



ELSEVIER

Virus Research 53 (1998) 187–196

Virus
Research

The nucleotide sequence and genome organization of the whitefly transmitted sweetpotato mild mottle virus: a close relationship with members of the family *Potyviridae*¹

D. Colinet *, J. Kummert, P. Lepoivre

Faculté Universitaire des Sciences Agronomiques, Unité de Phytopathologie, 13, Avenue Maréchal Juin, 5030 Gembloux, Belgium

Received 11 September 1997; accepted 24 November 1997

Abstract

Primers corresponding to conserved regions in the RNA-dependent RNA polymerase and the RACE procedure led to the cloning of the complete sweetpotato mild mottle virus (SPMMV) RNA genome. The assembled SPMMV genomic sequence was 10 818 nucleotides in length with a polyadenylated tract at the 3' terminus. The structure and organization of the SPMMV genome appear to be similar to those of potyviruses and rymoviruses. A 5' untranslated region, rich in A and U residues, is present between nucleotides 1 and 139. A putative initiation codon, at nucleotides 140–142, marks the beginning of a large open reading frame (ORF) which ends in UAA at positions 10 508–10 510. A 308-nucleotide untranslated region is present between the termination codon of the ORF and the beginning of the 3' polyadenylated region. Almost all known potyvirus motifs are present in the polyprotein of SPMMV. However, motifs in the putative helper-component and coat protein of SPMMV are incomplete or missing, which may account for its vector relations. Despite similarities with rymoviruses, potyviruses and, to a lesser extent, bymoviruses, comparative sequence analyses demonstrated that SPMMV belongs to a distinct genus of the family *Potyviridae*. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: SPMMV; Whitefly transmission; *Potyviridae*-specific primers; RACE; Complete nucleotide sequence; *Ipomovirus*

1. Introduction

The family *Potyviridae* was divided into three genera (Barnett, 1991; Shukla et al., 1994). The genus *Potyvirus*, with potato virus Y (PVY) as the type member, is the largest with more than 180

* Corresponding author. Tel.: + 32 81 622432; fax: + 32 81 610126; e-mail: colinet@fsagx.ac.be

¹ The nucleotide sequence data reported in this paper have been submitted to the EMBL nucleotide sequence database and have been assigned the accession number Z73124.

members. Potyviruses are transmitted non-persistently by many aphid species, while bymoviruses (type species: barley yellow mosaic virus (BaYMV)) and rymoviruses (type species: ryegrass mosaic virus (RGMV)) are transmitted by a fungus and mites, respectively (Brunt, 1992).

Potyvirus RNA genomes have a protein covalently linked at their 5' extremity (viral genome-linked protein, VPg) and a 3' poly(A) tail. They contain one long open reading frame (ORF) coding for a large polyprotein (Shukla et al., 1994). Brome streak mosaic rymovirus (BrSMV) has a similar genome organization (Götz and Maiss, 1995). The genome organization of the large RNA1 of the bipartite bymovirus BaYMV corresponds to the 3'-three quarters of the potyvirus and rymovirus genomes, whereas the smaller bymovirus RNA2 molecule codes for a polyprotein which has no counterpart in either genus (Kashizaki et al., 1990, 1991).

The family *Potyviridae* also contains some unassigned viruses, among which is the whitefly transmitted sweetpotato mild mottle virus (SPMMV). SPMMV was temporarily and tentatively assigned to a suggested distinct genus, *Ipomovirus* (Barnett, 1991; Shukla et al., 1994), awaiting analyses of genome organization and sequence data. SPMMV was isolated in East Africa from sweetpotatoes showing leaf mottling, veinal chlorosis, dwarfing and poor growth (Hollings et al., 1976). It has not been adequately characterized so far, but the morphology of its virions and cytoplasmic inclusions are similar to those of potyviruses (Moyer and Salazar, 1989). Virus particles are filamentous and contain one species of coat protein of molecular mass 37.7 kDa (Hollings et al., 1976). Recently, we reported the sequence of ≈ 2100 nucleotides derived from the 3' end of the SPMMV genome by family-specific RT-PCR primers (Colinet et al., 1996). Limited sequence identity in the coat protein core with other members of the family *Potyviridae* suggested that SPMMV should be assigned to a distinct genus (Colinet et al., 1996).

To establish similarities, in terms of genome organization and evolutionary relationships, between the whitefly transmitted SPMMV and other members of the family *Potyviridae*, and as a first

step towards genetic manipulation, we undertook the cloning and sequencing of SPMMV RNA. To our knowledge, this is the first report of the complete nucleotide sequence of a member of the family *Potyviridae* transmitted by the whitefly *Bemisia tabaci*.

2. Material and methods

2.1. Virus isolates and RNA purification

SPMMV was received from Dr Salazar (International Center of Potato, Lima, Peru) and propagated in *Nicotiana benthamiana*. Virions were partially purified from symptomatic leaves of *N. benthamiana* following the procedure of Hollings et al. (1976) and virus RNA was further isolated using TriPure Isolation Reagent (Boehringer).

2.2. cDNA synthesis and PCR amplification

Before PCR amplification with the degenerate POT5 (5'-GCAGGATCCAAYATHATHGARA-ATGG-3') and specific MVPOL2 (5'-ATTGATCCCAACTGTCCATG-3') primers, oligo(dT) primed single-stranded cDNA template was synthesized using the Superscript Preamplification System (Gibco BRL). After denaturation for 5 min at 94°C, the following thermal cycling scheme (Biometra cycler) was used for 35 cycles: template denaturation at 94°C for 30 s, annealing at 50°C (cycles 1–5) or 56°C (cycles 6–35) for 1 min and elongation at 72°C for 2 min. The amplified fragment was cloned into pCRTMII using the TA Cloning Kit from InVitroGen.

The RACE method (Frohman, 1990) was used to clone four cDNA fragments corresponding to the 5' terminal half of the genome of SPMMV. SPMMV-specific oligonucleotides primed cDNAs and further amplification of target cDNAs were performed using the Gibco BRL 5' RACE System for Rapid Amplification of cDNA Ends as described by the manufacturer. Amplified fragments were cloned into pCRTMII using the TA Cloning Kit from InVitroGen.

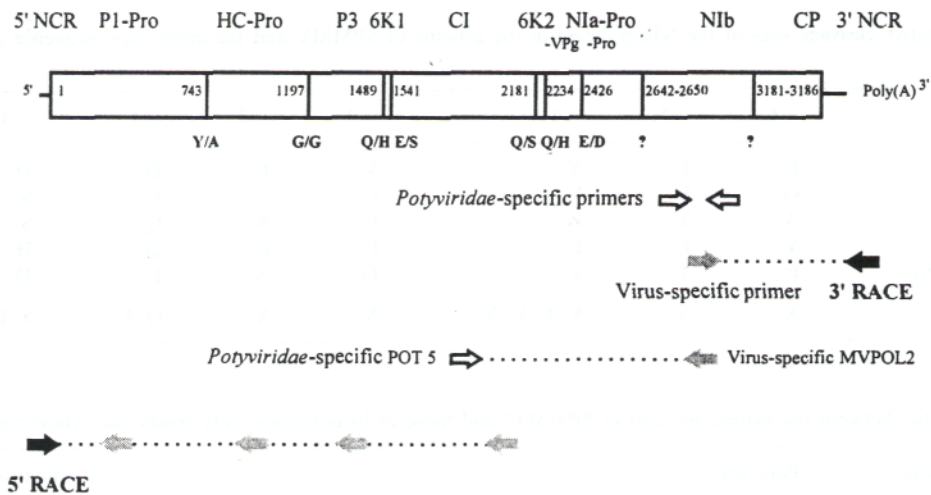


Fig. 1. Tentative genetic map of SPMMV genome showing the relative positions of the *Potyviridae*-specific primers (open arrows) used previously (Colinet et al., 1996), the degenerate primer POT5 (open arrow) based on a selected *Potyviridae* conserved region, the specific SPMMV primers (gray arrows) designed to clone the complete SPMMV genome and the RACE adapter primers (black arrows). Mature proteins and the dipeptide and location of the putative cleavage sites are marked. P1-Pro, first protein/protease; HC-Pro, helper component protease; P3, third protein; 6K1, first 6K peptide; CI, cytoplasmic inclusion protein; 6K2, second 6K peptide; N1a-Pro, nuclear inclusion a protein/protease; Nib, nuclear inclusion b protein (RNA polymerase, CP: coat protein.

2.3. DNA sequencing and analysis

Before sequencing, a fragment corresponding approximately to the 5' three-quarter part of the genome was amplified using MVNC1 (5'-CTATCATGGTGTCAGC-AAACGCAC-3') and MVPOL3 (5'-CCATGGTCCTTGCAAATGACACTT-3') primers following MVPOL2-primed single stranded-cDNA synthesis. The Boehringer ExpandTM Reverse Transcriptase and ExpandTM Long Template PCR Systems were used according to the manufacturers instructions. The following thermal cycling scheme was used for 35 cycles (MJ Research Cycler): template denaturation at 94°C for 10 s, annealing at 62°C for 30 s and elongation at 72°C for 6 min (cycles 25–35: +20 s at each cycle). The amplified fragment was excised from a 1% agarose gel, eluted with the QIAEX Gel Extraction Kit from QIAGEN and cloned into pCRTMII using the TA Cloning Kit from InVitroGen.

A mixture of ten independent PCR products was fragmented by sonication and the resulting 500 bp to 1 kbp fragments were excised from a 1% agarose gel and eluted as above. The DNA

extremities were repaired (Sambrook et al., 1989) and the blunt-ended fragments cloned into pZErOTM-1 vector (InVitroGen) digested with *EcoRV*. Transformed bacterial colonies were screened with digoxigenin-labeled amplified fragment. Digoxigenin labeling and hybridization were performed using the DIG DNA Labeling and DIG Luminescent Detection Kits from Boehringer according to the manufacturers instructions. Double-stranded DNA sequencing by the dideoxy chain termination method was performed, using either T7 DNA polymerase (Pharmacia) or FidelityTM DNA Sequencing System (Appligene) according to the manufacturers instructions.

Sequence analyses employed DNASIS and the University of Wisconsin Genetics Computer Group (GCG) sequence analyses software package version 8.0 (Devereux et al., 1984). Sequence comparisons were performed with Genbank and EMBL databases. Multiple sequence alignments were done using CLUSTAL V (Higgins and Sharp, 1989; Higgins et al., 1991), which calculated a phylogenetic neighbor-joining tree. Group-

Table 1

Five of the putative cleavage sites of the NIa protease in the genome of SPMMV and the amino acid sequence adjacent to the proposed sites

	-6	-5	-4	-3	-2	-1	+1
P3/6K1	E	L	Y	V	E	Q	H
6K1/VI	G	S	V	G	I	E	S
CI/6K2	A	Y	L	E	S	Q	S
6K2/NIa	Y	E	F	I	E	Q	H
NIa-VPg/NIa-Pro	P	T	V	D	A	E	D
Consensus	X	X	V, F, L, Y	X	X	Q, E	S, D, H

Table 2

Sequence identities between the mature proteins of SPMMV^a and those of bymoviruses, potyviruses and rymoviruses

	Bymovirus		Potyvirus					Rymovirus				
	BaYMV	BaMMV	BCMV	JGMV	PepMoV	PPV	PSbMV	PVY	TEV	TuMV	TVMV	BrSMV
P1			<10	15	11	15	<10	11	10	11	12	21
HC			19	18	14	18	19	18	23	19	21	14
P3	<10	<10	11	13	11	15	11	16	16	14	14	18
6K1	20	<10	23	13	21	19	25	20	23	17	17	25
CI	22	22	27	27	28	28	28	29	28	28	27	30
6K2	<10	<10	13	<10	20	13	19	13	19	15	13	24
NIa	13	16	22	22	20	22	21	20	23	23	20	30
NIb	32	34	40	42	36	40	40	39	43	39	39	41

Values are expressed as a percent.

^aWith the exception of the coat protein.

ings were considered significantly supported when they were found in at least 95% of 1000 bootstrap replicates. The following members of the family *Potyviridae* were used for sequence comparisons: *Bymovirus*: BaYMV (Kashiwazaki et al., 1990) and barley mild mosaic virus (BaMMV; Meyer and Dessens, 1996), *Rymovirus*: BrSMV (Götz and Maiss, 1995), *Potyvirus*: bean common mosaic virus (BCMV; Fang et al., 1995), johnsongrass mosaic virus (JGMV; Gough and Shukla, 1993), pepper mottle virus (PepMoV; Vance et al., 1992), plum pox virus (PPV; Lain et al., 1989), pea seedborne mosaic virus (PSbMV; Johansen et al., 1991), PVY (Robaglia et al., 1989), tobacco etch virus (TEV; Allison et al., 1986), turnip mosaic virus (TuMV; Nicolas and Laliberté, 1992) and tobacco vein mottling virus (TVMV; Domier et al., 1986).

3. Results and discussion

3.1. Cloning of SPMMV RNA genome

SPMMV has shown itself to be very recalcitrant, not only to purification but also to molecular characterization. Although no sequence information was available for viruses taxonomically related to SPMMV, the particular problems associated with this virus were overcome by designing broad spectrum PCR primers derived from conserved regions among members of the family *Potyviridae* (Colinet et al., 1996) and using the RACE procedure (Frohman, 1990). The strategy employed to amplify the complete SPMMV genomic RNA is shown in Fig. 1. A conserved region, ATNIIENG, was defined in the cytoplasmic inclusion protein and POT5, a *Potyviridae*-

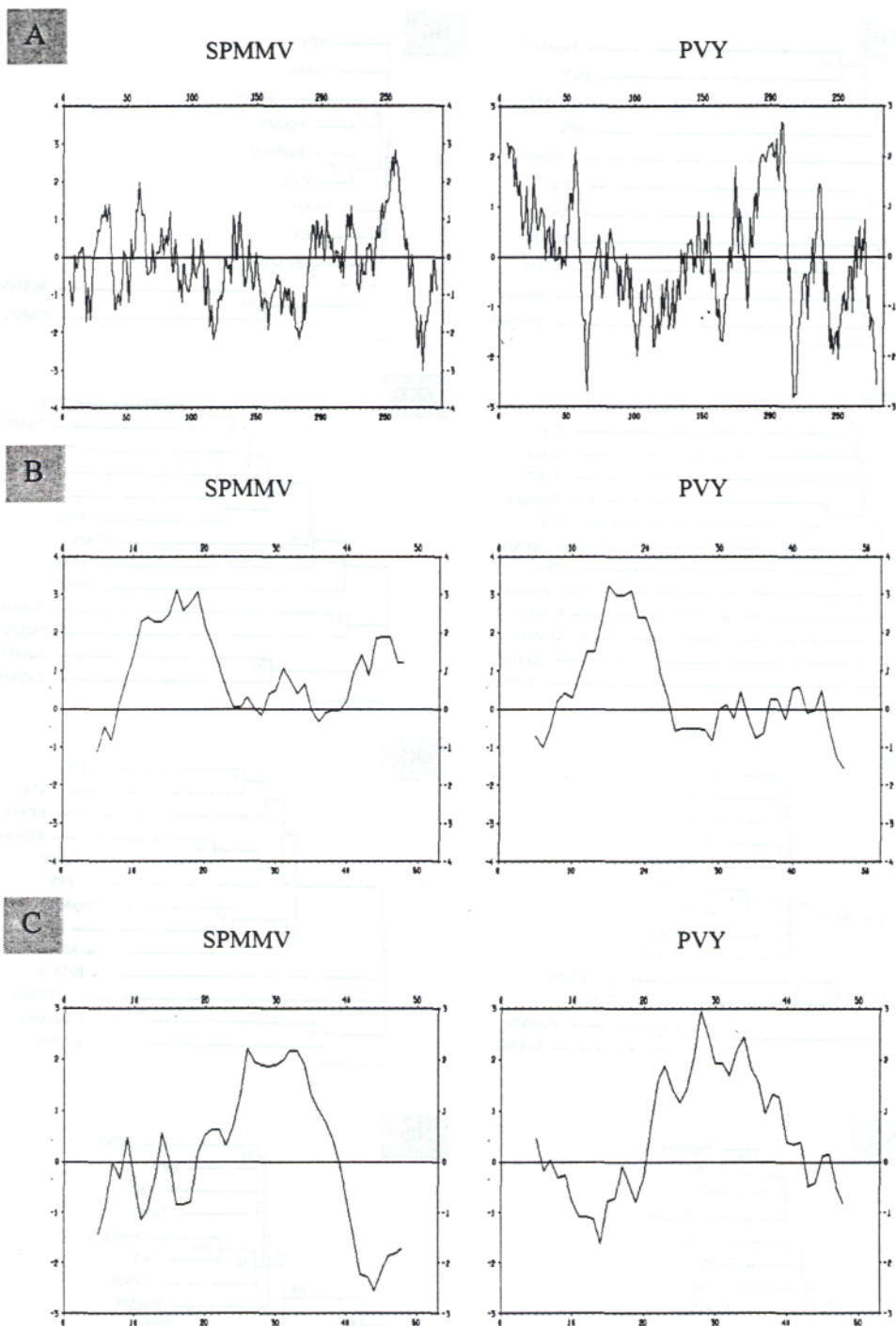


Fig. 2. Kyte-Doolittle plots of SPMMV and PVY proteins. (A) P3. (B) 6K1. (C) 6K2. X-axis, residue position; Y-axis, hydrophathy. Hydrophobic domains are above the X-axis, hydrophilic domains are underneath. The profiles were calculated using pepwindow and a window size of 9.

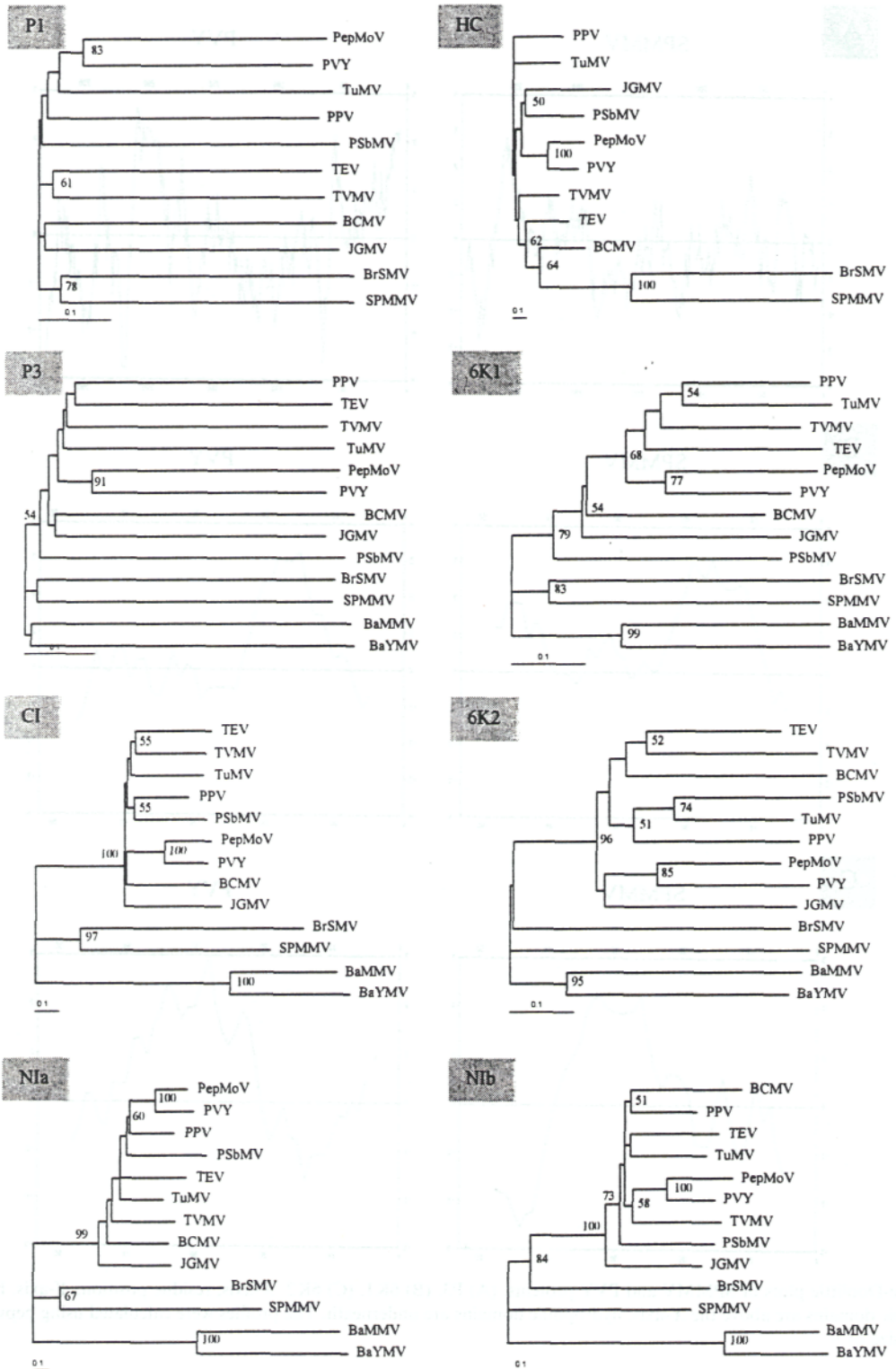


Fig. 3.

specific 26-mer oligonucleotide with a 36-fold degeneracy, was designed. A large 3056 bp fragment was amplified using the *Potyviridae*-specific primer POT5 together with the SPMMV-specific primer MVPOL2 designed from the known sequence of the 3' terminal part of the SPMMV genome. The approximate 5' half of the SPMMV genome was cloned using the RACE procedure (Frohman, 1990). The sequence of the complete RNA genome of SPMMV was then determined.

3.2. Genome organization

The complete nucleotide sequence of SPMMV consists of 10 818 nucleotides excluding the 3'-terminal poly(A) tail. SPMMV RNA is the longest among members of the family *Potyviridae*. Computer analysis of the RNA revealed one large ORF encoding a protein of 3456 amino acids. A 5' non-coding region of 139 nucleotides preceding the ORF was found to be rich in A and U residues (68.6%). The first AUG, at nucleotides 140–142 in-frame with the ORF, was considered to be the initiation codon. A second AUG, at positions 305–308 in frame with the ORF, is located further from the 5' end than that of other members of the family *Potyviridae*, suggesting that is not an initiation codon. Moreover, base composition between the AUGs is rich in G and C residues and is not typical for the 5' non-coding region of potyviruses. The termination codon UAA is located at positions 10 508–10 510.

A tentative map of the polyprotein of SPMMV is shown in Fig. 1, including the predicted cleavage sites and functional proteins. A comparison with potyviruses and a rymovirus showed that SPMMV has a similar genome organization, whereas the ORF located on RNA2 of the bymovirus genome differs considerably from the 5' end of the potyvirus genome. Three proteinases are involved in the complete processing of the potyviral protein. The NIa-Pro proteinase is responsible for several of the polyprotein processing

events (Shukla et al., 1994). A seven amino acid block is sufficient to define the cleavage site recognized by NIa-Pro which usually cleaves at V-2X-Q/(A, E, G, S, T or V) sequences (Shukla et al., 1994). Virus-specific variations may include hydrophobic residues other than valine at -4 position, glutamic acid instead of glutamine at -1 position and residues other than A, E, G, S, T and V at +1 position. Based on these findings, five potential NIa-Pro proteinase heptapeptide cleavage sites were found in the SPMMV polyprotein (Table 1). The sequences of these cleavage sites agree with consensus potyviral sequences though unusual dipeptides are proposed between P3 and 6K1, 6K2 and NIa-Pro and VPg and Pro domains of NIa-Pro. From these, the two first cleavage sites contain glutamine at -1, but atypically contain histidine at position +1, while the third is unusual in having glutamic acid instead of glutamine at position -1 and aspartic acid at position +1. Residues other than A, S, G, T or V at position -1 were also found in PepMoV (Vance et al., 1992) and the rymovirus BrSMV (Götz and Maiss, 1995), for example. Exchange of valine for tyrosine, leucine or phenylalanine at -4 position is unusual but differences in this position have been reported (Götz and Maiss, 1995). The cleavage sites between NIa-Pro and NIb and between NIb and the coat protein could not be determined precisely. On the basis of sequence alignments, their location is proposed between positions 2642 and 2650 and positions 3181 and 3186, respectively.

It has been shown that complete processing of the potyviral polyprotein also requires the proteolytic activity of HC-Pro (Carrington et al., 1989) and P1-Pro (Mavankal and Rhoads, 1991; Verchot et al., 1991). The consensus HC-Pro cleavage sequence is YXVG/G (Carrington and Herndon, 1992). A similar sequence (YKVG/G) is found in the polyprotein of SPMMV at positions 1191–1197. The precise location of the P1-Pro cleavage site has now been established (Mavankal and

Fig. 3. Neighbor-joining trees produced using CLUSTAL V to demonstrate the taxonomic relationships between the mature proteins of SPMMV, with the exception of the coat protein and those of members of the *Potyviridae* (as used in Table 2). Horizontal distances are proportional to percent divergence of tip species from nodes (scale given), vertical distances are arbitrary. Numbers on the branches are the results from bootstrapping using CLUSTAL V for 1000 replicates. Only bootstrap values > 50% are shown.

Rhoads, 1991) and the conserved cleavage site motif (I,M)-X-(Q,H)-(F,Y)/S has been proposed (Shukla et al., 1994). The dipeptide Y/A is found at positions 742–743 and the cleavage site surrounding sequence (IQFY/A) is similar to the consensus motif.

3.3. Comparison of proposed SPMMV mature proteins with those of other members of the family *Potyviridae*

Based on the presumptive location of cleavage sites and on sequence alignments, nine mature proteins are predicted for SPMMV (Fig. 1). The percentage identities of the SPMMV proteins, with the exception of the coat protein, with corresponding proteins of other members of the family *Potyviridae* are listed in Table 2.

The first protein (P1-Pro) generated from the N-terminus of the potyviral polyprotein differs most from the corresponding proteins of other *Potyviridae*. The SPMMV P1-Pro protein shares limited identity (from < 10 to 21%) with those of potyviruses and the rymovirus BrSMV, which are much smaller. The conserved catalytic triad composed of amino acids histidine, aspartic acid and serine representing the active site of the P1-Pro proteinase (Verchot et al., 1991, 1992) is located in the P1-Pro of SPMMV at positions 640, 651 and 689, respectively. Moreover, the conserved active site sequence GXSG found in all potyviruses (Shukla et al., 1994) and the rymovirus BrSMV is also present in the P1-Pro of SPMMV at positions 687–690.

The SPMMV HC-Pro-like protein showed 14–23% identity with HC proteins of four potyviruses and the HC-Pro-equivalent of the rymovirus BrSMV, whereas sequence identity ranged from 38 to 65% between the potyviruses and was \leq 20% between BrSMV and the potyviruses. The conserved motif GYCY-72X-H described for HC-Pro (Oh and Carrington, 1989) is present in the corresponding protein of SPMMV, with cysteine and histidine, essential for the proteinase active site of HC-Pro, at positions 1083 and 1155, respectively. The distance between C and H is 73 rather than 72 amino acids. The cysteine cluster C-8X-C-13X (or 7X-C-5X)-C-4X-C-2X-C in the

N-terminal third of the HC-Pro protein of potyviruses is similar to the zinc finger motifs of several nucleic acid-binding proteins (Robaglia et al., 1989; Shukla et al., 1994). This motif is probably involved in aphid transmission of potyviruses and/or cell-to-cell movement (Atreya et al., 1992). A reduced cluster of C₇₉₄-2X-C₇₉₇ was found in the polyprotein of SPMMV, resembling in this respect, the experiences with BrSMV (Götz and Maiss, 1995).

Sequence identity in the third protein (P3) between SPMMV and other members of the family *Potyviridae* was < 20%. The SPMMV P3 protein contains strong hydrophobic domains (Fig. 2), as has been found for other members of the family *Potyviridae* (Meyer and Dessens, 1996).

The two SPMMV 6K peptides showed < 10–25% identity with other members of the family *Potyviridae* (Table 2), whereas sequence identity was > 25% between distinct potyviruses. The 6K1 and 6K2 proteins of SPMMV have highly hydrophobic hydropathy profiles very similar to those of other members of the family *Potyviridae* (Fig. 2).

The nucleotide binding consensus sequence motif G-2X-G-X-GKST is present in the cytoplasmic inclusion (CI) proteins of all sequenced members of the family *Potyviridae* and is almost perfectly conserved in the polyprotein of SPMMV at positions 1631–1638. SPMMV CI shared 27–29% sequence identity with potyviruses, 30% identity with BrSMV and 22% identity with BaYMV and BaMMV (Table 2). For comparison, sequence identity between either bymovirus or rymovirus and other members of the family *Potyviridae* was < 30%, whereas potyvirus CIs showed generally > 50% identity and bymovirus CIs shared 44% identity.

The NIa-Pro protein of SPMMV, such as that of all members of the family *Potyviridae* sequenced to date, can be divided into two domains, VPg and Pro. The sequence identity with other members of the family *Potyviridae* ranged from 13 (with BaYMV) to 30% (with BrSMV), in contrast to the 40 to 67% found between potyviruses and the 37% identity shared between BaYMV and BaMMV. A conserved triad of the amino acids histidine, aspartic acid and cysteine in the consen-

sus motif H-38X-D-71X-GDCG represents the active site of the NIa-Pro Proteinase (Dougherty et al., 1989). The amino acid sequence of SPMMV includes these three amino acids, though the distance between the amino acids histidine and aspartic acid is slightly different (H₂₄₅₇-36X-D₂₄₉₄-71X-GDC₂₅₆₆G).

The SPMMV nuclear inclusion b (NIB) protein shared 32–43% with that of other members of the family *Potyviridae* (Table 2), indicating that NIB is the most conserved protein among members of the family *Potyviridae*. The conserved motif (S,T)G-3X-T-3X-N(S,T)-(18-37)X-GDD, proposed to be the active site of the RNA-dependent RNA polymerase (Domier et al., 1987; Koonin, 1991), is perfectly conserved in the sequence of SPMMV at positions 2959–2998.

In conclusion, sequence analyses revealed strong similarities between the whitefly transmitted SPMMV and members of the family *Potyviridae*, in terms of structure, genome organization and in the presence of consensus motifs. However, the zinc-finger motif, which is suspected to be involved in aphid transmission of potyviruses, is not complete in SPMMV. Furthermore, it is noteworthy that the motif DAG, involved in aphid transmission of potyviruses, is not found in the coat protein of SPMMV (Colinet et al., 1996). Mutagenesis studies would help to determine if such modifications in SPMMV HC-Pro and CP reflect its transmission by a different vector. The complete cloning and sequencing should permit a better understanding of the virus life cycle at the molecular level, following the production of a full-length infectious cDNA transcript and the introduction of defined mutations in the SPMMV genome.

Limited sequence identity in the coat protein core with that of potyviruses, bymoviruses, rymoviruses (Colinet et al., 1996) and members of the recently proposed genus *Macluravirus* (Badge et al., 1997) suggested that SPMMV should be assigned to a distinct genus of the family *Potyviridae*. This is consistent with the low levels of identity observed between SPMMV and other members of the family *Potyviridae* (Table 2) and the phylogenetic trees (Fig. 3) obtained for the other mature proteins. The eight potyviruses have

different degrees of relationship but they are more closely related to each other than to BaYMV, BaMMV, BrSMV or SPMMV. Similarly, the two bymoviruses are more closely related to each other than to any other member of the family *Potyviridae*. The alignment of the eight mature proteins showed less similarity between SPMMV and other members of the family *Potyviridae* than between potyviruses or bymoviruses. The only exception is P3, which displayed less sequence identity between BaYMV and BaMMV than between SPMMV and BrSMV. SPMMV seems somewhat closer to the rymovirus BrSMV (Fig. 3). However, the levels of identity are too small to warrant the grouping of both viruses in the same genus. It therefore seems appropriate to allocate the whitefly transmitted SPMMV to a distinct genus, as suggested previously and named *Ipomovirus*.

Acknowledgements

We thank Dr Salazar (International Centre of Potato, Lima, Peru) for supplying the SPMMV isolate. This work was financially supported by the 'Fonds National de la Recherche Scientifique', Bruxelles, Belgium.

References

- Allison, R., Johnston, R.E., Dougherty, W.G., 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: Evidence for the synthesis of a single polyprotein. *Virology* 154, 9–20.
- Atreya, C.D., Atreya, P.L., Thornbury, D.W., Pirone, T.P., 1992. Site-directed mutations in the potyvirus HC-Pro gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. *Virology* 191, 106–111.
- Badge, J., Robinson, D., Brunt, A., Foster, G., 1997. 3'-Terminal sequences 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.
- Barnett, O.W., 1991. *Potyviridae*, a proposed family of plant viruses. *Arch. Virol.* 118, 139–141.
- Brunt, A.A., 1992. The general properties of potyviruses. In: Barnett, O.W. (Ed.), *Potyvirus Taxonomy Arch. Virol. Suppl* 5. Springer, Wien, New York, pp. 3–16.

- Carrington, J.C., Freed, D.D., Sanders, T.C., 1989. Autocatalytic processing of the potyviral helper component proteinase in *Escherichia coli* and in vitro. *J. Virol.* 63, 4459–4463.
- Carrington, J.C., Herndon, K.L., 1992. Characterization of the potyviral HC-Pro autoproteolytic cleavage site. *Virology* 187, 308–315.
- Colinet, D., Kummert, J., Lepoivre, P., 1996. Molecular evidence that the whitefly transmitted sweetpotato mild mottle virus belongs to a distinct genus of the *Potyviridae*. *Arch. Virol.* 141, 125–135.
- Devereux, J., Haeberli, P., Smithies, O., 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.
- Domier, L.L., Franklin, K.M., Shahabuddin, M., Hellmann, G.M., Overmeyer, J.M., Hiremath, S.T., Siaw, M.F.E., Lomonosoff, G.P., Shaw, J.G., Rhoads, R.E., 1986. The nucleotide sequence of tobacco vein mottling virus RNA. *Nucleic Acids Res.* 14, 5417–5430.
- Domier, L.L., Shaw, J.G., Rhoads, R.E., 1987. Potyviral proteins share amino acid sequence homology with picorna-, como-, and caulimoviral proteins. *Virology* 158, 20–27.
- Dougherty, W.G., Parks, T.D., Cary, S.M., Bazan, J.F., Fletterick, R.J., 1989. Characterization of the catalytic residues of the tobacco etch virus 49 kDa proteinase. *Virology* 172, 302–310.
- Fang, G.W., Allison, R.F., Zambolim, E.M., Maxwell, D.P., Gilbertson, R.L., 1995. The complete nucleotide sequence and genome organization of bean common mosaic virus (NL3 strain). *Virus Res.* 39, 13–23.
- Frohman, M.A., 1990. RACE: Rapid amplification of cDNA ends. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), *PCR Protocols: A Guide to Methods and Applications*. Academic Press, San Diego, pp. 28–38.
- Gough, K.H., Shukla, D.D., 1993. Nucleotide sequence of johnsongrass mosaic potyvirus genomic RNA. *Intervirology* 36, 181–192.
- Götz, R., Maiss, E., 1995. The complete nucleotide sequence and genome organization of the mite transmitted brome streak mosaic rymovirus in comparison with those of potyviruses. *J. Gen. Virol.* 76, 2035–2042.
- Higgins, D., Sharp, P., 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.* 5, 151–153.
- Higgins, D., Bleasly, A., Fuchs, R., 1991. ClustalV: Improved software for multiple sequence alignments. *CABIOS* 8, 189–191.
- Hollings, M., Stone, O.M., Bock, K.R., 1976. Purification and properties of sweetpotato mild mottle, a whitefly borne virus from sweetpotato in East Africa. *Ann. Appl. Biol.* 82, 511–528.
- Johansen, E., Rasmussen, O.F., Heide, M., Borkhardt, B., 1991. The complete nucleotide sequence of pea seed-borne mosaic virus RNA. *J. Gen. Virol.* 72, 2625–2632.
- Kashiwazaki, S., Minobe, Y., Minobe, T., Hibino, H., 1990. Nucleotide sequence of barley yellow mosaic virus RNA 1: A close evolutionary relationship with potyviruses. *J. Gen. Virol.* 71, 2781–2790.
- Kashiwazaki, S., Minobe, Y., Hibino, H., 1991. Nucleotide sequence of barley yellow mosaic virus RNA 2. *J. Gen. Virol.* 72, 989–993.
- Koonin, E.V., 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* 72, 2197–2206.
- Lain, S., Riechmann, J.L., Garcia, J.A., 1989. The complete nucleotide sequence of plum pox potyvirus RNA. *Virus Res.* 13, 157–172.
- Mavankal, G., Rhoads, R.E., 1991. In vitro cleavage at or near the N-terminus of the helper component protein in the tobacco vein mottling virus polyprotein. *Virology* 185, 721–731.
- Meyer, M., Dessens, J.T., 1996. The complete nucleotide sequence of barley mild mosaic virus RNA1 and its relationship with other members of the *Potyviridae*. *Virology* 219, 268–273.
- Moyer, J.W., Salazar, L.F., 1989. Virus and virus-like diseases of sweetpotato. *Plant Dis.* 73, 451–455.
- Nicolas, O., Laliberté, J.F., 1992. The complete nucleotide sequence of turnip mosaic potyvirus RNA. *J. Gen. Virol.* 73, 2785–2793.
- Oh, C.S., Carrington, J.C., 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. *Virology* 173, 692–699.
- Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manificier, S., Casse-Delbart, F., 1989. Nucleotide sequence of potato virus (N strain) genomic RNA. *J. Gen. Virol.* 70, 935–947.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, New York.
- Shukla, D.D., Ward, C.W., Brunt, A.A., 1994. *The Potyviridae*. CAB International, Wallingford.
- Vance, V.B., Moore, D., Turpen, T.H., Bracker, A., Hollowell, V.C., 1992. The complete nucleotide sequence of pepper mottle virus genomic RNA: Comparison of the encoded polyprotein with those of other sequenced potyviruses. *Virology* 191, 19–30.
- Verchot, J.M., Koonin, E.V., Carrington, J.C., 1991. The 35 K protein from the N-terminus of a potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* 185, 527–535.
- Verchot, J.M., Herndon, K.L., Carrington, J.C., 1992. Mutational analyses of the tobacco etch potyviral 35 K proteinase: Identification of essential residues and requirements for autoproteolysis. *Virology* 190, 298–306.