microRNAs: historical aspects, biogenesis, mode of action, biological functions

Hervé Seitz
(herve.seitz@igh.cnrs.fr)

IGH (CNRS and université de Montpellier)

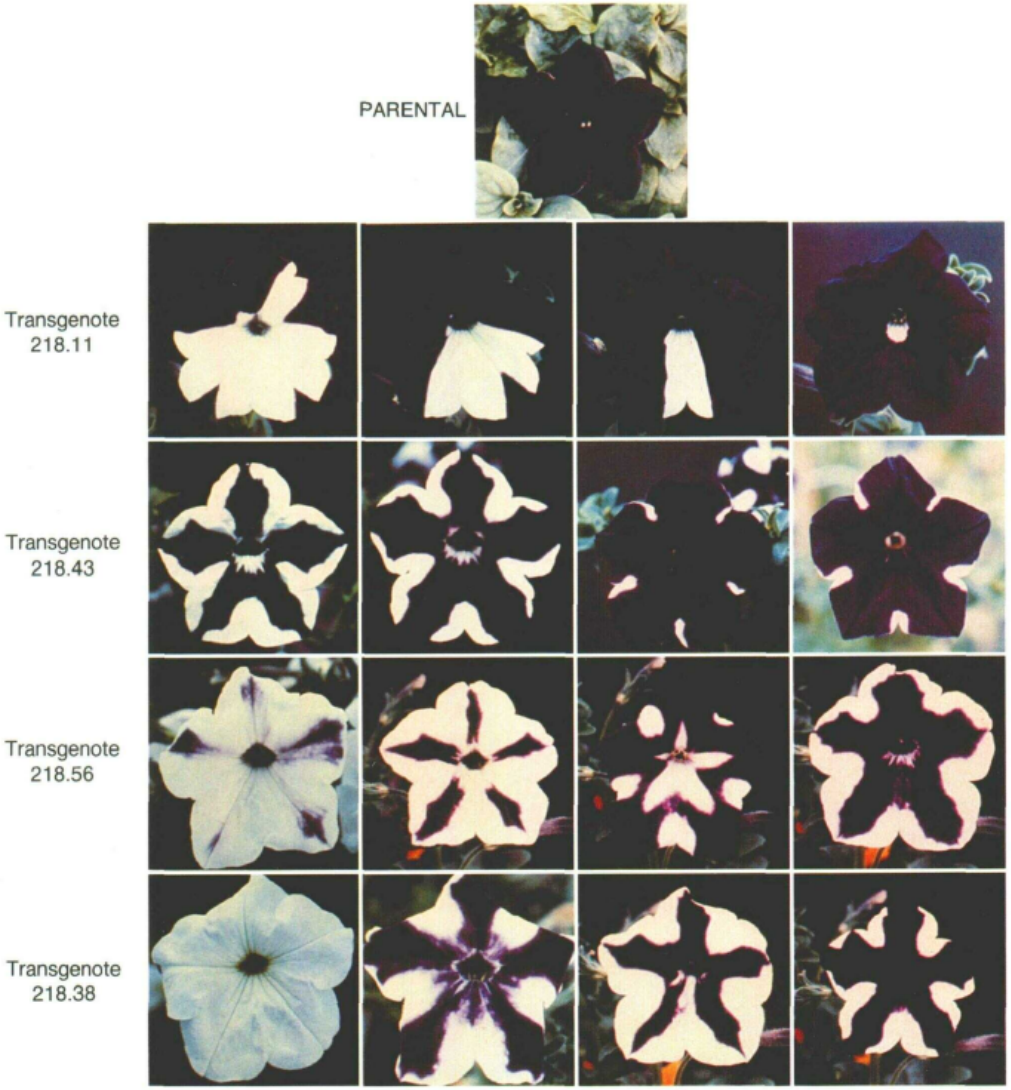
December 5, 2024

This slideshow is available at:
https://www.igh.cnrs.fr/equip/Seitz/UM2024.pdf

Initial discoveries

(taken from van der Krol *et al.*, 1990)

Initial discoveries



(taken from Napoli *et al.*, 1990)

Initial discoveries

Also observed in Tobacco, Rape, Tomato, *Arabidopsis thaliana*.

Initial discoveries

Also observed in Tobacco, Rape, Tomato, *Arabidopsis thaliana*.

Named “homology-dependent silencing”: repressed sequences are homologous to the triggering transgene.

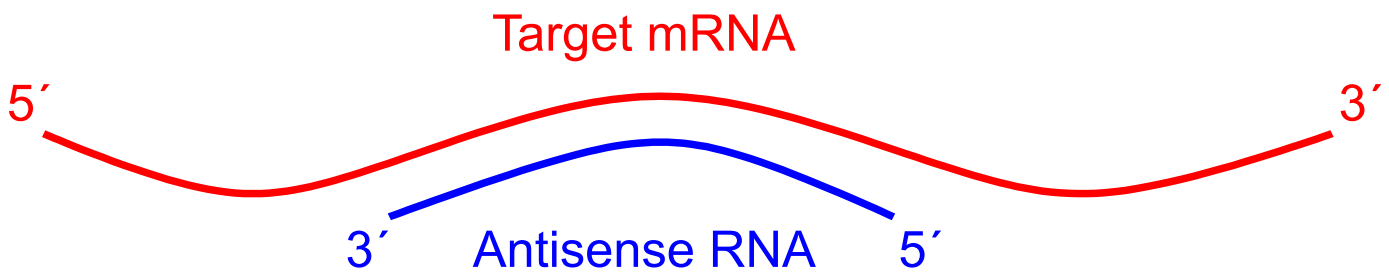
Initial discoveries

Also observed in Tobacco, Rape, Tomato, *Arabidopsis thaliana*.

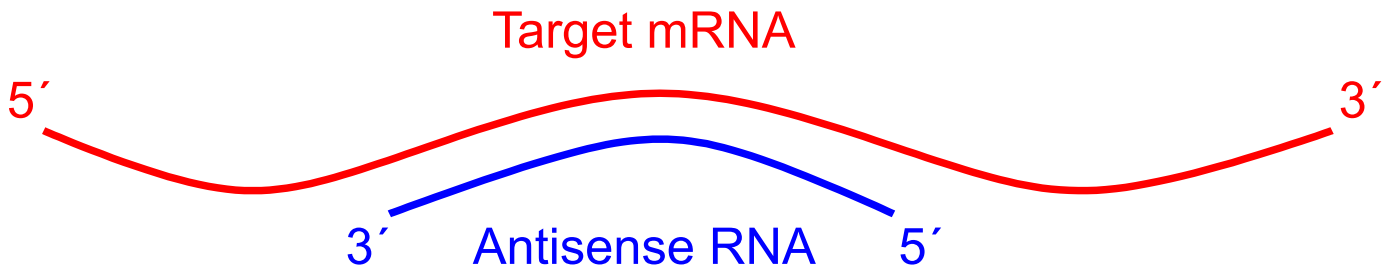
Named "homology-dependent silencing": repressed sequences are homologous to the triggering transgene.

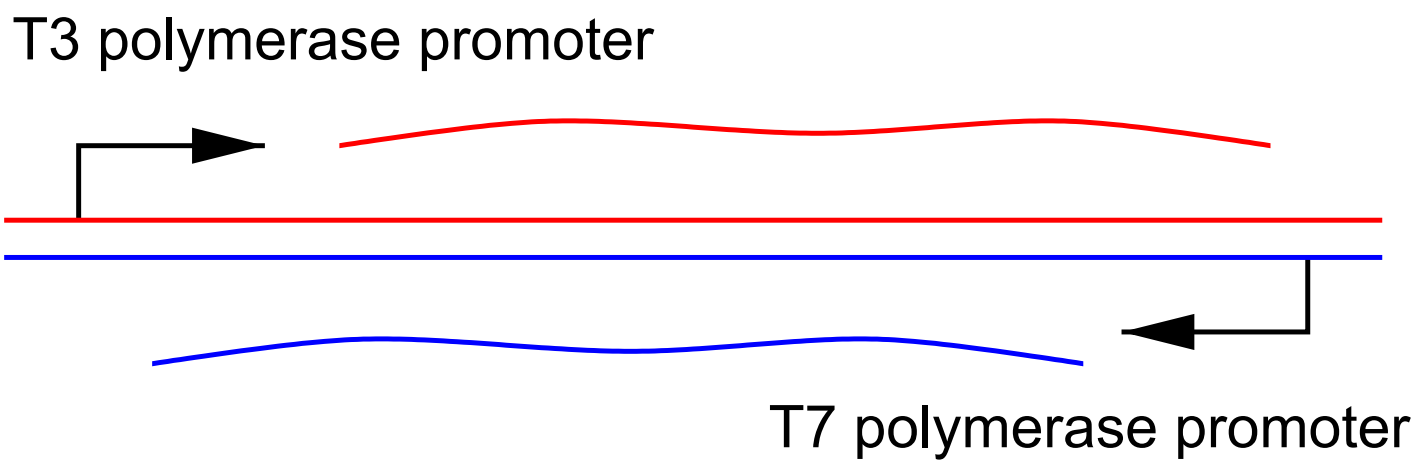
Homology sensing: nucleic acids are probably involved (proteins would not have a universal homology-sensing property) (Baulcombe, 1996).

Identification of the trigger



Identification of the trigger





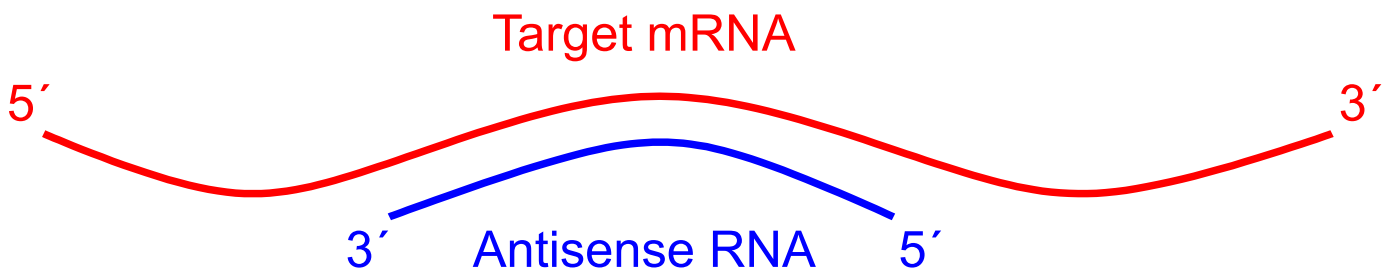
Untreated worms

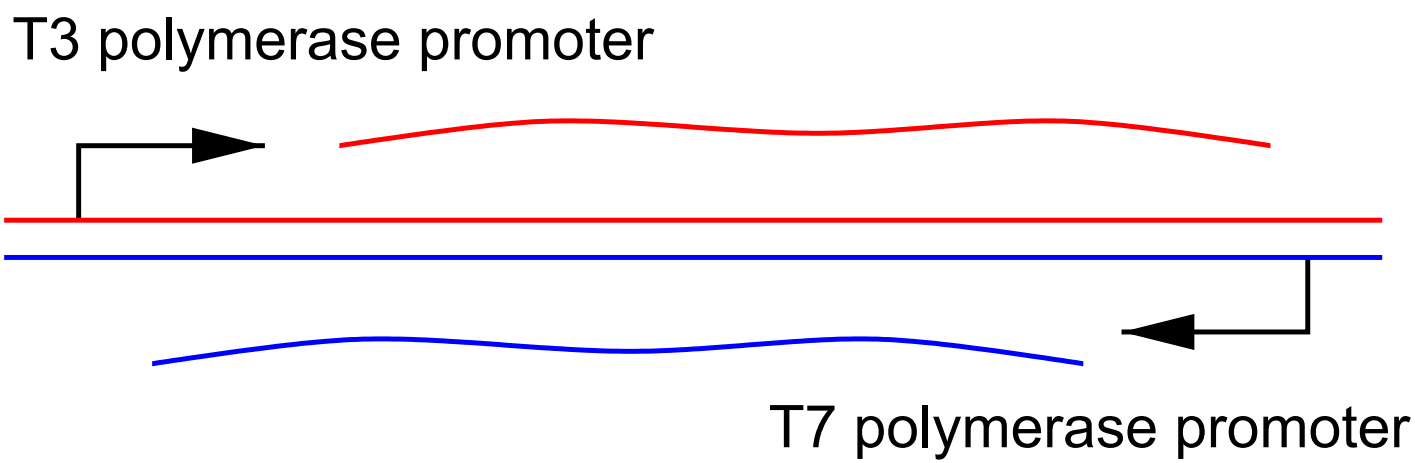
(video by Darryl Conte Jr., Worcester, MA, USA)

After injection of *unc-22* double-stranded RNA

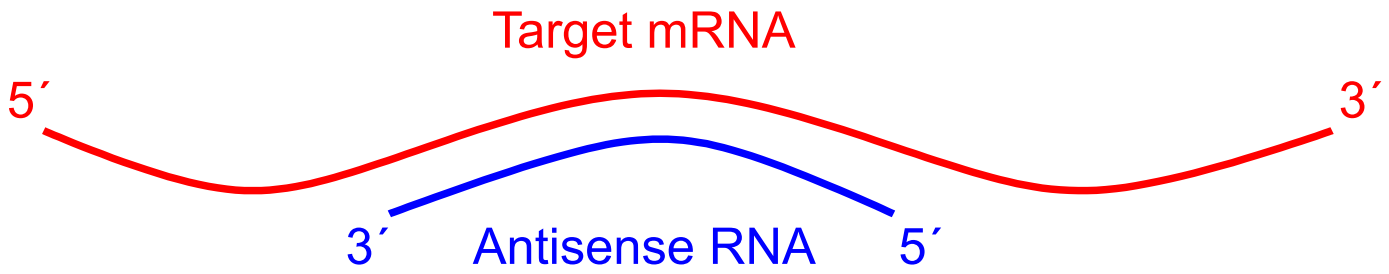
(video by Darryl Conte Jr., Worcester, MA, USA)

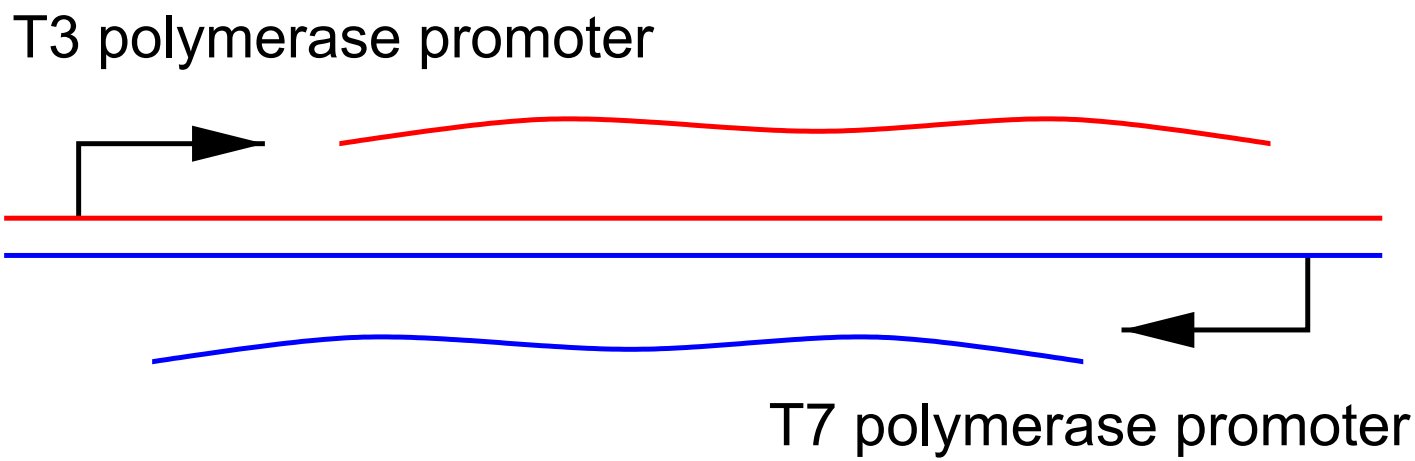
Identification of the trigger





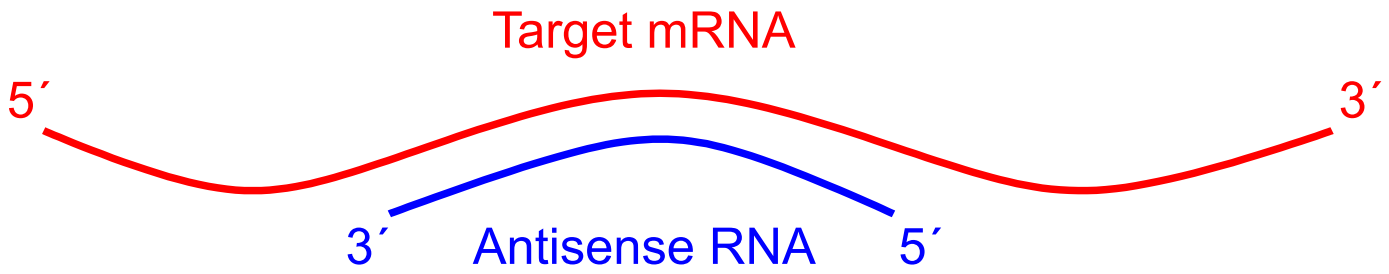
Identification of the trigger

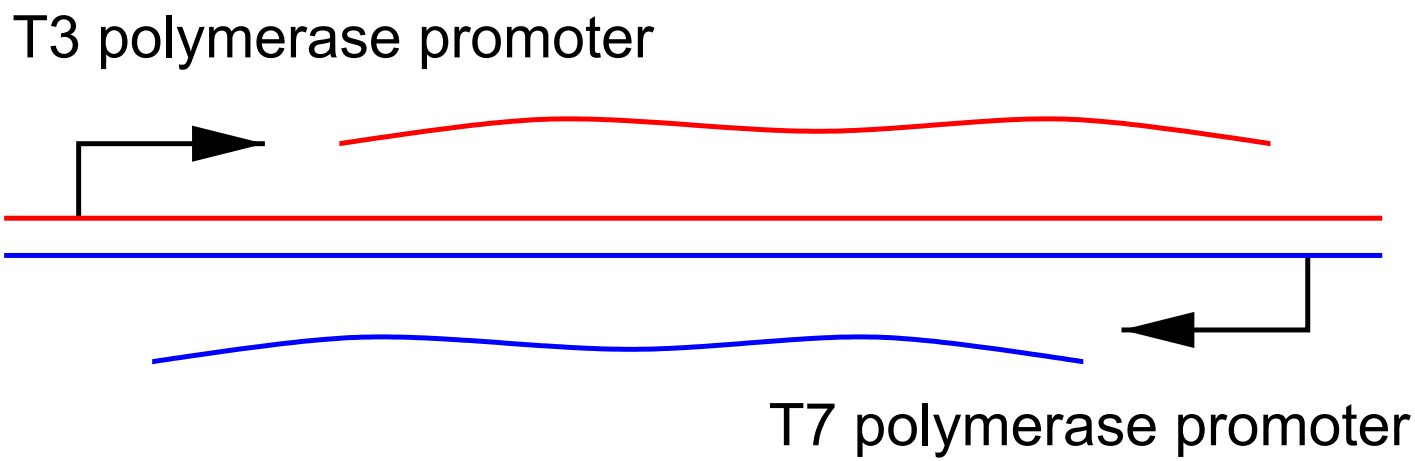




$\longrightarrow$ The trigger: double-stranded RNA (Fire *et al.*, 1998).

Identification of the trigger





$\longrightarrow$ The trigger: double-stranded RNA (Fire *et al.*, 1998).

RNAi-associated small RNAs: “siRNAs”

In vitro recapitulation in *Drosophila* embryo lysate ($\longrightarrow$ biochemistry).

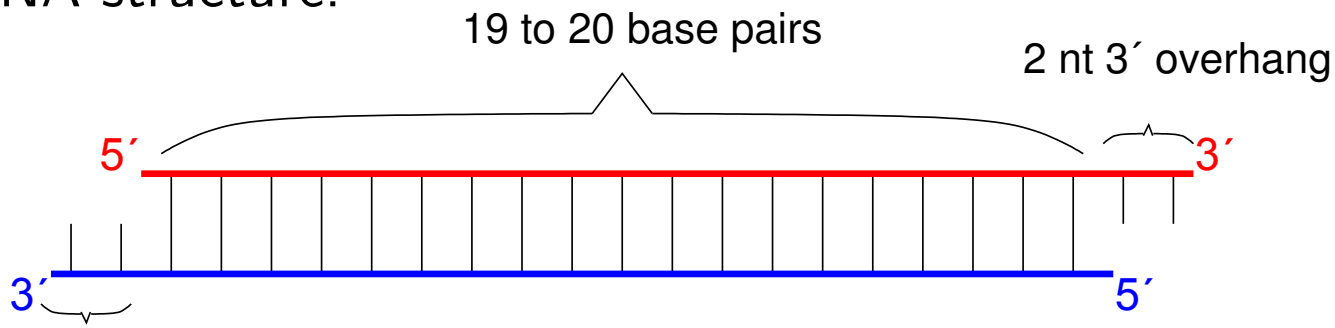
RNAi-associated small RNAs: “siRNAs”

In vitro recapitulation in *Drosophila* embryo lysate ($\longrightarrow$ biochemistry).

In plants (Hamilton and Baulcombe, 1999), then in the *in vitro Drosophila* system (Zamore *et al.*, 2000) and in cultured *Drosophila* cells (Hammond *et al.*, 2000): observation of small double-stranded RNAs ($\approx$ 20–25 bp) homologous to both the trigger and the target: “siRNAs” (small interfering RNAs).

RNAi-associated small RNAs: “siRNAs”

siRNA structure:

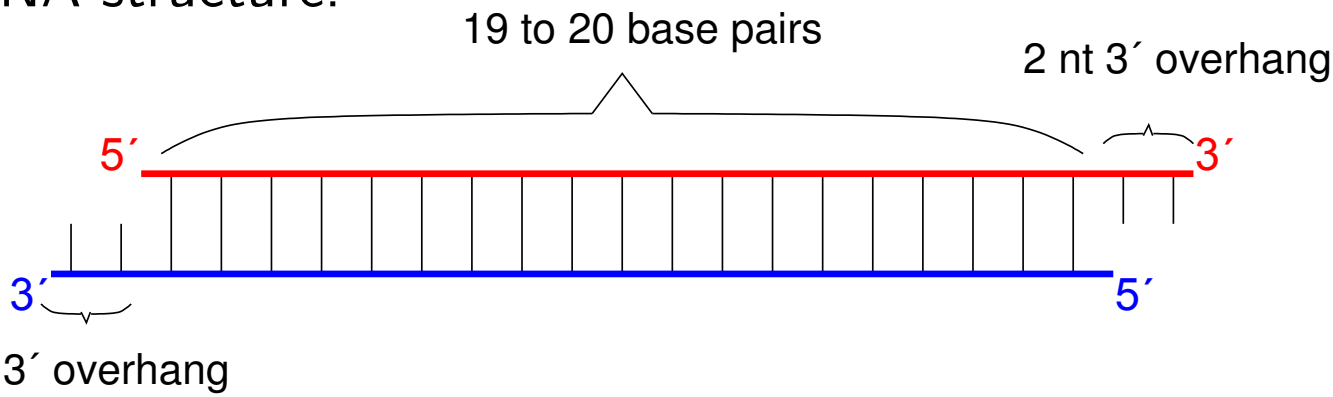


2 nt 3´ overhang

5´ monophosphate, 3´ hydroxyl

RNAi-associated small RNAs: “siRNAs”

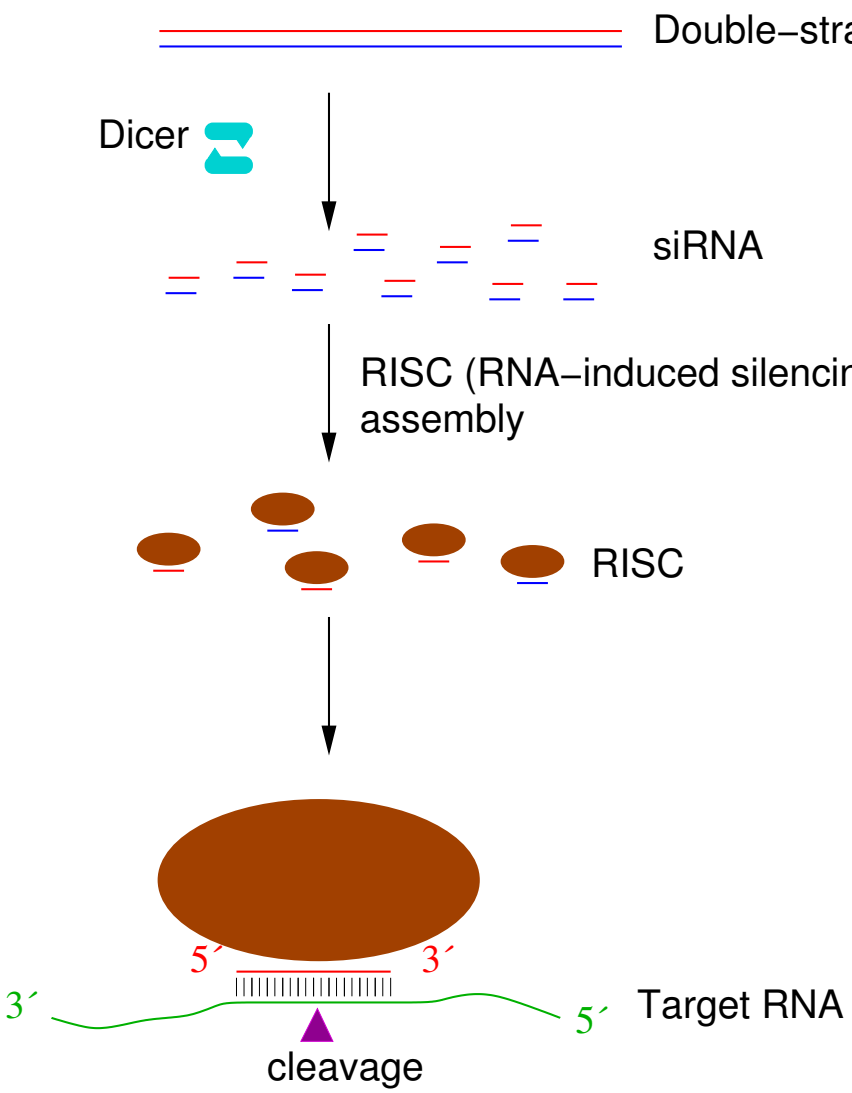
siRNA structure:



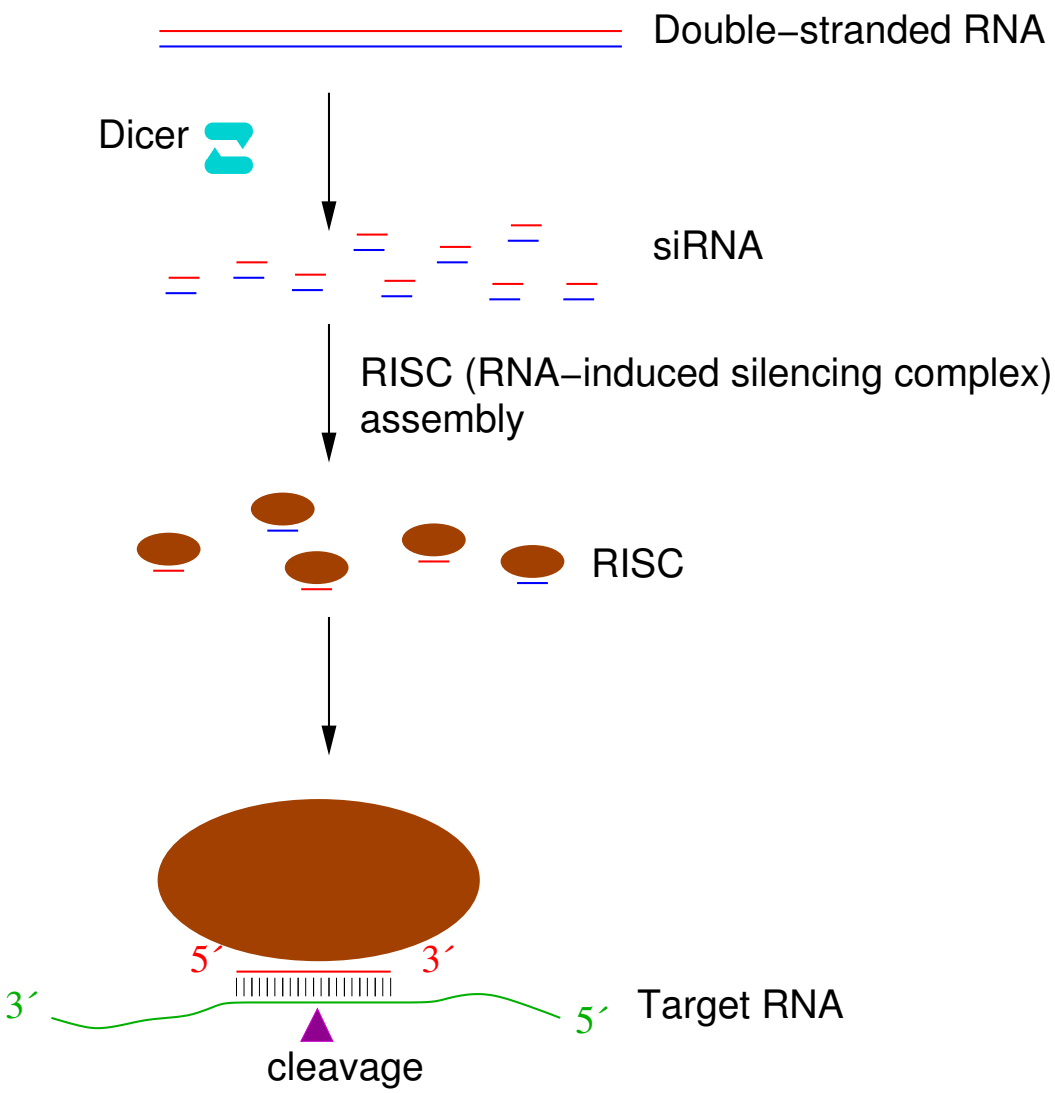
5´ monophosphate, 3´ hydroxyl

siRNAs come from the digestion of the double-stranded RNA trigger by the “Dicer” enzyme (member of the RNase III family) (Bernstein *et al.*, 2001).

Mechanism of the RNAi response

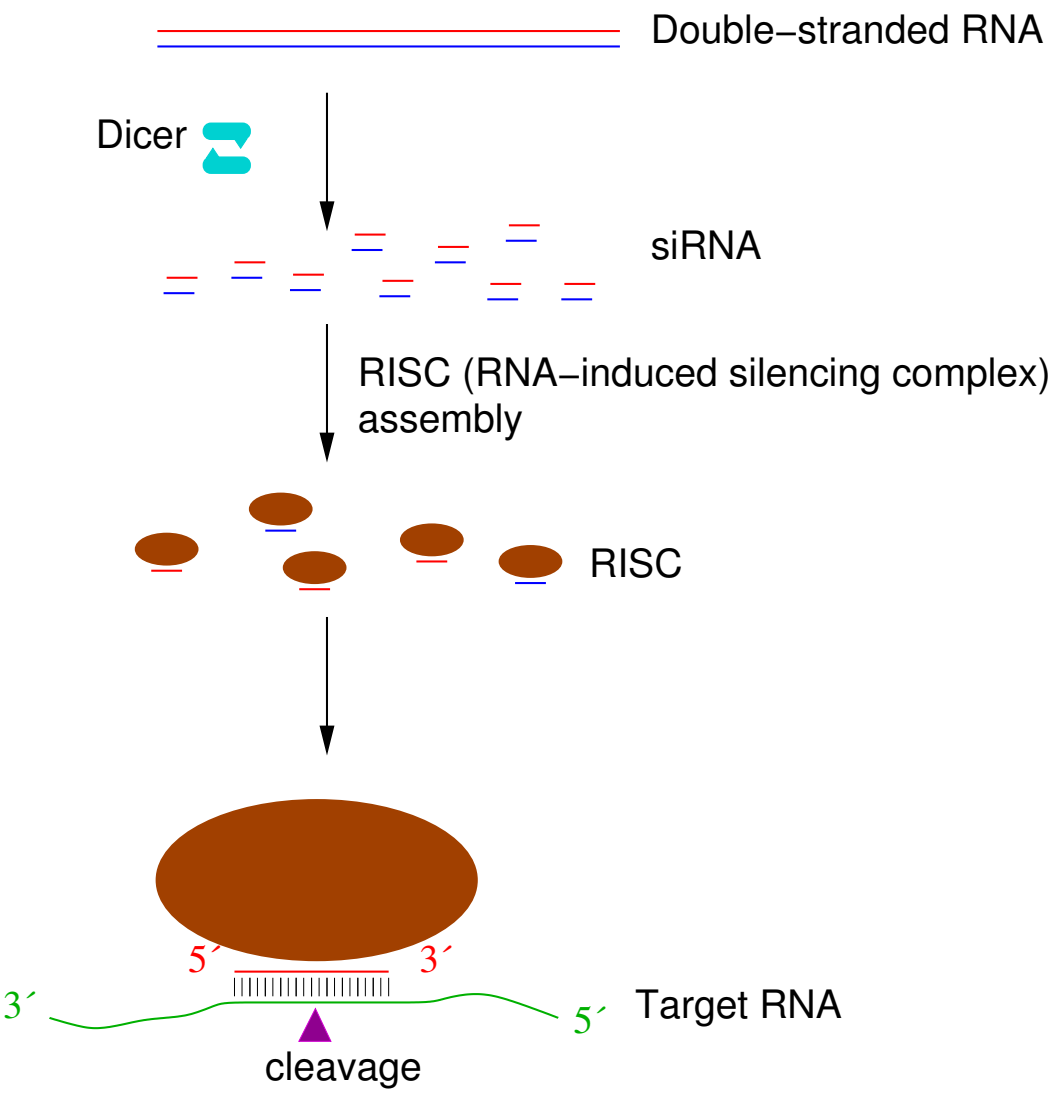


Mechanism of the RNAi response



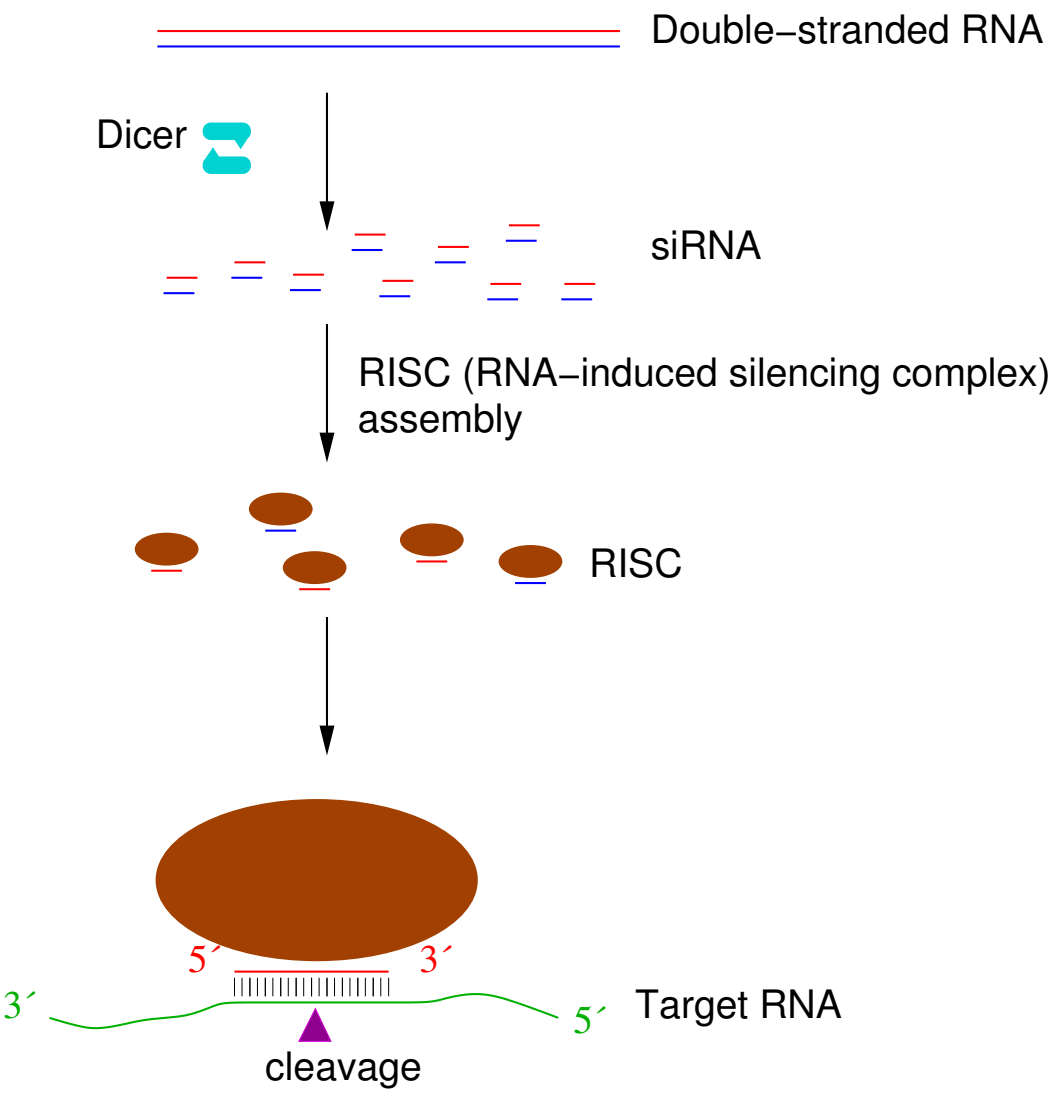
Cleavage between nucleotides (nt) facing guide nt 10 and 11.

Mechanism of the RNAi response



Cleavage between nucleotides (nt) facing guide nt 10 and 11.
Extensive complementarity is needed.

Mechanism of the RNAi response



Cleavage between nucleotides (nt) facing guide nt 10 and 11.
Extensive complementarity is needed ($\neq$ perfect!).

Applications of RNAi

Applications: specific repression of genes of interest.

Applications of RNAi

Applications: specific repression of genes of interest.
Problem in mammals: long dsRNAs trigger the interferon response. Use siRNAs instead (Elbashir *et al.*, 2001).

Applications of RNAi

Applications: specific repression of genes of interest.
Problem in mammals: long dsRNAs trigger the interferon response. Use siRNAs instead (Elbashir *et al.*, 2001).

Clinical application: in progress (targeting dominant mutant alleles).

Applications of RNAi

Applications: specific repression of genes of interest.
Problem in mammals: long dsRNAs trigger the interferon response. Use siRNAs instead (Elbashir *et al.*, 2001).

Clinical application: in progress (targeting dominant mutant alleles).

In the lab: mammalian systems did not have any simple method for specific gene silencing.

Natural functions for RNAi

Natural functions for RNAi

In plants and insects: antiviral effect (see session of October 19).

Natural functions for RNAi

In plants and insects: antiviral effect (see session of October 19). Apparent paradox in fungi: antiviral activity may have prompted the loss of RNAi in some species (Drinnenberg *et al.*, 2011).

Natural functions for RNAi

In plants and insects: antiviral effect (see session of October 19). Apparent paradox in fungi: antiviral activity may have prompted the loss of RNAi in some species (Drinnenberg *et al.*, 2011).

Endogenous siRNAs were easily detectable in fungi, plants and *Cænorhabditis elegans*, but not in mammals or in *Drosophila* (yet the RNAi machinery is conserved too).

Natural functions for RNAi

In plants and insects: antiviral effect (see session of October 19). Apparent paradox in fungi: antiviral activity may have prompted the loss of RNAi in some species (Drinnenberg *et al.*, 2011).

Endogenous siRNAs were easily detectable in fungi, plants and *Cænorhabditis elegans*, but not in mammals or in *Drosophila* (yet the RNAi machinery is conserved too). 2008: Small RNA-Seq is sensitive enough to detect endogenous siRNAs in mammals and in *Drosophila*.

Natural functions for RNAi

In plants and insects: antiviral effect (see session of October 19). Apparent paradox in fungi: antiviral activity may have prompted the loss of RNAi in some species (Drinnenberg *et al.*, 2011).

Endogenous siRNAs were easily detectable in fungi, plants and *Cænorhabditis elegans*, but not in mammals or in *Drosophila* (yet the RNAi machinery is conserved too). 2008: Small RNA-Seq is sensitive enough to detect endogenous siRNAs in mammals and in *Drosophila*.

An apparent role in gene regulation and transposable element silencing (derepression in RNAi mutants); *Drosophila* mutants for the RNAi pathway are viable, but sick.

Natural functions for RNAi

In plants and insects: antiviral effect (see session of October 19). Apparent paradox in fungi: antiviral activity may have prompted the loss of RNAi in some species (Drinnenberg *et al.*, 2011).

Endogenous siRNAs were easily detectable in fungi, plants and *Cænorhabditis elegans*, but not in mammals or in *Drosophila* (yet the RNAi machinery is conserved too). 2008: Small RNA-Seq is sensitive enough to detect endogenous siRNAs in mammals and in *Drosophila*.

An apparent role in gene regulation and transposable element silencing (derepression in RNAi mutants); *Drosophila* mutants for the RNAi pathway are viable, but sick.

Nomenclature: "rasiRNAs" are not siRNAs (they are "piRNAs").

The discovery of microRNAs

1980's: Genetics: the *lin-4* gene is a repressor of *lin-14*.

The discovery of microRNAs

1980's: Genetics: the *lin-4* gene is a repressor of *lin-14*.

The *lin-4* genomic locus is not coding: its functional product is not a protein, but rather a 22 nt-long RNA (Lee *et al.*, 1993).

The discovery of microRNAs

1980's: Genetics: the *lin-4* gene is a repressor of *lin-14*.

The *lin-4* genomic locus is not coding: its functional product is not a protein, but rather a 22 nt-long RNA (Lee *et al.*, 1993).

It sequence is complementary to conserved repressive elements in the 3´ UTR of the lin-14 mRNA (Wightman *et al.*, 1993).

The discovery of microRNAs

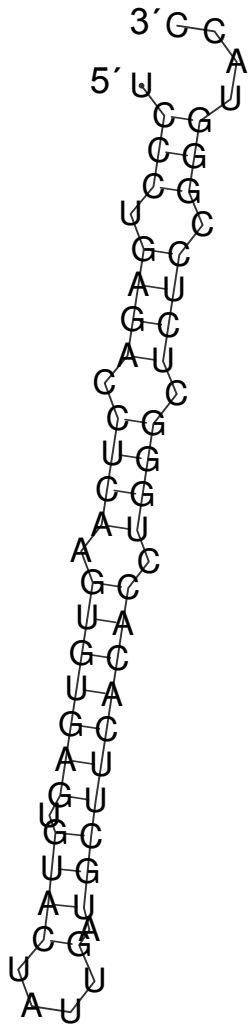
1980's: Genetics: the *lin-4* gene is a repressor of *lin-14*.

The *lin-4* genomic locus is not coding: its functional product is not a protein, but rather a 22 nt-long RNA (Lee *et al.*, 1993).

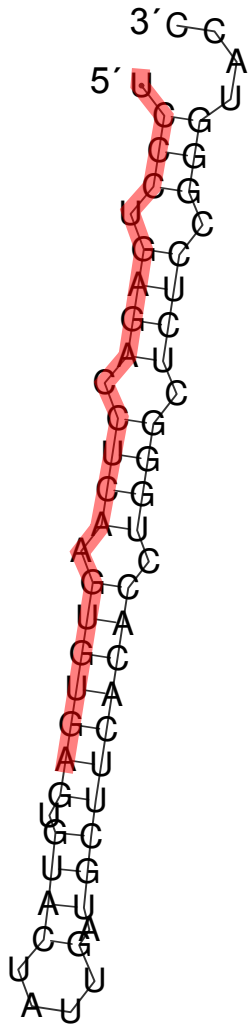
It sequence is complementary to conserved repressive elements in the 3´ UTR of the lin-14 mRNA (Wightman *et al.*, 1993).

A precursor for *lin-4*: stem-loop-folded RNA.

The discovery of microRNAs



The discovery of microRNAs



microRNAs everywhere...

2000: *let-7*: a novel small (21 nt) RNA involved in *C. elegans* development control.

microRNAs everywhere...

2000: *let-7*: a novel small (21 nt) RNA involved in *C. elegans* development control.
2000: *let-7* is conserved among triploblasts (Pasquinelli *et al.*, 2000).

microRNAs everywhere...

2000: *let-7*: a novel small (21 nt) RNA involved in *C. elegans* development control.
2000: *let-7* is conserved among triploblasts (Pasquinelli *et al.*, 2000).
T. Tuschl clones *Drosophila* embryo small RNAs while studying RNAi.

microRNAs everywhere...

2000: *let-7*: a novel small (21 nt) RNA involved in *C. elegans* development control.
2000: *let-7* is conserved among triploblasts (Pasquinelli *et al.*, 2000).
T. Tuschl clones *Drosophila* embryo small RNAs while studying RNAi.

Autumn 2001: identification of many similar small RNAs in nematode, *Drosophila* and cultured human cells: the "microRNA" family ("miRNA") (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001 and Lee and Ambros, 2001).

microRNAs everywhere...

2000: *let-7*: a novel small (21 nt) RNA involved in *C. elegans* development control.
2000: *let-7* is conserved among triploblasts (Pasquinelli *et al.*, 2000).
T. Tuschl clones *Drosophila* embryo small RNAs while studying RNAi.

Autumn 2001: identification of many similar small RNAs in nematode, *Drosophila* and cultured human cells: the "microRNA" family ("miRNA") (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001 and Lee and Ambros, 2001).
Common feature: genomic sequence is imperfectly self-complementary (stem-loop-folded RNA precursor ?).

microRNAs everywhere...

2000: *let-7*: a novel small (21 nt) RNA involved in *C. elegans* development control.
2000: *let-7* is conserved among triploblasts (Pasquinelli *et al.*, 2000).
T. Tuschl clones *Drosophila* embryo small RNAs while studying RNAi.

Autumn 2001: identification of many similar small RNAs in nematode, *Drosophila* and cultured human cells: the "microRNA" family ("miRNA") (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001 and Lee and Ambros, 2001).
Common feature: genomic sequence is imperfectly self-complementary (stem-loop-folded RNA precursor ?).

Since then: in metazoans, metaphytes, *Chlamydomonas*, viruses.

Mode of action

- Plants: most known targets exhibit near-perfect complementarity to the miRNA, they are cleaved between nucleotides paired to miRNA nt 10 and 11 (behave like siRNAs);

Mode of action

- Plants: most known targets exhibit near-perfect complementarity to the miRNA, they are cleaved between nucleotides paired to miRNA nt 10 and 11 (behave like siRNAs);
- Animals: a few rare examples (vertebrates: *HoxB8* and *Rtl1*; Yekta *et al.*, 2004 and Davis *et al.*, 2005); most known targets are very imperfectly complementary (no cleavage), they are repressed translationally.

Mode of action

- Plants: most known targets exhibit near-perfect complementarity to the miRNA, they are cleaved between nucleotides paired to miRNA nt 10 and 11 (behave like siRNAs);
- Animals: a few rare examples (vertebrates: *HoxB8* and *Rtl1*; Yekta *et al.*, 2004 and Davis *et al.*, 2005); most known targets are very imperfectly complementary (no cleavage), they are repressed translationally.

Location of imperfect binding sites: 3´ UTR, but $\exists$ exceptions !

microRNA maturation

Analogy with siRNAs

microRNA maturation

Analogy with siRNAs $\longrightarrow$ hypothesis: Dicer is responsible for microRNA maturation

microRNA maturation

Analogy with siRNAs $\longrightarrow$ hypothesis: Dicer is responsible for microRNA maturation $\longrightarrow$ indeed, it cleaves the last precursor ("pre-miRNA" stem-loop) into a miRNA/miRNA* duplex.

microRNA maturation

Analogy with siRNAs $\longrightarrow$ hypothesis: Dicer is responsible for microRNA maturation $\longrightarrow$ indeed, it cleaves the last precursor ("pre-miRNA" stem-loop) into a miRNA/miRNA* duplex.

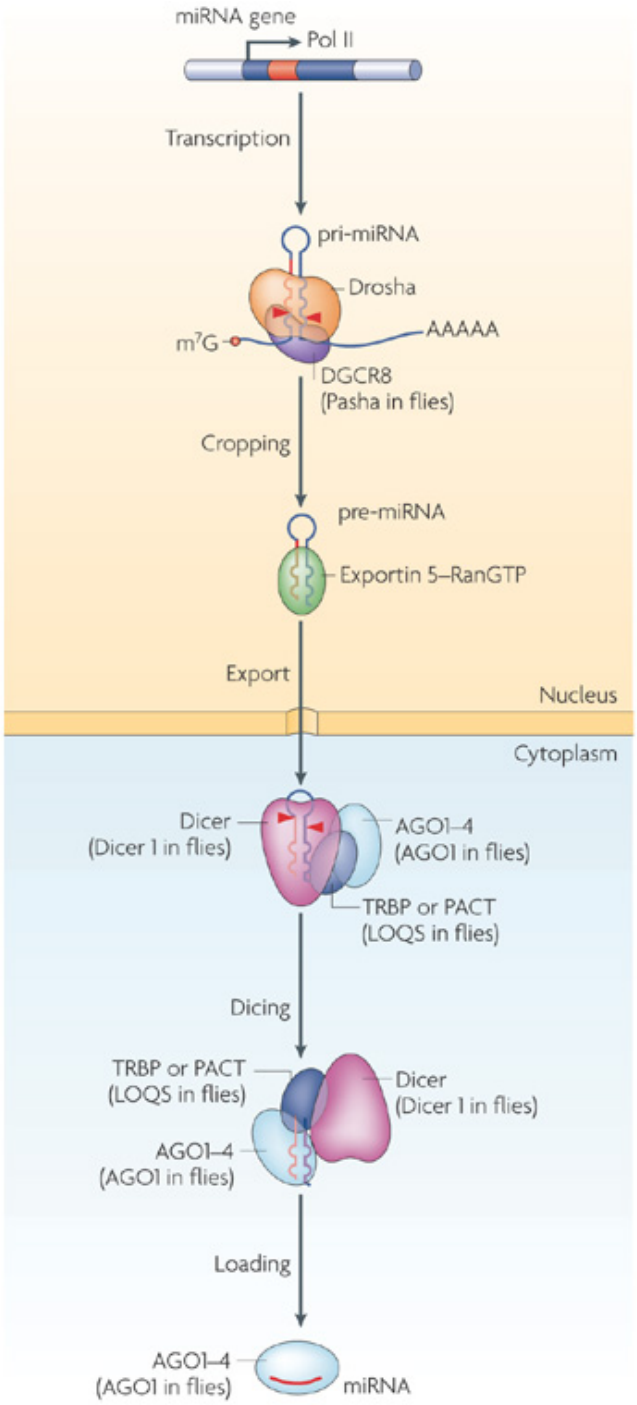
Previous step: by the Drosha (in animals) or Dicer (in plants) enzyme.

microRNA maturation

Analogy with siRNAs $\longrightarrow$ hypothesis: Dicer is responsible for microRNA maturation $\longrightarrow$ indeed, it cleaves the last precursor ("pre-miRNA" stem-loop) into a miRNA/miRNA* duplex.

Previous step: by the Drosha (in animals) or Dicer (in plants) enzyme.

There are a few exceptions.

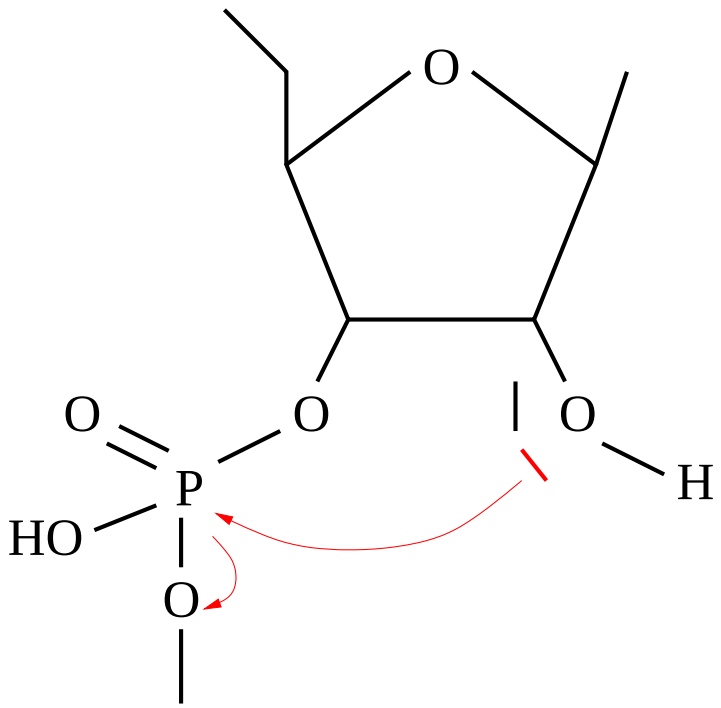


(taken from Kim *et al.*, 2009)

Why have miRNAs been ignored so long?

Why have miRNAs been ignored so long?

Too short: RNAs are very sensitive to degradation (nucleophilic attack by 2´ OH), and an RNA sample is always contaminated with degradation fragments (the bottom of the gel is traditionally ignored).



microRNA identification

microRNA identification

Large diversity (48,885 known miRNAs in 271 eukaryotic species and viruses: miRBase 22.1)

microRNA identification

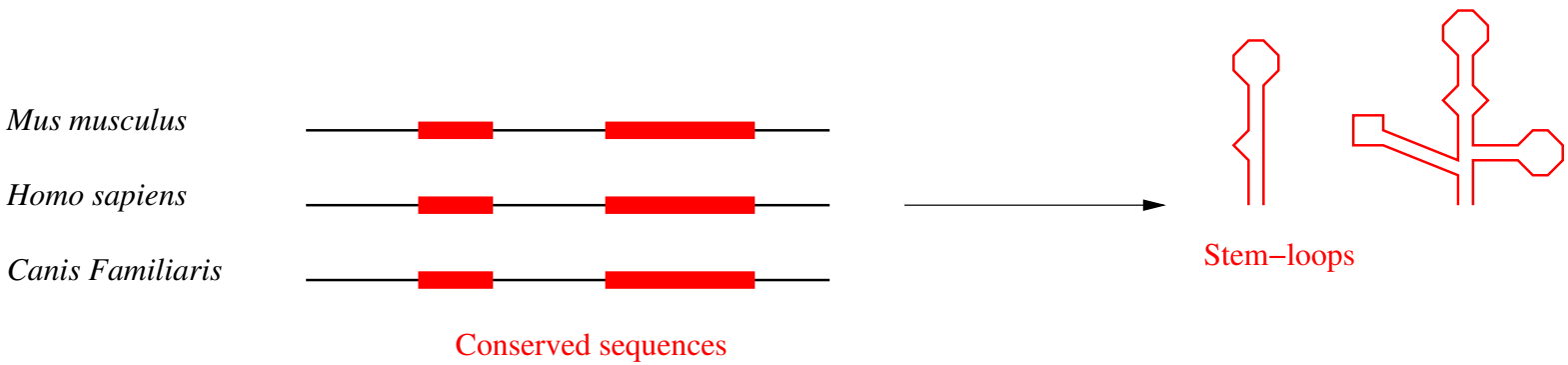
- Occasional discovery by genetics

microRNA identification

- Occasional discovery by genetics
- Experimental approach: cloning

microRNA identification

- Occasional discovery by genetics
- Experimental approach: cloning
- Computational approach: secondary structure prediction (not easy for plant miRNAs); selection (stem-loop stability, phylogenetic conservation).



microRNA identification

- Occasional discovery by genetics
- Experimental approach: cloning
- Computational approach: secondary structure prediction (not easy for plant miRNAs); selection (stem-loop stability, phylogenetic conservation).

Each approach is limited:

- Genetics: random, very rare.
- Cloning: abundance (weak expression, spatially or temporally restricted).
- Bio-informatics: choice of selection criteria is often arbitrary; does not identify miRNA extremities.

Differential transcriptomics or proteomics

Differential transcriptomics or proteomics

miRNA perturbation (*e.g.*, knockout) $\longrightarrow$ identify differentially expressed genes.

Differential transcriptomics or proteomics

miRNA perturbation (*e.g.*, knockout) $\longrightarrow$ identify differentially expressed genes.

Transcriptomics is currently more sensitive than proteomics; proteomics may capture more regulatory events.

Differential transcriptomics or proteomics

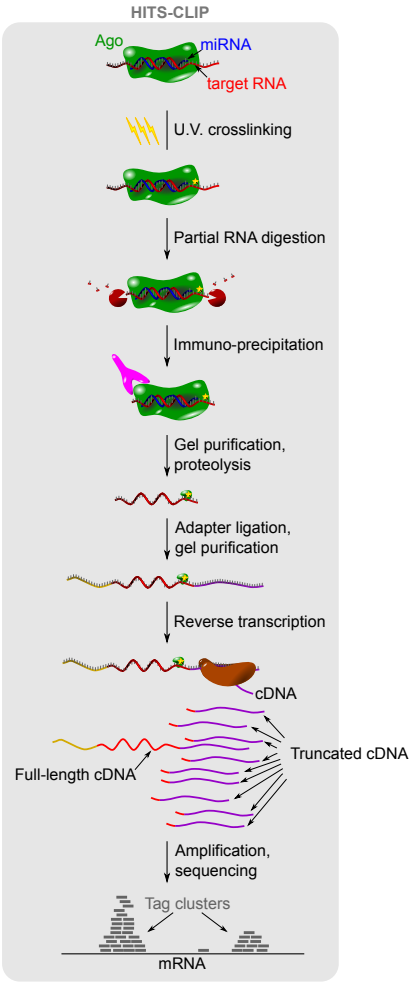
miRNA perturbation (*e.g.*, knockout) $\longrightarrow$ identify differentially expressed genes.

Transcriptomics is currently more sensitive than proteomics; proteomics may capture more regulatory events.

Identifies both direct and indirect targets.

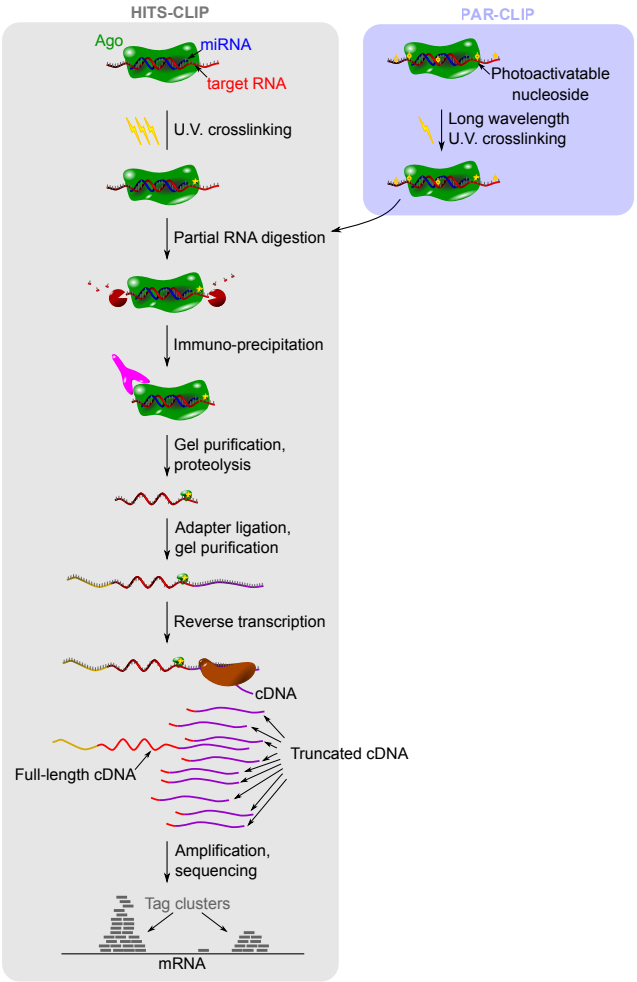
Identification of physical interactors

Identification of physical interactors



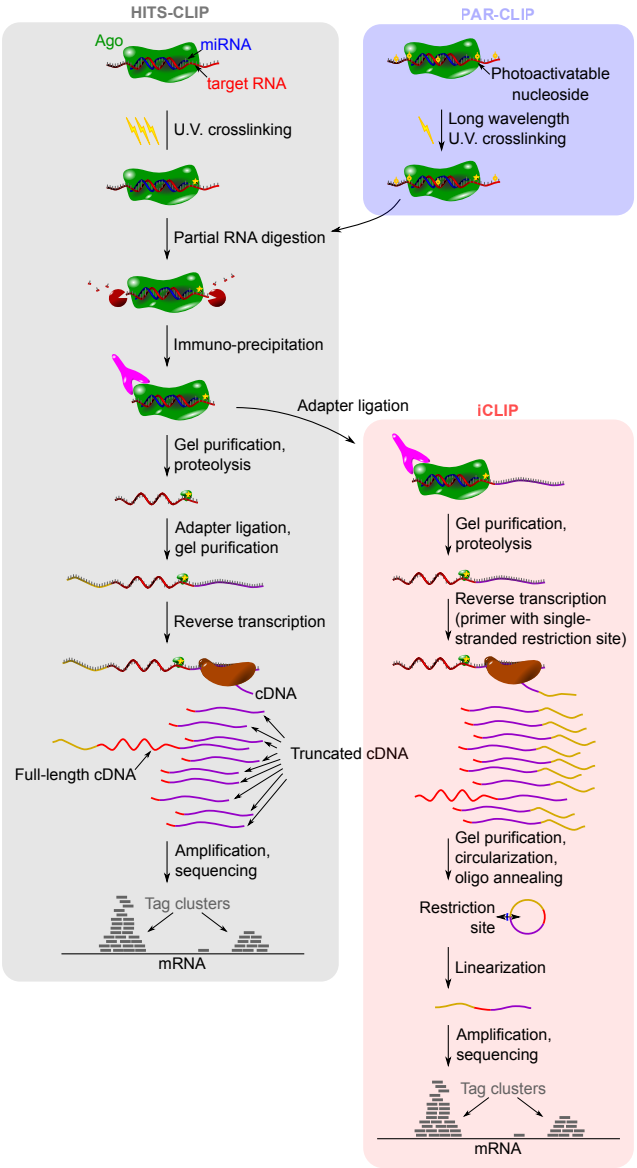
(taken from Mockly and Seitz, 2019)

Identification of physical interactors



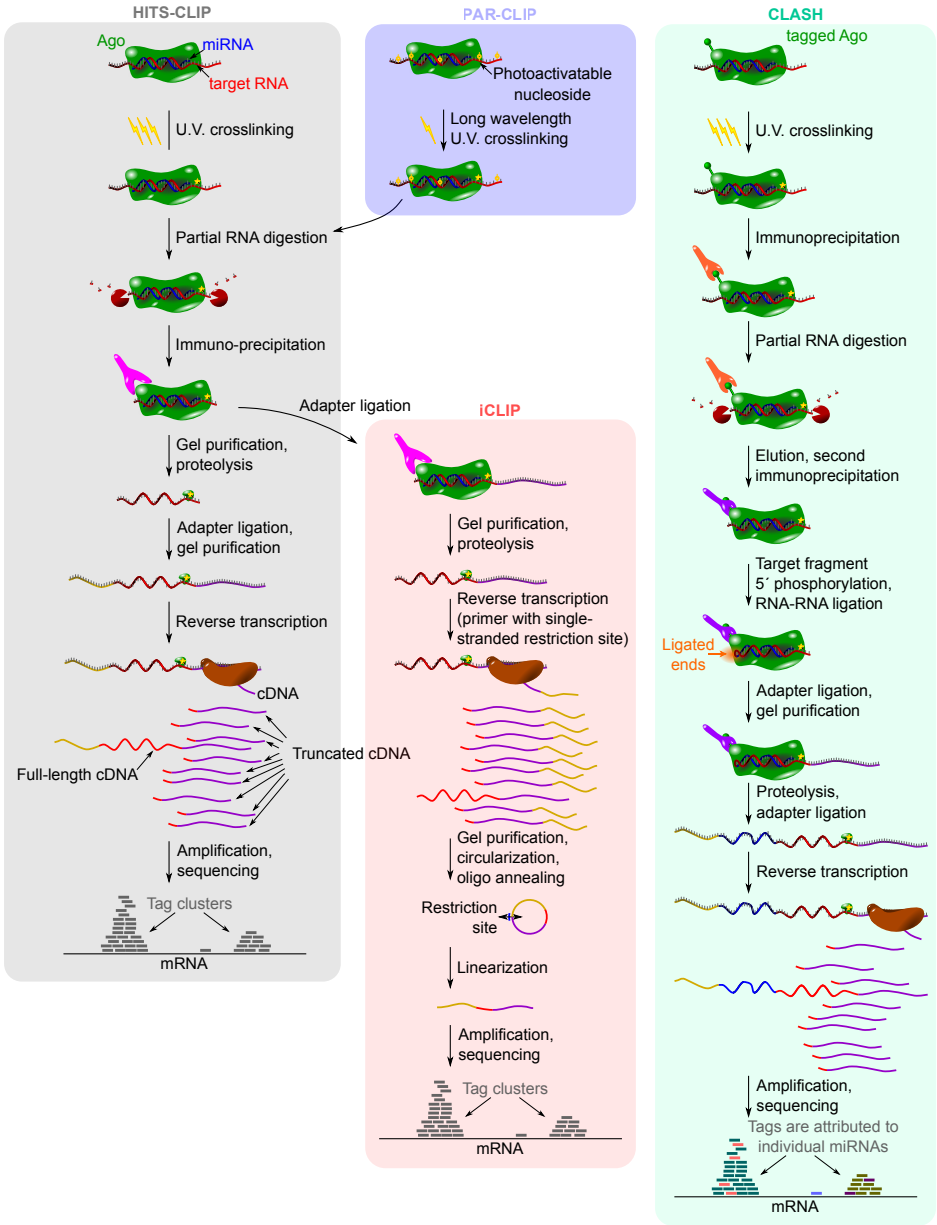
(taken from Mockly and Seitz, 2019)

Identification of physical interactors



(taken from Mockly and Seitz, 2019)

Identification of physical interactors



(taken from Mockly and Seitz, 2019)

Computational methods

Computational methods

Look for miRNA complementarity (usually: a seed match) in the transcriptome or 3´UTR-ome.

Computational methods

Look for miRNA complementarity (usually: a seed match) in the transcriptome or 3´UTR-ome.

($\longrightarrow$ would yield mostly false positives)

Computational methods

Look for miRNA complementarity (usually: a seed match) in the transcriptome or 3´UTR-ome.

($\longrightarrow$ would yield mostly false positives)

Further screening:

- look for accessible binding sites (Kertesz *et al.*, 2007)

Computational methods

Look for miRNA complementarity (usually: a seed match) in the transcriptome or 3´UTR-ome.

($\longrightarrow$ would yield mostly false positives)

Further screening:

- look for accessible binding sites (Kertesz *et al.*, 2007)
- look for favorable context (AU content, position in UTR, ...) (Agarwal *et al.*, 2015)

Computational methods

Look for miRNA complementarity (usually: a seed match) in the transcriptome or 3´UTR-ome.

($\longrightarrow$ would yield mostly false positives)

Further screening:

- look for accessible binding sites (Kertesz *et al.*, 2007)
- look for favorable context (AU content, position in UTR, ...) (Agarwal *et al.*, 2015)
- look for phylogenetically conserved sites (should capture known and unknown biochemical features) (Friedman and Burge, 2014)

Computational methods

Many programs available (PicTar, TargetScan, miRanda, rna22, Diana microT, ...).

Computational methods

Many programs available (PicTar, TargetScan, miRanda, rna22, Diana microT, ...).

Some are clearly better than others (rna22 $\approx$ random picking; Agarwal *et al.*, 2015).

Computational methods

Many programs available (PicTar, TargetScan, miRanda, rna22, Diana microT, ...).

Some are clearly better than others (rna22 $\approx$ random picking; Agarwal *et al.*, 2015).

Perhaps the best: TargetScan (and: frequently updated).

Computational methods

Many programs available (PicTar, TargetScan, miRanda, rna22, Diana microT, ...).

Some are clearly better than others (rna22 $\approx$ random picking; Agarwal *et al.*, 2015).

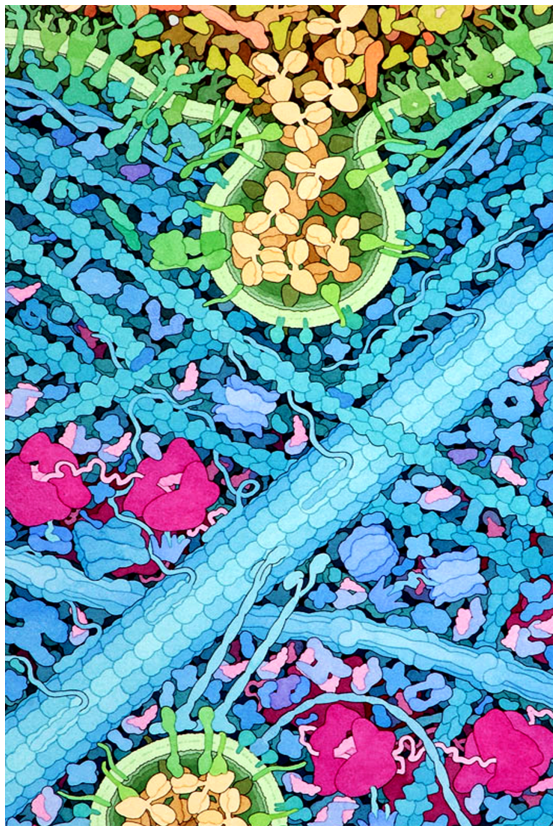
Perhaps the best: TargetScan (and: frequently updated). But: in their benchmark papers, the authors of TargetScan find that TargetScan is the best, the authors of PicTar find that PicTar is the best (Baek *et al.*, 2008; Selbach *et al.*, 2008) ...

A problem with high-throughput methods

A problem with high-throughput methods

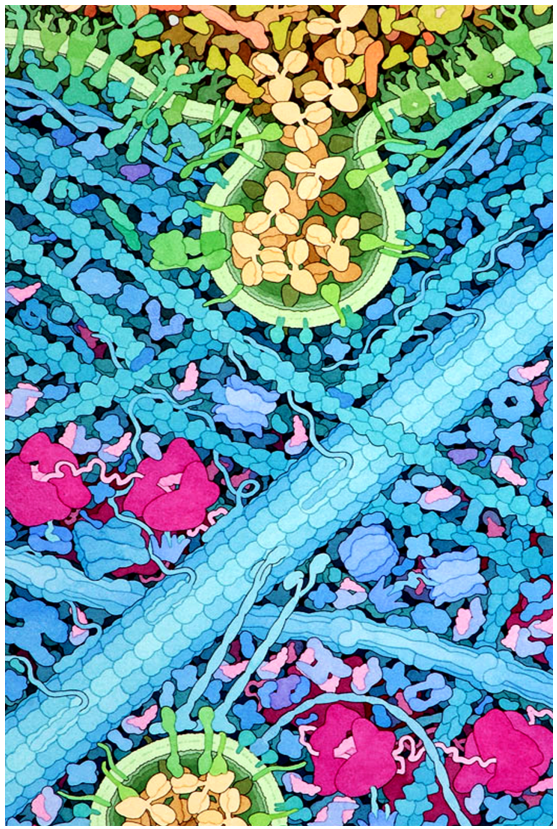
Showing that the miRNA binds (or: represses) the mRNA $\neq$ showing that this interaction plays a biological role

A problem with high-throughput methods



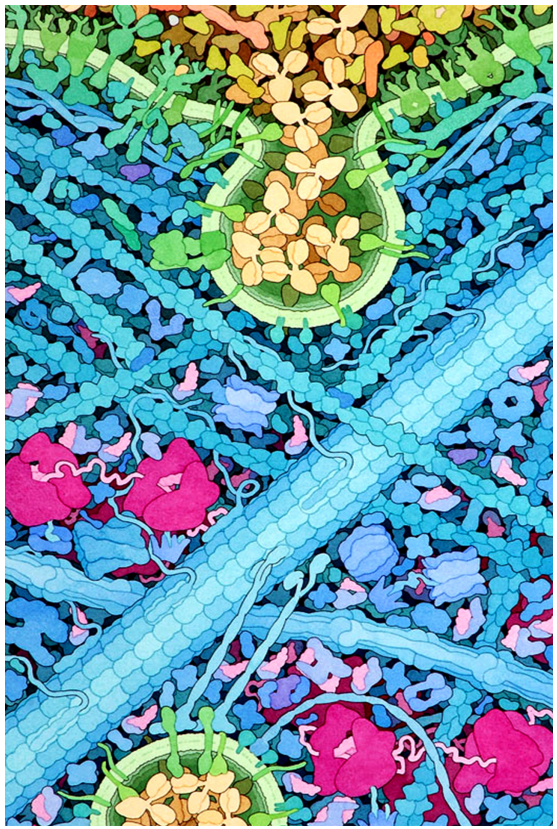
(image courtesy: Prof. D. Goodsell, Rutgers university)

A problem with high-throughput methods



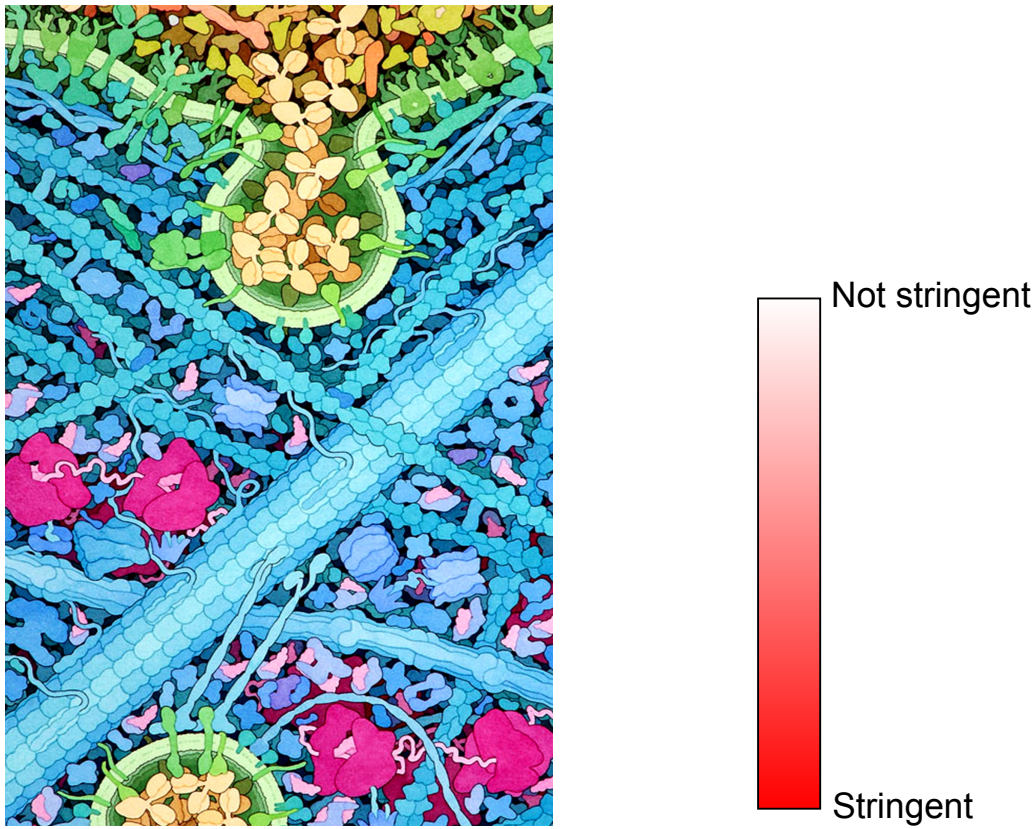
(image courtesy: Prof. D. Goodsell, Rutgers university)

A problem with high-throughput methods



(image courtesy: Prof. D. Goodsell, Rutgers university)
Physical interactions *do* occur in a cell.

A problem with high-throughput methods



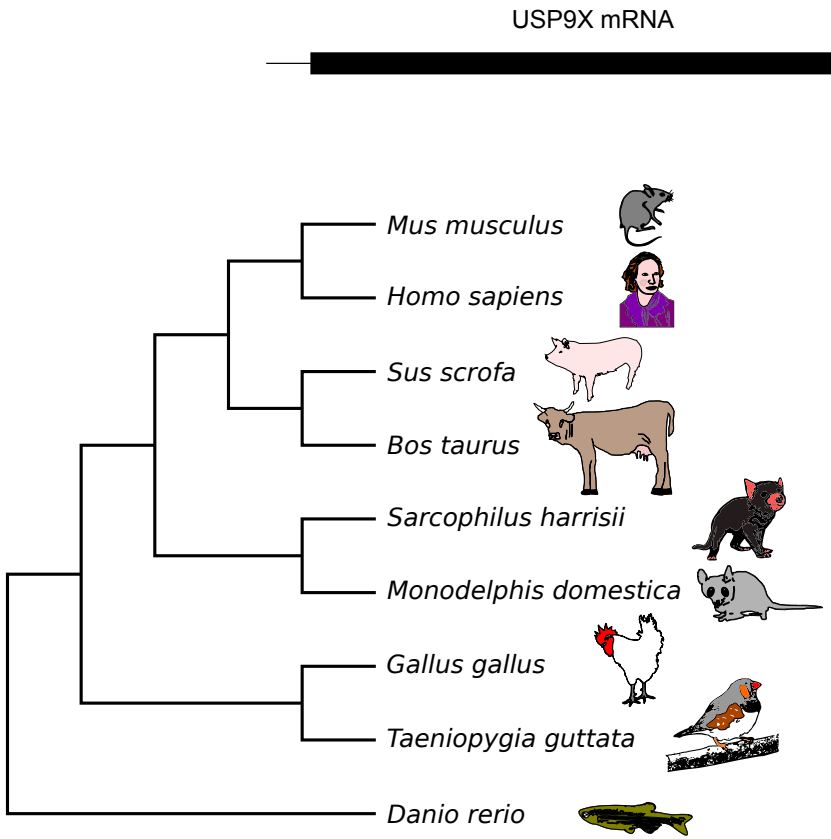
(image courtesy: Prof. D. Goodsell, Rutgers university)

Physical interactions *do* occur in a cell.

If it is conserved, it must be functional ($\longrightarrow$ prediction programs are useful even after a CLIP experiment).

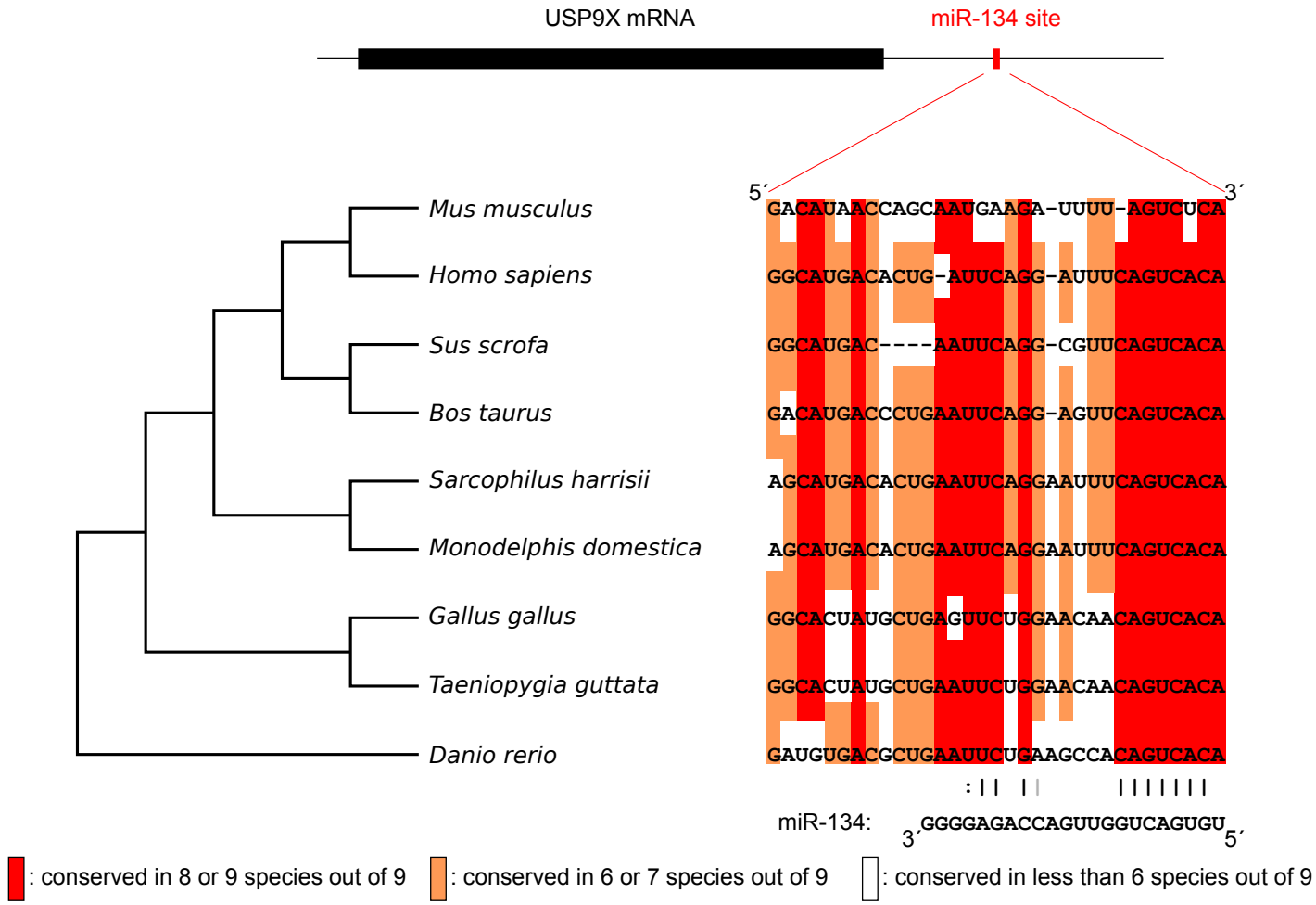
A problem with high-throughput methods

A problem with high-throughput methods



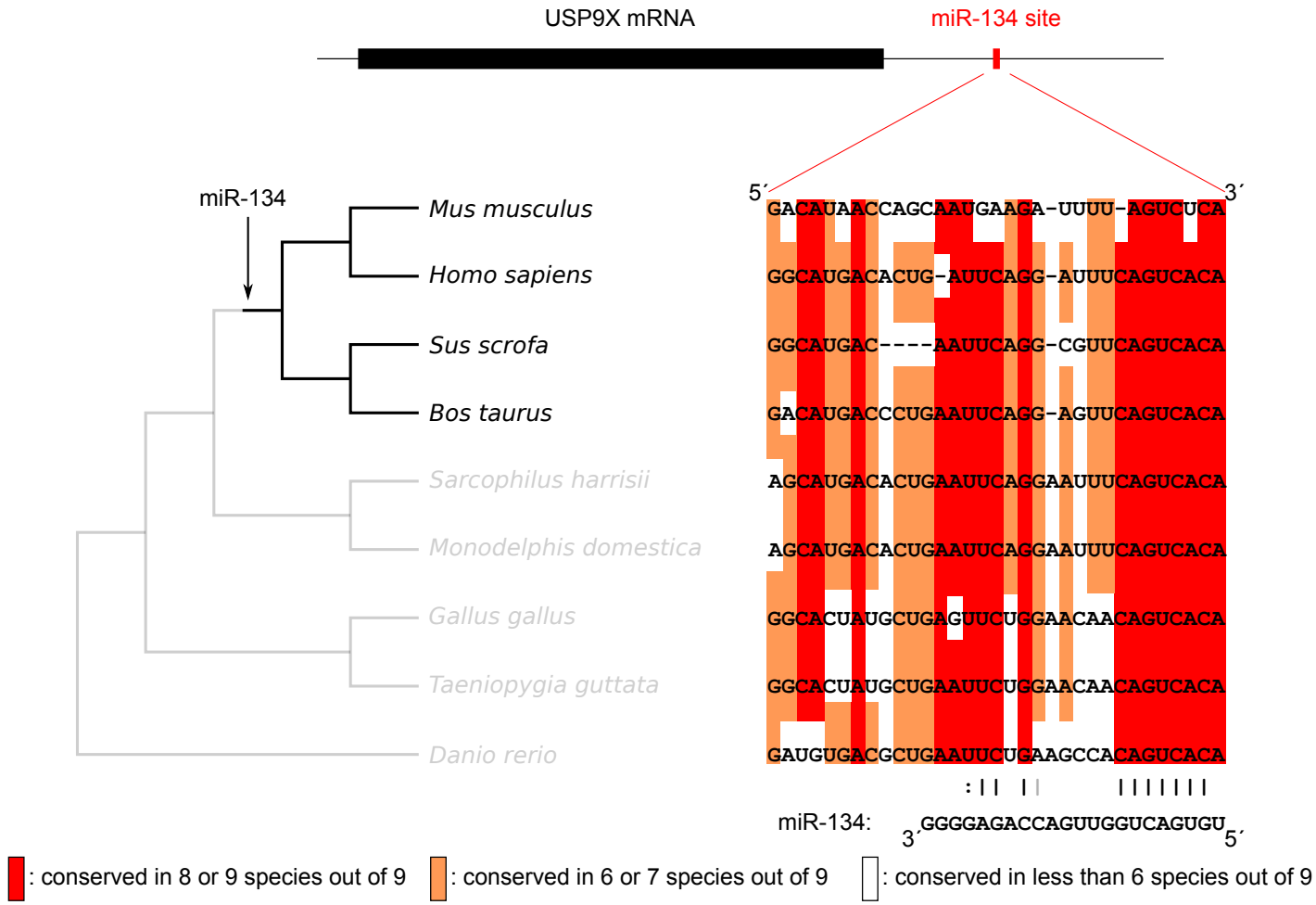
(taken from Pinzón *et al.*, 2017)

A problem with high-throughput methods



(taken from Pinzón *et al.*, 2017)

A problem with high-throughput methods



(taken from Pinzón *et al.*, 2017)

Conclusion

Conclusion

Computational predictions are now in good agreement with experimental identifications.

Conclusion

Computational predictions are now in good agreement with experimental identifications.

None of these methods can probe unambiguously for biological relevance.

Conclusion

Computational predictions are now in good agreement with experimental identifications.

None of these methods can probe unambiguously for biological relevance. Low-throughput genetics is required for a reliable assignment.

In vivo veritas: a large part of the miRNA literature certainly needs a cleanup (Seitz, 2024).