Article

Distinct Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the *C. elegans* Germline

Weifeng Gu,[1,8] Masaki Shirayama,[1,8] Darryl Conte, Jr.,[1,8,*] Jessica Vasale,[1] Pedro J. Batista,[1,3] Julie M. Claycomb,[1] James J. Moresco,[4] Elaine M. Youngman,[1] Jennifer Keys,[1] Matthew J. Stoltz,[1] Chun-Chieh G. Chen,[1] Daniel A. Chaves,[1,5] Shenghua Duan,[1] Kristin D. Kasschau,[6] Noah Fahlgren,[6] John R. Yates III,[4] Shohei Mitani,[7] James C. Carrington,[6] and Craig C. Mello[1,2,*]

[1]Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA 01606, USA
[2]Howard Hughes Medical Institute, Worcester, MA 01605, USA
[3]Gulbenkian PhD Programme in Biomedicine, Rua da Quinta Grande, 6, 2780-156, Oeiras, Portugal
[4]Department of Chemical Physiology, The Scripps Research Institute, La Jolla, CA 92037, USA
[5]Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal
[6]Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR 97331, USA
[7]CREST, Japan Science and Technology Agency and Department of Physiology, Tokyo Women's Medical University School of Medicine, Tokyo 162-8666, Japan
[8]These authors contributed equally to this work
*Correspondence: darryl.conte@umassmed.edu (D.C.), craig.mello@umassmed.edu (C.C.M.)
DOI 10.1016/j.molcel.2009.09.020

SUMMARY

Endogenous small RNAs (endo-siRNAs) interact with Argonaute (AGO) proteins to mediate sequence-specific regulation of diverse biological processes. Here, we combine deep-sequencing and genetic approaches to explore the biogenesis and function of endo-siRNAs in *C. elegans*. We describe conditional alleles of the Dicer-related helicase, *drh-3*, that abrogate both RNA interference and the biogenesis of endo-siRNAs, called 22G-RNAs. DRH-3 is a core component of RNA-dependent RNA polymerase (RdRP) complexes essential for several distinct 22G-RNA systems. We show that, in the germline, one system is dependent on worm-specific AGOs, including WAGO-1, which localizes to germline nuage structures called P granules. WAGO-1 silences certain genes, transposons, pseudogenes, and cryptic loci. Finally, we demonstrate that components of the nonsense-mediated decay pathway function in at least one WAGO-mediated surveillance pathway. These findings broaden our understanding of the biogenesis and diversity of 22G-RNAs and suggest additional regulatory functions for small RNAs.

INTRODUCTION

Regulatory pathways related to RNA interference (RNAi) utilize small RNAs to guide the sequence-specific modulation of gene expression, chromatin structure, and innate immune function (Ding and Voinnet, 2007; Moazed, 2009). Small RNA classes can be distinguished based on a number of factors, including mechanism of biogenesis, mode of regulation or function, and the Argonaute (AGO) proteins with which they interact (Ghildiyal and Zamore, 2009). AGO family members are structurally related to ribonuclease (RNase) H and bind to the ends of single-stranded small RNAs, presenting the central residues for base-pairing interactions with target nucleic acids (Parker et al., 2005; Song et al., 2004).

Diverse pathways have been implicated in the biogenesis and loading of various small RNA species onto their respective AGO proteins (Siomi and Siomi, 2009). Double-stranded RNA (dsRNA), including the stem-loop precursors of miRNAs, are processed into mature 5′ monophosphorylated short RNAs by Dicer, an RNase III-like enzyme. The biogenesis of PIWI-interacting RNAs (piRNAs) is less well understood and appears to be independent of Dicer. In flies, PIWI-mediated cleavage events appear to define the 5′ ends of new piRNAs (Brennecke et al., 2007; Gunawardane et al., 2007), whereas 3′ end maturation occurs through an undefined mechanism that ultimately results in the 2′-O-methylation of the 3′ residue (Klattenhoff and Theurkauf, 2008). In plants, fungi, and nematodes, silencing signals are amplified from target RNA by RNA-dependent RNA polymerases (RdRPs). In some cases, RdRPs and Dicer function in a concerted manner to synthesize and process dsRNA, producing siRNAs with 5′ monophosphate residues (Colmenares et al., 2007; Lee and Collins, 2007). In nematodes, RdRPs also catalyze the unprimed, de novo synthesis of 5′-triphosphorylated RNAs that appear to be loaded directly, without Dicer processing, onto members of an expanded clade of worm-specific AGOs (WAGOs) (Aoki et al., 2007; Pak and Fire, 2007; Sijen et al., 2007; Yigit et al., 2006).

In plants, flies, and mammals, several classes of small RNA species are derived from transposable elements and repeat sequences as well as a subset of nonrepetitive protein-coding sequences and pseudogenes (Czech et al., 2008; Ghildiyal

et al., 2008; Kasschau et al., 2007; Okamura et al., 2008; Tam et al., 2008; Watanabe et al., 2008). In many cases, these small RNAs are derived from loci capable of dsRNA formation, including transposable elements, inverted repeats, and bidirectionally transcribed regions. Naturally occurring small RNAs are often coincident with pericentric heterochromatin and have been implicated in the establishment and/or maintenance of heterochromatin and in centromere function (Hall et al., 2003; Verdel et al., 2004). In *C. elegans*, endogenous small RNAs (endo-siRNAs) have been reported to target several hundred loci, including protein-coding as well as noncoding loci (Ambros et al., 2003; Lim et al., 2003; Ruby et al., 2006).

Here, we demonstrate that the Dicer-related helicase DRH-3 is essential for the biogenesis of RdRP-derived small RNAs in *C. elegans*. We have named these small RNAs 22G-RNAs based on their strong propensity for a 5′ Guanosine (5′G) residue and a length of 22 nt. 22G-RNAs are abundantly expressed in the germline and are maternally deposited in embryos. Surprisingly, the majority of 22G-RNAs target unique genome sequences, including ~50% of the annotated coding genes in *C. elegans*. In addition to DRH-3, we show that the RdRPs, RRF-1 and EGO-1, and the tudor-domain protein EKL-1 are required for the biogenesis of 22G-RNAs. 22G-RNAs can be divided into two major systems based on the associated AGOs and their cofactors. One of these systems is dependent on RDE-3, MUT-7, and members of the WAGOs, including WAGO-1, while the second is dependent on the AGO CSR-1 and the nucleotidyl transferase CDE-1. The WAGO 22G-RNA system silences transposons, pseudogenes, and cryptic loci as well as certain genes, while the CSR-1 system functions to promote chromosome segregation (Claycomb et al., 2009). Finally, we demonstrate a role for components of the nonsense-mediated mRNA decay (NMD) pathway in 22G-RNA biogenesis. Our findings uncover a surprisingly rich maternal inheritance of small RNAs and raise many questions about the potential significance of these RNA species in the transmission of epigenetic information and genome surveillance.

RESULTS

Novel Alleles of *drh-3* Disrupt RNAi

A previous study identified DRH-3 as a Dicer-interacting factor required for germline RNAi and for viability (Duchaine et al., 2006). Animals homozygous for the *drh-3(tm1217)* deletion (a putative null allele) are infertile, and RNAi targeting *drh-3* results in a penetrant embryonic lethal phenotype with defects in chromosome segregation and in the production of both Dicer-dependent and Dicer-independent small RNA populations (Duchaine et al., 2006; Nakamura et al., 2007).

Our screens for RNAi-deficient (Rde) strains identified three additional alleles of *drh-3*. The homozygous mutants bearing these non-null alleles are viable at 20°C but are infertile at 25°C. Each allele alters a distinct amino acid within the HELICc domain of the putative helicase (Figure 1A), and, based on genetic tests, each behaves like a partial loss-of-function mutation. Consistent with previous work demonstrating that DRH-3 is required for germline RNAi (Duchaine et al., 2006), these *drh-3* point mutants were defective for RNAi targeting the maternal

gene *pos-1* (Figure 1B) and exhibited varying degrees of somatic RNAi deficiency. For example, the *drh-3(ne3197)* mutant was strongly resistant to RNAi targeting the somatically expressed basement membrane collagen *let-2*, while the *ne4253* and *ne4254* alleles are partially and fully sensitive, respectively (Figure 1B).

Although the *drh-3* point mutants exhibit a range of phenotypes that increase in penetrance at 25°C, they do not appear to be classic temperature-sensitive mutants. Null alleles of many germline factors, including several RNAi-pathway genes, cause similar conditional-sterile phenotypes that likely reflect an underlying temperature-dependent process in the germline that is uncovered in these mutant backgrounds. The spectrum of phenotypes observed in the *drh-3* point-mutant strains, including sterility, embryonic lethality, and high incidence of males (*him*) (Figures 1C and 1D), are similar to those observed for null alleles of several Mutator (Mut) class, Rde strains. And, as observed with Mut strains, the RNAi defect associated with the *drh-3* alleles was not affected by temperature: each allele was fully resistant to *pos-1* RNAi at the permissive temperature of 20°C (Figure 1A and data not shown).

To determine whether the *drh-3* point mutants showed small RNA defects similar to the deletion mutant (Duchaine et al., 2006), small RNA populations were isolated and analyzed by northern blot to detect previously characterized endo-siRNAs (Figure S1). All of the previously examined DRH-3-dependent endo-siRNAs were dramatically reduced or undetected in the *drh-3* point-mutant samples, while miRNA biogenesis was unaffected. These results are consistent with the notion that these alleles of *drh-3* cause a partial loss of *drh-3* function.

DRH-3 Is Essential for the Biogenesis of 22G-RNAs

In the course of analyzing the small RNA phenotypes associated with the *drh-3* mutants, we observed that a prominent small RNA species of ~22 nt was virtually absent in small RNA samples prepared from each of the *drh-3* mutants (Figure 2A). A second prominent small RNA species of ~21 nt appeared to be unaffected in *drh-3*. Neither small RNA species was altered in samples prepared from a *dcr-1(ok247)* deletion mutant (Figure 2A). Thin-layer chromatography experiments indicated that the 22 nt small RNAs have a 5′G and are resistant to terminator exonuclease, suggesting the presence of a 5′ cap or polyphosphate (Figure 2B). The 21 nt small RNAs are comprised of both 5′G and 5′ uracil (5′U) species, the latter of which we recently identified as the 21U-RNAs, piRNAs, associated with PRG-1 (Batista et al., 2008). In contrast to 21U-RNAs, the majority of the 22 nt small RNAs were sensitive to periodate (data not shown), indicating that the 3′ end is not modified (Ruby et al., 2006). Both 22 nt and 21 nt 5′G small RNAs were dramatically reduced in the *drh-3* mutants, whereas 5′U small RNAs were unaffected. These data are consistent with the notion that the 22 nt small RNAs represent an abundant pool of 5′-triphosphorylated products of endogenous RdRP (Aoki et al., 2007; Pak and Fire, 2007; Sijen et al., 2007).

To identify and characterize the DRH-3-dependent small RNAs on a genome-wide level, small RNAs between 18 and 26 nt were cloned from wild-type and *drh-3* mutant animals using a protocol compatible with cloning small RNAs bearing
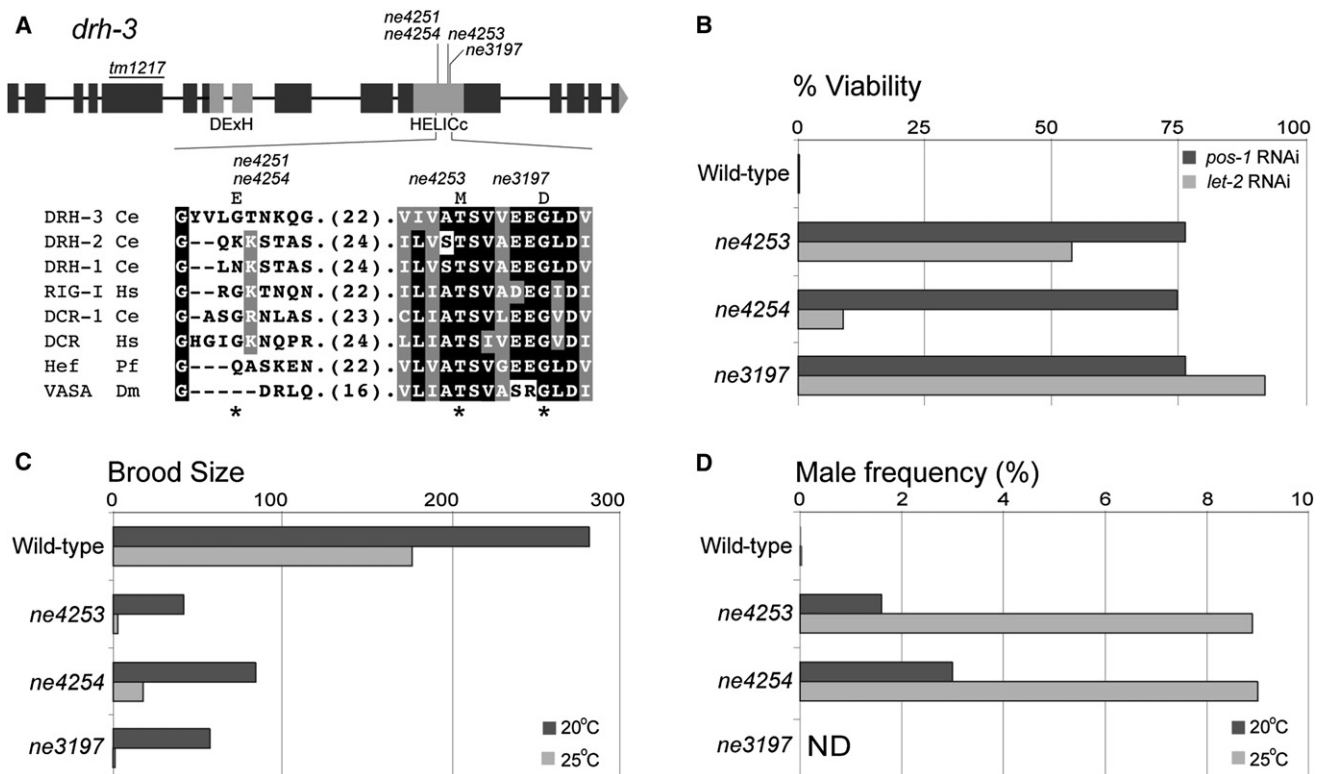
Figure 1. Hypomorphic Alleles of *drh-3* Are RNAi Deficient and Temperature Sensitive

(A) Schematic of the *drh-3* gene structure. Top panel, conserved DExH and HELICc domains and *drh-3* lesions; bottom panel, four missense alleles (indicated) map to the HELICc domain.

(B) RNAi-deficient phenotypes of *drh-3* mutants; fraction of viable embryos produced by animals exposed to *pos-1* RNAi food or viability of animals reared on *let-2* RNAi food at 20°C.

(C and D) Brood size and Him phenotypes of *drh-3* mutants at 20°C (dark bars) and 25°C (light bars). Mean brood size (top panel) of ≥ 10 hermaphrodites was determined by counting the number of embryos produced. Frequency of males among viable offspring was determined (bottom panel).

5′ triphosphates. Illumina sequencing of wild-type and *drh-3* libraries yielded 2.34 million and 4.33 million reads that perfectly match the *C. elegans* genome, respectively. Consistent with our biochemical analyses, both the size distribution and first nucleotide composition of small RNAs were dramatically altered in the *drh-3* sample (Figure 2C). First, wild-type reads peaked sharply at 21–22 nt, comprising 25% and 36% of the total reads, respectively. Whereas 21 nt reads had similar levels of 5′U and 5′G residues (11% and 10% of total reads, respectively), ∼60% of 22 nt reads started with 5′G (∼21% of total reads). Strikingly, reads with 5′G were strongly depleted from the *drh-3* mutant sample, resulting in a dramatic enrichment of reads with 5′U.

After removing structural RNA degradation products from the data set, about one-third of the wild-type reads matched to miRNAs (24.7%) and 21U-RNAs (11%) (Figure 2D). The most abundant class of small RNA reads from wild-type samples (34%) were antisense to protein-coding genes. The remaining small RNA reads were derived from transposons and other repetitive loci (∼10%) as well as nonannotated loci (∼20%). In contrast, the *drh-3* sample was strongly depleted of endo-siRNAs targeting protein-coding genes, pseudogenes, repeats, and nonannotated loci but enriched proportionately for reads matching miRNAs and 21U-RNAs. Together, our biochemical

and deep-sequencing data indicated that DRH-3 is essential for the biogenesis of an abundant class of endo-siRNAs expressed in *C. elegans*. Based on the propensity for 22 nt length and 5′G residue, we refer to these small RNAs as 22G-RNAs, which include what were previously identified as endogenous siRNAs (Ambros et al., 2003; Ruby et al., 2006).

22G-RNAs Are Enriched at Transcript Termini

While examining the distribution of 22G-RNAs targeting protein-coding loci, we observed that 22G-RNAs were most abundant toward the 3′ end of many protein-coding transcripts. For example, 22G-RNAs were clearly enriched at the 3′ ends of both *rrf-1* and *ama-1* transcripts in the wild-type sample and tapered toward the 5′ end (Figure 2E). Interestingly, 22G-RNAs were also enriched at the 5′ end of *rrf-1*. Remarkably, the remnant of *rrf-1* and *ama-1* 22G-RNAs that were cloned from *drh-3* mutants mapped almost exclusively to the 3′ end of these loci. This pattern might be expected if RdRP initiates 22G-RNA biogenesis at or near the 3′ end of a transcript.

To examine this on a larger scale, each transcript for which 22G-RNA reads occur in both wild-type and *drh-3* mutant samples was divided into 20 consecutive intervals of equal size. The total number of 22G-RNAs that map to each interval
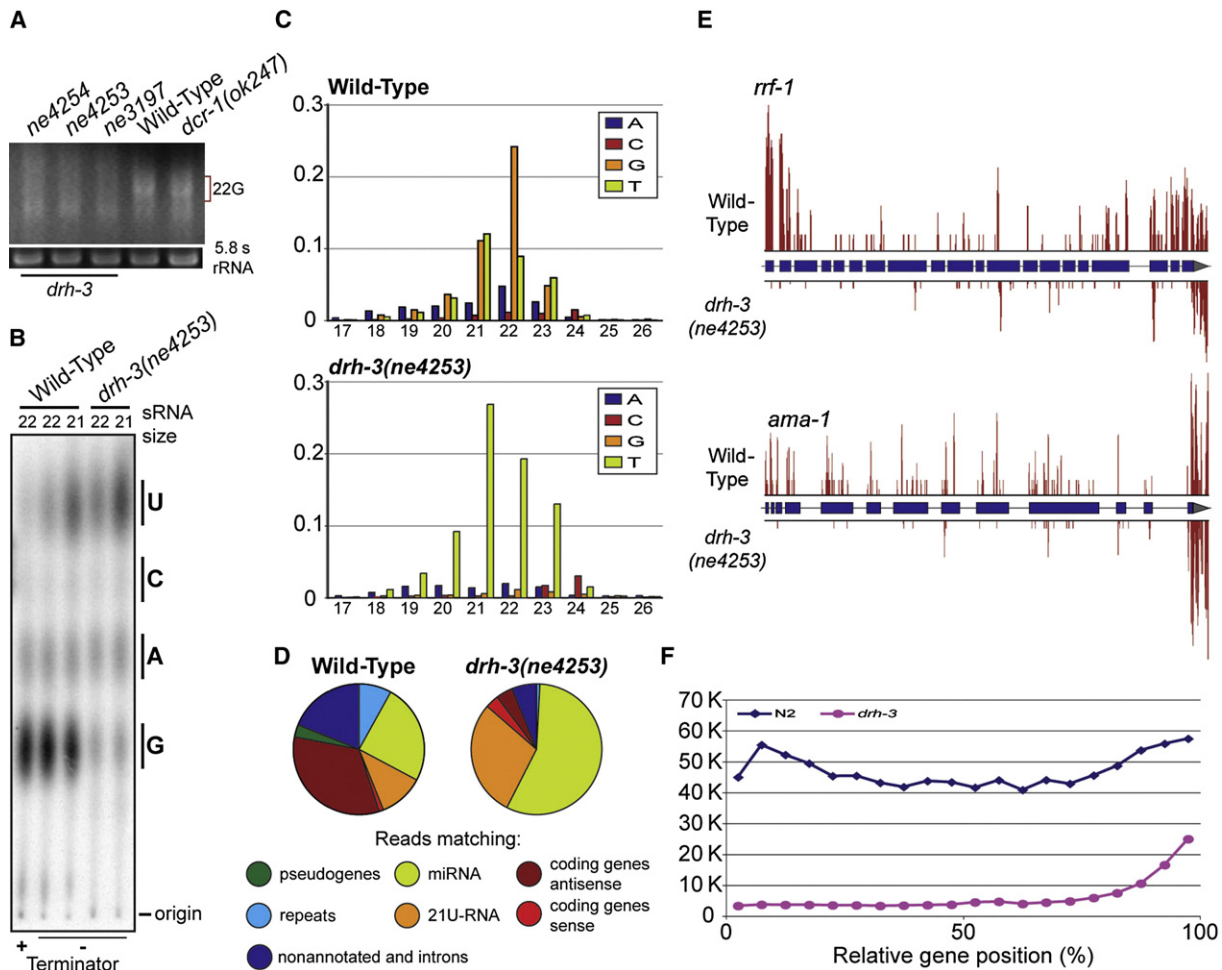
Figure 2. DRH-3 Is Required for the Biogenesis of 22G-RNAs

(A) Ethidium bromide staining of small RNAs. Twenty-two nucleotide position is denoted by bracket. Note persistence of ∼21 nt band in all samples. 5.8S rRNA was used as loading control.

(B) Thin-layer chromatography analysis of first nucleotide of gel-purified 22 and 21 nt RNA, as indicated, from wild-type and *drh-3* mutant samples. Bars (right), position of each nucleotide; "+," treated with Terminator exonuclease; "−," untreated.

(C) Length and first nucleotide distribution of genome-matching reads from wild-type and *drh-3* mutant small RNA libraries. Sense structural RNAs were excluded in the analysis.

(D) Distribution of reads that match indicated genome annotations sequenced in wild-type and *drh-3* mutant small RNA libraries.

(E) Examples of small RNA distribution within *rrf-1* and *ama-1*. Density of antisense reads indicated by vertical bars above (wild-type) or below (*drh-3*) gene structure.

(F) Genome-wide analysis of small RNA distribution within genes. Total antisense reads (y axis) plotted according to the relative position (%) within all genes (x axis; from 5′ to 3′).

was plotted for both wild-type and *drh-3*. In wild-type, 22G-RNAs were more abundant toward both the 5′ and 3′ termini of predicted transcripts (Figure 2F). This terminal distribution of 22G-RNAs represented a general trend for most target genes and was not caused by a few genes with a high number of small RNAs at either end (Figure S2).

Markedly different results were obtained when this analysis was applied to the *drh-3* mutant data. In *drh-3* mutants, while 22G-RNAs were greatly depleted, they were not depleted uniformly across their targets (Figures 2F and S2). The levels of 22G-RNAs within the first 15 bins (representing 75% transcript length) were disproportionately depleted, resulting in a marked

exponential trend in 22G-RNA levels over the remaining 5 bins, peaking at the 3′ end. This 3′ end enrichment was evident for about 86% of the 22G-RNA targets. Together, these results raise the possibility that 22G-RNA biogenesis is initiated at the 3′ end of most targets, and the wild-type activity of *drh-3* promotes the propagation of 22G-RNA biogenesis by RdRP along the template RNA.

Distinct Genetic Requirements for 22G-RNAs

To characterize the genetic pathways required for the biogenesis of 22G-RNAs, we examined a panel of RNAi-related mutants by northern blot analysis to detect small RNAs derived from
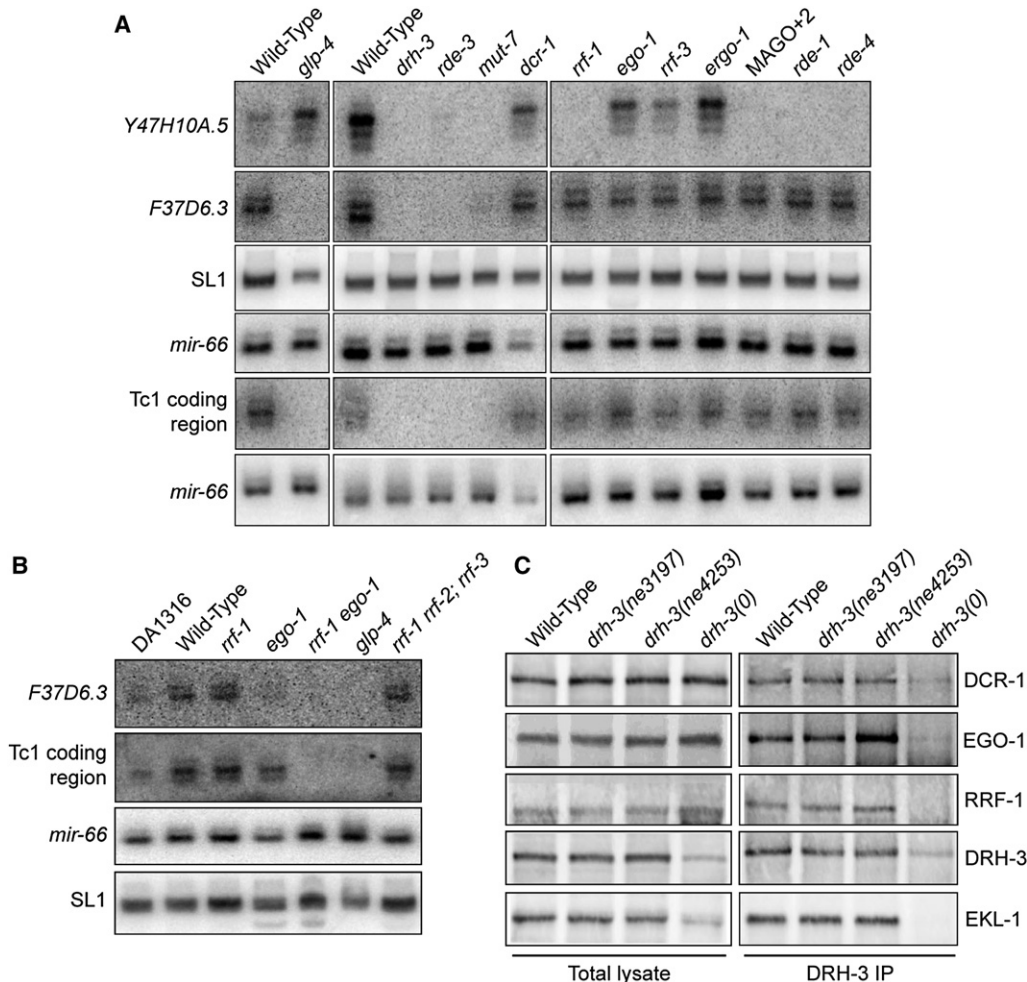
Figure 3. Distinct Requirements and Genetic Redundancy in 22G-RNA Pathways

(A) Northern blots of 22G-RNA expression in somatic gene *Y47H10A.5* and germline targets *F37D6.3* and Tc1 transposon in RNAi mutants. Loading controls: SL1 precursor and *mir-66*.

(B) Northern blots of germline 22G-RNAs in RdRP mutants. DA1316 is a congenic wild-type control for the *ego-1* and *rrf-1 ego-1* mutants. Loading controls: SL1 precursor and *mir-66*.

(C) Co-IP analyses of DRH-3 from wild-type and *drh-3* lysates. Total lysate (left panels) and DRH-3 IPs (right panels) were analyzed by western blot for candidate interacting proteins (indicated at right). *drh-3(0)*: the deletion allele (*tm1217*) control. Some DRH-3 protein detected in the *drh-3(0)* sample likely represents persistence of maternal product.

representative abundant 22G-RNA loci. In addition, we performed small RNA cloning experiments to identify 22G-RNA populations that are enriched in the germline, soma, or oocyte (Figure S3, Tables S1 and S2, and Supplemental Discussion). This exercise revealed that most 22G-RNAs are germline expressed and/or maternal (Figure S3). As expected from our deep-sequencing experiments, we found that 22G-RNAs enriched in both the soma (*Y47H10A.5*) and germline (*F37D6.3* and Tc1) were dependent on DRH-3. Likewise, RDE-3 and MUT-7 were required for 22G-RNAs targeting all three loci (Figure 3A), indicating that at least one aspect of the mechanism of 22G-RNA expression is shared by these somatic and germline loci.

Previous work has shown that some somatic 22G-RNA loci (e.g., *K02E2.6* and X-cluster) are dependent on the ERI endo-

RNAi complex as well as multiple AGO proteins that interact with secondary siRNAs produced by RdRP (Duchaine et al., 2006; Yigit et al., 2006). However, somatic 22G-RNAs derived from *Y47H10A.5* (Figure 3A) were independent of the ERI pathway genes *rrf-3* and *ergo-1*. Instead, *Y47H10A.5* 22G-RNAs fail to accumulate in exogenous (exo-) RNAi pathway mutants, including *rde-4*, *rde-1*, and *rrf-1* as well as a derivative of the multiple AGO mutant MAGO (Yigit et al., 2006) with two additional AGO mutations (MAGO+2), suggesting that *Y47H10A.5* 22G-RNAs could be triggered by dsRNA.

Germline 22G-RNAs appear to be independent of the exo-RNAi and ERI pathways. For example, mutations in *dcr-1*, *rde-4*, or *ergo-1* caused no visible depletion in 22G-RNA populations based on ethidium bromide staining of small RNA (Figure 2A and data not shown) and were not required for the

production of *F37D6.3* or Tc1 22G-RNAs (Figures 3A and S4). These findings suggest that 22G-RNA biogenesis is not triggered by dsRNA at these and many other targets. Despite the biochemical evidence indicating that germline 22G-RNAs are the products of RdRP, we observed near-wild-type levels of 22G-RNAs in each of the individual RdRP mutants as well as in the MAGO+2 mutant (Figures 3A and 3B), suggesting additional redundancy within the respective RdRP and AGO families of proteins.

DRH-3, the Tudor-Domain Protein EKL-1, and Two Partially Redundant RdRPs Form a Core Complex Essential for 22G-RNA Biogenesis

RRF-1 is required for RNAi in somatic tissues (Sijen et al., 2001), while EGO-1 is required for fertility and has been implicated in RNAi targeting some but not all germline-expressed genes (Smardon et al., 2000). The latter finding could be explained if RRF-1 is functionally redundant with EGO-1 and is expressed within an overlapping domain in the germline. This might also account for the persistence of germline 22G-RNA expression in the RdRP single mutants analyzed (Figure 3). A noncomplementation screen to generate the *rrf-1 ego-1* double mutant yielded a rearrangement, *neC1*, that disrupts *rrf-1* and results in a putative null allele of *rrf-1* linked to the *ego-1(om97)* nonsense allele (Figure S5). Northern blot analyses of small RNAs prepared from the *rrf-1(neC1) ego-1(om97)* double mutant revealed that germline 22G-RNAs fail to accumulate in animals null for both *rrf-1* and *ego-1* (Figure 3B), demonstrating that RRF-1 and EGO-1 function redundantly in the germline to produce 22G-RNAs.

Consistent with the overlapping functions of RRF-1 and EGO-1 for germline 22G-RNA biogenesis, both RRF-1 and EGO-1 interacted with DRH-3 in immunoprecipitation (IP) experiments (Figure 3C). A recent study demonstrated that DRH-3 interacts with RRF-1 and is required for RdRP activity in vitro (Aoki et al., 2007). We previously identified DRH-3 as a component of the ERI complex, which includes the tudor-domain protein ERI-5 (Duchaine et al., 2006). EKL-1 is a close homolog of ERI-5 and is required for fertility, RNAi, and chromosome segregation (Claycomb et al., 2009). Consistent with the phenotypic similarities between *drh-3*, *ego-1*, and *ekl-1* mutants (Duchaine et al., 2006; Rocheleau et al., 2008; Vought et al., 2005), EKL-1 also interacted with DRH-3 (Figure 3C). In addition, both EGO-1 and EKL-1 were among the most enriched proteins in DRH-3 IPs as assessed by multidimensional protein identification technology (MudPIT) (Figure S6). Although DCR-1 was detected in the DRH-3 IP by western blot, DCR-1 peptides were not identified in DRH-3 IP-MudPIT experiments. Combined with our deep-sequencing data (below), these data suggest that DRH-3, EKL-1, and RdRP form a core RdRP complex that is essential for the biogenesis of 22G-RNAs in *C. elegans*.

WAGO Family AGOs Function Redundantly and Interact with Germline 22G-RNAs

Previously, we demonstrated that the accumulation of secondary siRNAs generated by RdRP is dependent upon multiple redundant WAGO proteins. The previously described MAGO mutant was strongly defective for RNAi and failed to accumulate certain

endo-siRNAs (Yigit et al., 2006). However, the MAGO+2 mutant, which includes *ppw-2(tm1120)* and *C04F12.1(tm1637)* as well as the MAGO mutations (Yigit et al., 2006), continued to express normal levels of germline 22G-RNAs targeting *F37D6.3* and Tc1 (Figure 3A), indicating that additional AGOs interact with germline 22G-RNAs.

Therefore, we generated additional combinations of mutants within the WAGO clade (Figure 4A). This analysis identified other WAGO mutant combinations with germline RNAi defects. For example, *ppw-2(tm1120)*; *f58g1.1(tm1019)* double mutants were resistant to *pos-1(RNAi)* (data not shown), whereas the individual alleles are sensitive to *pos-1(RNAi)* (Yigit et al., 2006). A mutant lacking four of the branch III WAGOs, including *ppw-2(tm1120)*, *F55A12.1(tm2686)*, *F58G1.1(tm1019)*, and *ZK1248.7(tm1113)*, was resistant to germline RNAi (data not shown) but still produced normal levels of *F37D6.3* germline 22G-RNAs (Figure 4B, Quadruple). Deletion of the branch III WAGO, *wago-1(tm1414)*, resulted in a Quintuple AGO mutant with dramatically reduced germline 22G-RNAs (Figures 4B and 4C). Furthermore, the *wago-1(tm1414)* mutant alone showed a reduction in *F37D6.3* germline 22G-RNAs that was comparable to the Quintuple AGO mutant (Figure 4C), demonstrating that WAGO-1 plays a key role in germline 22G-RNA function. Transgenic lines expressing a GFP::WAGO-1 fusion under the control of the *wago-1* promoter revealed that WAGO-1 is expressed in the germline and localizes to perinuclear foci that resemble P granules (Figures 4D and 4E).

Finally, we generated a strain lacking all 12 of the WAGO genes (not including predicted pseudogenes). This duodecuple mutant (MAGO12) is viable and resistant to RNAi and exhibits a high frequency of spontaneous males and temperature-dependent sterility at 25°C (data not shown). Germline 22G-RNAs were undetectable by northern blot analysis in the MAGO12 strain (Figure 4C), demonstrating a clear dependence of the 22G-RNAs on WAGOs.

WAGOs, RDE-3, and MUT-7 Are Required for Germline 22G-RNA Silencing Pathways

To gain insight into the function of germline 22G-RNAs, we performed deep sequencing of small RNA populations from mutants with germline 22G-RNA defects. In addition, we generated transgenic animals that express a 3×FLAG::WAGO-1 fusion protein and deep sequenced the small RNAs that coprecipitate with WAGO-1. For each mutant, the fraction of reads matching coding genes, nonannotated loci, and repeat elements was reduced, with concomitant increases in the fraction of miRNA and 21U-RNA reads (Figure S7). Conversely, the fraction of reads matching to a particular set of coding genes, nonannotated loci, and repeat elements were enriched in the small RNA library prepared from the WAGO-1 immunoprecipitate, while miRNAs and 21U-RNAs were severely depleted (Figure 5A).

We next asked whether the reduction of 22G-RNAs in each mutant occurred globally or at particular loci. 22G-RNAs targeting protein-coding loci were nearly completely eliminated in the *drh-3* and *ekl-1* single mutants and in the *rrf-1 ego-1* double mutant. Gene-targeted 22G-RNAs were far less likely to be depleted in the *rde-3*, *mut-7*, and MAGO12 mutant samples (Figure 5B), with the notable exception that 22G-RNA species
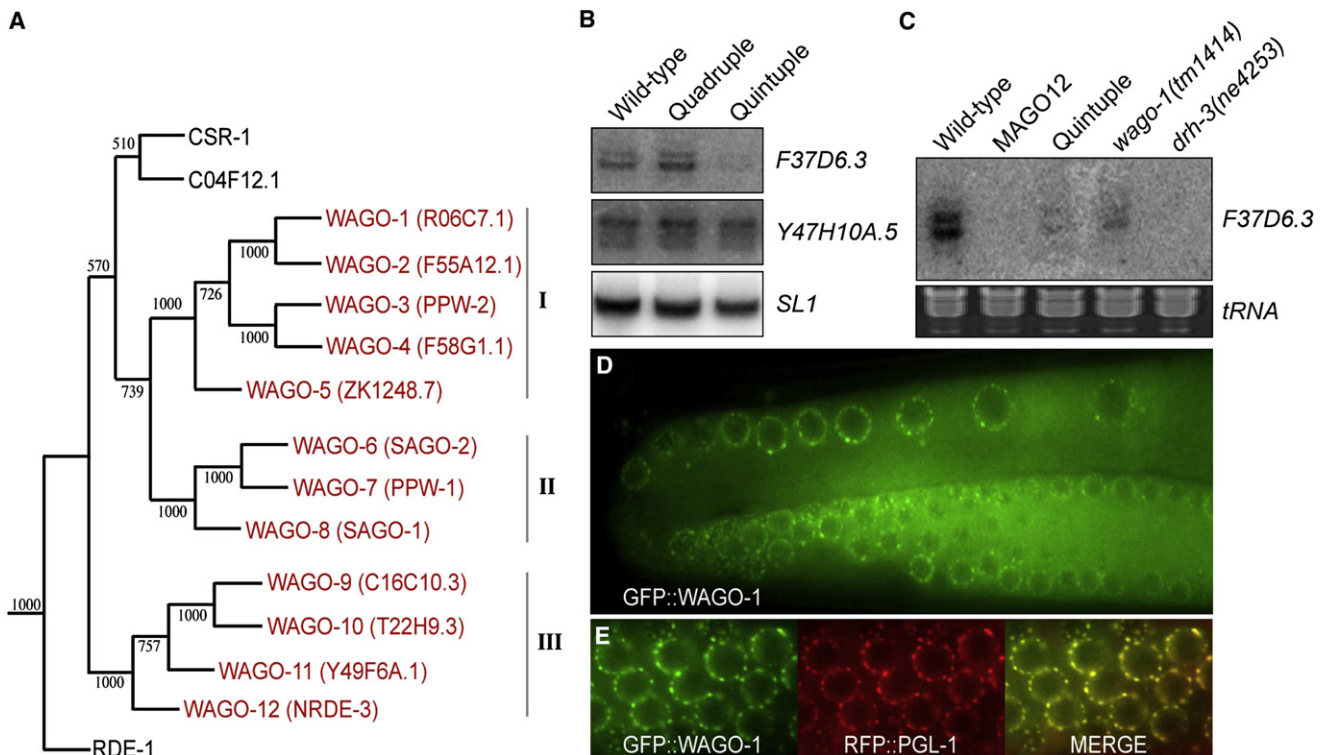
Figure 4. WAGO-1 and Highly Redundant WAGOs Required for Germline 22G-RNA Biogenesis

(A) Phylogenetic representation of WAGO proteins (see Supplemental Experimental Procedures). WAGOs deleted in MAGO12 are indicated in red.

(B and C) Northern blots of *F37D6.3* 22G-RNAs in multiple WAGO mutants. Loading control: SL1 precursor (B) and tRNA staining (C).

(D and E) Fluorescence microscopy of GFP::WAGO-1 and RFP::PGL-1, which colocalize to P granules in the germline.

targeting a subset of genes with normally very high 22G-RNA levels were strongly depleted in each of these mutants (Figure S8). In contrast, 22G-RNAs were largely unaffected in an *rde-4* mutant, which is required for ERI-class small RNAs (Duchaine et al., 2006; Lee et al., 2006; J.V., W.G., and C.C.M., unpublished data). The 22G-RNAs depleted in the *rde-3*, *mut-7*, and MAGO12 mutants were almost completely overlapping (Figure 5C and Table S3). Despite an overall reduction of 22G-RNA reads, a subset of 22G-RNA species was not depleted in *rde-3*, *mut-7*, and MAGO12 mutants (Figures 5B and S8); in fact, some were increased in proportion. These WAGO-independent 22G-RNA populations are associated with and dependent upon another germline-expressed AGO, CSR-1 (Figure S9) (Claycomb et al., 2009). The bimodal distribution of 22G-RNA loci indicates that at least two qualitatively distinct 22G-RNA pathways exist in the germline that depend on a core set of factors (DRH-3, EKL-1, and RdRP) whose small-RNA products interact with distinct AGOs.

Consistent with the requirement for WAGO-pathway components in exo-RNAi, the WAGO-associated 22G-RNAs appear to be involved in silencing their respective targets. Loci with the highest levels of 22G-RNAs in wild-type were consistently derepressed in the *drh-3* mutant, as assessed by semiquantitative PCR with reverse transcription (qRT-PCR) (Figure 5E) and Affymetrix tiling arrays (Figure S10). In contrast, CSR-1-associated 22G-RNAs do not appear to silence their targets (Claycomb

et al., 2009), consistent with the biological distinction between these pathways.

Previous work has shown that Tc elements are silenced in the germline by an RNAi mechanism (Ketting et al., 1999; Sijen and Plasterk, 2003; Tabara et al., 1999). Individual Repbase annotations, which include all major classes of transposons in *C. elegans*, were uniformly depleted of 22G-RNAs in *drh-3*, *ekl-1*, RdRP, *rde-3*, *mut-7*, and MAGO12 mutant samples, and most transposon 22G-RNAs were enriched in the WAGO-1 IP sample (Figure 5B). Transposon loci showed normal levels of 22G-RNAs in an *rde-4* mutant sample (Figure 5B) (Tabara et al., 1999). Thus, the transposon-silencing pathway in *C. elegans* consists of DRH-3, EKL-1, RdRPs, RDE-3, MUT-7, and multiple WAGOs, including WAGO-1.

The *drh-3* alleles described here display the hallmarks of Mut class, Rde mutants (Figure 1). Indeed, spontaneous mutants with phenotypes that revert at high frequency were cloned from the *drh-3* mutants, including a *dpy-5::Tc5* insertion. The frequency of reversion from Dumpy to wild-type, upon excision of Tc5 from *dpy-5* in *drh-3(ne4253)*, was similar to an allele of *mut-7(ne4255)* that was isolated in the same screen and almost 5-fold higher than the nonsense allele *mut-7(pk204)* (Figure 5D). Similar results were obtained with an *unc-22::Tc1* insertion (data not shown). Furthermore, Tc1 and Tc3 transcripts were derepressed in the *drh-3* mutant (Figure 5E), demonstrating that DRH-3 is required for transposon silencing.
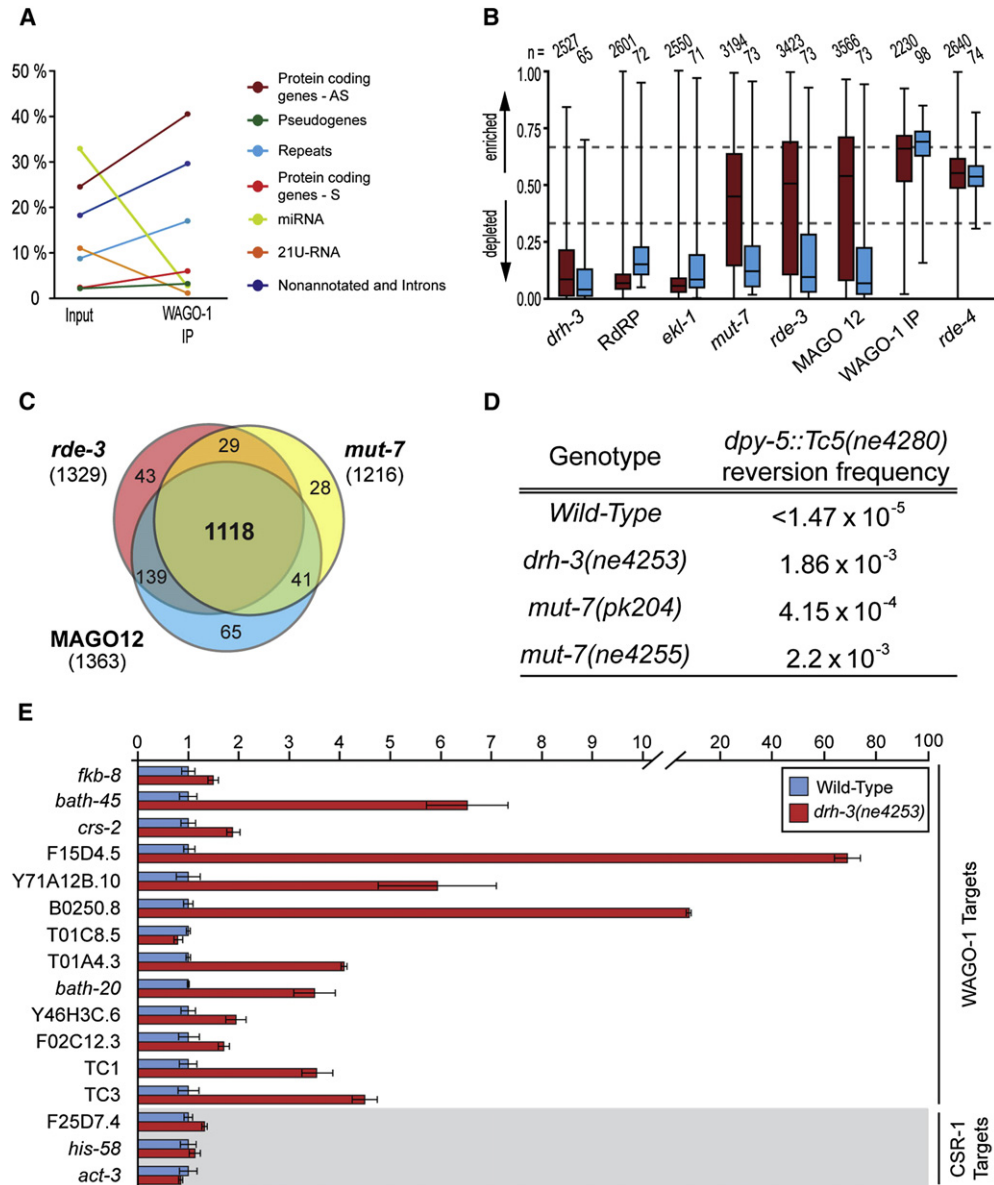
Figure 5. Deep-Sequence Analyses Identified at Least Two Distinct 22G-RNA Pathways in the Germline

(A) Graph depicting change in reads matching indicated genome annotations between input and WAGO-1 IP samples.

(B) Enrichment or depletion of 22G-RNAs derived from genes (red) or transposons (blue) in each mutant. Values approaching 1 indicate enrichment of 22G-RNAs, while values approaching 0 indicate depletion. Relative enrichment is calculated as ratio of mutant / (mutant + wild-type). WAGO-1 IP enrichment is calculated as WAGO-1 IP / (WAGO-1 IP + input). "n" loci with at least 25 reads per million total reads (not including structural) in either the wild-type or mutant sample were analyzed. The top and bottom of each box represent the 75th and 25th percentiles, respectively. The horizontal line within the box represents the median value. Dotted lines denote 2-fold enrichment (upper) and 2-fold depletion (lower).

(C) Venn comparison of genes depleted of 22G-RNAs (≥ 2-fold) in indicated mutants (loci below the lower dashed line in B).

(D) Frequency of reversion of *dpy-5::Tc5* as indicated (Ketting et al., 1999).

(E) Quantitative RT-PCR analysis of 22G-RNA target expression in *drh-3* mutant (red) relative to wild-type (blue). WAGO and CSR-1 (gray) targets are indicated. Error bars represent the standard deviation of the mean.

22G-RNAs and Surveillance

Approximately 15% of 22G-RNA reads were derived from non-annotated loci and were dependent on RDE-3, MUT-7, and MAGO12. These loci primarily correspond to unique intergenic sequences and could represent pseudogenes or cryptic loci that lack open reading frames and are unrecognizable by current bioinformatic approaches. In some cases, we could predict potential splicing patterns based on antisense reads spanning these nonannotated regions (Figure S11A). Consistent with this notion, 22G-RNAs derived from many loci annotated as
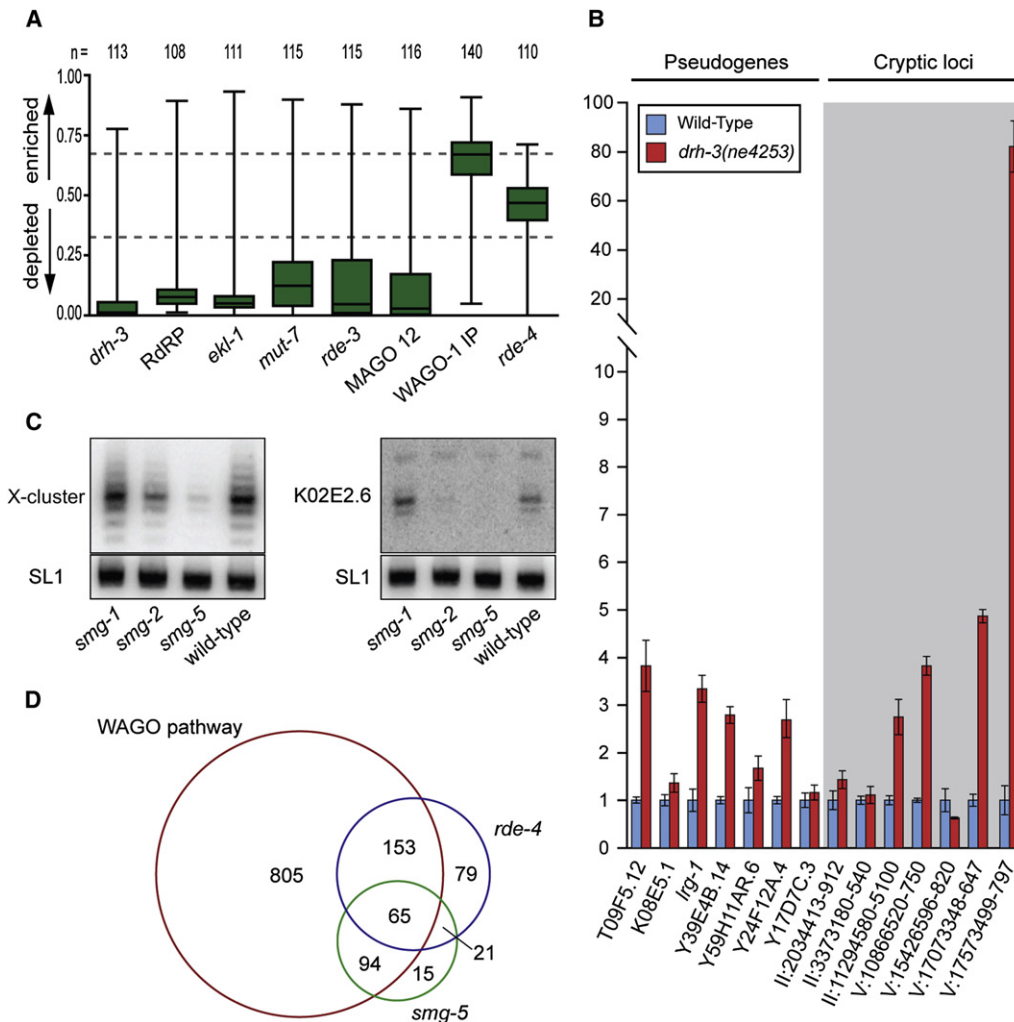
Figure 6. WAGO-Associated 22G-RNAs Define a Surveillance System

(A) Enrichment or depletion of 22G-RNAs derived from "n" annotated pseudogenes (as described in Figure 5B).

(B) Quantitative RT-PCR analysis of pseudogenes and cryptic loci targeted by WAGO pathway in *drh-3* mutant (red) relative to wild-type (blue). Error bars represent the standard deviation of the mean.

(C) Northern blots of 22G-RNAs in *smg* mutants grown at 20°C. Similar results obtained with mutants grown at 25°C (data not shown). Loading control: SL1 precursor.

(D) Venn comparison of genes depleted of 22G-RNAs (≥2-fold) in the indicated mutants. WAGO pathway represents intersection of *rde-3*, *mut-7*, and MAGO12 from Figure 5C.

pseudogenes were also depleted in *rde-3*, *mut-7*, and MAGO12 mutants (Figure 6A). As with annotated genes targeted by WAGO-associated 22G-RNAs, qRT-PCR and microarray analysis demonstrated that both pseudogenes and cryptic loci targeted by 22G-RNAs were desilenced in the *drh-3* mutant (Figures 6B and S12).

Upon closer inspection of protein-coding loci targeted by the WAGO pathway, we noted that the 22G-RNA profile often did not correspond to the annotated gene prediction (Figure S11B and Table S4). In many cases, 22G-RNAs mapped within predicted introns, suggesting that the corresponding introns were not spliced in the target RNA. In other cases, 22G-RNAs started or ended abruptly in the middle of the annotation and extended

well upstream or downstream of the gene prediction, suggesting that the annotation is incomplete or incorrect. Lastly, we noticed a number of WAGO target genes with intron annotations in 3′UTRs.

Because pseudogenes and genes with 3′UTR introns are expected to be targets of the NMD pathway, we asked whether 22G-RNA biogenesis was dependent on the PIN domain protein SMG-5, the Upf1 helicase SMG-2, and the phosphatidylinositol-kinase SMG-1 (Anders et al., 2003; Glavan et al., 2006; Grimson et al., 2004; Page et al., 1999). 22G-RNAs derived from X-cluster and *K02E2.6*, which has a 3′UTR intron, were reduced in the null mutant *smg-5(r860)* and to a lesser extent in the non-null *smg-2(r863)* (Figure 6C). *K02E2.6* and X-cluster 22G-RNAs were unchanged in the temperature-sensitive mutant *smg-1*
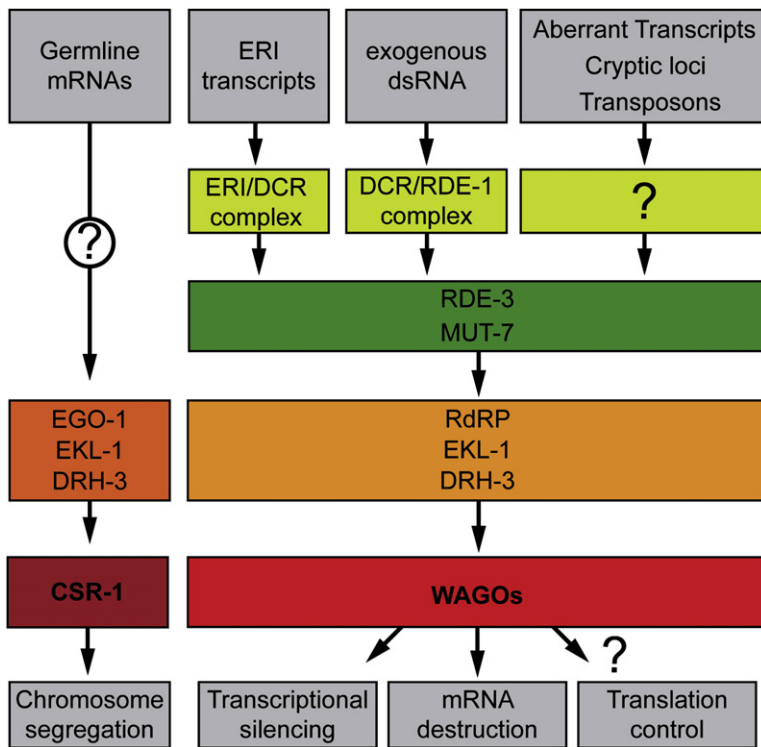
Figure 7. Model of Germline 22G-RNA Pathways Required for Genome Surveillance in *C. elegans*

DRH-3, EKL-1, and the partially redundant RdRPs, RRF-1 and EGO-1, form a core RdRP complex that functions in multiple 22G-RNA pathways (Figure 7). RDE-3, MUT-7, and members of the WAGO clade, in particular WAGO-1, define a general 22G-RNA surveillance system that silences transposable elements and aberrant transcripts. A second pathway is dependent on CSR-1, which promotes kinetochore structure and chromosome segregation (Figure 7) (Claycomb et al., 2009). Taken together, 22G-RNAs appear to engage targets that derive from both the actively expressed regions (CSR-1 associated), as well as the "silent" regions of the genome (WAGO-1 associated). These findings support a model in which the 22G-RNA pathways exert genome-scale surveillance important for maintenance of the germline.

Dicer-Dependent versus Independent 22G-RNA Biogenesis

It is formally possible that maternal DCR-1 is sufficient to generate primary siRNAs and that subsequent recruitment of the secondary RNAi machinery, dependent on RdRPs and WAGOs, results in a self-sustaining amplification cycle to produce 22G-RNAs in *dcr-1* zygotic mutants. However, extensive northern blot experiments demonstrate that ERI-dependent 22G-RNAs and miRNAs are depleted in the *dcr-1* mutant despite the maternal contribution of DCR-1, while germline 22G-RNAs are present at normal or, in some cases, elevated levels (Figure S4). Aside from miRNA and exo-RNAi complexes, the ERI complex appears to be the primary DCR-1 complex in *C. elegans* (Duchaine et al., 2006). We have shown that 22G-RNAs are largely independent of RDE-4, a DCR-1 cofactor in both exo-RNAi and ERI pathways (Tabara et al., 2002; Duchaine et al., 2006; Lee et al., 2006). Thus, we favor the model that the major 22G-RNA pathways are initiated in a DCR-1-independent fashion. New alleles of *dcr-1* will be important in order to resolve this issue in the future.

The Role of DRH-3 in 22G-RNA Biogenesis

How might Dicer-independent 22G-RNA biogenesis be triggered, and what role does DRH-3 play in this process? A recent report has shown that *drh-3* null mutant extracts are deficient in the in vitro synthesis of antisense small RNA by RRF-1 (Aoki et al., 2007). Our analysis of 22G-RNA levels in hypomorphic *drh-3* point-mutant alleles is consistent with a role for the conserved DRH-3 helicase domain in 22G-RNA biogenesis. Two of the three *drh-3* missense alleles alter highly conserved residues within the HELICc domain. The *drh-3(ne4253)* lesion (T834M) alters a residue that contacts RNA in the Vasa crystal structure; the *drh-3(ne3197)* lesion (G840D) alters a residue that maps to the interface between the HELICc and DExH domains and coordinates the water molecule that is thought to

(cc546) at both permissive and nonpermissive temperatures (Figure 6C and data not shown). These data suggest a role for SMG-2 and SMG-5 in 22G-RNA biogenesis that is distinct from their recognized role in NMD and that NMD per se is not required for 22G-RNA biogenesis.

Deep-sequence analysis of *smg-5* mutant small RNAs revealed that SMG-5 is required for the biogenesis of 22G-RNAs targeting 15% of WAGO-dependent 22G-RNA target genes (Figure 6C). Interestingly, SMG-5 was not required for most pseudogene-derived 22G-RNAs (data not shown). Furthermore, the few published endogenous targets of NMD do not appear to be 22G-RNA targets (data not shown). Roughly half of the SMG-5-dependent 22G-RNA loci overlap with RDE-4-dependent 22G-RNA loci (Figure 6D), which includes both ERI-dependent and ERI-independent 22G-RNA loci (J.V., W.G., and C.C.M., unpublished data). These findings indicate that multiple WAGO-dependent 22G-RNA pathways exist, which together define a general surveillance system that silences transposons and aberrant transcripts.

DISCUSSION

In this study, we have combined deep sequencing with the powerful genetics of *C. elegans* to identify and characterize an abundant class of endo-siRNAs that we call 22G-RNAs. In adult animals, 22G-RNAs are primarily germline expressed and are derived from unique sequences in the genome, including coding genes, transposons, pseudogenes, and nonannotated loci. Combining data from three small RNA libraries, including an AGO IP sample, we have identified 22G-RNAs antisense to over 50% of the annotated protein-coding genes.

be required for ATP hydrolysis (Sengoku et al., 2006). Both lesions are likely to abrogate ATPase and/or unwinding activity based on structural and biochemical studies with related proteins (Liang et al., 1994; Sengoku et al., 2006).

The enrichment of 22G-RNAs at the 3′ end of transcripts suggests that 22G-RNA biogenesis begins at the 3′ end of target RNAs, followed by cycles of 22G-RNA synthesis by RdRP and proceeding along the template toward the 5′ end. Interestingly, the 3′ localized 22G-RNAs are least diminished in the *drh-3* mutant, suggesting that these lesions do not prevent the initial loading of RdRP onto the template but rather interfere with the processivity of RdRP. DRH-3 could remove secondary structure from the template or facilitate transfer of the 22G-RNA to downstream WAGOs, allowing RdRP to initiate a second round of synthesis at the next available C residue in the template RNA.

Whatever the mechanism by which DRH-3 promotes 22G-RNA biogenesis, our data support the idea that RdRP is recruited to the 3′ end of target transcripts. Templates lacking a poly(A) tail are better substrates for RdRP in vitro (Aoki et al., 2007), suggesting that defective 3′ end formation may be one trigger for 22G-RNA biogenesis. In *Arabidopsis*, decapped, misspliced, and misterminated transcripts are recognized by, and activate, the RNA-silencing machinery (Gazzani et al., 2004; Herr et al., 2006). In fission yeast, two β-nucleotidyl transferases, Cid12 and Cid14, determine whether transcripts are recognized by the RdRP complex (Cid12) or the TRAMP/exosome complex (Cid14) (Bühler et al., 2007, 2008; Motamedi et al., 2004). In *cid14* mutants, transcripts that are normally turned over by the TRAMP surveillance pathway become substrates for the RdRP complex (Bühler et al., 2008), indicating that these pathways recognize a common feature. Perhaps the β-nucleotidyl transferase RDE-3 and the 3′-to-5′ exonuclease MUT-7 function in an exosome-like pathway that recognizes and processes the 3′ end of aberrant transcripts, providing a signal that recruits the RdRP complex (Figure 7) (Chen et al., 2005; Ketting et al., 1999; Lee et al., 2006; Tabara et al., 1999).

Multiple Distinct 22G-RNA Pathways

Based on studies in other model systems, we expected a significant fraction of germline 22G-RNAs to be derived from transposons. Indeed, transposable elements are targets of 22G-RNAs. However, 22G-RNAs derived from unique sequences, both genic and intergenic, comprise a major fraction of the 22G-RNAs that interact with WAGO-1. Furthermore, loci that produce the highest levels of 22G-RNAs appear to interact with WAGO-1 and are the most desilenced in the *drh-3* mutant. Remarkably, CSR-1 interacts with a nonoverlapping population of 22G-RNAs derived almost exclusively from protein-coding loci (Figure S9) (Claycomb et al., 2009). Loci targeted by CSR-1 produce fewer 22G-RNAs than WAGO-1 loci and are not desilenced in either *csr-1* or *drh-3* mutants (Claycomb et al., 2009). These findings indicate that these pathways are mechanistically or functionally distinct (Figure 7), a conclusion that is consistent with the genetically defined functions of CSR-1 and WAGO-1. Although both pathways seem to be important for an efficient response to foreign dsRNA (Yigit et al., 2006; Claycomb et al., 2009), it seems likely that the WAGO surveillance system is primarily involved, as WAGOs were shown to be limiting for RNAi and to interact

directly with the secondary 22G-RNAs in the amplification cycle (Yigit et al., 2006). *csr-1* mutants disrupt the perinuclear localization of the germline nuage (Claycomb et al., 2009) wherein WAGO-1 resides and hence could indirectly affect WAGO-1 function by interfering with its proper localization within these germline structures. Further genetic and biochemical studies will be necessary to dissect the relative contributions of the WAGO-1 and CSR-1 pathways to RNAi.

At least some of the specificity of CSR-1 and WAGO 22G-RNA pathways can be attributed to the involvement of distinct RdRP complexes. We have uncovered a role for RRF-1 in the germline, where it is redundant with EGO-1 in a surveillance pathway that regulates transposons, pseudogenes, and cryptic loci. However, EGO-1 alone is required for the 22G-RNAs that associate with CSR-1 (Claycomb et al., 2009). In addition, van Wolfswinkel et al. (2009) implicate the β-nucleotidyl transferase CDE-1 as a specificity factor for EGO-1 in the chromosome segregation pathway, as EGO-1 can also function with RDE-3 (a CDE-1 homolog) for the biogenesis of WAGO-associated 22G-RNAs (Figure 7).

Despite clear genetic redundancy among the WAGOs, we expect that individual WAGOs normally function in distinct pathways (Figure 7). Consistent with this idea, RDE-4 and SMG-5 are required for the biogenesis of distinct and overlapping subsets of WAGO-dependent 22G-RNAs. Domeier et al. (2000) showed that the exo-RNAi response is short-lived in *smg-2*, *-5*, and *-6* mutants. Our findings extend their work and provide molecular insight into the link between the NMD pathway and RNAi. Both studies connect 22G-RNA biogenesis to the translation apparatus and suggest that an alternative branch of the NMD pathway exists (Behm-Ansmant et al., 2007). Perhaps SMG proteins recognize a particular characteristic of 22G-RNA target transcripts and recruit the RdRP machinery. Our findings indicate that the signal is unlikely to be premature termination codons. Alternatively, NMD components could function as WAGO cofactors. WAGOs lack the catalytic residues important for Slicer activity and are not expected to cleave a target in vivo, suggesting that alternative turnover mechanisms are involved in silencing. It is interesting to note that the PIN domains of both SMG-5 and SMG-6 are structurally related to RNase H (Glavan et al., 2006), but only SMG-6 retains the catalytic residues important for the endonucleolytic cleavage that initiates NMD (Eberle et al., 2009; Huntzinger et al., 2008). If WAGOs do indeed lack catalytic activity, perhaps SMG-6 could provide the endonuclease activity that reinforces the 22G-RNA amplification cycle for this set of targets. Additional complexity and bifurcation/convergence of WAGO-dependent 22G-RNA pathways is likely to emerge as we identify new factors required for 22G-RNA biogenesis.

The Mutator phenotypes of *mut-7*, *rde-3*, and *drh-3* could, in part, result from defects in WAGO-dependent chromatin silencing (Figure 7). WAGO-12/NRDE-3 is a nuclear WAGO required for cotranscriptional silencing (Guang et al., 2008) but not for the accumulation of 22G-RNAs. However, mutants that block the biogenesis of 22G-RNAs prevent nuclear localization of NRDE-3 and exacerbate derepression of NRDE-3 targets, suggesting additional WAGOs are involved in a parallel, post-transcriptional silencing pathway. It will be of interest in the future

to dissect the potential role of 22G-RNAs in different chromatin-mediated silencing pathways.

Maternal Small RNAs

In several respects, including transposon control and DCR-1-independent small RNA biogenesis, the WAGO 22G-RNA system is analogous to the *Drosophila* and vertebrate piRNA pathways. Furthermore, we have shown that the WAGO and CSR-1 22G-RNA systems are maternal and that factors involved in these pathways localize to germline P granules (Claycomb et al., 2009), which are thought to function in the repression and storage of maternal mRNAs (Rajyaguru and Parker, 2009). This appears to be a common feature of germline small RNA pathways in animals, as PIWI family members also localize to P granules or nuage (Batista et al., 2008; Li et al., 2009; Malone et al., 2009). In each case, the localization of AGO proteins to P granules appears to be dependent on small RNA biogenesis. For the CSR-1 pathway, the P granule structure itself seems to be dependent on small RNA biogenesis (Claycomb et al., 2009). The close association between cytoplasmic P granules and nuclear pores would allow AGO systems to survey the entire transcriptome as RNAs exit the nucleus and enter the P granule, reinforcing both the biogenesis and regulatory functions of small RNAs.

Maternal small RNAs function in a number of epigenetic programs, from transposon silencing (Brennecke et al., 2008; Tam et al., 2008; Watanabe et al., 2008) to imprinting (Davis et al., 2005) to the maternal-zygotic transition (Giraldez et al., 2006; Lykke-Andersen et al., 2008). Our findings suggest that 22G-RNAs in *C. elegans* mirror the expression of many germline-expressed RNAs, including those destined for expression as well as silencing. Thus, 22G-RNAs and their AGO partners provide versatile regulators of both physical and epigenetic inheritance.

EXPERIMENTAL PROCEDURES

Worm Strains

C. elegans culture and genetics were essentially as described (Brenner, 1974). Strains and alleles used in this study are listed in Supplemental Experimental Procedures.

Molecular Biology

Standard methods were used for preparing and processing samples for RNA and protein analyses. Antibodies used in this study include: (1) affinity-purified anti-DRH-3 polyclonal; (2) anti-DCR-1 polyclonal (Duchaine et al., 2006); (3) affinity-purified anti-EKL-1 (Claycomb et al., 2009); (4) affinity-purified anti-RRF-1 polyclonal; (5) affinity-purified anti-EGO-1 polyclonal; (6) anti-FLAG M2 monoclonal (Sigma); (7) HRP-conjugated anti-rabbit IgG secondary antibodies (Jackson Laboratory). Additional details, including peptides used to raise DRH-3, RRF-1, and EGO-1 sera, are provided in Supplemental Experimental Procedures.

Small RNA Cloning Method

Small RNA amplicons were prepared essentially as described (Ambros et al., 2003; Lim et al., 2003), with some modifications. Libraries were sequenced using an Illumina Genome Analyzer at the Center for AIDS Research, UMass Medical School (Worcester, MA) or at the Center for Genome Research and Biocomputing, Oregon State University (Corvallis, OR). Detailed methods for cloning small RNAs are provided in Supplemental Experimental Procedures.

Bioinformatics

Small RNA sequences were processed and mapped to the *C. elegans* genome (WormBase release WS192) as well as Repbase (13.07) using Perl (5.8.6). Details are included in Supplemental Experimental Procedures.

Real-Time PCR and Tiling Microarray Analyses

qRT-PCR was performed as described (Batista et al., 2008). Primer sequences and additional details are provided in Table S5 and Supplemental Data. Tiling Analysis Software 1.1.2 (Affymetrix) was used to process raw tiling array data (three slides each for *drh-3(ne4253)* and wild-type) with the settings: bandwidth = 100; intensities = PM/MM. Custom Perl scripts were used to: convert Affymetrix probe coordinates (release WS170) to release WS192 coordinates; identify probe signals with p value < 0.1 that were present in both data sets; and calculate gene expression values, defined as the geometric mean of probe signals within a gene. Actin was used to normalize expression values prior to comparison.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Discussion, Supplemental Experimental Procedures, Supplemental References, 12 figures, and five tables and can be found online at http://www.cell.com/molecular-cell/supplemental/S1097-2765(09)00678-9.

REFERENCES

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr. Biol. *13*, 807–818.

Anders, K.R., Grimson, A., and Anderson, P. (2003). SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. EMBO J. *22*, 641–650.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. EMBO J. *26*, 5007–5019.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., 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.

Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Saulière, J., Wittkopp, N., and Izaurralde, E. (2007). mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. FEBS Lett. *581*, 2845–2853.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089–1103.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science *322*, 1387–1392.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71–94.

Bühler, M., Haas, W., Gygi, S.P., and Moazed, D. (2007). RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. Cell *129*, 707–721.

Bühler, M., Spies, N., Bartel, D.P., and Moazed, D. (2008). TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. Nat. Struct. Mol. Biol. *15*, 1015–1023.

Chen, C.C., Simard, M.J., Tabara, H., Brownell, D.R., McCollough, J.A., and Mello, C.C. (2005). A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr. Biol. *15*, 378–383.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van Wolfswinkel, J.C., Chaves, D.A., Shirayama, M., Mitani, S., Ketting, R.F., et al. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors target germline genes and are required for holocentric chromosome segregation. Cell *139*, 123–134.

Colmenares, S.U., Buker, S.M., Bühler, M., Dlakić, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Mol. Cell *27*, 449–461.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R., et al. (2008). An endogenous small interfering RNA pathway in Drosophila. Nature *453*, 798–802.

Davis, E., Caiment, F., Tordoir, X., Cavaillé, J., Ferguson-Smith, A., Cockett, N., Georges, M., and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. Curr. Biol. *15*, 743–749.

Ding, S.W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. Cell *130*, 413–426.

Domeier, M.E., Morse, D.P., Knight, S.W., Portereiko, M., Bass, B.L., and Mango, S.E. (2000). A link between RNA interference and nonsense-mediated decay in Caenorhabditis elegans. Science *289*, 1928–1931.

Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G., et al. (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343–354.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210.

Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A link between mRNA turnover and RNA interference in Arabidopsis. Science *306*, 1046–1048.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Natl. Rev. *10*, 94–108.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., et al. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077–1081.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75–79.

Glavan, F., Behm-Ansmant, I., Izaurralde, E., and Conti, E. (2006). Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. EMBO J. *25*, 5117–5125.

Grimson, A., O'Connor, S., Newman, C.L., and Anderson, P. (2004). SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans. Mol. Cell. Biol. *24*, 7483–7490.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science *321*, 537–541.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. Science *315*, 1587–1590.

Hall, I.M., Noma, K., and Grewal, S.I. (2003). RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. Proc. Natl. Acad. Sci. USA *100*, 193–198.

Herr, A.J., Molnàr, A., Jones, A., and Baulcombe, D.C. (2006). Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. Proc. Natl. Acad. Sci. USA *103*, 14994–15001.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA *14*, 2609–2617.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol. *5*, e57.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133–141.

Klattenhoff, C., and Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. Development *135*, 3–9.

Lee, R.C., Hammell, C.M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. RNA *12*, 589–597.

Lee, S.R., and Collins, K. (2007). Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs. Nat. Struct. Mol. Biol. *14*, 604–610.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., et al. (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.

Liang, L., Diehl-Jones, W., and Lasko, P. (1994). Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities. Development *120*, 1201–1211.

Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003). The microRNAs of Caenorhabditis elegans. Genes Dev. *17*, 991–1008.

Lykke-Andersen, K., Gilchrist, M.J., Grabarek, J.B., Das, P., Miska, E., and Zernicka-Goetz, M. (2008). Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. Mol. Biol. Cell *19*, 4383–4392.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522–535.

Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature *457*, 413–420.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789–802.

Nakamura, M., Ando, R., Nakazawa, T., Yudazono, T., Tsutsumi, N., Hatanaka, N., Ohgake, T., Hanaoka, F., and Eki, T. (2007). Dicer-related drh-3 gene

functions in germ-line development by maintenance of chromosomal integrity in Caenorhabditis elegans. Genes Cells *12*, 997–1010.

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E.C. (2008). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nat. Struct. Mol. Biol. *15*, 581–590.

Page, M.F., Carr, B., Anders, K.R., Grimson, A., and Anderson, P. (1999). SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast. Mol. Cell. Biol. *19*, 5943–5951.

Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science *315*, 241–244.

Parker, J.S., Roe, S.M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. Nature *434*, 663–666.

Rajyaguru, P., and Parker, R. (2009). CGH-1 and the control of maternal mRNAs. Trends Cell Biol. *19*, 24–28.

Rocheleau, C.E., Cullison, K., Huang, K., Bernstein, Y., Spilker, A.C., and Sundaram, M.V. (2008). The Caenorhabditis elegans ekl (enhancer of ksr-1 lethality) genes include putative components of a germline small RNA pathway. Genetics *178*, 1431–1443.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell *127*, 1193–1207.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. Cell *125*, 287–300.

Sijen, T., and Plasterk, R.H. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature *426*, 310–314.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465–476.

Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science *315*, 244–247.

Siomi, H., and Siomi, M.C. (2009). On the road to reading the RNA-interference code. Nature *457*, 396–404.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (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.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434–1437.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123–132.

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861–871.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., et al. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534–538.

van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.M., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135–148.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672–676.

Vought, V.E., Ohmachi, M., Lee, M.H., and Maine, E.M. (2005). EGO-1, a putative RNA-directed RNA polymerase, promotes germline proliferation in parallel with GLP-1/notch signaling and regulates the spatial organization of nuclear pore complexes and germline P granules in Caenorhabditis elegans. Genetics *170*, 1121–1132.

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.

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747–757.