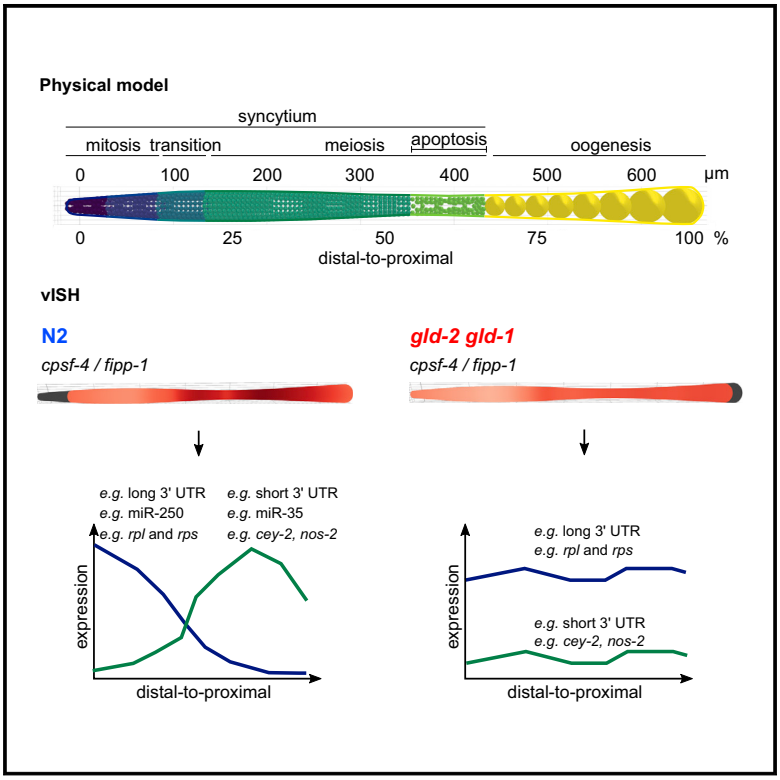
Resource

Developmental Cell

Spatiotemporal m(i)RNA Architecture and 3′ UTR Regulation in the *C. elegans* Germline

Graphical Abstract



Highlights

- Architecture of mRNA and miRNA expression at near single-cell resolution in worm gonad
- For hundreds of genes, 3′ UTR choice and length are regulated during development
- In *gld-2 gld-1* mutants, 3′ UTR choices are suppressed
- A 3D gonad model (SPACEGERM) for virtual *in situs* is presented

Authors

Asija Diag, Marcel Schilling, Filippos Klironomos, Salah Ayoub, Nikolaus Rajewsky

Correspondence

rajewsky@mdc-berlin.de

In Brief

Diag and Schilling et al. present a near single-cell gonad map of mRNAs and miRNAs and previously unannotated miRNAs and 3′ UTRs. By mutant analysis, regulators and mechanisms explaining highly dynamic spatiotemporal changes of 3′ UTR isoforms were identified. A mathematical, interrogable 3D gonad model allows comprehensive resolution of mRNA and miRNA expression in time (min) and space (∼1–100 cells).

Diag et al., 2018, Developmental Cell *47*, 785–800
December 17, 2018 © 2018 Elsevier Inc.
https://doi.org/10.1016/j.devcel.2018.10.005

CellPress

Developmental Cell
Resource

Spatiotemporal m(i)RNA Architecture and 3′ UTR Regulation in the *C. elegans* Germline

Asija Diag,[1,3] Marcel Schilling,[1,3] Filippos Klironomos,[1,2] Salah Ayoub,[1] and Nikolaus Rajewsky[1,4,*]
[1]Laboratory for Systems Biology of Gene Regulatory Elements, Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse10, 13125 Berlin, Germany
[2]Department of Pediatrics, Charité University Medical Center, Division of Oncology and Hematology, Augustenburger Platz 1, 13353 Berlin, Germany
[3]These authors contributed equally
[4]Lead Contact
*Correspondence: rajewsky@mdc-berlin.de
https://doi.org/10.1016/j.devcel.2018.10.005

SUMMARY

In animal germlines, regulation of cell proliferation and differentiation is particularly important but poorly understood. Here, using a cryo-cut approach, we mapped RNA expression along the *Caenorhabditis elegans* germline and, using mutants, dissected gene regulatory mechanisms that control spatiotemporal expression. We detected, at near single-cell resolution, >10,000 mRNAs, >300 miRNAs, and numerous unannotated miRNAs. Most RNAs were organized in distinct spatial patterns. Germline-specific miRNAs and their targets were co-localized. Moreover, we observed differential 3′ UTR isoform usage for hundreds of mRNAs. In tumorous *gld-2 gld-1* mutants, gene expression was strongly perturbed. In particular, differential 3′ UTR usage was significantly impaired. We propose that PIE-1, a transcriptional repressor, functions to maintain spatial gene expression. Our data also suggest that *cpsf-4* and *fipp-1* control differential 3′ UTR usage for hundreds of genes. Finally, we constructed a "virtual gonad" enabling "virtual *in situ* hybridizations" and access to all data (https://shiny.mdc-berlin.de/spacegerm/).

INTRODUCTION

Spatial and temporal restriction of gene expression is considered a crucial conserved mechanism for cellular and developmental programs such as specification of cell fates and compartmentalization. The functions of mRNA localization are diverse: it might lead to stabilization of the transcript, and it is thought to be more energy efficient by providing a template for multiple rounds of local translation (Jansen, 2001). Which mechanisms control mRNA localization? Previous studies suggested that alternative polyadenylation (APA) and hence 3′ untranslated regions (3′ UTRs) are important post-transcriptional regulators of spatially restricted gene expression and cell fate transition (Brumbaugh et al., 2018; Mayr, 2017). Additionally, an *in vivo* study of the *Caenorhabditis elegans* germline revealed that 3′ UTRs are the primary regulators of gene expression (Merritt et al., 2008). Furthermore, differential 3′ UTR usage can modulate the balance between proliferation and differentiation (Lackford et al., 2014; Mayr and Bartel, 2009; Sandberg et al., 2008; Shepard et al., 2011). Cells must decide, whether, when, where, and how fast to proliferate in order to keep the balance between proliferation and differentiation as improper regulation can lead to developmental defects and cancer. However, our understanding of how mRNA localization regulates this balance remains limited.

The *C. elegans* germline is a powerful *in vivo* model for studying the balance between cell proliferation and differentiation. The basic factors, molecular architecture, and processes are similar to those of other metazoans, and major players have been remarkably conserved during evolution. The germline is divided into different compartments: in the distal gonad arm close to the germline niche (distal tip cell [DTC]), proliferative germ cells are located, which form a syncytial tissue (Hirsh et al., 1976) (Figure 1A). At a defined distance from the niche, germ cells exit the mitotic cell cycle and enter meiosis (Figure 1A). This switch from proliferation to differentiation is termed mitosis-to-meiosis transition. Around the bend region, many germ cells undergo apoptosis (Gartner et al., 2008). Only certain germ cells become oocytes or sperm.

Previous studies already showed evidence that spatiotemporal restriction of RNA binding proteins (RBPs) in the *C. elegans* germline can regulate mRNA expression by binding to their 3′ UTR (Crittenden et al., 2006; Nousch and Eckmann, 2013). One important example is GLD-1, an RBP that binds multiple mRNAs including its own mRNA, thereby regulating the switch from proliferation to differentiation (Brenner and Schedl, 2016; Jungkamp et al., 2011; Theil et al., 2018). Additionally, GLD-2, the cytoplasmic poly(A)-polymerase (cytoPAP) in the *C. elegans* germline, accumulates around the pachytene stage and promotes meiotic entry by polyadenylation of mRNAs required for differentiation (Millonigg et al., 2014; Nousch et al., 2017).

Besides RBPs and 3′ UTRs being key players in regulating mRNA stability, translation, and localization, previous studies suggested that microRNAs (miRNAs) might control proliferation and differentiation in the *C. elegans* germline, too (Bukhari et al., 2012; Ding et al., 2008). MicroRNAs (miRNAs) belong to the class of small non-coding RNAs and are conserved post-transcriptional regulators of gene expression that bind mRNAs, primarily in their 3′ UTR (Bartel, 2018; Jens and Rajewsky,

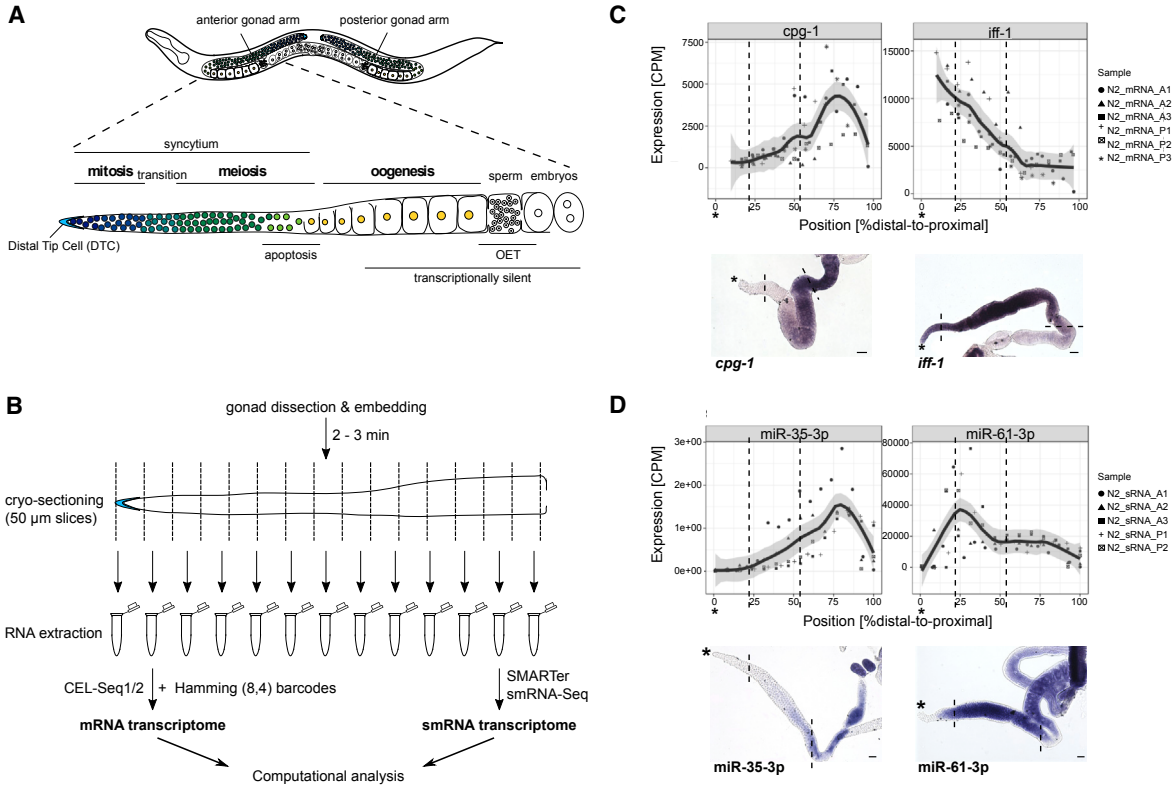


Figure 1. mRNAs and miRNAs Are Localized in the Germline

(A) Schematic overview of the *Caenorhabditis elegans* gonad. OET: oocyte-to-embryo transition.

(B) Schematic overview of the experimental approach.

(C) Spatial expression of *cpg-1* and *iff-1* from distal to proximal. n = 6 independent experiments (N2_mRNA_A1-A3 and N2_mRNA_P1-P3) for wild-type N2, LOESS ± standard error (SE). Corresponding *in situ* hybridization (ISH) images of *cpg-1* and *iff-1*. Asterisk: distal tip cell (DTC). Scale bar: 20 μm. Dashed lines represent the different zones in the germline.

(D) Spatial expression of mir-35-3p and miR-61-3p from distal to proximal. n = 5 independent experiments (N2_sRNA_A1-A3 and N2_sRNA_P1-P2) for wild-type N2, LOESS ± SE. Corresponding ISH images of mir-35-3p and miR-61-3p. Asterisk: DTC. Scale bar: 20 μm. Dashed lines represent the different zones in the germline.

See also Figures S1, S2, and S4.

2015). Usually, miRNA binding leads to transcript destabilization and/or translational inhibition. Bukhari and colleagues showed that loss of function of *alg-1* and *alg-2*, two miRNA-specific Argonaute proteins in *C. elegans*, leads to a reduced mitotic region and less proliferative cells in the *C. elegans* germline, indicating an important role of miRNAs in controlling germ cell biogenesis (Bukhari et al., 2012). So far, due to technical limitations such as low RNA content of the *C. elegans* germline and lack of sequencing protocols for low input materials, it has not been possible to gain a system-wide spatiotemporal resolved characterization of miRNA expression during germ cell proliferation and differentiation, with the exception of the oocyte-to-embryo transition (Stoeckius et al., 2014).

Here, we established an optimized version of the tomo-seq approach (Junker et al., 2014) and improved sequencing protocols for the *C. elegans* hermaphrodite germline (see STAR Methods). Thus, we were able to quantify mRNA and miRNA expression at near single-cell resolution, as a function of position along the germline. We capture *in vivo* RNA expression during the entire development of germ cells through proliferation and differentiation. With our approach, we were able to detect previously unannotated miRNAs with highly spatially restricted expression. As we also analyzed mutant strains, our data offer specific insights into mechanisms that are functionally important during germ cell development. We compared the spatiotemporal resolved gene expression of wild-type germline to the *gld-2 gld-1* double mutant germline, unravelling potential key players such as pie-1, a GLD-2 target, in the transition from proliferation to differentiation. By careful bioinformatics analysis of our data, we discovered also hundreds of unannotated 3′ UTRs that had escaped previous approaches probably because they are often specifically expressed. Furthermore, we discovered that differential 3′ UTR usage is widespread along the germline. Strikingly, this phenomenon, which is key for changing regulation of mRNAs across space and time, was perturbed in *gld-2 gld-1* mutants. With the exception of the cleavage and polyadenylation specificity factor (*cpsf-4*) and factor interacting with poly(A) polymerase (*fipp-1*), all other factors known to regulate alternative 3′ UTR usage were not perturbed in the mutant, arguing that the dynamic expression of these two

factors may contribute to differential 3′ UTR expression. To provide a user-friendly interface of our massive data and to analyze gene expression of different genes but compared to a "universal" germline reference coordinate system, we set out to create a 3D germline model. By collecting data from the literature, by mining our own microscopy data, and by mathematical modelling, we were able to create "SPACEGERM," a model that reflects germline gene expression. SPACEGERM can be interactively mined via the internet and can be used to systematically perform "virtual *in situs*" for >10,000 mRNAs and hundreds of miRNAs. In summary, we created a map of spatially and temporally resolved germline mRNA and miRNA expression, and our analysis provides crucial insights into mechanisms and function of RNA in the germline.

RESULTS

mRNAs and miRNAs Are Localized in the Germline

To investigate the spatiotemporal distribution of gene expression along the germline, we dissected and embedded gonads of young adult hermaphrodites in tissue freezing medium. This entire procedure takes only a few minutes, minimizing RNA degradation. We then cryo-sectioned the gonads into ~15 slices of 50 μm thickness (~100 cells/data point in the distal and 1 oocyte/data point in the proximal gonad arm) and performed RNA sequencing (RNA-seq) on each slice with different approaches for mRNAs and small RNAs (Figure 1B). Analysis of the sequencing data revealed that most reads matched known *C. elegans* transcripts (~75.8% for mRNAs and ~59% for small RNAs) (Figures S1A and S2A). Furthermore, quantified transcripts were in line with a poly(A) selection profile as expected due to the barcoded oligo(dT) primer used for capture (Figure S1B). Comparing pairs of biological and technical replicates confirmed that our experimental approach is highly reproducible (Pearson correlation coefficient: 0.96) (Figures S1C and S2B). Additionally, as a control, we performed our experimental approach with uncut gonads and compared it to a sliced gonad in order to investigate if the slicing had an influence on the measured gene expression. By averaging the measurement of gene expression across slices, we reconstructed *in silico* the gene expression profile of the uncut gonad (Figures S1D and S2C). Moreover, the ERCC spike-in abundance estimated from RNA-seq correlated well with the known concentrations (Figure S1E). Furthermore, we showed that our sequencing method is reliable, since it compared very well to other sequencing approaches such as poly(A)+-seq and ribosomal RNA depleted total RNA-seq (Figure S1F).

As our biological replicates were slightly shifted and compressed to each other due to different cutting start points, we aligned samples to a common coordinate system (see STAR Methods) before integrating the data of all replicates for downstream analyses. At this stage of the analysis, we also accounted for missing data in the very first slices. This is probably due to collecting empty slices before the actual specimen. Alternatively, this could reflect extremely low levels of total RNA in this region. A loss of material is unlikely as we observed the expected amount of ERCC spike-in reads. To determine the starting point of RNA expression, we used *in situ* data as a guide (see STAR Methods).

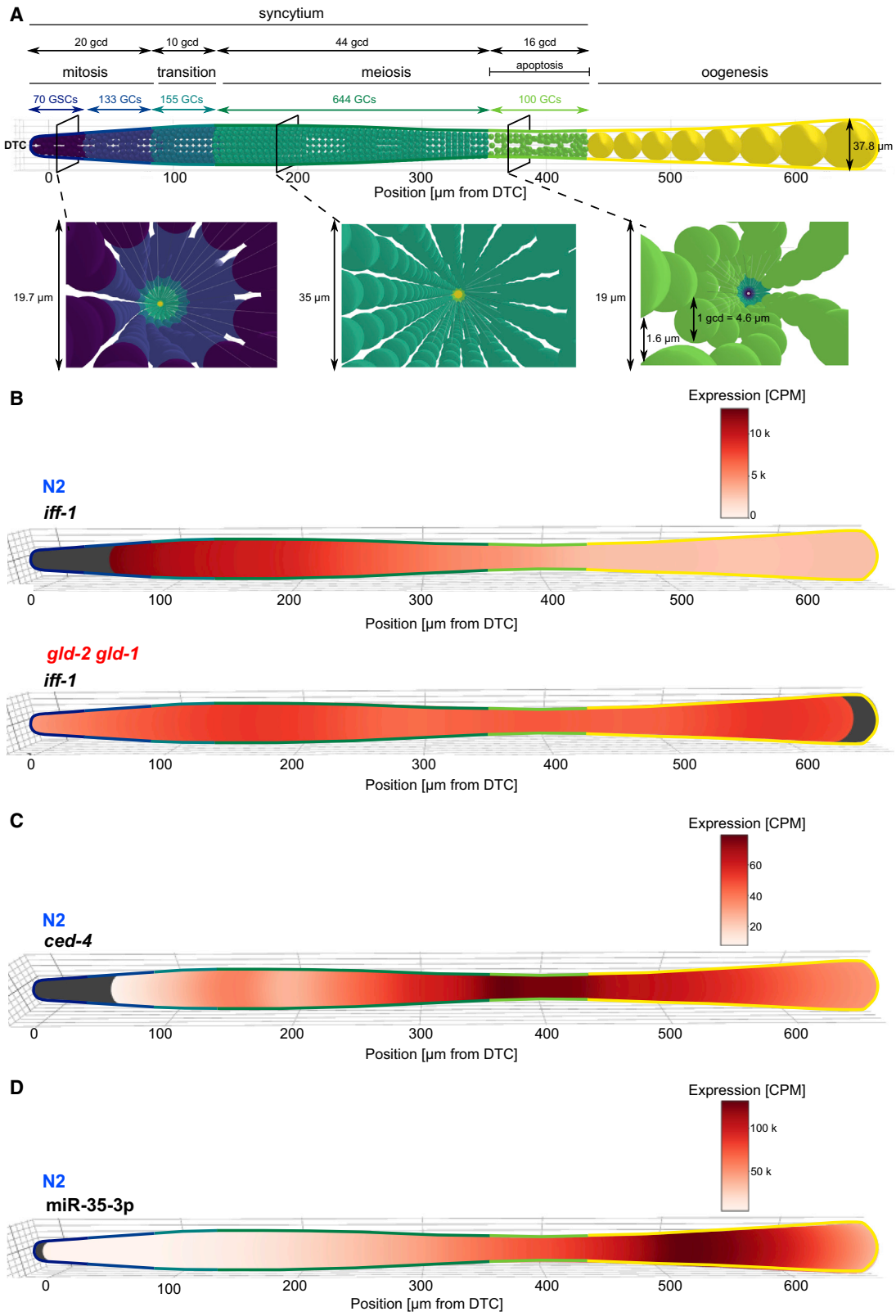
The *C. elegans* hermaphrodite contains two gonad arms, the anterior and the posterior gonad arm. We cut both arms to investigate any difference in gene expression between the arms. However, in concordance with the literature, we did not detect any difference above background between both arms (Figure S3A). Observable differences in gene expression between anterior and posterior gonad arm decreased with rising expression levels, arguing that they reflect noise. Hence, we treated the samples as biological replicates and could therefore further increase the statistical power of our analysis. Investigating the expression of mRNAs and miRNAs revealed that both RNA classes display distinct localization pattern across the germline (Figures 1C, 1D, S3B, S3C, and S3D). We confirmed the gene expression profiles with *in situ* hybridization (ISH) on the gonad (Figures 1D, 1E, S3A, S3B, and S3C). Altogether, these results demonstrate that our sequencing approach is reproducible and reliable and that the data reveal spatiotemporal organization of mRNAs and miRNAs throughout the germline.

A 3D Germline Model Reflects RNA Localization through Germ Cell Proliferation and Differentiation

We thought that it would be useful to construct a virtual model of the germline that can serve as a generalized framework on which expression data can be displayed and compared. For this, we systematically collected published data about the size, composition, and germ cell migration speed of each zone in the germline (Brenner and Schedl, 2016; Fox et al., 2011; Hansen and Schedl, 2013; Hirsh et al., 1976; Hubbard, 2007; Maciejowski et al., 2006; Wolke et al., 2007) and quantified our own gonad images (Table S1). Using these data, we were able to compute an *in silico* 3D physical germline model (STAR Methods and Figure 2A). Within the model, we defined the size of each zone and assigned the number of germ cells to each zone in the germline (Figure 2A). We used the 3D model as a guide to assign the different germline zones to our expression profiles. Thus, our 3D germline model integrates *in vivo* mRNA and miRNA expression throughout proliferation and differentiation of germ cells (Figures 2B–2D). Moreover, the model represents *in vivo* mRNA expression in perturbed systems such as the *gld-2 gld-1* double mutant (Figure 2B). In order to validate the assignment of the zones in our 3D model, we searched for apoptotic gene markers in the germline as most of the germ cells undergo apoptosis around the bend region. Indeed, we found apoptotic genes such as *ced-4*, having their highest expression precisely around the bend region, which starts at a distance of ~350 μm from the DTC (Figure 2C), in accordance with the assignment of the bend region in our model. Finally, our 3D model also represents miRNA localization throughout the germline (Figure 2D). Overall, we believe that the virtual germline may serve as a reference for future studies.

Spatial Gene Expression Is Perturbed in *gld-2 gld-1* Double Mutants

To determine whether mRNAs display a global localization pattern throughout the germline, we clustered the expression of germline-specific genes (Wang et al., 2009) by correlation (1 - Pearson's r). Clustering the expression data revealed that mRNAs are organized in groups with distinct localization patterns (Figure 3A). We observed many different gene clusters



(legend on next page)

along the germline. However, assigning these clusters to the zones in the germline by using SPACEGERM showed that most genes peaked in expression either in the mitotic or oogenesis region, whereas in the meiotic region, genes required for proliferation and mitosis slowly decreased and genes required for differentiation and oogenesis slowly increased in abundance.

To test whether mRNA localization is important for the transition between proliferation and differentiation, we performed the cryo-based method for the *gld-2 gld-1* double mutant, which possesses only one-third of meiotic entry, ending up in a solely proliferating and tumorous germline (Brenner and Schedl, 2016). Consistently, clustering the expression of germline-specific genes for the *gld-2 gld-1* mutant revealed that mRNA localization is perturbed compared to the wild-type (Figure 3A). In most cases, genes required for proliferation, e.g., *iff-1*, were expressed continuously throughout the germline, whereas genes required for embryogenesis were downregulated, e.g., *perm-2* and *perm-4*. However, clustering of the same genes in the *gld-2 gld-1* double mutant revealed that some genes still localize in the mutant (Figure S4A). In addition to the *gld-2 gld-1* double mutant, we further investigated the spatial gene expression of the *glp-1 (gf)* mutant germline, which possesses a prolonged proliferative zone. The mutant is temperature sensitive, resulting in an inducible tumorous phenotype. Hence, dissection of gonads from these mutants was impeded. Therefore, we only induced the phenotype for a short time, avoiding tumor development. As the induction of the mutation was very short, the spatial gene expression resembled more the wild-type germline (Figure S4B).

Based on the finding that mRNAs are localized in the wild-type germline and that this localization pattern is perturbed in the *gld-2 gld-1* mutant, we explored whether there are specific mRNAs localizing to a certain zone of the germline, i.e., proliferation or differentiation. We observed that many genes encoding for a ribosomal subunit (*rpl* and *rps* genes) had their highest expression in the distal gonad arm and decreased in expression in the proximal arm (Figure 3B). Interestingly, this was not the case for the *gld-2 gld-1* double mutant (Figure 3B). The *rpl* and *rps* genes had similar initial expression levels in the *gld-2 gld-1* double mutant as in the wild-type. However, the expression did not decrease in the proximal arm but stayed constant along the germline. This result confirms an important role of these genes in proliferation and was consistent with ISH images (Figures 3C and 3D). In contrast to the *rpl* and *rps* genes, we observed some genes that had their highest expression in the proximal arm in the wild-type while these genes were downregulated or completely absent in the *gld-2 gld-1* mutant. We identified *pie-1*, *cey-2*, and *nos-2* among these genes. The *pie-1* gene encodes for a maternal CCCH-type zinc-finger protein that is specific for oocytes and embryos (Merritt et al., 2008; Tenenhaus et al., 2001). Previous studies showed evidence that PIE-1 is a bifunctional protein that blocks the transcription of somatic transcripts during blastomere development, ensuring the germline fate and that it is required for the maintenance of *nos-2* and *cey-2* (Seydoux and Dunn, 1997; Seydoux et al., 1996; Tenenhaus et al., 2001). Additionally, it was shown that *pie-1* is a target of the cytoplasmic polymerase, GLD-2 (Kim et al., 2010). Consistent with these previously described findings, we observed in the *gld-2 gld-1* mutant, where GLD-2 is depleted, a strong downregulation of *pie-1* as well as *nos-2* and *cey-2* in accordance with our ISH images (Figures 3E and S4C). However, the *gld-2* single mutant did not show this strong downregulation (Figure S4C), indicating an important role of GLD-1 in this process as well. Together, these results suggest that *rpl* and *rps* genes are important for germ cell proliferation and *nos-2* and *cey-2* are important for differentiation, in which expression is maintained through PIE-1 expression.

Germline-Specific Small RNA Sequencing Identifies Previously Unannotated miRNAs

In order to investigate the spatially restricted expression of miRNAs in the germline, we used the SMARTer smRNA kit from Clontech. The kit has a very low level of bias as adapter ligation is completely abolished (Dard-Dascot et al., 2018). A side effect of the SMARTer kit, as reported by Dard-Dascot and colleagues, is the high frequency of side products such as adapter concatemers. However, standard preprocessing of small RNA-seq raw reads requires efficient adapter trimming, which removes such artefacts. While this approach has the potential to capture other small RNAs, we focused on miRNAs as this class of small RNAs was implicated in regulation of germ cell proliferation and differentiation (Bukhari et al., 2012). Of note, clustering the expression of all detected miRNAs in the germline revealed that miRNAs are organized spatially in the germline (Figure 4A).

Due to technical limitations, i.e., sequencing of small RNAs with very low input material ($\leq$1 ng total RNA), it is likely that miRNAs specifically expressed in the gonad or even limited to a specific region therein would have been missed by previous attempts to identify miRNAs as their signal would have been diluted out. As such highly specific miRNAs would be prime candidates for key regulators of spatial expression in the gonad, we screened our germline-specific small RNA-seq data for potential unannotated miRNAs. Therefore, we ran miRDeep2 (Friedländer et al., 2012) on our data. Indeed, we were able to predict 83 previously unannotated precursor miRNAs (Figure 4B and Table S2). In order to quantify these unannotated miRNAs, we included the mature and precursor sequences of the unannotated miRNA predictions in the miRBase21 reference and re-ran miRDeep2 for quantification of known and unannotated miRNAs in each slice separately. Unannotated miRNAs as well as known miRNAs were reproducibly quantifiable. Remarkably, most of

Figure 2. A 3D Germline Model Reflects RNA Localization throughout Germ Cell Proliferation and Differentiation

Figure360▷ For a Figure360 author presentation of Figure 2, see https://doi.org/10.1016/j.devcel.2018.10.005.

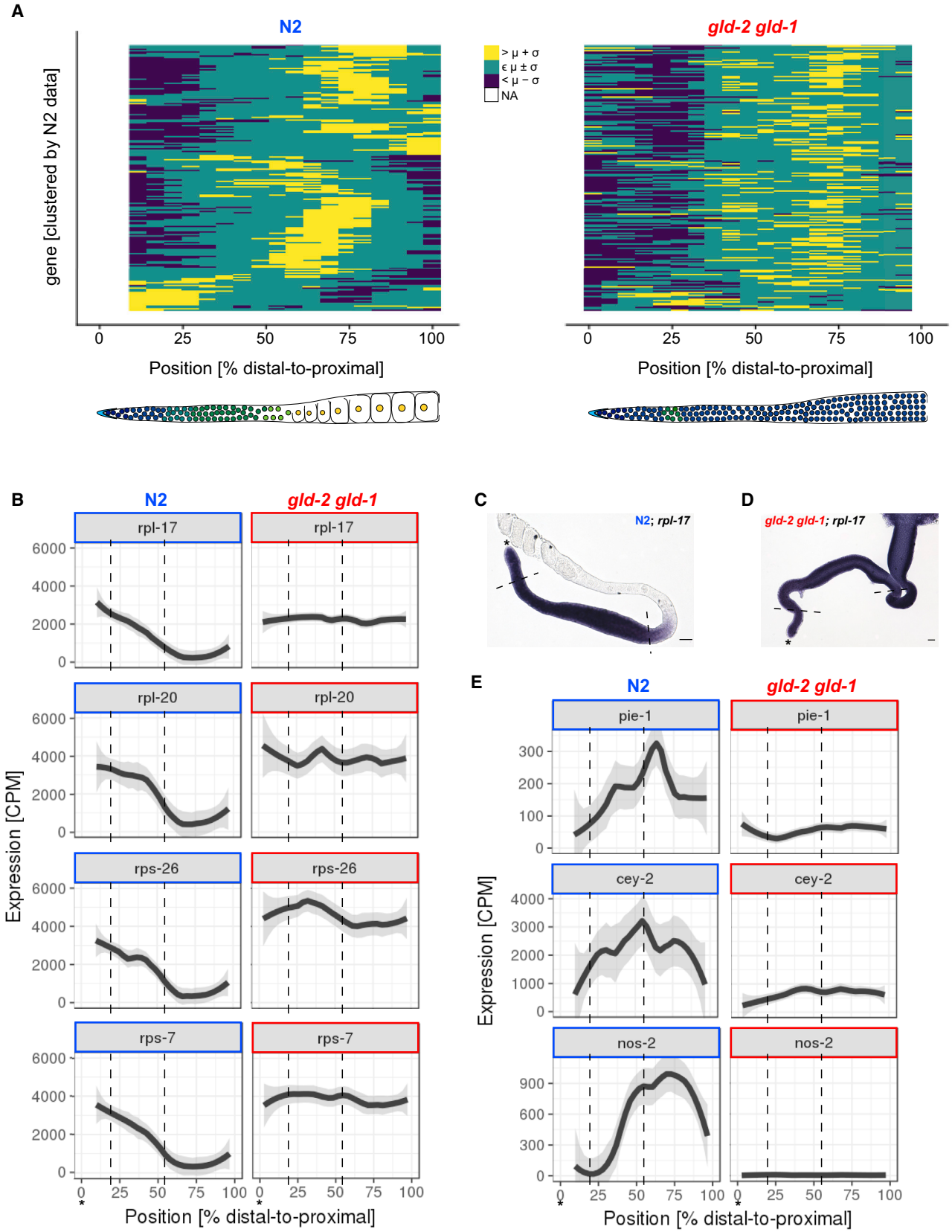
(A) 3D germline model with assigned sizes of each zone in germ cell diameter (gcd) and corresponding germ cell (GC) numbers. Three cross sections are shown at 70 μm, 200 μm, and 380 μm from the distal tip cell (DTC).

(B) 3D germline model representing *in vivo* expression of *iff-1* in N2 and *gld-2 gld-1* double mutant. Gray: no data.

(C) 3D germline model representing *in vivo* expression of *ced-4* in N2. Gray: no data.

(D) 3D germline model representing *in vivo* expression of miR-35-3p in N2. Gray: no data.

See also Figure S7 and Table S1.



(legend on next page)

the unannotated miRNAs were, when averaged over the germline, very lowly expressed (≤100 CPM), explaining why they might have been missed in previous studies. Unlike most known miRNAs, several unannotated miRNAs were primarily expressed in the distal part of the germline, suggesting a specific role for these miRNAs in proliferation (Figure 4B).

As low expression and distinct localization could be an indication of technical artefacts, we picked four (nov-1-3p, nov-63-3p, nov-72-3p, and nov-82-5p) out of the 83 unannotated precursor miRNAs with different expression levels for validation. All four candidates revealed a miRNA-like hairpin structure when folding their pre-miRNA sequence *in silico* with star and mature sequences extensively complementing each other (Figure 4C). Furthermore, the read coverage was miRNA-like with reads stacking up mostly on the mature sequence at aligned 5′ positions (Figures S5A–S5D). We validated the expression of three unannotated miRNAs out of the four either with TaqMan assay or with ISH experiments (Figures 4D and 4E). Interestingly, nov-63-3p revealed germline specificity as it was almost 8-fold enriched in the germline compared to the whole worm (Figure 4D). We could not validate nov-82-5p, maybe due to its low expression. Consistently, expression of miRNAs as measured by qPCR (raw C_Ts) correlates well with the CPMs determined with our sequencing approach (Figure S5E) Interestingly, the loci of nov-63-3p, nov-72-3p, and nov-82-5p were found covered by reads from DCR-1 PAR-CLIP data (Rybak-Wolf et al., 2014) and ALG-1 iPAR CLIP data (Grosswendt et al., 2014) previously published by our lab (Table S2), supporting the existence and functionality of these unannotated miRNAs.

Further, we investigated whether the unannotated miRNAs have distinct targets. For this purpose, we used the miRNA:mRNA chimera data generated previously in our lab (Grosswendt et al., 2014). The chimera data were generated using L3 staged worms, which lack a fully developed germline. Hence, we were not able to find almost any of our unannotated miRNA candidates in this dataset as we used young adults. However, nov-72-3p revealed an interesting unannotated interaction pattern as it interacts not only with mRNAs but also with other already known miRNAs (Figure S5). The strongest interaction was discovered between nov-72-3p and miR-52-5p. However, we did not observe any significant positive correlation between nov-72-3p and corresponding miRNAs throughout the gonad (data not shown). As we were able to validate nov-72-3p also in whole worm samples, the interaction between nov-72-3p and other miRNAs may play an important role in other tissues of the worm rather than the germline. One main issue with nov-72-3p was that we were not able to trace back the origin of this unannotated miRNA as the 20 nt of mature mRNA mapped to the *dpy-2* locus, whereas the first 17 nt also mapped to the locus of *rrn's*. However, mapping to the *dpy-2* locus revealed that the seed region of nov-72-3p is strongly conserved among other species (Figure S5C).

We also found chimeric reads for nov-1-3p. Intriguingly, nov-1-3p and other predicted unannotated miRNAs harbor a stretch of A's in the seed region, suggesting a potentially uncharacterized class of miRNAs. We validated nov-1-3p with small RNA ISH (Figure S4E). Overall, we detected a high fraction of previously unannotated miRNAs being localized throughout the germline.

A Germline-Specific miRNA Family Co-localizes with Its Predicted Targets

Based on the discovery that mRNAs and miRNAs revealed similar spatiotemporal expression patterns across the germline, we asked whether expression of miRNAs and their corresponding targets is co-localized, suggesting a putative miRNA:mRNA interaction. Instead of correlating each miRNA separately with putative targets, we correlated the summed family-wise expression with the corresponding expression of the putative target. An miRNA family was defined by the 6mer seed found in the 2–7 nt of the miRNA members. Putative targets were identified by their 3′ UTR carrying at least one 7mer seed for the 2–8 nt of the miRNA or one 6mer seed for the 2–7 nt of the miRNA, provided that opposite the first miRNA nucleotide was an A (Bartel, 2009). We investigated the correlation for the miR-35 family members that are known to be germline-specific (Miska et al., 2007) and that localize to the proximal gonad arm (Figures 1D, 2D, and S3D). Indeed, we showed that miR-35-3p is enriched in the germline compared to whole worm (Figure 4D). The analysis revealed that all miR-35 family members correlate positively with their targets, indicating a germline-specific interaction (Figure 5A). In contrast, miR-1-3p, a non-germline-specific miRNA that is expressed lowly in the germline compared to miR-35 family members (Figure 4D), did not reveal any prominent co-localization pattern with its targets, indicating an interaction outside of the germline (Figures 5A and 5B). Overall, our data suggest that germline-specific miRNAs co-localize with their targets, which is a necessity for *in vivo* interaction.

Hundreds of Previously Unannotated 3′ UTR Isoforms Detected in the Germline, and Hundreds of 3′ UTRs Are Switched during Development

A recent study in the *C. elegans* germline suggested that 3′ UTRs are the main regulators of gene expression and not promoters

Figure 3. Spatial Gene Expression Is Perturbed in *gld-2 gld-1* Double Mutants

(A) Hierarchical clustering of germline-specific genes by linear correlation (1 - Pearson's r) for N2 and *gld-2 gld-1* double mutant. μ: mean; σ: standard deviation. NA: no data.

(B) Spatial expression of two *rpl* genes (*rpl-17* and *rpl-20*) and two *rps* genes (*rps-26* and *rps-7*) in N2 and *gld-2 gld-1* double mutant from distal to proximal, respectively. n = 6 independent experiments for N2, and n = 4 independent experiments for *gld-2 gld-1* double mutant; LOESS ± standard error (SE). Asterisk: distal tip cell (DTC). Dashed lines represent the different zones in the germline.

(C) *In situ* hybridization (ISH) image of *rpl-17* in N2. Asterisk: DTC. Scale bar: 20 μm. Dashed lines represent different zones in the germline.

(D) ISH of *rpl-17* in *gld-2 gld-1* double mutant. Asterisk: DTC. Scale bar: 20 μm. Dashed lines represent the different zones in the germline.

(E) Spatial expression of *pie-1*, *cey-2*, and *nos-2* in N2 and *gld-2 gld-1* double mutant from distal to proximal, respectively. n = 6 independent experiments for N2, and n = 4 independent experiments for *gld-2 gld-1* double mutant; LOESS ± SE. Asterisk: DTC. Dashed lines represent different zones in the germline.

See also Figure S3.

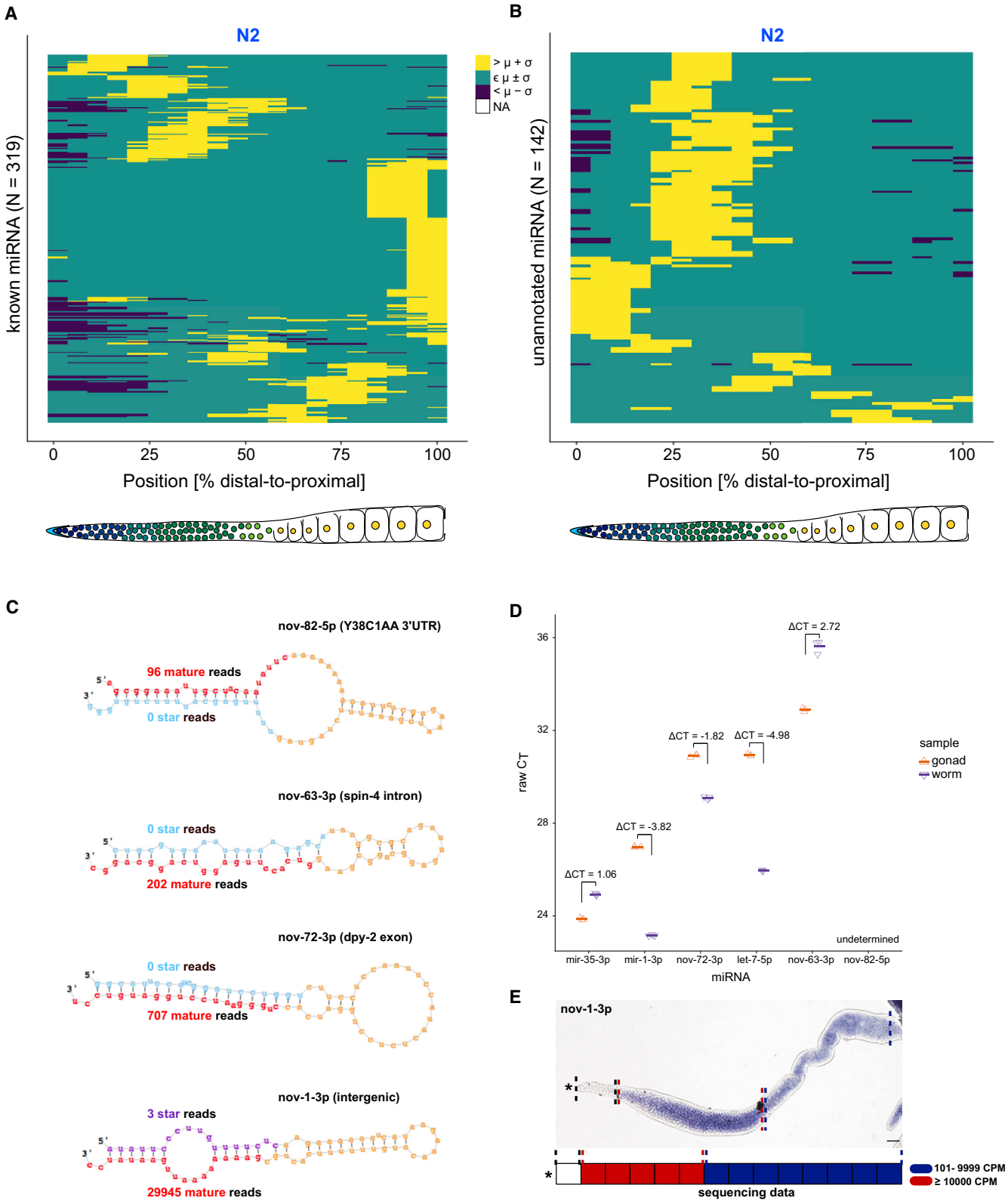


Figure 4. Germline-Specific Small RNA Sequencing Identifies Previously Unannotated miRNAs

(A) Hierarchical clustering of known miRNAs by linear correlation (1 - Pearson's r) for N2. μ, mean; σ, SD.

(B) Hierarchical clustering of unannotated miRNAs by linear correlation (1 - Pearson's r) for N2. μ, mean; σ, SD.

(legend continued on next page)

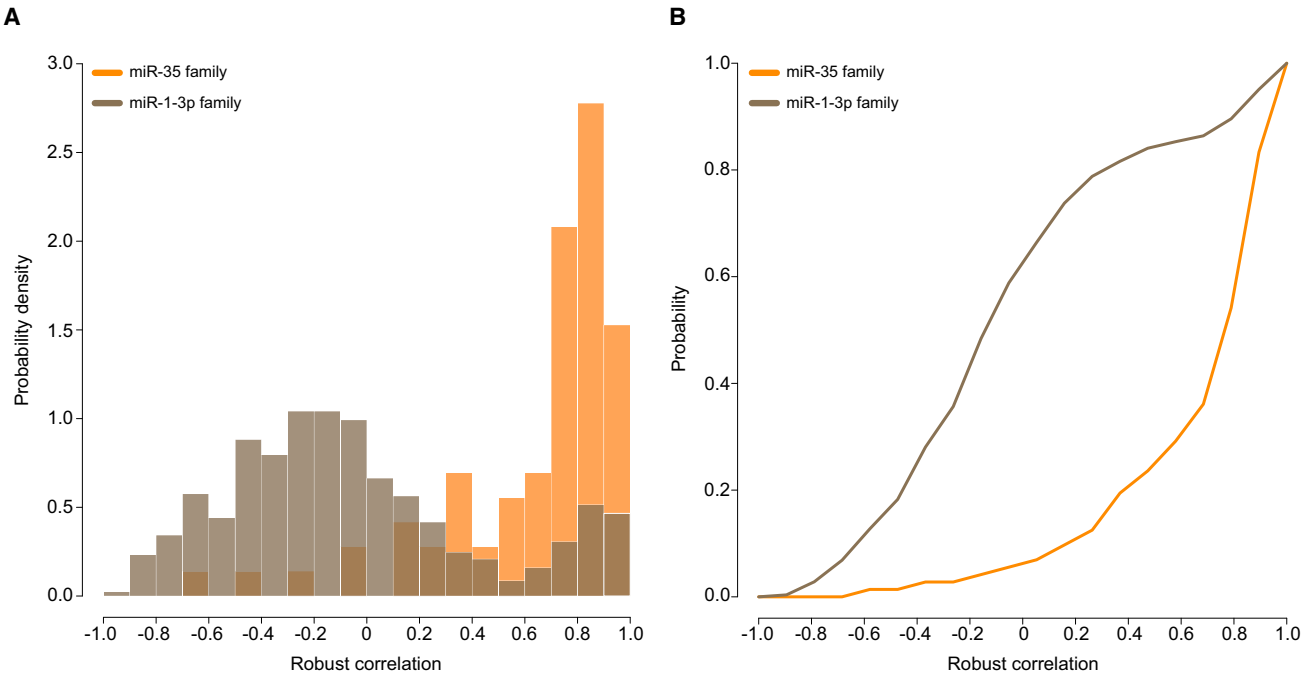


Figure 5. A Germline-Specific miRNA Family Co-localizes with Its Putative Targets

(A) Histogram of robust correlation of miR-35 and miR-1 family members with their putative mRNA targets, respectively.

(B) Density of robust correlation coefficients of miR-35 and miR-1 family members with their putative mRNA targets, respectively.

(Merritt et al., 2008). Additionally, short 3′ UTRs are mainly expressed in proliferating cells, whereas long 3′ UTRs are predominantly expressed in differentiating cells (Mayr and Bartel, 2009; Sandberg et al., 2008). Because the germline is divided in proliferating and differentiating cells, we aimed to determine whether genes expressing more than one isoform change 3′ UTR length across the germline. While most reads, as expected, map to genomic loci annotated as 3′ ends of protein coding genes, we observed several coverage peaks downstream of annotated genes, suggesting longer 3′ UTRs for these transcripts. Hence, we first extended the 3′ UTR annotation (STAR Methods and Figure 5A). We detected 499 intergenic peaks and assigned them to an upstream gene if the intergenic peak was less than 10 kb downstream (Figure S6A and Table S3). Of these intergenic peaks, we considered only the ones as valid 3′ UTRs that were less than 3 kb downstream of the assigned gene, leaving 419 candidates considered as previously unannotated 3′ UTRs (Table S3). Out of the 419 candidates, we picked 13 (Figure S6B) and validated nine of them by nested PCR (Figures S6C–S6F) or conventional PCR (Figures S6G and S6H), followed by Sanger-sequencing. After annotation of 3′ UTRs and further downstream analysis, we quantified the change of the (relative) 3′ UTR usage along the germline for 910 genes. Interestingly, we observed that some of these genes used predominantly the distal polyadenylation signal (PAS) in the distal gonad arm (longer 3′ UTR), while the proximal PAS (shorter 3′ UTR) was used mainly in the proximal arm (Figures 6B, 6C, S6G, and S6H). The switch occurred around the bend region of the gonad where almost 90% of the cells undergo apoptosis (Hansen and Schedl, 2013). As miRNAs from the miR-35 family, a germline-specific miRNA family, and other miRNAs have their highest expression around the bend region, i.e., late pachytene stage (Figures 1D, S3D, and 4A), it suggests that switching from distal-to-proximal PAS may be a mechanism to evade degradation of the transcript as longer 3′ UTRs usually harbor binding sites for miRNAs or other negative regulators.

In summary, we re-annotated the 3′ UTRs for more than 400 genes and showed that differential 3′ UTR usage takes place along the germline.

Differential 3′ UTR Isoform Usage Is Strongly Perturbed in the *gld-2 gld-1* Double Mutant

To better understand the mechanism by which differential 3′ UTR usage occurs across the germline, we examined potential 3′ UTR switching candidates in the *gld-2 gld-1* mutant. Surprisingly, the mutant revealed much fewer genes that switch isoform usage

(C) Four examples of identified unannotated miRNAs of different *C. elegans* genomic origin. Reduced miRDeep2 plots show the precursor hairpin structure and the coverage of mature (red), star (blue, violet), and loop (yellow) sequences.

(D) TaqMan assay validation of known miRNAs expressed in *C. elegans* (mir-35-3p, mir-1-3p, and let-7-5p) and unannotated miRNA predictions (nov-63-3p, nov-72-3p, and nov-82-5p). n = 3 independent experiments for gonad and whole worm sample, respectively. Triangles: raw C_T values; Bars: mean values.

(E) *In situ* hybridization (ISH) image of nov-1-3p with corresponding spatial sequencing data. Asterisk: DTC. Scale bar: 20 μm.

See also Figures S4 and S5 and Table S2.

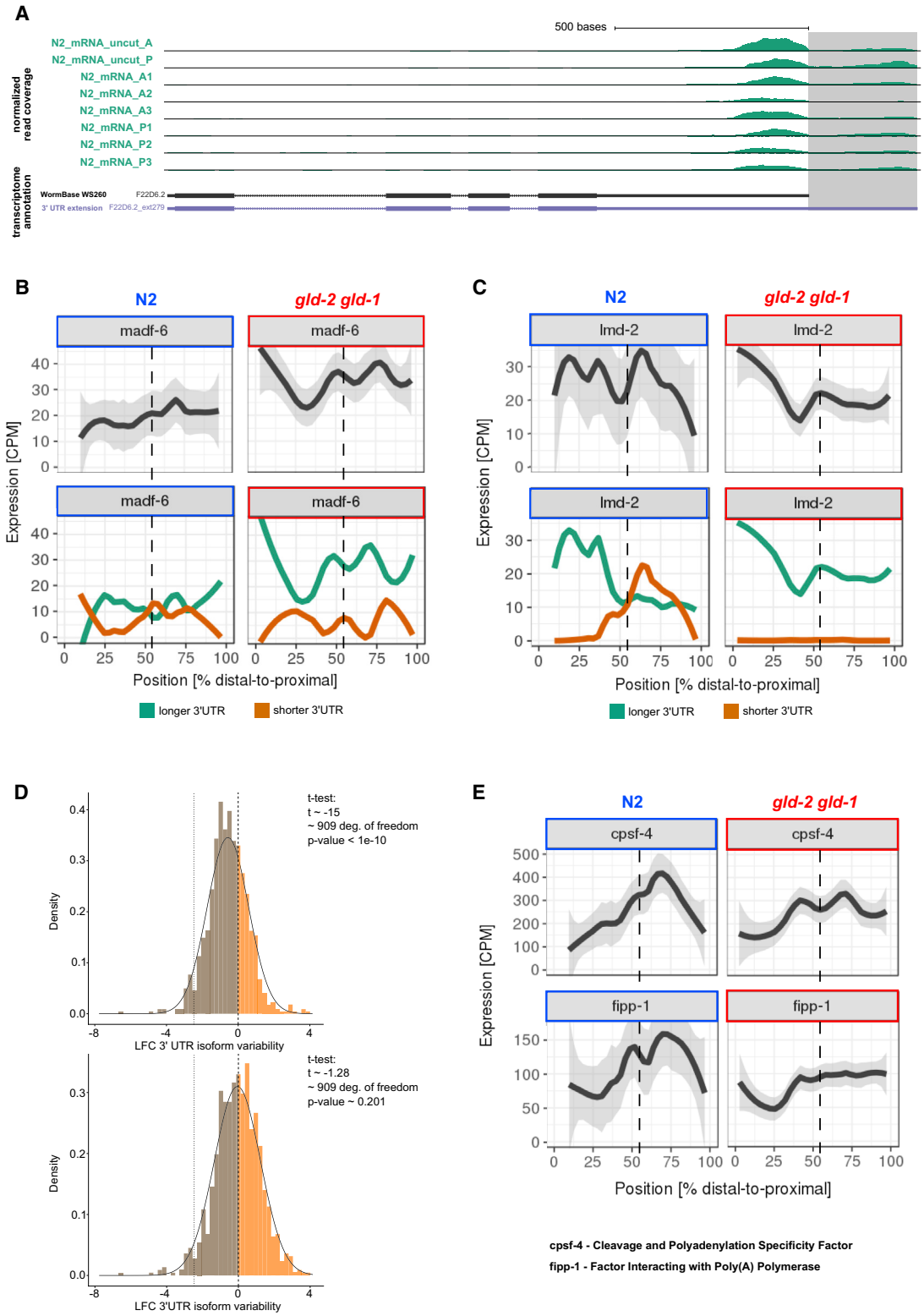


Figure 6. Differential 3′ UTR Isoform Usage in the Germline Is Perturbed in the *gld-2 gld-1* Double Mutant

(A) Genome browser track of an example gene with downstream extension of the annotated 3′ UTR.

(B) Spatial expression of *madf-6* in wild-type N2 and *gld-2 gld-1* double mutant from distal to proximal at gene and isoform level. n = 6 independent experiments for N2, and n = 4 for *gld-2 gld-1*; LOESS ± standard error (SE) for gene level, and LOESS only for isoform level. Longest 3′ UTR is marked in turquoise and shorter 3′ UTR in orange. Dashed line marks bend or loop region of the germline.

(legend continued on next page)

(Figure S6I). Moreover, the wild-type switch of differential isoform usage was generally impaired *gld-2 gld-1* (Figures 6B, 6C, 6D, S6I, S6J, and S6K). The switch did not occur in the *gld-2 gld-1* mutant, but instead, for some genes, only the distal PAS was used throughout the germline. Based on this finding, we explored whether factors involved in regulation of APA are perturbed in the mutant, too. We investigated two factors in more detail, *fipp-1* and *cpsf-4*. *Fipp-1* (human FIP1L1) and *cpsf-4* (human CPSF4L) are components of the cleavage and polyadenylation complex that recognize the canonical PAS (AAUAA) and interact with the poly(A) polymerase and other factors (Kaufmann et al., 2004). Hence, both factors play a key role in APA. We observed that in the wild-type, both factors increased expression toward the proximal arm and peak around the bend region (Figure 6E). This is in line with the switch from distal-to-proximal PAS around the bend region. Strikingly, the expression of both factors did not increase toward the proximal arm in the *gld-2 gld-1* mutant (Figure 6E). This suggests that the level of these factors may be important for the switch from distal-to-proximal PAS usage. Altogether, our data showed that differential 3′ UTR usage is highly regulated in the germline, indicating that the level of certain factors involved in APA may be important for this regulation.

SPACEGERM: A User-Friendly Interface for Exploring Spatial Expression in the Germline

We have shown how the spatiotemporal resolved expression data generated in this study provide new insights into the mechanistic coordination of fundamental processes in biology. Clearly, these data have the potential to inform a multitude of additional studies focusing on various specific biological questions. To enable other researchers to conveniently utilize our data and to provide a "universal" coordinate system, we developed SPACEGERM (spatial *C. elegans* germline expression of mRNA and miRNA), an interactive data visualization tool for exploring the spatiotemporal expression data in the germline (Figure S7). The tool allows the user to investigate the spatiotemporal expression of every gene, isoform, or miRNA detected in our datasets. The user can choose between wild-type and mutant samples and have a closer look at the raw data points or the smooth fits (LOESS) across all replicates. SPACEGERM also allows users to examine a set of genes by uploading an Excel file with gene IDs, again, for every genotype and gene type. Alternatively, one can investigate all genes detected with our sequencing approach, up to 500 genes at once. Furthermore, the user can also download an Excel file with information about genes, their average, their extreme values, locations, and cluster assignment. Finally, the reconstructed 3D germline can be explored concerning *in vivo* RNA expression throughout germ cell proliferation and differentiation ("virtual ISH" [vISH]).

DISCUSSION

By rapidly dissecting, shock-freezing, and cryo-cutting the *C. elegans* germline at 50 μm resolution and sequencing each slice separately, we created a spatially resolved RNA expression map of wild-type and mutant animal germlines. Additionally, we were able to reconstruct an *in silico* 3D germline model that can be used to perform vISH and/or interrogate RNA localization of almost all transcripts during germ cell proliferation and differentiation (Figure 2).

A Mechanistic Model of Spatial Gene Expression Regulation

In this study, we showed that mRNAs are organized spatiotemporally in the germline (Figures 1 and 3). We recovered the expression profile of *rpl* and *rps* genes, which encode for ribosomal protein and are mainly localized to the distal gonad arm and slowly decrease in expression toward the proximal arm (Figure 3) (West et al., 2018). Interestingly, the spatiotemporal gene expression was perturbed in the *gld-2 gld-1* double mutant. In particular, *rpl* and *rps* genes did not decrease in expression in *gld-2 gld-1* mutants but were constantly expressed throughout the germline (Figure 3). The *gld-2 gld-1* double mutant reveals only a third of the meiotic entry, i.e., germ cells fail to differentiate and instead proliferate constantly throughout the germline (Brenner and Schedl, 2016). However, genes involved in the deregulation of the balance between proliferation and differentiation in the *gld-2 gld-1* mutant remain poorly discovered. In this study, we propose that PIE-1 may be one potential key player that regulates the balance between proliferation and differentiation in the *C. elegans* germline (Figure 7A). PIE-1, a maternal CCCH-type zinc-finger protein and a repressor of RNA polymerase II-dependent gene expression that is important for germline cell fate determination, is specific for oocytes and embryos (Merritt et al., 2008; Seydoux and Dunn, 1997; Seydoux et al., 1996; Tenenhaus et al., 2001). A previous study showed that *pie-1* is a target of GLD-2, the main cytoPAP in the germline (Kim et al., 2010). Furthermore, Kim and colleagues showed that depletion of GLD-2 alone was sufficient to lower the abundance of most of its targets as these transcripts do not get polyadenylated and are therefore degraded (Kim et al., 2010). We showed that *pie-1* is mainly localized around the pachytene stage and that it has its highest expression during early oogenesis, exactly where the GLD-2 protein has its highest abundance (Millonigg et al., 2014). In accordance with this fact, the germline becomes transcriptionally silent from the late stage oogenesis (diakinesis)

(C) Spatial expression of *lmd-2* in wild-type N2 and *gld-2 gld-1* double mutant from distal to proximal at gene and isoform level. n = 6 independent experiments for N2, and n = 4 for *gld-2 gld-1*; LOESS ± SE for gene level, and LOESS only for isoform level. Longest 3′ UTR is marked in turquoise and shorter 3′ UTR in orange. Dashed line marks bend or loop region of the germline.

(D) Summary of log-CV (coefficient of variation) changes on a per-gene level between N2 and *gld-2 gld-1* double mutant. Lower panel shows the negative control with shuffled genotype assignments of the CVs. Left dashed line, the 5th percentile of the shuffled control normal fit. The colors refer to higher 3′ UTR isoform variability in N2 (brown) or in *gld-2 gld-1* (orange), respectively.

(E) Spatial expression of *cpsf-4* and *fipp-1* in wild-type N2 and *gld-2 gld-1* double mutant from distal to proximal. n = 6 independent experiments for N2, and n = 4 for *gld-2 gld-1*; LOESS ± SE. Dashed line marks bend or loop region of the germline.

See also Figure S6 and Table S3.

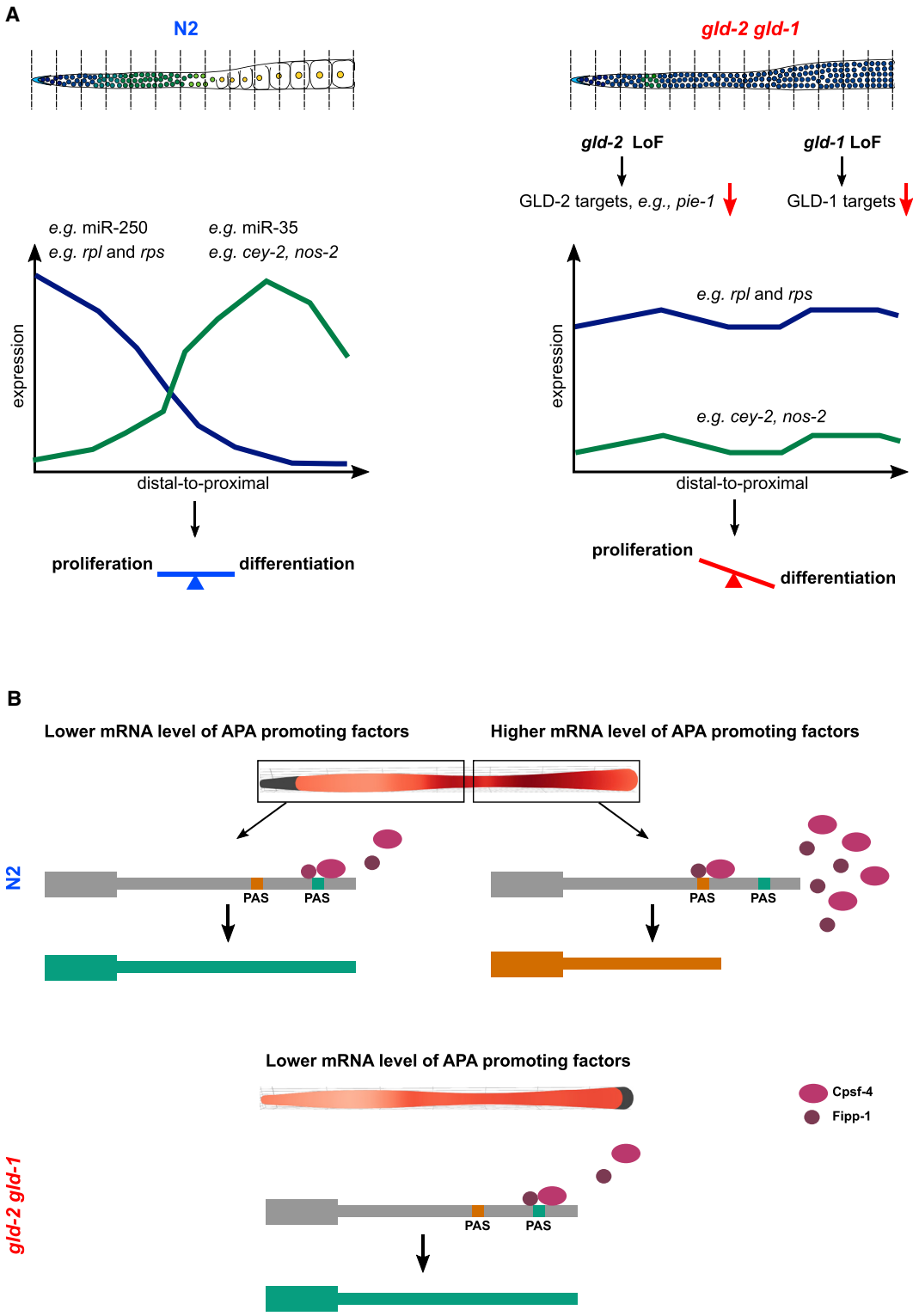


Figure 7. Model for Spatially Restricted Gene Expression and Differential 3′ UTR Isoform Usage in the Germline

(A) Schematic overview of mRNA and miRNA localization in wild-type N2 and mRNA localization in *gld-2 gld-1* double mutant germline, indicating putative regulators of spatially restricted gene expression. LoF, loss of function.

(B) Model for differential 3′ UTR isoform usage across the germline. Depending on the concentration of *cpsf-4* (vISH is shown) and *fipp-1*, two factors involved in alternative polyadenylation (APA), some genes use the longer 3′ UTR isoform in the distal gonad arm, while the shorter one is used in the proximal gonad arm. In *gld-2 gld-1* double mutants, only the longer 3′ UTR isoform is used.

up to the fourth cell-stage embryo (Evsikov et al., 2006; Stoeckius et al., 2014), suggesting that PIE-1 could play a key role in repressing the transcription as it does in the blastomere development. Indeed, many genes such as the *rpl* and *rps* genes decrease in expression toward the proximal gonad arm, supporting the hypothesis that PIE-1 is involved in transcriptional repression of these genes. In line with this hypothesis, when *pie-1* is downregulated in the *gld-2 gld-1* double mutant, *rpl* and *rps* genes are constantly expressed throughout the germline. Additionally, a previous expression study of PIE-1 in HeLa cells reported that PIE-1 can inhibit transcription directly, suggesting a conserved mechanism (Batchelder et al., 1999). As PIE-1 is also detected in the cytoplasm, mainly in association with P granules (Mello et al., 1996; Tenenhaus et al., 2001), it was suggested that PIE-1 is required for the maintenance of *nos-2* and possibly other class II mRNAs, RNAs that are associated with P granules (Tenenhaus et al., 2001). Indeed, our data revealed that *nos-2* and *cey-2*, two examples of class II mRNAs, increase in expression as *pie-1* expression increases in the proximal arm in the wild-type. The same mRNAs are downregulated in the *gld-2 gld-1* mutant. This suggests that the downregulation of these two mRNAs impede differentiation as the *gld-2 gld-1* mutant lacks differentiation, which results in a sterile phenotype. Interestingly, RNA interference (RNAi) of *pie-1* revealed many phenotypes, including a sterile phenotype as in the *gld-2 gld-1* mutant (Melo and Ruvkun, 2012). However, meiotic entry and differentiation are only perturbed in *gld-2 gld-1* double mutants, i.e., meiotic entry appears normally in mutants lacking either *gld-1* or *gld-2* (single mutants) (Brenner and Schedl, 2016). Hence, we cannot exclude other regulators and the regulation via *gld-1* as most of the GLD-1 targets are downregulated as well in our data. Additionally, the *gld-2* single mutant did not show decreased expression of *pie-1* and *cey-2* (Figure S3C). Furthermore, the relationship between *gld-1* in its role in the distal meiotic entry decision and its role further down in the germline is complex as loss of *cye-1* and *cdk-2*, two important regulators of the mitotic cell cycle, causes the germline tumors still to differentiate (Fox et al., 2011).

The Choice of 3′ UTR Is Strongly Regulated in the *C. elegans* Germline

Merritt and colleagues already reported that 3′ UTRs and not promoters are the main drivers of gene expression in the germline (Merritt et al., 2008). Other studies also suggested that proliferating cells use mainly the proximal alternative polyadenylation sginal (PAS; short 3′ UTR), while differentiating cells use predominantly the distal PAS (long 3′ UTR) (Mayr and Bartel, 2009; Sandberg et al., 2008; West et al., 2018). However, *in vivo* studies of differential 3′ UTR isoform usage still remain poorly investigated. Our sequencing approach does not offer the coverage and resolution to generally distinguish between different isoforms of one gene. This is because we sequence approximately 500-nt fragments while the mean 3′ UTR length of *C. elegans* transcripts is 211 nt (Jan et al., 2011; Mangone et al., 2010). Nonetheless, we succeeded in quantifying the change of the (relative) 3′ UTR usage along the germline for almost 1,000 genes. It is important to note that due to our technical limitations, this number is almost certainly only a fraction of all 3′ UTR length switches. Among these candidates, we observed some genes that mainly used the proximal PAS in the distal gonad arm, while the distal PAS was used in the proximal arm (Figures 6 and S6). Our data revealed that *cpsf-4* and *fipp-1* have their highest expression around the pachytene stage. This is exactly where the switch of differential 3′ UTR usage occurs (Figures 6 and S6). Those two genes are known as main regulators of APA. They are probably the key spatial regulators, as other APA factors are less abundant, less localized, and, importantly, not perturbed in the mutants (see SPACEGERM). Furthermore, we discovered that these two factors were constantly expressed in the *gld-2 gld-1* double mutant, leading to the failure of 3′ UTR length switching (Figures 6 and S6). Additionally, the expression levels of the APA factors in the mutant were similar to those in the very distal part of the gonad arm of the wild-type. Therefore, we speculate that the expression level of APA factors is crucial for the differential 3′ UTR usage and hence the switch between proliferation and differentiation during germline development. Lackford and colleagues already observed a similar phenomenon where APA depends on the level of different factors involved in APA such as Fip1 and CPSF (Lackford et al., 2014). It is thought that generally the distal PAS is stronger than the proximal one, leading to the predominant usage of the distal PAS if the level of APA factors is low (Lackford et al., 2014). Thus, we hypothesize that the spatial concentration of factors involved in APA is important for differential 3′ UTR usage along the germline and thereby for controlling proliferation versus differentiation (Figure 7B). Furthermore, the *gld-2 gld-1* mutant indicated that deregulation of the levels of APA factors perturb the differential 3′ UTR usage and therefore may disturb the proliferation and differentiation balance. In general, the *gld-2 gld-1* mutant showed globally decreased 3′ UTR variability in the germline compared to the wild- type. It remains still to be investigated which factors or pathways regulate the levels of the factors involved in APA and if their levels mirror the corresponding spatial protein expression. Furthermore, an approach is needed that increases the resolution to distinguish between different isoforms of a gene in a spatially resolved manner. This will help to determine the total number of genes that follow our differential 3′ UTR usage hypothesis (Figure 7B).

miRNA Expression Is Spatially Organized and Co-localizes with Germline Targets

Besides 3′ UTRs' being important regulators of gene expression in the *C. elegans* germline, previous studies also suggested that miRNAs control proliferation and differentiation in *C. elegans* (Bukhari et al., 2012; Ding et al., 2008). We comprehensively analyzed spatial expression of known miRNAs in the *C. elegans* germline. miRNAs were found in distinct localization patterns along the germline with the miR-35 family, the main miRNA family in the germline, being localized in the pachytene region (Figures 1 and 4), consistent with the results of McEwan and colleagues (McEwen et al., 2016). Furthermore, we showed that all members of the miR-35 family and their targets were co-localized throughout the germline, a requirement for functional interaction *in vivo* (Figure 5A). We speculate about a threshold function of the miR-35 family, where the miRNAs would keep the expression of their targets at a certain level below which protein production is inhibited (Mukherji et al., 2011; Sood et al., 2006). In contrast,

non-germline-specific miRNAs such as miR-1-3p did not show any predominant co-localization with their targets, suggesting an interaction outside of the germline (Figures 5A and 5B).

Identification of 83 Previously Unannotated miRNAs with Specific Spatial Localization

In addition, we discovered 83 previously unannotated precursor miRNAs and validated three of them (Figures 4 and S5). We note that some of these miRNAs have a limited spatial expression domain but are well expressed within this domain. This is probably the reason why they escaped detection in previous studies. Interestingly, we identified an unusual unannotated miRNA, nov-72-3p, that, as it was shown by chimera analysis, binds other miRNAs, creating a miRNA-miRNA duplex (Figure S5F). This phenomenon was predicted by Lai and colleagues in 2004 computationally but so far lacked experimental evidence (Lai et al., 2004). Their hypotheses were that the miRNA duplex could either stabilize the miRNA by protecting it from degradation or the miRNA could be tethered away from its targets and therefore stabilizing the mRNA targets. However, we were not able to define the genomic locus of nov-72-3p as the mature miRNA mapped antisense to the exon of the *dpy-2* locus, but the first 17 nt of the mature miRNA also mapped to ribosomal RNA transcripts. Hence, the locus remains still undetermined, impeding further analysis of nov-72-3p. Finally, we note that besides miRNAs we detected an overwhelming complexity of other small RNAs including endogenous siRNAs (small interfering RNAs) and piRNAs (piwi-interacting RNAs) with interesting spatiotemporal expression patterns. These data will be presented in a future study.

Public Availability of All Data via Interactive Web Application "SPACEGERM"

Finally, we developed an interactive data visualization tool, SPACEGERM, for exploring the spatiotemporal expression in the germline in well defined, "universal" coordinates, both as raw data and projected (vISH) on our 3D model (Figure S7). A complementary resource providing data from whole worm (hermaphrodite and male) cryo-sections is provided in the accompanying paper by Ebbing and collegaues (Ebbing et al., 2018).

Overall, we have presented a map of germline RNA at unprecedented spatial resolution. This near single-cell resolution was key for (1) discovering numerous of previously unannotated miRNAs and hundreds of previously unannotated 3′ UTRs, (2) beginning to interpret the spatial patterns, and (3) identifying, by comparison to mutant germlines, regulators and mechanisms that appear to play key roles in regulating germline biology. We believe that comparison to more mutants will dramatically improve our understanding of this system. Of course, many more measurements will need to be done as we currently only quantify RNA, and, even for RNA, we miss a lot of information—subcellular localization, methylation, polyadenylation states, and many more. However, we hope that our 3D model and data help to set a common reference that can be expanded in the future.

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
  - *C. elegans* Strains and Culture
- METHOD DETAILS
  - Embedding and Cryo-Sectioning
  - mRNA Library Preparation
  - Small RNA Library Preparation
  - Poly(A)+-Selected Library Preparation
  - Ribosomal RNA Depleted Total RNA Library Preparation
  - Probe Preparation for mRNA *In Situ* Hybridization (ISH)
  - mRNA ISH
  - Small RNA ISH
  - TaqMan® Assays
  - Nested PCR
  - Physical Gonad Model
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data Pre-processing
  - Mapping of Reads to the *C. elegans* Genome
  - 3' Extension of Transcript Annotation
  - Transcriptome Pre-processing
  - Isoform-Specific Transcript Abundance Estimation
  - Data Processing
  - Aligning Cryo-Cuts of Different Samples to a Single Coordinate System
  - Integration of Replicate Data
  - Data Analysis of the Small RNA Transcriptome
- DATA AND SOFTWARE AVAILABILITY
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at https://doi.org/10.1016/j.devcel.2018.10.005.

ACKNOWLEDGMENTS

We thank J.P. Junker and B. Tursun for helpful discussions. We also thank J.P. Junker for helping us to implement the tomo-seq technique. We thank the whole Rajewsky lab for discussions and support. We are grateful to M. Herzog for helping with the gonad dissection. We thank J. Hubbard and R. Ketting for helpful discussions. We thank A. Filipchyk for identifying the chimeric interactions of nov-72-3p in the small RNA sequencing libraries. We thank Jonathan Fröhlich for critically reading the manuscript and for helpful comments. A.D. was a member of the Computational Systems Biology (CSB) graduate school, which is funded by Deutsche Forschungsgemeinschaft (DFG). M.S. was funded by DFG and MDC. F.K. was funded by Deutsches Epigenom Programm (DEEP).

AUTHOR CONTRIBUTIONS

A.D. and N.R. conceived and designed the project. A.D. established and led the development of the project. A.D. designed and performed the experiments and wrote (with N.R.) the manuscript, with input from other authors. M.S. analyzed the mRNA data, designed the interactive data visualization tool, and constructed the 3D germline model. F.K. analyzed the small RNA data. S.A. helped with the establishment of the small RNA protocol. A.D. and N.R. led the interpretation of the data. N.R. supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 21, 2018
Revised: August 31, 2018
Accepted: October 4, 2018
Published: November 8, 2018

REFERENCES

Adiconis, X., Borges-Rivera, D., Satija, R., DeLuca, D.S., Busby, M.A., Berlin, A.M., Sivachenko, A., Thompson, D.A., Wysoker, A., Fennell, T., et al. (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. Nat. Methods *10*, 623–629.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq— a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215–233.

Bartel, D.P. (2018). Metazoan microRNAs. Cell *173*, 20–51.

Batchelder, C., Dunn, M.A., Choy, B., Suh, Y., Cassie, C., Shim, E.Y., Shin, T.H., Mello, C., Seydoux, G., and Blackwell, T.K. (1999). Transcriptional repression by the Caenorhabditis elegans germ-line protein PIE-1. Genes Dev. *13*, 202–212.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

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

Brenner, J.L., and Schedl, T. (2016). Germline stem cell differentiation entails regional control of cell fate regulator GLD-1 in Caenorhabditis elegans. Genetics *202*, 1085–1103.

Brumbaugh, J., Di Stefano, B., Wang, X., Borkent, M., Forouzmand, E., Clowers, K.J., Ji, F., Schwarz, B.A., Kalocsay, M., Elledge, S.J., et al. (2018). Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. Cell *172*, 106–120.e21.

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

Bystrykh, L.V. (2012). Generalized DNA barcode design based on Hamming codes. PLoS One *7*, e36852.

Crittenden, S.L., Leonhard, K.A., Byrd, D.T., and Kimble, J. (2006). Cellular analyses of the mitotic region in the Caenorhabditis elegans adult germ line. Mol. Biol. Cell *17*, 3051–3061.

Dard-Dascot, C., Naquin, D., d'Aubenton-Carafa, Y., Alix, K., Thermes, C., and van Dijk, E. (2018). Systematic comparison of small RNA library preparation protocols for next-generation sequencing. BMC Genomics *19*, 118.

Ding, X.C., Slack, F.J., and Grosshans, H. (2008). The let-7 microRNA interfaces extensively with the translation machinery to regulate cell differentiation. Cell Cycle *7*, 3083–3090.

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.

Dodt, M., Roehr, J.T., Ahmed, R., and Dieterich, C. (2012). FLEXBAR-flexible barcode and adapter processing for next-generation sequencing platforms. Biology (Basel) *1*, 895–905.

Ebbing, A., Vértesy, A., Betist, M.C., Spanjaard, B., Junker, J.P., Berezikov, E., van Oudenaarden, A., and Korswagen, H.C. (2018). Spatial transcriptomics of *C. elegans* males and hermaphrodites identifies sex-specific differences in gene expression patterns. Dev. Cell https://doi.org/10.1016/j.devcel.2018.10.016.

Evsikov, A.V., Graber, J.H., Brockman, J.M., Hampl, A., Holbrook, A.E., Singh, P., Eppig, J.J., Solter, D., and Knowles, B.B. (2006). Cracking the egg: molecular dynamics and evolutionary aspects of the transition from the fully grown oocyte to embryo. Genes Dev. *20*, 2713–2727.

Fox, P.M., Vought, V.E., Hanazawa, M., Lee, M.H., Maine, E.M., and Schedl, T. (2011). Cyclin E and CDK-2 regulate proliferative cell fate and cell cycle progression in the C. elegans germline. Development *138*, 2223–2234.

Friedländer, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. *40*, 37–52.

Gartner, A., Boag, P.R., and Blackwell, T.K. (2008). Germline Survival and Apoptosis.WormBook, 1–20 [https://doi.org/10.1895/wormbook.1.145.1].

Grosswendt, S., Filipchyk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. Mol. Cell *54*, 1042–1054.

Hansen, D., and Schedl, T. (2013). Stem cell proliferation versus meiotic fate decision in Caenorhabditis elegans. Adv. Exp. Med. Biol. *757*, 71–99.

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep. *2*, 666–673.

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K.J., Rozenblatt-Rosen, O., et al. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-seq. Genome Biol. *17*, 77.

Hirsh, D., Oppenheim, D., and Klass, M. (1976). Development of the reproductive system of Caenorhabditis elegans. Dev. Biol. *49*, 200–219.

Hubbard, E.J.A. (2007). Caenorhabditis elegans germ line: a model for stem cell biology. Dev. Dyn. *236*, 3343–3357.

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of Caenorhabditis elegans 3' UTRs. Nature *469*, 97–101.

Jansen, R.P. (2001). mRNA localization: message on the move. Nat. Rev. Mol. Cell Biol. *2*, 247–256.

Jens, M., and Rajewsky, N. (2015). Competition between target sites of regulators shapes post-transcriptional gene regulation. Nat. Rev. Genet. *16*, 113–126.

Jungkamp, A.C., Stoeckius, M., Mecenas, D., Grün, D., Mastrobuoni, G., Kempa, S., and Rajewsky, N. (2011). In vivo and transcriptome-wide identification of RNA binding protein target sites. Mol. Cell *44*, 828–840.

Junker, J.P., Noël, E.S., Guryev, V., Peterson, K.A., Shah, G., Huisken, J., McMahon, A.P., Berezikov, E., Bakkers, J., and van Oudenaarden, A. (2014). Genome-wide RNA tomography in the zebrafish embryo. Cell *159*, 662–675.

Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. EMBO J. *23*, 616–626.

Kim, K.W., Wilson, T.L., and Kimble, J. (2010). GLD-2/RNP-8 cytoplasmic poly(A) polymerase is a broad-spectrum regulator of the oogenesis program. Proc. Natl. Acad. Sci. USA *107*, 17445–17450.

Koller, M., and Stahel, W.A. (2011). Sharpening Wald-type inference in robust regression for small samples. Comput. Stat. Data Anal. *55*, 2504–2515.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics *23*, 1289–1291.

Lackford, B., Yao, C., Charles, G.M., Weng, L., Zheng, X., Choi, E.A., Xie, X., Wan, J., Xing, Y., Freudenberg, J.M., et al. (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. EMBO J. *33*, 878–889.

Lai, E.C., Wiel, C., and Rubin, G.M. (2004). Complementary miRNA pairs suggest a regulatory role for miRNA:miRNA duplexes. RNA *10*, 171–175.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

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.

Maciejowski, J., Ugel, N., Mishra, B., Isopi, M., and Hubbard, E.J.A. (2006). Quantitative analysis of germline mitosis in adult C. elegans. Dev. Biol. *292*, 142–151.

Mangone, M., Manoharan, A.P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S.D., Mis, E., Zegar, C., Gutwein, M.R., Khivansara, V., et al. (2010). The landscape of C. elegans 3'UTRs. Science *329*, 432–435.

Mayr, C. (2017). Regulation by 3'-untranslated regions. Annu. Rev. Genet. *51*, 171–194.

Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3' UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673–684.

McEwen, T.J., Yao, Q., Yun, S., Lee, C.Y., and Bennett, K.L. (2016). Small RNA in situ hybridization in Caenorhabditis elegans, combined with RNA-seq, identifies germline-enriched microRNAs. Dev. Biol. *418*, 248–257.

Mello, C.C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J.R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. Nature *382*, 710–712.

Melo, J.A., and Ruvkun, G. (2012). Inactivation of conserved C. elegans genes engages pathogen- and xenobiotic-associated defenses. Cell *149*, 452–466.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3' UTRs are the primary regulators of gene expression in the C. elegans germline. Curr. Biol. *18*, 1476–1482.

Millonigg, S., Minasaki, R., Nousch, M., Novak, J., and Eckmann, C.R. (2014). GLD-4-mediated translational activation regulates the size of the proliferative germ cell pool in the adult C. elegans germ line. PLoS Genet. *10*, e1004647.

Miska, E.A., Alvarez-Saavedra, E., Abbott, A.L., Lau, N.C., Hellman, A.B., McGonagle, S.M., Bartel, D.P., Ambros, V.R., and Horvitz, H.R. (2007). Most Caenorhabditis elegans microRNAs are individually not essential for development or viability. PLoS Genet. *3*, e215.

Mukherji, S., Ebert, M.S., Zheng, G.X.Y., Tsang, J.S., Sharp, P.A., and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. Nat. Genet. *43*, 854–859.

Nousch, M., and Eckmann, C.R. (2013). Translational control in the Caenorhabditis elegans germ line. Adv. Exp. Med. Biol. *757*, 205–247.

Nousch, M., Minasaki, R., and Eckmann, C.R. (2017). Polyadenylation is the key aspect of GLD-2 function in C. elegans. RNA *23*, 1180–1187.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rybak-Wolf, A., Jens, M., Murakawa, Y., Herzog, M., Landthaler, M., and Rajewsky, N. (2014). A variety of dicer substrates in human and C. elegans. Cell *159*, 1153–1167.

Sambrook, J., and Russell, D.W. (2006). Rapid amplification of 3' cDNA ends (3'-RACE). CSH Protoc., 2006.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science *320*, 1643–1647.

Seydoux, G., and Dunn, M.A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. Development *124*, 2191–2201.

Seydoux, G., Mello, C.C., Pettitt, J., Wood, W.B., Priess, J.R., and Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of C. elegans. Nature *382*, 713–716.

Shepard, P.J., Choi, E.A., Lu, J., Flanagan, L.A., Hertel, K.J., and Shi, Y. (2011). Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA *17*, 761–772.

Sood, P., Krek, A., Zavolan, M., Macino, G., and Rajewsky, N. (2006). Cell-type-specific signatures of microRNAs on target mRNA expression. Proc. Natl. Acad. Sci. USA *103*, 2746–2751.

Stoeckius, M., Grün, D., Kirchner, M., Ayoub, S., Torti, F., Piano, F., Herzog, M., Selbach, M., and Rajewsky, N. (2014). Global characterization of the oocyte-to-embryo transition in Caenorhabditis elegans uncovers a novel mRNA clearance mechanism. EMBO J. *33*, 1751–1766.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032–2034.

Tenenhaus, C., Subramaniam, K., Dunn, M.A., and Seydoux, G. (2001). PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of Caenorhabditis elegans. Genes Dev. *15*, 1031–1040.

Theil, K., Herzog, M., and Rajewsky, N. (2018). Post-transcriptional regulation by 3′ UTRs can be masked by regulatory elements in 5′ UTRs. Cell Rep. *22*, 3217–3226.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3–new capabilities and interfaces. Nucleic Acids Res. *40*, e115.

Wang, X., Zhao, Y., Wong, K., Ehlers, P., Kohara, Y., Jones, S.J., Marra, M.A., Holt, R.A., Moerman, D.G., and Hansen, D. (2009). Identification of genes expressed in the hermaphrodite germ line of C. elegans using SAGE. BMC Genomics *10*, 213.

West, S.M., Mecenas, D., Gutwein, M., Aristizábal-Corrales, D., Piano, F., and Gunsalus, K.C. (2018). Developmental dynamics of gene expression and alternative polyadenylation in the Caenorhabditis elegans germline. Genome Biol. *19*, 8.

Wolke, U., Jezuit, E.A., and Priess, J.R. (2007). Actin-dependent cytoplasmic streaming in C. elegans oogenesis. Development *134*, 2227–2236.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli, OP50	Caenorhabditis Genetics Center	WormBase ID: OP50; RRID: WB-STRAIN:OP50
Chemicals, Peptides, and Recombinant Proteins		
Affi-Gel Blue Gel	Bio-Rad	Cat#1537301
DIG RNA Labeling mix	Roche	Cat#11277073910
Anti-Digoxigenin-AP, Fab fragments	Roche	Cat#11093274910
NBT/BCIP Stock Solution	Roche	Cat#11681451001
ERCC RNA Spike-In Mix	Thermo Fisher Scientific	Cat#4456740
Critical Commercial Assays		
MessageAmp II aRNA Amplification Kit	Thermo Fisher Scientific	Cat#AM1751
Agencourt AMPure XP	Beckman Coulter	Cat#A63881
Agencourt RNAClean XP	Beckman Coulter	Cat#A63987
Custom TaqMan MicroRNA assay	Thermo Fisher Scientific	Cat#4398987
TruSeq small RNA Kit	Illumina	Cat#RS-200-0012
SMARTer smRNA-Seq Kit	Clontech	Cat#635031
TruSeq-stranded mRNA LT Kit	Illumina	Cat#RS-122-2102
NucleoSpin Gel and PCR Clean-up Kit	Macherey-Nagel	Cat#740609.250
Deposited Data		
Raw and analyzed data	This paper	GEO: GSE115884
Experimental Models: Organisms/Strains		
N2 *C. elegans* wild isolate	Caenorhabditis Genetics Center	WormBase ID: N2; RRID: WB-STRAIN:N2_(ancestral)
C. elegans: Strain: JK2879: *gld-2(q497) gld-1(q485)/hT2 [bli-4(e937) let-?(q782) qIs48]* (I;III)	Caenorhabditis Genetics Center	WormBase ID: JK2879; RRID: WB-STRAIN:JK2879
C. elegans: Strain: JK1743: *gld-2(q497)/ dpy-5(e61 unc-13(e-51) I*	Caenorhabditis Genetics Center	WormBase ID: JK1743; RRID: WB-STRAIN:JK1743
C. elegans: Strain: GC833: *glp-1(ar202) III*	Caenorhabditis Genetics Center	WormBase ID: GC833; RRID: WB-STRAIN:GC833
Oligonucleotides		
ISH primer, see Table S4	This paper	N/A
DIG-labeled LNA probes, see Table S4	Qiagen	N/A
TaqMan probes, see Table S4	Thermo Fisher Scientific	N/A
Nested PCR primer, see Table S4	Sambrook and Russell (2006) and This paper	N/A
Conventional PCR primer, see Table S4	This paper	N/A
Barcoded unanchored oligo(dT) primer, see Table S5	This paper	N/A
Software and Algorithms		
flexbar v2.5	Dodt et al. (2012)	https://github.com/seqan/flexbar/releases/tag/v2.5.0
STAR_2.5.1b	Dobin et al. (2013)	https://github.com/alexdobin/STAR/releases/tag/2.5.1b
sambamba v0.4.7	Tarasov et al. (2015)	https://github.com/biod/sambamba/releases/tag/v0.4.7
bedtools v2.23.0	Quinlan and Hall (2010)	https://github.com/arq5x/bedtools2/releases/tag/v2.23.0
kallisto 0.43.1	Bray et al. (2016)	https://github.com/pachterlab/kallisto/releases/tag/v0.43.1
bowtie2 (version 2.3.3.1)	Langmead and Salzberg (2012)	https://github.com/BenLangmead/bowtie2/releases/tag/v2.3.3.1
STAR (version 2.5.3a)	Dobin et al. (2013)	https://github.com/alexdobin/STAR/releases/tag/2.5.3a

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HTSeq (version 0.9.1)	Anders et al. (2015)	https://github.com/simon-anders/htseq/releases/tag/release_0.9.1
miRDeep2 (version 2.0.0.7)	Friedländer et al. (2012)	https://www.mdc-berlin.de/media/16706
samtools 0.1.19-96b5f2294a	Li et al. (2009)	http://archive.ubuntu.com/ubuntu/pool/universe/s/samtools/samtools_0.1.19.orig.tar.gz
Other		
Resource Website: SPACEGERM	This paper	https://shiny.mdc-berlin.de/spacegerm

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, Nikolaus Rajewsky (rajewsky@mdc-berlin.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

C. elegans Strains and Culture

All *C. elegans* strains were cultured by standard techniques (Brenner, 1974). Worms were maintained at 16°C on *E. coli* OP50-seeded nematode growth medium (NGM) plates. The following strains were used in this study: N2 Bristol wild type, *gld-2(q497) gld-1(q485)/hT2 [bli-4(e937) let-?(q782) qIs48]* (I;III), *gld-2(q497)/dpy-5(e61 unc-13(e-51) I* and *glp-1(ar202) III.* For all experiments, young adult or L4 staged hermaphrodites were used.

METHOD DETAILS

Embedding and Cryo-Sectioning

Gonads of wild type and mutants were dissected according to Francis and Nayak (Schedl lab) by cutting off the head or the tail with a syringe needle to extrude the anterior or posterior gonad arm, respectively. The gonad, still attached to the worm body, was transferred to a specimen mold (Tissue-Tek® cryomold®) filled with tissue freezing medium. This medium is very viscos, facilitating the stretching of the gonad and the separation from the worm body. Once the gonad was stretched, distal tip end and proximal end (at the end of oogenesis) were marked with AffiGel® blue beads (Bio-Rad). Following, the specimen mold was rapidly frozen at -80°C for 1 min and subsequently fixed in the cryotome to cut the gonad into slices of desired resolution (here 50 μm). Each slice of the gonad was collected in an individual LoBind Eppendorf® tube and immediately transferred to dry ice. RNA extraction of each slice was performed according Junker et al. (2014) with minor modifications. Briefly, 500 μl of self-made trizol were added to each slice including 0.5 μl GlycoBlue and 2 μl ERCC Spike-In mix (1:10,000-1:50,000). Samples were mixed thoroughly and incubated for 5 min at room temperature. Subsequently, 100 μl of chloroform were added to each sample, mixed well, incubated for 5 min at room temperature and centrifuged at 12,000 g for 15 min at 4°C. Following, the aqueous phase (~250-300 μl) was carefully transferred to a new LoBind Eppendorf® tube and 250-300 μl of isopropanol were added. Samples were mixed thoroughly and incubated over night at -20°C. Afterwards, samples were centrifuged for 10 min at 12,000 g and at 4°C, supernatant was removed and RNA pellet (should appear blue) was washed with 75% ethanol. Following the last centrifugation step at 7,500 g for 5 min and at 4°C, supernatant was removed and pellet was either resuspended in 1.2 μl barcoded oligo(dT) primer in order to proceed with mRNA library preparation or in 7 μl nuclease free water for small RNA library preparation. All experiments were performed in biological and technical triplicates for wild type and replicates for mutants for each gonad arm, i.e., anterior and posterior gonad arm.

mRNA Library Preparation

Reverse transcription and *in vitro* transcription (IVT) were performed with the Ambion™ MessageAmp™ II kit according to the CEL-seq method (Hashimshony et al., 2012) and the tomo-seq method (Junker et al., 2014), except that all purification steps were performed using Agencourt® AMPure® or RNAClean® XP beads according to CEL-seq2 minimizing loss of material (Hashimshony et al., 2016). In brief, after resuspending each slice in 1.2 μl barcoded oligo(dT) primer, first strand synthesis was performed for 2 h at 42°C. Following, slices were pooled and second strand synthesis was performed for 2 h at 16°C. The resulting cDNA was purified with Agencourt® AMPure® XP beads according manufacturer's instructions. Subsequently, IVT was performed for 13 h at 37°C to amplify the RNA which afterwards was fragmented and purified with RNAClean® XP beads according manufacturer's instructions. Library preparation was performed using the Illumina TruSeq® small RNA kit following the tomo-seq protocol (Junker et al., 2014). This included phosphatase and polynucleotide kinase (PNK) treatment, purification with RNAClean® XP beads, 3' adapter ligation, reverse transcription, PCR amplification and final purification with Agencourt® AMPure® XP beads. Unlike CEL-seq1/2 and tomo-seq, unanchored oligo(dT) barcodes used in this study were designed according to a Hamming [8,4] code allowing for barcode

correction after sequencing (Table S5) (Bystrykh, 2012). Anchored oligo(dT) barcodes according to CEL-seq and tomo-seq (Hashimshony et al., 2012; Junker et al., 2014) were used for uncut samples (N2_mRNA_uncut_A/P) and the first replicates of cut anterior and posterior gonad arm samples (N2_mRNA_A1/P1). Libraries (with 20%-30% of PhiX spike-in DNA) were sequenced on the NextSeq 500 in a paired end mode. Except for samples N2_mRNA_A1/P1 and N2_mRNA_uncut_A/P, the number of cycles for read 1 was decreased to sequence barcode and UMI only.

Small RNA Library Preparation

Small RNA libraries were only performed for the wild type N2 strain. Library preparation was performed for each slice separately using the SMARTer smRNA-Seq kit for Illumina from Clontech® according to manufacturer's instruction. Briefly, total RNA was extracted from each slice as described above. Following, total RNA was polyadenylated using a synthetic poly(A) polymerase. Subsequently, cDNA was synthesized and amplified according manufacturer's instructions. The PCR product was purified using the NucleoSpin Gel and PCR Clean-Up kit according manufacturer's instructions. Finally, size selection was performed with the BluePippin system (range: 148 bp-185 bp) in order to enrich for small RNAs. The small RNA libraries were pooled and sequenced on HiSeq 2500, 1 x 50 cycles.

Poly(A)+-Selected Library Preparation

For the poly(A)+-selected library, several gonads were dissected and pooled. Library preparation was performed with the Illumina TruSeq® stranded mRNA LT kit according to manufacturer's instruction. Sequencing was performed on the NextSeq 500.

Ribosomal RNA Depleted Total RNA Library Preparation

For the ribosomal RNA depleted (ribodepleted) total RNA library several gonads were dissected and pooled. Ribosomal RNA (rRNA) was depleted using an RNase H-based protocol adapted from Adiconis et al. (2013). The extracted total RNA from dissected gonads was mixed with the same amount of a DNA oligonucleotide pool consisting of a 50 nt long oligonucleotide mix covering the reverse complement of the entire length of each *C. elegans* rRNA (5S rRNA, 5.8S rRNA, 18S rRNA, 26S rRNA). Following, samples were incubated with 1 U of RNase H for 30 min at 45 °C and purified with RNAClean® XP beads according manufacturer's instructions. The purified RNA was treated with DNase according to the TURBO DNase rigorous treatment procedure (Thermo Fisher Scientific). After the final purification step with RNAClean® XP beads, rRNA-depleted RNA samples were fragmented and library preparation was performed with Illumina TruSeq® stranded mRNA LT kit. Paired end sequencing was performed on the NextSeq 500.

Probe Preparation for mRNA *In Situ* Hybridization (ISH)

Digoxigenin (DIG)-labeled anti-sense RNA probes were prepared by *in vitro* transcription using a PCR generated DNA template. PCR primers were designed using Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012) to amplify a 300-500 nt fragment from the cDNA prepared from whole worm samples. The T7 promoter sequence was added to the reverse primer to produce later an antisense probe by *in vitro* transcription. Primer sequences are provided in Table S4. PCR fragments were cleaned-up using Agencourt® AMPure® XP beads according to manufacturer's instructions and *in vitro* transcription was performed with 0.5-1 μg DNA template using the T7 RNA polymerase and a DIG-RNA labeling mix (Roche). Remaining DNA template was digested with DNase I and the RNA probe was precipitated with sodium acetate and ethanol for at least 30 min at -80°C. After centrifugation the RNA pellet was washed with 75% ethanol and probe integrity was checked on an agarose gel. The concentration of each RNA probe was adjusted to 50 ng/μl using 10 mM Tris-HCl/formamide solution (1:1).

mRNA ISH

Worms were washed several times in sperm salt buffer only (100 mM PIPES, pH 7.0; 90 mM NaCl; 50 mM KCl; 40 mM $CaCl_2$; 20 mM KH_2PO_4) and in the final step in sperm salt buffer containing levamisole. Up to 15 worms were transferred to a poly-L-lysine coated slide containing 8 μl of sperm salt and gonads were dissected as described above. After dissection 8 μl of 4% paraformaldehyde (PFA) were added to the dissected gonads, a cover slip was put on top and the slide was incubated for 2 min. Following, the slide was incubated on dry ice for at least 20 min and the coverslip was flipped away using a razor blade under the coverslip (freeze and crack method). Slides were immediately immersed in ice-cold 100% ethanol for 2 min, rehydrated in an ethanol series (90%, 70%, 50%, 20%), following washing with PBS containing 0.2% Tween 20 for 30 min. Permeabilization of gonads was achieved with proteinase K treatment (1 μg/ml) for 5 min. Slides were washed in PBS containing 0.1% Tween 20 (PBS-T), fixed for 20 min in 4% PFA, washed again with PBS-T, incubated in TEA buffer (aqua dest. containing 1.3% triethanolamine and 0.25% acetic anhydride; always prepared fresh), following final washing steps in PBS-T. Slides were prehybridized in prehybridization buffer (10 mM HEPES, pH 7.5; 600 mM NaCl; 50 mM DTT; 1 mM EDTA; 1 x Denhardt's solution; 100 μg/ml tRNA; 50% formamide) for 1 h at 50°C. Slides were hybridized over night at 50°C in hybridization buffer (prehybridization buffer containing 10% dextran sulphate) containing 0.5-1 μg/ml denaturated DIG-labeled antisense RNA probe (denaturation at 95°C for 10 min). Slides were washed at 50°C for 10 min with following solutions: posthybridization buffer (posthyb, 5 x SSC containing 50% formamide); 75% posthyb buffer + 25% 2 x SSC, 0.1% Triton X; 50% posthyb buffer + 50% 2 x SSC, 0.1% Triton X; 25% posthyb buffer + 75% 2 x SSC, 0.1% Triton X; 2 x SSC, 0.1% Triton X; 0.22 x SSC, 0.1% Triton X. Following, slides were washed in maleic acid buffer (11.6 g/l maleic acid; 9.76 g/l NaCl; 0.1% Triton X; pH 7.5) and afterwards incubated in 1% blocking solution (Roche) diluted in maleic acid buffer for 1 h. Slides were incubated in Anti-DIG-AP (Roche, 1:2500) over night at 4°C. After several washes with maleic acid buffer and

TMN buffer (0.1 M Tris-HCl, pH 9.5; 0.1 M NaCl; 50 mM $MgCl_2$; 1% Tween 20; always prepared fresh), the signal was developed using NBT/BCIP (diluted in TMN buffer) solution. Time of development depended on the expression of the corresponding RNA and took up to 24 h for very lowly expressed RNAs. The background was removed with dehydration and rehydration in an ethanol series (samples were fixed before in 4% PFA for 20 min again). For mounting, some μl of prolong gold (Invitrogen) were dropped on a coverslip and then inverted onto the slide. The edges were sealed with nail polish.

Small RNA ISH

Gonad preparation and the prehybridization procedure was the same as for mRNA ISH. The TEA buffer contained additionally 0.06 N HCl and 0.27 % acetic anhydride. For small RNA ISH, DIG-labeled LNA (Locked Nucleic Acid) probes (former: Exiqon, now: Qiagen) were used (Table S4) and the prehybridization and hybridization temperature was set according to manufacturer's instruction (20-25 °C below the melting temperature of the LNA probe). LNA probes were denaturated at 95°C for 1-5 min prior hybridization. Prehybridization (without probe) was done for 1 h and hybridization (with 10-25 nM of LNA probe) over night. Slides were washed several times with 2 x SSC buffer and with 0.2 x SSC buffer. Following, the slides were washed with PBS-T and incubated for 1 h in blocking solution (PBS-T containing 5% normal goat serum). Slides were incubated in Anti-DIG-AP (Roche, 1:2,000) over night at 4 °C. After several washes with PBS-T and TMN buffer, signal developing, background removal and mounting was performed according to mRNA ISH.

TaqMan® Assays

The TaqMan® assay was used to validate some of the unannotated miRNA predictions. TaqMan® probes were designed with the Custom TaqMan® Small RNA Assay Design Tool (ThermoFisher). TaqMan® assays were performed according manufacturer's instruction for gonad and whole worm samples. TaqMan® target sequences are provided in Table S4.

Nested PCR

Nested PCR was performed according to the Cold Spring Harbor Protocols (Sambrook and Russell, 2006). 0.5-1 μg of whole worm and gonad RNA were used as input RNA for the cDNA synthesis using TAP-VN as a primer. For the first nested PCR, 4 μl of 1:5 diluted cDNA was used. The first PCR was performed with a gene-specific forward primer and AP as a reverse primer. PCR products were purified using Agencourt® AMPure® XP beads and 10-20 ng of purified PCR were used for the second nested PCR. The second PCR was performed with a second gene-specific primer and MAP as a reverse primer. Annealing temperature was calculated using the NEB Tm Calculator (BioLabs). The PCR products from the second PCR were separated by agarose gel, purified and Sanger-sequenced to confirm the identity of the bands. Nested PCR was used for 3' UTR extension validation. Alternatively, conventional PCR by designing the forward primer in the second last exon (to distinguish from genomic DNA) and the reverse primer in the 3' UTR extension was used for validation (using whole worm RNA only). Primer sequences are provided in Table S4.

Physical Gonad Model

To be able to assign the relative distal-to-proximal coordinates used for the spatially resolved gene expression profiles, a physical model of the *C. elegans* germline was built using a custom R script. The following assumptions were made for that model: i) Cells are approximately spherical. ii) Germ cells form a single layer tube within the distal part of the gonad arm. iii) The diameter of the gonad is minimal under the constraint of encompassing all germ cells. This enables a direct conversion between the number of cells in a germ cell layer and the diameter of that cell layer (given the size of a single germ cell) using basic geometry:

$$\phi_G(d_L(l)) := \frac{\phi_g}{2\sin\left(\frac{\pi}{N_{g_L}(l)}\right)} \cdot 2 + \phi_g = \left(1 + \left[\sin\left(\frac{\pi}{N_{g_L}(l)}\right)\right]^{-1}\right)\phi_g,$$

were ϕ_g is the diameter of a germ cell, $l \in \mathbb{N}^+$ is the germ cell layer (one-based), $N_{g_L}(l)$ is the number of germ cells in layer l, $d_L(l)$ is the distance of the center of layer l to the distal tip cell (DTC) ($d_L(l) := (l - (1/2))\phi_g$) and $\phi_G(d)$ is the diameter of the gonad arm at distance d from the DTC.

The (modelled constant) diameter of a single germ cell was set to 4.6 μm (Maciejowski et al., 2006). Based on our own measurements and results by Hirsh and colleagues (Hirsh et al., 1976) the total length of a stretched-out gonad arm was defined as 650 μm. At this distance to the distal tip cell (DTC) (*i.e.,* at the proximal end), the gonad must fit a fully mature oocyte, while at the distal-most end only a single germ cell needs to be fit in the gonad arm. To get a rough estimate of the size of a fully matured oocyte, the number of cells per embryo (558) (Wolke et al., 2007) was multiplied with the volume of a single germ cell. Given the equality in diameter of embryonic cells and germ cells and the equality in volume of the mature oocyte and the embryo, this gives a direct estimate for the size of the oocyte. To be able to model the gonad diameter in-between those extreme boundaries, we measured four gonad arms based on microscopic images (Table S1). Using these measurements at discrete points, a spline fit was used to model the radius of the gonad

arm as a function of the distance to the DTC. Using this fit, the outline of the stretched-out gonad arm was modeled as a solid of revolution around the distal-to-proximal axis:

$$v_G(d_s, d_e) = \pi \int_{d_s}^{d_e} r_G(d)^2 \mathrm{d}d,$$

where d_s and d_e denote the distance to the DTC of the start and the end of the interval of interest, respectively, and $v_G(d_s, d_e)$ is the volume of the corresponding part of the gonad arm.

Based on the assumptions introduced above, the distal arm was filled with 1,002 germ cells in layers maximizing the number of cells per layer under the constraint given by the corresponding gonad diameter:

$$n_{g_L}(l) := \begin{cases} 0 & \text{if } \phi_L(l) < \phi_g \\ 1 & \text{if } \phi_L(l) < 2\phi_g \\ \left\lfloor \dfrac{\pi}{\arcsin\left(\dfrac{\phi_g}{\phi_L(l) - \phi_g}\right)} \right\rfloor & \text{otherwise} \end{cases},$$

where $l \in \{1, (\lfloor d_B / w_L \rfloor)\} \subset \mathbb{N}^+$ ($d_B = 389.3\ \mu m$ representing the distance to the DTC of the bend and $w_L := \phi_g$ the width of a germ cell layer) is the germ cell layer of interest (one-based), $n_{g_L}(l)$ is the number of germ cells in that layer, and $\phi_L(l)$ is the minimal diameter of the gonad in the interval containing germ cell layer l:

$$\phi_L(l) := \phi_{G_L}(l) = \min_{d \in [(l-1)w_L, lw_L] \subset \mathbb{R}} 2r_G(d),$$

where $r_G(d)$ is the radius of the gonad at distance d from the DTC according to the spline model.

The total number of distal germ cells was derived from the total number of distal germ cell layers which was determined by comparing the cumulative number of cells up to each potential layer to the expected number of germ cells $\tilde{N}_g := 1,000$:

$$N_g := n_{g_L}^{\Sigma}(N_L),$$

with

$$n_{g_L}^{\Sigma}(l) := \sum_{\lambda=1}^{l} n_{g_L}(\lambda),$$

and

$$N_L := \arg\min_{l=1}^{\left\lfloor \frac{d_B}{w_L} \right\rfloor} \left| \tilde{N}_g - n_{g_L}^{\Sigma}(l) \right|.$$

The mean distance in-between cells within the same layer resulting from this model was used as distance in-between germ cell layers:

$$w_L := \phi_g + \overline{d_g},$$

with

$$\overline{d_g} := \sum_{\lambda=1}^{\left\lfloor \frac{d_B}{w_L} \right\rfloor} \frac{d_{g_L}(\lambda)}{\left\lfloor \frac{d_B}{w_L} \right\rfloor},$$

where

$$d_{g_L}(l) := \frac{\phi_L(l)}{1 + \dfrac{1}{\sin\left(\dfrac{\pi}{N_g}\right)}} - \phi_g.$$

Germ layers were annotated functionally based on literature (Brenner and Schedl, 2016; Fox et al., 2011). The proximal gonad arm was filled with 8 oocytes, maximizing the diameter of each oocyte under the constraint of the corresponding gonad diameter. The proximal end of the distal germ cell layers and the distal end of the distal-most oocyte defined the boundaries of the loop region. Assuming steady-state with an apoptotic rate of 90% (Brenner and Schedl, 2016), the loop region was filled with 100 germ cells in layers (uniformly spread across the loop region).

In order to determine the germ cell migration time, migration speed data was used from Wolke and colleagues (Wolke et al., 2007). Those migration speeds were assigned to germ cell layers and oocytes as follows:

i) The switch to pachytene happens between the transition zone and the meiotic zone.

ii) The switch to 'proximal of the loop' happens at the bend (398.3 μm from the DTC).

iii) In between, the pachytene was split into thirds to assign 'early', 'mid', and 'late' pachytene.

iv) The speed of particles entering oocytes was assigned to the very last germ cell layer in the loop region.

v) Oocytes were assigned the speed of particles right after entering oocytes.

For each germ cell layer and oocyte, 10 speed values were sampled from a zero-truncated normal distribution with the corresponding mean and standard deviation reported by Wolke and colleagues (Wolke et al., 2007). Those speed samples were smoothened by calculating the running mean of 100 adjacent values (10 adjacent germ layers / oocytes). The smoothened speed samples were used to fit a natural spline function mapping the distance to the DTC $d[\mu m]$ to the migration speed $v(d)[\mu m/min]$. To translate migration speeds to migration times, for each distance (in 1 μm steps) along the gonad arm, the migration speed in $\mu m/min$ was calculated, inverted and summed up cumulatively to get the total migration time from the DTC to the given distance. These discrete measurements were used to fit the final spline model mapping the distance to the DTC $d[\mu m]$ to the migration time from the DTC $t(d)[h]$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data Pre-processing

Raw sequencing basecalls were demultiplexed and converted to FASTQ format using bcl2fastq v2.18.0.12 pooling reads across lanes (`--no-lane-splitting`). No adapter trimming was performed at this stage by not specifying adapter sequences in the sample sheet CSV file. To avoid masking of the short read 1 (barcode and UMI), the `--mask-short-adapter-reads=10` option was used. 3' reads (read 2) were annotated with their corresponding (corrected) barcode and UMI sequences (read 1) using custom scripts. Reads with identical barcode, UMI and sequence were collapsed and the unique reads were assigned to per-slice FASTQ files by barcode. Small RNA reads were subject to two rounds of 3' end trimming by flexbar v. 2.5 (Dodt et al., 2012): The first round to remove 3' adapters, the second to remove the poly(A)-tail added during the library preparation (using 10 A's as 'adapter sequence'). 3' nucleotides with low basecall quality scores were trimmed using flexbars `--pre-trim-phred=30` option and the 3 nucleotides 5' overhang introduced by the template-switching polymerase were trimmed using a custom awk script also discarding reads with a remaining length <18 nts.

Mapping of Reads to the *C. elegans* Genome

RNA-seq reads were mapped to the ce11/WBcel235 genome assembly using STAR_2.5.1b (Dobin et al., 2013) and an index with splice junction information from the Ensembl 82 transcriptome annotation. Alignments were sorted using sambamba v0.4.7 (Tarasov et al., 2015). Coverage tracks were generated using bedtools v2.23.0 (Quinlan and Hall, 2010) via the genomecov command specifying the `-split` and `-bg` options for splice-aware BedGraph output and splitting by strand using the `-strand` parameter. The total number of mapped reads per sample/splice was determined using the flagstat command of samtools 0.1.19-96b5f2294a (Li et al., 2009) and converted to the corresponding `-scale` parameter for `bedtools genomecov` for reads-per-million-mapped (RPM) normalization. Coverage BedGraph files were converted to BigWig format using bedGraphToBigWig v 4.

3' Extension of Transcript Annotation

The identification of downstream coverage peaks for 3' extension of the WS260 transcriptome was performed using a custom R script: For each protein coding gene, the intergenic distance to the next downstream protein coding, ncRNA, lincRNA, pseudogene, rRNA or snoRNA gene (on the same strand) was calculated. Intergenic regions longer than 10 kb were truncated and the RPM-scaled genome coverage per sample of those regions was extracted from the BigWig files generated before. The per-sample coverage vectors were averaged per genomic position and binarized into uncovered regions (< 5 RPM mean coverage) and covered regions ($\geq$ 5 RPM mean coverage). Covered regions with a length of $\geq$ 50 nucleotides were considered as coverage peaks. Per downstream intergenic region, the downstream-most coverage peak was selected for the 3' extension of the corresponding upstream gene. Only downstream extensions with a length up to 3 kb were considered for downstream analyses. For each gene with a downstream extension, all annotated transcript isoforms extending to the 3' most genomic position of the corresponding gene were kept and got their 3' UTRs extended by until the 3' position of the respective downstream peak. Those 3' extended transcripts were exported to a GTF file and merged with the WS260 transcriptome annotation using custom awk scripts.

Transcriptome Pre-processing

To enable the assignment of 3' end RNA-seq reads to transcript isoforms, the 3' extended WS260 transcriptome annotation was preprocessed using a series of custom R scripts: 3' A's were trimmed from all annotated transcripts as they would be indistinguishable from poly(A)-tails. The resulting transcripts were truncated to the 3' most 500 nucleotides. Transcript isoforms with the same genomic coordinates and internal structure were collapsed and enumerated by decreasing corresponding (max.) 3' UTR length.

Isoform-Specific Transcript Abundance Estimation

RNA-seq reads were assigned to transcripts using kallisto 0.43.1 (Bray et al., 2016): For 3' reads, an index of the collapsed transcriptome annotation described above was used. For full-length coverage reads (poly(A)+ and ribodepleted total RNA-seq libraries), an index of the full 3' extended transcriptome annotation was used. For all libraries, the `--bias` was passed to `kallisto quant`. For single end reads, additionally the `–single`, `–fragment-length=1` and `–sd=1` options were used. All libraries were sequenced with a first-strand-reverse stranded protocol. Thus, poly(A)+ and ribodepleted total RNA-seq samples were analyzed in `--rf-stranded` mode. The 3' reads, while presented to kallisto as single-end reads, originally were sequenced as read 2, therefore resembling first-strand-forward single-end data. Thus, for these libraries the `–fr-stranded` mode of `kallisto quant` was used. Per-isoform read counts were exported to TSV files using the `--plaintext` option.

Data Processing

The raw read counts per transcript isoform and slice/sample were further processed using a custom R script: Though the whole annotated transcriptome was quantified to check for specificity of the experimental and computational approach, downstream analyses were limited to protein coding transcripts only. For gene-level analyses, isoform-level read counts were summed across all isoforms of a given gene. To compensate for differences in sequencing depth, raw read counts were normalized to counts-per-million (CPM). For full-length coverage protocols (poly(A)+ and ribodepleted total RNA-seq) an additional correction for the transcript length was performed, resulting in transcripts-per-million (TPM) estimates. Slice-data were arranged from distal-to-proximal by the known order of their barcodes and assigned to a relative position scale representing each slice by its center and accounting for differences in the number of slices per sample.

Aligning Cryo-Cuts of Different Samples to a Single Coordinate System

As the start- and endpoint of gonad slicing was not precisely the same for all replicates, all slices of different replicates were aligned to a common coordinate system. This was achieved by comparing per-sample LOESS fits of abundance estimates across slices with *in situ* images of certain genes in the germline. Therefore, the gene profile of one replicate was fixed according to the corresponding *in situ* image and other replicates were aligned to the fixed replicate. This was done for approx. 30 gene profiles and the median of the shifting for those 30 profiles was calculated and used for all gene profiles.

Integration of Replicate Data

The aligned discrete per-gene/isoform spatial expression profiles of individual replicates were used to fit a continuous consensus profile using local regression (LOESS) with a span of 0.4 through a custom R script. Slices with less than 10,000 reads assigned to the transcriptome ('dropout-slices') were excluded from the fitting procedure. For visualization, 50 equidistant points along the distal-to-proximal axis were inferred from those fits. For downstream analyses, only 20 points were used to reflect the actual resolution of the data more conservatively. All data (incl. dropout-slices) are available through the interactive data exploration interface published alongside this study.

Data Analysis of the Small RNA Transcriptome

The trimmed libraries were first mapped with bowtie2 (version 2.3.3.1) (Langmead and Salzberg, 2012) using the parameters `--very-fast-local --phred33 --local` to the *E. coli* genome (NC_000913.3, K-12, MG1655) in order to remove *E. coli* RNA contamination. The cleaned-up libraries were then mapped with STAR (version 2.5.3a) (Dobin et al., 2013) to the WBcel235/ce11 genome assembly using the Ensembl 87 annotation and the parameters

`–alignIntronMax 140000 –alignSJDBoverhangMin 17`
`–alignSplicedMateMapLmin 30 --outFilterMultimapNmax 5`
`–outFilterMismatchNmax 2 --outFilterMatchNmin 17`
`–outFilterScoreMinOverLread 0 --outFilterMatchNminOverLread 0.`

Sense and antisense read counting on features was done using HTSeq (version 0.9.1) (Anders et al., 2015) with the parameters

`-a 0 -m intersection-nonempty --nonunique=all`
`–secondary-alignments=score` combined with `-s yes` for sense and with `-s reverse` for antisense counts. Known and previously unannotated miRNAs were identified separately using the cleaned-up libraries and the miRDeep2 algorithm (version 2.0.0.7) (Friedländer et al., 2012) with the miRBase21 reference. First, miRDeep2 was ran on the pooled libraries. Then, the unannotated miRNA predictions found were added to the miRBase21 reference. Consequently, the combined reference of known and unannotated miRNAs was used for a second run of miRDeep2 on each library separately and on the pooled library as well. This way we unified the expression estimates of known and unannotated miRNAs under a common measure of counts per million of mapped reads (CPM).

The miRNA-target correlation analysis used robust linear regression based on the MM-estimator in order to reduce the effect of outliers (Koller and Stahel, 2011). All miRNAs were divided into families based on their 2-7 nt 6mer seeds (reverse-complemented). Putative target genes were identified by counting miRNA 7mer seeds on all of their unique and longest 3' UTR isoforms. The 3' UTR isoform with the maximum number of 7mer seeds was taken as representative for that miRNA-target gene interaction. The miRNA 7mer seeds were chosen to be either the reverse-complement of the miRNA 2-8 nts or the reverse-complement of the miRNA 2-7 nts immediately followed by an A (Bartel, 2009). The control list of targets was generated by mutating the 3rd and 4th nucleotides of

these 7mer seeds. Robust linear regression was done by summing the LOESS smoothed CPMs among the miRNA family members on each LOESS point and using this summarized family-wise expression with the corresponding target smoothed expression. In order for a correlation to be considered we demanded that both the family-wise miRNA expression and the target expression were commonly non-zero in at least 25% of the LOESS points.

DATA AND SOFTWARE AVAILABILITY

The accession number for the raw and processed data reported in this paper is GEO: GSE115884.

ADDITIONAL RESOURCES

Our online tool SPACEGERM can be accessed via the following URL:
https://shiny.mdc-berlin.de/spacegerm/.