Complete genome sequence of a novel fusarivirus from the phytopathogenic fungus Corynespora cassiicola

Mingming Liu1,2 · Xintao Liu1 · Hui Zhao1 · Yunxia Ni1 · Min Jia1 · Peilin Hu1,2 · Hongyan Liu1,2 · Baoming Tian2

Received: 18 November 2021 / Accepted: 16 February 2022 / Published online: 9 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Corynespora cassiicola is an important phytopathogenic fungus that severely impairs crop production. Here, we report the molecular characterization of a novel positive-sense single-stranded RNA (+ssRNA) mycovirus, Corynespora cassiicola fusarivirus 1 (CcFV1), isolated from C. cassiicola strain 20200826-3-1. Excluding the poly(A) tail, the genome of the virus is 6491 nt in length and contains three putative open reading frames (ORFs). The large ORF1 encodes a polypeptide of 1524 aa with a conserved RNA-dependent RNA polymerase (RdRp) domain and a helicase (Hel) domain. BLASTp analysis showed that CcFV1 ORF1 has the highest similarity to Setosphaeria turcica fusarivirus 1 (StFV1, 50.45% identity, E-value 0.0). ORF2 encodes a polypeptide with a conserved chromosome segregation ATPase (Smc) domain. The smaller ORF3 encodes a polypeptide with an unknown function. Phylogenetic analysis based on the ORF1-encoded polypeptide showed that CcFV1 is phylogenetically related to members of the newly proposed family "Fusariviridae". Thus, we suggest that CcFV1 might be a novel member of the family "Fusariviridae", and is the first to be discovered in C. cassiicola.

Mycoviruses are ubiquitous in all major fungi and are transmitted intracellularly via cell division, sporogenesis, and cell-to-cell fusion [1]. Most fungal virus genomes consist of double-stranded RNA (dsRNA) or positive-sense single-stranded RNA (+ssRNA) [1]. Single-stranded circular DNA (ssDNA) viruses and negative-stranded RNA (-ssRNA) viruses have been reported recently in filamentous fungi [2, 3]. Mycoviruses with +ssRNA genomes are classified into nine families, including Alphaflexiviridae, Barnaviridae, Botourmiaviridae, Deltaflexiviridae, Endornaviridae, Gammaflexiviridae, Hypoviridae, Narnaviridae and Mito- viridae (https://talk.ictvonline.org/taxonomy). In 2014, the new family "Fusariviridae" was proposed. Its members are characterized by a +ssRNA genome with a length of 6-10 kb [4], and it was named after Fusarium graminearum virus-DK21 (FgV1) [5]. These viruses infect various phytopathogenic fungi, including Nigrospora oryzae [6], Neofusicoccum luteum [7], Botryosphaeria dothidea [8], Setosphaeria turcica [9], and Alternaria solani [10]. The detection of fusariviruses in an edible fungus, Auricularia heimuer, has been reported [11].

Corynespora cassiicola is a phytopathogenic fungus responsible for a devastating leaf disease [12] that has severely impaired the production of important crops. It has been found on plant leaves, stems, and roots, in nematode cysts, and on human skin [13–18]. C. cassiicola has a wide host range and has been reported to grow on at least 530 plant species of 380 genera [19].

So far, fungal viruses that infect C. cassiicola have rarely been reported. To date, only one unassigned dsRNA mycovirus has been reported to infect C. cassiicola [20]. In this study, we describe a novel fusarivirus isolated from the phytopathogenic fungus C. cassiicola. This is the first report of the complete genome sequence of a fusarivirus infecting C. cassiicola, and this mycovirus is therefore tentatively named "Corynespora cassiicola fusarivirus 1" (CcFV1).
Provenance of the virus material

The seven isolates of *C. cassiticola* used in this study were preserved at the Institute of Plant Protection, Henan Academy of Agricultural Sciences, China (Supplementary Table S1). For total RNA extraction, *C. cassiticola* strains were grown on potato dextrose agar (PDA) plates at 28°C in the dark for 8 days. Total RNA was extracted from 0.2 g of mycelium from each isolate using an RNAiso Kit (Takara, Dalian, China), following the manufacturer’s instructions. The RNA concentration of each preparation was adjusted to 200 ng/µl, and 15-µl aliquots from all preparations were pooled together.

The mixed sample was sent to Shanghai Bohao Biotechnology Corporation for high-throughput sequencing. Total RNA was depleted of rRNA using a Ribo-Zero™ technology (Takara, Dalian, China), following the manufacturer’s instructions. The RNA concentration of each preparation was adjusted to 200 ng/µl, and 15-µl aliquots from all seven isolates were pooled together.

To test for the presence of the newly discovered virus in the mixed sample, a paired-end (PE150) sequencing library was prepared and sequenced using an Illumina HiSeq 2500 platform. Clean reads to be used for data analysis were obtained by removing unqualified raw reads based on default parameters. These clean reads were assembled *de novo* using the scaffolding contig algorithm in CLC Genomics Workbench (CLC bio v6.0.4, Aarhus, Denmark) to obtain the primary unigenes. The final unigene sequences were obtained by applying CAP3 EST splicing software [21] for the second splicing of the primary unigenes. The final unigenes were used to search the NCBI GenBank database using BLASTx to identify homologous viral sequences. Using these methods, we identified the 6422-nt-long contig2308, assembled from 9415 related reads, as the genome sequence of a novel virus.

cDNAs of the seven *C. cassiticola* isolates were synthesized following the instructions of the PrimerScript II™ 1st Strand cDNA Synthesis Kit (TaKaRa, Dalian, China). For the presence of the newly discovered virus in each isolate, a virus-specific primer pair designed based on the contig2308 sequence was used for reverse transcription PCR (RT-PCR). A specific amplicon from contig2308 was detected in *C. cassiticola* strain 20200826-3-1, which was isolated from the diseased stem of a sesame plant from the city of Zhoukou, Henan Province, China.

The nucleotide sequence of contig2308 was verified by sequencing amplified fragments of the cDNA of strain 20200826-3-1, using 16 pairs of specific primers (Supplementary Table S2). To complete the 5’- and 3’-terminal genomic sequences of CcFV1, rapid amplification of cDNA ends (RACE) was performed using a SMARTer RACE 5’/3’ Kit (TaKaRa, China), using the specific primers R_GSP_Contig2308 and F_GSP_Contig2308, respectively. The procedures were performed according to the user manual provided with the kit. All PCR products were purified using a FastPure Gel DNA Extraction Kit (Vazyme, China), cloned into the pMD™18-T vector, and then introduced into chemically competent *E. coli* JM109 cells (TSINGKE, China) for propagation. At least three recombinant bacterial clones were selected and sent to Sangon Biotech for sequencing to verify the accuracy of the CcFV1 nucleotide sequence.

The ORF Finder and CD-search tools of NCBI (https://www.ncbi.nlm.nih.gov/) were used to search for ORFs and conserved motifs, respectively. The sequences obtained from the clones were assembled using DNAMAN software (Lynnon Biosoft, Canada). Multiple alignments of the amino acid sequence were performed using Clustal X and DNA-MAN software. The corresponding phylogenetic trees were constructed by the maximum-likelihood (ML) method with 1000 bootstrap replicates, using the Jones–Taylor–Thornton (JTT) model in MEGA (7.0 version) [23]. The GC content was determined using Novopro (https://www.novopro.cn/tools/gc-content.html). Expasy was used to calculate the protein molecular mass (Mₚ) and isoelectric point (pI) [24]. A schematic diagram of the genome organization was drawn using Illustrator for Biological Sequences [25].

Sequence properties

The full-length genome sequence of CcFV1 (GenBank accession number OL456888) is 6491 nt long, excluding the poly(A) tail, and has a GC content of 43.54%. The positive-strand RNA genome contains two main ORFs (ORF1 and ORF2) and a smaller ORF3. The lengths of the 5’-UTR and 3’-UTR are 18 nt and 47 nt, respectively (Fig. 1A). A BLASTx search against the NCBI non-redundant protein sequence database showed that the three viral sequences showing the highest degree of similarity to CcFV1 were those of the RNA-dependent RNA polymerases (RdRps) of Setosphaeria turcica fusarivirus 1 (StFV1, 50.10% identity, E-value 0.0), Erysiphe necator-associated fusarivirus 2 (EnFV2, 52.23% identity, E-value 0.0), and Erysiphe necator-associated fusarivirus 1 (EnFV1, 50.10% identity, E-value 0.0). All three of these viruses belong to the newly proposed family *Fusariviridae*.

ORF1 (nt 19-4593) encodes a polypeptide of 1524 amino acids (aa) with an approximate Mr of 172.59 kDa and an isoelectric point of 8.98. A BLASTp search against the NCBI non-redundant protein sequence database showed that this protein is 50.45%, 51.26%, and 49.55% identical to the RdRps of StFV1, EnFV2, and EnFV1, respectively, and the top 10 viruses with the highest degree of matching to this polypeptide were all fusariviruses (Table 1).

Using the CD search program on the NCBI website, two conserved sequence domains, encoding an RdRp (RdRp_1, pfam00680, E-value 7.80e-09, aa positions 479-741) and
A novel fusarivirus from *Corynespora cassiicola*

A helicase (Helicase_C, pfam00271, E-value 2.77e-09, aa positions 1229-1341), were identified in the ORF1-encoded polyprotein, and these have also been identified in other fusariviruses. The aa sequence similarities in the RdRp and Hel regions indicated that CcFV1 is a fusarivirus. Multiple alignments and comparisons of the viral RdRp domains between CcFV1 and other selected fusariviruses revealed eight conserved motifs that are characteristic of members of the proposed family "Fusariviridae" (Supplementary Fig. S1).

**ORF2** (nt 4645-6144) encodes a 499-aa polyprotein of 56.68 kDa with a pI of 9.06. A BLASTp search revealed...
only eight similar proteins, all of which were fusarivirus proteins (Table 1). The polyprotein encoded by ORF2 showed the highest sequence similarity to EnFV2 (35.42% identity, E-value 2.00E-87). A conserved domain spanning aa positions 89 to 246 was identified in this protein, which corresponded to a chromosome segregation ATPase (Smc, COG1196, E-value 1.50E-07) and was also predicted in other fusariviruses (Fig. 1A).

ORF3, covering nt positions 6145 to 6444, was predicted to encode a hypothetical polypeptide of 11.52 kDa with a pI of 9.23. A homology search indicated no aa sequence similarity to any other known proteins or conserved domains. Unlike those of StFV1, EnFV2, and EnFV1, ORF3 and ORF2 of CcFV1 are in a consecutive arrangement without overlapping regions (Fig. 1A).

To examine the relationships between CcFV1 and other mycoviruses, we performed phylogenetic analysis based on an alignment of the ORF1-encoded polyprotein containing the RdRp and RNA helicase (Hel) domains (Fig. 1B). The sequences of selected members of the family Hypoviridae were included as outgroups. The results of the phylogenetic analysis showed that CcFV1 clustered with previously reported fusariviruses.

The species demarcation criteria for fusariviruses have not been established [9], and therefore, whether CcFV1 should be considered a member of a novel species is uncertain. However, the RdRp gene product of CcFV1 is only 50.45% identical to that of the most closely matching virus (StFV1), and CcFV1 is the first fusarivirus isolated from C. cassiciola. Thus, it is likely that CcFV1 will be considered a member of a novel species in the proposed family "Fusariviridae".

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00705-022-05428-4.

**Acknowledgments** We are extremely grateful to Dr. Jing Wang for her valuable suggestions and help with sequence analysis.

**Funding** This work was supported by the China Agriculture Research System of MOF and MARA [Grant number CARS-14] and the Key Project of Science and Technology of Henan Province [Grant number 201300110600].

**Declarations**

**Conflict of interest** All 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. Ghabrial SA, Castón JR, Jiang D et al (2015) 50-plus years of fungal viruses. Virology 479–480:356–368. https://doi.org/10.1016/j.virol.2015.02.034

2. Yu X, Li B, Fu Y et al (2010) A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. Proc...
A novel fusarivirus from *Corynespora cassiicola*

1. Silva WPK, Wijesundera RLC, Karunanayake EH et al (2000) New hosts of *Corynespora cassiicola* in Sri Lanka. Plant Dis 84:202. https://doi.org/10.1094/pdis.2000.84.2.202d
2. Madriz-Ordeana K, Jrgensen HJL, Nielsen KL, Thordal-Christensen H (2016) First report of Kalanchoe leaf and stem spot caused by *Corynespora cassiicola* in Denmark. Plant Dis 101(101):505–505. https://doi.org/10.1094/PDIS-08-16-1173-PDN
3. Liu L, Xie J, Cheng J et al (2014) Fungal negative-stranded RNA virus that is related to bornaviruses and nyaviruses. Proc Natl Acad Sci 111:12205 LP-12210 LP. https://doi.org/10.1073/pnas.1401786111
4. Zhang R, Liu S, Chiba S et al (2014) A novel single-stranded RNA virus isolated from a phytopathogenic filamentous fungus, *Rosellinia necatrix*, with similarity to hypo-like viruses. Front Microbiol 5:1–12. https://doi.org/10.3389/fmicb.2014.00360
5. Kwon SJ, Lim WS, Park SH et al (2009) Erratum: Molecular characterization of a dsRNA mycovirus, *Fusarium graminearum* virus-DK21, which is phylogenetically related to hypoviruses but has a genome organization and gene expression strategy resembling those of plant potex-like viruses. Mol Cells 28:73–74. https://doi.org/10.1007/s10059-009-0112-1
6. Zhong J, Zhao SQ, Li GF et al (2016) A novel fusarivirus isolated from the phytopathogenic fungus *Nigrospora oryzae*. Virus Genes 52:891–895. https://doi.org/10.1007/s11262-016-0317-5
7. Marais A, Nivault A, Faure C et al (2018) Molecular characterization of a novel fusarivirus infecting the plant-pathogenic fungus *Neofusicoccum luteum*. Arch Virol 163:559–562. https://doi.org/10.1007/s00705-017-3620-x
8. Liu W, Hai D, Ma F et al (2020) Molecular characterization of a novel fusarivirus infecting the plant-pathogenic fungus *Botryosphaeria dothidea*. Arch Virol 165:1033–1037. https://doi.org/10.1007/s00705-020-04554-1
9. Gao Z, Cai L, Liu M et al (2021) A novel previously undescribed fusarivirus from the phytopathogenic fungus *Setosphaeria turcica*. Arch Virol 166:665–669. https://doi.org/10.1007/s00705-021-04954-x
10. Gong W, Liu H, Zhu X et al (2021) Molecular characterization of a novel fusarivirus infecting the plant-pathogenic fungus *Alternaria solani*. Arch Virol 166:2063–2067. https://doi.org/10.1007/s00705-021-05105-y
11. Li X, Sui K, Xie J et al (2020) Molecular characterization of a novel fusarivirus infecting the edible fungus *Auricularia heimuer*. Arch Virol 165:2689–2693. https://doi.org/10.1007/s00705-021-05105-y
12. Silva WPK, Wijesundera RLC, Karunanayake EH et al (2000) New hosts of *Corynespora cassiicola* in Sri Lanka. Plant Dis 84:202. https://doi.org/10.1094/pdis.2000.84.2.202d
13. Carris LM, Glawe DA, Gray LE (1986) Isolation of the soybean pathogens *Corynespora cassiicola* and *Phialophora gregata* from cysts of *Heterodera glycines* in Illinois. Mycologia 78:503–506
14. Sato R, Kitazawa K (1980) Occurrence of soybean root rot caused by *Corynespora cassiicola* (Berk. & Curt.) Wei in Hokkaido. Jpn J Phytopathol 46:193–199
15. Dornon LJ, Schlub RL, Pernezny K, Datnoff LE (2009) Host specialization and phylogenetic diversity of *Corynespora cassiicola*. Phytopathology 99:1015–1027. https://doi.org/10.1094/PHYTO-99-9-1015
16. Kusakari S, Okada K, Nakasone W, Tanaka Y (2009) *Corynespora leaf spot* of Perilla (*Perilla frutescens* Brit.) caused by *Corynespora cassiicola* (Berk. & Curt.). We. Jpn J Phytopathol 57:737–740
17. Yan XX, Yu CP, Fu XA et al (2016) CAR9 mutation linked to *Corynespora cassiicola* infection in a Chinese patient. Br J Dermatol 174:176–179. https://doi.org/10.1111/bjd.14082
18. Smith LJOY (2008) Host range, phylogenetic, and pathogenic diversity of *Corynespora cassiicola* (Berk. & Curt.). Wei. Ph.D. dissertations, University of Florida, Gainesville, Florida, USA
19. Wang Y, Zhao H, Xue C et al (2020) Complete genome sequence of a novel mycovirus isolated from the phytopathogenic fungus *Corynespora cassiicola* in China. Arch Virol 165:2401–2404. https://doi.org/10.1007/s00705-020-04723-2
20. Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877. https://doi.org/10.1101/gr.9.9.868
21. Thompson JD, Gibson TJ, Plewniak F et al (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882. https://doi.org/10.1093/nar/25.24.4876
22. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054
23. Gasteiger E, Hoogland C, Gattiker A et al (2005) Protein identification and analysis tools on the ExPASy server. Proteom Protoc Handb. https://doi.org/10.1385/1592598900
24. Liu W, Xie Y, Ma J et al (2015) IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics 31:3359–3361. https://doi.org/10.1093/bioinformatics/btv362

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