Detection and diversity of grapevine virus L from a *Vitis* cultivar collection in Stellenbosch, South Africa

David A. Read [1,2*], Genevieve D. Thompson [1,3], Dirk Swanevelder [1] and Gerhard Pietersen [2]

[1] Agricultural Research Council (ARC) - Biotechnology Platform, 100 Old Soutpan Road, Onderstepoort, Pretoria, 0110, South Africa

[2] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

[3] Gene Vantage, 34 Monte Carlo Crescent, Kyalami Business Park, Johannesburg, 1684, South Africa

* - Corresponding author

Email: david.read@fabi.up.ac.za

ORCID – David A. Read

0000-0002-6614-9946

Abstract

A total of 229 *Vitis* cultivars were sampled from a collection in Stellenbosch, Western Cape, South Africa and subjected to an RNAtag-seq workflow and Illumina HiSeq 2500 sequencing. Following *de novo* assembly and initial BLASTn analysis of the resulting reads showed that 85 cultivars were infected with grapevine virus L (GVL), with a total of 96 complete/near complete genomes. This is the first time that GVL has been detected in South Africa. Phylogenetic analyses of the amino acid sequences of ORF1 showed that GVL from this study is diverse, grouping with references from Croatia, USA and China, as well as forming a unique South African phylogroup.

Keywords

South Africa, grapevine virus L, RNAtag-seq

Vitis vinifera (grapevine) is associated with an exceptionally high diversity of viral pathogens, with close to 90 viruses from 34 genera (Fuchs, 2020). Nine species of the *Vitivirus* genus (*Betaflexiviridae*) are known to infect *Vitis*. These have +ssRNA genomes of ±7,500 nt with *Grapevine virus A* (GVA) being the exemplar species (Maree et al., 2020) and many have

been shown to be transmitted by various species of mealybug (Le Maguet et al., 2012). *Vitivirus* genomes encode for five open reading frames (ORFs), with ORF1 encoding for the replication associated proteins (RAP) and ORF2 for the "22-kDa" protein, which is thought be associated with mealybug transmission (Galiakparov et al., 2003). Cell to cell movement is facilitated by the movement protein (MP) and the coat protein (CP), which are expressed by ORFs 3 and 4 respectively. ORF 5 encodes for a nucleic acid binding protein (NABP) that functions as suppressor of RNA silencing (Galiakparov et al., 2003). Grapevine virus L (GVL) is a recently described betaflexivirus and putative vitivirus, with a typical genome organization for the genus (Debat et al., 2019). GVL was simultaneously detected from geographically diverse regions, including Canada, New Zealand, Croatia, and China (Debat et al., 2019). Since then, GVL has been detected in Tunisia (Ben Amar et al., 2020), USA (Diaz-Lara et al., 2019; Alabi et al., 2020), and Turkey (IlbağI et al., 2021). Despite not yet being formally recognized as a species (Maree et al., 2020), it appears to have a global distribution (Ben Amar et al., 2020).

A total of 229 samples, in the form of leaf petioles, were collected from the *Vitis* cultivar collection at the Agricultural Research Council's (ARC) Nietvoorbij campus, Stellenbosch, South Africa. Each sample was collected from a different *Vitis* species or cultivar of *V. vinifera* and represents between 1 and 5 pooled vine replicates per sample (Supplementary Table 1). Petioles were macerated using Bioreba extraction bags (Bioreba, Reinach, Switzerland) and RNA isolated using a modified cetyltrimethylammonium bromide protocol (White et al., 2008). RNA quality control was performed using both the NanoPhotometer N60 (Implen, Munich, Germany) and Qubit 3 (RNA BR assay, Thermo Fisher Scientific, Waltham, MA, United States). RNAtag-seq libraries were then prepared according to Shishkin et al., (2015), with multiplexing of up to 32 samples per library. The resulting libraries were run on a HiSeq 2500 instrument (Illumina, San Diego, CA, United States) at the ARC's Biotechnology Platform (Onderstepoort, Pretoria, South Africa), as paired-ends (2x125nt) using TruSeq V4 chemistry (Illumina, San Diego, CA, United States).

The resulting datasets were demultiplexed using the Je suite (Girardot et al., 2016) and trimmed using CLC Genomics Workbench 9 (Qiagen Bioinformatics, Aarhus, Denmark), with the following parameters: Minimum read length = 20bp, quality limit = 0.05 and adapter trimming with Illumina universal: 5' AGATCGGAAGAG 3' and RNAtag-seq adapters: 5' TACACGACGCTCTTCCGATCTNNNNNNNNT 3'. *De novo* assembly of trimmed reads was performed using metaSPAdes within SPAdes 3.14.0 (Nurk et al., 2017). Viral contigs were identified using BLASTn (Altschul et al., 1990) and the NCBI Refseq database for viruses.

Contigs showing homology to GVL were identified and associated with 85 cultivars of *V. vinifera*, *V. shuttleworthii* and *V. labrusca* or interspecific hybrids thereof, with a total of 96 complete genomes, with some variants being present as mixed multiple infections. Cultivars, together with the GVL genome length and coverage, as well as NCBI GenBank accession numbers are listed in Supplementary Table 1. The amino acid sequences of the putative RNA-dependent RNA polymerase gene (RdRp) product were determined using ORF finder (Wheeler et al., 2003). Sequence alignments were performed using ClustalW within BioEdit 7.2.5 (Hall, 1999). The best-fit maximum likelihood (ML) model was determined using MEGA X (Stecher et al., 2020). The ML phylogeny was generated using the Le Gascuel model with 1000 bootstrap replicates (Le and Gascuel, 2008).

GVL genome lengths varied from 7534 to 7637nt with a modal length of 7594nt. The ML phylogeny illustrated four phylogroups (Supplementary Figure 1), three of which contained variants from this study. Phylogroup I contained the largest number of variants, at 68 (71%) and clustered with extant variants from Croatia and Tunisia. Phylogroup III contained 19 variants but did not contain extant references, suggesting that these are novel variants of GVL. Phylogroup IV was associated with reference sequences from China and the USA (MH248020 and MH643739) and contained the fewest variants from this study, at 9. Variants clustering with the TX-WAT references (phylogroup II) from the USA (Alabi et al., 2020) were absent. MW309717, MW309747 and MW309807 were selected as representatives for phylogroups I, III and IV respectively and used to generate an abridged phylogeny (Figure 1). Pairwise comparisons showed that the genomes of the variants within phylogroup I all shared >97% nucleotide identity and >99% within phylogroups III and IV.

Due to the phylogenetic isolation of phylogroup III, the average amino acid identities (AAI) of the putative genes from MW309747 against those of the most homologous GVL sequence variant were determined. RdRp: 83.7% AAI with MH248020; 22kDa protein: 62.1% with MH643739; MP: 84.4% with MT269789; CP: 91.5% with MT269788 and NABP: 89.1% with MT319083. Although this study represents diverse putative strains, they all represent isolates of GVL as the shared amino acid sequences of the cognate coat protein genes are >80%, which is above the species demarcation threshold for vitiviruses (Adams et al., 2004).

Potential recombination events were determined using RDP (Martin and Rybicki, 2000) within RDP4 (Martin et al., 2015). This suggested that phylogroup IV consists of recombinant variants originating from a single event with a member of phylogroup III acting as the major parent and phylogroup I as the minor parent (Supplementary Figure 2). Since this putative

recombinant has already been detected in China, USA and now also in South Africa, it appears to be well established within the global GVL population.

The presence of GVL was confirmed for each sample, using the two-step RT-PCR protocol described by Read et al., (2020) using the following primer pair GVL-CP-F (5' GGA GAC CTC AGT TGA AGT CTA GGT 3') and GVL-CP-R (5' TAA CGT AAT TAC CCT CGT TTT CGG T 3'), targeting conserved binding sites and amplifying a 226bp amplicon portion of the coat protein gene. All samples tested with RT-PCR yielded a band of the expected size, with the amplicons associated with 10-07 Cornichon violet, 18-01 Lindley and 30-05 Schonberger, selected at random and sequenced with Sanger in the forward direction, to confirm GVL identity (Inqaba Biotechnical Industries, Pretoria, South Africa).

This study represents the first detection of GVL in South Africa and adds significantly to the existing knowledge regarding GVL genomics. With GVL being a recently described virus, there is currently no information relating to symptomatology, complicated by the fact that GVL variants from this study were part of complex populations of diverse viruses. When present as single infections, grapevine vitiviruses are generally very mild, with no specific foliar symptoms, or are asymptomatic (Rowhani et al., 2018). An apparent synergy has been shown to exist between some of the grapevine vitiviruses and grapevine leafroll-associated viruses (GLRaV), which results in both the increased severity of symptoms and increased titre of the vitivirus component (Rowhani et al., 2016). The interaction appears to be specific (Rowhani et al., 2018), with the most pronounced potentiation of symptoms shown for the GVA/GLRaV-3 association (Credi and Babini, 1997). The mutualistic co-transmission of certain grapevine ampelo- and vitiviruses has also been hypothesized (Hommay et al., 2008), further illustrating the intricate associations between these viruses. In this study, GLRaV-3 was present in all viral populations associated with GVL, with GLRaV-2 also a common co-infector. This is unsurprising, as GLRaV-3 is the most common and widespread virus of grapevine in South Africa (Maree et al., 2013). Most populations were also associated with at least one additional vitivirus, with GVA and grapevine virus B among those most common (data not shown). The association of GVL with any of the GLRaV viruses, if any, will require extensive further research, as well the determination of possible mealybug vectors.

Given how widespread GVL is across the sample site of this study it is likely that it has been present in South Africa for an extended period of time. Furthermore, the fact that these variants group into four phylogroups suggests that GVL has been brought in through multiple introductions. The sequence divergence of GVL, relative to other vitivirus species, has probably been behind the failure of its detection by targeted diagnostic techniques such as RT-

PCR, due to variation in critical primer binding sites (Diaz-Lara et al., 2020). This limitation has been removed with the rapid adoption of next-generation sequencing technologies for the characterization of plant virus genomes, which has led to the discovery of many novel plant viruses over the last decade (Roossinck, 2017). Generally, the virology of grapevines in South Africa is complex and there is need to maintain a perspective of viruses affecting the industry, as well as having a comprehensive set of diagnostic tools for their detection. This is especially important when maintaining the "virus-free" status of planting material that will be distributed to producers (Jooste et al., 2015).

Declarations

Funding: David Read is grateful for the financial support provided by the National Research Foundation of South Africa, under grant UID 104901.

Conflict of interest: All authors declare that they have no conflict of interest.

Availability of data and material: The data that support the findings of this study, are openly available in NCBI public databases.

Code availability: Not applicable

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

References

Adams, M.J., Antoniw, J.F., Bar-Joseph, M., Brunt, A.A., Candresse, T., Foster, G.D., Martelli, G.P., Milne, R.G., Zavriev, S.K., & Fauquet, C.M. (2004). The new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation. *Archives of Virology*, *149*, 1045–1060

Alabi, O.J., Appel, D.N., McBride, S., Rwahnih, M.A., & Pontasch, F.M. (2020). Complete genome sequence analysis of a genetic variant of grapevine virus L from the grapevine cultivar Blanc du Bois. *Archives of Virology, 165*, 1905–1909

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molelcular Biology, 215*, 403–410

Ben Amar, A., Daldoul, S., Zemni, H., Wetzel, T., Olmos, A., & Ruiz-García, A.B. (2020). First report of Grapevine virus L in grapevine in Tunisia. *Plant Disease, 104*, 12

Credi, R., & Babini, R. (1997). Effects of virus and virus-like infections on growth, yield, and fruit quality of Albana and Trebbiano Romagnolo grapevines. *American Journal of Enology and Viticulture, 48*, 7–12

Debat, HJ., Zavallo, D., Brisbane, R.S., Voncina, D., Almeida, R.P.P., Blouin, A.G., Al Rwahnih, M., Gomez, T.G., & Asurmendi, S. (2019). Grapevine virus L: A novel vitivirus in grapevine. *European Journal Plant Pathology, 155*, 319– 328

Diaz-Lara, A., Brisbane, R.S., Aram, K., Golino, D., & Al Rwahnih, M. (2019). Detection of new vitiviruses infecting grapevine in California. *Archives of Virology, 164*, 2573–2580

Diaz-Lara, A., Erickson, T.M., Golino, D., & Al Rwahnih, M. (2020). Development of a universal RT- PCR assay for grapevine vitiviruses. *PLoS ONE, 15 (9),* e0239522

Fuchs, M. (2020). Grapevine viruses: a multitude of diverse species with simple but overall poorly adopted management solutions in the vineyard. *Journal of Plant Pathology, 102*, 643–653

Galiakparov, N., Tanne, E., Sela, I., & Gafny, R. (2003). Functional analysis of the Grapevine Virus A genome. *Virology, 306*, 42–50

Girardot, C., Scholtalbers, J., Sauer, S., Su, S.Y., & Furlong, E.E. (2016). Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics, 17(1)*, 419

Hall, T.A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95–98

Hommay, G., Komar, V., Lemaire, O., & Herrbach, E. (2008). *Grapevine virus A* transmission by larvae of *Parthenolecanium corni. European Journal of Plant Pathology, 121*, 185–188

IlbağI, H., Panailidou, P., Gagiano, E., Pietersen, G., Maliogka, V., Katis, N., & Çıtır, A. (2021). First report of grapevine virus L in grapevine in Turkey. *Journal of Plant Pathology, 103,* 343

Jooste, A.E.C., Molenaar, N., Maree, H.J., Bester, R., Morey, L., de Koker, W.C., & Burger J.T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. *European Journal of Plant Pathology, 142*, 363–375

Le, S.Q., & Gascuel, O. (2008). An improved general amino acid replacement matrix. *Molecular Biology and Evolution, 25(7),* 1307–1320

Le Maguet, J., Beuve, M., Herrbach, E., & Lemaire, O. (2012). Transmission of six ampeloviruses and two vitiviruses to grapevine by *Phenacoccus aceris. Virology, 102,* 717-723

Maree, H.J., Almeida, R.P.P., Bester, R., Chooi, K.M., Cohen, D., Dolja, V.V., Fuchs, M.F., et al., (2013). Grapevine leafroll-associated virus 3. *Frontiers in Microbiology, 4*, 82

Maree, H.J., Blouin, A.G., Diaz-Lara, A., Mostert, I., Al Rwahnih, M., & Candresse, T. (2020). Status of the current vitivirus taxonomy. *Archives of Virology, 165*, 451–458

Martin, D. & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics, 16*, 562–563

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution, 1*, vev003

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P.A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research, 27,* 824–834

Read, D.A., Roberts, R., & Thompson, G.D. (2020). First report of alstroemeria mosaic virus in South Africa. *Australasian Plant Disease Notes, 15*, 40

Roossinck, M.J. (2017). Deep sequencing for discovery and evolutionary analysis of plant viruses. *Virus Research, 239*, 82–86

Rowhani, A., Uyemoto, J. K., Golino, D., Daubert, S. D., & Al Rwahnih, M. (2016). Viruses involved in graft- incompatibility and decline. In B. Meng, M. Fuchs, G. Martelli, & D. Golino (Eds.), Grapevine viruses: Molecular biology, diagnostics, and management, Chapter 12 (pp. 289–302). New York: Springer.

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

Shishkin, A.A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., Chen, J., et al., (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. *Nature Methods, 12*, 323–325

Stecher, G., Tamura, K., & Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular Biology and Evolution, 37*, 1237–1239

Wheeler, D.L., Church, D.M., Federhen, S., Lash, A.E., Madden, T.L., Pontius, J.U., Schuler, G.D., Schriml, L.M., Sequeira, E., Tatusova, T.A., & Wagner, L. (2003). Database resources of the National Center for Biotechnology. *Nucleic Acids Research, 31*, 28–33

White, E., Venter, M., Hiten, N.F., & Burger, J.T. (2008). Modified Cetyltrimethylammonium bromide method improves robustness and versatility: The benchmark for plant RNA extraction. *Biotechnology Journal, 3,* 1424–1428

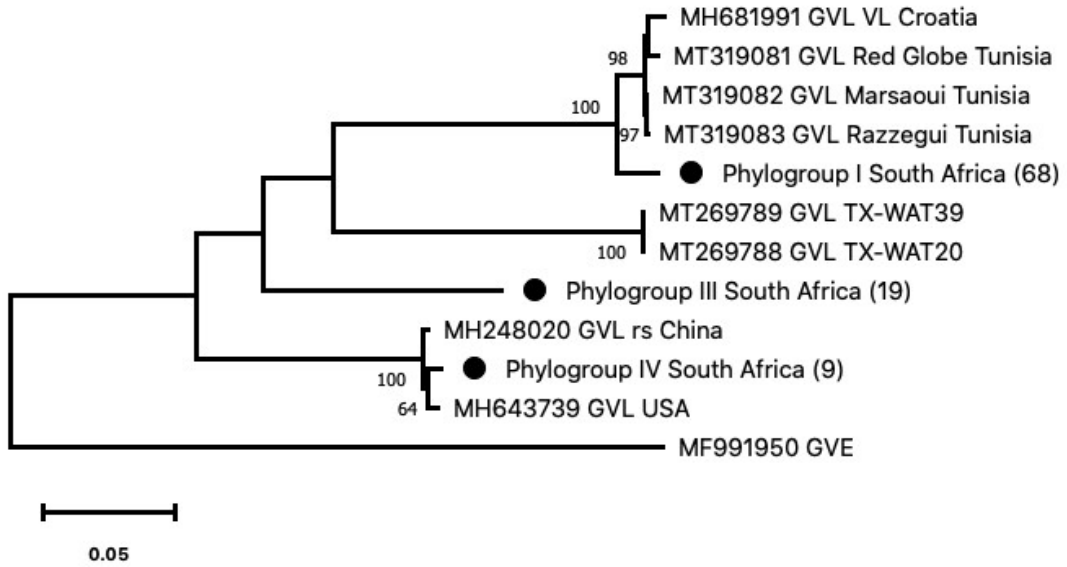

Figure 1: Abridged maximum-likelihood phylogeny for ORF1 of the grapevine virus L (GVL) variants from this study (indicated by solid circles), together with that of that of available references. The numbers of homologous South Africa GVL variants are shown in brackets.