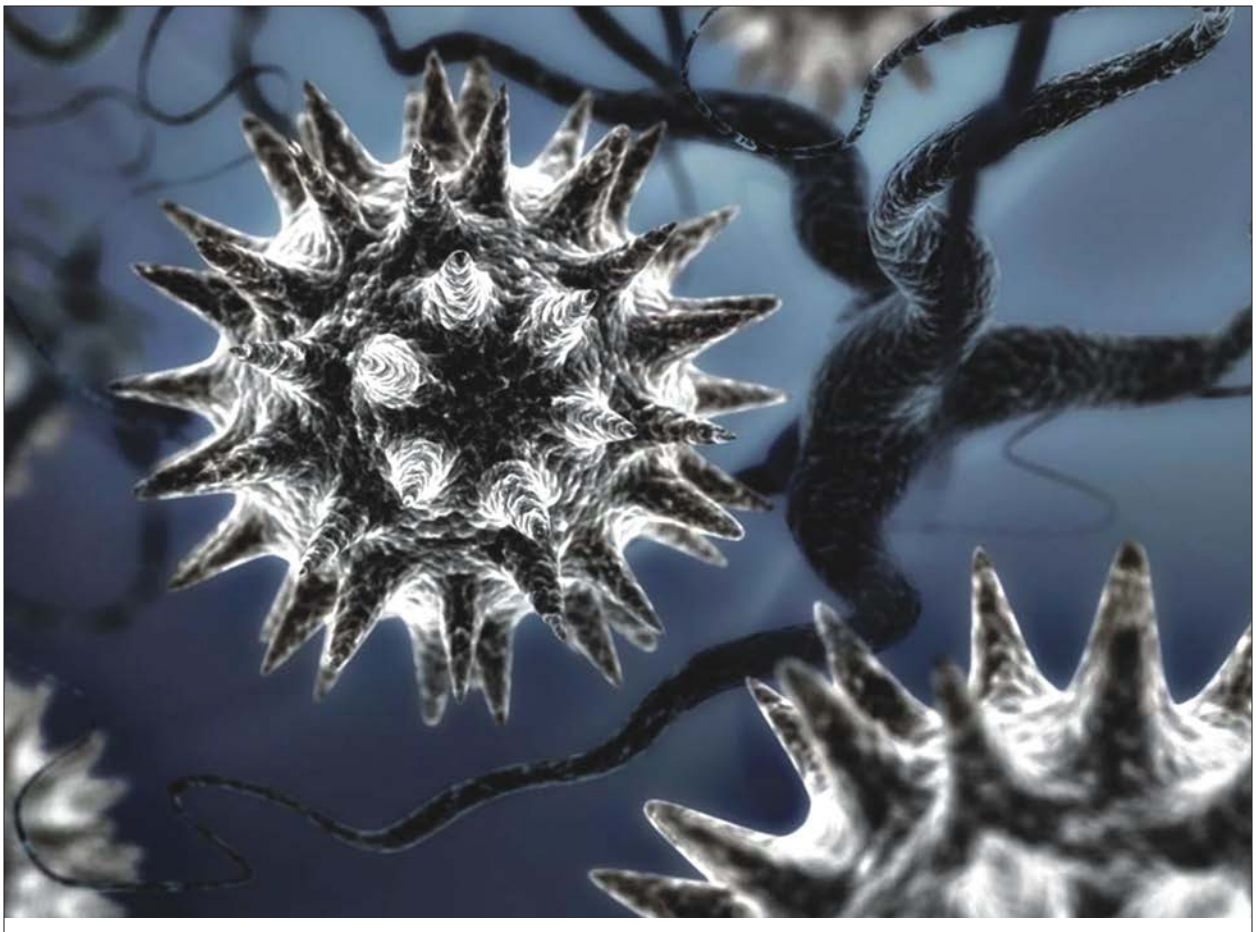

International Journal of Virology

ISSN 1816-4900

Academic Journals Inc.

www.academicjournals.com

International Journal of Virology 5 (4): 143-153, 2009
ISSN 1816-4900

Genome Sequencing, Comparison and Phylogenetic Analysis of *Citrus yellow mosaic virus* Isolates Originating from Different Citrus Species in India

[1]K.N. Gupta, [1]V.K. Baranwal, [1]B.K. Prasanna, [2]J. Singh,
[3]Q.M.R. Haq and [4]K. Gopal
[1]Plant Virology Unit, Division of Plant Pathology,
Indian Agricultural Research Institute New Delhi-110012, India
[2]College of Biotechnology, Sardar Vallabhbhai Patel University of Agriculture and Technology, Modipuram,Meerut-250110
[3]Department of Biosciences,Jamia Millia Islamia (A Central University), New Delhi-110025, India
[4]AICRP on Tropical Fruits (Citrus) Tirupati-517502, India

Abstract: A study was undertaken to find out the genomic variability in *Citrus yellow mosaic virus* (CMBV), a bacilliform virus under the genus *Badnavirus*, infecting different citrus species in India and their phylogenetic relationship with other badnaviruses. Comparison of genome sequences of four isolates of CMBV infecting different species of citrus with previously sequenced three CMBV isolates indicated variability in coding region of ORFs 1, 2 and 3 of all the CMBV isolates infecting same or different citrus species with highest variability in coding region of ORF 3. Coding region of ORF 4, 5, were also highly variable in CMBV isolates but they were highly conserved in CMBV isolates infecting Acid lime. ORF 6 was comparatively conserved and was identical in CMBV isolate infecting Acid lime. All the CMBV isolates shared maximum identity with *Cacao swollen shoot virus* (CSSV) in ORF 1 and 3 indicating that CMBV isolates are more closely related to CSSV than other badnaviruses. This study has implication in determining the diversity and diagnosis of CMBV.

Key words: *Citrus yellow mosaic virus*, *Badnavirus*, *citrus*, *Phylogeny*, *bacilliform DNA virus*

INTRODUCTION

Citrus yellow mosaic virus (CMBV), a bacilliform non-enveloped ds DNA virus causing yellow mosaic disease in citrus (Fig. 1) is grouped under the genus *Badnavirus* of family *Caulimoviridae* (Ahlawat *et al*.,1996; Huang and Hartung, 2001) and is widely distributed in India particularly on sweet orange (*Citrus sinensis* (L) Osbeck) and pummelo (*Citrus grandis* (L) Osbeck) (Ahlawat *et al*., 1996; Ahlawat, 1997). A PCR based diagnostic using primer from RT and RNase H region of CMBV has been developed (Baranwal *et al*., 2003). The virus is graft transmissible and vectored through mealy bug *Planococcus citri*. Huang and Hartung (2001) had published the first complete genome of CMBV consisted of 7559 bp and 6 ORFs. Recently complete genome sequences of two CMBV isolates infecting Acid lime and Pummelo were analyzed (Bora *et al*., 2009). Earlier we reported CMBV associated with Rangpur lime rootstock (Baranwal *et al*., 2003). New badnaviruses such as *Taro Bacilliform*

Corresponding Author: V.K. Baranwal, Plant Virology Unit, Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi-110012, India

Fig. 1: Mosaic symptoms induced by *Citrus mosaic virus* on Sathgudi sweet orange (Nagri) (CMBV SON)

virus (TaBV) *Bougainvillea spectabilis chlorotic vein banding virus* (BsCVBV) and *Dracaena mottle virus* (DrMV) have been also reported (Yang *et al*., 2003; Tsai *et al*., 2005; Lei *et al*., 2007) and their complete genome sequences are now available in GenBank. Here, we report variability in complete genome sequences of four CMBV isolates, two from Sathgudi Sweet orange (*Citrus sinensis* (L) Osbeck) and one each from Acid lime (*Citrus aurantifolia* (Christm.)) Swing and Rangpur lime (*Citrus limonia* (Osb.) and their comparison with those of other available isolates of CMBV and badnaviruses.

MATERIALS AND METHODS

Sample Collection

CMBV isolates were collected from different geographical locations in Andhra Pradesh in 2006. CMBV isolates from Sathgudi sweet orange (budded) were collected from two locations viz., Nagri (Chittor district) and Pulvendula (Cudappa, district) separated from each other by a distance of about 250 km. CMBV infected acid lime and Rangpur lime (nucellar) were collected from Tirupati district, about 40 km away from Chittor district. Further study on theses CMBV isolates was carried out from 2006 to 2008.

Electron Microscopy

The leaf-dip procedure was used for detection of virus particles associated with mosaic disease of Sweet orange (Nagri), Sweet orange (Pulvendla), Acid lime and Rangpur lime (Tirupathi). Leaf-dip preparations stained with uranyl acetate were examined under electron microscope (JEOL 100 CX-11) at the Advanced Centre for Plant Virology, Indian Agricultural Research Institute, New Delhi.

PCR

For full length genome amplification of CMBV infecting different citrus species nineteen pairs of primers were designed and synthesized (Qiagen Operon, GMBH, Germany). Out of nineteen pairs of primers, seven to eight sets of overlapping primers (17-22 nucleotides) (Fig. 2, Table 1) were used for full length genome amplification of CMBV isolates. Total DNA was extracted from infected citrus tissues using a simplified DNA extraction protocol (Baranwal *et al*., 2003). Viral DNA was amplified (~800 to ~1500 bp) using Taq DNA polymerase. Five microliter DNA was used in a 50 µL PCR reaction mix containing 0.2 µm

Table 1: Details of primer sets used for PCR amplification of full length genome of CMBV isolates (primer number based on Accession no AF347695)

Set No size	Primers and their sequences	Nucleotide No.	Expected amplicon (bp)
1	CMBV 1F-TGGTATCAGAGCTTGGTTAT	20	~1015
	CMBV1016-TTGTAAGCGTAGAAGGTA	18	
2A	CMBV897 F-AACCCCAGCAAGGCTCATCAAC	22	~1075
	CMBV1972R- CAATCATGTTTCTTGTATCCAC	22	
2B	CMBV 982- TGGTACCTTCTACGCTTACAA	21	~990
	CMBV1972R-CAATCATGTTTCTTGTATCCAC	22	
3	CMBV 1948F- TGGATACAAGAAACATGATTG	22	~1492
	CMBV 3440R- GAATCACAAGTAAGCCTCTC	20	
4A	CMBV3309F- TGATGGTCGTGAGGGTACTCA	21	~1062
	CMBV 4371R- TCCTGCTGTTGCTGTAAC	18	
4B	CMBV 3418F-TTGGAGAGGCTTACTTGTG	19	~953
	CMBV 4371R-TCCTGCTGTTGCTGTAAC	28	
5A	CMBV 4350F-ATTGTTACAGCAACAGCAGG	20	~1237
	CMBV 5587R-CGCTACCTGCTGAAAGCCAC	20	
5B	CMBV4222F- ACCACTCAGAGAGCTCGCTTACA	23	~ 982
	CMBV 5204 R- CCCAATACTTCATAGGCTCTTC	22	
6A	CMBV 5530F-AGGGGAGCACAATATTTTCC	20	~1032
	CMBV6562R-ATCTGGACAGAGCATCAGCC	20	
6B	CMBV 4998 F- CAACACCAGGCTTGCTGCACC	21	~1306
	CMBV 6205R- CATGCATCCATCCGTTTCG	19	
7A	CMBV 5894 F - TTCACAAAGGGCTTATCAAG	20	~1005
	CMBV 6899 R- GCCACCAGTTGTCTTGCTGA	20	
7B	CMBV 6542F-TGGCTGATGCTCTGTCCAG	19	~1035
	CMBV 18R-AACCAAGCTCTGATACC	17	
8	CMBV 6713 F- AGATTAGATCACCTTTAGCG	20	~833
	CMBV 230R- AGATTAGATCACCTTTAGCG	20	

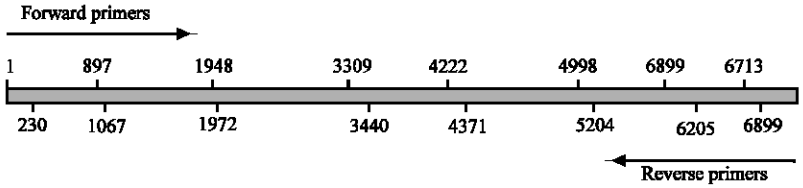

Fig. 2: Schematic representation of location of primers on the CMBV GENOME (Accession no AF347695) for amplification full genome of Citrus mosaic virus on Sathgudi sweet orange (Nagari) (CMBV SON)

each of forward and reverse primer of CMBV, Taq DNA polymerase 5 U (Sigma, USA), 5 μL of 10xPCR buffer, dNTPs each 10 mM and $MgCl_2$ 25 mM. Samples were amplified for 30 cycles, using a Mastercycler (Eppendorf, Germany). Each cycle consisted of denaturation at 94°C (30 sec), primer annealing at 53-54°C for 60 sec and extension at 72°C (60 sec), with a final extension of 10 min at 72°C. 10 μL of amplified product were separated by electrophoresis in a 1% agarose gel containing ethidium bromide at a concentration of 0.5 μg mL and photographed under UV illumination with an imaging system (BioRad XR documentation system).

Sequencing and Phylogenetic Analysis

PCR products were directly sequenced or cloned in pGEM-T Easy vector. For cloning of PCR products, amplified DNA was eluted by using gel extraction kit (SIGMA GenElute™ Gel Extraction Kit, USA). The eluted DNA from each amplicon was ligated into pGEMT easy

vector and transformed into competent *Escherischia coli* (DH5α strain) cells using standard molecular biology protocols. Restriction analysis or colony PCR was employed to confirm positive clones. At least two PCR products or clones of PCR products were sequenced in both the directions with 100% identity in the overlapping regions either at MICROSYNTH DNA sequencing facility (Switzerland) or at Department of Biochemistry, South Campus, Delhi University, Delhi (India). Sequences of overlapping cloned/PCR fragments were assembled using BioEdit sequence alignment editor version 5.09.04 (Hall, 1999). Sequences were analyzed by Blast (http:/www.ncbi.nlm.nih.gov/blast/). Sequences of other Badnaviruses for comparison were obtained from http:/www.ncbi.nlm.nih.gov/blast/. Sequence identity matrix and other basic analysis were carried out using Bioedit Sequence Alignment Editor Version 5.0.9. Multiple sequence alignments were generated using CLUSTAL W. Full genome sequences of four CMBV isolates were compared with those of three other CMBV isolates designated as CMBVSOH, CMBVAL2, CMBVPM infecting sweet orange, acid lime and pummelo respectively and 15 badnaviruses available in GenBank (Table 1). Sequence phylograms were constructed using PHYLIP package (boot-strap analysis with 1000 replicates) and rooted trees were generated using TREEVIEW software (Page, 1996).

RESULTS AND DISCUSSION

Association of bacilliform virus particles were observed in the symptomatic leaf samples of all the four citrus species. There were very few virus particles when seen under electron microscope. The particles observed were typical bacilliform shaped and measured 110-140×25-30 nm in size (Fig. 3). Full genome amplification of CMBV isolates were obtained by amplification using 7/8 sets of overlapping primer pairs (Fig. 4). After aligning all the eight sequences of PCR products, full length genome sequences were obtained. Complete genome of CMBV Sweet orange Nagri (CMBVSON, FJ617224) and CMBV Sweet orange Pulvendula (CMBVSOP, EU708316) were 7558 bp and 7497 bp in size respectively while those of CMBV acid lime Tirupati (CMBVAL1, EU 7081317) and CMBV Rangpur lime Tirupati (CMBVRL, DQ875213) were 7498 and 7522 bp in size respectively. Genome sequences of 4 CMBV isolates had non coding region ranging from 694 to 731 bp. The 3 other isolates of CMBV infecting Sweet orange, Acid lime and pummelo available in GenBank (AF347695; EU489744,

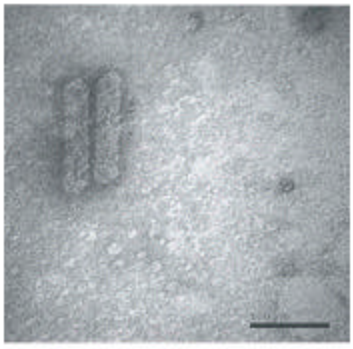

Fig. 3: Electron micrograph of CMBV associated with Rangpur lime (Tirupathi) in leaf dip preparation

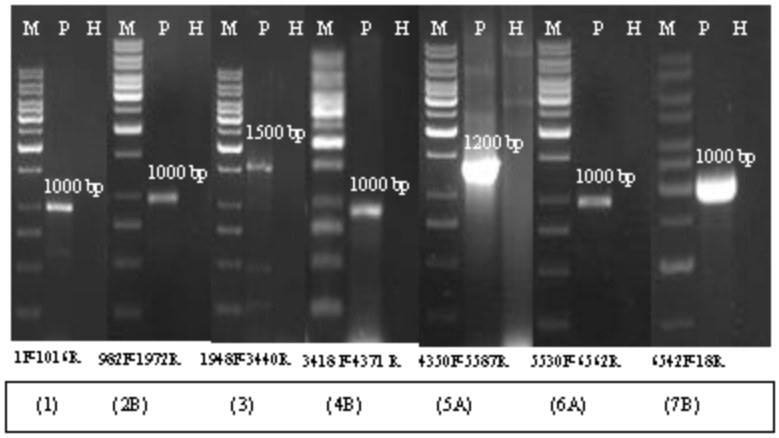

Fig. 4: Amplification of CMBVSOP with different sets of Primers. Lane M-1 kb DNA ladder, Lane P-CMBVSOP and Lane H-Healthy. Numbers in paranthesis indicates primer sets as indicated in Table 1

EU489745) have 7559, 7437 and 7487 bp in their genome with non coding region varying from 661 to 731 bp. Comparative analysis reveals that different CMBV isolates shared 87.4 to 98.9% identity in nucleotide sequence of their full genome (Table 2).

Genome architecture of CMBV isolates in the study is similar to CMBV genome reported previously. ORF finder in http://www.ncbi.nil.gov.blast was used to find out the size and number of ORFs in each isolate of CMBV. There are six open reading frames present on plus strand. ORF 1 encodes 143 amino acids in all the CMBV isolates and shows identity of 94.4 to 99.3%. ORF 2 codes for 137 amino acids in all the CMBV isolates except CMBV RL in which it encoded 138 amino acids. 89.0 to 99.2% identity is observed in ORF 2 of all isolates. Functions of ORF 1 and 2 are not known. ORF 3 is the largest ORF and codes for proteins such as movement protein, coat protein, aspartic protease, RNase H and reverse transcriptase observed by Huang and Hartung (2001). Number of amino acid encoded by nucleotides in ORF 3 varied from 1774 in the CMBVSON isolate to 1983 in CMBVSOH and identity of ORF 3 ranged from 83.7 to 98.6%. Deletions were observed in different parts of ORF 3 of all the isolates. However, the most significant deletion was observed in CMBVSON which showed a deletion of 199 amino acids at N-terminal end of coding region of ORF 3. A cystein rich zinc finger like RNA binding site ($CXCX_2CX_4HX_4C$) and a second cystein rich region ($CX_2CX_1CX_2CX_4CX_2C$) are present in putative coat protein region of ORF 3. ORF 3 also contains aspartate protease, reverse transcriptase and RNase H which are conserved among all pararetroviruses. ORF 4, ORF 5 and ORF 6 overlapped with ORF 3 except for isolate CMBVSON where ORF 6 did not overlap with ORF 3. Functions of proteins coded for by ORF 4, ORF 5 and ORF 6 are not known. ORF 4 overlaps with ORF 3 in its N terminal end. ORF 4 was highly variable in different CMBV isolates as it coded for 85 to 181 amino acids with identity in the range of 28.4 to 100%. The potential protein encoded in ORF 5 was also variable (95 to 103 amino acids) with identity of 69.0 to 100% in different isolates. ORF 6 was conserved like ORF 1 and ORF 2 and encodes 154 amino acids in all isolates except CMBVSON where it encodes 153 amino acids and did not overlap with ORF 3. Identity in amino acids of ORF 6 varies from 87.5 to 100% in all the CMBV isolates. Phylogenetic study

Table 2: Percent identity of full genome nucleotide sequences and amino acid sequences encoded in different ORFs of CMBV isolates and their comparison with other badnaviruses

Viruses	Nucleotide sequence identity (%)	ORFI		ORFII		ORFIII		ORFIV		ORFV		ORFVI	
		Size	Identity (%)	Size	Identity (%)	Size	Identity (%)	Size	Identity (%)	Size	Identity (%)	Size	Identity (%)
CMBV Isolates													
CMBVSOH	87.4-96.0	143	95.1-98.6	137	89.7-99.2	1983	87.1-95.9	178	34.6-87.1	95	71.1-90.5	154	89.6-96.1
CMBVSON	87.5-96.0	143	95.1-98.6	137	89.1-99.2	1774	83.7-87.1	164	28.4-87.1	103	69.5-95.1	153	87.2-96.1
CMBVSOP	88.1-90.2	143	97.2-99.3	137	89.0-92.7	1967	83.7-93.8	85	28.4-38.6	103	71.4-95.1	154	87.8-95.4
CMBVAL1	87.8-93.0	143	95.1-99.3	137	89.0-91.9	1968	84.9-95.9	181	38.6-100	95	71.5-100	154	88.4-100
CMBVAL2	87.6-98.9	143	95.1-98.6	137	90.5-97.8	1976	84.4-98.6	181	38.6-100	95	71.5-100	154	88.4-100
CMBVRL	88.6-91.6	143	95.8-99.3	138	89.1-93.4	1954	83.4-94.3	148	29.9-55.8	95	69.0-90.5	154	87.2-89.6
CMBVPM	87.4-98.9	143	94.4-98.6	137	89.7-97.8	1979	84.2-93.8	181	37.5-98.3	97	69.5-96.9	154	87.8-98.7
Otherbadnaviruses													
CSSV	48.4-49.1	143	53.8-55.9	145	18.2-19.4	1816	49.3-50.3	113	03.0-07.2	131	6.8-9.1	-	-
BSV	36.5-36.9	176	19.8-20.4	132	19.7-21.7	1900	30.7-33.0	-	-	-	-	-	-
BSAuV	36.5-36.9	176	19.8-20.4	132	19.7-21.7	1900	30.7-33.0	-	-	-	-	-	-
BSTrV	36.1-36.4	176	22.0-22.5	134	19.4-20.8	1709	33.8-34.5	-	-	-	-	-	-
BSGDV	36.2-36.4	177	22.0-22.5	134	19.4-20.8	1709	33.8-34.5	-	-	-	-	-	-
BSGFV	36.4-37.0	177	19.3-19.8	134	17.4-18.1	1832	32.6-34.7	-	-	-	-	-	-
BSMyV	38.2-38.6	176	19.8-20.4	132	19.7-21.7	1900	30.7-33.0	-	-	-	-	-	-
BSOLV	37.6-37.9	175	22.0-22.5	112	19.3-20.8	1832	31.8-34.0	-	-	-	-	-	-
BsCVBV	33.6-34.2	132	16.3-16.9	147	16.0-17.3	2252	28.1-30.4	198	04.0-07.0				
ComYMV	37.2-37.7	200	17.3-17.8	135	12.9-14.9	1886	33.4-33.8	-	-	-	-	-	-
DaBV	47.9-48.4	143	34.4-35.8	125	31.8-34.7	1295	24.4-27.8	59	03.0-0.03	-	-	-	-
DrMV	41.6-42.6	149	34.6-35.3	131	25.3-26.9	1916	33.6-36.5	103	09.0-12..1	91	04-13.8	139	21.2-22.7
KTSV	36.4-37.0	173	14.7-15.3	124	20.5-22.6	1941	30.0-32.5	-	-	-	-	-	-
SCBV	34.9-35.3	185	16.4-17.0	123	21.1-24.6	1912	26.8-28.9	-	-	-	-	-	-
TaBV	37.9-38.1	146	33.0-34.6	144	21.9-23.9	1881	30.7-33.7	-	-	-	-	-	-

-: Absent, Accession number of CMBV and other badnavirus sequences viz., *Banana streak virus Acuminata* (BSAuV), DQ092436; *Banana streak virus* (BSV), NC_008018; *Banana streak virus Tr* Virus (BSTrV), DQ859899; *Banana streak GD virus* (BSGDV), DQ451009; *Banana streak OL virus* (BSOLV), NC_003381 ; *Banana streak Mysore virus* (BSMyV), NC006955 ; *Banana streak GF virus* (BSGFV), NC_007002; *Cacao swollen shoot virus* (CSSV), NC001574; *Citrus yellow mosaic virus SON* (CMBVSON) FJ617224; *Citrus yellow mosaic virus SOH* (CMBVSOH), AF347695; *Citrus yellow mosaic virus* RL (CMBVRL), DQ875213; *Citrus yellow mosaic virus* PM (CMBVPM), EU489745; *Citrus yellow mosaic virus* AL 2 (CMBVAL 2), EU489744; *Citrus yellow mosaic virus* AL1 (CMBVAL1), EU7081317; *Citrus mosaic virus* SOP (CMBVSOP) , EU489744; *Commelina yellow mottle virus* (ComYMV), X52938; *Dioscorea bacilliform virus* (DaBV), DQ822074; *Dracaena mottle virus* (DrMV), DQ473478; *Kalanchoe top-spotting virus* (KTSV), NC-004540; *Sugarcane bacilliform virus* (SCBV), NC-003031; *Taro bacilliform virus* (TaBV), AF357836; *Bougainvillea spectabilis chlorotic vein bending virus* (BsCVBV) AY532653

indicated that CMBV infecting sweet oranges grouped together in ORF 1 while CMBV infecting Acid lime, Rangpur lime and Pummelo grouped together separately. However, the grouping of CMBV infecting different citrus species did not group in the same manner in ORF 2 and 3 (Fig. 4). A promoter element like TATATAA box is present in the intergenic region. Upstream of the TATA box are CACAAT and TGACG sequences similar to that found upstream of the TATA box in the 35S promoter of *Cauliflower mosaic virus* (CaMV) (Odell *et al*., 1985). Downstream of the putative CMBV promoter was a possible polyadenylation signal (AATAAA) Boeke and Corces (1989) similar to that previously reported in CMBV and also seen in other badnaviruses (Geering *et al*., 2000). Most of the badnaviruses have three ORFs except CMBV (6 ORFs), *Cacao swollen shoot virus* (CSSV) (5 ORFs), *Dracaena mottle virus* (DrMV) (7 ORFs), *Dioscorea bacilliform virus* (DaBV) and *Bougainvillea spectabilis chlorotic vein banding virus* (BsCVBV) (4 ORFs). Comparative sequence analysis of complete nucleotide genome of CMBV isolates with other badnaviruses showed that maximum identity of 48.4-49.1% was observed with CSSV followed by DaBV, (47.9-48.4 %) and DrMV (41.6-42.6 %). CSSV showed maximum identity in ORF 1 and ORF 3 with CMBV isolates but DaBV and DrMV showed more identity in ORF II with CMBV isolates in comparison to CSSV. CMBV isolates also have some identity in ORF 4, 5 and 6 with that of DrMV with maximum identity in ORF 6 (21.2 - 22.7%). Phylogenetic tree analysis indicated that different isolates of CMBV grouped differently in ORF 1, 2 and 3 but all the isolates of CMBV were branched out together with CSSV in ORF 1 and they in turn grouped with BsCVBV and DrMV and TaBV (Fig. 5). In ORF 2 CMBV isolates grouped with CSSV, DrMV and DaBV and they in turn grouped with TaBV and other badnaviruses (Fig. 5).

However, phylogenetic analysis of ORF 3 of badnaviruses indicated that only CSSV and DaBV grouped with CMBV isolates and these in turn grouped with DrMV, TaBV and BsCVBV which formed a separate group with a common tree with *Commelina yellow mottle virus* (ComYMV), the type species of genus *Badnavirus* and other badnaviruses such as species of *Banana streak virus*, Kalanchoe top-spotting virus (KTSV) and *Sugarcane bacilliform virus* (SCBV) (Fig. 5). All the CMBV isolates shared maximum identity with *Cacao swollen shoot virus* (CSSV) in ORF 1 and 3 indicating that CMBV isolates are more closely related to CSSV than other badnaviruses. Among all CMBV isolates CMBVSON isolate is quite distinct as it has smaller ORF 3 and 6 is not overlapped with ORF 3.

Citrus mosaic was first reported in Sathgudi Sweet orange in 1975 but association of bacilliform virus particle with the disease was confirmed by Ahlawat *et al*. (1996) in Sweet orange and Pummelo. Its genome was characterized as double stranded DNA and its complete nucleotide sequence was determined by Huang and Hartung (2001). The genome of CMBV infecting Sathgudi Sweet orange consisted of 7559 nucleotides and had 6 ORFs (Huang and Hartung, 2001). When CMBV genome was first characterized, the viral DNA was obtained from the purified preparation of virus from the infected citrus plants ((Huang and Hartung, 2001). However, in the present study, total genomic DNA was isolated from citrus infected with CMBV using a simplified DNA extraction protocol (Baranwal *et al*., 2003). Bora *et al.* (2009) determined the nucleotide sequences of CMBV genome infecting Acid lime and Pummelo which were 7473 and 7487 bp. Present study demonstrated that the nucleotide size of genome of CMBV infecting different citrus species was variable and ranged from 7497 to 7558 bp. The variability in the genome size of CMBV isolates was due to deletion/addition events at several places in the genome particularly in ORF 3 and intergenic regions. The deletion or addition of nucleotides may be due to recombination events occurring naturally in viruses and may not depend on host species. Generally the ORF 4, 5 and 6 are contained in ORF 3 in all CMBV isolates except the one infecting Sweet orange from

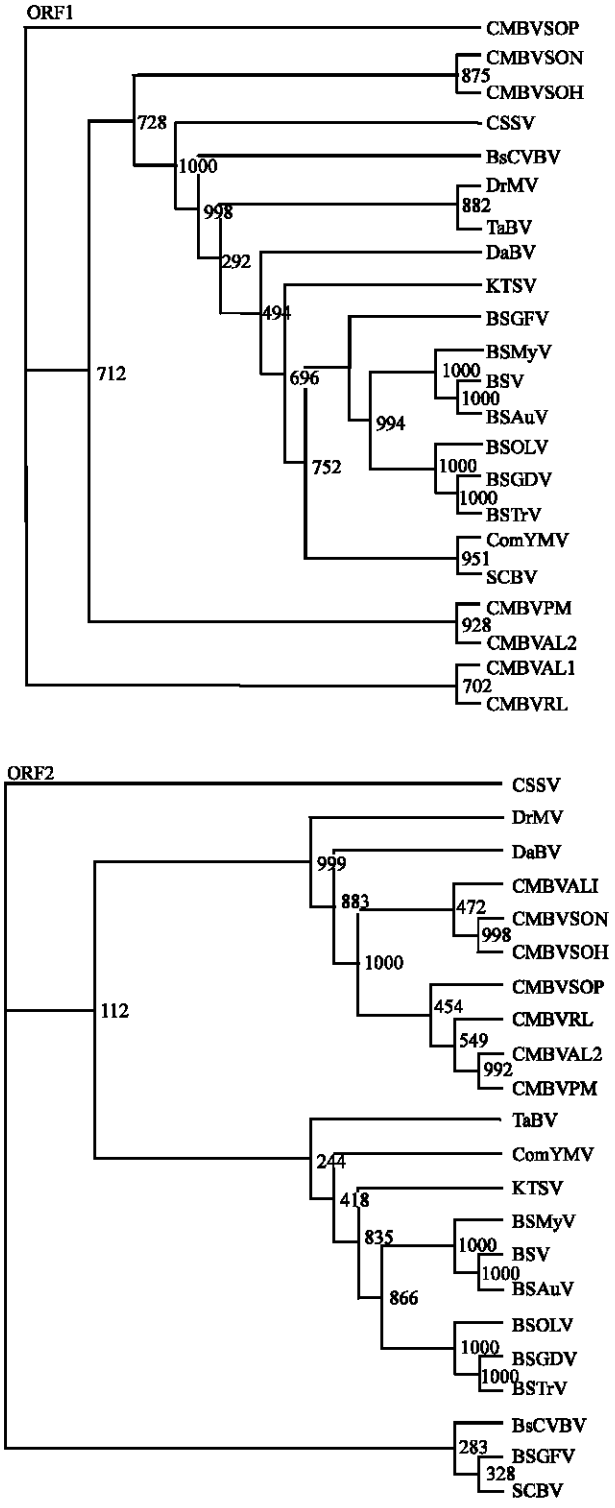

Fig. 5: Continued

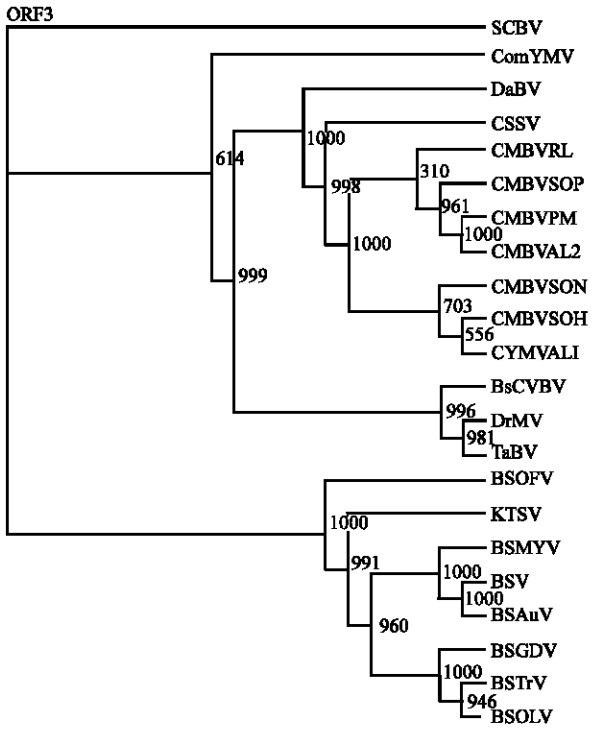

Fig. 5: Cluster dendogram showing the relationship of ORFs 1, 2 and 3 encoded by nucleotides in the genome of *Citrus yellow mosaic virus* isolates and other badnaviruses. The dendogram was constructed with bootstrapping (1000 replicates) using TREEVIEW Software. The number at nodes refer to number of times branching was supported. Accession number for compared CMBV ORFs and other badnavirus sequences; Banana streak Acuminata virus (BSAuV), DQ092436; *Banana streak virus* (BSV), NC_008018; *Banana streak Tr virus* (BSTrV), DQ859899; *Banana streak GD virus* (BSGDV), DQ451009; *Banana streak OL virus* (BSOLV), NC-003381; *Banana streak Mysore virus* (BSMyV),NC006955; *Banana streak GF virus* (BSGFV), NC-007002; *Cacao swollen shoot virus* (CSSV), NC001574; *Citrus yellow mosaic virus* SON (CMBVSON), FJ617224; *Citrus yellow mosaic virus* (CMBVH), AF347695; *Citrus yellow mosaic virus* RL (CMBVRL), DQ875213; *Citrus yellow mosaic virus* PM (CMBVPM), EU489745; *Citrus yellow mosaic virus* AL1 (CMBVAL1), EU7081317; *Citrus yellow mosaic virus* AL2 (CMBVAL2), EU489744; *Citrus mosaic virus* SOP (CMBVSOP), EU708316; *Commelina yellow mottle virus* (ComYMV), X52938; *Dioscorea bacilliform virus* (DaBV), DQ822074; *Dracaena mottle virus* (DrMV), DQ473478; *Kalanchoe top-spotting virus* (KTSV), NC-004540; *Sugarcane bacilliform virus* (SCBV), NC-003031; *Taro bacilliform virus* (TaBV), Australia;AF357836: *Bougainvillea spectabilis chlorotic vein banding virus* (BsCVBV), AY532653

Nagri (CMBVSON) where ORF 6 is not contained in ORF 3 and does not show much diversity compared to ORFs 4 and 5. It is interesting to note that there was 100% identity in coding region of ORF 4, 5 and 6 of two CMBV isolates infecting Acid lime (CMBVAL1 and CMBVAL2). However, ORFs 4, 5 and 6 were not conserved in CMBV isolates infecting sweet

orange (CMBVSOH, CMBVSON and CMBVSOP). Genomic heterogeneity is a common feature with many badnaviruses (Geijskes *et al*., 2002; Yang *et al*., 2003; Muller and Sacky, 2005) and our study also demonstrated that there are sequence variability in coding regions of ORF 1 2 and 3 of CMBV infecting same citrus species or different citrus species. In recent times, new badnaviruses such as BsCVBV, DaBV, DrMV, KTSV and TaBV have been characterized from different countries but none of them showed closer phylogenetic relationship with CMBV than CSSV. Genome characterization of CMBV isolates will be useful for future studies on functions of viral proteins. It will also help in developing strategies for its management. The promoter elements in intergenic region of CMBV have potential for its use in transgene expression in citrus and other crops. Sequencing will also be useful in designing primers for detection of different isolates of CMBV in PCR.

ACKNOWLEDGMENTS

This study was supported by Department of Biotechnology, Government of India. We are also grateful to the Head, Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi-12 for providing necessary facilities.

REFERENCES

Ahlawat, Y.S., 1997. Viruses, greening bacterium and Viroids associated with citrus (*Citrus species*) decline in India. J. Agric. Sci., 67: 51-57.

Ahlawat, Y.S., R.P. Pant, B.E.L. Lockhart, M. Srivastava, N.K. Chakraborty and A. Varma, 1996. Association of a badnavirus with citrus mosaic disease in India. Plant Dis., 80: 590-592.

Baranwal, V.K., S. Majumder, Y.S. Ahlawat and R.P. Singh, 2003. Sodium sulphite yields improved DNA of higher stability for PCR detection of Citrus yellow mosaic virus from citrus leaves. J. Virol. Methods, 112: 153-156.

Boeke, J.D. and V.G. Corces, 1989. Transcription and reverse transcription of retrotransposons. Annu. Rev. Microbiol., 43: 403-434.

Bora, B.K., A.M.A. Johnsom, G.D.V.R. Sai and I. Dasgupa, 2009. Sequencing and computational analysis of complete genome sequences of Citrus yellow mosaic badnavirus from acid lime and pummelo. Virus Genes, 39: 137-140.

Geering, A.D.W., L.A. McMichael, R.G. Dietzgen and J.E. Thomas, 2000. Genetic diversity among banana streak virus isolates from Australia. Phytopathology, 90: 921-927.

Geijskes, R.J., K.S. Braithwaite, J.L. Dale, R.M. Harding and G.R. Smith, 2002. Sequence analysis of an Australian isolate of sugarcane bacilliform badnavirus. Arch. Virol., 147: 2393-2404.

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid. Symp. Ser., 41: 95-98.

Huang, Q. and J.S. Hartung, 2001. Cloning and sequence of an infectious clone of citrus yellow mosaic virus that can infect sweet orange via *Agrobacterium* mediated inoculation. J. General Virol., 82: 2549-2558.

Lei, S., G. Shang, H. Yanwei, J. Chaoqun and W. Dickson., 2007. Complege genomic sequence of Dracaena mottle virus, a distinct badnavirus. Virus Genes, 35: 423-429.

Muller, E. and S. Sacky, 2005. Molecular variability analysis of five new complete *Cacao swollen* shoot virus genomic sequences. Arch. Virol., 150: 53-66.

Odell, J.T., F. Nagy and H.C. Nam, 1985. Identification of DNA sequences required for the activity of the mosaic virus 35S promoter. Nature, 313: 810-812.

Page, R.D.M., 1996. TreeView: An application to display phylogenetic trees on personal computers. Comp. Applied Biosci., 12: 357-358.

Tsai, C.H., H.J. Su, Y.C. Liao and T.H. Hung, 2005. First report of bougainvillea spectabilis chlorotic vein-banding virus infecting bougainvillea plants in Taiwan. Plant Dis., 89: 1363-1363.

Yang, I.C., G.J. Hafner, J.L. Dale and R.M. Harding, 2003. Genomic characterization of taro bacilliform virus. Arch. Virol. 148: 937-949.