RNA-dependent RNA polymerases in the black-legged tick produce Argonaute-dependent small RNAs and regulate genes

Canran Feng[1], Kyosuke Torimaru[1], Mandy Yu Theng Lim[2,3], Li-Ling Chak[2] , Kosuke Tsuji[1], Tetsuya Tanaka[4], Junko Iida[1] and Katsutomo Okamura[1,2,3]

[1]Nara Institute of Science and Technology, 8916-5 Takayama Ikoma Nara 630-0192, Japan

[2]Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore 117604, Singapore

[3]School of Biological Sciences, Nanyang Technological University
60 Nanyang Drive, Singapore 639798, Singapore

[4]Laboratory of Infectious Diseases, Joint Faculty of Veterinary Medicine, Kagoshima University, 1-21-24 Korimoto, Kagoshima, 890-0065, Japan

Author for correspondence
email: okamurak@bs.naist.jp
TEL: +81-743-72-5410

Abstract

Small regulatory RNAs (sRNAs) are involved in anti-viral defense and gene regulation. Although RNA-dependent RNA Polymerases (RdRPs) are known to produce sRNA in nematodes, plants and fungi, whether they play roles in sRNA biogenesis in other animals remains controversial. In this study, we study sRNAs in the ISE6 cell line, which is derived from the black-legged tick, an important vector of human and animal pathogens. We identify abundant classes of ~22nt sRNAs that require specific combinations of RdRPs and sRNA effector proteins (Argonautes or AGOs). RdRP-dependent sRNAs are mainly derived from sense and antisense strands of RNA polymerase III-transcribed genes and repetitive elements. Unlike *C. elegans* sRNA pathways, 5'-tri-phosphorylated sRNAs are not detected, suggesting that the tick pathways are distinct from the pathways known in worms. Knockdown of one of the RdRPs unexpectedly results in downregulation of a subset of viral transcripts, in contrast to their upregulation by AGO knockdown. Furthermore, we show that knockdown of AGO/RdRP causes misregulation of protein-coding genes including RNAi-related genes, suggesting feedback regulation. Luciferase assays demonstrate that one of the RdRP-regulated genes, the MEK1 ortholog IscDsor1 is regulated through its 3'UTR, where a putative sRNA target site resides. These results provide evidence that arachnid RdRPs are important sRNA biogenesis factors, and the discovery of novel pathways underscores the importance of characterizing sRNA biogenesis in various organisms to understand virus-vector interactions and to exploit RNAi for pest control.

Significance statement

RNA-dependent RNA Polymerases (RdRPs) are essential for biogenesis of small regulatory RNAs (sRNAs) in many organisms such as plants and fungi, but its general importance in animals besides nematodes remains controversial because experimental evidence is lacking. By using a tick cell line, we demonstrate that RdRP-dependent sRNAs are abundantly expressed and tick RdRPs regulate gene expression. These results indicate that ticks have unexpectedly complex sRNA biogenesis pathways that are essential for proper gene regulation. Because ticks are important vectors of human and animal pathogens, and their sRNA pathways control tick-borne viruses, studying the novel tick sRNA pathways may provide important information for a better understanding of vector-virus interactions and for developing RNA-based pesticides by utilizing tick's RNA silencing mechanisms.

Introduction

Foreign nucleic acids such as transposable elements (TEs), phages and viruses pose constant threats to host cells. To inactivate invading agents, cells are equipped with defense mechanisms that use short fragments of nucleic acids to distinguish those foreign nucleic acids from their own genetic materials and silence them (1).

In prokaryotes, a common defense mechanism involves CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) that are arrays of short sequences derived from phages and produce short RNAs against the foreign nucleic acids (2). Each of the short phage sequences produces a guide RNA that binds to the effector protein typically to degrade foreign DNA or RNA in a sequence-specific manner. Another major class of defense systems using small RNAs (sRNAs) involves another family of sRNA binding proteins known as Argonautes (AGOs). The AGO system is widely conserved in various organisms and can be found in both eukaryotes and prokaryotes (3).

The large diversity of sRNA pathways is assumed to be a result of a relentless arms race between the host cell and the invading nucleic acids. In bilaterian animals, three distinct AGO-dependent sRNA pathways, namely the PIWI-interacting RNA (piRNA), small interfering RNA (siRNA) and microRNA (miRNA) pathways, are widely present (4). These pathways were initially characterized in a few model animals including *Drosophila, C. elegans* and mice (5). Expression of the piRNA pathway components is virtually restricted within gonads in these organisms and they appear to be specialized in silencing transposons while also playing roles in gene regulation (6, 7). However, recent studies argue against the notion that active piRNA production is generally confined in gonads as abundant piRNAs have been detected in somatic tissues in a wide variety of arthropods, even in some insects (8–12).

The siRNA pathway is believed to be a major mechanism to control viruses in insects (13). Mutants of the core siRNA factors exhibit elevated sensitivity to various viruses (14) and siRNAs are efficiently produced in infected cells via a specific processing mechanism (15, 16). Interestingly, the siRNA factors are among the most rapidly evolving genes potentially because they must catch up with the rapidly changing viruses. Due to the rapid evolution, the siRNA pathway shows significant evolutionary diversity (17, 18). This was first recognized by a comparison between the insect and nematode RNAi mechanisms. The clearest difference is the essential involvement of RNA-dependent RNA polymerases (RdRPs) in worms but not in flies for robust RNA silencing (19–21). RdRPs are known to produce various small RNAs during RNA silencing in fungi and plants (22, 23) whereas the gene was believed to have been lost in the animal lineage until genome sequence analyses of non-model organisms started to discover RdRP genes in a broad range of animals (18, 24, 25).

The sequences of most RdRP genes support that the genes had been vertically transferred and were not introduced to the animals by the horizontal transfer, suggesting that RdRPs may have conserved biological roles in those lineages (26). The findings that the organisms with no piRNA pathway genes always retain the RdRP genes have led to the notion that the ping-pong amplification mechanism in the piRNA pathway and RdRP-dependent small RNA production pathway have overlapping roles in reinforcing silencing activity against transposable elements (TEs) (Lewis et al., 2018; Mondal et al., 2018b, 2018a, 2020). On the other hand, a past study found no evidence for the production of secondary sRNAs that have features seen with *C. elegans* secondary siRNAs, such as 5'-triphospophate groups and 5'-purine enrichment (26). Experimental evidence for the functional involvement of RdRPs in sRNA biogenesis in animals outside of Nematoda is currently lacking.

To directly test if RdRPs are involved in the production of sRNAs in arachnids, we use a cell line from the model tick *Ixodes scapularis*. Using the cell line, we provide experimental evidence that abundant classes of RdRP-dependent sRNAs regulate the expression of genes in tick cells. There are distinct classes of sRNAs produced through the activity of at least two different RdRP genes. We further demonstrate the involvement of RdRP-dependent sRNAs in gene regulation using global gene expression analysis and sensor assays. Knockdown of one of the RdRPs unexpectedly resulted in a reduction of specific viral transcripts, suggesting that the RdRP is required for maintaining the level of viral transcripts. These results demonstrate that tick RdRPs are essential for the biogenesis of specific sRNAs, and play roles in gene regulation and controlling viral transcript levels. This study unveils previously overlooked pathways that are potentially broadly conserved in arachnids, and provides experimental and transcriptomics resources for analysis of tick sRNA pathways.

Results

RNAi factors are diversified in the *Ixodes* genome

The broad presence of recognizable RdRP genes places arachnids in a unique position in the arthropod lineage (11). We hypothesized that arachnids might have previously unrecognized small regulatory RNA pathways fueled by the enzymatic activity of RdRPs to produce antisense RNA molecules.

To identify RdRP genes that were expressed in ISE6 cells, we first performed transcriptome assembly by sequencing ribosomal RNA-depleted RNAseq libraries (Table S1, Sheet7). In the assembled transcriptome, we found 3 genes that were similar to known *C. elegans* RdRP genes (Gene IDs and contig sequences are on Table S1, Sheet4). These sequences were predicted to contain the entire sequence of the conserved RdRP domain, suggesting that they were genuine RdRP genes (Figure 1A and B). We named them IscRdRP1,3 and 4. Nematode RdRP genes could be divided into two groups based on the presence or absence of an extra loop sequence. The presence of the putative loop defines the RRF-1 group, which is associated with the activity to produce short antisense RNAs presumably by a primer-independent manner (27). We aligned the amino acid sequences of IscRdRP1 and IscRdRP3 with the worm homologs (Figure S1). This analysis indicated that the catalytic residues (28) were conserved in the tick RdRPs but they lacked the RRF-1 specific loop, suggesting that they were more closely related to the ancient RdRP groups. These results suggested that tick RdRPs might participate in molecular machineries to produce antisense RNA molecules, which may be distinct from those discovered in nematodes. Two (RdRP1 and RdRP3) out of the three RdRP genes were expressed at >10 TPM in our transcriptome data (Figure 1C), and in this present study, we characterized these two RdRPs.

sRNA pathways often employ specific AGO proteins as their effectors (29). We tried to retrieve the AGO genes present in the ISE6 transcriptome. By using fly AGO/PIWI protein sequences as baits, blast analysis identified 6 contigs as potential AGO genes (Figure 1A and S2). We first verified the expression of these genes by qPCR and tested the specificity of qPCR primers by introducing dsRNAs against cognate genes from regions that did not overlap with the qPCR amplicons (Figure S3A). In all cases, we observed a strong (40-90%) reduction in the expression levels of the tested genes upon knockdown, confirming the specificity of the qPCR assay. Therefore, we concluded that these tested genes were indeed expressed in ISE6 cells.

Among the tested genes, we found a contig in our transcriptome data that matched two annotated genes (ISCI012408 and ISCI004800). This contig was potentially derived from two genomic loci with very similar sequences and the sequences similar to ISCI012408 and ISCI004800 were next to each other at both loci (Figure S3B). This contig showed the

highest similarity to the *Drosophila* AGO3 (dAGO3) (Figure 1A), and raised a possibility that these two ISCI entries represented fragments of a dAGO3 homolog. We consider these two ISCI entries a single gene throughout the manuscript and renamed this IscAGO3 because knockdown using dsRNAs derived from either of the annotated sequences resulted in depletion of the fragments from both ISCI012408 and ISCI004800 (Figure S3C). We found another gene belonging to the PIWI-clade (IscAub). Other genes were similar to AGO-clade Argonautes, which were identified in a previous study (30); an ortholog of the miRNA-class AGO (IscAgo-78) and three other genes that were relatively distant from miRNA AGOs (IscAgo-16, IscAgo-30 and IscAgo-96) (Figure 1A). The predicted protein sequences of Ago-16, Ago-30 and Ago-96 contained the entire PIWI domain (Figure 1B) and their catalytic residues were also conserved (Figure S2), suggesting that they were functional slicer enzymes (31).

To examine the expression levels of these AGO genes in ISE6 cells, we analyzed our total RNAseq data (Figure 1C). The PIWI-related genes (IscAGO3 and IscAub) were highly expressed (>90 TPM) in ISE6 cells, which were assumed to be derived from the neural lineage (32). Although PIWI proteins were previously believed to be confined in the animal gonad in general (33), our observation was consistent with the recent findings that supported the broader presence of active piRNA mechanisms in somatic tissues in arachnids (8, 11). Expression of the PIWI proteins was further confirmed by Western blotting using antibodies against asymmetric di-methyl-Arginine, a post-translational modification that is conserved among PIWI proteins (34). The dsRNA against IscAub decreased the signal, suggesting that IscAub was the major PIWI protein modified with asymmetric di-methyl-arginines in ISE6 cells (Figure S3D). The other four AGO genes were also highly expressed (10-250 TPM, Figure 1C).

Therefore, our RNAseq data indicated that components of multiple small RNA pathways including distinct PIWI/AGO genes as well as RdRP genes were expressed in ISE6 cells.

Subcellular localization of AGO and RdRP products

To study the localization of the homologs of AGOs and RdRPs, we cloned the putative ORFs of Ago-16, Ago-30, RdRP1 and RdRP3 into a mammalian expression vector with an N-terminal EGFP tag. Using the plasmids, we transfected HEK293T cells and analyzed their subcellular localization by confocal microscopy. Successful expression of the fusion proteins was confirmed by Western blotting analysis (Figure S3E). Transfected cells were fixed and observed by confocal microscopy and signals were mainly detected in the cytoplasm for all of the RdRP/AGO constructs (Figure 1D). Although such a heterologous experimental system might not accurately reflect their natural subcellular localization as seen with

mislocalization of PIWI proteins expressed in cells lacking an active piRNA processing pathway (35), our results suggested that these proteins could localize in the cytoplasm at least under certain conditions.

The sRNA repertoire of ISE6 cells

To understand the tick small RNA repertoires, we performed sRNAseq analysis using total RNA samples extracted from ISE6 cells (Table S1). To understand their biogenesis mechanisms, we also generated sRNA libraries from ISE6 cells after knocking down the AGO/PIWI/RdRP genes, and each of the libraries yielded ~13-20 million reads that could be mapped to the ISE6 genome (Table S1). sRNA reads in these libraries showed a bimodal distribution with peaks at ~22nt and in the 26-29nt range, which typically represent miRNAs/siRNAs and piRNAs, respectively (Figure S4A).

To seek clues to the functions and biogenesis mechanisms of the sRNAs, we categorized sRNA reads based on their genomic origins (Figure 2A). The sRNA reads were sequentially mapped to the reference sequences in different categories, including miRNAs, RNA polymerase III (RNAP III) transcribed genes, rRNA, snoRNAs, protein-coding genes and repetitive sequences (See Table S1 sheet2 for the details of the reference sequences). In flies, the sRNA population is dominated by miRNAs and piRNAs, and the percentages of other sRNA species including endogenous siRNAs tend to be small (36). In the control library transfected with GFP dsRNA, ~15% of the library was comprised of known miRNAs that were annotated in miRBase (ver 22). As this class of sRNAs was strongly reduced upon the knockdown of Ago-78, which encoded the miRNA AGO ortholog (Figure 1A), this result confirmed the major role of Ago-78 in the miRNA pathway. We did not observe strong effects on miRNAs when other AGOs were knocked down (Figure 2A, Figure S4B), suggesting that other AGOs might support functions of other sRNA classes. Repeats produce multiple classes of sRNAs. More than 40% of 22nt and 25-30nt species were derived from repeats (Figure S4C), and they might represent repeat-associated siRNAs and piRNAs as seen in the fly system (5). Indeed, the 25-30nt species derived from repeats were strongly decreased when PIWI genes were knocked down (Figure S4D, 25-30nt). On the other hand, the 22nt species showed no strong reduction upon knockdown of any of the factors (Figure S4D, 22nt), suggesting that there might be multiple mechanisms producing them as discussed later.

We also found an abundant group of 21-22nt sRNA reads derived from various genes that were known to be transcribed by RNAP III (Figure 2A). The read counts of sRNAs in this category accounted for ~9% of the control library, which was nearly as abundant as miRNAs (~15%) as a class. The sRNAs were mapped to both sense and antisense directions with respect to the direction of transcription of their host genes, excluding the possibility that they

were mere degradation products of abundant RNAP III transcripts (Figure 2B). Furthermore, these small RNAs were virtually eliminated when RdRP1 was knocked down, indicating that ISE6 cells possess molecular mechanisms to produce sRNAs that are different from those known in *Drosophila* or *C. elegans* (Figure 2A).

Chemical structures of tick sRNAs

The chemical structures of 5'- and 3'-terminal nucleotides of sRNAs often reflect their biogenesis mechanisms because processing and modifying enzymes leave characteristic functional groups at these ends (37, 38). In general, AGO-bound sRNAs possess 5'-mono-phosphate groups that are recognized by the 5' binding pocket of the MID domain (39), with a notable exception of nematode secondary siRNAs which possess 5'-tri-phosphate groups (29). piRNAs, fly siRNAs and plant miRNAs have 2'-O-methyl groups at their 3' nucleotides, whereas animal miRNAs carry hydroxyl groups at the equivalent position (40). The 2' modification status at the 3'-nucleotide could be analyzed by oxidizing RNA samples with a periodate, as the presence of vicinal free 2'-, 3'-OH species makes the RNA molecule amenable to oxidization and resulting oxidized RNA molecules lack a 3'-OH group that is required for the 3' linker ligation for sRNA library construction (41, 42). Although piRNA species were efficiently enriched in our oxidized sRNA library (Figure S4E), small RNAs from RNAP III-transcribed genes were depleted, indicating that the latter had free 2'-, 3'-OH groups (Figure 2C). To further support this conclusion, we verified the results by Northern blotting. ß-elimination of the 3' nucleotides occurs when oxidized RNA species are incubated in an alkaline solution, resulting in faster migration ß-eliminated RNA species on the denaturing gel (Horwich et al., 2007). After ß-elimination, piRNAs remained at the same size, while miRNAs migrated more rapidly, consistent with the previously known 3' structures of their counterparts in other animals (Figure 2D). We observed faster migration of small RNAs from RNAP III-transcribed genes after ß-elimination, confirming the conclusion that they had 2'-OH species at the 3'-nucleotide. Although this was different from the known structure of the fly siRNA, recent reports also showed that TE-derived sRNAs in arachnids had free 2'-OH at their 3'-ends (8, 10, 11, 26).

We also analyzed the 5' chemical structures of the sRNAs. The standard sRNA cloning protocol for next generation sequencing is selective for 5'-mono-phosphorylated species by taking advantage of the substrate specificity of the RNA ligase (44). The efficient inclusion of the sRNAs derived from RNAP III-transcribed genes in our libraries suggested that they harbored monophosphate groups at their 5' ends. To confirm this, we prepared an sRNA library after removing sRNA species with 5'-monophosphate groups by Terminator exonuclease (45) followed by dephosphorylation and re-phosphorylating the 5' ends using T4 polynucleotide kinase, allowing the resulting libraries to enrich 5' di- or tri-phosphorylated

sRNAs (Figure 2C, bottom). sRNAs from RNAP III-transcribed genes were not enriched in 5' mono-P-depleted observed in this library when compared with the regular 5'-mono-P-enriched library, supporting the hypothesis that these sRNAs were 5'-mono-phosphorylated. Due to the negative charge carried by the phosphate groups, 5'-phosphorylated species show faster migration on the denaturing gel (46). When the RNA samples were treated with a phosphatase, we observed a slight delay of all the sRNAs tested, indicating that they carried phosphate groups (Figure 2D).

Taken together, we concluded that the novel small RNA species from RNAP III transcribed genes carried a 5'-mono-phosphorylated group and were not modified at the 2'-position of the 3'-nucleotide.

sRNA production from RNAP III-transcribed genes is evolutionarily conserved in ticks

If RdRP-dependent sRNAs play important biological roles, one would expect the production of similar sRNA species to be conserved in evolutionarily distant tick species. We reanalyzed sRNA libraries from the Asian longhorned tick (*H. longicornis*) (47, 48). Phylogenetic analysis suggested that *H. longicornis* and *Ixodes* species shared the last common ancestor ~200 million years ago (49). We identified genomic regions that showed the highest similarities to the *I. scapularis* RNase P, RNase MRP and SRP RNA genes, and found that sRNAs were mapped to both strands of these loci (Figure 2E and S5). Importantly, they showed peaks at 22nt on both strands, suggesting that they were produced by specific processing machineries (Figure 2F).

Therefore, the production of 22nt species from RNAP III transcribed genes was broadly conserved in ticks. Furthermore, the presence of similar sRNA species in libraries made from tick animals and saliva suggested that the sRNA production was not restricted in cultured cell lines (Figure S5).

Biogenesis of sRNAs from RNAP III-transcribed genes

The high abundance of sRNAs from RNAP III-transcribed genes suggested there were mechanisms efficiently producing them. As mentioned above, these sRNAs were strongly reduced in RdRP1-depleted cells (Figure 2A and S6A). We verified the results using a set of independently prepared RNA samples by Northern blotting (Figure S6B). Although RdRP1 knockdown strongly reduced the ~22nt species derived from an SRP gene and a tRNA gene, RdRP3 knockdown did not affect the abundance of these sRNAs, suggesting that the two RdRPs had distinct functions. The PIWI-clade proteins and the miRNA effector Ago-78 appeared to play no significant roles as no discernible changes in the abundance of sRNAs from RNAP III-transcribed genes were seen upon Ago-78 knockdown (Figure S6B, SRP RNA- and tRNA-derived small RNA), while their respective client sRNAs were reduced

(Figure S6B, ~27-28nt species from piRNA-family-423-A; ~22nt species of miR-8 and miRNA-candidate-1). When we were analyzing the TE-derived piRNA-family-423-A, we noticed that there was a less abundant species at ~22nt, whose expression was reduced in RdRP1-knockdown cells (Figure S6B, second panel). This indicated that some repeats produce 22nt sRNAs in addition to piRNAs, and at least some of the 22nt repeat-derived sRNAs were RdRP1-dependent (see below for detailed genome-wide analysis).

Knockdown of Ago-16 and Ago-30 affected the expression of certain sRNA species (Figure S6A and B). The SRP RNA locus exhibited an interesting pattern. Knockdown of Ago-16 reduced the sRNA from the sense strand, while knockdown of Ago-30 resulted in a reduction of the antisense sRNA (Figure S6A and B). Distinct effects by knockdown of Ago-16 or Ago-30 were seen with many sRNA species, suggesting that these two AGOs were required for biogenesis of overlapping but different sets of sRNAs. The sRNA species may regulate levels of host ncRNA species. However, the level of the 300nt product of SRP RNA showed no clear difference between control and any of the knockdown samples (Figure S6B, second panel from the bottom). The effects of sRNAs on their host transcripts remain unclear.

Taken together, the results demonstrated that tick sRNA pathways produce RdRP1-dependent sRNAs whose stable expression require specific AGOs, arguing that these molecular pathways represent novel regulatory RNA pathways.

Various sRNAs are produced from coding genes

Although the fraction of sRNAs that mapped to coding exons was small (Figure 2A and Table S2), the production of sRNAs from both strands suggested the involvement of RdRPs. We hypothesized that individual mRNAs were recognized by different sRNA processing machineries. To test this hypothesis, we analyzed sRNA reads mapping to individual annotated protein-coding genes and collected genes that produced sRNA reads at >35RPM on average in the knockdown libraries (Figure 3A and Supplementary Data). When the sRNA levels in AGO- or RdRP-KD libraries from each locus were analyzed, most of the loci showed a strong reduction (>40%) in at least one library compared to the GFP-KD control (28 out of 39 loci, Figure 3B and Supplementary Data). sRNAs from some loci were reduced in more than one sample, and frequent overlaps were seen between the RdRP3-Ago-16 and Aub-AGO3 combinations (Figure 3A). On the other hand, no locus showed reductions both in RdRP1- and RdRP3-knockdown libraries. These results suggested that certain combinations of factors formed sRNA processing pathways and the two RdRPs belonged to different pathways.

The size distributions of sRNAs from individual loci roughly corresponded to their processing dependencies, where 22nt and 25-29nt peaks tended to be RdRP- and PIWI-

dependent, respectively (Figure 3B, lower panels, Supplementary Data). A subset of loci that had both 22nt and 25-29nt peaks showed changes in the level of 22nt species upon knockdown of the PIWI genes (Figure 3B, lower panel). This suggested that these sRNA pathways might interact with each other, and indicated that the two pathways were active in the same cells although the ISE6 culture might be a mixture of multiple cell types. ISCI012234, which encodes a homolog of histone H1, produced the highest number of sRNAs and the sRNAs were dependent on RdRP3 and Ago-16 (Figure 3B). However, the expression of this mRNA showed no consistent changes (see below and Table S3). Therefore molecular functions of RdRP3-dependent sRNAs remained unknown.

Global analysis of sRNAs from protein-coding genes revealed that multiple biogenesis mechanisms were involved in the production of coding gene-derived sRNAs. The fact that each RdRP was involved in the production of sRNAs from a small number of loci suggested that the RdRPs selectively recognize their substrates for sRNA production.

sRNAs produced from repeats

A common role for metazoan RNAi/piRNA pathways is silencing of TEs (29). To study TE-derived sRNAs, repetitive sequences were identified by the RepeatModeler2/RepeatMasker pipeline (50) and a genome-wide annotation of the repetitive sequences was obtained (See Materials and Methods). To test which of the sRNA factors might be working together within the same sRNA biogenesis pathways, we counted the numbers of TEs whose sRNAs were commonly reduced in multiple knockdown libraries (Figure 4A). We used 67 TEs that produced abundant sRNA reads (>800 rpm on average in the knockdown libraries) for this analysis. As expected, a large number of repeats produced sRNAs that were reduced (<60%) upon knockdown of the PIWI genes (30 out of the 67 TEs examined), and many of these showed reduced levels of sRNAs in all of the three PIWI-family knockdown libraries (dsAub, dsAGO3-1 and dsAGO3-2, 13 out of the 30 TEs producing PIWI-family dependent sRNAs). Large overlaps were seen with Ago-16-RdRP3 and Ago-30-RdRP1 combinations, suggesting that the AGOs and RdRPs might work together to produce repeat associated-sRNAs belonging to the same classes. Interestingly, very few overlaps were observed between the three groups, similar to the observation with the sRNAs from coding genes (Figure 3). All these results again suggested that these groups of sRNA processing factors represent sRNA production pathways that largely independently operate to produce their own classes of sRNAs.

To gain further insights, we analyzed individual repeat families (Figure 4B). When the read counts were plotted for each TE family, we noticed that repeat families producing larger numbers of sRNAs tended to be RdRP-dependent whereas families that produced PIWI-dependent sRNAs tended to produce fewer sRNAs (Figure 4B, upper). When their sRNA

sizes were analyzed, families that produced RdRP1- or RdRP3-dependent sRNAs showed clear 22nt peaks and families that produced PIWI-dependent sRNAs showed peaks at ~25-28nt (Figure 4B, bottom). The peaks at the expected sizes of their corresponding classes were strongly reduced by knockdown of the RdRP or PIWI protein, confirming the specific roles of these sRNA processing machineries in producing the respective classes of sRNAs.

The expression levels of repeats were analyzed after knocking down Ago-16, RdRP1 or RdRP3 (Table S4). To our surprise, very few repeats were misregulated. The most significantly misregulated repeat in the Ago-16 knockdown libraries was rnd-6_family-4937, which was also most significantly misregulated in the RdRP3 knockdown libraries. As this repeat produced the second highest number of RdRP3-Ago-16 dependent sRNAs (Figure 4B), these results suggest that the RdRP3-Ago-16 axis may silence repeats.

We found no repeat family showing a decrease of sRNAs in both RdRP1- and RdRP3-knockdown libraries, again supporting the idea that these sRNA processing pathways are largely independent. We occasionally observed repeats whose 22nt peaks were reduced (e.g. rnd-1_family-1111TE, Supplementary Data) or increased (e.g. rnd−5_family−5812, Figure 4B and Supplementary Data) upon Aub knockdown in addition to the reduction of the 25-28nt piRNA peaks. Therefore, interactions between these pathways should not be ruled out.

Roles for sRNA factors in controlling viral RNA levels

Another main biological role that is associated with the RNAi pathway is antiviral defense (51). In ticks, a previous study demonstrated that some AGOs studied in the present study were involved in controlling tick-borne viral pathogens in another tick cell line IDE8, suggesting interactions between the tick RNAi pathway and tick-borne viruses (30). We wondered if the RdRPs played a role in controlling viral transcript levels in ISE6 cells.

In a recent study, a set of viruses were identified in the ISE6 culture by next-generation sequencing analysis of putative viral particles enriched from the ISE6 culture (52). These viruses were assumed to be persistently present in ISE6 cultures. We profiled sRNAs derived from the viral sequences in the sRNAseq libraries made after knocking down the sRNA processing factors. In the control library, sRNAs mapping to the viral genomes were abundantly present (Figure 5A). The reads were distributed across the entire genomes with no strong enrichment in particular regions. The size distribution of the mapped reads showed a strong peak at 22nt without a recognizable peak at the piRNA size (Figure 5B). This was consistent with the vsiRNA seen in TBEV-infected ISE6 cells (30) while suggesting a difference from insect piRNA systems where the production of vpiRNAs was observed upon infection with various mosquito-borne viruses (53).

We sought to determine if vsiRNAs were dependent on any of the Argonautes or RdRPs. Interestingly, vsiRNAs were still produced upon knockdown of any of the RNAi factors (Figure 5C, upper panel). This suggested that these sRNA factors were dispensable for vsiRNA production although functional redundancy between the factors might hinder the detection of the effects. We tested if the knockdown of these factors had any effects on the abundance of viral transcripts (Figure 5C lower panel). In our total RNAseq libraries, upregulation of some of the viruses upon knockdown of Ago-16 was observed (Figure 5C, lower panel), consistent with the previous study using TBEV/LGTV (30). Furthermore, we observed unexpected downregulation of some viruses upon knockdown of RdRP1 (Figure 5C, lower panel). While RdRP1 did not appear to play detectable roles in the production of vsiRNA, it might have a role in maintaining the levels of viruses. To understand the precise mechanisms underlying this phenomenon, further investigation will be needed in the future.

Roles for sRNA factors in gene regulation

To clarify whether the new small RNA pathways described here had any roles in gene regulation, we analyzed the total RNAseq data of ISE6 cells after knocking down Ago-16, RdRP1 or RdRP3 (Figure 6).

Upon knockdown of these genes, we detected 47-84 genes to be differentially expressed compared to the control GFP KD sample (adjusted p-value <0.05, Figure 6A-C and Table S3). GO-term analysis revealed enrichment of the biological process categories related to RNAi and response to dsRNAs upon knockdown of Ago-16 or RdRP3 (Figure 6D and F, Supplementary PDF). In particular, Dicer homologs were upregulated in both libraries, while AGO homologs were up and down-regulated in RdRP3 and Ago-16 knockdown libraries, respectively. Although the possibility of off-target effects of the introduced Ago-16 dsRNAs on their homologs could not be excluded, these results strongly suggested auto-regulation of the genes in sRNA-related pathways. Upon RdRP1 knockdown, stress-response-related genes were often down-regulated (Supplementary PDF).

Misregulation of the gene Dsor1 homolog (ISCI005428, hereafter IscDsor1) upon RdRP1-knockdown caught our attention because it was most strongly upregulated in this dataset (Figure 6B highlighted by blue). The annotation of *Ixodes* genes was incomplete and gene models generally lacked UTRs. We noticed that there was a strong peak of RdRP1-dependent sRNAs in the downstream region of the IscDsor1 CDS (Figure 7A). The total RNAseq data showed continuous signals for ~4kb after the IscDsor1 coding region, suggesting that the signal represented the 3' UTR of IscDsor1 (Figure 7A). Consistent with this idea, RT-PCR using primers that bind the 3' end of the CDS and the 3' end of the putative 3' UTR yielded products having the correct sequence in a reverse-transcription-

dependent manner (Figure S7). The region where a large number of sRNAs were mapped were repetitive and corresponded to the rnd-1_family-272 sequence in our repeat annotation, which showed similarity to LTR/Gypsy family transposons (Table S1). Therefore, the sRNAs targeting IscDsor1 might be produced from other copies of this TE and act in trans.

The reciprocal changes in the targeting sRNAs and the target mRNA suggested direct regulation (Figure 7B and C). We first verified that IscDsor1 was upregulated in RdRP1 knockdown cells by qPCR (Figure 7C, qPCR panel). A statistically significant increase in the IscDsor1 level was also observed upon Ago-96 knockdown, in addition to RdRP1 knockdown, suggesting that Ago-96 might also be involved in the regulation of IscDsor1 (Figure 7C). To test if RdRP1 regulates IscDsor1 through its 3'UTR, we cloned IscDsor1 3'UTR after the firefly luciferase coding region of the pmirGLO/Fer-Luc2/Act-hRluc vector (54). After depleting RdRP1 or RdRP3 in ISE6 cells, we transfected the IscDsor1 3'UTR sensor plasmid and performed dual luciferase assays to detect up-regulation of sensors upon knockdown of RdRPs (Figure 7D). Upon knockdown of RdRP1, we detected ~3-fold upregulation ($p=0.002$) of the sensor expression, whereas RdRP3 knockdown had no effect. These results demonstrated that RdRP1 regulates IscDsor1 through the 3'UTR sequence, presumably by producing endogenous siRNAs targeting it.

In summary, our results showed that the components of the sRNA pathway play roles in the regulation of mRNA expression primarily to regulate genes related to sRNA pathways. Some specific genes including IscDsor1 appear to be targeted by sRNAs produced in these pathways. While the biological significance of the regulatory mechanisms in normal development and virus-tick interactions needs to be studied in the future, this study unveiled novel and unexpected gene regulatory mechanisms involving tick-specific sRNA factors.

Discussion

Novel small regulatory RNA pathways dependent on RdRPs in ticks

In contrast to the established roles of RdRPs in plants, fungi and worms, their roles remain unclear in other animals. Although RdRP genes were found in many arthropods, their roles in sRNA production were not experimentally demonstrated mainly due to the lack of suitable experimental systems (Lewis et al., 2018; Mondal et al., 2018b, 2018a, 2020). Analysis of sRNA chemical structures of sRNAs from various animals possessing RdRP genes did not find evidence for the production of sRNAs containing terminal structures similar to those of RdRP-dependent sRNAs in worms(26). Based on these results, the importance of RdRPs in animals outside of Nematoda remained controversial.

In the present study, we demonstrated the presence of abundant classes of RdRP-dependent sRNAs in tick cells. Some of them are expressed as highly as the most abundant miRNA genes expressed in the cell line (Figure 2A), implying that they played important biological roles. The conservation of the catalytic site in the tick RdRPs suggested they are active enzymes potentially producing antisense RNA species (27, 28).

A large fraction of RdRP-dependent small RNAs was derived from RNAP III-transcribed genes (Figure 2A), pointing to a potential functional link. Transcription by RNAP III is terminated by the presence of a stretch of 5 or more Ts on the non-template strand (55) and the presence of short U-tails is a signal recognized by a quality control mechanism getting rid of defective RNAP III products that did not undergo normal processing events to remove the tails (56–58). Therefore, these U-tails may be a signal for RdRP1 to produce their antisense RNAs. The 3'-oligouridylation often acts as a signal for RdRP-dependent sRNA production in *C. elegans* in artificial RNAi or silencing of rRNA transcription upon erroneous pre-rRNA production (59, 60). However, the distance between positions of abundant small RNAs and the 3'-ends of RNAP III transcripts showed no obvious trends in contrast to the expectation that the antisense RNA production may be initiated at a certain distance from the U-tails (61). Additional evidence to support this hypothesis is lacking. Alternatively, RdRPs might physically interact with RNAP III during transcription, similarly to how RDR2 in *Arabidopsis* recognizes RNAP IV products to produce their antisense strands (28). The functional links between RdRP1 and the RNAP III machineries remain unclear, and this deserves further investigation.

Molecular mechanisms and biological roles of sRNA-mediated gene regulation in ticks

The biological roles for endo-siRNA pathways have been studied in a few animal species. In *Drosophila melanogaster*, the endogenous roles for the RNAi components were unclear due to the lack of the obvious phenotypes of mutants lacking the RNAi factors (62–64), but later,

major roles in reproductive isolation and speciation via regulating male fertility were documented by studies of multiple *Drosophila* species (65, 66). In worms, more diverse functions have been discovered ranging from transposon silencing in the germline to transgenerational regulation of gene expression (67).

What might be the roles for the tick endogenous sRNA pathways? The production of antisense sRNAs from RNAP III-dependent genes appears to be conserved in the two tick species*, I. scapularis* and *H. longicornis*, suggesting that this pathway has a conserved role (Figure 2E-F). However, we observed no discernible effects of RdRP1-KD on the expression levels of the RNAP III product SRP RNA, suggesting that the sRNAs may play roles independently of gene regulation in cis (Figure 3B). In fission yeast, genes transcribed by RNAP III are involved in the compartmentalization of the genome in the nucleus by defining boundaries between regions with distinct chromosomal states (68, 69). It is interesting to speculate that the tick sRNA pathway might also contribute to the higher order organization of the genome in the tick nucleus. As demonstrated with sRNAs IscDsor1, the sRNAs can down-regulate mRNAs containing sRNA targets in some cases (Figure 7).

The function of RdRP3-dependent pathway is even more mysterious. RdRP3 knockdown caused misregulation of 63 genes (Table S3), many of which were also misregulated in Ago-16 knockdown (Figure 6). This was consistent with the overlapping dependencies of sRNAs on these factors (Figures 3 and 4). The misregulated genes included many sRNA-related factors including Dicer homologs. However, as these loci do not produce abundant RdRP3-dependent sRNA species, how the RdRP3-Ago-16 axis controls gene expression remains unknown. It is important to identify direct targets and understand the effects of sRNAs against targets by studying gene regulation using multiple approaches including proteomics and analysis of histone modifications. Characterization of biochemical properties of the AGOs may also provide clues to the molecular functions for the novel sRNA species. In addition to molecular functions, biological roles for RdRP-dependent sRNAs need to be investigated especially in the in vivo context.

Crosstalk between viral replication and RdRP-dependent sRNA pathways

One of the main roles for invertebrate RNAi pathways is the antiviral response. This was first established in flies by showing RNAi-defective mutant flies were more susceptible to viral infection (70). However, as the *Drosophila* genome has no recognizable RdRP gene, interactions between viral infection and RdRP-dependent pathways could not be studied. In *C. elegans*, the antiviral roles for the RNAi factors were demonstrated using a heterologous system by utilizing the broad host selectivity of the Flock House Virus (FHV) (71). Later, the discovery of Orsay virus, a viral pathogen naturally infecting worms, opened the possibility of studying virus-host interactions in the nematode, leading to the discovery of a polymorphism

in the *drh-1* (Dicer-Related Helicase-1) locus that correlated with the sensitivity of the worm to the Orsay virus infection (72, 73). However, *C. elegans* mutants of RdRPs were not used in these studies to clarify their roles in the antiviral defense system. Therefore, research regarding the interactions between RdRP-dependent pathways and viruses in animals is in its infancy. However, this is clearly an important area of research because ticks, which we demonstrated to have active RdRP genes, are pests of agricultural, veterinary and medical importance (8, 74, 75).

In *Arabidopsis*, a mutant of the RdRP gene RDR6 exhibited normal antiviral responses, suggesting that RDR6 is dispensable (23, 76). On the other hand, in tobacco, knockdown of the RDR6 gene caused higher susceptibility to viruses especially at high temperatures (77). Viruses often encode proteins to suppress hosts' RNAi mechanisms (Viral Suppressors of RNA silencing or VSRs), and complex interactions between VSRs and RNAi factors often complicate the interpretation of experimental data (78). Therefore, the contribution of VSRs needs to be taken into consideration to understand interactions between viruses and host or vector cells. The knowledge regarding the RdRP-dependent sRNA pathways obtained here will help us untangle the complex interactions at the molecular level.

Various signaling pathways, including the ERK pathway, play roles in antiviral immunity (79). In *Drosophila* cells, the ERK activity is quickly upregulated upon infection of various viruses, and the ERK pathway activity is important for restricting viral replication in cultured *Drosophila* and mosquito cells as well as in the fly gut. It is interesting that one of the integral factors in the ERK pathway, Dsor1 is upregulated upon RdRP1 knockdown in tick cells (Figure 7). On the other hand, the persistently infecting viruses were reduced upon RdRP1 knockdown, which is the opposite result if we expect RdRP1 to have antiviral roles (Figure 5). These results may reflect complex interactions between the host sRNA factors and viruses. Tick cell lines have been used as an in vitro model to study virus-vector interactions for tick-borne viruses (80). Based upon the previously established tools, we provide a genomics platform to study the ticks' sRNA pathways and viruses by annotating endogenous sRNAs and their biogenesis pathways. For convenience, we made the genomics resources available to the research community, including ISE6 and *H. longicornis* UCSC genome browser assembly tracks, RNAseq mapping data tracks and the sRNA size distribution charts. The UCSC assembly hubs with the RNAseq mapping tracks are available at https://data.cyverse.org/dav-anon/iplant/home/okamuralab/trackhub/Isc_ISE6/IscaI1_hub.txt and https://data.cyverse.org/dav-anon/iplant/home/okamuralab/trackhub/ucscgb_haeL2018/hubHaeL2018.txt.Together with the initial characterization of the sRNA pathways presented in this study, these resources

will facilitate studies related to gene regulation and virus-vector interactions that are mediated by sRNAs. Furthermore, as RNAi technologies depend on sRNA factors, detailed understanding through the characterization of the sRNA factors may pave the way for the development of RNAi-based pesticides.

Materials and methods

Tick cell culture and dsRNA transfection

Ixodes scapularis embryonic 6（ISE6) cells were obtained from ATCC and cultured according to the published protocol at 34 degrees C (81). The dsRNA transfection was performed using Effectene (QIAGEN). Cells were seeded at 1x10^6 /ml in 2ml fresh L-15B medium on a 6-well plate. 400ng dsRNA was diluted in 100ul Buffer EC and 3.2ul Enhancer was added. The mixture was incubated for 2-5min at room temperature. Then, 10ul Effectene was added and incubated for 5-10min at room temperature. The mixture was added to the culture and the cells were incubated for 7-10 days. After the first incubation, the dsRNA transfection procedure was repeated again and incubated for 7-10 days to ensure the maximal efficacy of RNAi.

Plasmids and dsRNA production

cDNA was amplified using the total RNA of ISE6 cells as a template. Contaminating genomic DNA was removed by treating total RNA samples with the Turbo DNA-free kit (Ambion) and cDNA was synthesized using Superscript III (Invitrogen) according to the manufactures' instructions. The cDNA encoding Ago-16, Ago-30, RdRP1 and RdRP3 were amplified using primers listed in Table S5, and clones were obtained by inserting the amplified cDNAs in the NotI-XbaI sites of pEGFP with a modified multiple-cloning-site sequence (82). The sequences of the inserts were verified by sequencing.
For the preparation of templates for dsRNA synthesis, ~500bp fragments of small RNA factors were amplified from cDNA of ISE6 cells using primers listed in Table S1 using iProof High-Fidelity Master Mix (Bio-Rad) and the amplicons were treated with XhoI to insert them in the XhoI site of pLitmus28i (NEB). To obtain templates for in vitro transcription, LitmusA and LitmusB primers were used (Table S5). 5ul of the PCR product was used in a 20ul in vitro transcription reaction using Megascript T7 kit (Ambion). dsRNA was purified by Phenol/Chloroform extraction followed by ethanol precipitation as described previously (82).

RNA extraction and RNAseq library preparation

Total RNA was extracted from ISE6 cells using Trizol-LS (Invitrogen) according to the manufacturer's instructions. For small RNA libraries of ISE6 cells depleted of small RNA factors and their control samples, 1ug total RNA was used for library construction using the TruSeq Small RNA Library Preparation kit (Illumina). For 5'-tri-P libraries, the small RNA fraction (~15-35nt) was isolated by gel extraction and RNA species bearing 5'-OH were monophosphorylated by T4 polynucletide kinase (NEB). This was followed by treatment by a terminator exonuclease (Epicentre), dephosphorylation by Calf Intestine Phosphatase (NEB)

and re-phosphorylation by T4 polynucletide kinase (NEB). For the oxidized small RNA library, the small RNA fraction (~15-35nt) from 50ug total RNA was isolated by gel extraction and treated with 25mM $NaIO_4$ dissolved in 60mM Borax buffer and incubated for 30 minutes at room temperature in the dark. Both 5'-tir-P enriched and oxidized samples were subjected to a small RNA library construction by the TruSeq Small RNA Library Preparation kit (Illumina). The resulting libraries were sent to BGI (Hongkong) for sequencing on a Hiseq2000.

For total RNAseq analysis, three sets of knockdown experiments were performed independently, and total RNA samples extracted by Trizol-LS (Invitrogen) were sent to BGI (Hongkong) for ribosomal RNA depletion using Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat) (Epicentre) followed by library construction using non-stranded (Replicate 1) or stranded (Replicates 2 and 3) TruSeq mRNA Library Prep Kit (Illumina). The resulting libraries were sequenced on Hiseq4000.

Identification of repetitive sequences

Repetitive sequences in the ISE6 genome (Iscal1; https://metazoa.ensembl.org/Ixodes_scapularis_ise6/Info/Index) were identified using the RepeatModeler2 pipeline (83). The reference TE sequences were downloaded from Repbase < https://www.girinst.org/server/RepBase/> and the RepBaseRepeatMaskerEdition-20181026.tar file was used for annotation.

Analysis of small RNAseq data

Small RNA libraries were analyzed following the previously established pipeline (84). Adaptor sequences were removed by using fastx_clipper and collapsed by fastx_collapser and converted by fasta_formatter (The tools were downloaded from <http://hannonlab.cshl.edu/fastx_toolkit/commandline.html>). The fasta files were used for mapping using the ISE6 genome sequence (Iscal1) (85) using bowtie1.3.0 < http://bowtie-bio.sourceforge.net/index.shtml> without allowing any mismatches. For identification of sRNA sources, genome-mapping reads were mapped to the following reference sequences: 1. miRNAs (miRbase Release 22.1) (86), RNAP III transcripts (downloaded from RNAcentral) (87), rRNAs from NCBI and mRNAs (Vectorbase Iscal1.0) (85). As 5S rRNA loci are separated from the rDNA repeats and transcribed by RNAP III, we added the 5S rRNA sequences to the "RNAP III" group. The reference sequences are summarized in Table S1.

Phylogenetic tree construction

Based on multiple sequence alignment with MUSCLE, we used ModelFinder (88) to determine the best-fit model and obtained branch supports with the ultrafast bootstrap (89) implemented in the IQ-TREE software (90).

Analysis of total RNAseq data

Total RNAseq libraries were analyzed as described previously (91). The adaptor sequences were trimmed using Cutadapt (92) with the default quality cutoff value (20). Gene expression was quantified by salmon (93) using the Vectorbase Iscal1.0 annotation or a fasta file containing the viral genome sequences persistently present in the ISE6 culture (Table S2; Nakao et al., 2017). Genome mapping was performed by bowtie2 <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml> with the default setting. Differential gene expression analysis was done using the DESeq2 package (94), with the cut-off of adjusted-p-value set to 0.05. GO-term enrichment analysis of misregulated genes was conducted using clusterProfiler (95), with GO terms obtained from eggNOG-mapper (96) and TRAPID (97). The transcriptome was assembled using Trinity (98). A blast database was constructed using the assembled cDNA sequences, and homologs of RNAi factors were identified by tblastn https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=Download using the bait sequences (RRF-1, EGO-1 and RRF-3 from *C. elegans*; AGO1, AGO2, Aub, AGO3 and PIWI from *D. melanogaster*). MUSCLE (Figure 1 and Figure S1) or CLUSTAL-O (Figure S2) was used for alignment.

Northern blotting

Northern blotting was done as described previously (82). Briefly, 10ug total RNA samples were separated on a 15% Sequagel (National Diagnostics) and transferred onto a positively charged nylon membrane and hybridized using DNA or LNA probes. The probe sequences are listed in Table S5.

Western blotting

Western blotting was performed as previously described (99). Table S5 lists antibodies used in this study.

Accession number

The small RNA library data produced for this study are deposited at NCBI SRA under GSE183810.

Acknowledgments

The authors are grateful to members of the K.O. laboratories in TLL and NAIST for discussion. We thank Dr. Masami Shiimori for critically reading the manuscript. We thank ATCC for ISE6 cells and BGI for Illumina sequencing. Research in K.O.'s group was supported by the National Research Foundation, Prime Minister's Office, Singapore under its NRF Fellowship Programme (NRF2011NRF-NRFF001-042), Temasek Life Sciences Laboratory core funding and the JSPS Fund for the Promotion of Joint International Research (Returning Researcher Development Research, 17K20145). Work in the T.T.'s group was supported by Takeda Science Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of these agencies.

Author contributions: C.F., K.To., L.-L.C., and K.O. planned and performed bioinformatics analysis. M.L., K.Ts., T.T., J.I. and K.O. planned and performed experiments. C.F. analyzed transcriptome data and designed analyses. K.Ts performed experiments using HEK293T cells. M.L. generated total RNAseq libraries, cloned RdRP/AGO cDNAs and characterized the proteins. L.-L.C. processed sequencing data and performed initial characterization of mRNAs and sRNAs. K.To. identified repeats and performed sRNA analysis. T.T. produced tick plasmids and designed *H. longicornis* analysis. J.I. performed other experiments including Western blots, Northern blots and qPCR. K.O. produced sRNA libraries. K.O. and T.T. provided funds. C.F. and K.O. wrote the initial draft. All authors edited and approved the manuscript.

References

1. S. A. Gorski, J. Vogel, J. A. Doudna, RNA-based recognition and targeting: sowing the seeds of specificity. *Nat. Rev. Mol. Cell Biol.* **18**, 215–228 (2017).
2. P. Mohanraju, *et al.*, Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science (80-.).* **353**, aad5147 (2016).
3. H. Cerutti, J. A. Casas-Mollano, On the origin and functions of RNA-mediated silencing: from protists to man. *Curr. Genet.* **50**, 81–99 (2006).
4. A. Grimson, *et al.*, Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**, 1193–1197 (2008).
5. K. Okamura, E. C. Lai, Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biol* **9**, 673–678 (2008).
6. M. C. Siomi, K. Sato, D. Pezic, A. A. Aravin, PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **12**, 246–258 (2011).
7. D. M. Ozata, I. Gainetdinov, A. Zoch, D. O'Carroll, P. D. Zamore, PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89–108 (2019).
8. M. Mondal, K. Mansfield, A. Flynt, siRNAs and piRNAs collaborate for transposon control in the two-spotted spider mite. *RNA* **24**, 899–907 (2018).
9. M. Mondal, J. K. Brown, A. Flynt, Exploiting somatic piRNAs in Bemisia tabac i enables novel gene silencing through RNA feeding. *Life Sci. Alliance* **3**, e202000731 (2020).
10. M. Mondal, P. Klimov, A. S. Flynt, Rewired RNAi-mediated genome surveillance in house dust mites. *PLOS Genet.* **14**, e1007183 (2018).
11. S. H. Lewis, *et al.*, Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat. Ecol. Evol.* **2**, 174–181 (2018).
12. O. S. Akbari, *et al.*, The Developmental Transcriptome of the Mosquito Aedes aegypti , an Invasive Species and Major Arbovirus Vector. *G3: Genes|Genomes|Genetics* **3**, 1493–1509 (2013).
13. S. W. Ding, RNA-based antiviral immunity. *Nat Rev Immunol* **10**, 632–644 (2010).
14. X.-H. Wang, *et al.*, RNA Interference Directs Innate Immunity Against Viruses in Adult Drosophila. *Science (80-.).* **312**, 452–454 (2006).
15. Q. Wu, *et al.*, Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *Proc Natl Acad Sci U S A* **107**, 1606–1611 (2010).
16. J. A. Mondotte, *et al.*, Immune priming and clearance of orally acquired RNA viruses in Drosophila. *Nat. Microbiol.* **3**, 1394–1403 (2018).
17. D. J. Obbard, F. M. Jiggins, D. L. Halligan, T. J. Little, Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr Biol* **16**, 580–585 (2006).
18. D. J. Obbard, K. H. Gordon, A. H. Buck, F. M. Jiggins, The evolution of RNAi as a

defence against viruses and transposable elements. *Philos Trans R Soc L. B Biol Sci* **364**, 99–115 (2009).

19. D. S. Schwarz, G. Hutvagner, B. Haley, P. D. Zamore, Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. *Mol Cell* **10**, 537–548 (2002).
20. T. Sijen, *et al.*, On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**, 465–476 (2001).
21. A. Smardon, *et al.*, EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. *Curr. Biol.* **10**, 169–178 (2000).
22. C. Cogoni, G. Macino, Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. *Proc. Natl. Acad. Sci.* **94**, 10233–10238 (1997).
23. T. Dalmay, A. Hamilton, S. Rudd, S. Angell, D. C. Baulcombe, An RNA-Dependent RNA Polymerase Gene in Arabidopsis Is Required for Posttranscriptional Gene Silencing Mediated by a Transgene but Not by a Virus. *Cell* **101**, 543–553 (2000).
24. S. Kurscheid, *et al.*, Evidence of a tick RNAi pathway by comparative genomics and reverse genetics screen of targets with known loss-of-function phenotypes in Drosophila. *BMC Mol Biol* **10**, 26 (2009).
25. B. T. Nganso, N. Sela, V. Soroker, A genome-wide screening for RNAi pathway proteins in Acari. *BMC Genomics* **21**, 791 (2020).
26. N. Pinzón, *et al.*, Functional lability of RNA-dependent RNA polymerases in animals. *PLOS Genet.* **15**, e1007915 (2019).
27. P. Sarkies, *et al.*, Ancient and Novel Small RNA Pathways Compensate for the Loss of piRNAs in Multiple Independent Nematode Lineages. *PLOS Biol.* **13**, e1002061 (2015).
28. V. Mishra, *et al.*, Assembly of a dsRNA synthesizing complex: RNA-DEPENDENT RNA POLYMERASE 2 contacts the largest subunit of NUCLEAR RNA POLYMERASE IV. *Proc. Natl. Acad. Sci.* **118**, e2019276118 (2021).
29. K. Okamura, Diversity of animal small RNA pathways and their biological utility. *Wiley Interdiscip. Rev. RNA* **3**, 351–368 (2012).
30. E. Schnettler, *et al.*, Induction and suppression of tick cell antiviral RNAi responses by tick-borne flaviviruses. *Nucleic Acids Res.* **42**, 9436–9446 (2014).
31. N. T. Schirle, I. J. MacRae, The crystal structure of human Argonaute2. *Science (80-.).* **336**, 1037–1040 (2012).
32. J. D. Oliver, A. S. O. Chávez, R. F. Felsheim, T. J. Kurtti, U. G. Munderloh, An Ixodes scapularis cell line with a predominantly neuron-like phenotype. *Exp. Appl. Acarol.* **66**,

427–442 (2015).
33. B. Czech, *et al.*, piRNA-Guided Genome Defense: From Biogenesis to Silencing. *Annu. Rev. Genet.* **52**, 131–157 (2018).
34. Y. Kirino, *et al.*, Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat Cell Biol* **11**, 652–658 (2009).
35. R. Yashiro, *et al.*, Piwi Nuclear Localization and Its Regulatory Mechanism in Drosophila Ovarian Somatic Cells. *Cell Rep.* **23**, 3647–3657 (2018).
36. J. Wen, *et al.*, Diversity of miRNAs, siRNAs, and piRNAs across 25 Drosophila cell lines. *Genome Res* **24**, 1236–1250 (2014).
37. J. Pak, J. M. Maniar, C. C. Mello, A. Fire, Protection from Feed-Forward Amplification in an Amplified RNAi Mechanism. *Cell* **151**, 885–899 (2012).
38. Z. Yang, HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res.* **34**, 667–675 (2006).
39. F. Frank, N. Sonenberg, B. Nagar, Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818–822 (2010).
40. R. H. Huang, Unique 2′-O-Methylation by Hen1 in Eukaryotic RNA Interference and Bacterial RNA Repair. *Biochemistry* **51**, 4087–4095 (2012).
41. H. Seitz, M. Ghildiyal, P. D. Zamore, Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA strands in flies. *Curr Biol* **18**, 147–151 (2008).
42. M. Ghildiyal, *et al.*, Endogenous siRNAs Derived from Transposons and mRNAs in Drosophila Somatic Cells. *Science (80-.).* **320**, 1077–1081 (2008).
43. M. D. Horwich, *et al.*, The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol* **17**, 1265–1272 (2007).
44. M. Hafner, *et al.*, Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **44**, 3–12 (2008).
45. W. Gu, *et al.*, Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Mol Cell* **36**, 231–244 (2009).
46. J. Pak, A. Fire, Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science (80-.).* **315**, 241–244 (2007).
47. M. Nawaz, *et al.*, miRNA profile of extracellular vesicles isolated from saliva of Haemaphysalis longicornis tick. *Acta Trop.* **212**, 105718 (2020).
48. W. Liu, *et al.*, Analysis of microRNA expression profiles dynamic in different life stages of Haemaphysalis longicornis ticks by deep sequencing of small RNA libraries. *Ticks Tick. Borne. Dis.* **11**, 101427 (2020).
49. N. Jia, *et al.*, Large-Scale Comparative Analyses of Tick Genomes Elucidate Their Genetic Diversity and Vector Capacities. *Cell* **182**, 1328-1340.e13 (2020).

50. J. M. Flynn, *et al.*, RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci.* **117**, 9451–9457 (2020).
51. B. C. Bonning, M.-C. Saleh, The Interplay Between Viruses and RNAi Pathways in Insects. *Annu. Rev. Entomol.* **66**, 61–79 (2021).
52. R. Nakao, *et al.*, Putative RNA viral sequences detected in an Ixodes scapularis-derived cell line. *Ticks Tick. Borne. Dis.* **8**, 103–111 (2017).
53. Z. Guo, Y. Li, S.-W. Ding, Small RNA-based antimicrobial immunity. *Nat. Rev. Immunol.* **19**, 31–44 (2019).
54. E. P. Hernandez, K. Kusakisako, T. Hatta, T. Tanaka, Characterization of an iron-inducible Haemaphysalis longicornis tick-derived promoter in an Ixodes scapularis-derived tick cell line (ISE6). *Parasit. Vectors* **12**, 321 (2019).
55. A. G. Arimbasseri, R. J. Maraia, Mechanism of Transcription Termination by RNA Polymerase III Utilizes a Non-template Strand Sequence-Specific Signal Element. *Mol. Cell* **58**, 1124–1132 (2015).
56. D. Ustianenko, *et al.*, TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. *EMBO J.* **35**, 2179–2191 (2016).
57. M. M. Reimão-Pinto, *et al.*, Molecular basis for cytoplasmic <scp>RNA</scp> surveillance by uridylation-triggered decay in *Drosophila*. *EMBO J.* **35**, 2417–2434 (2016).
58. A. Łabno, *et al.*, Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Res.*, gkw649 (2016).
59. H.-Y. Tsai, *et al.*, A Ribonuclease Coordinates siRNA Amplification and mRNA Cleavage during RNAi. *Cell* **160**, 407–419 (2015).
60. X. Zhou, *et al.*, RdRP-synthesized antisense ribosomal siRNAs silence pre-rRNA via the nuclear RNAi pathway. *Nat. Struct. Mol. Biol.* **24**, 258–269 (2017).
61. H.-Y. Tsai, *et al.*, A Ribonuclease Coordinates siRNA Amplification and mRNA Cleavage during RNAi. *Cell* **160**, 407–419 (2015).
62. Y. S. Lee, *et al.*, Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways. *Cell* **117**, 69–81 (2004).
63. K. Okamura, A. Ishizuka, H. Siomi, M. C. Siomi, Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* **18**, 1655–1666 (2004).
64. Q. Liu, *et al.*, R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science (80-.).* **301**, 1921–1925 (2003).
65. J. Wen, *et al.*, Adaptive regulation of testis gene expression and control of male fertility by the Drosophila hairpin RNA pathway. [Corrected]. *Mol Cell* **57**, 165–178 (2015).

66. C.-J. Lin, *et al.*, The hpRNA/RNAi Pathway Is Essential to Resolve Intragenomic Conflict in the Drosophila Male Germline. *Dev. Cell* **46**, 316-326.e5 (2018).
67. O. Rechavi, I. Lev, Principles of Transgenerational Small RNA Inheritance in Caenorhabditis elegans. *Curr. Biol.* **27**, R720–R730 (2017).
68. O. Iwasaki, A. Tanaka, H. Tanizawa, S. I. S. Grewal, K. Noma, Centromeric Localization of Dispersed Pol III Genes in Fission Yeast. *Mol. Biol. Cell* **21**, 254–265 (2010).
69. K. Noma, H. P. Cam, R. J. Maraia, S. I. S. Grewal, A Role for TFIIIC Transcription Factor Complex in Genome Organization. *Cell* **125**, 859–872 (2006).
70. X. H. Wang, *et al.*, RNA interference directs innate immunity against viruses in adult Drosophila. *Science (80-.).* **312**, 452–454 (2006).
71. R. Lu, *et al.*, Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans. *Nature* **436**, 1040–1043 (2005).
72. M.-A. Félix, *et al.*, Natural and Experimental Infection of Caenorhabditis Nematodes by Novel Viruses Related to Nodaviruses. *PLoS Biol.* **9**, e1000586 (2011).
73. A. Ashe, *et al.*, A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* **2** (2013).
74. R. Lani, *et al.*, Tick-borne viruses: A review from the perspective of therapeutic approaches. *Ticks Tick. Borne. Dis.* **5**, 457–465 (2014).
75. Y. Garbian, E. Maori, H. Kalev, S. Shafir, I. Sela, Bidirectional Transfer of RNAi between Honey Bee and Varroa destructor: Varroa Gene Silencing Reduces Varroa Population. *PLoS Pathog.* **8**, e1003035 (2012).
76. P. Mourrain, *et al.*, Arabidopsis SGS2 and SGS3 Genes Are Required for Posttranscriptional Gene Silencing and Natural Virus Resistance. *Cell* **101**, 533–542 (2000).
77. F. Qu, *et al.*, RDR6 Has a Broad-Spectrum but Temperature-Dependent Antiviral Defense Role in Nicotiana benthamiana. *J. Virol.* **79**, 15209–15217 (2005).
78. T. Csorba, L. Kontra, J. Burgyán, viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology* **479**–**480**, 85–103 (2015).
79. J. Xu, *et al.*, ERK signaling couples nutrient status to antiviral defense in the insect gut. *Proc. Natl. Acad. Sci.* **110**, 15025–15030 (2013).
80. L. Bell-Sakyi, E. Zweygarth, E. F. Blouin, E. A. Gould, F. Jongejan, Tick cell lines: tools for tick and tick-borne disease research. *Trends Parasitol.* **23**, 450–457 (2007).
81. U. G. Munderloh, T. J. Kurtti, Formulation of medium for tick cell culture. *Exp. Appl. Acarol.* **7**, 219–229 (1989).
82. K. Okamura, W. J. Chung, E. C. Lai, The long and short of inverted repeat genes in animals: microRNAs, mirtrons and hairpin RNAs. *Cell Cycle* **7**, 2840–2845 (2008).

83. J. M. Flynn, *et al.*, RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci.* **117**, 9451–9457 (2020).
84. L. Zhou, *et al.*, Importance of miRNA stability and alternative primary miRNA isoforms in gene regulation during Drosophila development. *Elife* **7** (2018).
85. J. R. Miller, *et al.*, A draft genome sequence for the Ixodes scapularis cell line, ISE6. *F1000Research* **7**, 297 (2018).
86. A. Kozomara, M. Birgaoanu, S. Griffiths-jones, miRBase : from microRNA sequences to function. **47**, 155–162 (2019).
87. B. A. Sweeney, *et al.*, RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res.* **49**, D212–D220 (2021).
88. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
89. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
90. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
91. L. Zhou, *et al.*, Importance of miRNA stability and alternative primary miRNA isoforms in gene regulation during Drosophila development. *Elife* **7** (2018).
92. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* **17** (2009).
93. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
94. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
95. T. Wu, *et al.*, clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov.* **2**, 100141 (2021).
96. C. Carlos, A. Hernández-Plaza, I. Letunic, P. Bork, J. Huerta-Cepas, eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *bioRxiv* **https://do** (2021).
97. F. Bucchini, *et al.*, TRAPID 2.0: a web application for taxonomic and functional analysis of de novo transcriptomes. *Nucleic Acids Res.* (2021) https:/doi.org/10.1093/nar/gkab565.
98. M. G. Grabherr, *et al.*, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

99. M. Y. Lim, *et al.*, The Drosophila Dicer-1 Partner Loquacious Enhances miRNA Processing from Hairpins with Unstable Structures at the Dicing Site. *Cell Rep* **15**, 1795–1808 (2016).

Supplementary Tables

Table S1. General bioinformatics information.

Sheet1: Statistics of the sRNAseq libraries.

Sheet2: Reference sequences used for classification of sRNA origins.

Sheet3: ISCI IDs of genes with sequences that matched with the dsRNA sequences used for the knockdown.

Sheet4: Trinity contigs and Gene IDs of the AGO/RdRP genes analyzed in this study.

Sheet5: References for amino acid sequences of AGO/RdRPs used for the phylogenetic analysis.

Sheet6: Repeat families identified by RepeatModeler2/RepeatMasker

Sheet7: Statistics of the total RNAseq libraries.

Sheet8: List of the *H. longicornis* sRNA libraries analyzed in this study.

Table S2. Summary of sRNA analysis

Sheets1-10: Related to Figure 2A and Figure S4. sRNA read counts for each category are shown. Each sheet reports the numbers in each of the knockdown libraries. The sum of sense and antisense reads (C18-C30), the number of sense reads (S18-S30) and the number of antisense reads (AS18-AS30) that matched the reference sequence in the category in the header row are reported for each length.

Sheet11: Normalized counts of sRNAseq and total RNAseq reads mapping to the persistently present viruses.

Table S3. Summary of DGE analysis

The results of DGE analysis using the total RNAseq libraries by the Salmon-DESeq2 pipeline are summarized. Sheet1: Ago-16 KD vs GFP control, Sheet2 RdRP1 KD vs GFP control, Sheet3: RdRP3 KD vs GFP control, Sheet4: TPM values for individual libraries.

Table S4. Summary of expression analysis for repeats

The results of DGE analysis using the total RNAseq libraries by the Salmon-DESeq2 pipeline are summarized. Sheet1: Ago-16 KD vs GFP control, Sheet2 RdRP1 KD vs GFP control, Sheet3: RdRP3 KD vs GFP control, Sheet4: TPM values for individual libraries. These analyses were done using a reference file containing both protein-coding genes and repeats and data only for repeats are shown.

Table S5. Materials used in this study.

The workbook contains the information of oligos (Sheet1), cell lines (Sheet2) and antibodies (Sheet3).

Supplementary Data

Links to the sortable tables of CDS- and TE-derived sRNAs in the knockdown library can be found in TE_cDNA_links.html. On the Normalized and Relative levels pages, normalized read counts (RPM) and relative levels compared to the control (dsGFP) library were used.

Supplementary PDF

Contains dot plots and Gene-GO networks visualizing the results of GO enrichment analysis for misregulated genes upon Ago-16/RdRP1/RdRP3 knockdown.

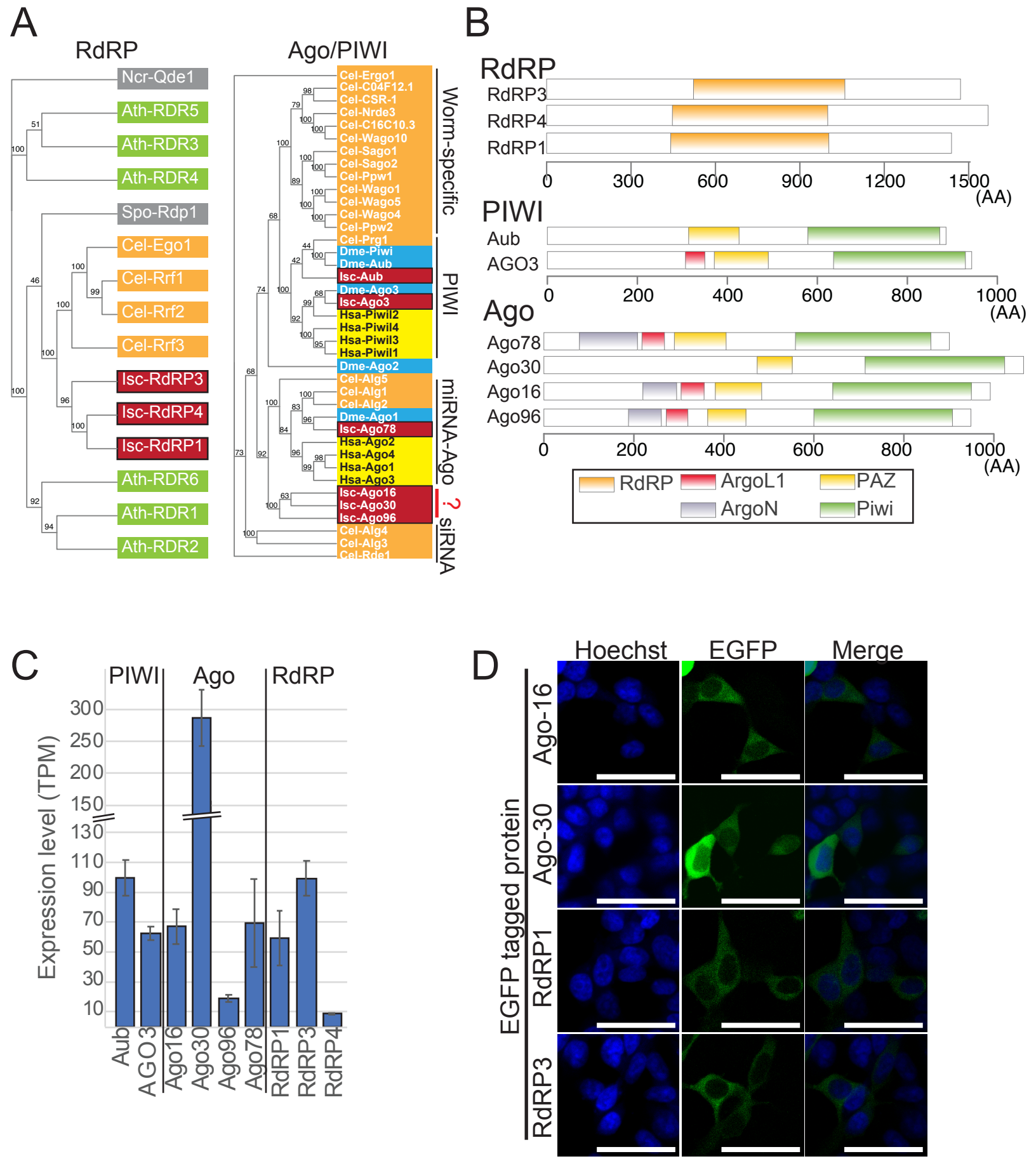

A
RdRP
Ncr-Qde1
Ath-RDR5
Ath-RDR3
Ath-RDR4
Spo-Rdp1
Cel-Ego1
Cel-Rrf1
Cel-Rrf2
Cel-Rrf3
Isc-RdRP3
Isc-RdRP4
Isc-RdRP1
Ath-RDR6
Ath-RDR1
Ath-RDR2
Ago/PIWI
Cel-Ergo1
Cel-C04F12.1
Cel-CSR-1
Cel-Nrde3
Cel-C16C10.3
Cel-Wago10
Cel-Sago1
Cel-Sago2
Cel-Ppw1
Cel-Wago1
Cel-Wago5
Cel-Wago4
Cel-Ppw2
Cel-Prg1
Dme-Piwi
Dme-Aub
Isc-Aub
Dme-Ago3
Isc-Ago3
Hsa-Piwil2
Hsa-Piwil4
Hsa-Piwil3
Hsa-Piwil1
Dme-Ago2
Cel-Alg5
Cel-Alg1
Cel-Alg2
Dme-Ago1
Isc-Ago78
Hsa-Ago2
Hsa-Ago4
Hsa-Ago1
Hsa-Ago3
Isc-Ago16
Isc-Ago30
Isc-Ago96
Cel-Alg4
Cel-Alg3
Cel-Rde1
Worm-specific
PIWI
miRNA-Ago
?
siRNA
B
RdRP
RdRP3
RdRP4
RdRP1
0
300
600
900
1200
1500
(AA)
PIWI
Aub
AGO3
0
200
400
600
800
1000
(AA)
Ago
Ago78
Ago30
Ago16
Ago96
0
200
400
600
800
1000
(AA)
RdRP
ArgoL1
PAZ
ArgoN
Piwi
C
PIWI
Ago
RdRP
Expression level (TPM)
300
250
200
150
130
110
90
70
50
30
10
Aub
AGO3
Ago16
Ago30
Ago96
Ago78
RdRP1
RdRP3
RdRP4
D
Hoechst
EGFP
Merge
EGFP tagged protein
Ago-16
Ago-30
RdRP1
RdRP3

Figure 1

Figure 1. Characterization of *Ixodes* RNAi factors

(A) Phylogenetic analysis of RdRP and AGO genes. The sequences of AGO and RdRP homologs from representative organisms (*Ixodes, C. elegans, Arabidopsis thaliana, Neurospora crassa and Schizosaccharomyces pombe for RdRP; Ixodes, Drosophila, C. elegans and human for Ago/PIWI*) were aligned using MUSCLE. Gene names are surrounded by colored rectangles depending on the species (Fungi: gray, *A. thaliana*: green, *C. elegans*: orange, *Drosophila*: blue, human: yellow, *Ixodes*: dark red).

(B) Protein domains found in *I. scapularis* RdRP, PIWI and AGO genes. The domains were identified using the CD-search tool (NIH/NLM/NCBI).

(C) Expression of tick RNAi factors analyzed by RNAseq. The dataset of ISE6 total RNAseq from the control sample (dsGFP transfection) was used to determine the expression levels of the indicated genes. Expression levels are shown as TPM and the averages and standard deviations were used for the bar chart (n=3).

(D) Subcellular localization of tick AGOs and RdRPs in HEK293T cells. HEK293T cells were transfected with the indicated plasmids and fixed. The cells were stained with Hoechst and observed by confocal microscopy and the bars indicate 50um.

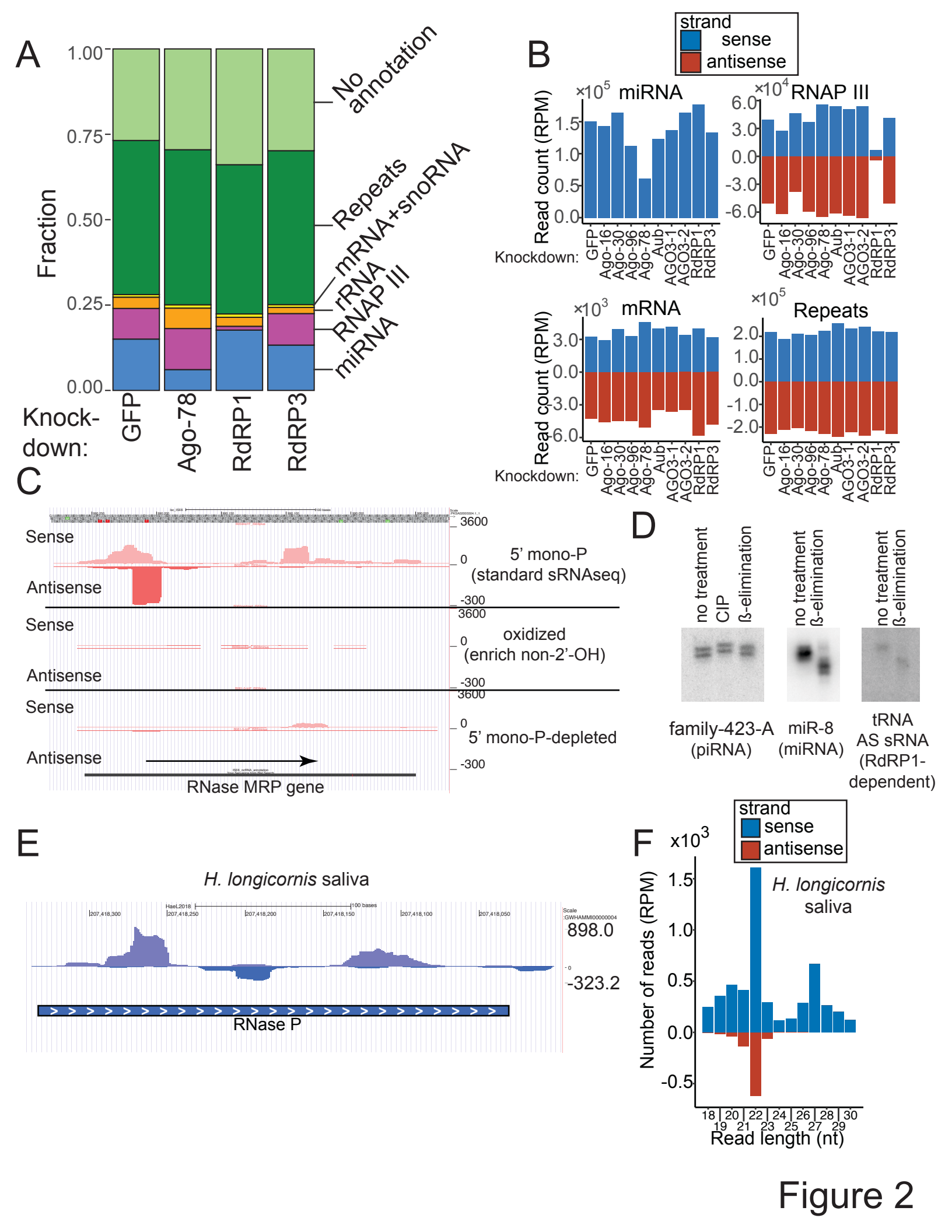

A
Fraction
1.00
0.75
0.50
0.25
0.00
No annotation
Repeats
mRNA+snoRNA
rRNA
RNAP III
miRNA
Knock-down:
GFP
Ago-78
RdRP1
RdRP3
B
strand
sense
antisense
miRNA
RNAP III
mRNA
Repeats
Read count (RPM)
Knockdown:
GFP
Ago-16
Ago-30
Ago-96
Ago-78
Aub
AGO3-1
AGO3-2
RdRP1
RdRP3
C
Sense
Antisense
5' mono-P
(standard sRNAseq)
oxidized
(enrich non-2'-OH)
5' mono-P-depleted
3600
0
-300
RNase MRP gene
D
no treatment
CIP
ß-elimination
family-423-A
(piRNA)
miR-8
(miRNA)
tRNA
AS sRNA
(RdRP1-
dependent)
E
H. longicornis saliva
898.0
-323.2
RNase P
F
strand
sense
antisense
H. longicornis
saliva
Number of reads (RPM)
Read length (nt)

Figure 2

Figure 2. Characterization of endogenous small RNA populations in ISE6 cells.
(A) Genomic origins of sRNA populations in ISE6 cells. The fractions of sRNAs derived from miRNAs (blue), RNAP III transcribed genes (pink), rRNAs (orange), snoRNAs and mRNAs (yellow) and repetitive sequences (green) are shown. sRNA reads whose genomic loci lack a gene annotation were shown as “No annotation” (light green). The percentage of reads falling within each category was calculated by dividing the number of reads in the category by the number of reads mapping to the ISE6 genomic sequence. Note the dramatic decrease of the “RNAP III” group (purple) in the RdRP1-KD sample, suggesting that this category included abundant RdRP1-dependent sRNAs.
(B) Normalized counts (RPM) of sense and antisense reads that mapped to the reference sequences in miRNAs, RNAP III transcripts, mRNAs and repeats in the indicated knockdown libraries are shown in blue and red, respectively. Negative values were given to antisense read counts.
(C) Example of small RNAs in the “RNAP III” category. A UCSC screenshot of the locus of the RNAP III transcribed gene, RNase MRP RNA gene is shown. On the track of “standard sRNAseq library”, reads corresponding to an abundant 22nt species are mapped on the antisense strand of the RNase MRP RNA locus, and there are reads throughout both sense and antisense strands of this locus. To analyze chemical structures at their 3’ nucleotides, we generated small RNA libraries after oxidizing RNA samples to enrich for small RNA species containing chemical modifications at the 2’-position. Nucleotides with free 2’-, 3’-OH groups react with periodate to form dialdehydes, which are not compatible with the 3’-linker ligation. Therefore, species with 2’-O-me modification will be overrepresented in the oxidized library. Small RNAs from RNAP III-transcribed genes tended to be depleted in the oxidized library suggestive of the presence of vicinal hydroxyl groups at the 3’-nucleotide of the sRNAs. The phosphorylation status of the 5’-end could also be analyzed by taking advantage of the 5’-mono-P specificity of T4 RNA ligase, which is used for library construction. For standard sRNA library construction, the 5’-linker ligation step strongly enriches for 5’-mono-P species. For the “5’-tri-P” library, small RNAs with 5’-di-P or 5’-tri-P groups were enriched by the removal of small RNAs with 5’-mono-P and 5’-OH groups by terminator exonuclease (See Materials and Methods). No enrichment was seen with the sRNA species from RNAP III-dependent loci in this library, suggesting that the sRNAs were 5’-mono-phosphorylated.
(D) 5’- and 3’-states of a piRNA (TE- family-423-A), an antisense sRNA from a RNAP III dependent gene (tRNA ISCW004624) and a miRNA (miR-8) were verified by Northern blotting. The removal of phosphate groups at the 5’-end causes a delay in the migration on the gel. The alkaline treatment removes the oxidized RNA with dialdehydes at the 3’ ends,

resulting in a faster migration of the RNA on the gel (ß-elimination). The results confirmed that sRNAs from RNAP III-transcribed loci had 5'-mono-P, and 2'- 3'-OH groups.
(E, F) Antisense sRNAs are produced from the RNAP III-dependent gene RNase P in longhorned tick animals. sRNA mapping at the RNase P locus is shown (E). The small RNA library was made using RNA samples extracted from purified extracellular vesicles in saliva. The size distribution of sRNAs mapping to the representative RNAP III-dependent genes (RNase P, RNase MRP and SRP RNA) are shown. RPM normalized sense and antisense read counts are shown in blue or red bars, respectively. Note that strong peaks at 22nt were observed in saliva and tick animals at different developmental stages at least on the antisense strand, indicating the conservation of an sRNA production mechanism similar to those observed in ISE6 cells (Figure S5).

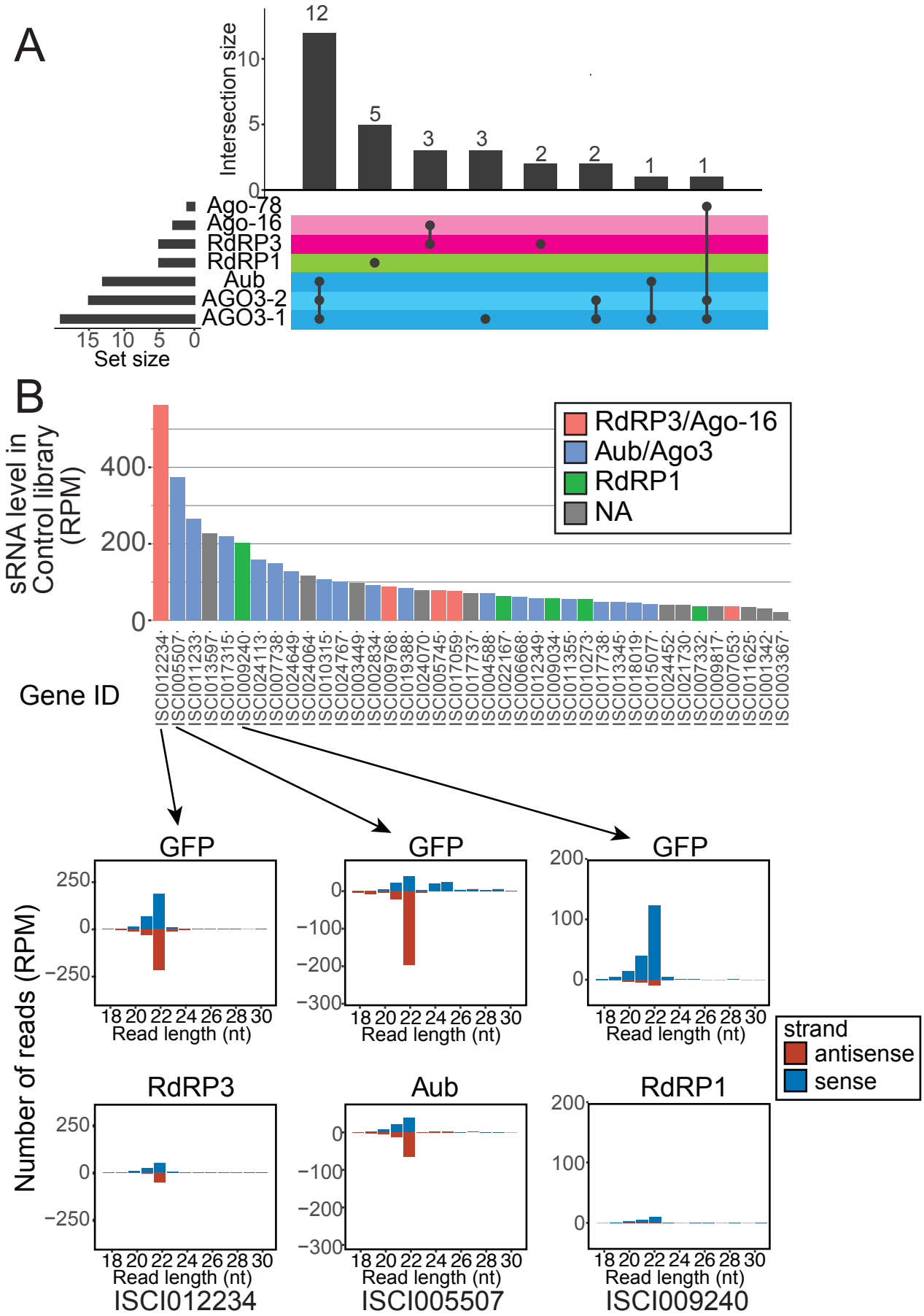

A
Intersection size
12
5
3
3
2
2
1
1
Ago-78
Ago-16
RdRP3
RdRP1
Aub
AGO3-2
AGO3-1
Set size
B
sRNA level in Control library (RPM)
RdRP3/Ago-16
Aub/Ago3
RdRP1
NA
Gene ID
GFP
RdRP3
Aub
RdRP1
Number of reads (RPM)
Read length (nt)
strand
antisense
sense
ISCI012234
ISCI005507
ISCI009240

Figure 3

Figure 3. sRNAs from coding regions

(A) UpSet plot of RNAi factor dependencies for sRNA-producing coding genes. The genes that had >35 RPM on average in the ten sRNA libraries were considered. If the sRNA reads from the locus were reduced by >40% in the knockdown library, the sRNA was judged as "dependent" on the factor that was knocked down. The number of dependent genes in each group is shown. Note that many genes were dependent on the Aub-AGO3 (blue) and RdRP3-Ago-16 (pink) combinations.

(B) Read counts from protein-coding sequences in the control GFP-KD library are shown in the bar chart. sRNA reads derived from the annotated coding regions that had >35 RPM on average in the ten knockdown sRNA libraries are shown. If small RNA reads were reduced by >40% upon knockdown of one of the indicated sRNA factors, the gene was judged dependent on the factors. sRNA read size distributions of a representative locus in each group are shown in the bar charts below.

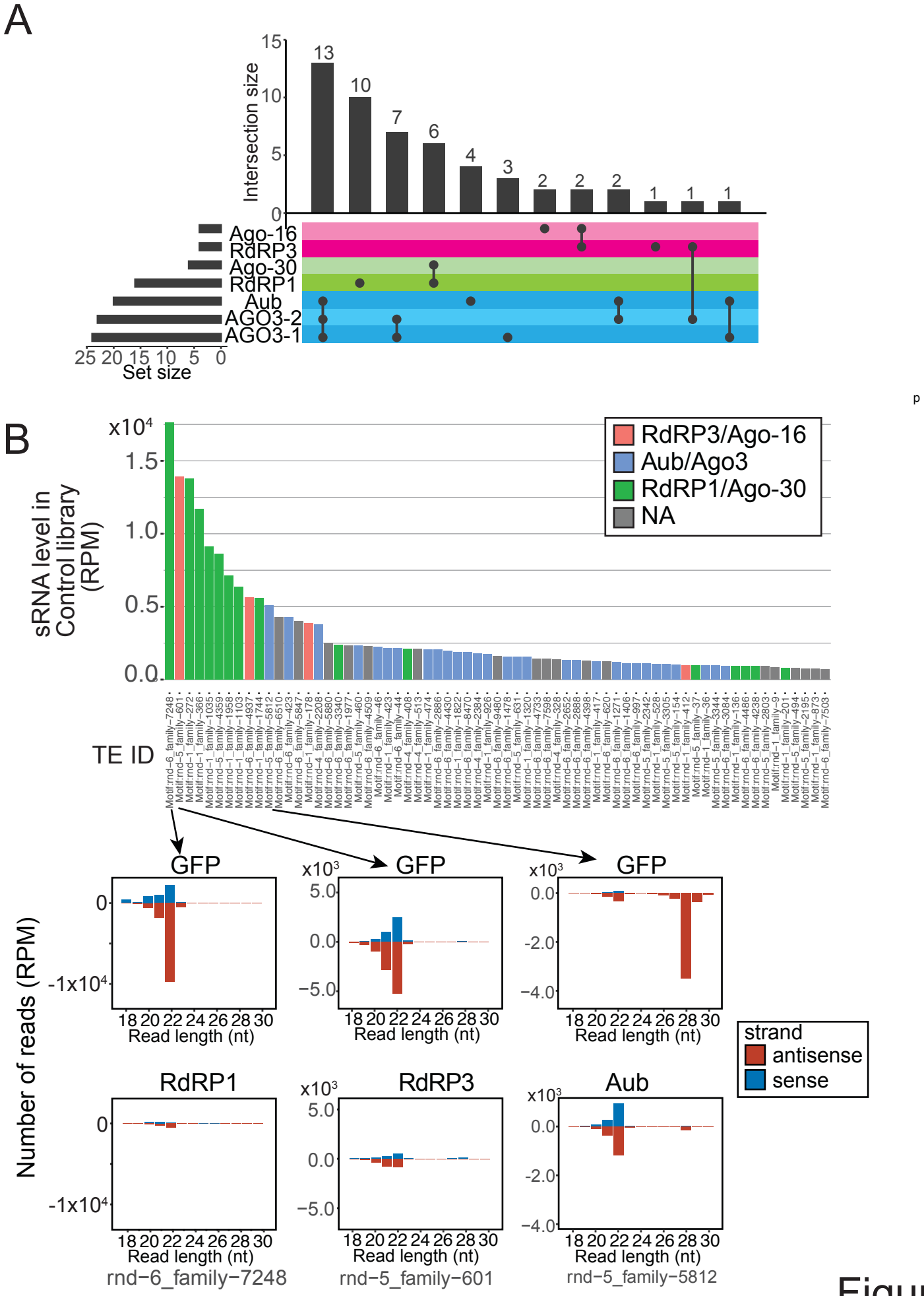

A
Intersection size
15
10
5
0
13
10
7
6
4
3
2
2
2
1
1
1
Ago-16
RdRP3
Ago-30
RdRP1
Aub
AGO3-2
AGO3-1
25 20 15 10 5 0
Set size
B
x10^4
1.5
1.0
0.5
0.0
sRNA level in Control library (RPM)
RdRP3/Ago-16
Aub/Ago3
RdRP1/Ago-30
NA
TE ID
GFP
GFP
GFP
RdRP1
RdRP3
Aub
Number of reads (RPM)
x10^3
5.0
0.0
−5.0
-2.0
−4.0
-1x10^4
18 20 22 24 26 28 30
Read length (nt)
strand
antisense
sense
rnd−6_family−7248
rnd−5_family−601
rnd−5_family−5812

Figure 4

Figure 4. Small RNAs from repeats

(A) UpSet plot of RNAi factor dependencies for sRNA-producing repeatss. The repeatss that had >800 RPM on average in the ten sRNA libraries were considered. If the sRNA reads from the locus were reduced by >40% in the knockdown library, the sRNA was judged as "dependent" on the factor that was knocked down. The number of dependent genes in each group is shown. Note that many genes were dependent on the Aub-AGO3 (blue), RdRP3-Ago-16 (pink) and RdRP1-Ago-30 (green) combinations.

(B) Read counts of repeat-associated sRNAs in the control GFP-KD library are shown in the bar chart. Repeats that had >800 RPM on average in the ten knockdown sRNA libraries are shown. Bars are color-coded based on their processing factor dependency. If small RNA reads were reduced by >40% upon knockdown of one of the indicated sRNA factors, the gene was judged dependent on the factors. sRNA read size distributions of a representative repeat in each group are shown in the bar charts below.

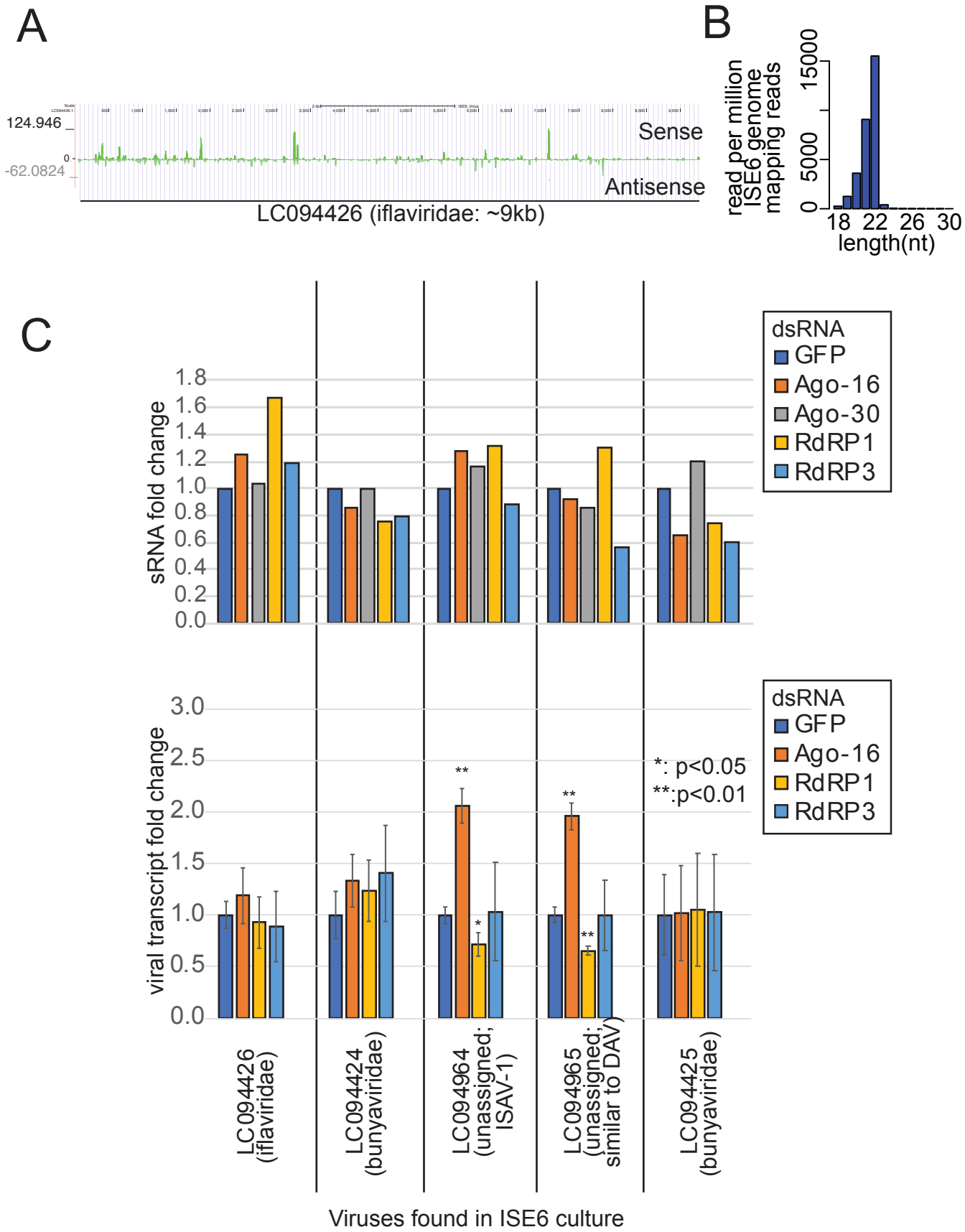

A
124.946
0
-62.0824
Sense
Antisense
LC094426 (iflaviridae: ~9kb)
B
read per million ISE6 genome mapping reads
0
5000
15000
18
22
26
30
length(nt)
C
sRNA fold change
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
dsRNA
GFP
Ago-16
Ago-30
RdRP1
RdRP3
viral transcript fold change
3.0
2.5
2.0
1.5
1.0
0.5
0.0
**
*
**
**
*: p<0.05
**:p<0.01
dsRNA
GFP
Ago-16
RdRP1
RdRP3
LC094426 (iflaviridae)
LC094424 (bunyaviridae)
LC094964 (unassigned; ISAV-1)
LC094965 (unassigned; similar to DAV)
LC094425 (bunyaviridae)
Viruses found in ISE6 culture

Figure 5

Figure 5. sRNAs from persistently infecting viruses

(A) sRNA reads from non-treated ISE6 cells were mapped to the 5 viral genomes that were known to be present in ISE6 cultures. As a representative virus, LC094426 (iflaviridae) is shown. Read density was calculated at each base and normalized for the number of reads mapping to the ISE6 genome.

(B) Size distribution of viral sRNAs. Viral sRNAs of each length were counted and normalized to the number of reads mapping to the ISE6 genome (reads per million genome mapping reads). The result of the non-treated control cells is shown.

(C) Changes in the sRNA (Upper, n=1) and the viral transcript (Lower, n=3) abundances in ISE6 cells after knocking down AGO or RdRP. sRNA libraries and total RNAseq libraries were analyzed to quantify the abundances of viral sRNAs and viral transcripts in the indicated libraries. Fold change values were calculated based on reads per million genome mapping reads (for small RNAs) and TPM (for viral transcripts). The averages and the standard deviations are shown.

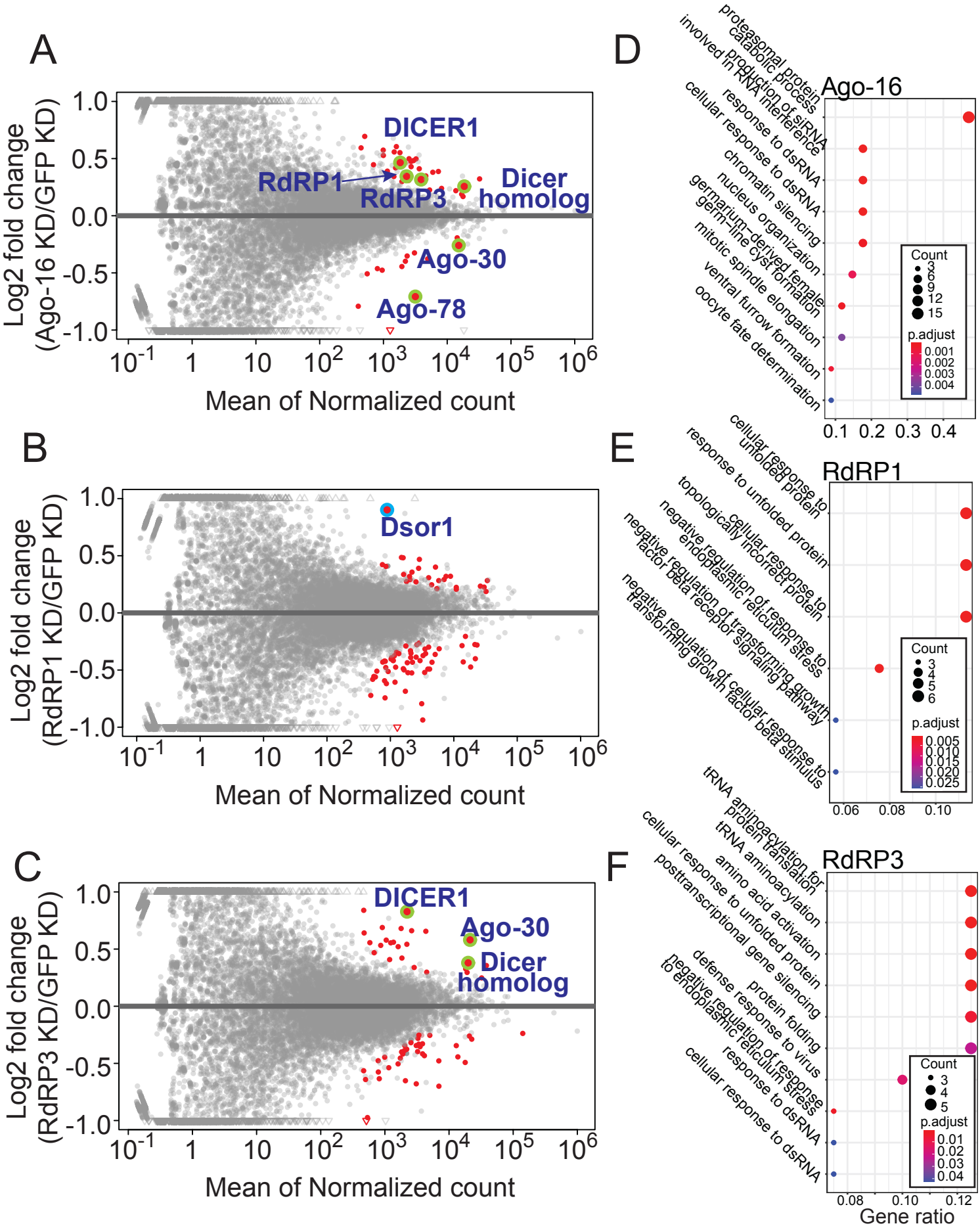

A
Log2 fold change (Ago-16 KD/GFP KD)
DICER1
RdRP1
RdRP3
Dicer homolog
Ago-30
Ago-78
Mean of Normalized count
B
Log2 fold change (RdRP1 KD/GFP KD)
Dsor1
Mean of Normalized count
C
Log2 fold change (RdRP3 KD/GFP KD)
DICER1
Ago-30
Dicer homolog
Mean of Normalized count
D
Ago-16
proteasomal protein catabolic process
production of siRNA involved in RNA interference
response to dsRNA
cellular response to dsRNA
chromatin silencing
nucleus organization
germarium-derived female germ-line cyst formation
mitotic spindle elongation
ventral furrow formation
oocyte fate determination
Count
p.adjust
E
RdRP1
cellular response to unfolded protein
response to unfolded protein
cellular response to topologically incorrect protein
negative regulation of response to endoplasmic reticulum stress
negative regulation of transforming growth factor beta receptor signaling pathway
negative regulation of cellular response to transforming growth factor beta stimulus
Count
p.adjust
F
RdRP3
tRNA aminoacylation for protein translation
tRNA aminoacylation
cellular response to amino acid activation
posttranscriptional gene silencing
protein folding
defense response to virus
negative regulation of response to endoplasmic reticulum stress
cellular response to dsRNA
Count
p.adjust
Gene ratio

Figure 6

Figure 6. Roles for the RdRP-dependent pathways in gene regulation

(A-C) MA-plots of pairwise comparisons between the control (dsGFP, n=3) and the knockdown sample (n=3). Red dots indicate genes differentially expressed in the knockdown samples (adjusted $p < 0.05$) and RNAi-related gene names are indicated if they are differentially expressed and highlighted by green circles. The point representing Dsor1 was highlighted by a blue circle.

(D-F) GO-term analysis of misregulated genes revealed in the above analysis. Results of biological process analysis are shown. The size and color represent the number of genes in the GO category and the significance of enrichment, respectively, as indicated in the legend. Results of molecular function and cellular component groups analyses are shown in Supplementary PDF.

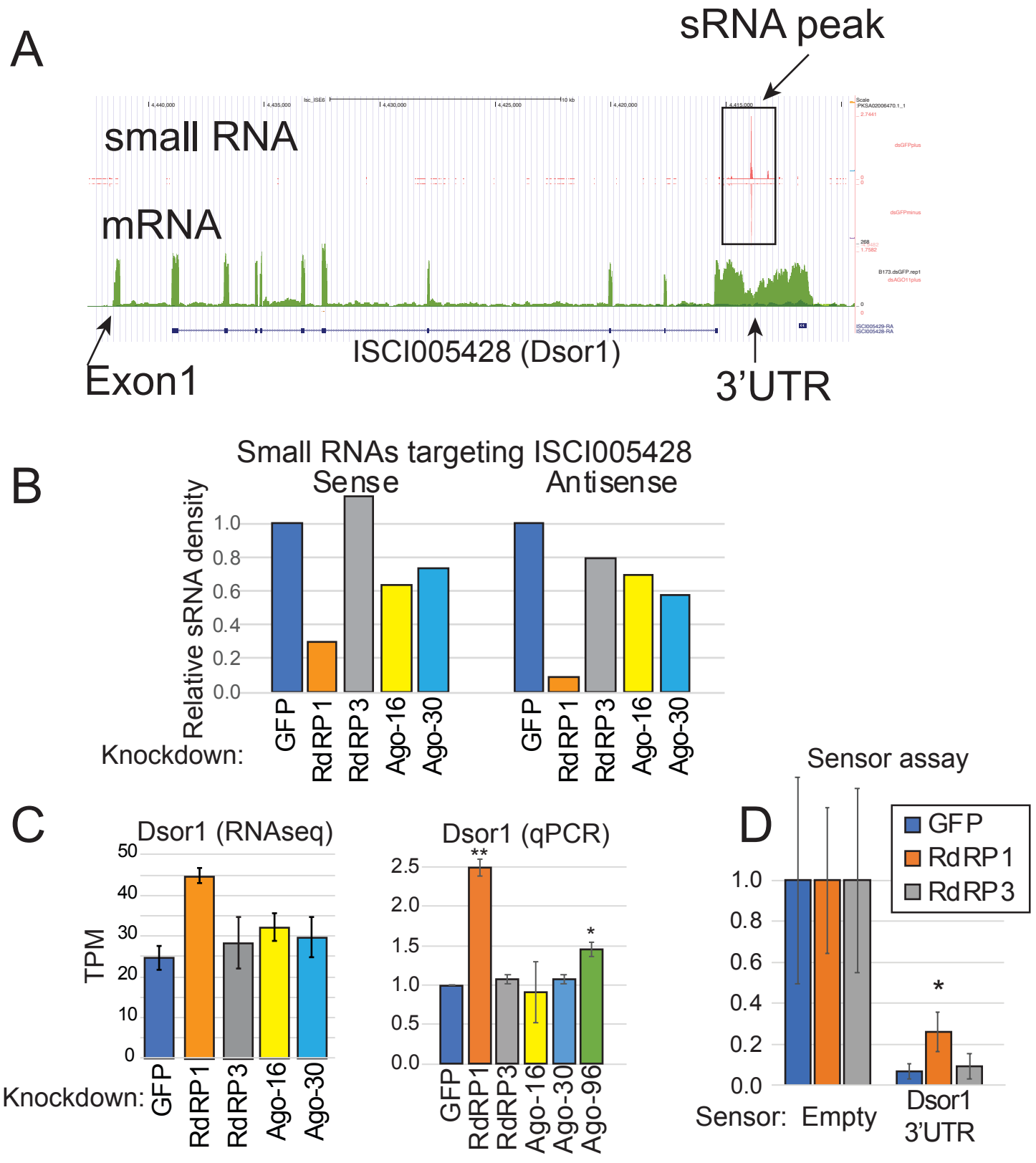

A
sRNA peak
small RNA
mRNA
ISCI005428 (Dsor1)
Exon1
3'UTR
B
Small RNAs targeting ISCI005428
Sense
Antisense
Relative sRNA density
1.0
0.8
0.6
0.4
0.2
0.0
Knockdown:
GFP
RdRP1
RdRP3
Ago-16
Ago-30
C
Dsor1 (RNAseq)
TPM
50
40
30
20
10
0
Knockdown:
GFP
RdRP1
RdRP3
Ago-16
Ago-30
Dsor1 (qPCR)
2.5
2.0
1.5
1.0
0.5
0.0
**
*
GFP
RdRP1
RdRP3
Ago-16
Ago-30
Ago-96
D
Sensor assay
GFP
RdRP1
RdRP3
1.0
0.8
0.6
0.4
0.2
0.0
*
Sensor:
Empty
Dsor1 3'UTR

Figure 7

Figure 7. Dsor1 is a target of the RdRP1-dependent sRNA pathway

(A) The gene structure of the RdRP1-regulated gene Dsor1 (ISCI005428). The original annotation in the Vectorbase (ISCI005428) lacks the UTRs but reads corresponding to the extending transcript beyond the coding region are visible on the RNAseq mapping data. Within its 3'UTR, there is a high peak of small RNAs ('sRNA peak").

(B) Small RNA reads mapping to the sense or antisense strands of Dsor1 locus was quantified. Dramatic reduction of small RNAs mapping to both sense and antisense strands was seen upon RdRP1 knockdown.

(C) Expression levels of Dsor1 mRNA as analyzed by the salmon pipeline. The averages and standard deviations of TPM values are shown (n=3). The results were verified by qPCR (Right panel, n=4).

(D) The Dsor1 3'UTR was cloned in the pmirGLO/Fer-Luc2/Act-hRluc vector (Hernandez et al. 2019), and ISE6 cells that were soaked with the indicated dsRNA were transfected with the sensor plasmid. The ratio between firefly luc and Rluc was normalized to that of empty sensor, and the means and standard deviations are shown in the chart (n=7). The experiments were performed twice on different days and the results from the two experiments were combined. Student's t-test was used to calculate p-values.

RdRP

Which paper identified

Active site

```
Ncr-Qde1        AHFPSDIQRVRAVFKPELHS--LKD----VIIFSTKGDVPLAKKLSGGDYDGDMAWVCWD 1018
Ath-RDR5        GLHFGDIHILKATYVKALEE--YVGNSKFAVFFPQKGPRSLGDEIAGGDFDGDMYFISRN 693
Ath-RDR3        GLHFGDIHVLKATYVKALED--YVGNAKFAVFFPQKGPRSLGDEIAGGDFDGDMYFISRN 704
Ath-RDR4        GLHFGDIHILKATYVKSLEQ--YVGNSKYGVFFPQKGPRSLGDEIAGGDFDGDMYFISRN 693
Spo-Rdp1        SLHPGDVRVCKAVRCDELMH--LKN----VIVFPTTGDRSIPAMCSGGDLDGDEYTVIWD 910
Cel-Ego1        CIVAGDVRIFEAVDIPELHH--MCD----VVVFPQHGPRPHPDEMAGSDLDGDEYSIIWD 985
Cel-Rrf1        CIVPGDVRIFEAVDIPELHH--MCD----VVVFPQHGPRPHPDEMAGSDLDGDEYSVIWD 958
Cel-Rrf2        GIVPGDVRIFEAVDIPELHH--LCD----VVVFPQHGPRPHPDEMAGSDLDGDEYSVIWD 939
Isc-RdRP3       CIHPGDIRLLEAVDVPALSH--IRD----CVVFPQIGKRPHPDEMAGSDLDGDEYSVIWY 937
Ath-RDR6        CLHPGDVRILEAVDVPQLHH--MYD----CLIFPQKGDRPHTNEASGSDLDGDLYFVAWD 874
Ath-RDR1        CLHPGDVRVLQAVNVPALNH--MVD----CVVFPQKGLRPHPNECSGSDLDGDIYFVCWD 808
Ath-RDR2        CLHPGDIRVLDAIYEVHFEEKGYLD----CIIFPQKGERPHPNECSGGDLDGDQFFVSWD 841
Cel-Rrf3        CHVPGDVRVFDAVWQPALAH--LVD----VVVFPQHGPRPHPDEMAGSDLDGDEYSIIWD 1145
Isc-RdRP4       CLHPGDVRKFTAIDEPRLHH--VVD----CIVFPGQGHRPHPDEMAGSDLDGDEYIVIWE 874
Isc-RdRP1       CLHPGDVRKFTAVDVPALHH--VKD----CVVFPARGPRPHPNEMAGSDLDGDEYVVIWK 880
                    .*:.   *     :      .     :.*.  *  .     :*.* ***   : .
```

•

•

•

•

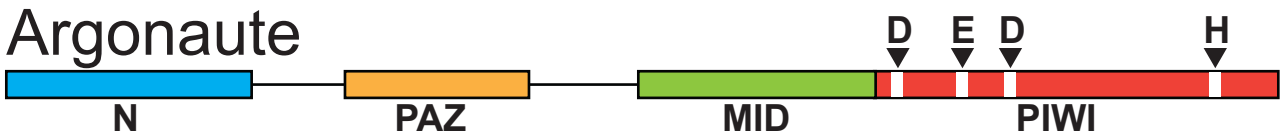
Putative loop

```
Ncr-Qde1        NLVDQSKQGIVFNEASWA------------QLRRELLGGALSLP--DPMYKSDSWLGRGE 1158
Ath-RDR5        DALDAPKKGDKVY-LPNK-------------LKPDIFPHYMERD---KKFQSTSILGL-- 829
Ath-RDR3        DALDAPKKGAKVD-LPPD-------------LEIKNFPHYMERD-PKRDFRSTSILGL-- 842
Ath-RDR4        DAIDAPKTGTEVN-LPLD-------------VKVDLFPHYMERN---KTFKSTSILGL-- 829
Spo-Rdp1        KAVDFAKSGVACK-MQAK-------------YHPKRYPDFMQKT-KTRSFRSETAVGK-- 1034
Cel-Ego1        QAVDFTKSGQPPDELERKWRKDEETGEMIPPERAERVPDYHMGNDHTPMYVSPRLCGK-- 1117
Cel-Rrf1        QAVDFTKSGQPPDPLETKWRADPVTFEVIPPENPERIPDFHMGNERSPMYVSPRLCGK-- 1090
Cel-Rrf2        QAVDFSKSGKPPDELQTTWKTDDATGEMIPPERAERVPDYHVGSDHMPKYVSPRLCGK-- 1171
Isc-RdRP3       ECVDFAKTGSSTR-LAQS-------------EKPPRYPDFMEKHAEKYTYPSKRALGL-- 1057
Ath-RDR6        TAVDFPKTGKIVS-MPFH-------------LKPKLYPDFMGKE-DYQTYKSNKILGR-- 994
Ath-RDR1        TAVDFPKTGVAAV-IPQH-------------LYVKEYPDFMEKP-DKPTYESKNVIGK-- 926
Ath-RDR2        RAVDFAKTGAPAE-MPYA-------------LKPREFPDFLERF-EKPTYISESVFGK-- 959
Cel-Rrf3        VAVDFPKSGVPAEPLSSF-------------EQCEMTPDYMMSG-GKPMYYSTRLNGQ-- 1262
Isc-RdRP4       ICLDFAKSGRTAY-LRRE-------------ERPMFYPDFMEKGSHKISYRSDRVLGL-- 994
Isc-RdRP1       TCLDFAKTGETSY-LNKD-------------EKPIQYPDFMEKGSSKDTYRSKRVLGH-- 1000
                  :* .* *                                          : *    *
```

Figure S1

Figure S1. Alignment of RdRP genes.

The sequences of RdRP homologs from Fission yeast (Spo), Neurospora (Ncr), Arabidopsis (Ath), worm (Cel) and tick (Isc) were aligned by MUSCLE <https://www.ebi.ac.uk/Tools/msa/muscle/> and the regions near the catalytically active site and the Rrf1-specific putative loop are shown. The accession numbers of the gene sequences can be found in Table S1.




```
dAgo2        -KNTMYIGADVTHPSPDQREIPSVVGVAASHDPYGASYNMQYRLQRG------ALEEIED      1009
hAgo2        QQPVIFLGADVTHPPAGDGKKPSIAAVVGSMDAHPNRYCATVRVQQH------RQEIIQD      641
dAgo1        NEPVIFLGADVTHPPAGDNKKPSIAAVVGSMDAHPSRYAATVRVQQH------RQEIIQE      765
IscAgo-78    NEPVIFLGADVTHPPAGDNKKPSIAAVVGSMDAHPSRYAATVRVQQH------RQEIIQD      683
IscAgo-16    RKPVIIIGADVSHPAPGDRIRPSIAACVGSLDSIPSKYRASIRVQLEDQEAVARVEMIKD      772
IscAgo-30    EEPVIILGADVNHPAPGRSNHPSYAALVGSLDSCPSKYHASVRIQRTSANS-NEREIIKD      772
IscAgo-96    SKPVIVMGADVTHPGAKEFNRPSIAAVVASTDRFAFRYITAFRIQKQNMEVKARVEIIED      731
IscAub       LTKMMCVGYDTYHDSR--QKGLSAGGFVASLNRTLTRWYSRVSFHQT------HQELGSA      698
IscAGO3      MHNVMVIGIDVYHDIT--RGRQSVMGFVASMNQSLTRWFSKCAFQEP------GKELVNC      754
                 : :* *. *         *  . ..* :     :     .:          *  .

dAgo2        MFSITLEHLRVYKEYR-NAYPDHIIYYRDGVSDGQFPKIKNEELRCIKQACDKVG----C      1064
hAgo2        LAAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQQVLHHELLAIREACIKLE--KDY      698
dAgo1        LSSMVRELLIMFYKSTGGYKPHRIILYRDGVSEGQFPHVLQHELTAIREACIKLE--PEY      823
IscAgo-78    LASMVKELLIQFYKST-RFKPNRIIFYRDGVSEGQFQQVLHHELLAVREACMKLE--ADY      740
IscAgo-16    LSGMVIELLKAFRE-ATRHKPEHIIFYRDGVSEGQFAEVRDLELQAIRDACLSLQPDGSF      831
IscAgo-30    LKGMVKEALRAYYIKT-HQKPRKIIFYRDGVSEGQFAEVLNHELPALRQACKELE--DGY      829
IscAgo-96    MKSIAKELLLGFYIANNQVRPDKILFYRDGVSEGQFRQVLDHELAAIRAACMELE--AGY      789
IscAub       LKTHMALSMKQYQEENDGAVPERLLFFRDGVSDGQLLQVQEWEVGQIQSLLTEMF--PGR      756
IscAGO3      IKIAMLEAIVKYYEVN-HKHPDRIFVFRDGVGDGQLSYVSDYEIEQLVQSFVNVS--PDY      811
             :       :  :        * ::: :*****.:**:  : . *:  :     .:

dAgo2        KPKICCVIVVKRHHTRFFPSGDVTTSNKFNNVDPGTVVDRTIVHPNEMQFFMVSHQAIQG      1124
hAgo2        QPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVDTKITHPTEFDFYLCSHAGIQG      758
dAgo1        RPGITFIVVQKRHHTRLFCAEKKEQSGKSGNIPAGTTVDVGITHPTEFDFYLCSHQGIQG      883
IscAgo-78    KPGITFVVVQKRHHTRLFCSDKKEQIGKSGNIPAGTTVDLGITHPTEFDFYLCSHAGIQG      800
IscAgo-16    KPPVTFIVVQKRHHTRFMPTNDRDGVGKARNVPPGTTVDTVVTHPVDFDFFLCSHYGIQG      891
IscAgo-30    TPAIVFILVQKRHSTRFMPKYQQDGVGRFNNVPPGTTVDRIVTHPKDFDFFLCSHAGIQG      889
IscAgo-96    EPGITFLTVQKRHHTRFMPENRRDGCGKSGNIPPGTTIDTTVTHPVDFDFFLCSHFGIQG      849
IscAub       EPQLAFIVVTKRIAARFFGTG----RGAFQNPLPGTVIDTNVTRPERYDFYLVSQSVRQG      812
IscAGO3      KPSIAVVVVQKRINTRIFARL--NGGRELDNPTPGTVVDHEVTRRDWCDFFLVSQKVRQG      869
              * :  : * **  :*::            *   **.:*  :.:    :*:: *:   **

dAgo2        TAKPTRYNVIENTGNLDIDLLQQLTYNLCHMFPRCNRSVSYPAPAYLAHLVAARGRVYLT      1184
hAgo2        TSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLV      818
dAgo1        TSRPSHYHVLWDDNHFDSDELQCLTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLV      943
IscAgo-78    TSRPSHYHVLWDDNQFSADELQCLTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLV      860
IscAgo-16    TSKPAHYYVVHDDYNFSSDDLQKLSYYLCHTYARCARSVSIPAPVYYAHLAAFRAKEHIF      951
IscAgo-30    TSRPTHYYVLHDDVGFQADELQSLTFYLCHTYARCPRSVSIPAPAYYAHWVAFRANQHAV      949
IscAgo-96    TSRPAHYYVLWDDNEFTADALQKLTYGLCHTYARCARSVSIPVPVYYAHHATQRAKCYVD      909
IscAub       TVAPTHFNVIHDTTTLKPEHMQRLSYKLTHLYFNWPGTIRVPAPCQYAHKLAFLAGQSLH      872
IscAGO3      TVSPTHYIVVRNTTELSPDQMQRLAYKLTHLYYNWPGTIRVPAPCQVGPAPGMSISRKVT      929
             *  *::: *: :   :  : :* *:: * * : .   ::  *.*   .
```

Figure S2

Figure S2. Alignment of AGO genes.

The sequences of AGO homologs from *Drosophila* (dAGOs), human (hAGOs) and tick (IscAGOs) were aligned by CLUSTAL Omega < https://www.ebi.ac.uk/Tools/msa/clustalo/> and the region containing the catalytic residues is shown. The conserved DEDH residues are highlighted in red. The accession numbers of the gene sequences can be found in Table S1.

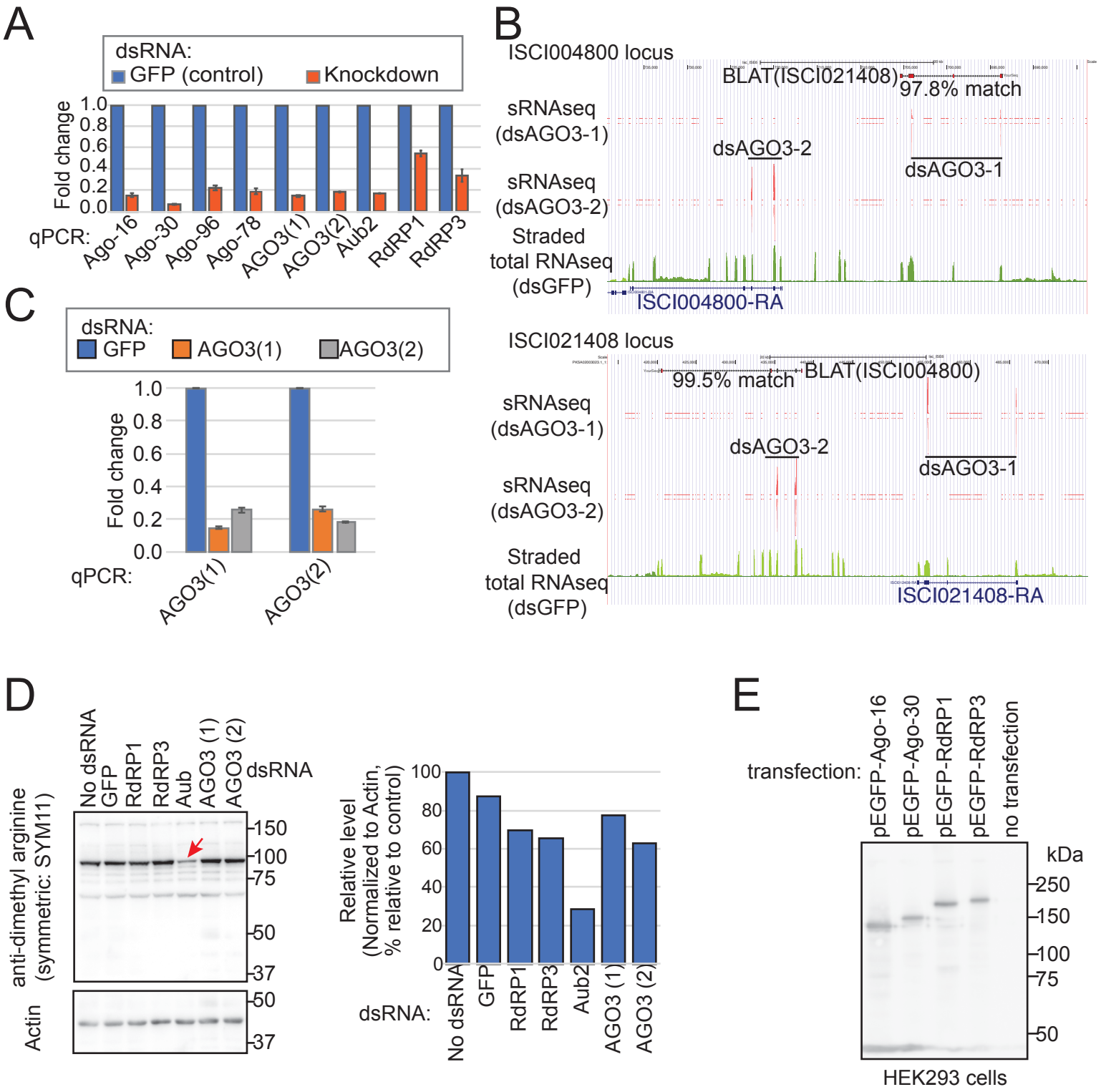

A
dsRNA:
GFP (control)
Knockdown
Fold change
1.0
0.8
0.6
0.4
0.2
0.0
qPCR:
Ago-16
Ago-30
Ago-96
Ago-78
AGO3(1)
AGO3(2)
Aub2
RdRP1
RdRP3
B
ISCI004800 locus
BLAT(ISCI021408)
97.8% match
sRNAseq (dsAGO3-1)
dsAGO3-2
dsAGO3-1
sRNAseq (dsAGO3-2)
Straded total RNAseq (dsGFP)
ISCI004800-RA
ISCI021408 locus
99.5% match
BLAT(ISCI004800)
sRNAseq (dsAGO3-1)
dsAGO3-2
dsAGO3-1
sRNAseq (dsAGO3-2)
Straded total RNAseq (dsGFP)
ISCI021408-RA
C
dsRNA:
GFP
AGO3(1)
AGO3(2)
Fold change
1.0
0.8
0.6
0.4
0.2
0.0
qPCR:
AGO3(1)
AGO3(2)
D
anti-dimethyl arginine (symmetric: SYM11)
No dsRNA
GFP
RdRP1
RdRP3
Aub
AGO3 (1)
AGO3 (2)
dsRNA
150
100
75
50
37
Actin
50
37
Relative level (Normalized to Actin, % relative to control)
100
80
60
40
20
0
dsRNA:
No dsRNA
GFP
RdRP1
RdRP3
Aub2
AGO3 (1)
AGO3 (2)
E
transfection:
pEGFP-Ago-16
pEGFP-Ago-30
pEGFP-RdRP1
pEGFP-RdRP3
no transfection
kDa
250
150
100
75
50
HEK293 cells

Figure S3

Figure S3. Characterization of AGO/RdRP proteins

(A) Verification of gene knockdown. RNA samples isolated from ISE6 cells transfected with the indicated dsRNAs were used for qPCR verification. The qPCR primers were designed to quantify the expression of the cognate gene but avoided the region corresponding to the introduced dsRNA sequence. The values were normalized by the Actin expression and further normalized to the value in the dsGFP control sample. Four technical replicates were made and the averages and the standard deviations are shown.

(B) ISCI021408 and ISCI004800 correspond to two fragments of the IscAGO3 gene. There are two loci in the ISE6 assembly that match with ISCI021408 and ISCI004800 genes at high homologies (>97.8%) and the gene structures were conserved in the two loci, suggesting that these two loci may represent a very recent gene duplication or a genome assembly artifact.

(C) qPCR verification of AGO3 knockdown against ISCI021408 (dsAGO3-1) and ISCI004800 (dsAGO3-2). Note that transfection of either dsRNA construct resulted in strong reduction of both ISCI021408 and ISCI004800 qPCR amplicons, indicating the dsRNA constructs reduce both genes. Based on the observations in (B), we consider the two loci as a single gene, IscAGO3.

(D) Western blot analysis for the PIWI proteins detected by anti-symmetric dimethyl arginine antibody, SYM11. ISE6 cells were treated with the indicated dsRNAs and total protein was extracted and separated on an SDS-PAGE gel. The proteins were detected by SYM11, which recognizes symmetric dimethyl arginine that is characteristic of PIWI proteins and some other RNA binding proteins (Kirino et al. 2009). The quantified values of the band at ~100kDa were normalized to the value of the Actin band and shown in the bar chart next to the Western blotting panel. Note the strong reduction of the 100kDa band upon Aub KD, suggesting that this band corresponded to the Aub protein.

(E) HEK293T cells were transfected with the indicated construct and total protein samples were subjected to Western blot analysis using the GFP antibody. The bands were detected at the predicted sizes.

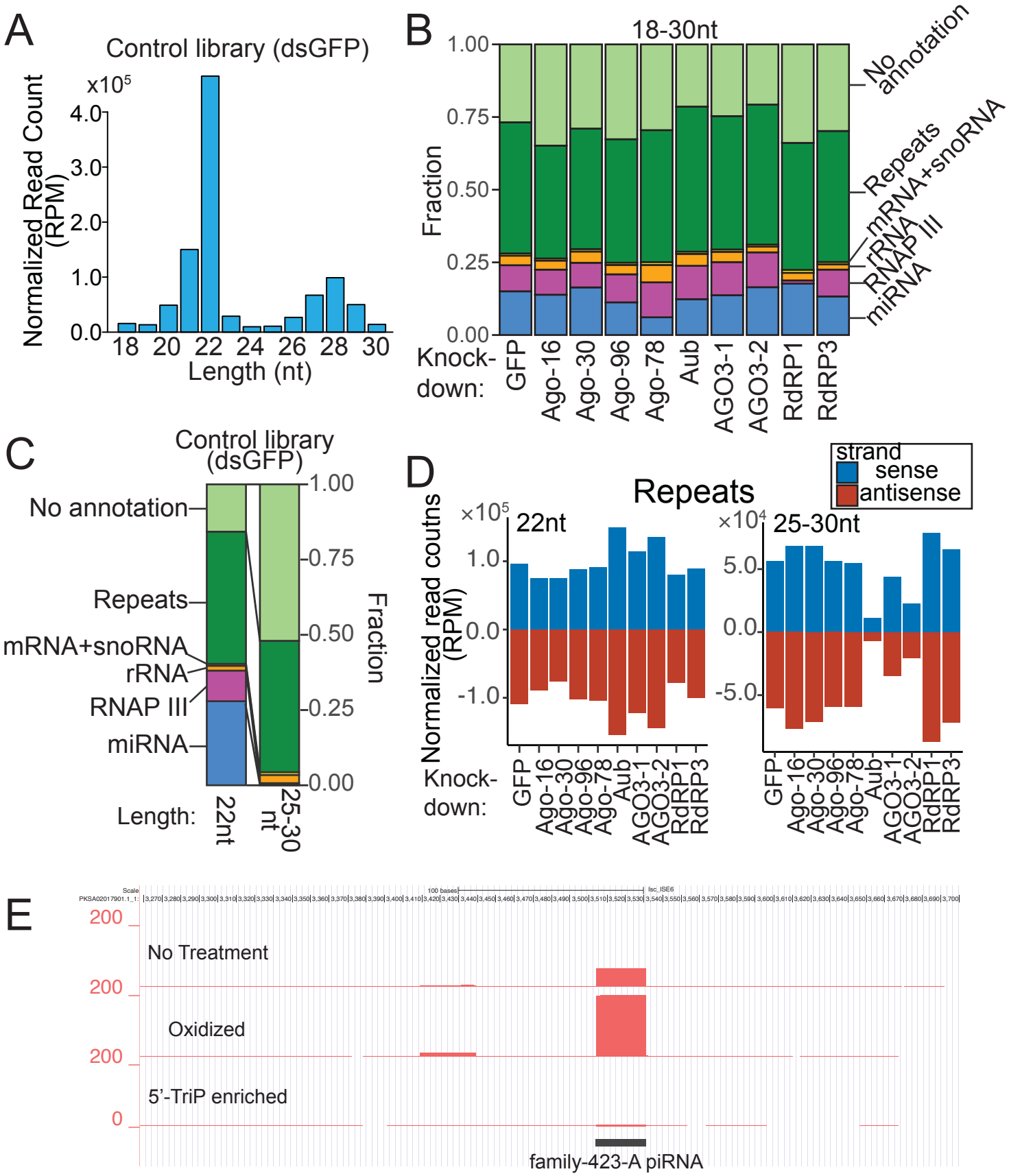

A
Control library (dsGFP)
x10^5
Normalized Read Count (RPM)
4.0
3.0
2.0
1.0
0.0
18 20 22 24 26 28 30
Length (nt)
B
18-30nt
Fraction
1.00
0.75
0.50
0.25
0.00
No annotation
Repeats
mRNA+snoRNA
rRNA
RNAP III
miRNA
Knock-down:
GFP
Ago-16
Ago-30
Ago-96
Ago-78
Aub
AGO3-1
AGO3-2
RdRP1
RdRP3
C
Control library (dsGFP)
No annotation
Repeats
mRNA+snoRNA
rRNA
RNAP III
miRNA
1.00
0.75
0.50
0.25
0.00
Fraction
Length:
22nt
25-30 nt
D
Repeats
strand
sense
antisense
22nt
25-30nt
×10^5
×10^4
Normalized read coutns (RPM)
1.0
0.0
-1.0
5.0
0.0
-5.0
Knock-down:
GFP
Ago-16
Ago-30
Ago-96
Ago-78
Aub
AGO3-1
AGO3-2
RdRP1
RdRP3
E
No Treatment
Oxidized
5'-TriP enriched
200
200
200
0
family-423-A piRNA

Figure S4

Figure S4. Compositions of sRNA libraries

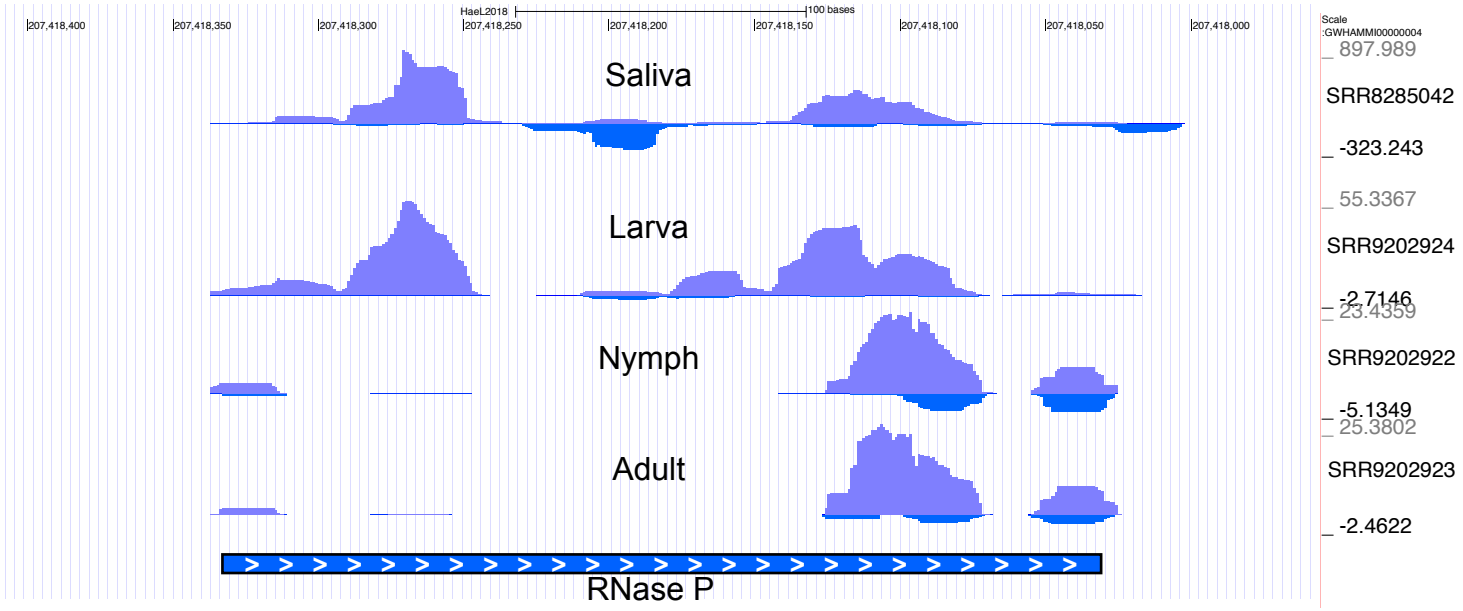
(A) Size distribution of sRNA reads. Genome-matching sRNA reads of each length in the control (dsGFP) library were counted and normalized by the number of reads of all lengths and expressed as RPM (Reads Per Million mapped reads). (B) Percent stacked-bar-chart showing the composition of sRNAs in the knockdown libraries. The fractions of sRNAs derived from miRNAs (blue), RNAP III transcribed genes (pink), rRNAs (orange), snoRNAs + mRNAs (yellow), repeats (green) and sRNAs that did not belong to any of the categories (light green) are shown. 18-30nt reads were used. Note that the same bars are shown in Figure 2A for dsGFP, dsAgo-78, dsRdRP1 and dsRdRP3.

(C) Percent stacked-bar-chart showing the composition of sRNAs in the control (dsGFP) library. 22nt (left bar) and 25-30nt (right bar) reads were used.

(D) Normalized counts (RPM) of sense and antisense 25-30nt reads that were derived from repeats in the indicated knockdown libraries are shown in blue and red, respectively. Negative values were given to antisense read counts.

(E) UCSC Genome Browser screenshot of the TE family-423 region. A strong peak of a piRNA (family-423-A) was observed in the control library, which was prepared with the standard protocol. This peak was even higher in the library prepared after oxidization of the total RNA, supporting that this RNA species have a 2'-O-me group similar to piRNAs in other organisms. These species were depleted when 5'-mono-phosphorylated species were removed, suggesting that this RNA species had a 5'-mono-phosphate group.

A


HaeL2018
100 bases
207,418,400
207,418,350
207,418,300
207,418,250
207,418,200
207,418,150
207,418,100
207,418,050
207,418,000
Scale
:GWHAMMI00000004
Saliva
897.989
SRR8285042
-323.243
Larva
55.3367
SRR9202924
-2.7146
Nymph
23.4359
SRR9202922
-5.1349
Adult
25.3802
SRR9202923
-2.4622
RNase P

B

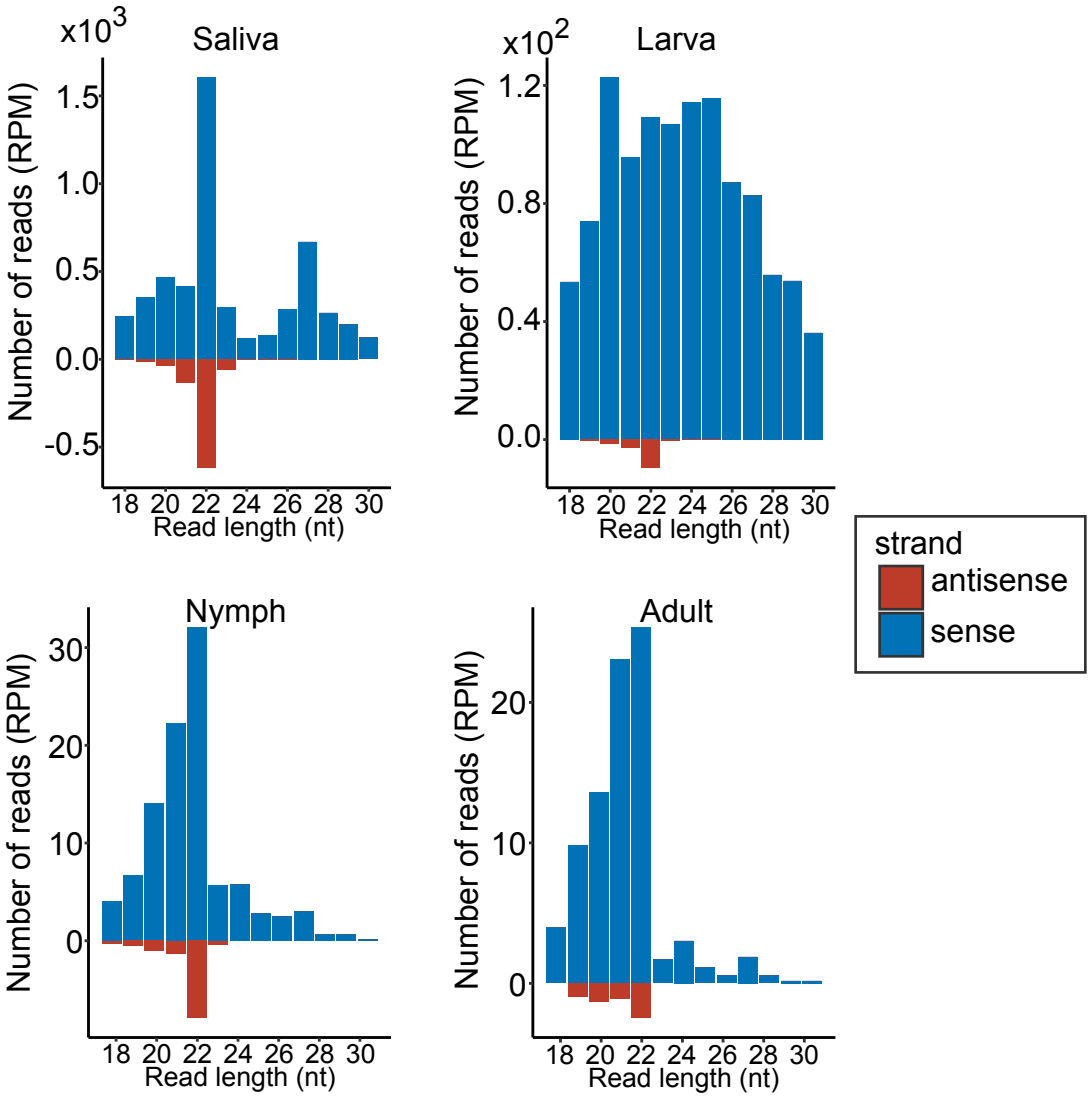

Saliva
x10^3
Number of reads (RPM)
Read length (nt)
Larva
x10^2
Nymph
Adult
strand
antisense
sense

Figure S5

Figure S5. Conservation of 22nt-sRNA production from RNAP III transcribed genes in *H. longicornis*

(A) UCSC genome browser screenshot of the RNase P locus in the *H. longicornis* genome. sRNA mapping density is shown. The upper (purple) and lower (blue) signals represent sense and antisense reads. The data are from previous studies (Nawaz et al. 2020; Liu et al. 2020).

(B) Size distribution of sRNAs mapping to the representative RNAP III-dependent genes (RNase P, RNase MRP and SRP RNA) of *H. longicornis*.

For both (A) and (B), the saliva data are also shown in Figure 2E-F.

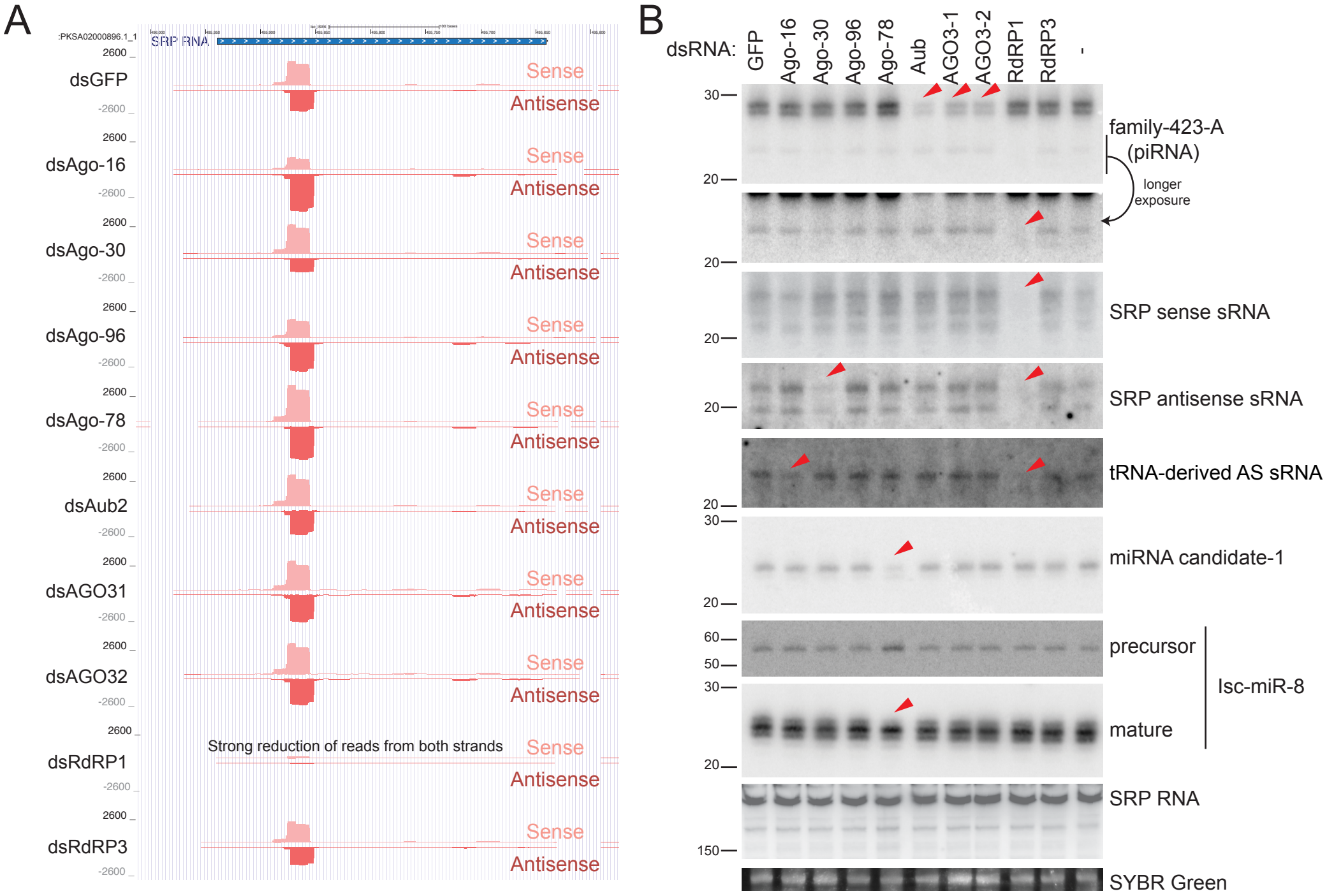
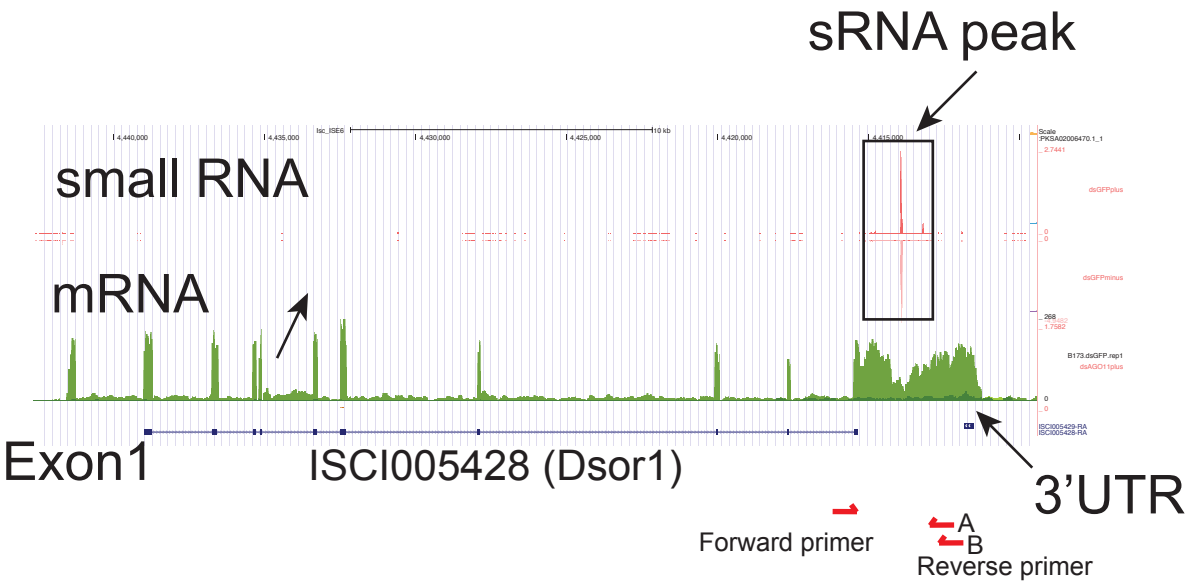

A
:PKSA02000896.1_1
SRP RNA
2600
-2600
dsGFP
dsAgo-16
dsAgo-30
dsAgo-96
dsAgo-78
dsAub2
dsAGO31
dsAGO32
dsRdRP1
dsRdRP3
Sense
Antisense
Strong reduction of reads from both strands
B
dsRNA:
GFP
Ago-16
Ago-30
Ago-96
Ago-78
Aub
AGO3-1
AGO3-2
RdRP1
RdRP3
-
30
20
60
50
150
family-423-A
(piRNA)
longer
exposure
SRP sense sRNA
SRP antisense sRNA
tRNA-derived AS sRNA
miRNA candidate-1
precursor
mature
Isc-miR-8
SRP RNA
SYBR Green

Figure S6

Figure S6. Biogenesis of small RNAs from RNAP III-transcribed genes

(A) UCSC screenshot at the SRP RNA locus shown as an example of a locus producing RdRP1-dependent sRNAs. The tracks show sRNA mapping densities in the indicated libraries from the sense (upward) and antisense (downward) strands of the locus with respect to the direction of the transcription.

(B) Verification of sRNA library analysis by Northern blotting of representative small RNA species. 10ug total RNA from each of the indicated knockdown samples was loaded, and the membrane was incubated with the indicated probes. The probe sequences can be found in Table S5. The >150nt (SRP RNA), ~50-60nt (pre-mir-8) and ~20-30nt (other panels) areas of the blots are cropped and shown. miRNA-candidate-1 was not in miRBase but identified by visual inspection of mapping data as well as RNA secondary structure prediction for its precursor structure. The full analysis for novel miRNA genes will be published elsewhere. The SYBR Green II staining result is shown as the loading control. Red arrowheads show the bands that showed a clear reduction in the knockdown lanes.


sRNA peak
small RNA
mRNA
Exon1
ISCI005428 (Dsor1)
3'UTR
Forward primer
A
B
Reverse primer

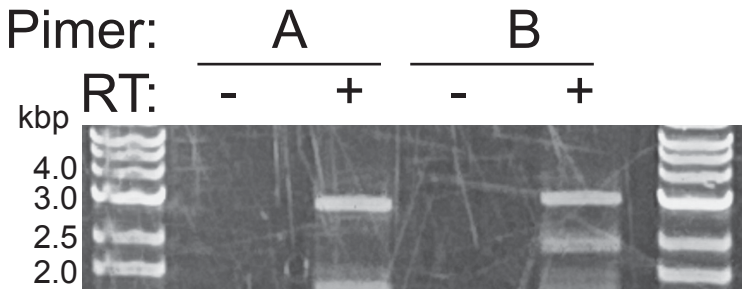

Pimer:
A
B
RT:
-
+
-
+
kbp
4.0
3.0
2.5
2.0

Figure S7

Figure S7. RT-PCR verification of Dsor1 3'UTR

(A) UCSC genome browser screenshot of the Dsor1 (ISCI005428) locus in the *I.scapularis* genome. sRNAseq read density (upper) and stranded total RNAseq read density from the control libraries are shown. The positions of RT-PCR primers used in (B) are shown. The forward primer was designed against the 3'-end of the coding region, and the reverse primers were designed against the putative 3'UTR sequence.

(B) RT-PCR results. Using the primers illustrated in (A), RT-PCR was performed with (+) or without (-) adding reverse transcriptase in the RT reaction. Both primer pairs amplified the expected fragments in a reverse-transcription dependent manner, confirming the extended sequence represents the 3'UTR of Dsor1. The band amplified with pair B was cloned and sequenced, and used for the luciferase assays shown in Figure 7D.