Molecular characterization of a novel polymycovirus from the phytopathogenic fungus *Setosphaeria turcica*

Zhongnan Gao1 · Manyu Zhang1 · Tingting Yu1 · Xiaoyan Wang2 · Xin Wang3 · Hongliu An1 · Songbai Zhang1 · Minghong Liu2 · Shouguo Fang1

Received: 30 January 2021 / Accepted: 6 April 2021 / Published online: 24 May 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract

A putative polymycovirus tentatively named "Setosphaeria turcica polymycovirus 1" (StPmV1) was discovered in the phytopathogenic fungus *Setosphaeria turcica*. StPmV1 has a genome comprising five double-stranded RNAs (dsRNAs). dsRNA1, 2, and 3 each encode a protein sharing significant similarity but lower than 64% sequence identity to the corresponding proteins of other polymycoviruses. dsRNA4 and 5 each encode a protein with a sequence that is not conserved among polymycoviruses. However, the protein encoded by dsRNA4 is rich in proline (P), alanine (A), and serine (S) residues, which is a feature shared by the so-called PAS-rich proteins encoded by all polymycoviruses. Phylogeny reconstruction using the RNA-dependent RNA polymerase (RdRp) sequences of accepted or putative polymycoviruses revealed that StPmV1 is most closely related to Plasmopara viticola lesion associated polymycovirus 1 (PvaPolymyco1), a putative polymycovirus recovered from the phytopathogenic oomycetes *Plasmopara viticola*. These data suggest that StPmV1 may represent a novel species of the genus *Polymycovirus*, family *Polymycoviridae*. To our knowledge, this is the first polymycovirus reported from *S. turcica*.

Polymycoviruses are a family of double-stranded RNA (dsRNA) viruses that infect a wide range of fungi (https://talk.ictvonline.org/taxonomy). Unlike most other dsRNA viruses, polymycoviruses do not form isometric virions [1–9]. Instead, they encode a protein that is rich in proline (P), alanine (A), and serine (S) residues. In the case of Colletotrichum camelliae filamentous virus 1 (CcFV1), this PAS-rich protein (PASrp) encapsulates polymycoviral dsRNAs into filamentous particles [3]. In most other cases, PASrp coats, but does not encapsulate, viral dsRNAs [1, 2, 4–8]. Moreover, polymycoviruses have an RNA-dependent RNA polymerase (RdRp) that more closely related to the RdRps of positive-sense single-stranded RNA (+ssRNA) viruses than to those encoded by other dsRNA viruses. At least some polymycoviruses are infectious as naked genomic dsRNAs [1, 3]. These features have led to the idea that polymycoviruses may represent an evolutionary intermediate between dsRNA and +ssRNA viruses and between capsidated and capsidless viruses [1, 3, 8, 10]. Polymycoviruses show great flexibility in their genome arrangement. For example, Aspergillus fumigatus tetramycovirus 1 (AfuTmV-1), the type member of the family *Polymycoviridae*, has a genome consisting of four dsRNAs [1], whereas the genomes of many other polymycoviruses have 1-3 additional dsRNAs that encode proteins that are not conserved among polymycoviruses [2–9]. Although the effects of most polymycoviruses on the growth and pathogenicity of their host fungi remain poorly characterized, some polymycoviruses have been found to confer hypo- or hypervirulence to their hosts [1–4].

Handling Editor: Robert H.A. Coutts.

Zhongnan Gao and Manyu Zhang have contributed equally to this work.

Songbai Zhang
yangtze2008@126.com

Minghong Liu
lmh859@163.com

Shouguo Fang
sg-fang@hotmail.com

1 Hubei Engineering Research Center for Pest Forewarning and Management, Yangtze University, Jingzhou 434025, Hubei, China
2 Zunyi City Company, Guizhou Tobacco Company, Zunyi 563000, Guizhou, China
3 Fujian Province Key Laboratory of Plant Virology, Institute of Plant Virology, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China
Setosphaeria turcica is a fungal pathogen of corn as well as many other important crop plants. Four mycoviruses have been discovered in *S. turcica*: an ambiguivirus, a hypovirus, and a mitovirus, which were identified by searching the transcriptome data of *S. turcica*, and a fusarivirus, which was discovered by conventional cloning experiments [11, 12]. In this sequence note, we report the fifth mycovirus of *S. turcica*, a polymycovirus tentatively named "Setosphaeria turcica polymycovirus 1" (StPmV1).

Provenance of the virus material

*S. turcica* strain ZY17 was isolated from a corn leaf showing northern leaf blight symptoms from Zaoyang, Hubei province, China, in 2019. After growth for 6 days on a potato dextrose agar (PDA) plate at 28°C, mycelia (~0.2 g) of ZY17 were harvested and subjected to dsRNA extraction using a procedure described previously [13]. The dsRNA molecules in the extract were separated by electrophoresis in a 1% agarose gel. Four discrete bands were observed in the gel after ethidium bromide (EB) staining (Fig. 1A). Assuming that dsRNAs have a migration speed comparable to those of dsDNAs, this banding pattern suggests the presence of multiple dsRNA elements ranging in size between 1 and 2.5 kb. To sequence these dsRNAs, each band was excised from the gel. dsRNA from each gel slice was recovered using a TIANgel Midi Purification Kit (Tiangen, China) and reverse transcribed using the tagged random primer 5'-CGATCGATCATGATGGAATGNNNNN-3'. The resulting complementary DNA (cDNA) was amplified using the primer 5'-CGATCGATCATGATGGAATG-3'. PCR amplicons longer than 500 bp were each ligated to the vector pMD-20T (TAKARA), and the recombinant plasmids were introduced into *Escherichia coli* DH5a by transformation. For each transformation, 8-14 randomly selected colonies were sent to Beijing Genomics Institute (BGI, China) for Sanger sequencing. The sequences were analyzed and assembled using DNASTAR. The 5'- and 3'- terminal sequences of each dsRNA were determined by RNA-ligase-mediated rapid amplification of cDNA ends (RLM-RACE) [14]. The results showed that the smallest band was formed by two distinct dsRNAs, 1179 bp and 1115 bp in length. Each of the three larger bands represents a single dsRNA with a size of 2427, 2259, and 1994 bp, respectively. The five dsRNAs are named dsRNA1-5 in decreasing order of size, and their sequences were deposited in the GenBank database under the accession numbers MW429374.1-MW429378.1. Each nucleotide position on each dsRNA was sequenced at least four times.

![Fig. 1](image-url) Identification and molecular characterization of Setosphaeria turcica polymycovirus 1 (StPmV1). (A) Agarose gel electrophoresis of the dsRNAs extracted from *Setosphaeria turcica* strain ZY17. The positions of StPmV1 dsRNA1-5 are indicated. M indicates a DNA marker (DL15000, TAKARA). (B) The genome organization of StPmV1. Each dsRNA is represented by a horizontal line. The open reading frame (ORF) detected in each dsRNA is represented by a green box, and the nucleotide positions of the initiation and termination codon of each ORF are shown above the box. The relative positions of the conserved domains detected on each protein of StPmV1 are shown inside the boxes. (C) The 5’ and 3’ termini of the coding strand of StPmV1 dsRNA1-5.
**Sequence properties**

In accordance with the observation that polymycoviruses have a genome rich in GC pairs, the GC content of StPmV1 dsRNA1-5 is 60.90%, 61.22%, 61.08%, 62.34%, and 62.06%, respectively [1–9]. Using the ORF Finder program available at the National Center for Biotechnology Information (NCBI), a single open reading frame (ORF) was detected on each dsRNA. For simplicity, these ORFs are called ORF1-5, and the proteins encoded by them are named P1-P5, respectively. ORF1-5 are preceded by a 5′ untranslated region (UTR) of 23, 71, 52, 98, and 23 nt and followed by a 3′-UTR of 100, 94, 97, 118, and 159 nt, respectively (Fig. 1B). Supporting the idea that dsRNA1-5 are genomic components of the same virus, their terminal sequences are very similar: the coding strand of all five dsRNAs has an AU-rich 5′ terminus preceded by the dinucleotide GC and a GC-rich 3′-terminus ending with four U nucleotides (Fig. 1C).

P1 has a molecular mass of 84.28 kDa. A search of the NCBI conserved domain database revealed an RdRp domain (cd01699, RNA_dep_RNAP) spanning amino acid positions 466–603 [15–17]. As has been observed in other polymycoviruses, this domain contains a GDNQ motif that is common to the RdRps of -ssRNA viruses of the order Mononegavirales (Fig. 2A) [1–9, 19]. These data suggest that P1 is the RdRp of StPmV1. Indeed, a BLASTp search of the NCBI protein database found that P1 shares 63.29%, 57.63%, and 55.96% sequence identity with the RdRp of Plasmopara viticola lesion associated polymycovirus 1 (PvaPolymyco1), PvaPolymyco2, and Magnaporthe oryzae polymycovirus 1 (MoPmV1) and 25.53%–53.70% identity with the RdRps of other polymycoviruses [9, 19].

P2 has a molecular mass of 74.96 kDa. No conserved domain was found in P2. However, P2 shows 29.48%–47.49% sequence identity to dsRNA2-encoded proteins from many, although not all, polymycoviruses. The N-terminus of P2 has the sequence MADLTRL, which, as part of a signal peptide, is found at the N-termini of dsRNA2-encoded proteins of all polymycoviruses, as noted by Kotta-Loizou and Coutts, who proposed that polymycoviruses may use the protein encoded by dsRNA2 to anchor their replication machinery to a membranous compartment within the host cell [4]. Supporting this, a transmembrane helix from amino acid residue 279 to residue 298, was strongly predicted in P2 using the TMpred server (data not shown) [20]. Notably, a hypothetical protein encoded by an unnamed dsRNA (GenBank accession no. ACL80752) from a fungus of the genus Alternaria showed 60.34% amino acid sequence identity to P2.

P3 has a molecular mass of 65.68 kDa and is predicted to contain a methyltransferase domain (pfam13649) spanning amino acid residue position 131 to position 236. This methyltransferase domain has been detected in the dsRNA3-encoded proteins of all polymycoviruses [1–9]. Although experimental data are lacking at present, the presence of this domain suggests that polymycoviruses may use dsRNA3-encoded proteins to cap their genomic or messenger RNAs [1]. Like P2, P3 shows significant similarity to cognate proteins from a subset of polymycoviruses. The most similar protein, encoded by dsRNA3 of PvaPolymyco1, shares 59.54% amino acid sequence identity with P3 [9].

P4 has a molecular mass of 33.57 kDa. This 320-amino-acid-long protein has 31, 38, and 43 P, A, and S residues, respectively, suggesting that it is the PASrp of StPmV1. Although common to all polymycoviruses, PASrp does not have a conserved sequence [1–9, 21]. Consistent with this, a BLASTp search using P4 as a query returned no significant hits. Interestingly, however, a provisional domain described as “large tegument protein UL36” (cl33720) was found on P4, spanning amino acid positions 6–229.

P5 has a molecular mass of 33.81 kDa. This protein does not show similarity to any known proteins, nor does it contain any conserved domains. Thus, the function of P5 cannot be predicted at present.

To investigate the relationships between StPmV1 and other polymycoviruses, the RdRp sequences of accepted or putative polymycoviruses were used for phylogeny reconstruction with IQ-TREE [22, 23]. The RdRp of Hadaka virus 1 (HadV1), a polymycovirus-related virus, was used as an outgroup. The maximum-likelihood tree returned from IQ-TREE divided polymycoviruses into four major clusters, although the relationships between these clusters are not well resolved (Fig. 2B). StPmV1 was placed in cluster 4. As a comparison, AfuTmV-1, the type member of the family Polymycoviridae, was placed in cluster 3 [1]. CcFV1, the only polymycovirus known to form filamentous virions, was placed in cluster 2 [3]. Beauveria bassiana polymycovirus 1 (BbPmV1), a polymycovirus infecting the entomopathogenic fungus Beauveria bassiana, was placed in cluster 1. Within cluster 4, StPmV1 formed a monophyletic branch with PvaPolymyco1. Interestingly, PvaPolymyco3 and PvaPolymyco5, which were recovered from the same host as PvaPolymyco1, were not included in this branch [9].

Overall, StPmV1 has a genome with features characteristic of polymycoviruses. However, StPmV1 encodes an RdRp sharing <64% sequence identity with the RdRp of other polymycoviruses. StPmV1 may infect S. turcica, a fungus from which no polymycovirus has been reported. These data suggest that StPmV1 represents a novel species of the genus Polymycovirus, family Polymycoviridae. Because some polymycoviruses have been shown to affect the pathogenicity of their hosts, further studies on the effects of StPmV1 on S. turcica would be of interest.
Fig. 2 The evolutionary relationships between Setosphaeria turcica polymycovirus 1 (StPmV1) and other polymycoviruses. (A) Conserved RdRp motifs detected on the protein encoded by dsRNA1 of StPmV1. The alignment was performed with CLUSTAL_X. Three conserved RdRp motifs corresponding to motifs IV, V, and VI are shown. The asterisks indicate identical amino acid residues. Colors indicate highly conserved or related residues. Numbers within the brackets indicated the number of amino acid residues not shown. (B) A phylogeny of polymycoviruses reconstructed using IQ-TREE with the LG+F+I+G4 model. StPmV1 is indicated in red. The names of other polymycoviruses and Hadaka virus 1 (HadV1), which was used as an outgroup, are shown in black and preceded by the GenBank accession numbers of their RdRp sequences. The virus names and their abbreviations are as follows: Penicillium brevicompactum tetramycovirus 1 (PbTmV1), Beauveria bassiana polymycovirus 1 (BbPmV1), Sclerotinia sclerotiorum tetramycovirus 1 (SsTmV1), Phaeoacremonium minimum tetramycovirus 1 (PmTmV1), Beauveria bassiana polymycovirus 2 (BbPmV2), Magnaporthe oryzae polymycovirus 1 (MoPmV1), Plasmopara viticola lesion associated polymycovirus 3 (PvaPolymyco3), Plasmopara viticola lesion associated polymycovirus 5 (PvaPolymyco5), Plasmopara viticola lesion associated polymycovirus 1 (PvaPolymyco1), Penicillium digitatum polymycoviruses 1 (PdPmV1), Penicillium janthinellum polymycovirus 1 (PjPmV1), Botryosphaeria dothidea virus 1 (BdV1), Aspergillus fumigatus polymycovirus 1 (AfPmV1), Aspergillus niger tetramycovirus 1 (AsPmV1), Cladosporium cladosporioides virus 1 (CcV1), Fusarium redolens polymycovirus 1 (FrPmV1), and Colletotrichum camelliae filamentous virus 1 (CcFV1)
A novel polymycovirus from Setosphaeria turcica

Acknowledgments This work was supported by the Science and Technology Project of Guizhou Tobacco Company (No. 201921), the National Natural Science Foundation of China (no. 31972243), and the Opening Fund of Key Lab of Biopesticide and Chemical Biology, Ministry of Education, Fujian Agriculture and Forestry University. We are grateful to Dr. Zhenguo Du at the Institute of Plant Virology, Fujian Agriculture and Forestry University, for his help in preparing the manuscript.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

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

References

1. Kanhaiyuwa L, Kotta-Loizou I, Ozkan S, Gunning AP, Coutts RHA (2015) A novel mycovirus from Aspergillus fumigatus contains four unique dsRNAs as its genome and is infectious as dsRNA. Proc Natl Acad Sci USA 112:9100–9105
2. Zhai LF, Xiang J, Zhang MX, Fu M, Yang ZK, Hong N, Wang GP (2016) Characterization of a novel double-stranded RNA mycovirus conferring hypovirulence from the phytopathogenic fungus Botryosphaeria dothidea. Virology 493:75–85
3. Jia H, Dong KL, Zhou LL, Wang GP, Hong N, Jiang DH, Xu WX (2017) A dsRNA virus with filamentous viral particles. Nat Commun 8:168
4. Kotta-Loizou I, Coutts RHA (2017) Studies on the virome of the entomopathogenic fungus Beauveria bassiana reveal novel dsRNA elements and mild hypovirulence. PLoS Pathog 13:e1006183
5. Niu YH, Yuan YZ, Mao JL, Yang Z, Cao QW, Zhang TF, Wang SQ, Liu DL (2018) Characterization of two novel mycoviruses from Penicillium digitatum and the related fungicide resistance analysis. Sci Rep 8:5513
6. Nerva L, Forgia M, Ciuffo M, Chitarra W, Chiapello M, Vallino M, Vares GC, Turina M (2019) The mycovire of a fungal collection from the sea cucumber Holothuria polii. Virus Res 273:197737
7. Mahillon M, Decrois A, Liénard C, Bragard C, Legrève A (2019) Full genome sequence of a new polymycovirus infecting Fusarium redolens. Arch Virol 164:2215–2219
8. Sato Y, Jamal A, Kondo H, Suzuki N (2020) Molecular characterization of a novel polymycovirus from Penicillium janthinellum with a focus on its genome-associated PAsrp. Front Microbiol 11:592789
9. Chiapello M, Rodríguez-Romo J, Ayllón MA, Turina M (2020) Analysis of the virome associated to grapevine downy mildew lesions reveals new mycovirus lineages. Virus Evol 6:veaa058
10. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH (2020) Global Organization and Proposed Megataxonomy of the Virus World. Microbiol Mol Biol Rev 84(2):e00061-e119
11. Gao ZN, Cai LN, Liu MH, Wang XY, Yang JG, An HL, Deng QC, Zhang SB, Fang SG (2021) A novel previously undescribed fusarivirus from the phytopathogenic fungus Setosphaeria turcica. Arch Virol. https://doi.org/10.1007/s00705-021-04954-x
12. Gilbert KB, Holcomb EE, Allscheid RL, Carrington JC (2019) Hiding in plain sight: New virus genomes discovered via a systematic analysis of fungal public transcriptomes. PLoS ONE 14(7):e0219207. https://doi.org/10.1371/journal.pone.0219207
13. Deng QC, Wang HR, Li C, Li P, Fang SG, Yang S, Yan F, Zhang SB, Chen Z (2017) The complete genomic sequence of a novel alphapartitivirus from Bipolaris maydis, the causal agent of corn southern leaf blight. Arch Virol 162:2433–2436
14. Liu X, Gorovska MA (1993) Mapping the 5′ and 3′ ends of Tetrahymena thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). Nucleic Acids Res 21:4954–4960
15. Lu SN, Wang JY, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GB, Song JS, Thakni N, Yamashita RA, Yang GZ, Zhong DC, Zheng CJ, Lanzczky CJ, Marchler-Bauer A (2020) CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res 48:D265–D268
16. Bruenn JA (1993) A closely related group of RNA-dependent RNA polymerases from double-stranded RNA viruses. Nucleic Acids Res 21:5667–5669
17. Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. Crit Rev Biochem Mol Biol 28:375–430
18. Poeh O, Blumberg BM, Bouguelleret L, Tordo N (1990) Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: Theoretical assignment of functional domains. J Gen Virol 71:1153–1162
19. Hu G, Kurgan L (2019) Sequence similarity searching. Curr Protoc Protein Sci 95:e71
20. Hofmann K, Stoffel W (1993) TMBase-A database of membrane spanning proteins segments. Biol Chem Hoppe-Seyler 374:166
21. Sato Y, Shamsi W, Jamal A, Bhatti MF, Kondo H, Suzuki N (2020) Hadaka virus 1: a capsidless eleven-segmented positive-sense single-stranded RNA virus from a phytopathogenic fungus Fusarium oxysporum. MBio 11:e00450-e520
22. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274
23. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol 37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015

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