AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENOME REPLICATION AND REGULATION OF VIRAL GENE EXPRESSION

Replication-Dependent Biogenesis of Turnip Crinkle Virus Long Noncoding RNAs

Shaoyan Zhang,[a] Rong Sun,[a] Camila Perdoncini Carvalho,[a] Junping Han,[a] Limin Zheng,[a] Feng Qu[a]

[a]Department of Plant Pathology, Ohio Agricultural Research and Development Center, The Ohio State University, Wooster, Ohio, USA

ABSTRACT Long noncoding RNAs (lncRNAs) of virus origin accumulate in cells infected by many positive-strand (+) RNA viruses to bolster viral infectivity. Their biogenesis mostly utilizes exoribonucleases of host cells that degrade viral genomic or subgenomic RNAs in the 5′-to-3′ direction until being stalled by well-defined RNA structures. Here, we report a viral lncRNA that is produced by a novel replication-dependent mechanism. This lncRNA corresponds to the last 283 nucleotides of the turnip crinkle virus (TCV) genome and hence is designated tiny TCV subgenomic RNA (ttsgR). ttsgR accumulated to high levels in TCV-infected *Nicotiana benthamiana* cells when the TCV-encoded RNA-dependent RNA polymerase (RdRp), also known as p88, was overexpressed. Both (+) and (−) strand forms of ttsgR were produced in a manner dependent on the RdRp functionality. Strikingly, templates as short as ttsgR itself were sufficient to program ttsgR amplification, as long as the TCV-encoded replication proteins p28 and p88 were provided in *trans*. Consistent with its replicational origin, ttsgR accumulation required a 5′ terminal carmovirus consensus sequence (CCS), a sequence motif shared by genomic and subgenomic RNAs of many viruses phylogenetically related to TCV. More importantly, introducing a new CCS motif elsewhere in the TCV genome was alone sufficient to cause the emergence of another lncRNA. Finally, abolishing ttsgR by mutating its 5′ CCS gave rise to a TCV mutant that failed to compete with wild-type TCV in *Arabidopsis*. Collectively, our results unveil a replication-dependent mechanism for the biogenesis of viral lncRNAs, thus suggesting that multiple mechanisms, individually or in combination, may be responsible for viral lncRNA production.

IMPORTANCE Many positive-strand (+) RNA viruses produce long noncoding RNAs (lncRNAs) during the process of cellular infections and mobilize these lncRNAs to counteract antiviral defenses, as well as coordinate the translation of viral proteins. Most viral lncRNAs arise from 5′-to-3′ degradation of longer viral RNAs being stalled at stable secondary structures. Here, we report a viral lncRNA that is produced by the replication machinery of turnip crinkle virus (TCV). This lncRNA, designated ttsgR, shares the terminal characteristics with TCV genomic and subgenomic RNAs and overaccumulates in the presence of moderately overexpressed TCV RNA-dependent RNA polymerase (RdRp). Furthermore, templates that are of similar sizes as ttsgR are readily replicated by TCV replication proteins (p28 and RdRp) provided from nonviral sources. In summary, this study establishes an approach for uncovering low abundance viral lncRNAs, and characterizes a replicating TCV lncRNA. Similar investigations on human-pathogenic (+) RNA viruses could yield novel therapeutic targets.

KEYWORDS Turnip crinkle virus, long noncoding RNA, plant viruses, positive-sense RNA virus

Viruses with single-stranded, positive sense (+) RNA genomes are among the most common pathogens of humans, animals, and plants (1). Well-known examples of human-pathogenic (+) RNA viruses encompass poliovirus, dengue virus (DENV), Zika virus,

Citation Zhang S, Sun R, Perdoncini Carvalho C, Han J, Zheng L, Qu F. 2021. Replication-dependent biogenesis of turnip crinkle virus long noncoding RNAs. J Virol 95:e00169-21. https://doi.org/10.1128/JVI.00169-21.

Editor Anne E. Simon, University of Maryland, College Park

Address correspondence to Feng Qu, qu.28@osu.edu.

Received 31 January 2021
Accepted 14 June 2021
Accepted manuscript posted online 23 June 2021
Published 25 August 2021

and most recently severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that is responsible for the ongoing COVID-19 pandemic (2–5). Plant-infecting (+) RNA viruses include tobacco mosaic virus (TMV), the first virus ever discovered, and many other agriculturally important crop pathogens. Despite their vastly different host tropism, (+) RNA viruses share many important characteristics. For example, they all use their genomic RNAs as mRNA to direct the translation of at least some of the virus-encoded proteins and replicate their genomes via negative-strand (−) replication intermediates. As a result, research on both animal and plant-infecting (+) RNA viruses has historically contributed to our understanding of these shared characteristics and informed on strategies for control and management of diseases caused by these viruses.

A recently rediscovered characteristic shared by animal and plant-infecting (+) RNA viruses is the accumulation of long noncoding RNAs (lncRNAs) of virus origin in cells they infect. lncRNAs are generally defined as RNAs that are more than 200 nucleotides (nt) long and yet lack recognizable coding capacities (6, 7). Viral lncRNAs have been observed in (+) RNA virus-infected plants for a long time (8–12). Similarly, many animal viruses of the family *Flaviviridae* were also reported to accumulate lncRNAs, designated subgenomic flavivirus RNAs [sfRNAs] in infected cells (13–15). These viral lncRNAs are almost always coterminal with the 3′ ends of the genomic RNAs of cognate viruses. However, the questions of how these viral lncRNAs are produced and what roles they play in viral infections remained unresolved until 2008, when sfRNAs, and lncRNAs of plant-infecting red clover necrotic mosaic dianthovirus (RCNMV), were independently established as products of stalled 5′-to-3′ degradation of longer viral RNAs by exoribonucleases encoded in host cell genomes, with the stalling caused by well-defined RNA secondary structures within these lncRNAs (16, 17).

Extensive follow-up investigations have not only revealed the existence of similar viral lncRNAs in many more viruses but also identified the primary 5′-to-3′ exoribonucleases—Xrn1 in animals and possibly XRN4 in plants—responsible for their biogenesis (18–27). The specific RNA structures that act to stall exonuclease degradation have been thoroughly examined and, in a few cases, the crystal structures resolved (19, 28–32). These lncRNA-resident structures have complex modular features, including multiple stem-loops interspersed with frequent and often long-distance pseudoknot interactions.

Perhaps the most astounding revelation from these investigations is that these viral lncRNAs, despite their replication-independent biogenesis, all have proviral functions during the life cycles of their parental viruses (33). These functions ranged from mitigating host antiviral defenses, such as RNA interference (RNAi), interferon signaling, and other forms of innate immunity, to regulating the translation of various virus-encoded proteins (17, 19, 23). Preservation of these important functions by viral lncRNAs in diverse (+) RNA viruses raises the following interesting question: could some viruses have evolved the ability to produce similar lncRNAs via *de novo* synthesis and/or amplification by virus-encoded replication proteins? Namely, could functional lncRNAs of some viruses be products of virus-mediated replication or transcription?

Here, we report the identification and characterization of a lncRNA derived from turnip crinkle virus (TCV). TCV is a small plant-infecting (+) RNA virus with a genome of 4,054 nucleotides (nt), encoding 5 proteins (34). The two proteins directly translated from the genomic RNA (gRNA) are the p28 auxiliary replication protein and the p88 RNA-dependent RNA polymerase (RdRp), with the latter being the product of infrequent translational readthrough overcoming the p28 stop codon (Fig. 1a). TCV also produces two subgenomic RNAs (sgRNAs) during cellular infections that encode the movement proteins p8 and p9 and coat protein (CP) p38, respectively. p38 is also the suppressor of RNA silencing encoded by TCV (35–37). The existence of a TCV-encoded lncRNA, designated tiny TCV subgenomic RNA (ttsgR), was first unveiled by overexpressing p88 in cells replicating a TCV replicon (38). In the current study, we further examined the biogenesis of ttsgR independent of a replicating TCV replicon and discovered a replication-based mechanism for ttsgR production. Our results revealed an

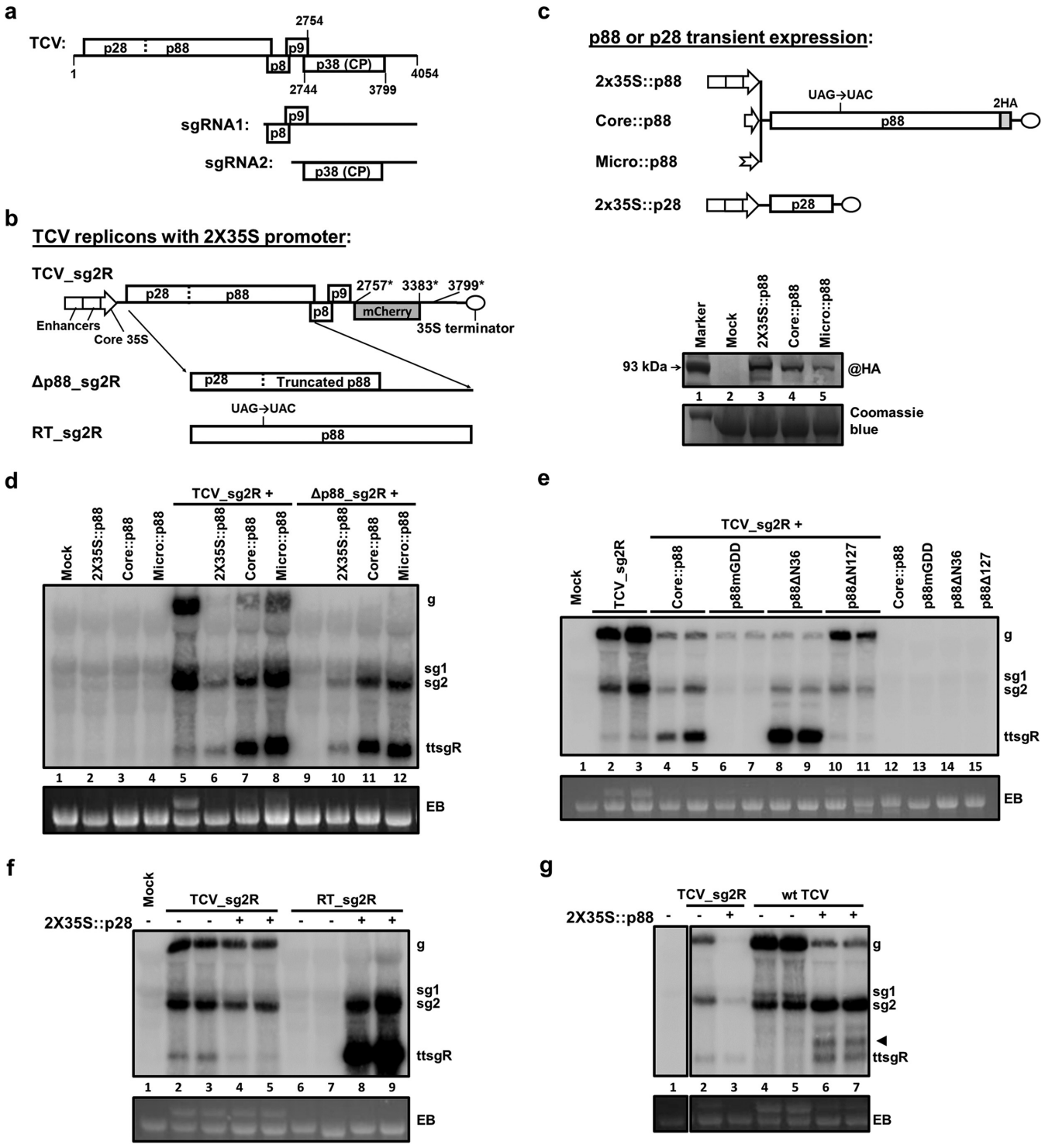


FIG 1 Moderately overexpressed p88 stimulates abundant ttsgR accumulation. (a) Schematic representation of the TCV genome with encoded proteins shown as rectangles. The two sgRNAs, namely, sgRNA1 and sgRNA2, are produced during cellular infections and serve as mRNA for p8 and p9 and for p38 (CP), respectively. (b) TCV replicons used in the current study. TCV_sg2R encodes wild-type (wt) p28 and p88 but expresses the red fluorescent mCherry instead of p38 from sgRNA2. Δp88_sg2R contains a 4-nt deletion in the middle of p88 ORF, making it incapable of translating a functional p88. Conversely, RT_sg2R contains a UAC codon in place of the p28 stop codon (UAG), rendering it incapable of translating p28 and yet overexpresses p88. All three replicons use the strong 2X35S promoter of CaMV to drive the transcription of infectious TCV RNA. (c) Constructs that expresses TCV p88 and p28 independent of replication. Three promoters, namely, 2X35S, Core35S, and Micro, were used to drive the expression of p88. Note that p88 expressed from these constructs contain a C-terminal, duplicated hemagglutinin (2HA) tag that permits protein detection with an HA antibody. This tag does not affect the function of p88 (38). In contrast, the p28-expressing construct contained a 2X35S promoter preceding the p28 ORF but no epitope tag at the C terminus.

(Continued on next page)

alternative viral lncRNA biogenesis strategy that does not involve 5′-to-3′ exonuclease degradation of longer viral RNAs.

RESULTS

ttsgR accumulates to high levels in the presence of moderately overexpressed p88. We recently reported that replication of a TCV replicon, TCV_sg2R, was inhibited by the overexpression of p88, the TCV-encoded RdRp (38). TCV_sg2R encodes the red fluorescent mCherry protein in place of TCV p38 (38, 39) (Fig. 1b), permitting convenient monitoring of its replication with confocal microscopy. Although TCV_sg2R was unable to produce the p38 RNA silencing suppressor (35), it replicated robustly in *N. benthamiana* cells when supplemented with p19 of tomato bushy stunt virus (TBSV). However, TCV_sg2R replication was repressed by p88 translated from a separate, nonreplicating mRNA (38); and this repression inversely correlated with the p88 protein levels (38–40). Intriguingly, ttsgR, a small TCV-derived RNA present at low levels in regular TCV_sg2R infections (Fig. 1d, lane 5), became highly abundant in the presence of moderate levels of p88 provided with the Core::p88 construct (Fig. 1c, bottom, lane 4; and Fig. 1d, lane 7). The Core promoter in this construct contained the last 99 nt of the cauliflower mosaic virus (CaMV) 35S promoter, devoid of the upstream enhancer (Fig. 1b). It drove a lower expression of p88 than that of the 2X35S::p88 construct (Fig. 1c, bottom, lanes 3 and 4). As a result, Core::p88 was less potent than 2X35S::p88 at repressing TCV_sg2R replication (Fig. 1d, lanes 6 and 7). Conversely, it was more competent at complementing the replication of the p88-defective Δp88_sg2R replicon (Fig. 1d, lanes 10 and 11; note that the complementation was inefficient for gRNA due to p88 overexpression), indicating that lowering p88 expression made it more replication competent (38). Accordingly, the high abundance of ttsgR in lanes 7 and 11 of Fig. 1d, in comparison with lanes 6 and 10, could be interpreted as preferential replication/transcription of shorter TCV RNAs in the presence of moderately overexpressed p88. This interpretation would be consistent with the concurrent overaccumulation of TCV sgRNA2 (Fig. 1d, lanes 6, 7 and 10, 11).

If this interpretation is correct, we would expect the level of ttsgR to increase even more if p88 levels were further decreased. To test this prediction, we next expressed p88 using a tobacco promoter known to be microspore specific (41) and, hence, only minimally active in *N. benthamiana* leaf cells (Fig. 1c, top, Micro::p88). The Micro::p88 construct indeed expressed a very low level of p88 protein (Fig. 1c, bottom, lane 5), although it was probably still higher than that p88 produced by readthrough translation from replicon genomes. Accordingly, its repression of TCV_sg2R replication was weaker than that of Core::p88 (Fig. 1d, lane 8). However, it was only marginally better at complementing the replication of Δp88_sg2R (Fig. 1d, lane 12). Nevertheless, in both cases, the ttsgR levels, along with that of sgRNA2, visibly increased. These results were consistent with the involvement of the p88 RdRp activity in ttsgR biogenesis.

To directly assess whether the RdRp activity of p88 was needed for ttsgR accumulation, we created p88mGDD, an RdRp-null mutant of p88, by changing the highly conserved GDD motif to VAA in the Core::p88 background. As shown in Fig. 1e (lanes 6 and 7), expression of p88mGDD caused a dramatic reduction of TCV_sg2R gRNA and sgRNAs and disappearance of ttsgR, recapitulating only the repressive activity of overexpressed p88 (38). This result contrasted with the selective repression of gRNA by Core::p88 that was accompanied by preferential accumulation of ttsgR and sgRNA2 (lanes 4 and 5) and, hence, implicating the RdRp activity of p88 in ttsgR overaccumulation.

FIG 1 Legend (Continued)

The image beneath the diagrams shows a Western blotting detection of the 2HA-tagged p88 from cells of *N. benthamiana* receiving the three p88 constructs. (d) Repression and replicational complementation of TCV replicons by p88 are accompanied by preferential overaccumulation of ttsgR and sgRNA2. The Northern blotting autoradiograph shows the accumulation levels of TCV gRNA (g), sgRNAs (sg1 and 2), and ttsgR in the total RNA samples extracted from *N. benthamiana* leaves that were treated with various constructs or combinations thereof. Note that a TBSV p19-expressing construct was always included to mitigate RNA silencing-mediated mRNA degradation. EB, ethidium bromide stained agarose gel serving as the loading control. (e) The RdRp activity of p88 is required for ttsgR overaccumulation. (f) p88 overexpression from a defective replicon (RT_sg2R) caused ttsgR overaccumulation upon replicational complementation by *trans*-supplied p28. (g) wt TCV also responded to p88 overexpression with the overaccumulation of ttsgR (and another lncRNA denoted with an arrowhead).

To further interrogate the role of p88 RdRp in ttsgR biogenesis, we additionally tested two deletion mutants of Core::p88, namely, p88ΔN36 and p88ΔN127, that removed the N-terminal 36- and 127-amino acid (aa) residues, respectively, from the p88 open reading frame (ORF) (38). We previously established that while p88ΔN36 retained both the repressive and complementing activities of p88 to a substantial extent, p88ΔN127 was modestly repressive but incapable of complementing p88-defective replicons. As shown in Fig. 1e, preferential overaccumulation of ttsgR occurred in the presence of p88ΔN36 (lanes 8 and 9) but not p88ΔN127 (lanes 10 and 11). Together, these results provided further support for the involvement of p88 RdRp activity in ttsgR biogenesis.

Additional observations corroborated the importance of p88 in ttsgR production. First, although p88 encompasses p28 at its N terminus, overexpression of the entire p88, rather than p28, was necessary for ttsgR overaccumulation (Fig. 1f, lanes 2 to 5). Second, for ttsgR overaccumulation to occur, p88 overexpression did not have to be from a nonreplicon source. Instead, abundant ttsgR (and sgRNA2) accumulated when the RT_sg2R mutant replicon was made to replicate with p28 provided in *trans* (Fig. 1f, lanes 6 to 9). The RT_sg2R mutant lacked the p28 stop codon, causing it to sacrifice p28 production in favor of p88 overproduction (Fig. 1b). As a result, its replication could occur only in the presence of p28 provided in *trans*. Third, ttsgR overaccumulation was triggered by p88 overexpression during the replication of not only TCV_sg2R but also wild-type (wt) TCV (Fig. 1g, lanes 4 to 7). Interestingly, another TCV lncRNA slightly larger than ttsgR also accumulated in wt TCV infections (arrowhead in Fig. 1g), the significance of which will be discussed.

ttsgR is unlikely to be a degradation product of longer TCV RNAs. Results in the previous section strongly hinted that ttsgR overaccumulation was caused by the RdRp activity of overexpressed p88. If true, this conclusion would be in stark contrast with lncRNAs associated with RCNMV and tobacco necrosis virus-D (TNV-D), which were both products of exonuclease-mediated 5′-to-3′ degradation of larger viral RNAs (17, 19). This contrast is noteworthy because TCV is highly similar to both RCNMV and TNV-D in terms of genome organization and sequence homology of replication proteins, with all three of them belonging to the family *Tombusviridae*. It in turn demanded more rigorous scrutiny of our observations.

The lncRNA of TNV-D, known as small viral RNA (svRNA), primarily arises from the degradation of TNV-D sgRNA1, one of its two sgRNAs (19). Incidentally, ttsgR overaccumulation was frequently accompanied by the preferential overproduction of TCV sgRNA2. To determine whether ttsgR abundance depended on either of the two TCV sgRNAs, we tested three TCV_sg2R mutants, namely, Δsg1, Δsg2, and Δsg1&2, in which the production of sgRNA1, sgRNA2, or both, were abolished through mutagenesis of sgRNA promoters (42, 43) (Fig. 2a). As shown in Fig. 2b, none of the three mutants abolished p88-dependent ttsgR accumulation, despite the loss of either or both sgRNAs in the corresponding mutants. Therefore, ttsgR was unlikely the degradation product of TCV sgRNAs.

We next examined whether ttsgR could have been derived from degradation of TCV gRNA. For this purpose, we generated three new constructs, as follows: [p28stop]_sg2R, sg2R_Temp, and ttsgR_Temp (Fig. 2c). Note that while [p28stop]_sg2R was a full-length replicon, the extra stop codon at aa position 36 of p28 prevented the production of both p28 and p88, making it incapable of launching replication by itself. The sg2R_Temp and ttsgR_Temp constructs were likewise incapable of self replication because their transcripts, 1,671 and 669 nt, respectively, were 3′ coterminal fragments of TCV_sg2R gRNA with their 5′ ends located at nucleotide positions 2468 and 3386, which are both downstream of the p28 and p88 ORFs (Fig. 2c). However, when p28 and p88 were provided in *trans* (with constructs 2X35S::p28 and Core::p88) (Fig. 1c), all three constructs produced TCV-specific RNAs of expected sizes (Fig. 2d). Importantly, accumulation of these RNAs depended on the presence of both p28 and p88 and required the GDD motif of p88 RdRp (Fig. 2d). Most significantly, all three templates additionally accumulated ttsgR in a manner dependent on the presence of both p28

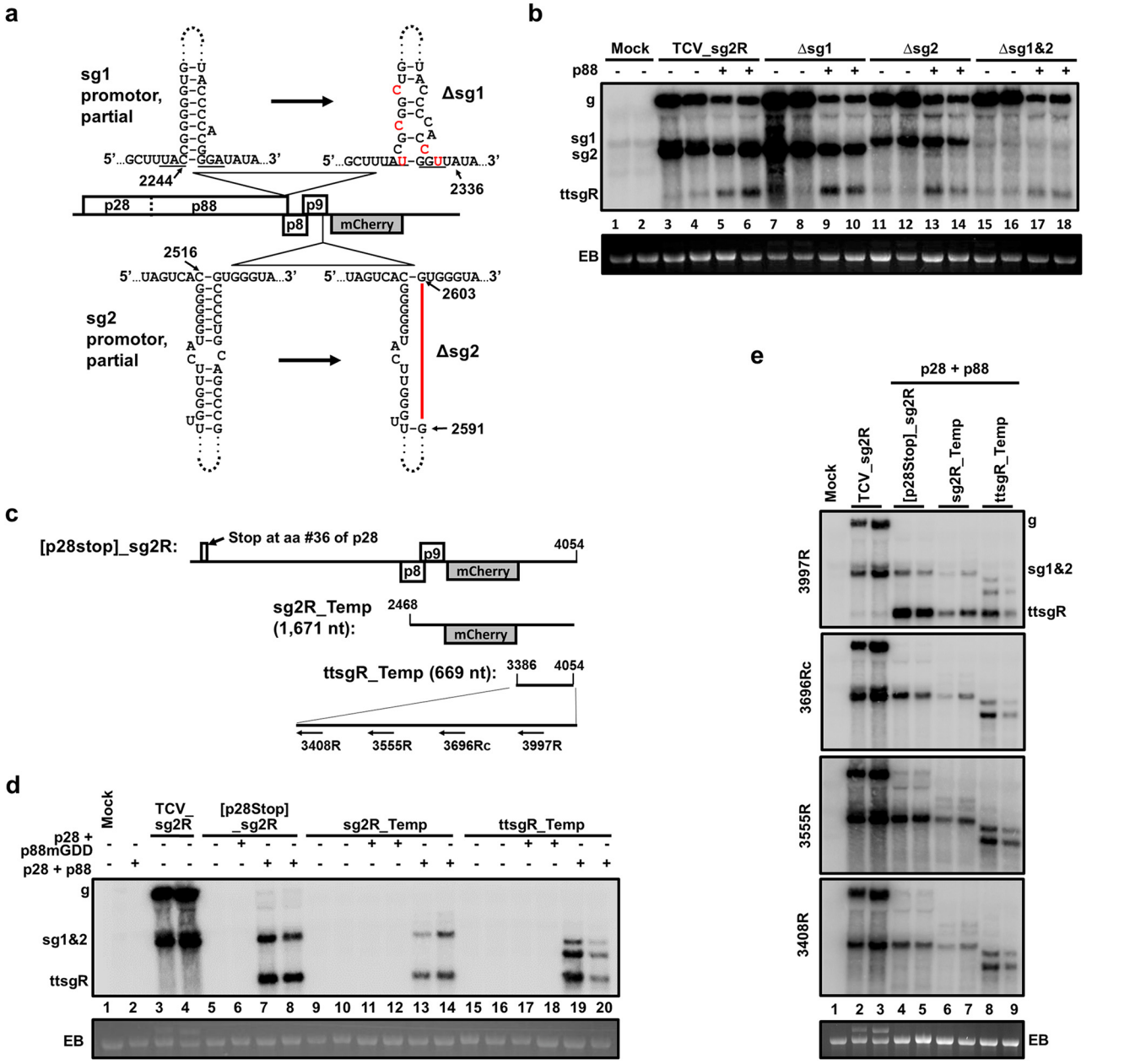
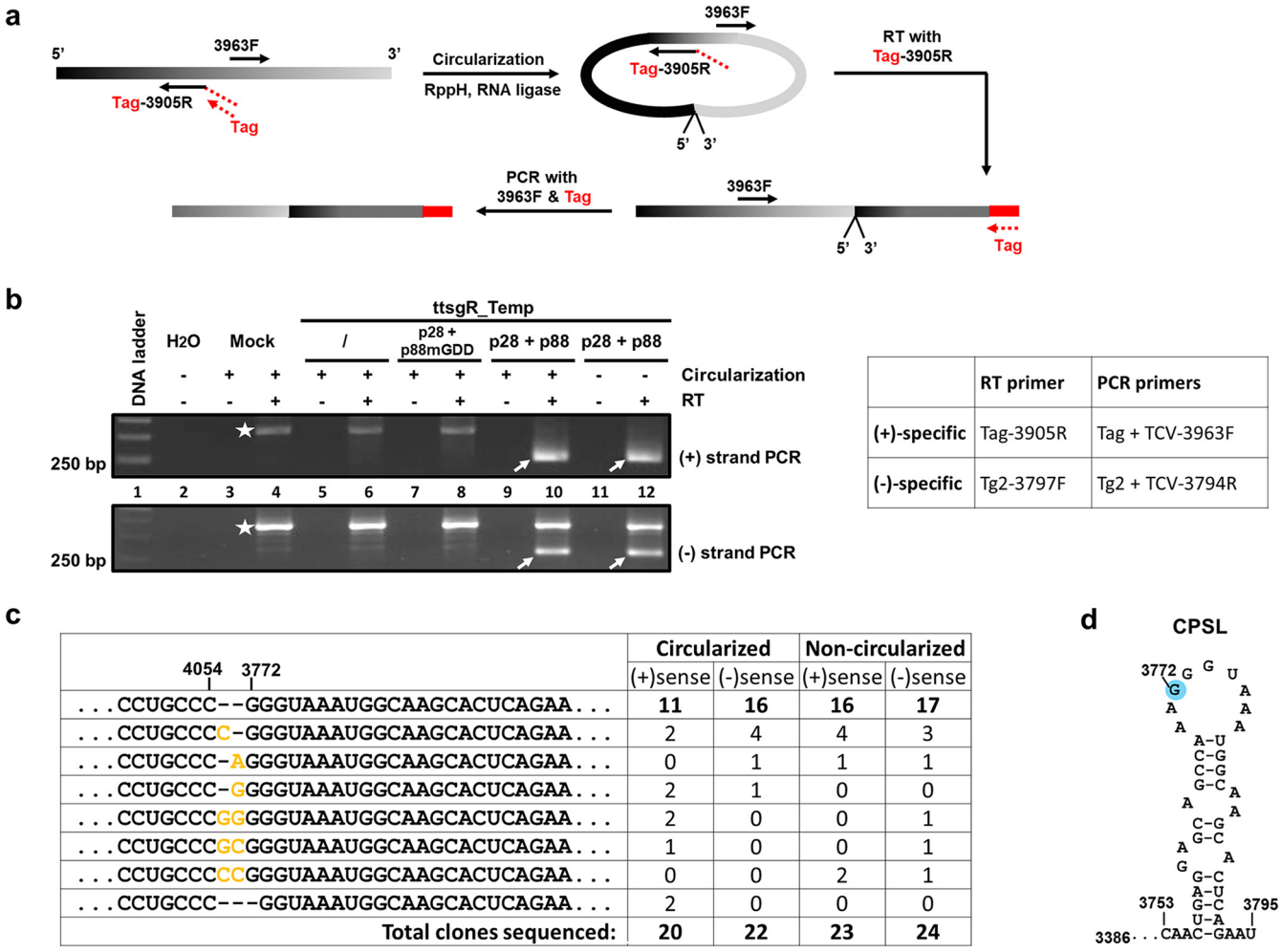


FIG 2 ttsgR is unlikely a product of 5′-to-3′ degradation of longer TCV RNAs. (a) Schematic depiction of Δsg1 and Δsg2 mutants of TCV_sg2R. The bottom portion of the stem-loop structure identified previously as the sgRNA1 promoter is shown on the top left. The Δsg1 mutant, with the five single-nucleotide mutations shown in red letters, is represented to the right. Similarly, the partial sgRNA2 promoter stem-loop is at the bottom left, with the corresponding Δsg2 mutant shown to the right, in which a deletion of 12 nt is denoted as a red line. (b) ttsgR accumulation was not abolished in Δsg1, Δsg2, and Δsg1&2 mutants, even though the levels of the corresponding sgRNAs were drastically diminished, as evidenced by Northern blotting results. (c) Schematic representation of three template constructs that transcribe 3′-coterminal TCV (TCV_sg2R)-derived RNAs of various lengths, with none of them capable of translating the replication proteins. (d) The three defective templates templated the production of ttsgR, in addition to RNAs of template lengths, in the presence of *trans*-supplied p28 and p88, but not p28 and p88mGDD. Note that the [p28stop]_sg2R template accumulated very low levels of full-length RNA but an easily detectable level of sgRNA2 and ttsgR. This outcome was expected, as overexpressed p88 complemented the replication of TCV gRNA poorly. (e) Preliminary mapping of the 5′ terminus of ttsgR by Northern blotting with four different probes.

and p88. Therefore, ttsgR was independent of the degradation of full-length TCV gRNA. On the contrary, ttsgR biogenesis strictly required the RdRp activity of p88 (Fig. 2d) and could occur with a template as short as the last 669 nt of TCV gRNA.

We next tried to pinpoint the 5′ terminus of ttsgR by subjecting these RNA samples to Northern blot hybridizations with four different oligonucleotide probes (Fig. 2c,



	RT primer	PCR primers
(+)-specific	Tag-3905R	Tag + TCV-3963F
(-)-specific	Tg2-3797F	Tg2 + TCV-3794R

FIG 3 The 5′ terminus of ttsgR maps to nucleotide position 3772 of TCV gRNA. (a) Step-by-step depiction of the circularization RACE. Note that (i) a non-TCV tag was fused to the 5′ end of the RT primer to enhance the strand specificity; and (ii) the two PCR primers (Tag and 3963F) are back-to-back from each other, and hence, a PCR product is only possible when the template is circularized or in dimeric form. (b) PCR products of both (+) and (−) strands arisen from circularization RACE. The stars on both agarose gels denote nonspecific products present even in mock-treated samples. The primers used are shown to the right. (c) Sequencing results of cloned PCR products (arrows in b) delimit the starting and ending nucleotide of ttsgR. (d) The position of the ttsgR 5′ terminus in a previously identified origin-of-assembly stem-loop within the CP ORF (CPSL).

bottom). As shown in Fig. 2e, ttsgR was readily detectable with the 3997R oligonucleotide probe, but not with the 3696Rc probe or two other probes that hybridized to more upstream positions of TCV gRNA. These experiments narrowed down the 5′ end of ttsgR to a position approximately between positions 3690 and 3997 of TCV, suggesting a ttsgR size of no more than 365 nt.

ttsgR initiates at a discrete position of TCV RNA, and it exists in both (+) and (−) strand forms. To ultimately prove ttsgR as the product of p88 RdRp activity, we needed to map its terminal sequences and ideally demonstrate that a template with the exact termini of ttsgR is amplified in the presence of p88 (and p28). To these ends, we first mapped the termini of ttsgR using a circularization rapid amplification of cDNA ends (RACE) procedure (44). Briefly, this procedure used T4 RNA ligase to connect the 5′ end of an RNA to its own 3′ end and, hence, circularize the RNA (Fig. 3a). The circularized RNA was then subjected to reverse transcription-PCR (RT-PCR) with divergent primers, followed by sequencing to discern the sequence of the 3′-to-5′ junction (Fig. 3a). Hence, this approach would allow us to map the 5′ and 3′ ends of ttsgR simultaneously. To further determine whether both (+) and (−) strands of ttsgR were present in the samples, we adopted a strand-specific (ss) RT-PCR procedure that utilized nonviral sequence tags (Tag and Tg2 in Fig. 3a and b) (sequences available upon request) to minimize the contamination of opposite strands (45).

As shown in Fig. 3b, specific, RT-dependent fragments were detected in samples in which ttsgR_Temp was administered along with p28 and p88 but not p28 and p88mGDD (compare lanes 8, 10, and 12). Furthermore, both (+) and (−) strands of ttsgR were successfully amplified. Surprisingly, RNA circularization was unnecessary for the amplification of both (+) and (−) strands, indicating that at least some of the ttsgR molecules existed as tandem dimers. The presence of (−) ttsgR, coupled with detection of dimers, is consistent with ttsgR undergoing active replication in the presence of p28 and p88.

The PCR products highlighted by arrows (Fig. 3b) were purified and cloned. For each of the PCR products, more than 20 individual clones were sequenced. As shown in Fig. 3c, regardless of the sense of the RNA, or whether the RNA was circularized, more than 50% of clones in each category shared a 3′-to-5′ junction that connected the 3′ end CCC with GGG beginning with nucleotide position 3772 of TCV. The rest of the clones mostly contained 1 to 2 extra G or C between the two termini. It should be noted that because circularized and noncircularized RNA samples yielded similar sequencing results, most of the junction sequences were probably derived from dimeric RNAs. As a result, some of the RNA species, especially those with rare junction identities, might have replicated as a dimeric rather than monomeric form. Nevertheless, the fact that most clones had 3772 joining 4054 led us to conclude that the ttsgR is predominantly 283 nt long and most likely arose from RdRp-mediated amplification that initiated at nucleotide position 3772. Notably, this initiation site is located within the end loop of a stem-loop structure previously identified as important for the packaging of TCV gRNA in virions (46) (Fig. 3d), prompting further investigations to determine whether the same RNA structure was needed for ttsgR biogenesis.

The hairpin structure flanking the 5′ end of ttsgR is not needed for ttsgR biogenesis. That ttsgR initiated at a position within the end loop of a previously identified hairpin structure (46) prompted us to evaluate the importance of this structure in ttsgR biogenesis. This was first carried out in the ttsgR_Temp background, which allowed us to disrupt the structure extensively. As shown in Fig. 4a, the ttsgR_mL and -mR mutants each contained eight single-nucleotide mutations (in red letters) that disrupted nearly all base pairs in the stem, whereas the ttsgR_mLR combined the mutations in the first two mutants to restore the base pairing in the stem, but not the original sequence. However, none of the three mutants abolished ttsgR accumulation. Therefore, this stem-loop structure was not required for ttsgR production from the ttsgR_Temp template.

We next tried to destabilize this structure, designated CP-resident stem-loop [CPSL], in the contexts of the TCV_sg2R replicon and wild-type (wt) TCV. The new mCPSL mutant had to preserve the identity of amino acid residues of TCV CP (p38) and, hence, contained just six single-nucleotide changes (Fig. 4c). Nevertheless, the CPSL was predicted to be substantially weakened. However, these changes failed to compromise p88-dependent ttsgR accumulation in either TCV_sg2R or wt TCV infections. Therefore, the stem-loop structure surrounding the ttsgR initiation site is not needed for ttsgR biogenesis from replicating TCV either.

A template with 28 extra 5′-end nucleotides reveals a specific terminal motif for ttsgR. The precise mapping of the ttsgR 5′ end allowed us to test a shorter template, referred to as ttsgR_S, that begins with TCV nucleotide position 3744, with just 28 extra nucleotides upstream of the ttsgR initiation site at nucleotide position 3772 (Fig. 5a). As shown in Fig. 5b (lanes 6 to 7), ttsgR_S permitted p28/p88-dependent accumulation of RNAs similar to ttsgR in size (compare lanes 6 to 7 with 3 to 4). Curiously, a slightly larger RNA species, designated ttsgR*, was also detected (arrows in lanes 6 and 7). In addition, another substantially larger, but fainter, species was detected as well (diamonds in lanes 6, 7, 9, and 10), probably representing ttsgR/ttsgR* dimers. The termini of ttsgR and ttsgR* were mapped with divergent pairs of primers designed to amplify their dimeric forms and, hence, bypassing RNA circularization. Sequence analysis of the RT-PCR products verified that ttsgR* extended ttsgR by 32 nt at the 5′ end, initiating at a site within the vector plasmid that was 3 nt upstream of the ttsgR_S insert (Fig. 5c). ttsgR and ttsgR* each accounted for 4 of the 18 clones sequenced (Fig. 5c). The 5′ ends of the remaining 10 sequences congregated near that of ttsgR (3 clones) or ttsgR* (7 clones). These results indicated that replication proteins p28 and p88 successfully amplified 2 RNAs with 2

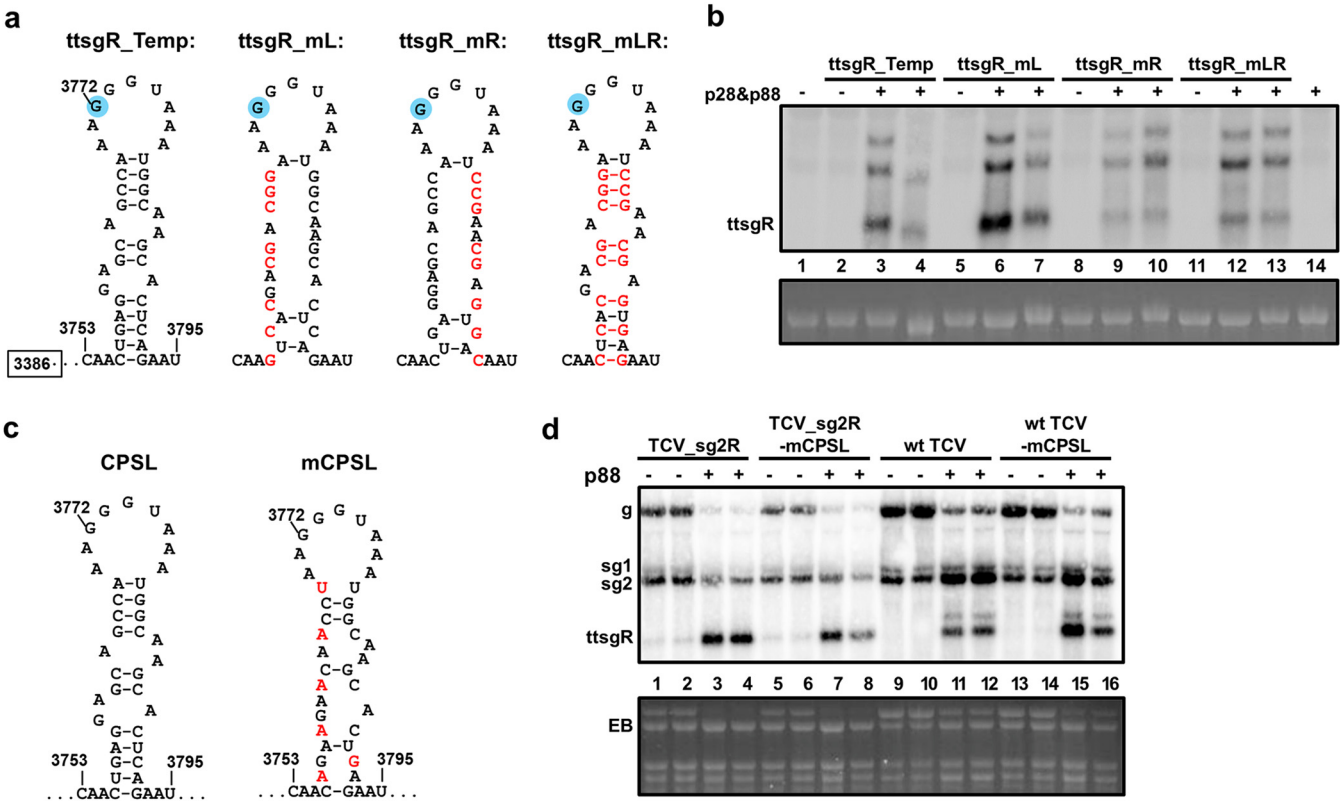


FIG 4 Disruption of the CPSL structure has no effect on ttsgR production from either the ttsgR_Temp (Fig. 2c), or TCV replicons. (a) Specific mutations (red letters) that disrupt CPSL in the ttsgR_Temp construct. (b) p28&p88-dependent ttsgR production from various ttsgR_Temp mutants detected with Northern blotting. (c) Mutations that disrupt CPSL in TCV replicons. (d) p88-dependent production of ttsgR from the mCPSL mutants of TCV_sg2R and wt TCV.

distinct initiation sites that were 32 nt apart. Interestingly, the first seven nucleotides of ttsgR and ttsgR*, namely, GGGUAAA and GAGAAAU, appeared to be strikingly similar; they began with GGG or GAG, followed by four A or U residues (Fig. 5c).

Closer inspection also revealed extensive sequence complementarity between 5′ and 3′ termini of ttsgR that could favor a double-stranded stem of eight GC-rich base pairs, with a 3-nt bulge (AAU) (Fig. 5a). We hence investigated whether this potentially stabilizing stem was critical for ttsgR accumulation, by first mutating the three basal base pairs (Fig. 5a). All three mutants were based on ttsgR_S and, hence, designated S_m5p, S_m3p, and S_m5p3p (Fig. 5a). Interestingly, changing GGG at the 5′ end of ttsgR to CCA (S_m5p) abolished ttsgR while appearing to leave the longer ttsgR* unchanged (Fig. 5b, lane 9 and 10). Sequence analysis verified that most of the clones harbored initiation sites at or near the upstream GAG, while none of them initiated at the original ttsgR initiation site (Fig. 5d). Therefore, the GGG or GAG motif at the 5′ end appeared to be important for ttsgR and ttsgR* accumulation.

Intriguingly, 3 additional clones derived from the S_m5p mutant had their initiation sites at or near nucleotide position 3802, which was 30 nt downstream of the authentic ttsgR initiation site (Fig. 5d, near bottom). Furthermore, the original sequence at this location, namely, CGGUAAUA, became CGGAAUA, GGGAAUA, or GGGUAAUA, in the three clones. Therefore, while a sequence signature with two Gs followed by four Us or As could also be used as an initiation site, the resulting replicating lncRNAs appeared to favor a GGG 5′ end. Together, three classes of ttsgR_S and S_m5p replication products were represented by multiple clones, and all of them possessed characteristic 5′ ends that could be summarized as $G_3(A/U)_4$, although the G_3 portion could tolerate some variations (e.g., GAG or GG). Importantly, sequences with the $G_3(A/U)_4$ consensus have been found at the 5′ termini of gRNA and sgRNAs of many viruses belonging to the genera of

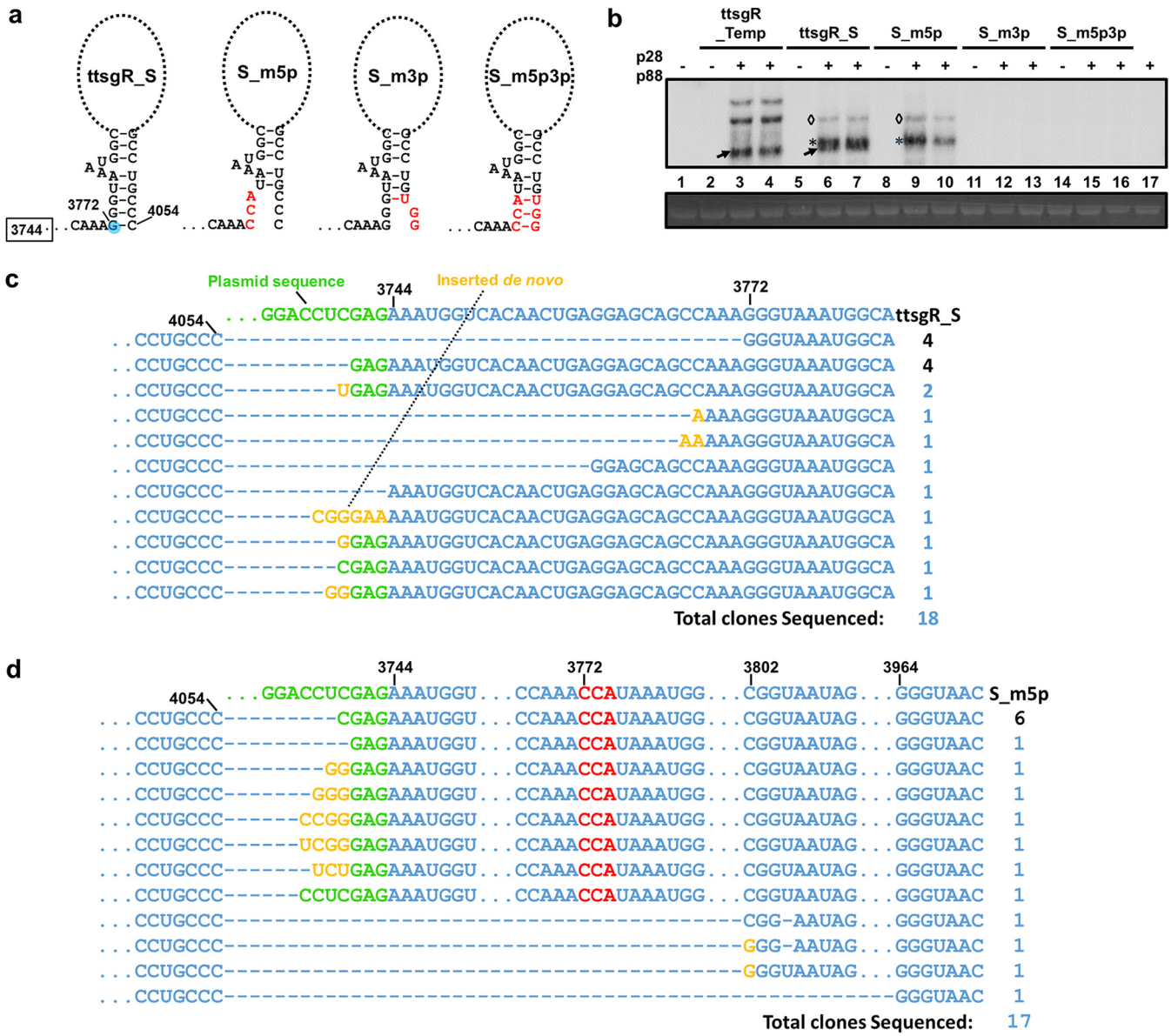


FIG 5 ttsgR_S, a template extending ttsgR by 28 nt at the 5′ end, permits efficient ttsgR replication by p28 and p88 provided in *trans* and facilitates the interrogation of potential 5′ to 3′ base-pair interaction. (a) Schematic representation of the ttsgR_S template, and three mutants of it, with mutated nucleotides in red. The putative terminal stems formed by eight predicted base pairs, and its disruption by mutations, are also depicted. (b) Northern blotting detection of ttsgR from five different templates. Arrows, ttsgR; stars, ttsgR*; diamonds, ttsgR/ttsgR* dimers. (c) The terminal nucleotide positions and identities of ttsgR variants arisen from the ttsgR_S template. (d) The terminal nucleotide positions and identities of ttsgR variants arisen from the S_m5p template. Data in c and d were obtained by sequencing the cloned products of RT-PCR primed with divergent (back-to-back) primers.

Alpha-, *Beta-*, and *Gammacarmovirus*, as well as other closely related viruses, and were collectively designated carmovirus consensus sequence (CCS) (47–49).

Perhaps unsurprisingly, changing CCC to UGG at the 3′ terminus (mutants S_m3p and S_m5p3p) completely abolished the accumulation of both RNA species (Fig. 5b, lanes 11 to 17). This result indicated that the identity of the CCC triplet was essential for the accumulation of both ttsgR and ttsgR*. The 3′ CCC triplet was previously found to be essential for the replication of TCV gRNA as well as a TCV-associated satellite RNA (50). Therefore, the fact that the same CCC motif was essential for ttsgR accumulation provided further corroboration for a replicational origin of ttsgR.

A ttsgR template with an authentic 5′ terminus directs efficient ttsgR synthesis. To further assess the involvement of any 5′-end RNA structure and stalled exonuclease degradation, in ttsgR accumulation, we then produced a third ttsgR template, ttsgR_S2,

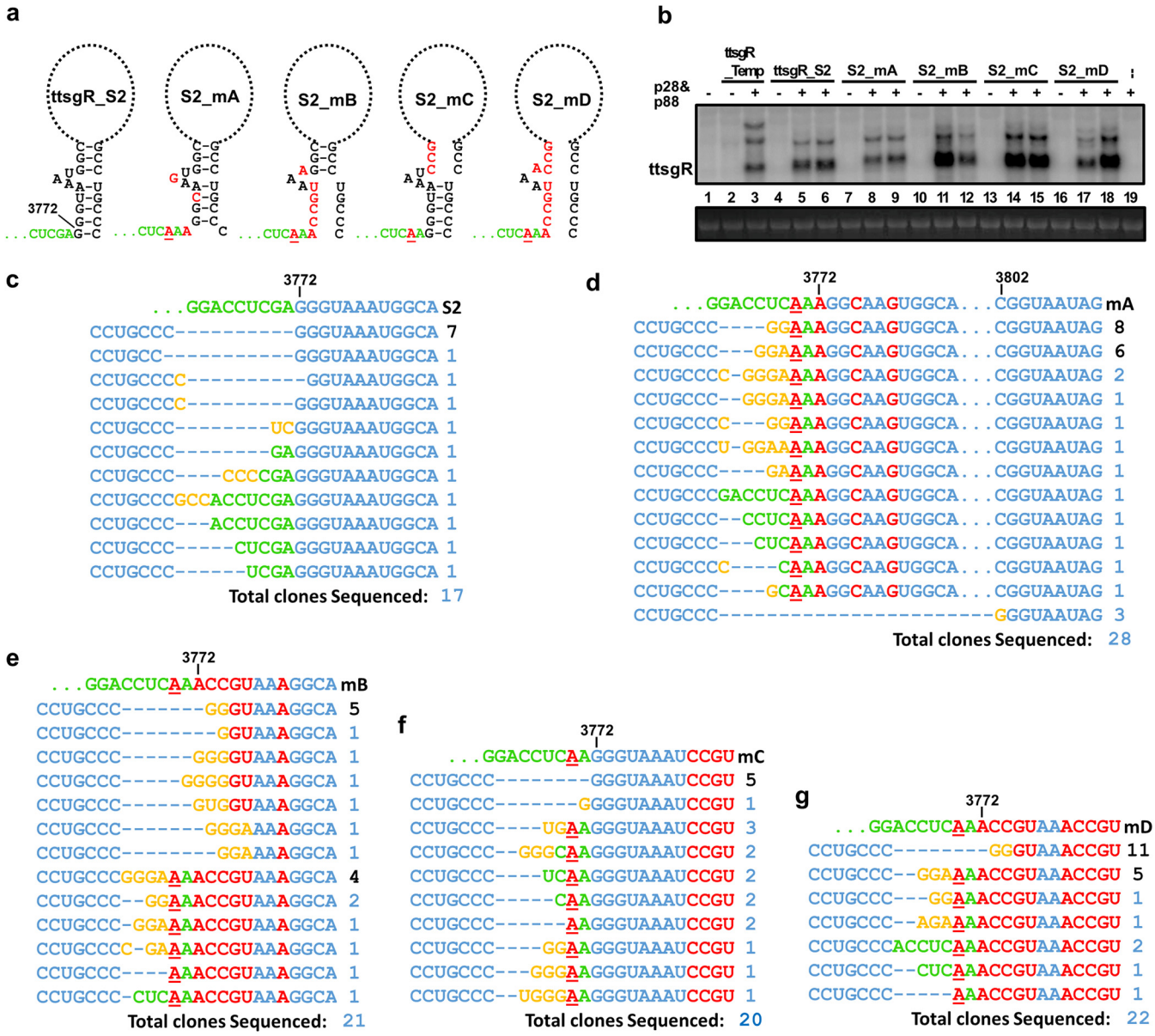


FIG 6 Interrogation of ttsgR_S2, a template with the exact ttsgR 5′ terminus, reveals $G_3(A/U)_4$ as a conserved 5′-end motif needed for replication of ttsgR variants. (a) Schematic representation of the ttsgR_S2 template, and four mutants of it, with mutated nucleotides in red. The putative terminal stems formed by eight predicted base pairs, and its disruption by mutations, are also depicted. (b) Northern blotting detection of ttsgR from different templates. (c to g) Terminal nucleotide positions and identities of ttsgR variants arisen from the ttsgR_S2 and S2_mA to S2_mD templates. The sequence data were obtained by sequencing the cloned products of RT-PCR primed with divergent (back-to-back) primers.

that has the same 5′ terminus as ttsgR itself (Fig. 6a, left). Note that imprecise transcriptional initiation from the CaMV 35S promoter could generate template transcripts that contain up to 10 vector-derived nucleotides (green letters in Fig. 6a). As shown in Fig. 6b (lanes 4 to 6), ttsgR_S2 permitted p28/p88-dependent accumulation of ttsgR. We further verified the identity of the accumulated lncRNAs by analyzing the sequences of RT-PCR products and found the expected ttsgR termini in 7 of the 17 clones. The remaining 10 clones, each representing one sequence, all contained slight variations at the 3′-to-5′ junction. Note that some of them might represent variants that replicated exclusively in dimeric forms. Together, these results demonstrated that a ttsgR template without any extra upstream TCV sequence or structure readily directed its own replication, and the replicated RNA is sufficiently stable to allow for consistent detection.

A 5′ CCS is critically important for ttsgR accumulation. Since the ttsgR_S2 template drove faithful amplification of ttsgR, it was next adopted as a template for more extensive interrogations of potential terminal motifs that route template RNAs to replication complexes containing p28 and p88. The ttsgR_S2 template also had the advantage of minimizing initiation at the alternate GAG triplet (Fig. 6c), permitting more definitive delineation of the ttsgR 5′ end. To determine whether the CCS motif at the ttsgR 5′ terminus tolerated certain variations, we generated four mutants in the sgR_S2 background, named S2_mA, S2_mB, S2_mC, and S2_mD (Fig. 6a). S2_mA contained three single-nucleotide mutations in the CCS motif, changing it from GGGUAAU to AGGCAAG. It also contained a G-to-A mutation in the nearby vector sequence, making it less likely to serve as an alternate initiation site (Fig. 6a). Surprisingly, this mutant templated the synthesis of ttsgR variants that migrated very slightly slower than ttsgR. Among the 28 cloned sequences analyzed, most (25) retained the 4 mutations introduced (red-colored letters), but none of them initiated at the mutated ttsgR initiation site (Fig. 6d). Instead, a majority of them acquired novel initiation sites by incorporating nontemplate, nonvector nucleotides that ranged from one to three Gs and up to two As (Fig. 6d, orange-colored letters). Together with the existing sequence (AAA) at the vector-insert junction, these extra nucleotides rebuilt the 5′ ends that varied from GGAAA (8 clones), GGAAAA (6 clones), to GGGAAAA (2 clones). Therefore, the mutated site had to be altered *de novo* in order to serve as the initiation site for ttsgR replication, with the resulting sites comprising at least two Gs, followed by at least three A/Us. Importantly, ttsgR variants with the initiation site at nucleotide position 3802 (following the numbering of full-length TCV genome) were again recovered in three independent clones, with the first nucleotide changed to G in all three clones (Fig. 6d, bottom). Combined with similar results in Fig. 5d, these findings provided additional evidence for the importance of the CCS motif.

This necessity of CCS was further reinforced by the next three mutants. The S2_mB mutant was originally designed to cause more extensive disruption to the putative long-distance terminal interaction shown in Fig. 5. It nevertheless also changed GGGUAAA to ACCGUAA (Fig. 5A, middle diagram). However, there were still four U/As following the sole G residue. This mutant replicated robustly in *N. benthamiana* cells in the presence of p28 and p88 (Fig. 6b, lanes 10 to 12). Among the 21 cDNA sequences analyzed, approximately one-half (11 clones; Fig. 6e, rows 2 to 8) initiated at the single G at position 3775, but most of them also acquired additional Gs, or underwent other small modifications (e.g., U to A at position 3776, rows 7 to 8), to restore the CCS motif. The other half of the sequences (10 clones), like those of S2_mA, adopted the upstream AAA as the initiation site but again acquired a varied number of Gs or Gs plus A to reconstitute CCS. Notably, in these 10 sequences, all of the original mutations were retained, suggesting that substantial 5′ to 3′ base pairing, with the possible exception of three terminal G:C pairs, was not essential for ttsgR replication.

Many of the descendants of the mutant S2_mC appeared to deviate from the above generalization. Note that this mutant did not change the original GGGUAAA initiation motif, and it supported the accumulation of high levels of ttsgR (Fig. 6b, lanes 13 to 15). All of the 20 clones sequenced retained the original mutations downstream of GGGUAAA, indicating that the identity of these nucleotides (and their potential base-pairing with the 3′ end as depicted in Fig. 6a) were nonessential for ttsgR synthesis. RT-PCR products with the original ttsgR 5′ terminus accounted for 5 of the 20 clones (Fig. 6f). The remaining 15 cDNAs contained insertions of up to 6 nt between the 3′ CCC and 5′ GGG, with many of them being nontemplate nucleotides. However, since the S2_mC mutant was a highly potent ttsgR template, producing substantially higher levels of ttsgR than even the ttsgR_S2 template, it is possible that those with wild-type 5′ ends accumulated predominantly in the monomeric form and, hence, were underrepresented in the pool of dimers we profiled. Conversely, those with insertions between CCC and GGG might have replicated only in the dimeric form.

Finally, the mutant S2_mD, which combined all mutations in S2_mB and S2_mC, also accumulated abundant ttsgR (Fig. 6b, lanes 16 to 18). Among the 22 clones

sequenced, 11 contained cDNA sequences that initiated at the single G at nucleotide position 3775 that preceded UAAA, but with 2 Gs added *de novo* (Fig. 6g). The other 11 sequences initiated from various upstream sites, with or without extra G/A added *de novo*. However, all of them contained at least three As in the vicinity of starting sites. Overall, the data obtained with this set of mutants provided strong evidence for a conserved 5′-end CCS (48, 49) and demonstrated that a CCS alone could initiate the transcription/replication of a TCV lncRNA.

The CCS motif is necessary and sufficient for initiating lncRNA synthesis from internal positions of TCV gRNA. Results in previous sections demonstrated that a CCS is needed for p28/p88 to initiate the replication of ttsgR (and its variants) from templates that were extended at the 5′ end by either 386 nt (ttsgR_Temp), 28 nt (ttsgR_S), or 0 nt (ttsgR_S2) of TCV-derived sequences. Nevertheless, when the CCS motif in these templates was abolished through mutagenesis, similar motifs were relatively easily reconstituted in plant cells with spontaneous mutations or addition of nontemplate nucleotides. However, it was not known whether a similar reconstitution could occur in full-length TCV replicons, e.g., TCV_sg2R or wt TCV. To resolve this question, we next tried to mutate the ttsgR-initiating CCS in both TCV_sg2R and wt TCV while maintaining the amino acid identity of CP. The first mutant m1 altered 3 nts within the motif, changing GGGUAAA to AGGCAAG (Fig. 7a). In both TCV_sg2R and wt TCV, these mutations caused ttsgR to disappear even in the presence of overexpressed p88 (Fig. 7b, compare lanes 7, 8 with 3, 4; and 15, 16 with 11, 12) (note also that lanes 9 and 10 showed very low levels of ttsgR in infections of wt TCV alone). Although a new, low abundance lncRNA was produced by the TCV_sg2R-m1 mutant (Fig. 7b, lanes 7 and 8), it was slightly larger than ttsgR, probably initiating at the nucleotide position 3740, from a GUGAAAA motif (see later). On the other hand, wt TCV produced another slower migrating lncRNA whose level was not affected by the m1 mutations (Fig. 7b, arrowheads). This lncRNA could have originated from GGGUUUA at nucleotide position 3290, within the TCV CP coding region replaced by that of mCherry in TCV_sg2R (see Discussion).

Next, we mutated each of the three m1 mutations separately, creating mutants m1a, m1b, and m1c (Fig. 7a). As shown in Fig. 7c, while mutating the first (m1a; lane 8) or the last (m1c; lane 12) nucleotide of the 7-nt motif had a negligible effect on ttsgR accumulation, the U-to-C change immediately next to GGG (m1b) alone accounted for most of the loss caused by m1 in both TCV_sg2R and wt TCV (Fig. 7c and d, lanes 6 and 10). Therefore, the nucleotide immediately following GGG (or GG) must be A or U in order to permit ttsgR initiation. Interestingly, while m1c itself had little effect on ttsgR abundance, another mutant obtained accidentally while producing m1c, referred as m1c-ii, caused a near-complete loss of ttsgR in TCV_sg2R (Fig. 7c, lane 16). Since m1c-ii differed from m1c by just 1 nt, namely, a deletion of A (Fig. 7a), this result suggested that a minimum of 3 U/A must follow GGG (or GG) in order to facilitate efficient ttsgR biogenesis. These results were also consistent with the sequencing results described in Fig. 5 and 6.

We then asked whether creation of a CCS motif at a different location of the TCV (and TCV_sg2R) genome was sufficient to cause the production of a new lncRNA. We hence mutated 4 nts between 3401 and 3407 of TCV, leading to a new CCS with the sequence of GGGUUAA (Fig. 7a). This set of m2 mutations were introduced into both TCV_sg2R and wt TCV, alone or with m1 mutations (Fig. 7a). As shown in Fig. 7e, the m2 mutations indeed led to the appearance a new lncRNA but also simultaneously depressed the accumulation of ttsgR (lanes 6 and 7; 14 and 15). As expected, combining the m1 and m2 mutations led to a more complete loss of ttsgR, with little impact on the new, m2-dependent lncRNA.

We also verified the identity of the new lncRNA by analyzing the progeny sequences obtained from *N. benthamiana* cells infected with the m1m2 mutant (in TCV_sg2R background). More than one-half of the clones (14 out of 23; data not shown) contained lncRNA sequences that initiated at the new CCS, whereas none of them initiated from the original site where the CCS was abolished. Interestingly, we again recovered sequences that initiated at nucleotide position 3802 (7 out of 23), verifying the frequent usage of this

a

mutant	3401-3410	3771-3781
wt	GGA GCC AAU G	AG GGU AAA UGG
m1		AA GGC AAG UGG
m1a		AA GGU AAA UGG
m1b		AG GGC AAA UGG
m1c		AG GGU AAG UGG
m1c-i		-G GGU AAG UGG
m1c-ii		AG GGU A-G UGG
m2	GGG UUA AAU G	
m1m2	GGG UUA AAU G	AA GGC AAG UGG

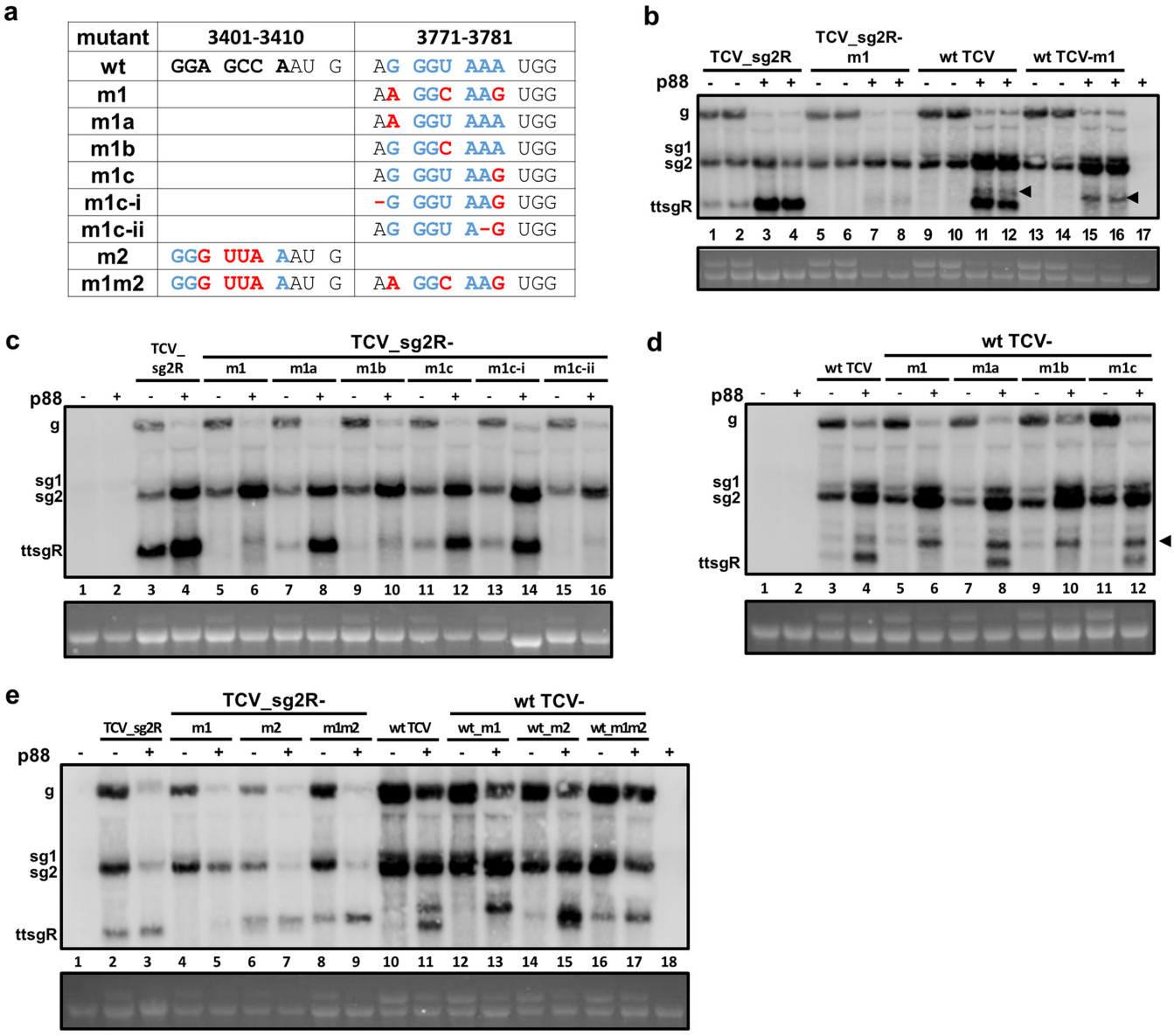


FIG 7 The $G_3(A/U)_4$ motif is necessary and sufficient to initiate the synthesis of ttsgR and similar TCV lncRNAs. (a) wt TCV sequences at nucleotide positions 3401 to 3410 and 3771 to 3781 and sequences of various mutants that altered one or both of these regions. The sequences are separated by spaces into codons in CP ORF. The wild-type GGGUAAA motif responsible for ttsgR initiation is highlighted blue. All mutated nucleotides are highlighted in red. (b, c, d, and e) Northern blots showing the impact of various mutations on the levels of ttsgR and related lncRNAs initiated from replicons. (b) TCV_sg2R-m1 and wt TCV-m1, both containing the same three single-nucleotide mutations within the GGGUAAA motif, suffered a near-complete loss of ttsgR, with or without p88. The two arrowheads denote an additional lncRNA present in wt TCV lanes, which could have initiated from a $G_3(A/U)_4$ motif within the TCV CP region that was by mCherry (nucleotide positions 3290 to 3296, GGGUUUA) in TCV_sg2R. (c) The impact of several single-nucleotide mutations (m1a, m1b, m1c, m1c-i, and m1c-ii) on the level of ttsgR in the TCV_sg2R background. (d) The impact of m1a, m1b, and m1c mutations on the level of ttsgR in the wt TCV background. Note the arrowhead denoting the position of the additional lncRNA unique to wt TCV. (e) Creation of a new $G_3(A/U)_4$ motif (m2, GGGUUAA at nucleotide position 3401 to 3407) was sufficient to program the production of a new lncRNA, in both TCV_sg2R and wt TCV backgrounds.

GGUAAU site as an initiation point for lncRNA. Similarly interesting were 2 sequences that initiated from nucleotide position 3740, although in the resulting lncRNA, the GUGAAAA motif was changed into GGGAAAA. Together these results demonstrated that a CCS motif was sufficient to launch the synthesis of a new lncRNA (or sgRNA) from an internal position of TCV gRNA, whose accumulation was bolstered by p88 overexpression.

The m1 mutant competes poorly with wt TCV in *Arabidopsis*. Finally, we wondered whether production of ttsgR was critical to efficient systemic infection of TCV in host plants. We first used the m1 mutant, in the wt TCV background, to infect *N. benthamiana* plants and closely monitored the infected plants for 6 weeks. Compared

TABLE 1 Fate of the m1 mutant in *N. benthamiana* and *Arabidopsis*

	No. of TCV clones					
			Arabidopsis			
	N. benthamiana		Wild-type Col-0		Mutant *dcl2 dcl4*	
Infected with:	wt	m1	wt	m1	wt	m1
wt TCV	4/4	0/4	3/3	0/3	5/5	0/5
m1	0/5	5/5	0/4	4/4	0/4	4/4
wt TCV + m1	9/26	17/26	18/18	0/18	18/19	1/19

with plants infected with the wt TCV control, these plants did not show any significant differences in the timing and severity of symptoms. The titers of viral RNA as determined using Northern blotting were very similar as well (not shown). Furthermore, the m1 mutations remained stable for at least four sequential passages in *N. benthamiana* and continued to cause symptoms indistinguishable from wt TCV. We also carried out competition experiments by infecting *N. benthamiana* plants with a mixture of m1 and wt TCV. Both variants were recovered at similar frequencies in the systemically infected leaves of six mix-infected plants (Table 1). Indeed, among a total of 26 clones sequenced, 17 were m1 and 9 were wt TCV, suggesting that the m1 mutant was not competitively disadvantaged in *N. benthamiana*.

Surprisingly, starkly different results were obtained when similar competition experiments were carried out in *Arabidopsis* plants, in both wild-type Col-0 and the double-knockout mutant *dcl2 dcl4* defective in antiviral RNA silencing (34, 51). Col-0 and *dcl2 dcl4* plants inoculated with m1 alone developed systemic symptoms indistinguishable from those inoculated with wt TCV or (m1 + wt TCV) mix (not shown). To determine the stability and competitiveness of the m1 mutant, systemically infected leaves were collected at 3 weeks postinoculation, from four different plants per inoculation group. Those from the same inoculation group were pooled and subjected to RNA extraction and RT-PCR. Individual plasmid clones containing RT-PCR fragments were then sequenced. While the m1 mutations remained stable in plants inoculated with m1 alone (Table 1), they all but disappeared in plants mix-inoculated with (m1 + wt TCV) (Table 1). Furthermore, RNA silencing did not appear to play a major role because, while all clones recovered from Col-0 plants contained the wt TCV sequence (Table 1) (18/18), those recovered from *dcl2 dcl4* were also mostly wt TCV (Table 1) (18/19). These results strongly suggest that ttsgR affords a host-specific competitive advantage to wt TCV in *Arabidopsis* but not in *N. benthamiana*.

DISCUSSION

Replication-based production of ttsgR. Viral lncRNAs have been reported for numerous (+) RNA viruses. They have been well investigated for a number of human-infecting flaviviruses, including DENV, West Nile virus (WNV), and Zika virus, which are collectively known as long noncoding subgenomic flavivirus RNAs (sfRNAs) (16, 28, 52). Viral lncRNAs have also been widely observed in (+) RNA plant virus infections, including BNYVV, RCNMV, and TNV-D (17, 19, 22). Most of these viral lncRNAs share the following characteristics: (i) they are typically less than 500 nt in length and often heterogenous at the 5′ end, (ii) they are derived from the 3′ termini of parental viruses, (iii) they are produced by exonuclease degradation of longer viral RNAs that stalls at various stable secondary structures, and (iv) they mostly function to benefit the parental viruses.

ttsgR shares most of these characteristics but is notably different in one key aspect, as follows: unlike the previously examined viral lncRNAs, ttsgR is produced by the viral replication process. This conclusion is well supported by several lines of evidence. First, ttsgR was readily produced from a template that was the same size as the final product, in the absence of any other longer viral RNA, but in the presence of viral replication proteins p28 and p88. Second, p28 and/or p88-mediated template stabilization

was unlikely the cause for ttsgR accumulation because replacing p88 with the RdRp-null p88mGDD completely abolished it. Third, both (+) and (−) strands of ttsgR were detectable, suggesting that the (+) template was copied into (−) intermediates through replication. Fourth, dimeric forms of ttsgR were readily detectable, which could only have been produced through inaccurate copying, possibly on noncovalently circularized templates (53, 54). Finally, production of ttsgR depended on the presence of the CCC motif at the 3′ end and a carmovirus consensus sequence (CCS) at the 5′ end. The CCC motif was previously established by others as essential for the replication of TCV gRNA and a satellite RNA associated with TCV (satC) (50, 55). Similarly, CCS was also found to be required for the (−)-to-(+) strand replication of satC and widely conserved among carmovirus genomic and subgenomic RNAs (47–49). Together, these findings implicate the TCV replication machinery in the biogenesis of ttsgR and unveil the first replicating viral lncRNA.

While replication-based production has not been reported for other viral lncRNAs, this production mode should not be surprising given the functional conservation of viral lncRNAs. Namely, if these lncRNAs play important proviral roles during infections, the mechanisms of their biogenesis can be expected to evolve independently, converging on similar end products. Another important point we wish to make is that none of the previously reported viral lncRNAs have been examined using an *in vivo* replication system that is independent of the replicating parental virus. It may not be unthinkable that some of those earlier lncRNAs could be produced/stabilized by a combination of replication and stalled degradation. Indeed, a recent report found that the (−) strand form of RCNMV lncRNA was detectable in infected tissues, suggesting that this lncRNA may have resulted from such a combination (56). Separately, one of the lncRNA species derived from Japanese encephalitis virus (JEV) was found to act as a potent template for JEV RdRp *in vitro* (57). Finally, a very recent study unveiled a striking example of a coding sgRNA produced by the exonuclease degradation pathway (47), further supporting the complementary roles of the two sgRNA biogenesis modes.

Role of p88 overexpression in ttsgR. It is worth noting that ttsgR accumulation was rather low in infections with wt TCV or TCV replicons that encode their own p28 and p88. However, its levels were dramatically elevated by p88 overexpression. We reported earlier that moderate p88 overexpression depressed the accumulation of TCV gRNA but elevated the accumulation of sgRNAs, especially sgRNA2 (38). The fact that both sgRNA2 and ttsgR were preferentially amplified by p88 overexpression hints at a regulative valve controlled by p88 protein levels. It is possible that higher p88 levels, which could occur at a late stage of TCV cellular infections, serve as a switch that downregulates gRNA replication in favor of sgRNA and ttsgR replication (transcription). How can this be achieved? We speculate that higher p88 levels may lower the p28/p88 ratio, favoring the formation of smaller membrane-encircled spherules in which RNA replication occurs (58). It is further possible that such smaller spherules act as constraints to limit the size of the RNA templates they enclose, leading to preferential amplification of smaller RNA species.

Possible roles of additional internal CCS motifs. Other CCS motifs present in the TCV gRNA internally, especially those within its 3′ half, also drove potent initiation of viral lncRNAs that overaccumulated in the presence of excess RdRp. In addition to the GGGUAAA motif at nucleotide position 3772 that is responsible for ttsgR accumulation, similar motifs at nucleotide position 3740 (GUGAAAA) and 3803 (GGUAAU) likewise programed the synthesis of lncRNAs, albeit at a lower abundance. A fourth motif at nucleotide position 3290 (GGGUUUA) likely caused the initiation of another yet-to-be characterized lncRNA seen in multiple Northern blots. Finally, creating a GGGUUAA motif at a new location (nucleotide position 3401) in TCV gRNA was enough to program the production of a novel lncRNA. Together, these observations reinforce the critical role of CCS in the replication of diverse TCV RNAs, including gRNA and both sgRNAs.

Unlike sgRNAs, replication of ttsgR and similar lncRNAs did not require strong secondary structures prior to their initiation sites (42, 43). Those strong RNA secondary structures, known as attenuation structures, cause premature termination of (−) strand

TABLE 2 CCS near the 3′ ends of other betacarmovirus genomes

Virus	GenBank accession no.	CCS sequence(s)	Location in genome (nt position)	Relative to CP stop codon	Predicted lncRNA size (nt)
CCFV	L16015.1	UGGUAUA	3677–3683	90 nt upstream	365
JINRV	JQ807988	UGGUAUA; GGGUUAU	3768–3774; 3856–3862	43 nt; 131 nt downstream	267; 179
HCRSV	KC876666.1	AGGAAUU	3616–3622	8 nt upstream	293

synthesis at uniform locations, thus ensuring precise 3′ termini for (−) strand sgRNAs, which coincide with CCS at the 5′ end of (+) strands. However, spontaneous premature termination of (−) strand synthesis could occur at random positions with low frequencies. With the presence of overexpressed p88 and the smaller replication-enabling spherules (discussed above), some of the spontaneously terminated (−) strands could become protected from degradation and serve as the templates for lncRNA synthesis. Importantly, the presence of CCS at the 5′ termini of all successful lncRNAs suggests that this motif is a key recognition signal of p88 RdRp.

Potential conservation of ttsgR among betacarmoviruses. Functional characterization revealed a host-specific, and RNA silencing-independent, competitive advantage afforded by having ttsgR. This functional relevance implies that the capacity to encode ttsgR through a CCS near the 3′ end of the genome might be conserved among different TCV isolates and possibly even in different betacarmovirus species. Consistent with this idea, we found that all TCV isolates deposited in GenBank possess a CCS at the equivalent positions, although the specific sequences differ slightly and include GGGUAAA in the U.S. isolate (GenBank accession no. M22445.2), versus GGGAAAA in an UK isolate (AY312063.1), and GGGAAAU in MK301398.1. Furthermore, as summarized in Table 2, the three other species in the genus *Betacarmovirus*, namely, cardamine chlorotic fleck virus (CCFV), hibiscus chlorotic ringspot virus (HCRSV), and Japanese iris necrotic ring virus (JINRV), all contain CCS motifs near the 3′ ends of their genomes that are predicted to direct the synthesis of lncRNAs of various lengths, although their specific sequences diverged substantially from that of TCV ttsgR. In short, production of one or more 3′ coterminal lncRNAs is probably a shared feature of betacarmoviruses.

In summary, this study presents a thorough characterization of ttsgR, a TCV-encoded lncRNA whose biogenesis depended on the replication machinery of the virus. This biogenesis mode differs from the stalled exonuclease degradation responsible for most of the previously reported viral lncRNAs, revealing a novel route of viral lncRNA generation. Functionally, ttsgR appears to afford a host-specific competitive advantage to wt TCV that is unrelated to RNA-silencing defense, although its specific functional mechanism awaits further characterization. Further dissection of this and other TCV and betacarmovirus lncRNAs is expected to clarify their functional significance and offer new targets for virus control and management.

MATERIALS AND METHODS

Constructs. Binary constructs that express TCV p28 and TBSV p19 from the strong 2X35S promoter were generated earlier (35, 39). Those harboring TCV replicons, including TCV_sg2R, Δp88_sg2R, and [p28stop]_sg2R, as well as wt TCV, were also described in earlier studies (35, 38, 39). Constructs designed to express various forms of p88, with all of them harboring a C-terminal duplicated HA tag, including 2X35S::p88, Core::p88, Core::p88ΔN36, and Core::p88ΔN127, were generated in a recent study (38). A new p88-expressing construct, Micro::p88, was produced by replacing the Core 35S promoter of the Core::p88 construct with a microspore-specific promoter derived from *Nicotiana tabacum* (41). The p88mGDD construct was produced by replacing the HpaI-SalI fragment in the p88 coding region of the Core::p88 construct with a synthesized DNA fragment (gBlock) in which the sequence encoding the GDD motif (GGAGACGAT) was altered to encode VAA (GTTGCAGCT) (Integrated DNA Technology, Coralville, IA). Sequences of all of the gBlock fragments and oligodeoxyribonucleotides used in this study are available upon request. The replacement was accomplished using Gibson Assembly ligation (New England BioLabs, Ipswich, MA).

Mutant replicon constructs Δsg1, Δsg2, and Δsg1&2 (all of them are derivatives of TCV_sg2R) were similarly created with synthesized DNA fragments (gBlocks) and Gibson Assembly cloning. Specifically, the gBlock for Δsg1 contained 5 single-nucleotide substitutions known to disrupt the stem-loop structure of the sgRNA1 promoter (42) (Fig. 2A). It was used to replace the SalI-BamHI fragment in TCV_sg2R. The gBlock for Δsg2 contained a 12-nt deletion that disrupts the stem-loop structure of the sgRNA2

promoter (42, 43) (Fig. 2A). It was used to replace the BamHI-AatII fragment in TCV_sg2R. The Δsg1 and Δsg2 gBlocks were combined in a three-fragment Gibson Assembly to replace the SalI-AatII fragment of TCV_sg2R, creating the Δsg1&2 mutant.

The m1, m1a, m1b, m1c, m2, and m1m2 mutants, in both TCV_sg2R and wt TCV backbones, were created by using PCR-generated overlapping DNA fragments that were subsequently incorporated into TCV_sg2R or wt TCV digested with AatII (located at nucleotide position 2600 of TCV cDNA) plus SpeI (nucleotide position 3950 of TCV cDNA) using multifragment Gibson Assembly. Sequences of the oligodeoxyribonucleotide primers used in these PCRs are available upon request.

The construct serving as the template for sgRNA2 sg2R_Temp was made by removing the first 2,600 nt of the TCV_sg2R insert with XhoI (located at the extreme 5′ end of TCV cDNA) plus AatII digestions, followed by refilling the gap with a gBlock fragment spanning nucleotide positions 2468 to 2600, and thus restoring the stem-loop structure of the sgRNA2 promoter. The template constructs ttsgR_Temp and ttsgR_S were produced by replacing the XhoI-SpeI fragment of TCV_sg2R with gBlock fragments encompassing nucleotide positions 3386 to 3950 and 3744 to 3950 of TCV gRNA, respectively. Constructs ttsgR_mL, ttsgR_mR, and ttsgR_mLR are variants of ttsgR_Temp, each made with gBlock fragments that contained mutations shown in Fig. 2D. Finally, the ttsgR_S2 and ttsgR_S2a templates, as well as mutants thereof, were generated using PCR. The identities of all new constructs were verified with Sanger sequencing.

ACKNOWLEDGMENTS

We are indebted to Anne Simon, as well as one of the anonymous reviewers, for the valuable suggestions concerning additional testing in *Arabidopsis*. We thank K. Andrew White for critically reading the manuscript and making insightful suggestions for revision and correction. We are grateful to the labs of Lucy Stewart, Peg Redinbaugh, and Sally Miller for generous equipment sharing.

This study was supported in part by an NSF award (number 1758912), a SEEDS grant from the Ohio Agricultural Research and Development Center, and Graduate Assistantships from OSU and OARDC to R.S., as well as tuition assistance to S.Z. and R.S. from the Department of Plant Pathology, OSU. F.Q. is also supported by a USDA CRIS project (number OHO01337).

REFERENCES

1. Flint SJ, Racaniello VR, Rall GF, Skalka AM, Enquist LW. 2015. Principles of virology. ASM Press, Washington, DC.
2. Razum O, Sridhar D, Jahn A, Zaidi S, Ooms G, Müller O. 2019. Polio: from eradication to systematic, sustained control. BMJ Glob Health 4:e001633. https://doi.org/10.1136/bmjgh-2019-001633.
3. Martina BEE, Koraka P, Osterhaus ADME. 2009. Dengue virus pathogenesis: an integrated view. Clin Microbiol Rev 22:564–581. https://doi.org/10.1128/CMR.00035-09.
4. Petersen LR, Jamieson DJ, Powers AM, Honein MA. 2016. Zika virus. N Engl J Med 374:1552–1563. https://doi.org/10.1056/NEJMra1602113.
5. Cevik M, Kuppalli K, Kindrachuk J, Peiris M. 2020. Virology, transmission, and pathogenesis of SARS-CoV-2. BMJ 371:m3862. https://doi.org/10.1136/bmj.m3862.
6. Palazzo AF, Koonin EV. 2020. Functional long non-coding RNAs evolve from junk transcripts. Cell 183:1151–1161. https://doi.org/10.1016/j.cell.2020.09.047.
7. Quinn JJ, Chang HY. 2016. Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet 17:47–62. https://doi.org/10.1038/nrg.2015.10.
8. Scholthof HB, Jackson AO. 1997. The enigma of pX: a host-dependentcis-acting element with variable effects on tombusvirus RNA accumulation. Virology 237:56–65. https://doi.org/10.1006/viro.1997.8754.
9. Kelly L, Gerlach WL, Waterhouse PM. 1994. Characterisation of the subgenomic RNAs of an Australian isolate of barley yellow dwarf luteovirus. Virology 202:565–573. https://doi.org/10.1006/viro.1994.1378.
10. 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.
11. Balmori E, Gilmer D, Richards K, Guilley H, Jonard G. 1993. Mapping the promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3. Biochimie 75:517–521. https://doi.org/10.1016/0300-9084(93)90056-x.
12. 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.
13. Urosevic N, van Maanen M, Mansfield JP, Mackenzie JS, Shellam GR. 1997. Molecular characterization of virus-specific RNA produced in the brains of flavivirus-susceptible and -resistant mice after challenge with Murray Valley encephalitis virus. J Gen Virol 78:23–29. https://doi.org/10.1099/0022-1317-78-1-23.
14. Lin K-C, Chang H-L, Chang R-Y. 2004. Accumulation of a 3′-terminal genome fragment in Japanese encephalitis virus-infected mammalian and mosquito cells. J Virol 78:5133–5138. https://doi.org/10.1128/jvi.78.10.5133-5138.2004.
15. Scherbik SV, Paranjape JM, Stockman BM, Silverman RH, Brinton MA. 2006. RNase L plays a role in the antiviral response to West Nile Virus. J Virol 80:2987–2999. https://doi.org/10.1128/JVI.80.6.2987-2999.2006.
16. 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.
17. Iwakawa H-O, 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.
18. Pallarés HM, Costa Navarro GS, Villordo SM, Merwaiss F, de Borba L, Gonzalez Lopez Ledesma MM, Ojeda DS, Henrion-Lacritick A, Morales MA, Fabri C, Saleh MC, Gamarnik AV. 2020. Zika virus subgenomic flavivirus RNA generation requires cooperativity between duplicated RNA structures that are essential for productive infection in human cells. J Virol 94:e00343-20. https://doi.org/10.1128/JVI.00343-20.
19. 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.
20. 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.
21. Flobinus A, Chevigny N, Charley P, 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.
22. 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.

23. 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-FI.
24. MacFadden A, O'Donoghue Z, Silva PAGC, 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.
25. Funk A, Truong K, Nagasaki T, Torres S, Floden N, Balmori Melian E, Edmonds J, Dong H, Shi P-Y, 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.
26. Silva PAGC, Pereira CF, Dalebout TJ, Spaan WJM, Bredenbeek PJ. 2010. An RNA pseudoknot is required for production of yellow fever virus subgenomic RNA by the host nuclease XRN1. J Virol 84:11395–11406. https://doi.org/10.1128/JVI.01047-10.
27. 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.
28. 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.
29. 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.
30. Akiyama BM, Laurence HM, Massey AR, Costantino DA, Xie X, Yang Y, Shi P-Y, 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.
31. Steckelberg A-L, 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.
32. Steckelberg A-L, 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.
33. 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.
34. Cao M, Ye X, Willie K, Lin J, Zhang X, Redinbaugh MG, Simon AE, Morris TJ, Qu F. 2010. The capsid protein of turnip crinkle virus overcomes two separate defense barriers to facilitate systemic movement of the virus in Arabidopsis. J Virol 84:7793–7802. https://doi.org/10.1128/JVI.02643-09.
35. Qu F, Ren T, Morris TJ. 2003. The coat protein of turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. J Virol 77:511–522. https://doi.org/10.1128/jvi.77.1.511-522.2003.
36. Qu F, Ye X, Hou G, Sato S, Clemente TE, Morris TJ. 2005. RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in Nicotiana benthamiana. J Virol 79:15209–15217. https://doi.org/10.1128/JVI.79.24.15209-15217.2005.
37. Qu F, Ye X, Morris TJ. 2008. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proc Natl Acad Sci U S A 105:14732–14737. https://doi.org/10.1073/pnas.0805760105.
38. Zhang S, Sun R, Guo Q, Zhang X-F, Qu F. 2019. Repression of turnip crinkle virus replication by its replication protein p88. Virology 526:165–172. https://doi.org/10.1016/j.virol.2018.10.024.
39. Zhang X-F, Sun R, Guo Q, Zhang S, Meulia T, Halfmann R, Li D, Qu F. 2017. A self-perpetuating repressive state of a viral replication protein blocks superinfection by the same virus. PLoS Pathog 13:e1006253. https://doi.org/10.1371/journal.ppat.1006253.
40. Guo Q, Zhang S, Sun R, Yao X, Zhang X-F, Tatineni S, Meulia T, Qu F. 2020. Superinfection exclusion by p28 of turnip crinkle virus is separable from its replication function. Mol Plant Microbe Interact 33:364–375. https://doi.org/10.1094/MPMI-09-19-0258-R.
41. Oldenhof MT, de Groot PFM, Visser JH, Schrauwen JAM, Wullems GJ. 1996. Isolation and characterization of a microspore-specific gene from tobacco. Plant Mol Biol 31:213–225. https://doi.org/10.1007/BF00021785.
42. Wang J, Simon AE. 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.
43. Wu B, 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.
44. Shivaprasad PV, Akbergenov R, Trinks D, Rajeswaran R, Veluthambi K, Hohn T, Pooggin MM. 2005. Promoters, transcripts, and regulatory proteins of mungbean yellow mosaic geminivirus. J Virol 79:8149–8163. https://doi.org/10.1128/JVI.79.13.8149-8163.2005.
45. Plaskon NE, Adelman ZN, Myles KM. 2009. Accurate strand-specific quantification of viral RNA. 4:e7468. https://doi.org/10.1371/journal.pone.0007468.
46. Qu F, Morris TJ. 1997. Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size. J Virol 71:1428–1435. https://doi.org/10.1128/JVI.71.2.1428-1435.1997.
47. 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.
48. Guan H, Carpenter CD, Simon AE. 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.
49. Guan H, Song C, Simon AE. 1997. RNA promoters located on (-)-strands of a subviral RNA associated with turnip crinkle virus. RNA 3:1401–1412.
50. Nagy PD, Carpenter CD, Simon AE. 1997. A novel 3′-end repair mechanism in an RNA virus. Proc Natl Acad Sci U S A 94:1113–1118. https://doi.org/10.1073/pnas.94.4.1113.
51. Zhang X, Zhang X, Singh J, Li D, Qu F. 2012. Temperature-dependent survival of turnip crinkle virus-infected Arabidopsis plants relies on an RNA silencing-based defense that requires DCL2, AGO2, and HEN1. J Virol 86:6847–6854. https://doi.org/10.1128/JVI.00497-12.
52. Zhang Y, Zhang Y, Liu Z-Y, Cheng M-L, Ma J, Wang Y, Qin C-F, Fang X. 2019. Long non-coding subgenomic flavivirus RNAs have extended 3D structures and are flexible in solution. EMBO Rep 20:e47016. https://doi.org/10.15252/embr.201847016.
53. Herold J, Andino R. 2001. Poliovirus RNA replication requires genome circularization through a protein–protein bridge. Mol Cell 7:581–591. https://doi.org/10.1016/S1097-2765(01)00205-2.
54. Li Z, Nagy PD. 2011. Diverse roles of host RNA binding proteins in RNA virus replication. RNA Biol 8:305–315. https://doi.org/10.4161/rna.8.2.15391.
55. Carpenter CD, Simon AE. 1996. In vivo restoration of biologically active 3′ ends of virus-associated RNAs by nonhomologous RNA recombination and replacement of a terminal motif. J Virol 70:478–486. https://doi.org/10.1128/JVI.70.1.478-486.1996.
56. Newburn LR, White KA. 2020. A trans-activator-like structure in RCNMV RNA1 evokes the origin of the trans-activator in RNA2. PLoS Pathog 16:e1008271. https://doi.org/10.1371/journal.ppat.1008271.
57. Chen Y-S, Fan Y-H, Tien C-F, Yueh A, Chang R-Y. 2018. The conserved stem-loop II structure at the 3′ untranslated region of Japanese encephalitis virus genome is required for the formation of subgenomic flaviviral RNA. 13:e0201250. https://doi.org/10.1371/journal.pone.0201250.
58. Ertel KJ, Benefield D, Castaño-Diez D, Pennington JG, Horswill M, den Boon JA, Otegui MS, Ahlquist P. 2017. Cryo-electron tomography reveals novel features of a viral RNA replication compartment. eLife 6:e25940. https://doi.org/10.7554/eLife.25940.