JCB Mini-Review

Micro-RNAs: small is plentiful

Helge Grosshans and Frank J. Slack

Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520

Two small temporally regulated RNAs (stRNAs)* of ~22 nucleotides regulate timing of gene expression during development of the nematode *C. elegans*. This regulation occurs at a posttranscriptional, presumably translational, level and is distinct from RNA interference (RNAi). One of the two stRNAs, *let-7*, as well as its target gene, *lin-41*, are highly conserved even in humans, suggesting a wide employment of stRNA-mediated gene regulation. Recent reports indicate that these two stRNAs are indeed likely to represent only the tip of an iceberg with hundreds or more of additional micro-RNAs (miRNAs) existing in metazoans. miRNAs might thus be previously underestimated key participants in the field of gene regulation.

Introduction

Small nontranslated RNAs perform a multitude of functions in the cell, e.g., transfer RNAs function as adapters in translation, small nuclear RNAs are involved in mRNA splicing, and small nucleolar RNAs direct the modification of ribosomal RNAs. Recently, two other groups of small RNAs have received much attention: small interfering RNAs (siRNAs) and stRNAs. Both types of RNAs are ~21–25 nt long and both mediate downregulation of gene expression. However, despite these and other similarities (for review see Banerjee and Slack, 2002), their respective modes of action appear to be quite different. siRNAs mediate RNAi, a process in which target mRNAs are degraded (Elbashir et al., 2001; for review see Sharp, 2001). stRNAs, in contrast, downregulate expression of target RNAs after translation initiation without affecting mRNA stability (Lee et al., 1993: Moss et al., 1997; Olsen and Ambros, 1999; Reinhart et al., 2000; Slack et al., 2000).

Although our understanding of the mechanisms of RNAi has evolved over the last few years, its physiological role remains to be explored. The opposite is true for stRNAs; although their mode of action is largely obscure, their role in the control of developmental timing in the nematode *C. elegans* has been well established. Lack of expression or overexpression of either of the two known stRNAs, *lin-4* and *let-7*, causes heterochronic phenotypes (for reviews see Banerjee and Slack, 2002; Rougvie, 2001), i.e., changes in the timing of developmental events.

A brief history of stRNAs

lin-4 and *let-7* RNAs both act as repressors of their respective target genes, *lin-14*, *lin-28*, and *lin-41* (Lee et al., 1993; Moss et al., 1997; Slack et al., 2000). Repression in all these cases requires the presence of stRNA complementary sequences in the 3′ untranslated regions (UTRs) of the target mRNAs, suggesting that mRNA/stRNA hybrids can form (Fig. 1, A and B; Reinhart et al., 2000; Slack et al., 2000; Wightman et al., 1991, 1993). The presence of these stRNA-binding elements is indeed sufficient to confer stRNA-dependent expression on unrelated reporter genes (Ha et al., 1996; Lee et al., 1993; Moss et al., 1997; Slack et al., 2000). How downregulation is achieved through the putative RNA/RNA hybrid is, however, unclear. Evidence with *lin-4* suggests that mRNA stability and polyadenylation level as well as translation initiation are not affected (Moss et al., 1997; Olsen and Ambros, 1999; Wightman et al., 1993).

In addition to sharing a similar size and, presumably, mode of action, *lin-4* and *let-7* are also both synthesized as longer precursors that, in silico, can be folded into stem-loop structures (Fig. 1 C). Processing into the mature forms involves the RNase III–like protein Dicer (*dcr-1* in *C. elegans*) as well as the AGO1/*rde-1* family members *agl-1* and *agl-2* (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). This is another notable parallel to RNAi, where siRNA biogenesis likewise requires Dicer and distinct members of the AGO1/*rde-1* family.

stRNAs constitute a subset of a larger group of miRNAs

Since functional homologues of *lin-4* were only found in related *Caenorhabiditis* species (Lee et al., 1993), an open question after the initial identification of *lin-4* almost a decade ago was whether stRNAs are unique to nematodes or more widely used. The latter idea proved correct upon the identification of a second *C. elegans* stRNA, *let-7* (Reinhart et al., 2000). Perfect homologues of *let-7* were found in various other organisms including the fly, zebrafish, and human and in some of these organisms, stage-specific expression of *let-7* was observed (Pasquinelli et al., 2000). In addition, putative orthologues of *let-7*'s major target, *lin-41*, were identified in vertebrates and invertebrates (Slack et

Address correspondence to F.J. Slack, Department of Molecular, Cellular, and Developmental Biology, Yale University, 266 Whitney Ave., New Haven, CT 06520-8103. Tel.: (203) 432-3492. Fax: (203) 432-6161. E-mail: frank.slack@yale.edu

*Abbreviations used in this paper: stRNA, small temporally regulated RNA; miRNA, micro-RNA; siRNA, small interfering RNA; RNAi, RNA interference.

The Journal of Cell Biology, Volume 156, Number 1, January 7, 2002 17–21
http://www.jcb.org/cgi/doi/10.1083/jcb.200111033

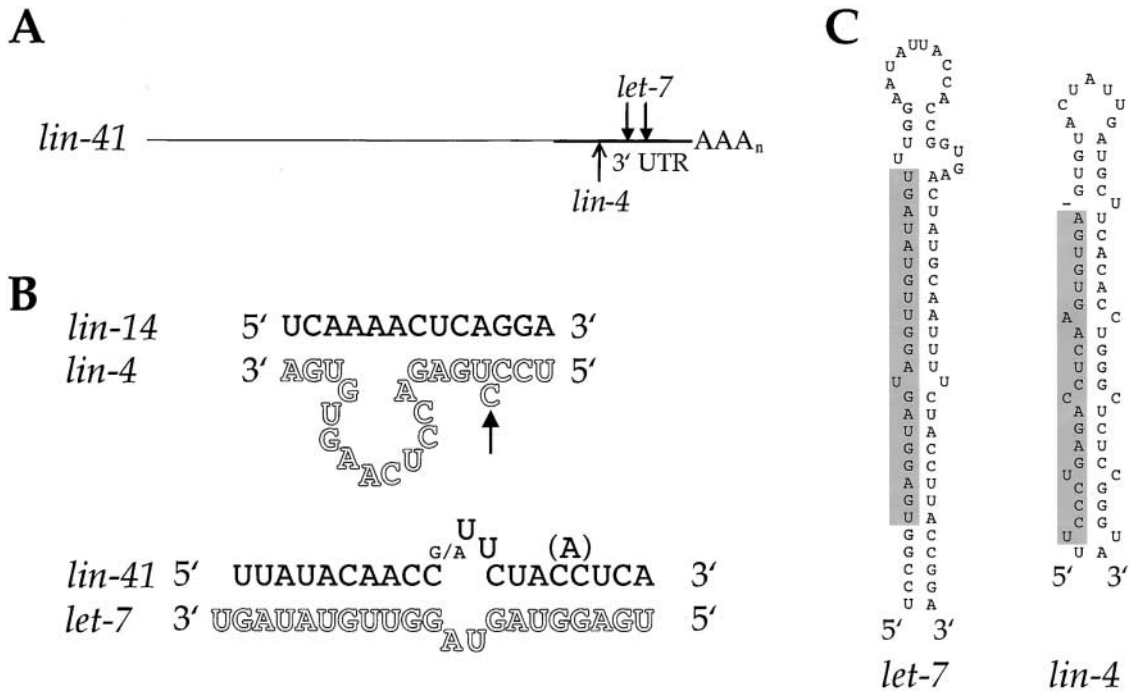

Figure 1. **Mature stRNAs form duplexes with the 3′ UTRs of their target mRNAs.** (A) *lin-41* is the major target of *let-7* and contains two slightly varying elements in its 3′ UTR that are complementary to the *let-7* stRNA as well as one element that is complementary to the second stRNA, *lin-4*. (B) Examples of potential RNA/RNA duplexes between *let-7* and *lin-41*, and *lin-4* and *lin-14*. In the case of *lin-41*, a consensus sequence for the two slightly varying 3′ UTR elements is shown. The arrow points to a conserved bulged C that is essential for *lin-4* function and might constitute a binding site for additional factors. The bulged A in *lin-41* is present in only one of the two elements and thus shown in brackets. (C) stRNAs and miRNAs are transcribed as precursors of ~60–70 nt that can be folded into a stem-loop structure. Processing releases a mature RNA of ~22 nt (shaded). For *lin-4* and *let-7*, which are shown here, the mature RNA is released from the 5′ arm of the precursor; however, miRNAs have been identified in which the 3′ arm (or even both arms) is stable.

al., 2000). Together, these data suggested that a posttranscriptional regulatory mechanism of gene expression might operate in a wide variety of other organisms besides *C. elegans*. This view has now received strong support through the identification of some 90 small, ~22 nt long RNAs from *C. elegans*, *Drosophila*, and human cell cultures (Table I), nine of which are found in more than one phylum (Table II; Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). To identify novel small RNAs, all three groups used a biochemical approach based on purification of RNAs after size selection (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001) or a bioinformatic approach (Lee and Ambros, 2001) centering on the conservation of intergenic regions of DNA between two related *Caenorhabditis* species. The expression of 2/3 of the thus identified RNAs was verified by Northern blot experiments (Table I).

Several lines of evidence suggest that these RNAs are indeed miRNAs and not degradation products of longer RNAs

Table I. **Summary of newly identified miRNAs**

	novel miRNAs		
	C. elegans	***Drosophila***	***Homo sapiens***
Lagos-Quintana et al., 2001	1	14	19
Lau et al., 2001	52	6	4
Lee and Ambros, 2001	15	3	2
Total[1]	57	20	24
Overlap[2]	11	4	2
Northern analysis (positive)[3]	34 (29)	15 (14)	19 (17)
5′ arm derived[4] (%)	28	36	58

99 unique sequences from the three organisms account for 91 novel miRNAs after correction for putative orthologues. Not included are *lin-4* and *let-7*, which were also identified.
[1]Number of unique sequences.
[2]Sequences identified by more than one group.
[3]Expression of the indicated number of miRNAs was investigated by Northern analysis. Numbers in brackets give the numbers of miRNAs for which a transcript of ~22 nt was detected in this analysis. 3 additional transcripts were verified as precursors only.
[4]Mature RNA derived from 5′ arm of precursor.

Table II. **miRNAs conserved across phyla**

Name	Sequence[1]	Homologues[2]
miR-1	UGGAAUGUAAAGAAGUAUGGAG	C, D, H
miR-2	UAUCACAGCCAGCUUUGA(G/U)G(U/A)GC[3]	C, D
miR-7	UGGAAGACUAGUGAUUUUGUUGU	D, H
miR-34	AGGCAGUGUGGUUAGCUGGUUG	C, D, H
miR-60	UAUUAUGCACAUUUUCUAGUUCA	C, D, H
miR-79	AUAAAGCUAGGUUACCAAAGCU	C, D
miR-84	UGAGGUAGUAUGUAAUAUUGUA	C, D, H
miR-87	GUGAGCAAAGUUUCAGG(U/A)GU[3]	C, D, H

[1]RNA sequences are deduced from cDNA sequencing; no RNAs have yet been sequenced.
[2]C, *C. elegans*; D, *D. melanogaster*, H, *H. sapiens*.
[3]Letters in parentheses indicate variations in otherwise identical miRNAs from different organisms or variant genes within one organism.

or siRNAs. First, all miRNAs are encoded in regions of the genome for which no genes had been predicted so far, although the biochemical protocols would be expected to yield a rather unbiased sample of RNAs of the appropriate size. Second, when genomic sequences surrounding the identified 22-nt RNAs were examined, computer analysis predicted miRNA precursors capable of forming stem-loop structures similar to those of the *lin-4* and *let-7* precursors (Fig. 1 C). Northern blot analysis also revealed precursors of appropriate sizes in some cases. Finally, in two cases where this was examined, an accumulation of these precursors could be detected in a *dcr-1* mutant, similar to the previously reported accumulation of pre–*lin-4* and pre–*let-7* stRNAs (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). In their entirety, these data thus strongly suggest that *lin-4* and *let-7* are just two of a vast group of miRNAs.

Lau et al. (2001) used a biochemical procedure that was tailored to identify Dicer products, i.e., siRNAs as well as stRNAs. Given further that size was the critical determinant used for selective RNA cloning by the other two groups (Lagos-Quintana et al., 2001; Lee and Ambros, 2001), it comes as a surprise that not a single siRNA was recovered. miRNAs rather than siRNAs might, under unstressed conditions, thus be the prevalent agent of small RNA-mediated regulation of gene expression. The small RNA processing machinery centered on Dicer might then, likewise, principally serve miRNA biogenesis. Lending credence to this claim is the finding that *dcr-1* mutants display heterochronic developmental defects and fertility defects (Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001).

miRNAs: conserved in size and structure, conserved in function?

Computer analysis suggested that some of the miRNAs were highly conserved in other organisms. In agreement with this idea, one miRNA was identified by all three groups independently, revealing its expression in *Drosophila*, *C. elegans*, and humans (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001), and >2/3 of the miRNAs identified from human Hela cells were found to be expressed in at least one other vertebrate species (Lagos-Quintana et al., 2001). In addition, most of the miRNAs identified in *C. elegans* were also found in a related nematode species, *C. briggsae* and, occasionally, in other, more distantly related species including man (Lau et al., 2001; Lee and Ambros, 2001). At least some miRNAs thus appear to be components of conserved pathways.

Are these pathways involved in developmental timing? They might be; after all, there is both precedence and room for function of novel miRNAs as stRNAs. Both *lin-4* and *let-7* are essential components of the *C. elegans* heterochronic pathway, with *lin-4* controlling transition from the first (designated L1) to the second (L2) larval stage (Chalfie et al., 1981; Lee et al., 1993) and *let-7* having a similar role in the L4 to young adult transition (Reinhart et al., 2000). No such switches have yet been identified for the L2-to-L3 and L3-to-L4 transitions, respectively; by analogy, leaving some attractive space for more stRNA function. Moreover, although the heterochronic pathway is well understood as far as larval development is concerned, little is known about embryonic development. Intriguingly, expression data suggest that many of the miRNAs are highly expressed in embryos with levels dropping after hatching (Lau et al., 2001).

However, the sheer number of miRNAs identified make it unlikely that they function exclusively in developmental timing. Also, although Northern blot analysis revealed some differences in miRNA expression levels when comparing embryos to adult animals, only a few miRNAs yielded signs of distinct up- or downregulation upon transition from one developmental stage to the next (Lagos-Quintana et al., 2001; Lau et al., 2001), a hallmark of *lin-4* and *let-7* (Feinbaum and Ambros, 1999; Reinhart et al., 2000).

If not developmental timing, what then? As noted earlier, the similarities between the control of spatial and temporal development are striking (Slack and Ruvkun, 1997; see also Banerjee and Slack, 2002). In a particularly revealing example, Hunchback and Krüppel, known for their roles in spatial patterning, were also found to define temporal identities of neuroblasts in the *Drosophila* central nervous system (Isshiki et al., 2001). It is thus conceivable that miRNAs might regulate genes involved in spatial organization of development and fertility. Although it may at first seem unlikely that key molecules in the well-understood pathway of spatial patterning may have been missed, one should consider the small size of the stRNAs that makes them poor targets in forward genetic screens.

Given the large number of miRNAs—and we predict that the screens have not yet been saturated—even more specialized functions are conceivable. Clearly, an important next step will be the identification of targets of all these small RNAs through genetical and bioinformatics approaches.

The role of the precursor in miRNA function

The abundance of stem-loop structures in pre-miRNAs, as well as the significant conservation of these structures in multiple organisms where homologous miRNAs were found (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001), suggests that these structures might be a prerequisite for miRNA maturation and possibly for function. One attractive hypothesis is that transcription as a precursor might allow another level of regulation of miRNA activity. Indeed, in one case Northern blot analysis suggested that the pre-miRNA might be continuously transcribed throughout development, but that maturation into the 22-nt form occurs only at distinct stages of development (Lau et al., 2001). This seems also to be true of *let-7* in sea urchin development (Pasquinelli et al., 2000).

The stem-loop structure of the pre-miRNA in combination with its size (~70 nt) also brings to mind another group of small RNAs, tRNAs. It is tempting to speculate that the structured pre-miRNA may actually be the substrate for nuclear export in which the dedicated tRNA nuclear export factor, Xpo-t (for review see Grosshans et al., 2000), might thus be involved through specific recognition of the pre-miRNA. Processing of the pre-miRNA through Dicer might then occur in the cytoplasm with subsequent targeting of the mature miRNA to its site of function.

Given what we know about the intricate coupling of maturation steps in other RNA biogenesis pathways, e.g., those that involve mRNA or tRNA (Wolin and Matera, 1999; Grosshans et al., 2000; Reed and Magni, 2001), we specu-

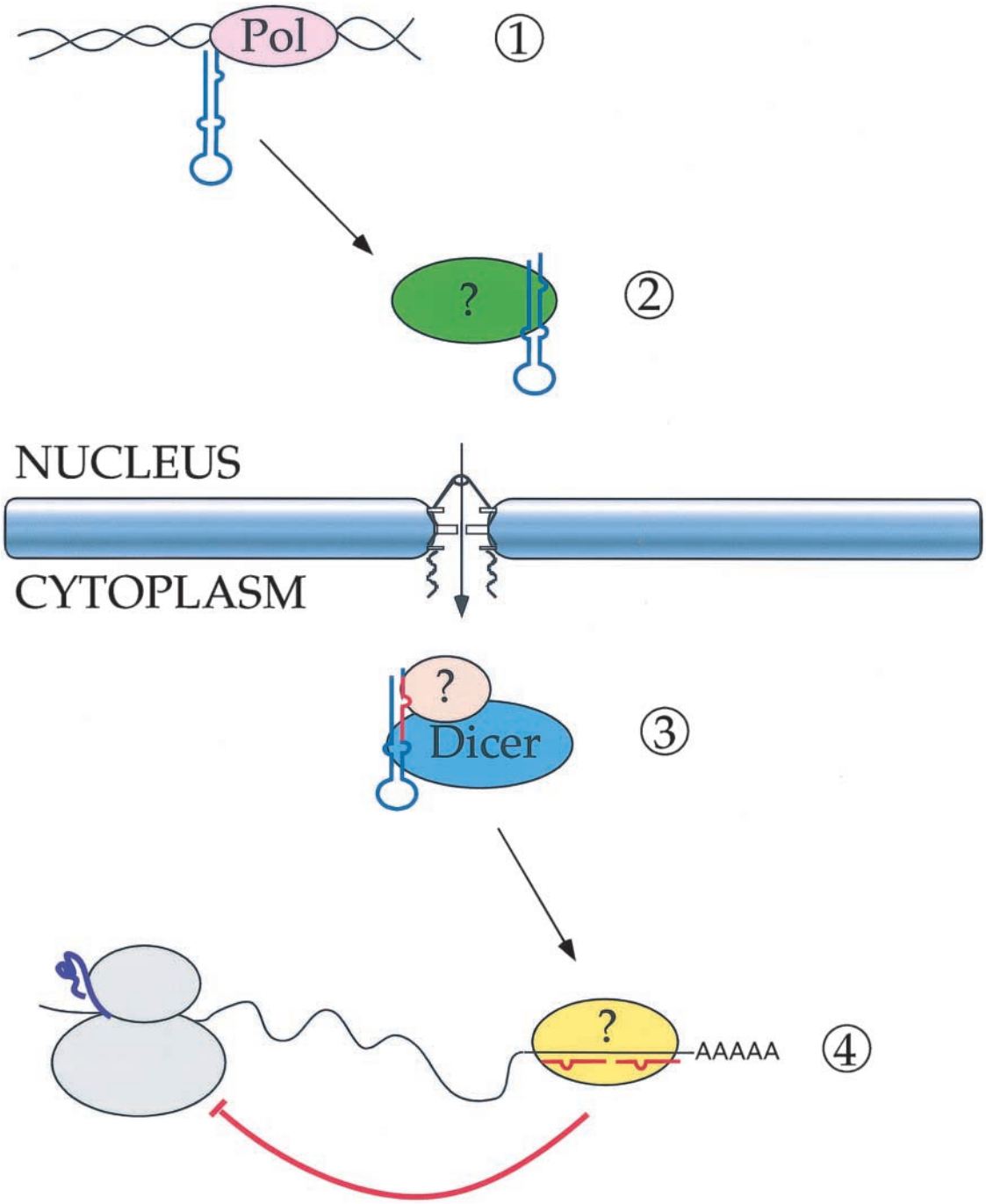

Figure 2. **Putative miRNA life cycle.** The model ilustrates important stages in the life cycle of a miRNA, all of which might be regulated. (1) miRNAs are transcribed as precursors by an, as yet, unidentified RNA polymerase (*Pol*). (2) Following its export to the cytoplasm by a likewise unidentified factor the precursor is processed (3) into the mature miRNA. (However, it should be noted that the site of pre-miRNA processing is currently unknown so that it is also possible that processing occurs in the nucleus and the mature miRNA is the substrate for nuclear export.) Pre-miRNA processing involves Dicer and, presumably, one or more cofactors such as AGL-1/AGL-2. These factors might then also be involved in delivery of the miRNA to its site of action, i.e., its target mRNA. (4) Binding of the miRNA to the mRNA 3′ UTR inhibits translation presumably at a point after translation initiation. (However, the nascent polypeptide chain (purple) exiting the ribosome in this model is solely for illustrative purposes, the presence of proteins stalled in elongation has not been demonstrated.) Downregulation of translation is thought to involve additional cofactors that presumably recognize the bulged-out nucleotides of the mRNA/stRNA duplexes.

late that biogenesis of miRNAs might indeed occur in a highly coordinated "channeling" pathway (Fig. 2). Nascent or freshly transcribed pre-miRNA might be recognized by the export machinery and, following nuclear export, be handed over to the processing machinery in the cytoplasm. Proteins that participate in the processing might also be involved in the delivery of the mature miRNAs to their sites of action so that miRNAs are handed over from one biogenesis factor to the next without intermittent release. Although no direct evidence is currently available to support this model,

it would well explain why miRNAs appear to be surprisingly stable. It would also allow for several levels of regulation—a feature one would anticipate for a regulatory pathway. The future identification of additional components of this pathway, as well as the subcellular localization of the known ones (such as Dicer), will help to establish the validity of such a model.

One mystery of miRNA biogenesis has been further complicated by the new data: how does Dicer recognize which side of the precursor molecule it has to excise? Although production of siRNAs by Dicer results in symmetric duplex structures, only one stem of the stRNA precursors remains stable (Fig. 1 C). The finding that with both *lin-4* and *let-7* the 5′ arm of the duplex survives suggested that the loop 3′ of the mature stRNA might be part of the solution. However, the majority of the newly identified mature miRNAs are derived from the 3′ arm of the stem-loop (Table I), and in one case mature miRNAs have even be identified from both arms. It is likely that accessory factors tell the Dicer machinery which arm of the stem-loop is to be preserved. Candidates for this job are the numerous members of the *rde-1* family. Mutations in two of these genes, *agl-1* and *agl-2*, have been shown to cause an accumulation of pre–*lin-4* similar to the phenotypes observed for mutants lacking Dicer, whereas pre-*let-7* levels remained unaffected (Grishok et al., 2001). It would therefore be highly interesting to see whether other pre-miRNA, in particular those where the mature miRNA is encoded on the 3′ arm, would accumulate in these mutants or in mutants of other members of the AGO1/*rde-1* family.

Future perspectives

The large number of miRNAs suggests that these RNAs might constitute an abundant component of the gene regulatory machinery. Although the first two such RNAs were discovered through their role in developmental timing, the sheer abundance of novel miRNAs as well as, in many cases, their apparent lack of developmental regulation suggest that miRNAs might be involved in other pathways as well. Clearly, we are only at the beginning of our exploration of miRNA versatility and function and many question marks remain regarding both miRNA biogenesis and function (Fig. 2). What are the functions and targets of all these miRNAs? Do all miRNAs function through a similar mechanism, and if so, what exactly is this mechanism? What are their cofactors? Which polymerase transcribes the miRNAs and how is their transcription regulated? Are there additional levels of regulation, e.g., processing or nuclear export? The abundance and conservation of miRNAs in animal genomes suggests that we have just begun to glimpse an important and so far underappreciated mechanism of gene regulation.

Submitted: 9 November 2001
Revised: 21 November 2001
Accepted: 26 November 2001

References

Banerjee, D., and F. Slack. 2002. Control of developmental timing by small temporal RNAs: a paradigm for RNA-mediated regulation of gene expression. *Bioessays.* In press.

Chalfie, M., H.R. Horvitz, and J.E. Sulston. 1981. Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell.* 24:59–69.

Elbashir, S.M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature.* 411:494–498.

Feinbaum, R., and V. Ambros. 1999. The timing of *lin-4* RNA accumulation controls the timing of postembryonic developmental events in *Caenorhabditis elegans*. *Dev. Biol.* 210:87–95.

Grishok, A., A.E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D.L. Baillie, A. Fire, G. Ruvkun, and C.C. Mello. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell.* 106:23–34.

Grosshans, H., G. Simos, and E. Hurt. 2000. Review: transport of tRNA out of the nucleus—direct channeling to the ribosome? *J. Struct. Biol.* 129:288–294.

Ha, I., B. Wightman, and G. Ruvkun. 1996. A bulged *lin-4*/*lin-14* RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. *Genes Dev.* 10:3041–3050.

Hutvagner, G., J. McLachlan, A.E. Pasquinelli, E. Balint, T. Tuschl, and P.D. Zamore. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science.* 293:834–838.

Isshiki, T., B. Pearson, S. Holbrook, and C.Q. Doe. 2001. *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell.* 106:511–521.

Ketting, R.F., S.E. Fischer, E. Bernstein, T. Sijen, G.J. Hannon, and R.H. Plasterk. 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15:2654–2659.

Knight, S.W., and B.L. Bass. 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science.* 293:2269–2271.

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

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

Lee, R.C., and V. Ambros. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science.* 294:862–864.

Lee, R.C., R.L. Feinbaum, and V. Ambros. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell.* 75:843–854.

Moss, E.G., R.C. Lee, and V. Ambros. 1997. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the *lin-4* RNA. *Cell.* 88:637–646.

Olsen, P.H., and V. Ambros. 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.

Pasquinelli, A.E., B.J. Reinhart, F. Slack, M.Q. Martindale, M.I. Kuroda, B. Maller, D.C. Hayward, E.E. Ball, B. Degnan, P. Muller, et al. 2000. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature.* 408:86–89.

Reed, R., and K. Magni. 2001. A new view of mRNA export: separating the wheat from the chaff. *Nat. Cell Biol.* 3:E201–E204.

Reinhart, B.J., F.J. Slack, M. Basson, A.E. Pasquinelli, J.C. Bettinger, A.E. Rougvie, H.R. Horvitz, and G. Ruvkun. 2000. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans. Nature.* 403:901–906.

Rougvie, A.E. 2001. Control of developmental timing in animals. *Nat. Rev. Genet.* 2:690–701.

Sharp, P.A. 2001. RNA interference—2001. *Genes Dev.* 15:485–490.

Slack, F., and G. Ruvkun. 1997. Temporal pattern formation by heterochronic genes. *Annu. Rev. Genet.* 31:611–634.

Slack, F.J., M. Basson, Z. Liu, V. Ambros, H.R. Horvitz, and G. Ruvkun. 2000. The *lin-41* RBCC gene acts in the C. elegans heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor. *Mol. Cell.* 5:659–669.

Wightman, B., T.R. Burglin, J. Gatto, P. Arasu, and G. Ruvkun. 1991. Negative regulatory sequences in the *lin-14* 3′-untranslated region are necessary to generate a temporal switch during *Caenorhabditis elegans* development. *Genes Dev.* 5:1813–1824.

Wightman, B., I. Ha, and G. Ruvkun. 1993. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin- 4* mediates temporal pattern formation in C. elegans. *Cell.* 75:855–862.

Wolin, S.L., and A.G. Matera. 1999. The trials and travels of tRNA. *Genes Dev.* 13:1–10.