Slicing of viral RNA guided by endogenous piRNAs triggers the production of responder and trailer piRNAs in *Aedes* mosquitoes.

Joep Joosten[1], Ronald P. Van Rij[1+] and Pascal Miesen[1+]

[1]Department of Medical Microbiology, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

[+] corresponding authors: ronald.vanrij@radboudumc.nl; pascal.miesen@radboudumc.nl

Key words: piRNA, *Aedes* mosquitoes, endogenous viral element, Zucchini, Nibbler, phasing

ABSTRACT

PIWI interacting (pi)RNAs are small RNAs mostly known to protect the genomes of animal germlines against transposable elements. In *Drosophila,* piRNA-mediated cleavage of transposon RNA triggers the release of a responder piRNAs via the ping-pong amplification cycle. Responder piRNA 3' end formation by the endonuclease Zucchini is coupled to the production of downstream trailer piRNAs, expanding the repertoire of piRNAs that target transposons. Intriguingly, in *Aedes aegypti* mosquitoes, somatic piRNAs are produced from viral RNA, implying a role of viral (v)piRNAs in antiviral immunity. Knowledge on how vpiRNA 3' ends are formed and whether viral RNA is subjected to trailer piRNA production, however, is lacking. To address these questions, we analyzed small RNA sequencing libraries from *Ae. aegypti* cells. We found that virus- and transposon-derived piRNAs have sharply defined 3' ends, and that uridine residues are enriched directly downstream of dominant piRNA sequences, both of which are characteristic features of Zucchini-like cleavage of precursor piRNAs. Next, we designed a piRNA reporter system based on Sindbis virus recombinants that harbor target sites for abundant endogenous piRNAs. These piRNAs guide cleavage of the viral RNA, which is subsequently processed into abundant responder piRNAs. Using this reporter virus system, we identified the *Ae. aegypti* orthologs of Zucchini, which is required for sharp 3' end formation of responder piRNAs, and Nibbler, a 3'-5' exonuclease involved in the trimming of a subset of piRNAs and miRNAs. Furthermore, we found that cleavage of viral RNA triggers the production of trailer piRNAs, thus expanding the piRNA sequence pool that is able to target viral RNA. Our results have important implications for understanding how autonomous piRNA production from viral RNA can be triggered by just a few cleavage events by genome-encoded piRNAs.

INTRODUCTION

Blood-feeding mosquitoes of the *Aedes* genus are responsible for the spread of arthropod-borne (arbo)viruses that cause severe diseases, such as dengue, Zika, chikungunya and yellow fever. Due to rising global temperatures and increased human mobility, the geographical distribution of these vector mosquitoes is continuously expanding, which has resulted in a rapid increase in the number of arbovirus epidemics in the last decennia (1). For efficient transmission to occur, arboviruses have to actively replicate in several mosquito tissues to eventually infect the salivary gland (2). Therefore, suppression of virus replication by the mosquito antiviral immune response strongly affects the efficiency of arboviral spread. The cornerstone of antiviral immunity in insects is the small interfering (si)RNA pathway, in which viral double stranded (ds)RNA is cleaved by Dicer-2 (3). The resulting cleavage products are processed into siRNAs, which provide sequence specificity to the endonuclease Argonaute 2 to direct the cleavage of single stranded viral transcripts.

Intriguingly, in *Aedes* mosquitoes, viral RNA is also processed by a somatically active PIWI interacting (pi)RNA pathway, suggesting that two independent small RNA pathways act in parallel to combat viral infections (4). The piRNA biogenesis machinery has been thoroughly characterized in the model organism *Drosophila melanogaster*, where a germline-restricted piRNA pathway defends the genome from parasitic genetic elements called transposons (5,6). piRNA biogenesis is initiated by the cleavage of genomically encoded piRNA precursors, which are rich in transposon remnants. Slicing of these precursor transcripts into pre-piRNAs is mediated either by a piRNA-guided PIWI protein or the endonuclease Zucchini (Zuc), which acts independently of small RNAs (7-9). Pre-piRNAs are loaded into the PIWI proteins Aubergine (Aub) and Piwi, where their 3' ends may be further trimmed by the exonuclease Nibbler (Nbr) and are finally 2'-*O*-methylated by Hen1 to generate mature piRNAs (10-14). Whereas Piwi translocates to the nucleus to silence transposons at the transcriptional level (15,16), Aub remains in the cytoplasm where it cleaves transposon mRNA with sequence complementarity to its associated piRNA (17,18). The resulting cleavage fragments are loaded into the PIWI protein Argonaute 3 (Ago3) and processed into responder piRNAs by Nbr-mediated trimming, followed by 2′-*O*-methylation by Hen1 (12,14). In turn, these responder piRNAs direct Ago3-mediated cleavage of piRNA precursors, triggering the production of new initiator piRNAs, completing the so-called ping-pong loop (17-20).

In *Drosophila,* Zuc-mediated generation of piRNA 3' ends often produces a downstream cleavage product which is preferentially loaded into Piwi, thus forming a new pre-piRNA. This mechanism results in phased processing of piRNA precursor transcripts into a string of piRNAs, named trailer piRNAs (8,9). Hence, while the ping-pong loop amplifies those piRNAs that initially recognized active transposons, phased piRNA production expands the piRNA repertoire to allow for more efficient repression of transposons. Historically, piRNAs derived from cluster transcripts and transposon mRNAs were termed primary and secondary piRNAs, respectively. Hereafter, we use the more intuitive names initiator or responder for ping-pong amplified piRNAs and trailer for piRNAs produced through phased biogenesis, as proposed in (5).

The *Ae. aegypti* piRNA pathway differs from that in *Drosophila* in three important ways: *i)* the pathway is active in somatic tissues as well as germline tissues (21,22), *ii)* it processes non-canonical substrates such as *de novo* produced viral RNA (23-25), and *iii)* mosquito piRNA clusters contain large numbers of endogenous viral elements (EVEs): sequences of non-retroviral RNA viruses inserted in mosquito genomes (26-28). These EVEs give rise to abundant piRNAs (26,29,30) and mediate antiviral defense (31-33).

In *Ae. aegypti* mosquitoes, the PIWI proteins Piwi5 and Ago3 are brought together by the TUDOR proteins Veneno and Yb and the RNA helicase Vasa to promote efficient ping-pong dependent production of piRNAs (34), yet, the mechanisms involved in specifying piRNA 3' ends is unknown. Moreover, while the *Ae. aegypti* piRNA pathway was recently proposed to efficiently generate trailer piRNAs (7), the molecular machinery responsible for phased piRNA production and whether viral RNA is also subjected to phased piRNA production is unclear.

It has been proposed that endogenous piRNAs encoded in the mosquito genome can trigger the production of piRNAs from viral sources (31,32). We hypothesized that through the mechanism of piRNA phasing, a single endogenous initiator piRNA could induce the production of a pool of viral piRNAs, thereby increasing the piRNA population that targets viral RNA. Here, we uncover that this interaction between an endogenous initiator piRNAs and viral RNA substrate indeed takes place in *Ae. aegypti*. Specifically, we demonstrate that *Ae. aegypti* piRNAs, both of transposon and viral origin, display sequence features indicative of a Zucchini-like biogenesis mechanism. Hence, we use a viral piRNA reporter system to show that AAEL011385 and AAEL005527, the *Ae. aegypti* orthologs of *Drosophila* Zuc and Nibbler, respectively, cooperatively determine piRNA 3' ends. Lastly, we show that cleavage

guided by a genomically encoded initiator piRNA triggers the production of additional trailer piRNAs from the viral genome, thus increasing the pool of piRNAs targeting the newly infecting virus. This mechanism may equip the *Aedes* piRNA pathway with an adaptive immune response that is able to adapt to newly encountered and continuously mutating viruses.

RESULTS AND DISCUSSION

Aedes aegypti piRNAs have sharp 3' ends

In *Drosophila*, piRNA 3' end formation is largely dependent on the cleavage of pre-piRNAs by the endonuclease Zucchini (Zuc). Zuc uses a sequence motif to preferentially cleave upstream of uridine residues *in vivo* (35-37), hence, piRNAs generated by Zuc generally have sharp 3' ends and the nucleotide directly downstream of the 3' end tends to be a uridine (+1U bias) (8,9). We examined whether these characteristics were present in our previously generated small RNA deep sequencing libraries from *Ae. aegypti* Aag2 cells (24) infected with Sindbis virus (SINV), a positive sense RNA virus of the *Togaviridae* family. We first analyzed transposon-derived piRNAs and found that individual piRNAs, defined by a shared 5' end, generally had the same length (Figure 1A). Specifically, for almost 60% of piRNAs the dominant length made up more than 75% of sequenced reads. We selected those piRNAs for which the dominant sequence length represented at least 75% of the reads and inspected the identity of the nucleotides downstream of that most abundant piRNA isoform. We found that the nucleotide position directly following the 3' end of the piRNA was biased for uridine (Figure 1B), strongly indicating that these piRNAs were generated by a mechanism that resembles Zuc cleavage in *Drosophila*. Strikingly, sharp 3' ends and +1U biases were also clearly visible for SINV-derived piRNAs, irrespective of the strand from which the piRNAs were produced (Figure 1C-D). These findings suggest that 3' ends of both (+) strand derived vpiRNAs, which are predominantly Ago3-associated and (-) strand derived vpiRNAs, which are mostly Piwi5-bound (24,34), are generated, at least in part, by Zuc-like cleavage events. Interestingly, we also observe sharp 3' ends and +1U biases for piRNAs generated from Phasi Charoen-like virus (Figure S1), a negative sense RNA virus from the *Phenuiviridae* family that persistently infects Aag2 cells (38). These findings indicate that Zuc-like piRNA biogenesis targets diverse RNA viruses.

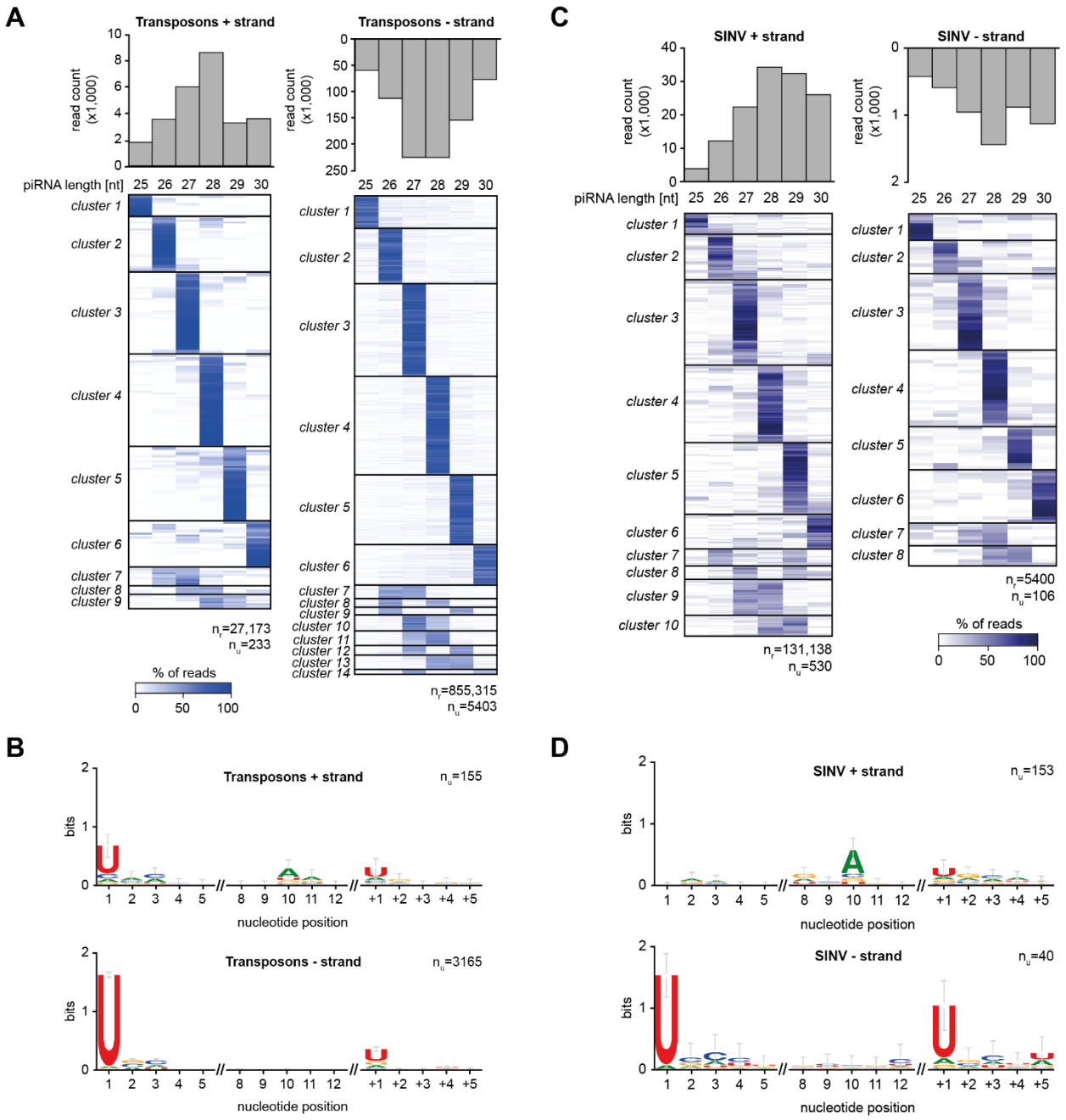


Figure 1. *Aedes aegypti* piRNAs have sharp 3' ends
(A) Heat map showing the relative size distribution of individual transposon-derived piRNAs, defined by a shared 5' end. Shades of blue indicate the percentage of reads contributing to the indicated read length, white represents absence of reads in the specific size class. The number of unique piRNAs (n_u) and the number of reads (n_r) that underlie the heat map are indicated. A minimum of 20 reads/unique piRNA was required to be included in the analysis.
(B) Nucleotide biases for the indicated nucleotide positions of transposon-derived piRNAs and the sequence at the genomic region directly downstream (+1 until +5) piRNA 3' ends. Only piRNAs from (A) that had a dominant length (at least 75% of reads) were considered in this analysis and only unique sequences were analyzed, disregarding read count for the individual piRNAs.
(C-D) The same analysis as for A and B was applied to piRNAs mapping to SINV.

Genomically encoded piRNAs are able to trigger production of virus-derived responder piRNAs

To study sequence determinants for vpiRNA 3' end formation, we designed a SINV-based

reporter system in which exogenous piRNA target sequences can be easily introduced. We modified SINV such that a duplicated subgenomic promoter drives the expression of a non-coding RNA sequence that harbors a target site for an abundant Piwi5-associated initiator piRNA (Figure 2A, Figure S2A). Initiator piRNA-guided recognition of the target site should trigger slicing by Piwi5 and subsequent processing of the resulting cleavage fragment into an Ago3-associated responder piRNA through the ping-pong amplification cycle. We designed viruses with target sites for two abundant endogenous Piwi5-associated initiator piRNAs: one derived from the *Ty3/gypsy* LTR retroelement *gypsy*73 (*g*73 - Figure 2A), and a second originating from an EVE of flaviviral origin (FV53, described in Supplemental text).

Our previous results indicate that both transposon- and SINV-derived piRNAs have a strong bias towards a uridine residue directly downstream their 3' ends (Figure 1B, D). To study the importance of +1U for viral responder piRNA production in *Aedes* mosquitoes, we introduced uridine residues at specified distances from the putative Piwi5 slice site in the viral reporter system (Figure 2A; viruses were named *g*73-25U, 28U, and 30U according to the responder piRNA 5' end-to-U distance). Responder piRNAs were readily detected by high resolution northern blotting in Aag2 cells infected with all reporter viruses (Figure 2B), but not in cells infected with control virus expressing GFP from the duplicated subgenomic promoter (SINV 3'GFP). These results indicate that endogenous piRNAs may instruct the cleavage of exogenous viral RNA and the subsequent production of responder piRNAs during acute infection.

Responder piRNA size did not fully reflect the distance between the Piwi5 cleavage site and the downstream uridine residue; while no clear differences in responder piRNA size are observed between *g*73-25U and g73-28U viruses, increasing the 5' end-to-U distance to 30 nt (*g*73-30U) results in a more diffuse distribution of responder piRNA 3' ends (Figure 2B). These data suggest that downstream uridines may not be absolutely required for determining the 3' ends of the reporter-derived responder piRNAs or that additional exonucleolytic trimming of pre-piRNA 3' ends mask a putative endonucleolytic cleavage event directly upstream of the uridine residues.

To discriminate between these two possibilities, we generated small RNA deep sequencing libraries from Aag2 cells infected with the various reporter viruses. We hypothesized that even if a Zuc mediated cleavage event would be masked by exonucleolytic trimming, the presence of downstream trailer piRNAs with 5' ends coinciding with the U residues would

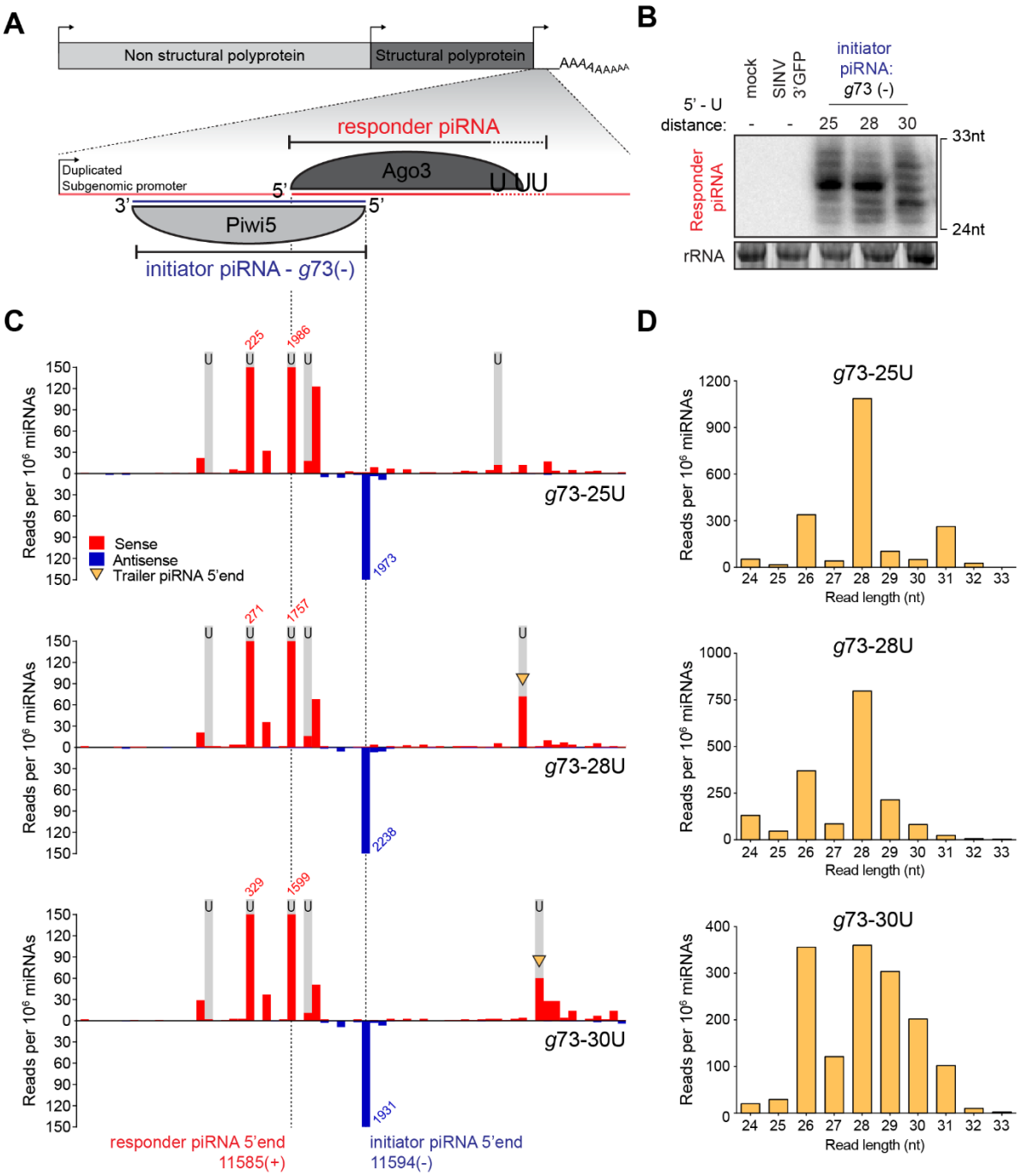


Figure 2. An endogenous piRNA is able to trigger production of virus-derived responder piRNAs
(A) Schematic representation of recombinant Sindbis reporter viruses. The enlarged view depicts the reporter locus expressed under the control of a duplicated subgenomic promoter. This non-coding RNA harbors a target site for a Piwi5-associated initiator piRNA derived from the *Ty3-gypsy7*3 retrotransposon (*g*73 - indicated in blue). Slicing of this target triggers the production of responder piRNAs (indicated in red) that are loaded into Ago3. The distance between the responder piRNA 5' end and the position of downstream uridine residues in the various viruses used in later experiments are shown.
(B) Northern blot analyses of responder piRNA production in cells infected with the indicated reporter viruses. Numbers indicate the distance between the 5' end of the responder piRNA and the first downstream uridine residue. SINV 3' GFP is a virus without an initiator piRNA target site and serves as a negative control, as do mock-infected cells. The positions marking 24 and 33 nt are inferred from an EtBr-stained small RNA size marker. EtBr-stained rRNA serves as loading control.
(C) Visualization of 5' ends of sense (red) and antisense (blue) piRNAs (24-33 nt) mapping to the reporter locus of the indicated viruses. Dashed lines indicate responder and initiator piRNA 5' ends and the positions of grey shading indicate the location of uridine residues on the sense strand of the various viruses. The red and blue numbers show piRNA counts (in reads per 10^6 miRNAs) that exceed the range of the y-axis; yellow arrowheads indicate the 5' ends of trailer piRNAs.
(D) Size distribution of responder piRNAs produced from the various *g*73-viruses, as determined by small RNA seep sequencing. Read counts were normalized to the number of miRNAs in each library.

indicate Zuc-mediated endonucleolytic cleavage was responsible for generating the responder piRNA 3' end. Mapping of piRNA 5' ends to the genomes of *g*73-targeted viruses reveals that virtually no antisense piRNAs other than the initiator piRNAs map to the reporter sequence (Figure 2C). These initiator piRNAs trigger the production of highly abundant sense responder piRNAs with characteristic 10 nt overlap of piRNA 5' ends, indicative of production by ping-pong amplification (Figure 2C). Responder piRNA size distribution for *g*73-targeted viruses generally recapitulates the results from the northern blot analysis, with reduced 3' end sharpness for the *g*73-30U virus (Figure 2D).

We noted that responder piRNAs <26 nt in size are inefficiently produced from any virus (including those with a responder 5' end-to-U distance of 25 nt: *g*73-25U and FV53-25/28U). As responder piRNAs are expected to associate with Ago3, we hypothesize that Ago3 may cover the first 25 nucleotides of pre-piRNAs, rendering them inaccessible to endo- or exonucleolytic processing. Indeed, small RNA data from Ago3 immunoprecipitation indicate a clear preference of Ago3 to bind piRNAs in the size range of 26-30 nt (Figure S3A). In line with this, responder piRNAs produced from all reporter viruses are predominantly 26-30 nt in size (Figure 2D).

Intriguingly, in viruses with a responder piRNA 5' end-U distance ≥28 nt, we uncover the production of a putative trailer piRNA (indicated with yellow arrowheads in Figure 2C) downstream of the responder piRNA. Even in viruses that produce responder piRNAs with diffuse 3' ends (e.g. *g*73-30U), the 5' end of the following trailer piRNA is sharply defined at the downstream uridine. These data suggest that downstream uridines may instruct endonucleolytic cleavage, thus coupling responder (pre-)piRNA 3' end formation to trailer piRNA 5' end formation. Accordingly, no trailer piRNA production is observed in cells infected with the *g*73-25U virus, in which we propose the downstream uridine is inaccessible for endonucleolytic cleavage. The diffuse responder piRNA size profile of *g*73-30U likely results from trimming of piRNA 3´ ends.

Responder piRNAs are produced through ping-pong mediated slicing

We previously identified Ago3 and Piwi5 as the core components of the ping-pong amplification loop in *Ae. aegypti* (24,34). We thus wanted to analyze which PIWI proteins are responsible for the generation of the responder piRNAs from our reporter virus. As expected,

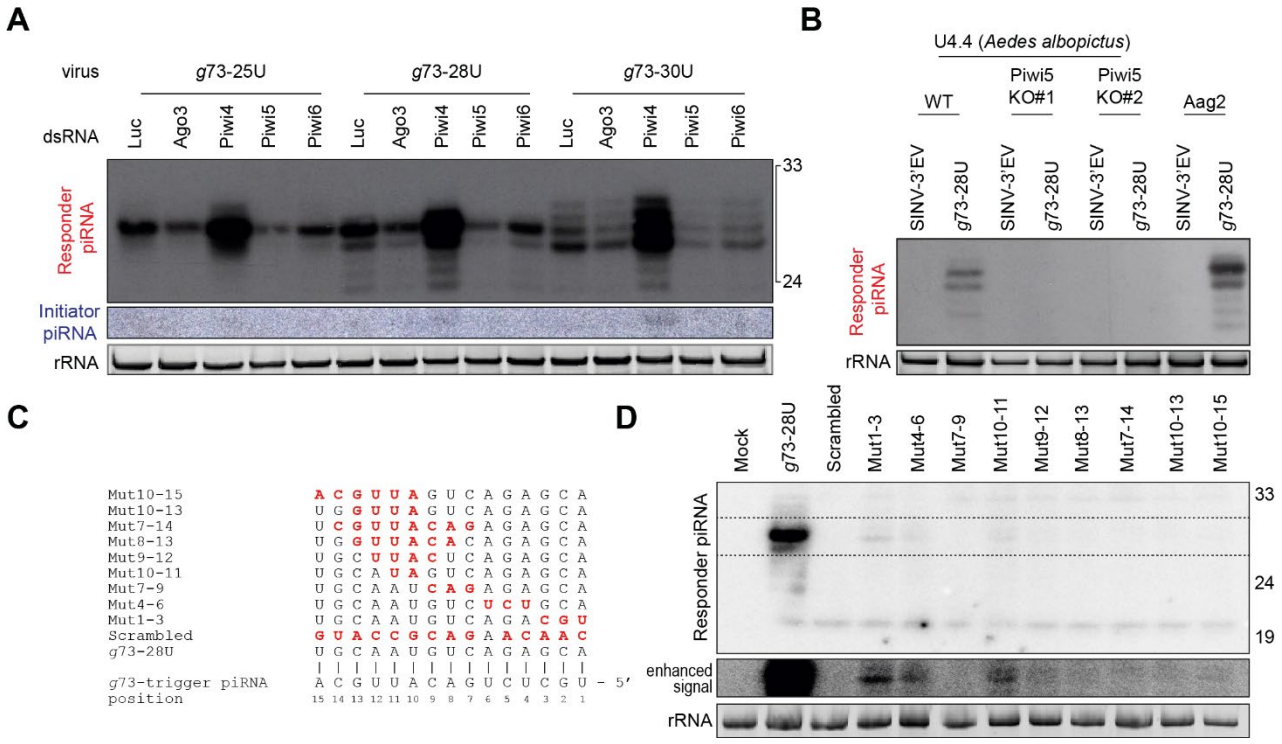


Figure 3. Responder vpiRNAs are produced through ping-pong mediated slicing
(A) Northern blot analysis of viral responder and *g*73-initiator piRNA levels in Aag2 cells in which indicated PIWI proteins are depleted by dsRNA transfection. EtBr stained-rRNA serves as loading control.
(B) Northern blot analysis of responder piRNA production in wildtype (WT) and Piwi5 knockout (KO) *Ae. albopictus* U4.4 cells and *Ae. aegypti* Aag2 cells infected with the *g*73-28U virus or a virus that does not express the reporter locus from the second subgenomic promoter (SINV-3' EV). EtBr-stained rRNA serves as a loading control.
(C) Magnification of the first 15 nucleotides of the *g*73 piRNA-target site, showing the mutations in the target site sequence of the various viruses. Red bold font indicates mutated residues that are mismatched with the *g*73 initiator piRNA and the numbers denote positions relative to the *g*73 initiator piRNA 5' end.
(D) Northern blot analysis of responder piRNA production in Aag2 cells infected with indicated (mutant) viruses. Responder piRNAs were detected using the 'minimal responder' probe that hybridizes to the last 18 nt at the 3' end of the responder piRNA (indicated in yellow in (C)), which are identical for all viruses. The dashed box denotes an area for which the contrast was adjusted to enhance weak responder piRNA signals (enhanced signal – middle panel). EtBr stained rRNA serves as loading control (bottom panel).

responder piRNA production was reduced upon knockdown of genes encoding the ping-pong partners Ago3 and Piwi5 and, to a lesser extent Piwi6 (Figure 3A, S4A). Surprisingly, responder piRNA production was increased upon Piwi4 knockdown, which may be explained by increased levels of the *g*73-derived initiator piRNA in Piwi4 knockdown (Figure 3A). The Piwi5-dependency of *g*73 responder piRNA production was further validated in Piwi5 knockout *Aedes albopictus* U4.4 cells (Figure 3B, S4B).

We next investigated targeting requirements for responder piRNA production. To this end, we introduced target site mutations into the seed region (nt 2-8) and around the putative slice site (nt 10-11, Figure 3C, S4C). Responder piRNA production was strongly depleted in viruses in which mutations were introduced in the seed sequence (Mut 1-3, Mut 4-6 and Mut 7-9), compared to a virus bearing the intact target site (*g*73-28U, Figure 3D). Similarly,

introducing mismatches around the slice site resulted in strongly reduced responder piRNA production. As viral RNA levels are virtually unchanged between all viruses, reduced responder piRNA production cannot result from differences in the amount of available substrate (Figure S4D). Weak responder piRNA production was observed in two seed mutants (Mut 1-3 and Mut 4-6) as well as one slice site mutant (Mut10-11) (Figure 3D). These data suggests that low level piRNA slicing may occur even in the absence of full complementarity in the seed region or the slice site, in line with findings in *Drosophila* (9) and mouse (39).

Zuc-mediated endonucleolytic cleavage defines piRNA 3'ends

Our previous results suggest the involvement of a Zucchini (Zuc)-like endonuclease in the generation of responder piRNAs 3' ends and trailer piRNA 5' ends. Hence, we used our viral responder piRNA reporter to identify the mosquito ortholog of Zuc. As the catalytic activity of Zuc lies in its phospholipase D (PLDc_2)-domain, we first identified all *Ae. aegypti* PLDc_2-domain containing proteins. The neighbor joining tree shows that AAEL011385 has the highest degree of similarity to the various *Z*ucchini orthologs (Figure 4A). The protein encoded by this gene contains a fully conserved catalytic H(X)K(X4)D (HKD)-motif (Figure S5A), suggesting that its endonucleolytic activity is conserved. Moreover, akin to Zuc orthologs in various other species (5,6), AAEL011385 localizes to the mitochondria in Aag2 cells (Figure 4B).

To our surprise, we found that AAEL011385 contains a sizeable insertion directly downstream of the catalytic HKD-motif (Figure S5A). Similarly, *Drosophila* Zuc also contains a (much smaller) insertion in the same location relative to mouse mitoPLD, suggesting it is a variable region beyond *Aedine* mosquitoes. This non-conserved region is part of a helix that sticks out of the Zuc core structure (36). Homology detection revealed that the large insertion is conserved in the *Culicidae* family, but not in *Drosophila*, suggesting it is a mosquito-specific insertion. Moreover, during cloning of the AAEL011385 gene, we found that the size of this insertion is increased in Aag2 cells (32 additional amino acids; Figure S5A).

We next aimed to functionally validate AAEL011385 as the ortholog of *DmZuc*. Indeed, knockdown of AAEL011385 in Aag2 cells resulted in a broader size distribution of viral responder piRNAs (Figure 4C, Figure S5B), which is consistent with reduced sharp endonucleolytic 3' end generation. Knockdown of the other PLDc_2 domain containing

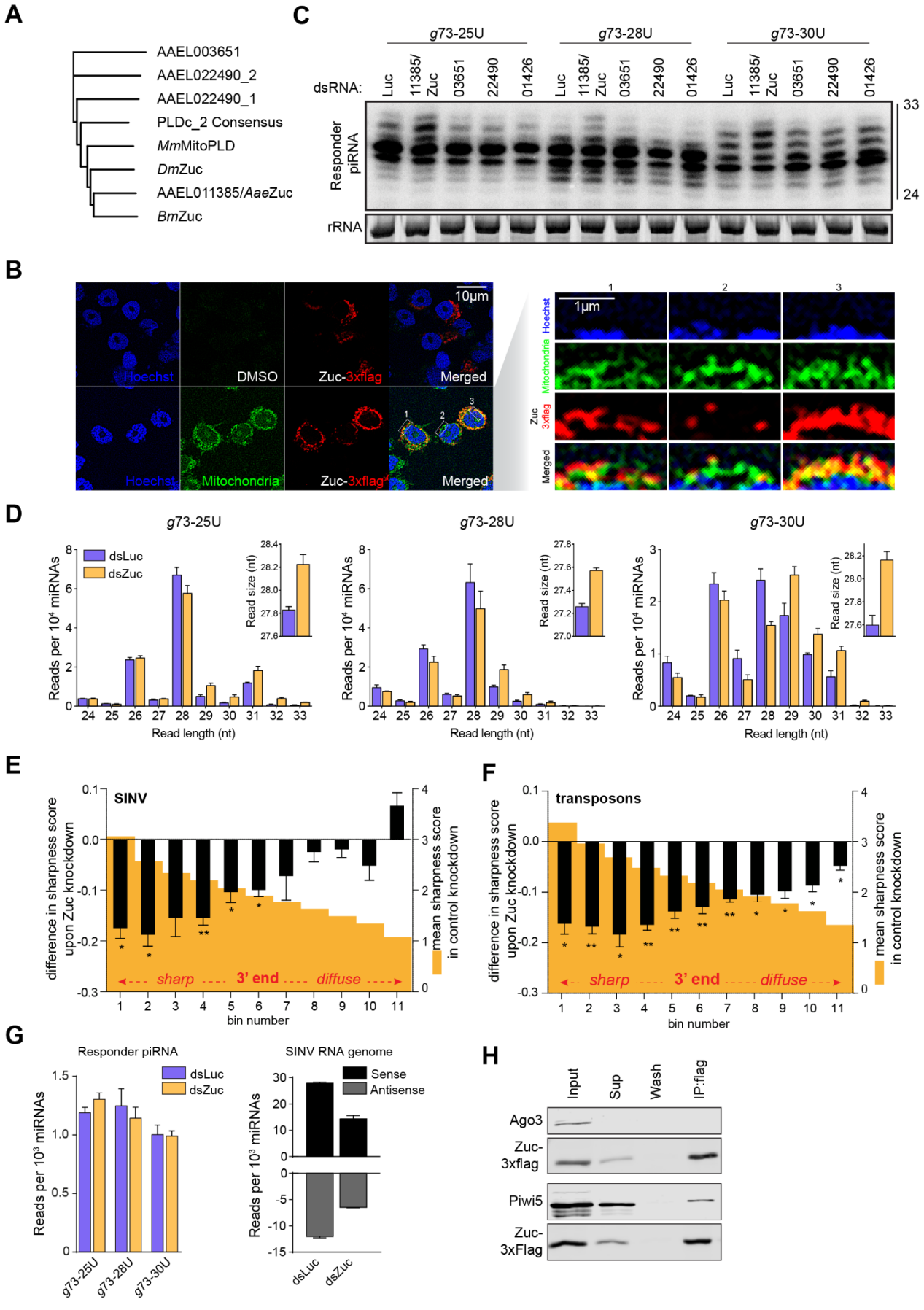


Figure 4. Zuc-mediated endonucleolytic cleavage defines vpiRNA 3'ends (previous page)
(A) Neighbor joining tree based on PLDc_2 domains from the PFAM database, *Ae. aegypti* PLDc_2 domain containing proteins and Zuc orthologs from *Drosophila (Dm* Zuc*),* silkworm (*Bm*Zuc) and mouse (*Mm*MitoPLD).
(B) Confocal microscopy images of Aag2 cells expressing C-terminally 3xflag tagged Zuc. Mitochondria were stained using Mitoview green. On the right, enlargements of the areas indicated by dashed boxes in the image on the left are shown, with the nuclei oriented at the bottom.
(C) Northern blot analysis of viral responder piRNA production in Aag2 cells upon dsRNA-mediated knockdown of *Ae. aegypti* PLDc_2 domain containing proteins and the PARN ortholog AAEL001426. Numbers indicate the VectorBase gene identifiers (without the AAEL0 prefix). The 24 and 33 nt size markers are inferred from an EtBr stained small RNA marker and rRNA stained. EtBr served as a loading control (same for (F)).
(D) Viral responder piRNA (11585(+)) size distribution as determined in small RNA deep-sequencing libraries from dsLuc and dsZuc treated cells. Read counts were normalized to the number of miRNAs in each library. The inset shows the average responder piRNA read size in Luc- and Zuc knockdown libraries. Bars and whiskers represent mean and SD, respectively.
(E) Read count of the responder piRNA in the reporter locus (left) and to the SINV genomic RNA (right), which is common for the three reporter viruses, in dsLuc and dsZuc treated Aag2 cells. Bars and whiskers show the mean +/- SD (left graph) and mean +/- SEM (right graph).
(F) A sharpness score was attributed to all viral piRNAs upstream of the artificial reporter cassette (see materials and methods). The maximum score ($\log_2(14) = 3.81$) would be reached if 100% of reads that map to an individual piRNA had the same length. For each of the different reporter viruses the sharpness score for the top most 275 piRNAs were analyzed upon control and Zuc knockdown. The piRNAs were ranked according to the score in the control condition and the means within 11 bins of 25 piRNAs were determined (orange shade). The difference of piRNA score upon Zuc knockdown is plotted as mean ± SEM. A two-sided student's t-test was applied to each bin, to assess whether its mean was significantly different from zero. The null hypothesis was rejected at $p<0.05$. * $P < 0.05$ and ** $P < 0.01$
(G) Same as (F), but for transposon piRNAs.
(H) Western blot of indicated PIWI proteins in Zuc-3×flag immunoprecipitations (IP) in Aag2 cells.

proteins (AAEL003651 and -22490) did not affect responder piRNA size (Figure 4C, Figure S5B), despite very efficient knockdown (88-94% and 96-98% for AAEL003651 and -22490, respectively, compared to the modest, 58-71% knockdown of AAEL011385/Zuc; Figure S5C). Because of the reduction in responder piRNA 3' end sharpness upon AAEL011385 knockdown and its mitochondrial localization, we formally identify AAEL011385 as the functional Zuc ortholog in *Ae. aegypti*.

Small RNA deep-sequencing of dsZuc-treated Aag2 cells recapitulated the viral responder piRNA size distribution seen by northern blotting, with a general increase in piRNA length upon knockdown of Zuc (Figure 4D). We propose that upon Zuc knockdown, the viral RNA is cleaved downstream of the uridine residue, either by a hitherto unknown endonuclease or by piRNA-PIWI ribonucleoprotein complexes, as in *Drosophila* (12). Subsequently, the responder pre-piRNA may be trimmed, giving rise to mature piRNAs with more diffuse 3' ends in a Zuc-independent manner.

To study the effects of Zuc knockdown on the 3' end sharpness of viral piRNAs outside of the reporter locus in more detail, we assigned a sharpness score to every individual piRNA, defined by a shared 5' end (see materials and methods for details). As expected, Zuc knockdown significantly reduced sharpness scores, especially of those piRNAs that in control conditions had the sharpest 3' ends and are therefore likely to be the most prominent Zuc substrates (Figure 4E). The same effect was observed for piRNAs that mapped to transposon sequences (Figure 4F). Moreover, Zuc knockdown resulted in an overall increase in size of

piRNAs produced from sense as well as antisense substrates of various origins, including retrotransposon subfamilies, mRNAs and viral RNAs (Figure S5F), suggesting that Zuc is involved in the formation of 3' end of piRNAs from all substrates.

We next assessed the global effects of Zuc depletion on vpiRNA levels. While the abundance of the *g*73-triggered artificial responder piRNA was largely unaffected upon Zuc knockdown (Figures 4G-left), Zuc depletion reduced the overall vpiRNA production from the SINV genomic and subgenomic RNA, which is common to the g73-25U, -28U and -30U viruses (Figure 4G-right).

At the heart of the ping-pong amplification machinery in *Ae. aegypti* are two PIWI proteins: Ago3, which predominantly binds sense piRNAs, and Piwi5, which mainly associates with antisense piRNAs (24,34). In *Drosophila*, a strong interaction between Zuc and Aub (the ortholog of Piwi5), as well as weak associations between Zuc and Ago3/Piwi have been observed (40-42). We thus evaluated the interaction between Zuc and the somatic *Aedes* PIWI proteins. While we readily detect Piwi5 in Zuc immunoprecipitates, we did not observe interaction of Zuc with Ago3 (Figure 4H), nor with Piwi4 and -6 (Figure S5E). The interaction between Zuc and Piwi5 may underlie the comparatively stronger +1U bias observed for piRNAs derived from antisense transcripts.

A subset of responder piRNAs undergoes Nibbler-mediated trimming

In *Drosophila*, piRNA 3' ends are generated by the concerted activities of two enzymes: the endonuclease Zuc (7-9) and the 3' – 5' exonuclease Nbr (12-14). Aside from its role in piRNA 3' end formation, Nbr is required for the trimming of microRNAs (miRNAs) (43,44). We set out to identify the functional *Ae. aegypti* Nbr ortholog by predicting all DEDDy-type 3' - 5' exonuclease domain-containing proteins, which were used in a phylogenetic analysis along with *Drosophila* Nbr (*DmNbr)*. This analysis identified AAEL005527 as a one to one ortholog of *DmNbr* (Figure 5A). To verify that AAEL005527 is indeed the functional orthologue of *Drosophila* Nbr, we assessed the effect of its depletion on trimming of two miRNA with heterogeneous 3' ends in *Ae. aegypti*: miR-34-5p and miR-184 (45,46). It has previously been shown that miR-34-5p undergoes extensive Nbr-mediated trimming in *Drosophila* (43,44). Indeed, knockdown of AAEL005527 significantly reduced trimming of both miR-34-5p and miR-184 (Figure 5B), confirming that AAEL005527 is indeed the functional ortholog of *Drosophila* Nbr.

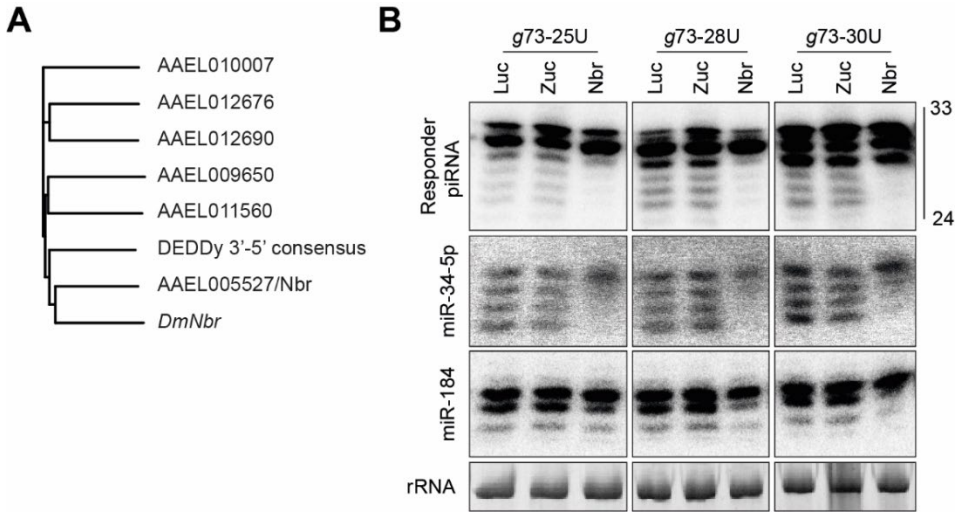


Figure 5. A subset of responder piRNAs undergoes Nbr-mediated trimming
(A) Neighbor joining tree based on the DEDDy 3' – 5' exonuclease domains identified in *Ae. aegypti*, along with the DEDDy consensus sequence (CDD) and the exonuclease domain of *DmNbr*.
(B) Northern blot analyses of viral responder piRNAs in Aag2 cells in which Zuc and *Aae*Nbr knockdown was established. miR-34-5p and miR-184 are two miRNAs with heterogenous 3' ends in *Ae. aegypti*. mir-34-5p has been shown to undergo Nbr-mediated trimming in *Drosophila* (43,44).

To evaluate the role of exonucleolytic trimming on responder vpiRNA 3' end formation in *Ae. aegypti*, we combined Nbr knockdown with infection using the *g*73-targeted reporter viruses. Knockdown was efficient and did not have major effects on viral RNA levels (Figure S6A-B). Interestingly, for all viruses tested, Nbr knockdown results in a specific loss of shorter (<27 nt) responder piRNA isoforms, but had no effect on larger isoforms (Figure 5B, Figure S6C). A similar reduction of shorter piRNA isoforms upon Nbr knockdown has previously been observed in *Drosophila* (13,14). These findings suggest that mosquito Nbr may trim pre-piRNAs generated through a Zuc-mediated endonucleolytic cut, but that only a minor fraction of such pre-piRNAs undergo trimming. Nbr-mediated trimming may be a general hallmark of aging piRNAs, as has recently been proposed for miRNAs (47).

The 3' - 5' exonucleases PNLDC1 and PARN-1 are responsible for trimming of piRNA 3' ends in *B. mori* and *C. elegans*, respectively (48,49). While PNLDC1 is not conserved in *Ae. aegypti*, a clear mosquito ortholog of PARN-1 can be identified: AAEL001426. Knockdown of this gene however, had no effect on responder piRNA 3' end formation in our viral reporter system (Figure 4C, Figure S5B).

Targeting by an endogenous piRNA triggers trailer piRNA production.

While *Ae. aegypti* displays strong signatures of phased piRNA production (7), it is currently unknown whether *de novo* produced RNA from cytoplasmic viruses is processed similarly

through piRNA phasing. This is especially interesting as the genomes of *Ae. aegypti* and *Ae. albopictus* mosquitoes contain large amounts of endogenous viral elements (EVEs). These non-retroviral sequence elements are enriched in piRNA clusters and, accordingly, give rise to abundant piRNAs (26), which may guide the slicing of cognate RNA from acutely infecting viruses. It has been recently shown that piRNAs derived from EVEs are indeed able to target and inhibit newly infecting viruses (31-33), yet, it remains unknown whether piRNA phasing may serve as a mechanism to expand the vpiRNA pool after an initial cleavage by an endogenous, possibly EVE-derived, piRNA.

Hence, we employed our viral reporter system to explore whether targeting by a genomically encoded piRNA may trigger phased piRNA biogenesis from an acutely infecting virus. To this end, we introduced an additional non-coding RNA sequence downstream of the *g*73- and FV53-derived initiator piRNA target sites, which we termed the trailer cassette. To direct sequential Zuc-mediated endonucleolytic cleavage, this cassette contains uridine residues at regularly spaced intervals in an RNA sequence that is otherwise devoid of uridines (U interval viruses, schematically shown in Figure 6C and Figure S7C). As a control, these uridines were replaced by adenosine residues to create a trailer cassette completely devoid of uridines (U desert viruses). In Aag2 cells infected with the *g*73 interval virus, we detected the first trailer piRNA using northern blotting (Figure 6A). Interestingly, we also observed the production of the first trailer piRNA in cells infected with the *g*73 desert virus (Figure 6A), suggesting a downstream uridine residue is not required for trailer piRNA production in *Ae. aegypti*. Trailer piRNA production from viruses targeted by the FV53 EVE-derived piRNA was also observed regardless of the presence of uridine residues in the trailer cassette (Figure S7A). Importantly, while viral RNA levels were similar (Figure 6B), no trailer piRNA production was observed in cells infected with viruses containing a scrambled target site or viruses lacking a trailer cassette (*g*73-28/FV25-28), indicating that, as expected, trailer piRNA production depends on initial targeting by the endogenous piRNA.

To assess phased piRNA biogenesis beyond the first trailer piRNA, we deep sequenced small RNAs produced in Aag2 cells infected with interval and desert viruses, as well as their respective scrambled control viruses. Mapping piRNAs to the trailer cassette reveals production of additional piRNAs in cells infected with the *g*73- and FV53-derived piRNA targeted viruses (Figure 6D, Figure S7D). Significantly fewer piRNAs were produced from the trailer cassette in viruses containing a scrambled target site. In contrast, vpiRNA production

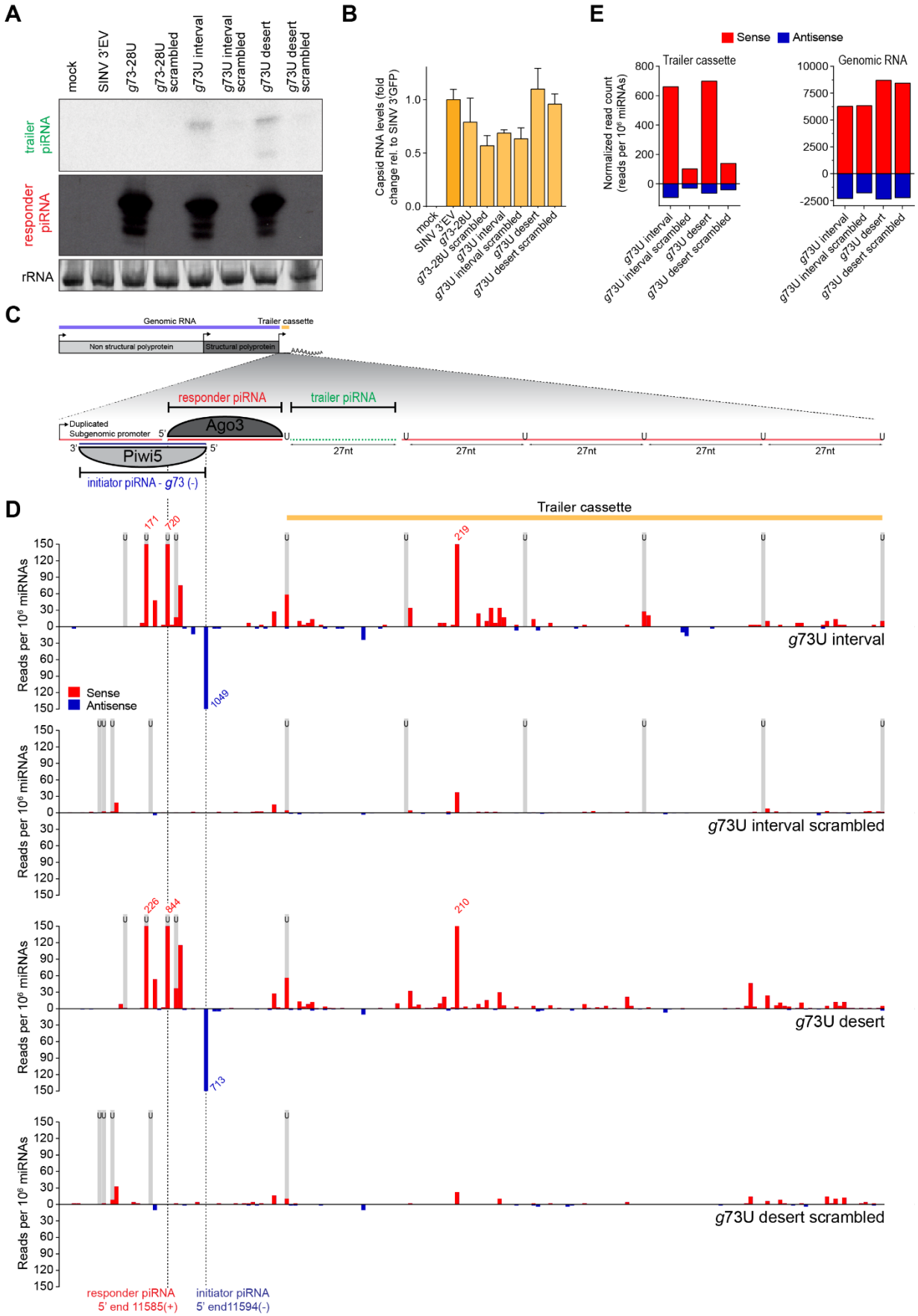


Figure 6. Targeting of viral RNA by an endogenous piRNA triggers trailer piRNA production (previous page).
(A) Northern blot analysis of the production of the first trailer piRNA from indicated viruses. As a control, the target site was scrambled to abolish targeting by the *g*73-derived piRNA. The remainder of the responder piRNA site as well as the trailer cassette are identical to the respective non-scrambled U interval and U desert viruses. As additional controls, a virus bearing an intact target site, but no trailer cassette (g73-28U), and a virus that contains no insert (SINV 3' EV) were used. The general structure of the *g*73 U interval virus is shown schematically in (C). rRNA stained with EtBr serves as a loading control.
(B) RT-qPCR analyses of viral capsid RNA production in Aag2 cells infected with indicated viruses. All data are shown as a fold change relative to cells infected with a control virus lacking any insert (SINV 3' EV). Bars and whiskers show the mean and SD, respectively.
(C) Schematic overview of the *g*73 U interval reporter virus. The inset shows a magnification of the non-coding reporter RNA expressed under control of the second subgenomic promoter. This reporter RNA contains a target site for a Piwi5-bound *g*73-derived initiator piRNA, which guides the production of an Ago3-associated responder piRNA. The downstream sequence makes up the trailer cassette and either contains regularly spaced uridine residues (*g*73U interval) or is completely devoid of uridine residues (*g*73U desert). Indicated in green is the first trailer piRNA, which was detected in (A).
(D) Normalized counts (in reads per 10^6 miRNAs) of piRNA 5' ends mapping to the initiator piRNA target site and trailer cassette of indicated viruses. 5' ends of initiator and responder piRNA are indicated by dashed lines. Numbers indicate counts that exceed the range of the y-axis and black boxes indicate the position of uridine residues. The position of uridine residues is indicated in grey.
(E) Total number of normalized sense (red) and antisense (blue) piRNA-sized (24-33 nt) reads mapping to the trailer cassette (left) and genomic RNA (right) of the indicated viruses. The areas of virus denoted as the trailer cassette and genomic RNA are shown in yellow and blue in (C).

from the SINV genome upstream of the artificial reporter and trailer cassettes is unaltered (Figure 6E, Figure S7E), indicating that there are no differences in sensitivity of these viruses for processing by the vpiRNA biogenesis machinery. As both the pattern and level of piRNA production is highly similar between U interval and U desert viruses (Figure 6D-E, Figure S7D-E), the presence of uridine residues to guide Zuc-mediated endonucleolytic cleavage appears to be dispensable for phased production of piRNAs from viral RNA, at least in the context of this reporter virus.

Conclusion

Altogether, our results indicate that during acute infection with a cytoplasmic RNA virus, genomically encoded piRNAs can initiate piRNA production from viral genomic RNA via the ping-pong amplification loop (Figure S8). The endonucleolytic and exonucleolytic activities of Zuc and Nbr are involved in maturation of the 3' ends of these newly produced piRNAs. Moreover, Zuc is also required for 3' end formation of endogenous and viral piRNAs, whereas Nbr trims miRNAs, like in *Drosophila*. Importantly, cleavage of viral RNA by an endogenous piRNA triggers the diversification of piRNA sequence pool by piRNA phasing. These findings suggest that few cleavage events by individual genome-encoded piRNAs are sufficient to launch a piRNA response of bona fide vpiRNAs.

MATERIALS AND METHODS

Cell culture, transfection and infection of Aag2 and U4.4 cells

Ae. aegypti Aag2 and Ae. albopictus U4.4 cells were grown in Leibovitz's L-15 medium (Invitrogen) supplemented with 10% fetal bovine serum (Gibco), 50 U/ml Penicillin, 50 µg/mL Streptomycin (Invitrogen), 1x Non-essential Amino Acids (Invitrogen) and 2% Tryptose phosphate broth solution (Sigma) at 25°C. Cell lines were maintained by splitting twice weekly according to confluency.

For knockdown experiments, ~1×10^6 cells were seeded in 6-wells plates and allowed to attach for 16-24 hrs. Subsequently, cells were transfected with gene-specific dsRNA and re-transfected 48 hrs later to ensure sustained knockdown. Three hrs after the second transfection, cells were infected with indicated viruses at multiplicity of infection (MOI) of 0.1 and RNA was harvested 72 hrs post infection. In experiments where no knockdown was performed, cells were infected 16-24 hrs after seeding at MOI = 0.1, followed by RNA extraction at 72 hrs post infection

Generation of reporter viruses

Target sites for initiator piRNAs and reporter responder piRNA were introduced into the recombinant Sindbis virus backbone to be expressed from a second subgenomic promoter. The previously described SINV-GFP (50,51) was digested using XbaI to remove the GFP-gene. The reporter locus was subsequently introduced downstream of the duplicated subgenomic promoter by ligation of annealed oligo's (see below), making use of the overhangs generated by the XbaI enzyme (initiator piRNA target site in red, downstream uridine position is indicated in **bold** font).

Primer name	Nucleotide sequence (5' - 3')
FW-FV53-25/28U	CTAGAGACCAGGACAGGCAATGAGACTGAGTACACGATTGAGGAGGCGCGAGCACAGAC**T**AGAACAG CAGAACGACGAGCAAAGGCGGCCGCAGCCT
RV-FV53-25/28U	CTAGAGGCTGCGGCCGCCTTTGCTCGTCGTTCTGCTGTTCTAGTCTGTGCTCGCGCCTCCTCAATCGT GTACTCAGTCTCATTGCCTGTCCTGGTCT
FW-FV53-28/31U	CTAGAGACCAGGACAGGCAATGAGACTGAGTACACGATTGAGGAGGCGCGAGCACAGACCAG**T**ACA GCAGAACGACGAGCAAAGGCGGCCGCAGCCT
RV-FV53-28/31U	CTAGAGGCTGCGGCCGCCTTTGCTCGTCGTTCTGCTGTACTGGTCTGTGCTCGCGCCTCCTCAATCGT GTACTCAGTCTCATTGCCTGTCCTGGTCT
FW-FV53-30/33U	CTAGAGACCAGGACAGGCAATGAGACTGAGTACACGATTGAGGAGGCGCGAGCACAGACCAGAA**T**AG CAGAACGACGAGCAAAGGCGGCCGCAGCCT
RV-FV53-30/33U	CTAGAGGCTGCGGCCGCCTTTGCTCGTCGTTCTGCTATTCTGGTCTGTGCTCGCGCCTCCTCAATCGT GTACTCAGTCTCATTGCCTGTCCTGGTCT
FW-*g*73-25U	CTAGAGACCAGGACAAAGCGCACATACAGTGCAATGTCAGAGCAGGCGCGAGCACAGAC**T**AGAACAG CAGAACGACGAGCAAAGGCGGCCGCAGCCT
RV-*g*73-25U	CTAGAGGCTGCGGCCGCCTTTGCTCGTCGTTCTGCTGTTCTAGTCTGTGCTCGCGCCTGCTCTGACAT TGCACTGTATGTGCGCTTTGTCCTGGTCT
FW-*g*73-28U	CTAGAGACCAGGACAAAGCGCACATACAGTGCAATGTCAGAGCAGGCGCGAGCACAGACCAG**T**ACAG CAGAACGACGAGCAAAGGCGGCCGCAGCCT
RV-*g*73-28U	CTAGAGGCTGCGGCCGCCTTTGCTCGTCGTTCTGCTGTACTGGTCTGTGCTCGCGCCTGCTCTGACA TTGCACTGTATGTGCGCTTTGTCCTGGTCT
FW-*g*73-30U	CTAGAGACCAGGACAAAGCGCACATACAGTGCAATGTCAGAGCAGGCGCGAGCACAGACCAGAA**T**AG CAGAACGACGAGCAAAGGCGGCCGCAGCCT
RV-*g*73-30U	CTAGAGGCTGCGGCCGCCTTTGCTCGTCGTTCTGCTATTCTGGTCTGTGCTCGCGCCTGCTCTGACAT TGCACTGTATGTGCGCTTTGTCCTGGTCT

To introduce U-interval and U-desert reporter sequences, the *g*73-28U and FV53-25/28U viruses were digested using NotI followed by ligation of annealed oligo's (see below).

Primer name	Nucleotide sequence (5' - 3')
FW-U-interval	GGCCTCGCAAGCACAAAGCCGAAGCAGACGAG**T**CGAACCAGGCAAGCAAAGCGACAACGA**T**AAGCA AAGCAACGGCAAAGCAAGCCAA**T**AACGAAAACAGCAAAGCAGACGAACAC**T**GC

RV-U-interval	GGCCGCAGTGTTCGTCTGCTTTGCTGTTTTCGTTATTGGCTTGCTTTGCCGTTGCTTTGCTTATCGTTG TCGCTTTGCTTGCCTGGTTCGACTCGTCTGCTTCGGCTTTGTGCTTGCGA
FW-U-desert	GGCCACGCAAGCACAAAGCCGAAGCAGACGAGACGAACCAGGCAAGCAAAGCGACAACGAAAAGCA AAGCAACGGCAAAGCAAGCCAAAAACGAAAACAGCAAAGCAGACGAACACAGC
RV-U-desert	GGCCGCTGTGTTCGTCTGCTTTGCTGTTTTCGTTTTTGGCTTGCTTTGCCGTTGCTTTGCTTTTCGTTG TCGCTTTGCTTGCCTGGTTCGTCTCGTCTGCTTCGGCTTTGTGCTTGCGT

Subsequently, viruses were grown as described previously (25).

Mutagenesis of target site mutant viruses

Target site mutations were introduced into the *g*73-28U virus backbone by mutagenesis PCR, using the primers shown below. PCR-products were DpnI-treated and In-fusion (Takara Biotech) was using to circularize the plasmid for transformation. After verification of the sequence by Sanger sequencing, viruses were grown as described previously (25).

Primer name	Nucleotide sequence (5' - 3')
FW_mut1-3	GTCAGACGTGGCGCGAGC
RV_mut1-3	CGCGCCACGTCTGACATTGC
FW_mut4-6	AATGTCTCTGCAGGCGCGAG
RV_mut4-6	GCCTGCAGAGACATTGCACTGT
FW-mut7-9	TGCAATCAGAGAGCAGGCGC
RV-mut7-9	TGCTCTCTGATTGCACTGTATGTGC
FW-mut10-11	CAGTGCATAGTCAGAGCAGGCG
RV-mut10-11	TCTGACTATGCACTGTATGTGCGC
FW-mut9-12	CAGTGCTTACTCAGAGCAGGCGC
RV-mut9-12	TCTGAGTAAGCACTGTATGTGCGCTTTG
FW-mut8-13	CAGTGGTTACACAGAGCAGGCGCG
RV-mut8-13	TCTGTGTAACCACTGTATGTGCGCTTTGTC
FW-mut7-14	CAGTCGTTACAGAGAGCAGGCGCGAG
RV-mut7-14	TCTCTGTAACGACTGTATGTGCGCTTTGTCCTG
FW-mut10-13	CAGTGGTTAGTCAGAGCAGGCGC
RV-mut10-13	TCTGACTAACCACTGTATGTGCGCTTTGTC
FW-mut10-15	CAGACGTTAGTCAGAGCAGGCGCGA
RV-mut10-15	TCTGACTAACGTCTGTATGTGCGCTTTGTCCT

dsRNA production for knockdown experiments

Gene-specific PCR products bearing T7 promoter sequences at both ends were *in vitro* transcribed using T7 polymerase. Either the T7 promoter sequence was introduced directly with the gene-specific PCR, or a universal GC-rich tag was introduced in the first PCR, to which the T7 promoter sequence was added in a second PCR. Oligos use to create the T7-promoter tagged PCR products are shown below (T7 promoter sequence in **bold** font, the universal GC-rich tag in underlined font):

Primer name	Nucleotide sequence (5' - 3')
F-T7-Luc	**TAATACGACTCACTATAGGGAGA**TATGAAGAGATACGCCCTGGTT
R-T7-Luc	**TAATACGACTCACTATAGGGAGA**TAAAACCGGGAGGTAGATGAGA
F-T7-Ago3	**TAATACGACTCACTATAGGGAGA**TGCTTACTCGTGTCGCGTAG
R-T7-Ago3	**TAATACGACTCACTATAGGGAGA**GGCATGGCAGATCCAATACT
F-T7-Piwi4	**TAATACGACTCACTATAGGGAGA**CGTGGAAGTCCTTCTTCTCG
R-T7-Piwi4	**TAATACGACTCACTATAGGGAGA**TGTCAGTTGATCGCTTCTCAA
F-T7-Piwi5	**TAATACGACTCACTATAGGGAGA**GCCATACATCGGGTCAAAAT
R-T7-Piwi5	**TAATACGACTCACTATAGGGAGA**CTCTCCACCGAAGGATTGAA
F-T7-Piwi6	**TAATACGACTCACTATAGGGAGA**CAACGGAGGATCTTCACGAG
R-T7-Piwi6	**TAATACGACTCACTATAGGGAGA**AATCGATGGCTTGATTTGGA
F-GC-AAEL011385/Zuc	GCCCGACGCCAGTGCGTTTCGATCG
R-GC-AAEL011385/Zuc	CGCCTCGGCCCCGTAATGAGCAACCC
R-GC-AAEL003651	GCCCGACGCGTGGTGCCTGCTTCG
F-GC-AAEL003651	CGCCTCGGCTGCCACATTTGTCCTAAGG
F-GC-AAEL022490	GCCCGACGCTTTCGCCCTGTTTCCC
R-GC-AAEL022490	CGCCTCGGCGATCATCACGTATCGGTCC
F-GC-AAEL001426	GCCCGACGCGACGATTCGCCAAACC
R-GC-AAEL001426	CGCCTCGGCCCATTTTCCTCTTCTTTCG

R-GC-AAEL005527/*Aae*Nbr	GCCCGACGCCCCCCGATAGTGAGCC
F-GC-AAEL005527/*Aae*Nbr	CGCCTCGGCCCAAAGCAGCTGTAAATCC
F-T7-universal primer	**TAATACGACTCACTATAGGGAGA**GCCCGACGC
R-T7-universal primer	**TAATACGACTCACTATAGGGAGA**CGCCTCGGC

Generation of Piwi5 knockout U4.4 cells

Description and characterization of Piwi5 knockout U4.4 cells is described elsewhere (manuscript in prep).

Phylogenetic analyses

The PFAM PLDc_2 domain (PF13091) consensus sequence was as input for iterative searches against the *Ae. aegypti* proteome using JackHMMR (https://www.ebi.ac.uk/Tools/hmmer/search/jackhmmer) to identify all *Ae. aegypti* PLDc_2 domain containing proteins. To generate a neighbor-joining tree, the PLDc_2 domain sequences from the identified genes, as well as those from Zuc orthologs in *Drosophila*, silkworm and mouse were aligned using M-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:mcoffee). One of the identified genes, AAEL022490, contains two PLDc_2 domains, both of which were used as separate inputs in the subsequent analysis. The same approach using the DEDDy type 5'-3' exonuclease (cd06141) CDD-consensus sequence was used for the identification *Ae. aegypti* Nbr.

For the alignment of full Zuc proteins, the entire sequence of *Zuc* as annotated on VectorBase (www.vectorbase.org), the Zuc sequence as determined by us from Aag2 cells and the sequences of Zuc orthologs in *Drosophila,* silkworm, mouse were aligned using M-Coffee with default settings and visualized using Jalview 2.11.0. The analyses of conservation of physicochemical properties at each position was performed in Jalview 2.11.0., and is based on the Analysis of Multiply Aligned Sequences (AMAS) method (52).

RNA isolation

Cells were lysed in 1 mL RNA-Solv reagent (Omega Bio-tek), followed by RNA extraction through phase separation and isopropanol precipitation. RNA integrity was evaluated on EtBr stained agarose gels, and RNA concentration was determined using the Nanodrop ND-1000.

Small RNA northern blotting

For small RNA northern blotting, 4-12 µg of RNA was size separated by denaturing urea polyacrylamide (15%) gel electrophoresis and transferred to nylon membranes using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) (53). ^{32}P labelled DNA probes (sequences shown below) were used to detect small RNAs.

Primer name	**Nucleotide sequence (5' - 3')**
g73-responder:	TTCTGGTCTGTGCTCGCGCCTGCTCTGACA
g73-initiator:	AAAGCGCACATACAGTGCAATGTCAGAGCA
FV53-responder:	TTCTGGTCTGTGCTCGCGCCTCCTCAATCG
Trailer piRNA:	CTCGTCTGCTTCGGCTTTGTGCTTGCG
miR34-5p:	CAACCAGCTAACCACACTGCCA
miR184:	GCCCTTATCAGTTCTCCGTCCA
minimal responder:	CTGGTCTGTGCTCGCGCC

To generate the line graphs shown in Supplementary Figure S5B and S5C, northern blot signals were quantified at the center of each lane. Afterwards, peaks were manually aligned to correct for minor variations in size separation between samples. The average of five neighboring pixels was plotted to smoothen the line graph.

Generation of small RNA deep sequencing libraries

Small RNA deep sequencing libraries were generated using the NEBNext Small RNA Library Prep Set for Illumina (E7560, New England Biolabs), using 1µg RNA as input. As piRNAs generally have 2'-*O*-methylated 3' ends, we performed the 3' adapter ligation for 18 hrs at 16°C to enhance the ligation efficiency of small RNAs bearing such modifications. The rest of the library preparation was performed in accordance with the manufacturers' instructions. Libraries were sequenced on an Illumina Hiseq4000 machine by the GenomEast Platform (Strasbourg, France).

Small RNA sequencing bioinformatic analyses

RTA 2.7.3 and bcl2fastq were used for image analysis and base calling, and initial quality control was performed using FastQC. All subsequent manipulations were performed in Galaxy (54). First, the FASTX Clip adapter software was used to clip 3' adapters from the small RNA sequence reads. Subsequently, reads were mapped onto the corresponding recombinant Sindbis virus genomes, transposable element sequences extracted from TEFAM (http://tefam.biochem.vt.edu; downloaded on 10-12-2019), *Ae. aegypti* transcripts (AaegL5.2 gene set, downloaded from VectorBase) and the Phasi Charoen like virus (PCLV) genome as sequenced from Aag2 cells (Genbank accession numbers KU936055, KU936056 and KU936057) (55). Reads were mapped using Bowtie allowing 1 mismatch, except for the mapping to interval and desert viruses and for 3' end sharpness analyses, in which cases no mismatches were allowed. Small RNA libraries were normalized to the number of reads mapping to published pre-miRNA sequences deposited in miRBase v.21 without allowing mismatches.

Size profiles were obtained by counting piRNA lengths after mapping with exception of the size profile of small RNAs obtained after PIWI IP (Figure S2E), for which raw reads prior to mapping were analyzed from our previously published data (34). For genome distribution plots, the number of piRNA 5' ends was determined for each position of the sequence to which small RNAs have been mapped.

For the analyses of piRNA production from the area of the non-coding reporter RNA downstream of the initiator/responder site versus the remainder of the viral genome shown in Figure 6E and Figure S7E, the SINV genome is defined as the area encoding the non-structural and structural polyproteins up to the start of the duplicated subgenomic promoter (nt 1-11385), and the trailer cassette as nt 11610-11753.

Sequence data have been deposited in the NCBI sequence read archive under SRA accession XXXXX.

To generate heat maps of the length profile of individual piRNAs, published small RNA sequencing data was re-analyzed (24). SAM files were converted into interval files and piRNA-sized reads (25-30 nt) were extracted and separated according to the strand they mapped to. For each individual mapped piRNA species, defined by a shared 5' end, the frequency of reads of different length was determined. Only piRNAs that were supported by at least 20 reads were considered for further analyses. The frequency of read lengths per individual piRNA was imported into Multiple experiment viewer (v 4.9.0; tm4) and k-means clustering based on Pearson correlation was performed using the indicated number of clusters and a maximum number of iterations of 500. The 'Construct hierarchical tree' option was enabled.

To analyze nucleotide biases within and downstream of each piRNAs, only piRNA species that were supported by at least 20 reads and had a dominant piRNA length (at least 75% of all mapped reads have the same length) were considered. Combining the Get flanks (v1.0.0)

and Get genomic DNA (v3.0.3) tools, the genomic sequence 5 nt downstream of that dominant piRNA length was extracted. Subsequently, each piRNA species was collapsed to one unique sequence and its corresponding downstream sequence. The sequence logo was generated from FastA files that contained either these unique piRNA or downstream sequences. The piRNA logos were cropped in Adobe illustrator to show nucleotide biases of piRNA positions 1 to 12 and +1 to +5.

To analyze the effect of Zuc knockdown on piRNA 3' ends, small RNAs mapping to the SINV genome up to the subgenomic promotor (nt 1-11385) were considered. For each of the six datasets (3x dsLuc and 3x dsZuc) available for the *g*73-25U virus, piRNA species were extracted by selecting only those piRNA start sites that were supported by at least fifty 25-30 nt reads in the combined six datasets. For these piRNA positions, all reads in the size range of 25-38 nt were then extracted from the original mapped data. As for the heat map analysis, the frequency of piRNA lengths was counted for each individual piRNA species. To filter out noise due to low read count, only the 275 most abundant piRNAs were considered, which corresponded to approximately 150 reads in the combined six datasets. From this frequency distribution, a sharpness score was calculated for each individual piRNA species based on the maximum entropy (all reads have the same length) minus the observed Shannon entropy: $S_{sharp} = S_{max}\text{-}S_{obs} = \log_2(14) - \sum_{n=25}^{38} p_n * \log_2 p_n$ (56). The mean sharpness score for each piRNA was determined for the three control and Zuc knockdown datasets and the 275 piRNAs were then ranked based on mean sharpness score in the control knockdown from high score (sharp 3' end) to lower scores (more diffuse 3' end). Within this ranking an average of the mean sharpness scores was then calculated for 11 bins of 25 individual piRNA positions each, both for the control and the Zuc knockdown. Finally, the difference in sharpness score (ΔS_{sharp}) between control and Zuc knockdown was determined for each bin. The entire analysis was repeated for the six datasets of the g73-28 and g73-30 viruses and the data was combined by calculating the mean + SEM of the ΔS_{sharp} per bin obtained for each type of reporter virus. To determine statistical significance, the ΔS_{sharp} was tested against the null hypothesis that there is no effect of Zuc knockdown on 3' end sharpness scores ($\Delta S_{sharp} = 0$). A two-sided student's t-test was performed per bin and the hypothesis was rejected at $p<0.05$. The sharpness score analysis for transposon mapping piRNAs was performed with the 4400 most expressed piRNAs (400 piRNAs per bin).

RT-qPCR

For RT-qPCR analyses, DNaseI (Ambion)-treated RNA was reverse transcribed using Taqman reverse transcriptase (Life Technologies) and PCR amplified in the presence of SYBR green, using the GoTaq qPCR system (Promega) according to the manufacturers' recommendations. Expression levels of target genes were normalized to the expression of the housekeeping gene lysosomal aspartic protease (LAP) for *Ae. aegypti* samples or Ribosomal Protein L5 (RpL5) for *Ae. albopictus* samples and fold changes were calculated using the $2^{\Delta\Delta CT}$ method (57). The following primers were used:

Primer name	Nucleotide sequence (5' - 3')
qF-SINV-Capsid	CTGGCCATGGAAGGAAAGGT
qR-SINV-Capsid	CCACTATACTGCACCGCTCC
qF-LAP	GTGCTCATTCACCAACATCG
qR-LAP	AACTTGGCCGCAACAAATAC
qF-*Aal*-RpL5	TCGCTTACGCCCGCATTGAGGGTGAT
qR-*Aal*-RpL5	TCGCCGGTCACATCGGTACAGCCA
qF-AAEL011385/Zuc	GATGACCCGTTTGGCC

qR-AAEL011385/Zuc	CATTGGCCACTGCACC
qF-AAEL003651	CCGTTGTGTTCAACAACC
qR-AAEL003651	GAGTTAACGGGACGTAATCC
qF-AAEL022490	CGGAATGCAGCATTCTGATG
qR-AAEL022490	GTTTCAGCTTTTCGTAGAATTTCC
qF-AAEL001426	GACACAACATGCTGTTGG
qR-AAEL001426	TGACGGCTTCGAACAC

Generation of Zuc expression plasmids

The sequence encoding the Zuc protein was amplified from Aag2 cDNA using CloneAmp HiFi PCR Premix (Takara) and cloned into the pAWG vector (The Drosophila Gateway Vector Collection, Carnegie Science) using In-fusion (Takara) technology.
The following primers were used to amplify the vector and insert:

Primer name	Nucleotide sequence (5' - 3')
FW-pAWG-INV	GACCCAGCTTTCTTGTACAAAGTGG
RV-pAWG-INV	GGTGTTTCTATCTCCTTCGAAGCC
FW-11385_C	GGAGATAGAAACACCATGATGTTGGAATCTGTGCCACG
RV-11385_C	CAAGAAAGCTGGGTCCTGCCTATCTCCCCCGATGG

Subsequently, pAW3F-Zuc was generated by inverse PCR of pAWG-Zuc, which introduced the 3×flag tag to replace the eGFP-tag, followed by In-fusion cloning (Takara Biotech). The primers used for this inverse PCR are shown below (underlined sequence makes up the 3×flag tag, double underlined sequence is the 15 nt overlap required for In-fusion cloning):

Primer name	Nucleotide sequence (5' - 3')
FW_pAW3F-11385	ATAAGGACGACGATGATAAGGACTACAAGGACGACGACGATAAGCACCGGTCCACGTGACG
RV_pAW3F-11385	CATCGTCGTCCTTATAATCCTTGTCATCGTCGTCTTTGTAATCGGTGGCGGAGCTCACC

Immunofluorescence

Approximately 5×10^5 Aag2 cells were transfected with 1 µg of the pUb-Zuc-3xflag expression vector using 1 µL X-tremeGENE HP DNA Transfection Reagent (Roche) according to the manufacturer's instructions. 48 hrs later, cells were fixed on coverslips using 4% paraformaldehyde for 10 minutes at room temperature. Fixed cells were permeabilized in PBS-Triton (0.25%) for 10 minutes and blocked in 10% normal goat serum/0.3M Glycine in PBS-Tween (0.1%). Antibody staining was performed using mouse anti-flag (1:200, Sigma, F1804, RRID: AB_262044) for 1 hr at room temperature, followed by goat anti-mouse IgG Alexa fluor 568 (1:200, Invitrogen, A-11004, RRID: AB_2534072) for 1 hr at room temperature. Following antibody staining, mitochondria were stained using 200 µM Mitoview Green (Biotium), according to the manufacturer's instruction. Lastly, Hoechst reagent was used to stain the nuclei and the coverslips were mounted onto microscope slides for imaging using the Zeiss LSM900 confocal microscope. In between all steps during the staining procedure, cells were washed three times with PBS.

Co-immunoprecipitation and western blot

For immunoprecipitation of Zuc, ~4.5×10^6 Aag2 cells were seeded and after ~16 hrs, 15 µg of the pAW3F-Zuc was transfected using X-tremeGENE HP DNA Transfection Reagent (Roche) according to the manufacturer's instructions. Cells were lysed 48 hrs after transfection using 300 µL lysis buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% Igepal CA-630 [Sigma], 10% Glycerol, 1x cOmplete Protease Inhibitor [Roche], 1 mM PMSF). After incubation for 1 hr at 4°C with end-over-end rotation, lysates were centrifuged

for 30 min at 15000 × g, 4°C. The supernatant was snap-frozen in liquid nitrogen and stored at -80°C for later use.

For immunoprecipitation, 15 µL M2-Flag bead slurry (Sigma) was equilibrated in lysis buffer, and, along with 450 µL dilution buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1x cOmplete Protease Inhibitor [Roche], 1 mM PMSF), added to the thawed lysate. After incubation for 2 hrs at 4°C with end-over-end rotation, beads were washed thrice using 500 µL dilution buffer, before harvesting immunoprecipitates by boiling for 10 min in 20 µL 2× sample buffer (120 mM Tris/Cl pH 6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol). After boiling, samples were diluted by adding 20 µL lysis buffer.

Samples were resolved on 10% polyacrylamide gels and blotted to nitrocellulose membranes. The following antibodies, generated in our laboratory, were used for western blotting: rabbit-anti-Ago3, -Piwi4, -Piwi5 and -Piwi6 (all at 1:500) (34,58). Additionally, mouse anti-flag (1:1000, Sigma, F1804, RRID: AB_262044) was used. Secondary antibodies were goat-anti-rabbit-IRdye800 [Li-cor; 926-32211, RRID: AB_621843] and goat-anti-mouse-IRdye680 [926-68070, RRID: AB_10956588].

ACKNOWLEDGEMENTS

We thank the members of the laboratory for critical discussion of this manuscript. Furthermore, we thank Bas Pennings for his help with cloning experiments. This work was financially supported by a Consolidator Grant from the European Research Council under the European Union's Seventh Framework Programme (grant number ERC CoG 615680) and a VICI grant from the Netherlands Organization for Scientific Research (grant number 016.VICI.170.090).

SUPPLEMENTARY MATERIAL

Figure S1

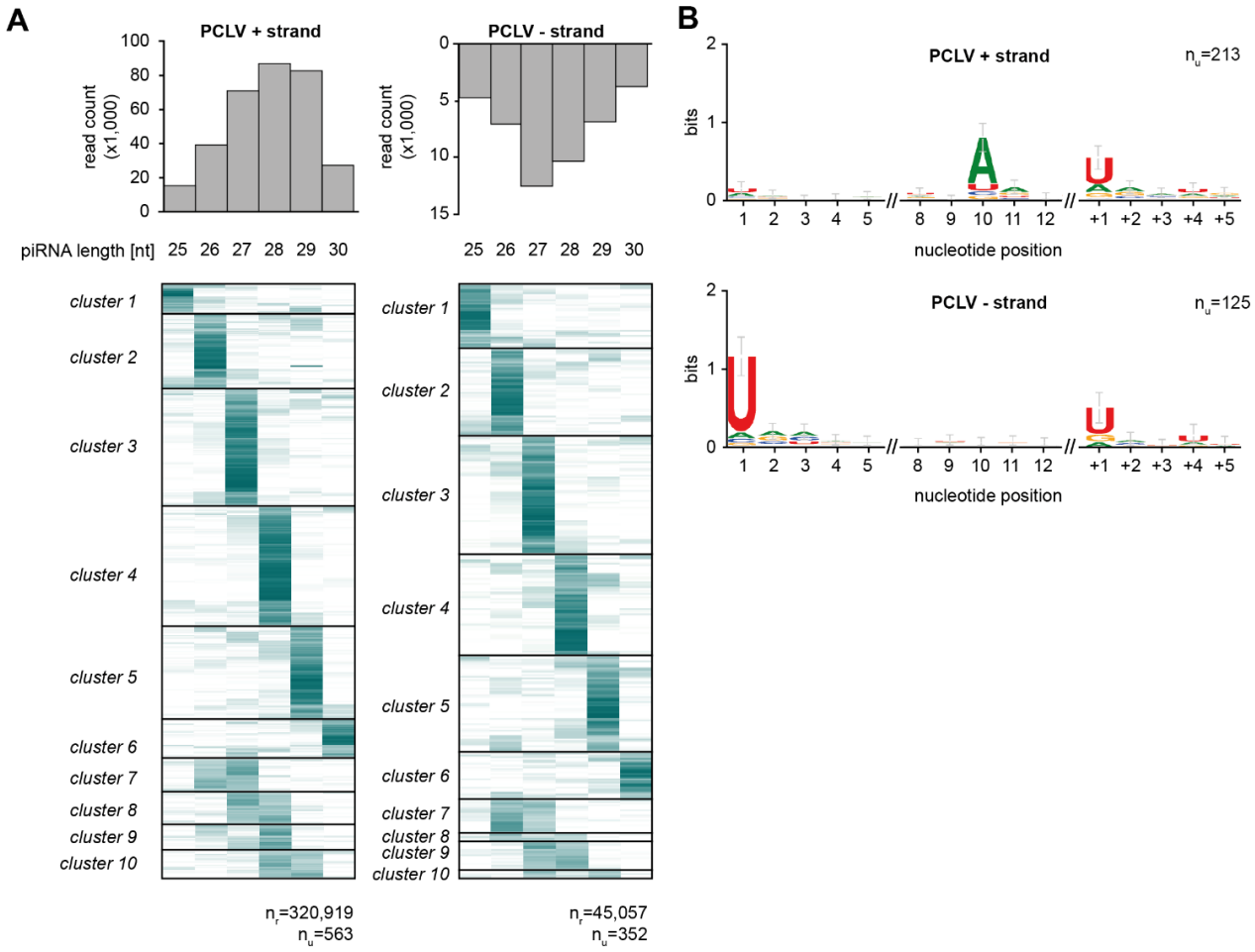


Figure S1. PCLV-derived piRNAs have sharp 3' ends
(A) Heat map showing the relative size distribution of individual piRNAs (defined by a shared 5' end) mapping to the genome of Phasi Charoen-like virus (PCLV). Shades of blue indicate the percentage of reads contributing to the indicated read length, white represents absence of reads of a specific size. The number of unique piRNAs (n_u) and the number of reads (n_r) that underlie the heat map are indicated. A minimum of 20 reads/unique piRNA was required to be included in the analysis.
(B) Nucleotide biases for the indicated nucleotide positions of PCLV-derived piRNAs and a downstream region. Only piRNAs from (A) that had a dominant length (determined by at least 75% of reads) were considered in this analysis and all reads were collapsed to unique sequences. The downstream region (+1 until +5) shows the bias at the genomic region downstream of the dominant piRNA sequences.

611 **Supplementary text and Figure S2**

612 **An EVE-derived piRNA has the potential to trigger responder piRNA production.**

613 Apart from the *g*73 piRNA triggered reporter virus described in the main text, we generated
614 a second set of viruses bearing target sites for an abundant Piwi5-associated piRNA derived
615 from an EVE of flaviviral origin (FV53 (26), Figure S2A). As the target site for the FV53-
616 derived initiator piRNAs can be recognized by two piRNA isoforms that align at their 3' end
617 and differ 3 nt in size, target cleavage may result in the production of two Ago3-bound
618 responder piRNAs isoforms that have an offset of 3 nt (Inset in Figure S2A).
619 Similar to the *g*73 target site bearing viruses (Figure 2B), responder piRNA 3' end sharpness
620 is moderately reduced as a function of increased distance between the initiator piRNA
621 cleavage site and the first downstream uridine residue (Figure S2B), suggesting a role for
622 exonucleolytic trimming in responder piRNA maturation.
623 For all viruses tested here, we observe abundant responder piRNAs as well as the production
624 of a first trailer piRNA (indicated with yellow arrowheads in Figure S2C). As the trailer piRNA
625 5' ends perfectly overlap with the uridine residue, we propose that this residue guides
626 cleavage of the viral RNA, simultaneously generating the responder pre-piRNA 3' end and
627 the trailer piRNA 5' end.
628 Aside from these responder and trailer piRNAs, additional abundant sense piRNAs are
629 produced from the sequence upstream of the initiator piRNA slice site (Figure S2C),
630 suggesting that other, partially complementary, initiator piRNAs may target the sequence and
631 instruct production of these additional sense piRNAs.
632

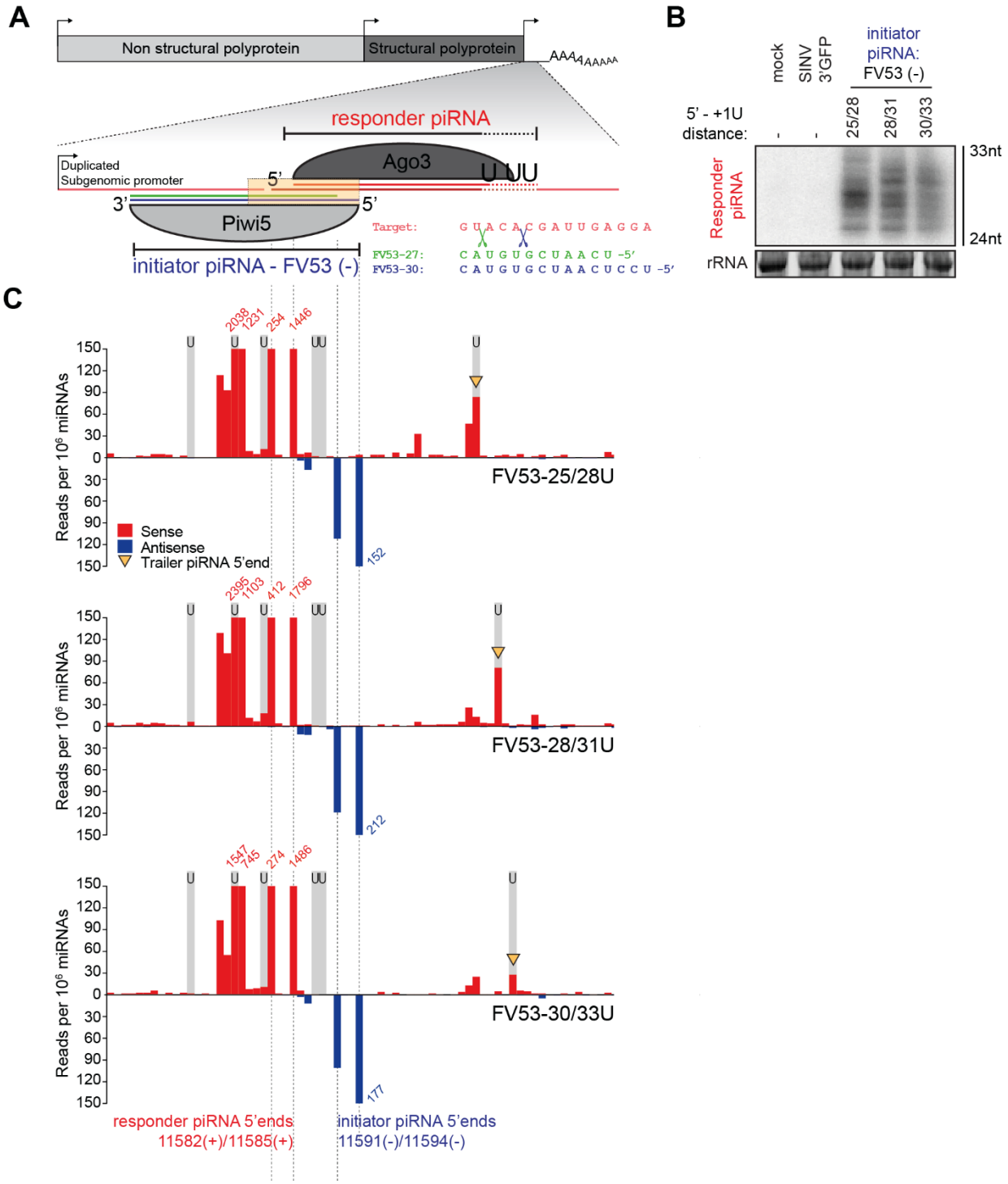


Figure S2. An EVE-derived piRNA triggers responder piRNA production from an acutely infecting virus RNA
(A) Schematic depiction of the SINV-based viral reporter system in which responder piRNA production is triggered by an initiator piRNA derived from an endogenous viral element (FV53). As the Piwi5-associated FV53 initiator piRNA is expressed as two isoforms that align at their 3' end (blue and green lines), endonucleolytic cleavage may generate two Ago3-bound responder piRNAs (red lines) that differ 3 nt in size. The inset shows a part of the viral target RNA (red), the 5' ends of the two FV53-piRNA isoforms (27-mer: green and 30-mer: blue) and the scissors represent their respective slice sites (green and blue).
(B) Northern blot analysis of responder piRNAs produced in Aag2 cells infected with indicated FV53-targeted reporter viruses. 24 and 33 nt size markers are inferred from EtBr staining of a small RNA marker, and EtBr stained rRNA serves as loading control.
(C) Visualization of 5' ends of sense (red) and antisense (blue) piRNAs (24-33 nt) mapping to the non-coding reporter RNA sequence. Predicted initiator and responder piRNA 5' ends are indicated by dashed lines and positions of uridine residues on the sense strand are denoted by light grey shading. Red and blue numbers indicate read counts (per 10^6 miRNAs) at positions where they exceed the y-axis range and yellow arrowheads indicate 5' ends of putative trailer piRNAs.

633 **Figure S3**

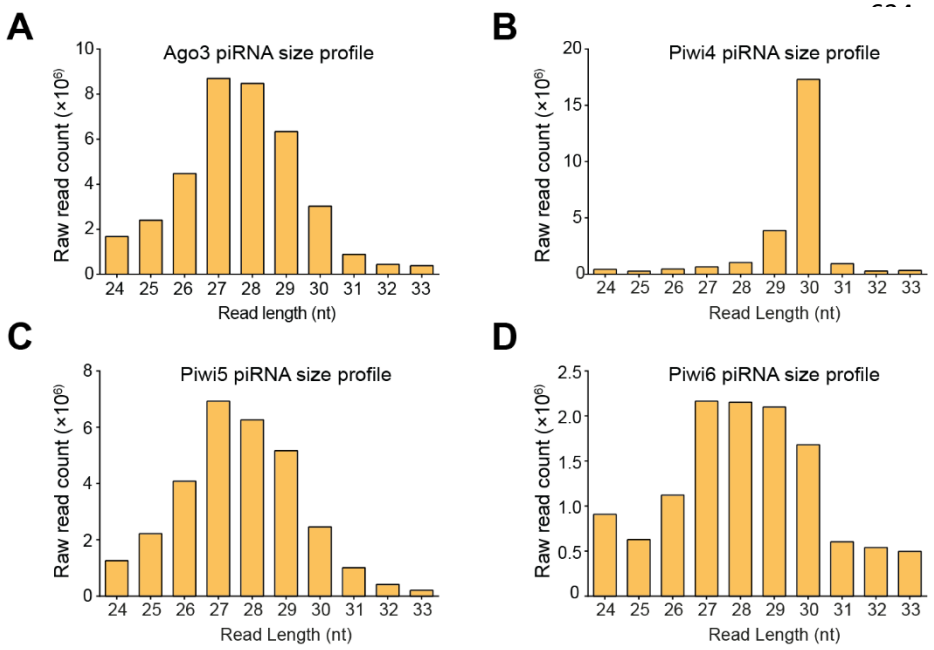


Figure S3. Size preferences of somatically expressed PIWI proteins
Size distribution of piRNA-sized (24-33 nt) small RNA reads in Ago3- (A), Piwi4- (B), Piwi5- (C), and Piwi6-IP (D) sequencing libraries. Piwi4-IP is dominated by two very abundantly expressed piRNAs (tapiR1/2- 30 and 29 nt in size, respectively), that are involved in the degradation of maternally provided transcripts during embryonic development in mosquitoes (58). Counts are raw, unmapped reads from our previously published data (34,58).

635 **Figure S4**

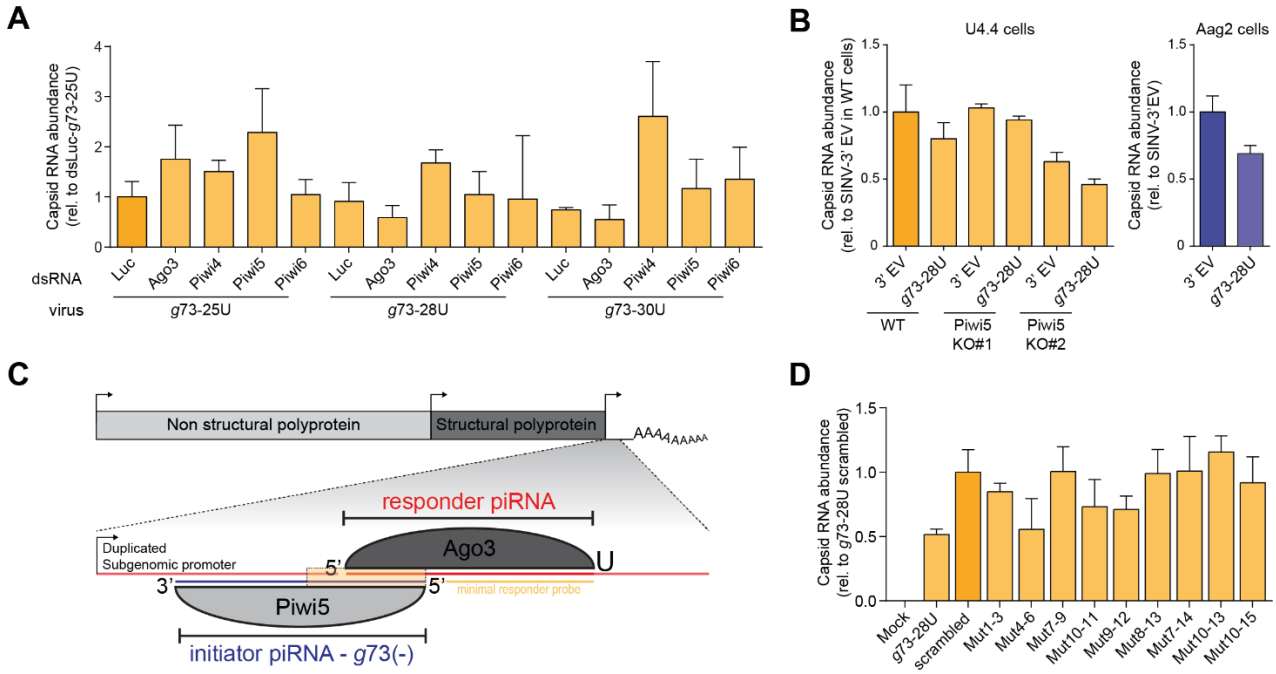
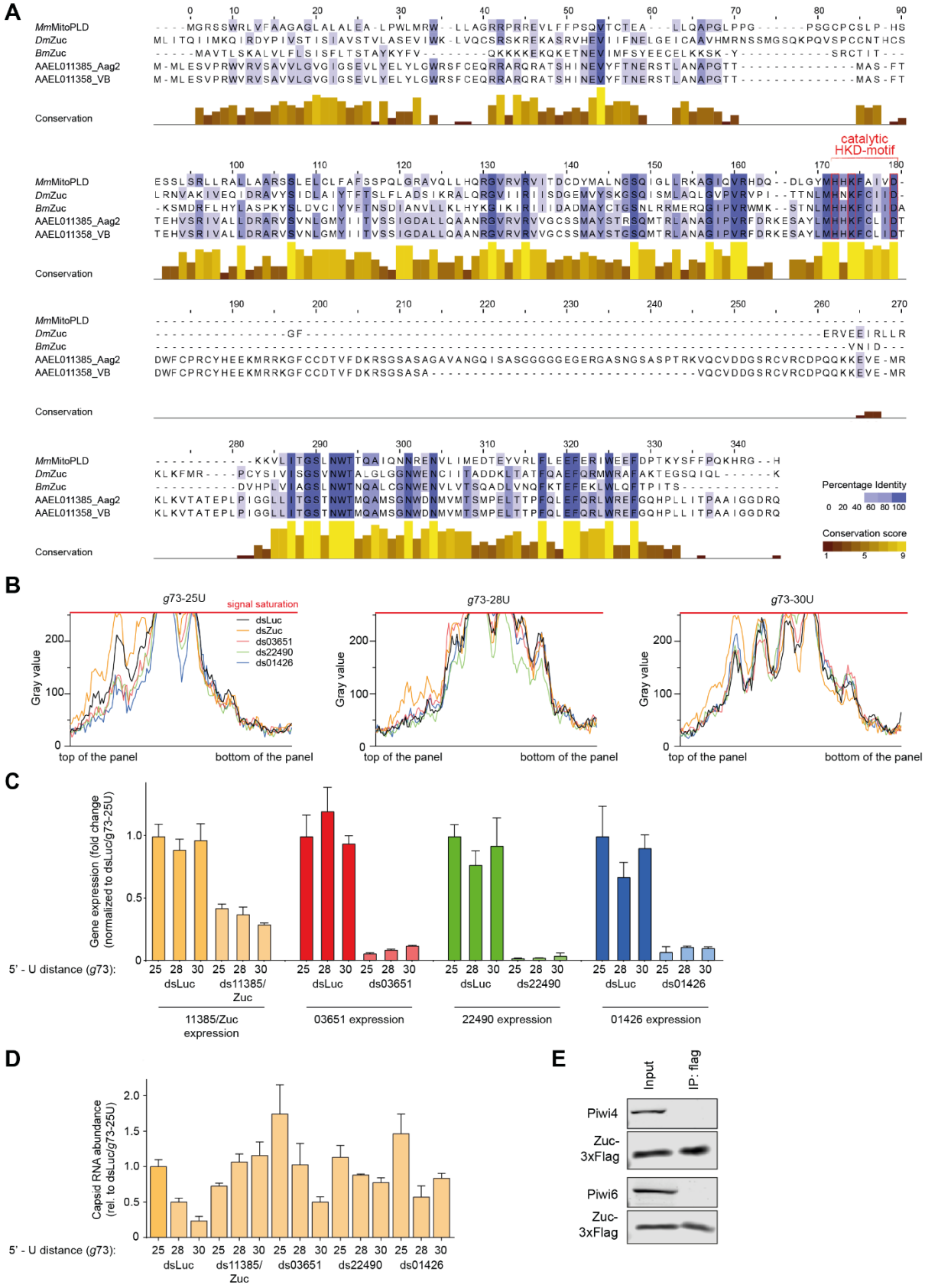


Figure S4. PIWI-knockdown does not affect virus replication
(A) RT-qPCR analyses of subgenomic capsid RNA abundance in the samples used in Figure 3A. RNA levels are shown as a fold change relative to dsLuc treated cells infected with *g*73-25U. Bars and whiskers show the mean and standard deviation (SD), respectively (same for B and D).
(B) RT-qPCR analyses of capsid RNA abundance in the indicated U4.4 (yellow) and Aag2 (blue) samples used in Figure 3B. Values shown are fold changes relative to WT cells infected with the SINV 3' EV control virus.
(C) Schematic representation of the *g*73-28U virus that was used to study the effect of target site mutations on responder piRNA production. Yellow shading indicates the area that is shown in Figure 3C.
(D) RT-qPCR analyses of capsid RNA levels in samples used in Figure 3D. All data are normalized to cells infected with the g73-28U scrambled virus.

636 **Figure S5**



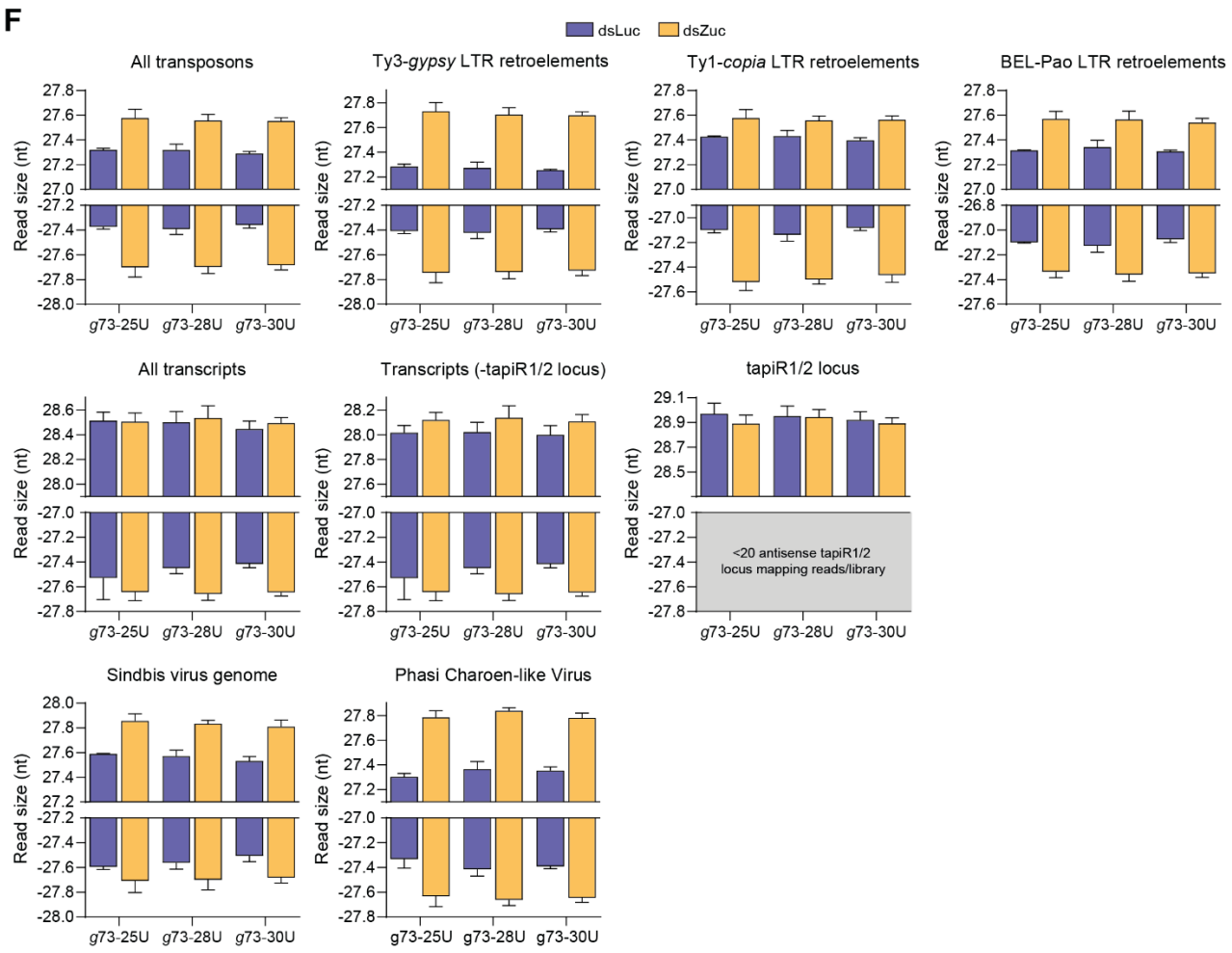


Figure S5. Mosquito Zuc is involved in 3' end formation of piRNAs from various substrates
(A) Multiple sequence alignment of Zucchini orthologs from *D. melanogaster* (*Dm*Zuc), *B. mori* (*Bm*Zuc) and *M. musculus* (*Mm*MitoPLD), and AAEL011385 as annotated in VectorBase (VB) and as sequenced from Aag2 cells. Residues shaded in blue are shared between ≥3 of the proteins, and the brown-yellow bars indicate conservation of physicochemical properties at each position.
(B) Quantification of the responder piRNA signal from the northern blot in Figure 4B, showing the responder piRNA size shift upon Zuc knockdown, analyzed for the three *g*73-derived piRNA targeted viruses separately. The X-axis represents the position on the northern blot from top to bottom, the Y-axis shows responder piRNA signal intensity.
(C-D) RT-qPCR analyses of the knockdown efficiency of indicated genes (C) and viral capsid RNA levels (D). Gene expression levels are shown as fold change relative to dsLuc treated cells infected with the *g*73-25U virus. Bars and whiskers depict mean and SD, respectively.
(E) Western blot analyses of Piwi4 and Piwi6 in the same Zuc-3xflag-IP material that was used in Figure 4H.
(F) Average size of piRNAs (24-33 nt) derived from the indicated substrates in small RNA deep sequencing libraries of dsLuc and ds*Zuc* treated Aag2 cells. The length profile of two piRNAs (tapiR1/2) involved in degradation of maternal transcripts during embryogenesis (58), is not affected by Zuc knockdown, suggesting that they are generated through an Zuc-independent, non-canonical mechanism. As virtually no antisense reads map to the tapiR1/2 locus, these data are not shown. Bars and whiskers represent mean and SD, respectively.

637

638 **Figure S6**

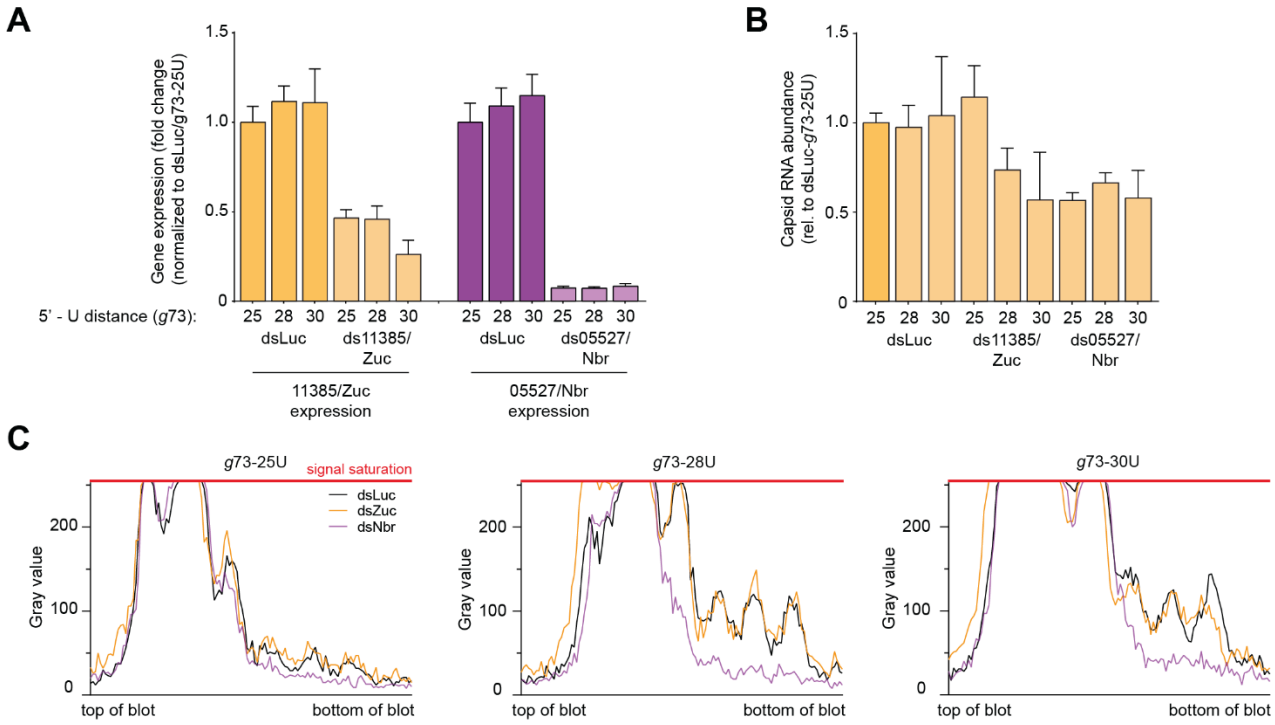


Figure S6. Nbr- and Zuc knockdown affect piRNA biogenesis
(A-B) RT-qPCR analyses of the knockdown efficiencies of the indicated genes (A) and viral capsid RNA levels (B). Expression values represent the fold change relative to dsLuc treated cells infected with the *g*73-25U virus. Bars and whiskers represent the mean +/- standard deviation.
(C) Quantification of the northern blot signal in Figure 5B, showing the effect of Nbr knockdown on responder piRNA size during infection with the indicated viruses. The Y-axis represents responder piRNA signal intensity, the X-axis represents the position on the northern blot from top to bottom.

639

Figure S7

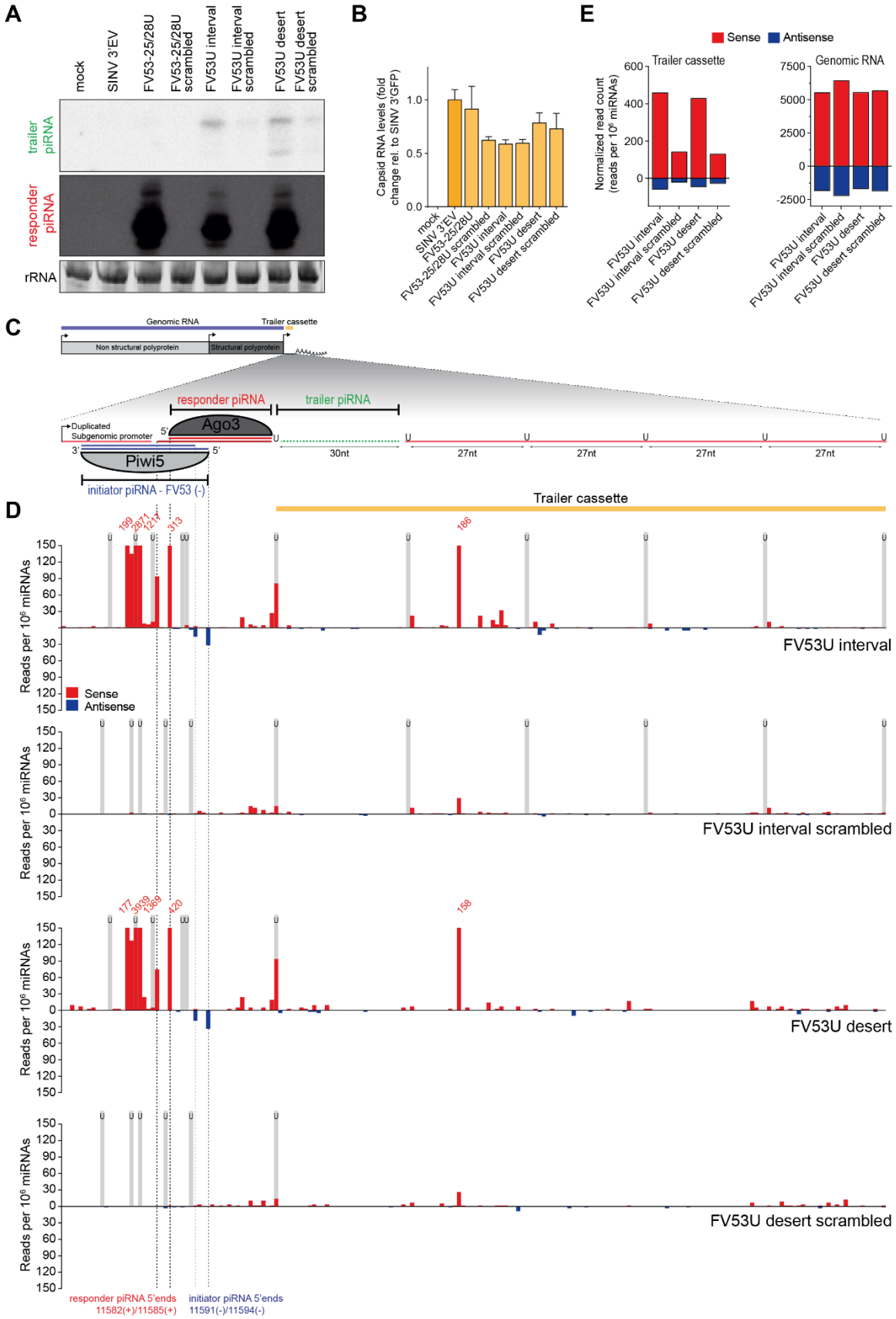


Figure S7. An EVE-derived piRNA triggers trailer piRNA production downstream of its target site (previous page)
(A) Northern blot analyses showing the production of the first putative trailer piRNA (marked in green in (C)) from the indicated viruses). As a control, the target site was scrambled to abolish targeting by the FV53-derived initiator piRNA. The remainder of the responder piRNA site, as well as the trailer cassette was identical to the U interval and U desert viruses. Additional controls include a virus that lacks the trailer cassette, but contains an FV53 piRNA-target site (FV25-28) and a control virus without any insert (SINV 3' EV). A schematic overview of the FV53 interval virus is shown in (C). rRNA stained by EtBr serves as loading control.
(B) Capsid RNA levels in Aag2 cells infected with indicated viruses as determined by RT-qPCR. All capsid RNA levels are shown as a fold change relative to the virus lacking an initiator piRNA target site (SINV 3' EV). Bars and whiskers denote the mean and SD, respectively.
(C) Schematic overview of the FV53 U interval virus. In this virus, a duplicated subgenomic promoter drives the expression of a non-coding reporter RNA (shown in the magnification), which contains a target site for two isoforms of a Piwi5-bound FV53-derived initiator piRNA (shown in blue). Slicing of the reporter RNA in the ping-pong amplification loop thus may give rise to two differently sized Ago3-associated responder piRNAs (shown in red). Downstream of this responder piRNA, the trailer cassette is located which either contains uridine residues at regularly spaced intervals (FV53 U interval) or is completely devoid of uridines (FV53 U desert). The first trailer piRNA which was detected in (A) is shown in green.
(D) Visualization of normalized counts of 5' ends of sense (red) and antisense (blue) piRNA-sized reads (24-33 nt) mapping to the non-coding reporter RNA. The 5' ends of the two FV53-derived initiator piRNA isoforms, as well as the 5' ends of putative responder piRNAs are indicated by dashed lines. Numbers indicate normalized read counts where they exceed the range of the y-axis and grey shading indicates the position of uridine residues.
(E) Quantification of the number of sense (red) and antisense (blue) piRNAs (in reads per 10^6 miRNAs) mapping to the trailer cassette (left) and genomic RNA (right) of the indicated recombinant Sindbis viruses. The areas of the viruses denoted as trailer cassette and genomic RNA are indicated in (C) in yellow and blue, respectively.

641

642 **Figure S8**

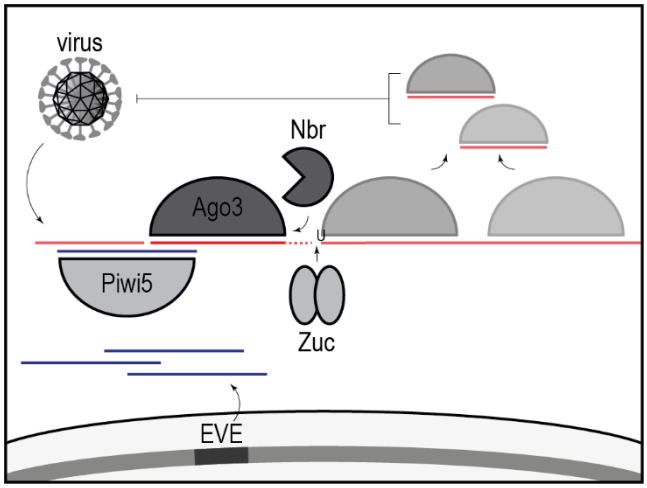


Figure S8. A model of piRNA-based adaptive immunity in mosquitoes
Genomically encoded endogenous viral elements (EVE) give rise to a pool of Piwi5-associated initiator piRNAs that have the potential to target newly infecting viruses. Upon acute infection with a virus containing a cognate sequence, EVE-derived piRNAs trigger the production of Ago3-bound responder piRNAs from the viral RNA, which are generated and maturated by the combined activities of Zuc and Nbr. Additionally, targeting of the viral RNA by the ping-pong machinery prompts the processing of the downstream RNA into additional trailer piRNAs, thus expanding the piRNA sequence pool that is able to target viral RNA.

643

REFERENCES

1. Kraemer, M.U.G., Reiner, R.C., Jr., Brady, O.J., Messina, J.P., Gilbert, M., Pigott, D.M., Yi, D., Johnson, K., Earl, L., Marczak, L.B. *et al.* (2019) Past and future spread of the arbovirus vectors Aedes aegypti and Aedes albopictus. *Nat Microbiol*, **4**, 854-863.
2. Franz, A.W., Kantor, A.M., Passarelli, A.L. and Clem, R.J. (2015) Tissue Barriers to Arbovirus Infection in Mosquitoes. *Viruses*, **7**, 3741-3767.
3. Bronkhorst, A.W. and van Rij, R.P. (2014) The long and short of antiviral defense: small RNA-based immunity in insects. *Curr Opin Virol*, **7**, 19-28.
4. Miesen, P., Joosten, J. and van Rij, R.P. (2016) PIWIs Go Viral: Arbovirus-Derived piRNAs in Vector Mosquitoes. *PLoS Pathog*, **12**, e1006017.
5. Ozata, D.M., Gainetdinov, I., Zoch, A., O'Carroll, D. and Zamore, P.D. (2019) PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet*, **20**, 89-108.
6. Czech, B. and Hannon, G.J. (2016) One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem Sci*, **41**, 324-337.
7. Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K. and Zamore, P.D. (2018) A Single Mechanism of Biogenesis, Initiated and Directed by PIWI Proteins, Explains piRNA Production in Most Animals. *Mol Cell*, **71**, 775-790 e775.
8. Han, B.W., Wang, W., Li, C.J., Weng, Z.P. and Zamore, P.D. (2015) piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science*, **348**, 817-821.
9. Mohn, F., Handler, D. and Brennecke, J. (2015) piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science*, **348**, 812-817.
10. Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P. and Zamore, P.D. (2007) The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Current biology : CB*, **17**, 1265-1272.
11. Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H. and Siomi, M.C. (2007) Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes & development*, **21**, 1603-1608.
12. Hayashi, R., Schnabl, J., Handler, D., Mohn, F., Ameres, S.L. and Brennecke, J. (2016) Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature*, **539**, 588-592.
13. Feltzin, V.L., Khaladkar, M., Abe, M., Parisi, M., Hendriks, G.J., Kim, J. and Bonini, N.M. (2015) The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in Drosophila. *Aging Cell*, **14**, 443-452.
14. Wang, H., Ma, Z., Niu, K., Xiao, Y., Wu, X., Pan, C., Zhao, Y., Wang, K., Zhang, Y. and Liu, N. (2016) Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in Drosophila. *Development*, **143**, 530-539.
15. Sienski, G., Donertas, D. and Brennecke, J. (2012) Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell*, **151**, 964-980.
16. Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A. and Toth, K.F. (2013) Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes & development*, **27**, 390-399.
17. Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H. and Siomi, M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science*, **315**, 1587-1590.
18. Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R. and Hannon, G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **128**, 1089-1103.

19. Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R. and Hannon, G.J. (2009) Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell*, **137**, 522-535.
20. Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M. *et al.* (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell*, **137**, 509-521.
21. Lewis, S.H., Quarles, K.A., Yang, Y., Tanguy, M., Frezal, L., Smith, S.A., Sharma, P.P., Cordaux, R., Gilbert, C., Giraud, I. *et al.* (2018) Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol*, **2**, 174-181.
22. Akbari, O.S., Antoshechkin, I., Amrhein, H., Williams, B., Diloreto, R., Sandler, J. and Hay, B.A. (2013) The developmental transcriptome of the mosquito Aedes aegypti, an invasive species and major arbovirus vector. *G3 (Bethesda)*, **3**, 1493-1509.
23. Morazzani, E.M., Wiley, M.R., Murreddu, M.G., Adelman, Z.N. and Myles, K.M. (2012) Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathog*, **8**, e1002470.
24. Miesen, P., Girardi, E. and van Rij, R.P. (2015) Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in Aedes aegypti mosquito cells. *Nucleic Acids Res*, **43**, 6545-6556.
25. Vodovar, N., Bronkhorst, A.W., van Cleef, K.W., Miesen, P., Blanc, H., van Rij, R.P. and Saleh, M.C. (2012) Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells. *PLoS One*, **7**, e30861.
26. Palatini, U., Miesen, P., Carballar-Lejarazu, R., Ometto, L., Rizzo, E., Tu, Z., van Rij, R.P. and Bonizzoni, M. (2017) Comparative genomics shows that viral integrations are abundant and express piRNAs in the arboviral vectors Aedes aegypti and Aedes albopictus. *BMC Genomics*, **18**, 512.
27. Whitfield, Z.J., Dolan, P.T., Kunitomi, M., Tassetto, M., Seetin, M.G., Oh, S., Heiner, C., Paxinos, E. and Andino, R. (2017) The Diversity, Structure, and Function of Heritable Adaptive Immunity Sequences in the Aedes aegypti Genome. *Current biology : CB*, **27**, 3511-3519 e3517.
28. Suzuki, Y., Frangeul, L., Dickson, L.B., Blanc, H., Verdier, Y., Vinh, J., Lambrechts, L. and Saleh, M.C. (2017) Uncovering the Repertoire of Endogenous Flaviviral Elements in Aedes Mosquito Genomes. *J Virol*, **91**.
29. Ter Horst, A.M., Nigg, J.C., Dekker, F.M. and Falk, B.W. (2019) Endogenous Viral Elements Are Widespread in Arthropod Genomes and Commonly Give Rise to PIWI-Interacting RNAs. *J Virol*, **93**.
30. Aguiar, E., de Almeida, J.P.P., Queiroz, L.R., Oliveira, L.S., Olmo, R.P., de Faria, I., Imler, J.L., Gruber, A., Matthews, B.J. and Marques, J.T. (2020) A single unidirectional piRNA cluster similar to the flamenco locus is the major source of EVE-derived transcription and small RNAs in Aedes aegypti mosquitoes. *RNA*, **26**, 581-594.
31. Suzuki, Y.B., A.; Miesen, P.; Frangeul, L.; Crist, A.B.; Merkling, S.H.; Fontaine, A.; Lequime, S.; Moltini-Conclois, I.; Blanc, H.; van Rij, R.P.; Lambrechts, L.; Saleh, M.C. (2020) Non-retroviral endogenous viral element limits cognate virus replication in Aedes aegypti ovaries. *BioRxiv*.
32. Tassetto, M., Kunitomi, M., Whitfield, Z.J., Dolan, P.T., Sanchez-Vargas, I., Garcia-Knight, M., Ribiero, I., Chen, T., Olson, K.E. and Andino, R. (2019) Control of RNA viruses in mosquito cells through the acquisition of vDNA and endogenous viral elements. *Elife*, **8**.
33. Cristina Crava, F.S.V., Elisa Pischedda, Rebecca Halbach,Umberto Palatini, Michele Marconcini, Annamaria Mattia, Seth Redmond, Yaw Afrane, Diego Ayala, Christophe Paupy, Rebeca Carballar-Lejarazu, Pascal Miesen, Ronald P. van Rij, Mariangela Bonizzoni. (2020) Immunity to infections in arboviral vectors by integrated viral sequences: an evolutionary perspective. *BiorXiv*.
34. Joosten, J., Miesen, P., Taskopru, E., Pennings, B., Jansen, P., Huynen, M.A., Vermeulen, M. and Van Rij, R.P. (2019) The Tudor protein Veneno assembles the ping-pong amplification complex that produces viral piRNAs in Aedes mosquitoes. *Nucleic Acids Res*, **47**, 2546-2559.

35. Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L. and Hannon, G.J. (2012) The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature*, **491**, 279-283.
36. Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J. *et al.* (2012) Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature*, **491**, 284-287.
37. Izumi, N., Shoji, K., Suzuki, Y., Katsuma, S. and Tomari, Y. (2020) Zucchini consensus motifs determine the mechanism of pre-piRNA production. *Nature*, **578**, 311-316.
38. Fredericks, A.C., Russell, T.A., Wallace, L.E., Davidson, A.D., Fernandez-Sesma, A. and Maringer, K. (2019) Aedes aegypti (Aag2)-derived clonal mosquito cell lines reveal the effects of pre-existing persistent infection with the insect-specific bunyavirus Phasi Charoen-like virus on arbovirus replication. *PLoS Negl Trop Dis*, **13**, e0007346.
39. Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., Antony, C., Sachidanandam, R. and Pillai, R.S. (2011) Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature*, **480**, 264-267.
40. Pane, A., Wehr, K. and Schupbach, T. (2007) zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. *Developmental cell*, **12**, 851-862.
41. Huang, H., Li, Y., Szulwach, K.E., Zhang, G., Jin, P. and Chen, D. (2014) AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification. *J Cell Biol*, **206**, 217-230.
42. Haase, A.D., Fenoglio, S., Muerdter, F., Guzzardo, P.M., Czech, B., Pappin, D.J., Chen, C., Gordon, A. and Hannon, G.J. (2010) Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. *Genes & development*, **24**, 2499-2504.
43. Han, B.W., Hung, J.H., Weng, Z., Zamore, P.D. and Ameres, S.L. (2011) The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to Drosophila Argonaute1. *Curr Biol*, **21**, 1878-1887.
44. Liu, N., Abe, M., Sabin, L.R., Hendriks, G.J., Naqvi, A.S., Yu, Z., Cherry, S. and Bonini, N.M. (2011) The exoribonuclease Nibbler controls 3' end processing of microRNAs in Drosophila. *Curr Biol*, **21**, 1888-1893.
45. Etebari, K., Osei-Amo, S., Blomberg, S.P. and Asgari, S. (2015) Dengue virus infection alters post-transcriptional modification of microRNAs in the mosquito vector Aedes aegypti. *Sci Rep*, **5**, 15968.
46. Li, S., Mead, E.A., Liang, S. and Tu, Z. (2009) Direct sequencing and expression analysis of a large number of miRNAs in Aedes aegypti and a multi-species survey of novel mosquito miRNAs. *BMC Genomics*, **10**, 581.
47. Reichholf, B., Herzog, V.A., Fasching, N., Manzenreither, R.A., Sowemimo, I. and Ameres, S.L. (2019) Time-Resolved Small RNA Sequencing Unravels the Molecular Principles of MicroRNA Homeostasis. *Mol Cell*, **75**, 756-768 e757.
48. Izumi, N., Shoji, K., Sakaguchi, Y., Honda, S., Kirino, Y., Suzuki, T., Katsuma, S. and Tomari, Y. (2016) Identification and Functional Analysis of the Pre-piRNA 3' Trimmer in Silkworms. *Cell*, **164**, 962-973.
49. Tang, W., Tu, S., Lee, H.C., Weng, Z. and Mello, C.C. (2016) The RNase PARN-1 Trims piRNA 3' Ends to Promote Transcriptome Surveillance in C. elegans. *Cell*, **164**, 974-984.
50. Hahn, C.S., Hahn, Y.S., Braciale, T.J. and Rice, C.M. (1992) Infectious Sindbis virus transient expression vectors for studying antigen processing and presentation. *Proc Natl Acad Sci U S A*, **89**, 2679-2683.
51. Saleh, M.C., Tassetto, M., van Rij, R.P., Goic, B., Gausson, V., Berry, B., Jacquier, C., Antoniewski, C. and Andino, R. (2009) Antiviral immunity in Drosophila requires systemic RNA interference spread. *Nature*, **458**, 346-350.
52. Livingstone, C.D. and Barton, G.J. (1993) Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Comput Appl Biosci*, **9**, 745-756.
53. Pall, G.S. and Hamilton, A.J. (2008) Improved northern blot method for enhanced detection of small RNA. *Nat Protoc*, **3**, 1077-1084.

786 54. Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., Nekrutenko, A. and Galaxy, T.
787 (2010) Manipulation of FASTQ data with Galaxy. *Bioinformatics*, **26**, 1783-1785.
788 55. Maringer, K., Yousuf, A., Heesom, K.J., Fan, J., Lee, D., Fernandez-Sesma, A., Bessant, C., Matthews,
789 D.A. and Davidson, A.D. (2017) Proteomics informed by transcriptomics for characterising active
790 transposable elements and genome annotation in Aedes aegypti. *Bmc Genomics*, **18**, 101.
791 56. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo
792 generator. *Genome Res*, **14**, 1188-1190.
793 57. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time
794 quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408.
795 58. Halbach, R., Miesen, P., Joosten, J., Taskopru, E., Rondeel, I., Pennings, B., Vogels, C.B.F., Merkling,
796 S.H., Koenraadt, C.J., Lambrechts, L. *et al.* (2020) A satellite repeat-derived piRNA controls embryonic
797 development of Aedes. *Nature*, **580**, 274-277.

798