Int.J.Curr.Microbiol.App.Sci (2021) *10*(01): 1923-1936

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 10 Number 01 (2021)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2021.1001.224

Phylogeographic Analysis and Codon Usage Bias of Coat Protein Gene of Banana Bunchy Top Virus Isolates in Kerala

Darsana Dilip[1*], Vimi Louis[2], P.M. Namitha[2]

[1]Department of Plant Pathology, College of Horticulture, Kerala Agricultural University, Thrissur, Kerala, 680656, India
[2]Division of Plant Pathology, Banana Research Station, Kannara, Kerala Agricultural University, Thrissur, Kerala, 680652, India, India

**Corresponding author*

Keywords

Banana bunchy top virus, Genetic diversity, Phylogeny, Codon usage bias

Article Info

Accepted:
12 December 2020
Available Online:
10 January 2021

A B S T R A C T

Banana bunchy top disease (BBTD) caused by *Banana bunchy top virus* (BBTV) is one of the most devastating diseases in banana. The virus is reported to be evolving and causing serious economic loss in various banana growing parts of the world. The complete coding region and partial untranslated region (UTR) of coat protein (CP) gene was amplified from total DNA isolated from infected samples. The partial sequence of DNA-S of seventeen Kerala isolates were analysed for molecular diversity, codon usage bias and phylogeny were deduced. A large negative value of Fu and Li' D and F tests indicated a demographic expansion of the population. Being a DNA virus, BBTV showed very less diversity among the isolates in the coding and non-coding regions. Although, nucleotide diversity and mutation rate in the non-coding region was surprisingly higher than the coding region. In the non-coding region, the stem loop region was highly conserved responsible for binding to the master rep protein and facilitating rolling circle replication of the genome. Nevertheless, the phylogeographic analysis confirmed an association with the geography as the northern, southern and clusters were grouped separately with high bootstrap support. There was a very conspicuous codon usage bias (CUB) in the CP gene and evidently nineteen synonymous codons were favoured over the others. The role of mutation and selection pressure on CUB of the gene was established using neutrality plot and Nc plot.

Introduction

Viral diseases causing serious decline in banana yield have been reported since 20th century. However, India has witness serious deterioration in area, production and productivity of banana in 1970s due to infestation of *Banana bunchy top virus*(BBTV) alone especiallyin large plantations of high yielding cultivar Virupakshi (AAB), also known as Hill Banana in lower Pulneyhills, Tamil Nadu (Selvarajan *et al.*, 2010). Although, banana aphid, *Pentalonia nigronervosa* has been established as the vector of the virus, BBTV is mainly spread through vegetative

propagules of banana which is why it spreads to larger distances and cause huge economic lossannually.

BBTV is a circular single stranded DNA virus belonging to the genus *Babuvirus* in the family *Nanoviridae* (Vetten *et al.,* 2004). The genome is multipartite with six components of approximately 1 kb size designated as DNA-R, -S, -M, -C, -N and -U, encapsidated separately in isometric virion of 18–20 nmin diameter(Burns *et al.,* 1995; Harding *et al.,* 1993). DNA-R encodes for master replicator protein, whereas DNA-S encodes the viral coat protein (CP) for encapsulation of all the six genomic components. DNA-M, -N and -C encodes for movement protein, nuclear shuttle protein and cell cycle link protein respectively. The function of DNA-U has not yet deciphered (Wanitchakorn *et al.,* 1997).

Many hypotheses have been proposed to substantiate the evolutionary advantage of genome segmentation (Chao, 1988; Szathmary, 1992). Whilst some researchers have explained it on the basis of efficiency in replication due to the shorter length of the genome, others have attributed virion stability to segmented genomes (Nee, 1987; Ojosnegros *et al.,* 2011). Whatsoever, the potential fitness gain seems to be an increase in genetic diversity of the virus population forming quasi species. Nelson *et al.,* in 2008 explained consequences of this as elimination of harmful mutations and maintaining favourable ones through genomic reassortment (component exchange) among different isolates or even by inter or intra-molecular recombination (Hyder *et al.,* 2011).

The BBTV isolates are categorized into two groups, namely the Pacific Indian Oceans group (PIO) and the South East Asian group (SEA) based on nucleotide sequence identity of DNA-R (Karan *et al.,* 1994). All the Indian isolates have been categorised into the former. Later, similar clustering was reported based on other segments of the genome (Banerjee *et al.,* 2014; Selvarajan *et al.,* 2010). Being a DNA virus, the variability of BBTV supposedly is lower compared to RNA viruses infecting banana like *Banana bract mosaic virus* (BBrMV) and *Cucumber mosaic virus* (CMV). Apparently, Wanitchakorn *et al.,* (2000) reported significant variability of 1.77% in the amino acid sequences of BBTVCP within the SEA group. It was correlated withthe evolution during the extended period of time the virus has been present in this region. Similarly, BBTV has been widespread in India since 1943 and so genetic variability among isolates of the PIO groups also exists which was reported earlier (Selvarajan *et al.,* 2010). However, no such studies regarding the molecular diversity and phylogeographic analysis of Kerala isolates have not done so far.

Most of the diagnostics are based on the coat protein gene of BBTV be it nucleic acid based or serology based. In this context, it is inevitable to understand the molecular diversity of the BBTV isolates prevalent in Kerala, and decode the phylogeography based on their CP gene. In the present study, the coat protein gene of BBTV Kerala isolates were characterised and their diversity was investigated to try and describe the reasons for it. An attempt to understand the fitness gain as well as the evolutionary advantage in this multipartite DNA virus which theoretically should not be highly variable.

Materials and Methods

Media and bacteriological reagents were procured from Hi-Media, Mumbai, India. Other chemicals, antibody and reagents, were purchased from Sigma-Aldrich and ThermoFisher Scientific. The oligonucleotides were synthesised from AgriGenome, Kochi, Kerala.

Sample collection

A purposive sampling survey was conducted in ten districts representing Northern, Central and Southern zones of Kerala. Northern districts in Kerala *viz.*, Kasaragod, Kannur, Kozhikode and Wayanad were included in the northern zone, Thrissur, Palakkad and Ernakulam were included in the central zone and Kottayam, Kollamand Trivandrum were included in the southern zone. Young leaves of banana showing banana bunchy top symptoms were collected from different varieties grown in various location of Kerala. Infected samples were collected irrespective of the cultivar.

Direct Antigen Coating-Enzyme Linked Immunosorbant Assay (DAC-ELISA)

Samples were subjected to DAC-ELISA to confirm infection (Clark and Adams, 1977). Antigen was isolated in carbonate buffer(pH 9.2) and coated on to 96 welled micro titre plate (Tarsons Pvt Ltd) with 2 % polyvinyl pyrrolidone (PVP- K30).The BBTV specific primary antibody (NRCB, Trichy) was diluted to 1:1000 v/v in PBS-TPO buffer (PBS-T with 2% poly vinyl pyrrolidone and 0.2% bovine serum albumin) and added to the antigen coated wells post washing with PBS-T and incubated at 37 ºC overnight. 1:10000 (v/v) anti-rabbit IgG conjugated with alkaline phosphatase enzyme diluted in PBS-TPO was used as secondary antibody. The wells were incubated at 37 ºC for 2 h. Three washes with PBS-T were carried out to remove excess antibody and subsequently, 200 µl of substrate (1000 ppm of para-nitro phenyl phosphate disodium salt in diethanolamine, pH 9.8) was added and incubated at RT in dark for 30 min. The absorbance was recorded at 405 nm in ELISA reader. Two healthy samples as negative control were added in the assay. The test samples that showed higher than double the A405 value of negative controls were taken as positive.

DNA extraction, PCR and sequencing

Total DNA was isolated from seventeen representative samples that gave positive results in ELISA using modified CTAB method. In the presence of liquid nitrogen, 1g leaf sample was pulverized to fine powder and transferred to an autoclaved Oakridge tube. To this, 10 ml pre heated 2X extraction buffer (2 % CTAB, 100mM Tris, pH 8.0, 1.4 M NaCl, 20mM EDTA, 1 % PVP and 0.1 % β-mercaptoethanol) was added and incubated at 65 ºC for 10 min. The tube was mixed by inversion intermittently. 10 ml of 24:1 v/v chloroform isoamyl alcohol was added to this and mixed well. The aqueous phase was transferred into another Oakridge tube after centrifugation at 10000 rpm for 15 min at 4 ºC. Equal volume of ice-cold isopropanol was added and incubated at -20 ºC for 1 hafter mixing thoroughly. The precipitated DNA was collected by centrifugation at 10000 rpm for 15 min at 4 ºC. The pellet was washed with 70 % ice cold ethanol and was air dried. Pellet was dissolved in nuclease free water and quantified in Nanodrop. Quality of DNA was also checked in 0.8 % agarose gel.

Total DNA isolated from infected samples was used as template for Polymerase Chain Reaction (PCR) to amplify coat protein (CP) gene using BBTV CP specific forward and reverse primers (Wanitchakorn *et al.,* 2000). PCR reaction mixture contained 1X Taq buffer, 10 mM each dNTPs, 10 pmol each of BBTV CP specific forward and reverse primers, 2 units of Taq polymerase (Genei, Bangalore). The template was diluted appropriately to 50-100 ng. PCR was performed in a Thermal cycler (Eppendorf, Germany) under the following conditions: 94 °C for 3 min, followed by 35 cycles of 94 °C for 1 min, 60.8 °C for 1 min, 72 ºC for 1 min 30 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 in both directions. The sequences were processed and contigs were prepared. The processed sequences were submitted in NCBI GenBank and accession numbers MT174314 - MT174330 were obtained.

Diversity analysis

The sequences were aligned by MUSCLE and sequence homology was represented as colour coded matrix using SDT v 1.2 software. DNA polymorphism was detected in DnaSP v.6.12.01. DNA divergence and gene flow between the subpopulation (North, South and Central zones of Kerala) were analysed. Kerala isolate with GenBank accession number GU125413.1 was used as reference isolate. Analysis of rate of non-synonymous and synonymous substitutions per site between pair isolates were conducted using the Nei-Gojobori method (Nei and Gojobori, 1986). The values of $P < 0.05$ were considered significant at the 5 % level.

Phylogeographic analysis

The aligned sequences were subjected to phylogenetic analysis in MEGA X software (Kumar *et al.,* 2018) by Neighbour joining method. Evolutionary distances were computed using the Kimura 2-parameter method (Kimura, 1980) and represented in the units of the number of base substitutions per site. Bootstrap test with 1000 replications to signifying the confidence of the branches was conducted (Felsenstein, 1985). Neighbour-Join and BioNJ algorithms were applied to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach to obtain initial tree(s) for the exploratory analysis (Saitou and Nei, 1987).

Codon usage bias

Relative Synonymous Codon Usage (RSCU) and effective number of codons (Nc) were calculated using a web server http://genomes.urv.es/CAIcal/ as indicators of Codon usage bias (CUB) (Puigbò *et al.*, 2008). The effective number of codons (Nc) was plotted in the Y axis and GC value at the wobble position (GC3) was plotted in the X axis to obtain Nc plot. The average of GC at 1st and 2nd position of the codon (GC12) was plotted against GC3 in the X axis to obtain neutrality plot. Nc plot and neutrality plot were constructed to study the degree of relatedness of mutational pressure and natural selection on codon usage bias on CP gene of BBTV.

Results and Discussion

Amplification of coat protein gene

Seventeen representative samples after ELISA, were chosen and total DNA was isolated and CP gene of BBTV was amplified. An amplicon of ~1060 bp was obtained. The PCR products were purified and sequenced using gene specific forward and reverse primers. Contigs was blasted again NCBI database to confirm the gene was BBTVCP. The dataset consisted of seven isolates from northern zone, eight from central and two from southern zone of Kerala.

The sequences were aligned by MUSCLE and further *in silico* predictions and analyses were carried out. Various authors have designed primers for characterisation of coat protein gene of BBTV isolates collected from all the major banana growing countries. In order to amplify the complete circular segment, overlapping primers were designed (Das and Banerjee, 2018; Vishnoi *et al.,* 2009). This would provide a better understanding of the whole gene.

Diversity analysis

The pairwise sequence homology of BBTV Kerala isolates ranged from 98.50 % to 100 % (result not shown) indicating the nucleotide identity of the isolates were very high. Majority of sequence pairs were had a sequence identity of 99 %. Most variable among all the isolates were MT174317 isolated from banana cv. Nendran collected from Wayanad district, northern zone of Kerala. The pairwise amino acid sequence diversity ranged between 1.2-1.8 % only indicated the conservation of amino acid sequences in coding region of the CP gene (Fig.1). This dismissed the hypothesis of the isolates being mutating at a faster rate. However, Banerjee *et al.,* (2014) had reported a new PIO isolate which was allegedly formed from recombination. Two significant inter-component recombination signals were identified in DNA-S segment of BBTV isolate from Meghalaya. Curiously, five intragenomic recombination signals were observed in the DNA-R segment of the isolate all in the non-coding region. On comparing isolates from different parts of the world, Hu *et al.,* (2007) reported a diversity below 10 % among the isolates tested in the study. The study highlighted that the segments DNA-S, -R and -N of various other nanoviruses also followed similar trend and clustered together in principle co-ordinate analysis.

A total number of 41 polymorphic sites were identified in the CP gene of Kerala isolates. Nine mutations were found in the coding region. No insertion-deletion (In-Del) mutations were identified in the CP gene. Three conserved regions were identified in the gene with high statistical significance. Out of the three regions two were in the coding region and one in the non-coding region. The stem-loop region in the common region (CR-SL) (underlined) in the common region, was highly conserved in the non-coding region (AGCGCTGGGGCTTATTATTACCCCCAGCGCT). This is the region which loops out and the master rep protein binds to for beginning of the rolling circle replication. Intriguingly, the CR-SL of DNA-U of Pakistani isolates were identified to be the most diverse region in the genome segment (Hyder *et al.,* 2011). Many recombination events were identified in the region. The report also highlighted that similar sequence type exists in the CR-SL of DNA-R, -S and -M and -C which validated findings of the present study.

The nucleotide diversity (π) and mutation rate (θ) of the CP gene also suggested three conserved regions. However, mutation rate in the coding region was much lower than that in the non-coding region (Fig 2) validating results of previous studies. Fu and Li' D and F tests gave significant negative values (-2.147and -2.104 respectively) indicating a demographic expansion of the population after a bottleneck.

The ratio of rate of non-synonymous substitutions over the rate of synonymous substitutions, a reference isolate was selected from NCBI database. Substitution rates of both synonymous and non-synonymous were absent in most of the sequence pairs. All BBTV isolates investigated in a study by Banerjee *et al.,* (2014) depicted more amino acid sequence identity than nucleotide identity in CP gene segment. Interestingly, Ka/Ks values were extremely low or zero indicating negative or purifying selection. Rate of synonymous substitutions were comparatively higher among the subpopulations collected from Northern and Southern Kerala.

Noteworthy observation was that the genetic diversity among the subpopulations were high. Statistical population diversity indices like G_{ST} and N_{ST} values were very low and significant. The diversity among northern and

central isolated were least. However, frequent gene flow (F_{ST}) was observed between the subpopulations (Table 1). Previously, two Tamil Nadu isolates were analysed for diversity and compared to other PIO and SEA isolates deposited in NCBI GenBank (Selvarajan *et al.,* 2010). It was noticed that the CP gene encoded by DNA-S was comparatively variable that the DNA-R segment encoding for master rep protein (replicase).

Phylogeographic analysis

Phylogeny was interpreted using Neighbour-Joining method and a tree was constructed based on coat protein gene sequences of Kerala isolates (Fig. 3). From the phylogenetic tree, a very prominent demarcation of southern and northern isolates was evident with significant bootstrap support. However, isolates collected from central zone paired with isolates collected from both northern and southern zones.The isolate MT174321 from central Kerala, grouped with the northern zone isolates *viz.*, MT174317, MT174328, MT174315, MT174319, MT174314 and MT174318. Alternatively, isolates collected from central zone (MT174322, MT174325, MT174323, MT174326, MT174327) clustered together. However, MT174329 collected from southern Kerala clustered with the isolates from central zone of Kerala. Three isolates *viz.*, MT174330 (Kollam, South Kerala), MT174316 and MT174320 (Wayanad, North Kerala) formed outgroups. From the branch lengths, it was evident that accessions MT174314 and MT174317 from Kannur and Wayanad respectively (both from northern zone) are most divergent among the group. The observations strongly suggested the correlation between the phylogeny and geography. The phylogenetic analysis corroborated with the molecular diversity analysis between the subpopulations wherein the least diverse subpopulations were Kerala isolates from northern and central zones. Remarkably, founder's effect might have a big role in geographic differentiation between northern and southern isolates.

Hitherto, monophyletic clustering of BBTV isolates from Tripura based on the CP gene was reported similar to the present study(Das and Banerjee, 2018). Similarly, negligible diversity of CP gene among the Japanese and Chinese BBTV isolates belonging to SEA groups were also discussed in detail in few reports even though rate of mutation was discovered to be higher in them (He *et al.,* 2000; Furuya *et al.,* 2005).

Codon usage bias of CP gene

The effective number of codons (Nc), measure of codon usage bias (CUB), of the CP gene of BBTV Kerala isolates is 55.541 ± 0.77 indicative of codon bias (Table 2). Codons with RSCU values higher than one were present in the dataset also indicated strong CUB. Out of these 29 codons, 19 were high frequency codons with RSCU value greater than 1.4 (UUU, UUA, UUG, CUU, AGU, CCU, ACA, GCA, UAU, CAU, CAG, AAC, GAU, GAA, UGU, CGG, AGA, AGG, GGG) and the rest were designated as desirable codons (Table 3). Twelve among the high frequency codons had A or U in the wobble position. This certainly points out the role of nucleotide composition of the gene especially that of the wobble position in codon bias. The CP gene is AT rich with 57.00 ± 0.15 % adenine and thymine residues. Percentage of A and U in the wobble position (%AU3) of the codon was 54.63 ± 0.45 % whereas, %GC3 was 45.37 ± 0.45 %.

The role of mutational pressure and selection pressure in CUB was assessed by Nc plot and GC plot (Fig 4).

Table.1 Gene flow analysis of CO gene of isolates collected from northern, southern and central zones of Kerala

Population 1	Population 2	Hs	Ks	Gst	Nst	Fst	Dxy	Da
North Kerala	Central Kerala	0.981	7.657	0.009	0.128	0.127	0.008	0.001
North Kerala	South Kerala	1.000	10.222	0.047	0.118	0.118	0.008	-0.001
Central Kerala	South Kerala	0.964	6.286	0.059	0.182	0.181	0.009	0.002

Hs denotes the genetic diversity within the population; **Ks** denotes the rate of synonymous substitutions per site; genetic differentiation is denoted by $\mathbf{N_{ST}}$ and $\mathbf{G_{ST}}$ values and gene flow is denoted by fixation index, $\mathbf{F_{ST}}$. **Dxy** is the average number of nucleotide substitutions per site between populations and **Da** denotes the number of net nucleotide substitutions per site between populations. All the values were estimated using the DNaSP v. 6 program with 1,000 permutation tests

Table.2 Nucleotide parameters and effective number of codons of BBTV CP of Kerala isolates

Isolates	%A	%C	%T	%G	% GC3	%AU3	%GC12	Nc
MT174314	30.21	17.74	26.71	25.34	45.61	54.39	41.8	56.50
MT174315	30.21	17.35	26.90	25.54	45.03	54.97	41.8	58.40
MT174316	30.21	17.74	26.71	25.34	45.61	54.39	41.8	52.70
MT174317	30.41	17.15	27.10	25.34	43.86	56.14	41.8	51.10
MT174318	30.21	17.54	26.71	25.54	45.61	54.39	41.8	53.50
MT174319	30.41	17.54	26.71	25.34	45.03	54.97	41.8	52.50
MT174320	30.21	17.74	26.71	25.34	45.61	54.39	41.8	55.60
MT174321	30.21	17.54	26.71	25.54	45.61	54.39	41.8	54.60
MT174322	30.41	17.54	26.71	25.34	45.03	54.97	41.8	54.60
MT174323	30.21	17.74	26.71	25.34	45.61	54.39	41.8	52.30
MT174324	30.21	17.54	26.71	25.54	45.61	54.39	41.8	54.40
MT174325	30.21	17.74	26.71	25.34	45.61	54.39	41.8	55.60
MT174326	30.21	17.54	26.71	25.54	45.61	54.39	41.8	53.60
MT174327	30.21	17.54	26.71	25.54	45.61	54.39	41.8	54.10
MT174328	30.41	17.54	26.71	25.34	45.03	54.97	41.8	54.40
MT174329	30.21	17.54	26.71	25.54	45.61	54.39	41.8	56.00
MT174330	30.21	17.74	26.71	25.34	45.61	54.39	41.8	55.60
Average	**30.26**	**17.58**	**26.74**	**25.42**	**45.37**	**54.63**	**41.8**	**55.54**
SD	**0.08**	**0.15**	**0.10**	**0.10**	**0.45**	**0.45**	**7.11E-15**	**0.77**

Table.3 Relative Synonymous Codon Usage (RSCU) of BBrMV CP. High frequency codons are indicated by (*) next to RSCU value. High frequency (Hfr) codons are indicated by (#) next to the codon

Amino acid	Codon	RSCU	Amino acid	Codon	RSCU
Phenylalanine	UUU#	1.600*	Asparagine	AAU	0.400
	UUC	0.400		AAC#	1.600*
Leucine	UUA#	1.500*	Lysine	AAA	0.640
	UUG#	1.500*		AAG	1.360*
	CUU#	1.875*	Aspartic acid	GAU#	1.500*
	CUC	0.022		GAC	0.500
	CUA	0.022	Glutamic acid	GAA#	1.556*
	CUG	1.081*		GAG	0.444
Isoleucine	AUU	0.900	Cysteine	UGU#	1.624*
	AUC	1.200*		UGC	0.376
	AUA	0.900	Arginine	CGU	0.032
Valine	GUU	1.349*		CGC	0.513
	GUC	0.517		CGA	0.000
	GUA	1.067*		CGG#	1.636*
	GUG	1.067*		AGA#	1.636*
Serine	UCU	0.857		AGG#	2.182*
	UCC	0.429	Glycine	GGU	0.444
	UCA	0.857		GGC	0.444
	UCG	0.857		GGA	1.333*
	AGU#	1.714*		GGG#	1.778*
	AGC	1.286*	Histidine	CAU#	1.600*
Proline	CCU#	1.451*		CAC	0.400
	CCC	1.098*	Glutamine	CAA	0.000
	CCA	0.513		CAG#	2.000*
	CCG	0.938	Alanine	GCU	1.042*
Tryptophan	ACU	0.333		GCC	0.353
	ACC	1.000		GCA#	2.084*
	ACA#	2.333*		GCG	0.521
	ACG	0.333	Tyrosine	UAU#	1.556*
				UAC	0.444

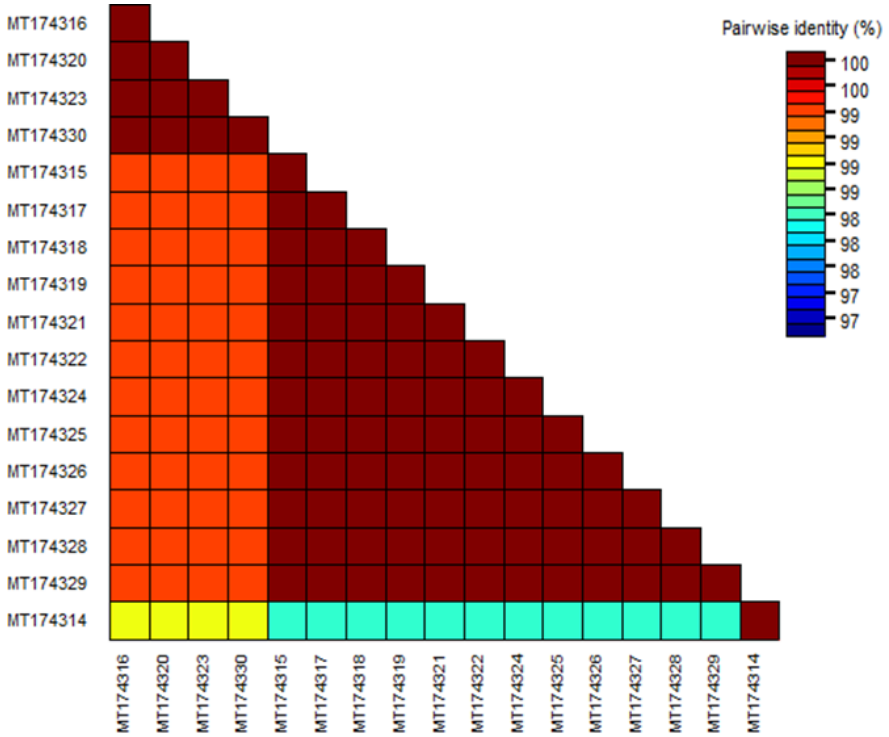

Fig.1 Amino acid sequence identity heat map denoting sequence identity of BBTV isolates

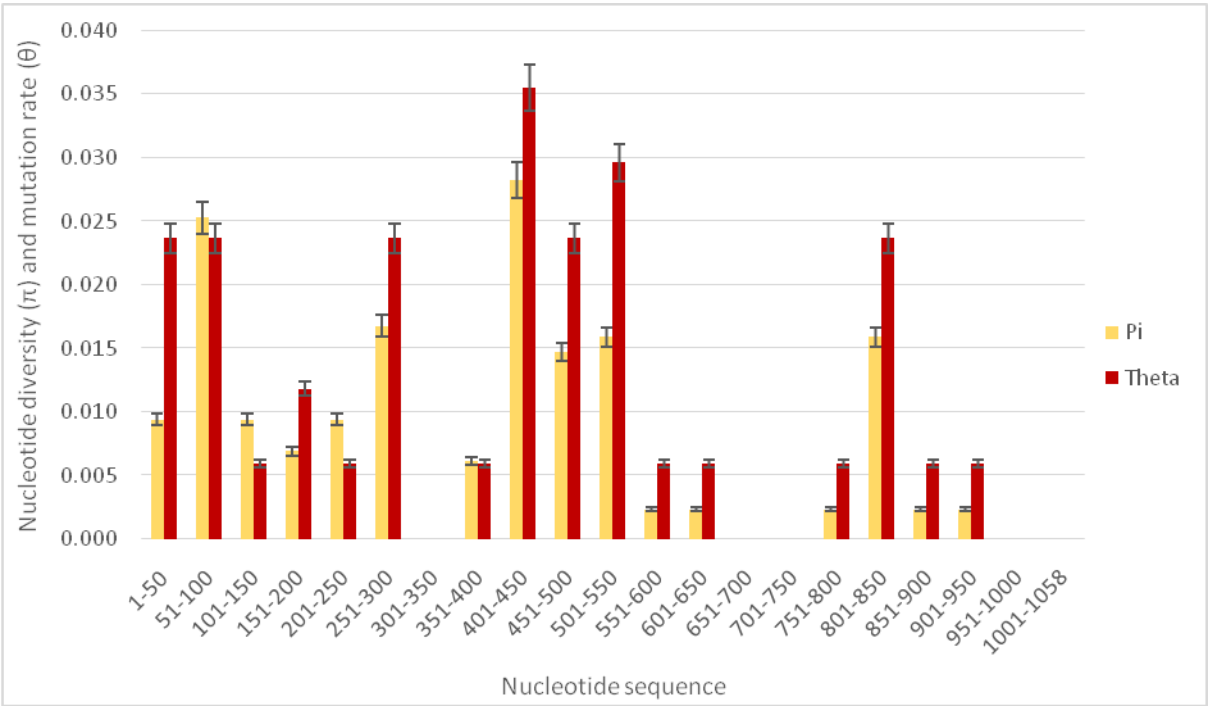

Fig.2 Nucleotide diversity and mutation rate of BBTV CP gene

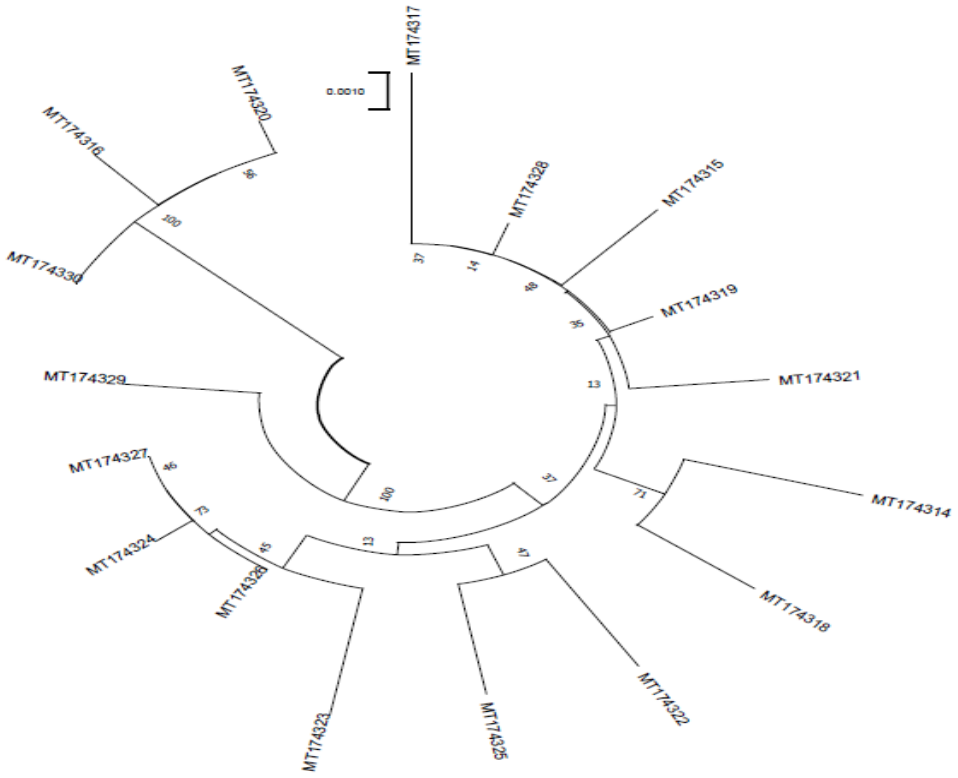

Fig.3 The evolutionary history of Kerala isolates was inferred using the Neighbor-Joining method and evolutionary distances computed using the Tajima-Nei method. The optimal tree with the sum of branch length = 0.04726888 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches

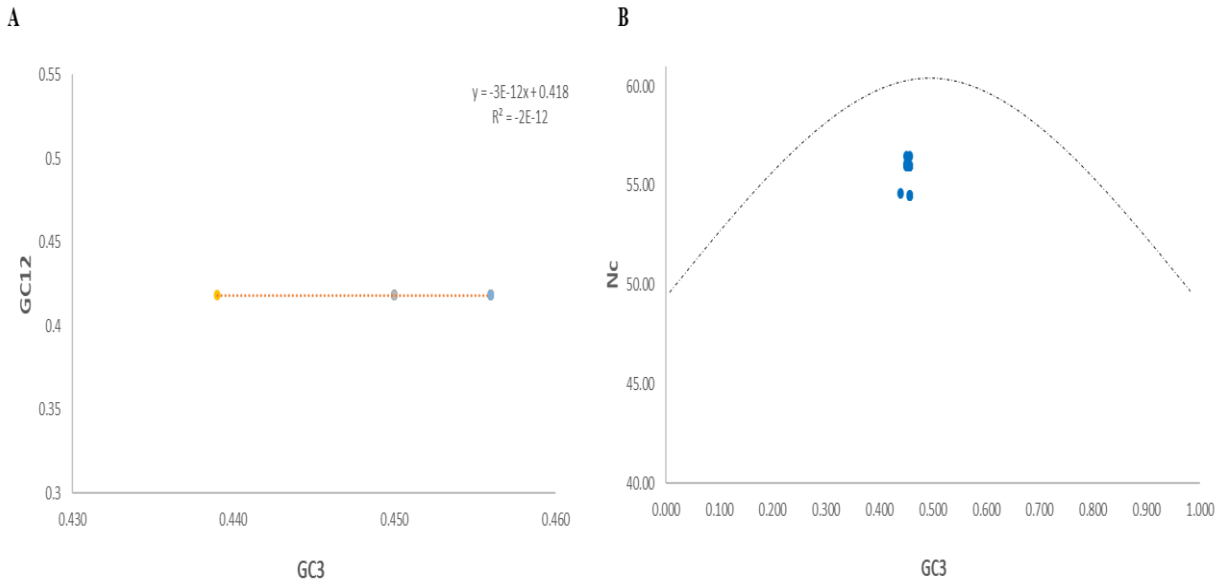

Fig.4A. Neutrality plot and **B.** Nc plot of BBTV CP gene of Kerala isolates**.** Neutrality plot with G+C in the first and second position of the codon (GC12) on the Y axis and G+C in the wobble position of the codon (GC3) on the X axis. Nc plot with GC3 on the X axis and effective number of codon (Nc) on the Y axis

The neutrality plot was linear and all the points were on the trendline depicting the role of both mutational and selection pressure on CUB (Fig 4A). However, The Nc plot depicted all the points below the standard curve indicative of the role of selection pressure on CUB (Fig 4B). If CUB was only influenced by nucleotide composition, the points would have been on the standard curve. Hitherto, CUB was established in viruses like and the role of mutational and selection pressures were identified in them. Codon usage pattern of DNA-S was inferred by Hu *et al.,* (2007).

In the multivariant analysis, DNA-S and DNA-R clustered together in distinct groups and exhibited variation other gene segments of BBTV and another genus in *Nanoviridae*. The study concluded that DNA-S of BBTV shows concerted evolution wherein paralogous genes within one species are more closely related to one another than to members of the same gene family in closely related species in *Nanoviridae* family.

In conclusion, the diversity of CP gene of BBTV was less between the subpopulations but high in each subpopulation. The Kerala isolates generated in the present study was homologous to the isolate submitted in NCBI database before a decade. Phylogenetic analysis proved the clustering of isolates based on the geography. The CUB of the CP gene was evident from the analysis and the role of mutation and selection on CUB was also established in this study. Furthermore, complete sequence of other genomic segments of BBTV Kerala isolates have to be assessed for molecular diversity, codon bias with phylogeographic comparisons and moreover, should be analysed for recombination and reassortment as reported in other PIO isolates, in order to reach a conclusion comprehensively.

Acknowledgment

DD acknowledges DST for INSPIRE fellowship. DD acknowledges Banana Research Station and College of Horticulture, Kerala Agricultural University for providing infrastructure and technical facilities.

References

Banerjee, A., Roy, S., Behere, G.T., Roy, S.S., Dutta, S.K. and Ngachan, S. V., 2014. Identification and characterization of a distinct banana bunchy top virus isolate of Pacific-Indian Oceans group from North-East India. Virus Res. 183, 41–49. https://doi.org/10.1016/j.virusres.2014.01.017

Burns, T.M., Harding, R.M. and Dale, J.L., 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

Chao, L., 1988. Evolution of sex in RNA viruses. J. Theor. Biol. 133, 99–112.

Clark, M. F. andAdams, A. N., 1977. Characteristics of the microplate method of Enzyme Linked Immunosorbent Assay for the detection of plant viruses. J. Gen. Virol. 34, 475-483.https://doi.org/10.1099/0022-1317-34-3-475

Das, T. and Banerjee, A., 2018. Distribution, molecular characterization and diversity of banana bunchy top virus in Tripura, India. Virus Disease 29, 157–166. https://doi.org/ 10.1007/s13337-018-0451-7

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39, 783-791.

Furuya, N., Kawano, S. and Natsuaki, K.T., 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

Harding, R.M., Burns, T.M., Hafner, G., Dietzgen, R.G. and Dale, J.L., 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

He, Z. F., Li, H. P., Xiao, H. G. and Fan, H. Z., 2000, Cloning and sequencing of DNA component 1 of two BBTV strains. Acta Phytopathol Sin. 30, 364–369.

Hu, J. M., Fu, H. C., Lin, C. H., Su, H. J. and Yeh, H. H., 2007. Reassortment and Concerted Evolution in Banana Bunchy Top Virus Genomes. J. Virol. 81, 1746–1761. https://doi.org/10.1128/jvi.01390-06

Hyder, M. Z., Shah, S. H., Hameed, S. and Naqvi, S. M. S., 2011. Evidence of recombination in the Banana bunchy top virus genome. Infect. Genet. Evol. 11, 1293–1300. https://doi.org/10.1016/j.meegid.2011.04.015

Karan, M., Harding, R. M. and Dale, J. L., 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

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16, 111-120.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018, MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol. 35, 1547-1549.

Nee, S., 1987. The evolution of multicompartmental genomes in viruses. J. Mol. Evol. 25, 277–281.

Nei, M. and Gojobori, T., 1986, Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 3, 418-426.

Nelson, M. I., Viboud, C., Simonsen, L., Bennett, R. T., Griesemer, S. B., George, K., Taylor, J., Spiro, D. J., Sengamalay, N. A., Ghedin, E., Taubenberger, J. K., Holmes, E. C., 2008. Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. PLoS Pathog. 4 (2), e1000012.

Ojosnegros, S., Garcı´a-Arriaza, J., Escarmı´s, C., Manrubia, S.C., Perales, C., Armando, A., Mateu, M.G. and Domingo, E., 2011. Viral genome segmentation can result from a trade-off between genetic content and particle stability. PLoS Genet. 7 (3), e1001344, doi: 10.1371/journal.pgen.1001344.

Puigbò, P., Bravo, I. G., Garcia-Vallvé, S., 2008. E-CAI: A novel server to estimate an expected value of Codon Adaptation Index (eCAI). BMC Bioinformatics, 9. https://doi.org/10.1186/1471-2105-9-65

Saitou, N. and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol, 4, 406-425.

Selvarajan, R., Mary Sheeba, M., Balasubramanian, V., Rajmohan, R., Dhevi, N.L. and Sasireka, 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

Szathmary, E., 1992. Viral sex, levels of selection, and the origin of life. J. Theor. Biol. 159, 99–109.

Vetten, H. J., Chu, P. W., Dale, J. L., Harding, R. M., Hu, J., Katul, L.,

Kojima, M., Randles, J. W., Sano, Y. and Thomas, J. E., 2004, In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA, editors. Virus taxonomy- eighth report of the International Committee on Taxonomy of Viruses. London, Academic Press, 343–352.

Vishnoi, R., Raj, S.K. and Prasad, V., 2009. Molecular characterization of an Indian isolate of Banana bunchy top virus based on six genomic DNA components. Virus Genes 38, 334–344. https://doi.org/10.1007/s11262-009-0331-8

Wanitchakorn, R., Harding, R. M. and Dale, J. L., 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., Harding, R. M. and Dale, J. L., 2000. Sequence variability in the coat protein gene of two groups of banana bunchy top isolates. Arch. Virol. 145, 593–602. https://doi.org/10.1007/s007050050049

How to cite this article:

Darsana Dilip, Vimi Louis, Namitha, P.M. 2021. Phylogeographic Analysis and Codon Usage Bias of Coat Protein Gene of Banana Bunchy Top Virus Isolates in Kerala. *Int.J.Curr.Microbiol.App.Sci.* 10(01): 1923-1936.
doi: https://doi.org/10.20546/ijcmas.2021.1001.224