

Complete nucleotide sequence of *Alfalfa mosaic virus* isolated from alfalfa (*Medicago sativa* L.) in Argentina

Verónica Trucco · Soledad de Breuil ·
Nicolás Bejerman · Sergio Lenardon ·
Fabián Giolitti

Received: 12 November 2013 / Accepted: 25 January 2014 / Published online: 8 February 2014
© Springer Science+Business Media New York 2014

Abstract The complete nucleotide sequence of an *Alfalfa mosaic virus* (AMV) isolate infecting alfalfa (*Medicago sativa* L.) in Argentina, AMV-Arg, was determined. The virus genome has the typical organization described for AMV, and comprises 3,643, 2,593, and 2,038 nucleotides for RNA1, 2 and 3, respectively. The whole genome sequence and each encoding region were compared with those of other four isolates that have been completely sequenced from China, Italy, Spain and USA. The nucleotide identity percentages ranged from 95.9 to 99.1 % for the three RNAs and from 93.7 to 99 % for the protein 1 (P1), protein 2 (P2), movement protein and coat protein (CP) encoding regions, whereas the amino acid identity percentages of these proteins ranged from 93.4 to 99.5 %, the lowest value corresponding to P2. CP sequences of AMV-Arg were compared with those of other 25 available isolates, and the phylogenetic analysis based on the CP gene was carried out. The highest percentage of nucleotide

sequence identity of the CP gene was 98.3 % with a Chinese isolate and 98.6 % at the amino acid level with four isolates, two from Italy, one from Brazil and the remaining one from China. The phylogenetic analysis showed that AMV-Arg is closely related to subgroup I of AMV isolates. To our knowledge, this is the first report of a complete nucleotide sequence of AMV from South America and the first worldwide report of complete nucleotide sequence of AMV isolated from alfalfa as natural host.

Keywords *Alfalfa mosaic virus* · Complete genome · Phylogenetic relationships · Alfalfa

Alfalfa mosaic virus (AMV) is the type member of the genus *Alfamovirus* in the *Bromoviridae* family [1]. Its genome consists of three single-stranded RNA molecules (RNA1 to RNA3) of plus-sense polarity, which are encapsidated into B, M and Tb components, respectively [2]. RNA1 encodes the protein 1 (P1), which contains an N-terminal methyltransferase-like domain and C-terminal helicase-like domain; RNA2 encodes the protein 2 (P2), which contains a polymerase-like domain [3]; RNA3 is bicistronic and encodes the movement protein (MP) and the coat protein (CP). CP is translated from a subgenomic messenger (RNA4), which is synthesized during the replication of RNA3 [4]. AMV was first reported in alfalfa (*Medicago sativa* L.) from the USA by Weimer [5] in 1931. It is a worldwide distributed virus with a very wide host range [6]. AMV is transmitted by aphids in the non-persistent manner and also by seeds and pollen of different plant species, including alfalfa [6, 7]. Phylogenetic analysis of CP genes has revealed that AMV isolates can be clustered in two subgroups (I and II) according to their geographic origin [8].

V. Trucco · S. de Breuil · N. Bejerman · S. Lenardon ·
F. Giolitti (✉)
Instituto de Patología Vegetal (IPAVE), Centro de
Investigaciones Agropecuarias (CIAP), Instituto Nacional de
Tecnología Agropecuaria (INTA), Camino 60 cuadras Km. 5.5,
X5020ICA Córdoba, Argentina
e-mail: giolitti.fabian@inta.gob.ar

S. Lenardon
e-mail: lenardon.sergio@inta.gob.ar

S. de Breuil
Consejo Nacional de Investigaciones Científicas y Técnicas
(CONICET), Buenos Aires, Argentina

S. Lenardon
Cátedra de Fitopatología, Facultad de Agronomía y Veterinaria,
Universidad Nacional de Río Cuarto, 5800 Río Cuarto, Córdoba,
Argentina

To date, no sequence information about AMV has been reported from Argentina and although a large number of partial sequences of AMV are available in GenBank, only four complete genomes of AMV (from USA, China, Italy and Spain) have been sequenced. Interestingly, none of these sequences were obtained from alfalfa [8–13], (Zhiyou Du and Jishuang Chen: pers. comm., 2012, Bol John: pers. comm., 2013). In Argentina, alfalfa is a primary forage crop and an essential component of the animal production chain. In 2010, a rhabdovirus infecting this crop was associated with shortened internodes (bushy appearance), leaf puckering and vein enation symptoms [14]. Later on, disease plants which exhibited shortening of internodes, chlorosis at the margins and ribs of the leaflet and varying-sized vein enations on its abaxial surfaces were diagnosed as co-infected by rhabdovirus and AMV (Lenardon Sergio: pers. comm., 2013). The infected plants produce significant losses and reduce the useful economic life of the alfalfa [15]. A similar disease complex was identified in Saudi Arabia which was due to a double infection of AMV and Alfalfa enation virus (*Nucleorhabdovirus*) [16]. The symptoms of AMV on alfalfa range from imperceptible to severe, depending on the AMV strain and alfalfa genotype, with the typical being chlorosis, mosaics, ringspots leaf blades malformations [17].

Here we report the complete nucleotide sequence of an Argentine AMV isolate (AMV-Arg) from alfalfa and compare it with the complete sequences available for other isolates. We also analyze the sequence identities and the phylogenetic relationships of the CP gene of between the AMV isolate from Argentina and compared it with that of other geographically diverse isolates. This information will be useful for future studies addressing disease management.

Samples of alfalfa plants showing the symptoms described were collected from the rural area of Manfredi (Córdoba

province), pooled and the viral particles were enriched according to Massah et al. [18]. Electron microscope observations from the enriched preparation revealed the presence of viral particles resembling those of AMV and a rhabdovirus, and also some spherical particles of a possible virus not yet identified. Total RNA, from the enriched preparation, was extracted with RNeasy Plant Minikit (QIAGEN GmbH, Hilden, Germany) and its sequences were determined by pyrosequencing with 454-ROCHE technology in INDEAR (Rosario, Santa Fe, Argentina). The raw data were processed using the pipeline ngs_backbone 1.4.0 [19] to remove adaptor barcode and low-quality regions. The cleaned sequences were de novo assembled with the software package iAssembler v1.30 [20] and the identity of individual contigs was analyzed using BLASTn and BLASTx. Then, the cleaned reads were mapped against the most related AMV genome (NCBI accession FN667965/6/7) using the software ngs_backbone 1.4.0. The whole-genome consensus sequence of the AMV-Arg isolate was determined with the SAMtools mpileup command and the obtained bam files were visualized and analyzed with the Tablet software [21, 22]. A total of 188,342 reads with an average length of 391pb were mapped to generate the whole genome of AMV-Arg; 126,734 reads corresponding to RNA1, 42,336 to RNA2, and 19,272 to RNA3, with coverages of 13,602, 6,384 and 3,697, respectively. The percentages of nucleotide (nt) and amino acid (aa) sequence identity of the AMV-Arg isolate with available AMV sequences were calculated by ClustalW [23] and implemented in the BioEdit 7.0.9.0 software [24]. The phylogenetic analysis of CP gene was carried out using neighbor-joining (NJ) and maximum likelihood (ML) methods, and ClustalX2, MEGA 5.0 and PhyML 3.0 programs.

The complete nt sequence of AMV-Arg comprised 3,643 nt for RNA1, 2,593 nt for RNA2, and 2,038 nt for

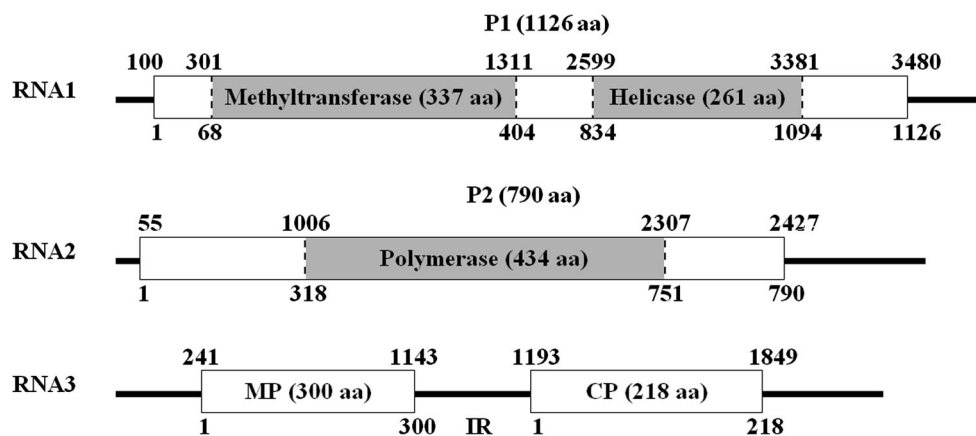


Fig. 1 Schematic representation of the genome organization of AMV-Arg. **Bold lines** represent the RNAs; **boxes** represent the ORFs. The methyltransferase-, helicase- and polymerase-like domains are highlighted and the number of amino acids that compose them and

each protein (Replicase P1, Polymerase P2, MP and CP) are shown in parentheses. The positions where each protein begins and ends are indicated above (nucleotide) and below (amino acid) the **boxes**. **IR** intergenic region

Table 1 Comparison of each encoding region and the whole nucleotide sequence of AMV isolated from Argentina with those of other four isolates from China, Italy, Spain and USA

AMV isolate origin	RNA1		P1 gene					
	nt	nt%	nt	nt%	aa%			
Argentina	3,643	–	3,381	–	–			
China	3,643	99.1	3,381	99.0	99.5			
Italy	3,643	98.8	3,381	98.9	98.8			
Spain	3,643	97.0	3,381	96.9	98.5			
USA	3,644	97.7	3,381	97.8	98.4			
AMV isolate origin	RNA2		P2 gene					
	nt	nt%	nt	nt%	aa%			
Argentina	2,593	–	2,373	–	–			
China	2,595	95.9	2,373	95.6	94.8			
Italy	2,593	98.1	2,373	98.1	98.8			
Spain	2,594	97.0	2,373	96.8	98.1			
USA	2,593	96.1	2,373	95.9	93.4			
AMV isolate origin	RNA3		MP gene			CP gene		
	nt	nt%	nt	nt%	aa%	nt	nt%	aa%
Argentina	2,038	–	903	–	–	657	–	–
China	2,041	97.9	903	97.3	99.0	657	98.3	98.6
Italy	2,038	98.0	903	97.5	97.3	657	97.9	98.6
Spain	2,037	96.0	903	94.8	96.7	666	93.7	94.6
USA	2,141	96.7	903	97.7	97.0	666	96.7	95.9

GenBank accession and hosts for four AMV isolates: China: HQ316635 to 37 (host: potato); Italy: FN667965 to 67 (host: lavender); Spain: FR715040 to 42 (host: cape honeysuckle); USA: L00163; X01572; K03542 (host: clover). The length (nt) of the three RNAs and each encoding region for all AMV isolates are indicated. The highest (bold) and lowest (italics) identity percentages are indicated

RNA3 (Accession numbers: KC881008 to 10, respectively) with a structure similar to those previously described for this virus [4] (Fig. 1). RNA1 encodes a P1 of 1126 aa, RNA2 encodes a P2 of 790 aa, and RNA3 encodes the MP and CP of 300 aa and 218 aa, respectively. The complete genome sequence and each encoding region of AMV isolated were compared with those of four isolates from China, Italy, Spain and USA, which have been completely sequenced. AMV-Arg isolate exhibited nt sequence identity of between 97.0 and 99.1 % for RNA1, 95.9–98.1 % for RNA2, and 96.0–98.0 % for RNA3 (Table 1). Identity percentages of nt for P1, P2, MP and CP genes ranged between 93.7 and 99.0 %, and the deduced aa sequence identities ranged from 93.4 to 99.5 % (Table 1).

Furthermore, we studied the identity percentages of CP nt and aa sequences and the phylogenetic relationships of CP nt sequence of AMV-Arg isolate with those of other 25 isolates available in GenBank. The comparison of CP nt sequences ranged between 77.8 and 98.3 % of identity with isolates of subgroup I and between 92.4 and 94.1 % with those of subgroup II, showing the highest value with the Chinese isolate (HQ316637) (data not shown). The CP

deduced aa sequence showed identity values from 95.4 to 98.6 % with isolates of subgroup I and from 91.4 to 94.5 % with those of subgroup II, sharing the highest percentage with four isolates, two from Italy (Y09110 and FN667967), one from Brazil (FJ858265) and the other one from China (HQ316637) (data not shown), which were isolated from tomato, lavender, alfalfa and potato, respectively. The topology of the phylogenetic trees of CP gene obtained with the NJ (Fig. 2) and ML (data not shown) methods indicates that the AMV-Arg isolate has the closest evolutionary relationships with those belonging to subgroup I of AMV [8], which includes all AMV CP genes isolated from alfalfa. As shown in Fig. 2, 73 % of AMV isolates were clustered in subgroup I, whereas the remaining 27 % were clustered in subgroup II, indicating that subgroup I is more heterogeneous than subgroup II, not only in host diversity but also in geographic origin. In addition, we analyzed the aa composition of CP and found the aa phenylalanine, glycine, glutamine and glutamic acid at positions 67, 84, 176 and 214, respectively, which is consistent with those aa that characterize the AMV subgroup I [8]. These results suggest that the differences in geographic origin and host

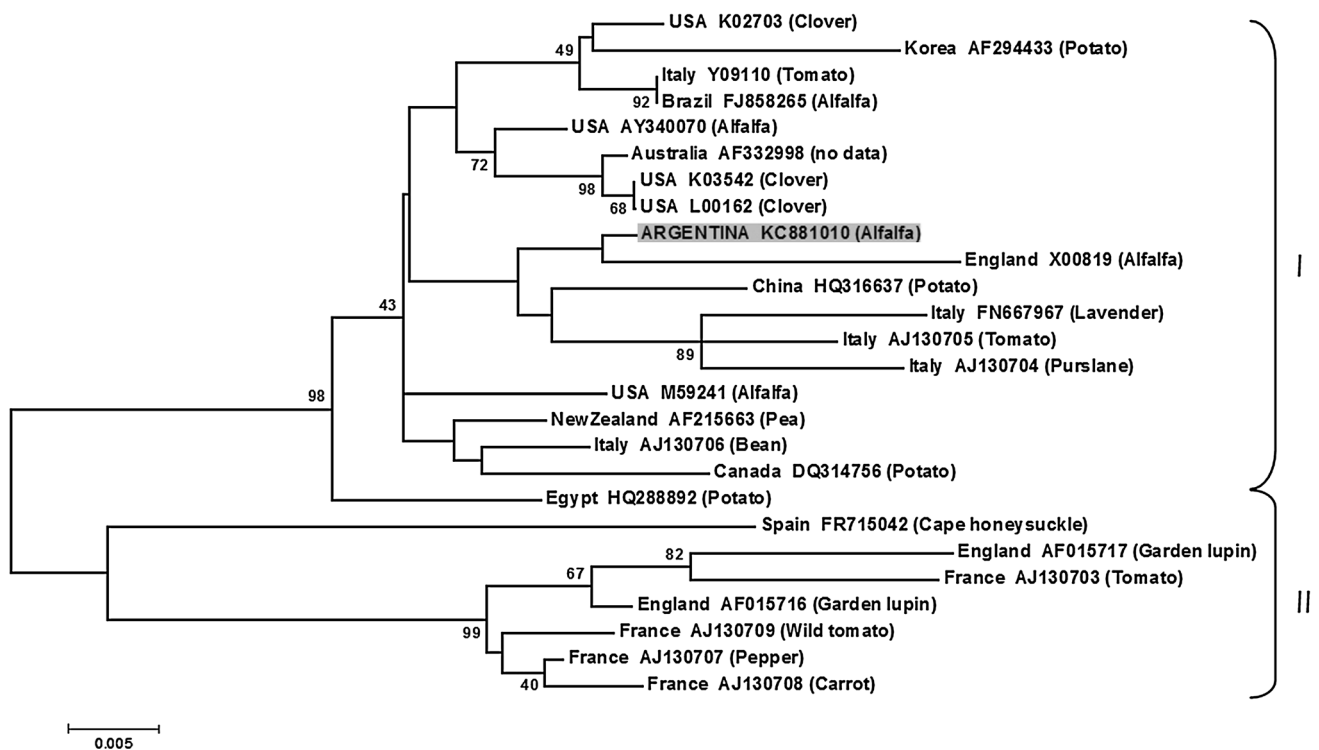


Fig. 2 Phylogenetic relationship based on CP gene of the AMV-Arg isolate with other 25 AMV CP sequences. Bootstrap values are shown as percentages of 1,000 replications (values below 40 % are not

shown). The *scale bar* indicates pairwise nucleotide distance. The numbers I and II correspond to the two subgroups of AMV. Origin, access number and host are indicated for each CP nt sequence

are not related to variability in the CP sequence, which is, in fact, highly conserved.

The AMV-Arg isolate belongs to Subgroup I of AMV and is highly similar to other isolates of this virus that are completely sequenced. To our knowledge, this is the first report of a complete nucleotide sequence of AMV from South America and the first worldwide report of a complete nucleotide sequence of AMV isolated from alfalfa as natural host.

References

- ICTV: International Committee on Taxonomy of Viruses (2012), <http://ictvonline.org/virusTaxonomy.asp>. Accessed 6 Nov 2013
- E.M.J. Jaspars, in *Molecular Plant Virology*, ed. by J.W. Davies (CRC Press, New York, 1985), p. 155
- J.F. Bol, *J. Gen. Virol.* **80**, 1089 (1999)
- C.H. Smit, E.M.J. Jaspars, *Virology* **117**, 271 (1982)
- J.L. Weimer, *Phytopathology* **21**, 122 (1931)
- E.M.J. Jasper, L. Bos, Alfalfa mosaic virus. (Association of Applied Biologists Description of Plant Viruses N° 229, 1980), <http://www.dpvweb.net/dpv/showdpv.php?dpvno=229>. Accessed 29 April 2013
- M.R. McLaughlin, Fungal pathogens of aphids, in *Aphid Plant Interaction: Populations to Molecules*, ed. by D.C. Peters, J.A. Webster, C.S. Chloubers (Oklahoma Agricultural Experiment Station, Stillwater, 1991), p. 318
- G. Parrella, C. Lanave, G. Marchoux, M.M. Finetti Sialer, A. Di Franco, D. Gallitelli, *Arch. Virol.* **145**, 2659 (2000)
- B.J. Cornelissen, F.T. Brederode, R.J. Moormann, J.F. Bol, *Nucleic Acids Res.* **11**(5), 1253 (1983)
- B.J. Cornelissen, F.T. Brederode, G.H. Veeneman, J.H. van Boom, J.F. Bol, *Nucleic Acids Res.* **11**(10), 3019 (1983)
- K. Langereis, M.A. Mugnier, B.J. Cornelissen, L. Pinck, J.F. Bol, *Virology* **154**(2), 409 (1986)
- G. Parrella, N. Acanfora, M.G. Bellardi, *Plant Dis.* **94**(7), 924 (2010)
- G. Parrella, N. Acanfora, A.F. Orílio, J. Navas-Castillo, *Arch. Virol.* **156**(6), 1049 (2011)
- N. Bejerman, C. Nome, F. Giolitti et al., *Plant Dis.* **95**, 771 (2011)
- S. Lenardon, J. Pérez Fernández, D. Basigalup, Descubren una nueva enfermedad en alfalfa. (INTA Informa, 2010), <http://intainforma.inta.gov.ar/?p=2693>. Accessed 9 Sep 2013
- A.A. Cook, A.C. Wilton, *FAO Plant Prot. Bull.* **32**(4), 139 (1984)
- J.H. Graham, K.W. Kreitlow, L.R. Faulkner, Disease, in *Alfalfa Science and Technology*, ed. by C.H. Hanson (American Society of Agronomy, Madison, 1972), pp. 497–526
- A. Massah, K. Izadpanah, A.R. Afsharifar, S. Winter, *Arch. Virol.* **153**, 1041 (2008)
- J.M. Blanca, L. Pascual, P. Ziarsolo, F. Nuez, J. Cañizares, *BMC Genomics* **12**, 285 (2011)
- Y. Zheng, L. Zhao, J. Gao, Z. Fei, *BMC Bioinform.* **12**, 453 (2011)
- H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G.T. Marth, G.R. Abecasis, R. Durbin, *Bioinformatics* **25**(16), 2078 (2009)
- I. Milne, G. Stephen, M. Bayer, P.J.A. Cock, L. Pritchard, L. Cardle, P.D. Shaw, D. Marshall, *Brief Bioinform.* **14**(2), 193 (2013)
- J.D. Thompson, D.G. Higgins, T.J. Gibson, *Nucleic Acids Res.* **22**, 4673 (1994)
- T.A. Hall, *Nucl. Acid. Symp. Ser.* **41**, 95 (1999)