



Published as: *Mol Cell*. 2008 July 11; 31(1): 67–78.

PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*

Pedro J. Batista^{8,1,3}, Graham Ruby^{8,4}, Julie M. Claycomb¹, Rosaria Chiang⁴, Noah Fahlgren^{5,6}, Kristin D. Kasschau^{5,6}, Daniel A. Chaves¹, Weifung Gu¹, Jessica J. Vasale¹, Shenghua Duan¹, Darryl Conte Jr¹, Shunjun Luo⁷, Gary P. Scroth⁷, James C. Carrington^{5,6}, David P. Bartel^{2,4}, and Craig C. Mello^{1,2}

¹Program in Molecular Medicine University of Massachusetts Medical School, Worcester, MA 01605, USA.

²Howard Hughes Medical Institute

³Gulbenkian PhD Programme in Biomedicine, Rua da Quinta Grande, 6, 2780–156, Oeiras, Portugal.

⁴Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

⁵Center for Gene Research and Biotechnology, Oregon State University Corvallis, Oregon 97331

⁶Department of Botany and Plant Pathology Oregon State University Corvallis, Oregon 97331

⁷Illumina, Inc., Hayward, CA 94545

SUMMARY

In metazoans Piwi-related Argonaute proteins have been linked to germ-line maintenance, and to a class of germ-line enriched small RNAs termed piRNAs. Here we show that an abundant class of 21-nucleotide small RNAs (21U-RNAs) are expressed in the *C. elegans* germ line, interact with the *C. elegans* Piwi-family member, PRG-1, and depend on PRG-1 activity for their accumulation. The PRG-1 protein is expressed throughout development and localizes to nuage-like structures called P-granules. Although 21U-RNA loci share a conserved upstream sequence motif, the mature 21U-RNAs are not conserved and, with few exceptions, fail to exhibit complementarity or evidence for direct regulation of other expressed sequences. Our findings demonstrate that 21U-RNAs are the piRNAs of *C. elegans* and link this class of small RNAs and their associated Piwi Argonaute to the maintenance of temperature-dependent fertility.

INTRODUCTION

Diverse organisms utilize sequence-specific gene regulatory pathways that share features with RNA interference (RNAi). The effector complex in all RNAi-related pathways consists of a single-stranded small RNA, and a member of the AGO protein family, which binds small-RNA termini, leaving internal nucleotides accessible for base-pairing interactions with target sequences. In canonical RNAi pathways, double-stranded RNA (dsRNA) is processed

© 2008 Elsevier Inc. All rights reserved.

Corresponding Authors: Craig C. Mello <craig.mello@umassmed.edu> and David P. Bartel <dbartel@wi.mit.edu>.

⁸These authors contributed equally

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

by members of the Dicer family of multifunctional ribonucleases into 21 to 24 nucleotide (nt) short-interfering (siRNAs) that interact with and guide AGO proteins to complementary target sequences in the cell [reviewed in (Hutvagner and Simard, 2007)].

Most animals have an additional AGO sub-family called Piwi. *C. elegans* has two Piwi-related genes (named *prg-1* and *prg-2*) that like Piwi family members from a number of animal species, have been implicated in germ-line maintenance and fertility (Klattenhoff and Theurkauf, 2008). Two classes of Piwi-interacting RNAs (piRNAs) have been identified, including (i) repeat-associated piRNAs (originally annotated as rasiRNAs) that appear to target transposons, and (ii) a second more mysterious class of piRNAs with no known targets (Lin, 2007). This class of piRNAs is extremely abundant in small-RNA fractions isolated from pachytene-stage mouse spermatocytes: over 80,000 distinct species are derived from large genomic clusters of up to 200kb (Aravin et al., 2006; Grivna et al., 2006; Girard et al., 2006; Girard et al., 2006; Lau et al., 2006). These clusters exhibit a marked strand asymmetry, as though the piRNAs within a region are all processed from one large transcript or two divergent transcripts.

Studies in *C. elegans* have identified several classes of endogenously expressed small RNAs (Ambros et al., 2003; Ruby et al., 2006) but which, if any, of these represent piRNAs has yet to be determined. One class of small RNAs, termed 21U-RNAs, shares several characteristics with the piRNAs of flies and mammals, including an overwhelming bias for a 5' uracil, a 5' monophosphate, and a 3' end that is modified and resistant to periodate degradation (Ruby et al., 2006; Ohara et al., 2007; Saito et al., 2007; Horwich et al., 2007; Kirino and Mourelatos, 2007). However, 21U-RNAs are shorter than piRNAs in flies and mammals, and their genomic organization is very different, with 21U-RNAs deriving from what appear to be thousands of individual, autonomously expressed loci broadly scattered in two large regions of one chromosome.

Here we show that 21U-RNAs are expressed in the germ line and that their accumulation depends on the wild-type activity of PRG-1. We show that PRG-1 localizes to germ-line P-granules and that 21U-RNAs co-immunoprecipitate with PRG-1 from worm lysates. Our analysis identifies many new 21U-RNAs, bringing the total number of 21U-RNA loci to 15,722 and confirms the expression of many 21U-RNA loci previously predicted based only on the presence of an upstream sequence motif. Like the abundant pachytene piRNAs found in mammals, 21U-RNAs encode remarkable sequence diversity and yet lack obvious targets. Although we identify one example of a transposon-directed 21U-RNA, our findings suggest that piRNA complexes of worms, charged with the remarkable sequence diversity encoded by 21U-RNAs, are likely to provide other essential germ-line functions.

RESULTS

Identification of over 15,000 unique 21U-RNA species in *C. elegans*

We used Solexa sequencing technology (Seo et al., 2004) to generate 29,112,356 small-RNA cDNA reads that perfectly matched the *C. elegans* genome. Among these we identified 971,981 reads from 15,458 unique loci with properties similar to previously defined 21U-RNA loci (Ruby et al., 2006). These new reads matched 95.1% of the 5,454 previously sequenced 21U-RNAs and 78.3% of the 10,644 previously predicted 21URNAs (Ruby et al., 2006) and brought the total number of unique experimentally confirmed 21U-RNA loci to 15,722.

A common characteristic of 21U-RNA loci is the presence of an upstream sequence motif (Figure 1A, (Ruby et al., 2006)). As previously observed, RNA species 21nts in length could be separated into two distinct sets based on the motif scores of their genomic loci (Figure

1B). Species with a high motif score also tended to exhibit the other essential features, including 21nt length and 5' U nucleotide, that together define the 21U-RNA class (Supplemental Figure 1 A-C).

21U-RNAs with strong upstream motif matches were concentrated in two broad regions along chromosome IV (Figure 1C and Ruby et al., 2006). Supporting the potential importance of this motif in 21U-RNA biogenesis, the motif score strongly correlated with the magnitude of 21U-RNA, as indicated by the number of sequenced reads in our data sets (Figure 1D). Despite the presence of many high-scoring 21U-RNA motifs in orthologous regions of the *C. briggsae* genome, the 21U-RNA sequences themselves were not conserved. Even in rare cases in which the core of the upstream motif was perfectly aligned to a high-scoring motif within a syntenic region of the *C. briggsae* genome (Blanchette et al., 2004), the sequence of the consequent 21U-RNA was essentially nonconserved (Figure 1E). Only approximately 6% of the 21U-RNA loci and/or motifs were unambiguously aligned within syntenic regions in *C. briggsae*. In these few cases, this was often due to overlap with annotated coding exons, which rarely contain 21U-RNAs (Supplemental Figure 1D). The only portion of the 21U-RNA flanking regions with elevated conservation frequencies above background was the 8nt core of the upstream motif (Supplemental Figure 1E).

21U-RNAs are expressed in the *C. elegans* germ line

The developmental dynamics of 21U-RNA expression were examined by Northern blot analysis using probes specific for 21U-RNA-1 and 21U-RNA-3442. Both small RNAs were expressed at low levels from the L1 to L3 stage, began to accumulate to high levels during the L4 stage, and reached maximal expression in the young-adult and gravid-adult stages (Figure 2A). This pattern of expression correlated with the proliferation of the germ line, and was consistent with a germ-line origin. Both RNAs were expressed at approximately equal levels in male- or female- enriched populations (Figure 2B), but were absent in RNA samples prepared from germ-line-deficient *glp-4(bn2)* and *eft-3(q145)* mutant populations (Figure 2B). Finally, both small RNAs were present in embryos (Figure 2A), which may reflect maternal and/or paternal loading.

High throughput sequencing indicated that the developmental expression profile for the entire class of 21U-RNAs, was identical to that of 21U-RNA-1 and 21U-RNA-3442 (Figure 2C). The number of sequenced reads for each 21U-RNA species increased dramatically in late larval and adult stages. Furthermore, the number of reads was reduced (130 fold), from 5.8% to just 0.04% of total reads, in animals lacking a germline (Figure 2C). Adult hermaphrodites switch to an exclusively female mode of gametogenesis and store only 200–300 mature sperm. The relative abundance of various individual 21U-RNA species was comparable between male and adult hermaphrodite populations, suggesting that very similar 21U-RNA populations are present in germlines undergoing to oogenesis and spermatogenesis.

PRG-1 is expressed in the germ line and required for 21U-RNA accumulation

To examine whether the accumulation of 21U-RNA-1 and 21U-RNA-3442 was dependent on known components of the RNAi machinery, we systematically examined RNA prepared from mutant strains lacking specific components of the RNAi pathway. The accumulation of 21U-RNAs did not require the wild-type activities of any of the previously described RNAi pathway components, including DCR-1 (Figure 3A Left and Supplemental Figure 2).

To determine if accumulation of 21U-RNAs is dependent of any AGO proteins, we also analyzed mutant strains representing all of the *C. elegans* AGO family members, including several multiple-mutant strains. Only *prg-1* mutants lacked 21U-RNA-1 and 21U-

RNA-3442 (Figure 3A Right and data not shown). Strains mutant for *prg-2*, a nearly identical homolog of *prg-1*, did not exhibit defects in 21U-RNA expression (Figure 3A Right). We observed no defects in miRNA expression or in the expression of any of several other species of endogenous siRNA examined in these mutants (data not shown). Moreover, *prg-1* mutants exhibited a wild-type RNAi response to foreign dsRNA (data not shown). These findings suggested that *prg-1* was defective specifically in the 21URNA pathway.

Consistent with the genetic requirement of *prg-1* for 21U-RNA accumulation, the stage-specific expression of PRG-1 protein was coincident with that of 21U-RNA-1 and 21U-RNA-3442. PRG-1 levels were reduced in L1/L2 and L2/L3 worms when compared with L4 worms, as well as young and gravid adults (Figure 3B). As observed for 21URNAs, we could also detect the PRG-1 protein in embryo extracts, and we were unable to detect PRG-1 in the *glp-4(bn2)* mutant strain, suggesting that this protein is expressed in the germline. PRG-1 was also present in protein extracts from both female- and male- enriched populations. Curiously, the expression of *prg-1* was reduced in wild type worms cultured at 25°C (Figure 3B). Analysis of the expression of the *prg-1/prg-2* mRNA by real-time PCR revealed an expression pattern similar to that observed for the PRG-1 protein. The only exception observed was in the embryonic stage (Figure 3B). Although we could detect a high level of the PRG-1 protein in embryos, the mRNA was almost undetectable, supporting the idea that PRG-1 complexes in embryos are parentally derived.

In wild type worms, we observed a striking localization of PRG-1 in the cytoplasm and in prominent cytoplasmic structures in germ cells at nearly all stages of germ-line development. In both hermaphrodites and males, PRG-1 formed perinuclear foci in both the mitotic and meiotic zones of the germ line (Figure 3C-D). In mature oocytes the staining persisted but PRG-1 foci lost their perinuclear association and became dispersed in the cytoplasm (Figure 3C and data not shown). In males all PRG-1 staining disappeared abruptly as spermatids matured (Figure 3D). The pattern of PRG-1 localization, including its localization during embryogenesis (Figure 3E-F), resembled that of P granules, which are components of the *C. elegans* germ-line cytoplasm, or nuage (Strome and Wood, 1982) (Strome, 2005). Indeed, the localization of PRG-1 perfectly overlapped throughout development the localization of the previously described P-granule component, PGL-1 (Kawasaki et al., 1998) (Figure 3G).

21U-RNAs depend on and interact physically with PRG-1

To determine whether PRG-1 is required more broadly for 21U-RNA accumulation, we performed high-throughput sequencing analysis on small-RNA populations prepared from *prg-1* mutant animals and from wild-type animals reared at 20°C. For wild-type animals, approximately 11% of the 1,789,450 genome-matching reads corresponded to the 21U-RNAs, whereas for *prg-1* mutant animals less than 0.05% of the 1,774,442 genome-matching reads corresponded to 21U-RNAs (Figure 4A). This dramatic reduction in 21U-RNAs resembled that observed in animals lacking a germ line altogether (Figure 4B). However, *prg-1* animals maintained at 20°C were fertile and exhibited nearly wild-type levels of another class of germ-line enriched small RNAs, the endogenous siRNAs (Figure 4C). These findings indicate that *prg-1* is required for the accumulation of the entire 21U-RNA class of small RNAs.

To examine whether the 21U-RNAs physically interact with PRG-1, we immunoprecipitated the PRG-1 protein complex along with associated RNA. Both 21URNA-1 and 21U-RNA-3442 co-precipitated with the PRG-1 immune complex but not with precipitates recovered using pre-immune serum (Figure 4D). Small RNA species that did not require PRG-1 activity for accumulation, such as miR-66, were not detected in PRG-1

immunoprecipitates (Figure 4D). In contrast, we found that ALG-1/ALG-2 AGO-associated immune complex contained miR-66 but not 21U-RNA-1 or 21U-RNA-3442 (Figure 4D).

Biochemical analysis of small RNAs recovered in the PRG-1 IP complex demonstrated a strong bias for small RNAs with 5' U (>91%) compared to the total input population, which was enriched for 5' G (>70%; Figure 4E). Similarly, deep sequencing of small RNA libraries prepared from the IP sample demonstrated a dramatic enrichment for 21nt RNAs with 5' U in the PRG-1 complex (Figure 4F). In addition, 21mers with high-scoring motif matches were dramatically enriched in the IP sample (Figure 4G), and mapped comprehensively across the previously described 21U-RNA clusters on chromosome IV (Figure 4H). No other RNA species was significantly enriched in the PRG-1 IP. The above observations suggest that PRG-1 specifically binds 21U-RNAs to form a complex important for germ-line function and fertility.

***prg-1* mutants exhibit a broad spectrum of germ-line defects**

A previous study demonstrated that RNAi targeting *prg-1* and *prg-2* leads to reduced fertility (Cox et al., 1998). In this previous study it was not possible to target these two genes separately because of their high level of sequence identity. However, our examination of the phenotypic contributions of recently identified probable null alleles, revealed that most, if not all, of the germ-line defects result from the absence of *prg-1*. For example, *prg-2* mutants exhibited wild-type brood sizes at both 20°C and 25°C (Figure 5A) as well as normal numbers of morphologically wild-type germ cells (compare Figure 5B-C). In contrast, *prg-1* mutants exhibited dramatically reduced fertility at both temperatures (Figure 5A). Consistent with this phenotype, two different *prg-1* mutant strains and a *prg-1 prg-2* double-mutant strain all exhibited a significant reduction in the total number of germ nuclei populating the adult gonad (Figure 5D-F). The numbers of germ nuclei were reduced in each zone, but were most dramatically reduced in the mitotic zone in these mutants. The reduction in germ cell numbers was observed at all temperatures, and thus does not by itself explain the sterility of *prg-1* mutants at 25°C.

Although *prg-1* mutants exhibit temperature-dependent sterility, they do not appear to encode thermo-labile products. Rather, both alleles examined in this study are likely to represent null mutations (Cuppen et al., 2007) (Supplemental Figure 3A). As expected for null-mutants, the PRG-1 protein was either absent or truncated in these mutant strains at all temperatures (Supplemental Figure 3B). Furthermore, the 21U-RNA depletion associated with *prg-1* mutants was observed at all temperatures examined, including the semi-permissive temperatures of 15°C and 20°C. These findings suggest that in addition to their role in maintaining proper germ-cell numbers at all temperatures, PRG-1/21U-RNA complexes may function at higher temperatures to facilitate an otherwise temperature-dependent germ-line process required for normal fertility.

Temperature-shift experiments demonstrated that the temperature-sensitive period of *prg-1* mutants occurs during the adult stage. The fertility of animals shifted down from 25°C as young adults was substantially rescued, to an average of 40 progeny (n=10). Conversely, maintaining animals at 15°C during the L1 to adult stage, when the germ line is proliferating most rapidly, did not significantly rescue the fertility defect. These results suggest that, the germ cells produced in *prg-1* null mutant animals (that entirely lack PRG-1 protein expression), are deficient in a process important for their functionality (ovulation or fertilization) at elevated temperature.

To examine the relative contribution of defects in sperm vs oocytes to the reduced fertility of *prg-1* mutants, mutant hermaphrodites raised at 25°C were mated to wild-type males. The temperature-dependent sterility of *prg-1* was partially rescued, as the average number of

prg-1 progeny produced by animals reared at 25°C was 3 (n=10), but this number increased to 19 (n=10) when *prg-1* mutants were mated with wild type males. These findings suggested that the fertility defects of *prg-1* hermaphrodites stem, in part, from defects in the production and/or functionality of both the male and female gametes.

In summary, *prg-1* mutants exhibit dramatically reduced germ-cell numbers at all temperatures, and the gametes produced are markedly more sensitive to temperature than are those of wild-type animals. For example, at 25°C, wild-type animals produce ~200 progeny, about two thirds of the brood size observed at 20°C, while *prg-1* mutants produce an average brood size of only 3 progeny at 25°C, less than one tenth the brood size of 40 observed at 20°C. This reduction in brood size at higher temperature is approximately correlated with the reduction in the number of embryos observed, consistent with the idea that ovulation or fertilization are impaired at higher temperature.

***prg-1* mutants exhibit surprisingly subtle changes in gene expression**

On Chromosome IV hundreds of protein-encoding genes are interspersed with intergenic and intronic 21U-RNA loci over genomic regions that are millions of base pairs in length. Therefore, tiling arrays were used to profile changes in gene expression to determine whether the absence of 21U-RNAs in *prg-1* mutants might cause significant perturbations of gene expression either on this autosome or elsewhere. We found that *prg-1* and wild-type animals have broadly similar patterns of gene expression. Notably, genes located near 21U-RNA loci, including genes located within and around the major clusters of 21U-RNA loci on Chromosome IV, were not significantly altered in their expression (Figure 6A). Among 88 groups of developmentally co-regulated genes, also referred to as gene 'mountains' (Kim et al., 2001), 66 were essentially unchanged between the wild-type and *prg-1* strains (Figure 6B). Among the 16 mountains with decreased expression in *prg-1* mutants, were several mountains with germ-line functions such as cell division and oogenesis. Among the 6 mountains with increased expression was one containing spermatogenesis-related genes.

In *C. elegans*, a large class of RdRP-derived endogenous siRNAs (endo-siRNAs) target transposons and repetitive sequences as well as numerous protein-encoding genes (Ambros et al., 2003; Ruby et al., 2006; Gu and Conte, in preparation). Although PRG-1 does not appear to interact directly with small RNAs of this type (Figure 6C and Supplemental Table 2-3), we wondered whether 21U-RNAs might be linked, perhaps indirectly, to changes in the patterns of endo-siRNA expression. In many instances, changes in endo-siRNA levels correlated inversely with changes in gene expression from the corresponding interval (Figure 6D and Supplemental Table 4). However, the regions with significant changes in endo-siRNA levels were not correlated with regions containing 21U-RNAs or sequences with extended sequence similarity to 21U-RNAs.

One curious exception to this finding was the transposon Tc3, within which resides a single 21U-RNA. Found in all 22 Tc3 genomic loci, 21U-RNA-15703 overlaps the 3' inverted repeat (IR) downstream of and in the same orientation as the transposase gene (Figure 6E). This sequence was identified three times among 2 million reads in our small-RNA library prepared from the PRG-1 immune complex, an apparent enrichment when compared to only 12 reads in over thirty million from the remaining non-IP-associated data set. Examination of the endo-siRNA profile across a representative Tc3 element revealed two types of endo-siRNA reads. The first were antisense to the transposase gene and were unaffected in *prg-1(tm872)* mutants (Figure 6F). The second were directed, with a marked strand asymmetry, toward the Tc3 IR regions and were severely depleted in *prg-1(tm872)* mutants (Figure 6F). Neither the IR-directed, nor the transposase-directed siRNAs exhibited co-immunoprecipitation with PRG-1 (Figure 6G). Although the numbers of endo-siRNAs targeting the transposase gene were not significantly reduced in *prg-1*, we nevertheless

observed a 3- to 4-fold up-regulation of the Tc3 transposase mRNA (Figure 6H). Up regulation of the transposon mRNA as well as a greater than 100-fold increase in Tc3 transposition frequency were also observed for two different *prg-1* mutant alleles in a parallel study (Das et al.,; See Discussion).

DISCUSSION

AGO-protein/small-RNA complexes mediate biological activities that, so far, fall into the two broad categories of genomic surveillance and gene regulation. Several studies suggest that a metazoan-specific branch of the AGO family, called the Piwi AGOs, have become specialized to provide surveillance functions required for germ-line maintenance in animals (reviewed in Aravin et al., 2007). *C. elegans* contains one of the largest and best studied families of AGO proteins. Yet, beyond a general requirement for fertility (Yigit et al., 2006), the function of *C. elegans* Piwi-related AGOs and the nature of their small RNA co-factors had not been explored. We have shown that PRG-1, a Piwi sub-family AGO, interacts with 21U-RNAs, which are encoded by over 15 thousand genomic loci broadly clustered in two regions of Chromosome IV. These findings link this unusual class of small-RNAs to an RNAi-related pathway and suggest that PRG-1 and 21URNAs form an RNP complex required for proper germ-line development. The sequence repertoire of 21U-RNAs appears to be more diverse than expected by chance, and with one possible exception discussed below, likely sequence-specific targets for 21U-RNAs are not found in the *C. elegans* genome.

piRNAs in worms, flies and mammals

Piwi AGOs bind small RNAs (piRNAs) with the following characteristics: a Dicer-independent biogenesis, a 5' end with a monophosphate and a strong bias for Uracil, and a 3' end that is modified and resistant to periodate degradation (Klattenhoff and Theurkauf, 2008). The *C. elegans* 21U-RNAs share these characteristics but also exhibit several other unique properties (Ruby et al., 2006). Perhaps the most remarkable distinction is that 21U-RNAs originate from thousands of loci that frequently share a common upstream motif and are clustered in two large regions of one autosome. Within these two large regions of 2-million and 4-million base pairs respectively, the 21U-RNA loci are interspersed on both strands and rarely overlap with each other, repeat elements, or coding regions. Instead they localize to introns and intergenic regions within these chromosomal regions at an average density of one 21U-RNA locus every 200–300 bp.

In other organisms, piRNAs lack discernable upstream motifs and are often found in much smaller clusters dispersed on all chromosomes. In flies, a sub-group of piRNAs, originally termed repeat-associated siRNAs (rasiRNAs), are derived primarily from within repeats and transposons and appear to target transposons for silencing (Brennecke et al., 2007; Gunawardane et al., 2007; Saito et al., 2006). Furthermore, unlike 21URNAs, repeat-associated piRNAs derived from opposite strands frequently overlap.

In mammals, two types of piRNA clusters have been identified based on their temporal expression during spermatogenesis. Similar to *Drosophila* rasiRNAs, piRNAs expressed prior to meiotic pachytene in mice are derived from repeat and transposon rich clusters. These rasi-like piRNAs interact with the MILI AGO, which is expressed in the same developmental stages (Aravin et al., 2007). During pachytene, a second type of piRNA becomes abundant, which is derived from clusters that differ from both 21URNA clusters and rasiRNA clusters. These pachytene piRNA clusters span tens of thousands of bases—the length of a typical pre-mRNA transcript. Within these clusters the piRNAs exhibit remarkable strand bias, as though all the piRNAs within a region are processed from a single RNA-Polymerase II transcript or from two divergent transcripts (Aravin et al., 2006; Girard

et al., 2006; Grivna et al., 2006; Lau et al., 2006). In contrast, neighboring 21U-RNA loci, even those within the same intron of an annotated gene, appear to have autonomous biogenesis, each with their own 5' motif and deriving from the opposite strand about as often as from the same strand.

Despite these striking differences, mammalian pachytene piRNAs are similar to 21U-RNAs in one very intriguing way. Both types of small RNA encode tremendous sequence diversity and yet seem to lack obvious targets. In general, 21U-RNAs do not match repeat sequences or protein coding genes with a frequency any higher than that expected by chance.

PRG-1 complexes exhibit a conserved localization in germ-line nuage

We have shown that the PRG-1 protein localizes to the germ-line nuage, called P granules, in *C. elegans*. In other animals, Piwi AGOs show similar localization. In both *Drosophila* (AGO3 and Aubergine), and zebrafish (Ziwi), Piwi proteins localize to peri-nuclear nuage structures (Brennecke et al., 2007; Houwing et al., 2007). A third Piwi protein from *Drosophila*, Piwi itself, exhibits a more complex distribution, localizing to the nuclei of both germ cells and somatic cells (Brennecke et al., 2007; Cox et al., 2000). In mice, the localization of Miwi and Mili, has been analyzed, and although their expression peaks at different times, both are cytoplasmic proteins present in developing spermatids but absent in mature sperm (Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004).

A striking feature of PRG-1 localization was its presence in P granules throughout development. In germ-line stem cells and developing gametes of *C. elegans*, P granules are localized in a perinuclear pattern and are often found in apposition to nuclear pores (Pitt et al., 2000). They are thought to function in the sorting and storage of messages involved in (i) gametogenesis and (ii) subsequent parentally programmed zygotic development (Strome, 2005). In the fertilized egg and early embryo, the P granules dissociate from the nuclear periphery and are distributed in the cytoplasm. In the male germ line, P granules are present in dividing stem cells as well as meiotic spermatocytes but rapidly disappear as the spermatids mature. Finally, similar to other organisms, where piRNA expression correlates tightly with the expression of their Piwi-class AGO binding partners (Aravin et al., 2006; Girard et al., 2006; Houwing et al., 2007), the expression of 21U-RNAs closely correlated with the expression of PRG-1.

A potential role for 21U-RNAs in Tc3 silencing

In *C. elegans*, members of an expanded worm-specific AGO clade (the WAGOs) are required for the majority of transposon silencing, and appear to function with RdRP-derived siRNAs (Tijsterman et al., 2002). Surprisingly, the silencing of a single transposon family, Tc3, appears to depend on both WAGO family members (Vastenhouw et al., 2003) and on PRG-1 (Das et al., in press).

We found a single 21U-RNA, 21U-RNA-15703, that mapped to Tc3. This 21URNA appeared enriched among small RNAs recovered from the PRG-1 immune complex, but was located downstream of the transposase 3'UTR in the sense orientation and thus could not directly silence the transposase mRNA. Interestingly, 21U-RNA-15703 was located just upstream of a series of siRNAs associated with the Tc3 inverted repeats (IR). The production of IR-associated siRNAs depended on PRG-1, but also required the activities of two RdRPs and of an AGO in the WAGO clade (Data not shown).

The production of the PRG-1-dependent IR-associated siRNAs could be explained by a two-step model similar to one previously described for RDE-1-directed silencing in *C. elegans* (Yigit et al., 2006; Sijen et al., 2007; Pak and Fire, 2007). If a PRG-1 complex containing 21U-RNA-15703 were to cleave a target RNA that extended into Tc3 from the downstream

genomic region (Figure 7A), it could create a template for the RdRP-dependent synthesis of the secondary IR-associated siRNAs. How the loss of these IR-associated siRNAs might lead to activation of Tc3 in *prg-1* mutants remains unclear. Perhaps their loss leads to alterations in chromatin structure in the IRs or to changes in the expression of IR-associated regulatory transcripts. Such changes could explain the 3- to 4-fold increase in transposase mRNA levels observed by qRT-PCR, and might also render the IR genomic regions more accessible for transposase-directed endonucleolytic cleavage. The notion that PRG-1 may serve as an upstream AGO capable of triggering secondary siRNA production has implications for how other 21URNs may function, and could explain how loss of an exceptionally low-abundance 21U-RNA could cause the 100-fold increase in transposition of Tc3 (Das et al., in press).

A conserved function for piRNA complexes in maintaining pluripotency

Despite differences in their expression and the types of clusters from which they derive, our findings suggest that the overwhelming majority of 21U-RNAs and the abundant pachytene piRNAs of mammals share some intriguing similarities. Perhaps most notably, they share the confounding feature that, with few exceptions, they lack recognizable targets upon which they might specifically act. Although a number of genes, like Tc3, exhibit changes in expression in *prg-1* mutants, these changes could easily reflect alterations that arise indirectly. Overall, our analyses suggest that there is no correlation between genes whose expression is altered in *prg-1* mutants and the proximity of those genes to 21U-RNA loci.

One possible model to explain this paradox is to imagine that PRG-1/21U-RNA complexes may base-pair imperfectly with targets. A precedent for this already exists with animal miRNAs and most of their targets, for which pairing to miRNA seed nucleotides 2–8 is often sufficient for target recognition (Grimson et al., 2007). However, if similar partial matches were sufficient for piRNA-mediated regulation, then the entire transcriptome could potentially be placed under 21U-RNA-directed regulation. Perhaps 21U-RNAs act collectively, through partial sequence matches, to negatively regulate gene expression broadly. For example germ-line expressed mRNA recognized by 21URNA/PRG-1 complexes could be stored in the cytoplasm (perhaps within P-granules) until a secondary factor releases repression (Figure 7B). Such a mechanism would require the maintenance of sequence diversity within the 21U-RNA family, as a whole, rather than conservation of specific 21U-RNA sequences.

Out of more than 15,000 different 21U-RNAs encoded in *C. elegans*, only one transposon-directed 21U-RNA was identified, strongly suggesting that transposon silencing is not the only function mediated by this ancient metazoan-specific group of AGOs. It is interesting to note that many mammals, including humans, have, at great apparent cost to their fitness (Werdelin and Nilsson, 1999), derived morphological adaptations that place the male germ-line external to the body cavity. Perhaps this adaptation is necessary to facilitate the same temperature-sensitive process in gametogenesis that is also facilitated in part by PRG-1.

EXPERIMENTAL PROCEDURES

Worm Strains

The Bristol strain N2 was used as the standard wild-type strain. Alleles used in this study are listed below, grouped by chromosome: LGI: *glp-4(bn2)*, *prg-1(tm872)*, *prg-1(pk2298)*, *rde-3(ne3364)*, *ego-1(om71)*, *rff-1(ok589)*, *rff-2(pk2040)*; LGII: *rff-3(pk1426)*; LGIII: *dcr-1(ok247)*, *rde-4(ne299)*, *mut-7(ne311)*, *eft-3(q145)*, *qC1[nels(myo2::avr-15, rol-6, unc-22(RNAi))]*; LGIV: *fem-1(hc17)*, *prg-2(ok1328)*, *prg-2(tm1094)*; LGV: *fog-2(q71)*.

AGO deletions described in (Yigit et al., 2006) were also assayed for levels of 21U-RNA-1 and 21U-RNA-3442. *C. elegans* culture and genetics were as described in (Brenner, 1974).

Antibody Generation

Anaspec generated and purified the PRG-1 antibody in rabbits using the following peptides: RSGGSNNSGGKDQKYL and RQQGQSKTGSSGQPQKC.

Biochemistry and Molecular Biology

Protein and RNA purifications were performed as described in (Hutvagner et al., 2004) and (Duchaine et al., 2006), respectively. Antibodies used in this study are as follows: (1) monoclonal antibody anti-AFP 3E6 (Qbiogene), (2) an affinity-purified polyclonal anti-PRG-1 antibody (3), HRP-conjugated secondary antibody (Jackson ImmunoResearch), anti tubulin (Accurate Chemical). Northern blot analysis was performed as in (Duchaine et al., 2006). A more detailed description can be found in the Supplemental methods.

Quantitative Real-time PCR

Real-time PCR was performed using Superscript III Reverse Transcriptase (Invitrogen). and Applied Biosystems SYBR Green PCR Master mix according to the supplier's instructions. Primer sequences are available upon request.

Immunostaining and Microscopy

Gonads were prepared for indirect immunofluorescence as in (Pasierbek et al., 2001) and incubated with primary antibody (K76 (Wood et al., 1984) and the anti-PRG-1 antibodies described above) overnight at 4°C. Cy-3 anti-mouse IgM, and FITC or TRITC anti-rabbit secondary antibodies (Jackson ImmunoResearch) were used to detect K76 anti-PGL-1 and anti-PRG-1, respectively. Slides were mounted in Vectashield with DAPI (Vector Labs). All images were collected using a Hamamatsu Orca-ER digital camera mounted on a Zeiss Axioplan 2 microscope and with Openlab software.

Small RNA Cloning

Small endogenous *C. elegans* RNAs from embryos, L1, L2, L3, L4, dauer, mixed-stage, *glp-4* young adults, *prg-1(tm872)*, *fog-2(q71)*, and wild-type control worms were cloned using a protocol derived from (Lau et al., 2001). Endo-siRNA libraries generated from wild type and *prg-1(tm872)* were created as described in (Gu and Conte, in preparation). To generate small RNA libraries from PRG-1 immuno-complexes, PRG-1 IPs were performed on 70mg of total wild type protein as described in (Duchaine et al., 2006). Total RNA was extracted from a fraction of worms equivalent to the one used for the PRG-1 IPs. These small RNA libraries were constructed as described in (Ambros, et al, 2003). PCR products generated for all the samples described above were sequenced on a Solexa sequencing platform (Illumina, Inc.) (Seo et al., 2004). Detailed description of the cloning protocols, as well as data analysis can be found in the supplemental methods.

Thin Layer Chromatography

Small RNAs in the 18nt to 26nt range, obtained from total RNA and the RNA fraction that co-immunoprecipitates with PRG-1, were gel purified, treated with Calf Intestinal Alkaline Phosphatase (NEB) in the presence of 1U of Super RNase Inhibitor (Ambion) and labeled at the 5' end with T4 Polinucleotide Kinase in the presence of γ ATP. The 5' end labeled RNAs were Gel purified and incubated with nuclease P1 (USBiological). Samples were loaded on a TLC plate and run on a 0.5M Lithium Chloride solution.

Tiling Microarray Procedures

Total RNA was extracted as described above and prepared using the RiboPure total RNA isolation kit (Ambion). Labeling reactions were performed following the manufacturer's protocols with the GeneChip WT Double-Stranded cDNA Synthesis Kit (Affymetrix), GeneChip Sample Cleanup Module (Affymetrix) and the GeneChip WT Double Stranded DNA Terminal Labeling Kit (Affymetrix). Array hybridization to GeneChip *C. elegans* Tiling 1.0R chips was done using standard Affymetrix protocols and reagents. Signal values for each array probe were calculated using Affymetrix Tiling Analysis Software 1.1.2 (bandwidth: 30; intensities: PM/MM) with three replicates of *prg-1(tm872)* experimental datasets and three control wild-type. Probe overlap with annotations was assessed using the Affymetrix-provided ce4 coordinate, which indicates the genomic position matching the center of the array probe.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank our lab mates for many helpful discussions and comments on the manuscript, Fan Zhang for her early efforts on this project, Eric Miska for sharing unpublished data, R. Ketting, the CGC and the *C. elegans* Gene Knockout Consortium for providing strains. P.J.B. is supported by a predoctoral fellowship from Fundação para Ciência e Tecnologia (SFRH/BD/11803/2003), Portugal. D.A.C. is supported by a predoctoral fellowship from Fundação para Ciência e Tecnologia (SFRH/BD/17629/2004/H6BM), JMC is an HHMI fellow of the LSRF. C.C.M. and D.P.B. are Howard Hughes Medical Institute Investigators. This work was funded in part by the National Institutes of Health (GM58800).

REFERENCES

- Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D. MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr Biol*. 2003; 13:807–818. [PubMed: 12747828]
- Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M, Tuschl T. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*. 2006; 442:203–207. [PubMed: 16751777]
- Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*. 2007; 316:744–747. [PubMed: 17446352]
- Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AF, Roskin KM, Baertsch R, Rosenbloom K, Clawson H, Green ED, Haussler D, Miller W. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res*. 2004; 14:708–715. [PubMed: 15060014]
- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*. 2007; 128:1089–1103. [PubMed: 17346786]
- Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics*. 1974; 77:71–94. [PubMed: 4366476]
- Cox DN, Chao A, Baker J, Chang L, Qiao D, Lin H. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev*. 1998; 12:3715–3727. [PubMed: 9851978]
- Cox DN, Chao A, Lin H. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development*. 2000; 127:503–514. [PubMed: 10631171]
- Cuppen E, Gort E, Hazendonk E, Mudde J, van de Belt J, Nijman IJ, Guryev V, Plasterk RH. Efficient target-selected mutagenesis in *Caenorhabditis elegans*: toward a knockout for every gene. *Genome Res*. 2007; 17:649–658. [PubMed: 17416746]
- Deng W, Lin H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev Cell*. 2002; 2:819–830. [PubMed: 12062093]

- Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte DJ, Pang K, Brownell DR, Harding S, Mitani S, Ruvkun G, Yates J. R. r. Mello CC. Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell*. 2006; 124:343–354. [PubMed: 16439208]
- Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*. 2006; 442:199–202. [PubMed: 16751776]
- Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell*. 2007; 27:91–105. [PubMed: 17612493]
- Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev*. 2006; 20:1709–1714. [PubMed: 16766680]
- Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science*. 2007; 315:1587–1590. [PubMed: 17322028]
- Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD. The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol*. 2007; 17:1265–1272. [PubMed: 17604629]
- Houwing S, Kamminga LM, Berezikov E, Cronenbold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RH, Hannon GJ, Draper BW, Ketting RF. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell*. 2007; 129:69–82. [PubMed: 17418787]
- Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol*. 2007
- Hutvagner G, Simard MJ, Mello CC, Zamore PD. Sequence-specific inhibition of small RNA function. *PLoS Biol*. 2004; 2:E98. [PubMed: 15024405]
- Kawasaki I, Shim YH, Kirchner J, Kaminker J, Wood WB, Strome S. PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans*. *Cell*. 1998; 94:635–645. [PubMed: 9741628]
- Kim SK, Lund J, Kiraly M, Duke K, Jiang M, Stuart JM, Eizinger A, Wylie BN, Davidson GS. A gene expression map for *Caenorhabditis elegans*. *Science*. 2001; 293:2087–2092. [PubMed: 11557892]
- Kirino Y, Mourelatos Z. The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. *RNA*. 2007; 13:1397–1401. [PubMed: 17652135]
- Klattenhoff C, Theurkauf W. Biogenesis and germline functions of piRNAs. *Development*. 2008; 135:3–9. [PubMed: 18032451]
- Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W, Lin H, Matsuda Y, Nakano T. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development*. 2004; 131:839–849. [PubMed: 14736746]
- Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*. 2001; 294:858–862. [PubMed: 11679671]
- Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE. Characterization of the piRNA complex from rat testes. *Science*. 2006; 313:363–367. [PubMed: 16778019]
- Lin H. piRNAs in the germ line. *Science*. 2007; 316:397. [PubMed: 17446387]
- Ohara T, Sakaguchi Y, Suzuki T, Ueda H, Miyauchi K, Suzuki T. The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. *Nat Struct Mol Biol*. 2007; 14:349–350. [PubMed: 17384646]
- Pasierbek P, Jantsch M, Melcher M, Schleiffer A, Schweizer D, Loidl J. A *Caenorhabditis elegans* cohesion protein with functions in meiotic chromosome pairing and disjunction. *Genes Dev*. 2001; 15:1349–1360. [PubMed: 11390355]
- Pitt JN, Schisa JA, Priess JR. P granules in the germ cells of *Caenorhabditis elegans* adults are associated with clusters of nuclear pores and contain RNA. *Dev Biol*. 2000; 219:315–333. [PubMed: 10694425]

- Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell*. 2006; 127:1193–1207. [PubMed: 17174894]
- Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev*. 2006; 20:2214–2222. [PubMed: 16882972]
- Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H, Siomi MC. Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes Dev*. 2007; 21:1603–1608. [PubMed: 17606638]
- Seo TS, Bai X, Ruparel H, Li Z, Turro NJ, Ju J. Photocleavable fluorescent nucleotides for DNA sequencing on a chip constructed by site-specific coupling chemistry. *Proc Natl Acad Sci U S A*. 2004; 101:5488–5493. [PubMed: 15064393]
- Strome S. Specification of the germ line. *WormBook*. 2005:1–10. [PubMed: 18050414]
- Strome S, Wood WB. Immunofluorescence visualization of germ-line-specific cytoplasmic granules in embryos, larvae, and adults of *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 1982; 79:1558–1562. [PubMed: 7041123]
- Tijsterman M, Okihara KL, Thijssen K, Plasterk RH. PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of *C. elegans*. *Curr Biol*. 2002; 12:1535–1540. [PubMed: 12225671]
- Vastenhouw NL, Fischer SE, Robert VJ, Thijssen KL, Fraser AG, Kamath RS, Ahringer J, Plasterk RH. A genome-wide screen identifies 27 genes involved in transposon silencing in *C. elegans*. *Curr Biol*. 2003; 13:1311–1316. [PubMed: 12906791]
- Werdelin L, Nilsson A. The evolution of the scrotum and testicular descent in mammals: a phylogenetic view. *J Theor Biol*. 1999; 196:61–72. [PubMed: 9892556]
- Wood WB, Schierenberg E, Strome S. Localization and determination in early embryos of *Caenorhabditis elegans*. *Molecular Biology of Development*. 1984:37–49.
- Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell*. 2006; 127:747–757. [PubMed: 17110334]

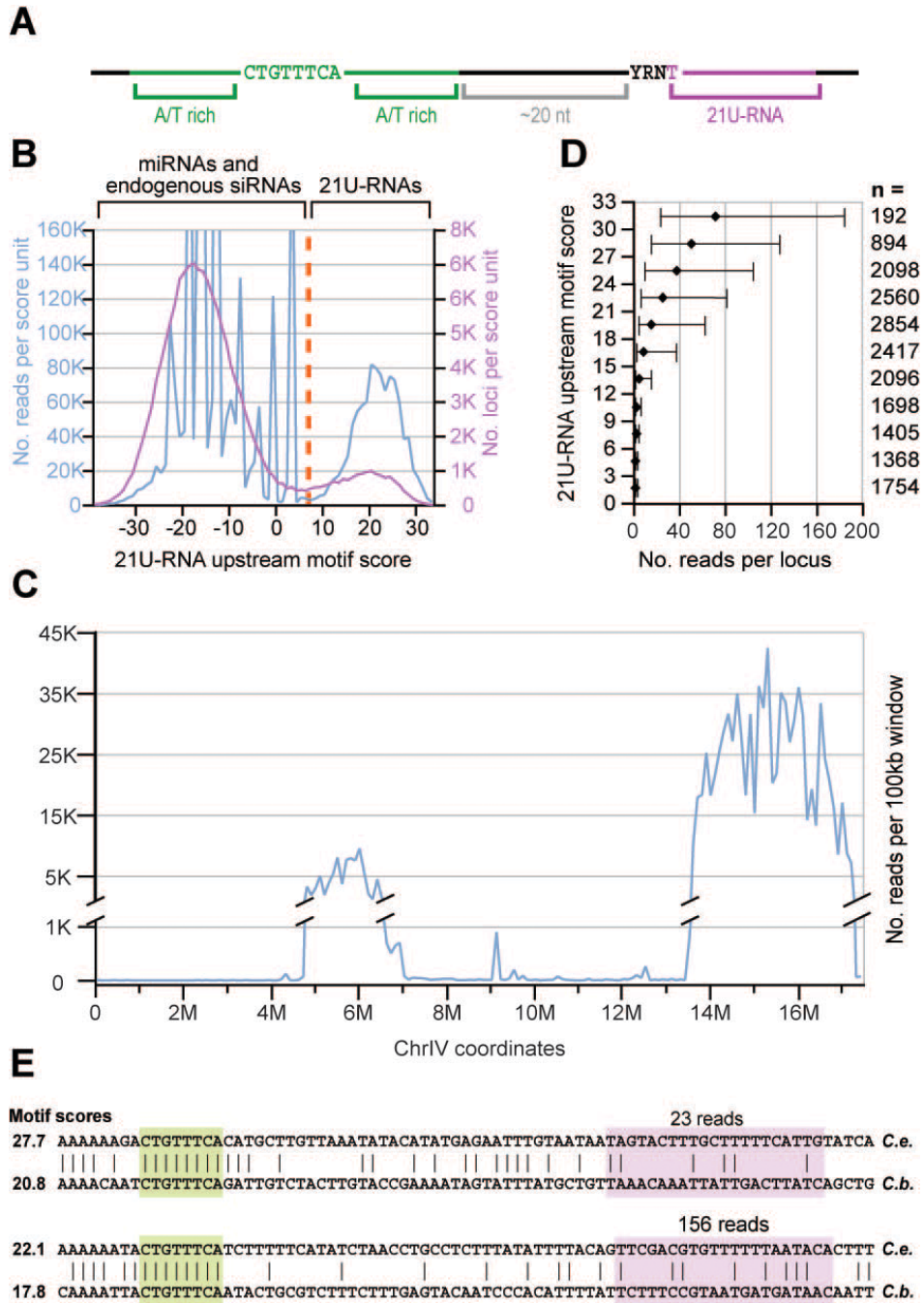


Figure 1. 21U-RNAs can be distinguished from other RNA species by their lengths and upstream motif matches

(A) A schematic representation of the 21U-RNA upstream motif as described previously (Ambros et al., 2003; Ruby et al., 2006). (B) The frequency of 21nt RNA reads (blue) or unique sequences (pink) versus upstream motif score. A cut-off score of 7 (orange) was used to define the 21U-RNA population. (C) The distribution of 21U-RNA reads across chromosome IV. Reads were classified as 21U-RNAs by their motif scores and normalized read counts were summed for each non-overlapping 100kb bin (blue). (D) The upstream motif score predicts the magnitude of 21U-RNA expression. For each three-bit bin of motif scores, the number of reads was determined for every observed 21U-RNA locus whose

motif score fell in that range. For each bin, the median read number was determined; error bars indicate the 25th and 75th percentiles. The number of loci in each bin is indicated by n. (E) Two 21U-RNA loci whose core upstream motifs are aligned (Blanchette et al., 2004). The core motif (green) and 21U-RNA (pink) are highlighted. The *C. briggsae* 21U-RNA was annotated based on the highest-scoring 5' end given the conserved core motif position. The number of reads for the 21U-RNA species from *C. elegans* is given above, and the motif score for each 21U-RNA ortholog is provided.

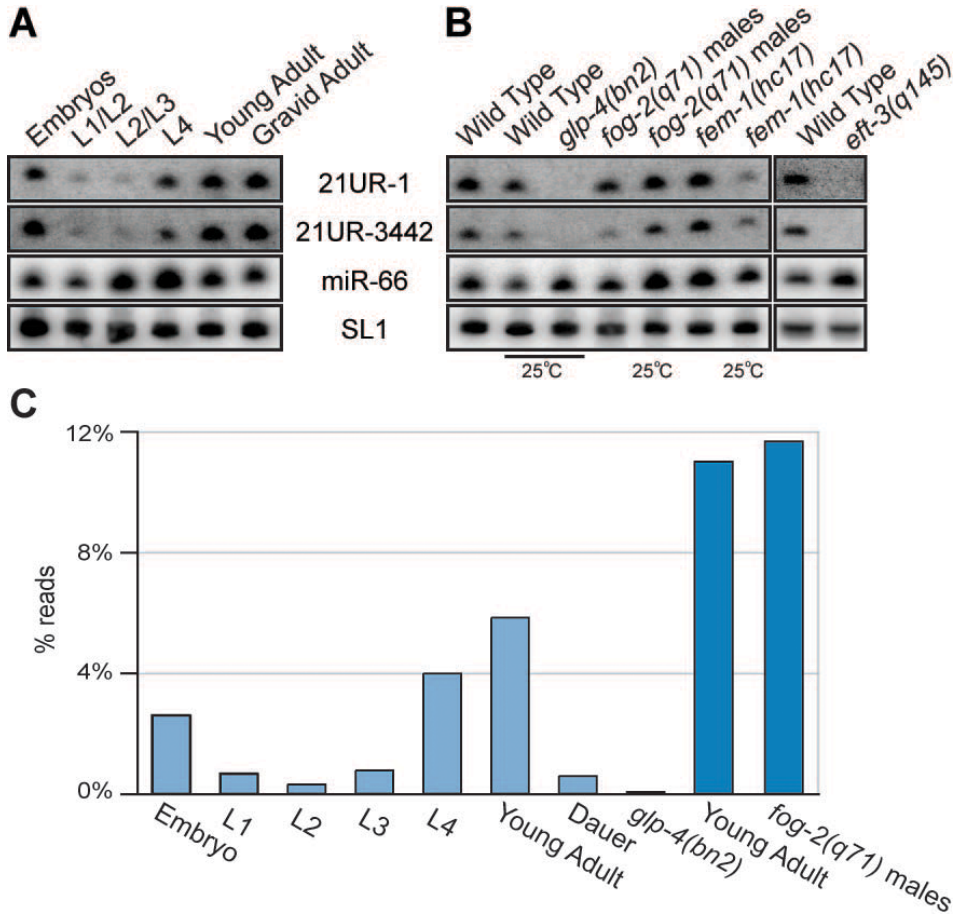


Figure 2. 21U-RNAs are expressed in the *C. elegans* germ line
 (A and B) Northern blot analysis of 21U-RNA-1 and 21U-RNA-3442 small RNAs. The SL1 precursor is shown as a loading control. In (A), RNA was generated from synchronized wild-type populations at distinct developmental stages. In (B) RNA generated from wild-type worms was compared to RNA obtained from strains with no germline: *glp-4(bn2)* and *eft-3(q145)*; a male-only population: *fog-2(q71)*; and a population of worms with no sperm: *fem-1(hc17)*. (C) The expression profile for the bulk population of 21U-RNAs as determined by large-scale sequencing. The percent of reads from each library with 21U-RNA identity are shown. Light blue libraries were prepared for sequencing with Rnl2(1–249) ligase, dark blue with T4 RNA ligase 1 (see methods).

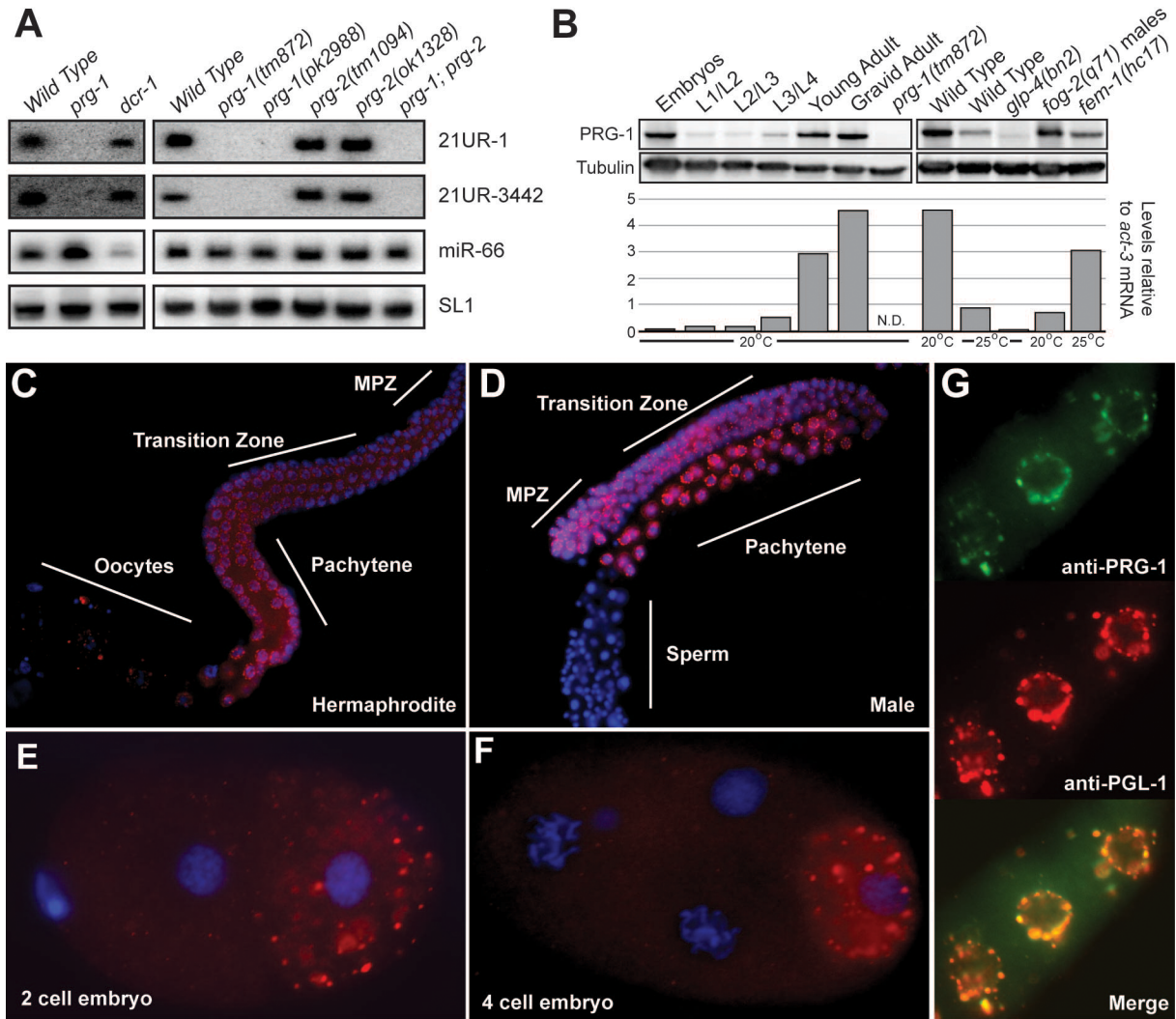


Figure 3. PRG-1 protein is expressed in the germ line and required for 21U-RNA accumulation (A) Northern blot analysis of 21U-RNA-1, 21U-RNA-3442 and miR-66 expression in wild-type and mutant strains. In the left panel RNA generated from *prg-1* and *dcr-1* homozygous populations was analyzed. In right panel RNA samples obtained from the *tm872* and *pk2298* alleles of *prg-1* and *tm1094* and *ok1328* alleles of *prg-2* as well as from the double mutant *prg-1(tm872) prg-2(tm1094)* were analyzed. The SL1 precursor serves as a loading control. (B) Western blot analysis of the PRG-1 developmental expression profile. A tubulin specific antibody was used as a loading control. On the left panel, protein lysate generated from wild-type populations at distinct developmental stages was analyzed. On the right panel, protein lysates from wild type worms were compared to protein lysates from strains with no germline: *glp-4(bn2)*; a male-only population; *fog-2(q71)* mutant worms; and a population of worms with no sperm: *fem-1(hc17)*. Quantitative Real-Time PCR analysis of the expression of the *prg-1/prg-2* mRNA is shown in the bottom panel. The PCR primers used recognize both *prg-1* and *prg-2*. Expression level is indicated relative to the levels obtained for actin (*act-3*) mRNA in each sample. (C-F) PRG-1 immunofluorescence (red) and DNA DAPI staining (blue) in dissected gonad arms from an adult hermaphrodite (C) and male (D), a two-cell embryo (E), and a 4-cell embryo (F). In (C and D) The mitotic (MPZ) and meiotic zones (Transition Zone plus Pachytene) are indicated, as are the proximal zones containing oocytes and sperm (respectively). (G) Dual immunofluorescence analysis of 3

oocytes in the proximal arm of a wild-type hermaphrodite gonad stained for PRG-1 and PGL-1 as indicated. Yellow represents overlap in the merged image (bottom panel).

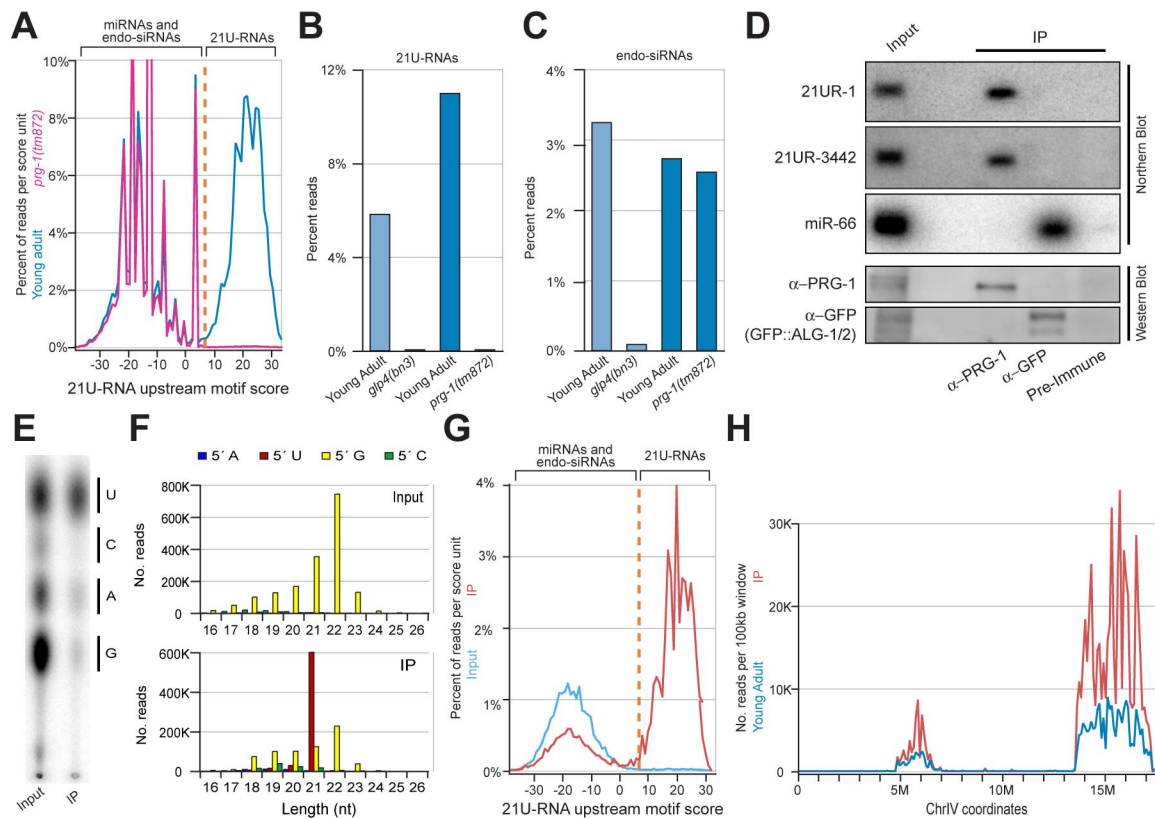


Figure 4. PRG-1 interacts with and is required for the accumulation of all 21URNAs

(A) The frequency of 21nt RNA reads from wild-type young adults (blue) and *prg-1(tm872)* young adult (pink) versus upstream motif score. A cut-off score of 7 (orange) was used to define the 21U-RNA population. (B) 21U-RNAs are absent in *glp-4(bn2)* and *prg-1(tm872)* mutant worms. Percent reads from each library with 21U-RNA identity are shown. Histogram bars colored as in Figure 2C. (C) Endogenous siRNAs are absent in *glp-4(bn2)* but not *prg-1(tm872)* mutant worms. Percent reads with 5' G nucleotides and complete antisense overlap with coding exons (Ambros et al., 2003) are shown. Histogram bars colored as in Figure 2C. (D) IP/Northern blot analysis of small RNAs in PRG-1 and GFP::ALG1/2 complexes. Immunoprecipitations were performed on lysates prepared from an otherwise wild-type transgenic strain carrying GFP-tagged ALG-1 and ALG-2. The top panels are Northern blots probed for associated small RNAs. The lower panels are Western blots probed as indicated. (E) Thin Layer Chromatography analysis of the first nucleotide of the small RNA population that co-immunoprecipitates with the PRG-1 protein. Bars show where the single nucleotides migrate. (F) The length and 5' nucleotide distribution of reads from the Input (top) and PRG-1 co-IP (bottom) libraries. (G) The frequency of 21nt RNA reads from the Input (blue) and PRG-1 co-IP (red) libraries versus upstream motif score. Plotted as in Figure 4B. (H) The distribution of 21U-RNA reads from the PRG-1 co-IP library (red) versus the young adult wild-type library prepared with T4 RNA ligase 1 (see methods; blue). Reads were classified as 21U-RNAs by their motif scores and normalized read counts were summed for each non-overlapping 100kb bin.

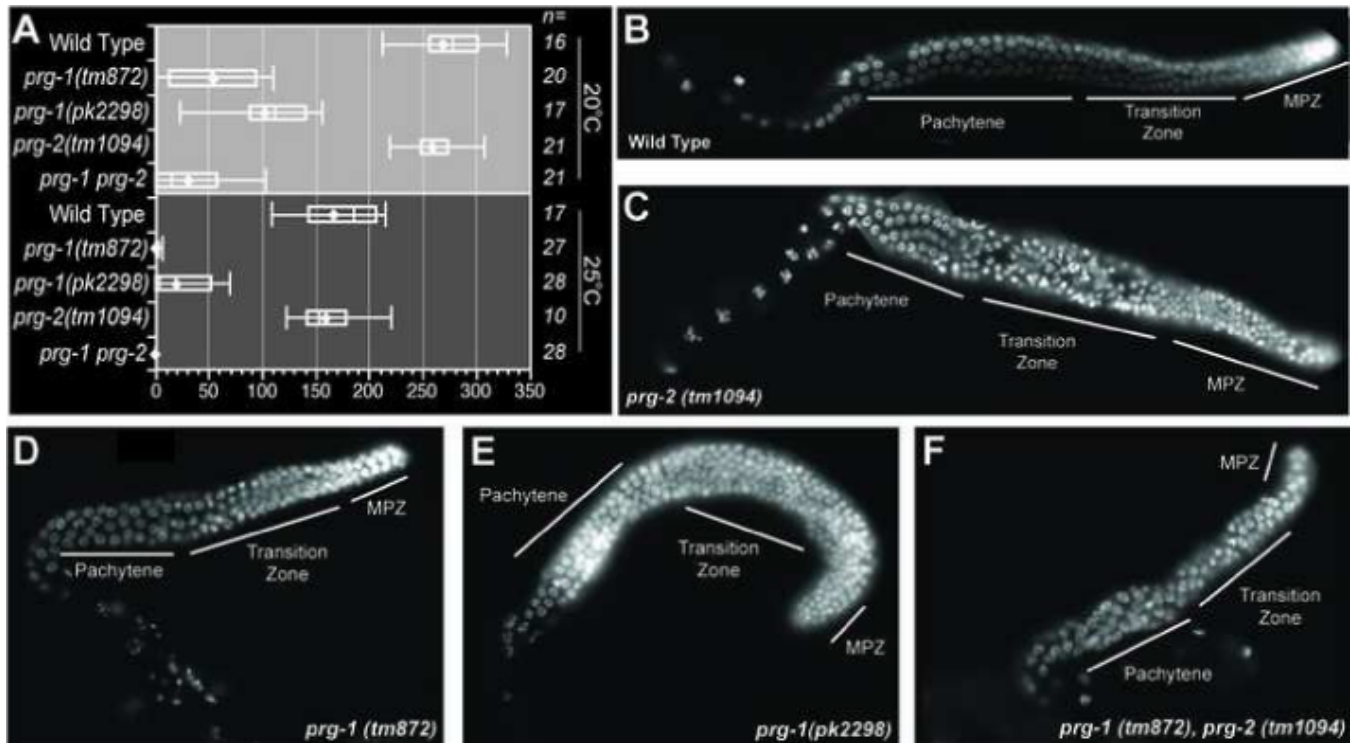


Figure 5. PRG-1 exhibits a broad spectrum of germ line defects

(A) Brood size analysis of *prg-1* and *prg-2* mutant strains. The brood size of 'n' individual animals for each strain was determined at 20°C and 25°C. Left and right lines represent highest and lowest values respectively. Left and right ends of each box represent the 75th and 25th percentile respectively, the diamond represents the average brood size and the vertical line inside the box represents the median value. (B - F) DAPI staining of excised gonads from wild-type, *prg-1* and *prg-2* strains (as indicated). Gonadal zones are indicated as in Figure 3.

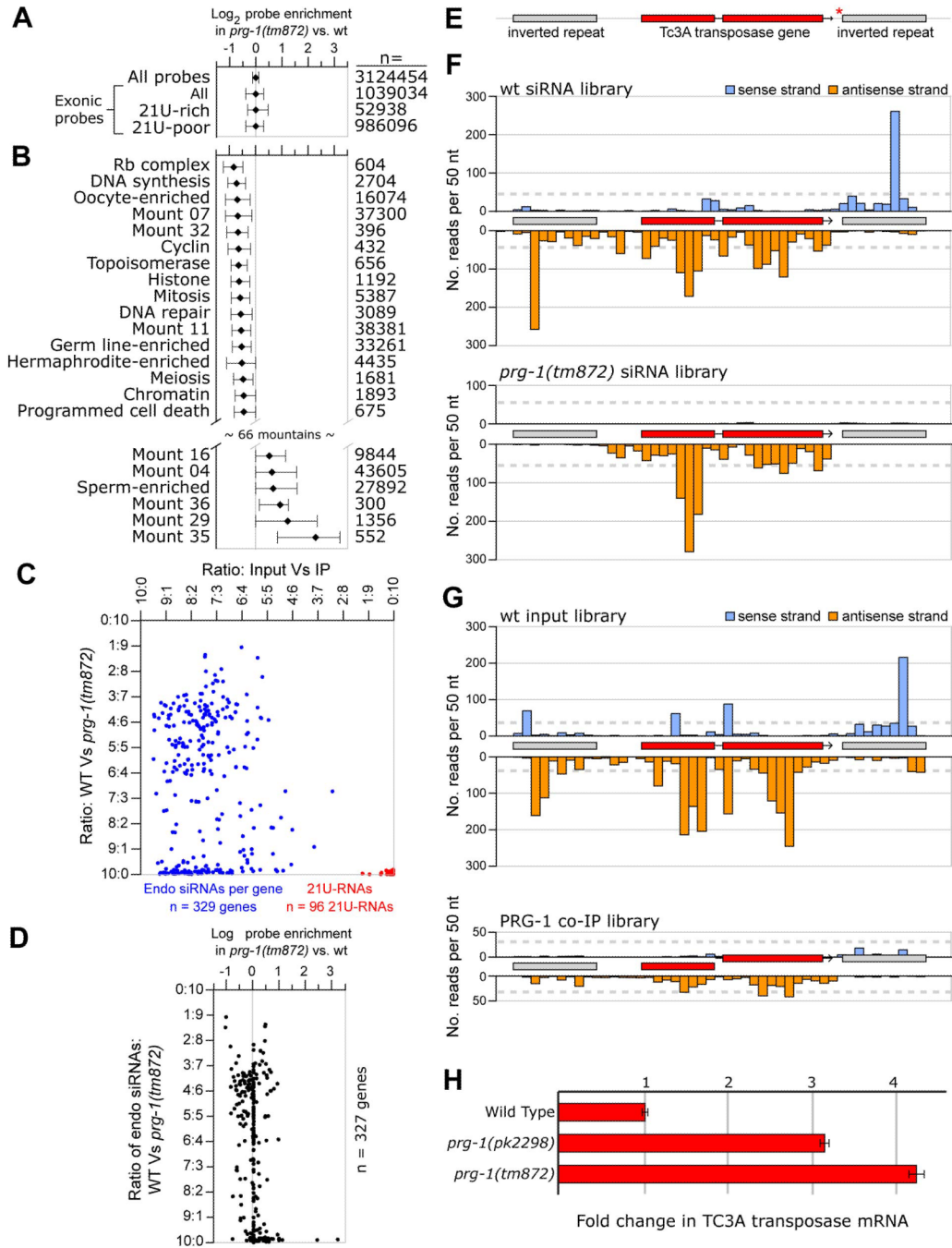


Figure 6. *prg-1* mutants exhibit surprisingly subtle changes in gene expression

(A) Gene expression was not preferentially affected in the 21U-rich portions of the *C. elegans* genome. Median values are shown for each of the indicated probe sets. Error bars indicate 25th and 75th percentile signal changes. The number of probes is indicated by 'n'. (B) The overall expression of some gene mountains was significantly altered in the *prg-1(tm872)* mutant. All probes overlapping the exons of all genes from each mountain were considered; the median signal changes from each probe set are shown (error bars indicate the 25th and 75th percentiles). All mountains from (Kim et al., 2001) are shown whose median expression was increased or decreased by 0.4 log₂ units or more. The number of probes is indicated by n. (C) Small RNA depletion in the *prg-1(tm872)* mutant versus

enrichment in the PRG-1 co-IP. The x axis indicates the ratio of read frequencies between the input versus PRG-1 co-IP libraries. The y axis indicates the ratio of antisense read frequencies between the wild-type and *prg-1(tm872)* mutant siRNA-enriched libraries. Each blue dot indicates the antisense read count for one gene whose wild type siRNA-enriched read count is 500. Each red dot indicates the read count for a 21U-RNA species with 200 reads from the young adult wild type library prepared with T4 RNA ligase 1 (see methods) and at least one read between the two libraries of each plot axis. (D) Changes to mRNA versus derivative siRNA levels per gene in the *prg-1(tm872)* mutants. Each point indicates a gene with 10 array probes and 500 antisense reads from the wt siRNA-enriched library overlapping annotated exons. The x axis is as in Figure 6A. The y axis is as in Figure 6C. (E) A schematic view of a full-length Tc3 transposon showing the inverted repeats (grey) and Tc3A transposase gene (red). The position of 21U-RNA-15703 is indicated by a red asterisk. (F) Density of reads mapping to the sense (blue) and antisense (orange) strands of the Tc3 element from Figure 6E. Reads per 50nt window are shown from the wild-type (top) and *prg-1(tm872)* mutant (bottom) siRNA-enriched libraries. Read counts are not normalized to the number of genomic matches. Dotted grey lines indicate 0.002% of each library. (G) Density of reads mapping to the sense (blue) and antisense (orange) strands of the Tc3 element from Figure 6E. Reads per 50nt window are shown from the input (top) and PRG-1 co-IP (bottom) libraries. Read counts are not normalized to the number of genomic matches. Dotted grey lines indicate 0.002% of each library. (H) qRT-PCR analysis of the expression of the *TC3A* mRNA. Primers recognizing *TC3A* mRNA were used in qRT-PCR on cDNA generated from wild-type, *prg-1(tm872)* and *prg-1(pk2298)* worms. Expression level is relative to actin (*act-3*) mRNA in each sample.

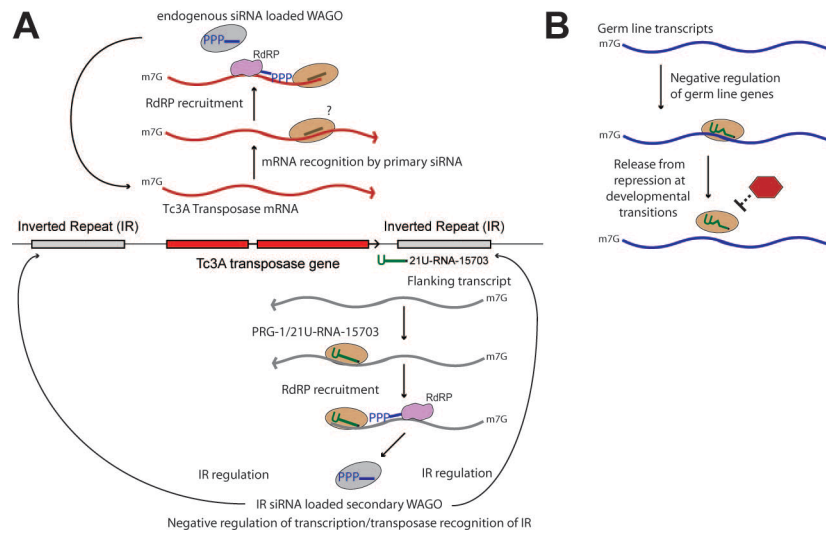


Figure 7. Models for 21U-RNA function
 (A) Regulation of TC3 inverted repeats by PRG-1/21U-RNA-15703 (B) Regulation of germ line transcripts by imperfect base pairing