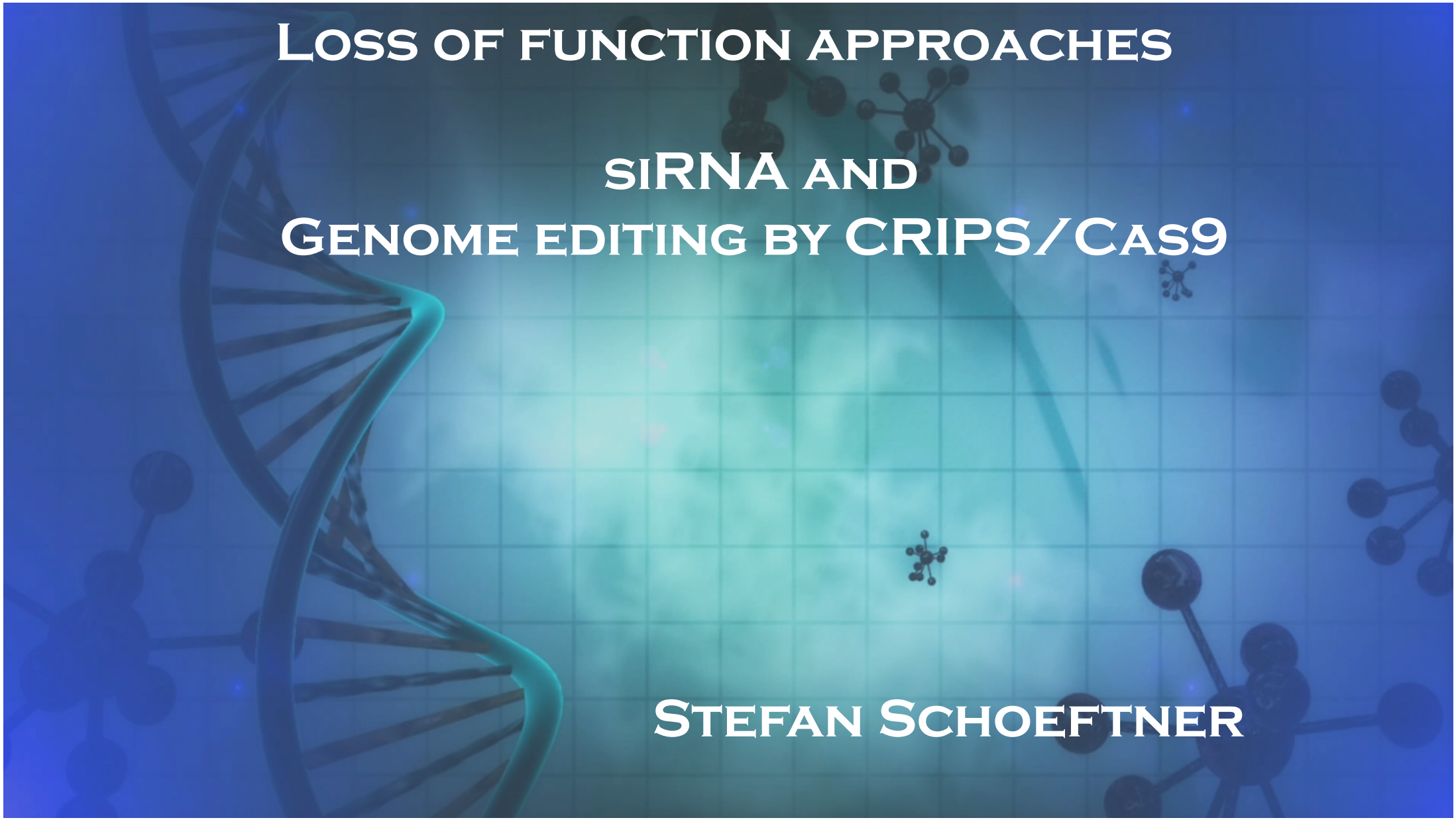
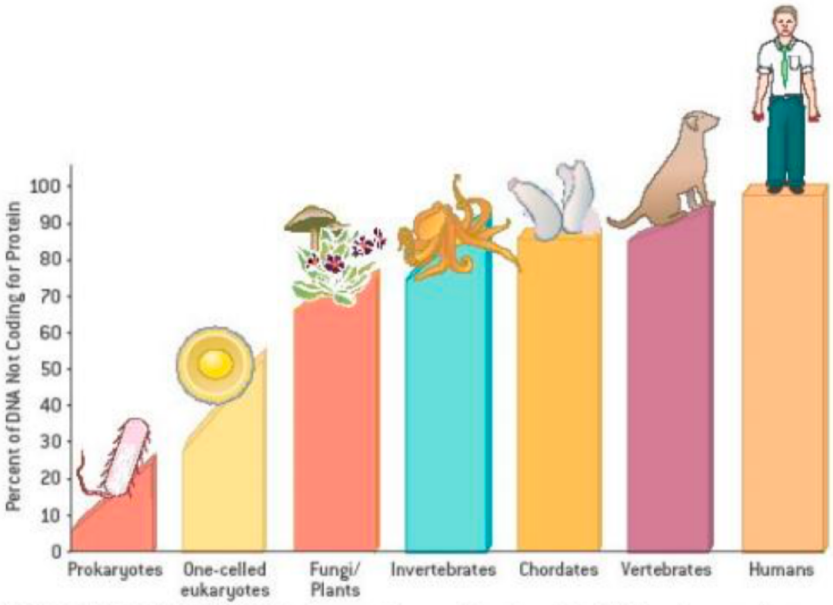
Loss of function approaches

siRNA and
Genome editing by CRIPS/Cas9

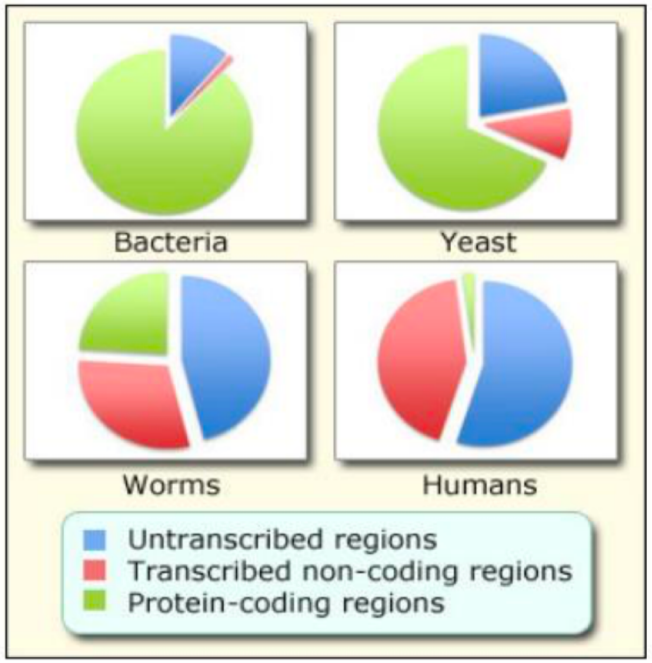


Stefan Schoeftner

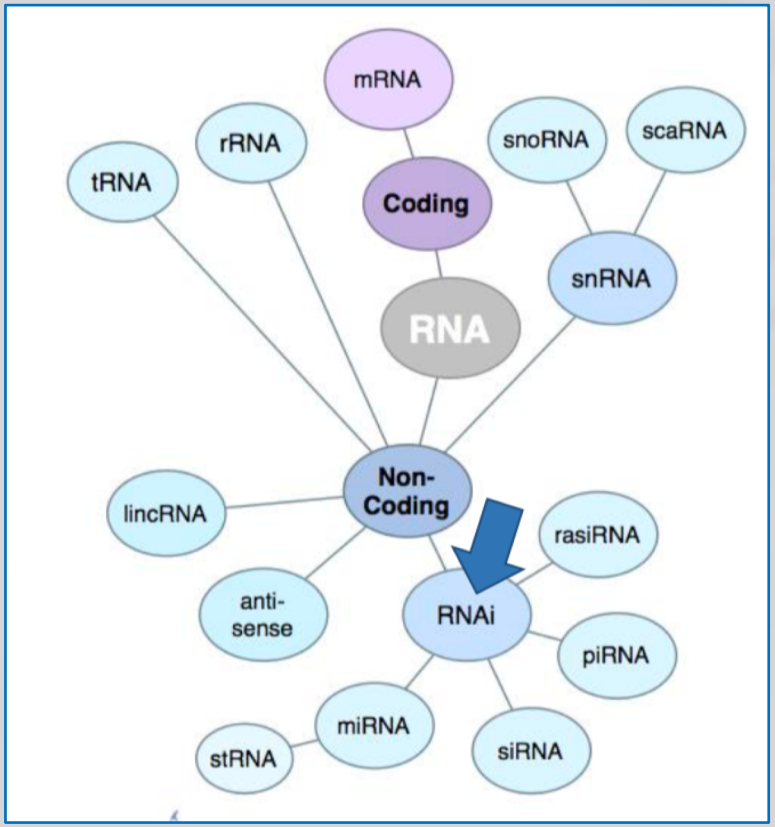
Introduction – Non coding sequences increase with complexity



NONPROTEIN-CODING SEQUENCES make up only a small fraction of the DNA of prokaryotes. Among eukaryotes, as their complexity increases, generally so, too, does the proportion of their DNA that does not code for protein. The noncoding sequences have been considered junk, but perhaps it actually helps to explain organisms' complexity.



Introduction – Non coding RNA (ncRNA) forms



Molecule	Function
mRNAs	messenger RNAs, code for proteins
rRNAs	ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
tRNAs	transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
snRNAs	small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
snoRNAs	small nucleolar RNAs, used to process and chemically modify rRNAs
scaRNAs	small cajal RNAs, used to modify snoRNAs and snRNAs
miRNAs	microRNAs, regulate gene expression typically by blocking translation of selective mRNAs
siRNAs	small interfering RNAs, turn off gene expression by directing degradation of selective mRNAs and the establishment of compact chromatin structures
Other non-coding RNAs	function in diverse cell processes, including telomere synthesis, X-chromosome inactivation, and the transport of proteins into the ER

Co-suppression of gene expression

The discovery of RNAi was preceded first by observations of transcriptional inhibition by antisense RNA expressed in transgenic plants.

Reports of unexpected outcomes in experiments performed by plant scientists in the United States and the Netherlands in the **early 1990s**.

Attempted to overexpress **chalone synthase** (anthrocyanin pigment gene) in petunia. (trying to darken flower color)

Caused **the loss** of pigment .

Further investigation of the phenomenon in plants indicated that the downregulation was due to **post-transcriptional inhibition of gene expression** via an increased rate of mRNA degradation.

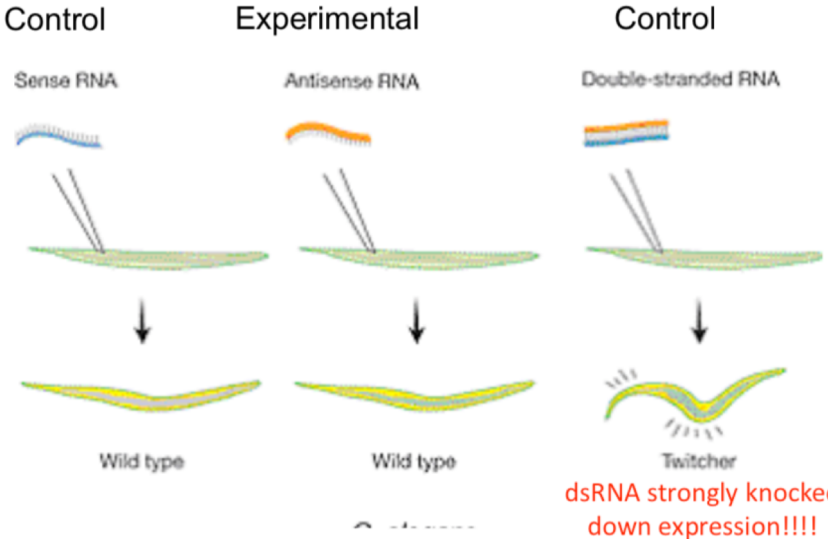
This phenomenon was called **co-suppression of gene expression**, **because suppressed expression of both endogenous gene and transgene** but the molecular mechanism remained unknown

RNA interference (RNAi)

Previously known as co-suppression or post transcriptional gene silencing (PTGS), now is known as **RNA interference (RNAi) a**s a process within living cells that moderates the activity of their genes.

Accidental Discovery of RNAi

- Goal: silence endogenous mRNAs with antisense RNA
- The *unc-22* gene encodes a myofilament protein.
- Decrease in *unc-22* activity is known to produce severe twitching movements.



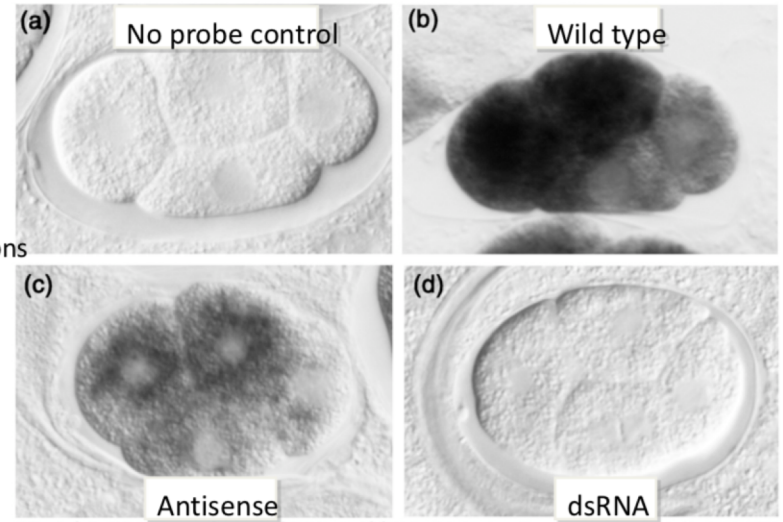
Fire *et al.* *Nature* 1997

Phenotypic effect after injection of ssRNA or dsRNA (*unc-22)* into the gonad of *C. elegans.*

RNA interference (RNAi)

Injection of dsRNA in *C. elegans* Shown To Cause Destruction of Specific mRNA

- Mello and colleagues, 1998
- Injection in gonads of dsRNA for *mex-3* (abundant RNA) gave much more efficient inhibition in embryos than antisense RNA
- dsRNA had to include exons; introns and promoter didn't work
- Effect was incredibly potent and even spread to other cells within the worm
- Termed 'RNA Interference'
- Incredibly useful as a tool for molecular biology



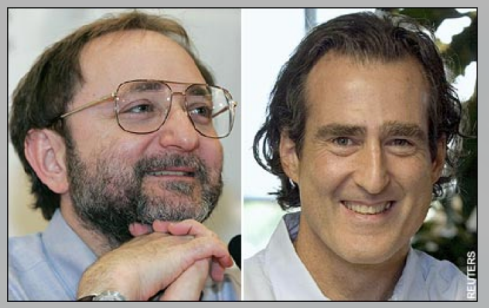
in situ hybridization four-cell stage embryo

Fire *et al. Nature* 1998

- dsRNA from mature mRNA elicits RNAi
- dsRNA from introns does not
- RNAi results in decreased mRNA levels
- RNAi is heritable (for a few generations)
- RNAi only requires a few molecules of dsRNA per cell
- RNAi is applicable to many different transcripts

RNA interference (RNAi)

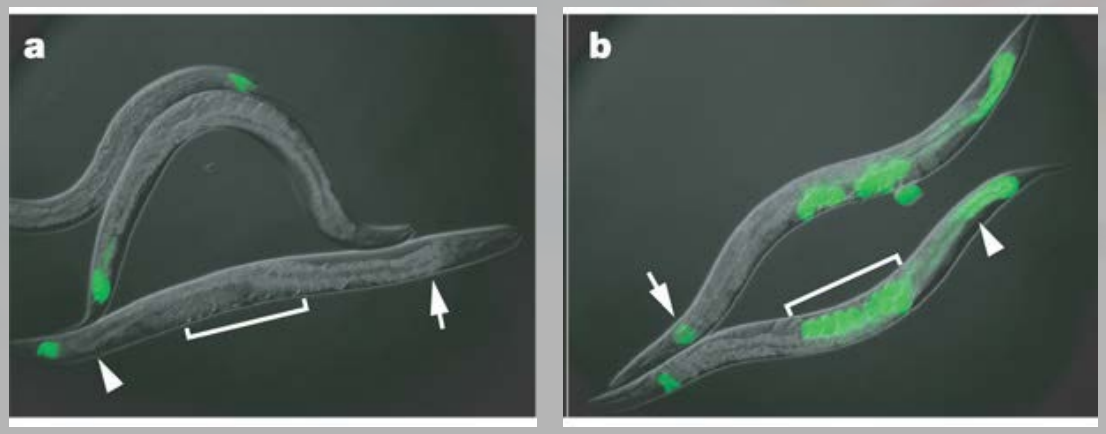
In **2006**, Andrew Fire and Craig C. Mello shared the Nobel Prize
in Physiology or Medicine for their work on RNAi in the nematode worm C. elegans.



Andrew Fire Craig Mello

RNAi in *C.elegans*

– Silencing of a green fluorescent protein (**GFP**) reporter in *C. elegans* occurs when animals feed on bacteria expressing GFP dsRNA (**a**) but not in animals that are defective for RNAi (**b**).

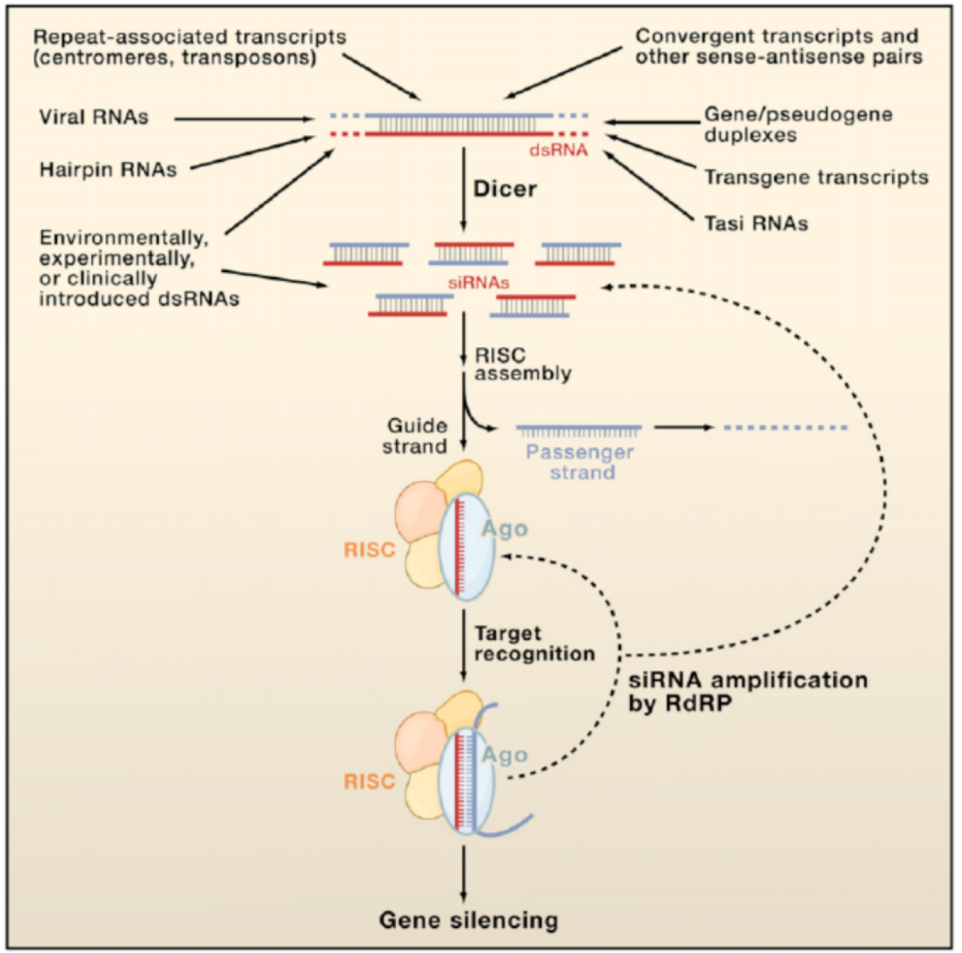


The lack of GFP-positive embryos in a (bracketed region) demonstrates the systemic spread

Sources of dsRNA

Sources of dsRNA

- Some dsRNAs have viral origin, but not all
- Genomic repetitive sequences also are source of siRNA
- Some even regulate other genes (ta-siRNA for trans-acting in plants)
- exo siRNAs (viral etc)
- endo siRNAs –the precursor has a nuclear phase (hairpins, sense-antisense transcripts etc)



Carthew and Sontheimer, Cell (2009) 136, 642-655.

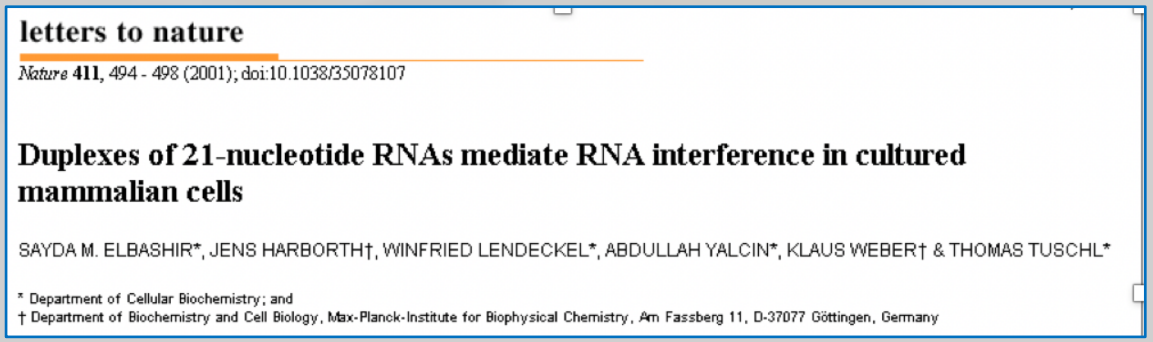
siRNA and miRNA

Two types of RNA molecules involved:
- **small interfering RNA (siRNA)**
- **microRNA (miRNA)**

They bind to other specific mRNAs and modulate their activity.

RNA interference has played an important role in defending cells against parasitic nucleotide sequences – viruses and transposons – but also in directing development as well as gene expression in general.

In 2001 first report of RNAi in MAMMALS



letters to nature

Nature **411**, 494 - 498 (2001); doi:10.1038/35078107

Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells

SAYDA M. ELBASHIR*, JENS HARBORTH†, WINFRIED LENDECKEL*, ABDULLAH YALCIN*, KLAUS WEBER† & THOMAS TUSCHL*

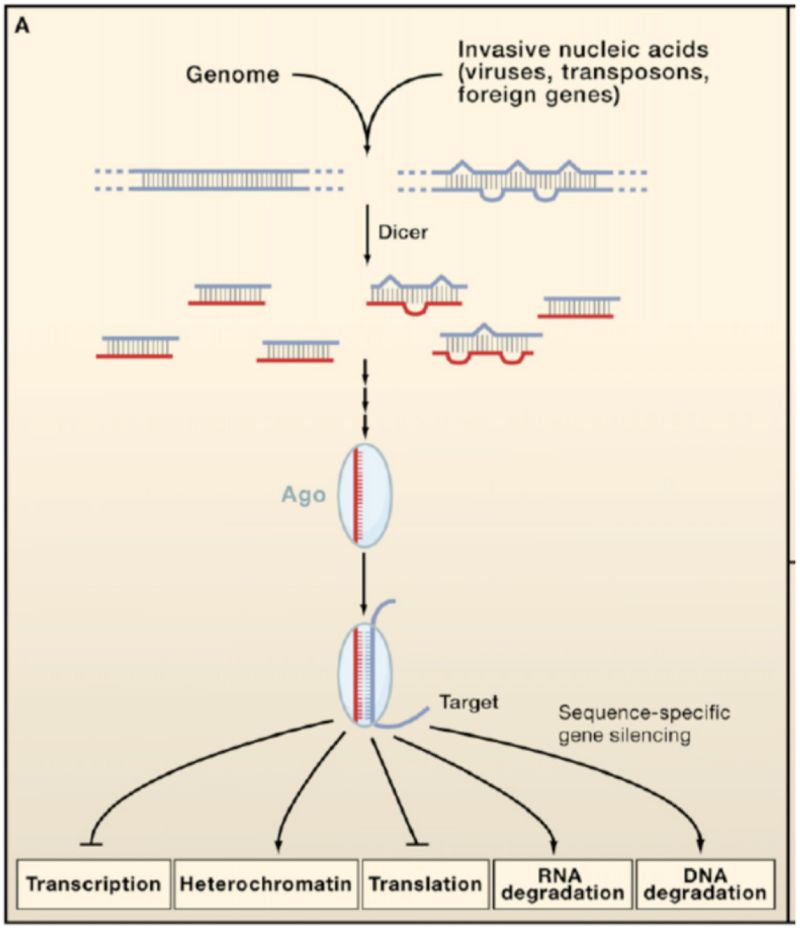
* Department of Cellular Biochemistry; and
† Department of Biochemistry and Cell Biology, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

siRNA and miRNA

Common Features of Pathways for siRNA and miRNA

Both pathways include:

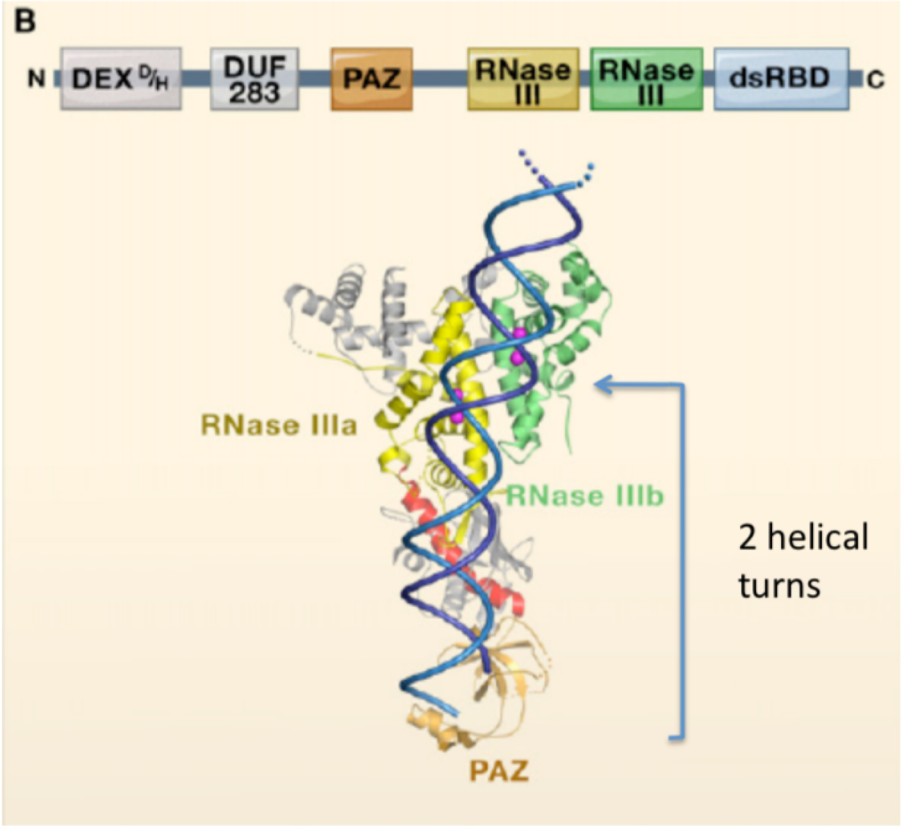
1. • dsRNA 'trigger'
2. • Dicer processing enzyme
3. • Argonaute (Ago)-containing complex to carry out effector function



Carthew and Sontheimer, Cell (2009) 136, 642-655.

DICER: Producer of Small (21-23 bp) RNA fragments

- Structure solved by Doudna and colleagues (2006)
- PAZ domain binds RNA end, RNase III domains cut RNA to produce 2 nt 3'-overhang
- Roles of other domains (not present in structure) remain unclear

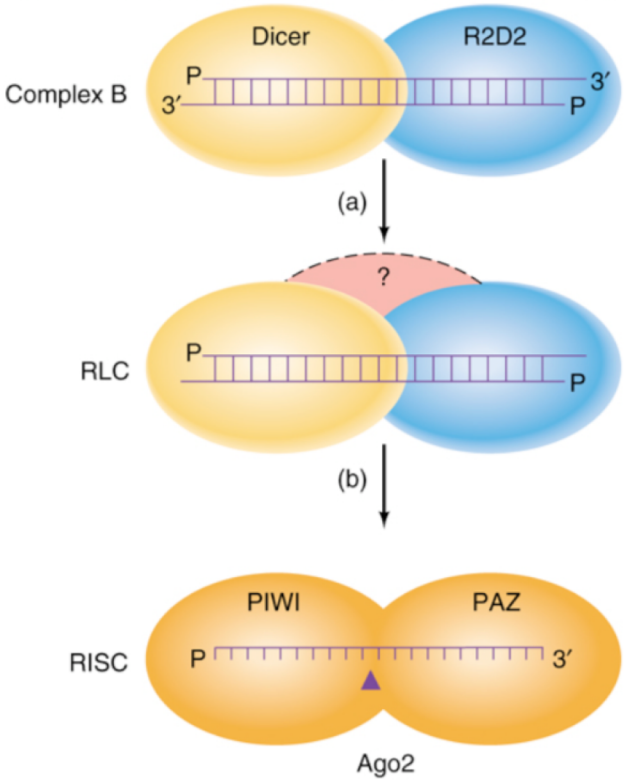


Carthew and Sontheimer, Cell (2009) 136, 642-655.

DICER partners and RISC

Assembly of the RNA-Induced Silencing Complex (RISC) Involves Additional Proteins

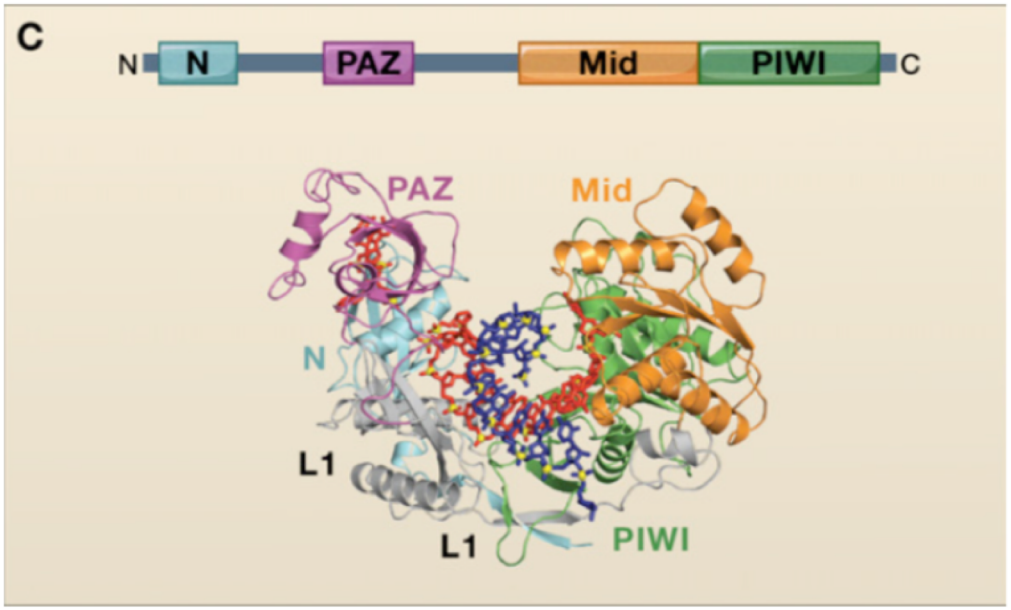
- Processing of dsRNAs into RISC requires accessory proteins: TRBP (R2D2 in *Drosophila*) forms complex with Dicer
- Other unknown proteins bind to form RISC Loading Complex
- Ago2 cleaves the passenger strand, leading to its ejection



ARGONAUTE: Central component of RISC

- One strand of the dsRNA produced by Dicer is retained in the RISC complex in association with Argonaute
- Structure first solved by Leemor-Tor and colleagues (2004), more recent structures by Patel and colleagues include RNAs mimicking guide ssRNA and target mRNA

- The PAZ domain has RNA 3' end binding activity
- In structure without mRNA, guide strand nucleotides 2-6 have bases exposed and available for base-pairing
- PIWI domain adopts RNase H fold and in some Ago proteins can cleave the 'passenger strand' : I.e. the mRNA



Carthew and Sontheimer, Cell (2009) 136, 642-655.

siRNA: Exogenous dsRNA molecules

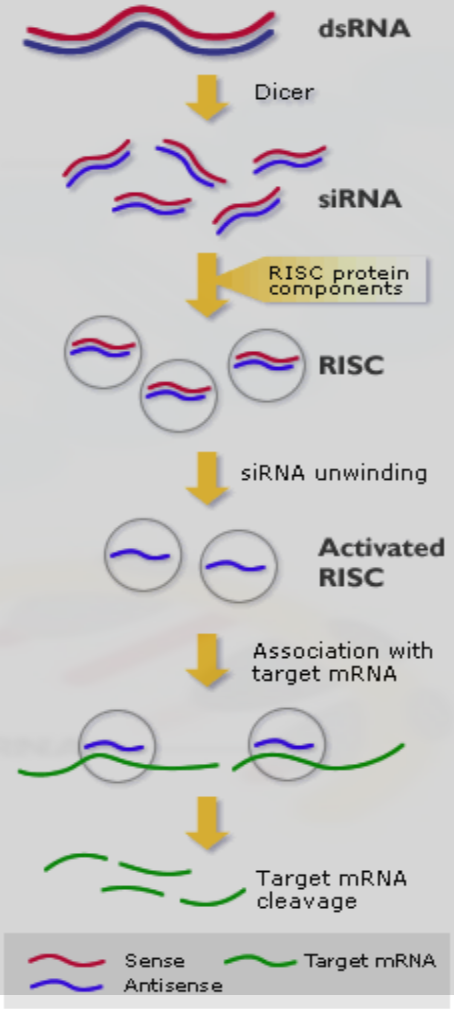
RNAi is controlled by **RISC** and is initiated by short dsRNA molecules in a cell's cytoplasm, where they interact with the catalytic RISC component argonaute.

dsRNAs is cleaved by the **Dicer enzyme** into short fragments of ~20 nucleotides that are called **siRNAs**.

Each siRNA is unwound into two single-stranded (ss) ssRNAs (**passenger** strand and the **guide** strand).

The passenger strand is degraded (red), and **the guide strand (blue) is incorporated into the RNA-induced silencing complex (RISC).**

The most well-studied outcome is post-transcriptional gene silencing, which occurs when the guide strand base pairs with a complementary sequence in a messenger RNA molecule (green) and induces **cleavage by Argonaute**, the catalytic component of the RISC complex.



miRNA: Endogenous RNA silencing

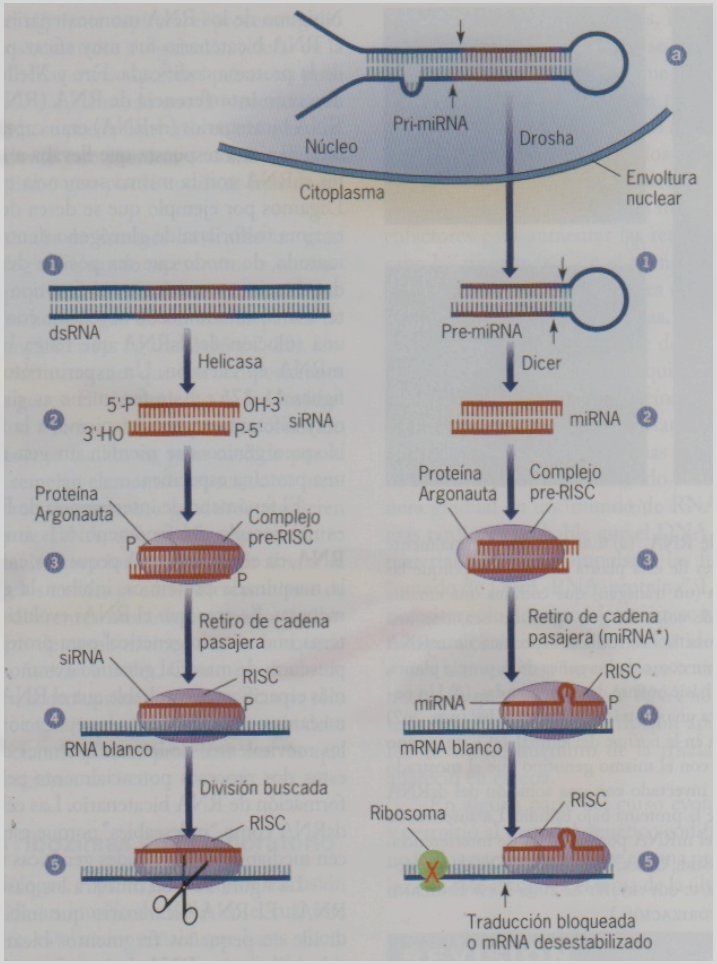
miRNAs are genomically encoded non-coding RNAs that regulate gene expression, particularly during development.

Mature miRNAs are structurally similar to siRNAs produced from exogenous dsRNA but **must undergo post-transcriptional modification**.

miRNA's are expressed from longer RNA-coding gene as a primary transcript (**pri-miRNA**) which is processed within the cell nucleus to a 70 bp stem-loop structure (**pre-miRNA**) by the microprocessor complex (RNase III **Drosha** and dsRNA binding protein DGCR8).

The dsRNA molecule is bound and cleaved by Dicer to produce the mature miRNA molecule that can be integrated into the RISC complex; thus, **miRNA and siRNA share the same cellular machinery downstream of their initial processing**.

miRNAs typically inhibit the translation of many different mRNAs with similar sequences. In contrast, siRNAs typically inhibit only a single, specific target.



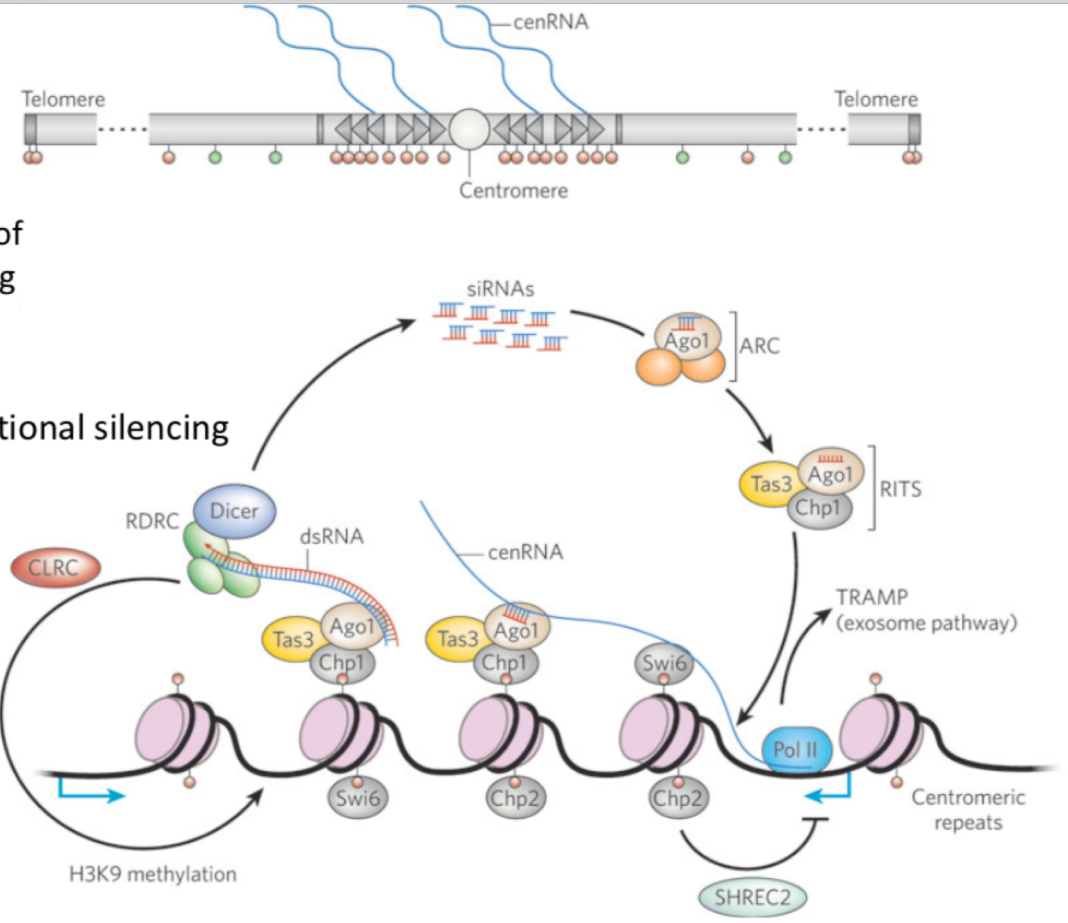
dsRNA in the nucleus: silencing by formation of heterochromatin

- Pathway best understood in *S. pombe*
- Silencing involves formation of heterochromatin and resulting transcriptional repression

the RNA-induced transcriptional silencing complex (RITS)

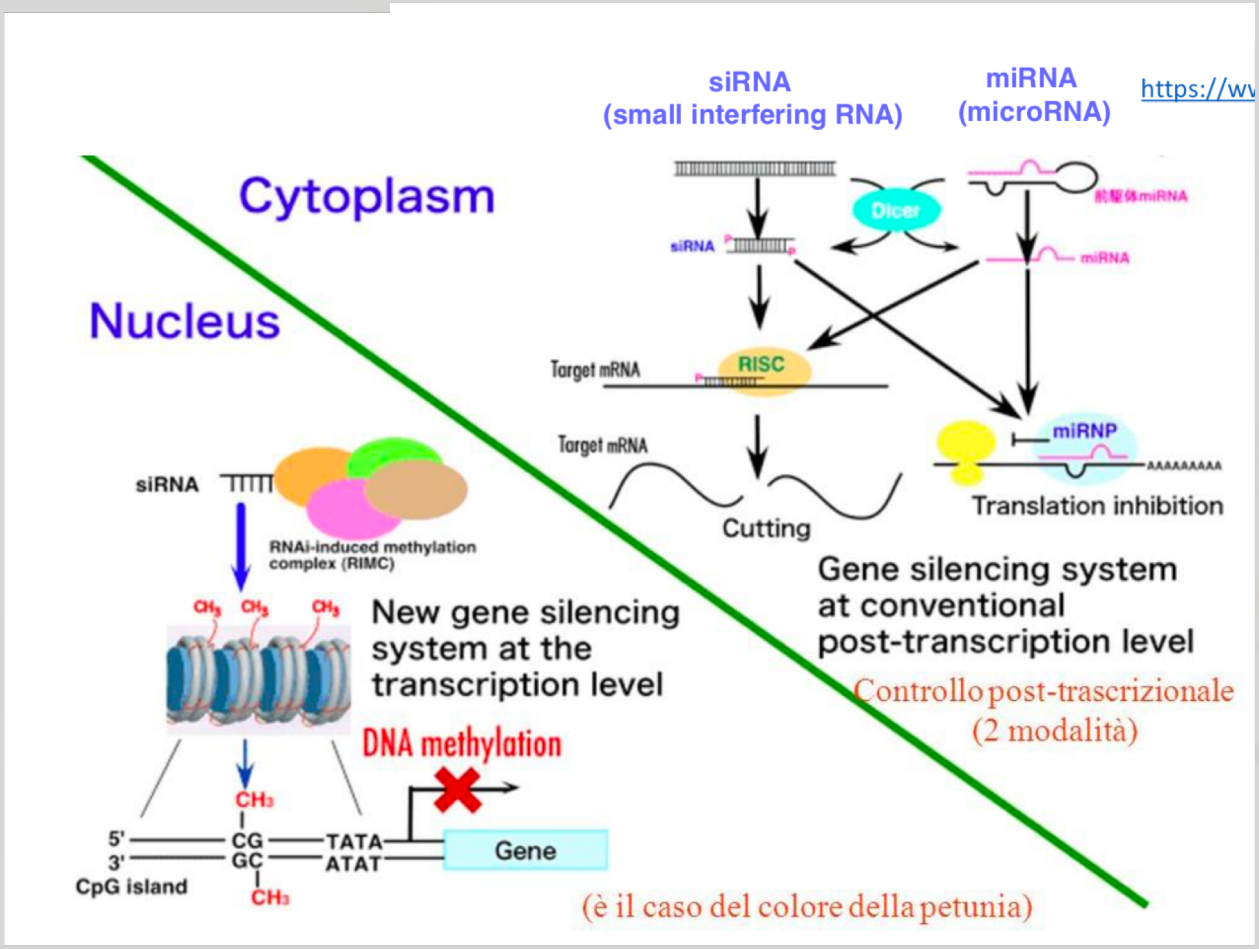
Clr4 methyltransferase complex (CLRC)

RNA-directed RNA polymerase complex (RDRC)

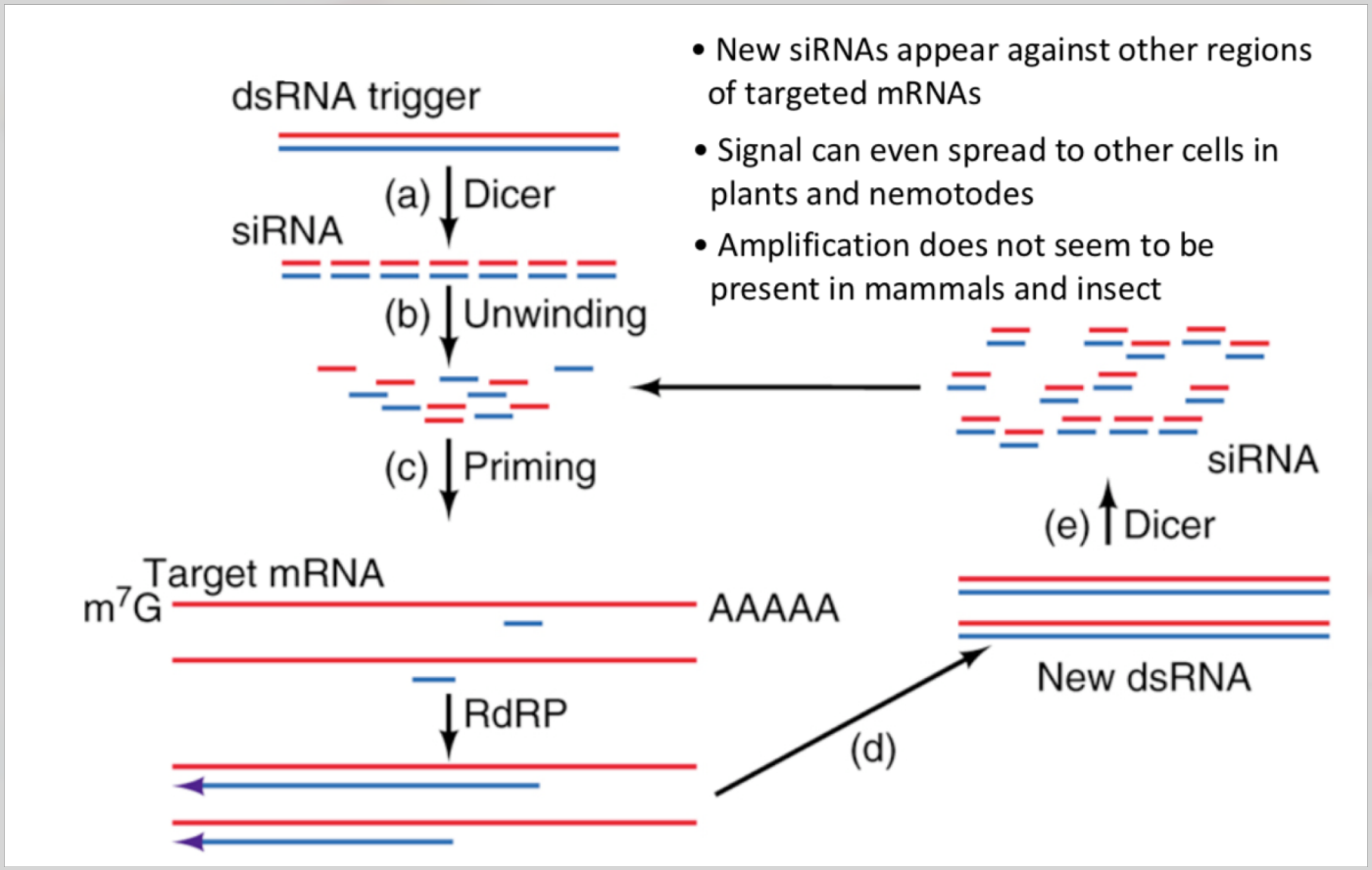


Moazed, *Nature* (2009) 457, 413-420

RNAi models


siRNA
(small interfering RNA)
miRNA
(microRNA)
https://wv
Cytoplasm
Nucleus
Dicer
siRNA
miRNA
Target mRNA
RISC
Target mRNA
Cutting
miRNP
Translation inhibition
siRNA
RNAi-induced methylation complex (RIMC)
New gene silencing system at the transcription level
DNA methylation
5'
3'
CG
GC
TATA
ATAT
CpG island
Gene
Gene silencing system at conventional post-transcription level
Controllo post-trascrizionale
(2 modalità)
(è il caso del colore della petunia)

In some organisms, siRNA signal is amplified and spread



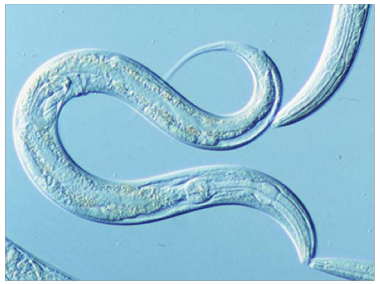
Biological functions - Immunity

In both juvenile and adult Drosophila, RNA interference is important in antiviral innate immunity and is active against pathogens such as Drosophila X virus.

A similar role in immunity may operate in C. elegans, as argonaute proteins are upregulated in response to viruses.

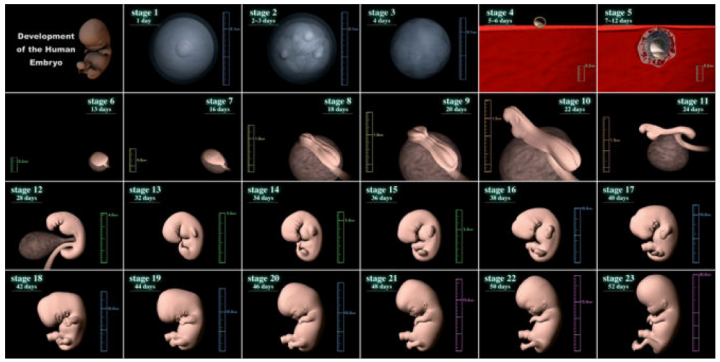
The role of RNA interference in mammalian innate immunity is poorly understood, and relatively little data is available.

However, the existence of viruses that encode genes able to suppress the RNAi response in mammalian cells may be evidence in favour of an RNAi-dependent mammalian immune response.



Biological functions – Gene regulation

Endogenously expressed **miRNAs**, including both intronic and intergenic miRNAs, are most important in **translational repression and in timing of morphogenesis and the maintenance of undifferentiated or incompletely differentiated cell types such as stem cells.**



The role of endogenously expressed miRNA in downregulating gene expression was first described in C. elegans in 1993.

In **plants**, the majority of **genes regulated by miRNAs are transcription factors**.

In many organisms, **including humans**, miRNAs have also been **linked to the formation of tumors and dysregulation of the cell cycle**. Here, miRNAs can function as both oncogenes and tumor suppressors.

RNA sequences (siRNA and miRNA) that are complementary to parts of a promoter can increase gene transcription, a phenomenon dubbed RNA activation.

Medical application– Gene regulation

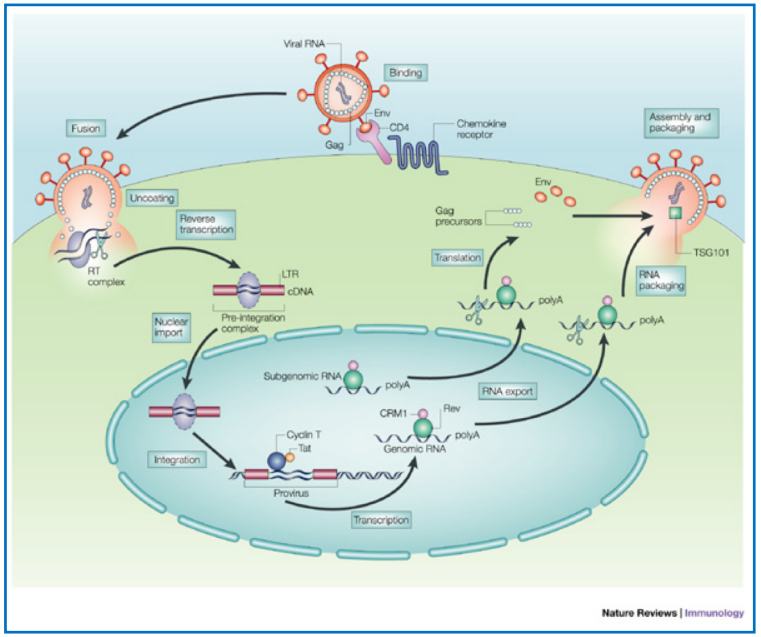
It is difficult to introduce long dsRNA strands into mammalian cells due to the interferon response, the use of siRNA mimics has been more successful.

First applications to reach clinical trials were: the treatment of macular degeneration and respiratory syncytial virus,

RNAi has also been shown to be effective in the reversal of induced liver failure in mouse models.

Other proposed clinical uses center on antiviral therapies:

- HSV type 2
- knockdown of host HIV receptors
- silencing of HIV, HAV,HBV and flu genes
- inhibition of measles viral replication.



Viruses like HIV-1 are particularly difficult targets for RNAi-attack because they are escape-prone, which requires combinatorial RNAi strategies to prevent viral escape.

RNAi
dalla teoria alla pratica di laboratorio

Come disegnare un siRNA in lab

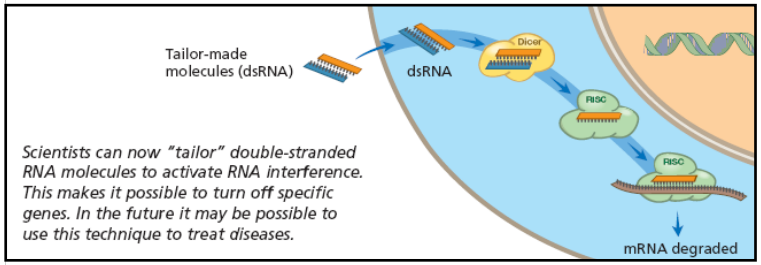
siRNA PROFILING TECHNOLOGIES

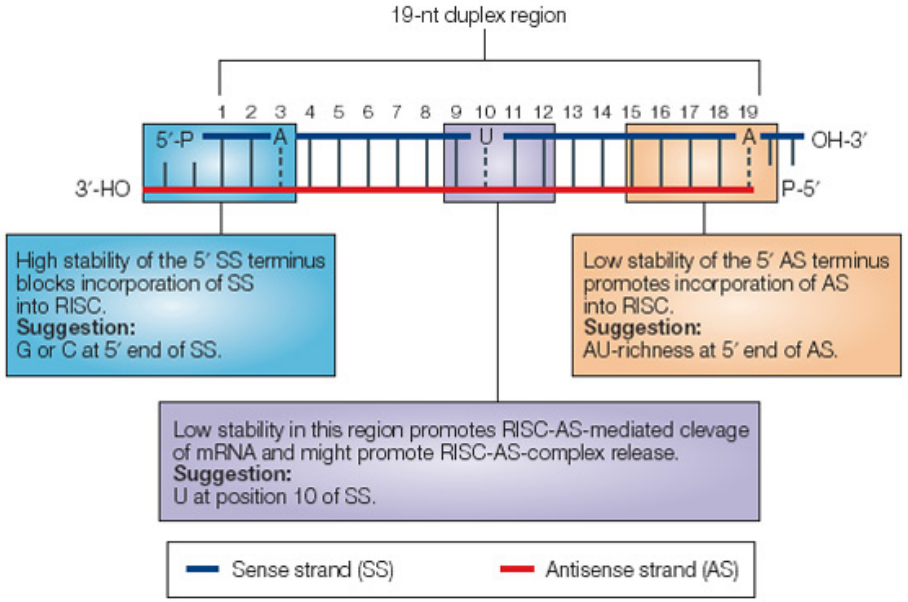
Disegno di siRNA

- il siRNA possiede un gruppo UU al terminale 3'
- 19 nucleotidi
- un contenuto in G/C < 50% è preferibile.

Algoritmo di nuova generazione

- Tm dell'siRNA
- Effetti della posizione nucleotidica
- Contenuto nucleotidico dei 3' overhangs
- Distribuzione nucleotidica
- Controllo della specificità

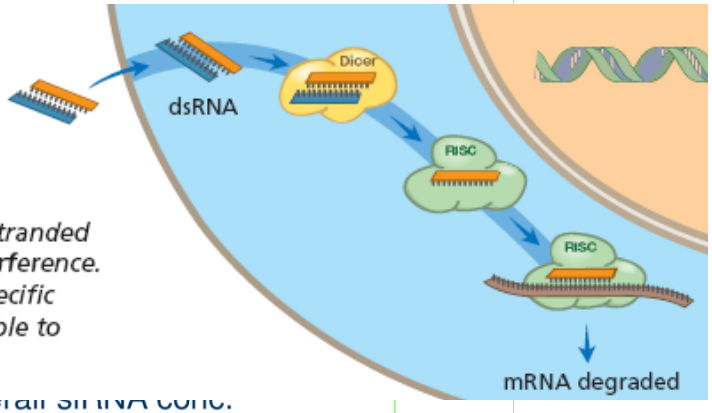




Come disegnare un siRNA in lab

Preventing Off-Target Effects

Overabundance of the siRNA activates the interferon pathway, as antiviral response



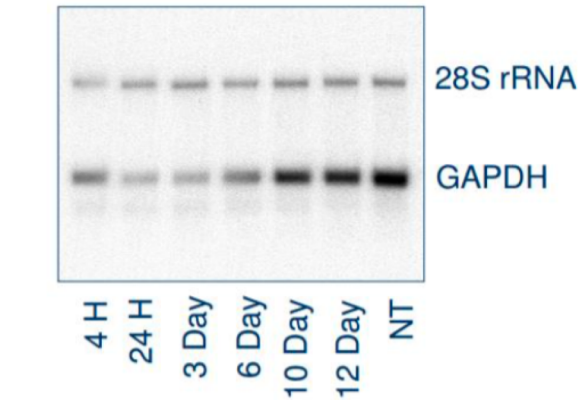
Low concent
- **chances o**
- **induction**

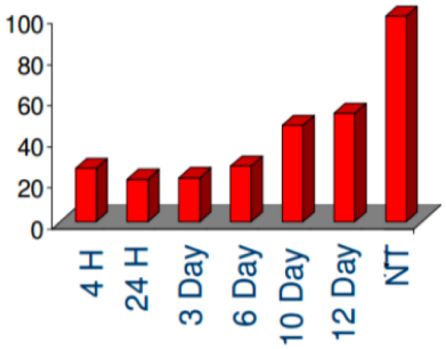
Scientists can now "tailor" double-stranded RNA molecules to activate RNA interference. This makes it possible to turn off specific genes. In the future it may be possible to use this technique to treat diseases.

It is curre
MIX of siRNAs that raise overall siRNA conc.

Verify specificity of RNAi effect by testing independent siRNAs to the same target

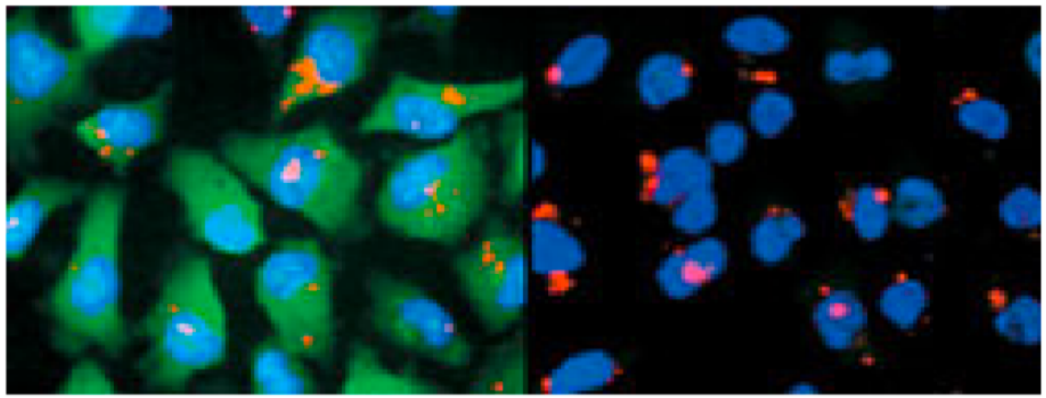
Durata del silenziamento transiente





CELLULE HeLa

In rosso: siRNA marcato
In blu: nuclei
In verde la proteina GAPDH



si RNA non specifico

si RNA contro GAPDH

Trattamento: 48h

Trasfezione con il siRNA: ottimizzazione delle condizioni

Prevenire effetti di spegnimento del target:

- **Basse concentrazioni** (~5-30 nM) di siRNA per minimizzare l'attivazione l'**interferon pathway** come risposta anti-virale
- E' preferibile usare un solo siRNA molto efficiente piuttosto che una miscela di siRNA meno potenti, la MIXTURE fa aumentare la concentrazione totale
- Usare RNAi specifici, dopo aver effettuato test di siRNA differenti sullo stesso mRNA bersaglio

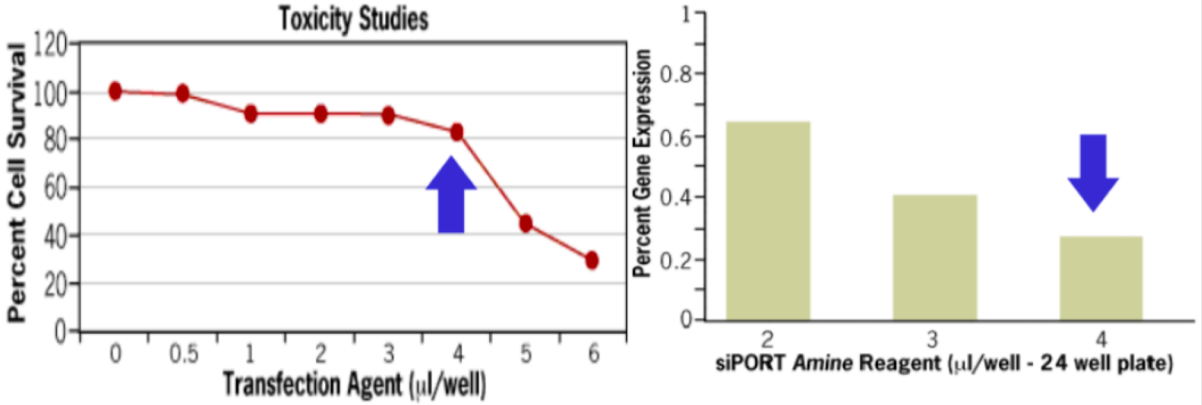
Trasfezione con il siRNA: ottimizzazione delle condizioni

Scelta del reagente trasfettante

Determining optimal amount of transfection reagent

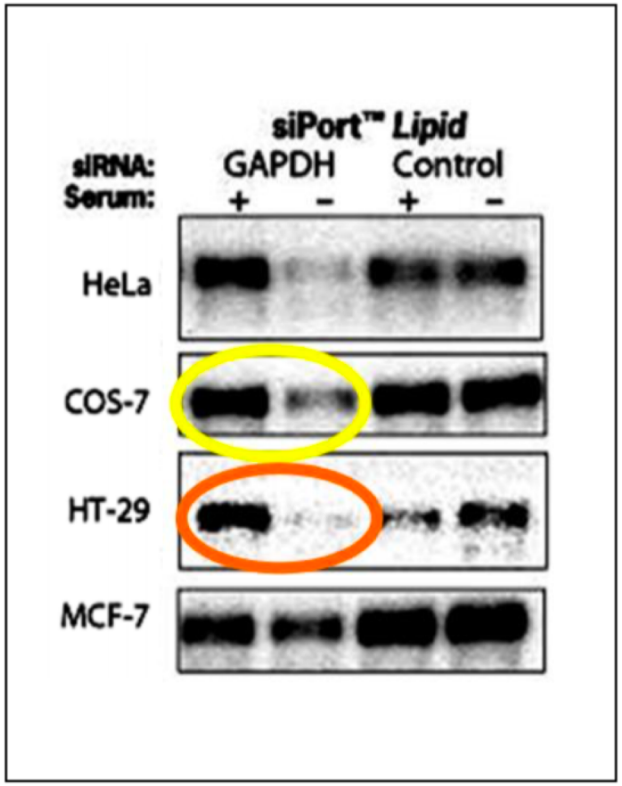
No cytotoxic

Effective



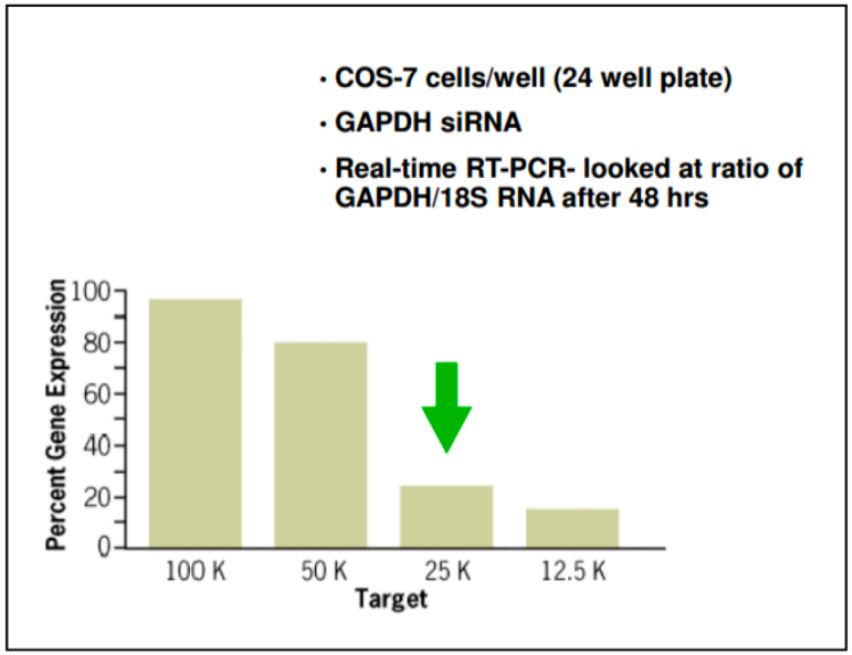
- COS-7 cells grown at 30K/well
- Transfected with GAPDH siRNA (10nM) after 24 hrs
- Real-time RT-PCR: GAPDH/18S rRNA (after 48 hrs)

Presenza ed assenza di siero

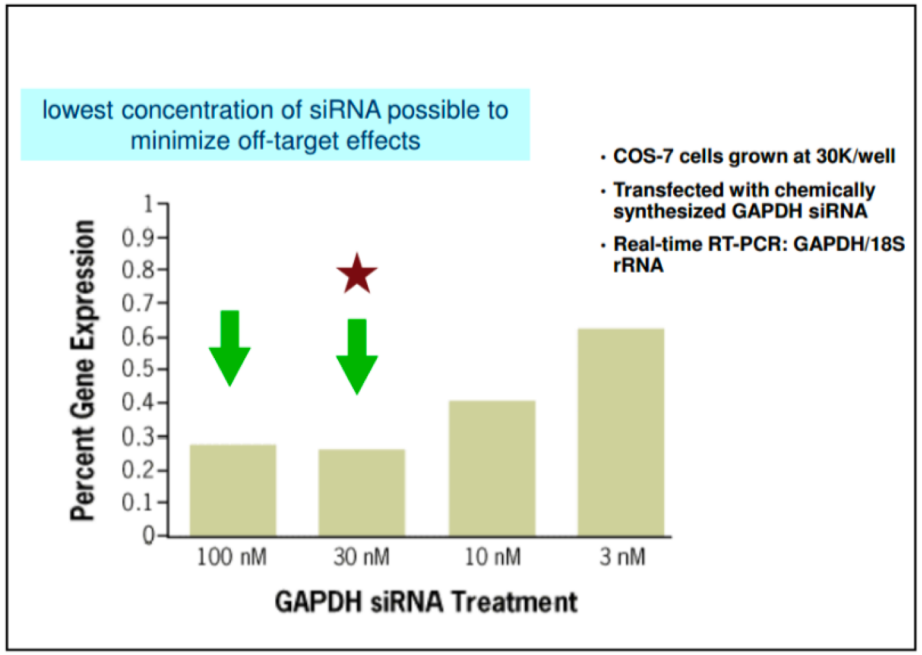


Trasfezione con il siRNA: ottimizzazione delle condizioni

Determinare la densità di semina delle cellule



Concentrazione di siRNA



Dal transiente alla trasfezione con vettori

PRO

- La trasfezione con siRNA è davvero molto efficiente in molti tipi di cellule
- Coi siRNA il silenziamento è immediato

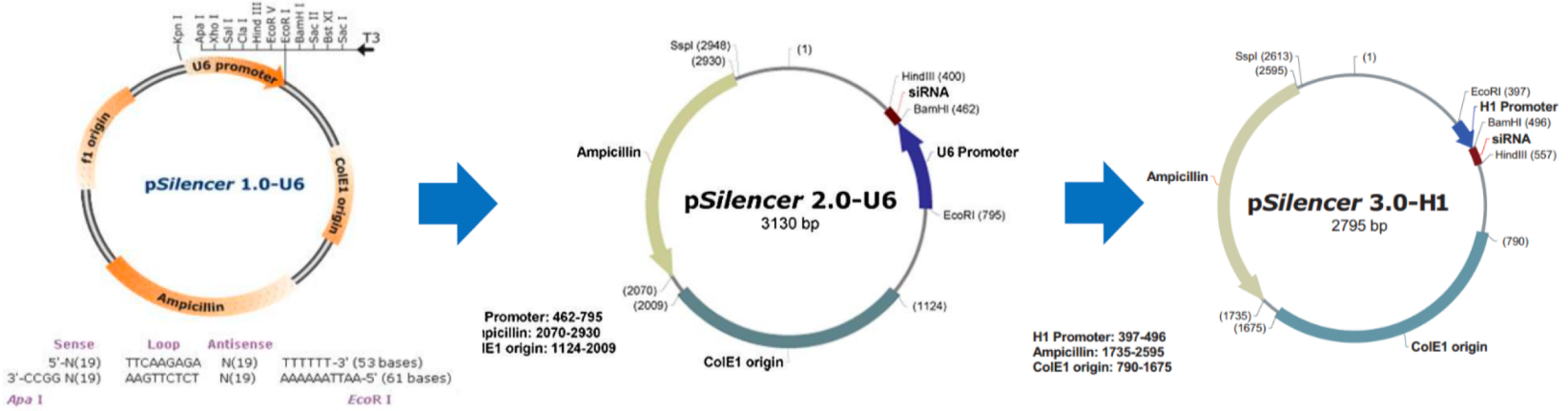
CONTRO

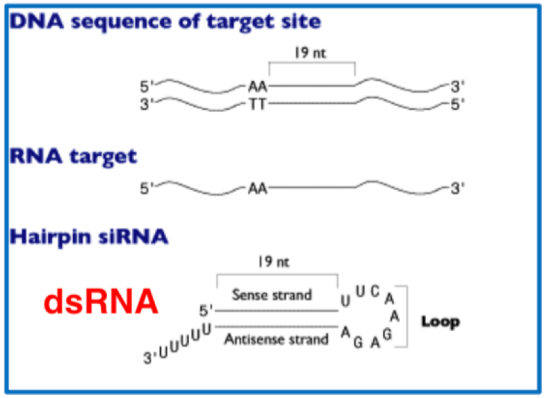
- Alcune cellule sono refrattarie alla trasfezione e la loro elettroporazione spesso causa morte cellulare
- I siRNA sono stabili, ma la trascrizione può risultare transiente se le cellule si duplicano molto in fretta diluendo il silenziamento e la vita media della proteina

Superamento del problema mediante……

Dal transiente alla trasfezione con vettori

Silenziamento tramite vettori a DNA



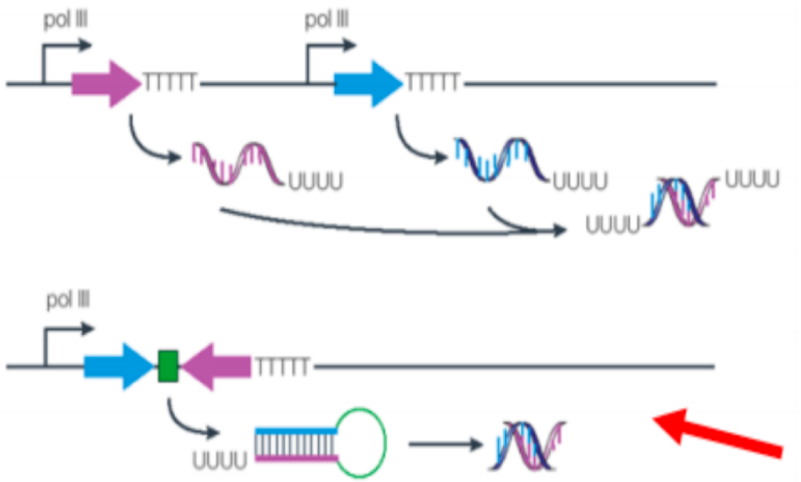


- Una sequenza stampo per un "hairpin siRNA" viene clonata in un opportuno vettore per trascrivere una molecola di RNA
- Produzione di siRNA *in vivo direttamente* nelle cellule trasfettate
- Trasfezione stabile nella linea cellulare di cui si vuol silenziare il gene target dell'RNAi
- **Silenziamento a lungo termine del gene target**

Dal transiente alla trasfezione con vettori

Sintesi di siRNA *in vivo*

- Nessuna sequenza richiesta dopo start site per la trascrizione
- TTTT: sufficiente per terminazione

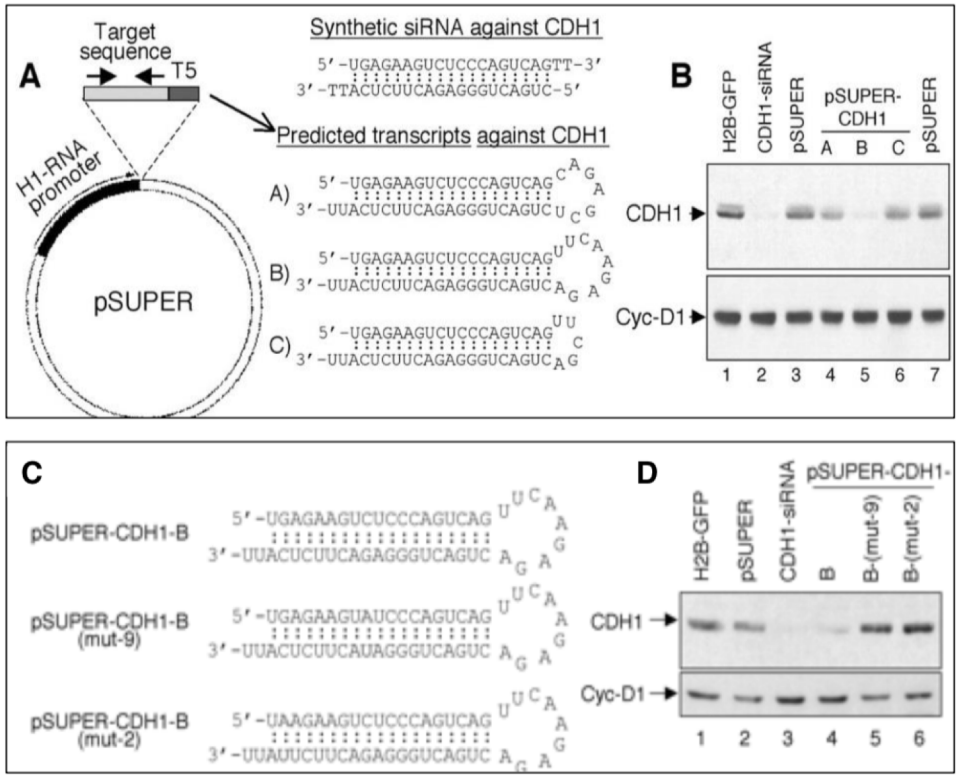


- **Clonati in vettori plasmidici** con promotori adatti per la produzione di RNA
- **Clonati in vettori virali**
  - **Oncoretrovirus**: MoMuLV o MSCV, le cellule devono duplicanti per poter essere infettate
  - **Lentivirus**: HIV-1, per infettare cellule quiescenti

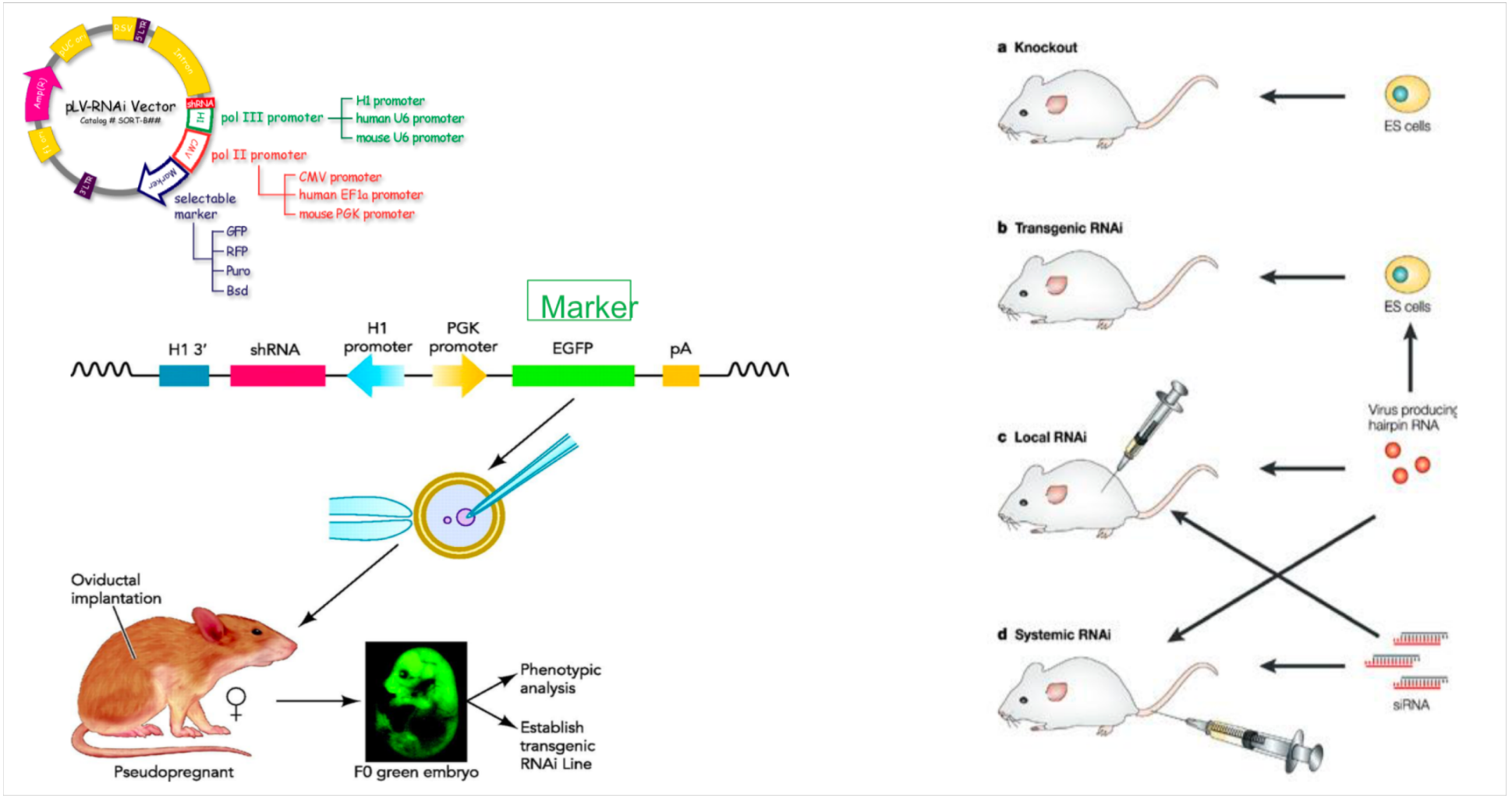
More efficiently processed by DICER!!

Espressione stabile di shRNA

Un sistema per l'espressione stabile di *short interfering RNA* in cellule di mammifero: vettore plasmidico



Espressione siRNA in vivo



Trasfezione con siRNA: le APPLICAZIONI

- Silenziamento genico specifico, efficiente e stabile nel tempo (economico e veloce)
- È un approccio di «genetica inversa»
- Screening delle funzioni genomiche (Genome-wide functional screenings)
- Terapia genica (es. antitumorale)
- Creazione di modelli per lo studio di agenti farmacologici (es. murini)
- Rivoluzione nello studio dei meccanismi di regolazione dell'espressione genica

siRNA library design

1. Grazie ai siRNA è possibile **silenziare uno alla volta** tutti i geni di un organismo.

2. Una tipica applicazione consiste **nell'identificare quali geni sono coinvolti in un certo processo**

3. Il punto di partenza è una **libreria di siRNA**, **specifica per un singolo gene del genoma**. Oggi esistono librerie in grado di coprire la maggior parte dei geni umani (≈ 20.000 siRNA).

siRNA "potenzialmente" funzionale:

- La regione target deve essere a valle del codone di inizio, ad una distanaza che varia da 50 a 100bp.
- Lunghezza compresa fra 19-22 bp.
- Contenuto in GC fra il 35-55%
- 2-nt 3´ overhangs di residui di uridina
- 5´-phosphate and 3´-hydroxyl group.

- **Stability**
- **Access to RISC**

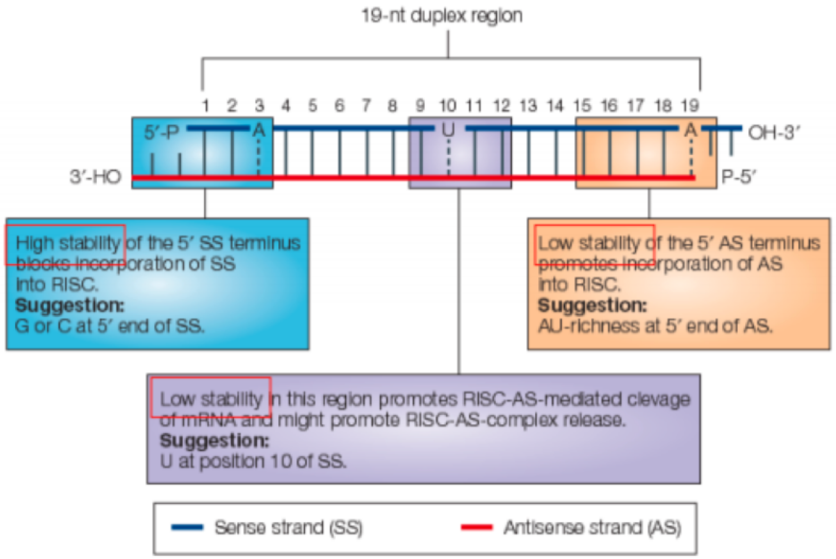
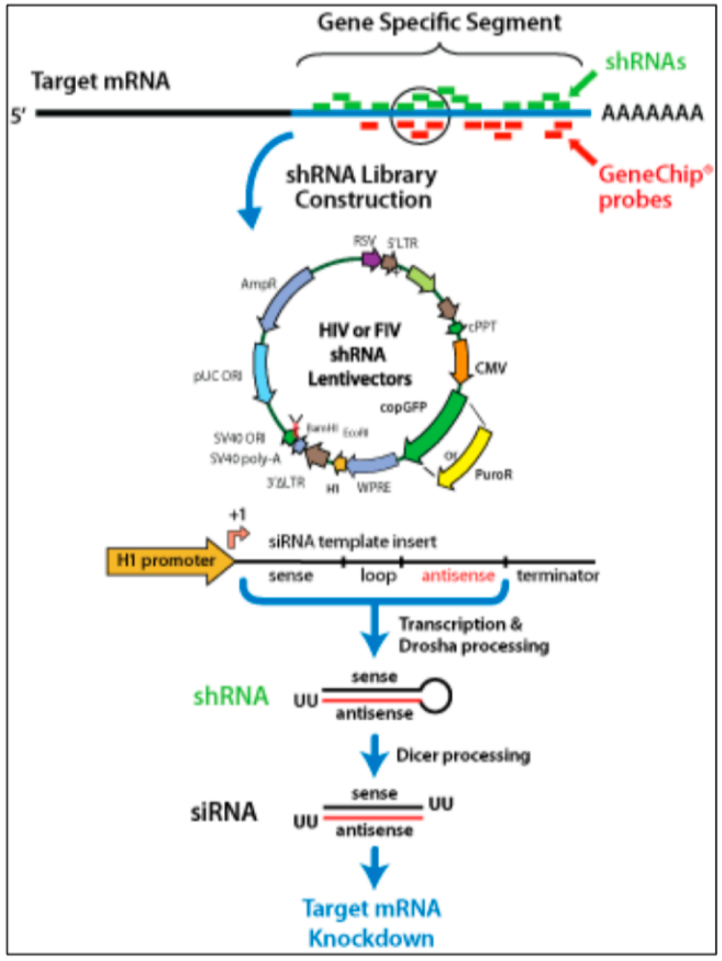


Figure 2 | **The generation of effective siRNA.** A small interfering RNA (siRNA) is a 21–23-nucleotide (nt) dsRNA that contains: a 19-nt duplexed region, symmetric 2–3-nt 3′ overhangs, and 5′-phosphate (P) and 3′-hydroxyl (OH) groups. The positions of each nucleotide in the 19-nt duplexed region of the sense strand are shown. On the basis of recently established design criteria, an effective siRNA has high stability at the 5′ terminus of the sense strand (blue box), lower stability at the 5′ antisense terminus (orange box) and at the cleavage site (purple box). In addition, the sequence-specific preferences at the following positions on the sense strand are important: the presence of an A at position 19, an A at position 3, a U at position 10 (BOX 2 lists other parameters). RISC, RNA-induced silencing complex.

Mittal, Nature Review Gentic, 2004

Rational siRNA design for RNA interference. Nature Biotechnology 22, 326 - 330 (2004)

Screening con siRNA

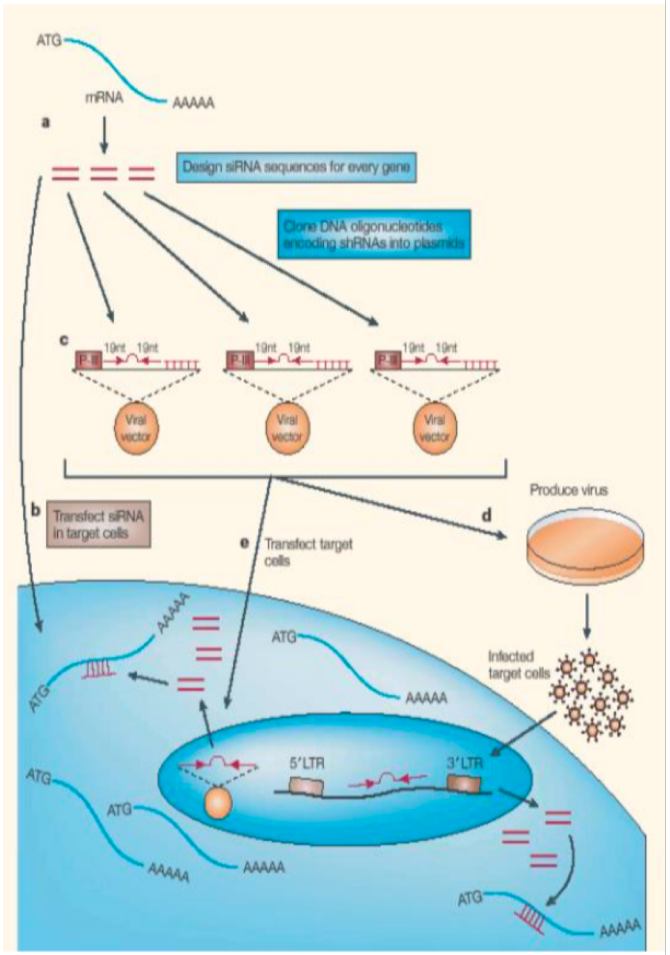


1. **Sintesi della libreria di siRNA, specifica per un singolo gene del genoma**. Oggi esistono librerie in grado di coprire la maggior parte dei geni umani (≈ 20.000 siRNA).

2. Trasfezione delle cellule con una libreria di siRNA diretti contro uno specifico gene target

3. Analisi espressione genica rispetto al controllo non trasfettato (Northern blotting; RT-PCR; gene-expression profiling) o ricerca della proteina analisi con saggi cellulari (FACS; ELISA)

4. **Identificazione del vettore con l'inserto in grado di inibire il gene target**



Limitazioni dei siRNA

Impossibile studiare geni essenziali
per la sopravvivenza cellulare (*housekeeping*) e sviluppo

Sviluppo di nuovi vettori
per l'espressione condizionale-inducibile dei shRNA

(tet OFF/ON H1 and U6 promoter system)

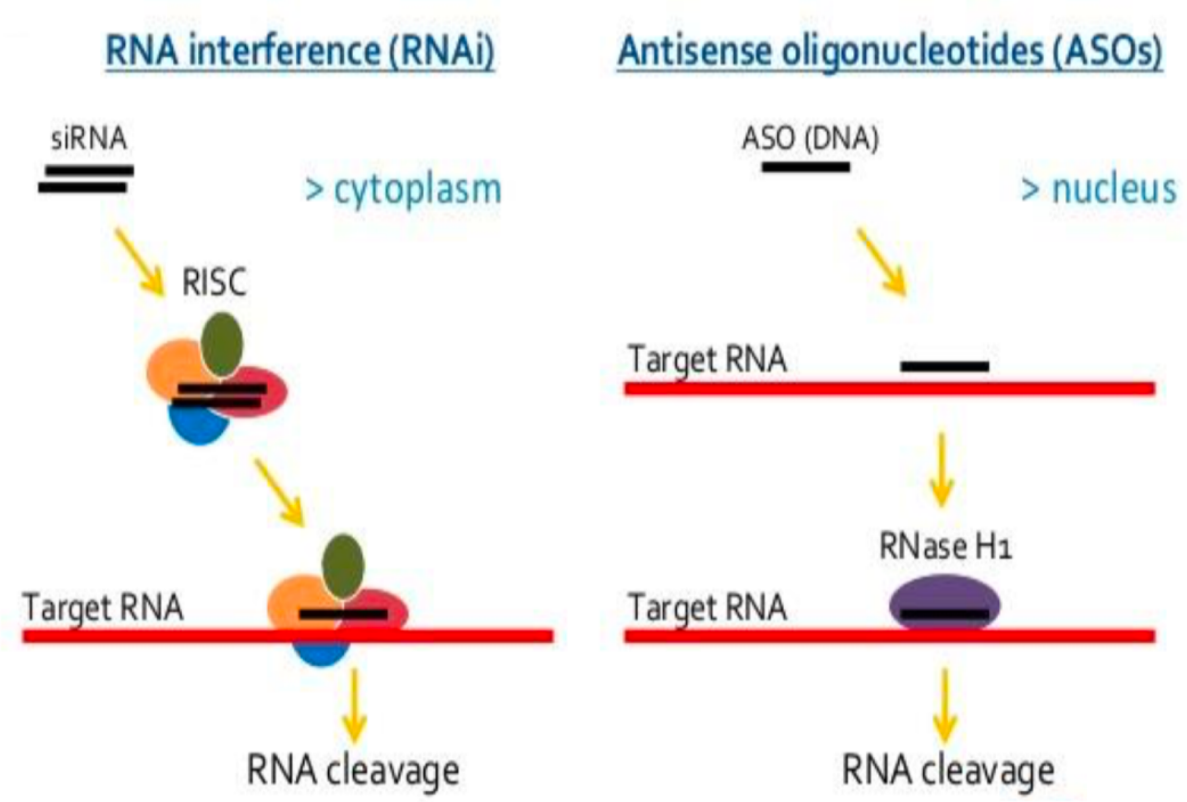
siRNA vs. oligonucleotidi antisense (a ssDNA)

Similarità

- Lunghezza
- Metodologia di *delivery* comune
- Induzione di silenziamento genico a livello post-trascrizionale
- Digestione di mRNA bersaglio da parte di endonucleasi
- Possibilità di stabilizzare con basi modificate
- Bio-distribuzione simile

Differenze

- Doppio filamento *vs.* singolo filamento
- Maggiore stabilità del siRNA
- Maggiore efficacia delle molecole in cellule in coltura
- Meccanismo d'azione mediato da RISC

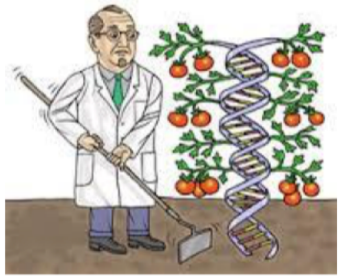


CRISPR/Cas technology

Introduction to CRISPR/Cas - Genetic Engineering

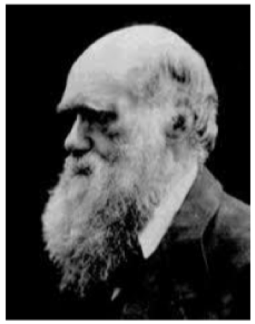
"The deliberate modification of the characteristics of an organism by manipulating its genetic material."

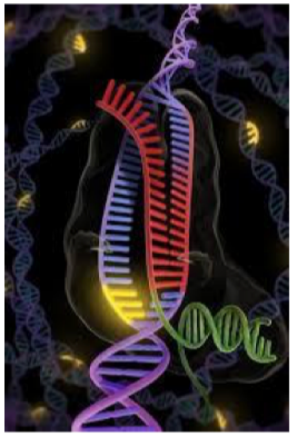
- Research
- Medicine (Protein/Enzyme production)
- Agriculture (Crops)
- Industrial Biotechnology (Biofuel production)
- Entertainment



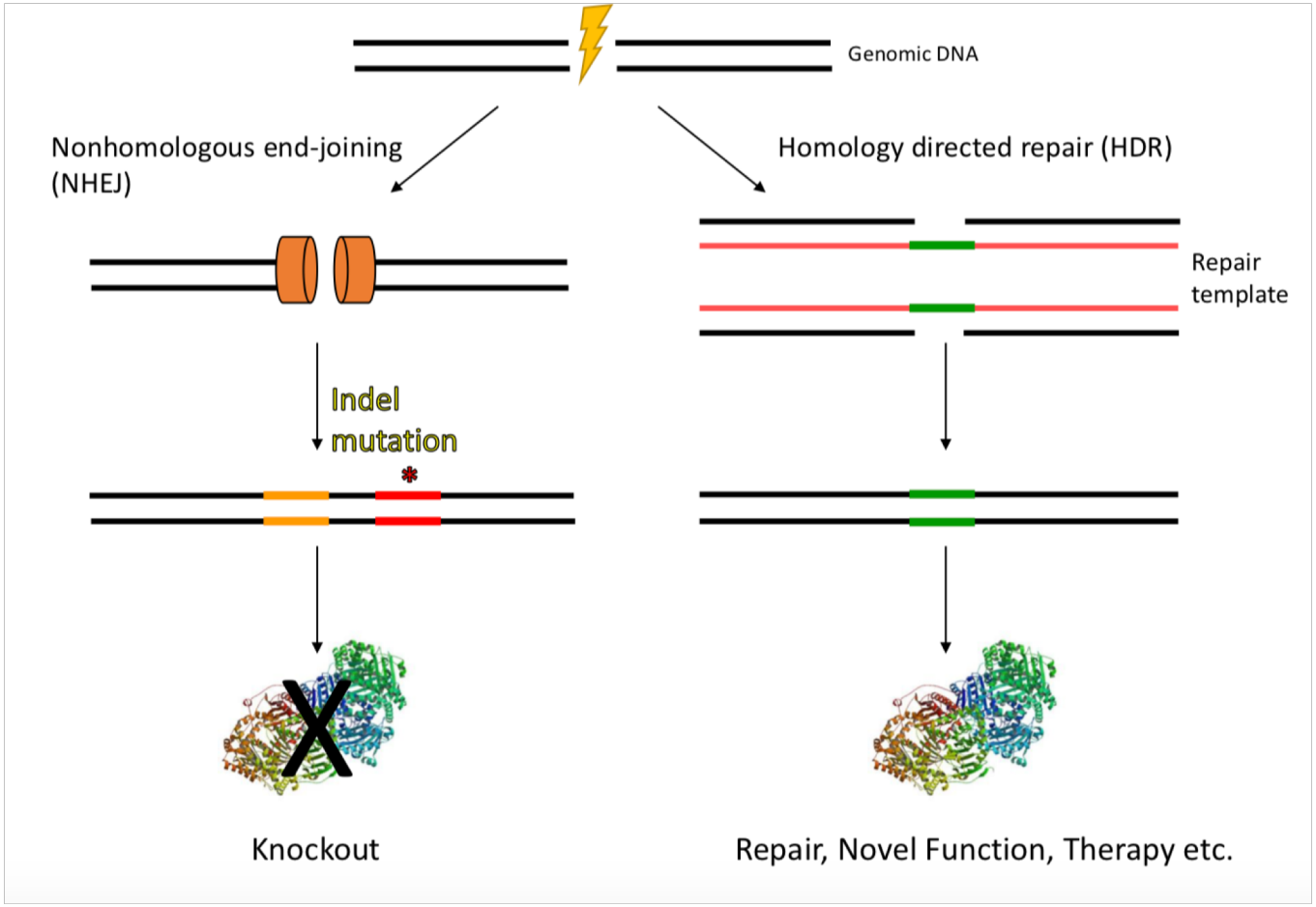
The way towars genetic Engineering

Rules	Information	Basics	Genome editing
1859 Darwin *"Origin of Species"*	1944 Avery–MacLeod–McCarty *DNA as the genetic material*	1970 Restriction Enzymes	Zink Fingers
1856-66 Mendel *"Mendelian inheritance"*	1953 Watson, Crick and Franklin *DNA structure*	1977 Sanger Sequencing	TALENs
1871 Mieska Nucleic acids	1961-1967 Genetic code	1983 PCR	CRISPRs
		2003 Human Genome Project	

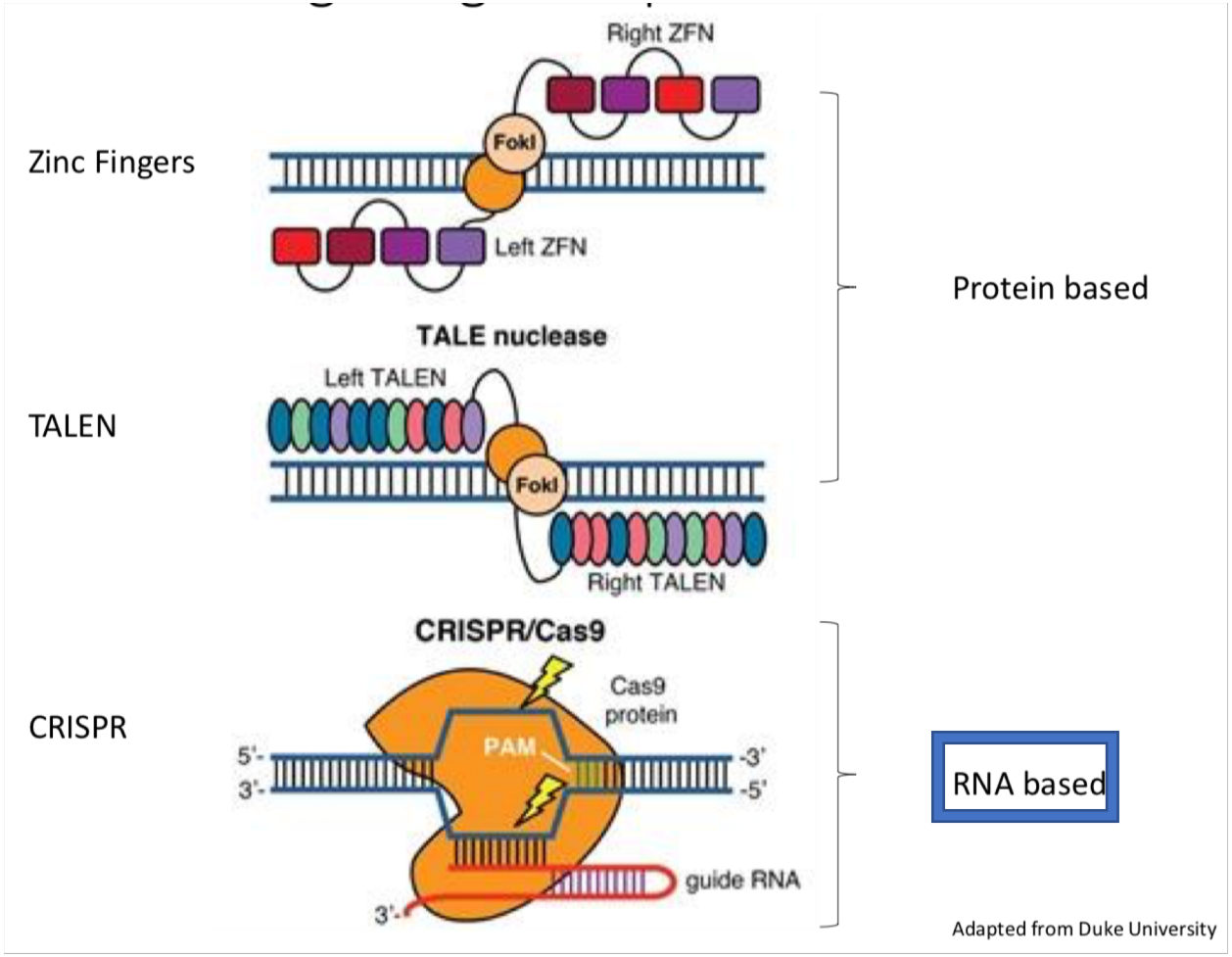




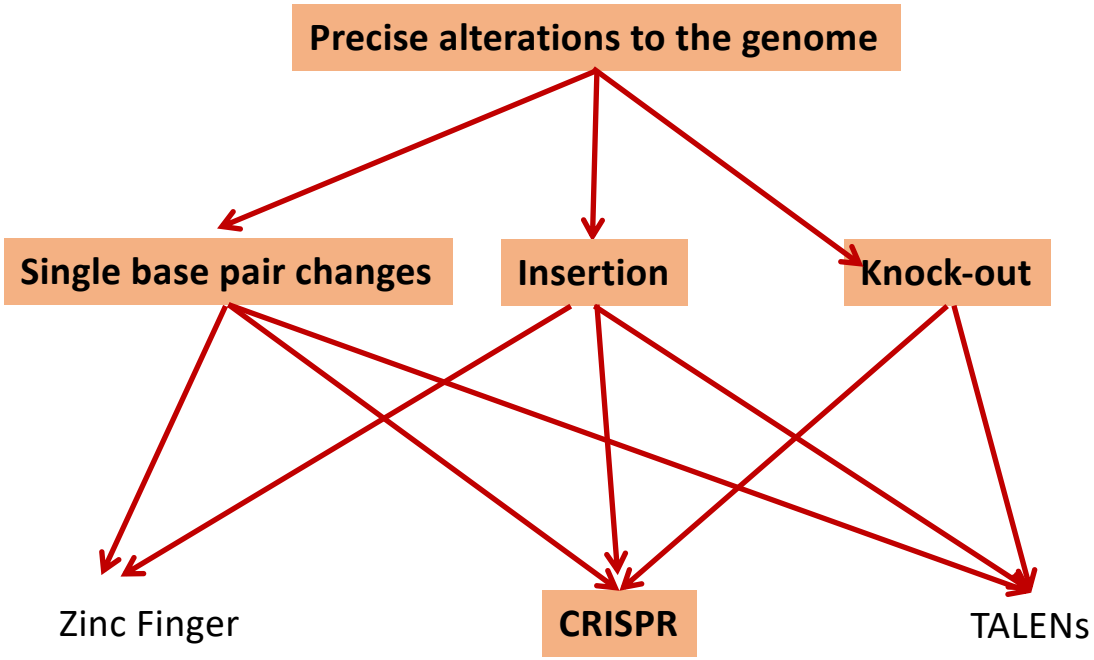
Genetic editing uses DNA repair pathways



Genome Editing using Site Specific Nucleases

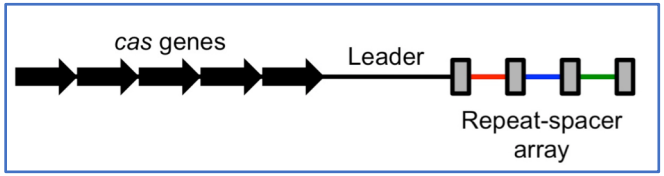


Genome Editing



CRISPR-Cas – Adaptive immune system in bacteria

A: CRISPR (**clustered regularly interspaced short palindromic repeats**) is a family of DNA sequences found within the genomes of prokaryotic organisms such as bacteria and archaea. These sequences are derived from DNA fragments from viruses that have previously infected the prokaryote and are used to detect and destroy DNA from similar viruses during subsequent infections. Hence these sequences play a key role in the antiviral defense system of prokaryotes.



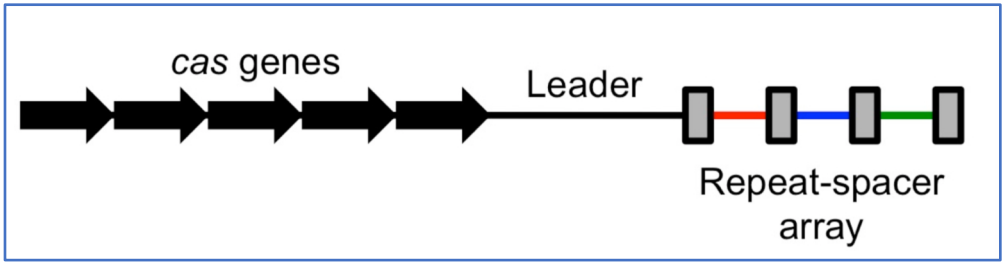
+

B. Cas9 (or "CRISPR-associated 9") is an enzyme that uses CRISPR sequences as a guide to recognize and cleave specific strands of DNA that are complementary to the CRISPR sequence.

Cas9 enzymes together with CRISPR sequences form the basis of a technology known as **CRISPR/Cas9** that can be used to edit genes within organisms

CRISPR-Cas – A. CRISPR locus

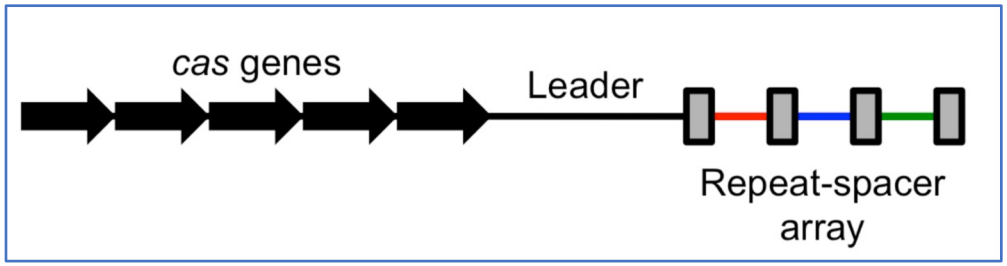
Simplified diagram of a CRISPR locus



The **three major components** of a CRISPR locus are shown: 1. ***cas* genes**, 2. **leader** and 3. **repeat-spacer array**. For the repeat-spacer array, repeats are shown as grey boxesc (typically range in size from 28 to 37 base pairs (bps), though there can be as few as 23 bp and as many as 55 bp), and spacers are colored bars

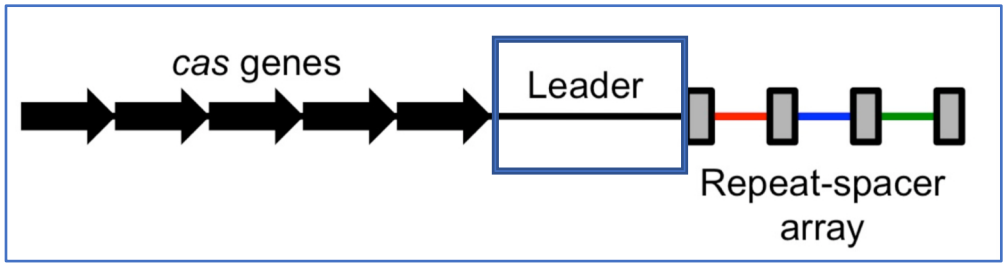
CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus



CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus

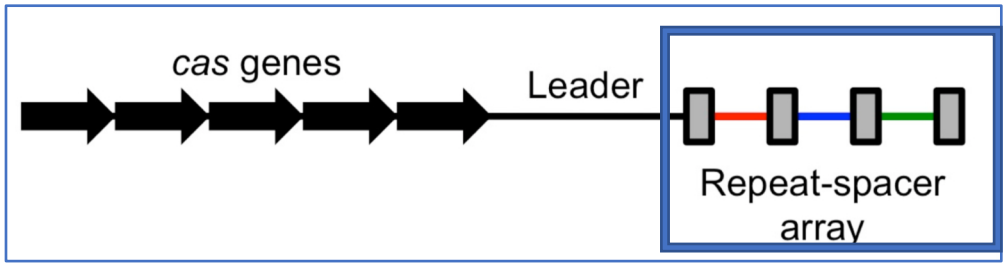


Leader Sequence

This sequence is an A-T reach sequence

CRISPR-Cas – A. CRISPR locus

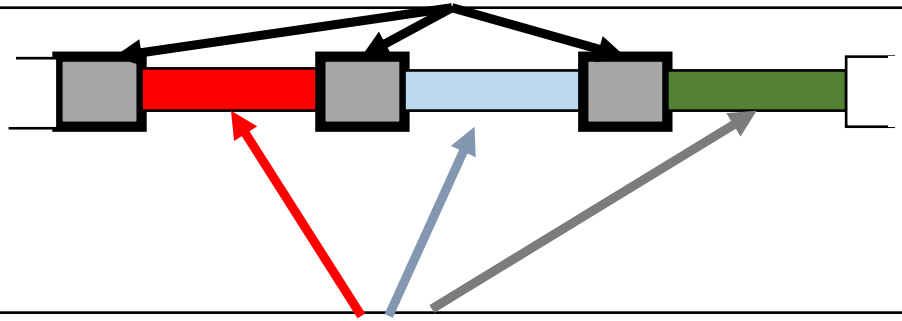
Simplified diagram of a CRISPR locus



REPEAT Spacer ARRAY

Shorts palindromic **repeats** (i.e. this is the same DNA sequence repeated in different places, gray boxes). These are part of the **bacterial genome** and typically range in size from 28 to 37 base pairs

Diagram of CRISPR locus in bacterial genome

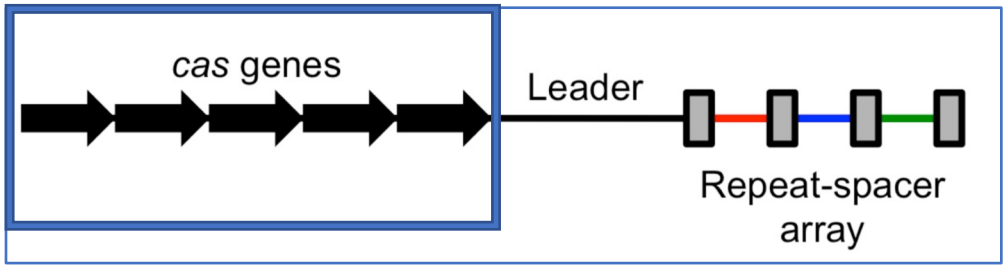


There are usually fewer than 50 units of the repeat-spacer sequence in a CRISPR array.

These bits/**spacers** are derived from **bacteriophage genome** (range 21 to 72 bp) and each one is different and these provide the guidance system for the adaptive immune system

CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus



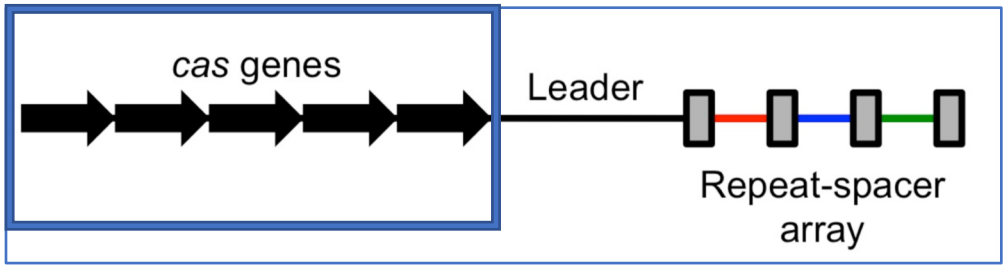
CAS genes

There are several other important regions of the bacterial DNA that are also always associated with the CRISPR locus and these provide the means for the palindromic repeat and the bacteriophage DNA sequences to actually destroy the bacteriophage.

These are called CRISPR Associated Sequences i.e. **Cas** genes

CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus



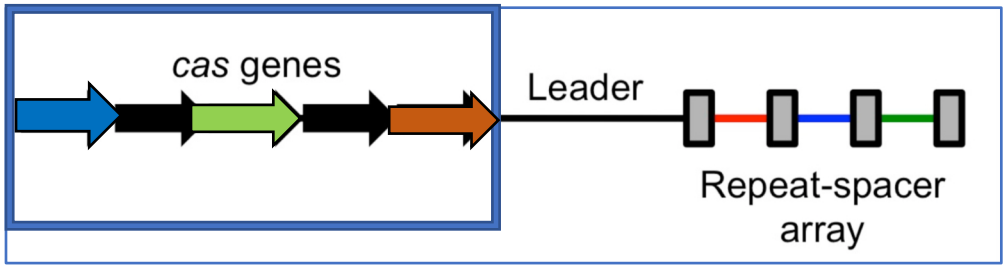
CAS genes

There are several other important regions of the bacterial DNA that are also always associated with the CRISPR locus and these provide the means for the palindromic repeat and the bacteriophage DNA sequences to actually destroy the bacteriophage.

These are called CRISPR Associated Sequences i.e. **Cas** genes

How does this genetic material in CRISPR locus then manage to kill bacteria ?

Simplified diagram of a CRISPR locus



CAS genes

The system can be slighty different in different types of bacteria but the best studies one is ***Streptococcus pyogenes*** so we will focus on that one

For the sake of simplicity lets focus on the 3 Cas genes (now colored arrows) most important for genetic engineering;

Codes for a **trans-activating CRISPR RNA (tracrRNA**) that will help in the process of ensuring the whole process only cuts bacteriophage DNA

Codes **for a protein** that is a nuclease that cuts DNA but only if it is given a very specific set of signals to do so (otherwise it would potentially damage the bacteria's own DNA). The most common one used in genetic engineering approaches is called Cas9. ; additional Cas1 and Cas2 are responsible for spacer geration

Codes for a very specific piece of RNA (**crRNA or guide RNA**) that will help in the process of ensuring the whole process only cuts bacteriophage DNA

For now lets not worry about the other genes in the Cas locus