



# Complete genome sequence of mimosa mosaic virus, a new sobemovirus infecting *Mimosa sensitiva* L.

Caterynne Melo Kauffmann<sup>1</sup> · Alessandra de Jesus Boari<sup>2</sup> · João Marcos Fagundes Silva<sup>3</sup> · Rosana Blawid<sup>4</sup> · Tatsuya Nagata<sup>1,3</sup>

Received: 11 October 2022 / Accepted: 18 November 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

## Abstract

A new sobemovirus, which we have named “mimosa mosaic virus” (MimMV), was found by high-throughput sequencing and isolated from a mimosa (*Mimosa sensitiva* L.) plant. The genome sequence was confirmed by Sanger sequencing and comprises 4595 nucleotides. Phylogenetic analysis based on the predicted amino acid (aa) sequences of the P2b protein (encoded by ORF2b) and the coat protein showed 52.7% and 31.8% aa sequence identity, respectively, to those of blueberry shoestring virus. The complete genome sequence of MimMV was less than 47% identical to those of other sobemoviruses. These data suggest that MimMV is a member of a new species in the genus *Sobemovirus*, for which the binomial name “*Sobemovirus mimosae*” is proposed.

Mimosa (*Mimosa sensitiva* L.) is a wild plant belonging to the family Fabaceae and subfamily Mimosoideae. In July 2015, we collected mimosa leaf samples showing virus-like symptoms (Fig. 1B) in Santo Antônio do Tauá in the state of Pará. The collected samples were analyzed by high-throughput sequencing (HTS). Samples were initially processed using a viral semipurification procedure [1]. Total RNA was extracted from the semipurified preparation using a ZR Plant RNA MiniPrep Kit (Zymo Research, Irvine, USA). The extracted RNA from mimosa plants was added to a pooled sample as part of a larger effort to find new viruses in several plants. (This pooled sample contained RNA from semipurified virus preparations from cowpea, black pepper, and patchouli in addition to mimosa plants.) Then, rRNA was removed from the pooled sample using a RiboZero kit for plants (Illumina, San Diego, USA), and a cDNA library

was prepared and sequenced using a Novaseq system (Illumina) at MacroGen with 100-bp paired-end reads (MacroGen Inc. Seoul, Korea).

The HTS reads were trimmed, and contigs were obtained by *de novo* assembly using SPAdes 4.2. [2]. The resulting contigs were analyzed with tBlastx against the virus genome database using Geneious v.8.1.9 (Biomatters Ltd, Auckland, New Zealand). Based on the contigs related to sobemovirus sequences, longer contigs were assembled by read mapping in Geneious. The longest contig assembled contained 4,448 nucleotides (nt) aligned with 4,999,563 reads (out of 35,191,002 total reads), with a mean coverage of 105,684.7. The specific primers SobemoR and SobemoF (Supplementary Table S1) were designed to amplify the RNA-dependent RNA polymerase (RdRp) gene region to detect sobemovirus sequences in individual mimosa samples used for the HTS analysis. To sequence the whole genome of the sobemovirus using the Sanger method, cDNA fragments were amplified by RT-PCR using RNA extracted from a single mimosa plant using specific primers (Supplementary Table S1) designed based on the HTS analysis. 5'- and 3'-RACE were performed as described previously [3]. The cDNA fragments were sequenced by the Sanger method.

The complete genome of the sobemovirus, which we have named “mimosa mosaic virus” (MimMV), contains 4595 nt (GenBank accession no. OP456085) as determined by Sanger sequencing. The genomic organization of MimMV is typical of members of the genus *Sobemovirus*

Handling Editor: Ralf Georg Dietzgen.

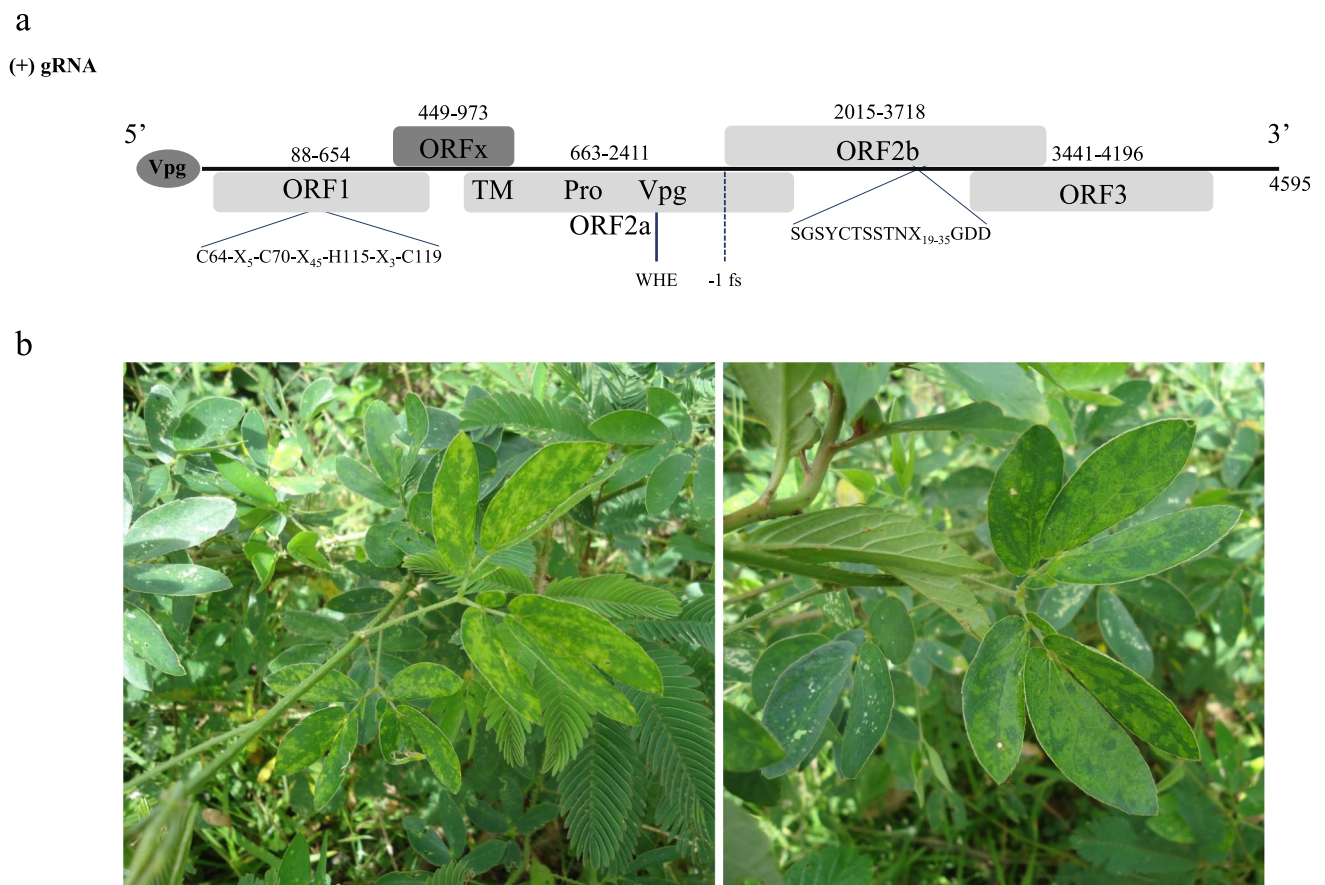
✉ Tatsuya Nagata  
tatsuya@unb.br

<sup>1</sup> Departamento de fitopatologia, Universidade de Brasília, Brasília, Federal District, Brazil

<sup>2</sup> Embrapa Amazônia Oriental, Belém, Pará, Brazil

<sup>3</sup> Departamento de Biologia Molecular, Universidade de Brasília, Brasília, Federal District, Brazil

<sup>4</sup> Departamento de Agronomia, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil



**Fig. 1** (a) Genome organization of mimosa mosaic virus (MimMV). Rectangles represent the relative positions of the open reading frames (ORFs). ORF1, movement protein; ORFx, protein X; TM, transmembrane domain; Pro, serine protease domain; VPg, virus protein, genome-linked; ORF2b, RNA-dependent RNA polymer-

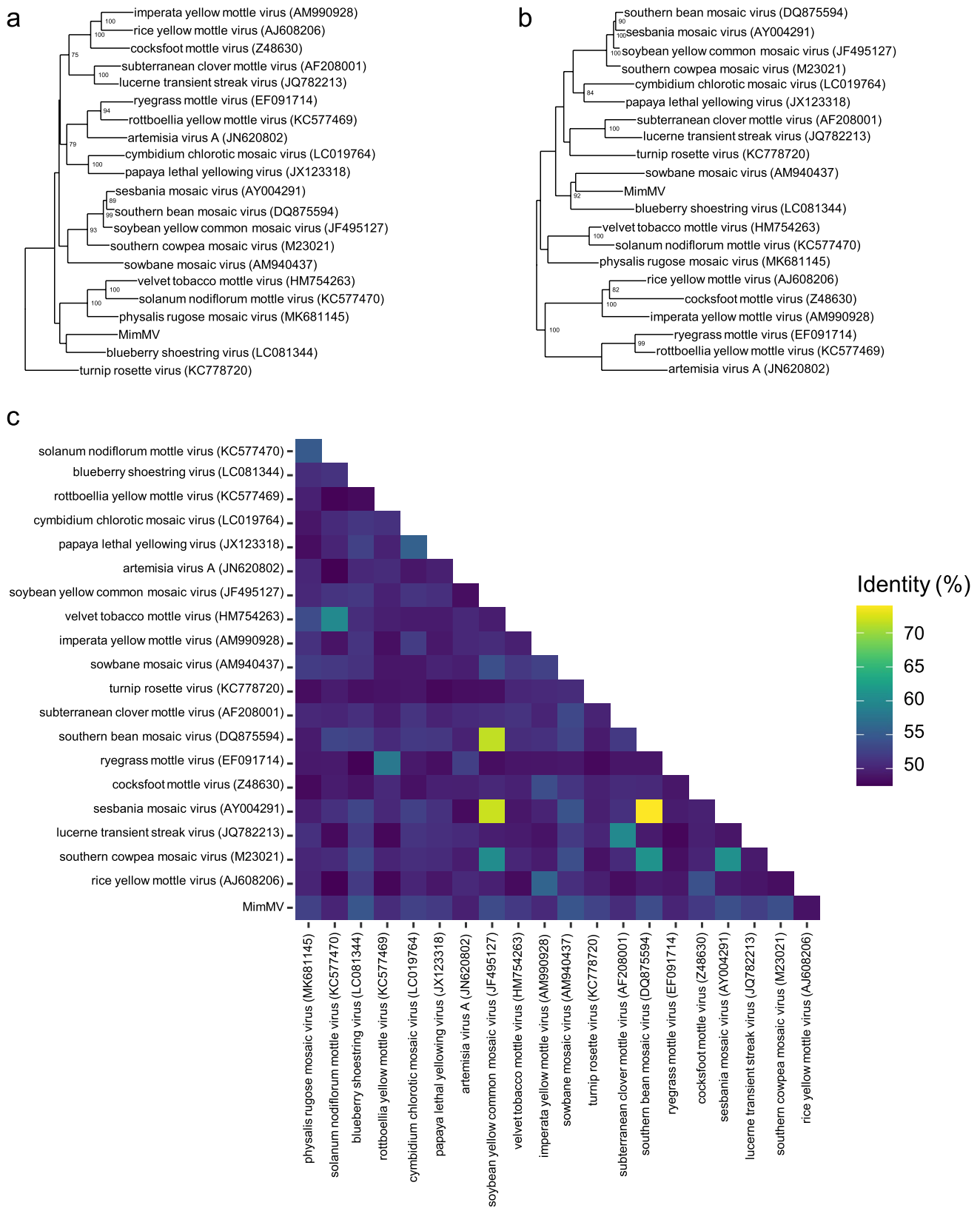
ase (RdRp); ORF3, coat protein. [C64-X<sub>5</sub>-C70-X<sub>45</sub>-H115-X<sub>3</sub>-C119] is the putative zinc-finger motif in ORF1. WHE is the VPg motif. (SGSYCTSSTNX<sub>19-35</sub>GDD) is the GDD motif of the RdRp. (b) Mosaic symptoms on leaves of mimosa

[4], with five open reading frames (ORFs) (Fig. 1B). The 5' end sequence is ACAAAA, which is present in sobemoviruses [4], and the 3' non-coding region is 399 nt long, slightly longer than those of other sobemoviruses (44 to 283 nt) [5]. P1 (encoded by ORF1) is a putative movement protein and has the conserved [C64-X<sub>5</sub>-C70-X<sub>45</sub>-H115-X<sub>3</sub>-C119] putative zinc-finger motif, which is involved in suppression of RNA silencing and cell-to-cell movement ability [4, 6]. The putative ORFx (174 aa) starts with a CTG codon [7]. P2a (ORFP2a) encodes a polyprotein with a putative transmembrane domain, a serine protease, and a genome-linked viral protein (VPg). In the region of overlap with the putative ORFx, two transmembrane alpha-helices (TMH) were predicted at aa positions 7-26 and 31-53, using TMHMM (<https://services.healthtech.dtu.dk>). The protease motif in P2a, TxxGxsG, was identified as TAQGFSG in MimMV, and the VPg motif (WAD/WGD/WNK), followed by a D/E-rich region [4, 8], was modified in MimMV as "WHEAEDEMDDAD" (Fig. 1A).

In other members of the genus, P2ab is expressed through a (-1) frameshift at a slippery sequence, UUUAAAC, which is followed by a stem-loop located seven nucleotides downstream of this site [9]. The frameshift site of MimMV is likely located at position 1909 in its genome. Using RNA-fold [10], a stem-loop was predicted four nucleotides downstream of the frameshift site of MimMV (Fig. 1A).

P2b encodes a putative RdRp and contains a conserved GDD motif (SGSYCTSSTNX<sub>19-35</sub>GDD). ORF3 encodes a putative coat protein (CP). The CP motif ACAA is located at position 3397 in the genome and may constitute the start site of the subgenomic RNA [4]. For genome recombination analysis, the complete genome sequence of MimMV and those of 21 sobemoviruses (20 of which have been assigned to a species and one for which a new species has been proposed), were analyzed using the RDP4 program [11], but no evidence signals of recombination was found for this isolate.

The MimMV genome sequence was compared to those of other sobemoviruses obtained from RefSeq. The sequences of the P2b and CP regions of the sobemoviruses were



**Fig. 2** Maximum-likelihood tree based on (a) P2b and (b) CP amino acid sequences of mimosa mosaic virus (MimMV) and members of the genus *Sobemovirus*. (c) Heatmap of pairwise sequence compari-

sons indicating the percent identity between the complete nucleotide sequences of MimMV and members of the genus *Sobemovirus*

extracted based on alignments made using MAFFT version 7 (<https://mafft.cbrc.jp/alignment/server>) and used for phylogenetic analysis. Phylogenetic trees were inferred using IQ-TREE2 [12] with the model Q.pfam+F+R4 for P2b and LG+F+I+G4 for CP, which were chosen using ModelFinder [13]. Phylogenetic analysis showed that blueberry shoestring virus (BBSSV) was the closest relative when the P2b region was used (Fig. 2a), and a rubus isolate of sowbane mosaic virus (SoMV) was closest to MimMV when the CP region was used (Fig 2b).

Pairwise sequence identity to sobemovirus sequences obtained from RefSeq was calculated using SDT v1.2. In the P2b region, BBSSV and tobacco velvet mottle virus (VTMoV) showed 52.7% and 43.3% aa sequence identity, respectively, to MimMV. In the CP region, BBSSV and the rubus isolate of SoMV showed 31.8% and 26.3% aa sequence identity, respectively. When complete genome sequences were compared, BBSSV was the closest relative to MimMV, with 46.8% nt sequence identity (Fig. 2c). The sequence identity between MimMV and BBSSV at the amino acid (52.7% in P2b and 31.8% in CP) and nucleotide (46.8% in the whole genome) level was low, despite BBSSV being the closest relative within the genus *Sobemovirus*. The demarcation criterion for establishing a new species in the genus *Sobemovirus* is < 75% nucleotide sequence identity in the complete genome sequence [14]. In conclusion, our results show that MimMV should be considered a member of a new species in the genus *Sobemovirus*, family *Solemoviridae*, and the binomial name “*Sobemovirus mimosae*” is suggested for this species.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00705-022-05683-5>.

**Acknowledgements** CMK and TN are CNPq fellows, Conselho Nacional de Desenvolvimento Científico e Tecnológico.

**Data availability statement** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors declare that they have no conflicts of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

1. Cali BB, Moyer JW (1981) Purification, serology, and particle morphology of two russet crack strains of sweet potato feathery mottle virus. *Phytopathology* 71:302–305
2. Blawid R, Silva JMF, Nagata T (2017) Discovering and sequencing new plant viral genomes by next-generation sequencing: description of a practical pipeline. *Ann Appl Biol* 170:301–314
3. Nicolini C, Inoue-Nagata AK, Nagata T (2015) Complete genome sequence of a proposed new tymovirus, tomato blistering mosaic virus. *Arch Virol* 160:609–612
4. Sömera M, Sarmiento C, Truve E (2015) Overview on sobemoviruses and a proposal for the creation of the family Sobemoviridae. *Viruses* 7:3076–3115
5. Desbiez C, Verdin E, Lecoq H (2021) Complete sequence of an isolate of snake melon asteroid mosaic virus confirms that it is a member of a distinct sobemovirus species. *Arch Virol* 166:2311–2313
6. Siré C, Bangratz-Reyser M, Fargette D et al (2008) Genetic diversity and silencing suppression effects of Rice yellow mottle virus and the P1 protein. *Virology* 375:1–12
7. Ling R, Pate AR, Carr JP et al (2013) An essential fifth coding ORF in the sobemoviruses. *Virology* 446:397–408
8. Olsper A, Peil L, Hebrard E et al (2011) Protein–RNA linkage and post-translational modifications of two sobemovirus VPgs. *J Gen Virol* 92:445–452
9. Tamm T, Suurväli J, Lucchesi J et al (2009) Stem-loop structure of cocksfoot mottle virus RNA is indispensable for programmed -1 ribosomal frameshifting. *Virus Res* 146:73–80
10. Lorenz R, Bernhart SH, HönerzuSiederdisenC, et al (2011) ViennaRNA Package 2.0. *Algorithms Mol Biol* 6:1–4
11. Martin DP, Murrell B, Golden M et al (2015) RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1:1–5
12. Minh BQ, Schmidt HA, Chernomor O (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37:1530–1534
13. Kalyaanamoorthy S, Minh BQ, Wong TK et al (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589
14. Sömera M, Fargette D, Hébrard E et al (2021) ICTV virus taxonomy profile: *Solemoviridae* 2021. *J Gen Virol* 102:001707

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.