Review

Small RNAs and the control of transposons and viruses in *Drosophila*

Ronald P. van Rij[1] and Eugene Berezikov[2]

[1] Department of Medical Microbiology, Radboud University Nijmegen Medical Centre, Nijmegen Centre for Molecular Life Sciences, PO Box 9101, 6500 HB Nijmegen, The Netherlands
[2] Hubrecht Institute and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

RNA interference (RNAi) – post-transcriptional gene silencing guided by small interfering RNA (siRNA) – is an important antiviral defense mechanism in insects and plants. Several recent studies in *Drosophila* identified endogenous siRNAs corresponding to transposons, to structured cellular transcripts and to overlapping convergent transcripts. In addition, one of these studies detected a large pool of Argonaute-2 associated siRNAs that mapped to the genome of flock house virus, a (+) RNA virus. Our bioinformatic analyses indicate that these viral siRNAs mapped in roughly equal proportions to both (+) and (-) viral RNA strands. These reports attribute an important function to RNAi in the defense against parasitic nucleic acids (viruses and transposable elements) and provide a novel mechanism for RNAi-based regulation of cellular gene expression. Furthermore, the detection of viral siRNAs of both (+) and (-) polarity implicates double-stranded RNA replication intermediates as the Dicer substrates that mediate antiviral defense.

RNA silencing pathways and antiviral defense

Invertebrates lack the innate and adaptive immune systems that mediate antiviral defense in mammals, yet they are capable of effectively fighting viral infections. RNA interference (RNAi) is an important antiviral defense mechanism in insects and plants (Figure 1) [1–4]. In *Drosophila*, viral double stranded RNA (dsRNA) is processed by the ribonuclease Dicer-2 (Dcr-2) (See Glossary) into viral small interfering RNAs (v-siRNAs), which are incorporated into the RNA induced silencing complex (RISC) and guide RISC onto mRNAs in a sequence-specific manner. Recognition of a complementary sequence triggers the endonucleic cleavage of the viral target RNA (slicer

Glossary

Argonaute: core component of RNA induced silencing complex, characterized by conserved PAZ and PIWI domains.
Cricket paralysis virus (CrPV): (+) RNA virus from the Dicistroviridae family (genus: Cripavirus) that infects invertebrate hosts.
Drosophila C virus (DCV): (+) RNA virus from the Dicistroviridae family (genus: Cripavirus) and a natural pathogen of *Drosophila melanogaster*.
Defective interfering (DI) RNAs: viral RNAs with large deletions of the viral genes that are required for replication and encapsidation; these RNAs are, therefore, completely dependent on a helper virus for replication. DI RNAs arise during replication of the parental virus through one or more premature termination and reinitiation events. *Cis*-acting RNA elements required for replication are retained in DI RNA, allowing for replication by the RdRP of the parental helper virus.
Dicer: ribonuclease of the RNase III family that generates small RNAs in the RNAi and miRNA pathways. Dicer acts like a molecular ruler: its PAZ domain binds the 3′ two-nucleotide overhang, and the length of the cleavage product is determined by the distance of the PAZ domain to the RNase III domains.
DNA transposon: mobile genetic element that moves directly from one position to another within the genome through a 'cut and paste' mechanism by a transposase enzyme (class II transposon).
flock house virus (FHV): (+) RNA virus with a bipartite genome from the Nodaviridae family (genus: Alphanodavirus) that infects invertebrate hosts.
Geminiviruses: plant viruses with a circular ssDNA genome that replicates in the nucleus via a double stranded DNA intermediate.
Loquacious (Loqs): protein with three dsRNA binding domains that partners with Dicer-1 for pre-miRNA processing and Ago-1 RISC loading.
LTR retrotransposons: see retrotransposons.
miRNA: small RNAs (21-23 nt in length) that are encoded in the genome as long primary transcripts, which are sequentially processed by the ribonucleases Drosha (in the nucleus) and Dicer-1 (in the cytoplasm) into mature miRNAs. miRNAs usually interact through imperfect basepairing with their target mRNAs and direct translational inhibition, but can also trigger degradation of the target mRNA.
Pararetrovirus: a nuclear dsDNA virus that uses reverse transcription for replication, but does not integrate in the genome of the host.
PAZ domain: nucleotide binding domain found in Dicer and Argonaute gene families that binds the 3′ two-nucleotide overhang of duplex siRNA or miRNA.
Phasing of siRNA: 21-nt periodicity of 5′ ends of siRNAs because of progressive dsRNA cleavage by Dicer from a defined start site.
Piwi-interacting RNA (piRNA): small RNAs of ~25-30 nt that associate with the Piwi class of Argonaute proteins. piRNAs seem to be exclusively expressed in somatic and germline cells of the *Drosophila* male and female germline, in which they are involved in control of transposon activity. piRNAs have a strong strand bias; those that associate with Piwi and Aubergine are predominantly antisense to transposons, those that associate with Ago-3 are predominantly sense. The biogenesis of piRNAs is independent from Dicer enzymes.
Piwi domain: protein domain that resembles the structure of RNase H. The piwi domain of Argonaute 2 is responsible for target mRNA cleavage. The piwi domain of Argonaute-1 lacks endonuclease activity.
R2D2: protein with two dsRNA binding domains that partners with Dicer-2 for dsRNA processing and RISC loading.
RNA dependent RNA polymerase (RdRP): enzyme that catalyzes the replication of RNA from an RNA template.
Retrotransposons: mobile genetic elements that move in the genome through a 'copy and paste' mechanism by reverse transcription (class I transposon). After transcription, transposon RNA is reverse transcribed into DNA, which is integrated at other locations in the genome. Retrotransposons consist of two sub-types: LTR transposons that contain long terminal repeats at each end, and non-LTR retrotransposons.
RNase H: Endonuclease with specificity for RNA in an RNA-DNA hybrid.
RNase III: Ribonuclease that cleaves dsRNA and usually generates 3′ two-nucleotide overhangs.
Stemloop: RNA (or DNA) hairpin structure formed through intramolecular base pairing, consisting of a double helix that ends in an unpaired loop.
Subgenomic RNA: viral RNA segment that is colinear with the 3′ proximal region of the genomic RNA of (+) RNA viruses, allowing expression of open reading frames (ORF) downstream or overlapping with upstream ORFs. Viral (-) RNA serves as a template for subgenomic RNA synthesis.
Transposon: DNA element that can insert itself at other non-homologous regions in the genome of a single cell. Transposition can cause mutations, deletions, duplications and changes in gene expression at the site of insertion or in nearby genes.
Untranslated regions (UTRs): the 5′ or 3′ terminal regions of an mRNA that are not translated into protein.

Corresponding author: van Rij, R.P. (r.vanrij@ncmls.ru.nl).

0966-842X/$ – see front matter doi:10.1016/j.tim.2009.01.003 Available online 18 March 2009

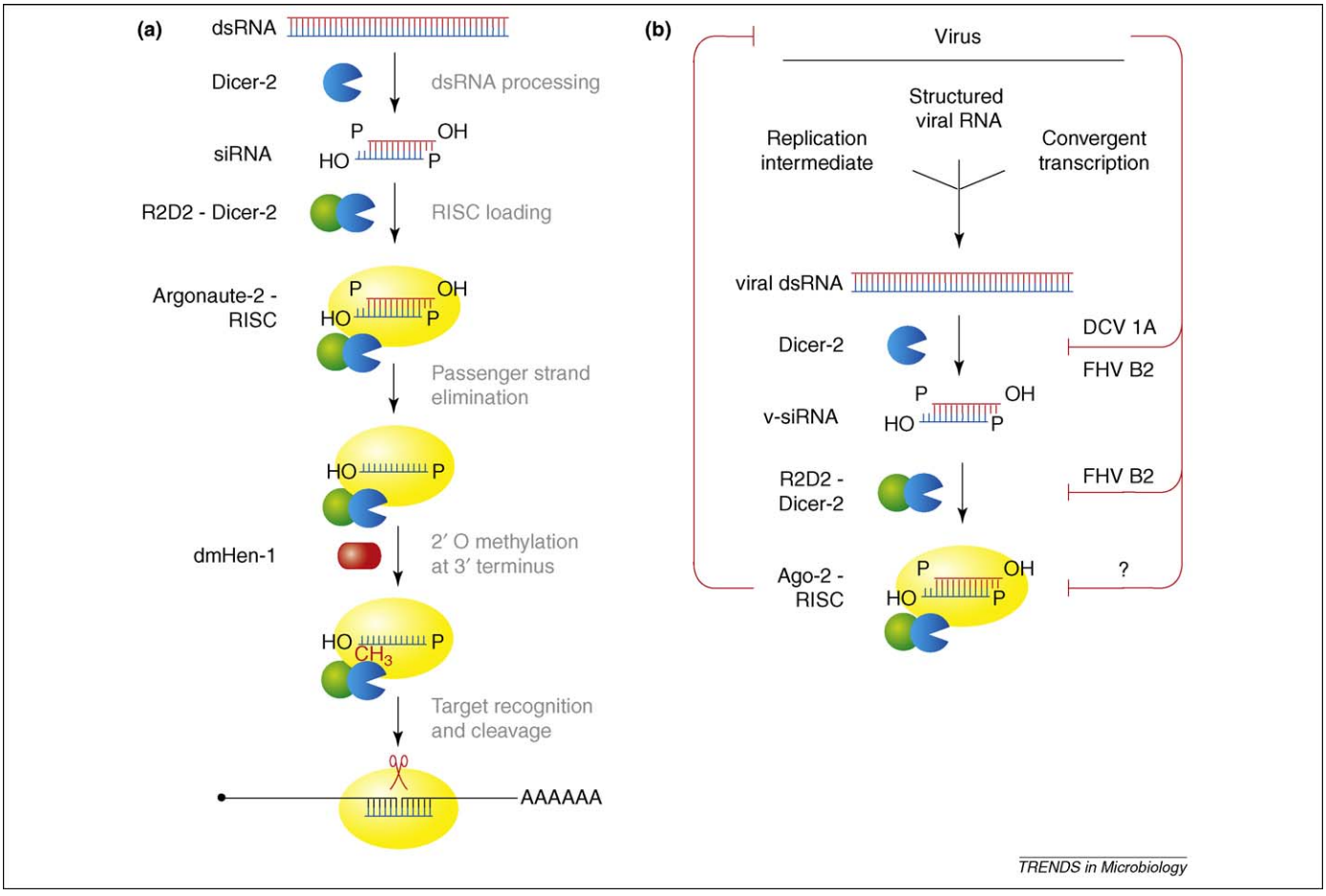

Figure 1. Mechanism and antiviral activity of the RNAi pathway in *Drosophila melanogaster*. **(a)** The ribonuclease Dicer-2 initiates the RNAi pathway by processing long dsRNA into 21-nt small interfering (si) RNA duplexes with 5′ phosphate groups and 3′ two-nt overhangs. The RISC loading complex, consisting of Dicer-2 and the dsRNA binding protein R2D2, delivers siRNAs into the RNA induced silencing complex (RISC), in which the RNA is bound by RISC's core catalytic component, Argonaute-2 (Ago-2). RISC is a multi-protein complex that, in addition to Argonaute-2, contains auxiliary factors such as tudor-staphylococcal nuclease (TSN) and vasa intronic gene (VIG) and the *Drosophila* orthologue of fragile-X mental retardation protein (FXR, also known as FMR1). One of the strands of the siRNA (passenger strand) is eliminated, whereas the other strand (the guide strand) is retained in RISC and modified with a 2′-O-methyl group at the 3′ terminus by the RNA methylase dmHen1 [17,18]. The guide strand directs RISC activity onto mRNAs in a sequence specific manner, and, upon recognition of a fully complementary sequence, triggers cleavage of the target RNA by the RNase H-like piwi domain in Argonaute-2 (slicer activity). See Refs [15,16] for more detailed descriptions of the biochemistry of RNAi. **(b)** In RNAi-mediated antiviral defense, viral dsRNA is processed into viral-siRNAs (v-siRNA) in a *Dicer-2–R2D2* dependent manner. These v-siRNAs are incorporated into Ago-2 RISC, which then targets viral RNA for degradation. Viral suppressors of RNAi (VSR) have evolved as a counter-defense, which can interfere with different steps of the RNAi pathway. See Box 1 for more details. Viral dsRNA replication intermediates are the predominant source of v-siRNAs in FHV infection (see Ref. [25] and our analyses in Figure 3). Adapted, with permission, from Ref. [38].

activity) by Argonaute-2 (Ago-2), the core catalytic component of RISC (Figure 1b and Box 1). Accordingly, *Drosophila melanogaster Dcr-2* and *Ago-2* mutants are hypersensitive to viral infection, resulting in increased mortality because of uncontrolled viral replication. As a counter-defense, viruses encode viral suppressors of RNAi (VSR) that allow them to replicate in the face of this antiviral system. For example, protein B2, the VSR from flock house virus (FHV), is essential for viral replication in wildtype flies, but dispensable for replication in RNAi mutants [2,3].

The *Drosophila Dicer* and *Argonaute* gene families diversified and acquired specialized functions during the evolution of three related pathways. First, the microRNA (miRNA) pathway – post transcriptional gene silencing mediated by small RNAs encoded in the genome – relies on Dcr-1 and its dsRNA binding partner Loquacious (Loqs) for miRNA biogenesis. Mature miRNAs are loaded into an Ago-1 containing RISC complex (miRISC) that is responsible for target recognition and effector function. Second, the RNAi pathway depends on Dcr-2 and its dsRNA binding cofactor R2D2 for dsRNA processing and siRNA loading into RISC, and on the slicer Ago-2 for effector function. Finally, the piwi-interacting RNA (piRNA) pathway is mediated by the piwi subfamily of the Argonaute family (consisting of *Ago 3*, *Piwi* and *Aubergine*) and controls transposable elements in the germline, guided by a specific class of endogenous small RNAs (piRNAs) [5,6].

Endogenous small RNAs thus instruct the miRNA and piRNA pathways, whereas the RNAi pathway seems to depend on exogenous siRNA sources (viruses or experimental dsRNA). It had, however, been unclear whether the RNAi pathway is solely dedicated to antiviral defense, or whether RNAi has additional regulatory activities in the cell. Several studies have now shown that Ago-2 is associated with a population of endogenous siRNAs (endo-siRNAs) in *Drosophila* [7–12] and in mouse oocytes [13,14]. These endo-siRNAs map to transposons and to specific cellular transcripts, implicating RNAi in the control of transposon activity and in the regulation of cellular gene expression. One study additionally detected a large pool of viral sequences because of a persistent infection of the

Box 1. RNAi as an antiviral defense mechanism: defense and counter-defense

RNAi is an important antiviral defense mechanism in plants and insects [4] (see Refs [40–43] for discussions on the mammalian system). Several lines of evidence support the antiviral function of RNAi in insects. *Dicer-2*, *R2D2* and *Ago-2* fly mutants are hypersensitive to infection by the (+) RNA viruses FHV, Drosophila C virus (DCV) and Cricket paralysis virus (CrPV) [1–3]. These mutants are unable to efficiently control virus replication, resulting in a dramatic increase in mortality. The antiviral function of RNAi is not limited to Drosophilid insects; depletion of *Ago-2* expression in *Anopheles gambiae* and *Aedes aegyptii* mosquitoes results in increased replication of, respectively, O'nyong nyong virus and Sindbis virus [44,45]. Detection of v-siRNA provides direct support for viral dsRNA cleavage by Dicer-2 in FHV and CrPV infections of adult *Drosophila* and in Alphavirus infections of several other insect species [2,3,45,46]. The dsRNA virus Drosophila X virus (DXV) is also under control of an antiviral RNAi response. Mutants for *Ago-2* and the Dicer-2 cofactor *R2D2* are hypersensitive to DXV infection, yet, paradoxically, *Dicer-2* mutants are not [47]. Whether *Drosophila* RNAi controls DNA virus replication has yet to be demonstrated. Of note, plant DNA viruses are controlled by RNAi; geminiviruses produce dsRNA through convergent transcription from opposite promoters [48], whereas a structured RNA leader element is the main source of v-siRNAs in the pararetrovirus Cauliflower mosaic virus [49].

Viral suppressors of RNAi (VSR) allow viruses to replicate in the presence of the antiviral activity of RNAi (Figure 1b in main text). The DCV 1A protein efficiently binds long dsRNA, preventing the production of siRNAs by Dicer-2 [1]. FHV B2 can bind both long dsRNA and siRNAs, and could thus inhibit RNAi via two distinct mechanisms: inhibition of Dicer-2 cleavage and sequestration of v-siRNAs, thereby preventing their incorporation into RISC [50–53]. Indeed, *Aliyari et al.* [25] recently showed that B2 binds viral dsRNA in infected cells, and, consequently, that v-siRNAs accumulate to high levels after infection with FHV lacking B2. They also provided support for a v-siRNA binding activity of B2 in infected cells. CrPV, the closest relative of DCV in the dicistrovirus family, also suppresses RNAi [3]; yet, dsRNA processing proceeds normally in extracts from CrPV infected cells [1]. CrPV might inhibit RNAi at a step downstream of Dicer-2, perhaps at the level of RISC activity, as was previously reported for the plant Cucumber mosaic virus 2b protein [54].

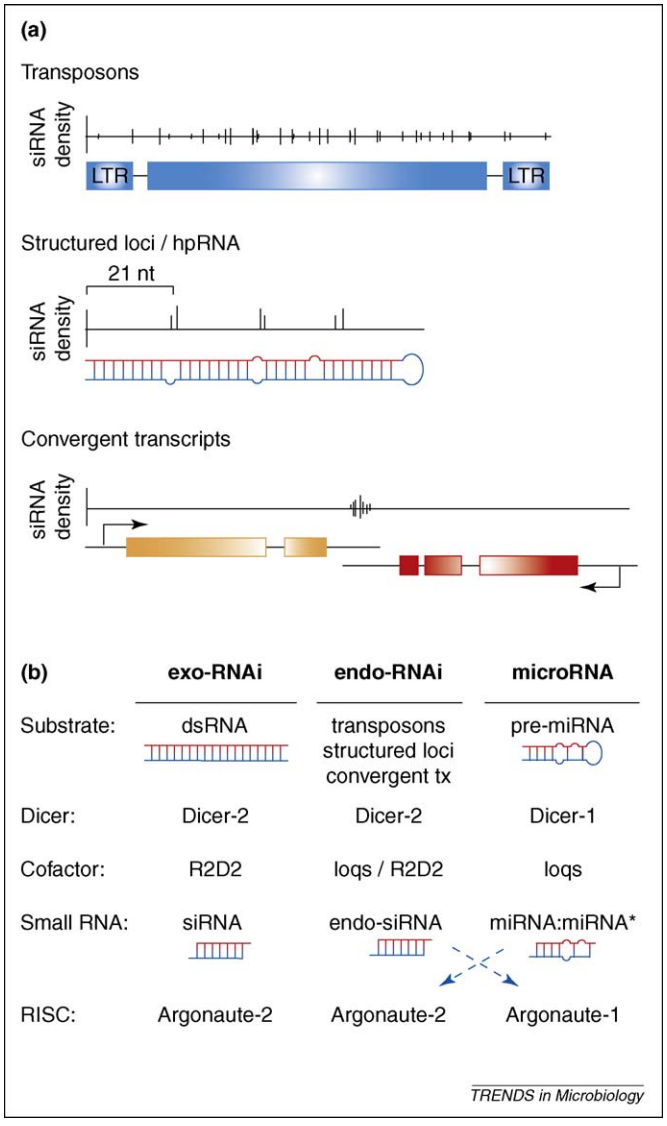

Figure 2. Biogenesis of endogenous siRNAs. **(a)** Schematic representation of three groups of endo-siRNA generating loci and their associated siRNA densities in *Drosophila*. Top panel: long terminal repeat (LTR) containing retrotransposons are the predominant source of endo-siRNAs. Transposon endo-siRNAs map across the entire transposon sequence without a bias for sense or antisense polarity. Middle panel: structured loci generate transcripts with long inverted repeats that have the potential to generate hairpin RNA (hpRNA) structures through intramolecular basepairing. hpRNA-derived endo-siRNAs arise from one genomic strand and show a phased pattern of 5′ Dicer-2 cleavage sites. Phasing of siRNA – 21-nt periodicity of 5′ ends of siRNAs because of progressive dsRNA cleavage from a defined start site – is not observed for other categories of endo-siRNAs. Bottom panel: many convergent transcription units in *Drosophila* overlap in their 3′ untranslated regions. These loci generate siRNAs from both genomic strands in the overlapping region, with a peak abundance at the center of the overlap. Endo-siRNA densities are represented as 5′ Dicer cleavage sites in the diagrams above the figures, the height of the bar correlating with the cloning frequency. Bars above the *x*-axis represent sense siRNAs; bars below the *x*-axis represent antisense siRNAs. Figure is not drawn to scale. **(b)** Schematic representation of the genetic requirements, Dicer substrates and small RNAs associated with exogenous RNAi (virus infection or experimental RNAi), endo-RNAi and miRNA pathways. Note that there can be some blurring between these pathways (dashed arrows). Some endo-siRNAs are loaded into Argonaute-1 [7], whereas some miRNAs, especially those with an extended dsRNA character, can load into Argonaute-2 in an *R2D2* dependent manner [39]. Abbreviation: convergent tx, convergent transcripts.

characterized cell line, providing the first in-depth view of the small RNA profile associated with antiviral defense in animals [7]. These reports attribute an important function to RNAi in the defense against parasitic nucleic acids (viruses and transposable elements) and shed light on the associated small RNAs.

Endogenous RNAi: transposons and cellular transcripts

Drosophila Dcr-2 cleaves dsRNA into specific 21 nt siRNAs that are loaded into Ago-2 and modified with a 2'-O-methyl group at their 3′ termini (Figure 1a) [15–18]. By deep sequencing small RNAs that were physically associated with Ago-2 or contained a 3′ methyl group, or by computational filtering for a 21-nt size, several groups identified endo-siRNAs [7–12]. Endo-siRNAs fall into three different categories: those derived from transposons, from structured cellular transcripts and from overlapping regions of convergent transcripts (Figure 2a).

Transposons

A large fraction of endo-siRNAs map to transposable elements, with numbers ranging from ~30% in fly heads, to 53% in ovaries and even up to 86% in the S2 cell line [7–9,11]. Endo-siRNAs map to three classes of transposons in *Drosophila*: long terminal repeat (LTR) retrotransposons, non-LTR retrotransposons and DNA transposons. However, LTR retrotransposons predominated the endo-siRNA

pool, also after correction for their abundance in the fly genome [8].

In general, endo-siRNA mapped across the entire transposon, without obvious hotspots or enrichment for sense or antisense polarity. A general mechanism for the biogenesis of transposon siRNAs is not immediately apparent. Possibly, two transposons are integrated as an inverted repeat in the genome, generating transcripts that can basepair to form Dicer substrates. Alternatively, a transposon could be integrated opposite to a cellular promoter that thus produces transcripts that are complementary to the transposon transcripts [19]. The terminal inverted repeats (TIR) of the *Tc1* DNA transposon in *C. elegans* were suggested to generate dsRNA through intra-molecular base-pairing [20]. Yet, the observation that siRNAs map across the *S-element* DNA transposon in *Drosophila*, without enrichment at the TIRs [8], indicates that other mechanisms for siRNA production are at work here.

Do transposon-derived siRNAs then contribute to control of transposon activity? Indeed, many endo-siRNA generating transposons are derepressed in somatic and germline tissues of RNAi mutant flies and in cell culture after knockdown of *Ago-2* or *Dcr-2* (Table 1) [7–9,11,21]. Perhaps not surprising, some of the most active transposons produce the most abundant endo-siRNAs. The LTR-containing *297*, *1731* and *mdg1* elements, for example, together contribute ~62% of the total endo-siRNA pool in S2 cells. Steady-state RNA levels of these elements increased five to eightfold upon *Dcr-2* knockdown [8]. Derepression of many other endo-siRNA generating transposons was more modest (Table 1).

The piRNA pathway was previously reported as an important pathway to silence retrotransposons in the germline [6]. Endo-siRNA clusters can overlap with piRNA clusters; the *Mdg-1* and *Stalker4* LTR retrotransposons and the non-LTR *F element*, for example, give rise to piRNA production [22], but are also the three most abundant sources of endo-siRNAs in ovaries [7]. Consequently, steady-state RNA levels of *Stalker4* and *F elements* were approximately two to threefold higher in *Ago-2* and *Dcr-2* mutants than in control flies [7]. No or only a mild derepression (~1.5-fold) of *Mdg-1* activity was observed in *R2D2*, *Dcr-2* and *Ago-2* mutant ovaries [6,7]. In striking contrast, *Mdg-1* RNA levels were almost 30-fold higher in *piwi* mutant ovaries [6]. Thus, both piRNA and endo-siRNA pathways contribute to transposon silencing in the germline, whereas endo-siRNAs control transposon activity in somatic tissues. The male sterility that is observed in *piwi*, but not *Dicer-2* or *Ago-2*, fly mutants implies that the piRNA pathway is the dominant mechanism to control transposons in the germline.

Structured loci

A remarkable novel source of endo-siRNAs are structured loci that can form extensive dsRNA structures through intramolecular basepairing (also termed hairpin RNA [hpRNA]) [7,9,12]. Two loci were independently identified and characterized by several groups [7,9,12], but several other loci also give rise to endo-siRNAs [12]. For example, *esi-1* can form a ~400 bp dsRNA hairpin structure, separated by a large 555-nt loop, through intramolecular basepairing of inverted repeat sequences. The *esi-2* locus consists of 20 palindromic ~260 nt repeats that have multiple possibilities to form dsRNA structures.

A possible function of endo-siRNAs from structured loci is the regulation of host gene expression through RNAi. The detection of Slicer products from the DNA damage response gene *mus308* in a region with extensive complementarity to a highly abundant *esi-2* siRNA provided proof of principle for such an activity. Accordingly, a (modest) increase in *mus308* expression was observed in RNAi mutants flies [7,12].

Convergently transcribed loci

Approximately one thousand gene pairs in the fly genome are positioned in opposite orientation and generate transcripts that overlap in their 3′ untranslated regions (UTRs) [7]. Overlapping genomic regions generate Dicer substrates if both transcripts meet and basepair to form dsRNA. Indeed, endo-siRNAs map to the overlapping genomic regions, but not to adjacent non-overlapping regions of the same genes. Whether these endo-siRNAs contribute to regulation of gene expression remains unclear. Little or no increase in expression of these endo-siRNA producing genes was observed in RNAi mutants [7,10].

Dcr-2 and R2D2 mediate siRNA biogenesis and Ago-2 loading in the canonical RNAi pathway (Figure 2b). Endo-siRNA biogenesis is indeed dependent on *Dcr-2*, but there seems to be a requirement for *loqs* for the biogenesis of structured loci endo-siRNAs [7,12]. This dependence, however, is not absolute; a specific *esi-1* endo-siRNA with an extensive dsRNA character shows a partial dependence on *R2D2*. Endo-siRNAs derived from transposons and con-

Table 1. Derepression of transposon activity in *Drosophila* RNAi mutants or in S2 cells after depletion of *Dcr-2* or *Ago-2*

Material	Derepressed transposons (%)[a]	Number analyzed	Fold repression (range)	Type	Refs
Ago-2 heads	100	5	2 - 5	LTR, non-LTR	Chung *et al.* [11]
Dcr-2 heads	40	5	3.5	LTR	
S2 cells, *Ago-2* k/d[c]	100	6	2 - 5	LTR, non-LTR	
S2 cells, *Dcr-2* k/d[c]	100	6	3 - 9	LTR, non-LTR	
Ago-2 ovaries	100	12	1.5 - 9	LTR, non-LTR	Czech *et al.* [7]
Ago-2 heads	46[b]	13	1.5 - 8.5	LTR, non-LTR	Ghildiyal *et al.* [8]
Dcr-2 heads	15[b]	13	2 - 7	LTR	
S2 cells, *Dcr-2* k/d[c]	88	8	1.5 - 8	LTR, non-LTR	
Dcr-2 carcasses without testes	71	7	1.5 - 7	LTR	Kawamura *et al.* [9]

[a]Percentage of analyzed transposons that show an increase in steady-state RNA levels of more than an arbitrary cut-off of 1.5-fold.
[b]More transposons are derepressed in *Ago-2* than in *Dcr-2* flies. This might be explained by the finding that some endo-siRNAs persist in *Dcr-2* mutants [8].
[c]k/d, knockdown of gene expression by RNAi.

vergent transcripts are much less dependent on *loqs* [7,11]. Perhaps the presence of internal bulges defines the dependence on *loqs* over *R2D2*. Structured loci hpRNA contain internal bulges, whereas convergent transcription generates perfect complementary dsRNA. Indeed, a transgenic construct encoding an inverted repeat that generates a perfect ~650 bp duplex RNA is fully dependent on *R2D2* for RNAi activity [23].

Exogenous RNAi: siRNA profile associated with antiviral defense

RNAi thus has multiple roles in the cell: regulation of cellular gene expression, control of transposons and defense against viruses. What, then, is the small RNA profile associated with virus infection? While profiling small RNAs of the *Drosophila* S2 cell line, Czech *et al.* [7] detected many siRNAs that map to FHV because of a persistent infection of their cell line. Seventeen percent of all Ago-2 associated siRNAs matched the viral genome. In agreement with the importance of RNAi (but not the miRNA pathway) for antiviral defense, only 0.6% of the Ago-1 associated RNAs were of viral origin [7].

v-siRNA profile during FHV infection

FHV is a positive (+) strand RNA virus with a bipartite genome (Figure 3a). As such, the viral (+) RNA serves as a template for negative (-) strand RNA synthesis by the viral RNA dependent RNA polymerase. The (-) RNA-1 subsequently serves as a template for synthesis of new (+) RNA progeny. As with all (+) RNA viruses, dsRNA is an essential intermediate in viral replication (Figure 4). Negative strand RNA synthesis, for example, produces a dsRNA replication intermediate that covers the entire genome. dsRNA replication intermediates are obvious candidates to serve as Dicer-2 substrates for v-siRNA production. In this case, roughly equal proportions of (+) and (-) v-siRNAs are expected. Other putative viral Dicer substrates are the structural RNA elements that control translation and replication of (+) RNA viruses. This scenario predicts that (+) v-siRNA predominate because viral (+) RNAs greatly outnumber (-) RNAs (50–100-fold) in infected cells [24].

To test these predictions, we aligned the 144.101 available Ago-2 associated FHV siRNAs to the viral genome [7] (Figure 3b,c). Small RNAs mapped in roughly equal proportions to both (+) and (-) viral RNA strands (42% (+) v-siRNA for RNA-1, 65.5% for RNA-2). RNA-2 contributed the majority of small RNAs in the total v-siRNA pool (70%). Corrected for the differences in size, RNA-2 generated 5.2-fold more v-siRNA than RNA-1 (13.8 v-siRNA per nt for RNA-1 and 72.4 v-siRNA per nt for RNA-2), which might partly be explained by the observation that (-) RNA-2 levels are ~2.5-fold greater than (-) RNA-1 levels [24]. These results indicate that viral dsRNA replication intermediates are the predominant source of v-siRNA production.

Small RNAs in RNA-1 mainly mapped to the 5′ terminal ~200 nt and to the 3′ terminal ~600 nt (respectively 24.2% and 64.1% of all RNA-1 v-siRNAs). In RNA-2, v-siRNA mapped to three hotspots: to the 5′ terminal ~235 nt (30.2%), to a more central region (nt 515 to 715; 62.1%) and to nt 1245 to 1330 (4.0%) in the 3′ terminal region (Figure 3b,c). The association with these hotspots is, however, not absolute. The v-siRNAs mapped across the viral genome, and only ~5% of the nucleotides were not represented in the total v-siRNA profile. Note that the subgenomic RNA-3 (corresponding to nt 2721–3107 of RNA-1) might contribute to the abundance of small RNAs in the 3′ terminal region of RNA-1. Yet, it cannot fully explain the 3′ terminal v-siRNA hot-spot because the region directly adjacent to RNA-3 (nt 2520–2720) was also an important source of v-siRNAs.

While this manuscript was under review, Aliyari *et al.* [25] described the v-siRNA profile in S2 cells abortively infected with FHV lacking the RNAi suppressor, B2. This virus (FHV ΔB2) has severe replication defects because of its inability to suppress viral dsRNA processing by Dcr-2, resulting in an extremely high abundance of v-siRNAs. As in our analyses, v-siRNAs mapped in roughly equal proportions to (+) and (-) viral RNA strands, and the density of v-siRNA in RNA-2 was higher than in RNA-1 [25]. However, the distribution of the v-siRNA across the viral genome was markedly different from our observations. The vast majority of v-siRNA map to the 5′ terminal ~400 nt of RNA-1 and, to a lesser extent, to the 5′ terminal ~200 nt of RNA-2. In addition to these hotspots, low abundant v-siRNAs mapped with a relatively uniform distribution across the viral RNAs. FHV ΔB2 replicates to extremely low levels, and produces little, if any, RNA-3, possibly explaining the lack of v-siRNAs mapping to the 3′ terminal region of RNA-1. Indeed, in wildtype FHV infection, the extreme predominance of v-siRNA at the 5′ terminal region of RNA-1 was lost, and the 3′ terminal region ~440 nt became the predominant source of v-siRNAs [25], in accordance with our analyses.

Local RNA structures derived from the viral genomic RNA are the predominant source of v-siRNAs in some (+) RNA viruses of plants [26]. Do structured RNA elements also contribute to the v-siRNA profile in FHV infection? Unfortunately, few RNA elements have been experimentally characterized in the FHV genome. However, a region with a potential to form an extensive stemloop structure [27]) gives rise to a peak of predominantly (+) v-siRNAs (asterisk in Figure 3b,c). The presence of this stemloop structure in the viral genome, and its potential to serve as a substrate for Dicer-2, awaits experimental validation.

RNA-2-derived defective interfering (DI) RNAs frequently arise in persistently infected cell lines, such as the one used in this study, and upon serial passage of FHV RNA in *Drosophila* cells [28,29]. These DIs usually comprise of three regions (~ nt 1–249, 517–728 and 1228–1400 of RNA-2 [28,29]) that show a remarkable correspondence to the regions that give rise to the majority of v-siRNA in RNA-2. Of note, DIs were previously shown to be an important source of v-siRNA in tombusvirus infection in plants [30]. It is, therefore, a reasonable hypothesis that DIs are an important source of RNA-2-derived v-siRNAs in these persistent infections. Further experiments in a more defined experimental setting are needed to test this hypothesis. Perhaps the appearance of DIs might contribute to the over-representation of RNA-2-derived v-siRNAs in the entire v-siRNA pool.

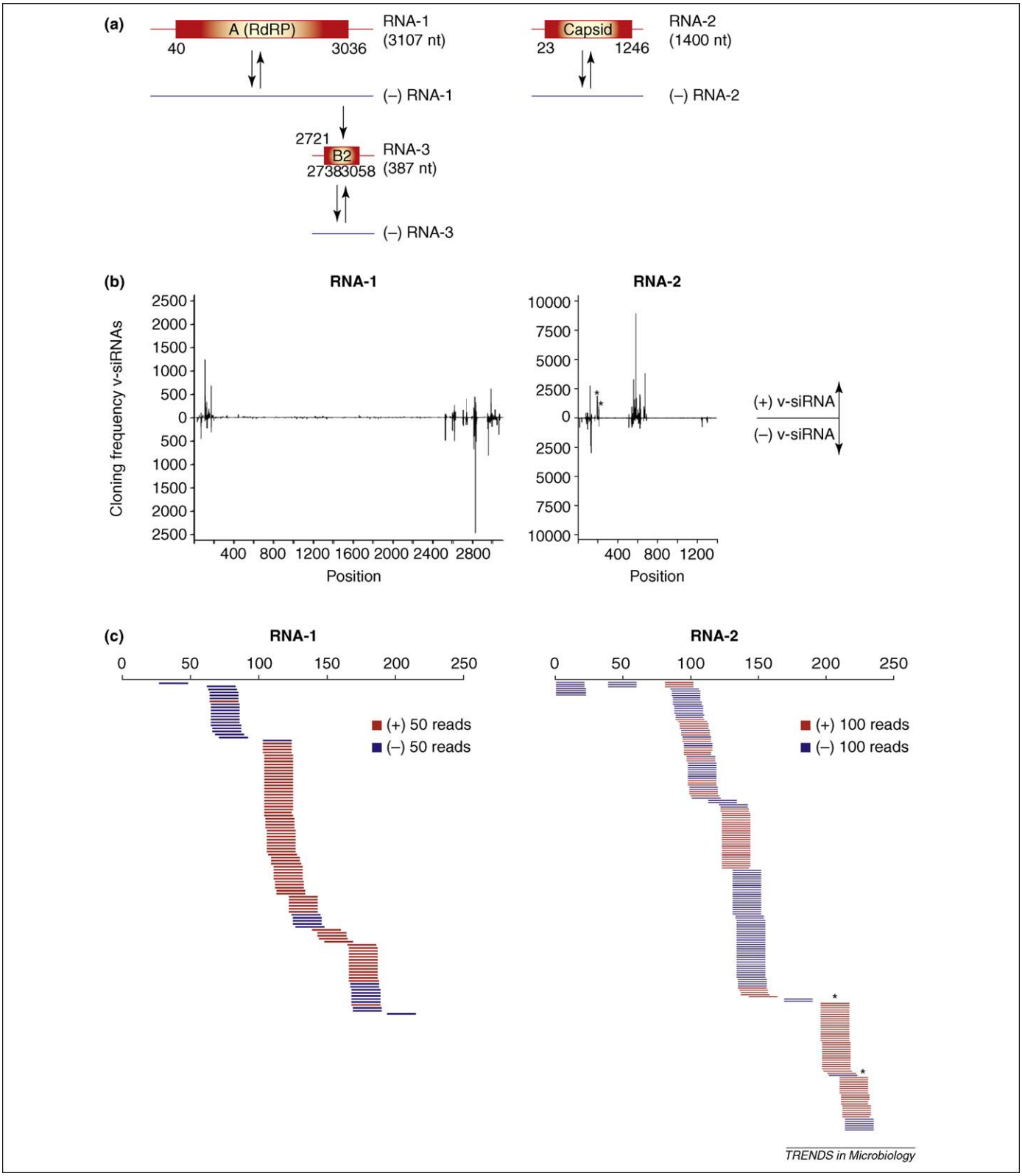

Figure 3. Viral siRNA profile in FHV infection. **(a)** Schematic representation of the FHV replication cycle. After viral entry, the host cell machinery translates RNA-1, generating protein A, the viral RNA dependent RNA polymerase (RdRP). Using the positive strand (+) viral RNA as a template, the RdRP generates negative strand (-) RNA, which subsequently serves as a template for the production of (+) RNA progeny. RNA-2 encodes the capsid protein, and depends on the viral RdRP for replication. Using RNA-1 as a template, the RdRP also generates a subgenomic RNA-3 that encodes the B2 protein, a viral suppressor of RNAi. (-) RNA-3 is produced during infection, and subsequently might serve as a template for synthesis of progeny (+) RNA-3 [29,32]. **(b)** Cloning frequency of Ago-2 associated viral small interfering RNAs (v-siRNA) from *Drosophila* S2 cells persistently infected with FHV. v-siRNAs were aligned to the viral genome and the position of the 5′ Dicer cleavage sites in the genome were represented by vertical lines, the height of the line correlating with the cloning frequency. Vertical lines above the *x*-axis indicate (+) v-siRNA; lines below the *x*-axis represent (-) v-siRNA. The asterisk indicates a region predicted to generate a stem-loop structure [27] that is enriched in (+) v-siRNAs. The Ago-1 associated v-siRNA profile was similar to the Ago-2 associated v-siRNAs with regard to size, strand bias and distribution over the viral genome (data not shown). **(c)** Alignment of v-siRNAs to the 5′ terminal 250 nt of RNA-1 and RNA-2. Red bars represent (+) v-siRNAs; blue bars represent (-) v-siRNAs. Each bar represents 50 (RNA-1) or 100 (RNA-2) individual reads. Accession numbers: NC004146 and NC004144 (viral genomic RNA) and gene expression omnibus GSE11086 (v-siRNAs) [7].

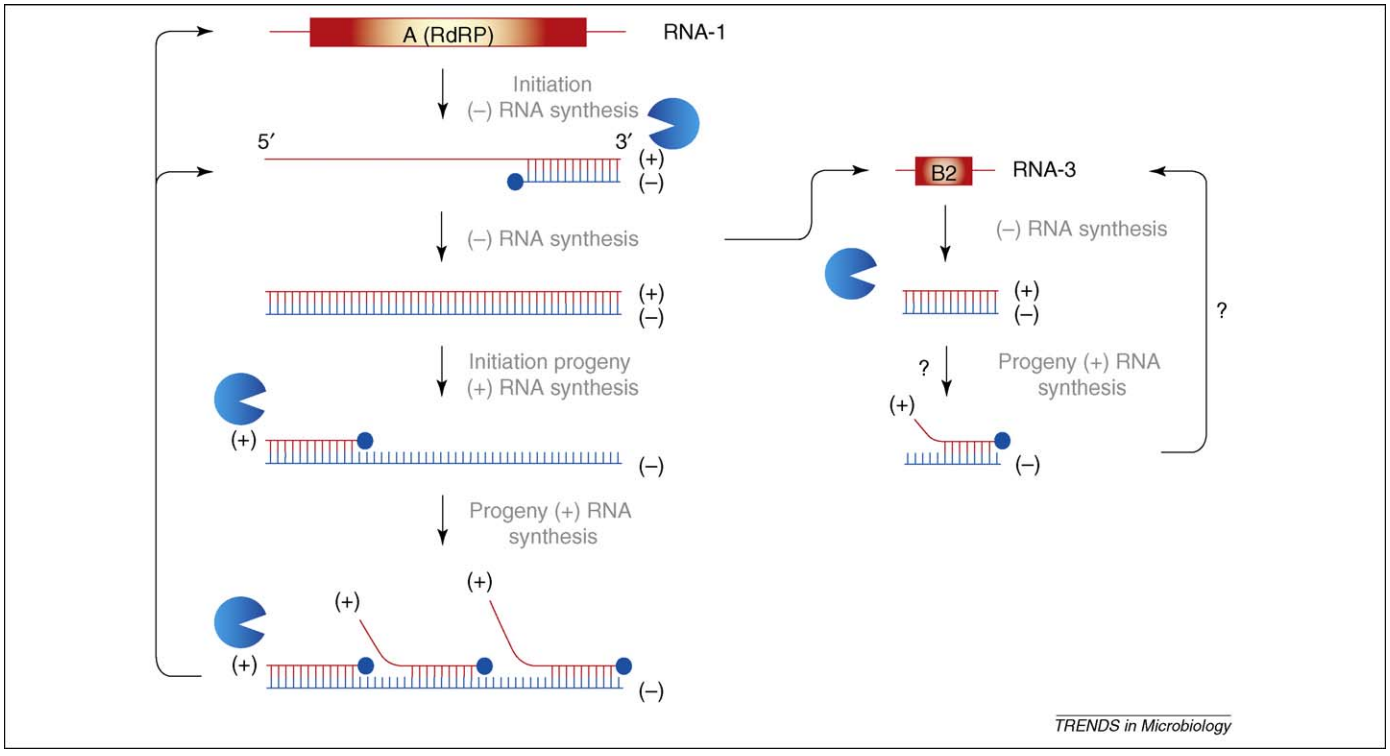

Figure 4. Model for production of FHV v-siRNAs by *Drosophila* Dicer-2. dsRNA is an essential intermediate in replication of (+) RNA viruses. Negative strand RNA synthesis produces a genome-length dsRNA replication intermediate. The viral (-) RNA serves as a template for multiple rounds of synthesis of progeny (+) RNA, which also generates stretches of viral dsRNA. FHV v-siRNAs of both (+) and (-) polarity predominantly map to the 5′ and 3′ terminal regions of RNA-1 (Figure 3b and Ref. [25]). These observations indicate that Dicer-2 cleaves dsRNA generated during initiation of the synthesis of viral (-) and progeny (+) RNA and during replication of subgenomic RNA-3. Blue dots represent the viral RNA dependent RNA polymerase (RdRP).

Biogenesis of v-siRNAs

The observation that v-siRNAs mainly map to the 5′ and 3′ terminal regions of RNA-1 can be explained by several, non-mutually exclusive, scenarios (Figure 4). First, the initiation of viral (-) RNA or progeny (+) RNA synthesis might provide a window of opportunity in which the viral dsRNA is prone to Dicer-2 digestion. For example, a hypothetical event of stalling, or even termination, of (-) RNA synthesis could lead to a prolonged exposure of a 3′ terminal dsRNA structure to Dicer-2. Internal RNA structures might be responsible for such a stalling event. In support of this hypothesis is the identification of an internal replication element (nt 2322 to 2501), predicted to fold into a long multi-branched stem-loop, directly upstream of the 3′ terminal RNA-1 hot-spot for v-siRNAs [31]. The notion that, by definition, B2 is absent during the initial round of (-) RNA synthesis might further contribute to the 3′ terminal v-siRNA hotspot. Second, the production and replication of RNA-3 [32] could generate dsRNA Dicer-2 substrates that overlap with the 3′ terminal region of RNA-1. Third, the terminal regions of the genome-length viral dsRNA replication intermediate might be particularly accessible to Dicer-2.

Biochemical analyses indicate that Dicer progressively cleaves from the termini of dsRNA structures, resulting in 'phasing' of siRNAs. The absence of pronounced phasing of v-siRNAs (Figure 3c) indicates that Dicer-2 processing does not initiate from a defined start site in the viral dsRNA. Thus, whereas the 5′ and 3′ terminal regions of RNA-1 and RNA-2 are important sources of v-siRNAs, Dicer-2 cleavage does not initiate at the ultimate termini. Perhaps structural features of the viral RNA, or association of viral RNA with viral or cellular proteins prevents the interaction of the Dicer-2 PAZ domain with the terminal nucleotides of the viral dsRNA. Heterogeneity at the ends of the viral dsRNA would provide an alternative explanation for the absence of phasing of v-siRNAs. We deem this possibility, however, unlikely, given the lack of length heterogeneity at the 5′ end of the viral RNA [33] and the absence in the FHV genome of a protruding variable-length poly(A) tail.

How does the biogenesis of v-siRNAs in the antiviral exo-RNAi pathway compare to the endo-siRNA pathway? Previous work indicated that *R2D2* mutant flies are hypersensitive to viral infection [3]. This result is at odds with the observations from Czech *et al.* [7] that knockdown of *loqs*, but not *R2D2*, results in a 2.5 to fivefold increase in FHV RNA levels. The cause of this discrepancy remains unclear. For now, it seems that perfect duplex siRNAs (such as those produced from experimental dsRNA, convergent transcripts and viral dsRNA) are loaded into RISC by an R2D2–Dcr-2 heterodimer. Internal bulges in an extended dsRNA structure, such as found in structured loci hpRNA, could define a requirement for *loqs* and *Dcr-2*. Whether Dcr-2–loqs and Dcr-2–R2D2 load their cargo into qualitatively distinct RISC complexes remains an intriguing possibility. The observation that exogenously provided siRNAs could not displace *loqs*-dependent endo-siRNA from RISC fits this possibility [9].

Conclusions and future perspectives

The *Drosophila* RNAi machinery is a master of multitasking. The endogenous RNAi pathway regulates cellular gene expression and contributes to control of transposon activity in soma and germline. By contrast, the exogenous

RNAi pathway is an important antiviral mechanism. Whether there is an interaction between the exo- and endo-RNAi pathways remains to be established. For example, does an acute viral infection displace endo-siRNAs? Does this lead to derepression of transposons and deregulation of endo-RNAi regulated cellular transcripts? Alternatively, there could be a residual pool of 'empty' RISC complexes awaiting their loading by exogenous sources of siRNAs, as suggested by Kawamura *et al.* [9]. Distinct RISC complexes dedicated to endo-RNAi and exo-RNAi would allow the cell to respond to viral infection without interfering with endo-RNAi regulated processes.

An important implication of these studies is the notion that Dcr-2 substrates can be more complex than previously anticipated. Biochemical studies indicate that Dicer progressively cleaves from termini of dsRNA and that cleavage is blocked by terminal extensions of ssRNA [34]. Extensive stretches of ssRNA flank the overlapping dsRNA regions in convergent transcripts, which nevertheless produce endo-siRNAs. Similarly, some endo-siRNAs seem to be excised from hpRNAs with extensive secondary structures [12]. Furthermore, the lack of phasing of siRNAs derived from transposons, convergent transcripts and viral sequences, indicates that internal initiation of dsRNA processing by Dcr-2 frequently occurs *in vivo*. These observations might prompt us to consider the possibility that cytoplasmic RNA viruses sacrifice some viral (+) RNA for v-siRNA production to manipulate gene expression of the host. For example, Dcr-2 could cleave structured RNA elements to generate v-siRNAs with extensive complementarity to specific host genes. These v-siRNAs could then inhibit expression of these genes via an RNAi or miRNA mechanism. This scenario resembles the strategy of mammalian nuclear DNA viruses that encode miRNAs for regulation of viral and host gene expression; this contributes to viral pathogenesis, for example, through modulation of antiviral immune responses [35].

Viral dsRNA replication intermediates seem to be the predominant source of v-siRNAs. This conclusion is difficult to reconcile with the observations that RNA viruses replicate exclusively in defined membranous structures. These structures are thought to shield the viral RNA from cellular ribonucleases and from cellular sensors for viral RNA that trigger innate immune responses (such as Retinoic acid inducible gene I (RIG)-like helicases). FHV, for example, replicates in spherical structures on the outer membrane of mitochondria, and FHV dsRNA is exclusively detected in these structures [24]. Similarly, Drosophila C virus (DCV), another target of *Drosophila* RNAi, replicates in a membranous compartment derived from the Golgi apparatus [36]. Is Dcr-2 capable of protruding into vesicles of such different origin? Alternatively, are there specific stages in the viral life cycle in which the viral dsRNA is exposed to the cytoplasm, in which Dcr-2 is thought to reside? In this regard, it will be of great interest to see how other classes of RNAi-controlled viruses, such as the (-) RNA virus VSV [37], are processed by Dicer.

Acknowledgements

We thank Carla Saleh and Marco Vignuzzi for critical reading of the manuscript. This work was financially supported by a fellowship from the Nijmegen Centre for Molecular Life Sciences to R.vR and by VIDI-fellowships from the Netherlands Organization for Scientific Research to R.vR. and E.B.

References

1 van Rij, R.P. *et al.* (2006) The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes Dev.* 20, 2985–2995
2 Galiana-Arnoux, D. *et al.* (2006) Essential function *in vivo* for Dicer-2 in host defense against RNA viruses in drosophila. *Nat. Immunol.* 7, 590–597
3 Wang, X.H. *et al.* (2006) RNA interference directs innate immunity against viruses in adult Drosophila. *Science* 312, 452–454
4 Ding, S.W. and Voinnet, O. (2007) Antiviral immunity directed by small RNAs. *Cell* 130, 413–426
5 Matranga, C. and Zamore, P.D. (2007) Small silencing RNAs. *Curr. Biol.* 17, R789–R793
6 Vagin, V.V. *et al.* (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313, 320–324
7 Czech, B. *et al.* (2008) An endogenous small interfering RNA pathway in Drosophila. *Nature* 453, 798–802
8 Ghildiyal, M. *et al.* (2008) Endogenous siRNAs derived from transposons and mRNAs in Drosophila Somatic Cells. *Science* 320, 1077–1081
9 Kawamura, Y. *et al.* (2008) Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* 453, 793–797
10 Okamura, K. *et al.* (2008) Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. *Nat. Struct. Mol. Biol.* 15, 581–590
11 Chung, W.J. *et al.* (2008) Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr. Biol.* 18, 795–802
12 Okamura, K. *et al.* (2008) The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* 453, 803–806
13 Tam, O.H. *et al.* (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 453, 534–538
14 Watanabe, T. *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453, 539–543
15 Tomari, Y. and Zamore, P.D. (2005) Perspective: machines for RNAi. *Genes Dev.* 19, 517–529
16 Sontheimer, E.J. (2005) Assembly and function of RNA silencing complexes. *Nat. Rev. Mol. Cell Biol.* 6, 127–138
17 Horwich, M.D. *et al.* (2007) The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* 17, 1265–1272
18 Saito, K. *et al.* (2007) Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3′ ends. *Genes Dev.* 21, 1603–1608
19 Waterhouse, P.M. *et al.* (2001) Gene silencing as an adaptive defence against viruses. *Nature* 411, 834–842
20 Sijen, T. and Plasterk, R.H. (2003) Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* 426, 310–314
21 Rehwinkel, J. *et al.* (2006) Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in *Drosophila melanogaster*. *Mol. Cell. Biol.* 26, 2965–2975
22 Brennecke, J. *et al.* (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* 128, 1089–1103
23 Forstemann, K. *et al.* (2005) Normal microRNA maturation and germline stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol.* 3, e236
24 Kopek, B.G. *et al.* (2007) Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol.* 5, e220
25 Aliyari, R. *et al.* (2008) Mechanism of induction and suppression of antiviral immunity directed by virus-derived small RNAs in Drosophila. *Cell Host Microbe* 4, 387–397
26 Molnar, A. *et al.* (2005) Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. *J. Virol.* 79, 7812–7818
27 Zhong, W. *et al.* (1992) Evidence that the packaging signal for nodaviral RNA2 is a bulged stem-loop. *Proc. Natl. Acad. Sci. U. S. A.* 89, 11146–11150

28 Li, Y. and Ball, L.A. (1993) Nonhomologous RNA recombination during negative-strand synthesis of flock house virus RNA. *J. Virol.* 67, 3854–3860
29 Zhong, W. and Rueckert, R.R. (1993) Flock house virus: down-regulation of subgenomic RNA3 synthesis does not involve coat protein and is targeted to synthesis of its positive strand. *J. Virol.* 67, 2716–2722
30 Szittya, G. *et al.* (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
31 Lindenbach, B.D. *et al.* (2002) Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. *J. Virol.* 76, 3905–3919
32 Eckerle, L.D. *et al.* (2003) Flock House virus subgenomic RNA3 is replicated and its replication correlates with transactivation of RNA2. *Virology* 317, 95–108
33 Ball, L.A. and Li, Y. (1993) cis-acting requirements for the replication of flock house virus RNA 2. *J. Virol.* 67, 3544–3551
34 Elbashir, S.M. *et al.* (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15, 188–200
35 Cullen, B.R. (2006) Viruses and microRNAs. *Nat. Genet.* 38 (Suppl), S25–S30
36 Cherry, S. *et al.* (2006) COPI activity coupled with fatty acid biosynthesis is required for viral replication. *PLoS Pathog.* 2, e102
37 Wilkins, C. *et al.* (2005) RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. *Nature* 436, 1044–1047
38 Van Rij, R.P. and Andino, R. (2008) The complex interactions of viruses and the RNAi machinery: A driving force in viral evolution. In *Origin and Evolution of Viruses* (Domingo, E. *et al.*, eds), pp. 161–181, Elsevier
39 Forstemann, K. *et al.* (2007) Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* 130, 287–297
40 Cullen, B.R. (2006) Is RNA interference involved in intrinsic antiviral immunity in mammals? *Nat. Immunol.* 7, 563–567
41 van Rij, R.P. and Andino, R. (2006) The silent treatment: RNAi as a defense against virus infection in mammals. *Trends Biotechnol.* 24, 186–193
42 Marques, J.T. and Carthew, R.W. (2007) A call to arms: coevolution of animal viruses and host innate immune responses. *Trends Genet.* 23, 359–364
43 van Rij, R.P. (2008) Virus meets RNAi. Symposium on antiviral applications of RNA interference. *EMBO Rep.* 9, 725–729
44 Keene, K.M. *et al.* (2004) RNA interference acts as a natural antiviral response to O'nyong-nyong virus (Alphavirus; Togaviridae) infection of Anopheles gambiae. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17240–17245
45 Campbell, C.L. *et al.* (2008) Aedes aegypti uses RNA interference in defense against Sindbis virus infection. *BMC Microbiol.* 8, 47
46 Sanchez-Vargas, I. *et al.* (2004) RNA interference, arthropod-borne viruses, and mosquitoes. *Virus Res.* 102, 65–74
47 Zambon, R.A. *et al.* (2006) RNAi is an antiviral immune response against a dsRNA virus in *Drosophila melanogaster*. *Cell. Microbiol.* 8, 880–889
48 Chellappan, P. *et al.* (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
49 Moissiard, G. and 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
50 Chao, J.A. *et al.* (2005) Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nat. Struct. Mol. Biol.* 12, 952–957
51 Lingel, A. *et al.* (2005) The structure of the flock house virus B2 protein, a viral suppressor of RNA interference, shows a novel mode of double-stranded RNA recognition. *EMBO Rep.* 6, 1149–1155
52 Lu, R. *et al.* (2005) Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature* 436, 1040–1043
53 Lakatos, L. *et al.* (2006) Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J.* 25, 2768–2780
54 Zhang, X. *et al.* (2006) Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis* Argonaute1 cleavage activity to counter plant defense. *Genes Dev.* 20, 3255–3268