ORIGINAL ARTICLE

First viruses infecting cockspur coral tree (*Erythrina crista-galli* L.): discovery of a putative novel capillovirus and a new host for the prune dwarf virus

Dijana Škorić[1] · Dorotea Grbin[1] · Martina Šeruga Musić[1] · Johan Rollin[2,3] · Sébastien Massart[2]

Abstract

A cockspur coral tree (*Erythrina crista-galli*, Fabaceae) from the collection of woody ornamentals of the Botanical Garden (Faculty of Science) in Zagreb showed conspicuous virus-like symptoms. The leaf tissue was analyzed by high throughput sequencing (HTS), revealing the presence of two viruses: prune dwarf virus (PDV, genus *Ilarvirus*) and an unknown virus belonging to the genus *Capillovirus* (family *Betaflexiviridae*). The complete sequence of PDV RNA3 of 2,129 nucleotides (nts) coding for the coat and movement proteins was obtained. The complete coding region spanning 6,483 nts was obtained for the capillovirus. It contained all expected open reading frames, and its maximum nucleotide identity was 42% with apple stem grooving capillovirus sequence (GenBank accession number LC143387). As this is far below the threshold for species delineation within the genus *Capillovirus*, we propose the name Erythrina capillovirus (ErCV) for this putatively new virus and a new virus species *Capillovirus ErCV* in the family *Betaflexiviridae*. RT-PCR confirms the presence of ErCV and PDV in symptomatic *E. crista-galli*, an unusual and exotic fabaceous host with no viruses recorded yet. Asymptomatic *Erythrina* plants were tested negative for the two viruses in RT-PCR, which, together with their presence in the symptomatic plant, suggests the capillovirus and/or PDV might play a role in eliciting observed symptoms and deserve further investigations.

Introduction

Erythrina crista-galli L. (Fabaceae), also known as the cockspur coral tree, is a deciduous woody plant flowering from late spring to early autumn with aesthetically pleasing

✉ Dijana Škorić
dijana.skoric@biol.pmf.hr

Dorotea Grbin
dorotea.grbin@biol.pmf.hr

Martina Šeruga Musić
martina.seruga.music@biol.pmf.hr

Johan Rollin
johan.rollin@doct.uliege.be

Sébastien Massart
sebastien.massart@uliege.be

[1] University of Zagreb, Faculty of Science, Department of Biology, Zagreb, Croatia

[2] Plant Pathology Laboratory, TERRA-Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium

[3] DNA Vision, Charleroi, Belgium

deep red flowers arranged in the inflorescences called racemes (Fig. 1a). The plant originated from wet subtropical habitats of South America (south Brazil, Argentina, Uruguay, Paraguay) and spread to the rest of the world as an ornamental tree. It is a good nitrogen fixator, as all plants in its family, but it is also rich in lectins and polyphenols interesting for the pharmaceutical industry. It has been used by people native to its area of origin for making folk remedies as it has anti-inflammatory and anti-microbial properties (Miño et al. 2002; Mitscher et al. 1988). To our knowledge, apart from the research on fungal endophytes contributing to its medicinal properties (Weber et al. 2005), there is no information about viruses, or any other microbe adversely affecting *E. crista-galli*.

Prune dwarf virus (PDV) is a common virus infecting various hosts in the plant genus *Prunus* grown for stone fruit production or as horticultural trees (Fulton 1970; Pallas et al. 2012). It is also able to infect wild hosts belonging to the *Prunus* genus. *Prune dwarf virus* is a species belonging to the genus *Ilarvirus* (family *Bromoviridae*). Ilarviruses have tripartite single-stranded plus RNA (ss+RNA)

Fig. 1 *Erythrina crista-galli* L. (Fabaceae) plants from the Botanical Garden of the Faculty of Science University of Zagreb: **a**) without symptoms and **b**) a stunted plant (specimen 6261J) with white striations, speckles and bleaching of the leaves and white striations of the flowers

genomes (Bujarski et al. 2012; Pallas et al. 2013). Genomic fragments are encapsidated in separate virus particles having icosahedral or bacilliform shapes (Fulton 1970; Pallas et al. 2013). RNA1 (3374 nucleotides, nts) and RNA2 (2593 nts) are monocistronic, having open reading frames (ORFs) coding for P1 and P2 protein, respectively (Ramptish and Eastwell 1997; Scott et al. 1998). RNA3 is bicistronic (2129 nts), containing ORFs for PDV movement (MP) and capsid proteins (CP) separated by a short intergenic region of about 70 nts (Bachman et al. 1994; Predajňa et al. 2017). A fourth subgenomic RNA (RNA4a) is packaged in viral particles with RNA3 (Bujarski et al. 2012; Pallas et al. 2013). Currently, the International Committee on Taxonomy of Viruses (ICTV) classification of PDV within its genus relies on the RNA3 conserved sequences of MP and CP (Simkovich et al. 2021). The CP and MP sequences have also been used for PDV detection and variability studies over the years (Massart et al. 2008; Predajňa et al. 2017; Ulubaş-Serçe et al. 2009; Vašková et al. 2000).

Despite its economic importance, PDV biology is far less studied than the other prominent bromoviruses like alfalfa mosaic virus (AMV), brome mosaic virus (BMV) or cucumber mosaic virus (CMV) (Kozieł et al. 2017). Nevertheless,

it has been known for a long time that PDV can be transmitted by vegetative propagation (grafting), seeds and pollen (Fulton 1970; Mink and Aichele 1984; Mink 1993). Along with the plant dwarfing phenotype, PDV may cause chlorosis and necrosis on leaves and malformations which may vary according to the plant host. In stone fruit production, yield can be reduced, and the fruits can be deformed due to PDV infection. Symptomless infections have also been recorded depending on the host, viral isolate or environmental conditions (Çağlayan et al. 2011; Fonseca et al. 2005). The host range of PDV is large, including species from the family Fabaceae (Fulton 1970; ICTVdB 2006). However, it has never been reported from *Erythrina crista-galli*, similar ornamental or endemic trees of the same genus like horticulturally important *E. variegata* (Wistler and Elevitch 2006).

Genus *Capillovirus* is one of the 15 currently recognized genera of *Betaflexiviridae* (Adams et al. 2012; Silva et al. 2022). It is a diverse family encompassing viruses forming long flexuous particles with monopartite polyadenylated ss+RNA genomes with short 5'- and 3'-untranslated regions (UTRs). Capilloviruses, as other *Betaflexiviridae* members, encode a long polyprotein performing the viral replication (Rep) with methyltransferase (Met), helicase (Hel) and RdRP

domains. Also, much shorter CP and MP proteins are encoded in the 3' downstream genome portion. The genus *Capillovirus* belongs to the subfamily *Trivirinae* whose genome organization entails the fused ORFs for Rep and CP whilst MP is in a different frame and covers the 3'-portion of the polyprotein and a small 5'-portion of CP ORFs (Adams et al. 2012). Apple stem grooving virus is one of the better-known capilloviruses. However, other viruses infecting pome, stone, citrus and other fruit tree species such as cherry virus A, citrus tatter leaf virus or mume virus A (MuVA) belong to this growing viral genus (Adams et al. 2012; Marais et al. 2018).

Capilloviruses have no known vectors, their host ranges are mostly limited to the original woody hosts, and the main transmission routes appear to be by grafting and seeds. The biology of each capillovirus is not equally known or clearly linked to the disease symptoms in the cultivated plant hosts of their origin (Adams et al. 2012). There is a surge of new virus discoveries from HTS studies nowadays. MuVA is a good example of such a new capillovirus discovered recently from *Prunus mume* also known as Japanese apricot (Marais et al. 2018). Novel virus species are sometimes difficult to classify under current taxonomic criteria (Adams et al. 2012). Thus, a recently proposed set of new criteria was proposed using *Betaflexiviridae* as a case study (Silva et al. 2022). Even though they are still not a part of the ICTV guidelines, it represents good tool for demarcating insofar-borderline species, as well as accommodating novel ones such as the putative new *Capillovirus* member described here.

In this study, one out of three *E. crista-galli* small trees belonging to the plant collection of the Botanical Garden of the Faculty of Science University of Zagreb, was found to have conspicuous virus-like symptoms most prominent in the leaves and flowers (Fig. 1b). Therefore, the presence and impact of viral pathogens in leaf samples was investigated by high throughput sequencing (HTS) and RT-PCR. Here, we describe a partial genomic sequence of PDV from symptomatic *E. crista-galli*, as well as almost a full genomic sequence of a proposed novel capillovirus.

Materials and methods

Plant samples

Three *Erythrina crista-galli* L. (Fabaceae) plants are part of the Botanical Garden (Faculty of Science University of Zagreb) collection. Out of three plants in the collection, one specimen (6261J, Melbourne 84) was symptomatic (Fig. 1b). The leaves had been initially sampled for virus screening in 2010 but no viruses were found, and symptoms persisted. Plants were sampled again in 2017 for HTS. PDV RT-PCR was done after the initial HTS data analysis in 2017-18 and repeated in 2022 when resampling was done for additional capillovirus targeted RT-PCR.

HTS sample preparation and sequencing

Fresh *E. crista-galli* symptomatic leaves (500 mg) were used (two samples from symptomatic plant) for the total nucleic acids extraction applying modified CTAB-protocol (Šeruga Musić et al. 2003). Nucleic acids were resuspended in 50 microliters of sterile nuclease free water and an aliquot of 20 microliters was treated with DNase (Promega, Madison, WI, USA) according to the manufacturer's instructions. After quality checking by NanoDrop 2000c spectrophotometer (Thermo Scientific, Carlsbad, CA, USA) and Bioanalyzer (Agilent, Santa Clara, CA, USA), the RNAs from *E. crista-galli* was used for the sequencing library preparation as previously described (Hanafi et al. 2020). HTS was performed on Illumina Nextseq 500 platform yielding a total number of 4.2 million reads (2×150 nts).

HTS data analysis and virus detection

The Geneious Prime software 2022.2.1 (Biomatters Ltd., Auckland, New Zealand) was used for the sequence analysis. The RNA-seq reads were paired, merged by using BBMerge from the BBtools suite and deduplicated. *De novo* assembly into contigs was done using the SPAdes software (Prjibelski et al. 2020) embedded in Geneious prime with default parameters and a k-mer of 51. The generated contigs were further annotated using TBLASTX on the refseq database of viral nucleotides sequences downloaded from NCBI (ftp://ftp.ncbi.nlm.nih.gov/refseq/release/viral). A single hit was retrieved for each contig. Contigs having homologies with plant viruses were further analysed by BLASTX on the non-redundant (nr) protein database from NCBI. In addition, VirHunter (as trained during the original publication Sukhorukov et al. 2022) was used with default parameters on the contigs with a size higher than 800 nts to confirm the identification and explore the presence of distant viruses.

Genome characterization of detected viruses

In order to extract the consensus genome of the detected viruses, read mappings (Minimap2, v.2.17) were performed by mapping the raw reads on the generated contigs (with iterations to extend the contigs). Mapping results were visually inspected for possible differences between the *de novo* contigs and consensus genome sequence (Li 2011) and to remove possible sequencing artefacts (IGV, 2.13.2).

For the newly discovered virus, the predicted ORFs were identified using Geneious Prime. The coat protein ORF was determined after alignment with 32 protein sequences from virus species belonging to *Capillovirus* genus and the closest *Trichovirus* genus in the subfamily *Trivirinae* of *Betaflexiviridae* family. The alignment was carried out using Muscle

3.8.425 (Edgar 2004) embedded in Geneious Prime with default parameters. Based on the alignment, the start codon was identified. A HMMER scan (Potter et al. 2018) was performed on the nearly complete genome recovered using HmmerWeb pfam 35.0 release hosted at EMBL's European Bioinformatics Institute (https://www.ebi.ac.uk/Tools/hmmer/).

The complete or nearly complete genomes of known virus species belonging to *Capillovirus* genus were downloaded from NCBI, selecting maximum 5 genome sequences per species (Table 2). The genomes were further aligned using Clustal Omega 1.2.2 (Sievers and Higgins 2018) embedded in Geneious Prime with default parameters. A phylogenetic tree was built using UPGMA method (Michener and Sokal 1957) embedded in Geneious tree builder using default parameter and 100 repetitions.

Confirmation of HTS results by RT-PCR and Sanger sequencing

Total RNA extracts from one symptomatic and two asymptomatic *E. crista-galli* plants were used to amplify PDV RNA3 genome segment portion by one-step RT-PCR (Superscript III-Platinum Taq, ThermoFisher Scientific, Carlsbad, CA, USA) using primers PD3-331F and PD3-1086R designed by Predajňa et al. (2017). Reverse transcription was performed at 60 °C for 15 min and the enzyme denatured at 94 °C for 2 min. PCR-cycles profiles were as follows: 94 °C, 15 s; 53 °C, 30 s; 68 °C, 1 min; 35 times; final extension was at 68 °C for 15 min. The obtained amplicon of 756 bp covers almost entire MP gene of PDV.

The same *E. crista–galli* RNA extracts were used for the amplification of the capillovirus by primers 5894F (5'-TTGGCCAGATGTTGAGATCCA-3') and 6273R (5'-GGCCTTCAGTCCTGAACAAC-3') constructed in this study covering the conserved portion of the capillovirus CP gene (genome positions from 5,894 to 6,273 nts). The reverse transcription was performed at 59 °C for 20 min, the enzyme denatured (94 °C for 2 min) and PCR cycle profiles were as described above except for the annealing step performed at 59 °C for 40 s. The capillovirus amplicon size was 380 bp and these amplicons were directly sequenced in both directions (Macrogen Europe, Amsterdam, the Netherlands).

TA-cloning was performed wherein amplicons from *E. crista-galli* were inserted into plasmid vector pTZ57R/T (Thermo Scientific™ InsTAclone™ PCR Cloning Kit, ThermoFisher Scientific, Carlsbad, CA, USA) and *E. coli* XL1-Blue cells (Invitrogen, Carlsbad, CA, USA) were transformed according to the manufacturers' instructions. Two clones from blue colonies were used for PCR verification of the inserts with the same set of primers as above and the same PCR-cycling conditions. Chimeric plasmids from two out of six clones were purified (PureLink Quick Plasmid Miniprep Kit, Invitrogen, Carlsbad, CA, USA) and sent to bidirectional Sanger sequencing (Macrogen Europe, Amsterdam, the Netherlands) with plasmid M13/pUC universal primers.

Phylogenetic analyses for PDV were done with MEGA11 (Tamura et al. 2021). The alignments between PDV genome contigs of RNA3 from HTS and MP sequences obtained from Sanger sequencing and the refseq PDV sequences (Table 1) were done using multiple sequence comparison by log-expectation (MUSCLE) alignment (Edgar 2004) embedded in Geneious with a maximum number of 8 iterations. Phylogenetic trees were inferred using the Neighbor Joining algorithm based on Tamura-Nei model with 500 bootstrap replicates.

Results

Disease symptoms in *E. crista-galli*

The symptoms in one out of three *E. crista-galli* plants in the Botanical Garden had been observed in 2010. The leaves showed speckling, bleaching and striations whilst the crimson flowers had prominent white striations in one out of two branches of the symptomatic plant (Fig. 1b). After cutting back the symptomatic branch, symptoms appeared in in the new flush and the remaining previously asymptomatic branch suggesting the presence of an infectious agent in this small tree. Symptoms persist and the symptomatic plant appears dwarfed in comparison to asymptomatic ones from the collection.

Molecular characteristics of PDV RNA3

In total, 4,132,702 reads were obtained by Illumina sequencing procedure, with the median read length of 151. Raw reads are publicly available in the NCBI Sequence Read Archive (SRA): BioSample Accession SAMN34578022 (https://www.ncbi.nlm.nih.gov/biosample/?term=SAMN34578022). The number of reads mapped to PDV genome were 11 for the PDV genomic segment RNA1 (32% of the genome was covered by sequencing reads), six for the segment RNA2 (covering 34% of the genome) and 57 for the segment RNA3 (covering 86% of the genome). By iterative mapping on the *de novo* contigs generated from these reads, a complete RNA3 segment (OP503942, 2,129 bp in length) of PDV genome encompassing its MP and CP genes was obtained (Fig. 2). The length of the intergenic MP-CP region was 72 nts as for many known isolates (Kozieł et al. 2017; Vašková et al. 2000) with four single nucleotide polymorphisms (Fig. 2) as compared to one of the Czech sweet cherry isolates (AF208737, Vašková et al. 2000).

The finding of PDV in symptomatic *E. crista-galli* was confirmed by RT-PCR with specific primers (Predajňa et al. 2017) covering almost entire PDV-MP gene. No amplicons were obtained from two asymptomatic *E. crista-galli* plants in the collection (results not shown). After cloning, two representative amplicons were Sanger sequenced. The sequences were 100% identical for two clones (OP503943) from the RT-PCR confirmation experiments whilst they had 99% nt and 100% amino acid (aa) identity with the same portion of sequence from a longer fragment obtained by HTS and resulted in partial coding sequences (CDS) covering 756 out of 882 bases of the MP (Fig. 2).

The only difference in MP with the phylogenetically closest Slovakian isolate (GenBank accession No. KU949336) from sweet cherry (Predajňa et al. 2017) is in the second aa wherein Ile is substituted by Val in our isolate (Fig. 2). PDV-CP gene had 96.3% nt identity and 97.3% amino acid (aa) identity with the same portion of sequence from Canada

(MZ221012) (Fig. 3b). The PDV-CP motif KPTTRSQSFA (aa 10-19) essential for the RNA binding (Vašková et al. 2000; Pallas et al. 2012) is also present and completely conserved (Fig. 2).

Phylogenetic grouping of the *E. crista-galli* PDV isolate

The Croatian PDV-MP sequences from *E. crista-galli* PDV from HTS and Sanger sequencing were compared to a range of PDV isolates originated from different parts of the world and types of hosts (Supplementary Table 1). Expectedly, the two Croatian sequences grouped together (Fig. 3a) and in the same branch with the Slovakian isolate from cherry (Predajňa et al. 2017). As PDV-CP sequence was obtained in this research only from HTS data, only that sequence was phylogenetically analyzed (Fig. 3b). It was the closest to a Chinese sequence from sweet cherry; however, it was in the same cluster with most of the isolates chosen for the
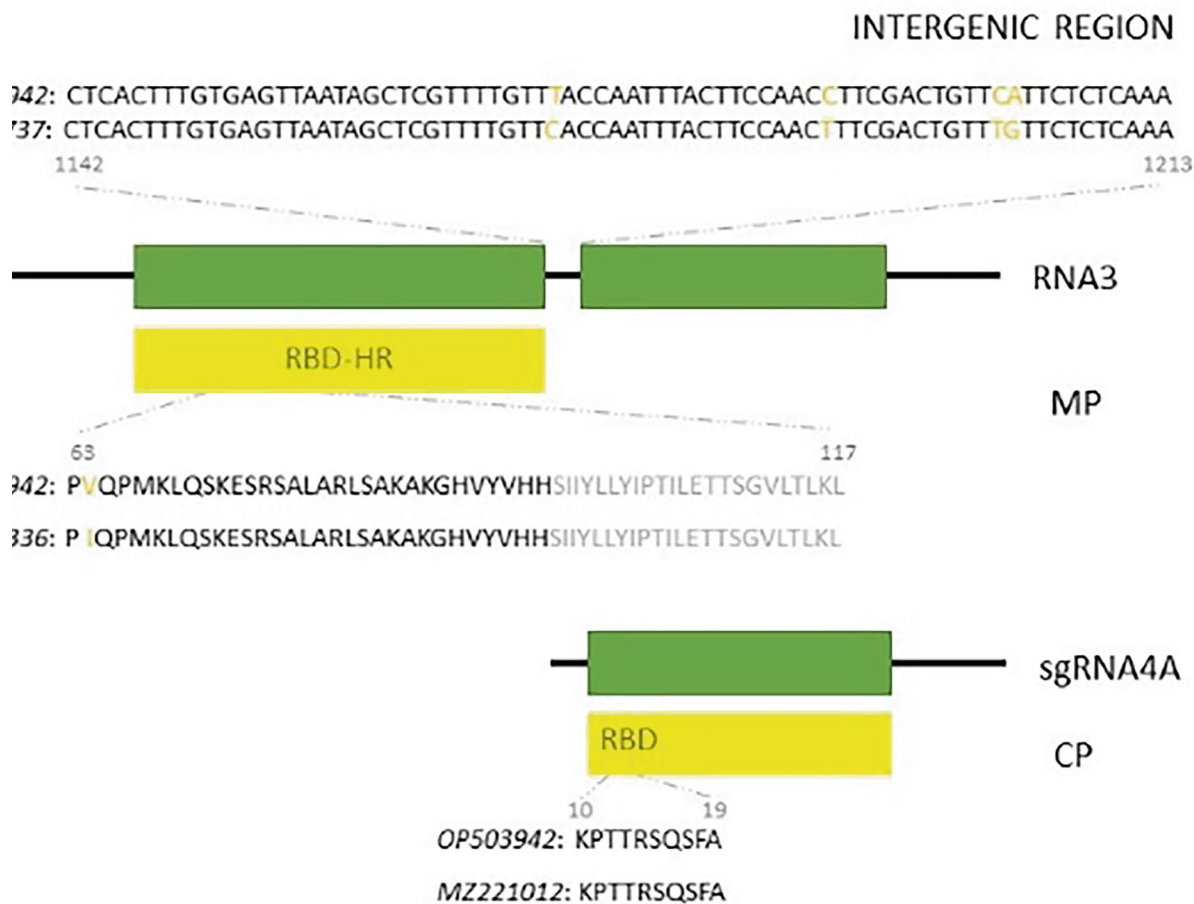
Fig. 2 Diagrams showing the RNA3 segment organization of prune dwarf virus (PDV) with coding sequences (CDS) for the movement (MP) and coat proteins (CP), amino acid contents of functionally important RNA binding domain (RBD) and hydrophobic region (HR), as well as a nucleotides in the intergenic region between MP and CP. GenBank accession number OP503942 denotes complete RNA3 sequence from *E. crista-galli*, AF208737 a cherry sequence from the Czech Republic (Vašková et al. 2000) and other two sequences are listed in the Supplementary Table 1
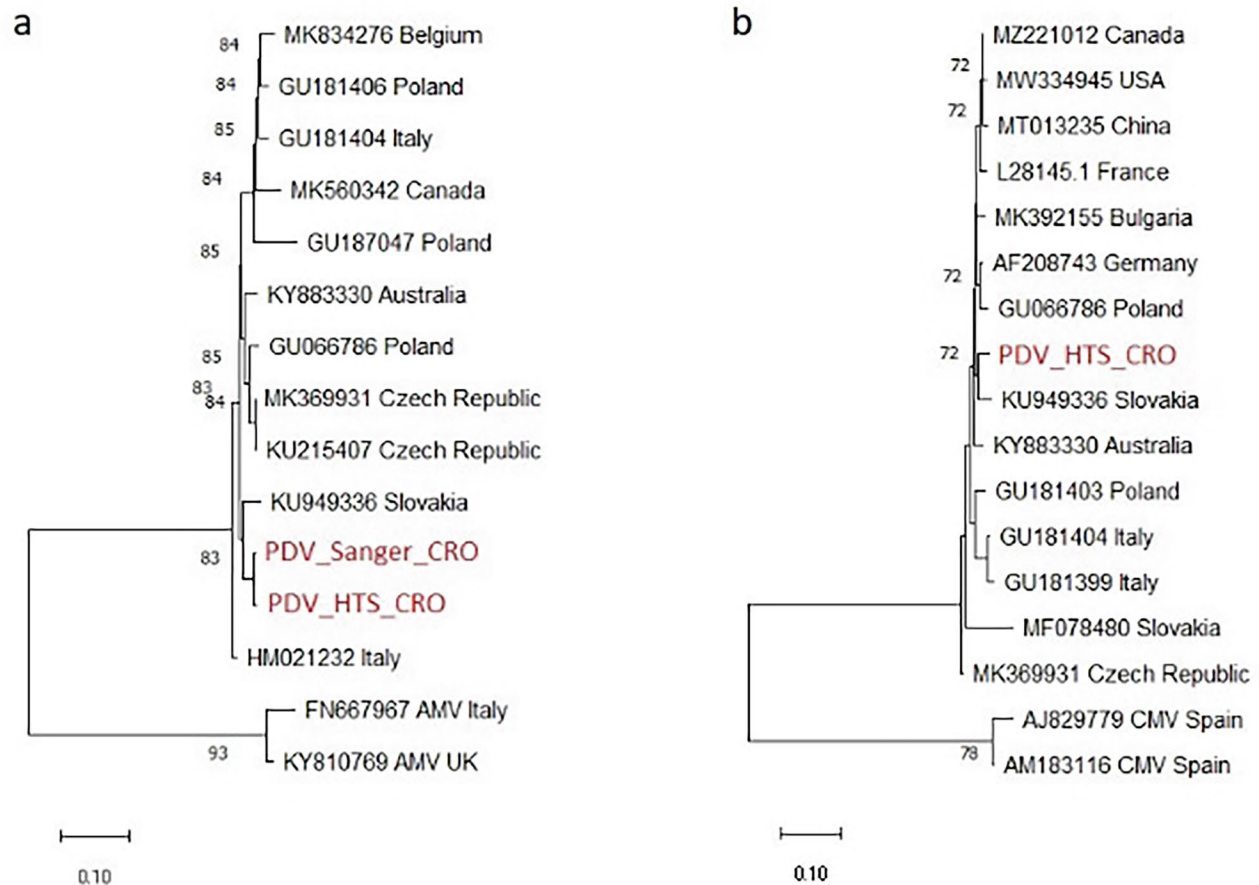
Fig. 3 Phylogeny reconstruction (Neigbour-Joining, Tamura-Nei substitution model) of the **a**)-PDV-MP gene from *E. crista-galli* (PDV_Sanger_CRO acc. No. OP503943 and PDV_HTS_CRO acc. No. OP503942) and **b**) CP gene (PDV_HTS_CRO). Bootstrap values >70 % (500 bootstrap resampling) are indicated on the branches and data on the PDV isolates and sequences used to root the trees are in the Table 1. The scale bar indicates a genetic distance

comparison (Supplementary Table 1) regardless of their geographical origin or host.

Molecular characterization and classification of a novel capillovirus from *E. crista-galli*

The number of reads assembled to the contig belonging to a putative new capillovirus species was 23,259 forming a contig of 6,483 nts with a mean coverage of 583.9. Almost a complete genome (GenBank acc. No. OQ067396) was obtained for this previously unknown virus of the family *Betaflexiviridae*. It spans 6,483 nts encompassing a polyprotein coding sequence of 6,326 nts (nucleotide positions 33 - 6,326) including the fused ORF for a CP of 654 nts (positions 5,673 – 6,326), 1071-nt-long MP (positions 4,754 – 5,824) in another frame and additional 32 and 155 nts of the respective 5'- and 3'-UTRs (Fig. 4). No other virus was detected and this result was confirmed by an independent analysis (unpresented results) using Virhunter, a recently developed algorithm to detect plant viruses in HTS datasets (Sukhorukov et al. 2022).

Phylogenetic analyses revealed a maximum nucleotide identity of 42% (at genome level) to the apple stem grooving virus (ASGV) (GenBank acc. No. LC143387) as its closest capillovirus. Higher degree of genetic relatedness of the new capillovirus with ASGV is also discernible from the dendrogram (Fig. 5) where the new virus forms a 100% statistically supported branch of its own in the same cluster with ASGV isolates from apples and yuzu. Similarly, other newly discovered viruses in this genus like yacon virus A (YVA) or breadfruit capillovirus 1 (BCV-1) also form their own branches. Interestingly, potato virus T (PVT) as a member of *Tepovirus* genus and apple chlorotic leafspot virus (ACLSV, genus *Trichovirus*) within the same *Trivirinae* subfamily are also in this group. Viruses found in the other cluster (Fig. 5.) are capilloviruses mostly found in the woody plants, including better known cherry virus A (CVA) and MuVA found in stone fruit trees.

A 309 bp-long sequence was obtained from the Sanger sequencing following the RT-PCR confirmatory test with
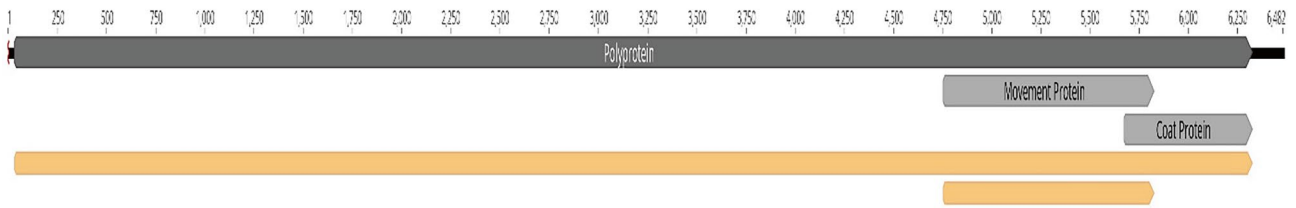
Fig. 4 Almost complete annotated genome (6,483 nucleotide-long) of a new capillovirus from *Erythrina crista-galli* as generated from the Geneoius Prime software

newly designed primers (GenBank acc. No. OQ428619). The amplicon sequence is 99.7% identical (308 out of 309 nts) to the corresponding CP portion of the new Erythrina capillovirus full genome sequence. The viral amplicon was confirmed only in one symptomatic *E. crista-galli* tree out of three in the collection, as for the PDV (results not shown).

The protein alignment analyses using 32 sequences of capilloviruses and trichoviruses (Supplementary Table 3) showed only a maximum of 39% aa identity between Erythrina capillovirus and the CP of the phylogenetically closest ASGV (aa sequence QVW10187.1). The majority of other CP sequences from the analysis were far less below this value with the lowest CP aa identity value of 17 % for CVA.

Fig. 5 Unrooted tree (UPGMA, 100 repetitions) of the new Erythrina capillovirus genome from Croatia (red star, GenBank acc. No. OQ067396) generated by Geneoius Prime software. The bar represents genetic distance. The sequences used for comparison are listed in the Supplementary Table 2. Virus acronyms or names are on the corresponding branches with the accession numbers
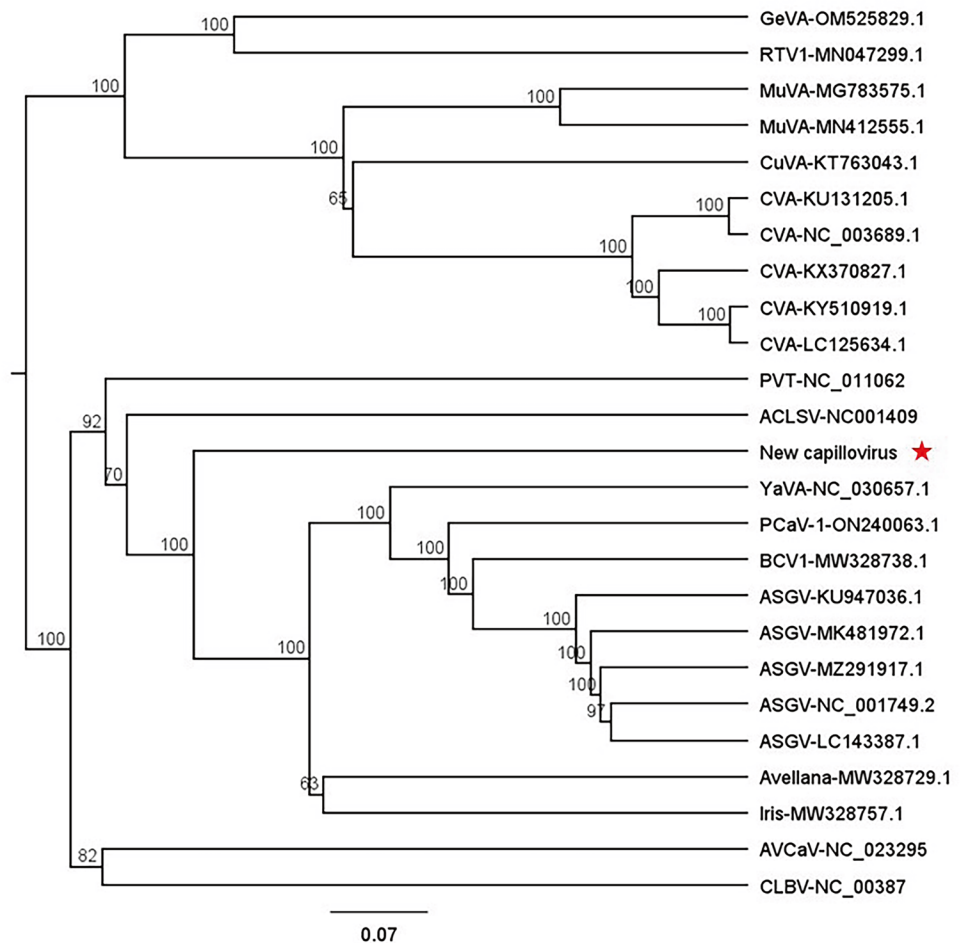
Discussion

PDV is an important virus adversely affecting the stone fruit trees (Çağlayan et al. 2011; Pallas et al. 2012), but its biology is not as well explored as for the other viruses in the family *Bromoviridae*, or genus *Ilarvirus* (Kozieł et al. 2017). A step forward has been done recently with the construction of a full-length cDNA infectious clone of a Canadian PDV sweet cherry isolate (Simkovich et al. 2021). Whilst this will provide means to push forward molecular and biological studies, there is probably still a lot to learn about the diversity and ecology of this virus. The virus has been recorded in Croatia and the countries of ex-Yugoslavia in stone fruits (Dulić-Marković and Ranković 1992a, b; Matić et al. 2008; Šarić and Velagić 1980; Zindović et al. 2013). However, in this study, we have identified a new isolate of PDV from a host outside the genus *Prunus*, where it has been most frequently reported worldwide. *E. crista-galli* (Fig. 1) belongs to the family Fabaceae known to encompass PDV hosts like *Sesbania exaltata* and *Crotalaria spectabilis* which were listed as assay species due to the local lesion appearing after inoculation (Fulton 1970). Nevertheless, no woody plant from this family has been recorded as a PDV host before this study.

The 2,129-nt-long RNA3 of our PDV isolate from *E. crista-galli* (Fig. 2) displayed typical molecular characteristics. For example, the more conserved MP gene of this bicistronic genome segment displayed RBD and HR domains (Kozieł et al. 2017; Vašková et al. 2000; Pallas et al. 2013; Predajňa et al. 2017). The substitution of isoleucine by valine in the second position of RBD (Fig. 2) is not likely to disturb its essential functions in viral cell-to-cell movement as both are hydrophobic side-chain amino acids. Phylogenetically, our PDV isolate groups with mostly European stone fruit isolates with the closest isolate corresponding to a Slovak sweet cherry isolate (Fig. 3). Also, the detectability of our isolate directly from the original host by PDV primers developed previously (Predajňa et al. 2017) corroborates these findings.

The data obtained here for the CP of *E. crista-galli* PDV isolate was compared with existing genomic information in the databases. The results suggest that its RNA binding activity through the KPTTRSQSFA motif is retained (Vašková et al. 2000; Pallas et al. 2012) as it is present and perfectly conserved (Fig. 2). The conservation of nucleotide and amino acid sequences, along with the persistence of this virus in the host and its detectability over the years suggest it is a functional CP forming virions. Phylogeny reconstruction based on the CP gene (Fig. 3b) was expected to be less informative, as this gene is more variable, and no correlation may be expected amongst isolates of the same geographical or host origin (Kozieł et al. 2017; Vašková et al. 2000;

Predajňa et al. 2017). However, in our analyses, it was similar to the more conserved MP gene (Fig. 3a). Even though the function of the intergenic MP-CP region of PDV RNA3 is unknown, it is worth noting that, unlike the diversity of found in the Slovakian group of PDV isolates (Predajňa et al. 2017), it is 72 nt-long in our isolate (Fig. 2) and highly similar to the described ones (Vašková et al. 2000).

The HTS data also enabled us the reconstruction of the complete coding sequence for a new virus (GenBank Acc. No. OQ067396) of 6,483 nts in length showing a typical organization for species of the genus *Capillovirus*, subfamily *Trivirinae* (Adams et al. 2012) within the family *Betaflexiviridae* (Fig. 4). Phylogenetic analysis of this nearly complete genome sequence confirmed that this virus has genetic relatedness with the other capilloviruses but forming its own lineage with 100% statistical support (Fig. 5) similarly to several other newly discovered capilloviruses. Its genome structure is typical of the capilloviruses (Adams et al. 2012) with annotated polyprotein, fused CP and MP in the second reading frame (Fig. 4). It has rather low nucleotide identity level (only 42%, Supplementary Table 3) with ASGV as the closest capillovirus.

The aa identity level of its CP with ASGV is also very low, only 39% (Supplementary Table 4). According to the still valid ICTV criteria for species delimitation within the *Betaflexiviridae* family (Adams et al. 2012), viruses with less than 72% nt or 80% aa identity in the CP or Rep genes should be considered different species. The new virus from *Erythrina* whose basic molecular characterization is performed here shows identity values far below these thresholds suggesting it should be considered a novel species. This would even be the case if the new set of criteria proposed recently (Silva et al. 2022), taking into consideration different evolutionary histories of Rep and CP, should be accepted. Therefore, we will propose to ICTV that this putatively new capillovirus should be named Erythrina capillovirus (ErCV) and considered as a new virus species with the name *Capillovirus ErCV*.

At this point, we do not have much biological data on the ErCV because it must be separated from PDV as the other constituent of the *E. crista-galli* virome. Also, the origin of PDV isolate and the ErCV in our collection is still enigmatic. They could have been even introduced by seeds because it is known that both viruses may be seed-borne (Fulton 1970; Mink and Aichele 1984; Mink 1993; Adams et al. 2012). A symptomatic seedling had been once obtained (Vanja Stamenković, personal communication) from the symptomatic plant described here (Fig. 1b) but was no more available when the potential viral etiology of the disease was explored through this study. Nevertheless, this hypothesis can be tested in future, as well as biological features of these two viruses.

This is the first analysis of *E. crista-galli* virome, an interesting fabaceous tree host exotic to the western world. The striking symptoms in this ornamental plant may even be considered aesthetically pleasing but they stunt the growth, deform the leaves and we have no information if the plant metabolism, including the level of important bioactive compounds (Miño et al. 2002; Mitscher et al. 1988; Weber et al. 2005) in the host, is adversely affected. Even though we have not determined the direct link between PDV, Erythrina capillovirus and the disease in *E. crista-galli*, the data gathered here certainly raise suspicion about the viral etiology of the disease. In addition, they reveal yet another new species of the growing *Capillovirus* genus and enrich knowledge on the diversity of PDV isolates worldwide putting them into new ecological niche occupied by fabaceous woody ornamentals as their previously unknown hosts.

Data availability Data is available freely as NCBI accession numbers are provided.

References

Adams MJ, Candresse T, Hammond J, Kreuze JF, Martelli GP (2012) Family *Betafexiviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus Taxonomy: classification and nomenclature of Viruses - Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, pp 920–941

Bachman EJ, Scott SW, Xin G, Vance VB (1994) The complete nucleotide sequence of prune dwarf ilarvirus RNA 3: implications for coat protein activation of genome replication in ilarviruses. Virology 201:127–131

Bujarski JJ, Figlerowicz M, Galitelli D, Roosinck MJ, Scott SW (2012) Family *Bromoviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus Taxonomy: classification and nomenclature of viruses – Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, pp 972–976

Çağlayan K, Ulubas-Serce C, Gazel M, Varveri C (2011) Prune dwarf virus. In: Hadidi A, Barba M, Candresse T, Jelkman W (eds) Virus and virus-like diseases of pome and stone fruits. American Phytopathological Society, St. Paul, pp 199–205. https://doi.org/10.1094/9780890545010.037

Dulić-Marković I, Ranković M (1992) Study on the prune dwarf virus-peach isolate. Zaštita Bilja 43:203–210

Dulić-Marković I, Ranković M (1992) Application of indirect ELISA technique for detecting prune dwarf virus and tobacco streak virus. Zaštita Bilja 43:219–225

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Fonseca F, Neto JD, Martins V, Nolasco G (2005) Genomic variability of prune dwarf virus as affected by agricultural practice. Arch Virol 150:1607–1619

Fulton RW (1970) Prune dwarf virus. C.M.I/A.A.B. Descriptions of Plant Viruses No. 19. https://www.dpvweb.net/dpv/showdpv/?dpvno=19 . Accessed on 25 Sept 2022

Hanafi M, Tahzima R, Ben Kaab S, Tamisier L, Roux N, Massart S (2020) Identification of divergent isolates of banana mild mosaic virus and development of a new diagnostic primer to improve detection. Pathogens 9:1045. https://doi.org/10.3390/pathogens9121045

ICTVdB Management (2006) 00.010.0.02.014. Prune dwarf virus. In: ICTVdB - The Universal Virus Database, version 4. Büchen-Osmond, C. (Ed), Columbia University, New York, USA. https://archive.ph/20060909153905/http://phene.cpmc.columbia.edu/ICTVdB/00.010.0.02.014htm. Accessed 30 Aug 2022

Kozieł E, Bujarski JJ, Otulak K (2017) Molecular biology of *prune dwarf virus*—a lesser known member of the *Bromoviridae* but a vital component in the dynamic virus–host cell interaction network. Int J Mol Sci 18:2733. https://doi.org/10.3390/ijms18122733

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993

Marais A, Faure C, Theil S, Candresse T (2018) Molecular characterization of a novel species of *Capillovirus* from Japanese apricot (*Prunus mume*). Viruses 10(4):144. https://doi.org/10.3390/v10040144

Massart S, Brostaux Y, Barbarossa L, César V, Cieslinska M, Dutrecq O, Fonseca F, Guillem R, Lavina A, Olmos A, Steyer S, Wetzel T, Kummert J, Jijakli MH (2008) Inter-laboratory evaluation of a duplex RT-PCR method using crude extracts for the simultaneous detection of prune dwarf virus and Prunus necrotic ringspot virus. Eur J Plant Pathol 122:539–547. https://doi.org/10.1007/s10658-008-9322-1

Matić S, Al-Rwahnih M, Myrta A, Đurić G (2008) Viruses of stone fruits in Bosnia and Herzegovina. Acta Hortic 781:71–74

Michener CD, Sokal RR (1957) A quantitative approach to a problem of classification. Evolution 11:490–499

Mink GI (1993) Pollen- and seed-transmitted viruses and viroids. Ann Rev Phytopathol 31:375–402

Mink GI, Aichele MD (1984) Detection of Prunus necrotic ringspot and prune dwarf viruses in *Prunus* seed and seedlings by enzyme-linked immunosorbent assay. Plant Dis 68:378–381

Miño J, Gorzalczany S, Moscatelli V, Ferrar G, Acevedo C, Hnatyszyn O (2002) Actividad antinociceptiva y antiinflamatoria de *Erythrina crista-galli* L. (Ceibo). Acta Farm Bonaerense 21:93–98

Mitscher LA, Gollapudi SR, Gerlach DC, Drake SD, Eduardo AV, Ward JA (1988) Erycristin, a new antimicrobial petrocarpan from *Erythrina crista-galli*. Phytochemistry 27:381–385. https://doi.org/10.1016/0031-9422(88)83103-0

Pallas V, Aparicio F, Herranz MC, Amari K, Sanchez-Pina MA, Myrta A, Sanchez-Navarro JA (2012) Ilarviruses of *Prunus* spp.: a continued concern for fruit trees. Phytopathology 102:1108–1120. https://doi.org/10.1094/PHYTO-02-12-0023-RVW

Pallas V, Aparicio F, Herranz MC, Sanchez-Navarro JA, Scott SW (2013) The molecular biology of ilarviruses. Adv Virus Res 87:139–181. https://doi.org/10.1016/B978-0-12-407698-3.00005-3

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD (2018) HMMER web server: 2018 update. Nucl Acids Res 46:W200–W204. https://doi.org/10.1093/nar/gky448

Predajňa L, Sihelská N, Benediková D, Šoltys K, Candresse T, Glasa M (2017) Molecular characterization of *prune dwarf virus* cherry

isolates from Slovakia shows their substantial variability and reveals recombination events in PDV RNA3. Eur J Plant Pathol 147:877–885. https://doi.org/10.1007/s10658-016-1055-y

Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020) Using SPAdes de novo assembler. Curr Protoc Bioinform 70:e102. https://doi.org/10.1002/cpbi.102

Rampitsch C, Eastwell KC (1997) The complete nucleotide sequenceof prune dwarf ilarvirus RNA-1. Arch Virol 142:1911–1918. https://doi.org/10.1007/s007050050210

Šarić A, Velagić Z (1980) Viruses and virus-like diseases of maraska sour cherry (*Cerasus acida* ssp. Maraska). Acta Phytopathol Acad Scient Hungar 15:367–369

Scott SW, Zimmerman MT, Ge X, MacKenzie DJ (1998) The coat proteins and putative movement proteins of isolates of Prunus necrotic ringspot virus from different host species and geographic origins are extensively conserved. Eur J Plant Pathol 104:155–161

Šeruga M, Škorić D, Botti S, Paltrinieri S, Juretić N, Bertaccini AF (2003) Molecular characterization of a phytoplasma from the aster yellows (16SrI) group naturally infecting *Populus nigra* L. 'Italica' trees in Croatia. For Pathol 33:113–125. https://doi.org/10.1046/j.1439-0329.2003.00312.x

Sievers F, Higgins DG (2018) Clustal Omega for making accurate alignments of many protein sequences. Protein Sci 27:135–145. https://doi.org/10.1002/pro.3290

Silva JMF, Melo FL, Elena SF, Candresse T, Sabanadzovic S, Tzanezakis IE, Blouin AG, Villamor DEV, Mollov D, Constable F, Cao M, Saldarelli P, Cho WK, Nagata T (2022) Virus classification based on in-depth sequence analyses and development of demarcation criteria using the *Betaflexiviridae* as a case study. J Gen Virol 103:001806. https://doi.org/10.1099/jgv.0.001806

Simkovich AJ, Li Y, Kohalmi SE, Griffiths JS, Wang A (2021) Molecular identification of prune dwarf virus (PDV) infecting sweet cherry in Canada and development of a PDV full-length infectious cDNA clone. Viruses. https://doi.org/10.3390/v13102025. (Viruses 13: 2025)

Sukhorukov G, Khalili M, Gascuel O, Candresse T, Marais-Colombel A, Nikolski M (2022) VirHunter: a deep learning-based method for detection of novel RNA viruses in plant sequencing data. Front Bioinform 2:867111. https://doi.org/10.3389/fbinf.2022.867111

Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mol Biol Evo 38:3022–3027

Ulubaş-Serçe Ç, Ertunç F, Öztürk A (2009) Identification and genomic variability of *prune dwarf virus* variants infecting stone fruit trees in Turkey. J Phytopathol 157:298–305. https://doi.org/10.1111/j.1439-0434.2008.01486.x

Vašková D, Petrzik K, Špak J (2000) Molecular variability of the capsid protein in the prune dwarf virus. Eur J Plant Pathol 106:573–580

Weber D, Gorzalczany S, Martino V, Acevedo C, Sterner O, Anke T (2005) Metabolites from endophytes of the medicinal plant *Erythrina crista-galli*. Z Naturforsch 60c:467–477. https://doi.org/10.1515/znc-2005-5-616

Wistler WA, Elevitch CR (2006) *Erythrina variegata* (coral tree). Species profiles for Pacific Island Agroforestry ver 3.1. https://agroforestry.org/free-publications/traditional-tree-profiles . Accessed 1 Sept 2022

Zindović J, Lanzoni C, Autonell CR, Ratti C (2013) First report of prune dwarf virus and prunus necrotic ringspot virus on peach in Montenegro. Plant Dis 97:1259. https://doi.org/10.1094/PDIS-12-12-1147-PDN