

Nematode Small RNA Pathways in the Absence of piRNAs

Maxim Zagoskin^{1,2,3*}, Jianbin Wang^{1,2,3,4*,#}, Ashley T. Neff¹, Giovana M. B. Veronezi¹, and
Richard E. Davis^{1,2#}

¹Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora,
CO

²RNA Bioscience Initiative, University of Colorado School of Medicine, Aurora, CO

³Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN

⁴UT-ORNL Graduate School of Genome Science and Technology, University of Tennessee, Knoxville,
TN

*These authors contributed equally

#Corresponding Authors, richard.davis@cuanschutz.edu, jianbin.wang@utk.edu

Abstract

Small RNA pathways play diverse regulatory roles in the nematode *C. elegans*. However, our understanding of small RNA pathways, their conservation, and their roles in other nematodes is limited. Here, we analyzed small RNA pathways in the parasitic nematode *Ascaris*. *Ascaris* has ten Argonautes with five worm-specific Argonautes (WAGOs) that are associated with secondary 5'-triphosphate small RNAs (22-24G-RNAs). These *Ascaris* WAGOs and their small RNAs target repetitive sequences (WAGO-1, WAGO-2, WAGO-3, and NRDE-3) or mature mRNAs (CSR-1, NRDE-3, and WAGO-3) and are similar to the *C. elegans* mutator, nuclear, and CSR-1 small RNA pathways. *Ascaris* CSR-1 likely functions to “license” gene expression in the absence of an *Ascaris* piRNA pathway. *Ascaris* ALG-4 and its associated 26G-RNAs target and appear to repress specific mRNAs during meiosis in the testes. Notably, *Ascaris* WAGOs (WAGO-3 and NRDE-3) small RNAs change their targets between repetitive sequences and mRNAs during spermatogenesis or in early embryos illustrating target plasticity of these WAGOs. We provide a unique and comprehensive view of mRNA and small RNA expression throughout nematode spermatogenesis that illustrates the dynamics and flexibility of small RNA pathways. Overall, our study provides key insights into the conservation and divergence of nematode small RNA pathways.

Keywords: Nematode, small RNA pathways, Argonautes, WAGOs, CSR-1, NRDE-3, ALG-4, piRNAs, 22G-RNA, 26G-RNA, RdRP, spermatogenesis, germline, and *Ascaris*

Introduction

Small RNAs contribute to many regulatory processes. They are associated with diverse functions including repressing foreign invaders and mobile elements, transcriptional regulation, mRNA translation and degradation, DNA repair, chromatin regulation and epigenetic inheritance, and ciliate genome rearrangements [1-5]. One of the first discoveries of small RNAs and their role in gene regulation was in the free-living nematode *Caenorhabditis elegans* [6, 7]. Subsequent studies in *C. elegans* have revealed a diverse and complex set of small RNAs, pathways, and associated Argonautes [8-11].

C. elegans small RNAs include miRNAs, 21U-RNAs (piRNAs) and small-interfering RNAs (siRNAs). *C. elegans* miRNAs are involved in post-transcriptional gene regulation through regulation of mRNA translation and degradation [12], but have also been associated with transcriptional activation [13]. *C. elegans* can generate siRNAs against foreign elements, but also has a large repertoire of endogenous siRNAs [8-11]. These *C. elegans* small RNAs are named based on their length and predominant first nucleotide of the RNA, e.g., 21U-, 22G- and 26G-RNAs. 21U-RNAs have a 5'-monophosphate and 3' 2'-O-methylation [14-16]. They are primarily thought to identify foreign (non-self) RNA elements. The identification of these foreign elements through RNA base-pairing leads to the subsequent synthesis of secondary siRNAs that repress their targets. 22G-RNAs have a 5'-triphosphate [16-20]. These are secondary siRNAs as they are generated in response to other small RNAs (21U-RNA, 26G-RNA, or other siRNAs) [20, 21]. Thus, primary siRNAs base-pair with transcript targets, marking or identifying them for the synthesis of the more abundant antisense, secondary 22G-RNAs by RNA-dependent RNA polymerases (RdRPs) [18-23]. 22G-RNAs silence mobile elements, pseudogenes, non-annotated loci, and select endogenous, germline genes [18]. They also serve to “activate or license”, “tune”, or repress gene expression [24-27]. 26G-RNAs have a 5' monophosphate [16, 19, 28, 29]. There are two classes of 26G-RNAs in *C. elegans*, one is testis-specific and associated with the Argonautes ALG-3/4 [30-32], the other is expressed in early embryos and associated with the Argonaute ERGO-1 [28, 29, 32, 33]. Like 21U-RNAs, 26G-RNAs trigger or act to prime the synthesis of secondary siRNAs (22G-RNAs) through base-pairing with targets [8-11]. Overall, *C. elegans* secondary 22G-RNAs are amplified responses to targets identified by 21U- and 26G-RNAs and other primary siRNAs.

Small RNAs are bound by effector Argonaute proteins. *C. elegans* Argonautes have undergone significant expansion and diversification. Twenty-seven Argonaute genes were originally described [23]; 19 are expressed (Julie Claycomb, personal communication). These include members of the AGO, PIWI, and WAGO (Worm-specific Argonautes) Argonaute clades [34]. The *C. elegans* AGO clade Argonautes (5 AGOs) interact with miRNAs and 26G-RNAs, the PIWI clade (1 AGO) with 21U-RNAs (the worm piRNA ortholog), and the WAGO clade (13 AGOs) with 22G-RNAs. Several of the WAGO clade Argonautes function in the nucleus regulating transcription and heterochromatin formation [3, 11, 35, 36]. The largest

expansion of *C. elegans* Argonautes is in the WAGO clade [23]. Many of these WAGOs are thought to have redundant functions in *C. elegans*.

Nematodes are an extremely diverse and abundant phylum adapted to a wide variety of lifestyles [37-39]. While extensive analyses of *C. elegans* Argonautes, small RNAs, and pathways have been carried out, relatively little is known regarding the conservation, divergence, or function of these pathways in other nematodes [8, 40, 41]. Nematodes have been divided into three major classes and five clades with *C. elegans* and its relatives as members of Clade V [42-45]. The nematode *Ascaris* is a parasite of humans (and pigs) infecting upwards of 800,000 people [46-48]. *Ascaris* is a Clade III nematode estimated to have diverged from *C. elegans* approximately 365-400 million years ago [49, 50]. Previous studies in *Ascaris* indicated that piRNAs and PIWI Argonautes are absent in *Ascaris* [51]. PIWI Argonautes and piRNAs play a key role in repressing mobile elements in the germline. *C. elegans* piRNAs have been proposed to serve in identifying and defining foreign genetic elements acting upstream of secondary siRNA pathways [52-56]. Thus, they have been described as defining self vs. non-self [57]. The WAGO-associated 22G secondary RNAs function in silencing and can maintain silencing of the foreign elements over generations in *C. elegans* [8, 11]. To counteract silencing, it has been proposed that 22G-RNAs associated with the *C. elegans* CSR-1 function to “license”, identify self, or protect germline genes from repression and allow for their expression [24, 26, 58]. Given the key role of piRNAs in many organisms, and their role in *C. elegans*, the absence of piRNAs and PIWI in *Ascaris* raises the question of how *Ascaris* small RNA pathways have adapted to the absence of piRNAs and PIWI (e.g., self vs non-self). Does *Ascaris* still need small RNA pathways to “license” gene expression without the presence of piRNAs? How are self vs non-self elements determined?

Here, we generated antibodies to *Ascaris* Argonaute proteins and carried out Argonaute IP and small RNA sequencing to characterize small RNA pathways in several developmental stages. *Ascaris* has Argonautes that are highly specific for binding miRNAs (AsALG-1) and 26G-RNAs (AsALG-4). AsWAGO Argonautes bind 5'-triphosphate small RNAs (22-24G-RNAs). These small RNAs target repetitive sequences including mobile elements, but they also target mRNAs and likely “license”, “tune”, and/or repress gene expression. Two *Ascaris* WAGOs change their genomic targets (mRNAs vs repetitive sequences) in different developmental stages. We exploited the long ~1 meter *Ascaris* male germline to obtain discrete regions of the testis and analyzed *Ascaris* Argonautes and their small RNAs throughout spermatogenesis. These analyses provide a unique and comprehensive timeline for the expression of Argonautes, their bound small RNAs and targets, and changes in expression of their corresponding genomic or mRNA targets throughout nematode spermatogenesis. Our study suggests that in the absence of piRNAs and with extensive evolutionary divergence from *C. elegans*, several *Ascaris* small RNA pathways and Argonautes appear to bind similar small RNAs, have similar targets, and in many cases appear to serve similar roles in both

nematodes. Therefore, the potential “licensing” of gene expression in another nematode by *Ascaris* CSR-1 does not appear to be a consequence or counter response to the presence of piRNAs. Overall, our studies illustrate the conservation, loss of Argonautes, and divergence of small RNA pathways that illustrate the flexibility and adaptability of Argonautes and small RNA pathways in nematodes.

Results

Ascaris Argonautes

We previously identified 10 *Ascaris* Argonautes [51]. Alignment and phylogenetic analyses of these Argonautes indicates the presence of 5 AGO clade and 5 WAGO (Worm-specific Argonautes) Clade Argonautes, but the absence of PIWI Argonautes (Figure 1A). Of the 5 *Ascaris* AGOs, two (AsALG-1 and AsALG-6) cluster with *C. elegans* miRNA associated Argonautes and one (AsALG-4) with *C. elegans* ALG-3/4 Argonautes. The two additional Argonautes (AsALG-5 and AsALG-7) cluster with 26G-RNA-like Argonautes or RDE-1 (Figure 1). Four of these five *Ascaris* AGOs have conserved RNase H, catalytic tetrad residues that confer slicer activity except for AsALG-6.

The 5 *Ascaris* WAGOs were named AsCSR-1, AsWAGO-1, AsWAGO-2, AsWAGO-3, and AsNRDE-3 based on expression pattern, their sequence and phylogenetic similarity to *C. elegans*, and the small RNAs associated with these Argonautes (see below)(Figure 1). Two *C. elegans* WAGOs function in nuclear RNAi and are expressed in different stages, HRDE-1 (germline) and NRDE-3 (soma) [59, 60]. Our phylogenetic analysis suggests one *Ascaris* WAGO, named AsNRDE-3, is related to these *C. elegans* nuclear WAGOs. It is expressed in zygotes and early embryos and present both in the cytoplasm and nuclei of embryos. We note that with only one *Ascaris* nuclear WAGO, AsNRDE-3 may also serve functions similar to *C. elegans* HRDE-1. Only AsCSR-1 and AsWAGO-3 have the RNase H, catalytic tetrad residues that confer slicer activity.

Ascaris has three RdRPs (Figure 1B and 4B). One is clearly orthologous to *C. elegans* RRF-3 (RdRP-3) and is expressed in the germline, oocytes, zygotes, and during early development. In the male germline, RdRP-3 RNA expression is highest at M5 when the initiation of 26G-RNAs occurs and likely plays a major role in their biogenesis. The other two *Ascaris* RdRPs are similarly orthologous to either *C. elegans* EGO-1 or RRF-1 and are expressed in the germlines (Table S1). No ortholog of *C. elegans* RRF-2 appears present in *Ascaris*. Additional *Ascaris* orthologous proteins likely associated with small RNA pathways in *C. elegans* are provided in Table S1.

Ascaris Argonautes are dynamically expressed throughout the male and female germlines, early development, larvae, and adult tissues (Figure 1B; rpkM values less than 5 are considered low or background levels of expression). The most abundant Argonaute in all these stages is the *Ascaris* CSR-1 ortholog, AsCSR-1 (Figure 1B), which is expressed in the male and female germlines, zygotes prior to pronuclear fusion (where the maternal to zygotic transition occurs) [61], and through the 4-cell stage of early development suggesting a pivotal role of AsCSR-1 and its small RNAs in germline development, the maternal-zygotic transition and early embryogenesis. AsCSR-1 is not expressed in late embryos, larvae or somatic tissues and the gene becomes heavily marked with H3K9me3 in 32-64 cell embryos when its

expression ceases (Figure S1). Two isoforms of CSR-1 protein are present in *C. elegans* that vary in the first exon and thus amino-terminus [62-65]. The *C. elegans* CSR-1a isoform with the extended amino-terminus is expressed during spermatogenesis [64]. RNA-seq, ISO-seq, or PRO-seq analysis in *Ascaris* did not identify RNAs that differ in exon 1 and the AsCSR-1 has the extended amino-terminal end similar to *C. elegans* CSR-1a. Other *Ascaris* WAGOs are also primarily expressed in germline and embryos, with AsWAGO-1 expression at higher levels than the others. Interestingly, AsWAGO-1 and AsWAGO-3 are also expressed in somatic tissues such as intestine and carcass (cuticle, body wall muscle, hypodermis, and some nervous tissue). Their targets and functions in these tissues remain to be determined.

For AGO-clade *Ascaris* Argonautes, the miRNA Argonautes (AsALG-1 and AsALG-6) are expressed in both the male and female germline, early development, and larvae but are present at much lower levels in somatic tissues including the muscle, intestine, and carcass. AsALG-6 is eliminated during *Ascaris* programmed DNA elimination [66-68]. The 26G-RNA Argonaute, AsALG-4, is highly and predominantly expressed in the male germline and highest during meiosis. AsALG-5 and AsALG-7 expression is in general low in the male and female germlines and early embryos through the comma stage, except for AsALG-7 in early spermatogenesis. The phylogenetic analysis in Figure 1A suggests they may be potential 26G-like Argonautes (Figure 1A). In some phylogenetic analyses (data not shown), AsALG-5 and AsALG-7 cluster with *C. elegans* ERGO-1 or RDE-1. However, *C. elegans* ERGO-1 binds 26G-RNAs that are 2'-O-methylated by HENN-1 but *Ascaris* lacks HENN-1 [51].

***Ascaris* Argonaute antibodies**

To identify small RNAs associated with the *Ascaris* Argonautes, we generated polyclonal antibodies (see Material and Methods) to AsALG-1 (miRNA), AsALG-4 (26G-RNA), and all five WAGO Argonautes (AsCSR-1, AsNRDE-3, AsWAGO-1, AsWAGO-2, and AsWAGO-3). These Argonautes were chosen as they represent the diversity of small RNA pathways in *Ascaris*, are the most highly expressed, and appeared orthologous to major *C. elegans* Argonautes. Western blot analyses show these antibodies specifically recognize the corresponding Argonaute proteins (Figure S2A). In addition, RNA and Northern blot analyses of the AsALG-1 and AsWAGO-1 IP indicate significant enrichment for miRNAs and 22G-RNAs supporting the specificity of these antibodies (Figure S2B). Immunoprecipitation and proteomic analysis of several of the Argonautes further demonstrated the specificity of the antibodies for their respective Argonaute, including known associated proteins in *C. elegans* (data not shown). Additional IP and small RNA sequencing data (see below) support the specificity of the antibodies.

Localization of *Ascaris* Argonautes in early embryos and the germline

Early embryo immunohistochemistry using the antibodies to *Ascaris* Argonautes indicates that AsCSR-1, AsWAGO-1, AsWAGO-2, AsWAGO-3, AsNRDE-3, AsALG-4, and AsALG-1 are present in both the

cytoplasm and nucleus in early embryos (Figure 1C). AsWAGO-2 and AsWAGO-3 localize to condensed mitotic chromosomes during anaphase of a programmed DNA elimination mitosis and are described in a separate manuscript [69]. AsCSR-1 does not localize to condensed mitotic chromosomes [69]. We further examined the nuclear and cytoplasmic localization of several Argonautes in early embryos using Western Blots (Figure S2A). AsWAGO-3, AsCSR-1, and AsNRDE-3 are present in both the cytoplasm and nuclei, but significantly greater in the cytoplasm. Notably, AsALG-1 is also present in nuclei and is associated with miRNAs. However, AsWAGO-2 is equally present in the cytoplasm and nucleus whereas AsWAGO-1 is around two-fold higher in the cytoplasm compared to the nucleus.

C. elegans Argonautes localize to a variety of germ granules (P granules, Mutator foci, Z granules, and SIMR foci) in the germline and early embryo [70]. We carried out *Ascaris* Argonaute antibody histochemistry on the male and female germlines and early embryo. Using a variety of conditions, we have not been able to observe germ granules in the germline or early embryo. Similarly, other antibodies to *Ascaris* proteins known to localize in germ granules in the early *C. elegans* embryo (Dcp1, Dcp2, DcpS, Cgh-1, eIF4E, and others) also do not identify discrete granules in *Ascaris* early embryos. While it is possible that optimal fixation and other conditions may not have been identified for *Ascaris* germ granules, we have carried out extensive and successful immunohistochemistry in *Ascaris* embryos [71]. It is possible that such granules and functional complexes are present, but do not form large enough complexes to be visualized. Interestingly, CSR-1 antibodies did not identify P-granules in early embryos of *C. briggsae* [65] and P-granules have not been examined in other nematodes.

***Ascaris* Argonaute antibodies IP specific small RNAs**

Ascaris Argonaute IP and small RNA sequencing demonstrate that the Argonautes associate with distinct small RNA classes that are complementary to different elements in the genome (Figure 2A). For all IP and small RNA sequencing experiments, we carried out two or more biological replicates. The Argonaute IPs and small RNAs obtained are highly reproducible and the small RNA sequencing per sample was over 30 million reads on average and often over 50 million (Table S2). In previous work we used a variety of approaches including preparing small RNA libraries that were both 5'-dependent and 5'-independent to determine if *Ascaris* small RNAs had 5'-monophosphate or 5' triphosphate termini [51]. In this study, most small RNA libraries were 5'-independent capturing all small RNAs regardless of their 5' phosphates (see Materials and Methods). However, AsALG-1 and AsALG-4 associated small RNA libraries were also constructed using 5'-dependent libraries to enrich for 5'-monophosphate RNAs (see Figure S3). *Ascaris* AsALG-1 specifically binds 5'-monophosphate miRNAs (many previously identified in [51]), AsALG-4 binds 5'-monophosphate 26G-RNAs, and the WAGOs typically bind 5'-triphosphate 22G-RNAs (22-24G-RNAs). Neither *Ascaris* miRNAs nor 26G-RNAs are represented at high abundance compared to the predominant 22G-RNAs (22-24G-RNAs) present in the total small RNA population (see Figure 2A, input). However,

these small RNA are highly enriched in AsALG-1 or AsALG-4 antibody IPs, respectively, demonstrating specificity and enrichment of the Argonaute IPs (Figure 2A). The majority of small RNAs in all *Ascaris* tissues and stages examined are the secondary, 5-triphosphate 22G-RNAs and they are primarily associated with AsWAGO-1. Much lower abundance 5-monophosphate 22G-RNAs are also present. Their relationship to the triphosphate 22G-RNAs could be as primary siRNAs or dephosphorylated triphosphate 22G-RNAs [72]. Argonaute IP from tissues where an Argonaute is expressed at low levels generates small RNA profiles reflecting the input or are random as seen in the ovary, and 4-cell IP of AsALG-4 (Figure 2A) or WAGO-2 and WAGO-3 in the testis (Figure 4, M6-M7), suggesting non-specific background for these IPs. Overall, the data indicate that the *Ascaris* antibodies are specific reagents for analysis of *Ascaris* Argonautes.

Ascaris chromosomes exhibit a relatively uniform distribution of genes [67] (Figure S4). However, repetitive sequences and mobile elements are biased toward DNA that is eliminated during programmed DNA elimination [66-68]. These eliminated sequences are at the ends of all germline chromosome (including 50–100 kb of subtelomeric repeat sequence that precedes the telomeres) and in the middle of a few chromosomes. The eliminated sequences include 30 Mb of a 120 bp tandem repeat that is present in internal and terminal regions of chromosomes. Based on a liberal prediction, we estimate that repetitive and mobile sequences constitute ~41% or 126 Mb of the *Ascaris* germline genome (see Materials and Methods).

The majority of the *Ascaris* small RNAs are complementary to repetitive sequences and/or mobile elements in the genome (Fig. 2A, input and Figure S4). However, simple and satellite repeats such as the 120 bp tandem repeats are typically not targeted by small RNAs. Small RNAs to telomeric sequences have been observed in *C. elegans* [73], but no *Ascaris* small RNAs to telomeric sequences were observed in our in-depth and comprehensive sequencing (Table S2). AsWAGO-1, AsWAGO-2 and AsNRDE-3 associated 22G-RNAs target the repetitive regions of the genome (5,243 repeats high-confident targets; see Materials and Methods and Table S3). We hereafter refer to this targeted repetitive sequence set as WAGO-repeats as they are targeted by *Ascaris* WAGO Argonautes. These WAGO-repeat regions are enriched with ≥ 10 -fold small RNA reads (relative to mean coverage) and cover ~9.5 Mb or 3.4% of the genome (Table S3 and see also Figure 5A-B). The majority (60-80%) of the AsWAGO-1, AsWAGO-2, and AsNRDE-3 associated small RNAs sequenced map to these WAGO-repeats (Figure 2A, bottom horizontal percentage bar). AsWAGO-1 and AsWAGO-2 small RNA targets largely overlap, with AsWAGO-1 and its small RNAs being expressed at much higher levels. AsNRDE-3 small RNAs and their targets show greater diversity compared to AsWAGO-1 and AsWAGO-2, particularly in the male germline during later meiotic stages where AsNRDE-3 associated small RNAs also target mRNAs (see below).

To identify *Ascaris* genes that are potentially targeted by small RNAs, we normalized the small RNA reads to mRNAs based on reads-per-kb-per-million (rpkm). We considered mRNAs with antisense small RNAs over 10 rpkm in a stage as high-confidence targets (Table S4). Small RNAs that target mRNAs are generally antisense and fully complementary to mature mRNAs and associated with AsCSR-1, AsWAGO-3, AsNRDE-3 (in the testis), and AsALG-4. The distribution of these small RNAs across mRNAs is typically uniform across transcripts or with a 5' bias (Figure S5). These mRNA targeting Argonautes bind distinct sizes of small RNAs, with AsNRDE-3, AsWAGO-3, AsCSR-1 and AsALG-4 associated small RNAs 22G-, 23G-, 24G-, and 26G-RNAs, respectively (Figure S6).

AsCSR-1 small RNAs appear to target most expressed mRNAs and the levels are often proportional to mRNA target expression in most stages (see Figure S7 and Table S4). These data are consistent with observations in *C. elegans* suggesting that AsCSR-1 likely functions to “license” gene expression [24, 26, 27, 63]. AsWAGO-3 appears orthologous to *C. elegans* C04F12.1 based on the phylogenetic analysis (Figure 1A). The targets of AsCSR-1 and AsWAGO-3 small RNAs are highly overlapping in many stages, particularly the 4-cell embryo (Figure 2B and Table S4). However, AsWAGO-3 is expressed at much lower levels compared to AsCSR-1 in germline and early embryos (Figure 1B). Notably, *Ascaris* small RNAs corresponding to genes in general do not target pre-mRNA introns.

To further examine the diversity and complexity of small RNAs associated with different *Ascaris* Argonautes in the testis, ovary, and early embryo (4-cell), we compared the levels of small RNAs in these IPs matching mRNAs and repeats using principal component analysis (PCA) (Figure 2B-C). Small RNAs to mRNAs are similar in the ovary, more divergent in the embryo, and highly divergent in the testes (Figure 2B). Small RNAs to repetitive sequences associated with Argonautes show a similar relationship (Figure 2C). These data indicate that small RNAs associated with *Ascaris* Argonautes are most diverse in the testes targeting both mRNAs and repeats. While AsCSR-1, AsWAGO-3, AsNRDE-3, and AsALG-4 small RNAs target different numbers of mRNAs in the testis (ALG-4 targets a much smaller number of mRNAs), many of these targeted mRNAs overlap (Figure 2D and Table S4). Repetitive sequences targeted by AsWAGO-1, AsWAGO-2, and AsNRDE-3 small RNAs also largely overlap, particularly in the testis (Figure 2E and Table S3).

Dynamics of *Ascaris* Argonautes and their associated small RNAs during spermatogenesis

Small RNAs and pathways have been extensively analyzed in the *C. elegans* female germline. In contrast, analysis of *C. elegans* small RNAs and pathways are more limited in the male germline [30-32, 74-76]. 21U-RNAs, 26G-RNAs, and 22G-RNAs are all known to play an important role in the *C. elegans* male germline. Our initial PCA analysis of small RNAs suggested a greater diversity of small RNAs in the testis (Figure 2B-C). *Ascaris* is sexually dimorphic and the extended male germline is ~1 meter in length. This

afforded us a unique opportunity to dissect and obtain large amounts of material from distinct regions of the testes, an attribute not available in *C. elegans*. We analyzed Argonautes, small RNAs, and target mRNA expression throughout *Ascaris* spermatogenesis to provide a comprehensive and in-depth analysis of small RNA pathways during nematode spermatogenesis.

We defined five regions of the *Ascaris* male germline based on nuclear morphology and comparison with *C. elegans* (see Materials and Methods) [77, 78]. These regions include the mitotic region, transition zone, early meiosis (pachytene), late meiosis I and II, and spermatids (Figure 3). We dissected and collected seven samples of the male germline (M1-M7) that cover both mitotic and meiotic regions and carried out Argonaute IPs and small RNA sequencing as well as RNA-seq. Total small RNA and RNA-seq profiles were also generated for spermatids (M8) (Figure 4C and Table S4). The RNA-seq expression data indicates that most *Ascaris* Argonaute RNAs are significantly expressed during spermatogenesis (Figure 4B) (except AsALG-5, see Figure 1B). Overall, expression is highest during early stages of spermatogenesis with expression decreasing during diplotene and diakinesis. The exception, however, is AsALG-4 with very low expression during early spermatogenesis and much higher expression during late meiosis (M6-M7).

The IP and small RNA sequencing data further illustrate the specificity of our Argonaute antibodies and reveals the types and targets of the small RNAs associated with each Argonaute during spermatogenesis (Figure 4D). The size distribution of the small RNAs also reveals distinct features of many AsWAGO associated siRNAs in the testis (Figure S6). This suggests flexibility in the biogenesis and processing of these small RNAs or changes in loading or binding properties [79, 80]. Notably, about 20-25% of all the AsCSR-1 small RNAs also begin with an adenine nucleotide. The low levels of AsWAGO-2 and AsWAGO-3 in M6-M7 result in poor IP and small RNA data. Similarly, the low levels of ALG-4 in early spermatogenesis result in small RNA IPs that look like input without distinct 26G-RNAs, whereas in M6-M7 where AsALG-4 is expressed, 26G-RNAs are clearly dominant in the IPs.

The targets of small RNAs associated with some Argonautes also change dramatically during spermatogenesis (Figure 4). For example, AsNRDE-3 targets mainly WAGO-repeats in early stages (M1-M5) but then targets mRNAs in late meiosis (M6-M7). AsALG-4 associated 26G-RNAs only appear in late meiosis (M6-M7) targeting male, meiosis-specific mRNAs. AsWAGO-2 and AsWAGO-3 do not IP any specific small RNAs in the M6-M7. Interestingly, two RdRPs are also upregulated in the pachytene stage (M5, Figure 4B). Overall, the data illustrate major Argonaute and small RNA changes during meiosis (later pachytene) and are consistent with a role for siRNAs in clearing most mRNAs prior to final sperm maturation (see below) as very few mRNAs are present in *Ascaris* spermatids [51, 61].

Repeat targets of small RNAs during spermatogenesis

Ascaris WAGO-1, WAGO-2, and NRDE-3 associated small RNAs extensively target repetitive sequences including mobile elements (defined as WAGO-repeats) during spermatogenesis. Many of these targets overlap with enriched H3K9me2/3 histone marks (Figure 5A-B). However, not all repetitive or mobile elements, H3K9me2/3 regions, and WAGO small RNA targets of the genome overlap. AsWAGO-2 and AsNRDE-3 associated repeat small RNAs decrease as spermatogenesis progresses, particularly in M6-M7 (Figure 5). In the testes, WAGOs target a subset of repetitive sequences at high levels in the mitotic germline. Most of these repetitive sequences are not expressed and thus appear silenced (Figure 5A and Tables S3). However, some WAGO-repeat regions are expressed, particular during pachytene (M4-M7, see Figure 5C and Table S3). Overall, small RNAs target many repetitive sequences or mobile elements during spermatogenesis.

Several previous studies defined *Ascaris* sequences with characteristics of mobile elements [81-84]. We observed significant RNA expression (based on RNA-seq and PRO-seq) in the male and female germline (but not the early embryo) for one of these, a non-LTR R4 retrotransposon that inserts into the ribosomal locus. R4 was previously described as mobile based on its presence in different genomic locations among individuals [82, 84]. In contrast to the expression of R4, our RNA-seq analyses suggest that most other predicted mobile elements (previous work and current predictions) are not highly expressed. However, PRO-seq suggests many of these mobile element loci are bi-directionally transcribed at low levels. This bi-directional transcription could lead to dsRNA substrates for Dicer to generate primary siRNAs, leading to subsequent secondary siRNA generation to further silence the loci. Consistent with this, high levels of AsWAGO-1, AsWAGO-2 and AsNRDE-3 associated 22G-RNAs are observed to target many of these elements.

mRNA targets of small RNAs during spermatogenesis

To define the timing and expression of Argonaute associated small RNAs and their mRNA targets during spermatogenesis, we first identified 5,526 genes targeted by at least one Argonaute (AsCSR-1, AsALG-4, AsNRDE-3, and AsWAGO-3) with antisense small RNAs over 10 rpkm in one or more of the spermatogenesis stages (see Table S4). These genes can be categorized into two broad groups based on their expression profiles. One group includes genes expressed throughout male germline development (M1-M7; 4,363 genes) whereas the other group are genes specifically expressed during meiosis (M5-M7; 1,163 genes, of which 232 are eliminated by programmed DNA elimination in early development [66-68]). Figure 6A shows an example of these two groups as neighboring genes on chromosome 1. The gene on the left (ag_00250, a chloride intracellular channel exc-4) is expressed throughout the male germline whereas the gene on the right (ag_00251, a gene encoding a hypothetical protein) is only expressed during meiosis. Both genes are targeted by AsCSR-1, and their mRNA levels are correlated with CSR-1 siRNA

levels across spermatogenesis. In contrast, the meiosis-specific gene *ag_00251* is expressed during pachytene (M5-M7) and targeted by *AsNRDE-3* in M5-M7 and *AsALG-4* associated 26G-RNAs in M6-M7. Notably, 26G-RNAs are not detected during most of early pachytene (M5, see Figure 4D).

Genome-wide analysis reveals *AsCSR-1* small RNAs target a large, comprehensive set of genes, and the number of genes targeted varies across developmental stages. In contrast, *AsALG-4* is expressed and targets a much smaller number of genes that are specifically expressed primarily during testis meiosis (Figure 6B). While *AsNRDE-3* does not target many mRNAs during early spermatogenesis (M1-M4), the number of mRNA targeted rises to a level similar to *AsCSR-1* in late spermatogenesis (M5-M7) (Figure 6B). In late meiosis (M6), the mRNA targets largely overlap between *AsCSR-1* and *AsNRDE-3* (Figure 6C and S7). While most *AsALG-4* mRNA targets are also targeted by *AsCSR-1* and *AsNRDE-3*, *AsCSR-1* and *AsNRDE-3* target many mRNAs not targeted by *AsALG-4* (Figure 6C).

The expression changes of Argonaute siRNAs and their targeted mRNAs identified two large groups mRNAs that are differentially regulated by these siRNAs (Figure 6D). Many of these *AsCSR-1*-only targets are differentially expressed during spermatogenesis including in the mitotic region and transition zone (Figure 6D). In contrast, *AsALG-4* and its small RNAs are specific to mRNAs expressed during meiosis (M5-M7) and notably levels of these mRNA targets decrease when the *AsALG-4* associated 26G-RNAs are expressed (Figure 6D). 26G-RNAs appear specific to mRNAs expressed during meiosis (~72%, 833 out of 1,163 meiosis-specific genes) (Table S4). In comparison, *AsCSR-1* small RNAs target most of the expressed mRNAs throughout spermatogenesis (~89%, 4,910 out of 5,526 expressed mRNAs) (Table S4), but both the levels of *AsCSR-1* and its small RNAs decrease significantly following M5. Notably, most *AsALG-4* targeted mRNAs are expressed and also targeted by *AsCSR-1* during pachytene (M5), but the corresponding 26G-RNAs are only expressed during later pachytene and meiosis (M6-7) (Figure 6D). Overall, our data suggests a model where *AsCSR-1* likely licenses, fine-tunes and represses the levels of all expressed mRNAs during early spermatogenesis and perhaps throughout spermatogenesis, while *AsALG-4* targets and likely represses meiosis-specific mRNAs. *AsNRDE-3* small RNAs also target mRNAs later in meiosis, preferably at the 5' end of the mRNAs (Figure 6E). Very low levels of mRNA are present in spermatids [51, 61]. The decrease in *AsCSR-1* levels during the later stages of spermatogenesis could result in a reduction in licensing or protection from repression facilitating a decrease in *AsCSR-1* target mRNAs. Prior to the formation of spermatids, *AsCSR-1* in concert with *AsALG-4* and *AsNRDE-3* small RNAs may also repress and facilitate the turnover of all mRNAs (Figure 6D).

AsALG-4 targeted genes are predominantly male meiosis-specific genes. Many of these genes, including major sperm protein genes, sperm motility genes, sperm specific family genes, class P genes, and pp1 phosphatases are also targeted by *C. elegans* ALG-3/4 [31, 74] indicating conservation of 26G-RNA

targets. However, many of the *Ascaris* 26G-RNA targets are unknown hypothetical genes and generally not conserved in other nematodes. New and novel genes have been shown to evolve at high levels in the testis [85-87]. The expression of these genes may require concerted control. AsALG-4 and its associated 26G-RNAs likely play a major role in clearing late meiosis mRNAs during spermatid formation, as the timing of these 26G-RNAs expression (M6-7) is in general later than their targeted mRNAs (M5-7), including AsALG-4 mRNA itself (Figure S8).

Plasticity of small RNA pathways during spermatogenesis

AsNRDE-3 largely targets repetitive sequences in the female germline, early embryo, and early stages of spermatogenesis. However, during the later stages of spermatogenesis, AsNRDE-3 small RNAs largely switch their targeting to mRNAs (Figure 7A). The percentage of mRNAs targeted by NRDE-3 increases from ~10% in the mitotic regions of the male germline to 50% in late pachytene (M6, Figure 4D and Figure 6B). Over 2,600 mRNAs are targeted by AsNRDE-3 during spermatogenesis, with 2,050 mRNAs targeted alone in M6 (Figure 6D). Most of these AsNRDE-3 targets overlap with AsCSR-1 targets (Figure 6C-D, S7 and Table S4). These data indicate a plasticity in *Ascaris* NRDE-3 targets (repetitive sequences vs mRNAs) during spermatogenesis and suggest an important role for AsNRDE-3 during the later stages of spermatogenesis. This plasticity of small RNA targets is also observed in AsWAGO-3 (Figure 7B). Although AsWAGO-3 expression is in general low, in germline tissues AsWAGO-3 targets WAGO-repeats while in early embryos it targets mostly mRNAs (Figure 7B). Overall, these data indicate that plasticity in *Ascaris* Argonautes targets (Figure 7C), with respect to the class of genomic elements, occurs in different developmental stages and tissues. These observations raise questions regarding the biogenesis of the small RNAs, their loading into the Argonautes, and how these processes change.

Discussion

Our understanding of nematode small RNA pathways comes from detailed studies in *C. elegans* [8-12, 70, 88]. However, nematodes represent a very diverse phylum adapted to a wide variety of lifestyles with species present in almost all ecosystems including many parasitic species [37]. Little is known regarding the conservation and functional role of small RNA pathways (other than miRNAs) in these divergent nematodes [51, 89, 90]. Here, we characterize Argonautes and their small RNAs in the parasitic nematode *Ascaris*. *Ascaris* is a Clade III nematode that diverged from *C. elegans* (Clade V) ~365-400 million years ago [49, 50]. *Ascaris* has a reduced set of expressed Argonautes compared to *C. elegans* (10 vs.19) including 5 AGO-clade and 5 WAGO-clade Argonautes (Figure 1A). *Ascaris* (as well as most non-Clade V nematodes) lacks piRNAs (21U-RNAs) and PIWI Argonautes [51, 90]. We generated specific polyclonal antibodies to seven *Ascaris* Argonautes and analyzed small RNAs associated with AsALG-1, AsALG-4, AsWAGO-1, AsWAGO-2, AsWAGO-3, AsCSR-1, and AsNRDE-3 in early embryos, the female germline, and throughout the male germline. Overall, several small RNA pathways and their functions appear conserved between *Ascaris* and *C. elegans* while others have diverged in function and targets or been lost.

The most abundant small RNAs in *Ascaris* are 22G-RNA secondary siRNAs. The majority are complementary to repetitive sequences and are associated with *Ascaris* WAGOs. However, other *Ascaris* WAGOs, including CSR-1, WAGO-3, and NRDE-3, also associate with 22G-24G RNAs that are complementary to mature mRNAs. Detailed analysis of AsCSR-1 and AsWAGO-3 small RNAs indicates they are primarily 24G-RNAs and 23G-RNAs, respectively (Figure S6), illustrating additional flexibility in average small RNAs sizes for their respective Argonautes. AsNRDE-3 and AsWAGO-3 small RNAs target mRNAs and/or repetitive sequences depending on the developmental stage. AsALG-1 binds miRNAs. AsALG-4 binds 26G-RNAs (like *C. elegans* testis ALG-3/4) that target mRNAs expressed primarily in the male germline during meiosis. AsALG-4 function may differ in *Ascaris* compared to *C. elegans*, whereas the AsCSR-1 pathway appears to function similarly as in *C. elegans* [24-27, 63, 91, 92] even in the absence of an *Ascaris* piRNA pathway.

A licensing CSR-1 pathway in *Ascaris* in the absence of a piRNA pathway. It has been suggested that a central role of piRNAs in *C. elegans* is to define “self” vs “non-self” (*C. elegans* genes vs foreign elements) [56]. *C. elegans* piRNAs have limited, miRNA-like complementarity to their targets (seed sequences, nucleotides 2-8, and additional 3' sequences, nucleotides 14-19), thus providing the potential for targeting a vast array of sequences [53, 93, 94]. It has been proposed that a role of the *C. elegans* CSR-1 pathway is for “licensing” gene expression, thus enabling expression of genes, to counteract promiscuous piRNA silencing of germline transcripts [24, 26, 58]. Given the absence of piRNAs and PIWI in *Ascaris*, we sought to determine the targets of AsCSR-1 small RNAs and thus the potential function of

the *Ascaris* CSR-1 pathway. AsCSR-1 represents the most abundant Argonaute mRNA in the *Ascaris* germline and early embryos. Its associated 24G-RNAs are complementary to almost all transcribed mature mRNAs. This suggests the AsCSR-1 pathway may also function in “licensing” gene expression in *Ascaris*.

C. elegans CSR-1 has also been proposed to “fine tune” gene expression of mRNAs loaded into oocytes [91], clear maternal mRNAs [25], and degrade mRNAs with non-optimal codons [95]. AsCSR-1 is present in oocytes and at its highest level during the maternal to zygotic transition which initiates at the one-cell stage in *Ascaris* prior to pronuclear fusion [61]. AsCSR-1 disappears early in *Ascaris* development at the 32-64 cell stage. These data are consistent with AsCSR-1 also playing a similar role as observed in *C. elegans* in fine tuning oocyte transcripts and clearing mRNAs during the maternal to zygotic transition. The balance among these functions of AsCSR-1 (licensing, fine tuning, and repressive) may be developmental stage and/or mRNA dependent and remains to be determined even in *C. elegans*.

***Ascaris* Argonautes and small RNAs in spermatogenesis.** Small RNA pathways and targets throughout the developmental stages of nematode spermatogenesis have not been analyzed. *Ascaris* is sexually dimorphic. We used the ~1 meter length of the male *Ascaris* germline to comprehensively examine mRNA, Argonautes and their associated small RNAs, and the small RNA targets throughout nematode spermatogenesis. A striking feature of *Ascaris* small RNA pathways is the unique expression of AsALG-4 and associated 26G-RNAs during the later stages of spermatogenesis (M6-M7). This is followed by a significant decrease in the mRNAs targeted by the 26G-RNAs (Figure 6). AsALG-4 and 26G-RNAs may be similar in function to pachytene piRNAs in mice that facilitate mRNA clearance in spermatids [96-99]. This function may be enhanced with AsCSR-1 and AsNRDE-3 that also appear to repress mRNAs during pachytene and later meiosis (see below). Thus, independently evolved small RNA pathways may have convergent functions and play key roles during pachytene of spermatogenesis.

In *C. elegans*, it has been proposed that 26G-RNAs likely identify spermatogenesis mRNA targets for generation of secondary 22G-RNAs [29-32, 74]. This does not appear to be the case in *Ascaris* as our current data further refine our previous finding [51] and indicate that AsALG-4 and 26G-RNA expression does not precede the appearance of secondary 22G-RNAs to the same mRNA targets. In contrast, our data suggests that 26G-RNAs appear much later than most of the small RNAs targeting mRNAs during spermatogenesis and 26G-RNAs are associated with the loss of mRNAs in late spermatogenesis and the formation of spermatids (Figure 6).

The current *Ascaris* data are consistent with a model where AsCSR-1 small RNAs likely license gene expression during early stages of spermatogenesis, the female germline, and early embryo similar to the licensing role of CSR-1 in *C. elegans* [24, 26, 27, 63]. We show that AsALG-4 26G-RNAs appear during

late stages of spermatogenesis commensurate with when their target mRNAs disappear. AsALG-4 26G-RNAs are uniformly distributed across their target transcripts (current data and [51]). This contrasts with the 5' and 3' end targeting of transcripts of *C. elegans* ALG-3/4 26G-RNAs during spermatogenesis [74]. Thus, AsALG-4 transcript targeting differs from *C. elegans* ALG-3/4 indicating flexibility and differences in targeting between nematodes. The AsALG-4 transcript targeting appears more similar to *C. elegans* ERGO-1 26G-RNAs that uniformly target across transcripts [74].

Strikingly, AsNRDE-3 22G-RNA targets are largely repetitive sequences during early spermatogenesis but during the later stages (M6-M7) their targets are predominantly mRNAs. *C. elegans* NRDE-3 is involved in the nuclear RNAi pathway and contributes to co-transcriptional repression (inhibiting elongation) [60, 100]. AsNRDE-3 small RNAs target the 5' ends of transcript in M6-M7. This may be associated with co-transcriptional repression of targets. The decrease in AsCSR-1 levels during the later stages of spermatogenesis may lead to loss of licensing or protection from repression also resulting in a decrease in AsCSR-1 target RNAs. AsCSR-1 could also play a repressive role as recently described [25, 95]. Notably, targets of AsCSR-1, AsALG-4, and AsNRDE-3 in M6-M7 largely overlap. *Ascaris* spermatids exhibit very low levels of mRNAs and small RNAs. We suggest that during the later stages of *Ascaris* meiosis these Argonautes, their changes in expression, and small RNAs act in concert to clear mRNAs in the formation of spermatids. Overall, complex and diverse small RNA pathways contribute to key patterns of gene regulation likely involving licensing during early spermatogenesis and transcriptional and post-transcriptional repression during the later stages of *Ascaris* spermatogenesis.

Plasticity of *Ascaris* WAGO-3 and NRDE-3 targets. *Ascaris* WAGO-1 and WAGO-2 associated small RNAs only target repetitive sequences. However, AsWAGO-3 and AsNRDE-3 small RNAs, depending on the stage and tissue, can target either repetitive sequences or mature mRNAs. AsWAGO-3 small RNAs in early embryos predominantly target mature mRNAs that overlap with AsCSR-1 (AsWAGO-3 RNA is expressed at much lower levels than AsCSR-1) whereas in the male gonad they largely target repetitive sequences. Sequence alignment and phylogenetic analyses suggests AsWAGO-3 appears similar to *C. elegans* C04F12.1 and clusters with CSR-1. *C. elegans* NRDE-3 IP and 22G-RNA analysis indicates the small RNAs target intergenic repetitive loci [101]. In *Ascaris*, NRDE-3 also targets primarily repetitive sequences during early spermatogenesis, the female germline, and in 4-cell embryos. However, during the later stages of spermatogenesis, a large percentage of NRDE-3 small RNAs target mRNAs. Secondary 22G-siRNAs in *C. elegans* are generated by RNA-dependent RNA polymerase following identification of target RNAs by primary siRNAs. The change in small RNA targets in AsWAGO-3 and AsNRDE-3 raises questions regarding how distinct types of targets are identified as substrates for the biogenesis of small RNAs, the RdRPs used to generate the 22G-RNAs, and the loading of these RNAs into distinct Argonautes in different stages [79, 80].

***Ascaris* primary siRNAs.** Several studies suggest that introduction of dsRNA leads to RNA interference in *Ascaris* embryos (Davis, R.E., unpublished), larvae [102, 103], the isolated and perfused adult intestine [104], and muscle [105] through injection of dsRNA into the pseudocoelom of adults. The key *C. elegans* RDE-1 Argonaute binds exogenous siRNAs derived from dsRNA. *Ascaris* appears to lack an RDE-1 Argonaute. What *Ascaris* Argonaute then acts on introduced dsRNA or endogenous dsRNA to generate primary siRNAs that induce the generation of secondary siRNAs? Most of our phylogenetic analyses suggest that AsALG-5 and AsALG-7 appear related *C. elegans* ALG-3/4, while in a few cases, they cluster with either RDE-1 or ERGO-1. AsALG-7 expression is highest in early stages of spermatogenesis. However, we observe very low levels of 26G-RNAs in early spermatogenesis suggesting AsALG-7 may not bind 26G-RNAs. The lack of HENN-1 and 26G-RNAs in other stages suggests that AsALG-5 and AsALG-7 are not likely orthologs of ERGO-1. It is possible that the widely expressed AsALG-7 might serve functions similar to *C. elegans* RDE-1. Overall, with no *Ascaris* piRNAs and 26G-RNAs not likely serving as primary siRNAs to initiate secondary siRNA generation, the origin of endogenous primary siRNAs in *Ascaris* that lead to the abundant 22G-RNAs is unclear. Recent data suggest that in some *C. elegans* stages or compartments the CSR-1 slicer activity [22] may initially function to target and cleave mRNAs thereby serving to identify RNAs for RdRPs to initiate the generation of secondary siRNAs for CSR-1 [95]. A similar mechanism may occur in *Ascaris*.

Evolutionary comparison of Argonautes and small RNA pathways in nematodes

A diverse array of repetitive sequences are targeted by 22G-RNAs in *Ascaris* and these are the most abundant small RNAs. Notably, *Ascaris* Argonautes that associate with small RNAs targeting repetitive sequences (AsWAGO-1, AsWAGO-2, and AsNRDE-3) do not have the catalytic tetrad for slicing activity. Many mobile elements in *Ascaris* are transcribed bidirectionally at low levels and may generate dsRNAs for primary siRNA generation that may then lead to secondary, amplified 22G-RNAs and suppression of mobile element expression. High levels of these 22G-RNAs target mobile elements in most developmental stages. Many, but not all of these loci exhibit repressive histone marks including H3K9me2/3 (Figure 5). The *Ascaris* genome lacks *de novo* DNA methyltransferases and cytosine DNA methylation appears absent in many nematodes [106, 107]. Thus, a key role of *Ascaris* small RNAs and their pathways is likely to silence mobile elements and other repetitive sequences through repressive histone chromatin marks.

Other Clade III-IV nematodes also target mobile elements in the absence of piRNAs. Although 22G-RNAs appear absent in Clades I & II [90], other types of small RNAs are generated to mobile elements in these clades. These small RNAs are likely derived from RdRPs generating dsRNA that are cleaved by Dicer similar to plant and fungal systems that generate monophosphate siRNAs [2]. Cytosine DNA methylation is present in some Clade I nematodes [107-109] and siRNAs may lead to DNA methylation and

transcriptional repression of mobile elements. Taken together, different nematodes appear to use diverse mechanisms to repress mobile elements [90].

Previous analysis of the evolutionary conservation of Argonautes and different types of small RNAs in nematodes suggest that the complexity of Argonautes and small RNA pathways in *Caenorhabditids* is not conserved throughout nematodes [8, 40, 41, 51, 89, 90, 106, 110-112]. The piRNA pathway is present only in Clade V nematodes; *C. elegans* CSR-1 and/or C04F12.1 orthologs are present in some Clade III, V and IV nematodes, but appear absent in Clade I and II nematodes; and nuclear Argonautes (HRDE-1 or NRDE-3) are present in many *Caenorhabditids* and some Clade III nematodes, but they appear absent in other Clades. Here, we have shown general conservation of *Caenorhabditid* CSR-1, ALG-3/4, and WAGO associated small RNAs and their targeting in *Ascaris* (Clade III). Secondary siRNAs (22G-RNAs or longer) are present in Clades III-V, but they are absent in Clades I-II. *Strongyloididae* (Clade IV) appear to have 27GA-RNAs instead of 22G-RNAs and the secondary siRNAs of *Globodera pallida* (Clade IV) are 22-26 nt in length. Differences in the sizes of small RNAs is also observed in *Ascaris* with discrete small RNAs (22-24G) associated with different WAGOs. Although orthologs of ALG-3/4 appear present in most nematodes, 26G-RNAs appear generally absent; *Ascaris* appears to be an exception with abundant 26G-RNAs during spermatogenesis meiosis. If 26G-RNAs in other nematodes are restricted to meiotic regions of the male germline as seen in *Ascaris*, they may not have been enriched in samples used in previous analyses of those nematodes.

22G-RNAs and transgenerational inheritance

Transgenerational inheritance in *C. elegans* is associated with piRNAs, 22G-RNAs, and histone H3K9me3 and H3K27me3 marks [11, 35, 88, 113]. *C. elegans* HRDE-1 is important for maintenance of inheritance over several generations whereas NRDE-3 appears limited to one generation. An interesting question is whether Argonautes and small RNA pathways are associated with transgenerational inheritance in *Ascaris* and other nematodes. Transgenerational inheritance in *C. elegans* has been associated with environmental information (odors, temperature, food, other stress, etc.), behavior, lifespan, and viral and bacterial pathogens [35, 88, 114-119]. In parasitic nematodes, transgenerational inheritance of information like these and particularly host immune responses would be clearly beneficial. For example, exosomes from the parasitic nematode, *Heligmosomoides polygyrus (bakeri)*, contain 22G-RNAs and a WAGO Argonaute, and can suppresses Type 2 innate responses and eosinophilia [111, 120]. Nematode Argonautes and small RNA pathways may play key roles in host-parasite interactions and communication among nematodes. As many of these nematode Argonautes and small RNA pathways are worm-specific and thus differ from their hosts, they could be potential targets for new therapies against parasitic nematodes.

Summary

Nematoda is a diverse phylum including free-living and parasitic species. *Ascaris* lacks piRNAs but maintains a CSR-1 pathway with small RNAs that target most expressed mRNAs. Thus, AsCSR-1 appears to act in a “licensing” pathway as observed in *C. elegans* in the absence of a piRNA pathway as well as serving other tuning or repressive functions. *Ascaris* ALG-4 associated 26G-RNAs target male meiosis-specific genes commensurate with their mRNA degradation and do not appear to act as primary siRNAs for targeting and generating secondary 22G-RNAs during spermatogenesis. Several *Ascaris* Argonautes, including AsNRDE-3 and AsWAGO-3 targets are stage dependent altering their targets between mRNAs and repeats during spermatogenesis and development. Our data significantly expand our understanding of the conservation, divergence, and flexibility of nematode Argonautes and small RNA pathways. Nematode Argonautes and small RNA pathways may have initially evolved for specific types of gene regulation leading to better fitness, but over time may have been co-opted for other functions in diverse nematodes.

Materials and Methods

Ascaris

Collection of *Ascaris* tissues, sperm, and embryonation were as previously described [51, 61]. Male and female germline tissues were washed in PBS, frozen whole or following dissection in liquid nitrogen, and stored at -80°C (see Figure S9, male germline). Lysates from frozen total germline tissue or dissected male germline were prepared by first grinding the frozen samples in a mortar and pestle in liquid nitrogen followed by homogenization in a metal dounce at 4°C in 20 mM Tris-HCl pH 7.9, 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT. Nuclei were lysed by addition of 20 volumes of 20 mM HEPES pH 7.6, 300 mM NaCl, 0.2 mM EDTA, 1 mM DTT, 7.5 mM MgCl₂, 1M urea, 1% NP-40(v/v), 200 U of NxGen RNase Inhibitor (Lucigen), and protease inhibitor cocktail (Roche) [121]. Fresh *Ascaris* 4-cell embryos were decoated and directly homogenized in a metal dounce and nuclei lysed as described above. Lysates were then used for total RNA isolation or immunoprecipitation.

Antibodies

We generated polyclonal antibodies to *Ascaris* fusion proteins for AsALG-1, AsALG-4, AsCSR-1, AsWAGO-1, AsWAGO-2, AsWAGO-3, and AsNRDE-3 and also to peptides for AsWAGO-2 and AsWAGO-3. AsALG-1 amino acids 1-165 (HQ611964), AsALG-4 amino acids 1-75 (HQ611965), AsCSR-1 amino acids 1-162 (HQ611969), AsWAGO-1 amino acids 1-113 (HQ611970) AsWAGO-2 amino acids 1-141 (HQ611971), AsWAGO-3 amino acids 49-202 (HQ611972), and AsNRDE-3 amino acids 45-246 (HQ611973) were fused in frame with GST using the Bam H1 site of pGEX-6P-1 (GE Healthcare), the proteins were expressed in *E. coli*, purified using Glutathione Sepharose 4B columns, and used as the immunogens for the initial boost to generate polyclonal rabbit antibodies (Covance). For additional boosts, we used the *Ascaris* proteins cleaved from GST-fusions bound to Glutathione Sepharose 4B by treatment with PreScission protease (GE Healthcare). AsWAGO-2 peptides (YASRRMFELTKQSRDDYRA and RGGQETSSSSSGEAGHLDI) and AsWAGO-3 peptides (VISYGRGKARRKEFPKQAGT and RENDLPRNGESVDWNRITD) were fused to KLH and used as immunogens to generate polyclonal rabbit antibodies (Pocono Rabbit Farm & Laboratory, Inc.). Antibodies were affinity purified using either the peptides linked to SulfoLink beads (Pierce) or the *Ascaris* proteins (without GST) linked to a mix of Affigel-10/15 (BioRad).

Embryo Immunohistochemistry

Ascaris embryo immunohistochemistry was carried as described [61] using a modified freeze-crack method to permeabilize and fix embryos. Briefly, decoated embryos were suspended in 50% methanol and 2% formaldehyde solution and were frozen and thawed 3 times using a dry ice/ethanol bath. The embryos were re-hydrated with 25% methanol in PBS pH 7.4 for 1 min. After washing twice with PBS pH 7.4, the embryos were incubated in signal enhancer solution (Invitrogen I36933) for 30 min at RT. The embryos

were then re-suspended in blocking solution (0.5% BSA in PBS pH7.4) for 30 min at RT, followed by overnight incubation in primary antibodies at 4°C, and then a 2 hr incubation in secondary antibodies (Invitrogen) at room temperature. Nuclei were stained with DAPI.

Cytological analysis of the *Ascaris* male germline

Regions of 1-5 cm were collected from 1% formaldehyde-fixed dissected gonads at defined distances from the distal tip. Germline fragments were placed on a glass slide with PBS pH 7.4 and gently rolled over with a glass rod to release germline tissue from the outer somatic sheath. Samples were next stained with DAPI and the slides mounted in Prolong anti-fade medium (Invitrogen). Regions were characterized and defined based on nuclear morphology and presence or absence of mitotic/meiotic structures as follows: 1. Mitotic: round-shape interphase nuclei with evenly distributed chromatin and defined nucleolus and the presence of mitotic metaphases and anaphases. 2. Transition zone: more compact chromatin with punctate aspects and irregular shape, no mitoses observed. 3. Meiosis I – Pachytene: Chromatin aspects similar to a “bowl of spaghetti”, increasing nuclear size from early (closer to TZ) to late (closer to meiotic progression). 4. Meiosis 1 & 2: increased chromatin compaction (diplotene), visualization of individual bivalents (diakinesis), further localization to the metaphase plate (metaphase I) and segregation to opposite poles with lagging material in the middle corresponding to sex chromosomes (anaphase I). While no clear observations of meiosis II prophase to anaphase were obtained, groups of 4 nuclei in very close proximity to each other indicated completion of the second meiotic division and formation of haploid nuclei. 5. Spermatids: dispersed, very small and compact nuclei.

Immunohistochemistry of the *Ascaris* germline was carried out on the same samples as described above. Tissue fragments were kept on the slide throughout the whole immunohistochemistry procedure, with careful pipetting of the solutions to avoid sample loss. Samples were first permeabilized in 0.2% Triton X-100/PBS for 15 min at room temperature, followed by 2 washes in PBS. Next, samples were suspended in blocking solution (1% BSA in PBS pH 7.4), covered with parafilm and incubated for 1 hr at room temperature in a wet chamber. Primary and secondary antibody (Invitrogen) incubations were performed overnight at 4 C and for 2 hr at RT, respectively, with slides covered with parafilm and placed in wet chamber. After staining of nuclei with DAPI, slides were mounted in anti-fade medium (Invitrogen) and kept in the dark at room temperature for 24 h before imaging.

Image Acquisition

Ascaris germline immunohistochemistry and DAPI-stained preparations were imaged on an Applied Precision DeltaVision microscope, using a 60X immersion objective and FITC/DAPI excitation filter set. Images were deconvolved with Applied Precision’s Softworx software and analyzed using Fiji software. *Ascaris* embryos immunohistochemistry images were captured on a Leica DM6B fluorescence

microscope, equipped with a Leica DFC 7000 T camera, using a 40X objective, FITC/DAPI excitation filter set and LASX software.

Argonaute Immunoprecipitation

Argonaute IPs were carried out using affinity purified rabbit antibodies pre-bound to Protein A Dynabeads (Fisher Scientific) (5-10 ug of antibody per 100 ul of beads) [122]. Germline or embryo whole cell lysates were mixed with Protein A Dynabeads and rotated overnight at 4°C. Protein A Dynabeads were recovered and washed with high-salt buffer (50 mM Tris-HCl, pH 7.4, 1M NaCl, 1mM EDTA, 1% Igepal CA-630, 0.1% SDS, 0.5% sodium deoxycholate) three times. The beads were resuspended in 250 ul of Proteinase K buffer containing 200 µg/ml Proteinase K and incubated for 1 hr at 37°C. RNA was extracted using Trizol LS (Invitrogen) adapted for small RNA recovery by precipitation using 2 volumes of isopropanol, 15 µg of GlycoBlue (Invitrogen), and 30 minutes incubation at -80°C followed by centrifugation at 18,500 x g for 35 minutes at 4°C. The RNA was treated with RppH (25 units, New England Biolabs) in Thermopol buffer for 5' independent cloning [123]. Both untreated and RppH-treated samples were done for AsALG-1 and AsALG-4 samples. Following RppH treatment, samples were repurified with Trizol LS (Invitrogen) (adopted for small RNAs extraction) and stored at -80°C.

RNA, small RNA libraries and sequencing

RNA was isolated using Trizol or LS Trizol (Invitrogen). Total RNA samples were fractionated into small RNA (< 200 nt) and long RNA (>200nt) using the Monarch RNA cleaner (New England Biolabs). Small RNA libraries were prepared using the Small RNA-Seq Library Prep Kit (Lexogen). Both 5'-phosphate and 5'-phosphate independent libraries (RppH-treated) [123] were prepared. Libraries from 4-cell embryos were also prepared using 18-30 nt gel purified RNA using the SMARTer smRNA-Seq Kit (Clontech) and NEXTflex-Small-RNA-Seq (New England Biolabs) with similar results. RNA >200 nt was treated with TURBO DNase (Ambion) and rRNA depletion carried out using RiboCop rRNA Depletion Kit for Human/Mouse/Rat (Lexogen). Long RNA (>200 nt) libraries were made using CORALL Total RNA-Seq Library Prep Kit (Lexogen). Both small and long RNA libraries were sequenced 150 nt from both ends using the Illumina NovaSeq 6000 System.

Small RNA data analysis

Bioinformatic processing of small RNA sequencing data was carried out as previously described [51]. Briefly, adaptor sequences were trimmed and then the reads collapsed to non-redundant datasets. The reads were then mapped to the *Ascaris* genome and other various datasets [124] to determine targets and their expression levels. To define complementary regions in the genome corresponding to AsWAGO-1 and AsWAGO-2 small RNAs, we first mapped reads from each Argonaute IP-small RNA library to the genome using bowtie [125]. We normalized the coverage for each library to 30 million reads, a number close to the

average number of raw input reads for the libraries. With an average length of ~22 nt for the small RNAs, this corresponds to ~2.5x coverage of the 280-Mb genome. As AsWAGO-1 and AsWAGO-2 small RNA targets are very similar, we merged genomic regions in close proximity (within 2,000 bp) if they have ≥ 20 fold average coverage (50 x) from any of the 20 AsWAGO-1 and AsWAGO-2 libraries (using mergeBed -d 2000). This merge is necessary as a locus is often not fully covered with high-levels of siRNAs. This approach defined an initial set of enriched loci (9,164 with length ≥ 50 bp). We then used the normalized reads (rpkm) mapped to these loci and a 10-fold enrichment in at least one library to define these genome regions as enriched for AsWAGO-1 and AsWAGO-2 small RNAs. These regions were further filtered to remove loci related to rRNA, tRNAs, miRNAs or mitochondrial DNA. Overall, we defined 3,948 AsWAGO-1 and AsWAGO-2 targeted (10x enriched) genomic regions that constitutes 8.1 Mb of sequence and named them WAGO-repeats. Using the same approach, we defined 2,912 (6.2 Mb) AsNRDE3 targets. The majority (~77%) of the sequences defined as AsNRDE-3 targets overlap with AsWAGO-1/2 targets. See Table S3 and S4 for the defined WAGO-repeats and mRNAs and their WAGO-associated small RNAs. PCA analysis on the small RNA libraries was done using DESeq2 plotPCA [126]. Heatmaps were generated using Treeview 3 [127].

Repetitive sequence identification.

Repetitive sequences were identified using a combination of homology-based and *de novo* approaches, including RepeatMasker [128], LTRharvest [129], RepeatScout [130], RepeatExplorer [131], dnaPipeTE [132], MGEScan-nonLTR [133], Helsearch [134], MITE-Hunter [135], SINEfinder [136], TEdenovo [137], and RepARK [138]. The final collection of repetitive sequences was filtered for redundancy with CD-hit and are available in within the repeat track at http://genome.ucsc.edu/s/jianbinwang/Ascaris_small_RNAs.

Data access.

The small RNA and RNA sequencing data is deposited to NCBI GEO database (accession number pending). The data is also available in UCSC Genome Browser track data hubs [139] that can be access with this link: http://genome.ucsc.edu/s/jianbinwang/Ascaris_small_RNAs.

Acknowledgements

We thank Richard Komuniecki, Bruce Bamber, Jeff Myers, and Routh Packing Co. for their support and hospitality in collecting *Ascaris* material. We thank Adam Wallace and Stella Kratzer for initial work on the purification, evaluation, and western blot analysis of the *Ascaris* Argonaute antibodies. We thank Diane Shakes and Diana Chu for advice and suggestions for the analysis of *Ascaris* spermatogenesis and members of the *C. elegans* small RNA research community for comments and feedback on the manuscript. This work was supported by NIH grants to R.E.D. (AI049558 and AI114054) and J.W. (AI155588) and startup funds from the University of Tennessee at Knoxville to J.W.

References

1. Bartel, D.P. (2018). Metazoan MicroRNAs. *Cell* 173, 20-51.
2. Gutbrod, M.J., and Martienssen, R.A. (2020). Conserved chromosomal functions of RNA interference. *Nat Rev Genet*.
3. Hourii-Zeevi, L., and Rechavi, O. (2017). A Matter of Time: Small RNAs Regulate the Duration of Epigenetic Inheritance. *Trends Genet* 33, 46-57.
4. Khanduja, J.S., Calvo, I.A., Joh, R.I., Hill, I.T., and Motamedi, M. (2016). Nuclear Noncoding RNAs and Genome Stability. *Mol Cell* 63, 7-20.
5. Ozata, D.M., Gainetdinov, I., Zoch, A., O'Carroll, D., and Zamore, P.D. (2019). PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet* 20, 89-108.
6. Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-854.
7. Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.
8. Almeida, M.V., Andrade-Navarro, M.A., and Ketting, R.F. (2019). Function and Evolution of Nematode RNAi Pathways. *Non-coding RNA* 5.
9. Billi, A.C., Fischer, S.E., and Kim, J.K. (2014). Endogenous RNAi pathways in *C. elegans*. *WormBook*, 1-49.
10. Grishok, A. (2013). Biology and Mechanisms of Short RNAs in *Caenorhabditis elegans*. *Advances in genetics* 83, 1-69.
11. Weiser, N.E., and Kim, J.K. (2019). Multigenerational Regulation of the *Caenorhabditis elegans* Chromatin Landscape by Germline Small RNAs. *Annu Rev Genet* 53, 289-311.
12. Ambros, V., and Ruvkun, G. (2018). Recent Molecular Genetic Explorations of *Caenorhabditis elegans* MicroRNAs. *Genetics* 209, 651-673.
13. Turner, M.J., Jiao, A.L., and Slack, F.J. (2014). Autoregulation of *lin-4* microRNA transcription by RNA activation (RNAa) in *C. elegans*. *Cell Cycle* 13, 772-781.
14. 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.
15. Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol Cell* 31, 79-90.
16. 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.
17. 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.
18. Gu, W., Shirayama, M., Conte, D., Jr., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J., et al. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol Cell* 36, 231-244.
19. Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* 315, 241-244.
20. 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.

21. 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.
22. 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.
23. 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.
24. Cecere, G., Hoersch, S., O'Keeffe, S., Sachidanandam, R., and Grishok, A. (2014). Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. *Nat Struct Mol Biol* **21**, 358-365.
25. Quarato, P., Singh, M., Cornes, E., Li, B., Bourdon, L., Mueller, F., Didier, C., and Cecere, G. (2021). Germline inherited small RNAs facilitate the clearance of untranslated maternal mRNAs in *C. elegans* embryos. *Nature communications* **12**, 1441.
26. Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., Jr., and Mello, C.C. (2013). The *C. elegans* CSR-1 Argonaute Pathway Counteracts Epigenetic Silencing to Promote Germline Gene Expression. *Dev Cell* **27**, 656-663.
27. Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013). Protection of Germline Gene Expression by the *C. elegans* Argonaute CSR-1. *Dev Cell* **27**, 664-671.
28. Gent, J.I., Lamm, A.T., Pavelec, D.M., Maniar, J.M., Parameswaran, P., Tao, L., Kennedy, S., and Fire, A.Z. (2010). Distinct Phases of siRNA Synthesis in an Endogenous RNAi Pathway in *C. elegans* Soma. *Mol Cell* **37**, 679-689.
29. Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D., Jr. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci U S A* **107**, 3582-3587.
30. Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* **107**, 3588-3593.
31. Conine, C.C., Moresco, J.J., Gu, W., Shirayama, M., Conte, D., Jr., Yates, J.R., 3rd, and Mello, C.C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in *C. elegans*. *Cell* **155**, 1532-1544.
32. Han, T., Manoharan, A.P., Harkins, T.T., Bouffard, P., Fitzpatrick, C., Chu, D.S., Thierry-Mieg, D., Thierry-Mieg, J., and Kim, J.K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* **106**, 18674-18679.
33. Pavelec, D.M., Lachowiec, J., Duchaine, T.F., Smith, H.E., and Kennedy, S. (2009). Requirement for the ERI/DICER complex in endogenous RNA interference and sperm development in *Caenorhabditis elegans*. *Genetics* **183**, 1283-1295.
34. Olina, A.V., Kulbachinskiy, A.V., Aravin, A.A., and Esyunina, D.M. (2018). Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry (Mosc)* **83**, 483-497.
35. Minkina, O., and Hunter, C.P. (2018). Intergenerational Transmission of Gene Regulatory Information in *Caenorhabditis elegans*. *Trends Genet* **34**, 54-64.
36. Xu, F., Guang, S., and Feng, X. (2018). Distinct nuclear and cytoplasmic machineries cooperatively promote the inheritance of RNAi in *Caenorhabditis elegans*. *Biol Cell* **110**, 217-224.
37. De Ley, P. (2006). A quick tour of nematode diversity and the backbone of nematode phylogeny. *WormBook*, 1-8.

38. Eisenhauer, N., and Guerra, C.A. (2019). Global maps of soil-dwelling nematode worms. *Nature* 572, 187-188.
39. van den Hoogen, J., Geisen, S., Routh, D., Ferris, H., Traunspurger, W., Wardle, D.A., de Goede, R.G.M., Adams, B.J., Ahmad, W., Andriuzzi, W.S., et al. (2019). Soil nematode abundance and functional group composition at a global scale. *Nature* 572, 194-198.
40. Britton, C., Laing, R., and Devaney, E. (2019). Small RNAs in parasitic nematodes - forms and functions. *Parasitology*, 1-10.
41. Hoogstrate, S.W., Volkens, R.J., Sterken, M.G., Kammenga, J.E., and Snoek, L.B. (2014). Nematode endogenous small RNA pathways. *Worm* 3, e28234.
42. Blaxter, M., and Koutsovoulos, G. (2015). The evolution of parasitism in Nematoda. *Parasitology* 142 Suppl 1, S26-39.
43. Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstraete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M., et al. (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature* 392, 71-75.
44. Smythe, A.B., Holovachov, O., and Kocot, K.M. (2019). Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evol Biol* 19, 121.
45. van Megen, H., van den Elsen, S., Holterman, M., Karssen, G., Mooyman, P., Bongers, T., and al., e. (2009). A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* 11, 927-950.
46. Easton, A., Gao, S., Lawton, S.P., Bennuru, S., Khan, A., Dahlstrom, E., Oliveira, R.G., Kepha, S., Porcella, S.F., Webster, J., et al. (2020). Molecular evidence of hybridization between pig and human *Ascaris* indicates an interbred species complex infecting humans. *eLife* 9.
47. Jourdan, P.M., Lamberton, P.H.L., Fenwick, A., and Addiss, D.G. (2018). Soil-transmitted helminth infections. *Lancet* 391, 252-265.
48. Wang, J., and Davis, R.E. (2020). *Ascaris*. *Curr Biol* 30, R423-R425.
49. Blaxter, M.L. (2009). Nematodes (Nematoda). In *The Timetree of Life*, S.B. Hedges and S. Kumar, eds. (Oxford University Press), pp. 247-250.
50. Xie, Y., Wang, S., Wu, S., Gao, S., Meng, Q., Wang, C., Lan, J., Luo, L., Zhou, X., Xu, J., et al. (2021). Genome of the giant panda roundworm illuminates its host shift and parasitic adaptation. *bioRxiv*, 2021.2005.2029.446263.
51. Wang, J., Czech, B., Crunk, A., Wallace, A., Mitreva, M., Hannon, G.J., and Davis, R.E. (2011). Deep small RNA sequencing from the nematode *Ascaris* reveals conservation, functional diversification, and novel developmental profiles. *Genome Res* 21, 1462-1477.
52. Ashe, A., Sapetschnig, A., Weick, E.M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs Can Trigger a Multigenerational Epigenetic Memory in the Germline of *C. elegans*. *Cell* 150, 88-99.
53. Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E.A. (2012). Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* 337, 574-578.
54. Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte, D., Jr., and Mello, C.C. (2012). *C. elegans* piRNAs Mediate the Genome-wide Surveillance of Germline Transcripts. *Cell* 150, 78-87.
55. Luteijn, M.J., van Bergeijk, P., Kaaij, L.J., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in *Caenorhabditis elegans*. *The EMBO journal* 31, 3422-3430.
56. Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., Jr., and Mello, C.C. (2012). piRNAs Initiate an Epigenetic Memory of Nonsel self RNA in the *C. elegans* Germline. *Cell* 150, 65-77.

57. Kasper, D.M., Gardner, K.E., and Reinke, V. (2014). Homeland security in the *C. elegans* germ line: insights into the biogenesis and function of piRNAs. *Epigenetics : official journal of the DNA Methylation Society* *9*, 62-74.
58. Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013). Protection of germline gene expression by the *C. elegans* Argonaute CSR-1. *Dev Cell* *27*, 664-671.
59. Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* *489*, 447-451.
60. 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.
61. Wang, J., Garrey, J., and Davis, R.E. (2014). Transcription in Pronuclei and One- to Four-Cell Embryos Drives Early Development in a Nematode. *Current biology : CB* *24*, 124-133.
62. Charlesworth, A.G., Lehrbach, I.J., Seroussi, U., Renaud, M.S., Molnar, R.I., Woock, J.R., Aber, M.J., Diao, A.J., Ruvkun, G., and Claycomb, J.M. (2020). Two isoforms of the essential *C. elegans* Argonaute CSR-1 differentially regulate sperm and oocyte fertility through distinct small RNA classes. *bioRxiv* <https://doi.org/10.1101/2020.07.20.212050>.
63. 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 are required for holocentric chromosome segregation. *Cell* *139*, 123-134.
64. Nguyen, A.D., and Phillips, C.M. (2021). Arginine methylation promotes siRNA-binding specificity for a spermatogenesis-specific isoform of the Argonaute protein CSR-1. *Nature Communications*.
65. Tu, S., Wu, M.Z., Wang, J., Cutter, A.D., Weng, Z., and Claycomb, J.M. (2015). Comparative functional characterization of the CSR-1 22G-RNA pathway in *Caenorhabditis* nematodes. *Nucleic Acids Res* *43*, 208-224.
66. Wang, J., Mitreva, M., Berriman, M., Thorne, A., Magrini, V., Koutsovoulos, G., Kumar, S., Blaxter, M.L., and Davis, R.E. (2012). Silencing of germline-expressed genes by DNA elimination in somatic cells. *Dev Cell* *23*, 1072-1080.
67. Wang, J., Veronezi, G.M.B., Kang, Y., Zagoskin, M., O'Toole, E.T., and Davis, R.E. (2020). Comprehensive Chromosome End Remodeling during Programmed DNA Elimination. *Curr Biol*.
68. Wang, J., Gao, S., Mostovoy, Y., Kang, Y., Zagoskin, M., Sun, Y., Zhang, B., White, L.K., Easton, A., Nutman, T.B., et al. (2017). Comparative genome analysis of programmed DNA elimination in nematodes. *Genome Res* *27*, 2001-2014.
69. Zagoskin, M., Wang, J., Veronezi, G.M.B., and Davis, R.E. (2021). Nematode Argonautes, small RNAs, and programmed DNA elimination. In Preparation or Submitted.
70. Sundby, A.E., Molnar, R.I., and Claycomb, J.M. (2021). Connecting the Dots: Linking *Caenorhabditis elegans* Small RNA Pathways and Germ Granules. *Trends Cell Biol*.
71. Kang, Y., Wang, J., Neff, A., Kratzer, S., Kimura, H., and Davis, R.E. (2016). Differential Chromosomal Localization of Centromeric Histone CENP-A Contributes to Nematode Programmed DNA Elimination. *Cell reports* *16*, 2308-2316.
72. Chaves, D.A., Dai, H., Li, L., Moresco, J.J., Oh, M.E., Conte, D., Jr., Yates, J.R., 3rd, Mello, C.C., and Gu, W. (2021). The RNA phosphatase PIR-1 regulates endogenous small RNA pathways in *C. elegans*. *Mol Cell* *81*, 546-557 e545.
73. Frenk, S., Lister-Shimauchi, E.H., and Ahmed, S. (2019). Telomeric small RNAs in the genus *Caenorhabditis*. *RNA* *25*, 1061-1077.
74. Almeida, M.V., de Jesus Domingues, A.M., and Ketting, R.F. (2019). Maternal and zygotic gene regulatory effects of endogenous RNAi pathways. *PLoS Genet* *15*, e1007784.

75. Bezler, A., Braukmann, F., West, S.M., Duplan, A., Conconi, R., Schutz, F., Gonczy, P., Piano, F., Gunsalus, K., Miska, E.A., et al. (2019). Tissue- and sex-specific small RNAomes reveal sex differences in response to the environment. *PLoS Genet* *15*, e1007905.
76. Tzur, Y.B., Winter, E., Gao, J., Hashimshony, T., Yanai, I., and Colaiacovo, M.P. (2018). Spatiotemporal Gene Expression Analysis of the *Caenorhabditis elegans* Germline Uncovers a Syncytial Expression Switch. *Genetics* *210*, 587-605.
77. Shakes, D.C., Wu, J.C., Sadler, P.L., Laprade, K., Moore, L.L., Noritake, A., and Chu, D.S. (2009). Spermatogenesis-specific features of the meiotic program in *Caenorhabditis elegans*. *PLoS Genet* *5*, e1000611.
78. Chu, D.S., and Shakes, D.C. (2013). Spermatogenesis. *Advances in experimental medicine and biology* *757*, 171-203.
79. Brosnan, C.A., Palmer, A.J., and Zuryn, S. (2021). Cell-type-specific profiling of loaded miRNAs from *Caenorhabditis elegans* reveals spatial and temporal flexibility in Argonaute loading. *Nature communications* *12*, 2194.
80. Gudipati, R.K., Braun, K., Gypas, F., Hess, D., Schreier, J., Carl, S.H., Ketting, R.F., and Grosshans, H. (2021). Protease-mediated processing of Argonaute proteins controls small RNA association. *Mol Cell*.
81. Aeby, P., Spicher, A., de Chastonay, Y., Muller, F., and Tobler, H. (1986). Structure and genomic organization of proretrovirus-like elements partially eliminated from the somatic genome of *Ascaris lumbricoides*. *EMBO J* *5*, 3353-3360.
82. Burke, W.D., Muller, F., and Eickbush, T.H. (1995). R4, a non-LTR retrotransposon specific to the large subunit rRNA genes of nematodes. *Nucleic Acids Res* *23*, 4628-4634.
83. Felder, H., Herzceg, A., de Chastonay, Y., Aeby, P., Tobler, H., and Muller, F. (1994). Tas, a retrotransposon from the parasitic nematode *Ascaris lumbricoides*. *Gene* *149*, 219-225.
84. Neuhaus, H., Muller, F., Etter, A., and Tobler, H. (1987). Type I-like intervening sequences are found in the rDNA of the nematode *Ascaris lumbricoides*. *Nucleic Acids Res* *15*, 7689-7707.
85. Rodelsperger, C., Ebbing, A., Sharma, D.R., Okumura, M., Sommer, R.J., and Korswagen, H.C. (2021). Spatial Transcriptomics of Nematodes Identifies Sperm Cells as a Source of Genomic Novelty and Rapid Evolution. *Mol Biol Evol* *38*, 229-243.
86. Rodelsperger, C., Prabh, N., and Sommer, R.J. (2019). New Gene Origin and Deep Taxon Phylogenomics: Opportunities and Challenges. *Trends Genet* *35*, 914-922.
87. Van Oss, S.B., and Carvunis, A.R. (2019). De novo gene birth. *PLoS Genet* *15*, e1008160.
88. Rechavi, O., and Lev, I. (2017). Principles of Transgenerational Small RNA Inheritance in *Caenorhabditis elegans*. *Curr Biol* *27*, R720-R730.
89. Holz, A., and Streit, A. (2017). Gain and Loss of Small RNA Classes-Characterization of Small RNAs in the Parasitic Nematode Family Strongyloididae. *Genome Biol Evol* *9*, 2826-2843.
90. Sarkies, P., Selkirk, M.E., Jones, J.T., Blok, V., Boothby, T., Goldstein, B., Hanelt, B., Ardila-Garcia, A., Fast, N.M., Schiffer, P.M., et al. (2015). Ancient and Novel Small RNA Pathways Compensate for the Loss of piRNAs in Multiple Independent Nematode Lineages. *PLoS biology* *13*, e1002061.
91. Gerson-Gurwitz, A., Wang, S., Sathe, S., Green, R., Yeo, G.W., Oegema, K., and Desai, A. (2016). A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. *Cell* *165*, 396-409.
92. Fassnacht, C., Tocchini, C., Kumari, P., Gaidatzis, D., Stadler, M.B., and Ciosk, R. (2018). The CSR-1 endogenous RNAi pathway ensures accurate transcriptional reprogramming during the oocyte-to-embryo transition in *Caenorhabditis elegans*. *PLoS Genet* *14*, e1007252.

93. Shen, E.Z., Chen, H., Ozturk, A.R., Tu, S., Shirayama, M., Tang, W., Ding, Y.H., Dai, S.Y., Weng, Z., and Mello, C.C. (2018). Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the *C. elegans* Germline. *Cell* **172**, 937-951 e918.
94. Zhang, D., Tu, S., Stubna, M., Wu, W.S., Huang, W.C., Weng, Z., and Lee, H.C. (2018). The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science* **359**, 587-592.
95. Singh, M., Cornes, E., Li, B., Quarato, P., Bourdon, L., Dingli, F., Loew, D., Proccacia, S., and Cecere, G. (2021). Translation and codon usage regulate Argonaute slicer activity to trigger small RNA biogenesis. *Nature Communications* <https://doi.org/10.1038/s41467-021-23615-w>.
96. Goh, W.S., Falciatori, I., Tam, O.H., Burgess, R., Meikar, O., Kotaja, N., Hammell, M., and Hannon, G.J. (2015). piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. *Genes Dev* **29**, 1032-1044.
97. Gou, L.T., Dai, P., Yang, J.H., Xue, Y., Hu, Y.P., Zhou, Y., Kang, J.Y., Wang, X., Li, H., Hua, M.M., et al. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res* **24**, 680-700.
98. Watanabe, T., Cheng, E.C., Zhong, M., and Lin, H. (2015). Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res* **25**, 368-380.
99. Zhang, P., Kang, J.Y., Gou, L.T., Wang, J., Xue, Y., Skogerboe, G., Dai, P., Huang, D.W., Chen, R., Fu, X.D., et al. (2015). MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. *Cell Res* **25**, 193-207.
100. Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* **465**, 1097-1101.
101. Zhou, X., Xu, F., Mao, H., Ji, J., Yin, M., Feng, X., and Guang, S. (2014). Nuclear RNAi contributes to the silencing of off-target genes and repetitive sequences in *Caenorhabditis elegans*. *Genetics* **197**, 121-132.
102. Islam, M.K., Miyoshi, T., Yamada, M., and Tsuji, N. (2005). Pyrophosphatase of the roundworm *Ascaris suum* plays an essential role in the worm's molting and development. *Infect Immun* **73**, 1995-2004.
103. Xu, M.J., Chen, N., Song, H.Q., Lin, R.Q., Huang, C.Q., Yuan, Z.G., and Zhu, X.Q. (2010). RNAi-mediated silencing of a novel *Ascaris suum* gene expression in infective larvae. *Parasitology research* **107**, 1499-1503.
104. Rosa, B.A., McNulty, S.N., Mitreva, M., and Jasmer, D.P. (2017). Direct experimental manipulation of intestinal cells in *Ascaris suum*, with minor influences on the global transcriptome. *International journal for parasitology* **47**, 271-279.
105. McCoy, C.J., Warnock, N.D., Atkinson, L.E., Atcheson, E., Martin, R.J., Robertson, A.P., Maule, A.G., Marks, N.J., and Mousley, A. (2015). RNA interference in adult *Ascaris suum*--an opportunity for the development of a functional genomics platform that supports organism-, tissue- and cell-based biology in a nematode parasite. *International journal for parasitology* **45**, 673-678.
106. Pratz, L., Rancurel, C., Da Rocha, M., Danchin, E.G.J., Castagnone-Sereno, P., Abad, P., and Perfus-Barbeoch, L. (2018). Genome-wide expert annotation of the epigenetic machinery of the plant-parasitic nematodes *Meloidogyne* spp., with a focus on the asexually reproducing species. *BMC genomics* **19**, 321.
107. Rosic, S., Amouroux, R., Requena, C.E., Gomes, A., Emperle, M., Beltran, T., Rane, J.K., Linnett, S., Selkirk, M.E., Schiffer, P.H., et al. (2018). Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity. *Nat Genet* **50**, 452-459.

108. Gao, F., Liu, X., Wu, X.P., Wang, X.L., Gong, D., Lu, H., Xia, Y., Song, Y., Wang, J., Du, J., et al. (2012). Differential DNA methylation in discrete developmental stages of the parasitic nematode *Trichinella spiralis*. *Genome Biol* *13*, R100.
109. Schiffer, P.H., Kroiher, M., Kraus, C., Koutsovoulos, G.D., Kumar, S., Camps, J.I., Nsah, N.A., Stappert, D., Morris, K., Heger, P., et al. (2013). The genome of *Romanomermis culicivorax*: revealing fundamental changes in the core developmental genetic toolkit in Nematoda. *BMC genomics* *14*, 923.
110. Buck, A.H., and Blaxter, M. (2013). Functional diversification of Argonautes in nematodes: an expanding universe. *Biochem Soc Trans* *41*, 881-886.
111. Chow, F.W., Koutsovoulos, G., Ovando-Vazquez, C., Neophytou, K., Bermudez-Barrientos, J.R., Laetsch, D.R., Robertson, E., Kumar, S., Claycomb, J.M., Blaxter, M., et al. (2019). Secretion of an Argonaute protein by a parasitic nematode and the evolution of its siRNA guides. *Nucleic Acids Res* *47*, 3594-3606.
112. Hunt, V.L., Hino, A., Yoshida, A., and Kikuchi, T. (2018). Comparative transcriptomics gives insights into the evolution of parasitism in Strongyloides nematodes at the genus, subclade and species level. *Scientific reports* *8*, 5192.
113. Woodhouse, R.M., and Ashe, A. (2020). How do histone modifications contribute to transgenerational epigenetic inheritance in *C. elegans*? *Biochem Soc Trans* *48*, 1019-1034.
114. Baugh, L.R., and Day, T. (2020). Nongenetic inheritance and multigenerational plasticity in the nematode *C. elegans*. *eLife* *9*.
115. Moore, R.S., Kaletsky, R., and Murphy, C.T. (2019). Piwi/PRG-1 Argonaute and TGF-beta Mediate Transgenerational Learned Pathogenic Avoidance. *Cell* *177*, 1827-1841 e1812.
116. Posner, R., Toker, I.A., Antonova, O., Star, E., Anava, S., Azmon, E., Hendricks, M., Bracha, S., Gingold, H., and Rechavi, O. (2019). Neuronal Small RNAs Control Behavior Transgenerationally. *Cell* *177*, 1814-1826 e1815.
117. Kaletsky, R., Moore, R.S., Vrla, G.D., Parsons, L.R., Gitai, Z., and Murphy, C.T. (2020). *C. elegans* interprets bacterial non-coding RNAs to learn pathogenic avoidance. *Nature* *586*, 445-451.
118. Klosin, A., Casas, E., Hidalgo-Carcedo, C., Vavouri, T., and Lehner, B. (2017). Transgenerational transmission of environmental information in *C. elegans*. *Science* *356*, 320-323.
119. Rechavi, O., Houry-Ze'evi, L., Anava, S., Goh, W.S., Kerk, S.Y., Hannon, G.J., and Hobert, O. (2014). Starvation-Induced Transgenerational Inheritance of Small RNAs in *C. elegans*. *Cell*.
120. Buck, A.H., Coakley, G., Simbari, F., McSorley, H.J., Quintana, J.F., Le Bihan, T., Kumar, S., Abreu-Goodger, C., Lear, M., Marcus, Y., et al. (2014). Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. *Nature communications* *5*, 5488.
121. Brugiolo, M., Botti, V., Liu, N., Muller-McNicoll, M., and Neugebauer, K.M. (2017). Fractionation iCLIP detects persistent SR protein binding to conserved, retained introns in chromatin, nucleoplasm and cytoplasm. *Nucleic Acids Res* *45*, 10452-10465.
122. Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., Konig, J., and Ule, J. (2014). iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* *65*, 274-287.
123. Almeida, M.V., de Jesus Domingues, A.M., Lukas, H., Mendez-Lago, M., and Ketting, R.F. (2019). RppH can faithfully replace TAP to allow cloning of 5'-triphosphate carrying small RNAs. *MethodsX* *6*, 265-272.
124. Wang, J. (2021). Genomics of the Parasitic Nematode *Ascaris* and Its Relatives. *Genes* *12*.
125. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* *9*, 357-359.

126. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* *15*, 550.
127. Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. *Bioinformatics* *20*, 3246-3248.
128. Smit, A.F.A., Hubley, R., and Green, P. (2013-2015). RepeatMasker Open-4.0.
129. Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* *9*, 18.
130. Price, A.L., Jones, N.C., and Pevzner, P.A. (2005). De novo identification of repeat families in large genomes. *Bioinformatics* *21 Suppl 1*, i351-358.
131. Novak, P., Neumann, P., Pech, J., Steinhaisl, J., and Macas, J. (2013). RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* *29*, 792-793.
132. Goubert, C., Modolo, L., Vieira, C., ValienteMoro, C., Mavingui, P., and Boulesteix, M. (2015). De novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the yellow fever mosquito (*Aedes aegypti*). *Genome Biol Evol* *7*, 1192-1205.
133. Rho, M., and Tang, H. (2009). MGEScan-non-LTR: computational identification and classification of autonomous non-LTR retrotransposons in eukaryotic genomes. *Nucleic Acids Res* *37*, e143.
134. Yang, L., and Bennetzen, J.L. (2009). Structure-based discovery and description of plant and animal Helitrons. *Proc Natl Acad Sci U S A* *106*, 12832-12837.
135. Han, Y., and Wessler, S.R. (2010). MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res* *38*, e199.
136. Wenke, T., Dobel, T., Sorensen, T.R., Junghans, H., Weisshaar, B., and Schmidt, T. (2011). Targeted identification of short interspersed nuclear element families shows their widespread existence and extreme heterogeneity in plant genomes. *Plant Cell* *23*, 3117-3128.
137. Flutre, T., Duprat, E., Feuillet, C., and Quesneville, H. (2011). Considering transposable element diversification in de novo annotation approaches. *PLoS ONE* *6*, e16526.
138. Koch, P., Platzer, M., and Downie, B.R. (2014). RepARK--de novo creation of repeat libraries from whole-genome NGS reads. *Nucleic Acids Res* *42*, e80.
139. Raney, B.J., Dreszer, T.R., Barber, G.P., Clawson, H., Fujita, P.A., Wang, T., Nguyen, N., Paten, B., Zweig, A.S., Karolchik, D., et al. (2014). Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. *Bioinformatics* *30*, 1003-1005.

Figure Legends

Figure 1. *Ascaris* has 10 Argonautes. **A.** Argonaute protein phylogenetic tree showing the relationship of *Ascaris* and *C. elegans* Argonautes. *Ascaris* Argonautes are in purple boxes. Labels indicate the type of Argonaute or its associated small RNA. GenBank accession numbers for these Argonaute protein sequences are available in Table S1. Argonautes highlighted in green represent a canonical fly PIWI and *Arabidopsis* AGO Argonaute. The Argonautes were aligned using Muscle and the tree generated using Neighbor Joining and Maximum Likelihood (1,000 bootstraps) with the bootstrap values illustrated. *Ascaris* proteins for which antibodies were generated are marked with Ab. **B.** RNA expression of *Ascaris* Argonautes and RNA-dependent RNA polymerases (RdRPs). The levels of expression (numbers are in rpkm) are shown in a heatmap with higher expression in red. Note that we view an rpkm value that is less than 5 as a background level of expression. The developmental stages are as follows: M1 to M8 are regions of the male germline, with M1 = early mitotic region, M2 = late mitotic region, M3 = transition zone, M4 = transition zone to early pachytene, M5 = pachytene, M6 = late pachytene, M7 = meiosis diplotene to diakinesis, and M8 = spermatids (see Figure 3, text, add Materials and Methods for a detailed description of the male germline stages); F1 to F5 are regions of the female germline, with F1 = mitotic region, F2 = early pachytene, F3 = late pachytene, F4 = diplotene, F5 = oocyte; zygote maturation stages prior to pronuclear fusion isolated from the uterus (E1-4; E4 is the stage passed from *Ascaris* and the host to the environment) [61]; embryo development (at 30° C) stages, with E5 = 24hr (1-cell), E6 = 46hr (2-cell), E7 = 64hr (4-cell), E8 = 96hr (16-cell), E9 = 116hr (32-64-cell), E10 = 7day (256-cell), larvae L1 (10-day) and L2 (21-day) [61]; and adult somatic tissues, with Mu = Muscle, In = intestine and Ca = carcass, which includes the cuticle, hypodermis, muscle, nervous system, and pharynx. **C.** *Ascaris* Argonaute immunohistochemistry in early embryos. DAPI (blue) in left panel and immunohistochemistry (green) in right panel.

Figure 2. *Ascaris* Argonautes bind specific sets of small RNAs. **A.** Size distribution, frequency, and targets of small RNAs associated with specific Argonautes. Small RNAs (18-30 nt) from input and Argonaute immunoprecipitation (IP) were plotted for the whole testis (male germline), whole ovary (female germline), and 4-cell embryo (60 hr). In these and all subsequent small RNA size distribution figures, small RNAs starting with A, C, G, and U of different sizes (x-axis) were plotted against their read frequencies (y-axis; raw reads in millions). Data are derived from two or more biological replicates combined. The small RNAs were categorized by their type (miRNA) or complementarity to different targets as shown in percentage bars below each size distribution plot. The target categories are sequences matching: 1) rRNAs or 2) tRNAs; 3) miRNAs; 4) repetitive sequences and mobile elements that are targets of AsWAGO-1, AsWAGO-2 and AsNRDE-3 (WAGO-repeats); 5) siRNAs antisense to mRNAs; siRNAs matching 6) introns or 7) intergenic regions; 8) small RNAs that have no full-length match to the genome (no match); 9) and sense small RNAs that correspond to mRNAs. **B** and **C.** Principal component analysis (PCA) of small RNAs associated with Argonautes in different stages showing the overall relationship among these small RNA libraries. **B** is PCA analysis of small RNAs that target mRNAs and **C** is small RNAs that target the WAGO-repeats. **D** and **E.** Venn diagram showing the Argonaute small RNA overlapping targets in the testis for mRNAs (**D**) and WAGO-repeats (**E**).

Figure 3. *Ascaris* male gonad regions and nuclear morphology. **A.** Schematic representation of regions of *Ascaris* male germline illustrated that correspond to *C. elegans* male germline and nuclear morphology. Corresponding regions between *C. elegans* and *Ascaris* are kept to scale; however, the length of the *Ascaris* male gonad is ~1 meter. Regions labeled are: 1) mitotic (pink), 2) transition zone (orange), 3) meiosis 1 – pachytene, 4) meiosis 1 & 2 (green), and 5) spermatids (blue). Regions of the *Ascaris* male germline collected for Argonaute IP and RNA-seq are labeled M1-M8, separated by blue vertical lines. **B.** Nuclear morphology of DAPI-stained regions of the *Ascaris* male gonad. **(a-d)** Overview of the mitotic region illustrating the progressive increase in gonad thickness. Mitotic stages are observed (metaphase, anaphase; arrows), confirmed by staining with CENP-A and the mitosis marker H3S10P (data not shown), as well as apoptotic nuclei (*). Larger and less condensed nuclei (*) are also detected, scattered through the mid-plane (dotted circles). **(e)** The transition zone exhibits more condensed and punctate nuclear morphology. **(f-h)** “Spaghetti bowl”-like nuclei characteristic of early **(f)**, middle **(g)** and late pachytene **(h)**.

(i-m) Stages of meiotic progression: diplotene (i); diakinesis, with distinguishable individual bivalents (j); metaphase I (k); anaphase I (l), note lagging sex chromosomes (inset, arrowhead); and second meiotic division resulting in 4 haploid nuclei (m, contour). (n) spermatids.

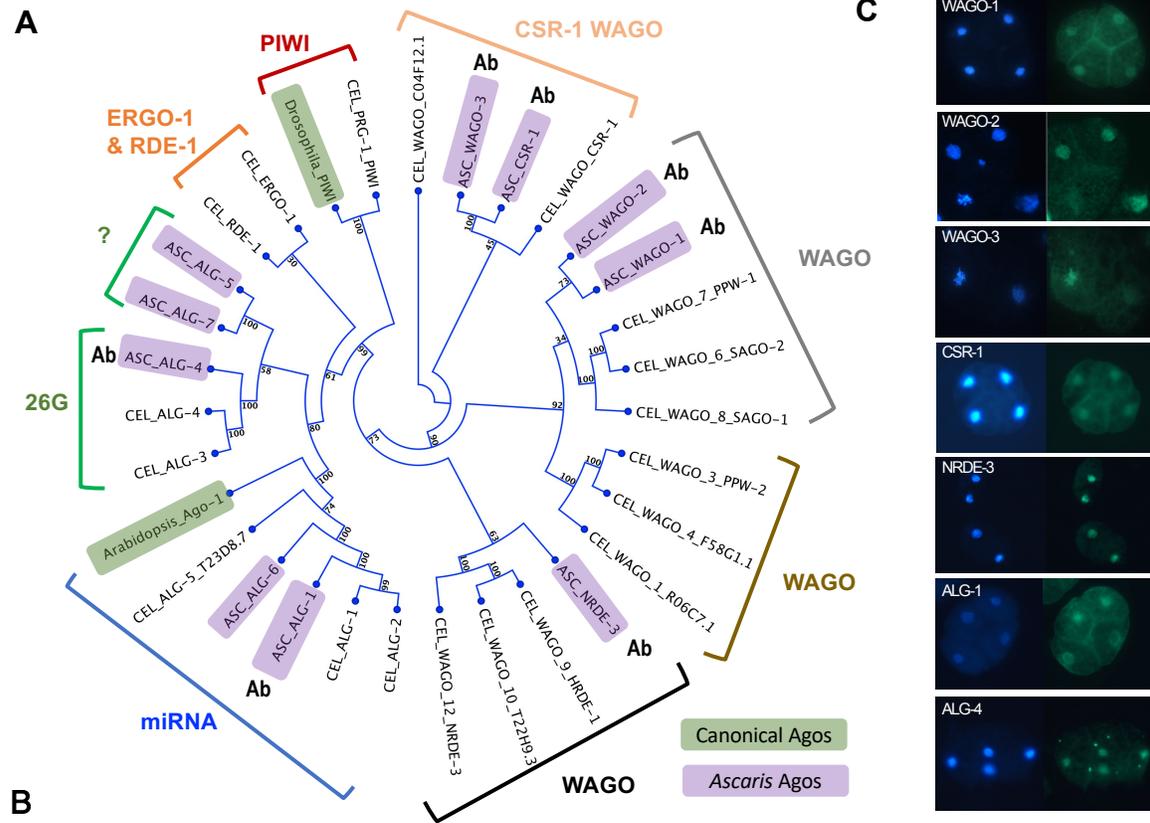
Figure 4. Small RNAs associated with *Ascaris Argonautes* during spermatogenesis. **A.** Illustration of *Ascaris* stages of spermatogenesis from Figure 3. **B.** RNA expression profiles of *Ascaris Argonautes* and RdRPs throughout spermatogenesis. The Argonautes are plotted in two groups based on their expression pattern or their targets. **C.** Small RNAs from mature spermatids (M8). The random size distribution pattern illustrates the low levels of *bona fide* small RNAs. **D.** Small RNAs during spermatogenesis (M1-M7). In **C** and **D**, the size distribution, frequency, and classification of small RNAs are as described as in Figure 2A.

Figure 5. AsWAGO-1, AsWAGO-2 and AsNRDE-3 associated small RNAs target repeats during spermatogenesis. **A.** Genome browser view of a region of chromosome 1 illustrating Argonaute associated small RNAs, RNA expression (RNA-seq), and H3K9me3 levels. Only spermatogenesis stages that exhibit changes in expression or changes in Argonaute small RNA or mRNA expression are illustrated. Note NRDE-3 associated small RNAs change their targets from repeats to mRNAs during pachytene and late meiosis (M5-M7). **B.** *Ascaris* WAGO repeat targets largely overlap with genomic H3K9me3 levels and/or the presence of transposable elements. **C.** siRNA levels to the top targets (top 1,776 WAGO repeat loci; see Table S3) for each WAGO during spermatogenesis. Heatmaps illustrate the standard Z-score (converted from rpk) showing changes in expression of siRNAs associated with AsWAGO-1, AsWAGO-2 and AsNRDE-3. The RNA expression (rpk) of the targets is also illustrated (right). Targeted loci were sorted based on the same order in the four heatmaps.

Figure 6. AsCSR-1 and AsNRDE-3 license and fine-tune mRNA expression during early spermatogenesis whereas AsCSR-1, AsNRDE-3, and AsALG-4 facilitate the clearance of mRNAs during late spermatogenesis. **A.** Genome browser view of a region of chromosome 1 illustrating AsCSR-1, AsNRDE-3 and AsALG-4 associated small RNAs and their mRNA target expression during spermatogenesis. **B.** AsCSR-1, AsNRDE-3 and AsALG-4 mRNA targets during spermatogenesis. AsCSR-1 small RNAs targets are pervasive in different stages and correlate with RNA expression, AsNRDE-3 small RNA targets switch from repeats to mRNAs in pachytene and late meiosis (M5-M7), whereas AsALG-4 is expressed and its associated 26G-RNAs target mRNAs that are primarily restricted to meiosis (M6-M7) in the male gonad. Shown are the number of mRNAs targeted by different Argonautes. The color indicates where these mRNA targets are most highly expressed in developmental stages using the same color scheme as Figure 1B. Note AsCSR-1 and AsNRDE-3 target a broad group of mRNAs, while AsALG-4 targets mostly testis-specific genes. **C.** AsCSR-1, AsNRDE-3 and AsALG-4 targeted mRNAs largely overlap during late meiosis. Venn diagrams showing the relationship between these targets in the M6 stage where all three Argonautes have a large number of targets. **D.** siRNA and targeted mRNA levels in M6 (Figure 6C) during spermatogenesis. Very low levels of small RNAs are in mature spermatids (Figure 4C); thus, a 0 value is used in M8 for the AGO IPs. Heatmaps illustrate the standard Z-score (converted from rpk) showing changes in siRNA and mRNA expression. Targeted genes were sorted based on the same order in the heatmap pairs. **E.** AsNRDE-3 siRNAs are enriched for sequence at the 5' ends of mRNAs. siRNA distribution on mRNAs changes in AsNRDE-3, AsCSR-1 and AsALG-4 during spermatogenesis (see also Figure S5).

Figure 7. Plasticity of *Ascaris* small RNA pathways. **A.** AsNRDE-3 bound small RNAs change their targets from genomic repetitive sequence to mRNAs during spermatogenesis (M5-M7). **B.** WAGO-3 bound small RNAs change their targets from genomic repetitive sequence in male mitotic germline to mRNAs in 4-cell embryos. **C.** Model for *Ascaris* small RNA pathways (see text) (Bolder lines indicate predicted stronger affects and the proportion of small RNAs and their repetitive or mRNA targets are shown as a percentage of all small RNAs below the horizontal lines).

Figure 1



| RNA-seq rpkm | Developmental stage | spermatogenesis to sperm | | | | | | | | oogenesis to oocyte | | | | | zygote | | | | early embryo | | | late embryo | | larvae | | soma | | | |
|--------------|---------------------|--------------------------|-----|-----|-----|-----|----|----|----|---------------------|----|-----|----|----|--------|-----|-----|-----|--------------|-----|----|-------------|----|--------|----|------|----|----|----|
| | | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | F1 | F2 | F3 | F4 | F5 | E1 | E2 | E3 | E4 | E5 | E6 | E7 | E8 | E9 | E10 | L1 | L2 | Mu | In | Ca |
| AGOs | ALG-1 | 21 | 23 | 31 | 28 | 12 | 1 | 1 | 5 | 93 | 54 | 56 | 12 | 3 | 10 | 13 | 18 | 28 | 26 | 39 | 42 | 42 | 58 | 78 | 40 | 8 | 0 | 3 | 2 |
| | ALG-4 | 3 | 2 | 1 | 3 | 34 | 63 | 40 | 4 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 4 | 3 | 3 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | ALG-5 | 8 | 5 | 6 | 4 | 2 | 0 | 0 | 1 | 10 | 4 | 5 | 1 | 2 | 3 | 3 | 3 | 6 | 8 | 13 | 7 | 13 | 13 | 8 | 4 | 1 | 1 | 0 | 1 |
| | ALG-6 | 24 | 28 | 29 | 21 | 12 | 2 | 1 | 5 | 30 | 16 | 17 | 5 | 5 | 23 | 22 | 26 | 34 | 21 | 25 | 18 | 5 | 2 | 1 | 2 | 0 | 1 | 0 | 1 |
| | ALG-7 | 40 | 44 | 25 | 17 | 20 | 6 | 3 | 2 | 10 | 8 | 8 | 4 | 3 | 9 | 8 | 9 | 15 | 14 | 14 | 10 | 17 | 15 | 10 | 9 | 2 | 2 | 1 | 3 |
| WAGOs | CSR-1 | 109 | 132 | 105 | 122 | 141 | 70 | 35 | 18 | 180 | 96 | 104 | 12 | 29 | 127 | 206 | 241 | 336 | 193 | 131 | 52 | 8 | 3 | 4 | 5 | 1 | 2 | 5 | 3 |
| | WAGO-1 | 80 | 81 | 47 | 92 | 65 | 8 | 4 | 11 | 17 | 12 | 13 | 6 | 9 | 88 | 89 | 80 | 125 | 106 | 74 | 51 | 23 | 10 | 6 | 5 | 1 | 8 | 28 | 65 |
| | WAGO-2 | 25 | 26 | 23 | 23 | 10 | 1 | 1 | 5 | 33 | 31 | 33 | 11 | 10 | 49 | 47 | 42 | 37 | 26 | 31 | 17 | 8 | 6 | 5 | 4 | 1 | 1 | 10 | 4 |
| | WAGO-3 | 15 | 16 | 15 | 15 | 9 | 2 | 1 | 3 | 9 | 5 | 5 | 2 | 1 | 2 | 2 | 3 | 3 | 11 | 10 | 10 | 21 | 21 | 19 | 15 | 3 | 3 | 30 | 27 |
| | NRDE-3 | 19 | 22 | 15 | 13 | 38 | 19 | 6 | 3 | 33 | 19 | 20 | 5 | 6 | 25 | 28 | 34 | 33 | 52 | 50 | 31 | 25 | 24 | 12 | 9 | 1 | 3 | 3 | 3 |
| RdRPs | RdRP-1 | 13 | 14 | 12 | 20 | 37 | 7 | 3 | 2 | 9 | 6 | 6 | 2 | 1 | 4 | 3 | 2 | 4 | 3 | 2 | 1 | 2 | 2 | 3 | 3 | 1 | 0 | 2 | 1 |
| | RdRP-2 | 6 | 6 | 4 | 4 | 3 | 0 | 0 | 1 | 3 | 3 | 2 | 2 | 0 | 1 | 1 | 1 | 3 | 2 | 2 | 2 | 4 | 4 | 4 | 3 | 0 | 0 | 1 | 0 |
| | RdRP-3 | 11 | 10 | 7 | 16 | 58 | 4 | 2 | 1 | 14 | 7 | 8 | 6 | 2 | 12 | 12 | 15 | 20 | 11 | 11 | 7 | 9 | 6 | 8 | 3 | 1 | 2 | 1 | 1 |

Figure 2

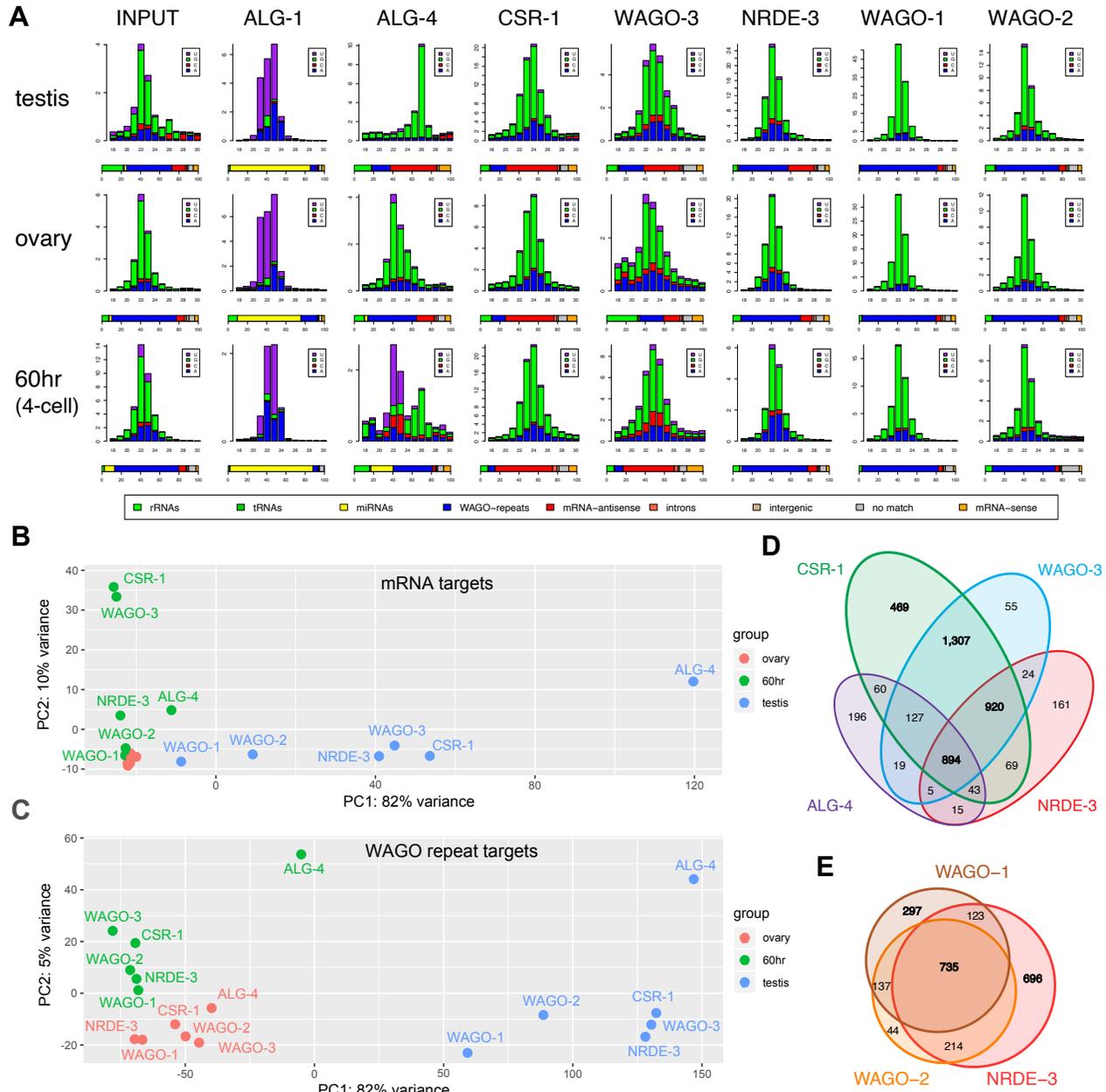


Figure 3

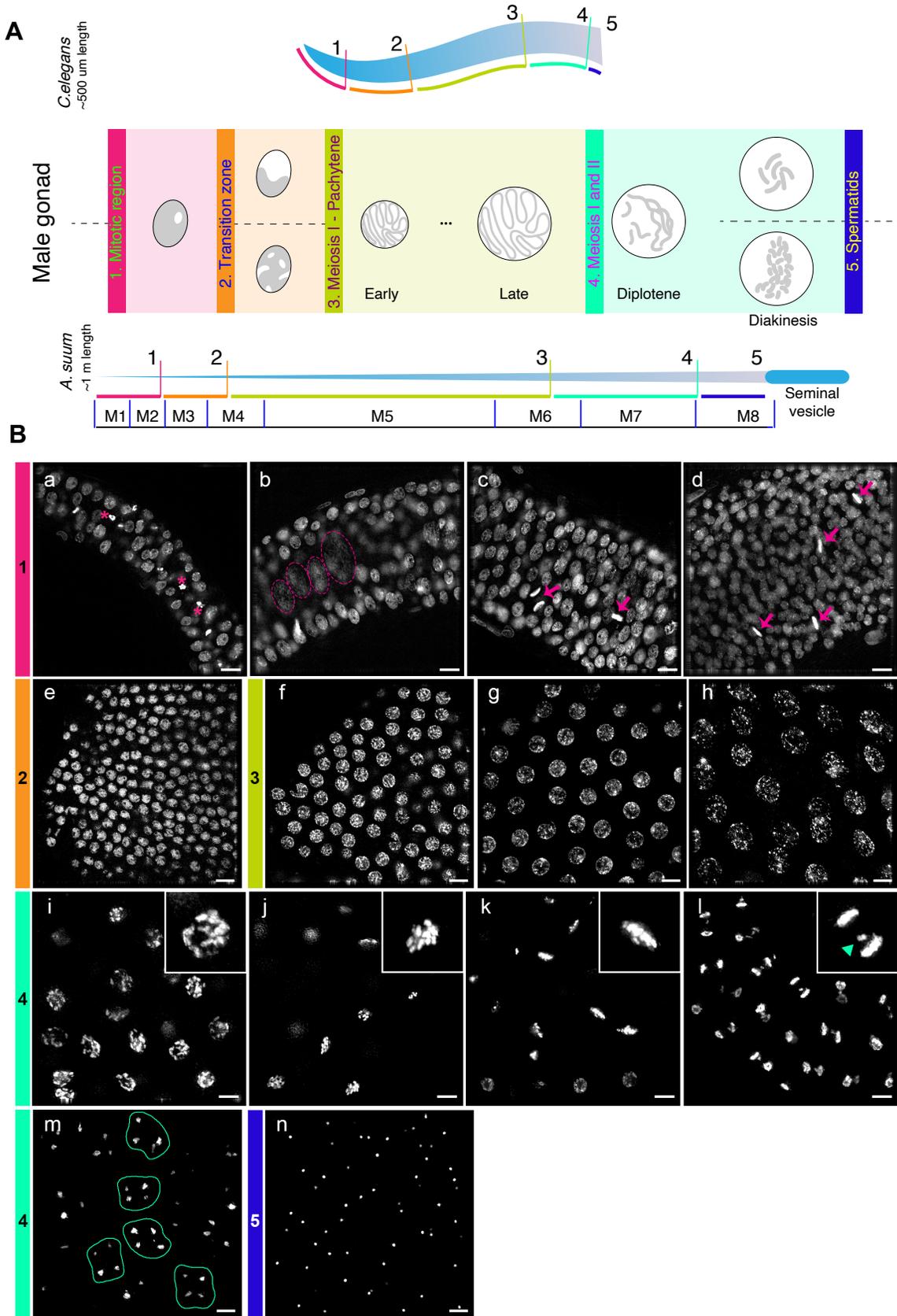


Figure 4

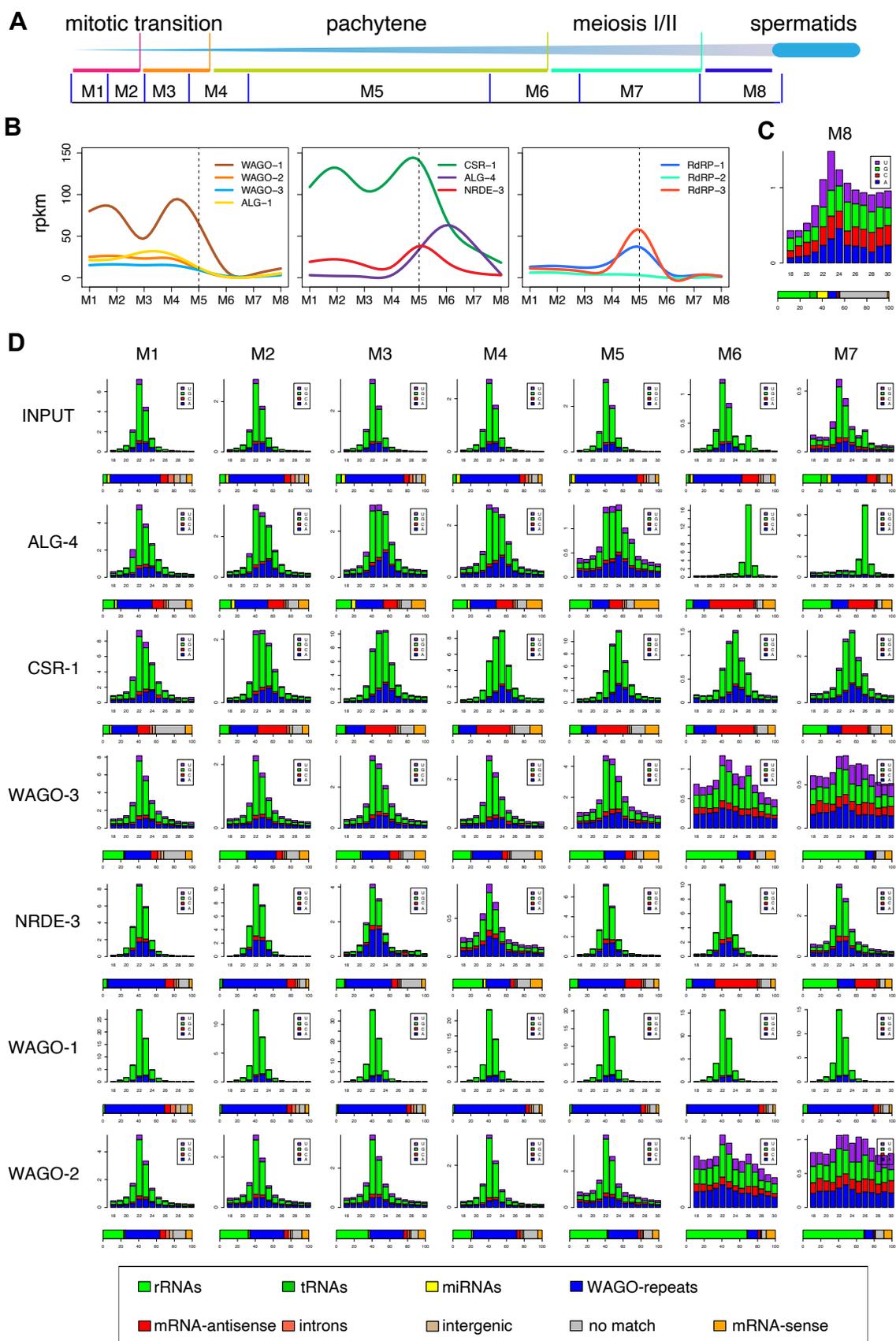


Figure 5

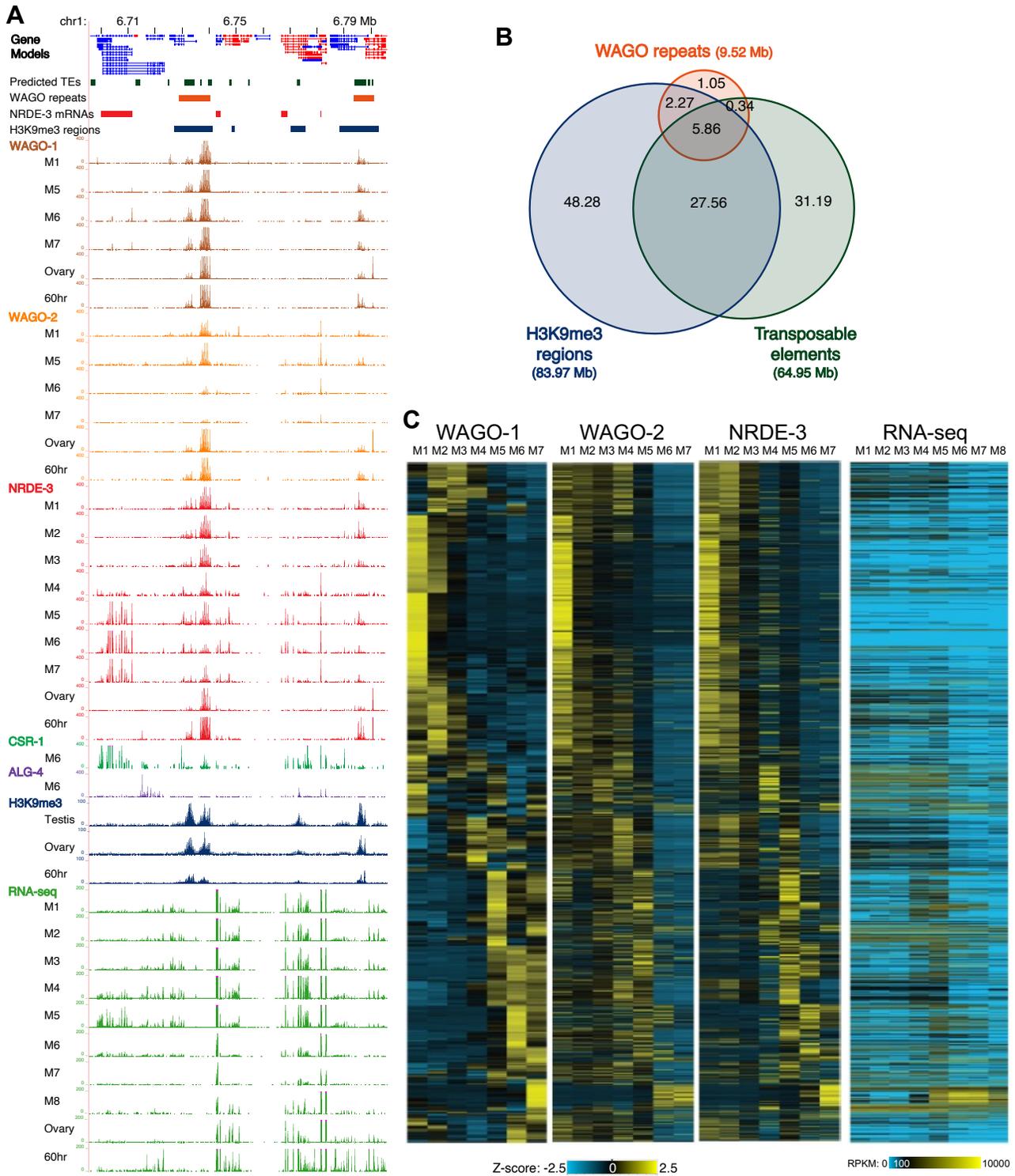


Figure 6

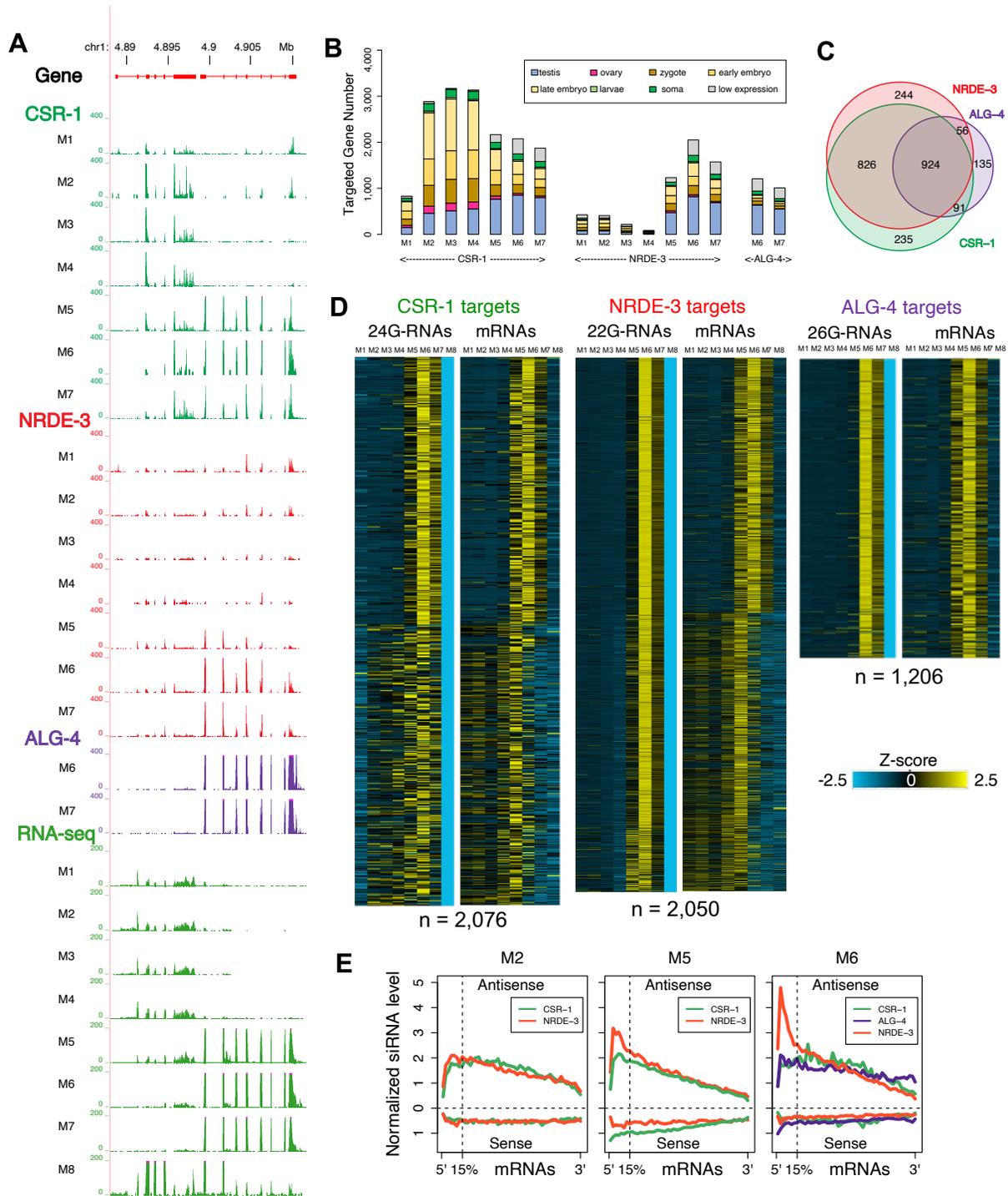


Figure 7

