Helper Virus Independent Transcription and Multimerization of A Satellite RNA Associated with Cucumber Mosaic Virus

Soon Ho Choi#, Jang-Kyun Seo#, Sun-Jung Kwon# and A. L. N. Rao*

Department of Plant Pathology & Microbiology, University of California, Riverside, California 92521-0122

# These authors contributed equally to this work

Running title: Helper virus independency of a satellite RNA of CMV

*Address correspondence to A.L.N. Rao, arao@ucr.edu.

Word count:

Abstract: 161

Text: 5293
SATellite RNAs are the smallest infectious agents whose replication is thought to be completely dependent on their helper virus (HV). Here we report that, when expressed autonomously in the absence of HV, a variant of satellite RNA (Q-satRNA) associated with Cucumber mosaic virus has a propensity to localize in the nucleus and transcribe, generating genomic and anti-genomic multimeric forms. The involvement of the nuclear phase of Q-satRNA was further confirmed by confocal microscopy employing in vivo RNA tagging and dsRNA labeling assays. Sequence analyses revealed that the Q-satRNA multimers formed in the absence of HV, compared to when HV is present, are distinguished by the addition of a template-independent hepta-nucleotide motif (HNM) at the monomer junctions within the multimers. Collectively, the involvement of a nuclear phase in the replication cycle of Q-satRNA not only provides a valid explanation for its persistent survival in the absence of HV but also suggests a possible evolutionary relationship to viroids that replicate in the nucleus.
INTRODUCTION

Subviral pathogens are the smallest known infectious molecules that manipulate the cellular systems of much higher organisms, including animals and plants, to replicate themselves (13, 28). Plant subviral pathogens can be divided into two major groups based on their replicability: Helper virus-dependent or independent subviral pathogens. The former includes plant satellite RNAs (satRNAs) while the latter includes viroids. While viroids autonomously replicate in susceptible cells, satRNAs have been shown to be dependent on their helper viruses (HV) for replication (12, 20, 48). Although satRNAs have no appreciable sequence homology with the HV genome, satRNAs not only utilize HV replicases for their replication but also interfere with HV replication and thereby modify disease-symptom expression on the host plants (28, 47, 49, 51).

satRNAs associated with *Cucumber mosaic virus* (CMV), the type member of the genus *Cucumovirus*, are among the earliest found and well studied subviral pathogens (24, 39, 41). CMV satRNAs have 5’ capped, noncoding single-stranded RNA genomes of 330-405 nucleotides (21, 48). They exhibit a high secondary structure with 50% intramolecular base pairing, sharing little or no sequence homology with their HV (21, 28, 48). During replication by HV, cucumovirus satRNAs have been shown to generate multimers of dimeric and tetrameric forms (32, 33). Some satRNAs (i.e. nepovirus satRNAs and sobemovirus satRNAs) generate multimeric intermediates by a rolling-circle mechanism and produce monomeric progeny by autocatalytic cleavage (19, 22). However, this is unlikely to be applicable to multimeric forms of cucumovirus satRNAs since no circular intermediates have been detected (33, 44). Although it was suggested
that CMV satRNAs can produce dimeric forms by self-ligation of double stranded RNA (dsRNA) monomeric forms (44), the mechanism involved in the production of multimeric forms is not well understood. Interestingly, CMV satRNA has been shown to survive up to 25 days without its HV (29, 38, 39). However, the molecular basis for this abnormal long-term survival of CMV satRNA still remains obscure although their high secondary structure was envisioned to contribute to this HV-independent survival (38, 39).

CMV satRNAs utilize CMV encoded proteins for their replication and encapsidation (41). The genome of CMV consists of three 5'-capped positive-strand RNAs. RNA 1 and 2 encode the non structural proteins, 1a and 2a, respectively, and are the major components of a functional replication complex (41). RNA2 also encodes another non-structural protein 2b, a well-characterized suppressor of RNA silencing (14). RNA3 is dicistronic encoding two proteins: a movement protein (MP) and a coat protein (CP), that are involved in the cell-to-cell and long distance movement of the virus (41). Both MP and CP are dispensable for RNA replication but are required for whole plant infection (41). CMV strains are classified into subgroups I and II (41). A notable feature that distinguishes CMV strains of subgroup II from those of subgroup I, is the presence of an additional RNA species, referred to as RNA5 (6). Molecular characterization of RNA5 revealed that it is a mixture of the 3’ terminal 307 and 304 nt regions of CMV RNAs 2 and 3, respectively (6, 11).

Since the discovery of satRNAs, these petite subviral molecules remain of interest to molecular biologists as relatively simple models for studying interactions between macromolecules in plant cells (28). A number of satRNAs have been shown to modulate the course of disease development in plants incited by their HVs (28), and
hence are of importance to practical agriculture (41, 48). Consequently, a majority of studies have focused on characterizing various strains of satRNAs, their relationship to HVs, symptom expression and origin (16, 27, 28, 41, 47-49, 51). Because of the inherent dependency on HV, most research on satRNA replication to date has been performed in the presence of HV using mechanical inoculation of either virion RNA or in vitro transcripts (47, 49, 51). Therefore, the major purpose of this work is to define in molecular terms the basis for HV-independent survival of satRNA. Consequently, in this study, we sought to examine the subcellular localization and the biological activities of a satRNA in the absence of its HV. Surprisingly, we found that when expressed in the absence of HV, a CMV strain Q (Q-CMV) satRNA (Q-satRNA) localized in the nucleus and was transcribed to generate multimers of genomic and antigenomic strands, a previously unrecognized novel feature that could account for the persistent survival of CMV satRNA in the absence HV, and place the replication cycle of satRNA in a new perspective.

MATERIALS AND METHODS

Construction of the Q-satRNA agrotransformant. For the Agrobacterium tumefaciens-mediated transient expression (agroinfiltration), the Q-SatRNA cDNA was amplified by PCR using a 5’ forward primer (GTTTTGTTTGTTAGAGAATTG) and a 3’ reverse primer (GGGTCTGTGGTAGGAATGATA). The resulting product was treated with T4 polynucleotide kinase (New England Biolabs, USA) and ligated into a binary vector pCassHDV digested with StuI and Ncol followed by mung bean treatment to
create blunt ends (4). The construction and characteristic features of agrotransformants corresponding to three genomic RNAs and RNA5 (Q-RNA5) of Q-CMV are as described previously (11).

**Agroinfiltration, mechanical inoculation and progeny analysis.** Following transformation into *Agrobacterium tumefaciens* strain EHA105, agrotransformants were infiltrated, into either *Nicotiana clevelandii* or *Nicotiana benthamiana* plants as described (4). Depending on the nature of experiments an agrotransformant of p19, a suppressor of RNA silencing was co-infiltrated. A Smal-linearized Q-satRNA cDNA plasmid (15) was used as a template for synthesizing capped transcripts using a T7 *in vitro* transcription system (11). For mechanical inoculation, an inoculum containing CMV virion RNA and/or the Q-satRNA *in vitro* transcripts (~100 μg/ml) was used. The total RNAs were extracted using TRIzol® reagent (Invitrogen, USA) according to manufacturer’s instructions. ³²P-labeled riboprobes used to detected CMV genomic RNAs were prepared as described previously (11). To prepare strand-specific Q-satRNA riboprobes, the Q-satRNA cDNA was amplified by PCR and inserted into pT3T7-lac vector (Stratagene, USA). The resulting clone was referred to as pT7T3-Q-Sat. T3 RNA polymerase transcripts from the *Eco*RI-linearized pT7T3-Q-Sat plasmid was used to specifically detect (+)-strand Q-satRNAs whereas T7 RNA polymerase transcripts from the *Hind*III-linearized pT7T3-Q-Sat was used to specifically detect (-)-strand Q-satRNAs. ³²P-labeled RNA size marker was synthesized using T7 RNA polymerase and millennium marker probe template according to the manufacturer protocol (Ambion, USA).
**In situ detection of dsRNA and confocal microscopy.** Immunofluorescence labeling of dsRNA forms of Q-satRNA in the leaves was performed essentially as described previously (5). For dsRNA detection, leaf sections were incubated with the monoclonal anti-dsRNA antibody J2 (Scicons, Hungary) diluted to 1:200 for 16 hrs in a moist chamber at 4°C. Samples were washed 5 times with PHEM buffer (60 mM PIPES, 25 mM HEPES, 5 mM EGTA, 2 mM MgCl$_2$, pH 6.9) for 5 min/wash, and then incubated with the Alexa Fluor 633 goat anti-mouse antibody (Invitrogen, USA) diluted to 1:100 for 2 hrs at room temperature followed by washing with PHEM (5 times, 5 min/wash). To stain nuclei, the samples were incubated with 300 nM of DAPI in PHEM for 5 minutes and washed with PHEM (5 times, 5 min/wash). The samples were finally mounted on slides using Vectashield™ mounting media (Vector Laboratories, USA). For RNaseIII treatment, samples were treated with 4 units of RNaseIII (Ambion, USA) in the supplied buffer for 2 h at 37°C and washed with PHEM buffer (3 times, 3 min/wash) prior to adding the J2 antibody. Confocal microscopy was performed using Leica TCS SP2 confocal microscopy (Leica, Germany). He/Ne lasers were used to detect Alexa Fluor 633 fluorescence and UV laser was used for detecting DAPI staining.

**RNA tagging assay.** The bacteriophage MS2-coat protein (MS2-CP) RNA-tagging assay is a specialized tool to monitor RNA trafficking in individual cells (26). A 23-nucleotide MS2 CP-binding site was engineered into the stem loop-C (SL-C) of Q-satRNA by the double-joint PCR method using Vent polymerase (New England Biolabs, USA) as described previously (46). Briefly, the 5’ region of Q-satRNA was amplified by PCR using a 5’ forward primer (5’-GTTTGTGTTGTTAGAGAATTGCGTAGAG-3’) and a
3' reverse primer (5'-GACATGGGTGATCCTCATGTTTAGCAGGGAGGTAACCCAC-3', the MS2 binding site sequence is underlined). The 3' region of Q-satRNA was amplified by PCR using a 5' forward primer (5'-AACATGAGGATCACCCATGTCCGGCGGTTGAGAGAC-3'; a sequence encompassing the MS2 binding site is underlined) and a 3' reverse primer (5'-CATGCCATGGGGTCTCGGTAGGAATGAT-3'; *Nco*I site is underlined). These two PCR products were mixed and amplified in a second round of PCR using a 5' forward primer (5'-GTTTTGTTTGTAGAGAATTGCGTAGAG-3') and a 3' reverse primer (5'-CATGCCATGGGGTCTCGGTAGGAATGAT-3'; *Nco*I site is underlined). The resulting PCR product was digested with *Nco*I and inserted into the binary vector pCassHDV digested with *Stu*I and *Nco*I. To generate the Q-satRNA transcripts with authentic 3' end, the resulting plasmid was digested with *Nco*I and treated with mung bean exonuclease prior to ligation. The resulting final plasmid was referred to as pQ-satRNA-MS2. The MS2 binding site was also inserted into 5' region of Q-CMV RNA5 by the same procedure as described above. The 5' region of Q-CMV RNA5 was amplified in PCR reaction using a 5' forward primer (5'-GTCCGAAGACGTTAAACTACACTCTC-3') and a 3' reverse primer (5'-GACATGGGTGATCCTCATGTTCAACACGTTTAGGACTTCAG-3', a sequence encompassing the MS2 binding site is underlined). Similarly, the 3' region of Q-CMV RNA5 was amplified using a 5' forward primer (5'-AACATGAGGATCACCCATGTCTTGCGCGGGAAACGGGT-3'; a sequence encompassing the MS2 binding site is underlined) and a 3' reverse primer (5'-CATGCCATGGGTCTCTTTGAGAAACCT-3', *Nco*I site is underlined). These two PCR
products were mixed and amplified during a second round of PCR using a 5' forward primer (5'-GTCCGAAGACGTTAAACTACACTCTC-3') and a 3' reverse primer (5'-CATGCCATGGTCTCCTTATGGAGAACCT-3'; Ncol site is underlined). The resulting PCR product was digested with Ncol, and subcloned to the binary vector pCassHDV as described above. The resulting final plasmid was referred to as pQ-RNA5-MS2. A region encompassing GFP-NLS-CP sequence (57) was amplified by PCR and ligated into the PZP binary vector (45). The resulting plasmid was referred to as PZP-GFP-NLS-CP. A sequence encompassing GFP and GFP-CP cassettes were also amplified by PCR and inserted into the PZP binary vector and the resulting clones were referred to as PZP-GFP and PZP-GFP-CP, respectively. A sequence encompassing cyan fluorescent protein (CFP)-NLS cassette was amplified by PCR using pSITE-1CA (10) as a template and a 5' forward primer (5'-ATGGTGAGCAAGGGCGAG-3') and a 3' reverse primer (5'-GGACCTTGATTTTCTTTCTTCTTCTTTGAGCCATCTTTGTACAGCTCGTCCATG-3', SpeI site is indicated in bold face and a sequence encompassing nuclear localization signal is underlined). The resulting PCR product was digested with SpeI and ligated into PZP vector digested with Stul and SpeI. The resulting final plasmid was referred to as PZP-CFP-NLS. The presence of engineered sequences was confirmed by sequencing. Confocal microscopy was performed as described above.

**Amplification of the Q-satRNA multimer junction sequences using RT-PCR.**

Following denaturation of 5 μg of total RNA by heating at 99°C for 5 min in 20% dimethyl sulfoxide (DMSO), cDNAs were synthesized using the M-MLV reverse transcriptase
(New England Biolabs, USA) using either primer Rv1 (5'-GTTTTGCTAGCGAACTGAGCGGGGG-3') or primer Rv2 (5'-TGATAAACATCCACGGAGAT-3'). The resulting cDNAs were amplified by PCR using either primer Fw1 (5'-GCGGAATTTCGAAAGAAAC-3') and primer Rv1 or primer Fw2 (5'-CCCTACCAGGACCCG-3') and primer Rv2, respectively. PCR was performed using following conditions: 3 min denaturation at 94°C followed by 40 cycles of 30 sec denaturation at 94°C, 30 sec annealing at 63°C and 30 sec-elongation at 72°C. The resulting PCR products were analyzed by electrophoresis in 2% agarose gel and either directly sequenced or sequenced after ligating into pGEM-T easy vector (Promega, USA).

RESULTS

Helper virus-independent transcription and multimerization of Q-satRNA. The Agrobacterium tumefaciens-mediated transient expression (agroinfiltration) not only facilitates synchronized delivery of multiple DNA-based transgene expression but also results in high-level expression of mRNAs independent of replication (4). Thus, a T-DNA based Q-satRNA construct (pQ-satRNA; Fig. 1A) was engineered to initiate transcription by the Cauliflower mosaic virus 35S promoter precisely at the authentic 5' end of Q-satRNA and terminate with the natural 3'CCC_{OH} through a self-cleaving Hepatitis delta virus (HDV) ribozyme (Fig. 1A).

During (+)-strand RNA virus replication, detection of (-)-strands indicates that the observed (+)-strands is the result of de novo synthesis. Therefore, before examining the
To test the biological activity of pQ-satRNA, *Nicotiana clevelandii* leaves were co-infiltrated with a mixture of transformed agrobacteria containing pQ-satRNA and all three genomic components of the HV [Q-CMV; (11)]. Control infiltrations were performed with the following transformed agrobacteria: (i) HV alone; (ii) empty vector (EV); (iii) pQ-satRNA with p19, a viral suppressor of RNAi from the tombusvirus (50); and (iv) pQ-RNA5 (302 nt sequence of Q-CMV RNA5) with p19. The reasons to include p19 in these assays are as follows: The satRNA is not a target for host RNA-dependent RNA polymerase (RdRp) (52). However, since transcriptional/post-transcriptional gene silencing is an inevitable consequence of agroinfiltration (30), ectopically expressed satRNA can be recognized as overexpressed RNA and targeted by RNAi machineries. Therefore, it is beneficial to include a silencing suppressor such as p19 when examining biological activities of satRNA in the absence of HV. Four days post-infiltration (dpi), total RNA was extracted from the infiltrated leaves and Northern blots were generated in triplicate (Fig. 1D) and hybridized with strand specific riboprobes: the first two blots are hybridized with a riboprobe specific for the (+)-strand and (-)-strand Q-satRNAs (Fig. 1D, top and middle panel, respectively), while the last blot was hybridized with a...
riboprobe specific for detecting the (+)-strands of the HV RNAs (Fig. 1D, bottom panel). As shown in Fig. 1D, the detection of the (+)- and (-)-strand progenies of Q-satRNA in the leaves co-infiltrated with HV and pQ-satRNA demonstrated that ectopically expressed Q-satRNA transcripts from the construct are biologically active.

Surprisingly, we detected both polarities of monomeric (1x) and multimeric (2x, 3x and 4x) forms of the Q-satRNA in the leaves expressing Q-satRNA alone (without HV) (Fig 1D, top and middle panels). Since the Q-satRNA riboprobes exhibited high degree of strand-specificity especially when the concentration of the Q-satRNA was lower (compare the hybridization signals for 25 ng or less by the strand specific riboprobes in Fig. 1 B, C), the detected monomeric and multimeric Q-satRNA (+)- and (-)-strand signals are authentic. No such RNA profile was seen in control plants expressing pQ-RNA5 alone (Fig. 1D, bottom panel). Re-probing of the Q-satRNA blots with HV-specific riboprobes failed to detect an HV RNA profile (Fig. 1D, bottom panel), arguing against contamination as a possible explanation for the Q-satRNA multimer formation.

Agroinfiltration is a DNA-based transient RNA expression system initiated in the nucleus (3, 55). Inherently, the replication of Q-satRNA does not involve DNA intermediates and therefore the nucleus is not a recognized site for the Q-satRNA replication. Thus, to mimic natural RNA-based inoculation and to examine whether the mechanical inoculation of Q-satRNA alone would result in multimer formation and (-)-strand synthesis, various concentrations of the (+)-strand monomeric Q-satRNA in vitro transcripts were mechanically inoculated without HV to the leaves of N. clevelandii followed by Northern blot hybridizations using the strand specific Q-satRNA riboprobes. Consistent with the agroinfiltration assays (Fig. 1D), both strands of the Q-satRNA
multimers (2x, 3x and 4x), that were absent in the Q-satRNA inocula transcripts, were
detected in the leaves mechanically inoculated with the Q-satRNA in vitro transcripts
alone (Fig. 1E). These observations confirmed that Q-satRNA is competent to
synthesize (-)-strands and form multimers in the absence of HV. However, we could not
detect any increase in accumulation levels of Q-satRNA in our time course analyses of
mechanically inoculate plants (data not shown). This suggests the synthesis of Q-
satRNA (-)-strands, in the absence of HV, is the result of de novo transcription but not
replication.

Nuclear localization of Q-satRNA in the absence of HV. It has recently been
shown that Q-satRNA is not a template for the known host encoded RdRps including
the RDR2 (52). Thus, we rationalize that HV-independent multimerization of Q-satRNA
observed in the above experiments (Fig. 1 B, C) might be occurring in the nucleus. To
examine this possibility, we performed a bacteriophage MS2-coat protein (MS2-CP)
RNA-tagging assay that allows visualizing the subcellular location of Q-satRNA in living
cells (26). This technique is based on the specific interaction between RNA molecules
bearing binding sites and RNA-binding proteins (such as MS2-CP) and allows to
visualize subcellular locations of RNA molecules in living cells when co-expressed with
an RNA-binding protein fused with GFP (such as MS2-CP–GFP) (26). Based on the
predicted in vitro secondary structure of Q-satRNA (23), we inserted a 23-nucleotide
MS2-CP-binding site into the stem loop-C (SL-C) of Q-satRNA (Fig. 2A). Similarly, the
MS2-CP binding site was engineered into the 5’ region of CMV RNA5 (Fig. 2B) that was
used as a control. The modified pQ-satRNA carrying the MS2 binding site (pQ-satRNA-
MS2) replicated efficiently when co-infiltrated with HV (Fig. 2C). Sequence analysis of the Q-satRNA-MS2 progenies recovered from systemic leaves confirmed the maintenance of engineered MS2 binding site in the progenies (data not shown) and suggested that the engineered insertion had no detectable affect on the the Q-satRNA replication. To monitor and visualize the subcellular location of Q-satRNA, a series of GFP control plasmids amenable for agroinfiltration were also constructed (Fig. 2D).

*N. benthamiana* leaves infiltrated with the above transformed agrobacteria in desired combinations were subjected to confocal microscopy at 3 dpi to monitor GFP signals. The results are summarized in Fig. 2 (E-I). The leaves infiltrated with control constructs of GFP, GFP-CP and GFP-NLS-CP showed the fluorescent signals in the expected subcellular compartments (Fig. 2E). Identical distribution patterns of GFP signals were observed when Q-satRNA was co-expressed with either GFP or GFP-CP or GFP-NLS-CP (Fig. 2F). Likewise, the fluorescent signals resulting from co-expression of Q-satRNA-MS2 and GFP were confined to the cytoplasm (Fig. 2G; top panel). However, when Q-satRNA-MS2 was co-expressed with GFP-CP or GFP-NLS-CP, most fluorescent signals were detected in the nuclei (Fig. 2G; middle and bottom panels). It is possible that the observed nuclear localization of Q-satRNA in these samples could be due to agroinfiltration in which RNA transcripts are primarily made in the nucleus. Therefore, to verify the nuclear localization of Q-satRNA in the absence of HV, we performed similar RNA-tagging assays using a control construct, pQ-RNA5-MS2 (Fig. 2B). The results are shown in Fig. 2H. In contrast to the GFP signals in the nucleus of the cells co-expressing Q-satRNA-MS2 and GFP-CP (Fig. 2G), the GFP signals resulting from co-expression of Q-RNA5-MS2 and GFP-CP were observed throughout.
the cytoplasm (Fig. 2H, middle panel). In addition, the GFP signals were detected both in the nucleus and cytoplasm when the Q-RNA5-MS2 and GFP-NLS-CP were co-expressed (Fig. 2H, bottom panel). The control infiltration involving the co-expression of Q-satRNA-MS2+GFP-CP+CFP-NLS revealed expected co-localization patterns for GFP and cyan fluorescent protein (CFP) signals (Fig. 2I). Taken together, these results confirm that Q-satRNA has a propensity to localize in the nucleus in the absence of its HV.

In situ detection of dsRNA forms of Q-satRNA in the nucleus. Double-strand RNA (dsRNA) is a generally accepted marker for intermediates during replication of positive-strand RNA viruses and their satRNAs (53). Since in the absence of HV, Q-satRNA accumulated both (+)- and (-)-strand monomeric and multimeric forms (Fig. 1D and E) and localized in the nucleus (Fig. 2G), we sought to examine whether dsRNA forms of Q-satRNA could be detected in the nucleus. Thus, we performed in situ assays for dsRNA detection using a monoclonal anti-dsRNA antibody (J2-Ab) (53). *N. benthamiana* leaves expressing Q-satRNA alone were incubated with J2-Ab and subjected to confocal microscopy at 3 dpi. We included the leaves infected with HV to serve as a positive control in this assay. The following four inocula were included as negative controls: (i) empty vector (EV); (ii) (-)-strand Q-satRNA representing a highly structured single strand RNA; (iii) Q-CMV RNA5 that forms a high secondary structure due to the presence of a 3' tRNA-like structure (TLS) and (iv) *Brome mosaic virus* RNA3 (B3) representing another single strand RNA. The results are summarized in Fig. 3. As expected, in controls and experimental samples, the DAPI treatment specifically stained
the nucleus emitting strong blue fluorescence (Fig. 3A-I). Since HV replication is entirely cytoplasmic (41), J2-Ab specific signals in the leaves expressing HV alone were found scattered in the cytoplasm, but not in the nucleus, as punctate bodies emitting red fluorescence (Fig. 3B). The absence of red fluorescence signals specific for the J2-Ab in the leaves infiltrated with four control inocula (Fig. 3 A, E-G) confirmed J2-Ab did not detect single strand RNA molecules even though they are highly structured. However, the localization of the J2-Ab specific signals in the nucleus was exclusively found in the plant cells expressing Q-satRNA alone (Fig. 3C). To confirm whether the red fluorescent signals detected by the J2-Ab were due to dsRNAs, N. benthamiana leaves infiltrated with respective samples were treated with RNase III (specific for dsRNA). As expected, red fluorescence signals from the J2-Ab in the leaves expressing HV or Q-satRNA alone were completely lost after RNase III treatment (Fig. 3H and I). To further verify that the nuclear localization phase associated with Q-satRNA is conserved in another CMV satRNA, a variant of CMV satRNA isolated from N. gluaca (referred to NG-satRNA) (35) was infiltrated into N. benthamiana leaves and processed for dsRNA detection. Analogous to Q-satRNA, the localization of J2-Ab specific signals in the nucleus was found in the cells expressing NG-satRNA alone (Fig. 3D). Collectively, our data suggest that, in the absence of HV, the nucleus is the preferred site for the synthesis of anti-genomic Q-satRNAs resulting in dsRNA formation.

The Q-satRNA multimers formed in the absence of HV exhibit unique junction sequences. To compare and characterize the junction sequences of the Q-satRNA multimers formed in the presence and absence of HV (Fig. 1D), we used a divergent
PCR approach (2) with two sets of primers (Fig. 4A and B). The resulting PCR products were either directly sequenced (Fig. 4C, Exp.1) or sequenced after subcloning into the pGEM-T vector (Fig. 4C, Exp. 2-4). The sequencing results revealed the formation of two classes of multimers. The Class 1 type was characterized by the presence of a junction sequence lacking a 3’ terminal C-residue of the first monomeric unit of head-to-tail repeats; whereas the Class 2 type was characterized by the addition of non-template GGGAAAA, referred to as the hepta nucleotide motif (HNM), to the 3’-end of the first monomeric unit of head-to-tail repeats (Fig. 4C). When total RNAs recovered from the leaves co-infiltrated with Q-satRNA and HV were subjected to RT-PCR using Fw1 and Rv1 primers, only the accumulation of the Class 1 type multimers was detected (Fig. 4C). Interestingly, similar RT-PCR using total RNA recovered from the leaves infiltrated with Q-satRNA alone revealed exclusive accumulation of the Class 2 type multimers (Fig. 4C). Similar results were obtained from the plants inoculated mechanically with either Q-satRNA and HV or Q-satRNA alone (Fig. 4C). The sequence analysis of the in vitro RNA transcripts used for mechanical inoculation terminated with expected authentic CCCOH suggesting the synthesis of the HNM was occurred de novo by an unknown mechanism (see Discussion). Furthermore, since the Q-satRNA multimers formed in the presence or absence of HV are clearly distinct, the detection of the Class 2 multimers represents a diagnostic feature for the Q-satRNA nuclear localization.

Evidence for nuclear localization of Q-satRNA in the presence of HV. Since Q-satRNA was observed to localize in the nucleus in the absence of HV, a question that
needs to be addressed would be: Does Q-satRNA localize to nucleus in the presence of HV? To find an answer to this question, we used two approaches. The first approach was based on our observations that the formation of the Class 2 multimers (Fig. 4B, C) is diagnostic for the nuclear localization of Q-satRNA. Therefore, we employed a divergent PCR assay to verify whether the Class 2 type multimers are found in the leaves co-expressing HV and Q-satRNA. The Class 2 type multimers were not detected using Fw1 and Rv1 primers in the leaves infiltrated with HV and Q-satRNA (Fig 4C). We reasoned that it could be due to that a majority of the accumulated multimers is the Class 1 and only the low amount of the Class 2 multimers accumulated in the presence of HV. Thus, we designed another primer set (Fw2 and Rv2; Fig. 4B) to amplify the junction region of only the Class 2, but not the Class 1 multimers (Fig. 4B). Using this primer set, we successfully amplified the junction sequences of only the Class 2 multimers from the leaves coinfiltrated with Q-satRNA and HV (Fig. 4B and C).

As a second approach the MS2-CP mediated RNA tagging assay was applied to the leaves co-infiltrated with a series of the transformed agrobacteria including controls (eg. HV+GFP) and test samples (eg. HV+Q-satRNA-MS2+GFP-CP). The results are shown in Fig. 5. As expected, the leaves infiltrated with the control constructs of HV+GFP, HV+GFP-CP or HV+GFP-NLS-CP expressed the fluorescent signals in expected subcellular compartments (i.e. throughout the cytoplasm) (Fig. 5A). Similar distribution patterns of the fluorescent signal were observed when Q-satRNA was co-expressed with HV+GFP, HV+GFP-CP or HV+GFP-NLS-CP (Fig. 5B). Likewise, as expected, the distribution of the fluorescent signals from co-expression of HV+Q-satRNA-MS2+GFP (Fig. 5C, top panel) was identical to a control (HV+GFP; Fig. 5A, top panel). However,
when Q-satRNA-MS2 was co-expressed with HV+GFP-CP or HV+GFP-NLS-CP, although the distribution of the fluorescent signals was detected in the cytoplasm, it also displayed a pattern distinct from that of control samples (Fig. 5C, bottom two panels). Since the Q-satRNA replication in the presence of HV is robust and most progenies accumulate throughout the cytoplasm, it is difficult to conclude whether Q-satRNA localized to the nucleus in the presence of HV. Therefore, we performed an additional assay by agroinfiltrating the test sample (HV+Q-satRNA-MS2+GFP-CP) along with a construct, CFP-NLS (Fig. 2D) that specifically localizes to the nucleus and emits blue fluorescent signals. The results shown in Fig. 5D clearly demonstrate that some fraction of Q-satRNA localized in the nucleus in the presence of HV. In all, our results clearly support that Q-satRNA does involve the nucleus during the replication cycles.

**DISCUSSION**

A most striking observation of this study is that Q-satRNA is able to localize in the nucleus and be transcribed to form multimers in the absence of its HV (Fig. 1 and 2). The application of the RNA tagging assay (Fig. 2) and the in situ detection of dsRNAs using J2-Ab as a probe (Fig. 3) convincingly demonstrated that Q-satRNA has a propensity to localize in the nucleus and some host enzyme activities (eg. polymerase II) might be involved in the transcription of anti-genomic strands and generating multimeric forms, as observed in viroids and HDV (12, 17). The intimacy between the nucleus and Q-satRNA that was not recognized before this study raises a possibility that this subcellular compartment is likely to play a critical role in the Q-satRNA replication cycle.
before entering the HV-dependent replication phase. We rationalize two scenarios for
the intrinsic involvement of the nuclear phase in the Q-satRNA replication cycle. First,
Q-satRNA is known to persistently survive up to 25 days without its HV (39). Thus, as
demonstrated here, nucleus could be the preferred subcellular compartment for satRNA
to protect itself against host defense before encountering a compatible HV. Second, as
discussed below, the nuclear localization might promote a template-independent
synthesis of the HNM to the primary transcripts of Q-satRNA to promote multimer
formation that subsequently function as efficient templates for HV-dependent replication.

Our cell biology based evidence (Figs. 2 and 5) showed that Q-satRNA localizes in
the nucleus but the question that needs to be addressed is: How does Q-satRNA get
transported into the nucleus? For viruses that replicate in the nucleus, the proteins
encoded by their genomes have been shown to contain nuclear localization signals
(54). Despite the lack of signals responsible for nuclear localization, a virus-based in
vivo system has shown that Potato spindle tuber viroid (PSTVd) localizes to the nucleus
(58). Recently, a protein isolated from PSTVd infected tomato plants, referred to as
VIRP1, has been shown to promote nuclear localization of PSTVd (34). In addition, a
previous study using the dissected PSTVd genome revealed that an evolutionarily
conserved sequence and/or the stem-loop structured spanning hairpin I and the center
conserved region are sufficient to direct import of PSTVd to the nucleus (1). Thus, it is
possible that host RNA binding proteins such as VIRP1 could mediate nuclear import of
Q-satRNA. Alternatively, as observed with PSTVd (1), a part of the Q-satRNA sequence
itself might function as a nuclear targeting signal.

One of the intriguing discoveries of our study is the presence of the HNM at the
The HNM is not encoded in the T-DNA construct and the presence of a HDV ribozyme would generate the Q-satRNA transcripts with authentic 3’CCCOH (Fig. 1A). Although T7 RNA polymerase is known to add non-template nucleotides during *in vitro* transcription (36), sequence analyses of the *in vitro* synthesized Q-satRNA transcripts used for mechanical inoculation did not reveal the presence of any non-template nucleotides. Therefore, the HNM found at the junction of the Class 2 type multimers is not a part of Q-satRNA and appears to originate by a different mechanism. In RNA viruses and their linear satRNAs, 3’ terminal sequences are critical for the (-)-strand synthesis (48). Therefore, to prevent their loss due to degradation by subcellular RNases, a rapid turnover and correction of 3’ termini in RNA viruses is mediated by a variety of mechanisms that include virus-encoded RdRps (8, 25), tRNA nucleotidyltransferase (42), recombination (8, 25), terminal transferases (40) and the cytoplasmic polyadenylation apparatus (43). In addition, the synthesis of non-template nucleotides at the 3’ end prior to initiating transcription by a viral RdRp was also observed (25). Alternatively, host encoded RdRps are also possible candidates for the addition of the HNM.

A wide range of satRNAs have been shown to form multimers in the presence of HV either by a rolling circle mechanism, as in circular satRNAs (7, 19, 31) or by a viral RdRp re-initiating prior to releasing the nascent strand as in linear satRNAs (9). Our study revealed that the multimers could also be formed in the absence of HV (Fig. 1D). We hypothesize that a primary step toward multimer formation is the addition of the non-template HNM to the 3’ end of the (+)-strand monomeric forms followed by the
ligation of the 5' end of the second monomeric unit. Recurrences of this process might convert the monomeric units into multimeric forms having the HNM at the junction. The detection of multimeric (-)-strands suggests that the HNM could serve as a potential initiation site for (+)-strand synthesis possibly by a host polymerase in the absence of HV.

In conclusion, our study provides a previously unrecognized nuclear phase associated with Q-satRNA. As shown in this study, HV-independency of Q-satRNA in multimer formation and nuclear localization suggests a possible evolutionary relationship to PSTVd, belonging to the family Pospiviroidae, which replicates in the nucleus by DNA-dependent RNA polymerase II via a rolling circle mechanism (18). Interestingly, it was reported that the replication of PSTVd was severely inhibited when the plants were co-infected with CMV and satRNA, whereas CMV alone had no effect on the replication of PSTVd (37, 56). This may indicate that PSTVd and CMV satRNA are competing for the same host enzyme during replication in the nucleus or for the same host protein involved in the nuclear import of the RNAs (i.e. VIRP 1). Therefore, it is of significant interest to examine which host proteins are involved in the nuclear import of Q-satRNA. This is the subject of our current research.

ACKNOWLEDGEMENTS

We thank Shou-Wei Ding for providing the Q-satRNA and CMV clones, Anne Simon for providing GFP-NLS-CP plasmid, David Carter of the Center for Plant Cell Biology and
Devinka Bumunusinghe for assistance with confocal immunofluorescence microscopy, Deb Mathews for providing a NG-satRNA clone and Bryce Falk and Deb Mathews for critical reading of the manuscript.

REFERENCES


FIGURE LEGENDS

FIG. 1 Multimerization of Q-satRNA in the absence of its HV. (A) Schematic diagram of a T-DNA based Q-satRNA construct (pQ-satRNA) for in vivo expression via agroinfiltration. The basal binary vector contains, in sequential order, a left border of T-DNA (LB), a double 35S promoter (35S), full-length cDNA of Q-satRNA, a ribozyme sequence (Rz) from hepatitis delta virus (HDV), a 35S terminator (Ter) and a right border of T-DNA (RB). The transcription initiation and HDV cleavage sites are shown. (B, C) Specificity of the Q-satRNA riboprobes. Duplicate blots containing indicated amounts of in vitro transcripts of the (+) and (-)-strand Q-satRNA were hybridized with 32P-labelled riboprobes designed to detect either (B) Q-satRNA (+)-strands or (C) Q-satRNA (-)-strands. (D) Northern blot analysis of total RNA recovered at 4 dpi from N. clevelandii leaves agroinfiltrated with the indicated cultures of agroconstructs along with p19 (a suppressor of RNA silencing) and hybridized with the riboprobes as shown. The positions of the Q-satRNA monomeric (1x), dimeric (2x), trimeric (3x) and tetrameric (4x) forms are indicated on the left of each panel. In the bottom panel, RNA1-RNA5 indicates HV progeny RNAs. rRNA represents loading control. All samples contained 20 μg of total RNA except in top and middle panels, in lanes indicated with EV+HV and pQ-satRNA+HV respectively contained 4 μg and 0.4 μg of total RNA. In the bottom panel, these two lanes contained 4 μg of total RNA. In each panel, a radiolabeled RNA size marker is shown on the right. (E) (Top and Middle panels) Northern blot analysis of total RNA recovered from N. clevelandii plants mechanically inoculated with the indicated concentrations of the Q-satRNA in vitro transcripts and probed with the indicated
riboprobes. Bottom panel shows the Northern blot hybridization of the Q-satRNA \textit{in vitro} transcripts used for mechanical inoculation.

**FIG. 2** Nuclear localization of Q-satRNA in the absence of HV. (A) The predicted \textit{in vitro} secondary structure of Q-satRNA and the stem-loop C (SL-C) is shown and a dotted circle indicates the location where MS2 binding site was engineered. (B) Schematic representation of pQ-R5-MS2 showing the location where MS2 binding site (indicated by a dotted circle) was engineered into Q-CMV RNA5. (C) Northern blot showing that the insertion of MS2 binding site into Q-satRNA has no significant affect on the Q-satRNA replication when co-infiltrated with HV. The blot was hybridized with a riboprobe to detect (+)-strand. The positions of the Q-satRNA monomeric (1x) and dimeric (2x) forms are indicated. rRNA represents loading control. (D) Schematic representation of agroconstructs used for ectopic expression of GFP alone or GFP-fused to MS2 CP or GFP-fused to NSL (nuclear localization signal) and MS2 CP. The basal binary PZP vector contains, in sequential order, a left border of T-DNA (LB), a double 35S promoter (35S), multiple cloning sites, a 35S terminator (T35S) and a right border of T-DNA (RB). The bent arrow indicates translation initiation site. The target genes were subcloned in-frame into StuI and SpeI sites. (E-I) Representative confocal microscopic images of \textit{N. benthamiana} leaves agroinfiltrated with either single or pairwise combinations of the indicated agrocultures. Fluorescent signals were observed in the epidermal cells at 3 dpi. In each panel, GFP expression (green), light mode, and merged images are shown. In addition, in panel I, CFP expression (blue) was shown. Bar= 20 μm.
FIG. 3 In situ evidence for the synthesis of the anti-genomic strands of Q-satRNA in the absence of HV. Immunolabeling of dsRNAs in the leaves infiltrated with the following agroconstructs: (A) Empty vector (EV); (B) HV; (C) Q-satRNA (D) NG-satRNA; (E) Q-satRNA (-)-strand; (F) Q-RNA5 and (G) BMV RNA3. To visualize the nuclei, the leaf segments were treated with DAPI (blue). dsRNAs in the leaf segments were probed using J2-Ab and then treated with Alexa Fluor 633 (emits red fluorescence) conjugated secondary antibody (see Materials and Methods section for the detailed methodology). The leaves infiltrated with the agroconstruct of HV (H) or Q-satRNA alone (I) were treated with RNase III prior to immunolabelling with J2-Ab. Bar, 10 µm.

FIG. 4 Analysis of the junction sequences in the Q-satRNA multimers formed in the presence or absence of HV. PCR products of the junction regions of the Q-satRNA multimers formed in the presence or absence of HV using the primer set Fw1+Rv1 (A) or Fw2+Rv2 (B). In each panel, the agarose gel analysis of the PCR products resulting from each primer set was shown on the left. On the right, schematic representation of the dimeric form of Q-satRNA and the location of each primer are shown. The primer Fw2 was designed to contain a guanine residue beyond the authentic terminal cytosine residue at the 3'-end of Q-satRNA. To assess the specificity of each primer set, 50 ng of pQSx2-ΔC (a Q-satRNA dimer construct lacking the terminal C-residue from the 3'-end of the first monomeric unit) and pQSx2-HNM [a Q-satRNA dimer construct having the HNM (GGGAAAAA) between monomers] were used in PCR reactions. (C) Summary of the sequence analyses of the junction region of the Q-satRNA multimers from the agroinfiltrated or mechanically inoculated leaves. The numbers on the right indicate the
number of cDNA clones with a specific sequence (shown on the left) among the total number of cDNAs that were sequenced.

FIG. 5  *In situ* evidence for the nuclear localization of Q-satRNA in the presence of HV. (A-D) Representative confocal microscopic images showing the distribution of GFP in the leaves expressing the indicated mixture of agroconstructs. The nature of agroconstructs, experimental conditions and confocal microscopy are as described in Fig. 2 legend. In all panels, GFP expression (green), light mode, and merged images are shown. In addition, in panel D, CFP expression (blue) was shown. Bar= 20 μm.
Fig. 3
Fig. 4

A

B

506
396
344
298
bp

Unit1
Unit2
Fw1 Fw1
Rv1 Rv1
5' 3'
- -GACCCGTTTT-- - -GACCCGTTTT--

B

506
396
344
298
bp

Unit1
Unit2
Fw2 Fw2
Rv2 Rv2
5' 3'
- -TCCCTACCAGGACCGTTTTGTTTGTT AGA-- - -TCCCTACCAGGACCGTTTTGTTTGTT AGA--

C

Inoculum Primers for PCR Sequence in the junction Agroinfiltration Mechanical inoculation

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>Primers for PCR</th>
<th>Sequence in the junction</th>
<th>Agroinfiltration</th>
<th>Mechanical inoculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-satRNA</td>
<td>Fw1+Rv1</td>
<td>TGAGAAAATTTTTGG</td>
<td>46/46</td>
<td>20/20</td>
</tr>
<tr>
<td>HV+Q-satRNA</td>
<td>Fw1+Rv1</td>
<td>TGAGAAAATTTTTGG</td>
<td>40/40</td>
<td>20/20</td>
</tr>
<tr>
<td>Q-satRNA</td>
<td>Fw2+Rv2</td>
<td>TGAGAAAATTTTTGG</td>
<td>23/23</td>
<td>20/20</td>
</tr>
<tr>
<td>HV+Q-satRNA</td>
<td>Fw2+Rv2</td>
<td>TGAGAAAATTTTTGG</td>
<td>15/15</td>
<td>15/15</td>
</tr>
</tbody>
</table>
Fig. 5