Arch Virol (2016) 161:2029–2032
DOI 10.1007/s00705-016-2854-3

ANNOTATED SEQUENCE RECORD

Complete genome sequence of a new enamovirus from Argentina infecting alfalfa plants showing dwarfism symptoms

Nicolás Bejerman[1,2] · Fabián Giolitti[1] · Verónica Trucco[1] · Soledad de Breuil[1,2] · Ralf G. Dietzgen[3] · Sergio Lenardon[1]

Received: 7 January 2016 / Accepted: 30 March 2016 / Published online: 11 April 2016

Abstract Alfalfa dwarf disease, probably caused by synergistic interactions of mixed virus infections, is a major and emergent disease that threatens alfalfa production in Argentina. Deep sequencing of diseased alfalfa plant samples from the central region of Argentina resulted in the identification of a new virus genome resembling enamoviruses in sequence and genome structure. Phylogenetic analysis suggests that it is a new member of the genus *Enamovirus*, family *Luteoviridae*. The virus is tentatively named "alfalfa enamovirus 1" (AEV-1). The availability of the AEV-1 genome sequence will make it possible to assess the genetic variability of this virus and to construct an infectious clone to investigate its role in alfalfa dwarfism disease.

In Argentina, alfalfa (*Medicago sativa* L.) is a primary forage crop and a major feed component in dairy and beef cattle production systems. In 2010, a rhabdovirus, named alfalfa dwarf virus (ADV), was found associated with alfalfa plants showing symptoms of shortened internodes (bushy appearance), leaf puckering and varying-sized vein enations on abaxial leaf surfaces [1]. The plants that showed these distinct symptoms were diagnosed as being co-infected by ADV and alfalfa mosaic virus (AMV), which led to significant yield losses and reduced the useful economic life of the crop [2]. Between 2010 and 2015, alfalfa dwarf disease had a prevalence of over 70 % in several growing regions of Argentina, and preliminary evaluations showed yield reductions of up to 30 % (S. Lenardon, personal communication). Deep sequencing of small RNAs of alfalfa plants collected in the central region of Argentina showing dwarfism symptoms revealed the presence not only of AMV and ADV but also of a virus related to pea enation mosaic virus 1 (PeMV-1), the type member of the genus *Enamovirus*, family *Luteoviridae* [3].

In this work, we describe the complete genome sequence of this new enamovirus associated with alfalfa dwarf disease in Argentina, which was tentatively named "alfalfa enamovirus 1" (AEV-1). Five samples of symptomatic alfalfa plants showing dwarfism were collected from the same field, located in Manfredi (central region of Argentina), and pooled, and total RNA was extracted using TRIzol Reagent (Life Technologies) according to manufacturer's instructions. This total RNA was sent to Fasteris Life Sciences SA (Plan-les-Ouates, Switzerland), where the bands located between 21 and 30 bp were excised, purified, processed and sequenced on an Illumina HiSeq 2000. Raw data were processed using the pipeline ngs_backbone 1.4.0 [4] to remove adaptor barcode and low-quality regions. The cleaned reads were assembled *de novo* using the software package Velvet v0.6.04 [5], and the identity of individual contigs was analyzed using BLASTn and BLASTx. Two assembled sequences (contigs) of 3,024 bp and 1,489 bp were obtained (Fig. 1a). The follow-up reverse

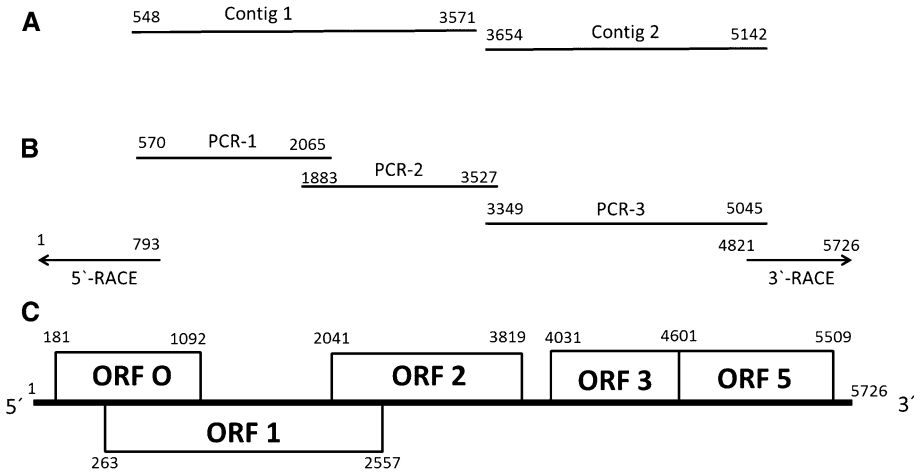
Electronic supplementary material The online version of this article (doi:10.1007/s00705-016-2854-3) contains supplementary material, which is available to authorized users.

✉ Nicolás Bejerman
nicobejerman@gmail.com

[1] IPAVE-CIAP-INTA, Av. 11 de Septiembre 4755, Córdoba 5020, Argentina

[2] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290 (1425 FBQ) CABA, Buenos Aires, Argentina

[3] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD 4072, Australia



D

	ORF O	ORF 1	ORF 1-2	ORF 3	ORF 5
PeMV-1-WSG	75.2/66.3	77.5/71.4	80.3/79.3	75.3/82.7	69.3/67.9
PeMV-1-ID	76.9/69.0	77.3/71.6	80.1/78.5	75.1/82.7	70.6/68.5
CVEV	41.5/37.4	43.6/39.6	49.6/50.8	47.2/44.2	48.5/45.3

Fig. 1 Diagrammatic representation of the contigs from *de novo* assembly of NGS data (**a**), the products of RT-PCR and RACE-PCR (**b**), and the putative genome organization of an alfalfa dwarfism-associated virus (AEV-1). Positions are marked at the extremities (**c**). Nucleotide (nt)/amino acid (aa) sequence identities (%) of AEV-1 ORFs to those of PeMV-1 and CVEV (**d**)

transcription PCR (RT-PCR), rapid amplification of cDNA ends (RACE), cloning, and Sanger sequencing confirmed the sequences of both contigs (Fig. 1b), closed the gap between the two contigs, and resulted in assembly of a full genome sequence (GenBank accession number KU297983) (Fig. 1c).

The complete sequence of AEV-1 consists of 5,726 nucleotides (nt). Five open reading frames (ORFs) were predicted (Fig. 1c), and the genome structure resembled that of PeMV-1 [6]. The first ORF, ORF 0, consists of 909 nt encoding a putative P0 of 303 amino acids (aa) with a calculated molecular mass of 33.7 kDa. In enamoviruses, P0 has been shown to function as an RNA silencing suppressor [7]. The second ORF, ORF1, which contains 2,292 nt, is predicted to be expressed by a ribosomal leaky scanning mechanism for a protein P1 (764 aa, 85 kDa). The third ORF, ORF2, which is translated by a −1 ribosomal frameshift from ORF 1 [8, 9], overlaps this ORF at its 5́ end and is predicted to produce an ORF1-ORF2 fusion protein, P1-P2 (1185 aa, 152.7 kDa). P1 and P1-P2 have a putative involvement in virus replication, while P1 is considered a serine-like protease, and the frameshift region (P2) of the P1-P2 protein is thought to contain the core domains of the viral RNA-dependent RNA polymerase (RdRP) [6]. A serine-protease-like domain (peptidase S39, pfam02122, P1 residues 322-526; E-value, 9.26e−59) in P1, and an RdRP domain (RdRP-4, pfam02123) resembling those from members of the genera *Luteovirus*, *Rotavirus* and *Totivirus* in P2 (P2 residues 152-542; E-value, 1.39e−76) were found when these protein sequences were analyzed. The fourth ORF, ORF3, consists of 567 nt, encoding a putative protein P3 (189 aa, 21.3 kDa). The fifth ORF, ORF5, is a putative in-frame readthrough product of ORF3 encoding a fusion protein P3-P5 (491 aa, 54.6 kDa). P3 is the coat protein (CP), whereas the CP readthrough extension (P5) of P3-P5 is thought to be an aphid transmission subunit of the virus [6]. When these protein sequences were analyzed, a luteovirid CP domain (Luteo_coat, pfam00894) in P3 (P3 residues 52-188; E-value, 3.87e−67) and a polerovirus readthrough protein domain (PLRV_ORF5, pfam01690) [8, 9] in P5 (P5 residues 25-277; E-value, 1.21e−94) were identified. AEV-1 has three untranslated regions (UTR): a 5′-UTR of 180 nt, a 3′-UTR of 217 nt, and an intergenic UTR of 211 bp between ORFs 2 and 3. Furthermore, a VPg bound at the 5′ end has been described in enamoviruses [8, 9]. These data indicate that AEV-1 represents the genome sequence of a virus that should be taxonomically classified in the family *Luteoviridae*. Viruses in the family *Luteoviridae* contain a single-stranded positive-sense RNA genome and are classified into three genera, *Enamovirus*, *Luteovirus* and *Polerovirus*. Unlike poleroviruses, enamoviruses do not encode a putative P4 movement protein, and luteoviruses lack a P0 gene [3]. Therefore, AEV-1 appears to be an enamovirus because it does not encode a P4 protein. The AEV-1 ORFs were compared to the predicted ORFs of the two complete PEMV-1 enamovirus genome sequences available in GenBank (NC_003629 and HM439775) and those of citrus vein enation virus (CVEV), a tentative enamovirus [10], in order to determine nucleotide and

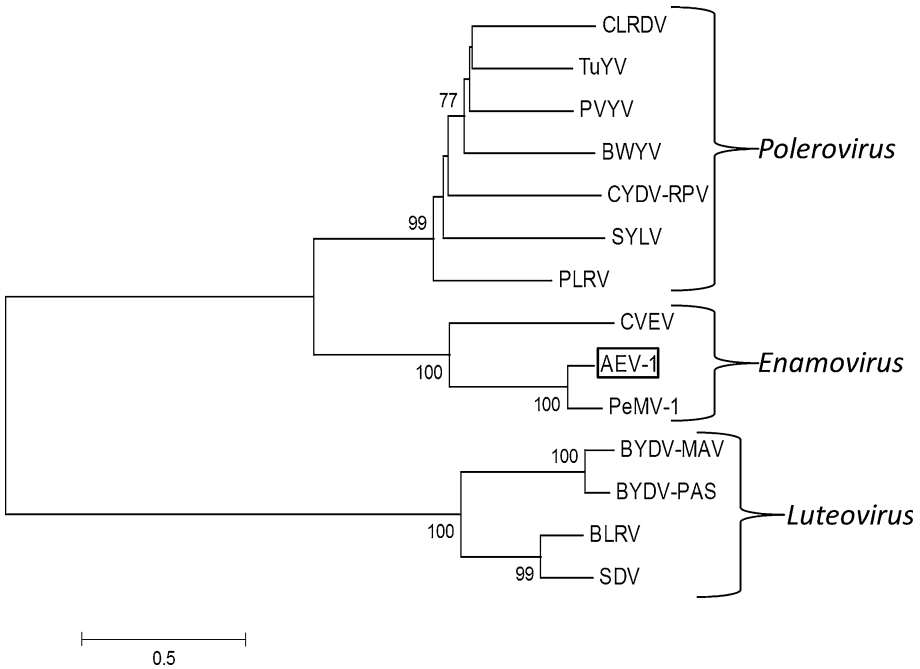


Fig. 2 Phylogenetic tree based on the alignment of the deduced amino acid sequences of the replicase protein (P1-P2) of AEV-1 and representative members of the family *Luteoviridae*. The tree was constructed in MEGA 6 using the neighbor-joining method. The values on the branches show the percentage of support out of 1000 bootstrap replications. The scale bar indicates the number of substitutions per base. The viruses used to construct the tree and their accession numbers are as follows: barley yellow dwarf virus-MAV (BYDV-MAV; NC_003680), barley yellow dwarf virus-PAS (BYDV-PAS; NC_002160), bean leafroll virus (BLRV; NC_003369), beet western yellows virus (BWYV; NC_004756), cereal yellow dwarf virus-RPV (CYDV-RPV; NC_004751), cotton leafroll dwarf virus (CLRDV; NC_014545), citrus vein enation virus (CVEV; NC_021564), pea enation mosaic virus-1 (PeMV-1; NC_003629), pepper vein yellows virus (PVYV; NC_015050), potato leafroll virus (PLRV; NC_001747), soybean dwarf virus (SDV; NC_003056), sugarcane yellow leaf virus (SYLV; NC_000874), turnip yellow virus (TuYV; NC_003743)

deduced amino acid sequence identities (Fig. 1d). The maximum nt sequence identity was 80.3, 80.1 and 49.6 %, respectively, whereas the maximum aa sequence identity was 82.7, 82.7 and 50.8 %, respectively (Fig. 1d). Therefore, the differences in aa sequence identity for each gene product were greater than 10 %, which is one of the criteria used by the International Committee on Taxonomy of Viruses to demarcate species in the genera *Polerovirus* and *Luteovirus* [3]. On the other hand, in a phylogenetic tree based on the P1-P2 fusion protein aa sequence of viruses of the family *Luteoviridae*, AEV-1 clustered with PeMV-1 in the enamovirus complex (Fig. 2). PeMV-1 infects several legume crops, including chickpea (*Cicer arietinum*), faba bean (*Vicia faba*), lentil (*Lens culinaris*) and pea (*Pisum sativum*) [11], but it does not infect alfalfa [12]. Taken together, these results suggest that AEV-1 may belong to a new species in the genus *Enamovirus* and would be the first member of this genus to infect alfalfa.

To determine the prevalence of AEV-1, we examined alfalfa field samples showing dwarfism symptoms that we collected between 2010 and 2015 during field trips to several regions of Argentina, by RT-PCR using specific primers designed from the polymerase coding sequence. Forward primer 5′ TGACCATGAGTCCACTGAGG 3′ and reverse primer 5′ AGTATCTTCCAACAACCACAT 3′ amplified a 451-bp diagnostic fragment [13]. AEV-1 was detected in 56 of the 85 samples collected (65.8 %) and was widely distributed across Argentina (Supplementary Fig. 1).

The availability of the AEV-1 genome sequence will make it possible to assess its genetic variability and construct an infectious clone to investigate the role of this virus in alfalfa dwarfism disease.

Acknowledgments This work was supported by the PNPV 1135022 project of INTA.

Compliance with ethical standards

Conflict of interest There is no conflict of interest.

References

1. Bejerman N, Nome C, Giolitti F, Kitajima E, de Breuil S, Pérez Fernández J, Basigalup D, Cornacchione M, Lenardon S (2011) First report of a Rhabdovirus infecting alfalfa in Argentina. Plant Dis 95:771
2. Trucco V, de Breuil S, Bejerman N, Lenardon S, Giolitti F (2014) Complete nucleotide sequence of *Alfalfa mosaic virus* isolated

from alfalfa (*Medicago sativa* L.) in Argentina. Virus Genes 48:562–565
3. Domier LL (2012) Family—Luteoviridae. In: Andrew MQK, Elliot L, Michael JA, Carstens EB (eds) Virus taxonomy, 9th edn. Elsevier, San Diego, pp 1045–1053
4. Blanca JM, Pascual L, Ziarsolo P, Nuez F, Cañizares J (2011) Ngs_backbone: a pipeline for read cleaning, mapping and SNP calling using next generation sequence. BMC Genomics 12:285
5. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo read assembly using de Bruijn graphs. Genome Res 18:821–829
6. Demler SA, de Zoeten GA (1991) The nucleotide sequence and luteovirus-like nature of RNA 1 of an aphid non-transmissible strain of Pea enation mosaic virus. J Gen Virol 72:1819–1834
7. Fusaro AF, CorreaRL Nakasugi K, JacksonC Kawchuk L, Vaslin MFS, Waterhouse PM (2012) The Enamovirus P0 protein is a silencing suppressor which inhibits local and systemic RNA silencing through AGO1 degradation. Virology 426:178–187
8. Martin RR, Keese PK, Young MJ, Waterhouse PM, Gerlach WL (1990) Evolution and molecular biology of luteoviruses. Annu Rev Phytopathol 28:341–363
9. Miller WA, Brown CM, Wang S (1997) New punctuation for the genetic code: luteovirus gene expression. Semin Virol. 8:3–13
10. Vives MC, Velázquez K, Pina JA, Moreno P, Guerri J, Navarro L (2013) Identification of a new enamovirus associated with citrus vein enation disease by deep sequencing of small RNAs. Phytopathology 103:1077–1083
11. Hagedorn DJ (1984) Compendium of pea diseases. The American Phytopathological Society, St Paul, p 60
12. Larsen RC, Kaiser WJ, Klein RE (1996) Alfalfa, a non-host of pea enation mosaic virus in Washington State. Can J Plant Sci 76:521–524
13. Tornos T, Cebrián MC, Córdoba-Sellés MC, Alfaro-Fernández A, Herrera-Vásquez JA, Font MI, Jorda MC (2008) First report of pea enation mosaic virus infecting pea and broad bean in Spain. Plant Dis 92:1469