Genome Sequence of Grapevine Virus K, a Novel Vitivirus Infecting Grapevine

Yeonhwa Jo,ᵃ Myung-Kyu Song,ᵇ Hoseong Choi,ᵃ Jae-Seong Park,ᵇ Jae-Wung Lee,ᵇ Sen Lian,ᶜ Bong Choon Lee,ᵈ Won Kyong Choᵃ

Department of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Koreaᵃ; Grape Research Institute, Chungbuk Agricultural Research and Extension Services, Okcheon, Republic of Koreaᵇ; College of Crop Protection and Agronomy, Qingdao Agricultural University, Qingdao, Shandong, Chinaᶜ; Crop Foundation Division, National Institute of Crop Science, RDA, Wanju, Republic of Koreaᵈ

ABSTRACT Here, we report the genome sequence of grapevine virus K (GVK), a novel single-stranded RNA virus identified from a transcriptome of grapevine. The genome of GVK is 7,476 nucleotides in length and encodes 5 open reading frames. GVK is a putative member of the genus *Vitivirus* in the family *Betaflexiviridae*.

Address correspondence to Won Kyong Cho, wonkyong@gmail.com.

The grapevine (*Vitis vinifera*) is a commercially important fruit crop worldwide that is used for wine production. Grapevines are clonally propagated by cutting and grafting. Thus, the infection rate of diverse viruses and viroids in grapevine is usually very high. To date, more than 70 different viruses infecting grape plants have been identified (1).

Recently, many novel viruses infecting grapevine have been identified with the help of next-generation sequencing techniques (2–4). Moreover, several recent studies have found that much plant transcriptome data contain viral sequences, which sometimes enables the assembly of the draft genome of a virus or viroid (5–7).

In the search for viruses and viroids from grapevine transcriptomes, we identified a novel virus from a transcriptome of the grapevine cultivar Lambrusco salamino (8). To summarize, three individual samples were pooled and used for total mRNA extraction. A cDNA library was prepared using the Illumina TruSeq RNA kit and single-end (85 bp) sequenced by the Illumina GAIIx platform. The raw data (3.3 Gb) for grapevine cultivar Lambrusco salamino deposited in the Sequence Read Archive (SRA) database (accession number ERR922631) were downloaded. The raw sequence reads were *de novo* assembled with Trinity version 2.0.6 with default parameters (9). The assembled contigs were analyzed with a BLAST search against the viral reference database (https://www .ncbi.nlm.nih.gov/genome/viruses).

We identified 143 contigs associated with viruses and viroids. Of these, three contigs were associated with a novel virus referred to as grapevine virus K (GVK) isolate Jo (GenBank accession number MF072319). GVK is a positive single-stranded RNA virus with a length of 7,476 nucleotides (nt). The obtained GVK genome sequence is a nearly full-length genome containing sequences for the 5′ and 3′ untranslated regions (80 and 77 nt in length, respectively). GVK encodes five open reading frames (ORFs), ORF1 (RNA-dependent RNA polymerase), ORF2 (unknown protein), ORF3 (movement protein), ORF4 (coat protein), and ORF5 (nucleic acid-binding protein). A BLASTp search using five GVK ORFs against NCBI's nonredundant protein database revealed that ORF1 of GVK was homologous to the known grapevine virus A (GVA) with 52% protein sequence identity. No protein showed sequence similarity to ORF2 of GVK. ORF3 and ORF4 for GVK showed sequence similarity to GVA (50% protein identity) and grapevine virus D (GVD) (97% protein identity). The ORF5 of GVK showed sequence similarity to GVD (96% protein identity) with the *E* value (9e⁻⁵⁹). Based on the BLASTp results, GVK

could be a potential member of the genus *Vitivirus* in the family *Betaflexiviridae* along with GVA and GVD. Taken together, we identified and *de novo* assembled a novel virus referred to as GVK, a putative member of the genus *Vitivirus* from the grapevine transcriptome.

Accession number(s). The full-genome sequence of grapevine virus K isolate Jo has been submitted to GenBank under the accession number MF072319.

REFERENCES

1. Martelli GP. 2014. Directory of virus and virus-like diseases of the grapevine and their agents. J Plant Pathol 96:1–136. https://doi.org/10.4454/JPP.V96I1SUP.
2. Al Rwahnih M, Daubert S, Golino D, Rowhani A. 2009. Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. Virology 387:395–401. https://doi.org/10.1016/j.virol.2009.02.028.
3. Giampetruzzi A, Roumi V, Roberto R, Malossini U, Yoshikawa N, La Notte P, Terlizzi F, Credi R, Saldarelli P. 2012. A new grapevine virus discovered by deep sequencing of virus- and viroid-derived small RNAs in Cv Pinot gris. Virus Res 163:262–268. https://doi.org/10.1016/j.virusres.2011.10.010.
4. Al Rwahnih M, Alabi OJ, Westrick NM, Golino D, Rowhani A. 2017. Description of a novel monopartite geminivirus and its defective subviral genome in grapevine. Phytopathology 107:240–251. https://doi.org/10.1094/PHYTO-07-16-0282-R.
5. Jo Y, Choi H, Kim SM, Kim SL, Lee BC, Cho WK. 2016. Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for *Apple stem grooving virus*. BMC Genomics 17:579. https://doi.org/10.1186/s12864-016-2994-6.
6. Jo Y, Choi H, Cho JK, Yoon JY, Choi SK, Cho WK. 2015. In silico approach to reveal viral populations in grapevine cultivar Tannat using transcriptome data. Sci Rep 5:15841. https://doi.org/10.1038/srep15841.
7. Jo Y, Choi H, Kim SM, Kim SL, Lee BC, Cho WK. 2017. The pepper virome: natural co-infection of diverse viruses and their quasispecies. BMC Genomics 18:453. https://doi.org/10.1186/s12864-017-3838-8.
8. Potenza E, Racchi ML, Sterck L, Coller E, Asquini E, Tosatto SC, Velasco R, Van de Peer Y, Cestaro A. 2015. Exploration of alternative splicing events in ten different grapevine cultivars. BMC Genomics 16:706. https://doi.org/10.1186/s12864-015-1922-5.
9. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. 2013. *De novo* transcript sequence reconstruction from RNA-Seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512. https://doi.org/10.1038/nprot.2013.084.