P16: *In silico* detection of a defective genomic RNA of *Grapevine leafroll-associated virus 4* strain 5 in High-Throughput Sequencing data

João Marcos Fagundes Silva[1], Rosana Blawid[1], Tatsuya Nagata[1], **Thor Vinícius Martins Fajardo[2*]**

[1]*Universidade de Brasília, Brasília-DF, Brazil,* [2]*Embrapa Uva e Vinho, Bento Gonçalves-RS, Brazil*

[*]*Corresponding author: thor.fajardo@embrapa.br*

INTRODUCTION

Grapevine leafroll disease (GLRD) is one of the most economically important viral diseases affecting grapevines (Maliogka et al., 2015). Most viruses associated with this disease belong to the family *Closteroviridae*, and are designated *Grapevine leafroll-associated virus*es (GLRaVs). The family *Closteroviridae* consists of viruses with a positive-sense ssRNA genome of ~15 kb and rod-shaped particles. GLRaV-1, -3 and -4 belong to the genus *Ampelovirus*, while GLRaV-2 and -7 belong to the genera *Closterovirus* and *Velarivirus*, respectively (Maliogka et al., 2015). These viruses exist within its host as a diverse cloud of related but distinct genomes known as a quasispecies (Lauring and Andino, 2010). The complexity of a quasispecies population is augmented by recombination events, which are also associated with the emergence of novel genotypes and defective genomic RNA molecules (dRNAs) (Bar-Joseph and Mawassi, 2013). In this work, we report the detection of a dRNA from a GLRaV-4 strain 5 in *Vitis vinifera* cv. Trajadura, containing multiple virus infections. This GLRaV-4 strain 5 dRNA is 9,295 bp long and is composed by the polyprotein, RNA-dependent RNA polymerase (RdRp) and a partial truncated divergent coat protein (CPd) ORFs, similar to a class 3 *Citrus tristeza virus* (CTV) dRNA. These dRNAs are composed by a large 5' terminus portion of the genomic RNA, containing the entire replicase complex, and usually a truncated 3' terminus ORF (Bar-Joseph and Mawassi, 2013).

MATERIALS AND METHODS

Seventeen grapevine samples collected from Brazil have been previously subjected to High-throughput sequencing (HTS) at Macrogen (Seoul, South Korea) or Eurofins Genomics (Huntsville, USA). The cDNA libraries were constructed from double-stranded (dsRNA) extracts, and sequencing was performed on the Illumina HiSeq 2000 platform (2 x 100 bp). The GLRaVs infecting these plants have already been described (Fajardo et al., 2017). For each GLRaV positive sample, the sequencing reads were aligned to the complete genome of the corresponding GLRaV with the program BWA (Li and Durbin, 2010). Variant calling was performed with LoFreq (Wilm et al., 2012). To screen for the presence of GLRaV derived dRNAs on these samples, we used the program ViReMa (Routh and Jonhson, 2014). This analysis revealed the presence of a dRNA from a GLRaV-4 strain 5 on a multiple virus-infected *V. vinifera* cv. Trajadura (S18-TRAJ).

RESULTS AND DISCUSSION

Due to vegetative propagation, grapevines often present mixed infections and complex viral pathosystems. By using HTS, we identified a dRNA from a GLRaV-4 strain 5 in *V. vinifera* cv. Trajadura coinfected with GLRaV-3, *Grapevine Red Globe virus* (GRGV), *Grapevine Syrah virus 1* (GSyV-1), *Grapevine rupestris stem pitting-associated virus* (GRSPaV), *Grapevine virus A* (GVA), *Grapevine fleck virus* (GFkV) and *Grapevine rupestris vein feathering virus* (GRVFV) (Fajardo et al., 2017). Recombinant dRNAs of GLRaVs are thought to be generated by template switching mechanisms, and are dependent on the parental sequence for essential viral functions, such as replication, encapsidation or systemic movement (Bar-Joseph and Mawassi, 2013). While retaining the entire replicase complex, the p5, HSP70h, HSP90h, coat protein (CP) and part of the CPd ORFs are absent in the GLRaV-4 strain 5 S18-TRAJ dRNA. Both parental full length GLRaV-4 strain 5 and GLRaV-3 genomes, as well as the other viruses infecting this host, may be providing the function of these proteins for efficient infection of GLRaV-4 strain 5 dRNA. The junction site of the GLRaV-4 strain 5 dRNA is located at nucleotides 8,785 and 13,313 of the genomic RNA. Similar dRNAs, composed by the entire replicase complex and a truncated 3' terminus ORF, have been described for the closterovirus CTV. These dRNAs can be mechanically transmitted to citrus plants and inoculated into *Nicotiana benthamiana* protoplasts, and are possibly self-replicating (Che et al., 2002). A high number of single nucleotide variants (SNVs) was annotated in our analysis (Fig. 1), suggesting a

diverse intrahost population of GLRaV-4 strain 5. However, it is not clear whether these variants are present in the full length or defective genome of GLRaV-4 strain 5. To our knowledge, this is the first report of a dRNA of GLRaV-4 strain 5.

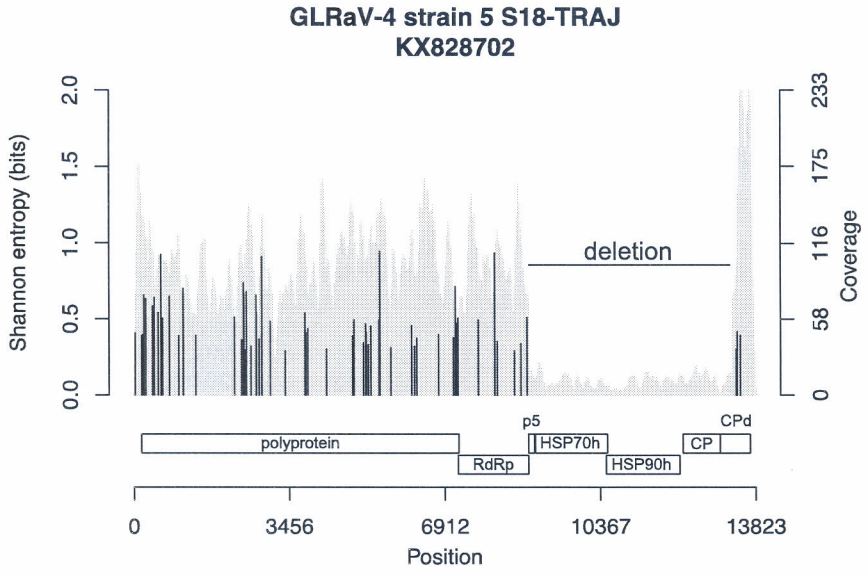

Fig. 1. Shannon entropy (left) of each single nucleotide variant (SNVs) annotated and coverage depth (right) of the full length genome of GLRaV-4 strain 5 (KX828702) on the S18-TRAJ sample, along with a schematic representation of the genome organization of this virus (below), showing the polyprotein, RNA-dependent RNA polymerase (RdRp), p5, HSP70h, HSP90h, coat protein (CP) and diverged coat protein (CPd) ORFs.

ACKNOWLEDGEMENTS

This work was supported by CNPq and EMBRAPA (project 02.13.14.002).

REFERENCES

Bar-Joseph, M., Mawassi, M. 2013. The defective RNAs of Closteroviridae. Front Microbiol., 4: 132.

Che, X., Mawassi, M., Bar-Joseph, M. 2002. A novel class of large and infectious defective RNAs of *Citrus tristeza virus*. Virology, 298(1): 133-145.

Fajardo, T.V.M., Silva, F.N., Eiras, M., Nickel, O. 2017. High-throughput sequencing applied for the identification of viruses infecting grapevines in Brazil and genetic variability analysis. Trop. Plant Pathol., 42: 250-260.

Lauring, A.S., Andino, R. 2010. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog., 6(7): e1001005.

Li, H., Durbin, R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics, 26(5): 589-595.

Maliogka, V.I., Martelli, G.P., Fuchs, M., Katis, N.I. 2015. Control of viruses infecting grapevine. Adv Virus Res., 91: 175-227.

Routh, A., Johnson, J.E. 2014. Discovery of functional genomic motifs in viruses with ViReMa – a Virus Recombination Mapper – for analysis of next-generation sequencing data. Nucleic Acids Res., 42(2): e11.

Wilm, A., Aw, P.P., Bertrand, D., Yeo, G.H., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., Nagarajan, N. 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res., 40(22): 11189-11201.