HAL
open science

Complete sequence of an isolate of snake melon asteroid mosaic virus confirms that it is a member of a distinct sobemovirus species

Cecile Desbiez, Eric Verdin, Hervé Lecoq

To cite this version:

Cecile Desbiez, Eric Verdin, Hervé Lecoq. Complete sequence of an isolate of snake melon asteroid mosaic virus confirms that it is a member of a distinct sobemovirus species. Archives of Virology, 2021, 166, pp.2311-2313. 10.1007/s00705-021-05109-8 . hal-03244292

HAL Id: hal-03244292

https://hal.inrae.fr/hal-03244292

Submitted on 9 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Complete sequence of an isolate of snake melon asteroid mosaic virus confirms that it is a member of a distinct sobemovirus species

C. Desbiez, E. Verdin, H. Lecoq

INRAE, UR0407 Pathologie Végétale, F-84140 Montfavet, France

Corresponding author: C. Desbiez
Email: cecile.desbiez@inrae.fr
Tel: 33 (0)4 32 72 28 07
Fax: 33 (0)4 32 72 28 42
ORCID number 0000-0001-8691-7423

Abstract

A virus tentatively named snake melon asteroid mosaic virus (SMAMV) was found in Sudan in cucurbit crops (10% of 600 samples) between 1992 and 2003. Biological and cytological properties, and sequence data on a 345-nt fragment suggested that SMAMV was a member of the genus *Sobemovirus*. However no complete sequence had been obtained and the relation between SMAMV and the acknowledged sobemoviruses had not been ascertained. In this work, we obtained the full-length sequence of a SMAMV isolate. The sequence was 4225-nt long, with a typical sobemovirus genetic organization. Sequence identity with other sobemoviruses was below 50%, both for the full-length genome and for individual proteins. These data confirm that SMAMV corresponds to a novel sobemovirus species.

24 The virus characterized in this study (Su-95-67) was obtained from a naturally infected snake melon
25 (*Cucumis melo* var. *flexuosus*) plant collected in December 1995 near Abou Rakham, Sudan [5]. The
26 sample was stored as dried leaf material on calcium chloride. One mg dry material was used for total
27 RNA extraction with PureLink Plant RNA reagent (Thermo Fisher Scientific, Waltham, NA). The total
28 RNA sample resuspended in 20 µL nuclease-free distilled water (Invitrogen, France) was sent to Fasteris
29 SA (Plan-les-Ouates, Switzerland) for small RNA library construction and deep sequencing. Library
30 preparation was performed using the Illumina TruSeq® Small RNA Sample Prep Kit before Single-read,
31 1 x 50 bp Illumina HiSeq High-Output sequencing. Trimmed sequences (21-24 nt) were assembled *de*
32 *novo* with CLC Genomics Workbench 7 (CLC Bio, Aarhus, Denmark) as described [12]. Blastn and blastx
33 searches were performed against GenBank. Five contigs presenting similarities with sobemoviruses
34 were used as templates for iterative mapping of reads to contigs with Geneious 8.1.9 (Biomatters, New
35 Zealand). To confirm sequences of low coverage areas, RT-PCR was performed with primers SMAMV-
36 1360-F+SMAMV-1860-R (5'-GTTCTTGTGGAGGTTCCTCA-3' + 5'-GCCCTCATAAGCGGACCA-3' ; positions
37 1356-1866) and SMAMV-10-F+SMAMV-880-R (5'-AAGCCGACTAAGAATAGTTAGC-3' + 5'-
38 CCACAGCCATATCACTATCGC-3' ; positions 7-885) ; amplified fragments were sent for Sanger
39 sequencing to GenoScreen (Lille, France). No contig showed similarity to satellites associated with
40 some sobemoviruses [1,9]. To confirm the sequences of the 5' and 3' ends, RNA was extracted from
41 purified virus [5] with TRI-reagent®. Extracted RNA was self-ligated (method 1) or ligated with a
42 nonviral primer (method 2), using T4 RNA ligase (NEB, Ipswich, USA) before RT-PCR with specific
43 primers SMAMV-140-R (5'-CAACTCAATCGATAATCTAATGGG-3') and SMAMV-4120-F (5'-
44 GACTATTCAAGAATTGTCCGT-3'), 100 to 150 nt from the expected 5' and 3' ends respectively (method
45 1), and SMAMV-4120-F and a primer complementary to the nonviral primer (method 2). Amplified
46 fragments were purified from agarose gel and sent to Sanger sequencing.
47 The sequence obtained was 4225 nt long (GenBank accession MT989351), similar to known
48 sobemoviruses ([1,9] and Table 1). Sequence identity with sobemoviruses was below 50% (Table 1).

49 The 5' end sequence was ACAAAA, the same as that for several sobemoviruses [9]. As in other
50 sobemoviruses, an ACAAA motif was present at position 3382 in the genome and may constitute the
51 start of a subgenomic RNA allowing the expression of the coat protein [9]. The 5' untranslated region
52 was 120 nt long (44 to 99 nt for other sobemoviruses), and the 3' noncoding region was 85 nt long (44
53 to 283 nt for other sobemoviruses). Secondary structure prediction with RNAfold [8] did not reveal any
54 specific structure for the noncoding regions (data not shown). SMAMV had a typical sobemovirus
55 genetic organization, with five putative open reading frames (ORFs) [9,10]. In the 5' part of the
56 genome, ORF1 (172 aa, 19.8kDa) and the putative ORFx (94 aa, 10.5 kDa) starting with a CTG codon [6]
57 had very low similarity with other viruses (Table 1). The conserved CxxC-(6-50)-CxxC putative zinc-
58 finger motif of sobemovirus P1 [9] was identified as CETC-(46)-CNHC in SMAMV. ORF2a (541 aa, 59.3
59 kDa) contained the putative transmembrane domain, serine protease and viral protein genome-linked
60 (VPg). The consensus protease motif in sobemovirus P2a TxxGxsG was identified as TAPRWSG in
61 SMAMV, and the WAG motif in the VPg [9] was modified to WSA. The large ORF2ab (966 aa, 108.6
62 kDa), as in other sobemoviruses, appears most likely expressed through a (-1) frameshift at a slippery
63 sequence UUUAAAC followed by a stem-loop structure located seven nucleotides downstream [9].
64 ORF3 corresponds to the putative coat protein (251 aa, 27.8 kDa), probably expressed through a
65 subgenomic RNA initiated at the internal motif ACAAA. The RdRp region (ORF2b) showed the highest
66 aa identity with other sobemoviruses (Table 1), but that identity did not exceed 50%. Multiple nt
67 sequence alignments were performed with MAFFT alignment software [4]. Evidence for recombination
68 explored with RDP, GeneConv, Chimaera, BootScan and SiScan included in RDP4 [7] did not reveal
69 recombination for SMAMV. The best substitution method, estimated with MEGA6 [11] for the aligned
70 nt and aa sequences, was applied for maximum-likelihood tree reconstruction (Figure 1 and data not
71 shown). In the complete nt sequence or in aa sequences, SMAMV did not appear particularly related
72 to any other sobemovirus. Indeed, some sobemoviruses present host-related clustering, but the basal
73 branches of the phylogenetic tree, where SMAMV attaches, are poorly resolved ([9] and Figure 1).

Sobemovirus radiation is supposed to have taken place circa 3000 years ago [3]. At that time, watermelon had already been cultivated in Sudan for nearly 2000 years [2].

The criteria for sobemovirus species demarcation are nt sequence identity below 75%, host range, and serological relatedness [1]. SMAMV host range is different from other sobemoviruses [5], and its full-length sequence shares less than 50% identity with other viruses. It can thus be considered as a distinct sobemovirus species. SMAMV has a narrow host range –in experimental conditions, it infected only melon and watermelon [5]- and has been found only in Sudan, but due to its seed transmissibility [5], it is important to avoid its accidental spread through germplasm exchanges.

Figure legends

Figure 1 : Maximum-likelihood unrooted tree obtained for full-length nucleotide sequences of sobemoviruses. Accession numbers for each virus correspond to the isolates listed in Table 1. The scale bar represents a genetic distance of 0.2. Bootstrap values above 60% (500 bootstraps) are indicated for each node.

Declarations

Funding: this study did not receive any specific funding

Ethical statement: This research did not involve any human participants and/or animals.

Conflict of interest: The authors declare that they have no conflict of interest.

References :

1. Truve E, Fargette D (2012) Sobemoviruses. In: King AM, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus Taxonomy. Elsevier/Academic Press, London, pp 1185-1189

2. Chomicki G, Schaefer H, Renner SS (2019) Origin and domestication of Cucurbitaceae crops: insights from phylogenies, genomics and archaeology. New Phytol 226: 1240-1255

99 3. Fargette D, Pinel-Galzi A, Sérémé D, Lacombe S, Hébrard E, Traoré O, Konaté G (2008)
100 Diversification of *Rice yellow mottle virus* and related viruses spans the history of agriculture
101 from the Neolithic to the Present. PLoS Path 4: e1000125
102 4. Katoh K, Misawa K, Kuma K-i, Miyata T (2002) MAFFT: a novel method for rapid multiple
103 sequence alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059-3066
104 5. Lecoq H, Dafalla G, Delecolle B, Wipf-Scheibel C, Desbiez C (2011) Snake melon asteroid
105 mosaic virus, a tentative new member of the genus Sobemovirus infecting cucurbits. Plant
106 Dis 95: 153-157
107 6. Ling R, Pate AE, Carr JP, Firth AE (2013) An essential fifth coding ORF in the sobemoviruses.
108 Virology 446: 397-408
109 7. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of
110 recombination patterns in virus genomes. Virus Evol 1: 1-5
111 8. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH (2004) Incorporating
112 chemical modification constraints into a dynamic programming algorithm for prediction of
113 RNA secondary structure. Proc Natl Acad Sci USA 101: 7287-7292
114 9. Somera M, Sarmiento C, Truve E (2015) Overview on Sobemoviruses and a Proposal for the
115 Creation of the Family Sobemoviridae. Viruses-Basel 7: 3076-3115
116 10. Somera M, Truve E (2017) Complete nucleotide sequence of Solanum nodiflorum mottle
117 virus. Arch Virol 162: 1731-1736
118 11. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary
119 Genetics Analysis version 6.0. Mol Biol Evol 30: 2725-2729
120 12. Verdin E, Wipf-Scheibel C, Gognalons P, Aller F, Jacquemond M, Tepfer M (2017) Sequencing
121 viral siRNAs to identify previously undescribed viruses and viroids in a panel of ornamental
122 plant samples structured as a matrix of pools. Virus Res 241: 19-28
123

124 Table 1 : sequence identity (%) between the SMAMV Su95-67 isolate and sobemoviruses, for the full-

125 length nucleotide sequence and for deduced amino-acid sequences of the different proteins

Accession	Sobemovirus	Genome size (nt)	Identity with SMAMV (%)					
			Complete seq (nt)	P1 (aa)	Px (aa)	P2a (aa)	RdRp (aa)	CP (aa)
NC_017914	Artemisia virus A (ArtVA)	4138	46.15	20.20	23.53	25.69	48.63	21.10
NC_029578	Blueberry shoestring virus (BSSV)	4020	44.23	15.63	16.46	24.80	43.85	23.91
NC_002618	Cocksfoot mottle virus (CfMV)	4082	45.17	22.09	16.44	27.65	46.97	22.17
NC_027123	Cymbidium chlorotic mosaic virus (CyCMV)	4083	48.06	22.50	18.67	28.71	50.48	30.87
NC_011536	Imperata yellow mottle virus (IYMV)	4547	44.00	11.43	17.57	22.48	46.44	21.58
NC_001696	Lucerne transient streak virus (LTSV)	4279	44.69	18.69	11.39	27.38	46.99	23.58
NC_018449	Papaya lethal yellowing virus (PLYV)	4145	46.93	19.66	19.18	26.25	46.53	32.61
NC_001575	Rice yellow mottle virus (RYMV)	4449	44.38	18.46	20.90	24.08	49.71	21.40
NC_027198	Rottboellia yellow mottle virus (RoYMoV)	4194	43.73	17.32	17.95	24.95	46.54	21.92
NC_003747	Ryegrass mottle virus (RGMoV)	4212	43.49	11.76	15.79	25.92	44.53	17.81
NC_002568	Sesbania mosaic virus (SeMV)	4148	47.11	18.03	12.31	25.34	47.40	30.30
NC_033706	Solanum nodiflorum mottle virus (SNMoV)	4297	45.90	18.52	13.04	27.57	44.53	24.89
NC_004060	Southern bean mosaic virus (SBMV)	4132	47.09	12.39	14.04	27.08	47.58	29.87
NC_001625	Southern cowpea mosaic virus (SCPMV)	4193	46.57	14.16	17.33	26.11	47.27	30.00
GQ845002	Sowbane mosaic virus (SoMV)	3983	46.28	13.49	16.22	29.71	47.08	27.85
NC_016033	Soybean yellow common mosaic virus (SYCMV)	4152	46.95	11.97	19.12	25.73	47.03	30.74
NC_004346	Subterranean clover mottle virus (SCMoV)	4258	45.51	23.48	11.69	28.16	45.98	29.78
NC_004553	Turnip rosette virus (TRoV)	4086	45.30	18.33	18.33	25.98	45.54	23.14
NC_014509	Velvet tobacco mottle virus (VTMoV)	4247	46.10	16.90	14.93	29.10	45.42	25.42

126

127

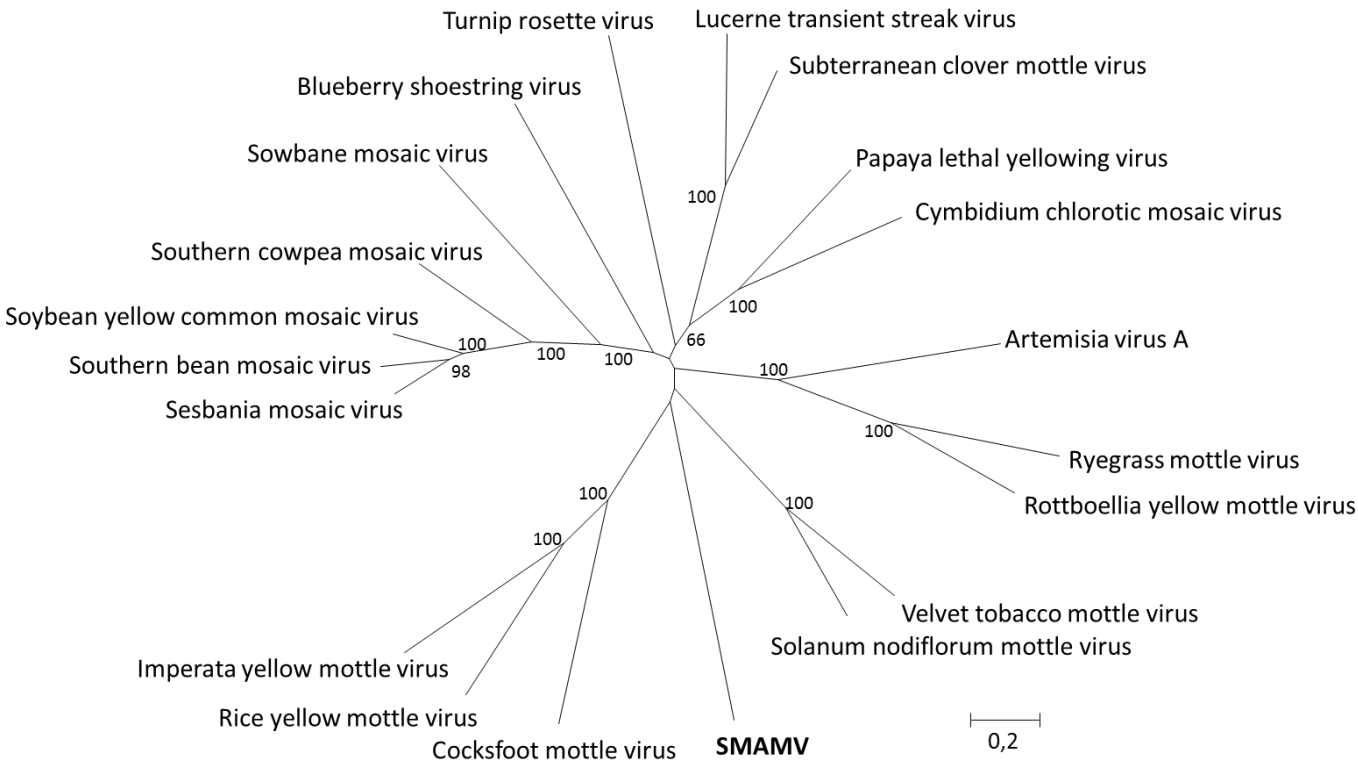
Turnip rosette virus
Lucerne transient streak virus
Subterranean clover mottle virus
Blueberry shoestring virus
Sowbane mosaic virus
Papaya lethal yellowing virus
Cymbidium chlorotic mosaic virus
Southern cowpea mosaic virus
Soybean yellow common mosaic virus
Southern bean mosaic virus
Sesbania mosaic virus
Artemisia virus A
Ryegrass mottle virus
Rottboellia yellow mottle virus
Velvet tobacco mottle virus
Solanum nodiflorum mottle virus
Imperata yellow mottle virus
Rice yellow mottle virus
Cocksfoot mottle virus
SMAMV
0,2
100
100
66
100
100
100
98
100
100
100
100
100
100

Figure 1