RNA interference

So that I don't get myself into any trouble, most of the following text is verbatim from the Ambion website, which nicely summarizes a lot of what I will talk about in class. I've edited it a bit and pasted in some figures to help, but I'm not keeping tract of what the website had and what I've inserted. If you want to see the original Ambion text, go to http://www.ambion.com/techlib/hottopics/rnai/.

The first experiments

Post-transcriptional gene silencing (PTGS), which was initially considered a bizarre phenomenon limited to petunias and a few other plant species, is now one of the hottest topics in molecular biology (1). In the last few years, it has become clear that PTGS occurs in both plants and animals and has roles in viral defense and transposon silencing mechanisms. Perhaps most exciting, however, is the emerging use of PTGS and, in particular, RNA interference (RNAi) — PTGS initiated by the introduction of double-stranded RNA (dsRNA) — as a tool to knock out expression of specific genes in a variety of organisms (reviewed in 1-3).

How was RNAi discovered? How does it work? Perhaps more importantly, how can it be harnessed for functional genomics experiments? This article will briefly answer these questions and provide you with resources to find in depth information on PTGS and RNAi research.

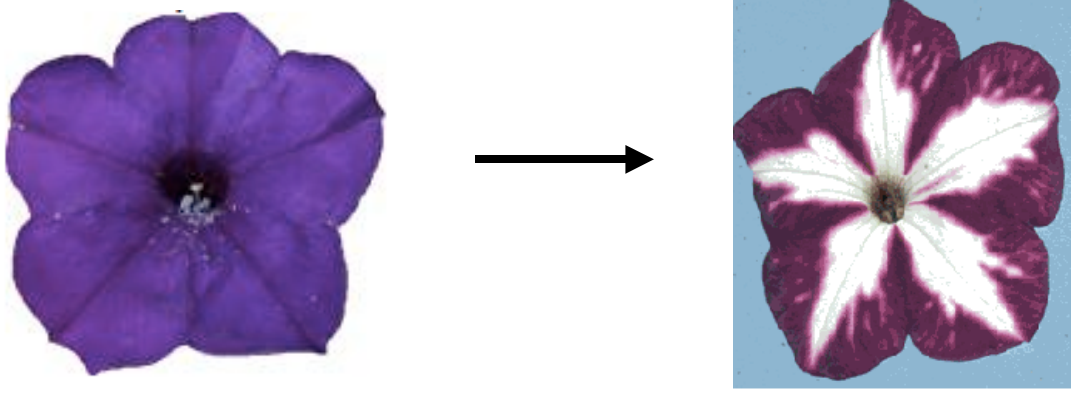

Figure 1. The attempt to overexpress chalone synthetase, an enzyme that produces anthocyanin pigment, resulted not in a darker petunia flower, but instead a loss of flower pigment.

More than a decade ago, a surprising observation was made in petunias. While trying to deepen the purple color of these flowers, Rich Jorgensen and colleagues introduced a pigment-producing gene under the control of a powerful promoter. Instead of the expected deep purple color, many of the flowers appeared variegated or even white. Jorgensen named the observed phenomenon "cosuppression", since the expression of both the introduced gene and the homologous endogenous gene was suppressed (1-5).

First thought to be a quirk of petunias, cosuppression has since been found to occur in many species of plants. It has also been observed in fungi, and has been particularly well characterized in Neurospora crassa, where it is known as "quelling" (1-3).

But what causes this gene silencing effect? Although transgene-induced silencing in some plants appears to involve gene-specific methylation (transcriptional gene silencing, or TGS), in others silencing occurs at the post-transcriptional level (post-transcriptional gene silencing, or PTGS). Nuclear run-on experiments in the latter case show that the homologous transcript is made, but that it is rapidly degraded in the cytoplasm and does not accumulate (1, 3, 6).

Introduction of transgenes can trigger PTGS, however silencing can also be induced by the introduction of certain viruses (2, 3). Once triggered, PTGS is mediated by a diffusible, trans-acting molecule. This was first demonstrated in Neurospora, when Cogoni and colleagues showed that gene silencing could be transferred between nuclei in heterokaryotic strains (1, 7). It was later confirmed in plants when Palauqui and colleagues induced PTGS in a host plant by grafting a silenced, transgene-containing source plant to an unsilenced host (8). From work done in nematodes and flies, we now know that the trans-acting factor responsible for PTGS in plants is dsRNA (1-3).

RNAi Is Discovered in Nematodes

The first evidence that dsRNA could lead to gene silencing came from work in the nematode Caenorhabditis elegans. Seven years ago, researchers Guo and Kemphues were attempting to use

antisense RNA to shut down expression of the *par-1* gene in order to assess its function. As expected, injection of the antisense RNA disrupted expression of *par-1*, but quizzically, injection of the sense-strand control did too (9).

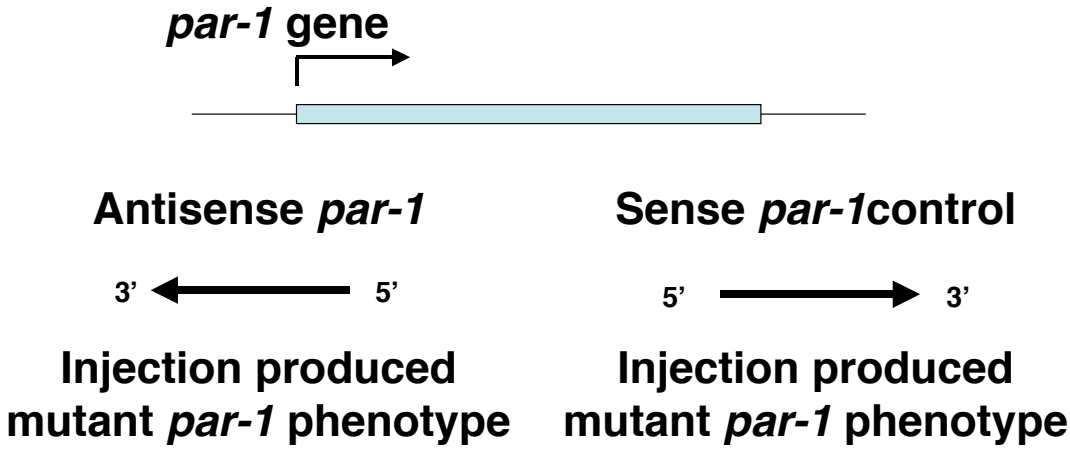

Figure 2. The Guo & Kemphues experiment.

This result was a puzzle until three years later. It was then that Fire and Mello first injected double-stranded RNA (dsRNA) — a mixture of both sense and antisense strands — into C. elegans (Figure 2) (10). This injection resulted in much more efficient silencing than injection of either the sense or the antisense strands alone. This is probably more info that you would like to look at, but I've inserted Table 1 from their paper, which shows of the effects that they saw for different genes. In most cases, injection of sense or antisense had little or no effect, but injection of dsRNA caused the same phenotype as mutants that had lost function in the gene being tested.

These investigators also showed that injection of dsRNA to a particular gene reduced the levels of the endogenous transcripts. This is shown in Figure 3 for the *mex-3* gene. These pictures are of four cell *C. elegans* embryos. A labeled antisense probe is used to detect the *mex-3* transcript in these embryos. Panel (a) shows an embryo that wasn't probed with antisense *mex-3*. Panel (b) shows an embryo probed with *mex-3* antisense to detect *mec-3* mRNA. The dark staining indicates that the cells are expressing lots of the transcript. Panel (c) is an embryo that came from a mother that had been injected with antisense *mex-3*. (Don't get confused by the use of antisense used in the injection of the mother and the use of an

antisense probe to detect the mRNA in the embryo.) You can see that the level of *mec-3* transcript is reduced compared with panel (b), but when dsRNA is injected into the mother her embyos (panel d) had no detectible *mec-3* RNA when probed with antisense RNA. Thus, injection of dsRNA appeared to affect RNA stability.

Table 1 Effects of sense, antisense and mixed RNAs on progeny of injected animals

Gene	segment	Size (kilobases)	Injected RNA	F_1 phenotype
unc-22				*unc-22*-null mutants: strong twitchers[7,8]
*unc22A**	Exon 21–22	742	Sense	Wild type
			Antisense	Wild type
			Sense + antisense	Strong twitchers (100%)
unc22B	Exon 27	1,033	Sense	Wild type
			Antisense	Wild type
			Sense + antisense	Strong twitchers (100%)
unc22C	Exon 21–22†	785	Sense + antisense	Strong twitchers (100%)
fem-1				*fem-1*-null mutants: femal (no sperm)[13]
fem1A	Exon 10‡	531	Sense	Hermaphrodite (98%)
			Antisense	Hermaphrodite (>98%)
			Sense + antisense	Female (72%)
fem1B	Intron 8	556	Sense + antisense	Hermaphrodite (>98%)
unc-54				*unc-54*-null mutants: paralysed[11]
unc54A	Exon 6	576	Sense	Wild type (100%)
			Antisense	Wild type (100%)
			Sense + antisense	Paralysed (100%)§
unc54B	Exon 6	651	Sense	Wild type (100%)
			Antisense	Wild type (100%)
			Sense + antisense	Paralysed (100%)§
unc54C	Exon 1–5	1,015	Sense + antisense	Arrested embryos and larvae (100%)
unc54D	Promoter	567	Sense + antisense	Wild type (100%)
unc54E	Intron 1	369	Sense + antisense	Wild type (100%)
unc54F	Intron 3	386	Sense + antisense	Wild type (100%)
hlh-1				*hlh-1*-null mutants: lumpy-dumpy larvae[16]
hlh1A	Exons 1–6	1,033	Sense	Wild type (<2% lpy-dpy)
			Antisense	Wild type (<2% lpy-dpy)
			Sense + antisense	Lpy-dpy larvae (>90%)‖
hlh1B	Exons 1–2	438	Sense + antisense	Lpy-dpy larvae (>80%)‖
hlh1C	Exons 4–6	299	Sense + antisense	Lpy-dpy larvae (>80%)‖
hlh1D	Intron 1	697	Sense + antisense	Wild type (<2% lpy-dpy)
myo-3-driven GFP transgenes¶				
myo-3::NLS::gfp::lacZ				Makes nuclear GFP in body muscle
gfpG	Exons 2–5	730	Sense	Nuclear GFP-LacZ pattern of parent strain
			Antisense	Nuclear GFP-LacZ pattern of parent strain
			Sense + antisense	Nuclear GFP-LacZ absent in 98% of cells
lacZL	Exon 12–14	830	Sense + antisense	Nuclear GFP-LacZ absent in >95% of cells
myo-3::MtLS::gfp				Makes mitochondrial GFP in body muscle
gfpG	Exons 2–5	730	Sense	Mitochondrial-GFP pattern of parent strain
			Antisense	Mitochondrial-GFP pattern of parent strain
			Sense + antisense	Mitochondrial-GFP absent in 98% of cells
lacZL	Exon 12–14	830	Sense + antisense	Mitochondrial-GFP pattern of parent strain

Each RNA was injected into 6–10 adult hermaphrodites (0.5×10^6–1×10^6 molecules into each gonad arm). After 4–6 h (to clear prefertilized eggs from the uterus), injected animals were transferred and eggs collected for 20–22 h. Progeny phenotypes were scored upon hatching and subsequently at 12–24-h intervals.

* to obtain a semiquantitative assessment of the relationship between RNA dose and phenotypic response, we injected each *unc22A* RNA preparation at a series of different concentrations (see figure in Supplementary information for details). At the highest dose tested (3.6×10^6 molecules per gonad), the individual sense and antisense *unc22A* preparations produced some visible twitching (1% and 11% of progeny, respectively). Comparable doses of double-stranded *unc22A* RNA produced visible twitching in all progeny, whereas a 120-fold lower dose of double-stranded *unc22A* RNA produced visible twitching in 30% of progeny. † *unc22C* also carries the 43-nucleotide intron between exons 21 and 22. ‡ *fem1A* carries a portion (131 nucleotides) of intron 10. § Animals in the first affected broods (layed 4–24 h after injection) showed movement defects indistinguishable from those of *unc-54*-null mutants. A variable fraction of these animals (25%–75%) failed to lay eggs (another phenotype of *unc-54*-null mutants), whereas the remainder of the paralysed animals did lay eggs. This may indicate incomplete interference with *unc-54* activity in vulval muscles. Animals from later broods frequently show a distinct partial loss-of-function phenotype, with contractility in a subset of body-wall muscles. ‖ Phenotypes produced by RNA-mediated interference with *hlh-1* included arrested embryos and partially elongated L1 larvae (the *hlh-1*-null phenotype). These phenotypes were seen in virtually all progeny after injection of double-stranded *hlh1A* and in about half of the affected animals produced after injection of double-stranded *hlh1B* and double-stranded *hlh1C*. A set of less severe defects was seen in the remainder of the animals produced after injection of double-stranded *hlh1B* and double-stranded *hlh1C*). The less severe phenotypes are characteristic of partial loss of function of *hlh-1* (B. Harfe and A.F., unpublished observations). ¶ the host for these injections, strain PD4251, expresses both mitochondrial GFP and nuclear GFP-LacZ (see Methods). This allows simultaneous assay for interference with *gfp* (seen as loss of all fluorescence) and with *lacZ* (loss of nuclear fluorescence). The table describes scoring of animals as L1 larvae. Double-stranded *gfpG* caused a loss of GFP in all but 0–3 of the 95 body muscles in these larvae. As these animals mature to adults, GFP activity was seen in 0–5 additional body-wall muscles and in the 8 vulval muscles. Lpy-dpy, lumpy-dumpy.

So how does injection of dsRNA lead to gene silencing? Many research groups have diligently worked over the last few years to answer this important question. A key finding by Baulcombe and Hamilton provided the first clue. They identified RNAs of ~25 nucleotides in plants undergoing cosuppression that were absent in non-silenced plants. These RNAs were complementary to both the sense and antisense strands of the gene being silenced (24).

Further work in Drosophila — using embryo lysates and an in vitro system derived from S2 cells — shed more light on the subject (3, 25, 26). In one notable series of experiments, Zamore and colleagues found that dsRNA added to Drosophila embryo lysates was processed to 21-23 nucleotide species. They also found that the homologous endogenous mRNA was cleaved only in the region corresponding to the introduced dsRNA and that cleavage occurred at 21-23 nucleotide intervals (26). Rapidly, the mechanism of RNAi was becoming clear.

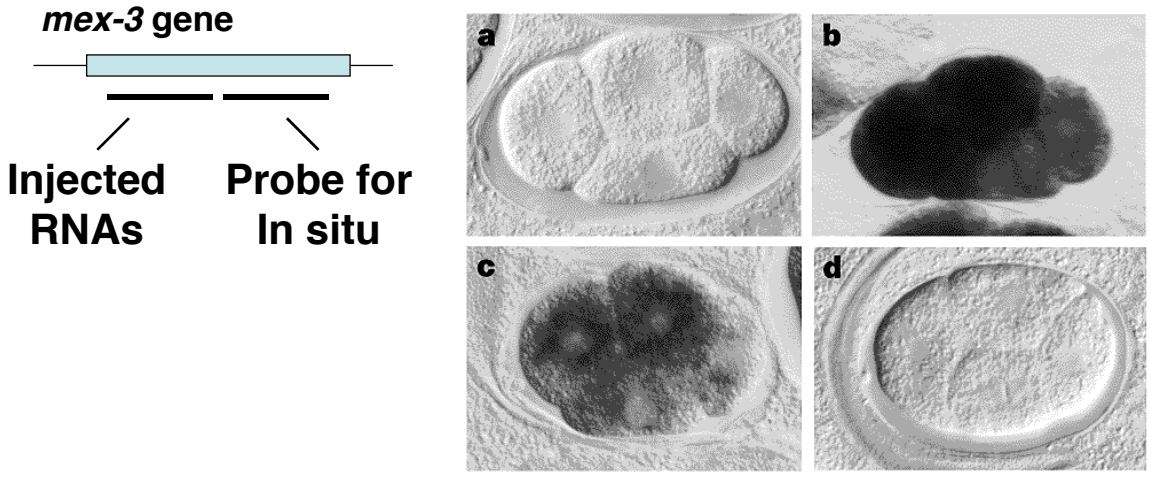

Figure 3. dsRNA affects RNA stability.

Current Models of the RNAi Mechanism

Both biochemical and genetic approaches (see "The Genes and Enzymes Involved in PTGS and RNAi" below for a discussion of genetic approaches used to understand RNAi) have led to the current models of the RNAi mechanism. In these models, RNAi includes both initiation and effector steps (27, see also a Flash animation of "How Does RNAi Work?", from reference 3).

In the initiation step, input dsRNA is digested into 21-23 nucleotide small interfering RNAs (siRNAs), which have also been called "guide

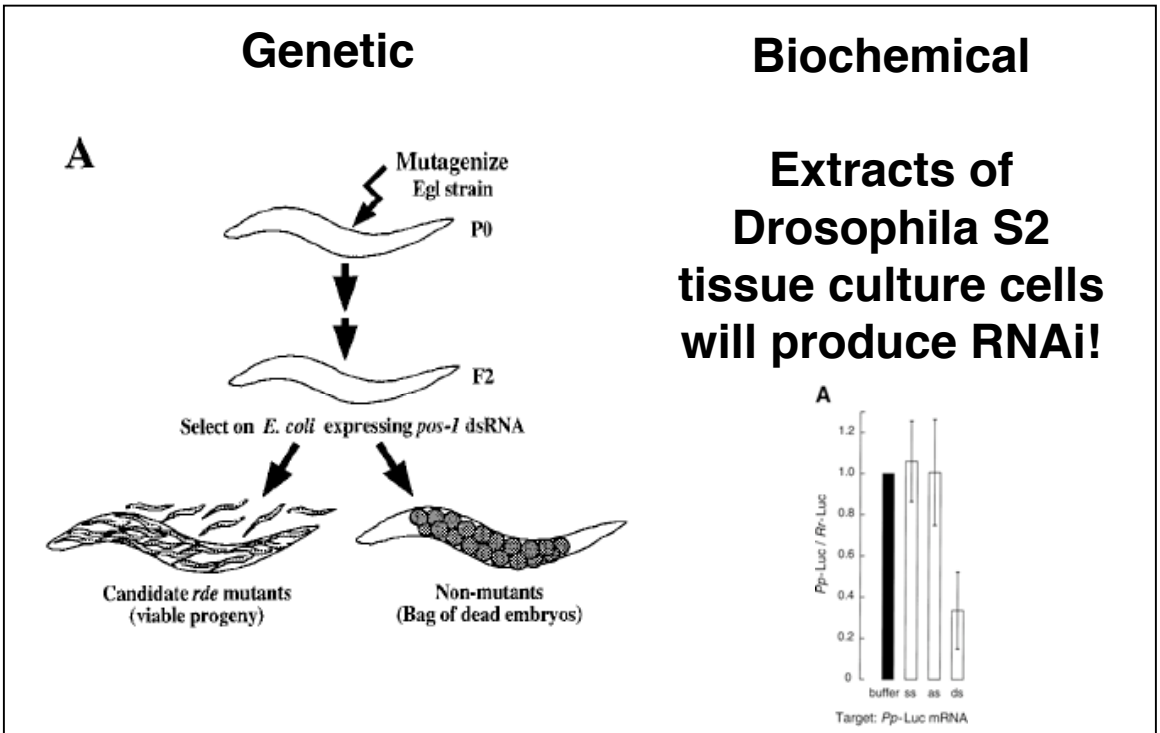

Figure 4. Genetic and biochemical approaches were used to identify molecules involved in RNAi.

RNAs" (reviewed in 3, 18, 27). Evidence indicates that siRNAs are produced when the enzyme Dicer, a member of the RNase III family of dsRNA-specific ribonucleases, processively cleaves dsRNA (introduced directly or via a transgene or virus) in an ATP-dependent, processive manner. Successive cleavage events degrade the RNA to 19–21 bp duplexes (siRNAs), each with 2-nucleotide 3' overhangs (27, 28).

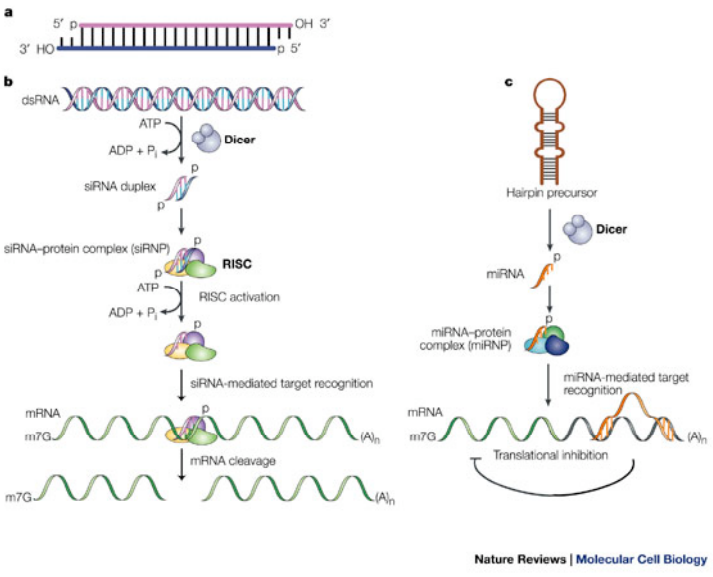

Figure 5. Roles of Dicer and the RISC complex in RNAi and miRNA function

In the effector step, the siRNA duplexes bind to a nuclease complex to form what is known as the RNA-induced silencing complex, or RISC. An ATP-depending unwinding of the siRNA duplex is required for activation of the RISC. The active RISC then targets the homologous transcript by base pairing interactions and cleaves the mRNA ~12 nucleotides from the 3' terminus of the siRNA (3, 18, 27, 29). Although the mechanism of cleavage is at this date unclear, research indicates that each RISC contains a single siRNA and an RNase that appears to be distinct from Dicer (27).

Because of the remarkable potency of RNAi in some organisms, an amplification step within the RNAi pathway has also been proposed. Amplification could occur by copying of the input dsRNAs, which would generate more siRNAs, or by replication of the siRNAs themselves (see "Possible Role for RNA-dependent RNA Polymerase" below). Alternatively or in addition, amplification could be effected by multiple turnover events of the RISC (3, 18, 27).

Possible Role for RNA-dependent RNA Polymerase

Genetic screens in *Neurospora, C. elegans*, and *Arabidopsis* have identified several genes that appear to be crucial for PTGS and RNAi. Several of these, including Neurospora *qde-1*, Arabidopsis SDE-1/SGS-2 and C. elegans *ego-1*, appear to encode RNA-dependent RNA polymerases (RdRPs). At first glance, it might be assumed that this is proof that an RdRP activity is required for RNAi. Certainly the existence of an RdRP might explain the remarkable efficiency of dsRNA-induced silencing if it amplifed either the dsRNA prior to cleavage or the siRNAs directly. But mutants of these genes have varying phenotypes, which makes the role of RdRP in RNAi difficult to discern (1, 3, 17, 18).

In *C. elegans ego-1* mutants ("ego" stands for "enhancer of glp-1"), RNAi functions normally in somatic cells, but is defective in germline cells where ego-1 is primarily expressed. In Arabidopsis SDE-1/SGS-2 mutants ("SGS" stands for suppressor of gene silencing), siRNAs are produced when dsRNA is introduced via an endogenously replicating RNA virus, but not when introduced by a transgene. It has been proposed that perhaps the viral RdRP is substituting for the Arabidopsis enzyme in these mutants (1, 3, 17, 18). Although no homolog of an RdRP has been found in flies or humans, an RdRP activity has recently been reported in Drosophila embryo lysates (30). One model of amplification, termed the

"random degradative PCR" model, suggests that an RdRP uses the guide strand of an siRNA as a primer for the target mRNA, generating a dsRNA substrate for Dicer and thus more siRNAs (27, 30). Evidence supporting this model has been found in worms, whereas experimental results refuting the model have been obtained from Drosophila embryo lysates (26, 27).

RNAi Initiators

Two *C. elegans* genes, *rde-1* and *rde-4* ("rde" stands for "RNAi deficient"), are believed to be involved in the initiation step of RNAi. Mutants of these genes produce animals that are resistant to silencing by injection of dsRNA, but silencing can be effected in these animals by the transmission of siRNA from heterozygous parents that are not silencing deficient. The C. elegans rde-1 gene is a member of a large family of genes and is homologous to the Neurospora *qde-2* ("qde" stands for "quelling deficient") and the Arabidopsis AGO1 genes ("AGO" stands for "argonaute"; AGO1 was previously identified to be involved in Arabidopsis development). Although the function of these genes in PTGS is unclear, a mammalian member of the RDE-1 family has been identified as a translation initiation factor. Interestingly, Arabidopsis mutants of AGO1, which are defective for cosuppression, also exhibit defects in leaf development. Thus some processes or enzymes involved in PTGS may also be involved in development (1, 3, 17, 18).

RNAi Effectors

Important genes for the effector step of PTGS include the *C. elegans rde-2* and *mut-7* genes. These genes were initially identified from heterozygous mutant worms that were unable to transmit RNAi to their homozygous offspring (16). Worms with mutated *rde-2* or *mut-7* genes exhibit defective RNAi, but interestingly, they also demonstrate increased levels of transposon activity. Thus, silencing of transposons appears to occur by a mechanism related to RNAi and PTGS. Although the *rde-2* gene product has not yet been identified, the *mut-7* gene encodes a protein with homology to the nuclease domains of RNase D and a protein implicated in Werner syndrome (a rapid aging disease) in humans (1, 3, 17, 18, 31). Perhaps this protein is a candidate for the nuclease activity required for target RNA degradation.

PTGS Has Ancient Roots

Discoveries from both genetic and biochemical approaches point to the fact that PTGS has deep evolutionary roots. Proposals have been put forth that PTGS evolved as a defense mechanism against transposons or RNA viruses, perhaps before plants and animals diverged (1, 3, 17, 18).

Interestingly, it was noted by many researchers that disruption of genes required for RNAi often causes severe developmental defects. This observation suggested a link between RNAi and at least one developmental pathway.

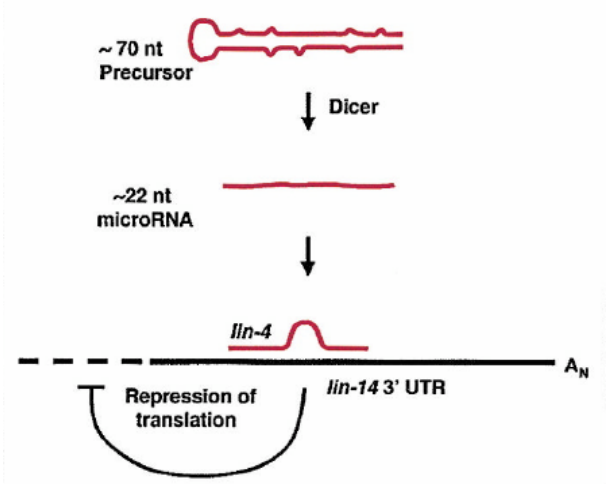

Figure 6. Model for *lin-4* inhibition of *lin-14*.

A group of small RNA molecules, known as micro RNAs (miRNAs), regulates *C. elegans* developmental timing through translational repression of target transcripts. Research indicates that the *C. elegans* *lin-4* and *let-7* miRNAs are generated from 70-nt transcripts following the folding of these longer transcripts into a stem-loop structure. The folded RNA molecules are cleaved to produce 22-nt stRNAs by the enzyme Dicer (called DCR-1 in *C. elegans*). Thus Dicer generates both siRNAs and stRNAs, and represents an intersection point for the RNAi and stRNA pathways (Figure 6.) (32-34). The *lin-4* miRNA inhibits the translation of the *lin-14* mRNA. It is this loss of *lin-14* expression that allows development to proceed. The *let-7* miRNA has a similar effect on *lin-28* translation.

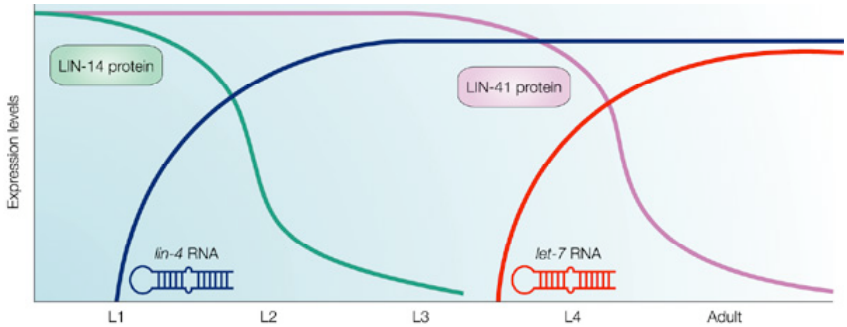

Figure 7. Model for how *lin-4* and *let-7* miRNAs regulate LIN-14 and LIN-41 proteins levels.

Recently, nearly 100 additional ~22 nt RNA molecules, termed microRNAs (miRNAs), were identified in Drosophila, *C. elegans*, and HeLa cells (35-38). Much like *lin-4* and *let-7*, these miRNAs are formed from precursor RNA molecules that fold into a stem-loop secondary structure. The newly discovered ~22 nt miRNAs are believed to play a role in regulation of gene expression, and at least two of them are known to require Dicer for their production (37). It appears that the use of small RNAs for both gene regulation and RNAi is a common theme throughout evolution.

Non-specific Gene Silencing by Long dsRNAs

While the natural presence of RNAi had been observed in a variety of organisms (plants, protozoa, insects, and nematodes), evidence for the existence of RNAi in mammalian cells took longer to establish. Transfection of long dsRNA molecules (>30 nt) into most mammalian cells causes nonspecific suppression of gene expression, as opposed to the gene-specific suppression seen in other organisms. This suppression has been attributed to an antiviral response, which takes place through one of two pathways.

In one pathway, long dsRNAs activate a protein kinase, PKR. Activated PKR, in turn phoshorylates and inactivates the translation initiation factor, eIF2a, leading to repression of translation. (39) In the other pathway, long dsRNAs activate RNase L, which leads to nonspecific RNA degradation (40).

A number of groups have shown that the dsRNA-induced antiviral response is absent from mouse embryonic stem (ES) cells and at least one cell line of embryonic origin. (41, 42) It is therefore possible to use long dsRNAs to silence specific genes in these specific mammalian cells. However, the antiviral response precludes the use of long dsRNAs to induce RNAi in most other mammalian cell types.

siRNAs Bypass the Antiviral Response

Interestingly, dsRNAs less than 30 nt in length do not activate the PKR kinase pathway. This observation, as well as knowledge that long dsRNAs are cleaved to form siRNAs in worms and flies and that siRNAs can induce RNAi in Drosophila embryo lysates, prompted researchers to test whether introduction of siRNAs could induce gene-specific silencing in mammalian cells (43). Indeed, siRNAs

introduced by transient transfection were found to effectively induce RNAi in mammalian cultured cells in a sequence-specific manner. The effectiveness of siRNAs varies — the most potent siRNAs result in >90% reduction in target RNA and protein levels (44-46). The most effective siRNAs turn out to be 21 nt dsRNAs with 2 nt 3' overhangs. Sequence specificity of siRNA is very stringent, as single base pair mismatches between the siRNA and its target mRNA dramatically reduce silencing (44, 47). Unfortunately, not all siRNAs with these characteristics are effective. The reasons for this are unclear but may be a result of positional effects (46, 48, 49). For current recommendations on designing siRNAs, see "siRNA Design".

Although the history and mechanism of RNAi and PTGS are fascinating, many researchers are most excited about RNAi's potential use as a functional genomics tool. Already RNAi has been used to ascertain the function of many genes in Drosophila, C. elegans, and several species of plants. With the knowledge that RNAi can be induced in mammalian cells by the transfection of siRNAs, many more researchers are beginning to use RNAi as a tool in human, mouse and other mammalian cell culture systems.

In early experiments with mammalian cells, the siRNAs were synthesized chemically (Ambion is one of several companies that offer custom siRNA synthesis). Recently, Ambion introduced a kit (the Silencer™ siRNA Construction Kit) to produce siRNAs by in vitro transcription, which is a less expensive alternative to chemical synthesis, particularly when multiple different siRNAs need to be synthesized. Once made, the siRNAs are introduced into cells via transient transfection. Due to differences in efficacy, most researchers will synthesize 3–4 siRNAs to a target gene and perform pilot experiments to determine the most effective one. Transient silencing of more than 90% has been observed with this type of approach (44-46, 48, 49).

So far, injection and transfection of dsRNA into cells and organisms have been the main method of delivery of siRNA. And while the silencing effect lasts for several days and does appear to be transferred to daughter cells, it does eventually diminish. Recently, however, a number of groups have developed expression vectors to continually express siRNAs in transiently and stably transfected mammalian cells (50-56). Some of these vectors have been

engineered to express small hairpin RNAs (shRNAs), which get processed in vivo into siRNAs-like molecules capable of carrying out gene-specific silencing (50, 53, 54, 56). The vectors contain the shRNA sequence between a polymerase III (pol III) promoter and a 4-5 thymidine transcription termination site. The transcript is terminated at position 2 of the termination site (pol III transcripts naturally lack poly(A) tails) and then folds into a stem-loop structure with 3' UU-overhangs. The ends of the shRNAs are processed in vivo, converting the shRNAs into ~21 nt siRNA-like molecules, which in turn initiate RNAi (50). This latter finding correlates with recent experiments in C. elegans, Drosophila, plants and Trypanosomes, where RNAi has been induced by an RNA molecule that folds into a stem-loop structure (reviewed in 3).

Another siRNA expression vector developed by a different research group encodes the sense and antisense siRNA strands under control of separate pol III promoters (52). The siRNA strands from this vector, like the shRNAs of the other vectors, have 5 thymidine termination signals. Silencing efficacy by both types of expression vectors was comparable to that induced by transiently transfecting siRNA.

The recent studies on RNAi have taken the research world by storm. The ability to quickly and easily create loss-of-function phenotypes has researchers rushing to learn as much as they can about RNAi and the characteristics of effective siRNAs. In the future, RNAi may even hold promise for development of gene-specific therapeutics. Much has been learned about this powerful technique, but additional information becomes available on an almost daily basis (see The RNA Interference Resource to learn about the very latest RNAi research and tools). It is not an understatement to say that the field of functional genomics is being revolutionized by RNAi.

See the Ambion link for the references.