Return to the RNAi World: Rethinking Gene Expression and Evolution (Nobel Lecture)**

Craig C. Mello

It's wonderful to be here today, I would like to start with the most important part, by saying thank you. First of all, I want to thank Andy Fire for being such a tremendous colleague, friend, and collaborator going back over the years. Without Andy I definitely wouldn't be here today. I need to thank the University of Massachusetts for providing for my laboratory, for believing in me, and for giving me not only a place and money, but great colleagues with whom to pursue my research. Without UMass and the great environment provided for me there, I probably would not be here today. And, of course my family; I'm not going to spend time now thanking them individually, but they know how important they are.

I'm going to talk today about *C. elegans* and the role of RNAi in *C. elegans* development. This animal is aptly named for its elegant simplicity (Figure 1). Only one millimeter in

Photo by Bob Goldstein

Figure 1.

length and yet capable of producing 300 progeny in three days by self-fertilization. One of the most beautiful things about *C. elegans*, immediately apparent upon viewing it under the microscope, is its transparency. Sydney Brenner recognized the importance of this attribute when deciding what organism to work on. As animals go, *C. elegans* is relatively simple, having only about a thousand cells in the adult organism. Indeed, the origin and fate of every cell, both in the embryo and adult, has been determined—an amazing accomplishment. At any stage of development, you can look at a cell and

know where that cell came from, tracing its origin back in time to the first division of the embryo.

It's a beautiful system. In fact, the researchers who work with *C. elegans* have their own lineage. Almost all of us can trace ourselves back to Sydney Brenner, who pioneered the modern genetic analysis of this organism. My particular ancestors, if you will, in the lineage of researchers are shown in Figure 2. I owe a tremendous amount of thanks to Dan Stinchcomb, for teaching me molecular biology and really being a fantastic mentor during my initial years in graduate school; Victor Ambros who, along with Dan, provided a wonderful joint laboratory at Harvard University where I did graduate work; and then Jim Priess, who taught me genetics, and was a tremendous mentor and a great friend out in Seattle where I conducted my postdoctoral research at the Fred Hutchinson Cancer Research Center. I owe a tremendous amount to these individuals. I'll show you more pictures of people I will need to thank as I go along.

I want to get down to the theme of my talk today, which really is, in part, about how we continually underestimate the complexity of life. It's the correction of these underestima-

[*] Dr. C. C. Mello
Howard Hughes Medical Institute
and
Program in Molecular Medicine
University of Massachusetts Medical School
373 Plantation Street, Worcester, MA 01605 (USA)
Fax: (+1) 508-856-2950
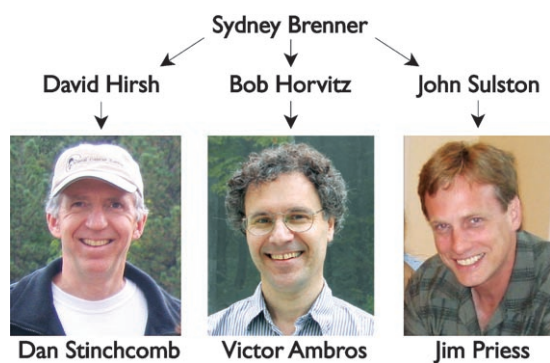E-mail: craig.mello@umassmed.edu

Figure 2.

tions that is quite often what this prize is really recognizing. As science progresses, our knowledge expands, we think we understand, and too often we become overconfident. The fact is, I think we almost always underestimate the complexity of life and of nature. Today has been a true celebration of that beauty and complexity. I attended the Physics talks and the Chemistry talks and it was just spectacular to contemplate. An embryonic universe 13.7 billion years old, originally on the scale of inches across, expanding in seconds through a mysterious process of inflation to occupy the nearly infinite dimensions of space. And to explore the workings of a polymerase at the atomic level, whose origins derive from a common ancestor of all life on Earth some 3.5 billion years ago.

These stories are so beautiful and stunning in their complexity. For every answer they provide, they raise a thousand new questions. And so one thing I'd like to accomplish in my talk, is to raise questions that I can't answer, to talk about the unknown some more. Andy's done such a wonderful job of introducing the subject, giving me the luxury of spending some time talking about potential implications and some of the things we don't know but would love to understand in the future. It is the unknown that inspires us and sparks our curiosity, and so I'd like to try to focus more on telling you about what we don't know and on speculating on what is possible.

If one looks carefully, the complexity of living things becomes strikingly clear. Consider for example the natural environment of *C. elegans*. Figure 3 is an electron micrograph taken by George Barron, who works on nematophagous fungi. The unfortunate worm shown here has become ensnared in a trap set by a fungus that preys on nematodes. It really is a jungle out there for these poor little animals; they struggle to survive, just like the rest of us. The soil is filled with hundreds of different species of these fungi that prey on worms as they're swimming around in the soil. These fungi can sense the motion or contact of a worm and, after the worm has entered its lariat, the fungus inflates it to constrict the snare around the animal, thereby trapping it. The fungus can then send hyphae into the worm to digest it. So, imagine that these poor elegant little animals are actually struggling to survive out there. Nature is filled with complexity that we don't appreciate. It is so tiny, that you would walk over
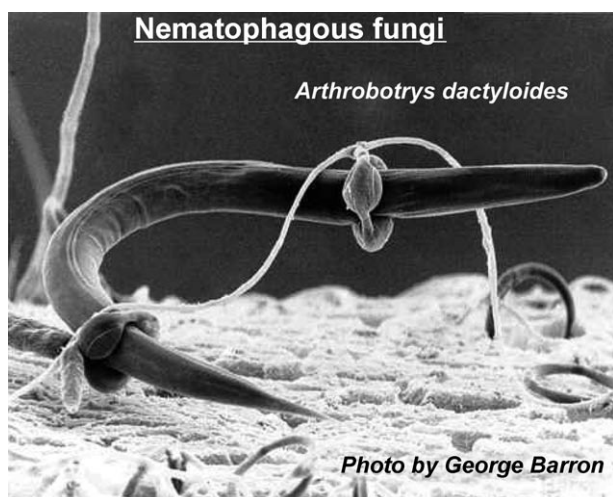
Craig C. Mello is Professor for Molecular Medicine (the Blais University Chair in Molecular Medicine) at the University of Massachusetts Medical School (UMMS). He was also designated an Investigator of the Howard Hughes Medical Institute in 2000, the third HHMI researcher selected at UMMS. The HHMI is a $13 billion medical research organization that employs more than 350 eminent researchers at 72 medical schools, universities, and research institutes worldwide. He obtained his BS in biochemistry from Brown University and his PhD in Cellular and Developmental Biology from Harvard University. He was a postdoctoral fellow at the Fred Hutchinson Cancer Research Center before moving to UMMS in 1995. In 1995 he also received the Pew Scholar Award in the Biomedical Sciences. Together with Andrew Fire, formerly of the Carnegie Institution of Washington, he received the 2006 Nobel Prize in Medicine for their discovery of RNA interference (RNAi). They demonstrated that a certain form of RNA had the unanticipated property of silencing—or interfering with—the expression of a gene whose coding sequence of DNA was similar to that of the RNA they tested. The RNAi mechanism—a natural response of an organism to double-stranded RNA, of which many viruses are comprised—destroys the gene products that a virus needs to replicate itself, essentially halting the progression of the invading viral infection. The discovery, which offers astounding potential for understanding and manipulating the cellular basis of human disease, has had two extraordinary impacts on biological science. One is as a research tool: RNAi is now the state-of-the-art method by which scientists can knock down the expression of specific genes in cells, to thus define the biological functions of those genes. But just as important has been the finding that RNA interference is a normal process of genetic regulation that takes place during development. Thus, RNAi has provided not only a powerful research tool for experimentally knocking out the expression of specific genes, but has opened a completely new and totally unanticipated window on developmental gene regulation. The work of Mello and Fire on RNAi, considered the formative paper on the topic in a 1998 edition of Nature, was hailed as "an electrifying discovery" and named the 2002 "Breakthrough of the Year" by Science magazine. In 2003, Mello and Fire received the prestigious National Academy of Sciences Award in Molecular Biology for their pioneering work as well as the Wiley Prize in the Biomedical Sciences from Rockefeller University; the fourth Annual Aventis Innovative Investigator Award at the Drug Discovery Technology World Conference; and, in 2004, the Warren Triennial Prize, the highest research honor bestowed by Massachusetts General Hospital. In 2005, they were both nominated to the National Academy of Sciences and were recipients of several additional honors including Brandeis University's Lewis S. Rosenstiel Award for Distinguished Work in Medical Research, the Canadian government's Gairdner International Award, and the Massry Prize. In 2006, the pair received the Paul Ehrlich and Ludwig Darmstaedter Prize, one of the highest and most internationally renowned awards conferred in Germany in the field of medicine. Most recently, C. Mello was named the inaugural recipient of The Dr. Paul Janssen Award for Biomedical Research by Johnson & Johnson. Given the fundamental and broad-based impact of RNAi, a patent—"Genetic Inhibition by Double-Stranded RNA" (US Patent 6,506,559 B1) issued to UMMS and Carnegie—is expected to have far-reaching licensing potential, both in the lab and in drug development. As both institutions were eager to bring RNAi to bear as broadly as possible to hasten genetic research, they developed a licensing policy by which companies can readily obtain (for a basic fee) a wide-ranging and non-exclusive license for scientists to use the technology for research. A significant number of companies have already licensed the invention and additional companies have expressed interest.

Figure 3.

numerous cell–cell signaling and differentiation events that give rise to the multicellular organism. These are exemplified by the distribution of the PIE-1 protein (Figure 4). PIE-1 tracks with, and is essential for, germline specification. As
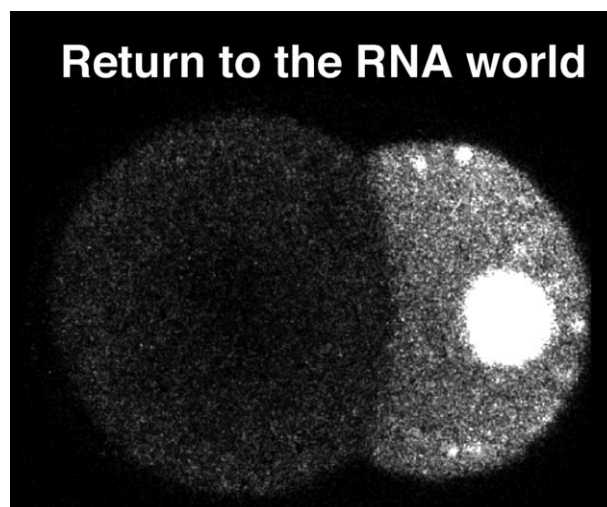

Figure 4.

millions, if not billions, of these little creatures in the soil every day on your way to work, never realizing the things that are happening there.

One of the great triumphs of biology was the discovery of the structure of DNA. The structure of DNA was first determined by Watson and Crick, who showed how two strands composed of four basic building blocks form polymers that intertwine in a beautiful helical staircase structure. This structure explains so much, really, about the basic biology of living things. It explains the segregation patterns, first described by Gregor Mendel, for certain genetic traits of pea plants. The structure alone, as Watson and Crick noted, suggests how the genetic material can be replicated. They stated in their famously brief paper in *Nature*[1] that, "*It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.*" The DNA strands are wrapped around each other and each can template the production of a perfect copy of the strand to which it's bound simply by unwinding and allowing the polymerase to copy it. In Roger Kornberg's talk, we heard about an RNA polymerase that can transcribe the DNA to produce RNA copies of the genetic information. These copies provide templates for the polymerization of the proteins through another elaborate and really beautiful process, called translation, that I certainly don't have time to describe today. I would hope that, if you're interested in these basic workings of the cell, which certainly I think everyone should be interested in, you should look at the literature and do some searching on the Internet to learn more about this process—it's truly amazing.

But, one of the problems with a discovery like this one, of DNA, is that we tend to become overconfident in the explanatory power of the discovery. Does the DNA sequence information control all of the events in the cell? Cells are constantly responding to their environment and to surrounding cells, and these external influences can alter the cell in heritable ways that do not require changes in the primary sequence information in the DNA. Consider the early *C. elegans* embryo. During these early divisions, maternal mRNA and protein products that are stored in the egg direct

shown in this image from a movie, PIE-1—in this case tagged with a glowing jellyfish protein—becomes localized after each division to the germline cell. In this two-cell embryo, PIE-1 protein is localized exclusively to the posterior cell where it is concentrated in the nucleus. This occurs through a fundamental developmental process called asymmetric (unequal) cell division. As a result of this process, the two daughter cells differ with respect to their content of maternally provided products, like PIE-1. These products, in turn, can direct the subsequent development of these cells such that, once differentiated in this way, these cells remain committed to their specific tasks in the animal through numerous rounds of cell division. These remarkably stable differentiation events can be maintained for the entire life of an organism without any underlying changes in the DNA sequence. The germline cells, which in *C. elegans* inherit PIE-1 protein, are the only cells that retain the potential to launch the developmental program again in the next generation.

How do developing cells, all with the same DNA content, lock in different programs of gene expression that are stable through so many rounds of cell division? One possibility, as I will discuss below, is that mechanisms related to those that mediate RNA interference have a role in this process. It has been suggested that the origin of life on Earth may have begun with self-replicating nucleic acid polymers that were more similar chemically to RNA than to DNA, a classic hypothesis referred to as the "RNA World" hypothesis. Hence, the provocative title in Figure 4, "Return to the RNA World," a world in which RNA molecules may have carried, and may still carry, genetic information. The direct ancestors through cell division of the *C. elegans* germ cells were primordial germ cells in the common metazoan (probably wormlike) ancestor of worms and humans, and going even farther back are direct descendants of the hypothetical

self-replicating RNA molecules that gave rise to all life on Earth some 3.5 billion years ago. We heard earlier today, in the physics lectures, that the temperature of the cosmic-background radiation is consistent with an age for the universe of 13.7 billion years. Thus, life on earth is about a quarter of the age of the universe. Living things and these mechanisms that we are talking about today are incredibly ancient. RNA interference itself is at least one billion years old. Biological mechanisms are far more constant than the positions of continents on our planet. That fact and the implicit concept of deep time are among the most profound discoveries of science.

Considering the possible origins of life in a world where information was stored in RNA polymers, and considering the remarkable sophistication of living things and the constancy of the basic and fundamental underlying mechanisms of biology, and finally, considering what we now know about RNA and RNA interference, it is perhaps a good time to reconsider the idea that genetic information is stored primarily in the nucleotide sequence of our DNA. In thinking about this, it is interesting to consider what previous scientists thought about the mechanism of inheritance before DNA and RNA were discovered. For example, in the late 1800s August Weismann, a famous naturalist and early thinker on mechanisms of inheritance, coined the term "biophore" to describe the hereditary agent.[2] Ernst Mayr, in describing Weismann's work in his book *The Growth Of Biological Thought*, characterizes Weismann's ideas as flawed. Weismann said that: 1) there is a special particle, the biophore, for each trait; 2) that these particles can grow and multiply independent of cell division; 3) that both the nucleus and cytoplasm consist of these biophores; 4) that a given biophore may be represented by many replicas in a single nucleus, including the germ cell; and 5) that during cell division the daughter cells may receive different kinds and numbers of biophores through unequal cell division.[3] Mayr concludes that "*As we now know (thanks to Mendel), postulates (2) and (5) are wrong and are responsible for the fact that Weismann was not able to arrive at a correct theory of inheritance.*"

Well, are they really wrong? If you try to apply Weismann's concepts to all genes or genetic traits they are clearly not adequate to explain inheritance. For example, Weismann's biophores could not explain the striking segregation patterns first observed by Mendel for the genetic traits of pea plants. Thus, yes, it would be wrong to apply Weismann's theory to define all genetic inheritance. But let's consider applying Weismann's theory to some traits, and then replace the term "biophores" with the term "siRNAs". Andy introduced siRNAs as these "small interfering RNAs", as we call them; the little chunks of RNA that go on and silence genes. If we put "siRNAs" into each place where "biophores" appears in Weismann's theory, we then have a very different situation: 1) there is a particle, containing siRNAs, for some traits; 2) these siRNAs can grow and multiply independent of cell division; 3) both the nucleus and the cytoplasm can contain the siRNAs; 4) a given siRNA may be represented by many replicas; and 5) that during cell division the daughter cells may receive different kinds and numbers of siRNAs through unequal cell division. And with

these changes, and in light of what we now know about RNAi (as will be discussed more below), it becomes clear that these postulates are not necessarily wrong. Weismann had some very good ideas and we shouldn't discard them out of hand. RNA may play a role in inheritance and evolution. I'll talk about a mechanism for RNA-directed inheritance toward the end of my talk. Furthermore, I'll suggest how natural variation in silencing levels could underlie heritable phenotypic variation upon which evolution could act.

To help introduce RNAi, I'm going to describe some movies that try to capture the essence of the RNAi process. Andy and I take hours to explain RNAi, but that won't do for today's television audience. It's a major problem for television programmers, as you may know. People watch with the remote control handy at all times, so you have to get your point across very quickly, before your viewer loses interest and clicks to another channel. Consequently, in the television industry, they're very good at making models and graphics that can show complex mechanisms like RNAi in just a few seconds.

So, here's what CBS Evening News came up with to try to explain RNAi. For the average viewer at prime time the attention span is about 15 seconds. So, here's what CBS Evening News came up with (Figure 5). In the movie, the
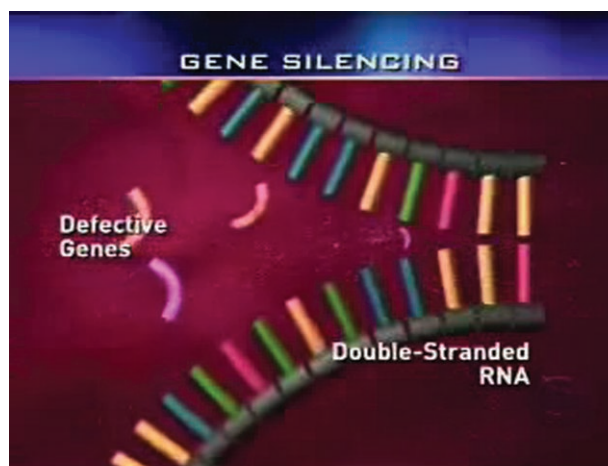

Image © 2005 CBS Evening News

Figure 5.

double-stranded RNA flies onto the scene, then opens at one end and begins to open and close as though it's chewing. Defective genes, shaped like colored cheese puffs, then begin to fly into the mouth of the RNA from the left. The RNA is literally eating the DNA for lunch. Now, Andy and I knew that RNA interference was something incredible when we started working on it, but we really didn't have any idea that it was this incredible. Of course there are a few details that are glossed over in this explanation.

Public broadcasting has the luxury of an audience that tends to have a bit more patience, and they came up with a 15-minute segment and another strategy, "the cop", to explain RNAi (Figure 6). They describe a little policeman who looks out for viruses and other misbehavior in the cell. When he

Image © 2005 PBS NOVA ScienceNOW

Figure 6.

sees double-stranded RNA he realizes something is not right. He then goes on to use "enzymatic Kung Fu" to destroy not only the dsRNA with that sequence, but all RNAs with that sequence that he encounters in the cell.

I like both of these movies because they illustrate a really important concept: that is, that RNAi is an active process, that there is an organismal response to the dsRNA.[4] We realized this at an early stage, because, first of all, as Andy mentioned, the silencing was heritable. RNA injected into an animal resulted in silencing that was transmitted to progeny and even transmitted through crosses for multiple generations via the egg or the sperm. So, the interference mechanism can be initiated in one generation and then transmitted in the germline. And, interestingly, RNAi is also systemic; RNA injected anywhere in the body, or even delivered by ingestion, can get into all the tissues, including the germline. So RNAi involves a transport mechanism, meaning it can be transferred from cell to cell in the body.

The inheritance properties and systemic nature of RNAi, along with its remarkable potency in *C. elegans*, all pointed toward an active organismal response to the double-stranded RNA. What we wanted to do immediately, upon realizing that there was an active response in the organism, was to find the genes in the animal that encode that response. Therefore, we set out to use the powerful genetics of *C. elegans* to look for mutant strains defective in RNAi. We imagined that these mutants would define genes required for the recognition of the foreign double-stranded RNA, genes required for the transport of the silencing signal from cell to cell, genes required for the amplification of silencing, and genes required for the silencing apparatus itself. Hiroaki Tabara (Figure 7) was the first person doing RNAi genetics in the world. He was a courageous postdoc who came to my lab to study cell development, but was willing to tackle something as unusual and as odd as RNAi. The screen that he did was very simple. Basically, he mutagenized animals, let them grow for two generations until mutations that had been induced would become homozygous and then, using a trick first developed by Lisa Timmons in Andy's lab,[5] he fed the worms *E. coli* expressing double-stranded RNA targeting an essential worm gene. According to this strategy, if the animals have an intact RNAi response, the RNAi would knock out the activity of the essential gene, causing lethality. Now, if by chance a mutant

Hiroaki Tabara

Figure 7.

exists in the population that lacks an RNAi response, then RNAi would not occur, and the corresponding animal and its progeny would be viable. Hiroaki used this very powerful genetic selection to identify mutants defective in RNAi, and his screen worked very, very well.

Hiroaki was able to identify numerous mutants. Some of these lacked the RNAi response and had no other obvious phenotypes, like *rde-1* and *rde-4*. However, some of his mutants had additional defects, including a very striking phenotype in which the transposons, which are self-replicating DNA elements present in the genomes of all organisms, became hyperactive, thus causing mutations by jumping from place to place in the genome. In addition, these same mutants had a reduced tendency to silence transgenes in the germline (transgenes are genes that are experimentally introduced into the organism). In normal worms, transgenes have the vexing property of becoming silent after injection into the animal. The same mutants with activated transposons also exhibited activation of transgenes in the germline.

These observations suggested that the normal physiological function of RNAi might be to defend cells against the potentially damaging effects of transposons and other foreign genetic elements (perhaps including viruses). However, there was a big problem with this relatively simple model. The *rde-1* and *rde-4* mutants, as I indicated earlier, had no other phenotypes. They were strongly deficient in RNAi in response to double-stranded RNA, but the transposon-silencing and the transgene-silencing mechanisms were still functioning in these mutant strains. These observations indicated to us, even at that very early stage of our analysis, the existence of some additional very interesting complexity. The rest of the science that I will discuss below really relates to our further investigation of this complexity, and to how these investigations led to the realization that related silencing pathways with distinct triggering mechanisms are at work in *C. elegans*. Keeping in mind that the Nobel committee was careful to recognize, very specifically, the initiation of gene silencing by

double-stranded RNA, I hope that we can look forward to future recognition of silencing that is triggered in other ways. For example, silencing driven from endogenous dsRNA-encoding genes, microRNAs, or silencing triggered by the introduction of transgenes.

Hiroaki cloned the *rde-1* gene and showed that it encodes a highly conserved protein that we now refer to as an Argonaute protein.[6] RDE-1 was an interesting protein for a couple of reasons. It had highly conserved domains found in related genes in organisms as diverse as plants and humans, and yet nothing was known about the enzymatic activities or the biological functions of these domains. This was a very exciting time in the laboratory. We at last had a gene that we knew was involved in the mechanism. Furthermore, previous work on one gene closely related to *rde-1* from *Drosophila* had linked this gene family to an epigenetic silencing pathway in the fruit fly,[7,8] and work in plants had linked a member of the family to the control of development.[9] Very shortly after our paper was published, Carlo Cogoni and Giuseppe Macino[10] published a very nice paper implicating an RDE-1 family member in silencing triggered by the introduction of a transgene in the fungus *Neurospora*. So from these findings in other organisms, and from Hiroaki's genetics, we hypothesized that there may be other types of triggers that initiate related silencing pathways either through natural developmental mechanisms or in response to transposons and transgenes.

A very exciting possibility occurred to us after we cloned *rde-1*. To explain this possibility, I first have to describe some fundamental facts about genes and how the amazingly successful genome-sequencing projects around the world have impacted biological research. Genes are composed of long sequences of nucleotides that specify the protein-coding potential and/or other functions of their gene products. The relationships between genes can be inferred by looking at the nucleotide sequence of the gene. For example, by using the nucleotide sequence to infer the protein-coding potential of all the known genes related to *rde-1*, it was possible to build what is called a phylogenetic tree (Figure 8), in which the most similar members of the gene family (often referred to as homologues) are closest to each other on the tree. Interestingly, it turns out that *rde-1* is a member of a large gene family, with 26 related genes in *C. elegans*. Similarly, there are multiple Argonaute genes in almost every organism. The organisms from which each gene in the tree is derived are indicated by a prefix as follows: *C. elegans* (Ce), humans (Hs), the plant *A. thaliana* (At), fruit fly (Dm), and fission yeast (Sp). The genes named in black represent the most highly conserved branch of the family, with members in plants, animals, and fungi. The green branch, often referred to as the Piwi family after its founding member, has family members in all animals (but not in plants or fungi). Finally, the red branch of the tree represents a *C. elegans*-specific subfamily of genes that are equally divergent from both the black and green families. This remarkable diversity of Argonaute genes raised the exciting possibility that different members of the family have become specialized in each organism to perform distinct functions. For example, *rde-1* according to our genetic studies is required for gene silencing in response to foreign dsRNA.
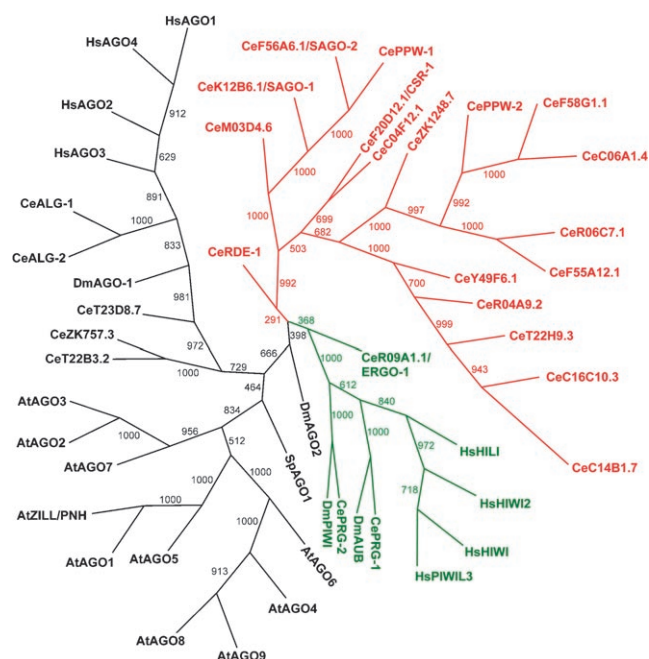

Figure 8.

Perhaps other members of the *C. elegans* Argonaute gene family mediate transposon and transgene silencing. Still others may function in developmental pathways related to RNAi. An outstanding graduate student, Alla Grishok (Figure 9), took on the task of trying to test these ideas.

To do this, Alla set out to inactivate members of the *C. elegans* Argonaute gene family. The first genes that she knocked out encoded two closely related members of the most highly conserved group of Argonautes, named *alg-1* and *alg-2* (see the black branch of the tree, Figure 8). When Alla knocked out these genes by RNAi, she observed a striking phenotype. But in order to explain the significance of her

Alla Grishok

Figure 9.

findings, I have to digress for a moment and tell you about some previous work that set the stage for Alla's discovery.

This previous work goes back to 1993 when, after several years of trying, Victor Ambros' lab succeeded in cloning the *lin-4* gene.[11] One of the reasons *lin-4* had been so hard to clone was that the gene was tiny and did not encode a protein. Instead, the *lin-4* gene appeared to encode two RNA products: an approximately 70-nucleotide-long RNA capable of forming a double-stranded RNA molecule with a hairpin-like structure, and a single-stranded 22-nucleotide RNA that appeared to be derived from this longer RNA (Figure 10).
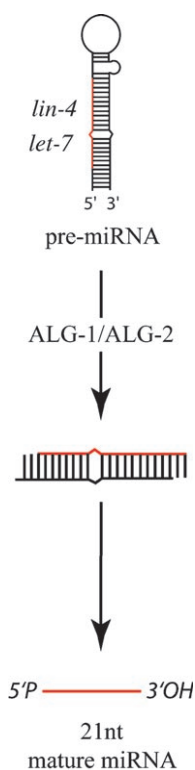
Figure 10.

This short RNA was capable of binding directly to sites in the transcript of the *lin-14* gene, a gene that is negatively regulated by *lin-4* during the normal course of worm development.

Even before we identified the RDE-1 protein, we were interested in the possibility of a relationship between the RNAi pathway and the *lin-4* pathway. Indeed, Hiroaki had raised the concern that RNAi-defective mutants could be hard to recover as viable strains since they might also cause disruption of the *lin-4* pathway. Making all of these possibilities even more exciting—while we were conducting genetic screens for RNAi-deficient strains, beautiful work was published by Hamilton and Baulcombe,[12] who linked small RNAs of about 21 nucleotides to viral gene silencing in plants, and by Gary Ruvkun's lab, who identified a second *lin-4*-like worm gene, *let-7*.[13] Whereas *lin-4* was a worm-specific gene, it turned out that the *let-7* gene had homologues in every animal, including humans. Remarkably, every single nucleotide in the 21-nucleotide mature *let-7* RNA products from the worm and human were identical to each other. The conservation of *let-7* initiated a gold rush to find small RNA-encoding genes, now referred to as micro-RNA genes, in the genomes of numerous organisms. But, despite all of the excitement, the relationship between RNA interference and micro-RNAs had not really been made yet. As Phil Sharp said at one meeting, "*It looks like a horse and smells like a horse*", but there was no molecular or genetic evidence that these pathways were linked.

While this exciting work was going on in worms and plants, biochemists were making rapid progress in reconstituting elements of the silencing pathway in *Drosophila* cell extracts. David Bartel's group along with Phil Zamore, Tom Tuschl, and Phil Sharp at MIT, and Greg Hannon's group at Cold Spring Harbor, spearheaded these efforts.[14, 15] They showed that activities present in *Drosophila* cells could process double-stranded RNA into tiny RNAs approximately 21 nucleotides long. Tom Tuschl and colleagues were the first to show that these small RNAs could silence gene expression in vertebrate cells.[16] Thus, genetic studies in worms had identified small RNAs as silencing agents beginning in 1993, experimental studies of virus infections in plants identified small RNAs accumulating in infected plants, biochemical studies in fly extracts identified small RNAs in extracts, and finally experimental studies identified silencing activity in cellular assays with vertebrate cells. But were these small RNA molecules only similar in size, or did their similarity extend to mechanism? The answer to that key question was still unknown.

Alla's work provided an answer. When Alla knocked out *alg-1* and *alg-2*, she observed a phenotype that was very similar to that observed when you knock-out *let-7*. To confirm this connection we collaborated with Gary Ruvkun and Amy Pasquinelli who had recently developed probes for following the processing of the *lin-4* and *let-7* precursor RNAs into their mature 21-nucleotide RNAs. In wild-type animals the precursor forms are barely detectable. However, they found that, after inactivation of *alg-1* and *alg-2*, this precursor accumulates to high levels, while the product, the mature 21-nucleotide RNA, is greatly diminished (Figure 11).[17]
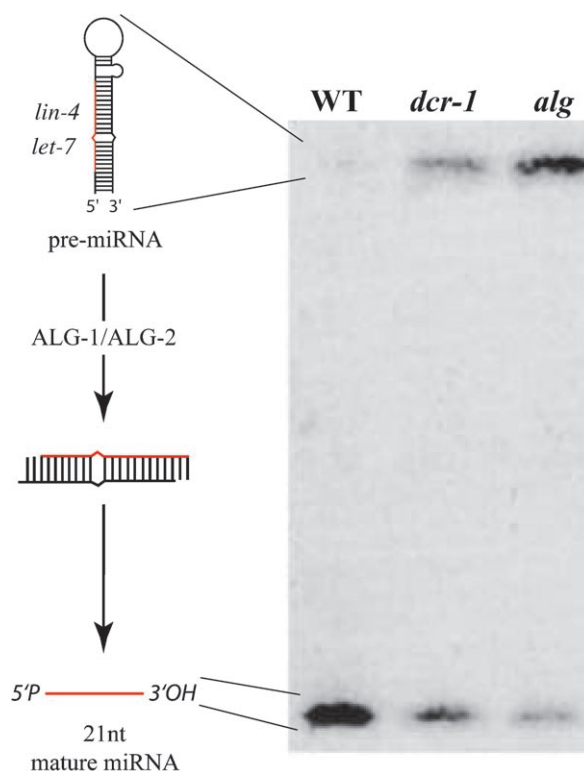
Figure 11.

We also looked at the involvement of the Dicer protein in this process. Dicer was identified by Greg Hannon's lab as a nuclease required for processing long double-stranded RNA into approximately 21-nucleotide fragments in *Drosophila* cells. We were able to show, as did several other groups,[18, 19] that when you knock out Dicer you also see defects in the processing of these micro-RNAs (Figure 11).

With these findings, the first link was established between RNA interference and a natural developmental mechanism for regulating gene expression. This was extremely exciting, and we envisioned a model (Figure 12), in which the RNAi
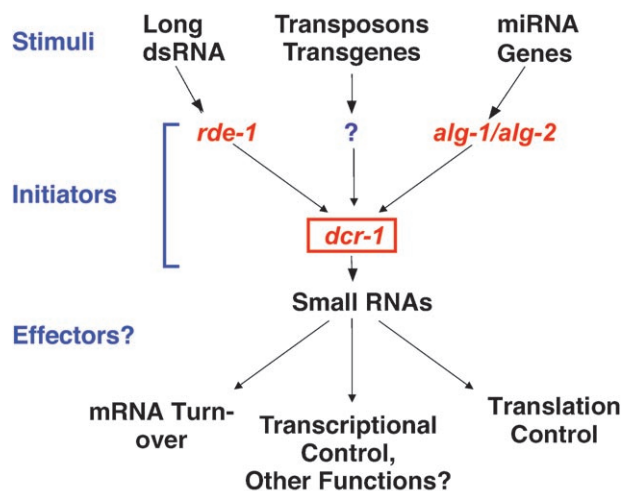

Figure 12.

and micro-RNA pathways utilized different members of the RDE-1 family and converged on Dicer. Downstream of Dicer, these pathways appeared to diverge again, through the action of unknown effectors that direct different types of silencing, including mRNA destruction, transcriptional silencing, and inhibition of translation. And yet we still had not identified the RDE-1 family member involved in transposon and transgene silencing.

At that time we thought of the RDE-1 family members (also known as Argonaute proteins) as initiators of the silencing pathways. Genetic studies had placed RDE-1 at an upstream step in the pathway and, as I showed you, ALG-1 and ALG-2 are required for processing the micro-RNA precursors. However, there was mounting evidence that these proteins might also function downstream in the silencing step. Definitive support for this idea came from Greg Hannon's group through a collaboration with Ji-Joon Song and Leemor Joshua-Tor at Cold Spring Harbor.[20] They showed that Argonaute proteins have structural similarity to an enzyme domain that can cut RNA, and they presented a model for how Argonaute proteins can bind the ends of the short RNAs and utilize the sequence information to find and destroy target mRNAs in the cell. These studies demonstrated that Argonaute proteins represent the long-sought "slicer" activity (or the "cop") that lies at the heart of the RNA-induced silencing pathway.

We were surprised to learn that RDE-1 was probably the slicer enzyme because our genetic studies had placed RDE-1 activity at an upstream step in the pathway. However, we realized that this observation could be explained if Argonaute proteins function more than once during RNAi in *C. elegans*. For example (Figure 13), we imagined that RDE-1 could function along with small RNAs derived from processing of the trigger dsRNA in an initial round of target mRNA
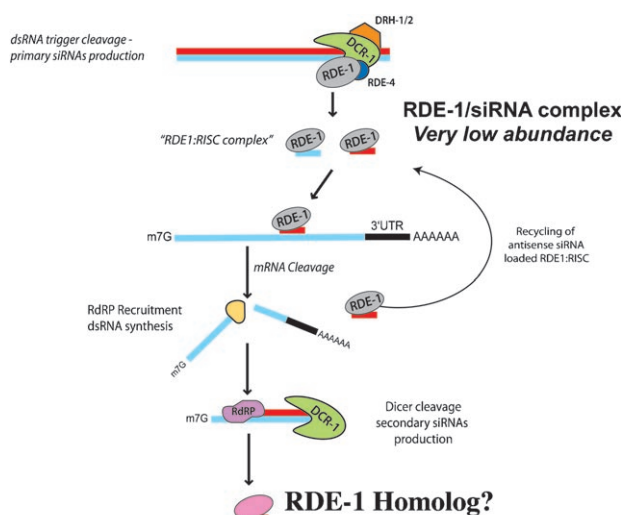

Figure 13.

cleavage. The cleaved target mRNA could then serve as a template for an RNA-dependent RNA polymerase that produces new siRNAs that could, in turn, interact with other Argonaute proteins to mediate efficient silencing of the gene.

Experimental tests of this model were recently published.[21] Surprisingly, we found that, rather than a single additional gene, multiple RDE-1 homologues function together to mediate silencing at the downstream step in the pathway. It was necessary to construct a strain containing six different Argonaute mutants in order to see a strong defect in RNAi. All of these functionally related genes reside within the expanded (red) family of Argonaute genes depicted in Figure 8. These downstream Argonaute proteins are limiting for RNAi. When they are over-expressed, RNAi is enhanced, and when they are inactive, RNAi is decreased. These observations suggest that these genes have been amplified in order to mediate efficient gene silencing.

The mechanism of silencing mediated by these downstream Argonaute proteins remains unknown. It could be through mRNA destruction, but comparison of members of this group of Argonautes to RDE-1 and other members of the family suggest that these downstream Argonaute proteins are not likely to have an intact RNA-cleaving nuclease domain. Our studies of these proteins indicate that they also function in endogenous silencing pathways (Figure 14), including pathways likely to have a role in silencing transposons, transgenes, and other genes at the chromatin level.[21]

The last concept I want to discuss relates to the question of how RNAi can interact with chromatin to silence genes, and the potential importance of this mechanism for gene regulation during both development and evolution. As indicated earlier in my talk, many of the genes involved in RNAi are also required for the silencing of transgenes in the germline. For example, the gene *mut-7* was identified in our screens for genes essential for RNAi,[6] but had also been identified in earlier studies as a gene required for transposon[22] and transgene silencing. While RNAi appeared to have a post-transcriptional effect, several studies suggested that transgene
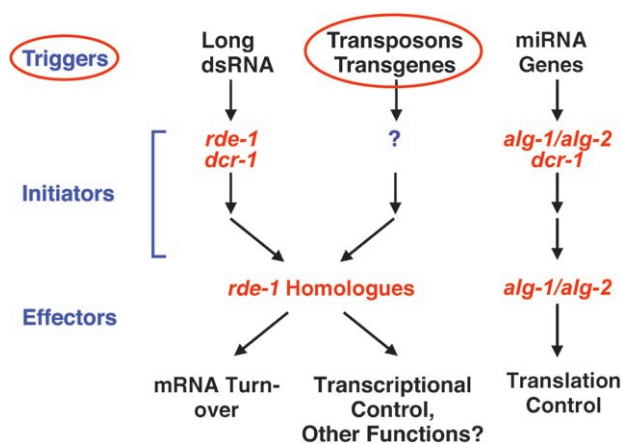
Figure 14.

Figure 15.

silencing involves regulation at the level of the DNA (or more precisely, the chromatin). For example, some of the genes required for transgene silencing in *C. elegans* were related to genes of the polycomb group that interact with chromatin to direct gene silencing in other organisms.[23] Beautiful direct evidence for a link between RNAi and chromatin silencing has more recently come from work in the fission yeast *S. pombe*, where a complex containing an Argonaute protein and known chromatin interacting factors has been shown to interact directly with silenced genes in the nucleus.[24]

To explain how RNAi could regulate DNA directly, I have to tell you a little bit about the physiological nature of DNA inside cells. Your DNA isn't just lying around by itself. The unit of packaging for DNA is a protein structure called the nucleosome. The DNA wraps around the nucleosome twice, and the nucleosomes are in turn wrapped up and packaged into even thicker fibers. Chromosomes are composed of these protein/DNA fibers, also referred to as chromatin. Partly, what's achieved by packaging the DNA into chromatin is a silencing effect. Structural studies on the nucleosome core suggest that short protein tails stick out past the DNA in such a way that they are readily accessible for modification.[25] The modification of these tails, and the resulting regulatory effects on gene expression, is turning out to be a fascinating subject— one that I'm sure this committee will need to consider in the future. Interestingly, mechanisms are now emerging that explain how small RNAs can guide the modification of these chromatin tails.[24] I will illustrate these mechanisms with a model that could not only explain chromatin-based silencing mediated by RNAi but also provide a mechanism for the RNA-mediated evolution concept mentioned earlier (Figure 15).

In this model the DNA (green line) is shown wrapped around the nucleosomes, which are in turn packaged into the higher-order chromatin structures. Modifications to the nucleosome tails that confer an active conformation are shown as four-pointed green stars, while silencing marks are shown as multipointed red stars. In the active conformation, the regulatory region of the gene, called the promoter, is free of nucleosomes and is shown bound by the RNA–polymerase complex, (the complex that produces messenger RNAs and the subject of this year's Nobel Prize for Chemistry). In the
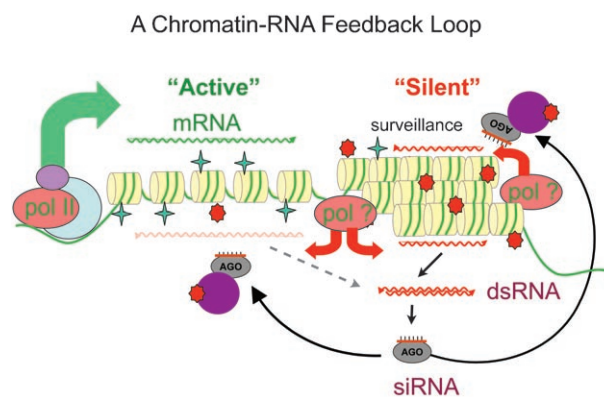
"silent" region, a different kind of polymerase activity is recruited. Instead of producing mRNA, this hypothetical polymerase produces transcripts that enter an RNAi-like silencing pathway. The silencing RNAs could arise by virtue of Dicer-mediated processing of double-stranded RNA. For example, double-stranded RNA could form as the result of bi-directional transcription within the region, or through recruitment of an RNA-dependent RNA polymerase that recognizes some feature of the surveillance RNA. Alternatively, it is also possible that short-interfering RNAs are made directly by transcription from nucleosomal DNA and are loaded onto Argonaute proteins without going through a dsRNA intermediate. Whatever the mechanism for generating the small RNAs, the resulting Argonaute/small RNA complexes could then interact through sequence-specific interactions with nascent surveillance transcripts, or directly with the DNA, to guide chromatin-modifying enzymes back to the locus to reinforce silencing. These silencing complexes could also function in trans to silence other genes with related sequence, such as repeated members of a transposon family.

The concept of transcription occurring not simply to express the gene, but also to regulate it, is extremely powerful. Silencing markers present at low levels within "active regions" could modulate gene expression by specifying the production of intermediate levels of silencing RNAs that in turn specify an intermediate level of gene expression. According to this model, the DNA is like the hardware in a computer and the RNA–chromatin interactions are like the software. The RNA, through interactions with the chromatin, determines not only what regions of the DNA are active, but also the level of activity. When the DNA is replicated and the chromatin is disassembled, the RNA can help reinstall the silencing markers, essentially programming the resulting daughter cells to adopt gene-expression patterns like the mother cell. As Weismann pointed out, asymmetric cell division could segregate this regulatory potential such that the daughter cells become different at one or many different loci.[2] This mechanism could help explain how a somatic cell nucleus can be reprogrammed to undergo embryonic development after transfer into an egg. Mechanisms like this could also help explain how cells are able to maintain their gene-expression programs for decades during an organism's life span.

However, and here is where evolution comes in, this RNA/chromatin feedback mechanism could also function within the germline. Chromatin markers in the germline could specify a level of surveillance-RNA expression that keeps some genes off entirely, and modulates others such that, when they are activated during somatic development, their level of expression is proportional to the amount of silencing RNA produced at the locus. The feedback loop is self-sustaining, but is likely to be subject to natural variations in levels. Upward or downward variations that occur naturally could be selected and transmitted in the germline from one generation to the next. This kind of reversible change in levels of gene expression could play an important role in helping organisms adjust to changes and variations in their environments.

We know from experiments in *C. elegans* that silencing induced by RNAi can be transmitted for multiple generations,[26] and that chromatin-modifying factors appear to play a role in this inheritance mechanism.[27] Given the existence of these phenomena, it is tempting to speculate that all genes might continuously sample, through natural variation, different levels of heritable siRNA–chromatin interactions. Variation of this type could have a major impact on "fitness" and evolution, thus providing a rapid mechanism for evolutionary change mediated through RNA–chromatin interactions, without any underlying changes in the DNA sequence. I will end by saying we simply don't know yet how important small RNAs will turn out to be during development and evolution. I encourage you all to think about the possibilities, to learn more about biology and RNAi, and if you get inspired and excited, please join the adventure and help explore the many unknowns that are still waiting to be addressed.

[1] J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 737–738.
[2] A. Weismann, *The Germ-Plasm: A Theory of Heredity*, Charles Scribner, New York, **1893**, vxiii, p. 477.
[3] E. Mayr, *The Growth of Biological Thought*, Harvard University Press, Cambridge, **1982**, p. 974.
[4] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, C. C. Mello, *Nature* **1998**, *391*, 806–811.
[5] L. Timmons, A. Fire, *Nature* **1998**, *395*, 854.
[6] H. Tabara, M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, L. Timmons, A. Fire, C. C. Mello, *Cell* **1999**, *99*, 123–132.
[7] J. E. Wilson, J. E. Connell, P. M. Macdonald, *Development* **1996**, *122*, 1631–1639.
[8] A. Schmidt, G. Palumbo, M. P. Bozzetti, P. Tritto, S. Pimpinelli, U. Schafer, *Genetics* **1999**, *151*, 749–760.
[9] K. Bohmert, I. Camus, C. Bellini, D. Bouchez, M. Caboche, C. Benning, *EMBO J.* **1998**, *17*, 170–180.
[10] C. Catalanotto, G. Azzalin, G. Macino, C. Cogoni, *Nature* **2000**, *404*, 245.
[11] R. C. Lee, R. L. Feinbaum, V. Ambros, *Cell* **1993**, *75*, 843–854.
[12] A. J. Hamilton, D. C. Baulcombe, *Science* **1999**, *286*, 950–952.
[13] B. J. Reinhart, F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz, G. Ruvkun, *Nature* **2000**, *403*, 901–906.
[14] P. D. Zamore, T. Tuschl, P. A. Sharp, D. P. Bartel, *Cell* **2000**, *101*, 25–33.
[15] S. M. Hammond, E. Bernstein, D. Beach, G. J. Hannon, *Nature* **2000**, *404*, 293–296.
[16] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, T. Tuschl, *Nature* **2001**, *411*, 494–498.
[17] A. Grishok, A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D. L. Baillie, A. Fire, G. Ruvkun, C. C. Mello, *Cell* **2001**, *106*, 23–34.
[18] G. Hutvagner, J. McLachlan, A. E. Pasquinelli, E. Balint, T. Tuschl, P. D. Zamore, *Science* **2001**, *293*, 834–838.
[19] R. F. Ketting, S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, R. H. Plasterk, *Genes Dev.* **2001**, *15*, 2654–2659.
[20] J. J. Song, S. K. Smith, G. J. Hannon, L. Joshua-Tor, *Science* **2004**, *305*, 1434–1437.
[21] E. Yigit, P. J. Batista, Y. Bei, K. M. Pang, C. N. Chen, N. H. Tolia, L. Joshua-Tor, S. Mitani, M. J. Simard, C. C. Mello, *Cell* **2006**, *127*, 747–757.
[22] R. F. Ketting, T. H. Haverkamp, H. G. van Luenen, R. H. Plasterk, *Cell* **1999**, *99*, 133–141.
[23] W. G. Kelly, A. Fire, *Development* **1998**, *125*, 2451–2456.
[24] A. Verdel, S. Jia, S. Gerber, T. Sugiyama, S. Gygi, S. I. Grewal, D. Moazed, *Science* **2004**, *303*, 672–676.
[25] T. J. Richmond, C. A. Davey, *Nature* **2003**, *423*, 145–150.
[26] A. Grishok, H. Tabara, C. C. Mello, *Science* **2000**, *287*, 2494–2497.
[27] N. L. Vastenhouw, K. Brunschwig, K. L. Okihara, F. Muller, M. Tijsterman, R. H. Plasterk, *Nature* **2006**, *442*, 882.