Arch Virol (2014) 159:1861–1863
DOI 10.1007/s00705-014-1992-8

ANNOTATED SEQUENCE RECORD

Complete genome sequence of a novel hypovirus infecting *Phomopsis longicolla*

Igor Koloniuk · Mohamed H. El-Habbak · Karel Petrzik · Said A. Ghabrial

Received: 7 November 2013 / Accepted: 14 January 2014 / Published online: 29 January 2014

Abstract The complete nucleotide sequence and genome organization of a hypovirus from the isolate ME711 of *Phomopsis longicolla* was determined and compared to sequences of members of the family *Hypoviridae*. The genome of the hypovirus, tentatively named Phomopsis longicolla hypovirus 1 (PlHV1-ME711), was determined to be 9760 nucleotides long, excluding the 3' poly (A) tail. The genome contains a single large open reading frame (ORF) encoding a polyprotein designated as P307. Its genomic organization is typical of members of the proposed genus *Betahypovirus* (Yaegashi et al. in Virus Res 165:143–50, 2012).

The members of family *Hypoviridae*, which infect fungal hosts, have monopartite ssRNA genomes ranging from 9 to 13 kilobases (kb) in size. Hypoviruses do not produce true virions, and their replication is known to occur in pleomorphic vesicles, containing replicative form dsRNA, in the cytoplasm of infected cells [2]. There are four recognized hypovirus species, whose members all infect *Cryphonectria parasitica*: *Cryphonectria hypovirus 1*, *2*, *3* and *4* (CHV1, CHV2, CHV3, CHV4). Recently, three other viruses were reported: Sclerotinia sclerotiorum hypovirus 1 (SsHV1), Fusarium graminearum hypovirus 1 (FgHV1) and Valsa ceratosperma hypovirus 1 (VcHV1) [1, 3, 4].

P. longicolla (teleomorph = *Diaporthe*) is the causal agent of soybean seed decay, which results in poor seed quality [6]. *P. longicolla* isolate ME711 was originally derived from a slow-growing sector in a culture of *P. longicolla* (designated here as Pl-KY) previously obtained from a soybean seed lot in Kentucky [5]. Mycelium from isolate ME711 was used for dsRNA extraction. A hypovirus-free virulent strain (*P. longicolla* Hobbs, ATCC 60325) was also included in this study as a control. The fungal cultures were maintained on PDAY plates at 24 °C in dark.

Species identification was based on sequencing of the genomic ITS region amplified with ITS1 and ITS4 primers [7]. Extraction of dsRNA was performed using CF-11 cellulose (Whatman, USA) as described previously [8]. cDNA was synthesized using SuperScript® II Reverse Transcriptase (Invitrogen, USA) and then cloned and sequenced as described previously [9]. Clones containing inserts were sequenced using a BigDye® Terminator v3.1 Cycle Sequencing Kit on an ABI 3730 DNA Analyzer (Applied Biosystems, USA). Sequence data were analyzed and assembled using CLC Main Workbench 6.7.1 (CLC bio, Denmark). The complete genomic sequence of PlHV1/ME711 was determined from a series of sequential RT-PCR steps. The 5' and 3' terminal sequences were amplified using a MINT Universal cDNA Synthesis Kit (Evrogen, Russia). Conserved domains of the putative functional proteins were determined by comparison of their sequences with those of other hypoviruses. Pairwise comparisons of nucleotide (nt) and deduced amino acid (aa) sequences were performed using the CLC Main Workbench 6.7.1 and Geneious 6.1.5 (Biomatters, New Zealand) software packages. Phylogenetic inference was conducted using the

I. Koloniuk (✉) · K. Petrzik
Department of Plant Virology, Institute of Plant Molecular Biology, Biology Centre of Academy of Sciences of the Czech Republic, v.v.i., Branišovská 31, 370 05 Ceske Budejovice, Czech Republic
e-mail: koloniuk@umbr.cas.cz

M. H. El-Habbak · S. A. Ghabrial
Plant Pathology Department, University of Kentucky, Lexington, KY 40546, USA

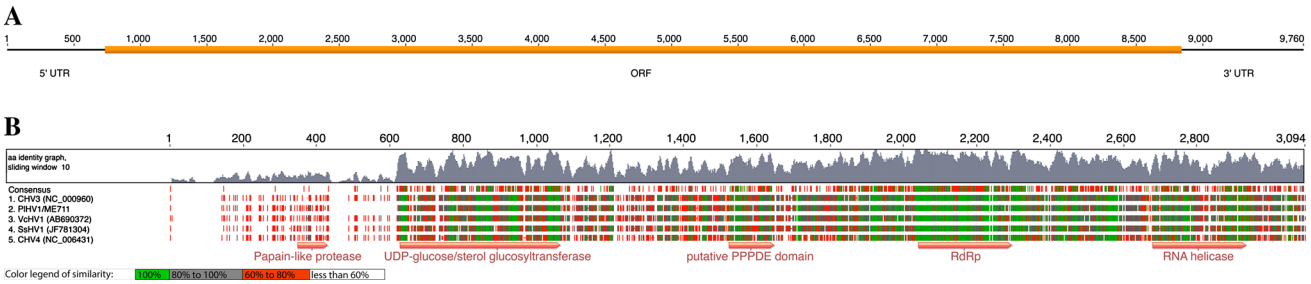

Fig. 1 (A) Genome organization of PlHV1/ME711. Domains of the predicted polyprotein are shown as shaded boxes. (B) Mean pairwise aa identity over all pairs [PlHV1 (KF537784), VcHV1(AB690372), SsHV1(JF781304), CHV3(AF188515), CHV4(AY307099)] calculated over a sliding window of 20 aa

maximum-likelihood method implemented in PhyML 3.0 and TreeDyn 198.3 at http://www.phylogeny.fr/ via the "A la Carte" mode with default settings. Branch support values (%) were estimated by the approximate likelihood ratio test (aLRT) with SH-like criteria, and those with values <50 % were collapsed. Alignments were conducted using Clustal Omega 1.2.0 as implemented at http://www.ebi.ac.uk/Tools/msa/clustalo/ with default settings.

The complete nucleotide sequence of the ME711 isolate of PlHV1 has been submitted to the GenBank database under accession number KF537784. Search for an ORF revealed a single 8123-nt-long ORF flanked by a 5' UTR of 732 nt and a 3' UTR of 904 nt. The 5' UTR of PlHV1/ME711 is thus so far the longest one among known hypoviruses, while the 3' UTR is almost as long as that of SsHV1/SZ-150 (1010 nt). The putative ORF starts at the AUG triplet (nt 733-735) and ends with a UAA termination codon at nt 8854-8856. The termination codon UAA is present also in VcHV1, CHV1 and CHV2, while ORFs of other hypoviruses are terminated with UAG. The predicted ORF of PlHV1-ME711 encodes a polyprotein of 2707 aa with a calculated molecular mass of 307.4 kDa. A multiple sequence alignment including CHV3, CHV4, SsHV1 and VcHV1 confirmed the existence of described functional domains: papain-like protease (Pro), UDP glucose/sterol glucosyltransferase (UGT), RNA-dependent RNA polymerase (Pol), and RNA helicase (Hel). After search of the P307 aa sequence against NCBI's Conserved Domain Database [10], a putative domain of permuted papain fold peptidases of dsRNA viruses and eukaryotes (PPPDE, pfam05903) was revealed between UTG and Pol (aa residues 1151 and 1266; Fig. 1). Proteins of the PPPDE family function as de-ubiquitinating and de-SUMOylating peptidases [11] and have not yet been characterized in detail.

Sequence analysis showed that the PlHV1/ME711 genome has the highest nt sequence identity (61.7 %) with CHV3 (AF188515). Pairwise nt sequence analysis of all hypoviruses clearly segregates PlHV1, SsHV1, VcHV1 and CHV3 into one distinct group. While they share over 50 % nt sequence identity with each other, their pairwise identity

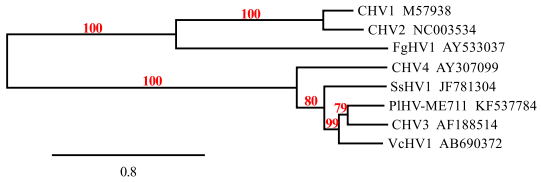

Fig. 2 Maximum-likelihood tree based on full-length sequence of PlHV1-ME711 and other definitive and probable members of the family *Hypoviridae*, highlighting the existence of two discrete clades

with CHV4 does not exceed 33 %. At the same time, CHV1 and CHV4 have 34 % nt sequence identity, although they have a slightly different genomic organization. Another group consists of CHV1 and CHV2, which share 58 % nt sequence identity. Phylogenetic analysis revealed that the ME711 isolate of PlHV1 was grouped into a distinct clade with VcHV1, SsHV1 and CHV3 (Fig. 2).

Protoplast preparation from the ME711 isolate of *Phomopsis longicolla* gave rise to several differently growing subcultures, from almost wild-type phenotype to stunted colonies with irregular colony edges (data not shown). It has been suggested that in some cases hypovirulence may be induced not by the virus itself, but by a satellite, like in the strain SZ-150 of *Sclerotinia sclerotiorum* infected with SsHV1/SZ-150 and harboring a satellite-like dsRNA element [3]. Thus, the biological properties and impact of PlHV1-ME711 on its fungal host will be addressed in future studies.

Based on the species demarcation criteria for hypoviruses proposed in the Ninth Report of the International Committee on Taxonomy of Viruses, PlHV1-ME711 represents an isolate of a new species, for which we propose the name "*Phomopsis longicolla hypovirus 1*".

Acknowledgments This work was supported by Postdok_BIOGLOBE (CZ.1.07/2.3.00/30.0032), co-financed by the European Social Fund and the state budget of the Czech Republic (to I.K and K.P) and KSEF-2178-RDE-013 from the Kentucky Science and Engineering Foundation (to S.A.G.). A part of this research was performed during the first author's postdoctoral tenure at the University of Kentucky.

References

1. Yaegashi H, Kanematsu S, Ito T (2012) Molecular characterization of a new hypovirus infecting a phytopathogenic fungus, *Valsa ceratosperma*. Virus Res 165:143–150. doi:10.1016/j.virusres.2012.02.008
2. Nuss D, Hillman B, Rigling D, Suzuki N (2011) Family *Hypoviridae*. In: King AMQ, Elliot L, Adams MJ, Carstens EB (eds) Virus taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, San Diego, pp 1029–1033
3. Xie J, Xiao X, Fu Y et al (2011) A novel mycovirus closely related to hypoviruses that infects the plant pathogenic fungus *Sclerotinia sclerotiorum*. Virology 418:49–56. doi:10.1016/j.virol.2011.07.008
4. Wang S, Kondo H, Liu L et al (2013) A novel virus in the family Hypoviridae from the plant pathogenic fungus *Fusarium graminearum*. Virus Res 174:69–77. doi:10.1016/j.virusres.2013.03.002
5. Gleason ML, Ghabrial SA, Ferriss RS (1987) Serological detection of *Phomopsis longicolla* in soybean seeds. Phytopathology 77:371. doi:10.1094/Phyto-77-371
6. Sinclair J (1993) Phomopsis seed decay of soybeans: a prototype for studying seed disease. Plant Dis 77:329–334
7. White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, UK, pp 315–322
8. Castillo A, Cottet L, Castro M, Sepúlveda F (2011) Rapid isolation of mycoviral double-stranded RNA from *Botrytis cinerea* and *Saccharomyces cerevisiae*. Virol J 8:38. doi:10.1186/1743-422X-8-38
9. Darissa O, Willingmann P, Adam G (2010) Optimized approaches for the sequence determination of double-stranded RNA templates. J Virol Methods 169:397–403. doi:10.1016/j.jviromet.2010.08.013
10. Marchler-Bauer A, Lu S, Anderson JB et al (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39:D225–D229. doi:10.1093/nar/gkq1189
11. Iyer LM, Koonin EV, Aravind L (2004) Novel predicted peptidases with a potential role in the ubiquitin signaling pathway. Cell Cycle 3:1440–1450