Title

Screening of some Croatian autochthonous grapevine varieties reveals a multitude of viruses, including novel ones

Journal

Archives of Virology, 163(8)

ISSN

0304-8608

Authors

Vončina, Darko
Almeida, Rodrigo PP

Publication Date

2018-08-01

DOI

10.1007/s00705-018-3850-6

Peer reviewed

Archives of Virology (2018) 163:2239–2243
https://doi.org/10.1007/s00705-018-3850-6

BRIEF REPORT

Screening of some Croatian autochthonous grapevine varieties reveals a multitude of viruses, including novel ones

Darko Vončina[1] · **Rodrigo P. P. Almeida**[2]

Received: 31 October 2017 / Accepted: 6 April 2018 / Published online: 21 April 2018

Abstract

Next-generation sequencing of total RNA samples from four Croatian autochthonous grapevine varieties revealed the presence of a novel virus in two grapevine accessions. The complete genome sequence of a novel virus, tentatively named "grapevine badnavirus 1" (GBV-1), was reconstructed from a *de novo*-assembled contig. GBV-1 has a genome of 7,145 nucleotides containing three ORFs with sequence similarity to other badnaviruses. In addition, several other viruses and viroids, including grapevine virus G, grapevine virus K/D, grapevine virus T, grapevine Roditis leaf discoloration-associated virus, grapevine yellow speckle viroids 1 and 2, and hop stunt viroid were detected and identified for the first time in Croatian grapevines in the course of this study.

Introduction

Grapevine is a widely grown fruit crop that is susceptible to a large number of pathogens. Since the discovery of grapevine fanleaf virus as the first grapevine-infecting virus over 50 years ago, more than 70 additional viruses have now been identified in this plant host. Next-generation sequencing (NGS), high-throughput sequencing, or deep sequencing, has become an important diagnostic tool for viruses. Although usually used to detect known plant viruses (e.g., read mapping to libraries), the technology may also be used to identify and detect novel ones. NGS technology was first applied in the discovery of grapevine viruses in 2009 [1], and since then, many other grapevine viruses have been identified using this approach. Another advantage is that NGS may provide insights into the overall sanitary status of a particular vine, which is especially helpful when plants harbor multiple viruses of economic importance. In the case of grapevines, subjecting one sample to NGS is cheaper, faster, and more reliable than using a panel of other pathogen detection approaches [2].

In this study, we selected four grapevine samples from the Croatian region of Kaštela (Central Dalmatia), which has a long viticultural tradition and numerous autochthonous grapevine varieties. In addition, previous studies have suggested that several varieties from this region had mixed infections with economically important viruses [3]. We report on the detection of several grapevine viruses and viroids in these four samples, including some not previously reported from Croatia, as well as a virus belonging to a putative new species.

Materials and methods

Four grapevine accessions maintained in a virus collection in Zagreb, each originating from the Kaštela region, were used in this study: VB-108 (variety Babica), VD-102 (var. Dobričić), VLJ-178 (var. Ljutun) and VVL-101 (var. Vlaška). For NGS analysis, total RNA was isolated in August 2015 from 0.1 g of leaf petioles using an RNeasy Plant Mini Kit (QIAGEN, Germany). The manufacturer's recommended protocol was modified by preheating the RLT buffer to 55 °C and dissolving 2.5% PVP-40 followed by filter sterilization. Host ribosomal RNA (rRNA) was

Handling Editor: Sead Sabanadzovic.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-018-3850-6) contains supplementary material, which is available to authorized users.

✉ Darko Vončina
dvoncina@agr.hr

[1] Department of Plant Pathology, University of Zagreb Faculty of Agriculture, 10 000 Zagreb, Croatia

[2] Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA

depleted using an Illumina Ribo-Zero rRNA Removal Kit (Plant) (Illumina, USA). Sequencing libraries were constructed at the Functional Genomics Lab at QB3-Berkeley Core Research Facility (UC Berkeley) on an Apollo 324™ with PrepX™ RNAseq Library Prep Kits (WaferGen Biosystems, USA). Thirteen cycles of PCR amplification were used for index addition and library fragment enrichment. Genomic data, 100 paired-end reads, were generated on a HiSeq4000 platform (Illumina, USA) at Vincent J. Coates Genomics Sequencing Laboratory (UC Berkeley). The FastQC program (Babraham Bioinformatics, UK) was used for initial data quality control, and reads were trimmed and filtered with Sickle using default parameters [4]. Trimmed reads were aligned to a grapevine genome sequence (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/) using Bowtie 2 ver. 2.2.9 [5], and unaligned (host genome filtered) reads were extracted, sorted using SAMtools ver. 1.3.1 [6], and converted to FASTQ file format using BEDTools ver. 2.26.0 [7] for analysis.

Unaligned reads were used for *de novo* assembly with SPAdes ver. 3.9.0 [8]; parameters used included k-mer values 21, 33, 55 and 75, and the careful option for reduction of mismatches and short indels. All assembled contigs were subjected to a remote BLAST nucleotide search for viruses and viroids with the cutoff e-value set at e-20 (query terms: "viruses NOT plants" and "viroids NOT plants"). Additionally, contigs larger than 1,000 nucleotides with no virus hits were manually checked using the BLAST nucleotide search tool. Contigs with no significant hits in the database were aligned to each other using MEGA [9]. A search for conserved domains and ORFs was done using NCBI's Conserved Domain Database and ORF Finder tool. For the novel badnavirus and vitiviruses detected for the first time in Croatia, phylogenetic trees were generated using codon-aligned nucleotide sequences of the RT/RNase H region and replicase/coat protein genes, respectively.

In addition to the *de novo* assembly, trimmed reads were subjected to i) mapping to a custom-made library of grapevine-associated viruses, mycoviruses and viroids (Supplementary Table 1) using Bowtie 2 ii) search by automated e-probe-based software Truffle [10], and iii) VirFind [11] with the default e-value (0.01) and the conserved domains search option. To confirm NGS-based results for the putative novel virus, together with viruses and viroids not previously reported in Croatia, plants were resampled in September 2017. Total RNA/DNA was isolated using an RNeasy/DNeasy Plant Mini Kit (QIAGEN, Germany) following the manufacturer's instructions. Based on sequences generated by NGS, specific primers were designed (Supplementary Table 2), and RT-PCR or PCR products were verified by Sanger sequencing at Macrogen (Korea).

Results

Once NGS reads were trimmed, 77,000,000 to 113,000,000 reads per sample were available for analysis, of which 70-87.5% aligned to the grapevine host genome. *De novo* assembly of data from grapevine accessions VB-108, VD-102, VLJ-178, and VVL-101 generated 436, 467, 459, and 251 contigs > 1,000 nt, respectively. Between 60 and 85 contigs from each sample resulted in hits to virus sequences deposited in the GenBank database. One contig of accession VLJ-178 of 7,220 nt in length and 931.0× coverage, and another from VVL-101 of 1,391 nt and 1.9× coverage had no hits using search parameters. Both contigs were identical in the overlapping region, representing a novel grapevine-associated virus, tentatively named "grapevine badnavirus 1" (GBV-1, GenBank accession. no. MF781082). Further analysis confirmed a circular genome of 7,145 nt and three ORFs: ORF1 encoding a 33.4-kDa hypothetical protein; ORF2, encoding a polypeptide of 1,846 amino acids with similarity to reverse transcriptase, ribonuclease H, cauliflower mosaic virus peptidase (A3), and a zinc-binding motif; and ORF3, encoding 35.6-kDa hypothetical protein (Fig. 1). The RT/RNase H region of the novel badnavirus showed highest nt sequence identity of 76% to grapevine Roditis leaf discoloration-associated virus (GRLDaV)

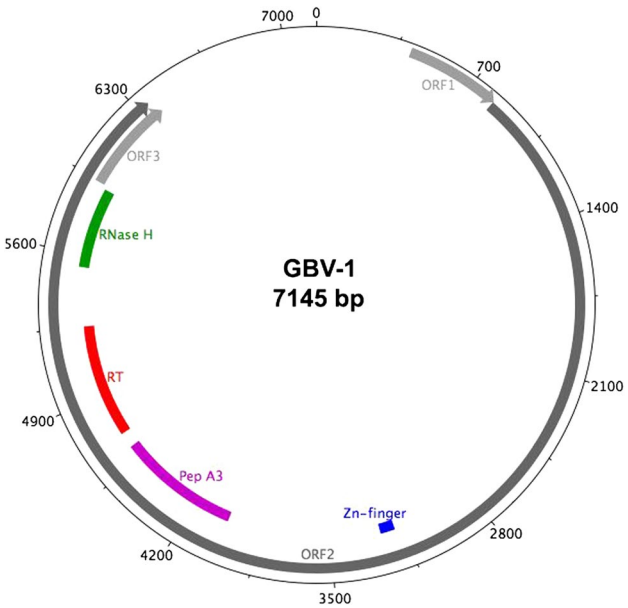

Fig. 1 Genome organization of grapevine badnavirus 1 (GBV-1, MF781082). Viral DNA is shown by a thick black circular arrow with the position of three open reading frames (ORF1-ORF3, grey), and conserved functional motifs (color) in the ORF2 polyprotein. Abbreviations: RNase H, ribonuclease H; RT, reverse transcriptase; Pep A3, cauliflower mosaic virus peptidase (A3); Zn-finger, Zinc knuckle. A visual genome summary was generated with DNAPlotter

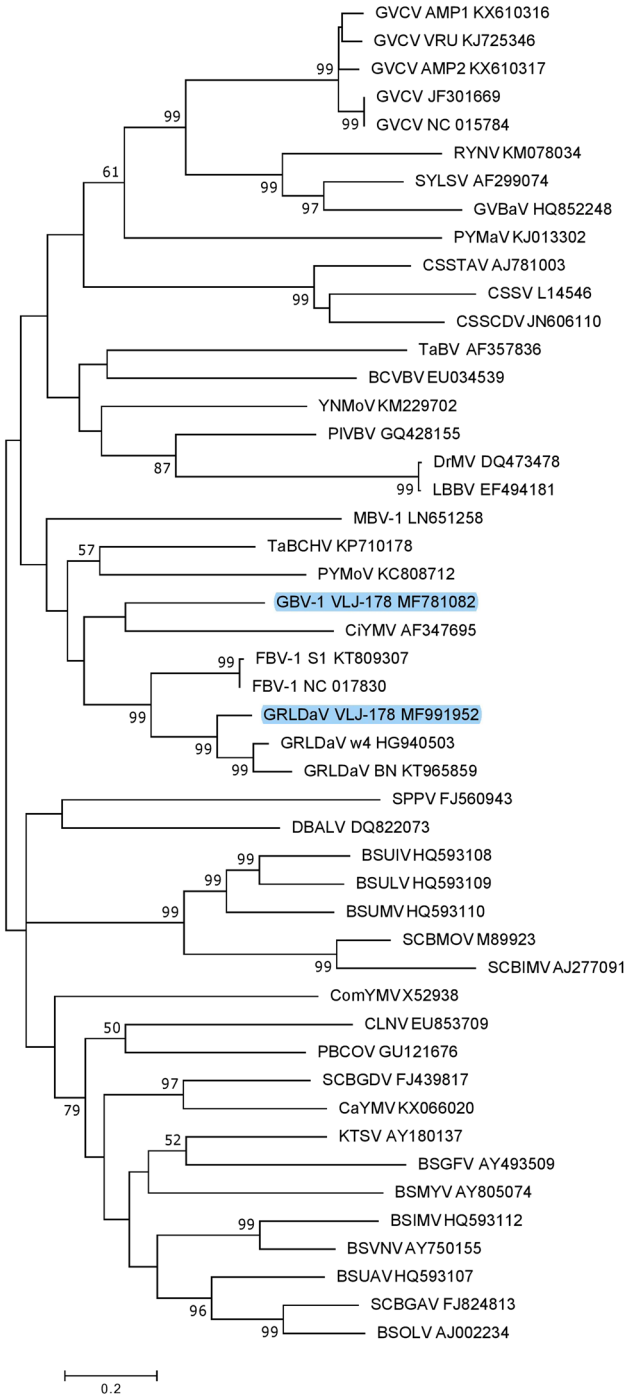

Fig. 2 Maximum-likelihood tree based on codon-aligned nucleotide sequences of the RT/RNase H region from members of the genus *Badnavirus*, with their respective GenBank accession numbers. Bootstrap values for 1000 replicates are shown on branches when there is > 50% support. A novel badnavirus, tentatively named "grapevine badnavirus 1" (GBV-1) and a Croatian isolate of grapevine Roditis leaf discoloration-associated virus (GRLDaV) are highlighted in blue. Abbreviations: BSGFV, banana streak GF virus; BSMYV, banana streak MY virus; BSIMV, banana streak IM virus; BSOLV, banana streak OL virus; BSUAV, banana streak UA virus; BSUIV, banana streak UI virus; BSULV; banana streak UL virus; BSUMV, banana streak UM virus; BSVNV, banana streak VN virus; BCVBV; bougainvillea chlorotic vein banding virus; CSSV, cacao swollen shoot virus; CSSCDV, cacao swollen shoot CD virus; CSSTAV, cacao swollen shoot Togo A virus; CaYMV, canna yellow mottle virus; CiYMV, citrus yellow mosaic virus; ComYMV, commelina yellow mottle virus; CLNV, cycad leaf necrosis virus; DBALV, dioscorea bacilliform AL virus; DrMV, dracaena mottle virus; FBV-1, fig badnavirus 1; GVBAV, gooseberry vein banding associated virus; GVCV, grapevine vein clearing virus; KTSV, kalanchoë top-spotting virus; LBBV, lucky bamboo bacilliform virus; MBV-1, mulberry badnavirus 1; PYMaV, pagoda yellow mosaic associated virus; PlVBV, pelargonium vein banding virus; PBCOV, pineapple bacilliform CO virus; PYMoV. piper yellow mottle virus; RYNV, rubus yellow net virus; SYLSV, spiraea yellow leafspot virus; SCBGAV, sugarcane bacilliform Guadeloupe A virus; SCBGDV, sugarcane bacilliform Guadeloupe D virus; SCBIMV - sugarcane bacilliform IM virus; SCBMOV, sugarcane bacilliform MO virus; SPPV, sweet potato pakakuy virus; TaBV, taro bacilliform virus; TaBCHV, taro bacilliform CH virus; YNMoV, yacon necrotic mottle virus

isolate w4 (HG940503) and several isolates of fig badnavirus 1 (FBV-1) (L9, KT809305; S1, KT809307; H1, KT809304), with 96% and 98% query coverage, respectively. Because the RT/RNase H nt sequence was less than 80% identical to those of other badnaviruses [12], GBV-1 should be considered a prototype of a new species in the genus. This was also confirmed by phylogenetic analysis, which positioned the novel virus in a separate clade (Fig. 2).

The survey also revealed several other viruses that have only recently been reported in grapevine. In all grapevine accessions, contigs of 7031 to 7552 nt in length and coverage between 4.1× to 7136× (MF781081, MF993573, MF993574, and MF993575) were variants of grapevine virus G (GVG) [13]. One contig from vine VLJ-178 (MF993576) of 7,428 nt and 280× coverage had 90% query coverage and 76% nt sequence identity to a sequence of grapevine virus K (GVK) [14] and 9% query coverage with 87% nt sequence identity with grapevine virus D (GVD) ORF1 and ORF2 (Y07764). Two other contigs from vine VLJ-178 (MG001925) and three from vine VB-108 (MG001926) showed the highest nucleotide sequence similarity to grapevine virus T (GVT) [15]. These contigs from vines VLJ-178 and VB-108, with no overlapping fragments, showed total query coverage of 87% and 88% with nucleotide sequence identity of 86% and 79%, respectively, to GVT isolate Cho (NC_035203).

In addition to the viruses discussed above, complete or almost complete genomes were retrieved from single contigs for several other viruses: grapevine leafroll-associated virus 3 (MF991951), grapevine virus A (MF979533), B (MF991949) and E (MF991950), grapevine rupestris stem pitting-associated virus (MF979534), and grapevine Roditis leaf discoloration-associated virus (GRLDaV, MF991952). In addition, complete genome sequences of grapevine yellow speckle viroid 1 (GYSVd-1, MF979527-29), 2 (GYSVd-2, MF979530) and hop stunt viroid (HSVd, MF979531-32) were also detected as a part of assembled contigs.

A detailed review of NGS data analysis by mapping with Bowtie 2, *de novo* assembly using SPAdes, VirFind and Truffle, and comparison with RT-qPCR, one-step RT-PCR, and PCR as well as ELISA results is given in Supplementary Table 3.

Discussion

The data generated by NGS technologies allow the use of different approaches to detect and study plant viruses. Despite the limited number of samples tested, a new virus was identified along with three recently reported ones (GVG, GVK/GVD, GVT). Additionally, three viroids and one virus (GRLDaV) detected in this study are reported for the first time from Croatian grapevines. Besides GBV-1, grapevine vein clearing virus (GVCV) and GRLDaV are, so far, the only known badnaviruses infecting grapevine [16, 17]. Unlike GRLDaV, GBV-1 has motifs of cauliflower mosaic virus peptidase (A3); it also lacks the arginine methyltransferase-interacting protein present in GVCV. Sequences related to recently reported GVG, GVK and GVT genomes found in this survey suggest that this is not a case of their random or sporadic occurrence. Additionally, nt sequence alignment of the replicase gene from Croatian GVG isolates showed between 95% and 99% identity, but only 73% identity when compared to isolates from New Zealand [13], suggesting large sequence divergence. The coat protein nt sequence of the Croatian GVK isolate was over 80% identical to the GVD isolates (KX828708, KY689027, JQ031715, JQ031716 and Y07764), suggesting that GVK and GVD could be different variants of the same virus (Supplementary Fig. 1). The high GVG and GBV-1 genome coverage as well as depth of coverage in accession VLJ-178 compared to the other accessions could be a consequence of sample cross-contamination or indexing issues. Additional verification tests with new nucleic acid extractions confirmed the presence of GVG in all grapevine accessions, and GBV-1 in VLJ-178 and VVL-101, suggesting that coverage differences may be related to other issues, including virus population variability among plants.

To the best of our knowledge, this study represents the first report of GRLDaV in Croatia. The contig of 7,157 nt and 856.9× coverage represents an incomplete genome, with highest query coverage of 98% and identity of 86% to strain BN (KT965859) from Italy. The presence of viroids infecting grapevine in Croatia was demonstrated previously using biological tests without identification at the species level [18]. This survey provides molecular evidence of three grapevine viroids (GYSVd-1 and 2, HSVd) being present in Croatia. All four tested grapevines showed a range of virus-like symptoms, ranging from short internodes and stunting to abnormal branching and the presence of zig-zag growth. Due to the complex sanitary status of each grapevine accession (i.e., at least 11 viruses and several viroids), the observed symptoms cannot be assigned to any particular virus or viroid.

NGS approaches applied in this study were successful in detecting more viruses and viroids than the RT-qPCR test. In addition to viruses that were not previously reported in Croatia, including GVB-1, several viruses previously known to infect these plants (such as GLRaV-3, GFkV, GFLV, etc.) were all confirmed by NGS data. The results of this study confirm the suitability of NGS technology for virus discovery purposes. However, it is important to stress the fact that biological information about most grapevine viruses is very limited or non-existent, and the ability to detect a large number of putative pathogens is of limited applied value if there are no studies describing their respective roles or how to reduce their impact on a crop of economic importance. Therefore, our next goal is to understand the impact, if any, of GBV-1 and several other viruses on grapevine production.

Depositories

Raw sequence data from different grapevine accessions are deposited in NCBI's BioProject database (Project ID PRJNA415169): VB-108, accession no. SAMN07816752; VD102, accession no. SAMN07816753; VLJ-178, accession no. SAMN07816754; VVL-101, accession no. SAMN07816755.

Acknowledgements The research leading to these results received funding from the European Union Seventh Framework Programme (FP7 2007-2013) under Grant agreement no. 291823 Marie Curie FP7-PEOPLE-2011-COFUND (The New International Fellowship Mobility Programme for Experienced Researchers in Croatia - NEWFELPRO). This article was prepared as a part of a project "Ecology of an emerging grapevine virus in Croatia and California (EcoGVCC)" which received funding through NEWFELPRO under Grant agreement no. 51. The work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. Special thanks to Dijana Škorić and Cecilia Prator for critical review of the manuscript, and Mladen Poletti Kopešić for technical support.

Compliance with ethical standards

Conflict of interest The authors (DV and RA) declare that they have no conflict of interest.

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

References

1. Al Rwahnih M, Daubert S, Golino DA, Rowhani A (2009) Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that included a novel virus. Virology 387:395–401

2. Al Rwahnih M, Daubert S, Golino D, Islas C, Rowhani A (2015) Comparison of next-generation sequencing versus biological indexing for the optimal detection of viral pathogens in grapevine. Phytopathology 105(6):758–763
3. Vončina D, Đermić E, Cvjetković B, Maletić E, Pejić I, Karoglan Kontić J (2009) Occurrence of the important grapevine viruses in population of Croatian autochthonous grapevine varieties from Kaštela region. In: 32nd World Congress of vine and wine, 7th general assembly of the OIV and Wine Kubanović, Veronika (ed.). Zagreb: Ministarstvo poljoprivrede, ribarstva i ruralnog razvoja:199–200
4. Joshi NA, Fass JN (2011) Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. https://github.com/najoshi/sickle
5. Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359
6. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing Subgroup (2009) The sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25:2078–2079
7. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842
8. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA (2013) Assembling genomes and mini-metagenomes from highly chimeric reads. Res Comp Mol Biol 7821:158–170
9. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874
10. Visser M, Burger JT, Maree HJ (2016) Targeted virus detection in next generation sequencing data using an automated e-probe based approach. Virology 495:122–128
11. Ho T, Tzanetakis IE (2014) Development of a virus detection and discovery pipeline using next generation sequencing. Virology 471–473:54–60
12. Geering ADW, Hull R (2012) Family Caulimoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus taxonomy. 9th report of the international committee on taxonomy of viruses. Elsevier Academic Press, London, pp 429–443
13. Blouin AG, Keenan S, Napier KR, Barrero RA, MacDiarmid RM (2017) Identification of a novel vitivirus from grapevines in New Zealand. Arch Virol. https://doi.org/10.1007/s00705-017-3581-0
14. Jo Y, Song M-K, Choi H, Park J-S, Lee J-W, Lian S, Lee BC, Cho WK (2017) Genome sequence of grapevine virus K, a novel vitivirus infecting grapevine. Genome Announc 5:e00994-17
15. Jo Y, Song M-K, Choi H, Park J-S, Lee J-W, Lian S, Lee BC, Cho WK (2017) Genome sequence of grapevine virus T, a novel foveavirus infecting grapevine. Genome Announc 5:e00995-17. https://doi.org/10.1128/genomeA.00995-17
16. Zhang Y, Singh K, Kaur R, Qiu W (2011) Association of a novel DNA virus with the grapevine vein clearing and vine decline syndrome. Phytopathology 101:1081–1090
17. Maliogka VI, Olmos A, Pappi PG, Lotos L, Efthimiou K, Grammatikaki G, Candresse T, Katis NI, Avgelis AD (2015) A novel grapevine badnavirus is associated with the Roditis leaf discoloration disease. Virus Res 203:47–55
18. Šarić A, Krajačić M, Škorić D (1996) Viroidi kao patogeni poljoprivrednih kultura u Hrvatskoj. 40. seminar iz zaštite bilja. Opatija, Glasnik Zaštite Bilja 1:6–7