Journal of Virology®

GENOME REPLICATION AND REGULATION OF VIRAL GENE EXPRESSION

Opium Poppy Mosaic Virus Has an Xrn-Resistant, Translated Subgenomic RNA and a BTE 3′ CITE

Muhammad Ilyas,[a]* Zhiyou Du,[a]* Anne E. Simon[a]

[a]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland, USA

Muhammad Ilyas and Zhiyou Du contributed equally to this work. Authorship was determined by the most recent contributor to the project listed first.

ABSTRACT *Opium poppy mosaic virus* (OPMV) is a recently discovered umbravirus in the family *Tombusviridae*. OPMV has a plus-sense genomic RNA (gRNA) of 4,241 nucleotides (nt) from which replication protein p35 and p35 extension product p98, the RNA-dependent RNA polymerase (RdRp), are expressed. Movement proteins p27 (long distance) and p28 (cell to cell) are expressed from a 1,440-nt subgenomic RNA (sgRNA2). A highly conserved structure was identified just upstream from the sgRNA2 transcription start site in all umbraviruses, which includes a carmovirus consensus sequence, denoting generation by an RdRp-mediated mechanism. OPMV also has a second sgRNA of 1,554 nt (sgRNA1) that starts just downstream of a canonical exoribonuclease-resistant sequence ($xrRNA_D$). sgRNA1 codes for a 30-kDa protein *in vitro* that is in frame with p28 and cannot be synthesized in other umbraviruses. Eliminating sgRNA1 or truncating the p30 open reading frame (ORF) without affecting p28 substantially reduced accumulation of OPMV gRNA, suggesting a functional role for the protein. The 652-nt 3′ untranslated region of OPMV contains two 3′ cap-independent translation enhancers (3′ CITEs), a T-shaped structure (TSS) near its 3′ end, and a *Barley yellow dwarf virus*-like translation element (BTE) in the central region. Only the BTE is functional in luciferase reporter constructs containing gRNA or sgRNA2 5′ sequences *in vivo*, which differs from how umbravirus 3′ CITEs were used in a previous study. Similarly to most 3′ CITEs, the OPMV BTE links to the 5′ end via a long-distance RNA-RNA interaction. Analysis of 14 BTEs revealed additional conserved sequences and structural features beyond the previously identified 17-nt conserved sequence.

IMPORTANCE *Opium poppy mosaic virus* (OPMV) is an umbravirus in the family *Tombusviridae*. We determined that OPMV accumulates two similarly sized subgenomic RNAs (sgRNAs), with the smaller known to code for proteins expressed from overlapping open reading frames. The slightly larger sgRNA1 has a 5′ end just upstream from a previously predicted $xrRNA_D$ site, identifying this sgRNA as an unusually long product produced by exoribonuclease trimming. Although four umbraviruses have similar predicted $xrRNA_D$ sites, only sgRNA1 of OPMV can code for a protein that is an extension product of umbravirus ORF4. Inability to generate the sgRNA or translate this protein was associated with reduced gRNA accumulation *in vivo*. We also characterized the OPMV BTE structure, a 3′ cap-independent translation enhancer (3′ CITE). Comparisons of 13 BTEs with the OPMV BTE revealed additional stretches of sequence similarity beyond the 17-nt signature sequence, as well as conserved structural features not previously recognized in these 3′ CITEs.

KEYWORDS 3′ CITE, BTE, cap-independent translation enhancer, exoribonuclease-resistant sites, umbravirus, Xrn1, subgenomic RNA

All assigned and likely members of the genus umbravirus (family *Tombusviridae*) (see Fig. 1 for a phylogenetic comparison) have a single, positive-sense genomic

Citation Ilyas M, Du Z, Simon AE. 2021. *Opium poppy mosaic virus* has an Xrn-resistant, translated subgenomic RNA and a BTE 3′ CITE. J Virol 95:e02109-20. https://doi.org/10.1128/JVI.02109-20.

Editor Rebecca Ellis. Dutch, University of Kentucky College of Medicine

Address correspondence to Anne E. Simon, simona@umd.edu.

* Present address: Muhammad Ilyas, USDA-ARS Beltsville, Maryland, USA; Zhiyou Du, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou, China.

Received 28 October 2020
Accepted 8 February 2021
Accepted manuscript posted online 17 February 2021
Published 12 April 2021

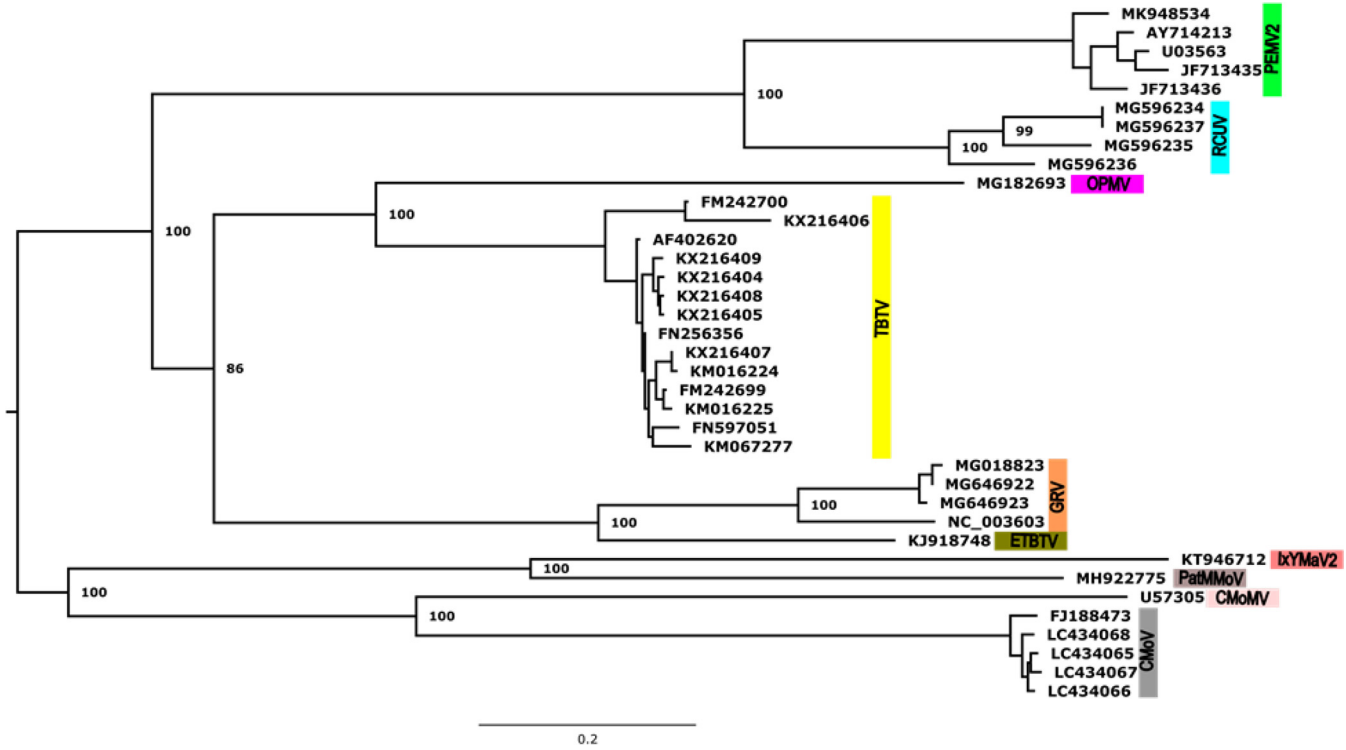


FIG 1 Maximum-likelihood analysis of the umbravirus genus based on 37 full-length umbravirus gRNA sequences available in GenBank. Viruses are highlighted in different colors. GenBank accession numbers are indicated. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to each branch. Vertical branches are arbitrary, and horizontal branches are proportional to calculated mutation distances. The tree is rooted at the midpoint. CMoMV, *Carrot mottle mosaic virus*; CMoV, *Carrot mottle virus*; ETBTV, *Ethiopian tobacco bushy top virus*; GRV, *Groundnut rosette virus; IxYMoV2*, *Ixeridium yellow mottle-associated virus 2*; OPMV, *Opium poppy mosaic virus*; PEMV2, *Pea enation mosaic virus 2*; PatMMoV, *Patrinia mild mottle virus*; RCUV, *Red clover umbravirus*; TBTV, *Tobacco bushy top virus*.

RNA (gRNA) of 4.0 to 4.3 kb with four open reading frames (ORFs). ORF1 encodes a replication-required protein in related viruses, and ORF2 encodes the RNA-dependent RNA polymerase (RdRp), which is synthesized following a programmed −1 ribosomal frameshift (−1 PRF) just upstream of the ORF1 termination codon (1, 2). ORF3 codes for a multifunctional, nucleocapsid-like long-distance movement protein that interacts with fibrillarin in the nucleus and coats the viral RNA protecting it from RNases and possibly RNA interference (RNAi) (3). In addition, ORF3 confers protection against nonsense-mediated decay (NMD), a cellular process that targets translated RNAs with long 3′ untranslated regions (UTRs) (4, 5). ORF4, which overlaps with ORF3, codes for the cell-to-cell movement protein (Fig. 2A) (1). Umbraviruses do not encode their own capsid protein or silencing suppressor and require coinfection with a helper virus for natural infections. In contrast, some umbraviruses are able to unilaterally infect laboratory hosts, including *Nicotiana benthamiana*, following mechanical inoculation or agrobacteria infiltration (6, 7).

Umbravirus translation has only been studied in detail for *Pea enation mosaic virus* 2 (PEMV2) and, to a lesser extent, for *Tobacco bushy top virus* (TBTV). PEMV2 gRNA expresses ORF1 and ORF2, and ORF3 and ORF4 are translated from a 3′ coterminal subgenomic RNA (sgRNA) by an unusual mechanism that does not involve leaky scanning (8, 9). Umbraviruses lack a 5′ cap and 3′ poly(A) tail, and like many uncapped plant viral RNAs, they harbor at least one cap-independent translation enhancer (CITE) in their 3′ UTR (2, 10–13). 3′ CITEs identified so far in plant plus-strand RNA viruses have been divided into seven distinct classes based on their sequences and secondary structures, with little discernible relationship among the classes (14–16). 3′ CITEs interact with one or more components required for translation initiation, and nearly all contain a sequence, usually found in the apical loop of a short hairpin, that is known or predicted to base pair with complementary sequence in or near the 5′ UTR. This long-distance

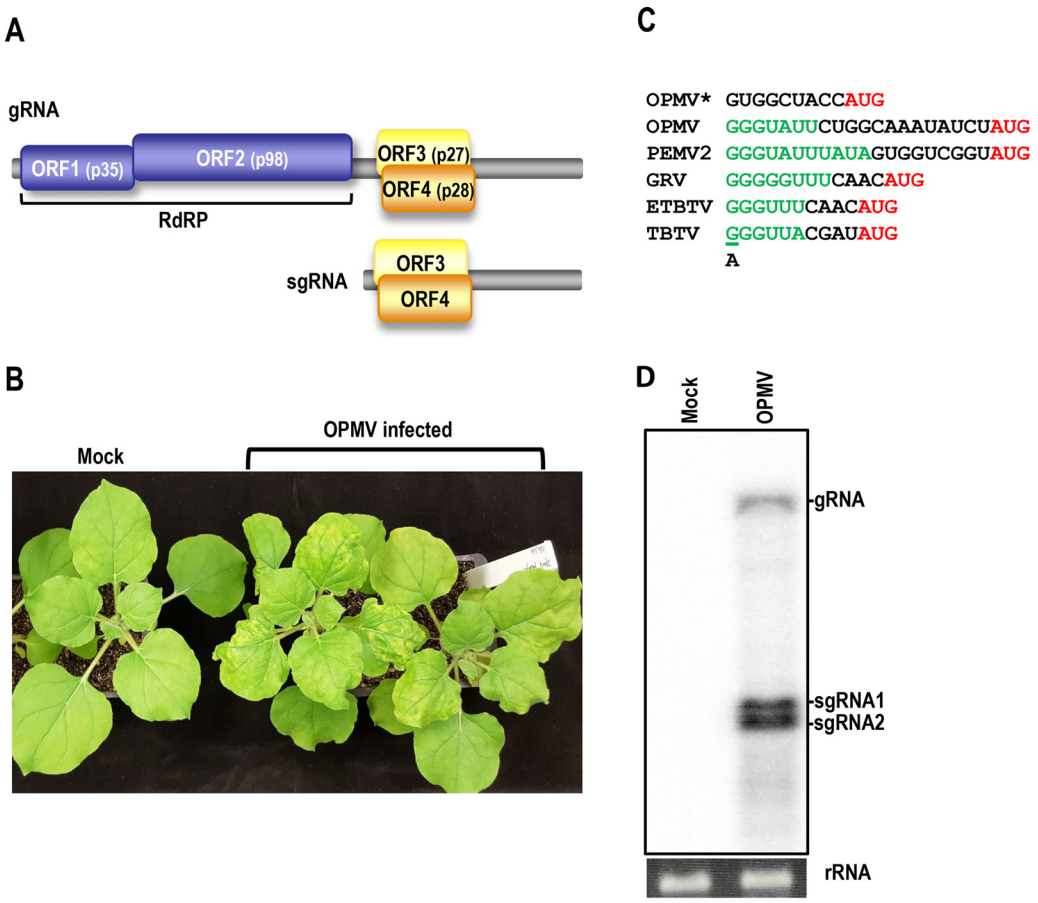


FIG 2 Basic properties of OPMV. (A) Genome organization of umbraviruses. Sizes of OPMV-encoded products are in parentheses. (B) Two *N. benthamiana* plants infected with OPMV and its uncharacterized helper virus showing mild leaf curling and yellowing. Mock, plant treated with inoculation buffer. (C) 5′ UTRs of OPMV and closely related umbraviruses. RCUV is not included, as a 5′ UTR sequence was not available. Carmovirus consensus sequence is in green and the ORF1 start codon is in red. OPMV*, previously reported OPMV 5′ UTR sequence. TBTV 5′ UTR starts with either G or A, depending on the isolate. (D) Northern blot of total RNA isolated from systemically infected *N. benthamiana* leaves following agroinfiltration with OPMV. Mock, infiltrated with empty binary vector. Ethidium bromide-stained 28S rRNA was used as a loading control.

interaction ensures that translation initiation complexes that organize at the 3′ end are relocated to the 5′ end, where translation initiates (17). Long-distance interactions connecting 3′ CITEs with the 5′ ends of gRNAs and/or sgRNAs have been confirmed for the Barley yellow dwarf virus-like translation element (BTE) of *Barley yellow dwarf virus* (BYDV), which interacts with translation initiation factor eIF4G (18, 19); Panicum mosaic virus-like translational enhancer (PTE) (20, 21), which interacts with eIF4E (22); TED of *Pelargonium line pattern virus* (23), which interacts with eIF4F (24); Y-shaped structure (YSS) of *Carnation Italian ringspot virus* (25, 26) and *Tomato bushy stunt virus* (27), which interact with eIF4F (25); I-shaped structure (ISS) of *Maize necrotic streak virus* (28), which interacts with eIF4F (29); and the kissing-loop T-shaped structure (kl-TSS) of PEMV2, which interacts with ribosomal subunits (12, 13).

Umbraviruses have a variety of known or putative 3′ CITEs in their 3′ UTRs, located both distally and proximally to the 3′ end. Three 3′ CITEs have been identified in the 3′ UTR of PEMV2, as follows: (i) a PTE that is not capable of an independent long-distance interaction; (ii) a kl-TSS adjacent to the PTE that forms the bridge with the 5′ end (12, 13); and (iii) a T-shaped structure (TSS), which folds into a form similar to tRNAs (11, 30, 31) and binds 60S ribosomal subunits (32). PEMV2 requires all three 3′ CITEs for translation of sgRNA reporter constructs *in vivo*, but only uses the kl-TSS and PTE for similar translation of gRNA reporter constructs (9, 33).

TBTV contains a BTE in a similar location as the kl-TSS and PTE of PEMV2 (10). In addition to umbraviruses, BTEs are also found in dianthoviruses, alphanecroviruses, and betanecroviruses in the family *Tombusviridae* and luteoviruses in the family *Luteoviridae* (15, 34). BTEs have three to six helices protruding from a central hub and can be recognized by an essential 17-nt conserved sequence, GAUCCCUGGGAAACAGG, that forms stem-loop 1 (SL1) by pairing the underlined bases (15). Around the hub are a number of nonconserved bases proposed to link the radiating helices (15, 18). Nucleotide changes outside the TBTV BTE caused structural changes in the BTE, resulting in reduced pathogenesis and indicating that CITE-surrounding regions are also important for enhancer efficiency (10).

Exoribonuclease-resistant RNAs (xrRNAs), which are discrete structures that can generate sgRNAs by blocking the progression of 5′ to 3′ cellular exoribonucleases, have been found in viruses from the *Flaviviridae*, *Luteoviridae*, and *Tombusviridae* families (35–42). Several umbraviruses were recently identified as containing one of the known exoribonuclease-resistant RNA consensus structures ($xrRNA_D$) in two locations (intergenic and 3′ UTR) (35, 39). $xrRNA_D$ sites are located upstream of the known sgRNA transcription start site in *Opium poppy mosaic virus* (OPMV), *Carrot mottle mimic virus* (CMoMV), and TBTV, but not *Groundnut rosette virus* (GRV), which was previously identified as having two similarly sized, uncharacterized sgRNAs (1, 7). Production of lengthy sgRNAs by exoribonuclease trimming would be unusual, and has led to speculation that such sgRNAs may be templates for translation and not simply function as noncoding RNAs (34). For this report, we developed OPMV as an additional system to study umbravirus sgRNAs and translation. We determined that OPMV accumulates two prominent, similarly sized sgRNAs. sgRNA2, the most prevalent sgRNA, codes for p27 and p28 movement proteins and is the orthologue of the sgRNA previously characterized in PEMV2. The slightly larger sgRNA1 has a 5′ end just upstream from the $xrRNA_D$ site, identifying this sgRNA as an exoribonuclease-resistant product. sgRNA1 expressed a p28 extension protein *in vitro* (p30) that cannot be synthesized by other umbraviruses, and inability to translate this protein was associated with reduced gRNA accumulation *in vivo*. OPMV has both a BTE 3′ CITE and the elements necessary to form a 3′ proximal TSS. Efficient translation of OPMV gRNA and sgRNA2 reporter constructs required a long-distance interaction between the BTE and 5′ UTR sequences. Mutational analysis indicated that, unlike PEMV2, only the BTE, and not the TSS, functions as a translation enhancer for both gRNA and sgRNA reporter constructs translated *in vivo*. Comparisons of 13 BTEs with the OPMV BTE revealed additional stretches of sequence similarity beyond the 17-nt signature sequence, as well as conserved structural features not previously recognized in these 3′ CITEs.

RESULTS AND DISCUSSION

Production of an infectious clone of OPMV. Leaf sap from a native poppy plant infected with OPMV and its uncharacterized helper virus (kindly provided under permit by Joe Tang, Plant Health and Environment Laboratory, Auckland, New Zealand) was mechanically inoculated to *N. benthamiana* leaves, and mild curling and yellowing of leaves was observed after 2 weeks (Fig. 2B). Total RNA isolated from infected plants along with OPMV primers specific for sequences at the 5′ and 3′ ends (from the full-length genomic sequence [GenBank accession number EU151723]) were used to synthesize full-length OPMV cDNA by reverse transcriptase PCR (RT-PCR). The cDNA was cloned downstream of the CaMV 35S promoter in an *Agrobacterium tumefaciens* binary vector, producing construct pCB301-OPMV*. Agroinfiltration of pCB301-OPMV* into *N. benthamiana* did not induce any visible symptoms, and OPMV gRNA and sgRNA could not be detected by Northern blots of total RNA extracted from infiltrated leaves (data not shown).

Sequence analysis of some closely related umbraviruses (GRV, PEMV2, TBTV, and *Ethiopian tobacco bushy top virus* [ETBTV]), revealed that most of their gRNA 5′ UTRs begin with a carmovirus consensus sequence (CCS) ($G_{1-3}A/U_{3-9}$; most commonly starts with three guanylates [43]). CCS are found at the 5′ ends of nearly all alphacarmovirus,

betacarmovirus, and gammacarmovirus gRNA and sgRNAs, as well as those of many other members of the *Tombusviridae*, but was absent from the reported full-length OPMV sequence (Fig. 2C). To determine if OPMV in the sap-inoculated plants had a different 5′ end sequence from the one reported, 5′ rapid amplification of cDNA ends (5′ RACE) was carried out on RNA isolated from the originally infected *N. benthamiana* plants. The RACE-derived sequence had both different and missing residues at the 5′ end (Fig. 2C), and included a canonical CCS. When the RACE-derived sequence was incorporated into the binary vector construct, followed by agroinfiltration into *N. benthamiana*, OPMV gRNA was detected 3 days later in infiltrated leaves by Northern blotting analysis of total extracted RNA (Fig. 2D). The sequence of the infectious OPMV construct is available in GenBank under accession number MG182693.

Mapping the 5′ ends of OPMV sgRNAs. Two prominent OPMV sgRNAs were also present along with the gRNA, which were labeled sgRNA1 and sgRNA2 (Fig. 2D). To map the transcription start sites of the two OPMV sgRNAs, total RNA was isolated from upper systemic leaves of *N. benthamiana* inoculated with sap of the original dried leaves. The RNA was used as the template for primer extension using a primer complementary to positions 2849 to 2865 of OPMV gRNA. As shown in Fig. 3A, reverse transcription reactions generated two strong-stop products. One extension product terminated with the cytidylate at position 2679, corresponding to a 3′ coterminal sgRNA of 1,554 nt (sgRNA1), and the second terminated with the guanylate at position 2791, corresponding to a 3′ coterminal sgRNA of 1,440 nt (sgRNA2).

The 5′ end sgRNA1, located in the intergenic region (IR) between the end of the RdRp ORF and the movement protein ORFs, was positioned two residues upstream of a canonical xrRNA$_D$ site that was predicted previously based on the presence of consensus sequences and structures (39). xrRNA$_D$ sites and other identified sites in animals and plants restrict movement of the 5′-to-3′ exoribonuclease Xrn1 (in mammals) or presumably Xrn4 (in plants) (44), which generates sgRNAs that function beneficially in virus accumulation, movement, and pathogenicity (41, 45–47). As shown in Fig. 3B, the OPMV xrRNA$_D$ structure contains the two required hairpins, the critical pseudoknot, and all highly conserved residues. The only difference between the OPMV xrRNA$_D$ and the consensus structure (39) is that the OPMV element has a 3 bp lower stem in the 5′ hairpin instead of the conserved 5 bp stem. Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) RNA structure mapping of the IR in full-length OPMV additionally supported the presence of xrRNA$_D$ (Fig. 3B). Other umbraviruses with a similarly placed, previously predicted xrRNA$_D$ are TBTV, CMoMV, and *Ixeridium yellow mottle-associated virus 2* (IxYMaV2) (39) (Fig. 4).

Curiously, GRV, the only other umbravirus identified with two closely spaced sgRNAs (1, 7), did not have a predicted xrRNA$_D$ in its IR. To determine if any sequences and/or structures are conserved in the IR of umbraviruses without xrRNA$_D$ sites that might correspond to a different exonuclease-resistant site, an IR alignment was generated for all umbraviruses (Fig. 5). IR length among umbraviruses was highly variable, ranging from 130 nt (*Patrina mild mottle virus* [PatMMoV]) to 337 nt (*Red clover umbravirus* [RCUV]). Some of the size variability stemmed from duplication and triplication of segments of less than 40 nt in length, with many IRs sharing common sequence stretches. Interestingly, all duplications and all shared sequence stretches (indicated by the horizontal arrows) were in umbraviruses that lacked a predicted xrRNA$_D$ element. The most striking feature found in these other umbravirus IRs was the scattered presence of similar hairpins with 5- or 6-base-pair stems (underlined in red) and "AAUHGA" (H = A or C or U) in the apical loop (highlighted in yellow) or UGGCU (highlighted in orange). All duplicated sequences included one of these hairpins, suggesting that extra copies of this element have a beneficial function. These hairpins and repeated sequences may comprise part of an alternative exonuclease-resistant sequence, as similarly sized small hairpin repeats were determined to be responsible for generation of some of the exoribonuclease-resistant subgenomic flaviviral RNAs (sfRNAs) in flaviviruses

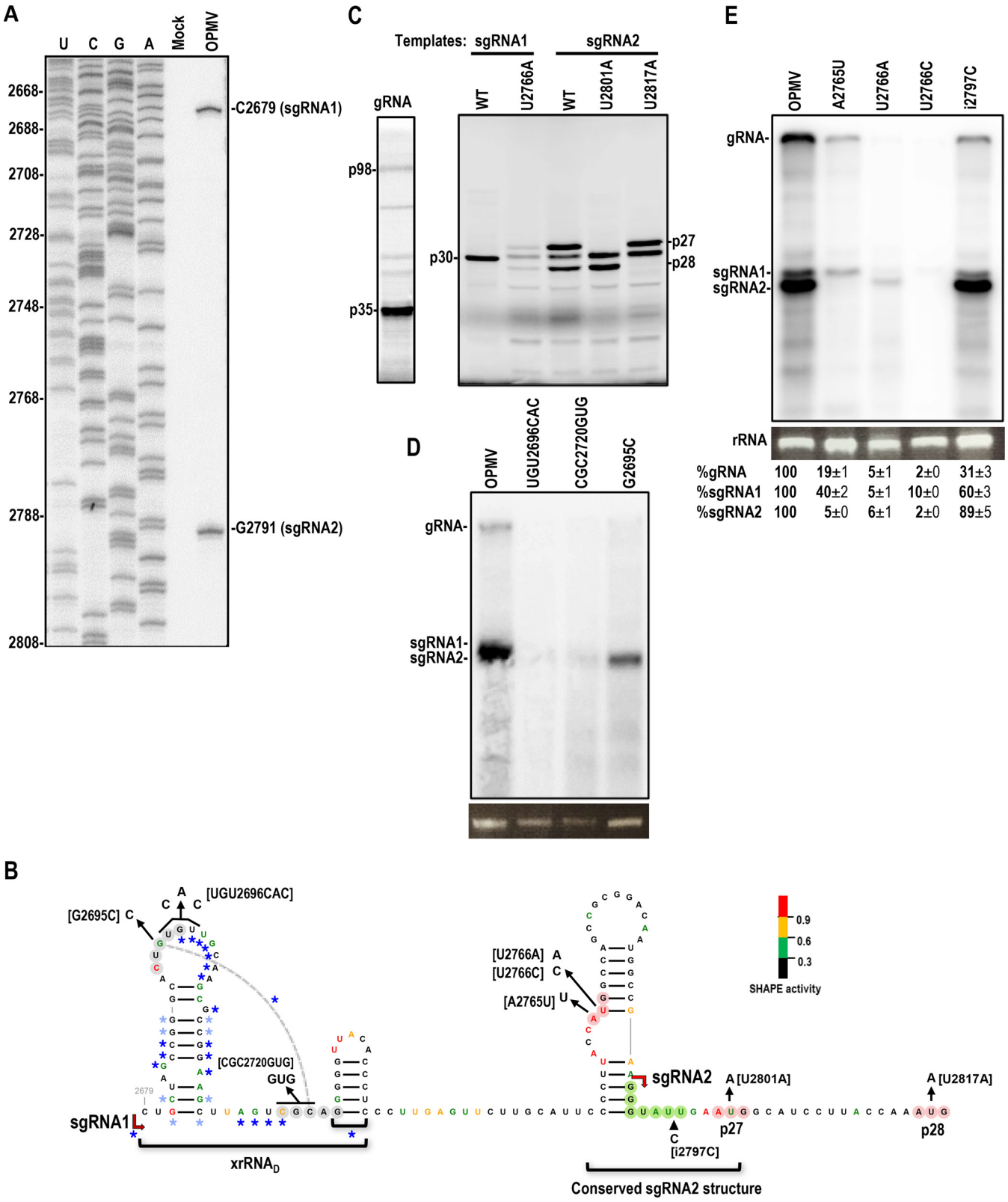


FIG 3 sgRNAs of OPMV and their encoded products. (A) Primer extension assay to map the transcription start sites of OPMV sgRNAs. Total RNA isolated from *N. benthamiana* that was either infected with OPMV (and its uncharacterized helper virus) or treated with transfection buffer (mock) was used. The first four lanes (U, C, G, and A) are sequencing ladders. Positions of the extension products are given. (B) Sequence and structure of a portion of the intergenic region of OPMV showing 5′ ends of the sgRNAs and their encoded products. Mutations used in panels B, C, and D are indicated, with construct names in brackets. Start codons for p27, p28, and p30 are shaded red. Bases are colored according to their *N*-methylisatoic anhydride (NMIA) reactivity in

(Continued on next page)

(48). Interestingly, the presence or absence of an xrRNA$_D$ element in an umbravirus is independent of their relationship within the phylogenetic tree shown in Fig. 1.

All umbraviruses have a CCS (shaded green) in positions comparable to OPMV position 2791, which corresponds to the 5′ end of known or putative sgRNA2 (8) (Fig. 5). Just upstream of each CCS was a structurally conserved hairpin consisting of two stems separated by an asymmetric interior loop, with the lower stem encompassing part of the CCS (Fig. 6). The upper portion of the hairpin and apical loop exhibited significant sequence conservation in 7 umbraviruses, including OPMV, with related PatMMoV and IxYMaV2 hairpins also sharing similar sequences in the upper stem. sgRNAs with 5′ CCS are thought to be generated by RdRp initiation on truncated minus-strand templates (49, 50), strongly suggesting that the two closely spaced sgRNAs in OPMV (and in likely GRV) are produced by different mechanisms.

Proteins translated by the OPMV sgRNAs *in vitro*. In PEMV2, the sgRNA corresponding to OPMV sgRNA2 is the bicistronic template for translation of ORF3 (p26) and ORF4 (p27) (8, 9). Since the positions and spacing of the start codons are highly conserved in umbraviruses (8), OPMV sgRNA2 should be the template for p27 and p28. OPMV sgRNA1 has a possible AUG translation initiation codon in a very good Kozak context located 86 nt downstream of its 5′ end (Fig. 5B). If ribosomes initiate translation at this AUG, a 30-kDa protein would be produced that contains a 17-amino-acid N-terminal extension of p28 (Fig. 5C).

To determine the *in vitro* translation products of OPMV gRNA, sgRNA1 and sgRNA2, *in vitro* synthesized gRNA and sgRNAs were subjected to translation in wheat germ extracts (WGE). As with PEMV2, OPMV gRNA was the template for ORF1 (p35), and additional slower-migrating products, including one that likely corresponds to ORF2 (p98), the −1 PRF product (Fig. 3C). OPMV sgRNA2 was unexpectedly the template for three major products. To determine which of these products corresponds to p27 and p28, both initiation codons were separately altered to AAG in sgRNA2 (Fig. 3B), followed by *in vitro* translation. When the p27 start codon was changed to AAG (U2801A), the slowest-migrating product was absent, defining this product as p27. When the p28 start codon was changed to AAG (U2817A), the fastest-migrating product was missing. This indicates that, as with PEMV2 (8), the products of OPMV ORF3 and ORF4 migrate aberrantly on SDS-PAGE gels following translation in WGE. Since neither mutation disrupted translation of the middle product, this product is not simply a processed variant. One possibility is that it represents initiation from an AUG located 23 nt downstream of the p28 start codon, in the same frame as p27, the slower migrating protein. An initiation codon at this location is not found in other umbraviruses, and thus the importance of this truncated product, if any, is unknown.

Translation of sgRNA1 in WGE produced a single product that migrated between p27 and p28, similar in location to the unknown product of sgRNA2 (Fig. 3C). When the p30 start codon was changed to AAG (U2766A), high levels of p30 were no longer generated, and very low levels of p27, p28, and either the unknown product or p30 were detected. The second AUG from the 5′ end of sgRNA1 is out of frame with both ORF3 and ORF4, and translation from this AUG would terminate after 6 codons and would not be detected on these gels. Analysis of all other umbravirus sequences in this region revealed that none are capable of coding for a similar ORF4 extension product as in-frame stop codons are located just upstream of all other ORF4 initiation

FIG 3 Legend (Continued)
selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) structure probing, with red being the most reactive and black being the least reactive. Structure was determined using SHAPE probing data and phylogenetic comparisons (see Fig. 4 and 6). Dark blue and light blue asterisks denote conserved residues (>90% or >75%, respectively) in xrRNA$_D$ elements (39). A dashed line connects residues (shaded gray) involved in a critical pseudoknot (noted with dark blue asterisk). The structure just upstream from the sgRNA2 transcription start site is conserved in all umbraviruses (see Fig. 6). Carmovirus consensus sequence (CCS) is shaded green. (C) SDS-PAGE gels of *in vitro* translated OPMV gRNA (left) or wild-type (WT) and mutant sgRNAs (right). Note that p27 and p28 migrate aberrantly in these gels. (D and E) Northern blots of total RNA isolated from *N. benthamiana* infiltrated with either wild-type (WT) or mutant OPMV with alterations in either the xrRNA$_D$ site (D) or the p30 initiation codon (E). i2797C is a single-base insert that terminates p30 translation at the adjacent downstream UGA. 28S rRNA served as a loading control. RNA bands intensities were measured using ImageJ. Data are from three experiments, and standard deviations are shown.

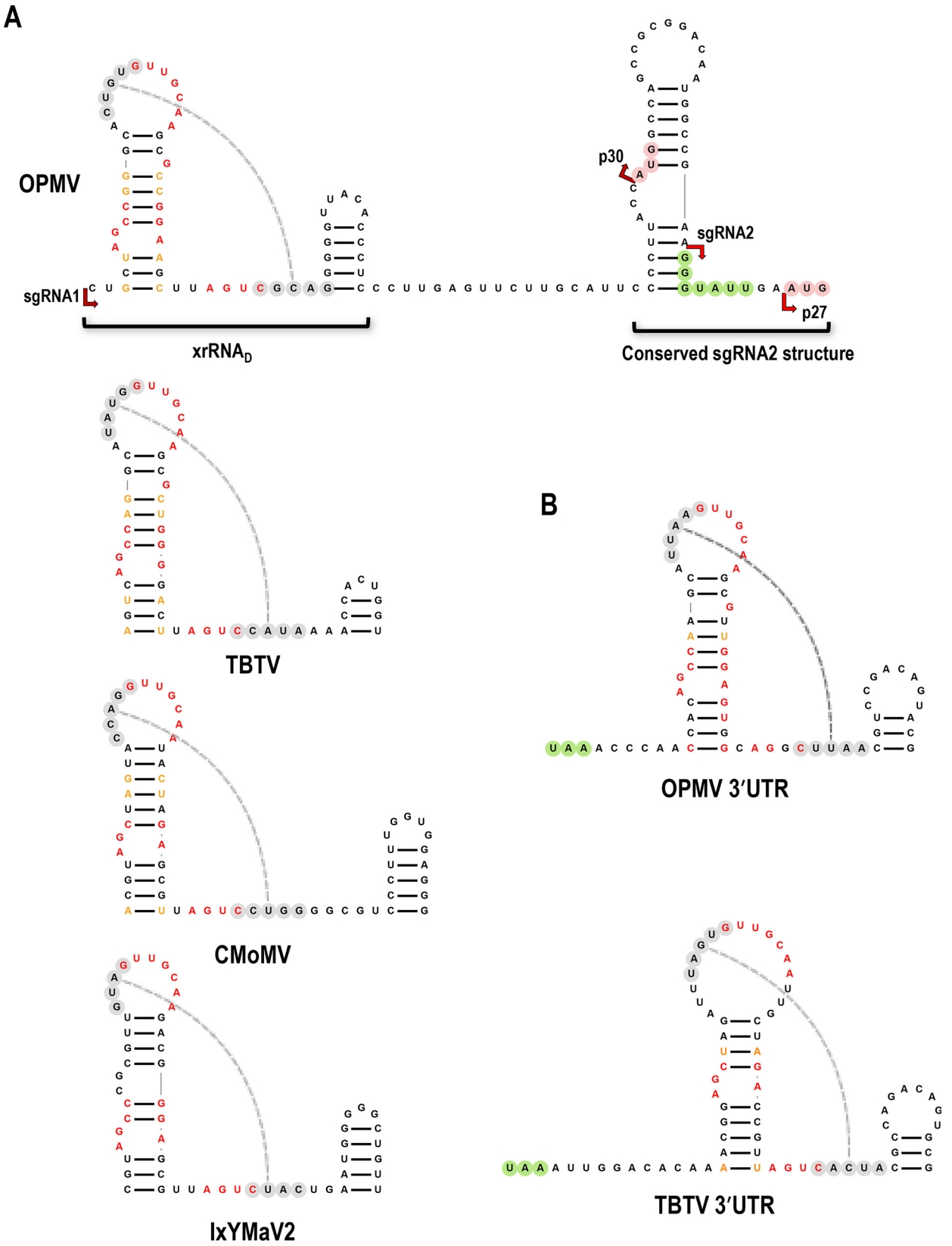


FIG 4 xrRNA$_D$ structures in umbraviruses. (A) xrRNA$_D$ structures in the intergenic region (IR) of four umbraviruses. Conserved bases are in red, and less conserved bases are in orange (39). A critical pseudoknot formed between residues shaded gray is shown by a dashed line. Start sites for the sgRNAs and some of their encoded products are shown for the OPMV structure. (B) OPMV and TBTV are the only umbraviruses with a second xrRNA$_D$ structure near the 5′ end of their 3′ UTR. ORF4 termination codon is shaded green.

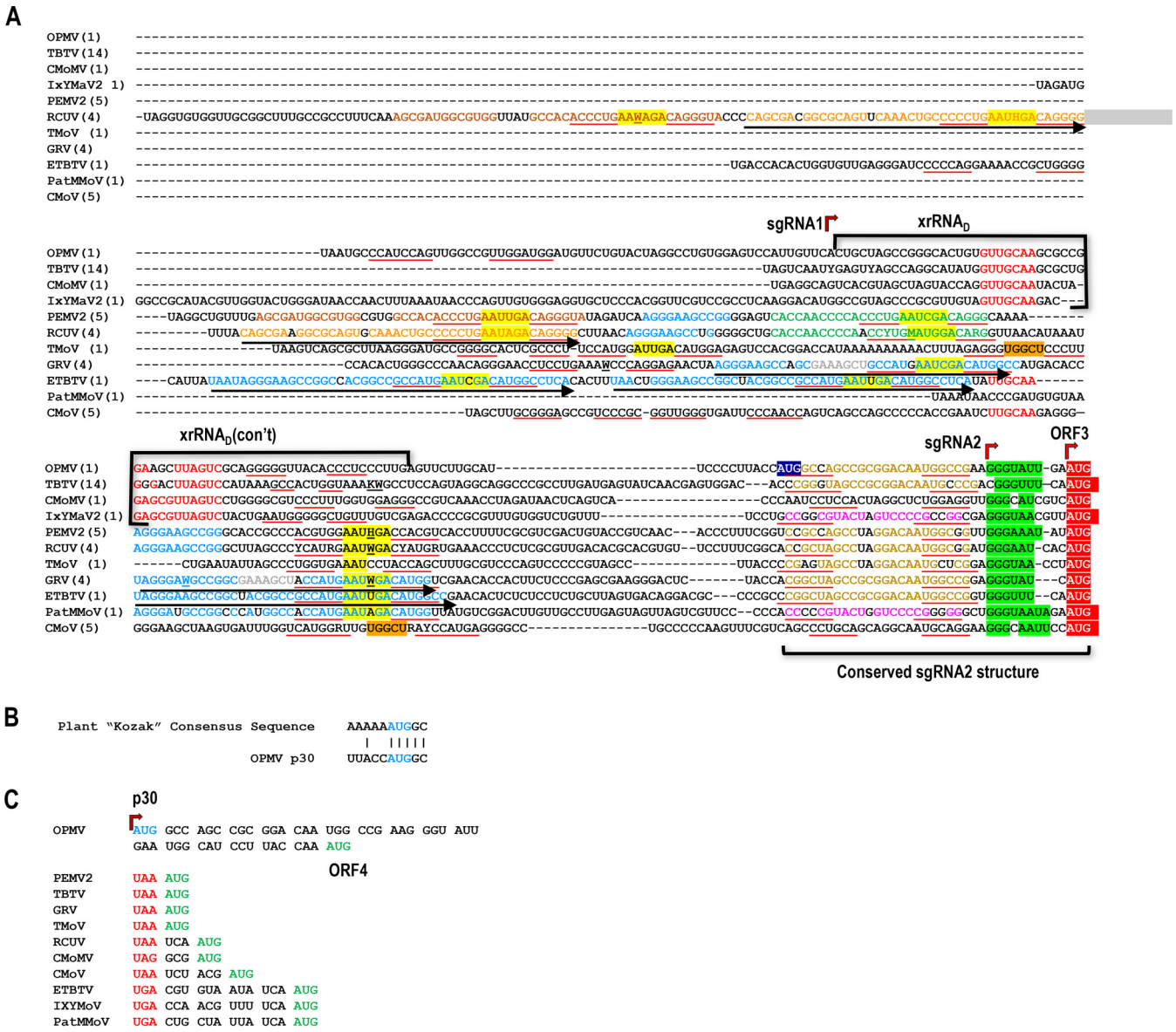


FIG 5 Umbravirus (assigned and unassigned) intergenic region (IR) alignment. (A) The numbers of isolates included in the alignment are given in parentheses. Degenerate nucleotide codes are as follows: Y (C or U), R (A or G), W (A or U), K (U or G), M (C or A), and H (A or C or U). Start sites of OPMV sgRNA1 and sgRNA2 are indicated. CCS at the beginning of OPMV sgRNA2 and equivalent sgRNAs are shaded green. Start codons of ORF3 and p30 are shaded red and blue, respectively. Stem regions of putative hairpins are underlined in red, and a highly conserved motif found in many apical loops (AAUHGA) is highlighted in yellow. A second conserved loop motif (UGGCU) is highlighted in orange. Other conserved sequences are colored alike. Duplicated stretches within the IR are denoted by arrows. The xrRNA$_D$ structure in four umbraviruses is denoted. The conserved sequence/structure in the vicinity of the sgRNA2 transcription start site is labeled, and structures are shown in Fig. 6. (B). Translation context of p30 initiation codon (blue). Critical purine at −3 and guanylate at +4 are present. (C) N-terminal sequence of p30, which is in frame with downstream ORF4. All other umbraviruses have an in-frame termination codon (in red) just upstream of ORF4 and thus cannot produce an N-terminal extended version of the ORF4 product.

codons (Fig. 5C). These results suggest that OPMV sgRNA1 may encode a unique translation product and, unlike sgRNA2, that it is not an efficient template for p27 and p28.

Biological significance of sgRNA1 *in planta*. To determine the importance of sgRNA1 for accumulation of OPMV *in planta*, three separate alterations were generated that disrupt highly conserved xrRNA$_D$ sequences (mutant G2695C) and/or the conserved pseudoknot (mutants GUG2696CAC and CGC2720GUG) (Fig. 3B). Full-length OPMV gRNA was delivered to *N. benthamiana* by agroinfiltration, and total RNA was isolated 72-h later. All mutations substantially reduced both sgRNA1 and gRNA levels, suggesting that sgRNA1 and/or its encoded product is important for OPMV replication or stability (Fig. 3D).

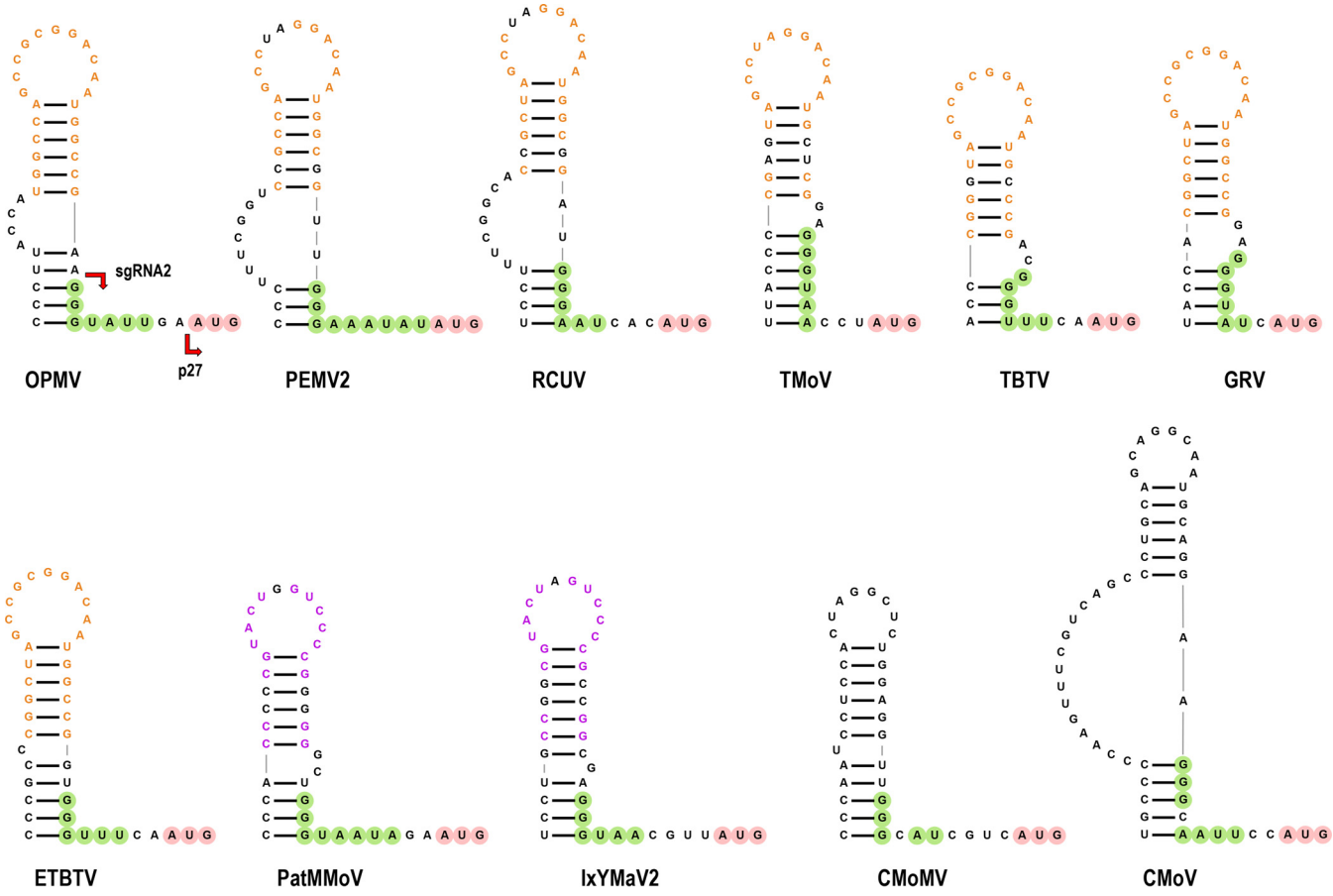


FIG 6 Proposed structure just upstream of the start site of umbravirus sgRNA2. Similar sequences are colored alike. CCS is shaded green and ORF3 start codon is shaded red.

To help determine if the encoded product is important for OPMV viability, the p30 initiation codon was altered in wild-type (WT) OPMV by changing A2765 to U (A2765U). A2765U reduced OPMV gRNA levels by 5-fold and sgRNA1 levels by only 60% (Fig. 3E). However, sgRNA2 levels were reduced by 20-fold, suggesting that A2765U may be affecting transcription of sgRNA2 by virtue of its location within the conserved element upstream of the sgRNA2 transcription start site (Fig. 3B). Since sgRNA2 encodes p27, which is necessary for efficient gRNA accumulation due to its blockage of NMD (5), reduced gRNA levels could result from decreasing the level of sgRNA2. To attempt to alter the initiation codon without negatively impacting the conserved structure that may be required for sgRNA2 synthesis, we used the observation that all other umbraviruses except for IxYMaV2 have a C:G pair located at the base of the 6-bp upper stem (Fig. 6). In OPMV, the "U" of the AUG initiation codon is paired with the G at this location. When this U was changed to a C, creating a C:G base pair at the base of the upper stem (U2766C), or to an A, creating an A-G mismatch (U2766A), accumulation of both OPMV gRNA and the sgRNAs decreased to near background levels (Fig. 3E). This result strongly suggests that a weak base pair is required at this location in OPMV for synthesis of sgRNA2. From these results, it cannot be determined if p30 is important for OPMV accumulation.

To suppress expression of p30 without changing the putative sgRNA2 promoter, a single base was inserted upstream of the p27 ORF start site (i2797C), introducing an in-frame translation termination codon that generates a truncated p30 but should not affect translation of p27 or p28 (Fig. 3B). OPMV containing i2797C decreased gRNA levels by 69% (compared with undetectable levels of gRNA when mutations disrupted the critical xnRNA$_D$ hairpin) and only had a modest effect on sgRNA2 levels (Fig. 3E).

Together, these results suggest that p30, despite its absence from other umbraviruses, may have a beneficial role in accumulation of OPMV gRNA in cells, and that sgRNA1 may have an additional beneficial function in virus replication or stability besides that of encoding p30.

OPMV gRNA and sgRNA reporter constructs. To identify translation elements in OPMV gRNA and sgRNA2, luciferase reporter constructs were generated for translation in *Arabidopsis thaliana* protoplasts. The initial gRNA reporter construct (5′20+3U) contained the OPMV 5′ UTR (20 nt) upstream of firefly luciferase (FLuc), followed by the full-length 3′ UTR (652 nt) (Fig. 7). 5′20+3U luciferase activity was approximately 40% of that produced by a previously studied PEMV2 reporter construct containing the gRNA 5′ 89 nt (the 20-nt 5′ UTR plus 69 coding nt) and the 3′ UTR (construct PEMV2-5′89+3U); this construct includes all elements necessary for the long-distance RNA-RNA interaction between the kl-TSS 3′ CITE and the 5′ end (13) (Fig. 7B). When the 5′ UTR of 5′20+3U was replaced by 18 random nucleotides (construct 5ΔU+3U), luciferase activity was reduced by 5.7-fold, suggesting that at least one important element is likely partially or fully within the OPMV 5′ UTR (Fig. 7C). To determine if additional downstream coding residues improve translation efficiency, constructs containing 38, 98, and 130 gRNA 5′ residues were generated and tested in protoplasts. Translation efficiency of 5′38+3U was similar to that of PEMV2-5′89+3U, and additional sequences did not further improve translation efficiency. Therefore, 5′38+3U was selected as the gRNA parental construct for further study.

For the sgRNA2 reporter constructs, the 5′ 27, 48, and 69 nucleotides were tested along with the full-length 3′ UTR (Fig. 7C). Luciferase activity produced by 5′27+3U in protoplasts was approximately 40% of PEMV2 5′89+3U (Fig. 7D). Additional downstream residues improved translational efficiency, with 5′69+3U activity similar to that of PEMV2 5′89+3U. Thus, 5′69+3U was selected as the parental sgRNA construct for further study.

Translation of OPMV gRNA and sgRNA reporter constructs depends on a BTE 3′ CITE and a long-distance RNA-RNA interaction with sequences proximal to the 5′ ends. SHAPE RNA structure probing combined with phylogenetic analyses were used to predict the structure of the 3′ terminal 652 nt in the 3′ UTR of OPMV (Fig. 8A). Key features of the structure include (i) a centrally located BTE 3′ CITE in a similar location as that of the PEMV2 PTE and kl-TSS 3′ CITEs and the TBTV BTE, (ii) a 3′ proximal TSS in the same location as the TSS in PEMV2 and several other umbraviruses composed of three hairpins (H4a, H4b, and H5) and two pseudoknots (ψ_2 and ψ_3) (Fig. 9), (iii) a pseudoknot (ψ_1) connecting one of the TSS hairpins (H5) with 3′ terminal sequences, which is also conserved in carmoviruses (51) and tombusviruses (52), and (4) a 3′ terminal hairpin, denoted the PR, which is structurally conserved in most umbraviruses and carmoviruses (11, 53). The TSS 3′ CITE, first characterized for carmovirus TCV (30, 32), is an unusual 3′ CITE in that it is missing an associated long-distance interacting sequence (54). The PEMV2 TSS was determined to function as a 3′ CITE for translation of sgRNA reporter constructs (11), and analogous hairpins and pseudoknots are found in OPMV, *Carrot mottle virus* (CMoV), PatMMoV, and TBTV (Fig. 9). As with PEMV2 and other umbraviruses, sequence in the apical loop of the PR is complementary to sequence in an asymmetric loop of a structure (the recoding stimulatory element [RSE]) just downstream from the −1 PRF site in all umbraviruses (2). This long-distance interaction is critical for efficient −1 PRF.

The OPMV BTE contains three stem-loops (SL1, SL2, and SL3), which radiate from a central hub atop a basal stem. The BTE has the canonical 17-nt BTE consensus sequence (GGGUCCUGGUAAACAGG) encompassing SL1, the hallmark of these 3′ CITEs (15, 55) (Fig. 8A). SL3 contains the sequence CUGCCAA in its apical loop, and the underlined sequence is frequently found in 5′ or 3′ sequences engaged in long-distance RNA-RNA interactions in carmoviruses and umbraviruses (15). The apical loop of the OPMV gRNA 5′-proximal hairpin (gH1) contains the sequence AUGGCA (Fig. 8A); therefore, these sequences likely represent the pairing partners that bring the BTE in proximity with the gRNA 5′ end. To support this prediction, the gRNA parental

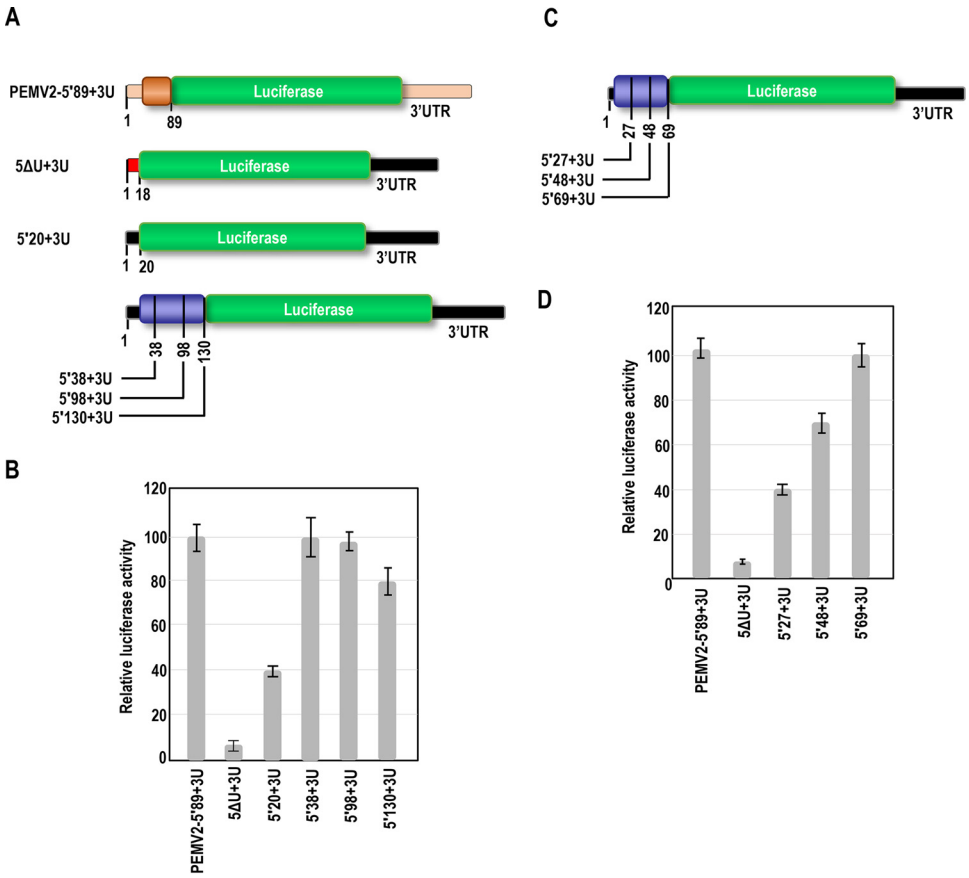


FIG 7 OPMV gRNA and sgRNA 5′ sequences necessary for efficient translation *in vivo*. (A) Constructs used to examine gRNA 5′ sequences required for efficient translation. PEMV2 control construct containing the PEMV2 5′ 89 nt and 3′ UTR (5′89+3U) was previously described (13). All OPMV constructs contain a full-length OPMV 3′ UTR. Blue box denotes OPMV coding region sequences. 5ΔU+3U contains 18 random nucleotides at the 5′ end. (B). Relative luciferase activity of constructs assayed in *Arabidopsis thaliana* protoplasts. Error bars denote standard deviation for three replicate assays. (C). Constructs used to assay OPMV sgRNA2 5′ sequences required for efficient translation. All constructs contain a full-length OPMV 3′ UTR. Blue box denotes coding region sequences. (D) Relative luciferase activity of constructs assayed in protoplasts. Error bars denote standard deviation for three replicate assays.

luciferase reporter construct 5′38+3U was mutated at a single site in the apical loops of BTE SL3 (mSL3) and gH1 (mgH1) such that combining the two mutations would be compensatory (Fig. 8A). Translation assays in protoplasts showed that both constructs with individual mutations produced 63% less luciferase than the parental construct (Fig. 8B). Combining the two mutations, which should reestablish the long-distance interaction, enhanced reporter activity to 98% of 5′38+3U levels. These results strongly suggest that the long-distance interaction between the OPMV BTE and the 5′ end of the gRNA is with the gRNA 5′ proximal hairpin.sgRNA2 also has a putative 5′ proximal hairpin with the apical loop sequence AUGGCA (Fig. 8A). To determine if the BTE similarly connects with this sequence in sgRNA2 (and possible sgRNA1), a mutation that would be complementary with BTE mutation mSL3 was generated in sgRNA parental construct 5′69+3U (msgH1) (Fig. 8A). BTE mutation mSL3 reduced translation of the sgRNA construct by 66%, and msgH1 reduced luciferase activity by 72% compared with the parental 5′69+3U. Combining the two mutations increased activity to 118% of parental levels (Fig. 8B). These results identify the components of the long-distance RNA-RNA interaction between the BTE and the 5′ ends of both sgRNA and sgRNA2, which are required for efficient translation of the two templates.

The BTE, but not the TSS, contributes to translation of the gRNA and sgRNA reporter constructs *in vivo*. In PEMV2, all three 3′ CITEs were required for translation of sgRNA reporter constructs *in vivo*, whereas only the kl-TSS and PTE were required

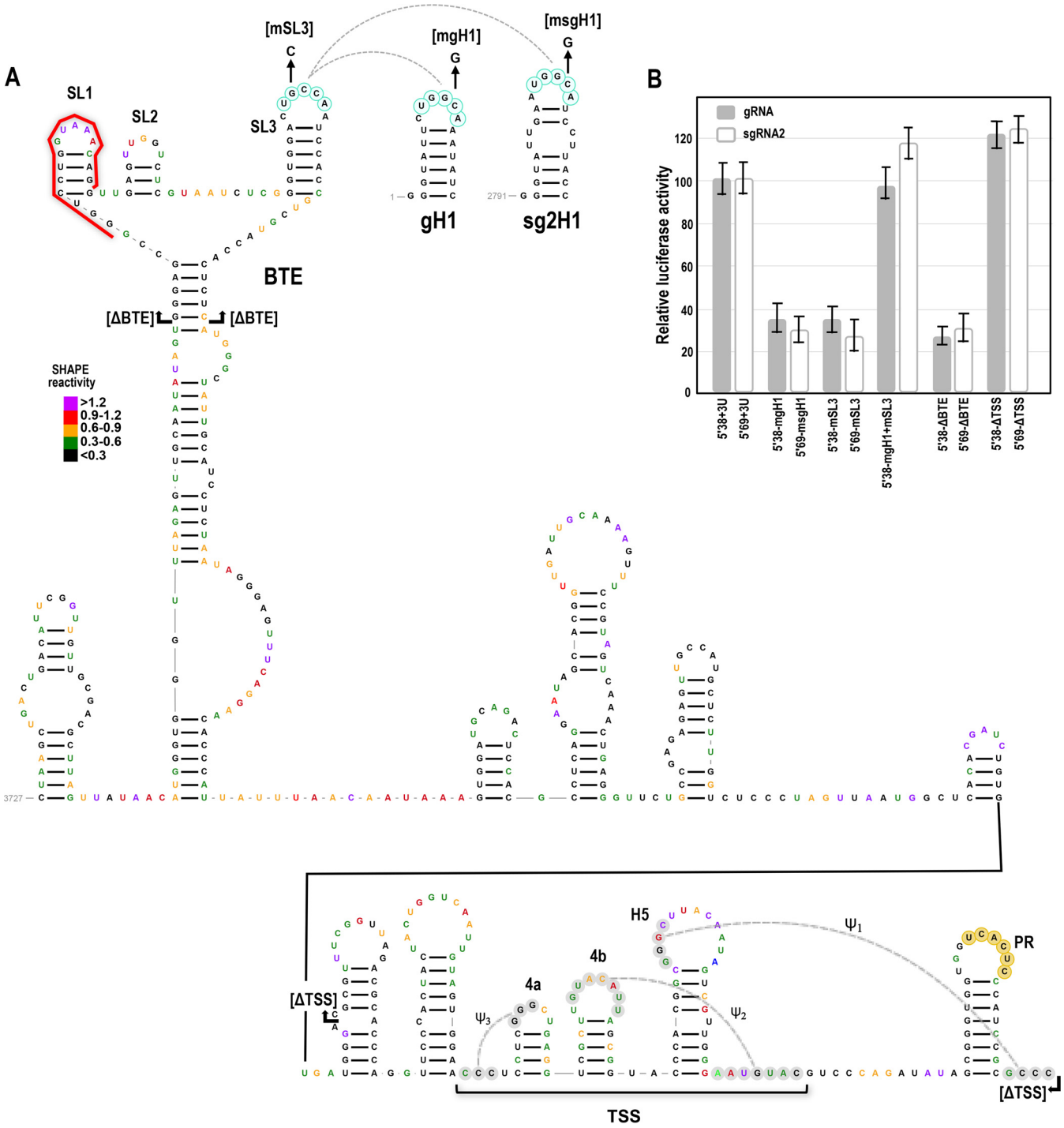


FIG 8 OPMV has a BTE 3′ CITE that connects with the gRNA and sgRNA2 5′ ends through a long-distance RNA-RNA interaction. (A) Proposed secondary structure of the 3′ UTR of OPMV. Residues are colored according to SHAPE NMIA reactivity, with purple being the most reactive and black being the least reactive. BTE and TSS 3′ CITEs are labeled. A red line denotes the BTE 17-nt signature sequence. Potential long-distance RNA-RNA interactions between the BTE and 5′ proximal hairpins in the gRNA (gH1) and sgRNA2 (sgH1) are denoted by dashed lines, with the signature motifs commonly found in carmovirus and umbravirus long-distance interactions shaded light blue. Alterations generated to assay for the interactions are shown with construct names in parentheses. Gray shaded sequences connected by dashed lines denote local tertiary interactions in the TSS and between the TSS and 3′ terminal sequences that are conserved in a subset of carmoviruses and umbraviruses (see Fig. 9). Hairpins and pseudoknots are labeled as previously described (11, 30). Sequence in the apical loop of the 3′ terminal hairpin that engages in the conserved interaction with a hairpin just downstream from the −1 PRF site is shaded tan. End points for ΔBTE and ΔTSS deletions in the parental gRNA luciferase construct 5′38+3U (generating 5′38+ΔBTE and 5′38+ΔTSS) and parental sgRNA construct 5′69+3U (generating 5′69+ΔBTE and 5′69+ΔTSS) are indicated. (B) Relative luciferase activity of constructs assayed in protoplasts. Error bars denote standard deviation for three independent assays.

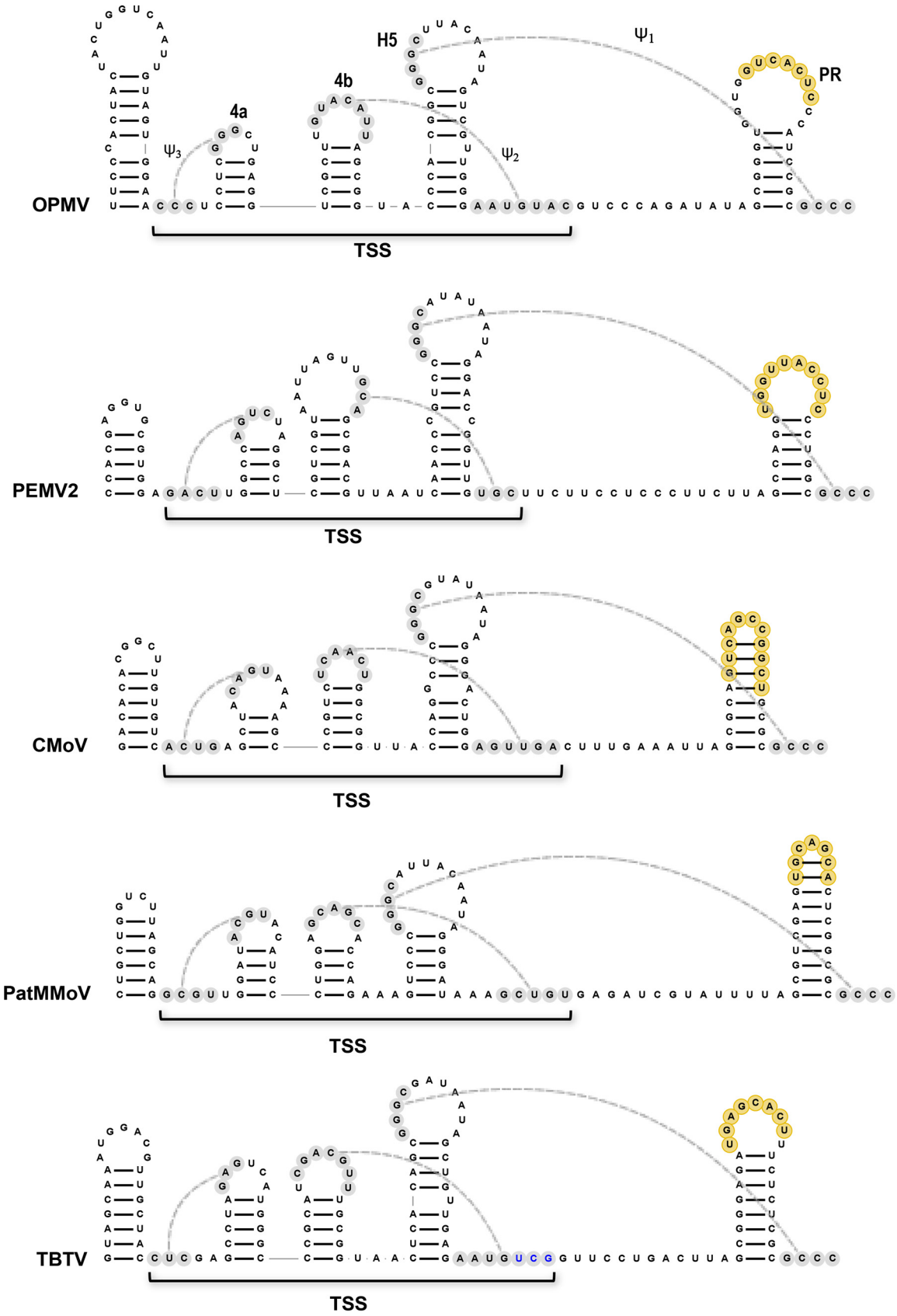


FIG 9 Conserved structures near the 3′ end of OPMV and some umbraviruses. Structure presented was confirmed by mutagenesis for PEMV2 (11). See legend to Fig. 8.

for translation of the gRNA reporter constructs (9, 33). To determine if both BTE and TSS are necessary for translation of the comparable OPMV gRNA and sgRNA reporter constructs, the upper portion of the BTE was deleted in gRNA construct 5′38+3U (5′38+ΔBTE) and sgRNA2 construct 5′69+3U (5′69+ΔBTE). 5′38+ΔBTE produced 28% of the luciferase activity of the parental construct (Fig. 8B), which was slightly below the luciferase activity in the absence of the long-distance interaction. Translation was also reduced in sgRNA construct 5′69+ΔBTE to a similar 32%. When the TSS was deleted in the gRNA construct (5′38+ΔTSS), luciferase activity was 21% higher than that of the parental construct. Similarly, deletion of the TSS in the sgRNA construct (5′69+ΔTSS) increased translation by a similar 24% over that of the parental construct. Results using these reporter constructs indicate that only the BTE is contributing to efficient translation of OPMV.

Comparative analyses of BTEs reveal additional structural and sequence similarities. In addition to OPMV and TBTV, GRV, ETBTV, PatMMoV, and IxYMaV2 all contain an element with the BTE 17-nt conserved sequence (Fig. 10). To identify additional conserved sequences and structures in BTE, a comparison was made between the umbraviral BTE and 7 additional BTE from viruses in the *Tombusviridae* and *Luteoviridae* (Fig. 10). As previously reported (15), there are three subclasses of BTE, which we have labeled subclasses A, B, and C. Subclass A are BTEs with three stem-loops (SLs), subclass B BTEs have two SLs, and subclass C BTEs have 5 SLs.

All BTEs contain the signature 17-nt sequence that forms SL1 (56). However, when represented with similar triangular 3-way junctions, a number of additional conserved BTE features, within and among the subclasses, became evident. For subclass A, these include: (i) the location of SL2, which is nearly always found (6 of 7 times) 2 nt downstream of SL1; (ii) the sequence G/AUAANNUCGGG is conserved just downstream of SL2, with the underlined guanylates forming the 5′ side base of SL3; and (iii) the sequence CCG U/C/A CGCAUCAC is conserved just downstream of SL3, with the underlined cytidylates forming the 3′ side base of SL3. In addition, just upstream of the 17-nt conserved sequence are 2 to 4 residues connected to a 6- or 7-nt base-paired stem. Little sequence or structural similarity is found below this stem region or within SL2 or SL3, with the exception of the base of SL3, as noted above. For subclass B, (i) UUUAACGGG is the conserved connector sequence between SL2 and SL3, with the underlined guanylates paired in the lower stem of SL2; (ii) just downstream of SL2 is the same conserved sequence (CCG U/A CGCAUCAC) as that found in subgroup A, with the underlined cytidylates again forming the lower stem of adjacent SL. The stem supporting the conserved upper portion of the BTE is more variable than that in subgroup A, consisting of 4 to 9 bp. Subgroup C, which is only found in the dianthovirus genus, contains the sequence "UCGGG" at the 5′ base of SL3 and "CCGUCG," with the pairing partners underlined, which are a subset of the conserved sequences in subgroup A.

The BYDV BTE binds with high affinity to eIF4G (18), the initiation factor that, along with eIF4E, is a component of the eIF4F cap-binding complex. Footprinting revealed residues protected by eIF4G binding, and these residues in the BYDV BTE and TNV-D BTE (subclasses A and B, respectively) are all within conserved sequences at the base of the hairpins (Fig. 10, boxes) (18). Since adoption of a compact native structure for the BYDV BTE required the presence of physiological concentrations of Mg^{2+} (18), it is likely that conserved residues connected to the upper supporting stem and drawn as unpaired in the BTE structures in Fig. 10, participate in noncanonical base pairings to stabilize the active structure and promote eIF4G binding.

Conclusions. We have developed OPMV as an additional umbravirus system for studying translation, revealing new information about sgRNAs generated by $xrRNA_D$ sites and BTE translation enhancers. A growing number of plant viruses, including dianthoviruses, benyviruses, cucumoviruses, sobemoviruses, machlomoviruses, tombusviruses, poleroviruses, and umbraviruses contain $xrRNA_D$-like structures or other elements that terminate Xrn4 5′-to-3′ exoribonuclease activity *in vivo* and/or *in vitro*, leading to the generation of sgRNAs (37–39, 57). With the exception of umbraviruses and poleroviruses, all such structures are located in 3′ UTRs generating noncoding, functional sgRNAs similar to those found in animal flaviviruses (46, 58). Our results confirm that the OPMV $xrRNA_D$ structure

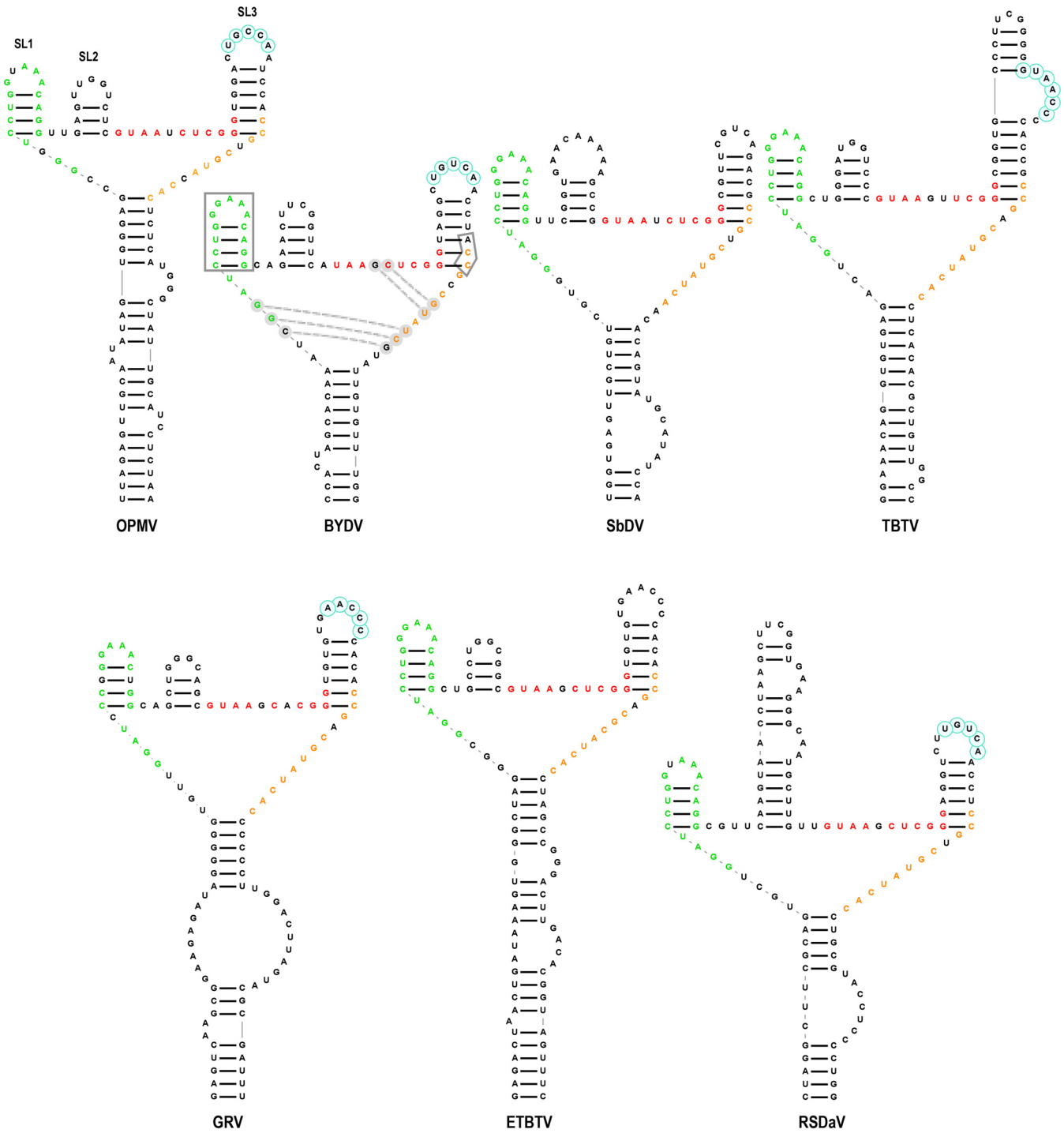
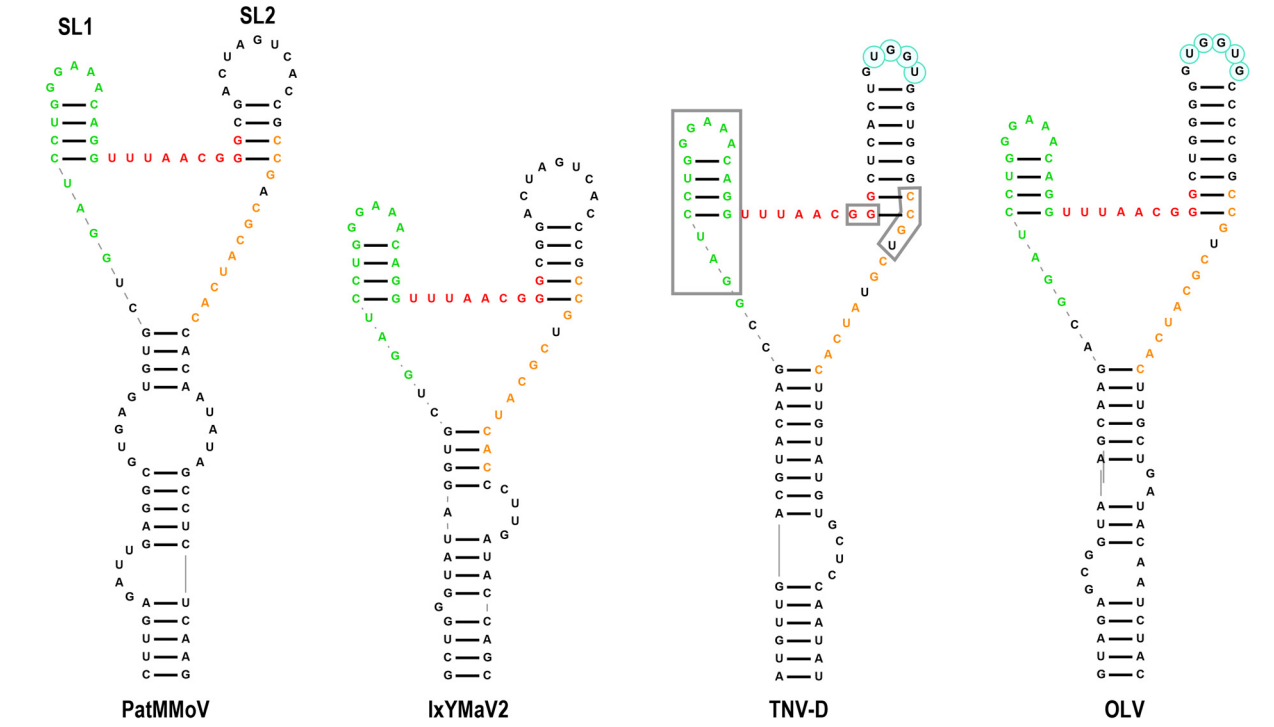


FIG 10 Comparative analysis of BTE structures. Proposed subclass A (three hairpins) includes BYDV, ETBTV, GRV, OPMV, *Rose spring dwarf-associated virus* (RSDaV), *Soybean dwarf virus* (SbDV), and TBTV. Subclass B (two hairpins) includes *Beet black scorch virus* (BBSV), IxYMaV2, OLV, PatMMoV, TNA-A, and TNVD. Subclass C (five hairpins) includes *Red clover necrotic mosaic virus* (RCNMV). All BTEs contain the 17-nt signature sequence (green), and the orange sequence is conserved in subclasses A and B. The red sequence connecting SL2 and SL3 in subclass A and SL2 and SL3 in subclass B is conserved within the subclasses. Bases that differ from consensus sequences within these motifs are in black. Additional base pairings for BYDV BTE are indicated, and these pairings are conserved for multiple BYDV isolates (18, 68). Boxed residues are protected by eIF4G binding from chemical modification (18). Apical loop residues known or predicted to be involved in the long-distance interactions with the 5′ end are shaded light blue (10, 13, 56, 69, 70).

Subclass B:



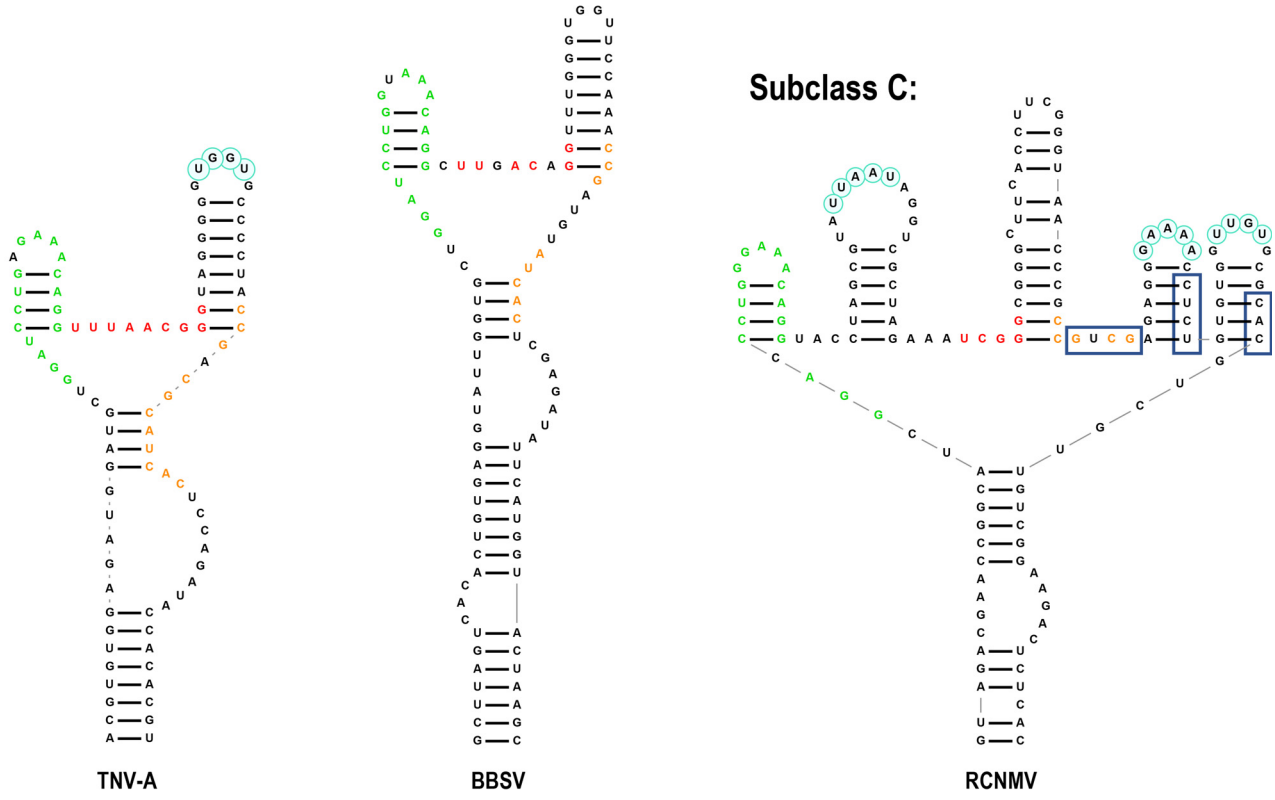


FIG 10 (Continued)

that was previously predicted (39) leads to the generation of an sgRNA (sgRNA1) *in vivo*. Curiously, only four of the 11 known or putative umbraviruses (OPMV, TBTV, CMoMV, and IxYMaV2) have an xrRNA$_D$ site at this IR location (Fig. 4A), and only two (OPMV and TBTV) have an additional xrRNA$_D$ site near the 5′ end of their 3′ UTR (Fig. 4B). Six of the remaining seven umbraviruses, including GRV, which is reported to also have two similarly sized sgRNAs (1, 7), contain various related sequence stretches and duplications near the same IR location, leading to the presence of between one and four versions of a similar hairpin with a conserved loop sequence (Fig. 4). It is reasonable to speculate that an unrelated structure exists in these umbraviruses that fulfills the same function as the xrRNA$_D$ site by terminating the degradation of the gRNA and producing an additional sgRNA.

Based on the location of an AUG in good context 87 nt downstream from the 5′ end, the OPMV Xrn4-generated sgRNA encodes a 30-kDa, N-terminal extension of the cell-to-cell movement protein p27 in WGE (Fig. 3). OPMV accumulation was significantly reduced in protoplasts when mutations disrupted either critical xrRNA$_D$-conserved sequences or translation of p30, suggesting that p30 may contribute to OPMV replication or stability in single cells. However, all other umbraviruses have in-frame termination codons just upstream of their ORF3 initiation codons, and thus if p30 is required for efficient accumulation of OPMV, this requirement is not shared by other umbraviruses. It is possible that sgRNA2, whose upstream hairpin and CCS is reminiscent of sgRNAs produced by TCV RdRp on truncated minus-strand templates (50, 59), is derived from sgRNA1 and not from the gRNA. However, mutations designed to disrupt a functional xrRNA$_D$ site did not eliminate detection of sgRNA2 (Fig. 3D). Further investigation will be needed to determine the role of sgRNA1 in the umbravirus infection cycle.

OPMV differs from PEMV2 by having a BTE 3′ CITE instead of a PTE/kl-TSS, but is similar by having a similar putative TSS near its 3′ end. Our results suggest that OPMV and PEMV2 differ in how they utilize the TSS for translation. Comparable sgRNA reporter constructs for PEMV2 required the TSS for efficient translation using 5′ sequences (9) whereas OPMV did not. One possibility is that insufficient sgRNA 5′ sequences were present in the sgRNA construct to allow for TSS use by the sgRNA reporter constructs. However, translation was still robust and comparable to translation of PEMV2 constructs in the absence of additional sequences (Fig. 7).

Finally, we have used this opportunity to examine all known BTE structures, which previously were thought to only share the conserved 17-nt signature sequence (15). By normalizing the structural depiction of BTEs, we have discovered an additional 12-nt conserved sequence in subclasses A and B that contain residues previously found to be protected by eIF4G (18), and additional conserved sequences within a particular subclass. The identification of these conserved sequences should assist in determining common structural and functional features within this important class of 3′ CITEs.

MATERIALS AND METHODS

Phylogenetic analysis of umbraviruses. For phylogenetic analysis of OPMV and other umbraviruses, full-length genome sequences of all available umbraviruses sequences in GenBank were aligned by MUSCLE (60). Using RAxML version 8.1 (61) and a GTR (general time-reversible) model, maximum-likelihood estimations were carried out with 1,000 bootstrap samples. A consensus phylogenetic tree was produced and edited in FigTree version 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and Inkscape version 1.0beta (https://inkscape.org).

5′ Rapid amplification of cDNA ends). 5′ RACE of the OPMV gRNA was carried out as previously described (62), with modifications. Briefly, total RNA from OPMV-infected *N. benthamiana* was used as the template in a reverse transcription reaction consisting of 1 μg of denatured RNA template, 1× first-strand buffer, 1 mM deoxynucleoside triphosphates (dNTPs), 0.5 μM OPMV-specific primer 1 (5′-CGATTTGGTCGTAGCTCGGT-3′), 20 U placental RNase inhibitor, and 100 U of Superscript III reverse transcriptase (Invitrogen) in a total volume of 20 μl. The reaction mixture was incubated at 42°C for 60 min. Following electrophoresis on a 1% agarose gel, cDNA migrating at the correct position was extracted. Poly(A) tails were added to the 3′ end of the cDNA in a reaction mixture containing 1× terminal transferase buffer, 0.2 mM dATP, and 10 U terminal transferase (New England Biolabs) in a total volume of 20 μl, which was incubated at 37°C for 15 min. PCR was performed in 1× PCR buffer (B9027; New England Biolabs), 2 mM dNTPs, 0.32 μM $(dT)_{17}$-adaptor primer (5′-AATGGATCCTCGAGCTCAAGCTTTTTTTTTTTTTTTT-3′), 0.64 μM adaptor primer (5′-AATGGATCCTCGAGCTCAAGC-3′), 0.64 μM OPMV-specific primer 2 (5′-TTGAATTCGGGTGGATTGCTCCTCAGAC-3′), and 2 U of Q5 DNA polymerase (New England Biolabs) in a total reaction volume of 50 μl, followed by cloning and sequencing of the amplified DNA.

Mapping sgRNA 5′ ends. The 5′ ends of OPMV sgRNAs were mapped by primer extension, as previously described (63). Briefly, 1 μg of total RNA isolated from OPMV-infected *N. benthamiana* at 20 h posttransfection was annealed to 0.5 pmol of 5′-end-[γ-^{32}P] ATP-labeled primer complementary to OPMV positions 2849 to 2865. *In vitro* synthesized transcripts of OPMV gRNA (3 pmol each dideoxynucleotide) were used as the template to produce sequencing ladders. SuperScript III reverse transcriptase (25 U) was added, and the reaction mixture was incubated at 52°C for 60 min. cDNA samples were resolved on 8% denaturing acrylamide gels. The gel was dried and exposed to a Fuji phosphorimager screen for 3 h, and the screen was scanned using an Amersham Typhoon fluorescent image analyzer.

Plasmid construction. Full-length OPMV gRNA and sgRNAs were PCR amplified using a T7 RNA polymerase promoter-containing forward primer incorporating an EcoRI restriction site and a reverse primer with a SmaI restriction site. The resulting PCR products were gel purified and cloned into pUC19, which was digested at compatible restriction sites to generate pUC19-gRNA, pUC19-sgRNA1, and pUC19-sgRNA2. OPMV cDNA was excised and cloned into *Agrobacterium tumefaciens* binary vector pCB301 downstream of the *Cauliflower mosaic virus* (CaMV) 35S promoter. Luciferase reporter constructs 5′20+3U, 5′38+3U, 5′98+3U, and 5′130+3U, containing the 5′ end 20, 38, 98, and 130 nt of OPMV gRNA, respectively, fused upstream of the firefly luciferase ORF, to which the OPMV full-length 3′ UTR was added downstream of the reporter ORF. Similarly, luciferase reporter constructs 5′27+3U, 5′48+3U, and 5′69+3U, containing the 5′ end 27, 48, and 69 nt of OPMV sgRNA2, respectively, fused upstream of the firefly luciferase ORF to which the OPMV full-length 3′ UTR was added downstream of the reporter ORF. Mutations were introduced using custom-designed oligonucleotide primers (Integrated DNA Technologies) and QuikChange one-step site-directed mutagenesis (64). All constructs were confirmed by sequencing (Eurofins Genomics).

Agroinfiltration and Northern blot analysis. *A. tumefaciens* strain GV3101 was transformed with OPMV binary vector constructs and cultured in the presence of antibiotics. The cells were harvested and resuspended in resuspension buffer (10 mM $MgCl_2$, 10 mM morpholineethanesulfonic acid [MES]-K [pH 5.7], and 100 μM acetosyringone). For all infiltrations of OPMV wild-type (WT) and mutant constructs, an optical density at 600 nm (OD_{600}) of 0.6 was mixed with the *Pothos latent virus* p14 silencing suppressor construct at an OD_{600} of 0.2 and infiltrated into fully expanded leaves of *N. benthamiana*. After 72 h, agroinfiltrated leaves were harvested, and TRIzol (Invitrogen) was used to isolate total RNA. For Northern blotting, RNA was subjected to electrophoresis through 2% agarose gels and transferred to a charged nylon membrane by capillary action in 4× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) buffer. After UV cross-linking, membranes were hybridized with the mixture of three [α^{-32}]dATP-labeled DNA oligonucleotides complementary to the sequence of OPMV at positions 3591 to 3630, 3634 to 3672, and 3676 to 3715.

SHAPE structure probing. SHAPE (selective 2′-hydroxyl acylation analyzed by primer extension) structure probing was performed using *in vitro* transcribed OPMV gRNA following phenol-chloroform extraction and ethanol precipitation. The resulting RNA was denatured at 65°C before being subjected to folding in RNA folding buffer [80 mM Tris-Cl (pH 8.0), 11 mM $Mg(CH_3COO)_2$, and 160 mM NH_4Cl] at 37°C for 20 min. Folded RNA was treated with either 15 mM *N*-methylisatoic anhydride (NMIA) dissolved in dimethyl sulfoxide (DMSO) or with the same volume of DMSO. For primer extension reactions, PET-labeled and 6-carboxyfluorescein (FAM)-labeled oligonucleotides complementary to OPMV gRNA were used in sequencing reactions (untreated gRNA) and sample reactions (NMIA- and DMSO-treated gRNA), respectively. The resulting cDNA products, along with ladders generated by Sanger sequencing, were sent to Macrogen, Inc., for fragment analysis. Trace files obtained from fragment analyses were processed and analyzed in QuShape (65). RNA secondary structures were initially predicted using mFold (66) and modified by phylogenetic comparisons.

***In vitro* RNA transcription and translation.** pUC19-gRNA constructs or luciferase reporter constructs were linearized with SmaI or SspI, respectively, and served as the templates for RNA transcription using T7 RNA polymerase. For *in vitro* translation, 10-μl translation mixtures contained 5 μl wheat germ extract (Promega), 0.5 pmol RNA template, 0.8 μl 1 mM amino acids mix (without methionine), 100 mM potassium acetate, and 0.5 μl [5 μCi] ^{35}S-methionine. The translation mixture was incubated at 25°C for 45 min and then resolved on a 10% SDS-PAGE gel. The gel was dried and exposed to Fuji phosphorimager screen for 3 h. The screen was subsequently scanned by an Amersham Typhoon fluorescent image analyzer. The intensity of radioactive bands was quantified using ImageQuant TL 8.1 (GE Lifesciences). All experiments were repeated at least three times.

Protoplast transfection and *in vivo* luciferase assays. *Arabidopsis thaliana* protoplasts were prepared from callus cultures and transfected using a polyethylene glycol-mediated transformation protocol as previously described (13). Briefly, 5 × 10^6 protoplasts were transfected with 20 μg of OPMV luciferase reporter transcripts and incubated under constant light for 20 h at 22°C. Cells were lysed at 18 h posttransfection, and luciferase activity was assayed with a dual-luciferase reporter assay system (Promega) using a Modulus microplate multimode reader (Turner BioSystems).

RNA structure drawing. All RNA structures were drawn using the RNA2Drawer online drawing tool at https://rna2drawer.app/ (67).

Data availability. The sequence of the infectious OPMV construct was deposited in GenBank under accession number MG182693.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (grant MCB-1818229 to A.E.S.).

We acknowledge Joe Tang, Plant Health and Environment Laboratory, Auckland, New Zealand, for providing the infected plant material. The University of Maryland made supercomputing resources (http://hpcc.umd.edu) available for conducting the research reported in this paper. We also thank K. Andrew White for very insightful comments on the manuscript.

REFERENCES

1. Taliansky ME, Robinson DJ. 2003. Molecular biology of umbraviruses: phantom warriors. J Gen Virol 84:1951–1960. https://doi.org/10.1099/vir.0.19219-0.
2. Gao F, Simon AE. 2016. Multiple *cis*-acting elements modulate programmed −1 ribosomal frameshifting in *Pea enation mosaic virus*. Nucleic Acids Res 44:878–895. https://doi.org/10.1093/nar/gkv1241.
3. Kim SH, MacFarlane S, Kalinina NO, Rakitina DV, Ryabov EV, Gillespie T, Haupt S, Brown JWS, Taliansky M. 2007. Interaction of a plant virus-encoded protein with the major nucleolar protein fibrillarin is required for systemic virus infection. Proc Natl Acad Sci U S A 104:11115–11120. https://doi.org/10.1073/pnas.0704632104.
4. Ryabov EV, Robinson DJ, Taliansky M. 2001. Umbravirus-encoded proteins both stabilize heterologous viral RNA and mediate its systemic movement in some plant species. Virology 288:391–400. https://doi.org/10.1006/viro.2001.1078.
5. May JP, Johnson PZ, Ilyas M, Gao F, Simon AE. 2020. The multifunctional long-distance movement protein of *Pea enation mosaic virus 2* protects viral and host transcripts from nonsense-mediated decay. mBio 11:e00204-20. https://doi.org/10.1128/mBio.00204-20.
6. May JP, Yuan X, Sawicki E, Simon AE. 2018. RNA virus evasion of nonsense-mediated decay. PLoS Path 14:e1007459. https://doi.org/10.1371/journal.ppat.1007459.
7. Taliansky ME, Robinson DJ, Murant AF. 1996. Complete nucleotide sequence and organization of the RNA genome of groundnut rosette umbravirus. J Gen Virol 77:2335–2345. https://doi.org/10.1099/0022-1317-77-9-2335.
8. Gao F, Alekhina OM, Vassilenko KS, Simon AE. 2018. Unusual dicistronic expression from closely spaced initiation codons in an umbravirus subgenomic RNA. Nucleic Acids Res 46:11726–11742. https://doi.org/10.1093/nar/gky871.
9. Gao F, Simon AE. 2017. Differential use of 3′ CITEs by the subgenomic RNA of Pea enation mosaic virus 2. Virology 510:194–204. https://doi.org/10.1016/j.virol.2017.07.021.
10. Wang DY, Yu CM, Liu SS, Wang GL, Shi KR, Li XD, Yuan XF. 2017. Structural alteration of a BYDV-like translation element (BTE) that attenuates p35 expression in three mild *Tobacco bushy top virus* isolates. Sci Rep 7:4213. https://doi.org/10.1038/s41598-017-04598-5.
11. Gao F, Kasprzak WK, Szarko C, Shapiro BA, Simon AE. 2014. The 3′ untranslated region of *Pea enation mosaic virus* contains two T-shaped, ribosome-binding, cap-independent translation enhancers. J Virol 88:11696–11712. https://doi.org/10.1128/JVI.01433-14.
12. Gao F, Gulay SP, Kasprzak W, Dinman JD, Shapiro BA, Simon AE. 2013. The kissing-loop T-shaped structure translational enhancer of *Pea enation mosaic virus* can bind simultaneously to ribosomes and a 5′ proximal hairpin. J Virol 87:11987–12002. https://doi.org/10.1128/JVI.02005-13.
13. Gao F, Kasprzak W, Stupina VA, Shapiro BA, Simon AE. 2012. A ribosome-binding, 3′ translational enhancer has a T-shaped structure and engages in a long-distance RNA-RNA interaction. J Virol 86:9828–9842. https://doi.org/10.1128/JVI.00677-12.
14. Miras M, Miller WA, Truniger V, Aranda MA. 2017. Non-canonical translation in plant RNA viruses. Front Plant Sci 8:494. https://doi.org/10.3389/fpls.2017.00494.
15. Simon AE, Miller WA. 2013. 3′ Cap-independent translation enhancers of plant viruses. Annu Rev Microbiol 67:21–42. https://doi.org/10.1146/annurev-micro-092412-155609.
16. Truniger V, Miras M, Aranda MA. 2017. Structural and functional diversity of plant virus 3′-cap-independent translation enhancers (3′-CITEs). Front Plant Sci 8:2047. https://doi.org/10.3389/fpls.2017.02047.
17. Das Sharma S, Kraft JJ, Miller WA, Goss DJ. 2015. Recruitment of the 40S ribosomal subunit to the 3′-untranslated region (UTR) of a viral mRNA, via the eIF4 complex, facilitates cap-independent translation. J Biol Chem 290:11268–11281. https://doi.org/10.1074/jbc.M115.645002.
18. Kraft JJ, Treder K, Peterson MS, Miller WA. 2013. Cation-dependent folding of 3′ cap-independent translation elements facilitates interaction of a 17-nucleotide conserved sequence with eIF4G. Nucleic Acids Res 41:3398–3413. https://doi.org/10.1093/nar/gkt026.
19. Treder K, Kneller ELP, Allen EM, Wang ZH, Browning KS, Miller WA. 2008. The 3' cap-independent translation element of *Barley yellow dwarf virus* binds eIF4F via the eIF4G subunit to initiate translation. RNA 14:134–147. https://doi.org/10.1261/rna.777308.
20. Chattopadhyay M, Kuhlmann MM, Kumar K, Simon AE. 2014. Position of the kissing-loop interaction associated with PTE-type 3'CITEs can affect enhancement of cap-independent translation. Virology 458–459:43–52. https://doi.org/10.1016/j.virol.2014.03.027.
21. Chattopadhyay M, Shi K, Yuan X, Simon AE. 2011. Long-distance kissing loop interactions between a 3′ proximal Y-shaped structure and apical loops of 5′ hairpins enhance translation of *Saguaro cactus virus*. Virology 417:113–125. https://doi.org/10.1016/j.virol.2011.05.007.
22. Wang Z, Parisien M, Scheets K, Miller WA. 2011. The cap-binding translation initiation factor, eIF4E, binds a pseudoknot in a viral cap-independent translation element. Structure 19:868–880. https://doi.org/10.1016/j.str.2011.03.013.
23. Blanco-Perez M, Perez-Canamas M, Ruiz L, Hernandez C. 2016. Efficient translation of *Pelargonium line pattern virus* RNAs relies on a TED-like 3′-translational enhancer that communicates with the corresponding 5'-region through a long-distance RNA-RNA interaction. PLoS One 11:e0152593. https://doi.org/10.1371/journal.pone.0152593.
24. Gazo BM, Murphy P, Gatchel JR, Browning KS. 2004. A novel interaction of cap-binding protein complexes eukaryotic initiation factor (eIF) 4F and eIF(iso)4F with a region in the 3′-untranslated region of satellite tobacco necrosis virus. J Biol Chem 279:13584–13592. https://doi.org/10.1074/jbc.M311361200.
25. Nicholson BL, Zaslaver O, Mayberry LK, Browning KS, White KA. 2013. Tombusvirus Y-shaped translational enhancer forms a complex with eIF4F and can be functionally replaced by heterologous translational enhancers. J Virol 87:1872–1883. https://doi.org/10.1128/JVI.02711-12.
26. Nicholson BL, White KA. 2008. Context-influenced cap-independent translation of *Tombusvirus* mRNAs *in vitro*. Virology 380:203–212. https://doi.org/10.1016/j.virol.2008.08.003.
27. Fabian MR, White KA. 2004. 5′-3′ RNA-RNA interaction facilitates cap- and poly(A) tail-independent translation of tomato bushy stunt virus mRNA: a potential common mechanism for *Tombusviridae*. J Biol Chem 279:28862–-28872. https://doi.org/10.1074/jbc.M401272200.
28. Nicholson BL, Wu B, Chevtchenko I, White KA. 2010. Tombusvirus recruitment of host translational machinery via the 3′ UTR. RNA 16:1402–1419. https://doi.org/10.1261/rna.2135210.
29. Liu Q, Goss DJ. 2018. The 3 mRNA I-shaped structure of maize necrotic streak virus binds to eukaryotic translation factors for eIF4F-mediated translation initiation. J Biol Chem 293:9486–9495. https://doi.org/10.1074/jbc.RA118.003377.
30. McCormack JC, Yuan X, Yingling YG, Kasprzak W, Zamora RE, Shapiro BA, Simon AE. 2008. Structural domains within the 3′ untranslated region of *Turnip crinkle virus*. J Virol 82:8706–8720. https://doi.org/10.1128/JVI.00416-08.
31. Zuo XB, Wang JB, Yu P, Eyler D, Xu H, Starich MR, Tiede DM, Simon AE, Kasprzak W, Schwieters CD, Shapiro BA, Wang YX. 2010. Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3′ UTR of turnip crinkle virus. Proc Natl Acad Sci U S A 107:1385–1390. https://doi.org/10.1073/pnas.0908140107.
32. Stupina VA, Meskauskas A, McCormack JC, Yingling YG, Shapiro BA, Dinman JD, Simon AE. 2008. The 3′ proximal translational enhancer of *Turnip crinkle virus* binds to 60S ribosomal subunits. RNA 14:2379–2393. https://doi.org/10.1261/rna.1227808.
33. Du Z, Alekhina OM, Vassilenko KS, Simon AE. 2017. Concerted action of two 3′ cap-independent translation enhancers increases the competitive strength of translated viral genomes. Nucleic Acids Res 45:9558–9572. https://doi.org/10.1093/nar/gkx643.

34. Nicholson BL, Lee PKK, White KA. 2012. Internal RNA replication elements are prevalent in *Tombusviridae*. Front Virol 3:279. https://doi.org/10.3389/fmicb.2012.00279.
35. Steckelberg AL, Vicens Q, Costantino DA, Nix JC, Kieft JS. 2020. The crystal structure of a *Polerovirus* exoribonuclease-resistant RNA shows how diverse sequences are integrated into a conserved fold. RNA 26:1767–1776. https://doi.org/10.1261/rna.076224.120.
36. Jones RA, Steckelberg AL, Vicens Q, Szucs MJ, Akiyama BM, Kieft JS. 2021. Different tertiary interactions create the same important 3D features in a distinct flavivirus xrRNA. RNA 27:54–65. https://doi.org/10.1261/rna.077065.120.
37. Flobinus A, Chevigny N, Charley PA, Seissler T, Klein E, Bleykasten-Grosshans C, Ratti C, Bouzoubaa S, Wilusz J, Gilmer D. 2018. Beet necrotic yellow vein virus noncoding RNA production depends on a 5′→3′ Xrn exoribonuclease activity. Viruses-Basel 10:137. https://doi.org/10.3390/v10030137.
38. Gunawardene CD, Newburn LR, White KA. 2019. A 212-nt long RNA structure in the *Tobacco necrosis virus-D* RNA genome is resistant to Xrn degradation. Nucleic Acids Res 47:9329–9342. https://doi.org/10.1093/nar/gkz668.
39. Steckelberg AL, Vicens Q, Kieft JS. 2018. Exoribonuclease-resistant RNAs exist within both coding and noncoding subgenomic RNAs. mBio 9:e02461-18. https://doi.org/10.1128/mBio.02461-18.
40. Steckelberg AL, Akiyama BM, Costantino DA, Sit TL, Nix JC, Kieft JS. 2018. A folded viral noncoding RNA blocks host cell exoribonucleases through a conformationally dynamic RNA structure. Proc Natl Acad Sci U S A 115:6404–6409. https://doi.org/10.1073/pnas.1802429115.
41. Pijlman GP, Funk A, Kondratieva N, Leung J, Torres S, van der Aa L, Liu WJ, Palmenberg AC, Shi P-Y, Hall RA, Khromykh AA. 2008. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. Cell Host Microbe 4:579–591. https://doi.org/10.1016/j.chom.2008.10.007.
42. Iwakawa HO, Mizumoto H, Nagano H, Imoto Y, Takigawa K, Sarawaneeyaruk S, Kaido M, Mise K, Okuno T. 2008. A viral noncoding RNA generated by *cis*-element-mediated protection against 5′′→3′ RNA decay represses both cap-independent and cap-dependent translation. J Virol 82:10162–10174. https://doi.org/10.1128/JVI.01027-08.
43. Guan H, Carpenter C, Simon A. 2000. Analysis of *cis*-acting sequences involved in plus-strand synthesis of a turnip crinkle virus-associated satellite RNA identifies a new carmovirus replication element. Virology 268:345–354. https://doi.org/10.1006/viro.1999.0153.
44. Nagarajan VK, Jones CI, Newbury SF, Green PJ. 2013. XRN 5′→3′ exoribonucleases: structure, mechanisms and functions. Biochim Biophys Acta 1829:590–603. https://doi.org/10.1016/j.bbagrm.2013.03.005.
45. Lee C-C, Lin T-L, Lin J-W, Han Y-T, Huang Y-T, Hsu Y-H, Meng M. 2015. Promotion of *Bamboo mosaic virus* accumulation in *Nicotiana benthamiana* by 5′→3′ exonuclease NbXRN4. Front Microbiol 6:1508. https://doi.org/10.3389/fmicb.2015.01508.
46. Clarke BD, Roby JA, Slonchak A, Khromykh AA. 2015. Functional non-coding RNAs derived from the flavivirus 3′ untranslated region. Virus Res 206:53–61. https://doi.org/10.1016/j.virusres.2015.01.026.
47. Peltier C, Klein E, Hleibieh K, D'Alonzo M, Hammann P, Bouzoubaa S, Ratti C, Gilmer D. 2012. Beet necrotic yellow vein virus subgenomic RNA3 is a cleavage product leading to stable non-coding RNA required for long-distance movement. J Gen Virol 93:1093–1102. https://doi.org/10.1099/vir.0.039685-0.
48. Zhang QY, Li XF, Niu XL, Li N, Wang HJ, Deng CL, Ye HQ, Huang XY, Chen Q, Xu YP, Dong HL, Li XD, Zhao H, Shi PY, Yuan ZM, Gong P, Fang XY, Qin CF, Zhang B. 2020. Short direct repeats in the 3′ untranslated region are involved in subgenomic flaviviral RNA production. J Virol 94:e01175-19. https://doi.org/10.1128/JVI.01175-19.
49. Jiwan SD, White KA. 2011. Subgenomic mRNA transcription in *Tombusviridae*. RNA Biol 8:287–294. https://doi.org/10.4161/rna.8.2.15195.
50. Wu BD, Oliveri S, Mandic J, White KA. 2010. Evidence for a premature termination mechanism of subgenomic mRNA transcription in a carmovirus. J Virol 84:7904–7907. https://doi.org/10.1128/JVI.00742-10.
51. Zhang JC, Zhang GH, McCormack JC, Simon AE. 2006. Evolution of virus-derived sequences for high-level replication of a subviral RNA. Virology 351:476–488. https://doi.org/10.1016/j.virol.2006.03.011.
52. Panaviene Z, Panavas T, Nagy PD. 2005. Role of an internal and two 3′-terminal RNA elements in assembly of tombusvirus replicase. J Virol 79:10608–10618. https://doi.org/10.1128/JVI.79.16.10608-10618.2005.
53. Simon AE. 2015. 3′UTRs of carmoviruses. Virus Res 206:27–36. https://doi.org/10.1016/j.virusres.2015.01.023.
54. Stupina VA, Yuan X, Meskauskas A, Dinman JD, Simon AE. 2011. Ribosome binding to a 5′ translational enhancer is altered in the presence of the 3′ untranslated region in cap-independent translation of turnip crinkle virus. J Virol 85:4638–4653. https://doi.org/10.1128/JVI.00005-11.
55. Koev G, Liu SJ, Beckett R, Miller WA. 2002. The 3′-terminal structure required for replication of *Barley yellow dwarf virus* RNA contains an embedded 3′ end. Virology 292:114–126. https://doi.org/10.1006/viro.2001.1268.
56. Wang Z, Kraft JJ, Hui AY, Miller WA. 2010. Structural plasticity of *Barley yellow dwarf virus*-like cap-independent translation elements in four genera of plant viral RNAs. Virology 402:177–186. https://doi.org/10.1016/j.virol.2010.03.025.
57. Dilweg IW, Gultyaev AP, Olsthoorn RC. 2019. Structural features of an Xrn1-resistant plant virus RNA. RNA Biol 16:838–845. https://doi.org/10.1080/15476286.2019.1592070.
58. Slonchak A, Khromykh AA. 2018. Subgenomic flaviviral RNAs: what do we know after the first decade of research. Antiviral Res 159:13–25. https://doi.org/10.1016/j.antiviral.2018.09.006.
59. Wang J, Carpenter CD, Simon AE. 1999. Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. Virology 253:327–336. https://doi.org/10.1006/viro.1998.9538.
60. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.
61. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
62. Anonymous. 2005. Rapid amplification of 5′ complementary DNA ends (5′ RACE). Nat Methods 2:629–630. https://doi.org/10.1038/nmeth0805-629.
63. Wang J, Simon A. 1997. Analysis of the two subgenomic RNA promoters for turnip crinkle virus *in vivo* and *in vitro*. Virology 232:174–186. https://doi.org/10.1006/viro.1997.8550.
64. Liu H, Naismith JH. 2008. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol 8:91. https://doi.org/10.1186/1472-6750-8-91.
65. Karabiber F. 2013. A peak alignment algorithm with novel improvements in application to electropherogram analysis. J Bioinform Comput Biol 11:1350011. https://doi.org/10.1142/S021972001350011X.
66. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415. https://doi.org/10.1093/nar/gkg595.
67. Johnson PZ, Kasprzak WK, Shapiro BA, Simon AE. 2019. RNA2Drawer: geometrically strict drawing of nucleic acid structures with graphical structure editing and highlighting of complementary subsequences. RNA Biol 16:1667–1671. https://doi.org/10.1080/15476286.2019.1659081.
68. Miller WA, Jackson J, Feng Y. 2015. *Cis*- and *trans*-regulation of luteovirus gene expression by the 3′ end of the viral genome. Virus Res 206:37–45. https://doi.org/10.1016/j.virusres.2015.03.009.
69. Guo L, Allen EM, Miller WA. 2001. Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. Mol Cell 7:1103–1109. https://doi.org/10.1016/s1097-2765(01)00252-0.
70. Shen RZ, Miller WA. 2004. The 3′ untranslated region of tobacco necrosis virus RNA contains a barley yellow dwarf virus-like cap-independent translation element. J Virol 78:4655–4664. https://doi.org/10.1128/jvi.78.9.4655-4664.2004.