Arch Virol (2002) 147: 1997–2007
DOI 10.1007/s00705-002-0851-1

Archives of Virology
Printed in Austria

Taxonomic position of sugarcane streak mosaic virus in the family *Potyviridae**

Brief Report

M. Hema[1], **P. Sreenivasulu**[1], and **H. S. Savithri**[2]

[1]Department of Virology, Sri Venkateswara University, Tirupati, Andhra Pradesh, India
[2]Department of Biochemistry, Indian Institute of Science, Bangalore, Karnataka, India

Received March 26, 2002; accepted May 16, 2002
Published online August 12, 2002

Summary. A cDNA library was generated from purified RNA of sugarcane streak mosaic virus – Andhra Pradesh (SCSMV-AP). Two overlapping clones covering 3160 nucleotides encoding partial CI, complete 6K2, VPg-NIa and NIb genes were sequenced. A comparison of this sequence along with the 3′ terminal 1315 nucleotides of SCSMV-AP determined earlier with the other members of the family *Potyviridae* indicated that it had only 30% identity at the amino acid level for the partial polyprotein open reading frame (ORF) with the members of *Ipomovirus* and *Tritimovirus* genera. Further, in the most conserved NIb region also there was only 40% identity with the type members of these genera. Based on this analysis, we suggest the taxonomic affiliation of SCSMV-AP into an undescribed new genus in the family *Potyviridae*.

*

SCSMV-AP, the pathogen causing mosaic disease of sugarcane in India was recently characterized [16]. Incidence of mosaic disease is almost 100% on sugarcane and considering the vast area under sugarcane cultivation, it results in significant yield losses [1, 19]. SCSMV-AP is a flexuous filamentous virus (890 × 15 nm) with monopartite ssRNA genome of approximate size 10 kb, encapsidated by coat protein subunits of M_r 40 kDa [16]. The virus induces pinwheel and laminated aggregate type of inclusions that are characteristic feature of members of the family *Potyviridae* [18]. It is transmitted through vegetative propagules

*The nucleotide sequence reported in this paper has been updated in GenBank and the accession number is Y17738.

(setts) under natural conditions. However, the vector for SCSMV-AP is yet to be identified. Natural infection of SCSMV-AP is confined to sugarcane alone and it infects sorghum under experimental conditions, but not any other tested Poaceous members including wheat (*Triticum aestivum* L.) [17, 18].

The *Potyviridae* is the largest family among plant viruses, whose members have positive-sense single-stranded RNA genomes, encoding a large polyprotein which gets proteolytically processed into functional proteins including P1, HC-Pro, P3, 6KI, CI, 6K2, VPg-NIa, NIb and coat protein (CP) [24]. Initially, the family was organized into four genera (*Potyvirus*, *Bymovirus*, *Rymovirus*, and *Ipomovirus*) on the basis of coat protein sequence comparisons and vector taxa [31]. Later, from an analysis of 3′-terminal sequences of aphid transmitted Narcissus latent virus and Maclura mosaic virus, which shared limited sequence identity with the members of genus *Potyvirus*, these viruses were included into a new genus *Macluravirus* [4]. Earlier, WSMV and BrSMV were classified with rymoviruses. However, since the two viruses shared less than 40% identity with rymoviruses and were transmitted by a different mite species (*Aceria tulipae*), they were separated into a new genus-*Tritimovirus* [27, 32, 21, 35]. The increasing availability of potyvirus sequence data presents new opportunities for comparisons that may reveal insights about the phylogeny and evolution of this important family of viruses. An accurate phylogenetic framework is an essential component of correct nomenclature and comparative understanding of the etiology, epidemiology, evolution and speciation of these viruses.

In most instances, coat protein and 3′ untranslated region (UTR) sequences can readily establish the taxonomic status of a particular virus in the family *Potyviridae* [9, 3, 14]. Previous studies on molecular characterization of SCSMV-AP showed that the deduced coat protein sequence has 93.6% identity and 3′ UTR has 94.3% identity with SCSMV-PAK [16, 15]. SCSMV-AP was therefore considered a strain of SCSMV-PAK which was earlier proposed as a tentative

Table 1. Percent identity of the partial ORF (C terminal 1420 amino acids) of SCSMV-AP with other members of the family *Potyviridae*

Virus	% Identity	Accession number
SCSMV-AP	100	Y17738 and this study
SCSMV-PAK	92	U75456
BrSMV	29	Z48506
WSMV	29	AF057533
SPMMV	30	Z73124
RGMV	20	Y09854
BaYMV	20	NC_002990
MacMV	21	U58771
PVY	25	NC_001616
SCMV	24	NC_003398
SrMV-SCI	23	U57359

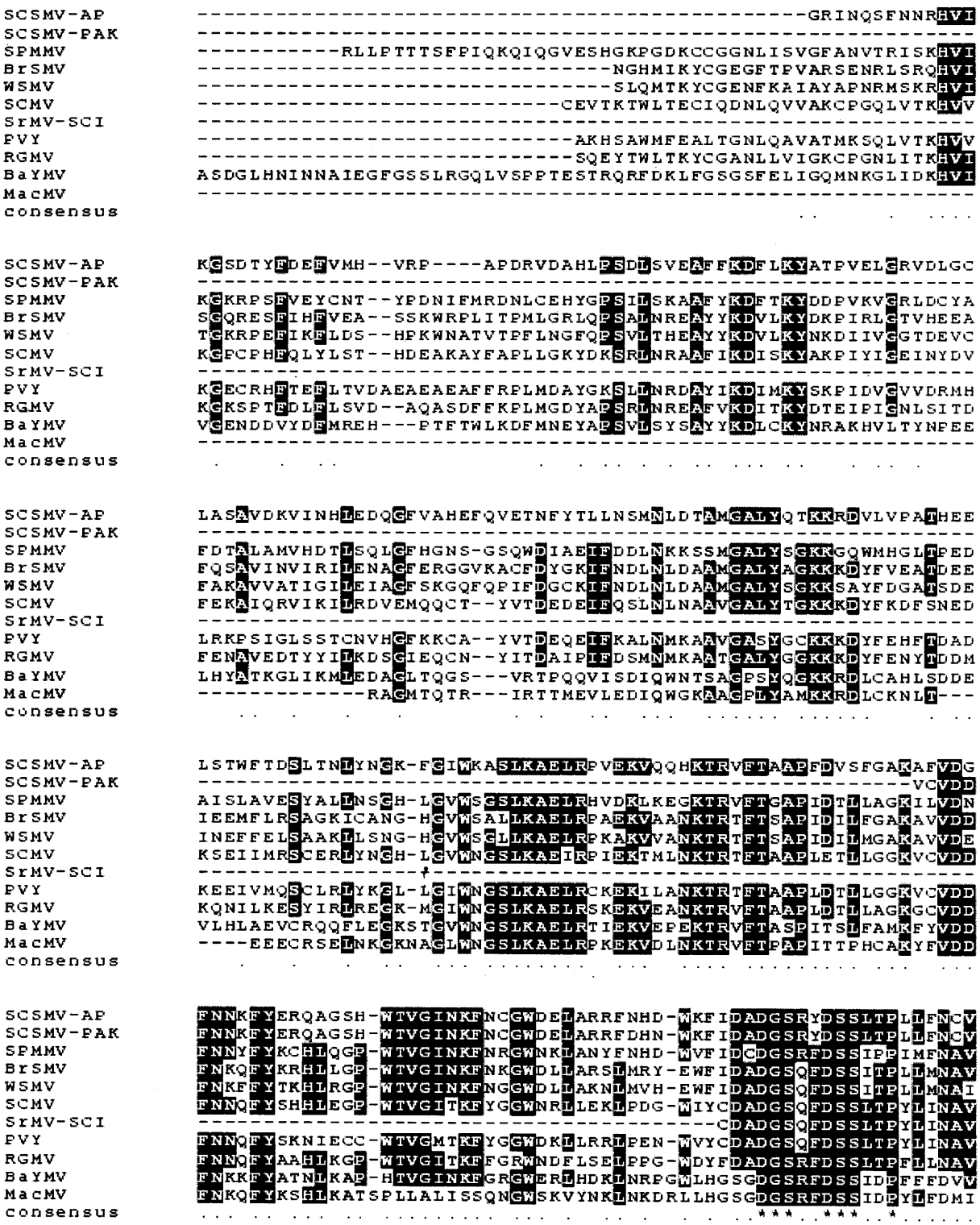
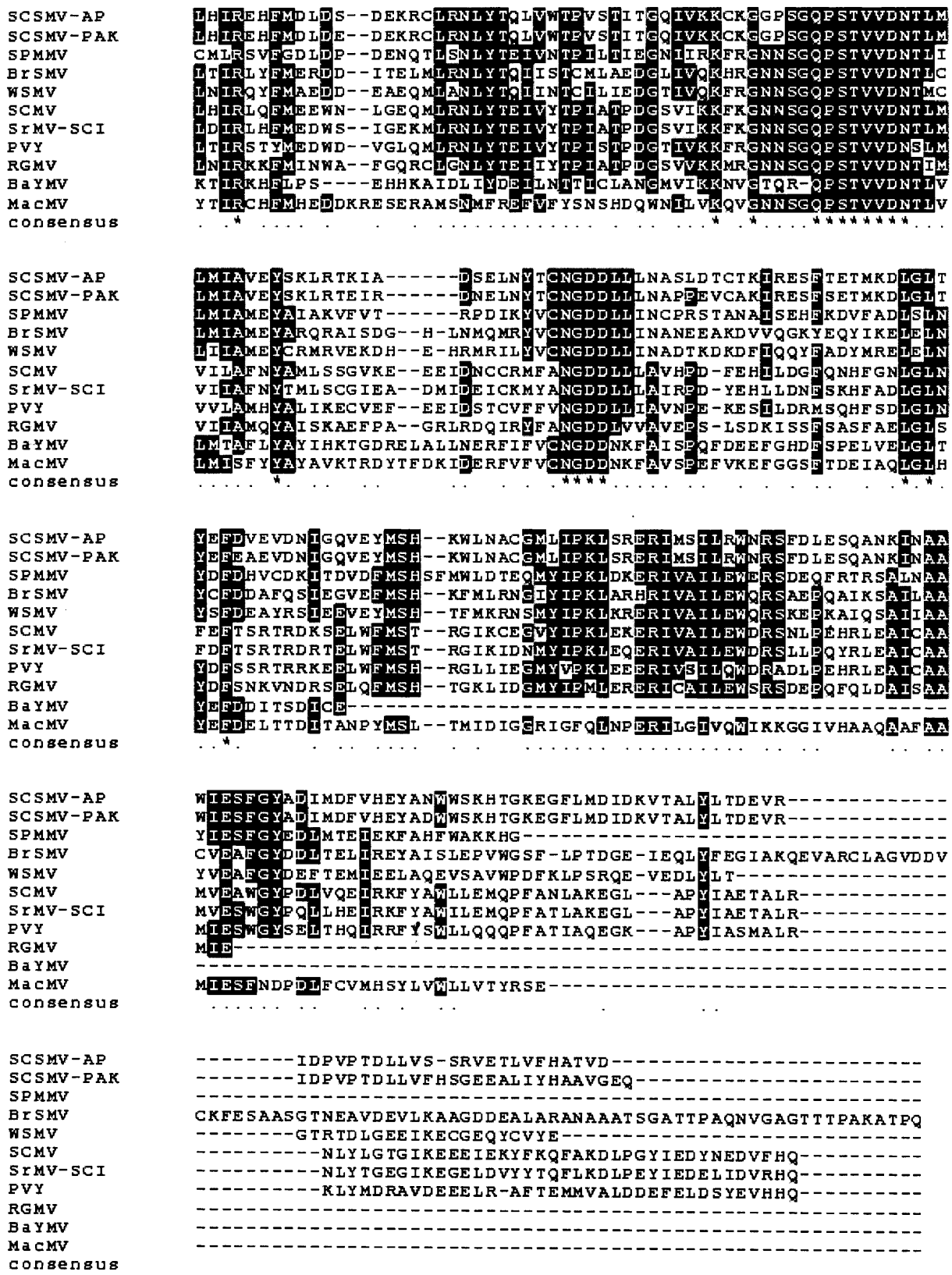


Fig. 1 (*continued*)


SCSMV-AP
SCSMV-PAK
SPMMV
BrSMV
WSMV
SCMV
SrMV-SCI
PVY
RGMV
BaYMV
MacMV
consensus

member of the genus *Tritimovirus* as it shared limited (∼ 30%) sequence identity with definitive members of this genus [16, 15]. Among the potyviruses, the most variable gene product is P1, followed by P3 and the most conserved gene product is NIb, the RNA dependent RNA polymerase [2]. So sequencing of NIb gene of SCSMV-AP could reveal its correct taxonomic position in the family *Potyviridae*. In this report, we present additional sequence of SCSMV-AP and utilize phylogenetic analysis to ascertain taxonomic placement of SCSMV.

SCSMV-AP was collected from commercial sugarcane with mosaic disease in Chittoor District of Andhra Pradesh (A.P.), India and propagated on *Sorghum bicolor* cv. Rio plants by periodic sap inoculation. SCSMV-AP was purified from infected sorghum leaves [16] and RNA was isolated as described by Zaitlin (1979) [38]. Virus specific oligonucleotide primer CD1 (5′ CTGTAGGCACTGGG TCAATCCTCA 3′) was designed based on the available sequence information of SCSMV-AP [18]. First strand complementary DNA was synthesized by reverse transcription with viral RNA as template and CD1 as primer using SuperScript II Reverse Transcriptase (GIBCO-BRL) as per the manufacturer's instructions. Second strand cDNA was then produced by RNase H/DNA polymerase method [13]. The cDNA was fractionated on a Sephacryl S-1000 column (Sigma) and ligated to *Hin*cII cut pUC 19. Cloned cDNA was transformed into competent *E. coli* DH5 α cells. Recombinant clones were screened by blue/white selection followed by colony hybridization using first strand cDNA as probe [29, 12]. Recombinant plasmids were isolated by alkaline lysis method [29] and analysed by restriction digestion. Two overlapping cDNA clones harboring 2.5 kb and 3 kb inserts (pSCSMV 32 and 57) were sequenced by following Sanger's dideoxynucleotide chain termination method [30] using ABI Prism automated DNA sequenator. Computer analysis of sequence data and compilation of amino acid sequences was performed using GCG set of programs (Genetics Computer Group, Wisconsin). Multiple alignments of the amino acid sequence data using pairwise distance measurements were carried out by CLUSTAL W [34]. Neighbor-joining trees [26] were generated from CLUSTAL W output using MEGA2 program [33].

The complete nucleotide sequence of the aligned cDNA clones revealed one large open reading frame (ORF) of 3160 nucleotides. This sequence was compiled with 3′ terminal 1315 nt sequence determined earlier [18] and the total deduced amino acid sequence of the partial ORF (1420 amino acids) was aligned with the corresponding sequences of selected members of *Potyviridae* (Table 1). As apparent from Table 1, SCSMV-AP and SCSMV-PAK (whose partial sequence of 650 amino acids only is available) did not show more than 30% identity with

Fig. 1. Multiple alignment of the deduced amino acid sequence of the NIb region of SCSMV-AP with other selective members of family *Potyviridae*. Sequences were aligned using CLUSTAL W program [34]. Shading was performed by the BOXSHADE program (ISRC Bioinformatics Group). Black boxes indicate residues that are identical. Asterisks indicate perfect matches within all sequences. The gaps introduced to optimize the sequence alignment are represented by dashes

any of the members of the family *Potyviridae* suggesting that these two viruses might belong to a distinct genus in the family. Based on amino acid sequence of coat protein of SCSMV-AP determined by us earlier, we have predicted a novel cleavage motif HATVD/G between NIb and CP [16]. Using this motif as a probe and the expected size of NIb to be 60 kDa, the cleavage motif between NIa/NIb was identified to be NLTIQ/G. It is interesting to note that these cleavage sites are not preceded by the conserved amino acids VXH observed in most of the members of the family *Potyviridae* [31]. Analysis of the derived amino acid sequence of NIb gene revealed the presence of the consensus motif GNNSGQPSTVVDNTLM and NGDDL which are characteristic of RNA dependent RNA polymerases of positive-strand RNA viruses [8, 20, 22].

The NIb region is the most conserved part of the genome in all the members of the *Potyviridae*. Hence a more detailed analysis was carried for this region. Alignment of the deduced amino acid sequences of NIb region of SCSMV-AP with other members is presented in Fig. 1. The viruses which were taken for comparison included at least one representative member from each genus in the family *Potyviridae* namely brome streak mosaic tritimovirus (BrSMV) [11], wheat streak mosaic tritimovirus (WSMV) [32], sweet potato mild mottle ipomovirus (SPMMV) [7], ryegrass mosaic rymovirus (RGMV) [10], barley yellow mosaic bymovirus (BaYMV) [6], maclura mosaic macluravirus (MacMV) [4], potato virus Y potyvirus (PVY) [25] and two potyviruses infecting sugarcane naturally viz. sugarcane mosaic potyvirus (SCMV) [5] and sorghum mosaic potyvirus (SrMV-SCI) [37] along with SCSMV-AP [17 and this study] and SCSMV-PAK [15]. Percentage sequence identity determined for SCSMV-AP with other members of the family at CP and NIb regions is presented in Table 2. In conformity with the earlier observations, SCSMV-AP and SCSMV-PAK showed maximum sequence identity of approximately 30% at CP region with WSMV and BrSMV, which are definitive members of the genus *Tritimovirus*. In the most conserved NIb region SCSMV showed only 41% and 39% identity with BrSMV and WSMV, respectively. Interestingly, it also showed 40% identity with SPMMV, a sole member of *Ipomovirus*, thus casting doubts on the earlier proposition that SCSMV-AP and SCSMV-PAK could belong to genus *Tritimovirus*. This was substantiated by the phylogenetic analysis and neighbor-joining trees generated from the CP and NIb of SCSMV-AP and those of selected members of the *Potyviridae* (Fig. 2). According to this phylogram, SCSMV-AP and SCSMV-PAK cluster not only with WSMV, BrSMV but also with SPMMV, therefore these two viruses probably belong to a new genus unrelated to these genera (Fig. 2). In general, strains of viruses have sequence identities $> 85\%$, members within genera have sequence identities $> 55\%$. It is therefore suggested that SCSMV-AP does not belong to any of the described genera in the family *Potyviridae*. Recently Rabenstein et al. (2002) [23] have reported the phylogenetic relationships, strain diversity and biogeography of tritimoviruses. Their analysis also suggests that SCSMV is not a tritimovirus and may represent a new genus within the family *Potyviridae*. Further, their analysis showed that Oat necrotic mottle virus (ONMV) is a definitive member of tritimovirus and it is not a rymovirus.

Table 2. Percent amino acid sequence identity at CP region (above the diagonal) and NIb region (below the diagonal) between SCSMV-AP and other selected members of the family *Potyviridae* after pairwise alignment of sequences using CLUSTAL W algorithms

	SCSMV-AP	SCSMV-PAK	BrSMV	WSMV	SPMMV	RGMV	BaYMV	MacMV	PVY	SCMV	SrMV-SCI
SCSMV-AP		94	29	30	25	20	11	7	14	19	16
SCSMV-PAK*	90		30	29	24	20	11	7	17	18	15
BrSMV	41	42		50	27	24	11	12	24	24	22
WSMV	39	38	62		20	20	10	4	25	23	18
SPMMV	40	42	41	44		21	16	15	22	18	19
RGMV	39	35	44	45	44		18	15	48	46	41
BaYMV	30	22	31	31	32	33		26	19	16	17
MacMV*	34	28	35	32	34	32	39		25	20	17
PVY	35	38	37	37	39	52	30	33		58	58
SCMV	35	36	37	38	39	56	31	35	59		70
SrMV-SCI*	33	33	33	34	34	42	20	28	59	78	

*Represents only partial NIb sequence availability in the database

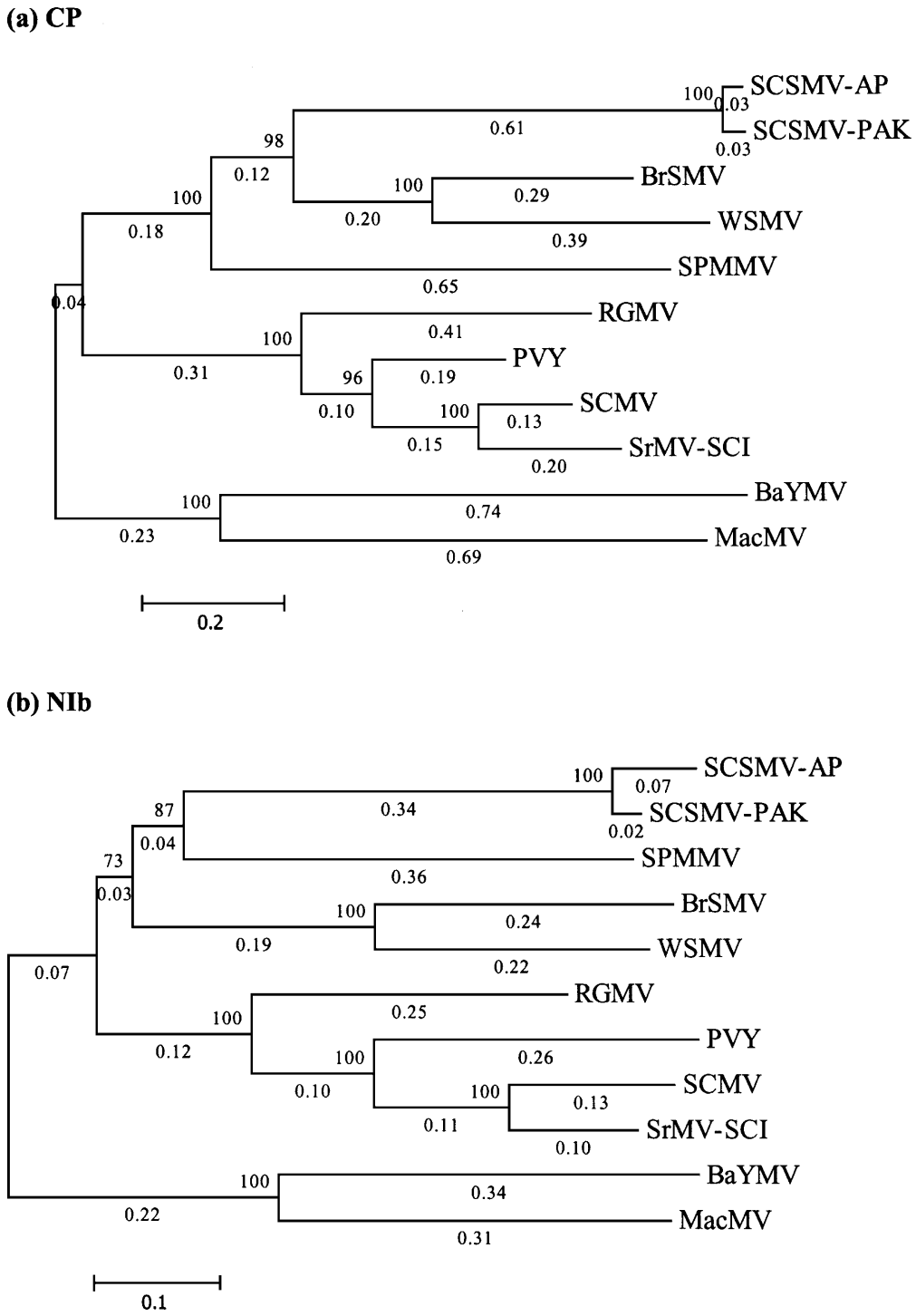


Fig. 2. Neighbor-joining relationship dendrograms illustrating the taxonomic position of SCSMV-AP among the members of the family *Potyviridae*. The trees were constructed using MEGA2 program [33] from multiple sequence alignment of (**a**) CP and (**b**) NIb obtained using CLUSTAL W [34]. Horizontal scale indicates sequence divergence with pairwise distance scores, vertical scale is arbitrary. The dataset was subjected to 100 bootstrap

In the past, the vector based taxonomy of the *Potyviridae* was thought to correlate well with sequence data and phylogenetic analysis [36]. However, questions have previously been raised concerning how these viruses are distinguished at the genus level [39]. Recent sequence data of several mite transmitted members of the *Rymovirus* genus indicated that they show significant sequence similarities with aphid transmitted potyviruses and have conserved motifs common to potyviruses [28]. Similarly whitefly transmitted SPMMV, the type member of *Ipomovirus* genus was shown to have similarity with mite transmitted tritimoviruses [7]. Aphid transmitted macluraviruses also showed similar behavior with bymoviruses [4]. Thus species within a genus could be vectored by the same taxa. However, vector taxa alone is not sufficient to classify a species into a genus.

The sequence analysis of SCSMV-AP presented in this paper clearly shows that it could belong to an undescribed new genus whose natural occurrence is confined to sugarcane alone and natural perpetuation is through setts. Further, sequence information on this virus and identification of the natural vector may reveal its relation with other genera, as of now it appears to be a member of a new undescribed genus in the family *Potyviridae*.

Acknowledgements

We are grateful to Dr. Claude M. Fauquet, Director ILTAB, Member Danforth Plant Science Centre, 975 North Warson Rd, St Louis MO63132, for carryingout an independent phylogenetic analysis of SCSMV-AP sequence with other members which confirms the conclusions presented here. Dr. M. Hema is a recepient of the research associateship (RA) from the Council of Scientific and Industrial Research (CSIR), New Delhi, India and she acknowledges the financial assistance from CSIR. The assistance by Kirthi Narayanaswamy and Mr. K. V. S. Badrinath in the analysis of sequence and phylogenetic tree construction is gratefully acknowledged. We thank DBT, India for providing the facilities for DNA sequencing and Bioinformatics at the IISc, Bangalore, India.

References

1. Agnihothri VP (1996) Current sugarcane disease scenario and management strategies. Ind Phytopathol 49: 109–126
2. Aleman-Verdaguer MEA, Goudou-Urbino C, Dubern J, Beachy RN, Fauquet C (1997) Analysis of the sequence diversity of the P1, HC, P3, NIb and CP genomic regions of several yam mosaic potyvirus isolates: implications for the intrapecies molecular diversity of potyviruses. J Gen Virol 78: 1253–1264
3. Atreya CD (1992) Application of genome sequence information in potyvirus taxonomy: an overview. In: Barnett OW (ed) Potyvirus Taxonomy New York. Arch Virol [Suppl] 5: 17–23
4. Badge J, Robinson DJ, Brunt AA, Foster GD (1997) 3′-terminal sequence of the RNA genomes of narcissus latent and maclura mosaic viruses suggest that they represent a new genus of the *Potyviridae*. J Gen Virol 78: 253–257
5. Chen J, Shi N, Cheng Y, Diao A (1999) Molecular analysis of barley yellow mosaic virus isolates from China. Virus Res 64: 13–21

6. Chen J, Chen JP, Adams MJ (2002) Characterization of potyviruses from sugarcane and maize in China. Arch Virol 147: 1237–1246
7. Colinet D, Kummert J, Lepoivre P (1996) Molecular evidence that whitefly-transmitted sweet potato mild mottle virus belongs to a distinct genus of the *Potyviridae*. Arch Virol 141: 125–135
8. Domier LL, Shaw JG, Rhoads RE (1987) Potyviral proteins share amino acid sequence homology with picorna-, como- and caulimoviral proteins. Virology 158: 20–27
9. Frenkel MJ, Ward CW, Shukla DD (1989) The use of 3′ non-coding sequence in the taxonomy of potyviruses: application to watermelon mosaic virus 2 and soyabean mosaic virus. J Gen Virol 70: 2775–2783
10. GenBank Accession number: Y09854
11. Götz R, Maiss E (1995) The complete nucleotide sequence and genome organization of the mite transmitted brome streak mosaic virus in comparison with those of potyviruses. J Gen Virol 76: 2035–2042
12. Grunstein M, Hogness DS (1975) Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. Proc Natl Acad Sci USA 72: 3961–3965
13. Gubler U, Hoffman BJ (1983) A simple and efficient method for generating cDNA libraries. Gene 25: 263–269
14. Habera LF, Berger PH, Reddick BB (1994) Molecular evidence from 3′terminus analysis that tobacco vein banding mosaic virus is a distinct member of the potyvirus group. Arch Virol 138: 27–38
15. Hall JS, Adams B, Parsons TJ, French R, Lane LC, Jensen SG (1998) Molecular cloning, sequencing and phylogenetic relationships of a new potyvirus; sugarcane streak mosaic virus and a reevaluation of the classification of the *Potyviridae*. Mol Phylogen Evol 10: 323–332
16. Hema M, Joseph J, Gopinath K, Sreenivasulu P, Savithri HS (1999a) Molecular characterization and interviral relationships of a flexuous filamentous virus causing mosaic disease of sugarcane (*Saccharum officinarum* L.) in India. Arch Virol 144: 479–490
17. Hema M, Venkatramana M, Savithri HS, Sreenivasulu P (1999b) Biological, antigenic and genomic relationships among the virus isolates causing mosaic disease of sugarcane in South India. Curr Sci 77: 698–702
18. Hema M, Savithri HS, Sreenivasulu P (2001) Sugarcane streak mosaic virus: occurrence, purification characterization and detection. In: Rao GP, Ford RE, Tosic M, Teakle DS (eds), Sugarcane pathology, vol. 2, Virus and phytoplasma diseases. Science Publishers, Inc, USA, pp 37–70
19. Jain RK, Rao GP, Varma A (1998) Present status of management of sugarcane mosaic virus. In: Hadidi A, Khetarpal RK, Koganezava H (eds) Plant virus disease control. American Publishing Company, Minnesota, pp 495–523
20. Koonin EV (1991) The phylogeny of RNA dependent RNA polymerases of positive strand RNA viruses. J Gen Virol 72: 2197–2206
21. Mayo MA (1999) Developments in plant virus taxonomy since the publication of the 6th ICTV Report. International Committee on Taxonomy of Viruses. Arch Virol 144: 1659–1666
22. Poch O, Sauvagent I, Delarue M, Tordo N (1989) Identification of four conserved motifs among the RNA dependent polymerase encoding elements. EMBO J 8: 3867–3874
23. Rabenstein F, Seifers DL, Schubert J, French R, Stenger DC (2002) Phylogenetic relationships, strain diversity and biogeography of tritimoviruses. J Gen Virol 83: 895–906
24. Riechmann JL, Lain S, Garcia JA (1992) Highlights and prospects of potyvirus molecular biology. J Gen Virol 73: 1–16

25. Robaglia C, Durand-Tardiff M, Tronchet M, Boudazin G, Astier-Manifacier S, Casse-Delbart F (1989) Nucleotide sequence of potato virus Y (N strain) genomic RNA. J Gen Virol 70: 935–947
26. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425
27. Salm SN, Rey MEC, Robertson NL, French R, Rabenstein F, Schubert J (1996a) Phylogenetic justification for splitting the rymovirus genus of the taxonomic family *Potyviridae*. Arch Virol 141: 2237–2242
28. Salm SN, Rey MEC, Robertson NL, French R, Rabenstein F, Schubert J (1996b) Molecular cloning and nucleotide sequencing of the partial genomes of Agropyron and Hordeum mosaic viruses, two members of the *Rymovirus* genus in the family *Potyviridae*. Arch Virol 141: 2115–2127
29. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
30. Sanger F, Nicklen S, Caulson AR (1977) DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74: 5463–5467
31. Shukla DD, Ward CW, Brunt AA (1994) The *Potyviridae*. CAB International, Wellingford
32. Stenger DC, Hall JS, Choi IR, French R (1998) Phylogenic relationships within the family *Potyviridae*: wheat streak mosaic virus and brome streak mosaic virus are not members of the genes *Rymovirus*. Phytopathology 88: 782–787
33. Sudhir Kumar, Tamura K, Ingrid BJ, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17: 1244–1245
34. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680
35. Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens E, Esters M, Lemon S, McGeoch D, Wickener RB, Mayo MA, Pringle CR, Maniloff J (eds) (1999) Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, New York
36. Ward CW, Weiller GF, Shukla DD, Gibbs A (1995) Molecular systematics of the *Potyviridae*, the largest plant virus family. In: Gibbs AJ, Calisher CH, Arenal G (eds), Molecular basis of virus evolution. Press Syndicate of the University of Cambridge, Cambridge, pp 447–500
37. Yang ZN, Mirkov TE (1997) Sequence and relationships of sugarcane mosaic and sorghum mosaic virus strains and development of RT-PCR based RFLPs for strain discrimination. Phytopathology 87: 932–939
38. Zaitlin M (1979) The RNA of monopartite plant viruses. In: Hall TC, Davies JW (eds), Nucleic acids in plants, vol. 2. CRC Press, Boca Raton, pp 31–64
39. Zettler FW (1992) Designation of potyvirus genera: A question of perspective and timing. Arch Virol [Suppl] 5: 235–237

Author's address: Prof. H. S. Savithri, Department of Biochemistry, Indian Institute of Science, Bangalore-560 012, Karnataka, India; e-mail: bchss@biochem.iisc.ernet.in