Discussion
V. DISCUSSION

1. Sampling and DNA Extraction

Sampling techniques are an important consideration when assaying for banana viruses. Virus concentration and symptom expression can vary considerably from one plant to another, and even in different leaves of the same plant. Our studies demonstrated that the concentration of banana bunchy top virus was higher in young leaves, and decreased with increasing age of the leaf. Hence samples for banana bunchy top detection were taken from young leaves accordingly.

Finding an efficient and effective procedure for the recovery of viral DNA from infected plant parts and carrier aphids proved critical to the success of subsequent research and analysis in the present study. The DNA isolation technique had to yield clean stable DNA devoid of any PCR inhibitors in the nucleic acid extracts. Comparison of the four methods tested, namely Dellaporta Method, Plant Miniprep Method using QIAGEN Plant DNA Extraction Kit, CTAB method and Leaf Dip Method, the QIAGEN Plant DNA Extraction Kit fulfilled all the desired criteria, however the time spent and cost incurred per sample proved to be the major determinant. In the Leaf Dip method, the concentration of DNA obtained was low and showed minimum reproducibility. The CTAB method produced satisfactory DNA, comparable to those obtained from DNA extracted from the QIAGEN plant DNA extraction kit. The CTAB method was eventually adopted for all further assays and investigations, as the yield and quality of the nucleic acid was satisfactory and cost per extraction for this method was comparatively lesser than the kit. The only disadvantage
proved to be the time duration for extraction which was 6 hours as compared to 35-45 minutes by the kit.

2. Cloning and Analysis

Both ClustalW and PRIMERDESIGN software provided suitable primers which were synthesized from SIGMA. Primers which tested positive were consistent in their performance once annealing temperatures standardized. PCR was standardized for amplification of each of the BBTV components and cloned into pTZ57 (a pUC19 based vector, from MBI Fermentas, Lithuania) using T-A cloning technique. Sequences derived from the clones were subjected to a variety of bioinformatics analysis programs.

2.1 Taxonomic Structure of the Family

BBTV virus has been classified as belonging to the family Nanoviridae (reclassified and recognized by the International Committee on Taxonomy of Viruses, Vetten et al., 2004), and are distinct from geminiviruses in being transmitted by aphids and having multipartite (6-11 segments) ssDNA genome components of ~1 kb encapsidated in icosahedral particles (Gronenborn et al., 2002). The proposed name, nanovirus, is derived from nanos in Greek meaning dwarf, as these viruses have among the smallest known virions and genome segment sizes and are known for the stunting effects caused in plants. There are two Genera in the Nanoviridae family:

**Babuvirus** (6 segments, with two unidirectionally transcribed ORFs on DNA1)

**Nanovirus** (8 or more monocistronic segments).

BBTV belongs to the Genus: Babuvirus with 6 monocistronic segments. It’s nucleic acid genetic makeup comprises of multipartite segments of closed circular single
stranded DNA, each segment about 1kb in size, with the general host being members belonging to the family *Musaceae*.

**Multipartite Genome Nature of the Virus**

Multipartite Genome genomes are segmented (genomes which are divided into two or more physically separate molecules of nucleic acid), wherein each genome segment is packaged into a separate virus particle. These discrete particles are structurally similar and may contain the same component proteins, but often differ in size depending on the length of the genome segment packaged.

Genome segmentation reduces the probability of breakages due to shearing, thus increasing the total potential coding capacity of the genome. However, the disadvantage of this strategy is that all the individual genome segments must be packaged into each virus particle, or the virus will be defective as a result of loss of genetic information. Separating the genome segments into different particles (the multipartite strategy), removes the requirement for accurate sorting, but introduces a new problem in that all the discrete virus particles must be taken up by a single host cell to establish a productive infection. This is perhaps the reason multipartite viruses are only found in plants. Many of the sources of infection by plant viruses, such as inoculation by sap-sucking insects or after physical damage to tissues, result in a large inoculum of infectious virus particles, providing opportunities for infection of an initial cell by more than one particle (Course notes, Xiong).

**2.2 Predicted genome organization**

The sequence analysis results obtained using nucleotide–nucleotide BLAST exhibited that the gene inserts in pTZ57. BT-1, pTZ57. BT-2, pTZ57. BT-3, pTZ57.
BT-4, pTZ57. BT-5 and pTZ57.BT-6 were of viral origin and represented the various components of the virus. PCR Analysis of infected samples revealed that the six components of Banana bimchy top nanovirus (BBTV DNA-1 to -6), were present in all BBTV infections tested.

Genome organization in geminiviruses comprised of a conserved 5' intergenic region that included the plus-strand origin of replication and the promoters for leftward and rightward transcription (Sunter et al, 1991). The IR also contains a hairpin with a 9-base pair loop sequence conserved among all geminiviruses that is cleaved during initiation and termination of RCR and is essential for geminivirus replication (Orozco and Hanley-Bowdoin, L. 1997). Almost all of the nano and circovirus DNAs share this nanonucleotide sequence 5'-'TTATTATTA-3’ as well as the modular pattern of genome organization with geminiviruses. The six components of BBTV were found to share a common genome organization consisting of non coding regulatory sequences and a single large open reading frame in the virion sense in each of the components (Fig. 23). The non coding regions harboured a major common region (CR-M) which spanned 92 nucleotides and shared 76% identity amongst the components. 3’ to this stretch was another intergenic region of homology within which was present, a highly conserved nucleotide sequence which formed a secondary hairpin structure called the stem loop. The loop sequence named CR-SL, has the highly conserved nanonucleotide sequence similar to all geminiviruses (Lazarowitz, 1992). Within the geminiviruses this conserved nonanucleotide sequence has been shown to be part of the viral origin of replication (Heyraud et al., 1993).
Fig. 23  Basic Modular structure of each BBTV Component

Each of the six DNA components associated with BBTV-Bg encodes at least one gene. Analysis of BBTV-Bg DNA component 1 (BBTV DNA-1) revealed the presence of two transcribed ORFs in the virion sense. The major DNA-1 gene encodes a replication initiation protein (Rep) and is further supported by studies conducted by Hafner et al., 1997b. However no exact function could be attributed to the smaller ORF, although sequence analysis predicted the ORF to code for a 5kDa protein. Analysis of BBTV-Bg DNA component 2 predicted two possible ORF with a protein prediction analysis of a 10kDa protein, although no functional domain could be assigned to component2. Protein prediction analysis and Multiple Sequence Alignment showed that BBTV-Bg DNA component3 encoded the viral coat protein; this is supported by studies conducted by Wanitchakorn et al. (1997). The functions of the genes encoded by BBTV-Bg DNA-4, -5 and -6 were extensively studied and predicted to possess the following functions. BBTV-Bg DNA4 encoded a viral movement protein; the protein
sequence of component-5 BBTV-Bg showed significant homology to a class of proteins known to bind to the tumor suppressor protein pRB (Retinoblastoma-protein). Presence of nuclear localization signals and conserved domain homology with other nanoviruses indicated that the putative protein encoded by Component 6 of BBTV-Bg was a nuclear shuttle protein.

**Fig. 24** Diagramatic Representation of the organization of the ssDNA genome of BBTV. The single ORF on the virus sense strand is shown clockwise

**Promoter and Transcription Factor Analysis** revealed a potential TATA box 3' of the CR-SL and associated polyadenylation signals at the 3' end of the ORF. Components 1, 2, 3, 4, 5 and 6, each had one large open reading frame in the virion sense located 3' of the stem-loop structure. Each ORF had potential polyadenylation signals associated with it and each polyadenylation signal had an associated GT-rich region containing the nucleotide sequence TGGG. The upstream (A(T)TGTA motif (Sanfaçon, 1994;
Rothnie et al., 1994), and the downstream motif comprising of a GT-rich stretch (Gil & Proudfoot, 1984; Conway & Wickens, 1985) containing the TTGGG sequence have been shown to be involved in correct 3'-end processing of CFDV Rep gene transcripts (Merits et al., 1995). These results provided further proof that the viral genes were expressed unidirectionally.

3. Sequence and Functional Analysis of components 1-6

The comparisons of the complete genome sequences of BBTV DNA components 1-6 from the Bangalore isolate with other sequences in the databank revealed that BBTV strains shared a 90% nucleotide sequence identity although DNA sequence of the Bangalore strain that causes banana bunchy top and other isolates did show a certain degree of divergence. The BBTV protein sequence shares about 40% identities, with SCSV and CDFV and more distant similarity with the geminiviruses. The amino acid sequences and motif structures are so similar; it is tempting to speculate that BBTV and other nanovirus proteins have evolved from a common ancestor. All viruses classified under the nanovirus family were found to share secondary structural features in the non-coding regions. The conserved loop sequence in BBTV, considered to be involved in viral replication (Lazarowitz et al., 1992), occurred in all the nanovirus and even shared some amount of similarity to the stem-loop found in geminiviruses. Moreover MDV, FBNYV, SCSV and BBTV have the same aphid vector in A. craccivora. In addition the nanoviruses also exhibit homology in protein motifs protein domains and hence protein functions.

3.1 Putative function of BBTV-1 protein

The largest open reading frame found in BBTV-1, called ORF1, potentially
encodes a 33 kDa protein possessing motifs typical found in the replication initiation proteins of gemini- and nanoviruses (Gorbalenya et al., 1990; Boevink et al., 1995; Katul et al., 1997). The major gene of BBTV DNA-1 was also found to contain a smaller internal gene in a +2 reading frame and seemed to corroborate Beetham results (Beetham et al., 1999). The nucleotide sequence of BBTV DNA1 is most closely related to the Rep-encoding components of the nanoviruses SCSV (Boevink et al., 1995), CFDV, MDV, (Sano et al., 1998) and FBNYV, (Katul et al., 1997, 1998). The BBTV-1 ORF was found to encode a putative replicase protein based on the presence of an NTP-binding motif GEGEGKS. However, it is more likely that this gene product functions as a Viral replicase as well as a Replication Initiator Protein (RIP) based on the presence of certain sequence motifs (Ilyina & Koonin, 1992), conserved within this protein. Protein domain analysis confirmed that the protein encoded by the BBTV 1 ORF was homologous to the family of viral replicases and the helicase domain of superfamily 3 (SF3) viral helicases, such as E1 proteins of papillomaviruses (Gorbalenya et al., 1990, 1993). Three conserved domains, typical for enzymes initiating replication in the rolling circle mode have been identified in the diverse Rep proteins. Similar to SF3 helicases, the ORF-1 rep domain contains a well conserved P-loop (motif A) for dNTP-binding and two conserved motifs (B and C) that probably determine helicase function. Rep proteins of circoviruses, nanoviruses and geminiviruses consist of two domains, specifically, a N-terminal nicking domain (with conserved motifs 1, 2 and 3, characteristic of viruses (circoviruses, nanoviruses, geminiviruses, bacteriophage f×-174) that utilize the rolling circle mechanism of
replication), and a **C-terminal SF3 helicase domain** (conserved motifs A, B and C), (Stegner, 2000; Boevink et al., 1995; Laufs et al., 1995; Saunders et al., 1999).

**Fig. 25a.** Domain Organization of BBTV Rep protein. The different domains have been identified using amino acid sequence homology, Protein Prediction servers from EMBL, SwissProt, and ExPASy.

**Fig. 25b.** Representation of 3-Dimensional structure of BBTV Replicase protein as obtained from Modeller 2.0 Ver. The four major domains of the protein have been highlighted.
Alignment of the primary sequence of BBTV plant nanovirus Rep with replication initiator proteins of viruses forming rolling-circle replicons, revealed the existence of three highly conserved motifs (RCR-1 through -3) within the N-terminal 120 amino acids (Fig. 25a and 25b, above). This conservation is maintained with initiator proteins of parvoviruses, animal circoviruses and geminivirus which all use rolling-circle mechanisms of DNA replication. The function of motif RCR-1 (FLTVPxC) is not known. Motif RCR-2 (HUHUUU), with U representing bulky hydrophobic amino acids, is likely, based on similarity to metalloenzymes, to be involved in the coordination of divalent cation (Mg$^{2+}$ or Mn$^{2+}$) through the two invariant histidine residues, the activity of which is necessary for RCR initiator proteins (Ilyina & Koonin, 1992). The motif may also be responsible for binding divalent Zn$^{2+}$ since a number of DNA binding proteins have been found to contain Zn atoms in different co-ordinations and were required for maintaining stability of the protein. InterProScan Analysis revealed the presence of a P-Loop motif, which is common to ATP- and GTP-binding proteins (Saraste et al., 1990).

Walker et al., (1997) have shown that hydrophobic residues like tyrosine contribute significantly to hairpin DNA binding and that mutation of tyrosine to phenylalanine in the Rep proteins of adeno-associated virus type (AAV-2), greatly reduces the ability of the protein to bind AAV hairpin DNA. Motif 3, (VxKYxxK), which corresponds to the catalytic site for DNA cleavage, contains a tyrosine residue that is conserved in all ssDNA Rep proteins and is required for the initiation (nicking) activity through the invariant tyrosine residue. A PROSITE Scan revealed that the putative active site tyrosine in a conserved amino acid environment,
was similar to that of replication initiator proteins of bacteriophages, prokaryotic plasmids, and the Rep proteins of geminiviruses. Rep proteins are known to possess **origin-specific DNA cleavage and nucleotidyl transfer activity** and the presence of a conserved tyrosine (Y78 of Rep) is essential for these reactions. The C-terminal motif encompassed the Tyr residue(s) which formed a covalent link with nicked DNA.

Studies on bacterially expressed African cassava mosaic virus Rep (Stanley, 1995), Wheat Dwarf Virus Rep, and TYLCV Rep demonstrated that it had a site-specific nicking activity within the invariant 9-nt sequence of the loop (TTATTATAC) (Guiterrez, 1999). A similar Rep binding domain within the intergenic region was identified in tomato leaf curl virus (Behjatnia, et al., 1998). The presence of the conserved nanonucleotide within the BBTV intergenic region would indicate that a similar site specific nicking activity is directly responsible for the initiation of (+) strand replication and the site where DNA replication starts *in vivo*.

Replication studies conducted in FBNYV indicated that Rep-2 initiated the replication of all non-*rep* components in addition to its cognate DNA. None of the other Rep proteins of FBNYV were able to trigger replication of any DNA other than its cognate, further suggesting that such conserved sequences (present on all components) may contain specificity elements of Rep2 recognition and can thus be equated to conserved repeat elements (iterons) in the non-*rep*-coding components of BBTV (i.e. components 2-6). Thus, as long as a nanovirus genome component contained a specific signal recognized by a given Rep protein, that particular Rep protein initiated its multiplication. Multiple Sequence Analysis of the genome components of BBTV revealed conserved sequences of about 70 nt shared by *rep* and
all DNAs encoding proteins other than Rep (non-rep components). This observation suggested that Rep protein might specifically recognize targets in the origin sequences common to all BBTV DNA components, indicating that BBTV DNA-1 encodes the ‘master’ viral replicase which was capable of acting on its cognate components to initiate DNA replication of integral BBTV viral components other than Rep.

Geminivirus Rep proteins bind to specific sequences (repeats/iterons) present in their non coding sequences and hydrolyses the phosphodiester bond between the seventh and eighth residues of the invariant nonamer 5'TTATTATAC 3' (Stanley, 1995; Laufs et al., 1995). Sequence analysis and Promoter analysis using TFSearch showed that the BBTV components’ Common Region contained motifs which were probably required for the control of gene expression and replication, notably the conserved TATA-box and poly-A motifs, repeat elements (iterons) and a putative stem–loop structure containing the highly conserved nanonucleotide ‘TTATTATAC’ that functions in the initiation of rolling circle replication.

Observations of the InterProScan from the European Molecular Biology Laboratory domain (URL:www.embl-heidelberg.de/predictprotein/.html), showed that the amino acid sequence of BBTV-C1 possessed distinct viral replication associated and RNA helicase domains and shared homology with other DNA binding, ATPase proteins indicating that this non structural protein possessed a DNA binding and helicase function. Most known DNA binding motifs include α-helical regions that recognize and contact DNA (Pabo et al., 1992). They show homology to the α-helical motifs of basic helix-loop-helix, homeodomain, zinc finger, or basic leucine zipper proteins (Harrison et al., 1991, Nelson et al., 1995). The secondary structure of BBTV
C-1 (as obtained from the program SAINT), showed the presence of 3 helices which were separated by a 5-amino acid loop, and resembled the helix-turn-helix motif, though our sequence comparison failed to uncover a nearby fourth helix characteristic of most helix-turn-helix DNA binding domains (Wintjens et al., 1996). The second set of predicted α-helices is located between BBTV amino acids 131 and 152. Several classes of DNA binding proteins, including members of the basic/helix-loop-helix, homeodomain, and basic/leucine zipper families, use α-helices for dimerization as well as DNA contacts (Murre et al., 1989). The significance of these predicted structures in DNA binding and/or protein interactions is yet to be investigated. Predict Protein analysis also revealed that the immature protein underwent post translational modifications such as N-glycosylation, cAMP- and cGMP-dependent protein kinase phosphorylation, Protein kinase C phosphorylation, Casein kinase II phosphorylation, Tyrosine kinase phosphorylation, N-myristoylation prior to evolving into the quaternary structure of the mature protein (Table 8).

The Rep protein probably binds to a specific recognition sequence within the nucleotide sequence through its DNA-binding domain. It has been shown that the Rep of circovirus PCV binds to the right leg of the stem loop structure, specifically the inner hexamer repeats (5'-CGGCAG-3') in the intergenic region (Mankertz et al., 2001; Steinfeldt et al., 2001). In the case of certain plasmids of the pT181 family, this interaction has been shown to result in structural changes within the sequence, such as DNA bending and cruciform extrusion (Koepsel and Khan, 1985; Jin et al., 1996). This, in turn, exposes the nick site in the DNA that is then cleaved by the Rep protein through its active tyrosine residue. However, in the case of the Rep proteins of the
pLS1 ssDNA plasmid family, there occurs only a transient attachment between the active tyrosine residue and the DNA (del Solar et al., 1998). Based on these observations it is possible that the BBTV Rep recognizes specific iterons in the intergenic region of each component and binds to it. This protein-DNA contact mediates hairpin formation and hence exposes the nanonucleotide loop which is subsequently cleaved by the Rep protein (Plate 22).

During rolling-circle replication, Rep remains bound covalently to the 5'-phosphate end, and the 3'-hydroxyl end thus generated becomes available for rolling-circle replication. After a full cycle of replication, the new origin sequence is generated, which is again hydrolysed by Rep. It is likely that the Rep protein, covalently bound to the 5' end of the DNA, stimulates DNA unwinding through its helicase domain thereby facilitating extension synthesis. Subsequently Rep ligates the nascent 3' end of DNA with the previously generated 5' end. In this way, a unit-genome length, circular, single-stranded DNA molecule, the mature viral genome, is processed (Plate 21, 22).

3.2 Putative function of BBTV component -2

No exact function could be proposed for the proteins encoded by component 2. Nucleotide sequence BLAST analysis revealed homology to other BBTV-2 isolates. Predict Protein analysis revealed that the putative protein possessed sites for the following post translational modifications as N-glycosylation, Protein kinase C phosphorylation and N-myristoylation. The protein prediction servers failed to provide any insight on the probable role of component 2.
3.3 Putative function of BBTV component -3

Both BLAST analysis and Multiple Sequence Analysis of the clone BT-C3 with corresponding sequences of other nanoviruses revealed that it shared maximum identity with BBTV capsid sequence at the nucleotide level (90%) as well as amino acid (92%) levels. Wanitchakorn et al. (1997) showed that DNA-3 encodes a putative viral coat protein. The virus particle is presumably comprised of 60 chemical subunits of a single protein species that made up the icosahedral structure of the coat protein. This protein migrated as a 19 kDa band in SDS-PAGE but had a MW of approximately 20 kDa as deduced from sequence analysis (based on PredictProtein) of the major ORF on DNA-3. Predict Protein Analysis also revealed an array of sites which potentially underwent post-translational modification such as N-glycosylation, cAMP- and cGMP-dependent protein kinase phosphorylation, Protein kinase C phosphorylation, Casein kinase II phosphorylation, Tyrosine kinase phosphorylation, N-myristoylation, Amidation (Table 8). This indicated that the coat protein eventually underwent further modifications before becoming the mature coat protein. Comparative analysis of the complete nucleotide sequences of the coat protein gene (DNA-3) of various strains of BBTV from different regions around the world revealed that the BBTV strains could be divided into two distinct isolates with those from Philippines, Taiwan and Vietnam forming the Asian group while the South Pacific group consisted of isolates from Australia, Burundi and Fiji. At the nucleotide level, the sequence of clone BT-C3 Bg (Bangalore) isolate showed close similarity to the South Pacific group. There was a significantly higher degree of divergence between the Asian isolates, which may indicate that BBTV has been present in this region for an extended period of time or
that there may have been multiple introductions of BBTV into bananas. A close study of the cluster dendrogram (Fig. 13) revealed that at the nucleotide level the sequences of BBTV DNA-3 were more similar between isolates from the same group than between members of the two different groups. At the amino acid level, the BBTV coat protein remained highly conserved, with a maximum of <3% sequence variation between all isolates in this study. The level of high conservation in the BBTV coat protein suggests that attempts to generate transgenic banana plants with resistance to both groups of BBTV isolates would probably be equally effective.

**Putative functions of BBTV- CP**

**Involvement of the BBTV Coat protein in Movement of the viral DNA**

The mechanism by which the viral DNA enters the plant cell nucleus during the early phase of infection is unknown. We suggest that the bunchy top viral CP, which packages the genomic DNA molecule into a viral particle, also transports this DNA into the host cell nucleus. To this end, the CP must itself be karyophilic, possessing a specific nuclear localization signal (NLS) sequence. A comparison of the amino acid sequences of CPs of bipartite geminiviruses with those of known karyophilic proteins (Melchior & Gerace, 1995) revealed a region homologous to the bipartite class of NLSs. This potential NLS, located at the amino-terminus of BBTV-CP, has the following amino acid sequence: 1-MARFPKKS\textbf{IKKRRVGRRKYGSKAA}-24 (see Table 9). Two basic domains of the bipartite signal known to be critical for its function are shown in bold boxed letters. This observation demonstrates that the NLS of the viral CP is capable of translocating the capsid protein across the plant nuclear membranes thereby facilitating nuclear entry of the viral genomes into plant host cell nuclei Kunik et
al.,(1998) have demonstrated TYLCV CP is transported into plant and insect-cell nuclei by an active process of nuclear import via a nuclear localization signal (NLS)-specific pathway.

The CPs of mastreviruses, curtoviruses and monopartite begomoviruses are essential for systemic movement within the plant (Briddon et al., 1989; Liu et al, 1997, 1999,1999a, 2000) whereas the CP of bipartite begomoviruses is not required for transport within the plant, but protects viral DNA during transmission and determines vector specificity (Briddon et al., 1989; Höfer et al., 1997). Since a separate component in BBTV was indicated to function as a movement protein, determining vector specificity and protecting the viral DNA during transmission could probably be attributed to the putative protein of BBTV C-3.

Another interesting aspect of nanovirus-like components was demonstrated by Saunders and Stanley,(1999), who demonstrated that DNA 1 (an autonomously replicating nanovirus-like DNA 1 component) of Ageratum yellow vein virus disease (AYVD), can accumulate in plants when co-inoculated with the Old and New World bipartite begomoviruses ACMV and TGMV, and with the curtovirus BCTV (Mansoor et al., 1999). They showed that DNA 1 and DNA 2 of AYVD evolved from nanovirus components by association with a begomovirus, necessitating an increase in size to enable encapsidation and systemic movement (Mansoor et al., 1999). Encapsulation by the begomovirus coat protein allowed adaptation of DNA 1 and DNA 2 from aphid transmission to whitefly transmission. This raises the intriguing observation that association of the nanoviral DNA to a specific coat protein can determine it’s vector specificity and hence dissemination.
The role of the CP in the establishment of infection is not clear. A study of the effect of mutations in the ORF of Tomato Yellow Leaf Curl Virus-CP showed that it was needed for viral single-stranded DNA accumulation and for successful infection in tomato (Wartig et al. 1997). Similar to TYLCV, the monopartite tomato leaf curl virus (TLCV) from Australia could not accumulate viral single-stranded DNA and spread in the absence of the CP (Ridgen et al. 1993). It should be noted that in all monopartite geminiviruses analyzed to date, the CP is required for viral induction of symptomatic systemic infection (Boulton et al. 1989; Briddon et al. 1989; Ridgen et al. 1993; Woolston et al. 1989). In the case of BBTV the CP too may be responsible for symptom production and spread.

Various modes of Coat Protein-viral DNA interaction during packaging and encapsidation

In most cases, by the late stages of virus infection when assembly of virus particles occurs, transcription of cellular genes has been reduced and large pools of virus genomes have accumulated. The overproduction of viral nucleic acids eases but does not eliminate the problem of specific genome packaging. Viruses also need to differentiate between viral nucleic acid and exclude host cell molecules during encapsidation. Virus must achieve the specificity required to select and encapsidate the virus genome from the large background of cellular nucleic acids. Like many other aspects of virus assembly, the way in which packaging is controlled is not well understood, but the key probably lies in the specific molecular interactions between the genome and the capsid. A conserved zinc finger motif in the coat protein of
Tomato leaf curl Bangalore virus was found responsible for binding preferentially to ssDNA in a sequence non-specific manner (Kirthi & Savithri, 2003).

Fig. 26 3-Dimensional structure representation of virus capsid protein complexed with double stranded region of genome DNA through Zn-finger domains. Pic. Courtesy: 3-D Molecular Designs

Putative nucleic acid binding domains were previously identified within these putative zinc finger regions (Fig. 26) in bromoviruses (Mas and Beachy, 1999). Motif analysis and InterPro analysis of BBTVC-3 amino acid sequences revealed sequence blocks which resembled putative zinc fingers motifs and could possibly harbour DNA binding domains.

Folding the genome in order to stuff it into such a confined space is quite a feat of topology, but is compounded by repulsion by the cumulative negative electrostatic charges on the phosphate groups of the nucleotide backbone resulting in the genome
resisting being crammed into a small space. ssDNA viruses overcome this difficulty by possessing **positively-charged amino acids** with **basic side-chains such as arginine & lysine** (Flint et al., 2000) in the coat protein which interact with the genome order to counteract this negative charge repulsion. Translation of the BBTVC-3 ORF revealed a basic Arg rich, N-terminal region which may be responsible for interacting with the packaged DNA. The other side of the packaging equation is the specific nucleotide sequences in the genome (the packaging signal) which permit the virus to select genomic nucleic acids from the cellular background. In general, viral genomes have packaging signals in their genome to which capsid proteins bind to initiate encapsidation, followed by the progressive assembly of the capsid around the nucleic acid or the insertion of the nucleic acid into the capsid. The **packaging signal** from a number of virus genomes has been identified (Adam et al., 1998; Wei et al., 1990; Alford et al., 2000). In TMV, the genomic RNA harbors an assembly signal sequence at nucleotides 5444–5518(of 6390) that serves as the site to which a small preformed disc of capsid protein attaches (Harrison et al. 1991). Presence of any such assembly signal sequence has yet to be investigated on BBTV-CP sequence.

**Packaging of the Virus**

The BBTV capsids are 20nm in diameter and were identified as possessing an icosahedral structure.

**Fig. 27a. Icosahedral Structure.** All viruses in the family Nanoviridae are known to possess the icosahedral structure.
The coat protein seems to consist of one type of structural protein present as 60 copies (in total) of viral coat protein BBTV C3 in a T=1 icosahedral capsid arrangement (William et al., 2000), although existence of more subunits in the coat protein cannot be ruled out.

Fig. 27b. Model representing the formation of the icosahedral viral capsid structure from each Coat protein subunit.

Icosahedral viruses have a rigidly defined amount of internal space to package nucleic acid and thus their genomes have a predetermined maximum size. Though the exact mechanism of packaging nucleic acid into the capsid is still a topic of debate, two general models have emerged. In the first model the viral nucleic acid associates with single capsid proteins or small preformed capsid subunits. The capsid then builds itself by polymerization around the nucleic acid. The second model requires assembly of nearly complete particles, which then draw the genome inside through a pore or a gap in the particle. A final structural rearrangement occurs in the capsid–protein complex (capsid shell often expands during this process to form the final nucleocapsid structure) to yield a structurally sound virus.
3.4 Putative function of BBTV component 4

Plant viruses have evolved different strategies for cell-to-cell movement. Virus encoded movement proteins participate actively in the intra- and intercellular transport of viral genomes to such an extent that movement protein dysfunction hinders viral infection (Boevink et al., 2005). One of the strategies employed by viruses involves the passage of complexes of the viral genome and movement proteins through the plasmodesmata, the membranous channels formed by prolongations of the ER membranes that interconnect cells in higher plants. Several of these virus-encoded proteins associate and assemble on the Endoplasmic Reticulum membranes (Peremyslov et al., 2004). The ER passes through Plasmodesmata in the form of the desmotubule and is intimately entwined with the actin cytoskeleton; therefore, the simplest and most efficient route for Movement Complexes of these viruses to take to Plasmodesmata from replication centers would be along the ER membrane (Peremyslov et al., 2004) although the path by which they reach the membrane has not yet been explored. TMV MP was found to demonstrate the ability to increase plasmodesmal size exclusion limit (SEL), a process referred to as gating (Wolf et al., 1989), and has been regarded as a key property of classic movement proteins. A hydrophobic core domain was found to be conserved in putative MPs of 17 plant virus genera (Melcher 2000; Mushegian and Koonin 1993). This hydrophobic domain contains a putative transmembrane helix, which is found to be conserved in cellular transport proteins.
BBTV-Component 4 was found to possess varying degrees of amino acid identity to certain 13kDa proteins encoded by SCSV-C1, FBNYV-C4, and MDV-C8 components, respectively (Fig. 16). Based on their similarity to the movement protein of the geminivirus MSV (Boulton et al., 1993; Dickinson et al., 1996), these 13 kDa proteins have been suggested to be movement proteins. Although none of these proteins had significant sequence similarity to the MSV MP, they all possess a hydrophobic domain of 25-30 amino acids in the N-terminal region, which is present in the MSV MP. Protein-Protein Blast, and analysis using Pfam revealed that component 4-ORF encoded a 30 residue hydrophobic domain with the hydrophobic residues Ala-Leu-Ser located in the middle of the sequence, indicating that this ORF encoded a transmembrane protein. The protein encoded by BBTV DNA-4 has been predicted by ExPasy’s Secondary Structure Prediction, and ‘SAINT’, (a secondary structure
prediction program) to have a hydrophobic β-sheet in the N terminus. Other prediction programs as SignalP 3.0 and 'Predotar' (Krogh et al., 2001, Small et al., 2004) which predicted transmembrane topology using a hidden Markov model also predicted the presence of N-terminal targeting sequence. These results suggest that BBTV DNA-4 encodes a putative virus movement protein. And the 29 amino acid N-terminal hydrophobic region of DNA-4 gene product is probably essential for specific localization to the cell periphery although further mutational analysis studies could confirm this.

Valera et al., (2004) demonstrated that the movement protein ‘p6’, of Beet yellows virus is inserted into ER membranes with its C-terminal hydrophilic domain facing the cytosol. Marc et al (2002) reported that CarMV movement protein ‘p9’ had two highly conserved hydrophobic domains forming transmembrane spanning β sheet and proposed a topological model in which CarMV movement protein ‘p9’ is anchored in the membrane with its N- and C-terminal tail segments interacting with its soluble, RNA-bound partner CarMV p7, to accomplish the viral cell-to-cell movement function. The genome RNA was targeted to the membrane through an interaction between p7 and p9.

Pioneering studies on plant viruses revealed that plasmodesmata allow the cell-to-cell trafficking of virally encoded proteins movement proteins (Lucas, 2006). This non-cell-autonomous protein (NCAP) pathway is similarly employed by the host to traffic macromolecules. Viral MP’s bind RNA/DNA in a sequence nonspecific manner to form nucleoprotein complexes (NPC). Host proteins are then involved in the delivery of MPs and NPC to the PD orifice and a role for the cytoskeleton has been
implied. Trafficking of NCAPs through the PD structure involves three steps in which the MP: (a) interacts with a putative PD docking complex, (b) induces dilation in the PD microchannels, and (c) binds to an internal translocation system for delivery into the neighboring cytoplasm. Viral genera that use this NCAP pathway have evolved a combination of a MP and ancillary proteins that work in concert to enable the formation of a stable NPC that can compete with endogenous NCAPs for the PD trafficking machinery. Incompatible MP-host protein interactions may underlie observed tissue tropisms and restricted infection domains. It is therefore likely that such a protein complex may exist during the course of a BBTV infection, which would facilitate cell-to-cell transport of the virus through the plasmodesmata channels.

Several movement proteins are shown to be phosphorylated during the infection process (Kawakami et al., 1999; Matsushita et al. 2000). Sequence analysis studies have shown that the BBTV-MP can be phosphorylated at a number of residues leading to speculation that the phosphorylated MP may be targeted for sequestration (Citovsky, 1993) by the plasmodesma associated MP phosphorylating kinases. These proteins are known to play a critical role in regulating the amount of MP available for viral RNP transport (Karpova et al., 1997).

3.5 Putative function of BBTV component 5

The Rep protein of wheat dwarf virus (WDV), a monopartite geminivirus, has been shown to interact with the human (Xie et al., 1995; Collin et al., 1996) and maize Rb proteins (Xie et al., 1996) through its LXCXE motif and binds specifically to the members of the retinoblastoma (Rb) family of proteins (Sidle, 1996), possibly through interactions with their pocket domain (Xie et al., 1995), (Fig.29). The LXCXE motif, is
conserved in animal DNA virus oncoproteins as well as in a number of geminivirus replication proteins including those from maize streak virus (MSV), (Mullineaux et al., 1984), Digitaria streak virus (DSV; Donson et al., 1987), tobacco yellow dwarf virus (TobYDV; Morris et al., 1992), Panicum streak virus (PSV; Briddon et al., 1992) and Bean yellow dwarf virus (BeYDV, Liu et al., 1999). Unlike the mastreviruses, the nanovirus Rep proteins do not possess a LXCXE motif. This motif is, however, contained in a protein encoded by a separate DNA component now identified as BBTV component 5. Sequence analysis of the BBTV C-5 showed that the protein encoded by this ORF contained an LXCXE motif and shared significant sequence identity with Rb binding-like protein as in Rb-binding proteins of animal DNA tumour viruses. The Rb tumour suppressor is the key regulatory factor of cell cycle progression at the G1/S boundary. Oncoproteins of certain mammalian DNA tumor viruses, such as E1A of adenovirus type 6, E7 of human papillomavirus type 16, or the large T antigen protein of simian virus 40, stimulate the entry of cells into S phase by interaction with pRB (Nagar et al., 1995) and are known to inactivate Retinoblastoma protein by forming a stable complex through the LXCXE motif. This drives the host cell cycle into S phase, creating a cellular environment suitable for replication and transcription of viral DNA. Blast Analysis and ClustalW of the putative protein of BBTV ORF5 indicated high percentage of homology with the 19 kDa proteins encoded by nanoviruses involved in host cell cycle manipulation such as the protein called ‘Clink’ (cell cycle link), a product of Component-10 of the nanovirus, Faba Bean Necrotic Yellow Virus (FBNYV DNA-10), which is capable of binding human Rb and enhancing replication of FBNYV Rep proteins when infected (Aronson et al., 2000). The interaction of the
BBTV-5 protein with the pRB probably abrogates the pRB-mediated block of cell cycle progression and may neutralize additional cell cycle regulators, in particular—the growth suppressor p53 (Neil et al., 1997).

Plate 29. Retinoblastoma-associated protein B domain interacting with a nine residue peptide containing the LXCXE motif [Accession number:PF01857] (Mittnach, 1998)

The crystal structure of the Rb pocket, shows that the LxCxE peptide binds a highly conserved groove on the B domain. The BBTV-C5 protein probably interacts in a similar manner and binds to the AB-pocket of Rb-protein. Retinoblastoma-like and retinoblastoma-associated proteins play a major role in cell cycle regulation. These proteins often act as tumor suppressors, and are potent inhibitors of E2F-mediated trans-activation (Mittnach, 1998; Dyson, 1998; Nevins 1992)

The proposed LXCXE motifs in BBTV were also found surrounded by many acidic amino acids (Fig.16), as is commonly observed in viral and cellular proteins which interact with Rb. In addition, a conserved sequence in the N-terminal region was
identified which bore homology to F-box motifs of F-Box proteins (Patton et al., 1998) (Fig. 16). The F-box, which mediates binding to SKP1, is a conserved domain found in a large number of proteins (Bai et al., 1996) and serves as an adapter to recruit various substrates for degradation through the ubiquitin-proteosome pathway (Hershko et al., 1998; Boyer, 1996). Aronson (2000) demonstrated that Clink, the protein encoded by FBNYV C10 contained distinct functional domains and bound to two cellular proteins, pRB and SKP1. Presence of similar domains on the BBTV-5 protein led us to the conclusion that its interaction with pRB may function by not only disrupting all of the pRB-E2F complexes in the cell, but may also inactivate pRB by targeting its degradation through the ubiquitin-proteosome pathway, thus represent an additional and more efficient mechanism of inactivating pRB. Finally, the BBTV-5 protein itself might be targeted for degradation, thereby restoring (part of) the host cell cycle control. A comparable degradation via ubiquitination of F-box proteins was recently described for yeast Cdc4p and Grrlp (Zhou, 1998) and MEKK protein of Dictyostelium (Chung, 1998).

The amino acid sequence of BBTV5 contained a basic amino acid cluster (R or K) in the N-terminal domain (Fig. 16, Table 9), which closely resembles the bipartite motif, a nuclear targeting signal (rrXXXXXXXXXXXrrrrr) (Dingwall & Laskey, 1991), where r is either R or K and x represents any amino acid residue. The presence of a nuclear targeting signal is also a common feature of tumour virus oncoproteins, and is consistent with the general observation that the Rb family are nuclear proteins (Ach et al., 1997, 1997a). Since BBTV DNA-5 shares similar motifs with FBNYV DNA-10, it
is likely that gene products of these components fulfill similar functions and that the
gene product of BBTV DNA-5 is an Rb-binding-like protein.

Furthermore, Xie et al. (1996) and Grafi et al. (1996) have reported the
isolation of plant Rb-like proteins from maize. These proteins contained a conserved
pocket domain and were able to interact with known Rb-binding proteins as well as
plant D-type cyclins. These findings further confirm the possibility of Rb-mediated
cell-cycle regulation in plant cells and also the ability of some plant viruses to
manipulate the host-cell cycle to enhance their replication. This corroborates the
evidence that BBTV encodes a gene that is capable of influencing host-cell cycling.
Interestingly, it has been shown in certain geminiviruses, that while point mutations
within this motif severely diminish or completely eliminate binding of these proteins to
Rb, these mutations may or may not affect virus replication. The RepA mutants of
BeYDV were shown to retain their ability to replicate and systemically infect Nicotiana
benthamiana and bean plants while replication of WDV mutants in wheat cell
suspensions was significantly reduced or eliminated (Xie et al., 1995; Liu et al., 1999).
Interestingly, the mastreviruses appear to have a different strategy for expressing the
protein required for cell-cycle manipulation and incorporate this function into their
RepA proteins, which have an LXCXE Rb-binding motif, while the begomoviruses
also incorporate this function into Rep although these proteins do not have an LXCXE
motif. The Rep proteins (AC1) of begomoviruses do not have an LXCXE motif, but
can still bind Rb in vitro (Ach et al., 1997). In addition, expression of the viral protein
was sufficient to induce the accumulation of the host DNA synthesis protein,
proliferating cell nuclear antigen (PCNA) (Castillo et al., 2003, Kelman, 1997), in
terminally differentiated cells of tobacco plants, which would suggest a role analogous to that of the animal virus oncoproteins. Hence, these facts may indicate that one of the essential functions of geminivirus Rep proteins is undertaken by the independent small proteins of MDV, FBNYV, SCSV and BBTV which encode proteins that can influence the accumulation of host DNA replication proteins.

Rb-binding proteins are responsible for altering the host cellular environment to facilitate viral DNA replication (Grafi et al., 1996), a function which would require the protein to be expressed early in the virus replication process. This hypothesis is supported by the results of Hafner et al. (1997 b) who demonstrated that BBTV DNA-5 is far more efficiently self-primed than any of the other BBTV DNA components. This suggest that, upon infection, DNA-5 is the first component to be converted to the transcriptionally active dsDNA replicative form (RF) and that the DNA-5 gene product would be produced at high levels in virtually any cell type resulting in the switching of these first infected cells to S-phase, an environment that enhances viral DNA replication.

3.6 Putative function of BBTV component 6

Cell-to-cell or long-distance movement involves a critical interface between the replication apparatus and the transport machinery, including the interactions between MPs and replication proteins or nascent genomes which may initiate the transport process and provide some degree of specificity for trafficking viral nucleic acids. BLAST analysis and Multiple Sequence analysis revealed that the amino acid sequence of BBTV DNA-6 were similar to FBNYV C8, MDV C6 and SCSV C4 and probably encoded equivalent proteins. Translation and CLUSTALW analysis of the putative
protein of component 6 indicated it to belong to the family of nuclear shuttle proteins. Nuclear shuttle proteins in FBNYV were preferentially targeted to the nucleus when expressed alone but, in the presence of the viral movement protein it was found to alter its behaviour by targeting to the cell periphery (Katul et al., 1998; Sano et al., 1998). Initial speculations were that BBTV movement would be similar to that of the monocot-infecting mastreviruses in which the coat protein acts as a nuclear shuttle protein. Instead, the characteristics of the translated proteins encoded by BBTV DNA-4 and DNA-6 seem to be more similar to those of the BL1 (movement) and BR1 (nuclear shuttle) proteins, respectively, of Squash Leaf Curl Virus and other Begomoviruses (Sanderfoot & Lazarowitz, 1995; Noueiry et al., 1994). The bipartite geminiviruses (begomoviruses) appear to encode two non-structural movement proteins, BV1 and BC1 which are not required for replication or encapsidation but are co-operatively essential for cell-to-cell and long-distance movement (Brough et al., 1988, Noueiry et al., 1994). Sanderfoot & Lazarowitz (1995) proposed that BR1 of squash leaf curl virus binds the replicated SqLCV ssDNA in the nucleus and shuttles the complex out to the cytoplasm where BL1 specifically binds the BV1–DNA complexes and directs them to the cell periphery, from where they are transported into the adjoining cells. The BR1 protein contains two basic NLSs within its N-terminal 100 residues and a domain(s) essential for its interaction with BL1 in the C-terminal region demonstrating that BR1(NSP) has a domain structure, with an N-terminal region required for nuclear targeting and a C-terminal region required for its interaction with BL1(Sanderfoot et al., 1996). This nuclear shuttle protein at equilibrium, localizes to nuclei in systemic leaves from infected pumpkin plants and in transfected tobacco
protoplasts (Pascal et al., 1994). However, when this equilibrium is perturbed, BR1 has been shown to shuttle from the nucleus to the cytoplasm when the second movement protein BL1 (responsible for virus cell-to-cell movement) is present (Sanderfoot et al., 1996). When coexpressed with BR1 in transfected tobacco protoplasts, BL1 traps BR1 in the cytoplasm and relocates it from the nucleus to the cell periphery (cortical cytoplasm), which is consistent with its proposed role in moving BR1-genome complexes to and across the cell wall Sanderfoot and Lazarowitz, 1995; Sanderfoot et al., 1996). The requirement of both proteins for intercellular movement was also recently demonstrated for bean dwarf mosaic virus, where mutation of the BV1 (NSP) and BC1 (MP) proteins restricted the cell-to-cell movement of viral DNA (Sudarshana et al., 1998). Results indicate that BBTV may utilize a system analogous to that of the begomoviruses with the BBTV DNA-6 protein acting as a nuclear shuttle protein (NSP) while the DNA-4 protein transports the NSP–DNA complexes to the cell periphery for intercellular transport. Interaction of the two proteins regulates the directionality of intracellular viral DNA transport. This is further supported by the fact that the amino acid sequence of BBTV6 contained residues with homology to other known NES (Table 7) and resembled the NES found in BR1 of SqLCV, HIV Rev, TFIIIA, and several other rapidly shuttling nuclear proteins (Lazarowitz, et al., 1999). As shown in Table 7, the sequence was hydrophobic in character and contained three leucine and one Isoleucine residues. Site-directed mutational studies have also shown that the leucine residues within the NES of HIV Rev and PKI are essential for its function in export (Fischer et al., 1995; Fridell et al., 1996a). Most rapidly shuttling proteins - the retroviral Rev protein, adenovirus E4 34-kD protein, and herpes simplex
virus ICP27, each of which contains a hydrophobic leucine-rich NES, or the influenza virus M1 matrix protein and nucleocapsid protein NP (Lazarowitz, et al., 1999), appears to contain a bidirectional signal (Whittaker et al., 1996) a fact which is not surprising as the replication cycle of each of these viruses involves nuclear events in which the virus must compete for export with host mRNAs and viral transcripts, and/or the viral genome must be exported to the cytoplasm in a regulated manner. Perhaps, analogous to influenza virus NP (Martin and Helenius, 1991) and SqLCV (Sanderfoot et al., 1996; Sanderfoot and Lazarowitz, 1996), the BBTV C-6 protein acts to export the viral genome (single-stranded DNA) from the nucleus to the cytoplasm, where the cooperative interaction of BBTV6 Nuclear Shuttle Protein with BBTV4 Movement Protein1 will direct the BBTV6-NSP-genome complex to the cortical cytoplasm and across the cell wall into adjacent uninfected cells to propagate infection. As in the case of influenza virus NP1-genome complex (Whittaker et al., 1996), the BBTV4-MP interaction with BBTVNSP-ssDNA complex may also act to prevent reimport of NSP-genome complexes into the nucleus. In the case of SqLCV, following interaction with BL1, BR1 targets to the perinuclear region of the cytoplasm and does not reenter the nucleus. Thus, it appears that following release from BL1 in uninfected cells, a post-translational modification of BR1, possibly in its phosphorylation state, may act to prevent reimport of BR1 into the nucleus together with the viral genome (Sanderfoot and Lazarowitz, 1996). However, more studies on biochemical identification and characterization of nuclear export receptors and other components of the export machinery are essential for understanding the details of the import-export mechanism and how this trafficking of macromolecules across NPCs is regulated.
4. Phylogenetic Inference

Analysis of the sequence variability of BBTV-C1 to -C6 reveals that BBTV isolates from more than eight different countries can be separated into two large groups, the South Pacific and the Asian groups. The South Pacific group contains isolates from Australia, Burundi, Egypt, Fiji, India, Tonga and Western Samoa, while the Asian group contains isolates from Vietnam, Philippines and Taiwan. The results suggest that, while the Asian isolates are clearly more closely related to each other than they are to the South Pacific group of isolates, they have undergone considerably more divergence than the South Pacific group. This phenomenon could reflect either the period of time in which these isolates have been evolving in bananas in the Asian region or that there has been more than one introduction of BBTV into bananas in the South East Asian region. The South Pacific group of isolates probably has a more recent common progenitor strengthening the probability that the movement of BBTV through the South Pacific, Australia, South Asia and Africa has been recent and derived from a single source.

There was greater variability within BBTV DNA-1 between the Asian isolates whereas within the South Pacific group of isolates, the CR-M and the ORF were more strongly conserved than the equivalent regions within BBTV DNA-1. The CR-M appears to be an important marker for the division of the two groups as it is strongly conserved within groups of isolates both within and between components though there is between 27% and 32% sequence difference between the Asian and South Pacific groups of isolates.
Protein-Protein BLAST and phylogenies of the nanovirus components supported the hypothesis that faba bean necrotic yellows virus (FBNYV) and milk vetch disease virus (MDV) are sister taxa; that subterranean clover stunt virus (SCSV) branched next; and that banana bunchy top virus (BBTV) is an outgroup to the three other species.

Comparison of the potential Rep proteins identified to date from BBTV, CFDV, FBNYV, MDV and SCSV showed that there is generally a high variability among Rep proteins, not only from distinct nanovirus species but also from a given virus (Fig.22a). High levels of identity (>80%) were observed only between the Rep protein of FBNYV-C7 and that of MDV-C10 (89.8%), and between the C1 Rep protein of FBNYV-Eg and the Rep proteins of MDV-C2 (88.3%) and SCSV-C2 (81.4%). The dendrogram indicated that BBTV Rep was related to CFDV-Rep with 54% identity. Neighbour-joining relationship dendrogram of all known putative Rep proteins identified from nanoviruses, with the Rep protein of porcine circovirus (PCV; accession no. U49186) and geminivirus used as an outgroup sequence (Fig. 22a), reflected the degree of similarity and hence functional equivalence shared by them.

Comparison of 16 Non-Rep proteins identified so far from BBTV, FBNYV, MDV and SCSV (Chu et al., 1993; Burns et al., 1995; Katul et al., 1997; Wanitchakorn et al., 1997; Sano et al., 1998) revealed various levels of identity between (Fig.30) the proteins that have been shown to, or, on the basis of sequence homology, suspected to, have similar functions. The putative 13.4 kDa protein encoded by BBTV-C4 had 49%, 48.3 and 21.8% amino acid identity to those encoded by SCSV-C1, FBNYV-C4, and MDV-C8, respectively. The putative protein encoded by
BBTV-C6 was strikingly similar to the proteins encoded by MDV-C6, SCSV-C4 and FBNV-C8 while the presence of a conserved LXCXE motif in the C-terminal one-third of the BBTV-C5 protein, as well as in its homologues from the other three nanoviruses, SCSV C-3, MVDV C-4, FBNV C-10, suggests that this protein probably shares the same function of interacting with a retinoblastoma (Rb)-like plant protein.

Fig. 30 Dendrogram of protein homologues of BBTV, FBNV, SCSV and MDV.

Highest identities were observed during BLAST and ClustalW Analysis between the putative protein of BBTV-C6 and those of SCSV-C4 (90.8%), FBNV (47.7%) and MDV-C6 (65.8%) (Fig. 30, Plate 16). The similarities in the putative movement proteins and putative Rb-binding proteins encoded by BBTV-C4 and -C5, respectively, and by their corresponding SCSV, MDV and FBNV components, were also highly significant, ranging from 55.1 to 47.7 % between BBTV and SCSV and CFDV, and from 43.8 to 23.6 % between BBTV and FBNV and MDV respectively. The putative protein encoded by BBTV-C3 had a molecular mass of 20 kDa, and
BLAST analysis showed 83.1, 55.5 and 23.7% amino acid identity to the capsid proteins (CPs) encoded by SCSV-C5, FBNYV, and MDV respectively. These results indicate that it is likely that the nanoviruses may share common epitopes in their coat proteins. From the above observation it appears that BBTV is most closely related to SCSV, intermediately related FBNYV and distantly related to MDV (BBTV identities of 23.6%).

The observation that there are very similar counterparts of the putative Rep components from MDV, FBNYV and SCSV in BBTV could indicate that these viruses may have evolved from a common origin. The diversity of their Rep components presumably resulted from recombination events. Based on the overall sequence similarities, BBTV and SCSV are more closely related to each other than MDV and FBNYV. The relatively low sequence similarity of BBTV and SCSV to the other viruses may reflect the geographical or evolutionary isolation of the former from MDV and FBNYV. The phylogeny of replication (Rep) proteins indicate that this small viral multi-gene family has evolved by a process of duplication and subsequent loss of Rep-encoding genome components, analogous to the "birth-and-death" process of evolution which has been described in eukaryotic multi-gene families. By contrast, repeated recombinalional events between components (Hughes, 2004) were found to have homogenized the non-coding portions of several components encoding unrelated components. For example, as result of recent recombination a portion of the non-coding region is virtually identical among SCSV components 1, 3, 4, 5, and 7. Thus, there is a process of concerted evolution of non-coding regions of Nanovirus genome
components, which raises the possibility that certain non-coding regions are subject to functional constraint.

Inferred phylogenetic relationships, based on genome sequence analyses, suggest that BBTV are related to the ssDNA geminiviruses of plants, ssDNA Circoviruses in animals, and appear to form a distinct group. The two known circoviruses, *Porcine circovirus* (PCV) and *Psittacine Beak and Feather Disease circovirus* (BFDV), are similar to nanoviruses in several ways; both have small, icosahedral particles, 17-22 nm in diameter, and small, circular, single-stranded DNA genomes; those of circoviruses are about 2 kb long (Mark & Gibbs, 1999), whereas nanovirus genomes are about 1 kb long. The two kinds of virus encode a replication initiator protein (Rep), with clear similarities between the sequences of these proteins (Fig. 22b). Circoviruses were thought to be a sister-group of nanoviruses, but phylogenetic analyses, which take account of the recombination, indicate that circoviruses evolved from a nanovirus (Meehan et al., 2001). This transferred DNA included the viral origin of replication and the N-terminal region of nanoviral Rep protein. Reps initiate rolling circle replication at a nonanucleotide sequence within a longer origin-of-replication sequence and the nonanucleotide sequences of circoviruses (Mark & Gibbs, 1999) match those of some nanoviruses. The proximity and the similarities between the ori sequences of circoviruses and nanoviruses indicate that these sequences evolved from a common ancestral sequence, and, more importantly, that the nanovirus DNA was transferred from a plant to a vertebrate.

Numerous geminivirus particles and non-geminivirus molecules of similar size to nanoviruses have been found in geminivirus-infected plants. Saunders &
Stanley,(1999) and Frischmuth et al.,(2001), have proposed that the diseases caused by nanoviruses are an accumulation of these geminivirus-associated molecules, which have become independent of the geminivirus for disease causation. Furthermore, interfamilial recombination between nano- and geminiviruses has been suggested from sequence analysis of ageratum yellow vein virus-associated molecules. Therefore, geminiviruses might even represent recombinant nanoviruses.

5. Promoter analysis

The transcription elements required for strong promoter activity were located within 239 bp of the translation start codon. Promoter and Transcription Factor Analysis using Tfsearch and Promoter search, two analysis tools which predicted transcription factor binding sites as well as detected promoter sequences, revealed that this region contained the TATA box. The sequence ‘GATAAG’ identified by Motif Analysis in the components showed homology to the I-box and G-box cis-elements during BLAST analysis and have been associated with strong expression in plant systems. The G-box core has been shown to interact with several plant transcription factors (reviewed in Katagiri & Chu, 1992), and is responsive to plant hormones such as abscisic acid, jasmonate and ethylene. The location of the I-box consensus with respect to the G-box, and in relation to the downstream TATA-box, implies that these motifs identified in the BBTV components may be functionally equivalent to those found in plant systems and probably involved in promoter activity. Furthermore, a 56 bp region, located 3' of the common stem-loop structure, also contributed to promoter activity. This region contained a 10 bp sequence, CATGACGTCA, which has strong homology with the 3' end of the 20 bp consensus incorporating the ocs element.
(Bouchez et al., 1989) of the Agrobacterium tumefaciens T-DNA promoters. This region also contains a sequence, ACGTCA, with homology to the hexamer motif of plant histone promoters (Mikami et al., 1987). The presence of this motif in the genomes of several other ssDNA plant viruses has been suggested to reflect an evolutionary conservation of transcription control mechanisms within this group (Morozov et al., 1994). The hexamer motif ‘ACGTCA’ (Nakayama et al., 1992) is most likely associated with strong promoter activity in undifferentiated, actively dividing cell types, as histone genes are expressed specifically in the ‘S’ phase of the cell cycle. The promoter for the CFDV ‘master Rep’ protein has been shown to possess phloem specific attributes (Rohde et al., 1995). Comparison of the BBTV Rep promoter indicated several similarities in sequence organization and the high incidence of the virus in banana phloem tissue could probably be attributed to this occurrence. This theory is supported by studies by Magee (1970), who reported disorganization of the phloem elements and surrounding parenchyma cells in banana at the point of initial BBTV infection, with subsequent hypertrophy and rapid cell division. However, certain cis-acting elements were also named to be involved in phloem-specificity in plant viruses (Hehn & Rohde, 1998).

6. Molecular Mechanisms of Pathogenesis and Replication of BBTV

The biochemical events involved in BBTV replication have not been determined. Available information from other nanoviruses suggests that replication is similar to that of geminiviruses in being completely dependent on the host cell's DNA replication enzymes, and occurring in the nucleus through double-stranded DNA intermediates by a rolling circle replication mechanism (Laufs et al., 1995; Bisaro,
Rolling circle replication (RCR, Plate 23) of DNA is characteristic of viruses possessing circular, single-stranded DNA genomes or single-stranded DNA (ssDNA) intermediates in their replication cycles (Ilyina & Koonin, 1992). Strategy of geminivirus DNA replication resembles that of prokaryotic ssDNA replicons (Baas and Janzs, 1988; Novick, 1998). Genetic entities that multiply their DNA via RCR, range from ssDNA plasmids of Archaeabacteria and Eubacteria, ssDNA phages (e.g., fX174), ssDNA viruses of plants (geminiviruses) to circoviruses and ssDNA viruses of birds and mammals (Saunders et al., 1999; Stenger et al., 2000).

The virus-encoded replication proteins involved in RCR of ssDNA perform ATP-dependent strand nicking-closing and topoisomerase I activities. Viral DNA replication is initiated by the ‘Master-Rep’ protein that interacts with common sequence signals on all the genomic DNAs (Timchenko et al., 1999, 2000; Horser et al., 2001). During RCR the Rep protein cleaves the consensus nonamer sequence TAGTATT\(_{1}AC\) located at the origin of replication (Timchenko et al., 1999), creating a 3′-OH terminus and, by analogy to geminivirus replication, is thought to prime viral (plus) strand DNA synthesis. This observation is consistent with the presence of conserved repeat sequences or iterons in the intergenic region as well as a similar nanonucleotide motif at the apex of the stem loop region in all BBTV genome components. The intergenic stem-loop structure in the BBTV DNA molecule was found to bear a striking similarity to that of the geminiviruses and the ori hairpin of single-stranded bacterial plasmids (Gruss & Ehrlich, 1989). This conserved loop sequence (Plate 21) probably contains the nick site where the replicating molecule becomes covalently attached to the active site tyrosine in motif 3 of the replication
protein thus forming the origin of initiation for the operation of a rolling circle model for DNA replication.

Thus summarizing the sequence of events, upon inoculation of the bunchy top virus into the host banana plant, the nuclear localisation signal in the N-terminus of the coat protein probably ensures that the virus is translocated to the nucleus. This is presumably followed by decapsidation of viral ssDNA, one of the first biosynthetic events is the conversion of the ssDNA genome into a dsDNA intermediate product (Saunders et al., 2000). Replication of the linear genome of BBTV would then involve the use of the semi-stable genomic 3'-hairpin structure as a primer, synthesis of new DNA by host delta DNA pol using genomic DNA as template, covalent closure then re-opening of the circular DNA product, and re-initiation at the new 3'-terminus with the latter as primer after "melting" of the newly-formed duplex and transient hairpin formation at the new 3' end (Plate 21). The other circular genomes replicate by a "rolling circle" model: nicking of one strand of the ccDNA RF allows use of an exposed 3'-terminus as a primer, which results in elongation of the "primer" strand, and displacement of the other. Ordinarily this results in greater-than- unit-length ss- and dsDNA forms being found in infected cells. The dsDNA forms now serves as the transcription template for the production of a small number of viral proteins, one of which acts as a replication initiator protein. Binding of the Replication protein may distort double stranded origin (dso) of the BBTV genome DNA, thereby exposing a single-stranded region containing nick site (Plate 22). The structural changes in may facilitate binding of host replication proteins to the dso. Alternatively, the Rep protein recognizes the (ori) origin and binds to a specific DNA sequence. It then catalyzes
DNA cleavage and ligation to initiate and terminate rolling circle replication, and acts as a DNA helicase during the elongation phase of replication. These functions were demonstrated by Hanley-Bowdoin et al. (1999) in the AL-1Rep proteins of geminiviruses. However the size limitation of the genome ensures that the virus transcribes only the factors required to initiate rolling circle replication and depends on plant nuclear DNA polymerases such as DNA primase, α-like and δ-like DNA polymerases, whose activities have been identified in plants (Coello and Vazquez-Ramos, 1995a; Garcia et al., 1997) to amplify their genomes.

Initiation of DNA replication during the rolling circle phase, requires the concerted action of the viral Rep protein and Rb binding like viral protein with cellular factors, and leads to the production of new dsDNA and ssDNA viral forms (Stenger et al., 2000; Heyraud et al., 1993; Stanley, 1995). AL1-rep protein also plays a central role in reprogramming mature plant cells to support DNA replication and recruiting host replication machinery to the origin (Kong, et al., 2000; Gutierrez et al., 2002). In nanoviruses, replication of the DNA components by cellular enzymes, is facilitated and enhanced by the action of an Rb-like binding protein, a nanovirus-encoded cell cycle modulator protein (Aronson et al., 2000).

**BBTV replication is triggered by a master Rep protein.**

One of the key steps during the initiation of DNA replication is origin recognition. Nanoviral Rep proteins contained several conserved motifs, (Motif I, II and III, Fig.9), which was found conserved in vast class of proteins involved in binding, initiation and termination of rolling circle DNA replication (RCR) Rep proteins.
The origin of DNA replication in geminiviruses has a modular architecture (Fontes et al., 1994a; Sanz-Burgos and Gutierrez, 1998) and falls into two major categories in geminiviruses: the mastrevirus-type origin, e.g. wheat dwarf virus (WDV; Kammann et al., 1991; Hofer et al., 1992; Schneider et al., 1992; Sanz-Burgos and Gutierrez, 1998), which consists of a large cis-acting region where the initiator Rep protein forms multiple complexes (Castellano et al., 1999), and the begomovirus-type origin, e.g. tobacco golden mosaic virus, containing one binding site for Rep (Fontes et al., 1992, 1994a,b; Orozco and Hanley-Bowdoin, 1998). The concept of a modular arrangement of specificity elements and a common initiation signal, recognized and acted on by Rep proteins, in a two-step process is easily transferable from the bipartite genome of some geminiviruses to the multipartite genome of the nanoviruses. In geminivirus infections, both +ve and −ve DNA strands are transcribed. By contrast, nanovirus transcription has been found to be (typically) unidirectional, with each genome component encoding only a single protein. The master replication protein thus binds to repeated sequences near the ori and, in combination with binding of host-encoded DnaA protein, melts an AT-rich region which contains the origin. Replication that follows, involves usual leading and lagging strand synthesis (Khan, 1999)

A common strategy of DNA viruses to overcome the limitation of a small genome, is to encode proteins that interact with components of the host transcriptional apparatus and cell cycle regulatory network (Nevins, 1992), to create an environment favorable for efficient replication of their genome. Viruses with small genomes do so by subverting the cell cycle control of the host and forcing cells into DNA synthesis or S phase and exploiting the DNA synthesis of the host. They encode proteins that target
basal transcription factors (Gruda et al., 1993), the histone transacetylase p300 (Eckner et al., 1994), and the tumor suppressors, pRb or pRb like proteins (Dyson et al., 1992) and p53 (Werness et al., 1990). These protein interactions cause quiescent cells to re-enter the cell division cycle and synthesize the enzymes necessary for viral DNA replication. The discovery that geminivirus AL1 protein also induces accumulation of host DNA replication machinery in mature cells (Nagar et al., 1995) and binds the plant pRb homologue pRBR (Ach et al., 1997a) suggested that geminiviruses use similar mechanisms to reprogram their plant hosts. Upon infection, double-stranded replicative intermediates are synthesized in the nuclei of infected cells and serve as transcription templates for viral gene expression. From these dsDNAs (RF) forms, host RNA polymerase would transcribe mRNAs encoding other proteins such as the movement and nuclear shuttle and coat proteins required for virus movement.

In animal cells, the human tumour suppressor retinoblastoma (pRB) protein, together with other members of the so-called 'pocket family' (p107 and p130), facilitate differentiation and negatively regulate the passage of cells through the G1 phase and the G1-S transit of the cell cycle by modulating the activity of the E2F-DP family of transcription factors (Weinberg, 1995). Plant pRb homologues (pRBR, retinoblastoma-related) have been cloned from maize (Grafi et al., 1996; Xie et al., 1996; Ach et al., 1997a), tobacco (Nakagami et al., 1999), Chenopodium and Arabidopsis (Fountain et al., 1999). Experiments showing that a maize pRb homologue is preferentially expressed in mature leaf tissue (Huntley et al., 1998) suggest that pRBR proteins may serve similar functions in plants. The pRb family members of plants and animals display strong sequence homology across a large central domain known as the A/B
pocket (Murray, 1997) and is involved in a variety of protein interactions by way of a conserved LXCXE motif (Moran, 1993; Dyson et al., 1992). A sophisticated mechanism allows the sequential phosphorylation of RB by CDK-cyclin complexes leading to the release of RB-bound E2F-DP factors needed to activate transcription of genes required for the G1-S transition and S-phase progression in eukaryotic systems (Mittnacht, 1998; Sherr and Roberts, 1999). Animal oncoviruses are able to bypass the normal RB control pathway in G1 by the action of one oncovirus-encoded protein whose LxCxE amino acid motif mediates binding to RB (Ludlow, 1993; Vousden, 1993). The finding that a bunchy top, component 5 viral protein, contains an LxCxE motif and was similar to Rb like binding protein, was a first clue as to which mechanism could be used by the bunchy top virus to induce a permissive cellular state. Lazarowitz et al., (1992) pointed out that bunchy top virus DNA replication seemed to be excluded from the actively dividing regions of the plant, called meristems. BBTV does not appear to be strictly phloem restricted, and have been detected outside the vascular system, in differentiated cells (Rushing et al., 1998; Nagar et al., 1995).

One possibility is that the virus does not necessarily induce cell proliferation but, rather, initiates a cellular state in which S-phase functions are up-regulated. This can be mediated through interference with the RBR pathway by allowing E2F-dependent activation of S-phase-specific genes. The presence of this Rb-binding protein would be expected to result in the switching of these first infected cells to S-phase, an environment that enhances viral DNA replication. This observation together with the identification of plant D-type cyclins (Dahl et al., 1995; Soni et al., 1995), which also contain an LxCxE motif, provides strong support for the notion that a plant
RB-related (RBR) pathway might exist in plants. In this context, it would appear that BBTV proteins actually induce dedifferentiation and re-entry into the cell cycle with a concomitant passage through S-phase.

**Regulation of Cell proliferation during viral replication**

ssDNA plant viruses probably encode functions to restore normal cell cycle control in the host cells, once a critical amount of viral genome products has accumulated, since neither geminiviruses nor nanoviruses cause uncontrolled cell proliferation. Participation of proteolysis has been described for different steps of cell cycle regulation (Townsley, 1998). An important core component, amongst others, of a particularly versatile class of ubiquitin-ligases is SKP1, which assemble with different F-box proteins (Koepp, et al., 1999, Krek et al., 1998). The F-box motif found in amino acid sequence of BBTV-5 has a conserved domain found in a large number of proteins and mediates binding to SKP1 (Bai, et al., 1996). The F-box serves as substrate-specific adapter subunits to recruit various substrates to a core ubiquitination complex; proteins once ubiquitinized are destined for degradation by the 26S proteasome (Bai, et al., 1996). The FBNYV 'Clink' protein was found to interact with an SKP1 homologue in plants, mediating the degradation of proteins in a later phase of the cell cycle and of virus infection (Aronson et al., 2000). Binding to Clink may render SKP1 inaccessible for other proteins involved in the control of mitosis. If this were the case, cell divisions would be inhibited and viral DNA replication might be achieved during a process of endoreduplication. In yeast, F-box proteins and the ubiquitin-proteasome pathway are involved in the control of the ploidy level (Kominami et al., 1997). The F-Box motif recognized in the amino acid sequence of BBTV-5 encoding the putative Rb-binding
protein itself might be targeted for degradation, thereby restoring part of the host cell cycle control.

**Fig 31.** A simplified representation of the Nanovirus DNA replication, viral gene expression and intercellular movement of viral DNA.

BBTV DNA replication occurs in two stages. First, the ssDNA is converted into dsDNA with the participation of cellular factors. The dsDNA serves as template for viral gene expression. Secondly, the dsDNA initiates the rolling circle phase, with the participation of viral and cellular factors, to produce new ssDNA products. These can (i) re-enter the DNA replication pool, (ii) associate with CP or (iii) be transported outside the nucleus and to the neighbouring cell, most probably through plasmodesmata, with the help of viral MPs.

**Multipartite nature of BBTV Genome and effect on Transmission**

Unlike the ssDNA plasmids, phages, and circoviruses, the genetic information of the nanoviruses is distributed over at least six different DNAs and the proteins are expressed separately from each of them. Each ssDNA is individually encapsidated
into small isometric virions, and virus transmission is accomplished only by insects (aphids). In this case infection occurs only when the particles having the divided genomes infect cells simultaneously.

7. Development of Diagnostic Assays

Micro propagated banana plants need to be indexed for BBTV and other viral diseases to detect and eliminate symptomless, yet infected plants from accidental multiplication. An effective and reliable diagnostic test for BBTV virus is thus essential for surveillance and monitoring, which are critical for implementation of disease control strategies. ELISA and Immunosorbent electron microscopy (ISEM) have been developed for the three most widespread viruses of banana, banana bunchy top virus, cucumber mosaic virus and banana streak virus (Lockhart and Olszewski, 1996; Ndowora and Lockhart, 1998). However, ELISA’s lack sufficient sensitivity to detect low-titer infections, and at times produce ambiguous results due to the occurrence of serological diversity among virus isolates. ISEM’s are complicated by labour, cost, time, equipment constraints and inability to index large number of samples (Mollina, 1998; Morgan, 1995).

With the objective of rapid and accurate diagnosis in consideration, two molecular based detection assays namely BBTV PCR Assay and BBTV SLOT/Blot Hybridization Assay were developed and standardized for simple routine detection assays. The DNA-based diagnostic tests developed during this study can return an accurate result in as little as 24 hours, a factor integral for timely containment and management of disease outbreaks. The PCR-based molecular assay was sensitive and capable of detecting BBTV from samples as little as 80ng banana leaf tissue, and was
found to be the most sensitive assay when compared to Dot Blot hybridization and ELISA while Dot Blot assays were in turn more sensitive than ELISAs. This parameter could be critical for accurate and timely detection of disease outbreaks so that disease control strategies as can be implemented as soon as possible to limit their spread. PCR assays were also used to confirm samples that were inconclusive in ELISA and Dot Blot tests. However, molecular methods such as PCR are not adequate for mass screening as they are highly dependent on availability of pure DNA without inhibitors, and are an expensive option.

For routine tests of large number of samples, Dot Blot assay was found to be most reliable, convenient and economically viable, as the probe and most hybridization chemicals could be re-used several times. The Digoxigen (DIG) labeled probes synthesized against BBTV ORF 1, BBTV ORF3 and BBTV ORF 5 were routinely used in Southern Blots, Dot or Slot Blots, and could be utilized for a period upto 1 year, when stored at 4°C. Assay studies revealed that the probe, when used at $10^6$ dilution could detect viral DNA from a $10^8$ dilution. Radioisotope labeled probes used earlier required special procedures and facilities for handling of radiochemicals which frequently have short half lives. Such facilities are expensive and at times difficult to obtain. Furthermore, radioactive decay limited the duration of a probe that could be used. By comparison, the non-radioactive probes developed did away with the problems associated with handling radioactive probes. Non-isotopic probes are more stable than isotopic probes, and have better signal-to-noise ratios: they also eliminated disposal problems associated with radioactive probes. In addition, hybridization times could be reduced by several hours by increasing the amount of probe used.
Nitrocellulose membranes are easy to handle with non radioactive detection systems. DIG probes have other several advantages over radioactive probes: the nucleotide probes possess long shelf-life since there is no radioactive decay. Chemiluminescent detection eliminates time-consuming handling, monitoring and clean up and the need to designate special "hot" areas on lab. Using chemiluminescent substrates, detection can be completed with a 15- to 30-min exposure to X-ray film, comparable to 1-4 days required for most $^{32}$P-labelled probes. A further advantage is that probes can be stored at 4°C or -20°C for many years with no apparent loss in sensitivity. Further, for the CDP-Star chemiluminescent substrate, Dioxetane, the light output reaches a maximum at 4 hours but then lasts for several days. This allows for multiple, long exposures whereas other chemiluminescent systems, utilizing HRP, only last for about 6-8 hours. Luminescent detection is fast and sensitive, and membranes can be easily stripped and reprobed for existence of other viruses in the same sample/Blot.

A Dot Blot Protocol was also standardized for detection of BBTV directly from the Plant sap using the above DIG labeled probes. Use of plant sap in Dot Blots even precluded the cumbersome, time consuming and expensive DNA extraction procedures. Both molecular diagnostic methods ensure detection at the nanogram level; even before the onset of symptoms, when viral titers are low. Application of virus indexing on mother stocks (prior to propagation of tissue culture plantlets) based on these molecular diagnostic probes, would minimize the risk of inadvertent propagation of virus infected plants in tissue culture and facilitate phytosanitary safety during germplasm movement, as well as ensure quality control. Although the DIG labeled BBTV component-3 probe produced equally efficient results, it was predominantly
used for verification analysis during recombinant Coat protein experimental studies. DIG-labelled BBTV Component 1 and 5 probes exhibited maximum efficiency in terms of detection as both components were found to be present in detectable concentrations, irrespective of the stage of pathogenesis. The probes are being routinely used for BBTV detection of sample (both infected and suspected) banana plants, infected tissue culture plants, and for indexing of both tissue cultured export samples as well as mother stocks prior to micropropagation by tissue culture.

Benefits to the industry due to availability of DNA-based diagnostic tools include:

- Rapid, definite and accurate diagnosis of Banana Bunchy Top disease ensures ‘clean’ planting material and limits the spread of the disease.
- Improved capability to determine the distribution of banana viral disease following an incursion, e.g. during eradication campaigns
- Improved capability to define the distribution of endemic banana diseases, e.g. when an inter-farm movement of the virus need to be prevented
- Ability to index and handle larger volumes of samples in a relatively shorter period

8. Cloning and Expression of Coat protein gene for Antibody Production

Antibodies raised against regions of the viral coat-protein epitopes have been exploited for virus detection especially in ELISA which enables rapid testing of large number of plant sap samples for the presence of virus. Serological testing requires large quantities of antisera in order to index candidate mother stock which in turn requires the production of large amounts of antigens for immunization purposes.
These antigens were produced by labor-intensive, recurrent purifications of the virus, increased in herbaceous hosts. The techniques used for their production and maintenance were also time consuming and expensive.

The technique of studying pathogens by cloning fragments of specific genes in bacterial cells by Mutasa-Gottgens et al., (2000) was utilized and BBTV viral coat protein was successfully cloned into a protein expression vector. Expression of the gene of interest in a heterologous system as bacteria enabled production of the antigen when required. The availability of cloned viral gene led to the production of purified proteins and greatly simplified the development of standard diagnostic assays. The recombinant vector when allowed to multiply in the host bacterial cell with an appropriate inductor ensured sufficient yields of the protein. With the yields optimized, the recombinant proteins could be used as (viral) antigens to evoke an antibody response by sensitizing in rabbits. A litre of bacterial culture yielded about 8-14mg of BBTV-CP, an amount sufficient for the sensitization of at least two rabbits. In addition, frozen stocks of the recombinant vector could be revived within 24 hours whenever the need to generate fresh antisera arose for antibody generation. This was in stark contrast to the lengthy procedure involved and time required (4-5 days) for the classical BBTV purification methods, which often gave poor or no yields before the virus could be used for inoculation; additionally, need for all expensive chemicals and reagents could be circumvented. We can thus conclude that this technology bears great potential for the serological diagnosis of BBTV virus in plants, and that producing antibodies against BBTV using a recombinant antigen is an advantageous procedure compared to the existing labor-intensive and expensive
procedures. Polyclonal and monoclonal antibodies raised against BBTV particles would permit sensitive serological detection of BBTV infected plants, when used in Western Blots, Double Antibody Sandwich (DAS)-ELISA, triple antibody sandwich (TAS-ELISA) or Tissue Blot ImmunoAssay (TBIA). These methods represent an additional tool for efficient BBTV-indexing of banana propagative and mother stock materials, and also for use in support of biological and molecular techniques.


Disease control is the largest single component in the cost of production of dessert banana and plantain in the tropics (Diekmann et al., 2001). Engineering resistance to banana bunchy top virus (BBTV) seems an obvious objective for banana transformation, since no natural resistance to this virus has been identified in the entire Musa gene pool.

The strategy based on Pathogen Derived Resistance was adopted for genetic improvement of Cavendish bananas. The transgene used in the present study was the BBTV Coat Protein gene and BBTV replicase gene for constructing viable plasmid vectors for future development of virus resistant transgenic plants. The transgene undergoes transcription and translation, resulting in high levels of protein. The protein inhibits disassembly of the infecting virus and forces the assembly/disassembly equilibrium towards assembly.

Availability of information regarding ‘Viral Coat Protein’ and ‘Replicase’ from the present study provided the impetus to construct two chimeric vectors using Banana Bunchy Top Viral genes for conferring resistance to Bunchy Top Virus by
Agrobacterium mediated transformation. The high level of conservation in the BBTV Replicase protein and coat protein suggested that attempts to generate transgenic banana plants with resistance to both Asian and Pacific groups of BBTV isolates would probably be effective.

Resistance induced by CP gene is thought to occur through protein-mediation, since a single copy of the transgene was inserted. Resistance so expressed is of moderate level against a broad range of related viruses. The transgene undergoes transcription and translation, resulting in high levels of proteins. The protein inhibits disassembly of the infecting virus and forces the equilibrium towards assembly. CPMR may also, in many virus–host combinations be caused by post transcriptional gene silencing (PTGS). Plants containing translatable coat protein constructs produced either complete resistance, or delayed infection followed by recovery. Cassidy and Nelson in 1995 showed that transgenic plants containing a coat protein transgene of Peanut stripe virus, for example, exhibited either resistance or delayed symptom development following inoculation.

Replicase-mediated resistance is complex and impedes the ability of the virus to infect plants by interfering with both virus replication and movement (Nguyen et al., 1996). TMV replicase-mediated resistance was shown to be protein-mediated in protoplasts (Carr et al., 1992) while research by Nejidat and Beachy (1990) found that transgenic tobacco plants containing a gene encoding the TMV 54 kDa protein exhibited transgene silencing. The silencing was shown to contribute to virus resistance. The availability of effective vector systems is a prerequisite for genetic manipulation of plants through recombinant DNA technology. The vector pCambia
2301 from the CAMBIA vector series (www.cambia.com), Centre for Application of Molecular Biology to International Agriculture) contained minimal heterologous sequences for plant transformation and selection of transformants and was provided with two different CaMV35S-driven and terminated plant selection genes: hptII encoding resistance to hygromycin and nptII encoding resistance to kanamycin. These genes have been subjected to site-directed mutagenesis to eliminate interfering restriction sites within the coding sequence by silent changes. Two different bacterial resistance markers were provided (kanamycin or chloramphenicol), allowing a broader range of Agrobacterium or E. coli strains to be used. The pUC18 poly linker within the lac Z fragment allows for blue/white screening of clones in E.coli cloning work. The initial pCambia2301 vector lacked the promoter and terminator sequences for plant expression of newly cloned genes. This was achieved by inserting the target (viral) gene into a plant expression cassette which was later restriction digested out and cloned in-frame into the plant expression binary vector pCambia 2301. The expression cassette present in the cloning vector pUC19 as pUC.IVS.Ubi-CaMV35polyA was a kind gift from Dr. Ashwath, Department of Ornamental Crops, IIHR. The pUC19 expression cassette comprised of an enhancer called IVS, the Maize polyubiquitin promoter, a LacZ multiple cloning site and the CamV35S polyA termination signal. Constructs incorporating maize Ubi-1 promoter would be useful for generating high-level gene expression of selectable markers and facilitate expression of biotechnologically important protein products in transgenic monocotyledonous plants.
Fig. 32 Digramatic Representation of the plant expression cassette bearing vector pUC19.IVS.Ubi.CaMVPolyA. The gene of interest, (BBTV Rep and BBTV CP) was cloned into the Nhel/SpeI site of the lacZ multiple cloning site of the above vector.

(a)  
```
<table>
<thead>
<tr>
<th>IVS enhancer</th>
<th>Ubi promoter</th>
<th>lac Z MCS</th>
<th>CaMV PolyA terminator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant expression cassette in pUC19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+ BBTV Rep ORF
```

(b)  
```
Ligation
```

Expression cassette region of plasmid pUC19.IVS.Ubi.BBTV-Rep.CaMVPolyA

(c)  
```
Ligation +
```

pCAMBIA 2301 vector construct

- CaMV35S promoter
- Lac Z alpha
- Multiple Cloning Site
- CaMV35S polyA
- Hygromycin R
- Plant selection gene
- pBR322 ori
- Kanamycin (R)
- Bacterial selection gene
- pBR322 bom
- pVS1 sta
- pVS1 rep
- T-Border (left)
- T-Border (right)
- NheI (729)
- SpeI (1,15)
- HphI (8)
- FasI (752)
- EcoRI (762)
- NcoI poly-A

Hygromycin R __

Pl^il setecled gene /

Bacterial

selection gene /
The constructed pCAM.IVS.BBTV vectors possessed the following characteristics. The chimeric vectors, pCAM.BBTV-CP and pCAM.BBTV-Rep, harbour a 520 bp fragment of the coat protein gene and an 890bp fragment of the Replicase gene respectively shown as red boxed in Fig 32 (a-d, above). Both genes are under the control of the Maize polyubiquitin promoter (Christensen et al, 1992), one of the strongest promoters tested so far in banana (Sagi et al, 1995a) and is flanked by the CaMV35SpolyA terminator. The GUS gene is present as a reporter which is used to assess the efficiency of transformation. The MCS is between gusA (proximal to the right border of T-DNA transfer) and Hygromycin (proximal to the left border). Thus transformed plants expressing GUS and showing hygromycin resistance would be the ones having the cloned genes. Additionally the intron present in the gusA coding sequence ensures that the gene is not expressed in bacteria (Ohta et al, 1990) but only transferred to plants. The lacZ gene facilitates blue/white screening while the pBR322 ori enables high copy replication in E.coli. and a broad host range ori for low copy,
stable replication in *Agrobacterium*. Finally, a hexahistidine ‘tail’ has been engineered at the carboxyl end of GUS, allowing for protein purification by immobilized metal ion affinity chromatography. The strong constitutive cauliflower mosaic virus (CaMV) 35S promoter, the nopaline synthase (nos) poly(A) and CaMV35S-PolyA regulatory sequence regions would be used for transcription initiation and termination of the GUS reporter and hygromycin antibiotic resistance gene respectively. The nptII genes would confer resistance to kanamycin in transformed plants.

The plant transformation vector created enables direct transformation by both particle bombardment and *Agrobacterium*-mediated transformation, precluding the use of binary vectors. The vectors are stable in *Agrobacterium* even if grown under non-selective conditions because of the presence of the ‘rep’ and ‘sta’ regions from pVS1 (Deblaere et al, 1994). *Agrobacterium*-mediated transformation of banana meristem offers remarkable advantages over direct gene transfer methodologies. The main attractiveness of this method is the ease and speed of transgenic plant regeneration, and the lack of requirement for specialized equipment or sophisticated *in vitro* regeneration skills. In addition, it reduces the copy numbers of the transgene, potentially leading to fewer problems with transgene co-suppression and instability (Gheysen et al., 1998; Hansen, 1999; Shibata, 2000).

The development of effective vector systems expressing BBTV transgenes would be a major advantage for future transformation studies on currently accepted cultivars of Banana and plantain and holds promise of generating Banana Bunchy Top virus resistant plants, thereby helping find an effective, durable and sustained control of a major disease of this crop.
The present study has helped to reveal information about genome organization, replication and infection process of Banana Bunchy Top Virus within its host. With the use of recombinant DNA technology, the complete viral genome was determined, its genes/protein coding regions analyzed and putative functions of the proteins have been predicted using various computational tools.

This study has also helped to generate oligonucleotide primers, which proved an integral component in subsequent PCR reactions. Sequencing and sequence analysis of the various amplified and cloned genes proved to be helpful in identifying the possible functions of the different viral proteins and thus helped in understanding the mode of replication of the virus in its host system.

The two molecular based detection assays namely BBTV PCR assay and BBTV SLOT/Blot assay were developed based on the molecular information gained about the virus. The specificity, accuracy and unambiguity associated with these assays would minimize the risk of inadvertent propagation of virus infected plants in tissue culture which would facilitate phytosanitary safety during germplasm movement and ensure quality control. It is believed that this diagnostic test could prove as an efficient commercial tool for virus indexing in the plant biotech industry concentrating on the micropropagation of commercially viable banana plants.

In order to augment conventional breeding and to avoid constraints imposed by pests and pathogens, transgenic approaches were considered. A chimeric plant transformation vector was constructed using BBT viral genes, which could potentially
be used for *Agrobacterium*-mediated transformation of banana. The availability of transformation systems that can complement conventional breeding programs is of great importance for the improvement of *Musa* spp. and could, in the near future be used for Banana and plantain transformation via *Agrobacterium*-mediated transformation to eventually confer resistance to Banana Bunchy Top Virus.

An attempt was also made to express and purify the BBTV coat protein for future antibody generation. This was achieved by cloning the viral coat protein of BBTV into pRSET protein expression vector. Expression of the coat protein gene was induced using IPTG for enhanced yet controlled expression. The recombinant protein was finally affinity chromatography purified using Ni-NTA agarose. The BBTV coat protein could be used as an immunogen for the generation of both polyclonal as well as monoclonal antibodies. The Polyclonal antibodies prepared against bunchy top viral proteins expressed in *E. coli.* would serve as serological probes which are relatively inexpensive to produce and negate the need to obtain, propagate, and purify virions for use as an immunogen. Producing antibodies against BBTV, using a recombinant antigen is an advantageous procedure compared to the labor-intensive virus increase in herbaceous hosts and the cumbersome and expensive virus purification, and bears great potential for the serological diagnosis of viruses in plants.

Overall results of these investigations comprising the genomic analysis of BBTV, development of molecular markers for viral indexing and experiments using *Agrobacterium* mediated transformation would certainly be helpful in the control and management of banana bunchy top virus infection thus minimizing huge economic losses in agriculture.