



**Molecular characterization and Phylogenetic
Analysis of Citrus *Mosaic Badna Virus* (CMBV)
Associated with Sathgudi sweet orange**

BY

Kailash N. Gupta

Scientist (Plant Pathology)

***Department of Plant Pathology, Jawaharlal Nehru
Krishi Vishwa Vidyalaya, Jabalpur, M. P., India,***

Citrus is grown in 140 countries of the world and is one of the choicest fruit having high consumer's preference both as fruit as well as its refreshing processed juice.

Introduction

Citrus

Family : Rutaceae

Tribe : Citreae

Sub Tribe: Citrineae

Genus: Citrus

Area in India: 0.56 mha

Production in India: 4.58 mt

Productivity in India: 12-13 tons/ha

(The low productivity of citrus in India is due to many abiotic and biotic stresses.)

World Productivity: 28-30 tons/ha.

Important members : Sweet Orange, Acid Lime, Rangpur Lime, Pummelo, Rough Lemon, Mandarin and Grape Fruit

Major Citrus growing areas in India



Important diseases: Citrus tristeza, Indian Citrus Ringspot, Citrus mosaic virus, Citrus Greening, Citrus exocortis, Citrus Gummosis, Citrus canker etc.

Citrus mosaic virus disease caused by Citrus mosaic Badnavirus (CMBV), is an important disease in citrus growing areas particularly southern part of India.

Importance of work: for determine variability in the viral genome.

First report in India -

Dakshinamurti and Reddy, 1975

– Sathgudi Sweet orange - Andhra Pradesh

Ahlawat, 1985

– Khasi Mandarin – N. E.India

Disease Incidence

10 to 70%

(Ahlawat *et al.*, 1996)

77% reduction in Fruit yield

10% reduction in Juice and ascorbic acid content

(Reddy G.S *et al.*, 1985)

CMBV

Ahlawat et al., 1996 did partial characterization of citrus mosaic virus (CMBV) which they name as citrus yellow mosaic badnavirus. However, as per 8th edition of ICTV, The virus has been named as *citrus mosaic virus*

(According to Huang and Hartung, 2001)

- Family: Caulimoviridae
- Genus: Badnavirus
- Non enveloped bacilliform virus
- 30 X 120-150nm
- Circular ds DNA
- 7559 bp.
- six ORFs
- Pararetrovirus
- Huang and Hartung, 2001, First confirm the full genome of citrus mosaic badnavirus in sathgudi sweet orange

Molecular characterization of *Citrus mosaic virus*



Symptoms induced by *Citrus mosaic virus* on sathgudi sweet orange under field condition.

Method Materials:

Collection of culture: Survey of sathgudi sweet orange (*Citrus sinensis* (L.) Osb) was done in the Nagri village of Chittor, district which is about 40 km away from Tirupati in Andhra Pradesh during November 2006. Two orchards of sathgudi sweet orange which was having approximately 440 plants were surveyed for presence of mosaic and greening disease symptoms. Symptomatic leaves from sathgudi sweet orange plants showing mosaic symptoms were collected and stored at -80 0C.

Graft transmission and establishment of culture: Bud sticks from plants showing mosaic symptoms were wedge grafted on 10 plants each of 1 year old healthy seedlings of sweet orange (*Citrus sinensis* (L) Osbeck and acid lime (*Citrus aurantifolia* (Christm)). The grafted and healthy plants were maintained in the insect proof glasshouse. The grafted plants were observed for about 6 months and plants showing symptoms were used for further studies.

Electron microscopy of CMBV: The leaf-dip preparation was used for detection of CMBV and stained with 20% uranyl acetate examined under transmission electron microscope (JEOLJEM-1011) at Plant virology unit, Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi 110012.


Isolation of total DNA from plant leaves

1. The total genomic DNA was isolated from healthy and CMBV infected leaf from glass house using DNeasy plant mini kit method (**Qiagen GmbH, Hilden, Germany**). The method is described below.


2. 100 mg of diseased and healthy leaf tissue for CMBV were taken separately for total DNA isolation. The samples were washed and ground to powder with the help of pestle and mortar in liquid nitrogen. The powdered samples were immediately transferred to autoclaved 1.5 ml eppendorf tubes.

3. To this 400 μ l of AP1 buffer and 4 μ l of RNase A (100mg/ml) were added and vortexed vigorously for proper mixing. The mixtures were incubated for 10 min at 65°C. 4. The tubes were inverted 2-3 times during incubation period. After 10 mins 130 μ l of AP2 buffer was added to lysates, properly mixed and kept in ice for 5 mins. Lysates were centrifuged at 12,000 rpm for 5 mins.


5. The supernatant lysates were transferred into the QIAshredder spin columns (lilac colour) placed in 2 ml collection tube and centrifuged at 12,000 rpm for 2 mins.



6. The flow through were transferred to a new eppendorf tubes without disturbing cell debris. To the flow through 1.5 volumes (of flow through) of AP3/E buffer was added and mixed properly by pipetting.



7. The 650 µl of mixtures were transferred into the DNeasy columns placed in a 2 ml collection tubes. 8. Centrifuged for 1 min at ≥ 8000 rpm and flow through was discarded. Step 8 was repeated for remaining mixture, flow through and collection tubes were discarded.



9. The DNeasy mini spin columns were placed in a new 2 ml collection tubes. 10. 500 µl of AW buffer was added to columns and centrifuged for 1 min at ≥ 8000 rpm. 11. Flow through was discarded and columns with collection tubes were reused in step 11.

12. 500 µl of AW buffer was added to columns and centrifuged for 2 min at maximum speed to dry the membrane. 13. The DNeasy columns were transferred to a new 2 ml microcentrifuge tubes and 75 µl of preheated (65°C) AE buffer was added directly into DNeasy column membrane.

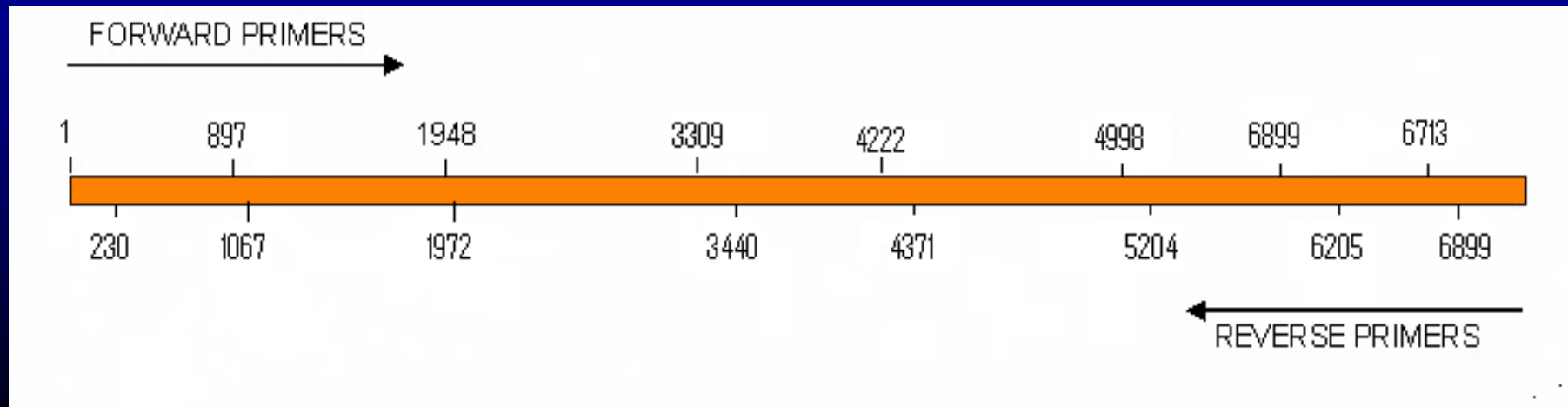


14. Columns were incubated at room temperature and centrifuged for 1 min at ≥ 8000 rpm to elute DNA from the membrane of the columns. 15. 75 µl of collected DNA was stored at -20°C for further use.

Primers for CMBV genome amplification

Eight sets of overlapping primers (17-22 nucleotides) were designed manually for full length genome amplification of citrus mosaic virus (CMBV) associated with sweet orange, Nagri isolate (CMBVSON). The full length genome sequence of CMBV available in GenBank (accession number AF3476695) was used for designing and synthesis of primers. The primers were synthesized at Biobasic, Canada.

Sweet orange – Andhra Pradesh (CMBVSOH) Accession Number : [AF3476695](#)



Details of primer sets used for PCR amplification of full length genome of CMBV sweet orange, Nagri (CMBVSON)

Primer Set Nos.	Name of the primer and their sequences	Nucleotide Nos.	Annea-ling temperature	Exten- sion time	Expected amplicon size based on CMBV sequence accession No. AF 347665 inGen bank
1	CMBV 1F- TGGTATCAGAGCTTGGTTAT CMBV1016- TTGTAAGCGTAGAAGGTA	20 18	54 ⁰ C	60s	1015 bp
2	CMBV897 F - AACCCCAGCAAGGCTCATCAAC CMBV1972R- CAATCATGTTTCTTGTATCCAC	22 22	54 ⁰ C	60s	1075bp
3	CMBV 1948F- TGGATAACAAGAAACATGATTG CMBV 3440R- GAATCACAAGTAAGCCTCTC	22 20	53 ⁰ C	90s	1492 bp
4	CMBV3309F- TGATGGTCGTGAGGGTACTCA CMBV 4371R- TCCTGCTGTTGCTGTAAC	21 18	50 ⁰ C	60s	1062bp
5	CMBV4222F- ACCACTCAGAGAGCTCGCTTACA CMBV 5204 R- CCCAATACTTCATAGGCTCTTC	23 22	53 ⁰ C	60s	982 bp
6	CMBV 4998 F- CAACACCAGGCTTGCTGCACC CMBV 6205R- CATGCATCCATCCGTTTCG	21 19	50 ⁰ C	90s	1306 bp
7	CMBV 5894 F - TTCACAAAGGGCTTATCAAG CMBV 6899 R- GCCACCAGTTGTCTTGCTGA	20 20	52 ⁰ C	60s	1005 bp
8	CMBV 6713 F- AGATTAGATCACCTTTAGCG CMBV 230R- AGATTAGATCACCTTTAGCG	20 20	54 ⁰ C	60s	~833 bp

PCR amplification

PCR mix and PCR profile used for amplification of CMBV

The composition of PCR mix used for amplification of genome of CMBV is given below,

Sterile distilled water	31.50 μl
10 x PCR buffer	5.00 μl
25mM MgCl₂	3.00 μl
10mM dNTPs	1.00 μl
Forward primer (100ng/μl)	2.00 μl
Reverse primer (100ng/μl)	2.00 μl
Taq polymerase (5units/μl)	0.50 μl
Template DNA	5.00 μl
Total PCR mix volume	50.00 μl

For amplification of the target DNA, following PCR profile was used in the Eppendorf Thermal cycler machine AG 22331 (Germany)

94 $^{\circ}$C	5 minutes	
94 $^{\circ}$C	30 second	
53 $^{\circ}$C - 62$^{\circ}$C*	1 minutes	} 30 cycles
72 $^{\circ}$C	1 minute	
72 $^{\circ}$C	10 minutes	

Cloning of amplified products

The vector pDrive 322 (3850bp) (Qiagen, Germany), was used for the cloning of eluted PCR products of CMBVSON. There is a single 3' and 5' U overhang in pDrive322 vector at the insertion sites (multiple cloning sites) which prevent self ligation of the plasmid vector

Ligation and transformation

Qiagen PCR Cloning plus kit (Qiagen, Germany) was used for ligation and Transformation of PCR amplified product

Ligation

i. After thawing 2x ligation master mix and pDrive322 cloning vector DNA, they were placed on ice.

ii. A ligation reaction mixture was prepared according to the following scheme

Component	Volume/reaction
Vector (50ng/μl)	1.0 μl
PCR product	4.0 μl
Ligtaion mix (2X)	5.0 μl
Total	10.0 μl

iii. The ligation reaction mixture was mixed and then incubated for 30 minutes at 4-16 °C.

iv. The ligation reaction mixture was then used for transformation or stored at -20 °C until use. (15-18 hours

Transformation

i. Two tubes of Qiagen EZ competent cells were thawed on ice. SOC medium was thawed and brought to room temperature.



ii. 5 μ l of ligation reaction mixture was added to the tube of Qiagen EZ competent cells, mixed gently, and incubated on ice for 5 minutes.



iii. The tube was heated in a 42°C in water bath or heating block for 30 seconds without shaking. iv. The tube was then incubated on ice for 2 minutes.



v. 250 μ l SOC medium was added to the tube and 100 μ l of each transformation mixture was then directly plated on to Luria agar plates containing ampicillin (100 μ g /ml), Xgal (80 μ g/ml and IPTG (50 μ m).



V1. The plate was then incubated at room temperature until the transformation mixture had been absorbed in to the agar. The plate and incubate was inverted at 37°C over night

Selection of transformants and screening for the positive clones

The transformants were selected on the basis of blue/white colony colour. The white colonies were selected and subsequently master plates were prepared using Luria agar containing Ampicillin (100 µg/ml), X-gal (200 µg/ml) and IPTG (0.1M) (Sambrook and Russell, 2001). In master plate individual white colonies were streaked in each grid using autoclaved tooth picks. In order to select positive clones carrying the desired insert, colony PCR was done.

Screaming of positive clones by colony PCR

Transferred small amount of bacterial cells from selected recombinant colonies (two to four) cells from individual grids were using sterile toothpicks to PCR tubes (200 µl) containing the PCR mixture bacterial cells served as template in PCR mix. Amplification conditions described earlier were followed and PCR products were electrophoresed in 1.0% agarose gel containing ethidium bromide.

Nucleotide sequencing of clones and *in silico* analysis

Clones obtained from PCR products of primer sets 2, 3, 4 and 5 to obtain full length sequence of CMBVSON

Clone	Sequence Number	Expected length (bp)
p-Drive CMBVSON 2	897F-1972R	1075
p-Drive CMBVSON 3	1948F-3440R	1492 bp
p-Drive CMBVSON 4	3309F-4371R	1062bp
p-Drive CMBVSON 5	4222F-5204 R	982bp

PCR Products obtained with 4 other primers sets 1, 6, 7,8 and used for full genome sequencing of CMBVSON

PCR product	Sequence Number	Length (bp)
CMBV N-1	1F-1016R	~1015 bp
CMBV N-6	4998 F-6205R	~1306 bp
CMBV N-7	5894 F-6899 R	~1005 bp
CMBV N-8	6713 F- 230R	~833 bp

Nucleotide sequencing of clones and *in silico* analysis

Selected recombinant clones of CMBV (CMBVSON-2, CMBVSON-3, CMBVSON-4, CMBVSON-5 with insert of approximate size ~1075, ~1492 bp, ~1062bp, ~982bp (Table 3.2) and PCR Product (CMBVSON-1, CMBVSON-6, CMBVSON-7, CMBVSON-8 (Table 3.3) with ~1015 bp, ~1306 bp, ~1005 bp, ~833 bp sizes were sequenced at Department of Biochemistry, DNA sequencing facility, South Campus, New Delhi

Sequences were analyzed by Blast (<http://www.ncbi.nlm.nih.gov/blast/>) (Altschul *et al.*, 1997). After sequence comparison checking homology all the sequences were aligned to get full length genome sequence by deleting the one of the overlapping regions between sequences. Protein coding region open reading frame (ORF) was searched by Bioedit as well as by ORF finder available at 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 (Hall, 1999). Multiple sequence alignments were generated using Clustal W (Thompson *et al.*, 1994). The full nucleotide and amino acids sequences of other badnaviruses for comparison were obtained from GenBank and multiple alignments at nucleotide and amino acid were generated.

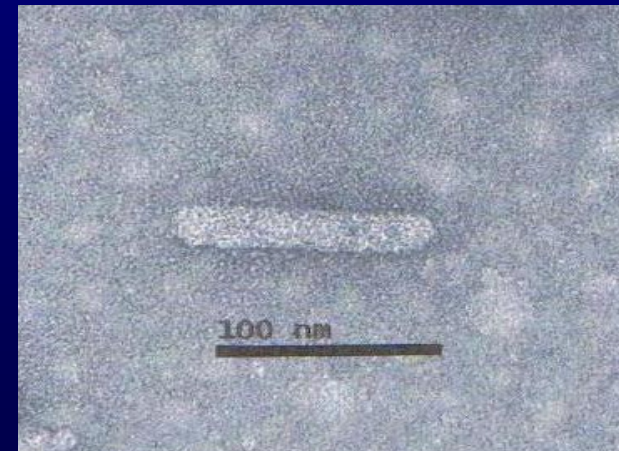
Results

Graft transmission and establishment of culture



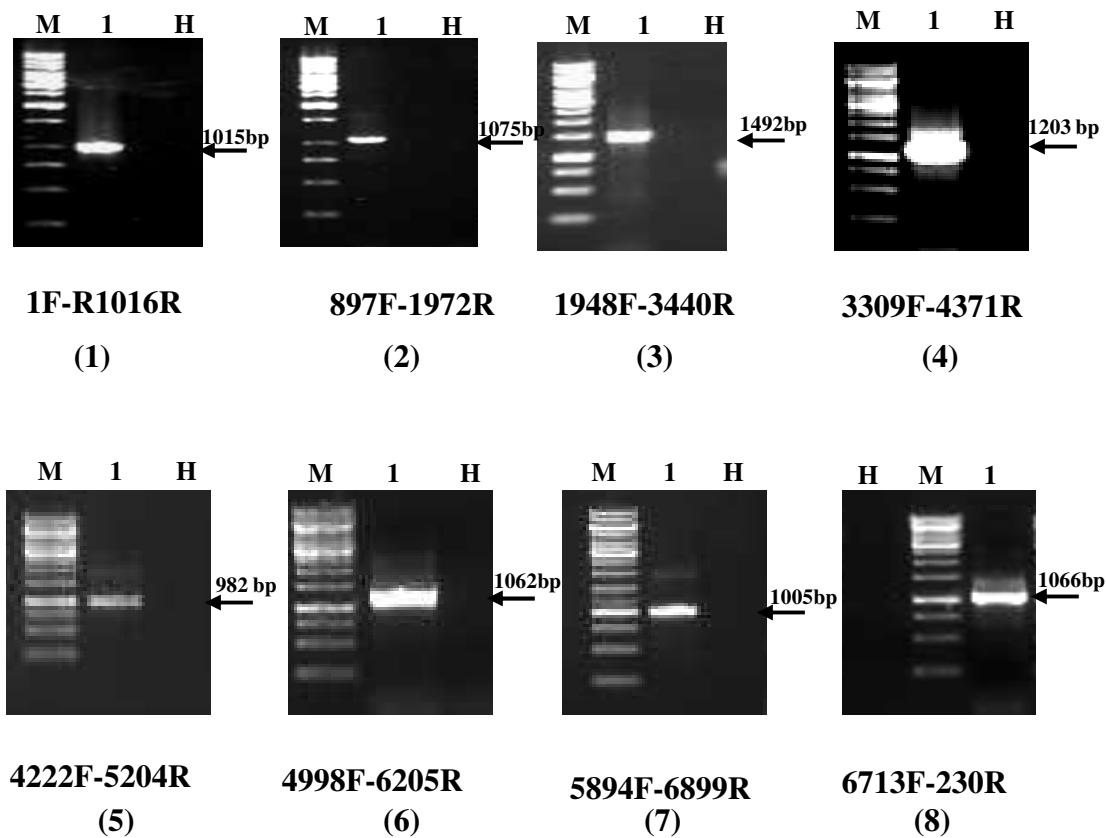
Symptoms developed on sweet orange citrus plant after graft inoculation of *Citrus mosaic virus* infecting sathgudi sweet orange plant .

Electron microscopy of CMBV:



Electron micrograph showing 120x30nm bacilliform particle of *citrus mosaic virus* (CMBV) in leaf dip preparation (magnification 1,00,000

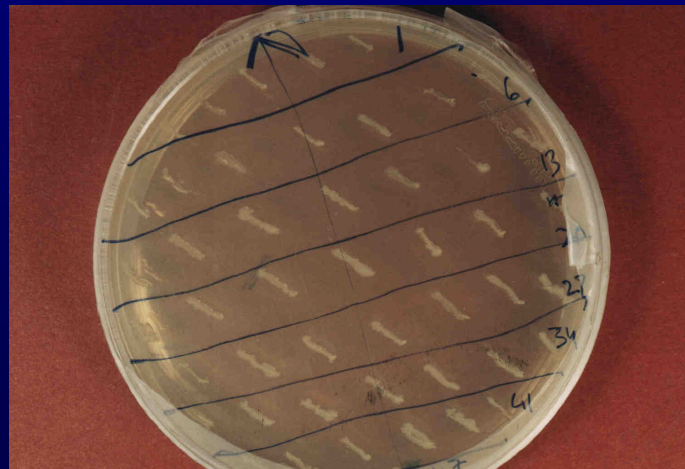
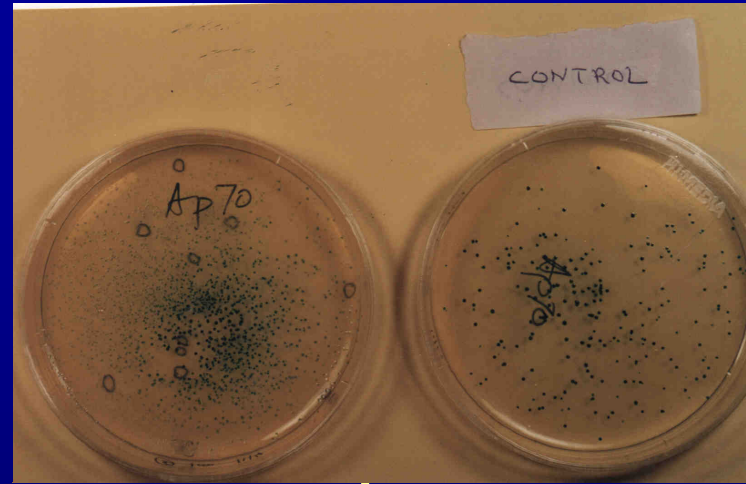
PCR amplification of CMBVSON



Gel electrophoresis of PCR products of genome of CMBVSON with eight sets of primers

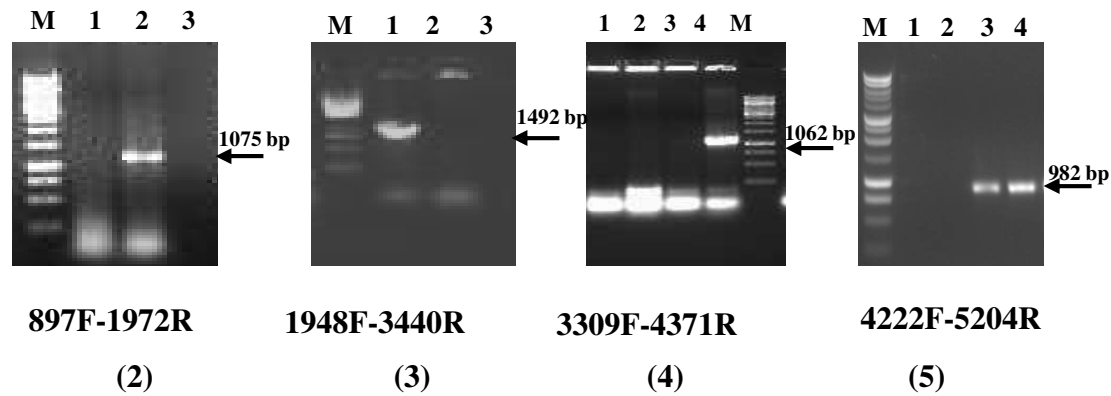
Cloning

Blue and white colonies



Master plate

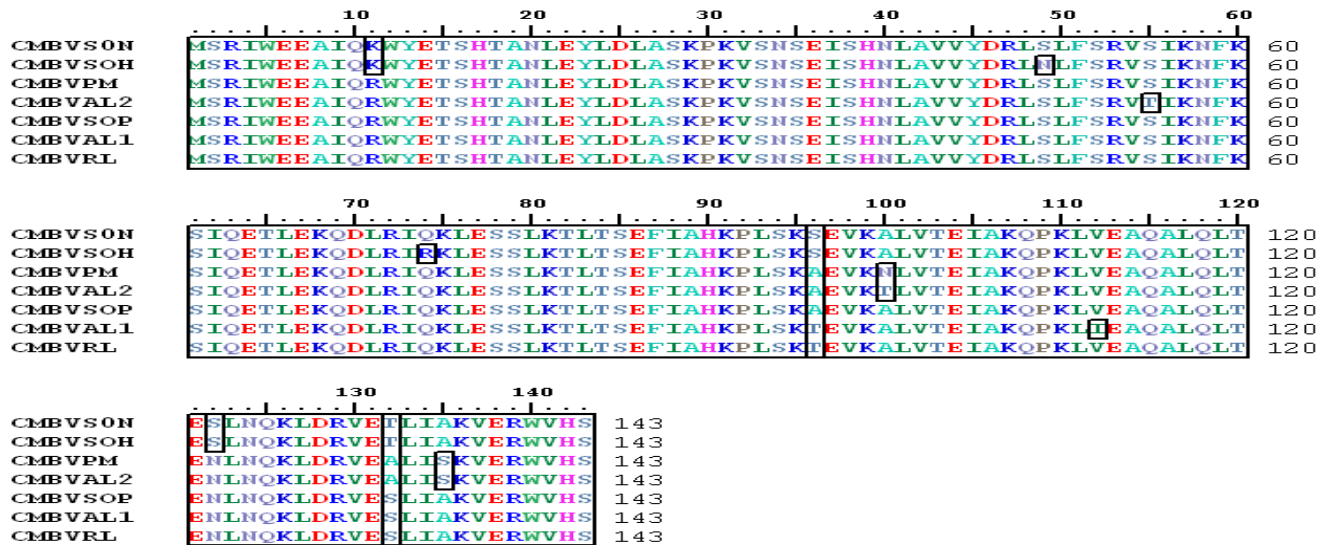
Colony PCR



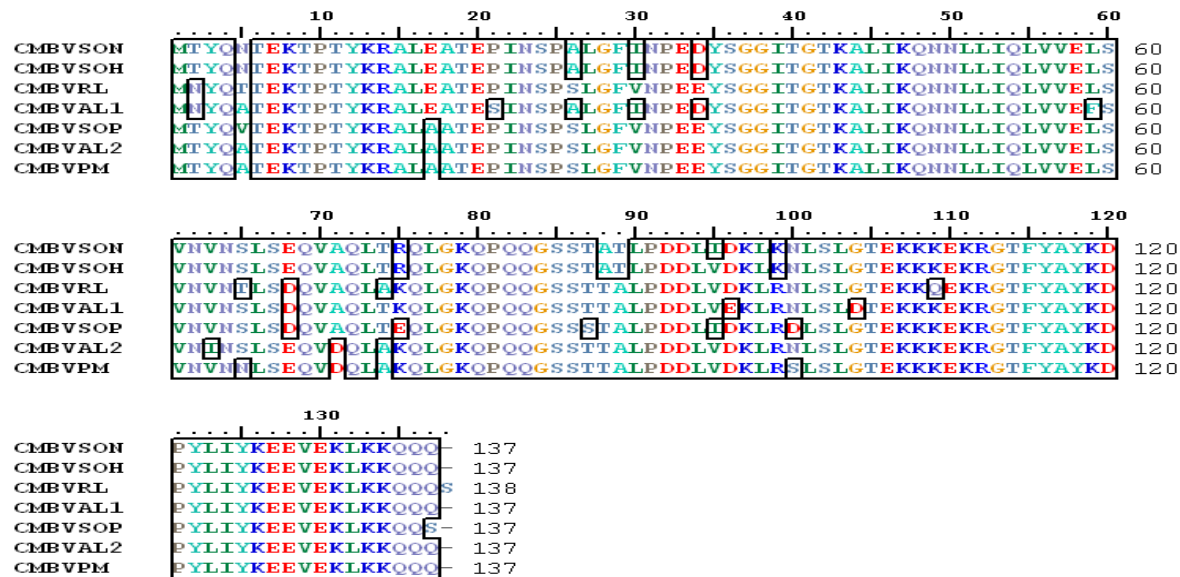
**Gel electrophoresis of PCR amplified products
from colonies of selected recombinants clones**

Multiple alignment of all the six ORFs (Amino acids sequence) of the genome of CMBVSON with other CMBV isolates.

ORF1



ORF2



Multiple alignment of all the six ORFs (Amino acids sequence) of the genome of CMBVSON with other CMBV isolates.

ORF3

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 10 to 60. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 70 to 120. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 130 to 180. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 190 to 240. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 250 to 300. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 310 to 360. Conserved regions are highlighted with boxes.

370 380 390 400 410 420

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 430 to 480. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 490 to 540. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 550 to 600. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 610 to 660. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 670 to 720. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 730 to 780. Conserved regions are highlighted with boxes.

Multiple alignment of ORF3 amino acid sequences for CMBVSON, CYMVS OH, CYMVAL1, CMBVAL2, CMBVSOP, CYMVRL, and CMBVPM. The alignment covers positions 790 to 840. Conserved regions are highlighted with boxes.

850 860 870 880 890 900

Multiple alignment of all the six ORFs (Amino acids sequence) of the genome of CMBVSON with other CMBV isolates.

ORF3

```

CMBVSON  NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  778
CMBVSOH  NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  778
CMBVAL1  NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  763
CMBVAL2  NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  770
CMBVSOP  NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  762
CMBVRL   NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  749
CMBVPM   NMLGPTTEKIIWQTWRMGYADEYENLVTTADGREGTQNILSQMRRVFSLEDPATGSTAVQD  770
    
```

```

          790      800      810      820      830      840
CMBVSON  EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  815
CMBVSOH  EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  815
CMBVAL1  EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  800
CMBVAL2  EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  807
CMBVSOP  EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  822
CMBVRL   EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  786
CMBVPM   EAYRDLERLTCDSV.....KHIVQYLNDFMRIAATKGRMFIG  807
    
```

```

          850      860      870      880      890      900
CMBVSON  PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  875
CMBVSOH  PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  875
CMBVAL1  PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  860
CMBVAL2  PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  867
CMBVSOP  PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  882
CMBVRL   PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  846
CMBVPM   PELS EKLWLMKPGDLGQRMKKAYEEKHPGNIIVGVCPRIILFAYKYLEGEGKDAAFRRSLKN  867
    
```

```

          910      920      930      940      950      960
CMBVSON  LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  935
CMBVSOH  LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  935
CMBVAL1  LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  920
CMBVAL2  LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  927
CMBVSOP  LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  942
CMBVRL   LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  906
CMBVPM   LSFCSSSIPGYYGGRSGEKRYGVRRTTTYKPKPHNHARIEKTKHLRNKKCKCYLGGEE  927
    
```

```

          970      980      990      1000     1010     1020
CMBVSON  SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  995
CMBVSOH  SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  995
CMBVAL1  SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  980
CMBVAL2  SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  987
CMBVSOP  SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  1002
CMBVRL   SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  966
CMBVPM   SHFARECPNDRRNKRVAMFEGLDLDPDCEIVS IDEGDDSDAIFSI SEGEETGAE EEOC  987
    
```

```

          1030     1040     1050     1060     1070     1080
CMBVSON  FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1053
CMBVSOH  FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1053
CMBVAL1  FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1038
CMBVAL2  FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1045
CMBVSOP  FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1060
CMBVRL   FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1024
CMBVPM   FVFQEECHGTYYLWGRGGYQDLVQIT...SKEIYYCOHEWEENQPIDDP...AHVRCYCPKRETT  1047
    
```

```

          1090     1100     1110     1120     1130     1140
CMBVSON  QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1113
    
```

```

CMBVSON  QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1113
CMBVAL1  QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1098
CMBVAL2  QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1105
CMBVSOP  QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1116
CMBVRL   QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1084
CMBVPM   QRARLHCKLCHITSCMLCGPPTYFNKKITVQIPQAPFNQKGLLQQQQQEYIAWCNNEIARI  1107
    
```

```

          1150     1160     1170     1180     1190     1200
CMBVSON  KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1173
CMBVSOH  KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1173
CMBVAL1  KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1158
CMBVAL2  KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1165
CMBVSOP  KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1176
CMBVRL   KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1144
CMBVPM   KEEVAFYKQLAQRRELQQLQLESRKELAGADSRRRKDKGVVIDEGSCYFNPETTRIVAF  1167
    
```

```

          1210     1220     1230     1240     1250     1260
CMBVSON  SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1233
CMBVSOH  SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1233
CMBVAL1  SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1218
CMBVAL2  SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1225
CMBVSOP  SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1236
CMBVRL   SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1204
CMBVPM   SDTQVTKTRPVKIMLYMDVVRMEIPGIPAFIVKAILDTGATTCCIDSRSPKDALEENSF  1227
    
```

```

          1270     1280     1290     1300     1310     1320
CMBVSON  VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1293
CMBVSOH  VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1293
CMBVAL1  VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1278
CMBVAL2  VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1285
CMBVSOP  VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1296
CMBVRL   VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1264
CMBVPM   VVNFSGINSKQVQKIKAGKMFINEHYFRIPYCYSEFMQIGDGIQLLGCNFRISMVGG  1287
    
```

```

          1330     1340     1350     1360     1370     1380
CMBVSON  VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1353
CMBVSOH  VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1353
CMBVAL1  VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1338
CMBVAL2  VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1345
CMBVSOP  VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1356
CMBVRL   VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1323
CMBVPM   VRLGHTITFYKQITSIINRLAAPLLKQEBEEKEBELHLEHRLIQEMVAYSTERPFVHF  1347
    
```

```

          1390     1400     1410     1420     1430     1440
CMBVSON  DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1413
CMBVSOH  DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1413
CMBVAL1  DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1398
CMBVAL2  DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1405
CMBVSOP  DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1416
CMBVRL   DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1383
CMBVPM   DQKFAGLIQELKAGQYIGEEPMKYWAKQVVCVCHLDIKNPDVIEDRPLKHVTPQMEESFR  1407
    
```

```

          1450     1460     1470     1480     1490     1500
CMBVSON  KHVEALLKIGAIRPKSRHRRTTALIVNSGTSIDPITGKEVKGRMVFVNYKRLNDLTKDK  1473
CMBVSOH  KHVEALLKIGAIRPKSRHRRTTALIVNSGTSIDPITGKEVKGRMVFVNYKRLNDLTKDK  1473
    
```

Multiple alignment of all the six ORFs (Amino acids sequence) of the genome of CMBVSON with other CMBV isolates.

CYMVAL1 KHVEALLKIGAIRP SKSRHRTTALIVNSGTSIDPITGKEVKGKERMFVFNKRLNDLTNKKD 1458
 CMBVAL2 KHVEALLKIGAIRP SKSRHRTTALIVNSGTSIDPITGKEVKGKERMFVFNKRLNDLTNKKD 1465
 CMBVSOP KHVEALLKIGAIRP SKSRHRTTALIVNSGTSIDPITGKEVKGKERMFVFNKRLNDLTNKKD 1476
 CYMVR1 KHVEALLKIGAIRP SKSRHRTTALIVNSGTSIDPITGKEVKGKERMFVFNKRLNDLTNKKD 1443
 CMBVPM KHVEALLKIGAIRP SKSRHRTTALIVNSGTSIDPITGKEVKGKERMFVFNKRLNDLTNKKD 1467

1510 1520 1530 1540 1550 1560
 CMBVSON QYSLPGIQTILQRLKGSTIFSKFDLKS GFHQVAMHPDSIEWTAFWVPSGLYEWLVMPPFGL 1533
 CYMVSOP QYSLPGIQTILQRLKGSTIFSKFDLKS GFHQVAMHPDSIEWTAFWVPSGLYEWLVMPPFGL 1518
 CMBVAL2 QYSLPGIQTILQRLKGSTIFSKFDLKS GFHQVAMHPDSIEWTAFWVPSGLYEWLVMPPFGL 1525
 CMBVSOP QYSLPGIQTILQRLKGSTIFSKFDLKS GFHQVAMHPDSIEWTAFWVPSGLYEWLVMPPFGL 1536
 CYMVR1 QYSLPGIQTILQRLKGSTIFSKFDLKS GFHQVAMHPDSIEWTAFWVPSGLYEWLVMPPFGL 1503
 CMBVPM QYSLPGIQTILQRLKGSTIFSKFDLKS GFHQVAMHPDSIEWTAFWVPSGLYEWLVMPPFGL 1527

1570 1580 1590 1600 1610 1620
 CMBVSON KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1593
 CYMVSOP KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1593
 CYMVAL1 KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1578
 CMBVAL2 KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1585
 CMBVSOP KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1596
 CYMVR1 KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1563
 CMBVPM KNAPAVFQRKMDHCFKGTAEAFIAVYIDDLVFSKTEDEHEEHLQIMLSICQKNGLILSPT 1587

1630 1640 1650 1660 1670 1680
 CMBVSON QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1651
 CYMVSOP QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1651
 CYMVAL1 QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1636
 CMBVAL2 QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1644
 CMBVSOP QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1654
 CYMVR1 QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1621
 CMBVPM QMKIAQAEIEFLGAIHKGGLIKLQPHIVQKLLFTFNKQLEEVKGLRSWLGILNYARSY 1647

1690 1700 1710 1720 1730 1740
 CMBVSON IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1711
 CYMVSOP IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1711
 CYMVAL1 IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1696
 CMBVAL2 IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1704
 CMBVSOP IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1714
 CYMVR1 IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1681
 CMBVPM IPHMGRLLSPLYAKVSPTGERRMNRQDWALIDKIRAQVQNLPALELPPADCFIIIEETDGC 1707

1750 1760 1770 1780 1790 1800
 CMBVSON MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1771
 CYMVSOP MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1771
 CYMVAL1 MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1756
 CMBVAL2 MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1764
 CMBVSOP MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1774
 CYMVR1 MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1741
 CMBVPM MDGWWGVCKWKLQYDPRSSSERVCAYASGKFNPPKSTIDAEIHAVMNSLHNFKIYYLDK 1767

1810 1820 1830 1840 1850 1860
 CMBVSON SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1774
 CYMVSOP SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1831
 CYMVAL1 SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1816

ORF3

CMBVAL2 SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1824
 CMBVSOP SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1834
 CYMVR1 SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1801
 CMBVPM SLCLRITDCQAILISFFNKS NVNKP SRVRWIAFTDFLTGLGIPVNIHIDGKNNHLADALS 1827

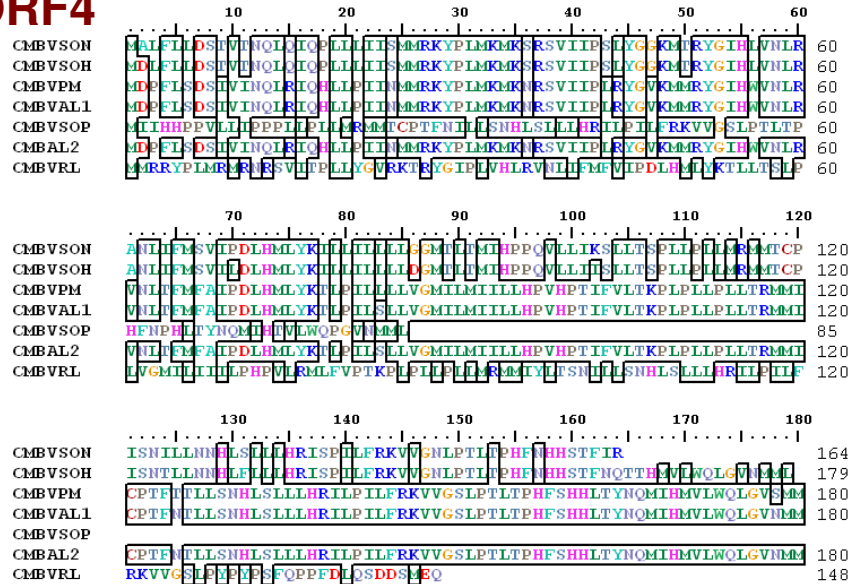
1870 1880 1890 1900 1910 1920
 CMBVSON LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1774
 CYMVSOP LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1891
 CYMVAL1 LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1876
 CMBVAL2 LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1884
 CMBVSOP LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1894
 CYMVR1 LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1861
 CMBVPM LVTGFFVFAEPQCQDKFDLGLKLEAALQEKKEAPQAMHEEYVSLIIRSADRTITLSLCSMR 1887

1930 1940 1950 1960 1970 1980
 CMBVSON DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1774
 CYMVSOP DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1951
 CYMVAL1 DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1936
 CMBVAL2 DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1944
 CMBVSOP DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1954
 CYMVR1 DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1921
 CMBVPM DSSSRITYS CRPGEKPEKALICEQKSCQSGDLGSAKTIVHSRSASHNQQNWWPSTSTNSL 1947

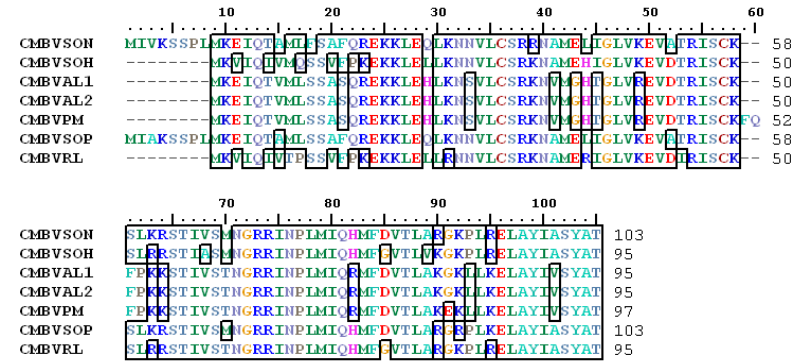
1990 2000 2010
 CMBVSON TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1774
 CYMVSOP TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1983
 CYMVAL1 TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1968
 CMBVAL2 TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1976
 CMBVSOP TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1986
 CYMVR1 TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1953
 CMBVPM TSEAKRQETTHMPINCPHAIGTTNSCVKWSSE 1979

Multiple alignment of all the six ORFs (Amino acids sequence) of the genome of CMBVSON with other CMBV isolates.

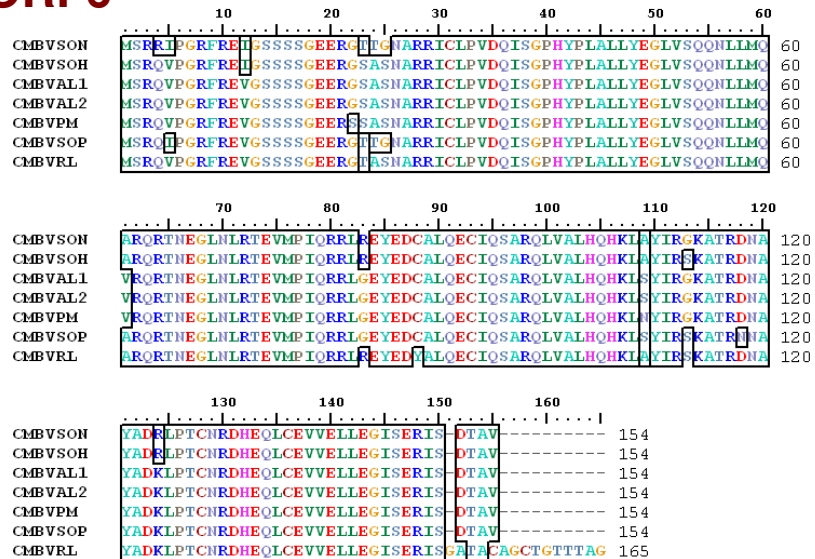
ORF4



ORF5



ORF6



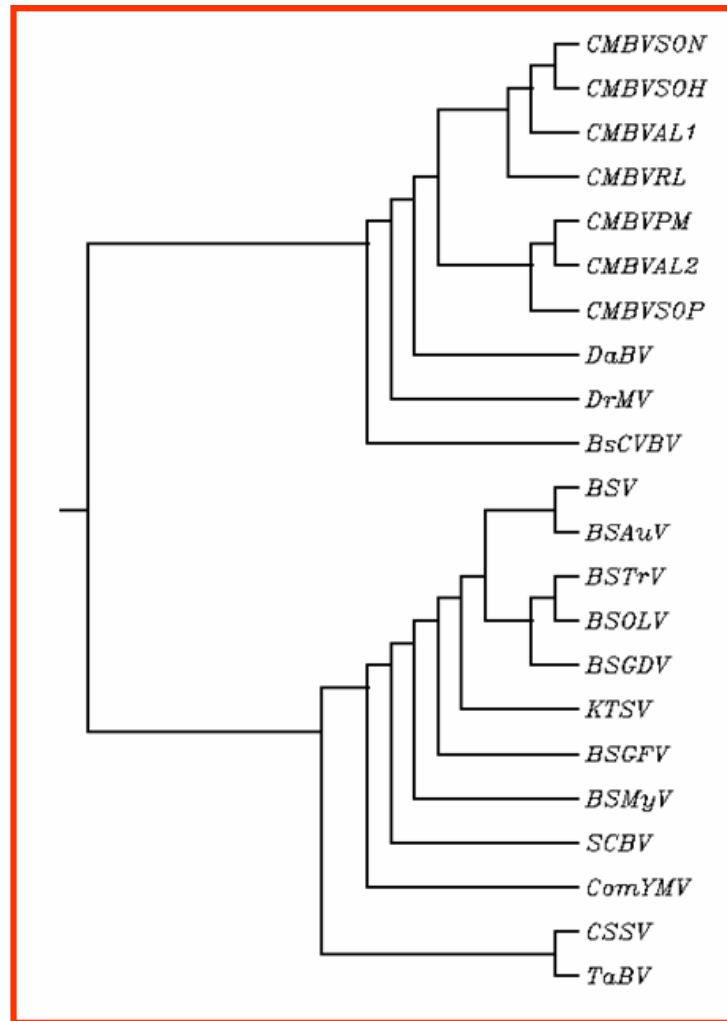
Different ORFs, their location and number of amino acids coded by each ORFs in CMBVSON genome

Name of ORF	Starting nucleotide	Stop nucleotide	Length (bp)	No. of Amino Acids	Molecular weight (kDa)
1	231	662	432	143	14.3
2	659	1072	414	137	13.7
3	1069	6393	5325	1774	17.74
4	1973	2467	495	164	16.4
5	3953	4264	312	103	10.3
6	6593	7057	465	155	15.3

Comparison of full length genome nucleotide sequence of CMBVSON with other badnaviruses

Isolates	CMBV	CMBV	CMBVS	CMEV	CMBV	CMBV	CMBV	BS				BSMy V	Com								
	SON	SOH	OP	AL1	AL2	RL	PM	BSV	TRYV	BSAuV	BSGDV		BSGFV	BSOLV	CSVV	KTSV	DrMV	YMV	DaBV	SCEV	TaBV
CMBVSON	100																				
CMBVSOH	96.0	100																			
CMBVSOP	90.1	88.1	100																		
CMEVAL1	91.7	93.0	88.1	100																	
CMEVAL2	87.7	87.6	90.2	88.0	100																
CMEVRL	89.1	88.9	89.9	91.6	88.2	100															
CMEVPM	87.5	87.4	90.2	87.8	98.9	88.6	100														
BSV	36.7	36.5	36.2	36.9	36.5	36.9	36.9	100													
BSTV	36.2	36.4	36.3	36.4	36.1	36.3	36.2	53.2	100												
BSAuV	36.7	36.9	36.7	36.9	36.5	36.9	36.5	100	53.3	100											
BSGDV	36.2	36.4	36.3	36.4	36.2	36.3	36.2	53.2	99.7	53.2	100										
BSGFV	36.8	37.0	36.6	37.0	36.4	36.9	36.5	48.6	47.5	48.6	47.4	100									
BSMyV	38.3	38.4	38.3	38.2	38.0	38.6	38.2	50.0	48.4	50.0	48.4	48.8	100								
BSOLV	37.7	37.9	37.7	37.9	37.6	37.8	37.7	55.0	93.1	55.0	93.0	43.0	49.7	100							
CSSV	48.4	48.6	48.9	49.1	48.9	48.8	49.0	37.4	36.8	37.4	36.8	38.0	38.1	38.3	100						
KTSV	36.9	37.0	36.7	36.8	36.4	36.7	36.4	53.4	53.2	53.4	53.2	47.8	47.8	55.3	37.3	10.0					
DrMV	42.2	42.1	41.7	42.3	41.4	42.5	41.7	38.7	37.0	38.7	36.9	39.0	39.3	38.6	38.9	38.3	100				
ComYMV	37.3	37.3	37.5	37.2	37.7	37.2	37.7	41.6	41.6	41.6	41.5	40.7	41.9	42.2	38.6	40.7	36.3	100			
DaBV	48.2	48.4	47.9	48.2	48.3	48.2	48.2	36.3	35.6	36.3	35.3	36.7	37.2	36.2	45.4	36.6	39.3	39.6	100		
SCEV	35.2	34.9	35.3	35.2	35.1	35.2	35.0	45.2	43.9	45.2	43.9	51.0	45.8	45.2	36.3	47.3	37.0	40.3	34.7	100	
TaBV	38.0	38.1	38.1	38.1	37.9	38.1	37.9	37.3	36.6	37.3	36.6	38.3	38.3	38.2	39.1	38.0	41.0	34.9	36.1	36.7	100

- 96% similarity with CMBVSOH
- 48.4 % similarity with CSSV



Phylogenetic tree of full length genomes CMBVSON other badnaviruses

Percent identity of nucleotide sequence and ORFs encoded by them in different of CMBV isolates and their comparison with other badnaviruses

Viruses	Nucleotide Sequence Identity %	ORFI		ORFII		ORFIII		ORFIV		ORFV		ORFVI	
		Size (aa)	Identity %	Size (aa)	Identity %	Size (aa)	Identity %	Size (aa)	Identity %	Size (aa)	Identity %	Size (aa)	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	1744	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 – 96.9	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 – 96.9	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
Other badnaviruses													
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	-	-
B SV	36.5 – 36.9	176	19.8 – 20.4	132	19.7 – 21.7	1900	30.7 – 33.0	-	-	-	-	-	-
B SAuV	36.5 – 36.9	176	19.8 – 20.4	132	19.7 – 21.7	1900	30.7 – 33.0	-	-	-	-	-	-
B STrV	36.1 – 36.4	176	22.0 – 22.5	134	19.4 – 20.8	1709	33.8 – 34.5	-	-	-	-	-	-
B SGDV	36.2 – 36.4	177	22.0 – 22.5	134	19.4 – 20.8	1709	33.8 – 34.5	-	-	-	-	-	-
B SGFV	36.4 – 37.0	177	19.3 – 19.8	134	17.4 – 18.1	1832	32.6 – 34.7	-	-	-	-	-	-
B SMyV	38.2 – 38.6	176	19.8 – 20.4	132	19.7 – 21.7	1900	30.7 – 33.0	-	-	-	-	-	-
B SOLV	37.6 – 37.9	175	22.0 – 22.5	112	19.3 – 20.8	1832	31.8 – 34.0	-	-	-	-	-	-
B sCVBV	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	-	-	-	-	-	-

Conclusions

1. A badnavirus designated as CMBVSON was confirmed in sathgudi sweet orange showing mosaic symptoms by electron microscopy and PCR.
2. The complete genome of CMBVSON consisted of 7558 nucleotide containing for 6 open reading frames (ORFs). High variability (87.4% to 96%) observed in ORF 4 and 5 infecting sathgudi sweet orange, rangpur lime acid lime and pummelo.
3. The intergenic region (non coding region) has 731 nucleotide. Intergenic region consists of putative promoter elements like TATATAA box, CACAAT sequence, TGACG sequence and AATAAA the polyadenylation signal sequence.
4. The molecular weight of full genome of CMBV sathgudi sweet orange isolate is 247.60 kDa with G+C Content 43.52% and A+T content 56.48%. The (A+T) and (G+C)ratio was 1.29.
5. The genome consists, Movement protein ,Aspartic Protease Reverse transcriptase and RNase H in ORF3.
6. Sequence information will also be useful in designing primers for detection of different isolates of CMBV in PCR and production disease free planting materials of citrus.

Thank You

Base composition of CMBVSON and its comparison with other CMBV isolates of citrus

CMBV Isolates	Base Composition				A+T (%)	G+C (%)	Molecular Weight in KDa	Total Nucleotides (bp)
	A	T	G	C				
CMBVSON	2326	1943	1668	1621	56.48	43.52	4586.7	7558
CMBVSOH	2337	1925	1666	1631	56.38	43.62	4586.9	7559
CMBVSOP	2322	1922	1666	1587	56.61	43.39	4549.0	7497
CMBVAL1	2312	1894	1675	1617	55.21	44.79	4549.0	7498
CMBVRL	2321	1620	1671	1910	56.25	43.75	4562.1	7522
CMBVAL2	2323	1877	1653	1620	56.19	43.79	4539.0	7473
CMBVPM	2323	1889	1648	1616	56.34	43.66	4541.0	7487

Comparison of ORFs of CMBVSON with CMBVSOH

