Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar
http://dx.doi.org/10.5424/sjar/2013111-3069

Spanish Journal of Agricultural Research 2013 11(1), 199-203
ISSN: 1695-971-X
eISSN: 2171-9292

Short communication. Molecular analysis of the genomic RNAs 1 and 2 of the first *Arabis mosaic virus* isolate detected in Spanish grapevines

I. López-Fabuel[1], A. Olmos[1], A. Bassler[2], L. Dupuis-Maguiraga[2], L. B. Torres[1, 2], E. Bertolini[1] and T. Wetzel[2]*

[1] *Centro de Protección Vegetal y Biotecnología. Instituto Valenciano de Investigaciones Agrarias (IVIA). Ctra. Moncada a Náquera, km 5. 46113 Moncada (Valencia), Spain*
[2] *RLP Agroscience. AlPlanta- Institute for Plant Research. Breitenweg 71. 67435 Neustadt an der Weinstrasse, Germany*

Abstract

The *Arabis mosaic virus* (ArMV) is one of the causative agent of the grapevine fanleaf disease, one of the most widespread and damaging viral diseases of grapevine. Recently, the ArMV has been detected in Spanish vineyards, and its determination and molecular characterization was undertaken. To this aim, the nucleotide sequence of the genomic RNAs 1 and 2 of the first isolate of ArMV infecting grapevine detected in Spain (ArMV-DU13) has been determined. The ArMV-DU13 genomic sequences were compared to the corresponding sequences of other isolates of ArMV, or nepoviruses. The most divergent genes among ArMV isolates were the X1 and VPg genes on the RNA 1, and the 2A gene on the RNA 2, with identity levels at the amino acid level of 78% (X1 and VPg) or 69% (2A) between the most distant isolates. Interestingly, the VPg genes were identical between the two grapevine isolates ArMV-Du13 and –NW, suggesting a possible implication of the host. The phylogenetic analysis of the RNA 2 showed that the Spanish isolate was close to *Grapevine fanleaf virus* isolates. The analysis of the full length RNA 2 suggests a recombination event between ArMV-DU13 and GFLV-GHu isolates between nucleotides 54 and 586 in the ArMV-DU13 isolate. Altogether, these results confirm the high variability between isolates of ArMV, and will be helpful to design more appropriate and reliable molecular diagnostic techniques for the control of this emerging virus in Spain.

Additional key words: ArMV-DU13; molecular characterization; phylogenetic analysis.

Arabis mosaic virus (ArMV), a member of plant virus genus *Nepovirus* of the family *Secoviridae* is transmitted by the nematode vector *Xiphinema diversicaudatum*, and has a wide plant host range (Sanfaçon *et al*., 2009). This virus is a causal agent of the grapevine fanleaf disease, one of the most damaging virus diseases affecting grapevine. The infection causes the decline or even the destruction of the vinestocks, severely affecting the productivity and the longevity of the plants, and is responsible of important economic losses (Mekuria *et al*., 2009). Nepoviruses have two positive sense, single-stranded genomic RNAs, RNA 1 and RNA 2, which are polyadenylated at their 3' end, and have a covalently attached small genome-linked viral protein (VPg) at their 5' end. To date, the complete nucleotide sequences of the genomic RNAs of ArMV have been described for only few isolates, the grapevine isolate NW (Wetzel *et al*., 2001 and 2004), the privet isolate Lv (Dupuis *et al*., 2008), and two isolates from barley (Genbank accession numbers GQ369526 -GQ369530). Additional complete sequences of RNAs 2 of ArMV have been reported for other isolates (Loudes *et al*., 1995; Imura *et al*., 2008; Vigne *et al*., 2008). In this study we describe the determination and molecular characterization of RNA 1 and RNA 2 of the Spanish

* Corresponding author: thierry.wetzel@agroscience.rlp.de
Received: 04-05-12. Accepted: 31-01-13.

Abbreviations used: ArMV (*Arabis mosaic virus*); BRV (*Blackcurrant reversion virus*); GCMV (*Grapevine chrome mosaic virus*); GFLV (*Grapevine fanleaf virus*); MP (movement protein); NTB (nucleotide-binding protein); ORF (open reading frame); PVX (*Potato virus X*); RDP (recombination detection program); VPg (genome-linked viral protein).

grapevine isolate ArMV-DU13, recently described as being mechanically transmitted from rooted cuttings onto *Chenopodium amaranticolor*, and producing systemic and symptomless infections (Abelleria *et al*., 2010).

To determine the full sequence of the viral genomic RNAs, double stranded (ds)RNAs were isolated as previously described (Moreno *et al*., 1990) from *C. amaranticolor* (provided by Dr. FJ Legorburu, NEIKER-Basque Institute for Agricultural Research and Development, Vitoria-Gasteiz, Spain). The nucleotide sequencing of RNA1 and RNA2 was carried out by an overlapping PCR fragments strategy, using in the initial steps primers described by Wetzel *et al*. (2001). Remaining gaps were amplified using sequence-specific primers, which were designed from subsequently obtained sequences. The 3' and 5' ends were determined by standard terminal desoxynucleotide transferase RACE using 5'/3' RACE Kit 2nd Generation (Roche). Each nucleotide was sequenced between 3 to 6 times, from independent reactions. Few ambiguities among nucleotide sequences were found and in those cases the consensus nucleotide was selected. All ambiguities were silent mutations except two, one in position 2,198 (E for G) and the other in position 2,656 (V for I). The nucleotide sequences of the RNA1 and RNA2 of ArMV DU13 were assembled and deposited at NCBI GenBank under Acc. No. JQ975057 and HQ834962 respectively.

The complete sequence of the ArMV-DU13 RNA 1 was 7336 nt long excluding the poly(A) tail. A single large open reading frame (ORF) was found encoding a 2285 amino acids polypeptide (MW 252499 = 252K). This putative ORF was preceded by a 230 nt 5' non-coding region, and followed by a 251 nt 3' non-coding region. The complete sequence of ArMV-DU13 RNA 2 was 3816 nt long excluding the poly(A) tail. A single large ORF was found encoding a 1111 amino acids polypeptide (MW 122307 = 122 K). This putative ORF was preceded by a 291 nt 5' non-coding region, and followed by a 192 nt 3' non-coding region.

A comparison between ArMV-DU13 and other ArMV isolates at the amino acid level for the coding sequences and at the nucleotide level for the non-coding regions is shown in Table 1. The 5' non-coding region of the RNA 2 of the DU13 isolate had a 28 nt insertion located 23 nt before the ATG start codon of the ORF, similar to the one in the NW isolate, when compared to the other isolates. In addition, in the 5' non-coding region, two nucleotide motifs (5'-GAGUUUAAGAAACUC-3') and (5'-TCCGTTAAGAGCGGA-3'), able to form stem-loops structures, were found repeated twice before the insertion. This differed from that observed for the nepoviruses *Grapevine deformation virus* and *Grapevine fanleaf virus* (GFLV) (Wetzel *et al*., 2001; Ghanem-Sabanadzovic *et al*., 2005), in which only the first motif was found repeated three times. The significance of these insertions and/or deletions is however unknown. For the *Grapevine chrome mosaic virus* (GCMV), the 5' non-coding region of the RNA-2 was shown to trigger a necrotic response on three *Nicotiana* species (Fernández *et al*., 1999). The experiments however, conducted in an heterologous system (PVX vector), showed that the stem and loop structures were dispensable for the induction of necrosis. On the other hand, the RNA2 5' non-coding region of the *Blackcurrant reversion virus* (BRV), which was found to efficiently mediate *in vivo* translation through an internal ribosomal entry site mechanism, contained little secondary structure, harbouring only one predicted stem-loop at the 5' end (Karetnikov & Lehto, 2007).

The comparisons between the coding regions of different isolates of ArMV (Table 1) revealed that the nucleotide binding protein (NTB) of the RNA1, and the movement protein (MP) of the RNA 2 were the most conserved genes on the ArMV genome. On the other hand, the X1 and VPg of the RNA1, and the 2A of the RNA 2 were the most variable genes on the ArMV genome, confirming previous reports (Wetzel *et al*., 2001; Ghanem-Sabanadzovic *et al*., 2005). Interestingly, while only 78% of identity was found between the VPg of ArMV-DU13 and –Lv, 100% identity was found between the VPg of the two ArMV isolates from grapevine (NW and DU13). It is unclear if this is coincidental, or if it reveals a specific host-pathogen interaction feature, like adaptation to its host for example (although the isolates were collected from different grapevine varieties —Tempranillo for the DU13 isolate, Pinot gris for the NW isolate— and geographical locations). The comparison of the sequences corresponding to the proteolytic cleavage sites revealed that the amino acids between which the cleavages putatively occurred were conserved for all the isolates. However, differences were found in the amino acid sequences upstream of the X2-NTB cleavage site between the different isolates. On the other hand, the sequences upstream the Protease-Polymerase cleavage site were the most conserved between the different ArMV isolates. The cleavage site between X2 and NTB of ArMV-NW was previously

Table 1. Sequence comparisons between the different isolates of ArMV. The ArMV-DU13 sequences were used as reference for pairwise comparisons, which are indicated in % of identity. Comparisons were done at the amino acid level for the coding sequences and at the nucleotidic level for the non-coding regions. The lengths of the different sequences used for comparison are shown in braquets. The Genbank accession numbers of the different sequences were: EU617326 (ArMV-Lv 1); AY303786 (ArMV-NW 1); GQ369526-8 (ArMV-barley 1,2,3); HQ834962 (ArMV-DU13 2); EU617327 (ArMV-Lv 2); AY017339 (ArMV-NW 2); X81815 (ArMV-L 2); X81814 (ArMV-U 2); AB279739 (ArMV-Bu 2); AB279741 (ArMV-Ly 2); AB279740 (ArMV-Na 2); EF426853 (ArMV-Ta 2), GQ369529-30 (ArMV-barley 1,2)

Isolates of ArMV	ArMV-DU13 RNA1							
	5'nc (230)	**X1** (415)	**X2** (190)	**NTB** (611)	**VPg** (23)	**Pro** (220)	**Pol** (825)	**3'nc** (251)
Lv 1	80 (229)	78 (417)	87 (190)	92 (611)	78 (23)	91 (220)	86 (824)	85 (247)
NW 1	87 (227)	82 (414)	93 (190)	95 (612)	100 (23)	94 (220)	93 (825)	89 (252)
Ba(1) 1	81 (228)	79 (417)	86 (190)	96 (611)	78 (23)	91 (220)	88 (824)	74 (298)
Ba(2) 1	81 (229)	80 (417)	85 (190)	94 (611)	73 (23)	90 (220)	88 (824)	74 (297)
Ba(3) 1	80 (228)	78 (417)	85 (190)	95 (611)	78 (23)	90 (220)	88 (824)	75 (298)

Isolates of ArMV	ArMV-DU13 RNA2				
	5'nc (291)	**2A** (259)	**MP** (346)	**CP** (505)	**3'nc** (192)
Lv 2	69 (261)	71 (267)	95 (346)	93 (505)	84 (194)
NW 2	70 (295)	88 (259)	96 (346)	94 (505)	86 (192)
L2	71 (260)	66 (233)	96 (346)	90 (505)	84 (196)
U2	72 (260)	68 (280)	93 (346)	91 (505)	84 (196)
Bu 2	71 (260)	63 (232)	95 (346)	91 (505)	83 (194)
Ly 2	70 (260)	63 (232)	95 (346)	92 (505)	84 (193)
Na 2	71 (260)	62 (233)	96 (346)	91 (505)	84 (194)
Ta 2	69 (260)	90 (259)	96 (346)	92 (505)	83 (187)
Ba(4) 2	73 (260)	72 (267)	96 (346)	91 (505)	85 (194)
Ba(5) 2	72 (260)	72 (267)	96 (346)	92 (505)	86 (195)

shown to be inefficient in *in vitro* experiments (Wetzel *et al*., 2008), suggesting that the efficiency of release of the mature viral proteins might have a regulatory function somewhere in the viral infection process. It could be postulated that the differences observed between the sequences around the proteolytic cleavage sites of the different isolates could result in faster or more efficient release of viral proteins. On the other hand, it is interesting to note that the most conserved cleavage sites are those of the protease and the polymerase genes, for which deficient cleavages would most likely be deleterious for the virus. Additional sequences of ArMV isolates from grapevine and other hosts, together with infectious clones of ArMV, would be needed to clarify these questions.

Due to the low number of complete RNA 1 sequences available in the databases, RNA 2 was used for phylogenetic analysis. The full length RNA 2 sequences were analysed by PhyML algorithm implemented into Geneious Pro 5.4.6 package (Biomatters, New Zealand) applying Kimura two parameter nucleotide substitution model (Kimura, 1980) and 1000 bootstrap. It showed that ArMV DU13 isolate was closest to ArMV-NW and –Ta (Fig. 1a). On another hand, the analysis of the 5' non-coding region of the RNA-2 placed ArMV-DU13 isolate halfway between the ArMV isolates (except ArMV-Bu) and the Ghu isolate of GFLV (not shown). Consequently, a recombination analysis was performed using the full length RNA 2 of near isolates (ArMV-Lv, ArMV-Ta and GFLV-Ghu) using the RDP3 package (Martin *et al*., 2010). The analysis showed that the sequences of the isolates ArMV-DU13 and GFLV-Ghu present statistical evidence of recombination in one region (Fig. 1b), with breakpoints located on the ArMV-DU13 genome at nt 54 (5'UTR region) and nt 586 (2A region) respectively, and the potential major parent ArMV-Lv isolate and the potential minor parent GFLV-Ghu isolate. The same

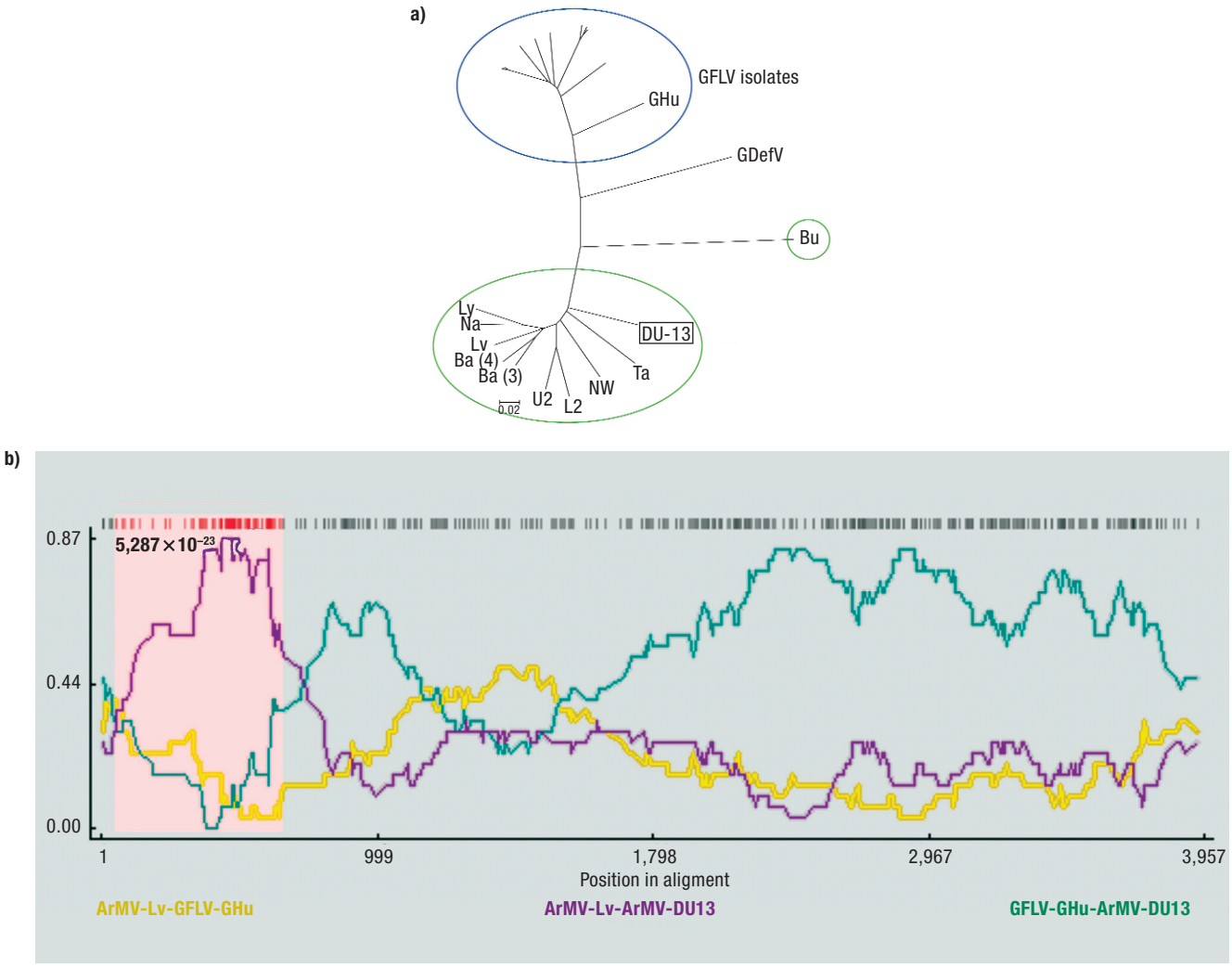

Figure 1. Sequence analysis of *Arabis mosaic virus* (ArMV) isolates. a) Phylogenetic relationship based on the nucleotide sequence of the full-length RNA 2 between ArMV DU13 and 22 of the most related nepoviruses. b) RDP analysis of ArMV-DU13, ArMV-Ta, ArMV-Lv and GFLV-GHu. A recombination event between ArMV-DU13 and GFLV-GHu is shown, with breakpoints located in ArMV-DU13 genome in 54 nt (position 58 in alignment) and 586 nt (position 637 in alignment) respectively and the potential major parent ArMV-Lv isolate (86.8%) and the potential minor parent GFLV-GHu isolate (89.2%).

result was obtained using different algorithms (P-Val = 5.287×10^{-23} for RDP, P-Val = 5.794×10^{-20} with GENECONV, P-Val = 4.122×10^{-24} with BootScan, P-Val = 4.443×10^{-15} with MaxChi, P-Val = 5.376×10^{-16} with Chimaera, P-Val = 1.027×10^{-24} with SiScan, and P-Val = 3.838×10^{-22} with 3Seq). The origin of the recombination, which could have occurred during a simultaneous infection of the same host, was similar to the recombination event reported between ArMV-Ta and GFLV-Ghu (Vigne *et al*., 2008).

Altogether, the information described in this study, which includes the molecular characterization of the first ArMV Spanish isolate, confirms the high variability within this species. This information will be helpful to design more appropriate and reliable molecular diagnostic techniques for the control of this emerging virus in Spain.

Aknowledgements

This work was supported by grants from the German Academic Exchange Service (DAAD, Projekt Nr 50749365), and by grants from the Spanish Ministry of Science and Education (AIB2010DE-00063) and by INIA project RTA2011-00067.

References

Abelleira A, Mansilla JP, Padilla V, Hita I, Cabaleiro C, Bertolini E, Olmos A, Legorburu FJ, 2010. First report of *Arabis mosaic virus* on grapevine in Spain. Plant Dis 94: 635.

Dupuis L, Cobanov P, Bassler A, Krczal G, Wetzel T, 2008. Complete genome sequences of a virulent isolate of *Arabis mosaic virus* from privet (*Ligustrum vulgare*). Arch Virol 153: 1611-1613.

Fernandez I, Candresse T, Le Gall O, Dunez J, 1999. The 5' noncoding region of grapevine chrome mosaic nepovirus RNA-2 triggers a necrotic response on three *Nicotiana* spp. Mol Plant Microbe Interact 12: 337-344.

Ghanem-Sabanadzovic NA, Sabanadzovic S, Digiaro M, Martelli GP, 2005. Complete nucleotide sequence of the RNA-2 of *Grapevine deformation* and *Grapevine Anatolian ringspot viruses.* Virus Genes 30: 335-345.

Imura Y, Oka H, Kimata K, Nasu M, Nakahama K, Maeda T, 2008. Comparisons of complete RNA-2 sequences, pathological and serological features among three Japanese isolates of *Arabis mosaic virus.* Virus Genes 37: 333-341.

Karetnikov A, Lehto K, 2007. The RNA2 59 leader of *Blackcurrant reversion virus* mediates efficient *in vivo* translation through an internal ribosomal entry site mechanism. J Gen Virol 371: 292-308.

Kimura M, 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.

Loudes AM, Ritzenthaler C, Pinck M, Serghini MA, Pinck L, 1995. The 119 kDa and 124 kDa polyproteins of *Arabis mosaic nepovirus* (isolate S) are encoded by two distinct RNA2 species. J Gen Virol 76: 899-906.

Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P, 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462-2463.

Mekuria TA, Gutha LR, Martin RR, Naidu RA, 2009. Genome diversity and intra- and interspecies recombination events in *Grapevine fanleaf virus*. Phytopathology 99: 1394-1402.

Moreno P, Guerri J, Muñoz N, 1990. Identification of Spanish strains of *Citrus tristeza virus* (CTV) by analysis of double-stranded RNAs (dsRNA). Phytopathology 80: 477-482.

Sanfaçon H, Wellink J, Le Gall O, Karasev A, van der Vlugt R, Wetzel T, 2009. *Secoviridae*: a proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus *Torradovirus*. Arch Virol 154: 899–907.

Vigne E, Marmonier A, Fuchs M, 2008. Multiple interspecies recombination events within RNA2 of *Grapevine fanleaf virus* and *Arabis mosaic virus.* Arch Virol 153: 1771-1776.

Wetzel T, Meunier L, Jaeger U, Reustle GM, Krczal G, 2001. Complete nucleotide sequences of the RNAs 2 of German isolates of *Grapevine fanleaf* and *Arabis mosaic nepoviruses.* Virus Res 75: 139-145.

Wetzel T, Beck A, Wegener U, Krczal G, 2004. Complete nucleotide sequence of the RNA 1 of a grapevine isolate of *Arabis mosaic virus*. Arch Virol 149: 989-995.

Wetzel T, Chisholm J, Bassler A, Sanfacon H, 2008. Characterization of proteinase cleavage sites in the N-terminal region of the RNA1-encoded polyprotein from *Arabis mosaic virus* (subgroup A nepovirus). Virology 375: 159-169.