Sackler Colloquia of the National Academy of Sciences

Biology of RNAi

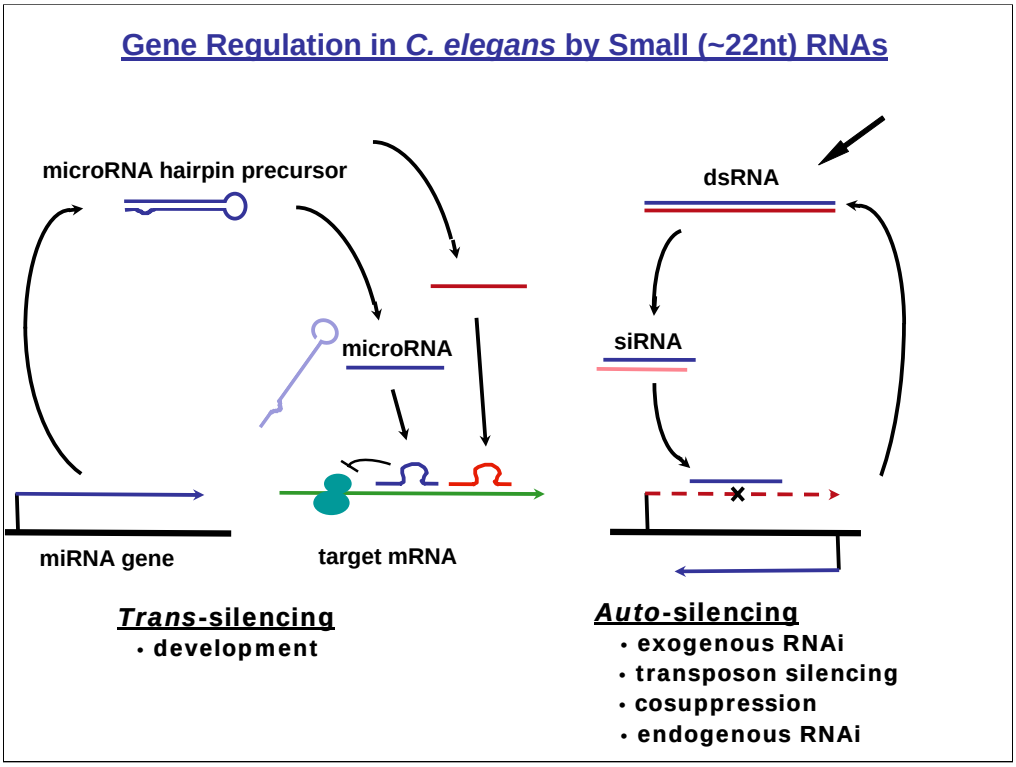
May 17-18 2004

Gene Regulation in *C. elegans* by Small RNAs

Victor Ambros

Dartmouth Medical School Department of Genetics

I am sincerely grateful to the NAS and the Sackler Foundation for sponsoring this meeting, and to Philip Sharp and Andy Fire for inviting me to speak. I learned a lot from the other speakers and participants. This was a very special meeting indeed. What follows is a summary of my presentation at the meeting.



One can think in terms of two chief modes of gene regulation in C. elegans that involve small RNAs of about 22 nt in length. These are thought to be chiefly negative regulatory mechanisms that "silence" gene expression. In the trans-silencing mode (1), a microRNA is transcribed from a noncoding gene as part of a precursor that forms a hairpin structure. That hairpin is processed by the Dicer enzyme to produce the mature ~22 nt microRNA. The microRNA is complementary to sequences in the 3' untranslated region (UTR) of target messenger RNAs (mRNAs). As is generally the case for animal microRNAs, this complementarity is partial, and the consequences of the action of the microRNA is thought to be repression of translation without destabilizing the mRNA (8). In the second, auto-silencing mode (2), the small interfering RNA (siRNA) is produced from a longer double-stranded precursor by Dicer, but the double-stranded RNA originates from the same sequence as the siRNA. The precise match between the siRNA and its target results in cleavage and degradation of the target. This is the familiar RNAi mechanism.

[NOTE: See last page of this presentation for a list of references.]

Gene regulation in *C. elegans* by Small RNAs

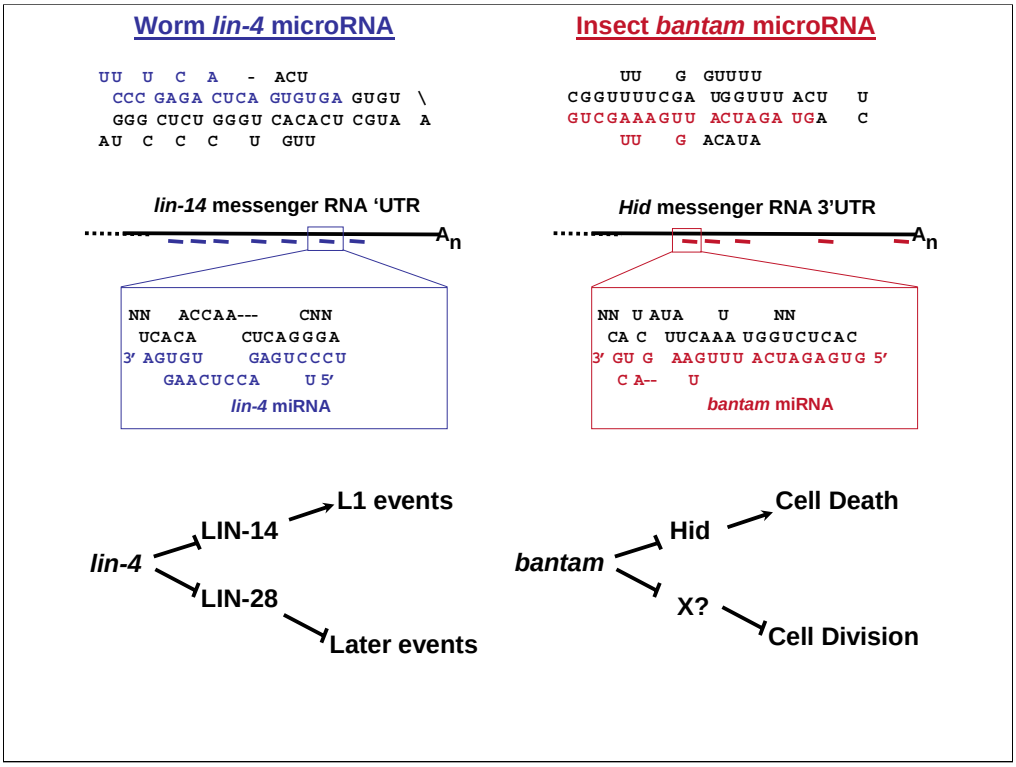
Overview of microRNA genes

- **Developmental timing mechanism**
  - **Collaboration within and among microRNA families**

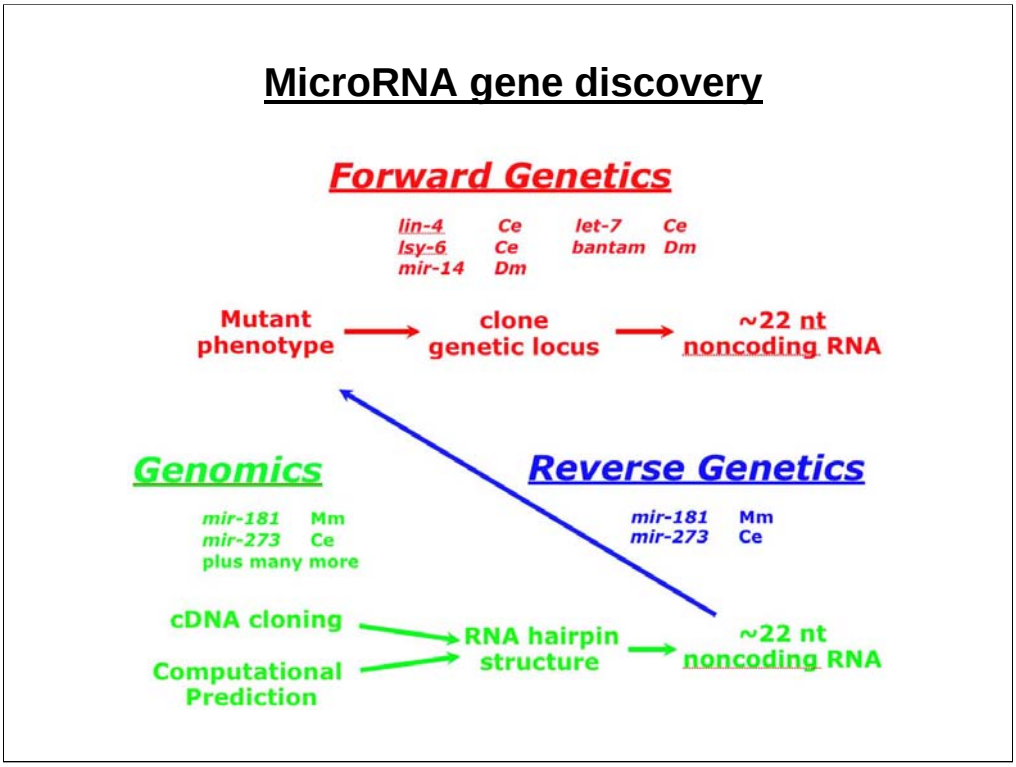
Other small RNAs in worms

- **Tiny non-coding RNAs**
- **Endogenous RNAi**

Here, we will discuss two major points about small regulatory RNAs in C. elegans. The first point will emerge from a summary of some of the genetic data underlying certain aspects of how microRNAs function to control developmental timing in C. elegans. The main idea here will be that microRNAs can (and may generally) collaborate to repress the expression of a given target mRNA. The second point regards the production of diverse small interfering RNAs (siRNAs) in normal worms, suggesting that RNAi mechanisms contribute to the general level of gene expression on a genome scale in C. elegans.



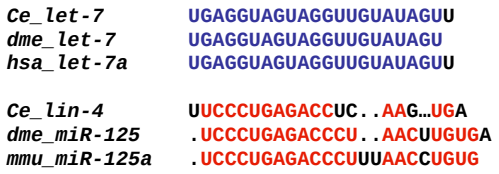
By way of introduction to microRNAs in animals, it should be noted that distinct microRNAs in different animals seem to function by similar mechanisms, but often in quite distinct contexts. For example, in worms, the lin-4 microRNA represses the translation of lin-14 and lin-28, and thereby reduces the level of these proteins at later stages of development (18, 23), permitting the timely choice of developmental events specific to later stages. By analogy, the bantam microRNA of drosophila controls choices between cell death (through regulation of the synthesis of the Hid cell death effector) or cell division (25).



Lin-4 and bantam are examples of microRNAs that were discovered by standard forward genetics followed by cloning of the mutant locus. Noteworthy among the microRNAs recently identified in C. elegans by forward genetics is the product of the lsy-6 locus, which functions in a pathway controlling the left-right asymmetry of a part of the worm nervous system (28). MicroRNAs have also been discovered by genomics--cDNA cloning and/or computational predictions based on comparative genomics (9-17, 30). The functions of microRNA genes discovered by genomics are determined by reverse genetics, which generally involve gene knockout (as will be discussed later), and/or over expression experiments (32).

- ***C. elegans* miRNA genes**

  ~ 110 distinct genes
  ~ 50% developmentally regulated
  ~ 30% conserved in insects/vertebrates



There are about 110 microRNA genes identified in C. elegans, and these probably represent most or all of the genes that produce relatively abundant and/or phylogenetically conserved microRNAs (9, 10, 12, 14). Some are remarkably well conserved phylogenetically, suggesting that they function in conserved pathways and/or they recognize multiple, phylogenetically-conserved targets.

- ***C. elegans* miRNA genes**

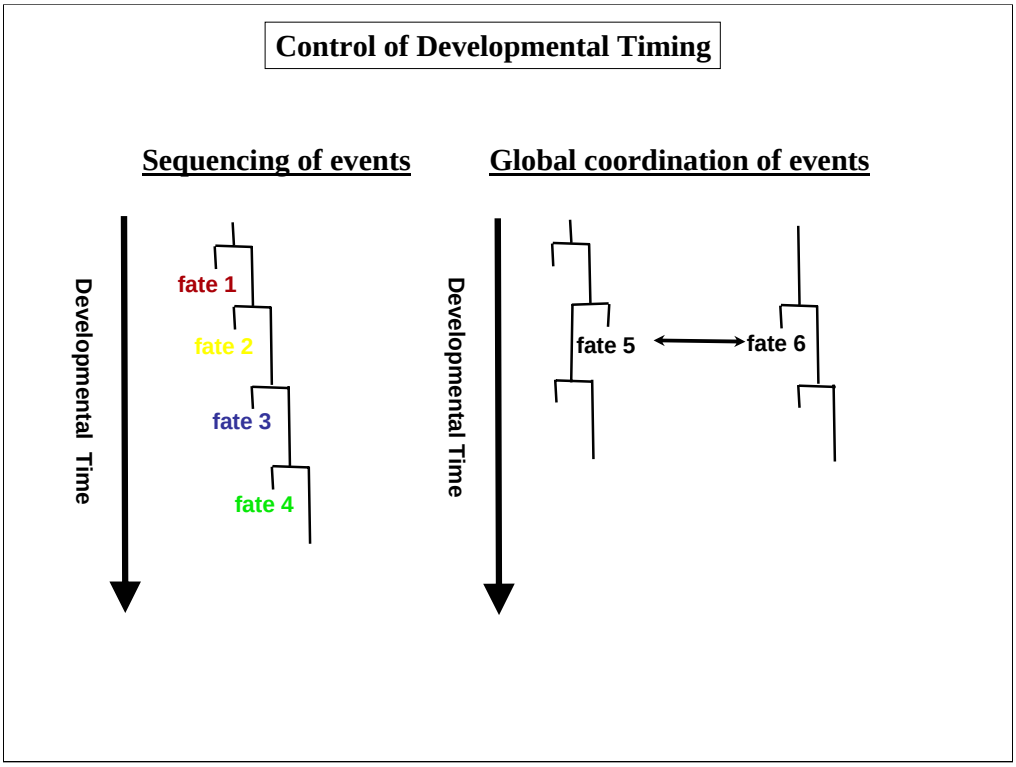
  ~ 110 distinct genes
  ~ 50% developmentally regulated
  ~ 30% conserved in insects/vertebrates

  microRNA families:

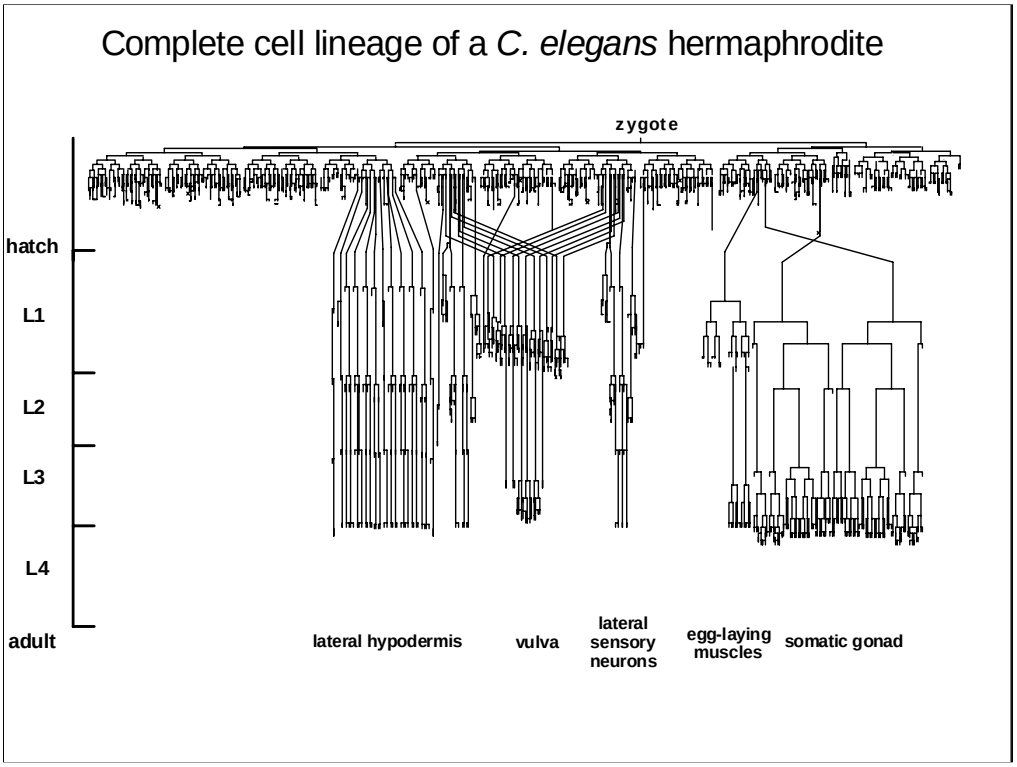
```
Ce_lin-4        UUCCCUGAGACCUCAAGUGUGA
Ce_miR-237      .UCCCUGAGAAUUCUCGAACAGCUU

Ce_let-7        UGAGGUAGUAGGUUGUAU.AGUU
Ce_mir-48       UGAGGUAGGCUCA.GUAG.AUGCGA
Ce_mir-84       UGAGGUAGUAUGUAAUAUUAUA
Ce_mir-241      UGAGGUAGGUGCGAGAAAUAA
```

A prominent characteristic of microRNAs is that they come in families; for example, let-7 of C. elegans is similar in sequence to three other genes, mir-48, mir-84 and mir-241 (9, 10, 12, 14). These four microRNAs share a block of identical sequence near their 5' ends, in a region thought to be particularly important for target recognition.

The next part of the discussion will address how these let-7 family microRNAs functionally interact to control the timing of developmental events in C. elegans.

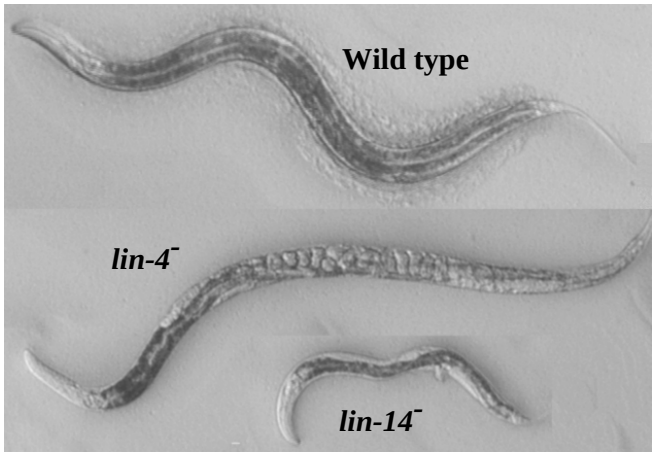


Developmental timing can be thought to involve two kinds of regulatory issues: First, within cell lineages, such as the idealized stem cell lineage shown at left, developmental events (or the expression of specific cell fates) must be regulated in such a way that they occur in the proper sequence. Second, as depicted on the right, cell fates must be coordinated throughout the animal. In C. elegans, these functions are managed by the heterochronic gene pathway, which interact to control sequences of developmental events within cell lineages, and, since these genes affect diverse cell types, they also contribute to the global coordination of events (3).

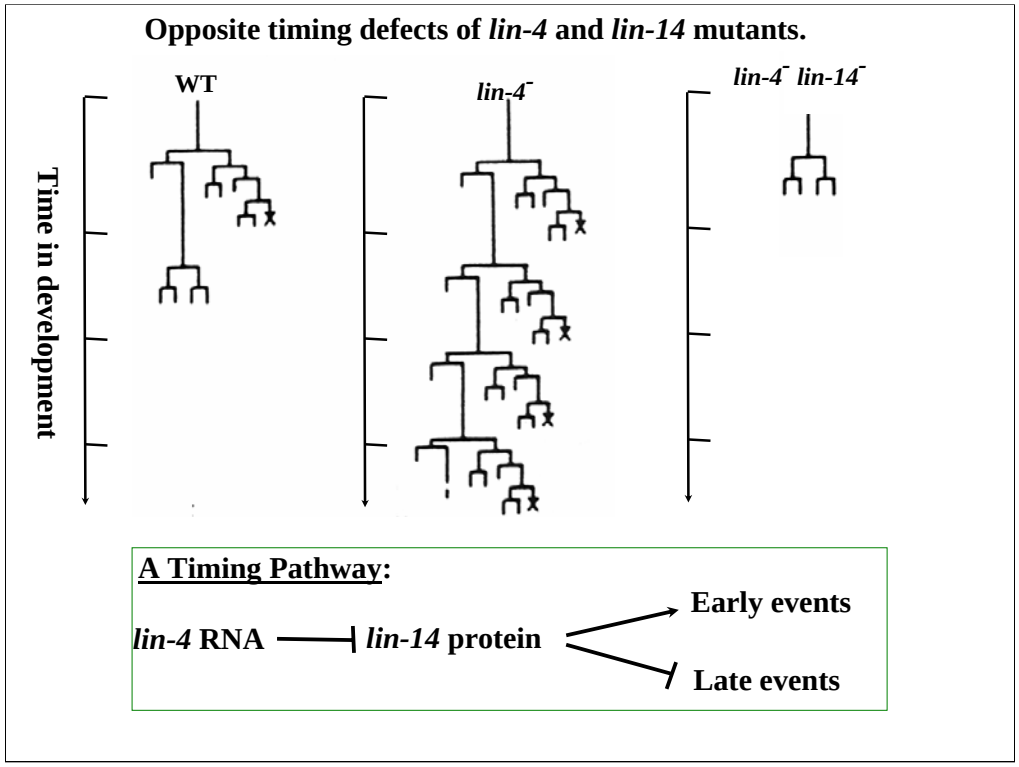


The description of the complete cell lineage of C. elegans (6) illustrates the precision with which developmental events are scheduled in the worm, and provides a a basis for the identification and characterization of mutants defective in developmental timing (7).

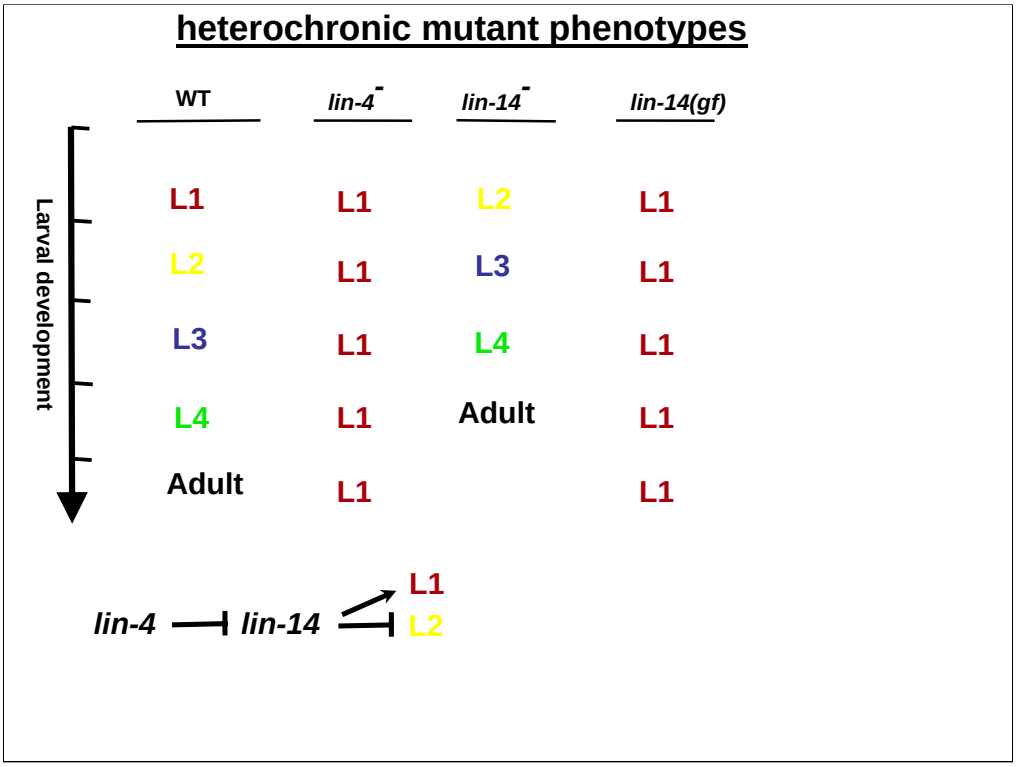
Opposite timing defects of *lin-4* and *lin-14* mutants



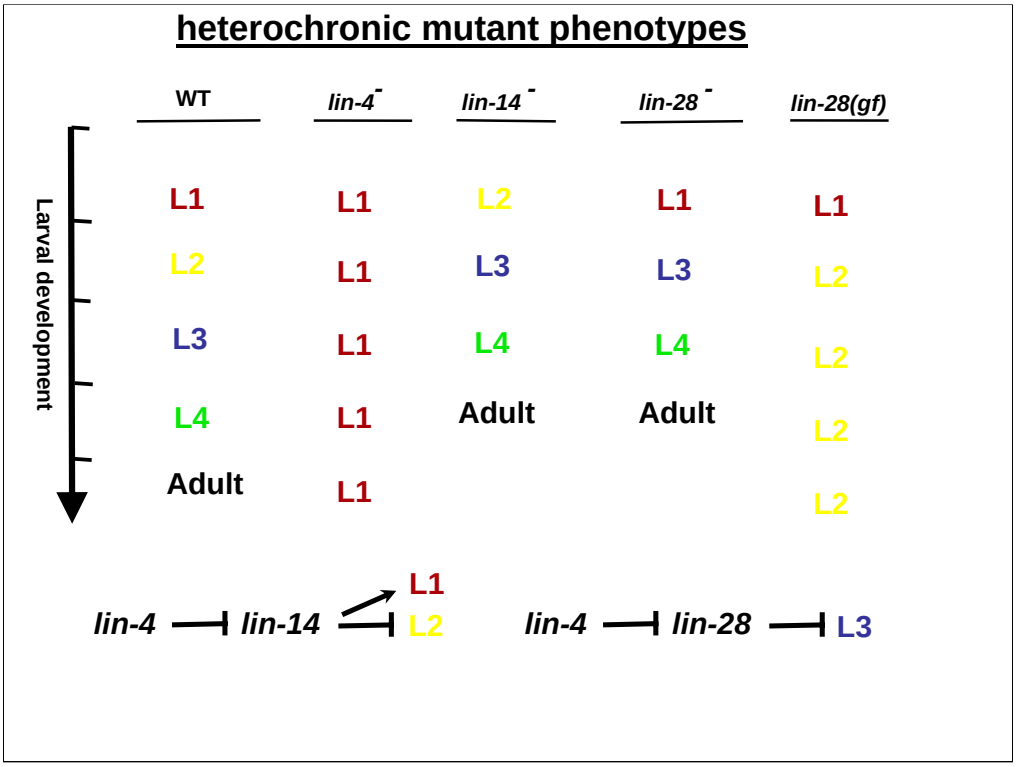
The study of developmental timing in C. elegans began in earnest in the lab of Bob Horvitz, through the genetic analysis of mutants with opposite developmental timing defects (7, 36).



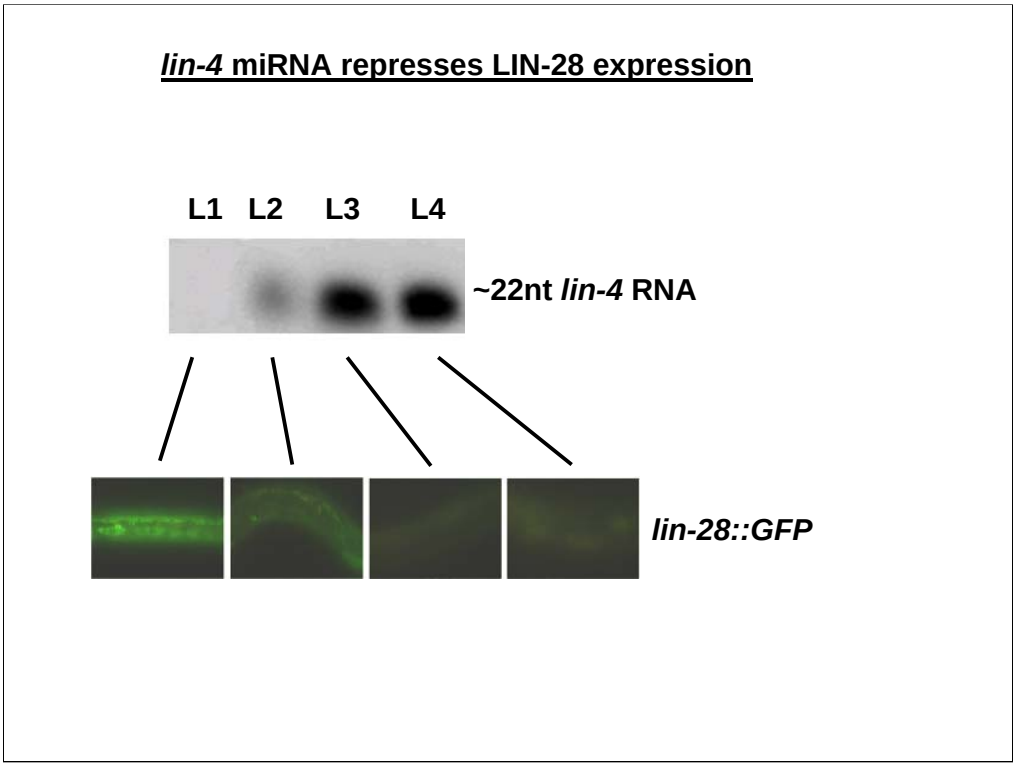
lin-4 loss-off-function(lf) mutants repeat first larval stage, whereas lin-14(lf) mutants skip the first larval stage (7). Epistasis analysis led to the model that lin-4 is a repressor off lin-14, and that lin-14 decreases during development to specify the progressive expression of L1 and L2 events (37).



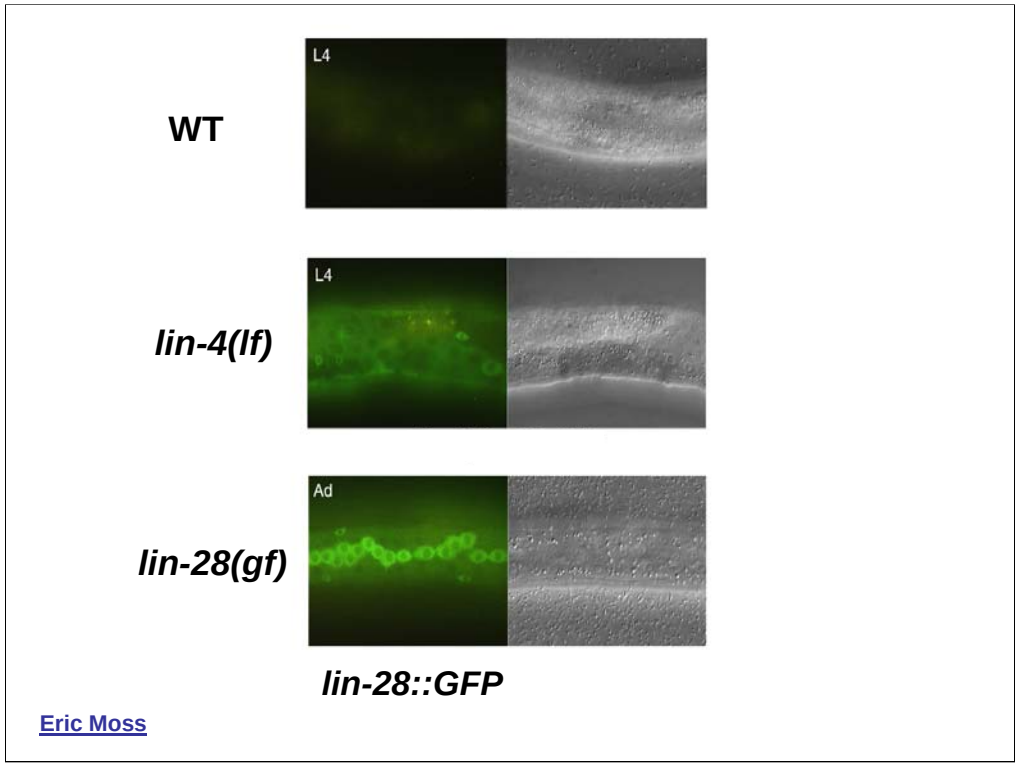
A shorthand depiction of this model, with the addition of the phenotype of lin-14 gain-of-function (gf) mutants, which is identical to that of lin-4(lf) animals. Lin-14(gf) mutations are deletions of lin-4 complementary elements from the lin-14 3' UTR (21).



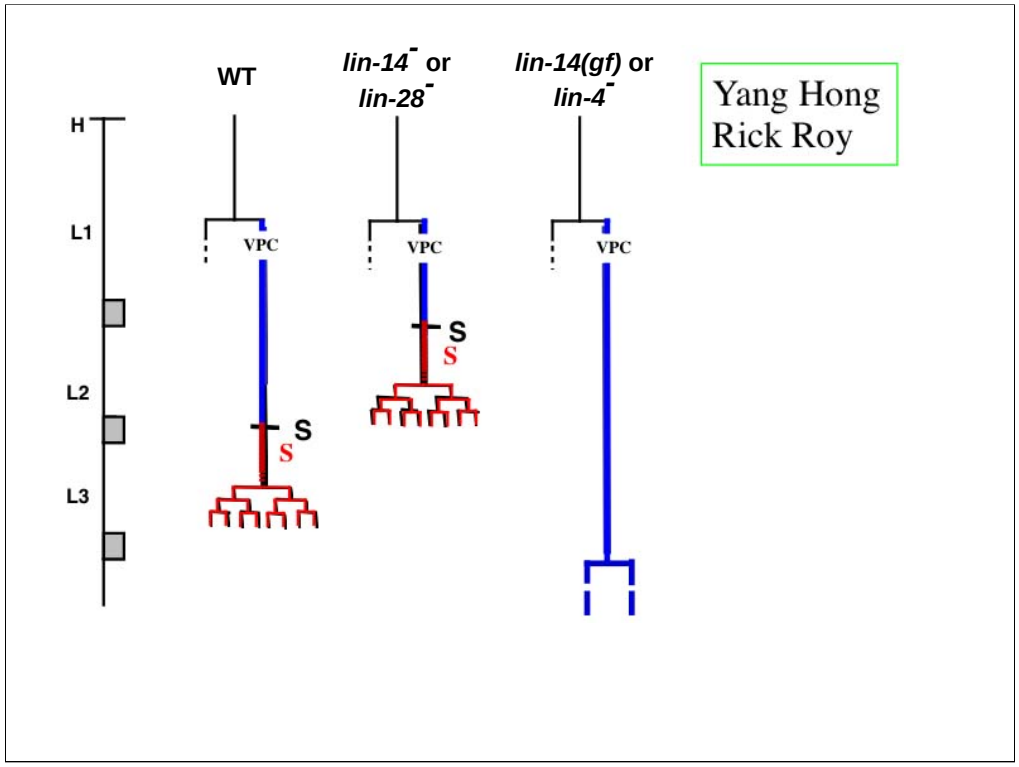
Expression of the sequence of events L1-L2-L3 involves lin-14 and another gene, lin-28. Lin-28 protein also decreases during development, and lin-28(lf) , mutants skip the L2. Lin-28(gf) mutation, which removes a lin-4-complementary element from the lin-28 3' UTR, results in the repetition of the L2 events (23). The model, can therefore be expanded to propose that lin4 represses both lin-14 and lin-28.



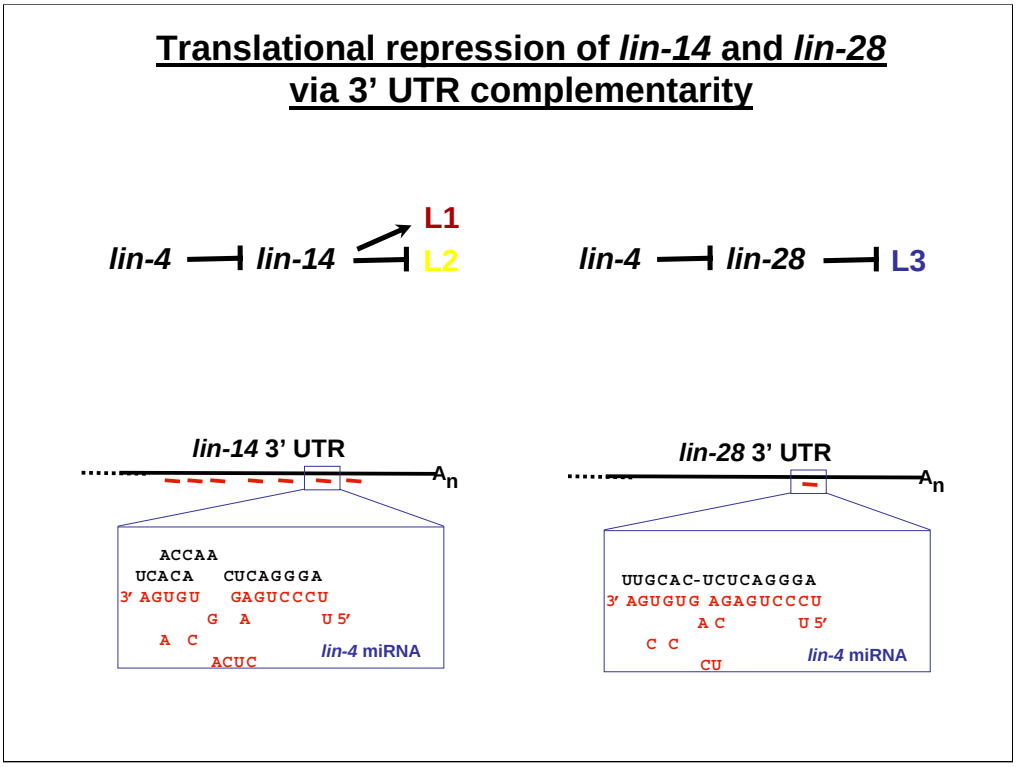
Here is a Northern blot demonstrating the up-regulation of lin-4 RNA after the first stage, and the correponding decrease in LIN-28 protein, visualized using a LIN-28::GFP fusion protein in living worms.



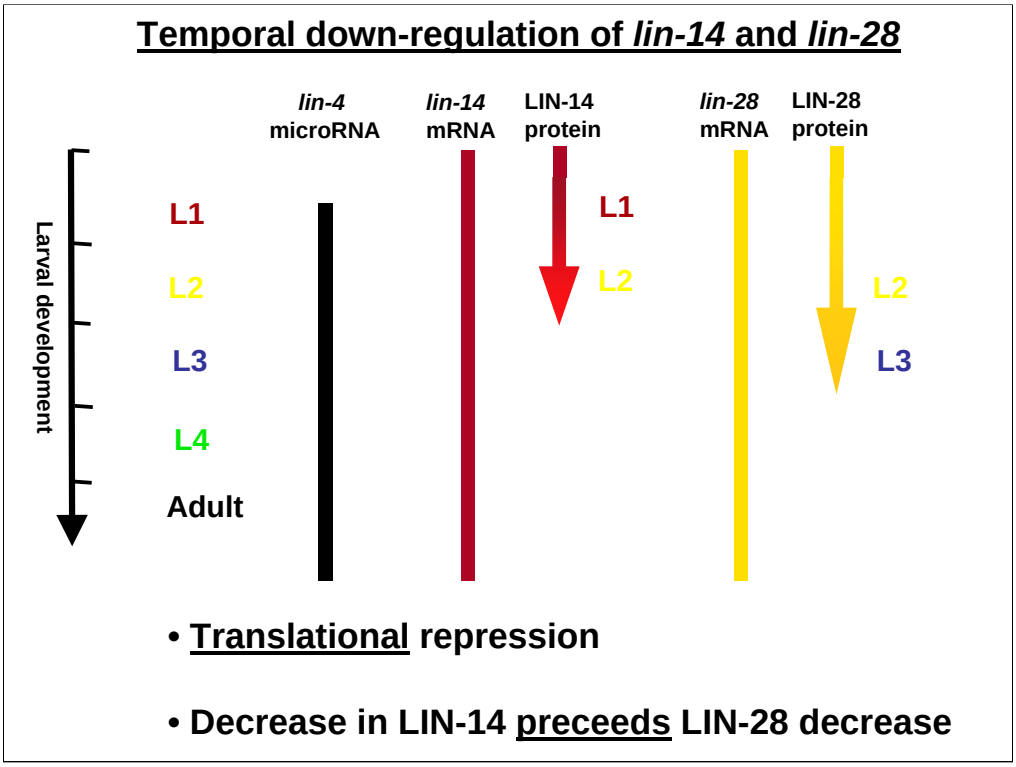
The same LIN-28::GFP fusion protein is not down-regulated in the absence of lin-4, or in the absence of the lin-4 complementary element in the lin-28 3' UTR.



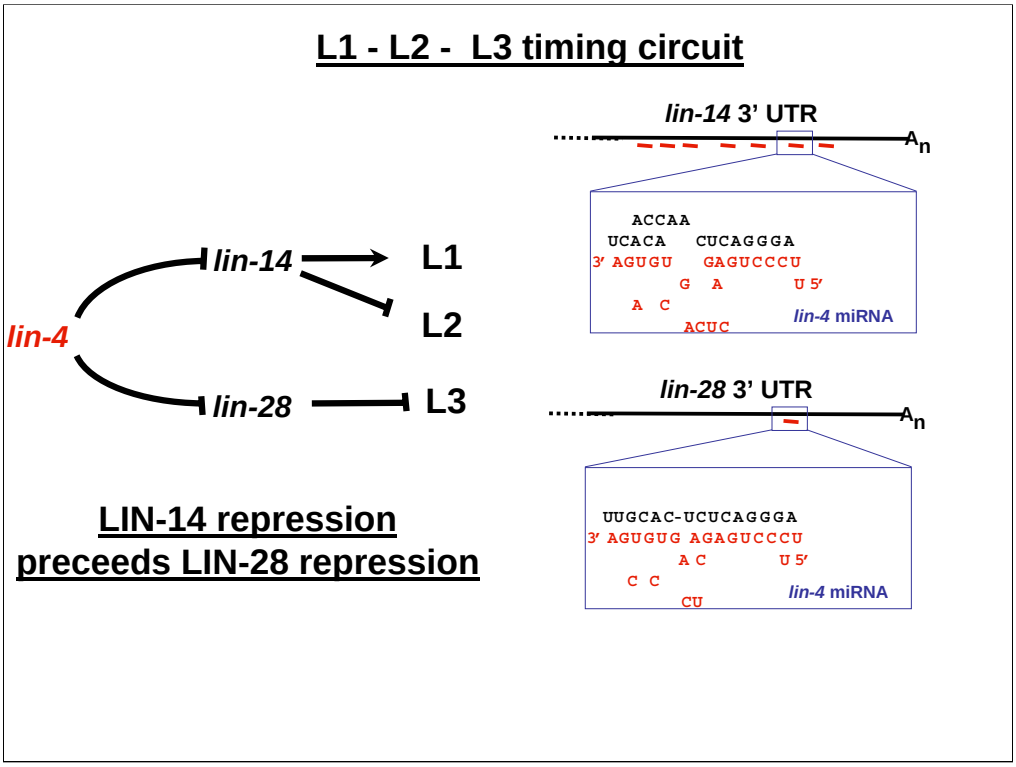
Here is an example of an L3 event controlled by the lin-4-lin-14-lin-28 heterochronic gene pathway. In this case, the G1-to-S transition in a particular cell type is controlled by the pathway. The target of the timing pathway is, in this case, expression of cki-1, a p27 homolog (34).



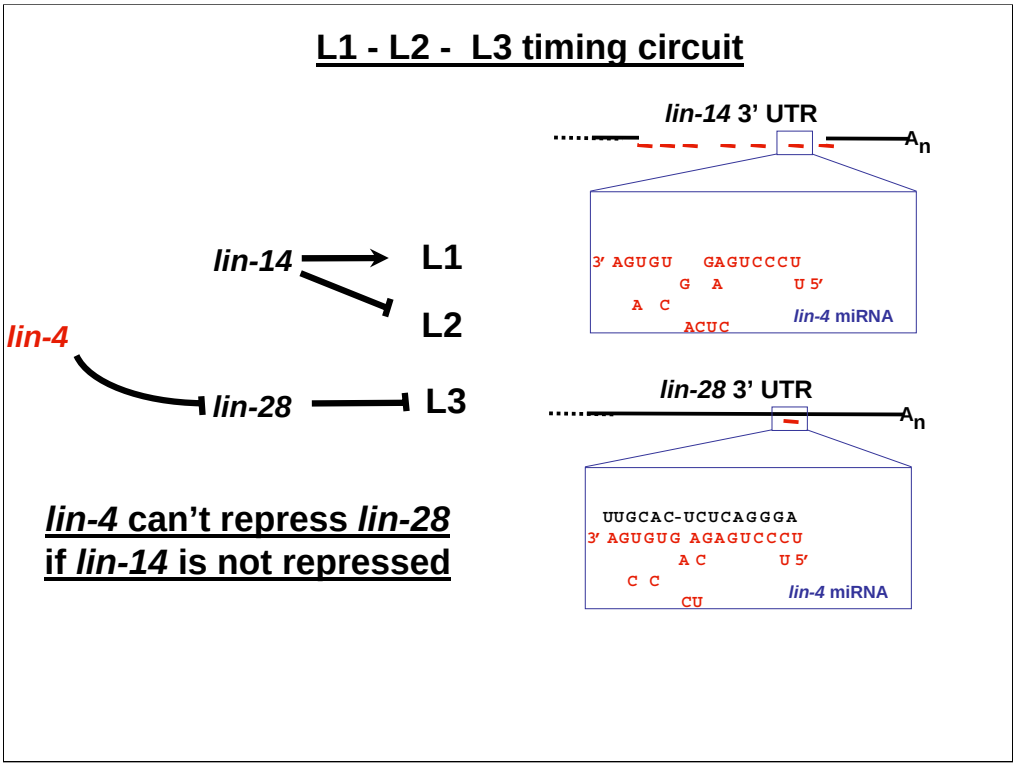
How does lin-4 coordinate the repression of lin-14 and lin-28 protein expression?



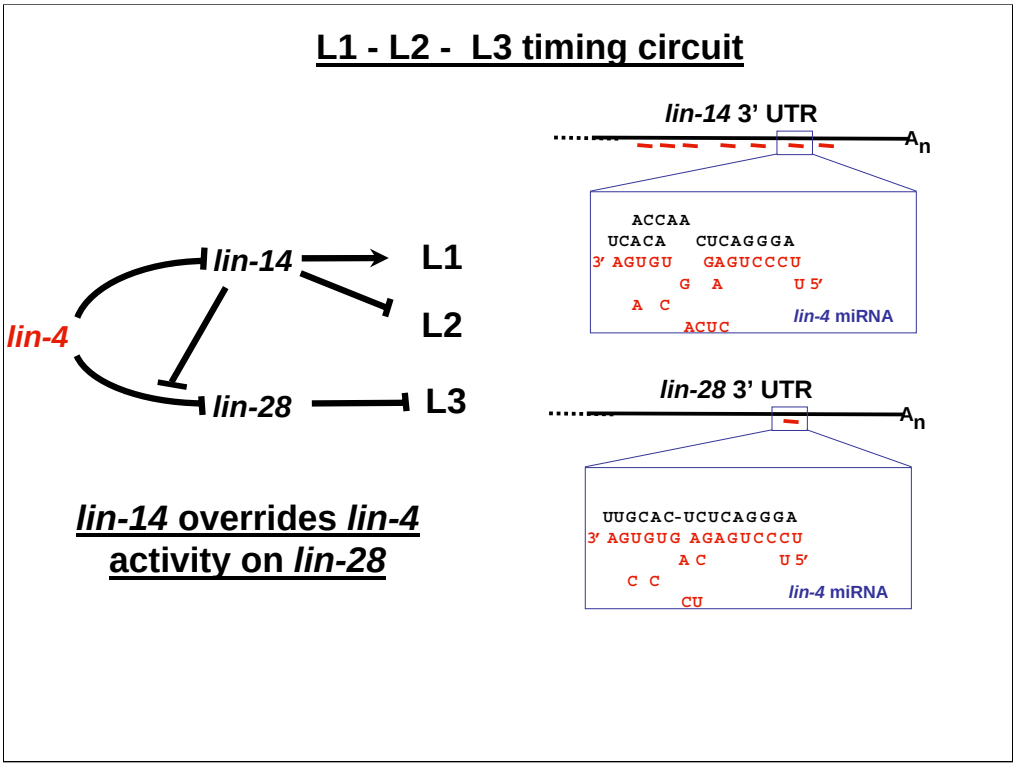
Note that in order for the L2 programs to follow the L1, and to properly precede the L3 programs, LIN-144 protein must be down-regulated prior to down-regulation of LIN-28.



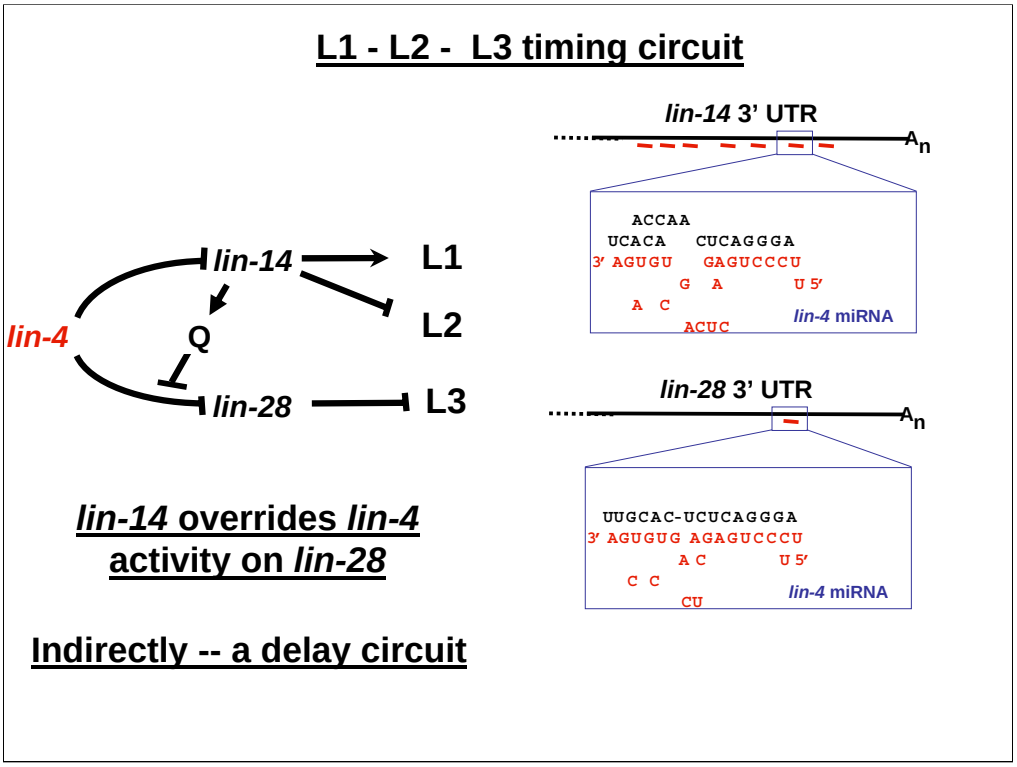
A simple model for lin-4 action in the coordination of L1 - L2 - L3 events.



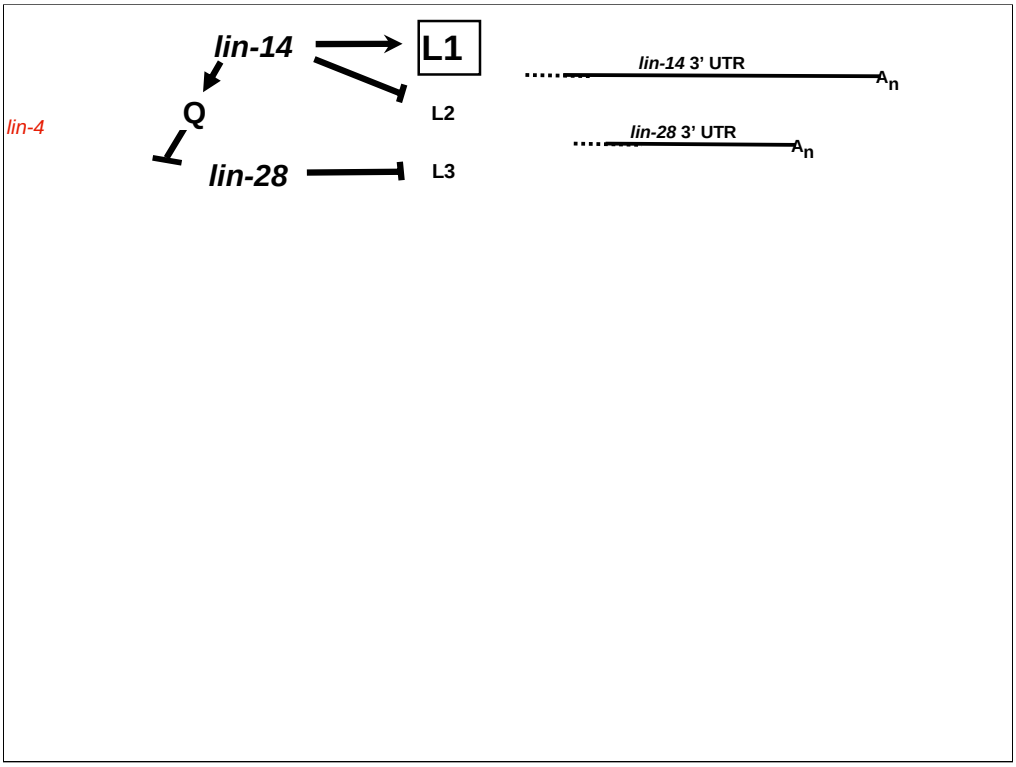
A provision must be added to the model to account for the observation that in a lin-14(gf) mutant, lin-28 is not repressed (23).



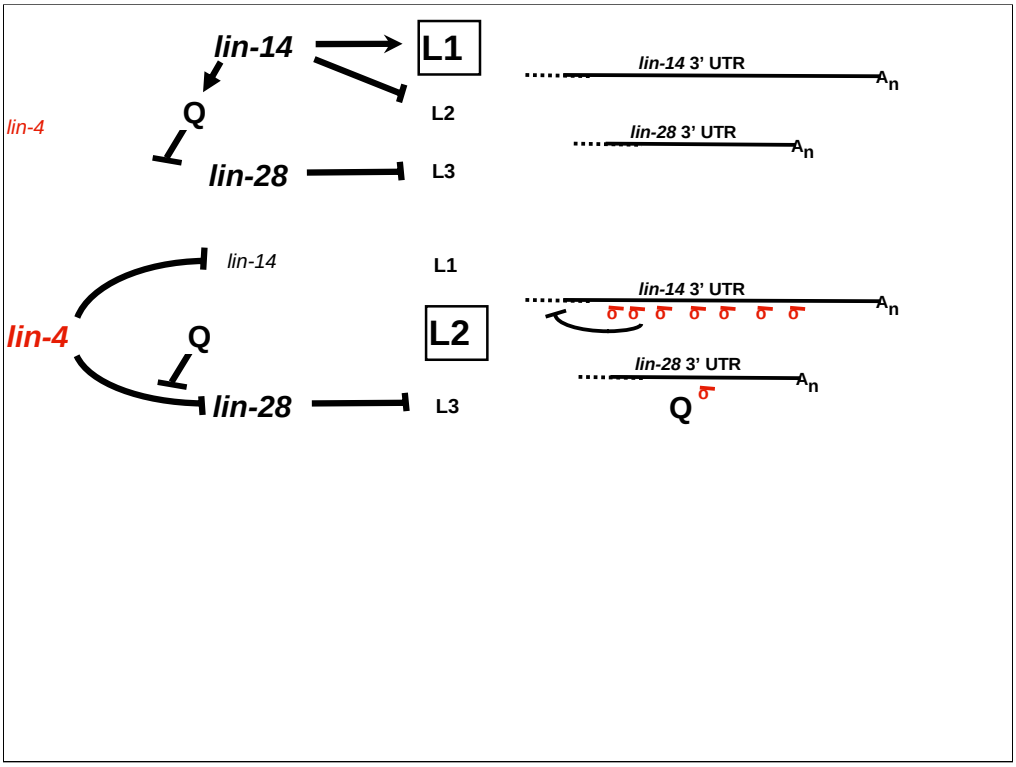
Therefore, a positive regulatory pathway is proposed whereby an activity controlled by lin-14 can overcome the activity of lin-4.



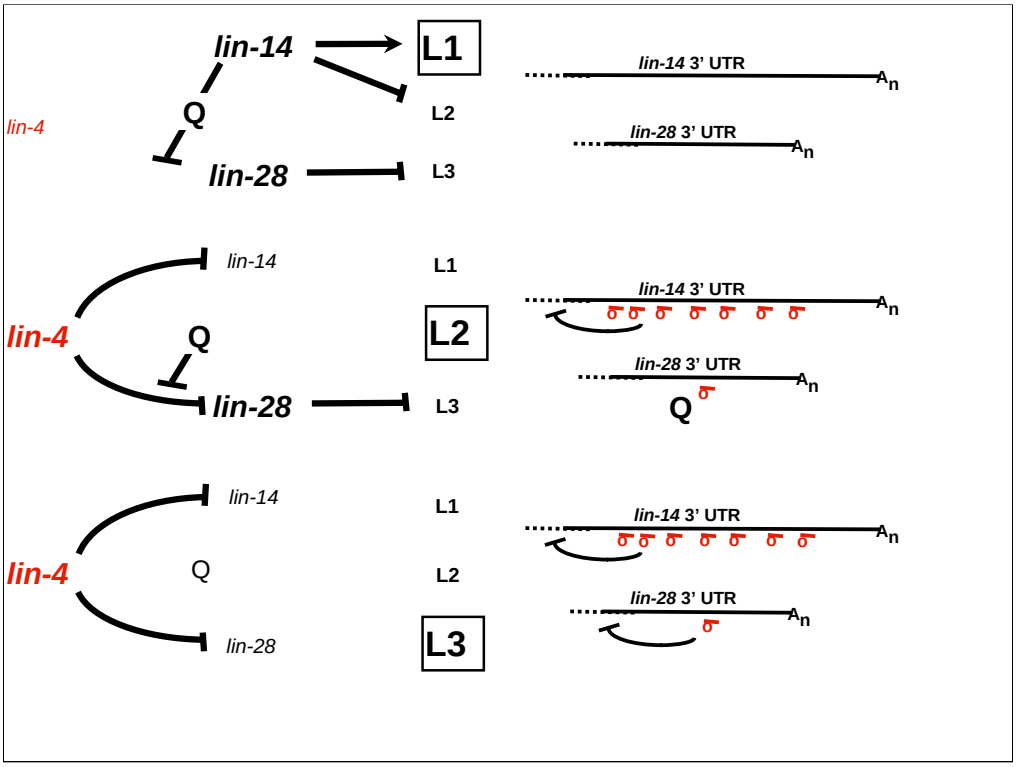
This pathway has been proposed in a somewhat different formulation by Moss and coworkers, who show that the pathway is fundamentally a repressive mechanism. Therefore a hypothetical negative regulatory intermediate component “Q” is proposed.



The model now takes on a dynamic behavior to account for L1 -L2 - L3 developmental events: First, lin-14 and lin-28 are abundant, and Q is active……

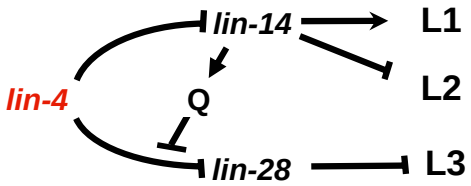


In the second larval stage, lin-4 microRNA accumulates, and lin-14 is repressed, allowing for the expression of L2 fates. Note that according to this model, lin-28 protein remains abundant on account of the activity of Q, which over-rides lin-4's ability to repress LIN-28 synthesis.



Finally, in the third stage, Q becomes inactivated, permitting the repression of lin-28 by lin-4.

L1 - L2 - L3 timing circuit



***lin-14* controls a delay in *lin-28* repression**

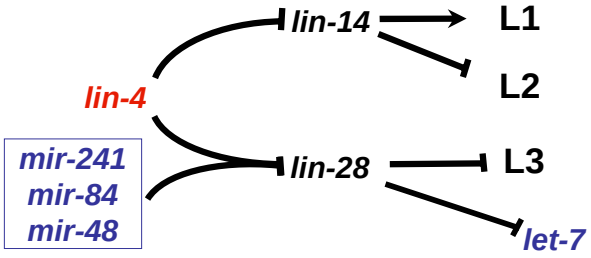
How does the “Q” pathway work to inhibit the efficacy of lin-4 microRNA?

(Note that lin-4 is expressed at normal levels in a lin-14(gf) mutant, so the block to lin-4 is not through control of lin-4 expression.)

Do *let-7* family microRNAs function in the L1 - to - L3 timing circuit?

Worm microRNA knockouts

Allison Abbott	V. Ambros
Ezequiel Alvarez-Saavedra	H.R. Horvitz
Eric Miska	H.R. Horvitz
Nelson Lau	D. Bartel

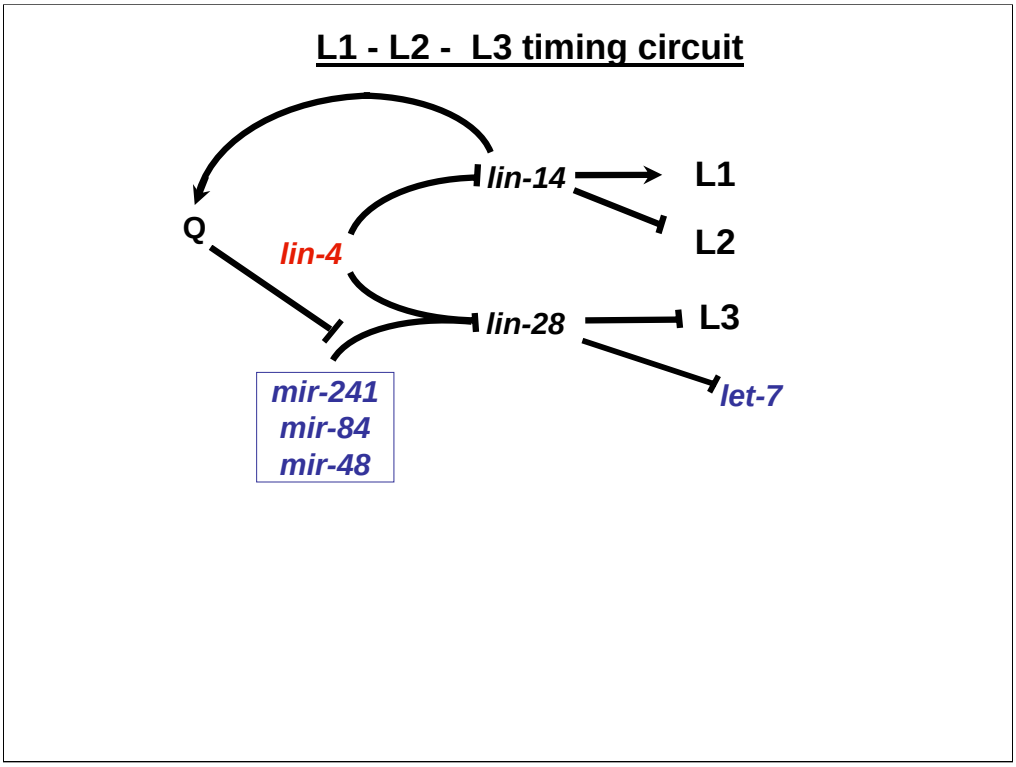


Part of the answer comes from considering the roles of other microRNAs that also repress lin-28, working together with lin-4.

This and the next few diagrams represent a summary of unpublished work from a collaboration between the Ambros lab at Dartmouth Medical School, the Bartel lab at MIT and the Whitehead Institute, and the Horvitz lab at MIT (and Howard Hughes Medical Institute).

In short, the current data suggests that lin-4 is not the only microRNA regulating LIN-28. Analysis of the phenotypes of mir-48, mir-84 and mir-241 knockout mutations indicate that these three let-7 family miRNAs act redundantly to supplement the activity of lin-4 after the L2 stage, and thereby they provide a critical impetus for the down-regulation of LIN-28 after the L2.

The fourth let-7 family member, let-7 itself, acts downstream of lin-28, while mir-48, mir-84 and mir-241 act upstream of lin-28. Thus, let-7 functions more directly to control late larval programs (L4 and Adult fates), while mir-48, mir-84 and mir-241 function to control L2 and L3 fates.

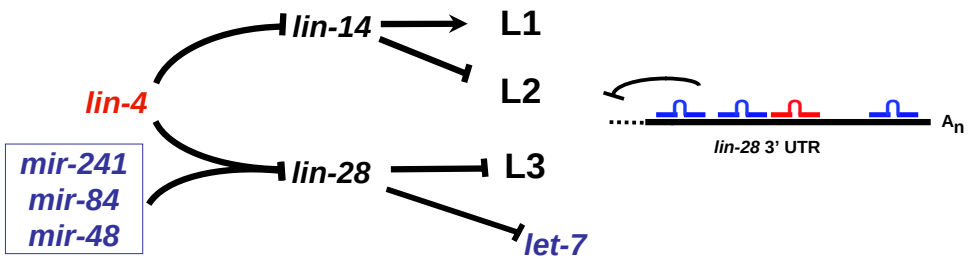


Recall that the hypothetical “Q” activity, is invoked to account for the fact that lin-14 activity blocks the repression of LIN-28 by lin-4, and that normally, LIN-14 is repressed before LIN-28. Q is provisionally proposed to regulate the activity of mir-48, mir-84 and mir-241. That hypothesis is currently being tested genetically.

- **A <u>family of microRNA genes</u> (eg, *let-7 family*) can collaborate to repress translation of a target messenger RNA.**

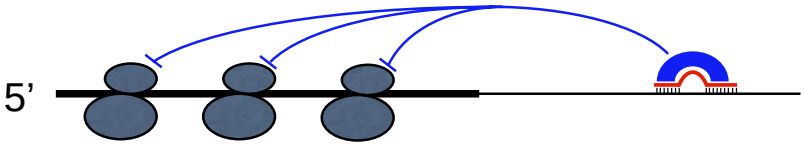
- **<u>Distinct microRNA genes</u> (eg, *lin-4* and *let-7* family) can collaborate to repress translation of a target.**

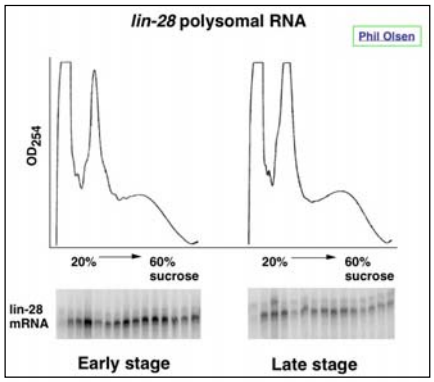
- **Developmental events can be scheduled through <u>progressive activation</u> of collaborating microRNAs**



These results illustrate principles of let-7 family microRNA function that we believe may be generally applicable to other microRNAs as well.

Post-initiation translational inhibition



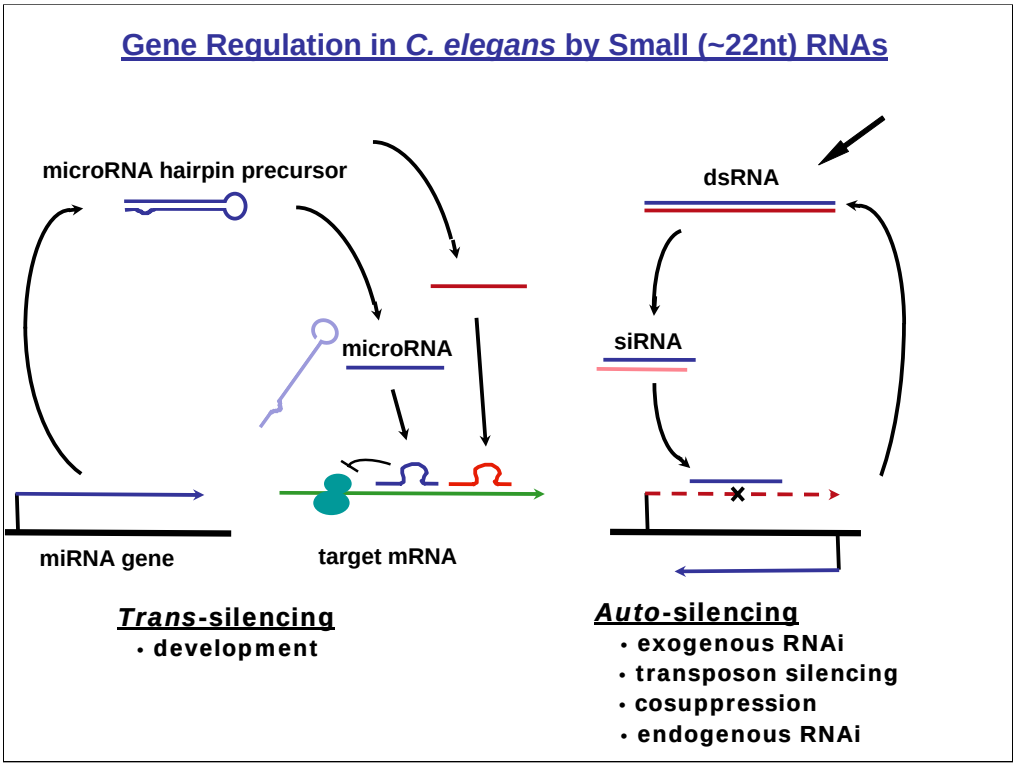


What is the mechanism of translational repression of lin-28 by lin-4 microRNA, in combination with let-7 family microRNAs? From experiments involving polyribosome fractionation, conducted in the Ambros lab (8), and in the lab of Eric Moss (46), the following model is suggested:

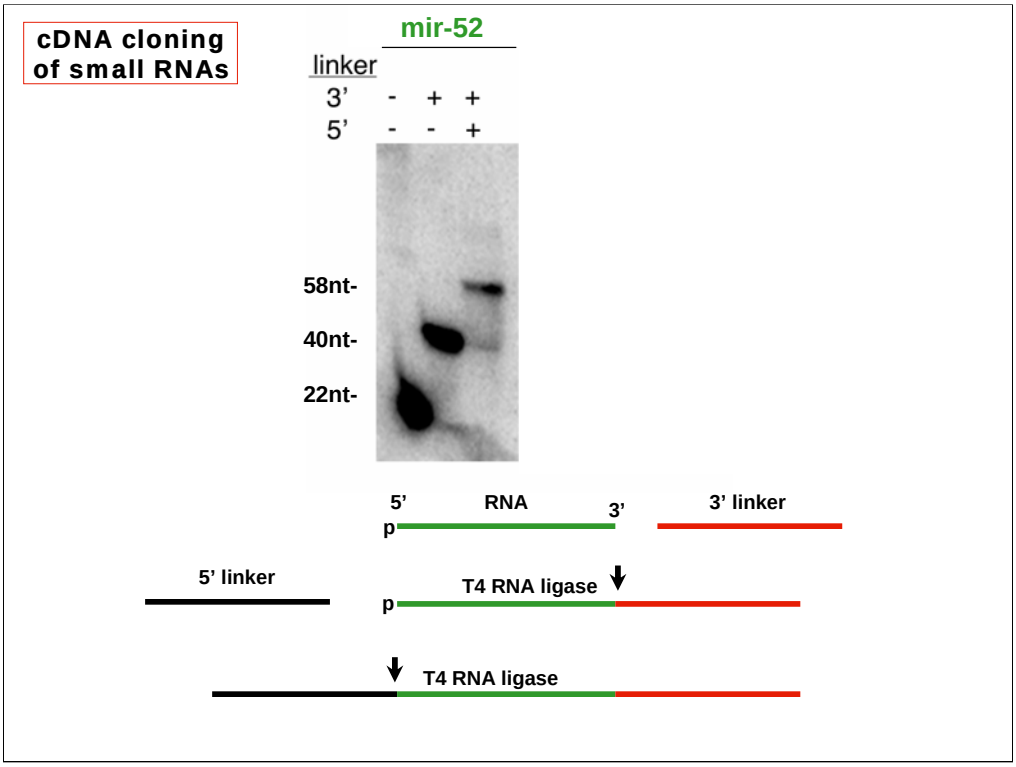
1) The microRNA(s) recruit to the 3’ UTR of the target mRNA a translational repressor protein (or protein complex).

2) The repressor does not impede loading of ribosomes onto the mRNA, or the progression of the ribosomes along the mRNA.

3) Somehow, the repressor blocks the production of stable protein product. This could be accomplished by co-translational proteolysis, or by tagging the protein for subsequent degradation.

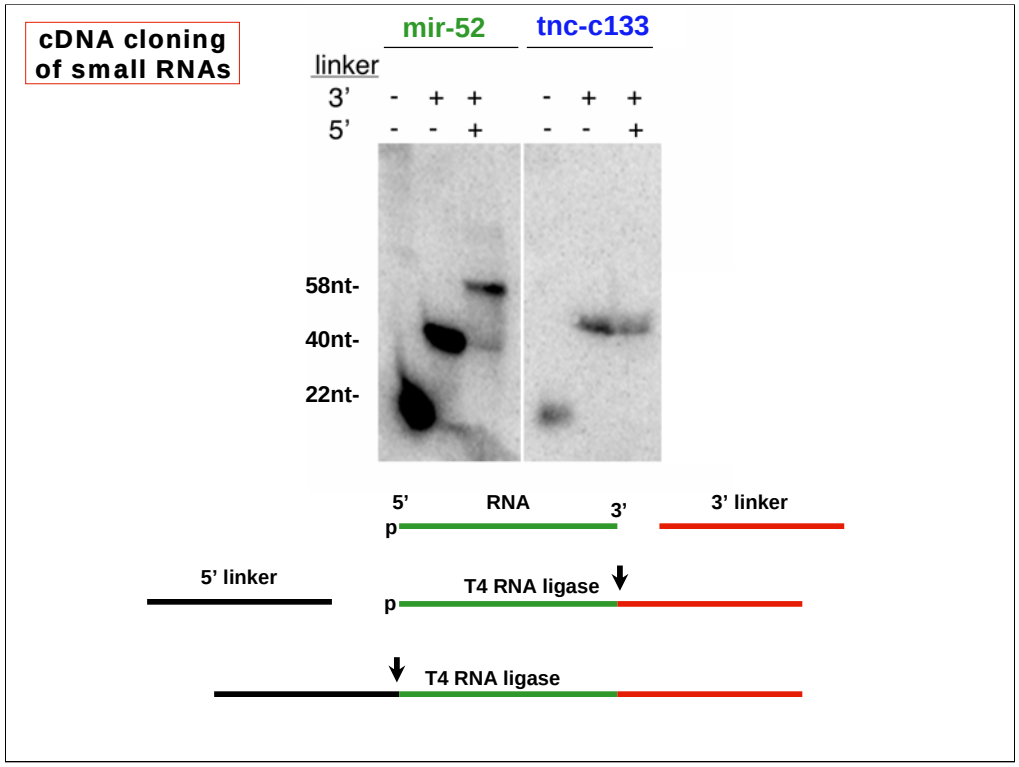


In the course of efforts to clone and sequence cDNAs corresponding to C. elegans microRNAs, we also identified endogenous small RNAs invovled in auto-silencing of worm genes (12). These RNAs include siRNAs invvolved in endogenous gene silencing on a genomic scale. We also found a third class of small RNA of about ~22 nt in length, called tncRNAs (tiny noncoding RNAs).

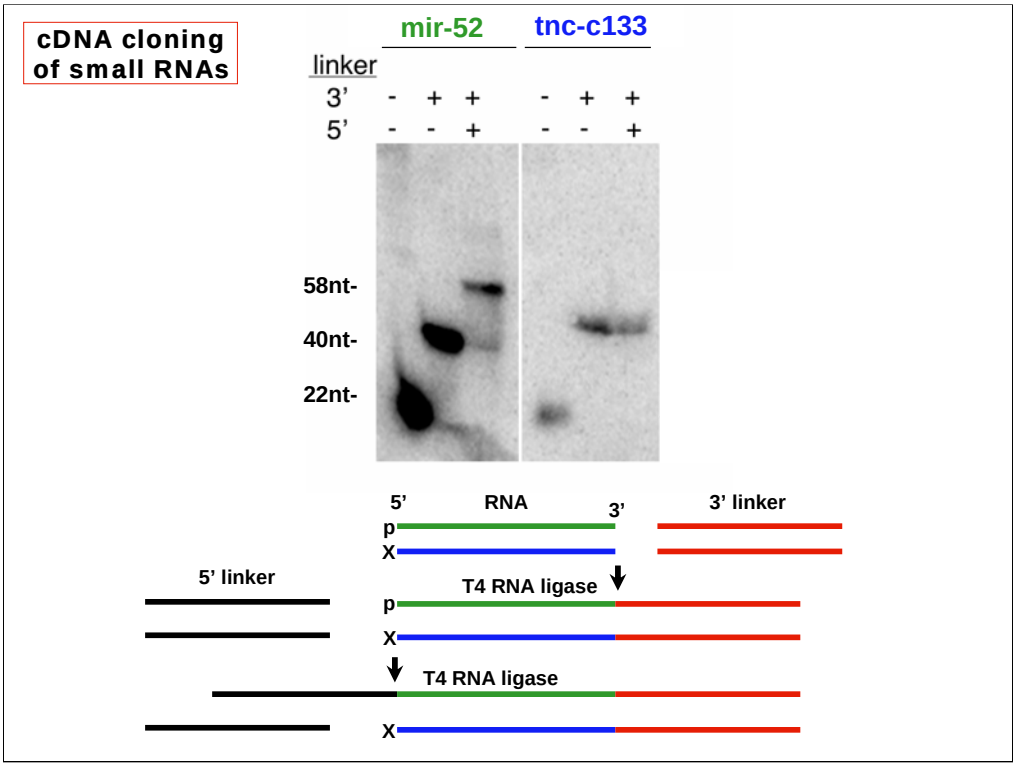


To understand how these other small RNAs are cloned, it should be noted that they require cDNA cloning methods that differ slightly from the methods tailored for cloning miRNA cDNAs.
In miRNA cloning, a linker oligonucleotide is affixed to the 3’ end of the RNA sample using T44 RNA ligase, and then a second linker is ligated to the 5’ end of the RNA sample (9, 11). Here is a northern blot that shows the molecular weight shift of a miRNA associated with each of these ligation reactions.

However…..



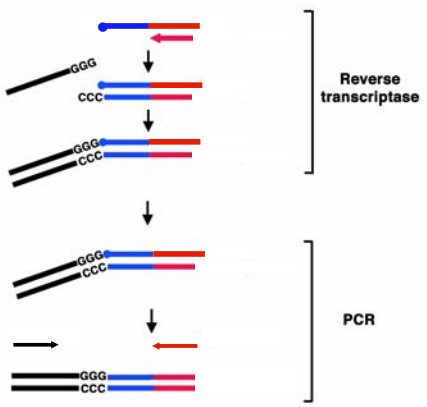
….certain ~22 nt RNA sequences are not amenable to this method (10, 12), as their 5' ends seem to be different from miRNAs, which have a 5' monophosphate. Note that…….



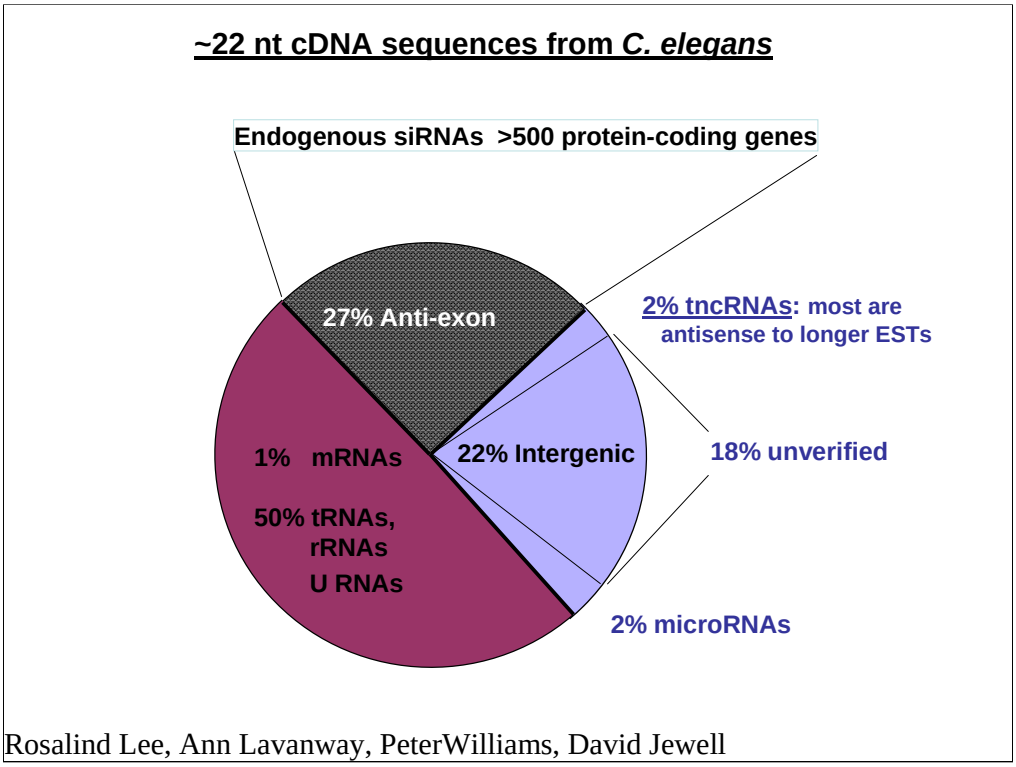
……..the tnc-c133 RNA is resistant to the ligation of the 5' linker.

In order to clone cDNAs corresponding to the tncRNA class of ~22nt RNAs……

cDNA cloning of small RNAs



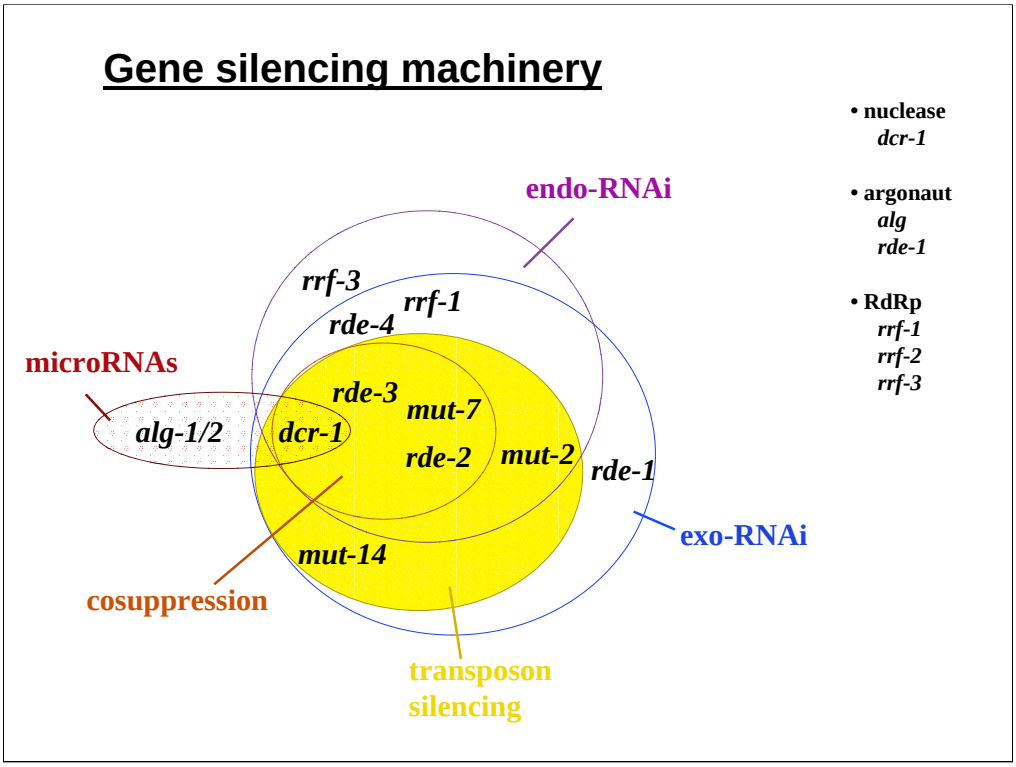
…we use a method that bypasses the 5' ligation step, and instead uses the Invitrogen "SMART" method for cloning the 5' sequences of RNAs (10, 12). In short, this method takes advantage of the intrinsic terminal transferase activity of RNAse H-defective Reverse Transcriptase, and an appropriate linker oligo for template switching. The result is a single-step reverse transcription reaction that produces a cDNA with knoen linkers at both ends.



In general, this method of cDNA cloning results in the identification of microRNAs, and two other classes of ~22nt RNAs (12):

1) Sequences antisense to protein-coding exons of diverse worm genes. These apparent endogenous siRNAs suggest that endogenous worm genes are actively silenced by RNAi in normal worms.

2) Tiny noncoding(tnc) RNAs that arise from transcription of intergenic sequences that do not apparently encode protein.



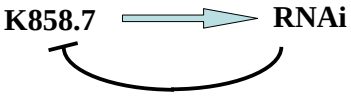
First, consider the endogenous siRNAs: In unpublished results, the Ambros lab finds that at least some of these siRNAs depend on some of the same RNAi machinery as exogenous RNAi, with some exceptions. Of particular note is the requirement for the RdRp homolog rrrf-3 for endogenous RNAi, whereas rrf-3 is not required for exogenous RNAi. In fact, the hypersensitivity of rrf-3 mutants to exogenous RNAi may be explained by the inactivation of endogenous RNAi and a consequent freeing-up of limiting shared components.

Regulatory consequences of endogenous gene silencing

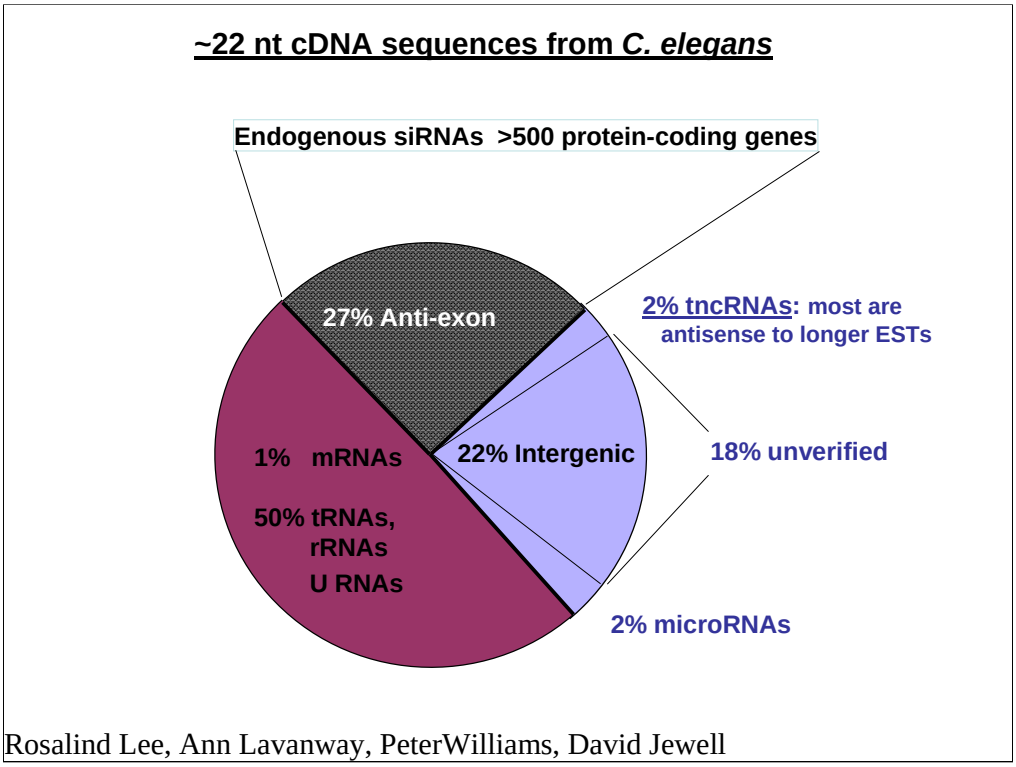
- **Contributes to setting the level of gene expression**
- **Competition with exogenous RNAi for (limiting) common factors**
- **At least three genes that we found to be silenced by endogenous RNAi….**

  K858.7 (oocyte enriched)
  D2096.8 (germline)
  C28A5.1 (germline)

…..are also *part of RNAi machinery* (Vasteneau et al...Plasterk, 2003)



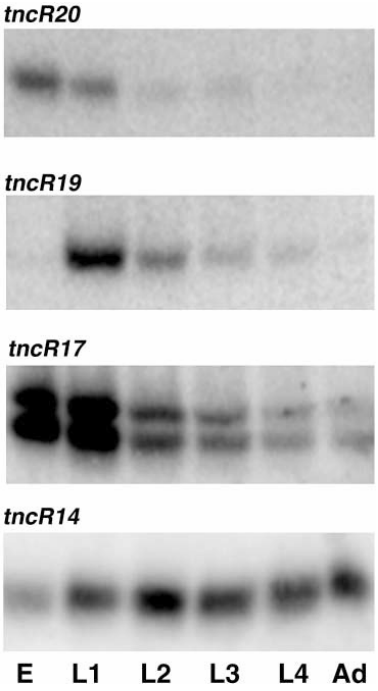
The diverse set of genes represented among the cloned siRNAs, suggests that the process of endogenous RNAi could significantly affect the expression levels of many important genes. Of particular note are certain genes that were among those that we identified as undergoing endogenous RNAi, but that also were found by the Plasterc group (49) to actually function as components of the RNAi machinery. This suggest as possible significant role for feedback regulation of RNAi pathways.



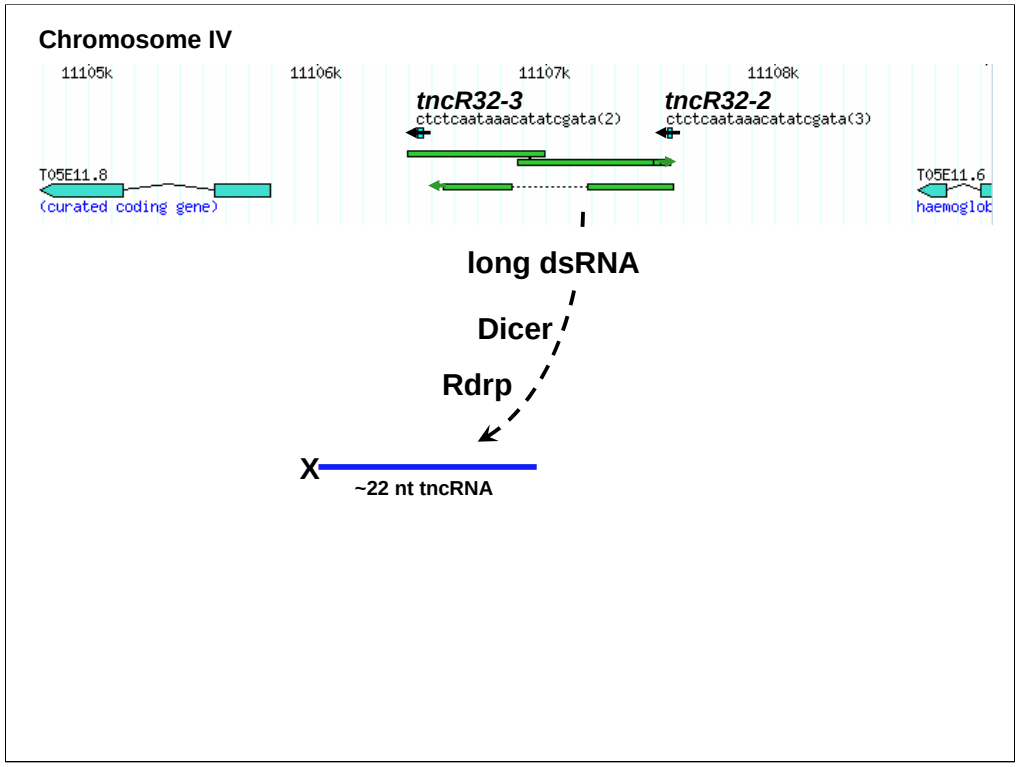
The tiny noncoding RNAs have characteristics of both microRNAs and endogenous siRNAs…..

First, like microRNAs, tncRNAs are transcribed from regions of the genome that doe not encode proteins.

Tiny noncoding RNAs (tncRNAs)



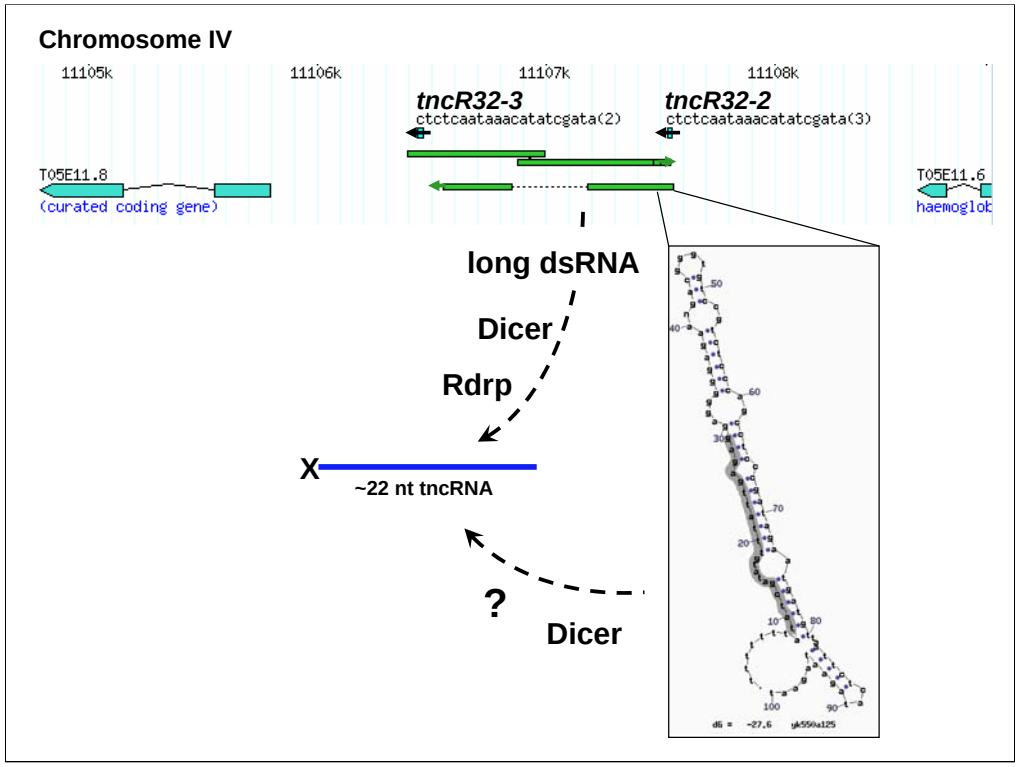
Second, like miRNAs, some tncRNAs are developmentally regulated.



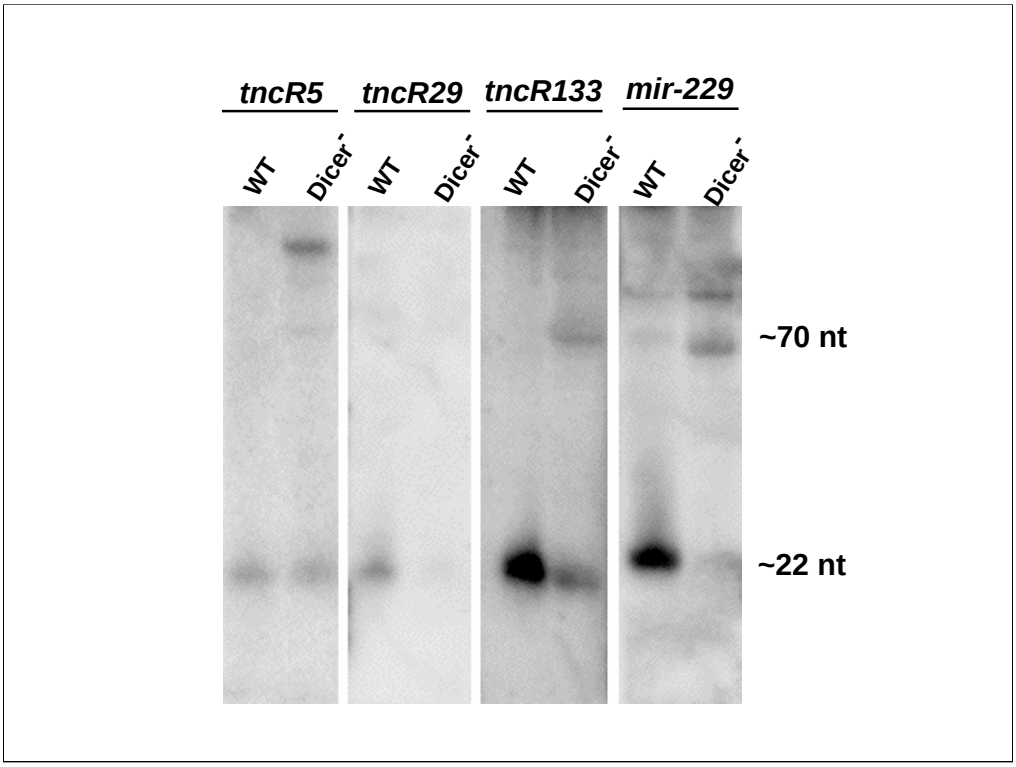
But unlike miRNAs, tncRNAs are not conserved phylogenetically

And, unlike miRNAs (which are produced by processing of a short hairpin transcript), some tncRNAs clearly come from regions of the genome documented to be transcribed bi-directionally, and hence they are likely produced from the dicer-dependent processing of a long duplex RNA.

However, in some cases…….

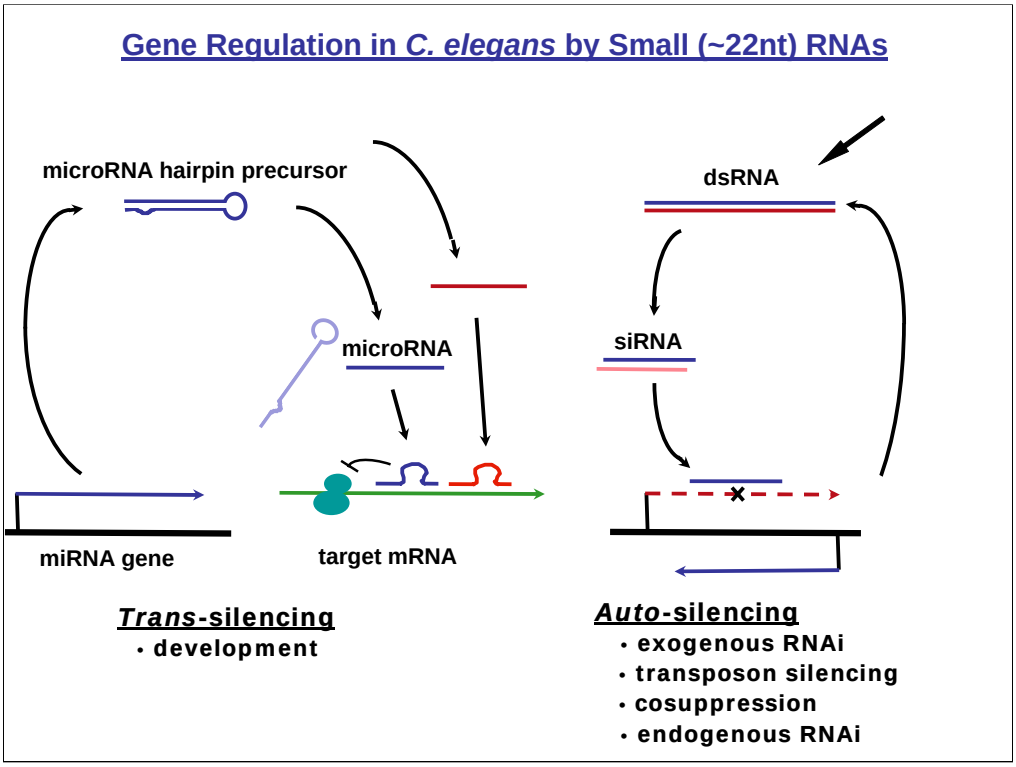


,……. the same region of the genome could be predicted to produce a hairpin RNA….



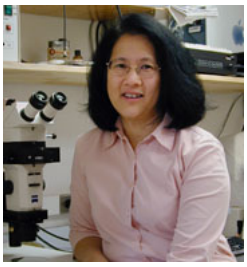
Indeed, some tncRNA sequences seem to be associated with longer, possibly hairpin, transcripts.

So tncRNAs could represent a type of rapidly evolving evolutionary precursors to new microRNAs.



In summary, the regulation of gene expression in C. elegans by ~22 nt RNAs include trans-silencing pathways that manage the specific control of certain mRNA targets on the level of translation (through microRNAs), and more general genome-wide pathways that set the level of mRNA from diverse genes though endogenous RNAi triggered by bidirectional transcription of protein-coding genes.

Small RNAs and endogenous gene silencing



Rosalind Lee

Chris Hammell

microRNAs and developmental timing

Allison Abbott

Allison Abbott	Ambros lab, Dartmouth
Eric Miska	Horvitz Lab, MIT
Ezequiel Alvarez-Saavedra	Horvitz Lab, MIT
Nelson Lau	Bartel lab, Whitehead

Peter Williams
Lorenzo Sempere
Nick Sokol
Tamara Zaytouni
Maria Ow

Ann Lavanway
David Jewell

Eric Moss
Phil Olsen

Much of the work on small RNA cloning and characterization in C. elegans is the work of Rosalind Lee (who, working with Rhonda Feinbaum, discovered the first microRNA, lin-4).

The genetic analysis of let-7 family microRNAs summarized is chiefly conducted by Allison Abbott, a postdoc in the Ambros lab, as part of a collaborative project by the Ambros, Bartel, and Horvitz labs to knock out all the work microRNA genes .

REFERENCES:

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004 Jan 23;116(2):281-97. Review.

2. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806-811.

3. Ambros, V. (2000). "Control of developmental timing in *Caenorhabditis elegans*." Current Opinion in Genetics and Development 10: 428-433.

4. Zeng, Y., Yi, R., and Cullen, B.R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc. Natl. Acad. Sci. USA 100, 9779-9784.

5. Doench, J.G., Peterson, C.P., and Sharp, P.A. (2003). siRNAs can function as miRNAs. Genes Dev. 17, 438-442.

6. Sulston, J., and Horvitz, H.R. (1977). "Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*." Developmental Biology 56: 110-156.

7. Ambros, V., and Horvitz, H.R. (1984). "Heterochronic mutants of the nematode *Caenorhabditis elegans*." Science 266: 409-416.

8. Olsen, P.H. and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev. Biol. 216, 671-680.

9. Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294, 858-862.

10. Lee, R.C. and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science 294, 862-864.

11. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science 294, 853-858.

12. Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr. Biol. 13, 807-818. [Summary] [Full Text]

13. Lai, E.C., Tomancak, P., Williams, R.W., and Rubin, G.M. (2003).