Supporting Information

Han et al. 10.1073/pnas.0906378106

SI Text

SI Results

Requirement of target mRNA transcript for 26G RNA biogenesis.

deps-1 is a gene whose 3'UTR appears to be targeted by a class I sperm 26G RNA (26G-S4) (Fig. S7A). Two alleles of *deps-1* (*bn121* and *bn124*) introduce premature stop codons into the gene and destabilize *deps-1* transcripts (Fig. S7A) (1). In both alleles, the expression of 26G-S4 is significantly depleted (>10-fold), while expression of other 26G RNAs that do not target *deps-1* (26G-S5, -S6) is not affected, supporting the requirement of *deps-1* transcript as a template for 26G-S4 production (Fig. S7B). We attempted to rescue 26G-S4 expression by crossing the *deps-1* mutants into the *smg-1(r861)* background, which stabilizes transcripts with premature stop codons that are degraded by the nonsense-mediated mRNA decay pathway (2). While the expression of *deps-1* mRNA in the double mutants is still below wild-type levels, we observed a noticeable increase in 26G-S4 expression in one (*deps-1(bn121); smg-1(r861)*), but not the other (*deps-1(bn124); smg-1(r861)*), double mutant (Fig. S7D); this is likely because the expression of *deps-1* mRNA remains below a threshold level for 26G-S4 synthesis or detection in the *deps-1(bn124); smg-1(r681)* double mutant (Fig. S7C).

5' monophosphorylation of 26G RNAs.

In *C. elegans*, during exogenous RNAi, a similar RdRP-mediated process programmed by *rrf-1* generates secondary siRNAs to amplify the silencing signal (3-5). These secondary siRNAs start with a guanine nucleotide and are triphosphorylated at 5' end (5'-PPP). However, although 26G RNAs require the RRF-3 RdRP, they are suitable substrates for T4 RNA ligase-mediated 5' linker ligation (Fig. S8), suggesting that most 26G RNAs possess a 5' monophosphate group (3, 4). These findings further support our original small RNA cloning procedure that identified the 26G RNAs based on a ligation reaction that selected small RNAs containing a 5' monophosphate group.

The biogenesis of 26G RNAs is likely Dicer-dependent.

The presence of a 5' monophosphate on a small RNA is a signature for Dicer processing with notable exceptions (*e.g.* 21U RNAs) (6). Because 26G RNAs appear to be 5' monophosphorylated, we asked if they are processed by Dicer. *C. elegans* encodes a single Dicer ribonuclease, *dcr-1*, which is essential for germline development and viability (7-9). Homozygous *dcr-1(ok247)* null animals that are produced by *dcr-1*/+ heterozygotes live until adulthood but exhibit pleiotropic defects including complete sterility, abnormal vulval structures, and unfertilized oocytes (7) (Fig. S9A). At the young adult stage, levels of both microRNAs (*let-7* and miR-1) and 26G RNAs (26G-O1, -O2, -O3, and -194) significantly decrease in the *dcr-1* null mutant, relative to the *dcr-1* heterozygous animals (Fig. S9C). The degree to which the expression of microRNAs and 26G RNAs is compromised correlates with the severity of *dcr-1* phenotypes (Fig. S9A) and with the level of maternal *dcr-1* mRNA remaining in the *dcr-1* mutant (Fig. S9B). Thus, *dcr-1* null animals with the severe phenotype of bursting exhibit the lowest level of maternal *dcr-1* mRNA, microRNAs, and 26G RNAs. Surprisingly, the expression of 21U RNAs (21UR-342 and 21UR-684), which was reported to be *dcr-1*-independent (6), also decreases to a similar degree as the microRNAs and 26G RNAs in the young adult *dcr-1* mutant (Fig. 9C). The germline itself is required for the biogenesis of 26G and 21U RNAs (Fig. 1E), but dispensable for the somatic

expression of the microRNAs *let-7* and miR-1. Therefore, we next asked if we could discriminate *dcr-1*-dependence among the three types of small RNAs during an earlier period of germline development, the L4 larval stage, when the cumulative effects of the *dcr-1* mutation on germline development are less severe (Fig. S9D). In agreement with the results in the young-adult stage, the *dcr-1* homozygotes at L4 stage again exhibited decreased expression of microRNAs (*let-7* and miR-1) and 26G RNAs (26G-S4, -S5, and -S6). However, the expression of 21U RNAs (21UR-342 and 21UR-684) does not decrease in the L4-stage *dcr-1* homozygotes, supporting previous observations of *dcr-1*-independence (6); on the contrary, their levels appear to increase. Taken together, our data suggest that the biogenesis of 26G RNAs likely requires *dcr-1*. However, because the 26G RNAs are generated in the germline, we cannot conclusively rule out the possibility that the decrease in 26G RNA expression may be an indirect consequence of defects in germline development exhibited by the *dcr-1* (-/-) mutant.

SI Methods

Sequence processing.

All raw sequences (consolidating both 454 and Solexa) were processed with a custom Perl script to remove linker sequences and then mapped against the WS190 *C. elegans* genome using BLAST (10). Sequences matching the genome with 0-2 mismatches were retained. Reads not matching the genome were mapped against Expressed Sequence Tags (EST) using BLAST to identify sequences that span exon-exon junctions. For reads matching more than one genomic locus, counts were normalized according to Ruby *et al.* (11). For example, if a sequence had 20 reads and matched 2 genomic loci, each locus was assigned 10 reads. For all endo-siRNA analyses, reads corresponding to microRNAs (11), 21U RNAs (6, 12), and putative degradation products of non-coding RNAs (*i.e.* rRNAs, tRNAs, snRNAs, snoRNAs) were identified and excluded.

Genomic mapping of 26G RNAs.

As outlined in Fig. S1, we applied sequential filters to retain 26G RNAs with ≥ 2 reads in the 11 sequenced libraries and mapped them sequentially to WormBase (WS190) and predicted gene models (Twinscan and Genefinder in WS190). Because 3'UTR regions are not well annotated, reads immediately downstream (within 500bp) of stop codons were annotated as overlapping with 3'UTR, which agrees well with the distribution of known 3' UTR lengths of annotated genes in WormBase (Fig. S2). The remaining intergenic 26G RNA sequences (23.3%) may also target genes yet to be identified by WormBase gene annotations and predictions.

Mapping 26G RNAs to the AceView transcriptome.

As outlined in Fig. S3, 25-27 nucleotide sequences from 454 and Solexa were aligned in a hierarchical order using the following gene models: 1. WormBase (WS190) non-coding models, including rRNA, tRNAs, miRNAs, snoRNAs and other small RNAs; 2. WS190 protein coding CDS models; 3. AceView transcriptome (13) (www.aceview.org, June 2009 version). AceView database integrates all cDNAs from the following sources: 1. NCBI GenBank and dbEST (14); 2. hand-edited cDNA sequencing traces from the groups of Kohara (13, 15), Vidal (16), Exelixis, Martin (17), Piano (18), and the NCBI Trace repository; 3. Waterston deep sequencing of the L2, L3, L4 and young adult hermaphrodite transcriptome (Solexa, SRA003622, (19)); 4. *C. elegans* 3'UTRome sequences (http://128.122.61.5/cgi-bin/UTRome/utrome.cgi). Up to 2 mismatches were allowed over the entire sequence alignment; sequences mapping multiple loci were normalized and loci with 1 read were filtered from the analyses.

The remaining sequences were aligned to the WS190 genome, distinguishing successively sequences in '26G nests', introns, and intergenic regions. 26G nests are genomic regions ultra-rich in 26G RNAs and defined operationally in the following way: 26G RNAs are counted at each genomic position and allowed to add up recursively with neighboring sites closer than 1 kb away. Any group reaching 100 reads was designated as a 26G nest and is extended until it is separated by at least 1 kb on either side from the nearest 26G site. There are 80 such regions genome-wide, their average size is 1.55 kb, about the size of a gene, and all but one are antisense to (often partial) gene models, many of which are non-coding.

Nighty-nigh percent of the 26G RNAs target genes: 19% target WormBase CDS, 8% structural RNA models (in this order: nuclear and mitochondrial rRNA, mir-64 and mir-229, small nucleolar RNA CeN65 and various Pro, Leu and Gly tRNAs), an additional 62% target mRNAs already annotated in NCBI AceView and the remaining 10% lie within 1 kb from an annotated gene, in one of the nests. 26G RNAs exhibit a strong antisense bias (90% are antisense to mRNAs, 8% are sense to structural RNAs, and the remaining 2% appear intergenic, intronic or ambiguous) (Fig. S3). In addition, almost all 26G RNAs targeting genes map to exons (91.0%) or span exon-exon junctions (7.5%), leaving only 1.5% in currently un-annotated exons or introns. This strongly suggests that mature mRNAs are the main targets of 26G RNAs. (Fig. S3).

Cluster analysis of 26G RNAs.
26G RNAs (≥ 2 total reads) were clustered using Cluster 3.0 software (copyright Stanford University, 1998-99) and visualized using Java TreeView (open source). Clusters of the class I sperm 26G RNAs and the class II oocyte/embryo 26G RNAs were extracted from Java TreeView.

Target analysis of 26G RNAs.
Targets of class I sperm 26G RNAs and class II oocyte/embryo 26G RNAs (extracted from clustering analysis) were annotated as spermatogenesis-enriched, oogenesis-enriched, germline-intrinsic, and "others" according to Reinke *et al.* (20). For microarray analyses, raw CEL data from Asikainen *et al.* (21) were downloaded from NCBI Gene Expression Omnibus (Series GSE8659) and processed with dChip software (22). Probe intensities corresponding to targets of sperm 26G RNAs were extracted from the CEL data. The 9 fold enrichment described in main text was derived as follows: out of 72 genes that are significantly upregulated in *eri-1* (21), 36 (50%) are 26G RNA targets, representing a 9 fold ernichment, as 1118 genes (5.5% of all protein coding genes) are targeted by 26G RNAs.

Sperm, oocyte, and embryo purifications.
For sperm isolation, we used the *him-8(e1489)* strain, which increases the percentage of XO males to ~37% of the population versus ~0.2% males in the N2 wild-type strain (23). Male worms from the *him-8(e1489)* strain were further isolated from hermaphrodites by filtering through a 35 μm nylon mesh filter as described (24), resulting in >95% males in the final sample. Isolated *him-8(e1489)* males were then subjected to 20,000 psi for 1 min, 3 times, to extrude and increase the yield of purified sperm. We used the *fer-1(hc1)* strain, which produces nonfunctional sperm at 25°C (25), to obtain purified unfertilized oocytes. The *fer-1(hc1)* worms grown at 25°C were disrupted briefly in a Waring blender to release more oocytes from the body cavity. Sample purity (>95%) was inspected by DAPI staining and microscopy. Isolation of embryos from gravid adult worms was performed as described (26).

Construction of small RNA sequencing library.
RNA oligos were purchased from Dharmacon and DNA oligos from Integrated DNA Technologies. Six Solexa libraries were constructed and sequenced on the 1G Genome Analyzer (Solexa/Illumina): N2 (mixed stage), sperm, oocyte, embryo, *eri-1(mg366)*, and *glp-4(bn2)* young adults (YA). Five 454 libraries (sperm, oocyte, N2, *eri-1(mg366)*, and *glp-4*) were sequenced on the Genome Sequencer FLX system (454/Roche).

RT-qPCR analysis of small RNA and mRNA levels.
Custom small RNA Taqman assays were designed and synthesized by Applied Biosystems (27). For each reaction, 50ng of total RNA was converted into cDNA with Multiscribe Reverse Transcriptase (Applied Biosystems) following the vendor's protocol. The resulting cDNAs were analyzed by a Realplex^2 thermocycler (Eppendorf) with TaqMan Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems). Relative expression levels of small RNAs were calculated based on 2^{-ct} method (28). For oocyte/embryo 26G RNA quantifications, miR-35 was used for normalization. For sperm 26G RNA quantifications, miR-1 was used for normalization. Gene targets of each class of 26G RNAs were selected based on 26G RNA cluster analysis (described below in supplementary computational methods). For quantification of mRNA levels, 250ng to 1μg of total RNAs was converted into cDNAs with Multiscribe Reverse Transcriptase (Applied Biosystems) following the vendor's protocol. cDNAs were analyzed by a Realplex^2 thermocycler (Eppendorf) using Power Sybr Green PCR master mix (Applied Biosystems). Relative mRNA levels were calculated based on 2^{-ct} method using *act-1* for normalization.

Oligos for RT-qPCR

Gene	Forward (5' to 3')	Reverse (5' to 3')
act-1	CCAGGAATTGCTGATCGTATGCAGAA	TGGAGAGGGAAGCGAGGATAGA
C04G2.8	CGTGCTTCGACTGCAAAGAAGA	TTCTGTTGGCTTCTGCTGCG
C32E8.4	GAGCAACTTCTGCCGAAGGAA	CTTCAGGTTCTCCTTGAGCG
C40A11.10	AATGGCTCCTTGAAAAGATCG	TACATTTCCGCCACGTTGAAA
deps-1	GAAGGCTATGGCCGAAGTTCG	CAATGCGGTAACGGACAGATTT
dlc-6	CCGAAGGTTAAGCCACGTCATT	CTGCCATTGTGTATCATAATCCG
E01G4.7	GCACAAGGTTTCGTTCTTGGTG	AGTGACATCCCTTCTGATCG
F39E9.7	CCCAGTGGCCCAATTAAACG	CCCACGGCTTGTTCTTTGACA
F43E2.6	TGTAGGCGACGAGACTGATCG	TGCCGATGTTTCTGAGATGTCTT
F55B11.1	TTGATCGAGTCTCACTTTCCG	AAAGTCCACTGGTTCGTGATGAAT
F55C9.5	ACCATTGGAGCACGTAAATCAA	GGTCCTAATAATAAAGTTGCGTCG
fbxa-65	ACTTACAAGGATCAAGAAAAGCG	CCTTGACCGCTATTCCGAGAAA
fbxb-37	ATCGAAAGATGGAATACAAACCG	GACAAACATCCATCACATTCTTCG
gska-3	CGAGCAGACGACTCTGTGGAA	TTATTGAAACGCACAGTCTTCTCG
iff-1	CGAAGACCATAGAGAGTATGTCCG	CGAGCATTGCTTCGGGAAAGTA
K02E2.6	CAGTGGTACAAGTGGGAGTAAACG	AATTGGCAAGTAACTGATTCCG
K03H1.12	CAAAATTGCCACTTGTGATTCG	TCCAGTGAAGAGTGTCAAGAACCA
msp-49	ATTAACTCCTCGGCTCGCCG	AGCTTCCTTTGGGTCGAGGAC
snf-6	GGATTGTTGGCTACTGGCCG	TCAAGCCAAAGGAAGCAAAGAA
sod-1	GATCTATGGTTGTTCATGCCG	CTTCTGCCTTGTCTCCGACTCC
ssp-16	GTCATCAAACAACAATGAGTACCG	GCTCCAGCAGTGCGAGTGAT
ssp-19	GCACCGAAGGAAGACAAGCTG	GAGCCACTGCAACAAAAGCG
T05E12.8	TTCCATTTGAGGATTTTGCTACG	ATTATTTGGATGGCAGCCGATG
T08B2.12	GAAACCAATGCTCCAGTTGATAC	GATGAAAGCGATGGACGAGAAG
T25G12.11	ACGTGCTTTCTGATTCACTCCG	CATGGGTGGGATGAGAGCAC
tax-2	GATTAATCCAAGACAAGTTCCTAAATTGAT	TTCAATTCTTGAACTCCTTTGTTTTC
Tc1	AACCGTTAAGCATGGAGGTG	CACATGACGACGTTGAAACC
Tc3	GAGCGTTCACGGAGAAGAAG	AATAGTCGCGGGTTGAGTTG
tdc-1	GAACTTCGTCAGAGATTCCCG	TCTCAACGGAAGAATGGGCTTC
U6	TGGAACAATACAGAGAAGATTAGCA	CTTCACGAATTTGCGTGTCAT
W05H12.2	GCTCAAGACCAGATAATGCTTGGA	CAATCCCAAAGATTCAATACCG
Y37E11B.2	AATGGAGACTCTTCTTCCACCCG	AGCGAAGGCATTGATCTTGGTT
Y7A5A.11	CCATTACTTTCAACATGCCG	TCCTTGTTCCAGCACTAGCAGA
Y82E9BR.20	CTCCCGCTTTCTTGATGTATTG	AGTCCGAACTCATCCAAAGCAG
ZC168.6	GTCCAGTTTATGGGTTCGTGGATG	AGTCTCTTCGGCTGGCACTTC
ZC328.1	GGGCGGTCATTTCTATTGTTTG	GCCAAATTGGTCCGTAATCTTGT
ZK484.5	CCGTCAGACAACTGCTCTCCTC	GGTTGGGCTGCTTCAGAGTC

Oligos for small RNA cloning

5' RNA adaptor:	5' GUUCAGAGUUCUACAGUCCGACGAUC 3'
3' RNA adaptor:	5' pUCGUAUGCCGUCUUCUGCUUGidT 3' p = phosphate; idT = inverted deoxythymidine
RT-primer (DNA):	5' CAAGCAGAAGACGGCATACGA 3'
P7 primer (DNA):	5' CAAGCAGAAGACGGCATACGA 3'
P5 long primer (DNA):	5' AATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGA 3'

Oligos for northern blotting

21UR-1	5' GCACGGTTAACGTACGTACCA /3StarFire/ 3'
26G-O1	5' TTGAAAATAATCTACCGTTTCTGAGC /3StarFire/ 3'
26G-O2	5' CATTTGCTGCAATTATGAGTCATAAC /3StarFire/ 3'
26G-O3	5' AAAAGTATCCGACTTTCGAGTTTGTC /3StarFire/ 3'
26G-O5	5' CCCCTCTTTTCTTCTGCATTCCCATC /3StarFire/ 3'
26G-O6	5' ATGAAATGCCAGATGAATCCTTCTAC /3StarFire/ 3'
26G-S1	5' AATTATGTATTCTCGTCCTCCATAGC /3StarFire/ 3'
26G-S5	5' TACCATGTCGCTCACTGCTGATCCAC /3StarFire/ 3'
cel-miR-35	5' ACTGCTAGTTTCCACCCGGTGA /3StarFire/ 3'
cel-miR-1	5' TACATACTTCTTTACATTCCA /3StarFire/ 3'

References:

1. Spike CA, Bader J, Reinke V, & Strome S (2008) DEPS-1 promotes P-granule assembly and RNA interference in C. elegans germ cells. *Development* 135(5):983-993.
2. Hodgkin J, Papp A, Pulak R, Ambros V, & Anderson P (1989) A new kind of informational suppression in the nematode Caenorhabditis elegans. *Genetics* 123(2):301-313.
3. Pak J & Fire A (2007) Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science* 315(5809):241-244.
4. Sijen T, Steiner FA, Thijssen KL, & Plasterk RH (2007) Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* 315(5809):244-247.
5. 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(24):5007-5019.
6. Das PP*, 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(1):79-90.
7. 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(5538):2269-2271.
8. Ketting RF*, et al.* (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev* 15(20):2654-2659.
9. Grishok A*, et al.* (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* 106(1):23-34.
10. Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403-410.
11. Ruby JG*, et al.* (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell* 127(6):1193-1207.
12. Batista PJ*, et al.* (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. *Mol Cell* 31(1):67-78.
13. Thierry-Mieg D & Thierry-Mieg J (2006) AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol* 7 Suppl 1:S12 11-14.
14. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, & Sayers EW (2009) GenBank. *Nucleic Acids Res* 37(Database issue):D26-31.
15. Kohara Y (2001) Systematic analysis of gene expression of the C. elegans genome. *Tanpakushitsu Kakusan Koso* 46(16 Suppl):2425-2431.
16. Reboul J*, et al.* (2001) Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in C. elegans. *Nat Genet* 27(3):332-336.
17. Waterston R*, et al.* (1992) A survey of expressed genes in Caenorhabditis elegans. *Nat Genet* 1(2):114-123.
18. Piano F, Schetter AJ, Mangone M, Stein L, & Kemphues KJ (2000) RNAi analysis of genes expressed in the ovary of Caenorhabditis elegans. *Curr Biol* 10(24):1619-1622.
19. Hillier LW*, et al.* (2009) Massively parallel sequencing of the polyadenylated transcriptome of C. elegans. *Genome Res* 19(4):657-666.
20. Reinke V, Gil IS, Ward S, & Kazmer K (2004) Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. *Development* 131(2):311-323.
21. Asikainen S, Storvik M, Lakso M, & Wong G (2007) Whole genome microarray analysis of C. elegans rrf-3 and eri-1 mutants. *FEBS Lett* 581(26):5050-5054.

22. Li C & Wong WH (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 98(1):31-36.
23. Brenner S, Hodgkin,J , Horvitz, R . (1979) Nondisjunction mutants of the nematode C. elegans. *Genetics* (91):67-94.
24. L'Hernault SW & Roberts TM (1995) Cell biology of nematode sperm. *Methods Cell Biol* 48:273-301.
25. Ward S, Argon Y, & Nelson GA (1981) Sperm morphogenesis in wild-type and fertilization-defective mutants of Caenorhabditis elegans. *J Cell Biol* 91(1):26-44.
26. Stiernagle T (2006) Maintenance of C. elegans. *WormBook*:1-11.
27. Chen C*, et al.* (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33(20):e179.
28. Nolan T, Hands RE, & Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. *Nat Protoc* 1(3):1559-1582.

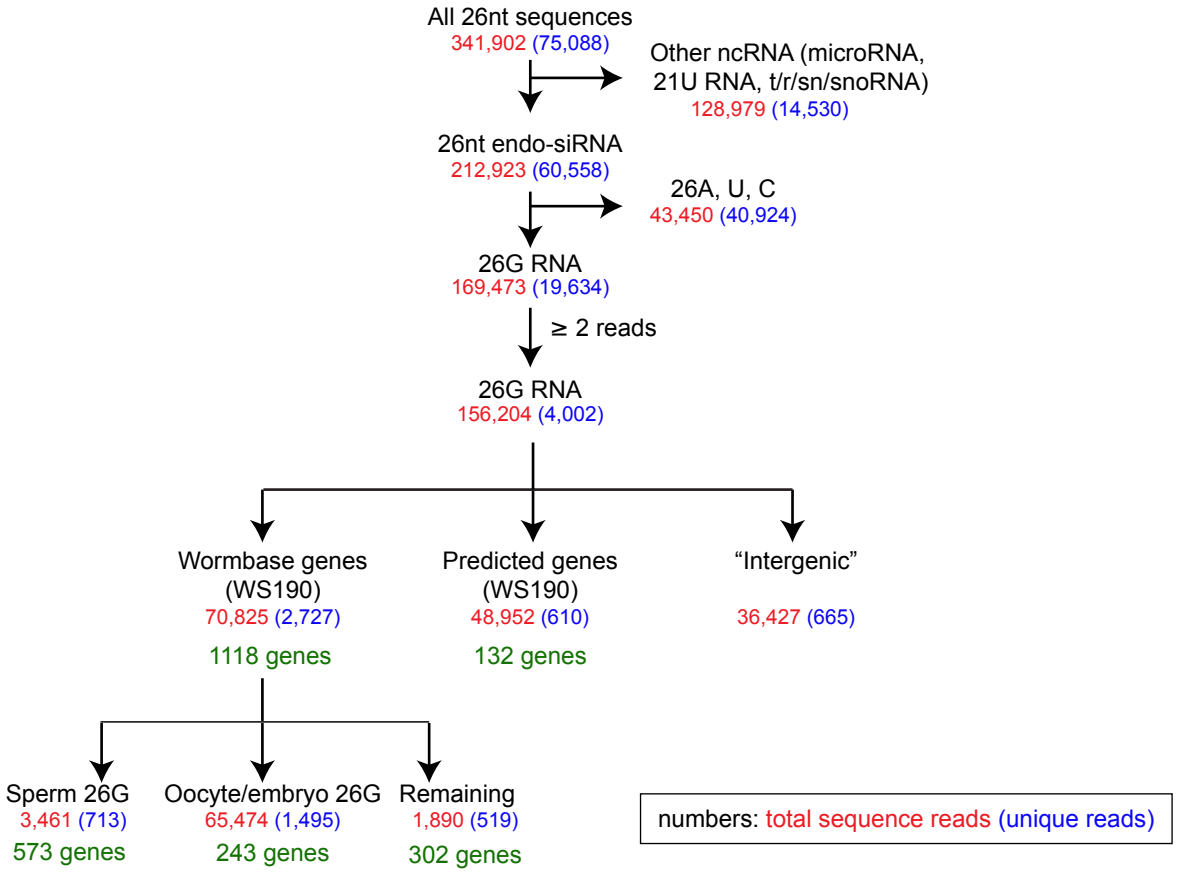


Fig. S1. Computational pipeline for 26G RNA annotations. All 26nt genome BLAST hits were extracted from our datasets. Sequences matching noncoding RNAs (i.e. tRNAs, rRNAs, snRNAs, snoRNAs) and other classes of small RNAs (microRNAs, 21U RNAs) were identified and excluded from the analyses. Two additional filters were applied to retain sequences starting with guanine and having ≥2 sequence reads. 26G RNAs mapping within 500bp downstream of WormBase gene annotations (WS190) and gene predictions (Twinscan, Genefinder predictions from WormBase) were sequentially annotated. In sum, 1,118 WormBase-annotated genes and 132 WormBase-predicted genes were identified to be targets of 26G RNAs. 26G RNAs derived from WormBase-annotated genes were further clustered into sperm 26G RNAs (with 573 gene targets) and oocyte/embryo 26G RNAs (with 243 gene targets).

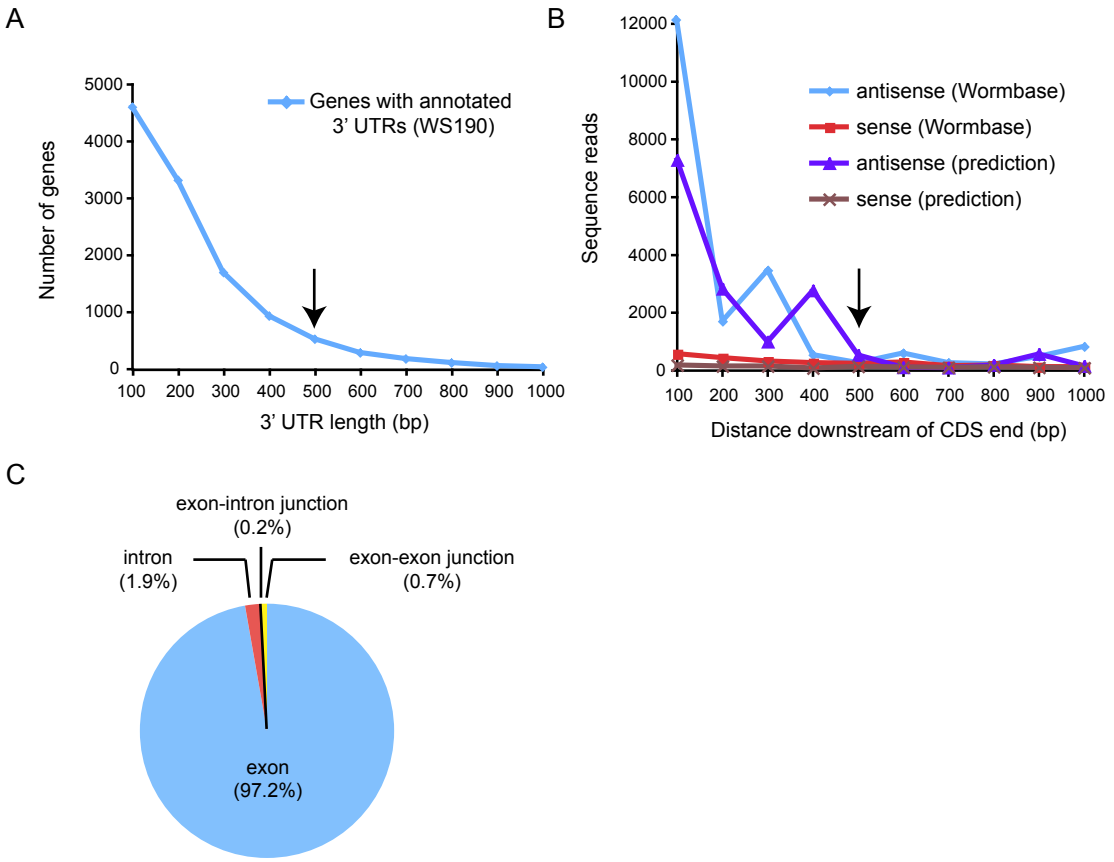
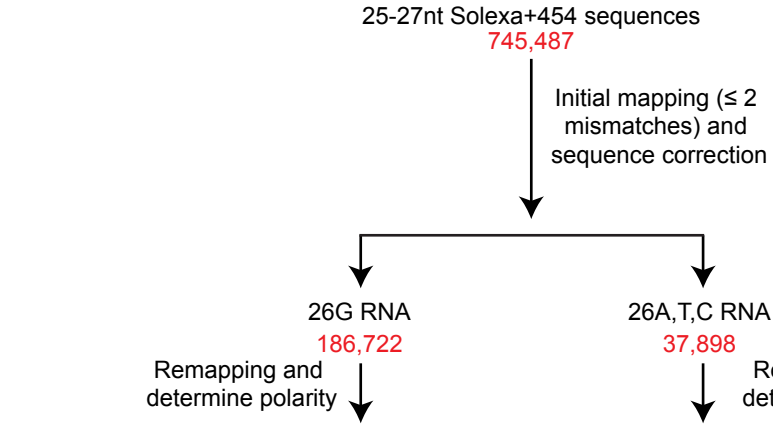


Fig. S2. Distribution and mapping of 26G RNAs. (A) The 3′ UTR length distribution of genes in WormBase. Arrow at 500nt indicates the 95% cutoff. (B) Number of 26G RNA reads that mapped within every 100bp up to 1Kb downstream of the ends of the coding sequences (stop codons) was plotted. The majority of reads are antisense to mRNAs and map within 500 bp (arrow) downstream of stop codons. (C) 26G RNAs mapping to exons and introns. 26G RNA counts matching exons, introns, exon-intron junctions and exon-exon junctions of WormBase genes were plotted. The majority of reads (97.9%) are derived from exons.



26G RNA	sense	anti-sense	ambi-guous	% anti-sense
WS190 ncRNA	15,819	13	0	0%
WS190 coding exons	1,651	25,305	2	93.9%
AceView exons	938	106327	870	98.9%
exon junction	122	8,897	0	98.6%
26G Nests	0	24,903		NA
intergenic or intronic	0	0	1,875	NA

26ATC RNA	sense	anti-sense	ambi-guous	% anti-sense
WS190 ncRNA	28,879	14	0	0%
WS190 coding exons	5,237	96	0	1.8%
AceView exons	1,539	261	134	14.5%
exon junction	403	13	0	3.1%
intergenic or intronic	0	0	1,322	NA

Fig. S3. Mapping 26G RNAs to the AceView transcriptome. 25-27nt Solexa and 454 sequences were extracted from our datasets and sequentially mapped (≤ 2 mismatches) to 1.) WS190 noncoding RNA model, 2.) WS190 CDS model, 3.) AceView transcriptome, and 4.) WS190 *C. elegans* genome. After correcting mismatches, 26G RNAs and 26A, 26T, and 26C RNAs were separated and remapped. 91% of the 26G RNAs target AceView messenger RNAs on the antisense strand, another 8% target WS190 noncoding RNAs on the sense strand, and only 1% map in intronic or intergenic regions, which probably correspond to still un-annotated transcribed regions. In contrast, 26A, 26T, and 26C RNAs mostly map to abundant noncoding RNAs with strong sense bias, indicating that they, like the remaining 26G RNAs targeting those genes, are likely degradation products.

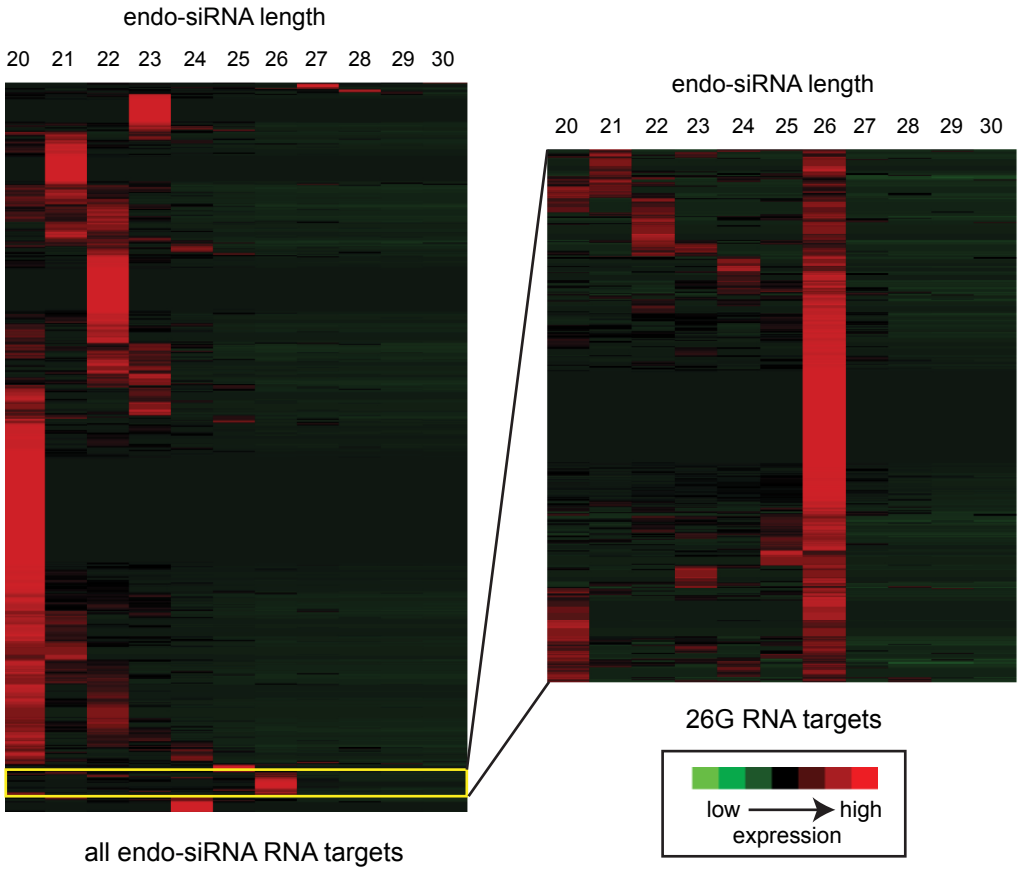


Fig. S4. 26G RNAs targets are a unique class of genes. Endo-siRNA targets (WormBase WS190) were clustered (left) based on the abundance of endo-siRNAs of different lengths. 26G RNA targets are predominantly targeted by 26G RNAs (right).

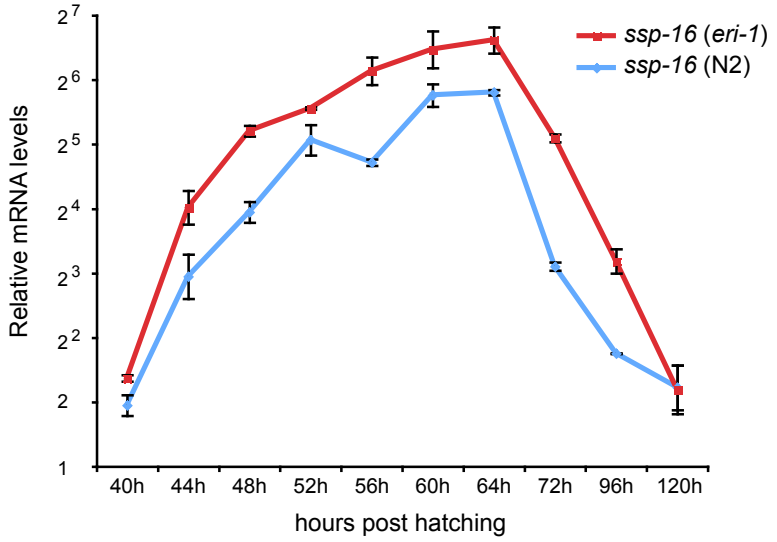


Fig. S5. *ssp-16* (a target of sperm 26G RNA) is de-repressed starting from spermatogenesis until young adulthood in the *eri-1* mutant. The X-axis represents hours post hatching at 20°C; the Y-axis indicates relative mRNA abundance in log2 scale. Relative mRNA levels were examined by RT-qPCR and normalized to *act-1*.

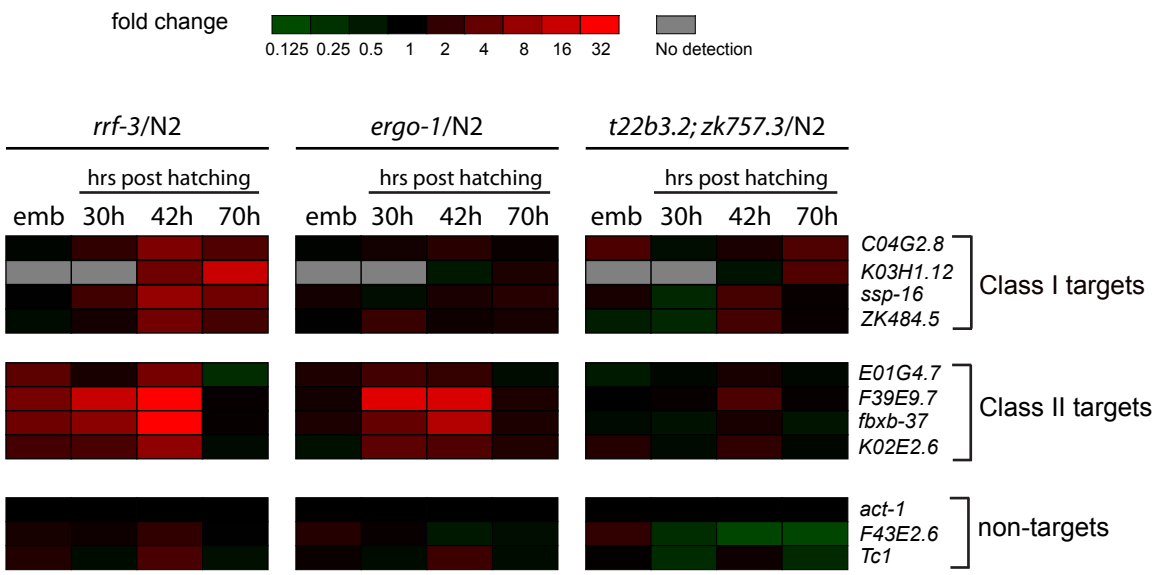


Fig. S6. Differential gene expression profiles of 26G RNA targets in N2, *rrf-3(pk1426)*, *ergo-1(tm1860)*, and the *t22b3.2(tm1155)*; *zk757.3(tm1184)* double mutant. The transcript levels of 4 targets of class I sperm 26G RNAs, 4 targets of class II oocyte/embryo 26G RNAs, and 3 non-targets were examined. For example, the class I targets *C04G2.8* and *K03H1.12* are 3-fold up-regulated at 70hrs in *t22b3.2(tm1155)*; *zk757.3(tm1184)* double mutant relative to N2 (right panel). Relative mRNA levels were examined by RT-qPCR and normalized to *act-1*. The fold up-regulation was represented according to the red-green color scheme shown (top panel).

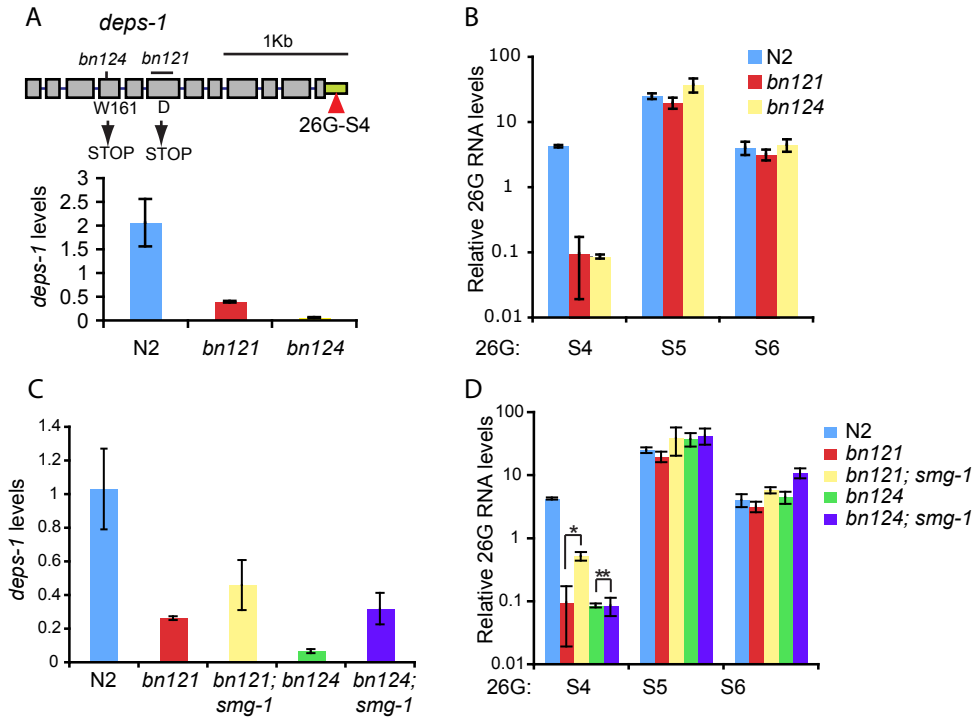
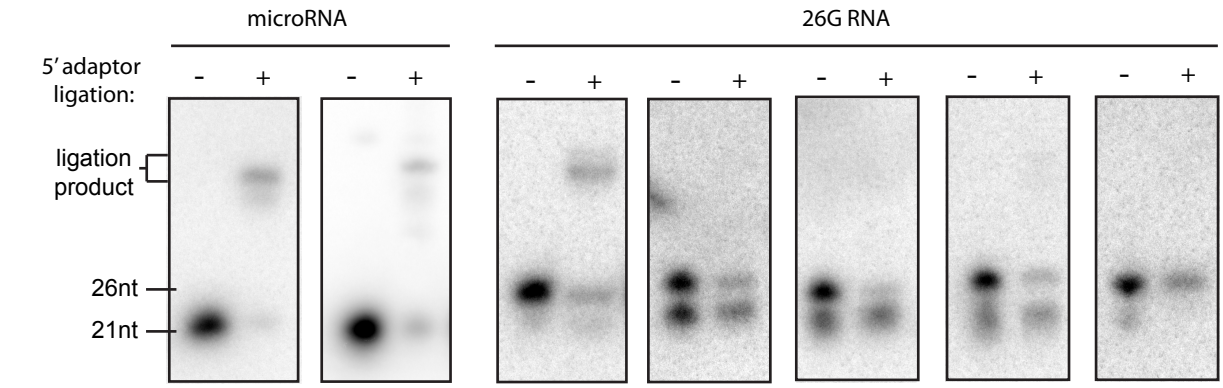


Fig. S7. Requirement of target mRNA transcript for 26G RNA biogenesis. (A) Two *deps-1* mutant alleles (*bn121* and *bn124*) harbor premature stop codons that destabilize the *deps-1* transcript. *deps-1* mRNA levels are measured by RT-qPCR and normalized to *act-1*. Error bars indicate standard deviation for replicates. (B) The expression of the class I 26G RNA 26G-S4, which is antisense to the *deps-1* 3'UTR (green), is compromised in the *deps-1* mutants, while the expression of other sperm 26G RNAs that do not target *deps-1* (26G-S5 and -S6) remains unchanged. 26G RNA levels were measured by RT-qPCR and normalized to miR-1. Error bars indicate standard deviation for replicates. (C) The expression of *deps-1* mRNA from the *deps-1* nonsense mutants (*bn121* and *bn124*) is partially restored in the nonsense decay mutant *smg-1(r861)*, but still falls below WT levels. (D) A noticeable increase of 26G-S4 levels, but not 26G-S5 and 26G-S6, is seen in the *deps-1(bn121)*; *smg-1(rr861)* double mutant, relative to the *deps-1(bn121)* single mutant (*).



	miR-1	miR-35	26G-O1	26G-O2	26G-O3	26G-O5	26G-O6
26nt ratio (- : +)	NA	NA	1 : 0.32	1 : 0.26	1 : 0.23	1 : 0.26	1 : 0.57
~21nt ratio (- : +)	1 : 0.08	1 : 0.17	1 : 0.75	1 : 0.59	1 : 0.92	1 : 0.81	1 : 0.23

Fig. S8. Depletion analysis indicates that 26G RNAs are suitable substrates for T4 RNA ligase-mediated ligation. Small RNAs (18-32nt) were isolated by PAGE and ligated to the 5′ RNA adaptor that preferentially selects for 5′ monophosphate substrates used in the small RNA cloning procedure. The ligation product and non-ligated small RNAs were resolved on 11% Urea-PAGE and subjected to northern blotting analysis. The ratio of small RNAs detected before and after linker ligation was quantified by ImageJ. 26G RNAs show similar levels of depletion after ligation compared to microRNAs miR-1 and miR-35, which are known to possess a 5′ monophosphate. For miR-1, miR-35, and 26G-O1, a higher band corresponding to ligation product can be detected. The ~21nt endo-siRNAs appear to be poor substrates for the linker ligation.

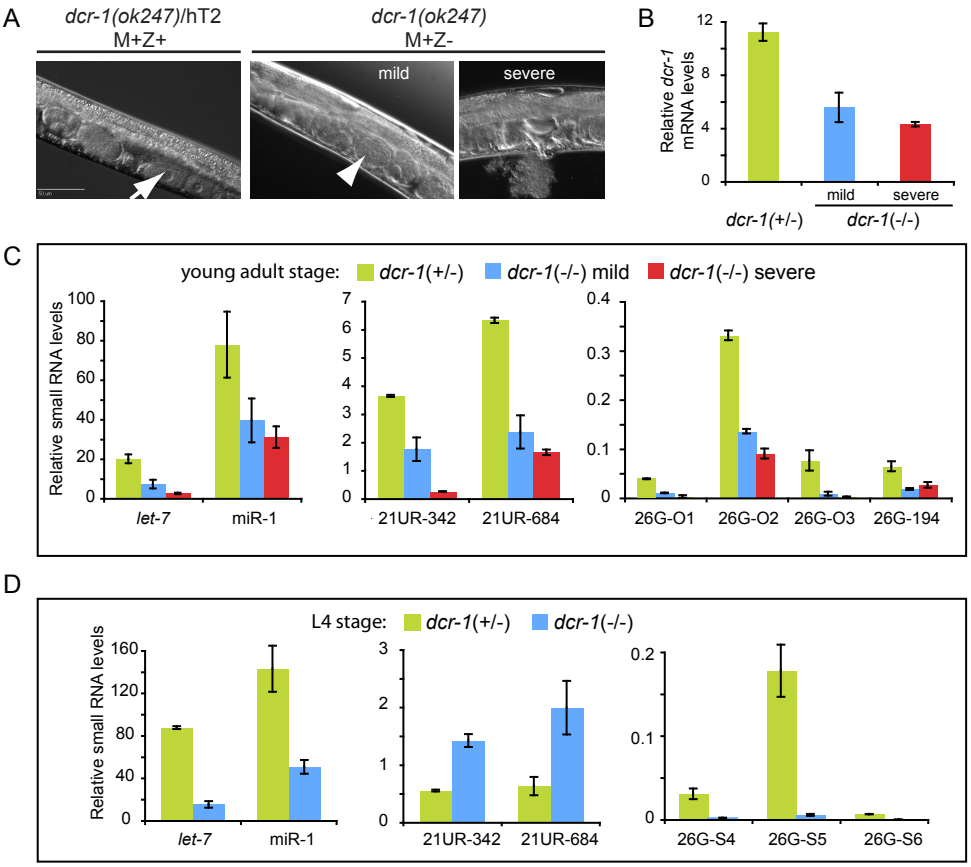


Fig. S9. Expression of 26G RNAs are likely *dcr-1*-dependent. (A) The population of *dcr-1(ok247)* null worms derived from heterozygous parents can be segregated into two groups: a "mild" phenotype group displaying abnormal oocytes (~100% penetrance) and a "severe" phenotype group displaying both abnormal oocytes and a bursting phenotype (<1% penetrance). An arrow indicates a normal oocyte in *dcr-1* (+/-) and an arrowhead indicates an abnormal oocyte in *dcr-1* (-/-). The *dcr-1* (+/-) heterozygotes retain both maternal and zygotic *dcr-1* mRNAs (M+Z+), while the *dcr-1* (-/-) homozygotes possess only maternal *dcr-1* mRNA inherited from the heterozygous parent (M+Z-). (B) *dcr-1* mRNAs are still present in *dcr-1*(-/-) animals due to maternal inheritance and slightly lower in the nulls displaying a "severe" phenotype versus the "mild" phenotype. The *dcr-1* mRNA levels were quantified by RT-qPCR and normalized to *act-1*. (C) At the young-adult stage, microRNAs (let-7 and miR-1), 21U RNAs (21UR-342, 684), and class II oocyte/embryo 26G RNAs (26G-O1, O2, O3 and 194) are all depleted in *dcr-1* (-/-) relative to *dcr-1* (+/-). Relative levels of small RNAs were quantified by Taqman RT-qPCR and normalized to *act-1*. (D) At L4 larval stage, microRNAs (let-7 and miR-1) and class I sperm 26G RNAs (26G-S4, S5, S6) are depleted in *dcr-1* (-/-) relative to *dcr-1* (+/-) while levels of 21U RNAs (21UR-342, 684) are slighted elevated in *dcr-1* (-/-).

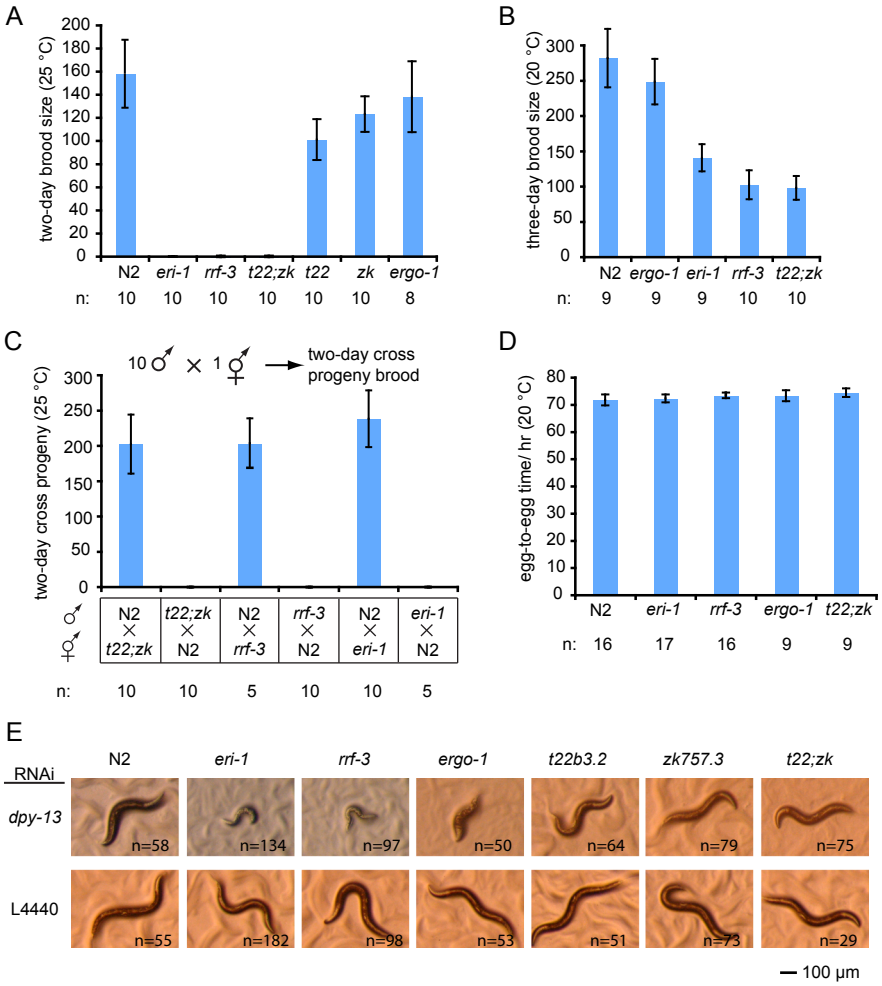


Fig. S10. Phenotypes of mutants defective in 26G RNAs. (A-B) The *t22b3.2(tm1155)*; *zk757.3(tm1184)* double mutant is sterile at 25°C and exhibits significant loss of fertility at 20°C. Synchronized worms were singled at L4 stage and progeny brood size was counted for the subsequent two days. N is the number of parents assayed. Error bars represent standard deviation. Alleles used in this assay: *eri-1(mg366)*, *rrf-3(pk1426)*, *ergo-1(tm1860)*, *t22b3.2(tm1155)*, *zk757.3(tm1184)*. (C) The *ts* sterility of *t22b3.2*; *zk757.3*, *eri-1*, and *rrf-3* can be fully rescued by crossing to WT males. For each cross, 10 males were crossed with 1 hermaphrodite, and two day cross progeny brood was scored. (D) N2, *t22b3.2(tm1155)*; *zk757.3(tm1184)*, *eri-1*, and *rrf-3* have similar egg-to-egg time at 20°C. The egg-to-egg time is the period it takes for fertilized eggs to develop into reproductive adults and produce fertilized eggs of their own. (E) The *t22b3.2(tm1155)*; *zk757.3(tm1184)* double mutant does not display an enhanced RNAi phenotype. Synchronized L1 worms of indicated genotypes were subjected to feeding RNAi of *dpy-13* or control vector. L4 and young adult worms were examined for the severity of dumpy phenotype. A moderate dumpy phenotype was observed in N2, *t22b3.2(tm1155)*, *zk757.3(tm1184)*, and the *t22b3.2(tm1155)*; *zk757.3(tm1184)* double mutant. In contrast, RNAi inactivation of *dpy-13* in *eri-1(mg366)*, *rrf-3(pk1426)*, and *ergo-1(tm1860)* generated a severe dumpy phenotype, indicating hypersensitivity to exogenous RNAi of *dpy-13*.