

## Systematic characterization of small RNAs associated with *C. elegans* Argonautes

Lei Liu<sup>1†</sup>, Xiaolin Wang<sup>1†\*</sup>, Wenfang Zhao<sup>1</sup>, Qiqi Li<sup>1</sup>, Jingxin Li<sup>1</sup>, He Chen<sup>2</sup> & Ge Shan<sup>1,3\*</sup><sup>1</sup>Department of Laboratory Medicine, The First Affiliated Hospital of USTC, the CAS Key Laboratory of Innate Immunity and Chronic Disease, School of Basic Medical Sciences, Division of Life Science and Medicine, University of Science and Technology of China, Hefei 230027, China;<sup>2</sup>Information Materials and Intelligent Sensing Laboratory of Anhui Province, Anhui University, Hefei 230601, China;<sup>3</sup>Department of Pulmonary and Critical Care Medicine, Regional Medical Center for National Institute of Respiratory Diseases, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou 310016, China

Received December 7, 2022; accepted December 28, 2022; published online May 5, 2023

Argonaute proteins generally play regulatory roles by forming complexes with the corresponding small RNAs (sRNAs). An expanded Argonaute family with 20 potentially functional members has been identified in *Caenorhabditis elegans*. Canonical sRNAs in *C. elegans* are miRNAs, small interfering RNAs including 22G-RNAs and 26G-RNAs, and 21U-RNAs, which are *C. elegans* piRNAs. Previous studies have only covered some of these Argonautes for their sRNA partners, and thus, a systematic study is needed to reveal the comprehensive regulatory networks formed by *C. elegans* Argonautes and their associated sRNAs. We obtained *in situ* knockin (KI) strains of all *C. elegans* Argonautes with fusion tags by CRISPR/Cas9 technology. RNA immunoprecipitation against these endogenously expressed Argonautes and high-throughput sequencing acquired the sRNA profiles of individual Argonautes. The sRNA partners for each Argonaute were then analyzed. We found that there were 10 Argonautes enriched miRNAs, 17 Argonautes bound to 22G-RNAs, 8 Argonautes bound to 26G-RNAs, and 1 Argonaute PRG-1 bound to piRNAs. Uridylated 22G-RNAs were bound by four Argonautes HRDE-1, WAGO-4, CSR-1, and PPW-2. We found that all four Argonautes played a role in transgenerational epigenetic inheritance. Regulatory roles of the corresponding Argonaute-sRNA complex in managing levels of long transcripts and interspecies regulation were also demonstrated. In this study, we portrayed the sRNAs bound to each functional Argonaute in *C. elegans*. Bioinformatics analyses together with experimental investigations provided perceptions in the overall view of the regulatory network formed by *C. elegans* Argonautes and sRNAs. The sRNA profiles bound to individual Argonautes reported here will be valuable resources for further studies.

**Argonaute, miRNA, 22G-RNAs, 26G-RNAs, piRNA, uridylation, transgenerational inheritance, interspecies**

**Citation:** Liu, L., Wang, X., Zhao, W., Li, Q., Li, J., Chen, H., and Shan, G. (2023). Systematic characterization of small RNAs associated with *C. elegans* Argonautes. *Sci China Life Sci* 66, 1303–1322. <https://doi.org/10.1007/s11427-022-2304-8>

### INTRODUCTION

Argonautes are a family of conserved proteins that contain N-terminal, Piwi-Argonaute-Zwille (PAZ), middle (MID), and PIWI domains, and generally they elicit regulatory functions by binding to different classes of small RNAs

(sRNAs) (Höck and Meister, 2008; Hutvagner and Simard, 2008). The number of Argonautes is often different in diverse species; for instance, the fission yeast *Schizosaccharomyces pombe* has 1, *Drosophila melanogaster* has 5, *Homo sapiens* and *Mus musculus* have 8, and *Caenorhabditis elegans* has 20 potentially functional Argonautes according to WormBase annotation (version WS285) (Höck and Meister, 2008; Sasaki et al., 2003; Yigit et al., 2006). Argonautes and their associated sRNAs form ribonucleoprotein complexes to

†Contributed equally to this work

\*Corresponding authors (Xiaolin Wang, email: [wxl20089@ustc.edu.cn](mailto:wxl20089@ustc.edu.cn); Ge Shan, email: [shange@ustc.edu.cn](mailto:shange@ustc.edu.cn))

play roles in events such as post transcriptional silencing (PTGS) and transgenerational epigenetic inheritance (TEI) (Bošković and Rando, 2018; Meister, 2013).

sRNAs are generally ~20–40 nt long, and can be divided into subclasses based on their features in biogenesis, length, functionality, etc. (Carthew and Sontheimer, 2009; Höck and Meister, 2008; Vaucheret, 2008). *C. elegans* has been at the forefront of sRNA study, since the discovery of the first microRNA (miRNA) to the milestone discovery in RNA interference (RNAi) pathway at the end of the last century (Fire et al., 1998; Ketting, 2011; Lee et al., 1993; Wightman et al., 1993). miRNAs are ~22 nt, and form an miRNA-induced silencing complex (miRISC) with Argonautes such as ALG-1 and ALG-2 in *C. elegans* (Ambros et al., 2003; Bartel, 2004; Liu et al., 2021; Si et al., 2023; Vasquez-Rifo et al., 2012). Another class of sRNAs is small interfering RNAs (siRNAs), and in *C. elegans*, refers mainly to 22G-RNAs and 26G-RNAs (Conine et al., 2010; Gent et al., 2010; Ruby et al., 2006; Stoeckius et al., 2009). Piwi-interacting RNAs (piRNAs) are another conserved sRNA class, and in *C. elegans*, known as 21U-RNAs (Batista et al., 2008). In *C. elegans*, 21U- and 26G-RNAs are the primary sRNAs, and they trigger the production of secondary 22G-RNAs (Ashe et al., 2012; Chen and Rechavi, 2022; Conine et al., 2010; Gent et al., 2010; Pak and Fire, 2007). 22G-RNAs are also produced from exogenous double-stranded RNAs (dsRNAs) in environmental feeding RNAi (Almeida et al., 2019; Chen and Rechavi, 2022; Tabara et al., 1999).

*C. elegans* probably has the most complex sRNA regulatory network in animals due to the presence of more functional Argonautes and diversified sRNA profiles (Stoeckius et al., 2009; Yigit et al., 2006). A renowned role of miRNAs is mediated by ALG-1 and ALG-2 to silence mRNAs (Bouasker and Simard, 2012; Chen and Rechavi, 2022). For siRNAs, a wide range of functions have been identified, and several Argonautes, such as CSR-1, are known to mediate the roles of siRNAs in *C. elegans* (Claycomb et al., 2009; Moore et al., 2019; Wedeles et al., 2013). HRDE-1 and WAGO-4 are two Argonautes that can bind to uridylated 22G-RNAs, and play roles in the TEI effect triggered by feeding RNAi (Buckley et al., 2012; Lev et al., 2019; Rechavi et al., 2014; Xu et al., 2018). In addition to feeding RNAi, *C. elegans* also demonstrates other phenomena of environmental RNAi, in which RNAs in bacterial food elicit physiological effects via the RNAi pathway in *C. elegans* (Kaletsky et al., 2020; Liu et al., 2012; Samuel et al., 2016). For 21U-RNAs, the Argonaute PRG-1 is essential for their functionality in surveilling germline transcripts (Das et al., 2008; Lee et al., 2012; Shen et al., 2018; Wang and Lin, 2021).

Although substantial understanding has been gained about *C. elegans* Argonautes and sRNAs, the profiles of sRNAs bound by each Argonaute require further disclosure. So far,

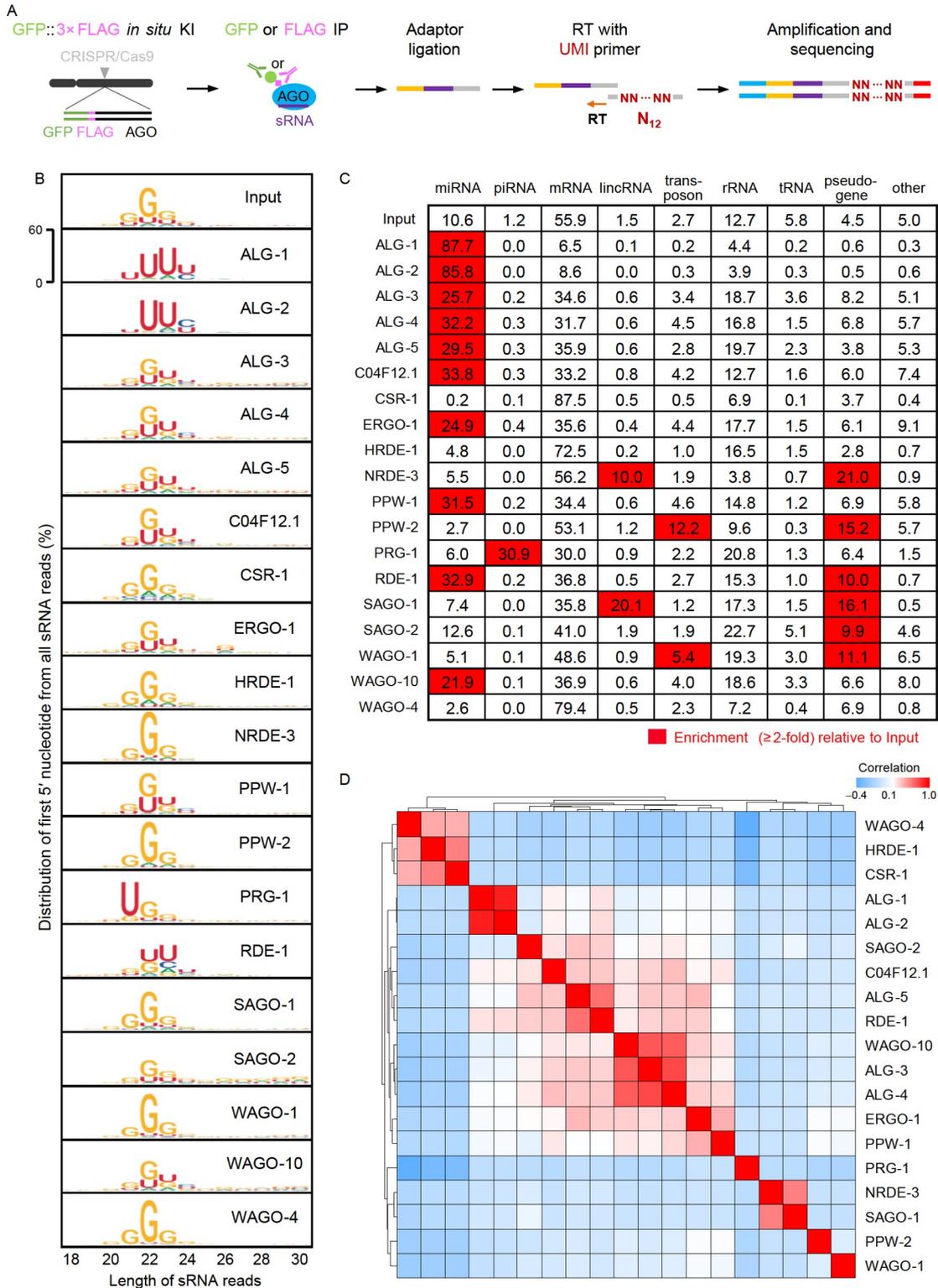
high-throughput RNA sequencing (RNA-seq) to reveal the sRNA binding profiles has been carried out for only four Argonautes, ALG-5, PRG-1, CSR-1, and PPW-2, at endogenous levels with innate spatiotemporal expression patterns (Brown et al., 2017; Nguyen and Phillips, 2021; Schreier et al., 2022; Shen et al., 2018). To gain further insight into Argonautes and sRNAs, we have applied CRISPR/Cas9 technology to generate *in situ* GFP and FLAG fusions for each of the 20 potentially functional Argonautes in *C. elegans*, and then performed RNA immunoprecipitation (RIP) followed with RNA-seq. Through bioinformatics analyses combined with experimental examinations, we have uncovered novel features and provided resources for future studies about Argonautes and their associated sRNAs in *C. elegans*.

## RESULTS

### Workflow to identify the associated sRNAs using knockin fusion of *C. elegans* Argonautes

*C. elegans* has as many as 27 paralogous Argonaute sequences in the genome, and 20 of them are potentially functional to encode Argonaute proteins according to WormBase annotation (current version: WS285) (Figure S1A in Supporting Information). Like most *C. elegans* proteins, specific antibodies for these Argonautes are currently unavailable. For the purpose to perform RIP with endogenously expressed Argonautes, we integrated two tags, GFP and FLAG, *in situ* into the N- or C-terminus of all 20 potentially functional Argonautes using CRISPR/Cas9 technology (Figure 1A) (Friedland et al., 2013). For CSR-1, an attempt to knockin (KI) both GFP and FLAG together failed, and we could only generate a FLAG::CSR-1 fusion. Genotyping with PCR was used to verify the successful knockin of the tags (Figure S1B in Supporting Information).

All strains were crossed with the wildtype N2 worms at least 6 times to clean up possible off-target or side effects of CRISPR/Cas9. By Western blots with an anti-FLAG antibody, we found that 19 out of the 20 Argonautes were expressed as fusion proteins (Figure S2A in Supporting Information). One of the 20 Argonaute genes, *wago-5*, appeared to be a pseudogene, as no fusion protein was detected from the *wago-5* gene, no matter whether the tags were knocked in the theoretical N- or C-terminus (Figures S1B and S2B in Supporting Information). We further evaluated the brood size and lifespan of the 19 tagged Argonaute strains, and found no significant alternation in either trait, when compared with the wildtype N2 worms (Figure S3A and B in Supporting Information). Furthermore, we chose 8 Argonautes with known expression patterns from previous publications, and examined the expression patterns of the corresponding KI strains (Figure S3C in Supporting In-



**Figure 1** Genome-wide analysis of small RNAs bound to *C. elegans* Argonautes. **A**, Experimental procedure of RIP and high-throughput RNA sequencing of 19 functional Argonautes (AGOs). In-frame GFP::3×FLAG tag is introduced into the endogenous locus of 19 AGOs with CRISPR/Cas9 technology. Then, the Argonaute-sRNA complexes are purified with an anti-GFP or anti-FLAG antibody, followed by preparing cDNA libraries with RT primers containing UMI and subjected to high-throughput sequencing. **B**, Seqlogo revealing first 5' nucleotide and size distribution (18–30 nt) of normalized sRNA reads from Input and 19 AGO IP samples. Input is the total sRNAs obtained from whole lysate of wildtype N2 worms. nt, nucleotide. **C**, Percentage of sRNA reads from Input and 19 AGO RIP samples according to annotated genomic loci. Sense reads are counted for miRNAs, piRNAs, and tRNAs. Both sense and antisense reads are counted for rRNAs. Antisense reads are counted for mRNAs, lincRNAs, transposons, and pseudogenes. Filled boxes (red) indicate the AGO-bound sRNA reads are at least 2-fold relative to Input, corresponding to distinct genomic coding region of the specific transcript category. **D**, Hierarchical clustering of sRNA profiles from 19 AGO RIP samples.

formation) (Batista et al., 2008; Buckley et al., 2012; Gu et al., 2009; Guang et al., 2008; Schreier et al., 2022; Vasquez-Rifo et al., 2012; Xu et al., 2018). It was found that all 8 Argonaute KIs demonstrated expression patterns in consistency with previous studies.

We then carried out RIP with the anti-GFP antibody using KI strains of 18 Argonautes (Figure 1A). Actually, for RIP experiments, either GFP or FLAG antibodies have been widely used for tagged proteins in *C. elegans* (Brown et al., 2017; Gu et al., 2009; Shen et al., 2018; Xu et al., 2018; Yu et al., 2017). We prioritized GFP in this study based on the well-established methodology with GFP-Trap Agarose beads for *in situ* KI strains in our laboratory (Yu et al., 2017). For FLAG::CSR-1 fusion, anti-FLAG antibody was used to perform RIP. sRNAs isolated from these RIP experiments were then reverse transcribed to complementary DNA (cDNA) with specific reverse primers containing unique molecular identifier (UMI) tags with 12 random bases during library construction, followed by next-generation RNA-seq (Smith et al., 2017).

### General features of sRNAs associated with *C. elegans* Argonautes

We first compared our sRNA profiles to the sRNA data from previous studies with four endogenously expressed Argonautes ALG-5, PRG-1, CSR-1, and PPW-2 (Figure S4A and B in Supporting Information) (Brown et al., 2017; Nguyen and Phillips, 2021; Schreier et al., 2022; Shen et al., 2018). Our data clustered well with previous studies (Figure S4A and B in Supporting Information). PRG-1 is known to function as an Argonaute for piRNAs, and thousands of piRNAs were primarily enriched in sRNAs from PRG-1 RIP in our study (Figure S4C in Supporting Information). These results indicated that our protocol was suitable for the systematic characterization of sRNAs bound to tagged Argonautes with endogenous expression.

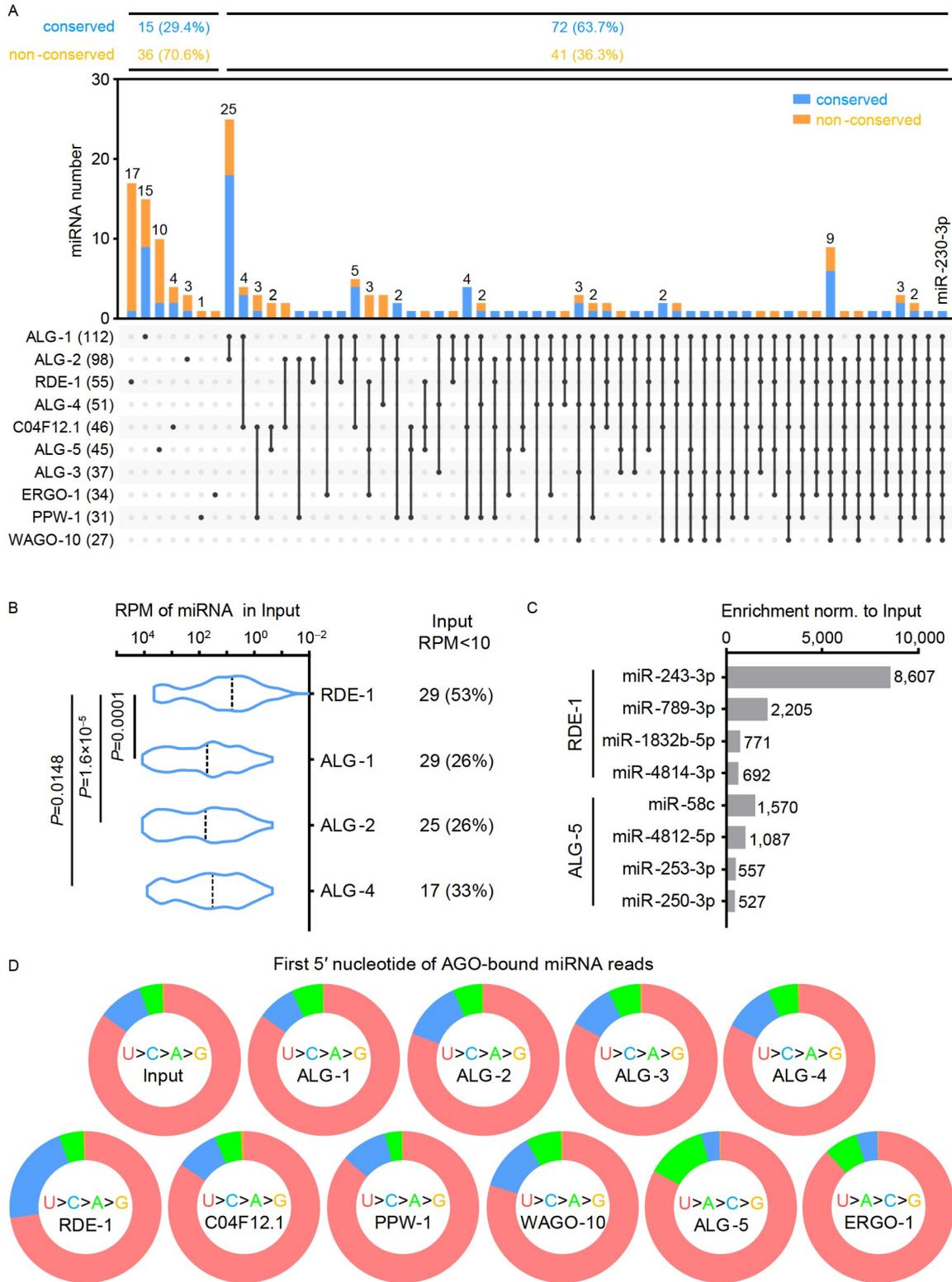
The length distribution and frequency of the first 5' nucleotide of sRNAs bound to each Argonaute were analyzed (Figure 1B; Figure S5 in Supporting Information). The percentage of reads corresponding to genomic regions, including miRNAs, piRNAs, mRNAs, lincRNAs, transposons, rRNAs, tRNAs, pseudogenes, and others that were unclassified, were also evaluated (Figure 1C; Figure S5 in Supporting Information). Several features were noticed from these simple examinations. First, PRG-1 was the only Argonaute that mainly bound to sRNAs of 21 nt and with U as the first base, which were 21U-RNAs (piRNAs) of *C. elegans* (Figure 1B and C; Figure S5 in Supporting Information). ALG-1 and ALG-2 mainly bound to miRNAs, with lengths of 22 or 23 nt and generally with U as the first base (Figure 1B and C; Figure S5 in Supporting Information). Except ALG-1 and ALG-2, all the other 17 Argonautes were

associated with some portion of 22G-RNAs (Figure 1B; Figure S5 in Supporting Information). 26G-RNAs were found with a low percentage in sRNA profiles of several Argonautes, such as ALG-3, ALG-4, and ERGO-1; ERGO-1 was the one with the largest portion of 26G-RNAs (Figure 1B; Figure S5 in Supporting Information).

Next, the 19 functional Argonautes were divided into several major clusters according to cluster analysis based on their associated sRNAs (Figure 1D). The first cluster comprised WAGO-4, CSR-1, and HRDE-1. Consistently, these three Argonautes are known to be expressed in the germline and regulate germline constitutive gene sets (Buckley et al., 2012; Charlesworth et al., 2021; Ortiz et al., 2014; Xu et al., 2018). The second one was ALG-1 and ALG-2, as Argonautes are primarily responsible for miRNA functions (Brown et al., 2017; Zhang et al., 2007). The third one included ALG-5 and RDE-1, and they shared a portion of miRNA partners. The fourth cluster was formed by WAGO-10, ALG-3, and ALG-4, and all of them are involved in spermatogenesis (Charlesworth et al., 2021). Finally, a cluster with NRDE-3 and SAGO-1 might be involved in the regulation of lincRNAs through binding to partially overlapped sRNAs (Figure 1C and D; Figure S5 in Supporting Information). The seven remaining proteins did not cluster with other functional Argonautes, and they might be individually unique according to their associated sRNAs. These clustering results were based on the profiles of sRNAs binding to each Argonaute, and thus were more related to functional characteristics, unlike the evolutionary relationships that were based on the amino acid sequences of important functional domains (Figure 1D; Figure S1A in Supporting Information).

### Argonautes and their corresponding miRNA loads

The first miRNA, *lin-4*, was discovered in *C. elegans* more than two decades ago (Ketting, 2011; Lee et al., 1993; Wightman et al., 1993). There were a total of 15 Argonautes that bound to miRNAs, and among them 10 Argonautes had some miRNA partners enriched for more than 2-fold relative to Input, and miRNAs bound to these Argonautes with enrichment  $\geq 3$ -fold and reads per million (RPM)  $\geq 10$  were further analyzed (Figure 2A; Figure S6A in Supporting Information). ALG-1 and ALG-2 bound to 112 and 98 miRNAs, respectively, out of the 293 miRNAs detected in Input and all the Argonaute RIP samples (Figure 2A; Figure S6A in Supporting Information). 25 miRNAs were shared by ALG-1 and ALG-2. 15 miRNAs were only associated with ALG-1 (ALG-1 specific), and not with any other Argonautes; 3 miRNAs were ALG-2 specific (Figure 2A). The features that ALG-1 bound to more miRNAs and more specific miRNAs than ALG-2 might partially explain why phenotypes such as growth and reproduction defects of *alg-1*



**Figure 2** Argonaute-associated miRNAs. A, miRNAs (RPM $\geq$ 10, enrichment $\geq$ 3-fold relative to Input) bound to 10 AGOs. The total numbers are indicated inside parentheses after each Argonaute. The bar figure shows number of miRNAs specific to each Argonaute or shared by different Argonautes (indicated with a vertical line to link the dots that represent the corresponding Argonautes). Conserved miRNAs are present in both *C. elegans* and *C. briggsae*, and miRNAs only in *C. elegans* are defined as non-conserved. B, Violin plots depicting levels of miRNAs in Input, for those RDE-1 and ALG-1/-2/-4 enriched miRNAs. The number and percentage of miRNAs with RPM<10 in Input are shown to the right. *P* values are calculated by the Mann-Whitney *U* test. C, Histogram revealing RDE-1 and ALG-5 associated miRNAs enriched >500-fold in the corresponding RIP samples relative to Input. Norm., normalized. D, Pie charts to show the distribution of the first 5' nucleotide of miRNA reads from Input and 10 AGO RIP samples.

mutants were more severe than those of *alg-2* (Brown et al., 2017; Bukhari et al., 2012; Vasquez-Rifo et al., 2012). We re-analyzed the mis-regulated mRNAs in *alg-1* or *alg-2* mutants from a previous study (GSE98935, Brown et al., 2017), and 931 mRNAs for *alg-1* and 906 mRNAs for *alg-2* demonstrated significant changes in expression levels (fold change  $\geq 2$  or  $\leq 0.5$ ,  $P < 0.05$ ) (Figure S7A in Supporting Information). These mRNAs were then analyzed as ALG-1 or ALG-2 targets. 525 mRNAs were targeted by both ALG-1 and ALG-2, 406 mRNAs were uniquely ALG-1-targeted, and 381 mRNAs were uniquely ALG-2-targeted (Figure S7A in Supporting Information). Furthermore, Gene Ontology (GO) analysis revealed that the 406 unique ALG-1 targets illustrated biological processes in defense response, cuticle development, response to external biotic stimulus, and so on. The 381 unique ALG-2 targets exhibited enrichment in biological processes such as peptidyl-serine modification, phosphorus metabolic process, and protein metabolic process, but without any GO term related directly to development (Figure S7B in Supporting Information). The difference in their mRNA targets could be linked directly to their distinct miRNA profile, and should be responsible for the distinct phenotypes between *alg-1* and *alg-2* mutants.

RDE-1 was the third Argonaute in terms of the number of different miRNAs loaded, with 55 miRNAs, and 17 of them were RDE-1 specific (Figure 2A). RDE-1 is known as an Argonaute that utilizes different types of sRNAs, including miRNAs, to screen the *C. elegans* transcriptome (Corrêa et al., 2010; Parrish and Fire, 2001). Seven Argonautes had a certain number of Argonaute-specific miRNAs, and ALG-3, ALG-4, or WAGO-10 did not have any specific miRNAs (Figure 2A). Interestingly, miRNAs specific for a particular Argonaute tended to be non-conserved (36 non-conserved in 51 miRNAs), while miRNAs shared by at least two Argonautes tended to be conserved (72 conserved in 113 miRNAs) (Figure 2A). Here, conserved miRNAs were those present in both *C. elegans* and another nematode species *C. briggsae*. It is worth noting that miR-230-3p was bound by all 10 Argonautes, and we speculated that it might be expressed in most types of cells and play regulatory roles through different Argonautes.

Out of the four Argonautes that bound to the largest number of miRNAs, RDE-1 demonstrated a significant tendency to bind less abundant miRNAs (Figure 2B). RDE-1, compared with ALG-1, ALG-2, and ALG-4, might regulate low-expressed or more stringent tissue-specific miRNAs. Another interesting point was that both RDE-1 and ALG-5 had 4 miRNAs with more than 500 times of enrichment, while the other Argonautes did not have any miRNAs with this kind of high enrichment (Figure 2C).

The first 5' nucleotide of the miRNAs bound by each Argonaute tended to be U, as U>C>A>G for eight Argonautes, and U>A>C>G for the other two Argonautes ALG-5

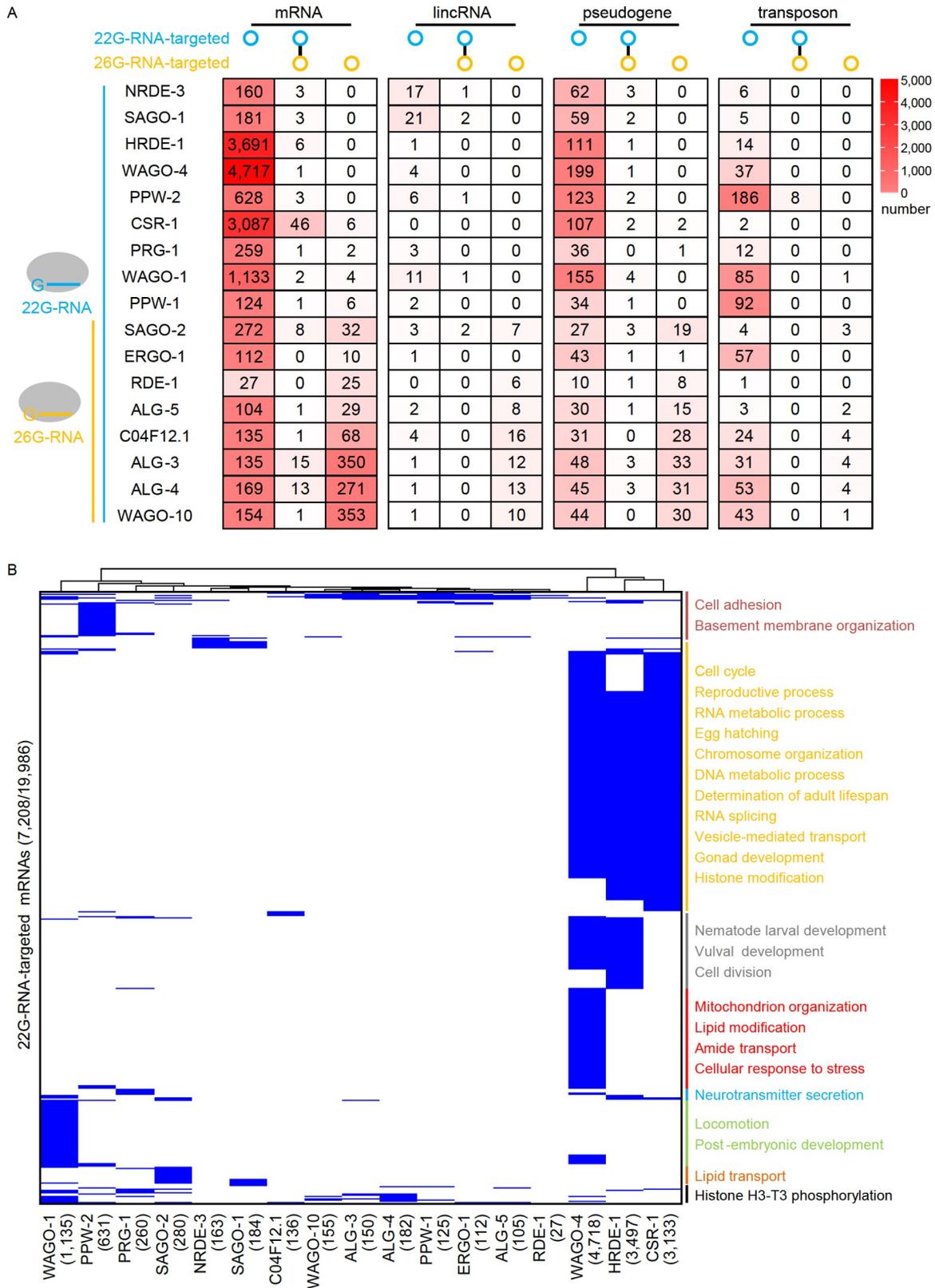
and ERGO-1 (Figure 2D). When the first two nucleotides of miRNAs bound to Argonautes were analyzed, U (C/A/G) was preferred compared with UU for each Argonaute (Figure S6B in Supporting Information). CA was the fourth represented 5' dinucleotide for essentially all the Argonautes, except for RDE-1 and ALG-5 (Figure S6B in Supporting Information). CA was the fifth for both RDE-1 and ALG-5, and CG or AG was the fourth for RDE-1 or ALG-5 (Figure S6B in Supporting Information).

PRG-1 bound to 9 miRNAs besides thousands of piRNAs, under the cutoff of  $\text{RPM} \geq 10$  and enrichment  $\geq 3$ -fold (Figures S4C and S6A in Supporting Information). However, 4 of the 9, miR-5549, miR-78, miR-8202, and miR-4936, which were not enriched by 18 other Argonautes, might actually be piRNAs. Genomic regions of these four "miRNAs" had motifs such as the Ruby motif that are signatures of piRNA coding sequences (Figure S6C in Supporting Information) (Batista et al., 2008; Ruby et al., 2006).

### 22G-RNAs and 26G-RNAs as endogenous siRNAs bound to Argonautes

Two major types of endogenous siRNAs are generated in *C. elegans*, which are 22G-RNAs and 26G-RNAs of ~22 and ~26 nt in length, respectively, with the first base usually as G (Ambros et al., 2003; Batista et al., 2008; Ruby et al., 2006; Vasale et al., 2010). As siRNAs, both 22G-RNAs and 26G-RNAs are produced by steps initiated from RNA-dependent RNA polymerase (RdRP) activity with the other transcripts as templates (Batista et al., 2008; Conine et al., 2010; Gent et al., 2010; Maniar and Fire, 2011). In some sense, for these siRNAs, the long template transcripts can be both recognized as origins and regarded as targets. Bioinformatics analyses revealed that except for ALG-1 and ALG-2, the other 17 functional Argonautes could bind a portion of 22G-RNAs, and 8 of them also bound to some 26G-RNAs (Figure 3A; Figure S8 in Supporting Information). Here, the cutoff for 22G-RNAs was sRNAs ranging from 21 to 23 nt, with a 5' guanosine,  $\text{RPM} \geq 5$ , and enrichment  $\geq 2$ -fold relative to Input. For 26G-RNAs, the cutoff was 26 nt sRNAs with a 5' guanosine,  $\text{RPM} \geq 5$ , and enrichment  $\geq 2$ -fold relative to Input.

Sequence features of 22G-RNAs associated with each Argonaute were then examined, and it appeared that CSR-1 and NRDE-3 had some bias to GN for the last two nucleotides at the 3' end (Figure S8A in Supporting Information). No other distinct sequence feature was observed for 22G-RNAs bound to the other Argonautes. There was also no obvious Argonaute specific sequence feature for 26G-RNAs (Figure S8B in Supporting Information). There is a class of 22G-RNAs called risiRNAs, which are complementary to 18S and 26S rRNA (Zhou et al., 2017). SAGO-1 and SAGO-2 were found to be the Argonautes that bound to the highest portion of risiRNAs (Figure S8C in Supporting Information).



**Figure 3** Targets of endo-siRNAs associated with *C. elegans* Argonautes. A, 22G-RNAs and 26G-RNAs bound to each Argonaute are divided into 4 groups according to their targets. Numbers of the genes that encode the targeted transcripts (mRNAs, lincRNAs, transposons, and pseudogenes) are listed in the table. B, Hierarchical clustering diagram and GO analysis of mRNAs targeted by AGOs associated 22G-RNAs. 7,208 is the total number of targeted mRNAs, and 19,986 is the number of annotated mRNAs according to WormBase annotation (WS285). Each blue line represents a single mRNA targeted by the AGO-bound 22G-RNAs. GO analysis is performed with the Gorilla web-server.

We then counted the numbers of targets, including mRNAs, lincRNAs, transcripts of pseudogenes, and transposons, of 22G-RNAs and 26G-RNAs corresponding to each Argonaute (Figure 3A). For mRNAs, four Argonautes, HRDE-1, WAGO-4, CSR-1, and WAGO-1, could each target more than 1,000 transcripts with the associated 22G-RNAs (Figure 3A). Among the eight 26G-RNA binding Argonautes, three of them, ALG-3, ALG-4 and WAGO-10 could each regulate more than 100 mRNAs through 26G-RNAs (Figure 3A). NRDE-3 and SAGO-1 targeted the most numbers of lincRNAs with 22G-RNAs, and ALG-3, ALG-4, C04F12.1, and WAGO-10 targeted the most lincRNAs with 26G-RNAs (Figure 3A). For pseudogene transcripts, each of the five Argonautes, HRDE-1, WAGO-4, PPW-2, CSR-1, and WAGO-1, targeted more than 100 of them (Figure 3A; Figure S9A in Supporting Information). For transposons, PPW-2 was ranked first by targeting 186 of them with 22G-RNAs, consistent with the previous observation that PPW-2 is a regulator in transposon silencing (Vastenhouw et al., 2003), and the number of transposons targeted by PPW-2 was more than 2-fold of the transposons targeted by the secondly ranked Argonaute PPW-1 (Figure 3A; Figure S9B in Supporting Information).

GO analysis and clustering of mRNA targets for each Argonaute based on the associated 22G-RNAs revealed that multiple Argonautes could regulate the same GO, while some Argonautes could regulate unique GO (Figure 3B). For example, cell adhesion genes were regulated by most of the Argonautes, while mitochondrion organization genes were targeted by WAGO-4 only (Figure 3B). Interestingly, the four Argonautes with the most mRNA targets fell into two separate clusters, with targets of WAGO-1 enriched in GO terms including locomotion and post-embryonic development, and targets of HRDE-1, WAGO-4, and CSR-1 enriched in GO terms such as RNA metabolic process and DNA metabolic process (Figure 3B). The three Argonautes ALG-3, ALG-4, and WAGO-10 targeting the most mRNAs with 26G-RNAs, could form a cluster with target genes enriched in GO terms such as peptidyl-Tyr dephosphorylation, peptidyl-Ser phosphorylation, and male gamete generation (Figure S10 in Supporting Information).

### Regulation of lincRNAs by siRNAs and Argonautes

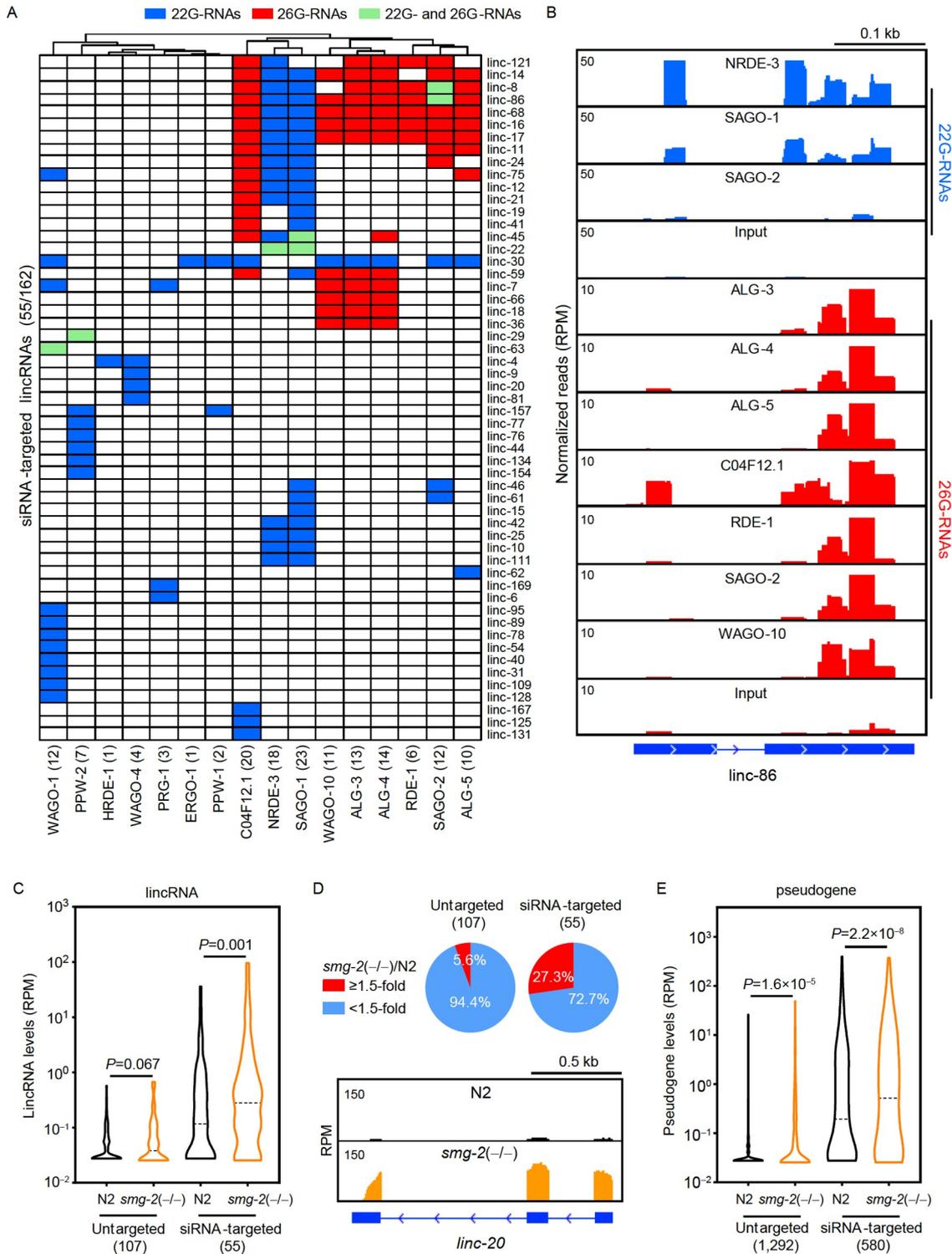
There are 170 lincRNAs annotated in the *C. elegans* genome, and multiple of them have been shown to play critical roles in diversified physiological events (Wei et al., 2019). In our RNA-seq data, antisense siRNAs corresponding to 117 lincRNAs were identified (Figure S11A in Supporting Information). A total of 55 lincRNAs were regulated by siRNAs bound to 16 Argonautes (Figure 4A). The 55 lincRNAs displayed less stage-specific expression than those not regulated by siRNAs (Figure S11B and C in Supporting

Information). NRDE-3, SAGO-1, C04F12.1, WAGO-1, and WAGO-10 targeted the most lincRNAs (Figure 4A). To give one example, among all lincRNAs, *linc-86* was targeted by 9, the highest number of Argonautes, through either 22G-RNAs or 26G-RNAs, and 22G-RNAs and 26G-RNAs were generated at overlapped loci (Figure 4B).

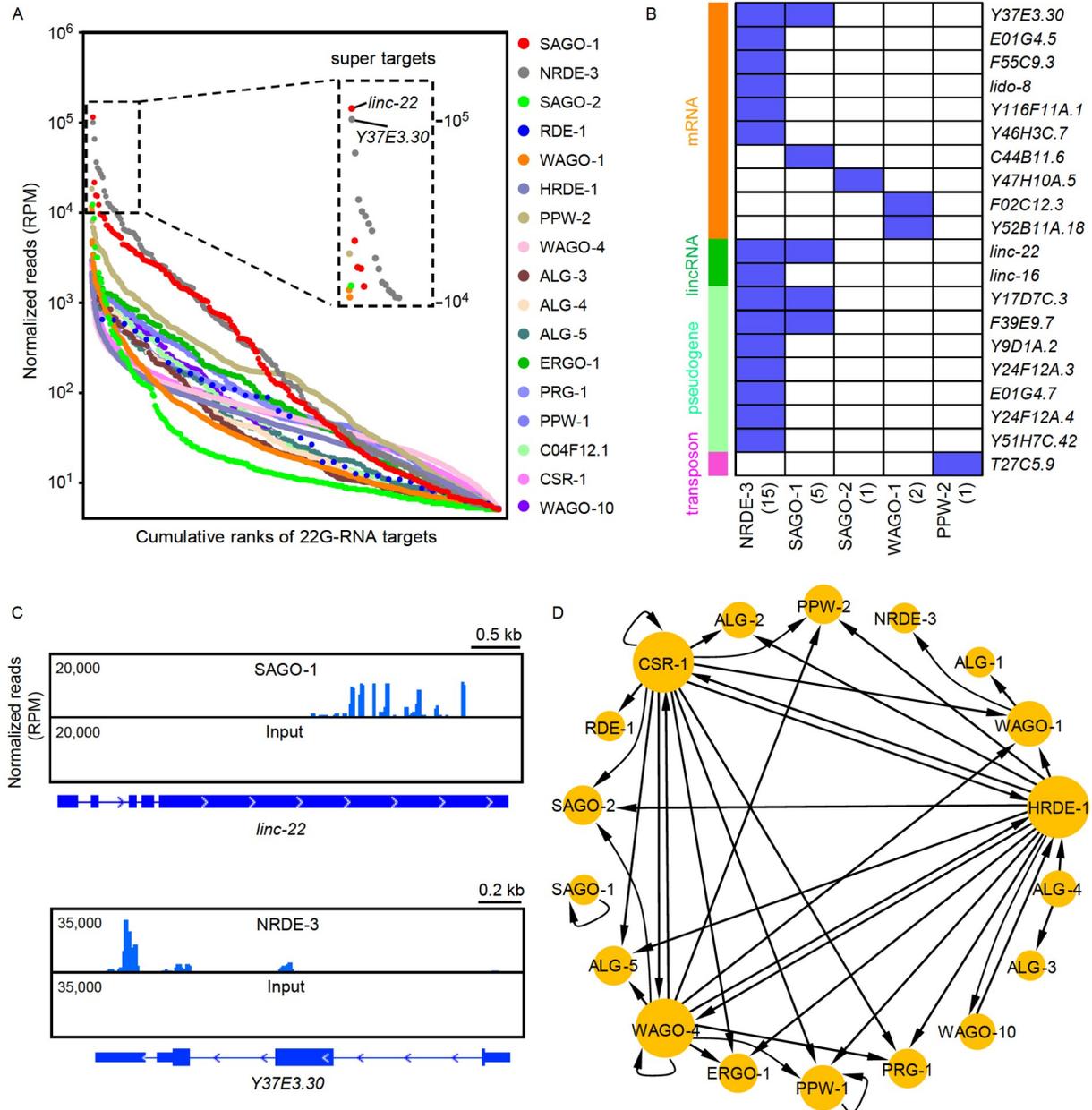
In eukaryotes, the nonsense-mediated mRNA decay (NMD) pathway is an important degradation pathway for mRNAs with premature stop codons (Arciga-Reyes et al., 2006; Scheer et al., 2016). In *C. elegans*, NMD genes such as *smg-2* are known to play roles in maintaining the silencing effect of RNAi (Domeier et al., 2000; Gu et al., 2009). LincRNAs and transcripts of pseudogenes as long non-coding RNAs (lncRNAs) share multiple features with mRNAs carrying premature stop codons, and we wondered whether NMD and RNAi could regulate these lncRNAs together. It was found that the expression levels of lincRNAs targeted by siRNAs loaded on Argonautes were significantly increased in *smg-2* mutant compared with wildtype N2, while those not targeted by siRNAs demonstrated a slight, statistically insignificant increase (Figure 4C). 27.3% lincRNAs targeted by siRNAs, while only 5.6% lincRNAs not targeted by siRNAs, had over 1.5-fold increased levels in *smg-2* mutant (Figure 4D). The expression of *linc-20* in the *smg-2* mutant, compared with the wildtype, showed the most striking increase (Figure 4D). We further examined the expression of pseudogenes in *smg-2* mutant. Levels of both siRNA-targeted and untargeted pseudogene transcripts were significantly increased in *smg-2* mutant, compared with N2 (Figure 4E). The fold changes of median levels between *smg-2* mutant and N2 were 1.09 for untargeted pseudogene transcripts and 2.35 for siRNA-targeted pseudogene transcripts (Figure 4E). These results indicated that levels of lincRNAs including lincRNAs and pseudogene transcripts were managed by NMD and the RNAi system together.

### Super targets of 22G-RNAs and potential inter-regulation among Argonautes

When reads of 22G-RNAs corresponding to each target of a particular Argonaute were plotted, it was noticed that 20 transcripts as targets of 5 Argonautes, SAGO-1, NRDE-3, SAGO-2, WAGO-1, and PPW-2, were enriched with an extraordinarily high amount of 22G-RNAs ( $\geq 10^4$  RPM) specific to each transcript (Figure 5A). Thus, these 20 transcripts could be viewed as super targets of the corresponding Argonaute in some way. Four of them were super targets for two Argonautes (Figure 5B). Two of them, *linc-22* as SAGO-1 target and *Y37E3.30* mRNA as NRDE-3 target had more than  $10^5$  RPM 22G-RNAs enriched (Figure 5A and C). Biological significance of the existence of these super targets remains for further evaluation.



**Figure 4** LncRNAs targeted by Argonautes associated endo-siRNAs. **A**, Hierarchical clustering diagram of lincRNAs targeted by AGO-associated endo-siRNAs. In total, 55 lincRNAs are targeted by AGO-associated endo-siRNAs, out of the 162 lincRNAs with corresponding antisense endo-siRNAs detected in Input and all the AGO RIP samples. Each blue or red box represents a single lincRNA targeted by the AGO-bound 22G-RNAs or 26G-RNAs. Green box represents a single lincRNA targeted by both AGO-bound 22G-RNAs and 26G-RNAs. **B**, Peak distribution of *linc-86* derived complementary 22G-RNAs (blue) and 26G-RNAs (red) from the corresponding AGO RIP. **C**, Violin plots revealing expression levels of siRNA-targeted and untargeted lincRNAs in wildtype N2 or *smg-2* mutant strain from previous data (GSE94077). *P* values are calculated by the Wilcoxon matched-pairs signed rank test. **D**, LincRNA profiling of *smg-2* mutant strain compared with wildtype N2 worms. Pie charts showing the portion of lincRNAs with Fold Change  $\geq 1.5$  (*smg-2/N2*,  $P < 0.05$ ) and Fold Change  $< 1.5$ . RNA-seq reads (GSE94077) of *linc-20* in N2 and *smg-2(-/-)* are shown. *P* values are calculated by the two-tailed Student's *t*-test. IGV, Integrative genomics viewer. **E**, Violin plots revealing relative levels of siRNA-targeted and untargeted pseudogene transcripts in wildtype N2 or *smg-2* mutant strain. *P* values are calculated by the Wilcoxon matched-pairs signed rank test.



**Figure 5** Super targets and regulatory network of Argonautes. **A**, Scatter plots of AGO-bound 22G-RNAs. The y-axis represents levels (in RPM) of AGO target-derived antisense 22G-RNAs, and the x-axis represents cumulative ranks of 22G-RNA targets. Super targets for the individual AGO are defined with the cutoff ( $\text{RPM} \geq 10^4$  in the corresponding RIP sample, enrichment  $\geq 2$ -fold relative to Input). **B**, Heatmap of siRNAs bound by five AGOs corresponding to super targets. **C**, Peak distribution of *linc-22* and *Y37E3.30* derived complementary 22G-RNAs from AGO RIPs. *linc-22* targeted by SAGO-1 and *Y37E3.30* targeted by NRDE-3 are the two most striking super targets with  $\text{RPM} > 10^5$ . **D**, Inter-regulation map demonstrating regulation among *C. elegans* Argonautes. Node size indicates the degree of interaction.

Notably, a regulatory network was formed among Argonautes themselves, through unidirectional-, bidirectional-, and self-regulation by Argonaute proteins and the associated endo-siRNAs corresponding to the Argonaute mRNAs (Figure 5D). One Argonaute, C04F12.1, did not participate in this network, and another Argonaute, SAGO-1, displayed only self-regulation. WAGO-4, CSR-1, and HRDE-1 formed the most links with the other Argonautes (Figure 5D). This network is part of the complex Argonaute-sRNA regulatory system of *C. elegans*.

### Uridylated 22G-RNAs and their associated Argonautes play roles in transgenerational inheritance

Uridylation at the 3' end of sRNAs is generally regarded as a signal for sRNA degradation in *C. elegans* (Heo et al., 2009; Lehrbach et al., 2009; Li et al., 2005). Our sequencing results revealed that four Argonautes, CSR-1, HRDE-1, WAGO-4, and PPW-2, bound to a portion of 22G-RNAs with one to three untemplated uridines (Figure 6A and B). Further analyses demonstrated that 22G-RNAs were the main class of

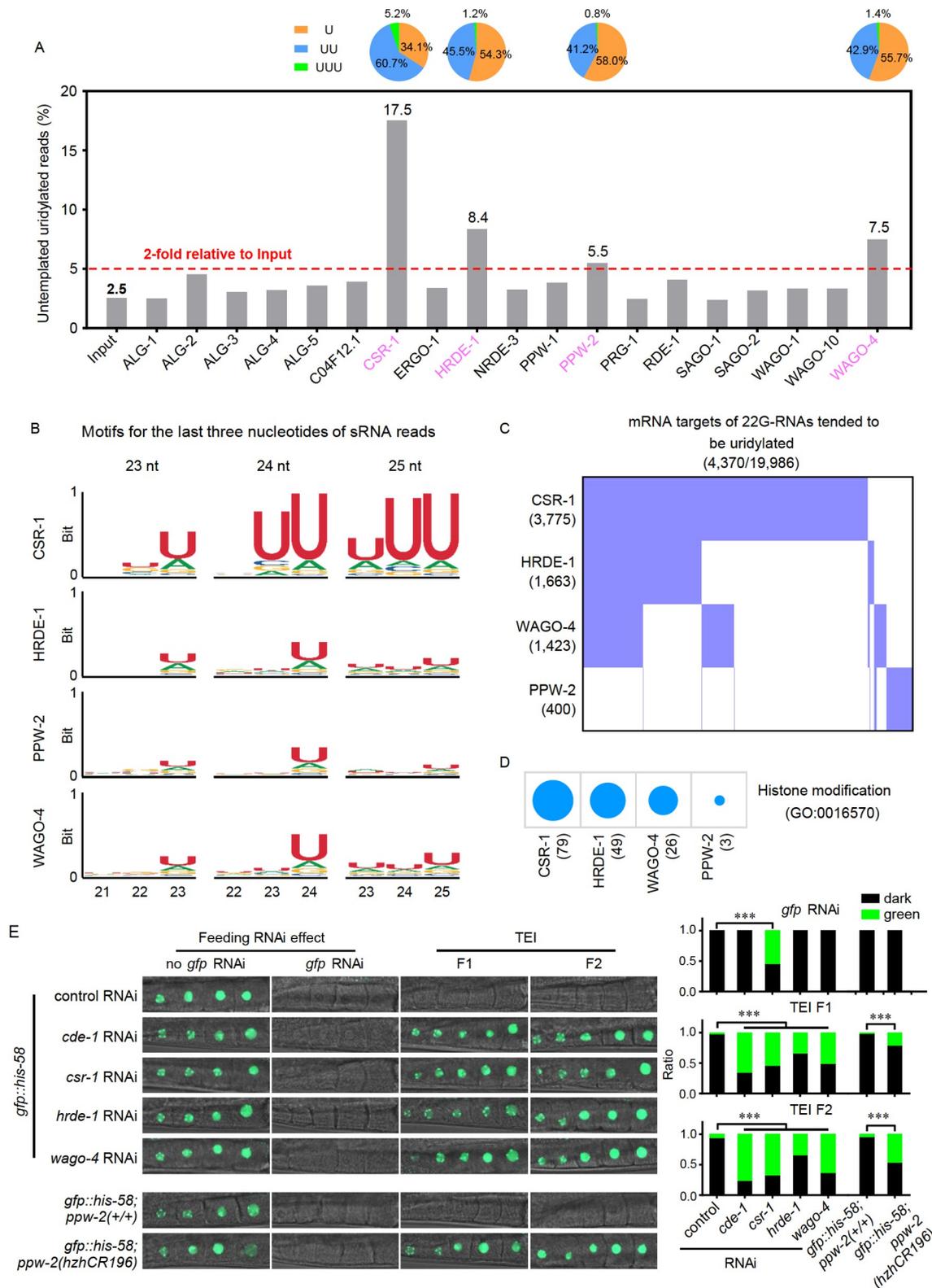
sRNAs with uridylation (Figure 6B; Figure S12A in Supporting Information). In detail, CSR-1 bound to 17.5% in terms of total reads and 70.7% in terms of distinct RNA molecules of uridylated 22G-RNAs; for HRDE-1, the portion was 8.4% and 70.7%, respectively; for WAGO-4, the portion was 7.5% and 47.2%, respectively; and for PPW-2, it was 5.5% and 55.6%, respectively (Figure 6A; Figure S12B in Supporting Information). For the 3,133 mRNAs corresponding to CSR-1 associated 22G-RNAs, CSR-1 targeted 2,642 (84.33%) mRNAs with uridylated 22G-RNAs (Figure S12C in Supporting Information). GO analysis revealed that the 2,642 mRNAs were enriched in biological processes such as mitotic cell cycle process, mitotic sister chromatid segregation, oogenesis, ubiquitin-dependent ERAD pathway, and spermatogenesis. While for the 491 mRNAs targeted by CSR-1 with non-uridylated 22G-RNAs, the enriched biological processes were ATP synthesis coupled proton transport, cellular oxidant detoxification, response to hydrogen peroxide, regulation of protein kinase activity, cellular response to unfolded protein, and axis specification (Figure S12D in Supporting Information). CSR-1 served as a protector of germline gene expression (Claycomb et al., 2009; Gu et al., 2009; Wedeles et al., 2013), and it seemed that targets of CSR-1 with uridylated 22G-RNAs were enriched for germline specific functions.

Previous studies have shown that CSR-1 and WAGO-4 could bind to uridylated 22G-RNAs (Lev et al., 2019; van Wolfswinkel et al., 2009; Xu et al., 2018). HRDE-1 and WAGO-4 are also known to play critical roles in TEI induced by feeding RNAi (Buckley et al., 2012; Hourri-Zeevi et al., 2020; Rechavi et al., 2014; Xu et al., 2018). However, tight links between TEI and Argonautes with the associated uridylated 22G-RNAs are still required. mRNAs targeted by the 22G-RNAs with uridylated isoforms associated to each of the four Argonautes varied from 400 to more than 3,700 (Figure 6C); interestingly, GO analysis of these four groups of mRNA targets revealed only one shared GO cluster, Histone modification (GO:0016570) (Figure 6D). TEI is well known to be associated with histone modifications (Gu et al., 2012; Klosin et al., 2017; Mao et al., 2015; Schwartz-Orbach et al., 2020). By using a germline specific reporter *mex-5p::gfp::his-58 (ustIs45)* (Xu et al., 2018), we observed that feeding RNAi against *gfp* led to GFP silencing, and this silencing persisted in F1 and F2 as the effect of TEI (Figure 6E) (Alcazar et al., 2008; Buckley et al., 2012). CDE-1 is known to regulate TEI as one of the three *C. elegans* poly-uridylation polymerases (PUPs), and *cde-1* RNAi could impede the TEI effect in a previous study and also in our assay (Figure 6E; Figure S12E in Supporting Information) (van Wolfswinkel et al., 2009). RNAi against *csr-1*, *hrde-1*, and *wago-4*, individually, also hindered the TEI effect (Figure 6E; Figure S12E in Supporting Information). RNAi of *csr-1* but not *hrde-1* or *wago-4* also had an inhibitory

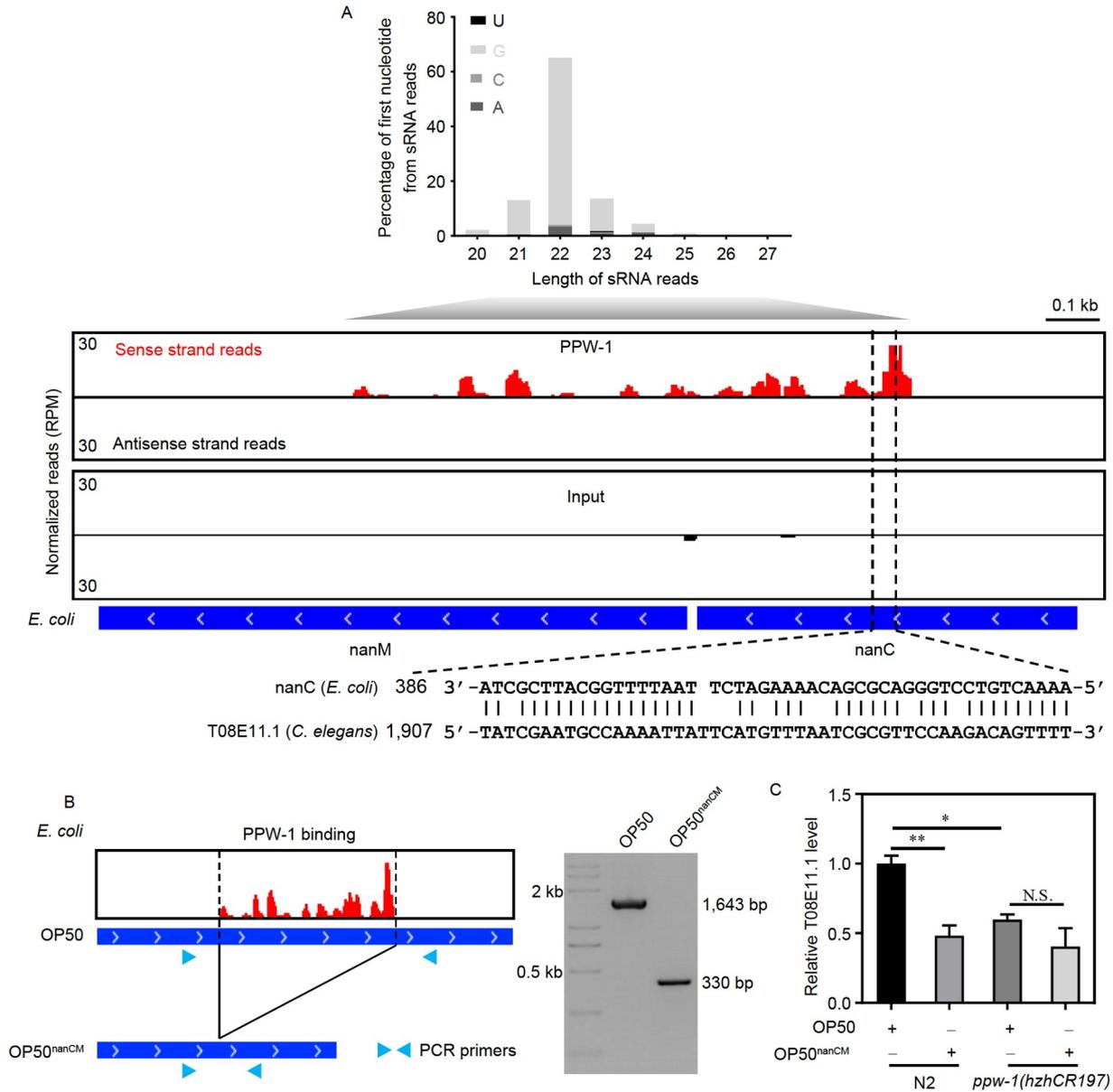
effect on the feeding RNAi itself (Figure 6E); *csr-1* is also involved in the effectiveness of feeding RNAi (Yigit et al., 2006). To assess the link between TEI and PPW-2, the newly identified Argonaute for uridylated 22G-RNAs, a *ppw-2* deletion mutant was generated. We found that *ppw-2* null worms were grossly normal, and the effect of feeding RNAi was not affected, while the TEI effect was significantly mitigated (Figure 6E). These results strongly indicated that all four Argonautes with binding ability to uridylated 22G-RNAs had regulatory roles in TEI.

### PPW-1 binds to sRNAs originating from *E. coli*

Increasing lines of evidence have demonstrated interspecies regulation of *C. elegans* gene expression by sRNAs originating from bacteria (Kaletsky et al., 2020; Liu et al., 2012; Samuel et al., 2016). For example, two *E. coli* sRNAs, OxyS and DsrA, which were induced under stress conditions, could generate sRNAs in *C. elegans* to regulate gene expression of this animal (Liu et al., 2012). Recently, P11, an sRNA of *Pseudomonas aeruginosa*, was found to elicit regulatory roles in *C. elegans* (Kaletsky et al., 2020; Moore et al., 2019). To assess potential Argonautes that mediated interspecies effects by binding to sRNAs with exogenous origins, we analyzed sRNAs bound to each Argonaute for those derived from *E. coli* strain OP50, which was the food source for *C. elegans* cultured in the laboratory. Under the cutoff of 2-fold enrichment ( $P < 0.05$ ), PPW-1 was the only Argonaute associated with 20–30 nt sRNAs of *E. coli* origin (Figure 7A; Figure S13A and B in Supporting Information). All these sRNAs were derived from the *E. coli* nanCM operon, *hdeA/B/D* mRNAs, and *lacZ/I* mRNAs (Figure 7A; Figure S13A and B in Supporting Information). We also noticed that these sRNAs mainly belonged to 22G-RNAs (Figure 7A; Figure S13A and B in Supporting Information). By BLAST alignment, one mRNA, *T08E11.1*, in *C. elegans* was found to harbor complementary sequences with sRNAs derived from the nanCM operon (nanCM-sRNAs) (Figure 7A). Actually, no other *E. coli* derived sRNAs shared continuous sequence similarity over 10 nt to any *C. elegans* transcript. We then generated a knockout OP50 strain of the nanCM operon (OP50<sup>nanCM</sup>), which was verified by genomic PCR and Sanger sequencing (Figure 7B). The control OP50 and OP50<sup>nanCM</sup> were respectively fed to wildtype N2 worms, and the levels of *T08E11.1* mRNA were examined. It was found that levels of *T08E11.1* mRNA decreased significantly in the group with OP50<sup>nanCM</sup> food (Figure 7C). When the OP50 or OP50<sup>nanCM</sup> was fed to *ppw-1(hzhCR197)* mutant worms, the levels of *T08E11.1* mRNA did not show a significant change between the two groups with OP50 and OP50<sup>nanCM</sup> food (Figure 7C). Levels of *T08E11.1* mRNA were also significantly lower in *ppw-1* null (*hzhCR197*) worms than those in N2, when fed with OP50 (Figure 7C). All these bioin-



**Figure 6** Uridylated 22G-RNAs associated with Argonautes. **A**, Percentages of reads for 22G-RNAs with untemplated uridine in Input and AGO RIP samples. Pie charts (top) indicate percentage of untemplated uridylylated reads with 3' end additions (U, UU, or UUU) in the corresponding AGOs. **B**, Motifs for the last three nucleotides of 23 nt-, 24 nt-, and 25 nt-sRNA reads in CSR-1, HRDE-1, PPW-2, and WAGO-4 RIPs. **C**, Heatmap of mRNAs targeted by 22G-RNAs with uridylylated isoforms (RPM $\geq$ 5, enrichment $\geq$ 2-fold relative to Input) associated with the four Argonautes. **D**, Histone modification (GO:0016570), the only overlapped biological process from GO analysis of the mRNA targets shown in **C**. **E**, Representative images of germline in *C. elegans*. *E. coli* expressing *gfp* dsRNA or other dsRNA corresponding to the specified gene were fed to *mex-5p::gfp::his-58 (ustIs45)* worms. Percentage of P0 (RNAi effect), F1 and F2 animals (TEI effect) expressing GFP is counted ( $n > 80$  worms). *P* values were calculated by the Chi-square test; \*\*\*,  $P < 0.001$ .



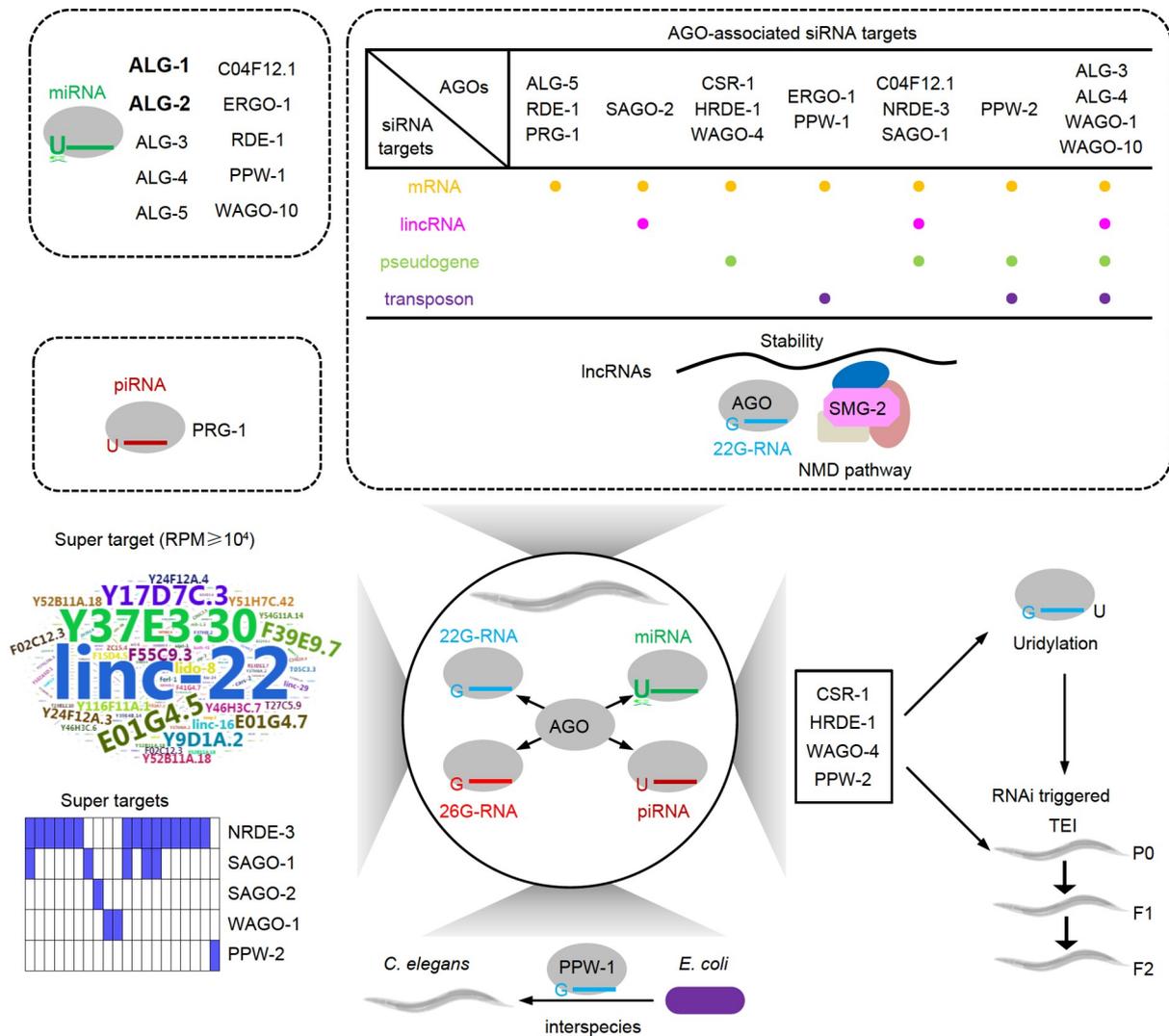
**Figure 7** PPW-1 binds to 22G-RNAs with *E. coli* origin and plays an interspecies regulatory role. **A**, 5' first nucleotide and size distribution of PPW-1 bound sRNA reads with *E. coli* origin (top). Peaks of sRNA bound to PPW-1 aligned to the *E. coli* genomic loci (middle). Red and black peaks represent sense and antisense strand reads mapping to *E. coli* genome (NC\_000913.3), respectively. Note, the PPW-1 associated sRNAs are actually antisense to *nanCM* RNA. Alignment of the *E. coli nanC* RNA sequence and *C. elegans T08E11.1* mRNA (bottom). **B**, Strategy of knockout the region in *E. coli* genome that corresponding to PPW-1 associated sRNAs (left) and gel image of PCR products from genotyping of the *E. coli* knockout strain OP50<sup>nanCM</sup> (right). bp, base pair. **C**, Relative *T08E11.1* mRNA level in wildtype N2 or *ppw-1(hzhCR197)* null worms with OP50 or OP50<sup>nanCM</sup> food examined by RT-qPCR. Error bars indicate SEM from three independent experiments. *P* values from two-tailed Student's *t*-test. N.S., not significant; \*, *P*<0.05; \*\*, *P*<0.01.

formatics and experimental results strongly indicated that *nanCM* RNA of *E. coli* might lead to the production of 22G-RNAs bound to PPW-1 in *C. elegans* to increase the expression of *T08E11.1*; both PPW-1 and *nanCM* sequences were required for this interspecies effect.

## DISCUSSION

We have performed experiments to identify sRNAs bound to

each of the 19 endogenously expressed functional Argonautes. The RNA-seq data obtained have allowed a systematic evaluation about the partnership between Argonautes and the sRNAs, including miRNAs, 21U-RNAs, 22G-RNAs, 26G-RNAs, and uridylylated 22G-RNAs (Figure 8). With further bioinformatics analyses and experimental evaluations, some regulatory functions and networks of these Argonaute-sRNA complexes have been uncovered (Figure 8). The *C. elegans* strains and sRNA profiles generated in this study are also rich resources.



**Figure 8** A simplified summary of *C. elegans* Argonautes, the associated sRNAs, and the corresponding functions.

For each major type of sRNAs, there are generally several functional Argonautes, except that 21U-RNAs are associated almost exclusively with PRG-1. There are also other types of sRNAs such as tRNA fragments (tRFs and tiRNAs), snoRNA-derived fragments (sdRNAs), and transcription start site-associated RNAs (TSSaRNAs) that are known to be critical in animal cells (Gu et al., 2022; Liu et al., 2022; Saw et al., 2021; Taft et al., 2010; Wajahat et al., 2021; Xue et al., 2020). Background studies about these sRNAs are currently lacking in *C. elegans*, and thus we have not performed analyses about them. For RNAs smaller than 20 nt, we have ignored them in in-depth bioinformatics analysis; thus, it is possible that some of these RNA molecules could be loaded to Argonautes to elicit functions.

Twenty super targets have been identified for 22G-RNAs loaded in 5 Argonautes, and these include 10 mRNAs, 2 lincRNAs, 7 pseudogenes, and 1 transposon. Among these super targets, mRNAs were under represented, while

lincRNAs and pseudogenes were overrepresented. Currently, we do not know why they are targeted by extraordinarily high levels of 22G-RNAs associated with specific Argonautes. None of the 10 mRNAs have known functions or associated phenotypes yet, which may be an indication that they may not be key protein coding genes, as thousands of coding genes have already been linked to critical biological events in *C. elegans*. It seems that these mRNAs as super targets may be managed by the endogenous RNAi system more like some lincRNAs and pseudogene transcripts. Besides, the NMD pathway and the RNAi system may synergize to regulate the levels of lincRNAs and pseudogene transcripts in *C. elegans*.

We have identified PPW-2 as a new Argonaute for uridylated 22G-RNAs. All four Argonautes for uridylated 22G-RNAs have positive roles in the TEI of feeding RNAi, and in this sense, uridylation of 22G-RNAs may be required in the establishment or maintenance of TEI. Uridylation is

generally considered as a modification for degradation (Heo et al., 2009; Lehrbach et al., 2009; Li et al., 2005; Scheer et al., 2016). These Argonautes may participate in TEI by binding to uridylated 22G-RNAs to facilitate the degradation of these sRNAs, although it is also possible that these Argonaute-uridylated sRNAs may actually be functional complexes in modulating elements involved in TEI.

PPW-1 is special in this study as the only Argonaute that binds to 22G-RNAs generated from exogenous *E. coli* RNAs. It seems that PPW-1 utilizes these sRNAs to increase the levels of *T08E11.1* mRNA, and the physiological function of this regulation is unknown. It remains for further investigation as whether this positive effect is achieved by the direct targeting of PPW-1-22G-RNA complex on *T08E11.1* mRNA or not. Previously, it was reported that two *E. coli* RNAs, OxyS and DsrA, can negatively regulate the expression of two *C. elegans* target genes, *che-2*, and *F42G9.6*, respectively (Liu et al., 2012). In both cases, tiny RNAs of 17–18 nt derived from OxyS and DsrA appear to function through the RNAi system. In this study, sRNAs with potential exogenous origin but smaller than 20 nt were not analyzed, partially due to the concerns about non-specific degradation of bacterial RNAs.

In this study, we have identified a snapshot of the regulatory network and provided resources for further investigations. The network of Argonautes and the associated sRNAs in *C. elegans* is extraordinarily complex. The regulatory network is also dynamic throughout the life cycle of *C. elegans*, and the expression of Argonautes and sRNAs is subjected to active and fine regulation. Future studies are still needed to portray the full image of the *C. elegans* Argonaute-sRNA system.

## CONCLUSIONS

We have inserted an *in situ* fusion tag for each of the 19 functional Argonautes in *C. elegans*, and performed experiments to isolate sRNAs bound to each Argonaute. Through bioinformatics and experimental characterizations, the features and functions of sRNAs loaded to these Argonautes are described. Novel insights about the roles of these ribonucleoprotein complexes in gene expression, transgenerational inheritance, and interspecies regulation are provided with supporting data. The worm Argonaute strains and the sRNA profiles corresponding to each Argonaute also provide support for future in-depth studies.

## MATERIALS AND METHODS

### *C. elegans* strains

All *C. elegans* strains were maintained on Nematode Growth

Medium (NGM) plates seeded with OP50 at 20°C unless otherwise stated (Brenner, 1974). N2 Bristol was obtained from the Caenorhabditis Genetics Center (CGC) and used as the wildtype strain. All worm strains generated or used in this study are listed in Table S1 in Supporting Information.

### Plasmid construction

All plasmids were constructed with restriction enzyme digestion and ligation or with recombination methods. For CRISPR/Cas9, the plasmid pDD162 expressing Cas9 II protein was used as our previous description (Wei et al., 2019; Yu et al., 2017). All single guide RNA (sgRNA) sequences used in this study were designed at <http://crispor.tefor.net/>. The 20-nt sgRNA sequence was inserted into the backbone pUC57 plasmid between the *EcoRI* and *HindIII* restriction endonuclease sites behind the U6 promoter. For homology recombination (HR) plasmids, 1 kb upstream and 1 kb downstream DNA sequences flanking the site of interest, and GFP or FLAG sequences were inserted into the plasmid pPD95.67 between the *SphI* and *ApaI* double-digested sites. Oligonucleotide sequences for primers used in plasmid construction are listed in Table S2 in Supporting Information.

### CRISPR/Cas9 knockin and knockout of *C. elegans*

The CRISPR/Cas9 system was carried out as previously described with minor modifications (Wei et al., 2019; Yu et al., 2017). For knockin strains, the plasmid pDD162 expressing Cas9 II protein ( $30 \text{ ng } \mu\text{L}^{-1}$ ), sgRNA plasmid ( $30 \text{ ng } \mu\text{L}^{-1}$ ), HR plasmid ( $30 \text{ ng } \mu\text{L}^{-1}$ ) and co-injection marker *Pmyo-2::mCherry* (pCFJ90) plasmid ( $15 \text{ ng } \mu\text{L}^{-1}$ ) were mixed and injected into 40 N2 adult worms. For knockout strains, pDD162 plasmid ( $30 \text{ ng } \mu\text{L}^{-1}$ ), two sgRNA plasmids ( $30 \text{ ng } \mu\text{L}^{-1}$  of each) corresponding to individual Argonaute, as well as co-injection marker *Pmyo-2::mCherry* (pCFJ90) plasmid ( $15 \text{ ng } \mu\text{L}^{-1}$ ) were injected into 40 N2 adult worms. Then, ~200 F1 worms with mCherry were singled out and F3 generation genotyping was confirmed by PCR of genomic DNA, followed by Sanger sequencing. All newly obtained strains were crossed with the wildtype N2 worms at least 6 times before further experiments.

### RNA immunoprecipitation

RIP with endogenously expressed Argonautes was carried out as previously described with minor modifications (Brown et al., 2017; Shen et al., 2018; Xu et al., 2018). Briefly, synchronized young adult or L4 stage worms (~500,000) were washed with M9 buffer for three times. Then worms were resuspended in lysis buffer ( $20 \text{ mmol L}^{-1}$  Tris-HCl (pH 7.5),  $200 \text{ mmol L}^{-1}$  NaCl,  $2 \text{ mmol L}^{-1}$   $\text{MgCl}_2$ ,

0.5% NP40, 1 mmol L<sup>-1</sup> DTT, 1×Protease-inhibitor cocktail (Roche, Switzerland) and 50 U mL<sup>-1</sup> RNasin Inhibitor (Promega, USA)), and then were sonicated on ice for 15 min using Bioruptor Plus sonication device. The worm lysates were spun at 14,000×g for 15 min at 4°C, and the supernatant was collected for immunoprecipitation. The supernatant of N2 worm lysates was saved as Input. The worm supernatant was incubated with anti-GFP beads (ChromoTek, Germany) or Protein G beads (Invitrogen, USA) pre-conjugated anti-FLAG antibody (Sigma-Aldrich, USA) for 4 h at 4°C. After being washed twice with washing buffer (50 mmol L<sup>-1</sup> HEPES-KOH (pH 7.5), 300 mmol L<sup>-1</sup> KCl, 0.5% NP-40, 1 mmol L<sup>-1</sup> DTT, 50 U mL<sup>-1</sup> RNasin Inhibitor, 1×Protease-inhibitor), beads were digested with 2 μL TURBO DNase (Thermo Fisher Scientific, USA) for 15 min at room temperature and 0.1 U μL<sup>-1</sup> RNase I (Thermo Fisher Scientific) for 3 min at 37°C. After being washed twice with high salt washing buffer (50 mmol L<sup>-1</sup> HEPES-KOH, pH 7.5, 500 mmol L<sup>-1</sup> KCl, 0.5% NP-40, 1 mmol L<sup>-1</sup> DTT, 50 U mL<sup>-1</sup> RNasin Inhibitor, 1×Protease-inhibitor), the Argonaute-sRNA complex was further digested with proteinase K (Invitrogen) for 10 min at 56°C, and the released RNAs were isolated with TRIzol reagents according to the manufacturer's instructions.

### Library preparation of RNA-seq

For library preparation, sRNAs were dephosphorylated by RNA 5' polyphosphatase (Lucigen, USA) for 30 min at 37°C and recycled with RNA clean up kit (Zymo Research, USA). Then, the purified sRNAs were ligated to DNA adapter at the 3' ends with T4 RNA ligase 2 (NEB, USA) for 45 min at 25°C, and RNA adapter at the 5' ends with T4 RNA ligase 1 (NEB) for 60 min at 25°C, sequentially. The ligated products were purified and reverse transcribed using the SuperScript III first-strand synthesis system (Invitrogen) with RT primers containing a random 12-mer UMI sequence according to the manufacturer's instructions. The cDNA libraries were PCR-amplified by Q5 High-Fidelity DNA polymerase (NEB) for ~15 cycles and separated on a 12% PAGE gel. Gel slices containing libraries (140 to 170 bp) were excised, purified and then subjected to high-throughput sequencing generating ~10 million reads using an Illumina NovaSeq 6000 platform (Novogene, Tianjin, China) with a 150-nt run length. Sequences of all primers are listed in Table S2 in Supporting Information.

### RIP-seq data analyses

For RIP-seq data analysis, the adapters were firstly trimmed with Cutadapt (-e 0.1 -O 5 -m 15) and the random 12-mer UMI for identifying PCR duplicates was further removed to obtain clean reads. Then, trimmed reads ranging from 18-nt

to 30-nt were aligned to the *C. elegans* genome (ce11) with Bowtie (-v 1) and Bedtools was used to count reads for genomic annotated locus. Sense reads were counted for miRNAs, piRNAs, and tRNAs, and antisense reads were counted for mRNAs, lincRNAs, pseudogenes, transposons, and rRNAs. For rRNAs, both sense and antisense reads were counted. RPM was used to calculate the corresponding levels for distinct RNA categories.

### miRNAs

The enriched mature miRNAs were determined with the criterion of RPM≥10 in IP samples and enrichment≥3 relative to the Input sample. For conservation analysis, miRNAs present in *C. elegans* and *C. briggsae* are classified as conserved, and the others are defined as non-conserved.

### piRNAs

The enriched piRNAs were determined with the criterion of RPM≥10 in IP samples and enrichment≥2 relative to the Input sample. For mis-annotated "miRNAs" which might be acute piRNAs, piRNA upstream large and small motifs in sequences from the genomic locus were depicted as previously described (Ruby et al., 2006).

### Endo-siRNAs (22G-RNAs and 26G-RNAs)

For endo-siRNA targets including mRNA, lincRNA, pseudogene, and transposon, 22G-RNAs (sRNAs ranging from 21 to 23 nt, with a 5' guanosine) and 26G-RNAs (26 nt sRNAs with a 5' guanosine) were used for further analyses. The enriched 22G-RNA-associated targets were determined with the cutoff (RPM of complementary 22G-RNAs≥5, and enrichment≥2 relative to Input). The enriched 26G-RNA-associated targets were determined with the cutoff (RPM of complementary 26G-RNAs≥5, and enrichment≥2 relative to Input).

### Super targets

For individual Argonaute, targets with the cutoff (RPM of complementary 22G-RNAs≥10<sup>4</sup> and enrichment≥2 relative to Input) were defined as super targets.

### *E. coli*-derived sRNAs

For sRNAs originating from *E. coli*, reads mapped to the *C. elegans* genome (ce11) were filtered out and the remaining reads (20–30 nt) were aligned to the *E. coli* genome (NC\_000913.3) with Bowtie allowing two mismatches. Peaks were called using cisGenome, by comparing Argonaute IP to Input with the cutoff (enrichment≥2 and *P*<0.05). The first nucleotide bias and size distribution were further analyzed from all sRNA reads forming peaks in the *E. coli* genome. The sequences of peak region in the *E. coli* genome are aligned to *C. elegans* transcripts using Blast with at least 10-nt continuous similarity to identify *C. elegans* homo-

logous or complementary targets.

#### Visualization of peaks

The distribution of peaks was generated with integrative genomics viewer (IGV).

#### GO analysis

GO analysis of mRNAs was performed using Gorilla web-server with default parameters (Eden et al., 2009).

#### Phylogenetic tree construction

The Argonaute protein sequences were downloaded from WormBase (WS285) and subjected to generate a phylogenetic tree using Mega X with the Neighbor-Joining (NJ) method (Kumar et al., 2018). For Argonautes with multiple isoforms, the amino acid sequences of the longest variants were selected for phylogenetic tree construction. Numbers above each node indicate the NJ bootstrap support values.

#### Western blotting

Proteins from whole worm lysates or IP samples were separated on the SDS-PAGE gels and then transferred to Nitrocellulose membranes (PALL Co., USA). Membranes were treated according to the ECL Western blotting protocol (GE Healthcare, USA). The following antibodies were used in Western blots: anti-FLAG (Sigma-Aldrich) and anti- $\beta$ -actin (TransGen, Beijing, China). Antibody validation is provided on the manufacturers' websites.

#### Lifespan

Lifespan assays were carried out as previously described with minor modifications (Liu et al., 2012). In brief, synchronized eggs of the distinct strains and N2 worms were plated onto NGM plates and grown to L4 stage at 20°C. Then, for each strain, L4 worms were transferred to an NGM plate containing 100 g mL<sup>-1</sup> ampicillin and 50 mmol L<sup>-1</sup> 5'-fluoro-2'-deoxyuridine (FUDR) during lifespan measurements. Three technical replicates with 20 worms were used for each strain. The hatching day was defined as day 1, and worms were scored as alive or dead every 2 days. The mean value of lifespan was computed and analyzed using the log-rank test.

#### Brood size

All KI strains and N2 worms were well-fed for at least 3 generations before brood size assay. After bleaching, eggs were allowed to hatch and grow until L4 stage at 20°C. Then 10 worms of different strains were placed individually on

NGM plates and allowed to lay eggs overnight. Every 24 h, adult worms were transferred to a new NGM plate until egg hatching stopped. The newly hatched F1 worms were also counted daily. At last, the mean value of brood size was computed and analyzed using unpaired Student's *t*-test.

#### PCR reactions

Total RNAs were extracted from worms with TRIzol LS reagent (Invitrogen) according to the manufacturer's protocol after three cycles of freezing at -80°C and thawing at room temperature. For PCR-mediated genotyping, the template of genomic DNA (gDNA) was isolated with phenol/chloroform extraction. 500 ng RNA was reverse transcribed into cDNA with HiScript® III RT SuperMix (Vazyme, Nanjing, China) according to the manufacturer's instruction. Real-time quantitative PCR (RT-qPCR) was performed with Genious 2×SYBR Green Fast qPCR Mix (ABclonal, Wuhan, China) on a QuantStudio 3 real-time PCR system (Thermo Fisher Scientific) according to standard procedures.  $\beta$ -actin mRNA was used as an internal control. All PCR products were confirmed by Sanger sequencing. All primers used in this study are listed in Table S2 in Supporting Information.

#### Feeding RNAi

Feeding RNAi assays were conducted as described previously (Liu et al., 2020; Timmons et al., 2001). Bacterial clones expressing dsRNAs against *cde-1*, *csr-1*, *hrde-1*, and *wago-4* were obtained from the Ahringer RNAi library and further validated by Sanger sequencing. HT115 bacteria expressing the empty vector L4440 was used as a negative control.

#### Assay of RNAi-triggered TEI

Synchronized embryos of the indicated treatment or genotypes were exposed to bacteria expressing *gfp* dsRNA. F1 and F2 generations were grown on NGM plates seeded with HT115 control bacteria.

#### Confocal imaging of *C. elegans*

Imaging was carried out as previously described with some modifications (Wei et al., 2019; Yu et al., 2017). *C. elegans* images were taken on a Zeiss confocal microscope with 63×1.40 NA oil-immersion objective with a resolution of 1,024×1,024 (Zeiss LSM 980, Carl Zeiss, Germany). The images used to count GFP signal ratio were captured by Axio Imager A2 (Carl Zeiss) with 1.0 s exposure at the young adult stage, followed by processing in the open-source image processing software, Image J (v1.53q). In Figure 6E, worms without fluorescence signals observed in the germline were

defined as “dark” worms, and worms with fluorescence signals in the germline were defined as “green”.

### *E. coli* strains

*Escherichia coli* OP50, used for the general culture of worms. OP50 with knockout of the putative operon nanCM (OP50<sup>nanCM</sup>) was obtained using  $\lambda$ -Red Recombinase System described previously (Datsenko and Wanner, 2000). OP50<sup>nanCM</sup> genotyping was confirmed with PCR of genomic DNA, followed by Sanger sequencing.

### Statistical analysis

In all experiments, Student's *t*-tests and Chi-square tests were used to calculate *P* values, as indicated in the Figure legends. For Student's *t*-tests, the values reported in the graphs represented averages of three independent experiments, with error bars showing SEM. For bioinformatics analysis, Mann-Whitney *U* tests or Wilcoxon matched-pairs signed rank tests were used to calculate *P* values, as indicated in the Figure legends.

### Accession numbers

All high-throughput sequencing data in this study have been deposited in the Gene Expression Omnibus (GEO) database and are available under accession number GSE212382. All experimental materials generated in this study are available upon request.

**Compliance and ethics** *The author(s) declare that they have no conflict of interest.*

**Acknowledgements** *This work was supported by the National Key Research and Development Program of China (2019YFA0802600), the National Natural Science Foundation of China (31930019, 31900442, 32200431), the Strategic Priority Research Program “Biological basis of aging and therapeutic strategies” of the Chinese Academy of Sciences (XDB39010400), and the China Postdoctoral Science Foundation (2022M713053). We gratefully thank Dr. Shouhong Guang (USTC) and Dr. Baolin Sun (USTC) for providing plasmids, strains, and experimental facilities. We thank the Bioinformatics Center of the USTC, Division of Life Sciences and Medicine, for providing supercomputing resources.*

### References

Alcazar, R.M., Lin, R., and Fire, A.Z. (2008). Transmission dynamics of heritable silencing induced by double-stranded RNA in *Caenorhabditis elegans*. *Genetics* 180, 1275–1288.

Almeida, M.V., Andrade-Navarro, M.A., and Ketting, R.F. (2019). Function and evolution of nematode RNAi pathways. *ncRNA* 5, 8.

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.

Arciga-Reyes, L., Wootton, L., Kieffer, M., and Davies, B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in

*Arabidopsis*. *Plant J* 47, 480–489.

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.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.

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.

Bošković, A., and Rando, O.J. (2018). Transgenerational epigenetic inheritance. *Annu Rev Genet* 52, 21–41.

Bouasker, S., and Simard, M.J. (2012). The slicing activity of miRNA-specific Argonautes is essential for the miRNA pathway in *C. elegans*. *Nucleic Acids Res* 40, 10452–10462.

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

Brown, K.C., Svendsen, J.M., Tucci, R.M., Montgomery, B.E., and Montgomery, T.A. (2017). ALG-5 is a miRNA-associated Argonaute required for proper developmental timing in the *Caenorhabditis elegans* germline. *Nucleic Acids Res* 45, 9093–9107.

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.

Bukhari, S.I.A., Vasquez-Rifo, A., Gagné, D., Paquet, E.R., Zetka, M., Robert, C., Masson, J.Y., and Simard, M.J. (2012). The microRNA pathway controls germ cell proliferation and differentiation in *C. elegans*. *Cell Res* 22, 1034–1045.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642–655.

Charlesworth, A.G., Seroussi, U., Lehrbach, N.J., Renaud, M.S., Sundby, A.E., Molnar, R.I., Lao, R.X., Willis, A.R., Woock, J.R., Aber, M.J., et al. (2021). Two isoforms of the essential *C. elegans* Argonaute CSR-1 differentially regulate sperm and oocyte fertility. *Nucleic Acids Res* 49, 8836–8865.

Chen, X., and Rechavi, O. (2022). Plant and animal small RNA communications between cells and organisms. *Nat Rev Mol Cell Biol* 23, 185–203.

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.

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 USA* 107, 3588–3593.

Corrêa, R.L., Steiner, F.A., Berezikov, E., and Ketting, R.F. (2010). MicroRNA-directed siRNA biogenesis in *Caenorhabditis elegans*. *PLoS Genet* 6, e1000903.

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.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97, 6640–6645.

Domeier, M.E., Morse, D.P., Knight, S.W., Portereiko, M., Bass, B.L., and Mango, S.E. (2000). A link between RNA interference and nonsense-mediated decay in *Caenorhabditis elegans*. *Science* 289, 1928–1930.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10, 48.

Fire, A., Xu, S.Q., Montgomery, M.K., Kostas, S.A., Driver, S.E., and

- Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.
- Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiácovo, M.P., Church, G.M., and Calarco, J.A. (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods* 10, 741–743.
- 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.
- Gu, H., Lian, B., Yuan, Y., Kong, C., Li, Y., Liu, C., and Qi, Y. (2022). A 5' tRNA-Ala-derived small RNA regulates anti-fungal defense in plants. *Sci China Life Sci* 65, 1–15.
- Gu, S.G., Pak, J., Guang, S., Maniar, J.M., Kennedy, S., and Fire, A. (2012). Amplification of siRNA in *Caenorhabditis elegans* generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. *Nat Genet* 44, 157–164.
- Gu, W., Shirayama, M., Conte Jr., D., 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.
- Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowicz, J., and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. *Science* 321, 537–541.
- Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* 138, 696–708.
- Höck, J., and Meister, G. (2008). The Argonaute protein family. *Genome Biol* 9, 210.
- Houri-Zeevi, L., Korem Kohanim, Y., Antonova, O., and Rechavi, O. (2020). Three rules explain transgenerational small RNA inheritance in *C. elegans*. *Cell* 182, 1186–1197.e12.
- Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22–32.
- 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.
- Ketting, R.F. (2011). The many faces of RNAi. *Dev Cell* 20, 148–161.
- 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.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35, 1547–1549.
- Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte Jr., D., and Mello, C.C. (2012). *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* 150, 78–87.
- 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.
- Lehrbach, N.J., Armisen, J., Lightfoot, H.L., Murfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in *Caenorhabditis elegans*. *Nat Struct Mol Biol* 16, 1016–1020.
- Lev, I., Toker, I.A., Mor, Y., Nitzan, A., Weintraub, G., Antonova, O., Bhonkar, O., Ben Shushan, I., Seroussi, U., Claycomb, J.M., et al. (2019). Germ granules govern small RNA inheritance. *Curr Biol* 29, 2880–2891.e4.
- Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. *Curr Biol* 15, 1501–1507.
- Liu, C., Cao, B., and Cao, X. (2022). Biogenesis, action and biological functions of an *Arabidopsis* 5' tRF, 5' tsR-Ala. *Sci China Life Sci* 65, 1050–1052.
- Liu, H., Wang, X., Wang, H.D., Wu, J.J., Ren, J., Meng, L., Wu, Q., Dong, H., Wu, J., Kao, T.Y., et al. (2012). *Escherichia coli* noncoding RNAs can affect gene expression and physiology of *Caenorhabditis elegans*. *Nat Commun* 3, 1073.
- Liu, S., Xie, S., Chen, H., Li, B., Chen, Z., Tan, Y., Yang, J., Zheng, L., Xiao, Z., Zhang, Q., et al. (2021). The functional analysis of transiently upregulated miR-101 suggests a “braking” regulatory mechanism during myogenesis. *Sci China Life Sci* 64, 1612–1623.
- Liu, X., Wang, X., Li, J., Hu, S., Deng, Y., Yin, H., Bao, X., Zhang, Q.C., Wang, G., Wang, B., et al. (2020). Identification of mecciRNAs and their roles in the mitochondrial entry of proteins. *Sci China Life Sci* 63, 1429–1449.
- Maniar, J.M., and Fire, A.Z. (2011). EGO-1, a *C. elegans* RdRP, modulates gene expression via production of mRNA-templated short antisense RNAs. *Curr Biol* 21, 449–459.
- Mao, H., Zhu, C., Zong, D., Weng, C., Yang, X., Huang, H., Liu, D., Feng, X., and Guang, S. (2015). The Nrde pathway mediates small-RNA-directed histone H3 lysine 27 trimethylation in *Caenorhabditis elegans*. *Curr Biol* 25, 2398–2403.
- Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* 14, 447–459.
- 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.e12.
- Nguyen, D.A.H., and Phillips, C.M. (2021). Arginine methylation promotes siRNA-binding specificity for a spermatogenesis-specific isoform of the Argonaute protein CSR-1. *Nat Commun* 12, 4212.
- Ortiz, M.A., Noble, D., Sorokin, E.P., and Kimble, J. (2014). A new dataset of spermatogenic vs. oogenic transcriptomes in the nematode *Caenorhabditis elegans*. *G3* 4, 1765–1772.
- Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* 315, 241–244.
- Parrish, S., and Fire, A. (2001). Distinct roles for RDE-1 and RDE-4 during RNA interference in *Caenorhabditis elegans*. *RNA* 7, 1397–1402.
- Rechavi, O., Houry-Zeevi, L., Anava, S., Goh, W.S.S., Kerk, S.Y., Hannon, G.J., and Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* 158, 277–287.
- 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.
- Samuel, B.S., Rowedder, H., Braendle, C., Félix, M.A., and Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proc Natl Acad Sci USA* 113, E3941–E3949.
- Sasaki, T., Shiohama, A., Minoshima, S., and Shimizu, N. (2003). Identification of eight members of the Argonaute family in the human genome. *Genomics* 82, 323–330.
- Saw, P.E., Xu, X., Chen, J., and Song, E.W. (2021). Non-coding RNAs: the new central dogma of cancer biology. *Sci China Life Sci* 64, 22–50.
- Scheer, H., Zuber, H., De Almeida, C., and Gagliardi, D. (2016). Uridylation earmarks mRNAs for degradation... and more. *Trends Genet* 32, 607–619.
- Schreier, J., Dietz, S., Boermel, M., Oorschot, V., Seistrup, A.S., de Jesus Domingues, A.M., Bronkhorst, A.W., Nguyen, D.A.H., Phillis, S., Gleason, E.J., et al. (2022). Membrane-associated cytoplasmic granules carrying the Argonaute protein WAGO-3 enable paternal epigenetic inheritance in *Caenorhabditis elegans*. *Nat Cell Biol* 24, 217–229.
- Schwartz-Orbach, L., Zhang, C., Sidoli, S., Amin, R., Kaur, D., Zhebrun, A., Ni, J., and Gu, S.G. (2020). *Caenorhabditis elegans* nuclear RNAi factor SET-32 deposits the transgenerational histone modification, H3K23me3. *eLife* 9, e54309.
- 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.e18.
- Si, F., Luo, H., Yang, C., Gong, J., Yan, B., Liu, C., Song, X., and Cao, X. (2023). Mobile ARGONAUTE 1d binds 22-nt miRNAs to generate phasiRNAs important for low-temperature male fertility in rice. *Sci China Life Sci* 66, 197–208.
- Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling

- sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* 27, 491–499.
- Stoeckius, M., Maaskola, J., Colombo, T., Rahn, H.P., Friedländer, M.R., Li, N., Chen, W., Piano, F., and Rajewsky, N. (2009). Large-scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression. *Nat Methods* 6, 745–751.
- Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99, 123–132.
- Taft, R.J., Simons, C., Nahkuri, S., Oey, H., Korbie, D.J., Mercer, T.R., Holst, J., Ritchie, W., Wong, J.J.L., Rasko, J.E., et al. (2010). Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. *Nat Struct Mol Biol* 17, 1030–1034.
- Timmons, L., Court, D.L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* 263, 103–112.
- van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell* 139, 135–148.
- Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte Darryl, J. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci USA* 107, 3582–3587.
- Vasquez-Rifo, A., Jannot, G., Armisen, J., Labouesse, M., Bukhari, S.I.A., Rondeau, E.L., Miska, E.A., and Simard, M.J. (2012). Developmental characterization of the microRNA-specific *C. elegans* Argonautes *alg-1* and *alg-2*. *PLoS ONE* 7, e33750.
- Vastenhouw, N.L., Fischer, S.E.J., Robert, V.J.P., Thijssen, K.L., Fraser, A. G., Kamath, R.S., Ahringer, J., and Plasterk, R.H.A. (2003). A genome-wide screen identifies 27 genes involved in transposon silencing in *C. elegans*. *Curr Biol* 13, 1311–1316.
- Vaucheret, H. (2008). Plant ARGONAUTES. *Trends Plant Sci* 13, 350–358.
- Wajahat, M., Bracken, C.P., and Orang, A. (2021). Emerging functions for snoRNAs and snoRNA-derived fragments. *Int J Mol Sci* 22, 10193.
- Wang, C., and Lin, H. (2021). Roles of piRNAs in transposon and pseudogene regulation of germline mRNAs and lincRNAs. *Genome Biol* 22, 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.
- Wei, S., Chen, H., Dzakah, E.E., Yu, B., Wang, X., Fu, T., Li, J., Liu, L., Fang, S., Liu, W., et al. (2019). Systematic evaluation of *C. elegans* lincRNAs with CRISPR knockout mutants. *Genome Biol* 20, 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.
- Xu, F., Feng, X., Chen, X., Weng, C., Yan, Q., Xu, T., Hong, M., and Guang, S. (2018). A cytoplasmic Argonaute protein promotes the inheritance of RNAi. *Cell Rep* 23, 2482–2494.
- Xue, Y., Chen, R., Qu, L., and Cao, X. (2020). Noncoding RNA: from dark matter to bright star. *Sci China Life Sci* 63, 463–468.
- Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C.G., 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.
- Yu, B., Wang, X., Wei, S., Fu, T., Dzakah, E.E., Waqas, A., Walthall, W.W., and Shan, G. (2017). Convergent transcriptional programs regulate cAMP levels in *C. elegans* GABAergic motor neurons. *Dev Cell* 43, 212–226.e7.
- Zhang, L., Ding, L., Cheung, T.H., Dong, M.Q., Chen, J., Sewell, A.K., Liu, X., Yates III, J.R., and Han, M. (2007). Systematic identification of *C. elegans* miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol Cell* 28, 598–613.
- Zhou, X., Feng, X., Mao, H., Li, M., Xu, F., Hu, K., and Guang, S. (2017). RdRP-synthesized antisense ribosomal siRNAs silence pre-rRNA via the nuclear RNAi pathway. *Nat Struct Mol Biol* 24, 258–269.

## SUPPORTING INFORMATION

The supporting information is available online at <https://doi.org/10.1007/s11427-022-2304-8>. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.