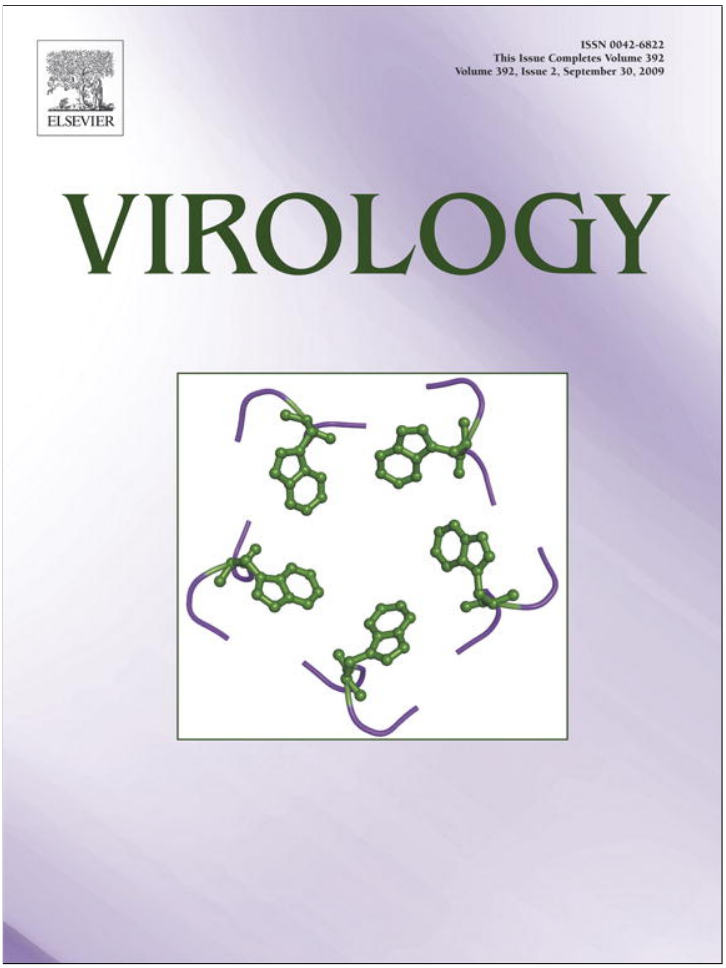


Virology 392 (2009) 203–214

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

# Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes

Livia Donaire [a], Yu Wang [b], Daniel Gonzalez-Ibeas [c], Klaus F. Mayer [b], Miguel A. Aranda [c], César Llave [a,*]

[a] Departamento de Biología de Plantas, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040-Madrid, Spain
[b] Institute for Bioinformatics and System Biology, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Inglostaedter Landstr. 1, D-85764 Neuherberg, Germany
[c] Departamento de Biología del Estrés y Patología Vegetal, Centro de Edafología y Biología Aplicada del Segura, CSIC, apdo. correos 164, 30100-Espinardo, Murcia, Spain

ARTICLE INFO

*Article history:*
Received 23 April 2009
Returned to author for revision 29 June 2009
Accepted 10 July 2009
Available online 7 August 2009

*Keywords:*
Plant–virus interactions
Plant virus
Viral small RNAs
RNA silencing

ABSTRACT

Plant virus infection involves the production of viral small RNAs (vsRNAs) with the potential to associate with distinct Argonaute (AGO)-containing silencing complexes and mediate diverse silencing effects on RNA and chromatin. We used multiplexed, high-throughput pyrosequencing to profile populations of vsRNAs from plants infected with viruses from different genera. Sense and antisense vsRNAs of 20 to 24 nucleotides (nts) spread throughout the entire viral genomes in an overlapping configuration; virtually all genomic nucleotide positions were represented in the data set. We present evidence to suggest that every genomic position could be a putative cleavage site for vsRNA formation, although viral genomes contain specific regions that serve as preferential sources of vsRNA production. Hotspots for vsRNAs of 21-, 22-, and 24-nt usually coincide in the same genomic regions, indicating similar target affinities among Dicer-like (DCL) enzymes. In the light of our results, the overall contribution of perfectly base paired double-stranded RNA and imperfectly base paired structures within single-stranded RNA to vsRNA formation is discussed. Our census of vsRNAs extends the current view of the distribution and composition of vsRNAs in virus-infected plants, and contributes to a better understanding of vsRNA biogenesis.



## Introduction

The final outcome of virus infections in plants depends on fine-tuned compatible and defensive interactions between hosts and viruses (Maule et al., 2002). Among the cellular mechanisms underlying the molecular basis of these interactions, RNA silencing plays a critical role by providing a complex matrix of gene regulation that target both host and viral genomes (Ding and Voinnet, 2007; Dunoyer and Voinnet, 2005). In plants, RNA silencing is triggered by RNA with double-stranded (ds) features which serve as a substrate for Dicer-like ribonucleases (DCL) to produce two major classes of small RNAs (sRNA): ~21-nucleotide (nt) microRNAs (miRNAs) and small interfering RNAs (siRNAs) of ~21 to 25 nts (Brodersen and Voinnet, 2006). Host-encoded RNA-dependent RNA polymerases (RDR) are required for the production of each siRNA class by converting single-stranded (ss) RNA into dsRNA (Kasschau et al., 2007). sRNAs associate with distinct Argonaute (AGO)-containing effector complexes to guide them to their RNA or DNA target molecules (Hutvagner and Simard, 2008; Sontheimer and Carthew, 2005; Vaucheret, 2008).

Plant viruses activate the RNA silencing machinery in infected cells through the formation of viral dsRNA by any of several mechanisms that include the activity of virus-encoded RNA polymerases, intermolecular base pairing between plus and minus strand viral RNAs, and imperfect folding of self-complementary sequences within viral ssRNA (Ding and Voinnet, 2007). In addition, at least three functional RDR enzymes have been recognized as antiviral effectors implicated in biosynthetic pathways of viral sRNAs (vsRNA), suggesting that RDRs use viral RNA as a template to synthesize negative complementary strands (Diaz-Pendon et al., 2007; Donaire et al., 2008; Qi et al., 2009). Upon virus infection, vsRNA of 20 to 25 nts in length are generated to guide autosilencing of viral RNAs as part of an antiviral self-defence response in plants (Ding and Voinnet, 2007; Pantaleo et al., 2007). Multiple *AGO* genes might be involved in anti-virus defence (Zhang et al., 2006) and, at least, two AGO proteins (AGO2 and AGO5) have been shown to bind vsRNAs (Takeda et al., 2008). Moreover, the regulatory potential of these molecules presumably involves functional interactions with host transcripts through perfect or near-perfect base pairing (Moissiard and Voinnet, 2006; Qi et al., 2009).

Compelling evidence indicates that the biogenesis of vsRNAs of different size classes involves the same DCL-dependent pathways responsible for the formation of endogenous siRNAs (Bouche et al., 2006; Gasciolli et al., 2005; Xie et al., 2004; Yoshikawa et al., 2005). In *Arabidopsis*, DCL4, DCL2 and DCL3 target viral genomes in a hierarchical fashion to yield vsRNAs of 21-, 22- and 24-nts, respectively. Antiviral immunity is conferred by DCL4-dependent,

* Corresponding author. Fax: +34 915 36 04 32.
*E-mail address:* cesarllave@cib.csic.es (C. Llave).

0042-6822/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2009.07.005

**Table 1**
Viruses and host plants used for construction of vsRNA libraries from virus-infected plants.

| Virus name | Acronym | Family | Genus | Genome type | Host | sRNA reads | vsRNA reads |
|---|---|---|---|---|---|---|---|
| *Melon necrotic spot virus* | MNSV | Tombusviridae | Carmovirus | (+)-ssRNA, monopartite | *Cucumis melo* | 48,023 | 27,291 |
| *Cymbidium ring spot virus* | CymRSV | Tombusviridae | Tombusvirus | (+)-ssRNA, monopartite | *Nicotiana benthamiana* | 34,998 | 22,468 |
| *Tobacco rattle virus* | TRV | Not assigned | Tobravirus | (+)-ssRNA, bipartite | *Arabidopsis thaliana* | 25,188 | 5224 |
| *Cucumber mosaic virus* | CMV | Bromoviridae | Cucumovirus | (+)-ssRNA, tripartite | *Arabidopsis thaliana* | 17,410 | 2464 |
| *Pepper mild mottle virus* | PMMoV | Not assigned | Tobamovirus | (+)-ssRNA, monopartite | *Nicotiana benthamiana* | 27,824 | 4411 |
| *Watermelon mosaic virus* | WMV | Potyviridae | Potyvirus | (+)-ssRNA, monopartite | *Cucumis melo* | 76,158 | 1473 |
| *Turnip mosaic virus* | TuMV | Potyviridae | Potyvirus | (+)-ssRNA, monopartite | *Arabidopsis thaliana* | 16,699 | 497 |
| *Potato virus X* | PVX | Flexiviridae | Potexvirus | (+)-ssRNA, monopartite | *Nicotiana benthamiana* | 17,309 | 651 |
| *Tomato yellow leaf curl virus* | TYLCV | Geminiviridae | Begomovirus | ssDNA, circular | *Solanum lycopersicum* | 17,737 | 1212 |

21-nt vsRNAs with DCL2 acting as a DCL4 surrogate (Blevins et al., 2006; Bouche et al., 2006; Deleris et al., 2006; Diaz-Pendon et al., 2007; Donaire et al., 2008; Fusaro et al., 2006). 24-nt vsRNAs produced by DCL3 might be related to the perception of non-cell autonomous silencing signals (Brosnan et al., 2007; Diaz-Pendon et al., 2007).

Several studies using RNA and DNA viruses revealed that vsRNAs originate from multiple genomic regions (Chellappan et al., 2004; Donaire et al., 2008; Ho et al., 2007; Moissiard and Voinnet, 2006; Molnar et al., 2005; Qi et al., 2009; Szittya et al., 2002). Although this observation may be diagnostic of a highly varied pool of vsRNAs in the infected tissue, the overall composition of the populations of vsRNAs generated by most plant viruses remains largely unknown. In this study, we use high-throughput DNA pyrosequencing to profile vsRNA across different plant virus genomes (Table 1). We report on a large and diverse population of vsRNAs and provide details on their origin, size, distribution and abundance.

## Results

### *Composition of the vsRNA populations in infected plants*

vsRNA populations from *Arabidopsis* plants infected with *Tobacco rattle virus* (TRV), *Turnip mosaic virus* (TuMV) or *Cucumber mosaic virus* (CMV), from *Nicotiana benthamiana* plants infected with *Cymbidium ringspot virus* (CymRSV), *Potato virus X* (PVX) or *Pepper mild mottle virus* (PMMoV), from *Cucumis melo* plants infected with *Melon necrotic spot virus* (MNSV) or *Watermelon mosaic virus* (WMV) and from *Solanum lycopersicum* plants infected with *Tomato yellow leaf curl virus* (TYLCV) were profiled using high-throughput pyrosequencing technology (Table 1). sRNA was prepared from infected tissue, ligated to 5′ and 3′ adapters, amplified by RT-PCR and subjected to multiplexed sequencing. Adapters were designed to specifically ligate vsRNAs containing 5′ monophosphate and 3′ hydroxyl ends, consistent with DCL-catalyzed cleavage products (Kasschau et al., 2007). A total of 281,346 reads with recognizable flanking adapter sequences and with lengths ranging between 17 and 28 nts were searched against the corresponding viral genomes. Only sequences that matched perfectly were further analyzed. A low proportion of reads containing single-position mismatches were discarded from further analyses. Each vsRNA read could be unambiguously assigned to one unique genome position. In total, 65,691 reads were considered as vsRNAs representing 23,055 unique, although frequently overlapping, sequences (Table 1).

vsRNAs between 17 to 28 nts were recovered from our libraries although reads in the range of 20 to 24 nts constituted 92% (59,984) of the total (Fig. 1). For most viruses tested, vsRNA of 21 nts was clearly the predominant class followed by 22-nt vsRNA, together accounting

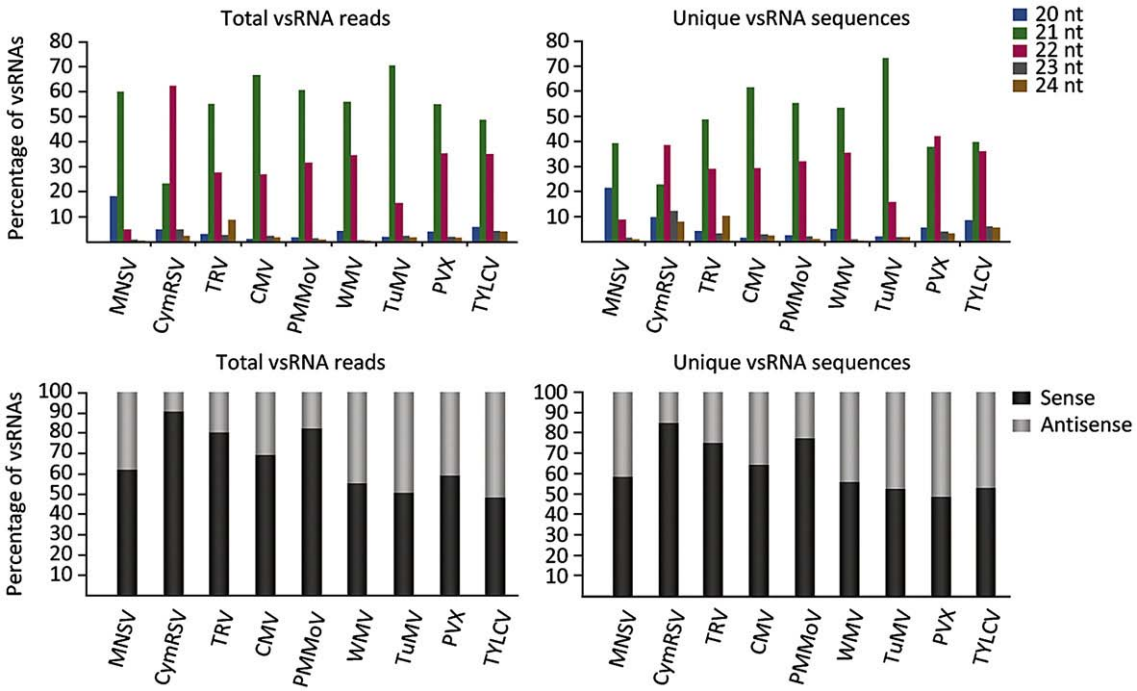



**Fig. 1.** Size (top) and polarity (bottom) distribution of sequenced vsRNAs from virus-infected plants. Histograms represent the percentage of total or unique vsRNA reads within each category.

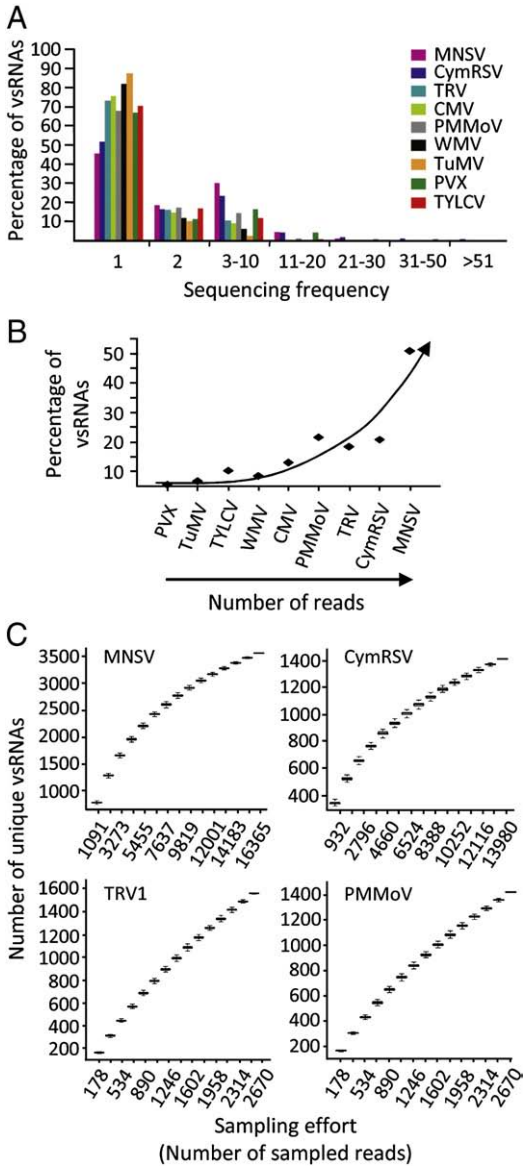



**Fig. 2.** Estimation of saturation of vsRNAs and species richness in the sequenced pools. (A) Sequencing frequencies of vsRNAs. Frequencies (*x*-axis) refer to as the number of times each unique vsRNA of 20 to 24 nts was sequenced within each library. (B) Percentage of vsRNAs within each library with respect to the maximum number of possible vsRNAs. Libraries are arranged on the *x*-axis from left to right according to their number of sequenced reads. (C) Quantification of species richness and sequencing effort. Shown are individual-based rarefaction curves of vsRNAs where the unique vsRNAs (*y*-axis) are plotted (Box-and-Whisker Plot) as a function of the sequencing effort (*x*-axis). Curves represent the means of repeated resampling of all pooled reads within each library. Note that the asymptote is reached only when all reads have been recovered. Data in (B) and (C) are relative to 21-nt vsRNAs, except for CymRSV that are to 22 nts.

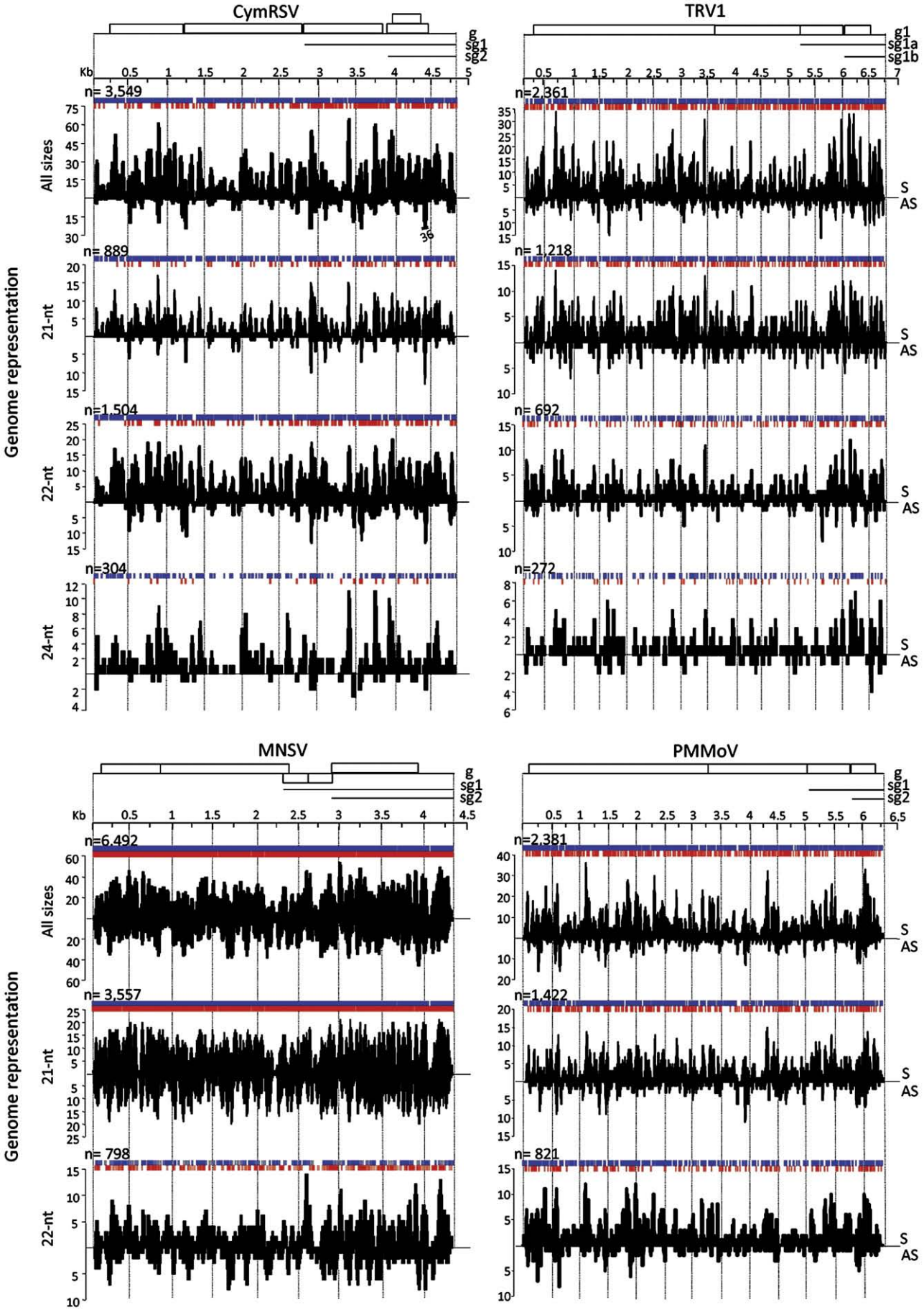
for 77.5% of total reads. The only exception was for CymRSV, that accumulated 22-nt vsRNAs to higher levels relative to 21-nt species (Fig. 1). Our deep-sequencing data conflicted with previous data obtained by low-scale sequencing of CymRSV-derived vsRNAs, showing that the vsRNAs were predominately of 20 and 21 nts, whereas the 22-nt species were underrepresented (Molnar et al., 2005). We reasoned that this difference was likely due to the low number of vsRNA sequences (228) analyzed and the cloning procedure used in that study (Molnar et al., 2005). The dominance of CymRSV-specific 22-nt vsRNAs might be a consequence of the preferential affinity of the p19 silencing suppressor of CymRSV to sequester sRNAs of 21 nts (Vargason et al., 2003). Alternatively, p19 might partially interfere with DCL4-mediated processing of CymRSV substrates, as it has been described for the p38 silencing suppressor of *Turnip crinkle virus* (TCV), which blocks DCL4 activity and promotes the preferential accumulation of DCL2-dependent, 22-nt vsRNA species (Deleris et al., 2006; Xie et al., 2004). In general, the strong bias in size distribution was consistent with the hierarchical action of DCL4, DCL2 and DCL3 in the biogenesis of vsRNAs (Bouche et al., 2006; Deleris et al., 2006) and suggested that all viruses tested were targeted by a common RNA silencing pathway regardless of their taxonomic origins.

vsRNA from plants infected with CymRSV, TRV, CMV or PMMoV showed asymmetrical distribution in strand polarity with dominance of sense vsRNAs compared to antisense species, while MNSV-derived sRNAs showed only a modest enrichment of sense species (Fig. 1). In contrast, vsRNAs from WMV, TuMV, PVX and TYLCV almost equally derived from positive and negative viral strands. Similar results were obtained from analysis with data sets containing unique vsRNA sequences (Fig. 1). For most viruses tested, vsRNAs displayed significant, albeit modest, differences in base composition at their 5′ termini ($\chi^2_3 > 8.0266$, $P < 0.0454$) with a preference to begin with an uridine (U) or an adenosine (A) and a clear tendency to avoid a guanidine (G). Analysis of total sRNA reads of 21, 22 or 24 nts revealed that vsRNAs of sense polarity with a 5′ terminal U were more abundant (~30%) for MNSV, CymRSV, TRV, CMV, PMMoV and TYCV while sense vsRNAs with a 5′ terminal U or A were equally represented in plants infected with WMV, TuMV and PVX. In contrast, antisense vsRNAs with a 5′ terminal U, A, or cytosine (C) were similarly represented in the overall vsRNA pool. The same distribution patterns were observed for unique vsRNA species and agree with the 5′ terminal base composition recently reported for *Tobacco mosaic virus* (TMV)-derived vsRNAs (Qi et al., 2009). Since loading of sRNAs into a particular AGO complex is preferentially dictated by their 5′ terminal nucleotides (Kim, 2008), our data suggested that vsRNAs have the potential to be recruited into multiple AGO-containing complexes.

The vast majority of the unique vsRNAs of 20 to 24 nts were sequenced in the range of 1 to 30 times. On average, 69% were only sequenced once, and 84% were represented in the sequenced pool by one or two reads (Fig. 2A). Highly repetitive reads that likely accounted for abundant vsRNA species were also recorded. The very low sequence frequencies suggested that libraries were not saturating, so deeper sequencing efforts are required in order to achieve saturation of unique vsRNAs species.

### *Profiling vsRNA diversity*

In order to study spatial distribution and sequence diversity within the populations of vsRNAs, unique vsRNA sequences were mapped across the corresponding viral genomes and the genomic coverage for each virus was calculated. For clarity, only the analysis of viruses that gave the most complete set of data is herein presented. Using all size classes, 100% of the MNSV genome was represented in the sequenced set of sense and antisense vsRNAs as all nucleotide positions in the genome were occupied by at least one unique sequenced vsRNA (Fig. 3). vsRNAs from plants infected with CymRSV or PMMoV encompassed about 98% of the corresponding viral genomes, while TRV, CMV or TYLCV genomes were covered by 80 to 96% (Fig. 3 and data not shown). Genomic coverage was also calculated using individual sizes, under the assumption that the vast majority of the most abundant 21- and 22-nt vsRNAs were authentic mature DCL4- and DCL2-cleavage products, respectively. 24-nt vsRNAs were investigated in CymRSV and TRV which generated sizeable populations of this class. Sense and antisense 21-nt vsRNAs produced by MNSV extended over 99.7% of the viral genome, while 95% of the CymRSV genome was reflected in the subset of sense and antisense 22-nt vsRNAs (Fig. 3). 21-nt vsRNAs extended over 90 and 80% of the TRV and CMV genome, respectively (Fig. 3 and data not shown). vsRNAs of



24 nts comprised about 53 and 62% of the TRV and CymRSV genomes, respectively (Fig. 3). These data suggested that DCL targeting occurred widespread at multiple as opposed to isolated positions along the virus genome to form a tapestry of overlapping vsRNAs. Also, as shown in Fig. 3, the profiles of 21-nt and 22-nt vsRNAs were largely similar, indicating a resemblance between DCL4- and DCL2-mediated targeting. For instance, regions with higher density of unique 21-nt vsRNA species overlapped with regions containing a larger number of unique 22-nt vsRNA sequences (Fig. 3). In conclusion, our results expanded beyond previous reports based on RNA hybridizations and small scale cloning of vsRNAs to highlight the genome-wide, massive formation of 20- to 24-nt vsRNAs in a characteristic overlapping configuration.

High-resolution mapping revealed that, in most cases, any given nucleotide position in the viral genome was occupied by numerous unique vsRNA species of both orientations (Fig. 3). In regions with higher density of unique species, vsRNAs overlapped with neighboring sequences to form a ladder of vsRNAs with their 5′ termini spaced at 1-nt increments (Fig. 4). This spacing was suggestive of DCL-mediated cleavage of viral dsRNA occurring at consecutive nucleotide positions along the viral genomes. If this is true and dsRNA comprises the full-length viral RNA, one might expect that the maximum number of unique vsRNAs of each size class generated by a particular virus would be solely constrained by the size of the genome. This theoretical number can be calculated simply as: $n = x - (y - 1)$, in which $x$ = genomic size (nts), and $y$ = sRNA size. Our data revealed that sequenced MNSV-derived, 21-nt vsRNAs, accounted for about 46% of the maximum theoretical, while vsRNAs of 22-nts from the CymRSV library represented nearly 27% (Fig. 2B). This percentage dropped for those viruses with few vsRNA counts in the sequenced pool but raised proportionally as the number of reads increased, suggesting that the theoretical number of all possible unique vsRNAs is potentially achievable under exhaustive sequencing of the vsRNA populations (Fig. 2B). In order to assess the effect of sampling effort on vsRNA species richness, we quantified unique vsRNA species in the sequenced pool using individual-based, species accumulation curves (Fig. 2C) (Gotelli and Colwell, 2001). We assessed graphically whether the sequencing effort was enough to retrieve all unique sequences of a given size class by means of a resampling scheme based on common methods for the measurement of species richness in biodiversity studies (Gotelli and Colwell, 2001). From each set of sequences (21-nt vsRNAs for MNSV, TRV and PMMoV; 22-nt vsRNAs for CymRSV) we randomly sampled one-fifteenth of the total reads and counted the number of unique sequences in this group. Then we repeated the sampling, this time with groups comprising two fifteenth of the observations and counted again the number of unique sequences. We repeated the procedure for groups up to the total number of sequences and built a curve of accumulated unique sequences, which should suggest an asymptote for larger enough sampling efforts. Our results in Fig. 2C clearly demonstrated that the number of unique sequences in our sequenced pool correlated with sampling efforts so as more individuals were sampled, more unique species were recorded. Most importantly, species richness hardly reached an asymptote, and only when practically the entire subset of reads had been sampled.

### *Distribution of vsRNA abundance*

We took advantage of direct high-throughput massive-parallel sequencing as a quantitative indicator of sRNA abundance (Kasschau et al., 2007; Lu et al., 2005; Rajagopalan et al., 2006). This approach is very convenient to achieve discrete measurement of specific sRNA species abundance within large mixtures of overlapping sRNAs. The abundance of vsRNAs within each sequenced pool was plotted in Fig. 5 using reads from all size classes as well as each of the most representative sizes independently. These analyses resulted in vsRNA abundance patterns that reflected a heterogeneous distribution across the entire viral genome and showed genomic regions with different densities of vsRNA reads (Fig. 5). A significant twofold enrichment of vsRNA reads within the subgenomic RNA-forming 3′ end region was observed for CymRSV, MNSV and TRV ($t < -3.6591$, $P < 0.0336$) (Fig. 5). In addition, hotspots of vsRNA accumulation were represented by sharp as well as broad peaks of vsRNA abundance scattered throughout the entire viral genome. These peaks clustered multiple reads representing several overlapping unique vsRNA sequences; sharp peaks denoted the presence of highly abundant vsRNA reads within the cluster (Fig. 5). Therefore, quantitative profiling of our data set support that each viral genome contains regions that serve as preferential sources of vsRNA production. In fact, for most viruses tested, the most prominent peaks of sequence diversity and abundance corresponding to vsRNAs of 21 nts usually localized within the same genomic regions as peaks corresponding to 22- or 24-nt vsRNAs (Figs. 3–5). This observation seems to indicate that all DCL activities generating vsRNAs show similar targeting affinities toward the same regions in a particular genome.

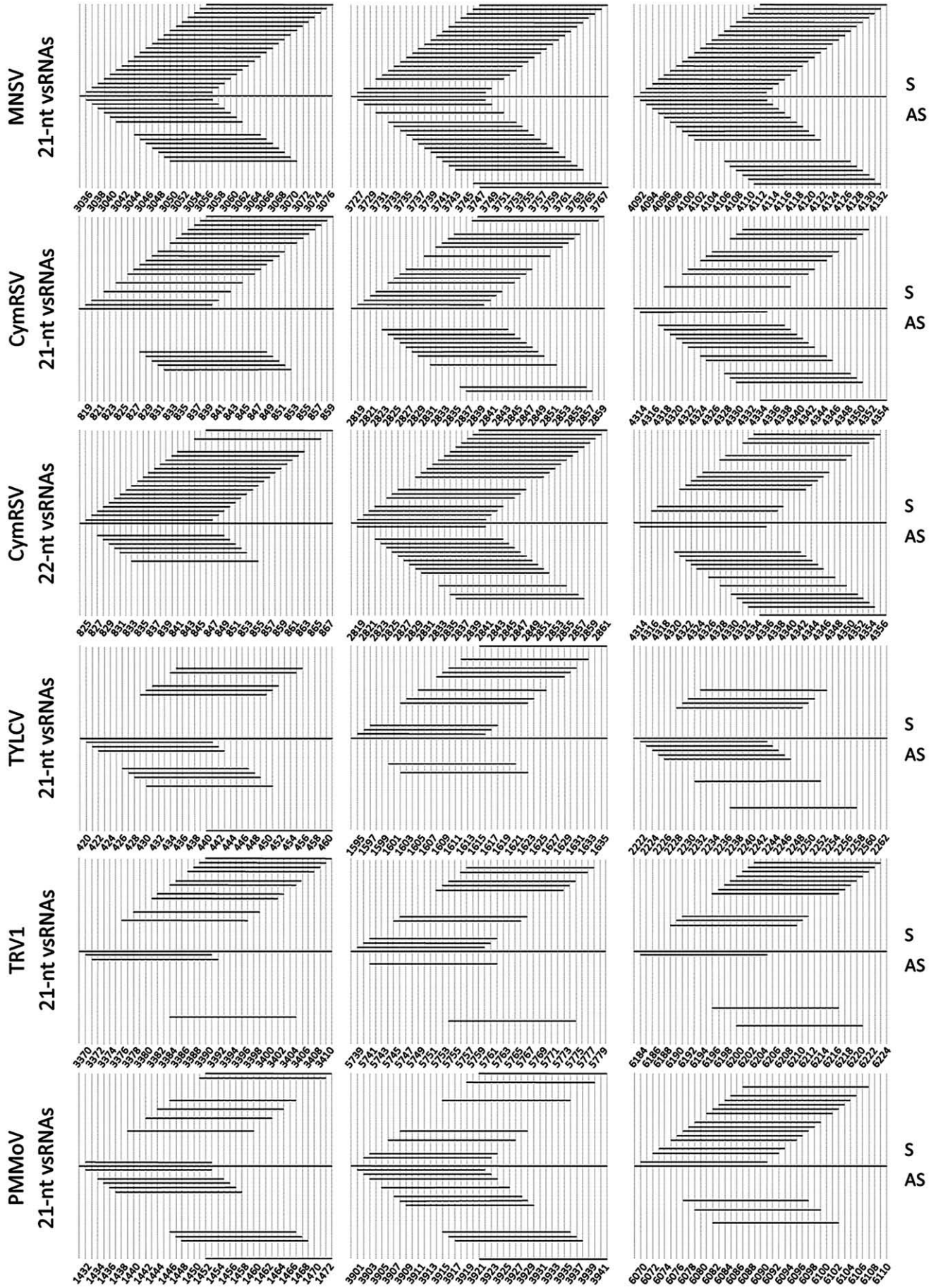
It was proposed that DCL enzymes function preferably on GC-rich regions, and that hotspots of vsRNA formation likely represented genomic segments with higher GC content (Ho et al., 2007). In our sequencing approach, there was a modest but significant bias in the GC content between vsRNAs and the corresponding virus genomes ($t > 2.12$, $P < 0.02$) (Supplementary Fig. 1). Also, the GC content was slightly higher in the hotspots compared to other viral genomic regions ($t > 1.75$, $P < 0.04$), supporting the previous view that GC-rich regions might provide a relatively more stable dsRNA substrate for DCL processing (Supplementary Fig. 1) (Ho et al., 2007).

We next used vsRNA abundance and positions occupied by each vsRNA to explore the possibility that a fraction of vsRNAs in each library were processed in particular phase registers. Identification of predominant phase registers along the virus genome or within specific genomic regions might be suggestive of a biosynthetic pathway involving consecutive processing of viral dsRNA from preferential defined termini, reminiscent of that for endogenous trans-acting (ta)-siRNAs (Allen et al., 2005; Axtell et al., 2006; Howell et al., 2007). We randomly set five sequence windows, each corresponding to 10 cycles of DCL processing, along the virus genome, and searched for representative in-phase positions for cycles of 21 or 22 nts. All vsRNA reads with a 5′ start coordinated into each of the possible 21 or 22 phasing registers within each window were determined (Howell et al., 2007). For all viruses tested, we could not detect a substantial enrichment of vsRNA reads in-phase at any particular register, suggesting a random distribution of vsRNAs falling into each of the possible registers (Supplementary Fig. 2). This observation is in good agreement with the ubiquitous spreading of vsRNAs.

### *Analysis of secondary structure within viral ssRNA*

It has been proposed that hotspots of vsRNAs might reflect preferential DCL activities on foldback structures within ssRNA (Molnar et al., 2005; Szittya et al., 2002). If this hypothesis is true, one might expect that sequences surrounding hotspots adopt stable

**Fig. 3.** Distribution of unique vsRNAs along viral genomes. A blue/red bar on the top of each graph stands for the genome coverage and measures the extent to which the viral genome was represented in the pool of unique vsRNA. Nucleotide genomic positions occupied by at least one unique read are indicated in blue (sense species) and red (antisense species). Graphs plot the number of unique vsRNA that hit every genomic position within each library. The number of unique vsRNA sequences (n) is indicated in each case. Bars above the axis represent sense (S) reads; those below represent antisense (AS) reads. Schematic representation of viral genomes including regions producing subgenomic RNAs are shown.



hairpin structures that are robustly predicted by *mfold*, as miRNA-like precursors (Jones-Rhoades and Bartel, 2004; Reinhart et al., 2002). To test the hypothesis, the vsRNA sequences were compared with predicted RNA secondary structure maps generated using *mfold* 3.2 (Zuker, 2003). The sequence and structure properties previously defined for recognition of miRNA-like precursors were used as constraints to search for putative foldback precursor of vsRNAs from viral ssRNA (Jones-Rhoades and Bartel, 2004; Meyers et al., 2008; Wang et al., 2004). As an illustrative example, we show data from CymRSV-derived secondary structures because CymRSV contained a large number vsRNA reads and nine hotspots of vsRNA accumulation were conspicuously identified along the virus genome. Each hotspot consisted of a cluster of unique, usually overlapping, vsRNAs where at least one vsRNA was sequenced more than 100 times. In general, sequences surrounding the vsRNA hotspots exhibited various degrees of intramolecular base pairing (Fig. 6). We determined the precise position of each of the vsRNA hits within the most favorable predicted stem-loops (Fig. 6). Some, but not all, of the most abundant vsRNA reads derived from hotspots 6 and 7 mapped within the longest base paired segment of the hairpin at a position consistent with cleavage at the center of the stem region (Fig. 6). Hotspots 3 and 5 also corresponded to regions predicted to form stable secondary structures, but they occupied positions within the foldbacks with limited base pairing, such that there were a large number of mismatched nucleotides and symmetric and asymmetric bulges (Fig. 6). Highly abundant reads at hotspots 1, 2 and 9 mapped in genomic regions predicted to fold into secondary structures containing an elevated number of G:U pairs, bulged or unpaired nucleotides (Fig. 6). The hotspot 8 corresponded to a region predicted to lack extensive base pairing (data not shown). Folding analysis of genomic sequences from PMMoV, TYLCV and TRV yielded similar results (data not shown) (Donaire et al., 2008). Taken together, some hotspots of vsRNAs occupied positions within the overall secondary structures that satisfied the size and structural criteria of stem–loops generating miRNAs in plants, while other hotspots occupied positions that were presumably suboptimal for DCL processing (Jones-Rhoades and Bartel, 2004; Meyers et al., 2008; Wang et al., 2004). In the latter case, regions with extensive base pairing within the predicted overall secondary structures were, in general, poorly represented in the vsRNA pool. These findings suggest that cleavage of foldbacks within viral ssRNA is unlikely to be the sole determinant to explain the formation of hotspots.

## Discussion

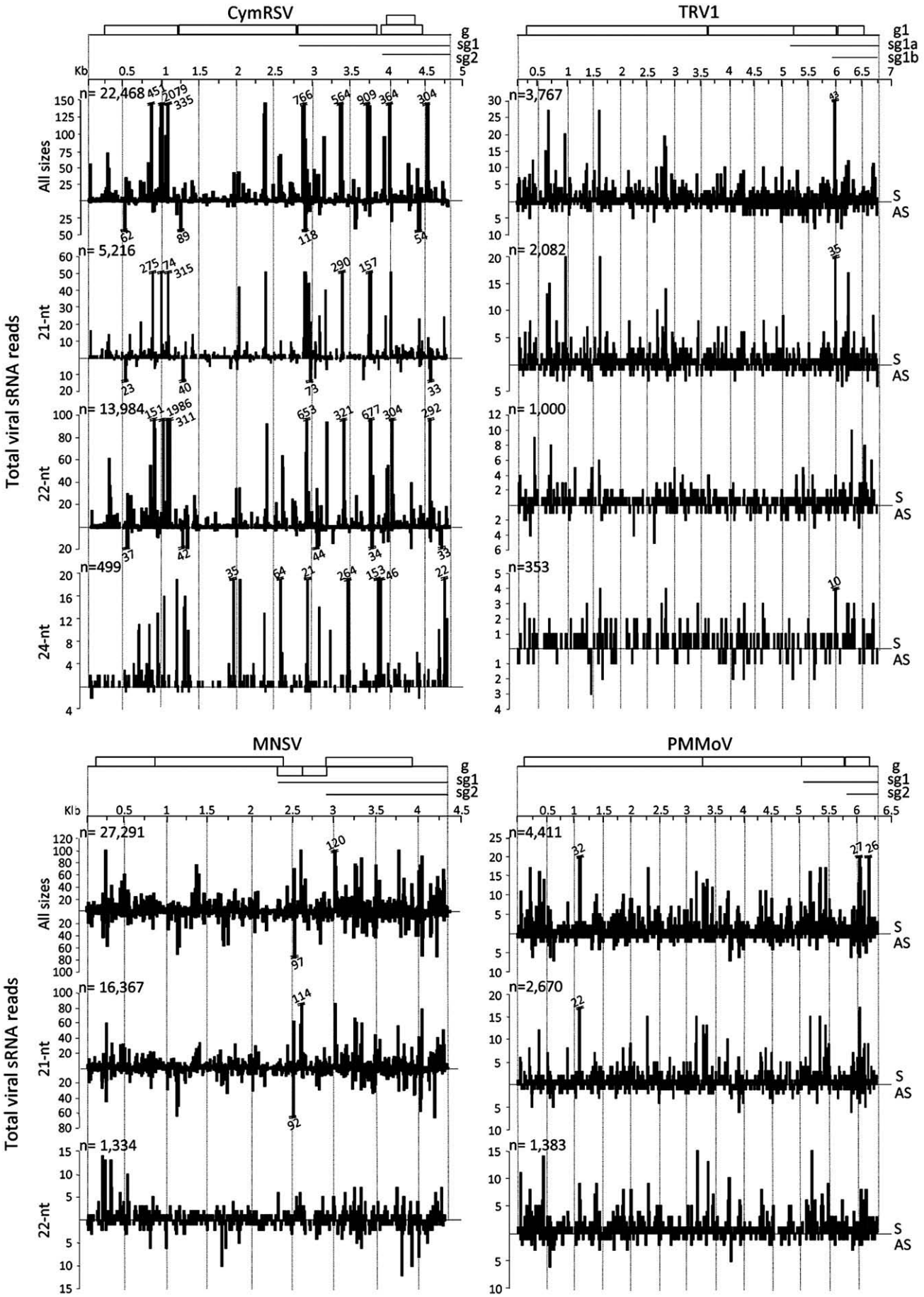
In this study, we used a variety of plant viruses with distinct genome organization, genome expression and replication strategies as well as different host plant species to present a general, comprehensive scenario that extends the current view of the composition and distribution of vsRNAs in infected plants. Populations of vsRNAs were abundant, diverse and reflected an effective and widespread targeting of viral genomes by several interconnected RNA silencing pathways, regardless of the host plant and virus tested.

Sequence analysis of the data set revealed that the total number of vsRNA reads varied between the nine source libraries. This is likely due to intrinsic differences in the replication and virus accumulation rates between all viruses tested and in the efficiency of the RNA silencing machinery to recognize and target each viral genome in the corresponding host. We presume that these two factors may drastically influence to what extent dsRNA is formed from viral RNA templates. In addition, the levels of vsRNA accumulation in the infected tissue largely reflect the mode of action of the silencing suppressor encoded by each virus (Li and Ding, 2006). This latter situation is well-illustrated, for instance, by the potyviral HC-Pro and the TRV-16K silencing suppressors which cause a reduction in the accumulation of vsRNAs that is more accentuated in the presence of HC-Pro compared to TRV-16K (Martinez-Priego et al., 2008). Likewise, the CMV-2b silencing suppressor specifically interferes with RDR1 to inhibit the production of 21-, 22- and 24-nt classes of secondary, CMV-derived vsRNAs in the infected tissue (Diaz-Pendon et al., 2007). Nevertheless, we cannot exclude that our sequencing strategy using a multiplexed format created a bias among the libraries probably due to differences in amplicon concentrations, as the total number of host sRNA also differed between libraries.

Production of sRNAs in eukaryotes evokes two well-differentiated mechanisms that involve i) DCL processing of structured RNA or dsRNA, and ii) RDR-mediated RNA synthesis directly from ssRNA (Voinnet, 2008). The presence of a 5′ monophosphate in vsRNAs, which is a characteristic feature of DCL-catalyzed cleavage products, can be inferred from ligation experiments with T4 RNA ligase, providing evidence that vsRNAs in our sequenced pool were DCL products. Moreover, the fact that the vsRNAs showed a clear bias in size distribution discounts the possibility that the vsRNA species sequenced in this study resulted mainly from non-specific RNA degradation. Our sequenced vsRNAs could be classified into the three major size classes predicted by the coordinated hierarchical actions of DCL4, DCL2 and DCL3 in vsRNA biogenesis (Bouche et al., 2006; Deleris et al., 2006). But, is dicing the sole mode of vsRNA production in plants? The accumulation of vsRNAs from several RNA and DNA plant viruses is severely compromised in loss-of-function *dcl2*/*dcl3*/*dcl4 Arabidopsis* mutants, indicating that DCL enzymes are major components of vsRNA biogenesis and antiviral defense (Bouche et al., 2006; Deleris et al., 2006; Diaz-Pendon et al., 2007; Donaire et al., 2008; Moissiard and Voinnet, 2006; Qu et al., 2008). In our study, deep-sequencing-based results regarding length and strand polarity distributions or spatial profiles are in tune with previous results using RNA blot analyses (Akbergenov et al., 2006; Chellappan et al., 2004; Diaz-Pendon et al., 2007; Donaire et al., 2008; Fusaro et al., 2006; Ho et al., 2006; Szittya et al., 2002; Xie et al., 2004), suggesting that our sequenced set of DCL-dependent vsRNAs fairly represents the entire population of vsRNA produced by each virus in the infected tissue. In *Caenorhabditis elegans*, secondary siRNAs are abundant and result from unprimed RNA synthesis by an RDR activity (Pak and Fire, 2007; Sijen et al., 2007). Therefore, RDR products carry 5′ di- or triphosphates and are unlikely to be recovered using 5′-ligation dependent cloning methods. Whether a fraction of vsRNAs in plants results from direct RDR-dependent secondary vsRNA biosynthesis will require further experimentation using cloning procedures that takes their triphosphorylated status into account.

Unique vsRNAs showed a genome-wide distribution. For viruses with the largest number of reads in the sequenced set, virtually all nucleotide positions in the genome were occupied by at least one unique vsRNA. Moreover, our results suggested that every single nucleotide position within the viral genome can be a putative cleavage site for vsRNA formation (see Fig. 4). DCL-catalyzed cleavage occurring at any of all the nucleotide positions along the viral genome would set a maximum number of unique vsRNA species of each size class to be produced by each particular virus. This theoretical number is only constrained by the size of the genome and reflects the enormous sequence diversity of the vsRNA population in the infected tissue. Extrapolating from sequencing frequencies and library saturation (evaluated using species richness within the sequenced pool), an

**Fig. 4.** Overlapping configuration of the vsRNA populations. Several representative regions showing overlapped vsRNAs are illustrated for some viruses tested. For presentation purposes, sense (S) and antisense (AS) vsRNA with their 5′ and 3′ termini, respectively, located within a 21- or 22-nt window sequence are only represented. Sense vsRNA are 5′–3′; antisense vsRNAs are 3′–5′. Gridlines correspond to 1-nt increments. Nucleotide coordinates are indicated below each scheme. Note that many more vsRNA sequences (not shown) were found surrounding the regions depicted in the figure.



asymptote representing the theoretical maximum number of unique vsRNA species generated by each virus might eventually be reached under exhaustive sampling of the entire vsRNA population.

Our sequence analysis indicated that vsRNAs were often biased to the genomic sense strand, as recently reported for TuMV- or TMV-derived vsRNA (Ho et al., 2007; Qi et al., 2009). Derivation of comparable quantities of sense and antisense vsRNAs (as observed for WMV-, TuMV-, PVX- and TYLCV-derived vsRNAs) from perfectly complementary dsRNA derived from intermolecular base pairing of positive and negative viral strands is relatively easy to envision. Conversely, preferential accumulation of vsRNAs of sense polarity indirectly supports a model by which folded RNA within viral ssRNA serves as a substrate for DCL cleavage (Molnar et al., 2005). It is a common inference that these two pathways of dsRNA formation are operational during initiation and maintenance of virus-induced gene silencing in plants (Ding and Voinnet, 2007; Voinnet, 2008). DCL enzymes might be recruited into limited foldback structures to generate primary vsRNAs, while secondary vsRNA production may engage DCL processing coupled to RDR-mediated synthesis of longer complementary viral RNA. In our study, dominance of vsRNA species originating from viral positive strands was well-illustrated for CymRSV, TRV and PMMoV, which accumulated over 80% vsRNAs of sense polarity. A key question, however, is whether the asymmetrical distribution of strand polarity accurately reflects a major contribution of secondary structures from viral genomic RNA to vsRNA formation. In other words, is the cleavage of structured RNA sufficient to explain the dominance of sense vsRNA species? Our analyses of secondary structure and vsRNA distribution revealed that putative foldback structures could be predicted along viral ssRNA and that vsRNAs mapping at the center of highly base paired regions within these structures could be detected. This observation indirectly supported the idea that secondary structures could be potentially targeted by DCL activities to generate a discrete number of sense vsRNA. However, in general, there was little correlation, if any, between regions of predicted local base pairing RNA and regions where positive vsRNAs originated. Indeed, such a correlation has never been demonstrated for any viral genome (Donaire et al., 2008; Ho et al., 2007; Qi et al., 2009). Consequently, we suspect that it is unlikely that processing of imperfectly base paired hairpins solely accounts for the excess of sense vsRNA. Nonetheless, further research is needed to elucidate if these proposed base paired structures exist *in vivo* and whether they are recognized and cleaved by DCL enzymes. At the current stage, we do not have experimental data to propose an alternative explanation to the bias in the polarity ratio but strand asymmetry differences between viruses could indicate differences in the relative contribution of distinct mechanisms of dsRNA formation between viruses and the participation of specific unidentified host and/or virus-encoded factors. Recently, it has been hypothesized that the nascent viral RNA strands resulting from RDR activities might be chemically modified to prevent the negative vsRNA strand from the vsRNA duplex from entering into an AGO complex (Qi et al., 2009). Alternatively, intrinsic structural and/or biochemical signatures particularly associated to the viral positive strand of the vsRNA duplex might favor the selective recruitment of sense molecules into AGO proteins.

Regardless of whether primary vsRNAs are formed from viral folded ssRNA, the widespread distribution of sense and antisense vsRNAs in an overlapping configuration and their biogenesis at consecutive positions along the virus genome conciliate with DCL-mediated processing of perfectly base paired, relatively long dsRNA as a principal contributor to vsRNA formation. This is consistent with the major role of several RDR-dependent pathways in the biogenesis of vsRNA and the maintenance of virus-induced RNA silencing in plants, a process linked to secondary siRNAs (Axtell et al., 2006; Diaz-Pendon et al., 2007; Donaire et al., 2008; Mourrain et al., 2000; Qi et al., 2009; Schwach et al., 2005; Yu et al., 2003). It is worth noting that there is no conflict between DCL processing of RDR-dependent, viral dsRNA substrates and enrichment of vsRNAs derived from positive RNA strands. This is exemplified by the fact that the largest bulk of vsRNAs of TRV or TMV, which is dominated by sense species, originates from the dicing of RDR products (Donaire et al., 2008; Qi et al., 2009). It would be interesting to determine whether the biogenesis of CymRSV-derived vsRNA is also RDR-dependent. Similarly, ta-siRNAs derived from *Arabidopsis TAS* transcripts seem to accumulate more sense than antisense species, even though they originate through DCL4-mediated processing of RDR6-dependend dsRNA (Allen et al., 2005; Axtell et al., 2006; Howell et al., 2007).

Despite their ubiquitous nature, regions with differential vsRNA density and diversity along each viral genome were identified. Why do some genomic regions serve as preferential sources of vsRNA formation? These areas likely hold structural features that ultimately influence accessibility, affinity or enzymatic activity of one or more components of the RNA silencing machinery required for dsRNA formation and subsequent processing into vsRNAs. For instance, vsRNAs were conspicuous at genomic regions producing subgenomic RNAs that may give rise to a substantial increase in RNA templates for synthesis of dsRNA. dsRNA substrates might also become available in the immediacy of sites for RDR template recognition and initiation of complementary-strand synthesis. RDR activities may be directed to the 3′ ends of viral RNA templates that lack molecular signatures associated to normally processed mRNA, such as those generated through random cleavage events, replication errors or vsRNA-guided cleavage (Allen et al., 2005; Axtell et al., 2006; Herr et al., 2006).

It is also tempting to speculate that vsRNA hotspots mirror a preferential DCL-mediated processing of selected regions containing hairpin secondary structure within viral ssRNA (Molnar et al., 2005; Szittya et al., 2002). In our study, genomic regions surrounding hotspots of vsRNA exhibited a number of possibilities to fold into relatively stable secondary structures. In general, highly repetitive reads were found in regions within the corresponding predicted secondary structures that lack extensive base pairing. In this situation, DCL4, DCL2 and DCL3 should act on imperfect duplexes containing a relatively higher degree of unpaired nucleotides compared to the canonical miRNA hairpin precursors. Although the affinity of these plant DCLs to folded RNA remains to be investigated, we find this scenario less probable as all these DCL enzymes are known to target relatively long, perfect dsRNA substrates, and they show only a residual processing activity on miRNA precursors (Bouche et al., 2006; Kasschau et al., 2007; Kurihara and Watanabe, 2004). In addition, vsRNAs display a clear tendency to begin with U, A and, to a lesser extent, C. This observation is in accord with the previous reporting of AGO proteins with preferred binding affinities for small RNAs having 5′ terminal U (AGO1), A (AGO2 and AGO4), and C (AGO5) (Mi et al., 2008; Montgomery et al., 2008; Takeda et al., 2008). Since the association of sRNAs with a particular AGO protein in plants is not apparently specified by the biogenesis pathways that produce the sRNAs or the structures of their precursors (Mi et al., 2008), it is reasonable to predict that vsRNAs bearing the appropriate 5′ sequence identities may be selectively loaded into multiple AGO complexes, as recently shown for CMV-derived vsRNA associated with AGO2 and AGO5 (Takeda et al., 2008). The low proportion of vsRNA beginning with a G in our data sets is consistent with the absence of AGO proteins known to prefer sRNAs having a 5′ terminal G. This is further supported by the over-representation of miRNA* sequences (arising

**Fig. 5.** Distribution of vsRNA abundance along viral genomes. The number of total reads with a 5′ terminus at each genomic position is plotted. For presentation purposes, maximum values plotted on the *y*-axis were adjusted to show average abundance values. Number on the top of each peak stands for their corresponding highest value. Figure content is as described in Fig. 3.

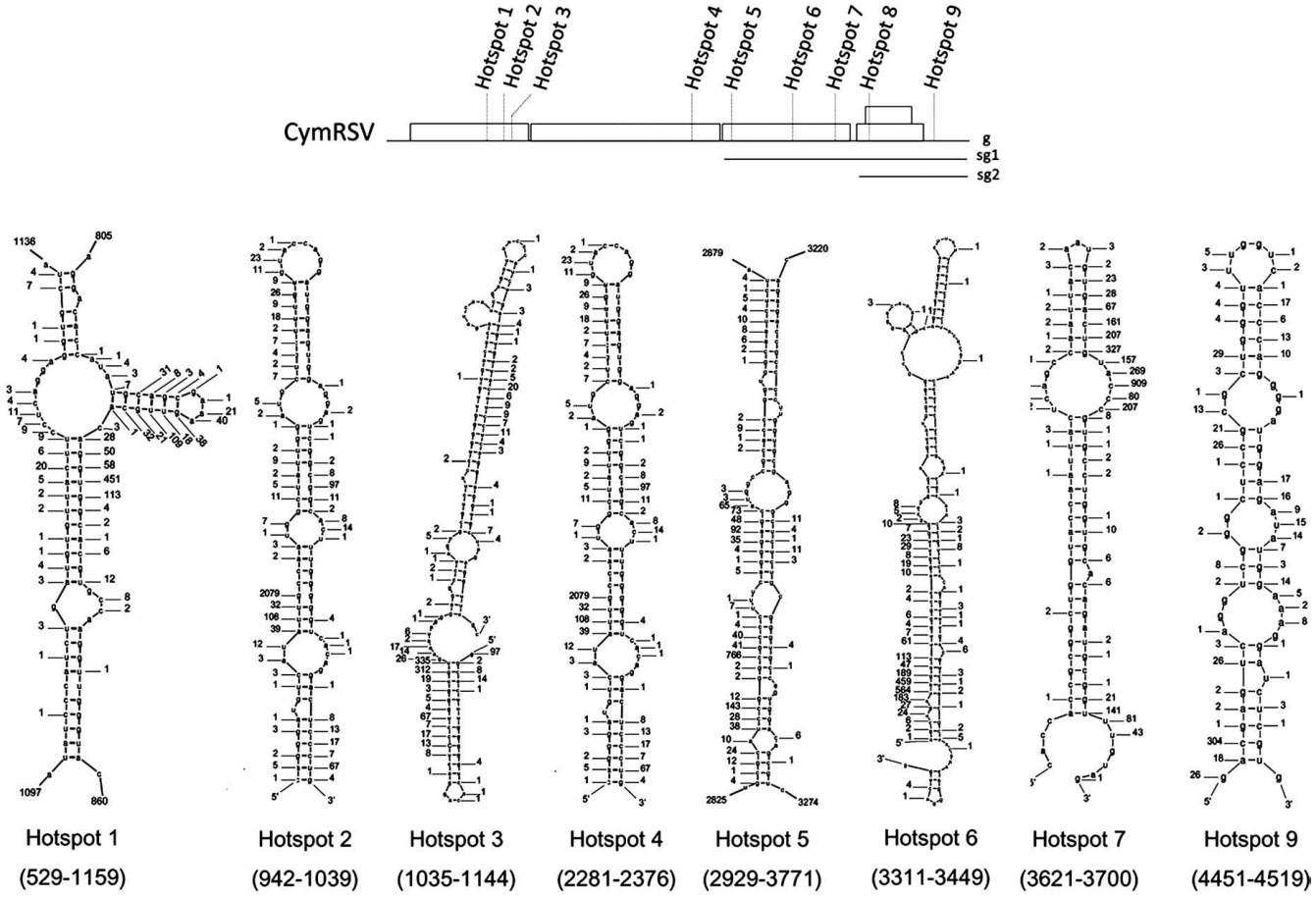



**Fig. 6.** Prediction of secondary structures within CymRSV genomic ssRNA. A schematic of the CymRSV genome indicating the location of each hotspot is shown. The precise position of the 5′ end of each sequenced vsRNA within the predicted foldback structure is indicated by a number, which refers to as the abundance of each vsRNA at each position. Coordinates of the genomic segment subjected to mfolding are given.

from the opposite arm in the miRNA precursor) with a G at their 5′ ends, which are not thought to be associated with AGO complexes in plants.

The manifest potential of plant viruses to produce a diverse and abundant pool of vsRNAs and their likely association with multiple AGO effector silencing complexes have profound implications in the cross-talk interaction between plant and viruses. First, prolific generation of vsRNAs resulting from different DCL activities could reinforce silencing of virus genomes by means of RNA turnover, translational repression, or silencing signaling (Brodersen et al., 2008; Ding and Voinnet, 2007; Mlotshwa et al., 2008; Pantaleo et al., 2007). Second, owing to base complementarity to host genes, vsRNAs could hold an intrinsic regulatory potential, contributing to infection efficacy and symptom expression (Ding and Voinnet, 2007; Dunoyer and Voinnet, 2005; Mlotshwa et al., 2008; Moissiard and Voinnet, 2006). Sorting of vsRNAs into distinct AGO complexes has important functional consequences given that different AGO proteins mediate diverse effects on RNA and chromatin (Hutvagner and Simard, 2008). vsRNA recruited by AGO1- and AGO10-containing complexes may direct cleavage of target mRNAs or inhibition of mRNA translation. vsRNAs with a 5′ terminal A might associate to AGO4 and direct DNA methylation and transcriptional gene silencing at specific genomic loci that share sequence complementarity with the vsRNA. Bioinformatic analyses for target prediction that used a mismatch/gap penalty scoring similar to that used for miRNA target prediction (Fahlgren et al., 2007) have been applied to identify hundreds of host genes as potential targets of vsRNAs (Moissiard and Voinnet, 2006; Qi et al., 2009; Donaire and Llave, unpublished results) (Supplementary Fig. 3). At the present time, the experimental evidence supporting a functional interaction between host mRNAs and vsRNAs is weak and only limited to a couple of genes among the bulk of predicted targets. Nevertheless, this finding suggests that, given the sequence diversity of the vsRNA population, an elevated number of host genes and their regulatory sequences might be targeted by vsRNA-mediated down-regulation during virus infection. The challenge ahead is to determine the extent of these functional interactions between vsRNAs and their targets in a biological perspective.

## Materials and methods

### *Virus and plant material*

Host plants were infected by mechanical inoculation with the following viruses (accession numbers are provided): MNSV (*AY122286*), CymRSV (*NC_003532*) TRV (*NC_003805*, *NC_003811*), CMV (*NC_002035*, *NC_002034*, *D10538*), PMMoV (*NC_003630*), WMV (*AY437609*), TuMV (*AB194802*), PVX (*NC_001455*) and TYLCV (*NC_004005*). Inoculated cotyledons from *C. melo*, upper noninoculated, systemically infected inflorescences from *Arabidopsis* or leaves from the remaining host plants were pooled (6 to 10 plants) and used for vsRNA analyses.

### *vsRNA amplification and sequencing*

Total RNA was extracted using Trizol reagent (Invitrogen) and ~300 μg were used for construction of sRNAs libraries as described (Kasschau et al., 2007) with the following modifications. The 3′ adapter was replaced by a pre-activated 5′ adenylated oligo (5′ rAppCTGTAGGCACCATCAAT3ddC 3′) (Integrated DNA technologies) to avoid the circularization of sRNAs. The chimeric RNA/DNA oligonucleotide 5′ adapters were described previously (Kasschau et al., 2007). A new adapter variant was generated by modification of the four-nucleotide identifier (3-1, ATCGTAGACGCCUGAUA). After each ligation step, sRNA was purified using 17% denaturing PAGE. The purified-ligated sRNA was reverse transcribed and the cDNA was amplified using Taq DNA polymerase (Perkin Elmer) and 3′ PCR FusionB and 5′ PCR FusionA primers (Kasschau et al., 2007). PCR primers contained the "A" and "B" tag sequences used by 454 Life Science during sequencing. DNA amplicons were gel-purified using 12% native polyacrylamide and eluted in 0.3 M NaCl as described (Donaire et al., 2008). Quantity and quality of DNA amplicons were measured using ND-1000 spectrophotometer (Nanodrop) and Experion Automated Electrophoresis System (BioRad), respectively. Same quantity of DNA amplicon from each library was pooled and sequenced by 454 Life Science technology (Lifesequencing, http://lifesequencing.com).

### *Data mining of the sRNA pool*

sRNA sequences were parsed from FASTA formatted files containing 281,346 reads from two 454 sequencing runs and assigned to specific libraries through identification of the sRNA/adapter boundaries and barcode analysis. The adapter sequences in the 454 sequencing reads were removed by using python scripts and the biopython library (http://biopython.org/). The viral genomic sequences were downloaded from NCBI (http://www.ncbi.hlm.nih.gov). vsRNA reads were mapped to the viral genomic sequences by using Blast and Vmatch (http://www.vmatch.de/). The potential RNA secondary structure was predicted using *mfold* version 3.2 (Zuker, 2003). Predictions were made using RNA sequences containing 50–200 nucleotides on either side of the vsRNA hotspot. In case no apparent local foldback structure was predicted for a given sequence, larger upstream and downstream sequences spanning the vsRNA were used for mfolding. All possible predicted variants of the same structure were analyzed, and those with the highest base pairing at the regions corresponding to the hotspots of vsRNA formation were further investigated. Criteria for recognition of candidate structured precursors are outlined elsewhere (Jones-Rhoades and Bartel, 2004; Wang et al., 2004).

### *Statistical analyses*

Data analysis was done using the SPSS 15.5 software. The individual-based rarefaction curves were computed using R 2.7.1 software (R Foundation for Statistical Computing) (http://www.R-project.org). The Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) accession number for the vsRNA sequences discussed in this paper is GSE16996.

## Acknowledgments

We are indebted to J. Seoane for their assistance with the analysis of species richness. We thank J.C. Carrington for critical reading of the manuscript. We also thank E.R. Bejarano, F. García-Arenal, J. Burgyán and S.P. Dinesh-Kumar for sharing plant and viral material. This wok was supported by PhD fellowships from the Comunidad de Madrid (to L.D.) and Ministerio de Educación y Ciencia (MEC, to D.G.I.) and by Grant BIO2006-13107 from the MEC, Spain.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2009.07.005.

## References

Akbergenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins Jr., F., Hohn, T., Pooggin, M.M., 2006. Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Res. 34, 462–471.

Allen, E., Xie, Z., Gustafson, A.M., Carrington, J.C., 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121, 207–221.

Axtell, M.J., Jan, C., Rajagopalan, R., Bartel, D.P., 2006. A two-hit trigger for siRNA biogenesis in plants. Cell 127, 565–577.

Blevins, T., Rajeswaran, R., Shivaprasad, P.V., Beknazariants, D., Si-Ammour, A., Park, H.S., Vazquez, F., Robertson, D., Meins Jr., F., Hohn, T., Pooggin, M.M., 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Res. 34, 6233–6246.

Bouche, N., Lauressergues, D., Gasciolli, V., Vaucheret, H., 2006. An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. EMBO J. 25, 3347–3356.

Brodersen, P., Voinnet, O., 2006. The diversity of RNA silencing pathways in plants. Trends Genet. 22, 268–280.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., Voinnet, O., 2008. Widespread translational inhibition by plant miRNAs and siRNAs. Science 320, 1185–1190.

Brosnan, C.A., Mitter, N., Christie, M., Smith, N.A., Waterhouse, P.M., Carroll, B.J., 2007. Nuclear gene silencing directs reception of long-distance mRNA silencing in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. 104, 14741–14746.

Chellappan, P., Vanitharani, R., Pita, J., Fauquet, C.M., 2004. Short interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts, and gene silencing targets specific viral sequences. J. Virol. 78, 7465–7477.

Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K.D., Carrington, J.C., Voinnet, O., 2006. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science 313, 68–71.

Diaz-Pendon, J.A., Li, F., Li, W.X., Ding, S.W., 2007. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in *Arabidopsis* is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. Plant Cell 19, 2053–2063.

Ding, S.W., Voinnet, O., 2007. Antiviral immunity directed by small RNAs. Cell 130, 413–426.

Donaire, L., Barajas, D., Martinez-Garcia, B., Martinez-Priego, L., Pagan, I., Llave, C., 2008. Structural and genetic requirements for the biogenesis of Tobacco rattle virus-derived small interfering RNAs. J. Virol. 82, 5167–5177.

Dunoyer, P., Voinnet, O., 2005. The complex interplay between plant viruses and host RNA-silencing pathways. Curr. Opin. Plant Biol. 8, 415–423.

Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Law, T.F., Grant, S.R., Dangl, J.L., Carrington, J.C., 2007. High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of miRNA genes. PLoS ONE 2, e219.

Fusaro, A.F., Matthew, L., Smith, N.A., Curtin, S.J., Dedic-Hagan, J., Ellacott, G.A., Watson, J.M., Wang, M.B., Brosnan, C., Carroll, B.J., Waterhouse, P.M., 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. EMBO Rep. 7, 1168–1175.

Gasciolli, V., Mallory, A.C., Bartel, D.P., Vaucheret, H., 2005. Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr. Biol. 15, 1494–1500.

Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Lett. 4, 379–391.

Herr, A.J., Molnar, A., Jones, A., Baulcombe, D.C., 2006. Defective RNA processing enhances RNA silencing and influences flowering of *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. 103, 14994–15001.

Ho, T., Pallett, D., Rusholme, R., Dalmay, T., Wang, H., 2006. A simplified method for cloning of short interfering RNAs from *Brassica juncea* infected with Turnip mosaic potyvirus and Turnip crinkle carmovirus. J. Virol. Methods 136, 217–223.

Ho, T., Wang, H., Pallett, D., Dalmay, T., 2007. Evidence for targeting common siRNA hotspots and GC preference by plant Dicer-like proteins. FEBS Lett. 581, 3267–3272.

Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D., Carrington, J.C., 2007. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting. Plant Cell 19, 926–942.

Hutvagner, G., Simard, M.J., 2008. Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell Biol. 9, 22–32.

Jones-Rhoades, M.W., Bartel, D.P., 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol. Cell 14, 787–799.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Carrington, J.C., 2007. Genome-wide profiling and analysis of *Arabidopsis* siRNAs. PLoS Biol. 5, e57.

Kim, V.N., 2008. Sorting out small RNAs. Cell 133, 25–26.

Kurihara, Y., Watanabe, Y., 2004. *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proc. Natl. Acad. Sci. U. S. A. 101, 12753–12758.

Li, F., Ding, S.W., 2006. Virus counterdefense: diverse strategies for evading the RNA-silencing immunity. Annu. Rev. Microbiol. 60, 503–531.

Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C., Green, P.J., 2005. Elucidation of the small RNA component of the transcriptome. Science 309, 1567–1569.

Martinez-Priego, L., Donaire, L., Barajas, D., Llave, C., 2008. Silencing suppressor activity of the Tobacco rattle virus-encoded 16-kDa protein and interference with endogenous small RNA-guided regulatory pathways. Virology 376, 346–356.

Maule, A., Leh, V., Lederer, C., 2002. The dialogue between viruses and hosts in compatible interactions. Curr. Opin. Plant Biol. 5, 279–284.

Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., Bowman, J.L., Cao, X., Carrington, J.C., Chen, X., Green, P.J., Griffiths-Jones, S., Jacobsen, S.E., Mallory, A.C., Martienssen, R.A., Poethig, R.S., Qi, Y., Vaucheret, H., Voinnet, O., Watanabe, Y., Weigel, D., Zhu, J.K., 2008. Criteria for annotation of plant microRNAs. Plant Cell 20, 3186–3190.

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., Chen, S., Hannon, G.J., Qi, Y., 2008. Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133, 116–127.

Mlotshwa, S., Pruss, G.J., Vance, V., 2008. Small RNAs in viral infection and host defense. Trends Plant Sci. 13, 375–382.

Moissiard, G., Voinnet, O., 2006. RNA silencing of host transcripts by Cauliflower mosaic virus requires coordinated action of the four *Arabidopsis* Dicer-like proteins. Proc. Natl. Acad. Sci. U. S. A. 103, 19593–19598.

Molnar, A., Csorba, T., Lakatos, L., Varallyay, E., Lacomme, C., Burgyan, J., 2005. Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. J. Virol. 79, 7812–7818.

Montgomery, T.A., Howell, M.D., Cuperus, J.T., Li, D., Hansen, J.E., Alexander, A.L., Chapman, E.J., Fahlgren, N., Allen, E., Carrington, J.C., 2008. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell 133, 128–141.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T.A., Vaucheret, H., 2000. *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell 101, 533–542.

Pak, J., Fire, A., 2007. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. Science 315, 241–244.

Pantaleo, V., Szittya, G., Burgyan, J., 2007. Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. J. Virol. 81, 3797–3806.

Qi, X., Bao, F.S., Xie, Z., 2009. Small RNA deep sequencing reveals role for *Arabidopsis thaliana* RNA-dependent RNA polymerases in viral siRNA biogenesis. PLoS One 4, e4971.

Qu, F., Ye, X., Morris, T.J., 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.

Rajagopalan, R., Vaucheret, H., Trejo, J., Bartel, D.P., 2006. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Dev. 20, 3407–3425.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., Bartel, D.P., 2002. MicroRNAs in plants. Genes Dev. 16, 1616–1626.

Schwach, F., Vaistij, F.E., Jones, L., Baulcombe, D.C., 2005. An RNA-dependent RNA polymerase prevents meristem invasion by Potato Virus X and is required for the activity but not the production of a systemic silencing signal. Plant Physiol. 138, 1842–1852.

Sijen, T., Steiner, F.A., Thijssen, K.L., Plasterk, R.H., 2007. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science 315, 244–247.

Sontheimer, E.J., Carthew, R.W., 2005. Silence from within: endogenous siRNAs and miRNAs. Cell 122, 9–12.

Szittya, G., Molnar, A., Silhavy, D., Hornyik, C., Burgyan, J., 2002. Short defective interfering RNAs of tombusviruses are not targeted but trigger post-transcriptional gene silencing against their helper virus. Plant Cell 14, 359–372.

Takeda, A., Iwasaki, S., Watanabe, T., Utsumi, M., Watanabe, Y., 2008. The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiol. 49, 493–500.

Vargason, J.M., Szittya, G., Burgyan, J., Tanaka Hall, T.M., 2003. Size selective recognition of siRNA by an RNA silencing suppressor. Cell 115, 799–811.

Vaucheret, H., 2008. Plant ARGONAUTES. Trends Plant Sci. 13, 350–358.

Voinnet, O., 2008. Use, tolerance and avoidance of amplified RNA silencing by plants. Trends Plant Sci. 13, 317–328.

Wang, X.J., Reyes, J.L., Chua, N.H., Gaasterland, T., 2004. Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. Genome Biol. 5, R65.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., Carrington, J.C., 2004. Genetic and functional diversification of small RNA pathways in plants. PLoS Biol 2, E104.

Yoshikawa, M., Peragine, A., Park, M.Y., Poethig, R.S., 2005. A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. Genes Dev. 19, 2164–2175.

Yu, D., Fan, B., MacFarlane, S.A., Chen, Z., 2003. Analysis of the involvement of an inducible *Arabidopsis* RNA-dependent RNA polymerase in antiviral defense. Mol. Plant-Microbe Interact. 16, 206–216.

Zhang, X., Yuan, Y.R., Pei, Y., Lin, S.S., Tuschl, T., Patel, D.J., Chua, N.H., 2006. Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis* Argonaute1 cleavage activity to counter plant defense. Genes Dev. 20, 3255–3268.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.