The perfect storm of tiny RNAs

Gary Ruvkun

The discovery of a previously unsuspected civilization of tiny RNAs within nearly all eukaryotes was a perfect storm that began to blow in 1992 with the discovery of the first 21-nucleotide microRNA and its role in translational control and built to a gale in 1999–2001 with the discovery of the equivalently sized small interfering RNAs (siRNAs) in RNA interference and the discovery of microRNAs conserved across phylogeny. It is now a 20-year storm and continues to intensify.

The elements of the storm converged from many points of the compass. There were satisfyingly off-center experimental systems and approaches, ranging from cell lineage control and gene silencing in worms to viral resistance in plants to heterochromatin formation in fungi to gene elimination in protozoa. The storm depended on a renaissance of developmental genetics to reveal the regulatory components of metazoan developmental control. It depended on the production and free availability of the first animal and plant full-genome databases, which did not ignore the regions of the genome that do not fit into the canon of protein-coding genes. It depended on the Internet to empower database searches which revealed the presence of conserved microRNAs and their locations in the genome. And it depended on a conserved protein-mediated pathway for the production and presentation of these tiny RNAs, to allow the genetic and biochemical discoveries from distantly related organisms to apply to each other, thus connecting up communities of biologists previously islands unto themselves.

The path to the discovery of the first microRNA and its role in translational control

I entered the tiny RNA world via molecular genetic characterization of *Caenorhabditis elegans* developmental control. The genes *lin-4* and *lin-14* were named for the cell lineage (*lin*) changes that mutations in these genes cause. The cell lineage of *C. elegans*—the patterns of division that generate the 959 cells of an adult worm—reveal it to be an animal with few enough cells to be denumerable and named, and a reproducible enough pattern of development to be described by a cell lineage diagram[1]. The collection of lineage mutations first identified by Horvitz, Sulston and Chalfie in the late 1970s held the promise of identifying a genetic control network for this cell lineage[2,3]. It was a combination of the use of the cell lineage in the analysis of developmental mutants, which afforded such an uncommonly digital view of development; the mutant collection; and a sense that the *C. elegans* field was poised to discover the molecular identity of dozens of control genes that brought me to the field in 1982 as a postdoc in Bob Horvitz's lab.

Victor Ambros, also a postdoc in the Horvitz lab, was just finishing his genetic analysis of heterochronic genes, mutations in which cause cell lineage changes that define a pathway for patterning the cell lineage in the temporal dimension[4–6]. The most compelling of the genes was *lin-14* because it had both gain-of-function and loss-of-function mutations with opposite cell lineage defects, and because Victor had already placed *lin-14* into a regulatory pathway with *lin-4*. Victor was very keen to learn the molecular identity of *lin-14* and offered to work together on its molecular analysis. I was elated because it was clear that I could learn a lot from Victor and that he would be fun to work with.

Developmental control genes with opposite gain-of-function and loss-of-function phenotypes were considered the keys to development in the lab at that time. The virtues of genes with such attributes were the mantra in Horvitz group meetings. More generally, gain-of-function mutations tend to reveal points of negative regulation in pathways; the activating mutations abrogate these negative feedbacks. Feedback regulation may be found in virtually any pathway, from the Krebs cycle to RNA interference (RNAi), and is a key element of control in those pathways. So abrogation of negative regulation via gain-of-function mutations is a fruitful approach to disentangling biological networks. As it turned out, the analysis of gain-of-function mutations in *lin-14* was a key component in the development of the microRNA field.

The *C. elegans* field was long on mutants but lacking a track towards molecular analysis. We needed a method to isolate the piece of DNA corresponding to a locus defined by classical genetics. For the molecular identification of *lin-14* (as well as *lin-4* and a slew of other genes defined by mutation), we developed a variant of restriction fragment length polymorphism (RFLP) mapping[7]. We knew that there were highly dispersed RFLPs caused by hundreds of transposon insertions in one strain of *C. elegans* that were not present in the standard *C. elegans* lab strain where all of the previous genetics had been done. So we mapped the particular few elements among these hundreds that were the closest in the genome to *lin-14*, allowing us to jump with transposon DNA probes right to the region. Thus, in one genetic mapping experiment, a 24-factor cross, we were about 99% of the way to *lin-14*.

The first evidence that *lin-14* had been identified came from the detection of DNA changes associated with *lin-14* alleles using probes from near the closest RFLPs[7]. Bruce Wightman, Joe Gatto, Thomas Burglin and I determined that the *lin-14* gain-of-function mutations affected the *lin-14* 3′ untranslated region (UTR), suggesting that the negative regulation of *lin-14* by *lin-4* was at the level of regulation of mRNA abundance or translation. Prema Arasu, Bruce Wightman and I found that the expression of LIN-14 protein is graded over time and that graded expression is disrupted in the *lin-4* and in *lin-14* gain-of-function mutants. Because the *lin-14* gain-of-function mutations mapped to the 3′ UTR and caused molecular defects similar to those of the *lin-4* loss-of-function mutations, the simplest model was that *lin-4* would regulate the *lin-14* 3′ UTR[8–10]. But both the Ambros and Ruvkun labs were envisioning a LIN-4

Gary Ruvkun is in the Department of Molecular Biology, Massachusetts General Hospital, Department of Genetics, Harvard Medical School, Fruit Street, Boston, Massachusetts 02114, USA.
e-mail: ruvkun@molbio.mgh.harvard.edu

regulatory protein that might engage the *lin-14* 3′ UTR directly, if we were very lucky.

The discovery in the Ambros lab that *lin-4* encodes a microRNA changed everything[11]. Just after the Ambros group had convinced themselves that *lin-4* might actually encode an RNA, Victor and I analyzed the *lin-4* and *lin-14* 3′ UTR sequences and detected complementarity between the two RNAs. It was a moment when years of work came together into a clear model. There are multiple elements in the *lin-14* 3′ UTR that are complementary to the *lin-4* microRNA. The two *lin-14* gain-of-function mutations fit with the sites complementary to the *lin-4* microRNA: the weaker gain-of-function allele deletes five of seven complementary sites, whereas the stronger allele removes them all[11,12]. And Ilho Ha showed that the *lin-4* microRNA binds to the *lin-14* mRNA *in vitro* and that a *lin-14* 3′ UTR with mutations in each of the *lin-4* complementary sites is no longer downregulated by *lin-4* (ref. 13).

But the sites complementary to the *lin-4* microRNA were not perfectly complementary—we predicted multiple bulges and loops both in the *lin-4* RNA strand and in the *lin-14* mRNA strand, so that these would be imperfect RNA duplexes, as in the secondary structures of that Ur of RNA biology, the ribosomal RNAs. As in the case of the ribosomal RNAs, these proposed bulges and loops were endorsed by phylogenetic conservation. Bruce Wightman and Ilho Ha fused the *lin-14* 3′ UTR onto a reporter gene and showed that it is necessary and sufficient to generate graded temporal expression of the reporter gene both in *C. elegans* and in *C. briggsae*, a nematode species about as distant from *C. elegans* as mouse and human are from each other[12]. Consistent with this conservation of *lin-14* 3′ UTR function, multiple elements in the *lin-14* 3′ UTR are conserved between *C. elegans* and *C. briggsae*, as is the *lin-4* RNA sequence[11,12]. In fact, all of the sequences in the *lin-14* 3′ UTR that are complementary to the *lin-4* RNA are conserved. The detailed features of these conserved elements in the *lin-14* 3′ UTR supported the existence of distinct RNA duplex structures—for example, bulged 'C'—rather than more perfect duplexes, which Ilho Ha showed had distinct functions[13].

These RNA duplex models, 3′ UTR fusion gene assays, and phylogenetic conservation experiments strongly supported the model that the *lin-4* microRNA directly binds to the *lin-14* mRNA 3′ UTR to downregulate its expression, but they did not assign any mechanism to that regulation. Back to back with the Ambros lab *lin-4* microRNA paper, Bruce Wightman, Ilho Ha and I showed that the *lin-4* microRNA regulates its target mRNA *lin-14* via translational control[12]. By monitoring LIN-14 protein and mRNA levels in the wild type and in the *lin-4* and the *lin-14* gain-of-function mutants, we showed that the *lin-4* RNA regulates not *lin-14* mRNA abundance, but rather translation of the *lin-14* mRNA[12]. *lacZ–lin-14* 3′ UTR fusion genes were also useful in this analysis. By monitoring β-galactosidase activity, a measure of protein production allowed by the *lin-14* 3′ UTR, and *lacZ* mRNA expression from a *lacZ–lin-14* 3′ UTR fusion gene in a wild-type or *lin-4* mutant background, Bruce and Ilho also showed that *lin-4* acts post-transcriptionally via the *lin-14* 3′ UTR to generate graded temporal expression[12]. This also showed that the *lin-4* microRNA does not depend on other *lin-14* mRNA sequences (the 5′ end, for example), constraining the mechanism of the *lin-4*·*lin-14* RNA duplex. This model was validated and extended by Olsen and Ambros, who showed that the translational control of *lin-14* in fact occurs after initiation because the *lin-14* mRNA that is not translated at late developmental stages when *lin-4* represses its translation is localized, paradoxically, to polysomes[14].

The second microRNA gene and its conservation across animal phylogeny

The discovery of the first microRNA did not trigger a gold rush. First, the heterochronic pathway was a rather parochial object of study; without homologs in other species, its generality was not obvious. And among developmental biologists there was a tendency to marginalize *C. elegans*. The problem was the cell lineage: the provisional conclusion that there were few cell-cell inductions in the lineage, based on just a few laser ablation experiments[1], were counter to a century of induction research in developmental biology on organisms from plants to frogs to mammals. It did not instantly recommend the worm as a model for development in other clades—there was much harrumphing about this by the aristocracy of developmental biology at Gordon Conferences. In the community of developmental biology, there was a sense that *C. elegans* was different enough that this *lin-4/lin-14* story could be an oddity of an odd creature. However, among the RNA processing, translational control and modification community, already steeped in deep conservation of ribosomal and splicing regulatory RNAs, universals were expected, and the *lin-4/lin-14* story quickly became part of the canon. It was also noticed by the bacterial natural antisense RNA community, as an example of a natural eukaryotic antisense RNA. In the regulatory RNA community, universals were expected and this expectation trumped any details of the particular organism from whence the genes came, which that community tended to ignore in any event.

The first hint of a more extensive microRNA world emerged after Brenda Reinhart and Frank Slack in my lab, in collaboration with Michael Basson and Bob Horvitz, discovered a second microRNA, *let-7*, that base pairs, again with bulges and loops, to the 3′ UTR of its major target mRNA, *lin-41* (refs. 15,16). This was a second example of a microRNA and the second example of a microRNA having one major target mRNA that it regulates via partially complementary elements in its 3′ untranslated region.

A microRNA world that extended across animal phylogeny emerged in 1999 and 2000 from Amy Pasquinelli's, Brenda Reinhart's and my work exploring the conservation of microRNAs[17]. The much longer ribosomal RNAs and the splicing regulatory U_n RNAs and tRNAs were already known to be conserved at the primary nucleotide sequence level over most eukaryotes, and, in the case of ribosomal RNAs, all the way to bacteria. And scores of protein-based developmental regulatory genes were known to be conserved across the animal kingdom. So even though *lin-4* and *let-7* were much smaller than any other RNAs, perhaps DNA segments encoding the fly or human orthologs of *lin-4* or *let-7* could be found in the emerging genome sequences of *Drosophila melanogaster* and human in 1999. Within about 30 seconds of conceiving the question—the time it took to paste in the *let-7* sequence into a query form and perform a BLASTN query over the network—we saw the answer in the BLASTN output: we found perfect conservation of the *let-7* mature microRNA sequence in the *Drosophila* and human genome sequence databases[17]. In addition, when we retrieved the sequence flanking these potential *let-7* orthologs from flies and humans, they folded into stem-loop precursors similar to those of *C. elegans lin-4* and *let-7*. We did not find any *lin-4* homologs in these database searches, though later, as microRNAs were directly sequenced, homologs with more substitutions and deletions were detected; for example, *miR-125* from humans and *Drosophila* is a probable *lin-4* homolog with four differences, too many to find in simple BLASTN full-genome searches.

The *let-7* match of 22 nucleotides in a human or fly genome could occur by chance, though the stem-loop prediction increased the confidence that these were bona fide *let-7* orthologs[17]. But it was also important to show on northern blots that *Drosophila* and mammals actually produce 22-nucleotide transcripts of these informatically detected genes. This we observed. But because we could assay for *let-7* conservation in the expressed RNAs of unsequenced organisms via a simple northern blot of total RNA, we sampled, in collaboration with many people who sent us RNA, a phylogenetically broad set

of animals and plants to explore the conservation and the origin of the gene. We found that the *let-7* microRNA is conserved in bilaterian animal species, including humans, but is not present in, for example, cnidarians, ctenophores, fungi, plants or basal eukaryotes[17]. In addition, the temporal regulation of *let-7* is also conserved, as is complementarity in the 3′ UTR of the conserved target gene *lin-41*, strongly suggesting an ancient function in temporal patterning[17]. The conservation of the *let-7* RNA argued for the generality of microRNAs.

The perfect conservation of *let-7* also presented a real conundrum: if the microRNAs pair with target mRNAs, but imperfectly, with particular regions of both microRNA and mRNA not in the RNA duplex, why are some microRNAs conserved over almost a billion years of evolution? In the ribosome, for example, regions that form RNA duplexes covary in evolution and are not as well conserved as some unpaired regions that constitute the active site of this ribozyme. One view was that if microRNAs have many targets, it would constrain the base-pairing regions because sequence covariation with dozens of targets could not occur[17]. However, we knew from the worm genetic analysis that there are just a few major targets for the known function of *let-7* in developmental patterning[16]. One possible explanation comes from the current view from bioinformatic target prediction programs of scores of mRNA targets for each microRNA: even in the case of *let-7*, there may be more targets regulated by *let-7* than are assayed by a worm phenotypic analysis of molting and of patterning of the cell lineage.

But then why are the non-base-paired regions so well conserved? As in the ribosome, the bulges and loops may be landing pads for the proteins or other ligands that might recognize the microRNA·mRNA duplex to, in turn, downregulate translation. So the bulges and loops, like the most conserved regions of the ribosomal RNA, may correspond to the unduplexed 'active site' of the microRNA·mRNA duplex in translational control. In fact, even in the less well conserved *lin-4* microRNA, we can see that the first few nucleotides of the unpaired loop region of the *lin-4* or *miR-125* microRNAs are conserved, suggesting that they are recognized as single strands.

The conservation of *let-7* across animal phylogeny and the fact that it was the second microRNA found meant that there might well be more to discover. In a set of papers that expanded the number of microRNAs enormously, the Ambros, Tuschl and Bartel labs used biochemical approaches to discover dozens more microRNAs[18–20]. In collaboration with Yonatan Grad, John Aach and George Church, John Kim and Gabe Hayes in my lab also conducted genome-scale informatic searches for microRNA genes conserved like *let-7* across animal phylogeny[21].

Because of the power of deep sequencing, we now have extensive descriptions of the small RNAs that are produced in animals, fungi, plants and protozoa. The conservation of about 30% of the microRNAs within animals (and a different set of microRNAs across plants), and the conservation of a small fraction of these as profoundly as *let-7*, is still unexplained. The mapping of microRNAs to clearly important target mRNA functions has been most productive in plants, where the nearly exact matches to target mRNAs have enabled accurate target prediction. The assignment of microRNA functions to pathways via target mRNA predictions is only now beginning in animals because target prediction with allowable bulges and loops is exponentially more difficult. As impressive as the genomic exploration of small RNAs has been, we may still only be viewing the tip of the iceberg. We have most of the data for small RNAs with 5′ monophosphates, the likely products of Dicer. But the piRNAs, small RNAs that associate with the PIWI subtype of Argonaute proteins, are generated by a process that does not depend on Dicer[22]. The small RNAs generated by other nucleases might not have 5′ phosphates. There is going to be another gold rush for these new classes of tiny RNAs.

Parallel universes of small RNAs and the protein cofactors that process and present them

When siRNAs of the same size as *lin-4* and *let-7* were discovered in 1999 by Hamilton and Baulcombe in plants[23] and in 2000 by Tuschl and colleagues in animals[24], the fields of RNAi and microRNAs suddenly converged. We immediately started to explore the action of the RNAi machinery in microRNA maturation and action. It seemed likely that the similarly sized microRNAs and siRNAs would use similar mechanisms. Amy Pasquinelli looked closely at the first RNAi-deficient mutants, *rde-1* and *rde-4*, but could not detect any heterochronic defects nor any change in *lin-4* or *let-7* microRNA activity or processing. Alla Grishok and Craig Mello then discovered that inactivation of 2 of 28 RDE-1 paralogs in the *C. elagans* genome causes a phenotype similar to the *let-7* lethality. Grishok and Pasquinelli showed that RNAi inactivation of these RDE-1 paralogs or of Dicer in *C. elegans* disrupts *lin-4* and *let-7* microRNA processing and activity, proving that the RNAi and microRNA pathways are related[25]. Phil Zamore's group also showed this for *let-7* and Dicer in *Drosophila*[26].

The current view is that microRNAs imperfectly base-paired to mRNAs, with bulges and loops, regulate translation, whereas microRNAs or siRNAs that are perfectly base-paired regulate mRNA degradation. But perfectly base-paired microRNA·mRNA duplexes in *Arabidopsis* work both through translational control and through mRNA degradation mechanisms[27,28]. And others have shown that there are transcript abundance responses to microRNAs at times as well[29]. Perhaps RNAi and microRNAs are more similar than we currently appreciate.

Saturation genetic analysis of the microRNA pathway by Devin Parry[30] and of the RNAi pathway by John Kim, Harrison Gabel, Ravi Kamath, Duo Wang, Maurice Butler, Sylvia Fischer and Scott Kennedy in my lab[31–34] has now revealed many of the protein cofactors that may mediate other steps in the process by which microRNAs and siRNAs engage their targets. As in any complex process, there may be many steps downstream of Dicer and Argonautes that interpret these small RNAs, and there may be paralogous pathways for other small RNAs, mediated by paralogs of the microRNA and siRNA cofactors, just as an Argonaute paralog mediates piRNA function[22]. A number of the microRNA pathway protein cofactors are annotated to be RNA-binding proteins that, for example, could mediate steps in the recognition of the microRNA·mRNA RNA duplex to downregulate translation. There is also increasing evidence of a cell biology of RNA regulation, specifically in P bodies and P granules[29] and perhaps in other vesicular elements[35], so, a more complex choreography is likely. Also supporting a cellular trafficking of tiny RNAs, cytoskeletal elements have emerged as strong candidates for activity in the microRNA pathway in both animals and plants[27,30], and RNA associates with the cytoskeleton at mitosis[36].

In the same way that the PIWI and Argonaute proteins have been found to shuttle small RNAs to their sites of action in cell, the protein cofactors discovered in genetic and RNAi screens for small RNA function may bind to and present small RNAs. The bound small RNAs would convey sequence information for translational control of mRNAs or modifications of DNA, for example. The protein cofactors might tether these guide RNAs at points of biological regulation: the ribosome, a replication fork, a synapse, a telomere, a DNA break, a dividing centrosome, a duplicating Golgi body or other inherited and regulated subcellular organelle. And an exploration of the cell biology of the tiny RNA protein cofactors could reveal points of regulation by tiny RNAs. The tiny RNAs that may bind to the protein cofactors revealed by the genetics to date could both suggest functions of small

RNAs, via their associations with cytoskeletal or candidate protein cofactors, and reveal new nexuses of small RNA regulation through their sequences.

RNAi was discovered in plants and worms probably because gene silencing by double-stranded RNA is so much more intense in plants and worms due to the amplification of siRNAs by RNA-dependent RNA polymerases and the systemic spread of these small RNAs[29]. But why do these clades use RNAi so much more intensively than others? One view is that the onslaught of viruses is somehow a larger problem for these soil-dwelling taxa. Another view is that small RNAs are spread systemically in plants and worms as a form of hormonal signaling, concerning either viral status or other informational onslaughts that are more endemic in plants and worms. The systemic spread of siRNAs in plants and worms does suggest the possibility that small RNAs constitute signaling molecules. There is evidence for microRNAs in the phloem of plants as well as in secretory vesicles of animal cells[35,37]. microRNAs as well as other tiny RNA signaling packets may constitute an intercellular signaling pathway that may reveal itself as our analyses become more sophisticated.

MicroRNAs and stem cells

One distinction of *C. elegans* and its close relatives in the animal kingdom is that there are only a dozen or so cell divisions from a totipotent fertilized egg to a fully formed adult. The genes affected by the heterochronic mutations that first revealed the tiny RNA world in worms mediate determinations of multipotent versus committed cell fates in the reprogramming of a cell from multipotency. The cell lineage analysis that was the great advance in early *C. elegans* genetic analysis was the lens through which these cell fate commitments were viewed. Perhaps it is the few divisions from these cell commitments to the fully formed larvae and later stages that allowed such major transformations in cell fate to be viable enough to survive and be classified by this lineage analysis. Similarly, the major organs of a plant derive from totipotent shoot apical meristem tissues and are just a few cell divisions away from this totipotency. As in *C. elegans*, the mutant phenotypes of plants defective in small RNA pathways of cell commitment could be interpreted because there are so few divisions from meristem to functional flower or leaf. In fact, the Argonaute and Dicer pathways emerged first from plant floral and leaf patterning genetics[38].

Other recent discoveries strongly favor the view that microRNAs act in the earliest divisions to specify the transition from totipotency to differentiation. Mammalian orthologs of the heterochronic gene *lin-28*, discovered and studied by Victor Ambros and Eric Moss[39], negatively regulate *let-7* maturation by binding to its stem-loop precursor during the transition from stem cell to differentiated cell[40,41]. Moreover, *lin-28* emerged from gene array studies of embryonic stem cells as one of the major transcripts downregulated upon stem cell differentiation[42]. And co-transfection of *lin-28* with just a few other transcription factors into committed cells is sufficient to transform those cells towards totipotence[43]. Finally, many papers point toward microRNA negative regulation of oncogenes and cell division, and there is an emerging view that most microRNAs are not expressed in totipotent cells but become more highly expressed as the cells commit to differentiated lineages. These studies strongly support the model that genes central to microRNA function are key in the transition from totipotency to commitment and cell differentiation.

Biodiversity and gene discovery in the genome era

So yes, *C. elegans* was among the stranger beasts to study. But not only have its technical advantages empowered many a research career, its RNA savantism has certainly launched a number of us beyond where we might have gone studying other organisms and other pathways. More generally, the view that particular taxa may have amplified or deleted functions present in others, and that these distinctions may be important in the evolution of these clades, is going to become more and more important as the genome era reveals the detailed specializations of taxa. For example, *C. elegans*, so proficient at RNAi, has a highly ramified set of Piwi Paz proteins (orthologs and paralogs of the Argonaute and PIWI subtypes) compared to other plants and animals; one might have been able to predict from the genome sequence some unique, evolutionarily important function for this amplified orthology group.

It is important for working biologists and for our patrons at the US National Institutes of Health and National Science Foundation and the equivalent bodies in other countries to realize that much of the tiny RNA revolution emerged from non-vertebrate genetic analysis—from plant biology, from worm genetics, from the fungi *Neurospora crassa* and *Schizosaccharomyces pombe*, from the protozoa *Tetrahymena thermophila* and from the trypanosomes[44–47]. Much of this research was funded generously over the past decade or two, reflecting well on the wisdom of the administrators and peer review system, as well as on the governments who funded these programs. It is impossible to calculate the economic value of this tiny RNA revolution—it is explosive and thus not predictable. But it does seem likely that small RNA pathways are going to intersect many of the processes that are disrupted in human and plant diseases, as well as in human and plant variation that we celebrate, and that viewing those pathways through the lens of small RNAs may break open those fields. We can expect to find small RNA action in synaptic signaling, in the regulation of cell division, in the regulation of organelle trafficking within cells, and so on.

And, of course, small RNAs are small molecules that can inactivate target genes and their products. So, small RNA drugs are likely to emerge as well. Even if small RNA drugs fail to be bioavailable or fail in pharmacokinetic features, it is very likely that traditional pharmaceutical drug development, targeting protein factors that act in small RNA pathways, is going to emerge as important. For example, small RNA pathways are limited to eukaryotes and appear to mediate many key activities of eukaryotes. It is likely that the drug-synthetic capacities of our bacterial competitors and commensals have evolved antibiotics and signals to engage the proteins and perhaps even the RNAs that are so highly conserved across the eukaryotes. Screening for such drugs is likely to represent a rich vein to mine.

But it is important to continue to explore the diversity of biology, and not become myopic about translating biological discovery to humans via, for example, more research on our closer relatives. It is very gratifying to see that comparative genomics now has become so inclusive of phylogenetic diversity—the availability of genome sequences is exposing a new generation of molecular biologists to the richness of biological diversity. Comparative genomics is going to launch the study of thousands and perhaps millions of species, as their genome sequences reveal particular gene complement signatures that may become associated with unique capacities. If the first few steps of the genome era stressed homology and conservation, the next steps may explore duplication and elaboration and deletion of gene pathways in evolution. Such gene duplication and deletion mechanisms are likely to have elaborated the richness of the multicellular biosphere during the relatively short time of the Cambrian explosion, and no doubt since then. And tiny RNA pathways, which appear to be unique to the eukaryotes, are likely to loom large in these specializations.

ACKNOWLEDGMENTS

Most important to the discoveries described here were the fantastic students, postdocs, technical, and administrative staff who created the lab. I would like to express my deepest gratitude to B. Wightman, I. Ha, P. Arasu, J. Giusto, J. Gatto, T. Burglin, B. Reinhart, A. Pasquinelli, F. Slack, S. Kennedy, D. Wang, J. Kim, H. Gabel, R. Kamath, S. Fischer, M. Butler, D. Parry, G. Hayes, X. Wu, C. Zhang, S. Garcia, C. Phillips and S. Curran on the tiny RNA team. An equal number of people on the aging, metabolism and Mars projects also contributed deeply to the intellectual and technical growth of the lab. I launched my lab in the Department of Molecular Biology at Massachusetts General Hospital under very special circumstances: we were completely funded by Hoechst from 1985 to 1992 and almost half funded by them for another decade. This level of patronage allowed us to embark on the study of an array of experimental problems that would have been much more difficult to tackle in a traditional grant-funded environment. H. Goodman deserves very special thanks for his decision to broaden this Hoechst-funded research beyond the endocrinology that was probably their original intent, and for his shrewd recruiting in so many fields. And I thank P. Leder for founding and settling the Department of Genetics at Harvard, my academic home, and my colleagues in the Department of Molecular Biology at Massachusetts General Hospital and the Department of Genetics at Harvard who taught me how to run a lab by their many examples of discovery and training of great scientists. I arrived at graduate school greener than green in 1976, and it was a combination of my fellow students, D. Hanahan, V. Sundaresan, W. Herr, G. Church and T. Wu, and my teachers, F. Ausubel, W. Gilbert, and R. Horvitz, who showed me how to become a scientist. During my postdoc, V. Ambros was my developmental genetics teacher and collaborator extraordinaire, our collaboration now extending over a career, and M. Finney was also a close collaborator, which extended to his postdoctoral work in my lab and current co-direction of our search for extraterrestrial genomes (SETG). So the tribes of my education, my lab and my academic environment were uniquely inspiring, supportive, and loads of fun. But my home tribe has been the wellspring of strength and joy: Natasha Staller is presumably the most sophisticated molecular geneticist among the world's art historians. As a historian of Cubism, she can see cultural inflection points that most do not, and after years of asking me about many details of our work, and reading most of our papers, she is also a very sophisticated biologist. From this vantage point, Natasha nudged me towards ambitious, high-risk projects, and her confidence in my abilities, and in the talents of my students and postdocs, incited a certain boldness. And I am truly grateful to Natasha and to our daughter Victoria for the joyous home life of our little tribe.

1. Sulston, J.E., Schierenberg, E., White, J.G. & Thomson, J.N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
2. Sulston, J.E. & Horvitz, H.R. Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **82**, 41–55 (1981).
3. Chalfie, M., Horvitz, H.R. & Sulston, J.E. Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell* **24**, 59–69 (1981).
4. Ambros, V. & Horvitz, H.R. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226**, 409–416 (1984).
5. Ambros, V. & Horvitz, H.R. The lin-14 locus of *Caenorhabditis elegans* controls the time of expression of specific postembryonic developmental events. *Genes Dev.* **1**, 398–414 (1987).
6. Ambros, V. A hierarchy of regulatory genes controls a larva-to-adult developmental switch in *C. elegans*. *Cell* **57**, 49–57 (1989).
7. Ruvkun, G. *et al.* Molecular genetics of the *Caenorhabditis elegans* heterochronic gene *lin-14*. *Genetics* **121**, 501–516 (1989).
8. Ruvkun, G. & Giusto, J. The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal switch during development. *Nature* **338**, 313–319 (1989).
9. Wightman, B., Bürglin, T.R., Gatto, J., Arasu, P. & Ruvkun, G. Negative regulatory sequences in the *lin-14* 3′-untranslated region are necessary to generate a temporal switch during *C. elegans* development. *Genes Dev.* **5**, 1813–1824 (1991).
10. Arasu, P., Wightman, B. & Ruvkun, G. Temporal regulation of *lin-14* by the antagonistic action of two other heterochronic genes, *lin-4* and *lin-28*. *Genes Dev.* **5**, 1825–1833 (1991).
11. Lee, R.C., Feinbaum, R.L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854 (1993).
12. Wightman, B., Ha, I. & Ruvkun, G. Post-transcriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855–862 (1993).
13. Ha, I., Wightman, B. & Ruvkun, G. A bulged *lin-4/lin-14* RNA duplex is sufficient for *Caenorhabditis elegans lin-14* temporal gradient formation. *Genes Dev.* **10**, 3041–3050 (1996).
14. Olsen, P.H. & Ambros, V. 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 (1999).
15. Reinhart, B.J. *et al.* The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906 (2000).
16. Slack, F.J. *et al.* 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 (2000).
17. Pasquinelli, A.E. *et al.* Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature* **408**, 86–89 (2000).
18. Lee, R.C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864 (2001).
19. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858 (2001).
20. Lau, N.C., Lim, L.P., Weinstein, E.G. & Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862 (2001).
21. Grad, Y. *et al.* Computational and experimental identification of *C. elegans* microRNAs. *Mol. Cell* **11**, 1253–1263 (2003).
22. Klattenhoff, C. & Theurkauf, W. Biogenesis and germline functions of piRNAs. *Development* **135**, 3–9 (2008).
23. Hamilton, A.J. & Baulcombe, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–952 (1999).
24. Elbashir, S.M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200 (2001).
25. Grishok, A. *et al.* Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34 (2001).
26. Hutvágner, G. *et al.* A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small temporal RNA. *Science* **293**, 834–838 (2001).
27. Brodersen, P. *et al.* Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**, 1185–1190 (2008).
28. Chen, X. A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. *Science* **303**, 2022–2025 (2004).
29. Carrington, J.C. & Ambros, V. Role of microRNAs in plant and animal development. *Science* **301**, 336–338 (2003).
30. Parry, D.H., Xu, J. & Ruvkun, G. A whole-genome RNAi screen for *C. elegans* miRNA pathway genes. *Curr. Biol.* **17**, 2013–2022 (2007).
31. Kim, J.K. *et al.* Functional genomic analysis of RNA interference in *C. elegans*. *Science* **308**, 1164–1167 (2005).
32. Kennedy, S., Wang, D. & Ruvkun, G. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* **427**, 645–649 (2004).
33. Wang, D. *et al.* Somatic misexpression of germline P granules and enhanced RNA interference in *C. elegans* retinoblastoma pathway mutants. *Nature* **436**, 593–597 (2005).
34. Fischer, S.E.J., Butler, M.D., Pan, Q. & Ruvkun, G. RNA duplex–mediated *trans*-splicing between independent mRNAs generates *C. elegans* ERI-6/7, a helicase that regulates RNAi. *Nature* (in the press).
35. Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).
36. Blower, M.D., Feric, E., Weis, K. & Heald, R. Genome-wide analysis demonstrates conserved localization of messenger RNAs to mitotic microtubules. *J. Cell Biol.* **179**, 1365–1373 (2007).
37. Buhtz, A., Springer, F., Chappell, L., Baulcombe, D.C. & Kehr, J. Identification and characterization of small RNAs from the phloem of *Brassica napus*. *Plant J.* **53**, 739–749 (2008).
38. Jacobsen, S.E., Running, M.P. & Meyerowitz, E.M. Disruption of an RNA helicase/RNAse III gene in *Arabidopsis* causes unregulated cell division in floral meristems. *Development* **126**, 5231–5243 (1999).
39. Moss, E.G., Lee, R.C. & Ambros, V. 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 (1997).
40. Viswanathan, S.R., Daley, G.Q. & Gregory, R.I. Selective blockade of microRNA processing by Lin28. *Science* **320**, 97–100 (2008).
41. Newman, M.A., Thomson, J.M. & Hammond, S.M. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **14**, 1539–1549 (2008).
42. Richards, M., Tan, S.P., Tan, J.H., Chan, W.K. & Bongso, A. The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells* **22**, 51–64 (2004).
43. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
44. Motamedi, M.R. *et al.* Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**, 789–802 (2004).
45. Cogoni, C. & Macino, G. Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature* **399**, 166–169 (1999).
46. Mochizuki, K., Fine, N.A., Fujisawa, T. & Gorovsky, M.A. Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell* **110**, 689–699 (2002).
47. Ngo, H., Tschudi, C., Gull, K. & Ullu, E. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. USA* **95**, 14687–14692 (1998).