Small RNAs and gene expression regulation

Roles for small RNAs in gene expression regulation

Three classes of small RNA can regulate genes by targeting transcripts in the cytoplasm. These are:

- **microRNAs (miRNAs)**, which are hairpin-derived RNAs with imperfect complementarity to targets and that cause translational repression;

- **small interfering RNAs (siRNAs)**, which have perfect complementarity to targets and cause transcript degradation;

- **PIWI-intereacting RNAs (piRNAs)**, which target transposon transcripts in animal germ lines.

RNA interference (RNAi)

- RNAi uses short antisense RNA to to degrade cytoplasmic mRNA by **post-transcriptional gene silencing (PTGS)**.

- RNAi pathways exist also in the nucleus where they can repress target genes at the transcriptional level by guiding epigenetic modification of chromatin: **transcriptional gene silencing (TGS).**

- In the nucleus RNAi pathways can also activate trancritpion by a mechanism known as **RNA activation (RNAa)**

- PTGS and TGS protects against viral infection, prevents transposon mobilization and regulates endogenous genes.

The Nobel Prize in Medicine 2006 to Andrew Fire and Craig Mello

"RNA interference – gene silencing by double-stranded RNA"

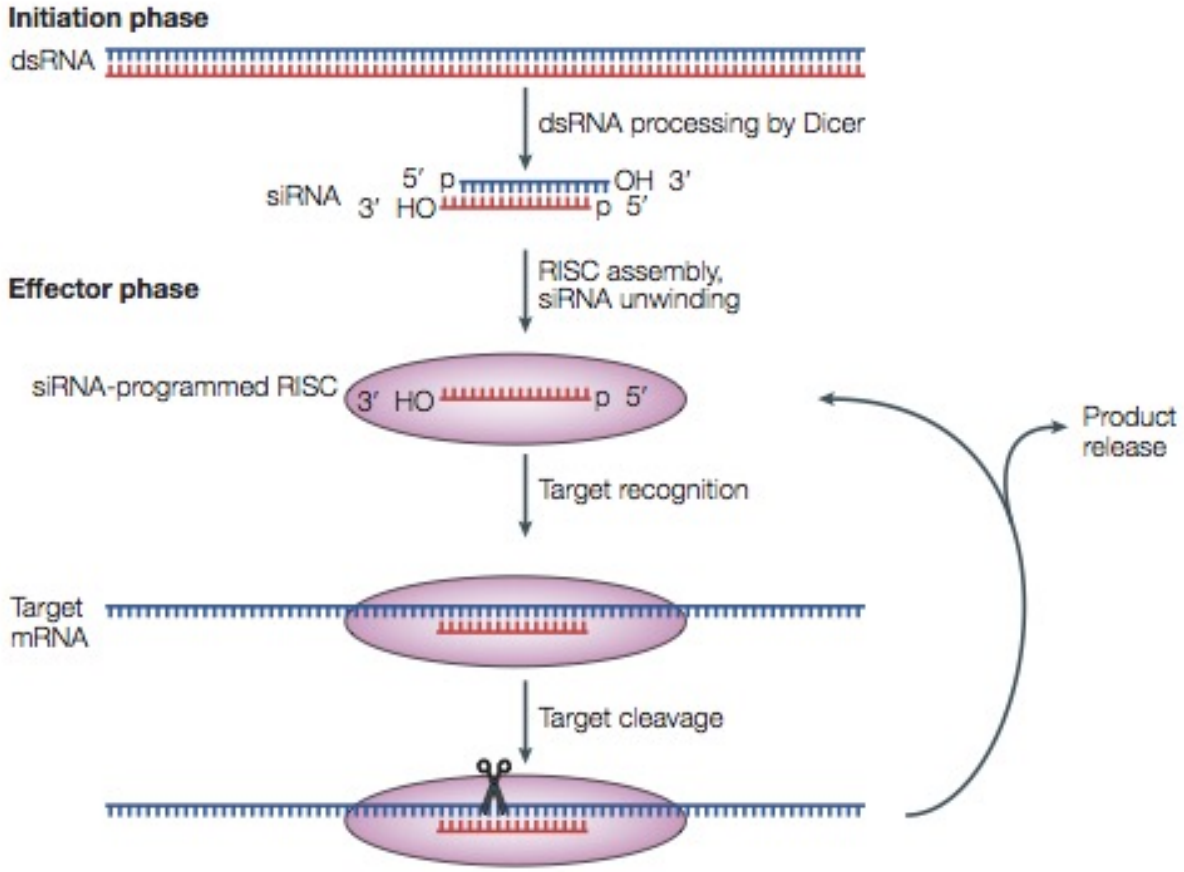
Nature. 1998 Feb 19;391(6669):806-11.

Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.

Fire A[1], Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC.

What is RNA interference (RNAi)?

RNA interference (RNAi) uses dsRNA to guide the destruction of specific mRNAs



The discovery of RNAi

Cosuppression

The overexpression of the **CHS (Chalcone synthase)** gene in petunia leads to lack of flower pigmentation instead of increasing it.

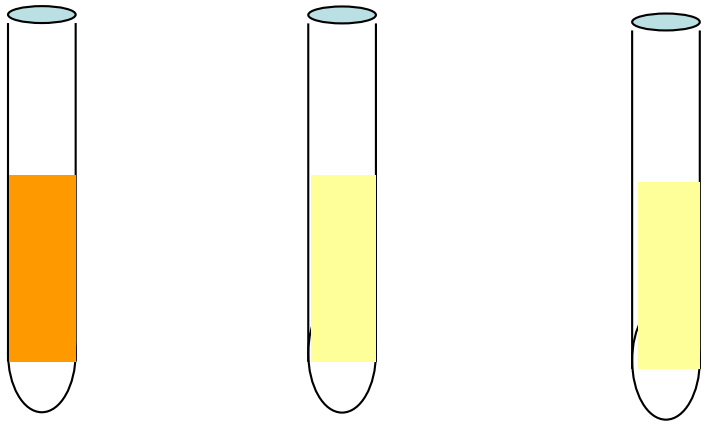
Napoli et al., (1990)
Jorgensen et al., (1996)

The transgene causes the suppression of both the exogenous and endogenous genes

The discovery of RNAi

Quelling

In *Neurospora crassa* the introduction of a transgene causes the silencing of the homologous endogenous gene **albino-1** (al-1) coding for a protein of the carotenoid biosynthetic pathway.



Romano e Macino (1992)
Cogoni et al., (1996)

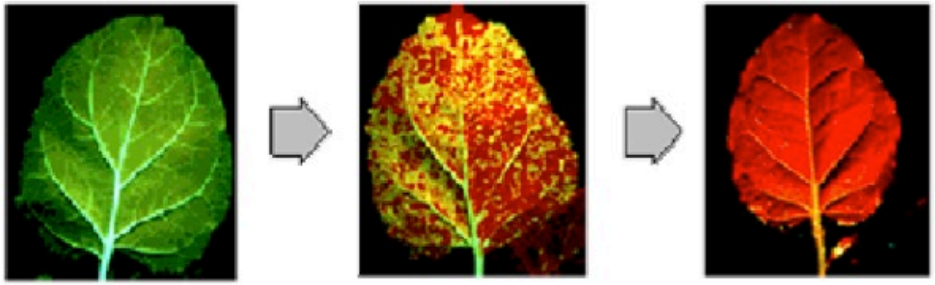
The transgene causes the *suppression* of both the exogenous and endogenous genes

The discovery of RNAi

Virus-induced gene silencing (VIGS)

Viral infection can silence both viral genes and endogenous genes sharing complementarity

PVX-GFP



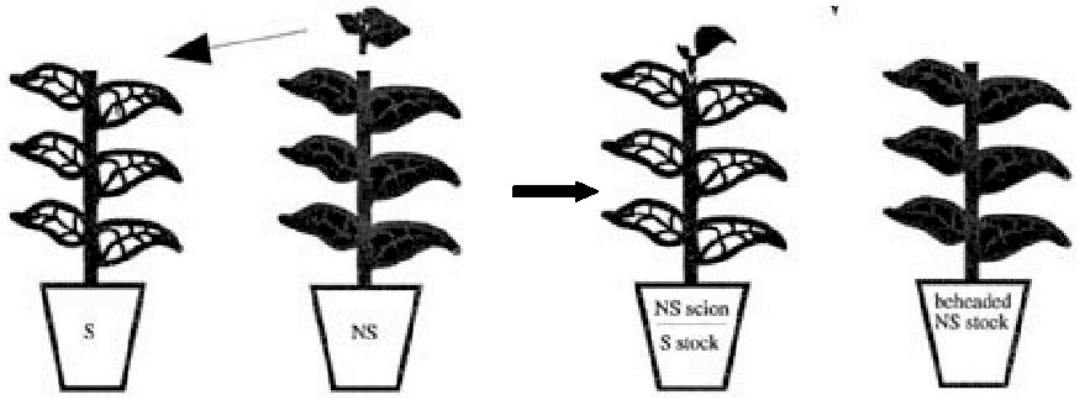
M. Teresa Ruiz, Olivier Voinnet, and David C. Baulcombe.The Plant Cell, Vol. 10, 937–946, June 1998

GFP expressing transgenic plant infected with GFP expressing potato virus X (PVX)

A mobile signal transmits gene silencing

In *Neurospora,* gene silencing can be transferred from nucleus to nucleus into heterocarotic strains (Cogoni et al., 1996)

In **plants**, PTGS is induced by a side shoot that is grafted onto a silent plant (Palauqui et al., 1997)



Nature 391, 806 - 811 (19 February 1998)

Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*

Andrew Fire*, SiQun Xu*, Mary K. Montgomery*, Steven A. Kostas*†, Samuel E. Driver‡ & Craig C. Mello‡

RNA interference (RNAi)

2006 RNA interference

- Discovered by accident
- An extremely useful tool for researchers

Shared the Nobel Prize in Physiology or Medicine for their work on RNA interference in the nematode worm *C. elegans*, which they published in 1998

Accidental Discovery of RNAi

... Goal: Silence mRNAs with an antisense RNA

... The unc-22 gene encodes a myofilament protein.

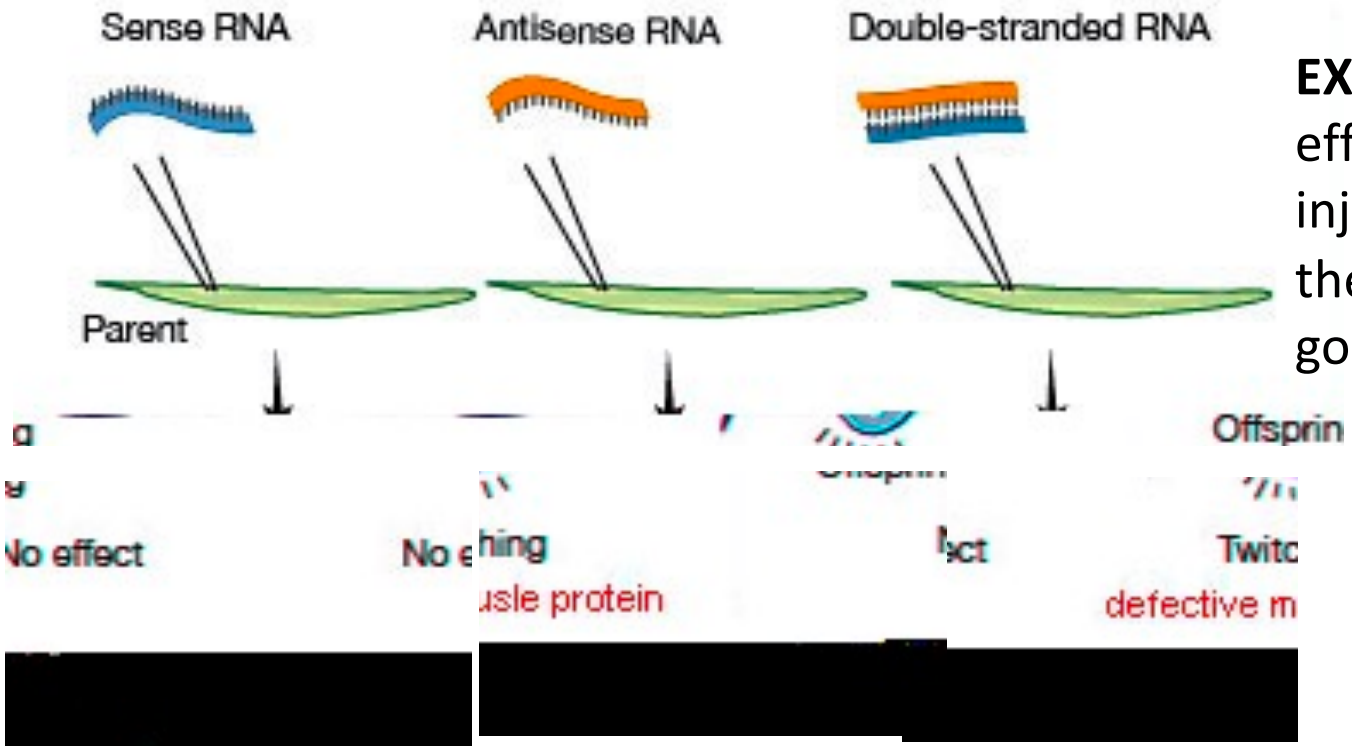
... The reduction of unc-22 activity is known to cause severe contractions.

Accidental Discovery of RNAi

... Goal: Silence mRNAs with an antisense RNA

... The unc-22 gene encodes a myofilament protein.

... The reduction of unc-22 activity is known to cause severe contractions.



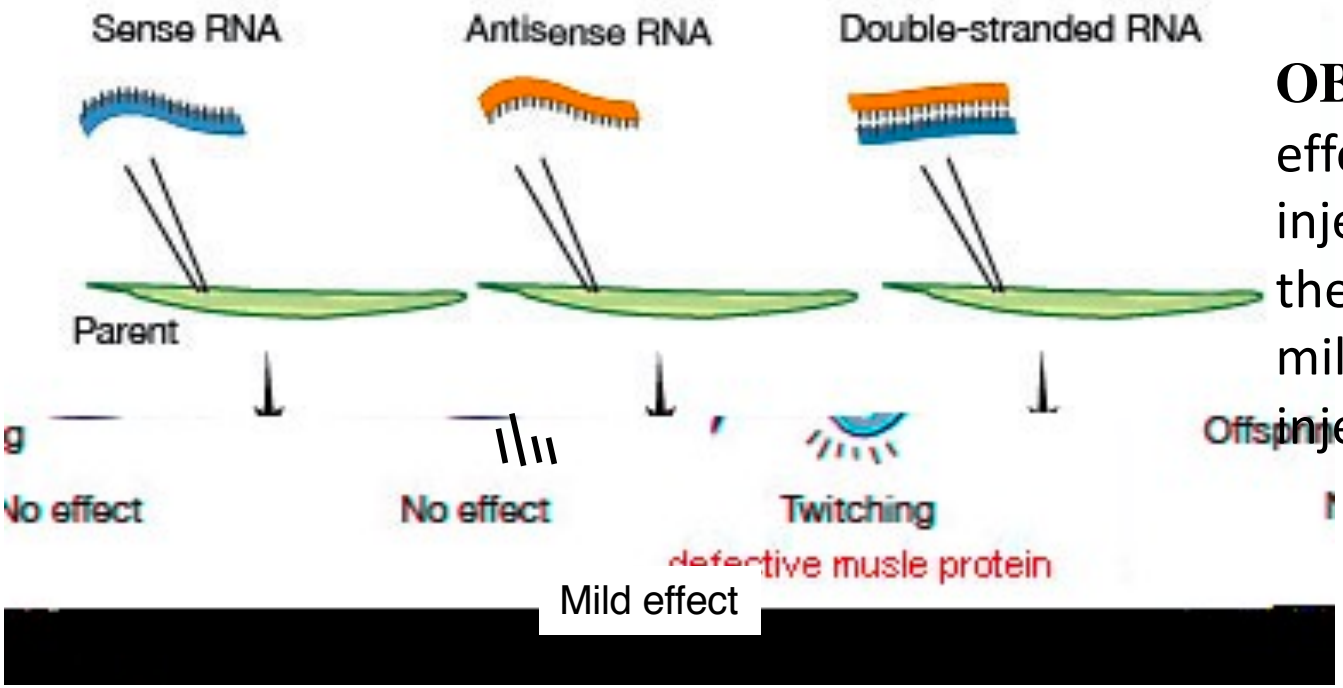
EXPECTED: The phenotypic effect is observed after the injection of ssRNA antisense to the unc-22 mRNA into the gonads of C. elegans..

Accidental Discovery of RNAi

... Goal: Silence mRNAs with an antisense RNA

... The unc-22 gene encodes a myofilament protein.

... The reduction of unc-22 activity is known to cause severe contractions.



OBSERVED: The phenotypic effect is observed after the injection of dsRNA (unc-22) into the gonads of C. elegans. A very mild effect is observed after the injection of ssRNA antisense.

In 1998, A. Fire and C. Mello (Nobel Prize in Medicine 2006) working on C. elegans discovered that not only by inserting a transgene, but also by inserting the gene in antisense or directly a double-stranded RNA (dsRNA) activates a gene silencing of the corresponding gene.

dsRNA induces PTGS

- Guo and Kemphues (1995), use antisense RNA to study the function of the par-1 gene in *Caenorhabditis elegans*. As expected, the injection of antisense abolishes par-1 expression.

- Fire and Mello (1998), inject dsRNA into *Caenorhabditis elegans* and get a much more effective mute than with antisense alone.

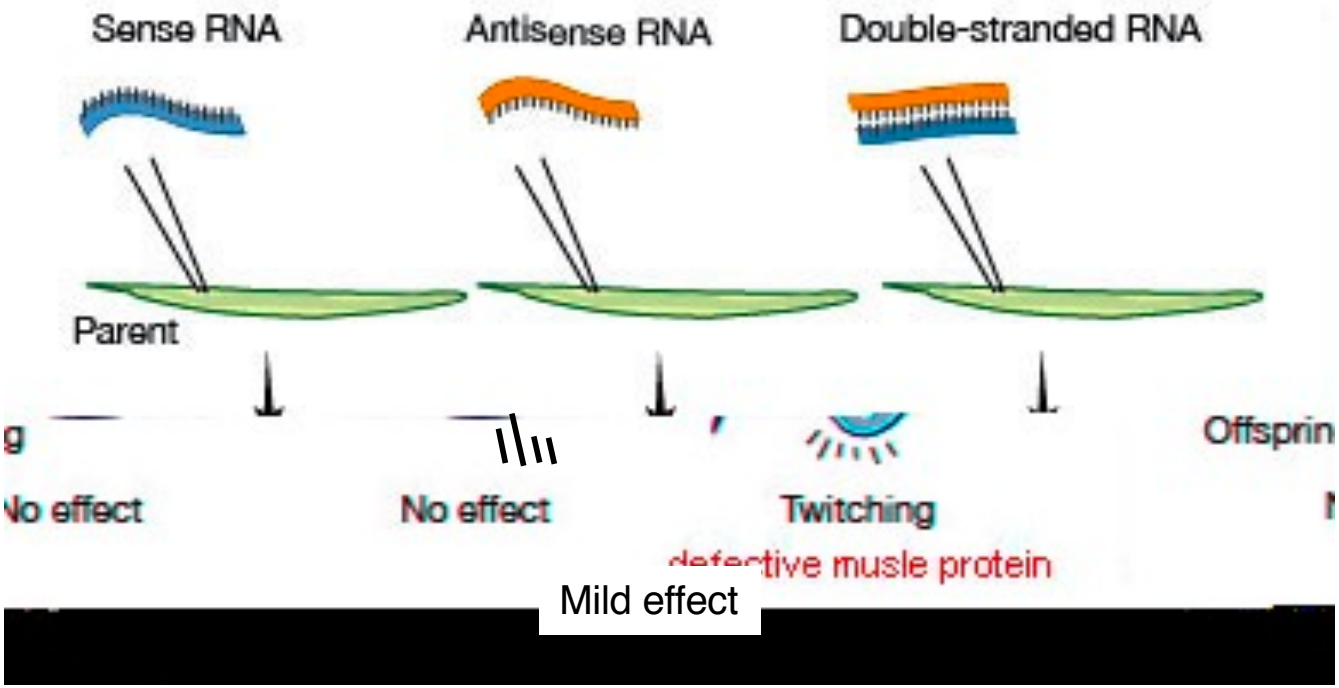
The observation came from what was intended by the authors to be a negative control!

Accidental Discovery of RNAi

... **Goal:** Silence mRNAs with an antisense RNA

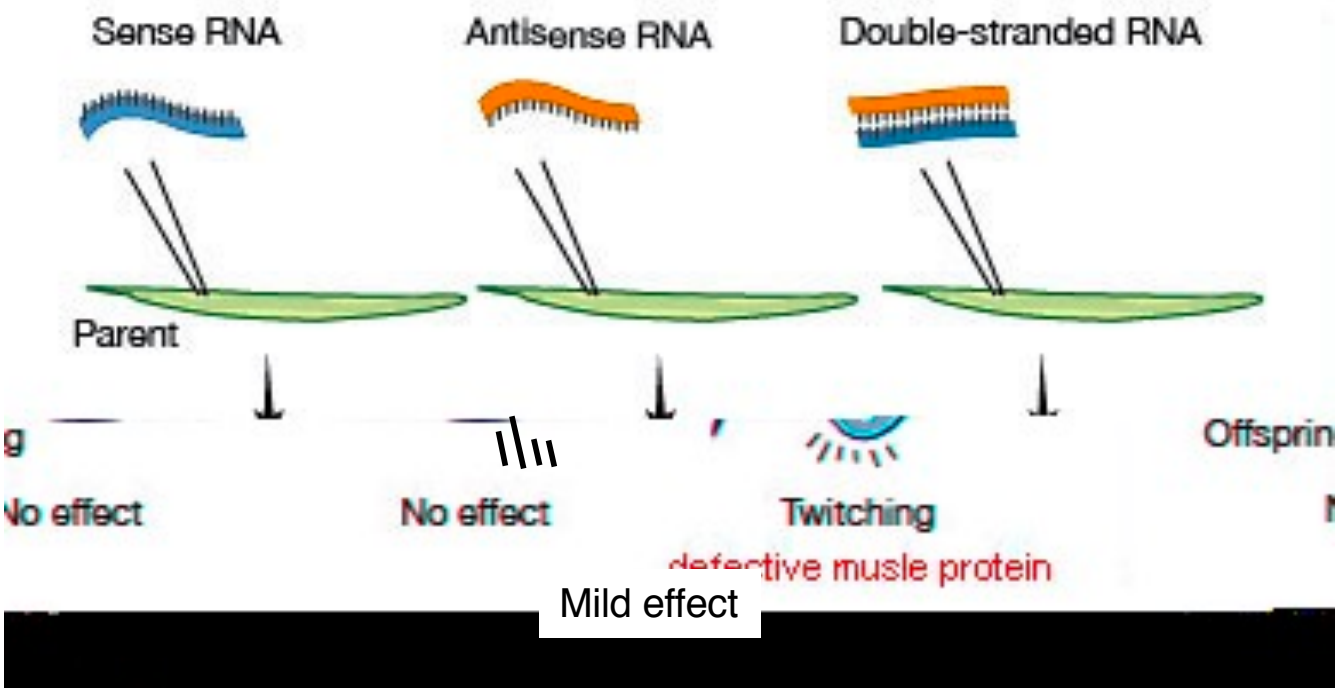
... The unc-22 gene encodes a myofilament protein.

... The reduction of unc-22 activity is known to cause severe contractions.



Scoperta accidentale dell'RNAi

- Goal: silenziare l'mRNAs con un RNA antisenso
- Il gene *unc-22* codifica per una proteina del miofilamento.
- La riduzione dell'attività di *unc-22* è nota causare severe contrazioni.



The term RNA interference was coined on this occasion. C. elegans was a useful experimental system because the evolutionary origin of all the cells of this organism is known and it is possible to inject RNA into embryos at early stages and observe the changes compared to the model during development.

Experiments of Fire and Mello in 1998

•Injection of dsRNA (sense and anti-sense strands) in C. elegans

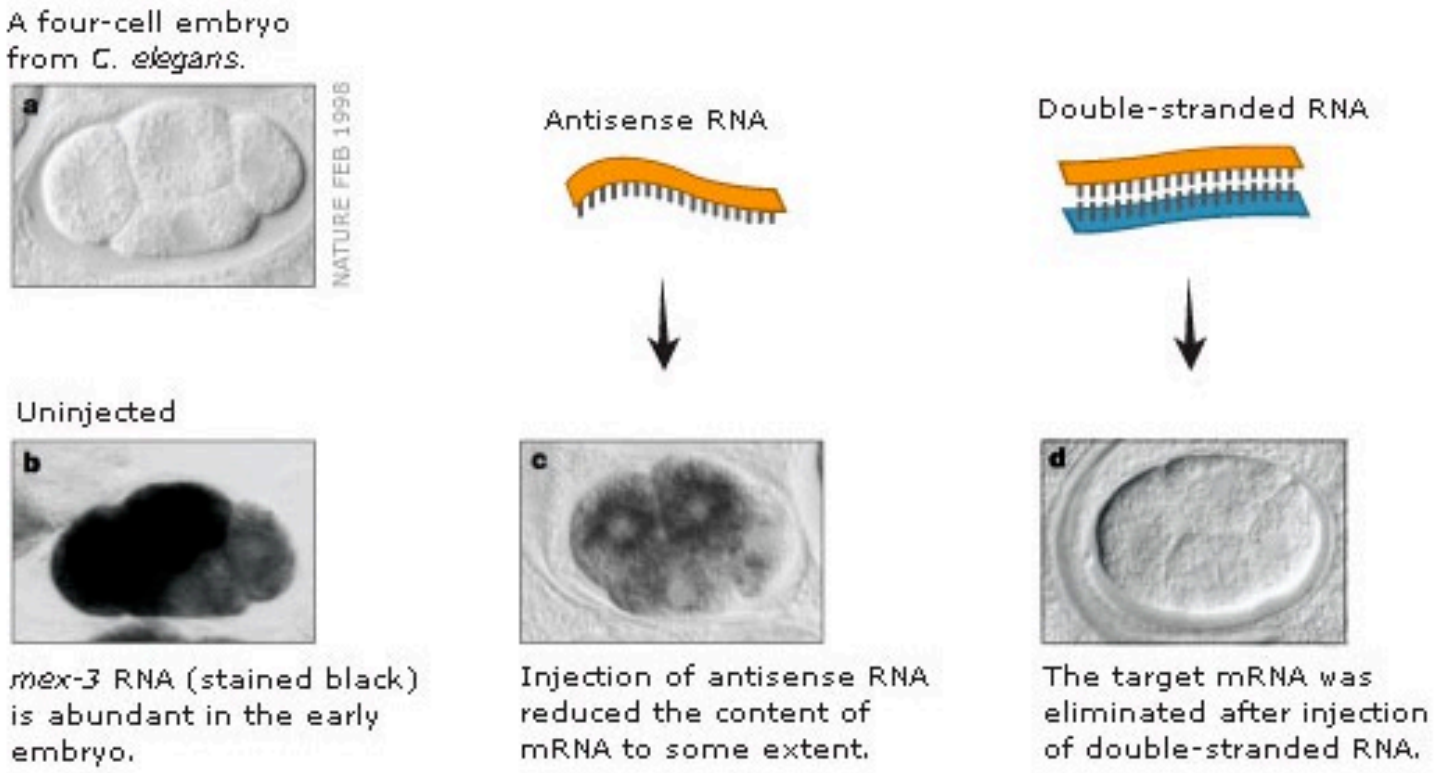
•Silencing more efficient than the injection of the anti-sense alone

•The dsRNA must include the exons; introns and the promoter do not work in silencing

•The silencing is due to the degradation of the target mRNA

•Only a few dsRNA molecules per cell were sufficient to silence the expression of the homologous gene. The phenomenon is transferred from one cell to another

•RNAi is applicable to many different transcripts



Important characteristics of RNA-mediated silencing

PTGS, quelling and RNAi have common critical components:

1.They are induced by dsRNA

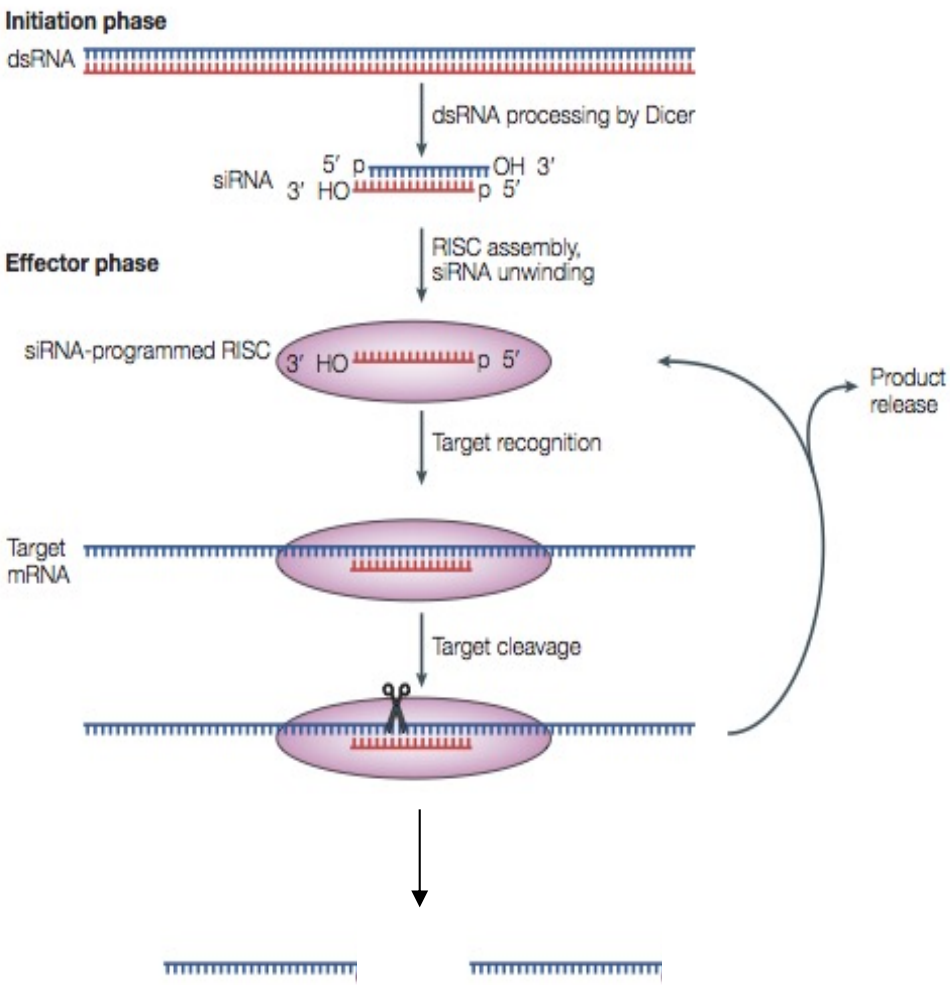
2.The target is degraded in a sequence-specific manner

3.The degradation machinery requires a set of proteins that are similar in structure and function in most organisms.

The general pathway of RNAi

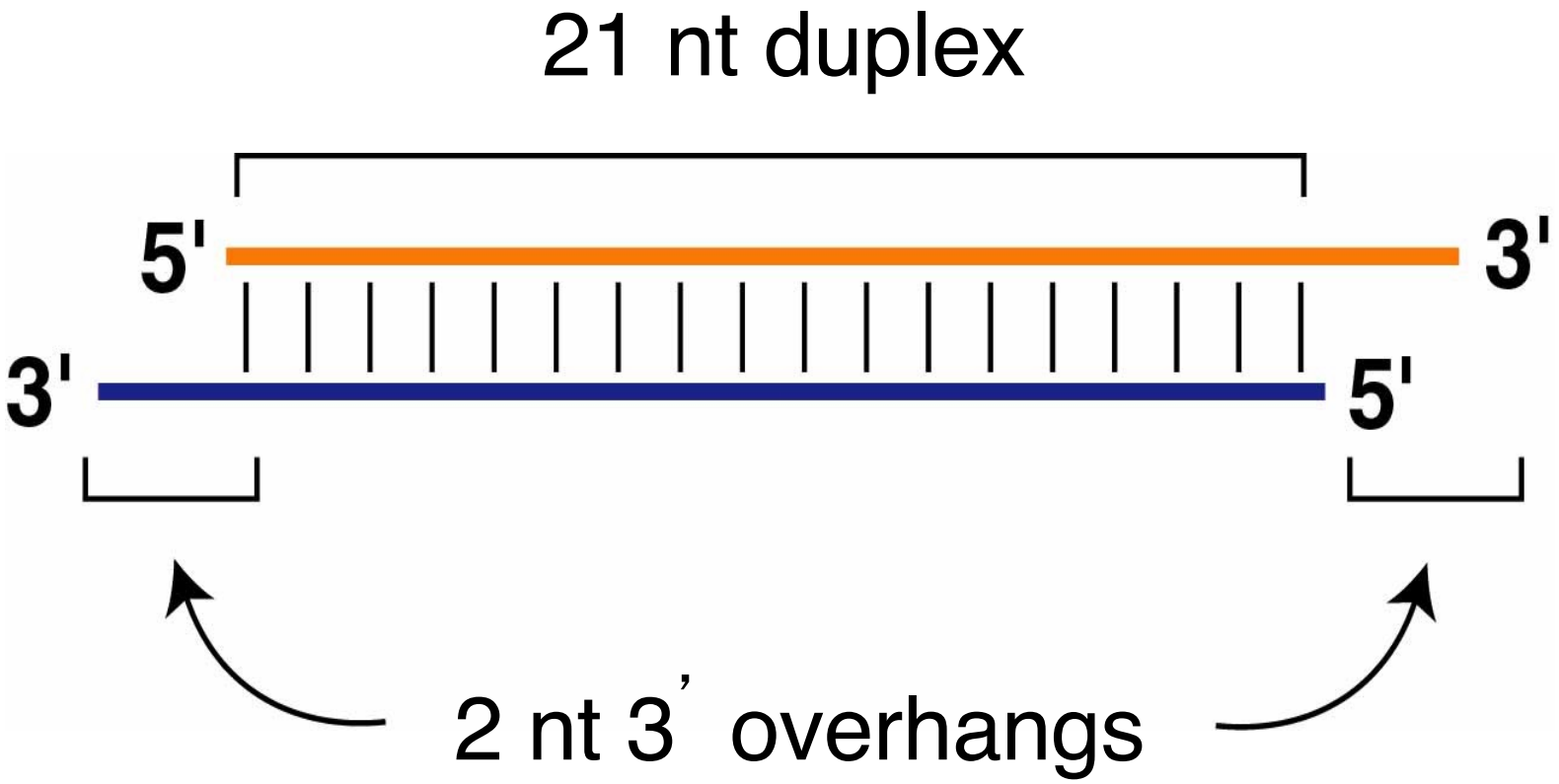
Initiation phase: the ribonuclease-III enzyme Dicer cleaves (ds)RNA molecules into 21–23-nt short interfering (si)RNA duplexes. siRNAs bear 5′-phosphate groups and 2-nt 3′-overhangs, both of which are important for subsequent siRNA-induced silencing complex (siRISC) assembly.

Effector phase: the siRNA becomes unwound and assembles into RISC. The activated effector complex recognizes the target by siRNA–mRNA base pairing, and then cleaves the mRNA strand with its endoribonuclease activity.

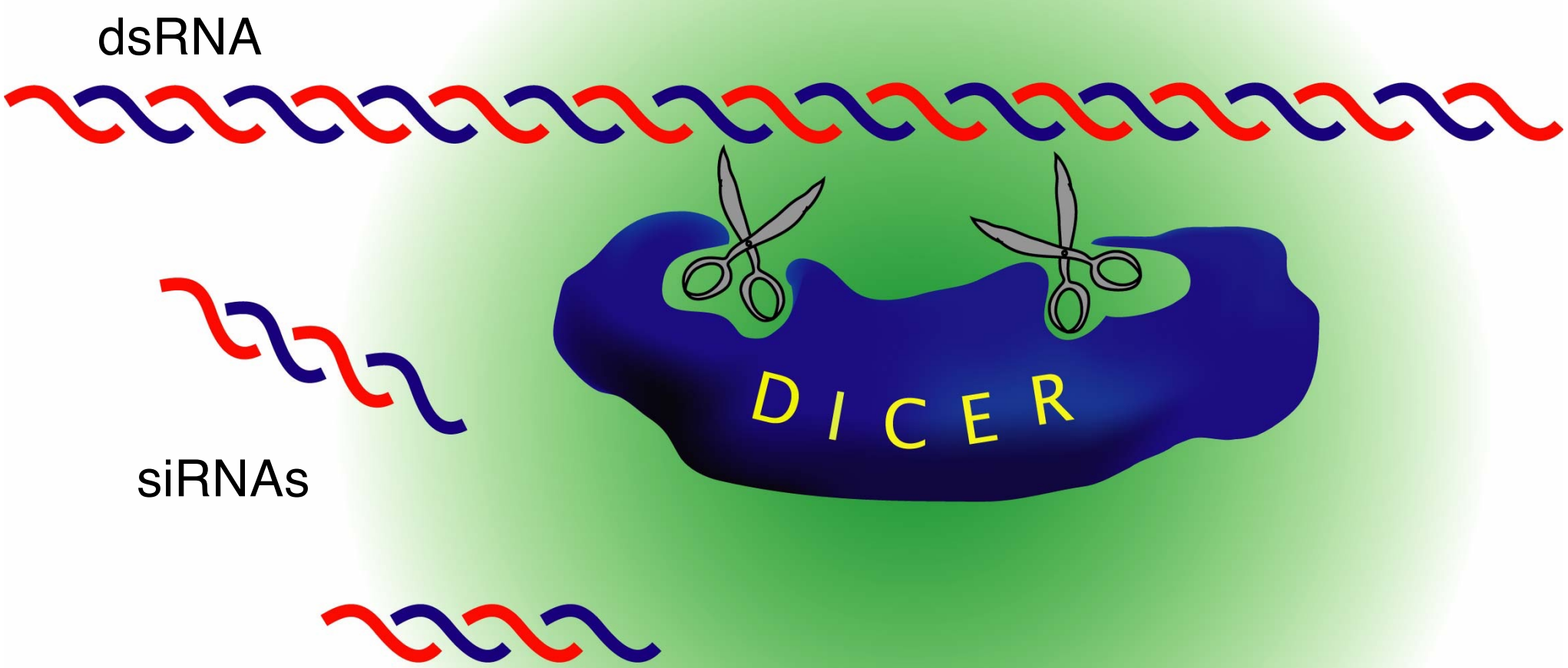


RNA degradation

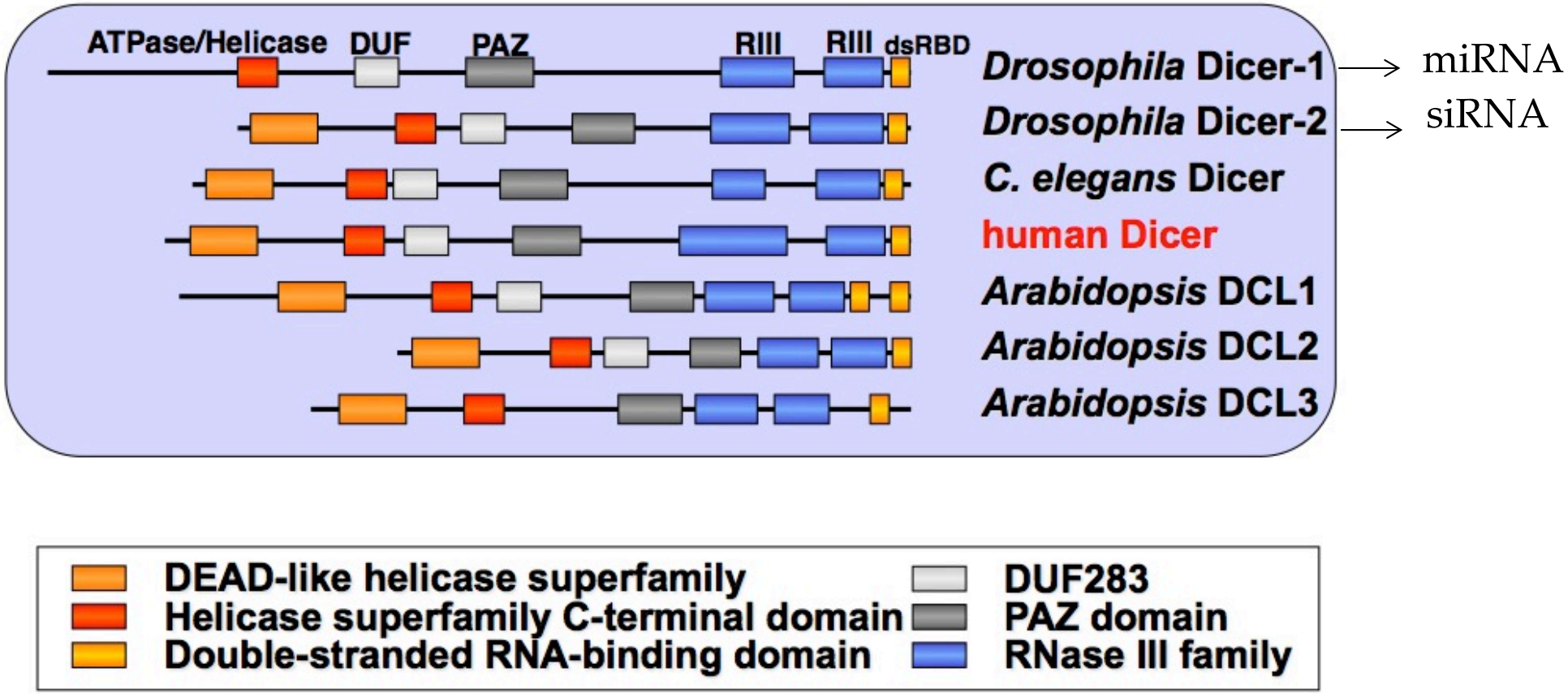
siRNAs have a well-defined structure



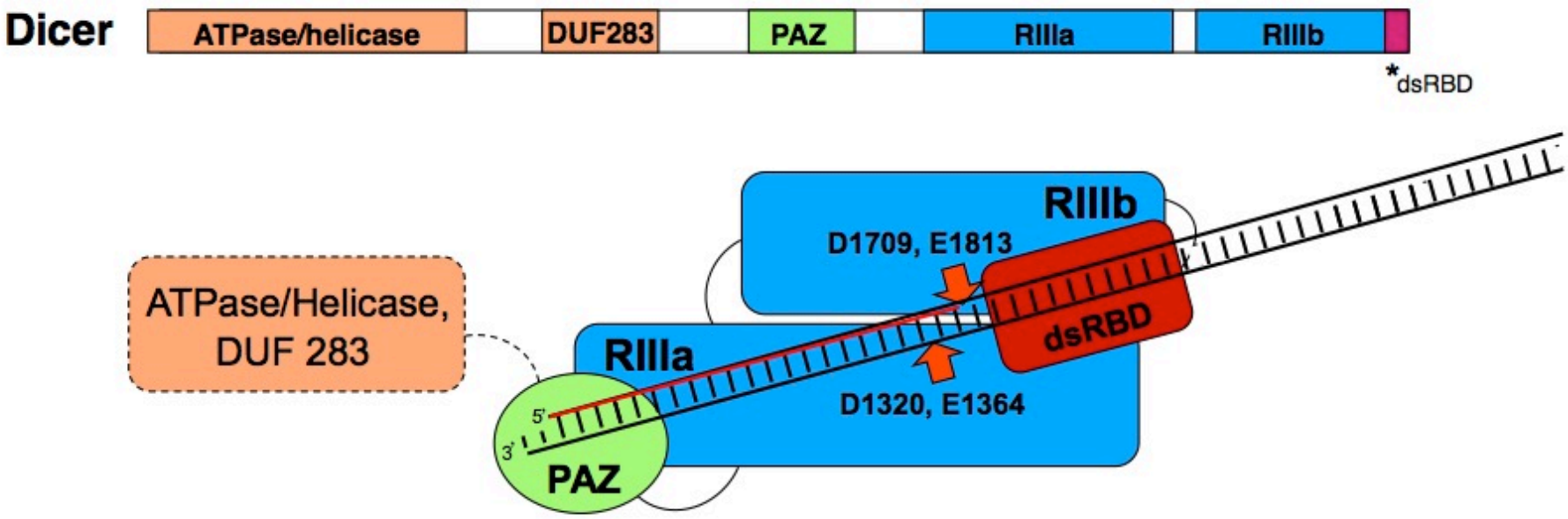
Dicer has two RNase III domains



Dicer: multi-domain nucleases involved in RNAi and microRNA biogenesis



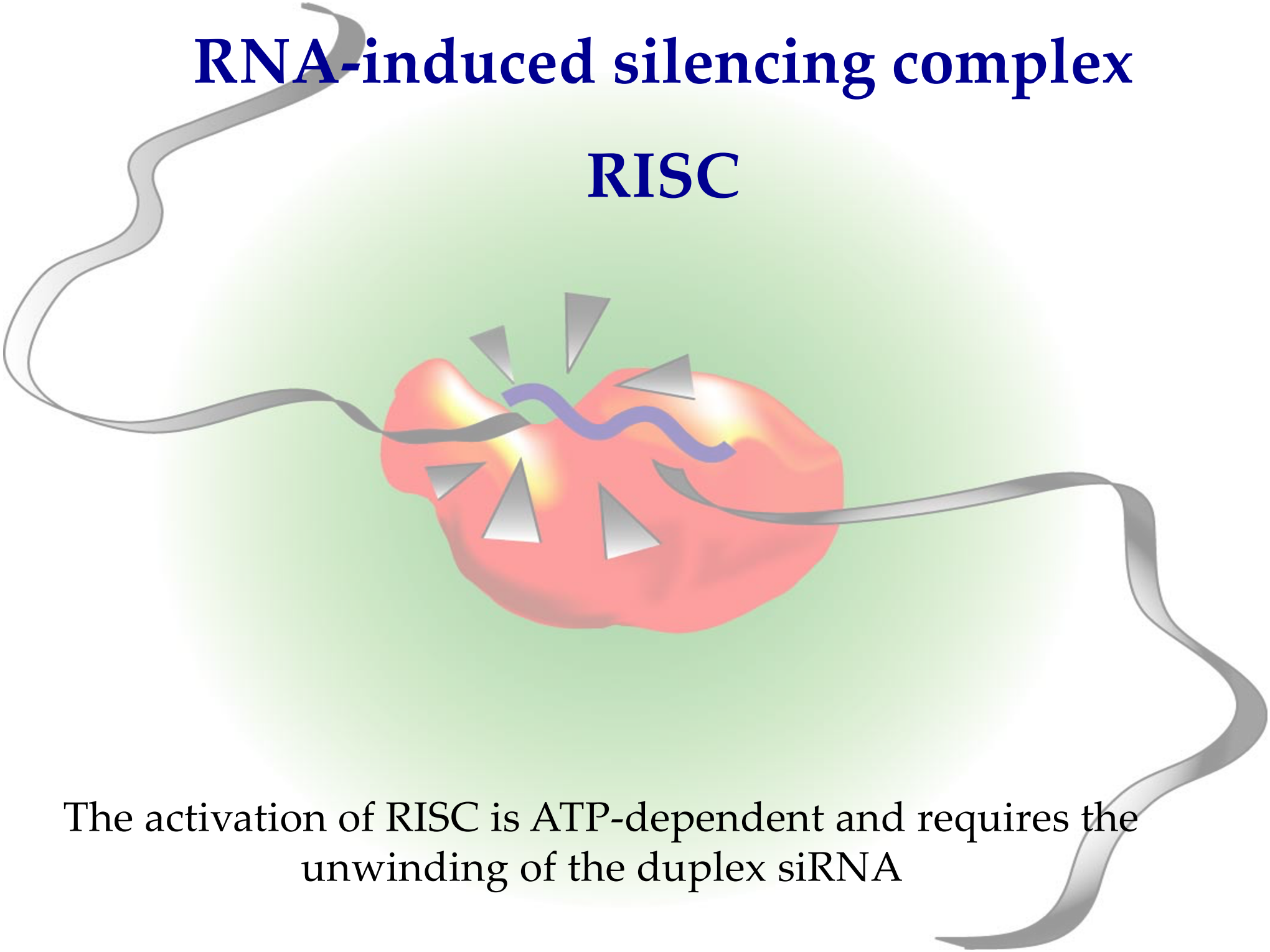
dsRNA processing by Dicer



- Dicer functions as a *monomer* (i.e., intra-molecular dimer).
- PAZ domain recognizes the 3'- overhang end.
- Dicer has a *single* processing center, with two independent catalytic sites.
- Each RNase III domain cuts one RNA strand in a *polar* way

RNA-induced silencing complex

RISC



The activation of RISC is ATP-dependent and requires the unwinding of the duplex siRNA

Genes involved in PTGS

- **RDE-1** (C. elegans), **QDE-**2 (N. crassa), **AGO-1** (A. thaliana) -> **PPD proteins**
  -They share two conserved domains: PAZ and PIWI
  - Mutants and have developmental defects and RNAi is abolished
  - Human homologue **eIF2C2** (**AGO2**)

- **RDE-4** (*C. elegans*) -> **dsRNA binding domain**

- **MUT-7** (*C. elegans*) -> **RNaseD domain**
-
- **QDE-3** (*N. crassa*) , MUT-6 (*C.elegans*) -> **RNA helicase**

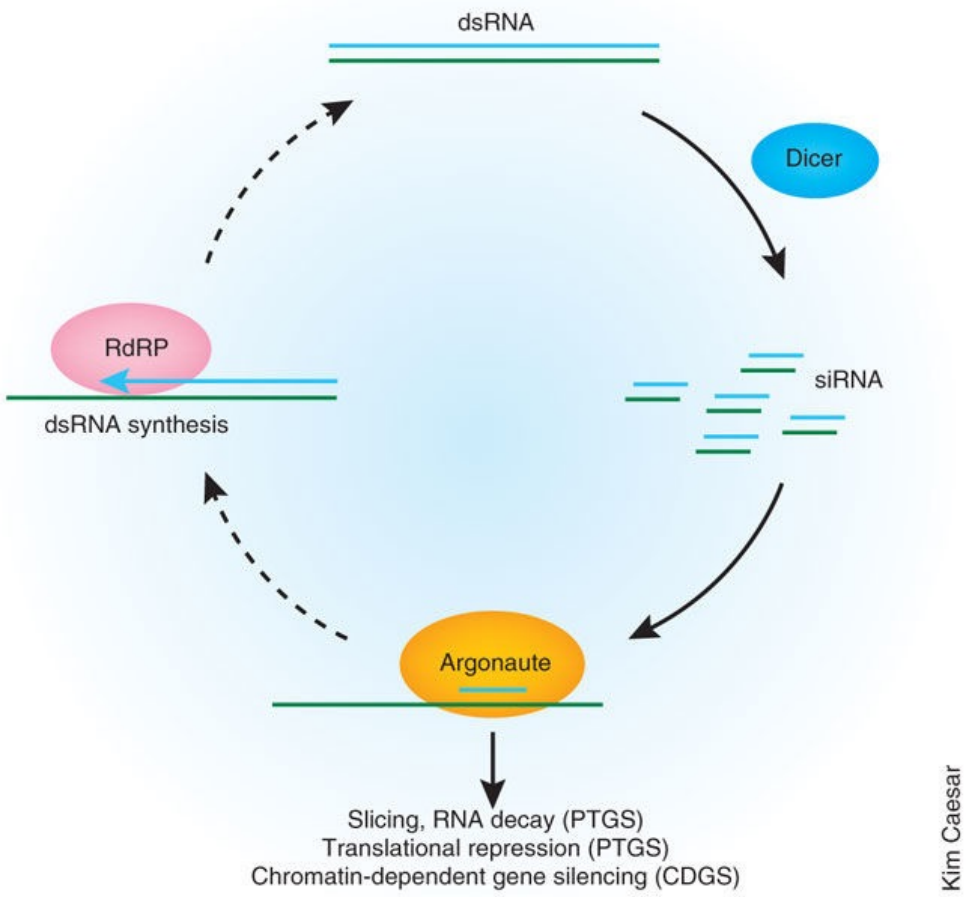
- **QDE-1** (*N. crassa*) , EGO-1 (*C.elegans*) -> **RdRP (RNA dependent RNA polymerase)**

RNA Dependent RNA Polymerase (RdRP)

RdRP found in plants, C. elegans and Drosophila m. but not in mammals

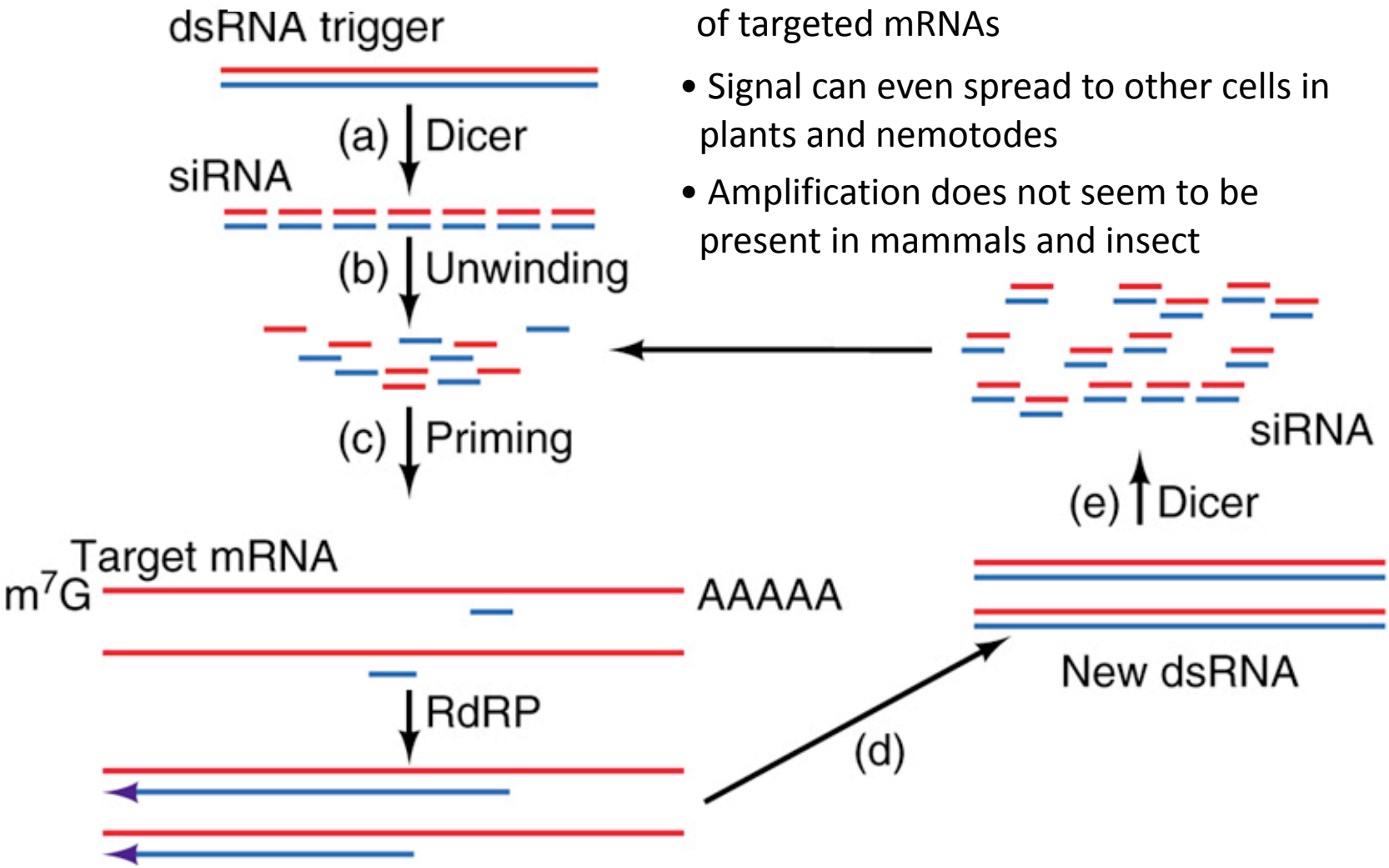
Copies one strand and resynthesizes dsRNA ex novo. The siRNAs function as primers on the messenger and the dsRNA is amplified

Can explain the duration of the phenomenon because it is responsible for the amplification of the siRNAs



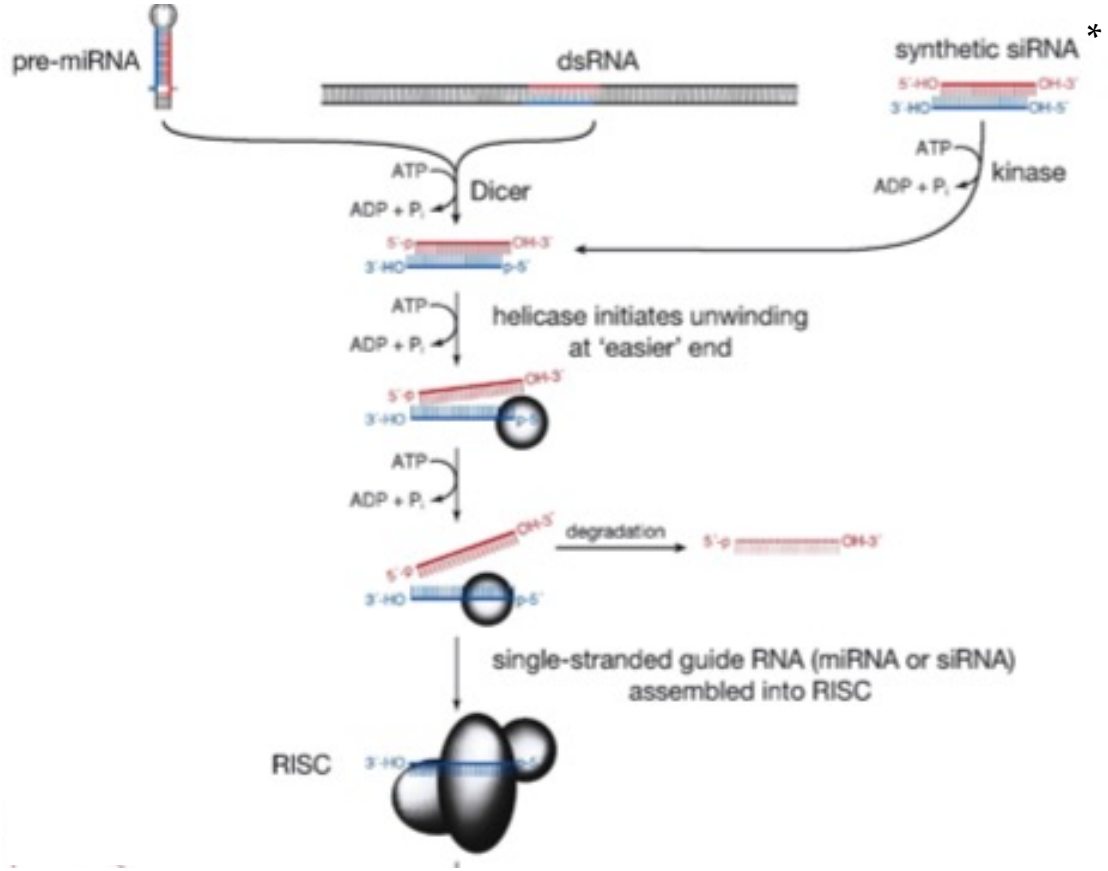
In Some Organisms, siRNA Signal Is Amplified and Spread

- New siRNAs appear against other regions of targeted mRNAs
- Signal can even spread to other cells in plants and nemotodes
- Amplification does not seem to be present in mammals and insect



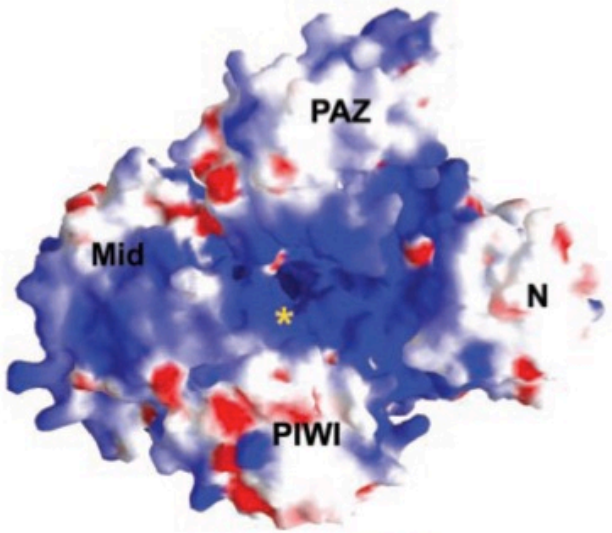
Asymmetry in the Assembly of the RNAi Enzyme Complex

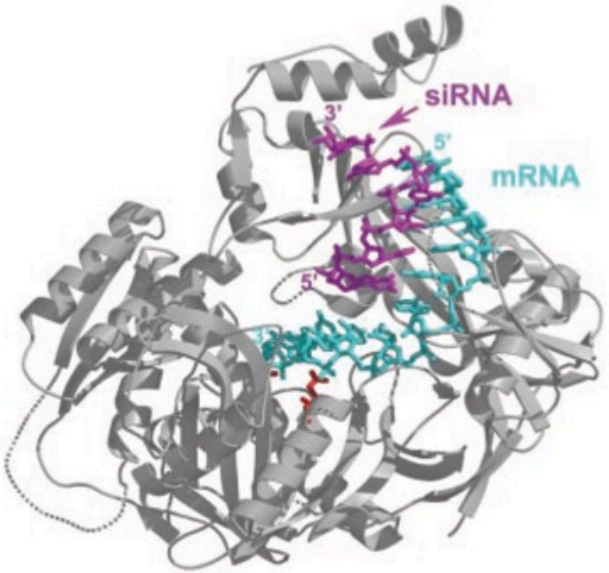
The two strands of an siRNA duplex are not equally eligible for assembly into RISC. Rather, both the absolute and relative stabilities of the base pairs at the 5′-ends of the two siRNA strands determine the degree to which each strand participates in the RNAi pathway.

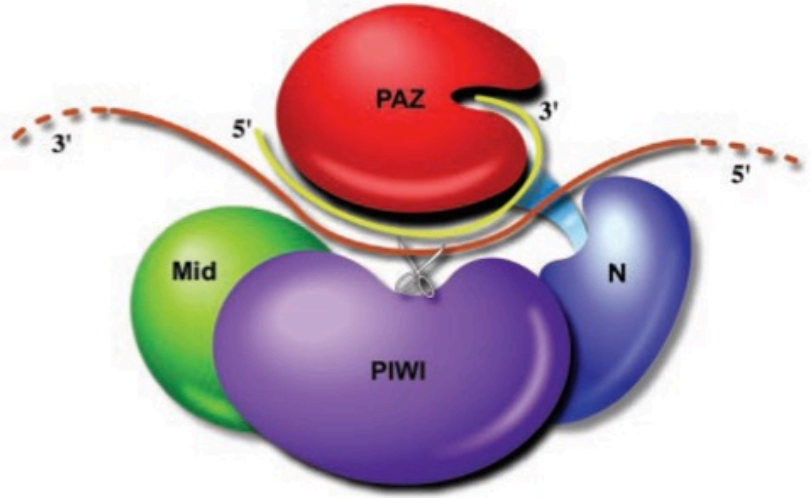


* Nucleotide can be chemically modified to get higher stability and strand specificity

The PIWI domains of Ago proteins harbour the endonuclease activity of RISC

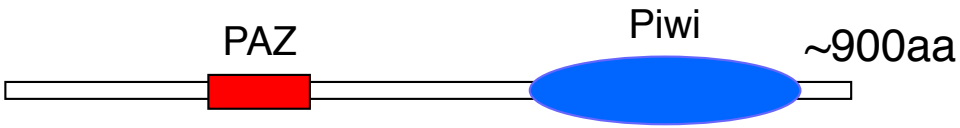


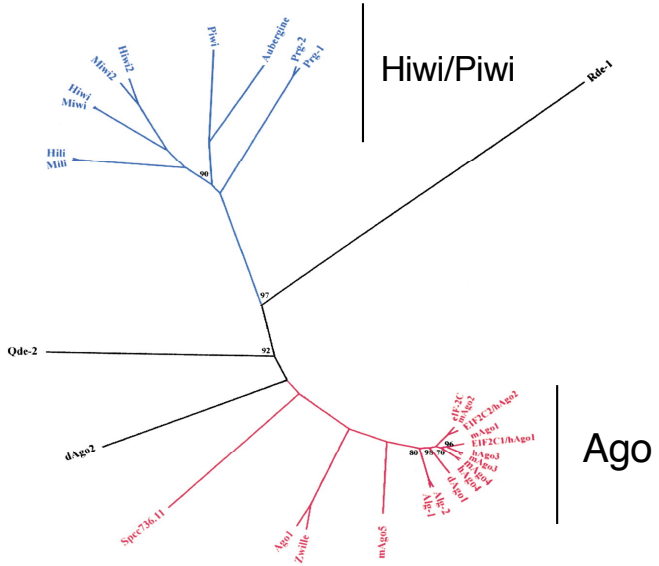




- **PAZ domain** anchors the 3'-ends of small RNAs
- **PIWI domain** is structurally similar to RNase H and can cleave complementary RNAs
- **The MID domain** anchors the 5' end of the small RNA

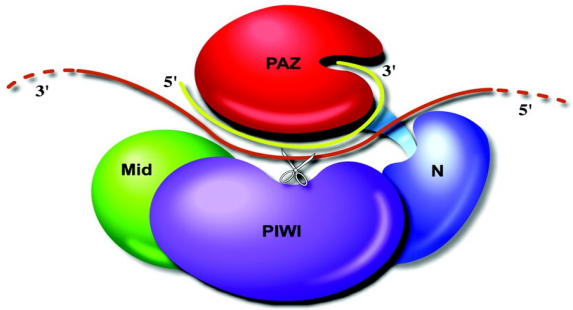
Family of Argonaute proteins





Carmell & Hannon, 2002

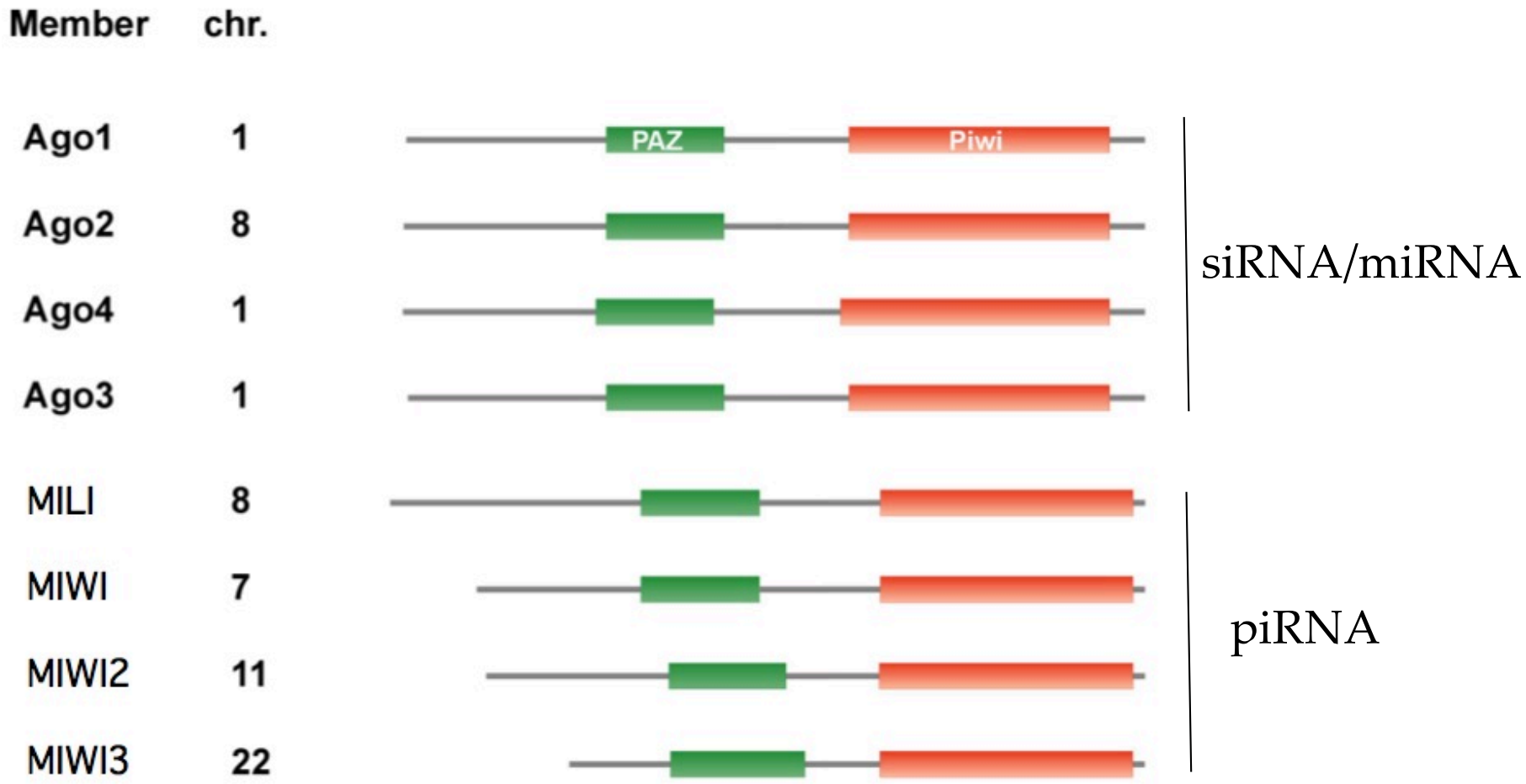
Ago proteins in different organisms	
Mammals	8
Drosophila	5
C. elegans	27
S.pombe	1
Archea	1



Role in RNAi/miRNAs:
- components of RISC/miRNPs
- bind siRNAs/miRNAs

Mammals:
- 4 Argonautes (Ago1-4)
- Ago2 is a "Slicer" (Piwi ~ RNaseH)

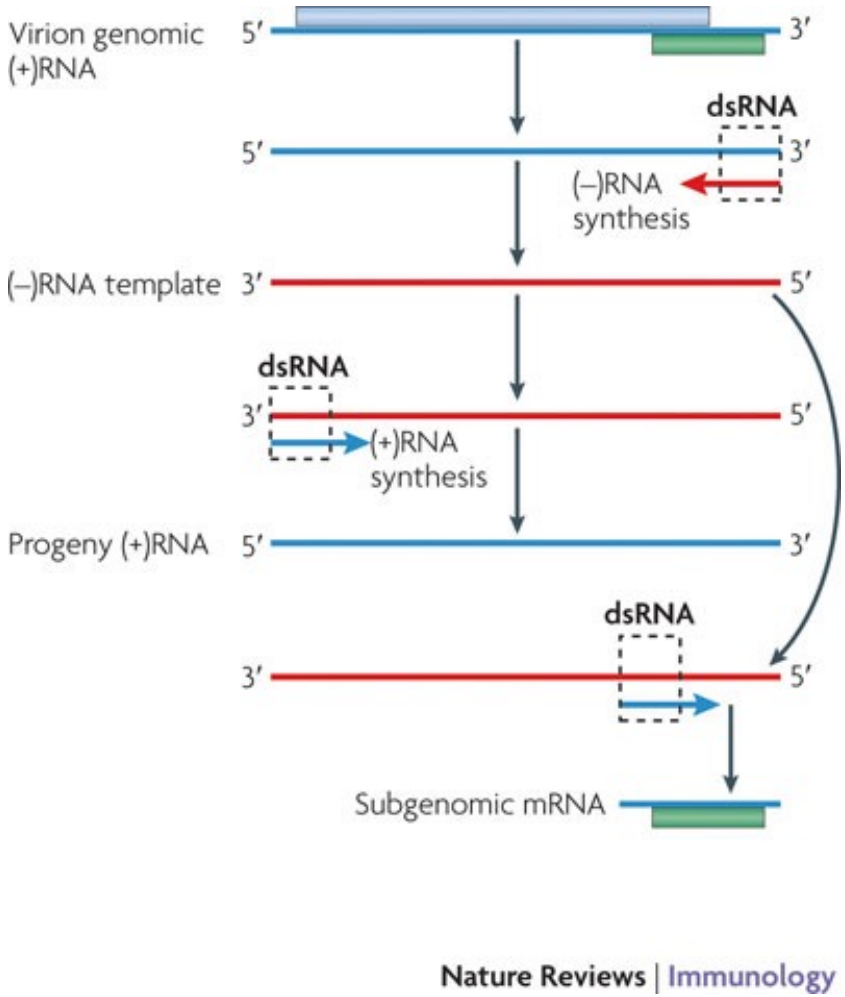
The human Argonaute protein family



RNA-based antiviral immunity

- The replication cycle of a positive-strand RNA virus includes multiple steps that yield double-stranded RNA. In addition, many organisms (plants, nematodes, insects) use viral RNA as template for **RNA-dependent RNA polymerase (RdRP)** to produce dsRNA.

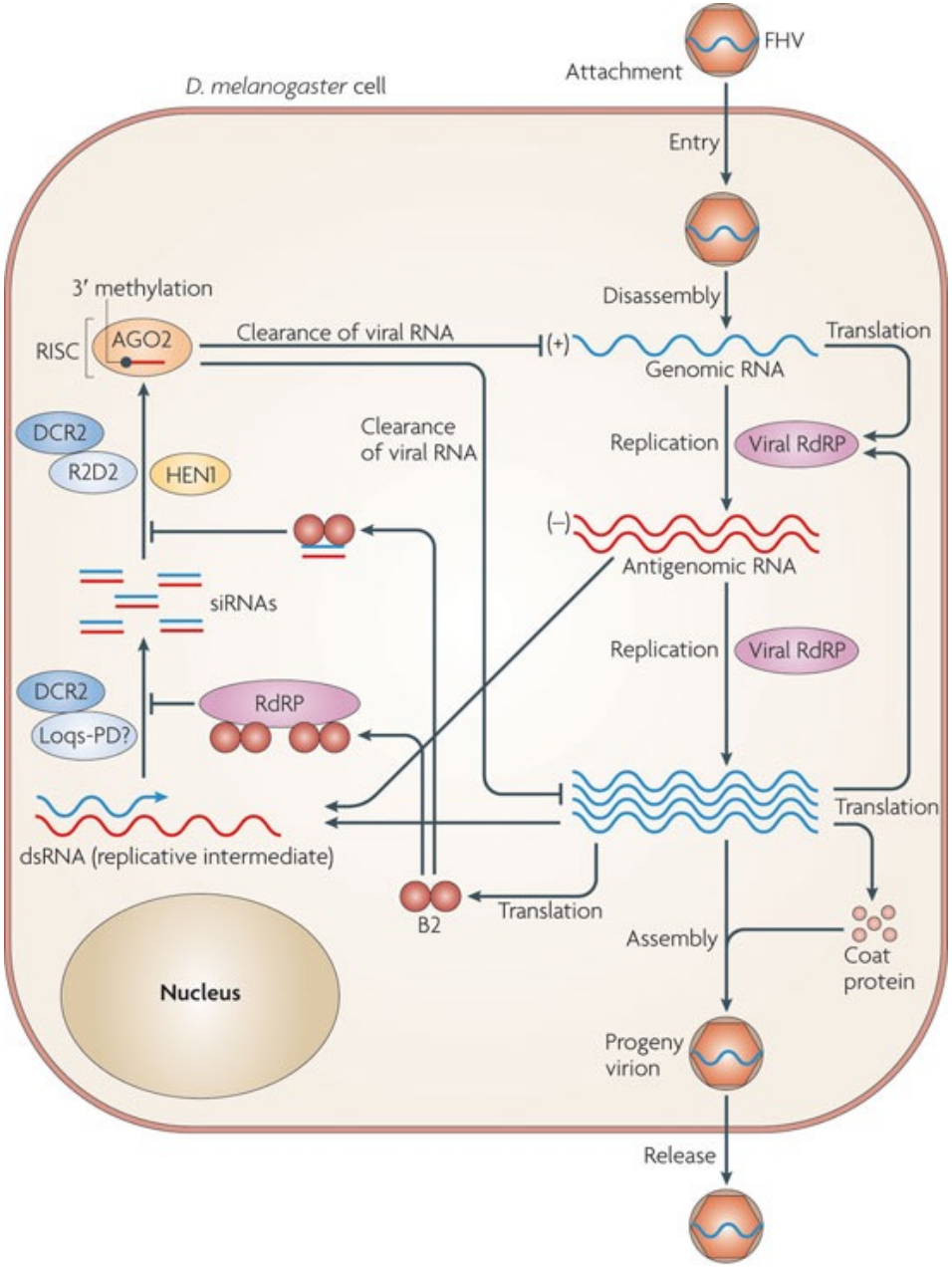
- RsdRP are also involved the amplification of antiviral response.



RNA-based antiviral immunity

D. melanogaster

Following entry and uncoating of flock house virus (FHV) virions, the genomic positive-strand RNA ((+)RNA) serves as a template for the synthesis of antigenomic negative-strand RNA ((−)RNA). The resulting dsRNA formed between the 5′-terminal nascent progeny (+)RNA and the (−)RNA template is recognized by Dicer 2 (DCR2) and cleaved into siRNAs, thereby triggering **RNA-based antiviral immunity**. As a counter-defence, FHV encodes a viral suppressor of RNA silencing, the B2 protein.



Published online: November 4, 2016

Article

SOURCE DATA | TRANSPARENT PROCESS | OPEN ACCESS

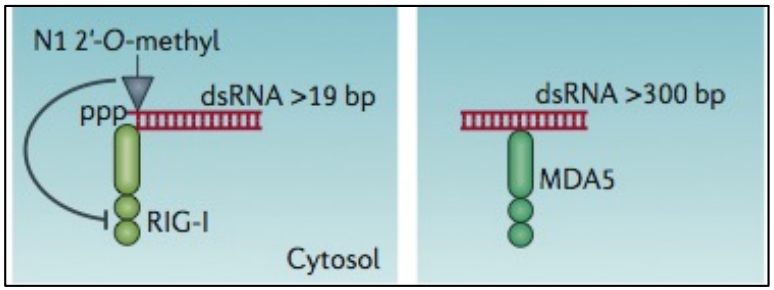
THE EMBO JOURNAL

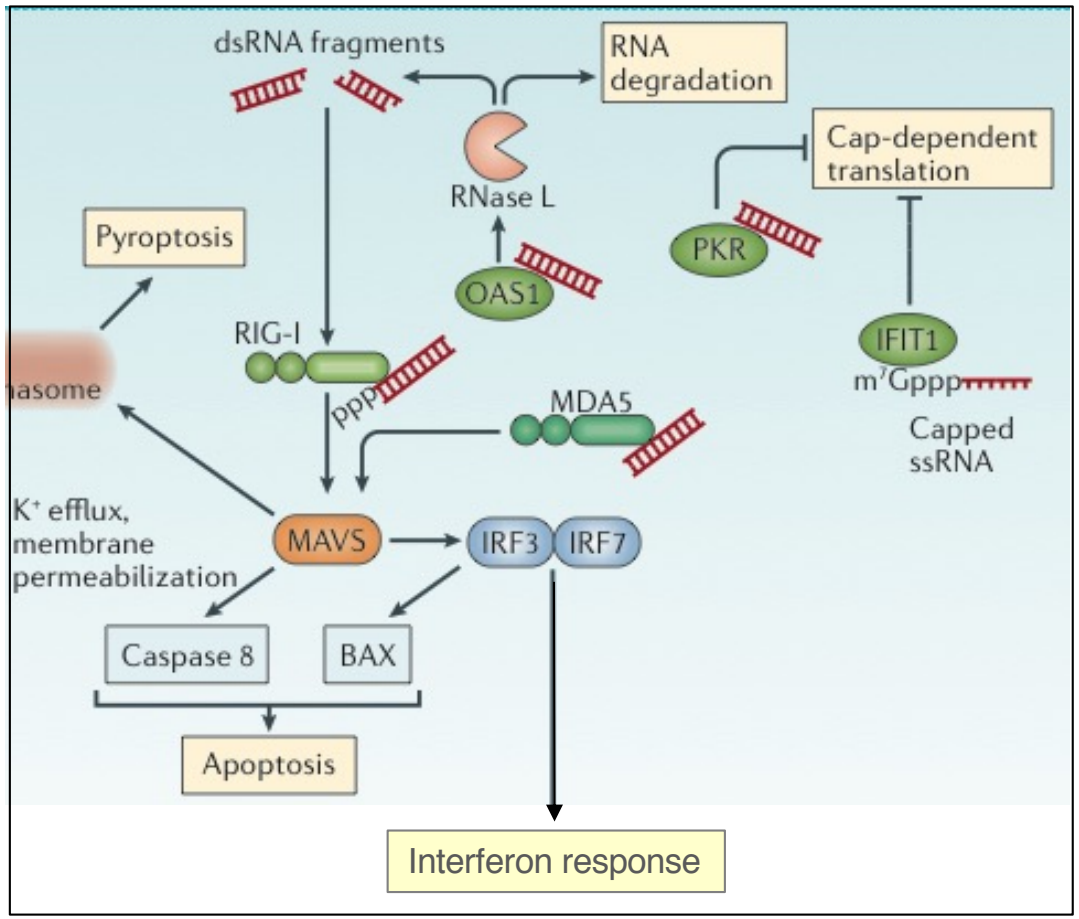
Inactivation of the type I interferon pathway reveals long double-stranded RNA-mediated RNA interference in mammalian cells

Pierre V Maillard[1,**], Annemarthe G Van der Veen[1], Safia Deddouche-Grass[1,†], Neil C Rogers[1], Andres Merits[2] & Caetano Reis e Sousa[1,*]

Transfection with long dsRNA specifically vaccinates IFN-deficient cells against infection with viruses bearing a homologous sequence.

Long dsRNAs are toxic for mammalian cells





dsRNAs must be long 21 nt in order to minimize interferon response in mammalian cells

2017 nature

LETTER

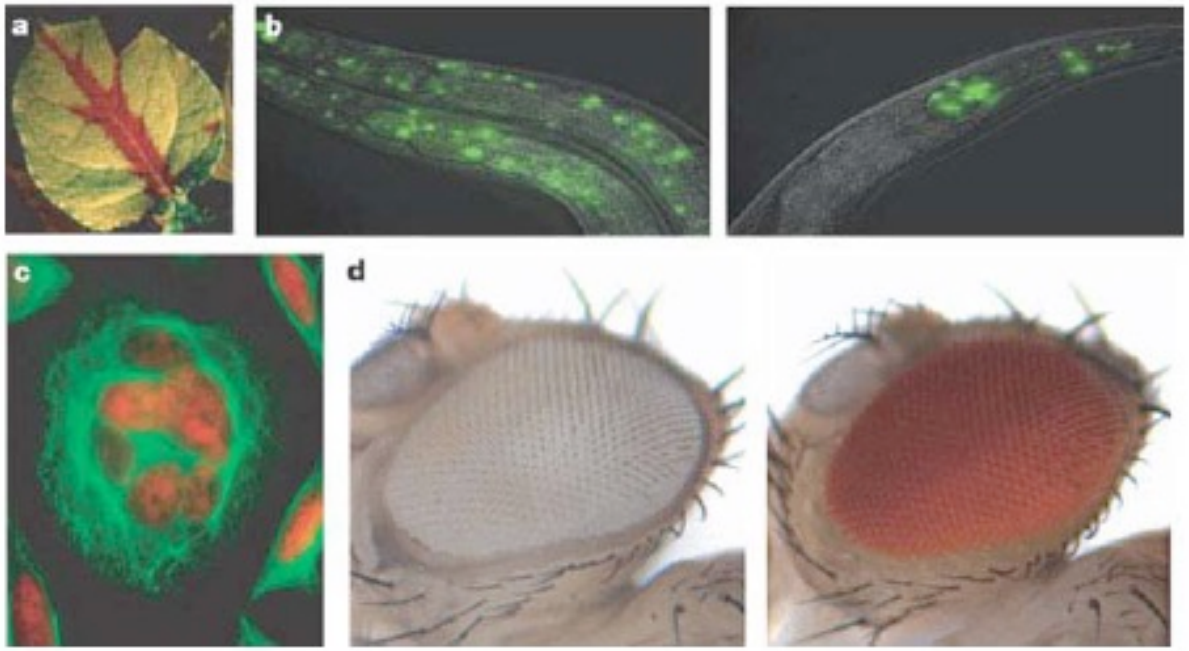
doi:10.1038/nature22990

RNase III nucleases from diverse kingdoms serve as antiviral effectors

Lauren C. Aguado[1]*, Sonja Schmid[1]*, Jared May[2], Leah R. Sabin[3], Maryline Panis[1], Daniel Blanco-Melo[1], Jaehee V. Shim[4], David Sachs[5], Sara Cherry[3], Anne E. Simon[2], Jean-Pierre Levraud[6] & Benjamin R. tenOever[1]

In contrast to the DNA-based viruses in prokaryotes, the emergence of eukaryotes provided the necessary compartmentalization and membranous environment for RNA viruses to flourish, creating the need for an RNA-targeting antiviral system[1,2]. Present day eukaryotes employ at least two main defence strategies that emerged as a result of this viral shift, namely antiviral RNA interference and the interferon system[2]. Here we demonstrate that Drosha and related RNase III ribonucleases from all three domains of life also elicit a unique RNA-targeting antiviral activity. Systemic evolution of ligands by exponential enrichment of this class of proteins illustrates the recognition of unbranched RNA stem loops. Biochemical analyses reveal that, in this context, Drosha functions as an antiviral clamp, conferring steric hindrance on the RNA-dependent RNA polymerases of diverse positive-stranded RNA viruses. We present evidence for cytoplasmic translocation of RNase III nucleases in response to virus in diverse eukaryotes including plants, arthropods, fish, and mammals. These data implicate RNase III recognition of viral RNA as an antiviral defence that is independent of, and possibly predates, other known eukaryotic antiviral systems.

RNAi: a tool for inhibithing gene expression in a sequence specific way



Hannon, G (2002) Nature 418, 244-251

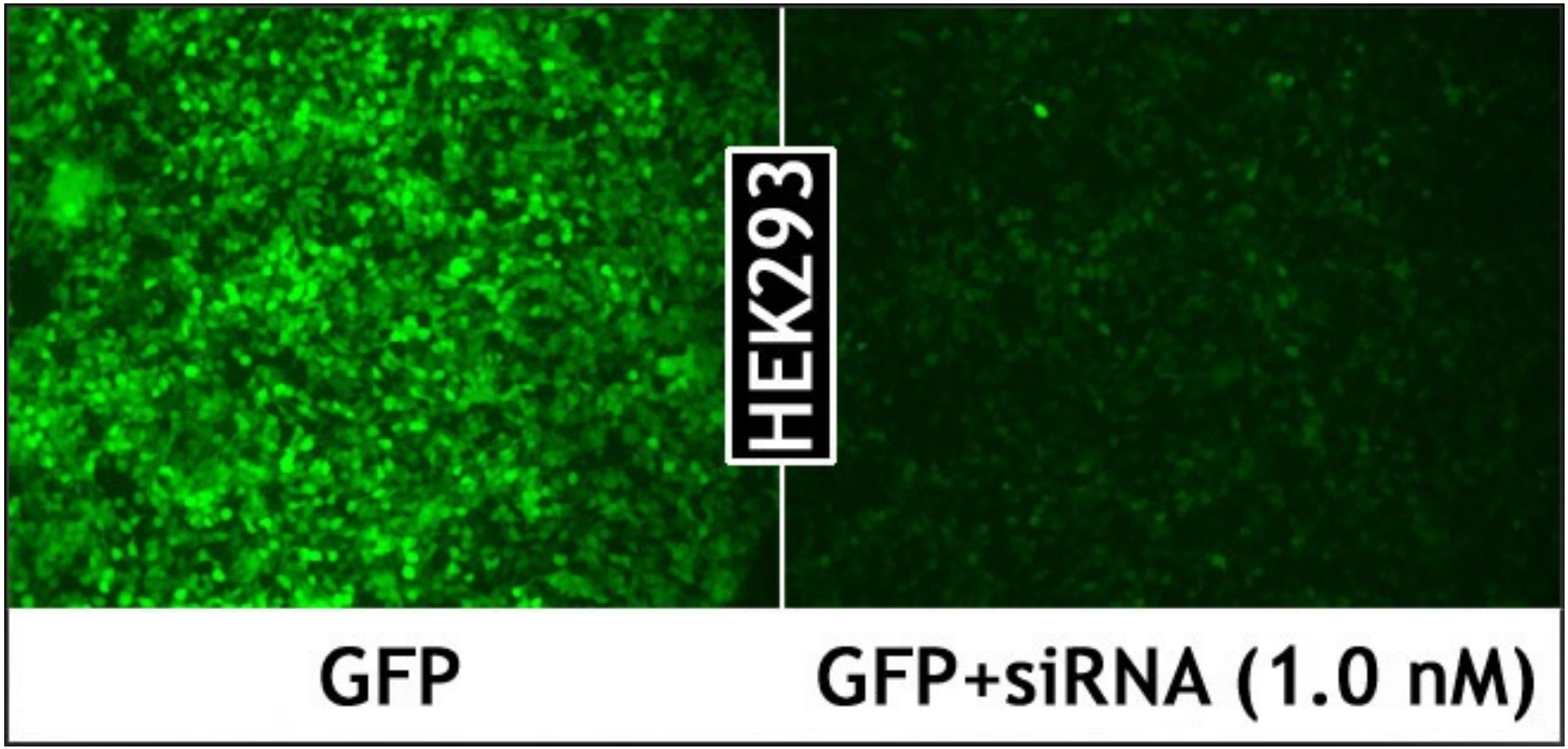
siRNAs as a tool for studying gene function

Short double-stranded RNAs are synthesized with a sequence complementary to the mRNA of interest and introduced into a cell or organism, where they are recognized by the RISC complex and activate the RNAi process.

Effect: drastic decrease in the expression of the target gene and by studying the effects of this decrease, the physiological role of the gene product can be deduced.

Since RNAi does not completely abolish gene expression, it is called a knockdown technique.

RNAi knockdown experiments can be used to study the functions of genes *in vivo*



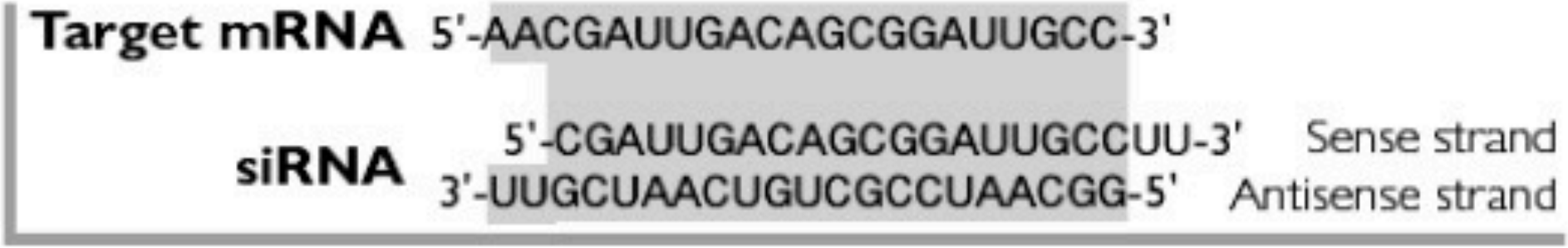
SYSTEMATIC GENOME-WIDE SCREENS OF GENE FUNCTION

Anne E. Carpenter and David M. Sabatini

By using genome information to create tools for perturbing gene function, it is now possible to undertake systematic genome-wide functional screens that examine the contribution of every gene to a biological process. The directed nature of these experiments contrasts with traditional methods, in which random mutations are induced and the resulting mutants are screened for various phenotypes. The first genome-wide functional screens in *Caenorhabditis elegans* and *Drosophila melanogaster* have recently been published, and screens in human cells will soon follow. These high-throughput techniques promise the rapid annotation of genomes with high-quality information about the biological function of each gene.

siRNA design

- 21-23nt
- 2-nt 3' overhangs (**UU overhangs**)
- G/C content: 30-50%.
- No base pair mismatch



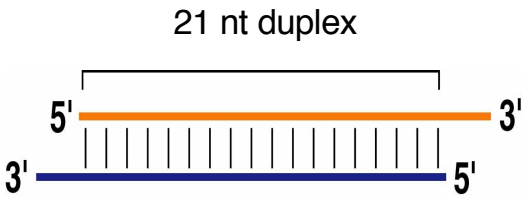
- Synthesized siRNA should not target introns, the 5'and 3'-end untranslated regions (UTR)
- Sequences within 75 bases of the start codon (ATG)

- BLAST : eliminate any target sequences with significant homology to other coding sequences.

How to provide siRNAs to a cell?

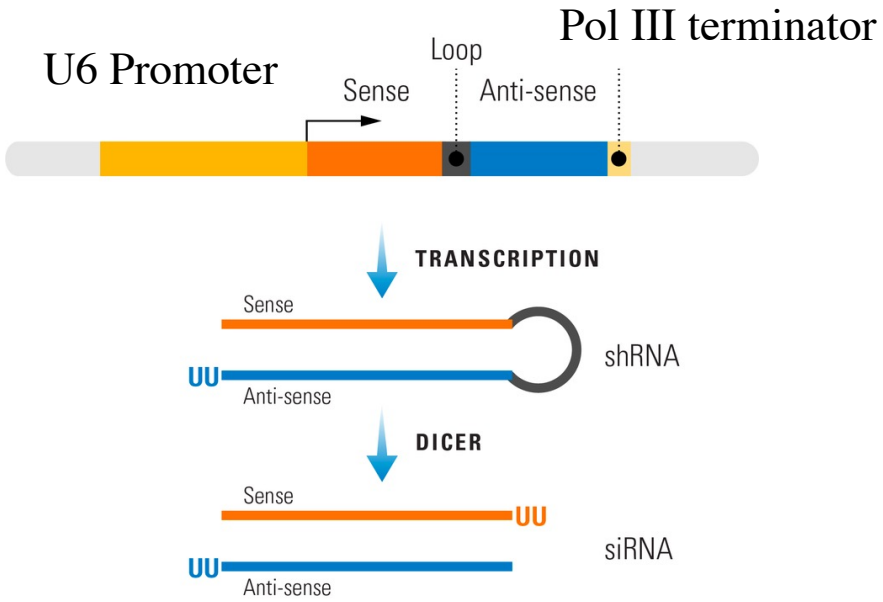
Sinthetic siRNAs

+ easy
+ many target sequences can be tested
- **but** expensive
- short-term effect (depends on the cell type)



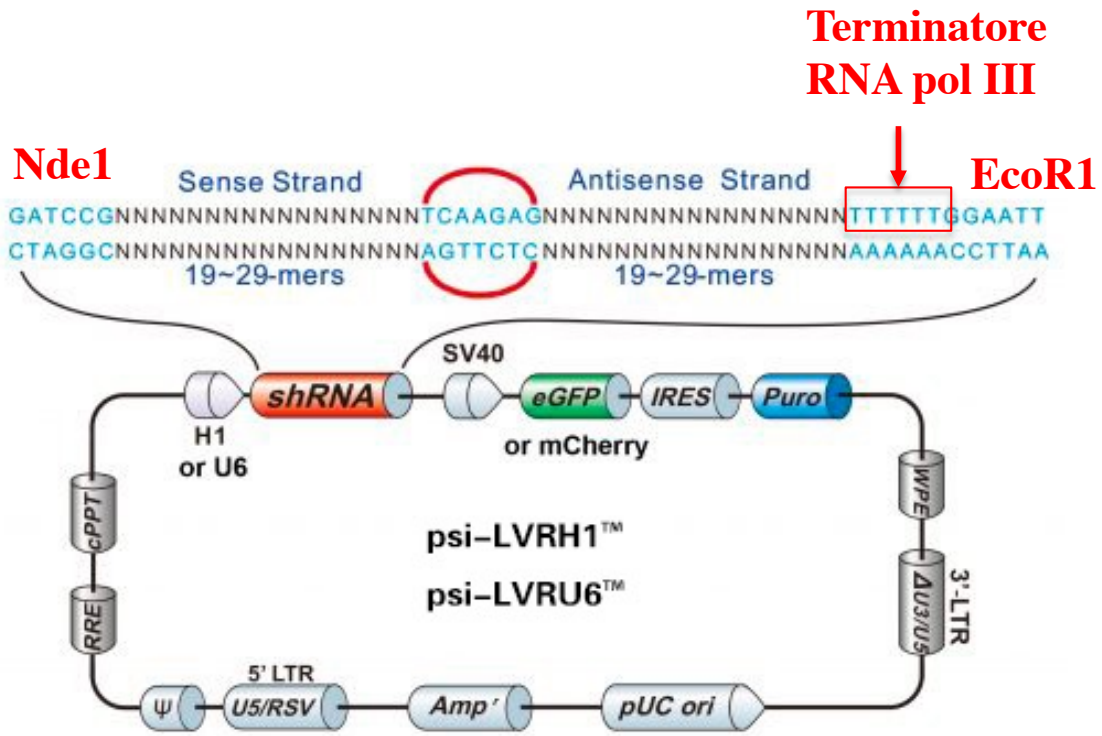
DNA vectors

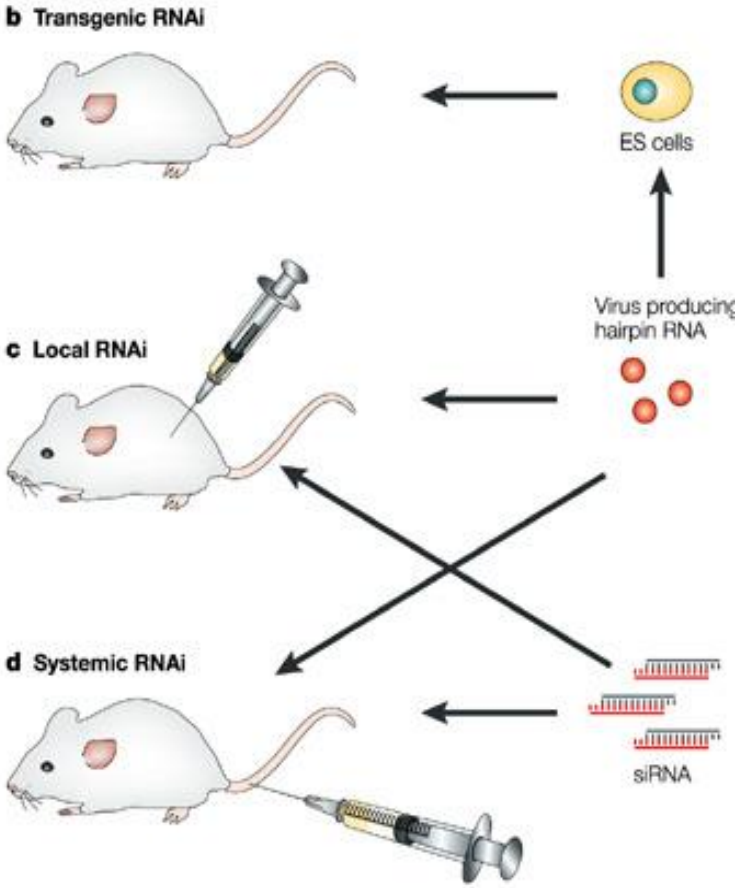
- more laborious
+ inexpensive
+ long-term effect



RNA Pol III promoter and terminator

A method for the stable expression of short interfering RNA in vivo: viral vector





Long term silencing of the target gene

siRNA in Biotechnology

Development of plants that produce low levels of natural toxins so that they can be introduced into the food chain.
The cotton seed is a particularly protein-rich and energetic product that has now become part of livestock feed, especially for dairy cows. However, these seeds contain a toxic product, gossypol, making them unusable. RNAi has been used to produce cotton stocks whose seeds contain reduced levels of delta-cadinene synthase, a key enzyme in 1 the production of gossypol, without affecting the production of the enzyme in other parts of the plant where gossypol is important in preventing damage from plant parasites.

Reduction of allergen levels in tomato plants and reduction of carcinogen precursors in tobacco plants.

Recent study of a new anti-mosquito.
Nanoparticles are used to administer double-stranded RNA to mosquito larvae at the time of feeding. RNA interference has been used to silence the gene responsible for the production of chitin, the main component of the exoskeleton in insects, crustaceans, and arachnids. Since dsRNA dilutes quickly, it cannot theoretically be administered directly and for this reason, the use of nanoparticles is resorted to.

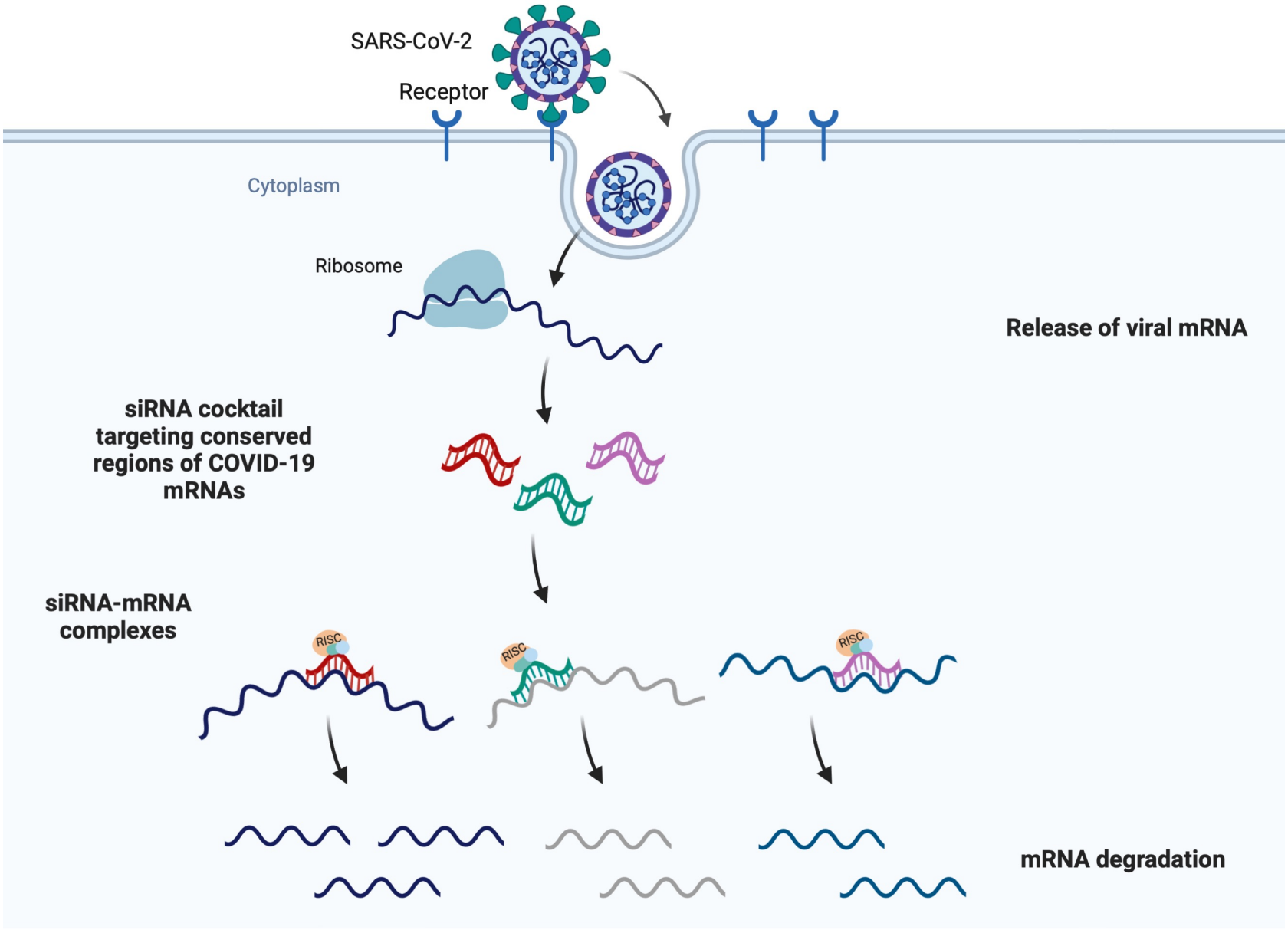
Therapeutic use of RNAi

- Hematological Diseases:
  - Alteration due to reduction, mutation, or absence of gene function

- Oncology:
  - Inhibit oncogenes
  - Increase the efficacy of chemotherapy and radiotherapy

- Infectious Diseases - Targeting Viruses (HCV and HIV):
  - Inhibit cellular and viral factors
  - Target the Reverse Transcriptase RNA, inhibit viral replication
  - Induce resistance to viral infection in infected organisms

Developing Breakthrough RNA Therapeutics for SARS-CoV-2



RNAi and Cancer

The first RNAi administered systemically for the treatment of cancer, Genasense (Genta, Inc.), targets the anti-apoptotic gene BCL2. This RNAi has shown promising results in clinical studies on metastatic melanoma when used in combination with conventional therapeutic chemotherapy.

Initially approved by the FDA but later put on hold.
Highly efficient mechanisms for the delivery of siRNA to specific cells will be particularly necessary and important for the success of metastatic carcinoma treatment.

RNAi for the treatment of neurodegenerative diseases

nature medicine

Silencing mutant SOD1 using RNAi protects against neurodegeneration and extends survival in an ALS model

G Scott Ralph[1], Pippa A Radcliffe[1], Denise M Day[1], Janine M Carthy[1], Marie A Leroux[1], Debbie C P Lee[1], Liang-Fong Wong[1], Lynsey G Bilsland[2], Linda Greensmith[2], Susan M Kingsman[1], Kyriacos A Mitrophanous[1], Nicholas D Mazarakis[1] & Mimoun Azzouz[1]

A synthetic, single-stranded DNA molecule designed to target the SOD1 gene.
An antisense oligonucleotide (ASO) that binds to the messenger RNA of the mutated SOD1 gene.

Tofersen is a targeted therapy that specifically addresses the root cause of SOD1-ALS. By reducing the production of the toxic SOD1 protein, tofersen may help slow disease progression. Tofersen is administered through intrathecal injection directly into the cerebrospinal fluid, ensuring delivery to the central nervous system.

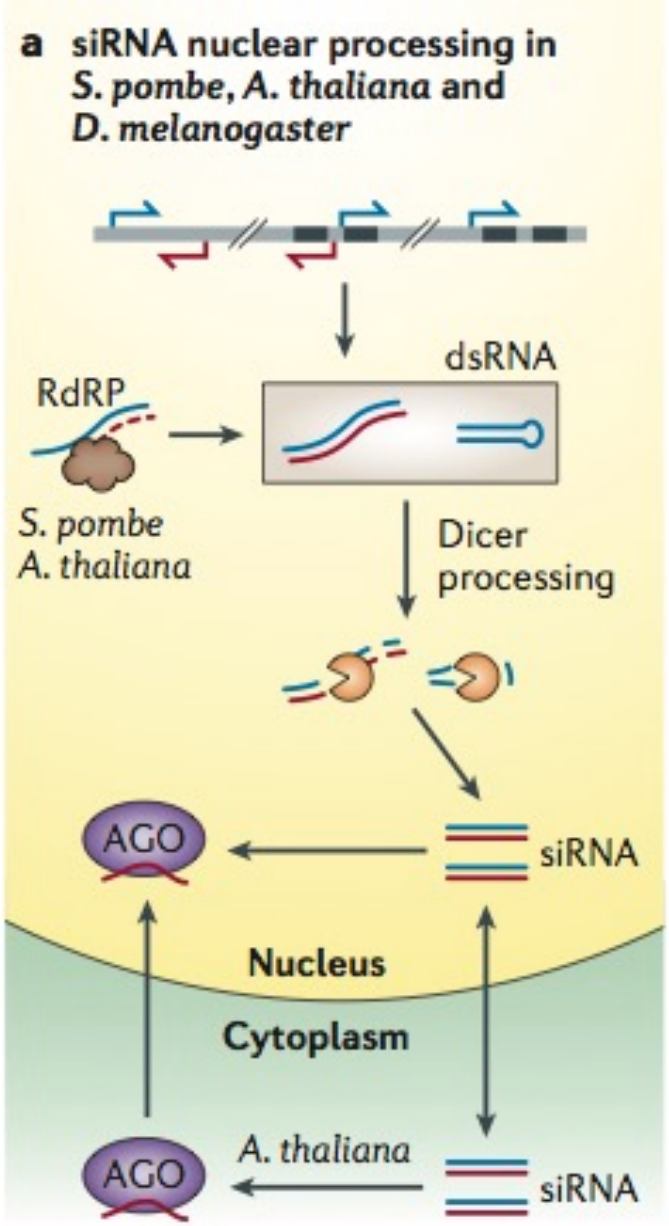
While tofersen shows promise, it's important to note that it is not a cure for ALS.

Ongoing research is exploring the potential of ASOs for other neurodegenerative diseases.

RNA interference in the nucleus

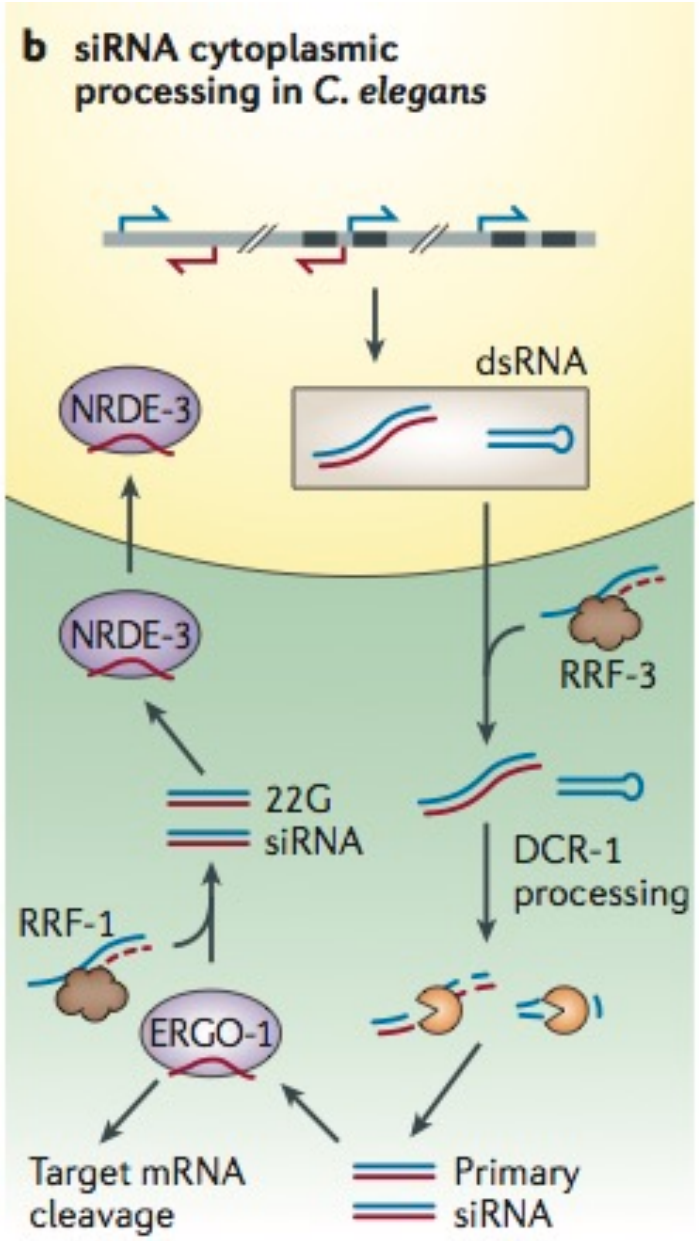
Biogenesis of nuclear small RNAs

siRNA processing takes place in the nucleus in *Schizosaccharomyces pombe* and *Drosophila melanogaster* and in the nucleolus in *Arabidopsis thaliana*. DsRNA can be produced by convergent transcription, complementary transcripts, structured loci or RNA-directed RNA polymerase (RdRP) activity in *A. thaliana* and *S. pombe*. Dicer proteins generate siRNAs that are loaded into an Argonaute protein (AGO). In *A. thaliana,* siRNAs are transported to the cytoplasm, where Argonaute is loaded and then imported into the nucleus.



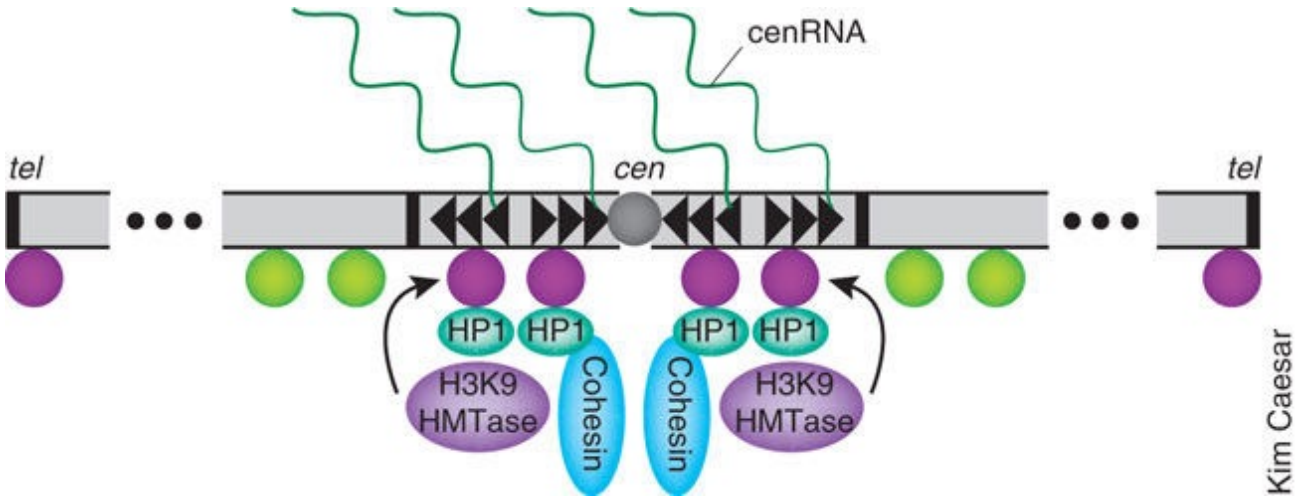
Biogenesis of nuclear small RNAs

In *Caenorhabditis elegans,* siRNA processing occurs in the cytoplasm in a two-step fashion. Primary trigger dsRNA arises from nuclear transcription or the RdRP activity of RRF-3, which acts on transcripts in the cytoplasm. Primary processing by DCR-1 produces primary 26-nucleotide siRNAs, which are loaded into the Argonaute ERGO-1. Loaded ERGO-1 can both facilitate post-transcriptional gene silencing (PTGS) in the cytoplasm and with RRF-1 can generate secondary 22G siRNAs. In the cytoplasm, secondary 22G siRNAs are loaded into the nuclear Argonaute NRDE-3, which is then transported into the nucleus.



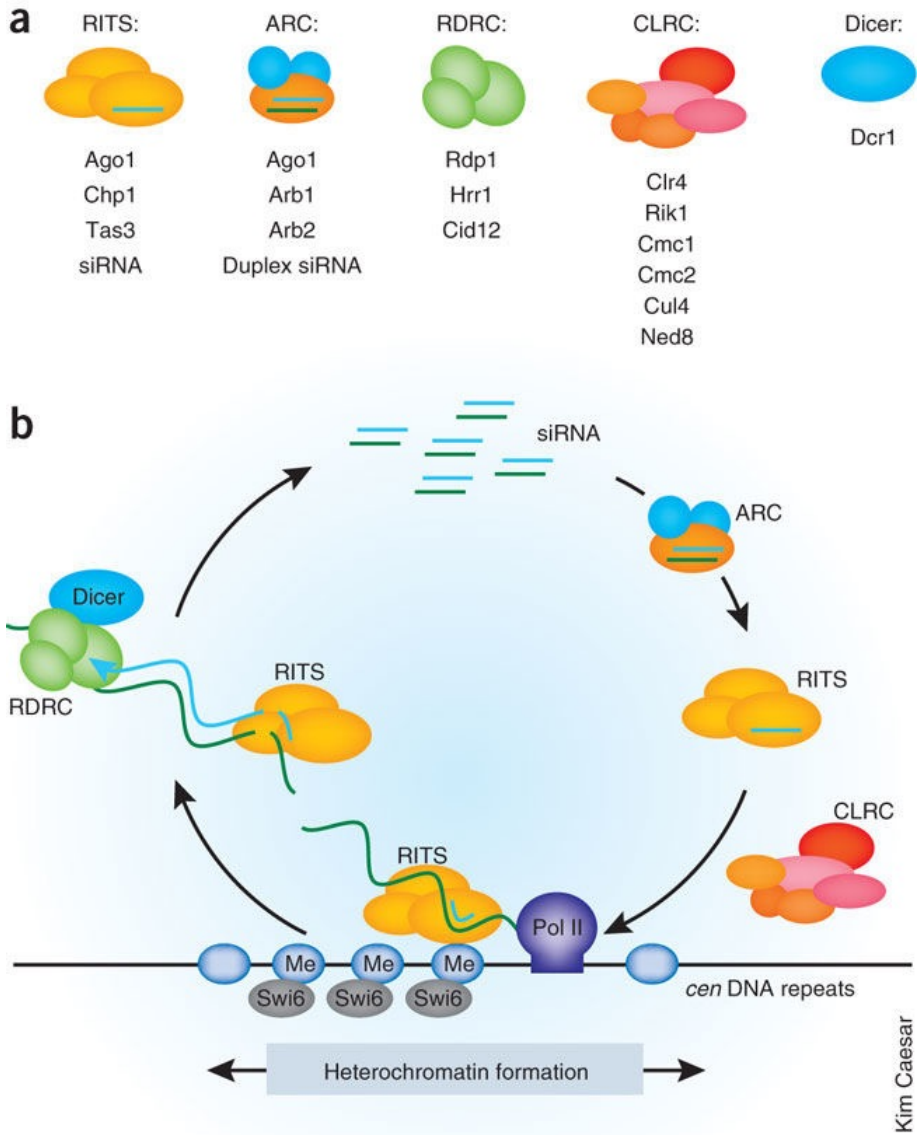
RNAi-mediated heterochromatin assembly in fission yeast

Heterochromatin is important in centromere function and the organization of chromosomes in the nucleus. In *S. pombe, RNAi* is required for the formation of constitutive heterochromatin at pericentromeric regions. These regions are highly enriched for H3K9me and are composed of varying numbers of repeat units that are bidirectionally transcribed to form cenRNAs.



RNAi-mediated heterochromatin assembly in fission yeast

The **RITS complex** mediates heterochromatin formation by associating with nascent transcripts via siRNA base-pairing, and with methylated H3K9 via the chromodomain of its **Chp1** subunit. dsRNA synthesis and siRNA generation occur in association with specific chromosome regions and may underlie *cis* restriction of siRNA-mediated silencing. The chromosome-associated siRNA synthesis loop is essential for the spreading of H3K9 methylation and silencing at the centromere. The coupling of the siRNA synthesis loop to H3K9 methylation forms a stable feedback loop that epigenetically maintains heterochromatin.



RNAi and transposones

Types of transposable elements

Transposable elements (TEs) are divided into Class I and Class II depending on their transposition mechanism:

- Class I elements are called **retrotransposons** because they use RNA as an intermediate that is reverse transcribed into DNA and integrated in the genome. Retrotransposomes are further divided in **autonomous retrotransposons** and **non-autonomous retrotransposons**.

- Class II **DNA transposons** encode a transposase that is required for their excision and insertion through a 'cut- and- paste' mechanism.

Silencing mechanisms that suppress TEs

A range of chromatin modifications suppress TE transcription, including modifications of histone tails, DNA methylation and alterations in chromatin packing and condensation.

Nucleosomes that are associated with TEs are enriched for **methylation of H3K9**, which is a signal for transcriptionally repressive chromatin

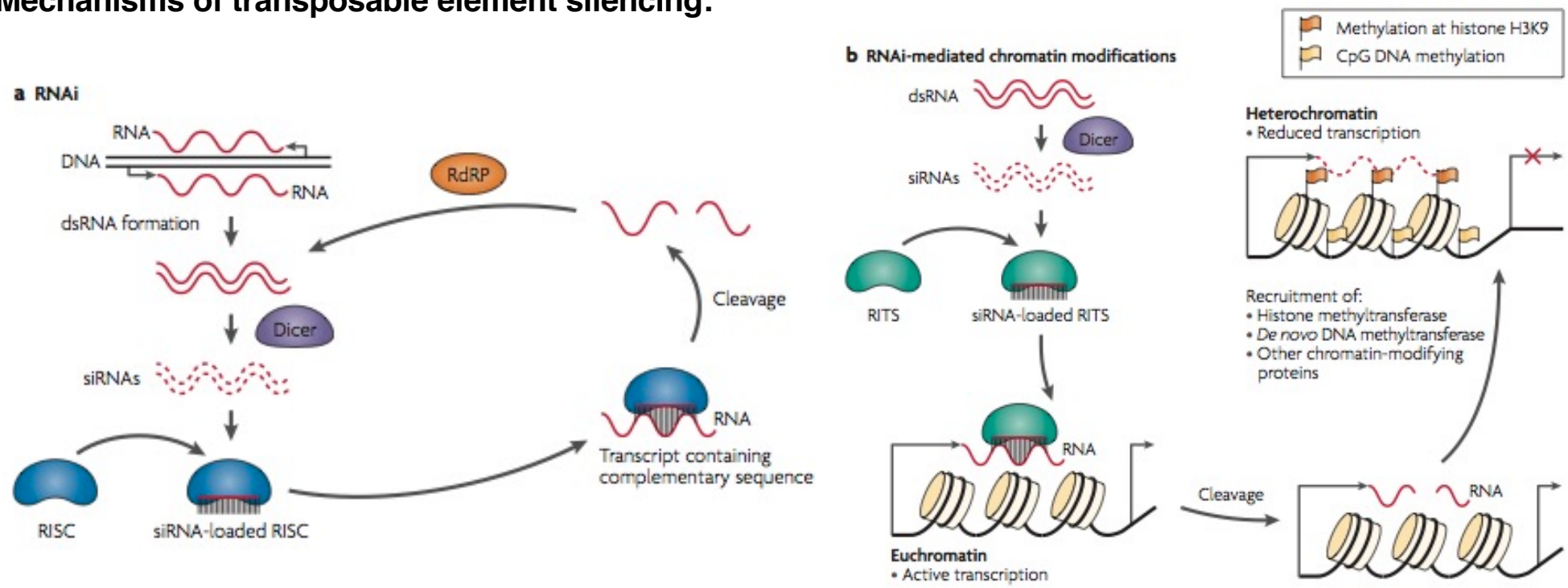
DNA methylation on cytosine residues is another important signal that represses TE transcription. In both plants and mammals, DNA methylation is copied to the new DNA strand upon DNA replication, providing a mechanism for inheritance of TE silencing.

De novo DNA methylation is also required for the epigenetic silencing of TEs at specific stages of mammalian development.

RNA-interference-mediated transposon silencing in the germline

Nuclear RNAi, **in the form of the piRNA pathway** in animals and various **siRNA pathways** in plants, is a front-line defence against transposable element mobilization. These elements have the ability to move and/or to multiply themselves to new positions in the genome, thereby posing a threat to the genomic stability of an organism.

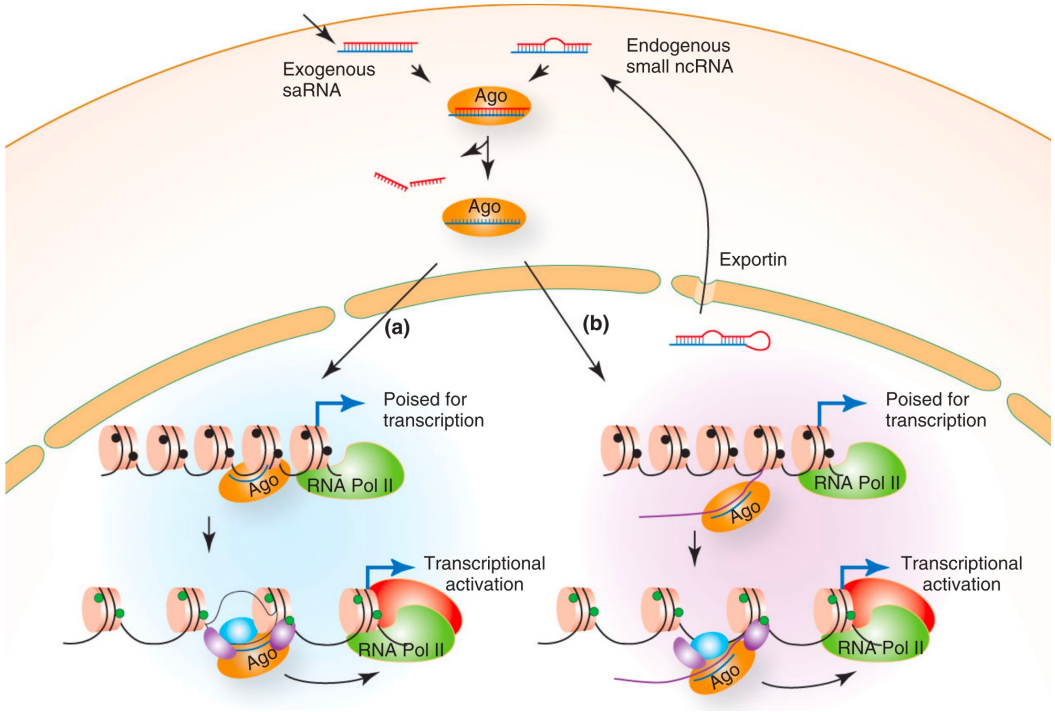
Mechanisms of transposable element silencing:



Small RNA and transcriptional upregulation

There is growing evidence that small RNAs can also serve as **activators** of gene expression by targeting gene regulatory sequences. This novel mechanism, known as **RNA activation** (**RNAa**), appears to be conserved in at least mammalian cells and triggered by artificially designed small RNAs. RNAa depends on Argonaute proteins, but possesses kinetics distinct from that of RNAi. Epigenetic changes are associated with RNAa and may contribute to transcriptional activation of target genes, but the underlying mechanism remains elusive.

It has been proposed that tha Ago-RNA complex may the bind to (a) complementary DNA sequences or (b) nascent cognate transcripts in promoters or 3′ flanking regions and further recruit histone modifiers, leading to an open chromatin structure and active transcription.



saRNA = small activating RNAs

small activating RNAs in clinical trial

Since its discovery in mid 2000s, improvements of saRNA design, synthetic chemistry and understanding of the biology have matured the way to apply RNAa. Indeed, MiNA therapeutics Ltd has conducted the first RNAa clinical trial for advanced hepatocellular carcinoma (HCC) patients with promising outcomes. A second company, Ractigen, recentlystarted preclinical trial with saRNAs for bladder cancer and SMA.

Drug	Company	Indication	Mechanism	Status
MTL-CEBPA plus sorafenib	MiNA	HCC	saRNA targeting *CEBPA*	Phase I/Ib
RAG-01	Ractigen	Bladder cancer	saRNA targeting p21 gene	Preclinical
RAG-06	Ractigen	SMA	saRNA targeting *SMN2*	Preclinical

PIWI-interacting RNAs (piRNAs)

PIWI-interacting RNA (piRNA)

- **PIWI-interacting RNAs (piRNAs)** are small, 21–30 nt single-stranded RNAs that associate with PIWI proteins in various organisms. It is their association with PIWI, but not AGO, proteins and their independence from Dicer that distinguishes piRNAs from siRNAs.

- **PIWI** proteins are a clade within the larger family of Argonaute proteins that is mostly specifically expressed in the germ line.

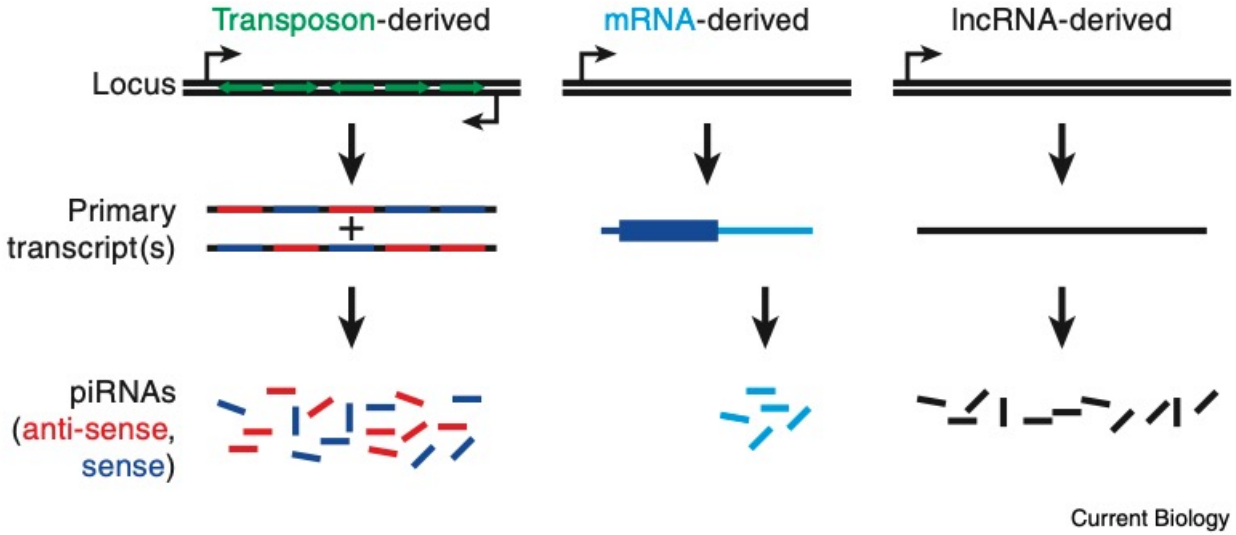
- In most animal species studied, PIWI proteins have been shown to repress 'non-self' sequences, such as transposable elements. However, non-transposon-related PIWI targets and piRNAs derived from regular mRNAs have been described as well.

- Some PIWI proteins act in the cytoplasm and may trigger degradation of mRNAs. However, some PIWI proteins also translocate to the nucleus on piRNA loading, where they can silence their targets at the transcriptional level.

- Some invertebrates use piRNAs to tackle viral infection

piRNA classification

piRNAs can originate from different sources, but piRNA function is only well understood for transposon-derived piRNAs.

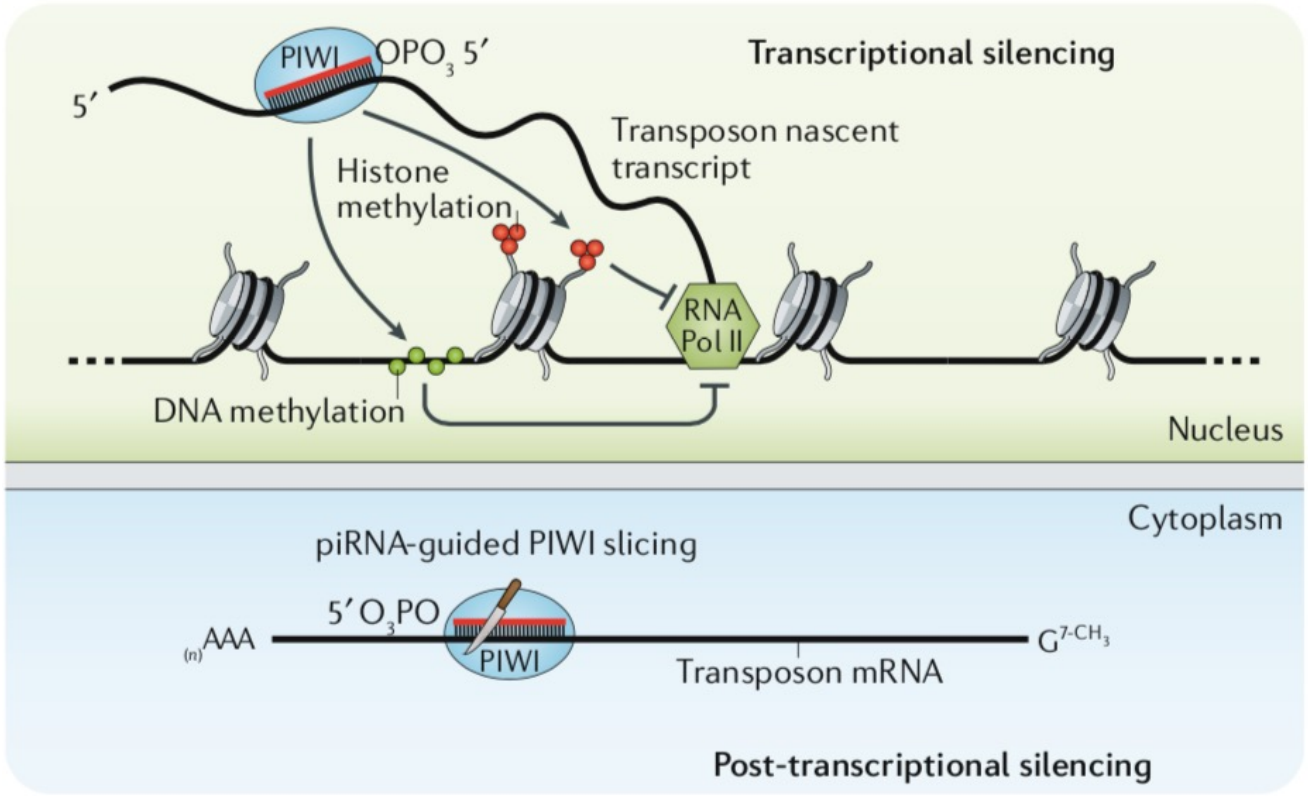


Several features of piRNAs in mice and humans have been reported.

1) are mainly present in germ cells;
2) are expressed in a regulated manner (PIWI protein has regulated expression);
3) map to the genome via clusters;
4) show a strong tendency for uridine (U) at the 5′ end;
5) have both sense and antisense directions;
6) are missing stem loop structures and have 5′-phosphate and 3′-OH groups;
8) Dicer biogenesis is independent of piRNAs;
9) interact with PIWI proteins.

They are one of the most abundantly expressed small RNAs, with every human spermatid containing approximately 1 million piRNAs.

piRNAs silence transposons transcriptionally and post-transcriptionally



Transcription of piRNA clusters

- piRNAs are produced from specific genomic loci termed ‘piRNA clusters’ that are transcribed by RNA Pol II as long non-coding transcripts before they are processed into piRNAs.

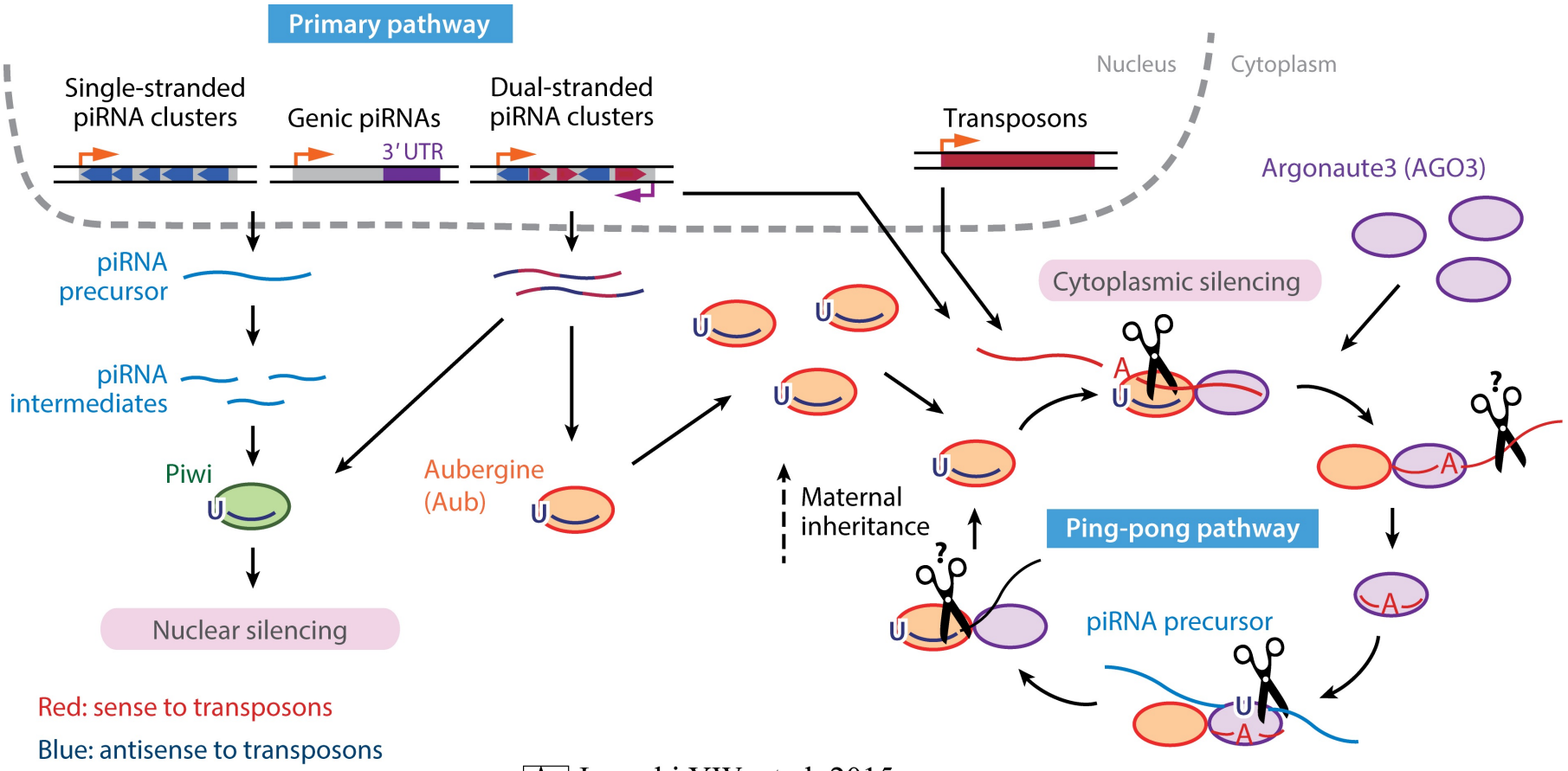
- Most piRNA clusters are bidirectional, lack clear promoter regions and carry H3K9 trimethylation marks that are generally associated with transcription repression.

- Most piRNA precursor transcripts lack classical mRNA features (that is, splicing and polyadenylation). They also use a specific variant of the nuclear export adaptor complex to reach the cytoplasm.

- piRNA precursor transcripts are transported to cytoplasmic perinuclear RNA granules called ‘*nuage*’, where they are processed into mature piRNAs.

Processing of piRNAs

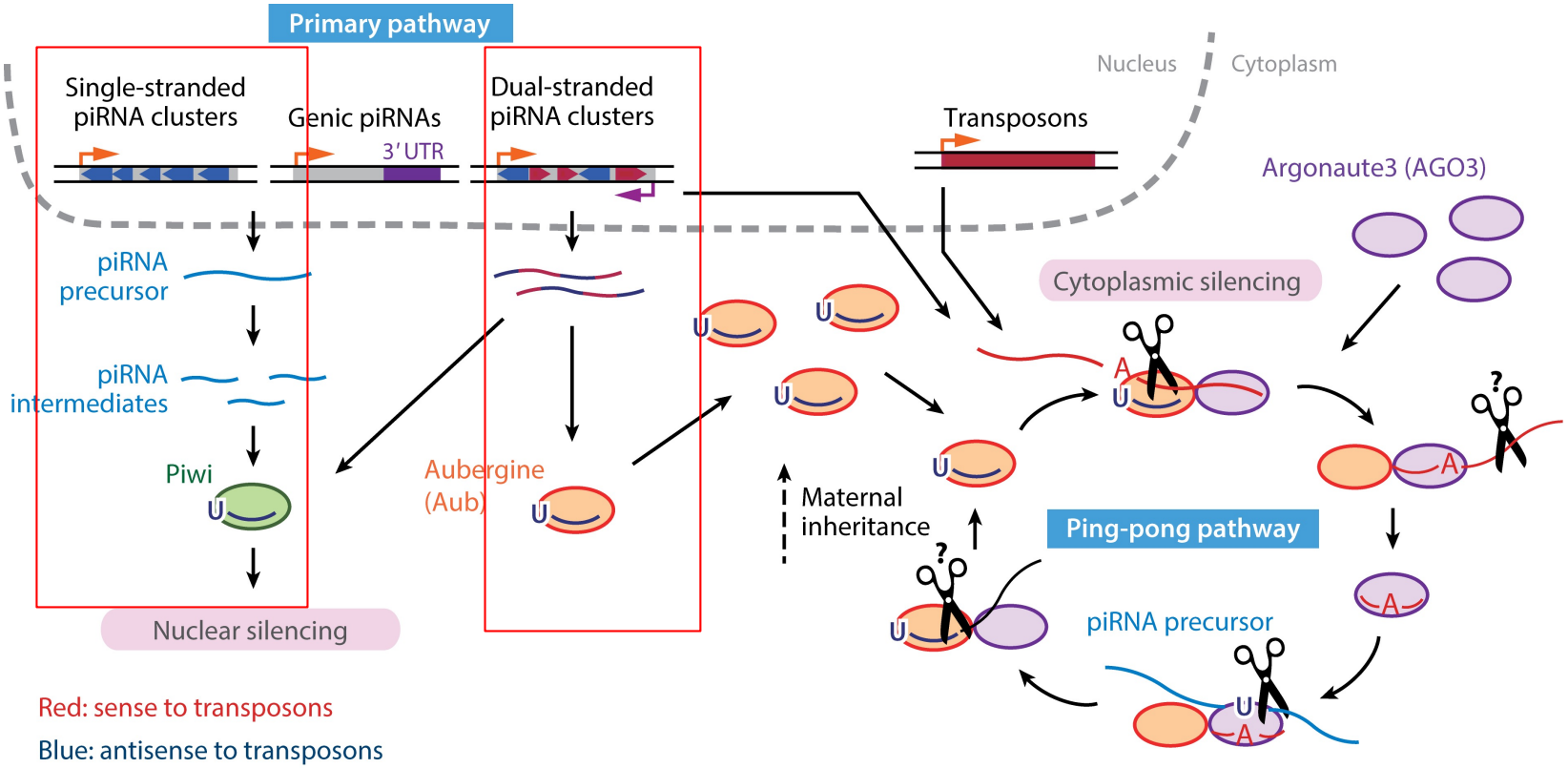
Two major pathways generate piRNAs: the **primary processing pathway** and the **ping-pong cycle** that amplifies secondary piRNAs. In *Drosophila* ovaries, the primary pathway operates in both germline and surrounding somatic cells, whereas the ping-pong cycle operates only in germline cells.



Iwasaki YW, et al. 2015.
Annu. Rev. Biochem. 84:405–33

Biogenesis of piRNAs

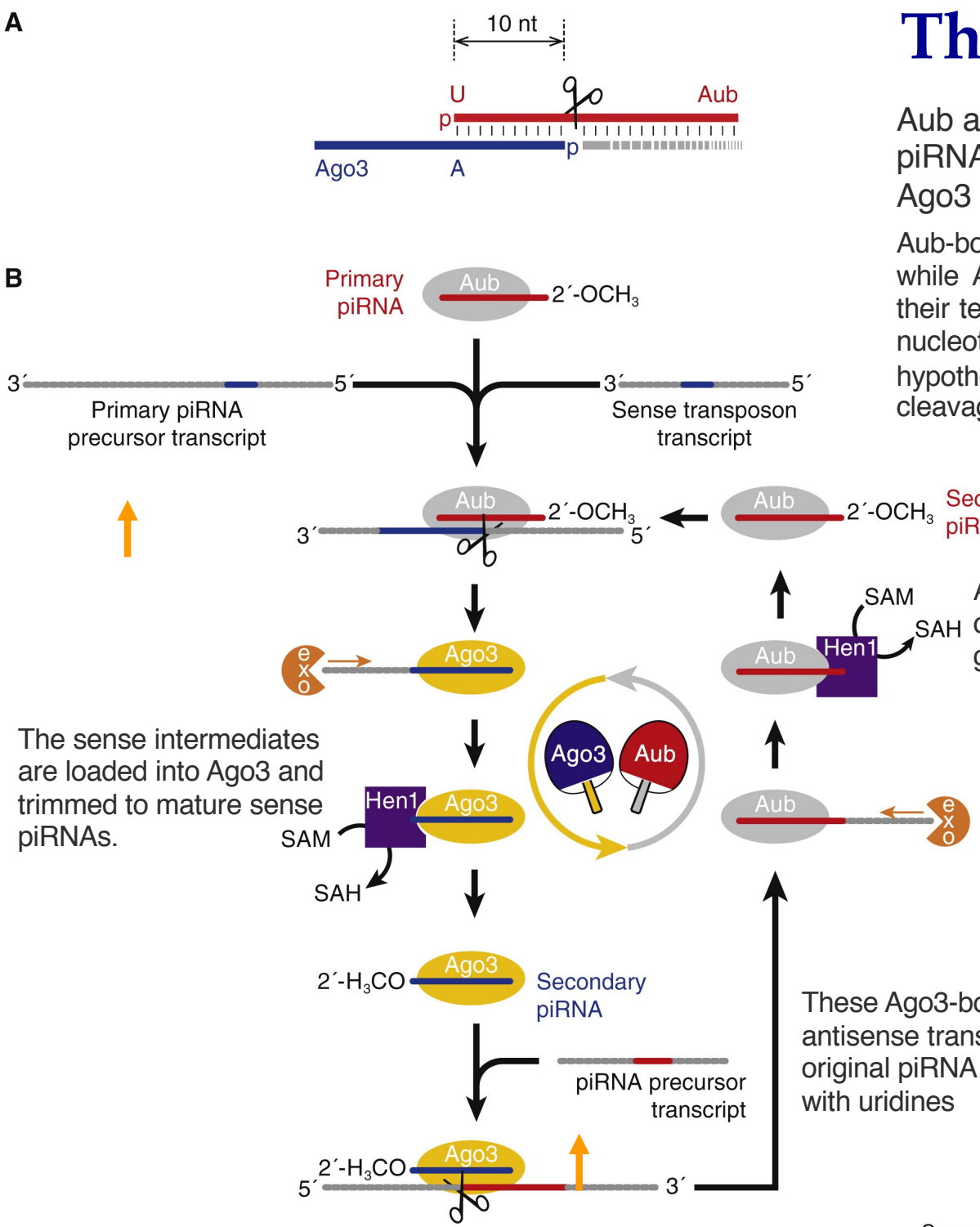
In the primary pathway, piRNAs are transcribed from genomic regions called piRNA clusters, processed, and loaded onto Piwi or Aub. Silencing takes place both in the cytoplasm and nucleus. Piwi performs transcriptional gene silencing in the nucleus. Together with AGO3, the Aub–piRNA complex serves as a trigger to start the ping-pong amplification pathway. The ping-pong pathway silences the target transposon sequence and amplifies the piRNA sequence at the same time. Note that some Aub–piRNA complexes are also maternally inherited.



The ping-pong cycle

Aub and Piwi seem to preferentially bind piRNAs in the antisense orientation while Ago3 tends to associate with sense piRNAs

Aub-bound antisense piRNAs typically start with a 5′ uridine while Ago3-bound sense piRNAs often have adenosine as their tenth nucleotide. Complementarity between the first ten nucleotides of Aub- and Ago3-bound piRNAs led to the hypothesis that sense piRNAs are generated by target cleavage directed by antisense piRNAs, and vice versa



Aub-bound, antisense piRNAs initiate the Ping-Pong cycle by cleaving transposon mRNA transcripts and generating sense-oriented piRNA intermediates.

The sense intermediates are loaded into Ago3 and trimmed to mature sense piRNAs.

These Ago3-bound, sense piRNAs can then bind and cleave the antisense transposon sequences present in the transcripts of the original piRNA cluster, producing piRNA intermediates that begin with uridines

piRNAs in cancer

A study showed that piRNA expression differs significantly across human somatic tissues. Variable degrees of heterogeneity in expression patterns are observed in different kinds of tissues. It has been reported that aberrant expression of piRNAs is a potential cancer-specifi signature and can be correlated with clinical features in malignant tissues, indicating an important role for piRNAs in several types of cancers.