PARTIAL SEQUENCE ANALYSIS OF A GRAPEVINE LEAFROLL-ASSOCIATED VIRUS 3 ISOLATE FROM SLOVAKIA

M. Glasa and L. Predajna

Institute of Virology, Slovak Academy of Sciences, Dubravská cesta 9, 84505 Bratislava, Slovakia

SUMMARY

Grapevine leafroll-associated virus 3 (GLRaV-3) is one of the agents responsible for leafroll disease of cultivated grapevines. To secure information on the molecular variability of GLRaV-3 isolates from central Europe, a 3,477 nucleotide long genome fragment coding for the end of a 55kDa protein, the complete capsid protein (CP), the minor CP (CPm), a 21kDa protein, and the N-terminal part of the p20A protein was sequenced from an atypical Slovak viral isolate, inducing no leafroll symptoms on the original Vitis vinifera host. In the sequenced ca. 3.5 kb region, the nucleotide identities with previously characterized isolates from different parts of the world ranged from 89.4 (Group 3 isolates) to 99.5% (Group 1 isolates). The Slovak SK04 isolate clustered within the largest Group 1, although in a slightly separated branch. Vegetative propagation by infected cuttings did not lead to changes of the dominant GLRaV-3 sequence, determined from the mother plant, when compared with the sequences obtained from the newly established 5-year-old self-rooted plants, indicating the high genetic stability of the SK04 viral population.

Key words: GLRaV-3, genome, grapevine, molecular diversity.

More than 60 viruses are associated with diseases of grapevines (Martelli and Boudon-Padieu, 2006) among which Grapevine leafroll-associated virus 3 (GLRaV-3), a member of the genus Ampelovirus, family Closteroviridae (Martelli et al., 2002). GLRaV-3 is one of the agents of the leafroll disease in grapevines, typically causing a downward rolling of the leaf blades, accompanied by a reddish or yellowish discoloration of the interveinal tissues, in the red-berried and white-berried varieties, respectively. Both the quantity and quality of the grapes can be adversely affected (Martelli et al., 2011). The virus is transmitted by coccid or pseudococcid mealybug vectors (Hemiptera) in a semi-persistent manner and by propagative materials (Martelli et al., 2011).

The GLRaV-3 genome consists of a single-stranded positive-sense RNA ca. 18,000 nucleotides (nt) in length, organized into 12 open reading frames (ORF) (Jrugula et al., 2010), characterized by significant genetic diversity among isolates (Wang et al., 2011; Sharma et al., 2011). To date, seven GLRaV-3 isolates from the USA, South Africa, and Chile have been completely sequenced (Ling et al., 2004; Engel et al., 2008; Maree et al., 2008; Jooste et al., 2010; Jrugula et al., 2010). European isolates have been incompletely analysed, for only the partial sequence of the RNA-dependent RNA polymerase (RdRp), heat-shock protein 70 homologue (HSP70), and coat protein (CP) genes (Turturo et al., 2005); or of the complete CP gene (Gouveia et al., 2011) have been obtained. In this work, a partial sequence of a Slovak isolate has been determined, targeting the 3’ proximal portion of the genome.

The GLRaV-3 SK04 isolate used in this study was isolated in January 2010 from a 25-year-old self-rooted vine of Vitis vinifera cv. Queen of the vineyards (an ancient white berry table cultivar) grown in a private garden in western Slovakia (GPS: 48°19.001’N, 17°28.918’E). The presence of common grapevine viruses in the plant was tested by DAS-ELISA, using commercial antisera against GLRaV-1 and GLRaV-3, Grapevine fleck virus (GFkV), Arabis mosaic virus (ArMV), and Grapevine fanleaf virus (GFLV) (all from Bioreba, Switzerland). Total RNAs were extracted from cortical scrapings from dormant canes using a NucleoSpin RNA Plant kit (Macherey-Nagel, Germany) following the manufacturer’s instructions. First strand cDNA was synthesized by reverse transcription of total RNAs, using random hexamer primers (Promega, USA) and PrimeScript™ reverse transcriptase (Takara Bio, Japan). Five primer pairs (Table 1) were designed and used to amplify five overlapping fragments encompassing a ca. 3.5 kb portion of the genome, coding for the end of a 55kDa protein, the complete CP, the minor CP (CPm), a 21kDa protein, and the N-terminal part of the p20A protein (nt position 13591-17140, based on the complete sequence EU259806).

PCR was performed using the proofreading TaKaRa Ex Taq™ polymerase (Takara Bio, Japan) and the am-
plified products were purified using the Wizard PCR Prep DNA purification system (Promega, USA). To obtain a master sequence characterizing the viral population present in the plant, purified RT-PCR products were directly sequenced by priming the sequencing reaction with the same oligonucleotides used for PCR. The nucleotide sequence reported herein has been deposited in the GenBank database under accession No. JQ267586.

To estimate the intra-isolate variability, the LR3-13591F/LR3-14576R-amplified PCR product was cloned into the pGEM-T Easy cloning vector (Promega, USA), and 10 randomly chosen clones were sequenced on both strands using universal pUC primers. Sequence analyses were performed using Molecular Evolutionary Genetics Analysis [MEGA v. 5.1; (Tamura et al., 2011)] and DNA Sequence Polymorphism software [DnaSP v. 5 (Librado and Rozas, 2009)]. For comparison, the available GLRaV-3 sequences were retrieved from GenBank.

Isolate SK04 was recovered from a grapevine that did not display typical leafroll symptoms, although the presence of GLRaV-3 was confirmed by DAS-ELISA, and was also infected by GFkV (Glasa et al., 2011) but not by ArMV and GFLV. As focusing on a short genome fragment could bias the analysis (and would be uninformative regarding possible recombination event/s), we have determined the nucleotide (nt) sequence of a 3.5 kb genome fragment of isolate SK04, encompassing three complete ORFs (ORF6, 7, and 8) and part of the adjacent regions.

Sequence analysis revealed a genome organization typical of GLRaV-3 and the strict co-linearity with the complete GLRaV-3 sequences previously obtained. The deduced sequences of CP, CPm, and 21kDa were 314, 478, and 185 amino acid (aa) in length, respectively, with predicted molecular mass of 34.930 (CP), 53.276 (CPm), and 21.355 (p20A). The nt and aa identities of SK04 with the representative isolates of five GLRaV-3 genetic groups (Gouveia et al., 2011) are shown in Table 2.

Estimation of the genetic distance, based on the complete CP sequence of SK04, has shown the closest identity with GLRaV-3 isolates from Portugal (accession No. HQ401017), Brazil (DQ680141) and the USA (GU983863 and HQ130315), reaching 99.7%. On the other hand, the nt identity with the most distant isolates, all from the USA (HQ130307, HQ130326, and HQ130331-32), was only 90.8%.

Phylogenetic analysis using nt sequences of the complete CP gene showed that isolate SK04 and 74 previously characterised GLRaV-3 isolates whose complete CP sequences are available in databases grouped into 5 phylogenetic clusters unambiguously defined with high bootstrap support (Fig. 1, named according to Gouveia et al., 2011). The Slovak SK04 isolate clustered in the largest group (Group 1), although in a slightly separated branch. This phylogenetic study has shown

Table 1. Primers used to amplify a ca. 3.5 kb region of the SK04 genome.

<table>
<thead>
<tr>
<th>Primer1</th>
<th>Sequence 5'→3'</th>
<th>Position2, amplified genome portion</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR3-13591F</td>
<td>TGGAGAAGCGTAGTTATG</td>
<td>13591-14576, 55kDa/CP</td>
</tr>
<tr>
<td>LR3-14576R</td>
<td>ACCATTCAGGGTCGCCGTG</td>
<td></td>
</tr>
<tr>
<td>LR3-14518F</td>
<td>CKGTAAGCGCTGATCTAGC</td>
<td>14518-14955, CP/CPm</td>
</tr>
<tr>
<td>LR3-14995R</td>
<td>GAGAGATCCTGAGACCTC</td>
<td>14921-16215, CPm</td>
</tr>
<tr>
<td>LR3-16123F</td>
<td>GAAGAGTTCACTCGCTAC</td>
<td>16123-16744, CPm/21kDa</td>
</tr>
<tr>
<td>LR3-16744R</td>
<td>CACACCCGCTATGGCCTTT</td>
<td></td>
</tr>
<tr>
<td>LR3-16608F</td>
<td>AGCCGATATAACGTCGGCG</td>
<td></td>
</tr>
<tr>
<td>LR3-17140R</td>
<td>GACGATTTAAGTCCTCCG</td>
<td>16608-17140, 21kDa/p20A</td>
</tr>
</tbody>
</table>

1orientation (F = forward, R= reverse), 2 numbered according to the full-length GLRaV-3 sequence EU259806.

Table 2. Percent identity at the nucleotide and amino acid level (calculated using the Kimura's two parameter method and p-distance model, respectively) between the functional products in the analysed region of SK04, and representative isolates belonging to 5 phylogenetically distinct GLRaV-3 groups.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>CP nt</th>
<th>CP aa</th>
<th>CPm nt</th>
<th>CPm aa</th>
<th>21kDa nt</th>
<th>21kDa aa</th>
<th>Whole ca. 3.5 kb fragment nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>GQ352631 (Group 1)</td>
<td>99.5</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>100</td>
<td>100</td>
<td>99.5</td>
</tr>
<tr>
<td>AF037268 (Group 1)</td>
<td>99.2</td>
<td>98.4</td>
<td>99.1</td>
<td>98.1</td>
<td>99.5</td>
<td>100</td>
<td>99.2</td>
</tr>
<tr>
<td>EU344894 (Group 5)</td>
<td>95</td>
<td>97.1</td>
<td>n.a.</td>
<td>n.a.</td>
<td>93.4</td>
<td>97.3</td>
<td>92.9</td>
</tr>
<tr>
<td>GQ352632 (Group 2)</td>
<td>93.1</td>
<td>95.5</td>
<td>92.3</td>
<td>90.8</td>
<td>94.3</td>
<td>97.3</td>
<td>92.9</td>
</tr>
<tr>
<td>HQ130311 (Group 4)</td>
<td>92.4</td>
<td>97.4</td>
<td>88.6</td>
<td>90.4</td>
<td>91.4</td>
<td>95.1</td>
<td>90</td>
</tr>
<tr>
<td>GQ352633 (Group 3)</td>
<td>91.6</td>
<td>96.2</td>
<td>88.3</td>
<td>89.5</td>
<td>90.7</td>
<td>94.6</td>
<td>89.4</td>
</tr>
</tbody>
</table>
the absence of geographical structuration within GLRaV-3. The presence of isolates from different countries, even from different continents, within each genetic group points to the mixing of the viral population due to international trade of the propagative materials. The intra-isolate diversity (π) of SK04, evaluated from LR3-13591F/LR3-14576R sequences of 10 randomly chosen clones, was 0.00125 (haplotype diversity H=0.778), which is indicative of the homogeneity of the viral population. Similar low values of within-isolate ge-

Fig. 1. Phylogenetic tree of GLRaV-3 isolates, generated from the SK04 sequence determined in this study (highlighted in bold), plus 74 complete CP nucleotide sequences retrieved from GenBank. Isolates are identified by their accession numbers; their geographical origins are indicated in brackets. The scale bar indicates a genetic distance of 0.01. Bootstrap values >70 (1000 bootstrap re-samplings) are indicated on the branches as percentages. The phylogenetic tree was inferred using the neighbour-joining algorithm implemented in MEGA, using Kimura 2 parameter genetic distances. The affiliation of the isolates to 5 GLRaV-3 groups is indicated to the right of the tree.
nestic variability had been observed for the majority of GLRaV-3 isolates analysed by Turturo et al. (2005).

No evidence of a recombination signal was found along the sequenced 3.5 kb region of isolate SK04, as determined by the Recombination Detection Program using the dataset including the previously determined complete GLRaV-3 sequences (RDP3; Martin et al., 2010). Likewise, the phylogenetic analysis separately targeting 3 different genes (CP, CPm, 21kDa) assigned SK04 to Group 1, therefore, it did not reveal any phylogenetic incongruence indicative of possible recombination (data not shown).

The original grapevine plant (from which isolate SK04 was recovered), had served in 2006 as a source of dormant cuttings, from which six new self-rooted plants were obtained and planted in the same locality. In September 2011, these 5-year-old plants tested positive for GLRaV-3 by DAS-ELISA. Thereafter, three leaf petioles were collected from each plants pooled, and total RNA, isolated as described above, was used as template for RT-PCR with primers LR3-13591F/LR3-14576R. The 985 bp fragments (Table 1) amplified from each of the six vines were directly sequenced. Interestingly, multiple alignments revealed no mutations among the six amplicons, and a 100% identity with the original SK04 sequence. Although this comparison was based on a short genome fragment, it suggests that following vegetative propagation from the original grapevine, the dominant viral population had not undergone detectable changes in the newly propagated plants. These data confirm the extremely high sequence stability of RNA viruses, despite their high mutation rates, due to strong purifying selection (Garcia-Arenal et al., 2001).

Vegetative propagation can lead to a genetic bottleneck, due to the possible uneven distribution of viral variants within the plant (D’Urso et al., 2000; Jridi et al., 2006). In our experiment, it was determined that there was an unchanged dominant viral population in all of six plant coming from the vegetative propagation of SK04. This suggests the homogenous repartition of the same viral population in the original vine, and the absence of selection pressure acting on the viral population in the new plants following the 5-year period.

ACKNOWLEDGEMENTS

This work was supported by grant VEGA 2/0030/10 from the Scientific Grant Agency of the Ministry of Education and Slovak Academy of Sciences.

REFERENCES


Martin D.P., Lemey P., Lott M., Moulton V., Posada D., Lefevre P., 2010. RDP3: a flexible and fast computer pro-


Received January 10, 2012
Accepted May 7, 2012