Virology 483 (2015) 275–283

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Complete genome sequence and integrated protein localization and interaction map for alfalfa dwarf virus, which combines properties of both cytoplasmic and nuclear plant rhabdoviruses

Nicolás Bejerman [a,b,*,1], Fabián Giolitti [a], Soledad de Breuil [a], Verónica Trucco [a], Claudia Nome [a], Sergio Lenardon [a], Ralf G. Dietzgen [b]

[a] Instituto de Patología Vegetal (IPAVE), Centro de Investigaciones Agropecuarias (CIAP), Instituto Nacional de Tecnología Agropecuaria (INTA), Camino a 60 Cuadras k 5,5, Córdoba X5020ICA, Argentina

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

ARTICLE INFO

Article history:
Received 15 March 2015
Returned to author for revisions
1 May 2015
Accepted 2 May 2015
Available online 21 May 2015

Keywords:
Genome structure
Cytorhabdovirus
Nucleorhabdovirus
Protein localization
Protein–protein interactions
Phosphoprotein

SUMMARY

We have determined the full-length 14,491-nucleotide genome sequence of a new plant rhabdovirus, alfalfa dwarf virus (ADV). Seven open reading frames (ORFs) were identified in the antigenomic orientation of the negative-sense, single-stranded viral RNA, in the order 3′-N-P-P3-M-G-P6-L-5′. The ORFs are separated by conserved intergenic regions and the genome coding region is flanked by complementary 3′ leader and 5′ trailer sequences. Phylogenetic analysis of the nucleoprotein amino acid sequence indicated that this alfalfa-infecting rhabdovirus is related to viruses in the genus *Cytorhabdovirus*. When transiently expressed as GFP fusions in *Nicotiana benthamiana* leaves, most ADV proteins accumulated in the cell periphery, but unexpectedly P protein was localized exclusively in the nucleus. ADV P protein was shown to have a homotypic, and heterotypic nuclear interactions with N, P3 and M proteins by bimolecular fluorescence complementation. ADV appears unique in that it combines properties of both cytoplasmic and nuclear plant rhabdoviruses.

Introduction

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 was found associated with alfalfa plants showing symptoms of shortened internodes (bushy appearance), leaf puckering and varying-sized vein enations on abaxial leaf surfaces (Bejerman et al., 2011). The plants that showed these distinct symptoms were diagnosed as being co-infected by a rhabdovirus and alfalfa mosaic virus (AMV), led to significant yield losses and reduced the useful economic life of the crop (Trucco et al., 2014). 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, pers. comm.).

Members of the family *Rhabdoviridae* can infect a wide range of hosts, including humans, livestock, fish, plants, and insects (Ammar et al., 2009; Jackson et al., 2005; Kuzmin et al., 2009). Six rhabdovirus genera are recognized by the International Committee on Taxonomy of Viruses (ICTV) in its 9th report (Dietzgen et al., 2011) and five additional genera have since been approved (ICTV Master Species List 2013v2). Plant-infecting rhabdoviruses are currently classified into two genera, *Nucleorhabdovirus* and *Cytorhabdovirus,* which are distinguished depending on whether the viruses elicit inclusions in the nucleus, bud from the inner nuclear envelope, and accumulate in perinuclear spaces, or whether they develop cytoplasmic viroplasms, undergo morphogenesis from cytoplasmic membranes, and accumulate in the cytoplasm, respectively (Jackson et al., 2005). Viruses of both genera are transmitted by insects, in which they also replicate and circulate. Plant rhabdoviruses have been reported worldwide, and can infect most major crops and a number of weed hosts in temperate, subtropical and tropical regions (Jackson et al., 2005; Redinbaugh and Hogenhout, 2005).

Plant rhabdoviruses have non-segmented negative-sense, single stranded RNA genomes of approximately 11–15 kb, that encode at least six proteins: nucleocapsid protein (N), phosphoprotein (P), movement protein (P3), matrix protein (M), glycoprotein (G) and RNA-dependent RNA polymerase (L) in the order 3′ N-P-P3-M-G-L 5′ (Dietzgen et al., 2011). The viral RNA is tightly encapsulated by the N protein, and this N-RNA complex, together with the P and L

* Corresponding author at: Instituto de Patología Vegetal (IPAVE), Centro de Investigaciones Agropecuarias (CIAP) Instituto Nacional de Tecnología Agropecuaria (INTA), Camino a 60 Cuadras k 5,5, Cordoba X5020ICA, Argentina. Tel.: +54 351 4973636.

E-mail addresses: n.bejerman@uq.edu.au, nicobejerman@gmail.com (N. Bejerman).

[1] Present address: Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD 4072, Australia. Tel.: +61 7 3346 6521.

http://dx.doi.org/10.1016/j.virol.2015.05.001
0042-6822/

proteins, forms a helical ribonucleoprotein (RNP) complex that is essential for virus replication. The M protein is thought to be responsible for condensation of RNP complexes into a skeleton-like structure (RNP-M core) during virion assembly, and G protein is thought to form trans-membrane spikes (Jackson et al., 2005).

Although more than 75 plant rhabdoviruses have been detected in a variety of host species based on their large bacilliform virions, the complete genome sequences of only 6 cytorhabdoviruses (Dietzgen et al., 2006; Heim et al., 2008; Ito et al., 2013; Schoen et al., 2004; Tanno et al., 2000; Yan et al., 2015) and 8 nucleorhabdoviruses (Bandyopadhyay et al., 2010; Heaton et al., 1989; Huang et al., 2003; Massah et al., 2008; Pappi et al., 2013; Reed et al., 2005; Revill et al., 2005; Tsai et al., 2005) have been determined. Thus, the sequence determination of the complete genomes of additional plant rhabdoviruses is sorely needed to support and clarify the taxonomy of this important group of viruses and lead to a better understanding of their evolutionary trajectories.

Initial electron microscopy and phylogenetic analysis of a partial fragment of the L gene suggested that the alfalfa-infecting rhabdovirus is a tentative member of the genus *Cytorhabdovirus*, which has been named alfalfa dwarf virus (ADV) (Bejerman et al., 2011). Here, we report and analyze the complete genome sequence of ADV, demonstrate the intracellular localization of its encoded proteins by transient expression as autofluorescent fusions in *Nicotiana benthamiana* leaves and establish a viral protein–protein interaction map. Unexpectedly, viral protein localization and interactions were not fully consistent with ADV being a cytorhabdovirus, whereas these properties have been in complete agreement for all other plant rhabdoviruses analyzed so far.

Results

Alfalfa dwarf virus genome sequence analysis

The negative-sense RNA genome of ADV is 14,491 nucleotides (nt) in length (GenBank accession number KP205452) and contains seven open reading frames (ORFs) in the anti-genome, positive-sense orientation (Fig. 1B). BlastX searches identified these ORFs as encoding the nucleocapsid protein (N; ORF1), phosphoprotein (P; ORF2), putative movement protein (P3; ORF3), matrix protein (M; ORF 4), glycoprotein (G; ORF 5), and RNA-dependent RNA polymerase (L; ORF 7) based on highest sequence identity scores with plant rhabdoviruses, whereas P6 (ORF6) did not have significant matches with any Genbank plant rhabdovirus entries (Table 1). The coding sequences are flanked by complementary 3′ leader (l) and 5′ trailer (t) sequences (the 22 terminal nt are complementary) revealing a genome organization of 3′ l-N-P-P3-M-G-P6-L-t 5′ (Fig. 1B). All ADV genes are separated by conserved gene junctions, which are composed of a polyadenyilation signal of the preceding gene, an intergenic region which varied in length from 3 nt to 18 nt, and a transcriptional start of the following gene (Table 2). ADV gene junctions were most similar to those of cytorhabdoviruses (Table 2). Amino acid (aa) sequence comparisons between the deduced ADV proteins (except P protein) and the corresponding sequences of other plant rhabdoviruses (Table 3) revealed the closest relationships to cytorhabdoviruses, in particular persimmon virus A (PeVA) (the full genome data for strawberry crinkle virus (SCV) is not accessible). However, ADV P protein shared similar sequence identities with both cyto- and nucleorhabdoviruses (Table 3). The levels of amino acid sequence identity between ADV and PeVA were low and ranged from 17.4% in the matrix proteins to 43.1% in the L polymerase proteins; ADV M protein was most similar to the M protein of lettuce yellow mottle virus (LYMoV) (Table 3).

The characteristics of proteins encoded by ADV genome as determined by predictive algorithms are shown in Table 1. The N gene contains a 1443-nt ORF, which encodes the nucleocapsid protein with predicted molecular weight of 53.2 kDa and an isoelectric point of 9.07; sequence identity between the N proteins of ADV and other cytorhabdoviruses ranged from 19.9% to 31.4% (Table 3). The P gene contains a 933-nt ORF, which encodes the phosphoprotein with a predicted molecular weight of 35.1 kDa and an isoelectric point of 4.68; sequence identity between the P proteins of ADV and other cytorhabdoviruses ranged from 14.7% to 19.7% (Table 3). The P3 gene contains a 720-nt ORF, which encodes the putative movement protein with a predicted molecular weight of 27 kDa and an isoelectric point of 9.92; sequence identity between the P3 proteins of ADV and other cytorhabdoviruses ranged from 14.7% to 18.5% (Table 3). The M gene contains a 570-nt ORF, which encodes the putative matrix protein with a predicted molecular weight of 20.4 kDa and an isoelectric point of 8.05; sequence identity between the M proteins of ADV and other cytorhabdoviruses ranged from 17.4% to 22.7% (Table 3).

The G gene contains a 1692-nt ORF, which encodes the glycoprotein with a predicted molecular weight of 63.05 kDa and an isoelectric point of 7.7; the sequence identity between the G proteins of ADV and other cytorhabdoviruses ranged from 19.4% to 29.5% (Table 3). The amino terminal region of the ADV G protein is hydrophobic, and its overall hydrophilicity plot (data not shown) is similar to those of the G proteins of LYMoV (Heim et al., 2008) and lettuce necrotic yellows virus (LNYV) (Dietzgen et al., 2006). In ADV G, a signal peptide consisting of 30 aa and containing a IGD=R (=indicates the predicted cleavage site) was identified. Additionally, six potential glycosylation sites (N-[P]-S/T-[P]-) were

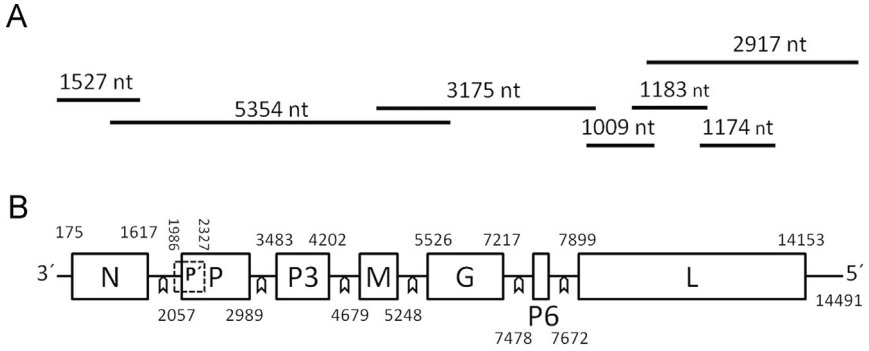

Fig. 1. Schematic diagram of ADV (A) genome sequencing strategy and (B) genome organization. The seven overlapping fragments amplified to cover the whole genome are indicated in (A). Open reading frames (ORFs) are shown as squares and N, P, P3, M, G, P6, L genes identified in (B). Numbers above or below each ORF, indicate ORF first and last nucleotides. Intergenic regions are indicated by arrowheads and the internal P gene ORF P′ is shown as dashed square.