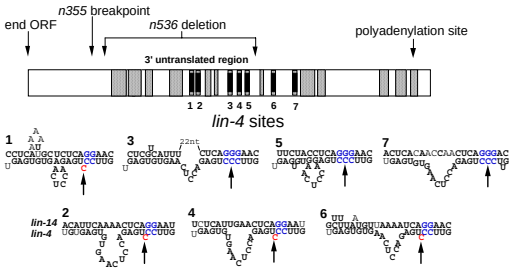
miRNA and RNAi Research in the Ruvkun Lab

John Kim, PhD
Sylvia Fischer, PhD
Devin Parry, PhD
Xiaoyun Wu, PhD
Meng Wang, PhD
Chi Zhang, PhD
Julia Huang, Ph.D.
Gabe Hayes
Maurice Butler
Sascha Russel
Harrison Gabel
Jinling Xu

miRNAs (miRNAs) are small, non-coding RNAs that are thought to regulate expression of target genes either, as in *Caenorhabditis elegans*, by blocking translation through annealing to the 3'UTR of the mRNA or, as in some cases in *Arabidopsis thaliana*, by directing degradation of the target mRNA. The first miRNA identified was the *C. elegans lin-4* miRNA which acts in the pathway of temporal patterning. We identified the second miRNA, *C. elegans let-7* that also acts in the temporal patterning pathway, and showed that it and the complementarity in its target gene are conserved in a wide range of species including *Drosophila* and human. Several hundred miRNAs have since been identified in mammals, fly, worm, and plants. So far, about 20% of the miRNAs are conserved among nematodes, flies and mammals, suggesting ancient roles in gene regulation.

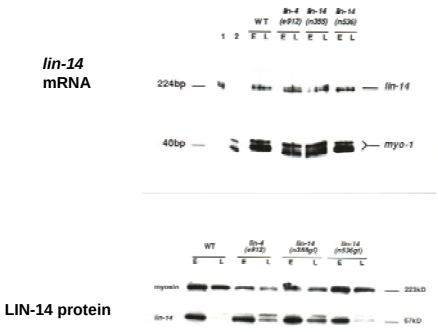
The *lin-4* RNA is complementary to 7 sites in the *lin-14* 3' UTR



lin-4 and *let-7* are the miRNAs that we know the most about because they have been intensively studied and their targets (*lin-14, lin-28, lin-41, hbl-1*) and other genes in the pathway have been genetically identified. The expression of the *lin-4* RNA late in the first larval stage downregulates the translation of *lin-14* mRNA via complementary 3' UTR elements to control the timing of developmental events; the expression of the *let-7* RNA at late larval stages similarly downregulates *lin-41* expression via 3' UTR complementary elements. These temporally regulated events include timing of hypodermal seam cell division and differentiation, timing of vulva development, and timing of remodeling of the synaptic connectivity of the DD neurons. Loss of function mutations in the *lin-14* gene suppress *lin-4* mRNA null mutations whereas *lin-14* gain of function mutations that delete *lin-4* complementary sites phenocopy the *lin-4* miRNA mutation, showing that *lin-14* is the major *lin-4* regulated target mRNA. *lin-14* encodes a nuclear localized protein that is expressed in cells of the hypodermis, intestine, body wall muscle, nerve ring and ventral nerve cord of embryos and early first larval stage worms. Transition to the second larval stage fates coincides with dramatic reduction in the level of LIN-14 protein in most tissues. The *lin-4* RNA binds to the *lin-14* 3' UTR in vitro and this binding is dependent on the *lin-4* complementary sites in the *lin-14* 3' UTR. Also, the *lin-14* 3' UTR is sufficient to confer *lin-4* dependent down-regulation on a reporter gene.

Rules that may be general for the mechanisms of miRNA regulation of target genes have emerged from biochemical studies of *lin-4* regulation of *lin-14* mRNA translation. The abundance of the *lin-14* target message does not significantly decrease upon expression of the *lin-4* miRNA, making it unlikely that transcription or mRNA stability are altered. Additionally, *lin-14* mRNA remains associated with polyribosomes even after LIN-14 protein production halts in response to *lin-4* RNA expression. We have found that the localization of miRNAs to polysomes is general to a wide range of

Post-transcriptional down-regulation of *lin-14* by *lin-4*



Bruce Wightman

mammalian miRNAs, suggesting that the *lin-4* findings are general to miRNAs.

In addition to controlling the timing of stage specific cell divisions, *lin-14* also regulates changes in neural function. For example, a set of GABAnergic motor neurons undergoes synaptic remodeling at the end of the L1 stage. In *lin-14* loss of function mutants, these neurons remodel prematurely. The precocious rewiring of these neurons is rescued by expression of *lin-14* in the affected GABAnergic neurons, indicating that LIN-14 protein can act cell-autonomously to direct the timing of rewiring. This finding is in accord with our demonstration of many neural miRNAs and expression of many miRNAs in neurons.

The generation and utilization of 22nt RNAs from longer dsRNA was first established in plant post-transcriptional gene silencing (PTGS) which is equivalent to RNAi. miRNAs are produced via a mechanism similar to the short interfering RNAs (siRNAs) that act in the RNAi pathway. siRNAs and miRNAs are found together in multiprotein complexes in *Drosophila* and human and also share factors required for function in *Arabidopsis*. siRNAs or mature miRNAs are produced from a longer double stranded RNA structure through cleavage by the RNase III-type enzyme Dicer; the double stranded precursor of miRNAs is usually a 80-100 nt imperfect hairpin RNA, while siRNAs can be generated from long dsRNAs or hairpin RNAs. In the second step, the now single stranded, 21-25 nt RNAs are incorporated into an effector complex named RISC or miRNP, and this complex, by base pairing with the target mRNA, blocks the expression of that target. A *C. elegans* member of the PIWI/PAZ family, RDE-1 (RNAi Defective), is essential for RNAi. The paralogs ALG-1 and ALG-2 regulate the maturation and function of the *lin-4* and *let-7* tiny RNAs. Inactivation of *alg-1* and *alg-2* cause heterochronic defects consistent with the dysregulation of *lin-4* and *let-7* maturation. The function of the PIWI and PAZ domains are unknown but the amino acid sequences are highly conserved from fungi to humans. A reasonable model is that the RDE-1 protein functions with the DCR-1 RNase to facilitate processing of siRNAs from dsRNA, whereas the ALG-1/ALG-2 proteins function with DCR-1 to mediate the maturation and function of miRNAs. Curiously, there are more than 25 PIWI/PAZ protein genes in *C. elegans* compared to about one quarter that number in other sequenced genomes. The hundreds of new miRNAs identified from plants, *C. elegans*, *Drosophila*, and human are predicted to be processed from imperfect stem-loop precursors analogous to those of *lin-4* and *let-7*, and therefore are likely to be processed by Dicer and some or all of the PIWI PAZ domain proteins.

After recognition of target mRNAs, it is likely that it is only their perfect vs. imperfect complementarity in animals that distinguishes mRNA degradation pathways from translational control pathways. Animal miRNAs usually show substantial non-complementarity in the middle of the miRNA:mRNA duplex and do not direct degradation of the mRNA but rather block its translation after translation initiation. By changing the degree of complementarity, miRNAs can function as siRNAs, and *vice versa*. Unlike animal miRNAs, plant miRNAs show almost perfect complementarity to known mRNAs. Thus target prediction for plants is far easier than in animals. siRNAs or some *Arabidopsis* miRNAs direct cleavage in the middle of the siRNA:mRNA duplex of target mRNAs that are highly complementary. An impressive list of plant regulatory genes have now been found to be targets of these miRNAs. There seems to be an enrichment for meristem patterning elements. It is unclear whether this is due to some feature involving oriented cell divisions in the plant or perhaps a connection between totipotency vs commitment at the meristem. It may be significant that miRNAs have been implicated by the isolation of the piwi mutations in stem cell vs. commitment to germ line fates in Drosophila and mice.

The identification of animal target mRNA genes by pure informatics is complicated by imperfect complementarity between miRNAs and the 3'UTR sequences of their targets. In the case with *lin-4* and *let-7*, their targets were first identified by genetic analyses. A major problem for the field is to improve rules of target mRNA prediction. Only few targets of miRNA control are known in animals other than *C. elegans*: *Drosophila hid*, a gene that induces apoptosis. The lethal phenotype of loss-of-function mutants of *C. elegans* Dicer (*dcr-1*) suggests that some miRNA activity or endogenous siRNA activity is required for viability. miRNAs could act in translational control pathways commonly utilized for regulation of maternal mRNAs in early development or of mRNAs tethered in dendritic regions of neurons proposed to mediate synaptic plasticity. Translational control may be just one type of gene regulation mediated by miRNAs. For example, base pairing of miRNAs to 5'UTR sites could impede translational initiation, to pre-mRNA splice sites could direct alternative splicing or to polyadenylation signals could alter mRNA stability. Even DNA sequences could be targets of miRNA regulation; duplex formation between miRNAs and DNA could alter transcription or replication either directly or by guiding modification of specific DNA sequences, as has been seen in *S. pombe*.

In *C. elegans*, miRNA complementary sites have so far only been found in the 3'UTRs of target genes. Repression of these mRNAs may require RNA binding proteins that recognize these RNA duplexes to either disrupt initiation of translation, control translation at a post-initiation step or decrease the stability of the nascent polypeptide. The Fragile X Mental Retardation Protein (FMRP) is part of the RISC complex together with miRNAs and siRNAs.

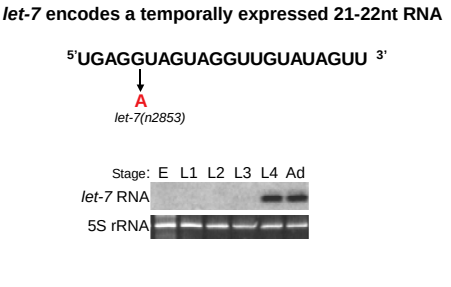
Thus, miRNAs might act in concert with RNA binding proteins, or act similar to RNA binding proteins, to suppress expression of their target mRNAs.

Another curious feature of RNAi, as well as PTGS, is the ability of the initiating signal to spread to distant regions of the organism. Presently the exact nature of the mobile element is unknown but it likely includes processed forms of the input RNA. The SID-1 protein with eleven transmembrane domains is required in *C. elegans* for spread of the RNAi signal. The fact that worms harbor transmembrane proteins that may shuttle RNA molecules between cells raises the question of the endogenous role of such proteins. Although deletion of this protein receptor gene results in no obvious developmental abnormalities, it is possible that related proteins could mediate transport of endogenous RNAs such as *lin-4* and *let-7*. The cellular origin as well as the final tissue distributions of these RNAs are yet to be identified.

Why regulate gene expression with miRNAs? The expression of a 70nt RNA and processing to the 22nt form is likely to be much faster than the transcription of a >1000nt typical messenger RNA that undergoes extensive processing (i.e. splicing, polyadenylation) to generate the mature protein-coding mRNA. miRNAs may mediate particularly rapid regulatory events, or example during the short cell cycles of developing embryos to in turn regulate the translation of pre-existing maternal mRNAs. Given that the 22nt siRNAs can act systemically with probable amplification in at least some organisms, it is possible that miRNAs may also undergo amplification and intercellular transport to control gene expression. Identification of the targets of the new miRNAs as well as elucidation of their mechanisms of action will ultimately help us understand the role of 22nt RNAs in gene regulation.

1. The *let-7* miRNA regulates *Caenorhabditis elegans* developmental timing

let-7 was identified in a genetic screen for mutations that suppress the precocious defect of the early acting heterochronic gene *lin-14*. The first evidence of a *let-7* heterochronic defect is at the L4-to-adult molt. *let-7* acts upstream of the heterochronic gene *lin-29*, a zinc finger transcription factor that specifies adult-specific patterns of cell lineage and cell differentiation. The retarded alae phenotype caused by *let-7* mutations was partially suppressed by precocious mutations in the genes *lin-41, lin-42, lin-14*, and *lin-28*.



By a combination of transgene complementation, RNA expression analysis, and mutant allele sequencing, we established that *let-7* encodes a miRNA. *let-7* expression is temporally regulated: expression was first detected at the early L3 stage, and high level expression was detected at the early L4 and adult stages. This expression profile is consistent with the *let-7* mutant phenotype, which affects development specifically in late larval and adult stages.

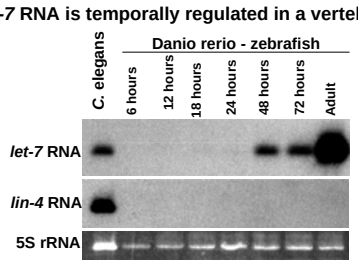
Five heterochronic genes contain sequences complementary to *let-7* in their 3' UTRs. The sites that are complementary to the *let-7* RNA in the *lin-41* 3' UTR mediate *let-7* dependent temporal down-regulation of a *lac-Z* reporter gene. A reporter bearing the *lin-41* 3' UTR was expressed in larval stage animals but not in adult wild-type animals, similar to the temporal regulation of the *lin-41* gene itself whereas a control reporter bearing the *unc-54* 3' UTR was expressed at all stages. The segments complementary to *let-7* are located adjacent to or overlapping segments complementary to the *lin-4* regulatory RNA in the *lin-14*, *lin-28*, and *lin-41* 3' UTRs, suggesting that the two RNAs could both regulate the expression of these target genes.

2. The *lin-41* RBCC gene acts between the *let-7* miRNA and the *lin-29* transcription factor

We performed F1 and F2 screens for mutations that suppress the lethality of *let-7(n2853)* and identified seven alleles of *lin-41*. The *lin-41* mutations in a *let-7*(+) background cause recessive precocious heterochronic expression of an adult-specific hypodermal cell fate. By genetic mapping, and candidate gene sequencing, we identified mutations associated with *lin-41* alleles in C12C8.3. Some are null alleles. LIN-41 is a member of a large family of RING finger proteins. RING finger proteins are subunits of the E3 ubiquitin ligase. Two sequences in the *lin-41* 3' UTR are complementary to the *let-7* miRNA. *lin-41* functions upstream of the *lin-29* transcription factor. During wild-type development, LIN-29 protein is first detected during the L4 stage. However, the *lin-29* mRNA accumulates much earlier, starting at the L2 stage. An attractive model is that the LIN-41 RING finger protein actively regulates the half life of the LIN-29 protein. Alternatively LIN-41 may bind to the *lin-29* mRNA and inhibit its translation during early larval stages. The *let-7/ lin-41* regulatory circuit is evolutionarily conserved. Both the *let-7* RNA and sequences complementary to it in the 3'UTRs of *lin-41* orthologs are maintained from worms to vertebrates.

Additionally, the temporal expression pattern of *let-7* RNA is conserved, indicating that this circuit may control developmental timing broadly across animal phylogeny.

3. Conservation across animal phylogeny of the sequence and temporal expression of the *let-7* miRNA

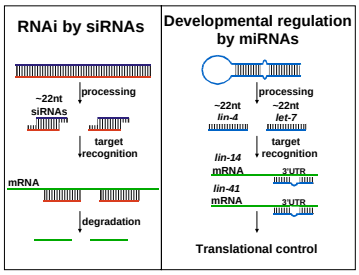


Amy Pasquinelli

We detected a *let-7* miRNA of ~21 nts. in samples from a wide range of animal species, including vertebrate, ascidian, hemichordate, mollusc, annelid and arthropod, but not in RNAs from multiple cnidarian and poriferan species, *S. cerevisiae*, *E. coli*, or *Arabidopsis*. *let-7* temporal regulation is also conserved: *let-7* RNA expression is first detected at late larval stages in *C. elegans* and *Drosophila*, at 48 hours post-fertilization in zebrafish, and in adult stages of annelids and molluscs. The *let-7* regulatory RNA may control late temporal transitions during development across animal phylogeny. Similar stem-loop secondary structures are predicted for precursor transcripts of *Caenorhabditae*, *Drosophila* and human *let-7* RNAs. The expression levels of the human *let-7* RNA varied among tissues, indicating possible cell type regulation of *let-7* expression. Considering that *let-7* is expressed late in animal development, it may be significant that the lowest level of human *let-7* is observed in bone marrow, which consists of a large proportion of immature cells. The conservation of sequence, of a longer structured precursor, of the 21 nucleotide length, of temporal regulation, and of complementary sites in the *lin-41* target in two ecdysozoan and one chordate species are strong evidence of a conserved function.

4. Genes and mechanisms related to RNA interference regulate expression of the miRNAs that control *C. elegans* developmental timing

let-7 and *lin-4* miRNAs are processed from larger transcripts of approximately 70 nucleotides that have the potential to fold into stable stem-loop structures. The Dicer protein mediates the processing of longer dsRNAs into the siRNAs which subsequently guide mRNA destruction.



We found that *dcr-1*, the *C. elegans* ortholog of *Drosophila* dsRNA processing ribonuclease Dicer is required for RNAi. The processing of the *lin-4* and *let-7* miRNAs from longer precursors also requires DCR-1. RDE-1 is a PIWI PAZ domain protein that is essential for RNAi but does not affect *let-7* or *lin-4* miRNA processing, nor does an rde-1 null mutation cause any obvious developmental defects. However, we found that the activities of two *C. elegans* homologs of *rde-1*, *alg-1* and *alg-2*, are essential for the processing and activity of the *lin-4* and *let-7* miRNAs. Inhibition of *dcr-1* or *alg-1* and *alg-2* causes heterochronic phenotypes that are consistent with their effects on *lin-4* and *let-7* RNA processing. These findings show that natural dsRNAs are processed into small regulatory RNAs via a mechanism analogous to that involved in processing the double-stranded RNAs that trigger RNAi.

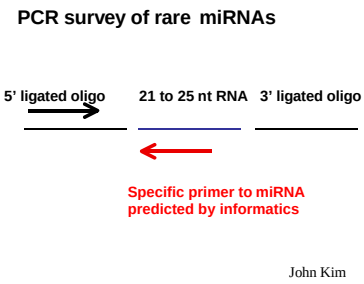
The RNAi and miRNA pathways trigger distinct outcomes: RNA destruction versus translation inhibition. The distinction between mRNA destruction by RNAi and inhibition of translation by the *lin-4* regulatory RNA is most likely due to a difference between perfect and imperfect complementarity, because plant miRNAs are nearly perfectly complementary to target mRNAs and cause their degradation, and engineered perfectly complementary sites in animal mRNAs causes their degradation rather than translational repression.

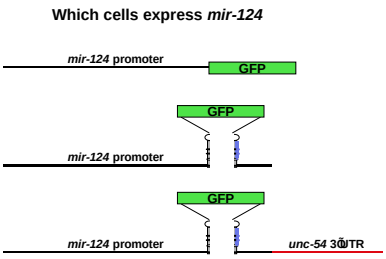
In addition to a function in RNAi, developmental functions have also been reported for members of the *piwi* and *ago1* families in both animals and plants. One feature that emerges from studies of these developmental phenotypes is that many of these genes regulate germ-cell and stem-cell functions.

C5. Computational and experimental identification of *C .elegans* miRNAs

Despite the bevy of miRNAs that has emerged from biochemical cloning, such screens are biased to abundant miRNAs. We developed computational methods that predict miRNAs encoded by the *C. elegans* genome independently of abundance, searching for properties of known miRNAs, including their length (typically 21-24 nt), precursor hairpin structure (typically ~70-90 nt, with multiple 1-4 nt bulges and mismatches). We improved prediction methods by focusing on miRNAs conserved across species or within a species.

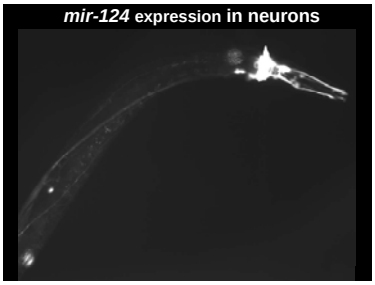
We surveyed the intergenic, non-coding regions of the *C. elegans* genome and identified sequences that would form a hairpin of ~70 nucleotides. We derived three sets of predictions from an initial set of 8,713 candidate *C. elegans* miRNA hairpins. Two algorithms were based on interspecies conservation of the candidate miRNAs in worm, fly, and human, while the third was based on homology to known miRNAs. We identified the majority of the known *C. elegans* miRNAs that were conserved in fly and human as well as 74 other candidate miRNAs that were conserved with the characteristic secondary structure of known miRNAs. Most of the candidate miRNAs were tested by Northern blotting and 3 showed strong 22 nt expression. We generated PCR-based libraries of miRNAs and surveyed those libraries by PCR (see Aim 1a). 20% of the tested 50 candidate miRNAs were positive by a PCR assay. In addition, all of the known miRNAs assayed were detected by the PCR assay. The miRNA candidates not detected by the PCR strategy may either be expressed at even lower levels or may be expressed under specific environmental conditions such that they are not represented in the original library constructed from mixed-stage worm RNA, or may be false positives. We estimated that the *C. elegans* genome encodes hundreds of miRNAs.

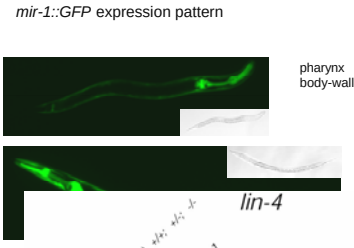




6. Promoter fusions of miRNAs and GFP

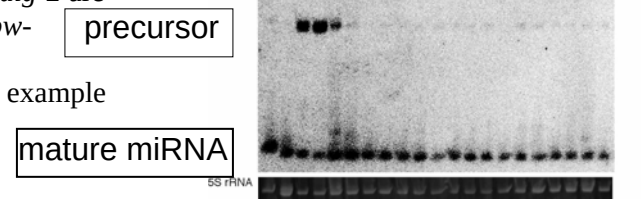
We have so far fused the upstream regions of the miRNA genes *lin-4*, *mir*-1, *mir*-124 to GFP and see very promising expression patterns. For example the *lin-4*::GFP fusion is strongly up regulated at the L1 stage when we see up regulation by Northern blotting. We can see *lin-4*::GFP in the seam cells that are affected by *lin-4* mutations but also in the developing ventral nerve cord. Neurons in the ventral nerve cord are known to rewire during the L1 stage in a *lin-14* dependent manner and it is *lin-14* regulation within those neurons that is key to the timing regulation of that rewiring. Similar studies of proper regulation a *let-7* GFP fusion have been published. The timing and pattern of these GFP expression genes suggests that it recapitulates the normal expression pattern of the corresponding miRNA. Most persuasive is our finding that the *mir*-1::GFP fusion is expressed only in muscle. Mammalian *mir*-1 is also muscle specific. We need to validate these studies using sensor genes engineered to have perfect complementarity to miRNAs as proposed in Aim 1b, but we are very optimistic about the possibility of learning the site of gene expression from these fusions. This is key in *C. elegans* where in situ hybridization has been disappointing for much less demanding mRNA *in situ*.





7. The role of other PIWI-PAZ family proteins in miRNA processing

More than 25 genes encoding PIWI-PAZ domain (Argonaute family) proteins appear in an unusual proliferation in *C. elegans*. Only the 2 worm piwi (*prg-1* and *prg-2*) and 2 worm argonaute subtypes (alg-1 and alg-2) are orthologous to genes in other animal and plant species. *alg-1* and *alg-2* are required for miRNA processing, while others, such as *rde-1* and *ppw-1* are required for RNAi. We hypothesized that other Argonaute family proteins might also function in miRNA processing, with for example particular Argonaute paralogs responsible for a subset of miRNAs.



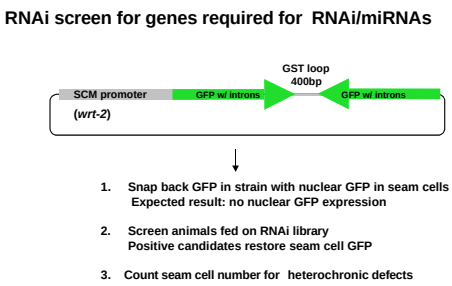
We examined the RNAi clones against the 19 PIWI-PAZ proteins available in the Ahringer RNAi library for effects on miRNA processing. *dcr-1* and *alg-1* mutant phenotypes were recapitulated by RNAi, but most clones produced no phenotype or a weak growth phenotype. We isolated RNA from animals grown on each of the 19 RNAi strains, as well as vector. We selected miRNAs from different sequence "families" and probed blots for *lin-4*, *let-7*, and *mir-1*, *-2*, *-66* and *–83*. All miRNAs showed defects in precursor processing when subjected to *dcr-1*, *alg-1* and *alg-2* RNAi, but not to RNAi of other Piwi paz genes. This result is

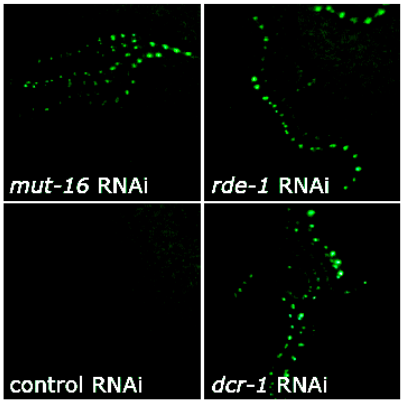
inconsistent with the hypothesis that other Argonaute homologs have roles analogous to that of ALG-1 and –2, though the possibility remains that they might play such a role in restricted tissues, for example rare stem cells.

8. Identification of factors required for down regulation of reporter gene expression through the *lin-14* 3'UTR. We used a *C. elegans* strain carrying a *rol-6(su1006)* gene under control of the *rol-6* promoter and the *lin-14* 3'UTR, containing seven *lin-4* complementary site. *rol-6(su1006)* is a dominant mutation, causing a Rol phenotype if the gene is under control of the *rol-6* 3'UTR. However, when controlled by the *lin-14* 3'UTR, the fusion gene does not cause a Rol phenotype in a wild type background. When DCR-1 is down regulated by RNAi, a Rol phenotype is observed, suggesting that suppression of expression of the transgene is dependent on a miRNA (or siRNA), presumably *lin-4*. We made a sub-library out of the whole genome RNAi library that targets 450 *C. elegans* genes predicted to encode RNA binding proteins, helicases, nucleases, genes known to be involved in RNA silencing in *C. elegans* and other systems as well as heterochronic genes. This cherry picked sub-library was screened for suppression of *lin-4* mediated repression of the fusion gene and identified nine RNAi clones that reproducibly suppress the non-Rol phenotype. For example, two of 41 KH domains are detected (T21G3.5 and R11A8.7), two helicases, T07D4.3 and the drh-1 helicase that interacts with the RNAi PIWI PAZ factor RDE-1, the pumilio protein W06B11.12, and the RNP1 proteins K08D10.4 and D1007.7 out of 176 RNP1 proteins in *C. elegans* are detected. None of these genes was previously shown to function in miRNA activity. To verify that the expression of the *rol-6*/*lin-14* 3' UTR fusion gene observed with the RNAi-suppressors is not a result of a general, *lin-4* independent, relief of transgene silencing, we have studied the expression of another transgene, *let-858::gfp::unc-54*, after RNAi of the suppressors. Two out of the nine RNAi suppressors also induce de-silencing of this transgene but 7 are specific to the *lin-14* 3' UTR. From this list pumilio paralogs are very interesting because of their genetic assignment to the sorting of mRNAs in Drosophila posterior patterning [100]. T21G3.5, tag-44, is the closest homolog of *gld-1*. GLD-1 is known to repress several *C. elegans* mRNAs, *e.g. tra-2*, by binding to their 3'UTRs.

9. A similar screen for factors necessary for RNAi as well as miRNA function.

This same cherry picked library was used to screen for factors that when RNAi inactivated cause defects the silencing of a GFP reporter gene by an engineered snapback dsRNA GFP. The starting strain for this screen is non green due to silencing of a GFP reporter gene by dsRNA from GFP. Feeding this strain dsRNA for dcr-1 potently activates the expression of GFP. By screening through the 450 cherry picked hits, we identified 11 more genes that affect RNAi. These include 3 hits that are expected: K12H4.8= *dcr-1*, K08H10.7=*rde-1*, T20G5.11= *rde-4*, and more hits that make sense: T22B3.2a=another PIWI PAZ, Y38A10A.6=DEAD/DEAH box helicase, R05D11.4=DEAD/DEAH box helicase, D1007.7=RNA-binding region RNP-1, D1046.1=RNA-binding region RNP-1, F26B1.2=human HNRPK, and others that are unannotated but that we are following up: T19B10.4A=unknown, F52G2.2=unknown. These genes need to be further validated, most especially because RNAi inactivation of RNAi components is unlikely to yield null phenotypes and is subject to some more trivial explanations as well---most worrisome is that RNAi of two genes tends to not work as well due to competition in this non linear process. We propose further tests of the validity of these results. But because dcr and rde-1 and rde-4 emerge from the analysis so cleanly, we think that many of the hits will survive these further filters. Some of the genes identified from the *lin-14* UTR search are also identified by the RNAi deficient search.



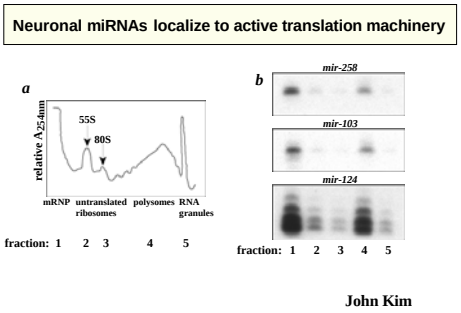


10. Identification of mammalian neural miRNAs that co-purify with polyribosomes

Regulatory pathways rich in evidence for translational control are most likely to use miRNAs. Particular mRNAs, as well as translational machinery including ribosomes and other non-coding RNAs, are localized to dendritic regions of neurons. Some of these mRNAs encode proteins such as kinases and translational control factors that are attractive candidates to mediate synaptic changes]. Elements in the 3' UTR of some mRNAs have been implicated in their localization. Synaptic activity activates the translation of these localized mRNAs.

miRNAs would fit well into this model, if their maturation, localization, or expression were modulated by synaptic activity.

We sought miRNAs from mammalian brain preparations and explored their regulation and possible function in translational control. We isolated miRNAs from rat cerebro-cortical dissociated cultures of E18 rat embryos. RNAs of 20 to 25 nt were purified, cloned, and sequenced. Eighty-four distinct miRNAs were identified; 48 were distinct from previously reported mammalian miRNAs and the remaining 38 matched those previously identified. The number of times a miRNA was cloned varied greatly, ranging from 71 clones for *mir-125* to a single clone for 32 of the 86 miRNAs; in all, 29 of the 86 miRNAs were isolated 4 or more times. Predicted precursors for 39 of the 48 newly identified rat miRNAs were found in rat genome traces and the assembled mouse genome sequence, and in nearly all of those cases, sequence conservation included the predicted precursor.



All 48 new miRNAs were examined by Northern blotting to total RNA from adult rat brain, and 23 of the miRNAs were detected. In general, the number of times a particular miRNA was cloned correlated with the intensity of the signal by Northern blotting. The remaining 25 may either be expressed at levels below the detection threshold for Northern blotting or may exhibit temporally restricted expression patterns. Of the 12 miRNAs examined, eight (*mir-103*, *-124a*, *-128*, *-224*, *-229*, *-233*, *-250*, *-258*) were expressed largely in the cortex and cerebellum, suggesting a neural-specific function. The eight neuronally expressed miRNAs were abundant in the polyribosome fraction, the site of active translation. These findings are consistent with observations in *C. elegans*: *lin-4* and its target *lin-14* mRNA both associate with polyribosomes, even after *lin-4* downregulates the translation of the *lin-14* mRNA into protein, indicating that repression occurs after the initiation of translation.

11. Computational identification of miRNA targets

Sequence mismatches between the miRNA and its target mRNA 3'UTR binding sites create bulges and loops in the heteroduplex. Therefore, finding miRNA targets based on short regions of complementarity between the miRNA and its target presents a particular computational challenge. Two additional parameters can be imposed to restrict the number of candidates: 1) multiple miRNA binding sites in the 3'UTRs. The *lin-4* target, *lin-14* mRNA, contains seven *lin-4* binding sites in its 3' UTR while the *lin-41* mRNA contains two *let-7* bindings sites in its 3'UTR. However, some bona fide targets are known to have only one complementary site, so this filter excludes some targets. 2) The conservation of *let-7* in animal phylogeny also corresponds with the conservation of its target gene, *lin-41*. Therefore, we demand that the targets of the conserved miRNAs also be conserved.

mir-124 3 CGUAAGUGGCGCACGGAAU 21 .... T19B10.3 4083 GCATTCGCAAGTGGCCTTG 4101	Ortholog of human GALACTOSIDASE, BETA 1 (GLB1), implicated in gangliosidosis
mir-1 1 AUGUAUGAAGAAAUGU-AAGGU 21 ... . C33G3.1 15471 TACCTGGATCTTTACGCTTCCG 15492	DYC-1 is expressed in muscle, where it has a dystrophin-related function
mir-1 2 UGUAUGAAGAAAUGUAAGGU 21 ... . C56C10.8 955 GCAAGCTT-TCTACGTTCCG 973	Similar to human transcription factor BTF3

Using an algorithm in which we asked for multiple *let-7* binding sites in the 3'UTR of potential targets, we have identified ~30 candidates including *lin-41*. A different search demanded that the targets of the conserved miRNAs also be conserved between *C. elegans* and *C. briggsae*. We have identified predicted targets of the conserved *mir*-1 and *mir*-124 miRNAs; satisfyingly many of the predicted targets make sense for neural function in the case of *mir*-124 and muscle function in the case of *mir*-1, where our promoter fusions show these miRNAs are expressed. Our target prediction informatics is still being developed and validated and is a major aim in this renewal.

12. *C. elegans eri-1* encodes an exonuclease that is a negative regulator of RNA interference

mRNAs expressed within the nervous system are refractory to RNAi; nearly all neuronally expressed genes that are have a loss of function uncoordinated phenotype are resistant to RNAi. We performed a genetic screen for mutants with enhanced sensitivity to dsRNAs. Such genes might be expected to normally inhibit the uptake or processing of dsRNAs, or inhibit the amplification, spreading or stability of siRNAs. *unc-47*, a probable GABA transporter, is expressed within the 26 *C. elegans* GABAnergic neurons. Animals carrying an integrated *unc-47*::GFP (green fluorescent protein) fusion gene show little or no decline in GFP fluorescence following feeding on bacteria expressing GFP dsRNA. We screened for mutant animals that exhibit a dramatic decrease in the number of neurons that express GFP following feeding on GFP dsRNA. Two of the strongest enhancers of RNAi define the gene enhanced RNAi-1 (*eri-1*). The Eri phenotypes of both *eri-1* alleles, which are predicted to be null alleles, are indistinguishable and both alleles are temperature sensitive (ts) sterile. *eri-1(mg366)* animals show enhanced sensitivity to *lin-1*, *dpy-13*, *daf-19*, *myo-2*, *hmr-1*, *unc-86*, and *daf-2* dsRNAs. Genetic mapping localized *eri-1* to T07A9.5 which bears a 3'-5' exonuclease domain and a SAP domain. Consistent with this gene annotation, injection of synthetic double stranded 25bp *unc-22* siRNAs with 2 bp 3' overhangs causes 10x more progeny to exhibit an *unc-22* phenotype in *eri-1* animals than in the wild type. In addition, injected *unc-22* siRNAs are more abundant in the progeny of *eri-1* animals. *In vitro*, *C. elegans* and human *ERI-1*, partially degrade a ds siRNA with 2bp 3' overhangs, but fail to degrade single stranded siRNAs or a single stranded siRNA hybridized to an mRNA. Thus *ERI-1* inhibits RNAi by degrading the 3' overhangs of ds siRNAs.

***eri-1* and *rrf-3* mutants are more sensitive to RNAi**

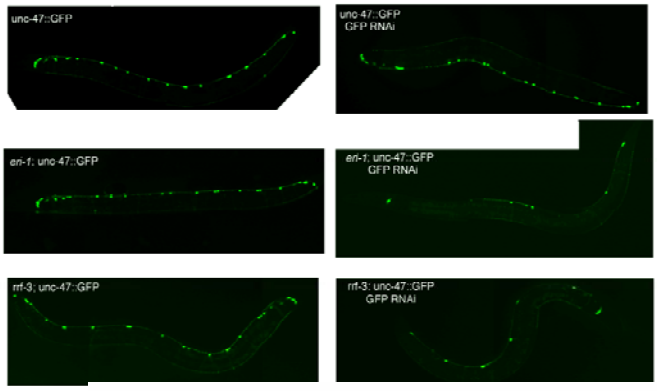
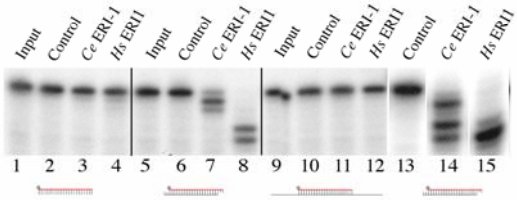
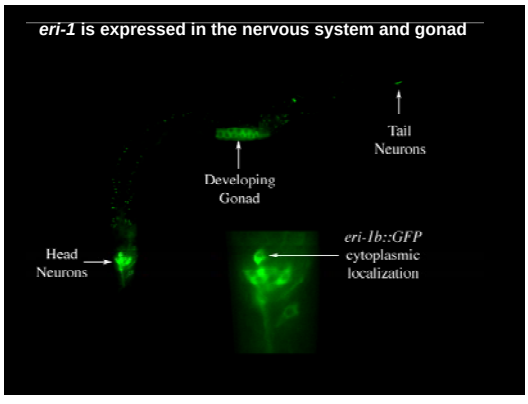


Fig. 3C



A functional *eri-1*::GFP fusion gene is expressed in a subset of neurons and in the developing somatic gonad. *eri-1* mutant animals stably expressing *eri-1b*::GFP are rescued for both enhanced RNAi and ts sterility, indicating that this fusion gene is functional and representative of endogenous *eri-1* expression. The high level expression of *ERI-1* in a subset of neurons may, at least in part, explain the relative inefficiency of RNAi within these neurons in wild type animals. Within neurons *ERI-1* is predominantly localized to the cytoplasm. In adult animals *ERI-1* is expressed in neurons and gonadal expression is restricted to the spermatheca.



eri-1 and *rrf-3* are likely to function in the same genetic pathway. The *C. elegans* genome encodes for four RNA dependent RNA polymerase (RdRPs). Two of these RdRPs are required for RNAi and function in the amplification of secondary siRNAs [91a]. Paradoxically inactivation of the RdRP *rrf-3* also causes enhanced RNAi, suggesting that this RdRP antagonizes the other RdRPs or by shunts siRNAs along a distinct pathway. Our genetic screen also identified an allele of *rrf-3*. Both mutants increase siRNA levels. *eri-1* animals also share several pleiotropic phenotypes with *rrf-3*; increased chromosome nondysjunction, as measured by the production of XO males, and temperature sensitive sterility. The *eri-1(mg366); rrf-3(pk1426)* double mutant does not exhibit additional enhanced RNAi phenotypes or synthetic developmental phenotypes compared to the single mutant animals. The ts sterility of *eri-1* null mutant animals is due to defective sperm development; the sterility can be rescued by mating to wild type males or *eri-1* males grown at the permissive temperature. DAPI stained *eri-1* gonads reveal a normal mitotic expansion of the germ line but sperm nuclei exhibit gross morphological defects. *rde-1*, *rde-4*, and *sid-1* fail to suppress the sperm defect in *eri-1* animals. Loss of *ERI-1* or RRF-3 may induce misregulation of endogenous miRNAs normally required for proper spermatogenesis. The *eri-1* ts sterility is suppressed by RNAi of the piwi factor

prg-1/prg-2, homologs of which mediate spermatogenesis in mammals and Drosophila. This supports a connection between piwi function and *eri-1*/rrf-3 roles in spermatogenesis. We will explore this more fully in Aim 4.

The chromosome nondysjunction of *eri-1* and rrf-3 mutants may be analogous to the involvement of an RNAi like machinery in centromere silencing of *S. pombe* (43b) - endogenous siRNAs or miRNAs produced by *eri-1*/piwi in for example the spermatheca may in turn regulate centromere structure for proper mitotic segregation of X and perhaps other chromosomes.

Of the more than 100 mutants identified in our enhanced RNAi screen 4 alleles define the genes *rrf-3* and *eri-1*, 5 additional alleles that share pleiotropic phenotypes with *eri-1* and *rrf-3*: increased chromosome nondysjunction, as measured by the production of XO males, and a 100% penetrant temperature sensitive (ts) defect in sperm formation. These mutants, like *eri-1* and *rrf-3*, show increased sensitivity to dsRNAs targeting a broad range of mRNAs; including mRNAs expressed both within neurons and in other tissues. We have mapped these 4 eri loci to small genetic intervals.

All of the enhanced RNAi loci identified in our genetic screens to date will be tested for utility in RNAi screens. Those mutant loci (or double mutants if necessary) that now enable RNAi in the nervous system will be used for RNAi screens, where we expect a neural focus such as suppression of *mir*-124 over-expression (Aim 3), or affecting miRNA functions in localized neuronal translation (Aim 3). *eri-1* is already being used for such screens.

13. *C. elegans* Retinoblastoma Pathway Genes Inhibit RNA Interference

C .elegans homologues of the retinoblastoma (Rb) tumor suppressor protein and associated cofactors specify vulval cell lineages during development. We have found that mutations in multiple Rb pathway components enhance RNA interference (RNAi) and cause somatic cells to express a germ line-specific gene. Particular gene inactivations that disrupt RNAi also suppress the Rb-mediated soma to germ line cell fate transformation and reverse the cell lineage transformations of Rb mutants. We suggest that Rb may act by a similar mechanism to transform mammalian cells. The Rb pathway regulates RNAi synergistically with other negative regulators of RNAi, the ERI-1 siRNase and the RRF-3 RNA-dependent RNA polymerase. The enhanced RNAi response in mutants defective in both the Rb and *eri-1* pathways extends the use of feeding RNAi screens to allow systematic analysis of neuronal signalling in behaviour and physiology.