3 Biotech (2021) 11:471
https://doi.org/10.1007/s13205-021-03017-x

ORIGINAL ARTICLE

Restriction-free cloning for molecular manipulation and augmented expression of banana bunchy top viral coat protein

Darsana Dilip[1] · Vimi Louis[2] · H. S. Savithri[3] · P. M. Namitha[2]

Received: 25 May 2021 / Accepted: 5 October 2021 / Published online: 22 October 2021

Abstract

Banana bunchy top virus (BBTV) causing bunchy top disease, is one of the most devastating diseases of banana and plantain. All the six genomic components of isolates from different parts of the world have been well characterised, with most of the studies focusing on replicase gene and coat protein gene. Overexpression of coat protein (CP) in *Escherichia coli* system can contribute significantly in structural as well as immunological studies. In the present investigation, the full length BBTV CP was cloned to pGEX-4T-2 expression vector and overexpressed in various *Escherichia coli* strains to obtain high quality and quantity of the CP. An augmented overexpression and stability of recombinant coat protein was achieved by molecular manipulation of the clone by restriction-free (RF) cloning platform. The RF cloning was employed to replace the thrombin cleavage site in the vector backbone, which was also present in the protein of interest, and to incorporate TEV protease site to cleave fusion protein at this specific site, and separate the affinity tag. The RF method allows direct transformation of the PCR product to undergo ligation in vivo and obtain the transformants thereby avoiding the restriction digestion and ligation of the product to the linearized plasmid. From a litre culture, 1.084 mg/ml of fusion protein with GST tag was obtained after GSH sepharose affinity column chromatography. The fluorescence spectra indicated partial disordered tertiary structure of the fusion protein. Cleavage of tag was attempted using TEV protease overexpressed and purified in the laboratory.

Keywords *Banana bunchy top virus* · Coat protein · Restriction-free cloning · Fluorescence spectroscopy

Introduction

Banana bunchy top virus (BBTV) causing banana bunchy top disease (BBTD) is designated as one of the 100 World's Worst Invasive Alien Species by Invasive Species Specialist Group (ISSG) of the World Conservation Union-International Union for conservation of nature (IUCN) (Lowe et al. 2000). It was first reported from Fiji in 1889 which was then reported in almost all banana growing areas in the Old World within a very short span of time (Magee 1927, 1953; Wardlaw 1961). At present, BBTD occurs in 35 Old World countries *viz.,* Africa, Asia, and Oceania and in Hawaii (Conant 1992; Kumar et al. 2011; Blomme et al. 2013). It causes, narrowing and yellowing of apex leaves which become brittle and bunch together. These leaves turn necrotic at a later stage. The infected plants become dwarf and bears malformed fruits or no fruits at all (Elayabalan et al. 2015; Kumar et al. 2014). BBTV was detected to spread through vegetative propagule, rhizome and even tissue culture plants and is also transmitted persistently by banana aphid, *Pentalonia nigronervosa* Coquerel (Drew et al. 1989; Anhalt and Almeida 2008).

BBTV is a multipartite virus with six 1.1 kb single-stranded (ss) DNA components encapsidated separately by its coat protein to form 18 nm diameter icosahedral particles (Harding et al. 1991). Initially, BBTV was included in luteovirus group (King et al. 2012) and later reclassified to genus *Babuvirus,* in *Nanoviridae* along with *Cardamom bushy dwarf virus* and *Abaca mosaic virus* (Mandal et al. 2004; Sharman et al. 2008). The genomic components of BBTV are designated as DNA-R, U3, S, M, C and N.

✉ Vimi Louis
vimi.louis@kau.in

[1] Department of Plant Pathology, College of Agriculture, Kerala Agricultural University, Thrissur, Kerala 680656, India

[2] Division of Plant Pathology, Banana Research Station, Kannara, Kerala Agricultural University, Thrissur, Kerala 680652, India

[3] Department of Biochemistry, Indian Institute of Science, New Biological Sciences Building, Bangalore 560012, India

DNA-R encodes for master replicase protein, with additional helicase, endonuclease and ligase domain aiding in rolling circle replication of all the genome components. DNA-S encode for the coat protein which encapsulate the genomic components. DNA-M codes for the movement protein that facilitates the cell to cell movement of the virus whereas, DNA-C, encodes the cell cycle link protein that alters the cell cycle to favour the replication of the virus in the differentiated cells of the host. This protein is similar to retinoblastoma like protein that helps in prolonging S phase of the cell cycle. DNA-N codes for the nuclear shuttle protein that transports the replicated genome to the cytoplasm for translation. Function of DNA-U3 is unknown (Burns et al. 1995; Dale et al. 1986; Harding et al. 1993; Wanitchakorn et al. 1997, 2000; Beetham et al. 1999; Vetten et al. 2005).

Based on the genomic characterisation, BBTV isolates are classified into South East Asian (SEA) group and Pacific Indian Oceans (PIO) group (Karan et al. 1994). The former comprises of isolates from China, Indonesia, Japan, Philippines, Taiwan and Vietnam whereas the latter includes isolates from Australia, Egypt, Hawaii, India, Myanmar, Pakistan, Sri Lanka and Tonga (Karan et al. 1994; Yu et al. 2012).

Hitherto, control of the disease has not been possible by conventional breeding as the source of resistance has not been traced. Various attempts of Pathogen derived resistance (PDR) targeting coat protein and replicase has been successful in laboratory but was proved to be ineffective in glass house or field conditions (Elayabalan and Kalaimughilan 2013). Attempts were made to produce transgenic plants that can interrupt viral replication or translation by RNA interference technology targeting DNA-R (Kumar et al. 2014; Tsao 2000; Shekhawat et al. 2012; Krishna et al. 2013; Elayabalan et al. 2017). However, acceptability and consistency of transgenic plants is debatable (Elayabalan and Kalaimughilan 2013). Consequently, early detection and supplying virus free planting materials to the stake holders is the most feasible measure to control the disease. Among the various serological and nucleic acid based diagnostic methods, Enzyme-linked immunosorbent assay (ELISA) is most preferred. Specific and sensitive serological assay can be developed only if quality antiserum raised against pure antigen is available (Hema et al. 2003).

The 20 kDa CP coded by DNA-S, apart from protecting the genome, assists in cell to cell movement of the virion to maintain the viral load in the phloem tissues, facilitating the survival of the host for a longer time. Reports also suggested the role of coat protein as silencing suppressors (Niu et al. 2009). The structure of BBTV CP may perhaps have some implications on its multifunctionality; however, the structure of BBTV CP has not been solved to date. Characteristically, multifunctional CPs are targeted for diagnosis due to its immunogenicity. Recombinant CP expressed in various hosts, have innumerable advantages especially in developing serological assays for detection of viruses and in structural biology (Rosano and Ceccarelli 2014; Schumann and Ferreira 2004). Attempts to express BBTV CP protein as fusion protein have been made earlier (Wanitchakorn et al. 1997). The current study aims at developing BBTV CP without affinity tag which can be used for immunisation and structural studies. Cleaving off the affinity tags are sometimes complicated due to various reasons and also expensive based on the enzymes used. Owing to restriction-free (RF) cloning process, we have tried to manipulate the protease cleavage site of the expression vector to facilitate the removal of the tag using a very specific TEV protease, without hindering the solubility of the protein.

Materials and methods

Media and bacteriological reagents were procured from Hi-Media, Mumbai, India. General biochemicals and accessories such as isopropyl-β-D-thiogalactopyranoside (IPTG), Nitrocellulose membrane, ECL solution, antibiotics, etc. used in this study were obtained from Sigma-Aldrich USA. GSH sepharose beads and thrombin were purchased from Novagen. The oligonucleotides were synthesised from Eurofins, Bangalore. DNA modifying enzymes and restriction enzymes were obtained from ThermoFisher and New England Biolabs. The bacterial strains used in the study are *Escherichia coli* DH5α, *E.coli* strains BL21 (DE3) pLysS, Rosetta (DE3) pLysS, C41 and the vectors were purchased from Invitrogen.

Sample collection

A purposive sampling survey was conducted in ten districts representing Northern, Central and Southern zones of Kerala. Young leaves of banana showing banana bunchy top symptoms were collected from different varieties grown in various location of Kerala.

DNA extraction, PCR and sequencing

Total DNA was isolated from five representative samples using modified CTAB method. In the presence of liquid nitrogen, 1 g leaf sample with midrib was pulverized to fine powder and homogenised with 10 ml pre heated 2X extraction buffer (2% CTAB, 100 mM Tris, pH 8.0, 1.4 M NaCl, 20 mM EDTA and 0.1% β-mercaptoethanol). The mixture was transferred to autoclaved Oakridge tube and incubated at 65 °C in water bath giving intermittent mixing. To the Oakridge tube, 10 ml of 24:1 v/v chloroform isoamyl alcohol was added and mixed in an end to end rotor for 10 min. Post centrifugation at 10,000 rpm for 15 min at 4 °C, the aqueous

phase was transferred to another Oakridge tube. To the aqueous phase, equal volume of ice cold isopropanol was added and after mixing thoroughly it was incubated at − 20 °C for 1 h. The precipitated DNA was collected by centrifugation at 10,000 rpm for 20 min at 4 °C. The pellet was washed with 70% ice cold ethanol and was blot dried. Pellet was dissolved in water and quantified in Nanodrop (ThermoFischer Scientific). Quality of DNA was also checked in 0.8% agarose gel.

DNA was used as template for polymerase chain reaction (PCR) to amplify coat protein (CP) gene using designed forward primer and reverse primers specific to CP gene (Table 1). EcoR1 Nhe1 restriction sites in the forward and BamH1 site in the reverse primer were incorporated (underlined). PCR reaction mixture contained 1X HF buffer, 10 mM dNTPs, 10 pmol each of BBTV CP specific forward and reverse primers, 6% dimethyl sulfoxide (DMSO), 2 units of phusion polymerase (New England Biolabs) and 50 ng of template. PCR was performed in a Thermal cycler (Biorad, USA) under the following conditions: 98 °C for 1 min, followed by 30 cycles of 98 °C for 30 s, 38.8 °C for 30 s, 72 °C for 25 s with a final extension for 10 min at 72 °C. The PCR products were analysed on 1.2% agarose gel and purified by PCR purification kit according to manufacturer's instructions (ThermoFisher Scientific). The PCR products were sequenced by Sanger dideoxy sequencing method.

Cloning of BBTV coat protein to pGEX-4T-2

The purified PCR product was gel eluted using GeneJET gel extraction kit (ThermoFisher Scientific) as per manufacturer's protocol. It was cloned to pUC19 cloning vector at EcoRI and BamHI site and transformed to *E. coli* DH5α by heat shock method for amplification. Confirmation of clone was carried out by restriction digestion of the plasmid isolated from the transformed colonies with EcoRI and BamHI (ThermoFisher Scientific) and sequencing.

The BBTV CP gene was amplified from BBTV CP/pUC 19 clone and the purified PCR product was cloned to SmaI site of pGEX-4T-2 using T4 DNA Ligase. The ligated mix was transformed to *E. coli* DH5α and plasmids were isolated from randomly selected colonies. The recombinant clones were screened by colony PCR with gene specific primers and restriction digestion. The orientation of the inserted gene was confirmed by BamHI digestion resulting in release of the gene (Supplementary Fig. 2). Plasmid isolation was performed by alkaline lysis method with slight modifications (Green and Sambrook 2016) and subjected to PCR for further confirmation. The clone was designated as pGEX/CP. After confirmation the recombinant BBTV CP/ pGEX-4T-2 plasmid was transformed to *E. coli* BL21(DE3) pLysS, Rosetta (DE3) pLysS and C41 strains for overexpression.

Restriction-free cloning of BBTV coat protein to pGEX-4T-2

Primers were designed for restriction-free (RF) cloning of BBTV CP to pGEX-4T-2 to eliminate thrombin cleavage site and introduce TEV protease cleavage site (underlined) between the target protein and glutathione-*S*-transferase (GST) tag (Table 1). Forward primer was designed with C terminus sequence of GST protein followed by TEV protease cleavage site and a few N terminal amino acid sequences of BBTV CP gene. The primers were used to amplify coat protein gene from total genomic DNA of BBTD infected banana. Reaction mix was prepared as mentioned earlier and PCR was performed in a Thermo cycler (Biorad, USA) with initial denaturation at 98 °C for 1 min followed by 30 cycles of denaturation at 98 °C for 30 s, annealing at 61.3 °C for 30 s and extension at 72 °C for 25 s with a final extension of 72 °C for 10 min. The PCR product was separated on 0.8% agarose gel and eluted using GenJet gel extraction kit (ThermoFischer Scientific) based on manufacturer's protocol. The purified PCR product was used as megaprimer for RF amplification of pGEX/CP clone which was used as template. 50 µl of PCR mix was prepared with 1X HF buffer, 10 mM dNTPs, 1 µg megaprimer, 6% dimethyl sulfoxide (DMSO), 2.5 units of phusion polymerase (New England Biolabs) and 500 ng of template. Sterile MilliQ water was added instead of megaprimers in a negative control. The RF reaction was as follows: a single denaturation step (95 °C, 30 s) was performed followed by 35 cycles of: denaturation (95 °C, 30 s),

Table 1 Details of primers used in this study for cloning to various vectors

Details	Forward primer	Reverse primer
BBTV CP specific primer for PCR	5′ GGG GAATTC GCTAGC ATG GCT AGG TAT CCG 3′	5′ CCC GGA TCC TCA AAC ATG ATA TG 3′
Megaprimers for RF cloning of BBTV CP gene to pGEX-4T-2	5′**C CAT CCT CCA AAA TCG GAT** GAG AAC CTG TAC TTC CAA GGT CAT ATG GCT AGG TAT CCG AAG 3′	5′TCAAAC ATGATATGT AAT TCTGTT CTGG 3′

annealing (55 °C, 1 min), and elongation (68 °C, 5 min) and a final elongation step of 7 min at 72 °C.

Following the RF reaction, the parental DNA strand was eliminated by Dpn1 digestion. 9 µl from both RF and control reactions were aliquoted to 0.6 µl Eppendorf tubes and 1 µl of Dpn1 (New England Biolabs) was added and incubation at 37 °C for 2 h. The enzyme was heat inactivated at 80 °C for 20 min on heat block after which it was transformed into DH5α strain of *E. coli*. Plasmid isolation from randomly selected white colonies was done using alkali lysis method. PCR using gene specific primers was performed to confirmed and designated as ΔpGEX/CP. It was transformed to BL21(DE3) pLysS, Rosetta (DE3) pLysS and C41 strains of *E. coli* for overexpression.

Overexpression of BBTV CP

For expression of BBTV recombinant coat protein (rCP), 1% overnight grown culture of three *E. coli* strains BL21 (DE3) pLysS, Rosetta pLysS and C41 harbouring pGEX/CP and ΔpGEX/ CP were inoculated separately to 50 ml Luria Bertani (LB) broth. The LB broth was supplemented with 50 µg/ml ampicillin and 50 µg/ml chloramphenicol to culture and overexpress recombinant BBTV CP in *E. coli* BL21 (DE3) pLysS and Rosetta pLysS but only 50 µg/ml ampicillin was added to culture and overexpress it in C41. The culture was grown at room temperature in shaker incubator for 4 h and induced with 1 mM isopropyl thiogalacto pyranoside (IPTG) and incubated at 16 °C for 12 h with continuous shaking. Cells were harvested by low speed centrifugation at 8000 rpm for 15 min at 4 °C. The supernatants were discarded and the pellets were resuspended in 50 mM Tris, pH 8.0, containing 200 mM NaCl and sonicated in ice bath using a Vibra Cell sonicator at 35% amplitude for 1 min. The soluble and insoluble fractions were analysed on 12% SDS-PAGE. Recombinant protein expressed from pGEX/CP was designated as rCP and that expressed from ΔpGEX/CP was designated as ΔrCP.

Purification of recombinant coat protein and quantification

Both rCP and ΔrCP clones were overexpressed in *E. coli* strains Rosetta pLysS and C41, respectively, grown in 1 l LB broth. The culture at log phase was induced with 0.5 mM IPTG and incubated at 16 °C for 12 h with constant shaking at 100 rpm. Cells were harvested by centrifugation at 8000 rpm for 20 min at 4 °C and resuspended in 40 ml of 50 mM Tris (pH 8.0) containing 200 mM NaCl, 1 mM EDTA (pH 8.0), 10% glycerol and 5 mM β-mercaptoethanol. The cells were lysed by sonication of two 5 min cycles and two 3 min cycles giving 6 s interval between 3 s pulse of 30% amplitude on ice bath. The debris was separated from crude lysate by centrifugation at 8000 rpm 4 °C for 30 min. 1 ml calibrated GSH sepharose beads were allowed to bind to the soluble fraction for 5 h at 4 °C in an end to end rotor. Column was packed with protein resin complex and washed off the unbound proteins with 10X bed volume of wash buffer (50 mM Tris, pH 8.0 and 200 mM NaCl) twice. Both rCP and ΔrCP were eluted with elution buffer containing 50 mM Tris, pH 8.0, 200 mM NaCl and 30 mM glutathione. The eluted fractions were pooled and dialysed against 500 ml of dialysis buffer (50 mM Tris; pH 8.0 containing 200 mM NaCl and 10% glycerol) to remove glutathione. The dialysed protein was quantified by Bradford's assay.

Purification of TEV protease for cleaving fusion protein

Recombinant plasmid containing TEV protease gene was obtained from IISER, Pune. It was transformed to *E. coli* BL21 (DE3) by heat shock method. Single colony from overnight grown culture of *E. coli* BL21 (DE3) with the recombinant plasmid containing TEV protease gene was inoculated to 25 ml of LB broth with 50 µg/ml of ampicillin. 1 ml of the culture at exponential growth phase was inoculated to 1 l LB containing 50 µg/ml of ampicillin and was induced using 0.5 mM IPTG at log phase. The culture was incubated at 25 °C overnight. The pellet was harvested by centrifugation at 8000 rpm, 4 °C for 20 min and resuspended in 30 ml lysis buffer (50 mM Tris; pH 8.0, 200 mM NaCl, 10% glycerol) and sonicated with 30% amplitude (3 s on 6 s off cycle) for 5 min. Two more sonication cycles of 3 min each was repeated after an interval of 5 min. Subsequently, the culture was centrifuged at 8000 rpm for 40 min at 4 °C and the supernatant was collected. The lysate was allowed to bind to 1 ml calibrated Ni^{2+}-NTA (nickel-nitrilotriacetic acid) beads (Millipore) in an end to end rotor for 6 h at 4 °C. The lysate–resin mix was loaded to the column, and after removing the flowthrough, beads were successively washed with Tris NaCl buffer (pH 8.0) containing 10 mM, 25 mM and 30 mM Imidazole to remove the unbound and weakly bound proteins. For elution, 50 mM Tris buffer of pH 8.0, containing 200 mM NaCl and 100 mM imidazole was used. Eluted samples were subjected to SDS-PAGE to confirm the presence of TEV protease at around 27 kDa. The respective fractions containing the pure protein was pooled and dialysed to remove imidazole and concentrated as per requirement.

Cleavage of fusion protein

Digestion of BBTV fusion protein (ΔrCP) expressed from ΔpGEX/CP was done using TEV protease to cleave off the GST tag. Digestion using 2.0, 1.0, 0.5 and 0.25 µg TEV

protease to obtain the tag less protein was attempted at 20 °C for 12 h and was separated by SDS-PAGE.

Characterisation of BBTV CP by fluorescence spectroscopy

The intrinsic fluorescence ΔrCP was measured in TECAN LS55 Luminescence Spectrometer (Perkin-Elmer). Spectra were recorded at protein concentrations of 0.1–0.5 µg in buffer containing 50 mM Tris, pH 8.0 and 200 mM NaCl. Blank was set with the same buffer without the fusion protein. The excitation wavelength was 280 nm and the emission was scanned between 300 and 400 nm.

In silico analysis of fusion protein

The sequence of fusion proteins (pGEX/CP and ΔpGEX/CP) was translated by ExPASY translate tool (Gasteiger et al. 2003). Protease cleavage sites in the amino acid sequence were identified using Peptide Cutter (Gasteiger et al. 2005). The ordered and disordered regions in the protein were determined by FoldIndex analysis (Prilusky et al. 2005).

Results

PCR amplification and cloning of BBTV CP gene

Primers specific to BBTV CP gene were designed manually and PCR reaction was carried out as mentioned in the methods section. Annealing temperature was standardised experimentally by gradient PCR as 38.8 °C. The expected amplicon length was 537 bp, corresponding to the complete coding region of BBTV CP gene (513 bp) and the extra nucleotides from the primers for restriction cleavage. An amplicon of ~ 540 bp was obtained after PCR (Supplementary Fig. 1). Non-specific amplifications and primer dimers were absent. The BBTV CP gene was initially cloned to pUC19 cloning vector at EcoR1 and BamH1 sites. The gene was also inserted into SmaI site of pGEX-4T-2 expression vector. The clones transformed to *E. coli* DH5α were confirmed by PCR, and restriction digestion of the plasmid (Supplementary Fig. 2). Amplicon of ~ 540 bp size was obtained after PCR using CP specific primers with pUC19/CP and pGEX/CP as templates. BamH1 recognition site was present before the SmaI site where the gene was cloned and the 3′ terminal of the gene had the same site incorporated to it. Hence, restriction digestion of pGEX/CP with BamH1 released insert of ~ 540 bp (Supplementary Fig. 2b).

RF cloning of BBTV CP

Restriction-free cloning of BBTV CP gene to pGEX-4T-2 was carried out to replace thrombin cleavage site between GST tag and the coat protein with TEV protease cleavage site. Using specific primers designed for RF cloning, BBTV CP gene was amplified from total genomic DNA. The 550 bp PCR product obtained was used as megaprimer for RF reaction (Fig. 1a). The complete recombinant pGEX plasmid with BBTV CP gene (pGEX/CP) was amplified using this megaprimer to obtain 5.6 kb product (Fig. 1b) which was transformed directly into DH5α cells after DpnI treatment for in vivo ligation. The confirmation of the recombinant clones after RF cloning was done by restriction digestion with BamHI. The BamHI recognition site upstream to GST tag was removed during RF cloning and hence ΔpGEX/CP was linearised (result not shown here). In contrast to the previous observation where, BamHI digestion released the insert from pGEX/CP clone (Supplementary Fig. 2b).

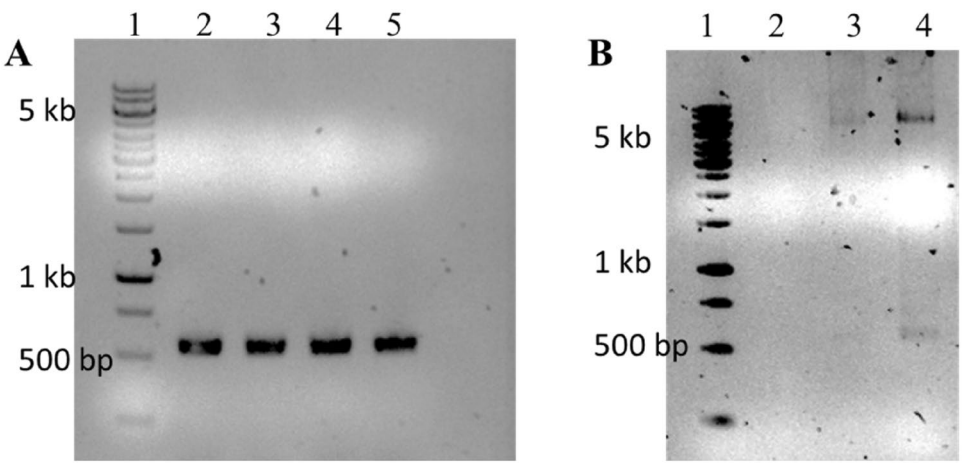

Fig. 1 Elimination of thrombin cleavage site and introduction of TEV protease using Restriction-free platform. **a** PCR amplification of BBTV CP gene for insertion of vector sequence and TEV protease cleavage site. **b** RF cloning of BBTV CP to pGEX-4T-2 with megaprimer

Overexpression and purification of BBTV rCP and ΔrCP

The pGEX-4T-2/CP clone was expressed in *E. coli* BL21 (DE3) pLysS, Rosetta pLysS and C41 bacterial systems, showed a band at ~ 45 kDa (25 kDa GST + 19 kDa CP) (Fig. 2). The expression of protein was induced by 1 mM IPTG at 16 °C and 30 °C. The overexpression of BBTV rCP was equally expressed in all the cell lines and was highly soluble at 16 ºC in LB broth. Hence for purification, Rosetta pLysS harbouring pGEX/CP was cultured. A very thick band at 25 kDa corresponding to GST protein was observed in the SDS-PAGE profile in the induced fraction indicating proteolysis between the target protein and the affinity tag. Nevertheless, when induction was carried out at 30 °C, degradation of protein was observed in SDS-PAGE (result not shown here).

The ΔpGEX/CP clone obtained after RF reaction was also expressed in all the three cell lines (Fig. 3). A band at 45 kDa was observed only in the induced fractions in SDS-PAGE profile. The expression was highest in *E. coli* C41 at 16 °C when induced with 0.5 mM IPTG. Therefore, ΔpGEX/CP was overexpressed in C41 and cultured in LB medium for purification. On comparing, augmented expression of BBTV CP was observed from ΔpGEX/CP clone in C41 indicating reduction in proteolysis of fusion protein.

Fusion proteins, rCP and ΔrCP, were purified by GSH sepharose column chromatography as described in the methods section. The soluble and insoluble fractions after sonication, along with flowthrough, wash and elutions were loaded on to 12% SDS gel and electrophorized. As observed earlier, the fusion protein (rCP) expressed at 45 kDa was seen to undergo proteolysis and separate as GST tag and the CP protein. However, after purification, the protein was seen to be unstable and only bands at 25 kDa corresponding to the GST tag was observed in SDS PAGE profile (Supplementary Fig. 3). On analysing the probable protease recognition sites

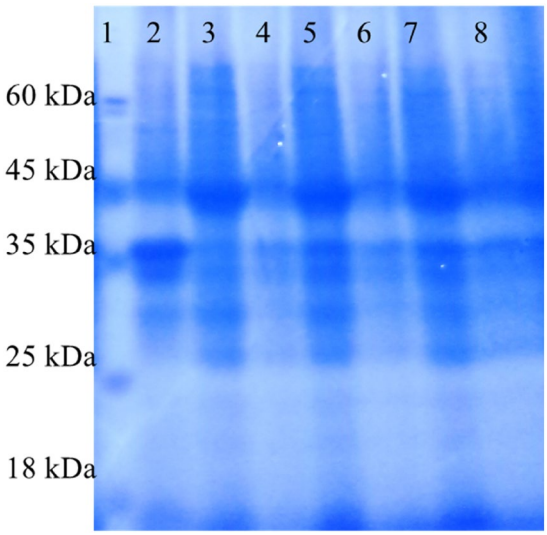

Fig. 3 Overexpression of BBTV ΔpGEX/CP. lane 1: Protein molecular mass marker, lane 2: and 3: insoluble and soluble fractions of induced BBTV CP in C41, lane 4: cell lysate of uninduced C41 harbouring ΔpGEX/CP, lane 5 and 6: induced and uninduced BBTV CP in BL21 (DE3) pLysS, lane 7 and 8: induced and uninduced BBTV CP in Rosetta pLysS

on the protein using Peptide Cutter, it was observed that the arginine and lysine rich N terminus of BBTV CP was prone to trypsin digestion (Supplementary Fig. 4). There have been reports of GST protein also contributing to degradation reasons of which are unidentified.

The ΔrCP was purified and high-intensity band was obtained at 45 kDa (Fig. 4). The elutions were pooled, dialysed and quantified by Bradford's assay. High concentration (5 ml ΔrCP of concentration 1.084 mg/ml) of pure protein was obtained from 1 l culture. Cleavage of tag was attempted using TEV protease, overexpressed and purified in-house (Supplementary Fig. 5), recognition site of which was incorporated through RF platform. The rCP without the tag was

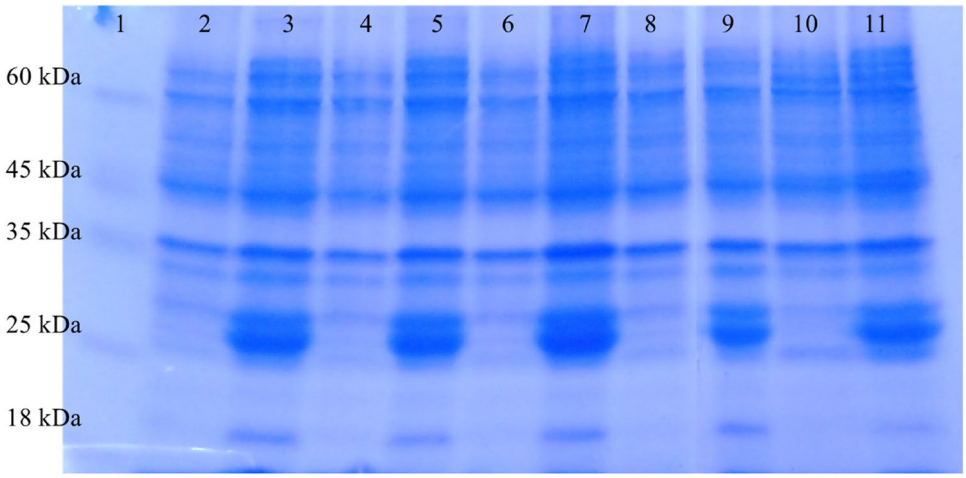

Fig. 2 Overexpression of BBTV pGEX/CP. lane 1: Protein molecular mass marker, lane 2 and 3: uninduced and induced BBTV CP in Rosetta pLysS, lane 4 and 5: uninduced and induced BBTV CP in BL21 (DE3) pLysS, lane 6 and 7: uninduced and induced BBTV CP in C41

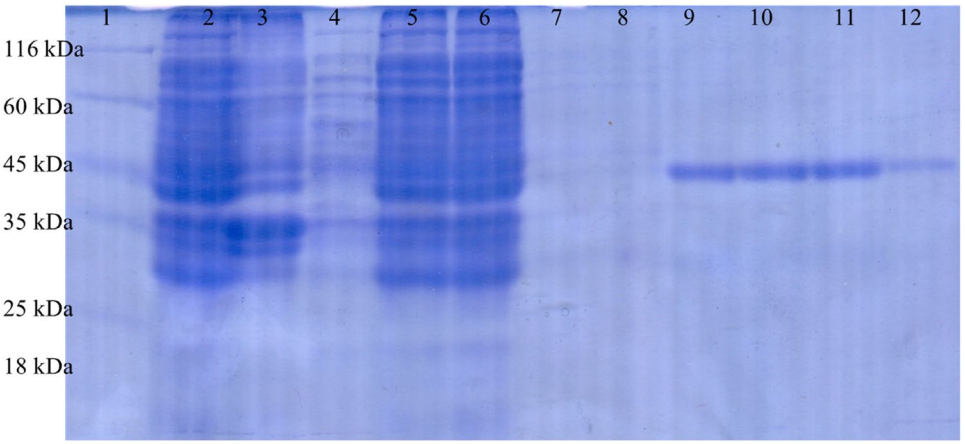

Fig. 4 Purification of BBTV CP (ΔrCP) overexpressed in C41. Lane 1: Protein molecular weight marker, lane 2: supernatant, lane3: pellet, lane 4: flowthrough, lane 5–6: washes, lane 7–12: elutions of BBTV CP fusion protein

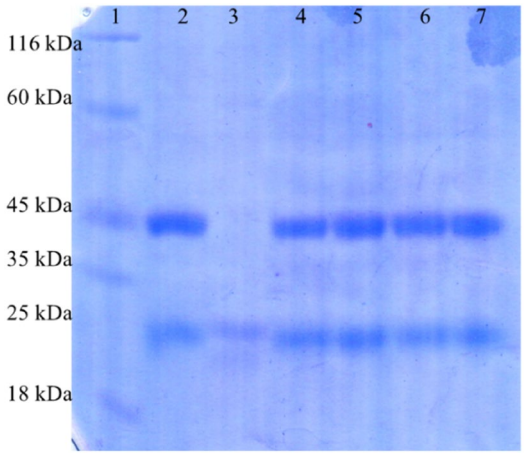

Fig. 5 Cleavage of BBTV CP fusion protein (ΔrCP) with TEV protease. Lane1: Protein molecular weight marker, lane2: BBTV CP-GST fusion protein, lane3: GST protein, BBTV CP fusion protein incubated with, lane 4: 2.0 µg, lane5: 1.0 µg, lane6: 0.5 µg, lane7: 0.25 µg purified TEV protease at 20 °C for 12 h

unstable above 4 °C (inferred from previous experiment) and TEV protease exhibits a reduction in its activity below 20 °C. Hence cleavage of fusion protein was unsuccessful (Fig. 5). In the present study, increasing the amount of TEV protease also did not aid to resolve this issue.

From the FoldIndex analysis, it is clear that both GST protein and BBTV CP are ordered except for the N terminus of CP (Supplementary Fig. 6). The TEV protease site lies in the disordered region in between the two highly ordered structures.

Characterization by fluorescence spectroscopy

Fluorescence spectrum of tagged BBTV CP was measured between 300 and 400 nm. There were two peaks one at 325 nm and 345 nm (Fig. 6). However, the intensity was less. From the spectrum of ΔrCP it can be deduced that protein is partially folded with some disordered region. This corroborates the FoldIndex analysis wherein the N terminal

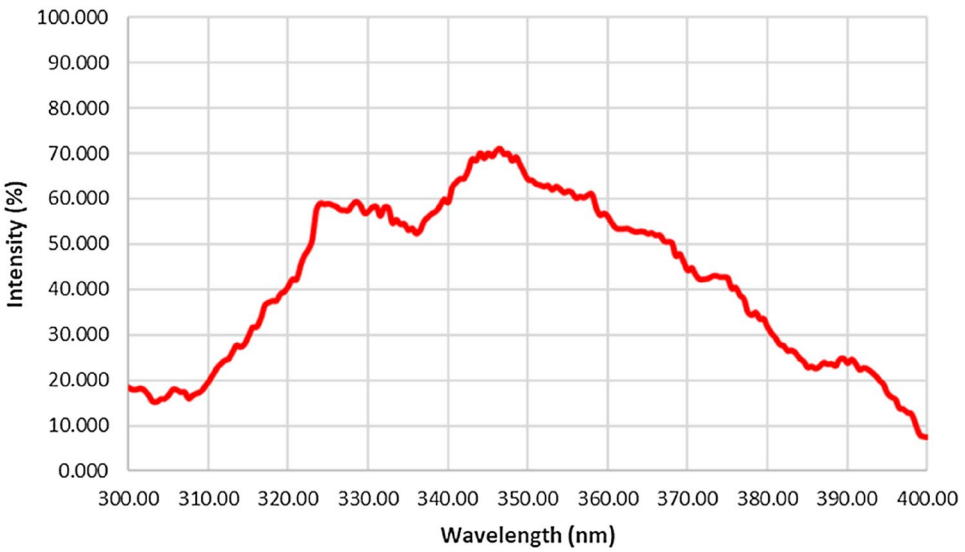

Fig. 6 Fluorescence spectrum of ΔrCP. The excitation measured between 300 and 400 nm

region of the protein is predicted to be disordered (Supplementary Fig. 6).

Discussion

In the present study, BBTV CP gene was cloned to various vectors by conventional method for amplification of the gene and overexpression of the recombinant protein. Previously, specific primers have been designed for nucleic acid-based detection of the virus in field samples and tissue culture plants (Mansoor et al. 2005; Mahadev et al. 2013) as well as for molecular characterisation of complete viral coat protein. The BBTV CP gene from various parts of the world have been characterised through PCR and sequencing of molecular clones in vectors like pTZ57R/T and InsT/A clone (Karan et al. 1994; Amin et al. 2008; Wickramaarachchi et al. 2016; Furuya et al. 2005; Selvarajan et al. 2010).

Cloning of BBTV CP complete coding region was attempted by Wanitchakorn et al. in pMAL-c2 expression vector for recombinant fusion protein overexpression (Wanitchakorn et al. 1997). In the study, however, the expression of BBTV CP-maltose binding (MBP) fusion protein was induced at 37 °C to yield high concentration of 60 kDa fusion protein. However, when BBTV CP was expressed in pET28a (+) vector, without any bulky affinity tag, the overexpression was observed to be low even after 6 h of induction at 37 °C (Arumugam et al. 2017). The small size and instability of protein was suspected to be the reasons for low level of overexpression. These results corroborate with the inferences from the present study that presence of bulky GST tag is contributing to the solubility as well as the stability of rCP and ΔrCP. However, the rCP was prone to proteolytic cleavage and separation of the GST tag during purification. Presence of trypsin cleavage sites were predicted by Peptide Cutter (Supplementary Fig. 1) at the extra amino acid sequences between the tag and the recombinant protein inserted as the result of cloning.

Restriction-free cloning also known as ligation independent cloning has been widely used for cloning of any target gene to any vector without addition of unnecessary sequences and moreover facilitating easy molecular manipulations (Aslanidis and Dejong 1990; Tillett and Neilan 1999; Unger et al. 2010). Through the RF platform, the thrombin as well as the trypsin cleavage sites were replaced with TEV protease cleavage site. According to many investigations, specificity of thrombin that cleaves off GST tag from fusion protein is not absolute (Jenny et al. 2003). From the map generated from Peptide Cutter thrombin cleavage site is present at 291st position of fusion protein apart from that present in the vector after the GST tag causing cleavage of protein (Supplementary Fig. 1). TEV protease overexpressed and purified in the laboratory, can be an efficient substitute for the expensive commercially available thrombin. The highly specific 27 kDa catalytic domain of nuclear inclusion protein (NIa) of *Tobacco etch virus*, recognises the amino acid sequence ENLYFQS/G and cleaves between glutamine (Q) and serine (S) or glycine (G) residues (Daros et al. 1999; Kapust and Waugh 2000; Nallamsetty et al. 2004).

Nevertheless, the attempt to cleave off the GST tag with TEV protease was unsuccessful at 20 °C. On the contrary, Nallemsetty et al. reported that TEV protease (mutated) is active at 4 °C (Nallamsetty et al. 2004). In the present study, large amount TEV protease also did not aid to resolve this issue. However, it is a well-known fact that some fusions proteins are intrinsically poor substrates for TEV protease. If the fusion protein exists in form of soluble aggregates or if the cleavage site is close to ordered structure of target protein causing steric hinderance make the cleavage site unavailable for the protease (Kapust and Waugh 2000). The N terminal of BBTV CP, rich in positively charged lysine (K) and arginine (R) residues are responsible for nucleotide binding (Kauffman and Karypis 2008). This disordered region predicted by FoldIndex, is positioned between the two ordered regions corresponding to the GST protein and the BBTV CP, providing a steric hinderance for the TEV protease. Addition of extra residues between TEV protease cleavage site and N terminus of BBTV CP might make the site available for proteolysis.

Fluorescence spectroscopy is a very important tool to help understand the tertiary structure of the protein which is not or cannot be characterised by crystallography (Lakowicz 1992). Globular proteins usually depict fluorescence emission maximum (λmax) at 330 nm in native condition as seen in case of alfalfa mosaic virus coat protein (Kan et al. 1986). The spectra of BBTV fusion protein deviated from this value indicating partially disordered tertiary structure of the protein. This corroborates the FoldIndex analysis wherein except N terminus the protein is predicted to be ordered. The result also reinstates that the reason for absence of cleavage of the tag at TEV protease site was due to the steric occlusion created by the folded structure.

Previously, viral coat proteins of few viruses have been characterised by florescence spectroscopy (Ksenofontov et al. 2013; Kumar et al. 2019; Anindya and Savithri 2003). Potato virus A coat protein exhibited a similar spectrum like that of BBTV CP with two maxima at 314 nm and ~ 330 nm (Ksenofontov et al. 2013). The λmax of 30 kDa recombinant coat protein of bean common mosaic potyvirus, was 345 nm similar to that observed in our study, indicating disorder in the protein (Kumar et al. 2019). It is experimentally demonstrated that surface exposed N and C terminus of coat protein of potyvirids are disordered facilitating in oligomerisation and assembly (Anindya and Savithri 2003). Correspondingly, in BBTV CP, the N terminal disordered region is likely to involve in DNA binding and assembly of the virus.

In conclusion, restriction-free cloning platform was used to manipulate the vector backbone for an improved overexpression and stability of recombinant BBTV CP. The high yield of quality protein is very important in immunological studies of the virus which was successfully achieved in the present study. However, further attempts will have to be made to remove the affinity tag and stabilise the 19 kDa recombinant CP for determining the structure of the virus.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13205-021-03017-x.

Acknowledgements Authors acknowledge Dr. Gayathri Panangat (Assistant Professor, IISER, Pune) for providing TEV protease clone and Ms. Sukanya Chakraborthy for valuable suggestions. HSS acknowledges National Academy of Sciences, India for the Senior scientist position.

Author contributions DD: investigation and validation of the experiments, writing of the manuscript. VL and HSS: resources, conceptualisation, supervision and editing the manuscript. NPM: data curation and editing the manuscript.

Funding DD acknowledge DST for INSPIRE fellowship (No. DST/INSPIRE Fellowship/2016/IF160765).

Declarations

Conflict of interest The authors have no conflict of interest to declare.

References

Amin I, Qazi J, Mansoor S, Ilyas M, Briddon RW (2008) Molecular characterisation of *Banana bunchy top virus* (BBTV) from Pakistan. Virus Genes 36:191–198. https://doi.org/10.1007/s11262-007-0168-y

Anhalt MD, Almeida RPP (2008) Effect of temperature, vector life stage, and plant access period on transmission of banana bunchy top virus to banana. Phytopathology 98:743–748

Anindya R, Savithri HS (2003) Surface-exposed amino- and carboxy-terminal residues are crucial for the initiation of assembly in *Pepper vein banding virus*: a flexuous rod-shaped virus. Virology 316:325–336. https://doi.org/10.1016/S0042-6822(03)00593-2

Arumugam C, Kalaimughilan K, Kathithachalam A (2017) Banana bunchy top viral coat protein (CP) gene expression studies at molecular level in Hill Banana cv. Sirumalai (AAB). Int J Curr Microbiol Appl Sci 6:398–411. https://doi.org/10.20546/ijcmas.2017.606.047

Aslanidis C, Dejong PJ (1990) Ligation-independent cloning of PCR products (Lic-PCR). Nucleic Acids Res 18:6069–6074

Beetham PR, Harding RM, Dale JL (1999) Banana bunchy top virus DNA-2 to 6 are monocistronic. Arch Virol 144:89–105

Blomme G, Ploetz R, Jones D, De Langhe E, Price N et al (2013) A historical overview of the appearance and spread of Musa pests and pathogens on the African continent: highlighting the importance of clean Musa planting materials and quarantine measures. Ann Appl Biol 162:4–26

Burns TM, Harding RM, Dale JL (1995) The genome organization of banana bunchy top virus: analysis of six ssDNA components. J Gen Virol 76:1471–1482. https://doi.org/10.1099/0022-1317-76-6-1471

Conant P (1992) Banana bunchy top disease, a new threat to banana cultivation in Hawaii. Proc Hawaii Entomol Soc 31:91–95

Dale JL, Phillips DA, Parry JN (1986) Double-stranded RNA in banana plants with bunchy top disease. J Gen Virol 67:371–375. https://doi.org/10.1099/0022-1317-67-2-371

Daros JA, Schaad MC, Carrington JC (1999) Functional analysis of the interaction between VPg-proteinase (NIa) and RNA polymerase (NIb) of tobacco etch potyvirus, using conditional and suppressor mutants. J Virol 73(10):8732–8740

Drew RA, Moisander JA, Smith MK (1989) The transmission of banana bunchy top virus in micropropagated bananas. Plant Cell Tissue Organ Cult 16:187–193

Elayabalan S, Kalaimughilan K (2013) Genetic engineering in banana and plantain. Adv Genet Eng 2:114. https://doi.org/10.4172/2169-0111.1000114

Elayabalan S, Subramaniam S, Selvarajan R (2015) Banana bunchy top disease (BBTD) symptom expression in banana and strategies for transgenic resistance: a review. Emirates J Food Agric 27:55–74. https://doi.org/10.9755/ejfa.v27i1.19197

Elayabalan S, Subramaniam S, Selvarajan R (2017) Construction of BBTV rep gene RNAi Vector and evaluate the silencing mechanism through injection of Agrobacterium tumefaciens transient expression system in BBTV Infected Hill Banana plants cv. *Virupakshi* (AAB). Indian J Nat Sci 42(7):12395–12403

Furuya N, Kawano S, Natsuaki KT (2005) Characterization and genetic status of *Banana bunchy top virus* isolated from Okinawa. Japan J Gen Plant Pathol 71:68–73. https://doi.org/10.1007/s10327-004-0162-3

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31(13):3784–3788. https://doi.org/10.1093/nar/gkg563

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, Totowa, pp 571–607

Green MR, Sambrook J (2016) Preparation of plasmid DNA by alkaline lysis with sodium dodecyl sulfate: minipreps. Cold Spring Harb Protoc. https://doi.org/10.1101/pdb.prot093344

Harding RM, Burns TM, Dale JL (1991) Virus-like particles associated with banana bunchy top disease contain small single-stranded DNA. J Gen Virol 72:225–230. https://doi.org/10.1099/0022-1317-72-2-225

Harding RM, Burns TM, Hafner G, Dietzgen RG, Dale JL (1993) Nucleotide sequence of one component of the banana bunchy top virus genome contains a putative replicase gene. J Gen Virol 74:323–328. https://doi.org/10.1099/0022-1317-74-3-323

Hema M, Kirthi N, Sreenivasulu P, Savithri HS (2003) Development of recombinant coat protein antibody based IC-RT-PCR for detection and discrimination of sugarcane streak mosaic virus isolates from Southern India. Arch Virol 148:1185–1193. https://doi.org/10.1007/s00705-003-0015-y

Jenny RJ, Mann KG, Lundblad RL (2003) A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. Protein Expr Purif 31:1–11. https://doi.org/10.1016/S1046-5928(03)00168-2

Kan JH, Van Resandt RWW, Dekkersf HPJM (1986) Fluorescence studies on the coat protein of alfalfa mosaic virus. J Biomol Struct Dyn 3:827–842. https://doi.org/10.1080/07391102.1986.10508464

Kapust RB, Waugh DS (2000) Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr Purif 19:312–318

Karan M, Harding RM, Dale JL (1994) Evidence for two groups of banana bunchy top virus isolates. J Gen Virol 75:3541–3546. https://doi.org/10.1099/0022-1317-75-12-3541

Kauffman C, Karypis G (2008) An analysis of information content present in protein–DNA interactions. Pac Symp Biocomput 13:477–488

King AMQ, Adams MJ, Lefkowitz EJ, Carstens EB (eds) (2012) Virus taxonomy: ninth report of the international committee on taxonomy of viruses. Elsevier Academic Press, New York

Krishna B, Kadu AA, Vyavhare SN, Chaudhary RS, Joshi SS, Patil AB, Sane P (2013) RNAi-mediated resistance against banana bunchy top virus (BBTV) in 'Grand Nain' banana. Acta Hortic 974:157–164. https://doi.org/10.17660/ActaHortic.2013.974.18

Ksenofontov AL, Paalme V, Arutyunyan AM, Semenyuk PI, Fedorova NV, Rumvolt R, Baratova LA, Järvekülg L, Dobrov EN (2013) Partially disordered structure in intravirus coat protein of *potyvirus Potato Virus A*. PLoS ONE. https://doi.org/10.1371/journal.pone.0067830

Kumar PL, Hanna R, Alabi OJ, Soko MM, Oben TT et al (2011) Banana bunchy top virus in sub-Saharan Africa: investigations on virus distribution and diversity. Virus Res 159:171–182

Kumar PL, Selvarajan R, Iskra-Caruana ML, Chabannes M, Hanna R (2014) Biology, etiology, and control of virus diseases of banana and plantain. Advances in virus research, 1st edn. Elsevier Inc, Amsterdam. https://doi.org/10.1016/bs.aivir.2014.10.006

Kumar S, Karmakar R, Garg DK, Gupta I, Patel AK (2019) Elucidating the functional aspects of different domains of bean common mosaic virus coat protein. Virus Res 273:197755. https://doi.org/10.1016/j.virusres.2019.197755

Lakowicz JR (1992) Topics in fluorescence spectroscopy: volume 3. Biochemical applications. Plenum Press, New York

Lowe S, Browne M, Boudjelas S, De Poorter M (2000) 100 of the world's worst invasive alien species a selection from the Global Invasive Species Database. (The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN)). http://www.issg.org/pdf/publications/worst_100/english_100_worst.pdf

Magee CJP (1927) Banana bunchy top disease of the banana. CSIResearch Bull, London

Magee CJP (1953) Some aspects of bunchy top disease of banana and other Musa spp. J Proc R Soc N S W 87:1–18

Mahadev SR, Thamilarasan SK, Kathithachalam A (2013) PCR detection of *Banana bunchy top virus* (BBTV) at tissue culture level for the production of virus-free planting materials. Int Res J Biol Sci Int Res J Biol Sci 2:2278–3202

Mandal B, Mandal S, Pun KB, Varma A (2004) First report of the association of a nanovirus with foorkey disease of large cardamom in India. Plant Dis 88:428

Mansoor S, Qazi J, Amin I, Khatri A, Khan IA (2005) A PCR-based method, with internal control, for the detection of Banana bunchy top virus in banana. Mol Biotechnol 30:167–169

Nallamsetty S, Kapust RB, Tözsér J, Cherry S, Tropea JE, Copeland TD, Waugh DS (2004) Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. Protein Expr Purif 38:108–115

Niu S, Wang B, Guo X, Yu J, Wang X (2009) Identification of two RNA silencing suppressors from banana bunchy top virus. Arch Virol 154:1775–1783. https://doi.org/10.1007/s00705-009-0515-5

Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL (2005) FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics 21:3435–3438. https://doi.org/10.1093/bioinformatics/bti537

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:1–17. https://doi.org/10.3389/fmicb.2014.00172

Schumann W, Ferreira LCS (2004) Production of recombinant proteins in *Escherichia coli*. Genet Mol Biol 27:442–453. https://doi.org/10.1590/S1415-47572004000300022

Selvarajan R, Sheeba M, Balasubramanian V, Rajmohan R, Dhevi NL, Sasirekha T (2010) Molecular characterization of geographically different banana bunchy top virus isolates in India. Indian J Virol 21:110–116. https://doi.org/10.1007/s13337-010-0020-1

Sharman M, Thomas JE, Skabo S, Holton TA (2008) Abaca bunchy top virus, a new member of the genus Babuvirus (family Nanoviridae). Arch Virol 153:135–147

Shekhawat UKS, Ganapathi TR, Hadapad AB (2012) Transgenic banana plants expressing small interfering RNAs targeted against viral replication initiation gene display high-level resistance to banana bunchy top virus infection. J Gen Virol 93:1804–1813. https://doi.org/10.1099/vir.0.041871-0

Tillett D, Neilan BA (1999) Enzyme-free cloning: a rapid method to clone PCR products independent of vector restriction enzyme sites. Nucleic Acids Res 27:e26

Tsao T (2000) Towards the development of transgenic banana bunchy top virus (BBTV)-resistant banana plants: interference with replication, Ph.D Thesis, Queensland University of Technology, Australia

Unger T, Jacobovitch Y, Dantes A, Bernheim R, Peleg Y (2010) Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. J Struct Biol 172:34–44. https://doi.org/10.1016/j.jsb.2010.06.016

Vetten HJ, Chu PW, Dale JL, Harding RM, Hu J et al (2005) Nanoviridae. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) Virus taxonomy: eighth report of the international committee on taxonomy of viruses. Academic Press, San Diego, pp 343–352

Wanitchakorn R, Harding RM, Dale JL (1997) Banana bunchy top virus DNA-3 encodes the viral coat protein. Arch Virol 142:1673–1680. https://doi.org/10.1007/s007050050188

Wanitchakorn R, Hafner GJ, Harding RM, Dale JL (2000) Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6. J Gen Virol 81:299–306. https://doi.org/10.1099/0022-1317-81-1-299

Wardlaw CW (1961) Mosaic, infectious chlorosis and other virus diseases: banana diseases, including plantains and abaca. Longmans, London, pp 116–145

Wickramaarachchi WART, Shankarappa KS, Rangaswamy KT, Maruthi MN, Rajapakse RG, Ghosh S (2016) Molecular characterization of *Banana bunchy top virus* isolate from Sri Lanka and its genetic relationship with other isolates. Virus Dis 27:154–160. https://doi.org/10.1007/s13337-016-0311-2

Yu NT, Zhang YL, Feng TC, Wang JH, Kulye M et al (2012) Cloning and sequence analysis of two banana bunchy top virus genomes in Hainan. Virus Genes 44:488–494