EGYPTIAN ACADEMIC JOURNAL OF
BIOLOGICAL SCIENCES
MICROBIOLOGY

G

ISSN
2090-0872

WWW.EAJBS.EG.NET

Vol. 14 No. 2 (2022)

Citation: *Egypt. Acad. J. Biolog. Sci. (G. Microbiolog) Vol.14 (2) pp.1- 18 (2022)*
DOI: 10.21608/EAJBSG.2022.248724

Egypt. Acad. J. Biolog. Sci., 14(2):1-18(2022)

Egyptian Academic Journal of Biological Sciences
G. Microbiology

ISSN: 2090-0872
https://eajbsg.journals.ekb.eg/

Bioinformatics Analyses of The Complete DNA Genome of An Egyptian Isolate of *Banana bunchy* Top Virus

Abdel Razek Fatma S.[1], El-Masry Samar S.[1*], Ibrahim S.D.[2], El-Arabi T.F.[1], Mahmoud E. Khalifa[3] and Sadik A.S.[1]

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

2-Agricultural Genetic Engineering Research Institute, Agricultural Research Center, 9 Gamaa st., P.O. Box, 12619, Giza, Egypt

3-Botany and Microbiology Department, Faculty of Science, Damietta University, Damietta, Egypt

*E.Mail: samar_sayed@agr.asu.edu.eg

ARTICLE INFO

Article History
Received: 2/5/202
Accepted:5/7//2022
Available:8/7/2022

***Keywords*:**
BBTD, BBTV, Bioinformatics analyzes, Genome, Domains, Nanoviruses.

ABSTRACT

Banana bunchy top disease (BBTD) caused by *Banana bunchy top virus* (BBTV) was recorded in the early twentieth century in Egypt. It is considered the most economic impact on banana yield productivity. Bioinformatics analyses of the complete genome of an Egyptian isolate of BBTV compared to overseas BBTV isolates or strains and other Nanoviruses were aimed. Banana leaf samples naturally infected with BBTD were collected from the open field and the presence of BBTV was detected *via* PCR. Bioinformatics analyzes of the BBTV-DNA genome were also studied. The experimental results showed that BBTV isolates of this study had a genome that consists of six encapsidated single-stranded DNA components (BBTV-DNA-R (LC468138, -U3 (LC468139), -S (LC468140), -M (LC468141), - C (LC468142) and –N (LC468143)) of ~ 1.1 kb in length, each with one open reading frame (ORF) in the virion-sense Two conserved regions (Common region-stem loop (CR-SL) and common region-M (RC-M)). TATA box, *i.e*., non-nucleotide potential TATA and polyadenylation signals adjacent to GC-rich regions that contain the TTG sequence were found. Differences ranging from 0.55 to 4.60% were recorded when BBTV components were compared to the most similar oversea BBTV strains of seven countries. Phylogenetic trees confirmed the genetic relationships among the high percentage of similarity between the compared BBTV strains. Families of Gene domains, restriction enzyme maps and differences between BBTV components and those of nano viruses were also discussed. The six BBTV-DNA components contain ORFs encoding the genes of rolling-circle replication initiation protein (*rp*, DNA-R), a protein of unknown function (DNA-U3), a coat protein peptide (*cp*, DNA-S), movement protein (*mv*, DNA-M), cell cycle link protein (*ccl*, DNA-C) and a nuclear shuttle protein (*ns*, DNA-N), respectively. Further studies should be done using molecular docking tools in a trial to find a suitable control strategy for BBTV.

INTRODUCTION

Viral diseases are considered as one of the most affecting diseases on the productivity of the banana plant due to the losses it causes to production as well as the quality of the banana fruits, in addition to the difficulty of exchanging the banana seedlings between the countries worldwide (Jones 2000, Amani and Avagyan 2014, Kumar *et al.,* 2015, Qazi 2016, Hamim *et al.,* 2017, Sila *et al.,* 2020 and Rahayuniati *et al.,* 2021a). Banana bunchy top virus (BBTV) has been registered as one of the world's 100 oldest known disease pathogens Lowe *et al.,* (2000).

BBTV particles are characterized by an icosahedral shape with a coat protein of about 20 KDa (Harding *et al.,* 1993, Nour El-Din *et al.,* 2005 and Rahayuniati *et al.,* 2021b). The viral genome was characterized as a circular single-strand DNA (ssDNA) (Harding et al 2000, Salama *et al.,* 2007 and Amin *et al.,* 2008) and belongs to the genus *Babuvirus* (Mandal 2010) and the family Nanoviridae (Vetten *et al.,* 2005, King *et al.,* 2012 and Vetten *et al.,* 2012). The BBTV particles contain multiple genomes of six single circular segments of ssDNA with lengths ranging from 1000 to 1100 nucleotides, each of them contains a stem-loop common region with a length of 69 nucleotides and they are referred to as DNA-1 to DNA-6 (Burns *et al.,* 1995, Xie and Hu 1995 and Harding *et al.,* 2000). These segments were renamed DNA-R, -U3, -S, -M, -C and –N, and the DNA-R segment was coded into two open reading frames and the rest of the others each encoded into one protein (Burns *et al.,* 1995, Beetham *et al.,* 1997, Hafner *et al.,* 1997). It has been shown that the three DNA segments DNA-R, -S and -M shown to be responsible for the genes of replication, coat protein and movement protein, respectively (Burns *et al.,* 1995).

Bioinformatics analysis of several overseas BBTV isolates showed that isolates or strains from the same area appeared common variation in the sequences of BBTV genome and used to be found together in a specific cluster among phylogenetic analysis (Su *et al.,* 2003, Vishnoi *et al.,* 2009, Selvaranjan *et al.,* 2010, Banerjee *et al.,* 2014). Moreover, high similarities among the nucleotide sequences and phylogenetic analyses were shown between BBTV isolates obtained from geographically neighboring countries in which an exchange of banana germplasm or suckers occurred compared with those that are somewhat far from each other (Molina and Kudagamage 2002).

BBTV isolates of different investigations were bioinformatics analyzed compared to the Pacific Indian Ocean BBTV group (Selvaranjan *et al.,* 2010 and Banerjee *et al.,* 2014). Genetic variability between BBTV isolates obtained from twelve banana cultivars exhibiting the characteristic symptoms of banana bunchy top disease (BBTD), collected from different regions of Assam, India, based on bioinformatics analysis was determined (Kakati et al 2018). They reported that BBTV isolates of Assams showed similar percentages ranging from 86 to 99 when compared to the Pacific Indian Ocean BBTV group.

At the level of satellite of BBTV, Yu *et al.* (2011) indicated that bioinformatics analysis of BBTV sequencing of Hainan isolates represents a satellite DNA component with 12 DNA sequence 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 was aimed at determining the complete nucleotide sequence of the circular ss-DNA genome of an Egyptian strain of BBTV followed by its bioinformatics analyses compared to BBTV overseas strains and other Nanoviruses.

MATERIALS AND METHODS

1. Source of Banana Samples:

Banana leaf samples of healthy and those exhibiting BBTD-like symptoms were

collected from different open field farms of Qaliobia Governorate in 2018. All banana leaves of different cultivars (Edward Cavendish, Williams Grand Nain and Maghrabi) were collected and put on plastic pages followed by storing at −20°C until used.

2. Presence of BBTV in Banana Collected Samples:

2.1 DNA Extraction:

DNA extracts were prepared from banana leaf samples by using the DNeasy Plant Mini Kit (Qiagen Santa Clarita, CA), following the manufacturer's instructions. The final DNA concentration was adjusted to 100 ng per µL for PCR amplification as a template.

2.2 PCR-detection of BBTV:

The presence of BBTV-DNA was detected in the DNA extracts of collected banana leaf samples that were used as templates for PCR based on the protocol described by Anandhi *et al.* (2007). One primer pair amplifying the full length of BBTV-DNA component 1 (DNA-R) (F1: 5'ATG GAG GAG AAG GAA AGA CAA3' and R1: TTG TCT TTC CTT CTC CTC CAT3') was designed and synthesized based on the nucleotide sequence of BBTV-DNA Australian isolate (S56276.1).

2.3 Electrophoresis of DNA Extracts and PCR Products:

Each DNA extract as well as PCR amplified products were visualized on UV light after agarose gel electrophoresis (1.5%) as described by **Sambrook *et al.* (1989)** in the presence of a DNA marker (100 bp DNA ladder or 1 Kb DNA ladder) and photographed using a Gel Documentation System (BIO-RAD).

3. Sequencing of BBTV-DNA Components:

For determining the nucleotide sequences of full-length BBTV-DNA components (DNA-R, DNA-U3, DNA-S, DNA-M, DNA-C and DNA-N), banana leaf samples infected with BBTV were dried in calcium chloride and cotton as shown in Fig. 1 as described by Prof. James Dale (Personal communication).

Fig 1. BBTV-infected banana leaf samples were dried using Calcium chloride and sent to Macrogen Company for determining the nucleotide sequences of BBTV-DNA components.

Additionally, six primer pairs specific to the six BBTV-DNA components were designed and synthesized based on the nucleotide sequences of the BBTV Australian strain (Harding *et al.*, 1993 and Burns *et al.*, 1995) flanking the complete nucleotide sequences as shown in Table 1. The PCR products of the six BBTV-DNA components were amplified as described by Harding *et al.* (1993) for DNA-R and Burns *et al.* (1995) for the other five DNA components. The PCR products were dried followed by sending plus the dried BBTV-infected banana leaf samples to Macrogen Company. The nucleotide sequences of the six BBTV-DNA components were determined by using Macrogen® (908 World Meridian Venture Center, #60-24, Gasan-dong, Geumchun-gu, Seoul 153-781, Korea).

Table 1. Primer pairs were designed and synthesized based on the sequences of BBTV-DNA components of Australian isolate used for PCR amplification of the six BBTV-DNA components.

BBTV components	Primer		Sites	Lengths (nts)	Accessions/References
	Codes	Sequences (5' …….. 3')			
DNA-1	F1	ATGGAGGAGAAGGAAAGACAA	673-693	21	S56276.1 Harding *et al* (1993)
(DNA-R)	R1	TTGTCTTTCCTTCTCCTCCAT	672-652	21	
DNA-2	F2	CTCCATCGGACGATGGAGGTTG	1018-1040	22	L41576.1 Burns *et al* (1995)
(DNA-U3)	R2	GCATCCAACGGCCCATA	1017-1000	17	
DNA-3	F3	GGGTATCTGATTATGTATCCTAAC	831-854	24	L41574.1 Burns *et al* (1995)
(DNA-S)	R3	GTACTTTGTCATAGTGT	830-813	18	
DNA-4	F4	TGGTATATGATTAGGTATCCTAACGA	788-814	26	L41575.1 Burns *et al* (1995)
(DNA-M)	R4	GTACTTTTGTCATAGTGT	797-781	18	
DNA-5	F5	AAGAGCCATGGAGTTCTGGGAATC	233-256	24	L41578.1 Burns *et al* (1995)
(DNA-S)	R5	CCGACGAGTGATTTCGGAAATCAC	232-208	24	
DNA-6	F6	TATTAGTAACAGCAACA	667-684	17	L41577.1 Burns *et al* (1995)
(DNA-N)	R6	CTAACTTCCAAGTCTCT	666-650	17	

4. Bioinformatics of BBTV-DNA Components:

Analyses of nucleotide sequences of BBTV-DNA components and its open reading frames were conducted using BLASTN 2.2.23 (http://www.ncbi.nlm.nih.gov/blast/) compared the BBTV overseas strains documented in GenBank with the lowest e-values and maximum identities. Multiple comparisons (pairwise alignment, multiple sequencing alignment, nucleotide percentages, restriction enzyme endonuclease maps (closest enzymes at 5' and 3' ends) were done as described by Jotun Hein Algorithms (Hein 1990) and the method given by Higgins and Sharp (1989).

RESULTS AND DISCUSSION

1. Isolation of *Banana bunchy* Top Virus:

It has been shown that BBTV is the causal agent of BBTD in bananas worldwide (Kumar *et al.*, 2009, Almedia *et al.*, 2009, Adegbola *et al.*, 2013, Wickramaarachchi *et al.*, 2016, Qazi 2016, Mpoki *et al.*, 2021, Singh *et al.*, 2022).

Results of Figure 2 showed some banana plants belonging to three cultivars (Edward Cavendish, Williams Grand Nain, and Maghrabi) exhibiting BBTD-like characteristic symptoms. These symptoms were represented as dark green streaks along the small veins on the underside of the leaves, midribs and stalks. The streaks consist of dots and dashes in a Morse code pattern which form J-shaped hooks where they join the midrib. In older banana plants dark green streaks were seen on flower bracts. Mature banana plants infected with BBTD may produce stunted and deformed bunches with worthless fruits. Banana black aphid (*Pentalonia nigronervosa* Conq.), the vector of BBTV, and the causal agent of BBTD, under open field conditions, was observed hiding in the sheaths of banana leaves (Fig 2).

2. Detection of BBTV Presence in Banana Collected Samples:

To achieve such an aim, total DNA was extracted and purified from collected banana cultivar samples. The purified DNA was electrophoresed on a 1.5% agarose gel to ensure its purity (1.86) and concentration (87 ng μL^{-1}) (Fig 3). DNA appeared with no digestion and in varying concentrations from one cultivar to another. Approximately 100 ng of each DNA extract was used as a template for PCR detection of BBTV in collected banana plant samples using one primer pair targeting BBTV-DNA-1 component (the master replication gene, BBTV-DNA-R). As shown in Figure 4, PCR products confirmed the presence of BBTV in all banana samples. No PCR product(s) was amplified from the healthy banana plant that was used as a negative control. This result is in harmony with that found by Xie and Hu (1995) and Harding *et al.* (2000).

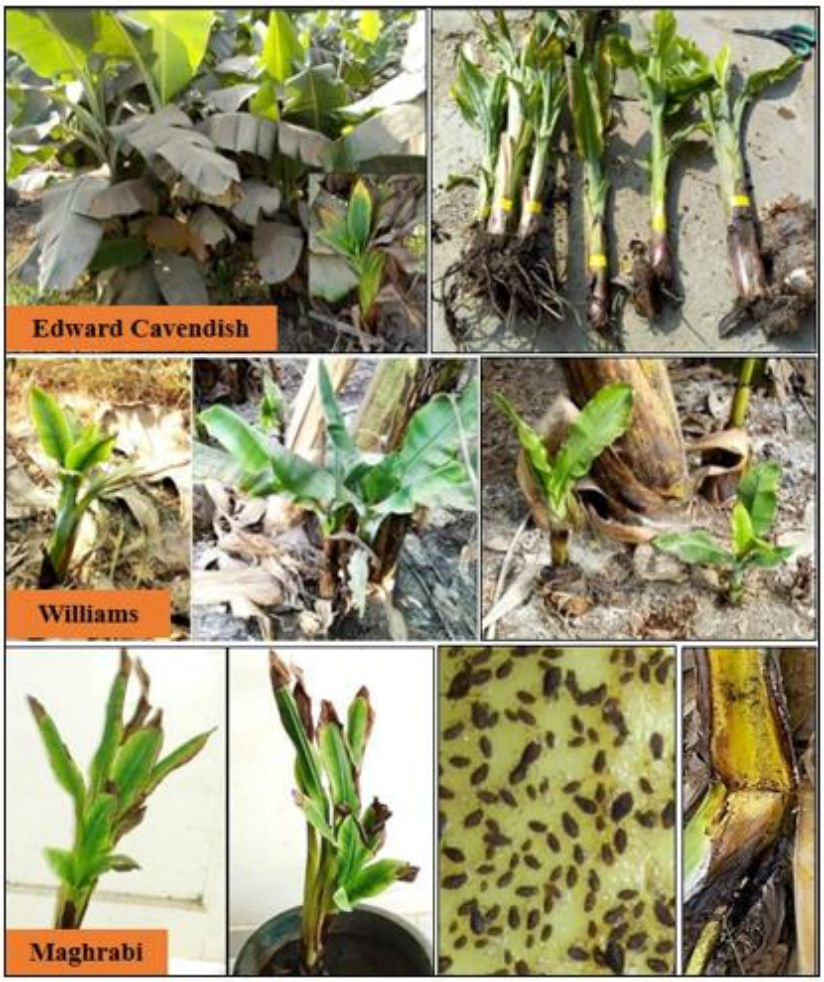

Fig 2. Banana plants cvs. Edward Cavendish, Williams and Maghrabi naturally infected with BBTD, collected from an open field private farm in Qaliobia Governorate, as a source for BBTV. Note, that the presence of *Pentalonia nigronervosa* Conq., the vector of BBTV occurred on the cigar and sheath leaves of cv. Williams banana plants.

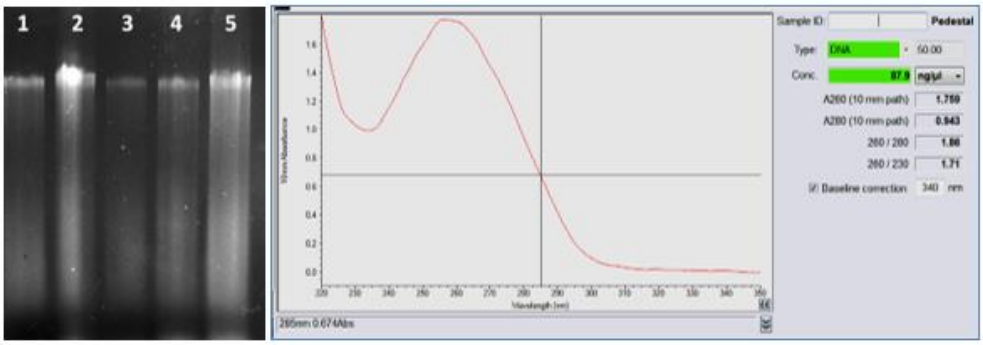

Fig 3. Agarose gel (1.5%) electrophoresis stained with ethidium bromide shows total DNA extracts prepared from five naturally BBTD-infected banana plant samples (Lanes, 1-5) collected from Qaliobia Governorate. Note: The purity of DNA extract and its concentration are illustrated.

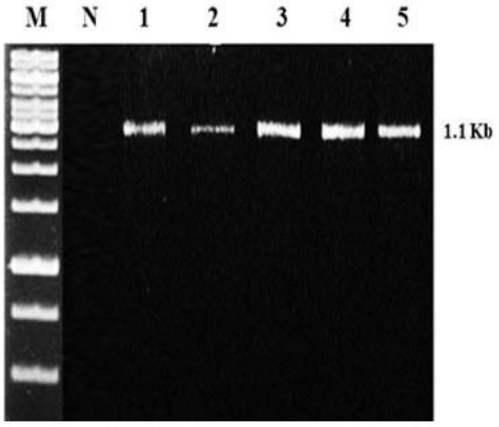

Fig 4. Agarose gel (1.5%) electrophoresis stained with ethidium bromide shows PCR detection of BBTV presence in naturally BBTD-infected banana plant samples (Lanes, 1-5) using a primer pair specific to the BBTV-DNA-R component. M: 100 bp Plus DNA Ladder. N: Negative control (PCR mixture with DNA of the healthy banana plant as a template).

4. PCR Amplification of BBTV-DNA Components:

The PCR-positive DNA extract was used as a template for PCR amplification of the six BBTV-DNA components using the six primer pairs with lengths ranging from 17 to 26 nucleotides. As shown in Figure 5, six DNA components (DNA-R, DNA-U3, DNA-S, DNA-M, DNA-C and DNA-N) with sizes of about 1.1 Kb were amplified and visualized after electrophoresis on a 1.5% agarose gel. No PCR product(s) was amplified from the negative control due to the absence of the BBTV-DNA template. Similar results were obtained by Harding *et al.* (1993) and Burns *et al.* (1995).

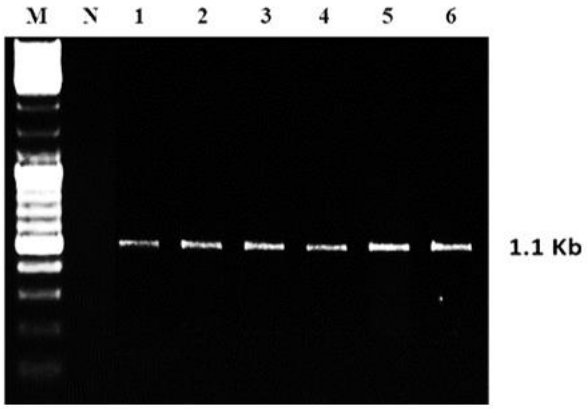

Fig 5. Agarose gel (1.5%) electrophoresis stained with ethidium bromide shows PCR products amplified from DNA extracted from banana plants naturally infected with BBTV using 6 primer pairs (Table 1) belonging to BBTV-DNA components, 1, 2, 3, 4, 5 and 6, respectively. M: 1 kb Plus DNA Ladder. N: Negative control (PCR mixture without DNA template).

5. Nucleotide Sequences of BBTV-DNA Components:

Following PCR, products of the six BBTV-DNA components were lyophilized and sent to Macrogen Company, Seoul, South Korea to determine their nucleotide sequences. BBTV-infected banana samples were just also sent as a source for DNA extraction if needed. The complete nucleotide sequences of the six BBTV DNA components were determined and documented in GenBank with accession numbers of LC468138.1 (DNA-1, represents DNA-R), LC468139.1 (DNA-2, represents DNA-U3),

LC468140 (DNA-3, represents DNA-S), LC468141 (DNA-4, represents DNA-M), LC468142 (DNA-5, represents DNA-C), and LC468143 (DNA-6, represents DNA-N). DNA-R) (Fig. 6). The length of DNA-R component was found to be 1109 nucleotides with 43% of G+C content. DNA-U3 consisted of 1056 nucleotides with a G+C content of 39%. DNA-S was found to have 1075 nucleotides long with 43% G+C content. DNA-M contained 1043 nucleotides and a G+C content of 40%. DNA-C contained 39% G+C and 61% A+T in a length of 1015 nucleotides. Finally, the sixth BBTV-DNA component (DNA-N) had a size of 1084 nucleotides and contained a G+C of 40%. The presence of nucleotide sequences of the six primer pairs that were used for PCR amplification and sequencing of the six BBTV-DNA components (DNA-R, -U3, -S, -M, -C and -N) was confirmed. Result in agreement with that of Burns *et al.* (1995).

Fig 6. Nucleotide sequences of the six BBTV-DNA components are documented in GenBank.

6. Genome Organization of BBTV-DNA Components:

The BBTV strain of this study had a genome that consists of six encapsidated single-stranded DNA components (BBTV-DNA-R, -U3, -S, -M, - C and –N) of ~ 1.1 kb in length, each with one open reading frame (ORF) in the virion-sense Table 2 and Figure 7. A conserved stem-loop structure (from nucleotide 1 to 30) was located at the 5' terminus of the large virion-sense ORF in all BBTV-DNA components. Two conserved common regions (CR) of the six BBTV-DNA components were scored. The 1st is called a common region-stem loop (CR-SL) starts from nucleotide number 1 to 44 (components DNA-R, DNA-S, DNA-M, DNA-C and DNA-N) or from nucleotide number 2 to 44 (component DNA-U3) in the non-coding region, with differences ranged from 4.54 to 13.6% among the six BBTV-DNA components (**Fig. 7**). The stem-loop structures (30 nucleotides) were incorporated in the CR-SL region and represent 68.18 % of it.

The 2nd is called common region-M (RC-M), which was recorded in five BBTV-DNA components (components DNA-R, DNA-U3, DNA-S, DNA-M and DNA-C) and absent in DNA-N component. This RC-M was divided into three regions named CR-M1, CR-M2 and CR-M3, which varied in their

presence among the five BBTV-DNA components (DNA-R, DNA-U3, DNA-S, DNA-m and DNA-N). CR-M2 and CR-M3 were absent in DNA-C component, whereas CR-M3 was not found in each of BBTV-DNA components DNA-U3 and DNA-S. The positions of the three CR-M regions among the five BBTV-DNA components also varied, CR-M1 and CR-M3 were located before the CR-SL in each of BBTV-DNA components DNA-S, DNA-M and DNA-C. CR-M1 and CR-M2 were located within the ORF in BBTV-DNA-R and after the ORF in DNA-U3 component. In BBTV-DNA component 3 (DNA-S), CR-M1 and CR-M2 were located after the ORF, and one of them (CR-M2) was before the CR-SL region. Finally, CR-M1 and CR-M2 in the BBTV-DNA component 4 (DNA-M) were located before and after the ORF, respectively.

Results in Figure 7 showed the presence of TATA box, *i.e*., non-nucleotide potential TATA and polyadenylation signals adjacent to GC-rich regions that contain the TTG sequence. The TATA box among the five BBTV-DNA components DNA-R, DNA-S, DNA-M, DNA-C and DNA-N was located upstream of the major ORFs of these components while being located with the ORF of DNA-U3 in BBTV-DNA component 2 (DNA-U3). In addition, a number of ORFs was demonstrated in BBTV-DNA-U3 and TATA boxes, or their associated polyadenylation signals were not associated with any major ORF in component-2 (DNA-U3). In other means, component-2 (DNa-U3) had a potential TATA box as well as polyadenylation signal location at the expected positions, and encoding a regulatory function of U3 protein.

Results in Table 2 showed that ORF of BBTV-DNA component-1, named DNA-R, is 837 nts and encodes a rolling-circle replication initiation protein. The ORF of BBTV-DNA component-2 (267 nts), named DNA-U3, encodes a protein of unknown function. The ORF of BBTV-DNA component-3 (528 nts), called DNA-S, represents a coat protein-peptide with an estimated molecular mass of 20 kDa. The ORF of BBTV-DNA segment-4 (309 nt), called DNA-M, represents the movement protein. The ORF of BBTV-DNA component-5 (471 nts), named DNA-C, encodes cell cycle link protein. Finally, the ORF of BBTV-DNA component-6 (390 nts), called DNA-N, encodes a nuclear shuttle protein.

Table 2. Genomic features of the six BBTV-DNA components under investigation.

Characters	**BBTV-DNA components**					
	DNA-R	**DNA-U3**	**DNA-S**	**DNA-M**	**DNA-C**	**DNA-N**
	LC468138	**LC468139**	**LC468140**	**LC468141**	**LC468142**	**LC468143**
Functions	**Replication associated protein**	**U3 protein**	**capsid protein**	**movement protein**	**cell-cycle link protein**	**nuclear shuttle protein**
Length	1109	1056	1075	1043	1015	1084
Stem	1-10	1-10	1-10	1-10	1-10	1-10
Loop	11-20	11-20	11-20	11-20	11-20	11-20
Stem	21-30	21-30	21-30	21-30	21-30	21-30
TATA-Box	51-59	256-264	195-203	258-266	188-196	223-231
ORF	127-963	143-409	213-740	279-587	240-710	281-670
CR-SL	1-44	2-44	1-44	1-44	1-44	1-44
CR-M1	743-777	455-486	813-882	195-214	990-1015	…….
CR-M2	910-929	758-829	1050-1075	781-850	…….	…….
CR-M3	1001-1105	…….	…….	1018-1043	…….	…….

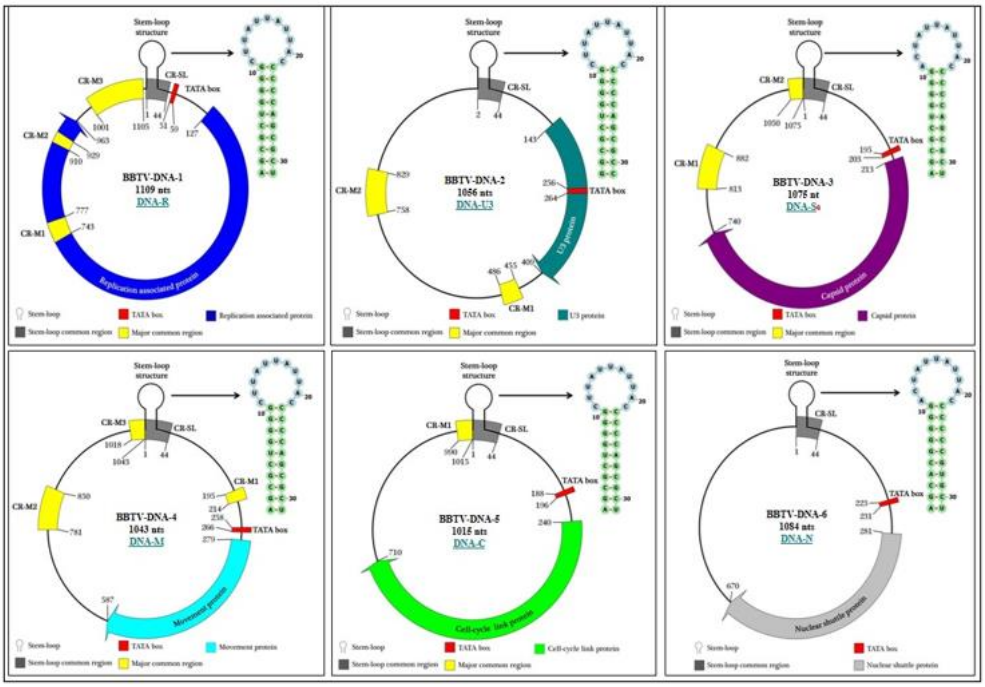

Fig 7. Organization map of BBTV components (DNA-R, DNA-U3, DNA-S, DNA-M, DNA-C and DNA-N)).

7. Bioinformatics Analyses of the six BBTV-DNA Components:

The concept of bioinformatics lies in the study of biological molecules in-depth with the aim of trying to know and understand the details of the origin of these molecules and their evolution and the genetic diversity between the serotypes and the interpretation of mutations that can occur as a result of exposure of these molecules to different mutagenic factors (Rana and Vaisla 2012). 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), and genetic protein structures and protein domains and their families, molecular networks and proteomics.

Multiple sequence alignments of BBTV-DNA-R component (LC468138.1), representing the *rep* gene encodes the replication-associated protein Table 3, and other overseas strains revealed that it is most similar to strains from six countries, *i.e.*, Egypt (AF416465.1, HQ259074.1 and AF102780.1), India (Delhi) (HM120718.1), SriLanka (JN250593.1), Rwanda (JQ820459.1), USA (NC_003479.1) and Australia (S56276.1) documented in GenBank. The identities between the BBTV-DNA-R and the eight most similar strains ranged from 97.39 to 98.56 %. Seven out of the eight strains had a similar length (1111 nts). The length of DNA component 1 (DNA-R) of the Egyptian strain of this study (LC468138.1) was different than that of the three Egyptian strains (accession number AF416465.1, 1111 nts, accession number HQ259074.1, 1108 nts and accession number AF102780.1, 1111 nts) compared to it Table 4. As shown in Table 3, BBTV-DNA component 2 (DNA-U3) encodes U3 protein, shared nucleotide sequence identities ranging from 92.5 to 98.85% with those of the eight most closely related strains (LC155097.1, KM607747.1, KM607826.1, KM607839.1, NC_003475.1, L41576.1, JX170764.1 and GQ214699.1) Table 4. The length of DNA-U3 (LC468139.1) varied among different strains of BBTV and ranged from 1056 to 1062 nts.

The complete sequence of BBTV-DNA component 3 (DNA-S) (LC468140.1) representing the *cp* gene encodes the capsid protein Table 3, shared identities ranging

from 97.18 to 98.98% when compared to the seven most similar overseas strains (NC_003473.1, L41574.1, JQ820455.1, JQ820461.1, KU759885.1, EF687856.1 and EU589459.1) documented in GenBank Table 4. A number of six out of the seven BBTV strains (NC_003473.1, L41574.1, JQ820455.1, JQ820461.1, KU759885.1, and EF687856.1) appeared to have the same length (1075 nts) of BBTV-DNA component 3 (DNA-S) (LC468140.1). While the Indian strain (EU589459.1) was one nucleotide shorter compared to other strains including the Egyptian strain of this study (LC468140.1).

BBTV-DNA component 4 (DNA-M) representing the *mv* gene that encodes the movement protein Table 3, was completely sequenced and documented in GenBank under the accession number LC468141.1. It has a length of 1043 nts and shared identities ranging from 91.71 to 95.40% when compared to the eight most similar overseas strains of BBTV DNA-4 (NC_003474.1, L41575.1, EU095948.1, JQ820462.1, JQ820456.1, EU190971.1, EU516323.1 and JQ820468.1) Table 4. Among the eight strains with the most similar DNA-M component, two strains (NC_003474.1& L41575.1) had the same length of 1043 nts, while the other six strains were 3 to 5 nts longer than the BBTV-DNA-4 under.

Results in Table 3 revealed that BBTV DNA-5 (DNA-C component) of the BBTV Egyptian strain, representing the *ccl* gene encodes the cell-cycle link protein, shared sequence identities of 96.16 to 98.33% with that of the eight most similar BBTV overseas strains (NC_003477.1, L41578.1, EU190969.1, JQ820469.1, JQ820457.1, JQ820463.1, KU759887.1 and EU051379.1). BBTV-DNA-C component of all strains including the one of the current study shared the same length of 1018 nts. DNA-C component of the Egyptian strain of this study (LC468142.1) shared high identities (98.33%) with two Australian strains (NC_003477.1 and L41578.1) Table 4.

The complete nucleotide sequence (1084 nts) of BBTV-DNA-6 (DNA-N component) of the Egyptian strain, representing the *ns* gene encodes the nuclear shuttle protein Table 3, was determined and deposited in GenBank with the accession number of LC468143.1. DNA-N component shared sequence identities ranging from 95.03 to 99.45% with that of the eight most similar overseas strains of BBTV documented in GenBank (AY948438.1, EU190970.1, KM607324.1, NC_003476.1, L41577.1, EU391633.1, KM607438.1 and KM607432.1) Table 4. DNA-N components of the eight most similar BBTV strains varied in their length (1080, 1087, 1089, 1089, 1089, 1096, 1090 and 1090 nts) compared to the Egyptian strain (1084 nts).

Table 3. Deduced amino acids and its ORfs of the six BBTV components.

Accession numbers	Products	aa seq.	Coding region	Protein ID
LC468138.1	Replication associated protein	278	127..963	BBJ34149.1
LC468139.1	U3 protein	88	143..409	BBJ34150.1
LC468140.1	Capsid protein	175	213..740	BBJ34151.1
LC468141.1	Movement protein	102	279..587	BBJ34152.1
LC468142.1	Cell-cycle link protein	156	240..710	BBJ34153.1
LC468143.1	Nuclear shuttle protein	129	281..670	BBJ34154.1

Table 4. Pairwise sequence identities between BBTV-DNA components of the Egyptian strain and most similar overseas strains with E-value (0.0).

Most similar BBTV strains documented in GenBank		Lengths (nts)	Query cover (%)	Identities (%)
Accession numbers	**Countries**			
BBTV-DNA-1 (DNA-R) (LC468138.1)				
AF102780.1	Egypt	1111	97	98.62
AF416465.1	Egypt	1111	100	98.56
HQ259074.1	Egypt	1108	100	98.56
NC_003479.1	Australia	1111	97	97.97
S56276.1	Australia	1111	97	97.97
JQ820459.1	Rwanda	1111	100	97.84
HM120718.1	India	1111	100	97.57
JN250593.1	Sri Lanka	1111	100	97.39
BBTV-DNA-2 (DNA-U3) (LC468139.1)				
LC155097.1	Egypt	1056	98	98.85
KM607747.1	Egypt	1057	98	97.70
KM607826.1	Tonga	1062	98	96.28
KM607839.1	Tonga	1062	98	96.10
NC_003475.1	Australia	1060	100	95.57
L41576.1	Australia	1060	100	95.57
JX170764.1	Pakistan	1062	94	93.03
GQ214699.1	Pakistan	1062	94	92.50
BBTV-DNA-3 (DNA-S) (LC468140.1)				
NC_003473.1	Australia	1075	100	98.98
L41574.1	Australia	1075	100	98.98
JQ820455.1	Malawi	1075	100	98.14
JQ820461.1	Rwanda	1075	100	98.05
KU759885.1	Democratic Republic of Congo	1075	98	97.92
EF687856.1	India	1075	100	97.40
EU589459.1	India	1074	98	97.18
BBTV-DNA-4 (DNA-M) (LC468141.1)				
NC_003474.1	Australia	1043	100	95.40
L41575.1	Australia	1043	100	95.40
EU095948.1	Pakistan	1046	100	92.10
JQ820462.1	Rwanda	1047	100	92.01
JQ820456.1	Malawi	1046	100	91.90
EU190971.1	India	1046	100	91.90
EU516323.1	India	1048	100	91.71
JQ820468.1	Rwanda	1046	100	91.71
BBTV-DNA-5 (DNA-C) (LC468142.1)				
NC_003477.1	Australia	1018	100	98.33
L41578.1	Australia	1018	100	98.33
EU190969.1	India	1018	100	98.03
KU759887.1	Democratic Republic of the Congo	1018	98	97.60
JQ820469.1	Rwanda	1018	100	97.54
JQ820457.1	Malawi	1018	100	97.54
JQ820463.1	Rwanda	1018	100	97.34
EU051379.1	India	1018	100	96.16
BBTV-DNA-6 (DNA-N) (LC468143.1)				
AY948438.1	India	1080	100	99.45
EU190970.1	India	1087	100	96.63
KM607324.1	Egypt	1089	98	96.22
NC_003476.1	Australia	1089	100	95.44
L41577.1	Australia	1089	100	95.44
EU391633.1	India	1096	100	95.03
KM607438.1	Tonga	1090	98	91.54
KM607432.1	Tonga	1090	98	91.35

Harding *et al.*, 1993 are identified a potential stem-loop structure in BBTV-DNA-R that had an 11-nucleotide loop sequence similar to that of Geminivirus. In BBTV-DNA components of the Egyptian strain, a stem-loop region was also recorded Figure 8, nine conserved nucleotides in between, 10 nucleotides representing the stem and 14 fully

conserved nucleotides. By comparing the six components (Fig. 8) a region of 25 nucleotides of the stem-loop region was identified plus 13 nucleotides at 3' of the stem-loop region. Results also showed that 8 differences representing 81% of identities, including one deletion in component 2, between the six components among the 44 stem-loop common regions were founded. It was noted that 14 nucleotides at the 5' prime belonging to the stem-loop region were fully conserved between the six BBTV components. Similarly, 13 nucleotides in the common region were also conserved between the six components. Forty-four nucleotides were conserved between the six BBTV-DNA components. Regarding component 2, a sequence conforming to the consensus polyadenylation signal was identified and was not associated with any major ORF in component 2. Therefore, the potential TATA box and polyadenylation signal in component 2 was not located at the expected position as extrapolated from the other five components, it is possible that this component does not encode a protein but may have a regulatory function (Burns *et al.*, 1995).

```
LC468140.1      AGCGCTGGGGACTATTATTACCCCCAGCGCTCGGGACGGGACAT-----GGG-CTAATGG
LC468141.1      AGCGCTGGGGCTTATTATTACCCCCAGCGCTCGGGACGGGACATCACGTGCT-TTAACAA
LC468139.1      GGCGCTGGGGCTTATTATTACCCCCAGCGC-CGGGACGGGACATG----GGC-TTTTTAA
LC468143.1      AGCACGGGGGACTATTATTACCCCCCGTGCTCGGGACGGGACATGAC--GTC-AGCAAGG
LC468142.1      AGCGCTGGGGCTTATTATTACCCCCAGCGCTCGGGACGGGACATCACGTGCAACTAACAG
LC468138.1      AGCGCTGGGGCTTATTATTACCCCCAGCGCTCGGGACGGGACATTTGCATCTATAAATAG
                 ** * ****  ************* * ** *************
```

Fig 8. Comparison between the CR-SL regions of the six components of BBTV-DNA.

Results in Figure 9 showed the genetic relationship between the six BBTV-DNA components of the Egyptian strain under investigation. It was noted that two components (DNA-R and DNA-U3) were located in one cluster, while the other four components lay in separate clusters.

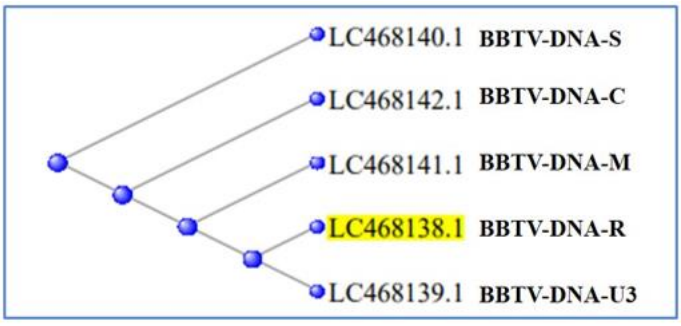

Fig. 9: Phylogenetic tree of the six BBTV-DNA components of an Egyptian strain.

Phylogenetic analyses (Fig. 10) were performed to reveal the genetic relationship between the nucleotide sequences of full-length BBTV-DNA component-1 (DNA-R) (LC 468138.1), encoding the Rep protein, and that of eight BBTV strains. The phylogenetic tree showed that the strain of this study was clustered in a clade with three strains from Egypt, HQ259074.1, AF416465.1 and https://www.ncbi.nlm.nih.gov/nucleotide/AF102780.1?report=genbank&log$=nucltop&blast_rank=3&RID=8ZEY0W70114AF 102780.1 which shared identities of 98.56, 98.56 and 98.62% with LC 468138.1 (Strain under investigation), respectively. While the other strains of Australia, India, SriLanka and the USA were grouped in a separate cluster.

Regarding the genetic relationship between BBTV-DNA component-2 (DNR-U3) (LC 468139.1) and those documented in GenBank, results also showed that different strains were grouped in four clusters (Fig 10). The 1st cluster included strains from Pakistan (JX170764.1 and GQ214699.1), the 2nd cluster included two strains from Australia (NC_003475.1 and L41576.1), the 3rd cluster included two strains from New Zealand (KM607826.1 and KM607839.1), while the 4th one included three strains (two from Egypt (LC155097.1 and LC 468139.1) and one from New Zealand (KM607747.1)).

Results in Figure 10 confirmed the very high percentage of similarity between the BBTV-DNA-S component (LC468140.1) of this research and that of two Australian strains (NC_003473.1 and L41574.1), as they fell together in a separate cluster. Data in Figure 10 proved the close genetic relationship between the BBTV-DNA-M component (LC468141.1) under investigation and BBTV-DNA-M of two Australian strains of (NC_003474.1, L41575.1) as they appeared together in the same cluster. The phylogenetic tree in Figure 10 in which DNA-C component of the Egyptian strain (LC468142.1) confirmed that it was clustered with those of the Australian strains (NC_003477.1 and L41578.1). A very high percentage of similarity (99.45%) between the BBTV-DNA-N component of the Egyptian strain (LC468143.1) and that of the Indian strain (AY948438.1) was recorded and they appeared together in a separate cluster. These results agree with that reported by some investigators (Sadik 1994, Harding *et al.*, 1993, Burns *et al.*, 1995, Amin *et al.*, 2008).

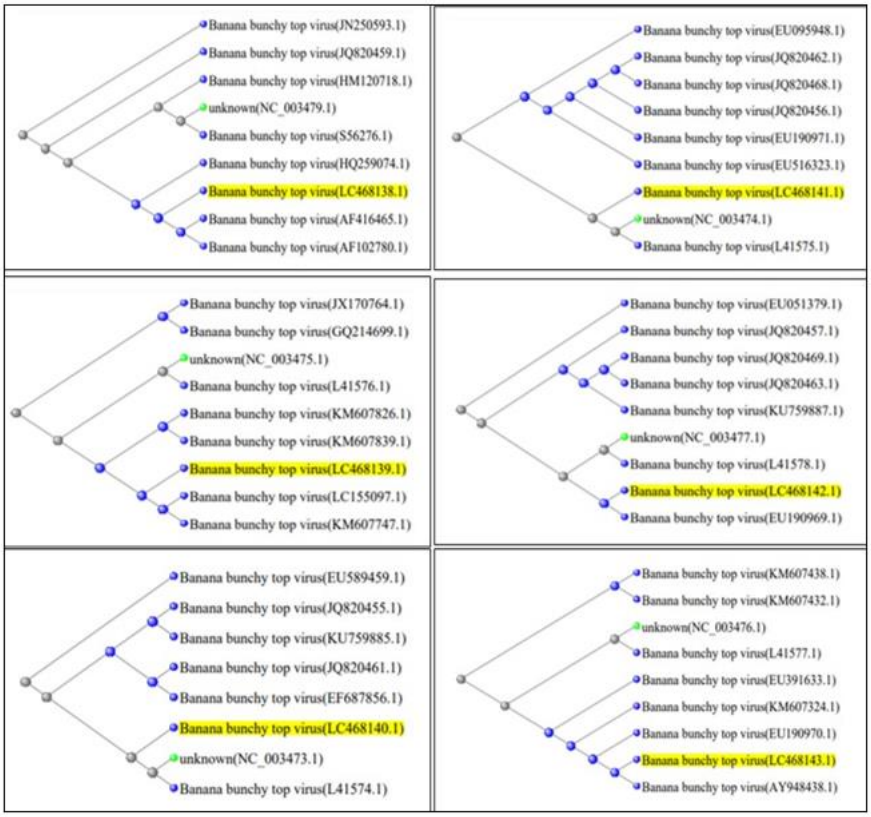

Fig 10. Phylogenetic trees of BBTV-DNA components (DNA-R, LC468138.1, DNA-U3, LC468139.1, DNA-S, LC468140.1, DNA-M, LC468141.1, DNA-C, LC468142.1, DNA-N, LC468143.1) compared to BBTV strains documented in GenBank.

Pairwise alignment between the six BBTV-DNA components and the most similar overseas BBTV-DNA strains of GenBank (Fig. 11) showed that 16 nucleotide differences, which represent 1.44%, were detected between the nucleotide sequences of *rep* genes of BBTV-DNA-1 (DNA-R) of the current study (LC468138.1) and that of the Egyptian BBTV strain (AF416465.1). Differences of 12 nucleotides representing 1.15% were recorded between BBTV-DNA-U3 of the strain under investigation and the most closely related strain (LC155097.1). Only 11 nucleotide differences (1.02%) of BBTV-DNA-S component were recoded between the DNA-S of the Egyptian strain of this study and that of the most closely related strain from Australia (NC_003473.1)**.** Results also showed the presence of 48 nucleotide differences (4.6%) in BBTV-DNA-M component of the Egyptian strain (LC468141.1) when compared to the Australian BBTV-DNA-M component (NC_003474.1). A number of 17 nucleotide differences (1.67%) in BBTV-DNA-C component of the Egyptian strain was recorded when compared with that of the two closely related Australian strains (NC_003477.1 and L41578.1)**.** Only 6 nucleotide differences (0.55%) were recorded in BBTV-DNA component 6 of the Egyptian strain (LC468143.1) when compared with that of the Indian BBTV-DNA component 6 (AY948438.1)**.**

The genetic relationships between the six components of BBTV-DNA of this study (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)) and BBTV isolates from Australia (L41574.1, L41575.1, L41576.1, L41577.1, L41578.1, S56276.1, NC_003475.1, NC_003476.1 and NC_468139.1) and Rwanda (JQ820459.1, JQ820462, JQ820469) documented in GenBank are shown in Figure 12. Six separate clusters have been obtained and each contains one component of BBTV-DNA genome of this study. It was noted that strains of Australia and Rwanda were found together in the same cluster, while the Egyptian BBTV-DNA components occurred in separate subclusters. This could be reflected in the strong genetic relationship between the Australian and Rwandan isolates, and one of them may have evolved from the same origin.

Fig 11. Pairwise sequence alignment between the nucleotide sequences 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) and the most similar BBTV overseas strains.

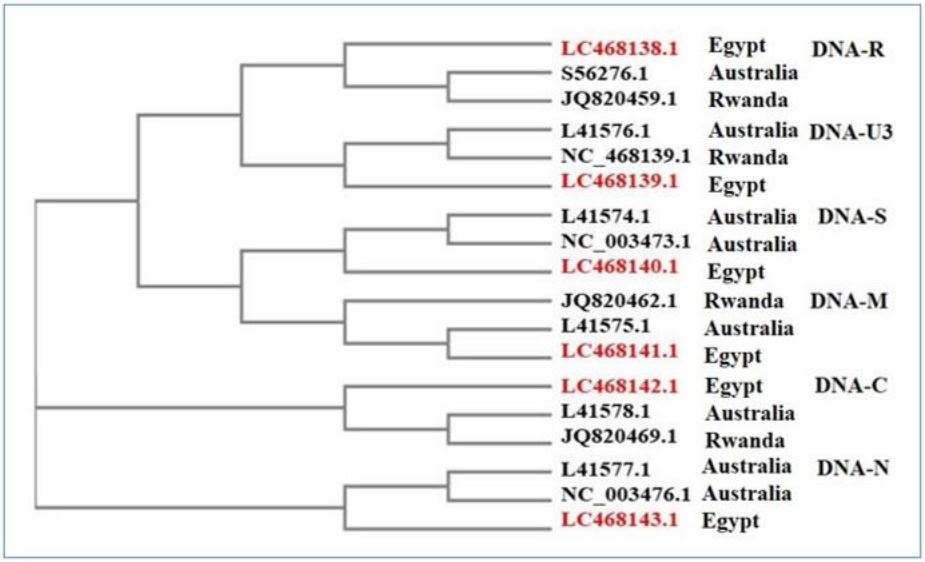

Fig 12. Phylogenetic tree of the six BBTV-DNA components (DNA-R (LC468138.1), DNA-U3 (LC468139.1), DNA-S (LC468140.1), DNA-M (LC468141.1), DNA-C (LC468142.1) and DNA-N (LC468143.1)) compared to six overseas BBTV strains, shows genetic relationship in between.

REFERENCES

Adegbola, RA; Ayodejim, O; Awosusi,OO; Atiri, GI; Kumar, PL (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.

Almeida, RP; Bennett, GM; Anhalt, MD; Tsai, CW; O'Grady, PA (2009) Spread of an introduced vector-borne banana virus in Hawaii. *Molecular Ecology*, 18, 136-46.

Amani, M; Avagyan, G (2014) Effect of polyethylene bunch cover on fungal diseases control of banana (*Musa acuminata* L.) in Iran. *International Journal of Farming and Allied Sciences*, 3, 1054-7.

Amin, I; Qazi, J; Mansoor, S; Ilyas, M; Briddon, RW (2008) Molecular characterisation of Banana bunchy top virus (BBTV) from Pakistan. *Virus Genes*, 36, 191-8.

Anandhi, J; Vijila, C; Viswanath, GS; Lokeswari, TS (2007). Screening banana plants for banana bunchy top virus with primers specific to Indian isolates. *Journal of Plant Diseases and Protection, 114*, 101-107.

Banerjee, A; Roy, S; Behere, GT; Roy, SS; Dutta, SK; Ngachan, SV (2014) Identification and characterization of a distinct banana bunchy top virus isolate of Pacific-Indian Oceans group from North-East India. *Virus Research*,183, 41-49.

Beetham, PR; Hafner, GJ; Harding, RM; Dale, JL (1997) Two mRNAs are transcribed from banana bunchy top virus DNA-1. *Journal of General Virology*, 78, 229-36.

Burns, TM; Harding, RM; Dale, JL (1995) The genome organization of banana bunchy top virus: analysis of six ssDNA components. *Journal of General Virology*, 76, 1471-82.

Hafner, GJ; Harding, RM; Dale, JL (1997) A DNA primer associated with banana bunchy top virus. *Journal of General Virology*, 78, 479-86.

Hamim, I; Green, JC; Borth,WB; Melzer, MJ; Wang,YN; Hu, JS (2017) First report of banana bunchy top virus in *Heliconia* spp. on Hawaii. *Plant Disease*, 101, 2153.

Harding, RM.; Burns, TM; Hafner, G; Dietzgen, RG; Dale, J L (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, RM; Sadik, AS; Bahieldin, A; Dale, JL (2000) A sensitive detection of banana bunchy top nanovirus using molecular genetic approaches. *Arab Journal of Biotechnology*, 3, 103-14.

Hein, J (1990) [39] Unified approach to alignment and phylogenies. *Methods in Enzymology,* 183, 626-645.

Higgins, DG; Sharp, PM (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *Bioinformatics, 5* 151-153.

Jones, DR (2000) Diseases of banana; abaca and Enset. CABI Publishing; Wallingford; UK.

Kakathi, N; Nath, PD (2018) Genetic diversity of banana bunchy top virus (BBTV) prevalent in Assam causing banana bunchy top disease. *International journal of current Microbiology and applied science*, 7(11), 1547-60.

Kakati, N., Dutta, P, Das, P, Nath, PD (2018). Compatibility of Entomopathogenous Fungi with Commonly used Insecticides for Management of Banana Aphid transmitting Banana bunchy Top virus (BBTV) in Assam Banana Production System. *International journal of current Microbiology and applied science, 7(11) 2507-2513.*

King, AM; Lefkowitz, E; Adams, MJ; Carstens, EB; editors (2011) Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses. NY; USA: Elsevier Academic Press.

Kumar, PL; Ayodele, M; Oben, TT; Mahungu, NM; Beed, FD; Coyne, DL; Londa, L; Mutunda, MP; Kiala, D; Maruthi, MN (2009) First report of Banana bunchy top virus in banana and plantain (Musa spp.) in Angola. *Plant Pathology*, 58, 402.

Kumar, PL; Selvarajan, R; Iskra-Caruana, ML; Chabannes, M; Hanna, R (2015) Biology, etiology and control of virus diseases of banana and plantain. *Advances in virus research,* 91, 229-269.

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. Auckland: Invasive Species Specialist Group; 2000 Dec 12.

Mandal, B (2010) Advances in small isometric multicomponent ssDNA viruses infecting plants. *Indian Journal of Virology*, 21, 18-30.

Molina, AB; Kudagamage, C (2002) The international network for the improvement of banana and plantain (INIBAP): PGR activities in South Asia. In: South Asia Net-work on Plant Genetic Resources (SANPGR) Meeting held on December 9-11 at Plant Genetic Resources Center (PGRC); Peradeniya; Sri Lanka; pp 1-7.

Mpoki, S; Mahuku, G; Mbanzibwa, DR; Mkamilo, G; Mark, D; Mosha, H; Pallangyyo, BD; Fihavango, M; Oresanya, A; Ogunsanya, P; Kumar, PL (2021) First report of banana bunchy top virus in banana and plantain (*Musa* spp.) in Tanzania. *Plant Disease*, 27, 1-7.

Nour El-Din, HA; Salama, MI; Barakat, AB; Saleim, AM; Sadik, AS (2005) Nucleotide sequence of BBTV-*cp* gene and using its fusion protein for producing specific polyclonal antibodies. *Arab Journal of Biotechnology,* 8, 353-368.

Qazi, J (2016) Banana bunchy top virus and the bunchy top disease. *Journal of General Plant Pathology*, 82, 2-11.

Rahayuniati, RF; Hartono, S; Somowiyarjo, S; Subandiyah, S; Thomas, JE (2021a) Characterization of banana bunchy top virus on Sumatra (Indonesia) wild banana. *Biodiversitas Journal of Biological Diversity*, 22(3):1243-1249.

Rahayuniati, RF; Subandiyah, S; Hartono, S; Somowiyarjo, S; Kurniawan RE;

Prakoso, AB; Crew, K; Vance, ME; Ray, JD; Thomas, JE (2021b) Recent distribution and diversity analysis on banana bunchy top virus of banana and alternative host in Indonesia. *Tropical Plant Pathology*, 46, 506-17.

Rana, J; Vaisla, KS (2012). Introduction to Bioinformatics. In: Bioinformatics-Tools & Applications; 1st Ed; Vaisla K.S. and Bharti R. (eds.). Printed by Charu Printers; Dehradun; USBD; Dehradun & BTKIT; Dwarahat**.**

Sadik, AS (1994). Studies on viruses affecting banana in Egypt. Ph.D. Thesis, in Agricultural Virology, Dept. Agric. Microbiol., Faculty of Agric., Ain Shams Univ., Cairo, Egypt. 199 pp.

Salama, MI; Khalil, SM; Abdel-Hamid, IA; Hanan, AN; Sadik, AS (2007) Partial nucleotide sequence of segment 1 encoding replicase gene of banana bunchy top nanovirus. *Egyptian Journal of Genetic and Cytology,* 36, 297-304.

Sambrook, J; Fritsch, EF; Maniatis, T (1989) Gel electophoresis of DNA In: Molecular Cloning: A Laboratory Mannual, Part 6, pp. 1-15, 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Selvarajan, R; Mary Sheeba, M; Balasubramanian, V; Rajmohan, R; Dhevi, N L; Sasireka, T (2010). Molecular characterization of geographically different banana bunchy top virus isolates in India. *Indian Journal of Virology*, *21*, 110-116.

Sila, S; Abadi, AL; Mudjiono, G; Astono, TH (2020) Banana Bunchy Top Virus (BBTV) on wild banana species in Kutai Kartanegara Regency. *EurAsian Journal of Biosciences*, 14, 2.

Singh,V; Adil, S; Quraishi, A (2022) Elimination of BBTV via a systemic in vitro electrotherapy approach. *Journal of Virological Methods*, 300, 114367.

Su, HJ; Tsao LY; Wu, ML; Hung, TH (2003) Biological and molecular categorization of strains of Banana bunchy top virus. *Journal of Phytopathology* ,151, 290-6.

Vetten, HJ; Chu, PW; Dale, JL; Harding, RM; Hu, J; Katul, L; Kojima, M; Randles, JW; Sano,Y; Thomas, JE (2005). Nanoviridae. Fauquet CM; Mayo MA; Maniloff J; Desselberger U; Ball LA; editors. Virus taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses. Academic Press.

Vetten, HJ; Dale, JL; Grigoras, I; Gronenborn, B; Harding, R; Randles, JW (2012) Family Nanoviridae in virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. pp. 395-404.

Vishnoi, R; Raj, SK; Prasad, V (2009) Molecular characterization of an Indian isolate of Banana bunchy top virus based on six genomic DNA components. *Virus Genes*, 38, 334-44.

Wickramaarachchi, WA; 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 Disease*, 27, 154-60.

Xie, WS; Hu, JS (1995) Molecular cloning; sequence analysis; and detection of banana bunchy top virus in Hawaii. *Phytopathology*, 85, 339-47.

Yu, NT; Feng, TC; Zhang, YL; Wang, JH; Liu, ZX.(2011) Bioinformatic analysis of BBTV satellite DNA in Hainan. *Virologica Sinica*, 26, 279-84.