SURVEY AND GENETIC DIVERSITY OF GRAPEVINE LEAFROLL ASSOCIATED VIRUS-2 IN ALGERIA

Arezki Lehad, Ilhem Selmi, Meriem Louanchi, Mouni Aitouada, Naima Mahfoudhi

Laboratoire de Protection des Végétaux, Institut National de la Recherche Agronomique de Tunisie, Rue Hedi Karray, 2049 Ariana, Tunisia.

ABSTRACT

Vineyards in western and center regions of Algeria were surveyed for the Grapevine leafroll-associated virus 2 (GLRaV-2). Analyses by DAS-ELISA and Reverse Transcription Polymerase Chain Reaction (RT-PCR) reveal 15, 8% prevalence. The genetic diversity of the GLRaV-2 population was studied by phylogenetic analyses of the HSP70h gene region of seven samples sequenced in this study and other sequences downloaded from GenBank. Results reveal segregation of the GLRaV-2 population into six distinct groups. An estimation of the ratio of non-synonymous substitutions per non-synonymous substitutions per synonymous site indicated that HSP70h gene evolve under positive selection. Similarity plot constructed with representative sequence from each group confirmed previous results. All Algerian isolates belong to group PN. As far as we know, this is the first characterization of GLRaV-2 isolates from Algeria.

Keywords: GLRaV-2, RT-PCR, genetic diversity, HSP70h gene.

INTRODUCTION

The grapevine leafroll disease represents the most widespread disease of Vitis vinifera worldwide (Martelli, 2014). It may be due to several virus belonging to the family of Closteroviridae which were designated as Grapevine Leafroll-associated Viruses (GLRaVs) and represented by GLRaV-1, -2, -3, -4, and -7 which were recognized as species and recently (Martelli et al., 2012) indicate that GLRaV-5, -6, -9, GLRaV-Pr, GLRaV-De, and GLRaV-Car are strains of GLRaV-4, considered before as distinct species. All these viruses belong to the family Closteroviridae, with GLRaV-2 belonging to the genus Closterovirus, GLRaV-7 belongs to a newly proposed genus Velariviruses and the other GLRaVs to the genus Ampelovirus. GLRaV-2 was described since 1984 and was associated with the grapevine leafroll disease (Gugerli et al., 1984). Later, it was associated to graft-incompatibility (Gref et al., 1995). The transmission vector for GLRaV-2 is unknown, although other members of the genus Closterovirus were transmitted by aphids (Karasev, 2000). GLRaV-2 is known to be transmitted by grafting with infected material. However, some isolates of GLRaV-2 were mechanically transmitted to herbaceous hosts like Nicotiana benthamiana (Goszczyński et al., 1996).

GLRaV-2 presents flexuous and filamentous particles of about 1600 nm length and its genome is a single strand, positive sense RNA with about 16500nt organized into nine ORF encoding at least 11 proteins (Zhu et al., 1998). Up to now, complete genome sequence was obtained for nine isolates (JX559644.1, NC_007448.1, DQ286725.2, AY881628.1, KF220376.1, FJ436234.1, JQ771955.1, AF314061.1 and NC_004724.1).

Several variants of GLRaV-2 have been described and characterized. (Zhu et al., 1998) and (Abou-Ghanem et al., 1998) described the first two variants from the varieties Pinot Noir and Semillon. Two years later, a new variant denoted GLRaV-2-H4 was discovered by (Ghanem-Sabanadzovic et al., 2000) on Vitis rupestris St. George. Further study reveals the presence of virus associated to GLRaV-2 isolated from cv Redglobe and associated to graft incompatibility (Rowhani et al.,
2000). Up to now, six variants were reported, the variant ‘Pinot Noir, 93/955, H4, BD, RG and PV20 (Zhu et al., 1998, Meng et al., 2005, Ghanem-Sabanadzovic et al., 2000, Rowhani et al., 2002, Angelini et al., 2004, Beuve et al., 2007, Bertazzon et al., 2010).

**MATERIAL AND METHODS**

**Virus Source:** The field study and sample collection were conducted in autumn 2010 and 2012 on table grape and wine grape collected in western (Ain Témouchent and Mascara) and central (Algiers, Tizi-Ouzou and Boumerdes) regions of Algeria. A total 584 samples were collected from individual vines from different varieties (30 varieties) including commercials (10 varieties) (445 samples) autochthonous vineyard (two varieties) (110 samples) and autochthonous Grapevine germplasm (18 varieties) (29 samples) of ITAF (Institute Technique de l’Arboriculture Fruitière et de la Vigne). Mature canes were randomly collected, one from each vine and stored at 4°C.

**Virus detection by double antibody sandwich-enzyme-linked immunosorbent assay (DAS-ELISA):** All collected samples were tested by DAS-ELISA (Clark and Adams, 1977), for the presence of GLRaV-2 using specific commercial polyclonal antiserum (Agridia, Bari, Italy). The extracts were obtained by macerating phloem tissues in the PBS-buffer (V:V). Optical density was recorded at 405 nm using an automatic microplate reader (Multiskan Ascent, Labsystems, Waltham, MA, USA). Samples with Optical density readings exceeding or equal to three times that of the healthy samples were considered positive.

**Virus detection by reverse transcription polymerase chain reaction (RT-PCR):** All positive samples were analyzed by RT-PCR. Total nucleic acids were extracted from 0.2g of cortical scraping of dormant cutting cane, which was powdered using liquid nitrogen. The powder was homogenized in 1 ml of grinding buffer (4 M guanidine thiocyanate, 0.2 M sodium acetate, 1 M potassium acetate, 0.025 mM EDTA, 25% PVP-40) mixed with 100 µl NLS 10% and denaturized at 70°C. The solution was centrifuged and the supernatant was recovered. The Total Nucleic Acid was precipitated by silica as described by (Svanella-Dumas et al., 2000) with some modifications. Two step protocol was used for the reverse transcription (RT) and amplification (PCR) of target RNA. Reverse transcription was performed using 1 µl of Moloney murine leukaemia virus reverse transcriptase (M-MLV 200 units/ µl), 4 µl of 5X Fs M-MLV buffer, 2 µl of DTT (0.1 M) and 0.5 µl of dNTPs (10 mM). The mixture was incubated at 39°C for 1h and at 70°C for 10 min. PCR was carried out with primer pair L2 F (5’-TAATTCCGGGTACATCCCACCT-3’) and U2 R (5’-GCCCTCGGCAACTAATGACAG -3’) which encompass 331pb located in the HSP70h gene. The DNA amplifications were carried out in 12,5 µl total reaction volume. 1.25 µl of cDNA were mixed with 11.25 µl of the amplification mixture, consisting of 10 mM each dNTP, 20 µM each primer, 50 mM MgCl2, 5 U/µl TaqDNA polymerase (Promega) and 10x Taq buffer. PCR cycling conditions, in a thermocycler, include denaturation of cDNA at 95°C for 5 min, followed by 35 cycles of 94°C/ 30s, 58°C/45s and 72°C/60 s. To end with, final elongation step was carried at 72°C during 7 min. The PCR products with positive and negative control used were provided from Institut Nationale de la Recherche Agronomique (INRAT) collection and PCR markers of 100pb were analyzed by electrophoresis on 1.5% agarose gel stained with ethidium bromide and visualized using an UV-transilluminator.

**Sequencing and sequence analyses:** Amplified products of 331pb from HSP70h gene region of GLRaV-2 of seven isolates were randomly chosen for sequencing, they were purified using the ExoSAP-IT purification kit. Direct sequencing was performed with the same primers used for RT-PCR on 3730xl DNA analyzer (Applied Biosystems) automated sequencer.

To investigate the different variants of GLRaV-2 population, multiple alignments of nucleotide sequences were performed using CLUSTAL W (Thompson et al., 1994), respectively, with default settings from MEGA6 software (Tamura et al., 2013) and comparison at the nucleotide level for HSP70h gene regions of Algerian isolates and representative sequences of the different phylogenetic groups described up to now was conducted (Table 2).

**Phylogenetic analysis:** Nucleotide sequences of the Algerian isolates and those downloaded from GenBank were aligned and evolutionary relationships among GLRaV-2 sequences inferred using the neighbor joining method (NJ) (Saitou and Nei, 1987) with 1,000 bootstrap replications. All these analyses were conducted in MEGA6 software with GLRaV-3 isolate BR5 (KF417599.1) used as an outgroup.

**Estimation of selection pressure and recombination analysis:** Gene and site-specific selection pressures over the entire alignment of data set for HSP70h were analyzed using the Datamonkey online (http://www.datamonkey.org/).
The ratio of nonsynonymous substitutions per nonsynonymous site (dN) to synonymous substitutions per synonymous site (dS), which is considered as an indicator of natural selection was calculated using two methods for detecting sites under positive selection: single-likelihood ancestor counting (SLAC), random-effects likelihood (REL). Depending on \( dN/dS \) values, the selection pressure was considered negative or purifying \( (dN/dS < 1) \), neutral \( (dN/dS = 1) \), or diversifying or positive \( (dN/dS > 1) \).

Recombinant events analysis of GLRaV-2 population was performed using the GARD algorithm (Kosakovsky et al., 2006) in the remote server Datamonkey (Delport et al., 2010). A similarity plot was constructed with Simplot software (downloaded from http://sray.med.som.jhmi.edu/SCRoftware/simplot/) using the isolate OR1 from group PN as reference sequence with a multiple sequence alignment of full genome sequences provided from the different representative groups (H4, BD, RG, and 93/955) except for PV20 group constructed with MEGA6 (Tamura et al., 2013). Currently, there is no full-length sequence for groups PV20.

**RESULTS**

**Prevalence of GLRaV-2:** DAS-ELISA test reveals that the GLRaV-2 presents prevalence of 15.8% in Algeria. The peak of prevalence is observed in Gros noir des Beni Abbas with 29%. The Alicante Bouchet presents 24.6%, King's Rubi (22.2%), Dattier de Beyrouth (18.7%), Valensi (14.5%), Chaucho Blanc and Muscat d'Alexandrie (10%), Cinsault (7.5%), Carignan (9.7%), and autochthonous collection (6.9%). The GLRaV-2 is absent in Merseguerra and Chasselas. The autochthones germplasm show to be free of GLRaV-2.

**Sequencing and Sequence Analysis:** The 331nt sequence fragment of the HSP70h Gene obtained by RT-PCR was cloned and sequenced for seven isolates of GLRaV-2 representing the first sequenced isolates provided from Algeria. Nucleotide sequences obtained were submitted to GeneBank (Table 2).

<table>
<thead>
<tr>
<th>Seq&gt;</th>
<th>ALG7</th>
<th>ALG9</th>
<th>ALG19</th>
<th>ALG22</th>
<th>ALG53</th>
<th>ALG55</th>
<th>ALG93</th>
<th>PV20</th>
<th>OR1</th>
<th>GRSLaV</th>
<th>BD</th>
<th>PN</th>
<th>93/955</th>
<th>GLRaV-2-SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALG7</td>
<td>ID</td>
<td>98%</td>
<td>99%</td>
<td>97%</td>
<td>98%</td>
<td>99%</td>
<td>99%</td>
<td>74%</td>
<td>98%</td>
<td>71%</td>
<td>75%</td>
<td>99%</td>
<td>86%</td>
<td>85%</td>
</tr>
<tr>
<td>ALG9</td>
<td>98%</td>
<td>ID</td>
<td>99%</td>
<td>98%</td>
<td>99%</td>
<td>98%</td>
<td>99%</td>
<td>74%</td>
<td>99%</td>
<td>72%</td>
<td>74%</td>
<td>99%</td>
<td>86%</td>
<td>85%</td>
</tr>
<tr>
<td>ALG19</td>
<td>99%</td>
<td>99%</td>
<td>ID</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>75%</td>
<td>99%</td>
<td>72%</td>
<td>75%</td>
<td>100%</td>
<td>87%</td>
<td>86%</td>
</tr>
<tr>
<td>ALG22</td>
<td>97%</td>
<td>98%</td>
<td>99%</td>
<td>99%</td>
<td>98%</td>
<td>98%</td>
<td>99%</td>
<td>74%</td>
<td>99%</td>
<td>72%</td>
<td>74%</td>
<td>99%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>ALG53</td>
<td>98%</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>98%</td>
<td>98%</td>
<td>99%</td>
<td>74%</td>
<td>99%</td>
<td>72%</td>
<td>75%</td>
<td>99%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>ALG55</td>
<td>99%</td>
<td>98%</td>
<td>99%</td>
<td>99%</td>
<td>98%</td>
<td>98%</td>
<td>99%</td>
<td>75%</td>
<td>99%</td>
<td>71%</td>
<td>75%</td>
<td>99%</td>
<td>87%</td>
<td>85%</td>
</tr>
<tr>
<td>ALG93</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>74%</td>
<td>100%</td>
<td>72%</td>
<td>75%</td>
<td>100%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>PV20</td>
<td>74%</td>
<td>74%</td>
<td>75%</td>
<td>74%</td>
<td>74%</td>
<td>75%</td>
<td>74%</td>
<td>1D</td>
<td>74%</td>
<td>72%</td>
<td>74%</td>
<td>100%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>OR1</td>
<td>98%</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>74%</td>
<td>74%</td>
<td>72%</td>
<td>74%</td>
<td>100%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>GRSLaV</td>
<td>71%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>72%</td>
<td>81%</td>
<td>81%</td>
<td>72%</td>
<td>71%</td>
<td>72%</td>
</tr>
<tr>
<td>BD</td>
<td>75%</td>
<td>74%</td>
<td>75%</td>
<td>74%</td>
<td>75%</td>
<td>75%</td>
<td>75%</td>
<td>74%</td>
<td>74%</td>
<td>81%</td>
<td>75%</td>
<td>74%</td>
<td>77%</td>
<td>77%</td>
</tr>
<tr>
<td>PN</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>99%</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>74%</td>
<td>100%</td>
<td>72%</td>
<td>75%</td>
<td>99%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>93/955</td>
<td>86%</td>
<td>86%</td>
<td>87%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>71%</td>
<td>74%</td>
<td>86%</td>
<td>86%</td>
<td>85%</td>
</tr>
<tr>
<td>GLRaV-2-SG</td>
<td>85%</td>
<td>85%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>86%</td>
<td>72%</td>
<td>77%</td>
<td>86%</td>
<td>85%</td>
<td>85%</td>
</tr>
</tbody>
</table>

Table 1. Percent identity between Algerian isolates and isolates from each group.

Table 2. Range of HSP70h nucleotide sequence identities within and between groups of GLRaV-2.
In order to access phylogenetic relationship between these isolates, comparison at nucleotide level was conducted with representative sequences from each group described by (Jarugula et al., 2010). The Algerian isolates present 97% to 100% identities. Comparison with the different groups reveal 98% to 100% identity with the isolates OR1 and PN, 74% to 75% identity with the isolates PV20 and BD, 71-72% identity with GRSLaV, 86-87% identity with isolates 93/955, and 85-86% identity with isolate GLRaV-2-SG.

The genetic distance, transformed to the percent identity between and within groups was calculated using MEGA6 software with default parameters reveals that the group PN presents 99% identity. The group H4 presents 95% identity, 96%, 98% and 100% identity for RG, BD and 93/955. The between percent identity reveal that the group PN presents 70-86% with all the other groups, 72-78% for group BD, 71-86%, 69-78%, 69-74% and 71-86% respectively for H4, RG, PV20 and 93/955 (Table 2).

Phylogenetic Analysis: Phylogenetic tree constructed using the neighbor joining method (NJ) implemented in MEGA6 software reveals that GLRaV-2 population cluster into six groups. Each of these Lineages were assigned to a reference isolate as described by previous study to maintain a standardized nomenclature of GLRaV-2 sequence variant groups. The groups PN and H4 represent the most important group of this population. The RG group contains isolates provided from the variety Red Glob obtained in California. The PV20, BD and 93/955 groups were less represented. Analyses reveal that all Algerian isolates collected from different regions, vineyards and varieties belong to the group PN. Further prospections in local varieties were needed to get a great understanding of genetic variation of GLRaV-2 in Algeria (Figure 1).

Similarity plot comparison of representative isolates from each group was conducted. Results reveal that the group PN is distant to the other groups and the groups H4 and 93/955 are more and less related, the same result was obtained for BD and RG groups which were found to be closely related (Figure 2). Currently, there are no full-length sequences from representative isolates of groups PV20.

Selection Pressure: Identifying evolutionary pressure represent a great deal because past environments that exerted these pressures can be different from present ones and these pressures are not unidirectional but rather the result of complex networks (Moury and Simon, 2011).
In order to study selection pressure occurring in a GLRaV-2 population, an estimation of the ratio of nonsynonymous substitutions per nonsynonymous site (dN) to synonymous substitutions per synonymous site (dS) was performed. Indeed, dN/dS ratio indicates the magnitude of selective pressure on each codon. When the ratio dN/dS>1 on a certain codon it indicates that the site evolve under positive selection and when dN/dS<1 in another site it indicates that this site evolve under negative selection. The mean value of dN/dS ratio obtained with SLAC and REL methods is upper than unity suggesting the occurrence of positive selection. Indeed, SLAC reveal mean ratio of dN/dS>1 (1.81) with Codon site under positive selection and non-negative selected codon site were found. Approximately, same results were obtained by REL method with mean ratio dN/dS=3.71 with 6 codon under positive selection and 6 codon under negative selection. Plotting of SLAC established by dN-dS for each codon, indicating the negative or positive selection, reveal that the number of codon site under positive selection is more important comforting previous result obtained in this study (Figure 3). Detection of evidence of putative recombination events in the HSP70h sequences of the GLRaV-2 population was performed with the genetic algorithms for recombination detection (GARD) available on the Datamonkey webserver. GARD analysis reveals no recombinant events within a HSP70h gene in the GLRaV-2 population.

Figure 2. Similarity plot constructed from a multiple alignment of six full-length sequences representing five well-defined variant groups of GLRaV-2 using SimPlot 3.2.

Figure 3. Plotting of single-likelihood ancestor counting (SLAC) based on 90 codon site from HSP70h gene of GLRaV-2 sequences.
DISCUSSION
Survey and genetic diversity of GLRaV-2 collected from different vineyards in the center and western regions of Algeria was established. This study represents the first comprehensive work on the prevalence and genetic diversity conducted in major grape-growing region of Algeria including an autochthonous germplasm collection with first sequencing of Algerian isolates. Phylogenetic analyses with isolates provided from different regions was conducted to make an update of molecular variability of GLRaV-2.

Grapevine leafroll associated virus 2 is present in all prospected regions. This may be due to the use of infected root stock which in majority provide from North America. Indeed, the vector transmission of GLRaV-2 was not described up to now (Martelli, 2014) and some grape nursery were found to present leafroll symptoms, allowing the assumption of infection by GLRaV-2 which is found almost associated with GLRaV-3 and propagation by infected shoot.

Lekikot et al. (2012) reported also the presence of GLRaV-2 in Algeria with less prevalence. The fact of this large propagation of GLRaV-2 may be due to the use of infected plant propagating materials. Indeed, observations made on some grape nursery during the prospection reveals the presence of leafroll symptoms. (Lekikot, 2012) reported that native varieties were more infected than the imported ones which may be due to large movements of infected material. Further studies on grape nursery were needed in order to understand the propagation of this virus in Algeria. Few studies were dedicated to the genetic diversity of GLRaV-2. Based on CP gene analyses, (Bertazzon et al., 2010) reported five clades. In the same year, (Jarugula et al., 2010) reported six lineages.

In this study, phylogenetic analyses of HSP70h gene performed reveal the presence of six lineages confirming previous studies. Indeed, comparison at the nucleotide level reveals that the described groups present less than 86% similarity between them and the similarity within groups reveal more than 95% identity. Furthermore, similarity plot reveals five distinct variants except for variant PV20 due to the absence of full length sequence. Our results were in concordance with those obtained by (Jarugula et al., 2010) which reported an interlineage sequence identities between PN, 93/955, and H4 lineages of 83 to 86% and between 'PV20', BD, and RG lineages were 68 to 80%.

Phylogenetic tree reveals that all Algerians isolates provided from a distant region belong to the Group PN containing in general American isolates. Tree reveals the presence of isolates provided from a distant region and cluster in the same clade. The same result was obtained by several authors who rejected the hypothesis that the phylogeny of GLRaV-2 population depend on geographical origin. Thus, (Jarugula et al., 2010) reported that isolates provided from distant vineyards show an important similarity and cluster in the same lineage. This observation may be due to the large movement of variants among root stock and cultivars. Indeed, a large survey on authoctonous and wild populations is needed for a better understanding of genetic variation of a GLRaV-2 population. However, Results show that the groups PN and H4 were the most important group constituted in the majority of North American isolates, this result may be due to the large sequences providing from this region.

The high genetic stability of viruses can be attributed to negative or purifying selection to maintain the functional integrity of the viral genome. However, several genera were reported as evolving under positive selection (the genera Aureusvirus, Carmovirus, Dianthovirus, Necrovirus and Tombuvirus). Thus (Jarugula et al., 2010) reported that HSP70h evolve under negative selection with dN/dS < 1. In our study, we found that this gene evolve under positive selection. Indeed, we obtained the mean ratio dN/dS greater than 1 (dN/dS= 1.81 for SLAC and dN/dS= 3.71). Results reveal also that some sites were under neutral selection and others under negative selection. Indeed, some sites in the HSP70h were exposed to mutation which were maintained in the population and other were under stability. To understand whether codon may be predisposed to mutation. The positive selection obtained may be explained by the changes on amino acid in some site of HSP70h gene, indicating specificity of each variant. It is interesting to know the effect of these mutations on plant virus interaction. Indeed, these mutations may explain the differences in virulence between the different variants. (Bertazzon et al., 2010) reported that isolates from group BD were less virulent and were unable to induce graft incompatibility and it rarely caused leafroll symptom. In contrast, the RG isolates appeared to be more virulent.

The absence of recombination events may be explained by the fact that each vine may be infected by only one
variant limiting genome exchange between distant variants. Indeed, results reveal the presence of only one variant in Algeria and the absence of natural vector of GLRaV-2 limit the large distribution of variants. Further studies of variant distribution are necessary.

REFERENCES
Moury, B. and V. Simon. 2011. dN/dS-based methods detect positive selection linked to trade-offs between different fitness traits in the coat protein of potato virus Y. Molecular biology and evolution. 28: 2707-17.

