



Sequence variability in the coat protein gene of Cowpea severe mosaic virus isolates from northeastern Brazil

José Evando A. Beserra Jr.^{1*}, Eduardo C. Andrade², Rosa F.R. Araújo Camarço¹, Aline K.Q. Nascimento¹ & José Albérico A. Lima¹

¹Laboratório de Virologia Vegetal, Departamento de Fitotecnia, Universidade Federal do Ceará, 60451-970, Fortaleza, CE, Brazil; ²Embrapa Mandioca e Fruticultura, 44380-000, Cruz das Almas, BA, Brazil

Author for correspondence: José Evando Aguiar Beserra Jr., e-mail: evandojr@yahoo.com

ABSTRACT

The comovirus *Cowpea severe mosaic virus* (CPSMV) causes a disease in cowpea (*Vigna unguiculata*) but, despite its importance, there are no studies on the genetic diversity of viral populations. We have determined the nucleotide sequences of part of the coat protein (CP) gene of six Brazilian isolates of CPSMV. Genomic fragments (521 nucleotides) were RT-PCR amplified, cloned, and their sequences were compared with each other and with other comoviruses. Sequence comparisons indicated a high degree of conservation for the CP gene, with 92-100% nucleotide and 97-100% amino acid sequence identity among the isolates. There was no correlation between geographical origin and sequence identity or phylogeny among the isolates.

Key words: *Comovirus*, *Vigna unguiculata*, CPSMV, sequence diversity.

RESUMO

Variabilidade de seqüências do gene da proteína do capsídeo de isolados do *Cowpea severe mosaic virus* do Nordeste do Brasil

O comovírus *Cowpea severe mosaic virus* (CPSMV) causa a doença em feijão-caupi (*Vigna unguiculata*) conhecida como mosaico severo, mas, apesar de sua importância, não há estudos de diversidade genética de populações virais. Foram determinadas as seqüências nucleotídicas de parte do gene da proteína do capsídeo (CP) de seis isolados brasileiros de CPSMV. Os fragmentos genômicos (521 nucleotídeos) foram amplificados por RT-PCR, clonados e suas seqüências foram comparadas entre si e com as seqüências de outros comovírus. As comparações de seqüências indicaram um elevado grau de conservação para o gene CP, com identidade de 92-100% (nucleotídeos) e 97-100% (aminoácidos) entre os isolados. Não houve correlação entre origem geográfica e identidade de seqüência ou filogenia entre os isolados.

Palavras-chave: *Comovirus*, *Vigna unguiculata*, CPSMV, diversidade de seqüência.

Cowpea (*Vigna unguiculata* (L.) Walp. subsp. *unguiculata*) is the most consumed bean in the Northern and Northeastern regions of Brazil, being one of the main components of the diet of the human population. According to the FAO, cowpea is one of the best alternatives for increasing the offer of plant-derived protein (Philips et al., 2003). Severe mosaic is among the most important diseases of cowpea in Brazil. The disease occurs in all major cowpea regions of the country, greatly reducing yields. *Cowpea severe mosaic virus* (CPSMV), family *Secoviridae*, genus *Comovirus*, the causal agent of severe mosaic, is outstanding for its high incidence and severity (Lima et al., 2005). Depending on the cultivar and time of inoculation, CPSMV can reduce cowpea yields by as much as 80% (Gonçalves et al., 1982; Booker et al., 2005). The genome of comoviruses consists of two single-stranded, positive-sense RNAs, denominated RNA1 or RNA-B (for bottom component) and

RNA2 or RNA-M (for middle component). Each RNA is separately encapsidated into icosahedral particles of ca. 28 nm in diameter. Both genomic RNAs are covalently linked at their 5' ends to a virus-encoded protein named VPg and are polyadenylated at their 3' ends (Hull, 2002).

Since the first report of the incidence of CPSMV isolate in Northeastern Brazil (Lima & Nelson, 1977), several studies have indicated some variability among isolates. However, the knowledge of this variability is restricted to biological and serological properties (Lin et al., 1981, 1984; Lima et al., 2005). Currently, four CPSMV serotypes are recognized in Brazil, and these can also be differentiated by host range (Lin et al., 1984). However, few studies at the molecular level have been conducted in Brazil and in other countries (Chen & Bruening 1992; Brioso et al., 1996; Souto et al., 2002; Nogueira et al., 2003; Camarço et al., 2009). Nucleotide sequence information is available for only two CPSMV isolates: one complete genome sequence of an isolate (CPSMV-DG) obtained from cowpea in the USA (Chen & Bruening, 1992), and a partial sequence from a Brazilian

*Present address: Universidade Estadual do Piauí, Campus Prof. Barros Araújo, 64600-000, Picos, PI, Brazil.

isolate (CPSMV-PR) obtained from soybean (*Glycine max* L.) (Souto et al., 2002).

The present study reports the sequence of the coat protein (CP) gene of six CPSMV isolates sampled in Northeastern Brazil. The nucleotide and deduced amino acid sequences of these isolates were compared with sequences of other comoviruses available at GenBank. The results indicate a low degree of sequence variability for this gene.

CPSMV isolates used in this study are part of the virus collection of the Plant Virus Laboratory of the Universidade Federal do Ceará (Camarço et al., 2009). The isolates were obtained from cowpea plants collected in different States of Northeastern Brazil: CPSMV-AL (Alagoas State), CPSMV-CE (Ceará State), CPSMV-MC (Piauí State), an isolate that infects cowpea cv. Macaibo, which is supposedly resistant to CPSMV (Lima et al., 1998), CPSMV-PE (Pernambuco State), and CPSMV-PI (also from Piauí State), plus an isolate obtained from *Crotalaria paulinea*, CPSMV-CROT (Maranhão State). Each isolate was maintained in cowpea plants, in individual anti-aphid screen cages, with temperatures varying from 23 to 31°C. The CPSMV-PR isolate (from Paraná State, in Southern Brazil), also used in the analyses, was previously described by Bertacini et al. (1998).

Total RNA was extracted from leaf samples infected with each virus isolate using the Trizol reagent (Invitrogen), according to the manufacturer's instructions. First strand cDNA was synthesized from viral RNA using M-MLV Reverse Transcriptase (Promega), according to the manufacturer's instructions. Viral fragments corresponding to a portion of the CP gene (3' end of CP_L plus 5' end of CP_S) were amplified by PCR using universal primers for the genus *Comovirus*: F [5'-GCA TGG TCC ACW CAG GT-3'] and R [5'-YTC RAA WCC VYT RTT KGG MCC ACA-3'], which anneal at positions 2471 and 3032 nt of RNA2, respectively. The primers used correspond to the conserved amino acid sequences AWSTQV, present in CP_L, and GFEN/RGPN, present in CP_S (Brioso et al., 1996). PCR amplification was performed with an initial heating at 94°C for 5 min followed by 30 cycles of denaturation (94°C/1 min), annealing (41°C/2 min) and extension (72°C/1 min), followed by a final extension at 72°C for 5 min. The amplified products were visualized in a 1% agarose gel stained with ethidium bromide under UV light.

PCR products with an estimated size of 500-600 bp were cloned using the pGEM-T Vector System I (Promega) according to the manufacturer's instructions. Competent cells of *Escherichia coli* JM109 were used for transformation. *E. coli* plasmid DNA was purified as described (Rott & Jelkmann, 2001) and was cleaved with *EcoR* I to liberate the insert. Inserts were sequenced in both directions using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer), according to the manufacturer's instructions.

The nucleotide and amino acid sequences corresponding to the CP gene from the CPSMV isolates were compared to those of other CPSMV isolates and of other species from the genus *Comovirus*. DNA-A nucleotide sequences were initially submitted to a BLAST search for preliminary species assignment (www.ncbi.nlm.nih.gov/blast). Additional pairwise comparisons were made with DNAMAN version 4.0 (Lynnon Biosoft), using the Optimal Alignment option with the following parameters: Ktuple = 2, Gap penalty = 7, Gap open = 10, Gap extension = 5. Nucleotide and deduced amino acid sequences were aligned using the Clustal W program (www.ebi.ac.uk/clustalw). Phylogenetic analysis was carried out using MEGA v. 4.0 (www.megasoftware.net).

So far, the genomes of only a few species of comoviruses have been at least partially sequenced. In most cases, the gene used for this analysis was the one encoding the capsid protein (CP). The CP sequence is currently considered the major criterion for demarcation of species in the genus *Comovirus*, despite the increasing use of the proteinase-polymerase region to study genetic variability in the order *Picornavirales* (Le Gall et al., 2005; Petrzik & Koloniuk, 2010).

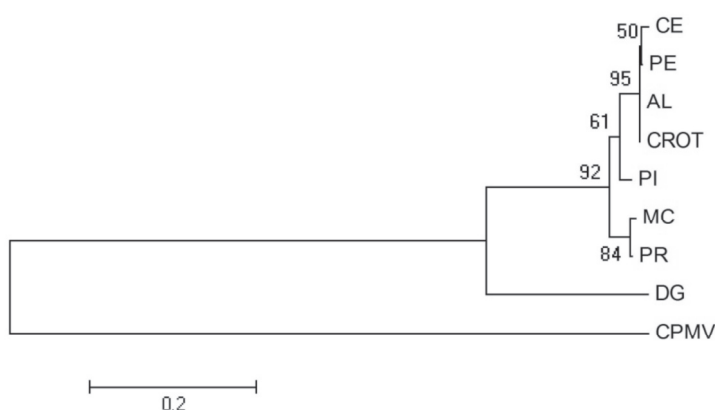
Three clones of each isolate, sequenced on both strands, were used to obtain the consensus nucleotide sequences for the isolates. The sequences (521 bp) from six CPSMV isolates were compared with each other and with other comoviruses (Table 1). Pairwise sequence comparisons revealed that the CP gene is quite conserved among the six CPSMV isolates, with 97-100% amino acid sequence identity (Table 1). The Northeastern CPSMV isolates form a monophyletic cluster together with the isolate from Paraná State (CPSMV-PR), with bootstrap confidence of 92% (Figure 1). The DG isolate collected in the USA was placed in a distinct branch of the Brazilian CPSMV isolates.

The low variability among the sequences is more evident when the MC and PR isolates are analyzed (Table 1). The MC isolate was obtained from the supposedly CPSMV resistant cowpea cv. Macaibo in Piauí State (Lima et al., 1998), while the PR isolate was obtained from soybean in Paraná State, more than 3000 km away (Bertacini et al., 1998). Even then, they shared 98% nucleotide sequence identity. Despite the similarity between these isolates, unfortunately there is no information about the ability of the PR isolate to overcome the resistance of cv. Macaibo.

It is also important to emphasize that the six isolates were collected in different years. The biggest collecting time interval is between the CE and PI isolates: the CE isolate was collected in 1977, while the PI isolate was collected in 2009. However, even with a 32-year span, the analyzed region of the CP genes remained remarkably conserved. This high conservation is also observed for other comoviruses, such as *Squash mosaic virus* (SQMV), which

TABLE 1 - Percentage of amino acid (above the diagonal) and nucleotide (below the diagonal) sequence identities between the CP genes of Brazilian CPSMV isolates, the North American isolate DG, and *Cowpea mosaic virus* (CPMV), the most closely related member from the genus *Comovirus*

	AL	CE	MC	PI	PE	CROT	PR	DG	CPMV
AL	--	98	98	100	100	100	98	86	39
CE	99	--	97	98	98	98	96	84	39
MC	93	92	--	98	98	98	98	86	37
PI	95	95	93	--	100	100	98	86	39
PE	99	99	93	96	--	100	98	86	39
CROT	100	99	93	95	99	--	98	86	39
PR	93	92	98	93	93	93	--	86	38
DG*	73	72	75	74	73	73	75	--	42
CPMV**	54	55	52	53	54	54	53	53	--

**FIGURE 1** - Phylogenetic analysis of *Cowpea severe mosaic virus* (CPSMV) isolates based on their CP nucleotide sequences. The tree was constructed by neighbor-joining method with MEGA 4.0. The numbers at each branch correspond to bootstrap confidence levels (2000 replications). Vertical lengths are proportional to the genetic distances between the isolates. Outgroup: *Cowpea mosaic virus* (CPMV) (GenBank access # NC_003550). CPSMV isolates include DG from the USA (NC_003544) and the Brazilian isolates AL (HM448424), CE (HM448426), CROT (HM448425), MC (HM450147), PR (AF263549), PE (GQ229416) and PI (HM450148).

showed a nucleotide sequence identity of 96.6% for the CP gene of five isolates from the USA, and *Radish mosaic virus* (RaMV), with a maximum value of 98.4% of amino acid sequence identity for the CP_L (Haudenschild & Palukaitis, 1998; Komatsu et al., 2008).

Not surprisingly, the analysis indicated that the Brazilian CPSMV isolates share a low sequence identity with the DG isolate from the USA (Table 1). The highest value observed was 75%, between DG and both the MC and PR isolates. The differences in nucleotide sequences were of little influence on the protein product, as the amino acid substitutions were mostly synonymous. In a total of 187 amino acids there were only eight substitutions. The PR, MC and CE isolates presented the largest number of substitutions: four, three and three, respectively (data not shown). Despite the low variability of the CP gene among the CPSMV isolates, there is a notable difference in their symptomatology on cowpea varieties (Camarço et al., 2009). Therefore, while the analyzed region of CP genes is conserved, other parts of the CP genes or of the viral genome may be more variable, and the viral determinants of symptomatology could be linked to these more variable regions.

ACKNOWLEDGMENTS

This research was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico – FUNCAP. The authors thank Dr. F. Murilo Zerbini for suggestions.

REFERENCES

- Bertacini PV, Almeida AMR, Lima JAA, Chagas CM (1998) Biological and physicochemical properties of Cowpea severe mosaic comovirus isolated from soybean in State of Paraná. *Brazilian Archives Biology Technology* 41:409-416.
- Brioso PST, Santiago LJM, Anjos JRN, Oliveira DE (1996) Identificação de espécies do gênero *Comovirus* através de “polymerase chain reaction”. *Fitopatologia Brasileira* 21:219-225.
- Booker HM, Umaharan P, McDavid CR (2005) Effect of *Cowpea severe mosaic virus* on growth characteristics and yield of cowpea. *Plant Disease* 89:515-520.
- Camarço RFEA, Nascimento AKQ, Andrade EC, Lima JAA (2009) Biological, serological and molecular comparison between

- isolates of *Cowpea severe mosaic virus*. *Tropical Plant Pathology* 34:239-244.
- Chen X, Bruening G (1992) Cloned copies of *Cowpea severe mosaic virus* RNAs: infectious transcripts and complete nucleotide sequence of RNA 1. *Virology* 191:607-618.
- Gonçalves MFB, Lima JAA (1982) Efeitos do *Cowpea severe mosaic virus* sobre a produtividade do feijão-de-corda. *Fitopatologia Brasileira* 7:547. (Resumo)
- Haudenschild JS, Palukaitis P (1998) Diversity among isolates of *Squash mosaic virus*. *Journal General Virology* 79:2331-2341.
- Hull R (2002) *Matthew's Plant Virology*. 4th Ed. London. Academic Press.
- Komatsu K, Hatada K, Hashimoto M, Ozeki J, Maejima K, Kagiwada S, Yamaji Y, Namba S (2008) Complete nucleotide sequence of a California isolate of Radish mosaic virus. *Archives Virology* 153:2167-2168.
- Le Gall O, Iwanami T, Karasev AV, Jones T, Lehto K, Sanfacon H, Wellink J, Wetzel T, Yoshikawa N (2005) Family *Comoviridae*. In: Fauquet CM (Ed.) *The eighth report of the International Committee on Taxonomy of Viruses*. San Diego CA. Academic Press., pp 807- 818.
- Lima JAA, Nelson MR (1977) Etiology and epidemiology of mosaic of cowpea in Ceará. *Plant Disease* 61:864-867.
- Lima JAA, Lima RCA, Gonçalves MAR, Sittolin IM (1998) Biological and serological characteristics of a genetically different *Cowpea severe mosaic virus* strain. *Virus Review and Research* 3:57-65.
- Lima JAA, Sittolin IM, Lima RCA (2005) Diagnose e estratégias de controle de doenças ocasionadas por vírus. In: Freire Filho FR, Lima JAA, Ribeiro VQ (Eds.) *Feijão-Caupi Avanços Tecnológicos*. Embrapa Informação Tecnológica. Brasília DF. pp. 403-459.
- Lin MT, Anjos RN, Rios GP (1981) Serological grouping of *Cowpea severe mosaic virus* isolates from Central Brazil. *Plant Disease* 71:435-438.
- Lin MT, Hill JH, Kitajima EW, Costa CL (1984) Two novel serotypes of *Cowpea severe mosaic virus*. *Plant Disease* 74:581-585.
- Nogueira MSR, Freire Filho FR, Brioso PST (2003) Diferenciação molecular entre os sorotipos I e II do *Cowpea severe mosaic virus* (CPSMV), através de RT-PCR. *Summa Phytopathologica* 29:73.
- Petrzik K, Koloniuk I (2010) Emerging viruses in the genus *Comovirus*. *Virus Genes* 40:290-292.
- Philips RD, McWatters KH, Chinnan MS, Hung Y, Beucha LR, Sefa-Dedeh S, Sakyi-Dawson E, Ngoddy P, Nnanyelugo D, Enwere J, Komey NS, Liu K, Mensa-Wilmot Y, Nnanna IA, Okeke C, Prinyawiwatkul W, FK Saalia FK (2003) Utilization of cowpeas for human food. *Field Crops Research* 82:193-213.
- Rott ME, Jelkmann W (2001) Characterization and detection of several filamentous viruses of cherry: Adaptation of an alternative cloning method (DOP-PCR) and modification of PCR extraction protocol. *European Journal Plant Pathology* 107:411-420.
- Souto ER, Almeida AMR, Bianchini A, Sartori F, Calves ES (2002) Análise molecular do segmento do RNA2 de *Comovirus* isolados de soja no Estado do Paraná. *Fitopatologia Brasileira* 27:525-527.

TPP 170 - Received 30 August 2010 - Accepted 18 January 2011
Section Editor: F. Murilo Zerbini