




Grapevine virus L: a novel vitivirus in grapevine

Humberto Debat  · Diego Zavallo · Reid Soltero Brisbane · Darko Vončina · Rodrigo P. P. Almeida · Arnaud G. Blouin · Maher Al Rwahnih · Sebastian Gomez-Talquenca · Sebastian Asurmendi

Accepted: 25 March 2019
© Koninklijke Nederlandse Planteziektenkundige Vereniging 2019

Abstract *Vitiviruses* are ssRNA(+) viruses in the family *Betaflexiviridae* (subfamily *Trivirinae*). There are currently 10 ICTV recognized virus species in the genus; nevertheless, the extended use of NGS technologies is rapidly expanding their diversity and official recognition of six more have been proposed recently. Here, we

present the characterization of a novel virus from grapevine, which fits the genomic architecture and evolutionary constraints to be classified within the *Vitivirus* genus. The detected virus sequence is 7607 nt long, including a typical genome organization of ORFs encoding a replicase (RP), a 22 kDa protein, a

Humberto Debat and Diego Zavallo contributed equally to this work.

Accession numbers Grapevine virus L sequences have been deposited in GenBank under accession numbers: MH248020 (GVL-RI), MH643739 (GVL-KA), MH681991 (GVL-VL), MH686191 (GVL-SB)

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10658-019-01727-w>) contains supplementary material, which is available to authorized users.

H. Debat (✉)
Instituto de Patología Vegetal, Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria (IPAVE-CIAP-INTA), X5020ICA Córdoba, Argentina
e-mail: debat.humberto@inta.gob.ar

D. Zavallo · S. Asurmendi
Instituto de Biotecnología, Centro de Investigación en Ciencias Veterinarias y Agronómicas (IB-CICVyA-INTA), 1686 Buenos Aires, Argentina

R. S. Brisbane
Foundation Plant Services, Davis, CA 95616, USA

D. Vončina
Department of Plant Pathology, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia

R. P. P. Almeida
Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA

A. G. Blouin
The New Zealand Institute for Plant & Food Research Limited, Private Bag 92169, Auckland 1142, New Zealand

M. Al Rwahnih
Department of Plant Pathology, University of California, Davis, CA, USA

S. Gomez-Talquenca (✉)
Estación Experimental Agropecuaria Mendoza, Instituto Nacional de Tecnología Agropecuaria (EEA-Mendoza-INTA), Luján de Cuyo, 5534 Mendoza, Argentina
e-mail: gomez.talquenca@inta.gob.ar

S. Asurmendi
CONICET, Buenos Aires, Argentina

movement protein, a coat protein (CP) and a nucleic acid binding protein. Phylogenetic analyses based on the predicted RP and CP proteins unequivocally place the new virus within the *Vitivirus* genus. Multiple independent RNAseq data confirmed the presence of the detected virus in berries at diverse developmental stages. Additionally, we detected, confirmed, and assembled virus sequences from grapevine samples of distinct cultivars from America, Europe, Asia and Oceania, sharing 74.4%–97.8% nt identity, suggesting that the identified virus is widely distributed and diverse. We propose the name grapevine virus L (GVL) to the detected *Vitivirus*.

Keywords *Vitivirus* · Grapevine · Virus discovery · *Betaflexiviridae*

Introduction

Vitiviruses have flexuous, non-enveloped, filamentous virus particles of 725–785 nm and 12 nm in length and diameter, respectively, with a nucleocapsid that is cross-banded and diagonally striated. Vitiviruses have a linear ssRNA(+) genome (~7.3–7.6 kb), with a methylated nucleotide cap at the 5' end and a 3' poly (A) tail (Du Preez et al. 2011; Adams et al. 2012). There are 10 species of vitiviruses recognized by the International Committee on Taxonomy of Viruses (ICTV), nevertheless, six new species have been proposed recently, with members of five of them infecting grapevine (*Vitis vinifera*) (Jo et al. 2017; Blouin et al. 2018a, b; Candresse et al. 2018; Diaz-Lara et al. 2018). Grapevine is the most prevalent natural host of vitiviruses, but they have also been reported from several crops such as mint (*Mentha x glaciensis*), arracacha (*Arracacia xanthorrhiza*) blue agave (*Agave tequilana*), kiwi (*Actinidia chinensis*) and blackberry (*Rubus* spp) (Tzanetakis et al. 2007; Blouin et al. 2012; Oliveira et al. 2017; Hassan et al. 2018). Vitiviruses appear to be latent in *V. vinifera* cultivars, and so far, only grapevine virus A and grapevine virus B have been consistently associated to grapevine diseases of the rugose wood complex (grapevine vitiviruses) or Shiraz disease (reviewed by Minafra et al. 2017). The etiological role of the recently described vitiviruses should be assessed in order to establish its relationships with known or unknown viral diseases. In addition, the synergistic effects between vitiviruses and other grapevine viruses appears to be significant (Rowhani et al.

2018). The availability of complete sequences of these viruses could allow the development of full-length infectious clones, fulfillment of the Koch's postulates, and improve our understanding of the biological role of these viruses (Goszczyński 2015). Here, we present the characterization of a novel vitivirus for which we tentatively propose as grapevine virus L (GVL). We were able to detect and characterize GVL in several distinctive *V. vinifera* cultivars from four continents.

Material and methods

Virus discovery and annotation

Virus discovery, confirmation, and annotation were implemented as described in (Al Rwahnih et al. 2015; Al Rwahnih et al. 2018; Blouin et al. 2018b; Diaz-Lara et al. 2018). RT-PCR assays were deployed as reported in Al Rwahnih et al. (2014). Briefly, raw RNA data from several Sequence Read Archive (SRA) accessions was downloaded from the National Center for Biotechnology Information database (NCBI). These files were processed in the following form: trimmed and filtered with the Trimmomatic tool, as implemented in <http://www.usadellab.org/cms/?page=trimmomatic>, and the resulting reads of each library were assembled de novo with Trinity v2.6.6 release with standard parameters (Haas et al. 2013). The obtained transcripts were subjected to bulk local blastx searches (E-value <1e⁻⁵) against a refseq virus protein database available at <ftp://ftp.ncbi.nlm.nih.gov/refseq/release/viral/viral.1.protein.faa.gz>. The resulting hits were explored manually and curated by iterative mapping using Bowtie2 available at <http://bowtie-bio.sourceforge.net/bowtie2/>. ORFs were predicted by ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) and annotated with the NCBI conserved domain search tool as implemented in <https://www.ncbi.nlm.nih.gov/Structure/cdd>.

Phylogenetic analyses

Phylogenetic insights were generated by multiple amino acid alignments of replicase proteins (BLOSUM62 scoring matrix) using as best-fit algorithm E-INS-i, which drives local alignment with generalized affine gap costs, and thus is applicable for RNA polymerases where diverse domains are dispersed among several

highly divergent regions, by MAFFT v7.394 as implemented in <https://mafft.cbrc.jp/alignment/software/>. Additionally, as determined by the auto-select parameter of MAFFT, coat protein alignments used the G-INS-i algorithm. Genomic nt pairwise identities were determined using MUSCLE with 8 iterations and Neighbor Joining clustering method as implemented in Geneious v8.1.9 (Biomatters, inc). The aligned proteins were subsequently used as input for FastTree 2.1.5 <http://www.microbesonline.org/fasttree/> maximum likelihood phylogenetic trees (best-fit model = JTT-Jones-Taylor-Thorton with single rate of evolution for each site = CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1000 resamples. SNPs were detected by the FreeBayes v0.9.18S tool with standard parameters implemented in <https://github.com/ekg/freebayes>.

RT-PCR assays

Total nucleic acid (TNA) were freshly extracted from leaf petioles tissues using the MagMax 96 Viral RNA isolation kit with the MagMax Express-96 magnetic particle processor (Thermo Fisher Scientific, USA). The final total RNA fraction was eluted in 100 µl of water. RT-PCR reactions mixes included 1.25 µl of each primer (10 µM), 5 µl Green GoTaq reaction buffer (Promega, USA), 1.1 µl MgCl₂ (25 µM), 0.5 µl dNTPs mix (10 µM), 12.39 µl water, 1.25 µl DTT (0.1 M), 0.125 µl GoTaq DNA polymerase (Promega, USA), 0.035 µl SuperScript II Reverse Transcriptase (Life Technologies, USA), and 2 µl TNA in a total reaction volume of 25 µl. The RT-PCR conditions were 35 cycles at 94 °C for 30 s, 56 °C for 45 s, and 72 °C for 1 min, with a final cycle of 7 min at 72 °C.

Results and discussion

Evidence of a novel virus in grapevine

In order to explore the potential grapevine viral landscape on High Throughput Sequencing (HTS)-generated publicly available libraries, we downloaded raw RNA data from several SRA accessions on the NCBI database, finding known virus in the vast majority of tested samples. The raw data were assembled de novo and the obtained transcripts were subjected to bulk local blastx searches (E-value <1e⁻⁵) against a refseq virus

protein database. The resulting hits were explored manually. One dataset, specifically NCBI Bioproject PRJNA400621, SRX3144921-SRX3144956, composed of RNA libraries from grape berry samples of *V. vinifera* cv. Riesling and cv. Cabernet sauvignon from Beijing, China (Chen et al. 2017), presented a 6155 nt long transcript from the assembled transcriptome of the SRX3144956 library. This sequence obtained a highly significant hit (E-value = 0; sequence identity 69%) with the replicase protein of grapevine virus E (GVE, isolate TvAQ7, YP_002117775.1; Nakaune et al. 2008). Iterative read mapping allowed to extend and polish the assembled contig into a 7607 nt long sequence, supported by 25,112 reads and a mean coverage of 162X, sharing the genome organization of vitiviruses (Fig. 1a; Table 1). The assembled virus sequence has a pairwise identity with GVE TvAQ7 (reference sequence) of 66.5% at the genomic RNA level. The predicted structure of the detected virus is formed by a 67 nt 5'UTR, followed by five ORFs and a 66 nt 3'UTR, excluding the poly (A) tail of unknown length. Both predicted UTRs are highly similar in length and size in relation to other vitiviruses, more precisely GVE and the proposed GVI and GVG (Supp. Figure 1). ORF1 (67–5158 nt coordinates) encodes a replicase protein (RP) of 1696 aa (192.1 kDa), sharing a 70.1% identity with GVE TvAQ7 replicase. We employed the NCBI conserved domain search tool to annotate the RP, which presents the characteristic domain architecture of vitiviruses. The 5' region of RP presents a viral methyltransferase domain (E-value = 1.76e-56; Pfam = pfam01660; 46–337 aa coordinates), involved in mRNA capping, followed by a DEXDc type viral RNA helicase (E-value = 1.47e-12; Pfam = pfam01443; 750–1047 aa coordinates). Interestingly, an Alkylated DNA repair dioxygenase domain (AlkB), of the 2OG-Fe(II) oxygenase superfamily, was identified within the helicase domain (E-value = 2.34e-06; COG = COG3145; 917–1032 aa coordinates) as previously described for an American isolate of GVE, GVG and GVI (Alabi et al. 2013; Blouin et al. 2018a, b). Finally, an RNA dependent RNA polymerase domain was found at the 3' region of the RP (RdRP_2; E-value = 3.47e-27; Pfam = pfam00978; 1313–1637 aa coordinates).

Although vitiviruses utilize a series of sub-genomic RNAs (sgRNA) as general strategy for expression of ORF2–5, it is worth mentioning that ORF2 (5155–5766 nt) of the detected virus overlaps in the tetranucleotide AUGA with ORF1. This overlapping cue

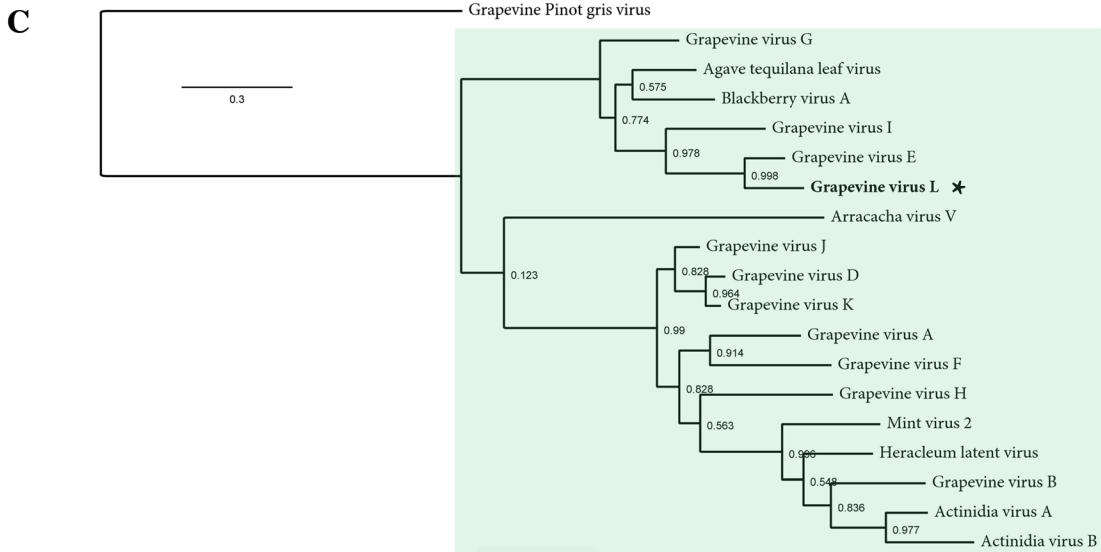
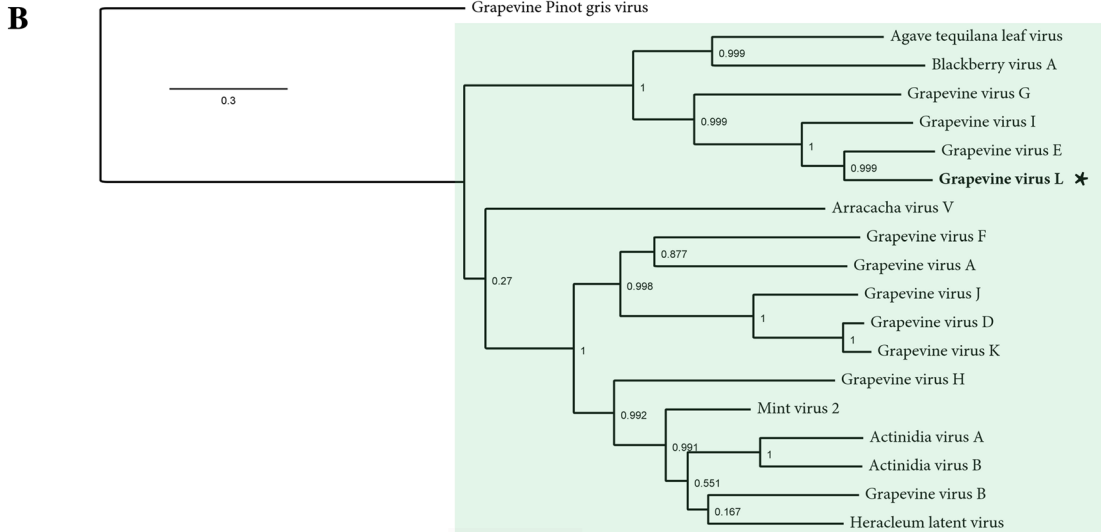
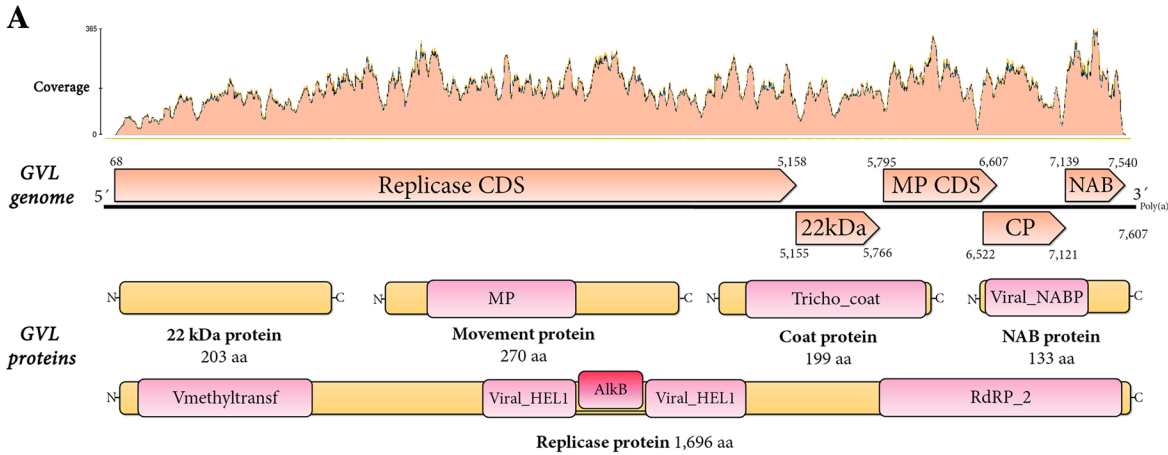


Fig. 1 **a** Molecular characterization of grapevine virus L (GVL). Mapping pattern of virus RNA reads from library SRX3144956 to the consensus putative GVL assembly indicating coverage per position. The genome architecture of GVL is characterized by a single ssRNA(+) sequence encoding for five ORFs arranged as 5'-UTR-RP-HP-MP-CP-NABP-UTR-3'. The predicted product of each ORF is depicted including the associated domains determined by NCBI-CDD. Abbreviations: RP, replicase protein; 22 kDa, 22 k Dalton hypothetical protein; MP, movement protein; CP, coat protein; NABP, nucleic acid binding protein. See the text for additional domain information. **b** Phylogenetic insights of putative grapevine virus L (GVL) in relation to accepted and proposed members of the genus *Vitivirus* based on MAFFT alignments (BLOSUM62 scoring matrix using as best-fit algorithm E-INS-i) of the predicted replicase and capsid protein (**c**) followed by maximum likelihood trees generated by FastTree (best-fit model=JTT-Jones-Taylor-Thorton with single rate of evolution for each site=CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1000 resamples). Accession numbers of the corresponding sequences: Actinidia virus A (JN427014), Actinidia virus B (NC_016404), mint virus 2 (AY913795), grapevine virus B (NC_003602), grapevine virus H (MF521889), grapevine virus F (NC_018458), grapevine virus A (NC_003604), grapevine virus K (NC_035202), arracacha virus V (NC_034264), grapevine virus E (NC_011106), *Agave tequilana* leaf virus (NC_034833), grapevine virus D (KX828708), grapevine virus G (MF405923), grapevine virus I (MF927925), heracleum latent virus (X79270), and grapevine Pinot gris virus (*Trichovirus*, NC_015782) used as outgroup. Scale bar represents substitutions per site. Node labels are FastTree support values. Asterisks indicate position of GVL

might probably function as a termination–reinitiation strategy signal, which was determined for caliciviruses and victoriviruses for coupled start-stop translation (Li et al. 2011). This putative AUGA signal between ORF1-ORF2, is present in the genome sequence of the recently proposed grapevine virus J (Diaz-Lara et al. 2018). The predicted ORF2 encodes a leucine rich (15.3% L) 22.8 kDa protein, of unknown function, which is the most divergent protein putatively encoded by the identified virus (44% identity and 98% coverage to the ORF2 encoded hypothetical protein of GVE, AHB08905). ORF3 (5795–6607 nt) encodes a putative movement protein (MP) of 270 aa (29.7 kDa) with a viral movement protein domain (E-value = 6.61e-07; Pfam = pfam01107) between the 26–195 aa coordinates. ORF4 (6522–7121 nt), partially overlapping with ORF3, encodes a coat protein (CP) 199 aa long (21.4 kDa) presenting a Trichovirus coat protein-like domain (Tricho_coat; E-value = 4.26e-45; Pfam = pfam05892; 37–199 aa coordinates). The predicted CP of the identified virus shares 77.49% identity with GVE-CP. Lastly, ORF5 (7139–7540 nt) encodes a 133

aa putative nucleic acid binding protein (15 kDa) with a nucleic acid binding domain (Viral_NABP; E-value = 1.40e-06; Pfam = pfam05515) between the 5–76 aa coordinates. This region shares significant identity with the CTV_P23 domain (E-value = 3.79e-06) of citrus tristeza virus (*Closteroviridae*) P23 protein which roles involve asymmetrical RNA accumulation, symptoms expression, suppression of RNA silencing as well as an increased viral movement (reviewed in Flores et al. 2013).

Sequence-based analyses supports GVL virus as a member of a new vitivirus species

The criteria for species demarcation in the family *Betaflexiviridae*, proposed by ICTV, is less than ca. 72% overall nucleotide identity and below 80% amino acid identity of CP or RP predicted proteins (Adams et al. 2012). Given that the detected virus shares 66.5% identity with GVE at the genomic RNA level, and a 70.1% and 78.4% identity with the respective RP and CP proteins (Supp. Figure 2), we tentatively suggest that the virus sequence represents a genome of a virus that could be a member of a new species, for which we propose the name grapevine virus L (GVL). In order to entertain this hypothesis, based on genetic criteria, we gathered evolutionary insights of the putative GVL sequence identified in cv. Riesling grapevines from China (GVL-RI) in the context of the *Betaflexiviridae* family. We generated a phylogenetic tree based on RP proteins corresponding to the 93 refseq *Betaflexiviridae* RP sequences available at NCBI. Unequivocally, the putative GVL RP clustered within the *Vitivirus* genus (Supp. Figure 3). To provide local tree topology and additional phylogenetic insights we generated maximum likelihood phylogenetic trees of both the RP and CP protein of accepted and proposed members of the *Vitivirus* genus. In both cases, putative GVL clustered within one of the two major sub-clades, made up of GVE, GVG and GVI, and more distantly, the *Agave tequilana* leaf virus (Fig. 1b-c). These results support the tentative assignment of the detected virus to the *Vitivirus* genus.

A preliminary incidence of GVL in cv Riesling datasets from China

In order to investigate the incidence of GVL in grapevine, we performed an in silico survey in publicly-available

Table 1 Size of the genome and genome products of grapevine virus L (GVL) compared to recognized or putative vitiviruses

Virus	Abbreviation	Accession #	Genome	5'UTR	RP	HP	MP	CP	NABP	3'UTR
Actinidia virus A	AcVA	JN427014	7566	NA	1714	219	297	198	105	137
Actinidia virus B	AcVB	NC_016404	7488	70	1707	231	290	198	106	83
<i>Agave tequilana</i> leaf virus	ATLV	NC_034833	6958	52	1539	151	267	199	126	86
arracacha virus V	AVV	NC_034264	7398	51	1705	NA	293	192	110	73
blackberry virus A	BVA	MG254193	7285	59	1676	148	272	199	109	18
grapevine virus A	GVA	NC_003604	7351	86	1707	177	278	198	90	68
grapevine virus B	GVB	NC_003602	7599	48	1707	179	322	197	123	148
grapevine virus D	GVD	MF774336	7479	79	1698	161	273	197	91	87
grapevine virus E	GVE	NC_011106	7564	57	1698	191	265	199	108	134
grapevine virus F	GVF	NC_018458	7551	90	1727	159	271	198	109	82
grapevine virus G	GVG	MF405923	7496	64	1703	154	286	201	117	53
grapevine virus H	GVH	MF521889	7446	97	1717	162	265	197	106	82
grapevine virus I	GVI	NC_037058	7507	68	1696	167	264	199	121	104
grapevine virus J	GVJ	MG637048	7390	95	1702	164	273	197	103	69
grapevine virus K	GVK	NC_035202	7476	80	1700	161	273	197	91	77
heracleum latent virus	HLV	X79270	3006	NA	253	NA	343	197	115	56
mint virus 2	MV-2	AY913795	3897	NA	500	209	329	209	111	85
grapevine virus L	GVL	MH248020	7607	67	1696	203	270	199	133	67
Vitivirus \bar{x}	NA	NA	7438	72	1692	177	286	198	108	89

Abbreviations: Accession #: NCBI Genbank accession number; Genome: Genome size (nt), 5'U: 5'UTR length (nt); RP: replicase protein length (aa); HP: hypothetical protein length (aa); MP: movement protein length (aa); CP: coat protein length (aa); NABP: nucleic acid binding protein length (aa); 3'UTR: 3'UTR length excluding the poly A tail (nt). \bar{x} indicates mean average values calculated within the genus

HTS libraries. We mapped the 1,014,833,524 raw reads available at the PRJNA400621 project (Chen et al. 2017) to the assembled GVL sequence, using as threshold two nt mismatches and a 22 nt seed and a cut-off value >10 reads hits per library with the Bowtie2 tool. None of the 18 RNA libraries of cv. Cabernet Sauvignon had the putative GVL. Interestingly, we were able to detect the virus in six additional cv Riesling libraries of the same study, obtained from independent biological samples conformed by grape berries sampled in three distinct developmental stages (E-L 35, E-L 36 and E-L 38) (Supp. Table 1). Moreover, while analyzing the intrinsic variability of the detected GVL sequences, we noticed that two of the samples presented 24 highly supported SNPs at frequencies >0.5 as detected by the FreeBayes v0.9.18S tool (Supp. Table 2). Given the low *p*-values of the variants and that 70% of the detected SNPs were silent, we speculate that these SNPs are not artifactual. In addition, in order to explore the distribution of GVL, we gathered data from grapevine samples from Croatia (Vončina and Almeida 2018), USA, and New Zealand.

GVL in Croatian cv Vlaska grapevines

A collection of HTS libraries obtained from petioles RNA samples of autochthonous Croatian grapevines (cv Babica, cv Vlaska, cv Dobricic and cv Ljutun; Bioproject PRJNA415169) had virus RNA reads similar to GVL. In turn, de novo assembly and curation of the putative GVL transcript resulted in a 7584 nt virus genome (GVL-VL, mean coverage of 675X), sharing the genomic organization and a 74.7% nt identity with GVL-RI (Supp. Table 3). The presence of this virus was subsequently confirmed by targeted RT-PCR (see below) of the HTS sequenced samples. Two PCR products were bidirectionally Sanger-sequenced and the obtained sequences shared a 100% identity with GVL-VL.

GVL in Katelin vine of Canadian origin

The sample from USA corresponded to a quarantine selection Katelin (KA), which was received in 2012

for inclusion in the Foundation Plant Services (FPS, UC-Davis, CA) collection. The vine was grown in a greenhouse and assayed for known grapevine viruses as described previously (Al Rwahnih et al. 2018). In addition, total RNA from the positive sample was high-throughput sequenced and the obtained ca. 11.5 million raw reads were filtered and de novo assembled with parameters described in Al Rwahnih et al. (2018). A contig with high similarity with GVL was identified by BLAST and subsequently refined, resulting in a 7591 nt virus draft genome (GVL-KA, MH643739), sharing the genomic architecture and a 97.8% nt identity with GVL-RI (Supp. Table 3).

GVL in cv sauvignon blanc from New Zealand

Lastly, the presence of GVL was confirmed in an HTS library derived from double-stranded RNA isolated by immunocapture as described previously (Blouin et al. 2016) from cv. Sauvignon Blanc, which was originally collected in April 2016 in the region of Marlborough, South Island, New Zealand. The virus was detected in only one of the 225 samples surveyed by HTS. The presence of the virus was confirmed by RT-PCR and Sanger sequencing. A draft genome with a partial truncation at 5' was assembled, encompassing 7365 nt, lacking both the 5'UTR and a ca. 184 nt coding region of the replicase (GVL-SB, mean coverage of 166X). The obtained sequence shared a 74.4% nt identity with GVL-RI. All draft genome sequences were annotated both structurally and functionally, showing consistent signatures of vitiviruses and a shared genome organization (Supp. Figure 4). In turn, sequence identity values were calculated indicating high sequence similarity in the predicted coat protein (92.5% to 99.5% aa identity) and significant divergence in the 22 kDa predicted protein (60.6% to 96.6% aa identity) (Supp. Table 3).

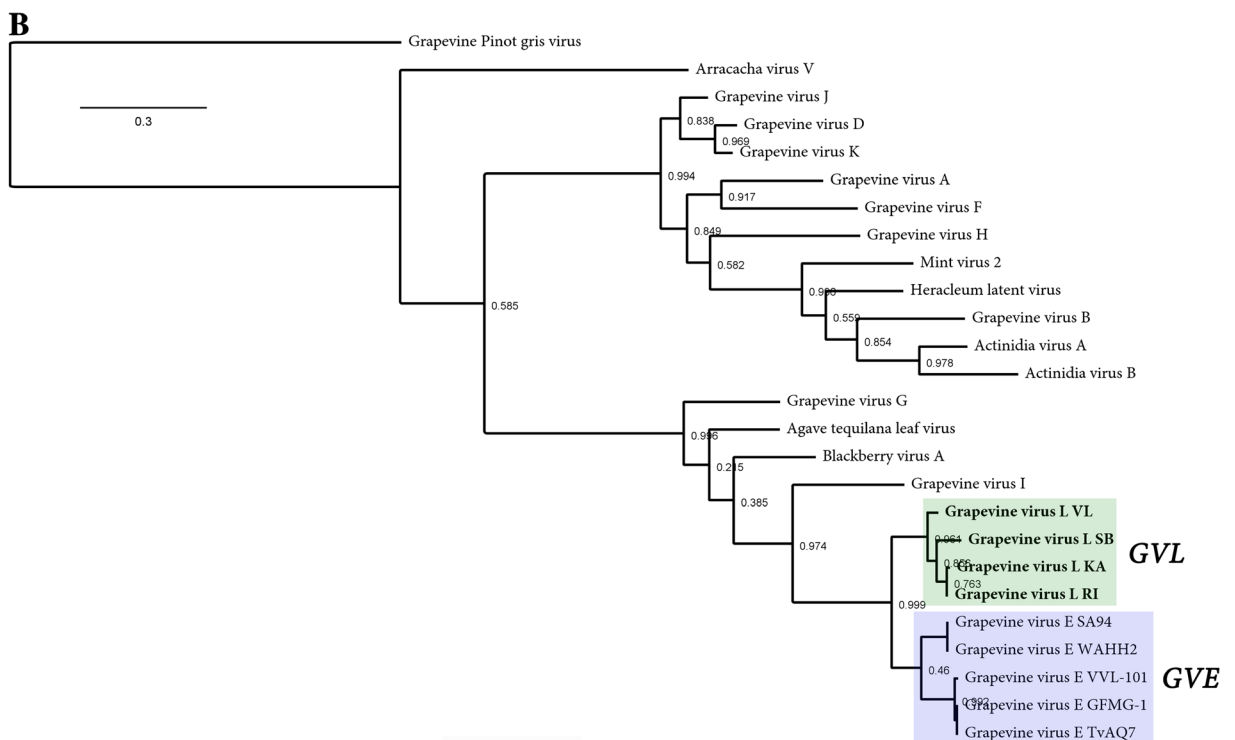
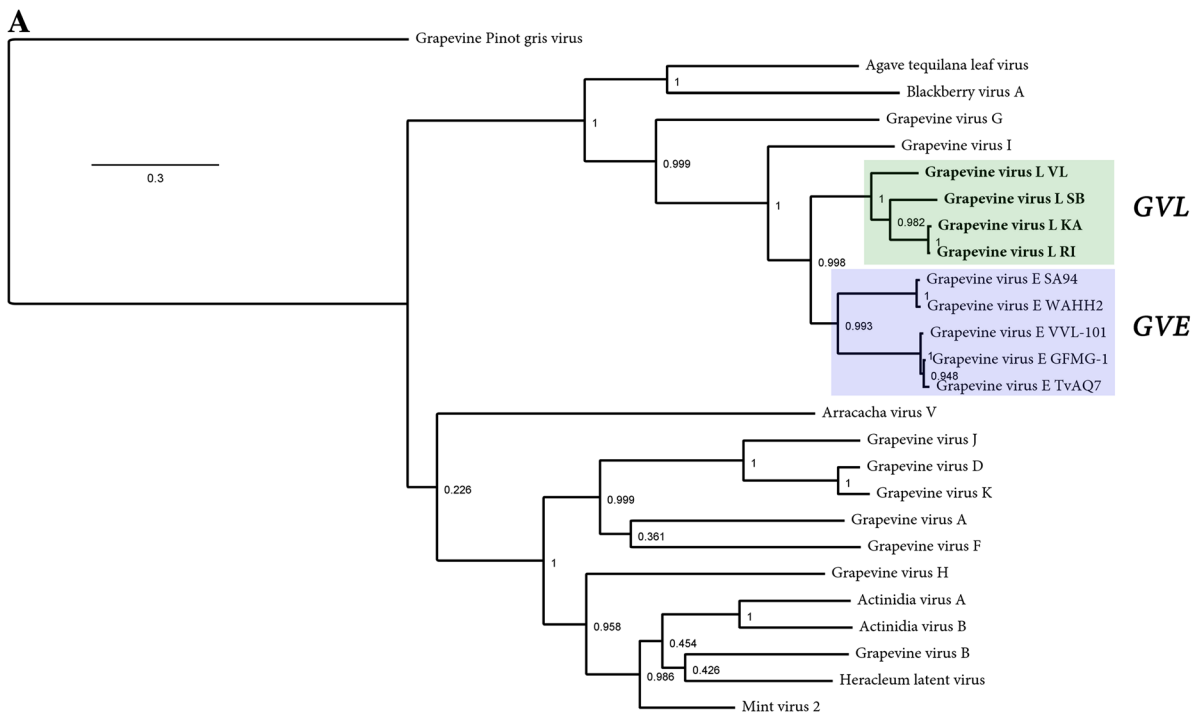
Implementing an RT-PCR assay to detect GVL

Further, an RT-PCR assay was deployed as reported in Al Rwahnih et al. (2014) to investigate the prevalence of GVL in grapevine samples from USA, using forward [AGTTGAAGTCTAGGTGCACAC] and reverse [GTACTCAGACTTCCCCGATCTA] primers, designed to target a conserved region encompassing the 3' end of the CP ORF and part of the NAB encoding sequence of GVL (Supp. Figure 5). The specificity of this primer pair in the detection of GVL was confirmed

by RT-PCR using TNA from vines infected with GVA, GVB, GVD, GVE, GVH, GVG, GVJ. Using the above primers in those assays, none of the grapevines that were infected with the non-target vitiviruses were found to be positive, whereas a sample of grapevine cv. Katelin tested positive. The development of this GVL specific RT-PCR assay will allow for the detection of GVL in field tests, and in clean-stock programs, facilitating a more effective control of this virus. Nevertheless, it is worth mentioning that additional GVL sequences obtained by non-targeting methods (e.g. HTS) are needed in order to confirm if the designed primers would be able to maintain its target in the context of the natural variability of the virus.

Assessing the evolutionary history of GVL

Given the significant intraspecific variability observed among the diverse GVL identified isolates, in the context of an overall interspecies affinity of GVL and GVE, we proceeded with additional genetic distance assessment and phylogenetic analyses. We integrated the genomic sequences and predicted products of four assembled GVL sequences and the five available, tentatively complete, genome sequences of GVE in multiple nucleotide and amino acid alignments (isolates TvAQ7, of *V. labruscana* from Japan; WAHH2, of *V. vinifera* cv Cabernet sauvignon from USA; SA94, of *V. vinifera* cv Shiraz from South Africa; VVL-101, *V. vinifera* cv. Vlaska from Croatia and GFMG-1 from China). We determined genetic distances of the five GVE isolates and the four GVL virus sequences at the level of genome (nt) and CP and RP (aa) (Supp. Figure 6), which are used as a species demarcation criteria in genus *Vitivirus*. Pairwise identities of the aforementioned genomic regions supported our predictions that the virus sequences presented in this study correspond to a new species. Further, maximum likelihood phylogenetic trees based both on RP and CP alignments mirrored the preceding tree topology presented in Fig. 1b and supported clade branching between GVL and GVE in the context of clustering within vitiviruses (Fig. 2a-b). We went one step forward and retrieved all GVE partial sequences available at GenBank. Besides the five GVE isolates mentioned before, additional sequences corresponded to the CP encoding ORF and a short region of the NAB ORF. Thus, we inspected the predicted products of the complete/nearly complete CPs of GVE (46 sequences) and generated



phylogenetic studies based on multiple amino acid alignments and maximum likelihood trees. The resulting trees mimicked our previous predictions,

unequivocally clustering all the GVL sequences in a separate clade, and GVE sequences divided into two sister groups (designated here as α and β) (Supp.

◀ **Fig. 2** Phylogenetic analyses showing relationships of four isolates of grapevine virus L (GVL) and five isolates of grapevine virus E based on MAFFT alignments (BLOSUM62 scoring matrix using as best-fit algorithm E-INS-i) of the predicted replicase (a) and capsid protein (G-INS-i best-fit algorithm) (b) followed by maximum likelihood trees generated by FastTree (best-fit model = JTT-Jones-Taylor-Thorton with single rate of evolution for each site = CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1000 resamples). Accession numbers of the corresponding sequences: grapevine virus E (GVE) isolate VVL-101(MF991950), GVE isolate WAHH2 (JX402759), GVE isolate SA94 (GU903012), GVE isolate GFMG-1 (KF588015), GVE isolate TvAQ7 (NC_011106). Additional sequences are indicated in Fig. 1b legend. Pinot gris virus was used as outgroup. Scale bar represents substitutions per site. Node labels are FastTree support values. GVL isolates are depicted in bold

Figure 7). GVE sequences were grouped into two well supported clusters, future studies should inquire whether the observed genetic differences are reflected in biological disparities on their hosts.

Final remarks

The availability of more RNA sequences corresponding to GVL would provide more insights into the evolutionary history and phylodynamics of GVL. The simultaneous detection of GVL in the Americas, Europe, Asia and Oceania in diverse *V. vinifera* cultivars suggest that this novel virus is widely distributed but showed a low incidence in all of the vineyards surveyed. The emerging biological properties of GVL remain elusive. Given that the samples harboring GVL were co-infected with several grapevine viruses and viroids, including grapevine rupestris stem pitting-associated virus (*Foveavirus*), grapevine Syrah virus-1 (*Marafivirus*), grapevine leafroll-associated virus 3 (*Ampelovirus*), grapevine yellow speckle viroid 1 & 2 (*Apscaviroid*), and hop stunt viroid (*Hostuviroid*), future studies should unravel whether GVL presence is associated to any specific symptoms in grapevine.

Acknowledgements This work was supported by project PEI131022 of the Instituto Nacional de Tecnología Agropecuaria (INTA) and by ANPCyT PICT 2015-1532 and PICT 2016-0429. The funders had no role in study design and analysis, decision to publish, or preparation of the manuscript. The research from New Zealand is funded by New Zealand Winegrowers Inc. and The Ministry of Business, Innovation & Employment. Many thanks to Alfredo Diaz Lara, Department of Plant Pathology, University of California, Davis, CA for his technical assistance.

Author contributions HD, DZ, SGT and SA designed the study. RSB, DV, RPPA, AGB and MAR, performed experiments, generated and analyzed data. HD integrated the data together with SA. HD and DZ wrote the initial draft of the manuscript and SGT, AGB, RSB, DV, RPPA, MAR and SA revised the manuscript. All authors approved the final version.

Compliance with ethical standards

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

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors. Therefore, informed consent was not required for this work.

References

- Adams, M. J., Candresse, T., Hammond, J., Kreuze, J. F., Martelli, G. P., Namba, S., et al. (2012). Family *Betaflexiviridae*. In A. M. Q. King, M. J. Adams, E. B. Carstens, & E. J. Lefkowitz (Eds.), *Virus taxonomy – Ninth report on the international committee on taxonomy of viruses* (pp. 920–941). London: Elsevier Academic Press.
- Al Rwahnih, M., Daubert, S., Islas, C., Golino, D., & Rowhani, A. (2014). Characterization of a fifth vitivirus in grapevine. *Journal of Plant Pathology*, *96*, 219–222.
- 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*, 758–763.
- Al Rwahnih, M., Rowhani, A., Westrick, N., Stevens, K., Diaz-Lara, A., Trouillas, F. P., et al. (2018). Discovery of viruses and virus-like pathogens in pistachio using high-throughput sequencing. *Plant Disease*, *102*, 1419–1425.
- Alabi, O. J., Poojari, S., Sarver, K., Martin, R. R., & Naidu, R. A. (2013). Complete genome sequence analysis of an American isolate of grapevine virus E. *Virus Genes*, *46*, 563–566.
- Blouin, A. G., Chavan, R. R., Pearson, M. N., MacDiarmid, R. M., & Cohen, D. (2012). Detection and characterisation of two novel vitiviruses infecting Actinidia. *Archives of Virology*, *157*, 713–722.
- Blouin, A. G., Ross, H. A., Hobson-Peters, J., O'Brien, C. A., Warren, B., & MacDiarmid, R. (2016). A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies. *Molecular Ecology Resources*, *16*, 1255–1263.
- Blouin, A. G., Chooi, K. M., Warren, B., Napier, K. R., Barrero, R. A., & MacDiarmid, R. M. (2018a). Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand. *Archives of Virology*, *163*, 1371–1374.
- Blouin, A. G., Keenan, S., Napier, K. R., Barrero, R. A., & MacDiarmid, R. M. (2018b). Identification of a novel vitivirus from grapevines in New Zealand. *Archives of Virology*, *163*, 281–284.

- Candresse, T., Theil, S., Faure, C., & Marais, A. (2018). Determination of the complete genomic sequence of grapevine virus H, a novel vitivirus infecting grapevine. *Archives of Virology*, *163*, 277–280.
- Chen, W. K., Bai, X. J., Cao, M. M., Cheng, G., Cao, X. J., Guo, R. R., et al. (2017). Dissecting the variations of ripening progression and flavonoid metabolism in grape berries grown under double cropping system. *Frontiers in Plant Science*, *8*, 1912.
- Diaz-Lara, A., Golino, D., & Al Rwahnih, M. (2018). Genomic characterization of grapevine virus J, a novel virus identified in grapevine. *Archives of Virology*, *163*, 1965–1967.
- Du Preez, J., Stephan, D., Mawassi, M., & Burger, J. T. (2011). The grapevine-infecting vitiviruses, with particular reference to grapevine virus a. *Archives of Virology*, *156*, 1495–1503.
- Flores, R., Ruiz-Ruiz, S., & Soler, N. (2013). Citrus tristeza virus p23: A unique protein mediating key virus–host interactions. *Frontiers in Microbiology*, *4*, 98.
- Goszczynski, D. E. (2015). Brief report of the construction of infectious DNA clones of south African genetic variants of grapevine virus a and grapevine virus B. *SpringerPlus*, *4*, 739.
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. *Nature Protocols*, *8*, 1494–1512.
- Hassan, M., Shahid, M. S., & Tzanetakis, I. E. (2018). Molecular characterization and detection of a novel vitivirus infecting blackberry. *Archives of Virology*, *163*, 2889–2893.
- Jo, Y., Song, M. K., Choi, H., Park, J. S., Lee, J. W., Lian, S., et al. (2017). Genome sequence of grapevine virus K, a novel vitivirus infecting grapevine. *Genome Announcements*, *5*, e00994–e00917.
- Li, H., Havens, W. M., Nibert, M. L., & Ghabrial, S. A. (2011). RNA sequence determinants of a coupled termination-reinitiation strategy for downstream open reading frame translation in *Helminthosporium victoriae* virus 190S and other victoriviruses (family *Totiviridae*). *Journal of Virology*, *85*, 7343–7352.
- Minafra, A., Mawassi, M., Goszczynski, D., & Saldarelli, P. (2017). Grapevine Vitiviruses. In: *Grapevine Viruses*. Meng B, Martelli GP, Golino DA, Fuchs M. eds. Springer International Publishing AG.
- Nakaune, R., Toda, S., Mochizuki, M., & Nakano, M. (2008). Identification and characterization of a new vitivirus from grapevine. *Archives of Virology*, *153*, 1827–1832.
- Oliveira, L. M., Orflio, A. F., Inoue-Nagata, A. K., Nagata, T., & Blawid, R. (2017). A novel vitivirus-like sequence found in *Arracacia xanthorrhiza* plants by high throughput sequencing. *Archives of Virology*, *162*, 2141–2144.
- Rowhani, A., Daubert, S., Arnold, K., Al Rwahnih, M., Klaassen, V., Golino, D., & Uyemoto, J. K. (2018). Synergy between grapevine vitiviruses and grapevine leafroll viruses. *European Journal of Plant Pathology*, *151*, 919–925.
- Tzanetakis, I. E., Postman, J. D., & Martin, R. R. (2007). Identification, detection and transmission of a new vitivirus from *Mentha*. *Archives of Virology*, *152*, 2027–2033.
- Vončina, D., & Almeida, R. P. (2018). Screening of some Croatian autochthonous grapevine varieties reveals a multitude of viruses, including novel ones. *Archives of Virology*, *163*, 2239–2243.