Archives of Virology (2024) 169:24
https://doi.org/10.1007/s00705-023-05921-4

ANNOTATED SEQUENCE RECORD

Complete genome sequence of a new putative member of the genus *Pelarspovirus* that infects *Quercus aliena* Blume in South Korea

Jeong A Kwon[1,2] · Se Eun Kim[1] · Davaajargal Igori[1,3] · Jun Ki Ahn[4] · Han Kyu Seo[5] · Yang Chan Park[6] · Jae Sun Moon[1,2]

Received: 26 September 2023 / Accepted: 24 October 2023 / Published online: 11 January 2024

Abstract

The complete nucleotide sequence of a newly discovered virus infecting *Quercus aliena* Blume, tentatively named "quercus leafroll virus" (QLRV), was determined through high-throughput and Sanger sequencing. The sequence comprises 3,940 nucleotides, has five open reading frames, and has a typical pelarspovirus genome organization, with neither 3' polyadenylation nor a 5' cap. The proteins encoded by QLRV share 17.9 to 44.2% amino acid sequence identity with known pelarspovirus proteins. The highest amino acid sequence identity values for the RNA-dependent RNA polymerase (RdRp) and coat protein were 67.5% and 55.2%, respectively, which are below the current thresholds for pelarspovirus species demarcation. On the basis of these results, we propose classifying QLRV as a new member of the genus *Pelarspovirus*, family *Tombusviridae*.

Plant viruses in the genus *Pelarspovirus* (family *Tombusviridae*) are non-enveloped, 28-34 nm in diameter, have capsids with $T = 3$ icosahedral symmetry, and possess a positive-sense RNA genome approximately 3.8-3.9 kb in length. All members of the family *Tombusviridae*, with the exception of members of the genus *Dianthovirus*, have a non-segmented genome that is neither 3'-polyadenylated nor 5'-capped [4]. Viruses belonging to the genus *Pelarspovirus* exhibit a number of distinctive characteristics. To translate their five gene products, they use five distinct reading frames [5]. The five ORFs encode two replication-supporting proteins, two movement proteins (MP), and a capsid protein (CP) [5]. One subgenomic RNA contains the genetic information for both movement proteins and the capsid protein [3]. In addition, some members of the genus *Pelarspovirus* have a non-canonical start codon (GUG or CUG) that initiates the expression of the ORF coding for MP2 [5]. The absence of an AUG codon in any frame between the AUG initiation codon of MP1 and the CP open reading frame is another typical feature of pelarspoviruses [5].

Quercus aliena Blume (family Fagaceae), commonly known as Oriental white oak, is native to eastern Asia and is a major deciduous tree in Korean natural forests [2]. The acorns of this species, which is widespread in eastern Asia, contribute to basic food chains in forest ecosystems [1]. The starchy nuts of mature trees may also help to mitigate long-term climate change through oak tree carbon sequestration [6].

In 2021, oak trees exhibiting virus-like symptoms, such as leaf rolling, leaf dwarfing, and browning of new leaves were observed at an elevation of 1,150 m at Changjuk-dong, Taebaek-si, and Gangwon-do, South Korea. Symptomatic leaves were collected, ground in the laboratory using liquid nitrogen, and stored at -80 °C. A WizPrep Plant RNA Mini Kit (Wizbiosolutions, Seongnam, South Korea) was used to

Handling Editor: Sead Sabanadzovic.

✉ Yang Chan Park
yang3151@nexbio.co.kr

✉ Jae Sun Moon
jsmoon@kribb.re.kr

1 Present Address: Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 34141, Republic of Korea

2 Department of Biosystems and Bioengineering, KRIBB School of Biotechnology, Korea University of Science and Technology, Daejeon 34113, Korea

3 Department of Biology, School of Mathematics and Natural Sciences, Mongolian National University of Education, Ulaanbaatar, Mongolia

4 Material & Component Convergence R&D Department, Korea Institute of Industrial Technology (KITECH), Ansan 15588, Korea

5 HK Genomics, Inc., Daejeon, Republic of Korea

6 NEXBIO Co., Ltd., Daejeon 34520, Republic of Korea

isolate total RNA. Ribosomal RNA was depleted from total RNA using a Ribo-Zero rRNA Removal Kit (Plant Leaf) according to the manufacturer's instructions (Epicentre, Madison, WI, USA). Next, a cDNA library was produced using a TruSeq RNA Sample Prep Kit (Illumina San Diego, CA, USA) and sequenced on an Illumina Novaseq 6000 platform (Illumina) by Macrogen (Seoul, South Korea). The resulting raw reads were filtered using the Trimmomatic program (http://www.usadellab.org/cms/?page=trimmomatic) to yield 595,758,264 high-quality reads. Trimmed high-quality reads were assembled *de novo* into contig sequences using the Trinity program, and each contig was newly annotated through a BLASTx search of the NCBI GenBank database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Among the identified plant virus contigs, one contig comprising 4,023 nucleotides (nt) was similar to pelarspovirus sequences available in the database. In particular, it shared 69.92% amino acid sequence identity (query coverage: 55%) with clematis chlorotic mottle virus (ClCMV, GenBank accession no. AYI50027). The sequence of the contig corresponding to the nearly full-length genome of the virus was confirmed by RT-PCR. For this, we designed four specific primer sets overlapping by more than 100 bp, based on the 4,023-nt contig sequence and made a new total RNA extract using a WizPrep Plant RNA Mini Kit (Supplementary Table S1). Reverse transcription (RT)-PCR performed using the four specific primer sets and SuPrimeScript RT-PCR Premix (GeNet Bio, Daejeon, South Korea) yielded bands only of the expected sizes from oak. All amplicons were purified using an AccuPrepPCR Purification Kit (Bioneer, Daejeon, Korea) and cloned, and at least three clones for each RT-PCR product were sequenced by Macrogen (Seoul, South Korea). CLC Main Workbench v.21 (Insilicogen, Yongin, South Korea) was used to assemble the resultant sequences. Because pelarspoviruses do not have a poly(A) tail at the 3′ end, total RNA was adenylated using a Poly(A) Polymerase Tailing Kit (Epicentre, Madison, WI, USA) before amplification. cDNA was synthesized from total RNA using a RevertAid RT Reverse Transcription Kit (Thermo Scientific, Waltham, MA, USA) and random N25 primers, and PCR products were obtained from known sequences using AccuPower ProFi Taq PCR PreMix (Bioneer). Next, 5′-terminal sequences were generated with the aid of a 5′-rapid amplification of cDNA ends (5′-RACE) kit (Invitrogen, Waltham, MA, USA). All amplicons were also cloned into a TA cloning vector, and at least six clones were sequenced in both directions by Macrogen. The sequences were then assembled using CLC Main Workbench v.21, revealing that the newly identified virus has a RNA genome size of 3,940 nt. This virus was named "quercus leafroll virus" (QLRV), and its genome sequence was deposited in the GenBank database under accession number OR193176.

The 3,940-nt-long genome sequence of QLRV contains five ORFs and lacks a poly-A tail (Fig. 1). Pairwise comparisons of QLRV RNA sequence regions with those of viruses belonging to 36 *Tombusviridae* species, using CLUSTAL W 2.1 [9], revealed 17.9-44.2% nt sequence identity (Supplementary Table S2). The putative p27 (27-kDa) protein encoded by the first ORF, which is 702 nt (234 codons) long, shares the highest amino acid sequence identity, 52%, with the p27 protein of clematis chlorotic mottle virus (ClCMV). ORF1 at nt positions 703-705 has an amber stop codon, UAG, that is likely read through to produce a 87-kDa replicase protein (ORF2). The sequence 5′-AAA UAG GGA-3′ (nt 700-708) in QLRV, which includes the amber stop codon, is a readthrough sequence matching the pattern (A/C/U)(A/U)A **UAG**(G/C)(G/A/U) [7]. The ORF2-encoded protein is the most similar to those of ClCMV and Rosa rugosa leaf distortion virus (RrLDV), with 67% and 51% aa sequence identity, respectively. ORF3 (nt 2254-2475,) encodes a movement protein (MP1) of 73 aa that is 61% identical to the MP1 protein of pelargonium chlorotic ring pattern virus (PCRPV). The current criteria for species demarcation in the genus *Pelarspovirus* are less than 75% amino acid sequence identity in the RdRp and CP to other members of the genus [5]. Since the similarity between

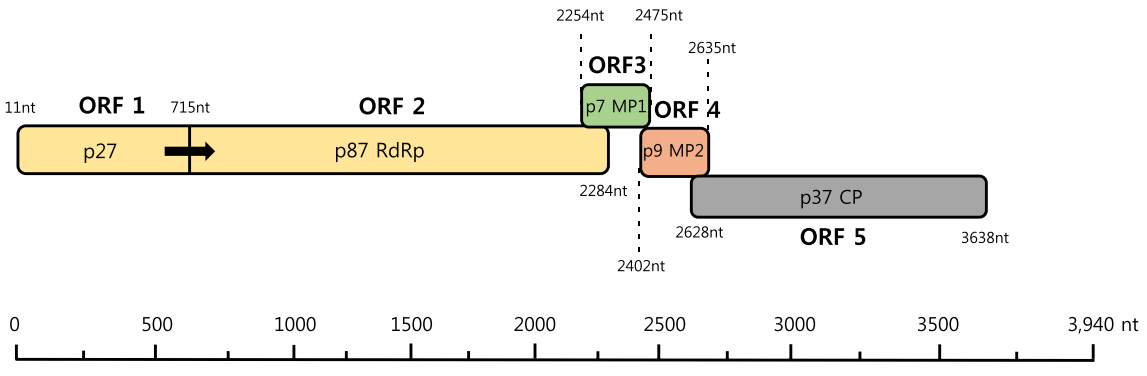

Fig. 1 Schematic representation of the genome organization of quercus leafroll virus (QLRV). ORFs 1, 2 and 4 are located in frame 2, ORF 3 is located in frame 1, and ORF 5 is located in frame 3. The arrow between ORF 1 and 2 indicates the readthrough site at an amber stop codon

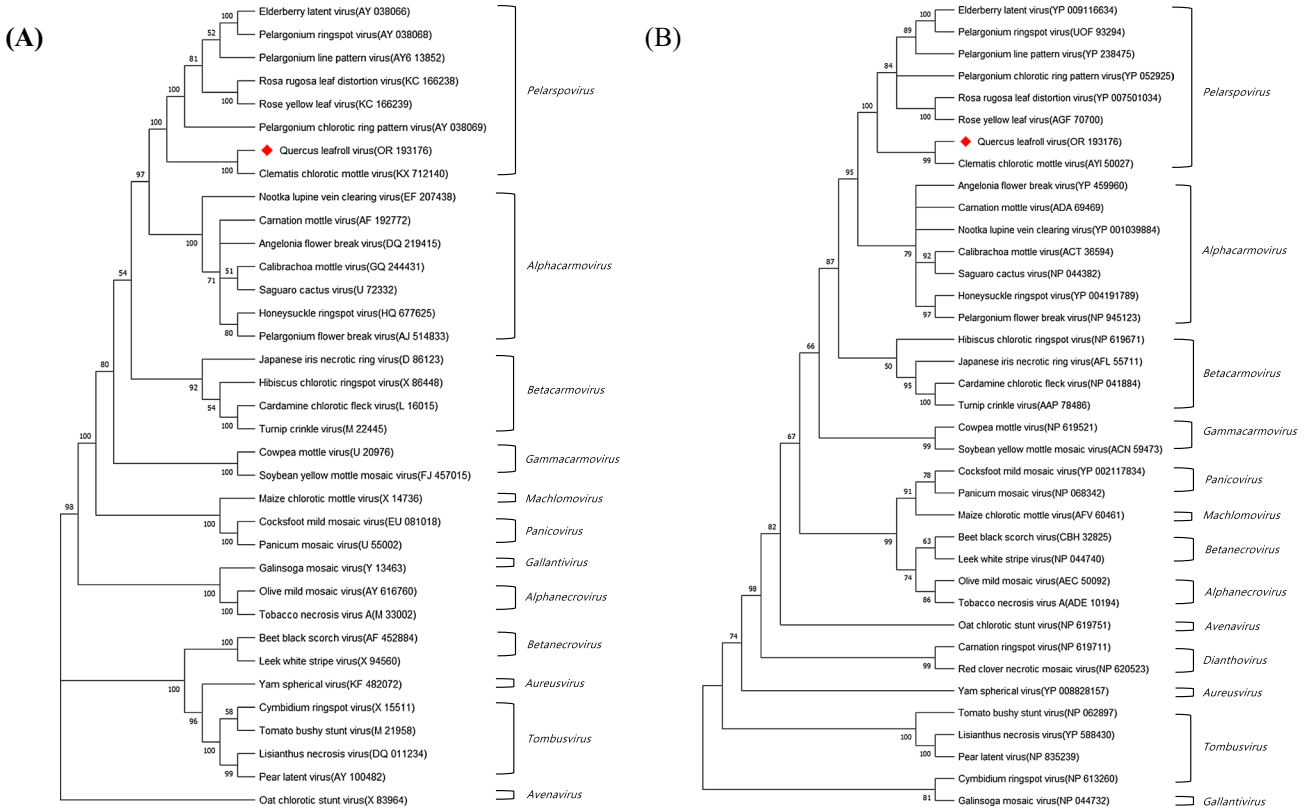

Fig. 2 Phylogenetic trees based on RdRp (A) and capsid protein (B) amino acid sequences of quercus leafroll virus (QLRV) and members of the family *Tombusviridae*. The trees were constructed by the maximum-likelihood method in MEGA 11, with 1000 bootstrap iterations. Numerical values indicate bootstrap percentage values. The accession numbers of the amino acid sequences of the proteins used to construct the tree are shown in parentheses

QLRV and other investigated pelarspoviruses is below the thresholds for both RdRp and CP, we conclude that QLRV is a representative of a new species in the genus *Pelarspovirus*.

To analyze the relationship of QLRV to other members of family *Tombusviridae*, we compared RdRp and CP amino acid sequences using MEGA 11 [8]. In the deduced phylogenetic trees shown in Fig. 2, QLRV is grouped with other members of the genus *Pelarspovirus*, again strongly suggesting that QLRV is a new member to the genus *Pelarspovirus* in the family *Tombusviridae*. This is the first report of a pelarspovirus infecting *Q. aliena* in South Korea.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00705-023-05921-4.

Acknowledgments This work was supported by IPET (Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries) and the Crop Viruses and Pests Response Industry Technology Development Program (321104-3), funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA), Republic of Korea. We thank Edanz (https://www.edanz.com/ac) for editing a draft of this manuscript.

Funding This study was funded by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET); through the Crop Viruses and Pests Response Industry Technology Development Program (321104-3), funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA), Republic of Korea. Also, we thank Barbara Goodson from Edanz (www.edanz.com/ac) for editing a draft of this manuscript.

Data availability The datasets used 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 conflict of interest.

Ethical approval This article does not contain any studies involving human participants or animals.

References

1. Chang G, Jin T, Pei J, Chen X, Zhang B, Shi Z (2012) Seed dispersal of three sympatric oak species by forest rodents in the Qinling Mountains, Central China. Plant Ecol 213:1633–1642
2. Lee CK, Lee SH, Park JH, Cho SE, Shin HD (2013) First report of oak anthracnose caused by *Apiognomonia errabunda* on oriental white Oak in Korea. Plant Dis 97:1121–1121

3. McLaughlin M, Lockhart B, Jordan R, Denton G, Mollov D (2017) Complete nucleotide sequence of clematis chlorotic mottle virus, a new member of the family *Tombusviridae*. Arch Virol 162:1373–1379
4. Rochon DRL, Russo M et al (2012) Virus taxonomy, 9th edn. Elsevier, Amsterdam
5. Scheets K, Jordan R, White KA, Hernandez C (2015) *Pelarspovirus*, a proposed new genus in the family *Tombusviridae*. Arch Virol 160:2385–2393
6. Shawn Overstreet SC, Chan-Ryul Park, Dowon Lee, Thomas Gradziel (2015) Acorn Production and Utilization in the Republic of Korea. In: Standiford RBP, Kathryn L., tech. cords. (ed). Gen. Tech. Rep. PSW-GTR-251. Berkeley, CA:U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station., Proceedings of the 7th California Oak Symposium: Managing Oak Woodlands in a Dynamic World, pp 265–271
7. Skuzeski JM, Nichols LM, Gesteland RF (1990) Analysis of leaky viral translation termination codons in vivo by transient expression of improved beta-glucuronidase vectors. Plant Mol Biol 15:65–79
8. Tamura K, Stecher G, Kumar S, Battistuzzi FU (2021) MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027
9. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680