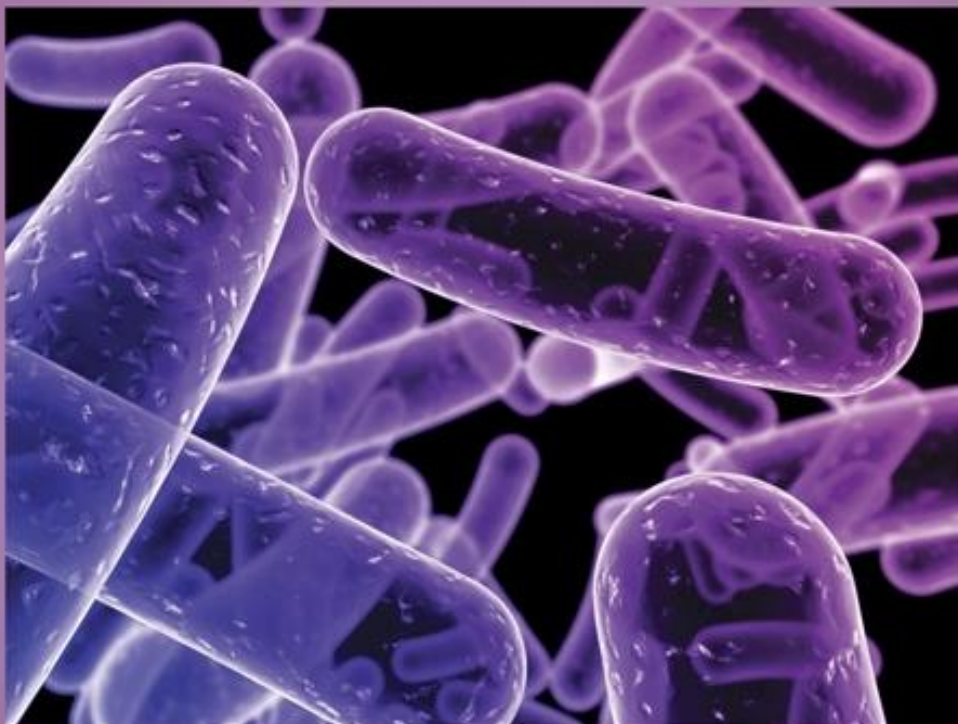




EGYPTIAN ACADEMIC JOURNAL OF  
**BIOLOGICAL SCIENCES**  
MICROBIOLOGY

G



ISSN  
2090-0872

WWW.EAJBS.EG.NET

**Vol. 14 No. 2 (2022)**



## Restriction Enzyme Maps of Circular Sequences of BBTV-DNA Components and Bioinformatics of Their Amino Acids

Abdel-Razek Fatma S.<sup>1</sup>, Ibrahim S.D.<sup>2</sup>, El-Masry Samar S.<sup>1</sup>, Faiesal Abeer A.<sup>3</sup> and Sadik A.S.<sup>1</sup>

1-Department of Agricultural Microbiology (Virology Laboratory), Faculty of Agriculture, Ain Shams University, P.O. Box 68, Hadayek Shobra, 11241, Cairo, Egypt

2-Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), 9 Gamaa St., P.O. Box12619, Giza, Egypt

3-Department of Basic and Applied Agricultural Sciences, Higher Institute for Agriculture Co-Operation, Shoubra EL-Kheima, Egypt

\*E. Mail: [samar\\_sayed@agr.asu.edu.eg](mailto:samar_sayed@agr.asu.edu.eg)

### ARTICLE INFO

Article History

Received:20/7/2022

Accepted:24/9/2022

Available:26/9/2022

#### Keywords:

BBTV, Restriction enzymes map, Domains, Nanovirus, Babovirus.

### ABSTRACT

In this study, NEBcutter is used in generating the restriction enzymes map of the circular sequence of BBTV-DNA components (DNA-R, DNA-U3, DNA-S, DNA-M, DNA-M and DNA-N) appearing with each cleavage code and enzyme name code. Results showed that the closest enzymes at the 5' end of the restriction sites of *HphI*, *ApoI*, *BtsIMuI*, *Fnu4HI*, *Hpy99I* and *BccI* were detected for the six BBTV-DNA components, receptively. While, the restriction sites of *Cac8I*, *AluI*, *Hpy166II*, *DpnII*, *AccI* and *AseI* were detected as the closest enzymes at 3' end for the same components, respectively. These BBTV-DNA components were encoding the genes of replication-associated protein (BBJ34149.1, 278 aa), U3 protein (BBJ34150.1, 88 aa), capsid protein (BBJ34151.1, 175 aa), movement protein (BBJ34152.1, 102 aa), cell-cycle link protein (BBJ34153.1, 156 aa) and nuclear shuttle protein (BBJ34154.1, 129 aa), respectively. Similarity ranged from 95 to 100% found when sequences producing significant alignments were done between the ORFs of BBTV-DNA components and the most similar BBTV overseas strains. The number and type of domains of the ORFs of the six BBTV components and their families were determined. Two domains were found for DNA-R (viral rep and RNA\_helicase domains). NA-U3 has a hypothetical protein belonging to un-characterized protein U3 domain, one domain called nanovirus\_coat the protein family membership is Nanovirus coat protein recorded for DNA-S, DNA-M has one domain Babovirus\_MP, DNA-C has a domain belonging to cell cycle protein and DNA-N has a domain belonging to nuclear shuttle protein domain.

### INTRODUCTION

Banana bunchy top virus (BBTV) the causal agent of banana bunchy top disease (BBTD) is considered the most effective viral disease on the productivity of bananas worldwide (Sadik 1994, Balakrishnan *et al.*, 1996, Lowe *et al.*, 2000, Almedia *et al.*, 2009, Adegbola *et al.*, 2013, Mpoki *et al.*, 2021 and Singh *et al.*, 2022).

BBTV was characterized by a circular single strand DNA (ssDNA) genome consisting of six components called DNA-R (BBTV-DNA-1) (Harding *et al.* 1993, Harding *et al.* 2000, Amin *et al.*, 2008), DNA-U3 (BBTV-DNA-2) (Sadik 1994 and Burns *et al.*, 1995), DNA-S (BBTV-DNA-3) (Burns *et al.*, 1995), DNA-M (BBTV-DNA-4) (Burns *et al.*, 1995), DNA-C (BBTV-DNA-5) (Burns *et al.*, 1995), and DNA-N (BBTV-DNA-6) (Burns *et al.*, 1995). BBTV was classified as a member of the genus *Babuvirus* (Mandal 2010) and the family Nanoviridae (Vetten *et al.*, 2005, King *et al.*, 2011 and Vetten *et al.*, 2012).

It has been shown that the six DNA components shown to be responsible for the genes of rolling-circle replication initiation protein, protein of unknown function, coat protein peptide with an estimated molecular mass of 20 kDa, the movement protein, cell cycle link protein and nuclear shuttle protein, respectively (Harding *et al.*, 1993, Sadik 1994, Burns *et al.*, 1995, Beetham *et al.*, 1997, Hafner *et al.*, 1997 and Nour El-Din *et al.*, 2005, Vetten *et al.* 2005, King *et al.*, 2011 and Vetten *et al.*, 2012).

Bioinformatics could be described as a multidisciplinary technology as it can include both computer science and information technology and biotechnology, which provides a great opportunity to analyze and interpret biological data, as well as modeling biological processes in addition to developing algorithms and related statistics (Su *et al.*, 2003, Selvaranjan *et al.*, 2010 and Banerjee *et al.*, 2014). They also revealed that

bioinformatics can be used to solve biological problems through computer technologies and applications, for example, bioinformatics is recently used to study genome sequences (nucleotide and amino-acid sequences), genetic protein structures and protein domains and their families, molecular networks and proteomics (Molina and Kudagamage 2002).

Yu *et al.*, (2011) indicated that bioinformatics analysis of BBTV sequencing of Hainan isolates, which represents a satellite DNA component with 12 DNA sequences motifs. They also predicted the structure, chemical properties and physical, signal peptide, phosphorylation, secondary & tertiary structures and functional domains of its encoding protein, and finally compared with the corresponding quantities in the replication initiation protein of BBTV-DNA component 1.

This study aimed to determine the restriction enzymes map of the circular sequence of the six BBTV-DNA components appearing in each cleavage code and enzyme name code. Analyses of deduced amino acids and their ORFs and types of domains and their superfamilies were also aimed.

## MATERIALS AND METHODS

Analyses of the six BBTV-DNA components which encode the genes of replication-associated protein (BBJ34149.1), U3 protein (BBJ34150.1), capsid protein (BBJ34151.1), movement protein (BBJ34152.1), cell-cycle link protein (BBJ34153.1) and nuclear shuttle protein (BBJ34154.1), respectively, were conducted using as follows:

BBTV-DNA components			
Restriction enzymes endonucleases maps	DNA-R	LC468138.1	<a href="https://nc2.neb.com/NEBcutter2/">https://nc2.neb.com/NEBcutter2/</a>
	DNA-U3	LC468139.1	
	DNA-S	LC468140.1	<a href="http://tools.neb.com/NEBcutter">http://tools.neb.com/NEBcutter</a>
	DNA-M	LC468141.1	
	DNA-C	LC468142.1	<a href="http://www.neb.com/rebase">http://www.neb.com/rebase</a>
	DNA-N	LC468143.1	
BBTV-Protein components			
Deduced amino acids and their ORFs	<a href="https://www.bioinformatics.org/sms2/orf_find.html">https://www.bioinformatics.org/sms2/orf_find.html</a>		
	<a href="https://www.ncbi.nlm.nih.gov/orffinder/">https://www.ncbi.nlm.nih.gov/orffinder/</a>		
Pairwise alignment	<a href="https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&amp;PAGE_TYPE=BlastSearch&amp;LINK_LOC=blasthome">https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&amp;PAGE_TYPE=BlastSearch&amp;LINK_LOC=blasthome</a>		
Multiple sequencing alignment	<a href="https://www.genome.jp/tools-bin/clustalw">https://www.genome.jp/tools-bin/clustalw</a>		
Types of domains and their superfamilies	<a href="https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi">https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi</a>		
	<a href="https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi">https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi</a>		
	<a href="https://myhits.isb-sib.ch/cgi-bin/motif_scan">https://myhits.isb-sib.ch/cgi-bin/motif_scan</a>		
	<a href="https://meme-suite.org/meme/tools/mem">https://meme-suite.org/meme/tools/mem</a>		

**RESULTS AND DISCUSSION**  
**Restriction Enzyme Maps of a Circular Sequence of BBTV-DNA Components:**

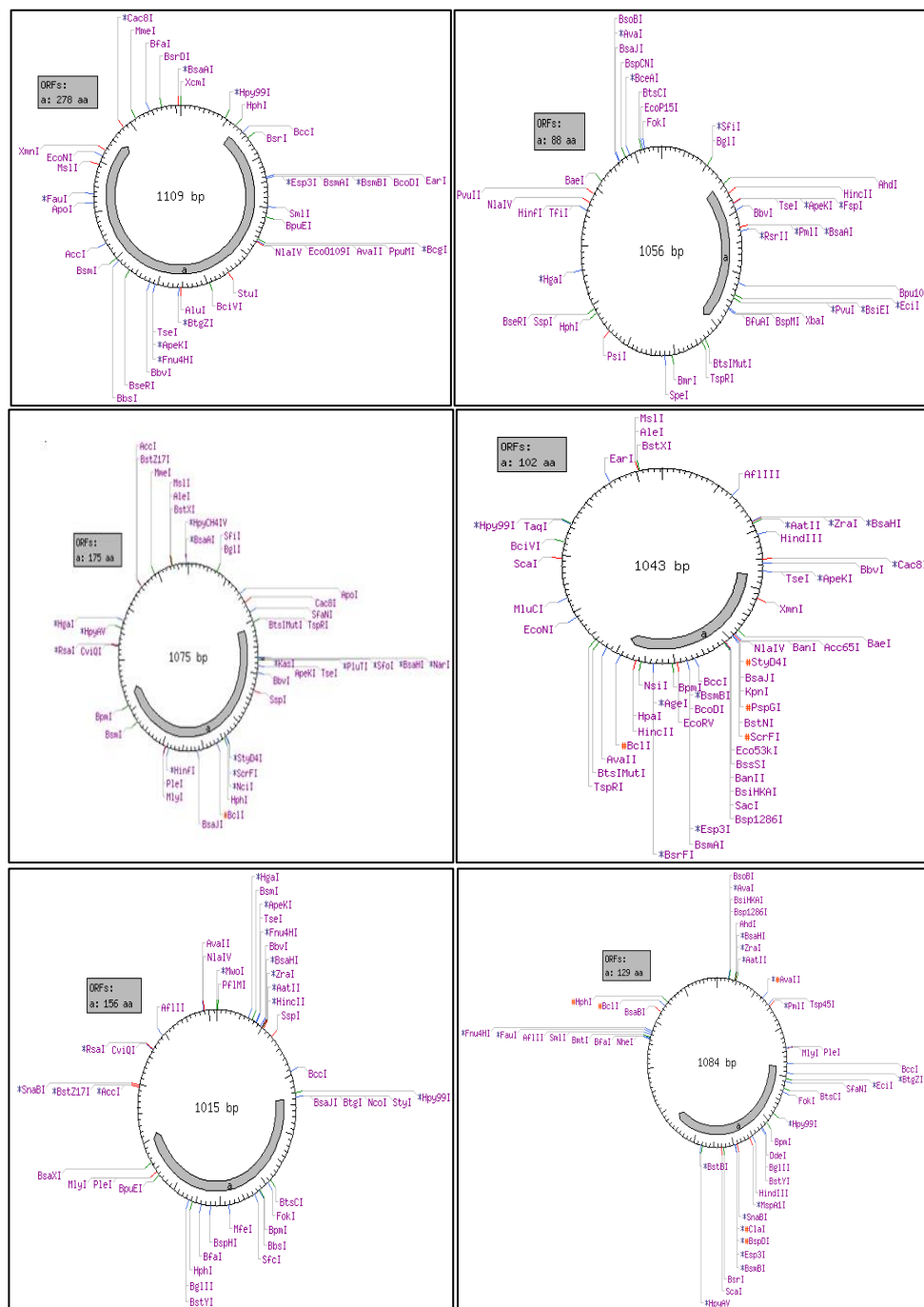
NEBcutter, version 1.0, is a program available via a web server that will accept an input DNA sequence and produce a comprehensive report of the restriction enzymes that will cleave the sequence (Vincze et al. 2003). It produces a variety of outputs including restriction enzyme maps, theoretical digests and links into the restriction enzyme database, and rebase.

Data in Table (1) and illustrated in Figure (1) show that NEBcutter was successfully used in generating the restriction enzymes map of the circular sequence of BBTV-DNA components (DNA-R (LC468138.1), DNA-U3 (LC468139.1), DNA-S (LC468140.1), DNA-M (LC468141.1), DNA-C (LC468142.1) and

DNA-N (LC468143.1)) appearing each of cleavage code and enzyme name code. These BBTV-DNA components were encoding the genes of replication-associated protein (BBJ34149.1), U3 protein (BBJ34150.1), capsid protein (BBJ34151.1), movement protein (BBJ34152.1), cell-cycle link protein (BBJ34153.1) and nuclear shuttle protein (BBJ34154.1), respectively. Both of the closest enzymes at the 5' end and 3' end occurred. At the level of closest enzymes at the 5' end the restriction sites of *HphI*, *ApoI*, *BtsIMutI*, *Fnu4HI*, *Hpy99I* and *BccI* were detected for the six BBTV-DNA components, respectively. While, the restriction sites of *Cac8I*, *AluI*, *Hpy166II*, *DpnII*, *AccI* and *AseI* were detected as the closest enzymes at 3' end for the same six BBTV-DNA components, respectively.

**Table (1): Closest enzymes at 5' and 3' ends of products of BBTV-DNA components.**

Components	Products	Amino acids (aa)	Coordinates	Proteins ID	Closest enzymes at 5' end	Closest enzymes at 3' end
LC468138.1	Replication-associated protein	278	127..963	BBJ34149.1	<i>HphI</i>	* <i>Cac8I</i>
LC468139.1	U3 protein	88	143..409	BBJ34150.1	<i>ApoI</i>	<i>AluI</i>
LC468140.1	capsid protein	175	213..740	BBJ34151.1	<i>BtsIMutI</i>	<i>Hpy166II</i>
LC468141.1	movement protein	102	279..587	BBJ34152.1	* <i>Fnu4HI</i>	# <i>DpnII</i>
LC468142.1	cell-cycle link protein	156	240..710	BBJ34153.1	* <i>Hpy99I</i>	* <i>AccI</i>
LC468143.1	nuclear shuttle protein	129	281..670	BBJ34154.1	<i>BccI</i>	<i>AseI</i>



**Fig. 1:** Restriction enzymes map generated by NEBcutter of a circular sequence of BBTV-DNA components (DNA-R: 1109, DNA-U3: 1056, DNA-S: 1075, DNA-M: 1043, DNA-C: 1015 and DNA-N: 1084 nts). Note, that both of cleavage code and of enzyme name code were available.

### Sequencing Analyses of Amino Acids Deduced from The Six Components of BBTV-DNA Compared to The Overseas Strains in GenBank:

The nucleotide sequences of the six components of BBTV (DNA-R, DNA-U3, DNA-S, DNA-M, DNA-C and DNA-N) were

subjected to analysis of their encoded amino acids and their function. Results in Figure (2) showed coding sequences (CDS) translation in reading frame 3 (837 nts/278 aa) on the direct strand extends from base 132 to base 968 of BBTV-DNA-1 (DNA-R) documented in GenBank (BBJ34149.1). Data in Figure (3)

show alignment sequences of ORF of BBTV-DNA-1 (BBJ34149.1) and only one difference in amino acids between the DNA-R and CAL90980.1 was recorded plus three differences overall when compared to the other four overseas BBTV strains.

```

1 ATGTTCCACCATCAACAATCCCACAACACTACCAGTGTAGGGAT
M F T I N N P T T L P V M R D
46 GAGATCAAATATATGGTATATCAAGTGGAGGGGACAGGAGGGT
E I K Y M V Y Q V E R G Q E G
91 ACTCGTCATGTGCAAGGATATGTCGAGATGAAGAGACGAAGTCT
T R H V Q G Y V E M K R R S S
136 CTGAAGCAGATGAGAGCCTTCTTCTGGCCACACCTTGAGAAA
L K Q M R A F F P G A H L E K
181 CGAAAGGGAAGCCAAGAAGAAGCGCGGTCTACTGTATGAAGGAA
R K G S Q E E | A R S Y C M K E
226 GATACAAGAAATCGAAGTCCCTTCGAGTTGGTGCATTAAATTG
D T R I E G P F E F G A F K L
271 TCATGTAATGATAATTTATTTGATGTCATACAGGATATGCGTGAA
S C N D N L F D V I Q D M R E
316 ACGCACAAAAGGCCTCTGGAGTATTTATATGATTGTCCTAAGACC
T H K R P L E Y L Y D C P N T
361 TTCGATAGAAGTAAGGATACATTATACAGAGTACAAGCAGAGATG
F D R S K D T L Y R V Q A E M
406 AATAAACGAAGCGGATGAATAGCTGGAGAACGCTTTTCAGTGGC
N K T K A M N S W R T S F S A
451 TGGACATCAGAAGTGGAGAATATCATGGCCGAGCCATGTCATCGG
W T S E V E N I M A Q P C H R
496 AGAATAAATTGGGTCTATGGCCCAATGGAGGAGAAGGAAAGACA
R I I W V Y G P N G G E G K T
541 ACGTATGCAAAACATCTAATGAAGACGAGGAATCGGTTTTATTCT
T Y A K H L M K T R N A F Y S
586 CCAGGAGGAAAATCATTGGATATATGTAGACTGTATAATTACGAG
P G G K S L D I C R L Y N Y E
631 GATATTGTTATATTTGATATTCGAAGATGCAAAGAGGATTATTTA
D I V I F D I P R C K E D Y L
676 AATTATGGGTTATTAGAAGAATTTAAGAATGGAATAATCAAGC
N Y G L L E E F K N G I I Q S
721 GGAAATATGAACCCGTTTGAAGATAGTGAATATGTCGAAGTC
G K Y E P V L K I V E Y V E V
766 ATTGTAATGGCTAACTTCCTCCGAAGGAAGGAATCTTTCTGAA
I V M A N F L P K E G I F S E
811 GATCGAATAAAGTTGGTTCTTGCTGA
D R I K L V S C *
    
```

**Fig. 2:** ORF2 CDS translation in reading frame 3 (837 nts/278 aa) on the direct strand extends from base 132 to base 968 of BBTV-DNA-1 (DNA-R) component documented in GenBank (BBJ34149.1).

BBJ34149.1	1	MFTINNPTTLPVMRDEIKYMVYQVERGQEGTRHVQGYVEMKRRSSLKQMRAFPFGAHLEK	60
AAZ07720.2	9	.....	68
AKD71380.1	9	.....V.....	68
CAL90980.1	9	.....	68
CAJ20054.1	9	.....G.....	68
ALA65728.1	9	.....	68
BBJ34149.1	61	RKGSQEARSYCMKEDTRIEGPFEPFGAFKLSNDNLFVDIQDMRETHKRPLEYLYDCPNT	120
AAZ07720.2	69	.....	128
AKD71380.1	69	.....	128
CAL90980.1	69	.....	128
CAJ20054.1	69	.....	128
ALA65728.1	69	.....R.....	128
BBJ34149.1	121	FDRSKDTLYRVQAE MNRTKAMNSWRTSFSAWTSEVENI MAQPCRRII WYVGPNGGEGKT	180
AAZ07720.2	129	.....	188
AKD71380.1	129	.....	188
CAL90980.1	129	.....N.....	188
CAJ20054.1	129	.....	188
ALA65728.1	129	.....	188
BBJ34149.1	181	TYAKHLMKTRNAPYSPGGKSLDICRLYNYEDIVIPDI PRCKEDYLN YGLLEEFKNGIIQS	240
AAZ07720.2	189	.....	248
AKD71380.1	189	.....	248
CAL90980.1	189	.....	248
CAJ20054.1	189	.....	248
ALA65728.1	189	.....	248
BBJ34149.1	241	GRYEPVLKIVEYVEVIVMANFLPKEGIFSEDRIKLVSC	278
AAZ07720.2	249	.....	286
AKD71380.1	249	.....	286
CAL90980.1	249	.....	286
CAJ20054.1	249	.....	286
ALA65728.1	249	.....	286

**Fig. 3:** Alignment sequences of ORF of BBTV-DNA-1 (DNA-R) (BBJ34149.1) compared to five overseas BBTV strains shows amino acid differences in between.

Regarding the BBTV-DNA-2 (DNA-U3) documented in GenBank (BBJ34150.1), data in Figure (4) show its ORF1 CDS translation in reading frame 3 (267 nts/88 aa) on the direct strand extends from base 153 to base 419. The lowest differences in amino

acids (two differences) between BBTV-DNA-2 (BBJ34150.1) and the POC750.1 were found when its amino acid sequence was compared to the five overseas BBTV strains (Fig. 5).

```

1 ATGACCGAAGGTCAAGGTAACCTGGTCAACATTATTCTGGCTTGCC
  M T E G Q G N W S T L F W L A
46 CAGCAAGATACACGAATTAATTTTATTAATTCGTAGGACACGTGGA
  Q Q D T R I N L L I R R T R G
91 CGGACCGAAATACTCTTGCATCTCTATAAATACCCTAATCCTGTA
  R T E I L L H L Y K Y P N P V
136 AAGGATAATTCGTCTCTCTCTCTCTGTCGTAAGGTGGTGTGCTAAGG
  K D N C S L S S V K V V V L R
181 CGGAAGATCGCCAGCGCGGATCGTCGGAACGACCTGCATCTAGAG
  R K I A S G D R R N D L H L E
226 AGGCGGCGAGGAACTACGGAGCGTATATCGGGTATTTATAG
  R R R G N Y G A Y I G Y L *
    
```

**Fig. 4:** ORF1 CDS translation in reading frame 3 (267 nts/88 aa) on the direct strand extends from base 153 to base 419 of BBTV-DNA-2 (DNA-U3) documented in GenBank (BBJ34150.1).

BBJ34150.1	1	MTEGQGNWSTLFWLAQQDTRINLLIRRTGRTEILLHLYKYPNPVKDNCSSSVKVVVLR	60
POC750.1	1	.....R.....	60
AQR06132.1	1	.....R.....A.....I.....K.....	60
ANV20585.1	1	.....R.....K.....S.....	60
ACR77724.1	1	.....R.....K.....S.....	60
ACH99204.1	1	.....R.....K.....S.....	60
BBJ34150.1	61	RKIASGDRRNDLHLERRRNYGAYIGYL	88
POC750.1	61	.....E.....	88
AQR06132.1	61	.....E.....	88
ANV20585.1	61	.....V...SEA...E... ..	89
BBJ34150.1	61	.....	77
ACH99204.1	61	.....V.....	77

**Fig. 5:** Alignment sequences of ORF of BBTV-DNA-2 (DNA-U3) (BBJ34150.1) compared to five overseas BBTV strains shows amino acid differences in between.

In the case of BBTV-DNA-3 (DNA-S) documented in GenBank (BBJ34151.1), data in Figure (6) show its ORF4 CDS translation in reading frame 3 (528 nts/175 aa) on the direct strand extends from base 213 to base 740, and encoding the capsid protein

gene of BBTV. Only one difference in between was recorded when alignment sequences of ORF of BBTV-DNA-3 (BBJ34151.1) compared to five overseas BBTV strains (Fig. 7).

```

1 ATGTCAGACAAGAAATGGCTAGGTATCCGAAGAAATCCATCAAG
  M F R Q E M A R Y P K K S I K
46 AAGAGCGGGTTGGGGCGGAAATGATGGCAGCAAGCGGCAAG
  K R R V G R R K Y G S K A A T
91 AGCCACGACTACTCGTCGTGAGGTCATATTTGGTTCCTGAAAAC
  S H D Y S S S G S I L V P E N
136 ACCGTCAAGGTATTTGGGATGAGCCTACTGATAAAACATTACCC
  T V K V F R I E P T D K T L P
181 AGATATTTTATCTGGAAAATGTTTATGCTTCTGTGCAAGGFG
  R Y F I W K M F M L L V C K V
226 AAGCCCGGAAGAATACTTCATTTGGGCTATGATCAAGAGTTCTTGG
  K P G R I L H W A M I K S S W
271 GAAATCAACCAGCCGACCACTCTGTTAAACTGGTATGTAAGTGGG
  E I N Q P T T C L E A R G L F
316 ATTAACCTGAACACAGCCATCTGTTAAACTGGTATGTAAGTGGG
  I K P E H S H L V K L V C S G
361 GAACCTGAAGCAGGACTCGCAACAGCAACATCAGATGTTGAATGT
  E L E A G V A T G T S D V E C
406 CTTTTCAGGAGACACACCGTGTGAGGAGAAATGTAACAGAGGTG
  L L R K T T V L R K N V T E V
451 GATTATTTATATTTGGCATTTCTATGTTAGTTCTGGAGTAAGTATA
  D Y L Y L A F Y C S S G V S I
496 AACTACCAGACAGAATTACATATCATGTTTGA
  N Y Q N R I T Y H V *
    
```

**Fig. 6:** ORF4 CDS translation in reading frame 3 (528 nts/175 aa) on the direct strand extends from base 213 to base 740 of BBTV-DNA-3 (DNA-S) documented in GenBank (BBJ34151.1).

BBJ34151.1	1	MFRQEMARYPKKSIKRRVGRRKYGSKAATSHDYSSSGSILVPENTVKVFRIEPTDKTLP	60
NP_604477.1	1	.....	60
ABK96868.1	1	...Q.....L.....	60
AEV53368.1	1	...Q.....L.....	60
AEF13039.1	1	...Q.....A.....L.....	60
ACZ59089.1	1	...Q.....L.....	60
BBJ34151.1	61	RYFIWKMFMLLVCKVKPGRILHWAMIKSSWEINQPTTCLEARGLFIKPEHSHLVKLVCSG	120
NP_604477.1	61	.....P.....	120
ABK96868.1	61	.....P.....	120
AEV53368.1	61	.....P.....D.....	120
AEF13039.1	61	.....P.....	120
ACZ59089.1	61	.....A.....P.....	120
BBJ34151.1	121	ELEAGVATGTS DVECLLRKTTVLRKNVTEVDYLYLAFYCSSGVSYNYQNRITYHV	175
NP_604477.1	121	.....	175
ABK96868.1	121	.....	175
AEV53368.1	121	.....	175
AEF13039.1	121	.....	175
ACZ59089.1	121	..... .....	175

**Fig. 7:** Alignment sequences of ORF of BBTV-DNA-3 (DNA-S) (BBJ34151.1) compared to five overseas BBTV strains shows amino acid differences in between.

For BBTV-DNA-4 (DNA-M) documented in GenBank (BBJ34152.1), results in Figure (8) showed its ORF2 CDS translation in reading frame 3 (309 nts/102 aa) on the direct strand extends from base 279 to base 587 and encoding the movement

protein gene of BBTV. A number of 16 differences were recorded when alignment sequences of ORF of BBTV-DNA-4 (BBJ34152.1) compared to five overseas BBTV strains were analyzed (Fig. 9).

```

1 ATGGCATTAAACAACAGAGCTGGTTAAACTATTCTTTGAATGGTTT
  M A L T T E L V K L F E | E W F
46 CTGTTCTTTGGAGCAATATTTATTGCGATTACAATATTATATATA
  L F F G A I F I A I T I L Y I
91 TTGTTGGTTTTGCTCTTTGAGGTACCCAGGTATATTAAGGAGCTC
  L L V L L F E V P R Y I K E L
136 GTGAGGTATTTGGTAGAATACCTGACCAGACTATGTGTATGGATG
  V R Y L V E Y L T R L C V W M
181 CAAAGGACGCAGTTGATGGAGGCAACTGGAGACGTAGATATCGGC
  Q R T Q L M E A T G D V D I G
226 AGAGGTATTGTGGAAGACAGACGGGATAAACAACCGTTGTCATA
  R G I V E D R R D K Q P V V I
271 CCAAATGCATCTCATGTTAACCTTCTCAACCCACATTGA
  P N A S H V N P S Q P H *
    
```

**Fig. 8:** ORF2 CDS translation in reading frame 3 (309 nts/102 aa) on the direct strand extends from base 279 to base 587 of BBTV-DNA-4 (DNA-M) documented in GenBank (BBJ34152.1).

BBJ34152.1	1	MALTTELVLKFFEFWFFFGAIFIAITILYILLVLLFEVPRYIKELVRYLVEYLRLCVWM	60
AEL32419.1	1	.....R.....RR...	60
AEL32422.1	1	.....R.....RR...	60
AKO70855.1	1	.....R.....K...Q.....RR...	60
AKO70966.1	1	.....R.....K...Q.....RR...	60
AKO70967.1	1	.....R.....V.....RR...	60
BBJ34152.1	61	QRTQLMEATGDVDIGRGIVEDRRDKQPVVIPNASHVNPSQPH	102
AEL32419.1	61	.....T.....E.....Q..A...HV.Q.I...N	102
AEL32422.1	61	.....T.....E.....Q..A...HV.Q.I...N	102
AKO70855.1	61	.....S.....EF.....Q..A...Q..Q.....	100
AKO70966.1	61	.....S.....EF.....Q..A...Q..Q.....	100
AKO70967.1	61	.....T.....E.....Q..A...HV.Q.I...N	102

**Fig. 9:** Alignment sequences of ORF of BBTV-DNA-4 (DNA-M) (BBJ34152.1) compared to five overseas BBTV strains shows amino acid differences in between.

At the level of BBTV-DNA-5 (DNA-C) documented in GenBank (BBJ34153.1), results in Figure (10) showed its ORF2 CDS translation in reading frame 3 (471 nts/156



aa) on the direct strand extends from base 240 to base 710 and encoding the cell cycle link protein gene of BBTV. Five differences in between were recorded when alignment

sequences of ORF of BBTV-DNA-5 (BBJ34153.1) compared to five overseas BBTV strains were analyzed (Fig. 11).

```

1 ATGGAGTTCTGGGAATCGTCTGCCATGCCTGACGATGTCAAGAGA
  M E F W E S S A M P D D V K R
46 GAGATTAAGGAAATATATTGGGAAGATCGGAAGAACTTCTGTTC
  E I K E I Y W E D R K K L L F
91 TCTCAGAAGTTGAAGAGCTATGTCTCAGAAGGATTCTTGCCTACGGA
  S Q K L K S Y V R R I L A Y G
136 GATCAGGAGGATGCCCTTGCTGGAGTGAAGGATATGAAGACTTCT
  D Q E D A L A G V K D M K T S
181 ATTATTCGCTATAGCGAATACTTGAAGAAACCATGTGTGGTAATT
  I I R Y S E Y L K K P C V V I
226 TGTGTGTGTAGCAATAAATCAATTGTGTATAGGTTAAACAGCATG
  C C V S N K S I V Y R L N S M
271 GTGTTCTTTTATCATGAATACCTTGAAGAACTAGGTGGTGATTAC
  V F F Y H E Y L E E L G G D Y
316 TCAGTATATCAAGATCTCTATTTGTGATGAGGTTCTCTCATCTTCA
  S V Y Q D L Y C D E V L S S S
361 TCGACAGATGAAGAAGATGTAGGAGTAATATACAGGAATGTTATC
  S T D E E D V G V I Y R N V I
406 ATGGCATCGACTCAAGAGAAGTTCTCTTGGAGTGATTGTTCAGAAG
  M A S T Q E K F S W S D C Q K
451 ATAGTAATTTCCAGACTATGA
  I V I S R L *

```

**Fig. 10:** ORF2 CDS translation in reading frame 3 (471 nts/156 aa) on the direct strand extends from base 240 to base 710 of BBTV-DNA-5 (DNA-M) documented in GenBank (BBJ34153.1).

BBJ34153.1	1	MEFWESSAMPDDVKREIKEIYWEDRKKLLFSQKLSYVRRILAYGDQEDALAGVKDMKTS	60
AKO70803.1	1	.....V.....	60
AKO70804.1	1	.....V.....L..	60
ARD25094.1	1	.....C.....	60
AGF92155.1	1	.....C.....L..	60
AKO70802.1	1	.....C.....V.....L..	60
BBJ34153.1	61	IIRYSEYLKPKCVVICCVSNKSIYRLNSMVFYFHEYLEELGGDYSVYQDLYCDEVLSSS	120
AKO70803.1	61	.....	120
AKO70804.1	61	.....	120
ARD25094.1	61	.....	120
AGF92155.1	61	.....	120
AKO70802.1	61	.....	120
BBJ34153.1	121	STDEEDVGVIIYRNVIMASTQEKFSWSDCQKIVIS	154
AKO70803.1	121	.....	154
AKO70804.1	121	.....	154
ARD25094.1	121	..E.....E.....	154
AGF92155.1	121	.. .E.....	154
AKO70802.1	121	.....	154

**Fig. 11:** Alignment sequences of ORF of BBTV-DNA-5 (DNA-M) (BBJ34153.1) compared to five overseas BBTV strains shows amino acid differences in between.

Finally, BBTV-DNA-6 (DNA-N) documented in GenBank (BBJ34154.1) was analyzed as shown in Figure (12) and its ORF2 CDS translation in reading frame 2 (390 nts/129 aa) on the direct strand extends from the base 281 to base 670 and encoding

nuclear shuttle protein gene of BBTV. One difference in between was recorded when alignment sequences of ORF of BBTV-DNA-6 (BBJ34154.1) compared to five overseas BBTV strains were analyzed (Fig. 13).

```

1 ATGGATTGGGCGGAATCACAAATCAAGACATGTACTCATGGATGC
  M D W A E S Q F K T C T H G C
46 GATTGGAAGAAGATATCATCGGATTCAGCCGATAATCGACAATAT
  D W K K I S S D S A D N R Q Y
91 GTACCATGGCTCGATTCTGGAGCTGGAAGAAAGTCGCCTCGCAAG
  V P C V D S G A G R K S P R K
136 GTACTTCTTAGATCTATTGAACTGTGTTTAACGGAAGCTTCAGC
  V L L R S I E T V F N G S F S
181 GGAAATAATAGGAACGTTCTGGATTCTCTACGTATCGATCAGA
  G N N R N V R G F L Y V S I R
226 GACGATGACGGAGAAATGCGTCCAGTACTCATAGTACCATTCGGA
  D D D G E M R P V L I V P F G
271 GGATATGGATATCATAATGATTTCTATTATTTCGAAGAAAGGGG
  G Y G Y H N D F Y Y F E G K G
316 AAAGTTGAATGTGATATATCATCAGATTATGTTGCGTCAGCCAGG
  K V E C D I S S D Y V A S A R
361 AATAGATTGGAGCAGAGACTTGGAAAGTTAG
  N R L E Q R L G S *
    
```

**Fig. 12:** ORF2 CDS translation in reading frame 2 (390 nts/129 aa) on the direct strand extends from base 281 to base 670 of BBTV-DNA-6 (DNA-N) documented in GenBank (BBJ34153.1).

BBJ34154.1	1	MDWAESQFKTCTHGC	60
AKO71098.1	1	.....A.....	60
AKO71071.1	1	.....A.....	60
AKO71116.1	1	.....A.....	60
NP_604479.1	1	.....A.....	60
AKO71035.1	1	.....A.....	60
BBJ34154.1	61	GNNRNRVGRFLYVSI	117
AKO71098.1	61	.....	117
AKO71071.1	61	.....	117
AKO71116.1	61	.....	117
NP_604479.1	61	.....	117
AKO71035.1	61	.....	117

**Fig. 13:** Alignment sequences of ORF of BBTV-DNA-6 (DNA-N) (BBJ34154.1) compared to five overseas BBTV strains shows amino acid differences in between.

Sequences producing significant alignments of ORF of BBTV-DNA-1 (DNA-R) encoding replication-associated protein gene (BBJ34149.1) compared to five overseas BBTV strains (AAZ07720.2, AKO71380.1, CAL90980.1, CAJ20054.1 and ALA65728.1) with query cover of 100% and E-value (0.0) was recorded in Table (2) with a similarity ranged from 99.64 to 100 %. This confirmed by the result of the genetic relationship (Fig. 14) between the ORF of DNA-R of BBTV-DNA-1 component and those most similar strains of BBTV in GenBank, as the DNA-R of this study was fell in a cluster with that of CAL90980.1.

When ORF of BBTV-DNA-2 (DNA-U3) (BBJ34150.1) was analyzed to the most five similar overseas BBTV strains, data of sequences producing significant alignments show the highest similarity (97.73%) was recorded against to RecName: Full=Uncharacterized protein U3 (P0C750.1) (Table 2). Results of phylogenetic tree of

ORF of DNA-U3 (BBJ34150.1) in Figure (14) based on the taxonomy names (Sq ID) show that it fell in a cluster with the Australian strain of BBTV (P0C750.1).

Data of sequences producing significant alignments (Table 2) show similarity ranging between 97.71 and 99.43 with a query cover of 100% and varied E-values were recorded between the ORF of BBTV-DNA-3 (DNA-S) (BBJ34151.1) and the most five BBTV strains in GenBank. The highest similarity was found when compared to coat protein (Banana bunchy top virus, NP\_604477.1) and fell together in a cluster (Fig. 14).

Regarding BBTV-DNA-4 (DNA-M) (BBJ34152.1), data in Table (2) show sequences producing significant alignments with similarity ranging between 87.25 and 88.24 % with query cover of 98-100% and varied E-value. The highest similarity was found when compared to coat protein (movement protein of Banana bunchy top

virus, AEL32419.1) and fell in a separate cluster (Fig. 14).

Data in Table (2) show sequences producing significant alignments with similarity ranging between 98.05 and 99.35% with varied E-value and query cover of 98% when the ORF of BBTV-DNA-5 (DNA-C) (BBJ34153.1) was analyzed compared to the most similar overseas BBTV strains. The highest similarity was found when compared to coat protein (cell cycle link protein of Banana bunchy top virus, AKO70803.1) and fell in an intermediate cluster (Fig. 14).

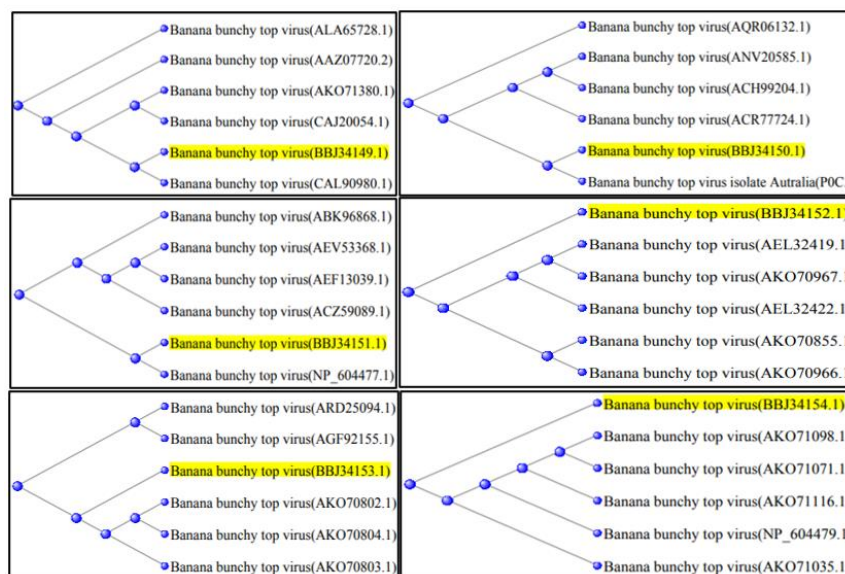
Sequences producing significant alignments (Table 2) show a similarity of 99.15% with varied E-value and query cover

of 95% between the ORF of BBTV-DNA-6 (DNA-N) (BBJ34154.1), and the five overseas representing the nuclear shuttle protein gene of BBTV and fell in a separate cluster (Fig. 14).

As a result of several studies of bioinformatics analyzes of the overseas strains of BBTV, it has been shown that the strains obtained from the same continental region showed simple variation in the nucleotide sequences of the BBTV genome, which provided the opportunity to be present together in the same group (Su *et al.* 2003, Selvarajan *et al.* 2010 and Banerjee *et al.* 2014).

**Table 2:** Sequences producing significant alignments of ORFs of BBTV-DNA components (DNA-R: 1109, DNA-U3: 1056, DNA-S: 1075, DNA-M: 1043, DNA-C: 1015 and DNA-N: 1084 nts) compared to some overseas BBTV strains.

Description	Blastp (protein-protein BLAST)		
	Query Cover (%)	Identities (%)	Accessions
<b>BBTV-DNA-1 (DNA-R) (BBJ34149.1)</b>			
replicase	100	100.0	AAZ07720.2
Replication associated protein	100	99.64	AKO71380.1
master replication associated protein	100	99.64	CAL90980.1
replicase protein	100	99.64	CAJ20054.1
replication initiation protein	100	99.64	ALA65728.1
<b>BBTV-DNA-2 (DNA-U3) (BBJ34150.1)</b>			
Full=Uncharacterized protein U3	100	97.73	P0C750.1
U3 protein	100	94.32	AQR06132.1
hypothetical protein	100	89.89	ANV20585.1
U3 protein	87	96.10	ACR77724.1
unknown	87	94.81	ACH99204.1
<b>BBTV-DNA-3 (DNA-S) (BBJ34151.1)</b>			
Coat protein	100	99.43	NP_604477.1
Coat protein	100	98.29	ABK96868.1
Coat protein	100	97.71	AEV53368.1
Coat protein	100	97.71	AEF13039.1
Coat protein	100	97.71	ACZ59089.1
<b>BBTV-DNA-4 (DNA-M) (BBJ34152.1)</b>			
movement protein	100	88.24	AEL32419.1
movement protein	100	88.24	AEL32422.1
movement protein	98	88.00	AKO70855.1
movement protein	98	88.00	AKO70966.1
movement protein	100	87.25	AKO70967.1
<b>BBTV-DNA-5 (DNA-C) (BBJ34153.1)</b>			
cell cycle link protein	98	99.35	AKO70803.1
cell cycle link protein	98	98.70	AKO70804.1
cell cycle link protein	98	98.05	ARD25094.1
cell-cycle link protein	98	98.05	AGF92155.1
cell cycle link protein	98	98.05	AKO70802.1
<b>BBTV-DNA-6 (DNA-N) (BBJ34154.1)</b>			
nuclear shuttle protein	90	99.15	AKO71098.1
nuclear shuttle protein	90	99.15	AKO71071.1
nuclear shuttle protein	90	99.15	AKO71116.1
NSP	90	99.15	NP_604479.1
nuclear shuttle protein	90	99.15	AKO71035.1



**Fig. 14:** Phylogenetic tree of ORF of BBTV-DNA components (DNA-R, DNA-U3: 1056, DNA-S: 1075, DNA-M: 1043, DNA-C: 1015 and DNA-N: 1084 nts) compared to some overseas BBTV strain, shows taxonomy names (Sq ID).

### Determining the type of domains and their families

The ORF of BBTV-DNA-1 (DNA-R) with a length of 278 aa, encoding the gene of replicas associated (BBJ34149.1) has two domains (Fig. 15). The first domain is the viral rep domain or putative viral replication protein which begins from 1-75 and belongs to the family of viral ORFs from various plant and animal ssDNA circoviruses. This domain has a biological process it has a role in DNA replication and protein-DNA covalent cross-linking. The second domain is RNA\_helicase started from 168 to 260 and belongs to a family that includes RNA helicases thought to be involved in duplex unwinding during viral RNA replication. This domain has a molecular function the function are endo-deoxyribonuclease activity, producing 5'-phosphomonoesters, nucleotidyl-transferase activity, RNA binding and RNA helicase activity.

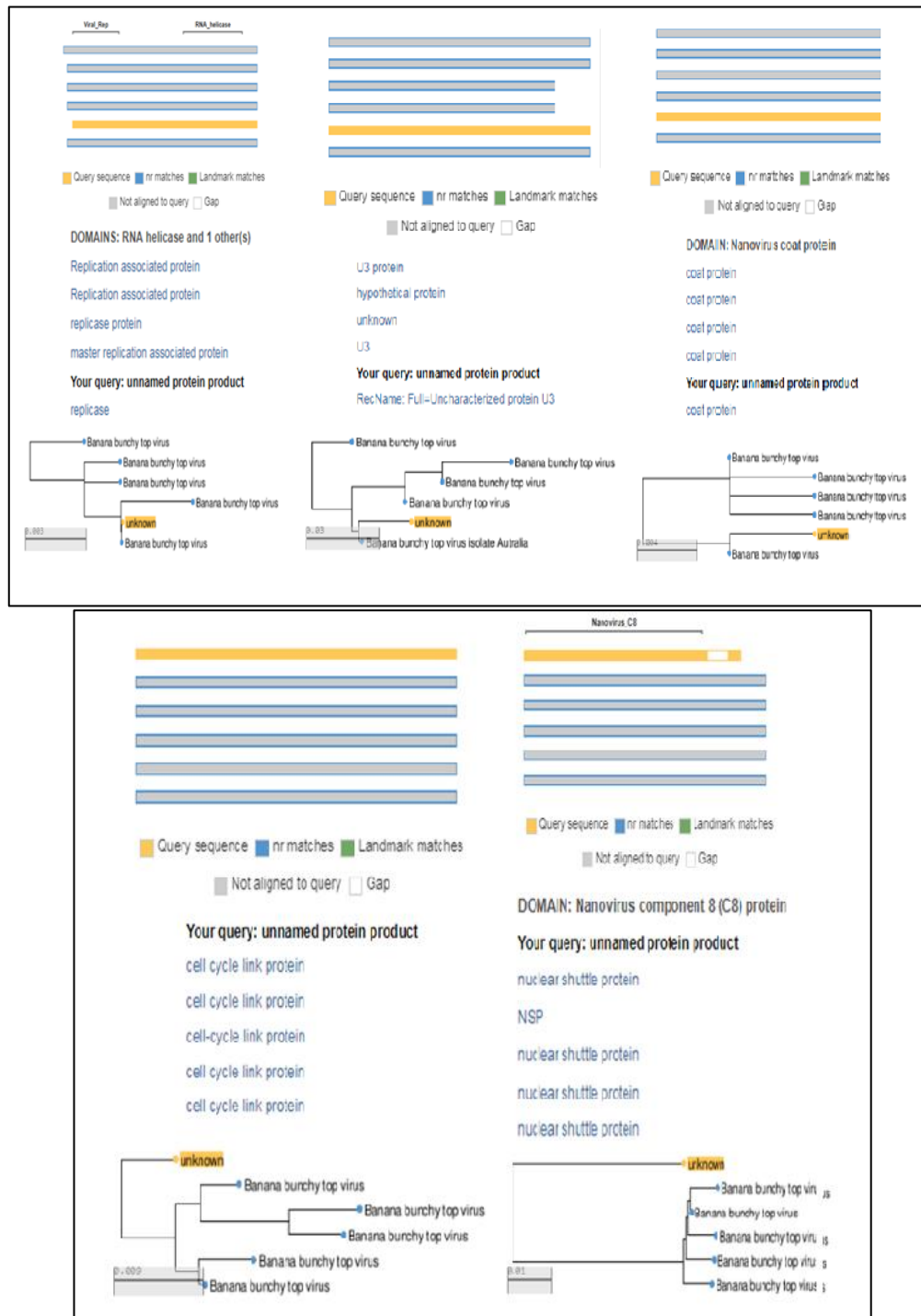
Data in Figure (15) show that BBTV-DNA-2 (DNA-U3) component which encodes the U3 protein based on its ORF with a length of 88 aa, has a hypothetical protein belonging to the uncharacterized protein U3 domain.

BBTV-DNA-3 which encodes the coat protein gene based on an ORF with a length of 175 aa has one domain called nanovirus\_coat the protein family membership is Nanovirus coat protein (Fig. 15). While BBTV-DNA-4 (DNA-M) has no motifs. This protein has one domain Babovirus\_MP beginning from 1-101 (Figure 15). The protein superfamily of this domain is a membership of movement and RNA silencing protein.

The type of ORF domain of BBTV-DNA-5 (DNA-C) component (BBJ34153.1) which represented the replication-associated protein gene of BBTV was belonging to the cell cycle protein domain belonging to cell cycle link protein Clink (Fig. 15). This segment has 155 aa, encoding cell cycle protein and has no motifs. Finally, BBTV-DNA-6 component (DNA-N) (BBJ34154.1) represented the nuclear shuttle protein gene of BBTV belonging to the nuclear shuttle protein domain (Fig. 15). Its ORF has 129 aa representing a protein that has no motifs. This protein has one domain beginning from 2-119 and refers to Nanovirus\_C8 (Nanovirus component protein).

These results are in harmony with that reported by Sadik (1994), Burns *et al.* (1995),

Su *et al.* (2003), Nour El-Din *et al.* (2005), Beetham *et al.* (1997), Hafner *et al.* (1997), Vetten *et al.* (2005), Salama *et al.* (2007), Amin *et al.* (2008), Mandal (2010), Selvaranjan *et al.* (2010), King *et al.* (2012) and Vetten *et al.* (2012).



**Fig. 15:** Analyses of ORFs of the six BBTV-DNA components (DNA-R1 (BBJ34149.1), DNA-U3 (BBJ34150.1), DNA-S (BBJ34151.1), DNA-M (BBJ34152.1), DNA-C (BBJ34153.1), and DNA-N (BBJ34154.1)) belonging to different domains and families.

## REFERENCES

- Adegbola, R.A., O. Ayodeji; O.O. Awosusi, G.I. Atiri and P.L. Kumar (2013). First report of banana bunchy top disease caused by Banana bunchy top virus in banana and plantain (*Musa spp.*) in Nigeria. *Plant Disease*, 97: 290.
- Almedia, R.P.; G.M. Bennett; M.D. Anhalt; C.W. Tsai and P.O. Grady (2009). Spread of an introduced vector-borne banana virus in Hawaii. *Molecular Ecology*, 18:136-146.
- Amin, I.; J. Qazi; S. Mansoor; M. Ilyas and R.W. Briddon (2008). Molecular characterization of Banana bunchy top virus (BBTV) from Pakistan. *Virus Genes*, 36:191-198.
- Balakrishnan, S.; C. Gokulapalan and S. Paul (1996). A widespread banana malady in Kerala, India. *Infomusa*, 5: 28-29.
- Banerjee, A.; S. Roy; G.T. Behere; S.S. Roy; S.K. Dutta and S.V. Ngachan (2014). Identification and characterization of a distinct banana bunchy top virus isolate of Pacific-Indian Oceans group from North-East India. *Virus Research*, 2: 41-49.
- Beetham, P.R.; G.J. Hafner; R.M. Harding and J.L. Dale (1997). Two mRNAs are transcribed from Banana bunchy top virus DNA-1. *Journal of General Virology*, 78: 229-236.
- Burns, T.M.; R.M. Harding and J.L. Dale (1995). The genome organization of banana bunchy top virus: analysis of six ssDNA components. *Journal of General Virology*, 76:1471-1482.
- Hafner, G.J.; R.M. Harding and J.L. Dale (1997). A DNA primer associated with Banana bunchy top virus. *Journal of General Virology*, 78: 479-486.
- Harding, R.M.; T.M. Burns; G. Hafner; R.G. Dietzgen and J.L. Dale (1993). Nucleotide sequence of one component of the Banana bunchy top virus genome contains a putative replicase gene. *Journal of General Virology*, 74: 323-328.
- Harding, R.M.; A. Bahieldin, A.S. Sadik and J.L. Dale (2000). A sensitive detection of Banana bunchy top nanovirus using molecular genetic approaches. *Arab Journal of Biotechnology*, 3: 103-114.
- King, A.M., E. Lefkowitz; M.J. Adams and E.B. Carstens (Eds.) (2011). Virus taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses Elsevier, 9: 21-26.
- Lowe, S.; M. Browne; S. Boudjelas and M. De Poorter (2000). 100 of the world's worst invasive alien species: a selection from the global invasive species database (Vol. 12). Auckland: Invasive Species Specialist Group.
- Lowe, S.; M. Browne; S. Boudjelas and M. De Poorter (2000). 100 of the World's worst invasive alien species. A selection from the Global Invasive Species Database. Published by The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN) 12 p.
- Mandal, B., (2010). Advances in small isometric multicomponent ssDNA viruses infecting plants. *Indian Journal of Virology*, 21: 18-30.
- Molina, A.B. and C. Kudagamage (2002). The international network for the improvement of banana and plantain (INIBAP): PGR activities in South Asia. In South Asia network on plant genetic resources (SANPGR) meeting held on December, pp. 9-11.
- Mpoki, S.; G. Mahuku; D.R. Mbanzibwa; G. Mkamilo; D. Mark; H. Mosha; B.D. Pallangyoyo; M. Fihavango; A. Oresanya; P. Ogunsanya and P.L. Kumar (2021). First report of Banana bunchy top virus in banana and plantain (*Musa spp.*) in Tanzania. *Plant Disease*, 27: 1-7.
- Nour El Din, Hanan; M.I. Salama; A.B. Barakat; A.M. Salem and A.S. Sadik (2005). Nucleotide sequence of

- BBTV-*cp* gene and using its fusion protein for producing specific polyclonal antibodies. *Arab Journal of Biotechnology*, 8: 353-368.
- Sadik, A.S. (1994). Studies on viruses affecting banana in Egypt. Ph.D., Dissertation in Agric. Virology, Dept. Agric. Microbiol., Faculty of Agric., Ain Shams Univ., Cairo, Egypt, 199 p.
- Selvarajan, R.; M.M. Sheeba; V. Balasubramanian; R. Rajmohan; N.L. Dhevi and T.Sasireka (2010). Molecular characterization of geographically different Banana bunchy top virus isolates in India. *Indian Journal of Virology*, 21: 110-116.
- Singh, V.; S. Adil and A. Quraishi (2022). Elimination of BBTV via a systemic *in vitro* electrotherapy approach. *Journal of Virological Methods*, 300:114367.
- Su, H.J.; L.Y. Tsao; M.L. Wu and T.H. Hung (2003). Biological and molecular categorization of strains of Banana bunchy top virus. *J. Phytopathol.*, 151:290–296.
- Vetten, H.J.; P.W. Chu; J.L. Dale; R.M. Harding; J. H; L. Katul; M. Kojima; J.W. Randles; Y. Sano and J.E.Thomas (2005). Nanoviridae. In: Fauquet C.M., Mayo M.A., Maniloff J., Desselberger U., Ball L.A. (eds) *Virus taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*. Academic Press, San Diego, pp. 343-352.
- Vetten, H.J.; J.L. Dale; I. Grigoras; B. Gronenborn; R. Harding and J.W. Randles (2012). Family Nanoviridae. In *Virus taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*, pp. 395-404.
- Vincze, T.; J. Posfai and R.J. Roberts (2003). NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic acids research*, 31(13): 3688-3691.
- Yu, N.; T. Feng; Y. Zhang; J. Wang and Z. Liu (2011). Bioinformatic analysis of BBTV satellite DNA in Hainan. *Virologica Sinica*, 26: 279-284.