View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by ICRISAT Open Access Repository

Arch Virol (1998) 143: 769–780

Archives of Virology
© Springer-Verlag 1998
Printed in Austria

Evidence that whitefly-transmitted cowpea mild mottle virus belongs to the genus *Carlavirus*

R. A. Naidu[1], **S. Gowda**[2], **T. Satyanarayana**[2], **V. Boyko**[2,*], **A. S. Reddy**[1], **W. O. Dawson**[2], and **D. V. R. Reddy**[1]

[1]Crop Protection Division, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India
[2]Citrus Research and Education Center (CREC), University of Florida, Lake Alfred, Florida, U.S.A.

Accepted October 7,1997

Summary. Two strains of whitefly-transmitted cowpea mild mottle virus (CPMMV) causing severe (CPMMV-S) and mild (CPMMV-M) disease symptoms in peanuts were collected from two distinct agro-ecological zones in India. The host-range of these strains was restricted to Leguminosae and Chenopodiaceae, and each could be distinguished on the basis of symptoms incited in different hosts. The 3′-terminal 2500 nucleotide sequence of the genomic RNA of both the strains was 70% identical and contains five open reading frames (ORFs). The first three (P25, P12 and P7) overlap to form a triple gene block of proteins, P32 encodes the coat protein, followed by P12 protein located at the 3′ end of the genome. Genome organization and pair-wise comparisons of amino acid sequences of proteins encoded by these ORFs with corresponding proteins of known carlaviruses and potexviruses suggest that CPMMV-S and CPMMV-M are closely related to viruses in the genus *Carlavirus*. Based on the data, it is concluded that CPMMV is a distinct species in the genus *Carlavirus*.

Introduction

Cowpea mild mottle virus (CPMMV) was first reported on cowpea in Ghana [3]. Subsequently it was reported from several tropical regions of Africa and Asia in a diverse range of plant species that include leguminous and solanaceous food crops (for a review see [17]). CPMMV is transmitted by the whitefly, *Bemisia tabaci*, in a non-persistent manner [20] and through seed in cowpea, soybean and French

*Present address: Institute of Medical Microbiology, University of Basel, Basel, Switzerland.

bean [3]. Naturally infected cowpea plants exhibit either a mild systemic mottle or are symptomless. Differences in symptom severity in cowpea plants make economic assessment of the disease difficult, but it is nonetheless considered a definite threat to the successful production of such legumenous crops as soybeans, peanuts, beans and mungbeans, all of which are intercropped with cowpeas in many tropical countries in Africa and Asia [6, 16, 23].

CPMMV has filamentous particles c. 650×15 nm in size with a coat protein of 31–33 kDa and a single-stranded RNA of Mr 2.5×10^6 [1]. Based on its physico-chemical properties, CPMMV was included as a tentative species in the genus *Carlavirus* [2]. However, CPMMV shows no serological relationship with any of the recognised species of the genus *Carlavirus*. Unlike virus particles of other carlaviruses, CPMMV particles occur in characteristic brush-like inclusions in infected plants [4]. Also, CPMMV is transmitted by whiteflies (*Bemisia tabaci*) whereas different species of the genus *Carlavirus* are transmitted by aphids [7].

CPMMV infects groundnuts (= peanuts) grown in different agro-ecological zones of India [15] and was shown to be serologically related to the CPMMV originally described by Brunt and Kenten [3]. Disease surveys and preliminary studies carried out under similar conditions have shown that a CPMMV isolate collected from peanuts grown in Mianpuri in Northern India incited mild mosaic mottling symptoms whereas the isolate collected by Iizuka et al. [15] from Hyderabad in Southern India cause much more severe disease symptoms. Therefore, the isolate described by Iizuka et al. [15] was designated as CPMMV-S and the one collected from Mainpuri as CPMMV-M. Both are transmitted by *B. tabaci* in a non-persistent manner. However, it was not clear if these isolates are strains of one virus or distinct but serologically related viruses. Therefore, we examined the genome organization of CPMMV-S and -M isolates in order to determine whether they fit in to the genus *Carlavirus* and to gain information and materials that would be useful for their detection and discrimination. In this paper we provide evidence, based on partial nucleotide sequence analysis and comparison of the deduced amino acid sequences with other carlaviruses, that the whitefly-transmitted CPMMV is a distinct species in the genus *Carlavirus*.

Materials and methods

Virus isolates and host-range studies

The CPMMV-S and -M isolates were maintained in groundnut (var. JL 24) initially by *B. tabaci* and subsequently by mechanical inoculations as described earlier [15, 20]. For host-range studies, ten plants of each indicator host plant species were sap-inoculated and maintained in glasshouse (25–30 °C). Both mechanically inoculated and subsequently developing leaves of these plants were assessed for symptoms and tested for the presence of virus by direct antigen coating form of enzyme-linked immunosorbent assay (DAC-ELISA) [14] with a polyclonal antiserum produced against a CPMMV isolate [15].

Virus purification

Systemically infected groundnut leaves were homogenized in cold 10 mM phosphate buffer, pH 8.0 containing 10 mM DIECA and 2% thioglycerol (1:4,w/v). The extract was filtered

through two layers of muslin, emulsified with chloroform (10%, v/v) and clarified at 5000 rpm for 10 min. The supernatant was subjected to centrifugation at 30,000 rpm for 60 min in a Beckman R45 rotor. The resulting pellet was suspended in 100 mM borate-phosphate buffer, pH 8.3 containing 200 mM urea (BPU) and clarified at 5000 rpm for 10 min. The supernatant was treated with Triton × 100 (0.1% final volume) and the virus was pelleted through 12 ml of 30% sucrose cushion (prepared in BPU) at 24,000 rpm for 90 min in a Beckman SW28 rotor. The pellet was suspended in BPU and subjected to 10–40% caesium chloride density gradient centrifugation for 1h at 30,000 rpm in a Beckman SW50 rotor. The opalescent virus band was collected and dialyzed overnight at 4 °C against BPU.The virus was pelleted at 30,000 rpm for 60 min in a Beckman R45 rotor and suspended in 100 mM phosphate buffer, pH 7.2.

cDNA synthesis and cloning

Viral RNA was extracted from SDS-disrupted purified virions with phenol/chloroform and concentrated by ethanol precipitation. Integrity and size estimations of viral RNA were done by electrophoresis in denaturing agarose gels [25]. Viral RNA was denatured at 65 °C for 5 min, annealed with random composition hexa-deoxyoligonucleotides and used as a template for cDNA synthesis [11]. The double-stranded (ds) cDNA was blunt-ended with T4 DNA polymerase and ligated into *Sma* I-cut pUC119. Clones containing virus-specific sequences were identified by Southern hybridization using ^{32}P labelled randomly primed first-strand cDNA of viral RNA as a probe.

The 3′ extremities of the viral RNA were cloned by RACE-PCR utilizing a commercial kit (Gibco BRL). First strand cDNA was made using the primer: 5′-GGCCACGCGTCGACTA GTAC$(T)_{17}$-3′ (supplied with the kit). PCR was done using the primer: 5′-CTACTACTACTA-GGCCACGCGTCGACTAGTAC-3 (supplied with the kit) and a second primer: 5′-ACTGGT-GGCACTCAAGG-3′ which was colinear to nucleotides 2 029–2 045 in CPMMV-S and nucleotides 2 030–2 046 in CPMMV-M.

Expression of the CPMMV-M CP gene in E. coli

The P32 ORF of CPMMV-M was amplified by PCR from a cDNA clone using an upstream primer: 5′-aggaccATGGAGTCTATCTTTG, which corresponds to the first 16 nt of the P32 ORF with extra nucleotides (lower case) added to create a *Nco I* site (underlined), and a down stream primer: 5′-gaagctcgagTCACTTCTTGGC GTGGTTG, which is complementary to the last 19 nucleotides of the P32 ORF with extra nucleotides (shown in lower case) added to create a *Xho I* site (underlined). The *Nco I/Xho I* fragment was sub-cloned into similarly digested protein expression plasmid vector pET 15b (Novagen). Protein expression, SDS-polyacrylamide gel electrophoresis (PAGE) and immunoblotting were performed as described by Ohshima et al. [21].

Sequence analysis

The double-stranded DNA was sequenced in both orientations by Taq cycle sequencing using an automated DNA sequencer (Perkin Elmer/Applied Biosystems). Sequences were assembled and analysed using programs in the University of Wisconsin GCG package [5]. The GAP, and CLUSTAL V [12] were used to align and compare nucleotide and deduced amino acid sequence identities (and similarities). Sequences of the following viruses belonging to the genus *Carlavirus* were used for comparisons (genomic sequence accession numbers given in parenthesis): potato virus M (PVM, X53062, D00515, D14449), carnation latent virus (CLV, X55331), Helenium virus S (HVS, D10454, D01119), lily symptomless virus (LSV, X15343), potato virus S (PVS-An, D00461), chrysanthemum virus B (CVB, S60150), poplar mosaic virus (PopMV, X65102), blueberry scorch virus (BlSV, L25658).

Results

Properties of CPMMV isolates

A comparison of host reactions of CPMMV-S and CPMMV-M are shown in Table 1. Each of the isolates infected a limited number of host plants belonging to Leguminosae and Chenopodiaceae and differed in the type of symptoms they induced in these plants. *Chenopodium amaranticolor* was a good differential host; CPMMV-S produced necrotic local lesions whereas CPMMV-M produced chlorotic spots.

Nucleotide sequence and genome organization

Initially, cDNA clones derived from random primed cDNA synthesis were utilized for dideoxy nucleotide sequencing. Based on the sequence information generated, oligonucleotide primers were designed for use in RT-PCR amplification of regions of CPMMV-S and CPMMV-M RNA that were not covered by the cDNA clones and for 3′RACE reactions. The nucleotide sequences at the 3′ end of the genomic RNAs of CPMMV-S (2511 nt) and CPMMV-M (2504) were deposited in the Gene Bank with accession numbers AF024628 and AF024629 respectively. CPMMV-S showed 70% sequence identity with CPMMV-M at the nucleotide level. Analysis of the coding capacity of each nucleotide sequence revealed five open reading frames (ORF) encoding possible proteins of Mr ($\times 10^{-3}$) 25.8 (P25), 11.7 (P12),

Table 1. Symptoms incited by CPMMV-S and CPMMV-M on different host plants

Symptoms[a]	CPMMV-S		CPMMV-M	
	local reaction	systemic reaction	local reaction	systemic reaction
Chenopodiaceae				
Chenopodium amaranticolor	Nll	–	CS	–
C. quinoa	CS	–	CS	–
Leguminosae				
Arachis hypogaea	–	Vc, Cr, Smmt, Pk	–	Vc, Mmmt
Glycine max	Nll, Vn	M, Cr, Mt	Nll, Vn	M, Cr, Mt
Vigna mungo	–	–	–	M
V. radiata	–	–	–	M
V. unguiculata	–	M	–	–
Pisum sativum	–	M	–	+
Macrotyloma uniflorum	–	–	–	N
Phaseolus vulgaris cv. Top crop	–	–	–	+

[a] *Cr* crinkle; *Cs* chlorotic spots; *M* mosaic; *Mmmt* mild mosaic mottling; *Mt* mottling; *N* necrosis; *Nll* necrotic local lesions; *Smmt* severe mosaic mottling; *Vc* veinal chlorosis; *Vn* veinal necrosis; + symptomless infection; − no symptoms and negative in ELISA tests

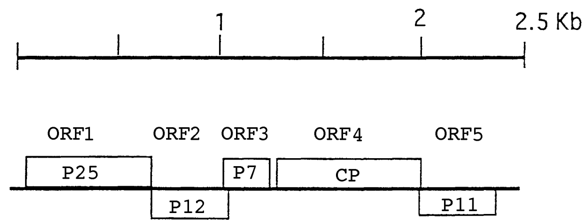

Fig. 1. Diagramatic representation of the genome organization, location of major ORFs and their putative protein products in the 3′-terminal region of CPMMV-S RNA. The genome organization of CPMMV-M (accession number AF024629) is similar to CPMMV-S (accession number AF024628)

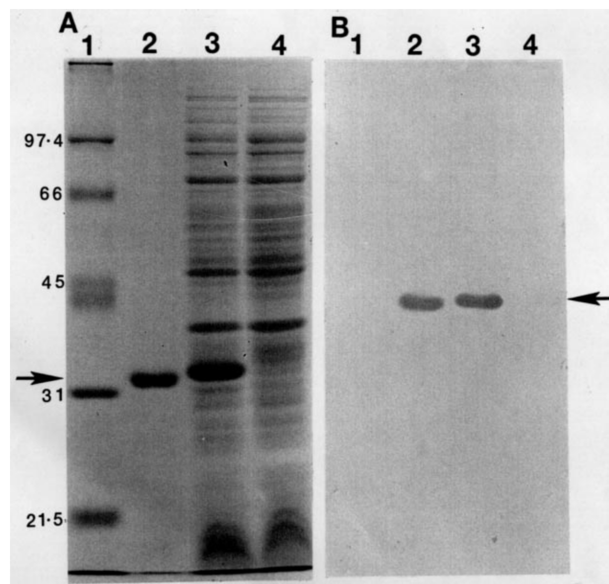

Fig. 2. SDS-PAGE of CPMMV-M coat protein expressed in *E. coli*. Following electophoresis, 10% SDS-polyacrylamide gels were either stained with Coomassie Brilliant Blue R-250 (**A**) or used for immunoblotting (**B**). *1* Protein MW markers (Bio-Rad) with sizes (in kDa) marked in the left margin; *2* coat protein from purified CPMMV-M virions; *3* total proteins from *E. coli* transformed with pET-15b containing P32 ORF; *4* total proteins from *E. coli* transformed with pET-15b alone. Arrow indicates position of the coat protein

7.7 (P7), 32.3 (P32) and 11.5 (P11) in CPMMV-S, and 25.7 (P25), 11.6 (P12), 7.7 (P7), 32.2 (P32) and 11.6 (P11) in CPMMV-M. The organization of these ORFs was similar for both CPMMV-S and CPMMV-M and is shown in Fig.1. In both the sequences, the P25, P12 and P7 form a characteristic "triple gene block" (TGB) of proteins similar to the TGB proteins in the genera *Carlavirus, Potexvirus, Hordeivirus* and *Furovirus*. The P7 substantially overlaps the P12 whereas the initiation codon of the P12 ORF is partially embedded in the termination codon of the P25 ORF. The P32 ORF was separated from P7 ORF by an intergenic region of 56 nt in CPMMV-S and 19 nt in the case of CPMMV-M. The initiation codon of the P11 in both the isolates was partially embedded in the termination codon of P32.

A.

```
CPMMV-S   MNELLSLLSDYNFKRLGHRLSDPIVINCVPGAGKTHLIRSLINSCDQFEAYTCSIPDPVNLSGRRIKHISEL--GESKKLVIIDEYQNLATVPNCAFAIF
CPMMV-M   MNELLAILSDYGFSRLGHRLSEPIVINCVPGAGKTHLINQLLEISDNFEAYTTSSPGQVNLNGKSIKHISEF--KVSNKLILIDEYQNLKEIPKDAFAIF
LSV       MDVLLSLLSEFGFERLSSELSLPIVVHSVPGGGKSSLIRKLINKDRRFSAYTFGLEDCESITGVRIKKAHAS--IPRSEFVVFDEYIE-GDTPPWAFAVF
PVS       MDVFLQVLNKYKFERVSSTLNKPIVVHSVPGAGKSSAIRELLKLDSRFECITRGRPDIPNLEGAFIK--AER--SGESKLLLVDEYIE-GPIPEDAFAIF
PVM       MDVIVDLLYKYKFERLSNKLVCPIVVHCVPGAGKSSLIRELLELDSRFCAYTAGVEDQPRLSGNWIRKWSGQ--QPEGKFVVLDEYTL-LTEVPPVFALF
BlSV      MDVLVKYLVKNNFERQSSNLTLPIVVHSVPGAGKSTLIRELITADPRFSALTFGKADPPNLRGAFIRSAFGA--TPYPEFTLIDEYLE-GEVLEGAFAVF
CVB       MDVFVKILSDFGFTRLSSSLHVPIVVHCMPGAGKSSCIRALLKADSRFVAYTLGVPDPCNLDGLCIKAFTGS--VDLKYFNILDEYARYNGDASDFFALF
PopMV     MDVLINKLASCNFSRTRNQVGKPCIINCVPGAGKSTLIRELLNSDSRFRAYTFGEADPKNLSGRRILPASELQNAPQGALIIIDEYTEGSWEPGKICAAF
          *. ..  *    * *    .  *  ... .**.**.  *  *.     *   *  .  .   . * *            . . ***          * *

CPMMV-S   GDPLQAHNCTLLPANFVCHKTQRFGKQTAELLRLLDFEISSELEDSVSIENIFEGEPEGQVICFEGEVINLLKAHNLDFATPCEVQGRTFDTVTFITSGD
CPMMV-M   GDPLQASKPLILPANFVCHTSRRFGKNTASLLSSLGFKVQSNLEDEVVKEDIFIGEPEGQIICFEVEVINLLKAHKLEFLTSCTSQGKSFKTVTFITTGL
LSV       ADPLQGGPGPVLRAHFIKRRSHRFGKCTAQLLNDLSYEVESDLADVVQIQGLYETDLQGTVVYYEACVGNLLRAHSVPAYYISEIRGQTFESVTFVTSEN
PVS       ADPLQSTAVSPHRAHFIKTLSHRFGKCTDSLLRDLGWDVQAEGQDSVQIADIFTVDPRETIVYFEPEVGELLRSHGVEASCIGEVRGATFEHVTFVTSEN
PVM       GDPIQSNTSAVQRADFVCSVSRRFGSATCGLLRELGWNVRSEKADLVQVSDIYTKDPLGKVVFSEEEVGCLLRSHGVEALSLQEITGQTFEVVTFVTSEN
BlSV      ADPLQGGPGIPLRAHFIKRISHRFGSQTSALLTSLGFEVEACGEDRVEIADIFKRDPVGVVLYHEKEIGVLLAAHSVEAYCIKEVRGQTFESVTFVTAES
CVB       GDPVQSPPRNLYRAHFKAVLSKRFGSCTAQLLRELGFEVESTKEDLVSIRGLYDFDPVGTVIYYEKEIGCLLRAHSIEAYEPEEVVGKTFETVTFVTAEN
PopMV     GDPIQLMGARHCLTDFVCNKTKRFGSSTCELLNGFGFEIHSEKEDICLVRDFFQVDAEGTVIAFESEVKDILARHFVEFEDICSIRGKTFEEVTFFTASN
          .**.*       ..*    . ***  *  **   .   .  .    *       ..  .     ..   *   .  .*  *  .           * .*  *** *.

CPMMV-S   -INGENRHLHLVCLTRHKHKLKILSPNSAYPRVL------
CPMMV-M   -ITAENRHEHLICLSRHISKLKILSPQAKYPSEE------
LSV       Y-PV-DRALAFQCLTRHRSSLLILSPNATYTAS-------
PVS       S-PLIDKASAFQCLTRHTKSLLILCPDATYTAA-------
PVM       S-PVINRAAAYQCMTRHRTALHILCPDATYTAA-------
BlSV      A-PVLDRALSFQCLTRHRKTLL-LCPNATYTAA-------
CVB       HIPAESRHLVYQCLTRHRSVLHLMTPDASYTST-------
PopMV     SIPEHLAGRLLPVLNPAQKQAHNCLPRCHFRPLLVILRLS
                      .           *    .
```

B.

```
CPMMV-S   MPLTAPPDHSRTFTYIAIGASICVGIFALTRHTTPIAGDNIHRLPFGGCYRDGTKAILYNRPNNSRLQTGGF--SKFEILTGVLTLTALIVILSSRRNLN
CPMMV-M   MPLTAPPDYTRTFTYLVVGLSICIGLFTLTRHTTPIAGDNIHRLPFGGCYRDGTKSILYNRPGLNSQSFLNL--GKLEALLAVIVISGLIIFSSSRKARV
LSV       MPLTPPPDYTRVYTALAIGASIAFFTGLITRNTLPSVGDLQHNLPHGGRYRDGTKSVEYCGPRKLNSVESG-SRWTFQPWLLVIVLVALIIALGRQGH-N
PVS       MPLTPPPNYTGLYIAAALGVSLAAVVALFTRSTLPIVGDSQHNLPHGGRYRDGTKAIDYFKPTKLNSVEPG-NYWYTQPWLLVILLVALICLSGRHAQ-C
PVM       MPLTPPPDFTKVYLSAALGVSLALVVWLLIRSTLPVVGDRDHNLPHGGWYRDGTKSVFYNSPGRLNSIEARKAPLLGQPWAIVVLLVLLIWASHKLGRPN
BlSV      MPLTPPPNYTQAITAACIGISLALLVGLLTRSTLPTVGDLQHNLPHGGRYRDGTKCIDYRGPAKLNSVEGH-GTWGTQPWLLVIVLVGAIILLSKRGT-R
CVB       MPLTPPPDHTKVLLVAAIGLSIVASILTYSRNTLPQVGDHSHLLPHGGVYKDGTKTIVYGGPRKLNSLEGG-FNLPVQPWFLVILLSAAIFLLSCRSG-H
PopMV     MPLSPPPSHSKTFLVAAAGLSLVLCLYTLTRSTLPGVGDNIHSLPHGGQYRDGTKSINYCSPGKSYPSSNLLRGGNFSALCAILLISGAILISYRFQPGA
          ***..**  .           * *.           * * *  **  * ** ** ****** . *  *                           .. .    *

CPMMV-S   CHVCGRAH--
CPMMV-M   CHVCGREH--
LSV       CRACGRSH--
PVS       CPRCNRVHSA
PVM       CRACAGSHT-
BlSV      TCQCGTTH--
CVB       RRVCGQCH--
PopMV     LSRCGVTH--
             *   *
```

C.

```
CPMMV-S   MFVDAHIKLL-GCLLVLTFLVVLYYINSSL----NSYCVLLLTGESFTISNCPVNEELVALIGEQREAFNCRNF--
CPMMV-M   MFAVESTKLLVASLFLLGF-ILLYCVNNYL----SNNCVLLLTGESFTVSNCKINQDLVDLIREQRDYLSCHN---
ClV       MQASGLI--LVALFSAVVSYLALLHLSS-----SSSSCVVVVTGESFRISGCDFTEEFIGFAKTLRV-AN-S--QP
LSV       MRSVA----LTLCAIIAGYLLVSNLQNVF----SPEVCTLVITGESIRINGCNLSPAHFRAISHLKV-LQ---IHL
PVS       MLPKMQPSAQCLIVFSLAFVLGWYVLRPG-----NTSCVLLITGESVRLVNCELTKDLVEAVLLRPL----KHL--
PVM       M----IVYVLVGLSAF-CIVL-------YLISQGQSDCVVLITGESVRVQGCRIDGEFGSVLSKLKP-FGCGSFRS
BlSV      MWNNP----LVIAIVVCCFLITLFAFRSF----DSQSCVIVLTGESIRVVGCSLSPEHIVALSHLKV-LQ ---VDL
CVB       M----SLSYLDLLLAFGCVLAVSVIVNCFLVSH--NNCVIEITGEAVRISGCTFDRTFVELVKGLKP-A-----RH
PopMV     MWSDSLVSRICVPIIVVCTSIALLNVVSF---RSECSCVVHISGAAIDIRGCSFTPDFIEYAKTLRV-FN-HRYQE
          *                                    * .  ..* .  .  *
```

Fig. 3 (*continued*)

D.

```
CPMMV-S   ME-------SIF---DLNKLLDDSMGTKEKTGVPKAAD----------GTLLP----------DVDAE--LQKRLDSLREFLRKTQSASEITNPGFELGR
CPMMV-M   ME-------SIF---DLSNLFDSGMATKEKTGVPKAAD----------GTPLP----------DIDAE--LQKRLDSLREYLRKTQSASEITNPGFELGR
CLV       MP---------------EKRLEEDVGSQGQTGPNNPPQQHQGEQSAVSGMVSVNEVDLRNLQNRVEEAEKFMQRFQKLKEFNSQNLTAGELKNGGFESGR
LSV       MQ-------S---R--PAQESGS------ASETPARGRPTPSDAPRDEPT-----------NYNNNAESLLEQRLTRLIEKLNAEKHNSNLRNVAFEIGR
PVS       MP-------P---KPDPS----------SSGEAPQAMQPAPP--PRAEGHM--YAQPEGP---GQNEEAMLEQRLIRLIELMATKRHNSTLSNISFEIGR
PVM       MG-------DSTKKAETAKDEGTSQERREARPLPTAADFEGKDTSENTDGRAA----------DADGEMSLERRLDSLREFLRERRGAIRVTNPGLETGR
BlSV      MP-------P---KEAPAQSSEGNIVAKEAGEVPPRVAPAAPHVPPQAPPLQPIIQPAGAVAANLVPENQLEQRLMNLIEVLKPQRHNSSLKNVAFEIGR
HVS       MP-------PKVAPESSDAVSSQEQPQRPPPATPPVPTPPPGRREEVG-DRAEDPILQ---------------RLESLTALLRSERSAVRVTNASFETGR
CVB       MP-------PKPAPGDNEGNASGSTPTPPPPPPARTAEEARLRLAEMEREREQEQLLEEMNSNTPAEDARNISRLTQLAALLRREQTNVHVTNMALEIGR
PopMV     MSGEQTEQISKDKAVAAEQARKEQIAEGKKAAESPEVERRKKNIAEIAKLNEKAREAKKQATEQEETTTSLLERFNLLKEWHLNQQVNNKVKNPAMESET
          *                                                                                  *.  *         . * . *

CPMMV-S   -PELKQSTFNSDKHTHIYGKWSIDQLSRIVPKKISNNMATAEEMAKVQIILEGLGVPTEHVAEVLLQVAIYCKDTSSSAYMDSSGTFDWKGGSILSDSVI
CPMMV-M   -PELKQSTFNSDRDTHIYGKWSIDQLRRIVPKKISNNMATAEEMAKVQIILEGLGVPSEHVDEVLLQVAIYCKDVSTSAYMDSSGTFDWKGGSILSDSVI
CLV       -PPAKISEHLRGSTANVFTRPSLDALQMMDFKPESNMIVTAEELAAITAKLEGLGVPTERSAPICWAVARYCANTSTSPYTDPKGVFEFPGGAITRDAVF
LSV       -PSLEPTSAMRRNPANPYGRFSIDELFKMKVGVVSNNMATTEQMAKIASDIAGLGVPTEHVASVILQMVIMCACVSSSAFLDPEGSIEFENGAVPVDSIA
PVS       -PSLEPTPEMRRNPENPYSRFSIDELFKMEIRSVSNNMANTEQMAQITADIAGLGVPTEHVAGVILKVVIMCASVSSSVYLDPAGTVEFPTGAVPLDSII
PVM       -PRLQLAENMRPDPTNPYNRPSIEALSRIKPIAISNNMATSEDMMRIYVNLEGLGVPTEHVQQVVIQAVLFCKDASSSVFLDPRGSFEWPRGAITADAVL
BlSV      -PALEPVPTMRRNPANPYGRFSIDELYKMDVQVVSNNMATTEQMAKISSAIAGLGVPTEQVADVILKMVVMCASVSSSVYLDPDGSIEFDGGAVPVDSIA
HVS       -PALQPTADMRGDVTNMYNRPSTDSLWAVKPKPISNNMATSEDMVKIKVALEGLGVPTEHITGIIYQMCFYCASTSSSSYQDPKGVFEWPGGAIMVDDVM
CVB       -PALQPPPNMRGDPTNMYSQVSSDFLWKIKPQRISNNMATSEDMVKIQVALEGLGVPTESVKEVIIRLVLNCANTSSSVYQDPKGVIEWDGGAIIADDVV
PopMV     EPAL--ADELKPDMSNLFARPTVTDLQKMKWNAESNKMATADDMAFIEAEFQSLGVPKENLAKVMWTLTRYCVGASSSQYLDPKGEEEKLCGGVTRAALI
           *           . . . .      *  .       ** .   .... .   . .**** *     .       *    *.*  . *. *   .    *..     .

CPMMV-S   AALRKDENTLRRVCRLYAPIT-----WNFMLTHKAPPSDWAAMGFKYSDRYAAFDCFDYVENPAAIQPAEGLIRKPTPSEKIAHNTYKRLALDRSNRNEH
CPMMV-M   AALRKDENTLRRVCRLYAPVT-----WNFMLTHKAPPSDWAAMGFKYADRYAAFDCFDYVENPAAIQPAEGLIRKPTPSEKIAHNTYKRIALDRSNRNEH
CLV       AVIR-EVTTLRAFCRAFAPIT-----WNQMLFAKSPPENWQAKGYTYETRYAAFDVFDFVQNPAAIQPLEGLLRIPTAEEKIAHATNKRLALDRNRRNAR
LSV       AIMKKHAG-LRKVCRLYAPIV-----WNSMLVRNQPPADWQAMGFQYNTRFAAFDTFDYVTNQAAIQPVEGIIRRPTSAEVIAHNAHKQLALDRSNRNER
PVS       AIMKNRAG-LRKVCRLYAPVV-----WNYMLVQNRPPSDWQAMGFQWNARFAAFDTFDYVTNGAAIQPVEGLIRRPTPEETIAHNAHKSMAIDKSNRNER
PVM       AVLKKDAETLRRVCRLYAPVT-----WNHMLTHNAPPADWAAMGFQYEDRFAAFDCFDYVENTAAVQPLEGLIRRPTPREKVAHNTHKDIAVRGANRNQV
BlSV      AIMKKEAG-LRKVCRLYAPVV-----WNLMLVKNQPPSDWQAMGYPKEARFAAFDTFDYVTNGAAIQPVEGLIRGPTPAECIAHNAHKRHALDRSNRNEK
HVS       GKVQEIAG-IRRVCRLYAPVT-----WNYMHIHDSPPSDWASMGFAPNVKYAAFDCFDYVENPAAVQPLGGVIPRPTRDEYVAYNAYKLIVLNKANNNDT
CVB       GVITKHST-LRKVCRLYAAVA-----WNYMHLQQTPPSDWSAMGFHPNVKYAAFDFFDYVENGAAIRPSGGIVPRPTRAEYVAYNTYKMLALNKANNNDT
PopMV     ACIKKRS-TLSVKCADFMRPSCGITCWSTTFLQKIGSP-RATLS---ETKFAAFDTFDFVMNPAAIQPLEGLIRSPTKAEIIANETHKRIALDRNANNER
          . .      .    *        *.          .  .  .     ..**** **.* * **..*   *..   **  * .*  . *     .       *
```

```
CPMMV-S   FANLNTEVTGGTQGPEISRNFNHAKK---
CPMMV-M   FGNLNTEVTGGTQGPEISRNFNHAKK---
CLV       FSSTDSLVTGGMYGKDIKTNF-NGSNNSD
LSV       LGSLETEYTGGVQGAEIVRNHRYANNG--
PVS       LANTNVEYTGGMLGAEIVRNHRNAIN--Q
PVM       FSSLNAEVTGGMNGPELTRDYVKSNRK--
BlSV      YGNLETEYTGGLQGAEIVRNHRNAGNGSA
HVS       YGNFSAQITGGRMGPTIEHNFNNANNKKQ
CVB       FGNFDSAITGGRQGPAIHNNLNNANNKTL
PopMV     FANLGSEITGGNFGCRVGTKWRESKCDNG
           ..      ***  *   .      .
```

E.

```
CPMMV-S   M------------------------LGYKRLSVLLYLCSNKIGCCLP------FDLC-VLIALKCDPSKI--------------NYGESSYAKKRRARAI
CPMMV-M   M------------------------LGYKKVAILLCLCVNKNGFELP------FELC-VLIALKSEPTLI--------------GQGRSSYARRRRAKSI
CLV       M------------------------RERKLRKQLEDLFKRFASVQHG------HSDCINIIIAKIKS----------------DQPGESKYARRRRAKSI
LSV       MSV----W---------GAWKPNTPVGYKELKSSEIIDTQIMDEALKRR--TTIVLC--LLS---AFPRDICRDILRRTSSHIVGLGRSRYARRRRALQI
PVS       MKA------------------------------------------ERL--EMLLLC--VYRLGYILPVDVCIKII--SVAQVSVQGRSTYSCKRRARSI
PVM       M---------------------------KDVTKVALLIARAM--------------CAS--SGTFVFELAFSITEYT---GRPLGGGRSKYARRRRAISI
BlSv      MLTKGMRWIDPIAMRSTGISKRSTLVGFKVQKSSAIIEMRAMDQHDQRRLLALIKVCRSF----NLYDVGVCIHILNKVPPREVGNGCSSYAAKRRAKSI
HVS       MD--------------------------KRNKANVVLSL-----------------CSMFASRGNCIPIPIVFNIYMRAFPKLVGRGTSTYARRRRARSI
CVB       MD---------------------------VIVK---MLIL------------------RKFVEQGNVCPIHLCVDIYKRAFPRSVNKGRSSYARRRRALEL
PopMV     M------------------------VNMRKVLALMQVFRERYDHKCD------FNFCDIAVSIVCRSELDF-----------INEPGLSNYAKRRRARRL
          *                                                                               * * *. .***  .

CPMMV-S   GRCHKCYRVSPGFY-FTKCCNGVNCVPGITYQRWVEEFIK----------------FGLK------RE
CPMMV-M   GRCHRCYRVNPKFY-FTKRRDGVTCVPGISYNQFVKDFIK----------------FGFK-----GSN
CLV       ARCPRCARVSPGFY-FTTRCDGKTCRPGLSARPDLLEFIGIDLC-----VRSK---------------
LSV       GRCERCYRVYPPVC--GSKCDNKTCRPGLSINTNVANYIDHGVTEVIPWISPHRGQFYLR------PK
PVS       GRCWRCYRVYPPVC--NSKCDNRTCRPGISPNFKVVTFIR-------GW-----------------SN
PVM       ARCHRCYRLWPPTV-FTTRCDNKHCVPGISYNVRVAQFIDEGVTEVIPSVIN-------K-------RE
BlSV      GRCERCYRVFPI---GASKCYNRTCVPGISYNEKVANFIRCGVTEVI----PHPG-FNF---------
HVS       LRCERCYRVYPPL-PFSKKCDNRTCVPGISYNIKVADFIKWGVTEVI----PHPG-FNF---------
CVB       GRCHRCYRVYPPLFPEISRCDNRTCVPGISYNSKVRDYILWGVTEVI----PHPG-YNF---------
PopMV     GRCVRCFRVNPGFY-FTKRCDGITCVPGISWNYDVEDYIKRGRV-----TGDRETPYNFSWIWISSWP
            ** .* *. *         *      * **..     . .*
```

Fig. 3 (see caption on p. 776)

Table 2. Comparison of deduced amino acid sequence of P25, P12, P7, P32 (CP) and P11 (3′ORF) of CPMMV-S and CPMMV-M with respective proteins of other carlaviruses[a]

Protein	Virus	CPMMV-S	BISV	CLV	CVB	HVS	LSV	PopMV	PVM	PVS
P25	CPMMV-S	100[b]	58	NA[d]	62	NA	65	60	61	61
		100[c]	40		43		44	41	41	40
	CPMMV-M	75	54	NA	58	NA	60	57	57	57
		66	37		41		38	38	40	41
P12	CPMMV-S	100	55	NA	59	NA	61	56	59	57
		100	40		44		47	46	39	41
	CPMMV-M	77	62	NA	71	NA	64	60	65	61
		64	47		44		47	48	43	45
P7	CPMMV-S	100	55	62	52	NA	42	44	63	57
		100	25	38	30		22	21	30	38
	CPMMV-M	69	50	60	60	NA	48	41	55	54
		49	30	25	29		22	15	24	33
P32 (CP)	CPMMV-S	100	64	62	63	62	62	55	69	63
		100	49	46	45	43	47	32	54	46
	CPMMV-M	96	64	59	62	62	62	54	69	61
		93	48	42	44	45	46	31	54	45
P11	CPMMV-S	100	62	57	57	57	54	64	57	57
		100	41	37	38	41	34	39	36	43
	CPMMV-M	78	65	60	60	61	59	69	58	63
		63	44	47	48	48	43	49	47	46

[a]See Materials and methods for carlaviruses used in the analysis. Amino acid similarity and identity were analyzed using the GAP program (gap weight 3.0; length weight 0.1) of the GCG sequence analysis soft ware package

[b]Percentage amino acid similarity

[c]Percentage amino acid identity

[d]*NA* Data not available for analysis

Fig. 3 (p. 774f.). Multiple sequence alignments of different genes of CPMMV-S and CPMMV-M with their homologues of other carlaviruses: (**A**) P25, (**B**) P12, (**C**) P7, (**D**) coat protein and (**E**) P11. Asterisks below the sequences represent identity at all positions, dots represent similar residues at all positions. Dashes indicate gaps introduced to give optimal alignment. The alignment was generated using Clustal V program and performed using the sequences of different carlaviruses listed in Materials and methods

Expression of P32 ORF in E. coli

In SDS-polyacrylamide gels, extracts from isopropyl-beta-D-thiogalactopyranoside (IPTG)-induced *E. coli* transformed with pET-15b containing P32 ORF were found to contain an additional protein band that was absent from the control cells transformed with pET-15b alone (Fig. 2A, lanes 3 and 4 respectively). This protein migrated with an apparent molecular weight of 32-kDa, which is in agreement with the size of the coat protein of a purified CPMMV-M preparation (Fig. 2A, lane 2). In immunoblotting assays, the 32-kDa protein reacted with polyclonal antiserum prepared against the purified virus (Fig. 2B, lanes 2 and 3). These results suggest that P32 ORF-encoded protein is indeed the virus coat protein.

Sequence similarities between CPMMV isolates and members of the genus Carlavirus

The percent identities and similarities (in parentheses) of the deduced amino acid sequences between the corresponding ORFs of CPMMV-S and CPMMV-M are as follows: P25-66.2 (75.3), P12-64.2 (77.4), P7-49.1 (69.1), P32-92.7 (95.8) and P11-63.0 (78.0). Thus there was a much greater degree of identity between the coat proteins than between other proteins.

Multiple sequence alignments and pairwise comparisons were made between deduced amino acid sequences encoded by different ORFs of the two CPMMV isolates with corresponding proteins encoded by several carlaviruses (Fig. 3a–e and Table 2). Both CPMMV-S and CPMMV-M contain several conserved amino acid sequences/domains in different ORFs that are present in corresponding proteins of carlaviruses. Both CPMM-S and CPMMV-M shared greater amino acid sequence similarities with carlaviruses than with potexviruses (data not shown) suggesting that they are more closely related to the genus *Carlavirus* than the genus *Potexvirus*. The presence of an additional ORF between the coat protein cistron and the 3′ non-coding region in CPMMV isolates further supports that they are organizationally more similar to carlaviruses than to potexviruses. Nevertheless, the much lower degree of shared identity between these five proteins with other carlaviruses suggest that both CPMMV-S and CPMMV-M are only distantly related to any of the well characterized carlaviruses.

Discussion

Information on the genome organization and the amino acid sequence alignments provide useful data for understanding the taxonomic status of a particular virus. The genome organization and sizes of the putative protein products encoded by each ORF of either CPMMV-S or CPMMV-M are similar to those of well characterized species of the genus *Carlavirus* [2]. In addition, multiple alignments of the deduced amino acid sequences of different ORFs with their homologues from other carlviruses revealed the presence of several conserved motifs in CPMMV-S and CPMMV-M sequences: the P25 has a putative

NTP-binding motif [10] containing the highly conserved G-GKT sequence (bold letters in Fig. 3a), the P12 contains a region of high homology centered about the consensus sequence GD–H-LP-GG-YRDGTK—Y (bold letters in Fig. 3b) [7], the P7 contains a highly homologous region between the two strictly conserved cysteine residues (bold letters in Fig. 3c), the coat protein contains a strong conserved motif "His-X_8Asp-X_{15}Thr-Gly-Gly" (bold letters in Fig. 3d), which is similar to a motif in cellular serine protease, in the C-terminal region [13], and the cysteine-rich P11 contains a conserved motif with the four cysteine residues in a pattern of C-X_2-C-X_{12}-C-X_4-C that characterizes the putative zinc finger motif [18] and an upstream basic domain of amino acids (bold letters in Fig. 3e). Another interesting feature is the presence of the consensus sequence C/TTTAGGT upstream of the ATG codon of coat proteins of both CPMMV-S and CPMMV-M, which is similar to chloroplast ribosome binding sites observed in all carlaviruses [8,19]. These characteristics support the inclusion of CPMMV in the genus *Carlavirus*.

However, the range of similarity values obtained in pairwise comparison of amino acids between different ORFs of CPMMV isolates and carlaviruses suggest the distinction of CPMMV as a seperate species in the genus *Carlavirus*. This is further supported by the lack of serological relationships with any of the 18 recognised members of the genus *Carlavirus* [2], and in CPMMV being transmitted by the whitefly vector, *B. tabaci*, instead of aphid vector like the rest of the carlaviruses. Even though these characteristics support the earlier suggestion [2,17] that CPMMV be placed under a separate subgroup in the genus *Carlavirus*, it might be necessary to think of a new genus once more data accumulate about other whitefly-transmitted CPMMV-like viruses.

Although serologically related to each other and transmitted by *B. tabaci*, CPMMV-S and CPMMV-M showed differences in symptoms incited in different host plants and in the size of the genomic RNA. Their coat protein amino acid sequences are 93% identical whereas the identity values are much lower in the other proteins. In analogy with the situation described for *Potyviridae* and some members of *Geminiviridae* [22, 26], 93% identity between coat protein amino acid sequences would suggest that CPMMV-S and CPMMV-M are strains of one virus and the lower similarity among other gene sequences argue that they are distinct viruses. A comparison between the rest of the sequences of the genomes would better establish their relationships. If they are considered as strains of one virus, the biological significance of the localized regions of high variation in non-structural genes and high identity in coat protein gene is not known. In addition, nucleotide sequence analysis of other whitefly-transmitted filamentous viruses occuring in several tropical countries which are reported to be serologically related to CPMMV would help to develop criteria for their classification either as distinct viruses or strains of CPMMV. This has great practical value in the context of increasing recognition of the economic importance of whitefly-borne plant virus diseases in the tropics.

Acknowledgements

The authors wish to thank Dr. M. A. Mayo, Scottish Crop Research Institute, Dundee, UK for his help with CLUSTAL V analysis and for useful comments, and Dr. K. L. Manjunath, CREC, University of Florida, Lake Alfred, for help in computer analysis of nucleotide sequences. This paper is submitted as Journal Article No. JA 2017 by the ICRISAT.

References

1. Antignus Y, Cohen S (1987) Purification and some properties of a new strain of cowpea mild mottle virus in Israel. Ann Appl Biol 110: 563–569
2. Brunt AA (1995) Genus *Carlavirus*. In: Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA Summers MD (eds) Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses. Springer, Wien New York, pp 474–478 (Arch Virol [Suppl] 10)
3. Brunt AA, Kenten RH (1973) Cowpea mild mottle, a newly recognized virus infecting cowpeas (*Vigna unguiculata*) in Ghana. Ann Appl Biol 74: 67–74
4. Brunt AA, Atkey PT, Woods RD (1983) Intracellular occurrence of cowpea mild mottle virus in two unrelated plant species. Intervirology 20: 137–142
5. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395
6. El-Hassan SM, Naidu RA, Ahmed AH, Murant AF (1997) A serious disease of groundnut caused by cowpea mild mottle virus in the Sudan. J Phytopathol 145: 301–304
7. Foster GD (1992) The structure and expression of the genome of carlaviruses. Res Virol 143: 103–112
8. Foster GD, Mills PR (1991) Occurrence of chloroplast ribosome recognition sites within conserved elements of the RNA genomes of carlaviruses. FEBS Lett 280: 341–343
9. Foster GD, Meehan BM, Mills PR (1990) A comparison of the nucleotide sequence homologies between isolates of the Andean and ordinary strains of potato virus S and their relationship to other carlaviruses. Virus Genes 4: 257–260
10. Gorbalenya AR, Koonin EV, Donchenko AP, Blinov VM (1988) A novel super family of nucleoside triphosphate binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. FEBS Lett 235: 16–24
11. Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. J Mol Biol 227: 1–8
12. Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSATAL V: improved software for multiple sequence alignment. CABIOS 8: 189–191
13. Hillman BI, Lawrence DM (1995) Carlaviruses. In: Singh RP, Singh US, Kohmoto K (eds) Pathogenesis and host specificity in plant diseases, Vol III: viruses and viroids. Elsevier, New York, pp 35–50
14. Hobbs HA, Reddy DVR, Rajeshwari R, Reddy AS (1987) Use of direct antigen coating and protein A coating ELISA procedures for detection of three peanut viruses. Plant Dis 71: 747–749
15. Iizuka N, Rajeshwari R, Reddy DVR, Goto T, Muniyappa V, Bharathan N, Ghenekar AM (1984) Natural occurence of a strain of cowpea mild mottle virus on groundnut (*Arachis hypogaea*) in India. Phytopathol Z 109: 245–253
16. Iwaki M, Thongmeearkom P, Honda Y, Prommin M, Deema N, Hibi T, Iizuka N, Ong CA, Saleh N (1986) Cowpea mild mottle virus occuring on soybean and peanut in Southeast Asian Countries. Technical Bulletin of Tropical Agricultural Research Center 21: 106–120

17. Jeyanandarajah P, Brunt AA (1993) The natural occurrence, transmission, properties and possible affinities of cowpea mild mottle virus. J Phytopathol 137: 148–156
18. Klug A, Rhodes D (1987) "Zinc fingers": a novel protein motif for nucleic acid recognition. Trends Biochem Sci 12: 464–469
19. Levay K, Zavriev S (1991) Nucleotide sequence and gene organization of the 3′-terminal region of chrysanthemum virus B genomic RNA. J Gen Virol 72: 2 333–2 337
20. Muniyappa V, Reddy DVR (1983) Transmission of cowpea mild mottle virus by *Bemisia tabaci* in a non-persistent manner. Plant Dis 67: 391–393
21. Ohshima K, Inoue AK, Shikata E (1993) Molecular cloning, sequencing, and expression in *Escherichia coli* of the potato virus Y cytoplasmic inclusion gene. Arch Virol 128: 15–28
22. Padidam M, Beachy RN, Fauquet CM (1995) Classification and identification of geminiviruses using sequence comparisons. J Gen Virol 76: 249–263
23. Rossel HW, Thottappilly G (1985) Virus diseases of important food crops in tropical Africa. IITA, Ibadan, Nigeria
24. Saitou N, Nei M (1987) The neighbour-joining method: A new method for constructing phylogenetic trees. Mol Biol Evol 4: 406–425
25. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
26. Shukla DD, Ward CW, Brunt AA (1994) The potyviridae. CAB International, Wallingford

Authors' address: Dr. R. A. Naidu, Crop Protection Division, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India.

Received February 13, 1997