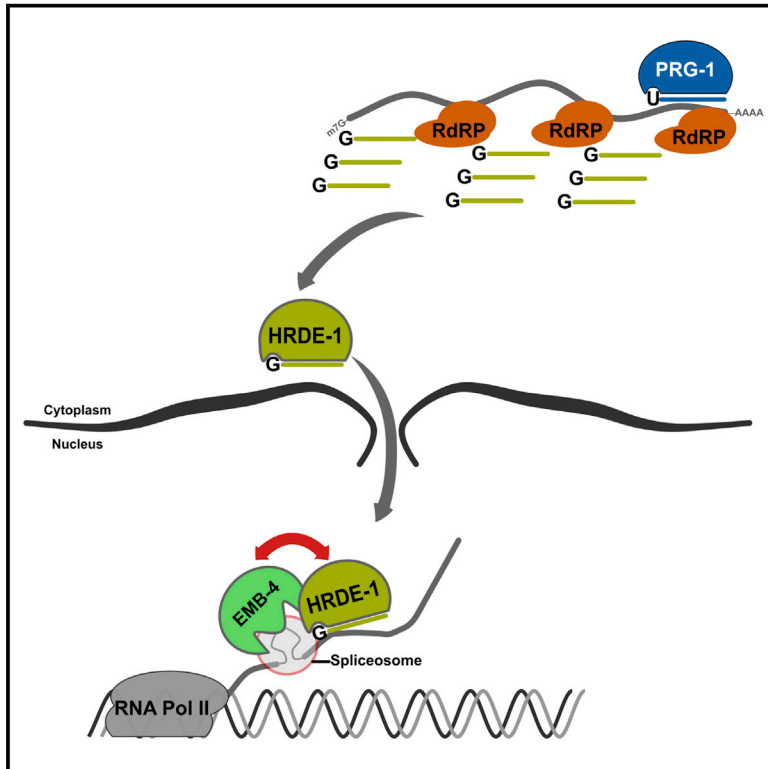


# Developmental Cell

## The Helicase Aquarius/EMB-4 Is Required to Overcome Intronic Barriers to Allow Nuclear RNAi Pathways to Heritably Silence Transcription

### Graphical Abstract



### Authors

Alper Akay, Tomas Di Domenico, Kin M. Suen, ..., Angus I. Lamond, Julie M. Claycomb, Eric A. Miska

### Correspondence

eric.miska@gurdon.cam.ac.uk

### In Brief

Nuclear RNAi is essential for heritable silencing in many organisms. Akay et al. show that RNA helicase EMB-4/AQR physically interacts with nuclear Argonaute HRDE-1 and bridges the small RNA pathway to transcriptional silencing. They further show that introns can form a barrier to transcriptional silencing, which EMB-4/AQR can overcome.

### Highlights

- Nuclear Argonaute HRDE-1 physically interacts with the helicase EMB-4/AQR
- EMB-4/AQR is required for heritable silencing of genes and transposable elements
- Intronic sequences form a barrier to nuclear RNAi pathway
- EMB-4/AQR removes intronic barriers for efficient silencing by HRDE-1



# The Helicase Aquarius/EMB-4 Is Required to Overcome Intronic Barriers to Allow Nuclear RNAi Pathways to Heritably Silence Transcription

Alper Akay,<sup>1,2</sup> Tomas Di Domenico,<sup>1,2,3</sup> Kin M. Suen,<sup>1,2</sup> Amena Nabih,<sup>4</sup> Guillermo E. Parada,<sup>1,3</sup> Mark Larence,<sup>5</sup> Ragini Medhi,<sup>1,2</sup> Ahmet C. Berkuyrek,<sup>1,2</sup> Xinlian Zhang,<sup>6</sup> Christopher J. Wedeles,<sup>4</sup> Konrad L.M. Rudolph,<sup>1,2,3</sup> Jan Engelhardt,<sup>7</sup> Martin Hemberg,<sup>1,3</sup> Ping Ma,<sup>6</sup> Angus I. Lamond,<sup>5</sup> Julie M. Claycomb,<sup>4</sup> and Eric A. Miska<sup>1,2,3,8,\*</sup>

<sup>1</sup>Wellcome Trust Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK

<sup>2</sup>Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK

<sup>3</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SA, UK

<sup>4</sup>Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

<sup>5</sup>Centre for Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK

<sup>6</sup>Department of Statistics, University of Georgia, Athens, GA 30602, USA

<sup>7</sup>Bioinformatics Group, Department of Computer Science, Interdisciplinary Center for Bioinformatics, University of Leipzig, Haertelstraße 16-18, Leipzig 04107, Germany

<sup>8</sup>Lead Contact

\*Correspondence: [eric.miska@gurdon.cam.ac.uk](mailto:eric.miska@gurdon.cam.ac.uk)

<http://dx.doi.org/10.1016/j.devcel.2017.07.002>

## SUMMARY

Small RNAs play a crucial role in genome defense against transposable elements and guide Argonaute proteins to nascent RNA transcripts to induce co-transcriptional gene silencing. However, the molecular basis of this process remains unknown. Here, we identify the conserved RNA helicase Aquarius/EMB-4 as a direct and essential link between small RNA pathways and the transcriptional machinery in *Caenorhabditis elegans*. Aquarius physically interacts with the germline Argonaute HRDE-1. Aquarius is required to initiate small-RNA-induced heritable gene silencing. HRDE-1 and Aquarius silence overlapping sets of genes and transposable elements. Surprisingly, removal of introns from a target gene abolishes the requirement for Aquarius, but not HRDE-1, for small RNA-dependent gene silencing. We conclude that Aquarius allows small RNA pathways to compete for access to nascent transcripts undergoing co-transcriptional splicing in order to detect and silence transposable elements. Thus, Aquarius and HRDE-1 act as gatekeepers coordinating gene expression and genome defense.

## INTRODUCTION

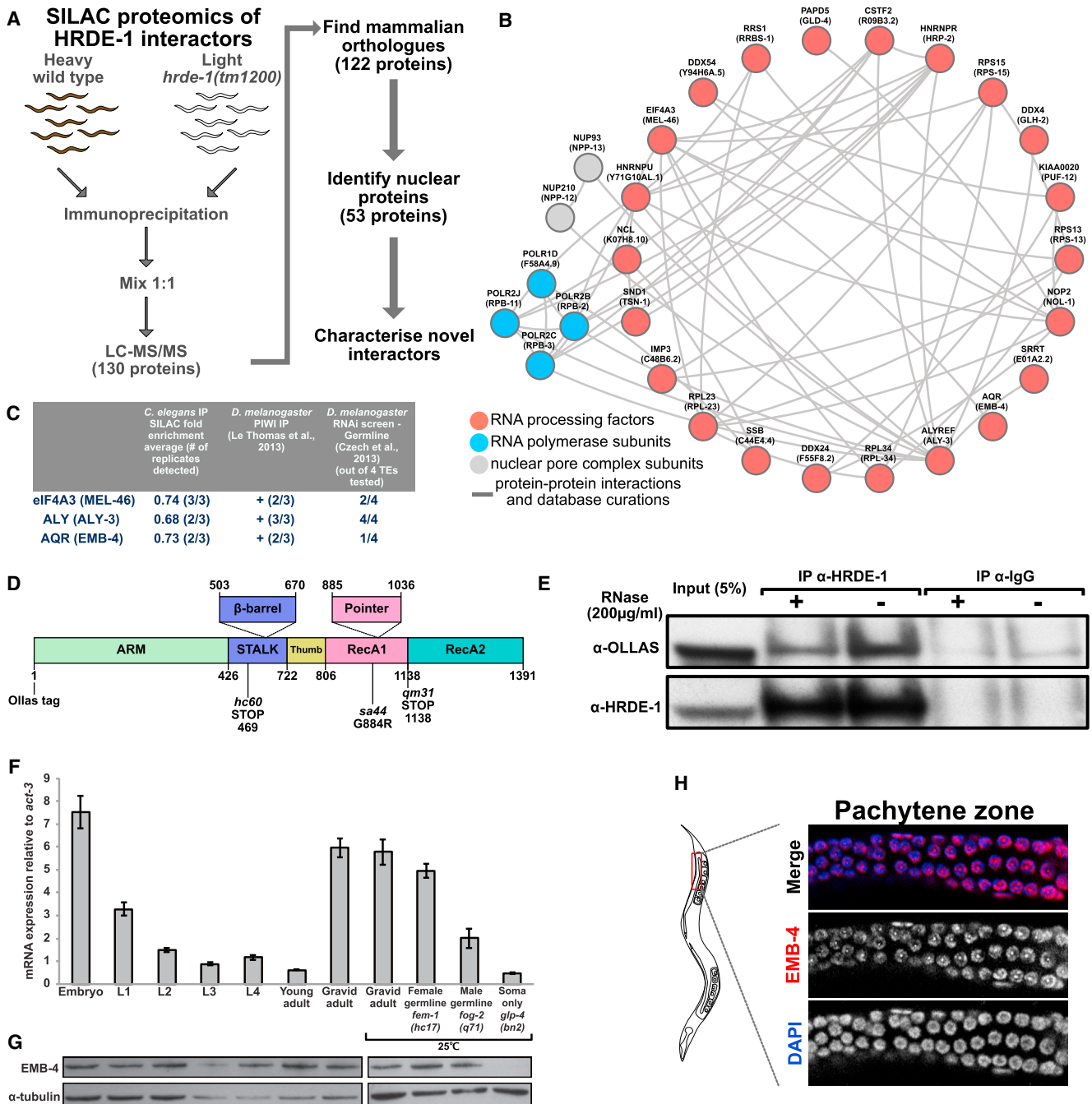
Transposable elements (TEs) are a feature of almost all eukaryotic genomes, which if left unchecked present a considerable danger to genome integrity and species survival. Therefore, organisms have evolved robust pathways to silence the expression of transposable elements and to restrict their mobility (Slotkin and Martienssen, 2007). Among these, DNA and chromatin modification and small RNA (sRNA)-mediated silencing

are ancient and conserved mechanisms. sRNA pathways in eukaryotes are all related to RNAi mechanisms (Fire et al., 1998): 21–32-nt sRNAs are bound by Argonaute superfamily proteins, interact with target RNAs through Watson-Crick base pairing, and initiate silencing of these targets. Such sRNA-mediated silencing can be post-transcriptional on mRNAs in the cytoplasm (PTGS) or co-transcriptional on nascent transcripts in the nucleus (coTGS). The latter provides the potential to couple sRNA-mediated silencing to DNA and chromatin-based gene regulatory pathways. The mechanisms by which such sRNA-mediated silencing affects DNA methylation and/or chromatin modification remain largely unknown, particularly in animals (Holoch and Moazed, 2015; Weick and Miska, 2014).

Most organisms have also evolved sRNA amplification systems to enhance sRNA-mediated silencing. The ancestral amplification mechanism is based on RNA-dependent RNA polymerases (RdRPs), which use target RNAs as templates to generate secondary sRNAs. Some animals, including mammals, have lost RdRPs but have instead evolved the Ping-Pong amplification system (Brennecke et al., 2007; Gunawardane et al., 2007). The Ping-Pong pathway amplifies Piwi-interacting RNAs (piRNAs), an animal-specific class of sRNAs, through double-stranded target RNA intermediates and Argonaute RNase activity (RNA slicing). piRNAs are named after a subfamily of Argonautes, the Piwi proteins (Cox et al., 1998). Piwi proteins and piRNAs act in TE silencing in the germline of many animals. Another distinction of piRNAs is that they act in *trans*, whereby piRNAs generated from genomic clusters silence TEs throughout the genome. piRNA clusters can be thought of as a memory of TEs within the genome of an organism analogous to the guide RNA cluster/CRISPR systems of prokaryotes (Wiedenheft et al., 2012). Like other sRNAs, piRNAs have been shown to act through PTGS and coTGS mechanisms in nematodes, insects, fish, and mice (Holoch and Moazed, 2015; Weick and Miska, 2014).

*Caenorhabditis elegans* has both PTGS and coTGS mechanisms in the soma and the germline (Weick and Miska, 2014).





**Figure 1. SILAC Proteomics Identifies Aquarius/EMB-4 as an Interactor of the Nuclear Argonaute HRDE-1**

(A) SILAC labeling and IP scheme for wild-type (heavy labeled) and *hrde-1(tm1200)* mutants (light labeled) using anti-HRDE-1 antibodies.

(B) Known protein interactions identified by the STRING database for mammalian orthologues of nuclear HRDE-1 interactors. Gray lines indicate known protein-protein interactions, red circles highlight known RNA-processing factors, blue circles highlight RNA pol II subunits, and gray circles highlight nuclear pore complex subunits.

(C) Aquarius and two exon-junction complex proteins eIF4A3 and ALY are detected in HRDE-1 IPs (second column, mean log<sub>2</sub> fold enrichment heavy/light and number of replicates detected) and in *D. melanogaster* PIWI IPs (third column, number of IPs detected/total number of IPs). The effect on TE desilencing in *D. melanogaster* upon RNAi knockdown (fourth column, number of TEs desilenced out of four tested).

(D) EMB-4 domain structure based on the mammalian homolog Aquarius and the position of *emb-4* mutations used in this study.

(E) Validation of protein-protein interactions between HRDE-1 and EMB-4 using anti-HRDE-1 antibodies to immunoprecipitate HRDE-1 complexes in CRISPR-tagged OLLAS-EMB-4 strain with or without RNase treatment (immunoglobulin G [IgG] negative control).

(legend continued on next page)

In the germline, sRNAs and piRNAs can initiate coTGS through a two-step mechanism. *C. elegans* piRNAs are 21-nt RNAs with a 5' uracil (21U-RNAs) that are bound by the PRG-1 Piwi protein in the germline cytoplasm (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). Once a PRG-1/piRNA complex has recognized a target RNA it recruits an RdRP-containing complex to generate 22-nt antisense sRNAs with a 5' guanine (22G-RNAs) (Bagijn et al., 2012; Pak and Fire, 2007). 22G-RNAs are then bound by the Argonaute HRDE-1 and imported into the nucleus (Ashe et al., 2012; Buckley et al., 2012). An HRDE-1/22G-RNA complex is thought to directly interact with nascent transcripts. Genetic screens have identified several additional factors that are required for HRDE-1-mediated coTGS including NRDE factors (NRDE-1/-2/-4), the H3K9 histone methyltransferases SET-25 and SET-32, and the heterochromatin protein 1 (HP1) homolog HPL-2 (Ashe et al., 2012; Buckley et al., 2012; Burkhart et al., 2011; Guang et al., 2008; Shirayama et al., 2012). However, how HRDE-1 links sRNA-mediated silencing to coTGS and chromatin modifications remains unknown.

Transcription and mobility of TEs is actively suppressed by this two-step piRNA/22G-RNA coTGS system in *C. elegans*. Interestingly, piRNA- and sRNA-mediated coTGS can last for multiple generations (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012) and can be the source of non-genetic transgenerational effects (Ashe et al., 2012; Jobson et al., 2015; Ni et al., 2016; Rechavi et al., 2014), and loss of genes in the piRNA and the 22G-RNA pathway result in a multi-generational loss of fertility (de Albuquerque et al., 2015; Buckley et al., 2012; Phillips et al., 2015; Simon et al., 2014).

Eukaryotic mRNA transcription is a complex, multi-step process. First, the RNA polymerase II holoenzyme assembles on genomic DNA at a transcription start site. Second, nascent transcripts are produced by the elongating RNA polymerase II. Third, nascent RNA transcripts are processed to mature mRNAs through the assembly of multiple large RNA protein (RNP) complexes to carry out 5' capping (Gonatopoulos-Pournatzis and Cowling, 2014), splicing (Carrillo Oesterreich et al., 2016), and 3' poly(A) tailing (Shatkin and Manley, 2000). All of these steps are required to protect transcripts from degradation and to ensure that mRNAs are successfully exported into the cytoplasm for protein translation. Of these co-transcriptional events splicing is probably the most complex of all, requiring more than 200 proteins and many non-coding RNAs (Nielsen, 2003; Wahl and Lührmann, 2015). Pre-mRNA splicing by the spliceosome is immediately followed by the assembly of a set of proteins known as the exon-junction complex (EJC). EJC functions in the export, localization, and translation of mRNAs. Assembly of different EJC components can determine the fate of mRNAs, and EJC can be considered a regulatory hub between transcription and translation (Le Hir et al., 2016).

Interestingly, most TEs are adapted to the host transcription and RNA-processing machineries, and exploit them for

their own expression and mobility. Furthermore, TEs have been domesticated in such a way that their regulation can contribute to gene regulation (Cowley and Oakey, 2013; Rebollo et al., 2012).

Here, we used a proteomics approach to identify protein interactors of the germline Argonaute HRDE-1 in *C. elegans*. One HRDE-1 interacting factor is the conserved RNA helicase Aquarius/EMB-4. HRDE-1 and Aquarius/EMB-4 act to silence an overlapping set of TEs and TE-containing genes. Surprisingly, removal of introns from a coTGS target removes the requirement for Aquarius/EMB-4 in sRNA-mediated silencing. Thus, Aquarius/EMB-4 activity allows HRDE-1 access to the nascent RNA transcripts.

## RESULTS

### SILAC Proteomics Identifies Protein Interactors of the Germline Nuclear Argonaute HRDE-1

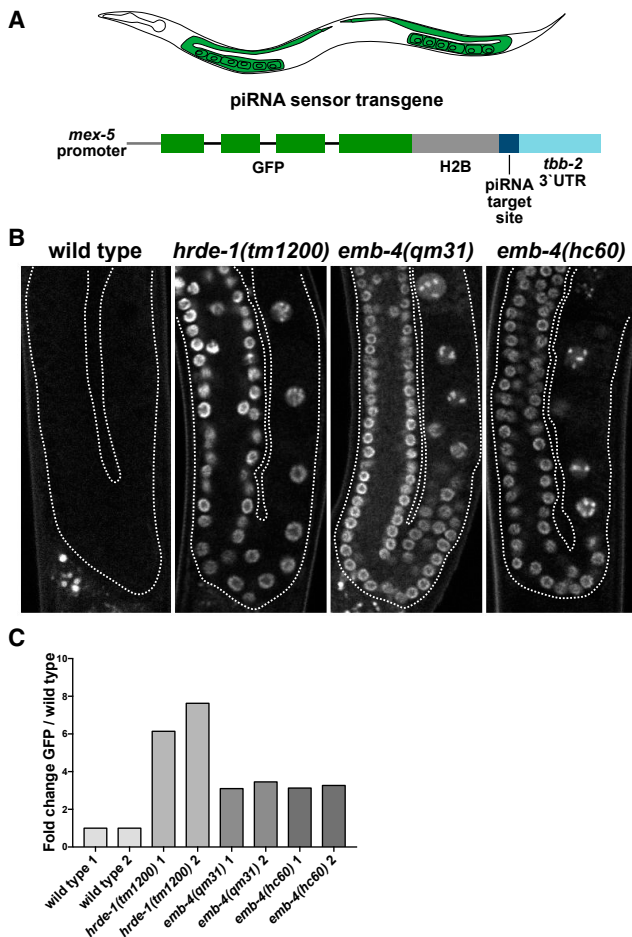
To further our understanding of coTGS in animals and as a complement to genetic approaches, we sought to identify proteins interacting with the nuclear Argonaute HRDE-1 by protein immunoprecipitation (IP) coupled with SILAC (stable isotope labeling of amino acids in cell culture) proteomics (Ong et al., 2002), using the SILAC labeling for nematodes approach (Akay et al., 2013; Fredens et al., 2011; Larance et al., 2011) (Figure 1A). We produced protein extracts from wild-type young adult *C. elegans* grown on “heavy” isotopes and *hrde-1(tm1200)* mutant animals grown on “light” isotopes. We chose the young adult stage, as the germline is then fully developed and endogenous HRDE-1 expression is at its peak (Ashe et al., 2012). We then immunoprecipitated a 1:1 mix of protein extracts of both genotypes using an anti-HRDE-1 antibody followed by liquid chromatography-tandem mass spectrometry (LC/MS-MS). Based on three biological replicates, LC/MS-MS identified 130 candidate HRDE-1 interacting proteins. For further analysis and to take advantage of prior proteomics data from mammals, we focused on 53 proteins with human orthologs that are also classified as being nuclear (Uhlén et al., 2015; UniProt Consortium, 2015). We used the STRING database (Szklarczyk et al., 2015) to identify known protein complexes between these factors. Twenty-eight of these 53 proteins were connected with each other based on known interactions, and the majority of the proteins represent RNA processing factors, RNA polymerase II subunits, and nuclear pore complex components (Figures 1B and S1A). In the fruit fly *Drosophila melanogaster*, the Piwi protein has a function analogous to that of HRDE-1 in coTGS in the germline (Weick and Miska, 2014). Interestingly, eight proteins identified in our HRDE-1 IPs were also found in *D. melanogaster* PIWI IPs. In addition, 13 of our HRDE-1 interactors were hits in genome-wide RNAi screens for regulators of transposable elements (Figure 1C and Table S1) (Czech et al., 2013; Le Thomas et al., 2013). However, the role of these interactors in the sRNA pathway is currently not

(F) qRT-PCR analysis of *emb-4* mRNA expression at different developmental stages (embryo to gravid adult), in animals lacking sperm (female germline *fem-1(hc17)*), in animals lacking oocytes (male germline *fog-2(q71)*), and in animals lacking a germline (soma only *glp-4(bn2)*). *fem-1(hc17)*, *fog-2(q71)*, and *glp-4(bn2)* are temperature-sensitive mutants and were grown at 25°C alongside the wild-type gravid adult control. Error bars represent SD.

(G) Western blot analysis of EMB-4 protein levels using anti-EMB-4 antibodies for the same conditions as in Figure 2A.

(H) Localization of EMB-4 protein in the germline of adult animals.

See also Figures S1 and S2.



**Figure 2. Aquarius/EMB-4 Is Required for the piRNA-Mediated Silencing of a Sensor Transgene**

(A) piRNA sensor transgene and its expression pattern in *C. elegans* germline. (B) Fluorescent microscope images of wild-type and mutant animal germlines with an integrated single-copy piRNA sensor transgene (germline boundaries are marked by white dotted lines).

(C) qRT-PCR analysis of GFP expression in animals described in Figure 3B (two biological replicates with at least two technical replicates).

See also Figure S3.

understood in any organism. Overall, our SILAC-based HRDE-1 proteomics identified a set of conserved proteins that might bridge sRNA pathways, RNA processing, and TE silencing.

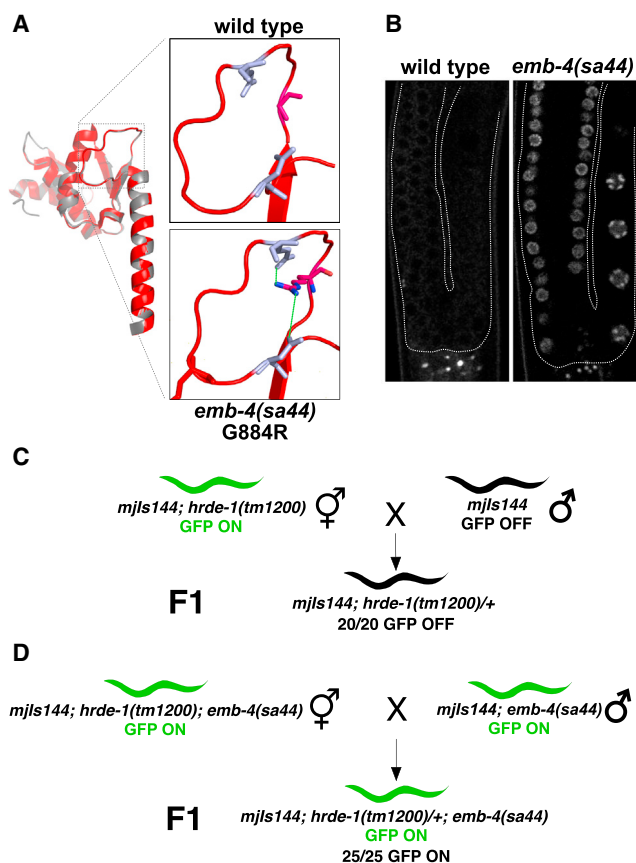
### The Aquarius RNA Helicase Ortholog EMB-4 Is Nuclear and Germline Enriched and Interacts with HRDE-1

Of the candidate HRDE-1 protein interactors, we were particularly curious about EMB-4 as its interaction with nuclear Argonaute proteins that appeared to be conserved in *D. melanogaster* (Figure 1C). EMB-4 is the *C. elegans* ortholog of Aquarius, or AQR, a large scaffolding protein that includes an Armadillo repeat and two RecA helicase domains (Figure 1D). Aquarius/EMB-4 is conserved in all eukaryotes with an intact RNAi pathway, including the fission yeast *Schizosaccharomyces pombe*, but not *Saccharomyces cerevisiae*, which lacks RNAi (Drinnenberg et al., 2011) (Figures S1A–S1C). As Aquarius has a conserved role in nascent

RNA processing in eukaryotes (Figure S1D) (De et al., 2015; Wahl and Lührmann, 2015), we wondered whether Aquarius/EMB-4 might provide new insights into the interface between the nuclear RNAi pathway and the general RNA-processing machinery. Therefore, we validated the interaction between HRDE-1 and EMB-4 independently. First, we generated an N-terminal epitope-tagged version of EMB-4, *in vivo*, using CRISPR/Cas9 genome engineering (Paix et al., 2015; Wiedenheft et al., 2012) (Ollas tag, Figure 1D). We then tested the interaction between HRDE-1 and EMB-4-Ollas using an anti-Ollas antibody (Park et al., 2008) and an anti-HRDE-1 antibody (Ashe et al., 2012), through IP followed by western blotting. As shown in Figure 1E, we found that EMB-4-Ollas interacts specifically with HRDE-1 in whole-cell extracts (IP  $\alpha$ -HRDE-1(–) RNase treatment, Figures S1E and S1F). A significant portion of the EMB-4/HRDE-1 interaction remains intact upon strong RNase treatment (Figure 1E, IP  $\alpha$ -HRDE-1(+) RNase treatment, Figure S1F), consistent with a model in which EMB-4 binds directly to HRDE-1 but with interaction being stabilized on the nascent RNA transcript. In addition, we validated this interaction using an anti-EMB-4 antibody and an FLAG-tagged version of HRDE-1 generated by MosSCI transgenesis (Frøkjær-Jensen et al., 2008; Shirayama et al., 2012) (Figures S1G and S1H). Next, we asked whether EMB-4 is co-expressed with HRDE-1 *in vivo*. We found that *emb-4* mRNA was highly enriched in the germline, embryo, and early larval stages of *C. elegans* (Figure 1F). Using an anti-EMB-4 antibody, we observed the same for the endogenous EMB-4 protein (Figure 1G). Taking advantage of temperature-sensitive germline mutants of *C. elegans*, we found that in gravid adult animals the majority of EMB-4 mRNA and protein is germline restricted (Figures 1F and 1G). This is of particular interest, as HRDE-1 expression is germline specific (Ashe et al., 2012). Within the germline, we find that EMB-4 is highly enriched in germline nuclei and closely associated with chromatin in the mitotic, transition, and pachytene region of the germline as well as in oocytes (Figures 1H, S2A, and S2B). Mutations in *hrde-1* do not affect the nuclear localization of EMB-4, in line with the general role of EMB-4 in pre-mRNA processing. Similarly, mutations in *emb-4* do not change the overall nuclear localization of HRDE-1 (Figure S2B). Our observations are in agreement with the findings of the related work in this issue of *Developmental Cell* by Tyc et al. (2017). We conclude that EMB-4 is expressed in the germline nuclei and is a bona fide HRDE-1 interacting protein.

### Aquarius/EMB-4 Is Required for piRNA-Mediated Co-transcriptional Gene Silencing

We examined whether, like the Argonaute HRDE-1, Aquarius/EMB-4 is required for coTGS in the germline. For this purpose we used a piRNA sensor strain we generated previously (Bagijn et al., 2012). The piRNA sensor (*mjls144*) is a transgene that drives the expression of a GFP histone H2B fusion protein through the germline-specific *mex-5* gene promoter and is inserted as a single copy on chromosome II. In addition, the piRNA sensor contains a 21-nt sequence perfectly complementary to the endogenous piRNA 21UR-1 (Figure 2A). In otherwise wild-type animals, the piRNA sensor is silenced by piRNA-mediated coTGS. However, mutations in the piRNA or the germline nuclear RNAi pathway, e.g., *prg-1* (Bagijn et al., 2012), *prde-1* (Weick et al., 2014), or *hrde-1* (Ashe et al., 2012), lead to reactivation



**Figure 3. RNA Helicase Domain of Aquarius/EMB-4 Is Required for the Establishment of Transcriptional Gene Silencing**

(A) Structural alignment of the RecA1 and the pointer domains of human Aquarius (white) with *C. elegans* EMB-4 (red). Blown-up region shows the location of the G884R substitution found in *emb-4(sa44)* strain. G884R affects the loop region of the helicase domain. Green lines show possible salt bridge interactions between the amino acids.

(B) Fluorescent images of wild-type and *emb-4(sa44)* germlines with the integrated piRNA sensor transgene (germline boundaries are marked by white dotted lines).

(C and D) Scheme of genetic crosses showing the effect of *emb-4(sa44)* mutation during the establishment of gene silencing. (C) In the control cross, a wild-type copy of *hrde-1(+)* in the F1 heterozygous animals re-establishes complete piRNA sensor silencing. (D) Wild-type copy of *hrde-1* fails to establish piRNA sensor silencing in the *emb-4(sa44)* homozygous background (number of assayed F1 progeny is indicated below each cross). See also Figure S4.

of the piRNA sensor. Taking advantage of two previously isolated alleles of *emb-4* (Checchi and Kelly, 2006; Katic and Greenwald, 2006), we tested the requirement of EMB-4 in piRNA-mediated gene silencing. We observed that two null mutants of *emb-4* desilenced the piRNA sensor in the germline of *C. elegans* as indicated by GFP expression in germline nuclei, as did *hrde-1* as the positive control, while the piRNA sensor was silenced in otherwise wild-type animals (Figures 2B and S3A). We also quantified mRNA levels of piRNA sensor expression in these mutant backgrounds, which confirmed the desilencing in *emb-4* and *hrde-1* mutant backgrounds (Figure 2C). Finally, we carried out an analogous experiment using an inde-

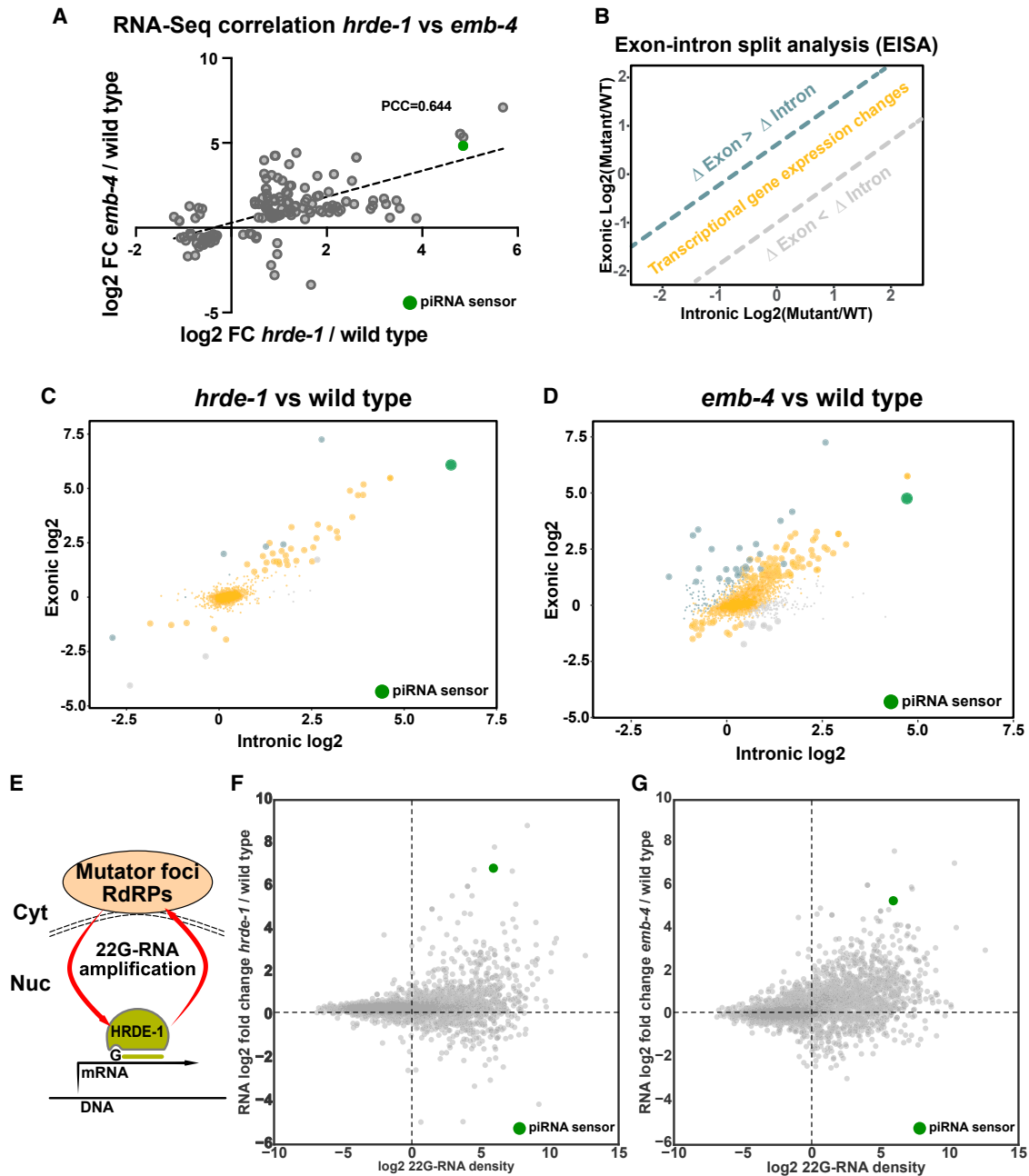
pendent transgene targeted by the piRNA pathway (*ccSi1504*, gift from C. Frøkjær-Jensen). *ccSi1504* is integrated on chromosome V and expressed under *smu-1* promoter, utilizes *smu-1* introns, and contains SV40 and EGL-13 nuclear localization signals. Both *emb-4* and *hrde-1* mutants desilenced the *ccSi1504* transgene similarly to the *mjls144* transgene (Figures S3B and S3C). Our results demonstrate that EMB-4 is an essential factor for piRNA-mediated coTGS in the germline.

### A Functional RNA Helicase Domain of Aquarius/EMB-4 Is Required to Establish Co-transcriptional Gene Silencing

The molecular function of Aquarius/EMB-4 *in vivo* remains unknown in any organism. However, previous studies identified Aquarius as intron binding and associated with spliceosome and EJC recruitment (Hirose et al., 2006; Ideue et al., 2007). *In vitro* experiments suggested a role for the ATPase activity of the RecA1 helicase domain of human Aquarius/AQR in RNA unwinding and spliceosomal complex assembly (De et al., 2015). We noted that the previously isolated allele *sa44* of Aquarius/EMB-4 in *C. elegans* is a missense mutation that induces a G884R substitution in the RecA1 helicase domain (Figure 3A) (Katic and Greenwald, 2006). *emb-4(sa44)* was identified in a genetic suppressor screen and lacks the embryonic lethality associated with null mutants of *emb-4* (Checchi and Kelly, 2006; Katic and Greenwald, 2006). Based on the structural alignment between human Aquarius/AQR, UPF1, and *C. elegans* EMB-4 (Figures S4A–S4C), G884R likely affects RNA binding and, thus, helicase activity (Figures 3A and S4D), without affecting the overall protein levels of Aquarius/EMB-4 (Figure S4E) (Chakrabarti et al., 2011; Cheng et al., 2007; De et al., 2015). Interestingly, *emb-4(sa44)* animals desilence the piRNA sensor similarly to *emb-4(null)* mutations (Figure 3B). Thus, Aquarius/EMB-4 helicase activity is required for co-transcriptional gene silencing. We next asked whether Aquarius/EMB-4 is required for the establishment and/or the maintenance of co-transcriptional gene silencing. Taking advantage of the *emb-4(sa44)* mutants, we addressed this using genetic crosses (Figures 3C and 3D). When *hrde-1(tm1200);mjls144* (piRNA sensor) animals with a desilenced piRNA sensor are crossed to *mjls144* animals generating F1 animals heterozygous for *hrde-1*, piRNA sensor silencing is restored (Figure 3C). However, when carrying out an analogous cross of *hrde-1(tm1200);mjls144;emb-4(sa44)* and *mjls144;emb-4(sa44)* animals, piRNA sensor silencing is not restored (Figure 3D). We conclude that Aquarius/EMB-4 is required for the establishment of co-transcriptional gene silencing.

### Aquarius/EMB-4 Is Required for sRNA-Dependent Co-transcriptional Silencing of Endogenous Genes

Next, we aimed to understand the role of Aquarius/EMB-4 on the endogenous transcriptome. We generated total RNA and sRNA expression profiles using high-throughput sequencing of wild-type animals and *hrde-1* and *emb-4* mutant animals. In our RNA sequencing (RNA-seq) data, when only the genes that show significant expression change in both *hrde-1* and *emb-4* mutants are considered, the majority of these genes show upregulation as opposed to downregulation and there is a significant correlation in gene expression levels between *emb-4* and *hrde-1*



**Figure 4. EMB-4 and HRDE-1 Are Required for Transcriptional Silencing of Genes and Transposable Elements**

(A) Log<sub>2</sub> fold change values of genes, which show significant expression change (adjusted  $p \leq 0.05$ ) in both *hrde-1* and *emb-4* mutants (151 genes in total), are plotted against each mutant background (dashed line, linear fit curve; PCC, Pearson's correlation coefficient; green dot, piRNA sensor transgene).

(B) Exon-intron split analysis (EISA) for comparison of transcriptional and post-transcriptional gene expression changes.

(C and D) Transcriptional and post-transcriptional gene expression changes in *hrde-1* (C) and *emb-4* (D) are colored as in (B). Significant gene expression changes are highlighted by larger dot size (large dots denote mRNA log<sub>2</sub> fold change  $\geq 1$ ,  $p \leq 0.05$ ). piRNA sensor transgene is highlighted by the green dot.

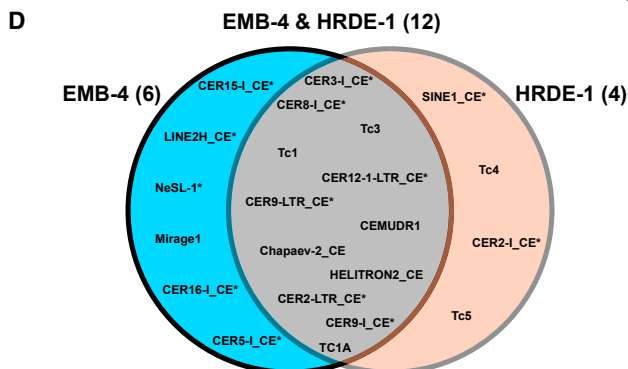
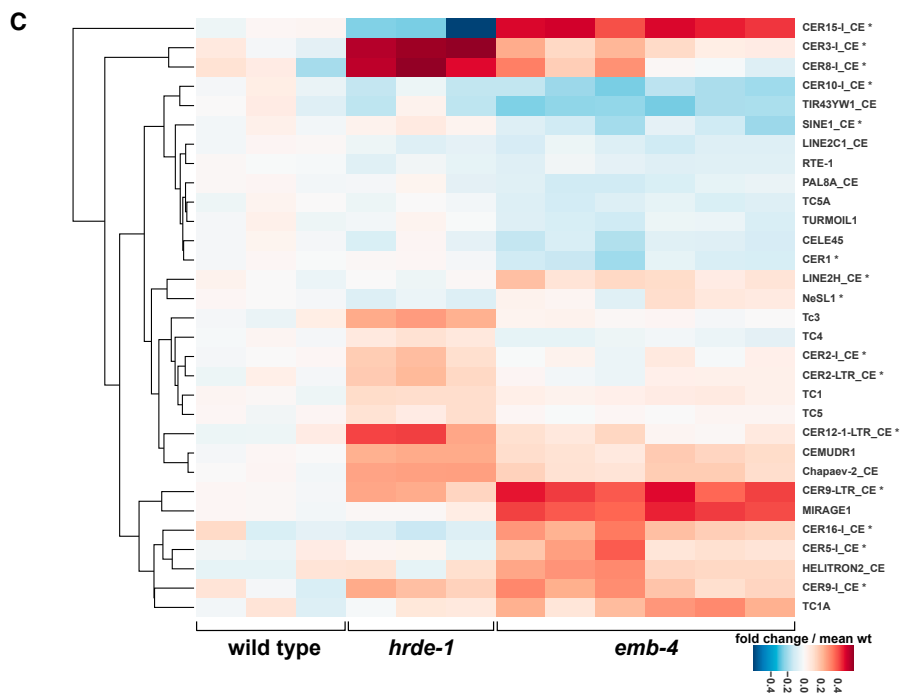
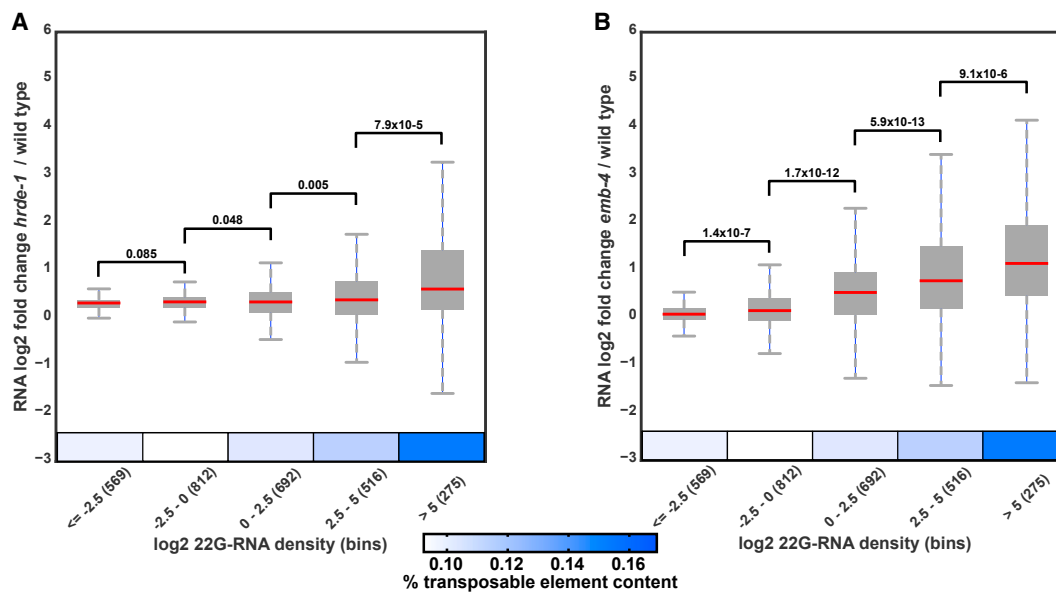
(E) Model showing 22G-RNA amplification in mutator foci, which requires HRDE-1 for 22G-RNA transport and stability.

(F and G) Comparison of RNA log<sub>2</sub> fold change in *hrde-1* and *emb-4* mutants with log<sub>2</sub> 22G-RNA density in wild-type animals (22G-RNA density = 22G-RNA count in HRDE-1 IP [wild-type]/RNA reads per kilobase per million mapped reads [wild-type]). piRNA sensor transgene is highlighted by the green dot.

See also Figure S5.

mutants (Figure 4A and Table S2, PCC = 0.644,  $p < 0.0001$ ). Next, we took advantage of the fact that total RNA-seq contains intronic sequence reads in addition to exonic sequence reads.

Considering intronic reads as a measure of nascent RNA transcription, one can therefore infer changes of transcription rate in addition to steady-state mRNA levels. Such a method was



(legend on next page)



recently reported as exon-intron split analysis (EISA) (Figure 4B) (Gaidatzis et al., 2015). Plotting the logarithm of the ratio of intronic reads from wild-type and mutant samples against the logarithm of the ratio of exonic reads of the same datasets, one can infer the type of gene expression change between the samples; that is, genes that vary in their transcription rate are aligned on the diagonal (yellow, Figures 4B–4D), positive post-transcriptional regulation is above the diagonal (blue, Figures 4B–4D), and negative post-transcriptional regulation is below the diagonal (gray, Figures 4B–4D). Comparing three biological replicates each of wild-type animals, *hrde-1*, and *emb-4* mutants (three replicates each of the *qm31* and *hc60* alleles, combined), we found that the majority of gene expression changes in *hrde-1* and *emb-4* mutants are due to changes in transcription rate (Figures 4C and 4D, percent transcriptional changes: 83% in *hrde-1* and 73% in *emb-4*; transcriptionally up/down: 34/5 in *hrde-1*, 100/16 in *emb-4*). Similarly, the piRNA sensor showed increased transcription rates in both *hrde-1* and *emb-4* mutants (Figures 4C and 4D, green dot). These findings are consistent with the known role of HRDE-1 in co-transcriptional gene silencing (Ashe et al., 2012; Buckley et al., 2012). We conclude that EMB-4 similarly affects transcription rate, which is consistent with a model of EMB-4 and HRDE-1 acting together in co-transcriptional gene silencing. In addition, we conclude that EMB-4 has no major impact on pre-mRNA splicing, as only a few genes showed a relative increase in intronic reads (Figure 4D, gray dots).

Furthermore, we considered sRNA expression alongside total RNA expression to focus on the direct targets of co-transcriptional gene silencing. As expected, *hrde-1* and *emb-4* mutants do not show any change in piRNA population (21U-RNAs), indicating that the upstream piRNA pathway remains intact in these mutant backgrounds (Figure S5A). HRDE-1-bound sRNAs are 22G-RNAs, antisense to their target RNAs and generated by RdRPs at perinuclear foci called mutator bodies (Figure 4E). HRDE-1 is required for the stability and amplification of 22G-RNAs in the germline (Ashe et al., 2012; Buckley et al., 2012; Sapetschnig et al., 2015). We therefore calculated the 22G-RNA density for genes with a matching 22G-RNA in HRDE-1 IPs, as the ratio of HRDE-1-bound 22G-RNAs in the wild-type and the expression levels of their target RNAs (Gerson-Gurwitz et al., 2016; Sapetschnig et al., 2015). We found that genes with high 22G-RNA density in wild-type animals tended to be upregulated in both *hrde-1* and *emb-4* mutants (Figures 4F and 4G).

### Aquarius/EMB-4 Is Required for Silencing of Transposable Elements

For further analysis we grouped all genes from Figures 4F and 4G into bins of increasing 22G-RNA density (Figures 5A and 5B, bins 1–5, 275–569 genes per bin) and found a number of correlating features. HRDE-1-bound 22G-RNAs are enriched in germline

22G-RNA targets and depleted from somatic 22G-RNA targets, as expected (Figure S5B) (Gu et al., 2009). In addition, 22G-RNA density correlates positively with piRNA targets and targets of the WAGO-1 pathway, which is known to overlap with HRDE-1 targets (Figure S5B) (Lee et al., 2012). Finally, ERGO-1, ALG-3/4, and CSR-1 Argonautes are all required for additional endogenous RNAi pathways that also produce 22G-RNAs (Claycomb et al., 2009; Conine et al., 2010; Vasale et al., 2010). We found that HRDE-1-bound 22G-RNAs were overall depleted from ERGO-1 and ALG-3/4 targets, and the density of HRDE-1-bound 22G-RNAs inversely correlated with CSR-1 targets (Figure S5B).

HRDE-1-bound 22G-RNAs are known to target transposable elements (Ni et al., 2014). Indeed, genes with the highest 22G-RNA density also have the highest TE content and show the strongest mRNA upregulation both in *hrde-1* and *emb-4* mutant animals (Figures 5A and 5B). We therefore asked whether EMB-4, like HRDE-1, regulates transposable element expression. When grouping TEs into families, we found that several DNA transposons and retro-elements were overexpressed in *hrde-1* and/or *emb-4* mutants (Figure 5C). Out of 22 TE families that show upregulation, 12 are co-regulated by EMB-4 and HRDE-1, whereas 6 are upregulated specifically in *emb-4* mutants and 4 are upregulated specifically in *hrde-1* mutants (Figure 5D). For instance, CER9-I\_CE, CER9-LTR\_CE, CEMUDR1, and Chapaev-2 are strongly induced in both *hrde-1* and *emb-4* mutants. Furthermore, this RNA induction is accompanied by significant loss of 22G-RNAs to the same locus in both mutant backgrounds (Figures 6A–6C). On the other hand, *bath-45* is a piRNA pathway target gene with multiple piRNA target sites on all of its exons. *bath-45* mRNA levels were upregulated in *hrde-1* but not in *emb-4* mutant animals (Figure 6D). Similarly two TEs, CER15-I\_CE and Mirage1, showed strong upregulation only in *emb-4* mutants (Figures 6E and 6F). Interestingly, CER15-I\_CE is also upregulated in *hrde-1* mutants when the animals are subjected to heat stress (Ni et al., 2016). Altogether, we conclude that HRDE-1 and Aquarius/EMB-4 act together to silence many endogenous 22G-RNA target loci, genes, and transposable elements.

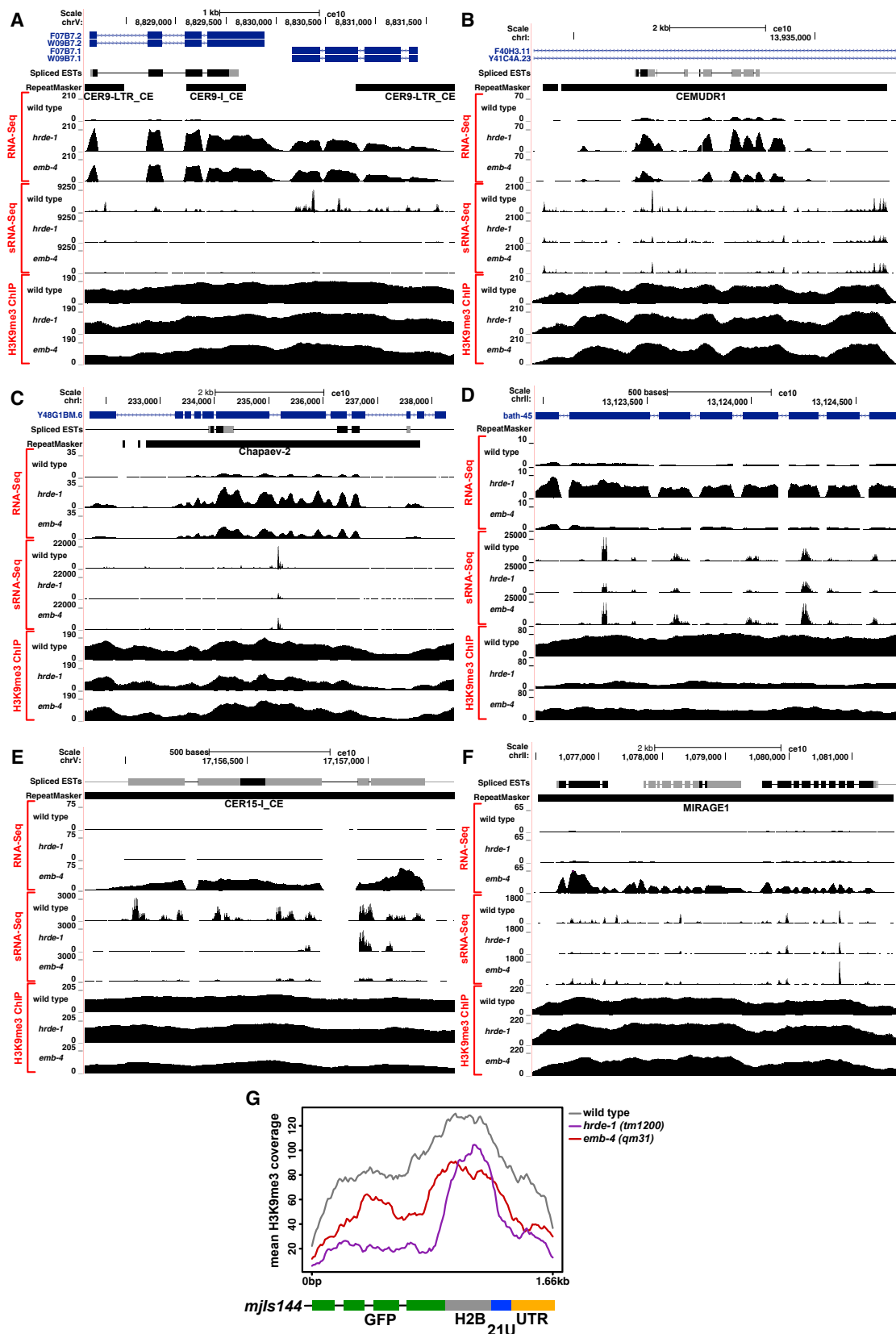
Chromatin-level silencing by histone modifications such as histone H3 lysine 9 trimethylation (H3K9me3) on the target loci has been proposed as a mechanism to maintain long-term silencing of sRNA targets in animals (Holoch and Moazed, 2015). In *C. elegans*, mutations in the histone methyltransferases *set-25* and *set-32* are sufficient to abolish the silencing of the piRNA sensor transgene (Ashe et al., 2012). We performed genome-wide H3K9me3 profiling by establishing the chipmentation method in *C. elegans* (Schmidl et al., 2015). Triplicate analysis of wild-type, *hrde-1*, and *emb-4* mutant animals show that H3K9me3 response is highly loci specific. For instance, H3K9me3 levels on the piRNA sensor, particularly on the GFP sequence, show strong reduction in *hrde-1* and *emb-4* mutant

### Figure 5. HRDE-1 and EMB-4 Are Required for the Suppression of Multiple Transposable Element Families

(A and B) Genes in Figures 4F and 4G are grouped into bins of increasing 22G-RNA density as shown on the x axis (number of genes in each bin is shown in parentheses, boxes show the lower and upper quartiles, red line shows the median, and p values of two-sample t test are shown above the box plots). Heatmap shows percent abundance of transposable elements in each 22G-RNA bin.

(C) Heatmap showing RNA fold change of transposable element families in wild-type, *hrde-1*, and *emb-4* mutants compared with mean wild-type levels. Asterisks indicate retro-element families.

(D) Venn diagram summarizing the overlap of upregulated transposable element families in *hrde-1* and *emb-4* mutant animals as shown in (C).



(legend on next page)

animals, which also correlate with the level of GFP mRNA upregulation in each mutant background (Figures 6G and 2C). Similarly, H3K9me3 levels are reduced on *bath-45*, *Chapaev-2*, and *CER15-I\_CE* in mutant backgrounds that show mRNA upregulation in these loci (Figures 6C–6E). In contrast, H3K9me3 levels do not show much change on *CER9-I\_CE*, *CEMUDR1*, or *Mirage1* loci. These observations are in line with recent studies showing that methyltransferase mutants affect certain *C. elegans* sRNA pathways more than others (McMurphy et al., 2017; Zeller et al., 2016) and that the correlation between H3K9me3 levels and gene silencing can be highly gene and loci specific (Kalinava et al., 2017; Lev et al., 2017).

### Aquarius/EMB-4 Is Specifically Required to Silence Spliced Transcripts

As Aquarius is known to bind introns and is required for RNP remodeling during spliceosome and EJC assembly, we wondered whether Aquarius/EMB-4 is specifically required to allow co-transcriptional gene silencing on nascent transcripts undergoing splicing. To test this we took advantage of the piRNA sensor, which requires HRDE-1 and EMB-4 for co-transcriptional gene silencing (Figures 2, 3, and 4). We previously characterized 22G-RNA populations for the piRNA sensor in detail (Sapetschnig et al., 2015): upon initial piRNA targeting of the sensor transgene in the 3' UTR, 22G-RNAs are generated proximal to the piRNA target site that are independent of HRDE-1 and the nuclear RNAi machinery and are not sufficient for piRNA silencing (Figure 7A, wild-type [blue bars]). Subsequently, HRDE-1-dependent 22G-RNAs spread along the whole length of the transcript, including the *GFP* coding region, to induce co-transcriptional gene silencing (Figure 7A, wild-type [green bars]). We find that loss of HRDE-1 or Aquarius/EMB-4 resulted in the loss of the majority of 22G-RNAs mapping specifically to the coding region of the piRNA sensor, consistent with HRDE-1 and Aquarius/EMB-4 acting together in nuclear RNAi (Figure 7A, *hrde-1* and *emb-4* [green bars]).

The piRNA sensor transgene (*mjls144*) has three synthetic introns within the *GFP* gene sequence. These same introns are generally used by the community to promote efficient transgene expression in *C. elegans*. To test whether Aquarius/EMB-4 is specifically required for co-transcriptional gene silencing of transcripts undergoing splicing, we removed two of the three introns from the piRNA sensor transgene (*mjls144*) to generate a new single-intron piRNA sensor integrated into the same chromosomal location (*mjls588*) (Figure 7B). We retained the single remaining intron, as an intronless transgene is unlikely to be expressed in *C. elegans* (Fire et al., 1990). As expected, the single-intron piRNA sensor (*mjls588*) was completely silenced in wild-type animals and fully desilenced in *hrde-1* mutant animals (Figures 7B–7D). In contrast, while the three-intron piRNA sensor required Aquarius/EMB-4 for silencing, the single-intron piRNA sensor did not (Figures 7B–7D). Importantly, GFP expression levels of the three-intron piRNA sensor and the single-intron

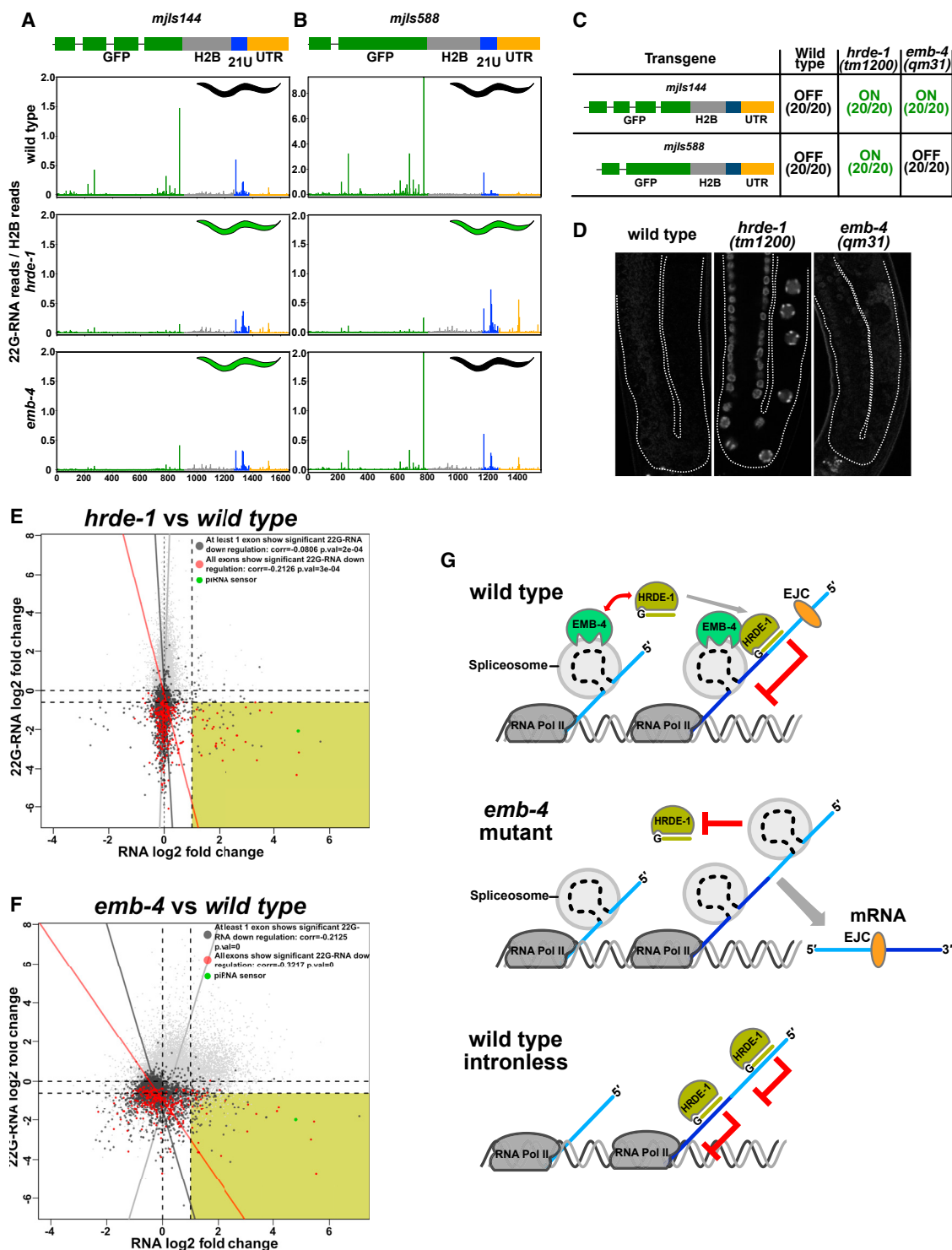
piRNA sensor were similar in *hrde-1* mutants, confirming that the difference in Aquarius/EMB-4 dependence was not simply due to different transcription rates (Figure 7D). Comparing 22G-RNA populations for both piRNA sensors we found that the single-intron piRNA sensor accumulated 5- to 6-fold more 22G-RNAs compared with the three-intron piRNA sensor in wild-type animals (Figures 7A and 7B, wild-type panels: y axis range is different between graphs), indicating that introns can potentially form a barrier to nuclear RNAi by limiting 22G-RNA levels. On the other hand, most endogenous *C. elegans* genes have, on average, four introns (Michael and Manyuan, 1999), and intronic sequences are also present in transposable elements (Sijen and Plasterk, 2003) (Figure 6, spliced expressed sequence tags [ESTs] track). Nuclear RNAi machinery needs to overcome the intronic barriers for efficient and complete silencing of target sequences. Aquarius/EMB-4 can help overcome the intronic barriers to silencing and lead to spreading of 22G-RNAs along transcripts. Indeed, we observed that gene desilencing in both *hrde-1* and *emb-4* mutants correlated with 22G-RNA depletion (Figures 7E and 7F, yellow quadrant). Furthermore, the negative correlation between 22G-RNA abundance and mRNA levels is stronger when all exons of a gene are targeted by 22G-RNAs (Figures 7E and 7F, red line [all exons targeted] versus gray line [single exon targeted]) when considering *hrde-1* and *emb-4* mutant animals.

Taken together, our results support a model in which the RNA helicase Aquarius/EMB-4 is required to provide the co-transcriptional silencing complex access to nascent transcripts undergoing splicing (Figure 7G). We conclude that pre-mRNA processing is a natural and powerful barrier to co-transcriptional gene silencing.

## DISCUSSION

Here we have identified the spliceosomal RNA helicase Aquarius/EMB-4 as an interactor of the germline nuclear Argonaute HRDE-1, which is required for heritable co-transcriptional gene silencing. We show that in *C. elegans*, Aquarius/EMB-4 expression is enriched in germ cell nuclei in adult-stage animals, which coincides with peak HRDE-1 expression in the germline. Our genetic experiments demonstrate that Aquarius/EMB-4 is required during the initiation step of silencing by HRDE-1 and that mutations in the helicase domain of Aquarius/EMB-4 are sufficient to impair its function in transcriptional gene silencing. Loss of Aquarius/EMB-4 leads to transcriptional desilencing of otherwise silenced genes, transposable elements, and piRNA sensor transgenes. We demonstrate that Aquarius/EMB-4 is specifically required to silence nascent transcripts undergoing splicing. In conclusion, we define a key interaction between sRNA pathways and the general transcription machinery. Aquarius/EMB-4 acts as a gatekeeper that permits sRNA pathways to monitor the nascent transcriptome. Interestingly, in a related article Claycomb's group reports that Aquarius/EMB-4

**Figure 6. Exemplary Upregulated Regions with mRNA, 22G-RNA, and H3K9me3 Chromatin IP Profiles in *hrde-1* and *emb-4* Mutants**  
(A–F) *CER9-LTR\_CE/CER9-I\_CE* retro-elements (A), *CEMUDR1* transposable element (B), *Chapaev-2* transposable element (C), piRNA target gene *bath-45* (D), *CER15-I\_CE* retro-element (E), and *Mirage1* transposable element (F). Repeatmasker tracks show transposable elements, and spliced EST tracks show evidence for canonical introns (UCSC genome browser). (G) H3K9me3 chromatin IP profile of piRNA sensor transgene.



**Figure 7. Aquarius/EMB-4 Acts to Remove Intronic Barriers to Transcriptional Gene Silencing**

(A and B) 22G-RNA profiles of the three-intron (A) and the single-intron (B) piRNA sensors in wild-type, *hrde-1*, and *emb-4* mutant animals (mean 22G-RNA abundance of three replicates) (y axis scale in wild-type animals of B is different from that of A; colors indicate different regions of the sensor transgene).

(C) Silencing of the three-intron piRNA sensor (*mjls144*) and the single-intron piRNA sensor (*mjls588*) in wild-type and mutant animals (20 animals assayed for each condition).

(D) Fluorescent microscope images of animals with the single-intron piRNA sensor (*mjls588*). Germline boundaries are marked by the white dotted lines.

(legend continued on next page)

can also act together with the CSR-1 Argonaute protein in *C. elegans* (Tyc et al., 2017).

Previously, other factors involved in the nuclear sRNA-mediated transcriptional gene silencing have been identified in *C. elegans* and *D. melanogaster* through genetic screens (Ashe et al., 2012; Czech et al., 2013; Handler et al., 2013; Lee et al., 2012). These factors include the NRDE factors in *C. elegans* that are required for nuclear RNAi by an as yet unknown mechanism (Guang et al., 2008, 2010), as well as several histone methyltransferases and chromatin factors common to both nematodes and flies (Ashe et al., 2012; Shirayama et al., 2012). Much less is known of the events preceding chromatin level changes. In *C. elegans*, HRDE-1 and other nuclear RNAi factors not only are required for transcriptional gene silencing but also affect the biogenesis and/or the stability of 22G-RNAs, which are the effector sRNAs for nuclear gene silencing (Sapetschnig et al., 2015). One possible explanation is that in the absence of these factors, interactions between the nuclear Argonaute HRDE-1 and its target RNAs are abolished, leading to 22G-RNA destabilization. It is known that for the microRNA pathway, such Argonaute-target interactions are required to stabilize the associated microRNAs (Chatterjee and Grosshans, 2009; Chatterjee et al., 2011). Considering that numerous factors are required for effective HRDE-1 function, it is possible that RNA itself harbors intrinsic features refractory to HRDE-1 targeting.

### Intronic Barriers to Transcriptional Gene Silencing

Mammalian Aquarius binds to introns at a certain distance from the intron branch point, in a sequence-independent manner (Hirase et al., 2006). Even though the details of such interactions have not been studied in *C. elegans*, EMB-4 has been shown to bind nascent transcripts (Shiimori et al., 2012). Thus, it is reasonable to hypothesize that introns can influence the function of Aquarius/EMB-4 during transcriptional gene silencing. Indeed, when we reduced the number of introns from three to one in our piRNA sensor transgene, the absence of Aquarius/EMB-4 no longer resulted in desilencing of the transgene, showing that Aquarius/EMB-4 needs intronic sequences to function. Together with our observation that the single intron piRNA sensor accumulates several-fold more 22G-RNAs compared with a three-intron piRNA sensor, we propose that introns and/or factors interacting with introns are the inhibitory signals for HRDE-1 silencing.

### Transcription, Introns, and Silencing

Our results add to the growing body of evidence implicating co-transcriptional processes in sRNA-mediated gene silencing. In *S. pombe* an sRNA-mediated coTGS mechanism is required for the silencing of centromeric repeats (Allshire and Ekwall, 2015; Martienssen and Moazed, 2015). Several factors that interact with nascent transcripts and non-essential splicing factors are required for efficient coTGS (Bayne et al., 2008, 2014). In addition, factors influencing the efficiency of RNA polymerase elongation are also important in coTGS mechanism in

*S. pombe* (Kowalik et al., 2015). In plants, similar to our results, intron-containing transgenes are protected from nuclear RNAi pathway in comparison with the strongly silenced intronless transgenes (Christie et al., 2011). One hypothesis is that introns or factors interacting with introns are capable of blocking the nuclear RdRPs found in yeast and plants (Dumesic and Madhani, 2013; Vermeersch et al., 2010). In contrast, in the pathogenic yeast *Cryptococcus neoformans* unspliced introns are a signal for sRNA biogenesis and silencing of genes and transposable elements (Dumesic et al., 2013). Clearly, coTGS mechanisms in different organisms have evolved, one way or the other, to incorporate co-transcriptional processes for the regulation of gene silencing. Unlike in yeast and plants, animals do not possess nuclear RdRPs, and although *C. elegans* relies on RdRPs for sRNA amplification, this process occurs in the cytoplasm (Phillips et al., 2012). Nuclear processes that lead to coTGS in nematodes, flies, and mammals instead rely on the transport of sRNAs from cytoplasm to the nucleus by Argonaute proteins. Our results show that in animal coTGS pathways introns can pose a barrier to transcriptional gene silencing, and we provide evidence that the conserved spliceosomal helicase Aquarius/EMB-4 is required to remove these inhibitory signals.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - SILAC Proteomics
  - LC-MS/MS and Analysis of Spectra
  - HRDE-1 SILAC IP Experimental Design, Statistical Rationale, and Data Analysis
  - HRDE-1/EMB-4 Co-immunoprecipitations and Western Blot Analysis
  - qRT-PCR Analysis of GFP Expression
  - qRT-PCR Analysis of *emb-4* Expression
  - Western Blot Analysis of EMB-4 Expression
  - Immunostaining of *C. elegans* Gonads and Embryos
  - RNA Sequencing
  - RNA Sequence Analysis
  - Histone H3K9me3 Chipmentation
  - Amplification and Sequencing of Chipmentation Libraries
  - Bioinformatic Analysis of H3K9me3 Chipmentation Data
  - Transgenic Animals
  - Structural Alignment of AQR and EMB-4
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

(E and F) mRNA expression levels correlate negatively with 22G-RNA abundance and correlation increases when more exons are targeted by 22G-RNAs (gray line, all genes; black line, genes that show significant 22G-RNA change in one exon; red line, genes that show significant 22G-RNA change in all exons). The x axis shows log<sub>2</sub> fold change of 22G-RNAs in mutants/wild-type and the y axis shows log<sub>2</sub> fold change of mRNA in mutants/wild-type (correlation coefficient and p values are shown on graphs).

(G) Model for Aquarius/EMB-4 function in intron-dependent transcriptional gene silencing.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.devcel.2017.07.002>.

## AUTHOR CONTRIBUTIONS

A.A. and E.A.M. conceived and designed the study and wrote the manuscript. A.A., K.M.S., A.N., M.L., R.M., C.J.W., and A.C.B. performed the experiments. A.A., T.d.D., G.E.P., M.L., A.C.B., X.Z., P.M., J.E., and K.L.M.R. analyzed the data. A.A., R.M., A.N., and J.M.C. generated reagents. M.H., J.M.C., P.M., A.I.L., and E.A.M. provided expertise and feedback.

## ACKNOWLEDGMENTS

We thank Kay Harnish and Sylviane Moss for high-throughput sequencing support, George Allen and Charles Bradshaw for core computing support, Nicola Lawrence and imaging facility for high-resolution microscopy support, and Media Kitchen for their support in providing reagents and media. A.C.B. was supported by an HFSP grant to E.A.M. (RPG0014/2015). This work was supported by Cancer Research UK (C13474/A18583, C6946/A14492), the Wellcome Trust (104640/Z/14/Z, 092096/Z/10/Z), and The European Research Council (ERC, grant 260688). The work of P.M. and X.Z. is supported by NIH grant R01GM113242 and NIH grant R01GM122080. R.M. was a Commonwealth Scholar, funded by the UK Government. J.M.C., A.N., and C.J.W. were supported by the CIHR (MOP-274660) and the Canada Research Chairs Program. A.I.L. was supported by a Wellcome Trust Programme Grant (108058/Z/15/Z) and M.L. was supported by 2013/RSE/SCOTGOV/MARIECURIE.

Received: December 21, 2016

Revised: May 14, 2017

Accepted: July 5, 2017

Published: August 7, 2017

## SUPPORTING CITATIONS

The following references appear in the Supplemental Information: Frøkjær-Jensen et al. (2016); Han et al. (2009).

## REFERENCES

- Akay, A., Craig, A., Lehrbach, N., Larance, M., Pourkarimi, E., Wright, J.E., Lamond, A., Miska, E., and Gartner, A. (2013). RNA-binding protein GLD-1/quaking genetically interacts with the mir-35 and the let-7 miRNA pathways in *Caenorhabditis elegans*. *Open Biol.* 3, 130151.
- de Albuquerque, B.F.M., Placentino, M., and Ketting, R.F. (2015). Maternal piRNAs are essential for germline development following *de novo* establishment of endo-siRNAs in *Caenorhabditis elegans*. *Dev. Cell* 34, 448–456.
- Allshire, R.C., and Ekwall, K. (2015). Epigenetic regulation of chromatin states in *Schizosaccharomyces pombe*. *Cold Spring Harb. Perspect. Biol.* 7, a018770.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11, R106.
- 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.
- 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.
- 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.
- Bayne, E.H., Portoso, M., Kagansky, A., Kos-Braun, I.C., Urano, T., Ekwall, K., Alves, F., Rappsilber, J., and Allshire, R.C. (2008). Splicing factors facilitate RNAi-directed silencing in fission yeast. *Science* 322, 602–606.
- Bayne, E.H., Bijos, D.A., White, S.A., de Lima Alves, F., Rappsilber, J., and Allshire, R.C. (2014). A systematic genetic screen identifies new factors influencing centromeric heterochromatin integrity in fission yeast. *Genome Biol.* 15, 481.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* 57, 289–300.
- Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128, 1089–1103.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94.
- 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.
- Burkhart, K.B., Guang, S., Buckley, B.A., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. *PLoS Genet.* 7, e1002249.
- Carrillo Oesterreich, F., Herzl, L., Straube, K., Hujer, K., Howard, J., and Neugebauer, K.M. (2016). Splicing of nascent RNA coincides with intron exit from RNA polymerase II. *Cell* 165, 372–381.
- Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell* 41, 693–703.
- Chatterjee, S., and Grosshans, H. (2009). Active turnover modulates mature microRNA activity in *Caenorhabditis elegans*. *Nature* 461, 546–549.
- Chatterjee, S., Fasler, M., Büssing, I., and Grosshans, H. (2011). Target-mediated protection of endogenous microRNAs in *C. elegans*. *Dev. Cell* 20, 388–396.
- Checchi, P.M., and Kelly, W.G. (2006). emb-4 is a conserved gene required for efficient germline-specific chromatin remodeling during *Caenorhabditis elegans* embryogenesis. *Genetics* 174, 1895–1906.
- Cheng, Z., Muhrad, D., Lim, M.K., Parker, R., and Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. *EMBO J.* 26, 253–264.
- Christie, M., Croft, L.J., and Carroll, B.J. (2011). Intron splicing suppresses RNA silencing in *Arabidopsis*. *Plant J.* 68, 159–167.
- 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.
- Cowley, M., and Oakey, R.J. (2013). Transposable elements re-wire and fine-tune the transcriptome. *PLoS Genet.* 9, e1003234.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p/b-value range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372.
- Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by Piwi are essential for stem cell self-renewal. *Genes Dev.* 12, 3715–3727.
- Czech, B., Preall, J.B., McGinn, J., and Hannon, G.J. (2013). A transcriptome-wide RNAi screen in the *Drosophila* ovary reveals factors of the germline piRNA pathway. *Mol. Cell* 50, 1–13.
- 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.
- De, I., Bessonov, S., Hofele, R., dos Santos, K., Will, C.L., Urlaub, H., Lührmann, R., and Pena, V. (2015). The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nat. Struct. Mol. Biol.* 22, 138–144.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Drinnenberg, I.A., Fink, G.R., and Bartel, D.P. (2011). Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* 333, 1592.
- Dumesic, P.A., and Madhani, H.D. (2013). The spliceosome as a transposon sensor. *RNA Biol.* 10, 1653–1660.
- Dumesic, P.A., Natarajan, P., Chen, C., Drinnenberg, I.A., Schiller, B.J., Thompson, J., Moresco, J.J., Yates, J.R., 3rd, Bartel, D.P., and Madhani, H.D. (2013). Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* 152, 957–968.
- Fire, A., Harrison, S.W., and Dixon, D. (1990). A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* 93, 189–198.
- Fire, A., Xu, S., 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.
- Fredens, J., Engholm-Keller, K., Giessing, A., Pultz, D., Larsen, M.R., Højrup, P., Møller-Jensen, J., and Færgeman, N.J. (2011). Quantitative proteomics by amino acid labeling in *C. elegans*. *Nat. Methods* 8, 845–847.
- Frøkjær-Jensen, C., Christian, F.-J., Wayne Davis, M., Hopkins, C.E., Newman, B.J., Thummel, J.M., Søren-Peter, O., Morten, G., and Jørgensen, E.M. (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* 40, 1375–1383.
- Frøkjær-Jensen, C., Jain, N., Hansen, L., Davis, M.W., Li, Y., Zhao, D., Rebora, K., Millet, J.R.M., Liu, X., Kim, S.K., et al. (2016). An abundant class of non-coding DNA can prevent stochastic gene silencing in the *C. elegans* germline. *Cell* 166, 343–357.
- Gaidatzis, D., Burger, L., Florescu, M., and Stadler, M.B. (2015). Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nat. Biotechnol.* 33, 722–729.
- 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.
- Gonatosopoulos-Pournatzis, T., and Cowling, V.H. (2014). Cap-binding complex (CBC). *Biochem. J.* 457, 231–242.
- 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.
- 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.
- 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.
- Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 315, 1587–1590.
- 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. USA* 106, 18674–18679.
- Handler, D., Meixner, K., Pizka, M., Lauss, K., Schmied, C., Gruber, F.S., and Brennecke, J. (2013). The genetic makeup of the *Drosophila* piRNA pathway. *Mol. Cell* 50, 762–777.
- Le Hir, H., Saulière, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.* 17, 41–54.
- Hirose, T., Ideue, T., Nagai, M., Hagiwara, M., Shu, M.-D., and Steitz, J.A. (2006). A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol. Cell* 23, 673–684.
- Hoffmann, S., Otto, C., Kurtz, S., Sharma, C.M., Khaitovich, P., Vogel, J., Stadler, P.F., and Hackermüller, J. (2009). Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.* 5, e1000502.
- Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* 16, 71–84.
- Hong, Y., Sonnevile, R., Agostinho, A., Meier, B., Wang, B., Blow, J.J., and Gartner, A. (2016). The SMC-5/6 complex and the HIM-6 (BLM) helicase synergistically promote meiotic recombination intermediate processing and chromosome maturation during *Caenorhabditis elegans* meiosis. *PLoS Genet.* 12, e1005872.
- Ideue, T., Sasaki, Y.T.F., Hagiwara, M., and Hirose, T. (2007). Introns play an essential role in splicing-dependent formation of the exon junction complex. *Genes Dev.* 21, 1993–1998.
- Jobson, M.A., Jordan, J.M., Sandrof, M.A., Hibshman, J.D., Lennox, A.L., and Baugh, L.R. (2015). Transgenerational effects of early life starvation on growth, reproduction, and stress resistance in *Caenorhabditis elegans*. *Genetics* 201, 201–212.
- Kalinava, N., Ni, J.Z., Peterman, K., Chen, E., and Gu, S.G. (2017). Decoupling the downstream effects of germline nuclear RNAi reveals that H3K9me3 is dispensable for heritable RNAi and the maintenance of endogenous siRNA-mediated transcriptional silencing in *Caenorhabditis elegans*. *Epigenetics Chromatin* 10, 6.
- Kammaing, L.M., van Wolfswinkel, J.C., Luteijn, M.J., Kaaij, L.J.T., Bagijn, M.P., Sapetschnig, A., Miska, E.A., Berezikov, E., and Ketting, R.F. (2012). Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of *Caenorhabditis elegans*. *PLoS Genet.* 8, e1002702.
- Katic, I., and Greenwald, I. (2006). EMB-4: a predicted ATPase that facilitates lin-12 activity in *Caenorhabditis elegans*. *Genetics* 174, 1907–1915.
- Kowalik, K.M., Shimada, Y., Flury, V., Stadler, M.B., Batki, J., and Bühler, M. (2015). The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. *Nature* 520, 248–252.
- Larance, M., Bailly, A.P., Pourkarimi, E., Hay, R.T., Buchanan, G., Coulthurst, S., Xirodimas, D.P., Gartner, A., and Lamond, A.I. (2011). Stable-isotope labeling with amino acids in nematodes. *Nat. Methods* 8, 849–851.
- Larance, M., Kirkwood, K.J., Xirodimas, D.P., Lundberg, E., Uhlen, M., and Lamond, A.I. (2012). Characterization of MRFAP1 turnover and interactions downstream of the NEDD8 pathway. *Mol. Cell. Proteomics* 11, M111.014407.
- Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* 150, 78–87.
- Lev, I., Seroussi, U., Gingold, H., Bril, R., Anava, S., and Rechavi, O. (2017). MET-2-dependent H3K9 methylation suppresses transgenerational small RNA inheritance. *Curr. Biol.* 27, 1138–1147.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930.
- 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.
- Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in *Caenorhabditis elegans*. *EMBO J.* 31, 3422–3430.
- Martienssen, R., and Moazed, D. (2015). RNAi and heterochromatin assembly. *Cold Spring Harb. Perspect. Biol.* 7, a019323.

- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12.
- McMurchy, A.N., Stempor, P., Gaarenstroom, T., Wysolmerski, B., Dong, Y., Aussanikava, D., Appert, A., Huang, N., Kolasinska-Zwierz, P., Sapetschnig, A., et al. (2017). A team of heterochromatin factors collaborates with small RNA pathways to combat repetitive elements and germline stress. *Elife* 6.
- Michael, D., and Manyuan, L. (1999). Intron-exon structures of eukaryotic model organisms. *Nucleic Acids Res.* 27, 3219–3228.
- Ni, J.Z., Chen, E., and Gu, S.G. (2014). Complex coding of endogenous siRNA, transcriptional silencing and H3K9 methylation on native targets of germline nuclear RNAi in *C. elegans*. *BMC Genomics* 15, 1157.
- Ni, J.Z., Kalinava, N., Chen, E., Huang, A., Trinh, T., and Gu, S.G. (2016). A transgenerational role of the germline nuclear RNAi pathway in repressing heat stress-induced transcriptional activation in *C. elegans*. *Epigenetics Chromatin* 9, 3.
- Nilsen, T.W. (2003). The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* 25, 1147–1149.
- Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376–386.
- Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. *Genetics* 201, 47–54.
- Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* 315, 241–244.
- Park, S.H., Cheong, C., Idoyaga, J., Kim, J.Y., Choi, J.-H., Do, Y., Lee, H., Jo, J.H., Oh, Y.-S., Im, W., et al. (2008). Generation and application of new rat monoclonal antibodies against synthetic FLAG and OLLAS tags for improved immunodetection. *J. Immunol. Methods* 331, 27–38.
- Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the *C. elegans* germline. *Genes Dev.* 26, 1433–1444.
- Phillips, C.M., Brown, K.C., Montgomery, B.E., Ruvkun, G., and Montgomery, T.A. (2015). piRNAs and piRNA-dependent siRNAs protect conserved and essential *C. elegans* genes from misrouting into the RNAi pathway. *Dev. Cell* 34, 457–465.
- Rebollo, R., Romanish, M.T., and Mager, D.L. (2012). Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* 46, 21–42.
- Rechavi, O., Houry-Ze'evi, 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.
- Sapetschnig, A., Sarkies, P., Lehrbach, N.J., and Miska, E.A. (2015). Tertiary siRNAs mediate paramutation in *C. elegans*. *PLoS Genet.* 11, e1005078.
- Schmidl, C., Rendeiro, A.F., Sheffield, N.C., and Bock, C. (2015). ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* 12, 963–965.
- Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: capping and polyadenylation. *Nat. Struct. Biol.* 7, 838–842.
- Shimori, M., Inoue, K., and Sakamoto, H. (2012). A specific set of exon junction complex subunits is required for the nuclear retention of unspliced RNAs in *Caenorhabditis elegans*. *Mol. Cell. Biol.* 33, 444–456.
- Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* 150, 65–77.
- Sijen, T., and Plasterk, R.H.A. (2003). Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* 426, 310–314.
- Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L.D., Mitchell, J., Sakaguchi, A., Miska, E.A., and Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to *Caenorhabditis elegans* Piwi mutants. *Cell Rep.* 7, 762–773.
- Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8, 272–285.
- Smit, A., Hubley, R., and Green, P. (2015). RepeatMasker Open-4.0. 2013–2015. Institute for Systems Biology. <http://repeatmasker.org>.
- Stempor, P., and Ahringer, J. (2016). SeqPlots - Interactive software for exploratory data analyses, pattern discovery and visualization in genomics. *Wellcome Open Res.* 1, 14.
- Stiernagle, T. (2006). Maintenance of *C. elegans*. *WormBook*, 1–11.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452.
- Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Tóth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* 27, 390–399.
- Tyc, K.M., Nabih, A., Wu, M.Z., Wedeles, C.J., Sobotka, J.A., and Claycomb, J.M. (2017). The conserved intron binding protein EMB-4 plays differential roles in germline small RNA pathways of *C. elegans*. *Dev. Cell* 42, this issue, 256–270.
- Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419.
- UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212.
- 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. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc. Natl. Acad. Sci. USA*.
- Vermeersch, L., De Winne, N., and Depicker, A. (2010). Introns reduce transitivity proportionally to their length, suggesting that silencing spreads along the pre-mRNA. *Plant J.* 64, 392–401.
- Wahl, M.C., and Lührmann, R. (2015). SnapShot: spliceosome dynamics I. *Cell* 161, 1474–e1.
- Wang, G., and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. *Curr. Biol.* 18, 861–867.
- Weick, E.-M., and Miska, E.A. (2014). piRNAs: from biogenesis to function. *Development* 141, 3458–3471.
- Weick, E.-M., Sarkies, P., Silva, N., Chen, R.A., Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E.A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in *C. elegans*. *Genes Dev.* 28, 783–796.
- Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338.
- Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C.C., Pettersen, E.F., Goddard, T.D., Meng, E.C., Sali, A., and Ferrin, T.E. (2012). UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J. Struct. Biol.* 179, 269–278.
- Zeller, P., Padeken, J., van Schendel, R., Kalck, V., Tijsterman, M., and Gasser, S.M. (2016). Histone H3K9 methylation is dispensable for *Caenorhabditis elegans* development but suppresses RNA: DNA hybrid-associated repeat instability. *Nat. Genet.* 48, 1385–1395.



## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Mouse monoclonal anti-EMB-4	This paper	5M19
Rabbit polyclonal anti-HRDE-1	(Ashe et al., 2012)	N/A
Rat monoclonal anti-OLLAS epitope tag	Novusbio	Cat#:NBP1-06713; RRID: AB_1625979
Rat monoclonal anti-alpha tubulin	Accurate Chemical	Cat#:YSRTMCA77P
Mouse monoclonal anti-alpha tubulin	SIGMA	DM1A; cat#: 9026; RRID: AB_477593
Rabbit polyclonal anti-Histone H3K9me3	Abcam	Cat#: ab8898; RRID: AB_306848
Goat anti-rabbit AlexaFluor 568	Molecular Probes	A11011; RRID: AB_143157
Goat anti-mouse AlexaFluor 647	ThermoFisher	A-21235; RRID: AB_141693
<b>Bacterial and Virus Strains</b>		
E. coli HB101	Caenorhabditis Genetics Center	N/A
<b>Critical Commercial Assays</b>		
TruSeq small RNA library preparation kit	Illumina	RS-200
NEBNext Ultra RNA library prep kit	NEB	E7530
Ribo-Zero rRNA Removal Kit	Illumina	MRZH11124
<b>Deposited Data</b>		
SILAC proteomics data of HRDE-1 IPs	This paper	Proteinexchange PXD004416
Small RNA and total RNA sequencing data	This paper	E-MTAB-4877
H3K9me3 ChIP-Seq data	This paper	E-MTAB-5662
HRDE-1 bound 22G-RNA data	(Sapetschnig et al., 2015)	GSE66344
<b>Experimental Models: Organisms/Strains</b>		
<i>C. elegans</i> : Strain SX1316 <i>mjls144 II</i>	(Bagijn et al., 2012)	N/A
<i>C. elegans</i> : Strain SX2000 <i>mjls144 II</i> ; <i>hrde-1(tm1200) III</i>	(Ashe et al., 2012)	N/A
<i>C. elegans</i> : Strain SX2929 <i>mjls144 II</i> ; <i>emb-4(qm31) V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX2930 <i>mjls144 II</i> ; <i>emb-4(hc60) V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3041 <i>mjls144 II</i> ; <i>emb-4(sa44) V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3073 <i>mjls588 II</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3074 <i>mjls144 II</i> ; <i>hrde-1(tm1200) III</i> ; <i>emb-4(sa44) V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3078 <i>mjls588 II</i> ; <i>hrde-1(tm1200) III</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3079 <i>mjls588 II</i> ; <i>emb-4(qm31) V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3117 <i>emb-4(mjSi92)</i>	This paper	N/A
<i>C. elegans</i> : Strain PD1504 <i>ccSi1504 V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3118 <i>hrde-1(tm1200) III ccSi1504 V</i>	This paper	N/A
<i>C. elegans</i> : Strain SX3179 <i>emb-4(qm31) V ccSi1504 V</i>	This paper	N/A
<i>C. elegans</i> : Strain VM285 <i>neSi21</i>	(Shirayama et al., 2012)	N/A
<b>Oligonucleotides</b>		
GFP <i>mjls144</i> RT-PCR Fwd: 5'-TCTGTCTCAGTGGAGAGGGTGA-3'	(Weick et al., 2014)	N/A
GFP <i>mjls144</i> RT-PCR Rev: 5'-TTTAAACTTACCCATGGAACAGG-3'	(Weick et al., 2014)	N/A
GFP <i>mjls588</i> RT-PCR Fwd: 5'-CGTACCATCTTCTCAAG-3'	This paper	N/A
GFP <i>mjls588</i> RT-PCR Rev: 5'-GATGTTCCGTCCTCCTT-3'	This paper	N/A
<i>cgh-1</i> RT-PCR Fwd: 5'-CCACCCAGGAAGAATTCTC-3'	(Weick et al., 2014)	N/A
<i>cgh-1</i> RT-PCR Rev: 5'-GGTAAGTCTCGCGTTTCTT-3'	(Weick et al., 2014)	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>emb-4</i> RT-PCR Fwd: 5'-TTCGTCCCCTGTCCATATC-3'	This paper	N/A
<i>emb-4</i> RT-PCR Rev: 5'-ATCGGCTTCTGGCCTAAAAT-3'	This paper	N/A
<i>act-3</i> RT-PCR Fwd: 5'-CCAAGAGAGGTATCCTTACCCTCAA-3'	This paper	N/A
<i>act-3</i> RT-PCR Rev: 5'-AAGCTCATTGTAGAAGGTGTGATGC-3'	This paper	N/A
Recombinant DNA		
pCFJ1416 (Psmu1:GFP2:smu1UTR) GFP2-GFP with altered codons and smu1 introns	This paper	(gift of Christian Frøkjær-Jensen)
CRISPR gRNA for <i>emb-4</i> N-term 5'-CAAGAAGCCGTGGTACTCG-3'	This paper	Dharmacon
Software and Algorithms		
MaxQuant quantitative proteomics software package	(Cox and Mann, 2008)	<a href="http://www.maxquant.org">http://www.maxquant.org</a> , version 1.3.0.5
RepeatMasker	(Smit et al., 2015)	<a href="http://www.repeatmasker.org/">http://www.repeatmasker.org/</a> , version open-4.0.5
Cutadapt	(Martin, 2011)	Version 1.9.0
STAR aligner	(Dobin et al., 2013)	Version v2.5.1b
SAMtools	(Li et al., 2009)	v1.3
featureCounts	(Liao et al., 2014)	v1.5.0-p1
DESeq2	(Love et al., 2014)	v3.2.2
Segemehl	(Hoffmann et al., 2009)	version 2.0
SeqPlots	(Stempor and Ahringer, 2016)	<a href="http://przemol.github.io/seqplots/">http://przemol.github.io/seqplots/</a>

**CONTACT FOR REAGENT AND RESOURCE SHARING**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Eric A. Miska ([eam29@cam.ac.uk](mailto:eam29@cam.ac.uk)).

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

*C. elegans* were grown under standard conditions at 20°C unless otherwise indicated. The wild-type strain was var. Bristol N2 (Brenner, 1974). *E. coli* strain HB101 was used as the food source on NGM plates (Stiernagle, 2006). Adult *C. elegans* animals were bleached to obtain synchronized L1 larvae population to grow synchronized animals used in experiments. Synchronized young adult animals were used in most experiments unless otherwise stated. All strains used are listed in Key Resources Table. Developmental stage of animals used in experiments are indicated in method details.

**METHOD DETAILS****SILAC Proteomics**

Bacterial and nematode growth conditions for SILAC experiments are previously described (Larance et al., 2011). Heavy (R10K8) labelled wild-type animals and medium (R6K4) labelled *hrde-1* mutant animals were grown to young adult stage, washed 3× with M9 buffer and lysed in native lysis buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% NP40, Roche complete protease inhibitor cocktail) by beat beating using zirconia beads and the PreCelys instrument (6,500 rpm, 3×30 s with 30 s intervals at 4°C). The lysate was kept on ice for 30 min and centrifuged for 10 min at 16,000 rcf at 4°C to remove insoluble material. BCA assay (Thermo Scientific) was used to determine protein concentration of the supernatant. 3-8 mg of total protein has been used for immunoprecipitations (IP) with 3 mg of anti-HRDE-1 antibody coupled dynabeads M270 (20 µg antibody / mg beads) for 1 hr at 4°C. Beads were washed 3× with wash buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl, 0.5 mM EDTA, Roche complete protease inhibitor cocktail) and equal amounts of beads from heavy and medium labelled IPs were mixed together prior to elution at the final wash. Elution was done by heating beads to 70°C for 10 min in LDS loading buffer. The eluted IP was loaded across multiple adjacent lanes (25 µl per lane) on 1 mm, 10-well, 4–12% (w/v) Bis-Tris NuPage gels using MES running buffer according to manufacturer's instructions but with the addition of 25 mM triscarboxyethylphosphine, and 50 mM N-ethylmaleimide in the LDS sample buffer. After electrophoresis at 150 V for 45 min, SYPRO Ruby staining was performed as per manufacturer's instructions (Invitrogen).

Protein bands of interest were excised and destained in 1 ml of 50% acetonitrile and 250 mM ammonium bicarbonate at room temperature for 45 min with shaking. The gel slice was dehydrated by incubation in 1 ml of 100% acetonitrile for 10 min at room

temperature. All solution was carefully removed prior to the addition of 50  $\mu$ l MS-grade trypsin (Promega) (12.5 ng/ $\mu$ l) in 100 mM NH<sub>4</sub>HCO<sub>3</sub> and incubation overnight at 37°C. Peptides were extracted by the addition of 0.1 ml of 5% formic acid and incubation at 37°C for 1 hr. Peptides were further extracted by the addition of 0.1 ml of 100% acetonitrile and incubation at 37°C for 1 hr. The gel slice was completely dehydrated by the addition of 0.5 ml of 100% acetonitrile and incubation at 37 °C for 10 min. The entire supernatant was then vacuum-dried.

### LC-MS/MS and Analysis of Spectra

Using a Thermo Fisher Scientific Ultimate 3000 RSLCnano UHPLC, 15  $\mu$ l of peptides in 5% (vol/vol) formic acid (final volume  $\sim$ 10  $\mu$ l) were injected onto an Acclaim PepMap C18 nano-trap column. After washing with 2% (vol/vol) acetonitrile, 0.1% (vol/vol) formic acid, peptides were resolved on a 50 cm X 75  $\mu$ m C18 EasySpray reverse phase analytical column with integrated emitter over a gradient from 2% acetonitrile to 35% acetonitrile over 220 min with a flow rate of 200 nl/min. The peptides were ionized by electrospray ionization at +2.0 kV. Tandem mass spectrometry analysis was carried out on a Q-Exactive mass spectrometer (Thermo Fisher Scientific) using HCD fragmentation. The data-dependent acquisition method used acquired MS/MS spectra on the top 30 most abundant ions at any one point during the gradient. The RAW data produced by the mass spectrometer were analysed using the MaxQuant quantitative proteomics software package (Cox and Mann, 2008) (<http://www.maxquant.org>, version 1.3.0.5). The MaxQuant output has also been uploaded to the ProteomeXchange Consortium under the same identifier given above. This version of MaxQuant includes an integrated search engine, Andromeda. Peptide and Protein level identification were both set to a false discovery rate of 1% using a target-decoy based strategy. The database supplied to the search engine for peptide identifications was the combined *C. elegans* and *E. coli* Swissprot and TrEMBL databases downloaded on the 12th July 2012. The mass tolerance was set to 7 ppm for precursor ions and MS/MS mass tolerance was set at 20 ppm. Enzyme was set to trypsin (cleavage C-terminal to lysine and arginine) with up to 2 missed cleavages. Deamidation of Asn and Gln, oxidation of Met, pyro-Glu (with peptide N-term Gln), phosphorylation of Ser/Thr/Tyr, and protein N-terminal acetylation were set as variable modifications. N-ethylmaleimide on Cys was searched as a fixed modification. The output from MaxQuant provided peptide level data as well as protein group level data. We used the protein groups as defined by the Maxquant package (Cox and Mann, 2008).

### HRDE-1 SILAC IP Experimental Design, Statistical Rationale, and Data Analysis

Three biological replicates were performed for SILAC-IP analysis of HRDE-1 and this level of replication was chosen based upon the variance detected in previous experiments using SILAC-IP analysis (Larance et al., 2012). To avoid disregarding low affinity binders, we used a low stringency cutoff such that a protein needed to have a H/M SILAC ratio  $>1$  in two out of three biological replicates in our data analysis with MaxQuant to eliminate non-specific binding proteins. This yields excellent removal of environmental contaminants (keratins, trypsin, antibody, etc.) that do not incorporate stable isotopes.

### HRDE-1/EMB-4 Co-immunoprecipitations and Western Blot Analysis

For HRDE-1 immunoprecipitations (Figures 1E, S1E, and S1F), animals were harvested 24 hours post-L4 stage in 30 mM HEPES, 100 mM potassium acetate, 2 mM magnesium acetate and 10% glycerol (DROSO buffer). To lyse the animals, samples were snap-frozen in liquid nitrogen as droplets and grinded to powder. *C. elegans* powder was then re-suspended in DROSO buffer supplemented with 0.1% NP-40 and further lysed by sonication. Lysates were subsequently cleared by centrifugation. 2 mg of proteins were used per IP at 4 mg/ml. For RNase treatment, lysates were either incubated with RNaseA or buffer for 30min at room temperature prior to addition of antibodies. 10  $\mu$ g antibody (normal IgG: SantaCruz Biotech, sc-2027; anti-HRDE-1: Genomic Antibody Tech, custom (Kamminga et al., 2012)) was added to the lysates and incubated overnight at 4°C with rotation. 30  $\mu$ l of proteinA/G-agarose beads (SantaCruz Biotech, sc-2003) were incubated for 2 hrs the next morning. Immunoprecipitates were washed 4 times with DROSO buffer and boiled with 2 $\times$  sample buffer for 5 min to elute. Samples were then analysed by western blotting (Anti-Ollas: Novusbio, NMP1-06713).

For EMB-4 immunoprecipitations (Figure S1H), 750  $\mu$ l of synchronized gravid adults were dounced using a metal wheaton dounce in DROSO 'complete' buffer (30 mM Hepes, 100 mM potassium Acetate, 2 mM Magnesium Acetate, 0.1% NP-40/Igepal, 2 mM DTT, 1 tablet/5mls Protease inhibitor (Roche), 1:100 Sigma Phosphatase Inhibitor 2, 1:100 Sigma Phosphatase Inhibitor 3.) until the worms and the embryos were no longer visible. Lysate was cleared by centrifugation for 10min at 13,000xg in a pre-cooled centrifuge (4°C). The concentration of the supernatant (total worm protein) was determined by Lowry assay using Bio-rad Lowry assay kit. Each IP was performed from 5 mg protein. Lysate was pre-cleared with 25  $\mu$ l protein A/G agarose bead slurry (Santa Cruz Biotech, beads are equilibrated in DROSO complete buffer prior to use) for one hour at 4°C on a rotator. 5  $\mu$ g (anti-Flag, Sigma Aldrich) or 50  $\mu$ l of EMB-4 antibody (specificity of EMB-4 antibodies are shown in Figure S1E) or buffer alone (no antibody control) was added to each IP sample and incubated for two hours on a rotator at 4°C. Immune complexes were recovered using 50 $\mu$ l of a 50% slurry of Protein-A/G agarose beads (Santa Cruz Biotechnology) and washed 6 $\times$ 5min at 4°C with DROSO buffer. Protein was eluted from beads and denatured by incubation in Thermofisher 2 $\times$  LDS sample buffer for 10min at 70°C. Input samples were prepared from the same lysate at a concentration of 2 $\mu$ g/ $\mu$ l using Thermofisher 2 $\times$ LDS sample buffer and reducing agent. Proteins were resolved by SDS-PAGE on Criterion Precast gradient gels (4-15%, Biorad) and transferred to Hybond-C membrane (Amersham Biosciences). The membrane was incubated overnight at 4°C with either: (i) affinity purified anti-EMB-4 (1:200), or anti-FLAG (Sigma Aldrich, 1:1000) in PBST-5% milk solution (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4, and 5% [w/v] dried milk). The membrane was incubated 2 h at room temperature with anti-mouse HRP-conjugated secondary antibody (Jackson

Immunoresearch) diluted 1:1,000 in PBST and then visualized by Luminata Forte Western HRP substrate. Full blots of all IPs are available in [Figure S6](#).

#### qRT-PCR Analysis of GFP Expression

cDNA was synthesised from purified and DNase treated total RNA (1  $\mu$ g) using the SuperScript II enzyme and random primers as described in the manual. qRT-PCR reactions were performed using Applied Biosystem Power SYBR Green Master Mix using primer sequences to amplify GFP sequence (primer 1: 5'-TCTGTCACTGGAGAGGGTGA-3', primer 2: 5'-TTTAACTTACCCATGGAA CAGG-3') and the endogenous germline control gene *cgh-1* for normalisation (primer 1: 5'-CCACCCCAGGAAGAATTCTC-3', primer 2: 5'-GGTAAGTCTCGGCGTTTCTT-3').

#### qRT-PCR Analysis of *emb-4* Expression

cDNA was generated from 1 mg of *C. elegans* total RNA using random hexamers with Superscript III Reverse Transcriptase (Invitrogen). qRT-PCR was performed using Applied Biosystems SYBR Green PCR Master mix with primers for *emb-4* (primer 1: 5'-TTCG TCCCCTGTTCCATATC-3', primer 2: 5'-ATCGGCTTCTGGCCTAAAAT-3') and for *act-3* (primer 1: 5'-CCAAGAGAGGTATCCTT ACCCTCAA-3', primer 2: 5'-AAGCTCATTGTAGAAGGTGTGATGC-3').

#### Western Blot Analysis of EMB-4 Expression

Proteins were resolved by SDS-PAGE on Criterion Precast gradient gels (4-15%, Biorad). and transferred to Hybond-C membrane (Amersham Biosciences). The membrane was incubated overnight at 4°C with either: (i) affinity purified anti-EMB-4, or anti alpha-tubulin (Accurate Chemical) antibodies diluted to 1:2000, in PBST-5% milk solution (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4, and 5% [w/v] dried milk). The membrane was incubated 1 h at room temperature with HRP-conjugated secondary antibodies (Jackson Immunoresearch) diluted to 1:5,000 in PBST and then visualized by Western Lightning ECL Kit from Perkin Elmer. Images were collected on a LAS-3000 Intelligent Dark-Box (Fujifilm).

#### Immunostaining of *C. elegans* Gonads and Embryos

For [Figures 2C](#) and [S2A](#), gonads were excised from gravid adult worms in 1x PBS on poly-L-lysine coated slides, frozen and cracked on dry ice for longer than 10 minutes, and fixed at -20°C for 5 min in each of the following (15 minutes total) respectively; 100% methanol, 50% methanol/50% acetone, and 100% acetone. All sample incubations were performed in a humid chamber. Samples were washed 2x 5min with 1xPBS, then 2x 5 mins with 1xPBS / 0.1% Tween-20. Samples are then blocked for one hour in 1xPBS/0.1% Tween-20 / 3%BSA (PBST+BSA) at room temperature, and then incubated with primary antibody (1:500) overnight at 4°C. Slides were washed 3x for 10 minutes with PBST, and then incubated for 1 hour in PBST+BSA. Secondary antibodies were from Jackson Immunoresearch and Molecular Probes. Incubations with anti-mouse secondary antibodies were performed for one hour in PBST+BSA at room temperature. Slides were washed 3x for ten minutes in PBST, 3x for 5 minutes in PBS and then incubated with DAPI (1:2500) for 10 minutes at room temperature. Finally, slides were washed in PBS 3x for 5 minutes then mounted in VectaShield (Vector Labs). All images were collected using Nikon Ti-S inverted microscope with NIS Element and AR software.

For [Figure S2B](#), gonad stainings were performed as described in ([Hong et al., 2016](#)). Adult animal gonads were dissected in dissection buffer (25mM HEPES pH7.4, 2mM MgCl<sub>2</sub>, 2mM CaCl<sub>2</sub>, 48mM KCl, 0.12 M NaCl, 0.2% Tween-20, 4mM levamisole) on cover slips and fixed 5min by equal volume of fixation buffer (25mM HEPES pH7.4, 2mM MgCl<sub>2</sub>, 2mM CaCl<sub>2</sub>, 48mM KCl, 0.12 M NaCl, 0.2% Tween-20, 4% formaldehyde). Glass cover slip were placed on poly-L-lysine coated glass slides and snap frozen in liquid nitrogen. Cover slips were removed to freeze crack the samples and glass slides were incubated in cold 50% acetone / 50% methanol for 10 min. Glass slides were washed 3X with 1% Triton-X100 PBS buffer, blocked with Image-enhancer (Lifesciences) for 20 min and blocked with blocking solution (PBS, 0.1% Tween-20, 1% BSA) for 20min. Slides were incubated with primary antibodies in blocking solution overnight at 4°C ( $\alpha$ -EMB-4 5M19 1:100,  $\alpha$ -HRDE-1 1:500), washed 3X with wash buffer (PBS, 0.1% Tween-20) and incubated for 2 hrs with secondary antibodies ( $\alpha$ -mouse Alexa 647 1:750 (Lifetech) and anti-rabbit Alexa 568 1:750 (Lifetech)). Slides were 3X with wash buffer and mounted using VectaShield DAPI mounting medium. Images were taken using a Leica SP8 confocal microscope using same laser settings between all slides imaged.

#### RNA Sequencing

Synchronised animals were grown to young adult stage at 20°C on HB101 seeded NGM plates. Animals were harvested and washed 3X in M9 buffer. Settled animals were mixed with Trisure reagent, bead beaten as described in proteomics experiments above and total RNA was isolated by a chloroform extraction.

For total RNA sequencing, Illumina Ribozero kit was used to remove ribosomal RNA from 1  $\mu$ g of total RNA prior to library preparation. RNA sequencing libraries were prepared using NEB Next Ultra library preparation kit. Small RNA sequencing performed by treating 5  $\mu$ g of total RNA with Epicentre 5' polyphosphatase to remove the 5' triphosphate from 22G-RNAs. After treatment, RNA is purified by phenol/chloroform extraction and 1  $\mu$ g of RNA is used to prepare small RNA libraries using Illumina TruSeq small RNA library preparation kit. Ribosomal depleted RNA and small RNA libraries are sequenced using Illumina HiSeq 1500 platform.

### RNA Sequence Analysis

The ce10/WS220 genome fasta file was obtained from the WormBase ftp server. Sequences for the piRNA sensor transgene and the piRNA sensor transgene with one intron were added as separate chromosomes when required. A GTF file containing annotations for genes and pseudogene for version ce10/WS220 of the *C. elegans* genome were downloaded from the UCSC table browser. To prevent multiple counts per read in the case of overlapping features, only the longest isoform of each gene was included in the analysis. A GTF file containing annotations for transposable elements was generated by running RepeatMasker (Smit et al., 2015) version open-4.0.5 in sensitive mode, run with rmbblastn version 2.2.27+ using RepeatMasker database version 20140131, against the ce10/WS220 genome fasta file. “Simple\_repeat” and “Low\_complexity” annotations were excluded from the analysis. Raw fastq small RNA sequencing files were processed by removing the Illumina TruSeq adaptor sequence using cutadapt v1.9 (Martin, 2011), with parameters “-minimum-length 18 -discard-untrimmed -a TGGGAATTCTCGGGTGCCAAGG”. Raw fastq RNA sequencing files were processed by removing the NEBNext adaptor sequence using cutadapt v1.9, with parameters “-a AGATCGGAA GAGCACACGCTGAACTCCAGTCAC”. Adaptor-trimmed small RNA sequencing reads were aligned to the ce10/WS220 *C. elegans* genome using STAR v2.5.1b (Dobin et al., 2013), with parameters “-outFilterMultimapNmax 50 -winAnchorMultimapNmax 50 -outFilterMismatchNmax 0 -limitBAMsortRAM 3100000000 -alignIntronMax 1 -alignEndsType EndToEnd -outSAMtype BAM SortedByCoordinate -runThreadN 6 -outBAMsortingThreadN 6 -readFilesCommand 'gunzip -c'”. Adaptor-trimmed RNA sequencing reads were aligned to the ce10/WS220 *C. elegans* genome using STAR v2.5.1b, with parameters “-outFilterMultimapNmax 5000 -winAnchorMultimapNmax 10000 -outFilterMismatchNmax 2 -alignEndsType EndToEnd -outSAMtype BAM Unsorted -runThreadN -readFilesCommand 'gunzip -c'”. Aligned RNA sequencing reads were sorted and indexed using samtools v1.3 (Li et al., 2009). Counts against the annotations in the GTF files were generated with featureCounts v1.5.0-p1 (Liao et al., 2014), with parameters “-T 6 -M -fraction”. Normalised counts, variance-stabilised counts, fold change values, and adjusted p-values were obtained using DESeq2 v3.2.2 (Love et al., 2014), called through a custom script.

We also used exon-intron split analysis (EISA) (Gaidatzis et al., 2015) to characterize the gene expression changes detected between *hrde-1* or *emb-4* null and wild-type strains. Both exonic and intronic read counts were quantified using FeatureCounts (Liao et al., 2014). When using EISA, we processed the counts and the annotation files by following the procedures described by (Gaidatzis et al., 2015).

We calculated the 22G-RNA density using previously published small RNA sequencing data obtained from HRDE-1 immunoprecipitations in wild-type and mutant animals normalised to library size (Sapetschnig et al., 2015). We used a cut-off of 22G-RNA reads in wild-type / *hrde-1* mutant control  $\geq 4$  for filtering out 22G-RNA reads that were unspecifically binding to anti-HRDE-1 antibody. We then used the following calculation 22G-RNA density = # of 22G-RNA reads in HRDE-1 IP of gene A / RPKM of the gene A.

We carried out exon level sRNA differential expression by filtering out genes that have zero mapped reads in all samples and normalising the samples by sample size using the Median Ratio Method (Anders and Huber, 2010) implemented in the R package DESeq2 to adjust for factors like the coverage and sampling depth. Next, we log transformed the exon read counts for each gene, performed a two-sample t-test on each exon independently and adjusted the p-values of testing results by false discovery rate using the Benjamini & Hochberg method (Benjamini and Hochberg, 1995).

### Histone H3K9me3 Chipmentation

H3K9me3 Chipmentation for wild-type, *hrde-1(tm1200)* and *emb-4(qm31)* young adult animals was performed according to (Schmidl et al., 2015) with slight modifications. Briefly, 50,000 animals were frozen and crushed in liquid nitrogen prior to 1% formaldehyde (SIGMA) fixation in 1XPBS. After 10 minutes of fixation at room temperature, 125 mM Glycine was added to quench the formaldehyde. Excess formaldehyde was washed twice with 1XPBS and once with the lysis buffer (50 mM HEPES pH:7.5, 150 mM NaCl, 0.1% TX100, 1 mM EDTA, protease and phosphatase inhibitors). The pelleted extract was resuspended in the lysis buffer and sonicated with a Bioruptor (15 SEC On, 90 SEC OFF, 10 cycles) to get 200-700 bps genomic DNA fragments at 4°C. 3 g of anti-H3K9me3 (Abcam, Ab8898) and 30 I Dynabeads Protein A (ThermoFisher) beads were incubated with 1 ml extract (Lysis buffer + 1% sarkosyl) for 12 hours on a rotating wheel at 4°C. The immunoprecipitated chromatin was washed with 150 mM, 500 mM and 1 M NaCl containing lysis buffer followed by LiCl and 1XTE pH:8.0 buffers, consecutively. Magnetic beads were resuspended in 1X tagmentation buffer with 1  $\mu$ l Tn5 transposase Tagment DNA Enzyme (Illumina Nextera DNA Prep Kit). Tagmentation was performed at 37°C for 10 minutes. Beads were then washed twice with 1XTE pH:8.0. Two rounds of elution buffer (150 mM NaCl, 10 mM Tris pH:8.0, 1% SDS, 1 mM EDTA) was used to elute the immunoprecipitated DNA at 65°C. Eluted DNA was RNase and Proteinase K treated for 1 hour at 37°C and 12 hours at 65°C, respectively. De-crosslinked DNA was purified with Invitrogen PCR Cleaning Kit and DNA concentration was determined with Qubit HS DNA. In addition to CHIP DNAs, input DNAs were prepared from 5 ng purified DNA after sonication via ChIP-tagmentation (Illumina Nextera DNA Prep Kit).

### Amplification and Sequencing of Chipmentation Libraries

1  $\mu$ l of each library was amplified in a 20  $\mu$ l qPCR reaction containing 0.2 M primers, 1X SyberGreen qPCR mix in StepOnePlus to determine the optimum number of PCR cycles in library preparation. Final libraries were prepared according to Illumina Nextera DNA Prep Kit with N cycles of PCR, where N is equal to the *C<sub>t</sub>* value obtained from the qPCR analysis. Sequencing was performed in HiSeq2500 with using single end 50 bp reads.

### Bioinformatic Analysis of H3K9me3 Chipmentation Data

Sequencing reads were aligned to the WS220/ce10 assembly of the *C. elegans* genome using Segemehl version 2.0 (Hoffmann et al., 2009) with 80% sequence similarity. SAMtools was used to convert aligned reads to BAM format. After mapping, the genome coverage was calculated for all individual sample replicates. Replicates were subsequently merged after normalisation by the library size factor  $s = c/g$  (with  $c$ : number of covered bases, and  $g$ : the size of the genome). Images for histone H3K9me3 enrichment on the piRNA sensor were prepared with SeqPlots (Stempor and Ahringer, 2016).

### Transgenic Animals

The *mjls588* allele was generated by removing the introns two and three from the GFP sequence in the plasmid pEM975 that is used to generate the *mjls144* allele. New plasmid is inserted on Chr II using the previously described MosSCI method (Frøkjær-Jensen et al., 2008) into the same location as in *mjls144* allele. *mjsi92* allele is generated by CRISPR tagging of endogenous *emb-4* N-term with the OLLAS epitope sequence using CRISPR gRNA (5'-CAAGAAGCCGTGGTGACTCG-3') and the repair template plasmid pEM2058 using Cas9 protein and RNA injections as described in Paix et al. (Paix et al., 2015).

### Structural Alignment of AQR and EMB-4

EMB-4 structure was determined by PHYRE-2 online prediction tool. Images were generated in Pymol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC) and mutagenesis was performed in Chimera (Yang et al., 2012).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analysis of RNA-Seq, sRNA-Seq and H3K9me3 data is explained in the sections on these methods. Normalised counts, variance-stabilised counts, fold change values, and adjusted p-values were obtained using DESeq2 v3.2.2, called through a custom script. DESeq uses the negative binomial distribution to perform differential expression analysis. For exon level sRNA analysis, we log transformed the exon read counts for each gene, performed a two-sample t-test on each exon independently and adjusted the p-values of testing results by false discovery rate using the Benjamini & Hochberg method. For the analysis of H3K9me3 ChIP data, the genome coverage of reads was calculated for all individual sample replicates. Replicates were subsequently merged after normalisation by the library size factor  $s = c/g$  (with  $c$ : number of covered bases, and  $g$ : the size of the genome).

Number of animals used in microscopy experiments are indicated in relevant figure legends. Error bars indicate standard deviation unless otherwise stated in figure legends.

## DATA AND SOFTWARE AVAILABILITY

SILAC proteomics, RNA-Seq, sRNA-Seq and H3K9me2 ChIP data is deposited to public databases as indicated in the [Key Resources Table](#).

The accession number for the SILAC proteomics data is Proteinexchange: PXD004416, the accession number for the RNA-Seq data is ArrayExpress: E-MTAB-4877, the accession number for the ChIP-Seq data is ArrayExpress: E-MTAB-5662.