Genome sequencing of two *Bell pepper endornavirus* (BPEV) variants infecting *Capsicum annuum* in Colombia

Secuenciación del genoma de dos variantes de *Bell pepper endornavirus* (BPEV) que infectan *Capsicum annuum* en Colombia

Laura Muñoz-Baena[1], Mauricio Marín-Montoya[2], and Pablo A. Gutiérrez[1]

ABSTRACT

Transcriptome analysis of chili and bell pepper samples from commercial plots in the municipalities of Santa Fe de Antioquia and El Peñol in the province of Antioquia revealed the presence of viral sequences with significant similarity to genomes of members of the genus *Endornavirus*. Assembly of the chili and bell pepper transcriptomes resulted in consensus sequences of 14,727 nt and 14,714 nt that were identified as *Bell pepper endornavirus* (BPEV). Both sequences were nearly identical by 99.9 % at both nucleotide and amino acid levels. The presence of BPEV was confirmed by RT-qPCR, RT-PCR and Sanger sequencing using RdRp-specific primers designed from the assembled sequences in ten independent random samples taken from the investigated bell pepper stands. The phylogenetic analysis of both BPEV variants and their affiliation within the genus *Endornavirus* is discussed. For our knowledge, this is the first study on this group of viruses in Colombia.

Key words: genomics, molecular biology, *Solanaceae*, virus diseases.

RESUMEN

El análisis de transcriptomas de ají y pimentón en muestras obtenidas de cultivos comerciales en Santa Fe de Antioquia y El Peñol en el departamento de Antioquia, reveló la presencia de secuencias virales con similitud significativa con miembros del género *Endornavirus*. El ensamblaje de los transcriptomas arrojó secuencias consenso de 14.727 nt y 14.714 nt que fueron identificadas como los genomas de la especie viral *Bell pepper endornavirus* (BPEV), al compartir con ésta un 99,9% de identidad a nivel de nucleótidos y aminoácidos. La presencia de BPEV fue confirmada en diez muestras foliares de pimentón mediante pruebas de RT-qPCR, RT-PCR y secuenciación parcial de la región RdRp con cebadores específicos diseñados en este trabajo. La relación filogenética de ambas variantes de BPEV con otros miembros del género *Endornavirus* es analizada y discutida. Para nuestro conocimiento, este es el primer trabajo reportado sobre este grupo de virus en Colombia.

Palabras clave: genómica, biología molecular, *Solanaceae*, enfermedades virales.

Introduction

Endornavirus is a recently approved (2006) genus within the family *Endornaviridae* that comprises viruses characterized by a naked double-stranded RNA (dsRNA) genome of about 9.8–17.6 kb, a single open reading frame (ORF) and a site-specific single chain break (nick) about 1.2-2.7 kbp from the 5' end (King *et al*., 2012). The encoded polypeptide is a large multidomain protein of about 3,000-6,000 amino acids that may contain methyltransferase (MT), viral RNA helicase (Hel), glucosyltransferase (GT) and RNA-dependent RNA polymerase (RdRp) domains; only the latter is observed across all species (Roossinck *et al*., 2011). Endornaviruses are normally found in low copy numbers (~100 copies/cell), seem to lack cell-to-cell movement and, apparently, can only be transmitted vertically through seeds and pollen in plants and spores in fungi or oomycetes (Straminipila) (Roossinck *et al*., 2011). In general, endornaviruses do not cause visible disease symptoms (Gibbs *et al*., 2000; Fukuhara *et al*., 2006) but some evidence suggests their involvement in cytoplasmic male sterility in broad bean (*Vicia faba*) (Grill and Garger, 1981; Pfeiffer, 1998) and hypovirulence in the plant pathogenic fungi *Helicobasidium mompa* (Osaki *et al*., 2006).

The first plant endornavirus genome sequence was obtained for *Oryza sativa endornavirus* (OsEV) identified in cultivated rice (*Oryza sativa* ssp. japonica), which was to become the type species of the genus (Moriyama *et al*., 1995; Fukuhara and Moriyama, 2008). The existence of endornaviruses infecting other plants was subsequently confirmed in wild rice (*Oryza rufipogon;* Moriyama *et al*., 1999*),* broad bean (*Vicia fava;* Pfeiffer, 1998), kidney bean (*Phaseolus vulgaris;* Wakarchuk and Hamilton, 1990) barley (*Hordeum*

Received for publication: 19 October, 2016. Accepted for publication: 15 March, 2017. Doi: 10.15446/agron.colomb.v35n1.60626

[1] Laboratory of Industrial Microbiology, Faculty of Sciences, Universidad Nacional de Colombia, Medellin (Colombia).
[2] Laboratory of Cellular and Molecular Biology, Faculty of Sciences, Universidad Nacional de Colombia, Medellin (Colombia). mamarinm@unal.edu.co

vulgare; Zabalgogeazcoa and Gildow, 1992), *Yerba mate* (*Ilex paraguariensis*; Debat *et al*., 2014), Malabar spinach (*Basella alba* L.; Okada *et al*., 2014), Grapevine (*Vitis vinif*era; Espach *et al*., 2012), Avocado (*Persea Americana*; Villanueva *et al*., 2012), cucurbits (Coutts, 2005) and various *Capsicum* species (Valverde and Gutiérrez, 2007; Jo *et al*., 2015). As indicated, endornaviruses are not exclusive to plants as they have been found infecting basidiomycetes: *Rhizoctonia solani* (Das *et al*., 2014), *Rhizoctonia cerealis* (Li *et al*., 2014), *Helicobasidium mompa* (Osaki *et al*., 2006); ascomycetes: *Rosellinia necatrix* (Yaegashi and Kanematsu, 2016), *Erysiphe cichoracearum* (Du *et al*., 2016), *Alternaria brassicicola* (Shang *et al*., 2015), *Sclerotinia sclerotiorum* (Khalifa and Pearson, 2014), *Tuber aestivum* (Stielow *et al*., 2011), *Gremmiella abietina* (Tuomivirta *et al*., 2009) and oomycetes within the genus *Phytophthora* (Hacker *et al*., 2005; Kozlakidis *et al*., 2010).

The first report regarding endornaviruses in *Capsicum annuum* described them as large molecular weight dsRNAs present in chloroplast fractions (Valverde *et al*., 1990), however, these molecules were only classified as the genomes of endornaviruses 17 years later after cDNA sequencing revealed their phylogenetic affinity to *Oryza sativa endornavirus* (OsEV) and *Oryza rufipogon endornavirus* (OrEV) in 2007 (Valverde and Gutiérrez, 2007). Endornaviruses are ubiquitous in *Capsicum* species (Okada *et al*., 2011), as confirmed by *in silico* analysis of pepper transcriptomes deposited in public databases (Jo *et al*., 2016) and Next-generation sequencing (NGS) studies of total small RNAs (Sela *et al*., 2012; Chen *et al*., 2015; Jo *et al*., 2015; Lim *et al*., 2015). In this work, we report the sequence of two *Bell pepper endornavirus* (BPEV) genomes naturally infecting *Capsicum annuum* in the province of Antioquia (Colombia). RT-PCR, RT-qPCR and Sanger sequencing using primers designed for this purpose confirmed the presence of BPEV in field samples.

Materials and methods

Next-generation sequencing

High-throughput sequencing of the *C. annuum* transcriptomes was performed on bulk leaf samples of chili pepper collected at the municipalities of Santa Fe de Antioquia and bell pepper from El Peñol (Antioquia, Colombia). After grinding the leaf tissue with liquid nitrogen, total RNA was extracted with the GeneJET Plant RNA Purification Mini kit (Thermo Fisher Scientific, USA) and rRNA depleted with the TruSeq Stranded Total RNA with Ribo-Zero Plant kit (Illumina, USA). Libraries were constructed using the TruSeq RNA Sample Preparation kit (Illumina, USA) and sequencing performed with the Illumina HiSeq 2000 system service provided by Macrogen (South Korea). Adapter sequences and low quality bases were removed with SeqTK prior analysis (https://github.com/lh3/seqtk). Sequence assembly was performed with Trinity (Grabherr *et al*., 2011) and endornaviral contigs were identified by a local BLASTN search using a database of endornavirus reference genomes. Genome assemblies were confirmed by mapping with Bowtie2 (Langmead and Salzberg, 2012) and checked for inconsistencies and assembly artifacts with Tablet (Milne *et al*., 2010). Consensus sequences were deposited in GenBank under accession codes KX977568 (BPEV from Bell pepper) and KX977569 (BPEV from Chili pepper).

RT-PCR and RT-qPCR detection

Primers BPEV_F (5'-AGG CTA AAT GTG CAC CTA AAA TTG G-3'; Tm = 60.3°C), BPEV_R (5'-TTT CTC AGC GAC TGC TGA CC-3'; Tm = 60.3°C) and qBPEV_R (5'-CTT TAC ACT GCC ATA ACA ACG C-3'; Tm = 58.5°C) were designed for specific amplification of BPEV using RT-PCR (BPEV_F and BPEV_R) or RT-qPCR (BPEV_F and qBPEV_R) using the assembled genomes as reference. Primer specificity was verified *in silico* using the program primer-BLAST (Ye *et al*., 2012). Experimental validation of primers was performed in ten random leaf samples collected in four commercial bell pepper fields from the municipalities of Marinilla and El Peñol (Antioquia). The NGS samples and a reaction mix lacking template cDNA were used as positive and negative controls, respectively. RNA was extracted from 100 mg of ground tissue using the GeneJET Plant RNA Purification kit (Thermo Fisher Scientific, Waltham, MA, USA) and eluted in 40 µL of DEPC treated water; the purity and concentration were determined by absorbance readings at 260 and 280 nm using a Nanodrop 2000C (Thermo Fisher Scientific). Retrotranscription was performed for 30 min at 50°C in 20 µL containing 200 U of Maxima Reverse Transcriptase (Thermo Fisher Scientific), 1X RT Buffer, 0.5 mM dNTP Mix, 100 pmol of specific reverse primer BPEV_R, 20 U de RiboLock RNase Inhibitor and 100-500 ng of total RNA. For the qPCR, the Maxima SYBR Green/ROX qPCR Master Mix (2X) kit (Thermo Fisher Scientific) was used in 25 µL of reaction containing 12.5 µL mix, 10 µL DEPC water, primers BPEV_F/ qBPEV_R at 0.3 µM and 50-100 ng cDNA. Samples were considered positive if they exhibited fluorescence values higher than the threshold before the 35th cycle (Schena *et al*., 2004). Primer specificity was verified by High Resolution Melting (HRM) in the 50 and 99°C range and compared to melting temperature (Tm) values from the positive controls (NGS samples). RT-PCR validation of primers BPEV_F/BPEV_R

was performed using cDNA obtained as described above. The amplification mix included 17.8 µL water, 1X enzyme buffer, 1.8 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM primers, 1 U Taq DNA polymerase (Thermo Fisher Scientific, USA) and 50-100 ng de cDNA in a final volume of 25 µL. The amplification consisted of an initial 3 min incubation at 95°C, followed by 35 cycles that included denaturation at 94°C (30 s), annealing at 52°C (1 min) and extension at 72°C (1 min); the amplification ended with an extension cycle at 72°C for 5 min. Amplicon size was determined by 1.8% agarose gel electrophoresis stained with GelRed 1X (Biotium, Hayward, CA, USA) in Bio Doc Analyze transilluminator (Biometra, Göttingen, Germany). Some sequences for qRT-PCR and RT-PCR amplification products were confirmed by the Sanger method using an ABI Prism 3730xl Sequencer (Applied Biosystems, Carlsbad, CA, USA) at Macrogen (South Korea). Sanger sequences were deposited in GenBank under accession codes KX977563-KX977567.

Bioinformatic analyses

The ORF coding for the endornaviral polyprotein was identified using the DNA translate program available at http://web.expasy.org. Identification of protein motifs was performed in the Conserved Domain Database at NCBI (Marchler-Bauer *et al*., 2015). Evolutionary analyses based on the complete polyprotein were constructed by the Jones-Taylor-Thornton model using a gamma distribution with five categories and a shape parameter of 1.09 (Jones *et al*., 1992). Evolutionary distances for the partial nucleotide sequences were computed using the Tamura 3-parameter method, rate variation among sites was modeled with a gamma distribution with a shape parameter of 1.2. Phylogenetic trees were calculated in MEGA7 using the Neighbor-Joining method with 1,000 bootstrap replicates (Kumar *et al*., 2016).

Results and discussion

Next-generation sequencing of the *C. annuum* sample of chili pepper from Santa Fe de Antioquia resulted in a transcriptome of 5,605,703 paired-end reads with a total of 1,121,140,600 nt. A local BLASTN search identified 1,033 reads (0.018%) with significant similarity to *Bell pepper endornavirus* (BPEV, NC_015781) with percent sequence identities between 82.5 and 100 percent (Fig. 1A). Similar results were obtained with the sample of bell pepper from El Peñol. In this case, a transcriptome consisting of 6,901,644 paired-end reads (1,380,328,800 nt) was obtained of which 2,936 were classified as associated to the BPEV genome (0.042%), with sequence identities ranging from 82.2 to 100 per cent (average of 95.2%) (Fig. 1B). Previous studies have shown the copy number of endornaviral sequences in *C. annuum* samples to be in the 0.01% to 0.18% range, which coincides with our results (Jo *et al*., 2016). As the species demarcation criteria within the genus *Endornavirus* has been set below 75% nucleotide sequence identity (King *et al*., 2012), it is clear that both viruses are BPEV isolates.

After confirming the presence of endornavirus in both samples, their genomes were assembled. The BPEV sequence from Santa Fe de Antioquia (BPEV_Santa_Fe) resulted in a consensus of 14,727 nt with an average sequence depth of 48.25X and 34 polymorphic sites (Fig. 2A). BPEV_Santa_Fe encodes a putative protein of 4,885 amino acids between nucleotide positions 26 to 14,683 that contains viral methyltransferase (M, cl03298), viral RNA helicase (H, pfam01443), glycosyltransferase (GT, cd03784) and RNA-dependent RNA polymerase (RdRp, cl03049) motifs at amino acid positions 327-562 (e-value: $3.98e^{-8}$), 1413-1649, (e-value: $2.23x10^{-6}$), 3,113-3,458 (e-value: $1.17x10^{-25}$) and 447-4,785 (e-value: $1.14x10^{-21}$), respectively. Amino acid polymorphisms were observed at positions

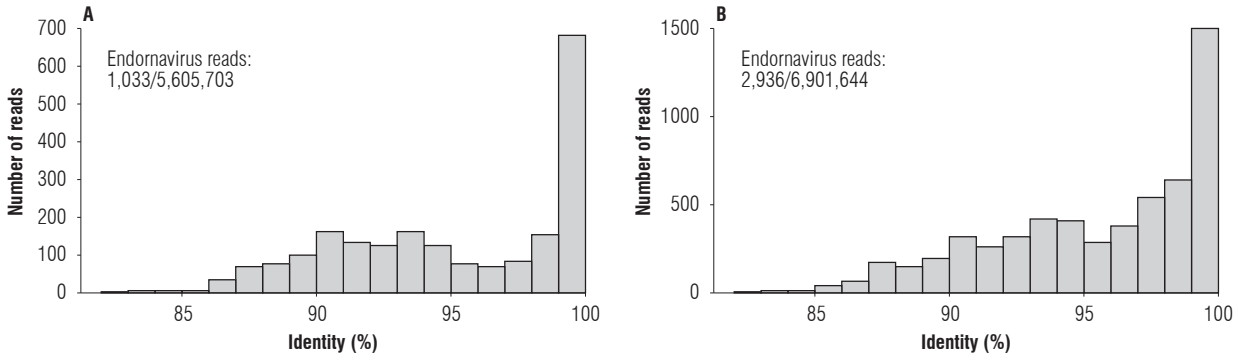

FIGURE 1. Frequency histograms showing the number of paired-end reads at different nucleotide sequence percent identity with respect to member of the genus *Endornavirus* in the *Capsicum annuum* transcriptomes from the municipalities of Santa Fe de Antioquia (A) and El Peñol (B) in Antioquia (Colombia).

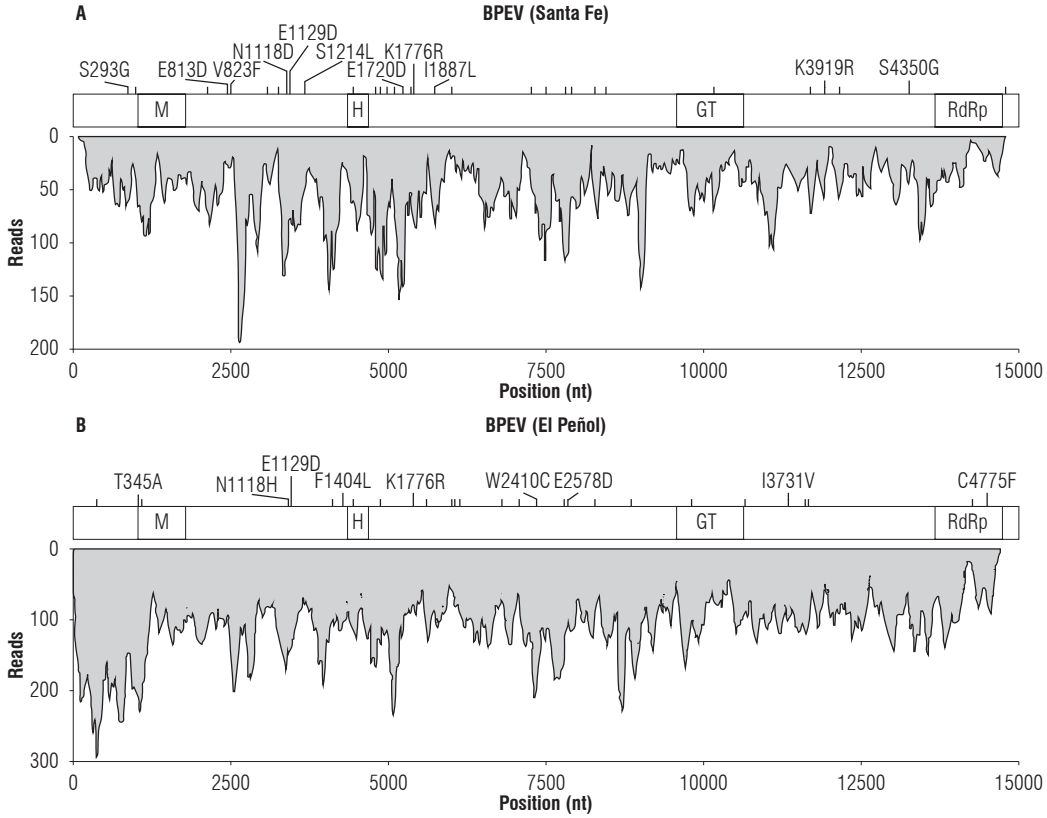

FIGURE 2. Sequence coverage and variability of the BPEV assemblies from Santa Fe de Antioquia (A) and El Peñol (B). The position of the open reading frame encoding for the endornaviral polyprotein is shown on top of each coverage graph and includes the location of predicted methyltransferase (M), helicase (H), glucosyltransferase (GT) and RNA-dependent RNA polymerase (RdRp) domains. Nucleotide polymorphic sites are indicated as vertical bars; amino acid changes are explicitly indicated when applicable.

S293G, E813D, V823F, N1118D, E1129D, S1214L, E1720D, K1776R, I1887L, K3919R and S4350.

The sample from El Peñol resulted in a contig of 14,714 nt (BPEV_Peñol) with average sequence depth of 105.4X (max 290X) and 32 polymorphic positions (Fig. 2B). An ORF, encoding a protein of 4,884 residues with a similar domain structure as BPEV_Santa_Fe was identified between positions 20 and 14,674. The following amino acid sequence variations resulting from polymorphisms in the assembly were observed: T345A, N1118H, E1129D, F1404L, K1776R, W2410C, E2578D, I3731V, C4775F. The consensus genomes of BPEV_Santa_Fe and BPEV_Peñol were practically identical as only eight nucleotide changes (T23C, T3647C, C3666T, T4826C, A5684T, A7720G, C10033T and G11781A) and three amino acid substitutions (L1214S, L1887I, K3919R) were observed; additionally, a leucine codon (GCT) is inserted in BPEV_Santa_Fe at position 11,464.

A survey of GenBank sequences using BPEV_Santa_Fe and BPEV_Peñol as query, resulted in significant matches with BPEV isolates Kyosuzu from Japan (99.2%, AB597230), Maor (99.3%, KP455654) and Yolo Wonder (88.3%, JN019858) from the United States, Healey from Canada (99.1%, KT149366), IS from Israel (99.2%, JQ951943) and Lj from China (98.1%, KF709944). Isolate Kyosuzu (BPEV_KS) was shown to have 100% incidence in tested bell pepper cultivars in Japan and some of its variants were found to infect other *C. annuum* genotypes as well as related *Capsicum* species such as *C. baccatum*, *C. chinense* and *C. frutescens* (Okada *et al*., 2011). BPEV_KS can be transmitted through seed but not by graft inoculations and hybridization studies after denaturing agarose gel electrophoresis detected a nick in the plus strand at position 880 (Okada *et al*., 2011). BPEV Yolo wonder (BPEV-YV) was the first BPEV to be sequenced (Okada *et al*., 2011). The other complete BPEV sequences have been only characterized *in silico*: isolate Maor was identified from transcript shotgun assembly as a contig with strong sequence identity to *Bell pepper endornavirus* (Jo *et al*., 2016); isolate Healey was identified in a NGS study of small RNAs extracted from the leaves of a pepper plant (cultivar Healey) with mild crinkling and chlorosis symptoms (Chen *et al*., 2015) and

BPEV_IS was identified in asymptomatic C. *annuum* L. cv. Yatir leaves in Israel by NGS sequencing of viral small RNA (Sela *et al*., 2012). There are not published reports on isolate Lj from China. Outside the BPEV group, the closest endornavirus species is Hot pepper endornavirus (HPEV, KR080326) a proposed species (Lim *et al*., 2015) also infecting *C. annuum* that shares 78.8% nucleotide sequence identity with BPEV_Santa_Fe and BPEV_Peñol. Due to the high similarity between the Colombian BPEV isolates and BPEV_KS, it is likely for them to share the same molecular and biological properties.

A phylogenetic analysis using complete endornaviral proteins confirmed the previous analysis (Fig. 3A). As expected, BPEV_Santa_Fe and BPEV_Peñol clustered within the group of Bell pepper endornaviruses (bootstrap of 100%) with HPEV as a sister species. The *Capsicum*-infecting endornaviruses are part of a larger clade of plant endornaviruses that includes PvEV-2, infecting common bean and, more distantly, HvEV infecting barley. A comparison of functional domains reveals that BPEV, HPEV, PvEV2 and HvEV share methyltransferase, helicase and RNA-dependent RNA polymerase domains; however, HvEV lacks a UDP-glycosyltransferase present in the others (Fig. 3B). It has been suggested that the UDP-glycosyltransferase domain was transferred to endornaviruses from marine and freshwater bacteria (Song *et al*., 2013) and it is likely that this event probably occurred after the split between HvEV and the PvEV2/BPEV/HPEV group.

A global analysis of the phylogenetic relationships among endornaviruses, reveal two additional clades infecting plants. One of these clades has *Vicia faba endornavirus* (VfEV) as a sole member and, in spite of having the largest

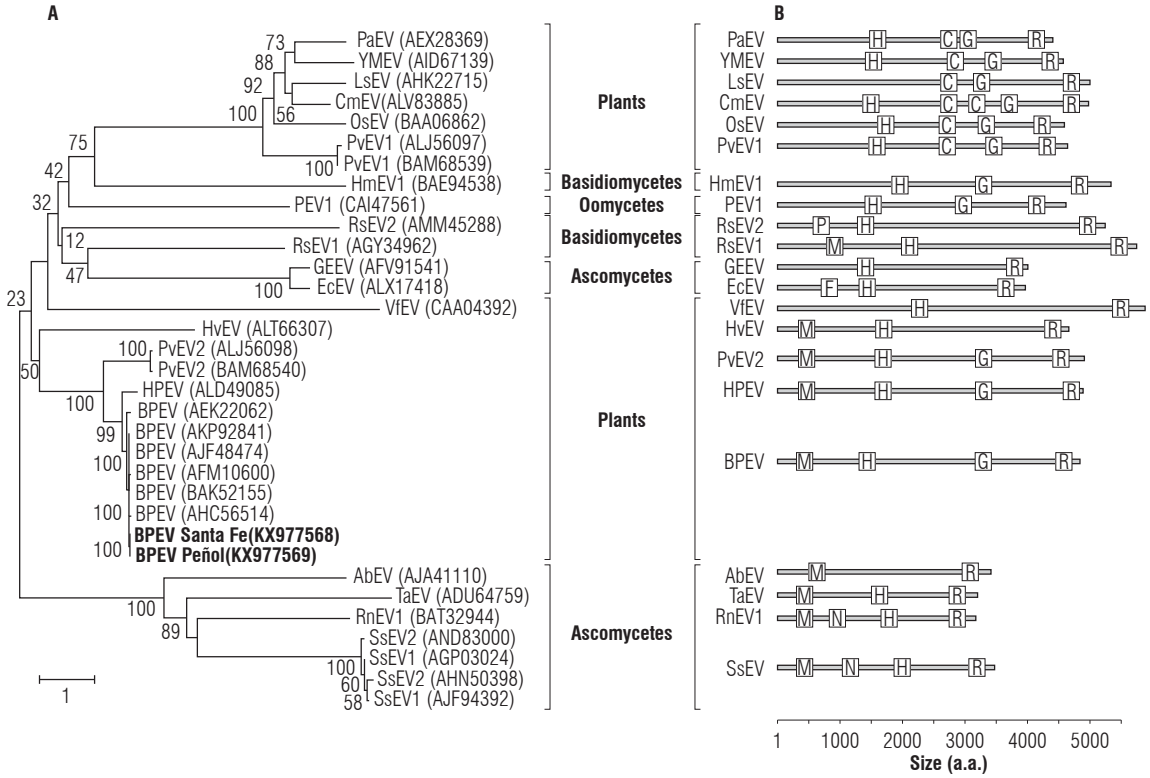

FIGURE 3. Evolutionary relationship of BPEV_Santa_Fe and BPEV_Peñol to members of the genus *Endornavirus*. A) Neighbor-joining analysis of complete endornaviral polyproteins infecting plants, fungi (Basidiomycetes and Ascomycetes) and Oomycetes (Straminipila). The tree is drawn to scale with branch lengths in units of number of amino acids substitutions per site as indicated in the bar at the bottom. B) Relative location of functional domains of the polyproteins depicted in the tree on the left. Species abbreviations: *Persea americana endornavirus,* PaEV; *Yerba mate endornavirus*, YMEV; *Lagenaria siceraria endornavirus*, LsEV; *Cucumis melo endornavirus*, CmEV; *Oryza sativa endornavirus*, OsEV; *Phaseolus vulgaris endornavirus 1*, PvEV1; *Helicobasidium mompa endornavirus*, HmEV1; *Phytophthora endornavirus 1*, PEV1; *Rhizoctonia solani endornavirus 2*, RsEV2; *Rhizoctonia cerealis endornavirus 1*, RcEV1; *Grapevine endophyte endornavirus 1*, GEEV; *Erysiphe cichoraceum endornavirus*, EcEV; *Vicia faba endornavirus*, VfEV; *Hordeum vulgare endornavirus*, HvEV; *Phaseolus vulgaris endornavirus 2*, PvEV2; *Hot pepper endornavirus*, HPEV; *Bell pepper endornavirus*, BPEV; *Alternaria brassicicola endornavirus*, AbEV; *Tuber aestivum endornavirus*, TaEV; *Rosellinia necatrix endornavirus 1*, RnEV1; *Sclerotinia sclerotiorum endornavirus* 1; SsEV1; *Sclerotinia sclerotiorum endornavirus* 2; SsEV2. Domain abbreviations: putative methyltransferase, M; viral helicase, H; glycosyltransferase, G; RNA-dependent RNA-polymerase, R; Glycosyltransferase sugar-binding, C; Phospholipase A2, P; Furin-Like, F; Neuromodulin, N and DEAD-like helicase, D. Bootstrap values are shown above branches.

polyprotein, has only two identifiable domains with Helicase and RdRp function; VfEV is more closely related to fungal endornaviruses. The third group of plant endornaviruses comprises PvEV1, OsEV, CmEV, LsEV, YMEV and PaEV, characterized by the lack of a methyltransferase domain and the presence of an additional glycosyltransferase sugar binding domain containing a DXD motif (cl19952; Song *et al*., 2013). Interestingly, members of the OsEV group have been shown to localize in microsomal fractions while members of the BPEV group localize in the chloroplast, a fact that could be explained by their different domain structure (Moriyama *et al*., 1996; Okada *et al*., 2013). The polyprotein phylogenetic analysis confirms the complex evolutionary history of endornaviruses and suggests at least two independent origins of plant endornaviruses from fungi (Roossinck *et al*., 2011; Okada *et al*., 2013). It is likely that plant endornaviruses were first transferred to the plant kingdom from endornaviruses infecting ascomycetes fungi, all sharing Methyltransferase domains. Radiation of endornaviruses to basidiomycetes, oomycetes (Straminipila) and other ascomycetes seem to derive from a genetic transfer from plants, an evolutionary event characterized by the loss of the methyltransferase domain. A second transfer to plants from basidiomycetes fungi seem to have occurred which later resulted in the incorporation of a second Glycosyltransferase domain. It is interesting to note that all the hosts of non-plant endornaviruses have a close ecological relationship with plants as either pathogens (*S. sclerotiorum*, *R. necatrix, A. brassicicola, E. cichoracearum, R. solani, R. cerealis, Phytophthora* sp., *H. longisporum*) or soil-inhabitants (*T. aestivum*).

To confirm the widespread presence of endornaviruses infecting *C. annumm* cultivars in Antioquia primers BPEV_F, BPEV_R and qBPEV_R were designed using the sequenced genomes as reference (Fig. 4A). Primer BPEV_F

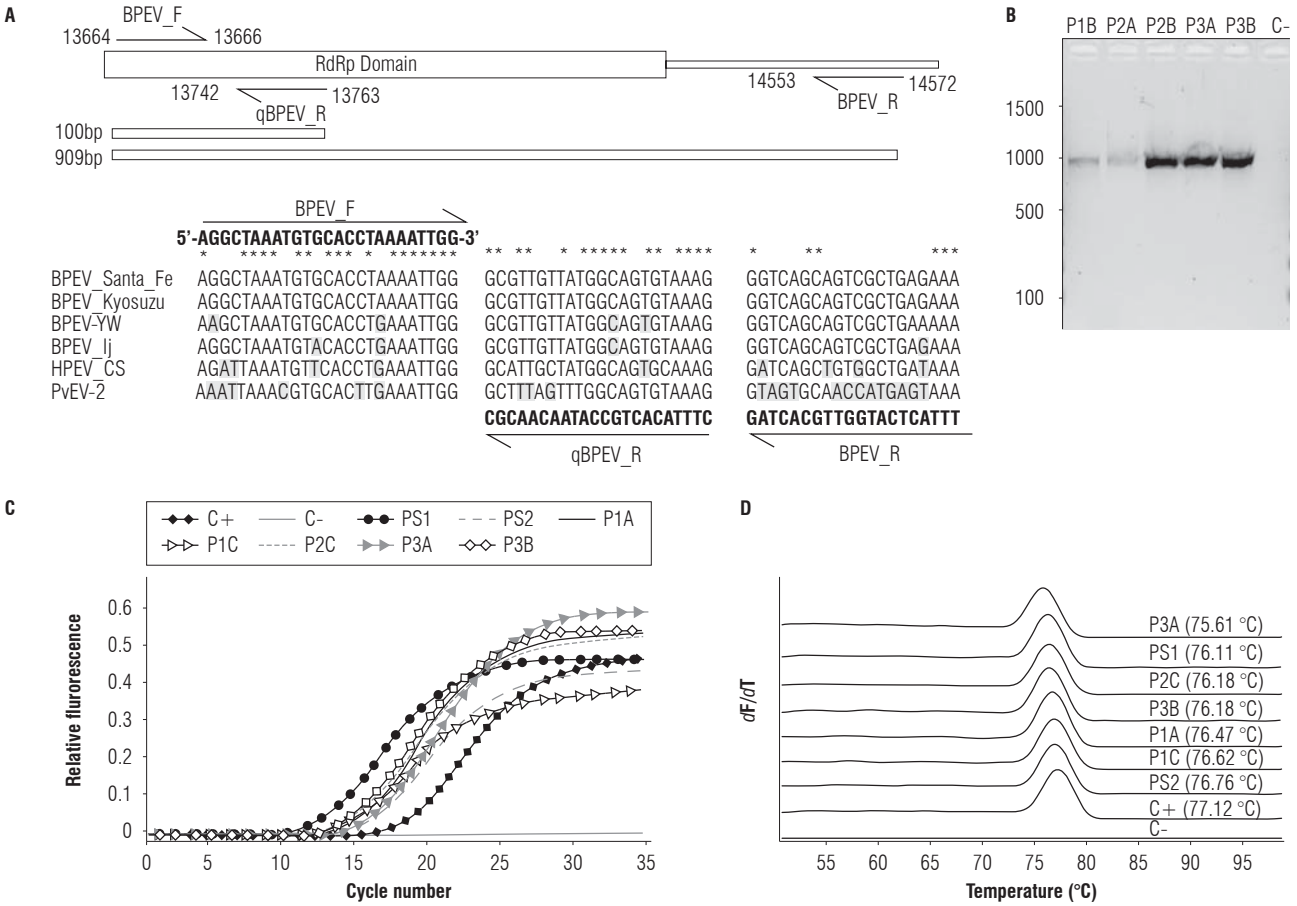

FIGURE 4. Molecular detection of BPEV in *C. annuum* samples from Marinilla and El Peñol. A. Location of the BPEV_F, BPEV_R and qBPEV_R primers described in the text with respect to the BPEV_Santa_Fe consensus genome and the RdRp motif. The sequence alignment with selected endornaviral genomes highlights variable positions with respect to each primer. B. RT-PCR detection of BPEV using the BPEV_F/R primer set showing the expected amplification product of 909 bp in a subset of the field samples. C. Validation of the BPEV_F/ qBPEV_R primer set in RT-qPCR. D. High Resolution Melting analysis of the amplification reactions shown in panel C. The positive control corresponds to the NGS sample sequenced from chili pepper.

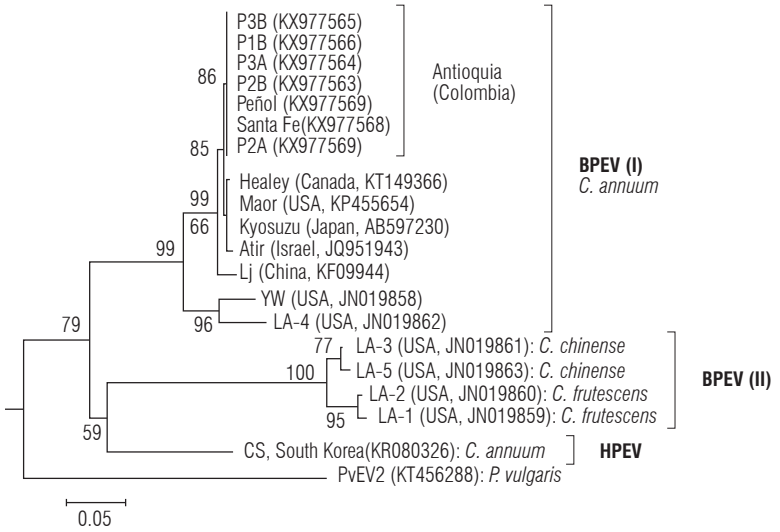

FIGURE 5. Evolutionary relationship between *Bell pepper endornavirus* isolates using partial nucleotide sequences of the RdRp region for a subset of the field samples. Sequences are labelled by the isolate name; location and GenBank accession codes are shown in parentheses. *Phaseolus vulgaris endornavirus 2* was used as outgroup. The tree was constructed by the Neighbor-Joining method using distances calculated by the Tamura 3-parameter method. The tree is drawn to scale with branch lengths in units of number of base substitutions per site as indicated in the bar at the bottom. Bootstrap values are shown above the branches.

binds to an internal segment of the RdRp domain which in combination with BPEV_R gives an amplification product of 909 bp useful in RT-PCR detection; in combination with primer qBPEV_R, BPEV_F amplifies a 100 bp segment suitable for RT-qPCR. Both primer sets were tested in *C. annuum* samples from different cultivation plots in Marinilla and El Peñol (Antioquia). In eight foliage samples, out of ten tested, RT-PCR gave amplification products with the expected size and sequence (Fig. 4B). Similar results were obtained by RT-qPCR using the BPEV_F/qBPEV_R primer set, in this case, sigmoidal amplification profiles were found in all the ten samples with Ct values in the 11.36-18.29 range (Fig. 4C). HRM analysis of the amplification reactions demonstrates that the amplification was specific, as evidenced by the presence of single denaturation peaks and their similar Tm values with respect to the NGS samples used as positive controls (Fig. 4D). Sanger sequencing confirmed the identity of the qRT-PCR amplicons. These results are in agreement with previous work suggesting a high incidence of BPEV in different *Capsicum* species such as *C. annuum*, *C. frutescens* and *C. chinens*e (Okada *et al*., 2011; Sela *et al*., 2012; Chen *et al*., 2015; Jo *et al*., 2015; Lim *et al*., 2015; Jo *et al*., 2016). It has been shown the phylogenetic relationship between BPEV sequences mirror the evolutionary history of their hosts and it is likely that BPEV was present in the common ancestor giving rise to these three *Capsicum* species but not in primitive *Capsicum* species such as *C. chacoense*, *C. annuum* var. glabriusculum and *C. pubescens* (Okada *et al*., 2011). With Colombia being part of the center of origin of *Capsicum* species, it would be interesting that future investigations address the appearance of endonarviruses within this plant genus as well as their incidence in domesticated and wild species.

Sequences derived from BPEV_F/BPEV_R amplicons (909 bp) were used to construct a phylogenetic tree with corresponding RdRp sequences of BPEV and another *Capsicum*-infecting endornaviruses like HPEV, with PvEV-2 as outgroup (Fig. 5). BPEV sequences form two clusters that correlated well with the infection host. BPEV group I comprises sequences infecting *C. annuum* cultivars and includes all the Colombian sequences reported in this work. Colombian sequences form a distinct group closely related to isolates Healey (Canada), Maor (USA), Kyosuzu (Japan) and Atir (Israel) and more distantly to isolate Lj from China. Two divergent BPEV sequences are observed within group I corresponding to isolates YW and LA-4 from the United States. BPEV group II comprises isolates LA-3 and LA-5 infecting yellow lantern chili (*C. chinense*) and isolates LA-1 and LA-2 infecting *C. frutescens*, a chili pepper species. Group II is more closely related to HPEV infecting *C. annuum* than to members of group I indicating that BPEV is a paraphyletic group as it does not include all sequences derived from the same common ancestor. A taxonomic revision by ICTV on the species comprising the BPEV clade is probably required.

Conclusions

NGS sequencing of the *C. annuum* transcriptome revealed the presence of *Bell pepper endornavirus* (BPEV) in Antioquia (Colombia) closely related to isolates with worldwide distribution. These results were confirmed by sequencing of the RdRp region using primers BPEV_F/BPEV_R and by RT-qPCR using primers BPEV_F/qBPEV_R. The results presented here are in agreement with previous genomic studies on BPEV. Further work should aim at characterizing the molecular features of BPEV in different *C. annuum* cultivating regions in Colombia and confirm completion of the 5' and 3' ends.

Acknowledgment

This work was funded by Universidad Nacional de Colombia (VRI Grant 34585) and International Foundation for Science (Sweden, Grant: C/4634-2).

Literature cited

Chen, B., M. Bernards, and A. Wang. 2015. Complete genome sequence of a *Bell pepper endornavirus* isolate from Canada. Genome Announc. 3(4), pii:e00905-15. Doi: 10.1128/genomeA.00905-15

Coutts, R.H. 2005. First report of an endornavirus in the *Cucurbitaceae*. Virus Genes 31(3), 361-362. Doi: 10.1007/s11262-005-3255-y.

Das, S., R.E. Falloon, A. Stewart, and A.R. Pitman. 2014. Molecular characterisation of an endornavirus from *Rhizoctonia solani* AG-3PT infecting potato. Fungal Biol. 118(11), 924-934. Doi: 10.1016/j.funbio.2014.08.003.

Debat, H.J., M. Grabiele, P.M. Aguilera, R. Bubillo, P.D. Zapata, D.A. Marti, and D.A. Ducasse. 2014. The complete genome of a putative endornavirus identified in yerba mate (*Ilex paraguariensis* St. Hil.). Virus Genes 49(2), 348-350. Doi: 10.1007/s11262-014-1096-2

Du, Z., W. Lin, P. Qiu, X. Liu, L. Guo, K. Wu, S. Zhang, and Z. Wu. 2016. Complete sequence of a double-stranded RNA from the phytopathogenic fungus *Erysiphe cichoracearum* that might represent a novel endornavirus. Arch. Virol. 161(8), 2343-2346. Doi: 10.1007/s00705-016-2911-y.

Espach, Y., H.J. Maree, and J.T. Burger. 2012. Complete genome of a novel endornavirus assembled from next-generation sequence data. J. Virol. 86(23), 13142. Doi: 10.1128/JVI.02538-12.

Fukuhara, T, R. Koga, N. Aoki, C. Yuki, N. Yamamoto, N. Oyama, T. Udagawa, H. Horiuchi, S. Miyazaki, Y. Higashi, M. Takeshita, K. Ikeda, M. Arakawa, N. Matsumoto, and H. Moriyama. 2006. The wide distribution of endornaviruses, large double-stranded RNA replicons with plasmid-like properties. Arch. Virol. 151(5), 995-1002. Doi: 10.1007/s00705-005-0688-5

Fukuhara, T. and H. Moriyama. 2008. Endornaviruses. 109-116. Mahy, B.W.J. and M.H.V. van Regenmortel (eds.). Encyclopedia of virology 3rd ed. Elsevier, Oxford, UK.

Gibbs, M.J., R. Ryuichi, H. Moriyama, P. Pfeiffer, and T. Fukuhara. 2000. Phylogenetic analysis of some large double-stranded RNA replicons from plants suggests they evolved from a defective single-stranded RNA virus. J. Gen. Virol. 81(1), 227-233. Doi: 10.1099/0022-1317-81-1-227.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, BW. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. 29(7), 644-652. Doi: 10.1038/nbt.1883

Grill, L.K. and S.J. Garger. 1981. Identification and characterization of double-stranded RNA associated with cytoplasmic male sterility in *Vicia faba*. Proc. Natl. Acad. Sci. USA. 78(11), 7043-7046. Doi: 10.1073/pnas.78.11.7043

Hacker, C.V., C.M. Brasier, and K.W. Buck. 2005. A double-stranded RNA from a *Phytophthora* species is related to the plant endornaviruses and contains a putative UDP glycosyltransferase gene. J. Gen. Virol. 86(5),1561-1570. Doi: 10.1099/vir.0.80808-0

Jo, Y., H. Choi, and W.K. Cho. 2015. *De novo* assembly of a *Bell pepper endornavirus* genome sequence using RNA sequencing data. Genome Announc. 3(2), pii: e00061-15. Doi: 10.1128/genomeA.00061-1

Jo, Y., H. Choi, J.Y. Yoon, S.K. Choi, and W.K. Cho. 2016. *In silico* identification of *Bell pepper endornavirus* from pepper transcriptomes and their phylogenetic and recombination analyses. Gene 575(2), 712-717. Doi: 10.1016/j.gene.2015.09.051

Jones, D.T., W.R. Taylor, and J.M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8(3), 275-282. Doi: 10.1093/bioinformatics/8.3.275

Khalifa, M.E. and M.N. Pearson. 2014. Molecular characterisation of an endornavirus infecting the phytopathogen *Sclerotinia sclerotiorum*. Virus Res. 189, 303-309. Doi: 10.1016/j.virusres.2014.06.010

King, A.M.Q., E. Lefkowitz, M.J. Adams, and E.B. Cartens. 2012. Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses. Elsevier, San Diego, CA, USA. Doi: 10.1016/B978-0-12-384684-6.00048-3

Kozlakidis, Z., N.A. Brown, A. Jamal, X. Phoon, and R.H. Coutts. 2010. Incidence of endornaviruses in *Phytophthora* taxon douglasfir and *Phytophthora ramorum*. Virus Genes 40(1), 130-134. Doi: 10.1007/s11262-009-0421-7

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33(7), 1870-1874. Doi: 10.1093/molbev/msw054

Langmead, B. and S. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9(4), 357-359. Doi: 10.1038/nmeth.1923

Li, W., T. Zhang, H. Sun, Y. Deng, A. Zhang, H. Chen, and K. Wang. 2014. Complete genome sequence of a novel endornavirus in the wheat sharp eyespot pathogen *Rhizoctonia cerealis*. Arch. Virol. 159(5), 1213-1216. Doi: 10.1007/s00705-013-1893-2

Lim, S., K.H. Kim, F. Zhao, R.H. Yoo, D. Igori, S.H. Lee, and J.S. Moon. 2015. Complete genome sequence of a novel endornavirus isolated from hot pepper. Arch. Virol. 160(12), 3153-3156. Doi: 10.1007/s00705-015-2616-7

Marchler-Bauer, A., M.K. Derbyshire, N.R. Gonzales, S. Lu, F. Chitsaz, L.Y. Geer, R.C. Geer, J. He, M. Gwadz, D.I. Hurwitz, C.J. Lanczycki, F. Lu, G.H. Marchler, J.S. Song, N. Thanki, Z. Wang, R.A. Yamashita, D. Zhang, C. Zheng, and S.H. Bryant. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43(Database issue):D222-226. Doi: 10.1093/nar/gku1221

Milne, I., M. Bayer, L. Cardle, P. Shaw, G. Stephen, F. Wright, and D. Marshall. 2010. Tablet-next generation sequence assembly visualization. Bioinformatics 26(3), 401-402. Doi: 10.1093/bioinformatics/btp666

Moriyama, H., T. Nitta, and T. Fukuhara. 1995. Double-stranded RNA in rice: a novel RNA replicon in plants. Mol. Gen. Genet. 248, 364-369. Doi: 10.1007/BF02191603

Moriyama, H., K. Kanaya, J.Z. Wang, T. Nitta, and T. Fukuhara. 1996. Stringently and developmentally regulated levels of a cytoplasmic double-stranded RNA and its high-efficiency transmission via egg and pollen in rice. Plant Mol. Biol. 31, 713-719. Doi: 10.1007/BF00019459

Moriyama, H., H. Horiuchi, T. Nitta, and T. Fukuhara. 1999. Unusual inheritance of evolutionarily-related double-stranded RNAs in interspecific hybrid between rice plants *Oryza sativa* and *Oryza rufipogon*. Plant Mol. Biol. 39(6), 1127-1136. Doi: 10.1023/A:1006118304093

Okada, R., E. Kiyota, S. Sabanadzovic, H. Moriyama, T. Fukuhara, P. Saha, M.J. Roossinck, A. Severin, and R.A. Valverde. 2011. Bell pepper endornavirus: molecular and biological properties, and occurrence in the genus *Capsicum*. J. Gen. Virol. 92(11), 2664-2273. Doi: 10.1099/vir.0.034686-0

Okada, R., C.K. Yong, R.A. Valverde, S. Sabanadzovic, N. Aoki, S. Hotate, E. Kiyota, H. Moriyama, and T. Fukuhara. 2013. Molecular characterization of two evolutionarily distinct endornaviruses co-infecting common bean (*Phaseolus vulgaris*). J. Gen. Virol. 94(1), 220-229. Doi: 10.1099/vir.0.044487-0

Okada, R., E. Kiyota, H. Moriyama, F. Toshiyuki, and R.A. Valverde. 2014. A new endornavirus species infecting Malabar spinach (*Basella alba* L.). Arch. Virol. 159(4), 807-809. Doi: 10.1007/s00705-013-1875-4

Osaki, H., H. Nakamura, A. Sasaki, N. Matsumoto, and K. Yoshida. 2006. An endornavirus from a hypovirulent strain of the violet root rot fungus, *Helicobasidium mompa*. Virus Res. 118(1-2),143-149. Doi: 10.1016/j.virusres.2005.12.004

Pfeiffer, P. 1998. Nucleotide sequence, genetic organization and expression strategy of the double-stranded RNA associated with the '447' cytoplasmic male sterility trait in *Vicia faba*. J. Gen. Virol. 79 (Pt 10), 2349-2358. Doi: 10.1099/0022-1317-79-10-2349

Roossinck, M.J., S. Sabanadzovic, R. Okada, and R.A. Valverde. 2011. The remarkable evolutionary history of endornaviruses. J. Gen. Virol. 92(11), 2674-2678. Doi: 10.1099/vir.0.034702-0

Schena, L., F. Nigro, A. Ippolito, and D. Gallitelli. 2004. Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. Eur. J. Plant Pathol. 110, 893-908. Doi: 10.1007/s10658-004-4842-9

Sela, N., N. Luria, and A. Dombrovsky. 2012. Genome assembly of bell pepper endornavirus from small RNA. J. Virol. 86(14), 7721. Doi: 10.1128/JVI.00983-12

Shang, H.H., J. Zhong, R.J. Zhang, C.Y. Chen, B.D. Gao, and H.J. Zhu. 2015. Genome sequence of a novel endornavirus from the phytopathogenic fungus *Alternaria brassicicola*. Arch. Virol. 160(7), 1827-1830. Doi: 10.1007/s00705-015-2426-y

Song, D., W.K. Cho, S.H. Park, Y. Jo, and K.H. Kim. 2013. Evolution of and horizontal gene transfer in the *Endornavirus* genus. PLoS One 8(5):e64270. Doi: 10.1371/journal.pone.0064270

Stielow, B., H.P. Klenk, and W. Menzel. 2011. Complete genome sequence of the first endornavirus from the ascocarp of the ectomycorrhizal fungus *Tuber aestivum* Vittad. Arch. Virol. 156(2), 343-345. Doi: 10.1007/s00705-010-0875-x

Tuomivirta, T.T., J. Kaitera, and J. Hantula. 2009. A novel putative virus of *Gremmeniella abietina* type B (Ascomycota: *Helotiaceae*) has a composite genome with endornavirus affinities. J. Gen. Virol. 90(9), 2299-2305. Doi: 10.1099/vir.0.011973-.

Valverde, R.A., S. Nameth, O. Abdalla, O. Al-Musa, P.R. Desjardins, and Dodds J.A. 1990. Indigenous double-stranded RNA from pepper (*Capsicum annuum*). Plant Sci. 67(2), 195-201. Doi: doi:10.1016/0168-9452(90)90243-H

Valverde, R.A. and D.L. Gutierrez. 2007. Transmission of a dsRNA in bell pepper and evidence that it consists of the genome of an endornavirus. Virus Genes 35(2), 399-403. Doi: 10.1007/s11262-007-0092-1

Villanueva, F., S. Sabanadzovic, R.A. Valverde, and J. Navas-Castillo. 2012. Complete genome sequence of a double-stranded RNA virus from avocado. J. Virol. 86(2), 1282-1283. Doi: 10.1128/JVI.06572-11

Wakarchuk, D.A. and R.I. Hamilton. 1990. Partial nucleotide sequence from enigmatic dsRNAs in *Phaseolus vulgaris*. Plant Mol. Biol. 14(4), 637-639. Doi: 10.1007/BF00027512

Yaegashi, H. and S. Kanematsu. 2016. Natural infection of the soilborne fungus *Rosellinia necatrix* with novel mycoviruses under greenhouse conditions. Virus Res. 219, 83-91. Doi: 10.1016/j.virusres.2015.11.004

Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. Madden. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13, 134. Doi: 10.1186/1471-2105-13-134

Zabalgogeazcoa, I.A. and F.E. Gildow. 1992. Double-stranded ribonucleic acid in Barsoy barley. Plant Sci. 83, 187-194.