groups. In contrast, CR-M, which is a sequence of about 90 nucleotides that is conserved, with some differences, between six different components of a Queensland isolate of BBTV (Burns, 1994) was highly conserved within each group but was very variable between the two groups. There is no obvious explanation for this but CR-M could be used to rapidly identify the origins and affinities of new isolates. At this time, there are no known biological characteristics that correlate with the two groups based on sequence presented here.

The first record of banana bunchy top disease was in Fiji in 1879. The history of the spread of the disease has been reasonably well documented perhaps because the symptoms are obvious (Dale, 1987). The disease was subsequently recorded in Formosa (Taiwan) in 1900, Egypt in 1901 and Ceylon (Sri Lanka) in 1913 (Dale, 1987). The disease also arrived in Australia in 1913 in a shipment of infected suckers (banana cuttings) from Fiji. This is believed to be the only occasion that BBTV was introduced into Australia but, in the following decade, the disease virtually destroyed the banana industry in its major production areas. Since that time, the disease has been recorded in progressively more banana-producing countries. However, the distribution of the disease is erratic through the Pacific, Asia and Africa. Interestingly, the virus is not present in the Americas or the Caribbean. Asia is the centre of origin for bananas and it is speculated that they moved from Asia to Africa from about 1000 to 3000 years ago and from there to the Americas about 500 years ago (Simmonds, 1959). Bananas are vegetatively propagated and therefore it can be assumed that viruses common in bananas when they moved from Asia to Africa and subsequently to the Americas, would have moved with them. However, BBTV does not occur in the Americas and has limited distribution in Africa and the only reports of the disease in Africa have been since 1900.

It is likely that BBTV began infecting bananas in the Asian Pacific regions after the initial movement of bananas from this region to Africa and the Americas. It is also possible that there are two sources of BBTV infections, one in Asia and another in the South Pacific. This is based on evidence that the two groups may have been separated for a considerable time as (i) there is little sequence difference between Australian isolates of BBTV DNA-1 and the virus has been in Australia for 80 years, (ii) BBTV was introduced into Australia from Fiji in infected suckers and now the sequences of BBTV DNA-

<table>
<thead>
<tr>
<th>Group</th>
<th>Full sequence (nucleotide)</th>
<th>CR-M (nucleotide)</th>
<th>ORF (nucleotide)</th>
<th>ORF (amino acid)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within South Pacific</td>
<td>19</td>
<td>3-5</td>
<td>1-7</td>
<td>1-3</td>
</tr>
<tr>
<td>Within Asian</td>
<td>30</td>
<td>2-0</td>
<td>3-2</td>
<td>1-9</td>
</tr>
<tr>
<td>Between groups</td>
<td>9-6</td>
<td>32-0</td>
<td>7-5</td>
<td>5-6</td>
</tr>
</tbody>
</table>
1 isolates from those two places have diverged by about 1% whereas (iii) BBTV DNA-1 (Australia and Fiji) differ from Asian isolates by nearly 10%.

There are some important questions that remain. For instance, do the sequences of other BBTV components confirm the existence of two groups of isolates; are there biological differences between the two groups of isolates; and what is the original host of BBTV?

References


(Received 5 April 1994; Accepted 27 July 1994)