Nuria Fontdevila Pareta
Phytobiomes Journal (Phytovirome Focus Issue)

Virome scanning of pear germplasm collections identifies a new Velarivirus and extends the geographical spread of three other pear viruses

Nuria Fontdevila Pareta[1], Carole Gailly[1], Arnaud G. Blouin[2], Beatrix Buchmann[3], Markus Buenter[3], Thierry Candresse[4], Nathalie Dubuis[2], Denis Kutnjak[5], Marc Lateur[6], Anja Pecman[5], Stéphan Steyer[7], Sébastien Massart[1*]

[1] University of Liège, Gembloux Agro-Bio Tech, Plant Pathology Laboratory, 5030, Gembloux, Belgium

[2] Agroscope, Plant Protection Department, 1260, Nyon, Switzerland

[3] Agroscope, Plant Health Service, 8820, Wädenswil, Switzerland

[4] Univ. Bordeaux, INRAE, UMR 1332 BFP, 33882, Villenave d'Ornon cedex, France

[5] National Institute of Biology, Department of Biotechnology and Systems Biology, 1000, Ljubljana, Slovenia

[6] Centre Wallon de Recherches Agronomiques (CRA-W), Plant & Forest Biodiversity & Breeding, 5030, Gembloux, Belgium

[7] Centre Wallon de Recherches Agronomiques (CRA-W), Crops & Forest Health, 5030, Gembloux, Belgium

[*]Correspoding author: Sébastien Massart; E-mail: Sebastien.Massart@uliege.be

Funding: European Union as part of the Marie Sklodowska-Curie Actions Innovative Training Network (H2020 MSCA-60 ITN) project INEXTVIR Grant Number: 813542; Slovenian Research and Innovation Agency core financing P4-0407 Project: J4-4553.

Abstract

In this study, an extensive virome investigation was performed on a germplasm collection of pear trees (*Pyrus communis* L.) from the Walloon Agricultural Research Centre (Gembloux, Belgium). In total,128 pear trees samples were analyzed as pools using high-throughput sequencing (HTS) techniques, and/or tested individually for targeted viruses by RT-PCR. During the virome survey, a novel velarivirus was identified in several asymptomatic trees while four known viruses were detected. Bioinformatics tools were used to assemble the genome of the new virus. The pear germplasm collection from Kozjanski Park (Slovenia) and a viral collection from Agroscope (Nyon, Switzerland) were also surveyed for the new pear virus and for three known viruses (CiVA, ARWV-1, and ARWV-2) to study their prevalence and geographic distribution. In Belgium, the new velarivirus was detected by RT-PCR in six of the 99 sampled trees (6%) and citrus virus A (CiVA) in 49 (49%) of them; in Slovenia four of the six trees sampled (67%) were positive for CiVA; and in Switzerland four of the nine trees sampled (44%) were positive for CiVA and 1 (11%) for apple rubbery wood virus 1 and 2 (ARWV-1 and -2). This study, combined pooled HTS analyses to maximize the number of germplasm tested and targeted RT-PCR tests on individual samples for accurate detection. It reports and describes a new velarivirus discovered in pear trees and first detections of CiVA in Belgium, Switzerland and Slovenia, and ARWV-1 and -2 in Switzerland.

Keywords: virome, plant virus, pear, velarivirus, high-throughput sequencing, citrus virus A, apple rubbery wood virus 1, apple rubbery wood virus 2, pyrus virus A

1 Introduction

Pear cultivation is a vital sector of the global horticultural industry, contributing to both domestic and international markets. Belgium is one of the top 10 global pear-producing countries, ranking just behind Turkey as the top producer in Europe in 2021. Worldwide, the area under pear cultivation decreased in 2015, but it has been increasing since then (FAO, 2024). In comparison, in the European Union the pear cultivation area has diminished, except in the Netherlands and Belgium. In these two countries, the most produced pear variety is 'Conference', representing 53% of the total pear production. The high production of only a few specific cultivars results in a narrow genetic diversity of cultivated crops, which can threaten the resilience of future production in the face of changes in biotic and abiotic stresses (Shahzad et al., 2021). Planting and using certified plant material in commercial orchards decrease the risk of having severe yield losses due to diseases of viral origin. Still, it should not be neglected that with changes in the environment, emerging viruses and diseases might become more common (Trebicki, 2020).

To tackle this issue, it is essential to better characterize the viruses that can infect pear trees and understand if environmental changes might affect their impact on pear cultivars. Beyond commercial cultivars, genetic diversity is preserved in germplasm collections. Germplasm collections are vital for conserving plant genetic diversity and as a source of genetic material and new traits for breeding, to ensure and sustain the future adaptability of food production. This is why it is necessary to characterize these resources, including their viral status. The Walloon Agricultural Research Centre (CRA-W) owns and develops an extensive collection of plant genetic resources, specializing in old cultivars, which are the basis of research to study their nutritional and biological properties and to promote their use in breeding programs. For pear, the CRA-W fruit tree collection includes more than 500 accessions conserved in ex-situ, and in-situ orchards.

Viral infections in fruit trees can be difficult to observe as symptoms vary greatly depending on the plant variety and the viral isolate (Katsiani et al., 2018; Maliogka et al., 2018). In addition, viruses in fruit trees are transmitted via vegetative propagation such as grafting, ensuring their transmission to the next generation, and raising the risk of accumulating multiple infections over time in a single plant. Hence, detecting plant viruses is essential, though challenging, for the safe propagation and cultivation of fruit trees. Molecular, serological, and biological assays are used to detect plant viruses, and each has its set of advantages and disadvantages (Boonham et al., 2014). High-throughput sequencing (HTS) technologies present great opportunities for virus discovery, detection, identification, and characterization in fruit trees, as they can potentially pinpoint every putative viral agent present in a sample without any prior knowledge of plant origin or symptomatology (Massart et al., 2014).

Pear vein yellows is the most common viral disease of pear. It is caused by apple stem pitting virus (ASPV) infection, without significant effects on growth and yield when ASPV is present alone (Jelkmann and Paunovic, 2011). However, when ASPV is present in a co-infection with some other viruses or phytoplasmas, significant growth reduction may occur (Yanase et al., 1989). Other viruses commonly infecting pear trees are apple stem grooving virus (ASGV) and apple chlorotic leaf spot virus (ACLSV) (Massart et al., 2011; Yaegashi et al., 2011).

In pear trees, a recent review showed that from 2011 to 2020 only two novel viruses have been discovered using HTS technologies (Hou et al., 2020). In addition, a new robigovirus was identified by HTS in pear trees, and tentatively named pomes virus Greece (PVGR) (Costa et al., 2022). This number is very low compared to apple trees, for which 15 new viruses have been discovered by HTS during the same period (Hou et al., 2020).

The family *Closteroviridae* includes viruses with long filamentous virions of 650 to 2,200 nanometers (nm), and large positive-sense single-stranded RNA genome (up to 19.3 kb). Their current taxonomy is based on the evolutionary histories of the three proteins used for their

classification: heat shock protein 70 homolog (HSP70h), RNA-dependent RNA polymerase (RdRp), and the coat protein (CP). Additionally, their genomes encode a duplicated, but divergent, copy of the capsid protein called the minor capsid protein (CPm) (Fuchs et al., 2020). Currently, species belonging to the family *Closteroviridae* (n=57) are classified in seven genera: *Ampelovirus*, *Bluvavirus*, *Closterovirus*, *Crinivirus*, *Menthavirus*, *Olivavirus*, and *Velarivirus*. Velariviruses infect primarily woody hosts and, in most cases, do not induce any apparent symptoms. Mechanical or seed transmission has not been reported for any virus from this genus. On the other hand, a vector has been identified for a single species: areca palm velarivirus 1 (ArPV1) was recently shown to be associated with the yellow leaf disease of Betel palm (*Areca catechu*) and to be transmitted by two mealybugs of the *Pseudococcideae* family (*Ferrisia virgata* and *Pseudococcus cryptus*) (Zhang et al., 2022).

The present study aimed to use HTS technologies for a better characterization of the viruses present in the CRA-W pear germplasm collection. An analysis of the diversity of viral infections in a small selection from two additional collections, in Switzerland and Slovenia, was also performed. The characterization of the virome of the selected collections led to the identification of a novel virus and the report and molecular detection of recently described viruses in new geographical locations. The newly described virus identified during the study was characterized following the optimized scientific and regulatory framework for the characterization and risk analysis of newly discovered plant viruses and viroids (Fontdevila Pareta et al., 2023).

2 Materials and methods

2.1 Plant material origin and outline of the tests conducted

The sampling strategy was designed to take into account the possible heterogenous distribution of plant viruses within a pear tree, following a previously published strategy (Kummert et al., 2004). For each tree, one leaf at each cardinal point at two different heights of the canopy was

collected. Therefore, a sample was comprised of eight leaves from one tree. Before extraction using the double-stranded RNA (dsRNA) protocol (see section 2.2.1), generally, four samples (0.75 g each) were mixed into a pool. Then, before sequencing, pools were tagged and mixed into libraries (see section 2.2.1). The collections screened are in the open air without protective nets. The collection at the Centre Wallon de Recherches Agronomiques (CRA-W) in Gembloux (Belgium) is distributed across various orchards, although for this study all samples were collected from the same orchard. In total, 128 pear trees from the CRA-W collection were sampled and analyzed, including 65 pear trees sampled and analyzed by HTS, in a total of 17 pools gathered in seven libraries, and, for the field survey, 99 pear trees were sampled and analyzed by reverse transcription PCR (RT-PCR). Notably, the trees showed no symptoms of viral infection at the time of sampling. Sampling in the collection of the CRA-W was done in June 2021 for samples that were analyzed with HTS and in May 2023 for samples analyzed by RT-PCR. To validate key detections by HTS, total RNAs (see section 2.2.2) from samples collected in June 2021 were also re-extracted and tested by RT-PCR two years later. Additionally, during spring 2023, the distribution of the newly discovered virus, tentatively named Pyrus virus A (PyVA) within a tree was studied by testing phloem, flowers, and leaves from the four cardinal points of a positive tree. From the collection at Kozjanski Park in Slovenia, 6 pear trees were sampled in May 2021 and tested by reverse transcription polymerase chain reaction (RT-PCR). From the collection at Agroscope in Switzerland, 9 pear trees were sampled in October 2022 and tested by RT-PCR.

2.2 Alien control strategy

An external alien control was used to monitor cross-contamination between samples and to aid in the differentiation between false and true positives, as proposed by Rong et al. (2023) for *Musa* spp. and in the guidelines for the use of HTS in the detection of plant pathogens and pests (Massart et al., 2022). As described by Massart et al. (2022), in plant virus diagnostics, an

external alien control corresponds to a plant sample containing one or several viruses (called alien viruses) that are not expected to be present in the tested samples. Thus, the detection of reads from an alien virus in a sample can be considered as a contamination from the alien control. In this study, leaves of a bean plant infected with three endornaviruses (Phaseolus vulgaris virus 1 (PvEV-1), 2 (PvEV-2), and 3 (PvEV-3)) were used as external alien control because their host range is restricted to *Phaseolus vulgaris*.

2.3 Extraction protocols and sequencing

2.3.1 Double-stranded RNA extraction and sequencing

Double-stranded RNA (dsRNA) was extracted from leaf tissue, reverse-transcribed, and amplified to be submitted to HTS, adapting the protocol described by Marais et al. (2018). The extraction protocol was adapted to upscale the extraction buffer and reagent volumes fourfold, so a total weight of 3 g of frozen sample was used for the extraction. In those cases where the number of samples per pool varied, the weight of each sample within the pool was adapted accordingly. Once the starting plant material was ground with liquid nitrogen, the powder was transferred to a 15 ml tube (Greiner bio-one International GmbH) containing the extraction buffer (4 ml of 2x STE, 280 µl of 20% SDS, 160 µl of sodium bentonite, and 5.7 ml of phenol-TE saturated). After adding the powdered sample to the extraction buffer, the tubes were agitated for 30 min on a horizontal shaker and centrifuged for 15 min at 3,000 g. Then, the aqueous phase was transferred to a 1.5 ml Eppendorf tube and centrifuged at 10,000 g for 20 min. After centrifugation, 1 ml of the aqueous phase was transferred to a new 1.5 ml tube, and the rest was kept at -80ºC. Then, dsRNAs were purified by two series of cellulose chromatography, between which a nuclease and proteinase K treatment was performed, as described in Marais et al. (2018).

Complementary DNA (cDNA) synthesis was performed for each sample by denaturing 5 µl of purified dsRNAs and 4.5 µl of diethyl pyrocarbonate (DEPC) treated water at 99ºC for 5 min

and then keeping the samples on ice for 1 min. Then, 2 µM of dodeca linkers (François et al., 2018) and DEPC-treated water were added for a total volume of 10.5 µl, and the samples were incubated at 95ºC for 5 min, followed by 1 min on ice. The next mix was composed of 1 mM of dNTPs, 1x reaction buffer, 10 U Superscript III Reverse Transcriptase (Invitrogen), 1 U/µl RNaseOUT Recombinant RNase Inhibitor (Invitrogen), and 10 mM DTT in a 20 µl volume reaction, which was added to each tube. After incubating at 25ºC for 10 min and at 42ºC for 60 min, the RT was inactivated at 70ºC for 5 min, followed by 2 min on ice. The cDNA was purified using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA), following the manufacturer's instructions. For the amplification step, 5 µl of cDNA were mixed with 1 µM of the multiplex identifier (MID) (François et al., 2018), 1x reaction buffer, 0.50 µl dNTPs, and 1.25 U of DreamTaq polymerase (Thermo Fisher Scientific) for a total 50 µl volume reaction. The tubes were heated at 94ºC for 1 min, at 65ºC for 0 s, 72ºC for 45 s, with a slope of 5ºC per second, followed by 40 cycles of 94ºC for 0 s, 45ºC for 0 s, 72ºC for 5 min (same slope); 72ºC for 5 min, and 37ºC for 5 min. The PCR products were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), following the manufacturer's instructions.

Library preparation to add the Illumina adapters and prepare the samples for sequencing was performed at the Center of Biomedical Research of Liège University (GIGA, Liège, Belgium) using the TruSeq PCR-free library preparation kit and sequenced with the Illumina Novaseq sequencing platform, with a read length of 150 base pairs (bp) paired-end.

2.3.2 Total RNA extraction and sequencing

Total RNAs were extracted using the RNeasy Plus Mini Kit (Qiagen, Valencia, CA), following the manufacturer's instructions. Before sequencing, purified RNAs were then treated with amplification grade DNase I (Invitrogen Life Technologies) by adding 1 µl DNase I (1 U/ µl) and 1 µl 10x DNase I reaction buffer for 1 µg of RNA sample in a 10 µl reaction. The samples

were incubated for 15 min at room temperature, and the DNase I was inactivated by adding 1 µl of 25mM EDTA solution and incubating at 65ºC for 10 min. Library preparation was performed at the Interdisciplinary Center of Biomedical Research of Liège University (GIGA, Liège, Belgium) with the TruSeq Stranded Total RNA Library Prep Plant (Illumina). After quantification and quality control, the prepared libraries were sequenced with the Illumina NextSeq 500 sequencing platform, with a read length of 150 base pairs (bp) paired-end.

2.4 HTS data analyses

After demultiplexing (Lebas et al., 2022), the sequencing reads' quality was checked using FastQC in Galaxy Europe (https://usegalaxy.eu) (The Galaxy Community, 2022). Then, the following analyses were done on Geneious Prime 2022 (Biomatters Ltd, Auckland, New Zealand). Reads were merged and cleaned by trimming the bases with quality below 30 and removing reads with length below 35 nucleotides, using BBDuk Adapter/Quality Trimming (Kechin et al., 2017) version 38.84. Duplicated reads were removed using Dedupe Duplicate Read Remover (Bushnell et al., 2017) version 38.84 with k-mer seed length set to 31. Following the quality control, trimming, and read cleaning steps, a de novo assembly of reads into contigs was performed using RNA SPAdes v. 3.15.5, with default parameters. Contigs of potential viral origin were annotated using tBLASTx against the viral RefSeq database from NCBI (nt) downloaded in April 2023 (release number 216). Mapping to reference genomes from NCBI or reconstructed genomes from *de novo* assembly was done using Geneious Prime 2022, allowing for 10 % mismatches for Citrus virus A (CiVA, *Coguvirus eburi*) (RNA 1: OR825541, RNA 2: OR825542) and apple rubbery wood virus 1 (ARWV-1, *Rubodvirus mali*) (segment L: OK398019, segment M: OK398020, segment S: OK398021); and 20% mismatches for apple stem pitting virus (ASPV) (NC_003462), apple chlorotic leaf spot virus (ACLSV) (NC_001409), phaseolus vulgaris endornavirus 1 (PvEV-1) (NC_039217), phaseolus vulgaris endornavirus 2 (PvEV-2) (NC_038422), phaseolus vulgaris endornavirus 3

(PvEV-3) (NC_040558), and putative virus Pyrus virus A (PyVA) (OR887735), to take into genetic variability within each virus species. The conserved protein domains and the protein functional analysis were predicted using InterProScan (Paysan-Lafosse et al., 2023).

To set a threshold for cross-contamination, the mapped reads per kilobase per million (RPKM) were calculated according to Mortazavi et al. (2008), allowing the normalization and comparison of the viruses' detection from each pool (Figure 1A). Based on RPKM values of the alien viruses, the cross-contamination ratios between each sample (from 1 to n samples) and the alien controls (ratio alien, $RA_{a\ 1\rightarrow n}$) was obtained by dividing the RPKM of the alien virus in the sample (true contamination / false positive; $RPKM_{a\ 1\rightarrow n}$) by the RPKM of the alien virus in the external alien control library (true positive; $RPKM_{a\ max}$) (Figure 1B). The ratios were calculated independently for the 3 alien viruses: PvEV-1, 2, and 3 (3n ratios in total). For each detected virus, a cross-contamination ratio was also calculated for each sample (ratio virus, $RV_{x\ 1\rightarrow n}$) by dividing the RPKM of the virus in the sample ($RPKM_{x\ 1\rightarrow n}$) by the maximal RPKM of the virus among the samples ($RPKM_{x\ max}$) (Figure 1C), which was considered as the likely source of contamination. Viruses with $RV_{x\ 1\rightarrow n}$ below the threshold for cross-contamination were considered likely false positives (FP) in the pool, and viruses with $RV_{x\ 1\rightarrow n}$ above the threshold were considered true positives (TP).

2.5 Phylogenetic tree reconstruction

Using the heat shock protein 70 homolog (HSP70h) amino acid(aa) sequence of known viruses within the family *Closteroviridae* and that from the reconstructed genome of PyVA a maximum likelihood phylogenetic tree was constructed on the Galaxy Europe server using MAFFT to generate the multiple alignment and IQ-TREE version 2.1.2 with 1,000 bootstrap replicates to reconstruct the tree (Minh et al., 2020). ModelFinder (Kalyaanamoorthy et al., 2017) was used to determine the best substitution model for the HSP70h amino acid sequences alignment

(rtREV+F+I+G4). The tree was visualized using the iTOL v6.8 tool (https://itol.embl.de/) (Letunic and Bork, 2021).

2.6 Molecular detection of viruses by reverse transcription PCR (RT-PCR)

Complementary DNA (cDNA) synthesis was performed for each sample by denaturing 2 µl of extracted total RNA (protocol described in section 2.2.2) and 2.5 µM of random hexamers (Invitrogen) for a reaction volume of 12 µl, at 65ºC during 5 min and then keeping the samples on ice during 1 min. 0.5 mM of dNTPs, 1x reaction buffer, 10 U Superscript III Reverse Transcriptase (Invitrogen), 2 U/µl RNaseOUT Recombinant RNase Inhibitor (Invitrogen), and 5 mM DTT in a 20 µl total reaction volume were added. After incubating at 25ºC for 5 min and 50ºC for 45 min, the RT was inactivated at 70ºC for 15 min, followed by 2 min on ice. Then 2 µl of cDNA, 1X reaction buffer, 0.2 µM forward primer (Table 1), 0.2 µM reverse primer (Table 1), 0.2 mM dNTPs, 2 mM MgCL2, and 0.1 U/µl Mango Taq DNA Polymerase (Bioline Reagents Ltd.) were added in a total reaction volume of 20 µl. The PCR products were visualized after migration in a 1% agarose gel.

The PCR products of one positive sample of CiVA from Slovenia, four positive sample of PyVA from Belgium, the positive sample of ARWV-1, and one of CiVA from Switzerland were purified using the Nucleospin Gel and PCR clean up (Macherey-Nagel, Germany), following the manufacturer's instructions, and sent to be sequenced by Sanger sequencing at Macrogen Europe BV (The Netherlands). The positive sample of apple rubbery wood virus 2 (ARWV-2, *Rubodvirus prosserense*) from Switzerland was purified from the agarose gel using the Wizard® SV Gel and PCR Clean-Up System (Promega) and sent to be sequenced by Sanger sequencing at Fasteris (Switzerland). More information regarding the samples tested can be found in Supplementary Table 1.

2.7 Grafting assays

To study the host range of PyVA, grafting assays were conducted in the field at the biological indexing facilities of Agroscope (Switzerland). The scions were taken from the tree with a single infection by PyVA (CRA-W accession number 626). Absence of other commonly found viruses, namely ASPV and ACLSV, was determined by HTS and RT-PCR (Table 1). In August 2022, these scions were grafted on the following indicators: Virgina crab apple, Lord Lambourne (*Malus domestica*), *Pyronia veitchii* (*Cydonia oblonga* x *Pyrus communis*), Beurré Hardy (*Pyrus communis*), Williams (*Pyrus communis*), A20 (*Pyrus communis*), Jules d'Airoles (*Pyrus communis*), and C7/1 (*Cydonia oblonga*) using four replicates for each variety. In March 2023, buds collected from the same original pear tree were also grafted by chip-budding on the following indicators: Pyrodwarf (*Pyrus communis*), M9 (*Malus domestica*), St. Julien (*Prunus domestica*), *Cydonia oblonga* and Gisela® 5 (*Prunus cerasus* x *Prunus canescens*). The grafted plants were kept under greenhouse conditions at the CRA-W (Belgium) and tested by RT-PCR for presence or absence of PyVA.

2.8 Transmission electron microscopy (TEM)

Purification of the particles of PyVA was performed according to the protocol from Pilotti et al. (1995), with some modifications. Briefly, 40 g of infected pear leaves and petioles (CRA-W accession number 626) were ground into powder using liquid nitrogen and a mixer (Sorvall Omni Mixer 17150 Homogenizer). Then, the powder was mixed with 6 volumes of extraction buffer (0.5 M Tris, pH 8.2, 5% v/v Triton, 4% v/v Polyclar AT, 0.5% w/v bentonite, 0.2% v/v β-mercaptoethanol). After 20 min of homogenization, the suspension was filtered through a double layer of cotton cloth, and the resulting filtrate was centrifuged at 4,000 rpm for 20 min. The supernatant was collected and transferred to an ultracentrifuge tube, and 5 ml of a 20% sucrose cushion (prepared in 0.1 M Tris, pH 8.2) were added. The tube was centrifuged at 40,000 rpm for 1 h and 30 min. The resulting pellet was incubated overnight at 4°C in 4 ml of

10x resuspension buffer (0.02 M Tris, pH 7.0, 1mM MgCl2). Three microliters were mixed with one volume of 0.1% of bovine serum albumin and one volume of 4% phosphotungstic acid (pH 6.0). Purified particles were observed by transmission electron microscopy (TEM) as described by Mahillon et al. (2023), using a Tecnai G2 Spirit microscope (FEI, Eindhoven).

3 Results

3.1 Viruses detected by high-throughput sequencing

After *de novo* assembly, between 29 and 1,537 contigs were obtained per pool for the 17 pools of samples prepared with the dsRNA virus enrichment protocol, with an average of 610 contigs per pool. An average of 191 contigs longer than 1,000 nucleotides was obtained, with an average N50 length of 1,495 (Supplementary Table 2). Using tBLASTx, four known and one unknown pear viruses were detected in the analyzed samples: ASPV, ACLSV, ARWV-1, CiVA, and a tentative novel *Closteroviridae* member which will be referred from this point on as Pyrus virus A (PyVA).

The cross-contamination ratios of the alien control ($RA_{1\ 1\rightarrow n}$, $RA_{2\ 1\rightarrow n}$, and $RA_{3\ 1\rightarrow n}$) ranged between 0 and 0.5%. Moreover, only 1 read (1.5 RPKM, 0.4% $RV_{ACLSV\ alien}$) from a pear virus (ACLSV) was found in the alien control, used in this case as a negative control, thus strengthening the set threshold of 0.5% for likely false positives (FP). Ratios below or equal to the threshold and above 0% ($0\%<RV_{x\ 1\rightarrow n}\leq 0.5\%$) were considered FP, ratios above the threshold were considered TP ($RV_{x\ 1\rightarrow n}>0.5\%$), and ratios of 0% ($RV_{x\ 1\rightarrow n}=0\%$) were considered true negatives (TN). In addition, confirmatory targeted molecular tests using RT-PCR were applied to the sequenced samples, which have been analyzed individually. It is worth mentioning that pool L7-2 had an unexpectedly high RA ratio (56.6%) for only one of the three alien viruses, namely PvEV-1 (Supplementary Table 2). This result was considered as aberrant and was discarded for two reasons: (i) no reads of PvEV-2 and PvEV-3 were observed for this pool (while the alien control showed more reads for these two viruses compared to PvEV-1),

and (2) the detected region of PvEV-1 (316 nt representing 2% of the genome) is not covered by any read in the alien control.

Out of the 17 pools analyzed, ASPV and ACLSV were detected by mapping in respectively 17 and 14 pools, while ARWV-1 and CiVA were both detected in six pools. PyVA was detected by mapping in two of the 17 pools analyzed (Table 2). The positive pools where CiVA and PyVA were detected by HTS contained at least one positive sample of the virus that was tested by RT-PCR during the field survey in 2023, although five pools that were negative for CiVA by HTS contained one or more samples found positive by RT-PCR and one pool that was considered a likely FP for PyVA by HTS contained one sample found positive by RT-PCR (Table 2). An extensive list of the sampled germplasm, the designed pools, and their corresponding libraries, as well as the tests performed and the viruses detected in each sample during the field survey, can be found in Supplementary Table 1.

To assemble the complete genome of PyVA, total RNA from leaves of samples included in pool L3-2 were extracted and Illumina sequenced (see section 2.2.2). Two genomic sequences of PyVA were reconstructed from cultivar (cv.) Jean Nicolas (tree Z14, accession 224; and tree Z15, accession 621) and showed 99.9% identity (Genbank OR887735-6). In addition, a nearly complete genomic sequence of a Belgian isolate of CiVA was also reconstructed from accession 224, with RNA 1 having a length of 6,663 nt (Genbank OR825541) and RNA 2 a length of 2,721 nt (Genbank OR825542).

3.2 A novel velarivirus infecting pear trees

3.2.1 Molecular and genomic characterization

The two assembled PyVA genomes include all coding regions, but, although attempted, the 3' and 5' UTRs could not be fully assembled. The assembled genome of isolate 621-BE has a length of 17,061 nucleotides (nt) and 3,527 of the RNASeq reads map on the genome, representing 0.09% of the total reads, with an average coverage depth of 31X. The assembled

genome of isolate 224-BE has a length of 17,142 nt and 1,228 RNASeq reads map on the genome, representing 0.03% of the total reads, with an average coverage depth of 11X. The size difference of the two isolates is due to incomplete sequencing of the 5' and 3' untranslated regions (UTRs). In the two assembled genomes the partial 5' UTR has a length of 49 nt. On the other hand, the partial 3' UTR of isolate 621-BE has a length of 209 nt and that of isolate 224-BE 290 nt. The two assembled genomes encode nine putative open reading frames (ORFs) (Figure 2A) and have a genomic organization resembling that of its closest relative, Malus domestica virus A (MdoVA), and of other velariviruses (Figure 2A). Similar to MdoVA, more reads were mapped at the 3' end of the genome (Koloniuk et al., 2020). Electron microscopy observation of semi-purified viral particles showed them to have a length of approximately 2000 nm (Figure 2B).

Within the complex ORF1a-ORF1b, two replication-associated domains, the methyltransferase (Mtr) (PF01660) and helicase (Hel) (PF01443), were identified in the product of ORF 1a. This ORF is 7,047 nt long (2,214 aa) and encodes a protein that would weigh 255 kDa. In ORF 1b, which is proposed to be translated through a +1 ribosomal frameshift, an RNA-dependent RNA polymerase (RdRp) (PF00978) conserved domain was identified. For most members of the family *Closteroviridae* the proposed frameshift motif for ORF 1b expression is "GUU_stop_C", which is present in all members of the genus *Velarivirus*, including PyVA (isolate 621-BE, nucleotides 7,091-7,097 "GUU_UGA_C"). However, since this potential shifting motif has not been proven experimentally, ORF 1b was annotated as a separate ORF. ORF 1b is 1,515 nt long (504 aa) and encodes a protein with a molecular weight of 58 kDa. Then there is the putative ORF 2 that would encode the small protein p4 of 102 nt (34 aa) with a molecular weight of 4 kDa. ORF 3 is 1,662 nt long and encodes a 524 aa (61 kDa) protein of the viral heat shock protein 70 homolog family (PF00012 HSP70h), and ORF 5 (1,554nt) encodes a protein of the viral heat shock protein 90 homolog family (PF03225, HSP90h) of 56

kDa (478 aa). ORF 4 (219 nt) is predicted to slightly overlap with ORF 3 and encodes the putative protein p9 of 9 kDa (72 aa). ORF 6 (1,029 nt) and ORF 7 (2,106 nt) respectively encode the structural capsid (CP) (PF01785) of 38 kDa (342 aa) and the minor capsid (CPm) protein of 75 kDa (648 aa). The last two ORFs, 8 (693 nt) and 9 (726 nt), encode the putative proteins p25 of 25 kDa (212 aa) and p26 of 26 kDa (225 aa).

3.2.2 Phylogenetic relationship within the family *Closteroviridae*

The newly identified virus fits the demarcation criteria, as well as the distinguishing properties, for viruses belonging to the genus *Velarivirus* and family *Closteroviridae* (Fuchs et al., 2020). The closest relative to PyVA is Malus domestica virus A (MdoVA), with 75% aa identity in the RdRp, 60% in the HSP70h and 41% identity in the CP (Table 3). The two p25 and p26 putative proteins are the most variable within the genus, showing <42% identity for the p25 and <73% for the p26 between all velariviruses (Supplementary Table 3). In addition, a phylogenetic tree using the HSP70h amino acid sequences of known members of *Closteroviridae* was constructed, which confirmed that PyVA was clustering with strong bootstrap support with other velariviruses (Figure 3).

3.2.3 Biological characterization of the new virus Pyrus virus A (PyVA)

To provide some insights in the biology of PyVA, its distribution within a tree, host range and preliminary geographic distribution were studied together with its graft transmissibility to pear trees. PyVA was successfully transmitted by chip budding to healthy plants of *Pyronia veitchii* (*Cydonia oblonga* x *Pyrus communis*) with a transmission rate of 25% (1/4), and to different pear cultivars (Beurré Hardy: 50% (2/4); Williams: 75% (3/4); A20: 75% (3/4); Jules d'Airoles: 50% (2/4)). None of the grafted plants infected with PyVA developed symptoms during their first year of growth. There was no graft transmission of PyVA to the following indicators: Virginia crab apple, Lord Lambourne (*Malus domestica*), C7/1 (*Cydonia oblonga*), Pyrodwarf (*Pyrus communis*), M9 (*Malus domestica*), St. Julien (*Prunus domestica*), and Gisela® 5

(*Prunus cerasus* x *Prunus canescens*). Within the original infected tree, PyVA was detected in all individual flower and leaf samples sampled from tree X18, and in phloem collected from branches facing North, East and West, suggesting that it is quite evenly distributed in that tree. Distribution within the orchard was rather sparse, as only six trees were found to be infected with PyVA, representing five different cultivars (Supplementary Figure 1). Out of the five trees of cultivar Beau Présent tested, only one was positive (1/5); of the three trees of cultivar Jean Nicolas tested, two were positive (2/3), the unique trees tested for cultivars Camberlain Blanc, Semis Henin, and Poire Grognet were positive as well. Moreover, except for the two trees of cultivar Jean Nicolas, the positives trees were not contiguous in the orchard.

3.3 Field survey by reverse transcription PCR (RT-PCR)

During the bioinformatic analyses of the HTS data, two known and widely distributed viruses (ASPV and ACLSV), one new virus (PyVA), and two recently discovered viruses (CiVA and ARWV-1) were detected in the tested Belgian samples. ARWV-1 had previously been identified in the CRA-W germplasm collection during a preliminary survey (Fontdevila Pareta et al., 2022), thus further efforts in Belgium focused on the distribution and prevalence of CiVA and PyVA. With the observed prevalence of ARWV-1 in the Belgian collection, the samples originating from Slovenia and Switzerland were also tested for ARWV-1. In addition, as one sample tested positive in Switzerland for ARWV-1 which is occasionally present in mixed infections with ARWV-2, it was also tested for ARWV-2.

PyVA was detected in six of the 99 trees sampled (6%) and CiVA in 49 of them (49%). In three trees (3%) there was a mixed infection of CiVA and PyVA, and 47 trees (47%) did not test positive for CiVA or PyVA (Supplementary Table 1). From the six pear trees sampled in Slovenia, four were positive for CiVA; and from the nine trees sampled in Switzerland, four were positive for CiVA and one for ARWV-1. Additionally, the sample that tested positive for ARWV-1 also tested positive for ARWV-2.

Specificity of the amplified PCR products of PyVA from Belgium, CiVA from Slovenia, and ARWV-1, -2, and CiVA from Switzerland was confirmed by Sanger sequencing. The partial RdRp sequence of CiVA from Slovenia, isolate KP-SLO-1 (Genbank OR825539), had an amplicon size of 655 nt and 99.1% nucleotide (nt) identity with RNA 1 of CiVA isolate B175 (Genbank MZ463039). From Switzerland, the partial CP sequence of ARWV-1, isolate 39652-CH (Genbank OR825538), had an amplicon size of 777 nt and 99.6% nt identity to segment S of ARWV-1 isolate BR-Mishima (Genbank MK936225), the partial CP sequence of ARWV-2, isolate 39652-CH (Genbank PP319005), had an amplicon size of 228 nt and 100% nt identity to segment Sa of ARWV-2 isolate CE30 (Genbank OP583932), and the partial RdRp sequence of CiVA, isolate 39653-CH (Genbank OR825540), had an amplicon size of 655 nt and 97.4% nt identity to RNA 1 of CiVA isolate P215/CIVA (Genbank MZ330076). From Belgium, partial CP sequences of PyVA from isolates 615-BE, 626-BE, and 638-BE (Genbank OR936022-4) had a length of 436 nt and 97.7% nt identity to isolate 224-BE (Genbank OR887735); while sequences from isolate 847-BE (Genbank OR936025) showed 98.2% nt identity to isolate 224-BE (Genbank OR887735).

4 Discussion

Scanning the virome of fruit tree germplasm collections has become possible with the evolution of HTS technologies and bioinformatic tools (Adams et al., 2018; Rott et al., 2017). Such approach is a useful tool to evaluate the viruses present in these germplasms before their potential use in breeding, provided the detection test is reliable.

In this context, an international initiative proposed guidelines for the reliable use of HTS technologies to detect plant pathogens and pests. One of the innovative aspects in these guidelines was the use of an alien control to monitor the levels of cross-contamination to differentiate between true and false positives (Massart et al., 2022). In this study, the dsRNA viral enrichment protocol, based on the binding properties of dsRNA to cellulose, was used to

analyze the virome of pear germplasms as it has proven very useful for large scale virome analyses (Marais et al., 2024; Schönegger et al., 2023). Because this protocol relies on numerous steps, there is an increased risk of cross-contamination between samples. Anticipating the potential uncertainty when identifying FP arising from cross-contamination events, an alien control strategy was adopted. To our knowledge, the present study is the first time this specific approach has been used in combination with the dsRNA viral enrichment protocol. To evaluate cross-contamination, a contamination ratio was calculated for the alien viruses ($RA_{1\rightarrow n}$) and the detected pear viruses ($RV_{x\ 1\rightarrow n}$). This ratio was inspired by Cont-ID, a tool designed to evaluate cross-contamination between samples/pools (Rollin et al., 2023), and created by reusing and adapting some of its principles to the dataset without the duplication step. In our case, the maximum cross-contamination background level using the alien viruses was between 0.4 and 0.5%, with the later value used as the cross-contamination threshold to distinguish true from likely false positives. In total, 14 detection events (related to ASPV, ACLSV and PyVA), out of a total of 48, were considered as likely false positive using this threshold.

To reinforce the reliability of detection, a second survey (based on RT-PCR tests) was carried out in 2023. All positive HTS results were confirmed for CiVA and PyVA (n=8) while five and one negative HTS results contained at least one positive sample in the pool for CiVA and PyVA, respectively. Therefore, total RNA was extracted from leaf material of individual trees from 2021 after 2-year storage and tested by RT-PCR for detection confirmation. CiVA was detected in nine of the 12 individual trees, confirming the presence of CiVA in those samples, and PyVA was detected in four of the seven individual trees, thus confirming its presence. Considering these results five pools considered as true negatives for CiVA and one as a likely false positive for PyVA based in HTS results were found to be true positives following RT-PCR analysis of the plants constituting these pools, challenging the diagnostic sensitivity of

the HTS test, although no absolute conclusion can be advanced since, for example, the heterogeneity of the virus's titer within tested trees is not known. Notably, dsRNA-based HTS indexing has been shown to have a lower sensitivity for negative-stranded RNA viruses (-ssRNA), such as ARWV-1 and CiVA (Marais et al., 2024; Schönegger et al., 2023), so that reliable detection of such viruses in fruit trees may necessitate the testing of individual trees - Marais et al., 2024) and not of pools of samples as was done here.

The virome study of the pear collection from the CRA-W in Belgium strengthened the assumption that there is a lower number of viruses that infect pear trees compared to other fruits trees, such as apple, given that only five viruses were detected in a total of 65 trees of diverse origins and genetic backgrounds analyzed by HTS. In contrast, a virome study of experimental and commercial apple orchards, of 18 cultivars, in British Columbia detected 21 plant viruses and one plant viroid by HTS (Xiao et al., 2022). Moreover, this study ratified the fact that pooling samples prior to extraction is a powerful approach to reduce the costs when performing large scale virome surveys, as seen by Fowkes et al. (2021) and Nyirakanani et al. (2021), although false negative can occur, underlining the importance of using an alien control to monitor the cross-contamination and accuracy of the test (Massart et al., 2022). For the most important detections, downstream individual testing on trees tested negative or positive by HTS on pooled samples, is recommended for a better virome characterization and an improved reliability of the findings.

This study describes a putative novel velarivirus tentatively named Pyrus virus A (PyVA) identified in pear (*Pyrus communis*) trees and the first report of CiVA in Slovenia, Switzerland, and Belgium. Moreover, it represents the first molecular evidence of presence of ARWV-1 and -2 in Switzerland. As a suggestion to the International Committee on Taxonomy of Viruses (ICTV), the authors tentatively propose the latinized form Velarivirus gembloutense, for its species name, after the region it was detected. Within the genus *Velarivirus*, the assembled

sequences of PyVA have less than 75 % aa identity to the RdRp, CP, and HSP70h products, which are the relevant gene products chosen by the ICTV for the species demarcation criteria of the genus (Fuchs et al., 2020). The size of the viral particle, and the genome structure and organization are similar to other viruses of the genus *Velarivirus* in the family *Closteroviridae*. Additionally, to our knowledge, this study provides the first TEM of a *Velarivirus*. There is a rather high diversity of genome organization and similarity levels within the genus. The most conserved ORF is ORF 1b, with a percentage of aa identity between 51% and 78% (Supplementary Table 3). In contrast, putative proteins p25 and p26 are the most divergent, with a percentage of aa identity lower than 42% and 73% respectively (Supplementary Table 3). Overall, the two closest viruses within the genus are PyVA and MdoVA, followed by Cordyline virus 1 and Cordyline virus 2, which could be due to their adaptations to similar plant hosts over time or their divergence from a common ancestor infecting the same plant host.

One of the characteristics of the family *Closteroviridae* is the hypothetical expression of the RdRp domain, encoded by ORF 1b, through a +1 ribosomal frameshift. This assumption has not been experimentally demonstrated; hence the identity of the frameshifting site remains speculative. Nevertheless, it has been proposed that in most members of this family the +1 ribosomal frameshift could occur at a conserved "GUU_stop_C" motif that includes the ORF 1a stop codon. This motif would promote a slippage of the ribosome from GUU to UUU (Agranovsky, 2016; Maia et al., 1996). To be cautious, in this case, ORFs 1a and 1b of PyVA were annotated individually but with a note explaining the alternative possibility of the RdRp domain to be encoded through a +1 ribosomal frameshift.

While other studies have shown a high genetic diversity within the *Closteroviridae* family (Liu et al., 2021), the set of sequences obtained in this study is too small and originating from the

same location, thus the sequences are expected to present low genetic variability. Additionally, the available data is not enough to draw any conclusions in this direction.

The characterization of PyVA was performed following the revised framework for the characterization of a novel plant virus or viroid discovered in an HTS dataset (Fontdevila Pareta et al., 2023). First, the first two steps of the framework, consisting of (i) the design of a detection test, confirmation of the detection and obtention of the genome sequence; and (ii) gathering contextual information and first notification and discussion with stakeholders, including the scientific community and plant protection agencies, were carried out. PyVA was suggested as a "no priority" virus since the infected trees in the field did not present symptoms, and there was no apparent rapid spread. To evaluate the association between the presence of the virus and symptoms in the plant host, as part of step three of the framework, a large-scale field survey and transmission assays were conducted. Again, no symptoms were observed on trees infected with PyVA alone or in mixed infection with other viruses (ASPV, ACLSV, and CiVA). Moreover, there were no symptoms observed on the graft inoculated trees with PyVA at Agroscope. It would be possible that symptoms may develop after some years, although it is unlikely as no visible symptoms on leaves and fruits were observed during sampling campaigns in 2021 and 2023. Monitoring of symptoms was also done during summer 2022, with no symptoms that could be linked to PyVA. Regarding transmission between different cultivars of *Pyrus communis*, even though it was possible, the transmission rate was not 100%. Other studies showed that there can be a high variability in the rate of transmission by grafting between fruit tree viruses (Khalili et al., 2023). The biological characteristics of the new virus and the results obtained in the course of the study argue in favor of a rather low phytosanitary risk of PyVA, indicating that it may not require immediate action from pest risk managers and authorities.

Moreover, the large-scale field survey assisted in the completion of data gaps for CiVA, ARWV-1, and ARWV-2 by providing insight into the prevalence of CiVA and the geographic distribution of CiVA, ARWV-1, and ARWV-2. This study reports the first detection of CiVA in Belgium, Switzerland, and Slovenia, and the first molecular evidence of presence of ARWV-1 and ARWV-2 in Switzerland. In citrus trees, CiVA has been found associated with disease symptoms (Beris et al., 2021; de Bruyn et al., 2022; Park et al., 2022), but its association with symptoms in pear trees has not been proven. CiVA was found with a high prevalence in collections in Belgium, Switzerland, and Slovenia, in trees showing no visible symptoms on leaves and/or fruits, although more samples from Switzerland and Slovenia should be tested to strengthen the risk evaluation process. Even though the collections targeted here were selected mainly to prioritize the study of the virome in ancient and local pear cultivars, commercial orchards could be studied to evaluate if a similar trend in the distribution and prevalence of these viruses is observed.

In conclusion, this study provides further evidence of the interest of virome survey of germplasm collections using HTS approaches while simultaneously showcasing the limitations that still exist. A proper evaluation of the virus infection status of these collections should become a cornerstone before evaluating their genetic potential. Besides detecting known viruses, including some poorly described ones, our study revealed an unknown virus in latent or asymptomatic infections, bringing valuable insights into the diversity and complexity of viral infections in pear trees. The downstream characterization of such new or poorly characterized viruses should accompany any virome survey as such additional studies can provide useful information to the stakeholders to evaluate the potential phytosanitary risk posed by the detected viruses.

Acknowledgments

The authors would like to thank Chantal Faure for her support with the training in the double-stranded RNA extraction protocol; Johan Rollin and Coline Temple for their help during sampling at the collection of the CRA-W (Gembloux, Belgium); Angelo Locicero, Vanessa Derycker, and Igor Lattenist for assisting in the total RNA extractions; and the personnel at Kozjanski Park (Slovenia) for the pear samples used in this study.

List of references

FAO, 2024. FAOSTAT [WWW Document]. Food Agric. Organ. U. N. URL https://www.fao.org/faostat/en/#data/QCL (accessed 11.5.23).

Adams, I.P., Fox, A., Boonham, N., Massart, S., De Jonghe, K., 2018. The impact of high throughput sequencing on plant health diagnostics. Eur. J. Plant Pathol. 152, 909–919. https://doi.org/10.1007/s10658-018-1570-0

Agranovsky, A.A., 2016. Closteroviruses: Molecular Biology, Evolution and Interactions with Cells, in: Gaur, R.K., Petrov, N.M., Patil, B.L., Stoyanova, M.I. (Eds.), Plant Viruses: Evolution and Management. Springer, Singapore, pp. 231–252. https://doi.org/10.1007/978-981-10-1406-2_14

Beris, D., Ioanna, M., Vassilakos, N., Theologidis, I., Rampou, A., Kektsidou, O., Massart, S., Varveri, C., 2021. Association of Citrus Virus A to Citrus Impietratura Disease Symptoms. Phytopathology 111, 1782–1789. https://doi.org/10.1094/PHYTO-01-21-0027-R

Bester, R., Karaan, M., Cook, G., Maree, H.J., 2021. First report of citrus virus A in citrus in South Africa. J. Citrus Pathol. 8. https://doi.org/10.5070/C481049000

Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J., Mumford, R., 2014. Methods in virus diagnostics: From ELISA to next generation sequencing. Virus Res., Plant Virus Epidemiology Symposium Ecology, evolution and control of plant viruses and their vectors 186, 20–31. https://doi.org/10.1016/j.virusres.2013.12.007

Bushnell, B., Rood, J., Singer, E., 2017. BBMerge – Accurate paired shotgun read merging via overlap. PLoS ONE 12. https://doi.org/10.1371/JOURNAL.PONE.0185056

Candresse, T., Lanneau, M., Revers, F., Macquaire, G., German, S., Dunez, J., Grasseau, N., Malinovsky, T., 1995. An immunicapture PCR assay adapted to the detection and the analysis of molecular variability of apple chlorotic leaf spot virus. Acta Hortic. 136–147. https://doi.org/10.17660/ActaHortic.1995.386.17

Costa, L.C., Hu, X., Malapi-Wight, M., Foster, J., McFarland, C., Hurtado-Gonzales, O.P., 2022. Identification of a novel robigovirus and a Prunus-infecting tepovirus in Pyrus communis and their transmissibility on Malus spp. Eur. J. Plant Pathol. 162, 275–288. https://doi.org/10.1007/s10658-021-02402-9

de Bruyn, R., Bester, R., Cook, G., Steyn, C., Breytenbach, J.H.J., Maree, H.J., 2022. Distribution and Genetic Diversity of Coguvirus eburi in South African Citrus and the

Development of a Real-Time RT-PCR Assay for Citrus-Infecting Coguviruses. Plant Dis. 106, 2221–2227. https://doi.org/10.1094/PDIS-11-21-2409-RE

Fontdevila Pareta, N., Khalili, M., Maachi, A., Rivarez, M.P.S., Rollin, J., Salavert, F., Temple, C., Aranda, M.A., Boonham, N., Botermans, M., Candresse, T., Fox, A., Hernando, Y., Kutnjak, D., Marais, A., Petter, F., Ravnikar, M., Selmi, I., Tahzima, R., Trontin, C., Wetzel, T., Massart, S., 2023. Managing the deluge of newly discovered plant viruses and viroids: an optimized scientific and regulatory framework for their characterization and risk analysis. Front. Microbiol. 14.

Fontdevila Pareta, N., Lateur, M., Steyer, S., Blouin, A. g., Massart, S., 2022. First reports of Apple luteovirus 1, Apple rubodvirus 1 and Apple hammerhead viroid infecting apples in Belgium. New Dis. Rep. 45, e12076. https://doi.org/10.1002/ndr2.12076

Fowkes, A.R., McGreig, S., Pufal, H., Duffy, S., Howard, B., Adams, I.P., Macarthur, R., Weekes, R., Fox, A., 2021. Integrating High throughput Sequencing into Survey Design Reveals Turnip Yellows Virus and Soybean Dwarf Virus in Pea (Pisum Sativum) in the United Kingdom. Viruses 13, 2530. https://doi.org/10.3390/v13122530

François, S., Filloux, D., Fernandez, E., Ogliastro, M., Roumagnac, P., 2018. Viral Metagenomics Approaches for High-Resolution Screening of Multiplexed Arthropod and Plant Viral Communities, in: Pantaleo, V., Chiumenti, M. (Eds.), Viral Metagenomics: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 77–95. https://doi.org/10.1007/978-1-4939-7683-6_7

Fuchs, M., Bar-Joseph, M., Candresse, T., Maree, H.J., Martelli, G.P., Melzer, M.J., Menzel, W., Minafra, A., Sabanadzovic, S., Report Consortium, I., 2020. ICTV Virus Taxonomy Profile: Closteroviridae. J. Gen. Virol. 101, 364–365. https://doi.org/10.1099/jgv.0.001397

Hou, W., Li, S., Massart, S., 2020. Is There a "Biological Desert" With the Discovery of New Plant Viruses? A Retrospective Analysis for New Fruit Tree Viruses. Front. Microbiol. 11, 2953. https://doi.org/10.3389/fmicb.2020.592816

Jelkmann, W., Paunovic, S., 2011. Apple stem pitting virus, in: Virus and Virus-Like Diseases of Pome and Stone Fruits, Virology. The American Phytopathological Society, p. 428.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285

Katsiani, A., Maliogka, V.I., Katis, N., Svanella-Dumas, L., Olmos, A., Ruiz-García, A.B., Marais, A., Faure, C., Theil, S., Lotos, L., Candresse, T., 2018. High-Throughput Sequencing Reveals Further Diversity of Little Cherry Virus 1 with Implications for Diagnostics. Viruses 2018 Vol 10 Page 385 10, 385–385. https://doi.org/10.3390/V10070385

Kechin, A., Boyarskikh, U., Kel, A., Filipenko, M., 2017. CutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. J. Comput. Biol. 24, 1138–1143. https://doi.org/10.1089/CMB.2017.0096/ASSET/IMAGES/LARGE/FIGURE3.JPEG

Khalili, M., Candresse, T., Koloniuk, I., Safarova, D., Brans, Y., Faure, C., Delmas, M., Massart, S., Aranda, M.A., Caglayan, K., Decroocq, V., Drogoudi, P., Glasa, M., Pantelidis, G., Navratil, M., Latour, F., Spak, J., Pribylova, J., Mihalik, D., Palmisano, F., Saponari, A., Necas, T., Sedlak, J., Marais, A., 2023. The Expanding Menagerie of Prunus-Infecting Luteoviruses. Phytopathology 113, 345–354. https://doi.org/10.1094/PHYTO-06-22-0203-R

Koloniuk, I., Přibylová, J., Fránová, J., Špak, J., 2020. Genomic characterization of Malus domestica virus A (MdoVA), a novel velarivirus infecting apple. Arch. Virol. 165, 479–482. https://doi.org/10.1007/s00705-019-04478-5

Komorowska, B., Malinowski, T., Michalczuk, L., 2010. Evaluation of several RT-PCR primer pairs for the detection of Apple stem pitting virus. J. Virol. Methods 168, 242–247. https://doi.org/10.1016/j.jviromet.2010.04.024

Kummert, J., Malice, M., Marbot, S., Lepoivre, P., Steyer, S., Oger, R., 2004. Sampling protocols and risk of error significance in molecular detection tests for fruit trees certification. Acta Hortic. 541–546. https://doi.org/10.17660/ActaHortic.2004.657.88

Lebas, B., Adams, I., Al Rwahnih, M., Baeyen, S., Bilodeau, G.J., Blouin, A.G., Boonham, N., Candresse, T., Chandelier, A., De Jonghe, K., Fox, A., Gaafar, Y.Z.A., Gentit, P., Haegeman, A., Ho, W., Hurtado-Gonzales, O., Jonkers, W., Kreuze, J., Kutjnak, D., Landa, B., Liu, M., Maclot, F., Malapi-Wight, M., Maree, H.J., Martoni, F., Mehle, N., Minafra, A., Mollov, D., Moreira, A., Nakhla, M., Petter, F., Piper, A.M., Ponchart, J., Rae, R., Remenant, B., Rivera, Y., Rodoni, B., Roenhorst, J.W., Rollin, J., Saldarelli, P., Santala, J., Souza-Richards, R., Spadaro, D., Studholme, D.J., Sultmanis, S., van der Vlugt, R., Tamisier, L., Trontin, C., Vazquez-Iglesias, I., Vicente, C.S.L., Vossenberg, B.T.L.H., Wetzel, T., Ziebell, H., Massart, S., 2022. Facilitating the adoption of high-

throughput sequencing technologies as a plant pest diagnostic test in laboratories: A step-by-step description. EPPO Bull. 52, 394–418. https://doi.org/10.1111/epp.12863

Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. https://doi.org/10.1093/nar/gkab301

Liu, Q., Zhang, S., Mei, S., Zhou, Y., Wang, J., Han, G.-Z., Chen, L., Zhou, C., Cao, M., 2021. Viromics unveils extraordinary genetic diversity of the family Closteroviridae in wild citrus. PLOS Pathog. 17, e1009751. https://doi.org/10.1371/journal.ppat.1009751

Mahillon, M., Brodard, J., Kellenberger, I., Blouin, A.G., Schumpp, O., 2023. A novel weevil-transmitted tymovirus found in mixed infection on hollyhock. Virol. J. 20, 17. https://doi.org/10.1186/s12985-023-01976-6

Maia, I.G., Séron, K., Haenni, A.-L., Bernardi, F., 1996. Gene expression from viral RNA genomes. Plant Mol. Biol. 32, 367–391. https://doi.org/10.1007/BF00039391

Maliogka, V.I., Minafra, A., Saldarelli, P., Ruiz-García, A.B., Glasa, M., Katis, N., Olmos, A., 2018. Recent Advances on Detection and Characterization of Fruit Tree Viruses Using High-Throughput Sequencing Technologies. Viruses 10, 436. https://doi.org/10.3390/v10080436

Marais, A., Faure, C., Bergey, B., Candresse, T., 2018. Viral Double-Stranded RNAs (dsRNAs) from Plants: Alternative Nucleic Acid Substrates for High-Throughput Sequencing, in: Pantaleo, V., Chiumenti, M. (Eds.), Viral Metagenomics: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 45–53. https://doi.org/10.1007/978-1-4939-7683-6_4

Marais, A., Gentit, P., Brans, Y., Renvoisé, J.P., Faure, C., Saison, A., Cousseau, P., Castaing, J., Chambon, F., Pion, A., Calado, G., Lefebvre, M., Garnier, S., Latour, F., Bresson, K., Grasseau, N., Candresse, T., 2024. Comparative performance evaluation of double-stranded RNA high-throughput sequencing for the detection of viral infection in temperate fruit crops. Phytopathology. https://doi.org/10.1094/PHYTO-12-23-0480-R

Massart, S., Adams, I., Al Rwahnih, M., Baeyen, S., Bilodeau, G.J., Blouin, A.G., Boonham, N., Candresse, T., Chandellier, A., De Jonghe, K., Fox, A., Gaafar, Y.Z.A., Gentit, P., Haegeman, A., Ho, W., Hurtado-Gonzales, O., Jonkers, W., Kreuze, J., Kutjnak, D., Landa, B.B., Liu, M., Maclot, F., Malapi-Wight, M., Maree, H.J., Martoni, F., Mehle, N., Minafra, A., Mollov, D., Moreira, A.G., Nakhla, M., Petter, F., Piper, A.M., Ponchart, J.P., Rae, R., Remenant, B., Rivera, Y., Rodoni, B., Botermans, M.,

Roenhorst, J.W., Rollin, J., Saldarelli, P., Santala, J., Souza-Richards, R., Spadaro, D., Studholme, D.J., Sultmanis, S., van der Vlugt, R., Tamisier, L., Trontin, C., Vazquez-Iglesias, I., Vicente, C.S.L., van de Vossenberg, B.T.L.H., Westenberg, M., Wetzel, T., Ziebell, H., Lebas, B.S.M., 2022. Guidelines for the reliable use of high throughput sequencing technologies to detect plant pathogens and pests. Peer Community J. 2. https://doi.org/10.24072/pcjournal.181

Massart, S., Jijakli, M.H., Kummert, J., 2011. Apple stem grooving virus, in: Virus and Virus-Like Diseases of Pome and Stone Fruits, Virology. The American Phytopathological Society, p. 428.

Massart, S., Olmos, A., Jijakli, H., Candresse, T., 2014. Current impact and future directions of high throughput sequencing in plant virus diagnostics. Virus Res. 188, 90–96. https://doi.org/10.1016/j.virusres.2014.03.029

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015

Minutolo, M., Cinque, M., Di Serio, F., Navarro, B., Alioto, D., 2023. Occurrence of apple rubbery wood virus 1 and apple rubbery wood virus 2 in pear and apple in Campania (southern Italy) and development of degenerate primers for the rapid detection of rubodviruses. J. Plant Pathol. 105, 567–572. https://doi.org/10.1007/s42161-023-01316-1

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 621–628. https://doi.org/10.1038/nmeth.1226

Nyirakanani, C., Bizimana, J.P., Kwibuka, Y., Nduwumuremyi, A., Bigirimana, V. de P., Bucagu, C., Lassois, L., Malice, E., Gengler, N., Massart, S., Bragard, C., Habtu, M., Brostaux, Y., Thonar, C., Vanderschuren, H., 2021. Farmer and Field Survey in Cassava-Growing Districts of Rwanda Reveals Key Factors Associated With Cassava Brown Streak Disease Incidence and Cassava Productivity. Front. Sustain. Food Syst. 5.

Park, J.-W., da Graça, J.V., Gonzalez, M., Louzada, E.S., Alabi, O.J., Kunta, M., 2022. First Report of Citrus Virus A in Texas Associated with Oak Leaf Patterns in Citrus sinensis. Plant Dis. 106, 2005. https://doi.org/10.1094/PDIS-03-21-0628-PDN

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D.H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D.A., Orengo, C.A., Pandurangan, A.P., Rivoire, C., Sigrist, C.J.A., Sillitoe, I., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Bateman, A., 2023. InterPro in 2022. Nucleic Acids Res. 51, D418–D427. https://doi.org/10.1093/nar/gkac993

Pilotti, M., Faggioli, F., Barba, M., 1995. Characterization of Italian isolates of pear vein yellows virus. Acta Hortic. 148–154. https://doi.org/10.17660/ActaHortic.1995.386.18

Rollin, J., Rong, W., Massart, S., 2023. Cont-ID: detection of sample cross-contamination in viral metagenomic data. BMC Biol. 21, 217. https://doi.org/10.1186/s12915-023-01708-w

Rong, W., Rollin, J., Hanafi, M., Roux, N., Massart, S., 2023. Validation of high throughput sequencing as virus indexing test for Musa germplasm: performance criteria evaluation and contamination monitoring using an alien control. PhytoFrontiers™. https://doi.org/10.1094/PHYTOFR-03-22-0030-FI

Rott, M., Xiang, Y., Boyes, I., Belton, M., Saeed, H., Kesanakurti, P., Hayes, S., Lawrence, T., Birch, C., Rast, H., 2017. Application of Next Generation Sequencing for Diagnostic Testing of Tree Fruit Viruses and Viroids. https://doi.org/10.1094/PDIS-03-17-0306-RE

Rott, M.E., Kesanakurti, P., Berwarth, C., Rast, H., Boyes, I., Phelan, J., Jelkmann, W., 2018. Discovery of Negative-Sense RNA Viruses in Trees Infected with Apple Rubbery Wood Disease by Next-Generation Sequencing. Plant Dis. 102, 1254–1263. https://doi.org/10.1094/PDIS-06-17-0851-RE

Schönegger, D., Moubset, O., Margaria, P., Menzel, W., Winter, S., Roumagnac, P., Marais, A., Candresse, T., 2023. Benchmarking of virome metagenomic analysis approaches using a large, 60+ members, viral synthetic community. J. Virol. 97, e01300-23. https://doi.org/10.1128/jvi.01300-23

Shahzad, A., Ullah, S., Dar, A.A., Sardar, M.F., Mehmood, T., Tufail, M.A., Shakoor, A., Haris, M., 2021. Nexus on climate change: agriculture and possible solution to cope future climate change stresses. Environ. Sci. Pollut. Res. 28, 14211–14232. https://doi.org/10.1007/s11356-021-12649-8

The Galaxy Community, 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 50, W345–W351. https://doi.org/10.1093/nar/gkac247

Trebicki, P., 2020. Climate change and plant virus epidemiology. Virus Res. 286, 198059. https://doi.org/10.1016/j.virusres.2020.198059

Xiao, H., Hao, W., Storoschuk, G., MacDonald, J.L., Sanfaçon, H., 2022. Characterizing the Virome of Apple Orchards Affected by Rapid Decline in the Okanagan and Similkameen Valleys of British Columbia (Canada). Pathogens 11, 1231. https://doi.org/10.3390/pathogens11111231

Yaegashi, H., Yoshikawa, N., Candresse, T., 2011. Apple chlorotic leaf spot virus in Pome Fruits, in: Virus and Virus-Like Diseases of Pome and Stone Fruits, Virology. The American Phytopathological Society, p. 428.

Yanase, H., Koganezawa, H., Fridlund, P.R., 1989. Correlation of pear necrotic spot with pear vein yellows and apple stem pitting, and a flexuous filamentous virus associated with them. ISHS Acta Hortic. 157–158. https://doi.org/10.17660/ActaHortic.1989.235.22

Zhang, H., Zhao, X., Cao, X., Khan, L.U., Zhao, R., Wang, H., Huang, X., 2022. Transmission of Areca Palm Velarivirus 1 by Mealybugs Causes Yellow Leaf Disease in Betel Palm (Areca catechu). Httpsdoiorg101094PHYTO-06-21-0261-R 112, 700–707. https://doi.org/10.1094/PHYTO-06-21-0261-R

Table 1. List of primers used in this study, including primer name, the virus and viral segment they target, their sequence, annealing temperature (Ta) and relevant reference, if any.

Virus	Primer name	Sequence (5' – 3')	Ta	Reference
Apple rubbery wood virus 1 segment S (AWRV-1)	ARWaV-1M479F ARWaV-1M1177R	ATCAATCTCTGTTTTCCCTTATGT TACCATACTTTTGAATCTTTGTGC	52 ºC	Rott et al., 2018
Apple rubbery wood virus 2 segment Sa (ARWV-2)	ARWV2-F1 ARWV2-R1	ATGTTGCATCACAGCTATTGGC ATTGTTCCATGCTGCCACAGAA	60 ºC	Minutolo et al., 2023
Citrus virus A RNA 1 (CiVA)	CiVA_1_2586F CiVA_1_1884R	CTAGGCACAAAGCTTGGTCAGAAG GTCTCCTCTTCATCTGACCTACCT	60 ºC	Designed for this study
Citrus virus A RNA 2 (CiVA)	CiVA_2_1F CiVA_2_285R	ATAACTTTTTTGTTAAAAAGC AATCTTGTTCCTTCACTAT	48 ºC	Bester et al., 2021
Pyrus virus A (PyVA)	PyVA-12722F PyVA-13206R	AGCAGCGAATGAATTGACACCAAA CGCCATCTGAGCCGTTTGATTATT	62ºC	Designed for this study
Apple chlorotic leaf spot virus (ACLSV)	ACLSV-A53-F ACLSV-A52-R	GGCAACCCTGGAACAGA CAGACCCTTATTGAAGTCGAA	56ºC	Candresse et al., 1995

Apple stem pitting virus (ASPV)	ASPF1CP	GGGTGTACTTTGAGGCAGTATT	55ºC	Komorowska et al., 2010
	ASPR3CP	GAGCGGATGCGGTACATCTGTAT		

Table 2. Presence or absence of CiVA and PyVA in the Belgian collection as determined by HTS and RT-PCR, For HTS and according to the threshold described above, pools with a ratio ($RV_{CiVA\ 1\rightarrow n}$, $RV_{PyVA\ 1\rightarrow n}$) below the threshold of 0.5% were considered likely false positives (~), pools with a ratio above the threshold were considered as positive (+), and pools with a ratio of 0% or without any mapped reads to the reference genome were considered as negative (-). For CiVA, detection was considered positive if at least one genomic RNA (RNA 1 or RNA 2) was positive (above the threshold). Abbreviations: Citrus virus A (CiVA), high-throughput sequencing (HTS), reverse transcription polymerase chain reaction (RT-PCR), the new virus Pyrus virus A (PyVA), reads per kilobase per million (RPKM), not analyzed (na). (*) the detection has been carried out after 2 years of storage at -20°C.

Library	Pool	Sample	CiVA		PyVA	
			HTS	RT-PCR	HTS	RT-PCR
			2021	2023/2021*	2021	2023/2021*
L1	L1-1	W20	+	+	-	-
		W18		+/+		-
		W10		+		-
		W13		+		-
	L1-2	W11	+	+	-	-
		W15		+		-
		W16		-		-
		W19		+		-
	L1-3	W2	+	-	-	-
		W4		+/+		-
		W5		-		-

		W7	+/+		-
L2	L2-1	W6	na	-	-
		W21	+/+		-
		W24	na		-
		W12	+/-		-
	L2-2	W14	+/-	-	-
		Y13	-		-
		Y15	+/+		-
		Y16	+/+		-
	L2-3	Y9	-	-	-
		Y11	-		-
		Y20	+/+		-
		Y21	na		-
L3	L3-1	Z1	na	~	na
		Z11	na		na
		Z12	+		-
		Z13	na		na
	L3-2	Z2	-	+	-/-
		Z3	-		-/-
		Z14	+		+/+
		Z15	-		+/+
L4	L4-1	Z4	na	-	na
		Z5	na		na
		Z6	na		na

		Z7		na	na
	L4-2	V14	-	na	- na
		V18		na	na
		V5		na	na
		V15		na	na
L5	L5-1	V6	-	na	- na
		V7		na	na
		V8		na	na
		V9		na	na
	L5-2	V10	-	na	- na
		V11		na	na
		V17		na	na
		V16		na	na
	L5-3	V12	-	na	- na
		V13		na	na
		X5		+/-	-
		X8		na	na
L6	L6-1	X14	-	-	~ -
		X10		-	-
		X12		-	-
		X9		+/+	+/+
	L6-2	X11	-	-	- -
		X20		-	-
		X22		+/+	-

		X23		na		na
L7	L7-1	X13	-	-	~	-
		X16		-		-
		X19		-		-
	L7-2	X18	-	-	+	+/+
		X21		-		-/-

Table 3. Percentages of amino acid (aa) identity of 10 proteins between the new virus (PyVA) and other members of the Velarivirus genus, obtained using a multiple sequence alignment tool (MAFFT). The genomes of cordyline virus 2 (NC_043453), cordyline virus 3 (NC_043107), and cordyline virus 4 (NC_043108) had partial sequences of the ORF 1a, and NC_001836 did not have an annotation for p9. Thus, they were not used for the comparison. A complete comparison of the percentage of identity between the proteins of the accepted velariviruses and PyVA is provided in Supplementary Table 3.

Genbank nº	Virus	1a	1b	p4	HSP70h	p9	HSP90h	CP	CPm	p25	p26
NC_027121	Areca palm velarivirus 1 (APV-1)	22%	56%	26%	40%	13%	23%	23%	13%	15%	11%
NC_038421	Cordyline virus 1 (CoV-1)	22%	55%	36%	43%	14%	28%	30%	14%	15%	12%
NC_043453	Cordyline virus 2 (CoV-2)	-	54%	24%	44%	24%	31%	31%	17%	15%	14%
NC_043107	Cordyline virus 3 (CoV-3)	-	57%	15%	46%	26%	31%	27%	13%	17%	9%
NC_043108	Cordyline virus 4 (CoV-4)	-	56%	16%	44%	25%	26%	30%	15%	21%	12%
NC_016436	Grapevine leafroll-associated virus 7 (GLRaV-7)	23%	55%	18%	42%	20%	28%	26%	11%	14%	10%
NC_001836	Little cherry virus 1 (LChV-1)	30%	60%	21%	48%	-	34%	21%	19%	13%	15%
NC_055599	Malus domestica virus A (MdoVA)	46%	75%	61%	60%	49%	52%	41%	26%	32%	27%

A) $$\mathrm{RPKM} = \frac{\text{number of mapped reads on the virus genome}}{\frac{\text{virus genome length}}{1{,}000} \times \frac{\text{total number of reads for the sample}}{1{,}000{,}000}}$$

B) $$\mathrm{RA}_{\mathrm{a}\ 1\ \rightarrow\ \mathrm{n}} = \frac{\mathrm{RPKM}_{\mathrm{a}\ 1\ \rightarrow\ \mathrm{n}}}{\mathrm{RPKM}_{\mathrm{a\ max}}} \times 100$$

C) $$\mathrm{RV}_{\mathrm{x}\ 1\ \rightarrow\ \mathrm{n}} = \frac{\mathrm{RPKM}_{\mathrm{x}\ 1\ \rightarrow\ \mathrm{n}}}{\mathrm{RPKM}_{\mathrm{x\ max}}} \times 100$$

Figure 1. A) Formula used to calculate the reads per kilobase per million (RPKM) applied for each detected virus in each sample. B) Formula of the cross-contamination ratio of the alien control for each alien virus. "RAa 1→n" is the ratio for each alien virus (PvEV-1, 2, and 3) in the corresponding sample (n samples in total), "RPKMa 1→n" is the RPKM value of the reads mapped to the alien virus reference genome for each of the n samples, and "RPKMa max" is the RPKM mapped to the alien virus reference genome in the alien control library (fixed for each alien virus). C) Formula of the cross-contamination ratio of the analyzed samples for each virus in each sample. "RVx 1→n" is the ratio for each virus (with x corresponding to the name of the virus), "RPKMx 1→n" is the RPKM value of the reads mapped to mapped to each virus reference genome independently in n samples, and "RPKMx max" are the highest RPKM mapped to a virus reference genome (variable for each virus).

147x47mm (300 x 300 DPI)

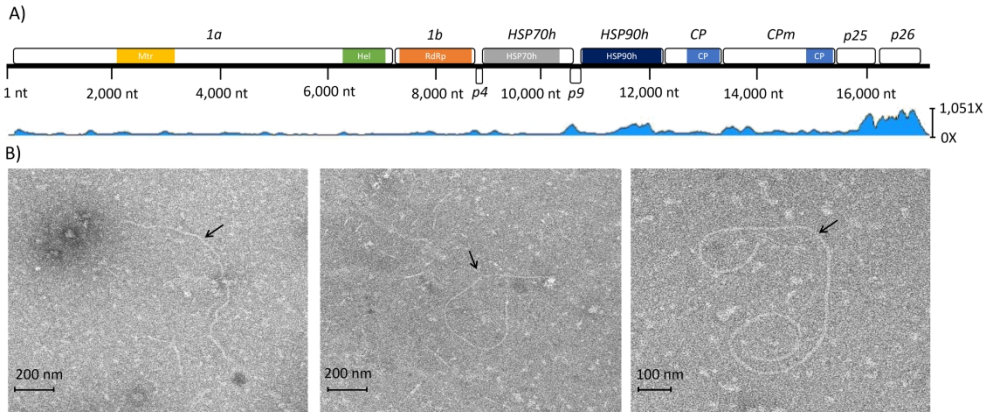

Figure 2. A) Graphic representation of the genomic organization of the new virus (PyVA, isolate 621-BE), with each box representing a predicted open reading frame (ORF) and the protein domains highlighted in different colors. The names in italics represent the products of the ORFs, the polyprotein 1a/1b complex, and the hypothetical proteins p4, p10p9, p27p25, and p28p26. The naming of the protein without an identified protein domain was done following the structure of MDoVA. The coverage of reads mapped to the genome of PyVA is shown in blue below the genome structure (maximum coverage = 1,051X). Abbreviations: methyltransferase (Mtr), helicase (Hel), RNA-dependent RNA polymerase (RdRp), heat-shock protein 70 homolog (HSP70h), heat-shock protein 90 homolog (HSP90h), capsid protein (CP), and minor capsid protein (CPm). B) Electron micrograph of three viral particles of the PyVA, marked with a black arrow. The particles were purified and observed by TEM, following the staining method described in section 2.6.

362x150mm (300 x 300 DPI)

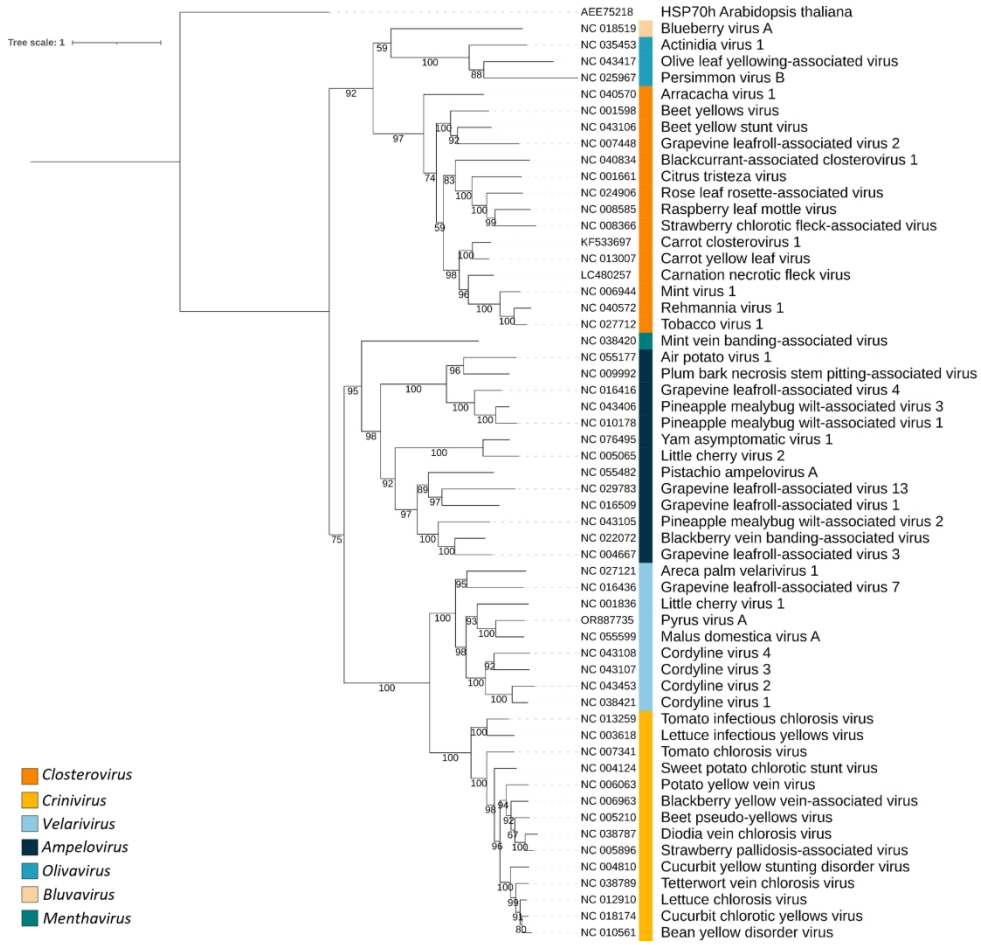

Figure 3. Maximum likelihood phylogenetic tree (rtREV+F+I+G4 substitution model, MAFFT alignment, 1000 bootstraps) based on the alignment of HSP70h amino acid (aa) sequences of members of the Closteroviridae family and of PyVA (highlighted in red). The phylogenetic analysis was performed using the Galaxy server and visualized using the iTOL v6.8 tool. The HSP70 sequence of Arabidopsis thaliana was used as outgroup to root the tree. Bootstrap values are shown for each branch, and colored labels represent the genus that each virus belongs to as shown in the legend on the left.

203x190mm (300 x 300 DPI)