L4.4

RNA stability & decay

RNA stability

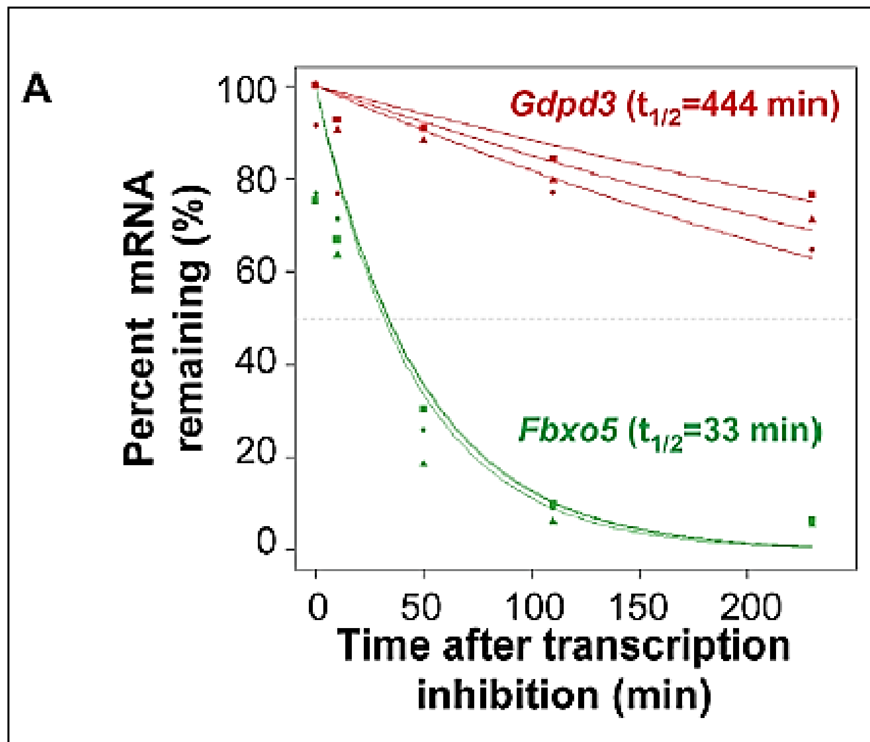
RNAs (both coding ad noncoding) display very **different half-lives**, in a quite wide range, which often are subjected to regulation

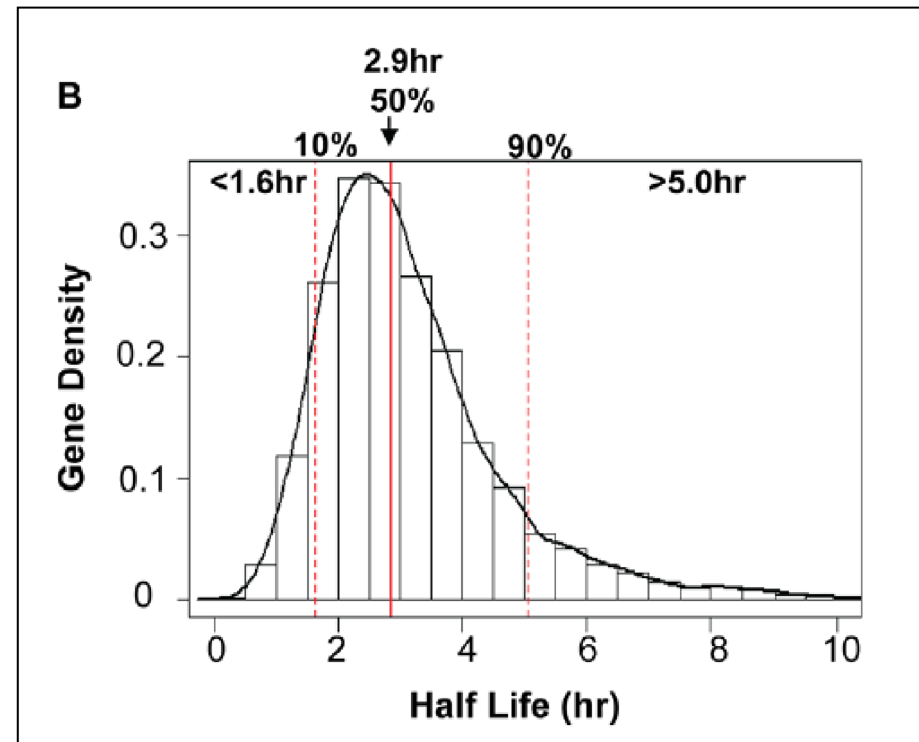
Intrinsic factors: *cis* sequences, most often within the 3’ UTR

Regulatory factors: RBP or Protein-small RNAs that can induce either stabilization or destabilization.

Lee et al., PLOS One, 5:11201 (2010)

Mouse myoblasts in culture treated with actinomycin D
Samples collected at 0, 10, 50, 110, 230 min
Total RNA → labeled → hybridized to Affymetrix Mouse Gene 1.0 arrays.



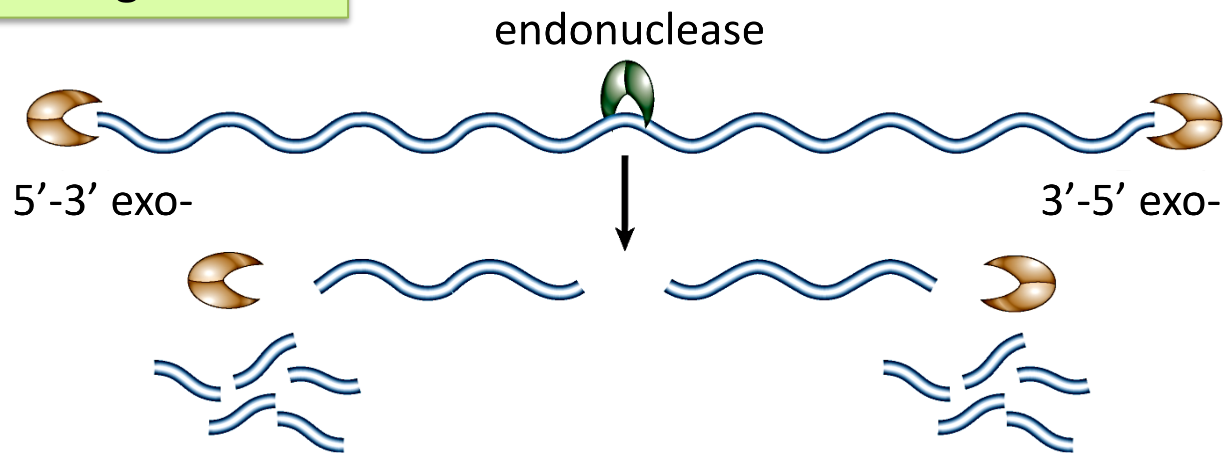


Analysis of mRNA decay rate in C2C12 cells.

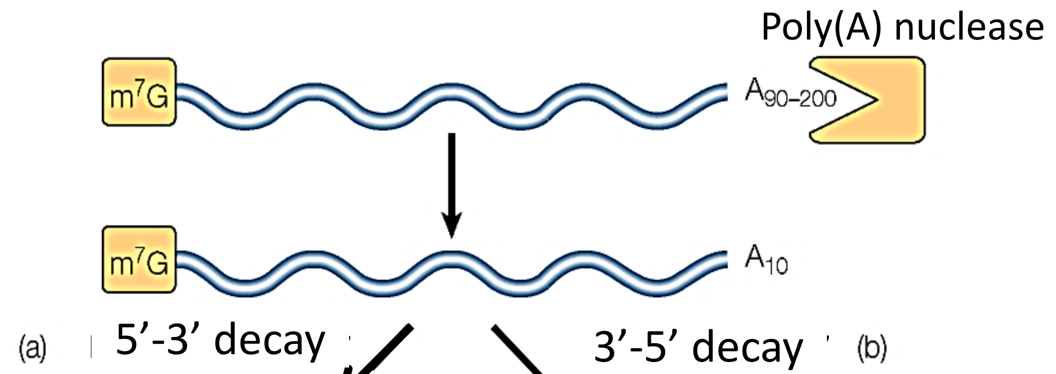
(A) Examples of mRNA decay curves at extremes. were derived by the nonlinear least squares method for a long and a short half life mRNA.

(B) Distribution of mRNA half and 90th-percentile values .

Pathways to RNA degradation



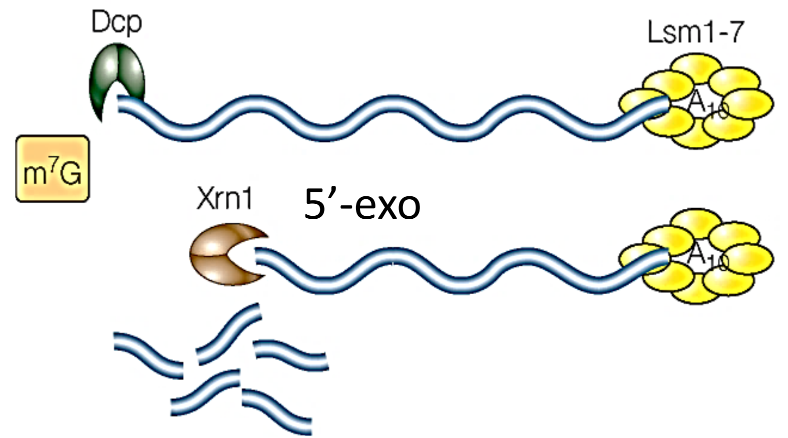
default



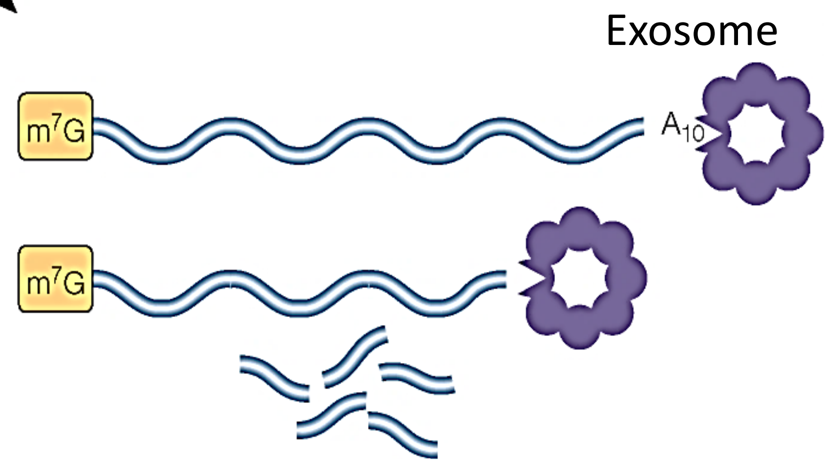
(a) 5′-3′ decay

3′-5′ decay (b)

decapping



Exosome



RNAs intrinsic *cis* regulatory sequences: Destabilizing (DE) or Stabilizing Elements (SE)

Examples of *trans*-acting regulators

RNA-binding protein	Function	RNA-binding domain	Mode of action	Modifications	Other functions
AUF1 (hnRNP D) and its four splice isoforms (p37, p40, p42, p45)	Usually destabilizing	RRM	Recruit the exosome; remodel mRNA to allow other proteins to bind	Phosphorylation allows isomerization by PIN1 leading to dissociation from RNA; interacts with 14-3-3 proteins	DNA binding
CUG-BP	Destabilizing	RRM	Recruits PARN; modulates ARE function	Phosphorylated by myotonic dystrophy protein kinase	Splicing; translation
ELAV proteins, for example, HuR and HuD	Stabilizing	RRM	Compete with destabilizing proteins for ARE-binding; might relocalize mRNAs away from decay machinery	CARM1-mediated methylation reduces stabilizing function	Translation; RNA localization
KSRP	Destabilizing	KH domain	Recruits decay enzymes: PARN and the exosome	Phosphorylation by p38-MAPK pathway leads to reduced RNA-binding affinity	Splicing
RHAU	Destabilizing	RNA helicase	Recruits decay enzymes: PARN and the exosome	Not known	Not known
TIA-1, TIAR	Translational silencing	RRM	Induce aggregation into stress granules	Phosphorylated by FAST	Alternative splicing
Tristetraprolin (TTP, TIS11, ZFP36), BRF1 (TIS11B, ZFP36L1), BRF2 (TIS11D, ZFP36L2)	Destabilizing	CCCH-type zinc finger	Recruit decay enzymes: CCR4, DCP1, PM-Scl75, RRP4	Phosphorylation by p38-MAPK pathway leads to association with 14-3-3 proteins	Transcription

Garneau *et al. Nature Reviews Molecular Cell Biology* **8**, 113–126 (2007) | doi:10.1038/nrm2104

nature REVIEWS MOLECULAR CELL BIOLOGY

RNA interference

From the regulatory point of view, the most interesting class is given by **micro-RNA and other small RNA-guided AGO proteins**

The pathways in which AGO proteins are involved are collectively called ***RNA interference***

RNA interference, small-interfering RNA, micro-RNA

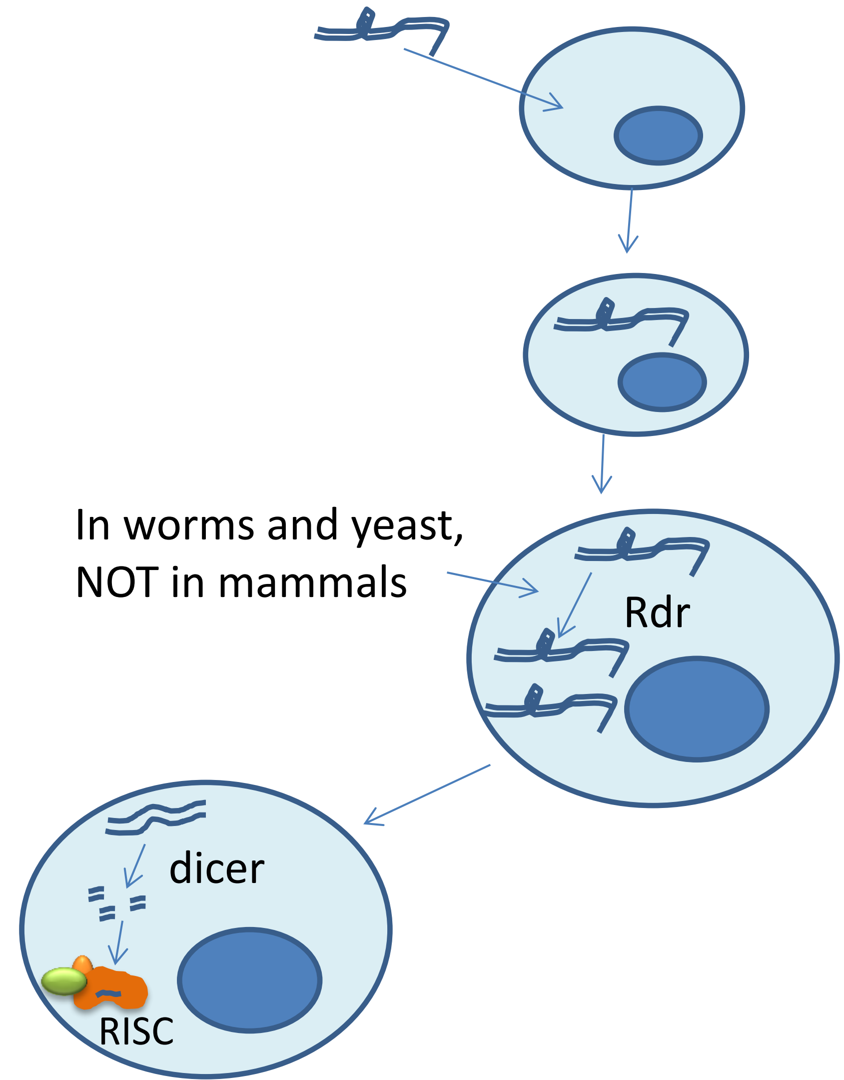
Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*

Andrew Fire and Craig Mello had the Nobel prize in 2006.

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

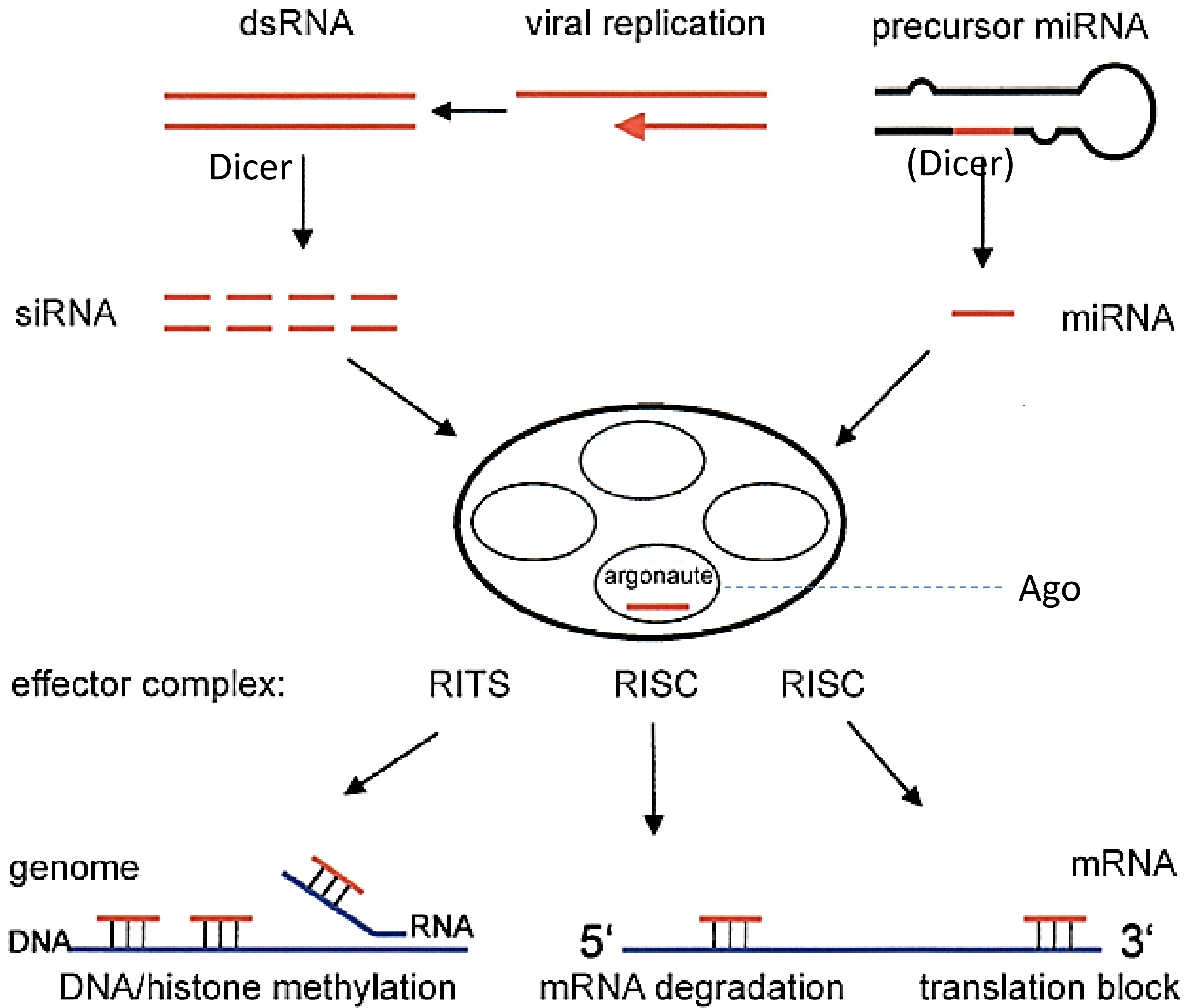
* *Carnegie Institution of Washington, Department of Embryology, 115 West University Parkway, Baltimore, Maryland 21210, USA*
† *Biology Graduate Program, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA*
‡ *Program in Molecular Medicine, Department of Cell Biology, University of Massachusetts Cancer Center, Two Biotech Suite 213, 373 Plantation Street, Worcester, Massachusetts 01605, USA*

Experimental introduction of RNA into cells can be used in certain biological systems to interfere with the function of an endogenous gene[1,2]. Such effects have been proposed to result from a simple antisense mechanism that depends on hybridization between the injected RNA and endogenous messenger RNA transcripts. RNA interference has been used in the nematode *Caenorhabditis elegans* to manipulate gene expression[3,4]. Here we investigate the requirements for structure and delivery of the interfering RNA. To our surprise, we found that double-stranded RNA was substantially more effective at producing interference than was either strand individually. After injection into adult animals, purified single strands had at most a modest effect, whereas double-stranded mixtures caused potent and specific interference. The effects of this interference were evident in both the injected animals and their progeny. Only a few molecules of injected double-stranded RNA were required per affected cell, arguing against stochiometric interference with endogenous mRNA and suggesting that there could be a catalytic or amplification component in the interference process.

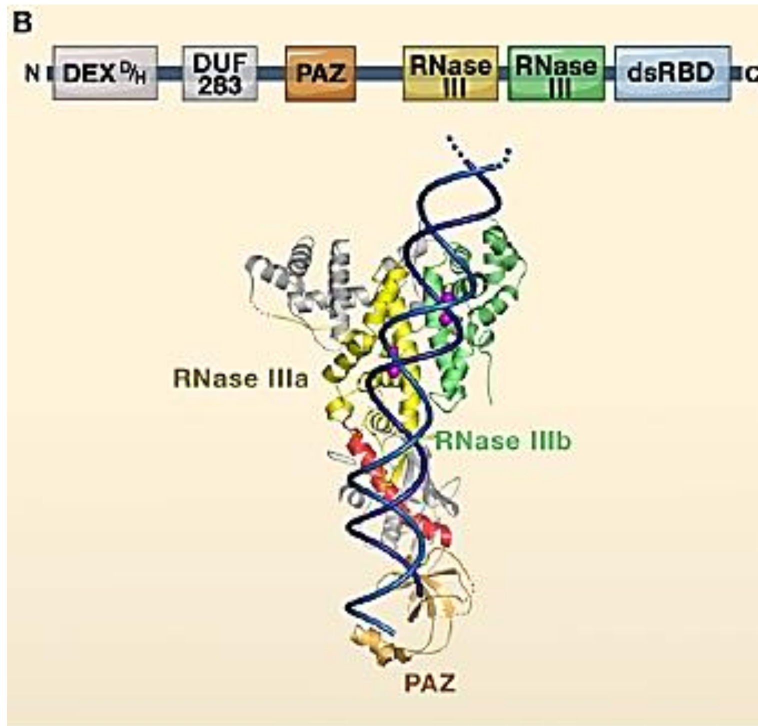


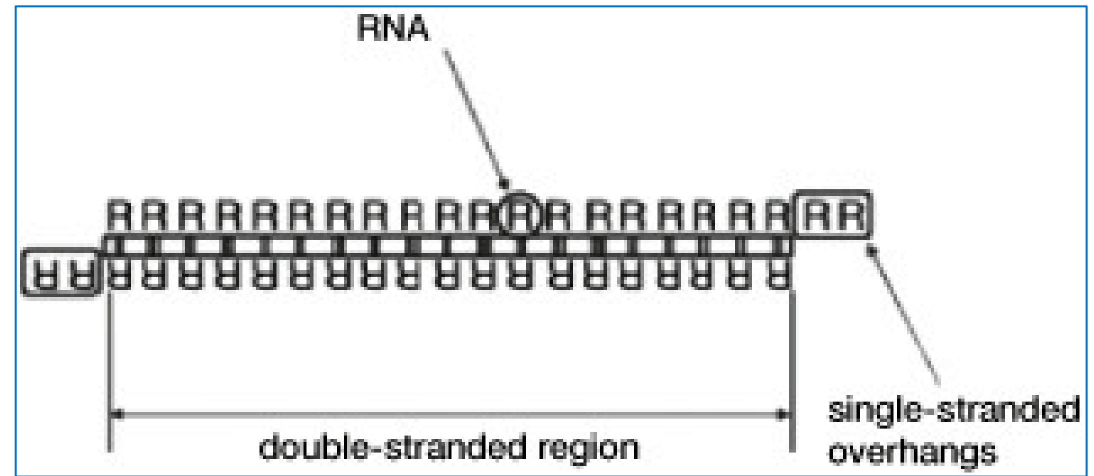
Nature © Macmillan Publishers Ltd 1998

RNA interference



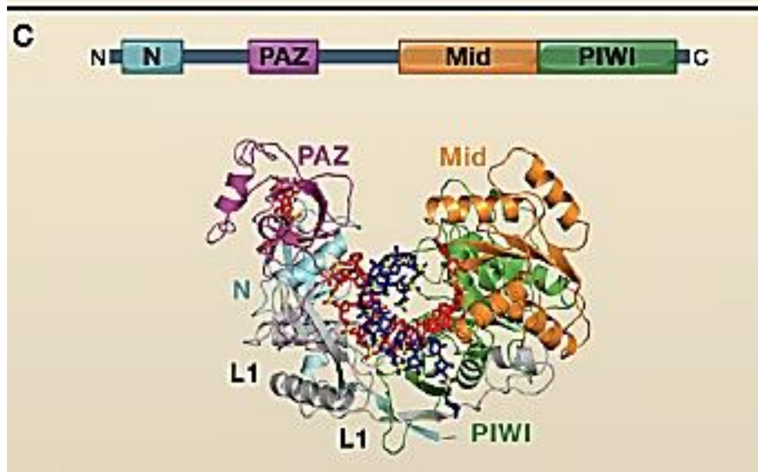
Carthew, 2009





DICER is a RNase III enzyme.
The PAZ domain binds the RNA duplex end, then the two catalytic subunits cleave leaving 2nt 3'-overhangs (5' monophosphate).

Mammals have a single DICER



Argonaute superfamily: 3 clades.

- ✓ Piwi clade: bind piRNAs
- ✓ Ago clade: bind siRNA and miRNA
- ✓ 3rd: only in Nematodes (to date)

8 Argonaute proteins (Ago) in H. Sapiens
Ago2 has endonuclease activity.

Different short noncoding RNA classes emerged from short-RNA-seq analyses

Table 1

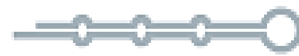
Classes of small non-coding RNAs

ncRNA	Abbreviation	Model organism(s)
microRNA	miRNA	Most eukaryotic organisms except fungi; viruses
miRNA-offset RNA	moRNA	*Ciona intestinalis*, mammals
Short hairpin-derived miRNA	shRNA-derived miRNA	*Mus musculus*
miRNA-like small RNA	milRNA	*Neurospora crassa*
mirtron	–	*D. melanogaster*, *C. elegans*, mammals
Piwi-interacting RNA	piRNA	*D. melanogaster*, mammals, *Danio rerio*, *Xenopus laevis*, *Bombyx mori*
21U-RNA	–	*C. elegans*
Endogenous small interfering RNA	endo-siRNA	*D. melanogaster*, *C. elegans*, mammals
tRNA-derived RNA fragment, tRNA-derived small RNA	tRF, tsRNA	Mammals
Promoter-associated small RNA, termini-associated small RNA	PASR, TASR	Mammals
Transcription start site-associated RNA, transcription initiation RNA	TSSa-RNA, tiRNA	Mammals, *Gallus gallus*, *D. melanogaster*
Splice-site RNA	spliRNA	Metazoans
snoRNA-derived RNA	sdRNA	Metazoans, *Arabidopsis thaliana*, *Schizosaccharomyces pombe*
QDE-2-interacting small RNA	qiRNA	*N. crassa*
Small vault RNA	svRNA	*Homo sapiens*

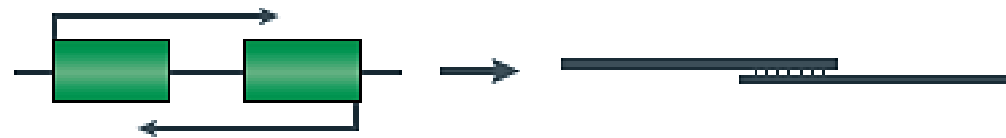
From Aalto & Pasquinelli, 2012

Endo-siRNA

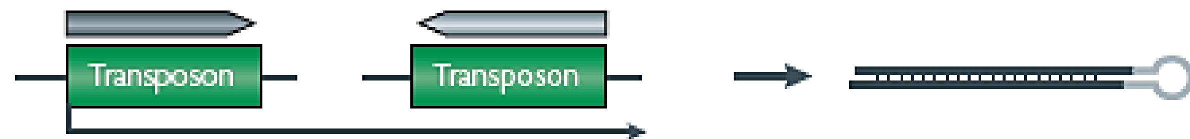
Structured loci



Convergent transcription

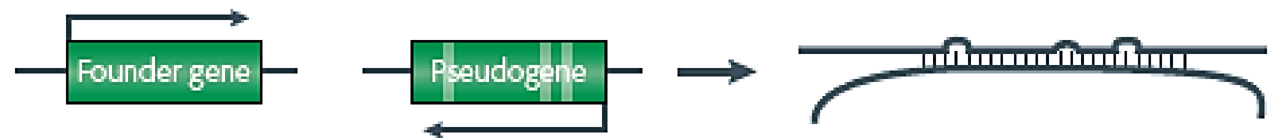


Read-through transcription of transposons in inverted orientation



Bidirectional transcription

Trans-interaction



Duplicated and inverted pseudogene copies

Figure 3 | **Genomic sources of dsRNA triggers for endogenous small interfering RNAs (endo-siRNAs) in flies and mammals.**

piRNA are small RNAs associated to the Piwi-subfamily of Argonaute proteins.

They have other specific features:

- Lack of dsRNA precursors
- Independence of Dicer

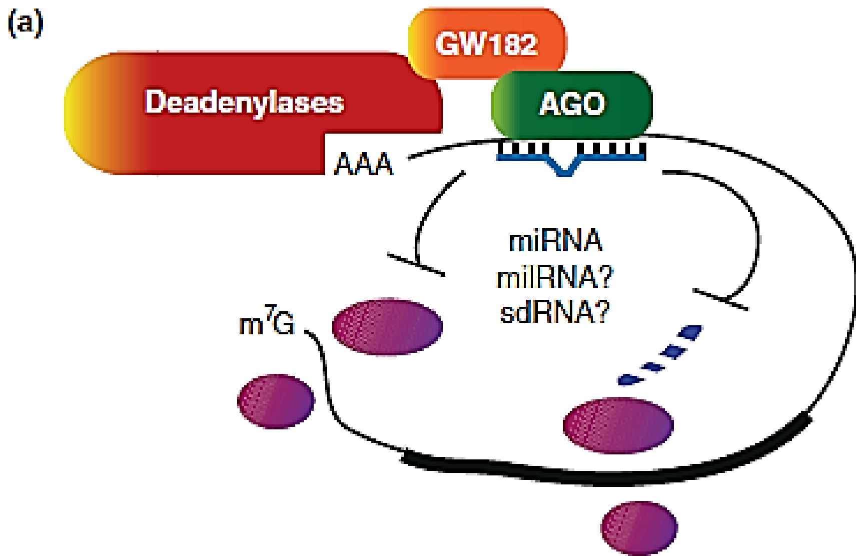
Primary piRNA produced from piRNA clusters.
Processed into **24-30 nt** piRNA that associate with Argonaute Piwi clade proteins.
piRNA pairs with transposon sense transcript and cleave them endonucleolitically.

piRNAs were first proposed to ensure germline stability by repressing transposons

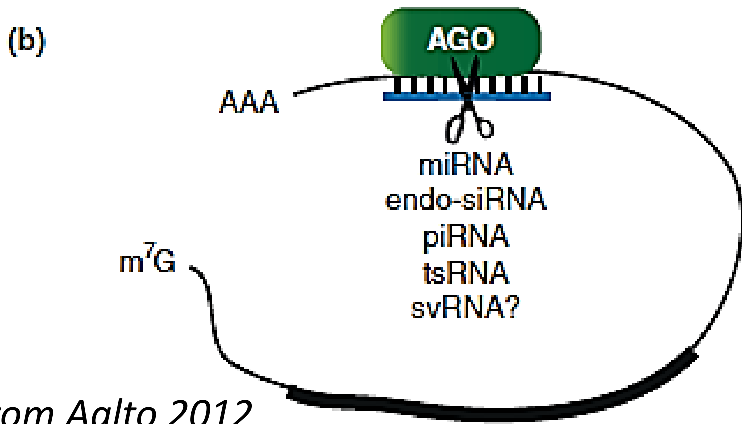
Mammalian piRNAs:
pre-pachytene and pachytene piRNAs
(according to the stage of meiosis in developing spermatocytes).

How do small interfering RNA exert their silencing effects?

Different mechanisms



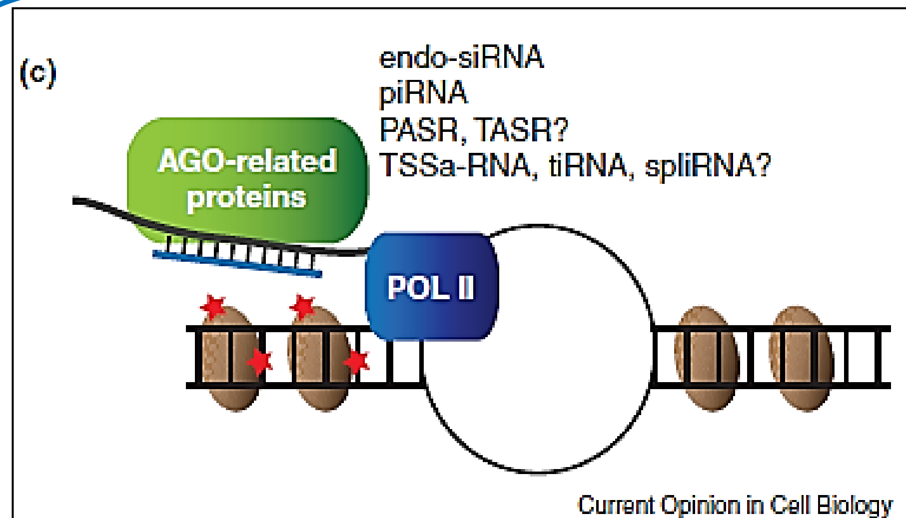
miRNAs partially pair to mRNA targets leading to mRNA degradation by GW182/deadenylases bound to Argonaute (AGO);
or miRISC inhibits translation initiation;
or stalls translation elongation;
or stimulates proteolysis of nascent peptides encoded by the target mRNA.



Perfect pairing of many types of sRNAs results in cleavage of the target mRNA by **Ago2**.

From Aalto 2012

Endo-siRNAs, piRNAs and possibly many of the sRNAs derived from protein-coding regions of the genome direct DNA and histone modifications (red stars) that regulate transcriptional activity



miRNA

Micro RNA are a family of small RNA that are transcribed from several locations in genomes.

They have a typical structure, making a stem-loop structure with some mismatches in the stem

MiRBase: http://www.mirbase.org/cgi-bin/browse.pl?org=hsa

Textbook

Cell

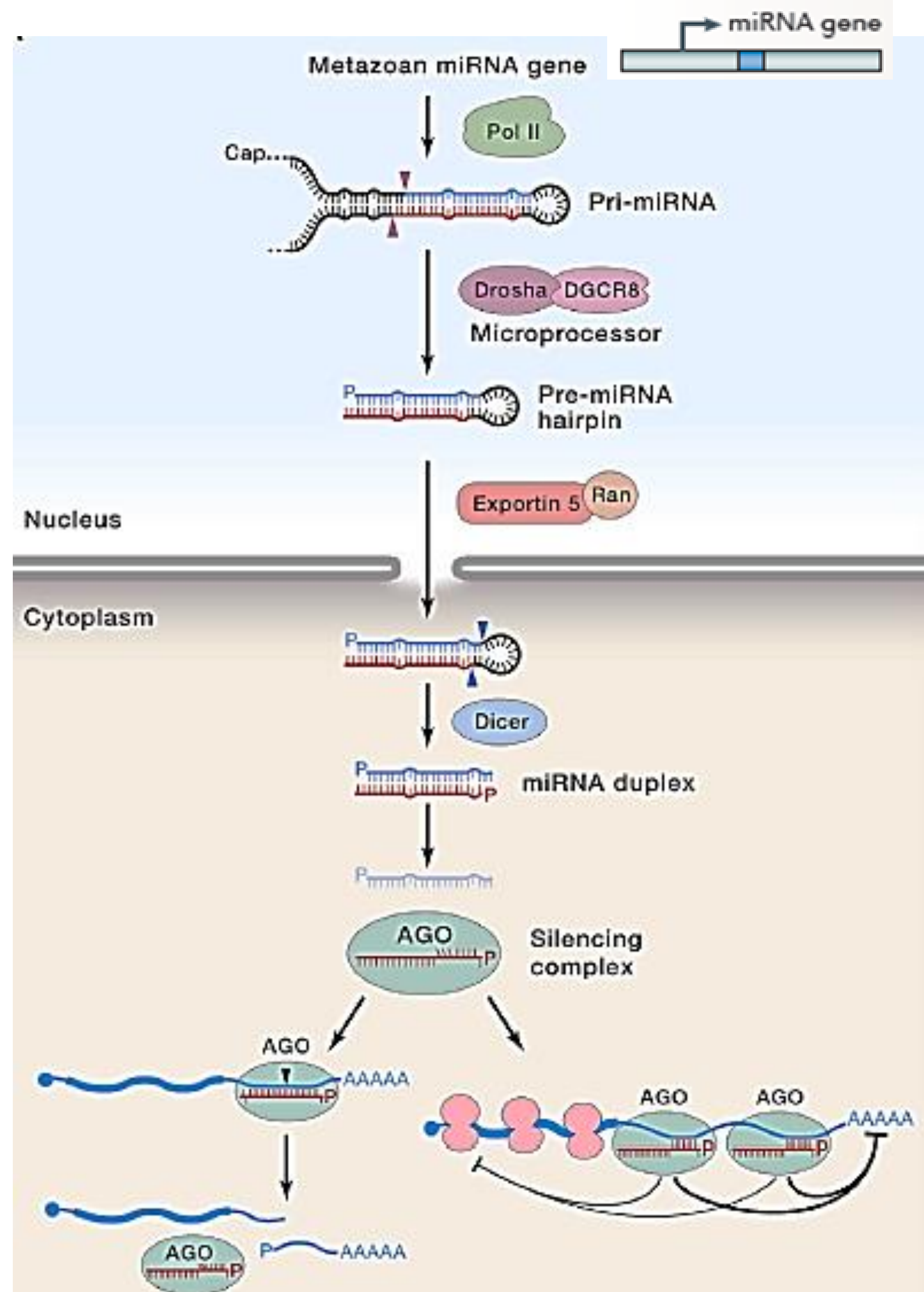
Leading Edge
Review

Metazoan MicroRNAs

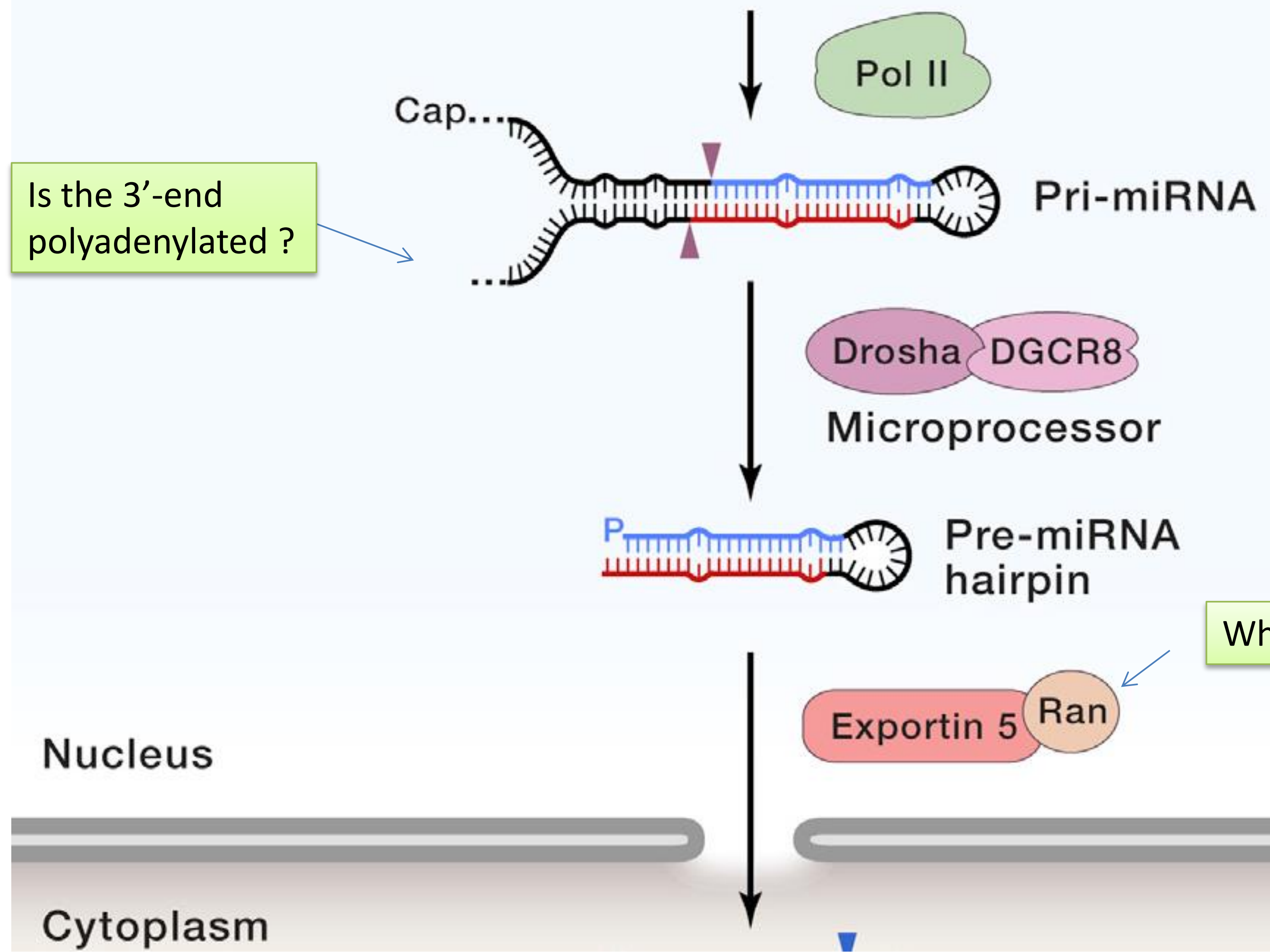
David P. Bartel[1,2,*]
[1]Howard Hughes Medical Institute and Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: dbartel@wi.mit.edu
https://doi.org/10.1016/j.cell.2018.03.006

MicroRNAs (miRNAs) are ~22 nt RNAs that direct posttranscriptional repression of mRNA targets in diverse eukaryotic lineages. In humans and other mammals, these small RNAs help sculpt the expression of most mRNAs. This article reviews advances in our understanding of the defining features of metazoan miRNAs and their biogenesis, genomics, and evolution. It then reviews how metazoan miRNAs are regulated, how they recognize and cause repression of their targets, and the biological functions of this repression, with a compilation of knockout phenotypes that shows that important biological functions have been identified for most of the broadly conserved miRNAs of mammals.

 © 2018 Elsevier Inc.



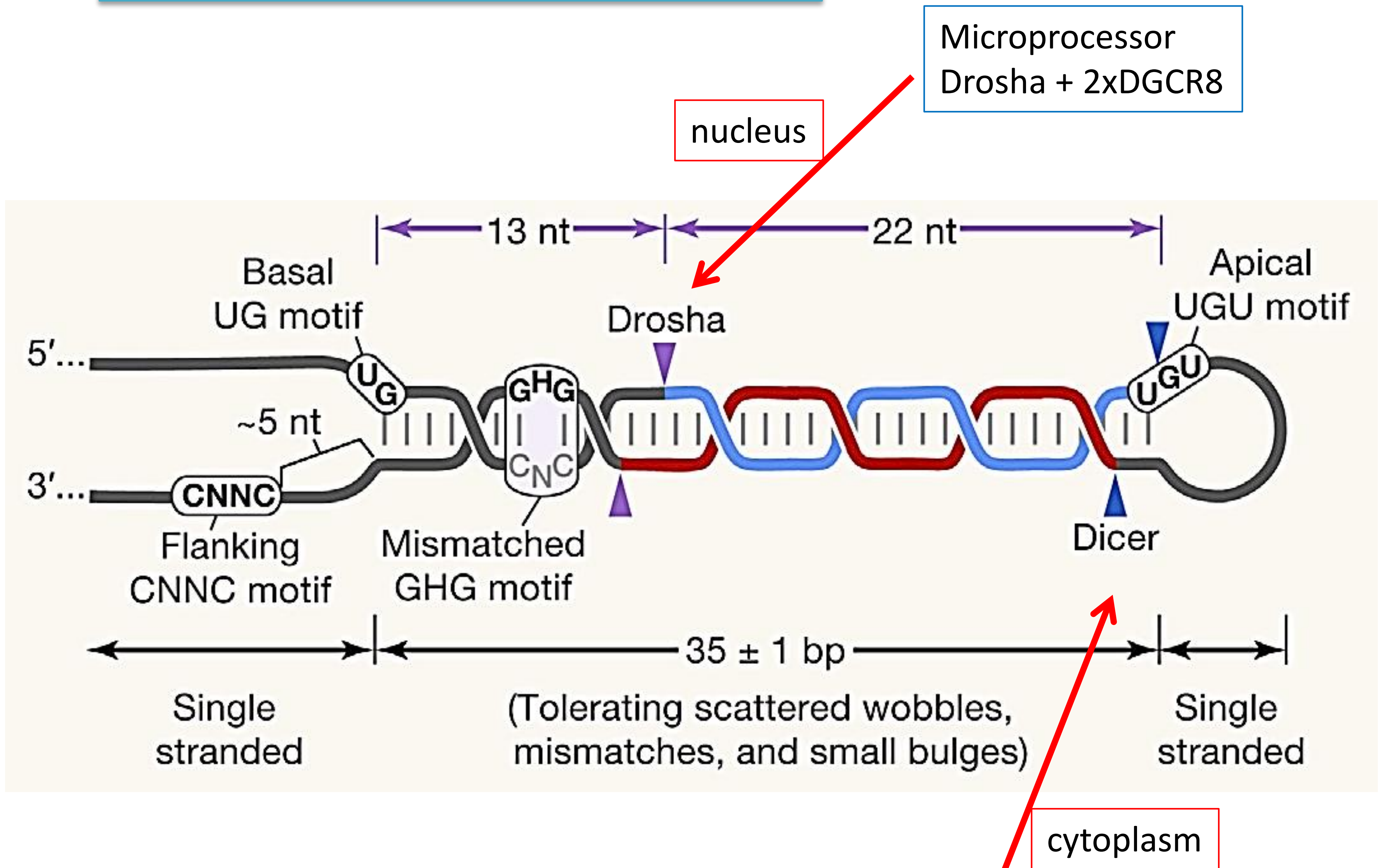
miRNA gene
Metazoan miRNA gene
Pol II
Cap...
Pri-miRNA
Drosha
DGCR8
Microprocessor
P
Pre-miRNA
hairpin
Exportin 5
Ran
Nucleus
Cytoplasm
P
Dicer
P
P
miRNA duplex
P
AGO
P
Silencing
complex
AGO
AAAAA
P
AGO
AGO
AAAAA
P
P
AGO
P
P
AAAAA



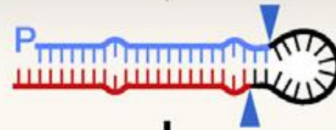
Is the 3’-end polyadenylated ?

What is Ran ?

Features defining a canonical pre-miRNA

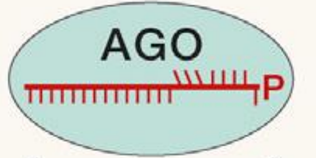


Cytoplasm



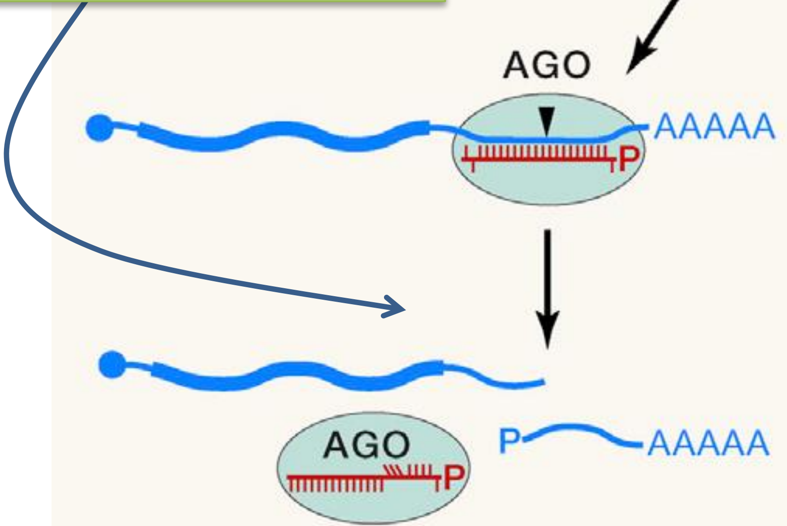
Dicer

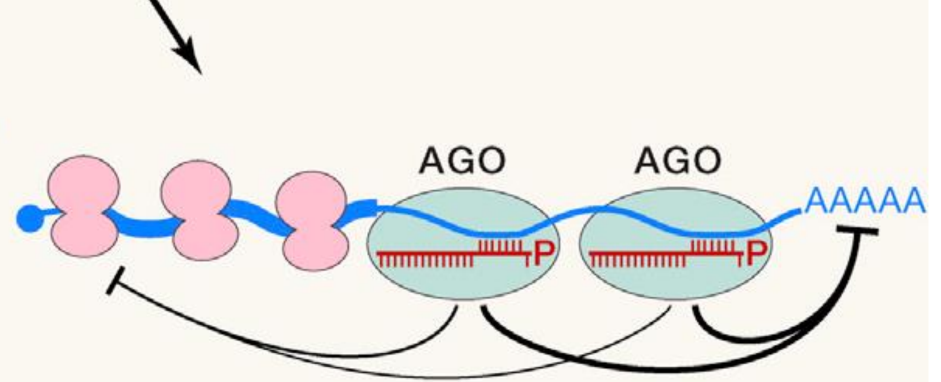
miRNA duplex

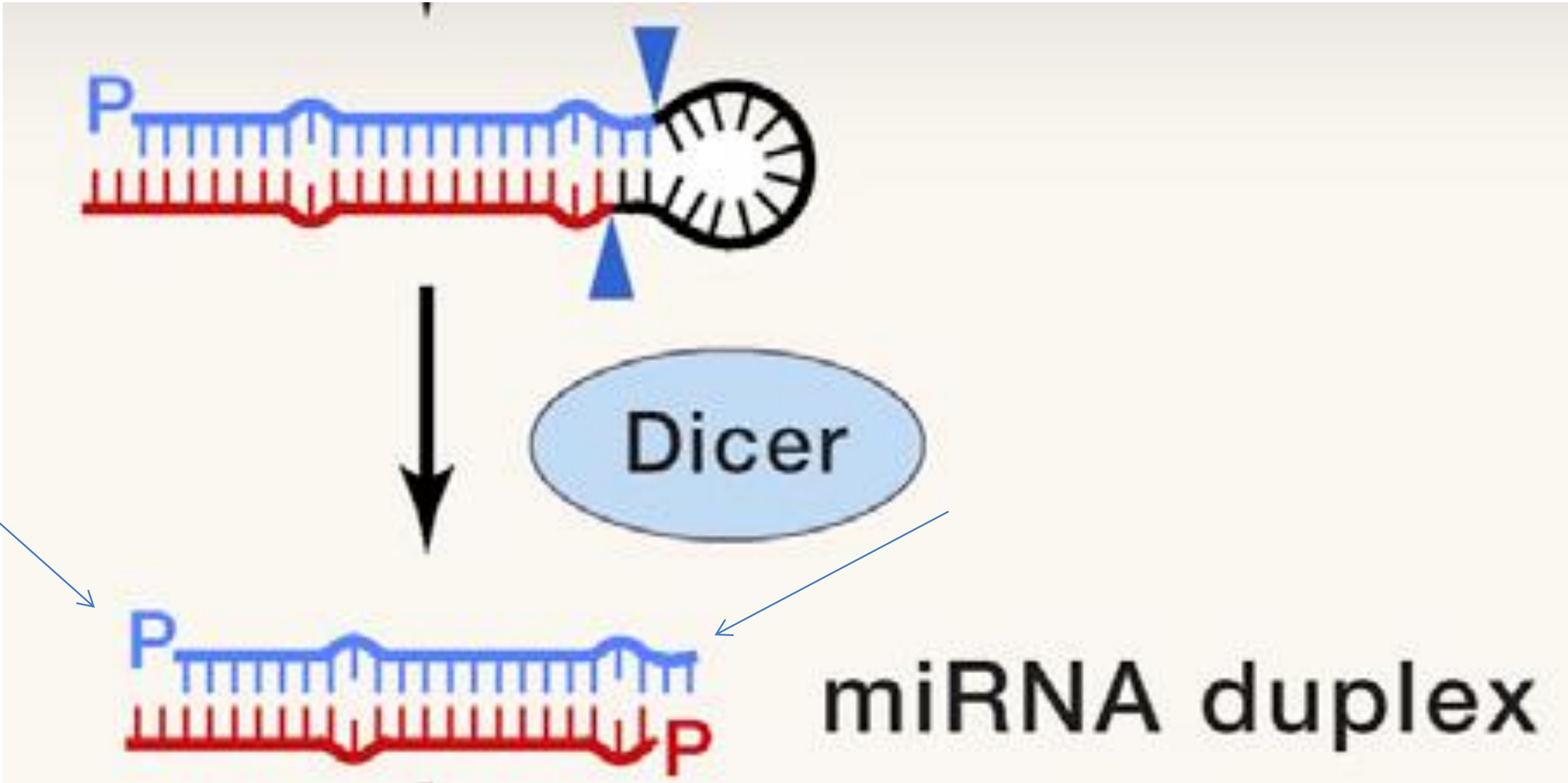


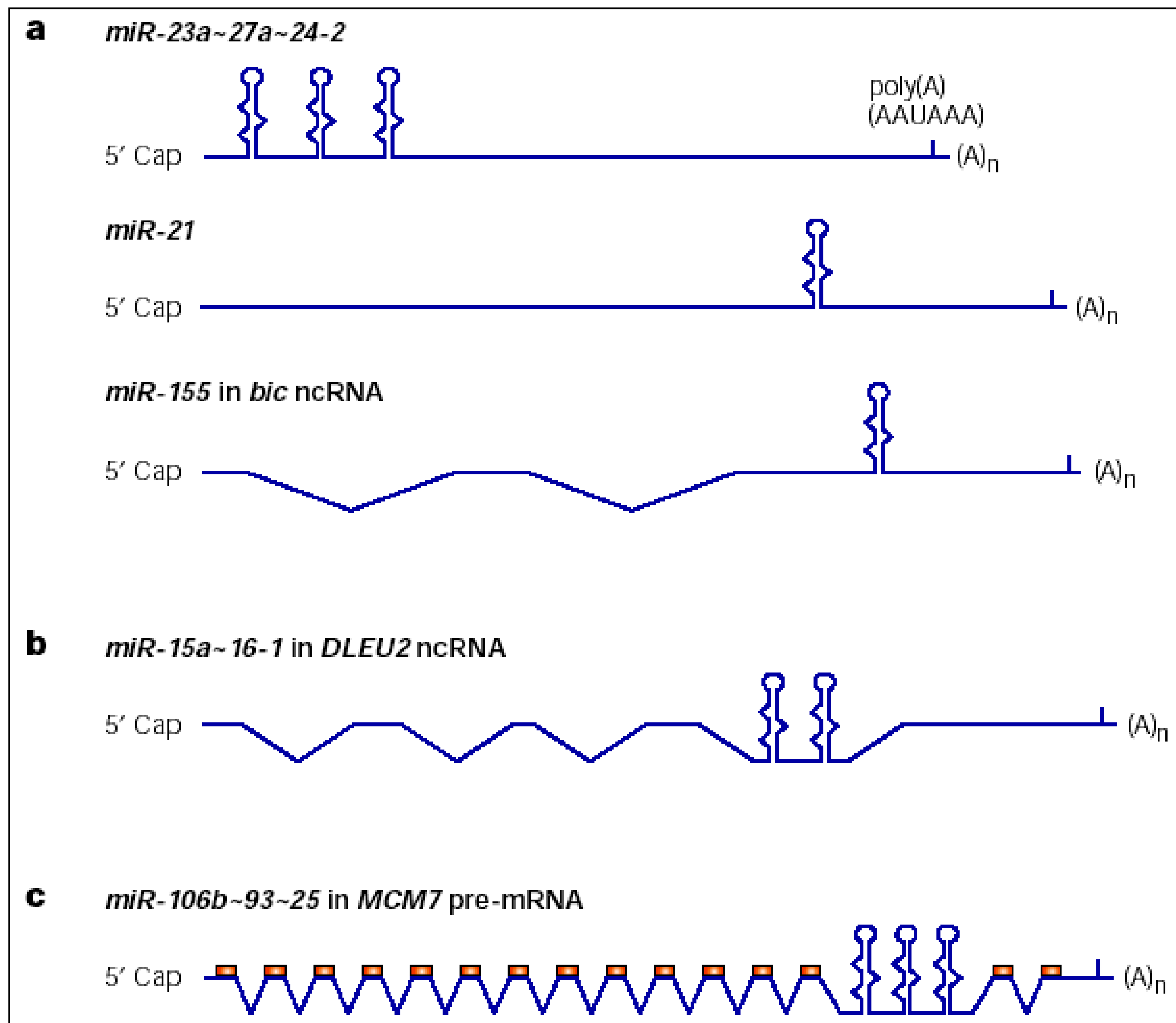
Silencing complex

Is slicing activity present in all Ago proteins in Mammals ?






P
Dicer
P
P
miRNA duplex



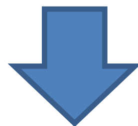
The structure of five pri-miRNAs. (This figure is not to scale**)**

a | Exonic miRNAs in non-coding transcripts

b | Intronic miRNAs in non-coding transcripts.

c | Intronic miRNAs in protein-coding transcripts.

miRNA biogenesis



A
Cannonical hairpins
Gene
Gene
Pri-miRNA transcript
Micro-
processor
P
or
P
Pre-miRNA
hairpin
Dicer
P
P
miRNA duplex
AGO
P
Silencing
complex
mRNA
repression

B
Mirtrons
Gene
Exon
Exon
Mirtron
Splicing
Exon
Exon
Mirtron lariat
Debranching

C
Endogenous shRNAs
Gene
Pol III
Dicer
PPP
P
miRNA duplex
VA RNA
Gene
Trimming
Dicer
P
Pre-miRNA
hairpin
miRNA
duplex

D
Chimeric hairpins
snoRNA
processing
Gene
ACA45 snoRNA
Processing
HVS
snRNA
pre-
snRNA
Trimming
Initial snRNA
processing
HVS pri-miRNA
Gene
Processing
Pol III
Gene
MHV pri-miRNA

E
miR-451
Gene
PARN
Cleavage
Micro-
processor
P
P
P
AAAAA
AGO
AGO
Pre-miR-451
Pri-miR-451

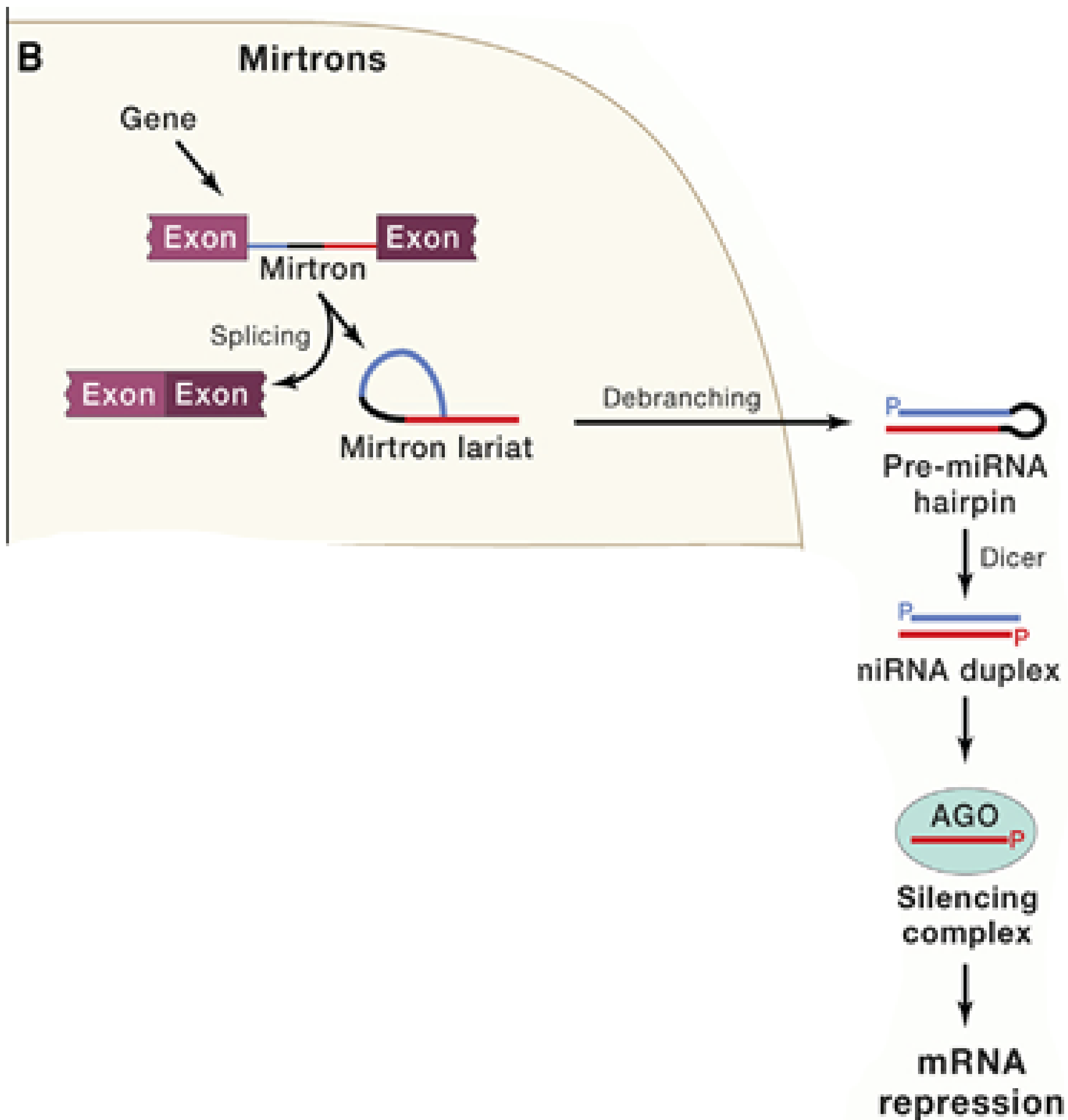
Genomic organization and transcription of miRNA genes

miRNAs are encoded in diverse regions of the genome including both protein coding and non-coding transcription units.

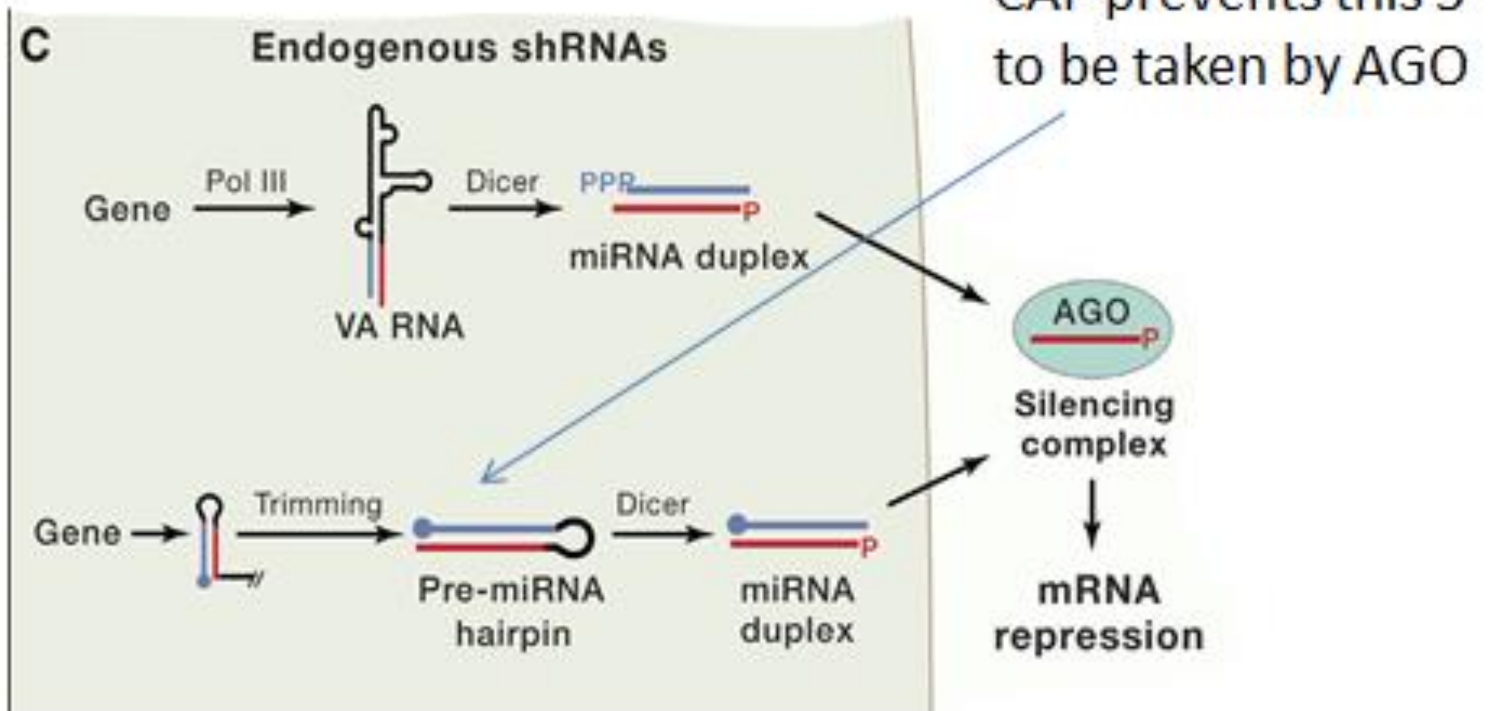
Approximately 50% of miRNAs are derived from noncoding RNA transcripts, while an additional ~40% are located within the introns of protein coding genes.

The majority of miRNAs are transcribed by RNA polymerase (RNA pol) II and bear a 7-methyl guanylate cap at the 5' end and poly (A) tail at the 3' end.

RNA pol III has also been demonstrated to generate the transcripts of a subset of miRNAs.

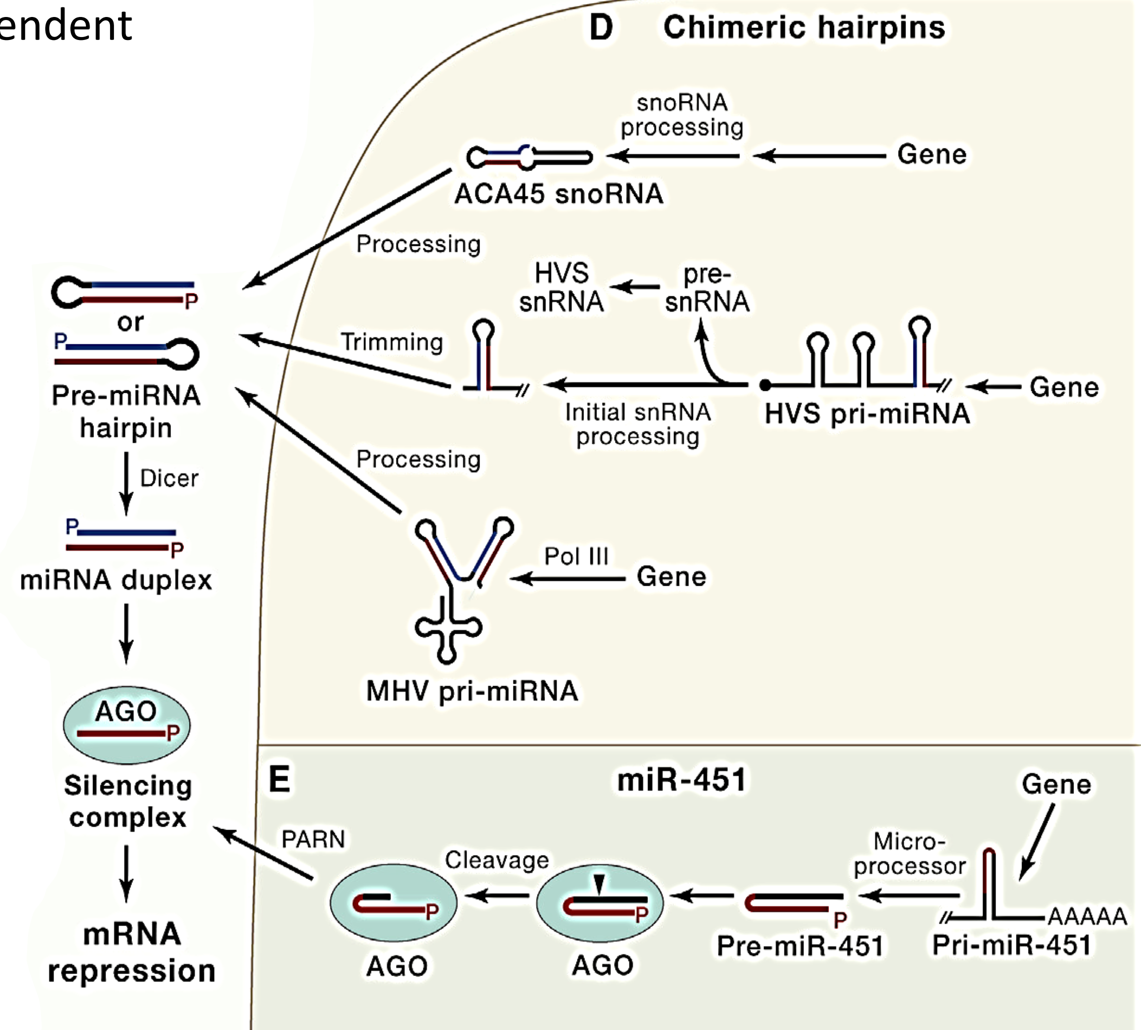


B
Mirtrons
Gene
Exon
Exon
Mirtron
Splicing
Exon
Exon
Mirtron lariat
Debranching
P
Pre-miRNA
hairpin
Dicer
P
P
niRNA duplex
AGO
P
Silencing
complex
mRNA
repression



Drosha-independent

Drosha-independent

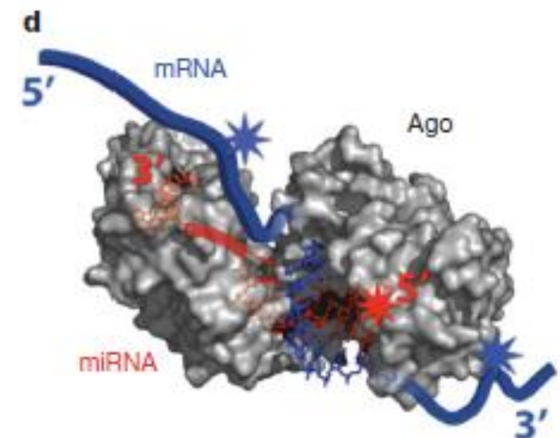


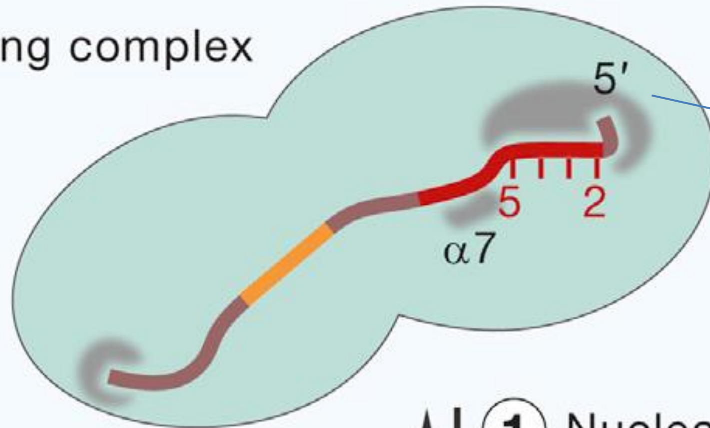
How is the micro-RNA recognized by AGO proteins ?

Which strand is kept as **guide** ?

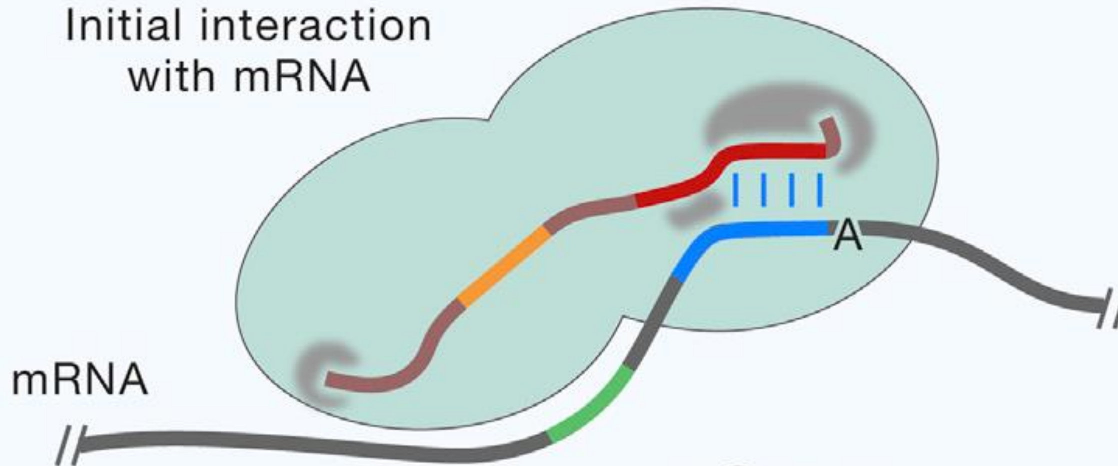
Apparently depends on two factors:

1. an «A» or «U» is preferred as first base (5’-phospho-)
2. the 5’-end with the lowest thermodynamic stability preferred





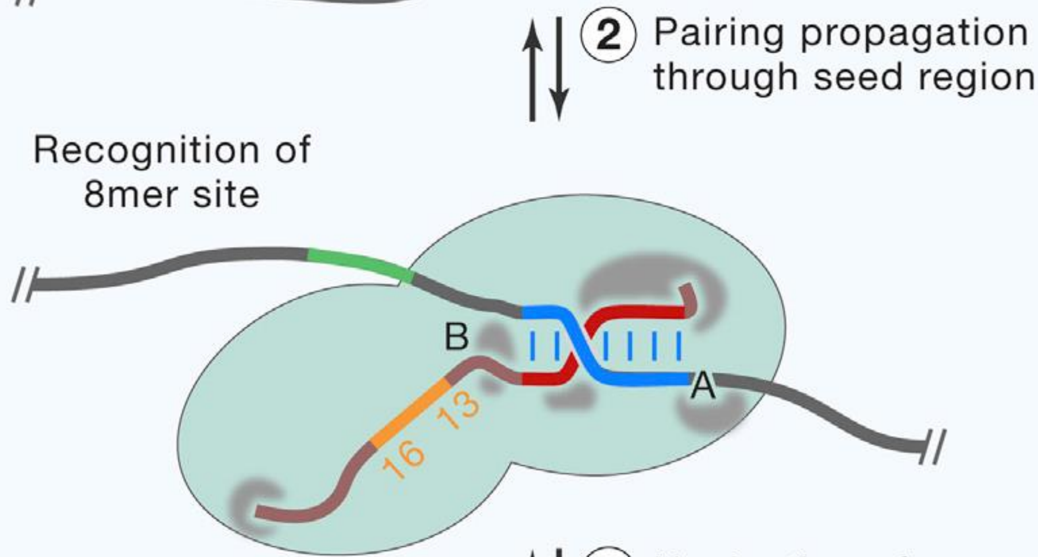
① Nucleation of seed pairing

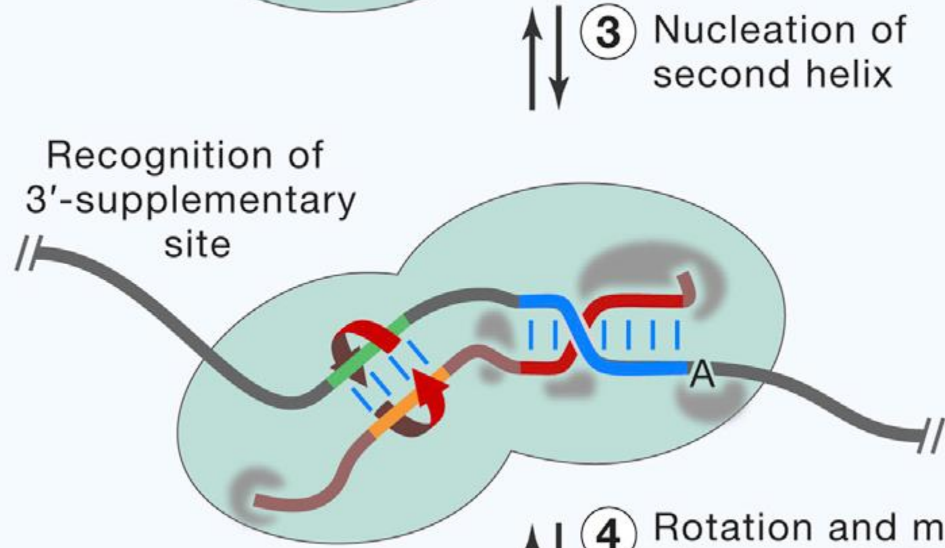


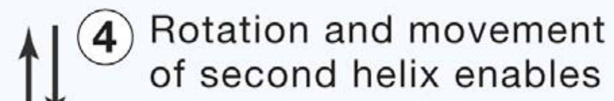
② Pairing propagation through seed region

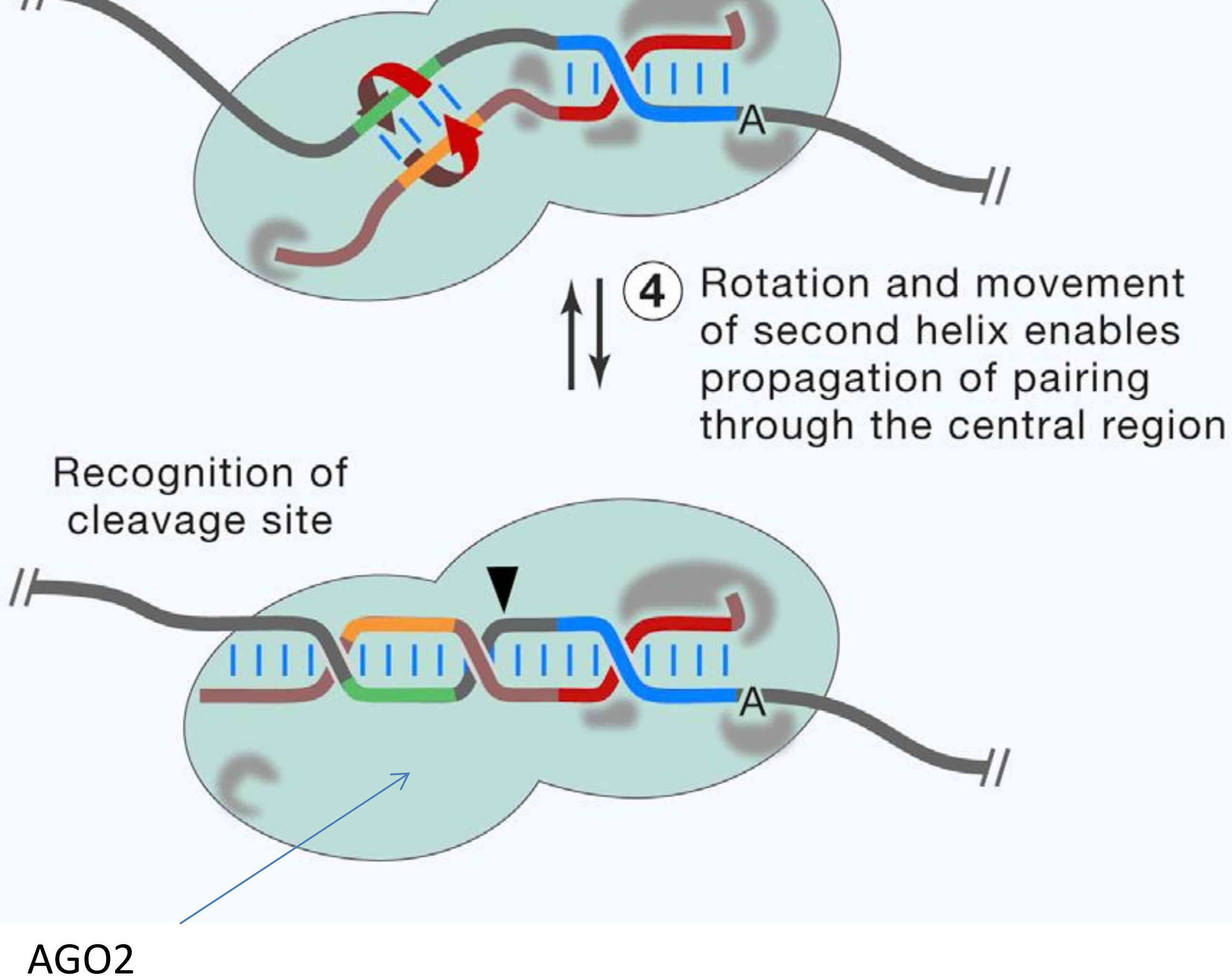
Phospho-required
«A» or «U» preferred

First nucleotide bound by AGO


2 Pairing propagation through seed region
Recognition of 8mer site
B
A
13
16


3 Nucleation of second helix
Recognition of 3′-supplementary site
A


4 Rotation and movement of second helix enables



AGO2

How do miRNAs direct target recognition ?

How do micro-RNAs direct target recognition ?

Examples from first discovered miRNA-mRNA pairs.
The 5’-end sequence of miRNA is the main determinant of targeting:
Continuous Watson-Crick pairing is required here (6-8 nt).
Target sequences mainly in mRNA 3’-UTR.

seed

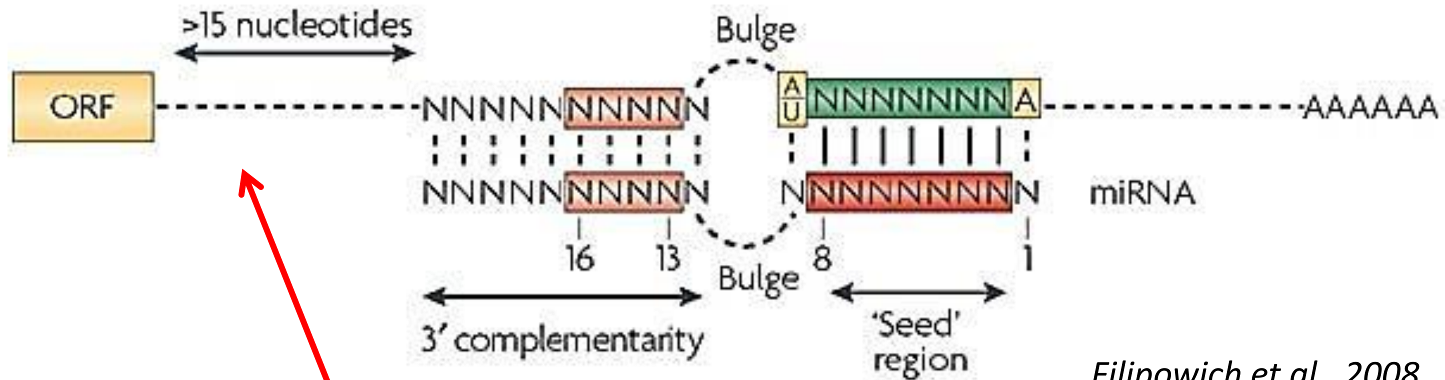
Example

```
HMGA2    5'  CCGACAUUCAAUUUCUACCUCA 3'
                    :  |  :|  ||||||||
let-7a   3'  UUGAUAUGUUGGAUGAUGGAGU 5'

NF2      5'  UACAAGAGAUUCUCCUGCCUCA 3'
              || |  |  |  :||  ||:|||||
let-7a   3'  UUGAUAUGUUGGAUGAUGGAGU 5'

E2F2     5'  GUGGGUGCU-CUGGGCUGAACCA  3'
              ||   | ||  |||::||||  |||
miR-24   3'  GACA-AGGACGACUUGACUCGGU  5'

DNMT3B   5'  UGGCAAAGAAGAUGUUUUGUGGUGCACUGAG 3'
               :||||      |||:|||:|||||||||
miR-148  3'   -UGUUU......CAAGACAUCACGUGACU- 5'
```



Filipowich et al., 2008

Nature Reviews | Genetics

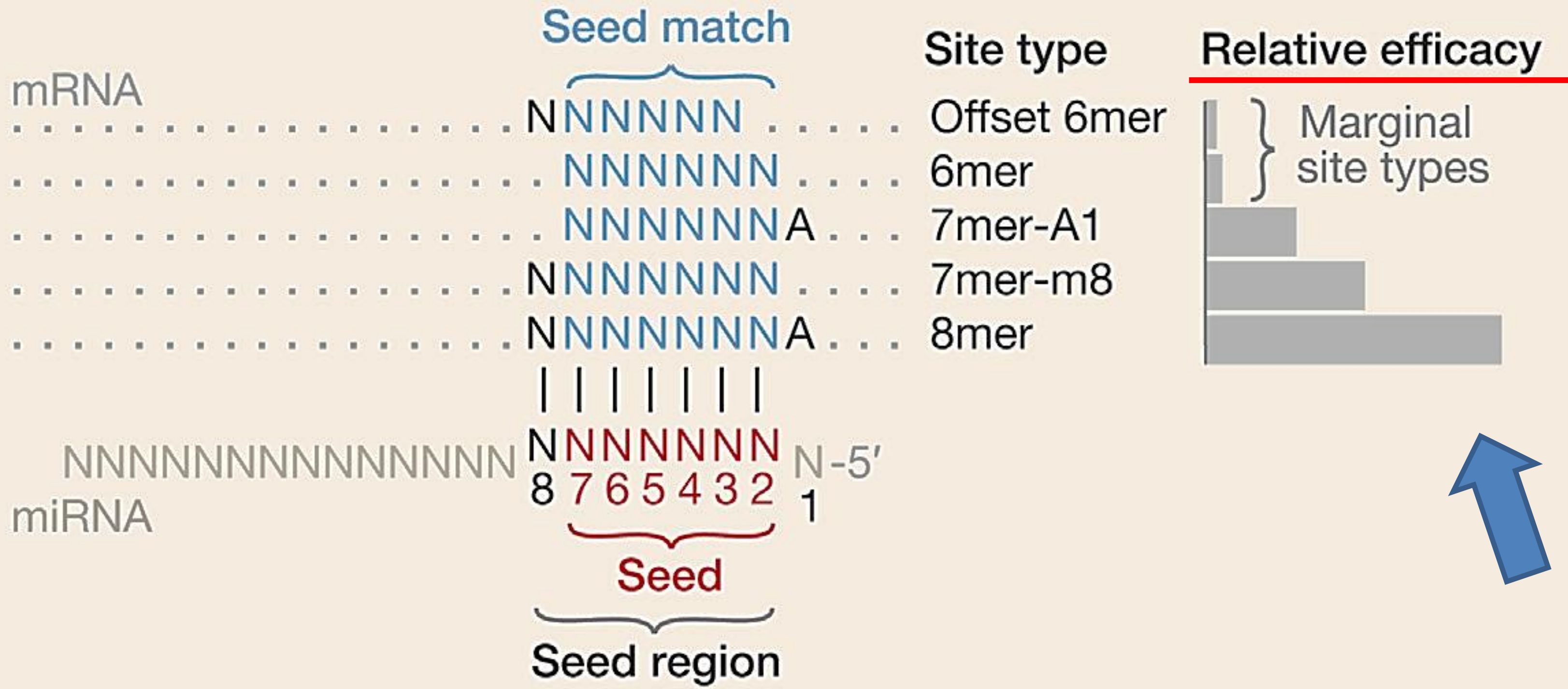
Position in 3'-UTR matters !

As a rule of thumb, miRNA recognize their targets using 6-8 nt at their 5'-end, followed by variable loop and some additional complementarity towards 3'-end.

Due to this variable recognition mode, each miRNA recognizes multiple mRNA (or other RNAs) with variable affinity (can calculate from duplex stability).

The union of possible RNA targets of a specific miRNA is called «regulome».

Several algorithms were developed to predict miRNA targets, and are available on websites. However, due to uncertainty of recognition modes, these algorithms work clearly in sub-optimal way.



Canonical sites of mammalian miRNAs. These canonical sites each have 6–7 contiguous Watson–Crick pairs (vertical lines) to the seed region of the miRNA (miRNA positions 2–8). Two of these sites also include an A at position 1. Relative site efficacy in mammalian cells is graphed to the right (log scale). The most effective canonical sites are 7–8 nt sites that include a perfect match to the miRNA seed (positions 2–7, red), whereas the 6 nt sites are the least effective.

B **Atypical canonical sites**

Supplementary pairing (≥ 3–4 pairs)

Short loop

Seed match

mRNA

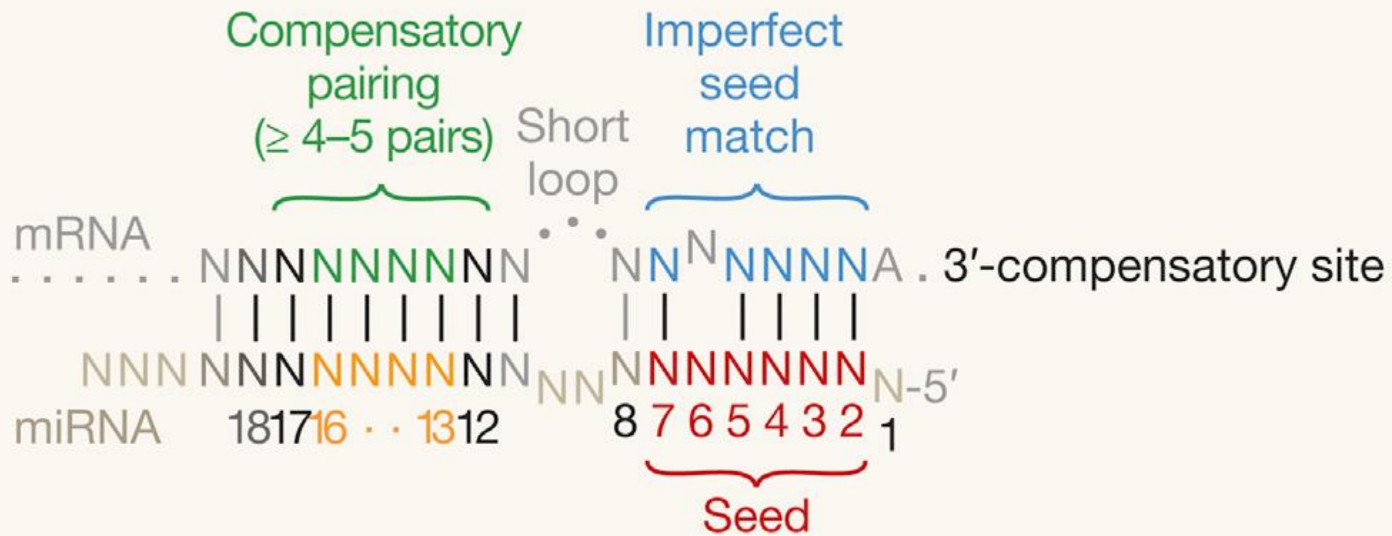
. . . . NNNNNNNNN NNNNNNNA . 3′-supplementary site

NNNNNNNNNNNNNNNNNNN-5′

miRNA 18 17 16 · · 13 12 8 7 6 5 4 3 2 1

Seed

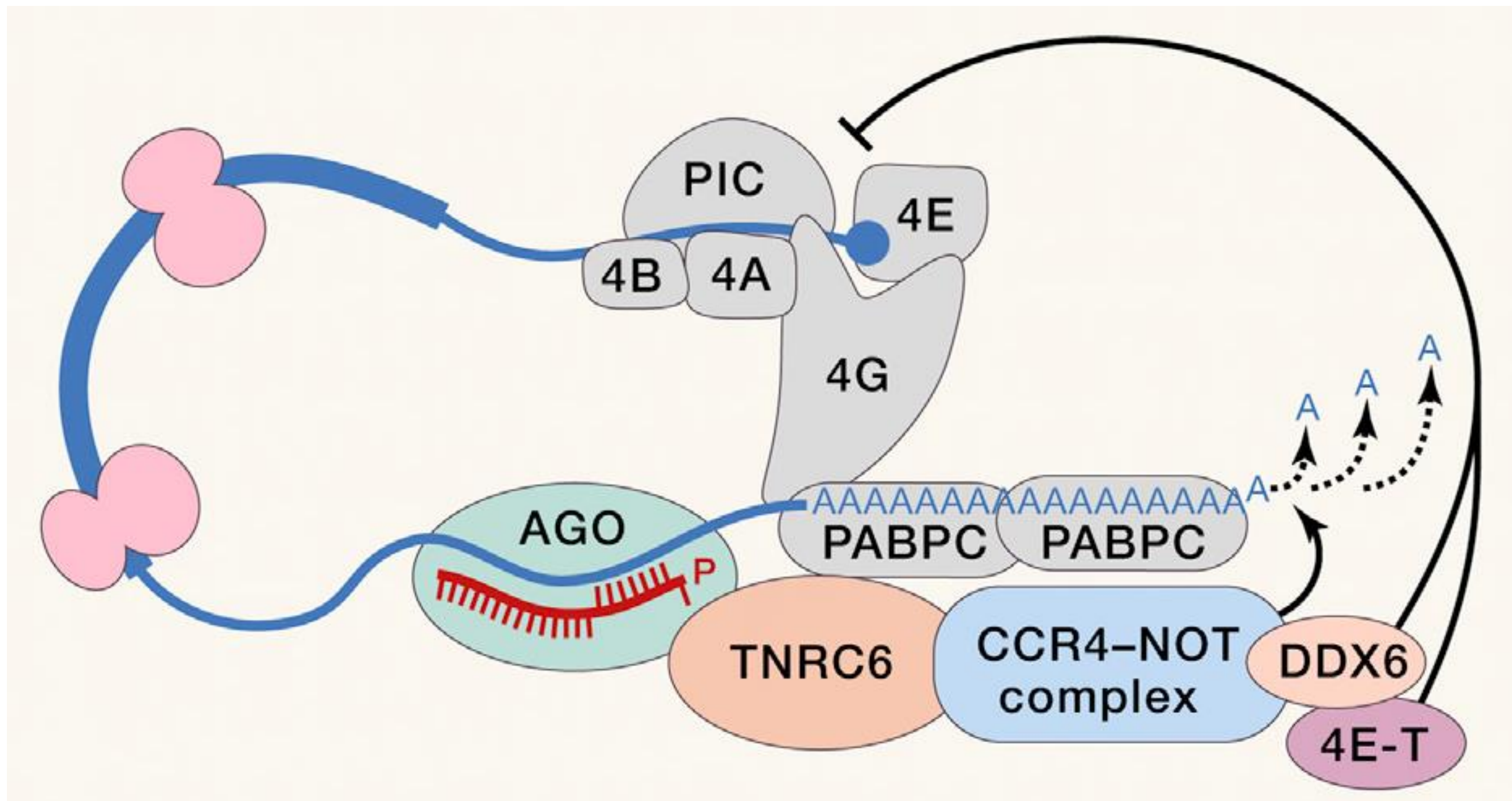
The 3’-supplementary site, an atypical type of canonical site.



The 3’-compensitory site, a functional type of noncanonical sites

How is post-transcriptional silencing performed ?

The Dominant Mechanisms of miRNA-Guided Repression in Bilaterian Animals



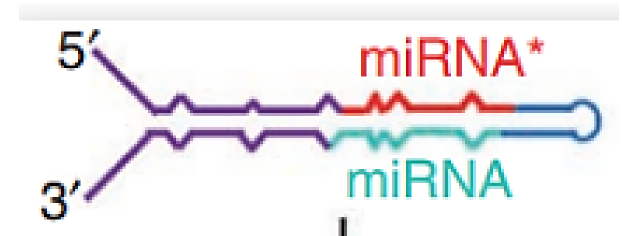
Slicing is determined only when maximal pairing between the micro-RNA and target RNA is present (unusual for miRNA)

This is the condition we see with siRNA (natural or laboratory)

miRNA nomenclature

Names/identifiers in database: hsa-mir-121 (first letter=species; numbering: sequential)
Gene: mir-121
Mature miRNA: miR-121

If there are distinct precursors / genes expressing the same miRNA, add suffix number:
hsa-mir-121-1
hsa-mir-121-2
Mature miRNAs with closely related sequences:
hsa-miR-121a
hsa-miR-121b



In some case, two functional miRNAs derive from the two arms of the same precursor.
If mature forms unbalanced quantitatively: miR-56 the major form (*guide*)
miR-56* the minor form (*passenger*)
If no data or no difference: miR-142-5p (from the 5' arm)
miR-142-3p (from the 3' arm)

(from MiRBase database - http://www.mirbase.org/help/nomenclature.shtml

500 to 600 **verified** miRNAs in Humans. Low % of total predicted.

Some are conserved down to Fish. Conservation very important to predict functional miRNAs.

miRNA divided in Families
– same or similar target recognition.

Each miRNA (family) recognizes hundreds of target mRNA (regulon)

The target sets typically show specifically enriched GO terms

Functions of miRNA

The typical mode of target recognition by miRNA, which is limited to few base pairs, suggests that each miRNA may recognize several mRNAs.

This is exactly what experimentally was seen. Each miRNA down-regulates the expression of a set of mRNAs (regulon).

Strikingly, these targets are enriched by one to several GO terms that are coherent with the physiological action of that miRNA.

Micro RNA have been intensively studied during last ten years.

- ✓ Several miRNAs expressed in each cell
- ✓ Strong tissue- and cell-specificity
- ✓ Developmentally and signal transduction regulated
- ✓ Individual miRNA deletion or overexpression: severe effects on cell physiology
- ✓ miRNA expression profile severely altered in pathologies
- ✓ miRNAs quite stable in serum: possible biomarkers for diseases.

Note: in Humans, 96 miRNA genes (88 families) are conserved among placental mammals but not in Vertebrates. A third of these families are in two paternally imprinted clusters.

Functions of miRNA

miRNA are involved in the control of the expression of virtually all genes

involved in most biological phenomena and processes

important in regulatory circuits for feed-forward and feed-back actions.

Many simple circuits involving one miRNA and one mRNA in cell fate decision processes

Many studies attribute the phenotype of KD/KO to a single miRNA/mRNA interaction

Quite unlikely unless proven

Main proof: remove or mutate MRE from target mRNA and see whether this will phenocopy the miRNA.

----

miRNAs and mRNAs (and most likely also a nuber of lncRNAs) constitute a **network**

while it is possible that individual deletion phenotype can be attributed to one prevalen target (as also in the case of overexpression , e.g. cancer) their physiological role has to be seen in the context of a complex network

How is expression of micro-RNA regulated ?

miRNA and regulatory networks

Regulation of miRNA expression: Transcription

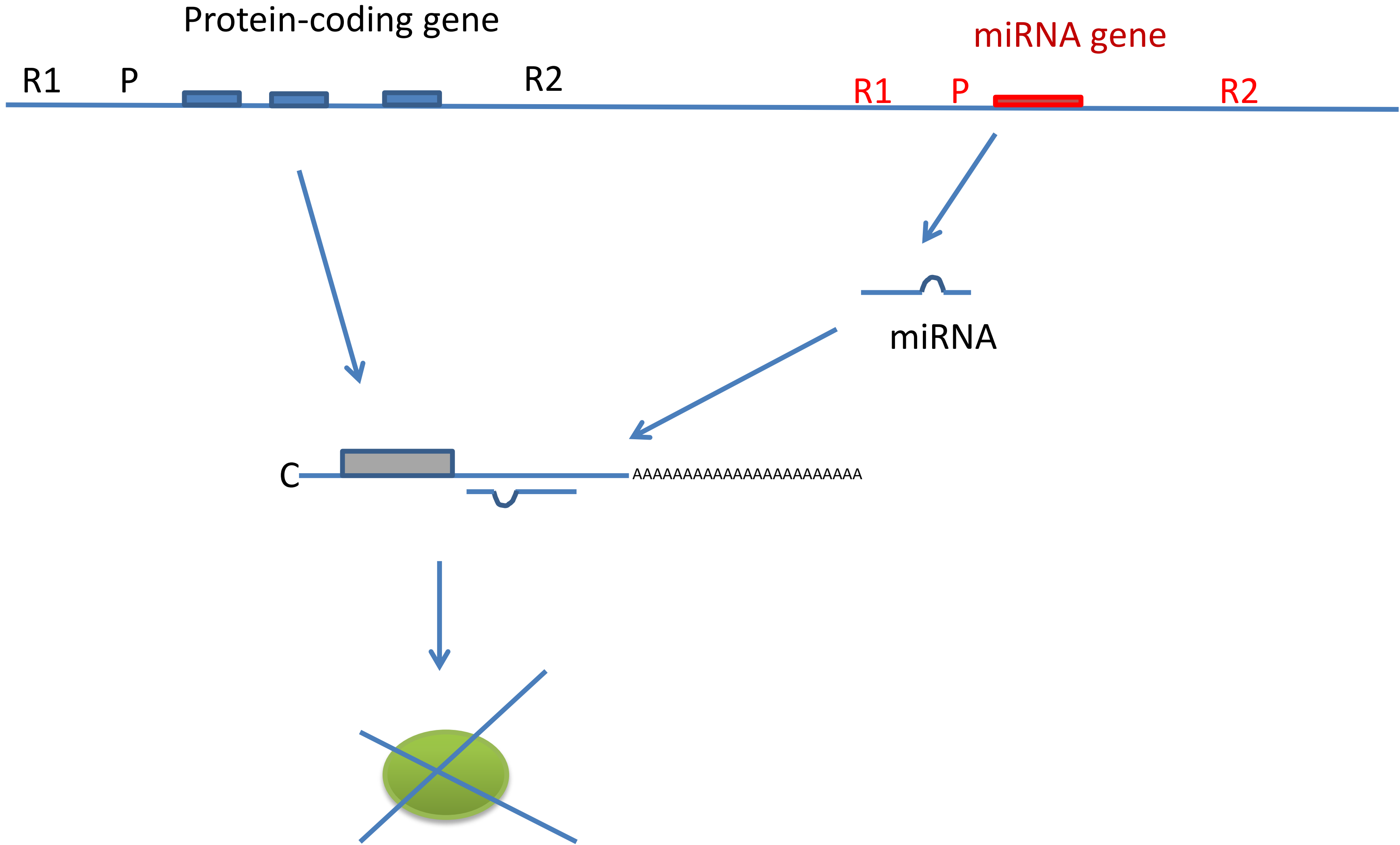
A recent large scale mapping of 175 human miRNA promoters through nucleosome positioning and chromatin immunoprecipitationon-genomic DNA microarray chip (or ChIP-onchip) analysis suggests that

the promoter structure of miRNA genes, including the relative frequencies of CpG islands, TATA box, TFIIB recognition, initiator elements, and histone modifications, is indistinguishable between the promoters of miRNA and mRNA.

Furthermore, DNA binding factors that regulate miRNA transcription largely overlap with those that control protein coding genes.

As a further proof, the expression of specific miRNA is also controlled by signal transduction pathways, in analogy to protein-coding genes.

A two-component regulatory circuit



Experimental:

Suppression of miRNA can be achieved by antisense technologies, especially using nonhydrolyzable oligos (anti-miR).

Over-expression: minigenes expressing pri-miRNA or pre-miRNA in constitutive or inducible vectors. Also synthetic miRNA mimics in transient transfection.

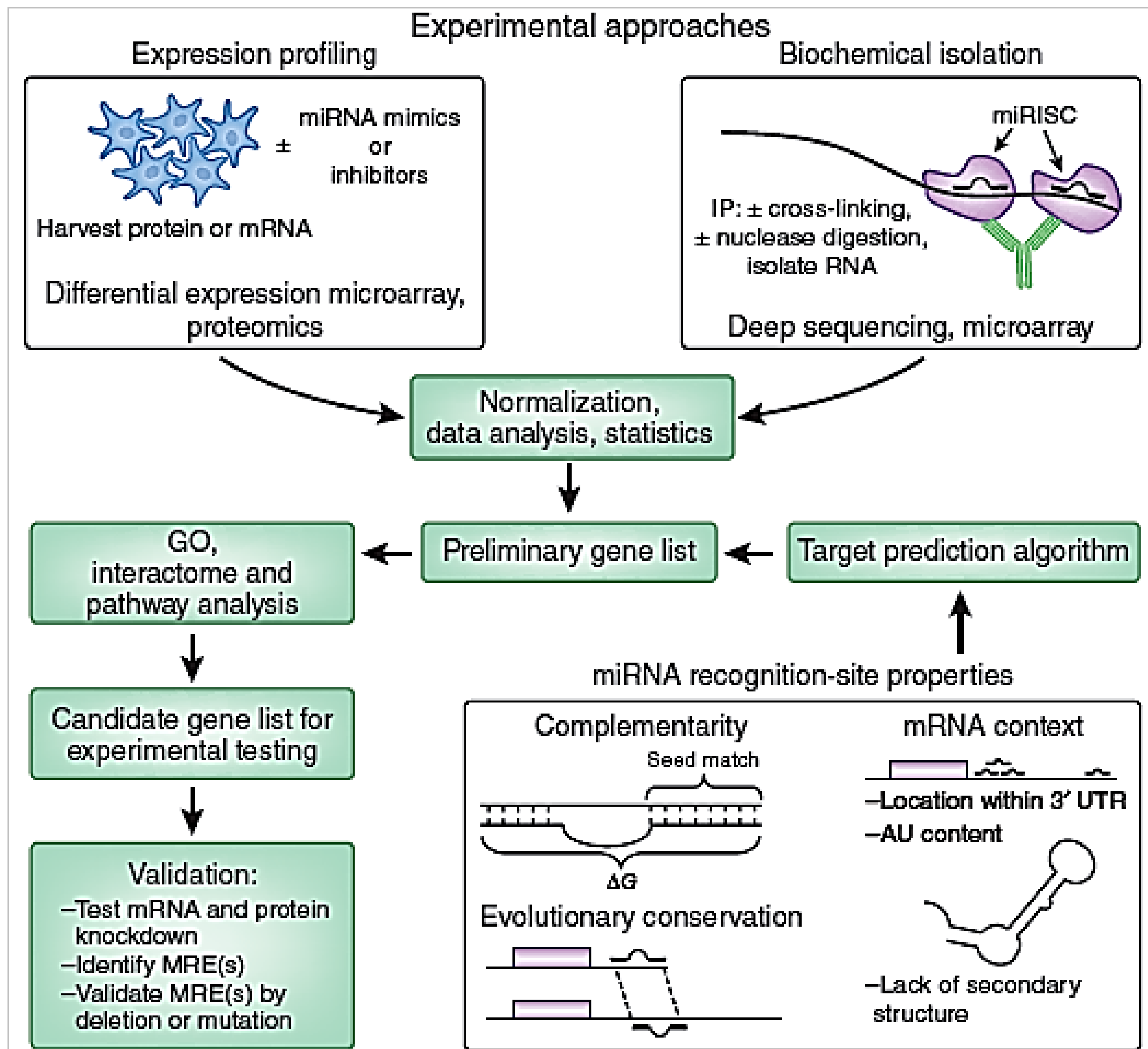
Modern: CRISPR-mediated deletion

Expression analysis:

RT-PCR methods available (problem: they are very short!)

Microarrays carrying probes for all known miRNAs available, also made of LNA oligos instead of normal oligos to increase hybridization specificity

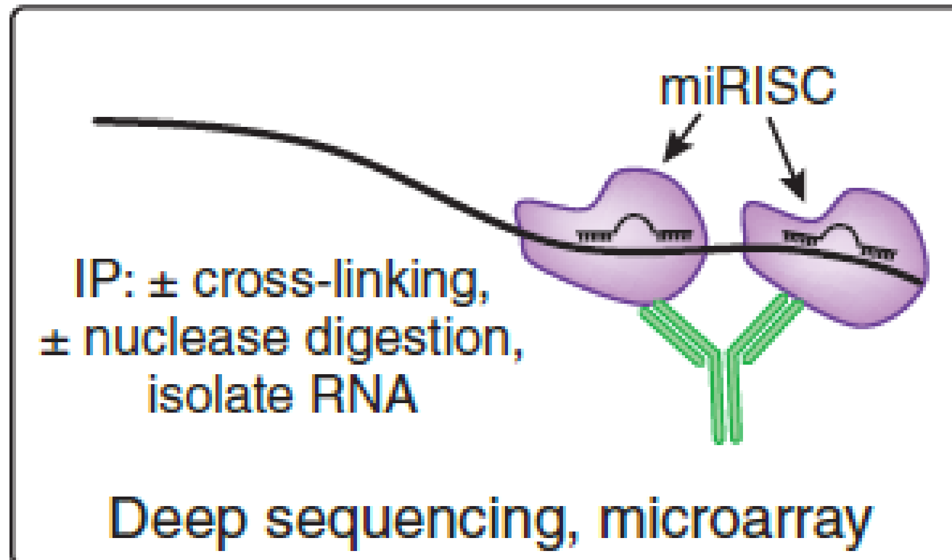
Specific protocol and application for **short RNA-seq NGS**.



MRE=miRNA Response Element

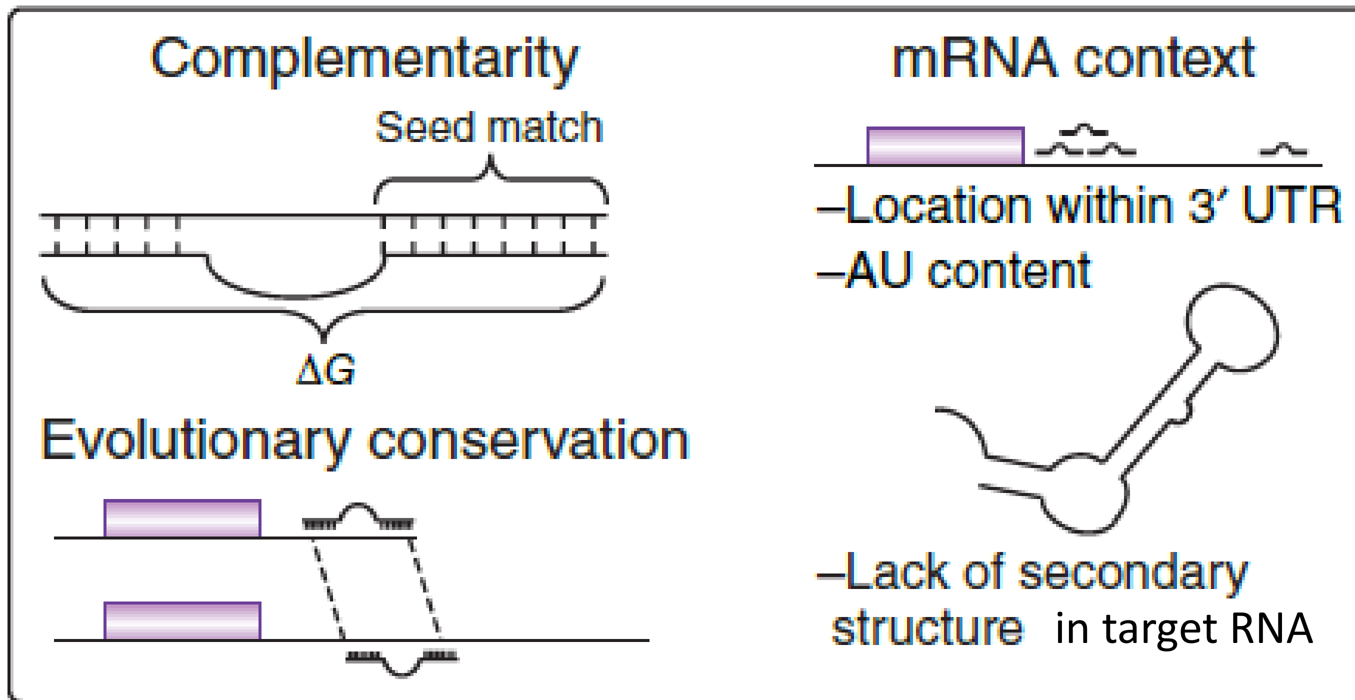
Expression profiling

- «antagomir»
- Vectors to express miRNAs
- miRNA mimics
- CRISPR deletion



- RIP
- CLIP
- HITS-CLIP
- PAR-CLIP

Bioinformatics



Number of MREs very important

Most evolved models take into account all these aspects: performance close to that of Ligase-mediated HITS-CLIP **(Your Research Paper No. 4 !**)

Example of the RNA-Protein immunoprecipitation methods (RIP, CLIP, HITS-CLIP and further methods)

Vol 460 | 23 July 2009 | doi:10.1038/nature08170 nature

ARTICLES

Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps

Sung Wook Chi[1], Julie B. Zang[1], Aldo Mele[1] & Robert B. Darnell[1]

MicroRNAs (miRNAs) have critical roles in the regulation of gene expression; however, as miRNA activity requires base pairing with only 6–8 nucleotides of messenger RNA, predicting target mRNAs is a major challenge. Recently, high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation (HITS-CLIP) has identified functional protein–RNA interaction sites. Here we use HITS-CLIP to covalently crosslink native argonaute (Ago, also called Eif2c) protein–RNA complexes in mouse brain. This produced two simultaneous data sets—Ago–miRNA and Ago–mRNA binding sites—that were combined with bioinformatic analysis to identify interaction sites between miRNA and target mRNA. We validated genome-wide interaction maps for miR-124, and generated additional maps for the 20 most abundant miRNAs present in P13 mouse brain. Ago HITS-CLIP provides a general platform for exploring the specificity and range of miRNA action *in vivo*, and identifies precise sequences for targeting clinically relevant miRNA–mRNA interactions.

IP using the anti-AGO 2AB antibody reveals different complexes

a Protein
IP: IgG 2A8
kDa
160
105
75
50
* Ago
IgG
IB: Ago

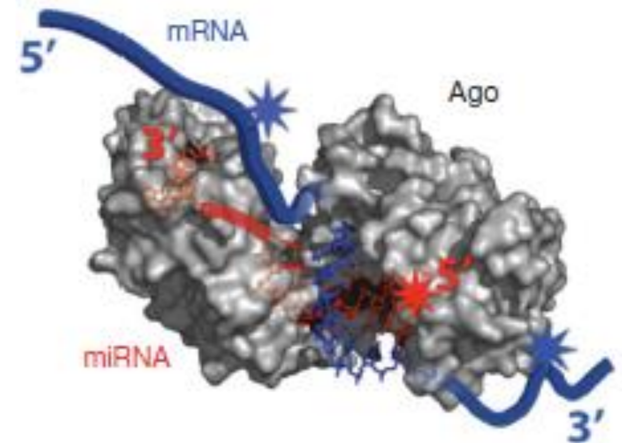
b RNA
IP: IgG 2A8
kDa
130
110
90
^{32}P-labelled RNA

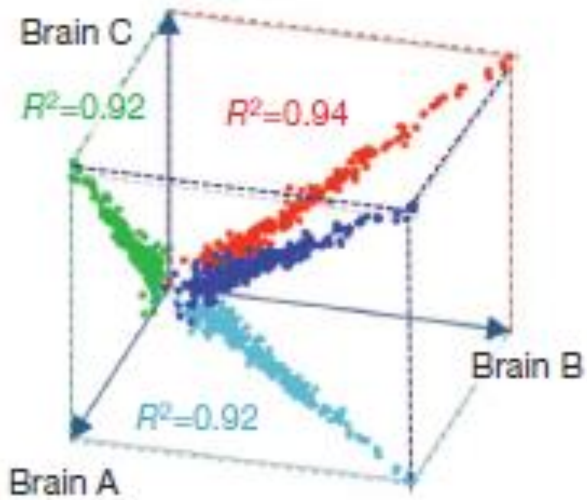
c cDNA
90 110 130 (kDa)
bp
200
mRNA
miRNA
50

NGS

Short NGS

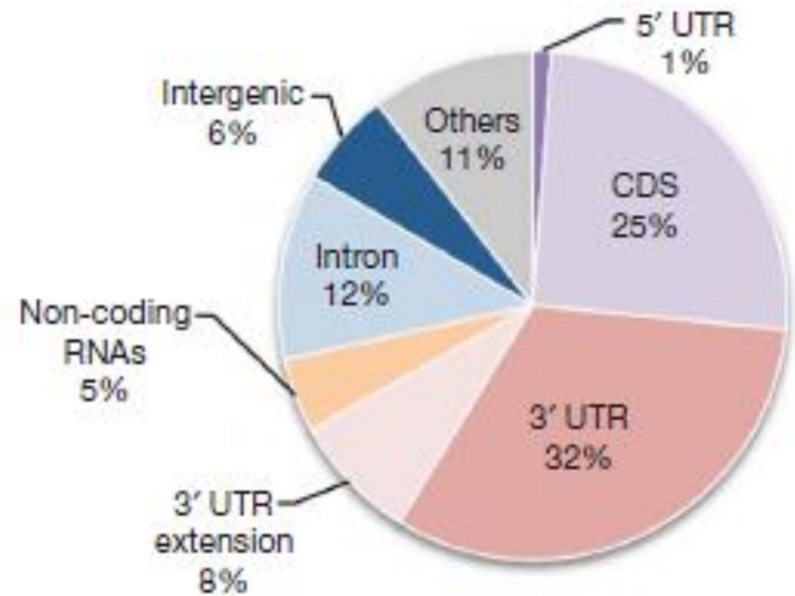
Extract bands, add 36-nt adapters, RT and PCR

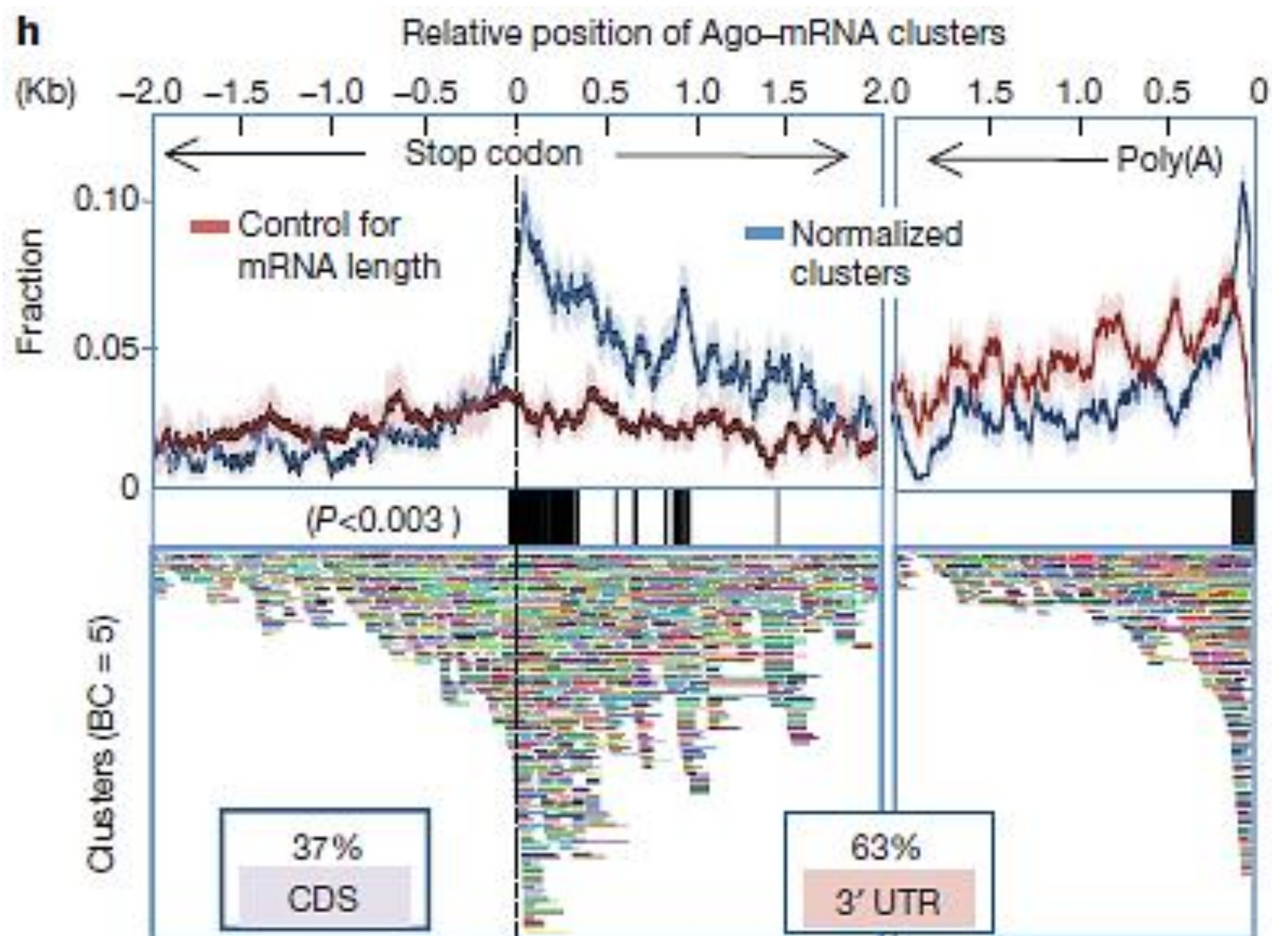




Three replicates from brain

Mapping of targets





h
Relative position of Ago–mRNA clusters
(Kb)
–2.0
–1.5
–1.0
–0.5
0
0.5
1.0
1.5
2.0
1.5
1.0
0.5
0
Stop codon
Poly(A)
Fraction
0.10
0.05
0
Control for mRNA length
Normalized clusters
(P<0.003)
Clusters (BC = 5)
37%
CDS
63%
3′ UTR

Main problems associated with CLIP and derivatives:

1. cross-linking requires intense, short-wavelength irradiation and this yields quite high non-specific results, due to AGO occasional «touch» with flanking mRNA sequences or completely unrelated RNAs.
2. two separate libraries must be generated, and the lists matched bioinformatically (guess level).
3. target (m)RNA sequence protected by AGO footprint is quite long and gives uncertainty to the possible target sequence.

Trying to bypass this problem, researchers developed PAR-CLIP (next slide)

Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP

Markus Hafner,[1,5] Markus Landthaler,[1,4,5] Lukas Burger,[2] Mohsen Khorshid,[2] Jean Hausser,[2] Philipp Berninger,[2] Andrea Rothballer,[1] Manuel Ascano, Jr.,[1] Anna-Carina Jungkamp,[1,4] Mathias Munschauer,[1] Alexander Ulrich,[1] Greg S. Wardle,[1] Scott Dewell,[3] Mihaela Zavolan,[2,*] and Thomas Tuschl[1,*]
[1]Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, 1230 York Avenue, Box 186, New York, NY 10065, USA
[2]Biozentrum der Universität Basel and Swiss Institute of Bioinformatics (SIB), Klingelbergstr. 50-70, CH-4056 Basel, Switzerland
[3]Genomics Resource Center, The Rockefeller University, 1230 York Avenue, Box 241, New York, NY 10065, USA
[4]Present address: Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, 13125 Berlin, Germany
[5]These authors contributed equally to this work
*Correspondence: mihaela.zavolan@unibas.ch (M.Z.), ttuschl@rockefeller.edu (T.T.)

RNA transcripts are subject to posttranscriptional gene regulation involving hundreds of RNA-binding proteins (RBPs) and microRNA-containing ribonucleoprotein complexes (miRNPs) expressed in a cell-type dependent fashion. We developed a cell-based crosslinking approach to determine at high resolution and transcriptome-wide the binding sites of cellular RBPs and miRNPs. The crosslinked sites are revealed by thymidine to cytidine transitions in the cDNAs prepared from immunopurified RNPs of 4-thiouridine-treated cells. We determined the binding sites and regulatory consequences for several intensely studied RBPs and miRNPs, including PUM2, QKI, IGF2BP1-3, AGO/EIF2C1-4 and TNRC6A-C. Our study revealed that these factors bind thousands of sites containing defined sequence motifs and have distinct preferences for exonic versus intronic or coding versus untranslated transcript regions. The precise mapping of binding sites across the transcriptome will be critical to the interpretation of the rapidly emerging data on genetic variation between individuals and how these variations contribute to complex genetic diseases.

UV, 365 nm

4SU 4SU 4SU 4SU RBP

immunoprecipitation

lysis, IP,
RNase T1 treatment,
T4 PNK, γ-^{32}P-ATP

^{32}P — U^{xl} — OH

SDS-PAGE
autoradiography

XL-RBP

electroelution
proteinase K treatment

^{32}P — U^{xl} — OH

cDNA library preparation,
PCR amplification

C
G

Solexa sequencing

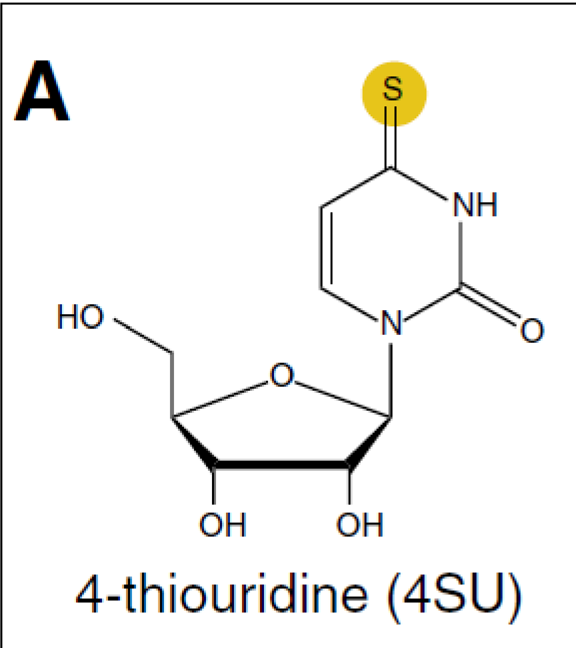
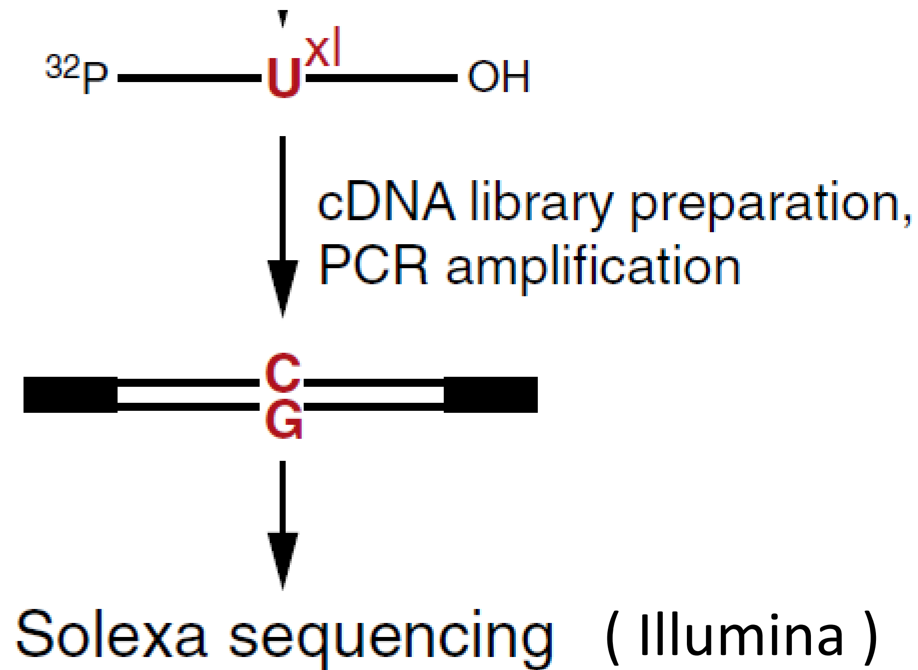


Figure 1. PAR-CLIP Methodology.

(A) Structure of photoactivatable nucleosides.

(C) Illustration of PAR-CLIP. 4SU-labeled transcripts were crosslinked to RBPs and partially RNase-digested RNA-protein complexes were immunopurified and size-fractionated. RNA molecules were recovered and converted into a cDNA library and deep sequenced.



The «U» that was directly cross-linked to the RBP is identified since it is converted to «C» and consequently the targets are univocally identified.

Major problems for these approaches is that they are indirect, i.e. they are based on the generation of **separate libraries**, one for miRNAs and the other for targets (mRNA, lncRNAs, etc). Matching is always based on complementarity searches.

For this reason, Darnell's group developed a strategy to ligate miRNAs and targets and sequence them together, called CLEAR-CLIP

nature COMMUNICATIONS

ARTICLE

Received 4 May 2015 | Accepted 12 Oct 2015 | Published 25 Nov 2015

DOI: 10.1038/ncomms9864

OPEN

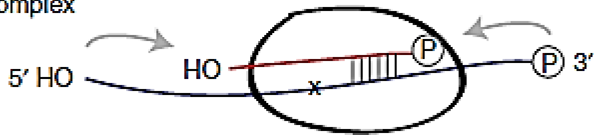
miRNA–target chimeras reveal miRNA 3′-end pairing as a major determinant of Argonaute target specificity

Michael J. Moore[1], Troels K.H. Scheel[2,3,4], Joseph M. Luna[1,2], Christopher Y. Park[1,5], John J. Fak[1], Eiko Nishiuchi[2], Charles M. Rice[2] & Robert B. Darnell[1,5]

Here we report a modified AGO HITS-CLIP strategy termed CLEAR (covalent ligation of endogenous Argonaute-bound RNAs)-CLIP, which enriches miRNAs ligated to their endogenous mRNA targets. CLEAR-CLIP mapped ca. 130,000 endogenous miRNA–target interactions in mouse brain and B40,000 in human hepatoma cells. Motif and structural analysis define expanded pairing rules for over 200 mammalian miRNAs. Most interactions combine seed-based pairing with distinct, miRNA-specific patterns of auxiliary pairing. At some regulatory sites, this specificity confers distinct silencing functions to miRNA family members with shared seed sequences but divergent 30-ends. This work provides a means for explicit biochemical identification of miRNA sites in vivo, leading to the discovery that miRNA 3′-end pairing is a general determinant of AGO binding specificity.

1. Tissue lysis, DNAse
2. RNAse treatment
3. AGO IP
4. High stringency washes

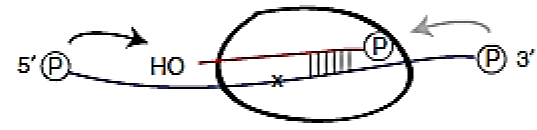
Purified AGO
ternary complex



x=cross-link

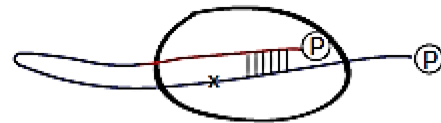
1. 5′ Phosphorylation (PNK, 3′Ⓟase minus)
2. Second stringent wash sequence

f

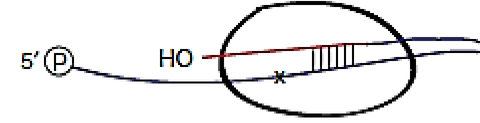


Ligation
(T4 RNA ligase I)

Major product (>90%)



Minor product (<10%)



1. 3′ Dephosphorylation (CIP)
2. 3′ linker addition (on-bead, truncated) RNA ligase 2, pre-adenylated linker)
3. PNK radiolabelling

SDS–PAGE, nitrocellulose transfer, cloning per
published HITS-CLIP protocol

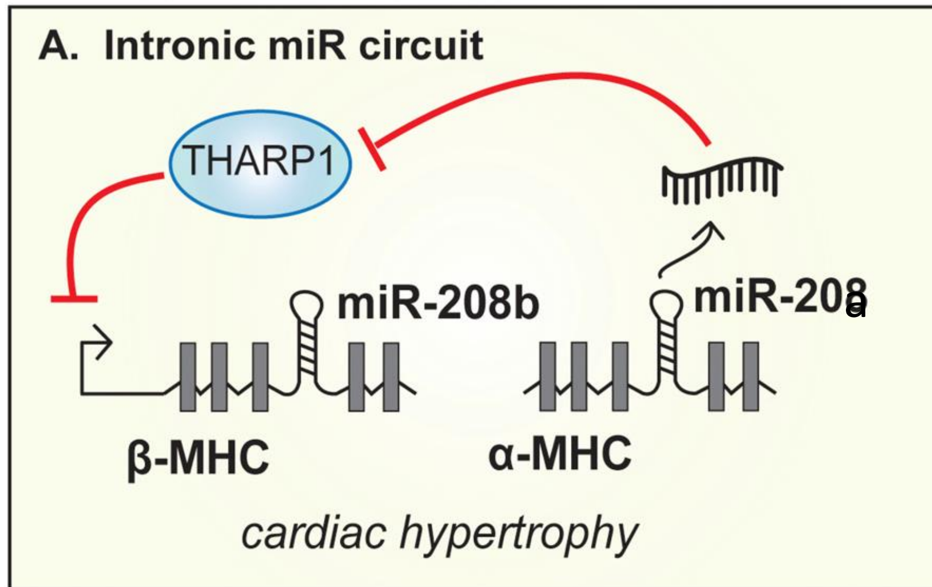
5′ ——— 3′ miRNA–target

5′ ——— 3′ target–miRNA

One library for NGS with joint target-miRNA

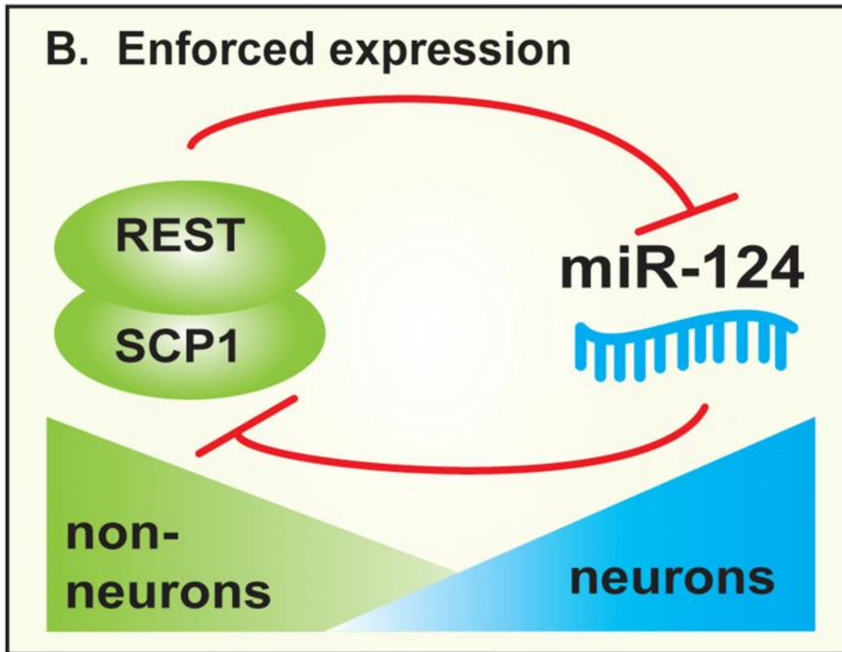
	Major 'miR-first' chimeras	Minor 'miR-last' chimeras
Mapped chimeras	363,953	39,413
Unique chimeras	166,580	10,458
Clustered events	130,120	6,619

miRNA regulatory circuits



The cardiac specific miR-208 family is encoded within the introns of myosin heavy chain (MHC) genes. miR-208a targets THARP1, and will reduce its level, thus increasing expression of β-MHC.

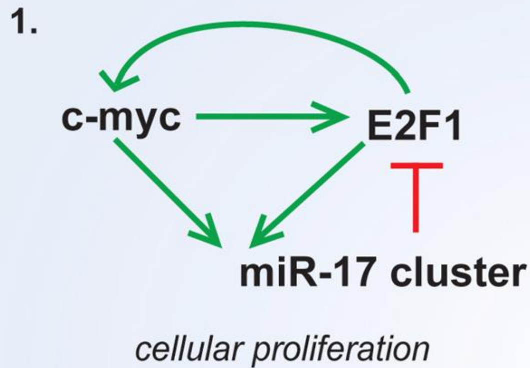
This is an auto-boosting circuit that is found altered in cardiac hypertrophy.

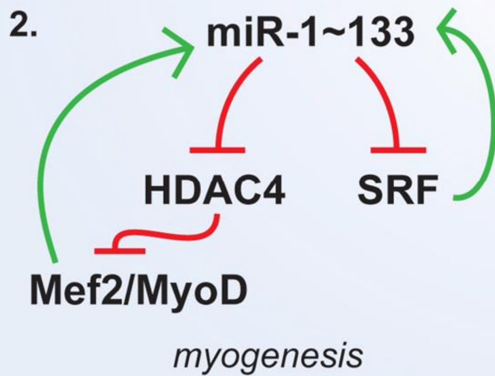


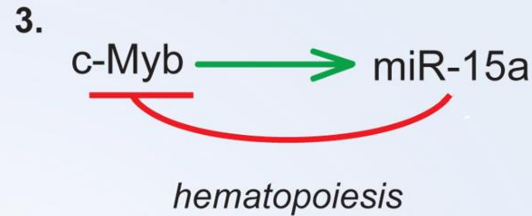
The expression of miR-124 is negatively regulated by the binding of the RE1 silencing transcription (REST) factor to the promoter in non-neuronal cells

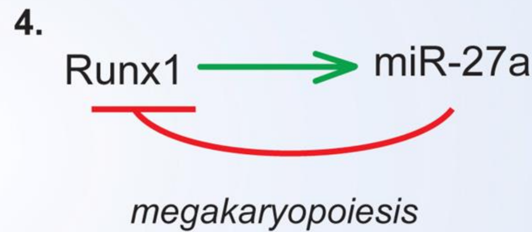
In neurons, miR-124 represses translation of SCP1 that is one component of the REST complex.

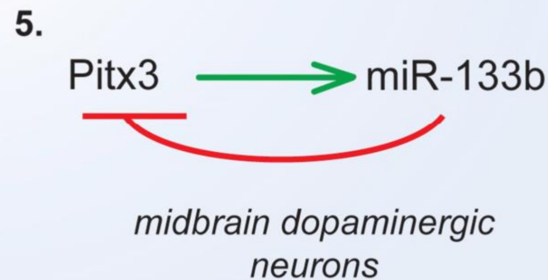
C. Transcription factor regulatory circuits







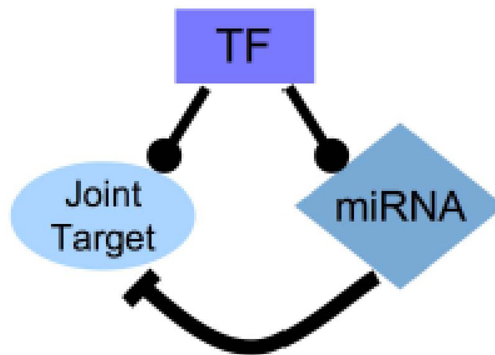


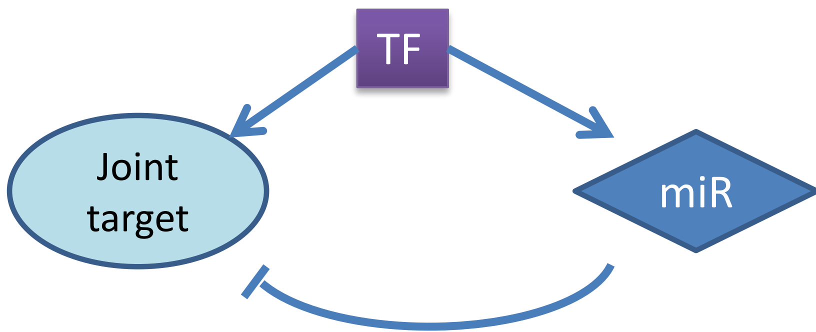


Examples of feed-back regulation of microRNA transcription through the repression of transcription factors.

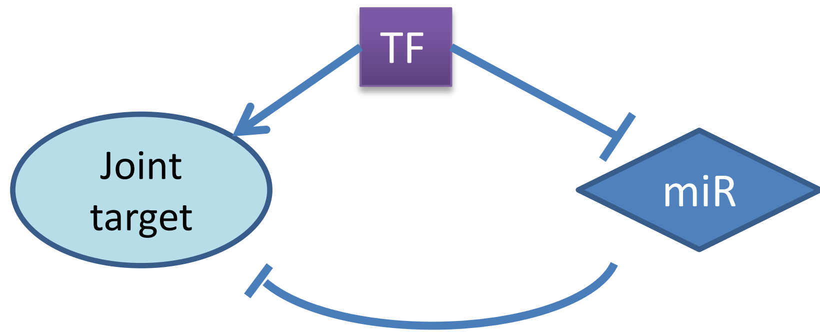
mRNAs that encode for Transcription Factors are very often controlled by miRNA.

This suggests a model where feed-forward and feed-back control loops exist.





uncoherent



coherent

Friard *et al. BMC Bioinformatics* 2010, 11:435
http://www.biomedcentral.com/1471-2105/11/435

BMC Bioinformatics

DATABASE Open Access

CircuitsDB: a database of mixed microRNA/transcription factor feed-forward regulatory circuits in human and mouse

Olivier Friard[1], Angela Re[2], Daniela Taverna[1,3,4], Michele De Bortoli[1,3], Davide Corá[1,5*]

Abstract

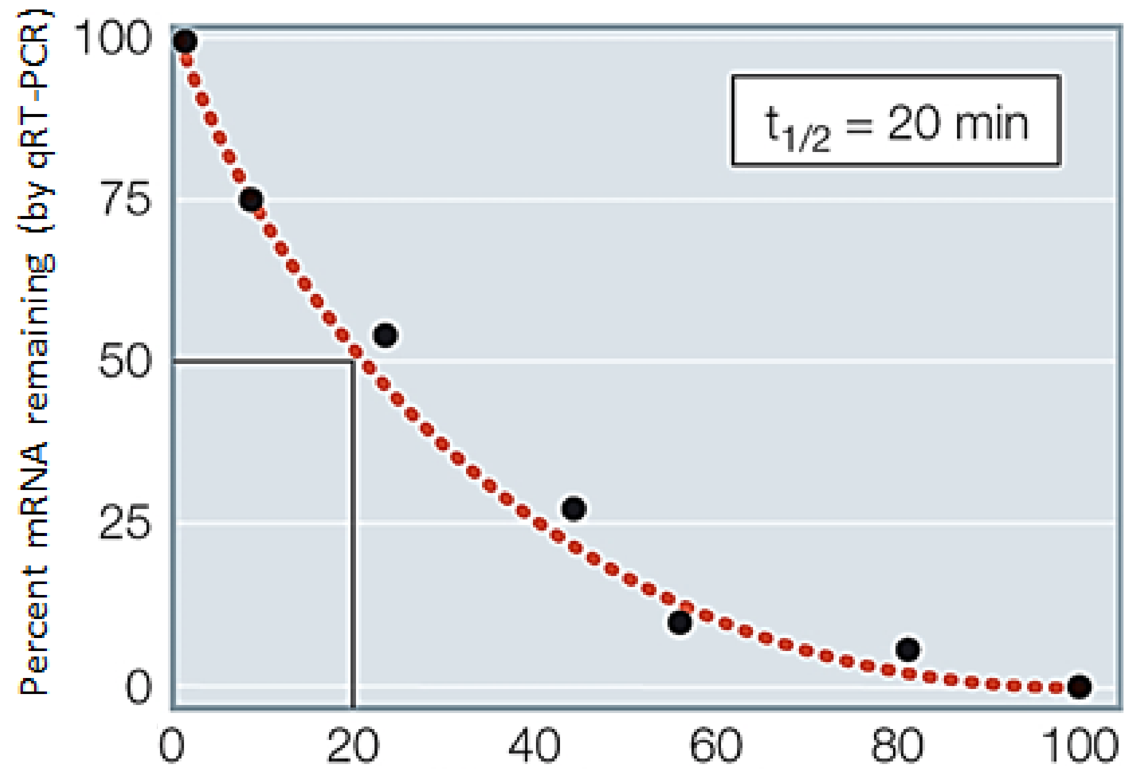
Background: Transcription Factors (TFs) and microRNAs (miRNAs) are key players for gene expression regulation in higher eukaryotes. In the last years, a large amount of bioinformatic studies were devoted to the elucidation of transcriptional and post-transcriptional (mostly miRNA-mediated) regulatory interactions, but little is known about the interplay between them.

Description: Here we describe a dynamic web-accessible database, `CircuitsDB`, supporting a genome-wide transcriptional and post-transcriptional regulatory network integration, for the human and mouse genomes, based on a bioinformatic sequence-analysis approach. In particular, `CircuitsDB` is currently focused on the study of mixed miRNA/TF Feed-Forward regulatory Loops (FFLs), i.e. elementary circuits in which a master TF regulates an miRNA and together with it a set of Joint Target protein-coding genes. The database was constructed using an ab-initio oligo analysis procedure for the identification of the transcriptional and post-transcriptional interactions. Several external sources of information were then pooled together to obtain the functional annotation of the proposed interactions. Results for human and mouse genomes are presented in an integrated web tool, that allows users to explore the circuits, investigate their sequence and functional properties and thus suggest possible biological experiments.

Conclusions: We present `CircuitsDB`, a web-server devoted to the study of human and mouse mixed miRNA/TF Feed-Forward regulatory circuits, freely available at: http://biocluster.di.unito.it/circuits/

How to study mRNA stability and decay

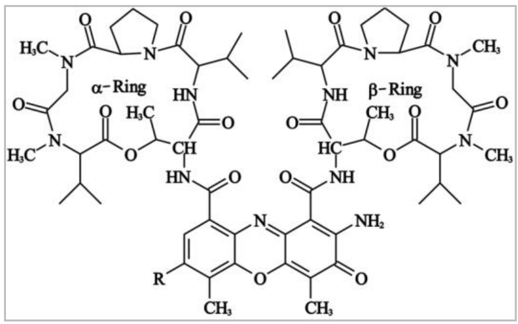
Measuring RNA half-life



- Blocking RNA Polymerase
- Labeling nascent RNA (pulse) and releasing

Cells treated with Actinomycin D

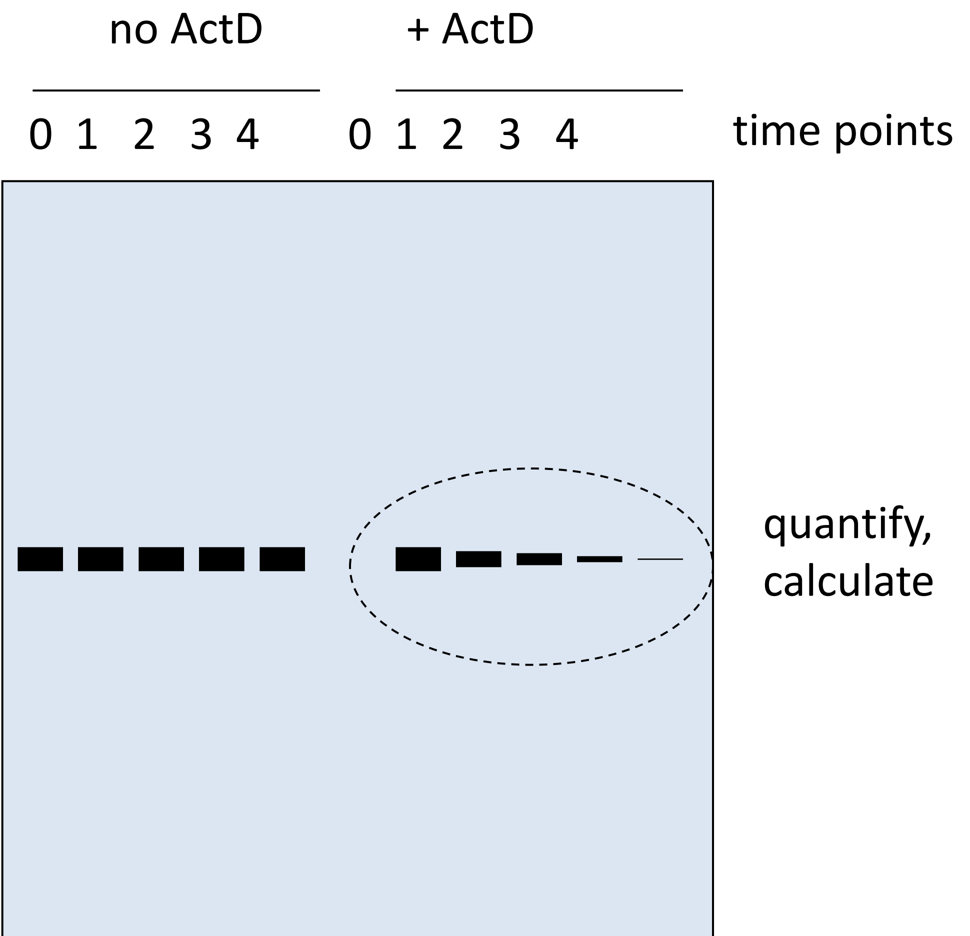
Older measure of mRNA half-life on single genes



RNA extracted at time intervals

Add radioactive protection probe

RNase (ss)



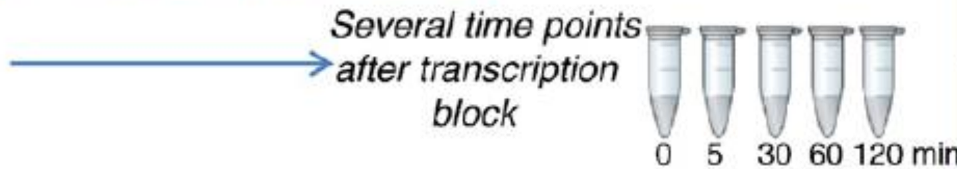
RNase Protection Assay (RPA)

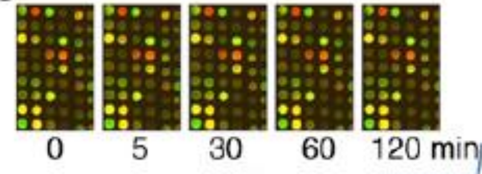
wiki

Genome-wide

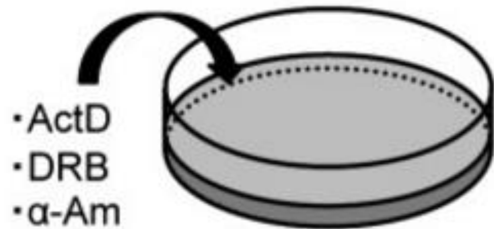
Blockers:
Actinomycin D (ActD),
5,6-dichloro-1–D-ribofuranosyl-benzimidazole (DRB)
α-amanitin (α-Am)

(b)- Transcription arrest





Global transcription arrest **Extract RNA** **Quantification**

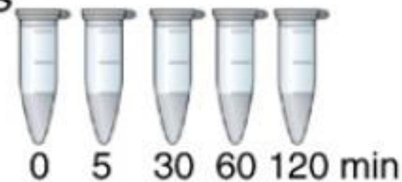


- DNA Microarray
- Deep sequencing

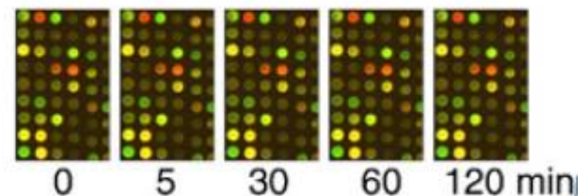
Pulse of modified UTP

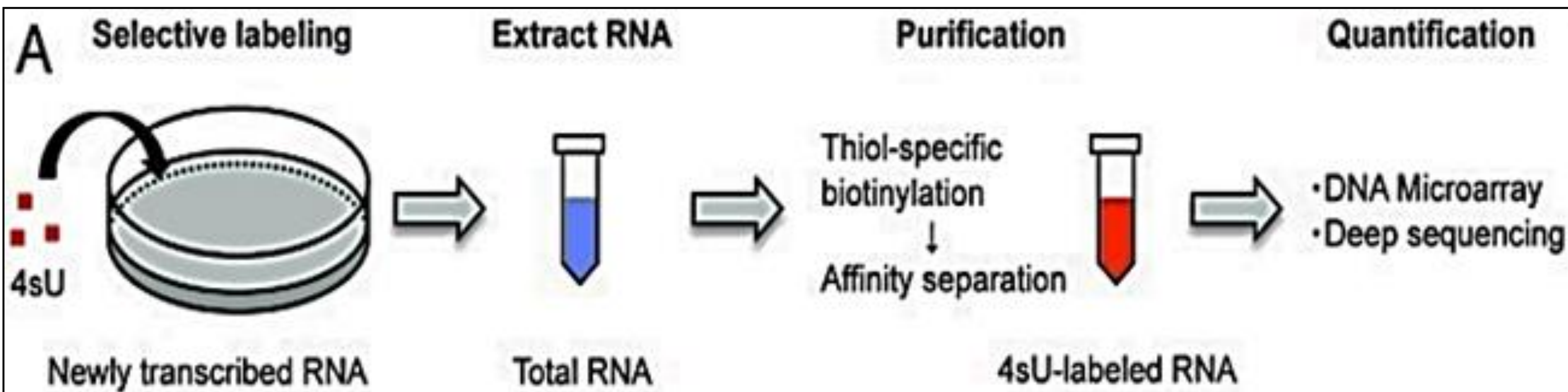
Chase

Several time points after precursor depletion



RNA *fractionation & labeling*





Letter Genome Res. 2003; 13(8):1863-72.

(First study human cells genome-wide: pre-NGS)

Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes

Edward Yang,[1,6] Erik van Nimwegen,[4,6] Mihaela Zavolan,[2] Nikolaus Rajewsky,[5] Mark Schroeder,[2] Marcelo Magnasco,[3] and James E. Darnell Jr.[1,7]

[1]Laboratory of Molecular Cell Biology, [2]Laboratory of Computational Genomics, [3]Laboratory of Mathematical Physics, and [4]Center for the Study of Physics and Biology, The Rockefeller University, New York, New York 10021-6399, USA; [5]Department of Biology and Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, USA

HepG2 cells (human liver carcinoma cell line)
+ primary cells (fibroblasts Bud8)

Actinomycin 2-3 hours

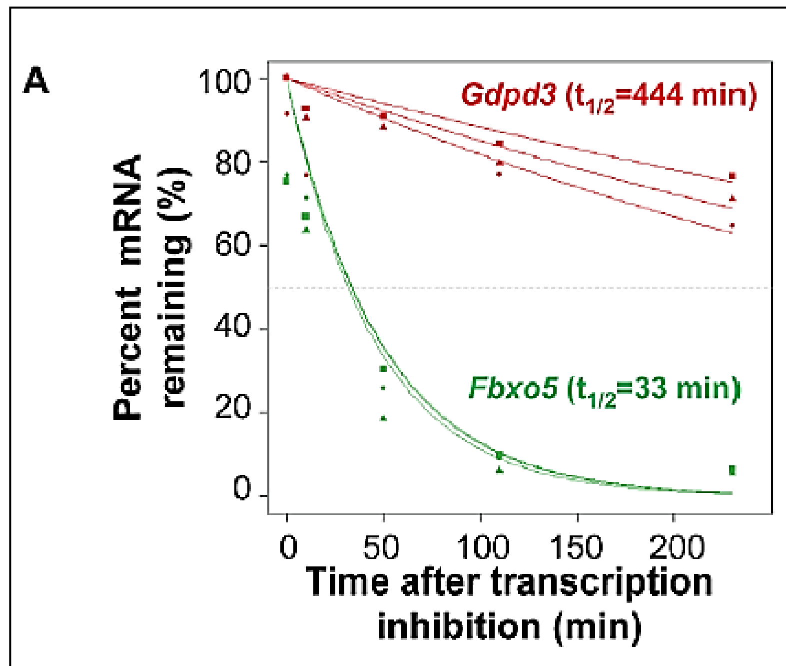
RNA extraction, labelling and → Affymetrix microarrays

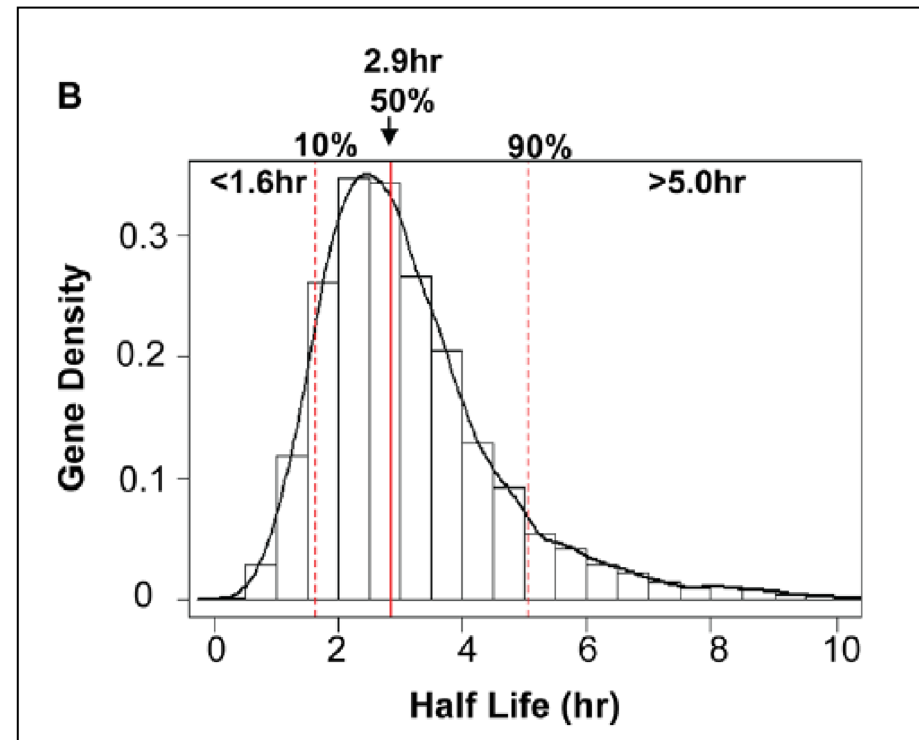
Decay rate estimates for 5,245 genes

The **median half-life** in both cell types is ~10 h, with wide range (0.5 hours up to “days”).

Lee et al., PLOS One, 5:11201 (2010)

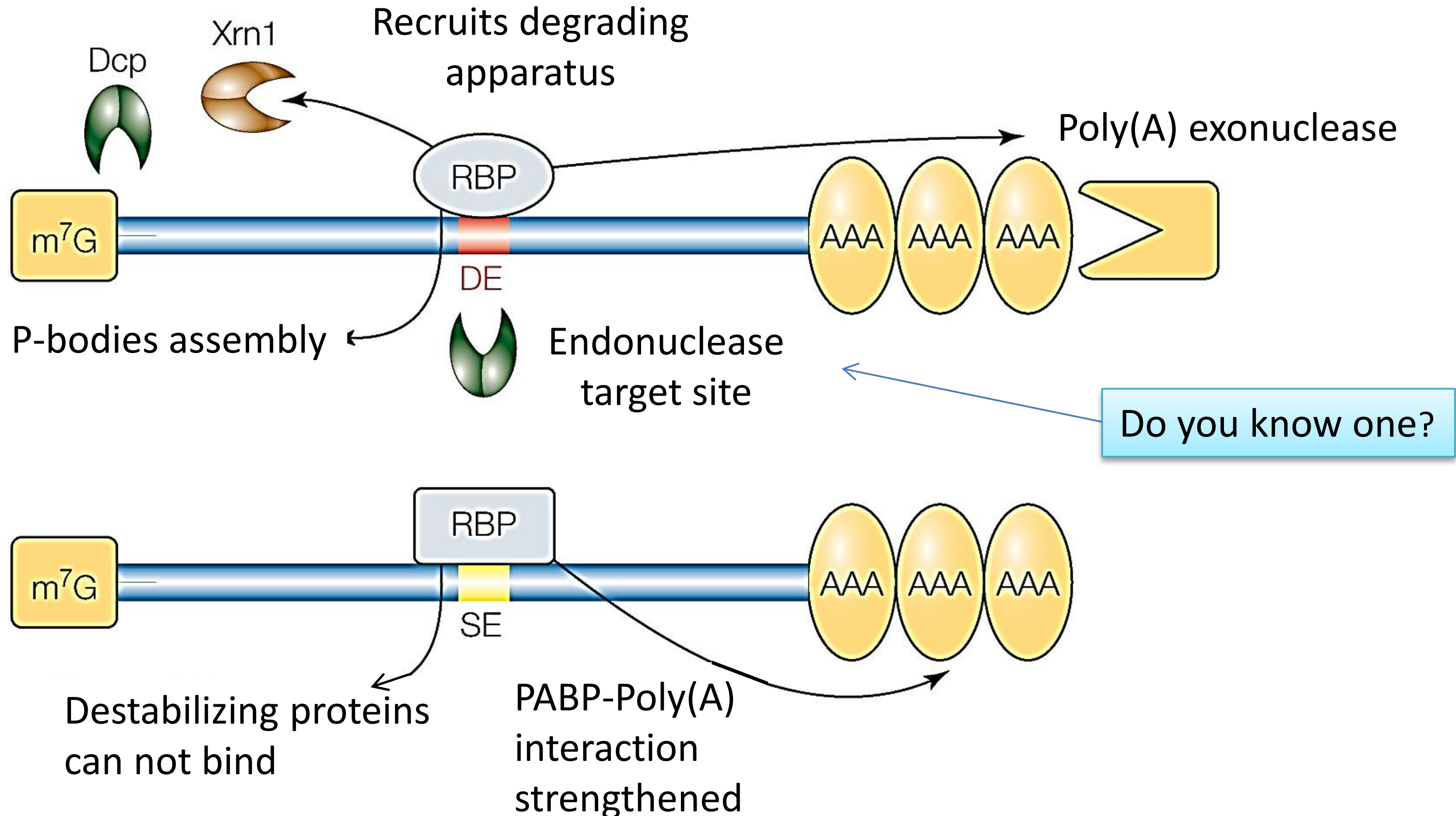
Mouse myoblasts in culture treated with actinomycin D
Samples collected at 0, 10, 50, 110, 230 min
Total RNA → labeled → hybridized to Affymetrix Mouse Gene 1.0 arrays.





Analysis of mRNA decay rate in C2C12 cells.
(A) Examples of mRNA decay curves at extremes. were derived by the nonlinear least squares method for a long and a short half life mRNA.
(B) Distribution of mRNA half and 90th-percentile values .

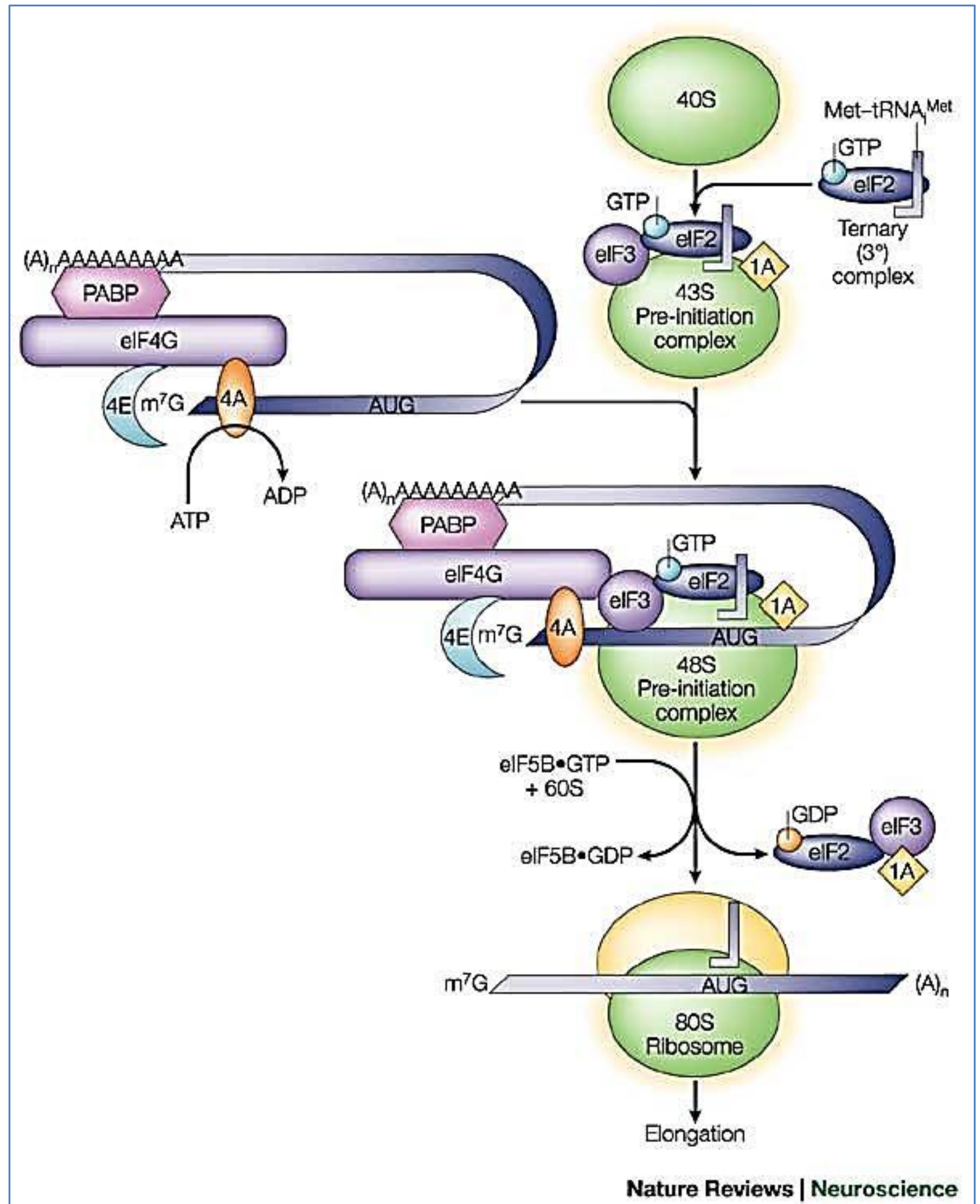
Stabilizing and de-stabilizing sequence motifs (cis-elements)



NOTE: Some cis-elements can function in both pathway, depending which RBP is expressed in specific cell contexts.

decapping or **polyadenylate shortening** strongly impairs translation, since interaction of cap and poly(A) is essential

Degradation and translation are competing events.



RNA-binding proteins and small RNA

Motifs imparting shorter or longer half-life may represent:

1 - protein-binding elements for regulatory RBPs
or
2 - targets for RNA-RNA interaction

In general, in the latter case, we have miRNA, siRNA or piRNA targets.

Let's have a look to most common protein regulators: ARE-binding proteins

Most known regulatory elements: the ARE elements

- are A/U-rich elements found in the **3'-UTR** of some mRNAs encoding cytochines, proto-oncogenes and growth factors
- are defined by their ability to promote rapid deadenylation-dependent mRNA decay
- their sequence requirements are only loosely conserved

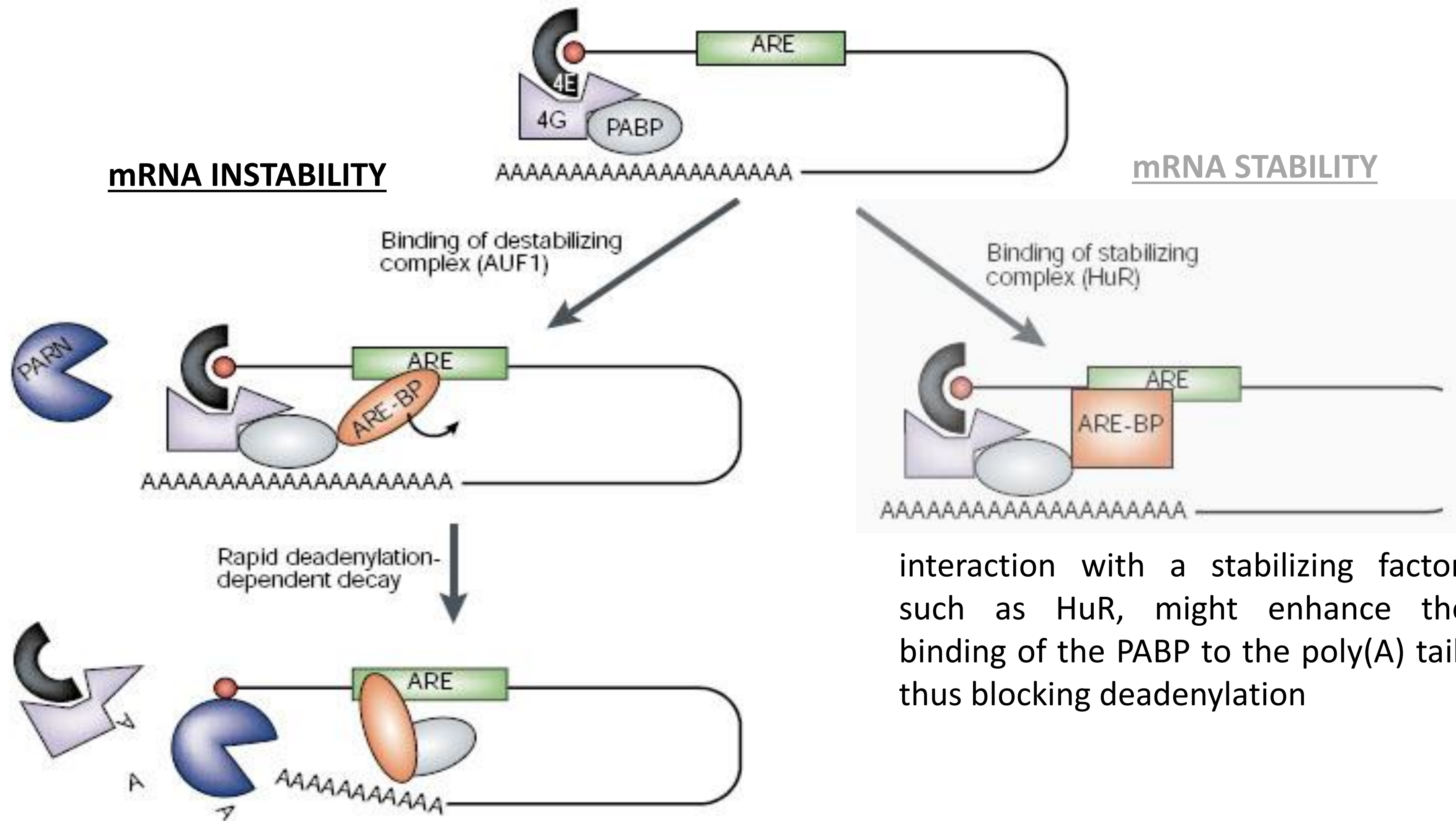
Group	Motif	Examples
I	WAUUUAW and a U-rich region	c-fos, c-myc
IIA	AUUUAUUUAUUUAUUUAUUUA	GM-CSF, TNF-α
IIB	AUUUAUUUAUUUAUUUA	Interferon-α
IIC	WAUUUAUUUAUUUAW	cox-2, IL-2, VEGF
IID	WWAUUUAUUUAWW	FGF2
IIE	WWWWAUUUAWWWW	u-PA receptor
III	U-rich, non-AUUUA	c-jun

Wilusz J.C. et al., 2001

ARE-binding proteins recognize these elements and, in conjunction with other proteins, will guide the mRNA to exosome degradation.

ARE-binding proteins

Many ARE-binding protein have been identified and have either *negative* or in some cases *positive* effect on processes such as stability, translation, subcellular localization of the mRNA



Destabilizing factor AUF1 might promote rapid deadenylation by reducing the affinity of the poly(A) binding protein (PABP) for the poly(A) tail

interaction with a stabilizing factor, such as HuR, might enhance the binding of the PABP to the poly(A) tail, thus blocking deadenylation

Wilusz J.C. et al., 2001

RNA-binding protein	Function	RNA-binding domain	Mode of action	Modifications	Other functions
AUF1 (hnRNP D) and its four splice isoforms (p37, p40, p42, p45)	Usually destabilizing	RRM	Recruit the exosome; remodel mRNA to allow other proteins to bind	Phosphorylation allows isomerization by PIN1 leading to dissociation from RNA; interacts with 14-3-3 proteins	DNA binding
CUG-BP	Destabilizing	RRM	Recruits PARN; modulates ARE function	Phosphorylated by myotonic dystrophy protein kinase	Splicing; translation
ELAV proteins, for example, HuR and HuD	Stabilizing	RRM	Compete with destabilizing proteins for ARE-binding; might relocalize mRNAs away from decay machinery	CARM1-mediated methylation reduces stabilizing function	Translation; RNA localization
KSRP	Destabilizing	KH domain	Recruits decay enzymes: PARN and the exosome	Phosphorylation by p38-MAPK pathway leads to reduced RNA-binding affinity	Splicing
RHAU	Destabilizing	RNA helicase	Recruits decay enzymes: PARN and the exosome	Not known	Not known
TIA-1, TIAR	Translational silencing	RRM	Induce aggregation into stress granules	Phosphorylated by FAST	Alternative splicing
Tristetraprolin (TTP, TIS11, ZFP36), BRF1 (TIS11B, ZFP36L1), BRF2 (TIS11D, ZFP36L2)	Destabilizing	CCCH-type zinc finger	Recruit decay enzymes: CCR4, DCP1, PM-Scl75, RRP4	Phosphorylation by p38-MAPK pathway leads to association with 14-3-3 proteins	Transcription

Entries in brackets indicate alternative protein names. AUF1, AU-rich binding factor-1; CUG-BP, CUG-binding protein; ELAV, embryonic lethal abnormal vision; FAST, Fas-activated serine/threonine kinase; hnRNP, heterogeneous nuclear ribonucleoprotein; KSRP, KH splicing regulatory protein; MAPK, mitogen-activated protein kinase; PARN, poly(A)-specific ribonuclease; RHAU, RNA helicase associated with AU-rich element; RRM, RNA-recognition motif; TIA-1, T-cell-restricted intracellular antigen-1; TIAR, TIA-1-related; TIS, TPA-induced sequence; ZFP, zinc finger protein.

Garneau *et al.* *Nature Reviews Molecular Cell Biology* **8**, 113–126 (2007) | doi:10.1038/nrm2104

nature REVIEWS MOLECULAR CELL BIOLOGY

RNA Binding Proteins are a growing class ...

Generic name: RBP (RNA binding Proteins) GO category: RNA-Binding

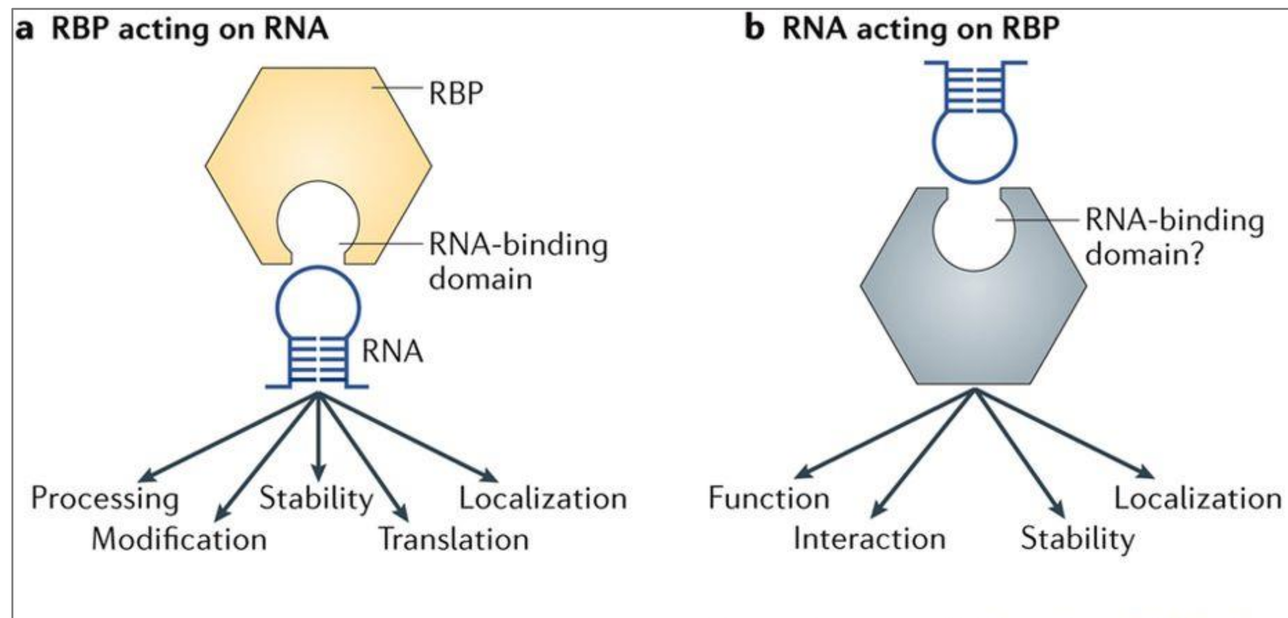
Recent studies used RIC (RNA Interactome Capture) identified an exceptional number of RBPs (860 from HeLa and 791 from HEK293). Many of these do not carry any of the known domains:

-RRM - RNA Recognition Motif

-KH

-DEAD box helicase

-Zn-fingers motifs



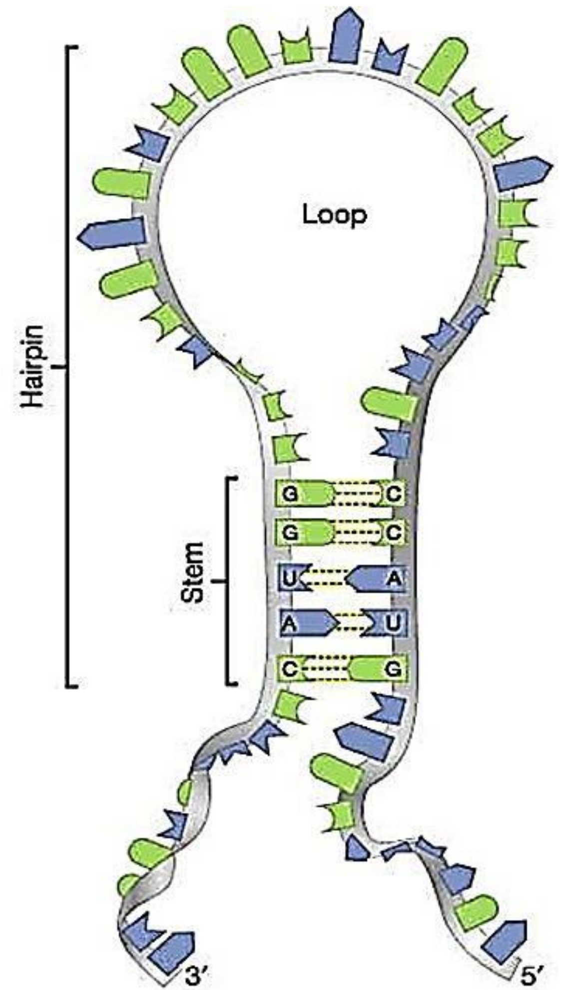
Form Hentze et al. Nature Reviews Molecular Cell Biology **volume 19**, pages 327–341 (2018)

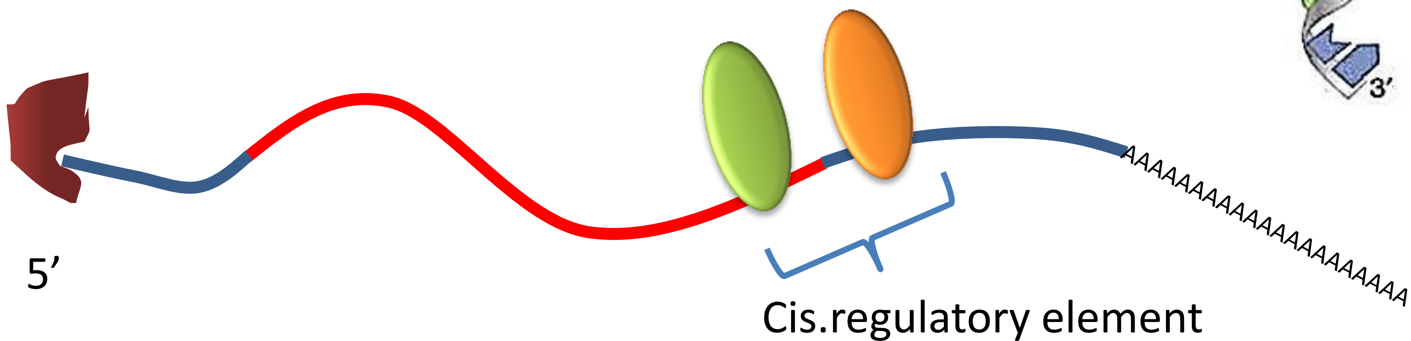
Identification of RNA binding proteins motifs

Specificity of RNA binding: both «sequence» and «structure» elements

Problems in predicting regulatory motifs:

- Localization (intron length)
- Sometimes dispersed elements
- Sometimes the structural component prevails upon pure sequence





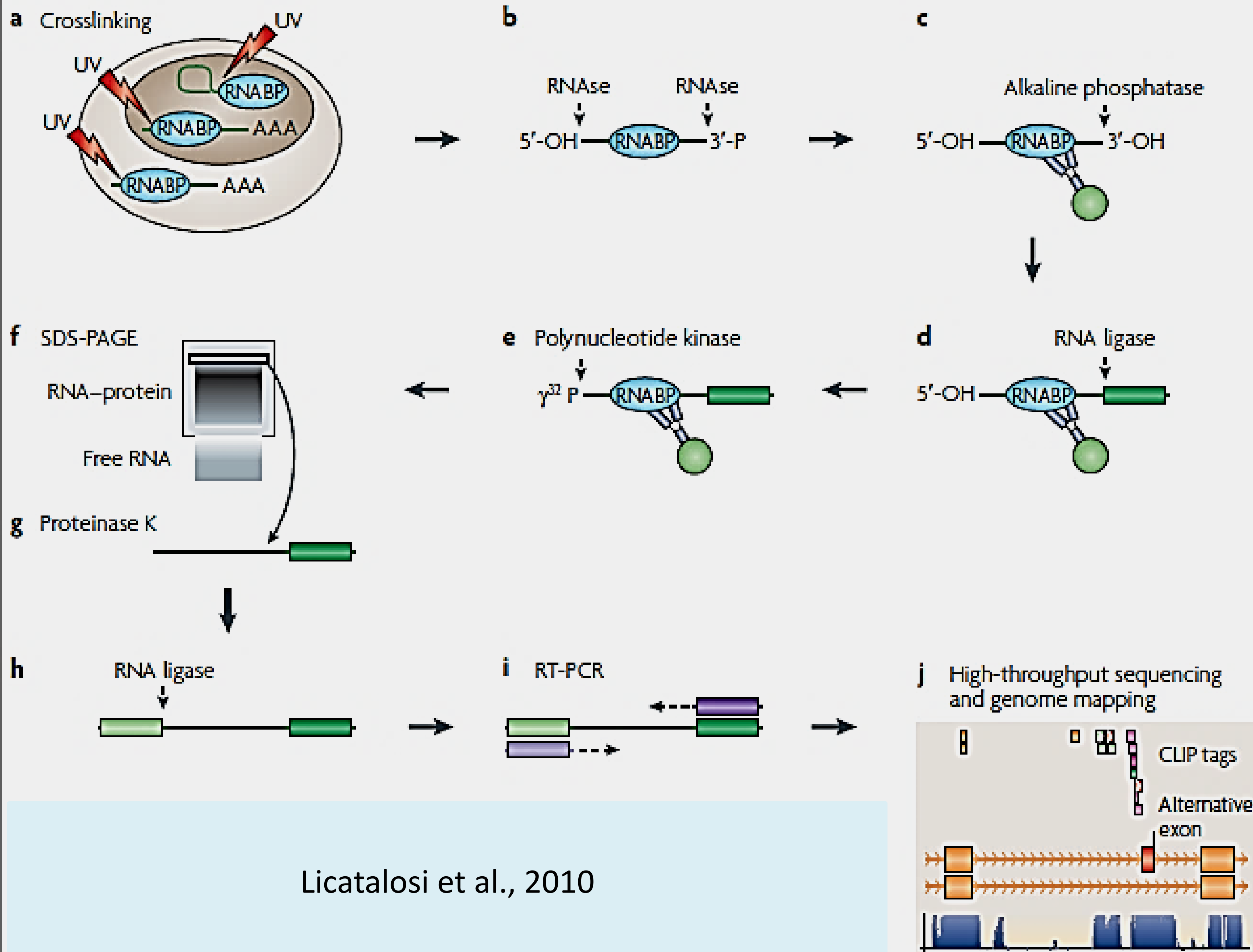
RRM: Sex-lethal http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1b7f

RRM: PTB
http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=2adc

KH: NOVA1 http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1dt4

Once Your Favorite Splicing Regulator is identified, you can look for:

- The RNA transcripts that are bound by YFSR
- the RNA element recognized by YFSR
- the effect of YFSR on the RNA it binds to



Licatalosi et al., 2010

Example:

Individual Splicing Factors HITS-CLIP profiles mapped to genome and compared to RNA-Seq profiles.

CLIP-seq reads mapping

Reference

RNA-seq reads