



Citation: G. Parrella, E. Troiano, A. Mignano (2024) Metagenomic analysis reveals the presence of prunus virus I in diseased *Clematis vitalba*: first record of this virus in Italy. *Phytopathologia Mediterranea* 63(1): 3-7. doi: 10.36253/phyto-14925

Accepted: December 22, 2023

Published: February 17, 2024

Copyright: ©2024 G. Parrella, E. Troiano, A. Mignano. This is an open access, peer-reviewed article published by Firenze University Press (<http://www.fupress.com/pm>) and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Editor: Arnaud G Blouin, Institut des sciences en production végétale IPV, DEFR, Agroscope, Nyon, Switzerland.

ORCID:

GP: 0000-0002-0412-4014
ET: 0000-0001-7755-4915
AM: 0000-0001-5365-7186

Short Notes

Metagenomic analysis reveals the presence of prunus virus I in diseased *Clematis vitalba*: first record of this virus in Italy

GIUSEPPE PARRELLA*, ELISA TROIANO, ANNA MIGNANO

Institute for Sustainable Plant Protection of the National Research Council (IPSP-CNR), Piazzale Enrico Fermi 1 – 80055 Portici (NA), Italy

*Corresponding author. giuseppe.parrella@ipspp.cnr.it

Summary. Prunus virus I (PrVI) was detected for the first time in *Clematis vitalba* in Italy using high-throughput sequencing and the complete genome of this isolate, named Clv-1, was assembled and characterized. The results of the bioinformatic analyses were further validated with RT-PCR assays using PrVI-specific primers and Sanger dideoxy sequencing. The Clv-1 genome included three RNA segments of nucleotide lengths of 3468 (RNA1), 2892 (RNA2) and 2225 (RNA3), with five predicted open reading frames. Phylogenetic analyses showed close relationships with other PrVI isolates from different geographical origins, including European and non-European countries. This new pathogen record extends the information on the geographical distribution of PrVI, and possibly reflects the international movement of infected clematis germplasm due to global trade. Further surveys on the presence and distribution of PrVI in weeds and crops, such as the two PrVI hosts sweet cherry and peach, are required in the countries where PrVI has been detected.

Keywords. PrVI, *Bromoviridae*, *Clematis* spp., high-throughput sequencing, emerging viruses.

INTRODUCTION

Clematis is a genus of perennial, chiefly climbing shrubs of *Ranunculaceae*, with 386 accepted species that are widely distributed, mainly in Asia and North America (POWO, 2023). Most *Clematis* spp. are cultivated for their attractive flowers, and more than 3000 cultivars have been obtained through breeding programmes (Mitrofanova *et al.*, 2021). The native European species, *Clematis vitalba* L., is a fast-growing climbing shrub which is often found in the vicinity of vineyards and orchards. It is considered a weed, and it can harbour many phytopathogens, including “flavescence dorée” phytoplasmas, which are often non-pathogenic for this plant (Angelini *et al.*, 2004). Approximately 11 different viruses have been reported to infect clematis plants. These viruses belong to: (i) *Bromoviridae*, including cucumber mosaic virus (CMV), tobacco streak virus (TSV), apple mosaic virus (ApMV) and prunus virus I (PrVI); (ii) *Virgaviridae*, including tobacco rattle virus (TRV); (iii) *Secoviridae*,

including tomato black ring virus (TBRV) and tomato ringspot virus (ToRSV); (iv) *Tombusviridae*, including tomato bushy stunt virus (TBSV), clematis chlorotic mottle virus (CICMV) and Moroccan pepper virus (MPV); and (v) *Tospoviridae*, tomato spotted wilt virus (TSWV) (Mitrofanova *et al.*, 2018; Salamom *et al.*, 2023).

In the winter of 2023, *C. vitalba* plants showing previously undescribed symptomatology (Figure 1), were observed in a public garden in Assisi (PG), Umbria province, Central Italy (N 43.068927; E 12.61576). The presence of PrVI was detected in virome analyses of *C. vitalba* plants, using high-throughput sequencing (HTS) and Sanger sequencing. The whole genome sequence of PrVI was obtained and compared with PrVI isolates previously reported from other countries.

MATERIALS AND METHODS

Plant material, HTS and bioinformatic analysis

Total RNAs were isolated from leaf tissues of the sample Clv-1 (Figure 1) using the Viral Gene-spin™ Viral DNA/RNA Extraction Kit (iNtRON Biotechnology, Inc., Seongnam, South Korea), following the manufacturer's instructions. The RNA purity, concentration and integrity were determined by NanoDrop™ 2000 (NanoDrop). Ribosomal RNAs were depleted using a TruSeq RNA Sample Prep Kit, and the remaining RNAs were used for construction of RNA-seq libraries, which were sequenced on an Illumina Novaseq 6000 platform with paired-end reads length of 150 bp. Quality control on the sequencing data was performed with the software FastQC (v. 0.11.5; <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>), then low quality bases and adapters were removed with the software BBDuk in the BBTools (v. 36; <http://jgi.doe.gov/data-and-tools/software-tools/bbtools>) package setting a minimum read quality of 25 and minimum read length of 35 bp. Taxonomic profiling of the reads were made with the software GAIA (v. 2.02; Paytuví *et al.*, 2019), which compare the reads against the databases Sila 132 (to identify ribosomal sequences) and NCBI nr (to identify viral sequences). The resulting filtered reads were used to assemble the viral genome by using the algorithms Metaspades and RNAViral implemented in SPAdes (v. 3.15.3; Bankevich *et al.*, 2012). BLASTn/BLASTx analysis of the contigs were carried out against local and online databases.

Validation of HTS data

Total RNAs extracted from the sample Clv-1 was used in RT-PCR, using virus-specific primers and sub-

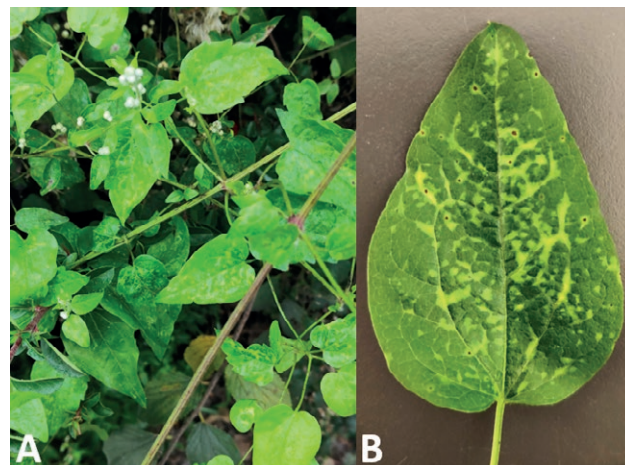


Figure 1. Symptoms in natural *Clematis vitalba* (A), and vein banding and pinpoint necrotic spot symptoms on a *C. vitalba* leaflet (B).

sequent Sanger sequencing of the amplicons to confirm the presence of the virus identified by HTS. Virus-specific primers were designed according to the contig sequences of the target virus obtained in this study (Supplementary Table 1).

The complete viral genomic sequence was obtained from sequence assemblies of the amplicons generated from, respectively, RT-PCR and 5'-3'-RACE reactions (2nd generation 5'/3' RACE kit; Roche). RT-PCR was conducted according to the manufacturer's protocol (ImProm-II™ Reverse Transcription System, Promega). The 25 µL RT-PCR reaction contained 2 µL of total RNAs, 0.5 µL of each primer (50 pmol µL⁻¹), 12.5 µL of 2× Master Mix, 0.5 µL of Enzyme Mix and 9 µL of distilled water. The thermal cycling conditions were: one cycle of reverse transcription at 42°C for 90 min, one cycle of denaturation at 94°C for 2 min, followed by 35 cycles of amplification at 94°C for 45 s, 52°C for 1 min and 72°C for 2 min, and a final cycle of 72°C for 10 min.

Based on the nucleotide sequences that were obtained, specific primers (Supplementary Table 1) were redesigned to amplify fragments of the 5'-upstream and 3'-downstream regions, corresponding to each segmented genome by 5' and 3'-RACE (rapid amplification of cDNA ends), as described by Parrella and Troiano (2022). Amplicons of the expected sizes were directly sequenced in both directions at Microsynth Seqlab GmbH (Göttingen, Germany).

In addition, total RNAs extracted from five other symptomatic *C. vitalba* plants and one asymptomatic *C. vitalba* plant, from the vicinity of Clv-1, were assessed in RT-PCR assays for the presence of the virus identified in Clv-1, using PrV-CP1/PrV-CP2 primers (Supplementary Table 1) following the protocol described above.

Phylogenetic analyses

Multiple sequence alignments were conducted using MUSCLE (Edgar, 2004) implemented in MEGA11. Phylogenetic trees were constructed using the best fit model for each alignment, using the maximum likelihood (ML) method in the MEGA11 (Tamura *et al.*, 2021) with 500 bootstrap replicates. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site.

RESULTS AND DISCUSSION

HTS of the Clv-1 sample yielded 25,409,477 raw reads. Using the results of the taxonomic classification, the trimmed reads were parsed to keep only those classified as “viral” or which were completely unclassified (i.e., 9,246,969 reads). Of the total trimmed reads, 0.068% were mapped on ilarvirus RNA1, 0.055% on ilarvirus RNA2 and 0.095% on ilarvirus RNA3. These reads were used to perform viral genome assembly with the two algorithms implemented in the Spades program. Two sets of filtered contigs were obtained: 4,661 contigs with the algorithms implemented in Metaspades, and 4,920 contigs with algorithms implemented in viral RNA.

BLASTx search showed presence of contigs with the greatest amino acid (aa) sequence identities (99–81%) to PrVI (*Iilarvirus*, *Bromoviridae*). Three contigs showing the greatest nucleotide identities with the PrVI genome were identified by BLASTn. These contigs represented almost full-length genomic sequences of the corresponding viruses. No other viral contigs belonging to other viruses were generated by HTS library. Based on the PrVI mapped contig sequences, primers were designed for each of the three RNAs (Supplementary Table 1). RT-PCR using these primers targeting the three different putative viral RNAs amplified products of the expected sizes (Figure 2). Sequences obtained from these amplicons were identical to the corresponding genome regions sequenced by HTS, confirming presence of PrVI in the *C. vitalba* plant. After verifying the sequences at the 5' and 3' ends, the sequences of the three genomic ssRNA segments of the Clv-1 isolate consisted of 3468 nucleotides (nt) for RNA1, 2892 nt for RNA2, and 2225 nt for RNA3. These sequences were deposited in GenBank with the accession numbers OR502867 (RNA1), OR602868 (RNA2) and OR602869 (RNA3).

RNA1 contains a single ORF (1a) with the ATG codon at position 29 and ending with a TAG codon at position 3301. It encodes the 121 kDa viral replicase protein (p1), consisting of 1090 amino acids (aa). The pro-

tein shares 98.9–99.6% aa identity and 99.0% similarity with other PrVI isolates in GenBank.

RNA2 is bicistronic with the two ORFs overlapping by 272 nt. The first ORF at the 5' end of RNA2 encodes the viral polymerase protein of 807 aa (p2a) with a predicted molecular mass of 92 kDa. The second ORF encodes a 205 aa protein (p2b) involved in cell-to-cell movement and posttranscriptional gene silencing, with a predicted molecular mass of 22 kDa. The p2a protein shares 98.2–98.6% aa identity and 98–99% similarity, while the p2b protein shares 95.6–99.5% identity and 98–99% similarity with other PrVI isolates in GenBank.

RNA3 contains two ORFs. The ORF 3a encodes the movement protein (MP) of 300 aa (p3a), and has an estimated molecular mass of 32 kDa. The p3a protein has 96.0–99.0% identity and 97–100% similarity with the movement proteins of other PVrI isolates. The ORF 3b encodes the coat protein (CP) of 217 aa long protein (p3b) and has a predicted molecular mass of 24 kDa. The p3b protein shares 98.6–99.5% identity and 99–100% similarity with the coat protein of other PrVI isolates.

The expected amplicons were obtained by RT-PCR with the primers Prv-CP1/PrV-CP2, using the RNAs extracted from five other *C. vitalba* plants showing similar symptoms of Clv-1, while no reaction product was obtained with RNA extracted from asymptomatic plants. Sanger sequencing on these amplicons revealed 100% identity among these sequences and with Clv-1 sequence obtained by HTS. These results may indicate a common origin of the infection among the different clematis plants tested, and a possible role in the spread of the virus by vegetatively propagated infected host material.

Maximum likelihood-based phylogenetic analyses of the three RNAs, performed with MEGA 11 under the best fit substitution models, always placed Clv-1 in the

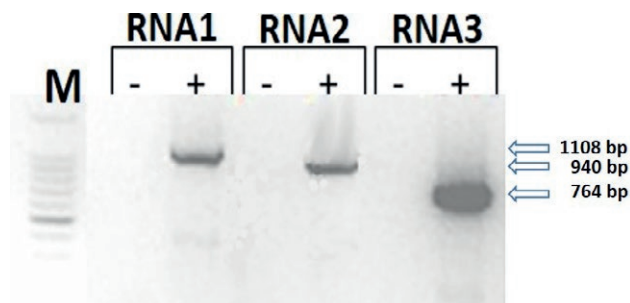


Figure 2. Agarose gel electrophoresis of the RT-PCR products obtained with the primers designed on the PrVI RNA1 (PrV-R1F/PrV-R1R), RNA2 (PrV-R2F/PrV-F2R), and RNA3 (PrV-R3F/PrV-R3R) (Supplementary Table 1). M = 100 bp ladder, - = negative control (healthy plant), + = RNA extracted from the infected *Clematis vitalba* plant.

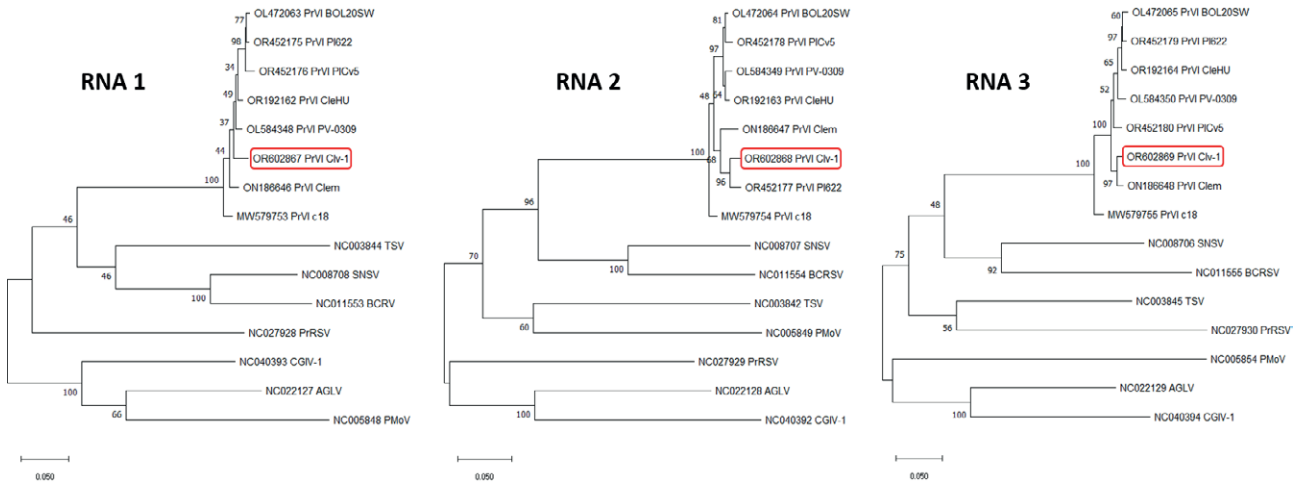


Figure 3. Phylogenetic relationships of the nucleic acid sequences of RNA1, RNA2, and RNA3 of PrVI isolates and ilarviruses belonging to subgroup I. Each label reports the GenBank accession number and the acronym of the virus: prunus virus I (PrVI), tobacco streak virus (TSV), strawberry necrotic shock virus (SNSV), blackberry chlorotic ringspot virus (BCRV), privet ring spot virus (PrRSV), cape gooseberry ilarvirus 1, ageratum latent virus (AGLV), and parietaria mottle virus (PMoV). Isolate names were also reported for PrVI. The position of the Italian isolate of PrVI (Clv-1) is highlighted with red boxes. Phylogenetic analyses were carried out using the maximum likelihood method in MEGA 11, with 500 bootstrap replicates based on the best substitution model found in MEGA 11. Bootstrap values at each node are shown.

same comparable well-supported clade for each RNA, comprising the recognized PrVI isolates within subgroup I in *Ilarvirus*. These results were confirmed by a high bootstrap value in all instances (Figure 3). In particular, for phylogenetic reconstruction based on RNA2, two well supported sub-clades were generated, with the Clv-1 isolate grouping with the two PrVI isolates from *C. vitalba*: isolate Clem identified in Russia and isolate Pl622, identified in Slovakia. The isolate c18 from *Picris echioides* (identified in Slovenia) was the most divergent, and was placed outside the clades of the PrVI isolates in all the three phylogenetic reconstructions, supported by 100% bootstrap values on each corresponding node (Figure 3).

The advances in sequencing metagenomic analysis have led to the discovery of new viruses, allowing the detection and prevention of emerging viruses. Using HTS analysis, PrVI was first described on asymptomatic sweet cherry (*Prunus avium*) in Greece (Orfanidou *et al.*, 2021), but was subsequently identified in the weed plant *P. echioides* (Rivarez *et al.*, 2022) and repeatedly in *Clematis* spp. plants, mainly in Eastern Europe, including Hungary, Slovenia, Slovakia, Croatia and Russia (Chirkov *et al.*, 2022; Rivarez *et al.*, 2022; Salamon *et al.*, 2023).

The recent discovery of PrVI in Italy described in the present study, although associated with different symptoms in *C. vitalba* plants from those described on the same host by previous authors (Chirkov *et al.*, 2022; Salamon *et al.*, 2023), confirms that this plant is an important natural host for this virus. Since PrVI was

only recently described, it is not clear what is its diffusion in nature or its impacts on cultivated plants. In addition, accidental dispersion of infected propagative material of *Clematis* cannot be excluded. This could explain the recent findings of PrVI on *Clematis* in different European countries in a short period. For these reasons, the development of an effective, sensitive and specific PrVI diagnostic method is required, which can be used in extensive monitoring in the areas where PrVI was recently described. Wild and cultivated plants, particularly stone fruit crops, should be assessed for PrVI infections.

ACKNOWLEDGEMENTS

This research was supported by a MISE CRESO grant (Viabio No. F/200095/01-03/X45). The Authors thank Dr Ettore Magaldi and Dr Rosaria Fusco (IPSP-CNR) for their administrative support.

LITERATURE CITED

- Angelini E., Squizzato F., Lucchetta G., Borgo M., 2004. Detection of a phytoplasma associated with Grapevine Flavescence dorée in *Clematis vitalba*. *European Journal of Plant Pathology* 110: 193–201.
- Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., ... Pevzner P.A., 2012. SPAdes: a new genome

- assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19(5): 455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Chirkov S., Zakubanskiy A., Sheveleva A., Zubkova N., Mitrofanova I., 2022. Detection and molecular characterization of viruses infecting clematis in Russia. *Journal of Plant Pathology* 105: 173–183. <https://doi.org/10.1007/s42161-022-01242-8>
- Edgar R. C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research* 32(5): 1792–1797.
- Mitrofanova I.V., Zakubansky A.V., Mitrofanova O.V., 2018. Viruses infecting main ornamental plants: an overview. *Ornamental Horticulturae* 24: 95–102.
- Mitrofanova I., Ivanova N., Kuzmina T., Mitrofanova O., Zubkova N., 2021. In vitro regeneration of *Clematis* plants in the Nikita Botanical Garden via somatic embryogenesis and organogenesis. *Frontiers in Plant Science* 12: 541171. <https://doi.org/10.3389/fpls.2021.541171>
- Orfanidou C.G., Xing F., Zhou J., Li S., Katis N., Malogka V.I., 2021. Identification and sequence analysis of a novel ilarvirus infecting sweet cherry. *Plants* 10: 514. <https://doi.org/10.3390/plants10030514>
- Parrella G., Troiano E., 2022. A New Iilarvirus Found in French Hydrangea. *Plants* 11(7): 944. <https://doi.org/10.3390/plants11070944>
- Paytuví A., Battista E., Scippacercola F., Aiese Cigliano R., Sanseverino W., 2019. GAIA: an integrated metagenomics suite. *bioRxiv* 804690. <https://doi.org/10.1101/804690>
- POWO (2023). Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://www.plantsoftheworldonline.org/> Retrieved 27 November 2023.
- Rivarez M. P. S., Pecman A., Bačnik K., Ferreira O. M. C., Vučurović A., ... Kutnjak D., 2022. In-depth study of tomato and weed viromes reveals undiscovered plant virus diversity in an agroecosystem. *bioRxiv* 2022.06.30.498278. <https://doi.org/10.1101/2022.06.30.498278>
- Salamon P., Nagyne-Galbacs Z., Demian E., Achs A., Alaxin P., ... Varallyay E., 2023. *Clematis vitalba* Is a Natural Host of the Novel Iilarvirus, Prunus Virus I. *Viruses*, 15: 1964. <https://doi.org/10.3390/v15091964>
- Tamura K., Stecher G., Kumar S., 2021. MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution* 38: 3022–3027.