



A low-abundance class of Dicer-dependent siRNAs produced from a variety of features in *C. elegans*

Thiago L. Knittel, Brooke E. Montgomery, Alex J. Tate, et al.

Genome Res. 2024 34: 2203-2216 originally published online December 2, 2024

Access the most recent version at doi:[10.1101/gr.279083.124](https://doi.org/10.1101/gr.279083.124)

References This article cites 106 articles, 36 of which can be accessed free at:
<http://genome.cshlp.org/content/34/12/2203.full.html#ref-list-1>

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner with a teal background. On the left, the text reads 'CRISPR and RNAi Genetic Screening. Your new superpower.' In the center, there is a white-bordered box containing the words 'LEARN MORE'. On the right, there is a photograph of a woman wearing a red superhero mask and a red cape over a white shirt. To the right of the photo is the Cellecta logo, which consists of a green molecular structure and the word 'CELLECTA' in white capital letters.

CRISPR and RNAi Genetic Screening.
Your new superpower.

LEARN MORE

CELLECTA

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Research

A low-abundance class of Dicer-dependent siRNAs produced from a variety of features in *C. elegans*

Thiago L. Knittel,^{1,4} Brooke E. Montgomery,^{1,4} Alex J. Tate,¹ Ennis W. Deihl,² Anastasia S. Nawrocki,² Frederic J. Hoerndli,^{2,3} and Taiowa A. Montgomery^{1,2}

¹Department of Biology, ²Department of Biomedical Sciences, Colorado State University, Fort Collins, Colorado 80523, USA;

³Cell and Molecular Biology Program, Colorado State University, Fort Collins, Colorado 80523, USA

Canonical small interfering RNAs (siRNAs) are processed from double-stranded RNA (dsRNA) by Dicer and associate with Argonautes to direct RNA silencing. In *Caenorhabditis elegans*, 22G-RNAs and 26G-RNAs are often referred to as siRNAs but display distinct characteristics. For example, 22G-RNAs do not originate from dsRNA and do not depend on Dicer, whereas 26G-RNAs require Dicer but derive from an atypical RNA duplex and are produced exclusively antisense to their messenger RNA (mRNA) templates. To identify canonical siRNAs in *C. elegans*, we first characterized the siRNAs produced via the exogenous RNA interference (RNAi) pathway. During RNAi, dsRNA is processed into ~23 nt duplexes with ~2 nt, 3'-overhangs, ultimately yielding siRNAs devoid of 5'G-containing sequences that bind with high affinity to the Argonaute RDE-1, but also to the microRNA (miRNA) pathway Argonaute, ALG-1. Using these characteristics, we searched for their endogenous counterparts and identified thousands of endogenous loci representing dozens of unique elements that give rise to mostly low to moderate levels of siRNAs, called 23H-RNAs. These loci include repetitive elements, putative coding genes, pseudogenes, noncoding RNAs, and unannotated features, many of which adopt hairpin (hp) structures reminiscent of the hpRNA/RNAi pathway in flies and mice. RDE-1 competes with other Argonautes for binding to 23H-RNAs. When RDE-1 is depleted, these siRNAs are enriched in ALG-1 and ALG-2 complexes. Our results expand the known repertoire of *C. elegans* small RNAs and their Argonaute interactors, and demonstrate that key features of the endogenous siRNA pathway are relatively unchanged in animals.

[Supplemental material is available for this article.]

Small interfering RNAs (siRNAs) are ~21–24 nt long noncoding RNA produced through cleavage of double-stranded RNA (dsRNA) by the endoribonuclease Dicer (Hamilton and Baulcombe 1999; Hammond et al. 2000; Zamore et al. 2000; Bernstein et al. 2001; Elbashir et al. 2001; Grishok et al. 2001; Ketting et al. 2001; Knight and Bass 2001). Argonaute proteins associate with siRNA duplexes, discarding one strand and retaining the other, which then acts as a guide to direct silencing of target messenger RNAs (mRNAs) (Grishok et al. 2001; Hammond et al. 2001).

Shortly after the discovery that dsRNA triggers RNA silencing in *Caenorhabditis elegans*, siRNAs were identified in plants for their role in silencing viruses and transgenes, revealing the specificity factor of the RNA interference (RNAi) pathway (Fire et al. 1998; Hamilton and Baulcombe 1999). In plants, beyond their roles in silencing viruses and transgenes, siRNAs also target endogenous coding genes, intergenic regions, and transposons (Llave et al. 2002; Kasschau et al. 2007). Similarly, in mice and flies, endogenous siRNAs originate from various sources, most notably transposons, suggesting a common role for siRNAs in silencing foreign elements (Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Okamura et al. 2008a,b; Tam et al. 2008; Watanabe et al. 2008).

C. elegans contains multiple distinct classes of endogenous small RNAs but canonical siRNAs like those found in other species have not been extensively studied. What is often classified as

siRNAs in *C. elegans* lacks key features characteristic of siRNAs in other species (Claycomb 2014). For example, 22G-RNAs do not require Dicer and appear to lack a dsRNA intermediate (Aoki et al. 2007; Pak and Fire 2007; Gu et al. 2009; Claycomb 2014). 26G-RNAs are similar to canonical siRNAs in their requirement for Dicer, but these small RNAs derive from an asymmetric dsRNA intermediate structure that lacks the ~2 nt, 3'-overhangs characteristic of Dicer products (Ambros et al. 2003; Ruby et al. 2006; Han et al. 2009; Conine et al. 2010; Gent et al. 2010; Vasale et al. 2010; Welker et al. 2010; Fischer et al. 2011; Blumenfeld and Jose 2016). Both 22G-RNAs and 26G-RNAs are produced antisense to their mRNA templates through the activity of RNA-dependent RNA polymerases, which further distinguishes them from siRNAs in mice and flies, but is a common feature of most siRNAs in plants and fungi (Cogoni and Macino 1999; Dalmay et al. 2000; Mourrain et al. 2000; Smardon et al. 2000; Sijen et al. 2001; Volpe et al. 2002; Aoki et al. 2007; Pak and Fire 2007; Gent et al. 2009, 2010; Gu et al. 2009; Han et al. 2009; Conine et al. 2010; Vasale et al. 2010).

Both exogenous RNAi and viral infection trigger a silencing response involving the production of exogenous canonical siRNAs in *C. elegans* (Ketting et al. 2001; Ashe et al. 2013). Furthermore, studies exploring the interplay between RNAi pathways and RNA editing by adenosine deaminases (ADARs) in *C. elegans* have identified numerous endogenous small RNAs with the hallmarks of canonical siRNA. For example, they appear to be

⁴These authors contributed equally to this work.

Corresponding author: tai.montgomery@colostate.edu

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.279083.124>.

© 2024 Knittel et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.html>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

processed from a dsRNA intermediate and are 22–24 nt long (Wu et al. 2011; Reich et al. 2018; Fischer and Ruvkun 2020). These siRNAs are highly upregulated in animals with loss of function in one or both ADAR genes, presumably because in wild-type animals the base changes introduced by RNA editing disrupt the secondary structure of dsRNA, preventing its recognition or processing by Dicer (Wu et al. 2011; Reich et al. 2018; Fischer and Ruvkun 2020).

Endogenous siRNAs may constitute a distinct class of small RNAs largely overlooked in *C. elegans*. Presumably, such siRNAs are rare in wild-type *C. elegans* since they have not emerged from high-throughput small RNA sequencing (sRNA-seq) studies characterizing small RNAs (Ruby et al. 2006). Nonetheless, even at low abundance, they could have important roles in gene regulation and thus their identification would be an important step toward a more complete understanding of the small RNA landscape of *C. elegans*. In this study, our aim was to identify and characterize these siRNAs in the context of their Argonaute binding partners.

Results

Molecular attributes of exogenous siRNAs in *C. elegans*

In *C. elegans*, siRNAs are processed from exogenous dsRNA but accumulate at very low levels compared to the secondary 22G-RNAs, which are produced downstream by RNA-dependent RNA polymerases (Sijen et al. 2001; Pak and Fire 2007). Because 22G-RNAs are produced antisense to an mRNA template, sRNA-seq can be used to distinguish primary siRNAs from secondary 22G-RNAs by their alignment in the sense orientation to target mRNAs (Sijen et al. 2001; Pak and Fire 2007). Although antisense reads are necessarily lost with this approach, we and others have used it to characterize exogenous siRNAs as predominantly 21–24 nt with a preference for a 5'U, A, or C (Zhang et al. 2012; Svendsen et al. 2019). However, the presence of 22 nt, 5'G-containing reads aligning to the sense strand of the mRNA in these libraries suggests contamination from 22G-RNAs, complicating the distinction between primary siRNAs and 22G-RNAs. To further refine the molecular attributes of exogenous siRNAs, we analyzed sRNA-seq libraries from *glp-4* mutants treated with RNAi against the germline-expressed gene, *pos-1* (Svendsen et al. 2019). Because *glp-4* mutants lack a germline when grown at 25°C, they express very little *pos-1* mRNA and thus have minimal capacity for 22G-RNA production (Beanan and Strome 1992). Therefore, *pos-1* aligning sequences will correspond almost exclusively to the primary siRNAs produced from the exogenous dsRNA (Fig. 1A). Consistent with our previous findings, primary siRNAs were predominantly 22–23 nt long and depleted of 5'G-containing sequences (Fig. 1B; Zhang et al. 2012; Svendsen et al. 2019).

22G-RNAs contain triphosphates at their 5' ends, whereas primary siRNAs contain monophosphates (Pak and Fire 2007). Enzymatic treatment of small RNAs to reduce triphosphates to monophosphates before library construction can be used to enrich for 22G-RNAs, and conversely, excluding this treatment can be used to enrich for primary siRNAs (Almeida et al. 2019). Pretreatment of small RNA libraries from *pos-1*-treated *glp-4*-mutants with RNA 5' polyphosphatase to reduce triphosphates to monophosphates had little impact on the relative abundance of *pos-1*-matching reads (Fig. 1C). This is further evidence that these libraries have very low levels of 22G-RNAs produced from *pos-1* and confirms that exogenous siRNAs are monophosphorylated.

We previously showed that exogenous siRNAs are modified at their 3' ends by the methyltransferase HENN-1 (Svendsen et al. 2019). We observed an approximately twofold enrichment of *pos-1*-matching reads in libraries pretreated with sodium periodate (NaIO₄), an oxidizing agent that selectively targets unmodified 3' bases, inhibiting 3' adapter ligation to these sequences (Fig. 1C; Seitz et al. 2008; Yu and Chen 2010; Svendsen et al. 2019). NaIO₄ treatment further enriched for 23 nt sequences and depleted 5'G-containing sequences, likely due to the loss of residual 22G-RNAs present in *glp-4* mutants (Fig. 1D).

Canonical siRNAs are processed from longer dsRNA into ~21–24 nt duplexes by the RNaseIII-like enzyme Dicer (Bernstein et al. 2001; Elbashir et al. 2001; Ketting et al. 2001). Like other RNaseIII cleavage products, siRNA duplexes typically have 2 nt, 3'-overhangs (Robertson et al. 1968; Elbashir et al. 2001). Although viral siRNA duplexes commonly have 2 nt overhangs in *C. elegans*, in cell-free extracts, dsRNA is often processed into duplexes with 3–4 nt overhangs (Welker et al. 2011; Ashe et al. 2013). To determine the most prevalent duplex configuration in exogenous RNAi, we identified all possible *pos-1* siRNA duplexes produced during RNAi. To minimize contamination of 22G-RNAs, we examined sRNA-seq data from libraries treated with NaIO₄, since these libraries had the lowest levels of 22 nt, 5'G-containing reads (Fig. 1D). Even for the most abundant potential duplexes, one strand was always present at low levels, suggesting that strand stabilization is very selective (Supplemental Fig. S1A). Among the possible 0–4 nt, 3'-end overhangs, configurations with 2 nt overhangs on either strand of the duplex were most prevalent, particularly when only including possible duplexes in which one strand had ≥50 reads (Fig. 1E,F).

Since complementary sequences may not originate from the same duplex, we cannot confirm Dicer's preference for siRNA duplexes with specific 3' overhangs in vivo. Moreover, stabilization of certain duplexes or single strands postdicing may skew results. However, our data suggest Dicer processes dsRNA imprecisely, producing various 3' overhangs, with 2 nt overhangs being the most common. In summary, we conclude that exogenous siRNAs have the following attributes: a predominant length of 23 nt, a bias against 5'G, a 5' monophosphate, a duplex intermediate with ~2 nt, 3' overhangs, and a 2'-O-methyl group at their 3' ends (Fig. 1G).

RDE-1 and ALG-1 associate with exogenous siRNAs

RDE-1 associates with at least some exogenous siRNAs and is essential for RNAi (Tabara et al. 1999; Yigit et al. 2006). To more comprehensively assess the extent to which RDE-1 binds exogenous siRNAs, we co-immunoprecipitated (co-IP'd) GFP::RDE-1-small RNA complexes from animals undergoing RNAi against both *nrf-1* and *oma-1*, genes chosen for their mild phenotypes (Davis et al. 2022). Animals in this experiment, unlike in the *pos-1* RNAi experiments above, were able to mount a 22G-RNA amplification response since the endogenous targets were present. Small RNAs matching *nrf-1* or *oma-1* in the cell lysate input fractions used for GFP::RDE-1 co-IPs were predominantly 22 nt and strongly biased for 5'G and were thus comprised mainly of 22G-RNAs (Fig. 1H; Supplemental Fig. S1B). In contrast, *nrf-1*- and *oma-1*-matching reads from the GFP::RDE-1 co-IP fractions were predominantly 23 nt and were depleted of reads containing 5'G (Fig. 1H; Supplemental Fig. S1B).

siRNA and microRNA (miRNA) duplexes are structurally similar except that miRNA duplexes typically have internal mispairs. While some miRNAs bind RDE-1, they primarily function with ALG-1 and ALG-2, and to a lesser extent with ALG-5 (Grishok

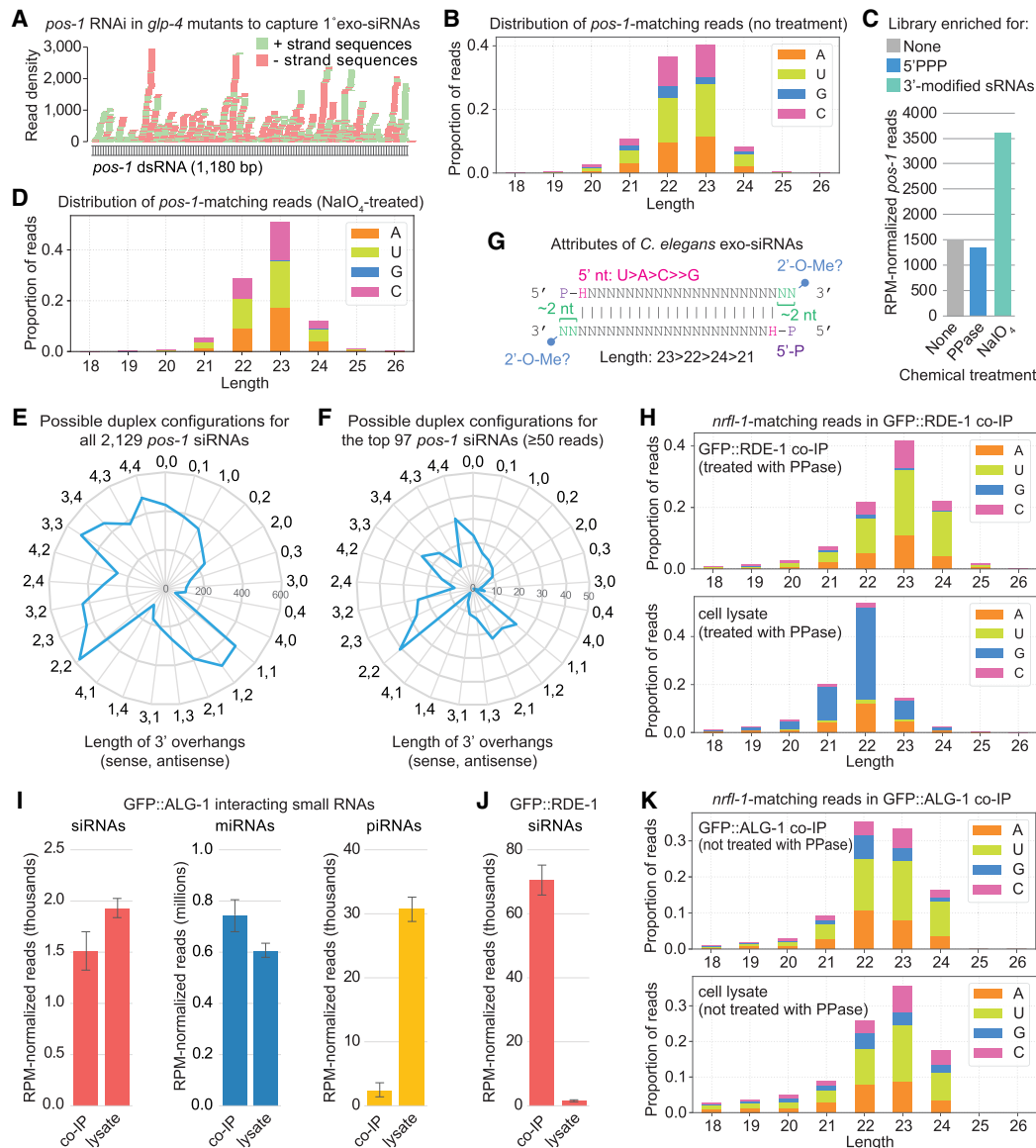


Figure 1. Attributes of exogenous siRNAs in *C. elegans*. (A) siRNA sequences matching exogenously delivered *pos-1* dsRNA in *glp-4* (*bn2*) mutants as identified using sRNA-seq. $n = 1$ biological replicate. Libraries from adults treated with *pos-1* RNAi and grown at 25°C to induce sterility. (B) Size and 5' nt distribution of *pos-1*-matching reads from the sRNA-seq libraries in A. (C) RPM-normalized *pos-1*-matching reads from untreated (none), RNA 5' polyphosphatase-treated (PPase, reduces 5' triphosphates), and NaIO₄-treated (blocks ligation of non-3'-end-modified small RNAs) sRNA-seq libraries. $n = 1$ biological replicate. RNA is the same as in the library in (A,B). (D) Size and 5' nt distribution of *pos-1*-matching reads from the NaIO₄-treated sRNA-seq library in C. (E) The numbers of *pos-1*-matching siRNA sequences with the possible duplex configurations indicated. Note that a single sequence could have multiple possible complementary strands with different 3'-overhangs. (F) As in E but limited to siRNAs with ≥50 reads. (G) Attributes of exogenous siRNAs in *C. elegans*. (H) Size and 5'-nt distribution of *nrfl-1*-matching reads from GFP::RDE-1 co-IP and cell lysate sRNA-seq libraries. $n = 2$ biological replicates. Data from one representative library for each condition are shown. Libraries from gravid adults treated with *nrfl-1* and *oma-1* RNAi. (I) RPM-normalized miRNA, piRNA, and *nrfl-1*-matching reads in GFP::ALG-1 co-IP and cell lysate sRNA-seq libraries. Libraries from gravid adults treated with *nrfl-1* RNAi. Error bars are standard deviation (SD) between three biological replicates. (J) RPM-normalized *nrfl-1*-matching reads in GFP::RDE-1 co-IP and cell lysate sRNA-seq libraries. Libraries from gravid adults treated with *nrfl-1* RNAi. Error bars are SD between three biological replicates. (K) Size and 5'-nt distribution of *nrfl-1*-matching reads from GFP::ALG-1 co-IP and cell lysate sRNA-seq libraries. Data from one representative library for each condition are shown.

et al. 2001; Corrêa et al. 2010; Brown et al. 2017; Seroussi et al. 2023). Since RDE-1 interacts with both siRNAs and miRNAs, this raises the possibility that other miRNA-associated Argonautes, such as ALG-1, may also bind siRNAs. To assess this, we sequenced small RNAs that co-IP'd with GFP::ALG-1 in animals undergoing *nrfl-1* RNAi. While *nrfl-1*-matching siRNAs were less enriched than miRNAs in GFP::ALG-1 co-IPs, based on reads per million (RPM)

normalization relative to the cell lysates, they were substantially more enriched than PIWI-interacting RNAs (piRNAs), which associate with PRG-1 and therefore represent carryover from the cell lysate (Fig. 1I; Batista et al. 2008; Das et al. 2008; Wang and Reinke 2008; Seroussi et al. 2023).

Despite their association with GFP::ALG-1, the relative abundance and enrichment of *nrfl-1*-matching siRNAs were

significantly lower compared to a parallel set of GFP::RDE-1 co-IPs, likely reflecting ALG-1's broader association with miRNAs (Fig. 1J; Brown et al. 2017; Seroussi et al. 2023). *nrf1-1* siRNAs were predominantly 23 nt in cell lysates used for GFP::RDE-1 and GFP::ALG-1 co-IPs (Fig. 1K; Supplemental Fig. S1C). These samples were not treated with RNA 5' polyphosphatase, hence the absence of 22G-RNAs in these libraries. siRNA reads were shifted slightly toward 22 nt sequences in GFP::ALG-1 co-IPs, but further toward 23–24 nt sequences in GFP::RDE-1 co-IPs, suggesting that the 23 nt bias may be driven by RDE-1 preferentially binding longer small RNAs (Fig. 1K; Supplemental Fig. S1C). Both Argonaute co-IPs displayed a lack of 5'G-containing siRNAs, implying either a shared bias or other factors influencing the 5' nt (Fig. 1K; Supplemental Fig. S1C). These results show that both RDE-1 and ALG-1 associate with exogenous siRNAs. However, since RDE-1 is specifically required for RNAi, the role of ALG-1 in this pathway, if any, is unclear (Tabara et al. 1999).

Identification of endogenous canonical siRNAs

Using the features of exogenous siRNAs defined above, we set out to identify their endogenous counterparts. We analyzed GFP::RDE-1 co-IP sRNA-seq libraries, expecting them to be enriched for endogenous siRNAs due to RDE-1's strong affinity for exogenous siRNAs. We filtered these libraries for 21–24 nt reads, a size range that would capture most exogenous siRNAs. A custom algorithm was used to identify complementary sequence pairs with 2 nt, 3' overhangs, which we predicted based on our analysis of exogenous siRNAs would most effectively capture endogenous siRNA duplexes. We further filtered for sequences with at least 10 reads and showing greater than twofold enrichment in GFP::RDE-1 co-IPs compared to cell lysates (Fig. 2A).

To broaden our search, we included additional RDE-1-associated sequences that aligned with loci producing the small RNAs identified above, without requiring that we identify a 2 nt, 3'-overhang duplex intermediate. This approach yielded 1196 candidate sequences perfectly matching 25,701 genomic loci (reflecting the repetitive nature of many of them, as described below) (Supplemental Table S1). Similar to exogenous siRNAs, these small RNAs were depleted of 5'G-containing sequences and enriched for 23 nt sequences (Fig. 2B). Cell lysates also showed this 23 nt enrichment and 5'G depletion in reads aligning to candidate regions (Fig. 2C). Given their length and bias against 5'G-containing sequences, we named these endogenous siRNAs 23H-RNAs, where H represents U, A, or C, following the nomenclature of other small RNA classes in *C. elegans* (Ruby et al. 2006; Gu et al. 2009; Han et al. 2009; Reich et al. 2018).

RNA 5' polyphosphatase treatment did not enrich for endogenous 23H-RNAs in wild-type sRNA-seq libraries, indicating that they do not contain a triphosphate group and are, therefore, likely monophosphorylated (Fig. 2D; Svendsen et al. 2019). In contrast, 23H-RNAs were moderately enriched in NaIO₄-treated libraries, indicating that their 3' ends are blocked, presumably by 2'-O-methylation (Fig. 2D; Svendsen et al. 2019). These results are consistent with the attributes we observed for exogenous siRNAs above.

23H-RNAs are produced from a variety of features, including hpRNAs and repetitive elements

Three genomic loci—F43E2.6, Y57G11C.1145, and *mir-5592*—accounted for ~98% of 23H-RNA reads in our GFP::RDE-1 co-IP sRNA-seq libraries (Fig. 3A; Supplemental Table S1). F43E2.6,

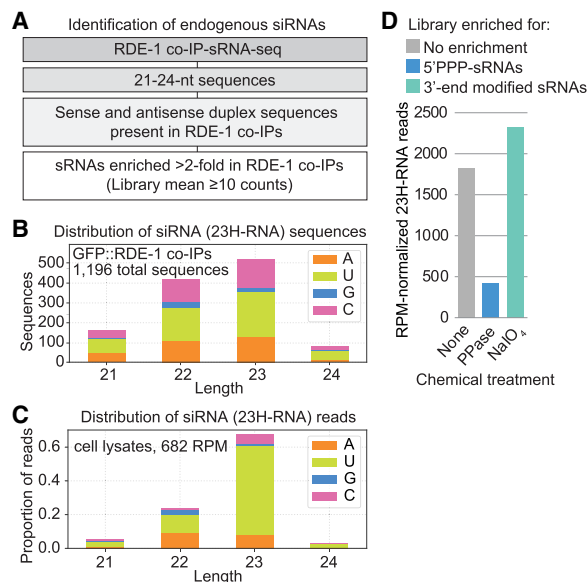


Figure 2. Identification of endogenous canonical siRNAs. (A) Canonical siRNA discovery flowchart. (B) Size and 5' nt distribution of all endogenous canonical siRNA (23H-RNA) sequences from GFP::RDE-1 co-IP sRNA-seq libraries. $n = 2$ biological replicates. Data from one representative gravid adult library are shown. (C) Size and 5' nt distribution of 23H-RNA reads in the cell lysates used for the GFP::RDE-1 co-IP in B. (D) 23H-RNA reads from untreated (none), RNA 5' polyphosphatase-treated (PPase, reduces 5' triphosphates), and NaIO₄-treated (blocks ligation of non-3'-modified small RNAs) sRNA-seq libraries. $n = 1$ biological replicate. Libraries from wild-type gravid adults grown at 25°C.

which is annotated as a coding gene but with no known function, produces siRNAs from both strands, indicating that there is overlapping transcription at the locus. However, insufficient high-throughput mRNA sequencing (mRNA-seq) reads prevented a detailed analysis (Fig. 3B; Warf et al. 2012; Reed et al. 2020). Y57G11C.1145 is a long hairpin RNA (hpRNA), reminiscent of hpRNAs found in flies and mice (Fig. 3A,C; Czech et al. 2008; Kawamura et al. 2008; Okamura et al. 2008b; Tam et al. 2008; Watanabe et al. 2008). Several other 23H-RNA loci form hairpin structures that range in length from tens to hundreds of base pairs (Fig. 3C).

mir-5592 is classified as a miRNA but is unusual in having a perfectly base-paired stem and in generating multiple distinct small RNAs (Fig. 3C; Supplemental Table S1). Most 23H-RNA candidates from the locus, including *mir-5592*, were depleted in animals with partial loss of function in *PASH-1*, a protein required for processing primary miRNA transcripts (Supplemental Fig. S2A; Supplemental Table S2; Montgomery et al. 2024). Similarly, most 23H-RNAs produced from the Y57G11C.1145 hairpin locus were also reduced in *pash-1* mutants (Supplemental Fig. S2A; Supplemental Table S2). However, 23H-RNAs produced from other hairpins, such as PALTA3 repetitive elements and the noncoding RNA *mcs-1*, were not depleted in *pash-1* mutants, nor were 23H-RNAs from non-hairpin-forming loci (Supplemental Fig. S2A; Supplemental Table S2; Hellwig and Bass 2008). It is possible that *mir-5592* and Y57G11C.1145 are incipient miRNAs, recognized by the processing machinery but not fully refined. *mir-5550*-derived small RNAs were also retained in our 23H-RNA list as they originate from a repetitive element, which is atypical of miRNAs but common among 23H-RNAs, as described below

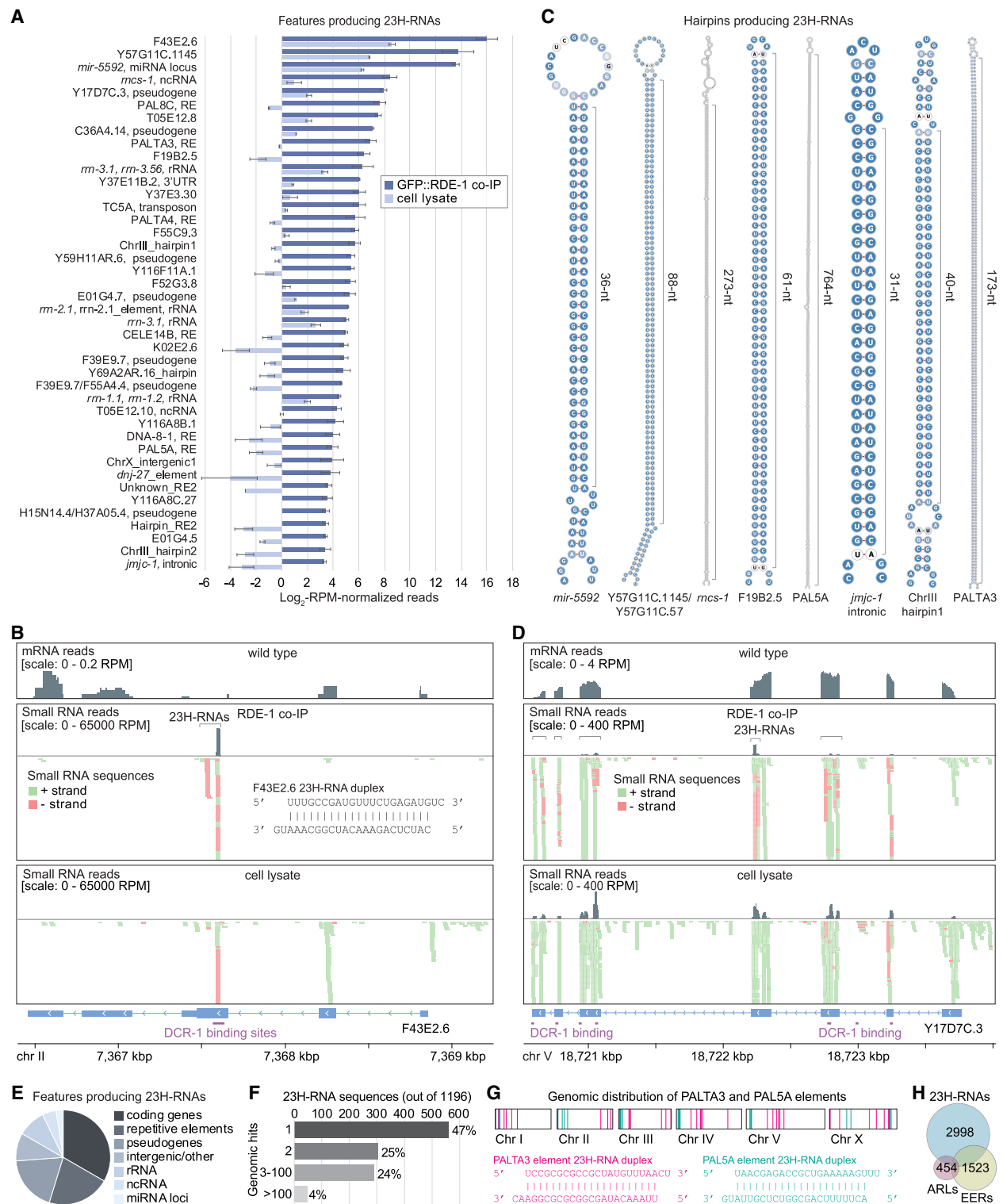


Figure 3. Genomic origins of 23H-RNAs. (A) Mean log₂-RPM-normalized 23H-RNA reads from GFP::RDE-1 co-IP and cell lysate sRNA-seq libraries. Only features producing 23H-RNAs with >10 RPM in GFP::RDE-1 co-IPs are shown. Error bars are SD between two biological replicates. Libraries are the same as in Figure 2A. (B) mRNA-seq read density from wild-type animals and sRNA-seq read density from GFP::RDE-1 co-IP and cell lysate libraries plotted along the F43E2.6 23H-RNA locus. Brackets span 23H-RNA clusters. The lower track shows individual sequences but is truncated due to space constraints. The sequence of the most abundant duplex is shown. Purple bars span DCR-1-binding sites based on DCR-1 PAR-CLIP data from Rybak-Wolf et al. (2014). (C) Secondary structures for several 23H-RNA loci that form hairpins. (D) As in B but for the gene Y17D7C.3. (E) The proportion of 23H-RNA loci with >10 RPM in GFP::RDE-1 co-IPs overlapping the indicated features. (F) 23H-RNA sequences are categorized by the number of perfectly matching genomic loci. (G) PALTA3 and PAL5A genomic distribution and the most abundant duplex from each repetitive element. (H) Overlap between 23H-RNA loci and EERs and ARLs. Numbers shown are for total features; 449 23H-RNAs and EERs overlap; 129 23H-RNAs and ARLs overlap; 93 EERs and ARLs overlap; and 24 in common to all three data sets.

(Supplemental Table S1). We identified reads from perfectly base-paired duplexes at several other miRNA loci but largely excluded them due to uncertainty in their classification.

23H-RNAs were often produced from among clusters of other small RNAs, but only the 23H-RNAs were enriched in GFP::RDE-1 co-IPs (Fig. 3B,D; Supplemental Fig. S2B). Putative protein-coding genes, such as F43E2.6, represented ~33% of 23H-RNA loci yielding >10 RPM in GFP::RDE-1 co-IPs, although it is unclear if they encode functional proteins (Fig. 3E). Another ~19% of these relatively high 23H-RNA-yielding loci are annotated as pseudogenes (Fig. 3E). The second largest source of 23H-RNAs was repetitive elements with ~28% produced from loci with >2 genomic copies and ~4% from loci with >100 copies (Fig. 3E–G). These loci are largely related to DNA transposons, such as CELE14B, DNA-8-1, PAL8A, PALTA3, PALTA4, PAL5A, and TC5A, pointing to a possible role for 23H-RNAs in transposon silencing (Fig. 3A,G).

Dicer is required for 23H-RNA biogenesis

Most loci producing relatively high levels of 23H-RNAs were identified as Dicer substrates in PAR-CLIP experiments, indicating that 23H-RNAs are processed through Dicer cleavage as anticipated based on their duplex structures (Fig. 3B,D; Supplemental Table S3; Rybak-Wolf et al. 2014). Additionally, total RPM-normalized levels of 23H-RNAs decreased by ~40% following *dcr-1* RNAi compared to control treatment (Supplemental Fig. S2C). Although most individual 23H-RNAs were modestly depleted following *dcr-1* RNAi, only a small subset passed our *P*-value threshold of 0.05, likely due to their low representation in our libraries (Supplemental Fig. S2D; Supplemental Table S4). We observed a similarly modest reduction in miRNA levels in these libraries (Supplemental Fig. S2E; Supplemental Table S4). The comparable effect of *dcr-1* RNAi on miRNAs, which are known Dicer substrates, and 23H-RNAs further supports a role for Dicer in 23H-RNA biogenesis (Grishok et al. 2001; Hutvagner et al. 2001; Ketting et al. 2001).

Dicer's helicase domain is required for processing certain endogenous small RNAs, particularly 26G-RNAs (Welker et al. 2010). The helicase domain is important for processing dsRNA containing blunt or single-stranded 5'-end overhangs, like 26G-RNAs, but not for dsRNA with single-stranded 3'-end overhangs, such as miRNAs (Welker et al. 2011). To determine if 23H-RNAs are dependent on Dicer's helicase domain, we analyzed sRNA-seq data from *dcr-1* mutants rescued with either wild-type *dcr-1* or a helicase mutant form of the gene (Welker et al. 2010, 2011). 23H-RNAs were only modestly depleted in animals containing the *dcr-1* helicase mutant, similar to some non-*dcr-1*-dependent small RNAs, in contrast to ERGO-1 class 26G-RNAs, which were almost completely lost (Supplemental Fig. S2F). Most 23H-RNA loci, including long dsRNA hairpins, produced only a few dominant duplexes, often only one or two (Fig. 3B,D,G; Supplemental Fig. 2B). Since the helicase domain is thought to promote Dicer processivity and 23H-RNAs are not strongly dependent on this domain, their biogenesis is likely not processive (Welker et al. 2011). This may explain why so few distinct 23H-RNAs are produced from each locus.

23H-RNA expression varies by developmental stage

We next compared the abundance of 23H-RNAs in gravid adults, the stage used for their identification, with L4 stage larvae, dissected distal gonads, and embryos. Most 23H-RNAs were significantly more abundant in embryos compared to adults (Supplemental Fig. S3A; Supplemental Table S5). dsRNA regions enriched in RNA ed-

iting also tend to be more highly expressed in embryos, which may reflect greater expression of transcripts forming dsRNA in embryos than in other developmental stages (Reich et al. 2018). Many 23H-RNAs also had elevated levels in the L4 stage relative to adults, and a small subset was more abundant in dissected gonads compared to whole adult animals (Supplemental Fig. S3B,C; Supplemental Table S5).

Overlap between 23H-RNAs and RNA-edited loci

We then compared 23H-RNA loci with RNA editing-enriched regions (EERs) from Reich et al. (2018) and ADAR-modulated RNA loci (ARLs) from Wu et al. (2011), which also produce siRNAs similar to 23H-RNAs but that are largely restricted to ADAR mutants (Wu et al. 2011; Reich et al. 2018). ADARs act posttranscriptionally, converting adenosine to inosine in dsRNA (Walkley and Li 2017). While extensive RNA editing typically prevents Dicer processing, limited editing can, in some instances, still permit siRNA production, albeit at low levels (Zamore et al. 2000; Scadden and Smith 2001; Warf et al. 2012). Clustering 23H-RNA loci that were within 1000 bp of each other resulted in 2998 distinct regions, of which, 449 overlapped with the 1523 EERs and 129 overlapped with the 454 ARLs (Fig. 3H; Supplemental Table S6); 2444 (~82%) 23H-RNA loci did not overlap with these previously reported siRNA clusters. Our focus on RNA editing-proficient animals likely explains the lack of overlap of 23H-RNAs with EER and ARL siRNA clusters, while the prior studies missed siRNA clusters from less edited sites, which presumably accounts for many of the loci identified here (Wu et al. 2011; Reich et al. 2018).

To check for RNA editing in our data, we realigned sRNA-seq reads from *pos-1* RNAi libraries and GFP::RDE-1 co-IP libraries allowing up to three genome mispairs. The two most abundant *pos-1* siRNAs had only ~3% and 0% of reads with A-to-G mismatches, which are indicative of RNA editing (Supplemental Fig. S4A). Approximately 22% of 23H-RNA reads from GFP::RDE-1 co-IP and cell lysate libraries had a single A-to-G mismatch and <2% had 2–3 mismatches (Supplemental Fig. S4B). miRNAs had only ~2.5% with 1 mismatch and <0.03% with 2–3 mismatches (Supplemental Fig. S4B). Exogenous *mtfl-1* siRNAs had similarly low levels, with only ~3% having 1 mismatch and ~0.1% having 2–3 mismatches (Supplemental Fig. S4B). sRNA-seq reads from most individual 23H-RNAs and miRNAs had <5% A-to-G mismatches, although some showed much higher levels (Supplemental Fig. S4C; Supplemental Table S7). However, many with seemingly high proportions of A-to-G mismatches were likely due to misalignment of more abundant reads from an alternative small RNA locus (Supplemental Fig. S4C). Nonetheless, for F43E2.6, for example, which was previously shown to produce siRNAs from an edited precursor, ~39% sense-strand 23H-RNA reads had 1–3 A-to-G mismatches, which were mostly indicative of a single edit and could be unambiguously assigned to the locus (Supplemental Fig. S4D; Warf et al. 2012). In contrast, only ~8%–9% of F43E2.6 antisense 23H-RNA reads had A-to-G mismatches (Supplemental Fig. S4D). The difference between these two strands of the same duplex may relate to the number of editable sites. Non-A-to-G mismatches accounted for 5%–7% of both F43E2.6 sense and antisense strand reads, which were likely due to 3'-end tailing and library preparation mutations (Supplemental Fig. S4D). These results indicate that endogenous 23H-RNAs can occasionally arise from lightly edited dsRNA (mostly a single A-to-G edit), whereas exogenous siRNAs are generally not produced from edited dsRNA.

RDE-1 has a high affinity for 23H-RNAs

Although RDE-1 primarily associates with miRNAs, ~14% of the total sRNA-seq reads from GFP::RDE-1 co-IPs corresponded to the 23H-RNA sequences annotated here, despite these small RNAs making up a tiny fraction of reads in the cell lysates (Fig. 4A). Although 23H-RNAs were a small percentage of total RDE-1 associated small RNAs, individual 23H-RNAs were strongly enriched in GFP::RDE-1 co-IPs compared to most miRNAs, with enrichment levels comparable to *nrf1-1* and *oma-1* exogenous siRNAs (Fig. 4B; Supplemental Table S8). This suggests that RDE-1 has a higher affinity for canonical siRNAs than for most miRNAs and other small RNAs.

Despite their strong association with RDE-1, 23H-RNAs were not depleted in *rde-1(ne219)* mutants but instead appeared somewhat elevated (Fig. 4C). The *rde-1(ne219)* allele is a point mutation and could possibly still produce a protein with the ability to bind and stabilize 23H-RNAs (Tabara et al. 1999). To confirm that 23H-RNAs are not dependent on *rde-1* for their stability, we generated a complete deletion allele of *rde-1* (*ram40*). Quantitative real-time PCR (qRT-PCR) analyses showed a modest increase in the level of an F43E2.6 siRNA in the deletion allele, confirming that *rde-1* is not required for the stability of 23H-RNAs (Fig. 4D).

In the absence of *rde-1*, 23H-RNAs might be stabilized through interactions with other Argonautes, such as ALG-1, which we showed also binds exogenous siRNAs. To identify other Argonautes that might bind and stabilize 23H-RNAs, we examined co-IP sRNA-seq data from Seroussi et al. (2023) for Argonautes involved in Dicer-dependent small RNA pathways (RDE-1, ALG-1, ALG-2, ALG-4, ALG-5, and ERGO-1) (Seroussi et al. 2023). Using DESeq2 for normalization (which involves calculating the ratio of each small RNA's counts to the geometric mean of counts for that small RNA across all samples and then normalizing by the median of these ratios), most 23H-RNAs were enriched in GFP::RDE-1 co-IPs in these data sets, as well as in co-IPs from the major miRNA Argonautes, ALG-1 and ALG-2 (Supplemental Fig. S5A–C; Supplemental Table S9; Love et al. 2014). A small subset of 23H-RNAs was also enriched in the germline-specific miRNA Argonaute, ALG-5 (Supplemental Fig. S5D; Supplemental Table

S9; Seroussi et al. 2023). Some were also enriched in ERGO-1, but not in ALG-4, both of which bind 26G-RNAs, which could relate to the overlap between ERGO-1 class 26G-RNA and 23H-RNA loci, as described below (Supplemental Fig. S5E,F; Supplemental Table S9).

When measured as a proportion of RPM-normalized reads, 23H-RNAs were only collectively enriched in GFP::RDE-1 co-IPs, relative to cell lysates, although this normalization method does not account for the distribution of individual 23H-RNAs across samples (Supplemental Fig. S5G). This suggests that while other Argonautes can bind 23H-RNAs, RDE-1 has a much higher affinity for this class of small RNAs. It is also important to note that the vastly different small RNA profiles of cell lysates and Argonaute co-IPs may introduce normalization artifacts, potentially confounding these results. Nonetheless, 23H-RNAs bind preferentially to RDE-1 but also broadly interact with ALG-1 and ALG-2, which we confirm below.

Competition between Argonautes for 23H-RNAs

We hypothesized that RDE-1 competes with other Argonautes for 23H-RNAs and tested this by assessing whether the binding affinities of ALG-1, ALG-2, ALG-5, and ERGO-1, which each showed enrichment for some 23H-RNAs, were elevated in the absence of RDE-1. We treated animals with *rde-1* RNAi and then co-IP'd each of the other Argonautes. By western blot, we detected an ~90% reduction in RDE-1 protein levels following *rde-1* RNAi, indicating that *rde-1*, despite being required for RNAi, was effectively knocked down (Supplemental Fig. S6A). In GFP::ALG-1 and GFP::ALG-2 co-IPs from animals treated with control RNAi, F43E2.6 23H-RNAs were enriched ~16-fold and approximately eightfold, respectively, as measured by qRT-PCR, confirming that both ALG-1 and ALG-2 bind 23H-RNAs (Supplemental Fig. S6B). Following *rde-1* RNAi, F43E2.6 23H-RNAs were further enriched to ~256-fold in GFP::ALG-1 co-IPs and ~100-fold in GFP::ALG-2 co-IPs, relative to cell lysates (Supplemental Fig. S6B). F43E2.6 23H-RNAs were not substantially enriched in GFP::ALG-5 or GFP::ERGO-1 co-IPs regardless of RNAi-treatment, indicating

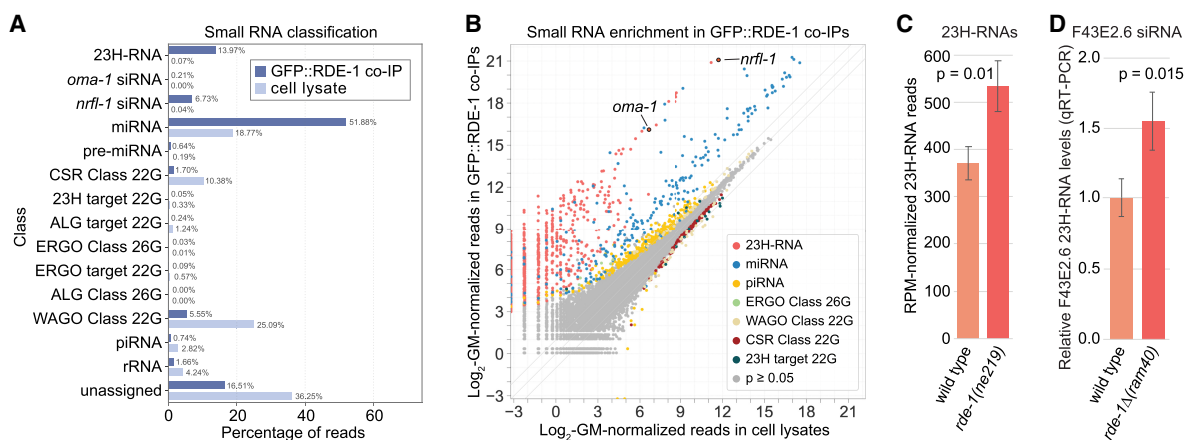


Figure 4. 23H-RNAs have a high affinity for RDE-1. (A) Percentages of RPM-normalized reads in GFP::RDE-1 co-IP and cell lysate sRNA-seq libraries corresponding to each class of small RNAs. $n = 2$ biological replicates. Libraries are the same as in Figure 2A. (B) Scatter plots display individual small RNA features as the average log₂-geometric mean-normalized reads in cell lysates (x -axis) and GFP::RDE-1 co-IPs (y -axis) colored by small RNA class. Exogenous siRNAs matching *nrf1-1* and *oma-1* are circled. (C) RPM-normalized 23H-RNA reads in wild-type and *rde-1(ne219)* mutant sRNA-seq libraries. Libraries from gravid adults. Error bars are SD between three biological replicates. The P -value was calculated using a two-sample t -test. (D) Relative levels of the most abundant F43E2.6 siRNA in wild-type and *rde-1Δ(ram40)* mutants as determined by qRT-PCR and normalized to miR-1 levels. RNA from gravid adults. Error bars are SD between three biological replicates. The P -value was calculated using a two-sample t -test.

that these Argonautes do not compete with RDE-1 for F43E2.6 23H-RNAs (Supplemental Fig. S6B).

To determine if ALG-1 and ALG-2 more generally compete with RDE-1 for 23H-RNAs, we performed sRNA-seq on co-IPs and cell lysates from control and *rde-1* RNAi-treated animals. While most small RNA reads (~97%–98%) from GFP::ALG-1 and GFP::ALG-2 co-IPs were miRNAs, a small fraction corresponded to 23H-RNAs (Fig. 5A,B). In animals treated with *rde-1* RNAi, 23H-RNA reads from GFP::ALG-1 and GFP::ALG-2 co-IPs were elevated ~4.5- and ~3.6-fold, respectively, relative to animals treated with control RNAi, based on RPM-normalization (Fig. 5A,B). In contrast, miRNA levels in GFP::ALG-1 and GFP::ALG-2 co-IPs were unchanged between *rde-1* and control RNAi, suggesting that there is competition specifically for 23H-RNAs and not miRNAs (Fig. 5A, B). Likewise, piRNA levels were unchanged, and given that they bind PRG-1, their presence in the co-IPs likely reflects the level of small RNA carryover from cell lysates rather than an authentic interaction (Fig. 5A,B). Although 23H-RNAs were present in the co-IPs far above what would be expected due to carryover from the cell lysates based on what we observed for piRNAs, they were not collectively enriched relative to the cell lysates, except in GFP::ALG-1 co-IPs from *rde-1* RNAi-treated animals (Fig. 5A,B). This likely relates to the high affinity of ALG-1 and ALG-2 for miRNAs, which were enriched ~2.5-fold in co-IPs relative to cell lysates (Fig. 5A,B).

Most individual 23H-RNAs were elevated in both GFP::ALG-1 and GFP::ALG-2 co-IPs relative to cell lysates based on geometric mean normalization, consistent with our earlier analysis (Supplemental Fig. S6C,D; Supplemental Table S10; Love et al. 2014). Many individual 23H-RNAs were enriched in co-IPs from

rde-1 RNAi-treated animals relative to control treatment but were largely unchanged in cell lysates (Supplemental Fig. S6C,D). These results suggest that ALG-1 and ALG-2 compete with RDE-1 for 23H-RNAs, although ALG-1 and ALG-2 seem to have similar or lower affinities for 23H-RNAs compared to miRNAs which may limit their interactions (Supplemental Fig. 6C,D). *rde-1*-matching siRNAs co-IP'd with GFP::ALG-1 and GFP::ALG-2 in the animals treated with *rde-1* RNAi, supporting our earlier findings that ALG-1 binds exogenous siRNAs and revealing that ALG-2 also binds siRNAs (Fig. 5C,D; Supplemental Fig. 6C,D).

Multiple small RNA pathways converge on 23H-RNA loci

During exogenous RNAi, RDE-1 directs target mRNAs into a secondary small RNA amplification pathway to produce 22G-RNAs (Sijen et al. 2001; Pak and Fire 2007). Additionally, RDE-1, in association with miR-243, directs the endogenous target Y47H10A.5 into the 22G-RNA pathway, leading to its silencing (Corrêa et al. 2010). To determine if endogenous 23H-RNAs also trigger 22G-RNA production, we searched sRNA-seq data sets from wild-type animals for 22 nt, 5'G-containing sequences produced from the 36 23H-RNA loci with annotated transcripts and >10 RPM 23H-RNA counts in GFP::RDE-1 co-IPs (Phillips et al. 2014). We identified 22 nt, 5'G containing small RNAs at each of these loci, and consistent with their designation as 22G-RNAs, they were almost entirely lost in *mut-14 smut-1* mutants, which are defective in 22G-RNA production (Fig. 6A; Supplemental Table S5; Phillips et al. 2014). Only a small number of these loci, which included some of the highest 23H-RNA-yielding loci, such as F43E2.6, Y57G11C.1145, and Y57G11C.57, were depleted of 22G-RNAs in *rde-1* mutants (Fig. 6B; Supplemental Table S5). This points to a modest or redundant role for RDE-1 in directing 22G-RNA formation from 23H-RNA targets.

Since many of the 23H-RNA loci were not substantially depleted of 22G-RNAs in *rde-1* mutants, we hypothesized that additional small RNA pathways target these loci. In *C. elegans*, piRNAs and ERGO-1 class 26G-RNAs also direct their targets into the 22G-RNA pathway (Han et al. 2009; Gent et al. 2010; Vasale et al. 2010; Bagijn et al. 2012; Lee et al. 2012). *piwi/prg-1* mutants, which lack piRNAs, produced higher levels of 22G-RNAs, based on sRNA-seq, possibly an artifact due to the loss of piRNAs and piRNA-dependent 22G-RNAs (Fig. 6C; Supplemental Table S5; Batista et al. 2008; Das et al. 2008; Wang and Reinke 2008). Only one 23H-RNA target, *jmj-1*, had substantially reduced levels of 22G-RNAs in *prg-1* mutants, indicating that piRNAs do not have a role in triggering 22G-RNA production from most 23H-RNA loci (Fig. 6C; Supplemental Table S5). In *ergo-1* mutants, which lack oogenic/embryonic 26G-RNAs, ~50% of the 23H-RNA loci analyzed were depleted of 22G-RNAs, which is consistent with many of these loci, such as Y17D7C.3

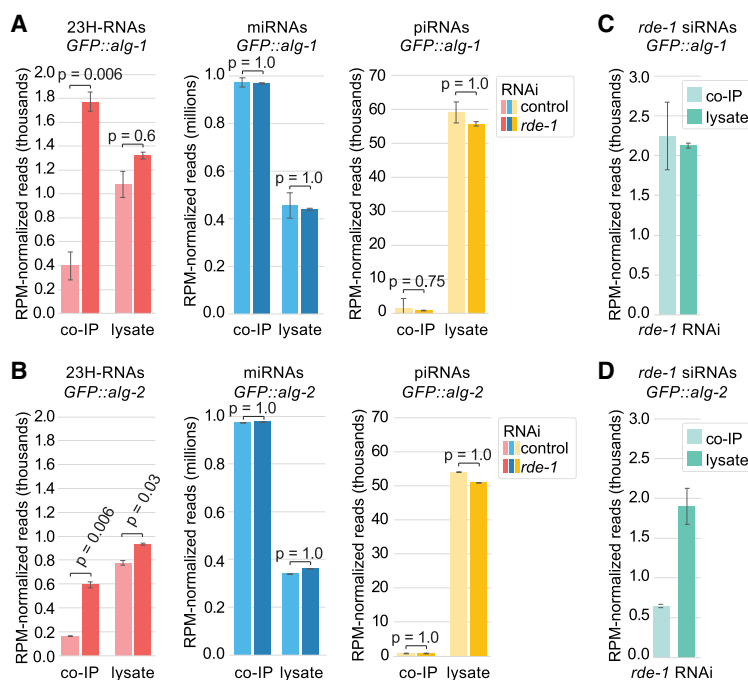


Figure 5. Competition between Argonautes for 23H-RNAs. (A,B) RPM-normalized reads corresponding to 23H-RNAs, miRNAs, and piRNAs from GFP::ALG-1 (A) or GFP::ALG-2 (B) co-IP and corresponding cell lysate sRNA-seq libraries. Two-sample *t*-tests were used to calculate the *P*-values. Bonferroni corrections were applied to account for multiple comparisons. Libraries from gravid adults treated with control (L4440) or *rde-1* RNAi. Error bars are SD between two biological replicates. (C,D) RPM-normalized *rde-1*-matching reads from GFP::ALG-1 (C) or GFP::ALG-2 (D) co-IP and corresponding cell lysate sRNA-seq libraries.

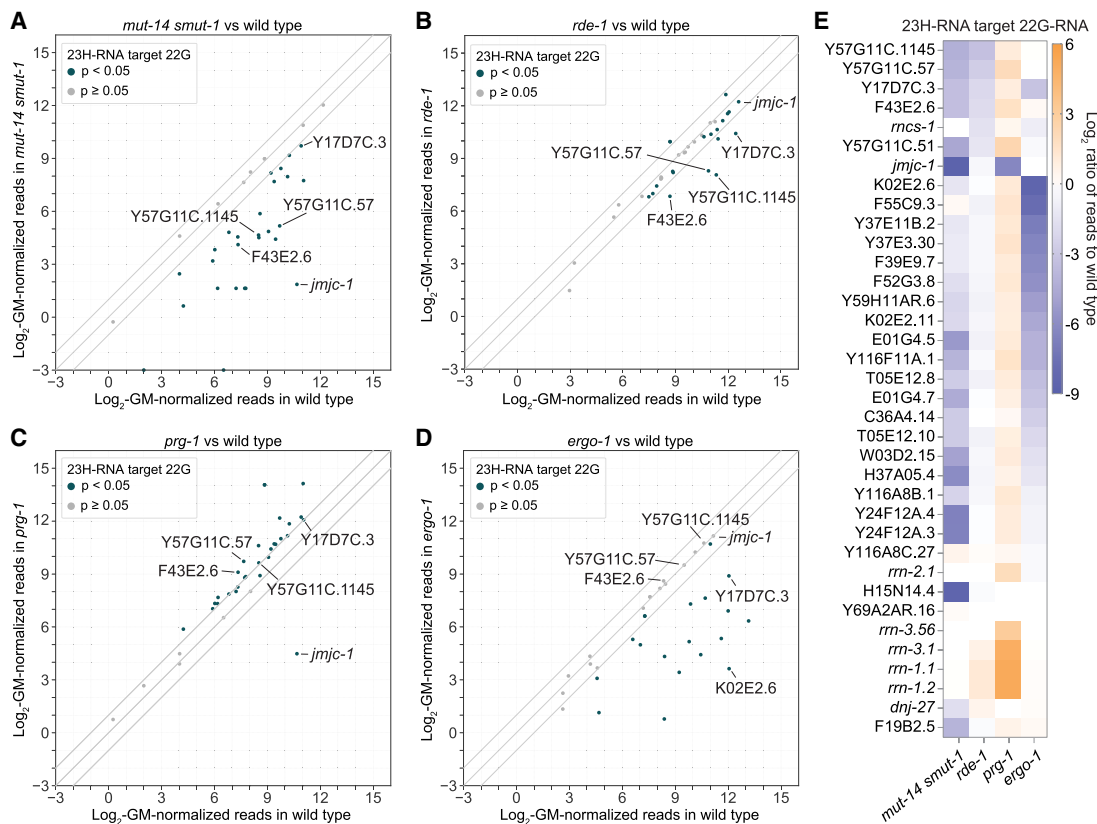


Figure 6. Genetic requirements for 22G-RNA formation from 23H-RNA loci. (A–D) Scatter plots display individual 23H-RNA loci as the average log₂-geometric mean-normalized 22G-RNA reads in wild-type (x-axis) and mutant (y-axis) animals. (A) *mut-14(mg464) smut-1(tm1301)*. (B) *rde-1(ne219)*. (C) *prg-1(n4357)*. (D) *ergo-1(tm1860)*. All libraries from gravid adults. $n = 3$ biological replicates for each except *mut-14(mg464) smut-1(tm1301)* ($n = 2$). (E) Average log₂ ratios of 22G-RNA reads for each of the top 23H-RNA loci (>10 RPM) in the indicated mutants to wild-type. Libraries are the same as in (A–D).

and K02E2.6, also being targeted by the ERGO-1 pathway (Fig. 6D; Supplemental Table S5; Han et al. 2009; Gent et al. 2010; Vasale et al. 2010; Fischer et al. 2011).

In contrast to 22G-RNAs, 23H-RNAs were not depleted in *mut-14 smut-1* mutants, indicating that the *Mutator* pathway involved in WAGO class 22G-RNA formation is not required for 23H-RNA formation (Supplemental Fig. S7A; Supplemental Table S5). Consistent with our earlier conclusion that RDE-1 is generally dispensable for 23H-RNA formation or stability, only five individual 23H-RNAs were significantly depleted in *rde-1* mutants (Supplemental Fig. S7B; Supplemental Table S5). None of the 23H-RNAs were significantly depleted in *prg-1* mutants and only two were affected in *ergo-1* mutants, indicating that the loss of 22G-RNAs we observed in these mutants, particularly *ergo-1*, is not due to the loss of 23H-RNAs (Supplemental Fig. S7C,D; Supplemental Table S5). Our results demonstrate that multiple small RNA pathways—specifically 23H-RNA, ERGO-1 class 26G-RNA, and WAGO class 22G-RNA pathways—converge on a common set of targets (Fig. 6E).

Fertility and fitness of *rde-1* mutants

rde-1 mutants did not display any noticeable developmental defects, nor did animals with a deletion of the F43E2.6 locus, which produces the most abundant 23H-RNAs (Supplemental Fig. S8A, B). We observed a slight reduction in fertility in *rde-1(ne219)* mutants at 25°C, consistent with previous findings (Supplemental

Fig. S8C; Seroussi et al. 2023). However, no significant reduction in progeny was observed in a distinct substitution allele of *rde-1*, *ne300*, nor in the *rde-1(ram40)* deletion allele (Fig. 7A; Supplemental Fig. S8C). Fertility was unchanged in F43E2.6 mutants as well (Supplemental Fig. S8C). The overall mobility and movement behavior of *rde-1* and F43E2.6 mutants also did not differ significantly from wild-type animals (Supplemental Fig. S8D–F). While loss of abundant 23H-RNAs could free up bandwidth in the RNAi pathway, deletion of F43E2.6 did not enhance RNAi against the essential fertility gene *hmr-1*, which responds weakly to RNAi in wild-type, but can be elevated in mutants with enhanced RNAi efficacy, such as *eri-5* (Supplemental Fig. S8G; Fischer et al. 2008; Thivierge et al. 2012).

We also explored the role of *rde-1* in regulating endogenous genes using mRNA-seq. None of the 23H-RNA loci were differentially expressed in *rde-1(ram40)* mutant gravid adults, although many had too few mRNA-seq reads to evaluate (Supplemental Table S11). Despite the prevalence of 23H-RNAs derived from repetitive elements, we did not observe substantial changes to transposon expression in *rde-1* mutants (Fig. 7B; Supplemental Table S11). It is possible that other small RNA pathways can compensate for the loss of *rde-1* function or that *alg-1* and *alg-2* offset the loss of *rde-1* in the 23H-RNA pathway. This could explain why *rde-1* mutants lack a discernible phenotype, aside from being RNAi-defective. The loss of both *alg-1* and *alg-2* causes embryonic lethality, which complicates further investigation (Grishok et al. 2001).

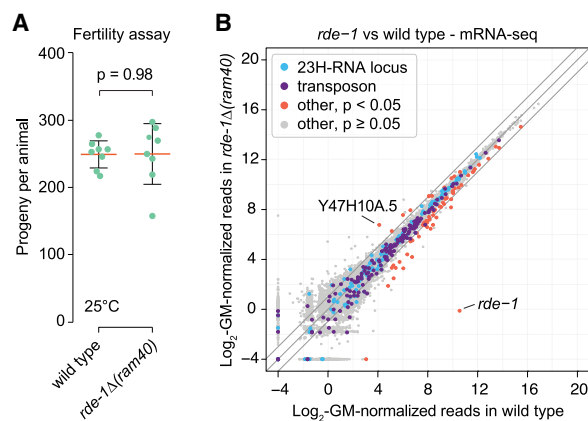


Figure 7. Fertility and differential mRNA expression in *rde-1* mutants. (A) Each data point is the number of progeny produced by a single wild-type or *rde-1* Δ (*ram40*) mutant animal grown at 25°C. The orange bars show the means and the black error bars show the 95% confidence intervals. A two-sample *t*-test was used to calculate the *P*-value. *n* = 8 animals per genotype. (B) The scatter plot displays individual transcripts as a function of the average log₂-geometric mean-normalized reads from wild-type (*x*-axis) and *rde-1* Δ (*ram40*) mutant (*y*-axis) mRNA-seq libraries colored by classification or *P*-value. None of the 23H-RNA loci had *P*-values < 0.05. *n* = 3 biological replicates for each strain. Libraries from gravid adults.

The most significantly upregulated transcript in *rde-1* mutants was Y47H10A.5, which is regulated by RDE-1 through its association with the miRNA miR-243 (Fig. 7B; Supplemental Table S11; Corrêa et al. 2010). In total, 70 transcripts were misexpressed in *rde-1* mutants, which were mostly downregulated. It is unlikely that all, if any, of these are direct targets of RDE-1 since RDE-1 represses its exogenous targets via the RNAi pathway, meaning endogenous targets would be expected to be upregulated in *rde-1* mutants (Fig. 7B; Supplemental Table S11). Among the differentially expressed transcripts in *rde-1* mutants were 11 genes with possible roles in innate immunity, all but one of which were downregulated (GO term enrichment, *P* = 0.001, Benjamini correction) (Supplemental Table S11). Additionally, three genes (*hsp-70*, F44E5.4, and F44E5.5) encoding heat shock proteins were downregulated and two genes (*vit-1* and *vit-2*) encoding the egg yolk precursor vitellogenin were also downregulated (Supplemental Table S11). It is possible that under less optimal growth conditions or when exposed to stress or xenobiotics, phenotypes related to these misexpressed genes could emerge in *rde-1* mutants.

Discussion

In this study, we characterized the endogenous canonical siRNAs of *C. elegans*. We hypothesized that these siRNAs were largely overlooked in earlier studies due to their relatively low abundance compared to other small RNA classes. Our analysis confirmed this by uncovering several hundred new small RNAs from RDE-1 co-IPs that exhibit the characteristics of canonical siRNAs, many of which were rare and often undetectable in cell lysates. Our results add to the already expansive repertoire of *C. elegans* small RNAs and demonstrate that endogenous siRNAs like those found in mice and flies are prevalent but typically at low levels in worms. We compiled these siRNAs, along with miRNAs, 21U-RNAs, 22G-RNAs, and 26G-RNAs into a GFF3-formatted annotation file compatible with *C. elegans* genome releases WS235–WS290+ and WS279 annotations, which can be utilized in most high-through-

put sequencing data analysis software (Supplemental Code) (updated versions will be available from MontgomeryLab.org). For optimal sensitivity, we recommend using RNA from embryos or RDE-1 co-IPs when analyzing 23H-RNAs.

RNA editing is antagonistic to RNAi and related pathways, likely because mispairs introduced by ADARs inhibit dsRNA processing by Dicer (Scadden and Smith 2001; Knight and Bass 2002; Tonkin and Bass 2003; Yang et al. 2005; Ohta et al. 2008; Heale et al. 2009; Sebastiani et al. 2009; Wu et al. 2011; Warf et al. 2012; Reich et al. 2018; Fischer and Ruvkun 2020). This could explain the rarity of most 23H-RNAs. Why some dsRNA is resistant to RNA editing and gives rise to 23H-RNAs is unclear.

It is intriguing that two of the hairpins we identified as 23H-RNA loci, *mir-5592* and Y57G11C.1145, are depleted of small RNAs in *pash-1* mutants, suggesting that they possess features that flag them for processing by the miRNA machinery. However, due to their strong association with RDE-1, their extended perfectly base-paired hairpins, and their tendency to produce multiple distinct small RNAs, we propose classifying them as 23H-RNAs. These loci are not conserved outside of *C. elegans* and may represent emerging miRNAs that have not yet been fully integrated into the miRNA pathway or regulatory networks. The differential dependence on *pash-1* for hairpin-derived small RNAs is noteworthy and could indicate intermediates evolving between 23H-RNAs and miRNAs.

The hyper-enrichment of 23H-RNAs and exogenous siRNAs in RDE-1 co-IPs, compared to most miRNAs, may relate to the near-perfect double-stranded nature of the duplex intermediates. Supporting this, nearly all the most highly enriched miRNAs in GFP::RDE-1 co-IPs derive from near-perfect duplexes (<3 mismatches) (Seroussi et al. 2023). While RDE-1 shows a high affinity for 23H-RNAs, ALG-1 and ALG-2 do not appear to prefer them over miRNAs. Whether this relates to the duplex structure is unclear. It is also possible that each of these Argonautes binds small RNA duplexes produced by Dicer indiscriminately, but that RDE-1 preferentially stabilizes duplexes for which it can cleave and discard the passenger strand, a necessary step specifically in the maturation of RDE-1-small RNA complexes (Steiner et al. 2009). Our findings that 23H-RNAs are enriched in ALG-1 and ALG-2 complexes following *rde-1* knockdown suggest that there is competition between the Argonautes for these small RNAs. It is possible that another factor normally assists in pairing 23H-RNAs with RDE-1, and in its absence, these siRNAs become more accessible to other Argonautes.

The mild phenotypes we observed in *rde-1* mutants suggest that under optimal growth conditions, RDE-1 is not important for development. The lack of a clear phenotype in *rde-1* mutants is reminiscent of the mild phenotypes originally observed in *Drosophila* AGO2 mutants, which appear superficially wild-type but have nonessential roles in germ cell development and male fertility, likely due to loss of siRNA activity (Okamura et al. 2004; Xu et al. 2004; Deshpande et al. 2005; Wen et al. 2015). Similar subtle phenotypes could emerge for *rde-1* mutants through more detailed analyses. We identified several genes with possible roles in innate immunity downregulated in *rde-1* mutants in our mRNA-seq data. It is possible that in addition to RDE-1's role in antiviral defense via the RNAi pathway, it has a role in regulating endogenous genes involved in innate immunity (Schott et al. 2005; Wilkins et al. 2005; Félix et al. 2011). Such a role could involve 23H-RNAs or miRNAs. In either case, it is possible that this or other roles for *rde-1* are masked by redundancy with *alg-1* and *alg-2*.

Future studies exploring 23H-RNAs will likely uncover their biological roles. For now, several knowledge gaps remain,

including whether 23H-RNAs act in *cis* or *trans*, if they function through their association with RDE-1, ALG-1, or ALG-2, and their roles in regulating endogenous genes. Additionally, it will be important to explore their biogenesis, sorting between Argonautes, and the mechanism by which they might regulate their targets. Our study provides a framework for addressing these questions and for further exploring the 23H-RNA pathway.

Methods

Strains

USC1080[*rde-1(rde-1(cmp133[(GFP + loxP + 3xFLAG)::rde-1]) V)*], WM27[*rde-1(ne219) V)*], WM45[*rde-1(ne300) V)*], PQ530[*alg-1(ap423 [3xFLAG::GFP::alg-1]) X)*], JMC203[*alg-2(tor139[GFP::3xFLAG::alg-2b]) II)*], JMC209[*alg-5(tor145[GFP::3xFLAG::alg-5]) I)*], and JMC211[*ergo-1(GFP::3xFLAG::ergo-1a) V)*] were previously described (Tabara et al. 1999; Aalto et al. 2018; Svendsen et al. 2019; Seroussi et al. 2023). TAM134[F43E2.6(*ram37*) II] and TAM151[*rde-1(ram40) V)*] were generated from wild-type (N2) *C. elegans* using CRISPR–Cas9-mediated genome editing (Gasiunas et al. 2012; Jinek et al. 2012; Cong et al. 2013; Mali et al. 2013). Additional information is in Supplemental Material.

RNAi

Synchronized L1 larvae were grown at 20°C on *E. coli* HT115 expressing dsRNA matching target genes (Kamath et al. 2003). Gravid adults were collected for RNA isolation or Argonaute co-IPs at 72 h. *pos-1* RNAi and associated data were from a previous experiment (Supplemental Table S12; Svendsen et al. 2019).

Protein-small RNA co-IPs

Co-IPs were performed on two or three biological replicates for each strain, as indicated. Animals were flash-frozen in liquid nitrogen and homogenized in cell lysis buffer. Cleared cell lysates were split into cell lysate and co-IP fractions. Proteins were co-IP'd with GFP-Trap Magnetic Agarose Beads (Proteintech gtm-100) for 1 h. Following three washes, beads were split between protein and RNA fractions. Protein fractions were heated at 95°C for 5 min in 1× Blue Protein Loading Dye (New England Biolabs B7703S). See Supplemental Material for additional details.

RNA isolation

RNA was isolated using TRIzol according to the manufacturer's recommendations except that two chloroform extractions were done (Invitrogen 15596018). RNA was precipitated in isopropanol overnight at –30°C in the presence of 20 µg glycogen.

Western blot

Protein fractions from GFP::3xFLAG::RDE-1 co-IPs and cell lysates were resolved on a Bolt 15-well 4%–12% Bis-Tris Plus Gel (Invitrogen NW04125BOX). Proteins were transferred to a nitrocellulose membrane and probed with FLAG (Sigma F3165) and actin (Abcam ab3280) antibodies. Blots were imaged on a FluorChem E Imaging System (ProteinSimple) and quantification was performed using ImageJ. The *P*-value was calculated using a two-sample *t*-test.

Small RNA sequencing

sRNA-seq libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina following the manufactur-

er's recommendations (New England Biolabs E7300S). See Supplemental Material for additional details.

Small RNA data analysis

sRNA-seq data were analyzed and plotted using the default configuration in tinyRNA (Tate et al. 2023). For analysis of A-to-G mismatches, the tinyRNA configuration was modified to allow for up to three genome mismatches during Bowtie alignment (end_to_end: 3) and the mismatch pattern was set to ADAR (counter_mismatch_pattern: ADAR), which only counts perfect genome-matching reads and mismatched reads with 1–3 A-to-G mispairs (Langmead et al. 2009; Tate et al. 2023). Reference genome sequences and annotations were from the *C. elegans* WS279 genome release (Davis et al. 2022). Custom Python scripts (Supplemental Code) were used to identify complementary sequences with 2 nt, 3'-overhangs enriched in GFP::RDE-1 co-IPs from tinyRNA alignment tables generated by modifying the counter option in the configuration file (counter_diags: True). 23H-RNA and other small RNA annotations are available in GFF3 file format as Supplemental Code. Hairpin structures were predicted and drawn with RNAfold (Gruber et al. 2008). The DESeq2 R package was used within the tinyRNA pipeline to compute normalized counts using the geometric mean method and perform statistical analysis using the Wald test (Love et al. 2014). Matplotlib, R, IGV, and Adobe Illustrator were also used for plotting and statistical analysis of data (Hunter 2007; Robinson et al. 2011; R Core Team 2021). See Supplemental Table S12 for a complete list of libraries.

qRT-PCR

Small RNA TaqMan qRT-PCR was done following the manufacturer's recommendations (Invitrogen 4331348) (Supplemental Material). Two-sample *t*-tests were used to compare differences between conditions. Microsoft Excel and GraphPad Prism were used to draw plots and perform statistical analysis.

Microscopy

Animals were imaged 48 h (for L4 stage larvae) or 72 h (for adults) after L1 synchronization. Animals were imaged on growth media on an Axio Imager Z2 Microscope (Zeiss) using a 5X objective.

Brood size assays

Individual animals were grown on OP50 and the numbers of progeny produced by each animal were counted each day until the cessation of egg laying. For the RNAi efficacy assay, 20 animals per replicate were treated with *hmr-1* RNAi, and the total progeny/replicate pool were counted at day 2 of adulthood. One-way ANOVA followed by Dunnett tests or two-sample *t*-tests were calculated in R or GraphPad Prism to compute *P*-values, as indicated (R Core Team 2021).

Behavior assays

For each assay, five 1-day-old hermaphrodites were transferred from 60 mm NGM plates with OP50 to 60 mm chemotaxis plates (CTX) (5 mM KH₂PO₄/K₂HPO₄ [pH 6.0], 1 mM CaCl₂, 1 mM MgSO₄, 2% agar) without food and left to roam for 2 min to adjust to the new plate. The plate was then placed on the Wormlab Imaging System (MBF Bioscience) and animal movements were recorded for 5 min (Roussel et al. 2014). Tracks were manually controlled for head and tail orientation blinded to genotype. Animals that showed no movement for 5 min or that left the field of view within 1 min were excluded from the analysis. Absolute speed (µm/s), total number of turns, and total number of reversals were

captured using the Wormlab imaging system and exported to GraphPad Prism for statistical analysis. One-way ANOVA followed by a Dunnett test was used to calculate *P*-values.

mRNA sequencing

Total RNA from gravid adults was DNase-treated with Turbo DNase (ThermoFisher AM2238) and depleted of rRNA using Ribo-Zero (Illumina 20020596). RNA-seq libraries were prepared with the Illumina TruSeq Stranded Total RNA kit (Illumina 20020596). Samples were sequenced on an Illumina NovaSeq X Plus (PE150) by Novogene.

mRNA data analysis

Adapter trimming and quality filtering were done with fastp (Chen et al. 2018). Transcript quantification was done with RSEM, using STAR for read mapping (Li and Dewey 2011; Dobin et al. 2013). DESeq2 was used for data normalization and statistical analysis (Love et al. 2014). R was used for plotting (R Core Team 2021). Data analysis and plotting scripts are available as [Supplemental Code](#). GO term analysis was done using the DAVID web server (Huang da et al. 2009; Sherman et al. 2022).

Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE254442.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

Thanks to Maritza Soto-Ojeda and Reese Sprister for help with media and solutions, and Josh Svendsen and Kristen Brown for help with sRNA-seq libraries. Thanks also to Dustin Updike for the use of his laboratory while we were analyzing data. Strains were provided by the CGC, which is funded by the National Institutes of Health (P40 OD010440). This work was supported by the National Institutes of Health (R35GM119775 to T.A.M. and R01NS115947 to F.J.H.).

References

- Aalto AP, Nicastro IA, Broughton JP, Chipman LB, Schreiner WP, Chen JS, Pasquini AE. 2018. Opposing roles of microRNA Argonautes during *Caenorhabditis elegans* aging. *PLoS Genet* **14**: e1007379. doi:10.1371/journal.pgen.1007379
- Almeida MV, de Jesus Domingues AM, Lukas H, Mendez-Lago M, Ketting RF. 2019. RppH can faithfully replace TAP to allow cloning of 5'-triphosphate carrying small RNAs. *MethodsX* **6**: 265–272. doi:10.1016/j.mex.2019.01.009
- Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D. 2003. MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr Biol* **13**: 807–818. doi:10.1016/S0960-9822(03)00287-2
- Aoki K, Moriguchi H, Yoshioka T, Okawa K, Tabara H. 2007. In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. *EMBO J* **26**: 5007–5019. doi:10.1038/sj.emboj.7601910
- Ashe A, Béliard T, Le Pen J, Sarkies P, Frézal L, Lehrbach NJ, Félix MA, Miska EA. 2013. A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *eLife* **2**: e00994. doi:10.7554/eLife.00994
- Bagijn MP, Goldstein LD, Sapetschnig A, Weick EM, Bouasker S, Lehrbach NJ, Simard MJ, Miska EA. 2012. Function, targets, and evolution of

- Caenorhabditis elegans* piRNAs. *Science* **337**: 574–578. doi:10.1126/science.1220952
- Batista PJ, Ruby JG, Claycomb JM, Chiang R, Fahlgren N, Kasschau KD, Chaves DA, Gu W, Vasale JJ, Duan S, et al. 2008. PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol Cell* **31**: 67–78. doi:10.1016/j.molcel.2008.06.002
- Beanan MJ, Strome S. 1992. Characterization of a germ-line proliferation mutation in *C. elegans*. *Development* **116**: 755–766. doi:10.1242/dev.116.3.755
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**: 363–366. doi:10.1038/35053110
- Blumenfeld AL, Jose AM. 2016. Reproducible features of small RNAs in *C. elegans* reveal NU RNAs and provide insights into 22G RNAs and 26G RNAs. *RNA* **22**: 184–192. doi:10.1261/rna.054551.115
- Brown KC, Svendsen JM, Tucci RM, Montgomery BE, Montgomery TA. 2017. ALG-5 is a miRNA-associated Argonaute required for proper developmental timing in the *Caenorhabditis elegans* germline. *Nucleic Acids Res* **45**: 9093–9107. doi:10.1093/nar/gkx536
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890. doi:10.1093/bioinformatics/bty560
- Chung WJ, Okamura K, Martin R, Lai EC. 2008. Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. *Curr Biol* **18**: 795–802. doi:10.1016/j.cub.2008.05.006
- Claycomb JM. 2014. Ancient endo-siRNA pathways reveal new tricks. *Curr Biol* **24**: R703–R715. doi:10.1016/j.cub.2014.06.009
- Cogoni C, Macino G. 1999. Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature* **399**: 166–169. doi:10.1038/20215
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819–823. doi:10.1126/science.1231143
- Conine CC, Batista PJ, Gu W, Claycomb JM, Chaves DA, Shirayama M, Mello CC. 2010. Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in *Caenorhabditis elegans*. *Proc Natl Acad Sci* **107**: 3588–3593. doi:10.1073/pnas.0911685107
- Corrêa RL, Steiner FA, Berezikov E, Ketting RF. 2010. MicroRNA-directed siRNA biogenesis in *Caenorhabditis elegans*. *PLoS Genet* **6**: e1000903. doi:10.1371/journal.pgen.1000903
- Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, et al. 2008. An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453**: 798–802. doi:10.1038/nature07007
- Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC. 2000. 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. doi:10.1016/S0092-8674(00)80864-8
- Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, et al. 2008. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol Cell* **31**: 79–90. doi:10.1016/j.molcel.2008.06.003
- Davis P, Zarowiecki M, Arnaboldi V, Becerra A, Cain S, Chan J, Chen WJ, Cho J, da Veiga Beltrame E, Diamantakis S, et al. 2022. Wormbase in 2022—data, processes, and tools for analyzing *Caenorhabditis elegans*. *Genetics* **220**: iyac003. doi:10.1093/genetics/iyac003
- Deshpande G, Calhoun G, Schedl P. 2005. *Drosophila argonaute-2* is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev* **19**: 1680–1685. doi:10.1101/gad.1316805
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21. doi:10.1093/bioinformatics/bts635
- Elbashir SM, Lendeckel W, Tuschl T. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**: 188–200. doi:10.1101/gad.862301
- Félix MA, Ashe A, Piffaretti J, Wu G, Nuez I, Béliard T, Jiang Y, Zhao G, Franz CJ, Goldstein LD, et al. 2011. Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol* **9**: e1000586. doi:10.1371/journal.pbio.1000586
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806–811. doi:10.1038/35888
- Fischer SEJ, Ruvkun G. 2020. *Caenorhabditis elegans* ADAR editing and the ERI-6/7/MOV10 RNAi pathway silence endogenous viral elements and LTR retrotransposons. *Proc Natl Acad Sci* **117**: 5987–5996. doi:10.1073/pnas.1919028117

- Fischer SE, Butler MD, Pan Q, Ruvkun G. 2008. Trans-splicing in *C. elegans* generates the negative RNAi regulator ERI-6/7. *Nature* **455**: 491–496. doi:10.1038/nature07274
- Fischer SE, Montgomery TA, Zhang C, Fahlgren N, Breen PC, Hwang A, Sullivan CM, Carrington JC, Ruvkun G. 2011. The ERI-6/7 helicase acts at the first stage of an siRNA amplification pathway that targets recent gene duplications. *PLoS Genet* **7**: e1002369. doi:10.1371/journal.pgen.1002369
- Gasiunas G, Barrangou R, Horvath P, Siksnys V. 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci* **109**: E2579–E2586. doi:10.1073/pnas.1208507109
- Gent JI, Schvarzstein M, Villeneuve AM, Gu SG, Jantsch V, Fire AZ, Baudrimont A. 2009. A *Caenorhabditis elegans* RNA-directed RNA polymerase in sperm development and endogenous RNA interference. *Genetics* **183**: 1297–1314. doi:10.1534/genetics.109.109686
- Gent JI, Lamm AT, Pavelec DM, Maniar JM, Parameswaran P, Tao L, Kennedy S, Fire AZ. 2010. Distinct phases of siRNA synthesis in an endogenous RNAi pathway in *C. elegans* soma. *Mol Cell* **37**: 679–689. doi:10.1016/j.molcel.2010.01.012
- Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, et al. 2008. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320**: 1077–1081. doi:10.1126/science.1157396
- Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha J, Baillie DL, Fire A, Ruvkun G, Mello CC. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**: 23–34. doi:10.1016/S0092-8674(01)00431-7
- Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL. 2008. The Vienna RNA Websuite. *Nucleic Acids Res* **36**: W70–W74. doi:10.1093/nar/gkn188
- Gu W, Shirayama M, Conte D, Vasale J, Batista PJ, Claycomb JM, Moresco JJ, Youngman EM, Keys J, Stoltz MJ, et al. 2009. Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol Cell* **36**: 231–244. doi:10.1016/j.molcel.2009.09.020
- Hamilton AJ, Baulcombe DC. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**: 950–952. doi:10.1126/science.286.5441.950
- Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**: 293–296. doi:10.1038/35005107
- Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. 2001. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**: 1146–1150. doi:10.1126/science.1064023
- Han T, Manoharan AP, Harkins TT, Bouffard P, Fitzpatrick C, Chu DS, Thierry-Mieg D, Thierry-Mieg J, Kim JK. 2009. 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in *Caenorhabditis elegans*. *Proc Natl Acad Sci* **106**: 18674–18679. doi:10.1073/pnas.0906378106
- Heale BS, Keegan LP, McGurk L, Michlewski G, Brindle J, Stanton CM, Caceres JF, O'Connell MA. 2009. Editing independent effects of ADARs on the miRNA/siRNA pathways. *EMBO J* **28**: 3145–3156. doi:10.1038/emboj.2009.244
- Hellwig S, Bass BL. 2008. A starvation-induced noncoding RNA modulates expression of Dicer-regulated genes. *Proc Natl Acad Sci* **105**: 12897–12902. doi:10.1073/pnas.0805118105
- Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57. doi:10.1038/nprot.2008.211
- Hunter JD. 2007. Matplotlib: a 2D graphics environment. *Comput Sci Eng* **9**: 90–95. doi:10.1109/MCSE.2007.55
- Hutvagner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, Zamore PD. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small temporal RNA. *Science* **293**: 834–838. doi:10.1126/science.1062961
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816–821. doi:10.1126/science.1225829
- Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot N, Moreno S, Sohrmann M, et al. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**: 231–237. doi:10.1038/nature01278
- Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of *Arabidopsis* siRNAs. *PLoS Biol* **5**: e57. doi:10.1371/journal.pbio.0050057
- Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H. 2008. *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453**: 793–797. doi:10.1038/nature06938
- Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH. 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* **15**: 2654–2659. doi:10.1101/gad.927801
- Knight SW, Bass BL. 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**: 2269–2271. doi:10.1126/science.1062039
- Knight SW, Bass BL. 2002. The role of RNA editing by ADARs in RNAi. *Mol Cell* **10**: 809–817. doi:10.1016/S1097-2765(02)00649-4
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25. doi:10.1186/gb-2009-10-3-r25
- Lee HC, Gu W, Shirayama M, Youngman E, Conte D Jr, Mello CC. 2012. *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* **150**: 78–87. doi:10.1016/j.cell.2012.06.016
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323. doi:10.1186/1471-2105-12-323
- Llave C, Kasschau KD, Rector MA, Carrington JC. 2002. Endogenous and silencing-associated small RNAs in plants. *Plant Cell* **14**: 1605–1619. doi:10.1105/tpc.003210
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550. doi:10.1186/s13059-014-0550-8
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* **339**: 823–826. doi:10.1126/science.1232033
- Montgomery BE, Knittel TL, Reed KJ, Chong MC, Isolehto IJ, Cafferty ER, Smith MJ, Sprister RA, Magelky CN, Scherman H, et al. 2024. Regulation of Microprocessor assembly and localization via Pasha's WW domain in *C. elegans*. bioRxiv doi:10.1101/2024.04.23.590772
- Mourrain P, Béclin C, Elmayer T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, et al. 2000. *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101**: 533–542. doi:10.1016/S0092-8674(00)80863-6
- Ohta H, Fujiwara M, Ohshima Y, Ishihara T. 2008. ADBP-1 regulates an ADAR RNA-editing enzyme to antagonize RNA-interference-mediated gene silencing in *Caenorhabditis elegans*. *Genetics* **180**: 785–796. doi:10.1534/genetics.108.093310
- Okamura K, Ishizuka A, Siomi H, Siomi MC. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* **18**: 1655–1666. doi:10.1101/gad.1210204
- Okamura K, Balla S, Martin R, Liu N, Lai EC. 2008a. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. *Nat Struct Mol Biol* **15**: 998. doi:10.1038/nmsb0908-998c
- Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, Lai EC. 2008b. The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**: 803–806. doi:10.1038/nature07015
- Pak J, Fire A. 2007. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **315**: 241–244. doi:10.1126/science.1132839
- Phillips CM, Montgomery BE, Breen PC, Roovers EF, Rim YS, Ohsumi TK, Newman MA, van Wolfswinkel JC, Ketting RF, Ruvkun G, et al. 2014. MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Curr Biol* **24**: 839–844. doi:10.1016/j.cub.2014.02.060
- R Core Team. 2021. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.
- Reed KJ, Svendsen JM, Brown KC, Montgomery BE, Marks TN, Vijayarath T, Parker DM, Nishimura EO, Updike DL, Montgomery TA. 2020. Widespread roles for piRNAs and WAGO-class siRNAs in shaping the germline transcriptome of *Caenorhabditis elegans*. *Nucleic Acids Res* **48**: 1811–1827. doi:10.1093/nar/gkz1178
- Reich DP, Tyc KM, Bass BL. 2018. *C. elegans* ADARs antagonize silencing of cellular dsRNAs by the antiviral RNAi pathway. *Genes Dev* **32**: 271–282. doi:10.1101/gad.310672.117
- Robertson HD, Webster RE, Zinder ND. 1968. Purification and properties of ribonuclease III from *Escherichia coli*. *J Biol Chem* **243**: 82–91. doi:10.1016/S0021-9258(18)99327-0
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26. doi:10.1038/nbt.1754
- Roussel N, Sprenger J, Tappan SJ, Glaser JR. 2014. Robust tracking and quantification of *C. elegans* body shape and locomotion through coiling, entanglement, and omega bends. *Worm* **3**: e982437. doi:10.4161/21624054.2014.982437
- Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP. 2006. Large-scale sequencing reveals 21U-RNAs and additional

- microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**: 1193–1207. doi:10.1016/j.cell.2006.10.040
- Rybak-Wolf A, Jens M, Murakawa Y, Herzog M, Landthaler M, Rajewsky N. 2014. A variety of dicer substrates in human and *C. elegans*. *Cell* **159**: 1153–1167. doi:10.1016/j.cell.2014.10.040
- Scadden AD, Smith CW. 2001. RNAi is antagonized by A→I hyper-editing. *EMBO Rep* **2**: 1107–1111. doi:10.1093/embo-reports/kve244
- Schott DH, Cureton DK, Whelan SP, Hunter CP. 2005. An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. *Proc Natl Acad Sci* **102**: 18420–18424. doi:10.1073/pnas.0507123102
- Sebastiani P, Montano M, Puca A, Solovieff N, Kojima T, Wang MC, Melista E, Meltzer M, Fischer SE, Andersen S, et al. 2009. RNA editing genes associated with extreme old age in humans and with lifespan in *C. elegans*. *PLoS One* **4**: e8210. doi:10.1371/journal.pone.0008210
- Seitz H, Ghildiyal M, Zamore PD. 2008. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. *Curr Biol* **18**: 147–151. doi:10.1016/j.cub.2007.12.049
- Seroussi U, Lugowski A, Wadi L, Lao RX, Willis AR, Zhao W, Sundby AE, Charlesworth AG, Reinke AW, Claycomb JM. 2023. A comprehensive survey of *C. elegans* argonaute proteins reveals organism-wide gene regulatory networks and functions. *eLife* **12**: e83853. doi:10.7554/eLife.83853
- Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W. 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* **50**: W216–W221. doi:10.1093/nar/gkac194
- Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**: 465–476. doi:10.1016/S0092-8674(01)00576-1
- Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM. 2000. 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. doi:10.1016/S0960-9822(00)00323-7
- Steiner FA, Okihara KL, Hoogstrate SW, Sijen T, Ketting RF. 2009. RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in *Caenorhabditis elegans*. *Nat Struct Mol Biol* **16**: 207–211. doi:10.1038/nsmb.1541
- Svendsen JM, Reed KJ, Vijayarathay T, Montgomery BE, Tucci RM, Brown KC, Marks TN, Nguyen DAH, Phillips CM, Montgomery TA. 2019. *hem-1/HEN1* promotes germline immortality in *Caenorhabditis elegans*. *Cell Rep* **29**: 3187–3199.e4. doi:10.1016/j.celrep.2019.10.114
- Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC. 1999. The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99**: 123–132. doi:10.1016/S0092-8674(00)81644-X
- Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM, et al. 2008. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**: 534–538. doi:10.1038/nature06904
- Tate AJ, Brown KC, Montgomery TA. 2023. tiny-count: a counting tool for hierarchical classification and quantification of small RNA-seq reads with single-nucleotide precision. *Bioinform Adv* **3**: vbad065. doi:10.1093/bioadv/vbad065
- Thivierge C, Makil N, Flamand M, Vasale JJ, Mello CC, Wohlschlegel J, Conte D Jr, Duchaine TF. 2012. Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. *Nat Struct Mol Biol* **19**: 90–97. doi:10.1038/nsmb.2186
- Tonkin LA, Bass BL. 2003. Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. *Science* **302**: 1725. doi:10.1126/science.1091340
- Vasale JJ, Gu W, Thivierge C, Batista PJ, Claycomb JM, Youngman EM, Duchaine TF, Mello CC, Conte D Jr. 2010. Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci* **107**: 3582–3587. doi:10.1073/pnas.0911908107
- Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833–1837. doi:10.1126/science.1074973
- Walkley CR, Li JB. 2017. Rewriting the transcriptome: adenosine-to-inosine RNA editing by ADARs. *Genome Biol* **18**: 205. doi:10.1186/s13059-017-1347-3
- Wang G, Reinke V. 2008. A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. *Curr Biol* **18**: 861–867. doi:10.1016/j.cub.2008.05.009
- Warf MB, Shepherd BA, Johnson WE, Bass BL. 2012. Effects of ADARs on small RNA processing pathways in *C. elegans*. *Genome Res* **22**: 1488–1498. doi:10.1101/gr.134841.111
- Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, et al. 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**: 539–543. doi:10.1038/nature06908
- Welker NC, Pavelec DM, Nix DA, Duchaine TF, Kennedy S, Bass BL. 2010. Dicer's helicase domain is required for accumulation of some, but not all, *C. elegans* endogenous siRNAs. *RNA* **16**: 893–903. doi:10.1261/rna.2122010
- Welker NC, Maity TS, Ye X, Aruscavage PJ, Krauchuk AA, Liu Q, Bass BL. 2011. Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. *Mol Cell* **41**: 589–599. doi:10.1016/j.molcel.2011.02.005
- Wen J, Duan H, Bejarano F, Okamura K, Fabian L, Brill JA, Bortolamiol-Becet D, Martin R, Ruby JG, Lai EC. 2015. Adaptive regulation of testis gene expression and control of male fertility by the *Drosophila* hairpin RNA pathway. *Mol Cell* **57**: 165–178. doi:10.1016/j.molcel.2014.11.025
- Wilkins C, Dishongh R, Moore SC, Whitt MA, Chow M, Machaca K. 2005. RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. *Nature* **436**: 1044–1047. doi:10.1038/nature03957
- Wu D, Lamm AT, Fire AZ. 2011. Competition between ADAR and RNAi pathways for an extensive class of RNA targets. *Nat Struct Mol Biol* **18**: 1094–1101. doi:10.1038/nsmb.2129
- Xu K, Bogert BA, Li W, Su K, Lee A, Gao FB. 2004. The fragile X-related gene affects the crawling behavior of *Drosophila* larvae by regulating the mRNA level of the DEG/ENaC protein pickpocket1. *Curr Biol* **14**: 1025–1034. doi:10.1016/j.cub.2004.05.055
- Yang W, Wang Q, Howell KL, Lee JT, Cho DS, Murray JM, Nishikura K. 2005. ADAR1 RNA deaminase limits short interfering RNA efficacy in mammalian cells. *J Biol Chem* **280**: 3946–3953. doi:10.1074/jbc.M407876200
- Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC. 2006. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**: 747–757. doi:10.1016/j.cell.2006.09.033
- Yu B, Chen X. 2010. Analysis of miRNA modifications. *Methods Mol Biol* **592**: 137–148. doi:10.1007/978-1-60327-005-2_10
- Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**: 25–33. doi:10.1016/S0092-8674(00)80620-0
- Zhang C, Montgomery TA, Fischer SE, Garcia SM, Riedel CG, Fahlgren N, Sullivan CM, Carrington JC, Ruvkun G. 2012. The *Caenorhabditis elegans* RDE-10/RDE-11 complex regulates RNAi by promoting secondary siRNA amplification. *Curr Biol* **22**: 881–890.

Received February 8, 2024; accepted in revised form October 3, 2024.