Tropical Plant Pathology
https://doi.org/10.1007/s40858-024-00647-3

SHORT COMMUNICATION

Analyzes of mealybug (*Pseudococcus longispinus*) virome reveal grapevine viruses diversity

Thor Vinícius Martins Fajardo[1] · Priscila Grynberg[2] · Roberto Coiti Togawa[2] · João Marcos Fagundes Silva[3] · Fabio Nascimento da Silva[4] · Osmar Nickel[1]

Received: 25 October 2023 / Accepted: 27 February 2024
© The Author(s), under exclusive license to Sociedade Brasileira de Fitopatologia 2024

Abstract

The long-tailed mealybug, *Pseudococcus longispinus*, is an important insect pest in grapevine growing areas in several countries, including Brazil. Metagenomic analysis of nucleic acids extracted from insect vectors makes it possible to study the diversity of insect viruses in addition to plant pathogenic viruses. In this study, insects (*Ps. longispinus*) were collected, and pooled throughout a plot of virus disease symptomatic vines, cultivated in growing beds, and analyzed by high throughput sequencing (HTS). The complete genome of grapevine leafroll-associated virus 2 and 3 (GLRaV-2 and -3) and a partial sequence of grapevine virus A (GVA) with two complete ORFs (coat protein and RNA-binding protein) were assembled from mealybug extracts and exhibited high nucleotide identities, up to 99%, with previously characterized homologous Brazilian isolates from grapevines. This information was validated by the detection of these viruses in the original symptomatic vines (N=76), from where mealybugs were collected, equivalent to an incidence of 34.2%, 89.5% and 36.8% for GLRaV-2, GLRaV-3 and GVA, respectively. Although one of these viruses is not transmitted by mealybugs (GLRaV-2), prospection of plant viruses infecting grapevine plants by analyzing the metagenome of insects could represent a relevant alternative to improve monitoring of viral diseases aiming at the management and control of viral diseases in vineyards or cultivation fields. This work is the first analysis of the *Ps. longispinus* virome in Brazil focusing on grapevine viruses.

Keywords Closterovirus · Ampelovirus · Vitivirus · Vector · High throughput sequencing

A total of 101 viruses have been identified in grapevines (*Vitis* spp.) worldwide thus far, belonging to 21 different families, and often found in mixed infections (Fuchs 2023). However, not all of these viruses represent a serious threat to global viticulture (Wu et al. 2023b). From an economic perspective and considering the adverse effects they can cause in infected vines, around a third of these viruses are important due to their high pathogenicity and involvement in widespread complex diseases such as infectious degeneration (fanleaf), rugose wood complex, and leafroll disease (Schoelz et al. 2021). In Brazil, at least 20 different viruses have already been reported infecting grapevines, the most detrimental viral diseases being related to leafroll disease, specifically caused by grapevine leafroll-associated ampelovirus 3 (GLRaV-3) and grapevine leafroll-associated closterovirus 2 (GLRaV-2) in addition to viruses associated with rugose wood complex (vitiviruses grapevine virus A (GVA) and grapevine virus B (GVB) and, grapevine rupestris stem pitting-associated foveavirus (GRSPaV)) (Basso et al. 2017).

Multiple viruses may accumulate in a grapevine over its life span, and individual viruses or combinations of different viruses may have detrimental effects on the physiology of the vine, yield, and berry quality, thus negatively affecting vineyard performance (Song et al 2021). Often, symptoms cannot be attributed to a single virus, as symptomatic vines usually have multiple virus infections or certain grapevine genotypes may be asymptomatically infected by viruses (Mannini and Digiaro 2017; Fuchs 2020).

✉ Thor Vinícius Martins Fajardo
thor.fajardo@embrapa.br

[1] Embrapa Uva e Vinho, Bento Gonçalves, RS 95.701-008, Brazil

[2] Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF 70.770-917, Brazil

[3] Institute for Integrative Systems Biology (I2SysBio), 46980 Paterna, Valencia, Spain

[4] Departamento de Agronomia, Universidade do Estado de Santa Catarina, Lages, SC 88.520-000, Brazil

Published online: 14 March 2024

Considering plant diseases, their development and their effects on yield result from the interaction between the host genotype, the pathogen genotype, and the environment in which the plant is found (environment consisting of climate, soil and crop conditions, occurrence of abiotic stresses, presence of insect vectors of viruses) (Rubio et al 2020). Mealybugs cause mechanical and feeding damages to plants and can be vectors of plant viruses, representing an important issue that results in economic losses. Through their feeding activities, they can reduce yield, quality, and productivity of vines (Ahmed et al. 2023).

Among several mealybugs, species in genera *Planococcus* and *Pseudococcus* are the most effective vectors of grapevine viruses, transmitting ampeloviruses and vitiviruses. Currently, many studies have identified several mealybugs, Pseudococcidae family (eg. *Heliococcus bohemicus*, *Pseudococcus affinis/Ps. viburni*, *Planococcus citri*, and *Pl. ficus*), and soft scales, Coccidae family (eg. *Neopulvinaria innumerabilis*), as vectors of several grapevine viruses (Vončina et al. 2024). Vector transmission of these viruses occurs in a semi-persistent manner, with no current evidence of a strict vector-virus species-specificity (Herrbach et al. 2017; Tsai et al. 2010). Important studies on the epidemiology of the ampelovirus GLRaV-3 and the vitiviruses GVA and GVB demonstrated that these viruses can be transmitted from vine to vine by *Ps. longispinus* (Kuniyuki et al. 2006; Tsai et al. 2010). Several mealybug populations, collected in grapevines or virus-infected vineyards, have already tested positive for grape viruses, e.g. GVA (La Notte et al. 1997). However, it is important to highlight that a biological virus-vector relationship is required for the latter to be considered a vector (Wistrom et al 2017).

The long-tailed mealybug, *Pseudococcus longispinus* (Insecta: Hemiptera, Pseudococcidae) is one of major vineyard mealybugs worldwide, including the Brazilian southern region, the largest grape production area in this country (Kuniyuki et al. 2006).

Vectors are responsible for transmitting viruses over short and long distances, thus, they are a component that must be considered in the epidemiology of diseases and, consequently, in disease management and control strategies (Grohs et al 2017). The management of grapevine viral diseases is based on prevention and suppression of the viral inoculum. The prevention of virus introduction in new vineyards is achieved by the use of healthy propagation material and reduction of the viral inoculum in infected vineyards is obtained by roguing diseased plants (Maliogka et al. 2015; Fuchs 2020). Therefore, for control and management strategies for grapevine viruses to be more effective, it will depend on a robust understanding of virus-vector transmission efficiency, virus-vector molecular interactions, how these viruses (ampelo- and vitivirus) are disseminated in vineyards (Herrbach et al. 2017) and, on the viral pathogen involved in the disease.

High throughput sequencing (HTS) represents a great advance for the study of grapevine viruses considering that it does not require prior information on the virus, its importance in terms of sensitivity, its wide application possibilities and reliability of the results. Finally, the implementation of HTS for diagnosis of viruses whether infecting grapevines or present in vectors or insects is extremely useful and elucidative (Fajardo et al 2020).

The objectives of this work were to survey metagenomic data of *Ps. longispinus* virome focusing on grapevine viruses and relate the presence of viruses found in insects with those in the vines from where the insects were collected.

A plot with approximately 450 plants (grapevine cutting accessions), maintained in a collection under greenhouse conditions at Embrapa Uva e Vinho, Brazil, in growing beds, most of them exhibiting virus-related symptoms, was evaluated. The observed symptoms were related with leafroll disease and rugose wood complex, two major diseases of grapevine, such as downward curling of the leaf edges and leaf reddening and green main veins or reddish leaves with abnormal appearance with different severity of symptoms (Fig. 1). Seventy-six plants (29 samples of cultivar Cabernet Franc, 24 of cv. Cab. Sauvignon and 23 of cv. Merlot.) exhibiting the described symptoms were chosen and the dormant canes sampled for viral indexing.

Subsequently, an intense infestation of mealybugs (pseudococcids) was observed in these plants (Fig. 1). Since mealybugs are known to act as virus vectors in grapevines (Herrbach et al. 2017), a study was conducted to determine whether the mealybug virome would reflect the virus-related phytosanitary status of diseased vines. For this purpose, 690 mg of pseudococcids, mainly adults, were collected using a fine paintbrush in the grapevine diseased symptomatic block (N=450). For metagenomics analysis, after grinding insects in liquid nitrogen, the total RNA was extracted using Trizol (Invitrogen) and the extraction buffer [0.5 M Tris-HCl pH 8.3, 0.14 M NaCl, 2% PVP-40 (w/v), 0.05% Tween-20 (v/v)], chloroform and isopropyl alcohol were used for phase separation and nucleic acid precipitation, respectively. The final pellet was dried, resuspended in autoclaved water and cleaned up with the RNeasy Plant Mini kit (Qiagen).

To molecularly identify the mealybug species, RT-PCR was carried out using primers defined by Beuning et al. (1999) for specific amplification of the *Pseudococcus longispinus* genome: L3 (5' CGGGTGTATACGTGTGTACG 3'), forward, and, L2 (5' CGTACGCGAAACGCATGTAA 3'), reverse. Amplification cycling (94 °C/15 min, and 35 cycles: 94 °C/50s, 48 °C/50s, 72 °C/50s) and, final extension of 72 °C/10 min, gel electrophoresis, cloning and purification of cloned recombinant plasmids from *Escherichia coli* were

Fig. 1 Representative sampled grapevines and mealybugs. Leafroll and virus-related foliar symptoms exhibited in cv. Cabernet Franc (**a, b**) and cv. Merlot (**c**) with different severity. Adults (approximately 3 to 5 mm long) and nymph individuals (**d**) of evaluated mealybug (*Pseudococcus longispinus*)

performed as described by Fajardo et al. (2023). Sanger nucleotide sequencing was performed with two clones.

Total RNA extracted from mealybugs was analyzed qualitatively ($A_{260}/_{280}$ = 2.13 and, $A_{260}/_{230}$ = 2.16) and quantitatively (476 µg) using NanoDrop 2000 spectrophotometer (Thermo Scientific) and used to perform high throughput sequencing (HTS). The sequencing library was generated using Illumina TruSeq stranded total RNA with Ribo-zero Plant kit (Illumina, USA) for plant ribosomal RNA depletion following manufacturer's recommendations and then subjected to HTS on Illumina HiSeq X-ten platform (USA).

To quantify grapevine transcripts from *Pseudococcus longispinus* RNA-Seq data, raw sequence data were analyzed for sequence quality, GC content, presence of sequencing adapters, overrepresented k-mers, and duplicated reads, using fastQC 0.12.0 (Andrews 2010). Reads were trimmed using Trimmomatic 0.40 (Bolger et al. 2014) in the 'adapters and SW' and 'adapters and MI' modes, with the following parameters PE -threads 8 -phred33 ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36, and PE -threads 8 -phred33 ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 MAXINFO:50:0.5 MINLEN:36, respectively. Clean reads were mapped against the grapevine transcriptome *Vitis vinifera* PN40024.v4.1 and its most recent annotation VCost.v3 (Canaguier et al. 2017) using Salmon v.1.2.0 (Patro et al. 2017) in the mapping-based mode running up to 8 threads, using the default k-mer of 31, with automatic library detection, bootstrapped and variance abundance estimates, sequence specific and GC bias corrections. The parameters were -l A --numBootstraps --numGibbsSamples -seqBias --gcBias -p 8 -validateMappings. The algorithm is a pseudo-aligner based on a two-phase inference quasi-mapping that makes use of massively-parallel stochastic collapsed variation

inference, and bias modeling to account for sequence-specific, position-specific, and fragment-GC content tendencies.

To carry out bioinformatic analysis for grapevine viruses, raw reads were first processed by Fastp (Chen et al. 2018) and the obtained sequences of *Ps. longispinus* were removed by Bowtie 2 program (Langmead and Salzberg 2012) based on the GenBank accession code GCA_900064475.1 (assembled genome of *Ps. longispinus*). The remaining reads were assembled by MEGAHIT v.1.2.9 (Li et al. 2016). The resulting contigs were classified by BLASTn (Camacho et al. 2009), first against a RefSeq complete genome database (O'Leary et al. 2016) and against the standard nucleotide (nt) database. For each identified virus species, the best contig (lowest e-value and highest bitscore) was extracted. Additionally, to obtain the complete genome of GLRaV-2, reads were trimmed with BBDuk v.39.01 with the parameters tbo qtrim=lr trimq=10 ktrim=r ref=adapters k=21, and mapped against the complete genome of GLRaV-2 isolate A1584e (accession: ON645914.1) using BBMap v.39.01 with the fast parameter, generating a consensus sequence with Geneious Prime (Dotmatics).

Nucleotide sequences obtained by HTS were deposited in the GenBank database and aligned with the viral reference sequences, some homologous foreign isolates retrieved from GenBank and/or other homologous Brazilian isolates of studied viruses obtained from vines, using the NCBI's BLASTN program (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The following previously characterized Brazilian isolates of grapevines viruses were included in the phylogenetic analyzes, GLRaV-2 complete coat protein genes (EU053125, EU053126, EU204909, EU204910, EU204911, EU204912 and KX774192), GLRaV-3 complete genomes (KX756669, KX701860, KX756668 and MK804765) and GVA complete coat protein genes (AF494187, KF667501, KX828703 and AY340581). Multiple nucleotide sequence alignments of complete coat protein genes (GVA and GLRaV-2) and complete genomes of GLRaV-3 were performed using ClustalX 2.1 (Larkin et al. 2007) and the phylogenetic relationships were determined from the aligned sequences by using the Neighbor-Joining (NJ) method with the Kimura 2-parameter model, and 1,000 bootstrap replications implemented in MEGA X software (Kumar et al. 2018).

Real-time RT-PCR (RT-qPCR TaqMan) was used to validate the presence of grapevine viruses identified by HTS in total RNA of original source of mealybugs and to detect grapevine viruses in the 76 sampled grapevines. Total RNA extraction from one gram of scrapings of grapevine dormant canes was performed by grinding them in liquid nitrogen and thereafter using a silica adsorption protocol (Rott and Jelkmann 2001). The oligonucleotides and probes used in the RT-qPCR reactions were previously defined by Osman et al. (2007) for GLRaV-2 and GLRaV-3 and by Osman and Rowhani (2008) for GVA. Components and cycling conditions of RT-qPCR reactions were previously described (Dubiela et al. 2013), consisting of presence/absence assays, using the TaqMan Fast Virus 1-Step Master Mix kit (Applied Biosystems) and the StepOnePlus Real-time PCR thermocycler System (Applied Biosystems). The reactions were analyzed by determining the Ct (cycle threshold), considering positive samples those with Ct ≤ to 35 and using the StepOne Software v2.3 (Applied Biosystems).

The incidence of mealybugs in the evaluated vine plot was estimated to around 60% of the plants with at least one mealybug. As expected, a DNA fragment of approx. 1,140 nucleotide (nt) length was amplified from total RNA of the same mealybug sample used for HTS. The obtained nucleotide sequence covering the partial sequence of the internal transcribed spacer (ITS) 1, complete sequence of 5.8S rRNA gene and partial sequence of ITS 2 with 1,138 nt was identified as *Pseudococcus longispinus* (GenBank accession OR502861). This sequence was aligned with homologous sequences and the highest nucleotide identity was 98.43% (query cover 100%) with *Ps. longispinus* from New Zealand (GenBank accession AF007264), thus confirming the initial morphological species identification.

Reads of *Vitis* origin were absent from the investigated samples, when analyzed as transcripts per million in relation to 40,979 transcripts in the grapevine transcriptome index (including differential splicing variants) (data not shown). Mealybug samples were collected from several grapevine genotypes of *V. vinifera*. The vine species is highly heterozygous and accumulates somatic mutations over time, as other clonally propagated crops (Zhou et al. 2019; Xiao et al. 2023). The final transcript abundance estimated by pseudoaligners is obtained by modeling sample-specific parameters and biases. Therefore, quasi-mapping approaches, which do not perform base-by-base alignment, as used in the current study, are adequate for heterozygous species and cultivars distinct from the one used to build the transcriptome index. It is concluded that the assembled viral contigs are the sequencing result of total RNA extracted from mealybugs and not from their host plant (grapevine).

A total of 41,986,322 paired-end reads were obtained by HTS from *Ps. longispinus* sample (NCBI BioProject ID: PRJNA1020893). Considering the grapevine virus isolates detected, the information of HTS data analyzes (number of mapped reads, % mapped and mean coverage depth) were 3,089, ca. 0.008 and, 26.32 (GLRaV-2), 12,954, 0.0239 and, 47.6 (GLRaV-3) and, 48, 0.0001 and, 8 (GVA), respectively. Long contigs of sizes 16,427 nt (GLRaV-2), 18,368 nt (GLRaV-3), and 1,029 nt (GVA) were obtained from mealybugs as follows: GLRaV-2, isolate PL-BR2 (GenBank accession OR640977), complete genome covering nine ORFs; GLRaV-3, isolate PL-BR (GenBank accession OR546046), with the complete viral genome organized in 13 ORFs, and

GVA, isolate PL-BR3 (GenBank accession OR613005), partial genome covering complete coat protein and RNA-binding protein genes. All obtained sequences of Brazilian isolates were submitted to BLASTn search, resulting in high nucleotide identities that ranged from 86% (KF220376) to 98% (MH814502), query cover 99-100% to GLRaV-2; 93% (MH814485) to 99% (MH814489), query cover 97-100% to GLRaV-3 and, 87% (MK404720) to 91% (OP752640), query cover 99% to GVA with the homologous viruses retrieved from GenBank database. These nucleotide sequence identities allowed the accurate identification of these viruses in the analyzed mealybug sample.

High nucleotide identities were also observed when comparing viral isolates obtained from mealybugs with other homologous Brazilian viral isolates, isolated from grapevines. Pairwise comparisons between homologous genomic regions of isolate PL-BR2 of GLRaV-2 and Brazilian isolate from vine (KX774192) showed 87.4% nt identity considering the complete genome; isolate PL-BR of GLRaV-3 and vine Brazilian isolates (KX756669, KX701860, KX756668 and MK804765) showed 93%-99% nt identities considering the complete genome and, lastly isolate PL-BR3 of GVA and Brazilian isolates from vine (AF494187, KF667501 and KX828703) showed 84%-90% nt identities considering the coat protein gene. In general, local isolates exhibited high identities regardless of whether isolated from mealybug or vines. Finally, pairwise comparison between Brazilian homologous partial coat protein genes (451 nt) of GVA, isolate PL-BR3, from *Ps. longispinus* and the same virus, isolate PC40, from *Planococcus citri*, the citrus mealbug (HM358052), collected from *V. vinifera*, showed 77.3% nt identity.

The coat protein gene sequence of GLRaV-2 obtained from *Ps. longispinus*, PL-BR2 isolate was compared to six strains (lineages) of this virus PN, 93/955, H4, BD, RG and PV20 (Jarugula et al. 2010), with seven other isolates from Brazilian grapevines and nine other retrieved from GenBank (Fig. 2). The PL-BR2 isolate clustered with the PN strain clade, as well as other Brazilian isolates (SE, IT, L/I) that were placed in this same phylogroup. The remaining Brazilian isolates (ISA-BR, M/C, RI, MH) clustered with the H4 strain clade (Fig. 2). Jarugula et al. (2010) suggested lack of clustering by geographical origin when analyzing the genetic variability of populations of GLRaV-2, a virus that has no biological vector. Grapevine is a vegetatively propagated crop, and the introduction of contaminated planting materials from different external sources and subsequent local dissemination of planting materials via multiple sources are common. This situation can contrast with the geographical delimitation of virus variants of some vector-borne viruses.

Phylogenetic analysis of GLRaV-3 isolates further confirmed identification of the virus, showing PL-BR isolate from *Ps. longispinus* clustered in the same branch with two other Brazilian isolates from grapevines (TRAJ-BR, GenBank KX756669 and NUB-BR¸ MK804765) and isolates from Canada, USA, South Africa and Chile, forming group I as defined by Maree et al. (2015) (Fig. 3). This suggests that the clustering of different GLRaV-3 isolates into the same clade is independent of the isolation source or host of the isolate, mealybug or vine. Five GLRaV-3 isolates from different countries (Israel and South Africa) including two Brazilian isolates (ISAB-BR, KX701860 and TC-BR, KX756668) from different vine cultivars and other four isolates from several countries (China, USA, South Africa and Croatia) were distributed in the distinct clusters forming groups II and III, respectively, of the phylogroup (supergroup) A (Fig. 3). Higher order groups were defined by Maree et al. (2015) as supergroups designated A to D, supergroup A includes variant groups I-V. Recently, genetic variability among a wide range of Brazilian isolates of GLRaV-3 collected from infected grapevines was reported (Fajardo et al. 2023).

GVA isolate PL-BR3 clustered in group II, and other Brazilian grapevine isolates (RS, IT-BA, SP and TRAJ2-BR), previously reported, clustered in group I with other foreign GVA isolates considered and classified by Wu et al. (2023a) as groups I and II, symptomless and symptomatic groups, respectively, based on the complete coat protein sequences used for the phylogenetic analysis (Fig. 4).

The diversity of the GLRaV-2, GLRaV-3 and GVA within the mealybugs samples was analyzed. Reads previously trimmed with BBDuk were aligned to the viruses' genomes with BBMap with the slow parameter, then the alignment was recalibrated and realigned near indels following the GATK pipeline (Van der Auwera and O'Connor 2020). Variant calling was performed with LoFreq v2 (Wilm et al. 2012). The diversity of each viruses was estimated by the sum of the Shannon entropy at each position normalized by contig length × 2 (which corresponds to the maximum entropy of the contig). The normalized entropy of GLRaV-2, GLRaV-3 and GVA was of 0.032, 0.013 and 0.005, respectively, which indicated that the sampled GLRaV-2 population is the most diverse.

Pseudococcus longispinus is known as vector of leafroll-associated ampeloviruses (GLRaV-3 and GLRaV-4 strains 5 and 9) and of rugose wood-associated vitiviruses of grapevine (GVA and GVB). Regarding GLRaV-2, although it has been detected in this mealybug species, it is not mealybug-transmitted, as demonstrated in transmission assays; a biological vector of GLRaV-2 is unknown (Herrbach et al. 2017; Ahmed et al. 2023). The mealybug-borne status of GLRaV-3 and GVA and non-mealybug-borne status of GLRaV-2 reflects the phylogenetic topology of these viruses (Tsai et al. 2010). Other studies have already reported the presence of insect viruses in mealybugs (Martinez-Mercado et al. 2022) and also plant viruses in insects using HTS or

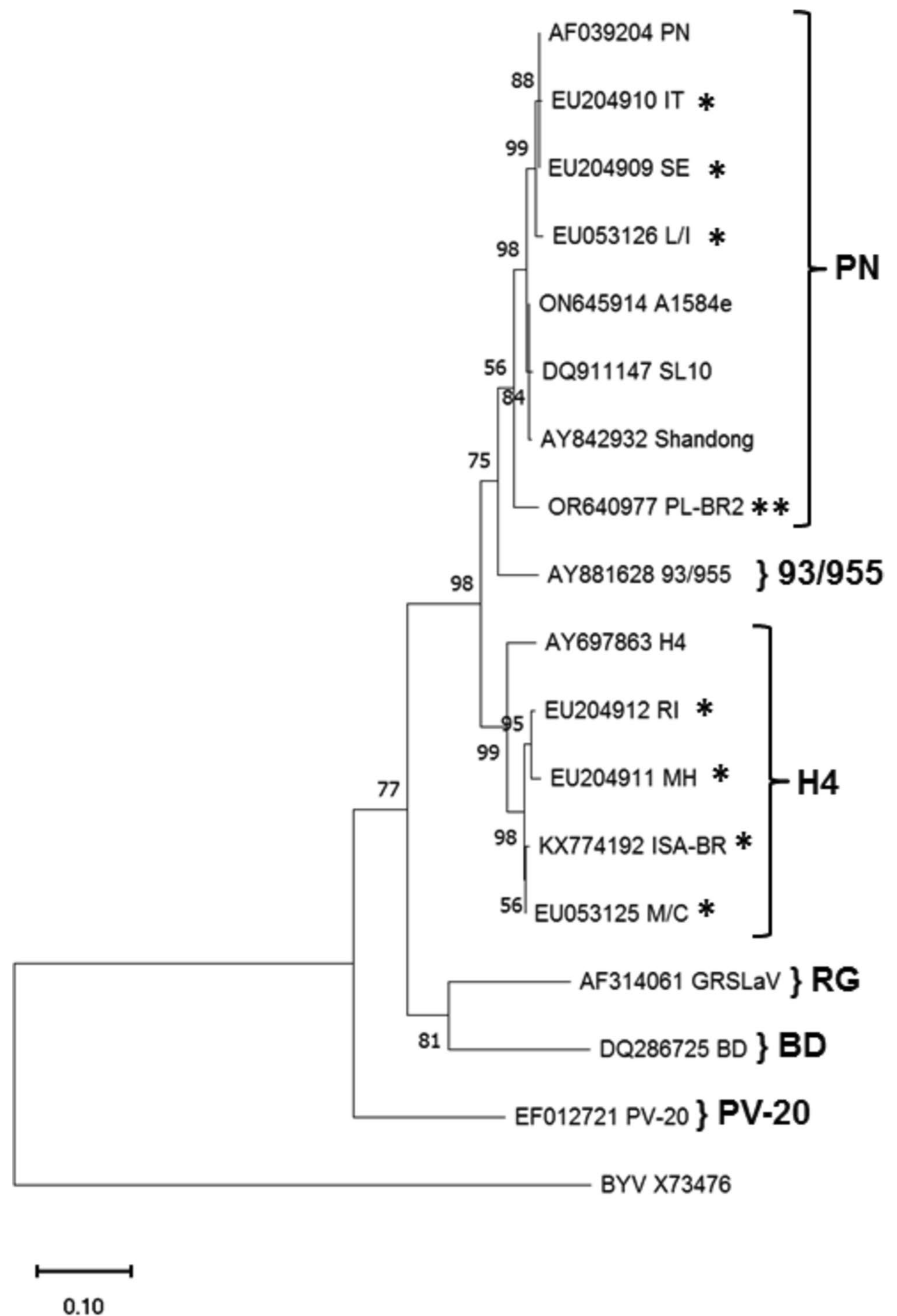

Fig. 2 Phylogenetic analysis of grapevine leafroll-associated virus 2 (GLRaV-2) based on complete coat protein nucleotide sequences. The Neighbor-joining (NJ) method with the Kimura 2-parameter model was used. Values at nodes are percent bootstrap support at 1,000 replications. Bootstrap percentage values are shown on the branches. Bars indicate the number of nucleotide substitutions per site. The tree is unrooted. Isolates include PL-BR2 (**) (GLRaV-2) from *Ps. longispinus* and seven previously obtained Brazilian homologous isolates from grapevines (SE, IT, L/I, ISA-BR, M/C, RI, and MH) (*), in addition to nine foreign isolates retrieved from GenBank. The phylogenetic groups indicated by brackets correspond to previously described GLRaV-2 (Jarugula et al. 2010). Nucleotide accessions and isolate names in GenBank are mentioned. For the external group of the tree, beet yellows virus (BYV, *Closterovirus*, X73476/NC_001598) and for type isolate, GLRaV-2 (AY881628/NC_007448) were used

RT-qPCR (Kaur et al 2016; McGreal et al 2019). Mealybugs are insects equipped with sucking mouth parts, acting like plant scale insects and aphids in sucking the fluids from leaves and stems of plants for the requirement of nutrients. Mealybugs are slow insects that transmit ampelo-, badna- and closteroviruses in a semi-persistent manner (Bhat and Rao 2020). However, for transmission to be effective, there must be a biological interaction between the virus and its vector, which does not occur between GLRaV-2 and, for example, mealybugs or aphids (phylloxera, *Daktulosphaira vitifoliae*) (Wistrom et al 2017). These authors mentioned that while, in rare cases, insects tested positive for GLRaV (-3 and -4LV), there was no evidence of virus transmission. Herrbach et al. (2017) mentioned that the natural vectors (if any) of GLRaV-2 are unknown but would be expected to be aphids based on the evolutionary history of the group/genus

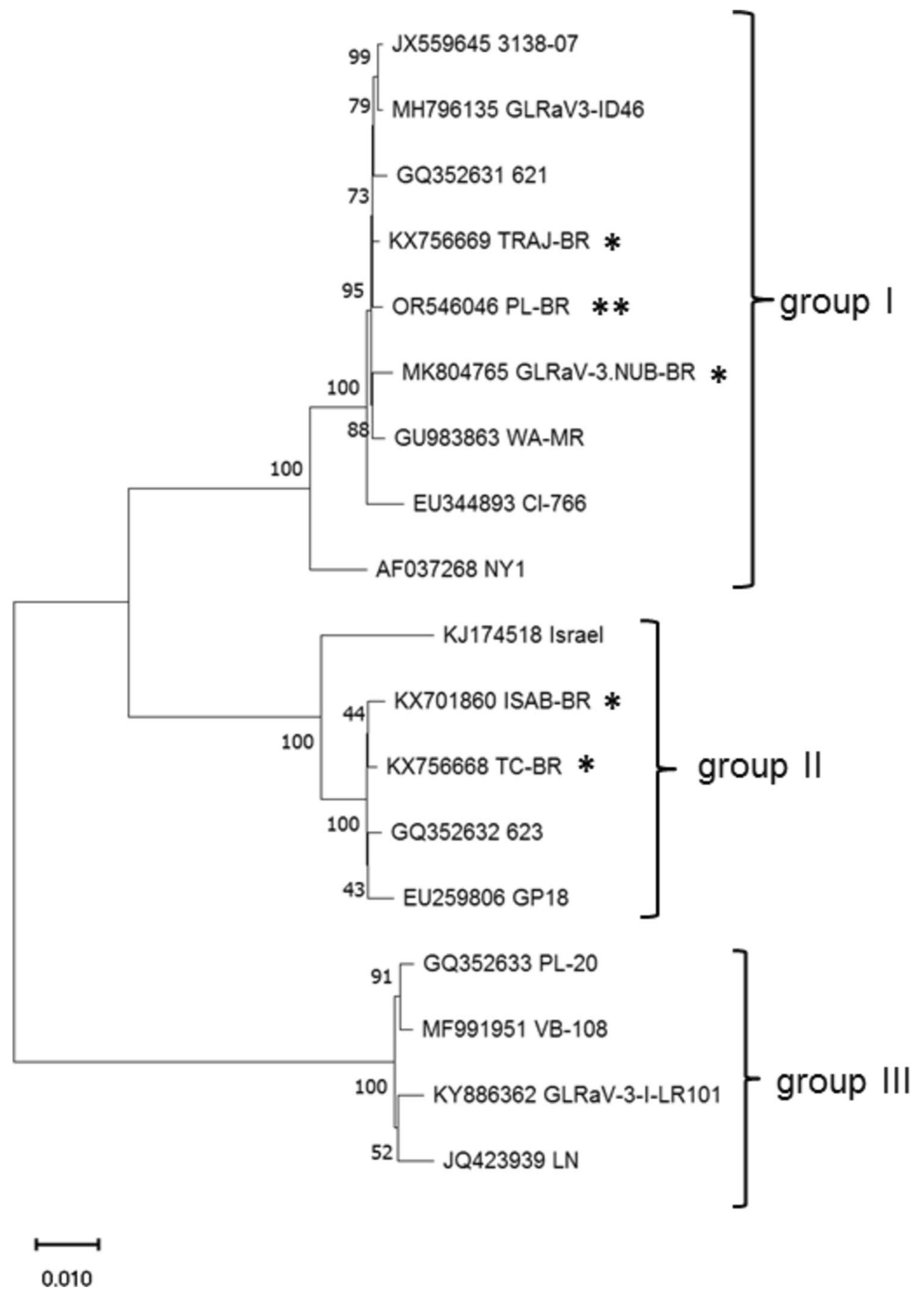

Fig. 3 Phylogenetic analysis of grapevine leafroll-associated virus-3 (GLRaV-3) based on complete genome nucleotide sequences. The Neighbor-joining (NJ) method with the Kimura 2-parameter model was used. Values at nodes are percent bootstrap support at 1,000 replications. The tree is unrooted. The phylogenetic groups indicated by brackets correspond to previously described phylogroups in the variant groups I-V included in supergroup A (Maree et al. 2015). The newly characterized isolate from mealybugs is indicated as PL-BR (**). Additional sequences of four Brazilian isolates, previously characterized, were TRAJ-BR, NUB-BR, ISAB-BR and TC-BR (*), that were phylogenetically related among three defined groups. For type isolate, GLRaV-3 (AF037268/NC_004667) was used

(Closterovirus). Thus, the presence of GLRaV-2 in this insect suggests that it may be a result of the sucking of vine sap and does not imply the biological ability to transmission of GLRaV-2 between the plants in the evaluated vine plot. These plants could have been infected by GLRaV-2 since the establishment of this experimental plot with imported and grafted vine cuttings.

Out of approximately 450 plants, 76 (16.9% of total) exhibiting symptoms of leafroll or rugose wood complex were visually evaluated and selected for indexing by RT-qPCR for GLRaV-2, GLRaV-3 and GVA. This could allow to establish a correlation between viruses detected by HTS in mealybugs sampled on vines and viruses found after indexing this group of plants. It was observed that 35.5% (27/76) of analyzed samples were GLRaV-2-infected grapevines (Ct of infected samples ranged from 17.70 to 34.65 and average of 30.40), 89.5% (68/76) of analyzed samples were GLRaV-3-infected grapevines (Ct of infected samples ranged from 15.95 to 34.93 and average of 25.04) and, 36.8% (28/76) of analyzed samples were GVA-infected

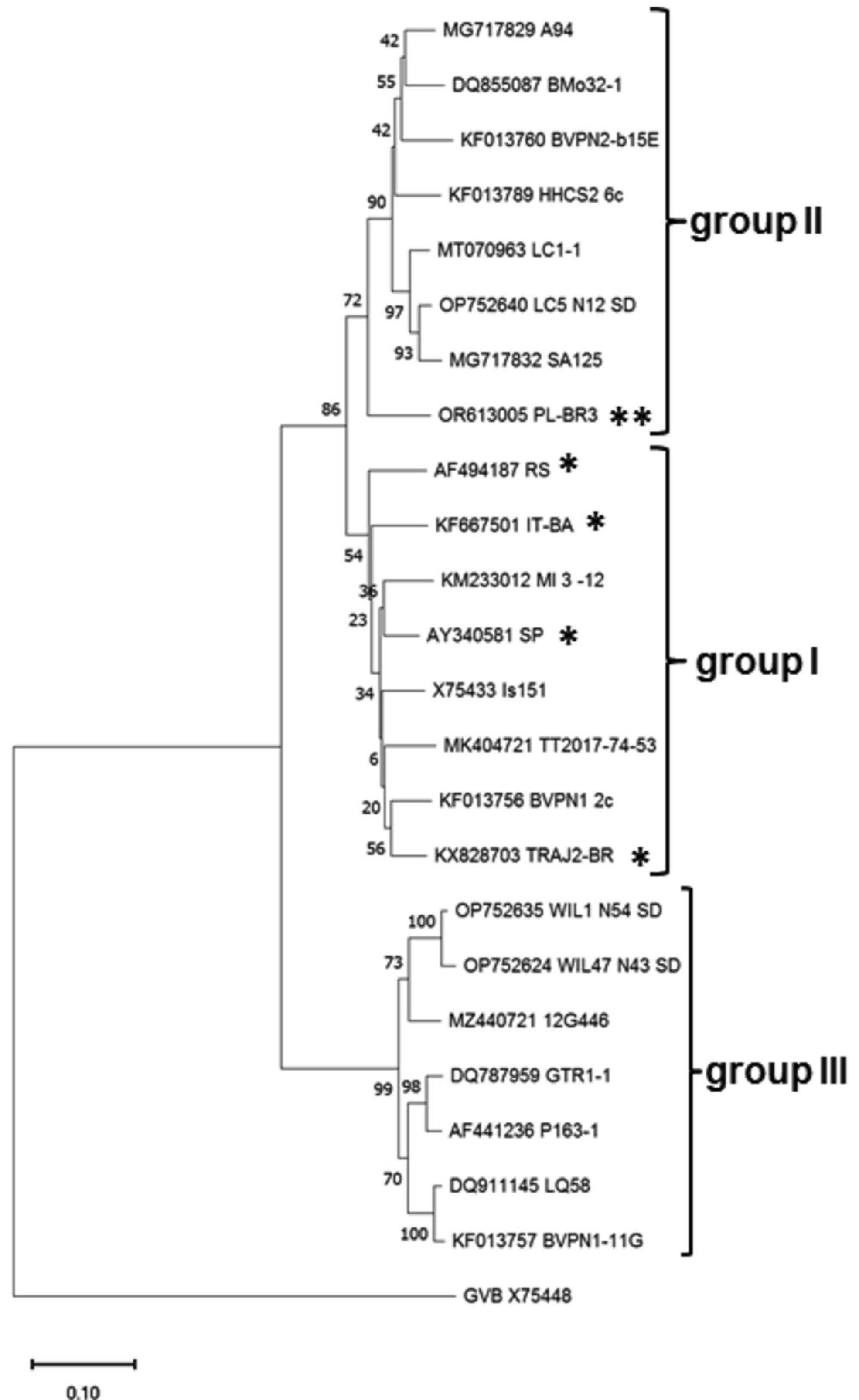

Fig. 4 Phylogenetic analysis of grapevine virus A (GVA) based on complete coat protein nucleotide sequences. The Neighbor-joining (NJ) method with the Kimura 2-parameter model was used. Values at nodes are percent bootstrap support at 1,000 replications. Bootstrap percentage values are shown on the branches. Bars indicate the number of nucleotide substitutions per site. The tree is unrooted. Isolates include PL-BR3 (**) (GVA) from *Ps. longispinus* and four previously obtained Brazilian homologous isolates from grapevines (IT-BA, TRAJ2-BR, SP and RS) (*), in addition to some foreign isolates retrieved from GenBank. The phylogenetic groups indicated by brackets correspond to previously described GVA phylogroups (Wu et al. 2023a). Nucleotide accessions and isolate names in GenBank are mentioned. For the external group of the tree, grapevine virus B (GVB, *Vitivirus*, X75448/NC_003602) and for type isolate, GVA (X75433/NC_003604) were used

grapevines (Ct of infected samples ranged from 19.69 to 34.90 and average of 31.51). RT-qPCR was also used to confirm the presence of grapevine viruses identified by HTS in the original total RNA sample, extracted from mealybugs, where all three viruses were detected with Ct values of 22.01 (GLRaV-2), 20.34 (GLRaV-3) and 20.80 (GVA).

These results validated those obtained by HTS, i.e. they determined presence of these viruses in the sampled mealybugs (pseucococcids) collected in the same plot of infected vines and also in those plants infested by *Ps. longispinus*. Considering the high incidence of viruses in the evaluated vines and the presence of viruses in mealybugs, as demonstrated by HTS and RT-qPCR, it is possible to suppose that *Ps. longispinus* was acting as a vector of GLRaV-3 and GVA, promoting the dissemination of these viruses within the evaluated block of plants, since plants are closer to each other in the sampled plot. This fact is partially corroborated by the high incidence of GLRaV-3 (89.5%) in the evaluated vines and by the increase in the number of plants exhibiting virus-like symptoms in the plot of investigated plants over the years (data not shown).

Determining the presence of viruses in a single sample of total RNA of mealybugs by RT-qPCR resulted in detection of three viruses (GLRaV-2 and -3 and GVA) similarly to the indexing result carried out on 76 plants from the analyzed grapevine plot. This could constitute a comprehensive approach of viral disease monitoring by evaluating insect vectors for viruses presence, thus enabling to generate information about the presence of viruliferous insects in the vineyard, rather than just sampling tissues of supposedly infected plants. Prospecting for viruses infecting grapevines by analyzing the metagenome of a mealybug vector sample collected inside a vineyard can also be an interesting alternative to generate valuable preliminary information on the health status of the vineyard and viral disease surveillance. It could even contribute the fact that mealybugs are slow moving insects, and sample collection could be done in different plants or part of plants (clusters, under the bark of the trunk) throughout the field to capture these insects. This also represents a possibility of improving the monitoring procedures of viral diseases for their management and control in vineyards and vine nurseries for producing high quality propagative material.

Lastly, a highly relevant epidemiological aspect, i.e. the insect vector that would be considered in this assessment as a related part of generation of selection pressure and viral variability, thus supporting and guiding analyzes. As observed, in some cases, there was no high correlation between the incidence level of a given virus in the sampled plot of vines and the possibility of assembly and obtaining the partial or complete genome of the same virus by HTS analysis. GVA and GLRaV-2 were detected in ca. 36% of plants, although it was possible to assembly only the partial genome of GVA (coat protein and RNA-binding protein genes) and the complete genome of GLRaV-2. In general, it was demonstrated that the viruses identified in the mealybug sample were also present in the infected vines in which they were collected.

Acknowledgements This study was supported by Embrapa, project 22.16.04.035.00.00. Authors thank Marcos F. Vanni for technical support and Vera M. Quecini (Embrapa Uva e Vinho, Brazil) for some bioinformatics HTS analyzes.

Author contribution Conceptualization: TVMF and ON; methodology: TVMF and ON; formal analysis and investigation: TVMF; HTS and bioinformatics analyzes: PG, RCT, JMFS and FNS; writing - original draft preparation: TVMF and ON; writing - review and editing: TVMF, PG, RCT, JMFS, FNS and ON; funding acquisition: TVMF; resources: TVMF, ON and FNS; supervision: TVMF.

Data availability All sequencing data obtained in this study were included in the manuscript and/or submitted to the GenBank database under the accession numbers OR502861, OR640977, OR546046 and, OR613005 and BioProject PRJNA1020893.

Declarations

Conflict of interest The authors declare no competing interests.

References

Ahmed AR, Apori SO, Karim AA (2023) Mealybug vectors: a review of their transmission of plant viruses and their management strategies. AIMS Agriculture and Food 8:736–761

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 23 Oct 2023

Basso MF, Fajardo TVM, Saldarelli P (2017) Grapevine virus diseases: economic impact and current advances in viral prospection and management. Revista Brasileira de Fruticultura 39:e-411

Beuning LL, Murphy P, Wu E, Batchelor TA, Morris BAM (1999) Molecular-based approach to the differentiation of mealybug (Hemiptera: Pseudococcidae) species. Journal of Economic Entomology 92:463–472

Bhat AI, Rao GP (2020) Transmission of viruses through mealybugs. Characterization of plant viruses. Springer Protocols Handbooks Humana Press, New York, pp 95–98

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421

Canaguier A, Grimplet J, Di Gaspero G, Scalabrin S, Duchêne E, Choisne N, Mohellibi N, Guichard C, Rombauts S, Le Clainche I, Bérard A, Chauveau A, Bounon R, Rustenholz C, Morgante M, Le Paslier MC, Brunel D, Adam-Blondon AF (2017) A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). Genomics Data 14:56–62

Chen S, Zhou Y, Chen Y, Gu J (2018) Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890

Dubiela CR, Fajardo TVM, Souto ER, Nickel O, Eiras M, Revers LF (2013) Simultaneous detection of Brazilian isolates of grapevine viruses by TaqMan real-time RT-PCR. Tropical Plant Pathology 38:158–165

Fajardo TVM, Bertocchi AA, Nickel O (2020) Determination of the grapevine virome by high-throughput sequencing and grapevine viruses detection in Serra Gaúcha, Brazil. Revista Ceres 67:156–163

Fajardo TVM, Peres CA, Nickel O (2023) Real-time RT-PCR high-resolution melting curve analysis to detect and differentiate

Brazilian variants of grapevine viruses. Ciência e Técnica Vitivinícola 38:188–195
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
Fuchs M (2023) Grapevine virology highlights: 2018-2023. Proceedings of the 20th congress of the international council for the study of virus and virus-like diseases of the grapevine, Thessaloniki, Greece, pp 18–26. Available at https://icvg.org/data/ICVG20Abstracts.pdf. Accessed 5 Mar 2024
Grohs DS, Almança MAK, Fajardo TVM, Halleen F, Miele A (2017) Advances in propagation of grapevine in the world. Revista Brasileira de Fruticultura 39:e-760
Herrbach E, Alliaume A, Prator CA, Daane KM, Cooper ML, Almeida RPP (2017) Vector transmission of grapevine leafroll-associated viruses. In: Meng B, Martelli GP, Golino DA, Fuchs M (eds) Grapevine viruses: molecular biology, diagnostics and management. Springer, Cham, pp 483–503
Jarugula S, Alabi OJ, Martin RR, Naidu RA (2010) Genetic variability of natural populations of Grapevine leafroll-associated virus 2 in Pacific Northwest vineyards. Phytopathology 100:698–707
Kaur N, Hasegawa DK, Ling K-S, Wintermantel WM (2016) Application of genomics for understanding plant virus-insect vector interactions and insect vector control. Phytopathology 106:1213–1222
Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35:1547–1549
Kuniyuki H, Gioria R, Rezende JAM, Willink CG, Novo JPS, Yuki VA (2006) Transmissão experimental do grapevine virus B pela cochonilha *Pseudococcus longispinus* Targioni-Tozzetti (Hemiptera: Pseudococcidae). Summa Phytopathologica 32:151–155
La Notte P, Buzkan N, Choueiri E, Minafra A, Martelli GP (1997) Acquisition and transmission of grapevine virus A by the mealybug Pseudococcus longispinus. Journal of Plant Pathology 79:79–85
Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357–359
Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948
Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W (2016) MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102:3–11
Maliogka VI, Martelli GP, Fuchs M, Katis NI (2015) Control of viruses infecting grapevine. Advances in Virus Research 91:175–227
Mannini F, Digiaro M (2017) The effects of viruses and viral diseases on grapes and wine. In: Meng B, Martelli GP, Golino DA, Fuchs M (eds) Grapevine viruses: Molecular biology, diagnostics and management. Springer, Cham, pp 453–482
Maree HJ, Pirie MD, Oosthuizen K, Bester R, Rees DJG, Burger JT (2015) Phylogenomic analysis reveals deep divergence and recombination in an economically important grapevine virus. PLoS One 10:e0126819
Martinez-Mercado MA, Jesús JLD, Galindo-Sánchez CE, Saavedra-Flores A, Carrillo-Tripp J (2022) Novel viral RNA genomes of the vine mealybug *Planococcus ficus*. Journal of General Virology 103:3
McGreal B, Sandanayaka M, Chooi KM, MacDiarmid R (2019) Development of sensitive molecular assays for the detection of grapevine leafroll-associated virus 3 in an insect vector. Archives of Virology 164:2333–2338
O'Leary NA, Wright MW, Brister JR, Ciufo S et al (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Research 44:D733–D745
Osman F, Rowhani A (2008) Real-time RT-PCR (TaqMan) assays for the detection of viruses associated with Rugose wood complex of grapevine. Journal of Virological Methods 154:69–75
Osman F, Leutenegger C, Golino D, Rowhani A (2007) Real-time RT-PCR (TaqMan) assays for the detection of Grapevine leafroll associated viruses 1–5 and 9. Journal of Virological Methods 141:22–29
Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods 14:417–419
Rott ME, Jelkmann W (2001) Characterization and detection of several filamentous viruses of cherry: adaptation of an alternative cloning method (DOP-PCR) and modification of an RNA extraction protocol. European Journal of Plant Pathology 107:411–420
Rubio L, Galipienso L, Ferriol I (2020) Detection of plant viruses and disease management: Relevance of genetic diversity and evolution. Frontiers in Plant Science 11:1092
Schoelz J, Volenberg D, Adhab M, Fang Z, Klassen V, Spinka C, Al Rwahnih M (2021) A survey of viruses found in grapevine cultivars grown in Missouri. American Journal of Enology and Viticulture 72:73–84
Song Y, Hanner RH, Meng B (2021) Probing into the effects of grapevine leafroll-associated viruses on the physiology, fruit quality and gene expression of grapes. Viruses 13:593
Tsai C-W, Rowhani A, Golino DA, Daane KM, Almeida RPP (2010) Mealybug transmission of grapevine leafroll viruses: An analysis of virus-vector specificity. Phytopathology 100:830–834
Van der Auwera GA, O'Connor BD (2020) Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media, USA
Vončina D, Jagunić M, De Stradis A, Diaz-Lara A, Al Rwahnih M, Šćepanović M, Almeida RPP (2024) New host plant species of grapevine virus A identified with vector-mediated infections. Plant Disease 108:125–130
Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N (2012) LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Research 40:11189–11201
Wistrom CM, Blaisdell GK, Wunderlich LR, Botton M, Almeida RPP, Daane KM (2017) No evidence of transmission of grapevine leafroll-associated viruses by phylloxera (*Daktulosphaira vitifoliae*). European Journal of Plant Pathology 147:937–941
Wu Q, Habili N, Kinoti WM, Tyerman SD, Rinaldo A, Zheng L, Constable FE (2023a) A metagenomic investigation of the viruses associated with Shiraz Disease in Australia. Viruses 15:774
Wu Q, Habili N, Tyerman SD, Rinaldo A, Little A, Constable FE (2023b) First detection of five previously unreported grapevine viruses in Australia. Australasian Plant Disease Notes 18:27
Xiao H, Liu Z, Wang N, Long Q, Cao S, Huang G, Liu W, Peng Y, Riaz S, Walker AM, Gaut BS (2023) Adaptive and maladaptive introgression in grapevine domestication. Proceedings of the National Academy of Sciences, USA 120:e2222041120
Zhou Y, Minio A, Massonnet M, Solares E, Lv Y, Beridze T, Cantu D, Gaut BS (2019) The population genetics of structural variants in grapevine domestication. Nature Plants 5:965–979

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.