


Complete nucleotide sequence and genome organization of peach virus D, a putative new member of the genus *Marafivirus*

Davaajargal Igori^{1,2}  · Seungmo Lim^{1,2} · Dasom Baek^{1,2} · San Yeong Kim³ · Euncheol Seo³ · In-Sook Cho⁴ · Gug-Seoun Choi⁴ · Hyoun-Sub Lim⁵ · Jae Sun Moon^{1,2}

Received: 11 November 2016 / Accepted: 17 January 2017 / Published online: 10 February 2017
© Springer-Verlag Wien 2017

Abstract The complete nucleotide sequence of peach virus D (PeVD) from *Prunus persica* was determined. The PeVD genome consists of 6,612 nucleotides excluding the 3' poly(A) tail and contains a single open reading frame coding for a polyprotein of 227 kDa. Sequence comparisons and phylogenetic analysis revealed that PeVD is most closely related to viruses in the genus *Marafivirus*, family *Tymoviridae*. The complete nucleotide and CP amino acid sequences of PeVD were most similar (51.1–57.8% and 32.2–48.0%, respectively) to members of the genus *Marafivirus*, suggesting that PeVD is a new member of this genus.

The family *Tymoviridae* comprises three genera (*Tymovirus*, *Marafivirus* and *Maculavirus*) of plant-infecting single-stranded positive-sense viruses that are grouped based on the similarity of their replication-associated polyproteins (RPs), which contain conserved domains for methyltransferase (MTR), papain-like protease (PRO), helicase (HEL) and RNA-dependent RNA polymerase (RdRp) [11]. The marafivirus genome is distinguished by the presence of a single long open reading frame (ORF) encoding an RP along with two coat proteins (CPs) and a conserved 16-nucleotide marafibox, the latter being analogous to the tymobox of tymoviruses, from which it differs by only two or three residues [2, 3, 5, 7, 11].

Peach (*Prunus persica*) is an economically important and widely grown deciduous tree in the family Rosaceae. Peach trees can be infected by a number of viruses and viroids, the most notable being apple chlorotic leaf spot virus (ACLSV), prunus necrotic ringspot virus (PNRSV), apricot pseudo chlorotic leaf spot virus (APCLSV), plum pox virus (PPV), plum bark necrosis stem pitting-associated virus (PBNSPaV), prune dwarf virus (PDV), tomato ringspot virus (ToRSV), peach latent mosaic viroid (PLMVd), and hop stunt viroid (HSVd). These viruses and viroids are members of the following genera and families: ACLSV and APCLSV, genus *Trichovirus* of family *Betaflexiviridae*; PNRSV and PDV, *Ilarvirus* (*Bromoviridae*); PPV, *Potyvirus* (*Potyviridae*); PBNSPaV, *Ampelovirus* (*Closteroviridae*); ToRSV, *Nepovirus* (*Secoviridae*); PLMVd, *Pelamoviroid* (*Avsunviroidae*); and HSVd, *Hostuviroid* (*Pospiviroidae*) [10, 13]. Nectarine virus M (NeVM), which belongs to the genus *Marafivirus* (family *Tymoviridae*), has recently been reported to infect nectarine (*P. persica* var. *nectarina*) [13]. Here, we report the complete genome sequence of peach virus D (PeVD), a new marafivirus infecting peach in South Korea.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-017-3255-y) contains supplementary material, which is available to authorized users.

✉ Jae Sun Moon
jsmoon@kribb.re.kr

- ¹ Biosystems and Bioengineering Program, University of Science and Technology, Daejeon 34113, Republic of Korea
- ² Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Republic of Korea
- ³ Cheongdo Peach Research Institute, Gyeongsangbuk-do Agricultural Research and Extension Services, Cheongdo 38315, Republic of Korea
- ⁴ Horticultural and Herbal Crop Environment Division, National Institute of Horticultural and Herbal Science, Rural Development Administration, Wanju 55365, Republic of Korea
- ⁵ Department of Applied Biology, Chungnam National University, Daejeon 34134, Republic of Korea

Leaves exhibiting symptoms of yellowing and slight mottling were collected from a yard-grown peach tree in Gyeongsangbuk-do, South Korea, in May 2015. Total RNA was extracted from the symptomatic leaves using a Wiz-Prep Plant RNA Mini Kit (Wizbiosolutions, Seongnam, South Korea). Plant ribosomal RNA (rRNA) was eliminated from the total RNA using a Ribo-Zero Plant Leaf rRNA Removal Kit (Epicentre, Madison, WI, USA). A library was constructed from an rRNA-depleted RNA sample using a TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA) and then sequenced at the Theragen Bio Institute (Suwon, South Korea) on an Illumina HiSeq 2500 instrument. Raw sequence reads were quality-filtered and assembled *de novo* using the software package Trinity v2.1.1. Assembled contigs were screened using BLAST to identify sequence similarities to other viral genomes in the GenBank database. Among the assembled contigs, two long contigs of 2,466 and 4,056 bp were related to marafiviruses and were designated as PeVD contigs (Fig. 1B). The 2,466- and 4056-bp contigs were assembled from 4700 reads (depth of coverage across the contig: maximum, 363; minimum, 1; mean, 189) and 3276

reads (maximum, 529; minimum, 1; mean, 79), respectively (Fig. 1A). BLASTx analysis showed that the PeVD contigs had the highest amino acid sequence similarity (56–61% identity) to NeVM (genus *Marafivirus*; GenBank accession no. KT273413). To confirm the PeVD contigs obtained by high-throughput paired-end sequencing, specific primer sets were designed based on the contig sequences (Supplementary Table 1). Total RNA was isolated from leaves exhibiting virus-like symptoms using a WizPrep Plant RNA Mini Kit. cDNA was synthesized from total RNA using RevertAid reverse transcriptase (Thermo Scientific, USA) with the 25-mer degenerate primer N25 and then amplified by PCR using AccuPower ProFi *Taq* PCR PreMix (Bioneer, Daejeon, South Korea) and the specific primer sets [4, 6]. All designed primer sets successfully produced overlapping PCR fragments of the expected size from the peach sample (Fig. 1C). The PCR products were purified and cloned into an RBC TA cloning vector (RBC Bioscience, Taipei, Taiwan), and the resulting plasmids were sequenced by GenoTech (Daejeon, South Korea). The generated sequences were consistent with those of the original contigs. The peach sample was also

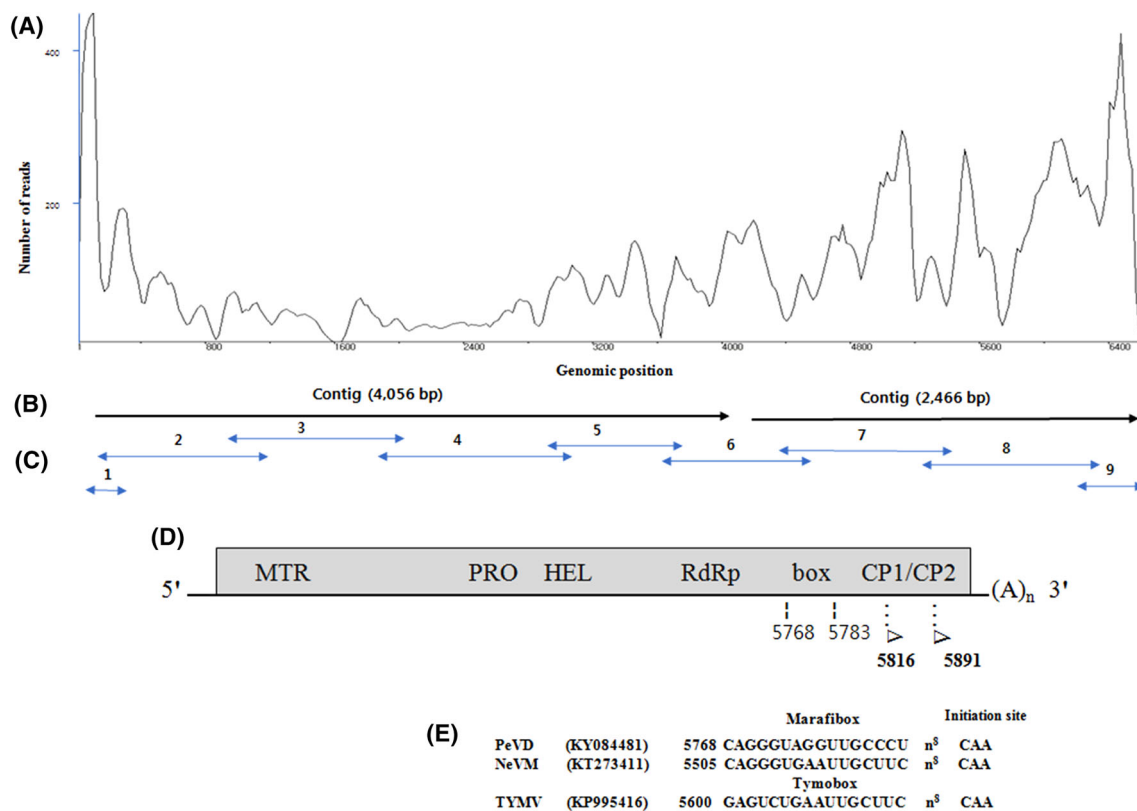


Fig. 1 (A) Coverage of the sequence reads on the complete genome sequence of PeVD. (B) Contig positions on the full-length genome. (C) Stepwise overlapping PCR fragments generated for sequence confirmation. (D) predicted genome organization of peach virus D (PeVD). The gray-shaded box represents a single open reading frame

(ORF) and the corresponding putative polyprotein. MTR, methyltransferase; PRO, papain-like protease; HEL, helicase; RdRp, RNA-dependent RNA polymerase; CP, coat protein; box, marafibox. (E) Sequence alignment of the PeVD 16-nucleotide marafibox and initiation site with the corresponding regions of NeVM and TYMV

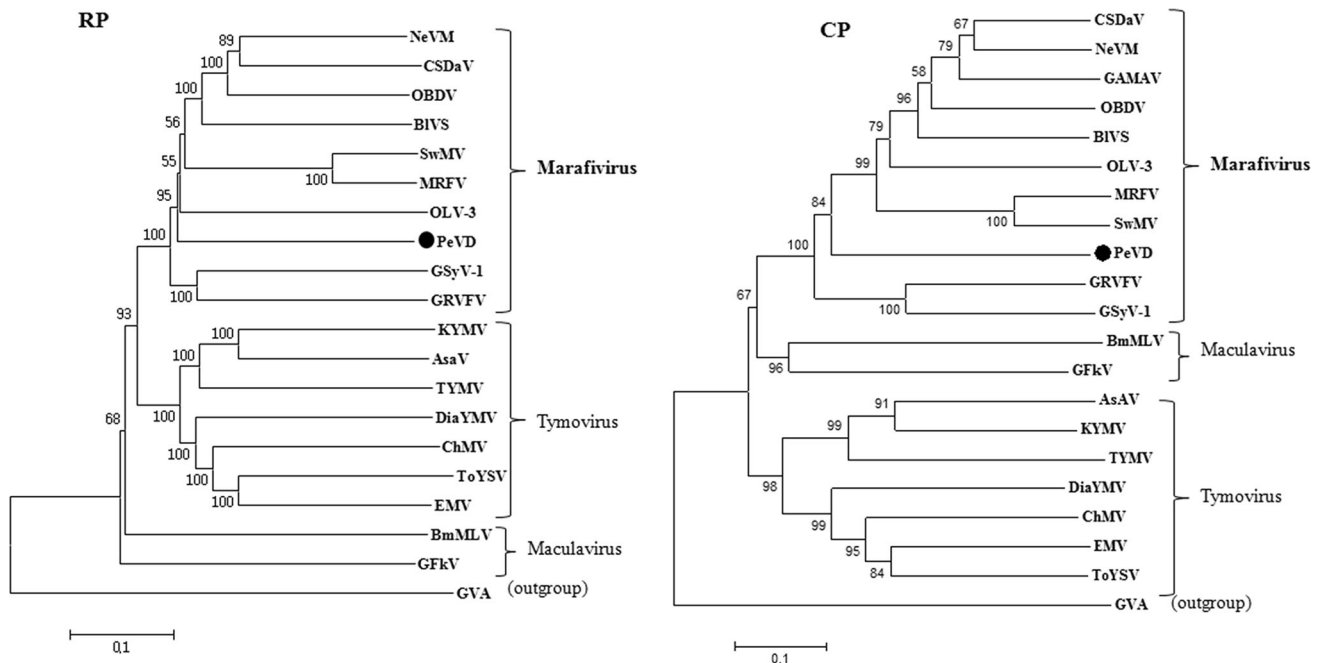


Fig. 2 Phylogenetic trees constructed from amino acid sequences of the replication protein (RP) and coat protein (CP) of peach virus D (PeVD) and viruses of the *Tymoviridae* genera *Tymovirus*, *Marafivirus* and *Maculavirus* using neighbor-joining method in MEGA6.0.

confirmed to be infected by PeVD. To complete the genomic sequence of PeVD, 5'- and 3'-terminal sequences were obtained using a 5'- and 3'-Rapid Amplification of cDNA Ends kit (RACE version 2.0; Invitrogen, Carlsbad, CA, USA) [9]. All overlapping fragments were assembled using the program DNAMAN v5.2.10 (Lynnon Biosoft, Quebec, Canada) [9]. The full sequence of PeVD was submitted to GenBank under accession number KY084481. Phylogenetic relationships of PeVD and other members of the family *Tymoviridae* were inferred from their predicted RP and CP amino acid sequences using the neighbor-joining method with 1,000 bootstrap replicates in MEGA v6.06 [12]. NC_003604 (grapevine virus A [GVA], genus *Vitivirus*, family *Betaflexiviridae*) was used as an outgroup sequence.

The complete genome of PeVD, excluding the 3' poly(A) tail, was found to comprise 6,612 nucleotides (nt) including a 295-nt 5' untranslated region (UTR) and a 3' UTR of 117 nt. Similar to other marafiviruses, the PeVD genome (GenBank accession no. KY084481) was found to contain a large ORF beginning with an AUG codon at nt positions 296–298 and ending with a UAA termination codon at positions 6,494–6,496. This ORF encodes a large precursor polyprotein of 2,066 amino acids (aa) with a predicted molecular mass of 226.5 kDa (p227) possessing a number of conserved domains such as viral methyltransferase (Pfam 01660), tymovirus endopeptidase (Pfam 05381), viral (superfamily 1) RNA helicase (Pfam 01443),

Support for nodes in the trees was assessed by bootstrapping with 1,000 replicates. Grapevine virus A (GVA; genus *Vitivirus*, family *Betaflexiviridae*) was used as an outgroup. The accession numbers and full virus names and are listed in Supplementary Table 3

RdRp (pfam 00978) and tymovirus CP (Pfam 00983) as well as putative CPs with deduced molecular masses of 23.7 and 21.2 kDa (Fig. 1D). Pairwise alignment of the complete nucleotide sequence of PeVD with other members of the family *Tymoviridae* revealed that PeVD has more nucleotide sequence similarity to marafiviruses (51.1–57.8% identity) than to tymoviruses (48.5–50.8%) or maculaviruses (44.4–47.6%) (Supplementary Table 2). In addition, the highest aa sequence similarity of the RP of PeVD was with marafiviruses (45.6–53.5%), followed by tymoviruses (40.8–43.9%) and maculaviruses (37.8–40.1%). When compared separately, the individual conserved domain regions MTR (aa 37–318), PRO (aa 768–864), Hel (aa 961–1192) and RdRp (aa 1504–1740) of PeVD were always more similar to those of marafiviruses than to those of other tymoviruses and maculaviruses (Supplementary Table 2) [1]. MTR and RdRp of PeVD were most similar (68.1% and 78.1% identity, respectively) to the corresponding domains of oat blue dwarf virus (OBDV; genus *Marafivirus*). PeVD PRO was most closely related (54.6%) to that of olive latent virus 3 (OLV-3; *Marafivirus*), while PeVD HEL was most similar (69.0%) to NeVM (*Marafivirus*). The aa sequence identities between the putative CP of PeVD and comparable proteins of other members of the family *Tymoviridae* ranged from 21.2% with turnip yellow mosaic virus (TYMV; *Tymovirus*) to 48.0% with NeVM [3]. A putative marafibox of PeVD was identified at nt positions

5,768–5,783, with a putative subgenomic RNA transcription site start site of CAA located at positions 5,792–5,794 [2, 3, 5, 7]. Multiple sequence alignment of the PeVD 16-nt marafibox and initiation site with corresponding regions of NeVM and TYMV revealed that the PeVD marafibox differs by six and nine nt from marafibox and tymobox sequences, respectively (Fig. 1E). It is a divergent marafibox, probably because the complete genome nt and CP aa sequences of PeVD show only 51.1–57.8% and 32.2–48.0% identity, respectively, to members of the genus *Marafivirus*. Phylogenetic analysis of aa sequences of RP and CP of PeVD and other members of family *Tymoviridae* revealed that PeVD is more closely related to marafiviruses than to members of other virus genera (Fig. 2) [1, 3]. The complete nt sequence identities between the genomes of PeVD and other members of the family *Tymoviridae* ranged from 44.4% (BmMLV; *Maculavirus*) to 57.8% (NeVM; *Marafivirus*), while the CP aa sequence identities varied from 21.2% (TYMV; *Tymovirus*) to 48.0% (NeVM) (Supplementary Table 2). According to current species demarcation criteria, members of distinct species in the genus *Marafivirus* should have less than 80% overall sequence identity and less than 90% capsid protein sequence identity to other members of the genus [8, 11]. Our results thus suggest that PeVD should be regarded as a member of a distinct species in the genus *Marafivirus*.

Acknowledgements This research was carried out with the support of “Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01252403),” Rural Development Administration, Republic of Korea.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with animals or human participants performed by any of the authors.

References

- Alabdullah A, Minafra A, Elbeaino T, Saponari M, Savino V, Martelli GP (2010) Complete nucleotide sequence and genome organization of Olive latent virus 3, a new putative member of the family *Tymoviridae*. *Virus Res* 152:10–18
- Edwards MC, Weiland JJ (2014) Coat protein expression strategy of oat blue dwarf virus. *Virology* 450:290–296
- Edwards MC, Zhang Z, Weiland JJ (1997) Oat blue dwarf marafivirus resembles the tymoviruses in sequence, genome organization, and expression strategy. *Virology* 232:217–229
- Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. *Gene* 25:263–269
- Hammond RW, Ramirez P (2001) Molecular characterization of the genome of *Maize rayado fino virus*, the type member of the genus *Marafivirus*. *Virology* 282:338–347
- Igori D, Hwang U, Lim S, Zhao F, Kwon S-Y, Moon J (2016) The first complete sequence and genome structure of daphne virus Y. *Arch Virol* 161:2905–2908
- Izadpanah K, Ping Zhang Y, Daubert S, Masumi M, Rowhani A (2002) Sequence of the coat protein gene of Bermuda grass etched-line virus, and of the adjacent ‘marafibox’ motif. *Virus Genes* 24:131–134
- King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2012) Virus taxonomy: ninth report of the international committee on taxonomy of viruses. Elsevier Academic Press, London, UK
- Lim S, Igori D, Zhao F, Yoo R, An T-J, Lim H-S, Lee S-H, Moon J (2015) Complete genome sequence of a tentative new caulimovirus from the medicinal plant *Attractylodes macrocephala*. *Arch Virol* 160:3127–3131
- Martelli GP, Flores R, Schneider B (2011) Classification of pome and stone fruit viruses, viroids and phytoplasmas. In: Hadidi A, Barba M, Candresse T, Jelmann J (eds) *Virus and virus-like diseases of pome and stone fruits*. The American Phytopathological Society, St. Paul, pp 13–16
- Martelli GP, Sabanadzovic S, Abou-Ghanem Sabanadzovic N, Edwards MC, Dreher T (2002) The family *Tymoviridae*. *Arch Virol* 147:1837–1846
- 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
- Villamor DEV, Mekuria TA, Pillai SS, Eastwell KC (2016) High-throughput sequencing identifies novel viruses in nectarine: insights to the etiology of stem-pitting disease. *Phytopathology* 106:519–527