Complete Nucleotide Sequence of an Endornavirus Isolated from Common Buckwheat (*Fagopyrum Esculentum)*

Ryo Okada
Ibaraki-ken Nogyo Sogo Center
Hiromitsu Moriyama (hmori714@cc.tuat.ac.jp)
Tokyo University of Agriculture and Technology https://orcid.org/0000-0001-7678-8015
Takashi Ogawara
Ibaraki-ken Nogyo Sogo Center
Heather Cizek
Louisiana State University
Rodrigo A Valverde
Louisiana State University

Research Article

Keywords: dsRNA, endornavirus, buckwheat , Fagopyrum esculentum

Posted Date: July 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-750508/v1

Abstract

A double-stranded RNA (dsRNA) of approximately 16 kbp was isolated from symptomless common buckwheat (*Fagopyrum esculentum*) plants. The size of the dsRNA suggested that it was the replicative form of an endornavirus. The dsRNA was sequenced, and it consisted of 15,677 nt containing a single open reading frame that potentially encoded a polyprotein of 5,190 aa. The polyprotein contained conserved domains for a viral methyltransferase, viral RNA helicase 1, MSCRAMM family adhesion SdrC, UDP-glycosyltransferase, and viral RNA dependent RNA polymerase 2. A site-specific nick in the plus strand was detected near the 5' end of the dsRNA. BLASTP analysis showed that the polyprotein shared the highest identity with the polyprotein of winged bean endornavirus 1. Results of phylogenetic analysis supported placing the novel virus from common buckwheat, which was provisionally named Fagopyrum esculentum endornavirus 1, in the genus *Alphaendornavirus* of the family *Endornaviridae.*

Introduction

Endornaviruses are viruses with linear single-stranded positive-sense RNA genomes ranging from approximately 9.7 to17.6 kb reported to infect plants, fungi, and oomycetes [22, 23]. Their genome codes for a single polyprotein which contains conserved domains. Although, true virions have not been reported, cytoplasmic vesicles have been associated with some endornaviruses [8, 17]. Endornaviruses are classified in the family *Endornaviridae* which contains two genera, *Alphaendornavirus* and *Betaendornavirus* [23]. In plants, endornaviruses do not cause apparent effects on the phenotype and are transmitted only vertically [2, 23]. Endornaviruses have been reported in many cultivated and wild plant species [2, 3, 5, 23].

Common buckwheat (*Fagopyrum esculentum* Moench) and Tartary buckwheat (*F. tataricum* Gaertner) are members of the family Polygonaceae. They are cultivated worldwide for their nutritional and medicinal values, which includes gluten-free and a broad spectrum of flavonoids with antioxidant effects [7, 18].

Symptomless common buckwheat plants cv. Kitawasesoba growing in the field of Ibaraki Agricultural Center in Kasama, Japan were tested for the presence of large dsRNAs by the spin column method [12] and found to contain an approximately 16 kbp dsRNA. Because this dsRNA was similar in size to dsRNAs found associated with many plant endornaviruses, we cloned and Sanger-sequenced random-primed cDNA and RACE products as described previously [11, 15]. Amino acid sequences were assembled and analyzed using Genetyx Ver9 (GENETYX corporation, Japan). Multiple sequence alignments of deduced amino acid sequences were conducted using Clustal X [20]. A Maximum Likelihood-based phylogenic tree was constructed using MEGA ver. 6.0 [19]. Bootstrap tests were performed with 1,000 resamplings. Preliminary sequence analyses yielded conserved domains typical of plant endornaviruses. Further sequence analyses suggested that a novel endornavirus was infecting Kitawasesoba common buckwheat. The complete sequence was submitted to the GenBank (Accession No. LC500285) and the virus was provisionally named Fagopyrum esculentum endornavirus 1 (FeEV1).

The nucleotide sequence of FeEV1 was determined to be 15,677 bp in length containing a single open reading frame (ORF) in one strand (designated as the plus strand), starting at nt 73 (the second AUG codon) and ending at nt 15,645. The ORF potentially encoded a polyprotein of 5,190 aa with an estimated molecular mass of 589 kDa (**Fig.1**). The first AUG codon was found at nt 22, however, it was not considered to be the favorable initiation codon. The 5'-UTR and 3'-UTR of FeEV1 consisted of 72 and 32 nt long, respectively. The 3'-UTR ended with 12 cytocines. A BLAST search using the aa sequence detected conserved domains of a putative viral methyltransferase (MT) (aa 311-454), viral RNA helicase 1 (HEL) (aa 1467-1712), UDP-glycosyltransferase (UGT) (aa 3262-3608), and viral RNA dependent RNA polymerase 2 (aa 4884-5116) (**Fig.1**). These four domains are also reported in some plant endornaviruses [10, 13, 15, 16]. In addition, a unique conserved domain, MSCRAMM family adhesion SdrC, was found in between the HEL and UGT domains (aa 1,945-2077). Features of this protein family include a YSIRK-type signal peptide at the N-terminus and a variable-length C-terminal region of Ser-Asp (SD) repeats followed by an LPXTG motif for surface immobilization by sortase [1]. FeEV1 also included the YSIRK motif at the N-terminus of the polyprotein (aa 28-32) and five SD motifs followed by the motif LPSTG (aa 1,950-1,954) in the MSCRAMM family domain. The aa sequence LPSTG may be the protease recognition site for the cleavage of the polyprotein. *In silico* analysis showed that a cysteine-rich region (CRR), identified in other endornaviruses and suggested as a candidate for a viral protease that can process the polyprotein into functional units [4, 15, 21], was also found in FeEV1. The polyprotein of FeEV1 had three CXCC signatures between MT and HEL domains (aa 773-776, 778-791, and 820-823). Like in the case of other plant endornaviruses [14, 15], a site-specific nick (nt 891) near the 5' region of the plus strand was detected in the FeEV1 dsRNA by 5' RACE experiments (**Fig.1**). The role of the nick is unknown, but it is thought to be involved in regulation of virus replication [13].

Seeds of 34 plant introductions of *F. esculentum* and two of *F. tataricum* were obtained from the USDA-ARS, Plant Genetic Resources Unit, Geneva, New York, USA and planted in the laboratory under fluorescent lights. Visual inspections of the plants did not show any virus-like symptoms. Foliar samples were collected from individual plants at maturity (flowering stage) and dsRNA extracted and analyzed by electrophoresis as described by Khankhum et al. [6]. Analyses of dsRNA extractions yielded dsRNAs of approximately 16 kb in 16 *Fagopyrum* plant introductions (from nine countries) but not in 20 others (**Supplementary Table 1**). DsRNA-positive plants were tested by RT-PCR using general endornavirus primers [24] and the endornavirus nature of the dsRNAs was confirmed. These results suggest that FeEV1 or perhaps other endornaviruses are present in the *F. esculentum* and *F. tataricum* germplasm.

BLASTP analysis showed that the polyprotein of FeEV1 shared the highest identity (45.84%) with that of winged bean endornavirus 1 (91% query cover). The genome organization of FeEV1 is similar to that of endornaviruses of winged bean (winged bean endornavirus 1), pepper (bell pepper endornavirus and hot pepper endornavirus), and common bean (Phaseolus vulgaris endornavirus 2). In addition, the phylogenetic tree of the RdRp places FeEV1 in a group including these four endornaviruses (**Fig. 2**). In conclusion, the data provided here support placing the virus isolated from common buckwheat as a novel species in the genus *Alphaendornavirus* in the family *Endornaviridae*.

Declarations

Acknowledgments

This investigation was partially supported by subsidy based on the three kinds of electric power laws for science and technology promotion of power plants and other nuclear energy facilities siting prefecture from Ministry of Education, Culture, Sports, Science and Technology. We wish to thank the United States Department of Agriculture, National Institute of Food and Agriculture, for partial support for this investigation.

References

1. Bae T, Schneewind O (2003) The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing. J Bacteriol 185:2910–2919. https://doi.org/10.1128/JB.185.9.2910-2919.2003
2. Fukuhara T (2019) Endornaviruses: persistent dsRNA viruses with symbiotic properties in diverse eukaryotes. Virus Genes 55:165–173. https://doi.org/10.1007/s11262-019-01635-5
3. Fukuhara T, Koga R, Aoki N, Yuki C, Yamamoto N, Oyama N, Undagawa T, Horiuchi H, Miyazaki S, Higashi Y, Takeshita M, Ikeda K, Arakawa M, Matsumoto N, Moriyama H (2006) The wide distribution of endornaviruses. Arch Virol 151:995–1002. doi:10.1007/s00705-005-0688-5
4. Hacker CV, Brasier CM, Buck KW (2005) A double-stranded RNA from a Phytophthora species is related to the plant endornaviruses and contains a putative UDP glycosyltransferase gene. J Gen Virol 86:1561–1570. doi:10.1099/vir.0.80808-0
5. Hershlag R, Escalante C, Rodrigues de Souto E, Kankhum S, Okada R, Valverde RA (2019) Occurrence of putative endornaviruses in non-cultivated plant species in south Louisiana. Arch Virol 164:1863–1868. doi:10.1007/s00705-019-04270-5
6. Khankhum S, Escalante C, Rodrigues de Souto E, Valverde RA (2017) Extraction and electrophoretic analysis of large dsRNAs from desiccated plant tissues infected with plant viruses and biotrophic fungi. Eur J Plant Pathol 147:431–441. https://doi.org/10.1007/s10658-016-1014-7
7. Kreft M (2016) Buckwheat phenolic metabolites in health and disease. Nutr Res Rev 29:30–39. https://doi.org/10.1017/S0954422415000190
8. Lefebvre A, Scalla R, Pfeiffer P (1990) The double-stranded RNA associated with the '447' cytoplasmic male sterility in *Vicia faba* is packaged together with its replicase in cytoplasmic membranous vesicles. Plant Mol Biol 14:477–490. doi:10.1007/BF00027494
9. Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25:1307–1320. doi:10.1093/molbev/msn067
10. Lim S, Kim KH, Zhao F, Yoo RH, Igori D, Lee SH, Moon JS (2015) Complete genome sequence of a novel endornavirus isolated from hot pepper. Arch Virol 160:3153–3156. doi:10.1007/s00705-015-2616-7

11. Okada R, Ichinose S, Takeshita K, Urayama SI, Fukuhara T, Komatsu K, Arie T, Ishihara A, Egusa M, Kodama M, Moriyama H (2018) Molecular characterization of a novel mycovirus in *Alternaria alternata* manifesting two-sided effects: Down-regulation of host growth and up-regulation of host plant pathogenicity. Virology 519:23–32. doi:10.1016/j.virol.2018.03.027
12. Okada R, Kiyota E, Moriyama H, Fukuhara T, Natsuaki T (2015) A simple and rapid method to purify viral dsRNA from plant and fungal tissue. J Gen Plant Pathol 81:103–107. https://doi.org/10.1007/s10327-014-0575-6
13. Okada R, Kiyota E, Moriyama H, Fukuhara T, Valverde RA (2017) Molecular and biological properties of an endornavirus infecting winged bean (*Psophocarpus tetragonolobus*). Virus Genes 53:141–145. https://doi.org/10.1007/s11262-016-1398-7
14. Okada R, Kiyota E, Moriyama H, Toshiyuki F, Valverde RA (2014) A new endornavirus species infecting Malabar spinach (*Basella alba* L.). Arch Virol 159:807–809. doi:10.1007/s00705-013-1875-4
15. Okada R, Kiyota E, Sabanadzovic S, Moriyama H, Fukuhara T, Saha P, Roossinck MJ, Severin A, Valverde RA (2011) Bell pepper endornavirus: molecular and biological properties, and occurrence in the genus Capsicum. J Gen Virol 92:2664–2673. doi:10.1099/vir.0.034686-0
16. Okada R, Yong CK, Valverde RA, Sabanadzovic S, Aoki N, Hotate S, Kiyota E, Moriyama H, Fukuhara T (2013) Molecular characterization of two evolutionarily distinct endornaviruses co-infecting common bean (*Phaseolus vulgaris*). J Gen Virol 94:220–229. doi:10.1099/vir.0.044487-0
17. Otulak-Koziel K, Koziel E, Escalante C, Valverde RA (2020) Ultrastructural analysis of cells from bell pepper (*Capsicum annuum*) infected with bell pepper endornavirus. Front Plant Sci 11:491. doi:10.3389/fpls.2020.00491
18. Sytar O, Brestic M, Zivcak M, Phan Tran LS (2016) The contribution of buckwheat genetic resources to health and dietary diversity. Curr Genomics 17:193–206. doi:10.2174/1389202917666160202215425
19. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. doi:10.1093/molbev/mst197
20. Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence align- ment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680. doi:10.1093/nar/22.22.4673
21. Tuomivirta TT, Kaitera J, Hantula J (2009) A novel putative virus of *Gremmeniella abietina* type B (Ascomycota: Helotiaceae) has a composite genome with endornavirus affinities. J Gen Virol 90:2299–2305. doi:10.1099/vir.0.011973-0
22. Uchida K, Sakuta K, Ito A, Takahashi Y, Katayama Y, Omatsu T, Mizutani T, Arie T, Komatsu K, Fukuhara T, Uematsu S, Okada R, Moriyama H (2021) Two novel endornaviruses co-infecting a phytophthora pathogen of asparagus officinalis modulate the developmental stages and fungicide sensitivities of the host oomycete. Front Microbiol 12:633502. doi:10.3389/fmicb.2021.633502

23. Valverde RA, Khalifa MA, Okada R, Fukuhara T, Sabanadzovic S, ICTV Report Consortium (2019) ICTV Virus Taxonomy Profile: *Endornaviridae*. J Gen Virol 100:1204–1205. doi: 10.1099/jgv.0.001277
24. Valverde RA, Sabanadzovic S, Rush MC (2011) Identification of *Oryza sativa endornavirus* in rice genotypes from breeding programmes in the United States. Plant Breeding 130:271–274. https://doi.org/10.1111/j.1439-0523.2010.01778.x

Figures

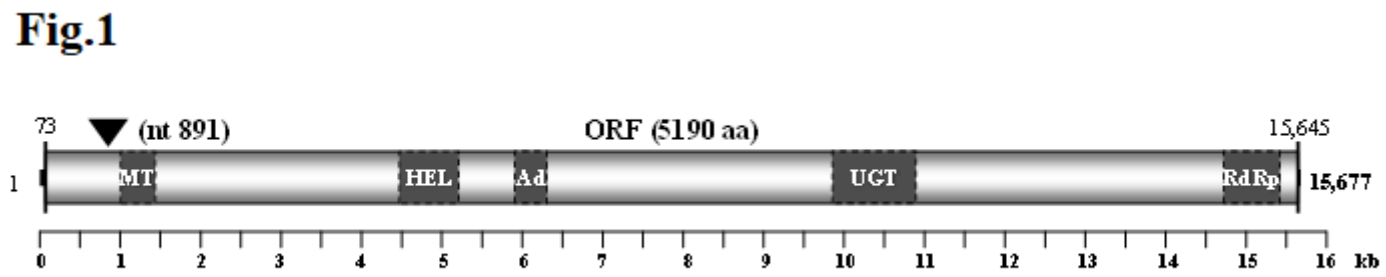

Figure 1

Schematic representation of the genome organization of Fagopyrum esculentum endornavirus 1 including the position of the nick in the plus strand (▼). The box represents the large ORF, whereas lines depict UTRs. MT, viral methyltransferase; HEL, viral RNA helicase 1; Ad, MSCRAMM family adhesin SdrC; UGT, UDP-glycosyltransferases and similar proteins; RdRp, viral RNA dependent RNA polymerase 2.

Fig.2

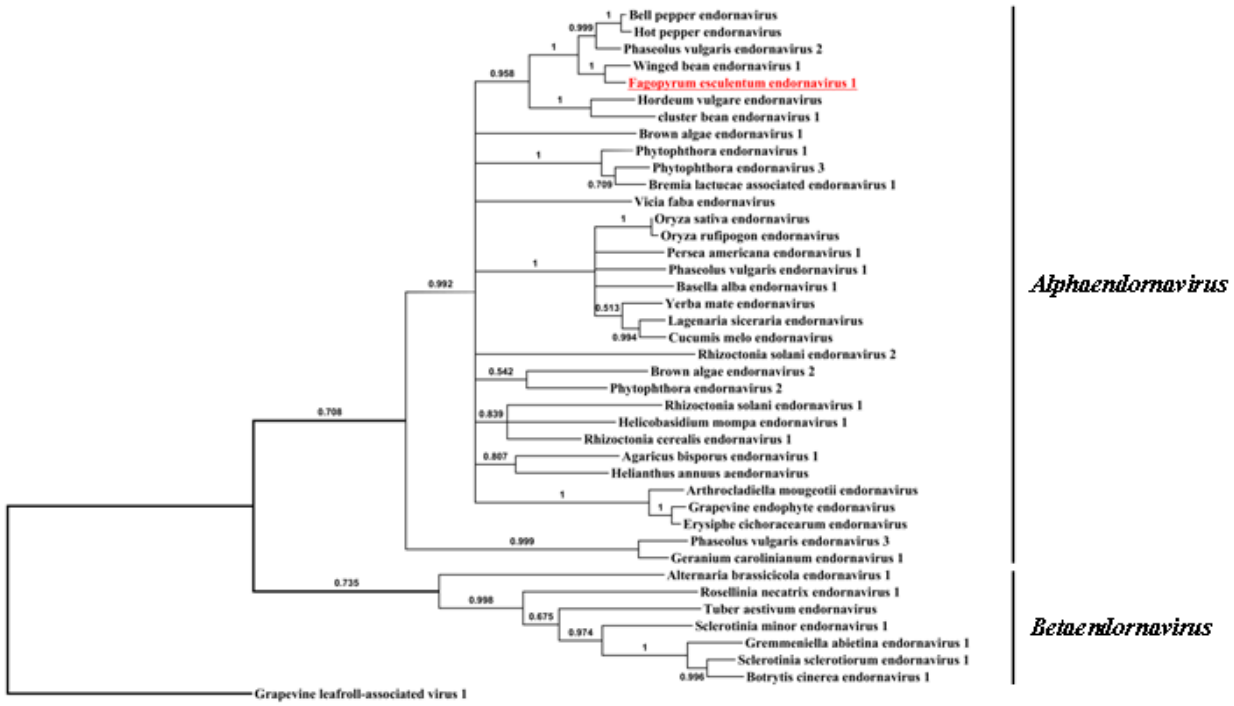

Figure 2

A Maximum Likelihood-based phylogenic tree using putative RdRp regions of Fagopyrum esculentum endornavirus 1 and related endornaviruses. The evolutionary history was inferred using the Maximum Likelihood method based on the Le and Gascuel model [9]. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 2.2569)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 4.1255% sites). Support for nodes was assessed by a reliability after 1,000 bootstrap iterations. Branches with less than 0.5 bootstrap values are collapsed. Grapevine leafroll-associated virus 1 is an ampelovirus and was used as an outgroup. GenBank accession numbers of the analyzed genes are provided in Supplementary Table 2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supptable.docx