BRIEF REPORT

Novel bird's-foot trefoil RNA viruses provide insights into a clade of legume-associated enamoviruses and rhabdoviruses

Humberto J. Debat[1] · **Nicolas Bejerman**[1,2]

Abstract

Here, we report the identification and characterization of two novel viruses associated with bird's-foot trefoil. Virus sequences related to those of enamoviruses (ssRNA (+); *Luteoviridae; Enamovirus*) and nucleorhabdoviruses (ssRNA (-); *Rhabdoviridae; Nucleorhabdovirus*) were detected in *Lotus corniculatus* transcriptome data. The genome of the tentatively named "bird's-foot trefoil-associated virus 1" (BFTV-1) is a 13,626-nt-long negative-sense ssRNA. BFTV-1 encodes six predicted gene products in the antigenome orientation in the canonical order 3′-N-P-P3-M-G-L-5′. The genome of the proposed "bird's-foot trefoil-associated virus 2" (BFTV-2) is 5,736 nt long with a typical 5′-PO-P1-2-IGS-P3-P5-3′ enamovirus genome structure. Phylogenetic analysis indicated that BFTV-1 is closely related to datura yellow vein nucleorhabdovirus and that BFTV-2 clusters into a monophyletic lineage of legume-associated enamoviruses. This subclade of highly related and co-divergent legume-associated viruses provides insights into the evolutionary history of the enamoviruses.

From an agronomic point of view, bird's-foot trefoil (*Lotus corniculatus* L.) is the most important *Lotus* species. This plant is a perennial legume that is mainly grown for fodder production in temperate regions and is considered one of the major forage legumes after lucerne (*Medicago sativa*) and white clover (*Trifolium repens*) [9]. It is a high-quality forage that can be grazed or cut for hay or silage and does not cause bloating in ruminants [13]. Intriguingly, to date, the only virus that has been identified to affect this crop is the widely distributed alfalfa mosaic virus (AMV), which

✉ Humberto J. Debat
debat.humberto@inta.gob.ar

✉ Nicolas Bejerman
bejerman.nicolas@inta.gob.ar

1 Instituto de Patología Vegetal, Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria (IPAVE-CIAP-INTA), X5020ICA Córdoba, Argentina

2 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

was found in a field study of Prince Edward Island, Canada, more than 35 years ago [16].

Numerous novel viruses, many of them not inducing any apparent symptoms, have been identified from different environments using metagenomic approaches, which reveals our limited knowledge about the richness of a continuously expanding plant virosphere, in which viruses have been found in every potential host assessed [18]. Genomic RNA molecules of these plant RNA viruses are often inadvertently co-isolated with host RNAs, and their sequences can be detected in plant transcriptome datasets [15, 17]. In a recent consensus statement report, Simmonds et al. [20] report that viruses that are known only from metagenomic data can be, should be, and have been incorporated into the official classification scheme of the International Committee on Taxonomy of Viruses (ICTV). Thus, the analysis of public data constitutes an emerging source of novel *bona fide* plant viruses, which allows the reliable identification of new viruses in hosts with no previous record of virus infections. Here, we have analyzed a transcriptomic dataset from bird's-foot trefoil, which is available in the NCBI Sequence Read Archive (SRA) database, and have identified and characterized a novel enamovirus and a nucleorhabdovirus associated with this crop.

The raw data analyzed in this study correspond to an RNAseq next-generation sequencing (NGS) library (SRA: SRS271068) associated with NCBI Bioproject

PRJNA77207. As described by Wang et al. [23], it was obtained by Illumina Hiseq 2000 sequencing of total RNA isolated from fresh flowers, pods, leaves, and roots from *L. corniculatus* collected in the Qinling Mountains, Shaanxi Province, China (BioSample: SAMN00759026). The 26,492,952 2×90-nt raw reads from the SRA were pre-processed by trimming and filtering with the Trimmomatic tool as implemented in http://www.usadellab.org/cms/?page=trimmomatic, and the resulting reads were assembled *de novo* with Trinity v2.6.6 release using standard parameters. *De novo* transcriptome assembly resulted in 55,473 transcripts, which were subjected to bulk local BLASTX searches (E-value < 1e-5) against a refseq virus protein database available at ftp://ftp.ncbi.nlm.nih.gov/refseq/release/viral/viral.2.protein.faa.gz. The resulting hits were analyzed manually. A 6,524-nt transcript showed a significant match (E-value = 0, identity = 60%) with the L protein of datura yellow vein nucleorhabdovirus (DYVV; [6]). Further similarity analyses resulted in the detection of three more DYVV-like transcripts of 3,953 nt, 2,354 nt and 1,432 nt (Supplemental Table 1). The tentative virus contigs were extended by iterative mapping of the raw reads from the SRS271068 library, using a method described recently by Singh et al. [21]. Briefly, this strategy employs BLAST/nhmmer to extract a subset of reads related to the query contig, uses the retrieved reads to extend it, and then repeats the process iteratively using the extended sequence as a query. Bowtie2 http://bowtie-bio.sourceforge.net/bowtie2/index.shtml was then employed for mean coverage estimation and reads per million (RPMs) calculations. The transcripts that were extended and polished now contained overlapping regions and were subsequently reassembled using the Geneious v8.1.9 (Biomatters Ltd.) alignment tool (cost matrix: > 93% similarity 5.0/-9.026) into a single 13,626-nt-long RNA contig with 57.7% overall sequence identity to DYVV. This resulting RNA sequence was supported by a total of 17,656 reads (mean coverage = 116.4X, reads per million (RPM) = 666.4, Supplemental File 1). To further advance in the characterization of this sequence, virus annotation was carried out as described by Debat [4]. In brief, virus ORFs were predicted with ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/), and the presence of domains and the architecture of translated gene products were determined by InterPro (https://www.ebi.ac.uk/interpro/search/sequence-search) and the NCBI Conserved Domain Database v3.16 (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Further, HHPred and HHBlits as implemented in https://toolkit.tuebingen.mpg.de/#/tools/ were used to complement the annotation of divergent predicted proteins by hidden Markov models.

The genome of the tentatively named "bird's-foot trefoil-associated virus 1" (BFTV-1) is a 13,626-nt-long negative-sense single-stranded RNA (GenBank accession number MH614262) that contains six open reading frames (ORFs) in the anti-genome, positive-sense orientation (Fig. 1A). BLASTP searches identified these ORFs as potentially encoding a nucleocapsid protein (N; ORF1, 462 aa), a phosphoprotein (P; ORF2, 344 aa), a movement protein (P3; ORF3, 325 aa), a matrix protein (M; ORF 4, 294 aa), a glycoprotein (G; ORF 5, 637 aa), and an RNA-dependent RNA polymerase (L; ORF 6, 2104 aa), based on highest sequence identity scores with nucleorhabdoviruses, in particular with black currant-associated nucleorhabdovirus (BCaRV; [24]), DYVV, and sonchus yellow net nucleorhabdovirus (SYNV; [12, 14]) (Table 1). The coding sequences are flanked by 3' leader (l) and 5' trailer (t) sequences, and the genome organization is 3' l–N–P–P3–P4-M–G–L–t 5' (Fig. 1A), which is the basic genome organization of plant rhabdoviruses [7] without any accessory genes. Like all plant rhabdoviruses, BFTV-1 genes are separated by intergenic "gene junction" regions, which are composed of the polyadenylation signal of the preceding gene, a short intergenic region, and the transcriptional start site of the following gene (Supplemental Table 2). Interestingly, BFTV-1 the consensus "gene junction" region sequence 3′-AUUCUUUUUGGUUGUA-5′ is identical to that of BCaRV, DYVV and SYNV (Supplemental Table 2). All BFTV-1-encoded proteins contain a classical mono- or bipartite nuclear localization signal (NLS) [7]. Importin-α-dependent nuclear localization signals were predicted using cNLS Mapper available at http://nls-mapper.iab.keio.ac.jp/, and the scores predicted an exclusively nuclear localization for the N, M and L proteins, while for P, P3 and G, the scores suggested localization to both the nucleus and cytoplasm (Table 1). Nuclear export signals were predicted using NetNES 1.1, available at www.cbs.dtu.dk/services/NetNES/, which predicted a leucine-rich nuclear export signal in the BFTNRV N protein near amino acid position 350 (data not shown), suggesting both nuclear import and export can occur. Similar results have been reported for the predicted products of DYVV [7]. Amino acid sequence comparisons between the deduced BFTV-1 proteins and the corresponding sequences of other nucleorhabdoviruses revealed high similarity of BFTV-1 to BCaRV, DYVV and SYNV. The BFTV-1N, G and L proteins were the most similar to those encoded by BCaRV, DYVV and SYNV, sharing 37.0-58.4% sequence identity, whereas P, P3 and M were the most divergent (Table 1).

Phylogenetic analysis based on the predicted replicase protein was done using MAFTT 7 https://mafft.cbrc.jp/alignment/software/ with multiple amino acid sequence alignments (BLOSUM62 scoring matrix), using E-INS-i (BFTV-1) or G-INS-I (BFTV-2) as the best-fit model. The aligned protein sequences were subsequently used as input in FastTree 2.1.5 (http://www.microbesonline.org/fasttree/) to generate phylogenetic trees by the maximum-likelihood method (best-fit model = JTT-Jones-Taylor-Thorton with
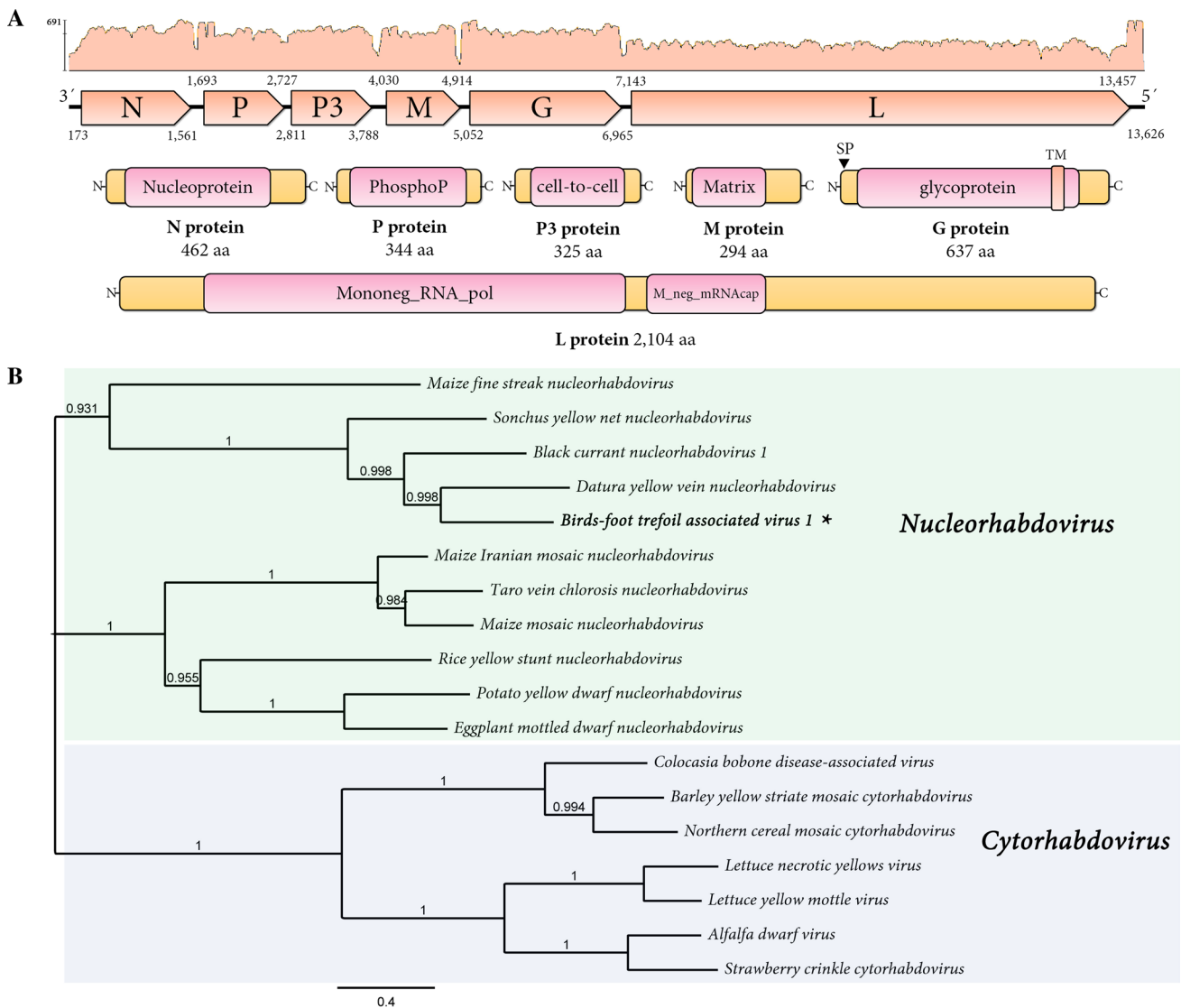
Fig. 1 Structural characterization and phylogenetic analysis of bird's-foot trefoil-associated virus 1 (BFTV-1) (**A**) Genome graphs depicting coverage, architecture, and predicted gene products of BFTV-1. The predicted coding sequences are shown in orange arrow rectangles, and start and end coordinates are indicated. Gene products are depicted in curved yellow rectangles, and the size in aa is indicated below the diagram. Predicted domains or HHPred best-hit regions are shown in curved pink rectangles. Abbreviations: N, nucleoprotein CDS; P, phosphoprotein CDS; P3, putative cell-to-cell movement protein CDS; M, matrix protein CDS; G, glycoprotein CDS; L, RNA dependent RNA polymerase CDS; TM, transmembrane domain; SP, signal peptide. (**B**) Maximum-likelihood phylogenetic tree based on amino acid sequence alignments of the L polymerase of BFTNRV and other plant rhabdoviruses. The tree is rooted at the midpoint; nucleorhabdovirus and cytorhabdovirus clades are indicated by green and blue rectangles, respectively. The scale bar indicates the number of substitutions per site. Node labels indicate Fast-Tree support values. The viruses used to construct the tree and their accession numbers (in parentheses) are as follows: black currant nucleorhabdovirus (MF543022), alfalfa dwarf virus (KP205452), barley yellow striate mosaic cytorhabdovirus (KM213865), colocasia bobone disease associated-virus (KT381973), datura yellow vein nucleorhabdovirus (KM823531), eggplant mottled dwarf nucleorhabdovirus (NC_025389), lettuce yellow mottle virus (EF687738), lettuce necrotic yellows virus (NC_007642), maize fine streak nucleorhabdovirus (AY618417), maize Iranian mosaic nucleorhabdovirus (DQ186554), maize mosaic nucleorhabdovirus (AY618418), northern cereal mosaic cytorhabdovirus (AB030277), potato yellow dwarf nucleorhabdovirus (GU734660), rice yellow stunt nucleorhabdovirus (NC_003746), sonchus yellow net virus (L32603), taro vein chlorosis virus (AY674964) and strawberry crinkle cytorhabdovirus (MH129615)

single rate of evolution for each site = CAT). Local support values were computed using the Shimodaira-Hasegawa test (SH) with 1,000 replicates. The resulting tree shows that BFTV-1 clusters together with other viruses of the genus *Nucleorhabdovirus* (Fig. 1B). BFTV-1 appears to have a close evolutionary relationship to BCaRV, DYVV and SYNV, which share a similar genomic organization to that described for BFTV-1 [8, 13, 24]. We speculate that

Table 1 Summary of BFTV-1-genome-encoded proteins

| ORF	Gene	Putative function	Lenght (aa)	MW (kD)	IEP[a]	Predicted NLS [cNLS Mapper score]	BLASTP e-value	Identity DYVV (%)	Identity BCaRV (%)	Identity SYNV (%)
1	N	Nucleocapsid protein	462	52.1	8.9	Bipartite [7.5]	0.0	58.4	50.1	46.4
2	P	Phosphoprotein	344	38.3	4.5	Bipartite [3.4]	3,00E–50	34.3	27.6	23.0
3	P3	Movement protein	325	36.7	8.7	Bipartite [3.4]	6,00E–98	42.6	28.9	22.2
4	M	Matrix protein	294	32.6	8.2	231LKKTLKRRKIA241 [9]	2,00E–54	32.6	32.0	27.7
5	G	Glycoprotein	637	71.1	6.8	Bipartite [3.3]	0.0	46.7	39.2	37.0
6	L	Polymerase	2,104	240.1	7.7	1638SSSKRKRTSI1647 [9.5]	0.0	50.1	48.9	43.2

[a]IEP, isoelectric point pH

BFTV-1, BCaRV, DYVV and SYNV might represent the most ancestral clade of nucleorhabdoviruses, since no accessory genes are present in their genomes. Taken together, our results suggest that BFTV-1 should be considered a member of a new virus species in the genus *Nucleorhabdovirus*.

In addition to the nucleorhabdovirus sequence, the BLASTX searches of the assembled transcriptome retrieved a 5,622-nt transcript which showed a significant sequence similarity (Supplemental Table 1, E-value = 0, identity = 74%) to the P1-P2 fusion protein of alfalfa enamovirus 1 (AEV-1; [2]). Mapping of the SRA raw reads extended the contig into a 5,736-nt RNA sequence (mean coverage = 61.6X, reads per million (RPM) = 147.1, Supplemental File 2) with 69.2% overall sequence identity to AEV-1. The final sequence of the tentative "bird's-foot trefoil-associated virus 2" (BFTV-2) consisted of 5,736 nt (GenBank accession number MH614261). Structural and functional annotation of the BFTV-2 virus sequence indicated a typical 5′-PO-P1-2-IGS-P3-P5-3′ enamovirus genome structure (Fig. 2A). The first ORF, ORF 0, consists of 909 nt, encoding a putative P0 protein of 303 aa with a calculated molecular mass of 33.8 kDa. In enamoviruses, P0 has been shown to function as an RNA silencing suppressor [11]. BFTV-2 P0 is most similar to that of the legume-infecting enamoviruses AEV-1 and pea enation mosaic virus-1 (PEMV-1), where an F-box-like motif located in the N-region of the predicted protein was identified (Supplemental Fig. 1A). Interestingly, this motif was not identified in the non-legume-infecting enamoviruses. The F-box-like motif is involved in suppression of silencing [11], and therefore, in those viruses, other viral proteins, rather than P0, may be involved in suppression of silencing. This hypothesis needs to be tested experimentally. The second ORF, ORF1, which is 2,283 nt long, is predicted to be expressed by a ribosomal leaky scanning mechanism to produce protein P1 (761 aa, 83.9 kDa). The third ORF, ORF2, which is translated by a -1 ribosomal frameshift from ORF 1, overlaps ORF1 at its 5′ end and is predicted to produce an ORF1-ORF2 fusion protein, P1-P2 (1183 aa, 131.4 kDa). The canonical motif for a −1 frameshift site is X_XXY_YYZ. A putative slippery sequence of the type G_GGA_AAC (Supplemental Fig. 2A), which is identical to that of PEMV-1 was detected at position 2,039 [5]. In addition, a 40-nt highly structured H-type pseudoknot (free energy = -12.80) was detected eight nucleotides downstream (7-nt spacer), immediately following the slippery sequence (Supplemental Fig. 2B), which was predicted using the KnotInFrame tool available at https://bibiserv2.cebitec.uni-bielefeld.de/knotinframe and visualized with the VARNA 3.93 applet (http://varna.lri.fr/). It is worth noting that the frameshifting pseudoknot in other poleroviruses and enamoviruses, including PEMV-1, is seven nt from the end of the shifty site (6-nt spacer), since the spacing between the slippery sequence and the

pseudoknot is important [1]. Further studies are needed to test the tentative prediction of this report. P1 and P1-P2 are putatively involved 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) [5]. A serine-protease-like domain (peptidase S39, pfam02122, P1 residues 314-515; E-value, 3.75e-42) in P1 and an RdRP domain (RdRP_4, pfam02123) resembling those of members of the genera *Polerovirus*, *Rotavirus* and *Totivirus* in P2 (P1-P2 residues 734-1116; E-value, 2.26e-61) were found when these protein sequences were analyzed (Fig. 2A). The fourth ORF, ORF3, consists of 570 nt, encoding a predicted P3 coat protein (190 aa, 21.2 kDa). The fifth ORF, ORF5, is a putative in-frame readthrough product of ORF3 encoding the fusion protein P3-P5 (506 aa, 56.1 kDa). The stop codon in BFTV-2 is in the canonical context for readthrough that has been observed in other enamoviruses (UGA-GGG; [10]). Moreover, between nt 4,601 and 6,665, 15 nt downstream of the stop codon, 11 CCNNN tandem repeat motif units were found, which are associated with readthrough of the stop codon of ORF 3 [3]. Furthermore, signatures in a cytidine-rich domain immediately after the stop codon (5' C-rich) and a branched stem-loop structure 684 nt downstream of the CP stop codon (CD-DRTE), which are required for ORF3 readthrough [25], were also identified in the genome sequence of BFTV-2 (Supplemental Fig. 3A-B). Both the BFTV-2 5' C-rich and CD-DRTE regions showed a high degree of sequence similarity to that of AEV-1, with the expected complementary bases of the two regions identical with the exception of a G-to-A SNP at the CD-DRTE, which is a non-complementary base in both BFTV-2 and AEV-1. 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 [5]. While the CP region of the predicted P3-P5 protein is more conserved in enamoviruses, some emerging motifs could be observed, such as a proline-rich stretch immediately following the readthrough region (Supplemental Fig. 1.B). This motif could be involved in the transmission of BFTV-2; nonetheless, mutational analysis coupled with virus transmission assays should be conducted to test this hypothesis. When these protein sequences were analyzed in detail, a luteovirid CP domain (Luteo_coat, pfam00894) in P3 (P3 residues 55-188; E-value, 1.63e-70) and a polerovirus readthrough protein domain (PLRV_ORF5, pfam01690) in P5 (P3-P5 residues 221-424; E-value, 1.16e-48) were identified.

The functional and structural results indicate that BFTV-2 should be taxonomically classified as a member of the family *Luteoviridae*. Viruses in the family *Luteoviridae* have a single-stranded positive-sense RNA genome and are classified into three genera, *Enamovirus*, *Luteovirus* and *Polerovirus*. Unlike poleroviruses, enamoviruses do not encode a P4 movement protein, and luteoviruses lack a P0 gene [8]. Therefore, BFTV-2 appears to be an enamovirus because it has a P0 gene and lacks a P4 gene. The BFTV-2 ORFs were compared to the predicted ORFs of AEV-1, PEMV-1, grapevine enamovirus 1 (GEV-1; [19]) and those of citrus vein enation virus (CVEV; [22]), in order to determine the level of nucleotide and deduced amino acid sequence similarity (Supplemental Table 3). The maximum nt sequence identity for each gene coding sequence (CDS) was 75.4, 77.1, 47.2 and 44.9%, respectively, whereas the maximum aa sequence identity for each gene product was 82.6, 85.3, 44.0 and 37.7%, respectively (Supplemental. Table 3). Therefore, the differences in aa sequence identity for each gene product were greater than 10%, which is one of the criteria used by the ICTV to demarcate species in the genera *Polerovirus*, *Luteovirus* [8], and *Enamovirus*. Consequently, BFTV-2 may belong to a new species in the genus *Enamovirus*. Furthermore, in a phylogenetic tree based on the P1-P2 fusion protein aa sequences of viruses of the family *Luteoviridae*, BFTV-2 clustered with AEV-1 and PEMV-1 in a monophyletic clade of legume-associated enamoviruses within the enamovirus complex (Fig. 2B), which suggest that all legume-infecting enamoviruses share a common ancestor. Further assessment of the genetic distance of each gene product of all reported enamoviruses (Supplemental Fig. 4) indicated an evident high similarity among the legume-associated enamoviruses, including BFTV-2. Unlike PEMV-1, which was first reported in pea, AEV-1 has been detected only in alfalfa, a member of the Faboideae subfamily of legumes, like bird's-foot trefoil. Thus, these legume viruses appear to have co-diverged. Additional legume viruses would be helpful to determine whether the evolutionary path of enamoviruses is characterized by co-divergence. AEV-1 is associated with a complex of viruses that is responsible for alfalfa dwarf disease [2]. It is interesting that BFTV-2 was detected together with a nucleorhabdovirus. Future studies should focus on determining if there might be any synergistic effect on the host caused by coinfection with enamoviruses and nucleorhabdoviruses – a subtle interaction scenario that may have important implications for management of associated diseases. For instance, the umbravirus PEMV-2 provides the movement protein [5, 8] function that is lacking in the enamovirus PEMV-1. No umbraviruses were identified in the bird's-foot trefoil dataset. While it is plausible that the lack of detection was associated with sensitivity or technical issues, another possibility is that the identified nucleorhabdovirus could be providing the movement functions to complement and thus allow the movement and replication of BFTV-2 in the host. It is worth mentioning that recent publications have reported enamoviruses (GEV-1 [19], AEV-1) lacking coinfecting umbraviruses. Interestingly, AEV-1 has also been found in alfalfa plants infected with the nucleorhabdovirus alfalfa dwarf virus (ADV) [2].

A

B
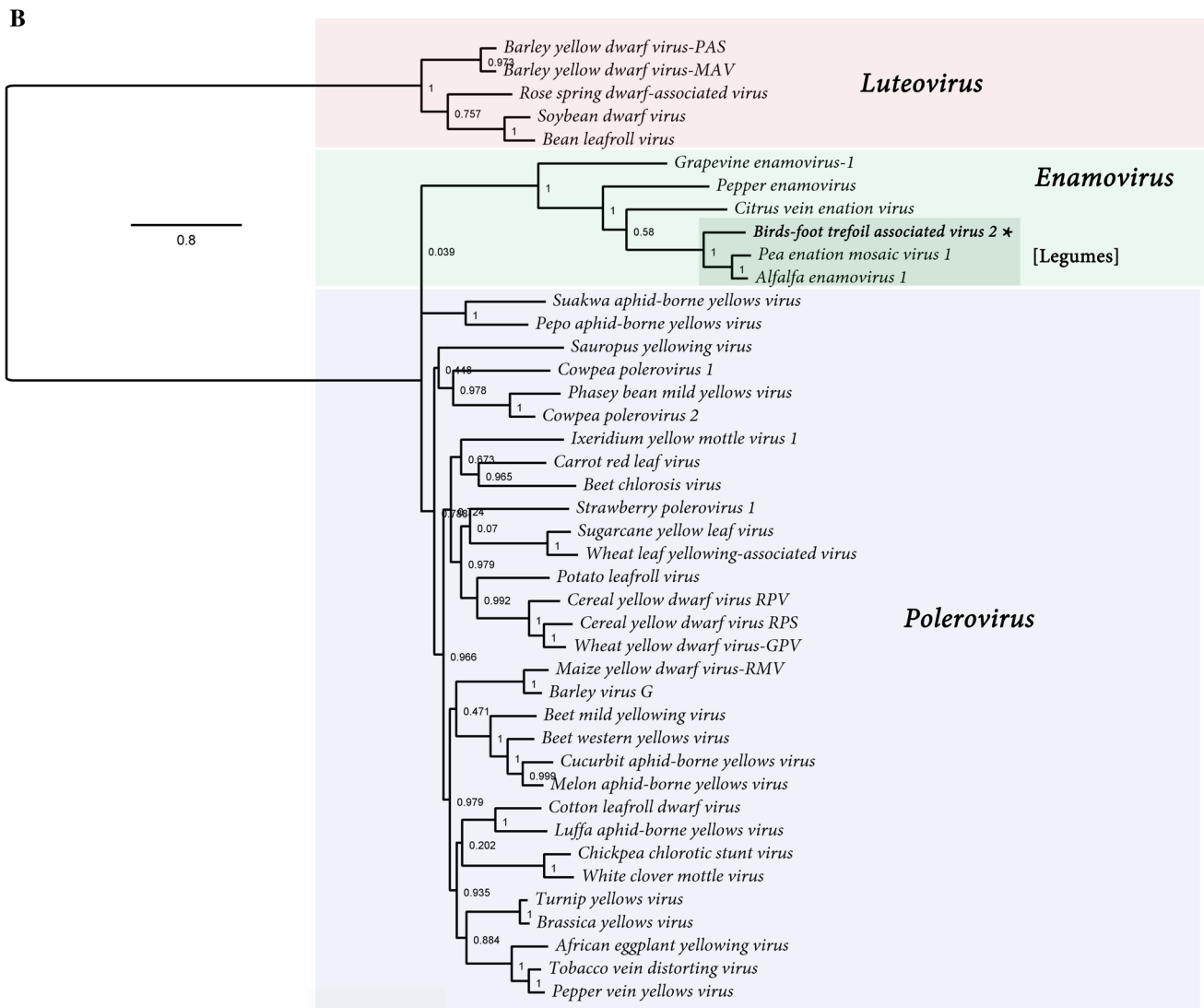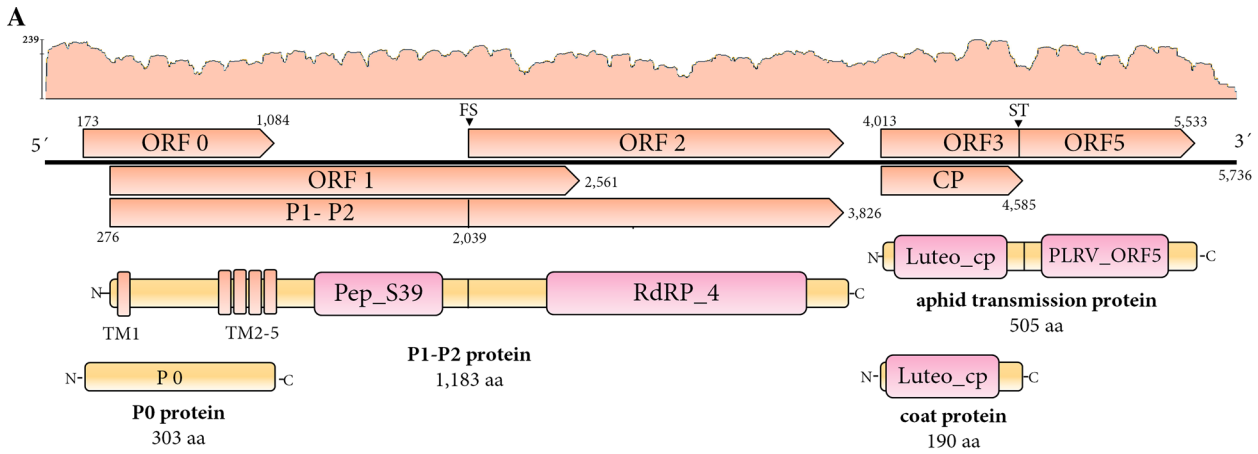
In summary, public SRA data represent an emerging potential source of genome sequences of novel plant viruses. By analyzing such data, we were able to identify and perform a molecular characterization of two novel viruses associated with bird's-foot trefoil. Our *in silico* investigation of NGS data permitted the reconstruction of coding-complete sequences with UTRs of typical length. However, based on the intrinsic effectiveness but restricted nature of our approach, these virus

◄**Fig. 2** Structural characterization and phylogenetic analysis of bird's-foot trefoil-associated virus 2 (BFTV-2) (**A**) Genome graphs depicting coverage, architecture, and predicted gene products of BFTV-2. The predicted coding sequences are shown in orange arrow rectangles, and start and end coordinates are indicated. Gene products are depicted in curved yellow rectangles, and the size in aa is indicated below the diagram. Predicted domains or HHPred best-hit regions are shown in curved pink rectangles. Abbreviations: CP, coat protein CDS; P1-P2, RNA-dependent RNA polymerase fusion protein CDS; FS, -1 ribosomal frameshifting signal; TM1-5, transmembrane domains; ST, signal of translation readthrough of a UGA stop codon; Pep_S39, peptidase S39 domain; RdRP_4, RdRP domain; Luteo_cp, luteovirid coat protein domain; PLRV_ORF5, polerovirus readthrough protein domain. (**B**) Maximum-likelihood phylogenetic tree based on amino acid sequence alignments of the P1-P2 polymerase of BFTV-2 and other enamoviruses. Members of the genus *Luteovirus* was used as the tree root; *Enamovirus*, *Polerovirus* and *Luteovirus* clades are indicated by green, blue, and pink rectangles, respectively. The scale bar indicates the number of substitutions per site. Node labels indicate FastTree support values. The viruses used to construct the tree and their accession numbers are provided in Supplementary Table 4

sequences should be presumed to have incomplete termini. Our analysis suggested that these viruses should be considered novel members of the genera *Nucleorhabdovirus* (BFTV-1) and *Enamovirus* (BFTV-2). The RNA virus sequences reported here provide, for the first time a glimpse of the virus landscape of this important crop. More importantly, the identification of these novel viruses provides evidence of a subclade of closely related and co-divergent legume-associated enamoviruses and provides new information about the evolutionary history of this genus of viruses. Future studies be done to should assess the prevalence of BFTV1 and BFTV-2 and investigate whether infection with these novel viruses is associated with specific symptoms.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Barry JK, Miller WA (2002) A − 1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. Proc Natl Acad Sci USA 99:11133–11138
2. Bejerman N, Giolitti F, de Breuil S, Trucco V, Dietzgen RG, Lenardon S (2016) Complete genome sequence of a new enamovirus from Argentina infecting alfalfa plants showing dwarfism symptoms. Arch Virol 161:2029–2032
3. Brown CM, Dinesh-Kumar SP, Miller WA (1996) Local and distant sequences are required for efficient readthrough of the barley yellow dwarf virus PAV coat protein gene stop codon. J Virol 70:5884–5892
4. Debat HJ (2017) An RNA virome associated to the Golden orb-weaver Spider Nephila clavipes. Front Microbiol 8:2097
5. 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
6. Dietzgen RG, Innes DJ, Bejerman N (2015) Complete genome sequence and intracellular protein localization of Datura yellow vein nucleorhabdovirus. Virus Res 205:7–11
7. Dietzgen RG, Kondo H, Goodin MM, Kurath G, Vasilakis N (2017) The family Rhabdoviridae: mono-and bipartite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. Virus Res 227:158–170
8. Domier LL (2012) Family Luteoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus taxonomy, ninth report of the international committee on taxonomy of viruses. Elsevier, Oxford, pp 1045–1053
9. Escaray FJ, Menendez AB, Garriz A, Pieckenstain FL, Estrella MJ, Castagno LN, Carrasco P, Sanjuan J, Ruiz OA (2012) Ecological and agronomic importance of the plant genus Lotus. Its application in grassland sustainability and the amelioration of constrained and contaminated soils. Plant Science 182:121e133
10. Firth AE, Brierley I (2012) Non-canonical translation in RNA viruses. J Gen Virol 93:1385–1409
11. Fusaro AF, Correa RL, Nakasugi K, Jackson C, Kawchuk L, Vaslin MF, Waterhouse PM (2012) The Enamovirus P0 protein is a silencing suppressor which inhibits local and systemic RNA silencing. Virol 426:178–187
12. Goodin MM, Austin J, Tobias R, Fujita M, Morales C, Jackson AO (2001) Interactions and nuclear import of the N and P proteins of sonchus yellow net virus, a plant nucleorhabdovirus. J Virol 75:9393–9406
13. Hannaway DB, Myers D (2004) Birdsfoot trefoil (*Lotus corniculatus* L.). Species selection information system. Oregon State University, Corvallis
14. Heaton LA, Hillman BI, Hunter BG, Zuidema D, Jackson AO (1989) Physical map of the genome of sonchus yellow net virus, a plant rhabdovirus with six genes and conserved gene junction sequences. Proc Natl Acad Sci USA 86:8665–8668
15. Kim H, Park D, Hahn Y (2018) Identification of novel RNA viruses in alfalfa (Medicago sativa): an alphapartitivirus, a deltapartitivirus, and a marafivirus. Gene 638:7–12
16. McDonald JG, Susuki M (1983) Occurrence of alfalfa mosaic virus in Prince Edward Island. Canadian Plant Disease Survey 63:47–50
17. Nibert ML, Vong M, Fugate KK, Debat HJ (2018) Evidence for contemporary plant mitoviruses. Virol 518:14–24
18. Roossinck MJ, Martin DP, Roumagnac P (2015) Plant virus metagenomics: advances in virus discovery. Phytopathol 105:716–727

19. Silva JMF, Al Rwahnih M, Blawid R, Nagata T, Fajardo TVM (2017) Discovery and molecular characterization of a novel enamovirus, Grapevine enamovirus-1. Virus Genes 53:667–671

20. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, Hull R (2017) Consensus statement: virus taxonomy in the age of metagenomics. Nat Rev Microbiol 15:161

21. Singh KS, Beadle K, Troczka BJ, Field L, Davies E, Williamson M, Bass C (2018) Extension of partial gene transcripts by iterative mapping of RNA-Seq raw reads. IEEE/ACM Trans Comput Biol Bioinform. https://doi.org/10.1109/tcbb.2018.2865309

22. 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. Phytopathol 103:1077–1086

23. Wang Y, Hua W, Wang J, Hannoufa A, Xu Z, Wang Z (2013) Deep sequencing of *Lotus corniculatus* L. reveals key enzymes and potential transcription factors related to the flavonoid biosynthesis pathway. Mol Genet Genom 288:131–139

24. Wu LP, Yang T, Liu HW, Postman J, Li R (2018) Molecular characterization of a novel rhabdovirus infecting blackcurrant identified by high-throughput sequencing. Arch Virol 162:2493–2494

25. Xu Y, Ju HJ, DeBlasio S, Carino EJ, Johnson R, MacCoss M, Heck MC, Miller WA, Gray SM (2018) A stem-loop structure in Potato leafroll virus ORF5 that is essential for read through translation of the coat protein ORF stop codon 700 bases upstream. J Virol 92:e01544-17