SUPPLEMENTAL MATERIALS

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[3], Anastasia S. Nawrocki[3], Frederic J. Hoerndli[2,3], and Taiowa A. Montgomery[1,2]*

[1]Department of Biology, Colorado State University, Fort Collins, CO 80523, USA

[2]Cell and Molecular Biology Program, Colorado State University, Fort Collins, CO 80523, USA

[3]Department of Biomedical Sciences, Colorado State University, Fort Collins, Colorado 80523, USA

[4]Equal contribution

*Lead Contact: tai.montgomery@colostate.edu

SUPPLEMENTAL METHODS

Strain Generation

The TAM134[F43E2.6*(ram37)* II] and TAM151[*rde-1(ram40)* V] strains were generated from wild type (N2) *C. elegans* using CRISRP-Cas9-mediated genome editing with purified Cas9 protein and synthesized guide RNAs (TAM134: CCUCCAACGAUCCACCACAG and AAUUCUUGAUAAAGUCCAGG; TAM151: GGACAUGUUUCAUCACUUUG and UGGAAUGAAUUUUUGAACC) (Integrated DNA Technologies). TAM134[F43E2.6*(ram37)* II] contains a substitution of a 2,483 bp sequence from positions -819 to 1,664, relative to the F43E2.6 start codon, with the sequence TTA. TAM151[*rde-1(ram40)* V] contains a 3,674 bp substitution from positions -10 to 3,664, relative to *rde-1* start codon, with the sequence AATCAAAAAACAAAAATCAAAAATTCATTC.

Protein-small RNA co-IPs

~11,000-20,000 animals for each replicates for each of the Argonaute co-IPs were flash-frozen in liquid nitrogen and ground in lysis buffer (50 mM Tris-Cl, pH 7.4, 100 mM KCl, 2.5 mM $MgCl_2$, 0.1% Igepal CA-630, 0.5 mM PMSF, and 1×Pierce Protease Inhibitor Tablets [Pierce Biotechnology, cat# 88266]). Cell lysates were cleared by centrifugation for 10 min at 12,000 RCF. Cleared lysates were split into cell lysate (input) and co-IP fractions. The co-IP fractions were incubated with GFP-Trap Magnetic Agarose Beads (Proteintech, cat# gtma-100) for 1 hour. Beads were washed in lysis buffer 3 times and split into RNA and protein fractions. Protein fractions were incubated at 95˚C for 5 min in 1× Blue Protein Loading Dye (New England Biolabs, cat# B7703S). RNA was purified using the RNA isolation method described in the main text.

Small RNA Sequencing

~16-30-nt RNA was size selected on 17% polyacrylamide/urea gels from untreated RNA or RNA pre-treated with RNA 5' polyphosphatase (Illumina, cat# RP8092H) or RNA 5' pyrophosphohydrolase (New England Biolabs, cat# M0356S) to reduce 5’ triphosphates to 5’ monophosphates to facilitate 22G-RNA ligation

(Almeida et al. 2019). Small RNA sequencing libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina following the manufacturer's recommendations with the 3' ligation step changed to 16˚C for 18 hours to improve capture of methylated small RNAs (New England Biolabs, cat# E7300S). Small RNA PCR amplicons were size selected on 10% polyacrylamide non-denaturing gels. Samples were sequenced on an Illumina HiSeq X or NovaSeq X Plus (PE150) by Novogene. Only the forward strand reads were analyzed.

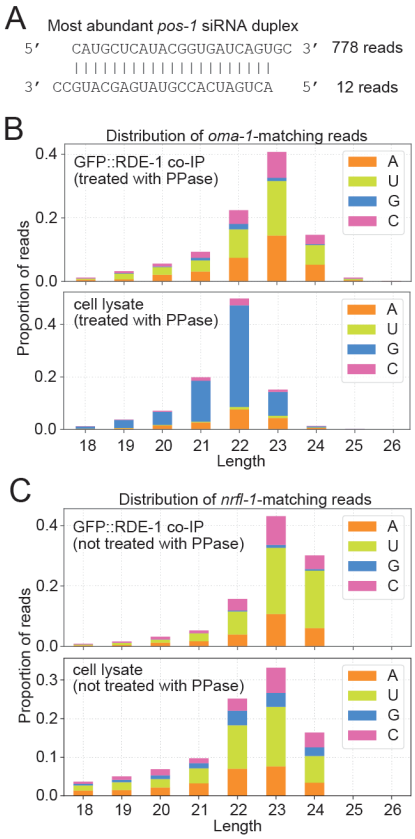
qRT-PCR

qRT-PCR of small RNAs was done with TaqMan reagents following the manufacturer's recommendations (Life Technologies, cat# 4331348). F43E2.6 siRNA and miR-1 cycle threshold (Ct) values were assayed using a CF96 Real-Time PCR Detection System (Bio-Rad). Ct values were averaged across three technical replicates for each of 3 biological replicates. Mean F43E2.6 siRNAs levels relative to miR-1 were calculated using the 2-ddCt method (Livak and Schmittgen 2001). Two-sample *t*-tests were used to compare differences between conditions. Microsoft Excel and GraphPad Prism were used for plotting and statistics.

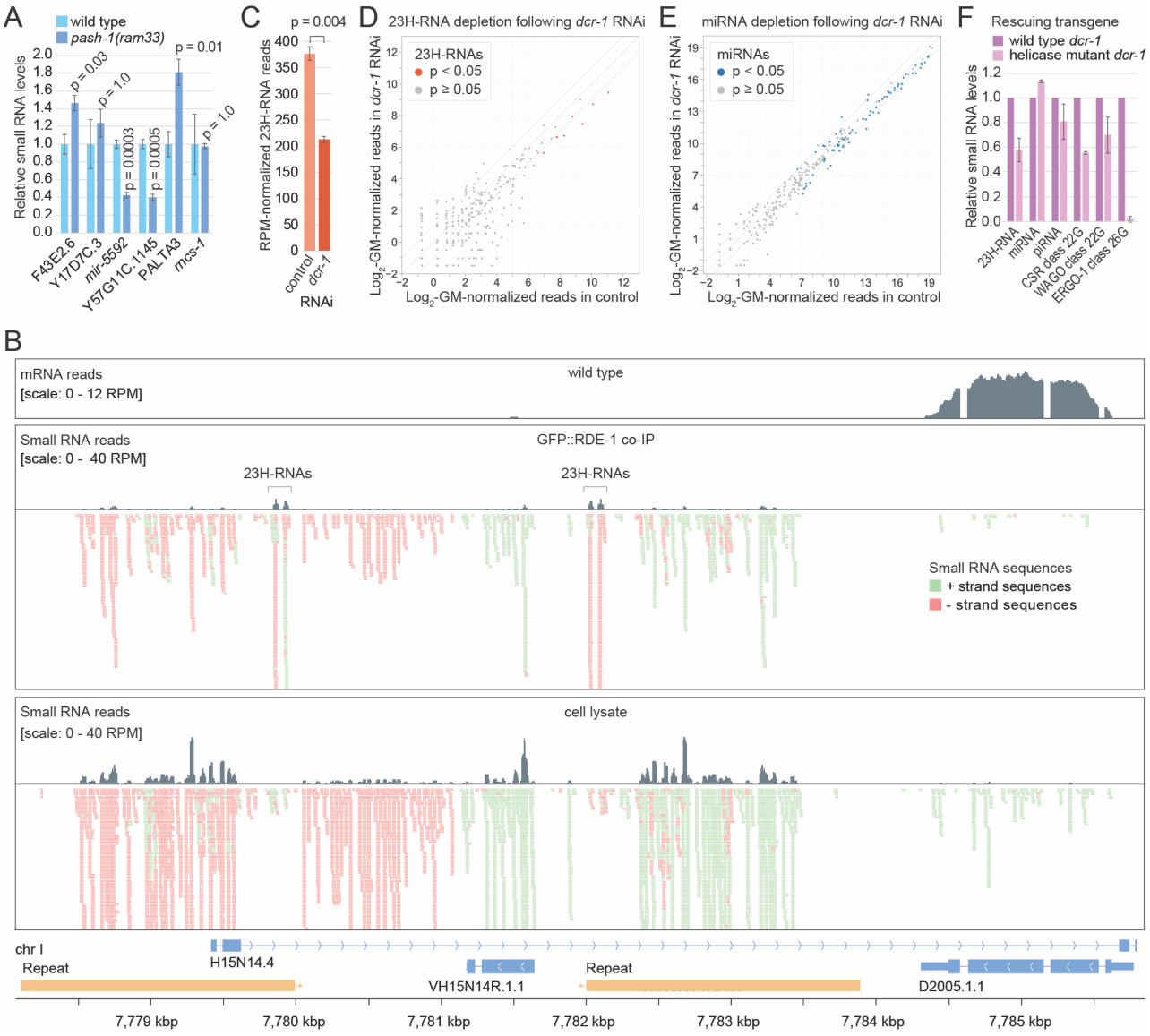
Brood size assays

To assess general fertility, the parental generation of the animals in which fertility was assayed was grown at 20°C until the L4 stage and then transferred to the experimental temperature (20˚C or 25˚C). Individual animals were grown on OP50 in individual wells of 24 well plates starting at the L4 stage and transferred to new wells every 24 hours. The number of progeny in each well was counted the following day. The process was repeated until the cessation of egg laying. To assess fertility for the RNAi efficacy assay, 20 animals per replicate were treated with *hmr-1* RNAi from L1 larval stage until day 2 of adulthood. The parental animals were then removed and the numbers of progeny were assessed by washing them off each plate ~24 hours later into 2 ml of M9 and then averaging the number of larvae in 5×10 μl aliquots and extrapolating the total. 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.

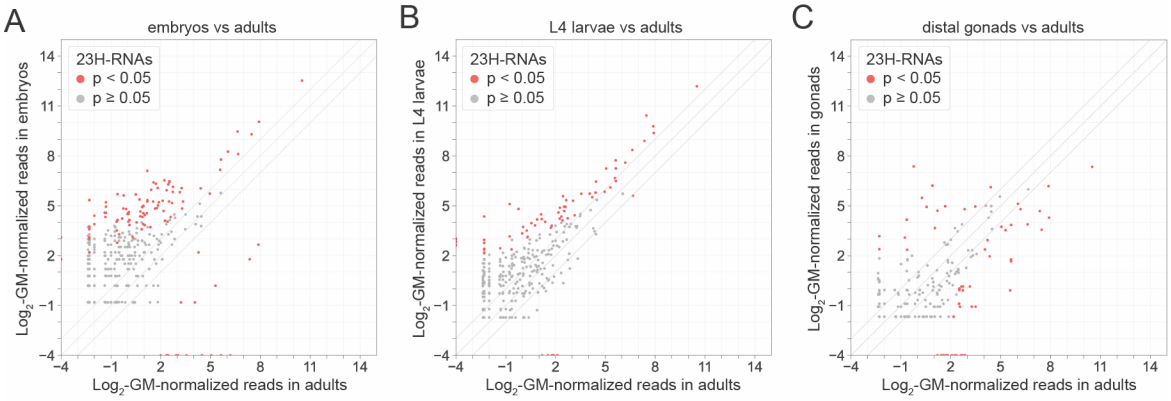
SUPPLEMENTAL FIGURES



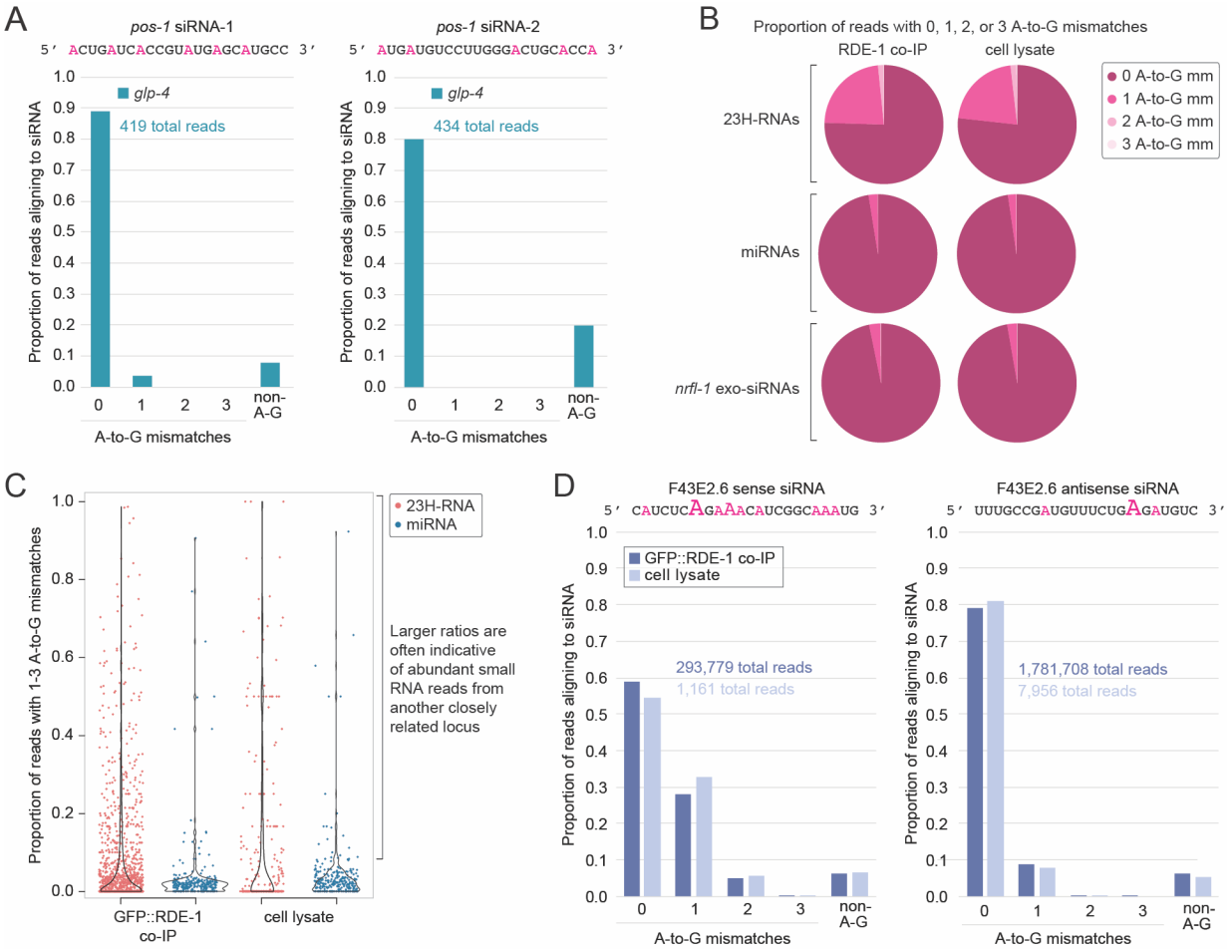
Supplement Figure S1. RDE-1 association with exogenous siRNAs. (A) The most abundant *pos-1*-matching siRNA duplex in sRNA-seq libraries. n=1 biological replicate. Libraries are from adults treated with *pos-1* RNAi and grown at 25˚C to induce sterility. (B) Size and 5'-nt distribution of *oma-1*-matching reads from GFP::RDE-1 co-IP and cell lysate sRNA-seq libraries. n=2 biological replicates. Data from one representative library shown. Libraries from gravid adults treated with *nrfl-1* and *oma-1* RNAi. RNA was treated with RNA 5' polyphosphatase (PPase) to capture 22G-RNAs prior to library preparation. (C) 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 shown. Gravid adults treated with *nrfl-1* RNAi. RNA was not treated with PPase prior to library preparation.



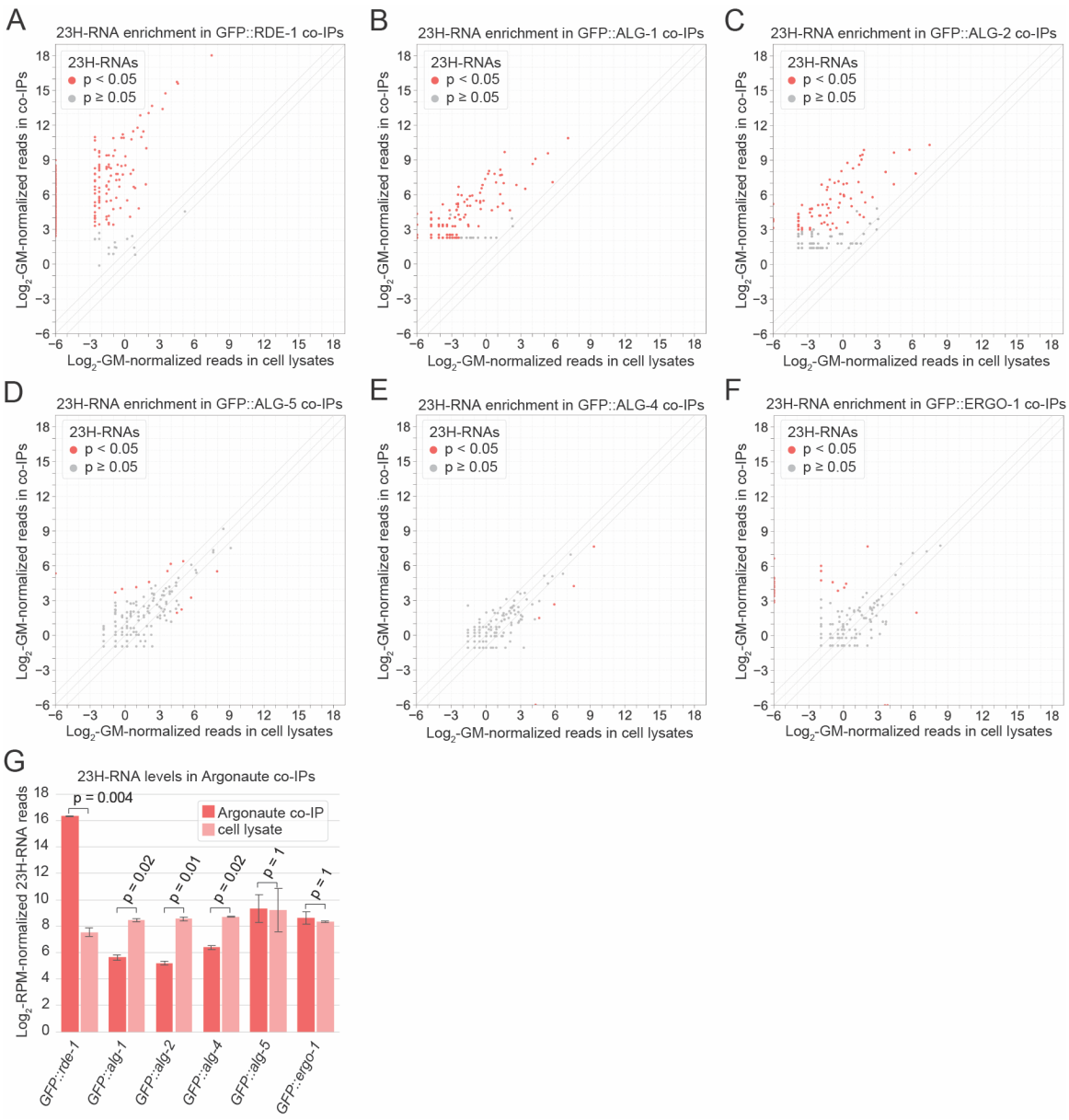
Supplemental Figure S2. 23H-RNA dependence on Dicer and Pasha. (A) Relative RPM-normalized levels of the indicated small RNAs in wild type and *pash-1(ram33)* mutants as measured by sRNA-seq. Error bars are SD between three biological replicates. The *P*-values were calculated using two-sample *t*-tests. A Bonferroni correction was applied to account for multiple comparisons. Gravid adults. (B) mRNA-seq read density from wild type animals and sRNA-seq read density from GFP::RDE-1 co-IPs and cell lysates plotted across an inverted repeat locus. Brackets indicate 23H-RNA clusters. The lower track shows individual sequences but is truncated at the end of the box due to space constraints. One of 2 biological replicates is shown. Gravid adults. (C) RPM-normalized 23H-RNA reads from sRNA-seq libraries wild type animals after control (empty L4440 vector) or *dcr-1* RNAi. Gravid adults. Error bars are SD between two biological replicates. The *P*-value was calculated using a two-sample *t*-test. (D-E) Scatter plots display individual 23H-RNAs (D) or miRNAs (E) as the average $\log_2$-geometric mean-normalized reads from sRNA-seq libraries from adult animals treated with control RNAi (L4440) (x-axis) and *dcr-1* RNAi (y-axis). n=3 biological replicates for each condition. (F) Relative levels of the various classes of small RNAs in *dcr-1* mutants rescued with wild type *dcr-1* or a *dcr-1* helicase mutant as determined by sRNA-seq. n=1 for wild type *dcr-1* rescue and n=2 for *dcr-1* helicase mutant rescue. Error bars for rescue with the *dcr-1* helicase mutant are SD of two biological replicates. Mixed stage animals.



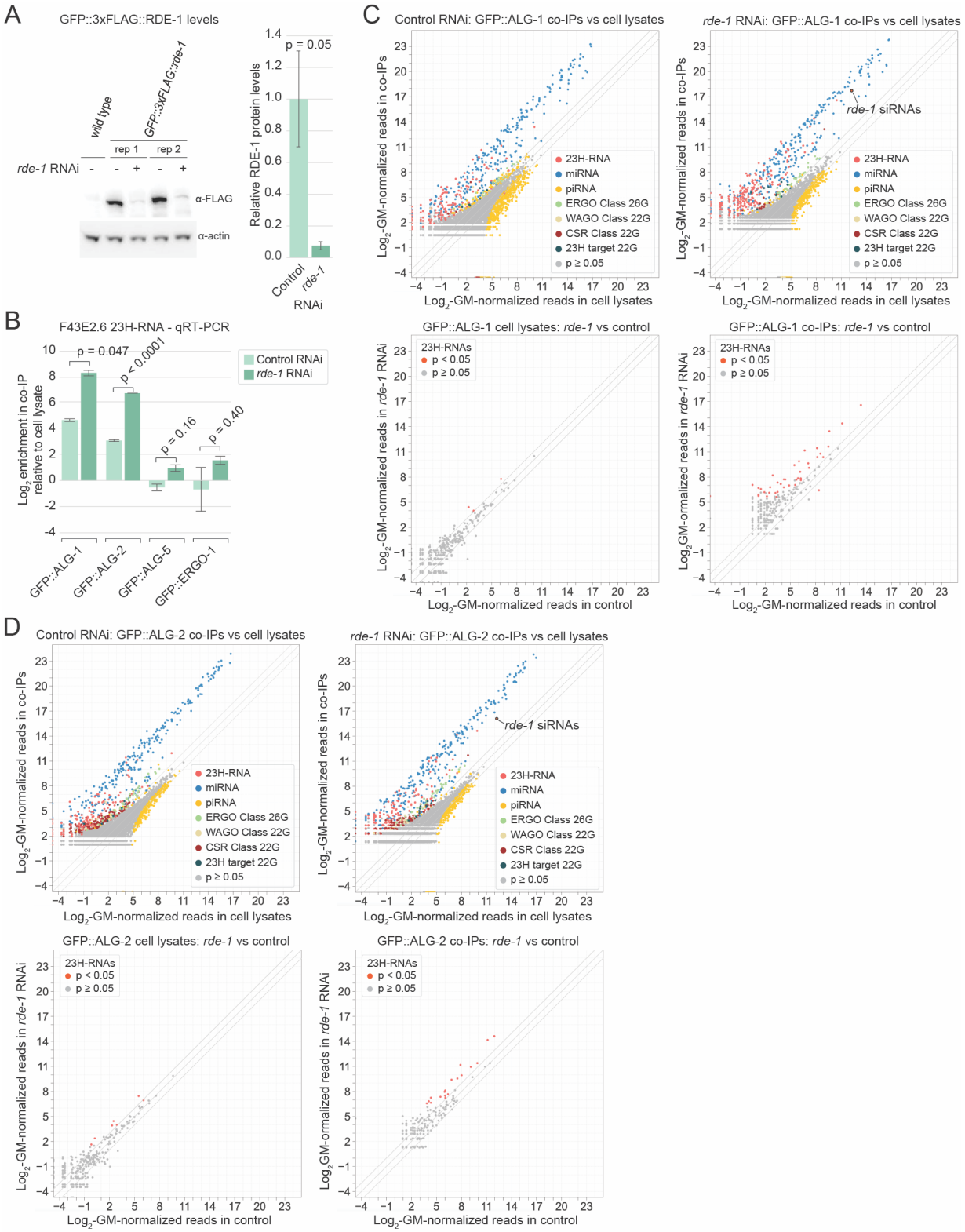
Supplemental Figure S3. 23H-RNA expression across development. (A) Scatter plots display individual 23H-RNAs as the average log_2-geometric mean-normalized sRNA-seq reads from gravid adults on the x-axis and embryos (D), L4 stage larvae (E), or dissected distal gonads (F) on the y-axis. n=3 biological replicates for each developmental stage except embryos for which n=2. Samples are from different experiments.



Supplemental Figure S4. RNA-editing signatures in exogenous siRNAs and endogenous 23H-RNAs. (A) The proportion of sRNA-seq reads from two different *pos-1* siRNAs with 0-3 A-to-G and total (1-3) non-A-to-G mismatches to the genome in *glp-4* mutants undergoing *pos-1* RNAi. Adenosines (As) are colored magenta and their heights are roughly proportional the frequency of A-to-G editing. n=1 biological replicate. Adult stage animals grown at 25˚C to induce sterility. (B) The proportions of 0, 1, 2, and 3 A-to-G genomic mispairs in 23H-RNA, miRNA, and *nrfl-1* exogenous siRNA reads in sRNA-seq libraries from GFP::RDE-1 co-IPs and cell lysates. One of 2 biological replicates is shown. Gravid adults. (C) Each data point in the violin plot shows the proportion of reads for a 23H-RNA or miRNA with 1-3 A-to-G genomic mispairs in sRNA-seq libraries from GFP::RDE-1 co-IPs and cell lysates. Libraries are the same as in (B). (D) The proportion of reads from the most abundant F43E2.6 23H-RNA duplex sense and antisense siRNAs with 0-3 A-to-G and total 1-3 non-A-to-G mismatches to the genome in sRNA-seq libraries from GFP::RDE-1 co-IPs and cell lysates. Adenosines (As) are colored magenta and their heights are roughly proportional to the frequency of A-to-G editing. One of 2 biological replicates is shown. Libraries are the same as in (B).

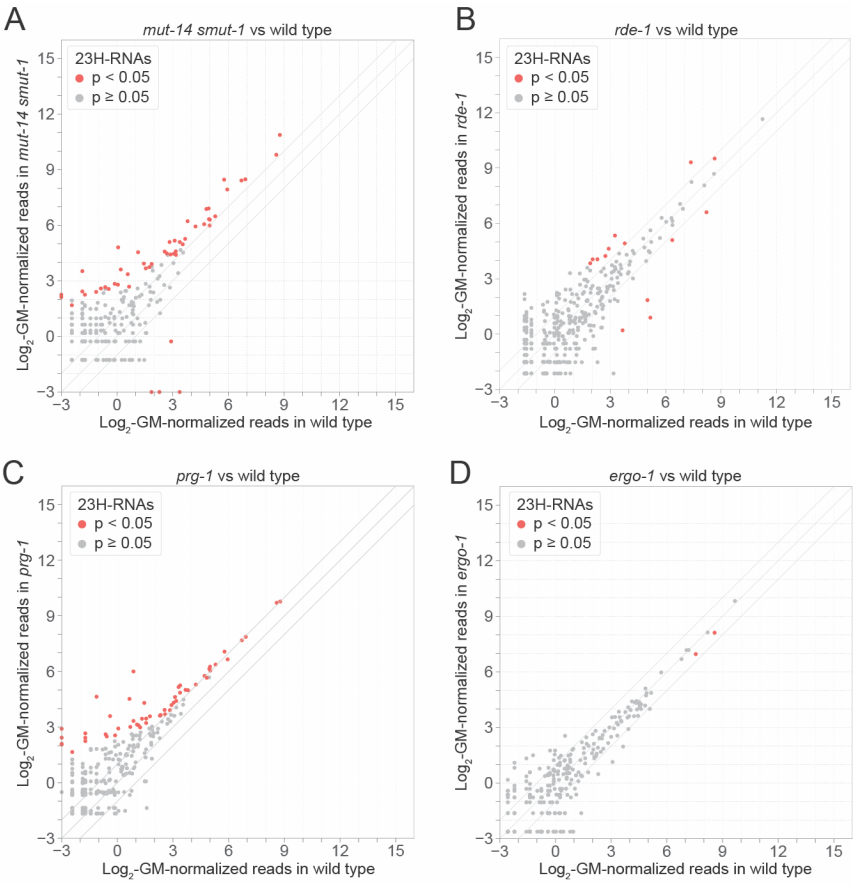


Supplemental Figure S5. 23H-RNA enrichment or depletion in Argonaute co-IPs. (A-F) Scatter plots display individual 23H-RNAs as the average $\log_2$-geometric mean-normalized 23H-RNA reads from cell lysate (x-axis) and Argonaute co-IP (y-axis) sRNA-seq libraries. (A) GFP::RDE-1 co-IPs. (B) GFP::ALG-1 co-IPs. (C) GFP::ALG-2 co-IPs. (D) GFP::ALG-4 co-IPs, L4 larvae. (E) GFP::ALG-5 co-IPs. (F) GFP::ERGO-1 co-IPs. n=2 biological replicates for each condition. Adult stage animals except for GFP::ALG-4 co-IPs and cell lysates which were L4 larval stage. (G) Total $\log_2$-RPM-normalized 23H-RNA reads in the indicated Argonaute co-IP and corresponding cell lysate sRNA-seq libraries. Data from the same libraries shown in (A-F). Error bars are SD between two biological replicates. The *P*-values were calculated using two-sample *t*-tests comparing co-IPs to cell lysates. A Bonferroni correction was applied to account for multiple comparisons.

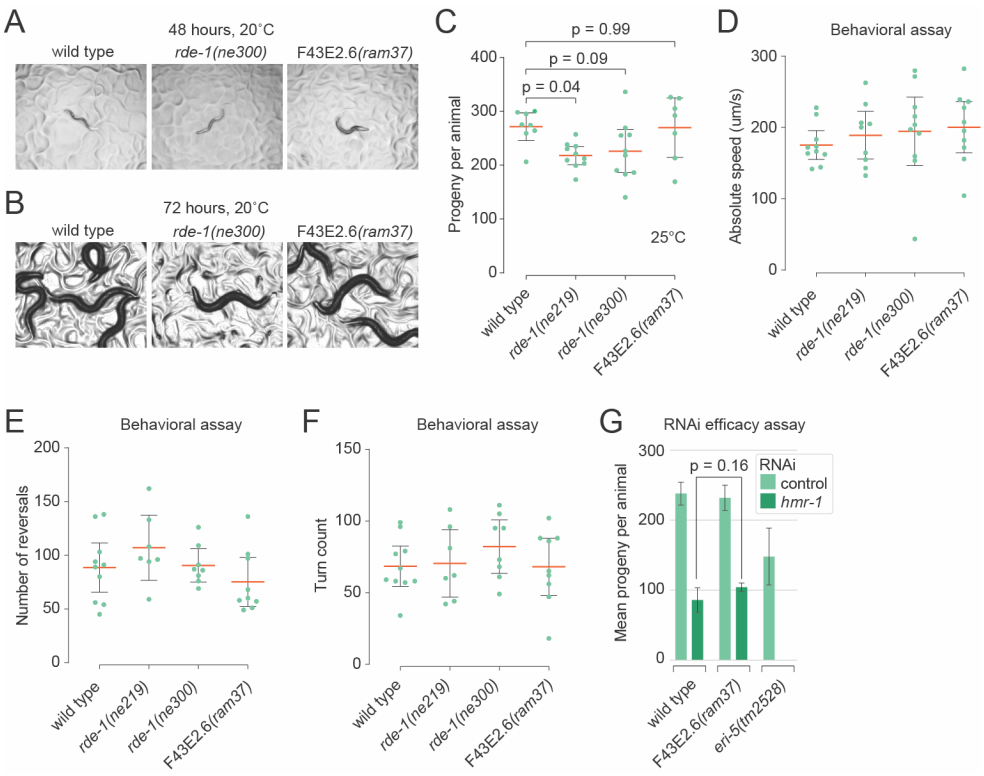


Supplemental Figure S6. Competition between Argonautes for 23H-RNAs. (A) Western Blot analysis of RDE-1 levels after control (L4440) or *rde-1* RNAi. Actin is shown as a loading control. The plot shows relative quantification of RDE-1 protein levels. Actin was used for normalization. Error bars are SD between two biological replicates. The *P*-value was calculated using a two-sample *t*-test. Gravid adults treated with control (L4440) or *rde-1* RNAi. (B) Log_2 enrichment of an abundant F43E2.6 23H-RNA in the indicated Argonaute co-IPs relative to cell lysates as determined by qRT-PCR. miR-1 levels were used for

normalization. Error bars are SD between two biological replicates. The *P*-values were calculated using two-sample *t*-tests comparing control and *rde-1* RNAi treated animals. A Bonferroni correction was applied to account for multiple comparisons. Gravid adults treated with control (L4440) or *rde-1* RNAi. (C-D) Scatter plots display individual small RNAs, colored by classification, as the average $\log_2$-geometric mean-normalized reads in sRNA-seq libraries from various combinations of GFP::ALG-1 (C) or GFP::ALG-2 (D) co-IPs and cell lysates. *rde-1*-matching siRNAs are circled. The same RNA was used for qRT-PCR in (B).



Supplemental Figure S7. Genetic requirements for 23H-RNAs. (A-D) Scatter plots display individual 23H-RNAs as the average log_2-geometric mean-normalized sRNA-seq reads from wild type on the x-axis and *mut-14(mg461) smut-1(tm1301)* (A), *rde-1(ne219)* (B), *prg-1(n4357)* (C), or *ergo-1(tm1860)* (D) on the y-axis. Gravid adults. Three biological replicates each, except *mut-14(mg461) smut-1(tm1301)* with only 2.



Supplemental Figure S8. Fertility and behavior of *rde-1* and F43E2.6 mutants. (A) Representative images of L4 stage larvae for each of the genotypes indicated. Animals were grown for 48 hours at 20˚C. (B) Representative images of adult stage animals for each of the genotypes indicated. Animals were grown for 72 hours at 20˚C. (C) Each data point is the number of progeny produced by a single animal for each of the genotypes indicated. The orange bars show the means and the error bars show the 95% confidence intervals. n=7-10 animals per genotype. One-way ANOVA followed by a Dunnett test was used for statistical analysis. Animals grown at 25°C. (D) Absolute speed (forward + reverse animal body movement). (E) Number of reversals. (F) Number of turns. For (D-F), n=9-10 animals per measurement. One of three independent experiments, each with ~10 animals per replicate, is shown. Measurements were taken over five minutes. All statistical analysis was done using one-way ANOVA followed by Dunnett tests. All *P*-values >0.4 for comparisons to wild type. (G) RNAi-efficacy assay. Mean progeny per animal for each of the genotypes indicated is shown following control (L4440) or *hmr-1* RNAi. Error bars are SD between three biological replicates. The *P*-value was calculated using a two-sample *t*-test.