



# Exoribonuclease-Resistant RNAs Exist within both Coding and Noncoding Subgenomic RNAs

Anna-Lena Steckelberg,<sup>a</sup> Quentin Vicens,<sup>a</sup> Jeffrey S. Kieft<sup>a</sup>

<sup>a</sup>Department of Biochemistry and Molecular Genetics and RNA BioScience Initiative, University of Colorado Denver School of Medicine, Aurora, Colorado, USA

**ABSTRACT** Many viruses produce protein-coding and noncoding subgenomic RNAs (sgRNAs) that are critical for infection. A recently discovered pathway for viral sgRNA production uses exoribonuclease-resistant RNAs (xrRNAs), discrete folded RNA elements that block the processive exoribonucleolytic degradation of RNA. xrRNAs are widespread in animal-infecting flaviviruses but had been found only in three members of the plant virus genus *Dianthovirus*. Also, xrRNAs had been found only in the 3' untranslated regions (3'UTRs) of viral RNAs, where they produce noncoding sgRNAs. The degree to which xrRNA elements exist in other viruses, the conservation of their ring-like fold, and the ability of xrRNAs to operate in diverse contexts were unknown. Using computational tools and biochemical assays, we discovered xrRNA elements pervading two large families of plant-infecting RNA viruses, demonstrating their importance and widespread utility. Comparison of the sequences and functional requirements suggests that all adopt the characteristic ring-like fold. Unexpectedly, many of these newly discovered xrRNAs are located in intergenic regions rather than 3'UTRs, and some are associated with the 5' ends of subgenomic RNAs that encode viral proteins. This suggests that xrRNAs are involved in the production of both coding and noncoding subgenomic RNAs and can operate as part of broader mechanisms to regulate RNA levels and protein expression. These discoveries expand the potential roles for xrRNAs and suggest that xrRNAs may represent a more general strategy for RNA maturation and maintenance than previously known.

**IMPORTANCE** During infection, viruses often produce subgenomic RNAs (sgRNAs) that either serve as the template for protein synthesis or act as “riboregulators” that interact with and influence the viral and cellular machinery. Recently, a mechanism for producing sgRNAs was found that depends on the presence of specifically structured RNA elements (xrRNAs). However, the degree to which this mechanism is used, where the elements are found, their structural diversity, and what types of sgRNAs are produced by this pathway were unclear. This article describes the discovery of these structured RNA elements in two large families of plant viruses and shows that they are used to produce both protein-coding sgRNAs and “riboregulatory” RNAs. These discoveries provide evidence that xrRNA-based RNA maturation pathways may be more widespread than previously anticipated and that they are involved in producing a variety of RNAs of diverse functions.

**KEYWORDS** *Infernal*, RNA structure, exoribonuclease resistance, plant viruses, viral RNA

Single-stranded positive-sense RNA viruses include pathogens that infect a wide range of animal and plant hosts, with significant human health, agricultural, and economic impact. During infection, these viruses must generate many copies of their full-length genomic RNA, but many also produce one or more infection-critical sub-

**Received** 6 November 2018 **Accepted** 12 November 2018 **Published** 18 December 2018

**Citation** Steckelberg A-L, 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>.

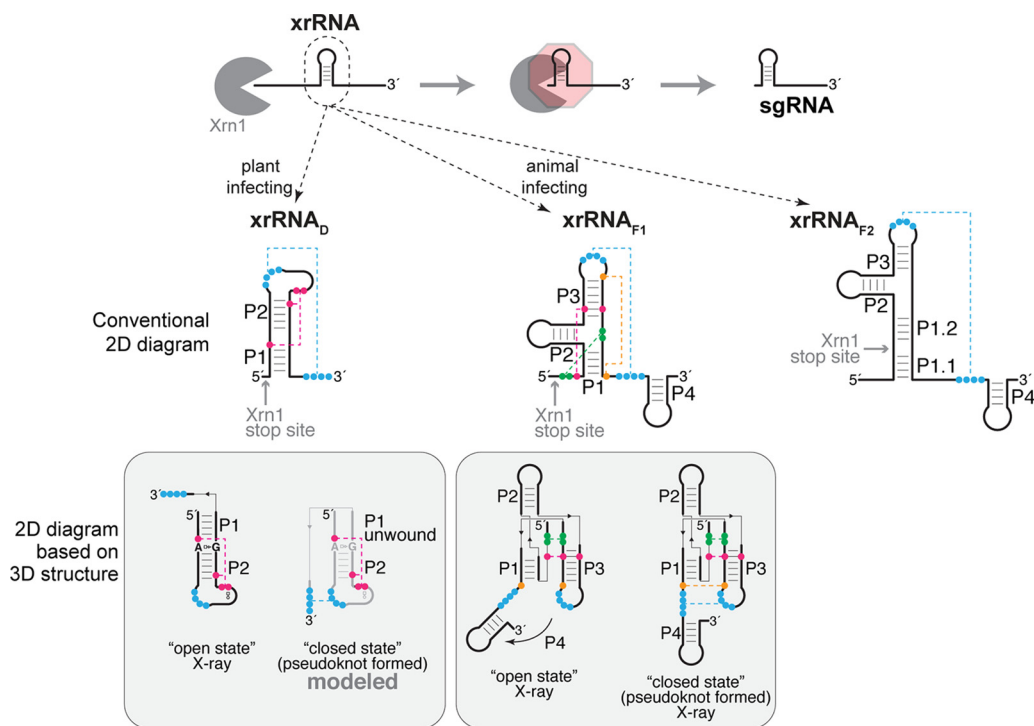
**Editor** Alexander Idnurm, University of Melbourne

**Copyright** © 2018 Steckelberg et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Quentin Vicens, [quentin.vicens@ucdenver.edu](mailto:quentin.vicens@ucdenver.edu), or Jeffrey S. Kieft, [jeffrey.kieft@ucdenver.edu](mailto:jeffrey.kieft@ucdenver.edu).

A.-L.S. and Q.V. contributed equally to this article.

This article is a direct contribution from a Fellow of the American Academy of Microbiology. Solicited external reviewers: Samuel Butcher, UW-Madison; W. Allen Miller, Iowa State University.



**FIG 1** An expanding repertoire of structured RNAs for blocking exoribonuclease degradation. (Top) xrRNAs adopt a three-dimensional structure that blocks the progression of 5'-to-3' exoribonucleases such as Xrn1 (gray). In the case of flaviviruses and dianthoviruses, xrRNAs are in the 3' UTR, resulting in accumulating noncoding sgRNAs. (Middle) Secondary structure diagrams of the two classes of xrRNAs from flaviviruses (xrRNA<sub>F1</sub> and xrRNA<sub>F2</sub>) (15, 22, 23) and of xrRNA<sub>D</sub> from dianthoviruses (26). Secondary structure features are labeled, and nucleotides involved in tertiary interactions are shown in colors connected by dashed lines (pseudoknot shown in blue). Experimentally determined Xrn1 stop sites are indicated. (Bottom) The boxes below each secondary structure contain diagrams reflecting the currently available three-dimensional structures (24–26). The A8-G33 pair is highlighted in the open state of the Sweet clover necrotic mosaic virus (SCNMV) xrRNA (far left).

genomic RNA species (sgRNA). sgRNAs can have different functions: they can encode and serve as the template for production of specific viral proteins or act as noncoding "riboregulators" that interact with and influence the cellular and viral machinery to benefit viral infection (1–6). Most viral sgRNAs are thought to be produced directly through transcription from internal subgenomic promoters or by premature termination during negative-strand synthesis (4). However, recent discoveries showed that some noncoding viral sgRNAs result from incomplete degradation of the genomic RNA in a pathway depending on discrete, compact RNA structures that block the progression of 5'-to-3' exoribonucleases (Fig. 1) (7–15). The full extent of this phenomenon and the diversity of RNA structures that can provide this ability are unknown, and to date all have been associated with noncoding RNA production.

Exoribonuclease-resistant RNA (xrRNA) elements were first identified in mosquito-borne flaviviruses (e.g., dengue virus, Zika virus, West Nile virus), where they protect the genome's 3' untranslated region (3' UTR) from degradation (8). The resultant decay intermediates accumulate and comprise biologically active viral noncoding sgRNAs (Fig. 1) (8, 9, 12, 16–21). xrRNAs are useful and important elements, as they are broadly found within the 3' UTRs of flaviviruses, including those that are tick-borne; are specific to insects; or have no known vector (15, 22, 23). Comparison and characterization of these diverse flaviviral xrRNA (xrRNA<sub>F</sub>) sequences revealed two classes; class I (xrRNA<sub>F1</sub>) is exemplified by mosquito-borne flaviviruses, whereas class II (xrRNA<sub>F2</sub>) is found in diverse flaviviruses (23). Extensive functional and high-resolution structural studies of xrRNA<sub>F1</sub> have shown that function is conferred by a specific three-dimensional (3D) fold containing an interwoven pseudoknot stabilized by conserved secondary and tertiary interactions; this creates an unusual ring-like conformation that protectively wraps

around the 5' end of the RNA structure (24, 25). Although aligned xrRNA<sub>F2</sub> sequences show conserved patterns, their 3D structures are unknown (23).

As with xrRNA<sub>F2</sub>, the structures of recently reported xrRNAs from most other viral clades remain unsolved (7, 13), with the exception of those found in a small genus of plant-infecting RNA viruses. Specifically, we recently characterized the structure and function of xrRNAs from the 3'UTRs of dianthoviruses, which are positive-sense RNA viruses in the *Tombusviridae* family; similarly to xrRNA<sub>F</sub>, dianthoviral xrRNAs (xrRNA<sub>D</sub>) function to produce a noncoding RNA derived from the viral 3'UTR (10, 26). xrRNA<sub>D</sub> also rely on a pseudoknot that forms a protective ring-like structure (26), but they have very different sequences and secondary structures from those of xrRNA<sub>F1</sub>, and the ring is formed by a different set of interactions (Fig. 1). The xrRNA<sub>D</sub> crystal structure is in an "open" conformation that likely represents a folding intermediate whose presence is necessary before the pseudoknot forms (26) (Fig. 1). Thus, we still do not know the full repertoire of secondary and tertiary interactions required to form and stabilize the exoribonuclease-resistant pseudoknot state of xrRNA<sub>D</sub>. In addition, because only 3 examples are known, the lack of diverse xrRNA<sub>D</sub> sequences prevents conclusions about the role, prevalence, and structural diversity of this fold.

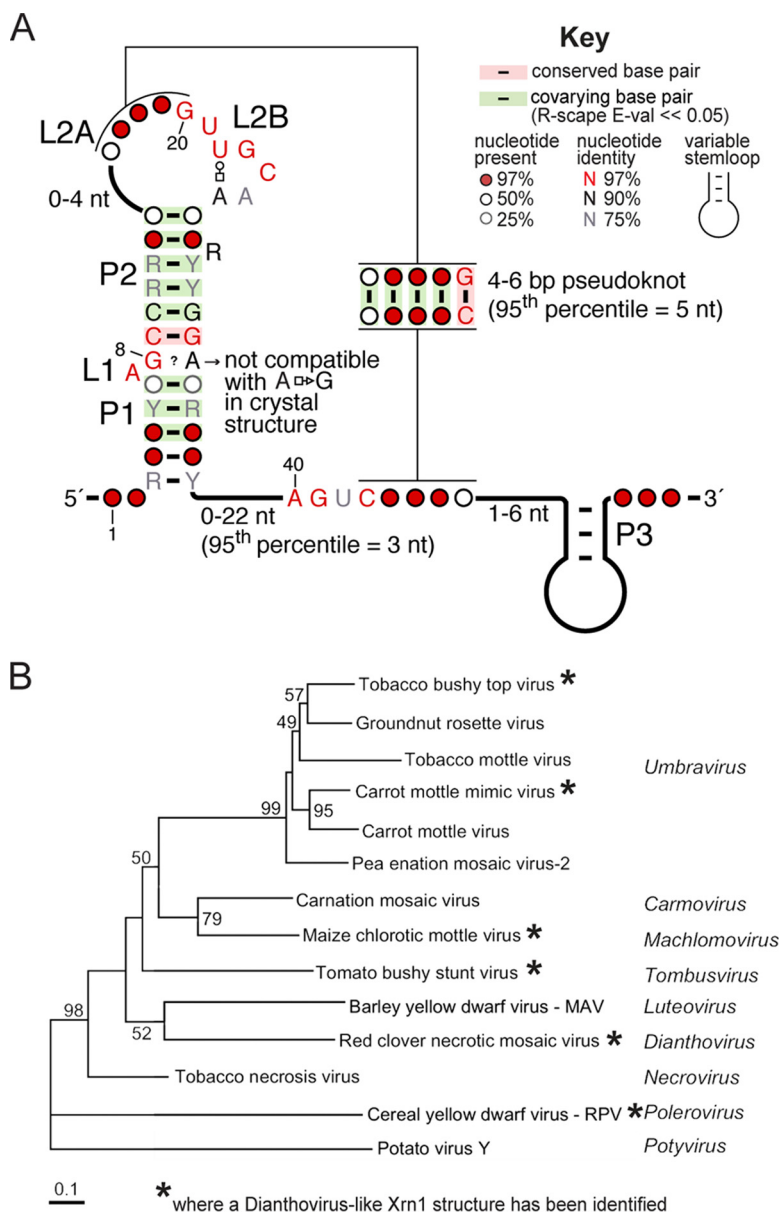
Because xrRNA<sub>F</sub> elements pervade the flaviviruses with associated sequence and structural diversity characteristics, it was puzzling that xrRNA<sub>D</sub> had been identified only in the three closely related members of the *Dianthovirus* genus. This observation raised the issue of whether xrRNAs similar to xrRNA<sub>D</sub> are more widespread and diverse than currently known; if so, it would indicate that they represent a more general and perhaps more important way to produce or protect viral RNAs than is currently recognized. This is true in a broader sense as well; the issue of how widespread xrRNAs are across biology remains open.

To begin to address both the specific issue of the presence of xrRNA<sub>D</sub> in other viruses and the more general issues about xrRNA diversity and distribution, we used a bioinformatic approach to search for more xrRNA<sub>D</sub> in a variety of plant viruses. We identified over 40 putative new xrRNA<sub>D</sub>-like elements in viruses belonging to the economically important *Tombusviridae* and *Luteoviridae* families. *In vitro* assays showed that these elements are indeed resistant to Xrn1, and analysis of these new xrRNAs revealed both conservation and variability. Surprisingly, we found many of these xrRNAs in intergenic regions of the viral genomic RNA, where they can be involved in the generation or maintenance of sgRNA species with protein-coding potential; hence, xrRNAs are not limited to noncoding RNA generation. These discoveries provide evidence that xrRNA-based RNA maturation pathways may be more widespread than previously anticipated and are involved in producing a variety of RNAs of diverse function.

## RESULTS AND DISCUSSION

To search for new xrRNA<sub>D</sub>-like elements, we used the *Infernal* software (S. R. Eddy laboratory), which enables screening of massive data sets of DNA sequences for conserved RNA secondary structure patterns with poor sequence conservation (27). Because the *Dianthovirus* genus has only three members (Red clover necrotic mosaic virus [RCNMV], Sweet clover necrotic mosaic virus [SCNMV], and Carnation ringspot virus [CRSV]) (26), we expanded our search to other plant-infecting positive-sense RNA viruses. The initial search within a library of viral reference genomes (see Materials and Methods) identified two potential sequences in the *Luteoviridae* corresponding to the poleroviruses wheat leaf yellowing-associated virus isolate JN-U3 (GenBank accession no. NC\_035451; *Infernal* E value = 0.00025, score = 44.3) and sugarcane yellow leaf virus (GenBank accession no. NC\_000874; *Infernal* E value = 6.5, score = 24.2). With these sequences added to the alignment, subsequent searches identified >40 candidates within the public repository of all available sequences for *Tombusviridae* and *Luteoviridae*, demonstrating how powerful this tool is for computationally identifying putative functional elements in viral RNAs (28).

A close inspection of all putative xrRNA<sub>D</sub>-like elements revealed that their predicted



**FIG 2** Widespread occurrence of Xrn1-resistant RNAs among plant viruses. (A) Consensus sequence and secondary structure of xrRNA<sub>D</sub> based on a comparative sequence alignment of 47 sequences of viruses belonging to the *Tombusviridae* and *Luteoviridae* families (shown in Fig. S1 in the supplemental material). Y = pyrimidine; R = purine. Non-Watson-Crick base pairs are shown using the Leontis-Westhof nomenclature (49). The numbering is that of the crystal structure of the SCNMV xrRNA (26). (B) Phylogenetic relationship between various plant viruses, based on the RNA-dependent RNA polymerase amino acid sequence (31). The viruses and corresponding genera in which we identified xrRNA<sub>D</sub> structures are marked by a star. Numbers at the nodes refer to bootstrap values as percentages obtained from 2,000 replications, shown only for branches supported by more than 40% of the data. Branch lengths are proportional to the number of changes. Further analysis will likely reveal xrRNA<sub>D</sub> elements in more of these viruses with additional sequence and structural variation.

secondary structures contain the same elements as were found in the known xrRNA<sub>D</sub>. Specifically, the assertion of formation of helices P1 and P2 and a predicted pseudoknot is supported by covariation data, which reveal little sequence conservation (R-scape [29] E values for the 12 covarying base pairs in the stems and the pseudoknot are within 3.10<sup>-4</sup> to 8.10<sup>-13</sup> (95<sup>th</sup> percentile = 1.10<sup>-12</sup>); Fig. 2A). Notably, our search criteria did not contain the pseudoknot interaction; thus, the fact that all putative xrRNA<sub>D</sub> sequences have the ability to form this functionally important element serves as internal

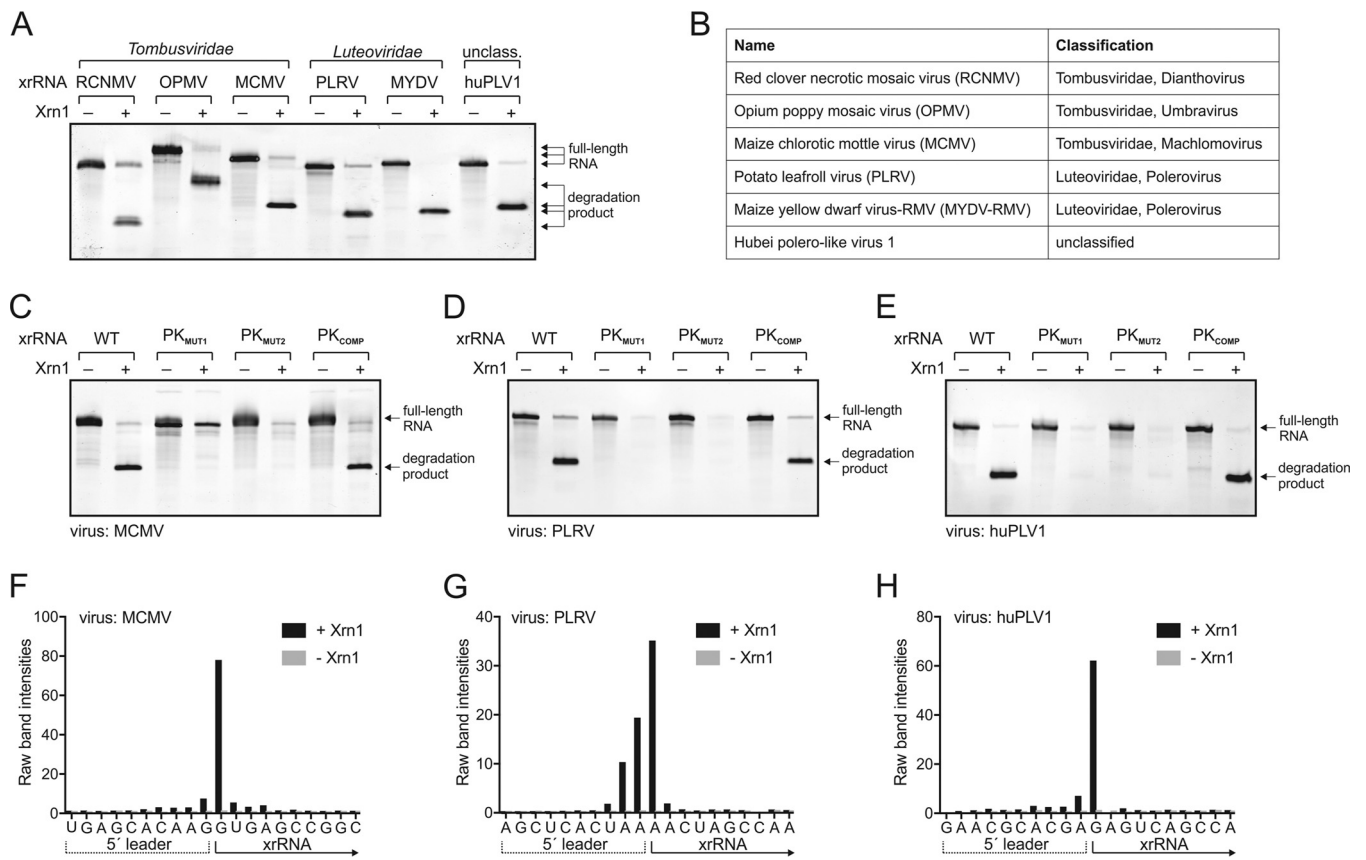
cross-validation. L1 and L2B are >97% conserved in sequence, consistent with their role in creating a specific folded motif that promotes pseudoknot formation (26). Also, two of the three nucleotides immediately upstream of the 3' side of the pseudoknot are >97% conserved, but their role is not obvious from the crystal structure of the open state. Likewise, the non-Watson-Crick A8-G33 base pair identified in the crystal structure (Fig. 1) cannot be reconciled with the predominant presence of G at position 8 and G/A at position 33 in all the other sequences. These observations support the previous assertion that the crystallized open state represents a folding intermediate of xrRNA<sub>D</sub> and that structural adjustments and additional interactions are present in the "closed" pseudoknot state.

Viruses in which we found putative novel xrRNAs include members of the *Tombusviridae* and *Luteoviridae* families. In the *Tombusviridae*, xrRNAs were found in the *Machlomovirus* and *Umbravirus* genera. In the *Luteoviridae* family, members of the *Polerovirus* and *Enamovirus* genera contain putative xrRNAs. We did not find putative xrRNAs in the *Luteovirus* genus despite its close relationship to the *Dianthovirus* or in the *Sobemovirus* genus, which is closely related to *Polerovirus*. It is possible that these viruses do not have xrRNAs or that they may have xrRNAs that are more divergent in sequence and secondary structure and thus would not be identified with our search criteria. We chose to remain conservative with respect to this search; future work may identify new elements in these viruses as well.

To experimentally determine if the computationally identified elements were authentic xrRNAs, we tested representative sequences from viruses of both families using our established *in vitro* Xrn1 resistance assay (11). Specifically, *in vitro*-transcribed and purified RNA sequences from opium poppy mosaic virus (OPMV), Maize chlorotic mottle virus (MCMV), Potato leafroll virus (PLRV), Maize yellow dwarf virus-RMV (MYDV-RMV), and Hubei polero-like virus 1 (HuPLV1) were challenged with recombinant Xrn1. All RNAs stopped Xrn1 degradation similarly to positive-control RCNMV xrRNA<sub>D</sub> (Fig. 3A and B), demonstrating that they are authentic xrRNAs that do not require additional *trans*-acting proteins for function. Moreover, mutations to disrupt the putative pseudoknot in the MCMV, PLRV, and HuPLV1 xrRNAs abolished Xrn1 resistance, while compensatory mutations that restore pseudoknot base pairing rescued the activity (Fig. 3C to E), verifying the functional importance of the pseudoknot in all of these examples. In addition, the mapped Xrn1 stop site is at the base of P1 in all newly identified xrRNAs, matching the xrRNA<sub>D</sub> stop site (Fig. 3F to H; see also Fig. S2 in the supplemental material) (26). Overall, the conserved secondary structure (Fig. 2A), the location of the exoribonuclease stop site, and the strict dependence on the pseudoknot for Xrn1 resistance suggest that these newly identified and tested xrRNAs use molecular folds and mechanisms similar to those seen with xrRNA<sub>D</sub>. By extension, the same is very likely true of the larger set that we computationally identified; thus, we classify them as members of the xrRNA<sub>D</sub> class of exonuclease-resistant RNA structures.

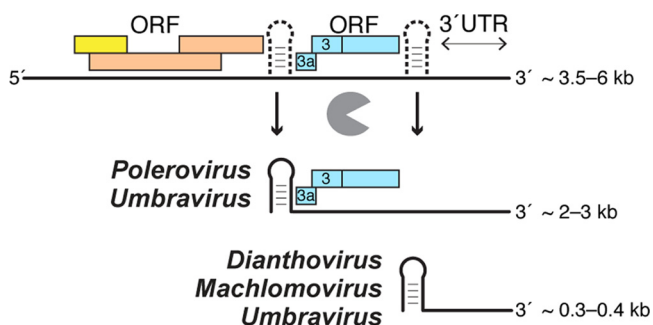
Although the newly identified xrRNA<sub>D</sub> elements share many features, there are notable structural differences in a subset of xrRNAs found in the *Tombusviridae* family (RCNMV, SCNMV, CRSV, OPMV, and MCMV). Specifically, these xrRNA<sub>D</sub> have a P3 stem-loop immediately downstream of the pseudoknot (Fig. 2A) (Table 1; see also Table S1 in the supplemental material). Although the presence of P3 was not recognized in the previous characterization of dianthovirus xrRNAs, the truncation analysis in that study showed that this part of the sequence is not required by xrRNA<sub>D</sub> for Xrn1 resistance *in vitro* (26). Consistent with this, an analogous stem-loop (P4) found in xrRNA<sub>F1</sub> is also dispensable *in vitro*; the crystal structure indicates that it may stabilize the pseudoknot through stacking interactions (Fig. 1) (24). Therefore, in xrRNA<sub>D</sub>, coaxial stacking of P3 on P1/P2 could similarly stabilize the RNA structure in the cell during infection, but it is not necessary in all contexts.

All previously known xrRNAs lie upstream of noncoding RNAs and lead to the generation of sgRNAs that do not encode proteins; however, the location of the newly discovered xrRNAs reveals unexpected variation and thus potential new roles for xrRNAs in general (Fig. 4). Surprisingly, only two of the newly identified xrRNAs are in



**FIG 3** Biochemical characterization of representative plant virus xrRNA<sub>D</sub> elements. (A) *In vitro* Xrn1 resistance assay of xrRNA<sub>D</sub> from various plant RNA viruses (Table 1). The xrRNA from RCNMV was included as a positive control. Arrows indicate the size of full-length RNAs and Xrn1-resistant degradation products. (B) Classification of viruses used in the experiments represented in panel A (Table 1). (C to E) *In vitro* Xrn1 resistance assay of wild-type (WT) and pseudoknot (PK) mutant versions of MCMV (C), PLRV (D), and HuPLV1 (E) xrRNAs. (F to H) Reverse transcription (RT) mapping of the Xrn1 stop site. Data represent distributions of RT products of Xrn1-resistant fragments of MCMV (F), PLRV (G), and HuPLV1 (H) degradation fragments. Experimentally validated stop sites are indicated on the secondary structure diagrams for all tested xrRNA<sub>D</sub> shown in Fig. S2.

the 3'UTR of the viral genome (Table 1). In MCMV, the first nucleotide of the P1 helix matches the 5' end of noncoding sgRNA2 (30); thus, as with the dianthoviruses, flaviviruses, and other xrRNAs, this new element probably blocks Xrn1 to generate noncoding sgRNAs derived from the 3'UTR. In contrast, for some members of the



**FIG 4** xrRNA<sub>D</sub> can produce or protect both coding and noncoding sgRNAs. The presence of xrRNA<sub>D</sub> in different contexts suggests an expanded role for these RNA elements. Full-length viral genomic RNA (top; colored boxes symbolize ORFs) can be processed by exoribonucleases that stop at xrRNAs (depicted as dashed structures) to yield both protein-coding sgRNAs (middle) and noncoding sgRNAs (bottom). Also, sgRNAs produced by subgenomic promoters could be “trimmed” or protected by xrRNAs (not shown). Only some umbraviruses (e.g., OPMV) possess two xrRNA<sub>D</sub> elements. Note that the ORF organization characteristics are not identical in all of these viruses; thus, this depiction should be considered conceptual. Details of the genetic organization and xrRNA location can be found in Fig. S3.

**TABLE 1** Selected set of plant viruses possessing an xrRNA<sub>D</sub><sup>a</sup>

Name	Abbreviation	Classification	GenBank accession no.	Length (nt)	Genomic location (nt) <sup>b</sup>	Genomic context
Red clover necrotic mosaic virus	RCNMV	<i>Tombusviridae</i> ; <i>Dianthovirus</i>	<a href="#">NC_003756</a>	3,890	3461–3504	3' UTR
Sweet clover necrotic mosaic virus	SCNMV	<i>Tombusviridae</i> ; <i>Dianthovirus</i>	<a href="#">NC_003806</a>	3,876	3446–3489	3' UTR
Maize chlorotic mottle virus (isolate KS1)	MCMV	<i>Tombusviridae</i> ; <i>Machlomovirus</i>	<a href="#">NC_003627</a>	4,437	4101–4143	3' UTR
Opium poppy mosaic virus (isolate PHEL5235)	OPMV	<i>Tombusviridae</i> ; <i>Umbravirus</i>	<a href="#">NC_027710</a>	4,230	3585–3629	3' UTR
Carrot mottle mimic umbravirus	CMoMV	<i>Tombusviridae</i> ; <i>Umbravirus</i>	<a href="#">NC_001726</a>	4,201	2664–2706	74 nt to AUG from ORF3
Chickpea chlorotic stunt virus	CpCSV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_008249</a>	5,900	3489–3534	11 nt to AUA from ORF3a; 129 nt to AUG from ORF3–5
Cowpea polerovirus 1 (isolate BE167)	CpPV1	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_034246</a>	5,845	3380–3425	11 nt to CUG from ORF3a; 129 nt to AUG from ORF3–5
Cotton leafroll dwarf virus	CoLRDV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_014545</a>	5,866	3451–3499	13 nt to CUG from ORF3a; 131 nt to AUG from ORF3–5
Cereal yellow dwarf virus-RPV	CYDV-RPV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_004751</a>	5,723	3566–3622	14 nt to AUU from ORF3a; 132 nt to AUG from ORF3–5
Maize yellow dwarf virus-RMV (formerly BYDV)	MYDV-RMV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_021484</a>	5,612	3335–3384	14 nt to ACG from ORF3a; 132 nt to AUG from ORF3–5
Potato leafroll virus	PLRV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_001747</a>	5,987	3509–3557	18 nt to AUA from ORF3a; 136 nt to AUG from ORF3–5
Sugarcane yellow leaf virus	ScYLV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_000874</a>	5,899	3467–3512	18 nt to CUG from ORF3a; 136 nt to AUG from ORF3–5
Beet Western yellows virus	BWYV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">NC_004756</a>	5,666	3346–3393	Defective ORF3a; 138 nt to AUG from ORF3–5
Beet Western yellows luteovirus (strain bwyv-1, isolate 28a)	BWYV	<i>Luteoviridae</i> ; <i>Polerovirus</i>	<a href="#">L39983</a>	973	341–389	135 nt to AUG from ORF3–5
Hubei polero-like virus 2 (strain QTM26674)	HuPLV2	<i>Unclassified</i>	<a href="#">NC_033229</a>	6,083	3706–3753	133 nt to AUG from ORF3–5
Hubei polero-like virus 1 (strain WHCC118254)	HuPLV1	<i>Unclassified</i>	<a href="#">NC_032224</a>	4,213	3357–3410	135 nt to AUG from ORF3–5

<sup>a</sup>Viruses are grouped by the genomic context of the xrRNA (last column). The complete list of sequences used for comparative sequence alignment is shown in Table S1. Smirnova et al. (30) was used as a reference for updated annotations of ORF3a. BYDV, Barley yellow dwarf virus; nt, nucleotide(s).

<sup>b</sup>xrRNA boundaries are defined as the first nucleotide of the P1 stem and the last nucleotide of the pseudoknot.

*Tombusviridae* family as well as for poleroviruses, xrRNA<sub>D</sub> is located in an intergenic region 5 to 20 nucleotides upstream from the translation start site of open reading frame 3a (ORF3a) and ~135 nucleotides from the start site of a readthrough protein encoded by ORF3 to ORF5 (ORF3–5) (our data suggest that ORF3a has not been annotated for all poleroviruses; see Table S1). ORF3a codes for protein P3a, which is essential for long-distance movement of the virus in plants (30). Translation of ORF3a initiates from sgRNA1 at a non-AUG codon (Table 1; see also Table S1) (30–32). The location of the xrRNA<sub>D</sub> upstream of the ORF3a start site implies that these xrRNAs, rather than functioning in noncoding RNA production, might act to produce or maintain protein-coding RNAs (Fig. 4).

The presence of xrRNA<sub>D</sub> in intergenic regions and upstream of protein-coding sequences suggests several possibilities for the role of xrRNAs in this new context. First, these intergenic xrRNA<sub>D</sub> could be used to produce sgRNAs similarly to the role ascribed to xrRNAs in 3'UTRs; that is, sgRNAs could be produced by incomplete degradation of full-length genomic RNAs without requiring a subgenomic promoter. Alternatively, precursor sgRNAs could be produced by transcription from a subgenomic promoter or from templates made by premature termination during negative-strand synthesis. These precursor RNAs could be subsequently “trimmed” by exonucleases to yield a mature sgRNA. In this scenario, the transcription start site could be at any distance from the 5' end of the mature sgRNA. Another possibility is that the sgRNA could be produced by transcription from a subgenomic promoter and that the role of the xrRNA would be to protect sgRNA from 5'-to-3' degradation by cellular exoribonucleases. Our

data suggest that all of the scenarios are possible and that the roles of intergenic xrRNA<sub>D</sub> may be different in different viruses. For example, unlike in MCMV, the 5' end of the xrRNA<sub>D</sub> sequence in PLRV (*Luteoviridae*; *Polerovirus*) does not correspond to the mapped 5' end of the sgRNA1. Rather, it is located 28 nucleotides upstream of the proposed exoribonuclease stop site. Furthermore, previous studies showed that PLRV sgRNA1 is likely generated by a replicative mechanism and thus that the xrRNA in PLRV probably does not directly function in the initial generation of sgRNA1 (33). This organization is also found in beet Western yellows virus (also a polerovirus) (34). In this case, the xrRNA<sub>D</sub> might be involved in a regulatory "trimming" step that alters the 5' end of existing transcripts, a process that may therefore occur only under certain conditions or at certain times in viral infection. Before testing this hypothesis, it would be crucial to determine if, when, and where xrRNA-dependent sgRNAs accumulate in infected cells and if there is variation in the 5' end within the population of a specific sgRNA. Moreover, mapping precise 5' ends of additional sgRNAs from different virus species (and comparing them to the location of xrRNA<sub>D</sub>) is needed to provide insight into the potential functions of xrRNAs during the generation, maintenance, and regulation of viral coding and noncoding sgRNAs.

That xrRNA<sub>D</sub> are at or near the 5' end of protein coding sgRNAs raises the issue of whether or not sgRNAs produced or maintained by intergenic xrRNA<sub>D</sub> can be translated. The 5' end of sgRNAs resulting from xrRNA-dependent halting of 5'-to-3' degradation would not have a modified nucleotide 5' "cap"; thus, sgRNAs lack the canonical translation initiation signal. However, since viruses of the *Tombusviridae* family use 3'-proximal cap-independent translation enhancers (3'-CITEs) to initiate translation, uncapped sgRNAs with xrRNAs on their 5' ends could potentially be translationally active (35, 36). In addition, it has been shown that some of these viruses use diverse 3'-CITEs and different amounts of various sgRNA species to fine-tune viral protein production during infection. Thus, these xrRNAs could be part of a larger RNA structure-dependent mechanism involving 3'-CITEs and xrRNAs in regulation of both the amount and translational activity of protein-encoding viral genomic RNAs and sgRNAs (37). Again, the details of such mechanisms would almost certainly differ in various viral species.

In contrast to the members of the *Tombusviridae* family, not all *Luteoviridae* members contain 3'-CITEs; thus, if xrRNA<sub>D</sub>-associated sgRNAs are translated, they must use a different mechanism of translation initiation. For example, in the *Polerovirus* genus, cap-independent translation is likely conferred by genome-linked proteins (VPg) that are covalently attached to the 5' end of the viral genome (W. Allen Miller, personal communication). The related sobemoviruses also have a VPg attached to their sgRNAs (38–41), but whether this is also true of polerovirus sgRNAs is currently unknown. Any sgRNAs resulting from xrRNA<sub>D</sub>-dependent exoribonuclease resistance would be expected to have a 5' monophosphate and not a VPg, raising the possibilities that such sgRNAs would be translationally inactive but could be maintained for some regulatory purpose and that there could be pools of translationally active and inactive sgRNAs. These possibilities all remain speculative; understanding the purpose of xrRNA<sub>D</sub> in each virus, how they relate to translation, and the existence of any larger trends will require ongoing detailed studies of diverse viruses.

Overall, our discoveries suggest that the roles of xrRNAs are more diverse than previously realized, depending on their genetic context. The presence of xrRNAs in various locations within viral genomes suggests that new xrRNA scaffolds may emerge from analyzing sgRNA 5' termini from other viruses; certainly, not all xrRNA elements were identified by the algorithm used here (5, 7, 42). Intriguing candidates for novel xrRNA identification include viruses with no obvious upstream promoter elements for sgRNA production and viruses in which putative promoter sequences do not seem to correspond well to the sgRNA 5' end (1, 5, 42, 43).

Many issues remain that pertain to understanding the structural/sequence requirements for Xrn1 resistance, the degree to which structural variation is tolerated, and how sequence diversity is integrated into similar folds (44). The now-expanded set of xrRNA<sub>D</sub>



**TABLE 2** Sequences used in the *in vitro* Xrn1 degradation assays<sup>a</sup>

RNA	Sequence (5'–3')
OPMV xrRNA	TAATACGACTCACTATAGGAATTGCCTCCACCAGTAACAAACCCAA <b>CC</b> CACACGCCAAGCATTAAAGTTGCAAGCGTTGGAGTGGCAGGCTT AACGTCGGACAGTACGACAACACTGCGG
MCMV xrRNA	TAATACGACTCACTATAGGTTCCAGGCCAGGGCTGGCAAATCATTGAGCACAAG <b>GT</b> GAGCCGGCATGAGGTTGCAAGACCCGGAAACAACC AGTCCTTCTGGCAGAGTCTGCAAA
PLRV xrRNA	TAATACGACTCACTATAgGCCACCACAAAAGAACTGAAGGAGCTCACTAAAAGTACCAAGCATAACGAGTTGCAAGCATTGGAAGT TCAAGCCTCGT
MYDV-RMV xrRNA	TAATACGACTCACTATAgTCCAGAAACAAAAGTTAAACAGAAAGCTCTCAAGTCAGCCAGGCAAATTCGAGTTGCAAGCACTGGATG ACCTAGTCTCGATA
HuPLV1 xrRNA	TAATACGACTCACTATAgGCCACCACAAAAGAAAGGAAGAACGCACGAGAGTCAGCCAAACAAACACAAGTTGCAAGTGTGGAGACT CATTCTAGTCTTGT

<sup>a</sup>The T7 promoter sequences are underlined, the leader sequence are indicated in italics, and the first protected nucleotides (experimentally validated as described here) are indicated in bold. Lowercase letters indicate extra nucleotides inserted to allow better transcription.

candidates provides a broader phylogeny for future bioinformatic and structural studies that will address these points.

## MATERIALS AND METHODS

**Computational search.** The published alignment performed with sequences from a total of three virus species (RCNMV, SCNMV, and CRSV) (26) was manually expanded in Ugene v. 1.29.0 (45) with two RCNMV variants (GenBank accession no. [J04357](#) and [AB034916](#)) retrieved from a standard nucleotide BLAST search for “somewhat dissimilar sequences” ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE\\_TYPE=BlastSearch](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch)). Sequences were aligned to the conserved 3D-based secondary structure, omitting the pseudoknot, and were exported in Stockholm format (see Fig. S4A in the supplemental material).

Using *Infernal* v. 1.1.2 (27) with default parameters, we searched for domains with similar structures and sequences within the complete reference genomes of viruses available from RefSeq, the NCBI Reference Sequence Database (<https://www.ncbi.nlm.nih.gov/refseq/>; downloaded on 10 January 2018). For subsequent iterations with *Infernal*, we searched the complete database of *Tombusviridae* and *Luteoviridae* available at GenBank (downloaded on 3 July 2018), using the alignment shown in Fig. S4B in the supplemental material.

In Ugene, we systematically added new hits from *Infernal* to the alignment only when they met the following criteria: (i) the sequence showed variation in more than 3 to 5 locations from the sequences already in the alignment; (ii) the *Infernal* E value was <0.05; (iii) the *Infernal* score was >10; (iv) the genomic context was coherent with that of the sequences already in the alignment. But a key objective in expanding the alignment further was also to analyze potential hits with a higher E value/a lower score, as they would often correspond to positive hits but with a larger sequence or structure variation. By the time the alignment reached a size of 10 to 12 sequences, we were able to retrieve most of the sequences that made it into the final alignment through further iterations of *Infernal* searches and manual addition to the alignment. Hits for unclassified viruses were also retrieved from large-scale transcriptomics data of invertebrate and vertebrate-associated RNA viruses using the deposited sequences (46, 47).

A statistical validation of the final proposed alignment of 47 sequences was performed using the latest version of R-scape available at <http://eddylab.org/R-scape/> (29) (last accessed on 17 August 2018). The corresponding conserved structure and sequence patterns were rendered using R2R v. 1.0.5 (48).

**Design of RNAs for *in vitro* assays.** The DNA templates used for *in vitro* transcription were gBlocks ordered from IDT and were cloned into pUC19 and verified by sequencing. RNA constructs for Xrn1 degradation assays contained the xrRNA sequence plus ~30 nucleotides of the endogenous upstream sequence (“leader sequence”) to allow loading of the exoribonucleases. Table 2 shows the sequences used in *in vitro* Xrn1 degradation assays.

**RNA preparation.** DNA templates for *in vitro* transcription were amplified by PCR using custom DNA primers (IDT) and Phusion Hot Start polymerase (New England BioLabs). Transcription reaction mixtures (2.5 ml) were assembled using 1,000- $\mu$ l PCR reaction mixtures as the template (~0.2  $\mu$ M template DNA, a 6 mM concentration of each NTP, 60 mM MgCl<sub>2</sub>, 30 mM Tris [pH 8.0], 10 mM dithiothreitol [DTT], 0.1% spermidine, 0.1% Triton X-100, T7 RNA polymerase, and 2  $\mu$ l RNasin RNase inhibitor [Promega]) and incubated overnight at 37°C. After inorganic pyrophosphates were precipitated by centrifugation, the reaction mixtures were ethanol precipitated and purified on a 7 M urea–8% denaturing polyacrylamide gel. RNAs of the correct size were excised, eluted overnight at 4°C into ~40 ml of diethylpyrocarbonate (DEPC)-treated Milli-Q filtered water (Millipore), and concentrated using Amicon Ultra spin concentrators (Millipore). Mutations were introduced using mutagenized custom DNA reverse primers (Table 3).

***In vitro* Xrn1 resistance assays.** RNA (4  $\mu$ g) was resuspended in 40  $\mu$ l 100 mM NaCl–10 mM MgCl<sub>2</sub>–50 mM Tris (pH 7.5)–1 mM DTT and refolded at 90°C for 3 min and then at 20°C for 5 min. A 3- $\mu$ l volume of recombinant RppH (0.5  $\mu$ g/ $\mu$ l stock) was added, and the samples were split into two 20- $\mu$ l reaction mixtures (with or without exoribonuclease). A 1- $\mu$ l volume of the recombinant Xrn1 (0.8  $\mu$ g/ $\mu$ l stock) was added where indicated. All reaction mixtures were incubated for 2 h at 30°C using a thermocycler. The degradation reaction mixtures were resolved on a 7 M urea–8% denaturing polyacrylamide gel and stained with ethidium bromide.

**TABLE 3** Primers used in this study<sup>a</sup>

Primer	Sequence (5'–3')
OPMV_WT_rev	5'-CCGCAGTTGTCGTA CTGTCGG-3'
OPMV_PKmut1_rev	5'-CCGCAGTTGTCGTA CTGTCGGACG <b>AA</b> TTGCCTGCCACTCCAACGC-3'
OPMV_PKmut2_rev	5'-CCGCAGTTGTCGTA CTGTCGGACGTTAAGCCTGCCACTCCAACGCTTGCAAC <b>AA</b> TTTGCTT GGCTGTGGTTGG-3'
OPMV_PKcomp_rev	5'-CCGCAGTTGTCGTA CTGTCGGACG <b>AA</b> TTGCCTGCCACTCCAACGCTTGCAAC <b>AA</b> TTTGCTT GGCT GTGGTTGG-3'
MCMV_WT_rev	5'-TGGCAGGACTCTGCCAGAAGG-3'
MCMV_PKmut1_rev	5'-TGGCAGGACTCTGCCAG <b>CTCC</b> ACTGGTTGTTCCGGTCTTGC-3'
MCMV_PKmut2_rev	5'-TGGCAGGACTCTGCCAGAAGGACTGGTTGTTCCGGTCTTGC <b>AA</b> GGAGATGCCGGCTCACC TTGTGCTC-3'
MCMV_PKcomp_rev	5'-TGGCAGGACTCTGCCAG <b>CTCC</b> ACTGGTTGTTCCGGTCTTGC <b>AA</b> GGAGATGCCGGCTCACC TTGTGCTC-3'
PLRV_WT_rev	5'-ACGAGGCTTGAACCTCCAATGC-3'
PLRV_PKmut1_rev	5'- <b>TG</b> CTGGCTTGAACCTCCAATGCTTGC-3'
PLRV_PKmut2_rev	5'-ACGAGGCTTGAACCTCCAATGCTTGC <b>AA</b> AGCAGTATGCTTGGCTAGTTTATAGT-3'
PLRV_PKcomp_rev	5'- <b>TG</b> CTGGCTTGAACCTCCAATGCTTGC <b>AA</b> AGCAGTATGCTTGGCTAGTTTATAGT-3'
MYDV-RMV_WT_rev	5'-TATCGAGACTAGGTCATCCAGTGC-3'
huPLV_WT_rev	5'-ACAAGACTAGAATGAGTCTCC-3'
huPLV_PKmut1_rev	5'- <b>TG</b> TTGACTAGAATGAGTCTCCAACACTTGC-3'
huPLV_PKmut2_rev	5'-ACAAGACTAGAATGAGTCTCCAACACTTGC <b>AA</b> CAACAGTTTGTGGCTGACTCTCG-3'
huPLV_PKcomp_rev	5'- <b>TG</b> TTGACTAGAATGAGTCTCCAACACTTGC <b>AA</b> CAACAGTTTGTGGCTGACTCTCG-3'

<sup>a</sup>Mutated residues are indicated in bold.

**Mapping of the exoribonuclease stop site.** To determine the Xrn1 stop site at single-nucleotide resolution, 30  $\mu$ g *in vitro*-transcribed RNA was degraded using recombinant RppH and Xrn1 as described above (the reaction volume was scaled up to 300  $\mu$ l, and 20  $\mu$ l of each enzyme was used). The degradation reaction mixture was resolved on a 7 M urea–8% polyacrylamide gel, and then the Xrn1-resistant degradation product was cut from the gel and eluted overnight at 4°C into ~20 ml of diethylpyrocarbonate (DEPC)-treated Milli-Q filtered water (Millipore) and concentrated using Amicon Ultra spin concentrators (Millipore). Once recovered, the RNA was subjected to reverse transcription (RT) using Superscript III reverse transcriptase (Thermo) and a 6-carboxyfluorescein (FAM) (6-fluorescein amidite)-labeled sequence-specific reverse primer (IDT) with a 20 (A) stretch at the 5' end to allow cDNA purification with oligo(dT) beads. The RT reaction volumes (5  $\mu$ l) contained 1.2 pM RNA, 0.25  $\mu$ l 0.25  $\mu$ M FAM-labeled reverse primer, 1  $\mu$ l 5 $\times$  first-strand buffer, 0.25  $\mu$ l 0.1 M DTT, 0.4  $\mu$ l 10 mM deoxynucleoside triphosphate (dNTP) mix, and 0.1  $\mu$ l Superscript III reverse transcriptase (200 U/ $\mu$ l) and were incubated for 1 h at 50°C. To hydrolyze the RNA template after reverse transcription, 5  $\mu$ l of 0.4 M NaOH was added and the reaction mixture incubated at 90°C for 3 min, followed by cooling on ice for 3 min. The reaction was neutralized by adding 5  $\mu$ l of acid quench mix (1.4 M NaCl, 0.57 M HCl, 1.3 M sodium acetate, pH 5.2), and then 1.5  $\mu$ l of oligo(dT) beads [Poly(A)Purist MAG kit (Thermo)] was added and the cDNA was purified on a magnetic stand according to the manufacturer's instructions. The cDNA was eluted in 11  $\mu$ l ROX-HiDi and analyzed on a model 3500 Genetic Analyzer (Applied Biosystems) for capillary electrophoresis. A Sanger sequencing (ddNTP) ladder of the undigested RNA was analyzed alongside each degradation product as a reference for band annotation.

**Data availability.** All data are available from us.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.02461-18>.

**FIG S1**, TIF file, 12.6 MB.

**FIG S2**, JPG file, 0.3 MB.

**FIG S3**, JPG file, 0.5 MB.

**FIG S4**, PDF file, 0.3 MB.

**TABLE S1**, PDF file, 1.2 MB.

## ACKNOWLEDGMENTS

We gratefully acknowledge David Farrell for information technology (IT) support; the members of the laboratory of J.S.K. for useful discussions; and David Costantino for careful reading of the manuscript.

This work was supported by NIH grants R35GM118070 and R01AI133348 (J.S.K.) and DFG STE 2509/2-1 (A.-L.S.). We declare that we have no conflicts of interest. A.-L.S., Q.V., and J.S.K. designed and analyzed the research; Q.V. performed the computational search; A.-L.S. performed the biochemical experiments; A.-L.S., Q.V., and J.S.K. wrote the paper.

## REFERENCES

- Miller WA, Shen R, Staplin W, Kanodia P. 2016. Noncoding RNAs of plant viruses and viroids: sponges of host translation and RNA interference machinery. *Mol Plant Microbe Interact* 29:156–164. <https://doi.org/10.1094/MPMI-10-15-0226-F1>.
- Sztuba-Solińska J, Stollar V, Bujarski JJ. 2011. Subgenomic messenger RNAs: mastering regulation of (+)-strand RNA virus life cycle. *Virology* 412:245–255. <https://doi.org/10.1016/j.virol.2011.02.007>.
- Miller WA, Koev G. 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* 273:1–8. <https://doi.org/10.1006/viro.2000.0421>.
- Jiwan SD, White KA. 2011. Subgenomic mRNA transcription in Tombusviridae. *RNA Biol* 8:287–294. <https://doi.org/10.4161/rna.8.2.15195>.
- Koev G, Miller WA. 2000. A positive-strand RNA virus with three very different subgenomic RNA promoters. *J Virol* 74:5988–5996. <https://doi.org/10.1128/JVI.74.13.5988-5996.2000>.
- Shen R, Miller WA. 2004. Subgenomic RNA as a riboregulator: negative regulation of RNA replication by Barley yellow dwarf virus subgenomic RNA 2. *Virology* 327:196–205. <https://doi.org/10.1016/j.virol.2004.06.025>.
- 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* 10:137. <https://doi.org/10.3390/v10030137>.
- Pijlman GP, Funk A, Kondratieva N, Leung J, Torres S, van der Aa L, Liu WJ, Palmenberg AC, Shi PY, 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>.
- Roby JA, Pijlman GP, Wilusz J, Khromykh AA. 2014. Noncoding subgenomic flavivirus RNA: multiple functions in West Nile virus pathogenesis and modulation of host responses. *Viruses* 6:404–427. <https://doi.org/10.3390/v6020404>.
- 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>.
- Chapman EG, Moon SL, Wilusz J, Kieft JS. 2014. RNA structures that resist degradation by Xrn1 produce a pathogenic dengue virus RNA. *Elife* 3:e01892. <https://doi.org/10.7554/eLife.01892>.
- Moon SL, Anderson JR, Kumagai Y, Wilusz CJ, Akira S, Khromykh AA, Wilusz J. 2012. A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability. *RNA* 18:2029–2040. <https://doi.org/10.1261/rna.034330.112>.
- Charley PA, Wilusz CJ, Wilusz J. 2018. Identification of phlebovirus and arenavirus RNA sequences that stall and repress the exoribonuclease XRN1. *J Biol Chem* 293:285–295. <https://doi.org/10.1074/jbc.M117.805796>.
- Clarke BD, Roby JA, Slonchak A, Khromykh AA. 2015. Functional noncoding RNAs derived from the flavivirus 3' untranslated region. *Virus Res* 206:53–61. <https://doi.org/10.1016/j.virusres.2015.01.026>.
- Schnettler E, Tykalova H, Watson M, Sharma M, Sterken MG, Obbard DJ, Lewis SH, McFarlane M, Bell-Sakyl L, Barry G, Weisheit S, Best SM, Kuhn RJ, Pijlman GP, Chase-Topping ME, Gould EA, Grubhoffer L, Fazakerley JK, Kohl A. 2014. Induction and suppression of tick cell antiviral RNAi responses by tick-borne flaviviruses. *Nucleic Acids Res* 42:9436–9446. <https://doi.org/10.1093/nar/gku657>.
- Schnettler E, Sterken MG, Leung JY, Metz SW, Geertsema C, Goldbach RW, Vlak JM, Kohl A, Khromykh AA, Pijlman GP. 2012. Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and mammalian cells. *J Virol* 86:13486–13500. <https://doi.org/10.1128/JVI.01104-12>.
- Göertz GP, Fros JJ, Miesen P, Vogels CB, van der Bent ML, Geertsema C, Koenraadt CJ, van Rij RP, van Oers MM, Pijlman GP. 2016. Noncoding subgenomic flavivirus RNA is processed by the mosquito RNA interference machinery and determines West Nile virus transmission by *Culex pipiens* mosquitoes. *J Virol* 90:10145–10159. <https://doi.org/10.1128/JVI.00930-16>.
- Schuessler A, Funk A, Lazear HM, Cooper DA, Torres S, Daffis S, Jha BK, Kumagai Y, Takeuchi O, Hertzog P, Silverman R, Akira S, Barton DJ, Diamond MS, Khromykh AA. 2012. West Nile virus noncoding subgenomic RNA contributes to viral evasion of the type I interferon-mediated antiviral response. *J Virol* 86:5708–5718. <https://doi.org/10.1128/JVI.00207-12>.
- Bidet K, Dadlani D, Garcia-Blanco MA. 2014. G3BP1, G3BP2 and CAPRIN1 are required for translation of interferon stimulated mRNAs and are targeted by a dengue virus non-coding RNA. *PLoS Pathog* 10:e1004242. <https://doi.org/10.1371/journal.ppat.1004242>.
- Liu Y, Liu H, Zou J, Zhang B, Yuan Z. 2014. Dengue virus subgenomic RNA induces apoptosis through the Bcl-2-mediated PI3k/Akt signaling pathway. *Virology* 448:15–25. <https://doi.org/10.1016/j.virol.2013.09.016>.
- Manokaran G, Finol E, Wang C, Gunaratne J, Bahl J, Ong EZ, Tan HC, Sessions OM, Ward AM, Gubler DJ, Harris E, Garcia-Blanco MA, Ooi EE. 2015. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science* 350:217–221. <https://doi.org/10.1126/science.aab3369>.
- Funk A, Truong K, Nagasaki T, Torres S, Floden N, Balmori Melian E, Edmonds J, Dong H, Shi PY, Khromykh AA. 2010. RNA structures required for production of subgenomic flavivirus RNA. *J Virol* 84:11407–11417. <https://doi.org/10.1128/JVI.01159-10>.
- MacFadden A, O'Donoghue Z, Silva P, Chapman EG, Olsthoorn RC, Sterken MG, Pijlman GP, Bredenbeek PJ, Kieft JS. 2018. Mechanism and structural diversity of exoribonuclease-resistant RNA structures in flaviviral RNAs. *Nat Commun* 9:119. <https://doi.org/10.1038/s41467-017-02604-y>.
- Akiyama BM, Laurence HM, Massey AR, Costantino DA, Xie X, Yang Y, Shi PY, Nix JC, Beckham JD, Kieft JS. 2016. Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* 354:1148–1152. <https://doi.org/10.1126/science.aah3963>.
- Chapman EG, Costantino DA, Rabe JL, Moon SL, Wilusz J, Nix JC, Kieft JS. 2014. The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production. *Science* 344:307–310. <https://doi.org/10.1126/science.1250897>.
- 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>.
- Nawrocki EP, Eddy SR. 2013. Computational identification of functional RNA homologs in metagenomic data. *RNA Biol* 10:1170–1179. <https://doi.org/10.4161/rna.25038>.
- Lim CS, Brown CM. 2017. Know your enemy: successful bioinformatic approaches to predict functional RNA structures in viral RNAs. *Front Microbiol* 8:2582. <https://doi.org/10.3389/fmicb.2017.02582>.
- Rivas E, Clements J, Eddy SR. 2017. A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat Methods* 14:45–48. <https://doi.org/10.1038/nmeth.4066>.
- Smirnova E, Firth AE, Miller WA, Scheidecker D, Brault V, Reinbold C, Rakontodrafara AM, Chung BY, Ziegler-Graff V. 2015. Discovery of a small non-AUG-initiated ORF in poleroviruses and luteoviruses that is required for long-distance movement. *PLoS Pathog* 11:e1004868. <https://doi.org/10.1371/journal.ppat.1004868>.
- Ryabov EV, Taliansky ME, Robinson DJ, Waterhouse PM, Murrant AF, de Zoeten GA, Falk BW, Vetter HJ, Gibbs MJ. 2012. Genus umbravirus, p 1191–1195. *In* *Virus taxonomy: Ninth report of the International Committee on Taxonomy of Viruses*. Academic Press, Inc, San Diego, CA.
- Stevens M, Freeman B, Liu HY, Herrbach E, Lemaire O. 2005. Beet poleroviruses: close friends or distant relatives? *Mol Plant Pathol* 6:1–9. <https://doi.org/10.1111/j.1364-3703.2004.00258.x>.
- Miller JS, Mayo MA. 1991. The location of the 5' end of the potato leafroll luteovirus subgenomic coat protein mRNA. *J Gen Virol* 72:2633–2638. <https://doi.org/10.1099/0022-1317-72-11-2633>.
- Miller WA, Dinesh-Kumar SP, Paul CP. 1995. Luteovirus gene expression. *Crit Rev Plant Sci* 14:179–211. <https://doi.org/10.1080/07352689509701926>.
- 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>.
- Miller WA, White KA. 2006. Long-distance RNA-RNA interactions in plant virus gene expression and replication. *Annu Rev Phytopathol* 44:447–467. <https://doi.org/10.1146/annurev.phyto.44.070505.143353>.
- Newburn LR, White KA. 2015. Cis-acting RNA elements in positive-strand RNA plant virus genomes. *Virology* 479–480:434–443. <https://doi.org/10.1016/j.virol.2015.02.032>.

38. Ghosh A, Rutgers T, Ke-Qiang M, Kaesberg P. 1981. Characterization of the coat protein mRNA of Southern bean mosaic virus and its relationship to the genomic RNA. *J Virol* 39:87–92.
39. Reinbold C, Lacombe S, Ziegler-Graff V, Scheidecker D, Wiss L, Beuve M, Caranta C, Brault V. 2013. Closely related poleroviruses depend on distinct translation initiation factors to infect *Arabidopsis thaliana*. *Mol Plant Microbe Interact* 26:257–265. <https://doi.org/10.1094/MPMI-07-12-0174-R>.
40. Hacker DL, Sivakumaran K. 1997. Mapping and expression of Southern bean mosaic virus genomic and subgenomic RNAs. *Virology* 234:317–327. <https://doi.org/10.1006/viro.1997.8667>.
41. Hébrard E, Pinel-Galzi A, Fargette D. 2008. Virulence domain of the RYMV genome-linked viral protein VPg towards rice rymv1-2-mediated resistance. *Arch Virol* 153:1161–1164. <https://doi.org/10.1007/s00705-008-0087-9>.
42. Johnston JC, Rochon DM. 1995. Deletion analysis of the promoter for the cucumber necrosis virus 0.9-kb subgenomic RNA. *Virology* 214:100–109. <https://doi.org/10.1006/viro.1995.9950>.
43. Scheets K. 2000. Maize chlorotic mottle machlomovirus expresses its coat protein from a 1.47-kb subgenomic RNA and makes a 0.34-kb subgenomic RNA. *Virology* 267:90–101. <https://doi.org/10.1006/viro.1999.0107>.
44. Toor N, Keating KS, Taylor SD, Pyle AM. 2008. Crystal structure of a self-spliced group II intron. *Science* 320:77–82. <https://doi.org/10.1126/science.1153803>.
45. Okonechnikov K, Golosova O, Fursov M, UGENE team. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28:1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>.
46. Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, Wang W, Eden JS, Shen JJ, Liu L, Holmes EC, Zhang YZ. 2018. The evolutionary history of vertebrate RNA viruses. *Nature* 556:197–202. <https://doi.org/10.1038/s41586-018-0012-7>.
47. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539–543. <https://doi.org/10.1038/nature20167>.
48. Weinberg Z, Breaker RR. 2011. R2R—software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics* 12:3. <https://doi.org/10.1186/1471-2105-12-3>.
49. Leontis NB, Westhof E. 2002. The annotation of RNA motifs. *Comp Funct Genomics* 3:518–524. <https://doi.org/10.1002/cfg.213>.