Small Silencing RNAs: Piecing Together a Viral Genome

Joël T. van Mierlo,[1] Koen W.R. van Cleef,[1] and Ronald P. van Rij[1,*]
[1]Department of Medical Microbiology, Radboud University Nijmegen Medical Centre, Nijmegen Centre for Molecular Life Sciences, Nijmegen Institute for Infection, Inflammation and Immunity, 6500 HB Nijmegen, The Netherlands
*Correspondence: r.vanrij@ncmls.ru.nl
DOI 10.1016/j.chom.2010.02.001

Virus-derived small interfering RNAs (siRNAs) are the hallmark of RNAi-based antiviral immunity. Wu and colleagues demonstrate how viral genomes can be assembled from these small RNA sequences. Their results provide an approach for virus discovery as well as important insights into how these siRNAs mediate antiviral defense.

Viral double-stranded RNA (dsRNA) is a potent activator of antiviral immunity. In vertebrates, it triggers a type I interferon-based innate immune response; in plants and invertebrates, it triggers an alternative antiviral immune response: RNA interference (RNAi) or RNA silencing (Ding and Voinnet, 2007; van Rij and Berezikov, 2009). This defense is initiated by the cleavage of viral dsRNA into virus-derived siRNAs (v-siRNAs) by a Dicer family member. Once bound to an Argonaute (Ago) protein in the RNA-induced silencing complex (RISC), v-siRNAs guide recognition and cleavage of complementary viral target RNAs by Ago, thereby restricting viral replication. v-siRNAs are thus specificity determinants of an antiviral effector complex.

For over a decade it was known that RNAi mediates antiviral defense in plants. In contrast, in insects such as fruit flies (*Drosophila melanogaster*) and mosquitoes, an antiviral activity of RNAi was demonstrated only in recent years. Accordingly, many aspects of antiviral RNAi in insects remain poorly understood. One of the major open questions regards the biogenesis of v-siRNAs. The available data in plants imply several potential viral substrates for Dicer: dsRNA replication intermediates of positive-strand (+) RNA viruses, structured RNA elements in single-stranded viral RNA, or overlapping convergent transcripts of DNA viruses (Figure 1A). With the increasing accessibility of massive parallel sequencing technology, we anticipate rapid progress in this field. Initial reports indicate that dsRNA replication intermediates are the major source of v-siRNA in infections with three distinct (+) RNA viruses: Flock house virus in *Drosophila* and Sindbis and West Nile viruses in *Aedes aegypti* and *Culex pipiens* mosquitoes, respectively (Aliyari et al., 2008; Brackney et al., 2009; Flynt et al., 2009; Myles et al., 2008; van Rij and Berezikov, 2009).

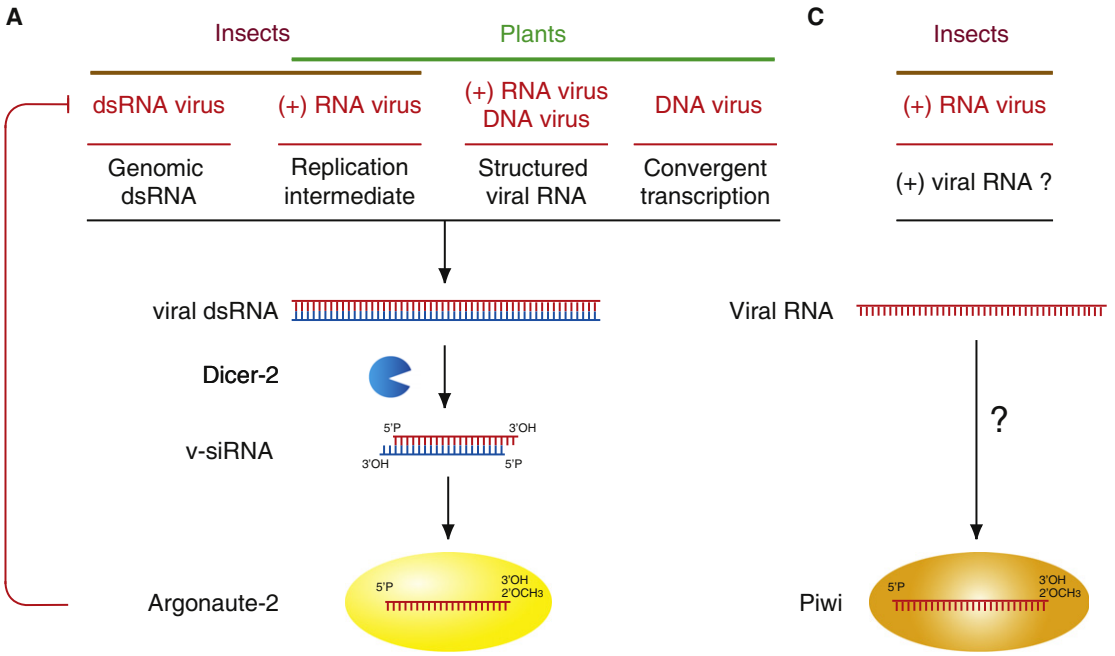
Having noted that v-siRNAs are often overlapping, two groups—one working on plant viruses (Kreuze et al., 2009), the other studying insect viruses (Wu et al., 2010)—have recently developed an approach for virus discovery that deduces viral genomic sequences from the enormous amounts of sequence data generated by deep sequencing of siRNA libraries. In a relatively straightforward yet powerful approach, both groups created long contiguous sequences using software specifically developed to assemble short reads and used BLAST searches to explore the origin of these sequences.

Both groups first validated the approach, one group using experimental infections of *Drosophila*, *C. elegans* nematodes, and mosquitoes (Wu et al., 2010), and one group using sweet potato plants (Kreuze et al., 2009), in some cases recovering complete viral genomes with high coverage. More striking, however, was the identification of unexpected or even novel viruses from unsuspected sources: seemingly healthy cell lines and symptomless mosquitoes and plants. For example, a subclone of the frequently used *Drosophila* S2 cell line is riddled with viruses; the authors recovered sequences from five distinct viruses belonging to four different virus families! Together, these papers identified seven previously unknown viruses from five distinct families (Figure 1B). For experimental biologists, these results are an important reminder to consider persistent virus infections as a source of irreproducibility and heterogeneity among different cell lines.

siRNA sequencing and assembly is an extremely powerful tool for virus discovery that does not require specific amplification or enrichment of viral nucleic acids and that is independent of the ability to culture the virus in vitro. Another advantage of the approach is that it enriches for small RNAs of viral origin, as it taps into a natural antiviral defense mechanism. Its dependence on the detection of significant similarity to known viruses in BLAST analyses might be a disadvantage of the technique. Another potential caveat is the lack of formal proof demonstrating that RNAi targets negative-strand (−) RNA viruses and DNA viruses in insects. Thus far, small RNA sequencing failed to identify significant amounts of v-siRNAs in vertebrate infections; therefore, the approach is unlikely to be successful in vertebrates.

An important application is the identification of novel viruses from hematophagous arthropods, such as mosquitoes and ticks, which have the potential to transmit pathogens to humans. In recent years, several novel arthropod-borne viruses were isolated from wild-caught mosquitoes, which included a mosquito genus that was not known to transmit arboviruses. With the reducing costs for massive parallel sequencing and the extreme sequence depth that allows multiplexed analyses, the approach may be used to comprehensively map the viral reservoir in relevant mosquito vectors.

The results of Wu et al. (2010) also provide important insights in v-siRNA biogenesis (Figure 1A). In (+) RNA virus



B

Viruses recovered through small RNA sequencing and assembly in insects.

Host	Virus	Family	Genome
Drosophila	Drosophila A virus	Tetraviridae	(+) RNA
	Drosophila C virus	Dicistroviridae	(+) RNA
	Noravirus	Unassigned	(+) RNA
	American Nodavirus	Nodaviridae	Bipartite (+) RNA
	Flock House Virus	Nodaviridae	Bipartite (+) RNA
	Drosophila Totivirus	Totiviridae	dsRNA
	Drosophila X virus	Birnaviridae	Bipartite dsRNA
	Drosophila Birnavirus	Birnaviridae	Bipartite dsRNA
Aedes aegypti	Mosquito Nodavirus	Nodaviridae	(+) RNA
	Sindbisvirus	Alphaviridae	(+) RNA

Novel Virus
Experimental infection

Figure 1. Simplified Scheme of an Antiviral RNAi Response

(A) An RNase III enzyme of the Dicer family processes viral dsRNA into viral siRNAs (v-siRNAs), which guide an Argonaute protein in the RNA-induced silencing complex (RISC) to viral target RNAs. Dicer-2 and Argonaute-2 are responsible for antiviral RNAi in insects.

(B) Overview of viruses that were recovered by small RNA sequencing and assembly in insects. Data are derived from *Drosophila* cell lines and adult *Aedes aegypti* mosquitoes. Red shading indicates the viruses that were used for experimental infections. Yellow shading indicates previously undescribed viruses. Since the novel *Drosophila* viruses were detected in cell lines, it is yet to be established whether these are natural fly pathogens.

(C) In a *Drosophila* cell line derived from the ovarian somatic sheet, Wu et al. detected viral piRNAs that corresponded to the (+) strand of two (+) RNA viruses, *Drosophila* C virus and American Nodavirus. How these viral piRNAs are generated and their role in antiviral defense is not known.

infection, viral (+) RNAs greatly outnumber (−) RNAs. A predominance of (+) v-siRNA is therefore expected if RNA structures are the major source of v-siRNA, as was observed for some plant viruses (Ding and Voinnet, 2007). However, in all four (+) RNA viruses analyzed, v-siRNAs mapped in roughly equal proportion to (+) and (−) viral RNA strands, in agreement with previous observations for three other (+) RNA viruses (Aliyari et al., 2008; Brackney et al., 2009; Flynt et al., 2009; Myles et al., 2008; van Rij and Berezikov, 2009). These results thus indicate that viral replication intermediates are the main substrates for Dicer-2 in insects. v-siRNAs generally mapped across the entire viral genome. One case of nonuniform distribution of v-siRNAs could be attributed to defective interfering RNA in persistent infections, as was proposed earlier (van Rij and Berezikov, 2009). Wu et al. also report v-siRNA profiles from another class of viruses; in three distinct dsRNA viruses, v-siRNAs mapped in proportions similar to (+) and (−) viral RNA strands, suggesting that in these infections, the viral genomic dsRNA is cleaved by Dicer-2.

Many viruses hide or sequester dsRNA to prevent activation of innate immune responses. Genomic viral dsRNA of *Birnaviridae*, for example, remains encapsidated in the viral capsid throughout the replication cycle. Likewise, (+) RNA viruses shield their replication intermediates in membrane vesicles. Nevertheless, these v-siRNA profiles imply that there is a window of opportunity, in space or time, in which the viral dsRNA is accessible to Dicer.

Perhaps the most tantalizing result of Wu et al. is the detection of viral PIWI-interacting RNAs (piRNAs) in an ovarian somatic sheet (OSS) cell line (Figure 1C). piRNAs differ from siRNAs in their size (~26–30 nt for piRNA; 21 nt for siRNAs), their Dicer-independent biogenesis, and their association with the PIWI subclass of the Ago family, which consists of Piwi, Ago-3, and Aubergine (Aub) (Ghildiyal and Zamore, 2009). PIWI proteins and their associated piRNAs are restricted to the fly's germline tissues, where they protect the genome from activation of transposons. The biogenesis of piRNAs is not fully understood. A pool of primary piRNAs arises from specialized genomic loci and engages in an amplification loop that involves Ago-3 and Aub in the presence of transposon target RNAs. This amplification loop gives rise to the signature of piRNAs: a strong 5′ uridine bias for Piwi and Aub-associated piRNAs and a bias for adenine at the tenth position for Ago-3-associated piRNAs. In contrast to Ago-3 and Aub, Piwi also resides in somatic cells of germline tissues of the fly, where it associates with piRNAs that do not engage in ping-pong amplification, but do present themselves with a 5′ uridine bias. Their biogenesis is not understood (Ghildiyal and Zamore, 2009).

The *Drosophila* OSS cell line resembles ovarian somatic cells in Piwi expression and piRNA profiles. Similar to the S2 cell line, the OSS cell line is riddled with (six distinct!) viruses. Two of these, *Drosophila* C virus (DCV) and American Nodavirus (ANV), seem to generate viral piRNAs. The size distribution of small RNAs, a sharp peak of 21 nt siRNAs and a distinct, broader peak of ~26–30 nt piRNAs with a 5′ uridine bias, implies that DCV and ANV are targets for both RNAi and piRNA pathways. Strikingly, two other (+) RNA viruses and two dsRNA viruses did not generate great numbers of piRNAs. Intrinsic to their biogenesis, transposon piRNAs have a strong strand bias. Viral piRNAs are predominantly (+) stranded, which likely reflects the higher abundance of (+) over (−) viral RNA.

Germline transmission is one of the most effective strategies for stable persistence of viruses in a host species. For the host, germline integrity is essential for proper development of offspring. The detection of viral piRNAs suggests that the (somatic) piRNA pathway protects the germline from invasion by viruses. The somatic follicle cells that surround the oocyte are perfectly positioned to perform such a task. An antiviral function of the piRNA pathway might explain the inability of DCV to invade the germline, despite its tropism for the epithelial sheath that surrounds the egg chamber. Many more questions remain. For example, what is the mechanism of viral piRNA biogenesis? How does the piRNA machinery discriminate viral RNA from abundant cellular mRNA? The paper by Wu et al. sets the stage for experimental dissection of the antiviral piRNA pathway, which may also provide insight into the biogenesis of primary transposon piRNAs. Finally, future studies will have to elucidate whether we can add the piRNA pathway to the arsenal of antiviral defenses in insects.

ACKNOWLEDGMENTS

We thank Marius van den Beek, Walter Bronkhorst, and Frank van Kuppeveld for discussions and Annemieke Jansens for critical reading of the manuscript. This work was financially supported by a fellowship from the Nijmegen Centre for Molecular Life Sciences, by a VIDI fellowship from the Netherlands Organization for Scientific Research (project number 864.08.003), and by Horizon Breakthrough fellowships from the Netherlands Genomics Initiative (project numbers 93519018 and 93518020).

REFERENCES

Aliyari, R., Wu, Q., Li, H.W., Wang, X.H., Li, F., Green, L.D., Han, C.S., Li, W.X., and Ding, S.W. (2008). Cell Host Microbe *4*, 387–397.

Brackney, D.E., Beane, J.E., and Ebel, G.D. (2009). PLoS Pathog. *5*, e1000502.

Ding, S.W., and Voinnet, O. (2007). Cell *130*, 413–426.

Flynt, A., Liu, N., Martin, R., and Lai, E.C. (2009). Proc. Natl. Acad. Sci. USA *106*, 5270–5275.

Ghildiyal, M., and Zamore, P.D. (2009). Nat. Rev. Genet. *10*, 94–108.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., and Simon, R. (2009). Virology *388*, 1–7.

Myles, K.M., Wiley, M.R., Morazzani, E.M., and Adelman, Z.N. (2008). Proc. Natl. Acad. Sci. USA *105*, 19938–19943.

van Rij, R.P., and Berezikov, E. (2009). Trends Microbiol. *17*, 163–171.

Wu, Q., Luo, Y., Lu, R., Lau, N., Lai, E.C., Li, W.-X., and Ding, S.-W. (2010). Proc. Natl. Acad. Sci. USA *107*, 1606–1611.