5.3 Post-transcriptional regulation

QC–Methylation – Stability – RNA interference

RNA processing

- **Capping**
- **Splicing**
- **Poly(A)**
- **RBP binding**
- **Editing ...**

....is RNA (better: RNP) finally ready for translation or function ?

Other post-transcriptional modifications to RNAs

- N^6-methyladenosine (m^6A) (the most studied today – the most frequent)
- N^1-methyladenosine (m^1A)
- 5-methylcytosine (m^5C)
- Pseudo-uridine (most abundant, mostly in rRNA and tRNA, but several mRNAs)
- 2'-O-methylnucleosides (methylation of 2'-OH of ribose)

Mainly in the nucleus, methylation is frequently co-transcriptional

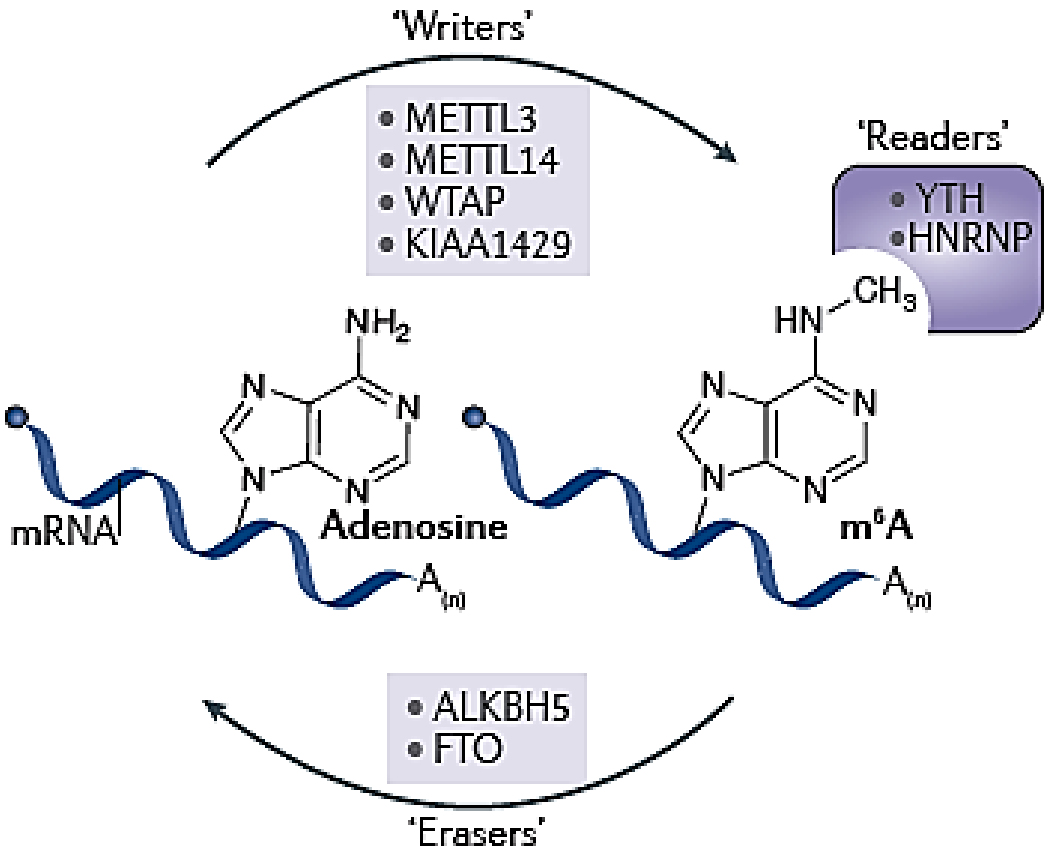


Figure 1 | **The writer, eraser and reader proteins of m^{6}A.**

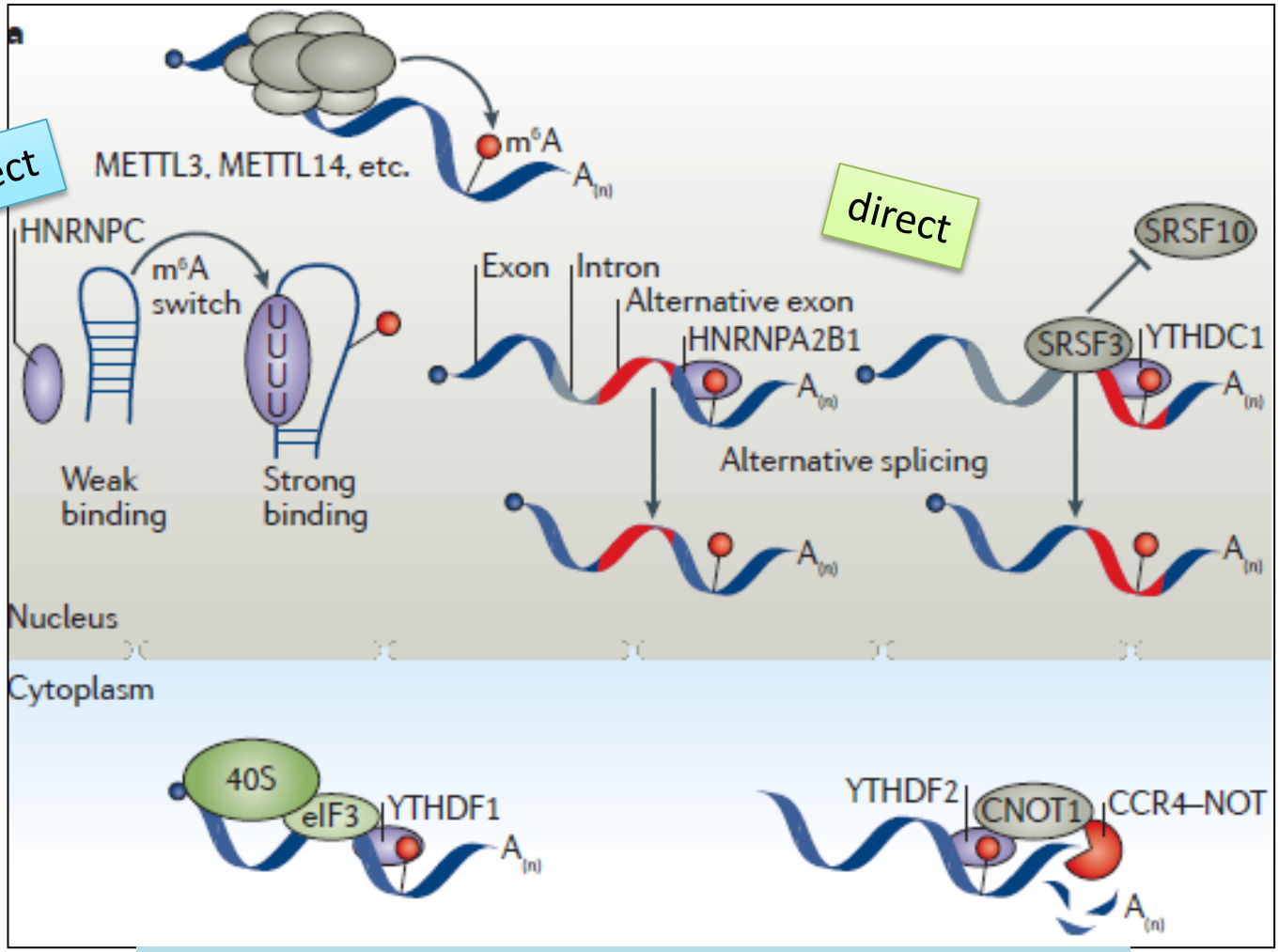
The YT521-B homology (YTH) domain family of proteins (YTHDF1, YTHDF2, YTHDF3 and YTHDC1) are direct readers of m^{6}A and have a conserved m^{6}A-binding pocket.

The heterogeneous nuclear ribonucleoprotein (HNRNP) proteins HNRNPA2B1 and HNRNPC selectively bind m6A-containing mRNAs.

Localized primarily in nuclear speckles

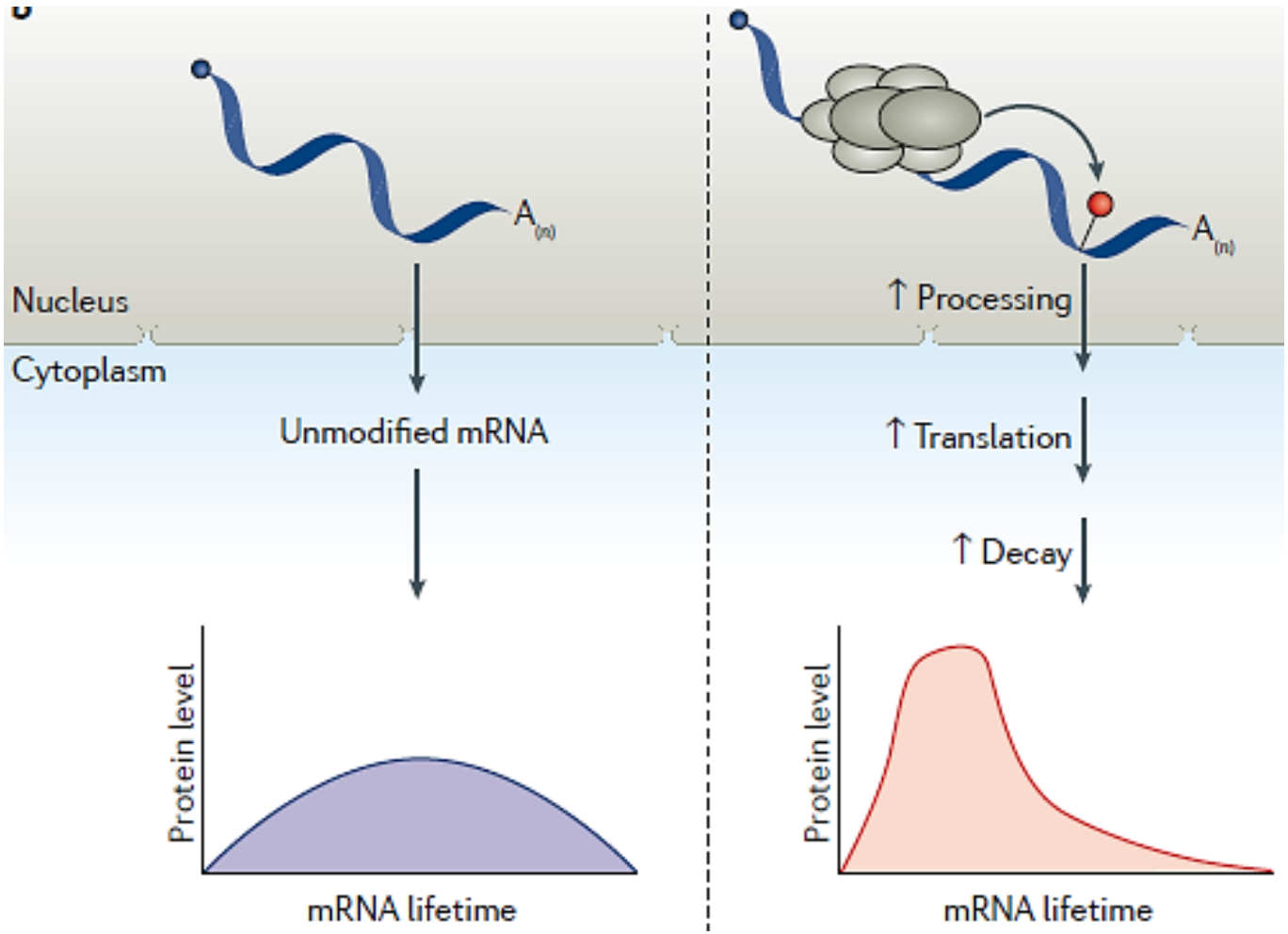
Consensus sequence:
« **RRACH** »

The methyl group at the *N*6 position of m^{6}A does not change Watson–Crick A•U base pairing but weakens duplex RNA by up to 1.4 kcal per mol
Consequence: “indirect” or “direct” effects on protein binding.

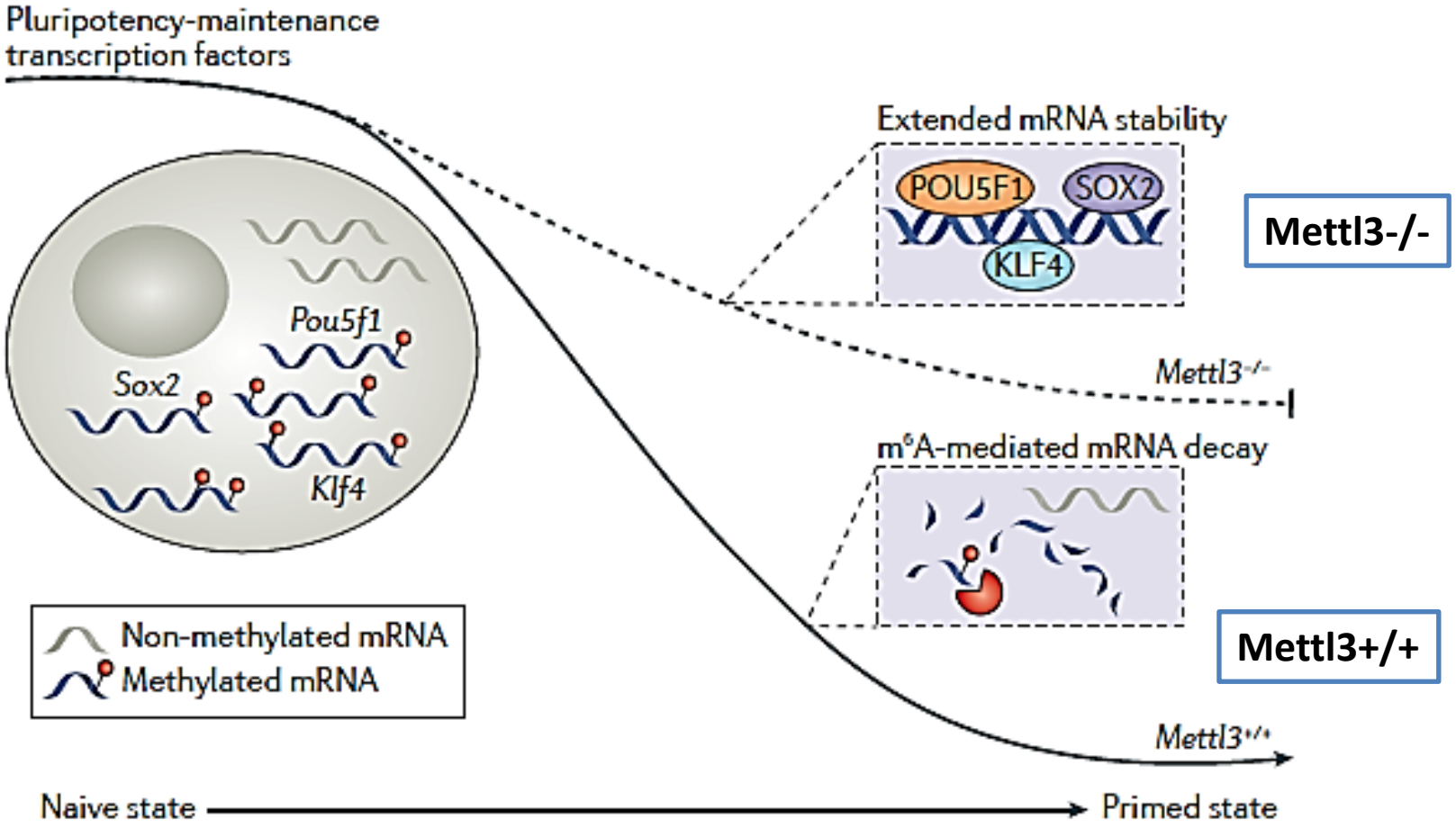


Apparently opposing effects of YTHDF1 and YTHDF2

All steps of RNA metabolism were found to be accelerated by m^6A



A general effect therefore is «sharpening» the window of expression



METTL3 is required for the transition of mouse ES cells **from a naive to the primed state**. During this process, the key pluripotency factor *Pou5f1*, *Klf4* and *Sox2* must be cleared. In mouse ES cells lacking *Mettl3*, this clearance is defective because non-methylated mRNAs are less subjected to decay, which prevents or delays the establishment of a differentiated transcriptome required to achieve a primed mouse ES cell state. *(from Zhao et al., 2017, mod.)*

For those of you who are interested in knowing more:

RNA PROCESSING AND MODIFICATIONS

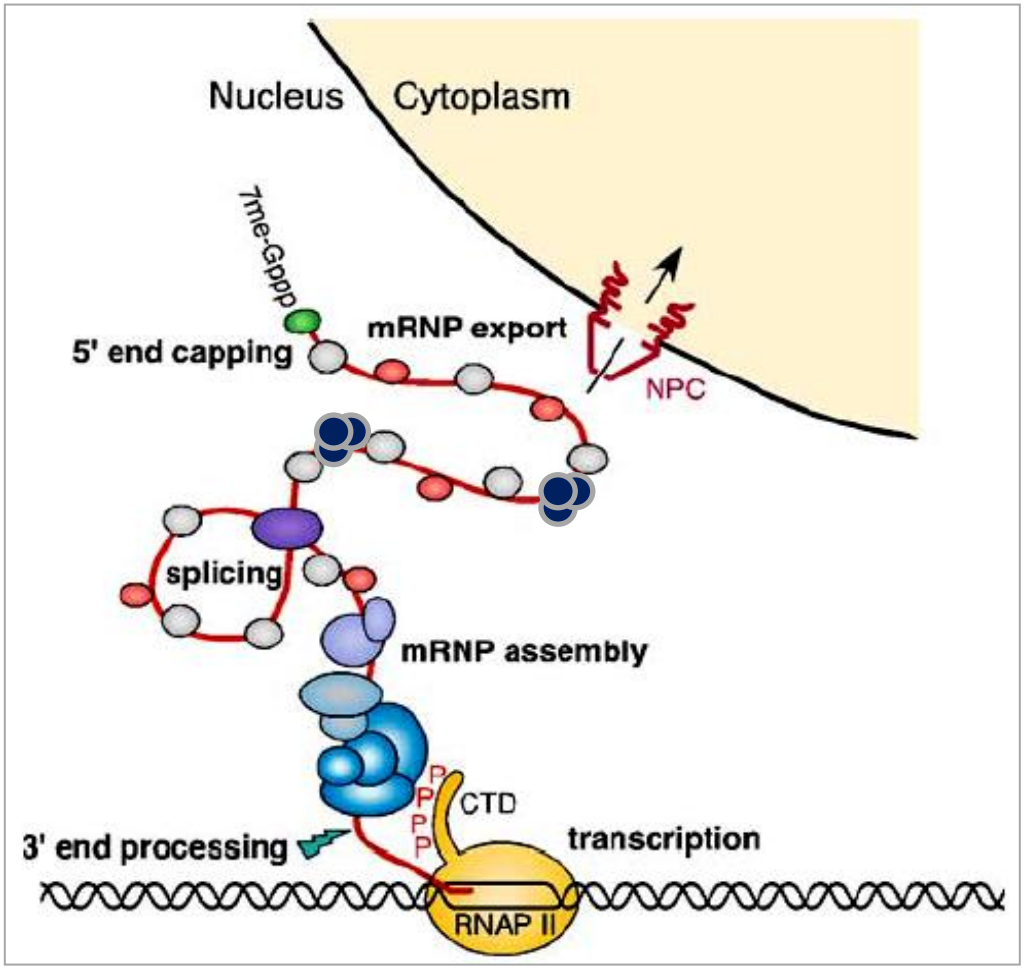
Post-transcriptional gene regulation by mRNA modifications

Boxuan Simen Zhao, Ian A. Roundtree and Chuan He

Abstract | The recent discovery of reversible mRNA methylation has opened a new realm of post-transcriptional gene regulation in eukaryotes. The identification and functional characterization of proteins that specifically recognize RNA N^6-methyladenosine (m^6A) unveiled it as a modification that cells utilize to accelerate mRNA metabolism and translation. N^6-adenosine methylation directs mRNAs to distinct fates by grouping them for differential processing, translation and decay in processes such as cell differentiation, embryonic development and stress responses. Other mRNA modifications, including N^1-methyladenosine (m^1A), 5-methylcytosine (m^5C) and pseudouridine, together with m^6A form the epitranscriptome and collectively code a new layer of information that controls protein synthesis.

REVIEWS

Transport to the cytoplasm



Nucleo-cytoplasmatic transport

- dedicated **exportins**
- accompanying proteins are loaded co-transcriptionally
- some accompanying proteins kept, other exchanged during transport

exon junction complex (EJC)

During the second step of splicing, the EJC is deposited approximately 20-24 nt from the 5' end upstream of the splice junction. The EJC is made up of several key proteins: RNPS1, Y14, SRm160, Aly/REF and Magoh, among others. These proteins have functions in **splicing** and **transport**. Important function in **NMD**.

Schematic view of the nuclear side of eukaryotic gene expression, from transcription to nuclear export. NPC Nuclear pore complex, CTD C-terminal domain of Rpb1, RNAPII RNA polymerase II.

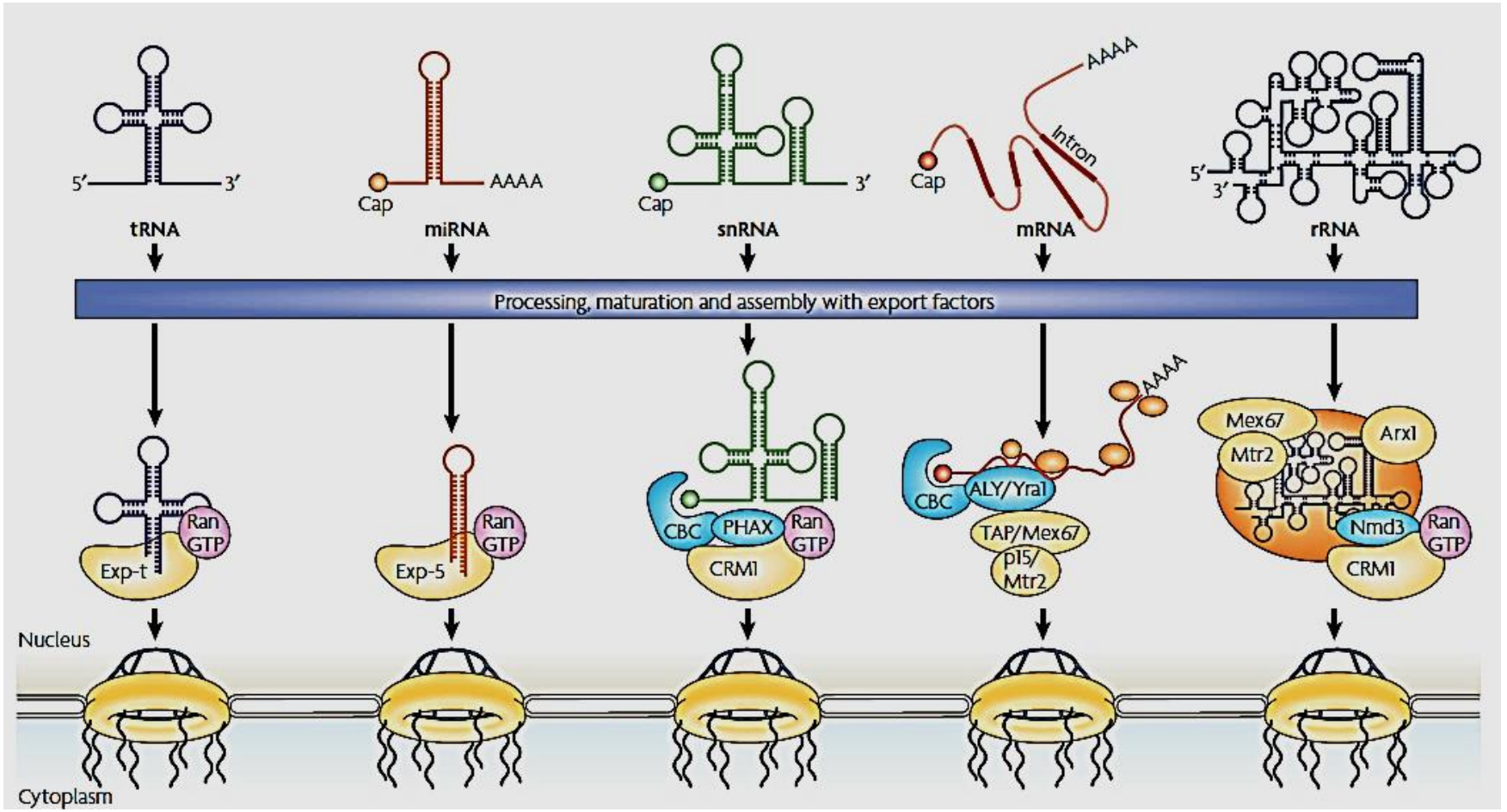
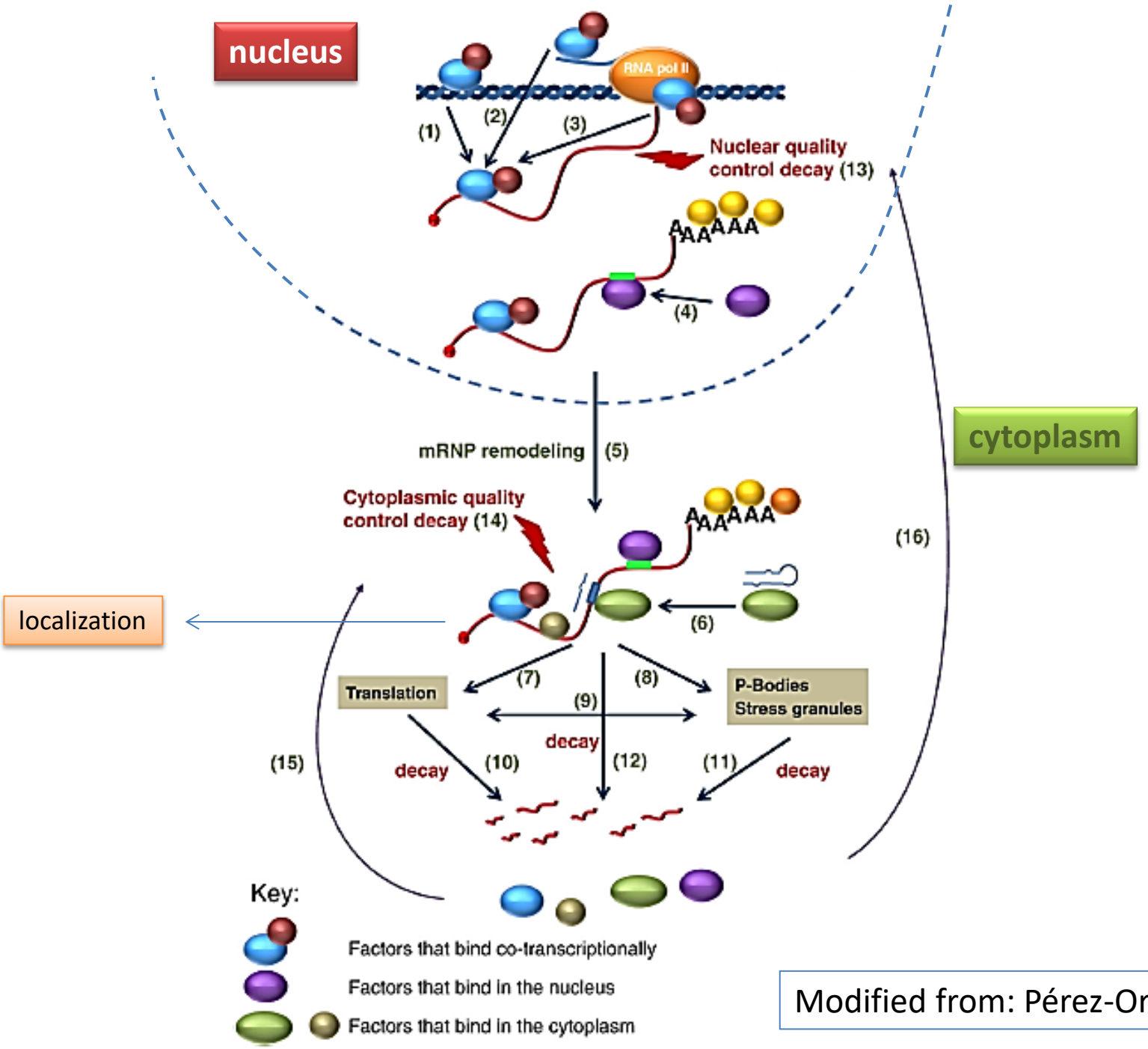
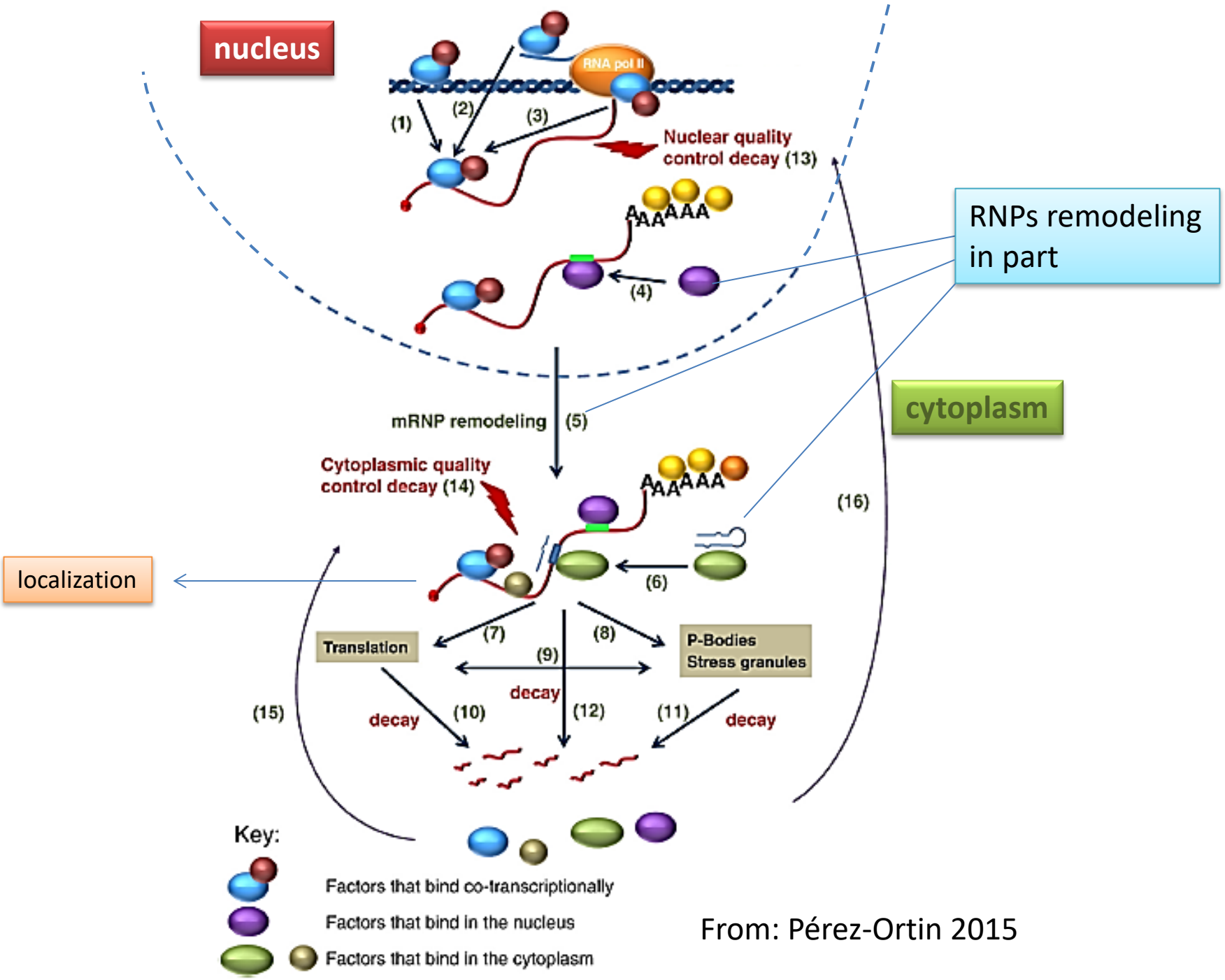


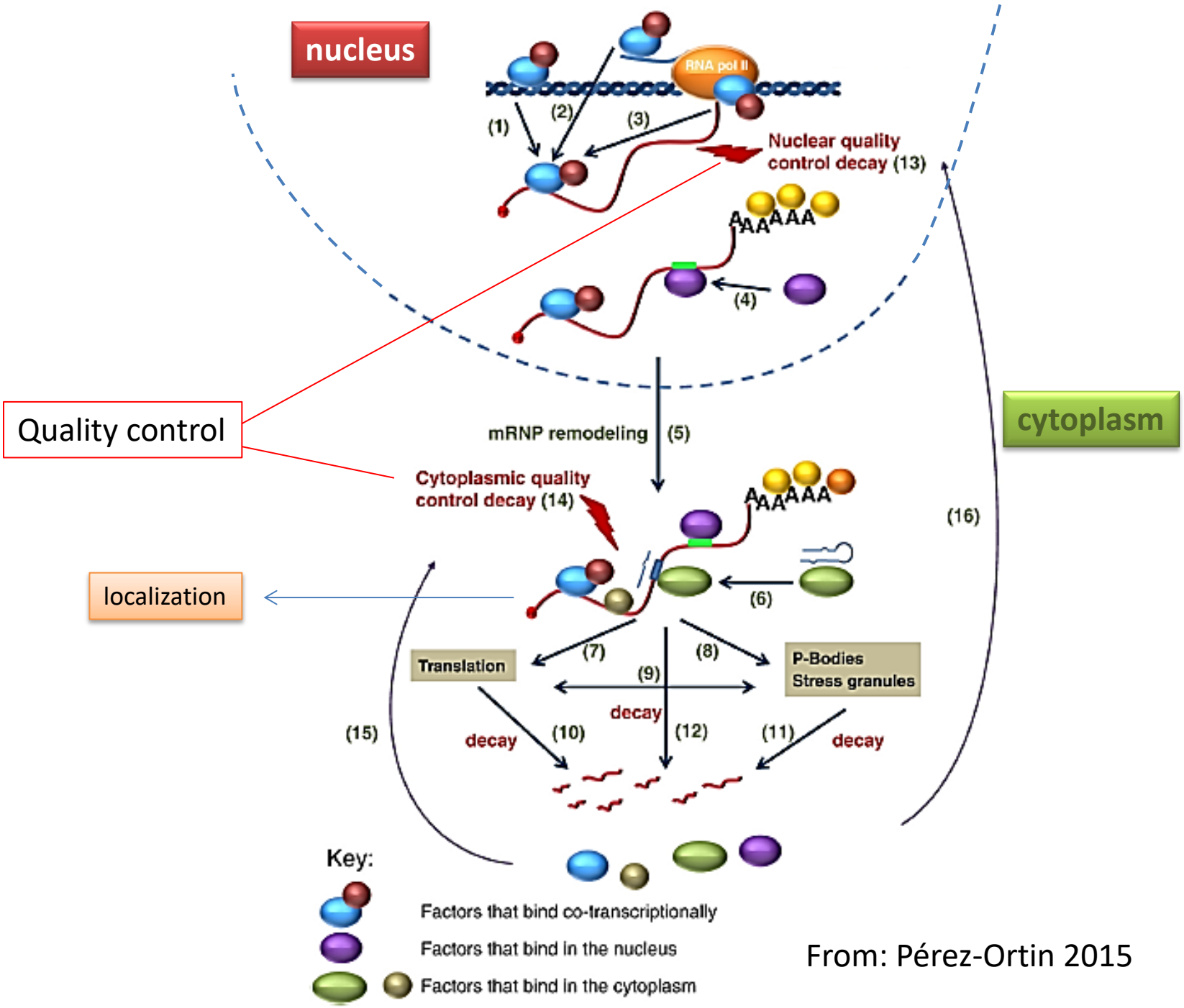
Figure 1 | **The different RNA export pathways.** The major RNA export routes are shown (tRNA, microRNA (miRNA), small nuclear (sn)RNA, mRNA, ribosomal (r)RNA). In each case, the primary RNA transcript is shown, as well as the transport-competent RNA after it has undergone processing, maturation and assembly with export factors (**export adaptors** are shown in blue, **export receptors** are shown in yellow). Prominent **structural motifs** in pre-RNAs are indicated (single/double-stranded RNA, loops, exons and introns, 5′ cap and 3′ poly(A) tail). For the mRNA export route, the names of both metazoan and yeast proteins are indicated, and mRNAs are shown with additional adaptor proteins and RNA binding general exporter in eukaryotes, CRM1, and two auxiliary exporters, Mex67–Mtr2 and Arx1, that have only been studied in yeast are depicted. CBC, cap-binding complex; Exp, exportin. (*From Kohler & Hurt, Nat Rev Mol Cell Biol, 8:761, 2007*).



Modified from: Pérez-Ortin 2015

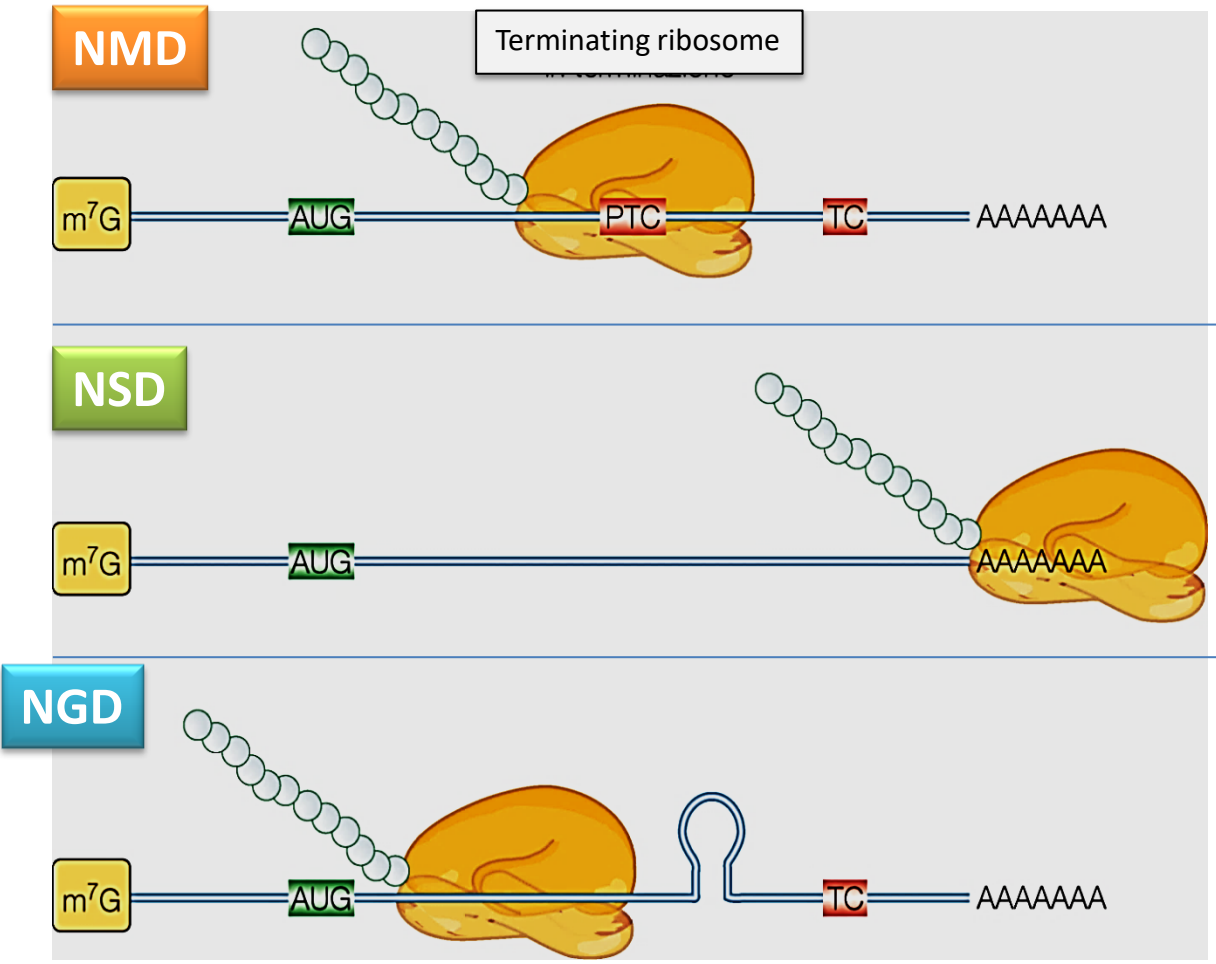


From: Pérez-Ortin 2015



From: Pérez-Ortin 2015

A regulatory role for RNA Quality Control pathways ?



1. Nonsense Mediated Decay
Premature termination codon

2. Non Stop Decay
Absence of termination codon

3. Non-Go Decay
Pausing by structure before TC

1. Nonsense-mediated decay

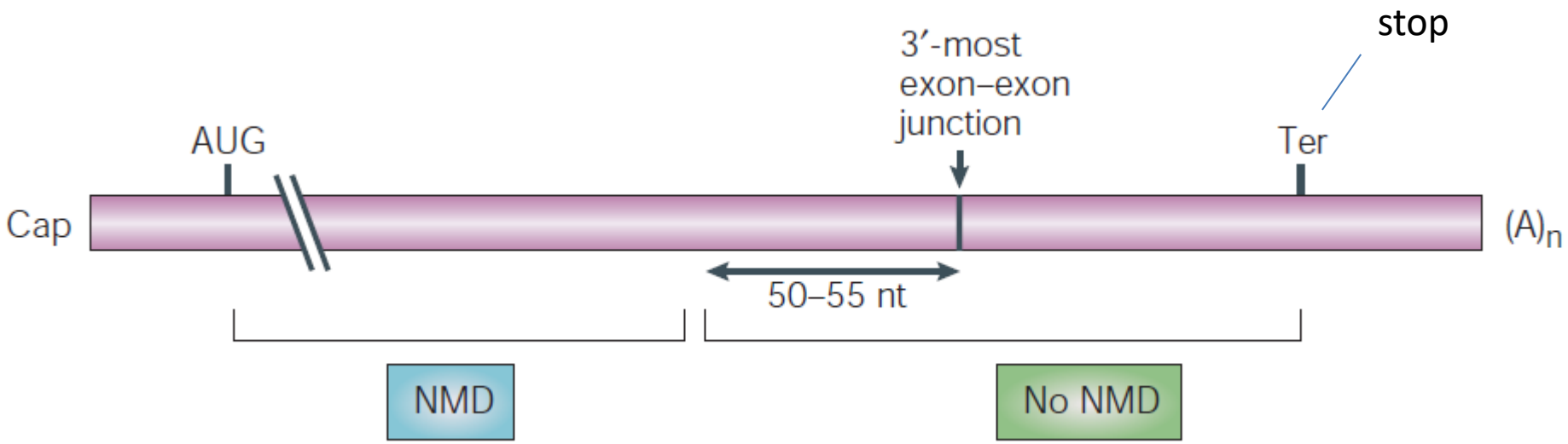
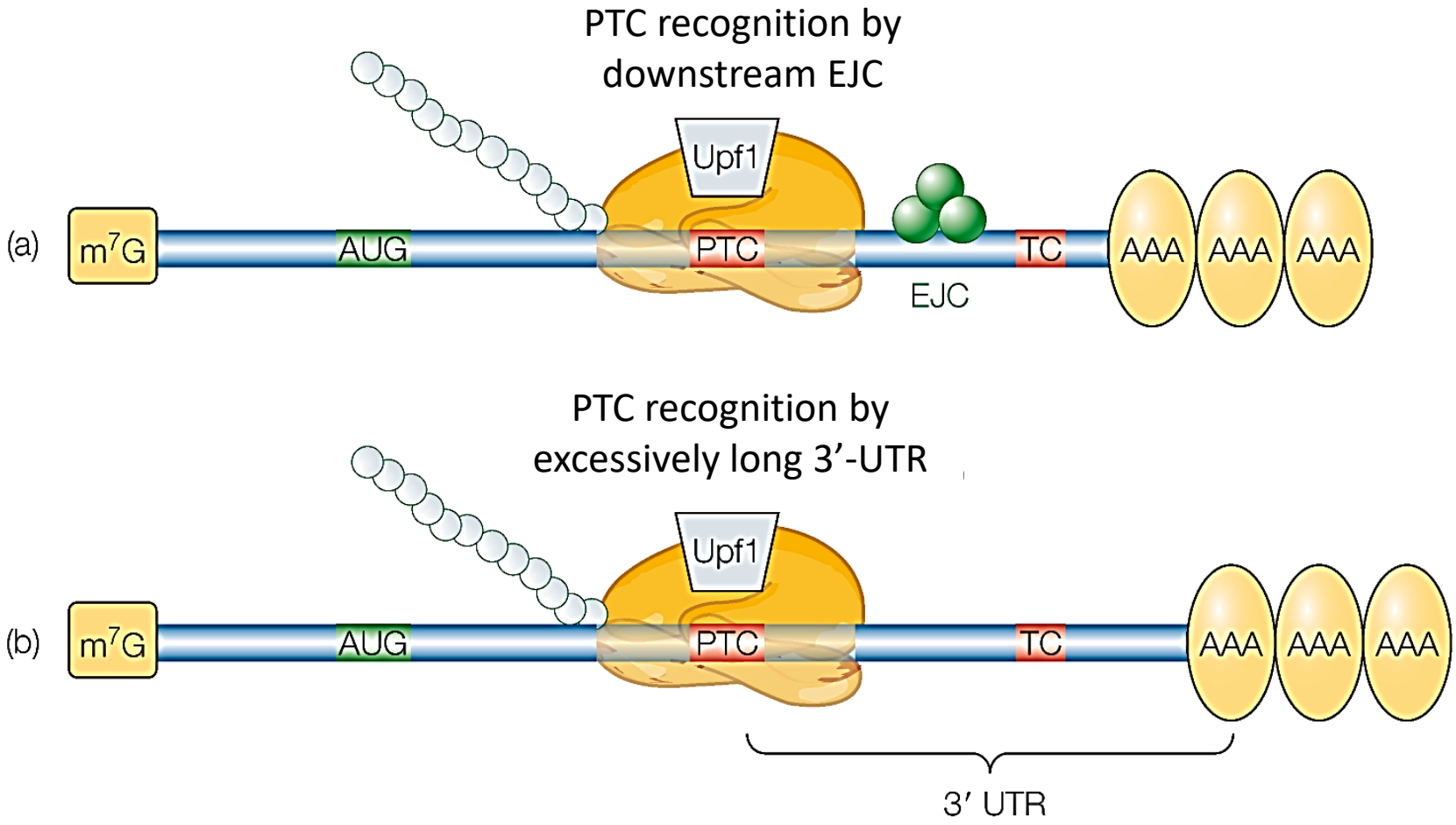


Figure 1 | **NMD and the 'position-of-an-exon–exon-junction' rule.** Only the 3′-most exon–exon junction within a generic mammalian mRNA is shown. Mammalian mRNAs generally have an average of 7–8 splicing-generated exon–exon junctions. A premature termination codon (PTC) that is located in the region indicated in blue, which is followed by an exon–exon junction more than 50–55 nucleotides (nt) downstream, elicits nonsense-mediated mRNA decay (NMD), whereas a PTC that is located in the region indicated in green fails to elicit NMD. The normal termination codon (Ter) usually resides within the 3′-most exon.

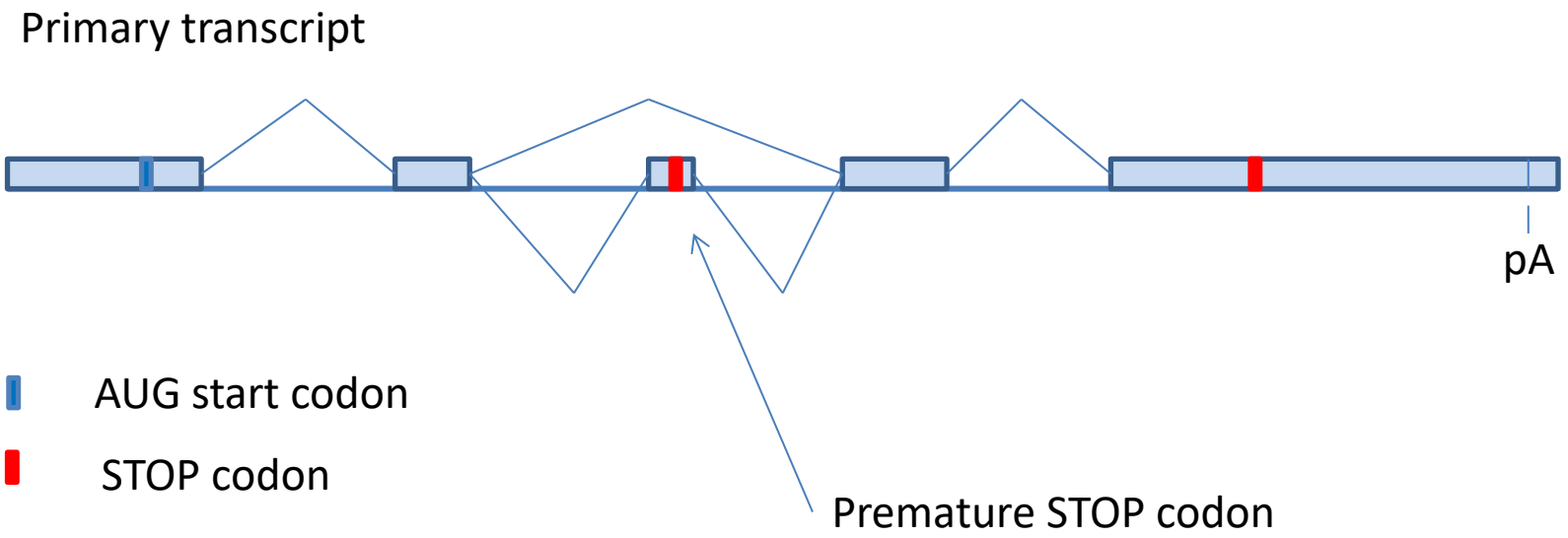
From: Maquat 2004, Nat Rev Mol Cell Biol 5:89-99.

NMD


PTC recognition by
downstream EJC
(a)
m7G
AUG
Upf1
PTC
EJC
TC
AAA
AAA
AAA
PTC recognition by
excessively long 3’-UTR
(b)
m7G
AUG
Upf1
PTC
TC
AAA
AAA
AAA
3′ UTR

What is the function of NMD ? Protection ? Regulatory ?

In Mammals, a number of Alternative Splicing events introduce a premature termination codon into the processed mRNA, not followed immediately by pA.



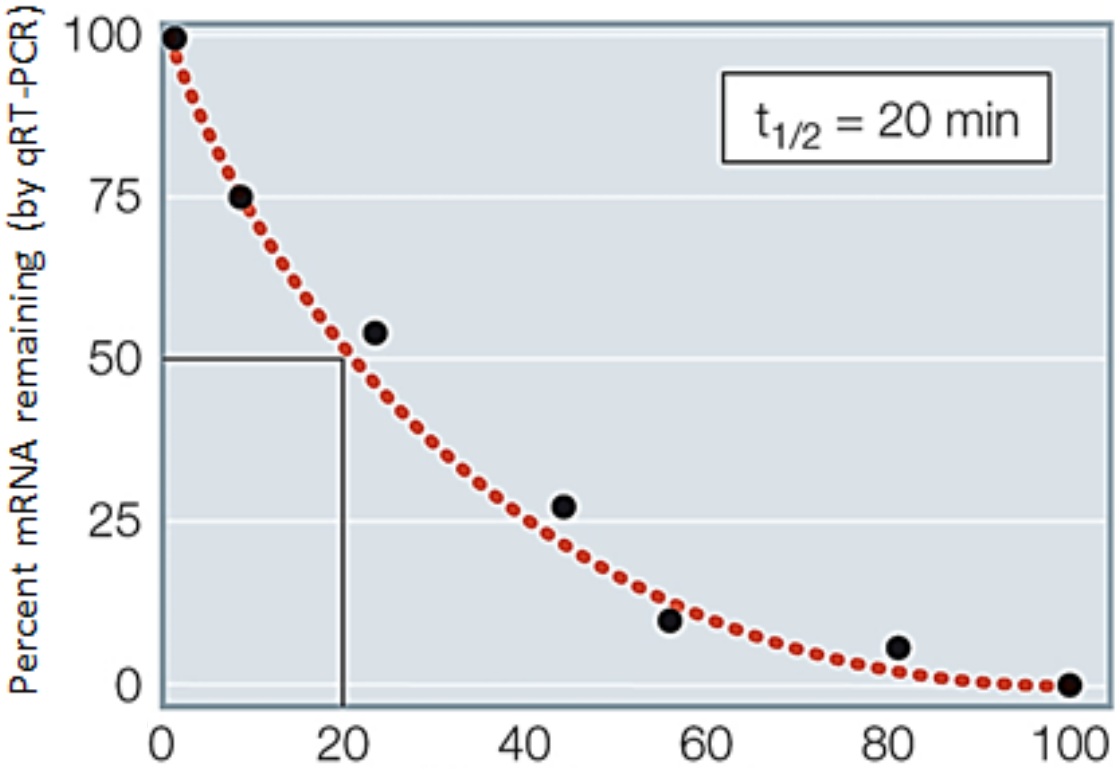
This induces NMD. It should be considered a silencing AS event.

RNA stability

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

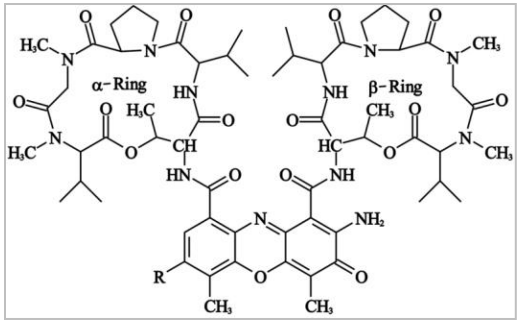
How can we measure RNA half-life ?

Measuring RNA half-life



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

Cells treated with Actinomycin D

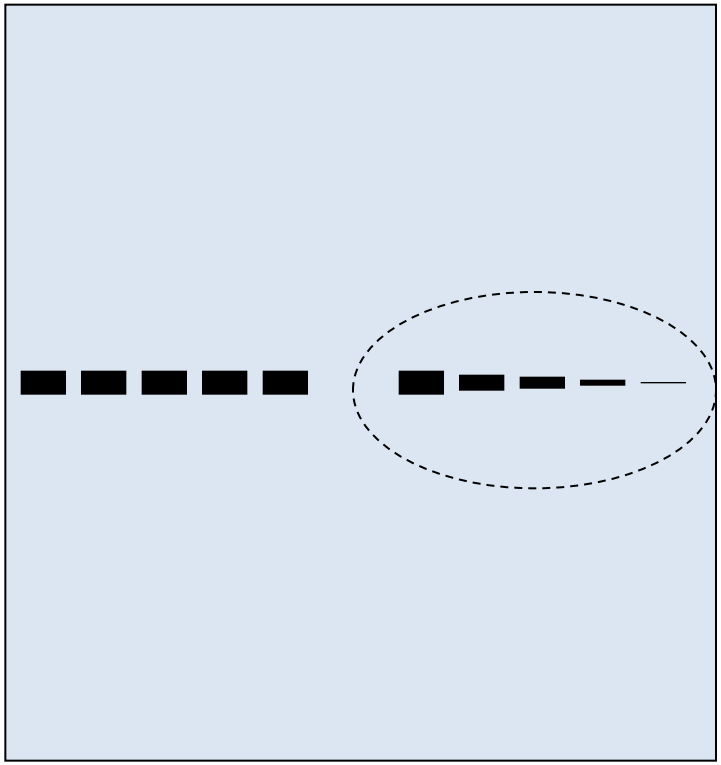


RNA extracted at time intervals

Add radioactive protection probe

RNase (ss)

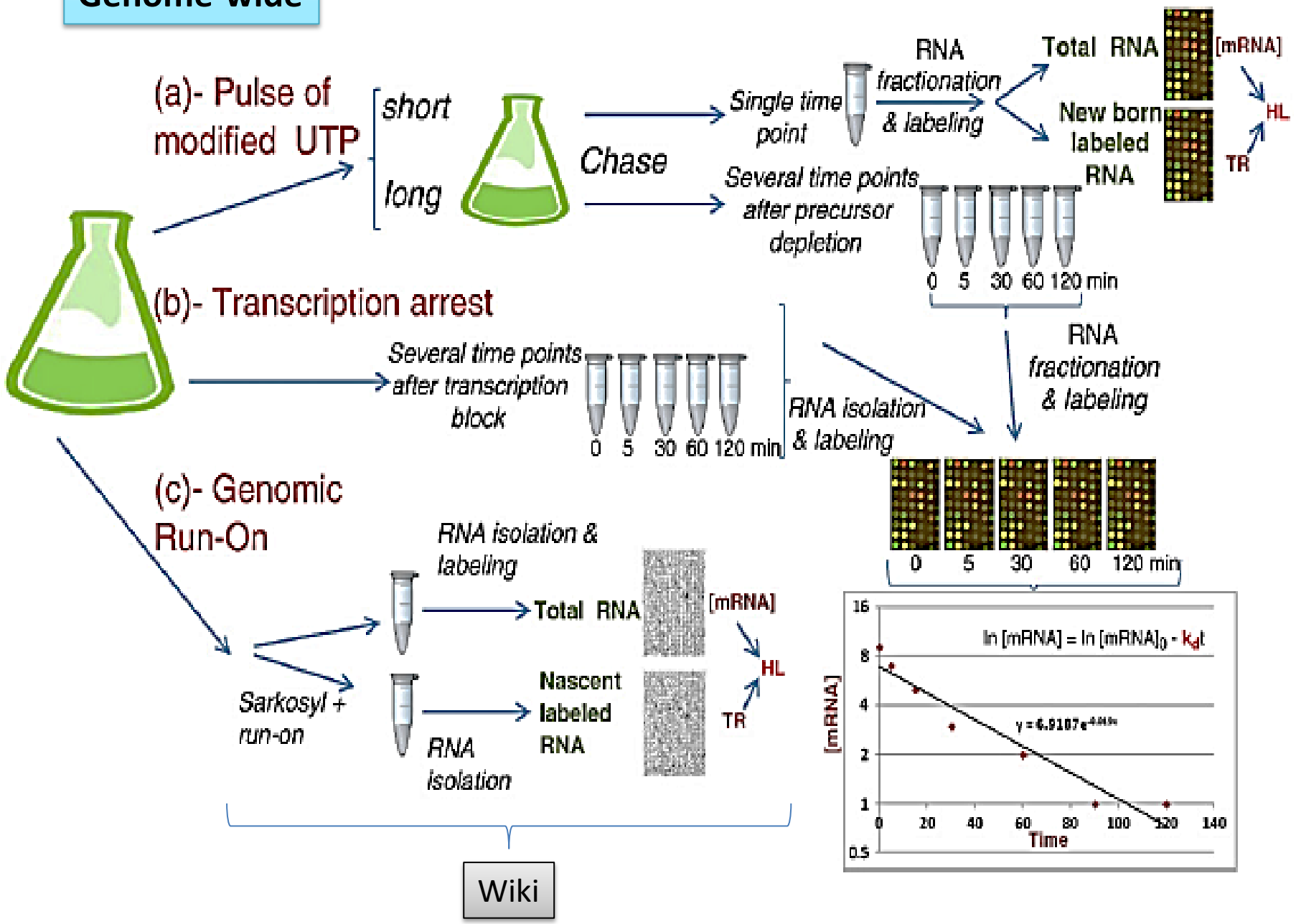
Older measure of mRNA half-life on single genes



RNase Protection Assay (RPA)

wiki

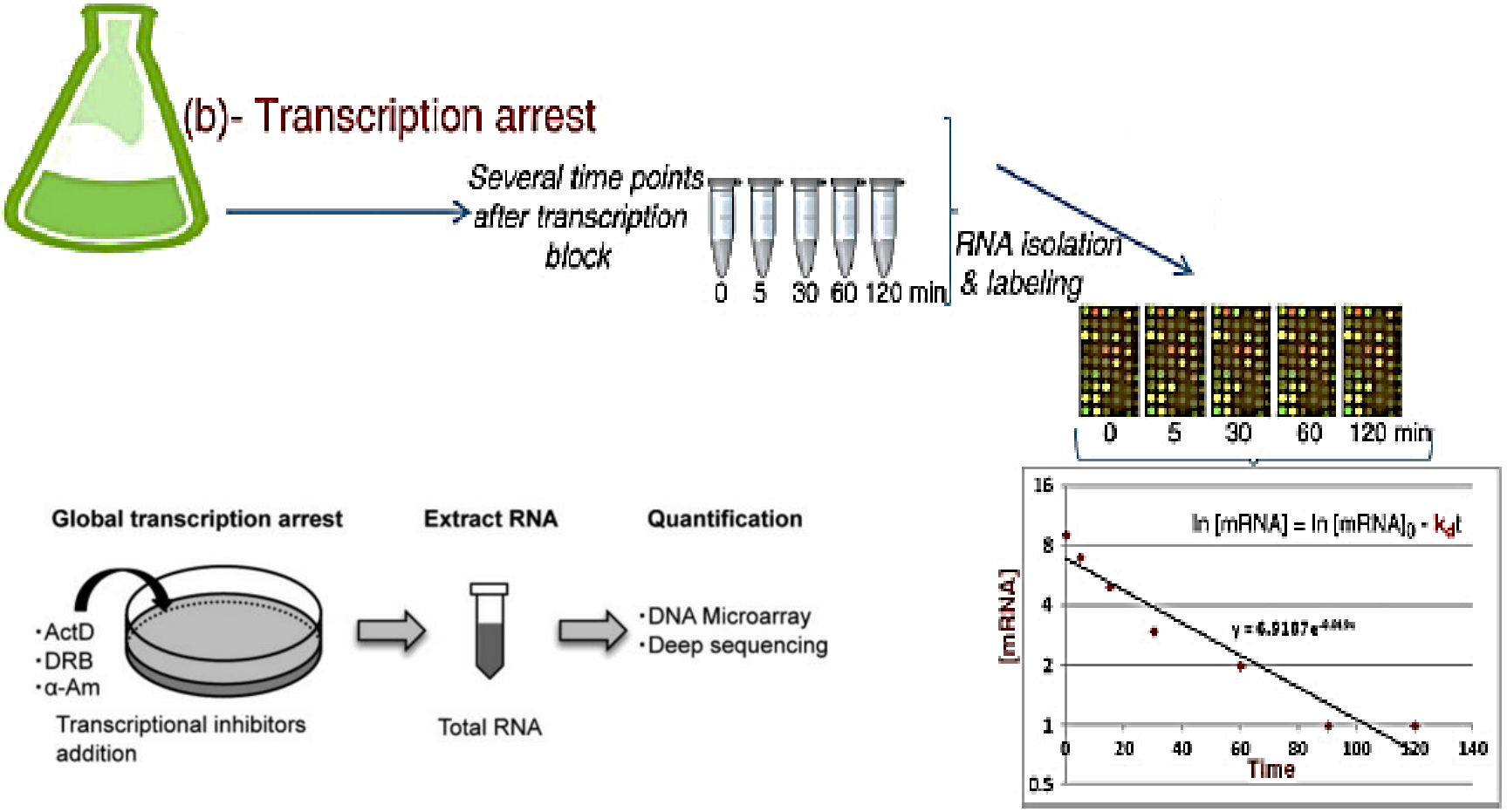
Genome-wide



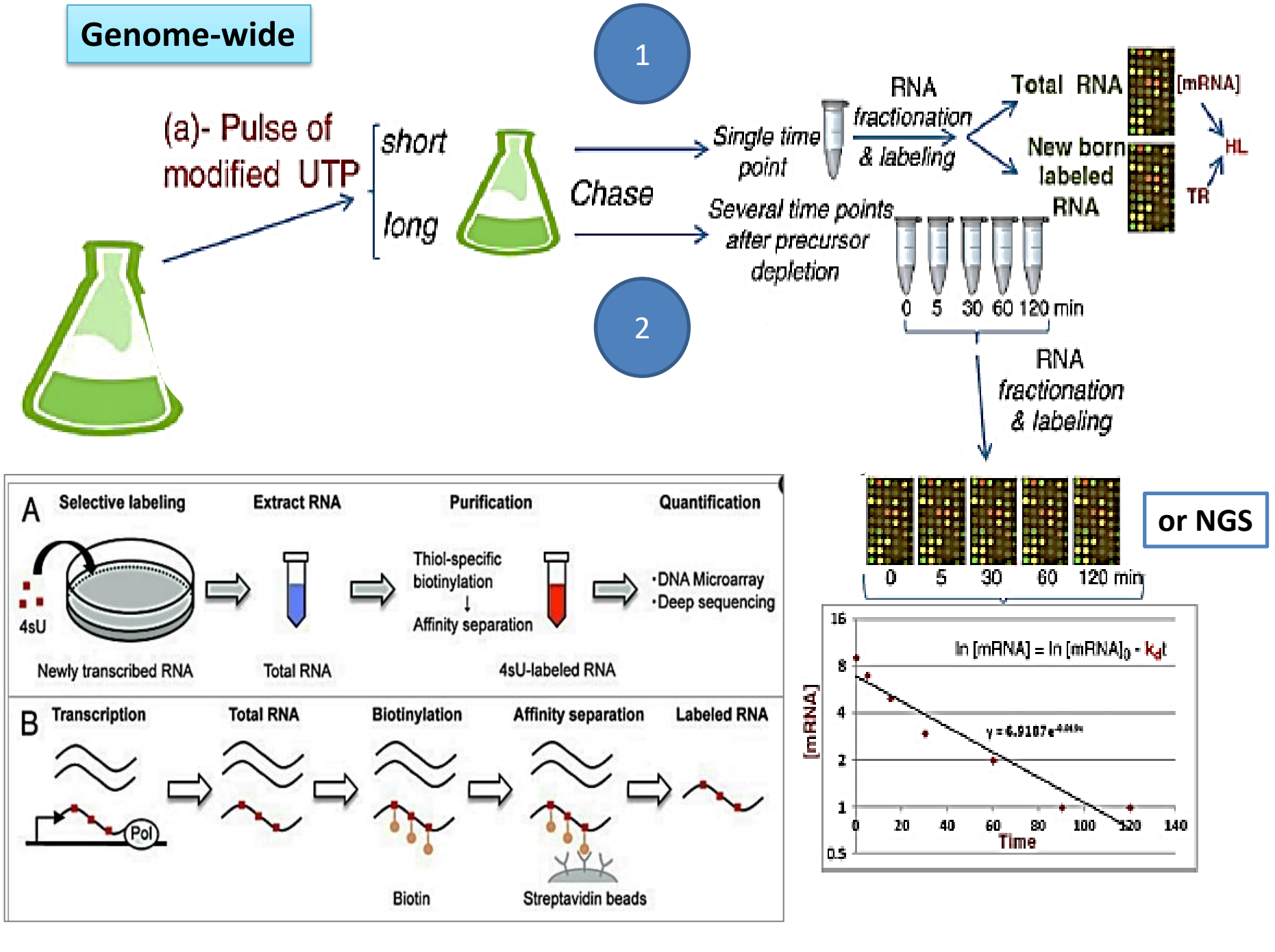
Wiki

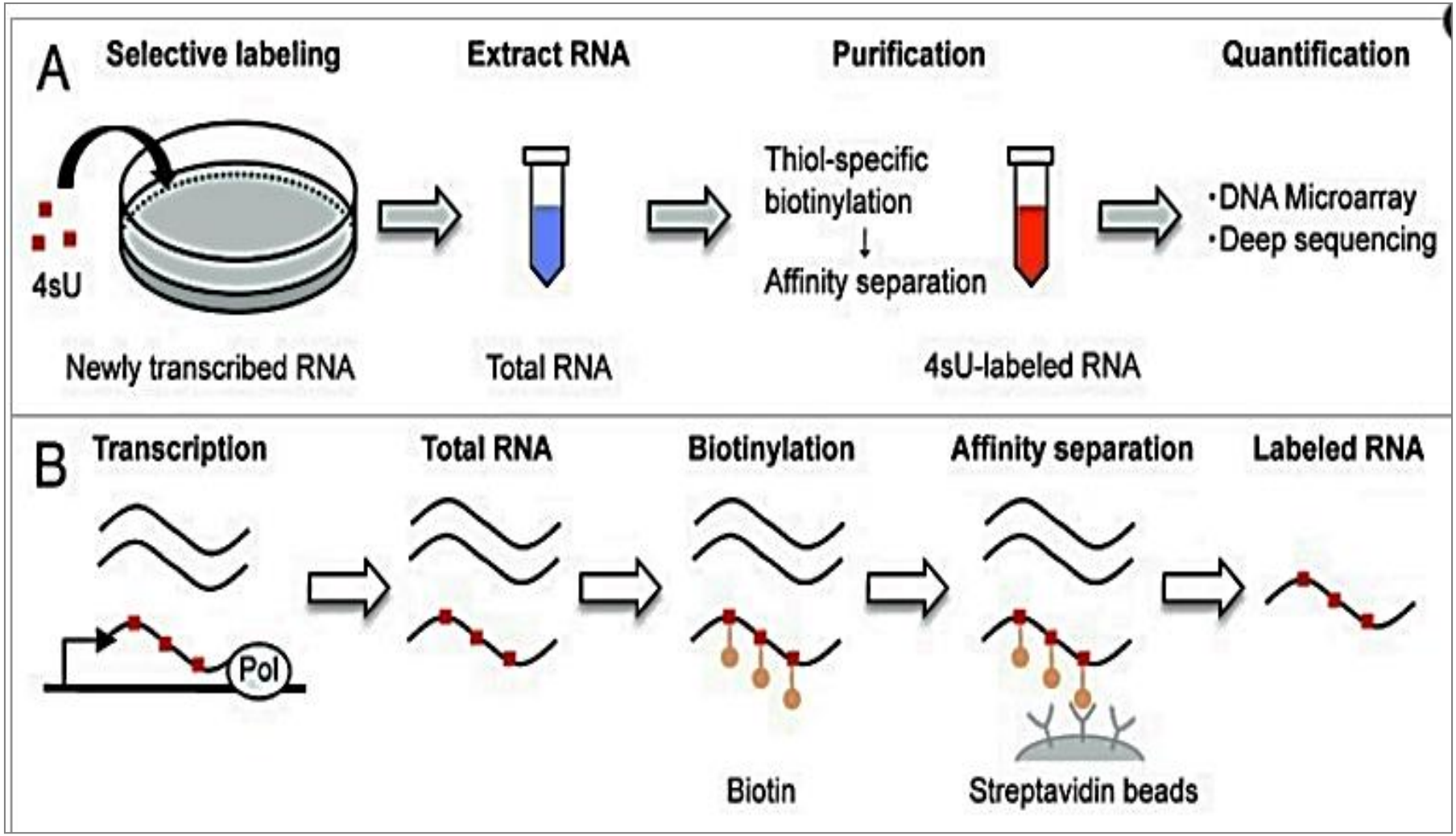
Genome-wide

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



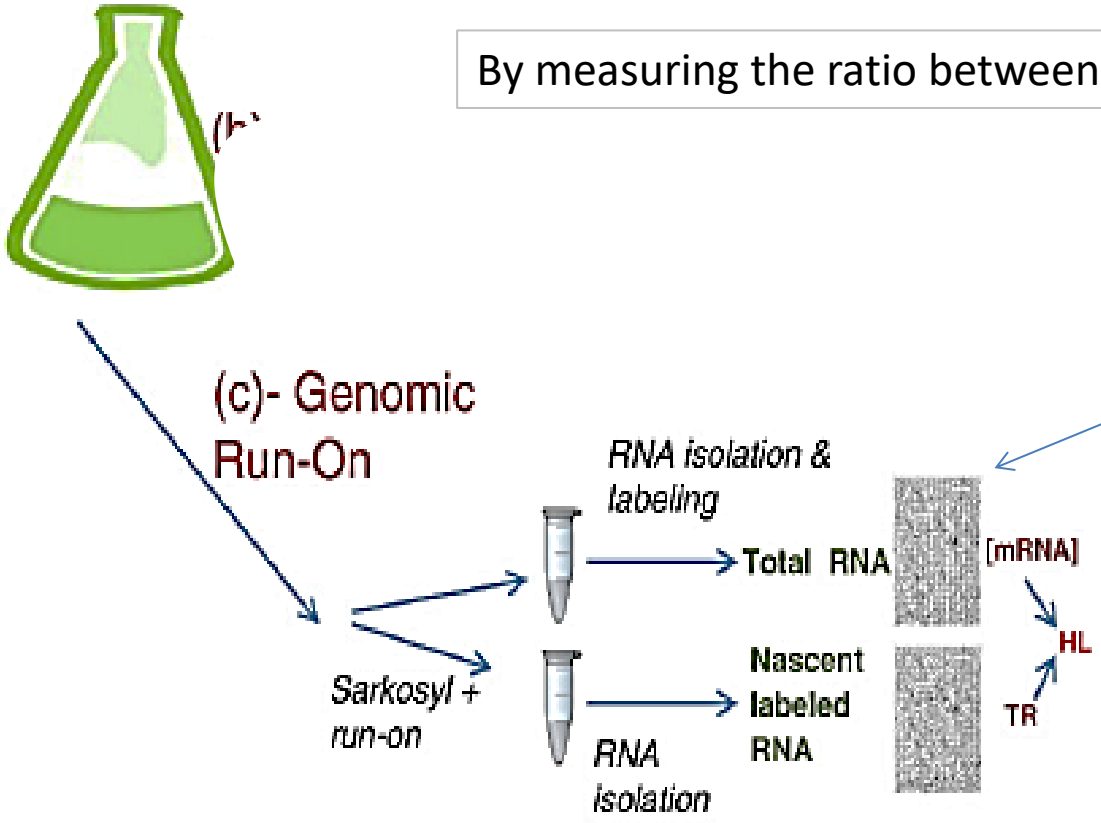
Genome-wide


1
(a)- Pulse of modified UTP
short
long
Chase
Single time point
RNA fractionation & labeling
Total RNA
New born labeled RNA
[mRNA]
HL
TR
Several time points after precursor depletion
0 5 30 60 120 min
2
RNA fractionation & labeling
A
Selective labeling
Extract RNA
Purification
Quantification
4sU
Thiol-specific biotinylation
Affinity separation
•DNA Microarray
•Deep sequencing
Newly transcribed RNA
Total RNA
4sU-labeled RNA
B
Transcription
Total RNA
Biotinylation
Affinity separation
Labeled RNA
Pol
Biotin
Streptavidin beads
0 5 30 60 120 min
or NGS
ln [mRNA] = ln [mRNA]0 - kdt
y = 6.9187e-0.445x
[mRNA]
Time


A
Selective labeling
Extract RNA
Purification
Quantification
4sU
Thiol-specific
biotinylation
↓
Affinity separation
•DNA Microarray
•Deep sequencing
Newly transcribed RNA
Total RNA
4sU-labeled RNA
B
Transcription
Total RNA
Biotinylation
Affinity separation
Labeled RNA
Pol
Biotin
Streptavidin beads

Genome-wide

By measuring the ratio between syntesis and steady-state level



or NGS

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

Letter

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

Although mRNA decay rates are a key determinant of the steady-state concentration for any given mRNA species, relatively little is known, on a population level, about what factors influence turnover rates and how these rates are integrated into cellular decisions. We decided to measure mRNA decay rates in two human cell lines with high-density oligonucleotide arrays that enable the measurement of decay rates simultaneously for thousands of mRNA species. Using existing annotation and the Gene Ontology hierarchy of biological processes, we assign mRNAs to functional classes at various levels of resolution and compare the decay rate statistics between these classes. The results show statistically significant organizational principles in the variation of decay rates among functional classes. In particular, transcription factor mRNAs have increased average decay rates compared with other transcripts and are enriched in "fast-decaying" mRNAs with half-lives <2 h. In contrast, we find that mRNAs for biosynthetic proteins have decreased average decay rates and are deficient in fast-decaying mRNAs. Our analysis of data from a previously published study of *Saccharomyces cerevisiae* mRNA decay shows the same functional organization of decay rates, implying that it is a general organizational scheme for eukaryotes. Additionally, we investigated the dependence of decay rates on sequence composition, that is, the presence or absence of short mRNA motifs in various regions of the mRNA transcript. Our analysis recovers the positive correlation of mRNA decay with known AU-rich mRNA motifs, but we also uncover further short mRNA motifs that show statistically significant correlation with decay. However, we also note that none of these motifs are strong predictors of mRNA decay rate, indicating that the regulation of mRNA decay is more complex and may involve the cooperative binding of several RNA-binding proteins at different sites.

[Supplemental material is available online at www.genome.org, and also at http://genomes.rockefeller.edu/~yange.]

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

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 (blocks RNA biosynthesis)

Samples collected at 0, 10, 50, 110, 230 min
Total RNA → labeled → hybridized to Affymetrix Mouse Gene 1.0 arrays.

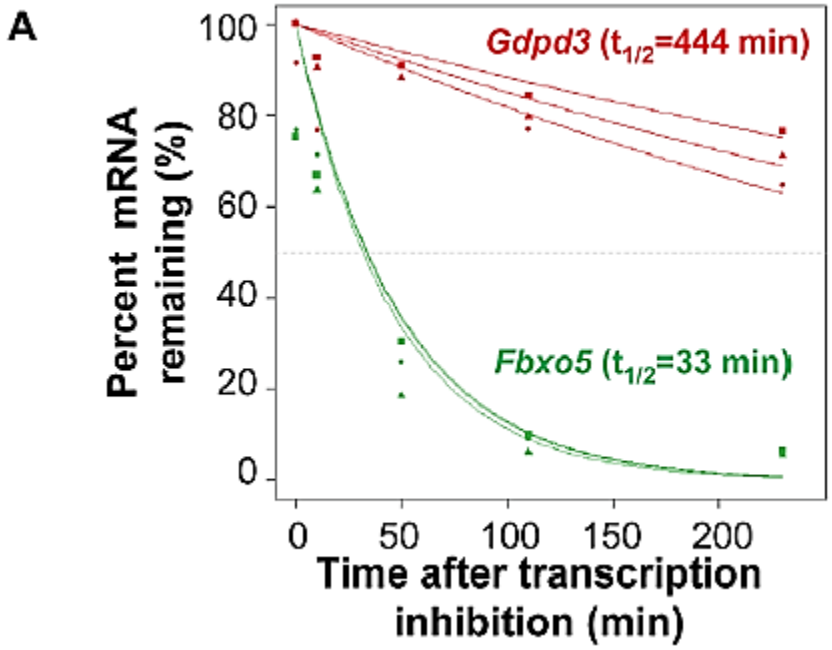
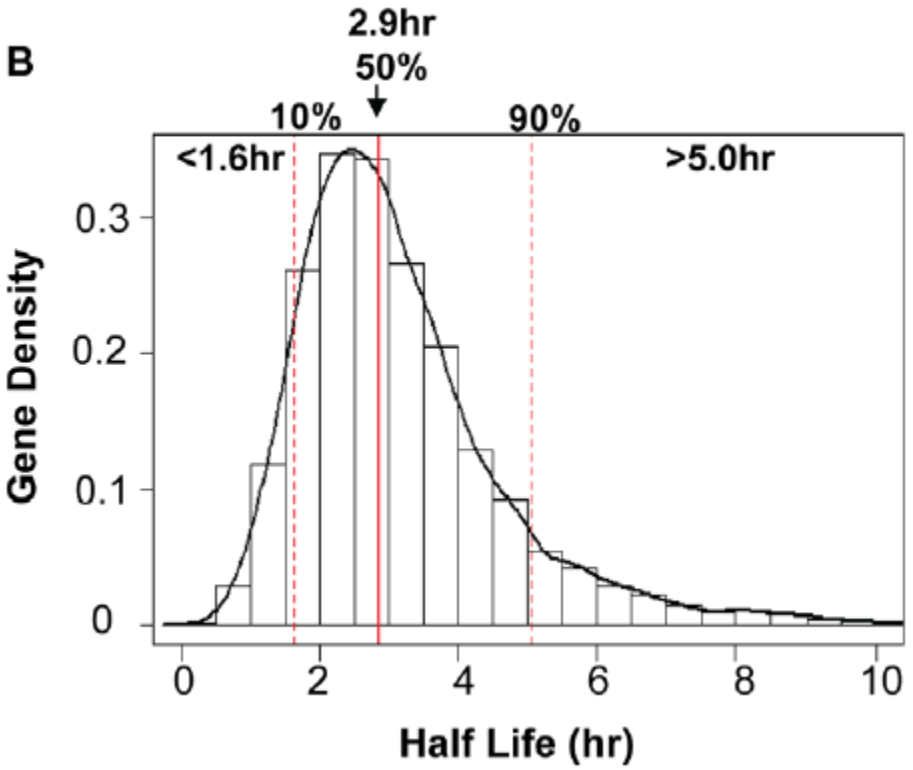


Figure 1. Analysis of mRNA decay rate in C2C12 cells.
(A) Examples of mRNA decay curves were derived by the nonlinear least squares method for a long and a short half life mRNA



(B) Distribution of mRNA half lives (see Dataset S1 for the complete list). The 10th-percentile and 90th-percentile values (indicated by red dotted lines) were used to select mRNAs with short and long half lives, respectively. The median value (2.9 hr) is indicated by a red line.

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

Questions:

1) Does different stability reflect **different functions** ?

Analysis in AMIGO (Gene Ontology)

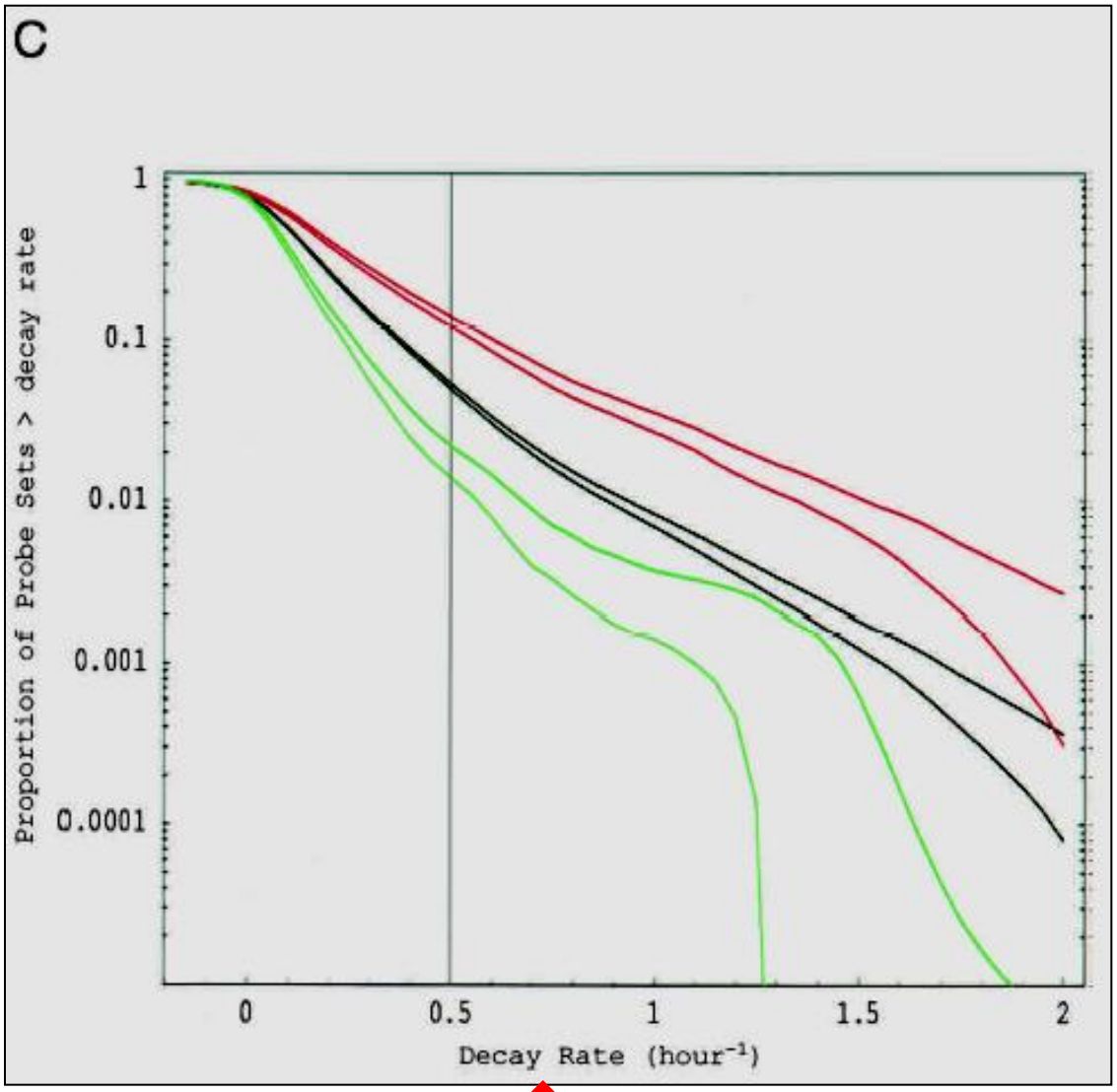
2) Are ***cis* elements** in RNA traceable that can be correlated to half-life ?

Analysis of «small word» over-representation

Gene Ontology

Figure 1 **Functional analysis** of decaying transcripts in human cells.

(C) Reverse cumulative distribution of decay rates for probe sets in different functional classes (HepG2 experiments). Decay rate r is shown horizontally, while vertically the fraction of probe sets with decay rates higher than r is plotted on a logarithmic scale. The pairs of lines show the 98% posterior probability intervals for the fraction at each value of r. (Red) GO process **transcription**; (black) all probe sets; (green) **biosynthesis**. The gray line indicates the decay rate r = 0.5/h, which is our cutoff for fast decay in PFDI.



Rate of decay

Decay rates were studied in functional groups, i.e. in G.O. categories.

Yang et al, 2003

Answer to our question 1) is:

RNAs half-lives are correlated to functions of the encoded protein

(*mRNA species that need rapid regulation must have short half-lives*)

…before answering to question 2 let's recall what we know about RNA decay mechanisms…

Searching for regulatory elements.

we can analyse short-living or long-living RNAs in order to find over-represented «motifs», that can suggest a stability regulatory sequence.

Yang et al. made this analysis using two different algorithms:

results are divided in 5’UTR, coding sequence, and 3’UTR

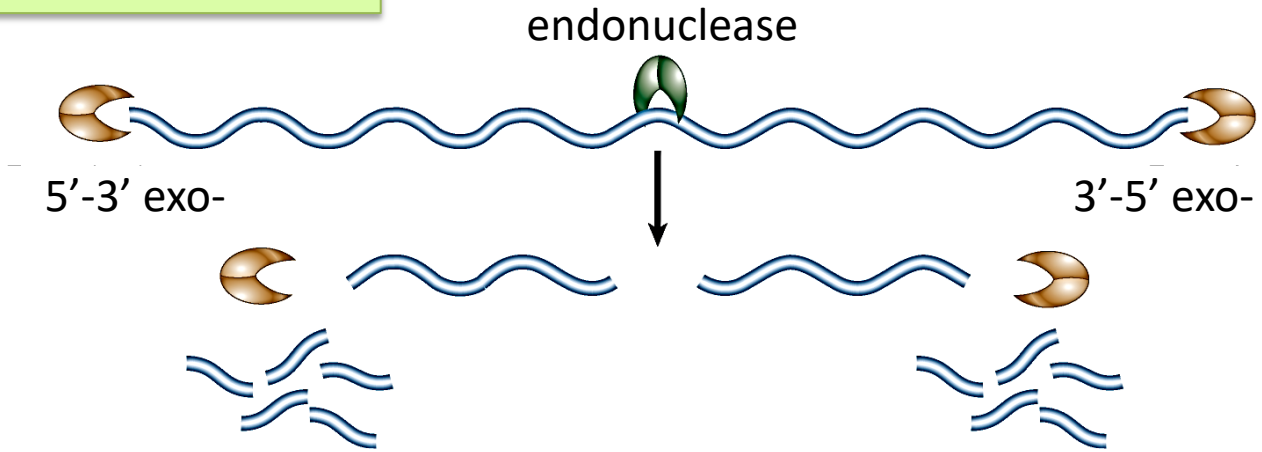
HepG2 motif summary

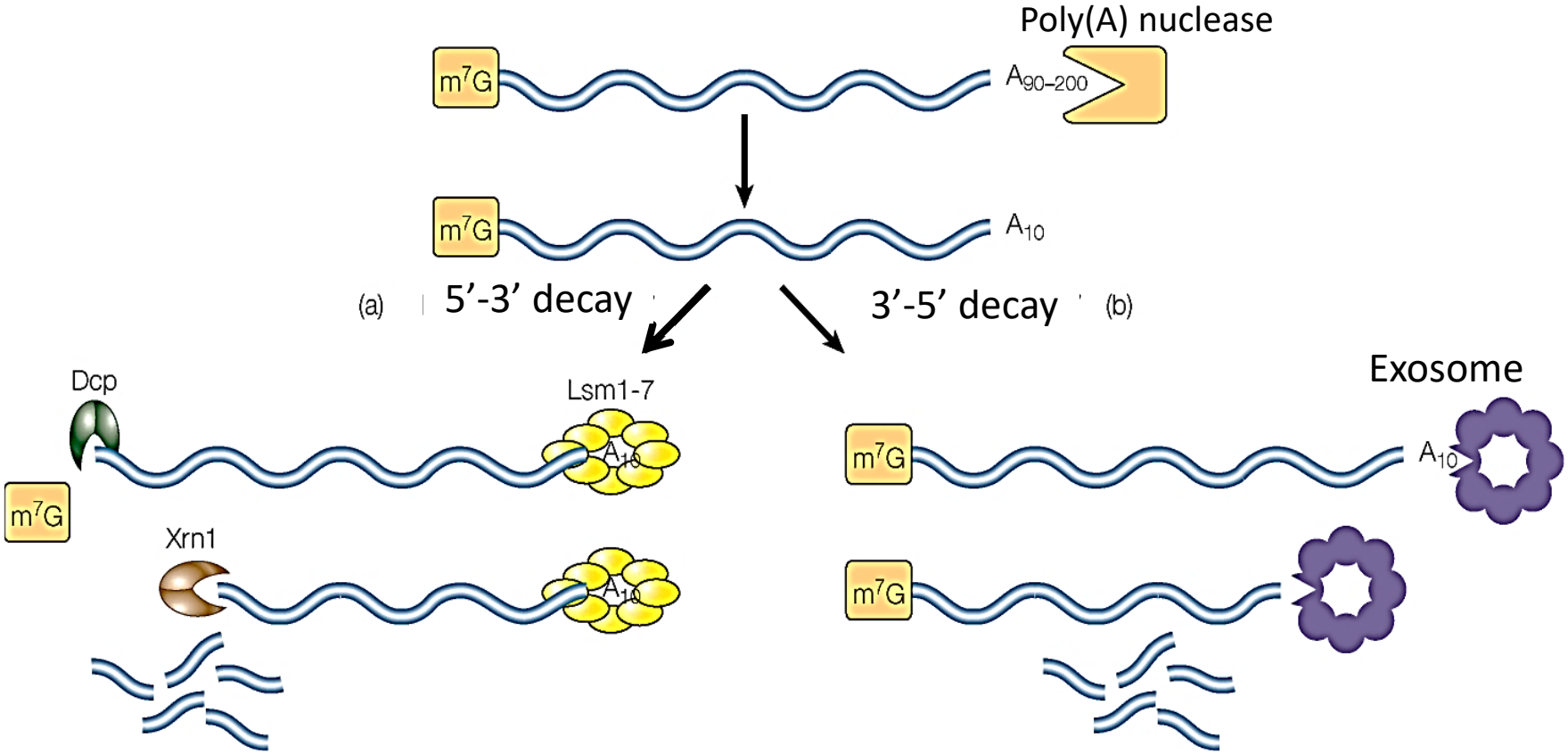
Known and novel motifs examined

		p-value for In-Group rate change			
Motif	Sequence	5'UTR	ORF	3'UTR	WHOLESEQ
Described Motifs					
1	(AT)ATTTA(AT)		2.02E-12	0	0
2A	ATTTATTTATTTATTTATTTA		na		
2B	ATTTATTTATTTATTTA		na		
2C	(AT)ATTTATTTATTTA(AT)		na		
2D	(AT)(2)ATTTATTTA(AT)(2)		na		
2E	(AT)(4)ATTTA(AT)(4)		3.50E-04	1.13E-13	1.38E-14
MEG	TTATTTATT			2.57E-07	1.00E-07
MEGSHORT	TATTTAT		5.03E-04	0	0
Undescribed Motifs					
H1	TTTTTTT			0	3.50E-20
H2	TTTTTAAA	3.83E-04	4.84E-04	0	0
H3	TTGTAAATA			4.77E-11	6.96E-10
B1	TTTTAAAT		1.48E-06	1.52E-11	1.57E-13
B2	TTTTAATTT				0.005
B3	AAATATTTT		0.004	6.05E-09	5.11E-10
B4	AATATTTTT			3.43E-09	6.77E-10
H-1	CCGCCTC		0.005		
H-2	CCAGCCTC		1.93E-08		4.18E-04
B-1	GGGCCTGG				
B-2	CCCAGCCCC		7.72E-04		

from Yang et al., 2003.

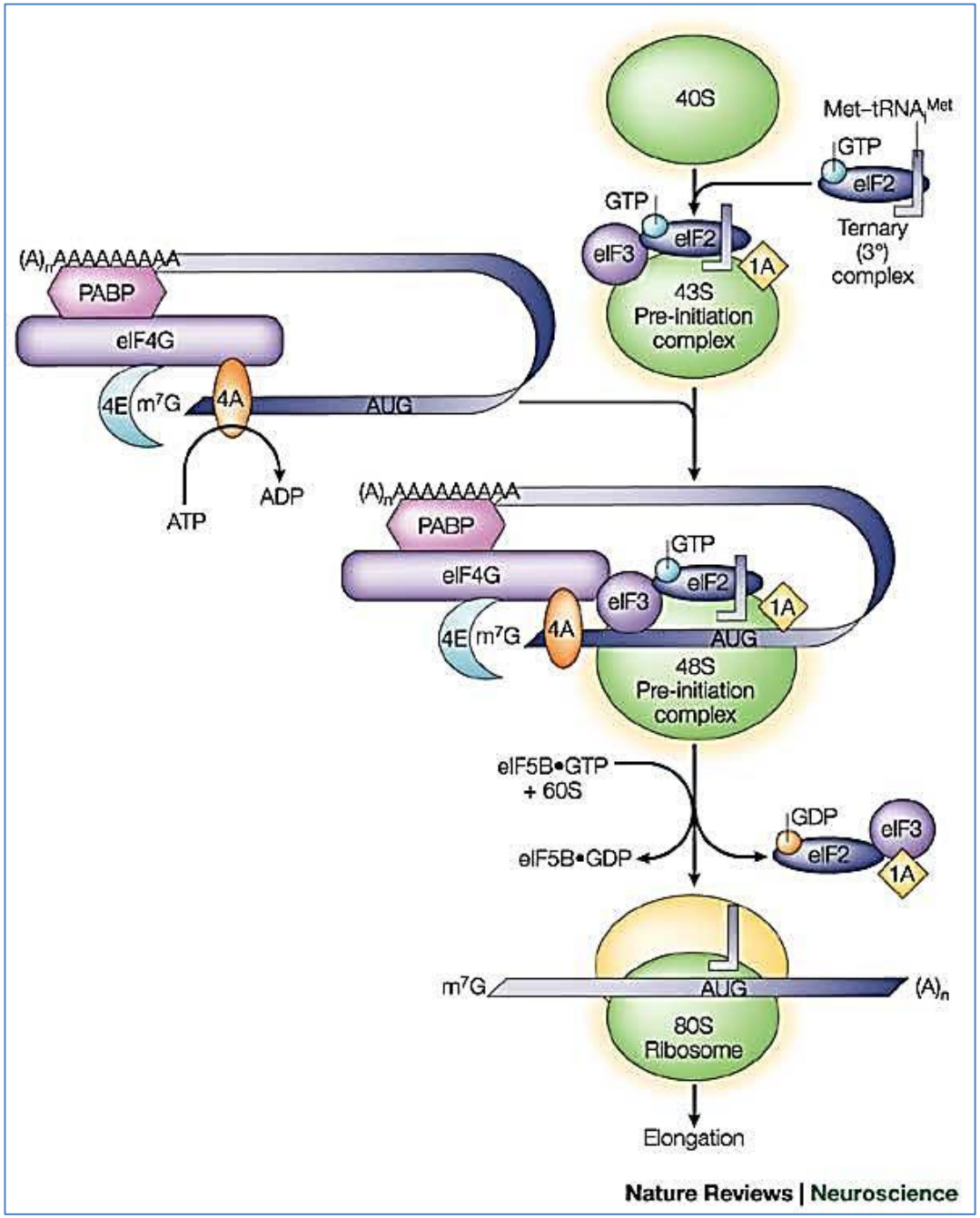
Pathways to RNA degradation



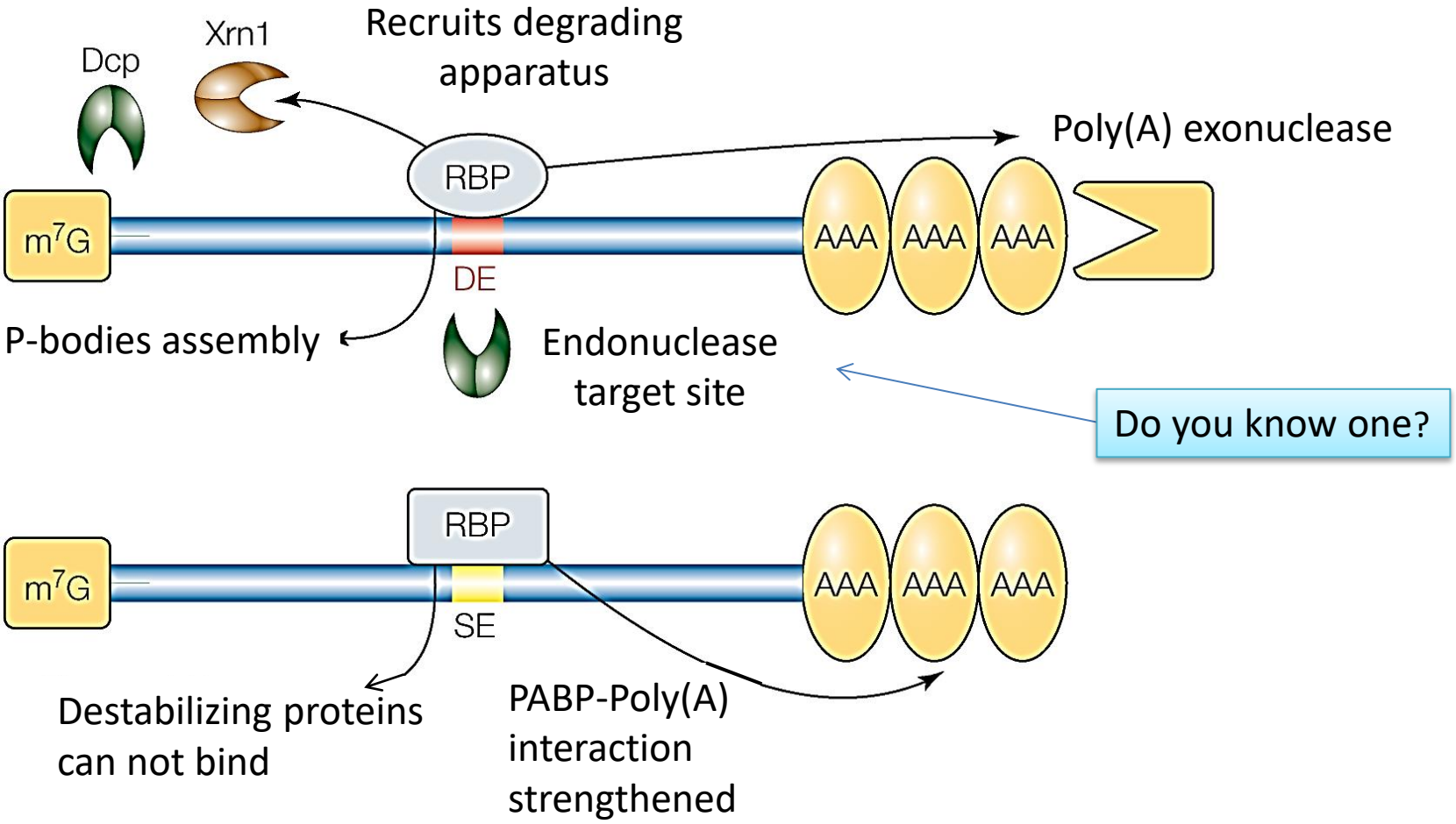


decapping or **polyadenylate shortening** strongly impairs translation, since interaction of cap and poly(A) is essential

Degradation and translation are competing events.



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.

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. These kind of elements are simpler to recognize by mapping the sequences back to the genome.

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

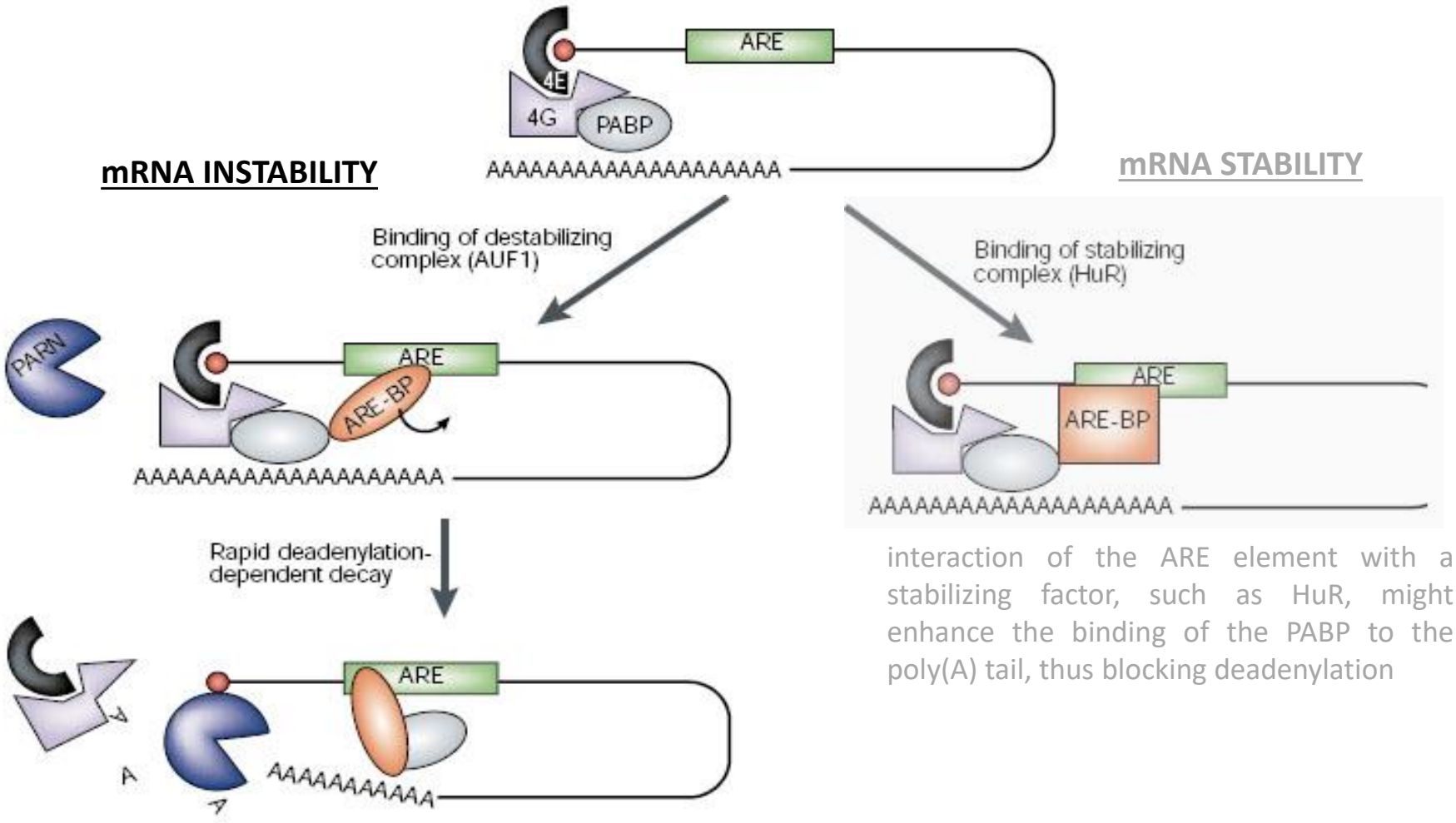
Group	Motif	Examples
I	WAUUUAW and a U-rich region	c-fos, c-myc
IIA	AUUUAUUUAUUUAUUUAUUUA	GM-CSF, TNF-α
IIB	AUUUAUUUAUUUAUUUA	Interferon-α
IIC	WAUUUAUUUAUUUAW	cox-2, IL-2, VEGF
IID	WWAUUUAUUUAWW	FGF2
IIE	WWWWAUUUAWWWW	u-PA receptor
III	U-rich, non-AUUUA	c-jun

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 and subcellular localization of the mRNA



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

interaction of the ARE element with a destabilizing factor, such as AUF1, might promote rapid deadenylation by reducing the affinity of the poly(A) binding protein (PABP) for the poly(A) tail

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 (February 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-guided AGO proteins**

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

RNA interference, small-interfering RNA, micro-RNA

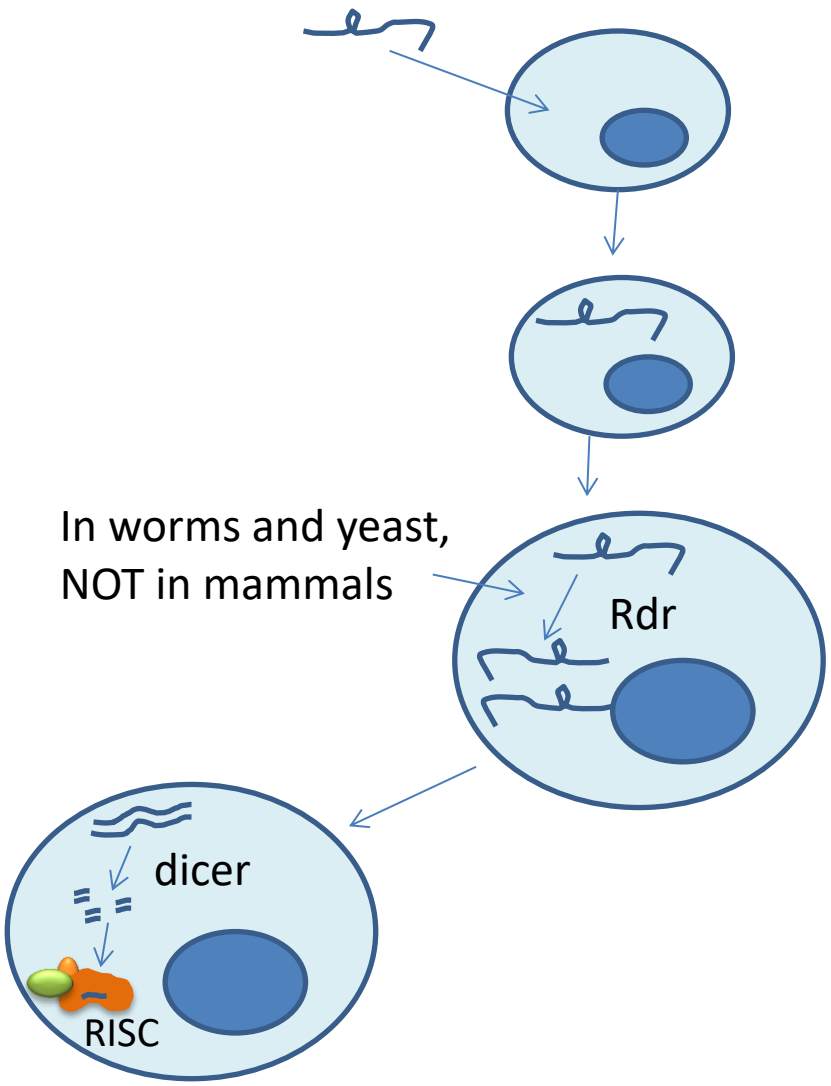
Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*

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

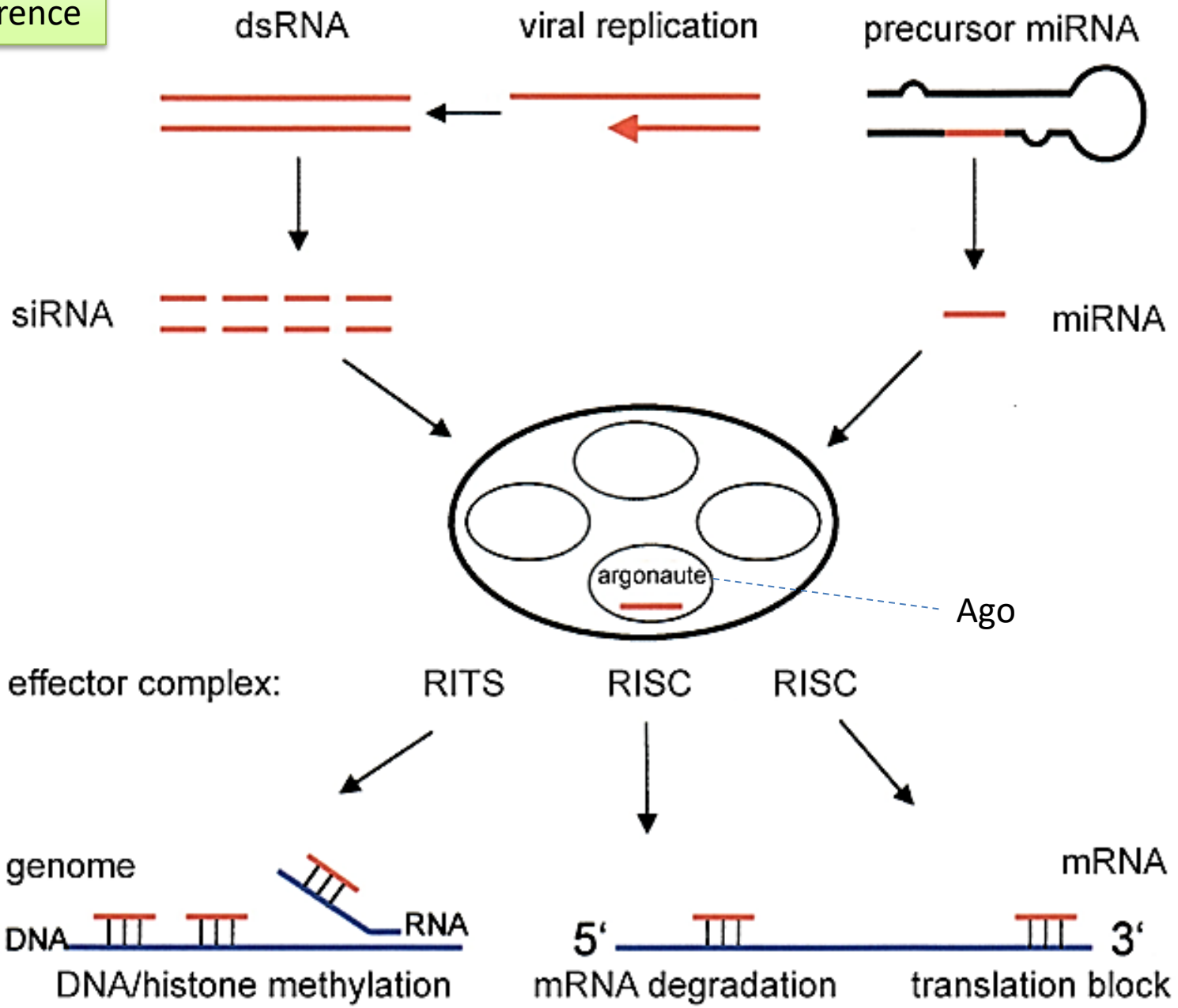
Andrew Fire*, SiQun Xu*, Mary K. Montgomery*, Steven A. Kostas*†, Samuel E. Driver‡ & Craig C. Mello‡

* *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.



RNA interference



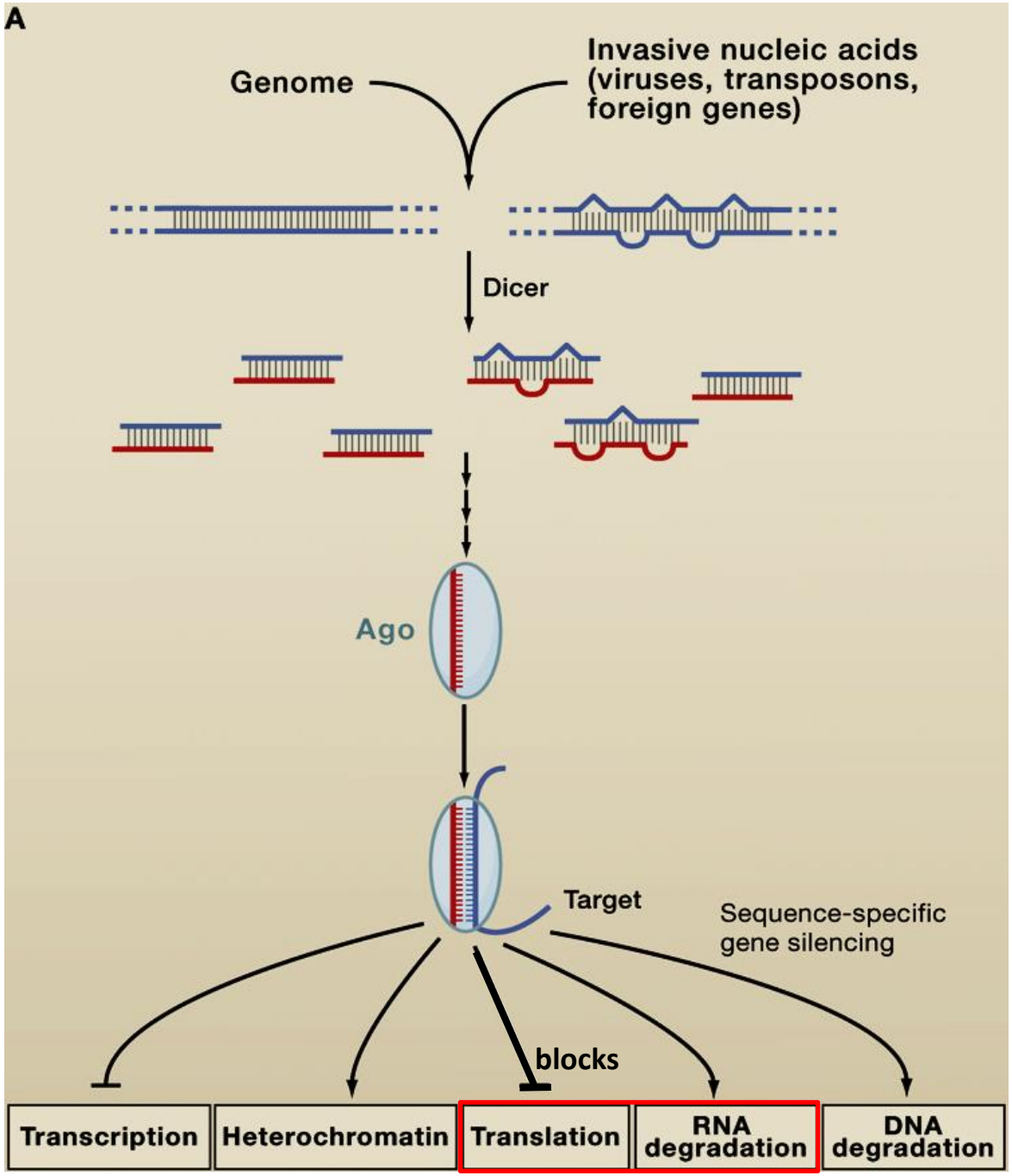
RNA interference

Endogenous and exogenous **double-stranded RNAs**

Dicer cleaves into 21-23nt pieces

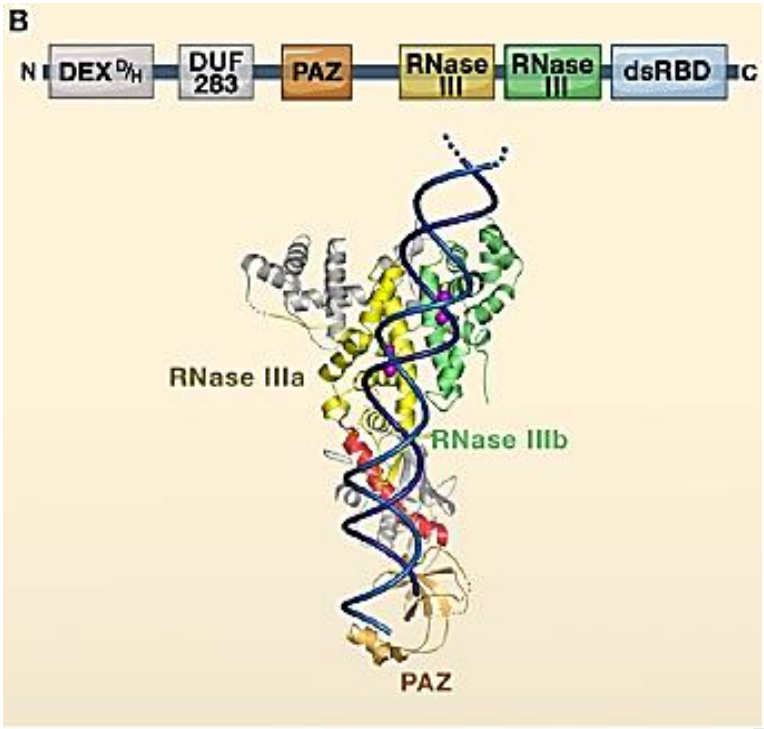
Si-RISC or mi-RISC

Effect



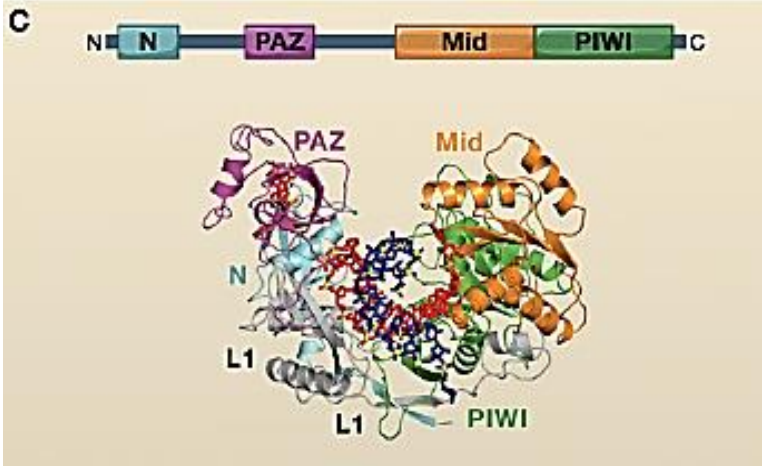
Cartew, 2009

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, whereas other classes have more enzymes.



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
Piwi domain has high affinity for ends with 3'-overhangs and in some case has endonucleolytic activity.
Ago enter a RISC assembly pathway: only one RNA strand stably associated, the other (passenger) is discarded.

Three major classes of **silencing small RNAs**.

- Micro RNA (miRNA, also miR, *but caution for confusion with MIR, a class of transposons*)
- Small interfering RNA (siRNA)
- Piwi-dependent RNA (piRNA)

- Some new classes emerged from NGS RNA-Seq experiment, but no clear evidence of function is available to date (see next table)

In the lab:

Gene silencing using siRNA. (and shRNA)

DICER product is like this

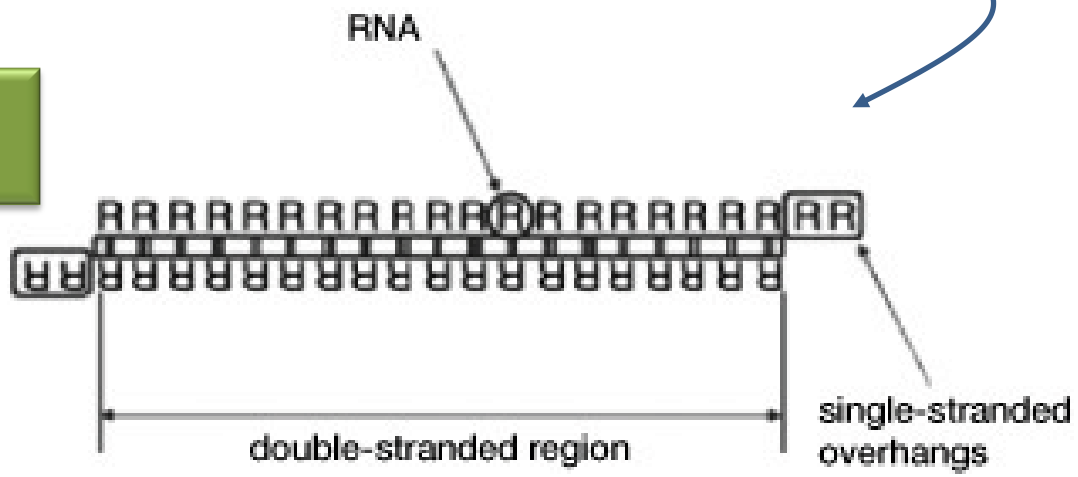


Table 1

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

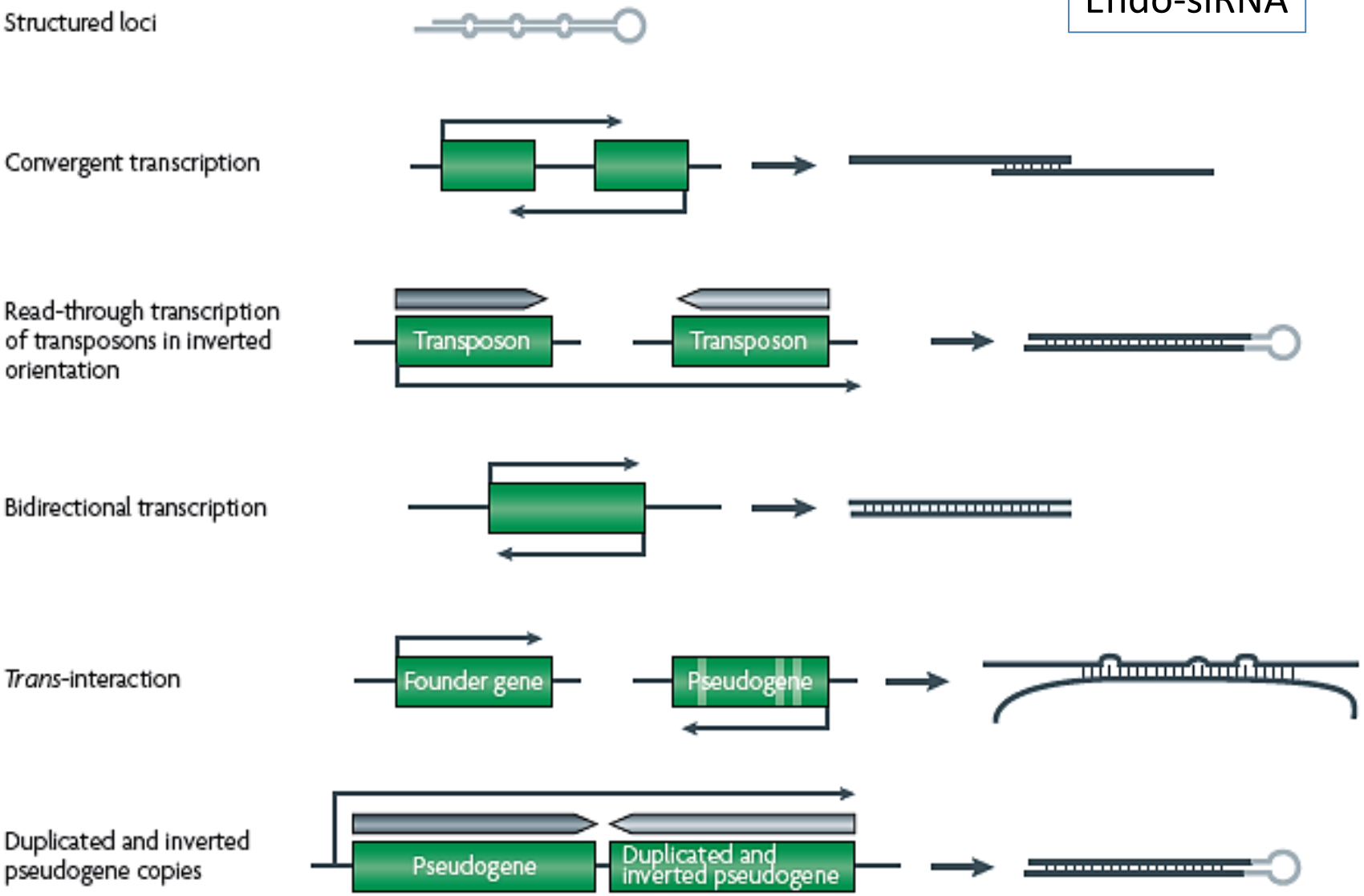


Figure 3 | **Genomic sources of dsRNA triggers for endogenous small interfering RNAs (endo-siRNAs) in flies and mammals.**

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

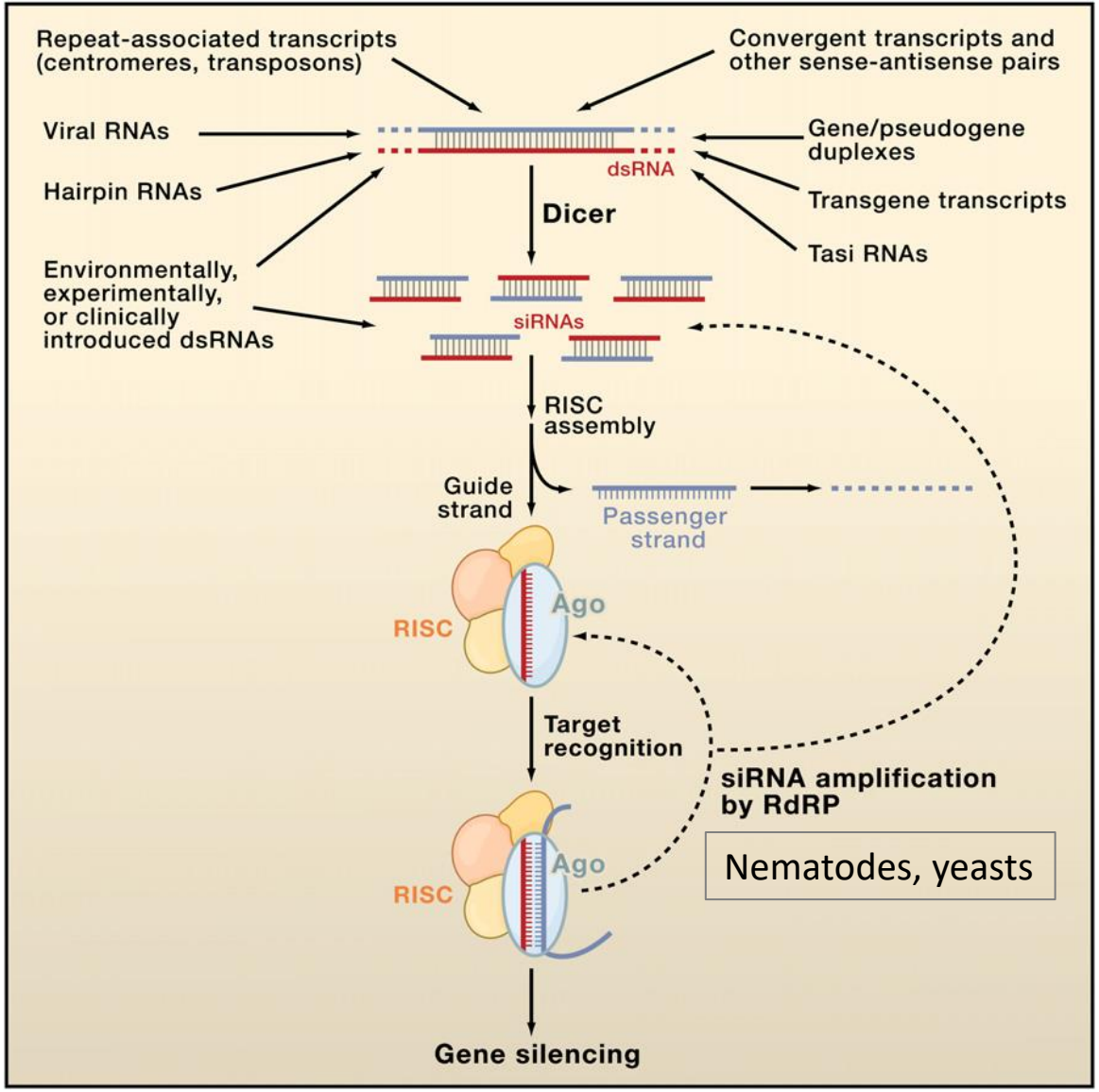
They have other specific features:

- 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 then pairs with transposon sense transcript and cleave them endonucleolitically.

piRNAs were first proposed to ensure germline stability by repressing transposons

Mammalian piRNAs can be divided into pre-pachytene and pachytene piRNAs, according to the stage of meiosis at which they are expressed in developing spermatocytes. like piRNAs in flies, pre-pachytene piRNAs predominantly correspond to repetitive sequences and are implicated in silencing transposons, such as L1 and intracisternal A-particle.

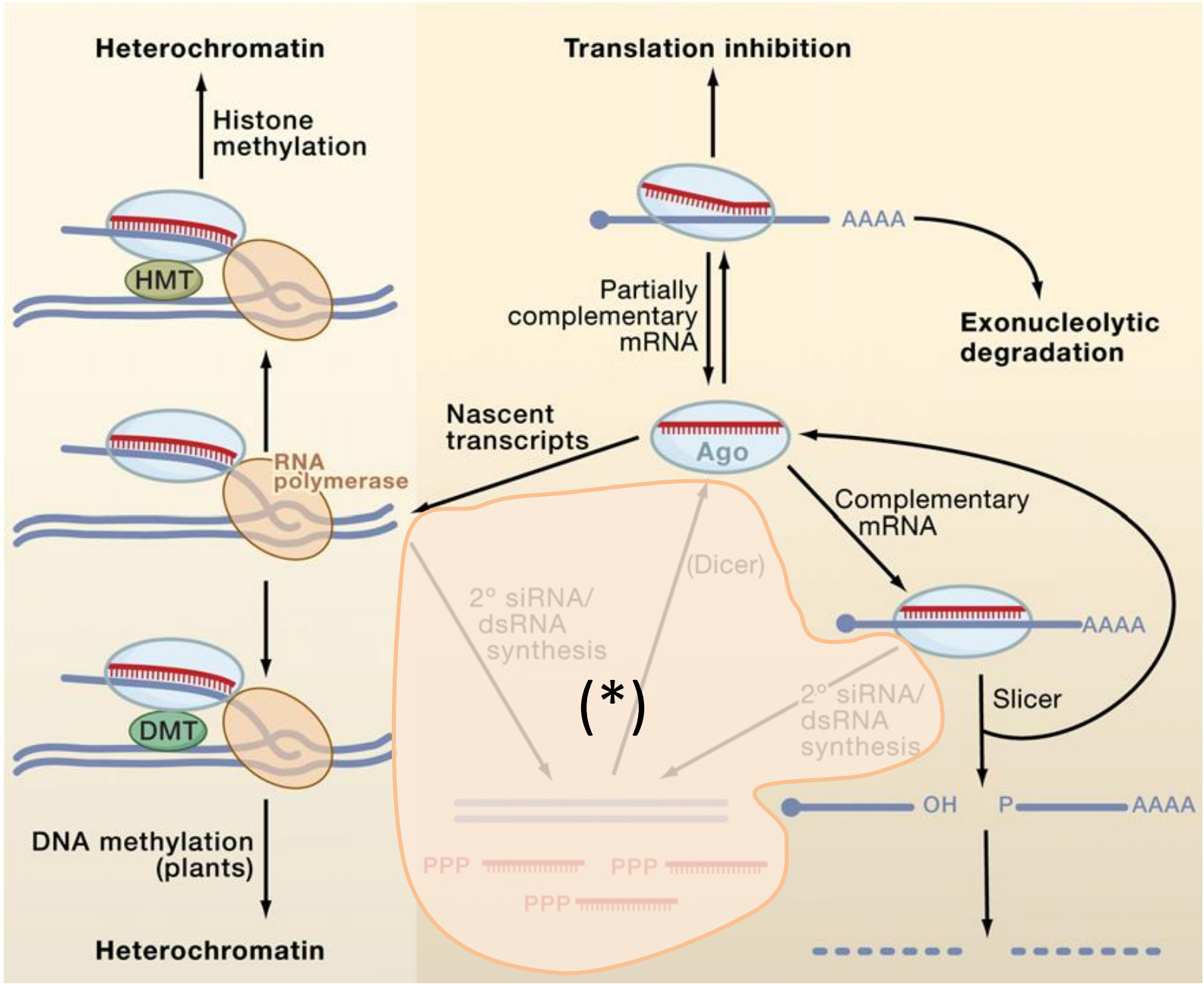


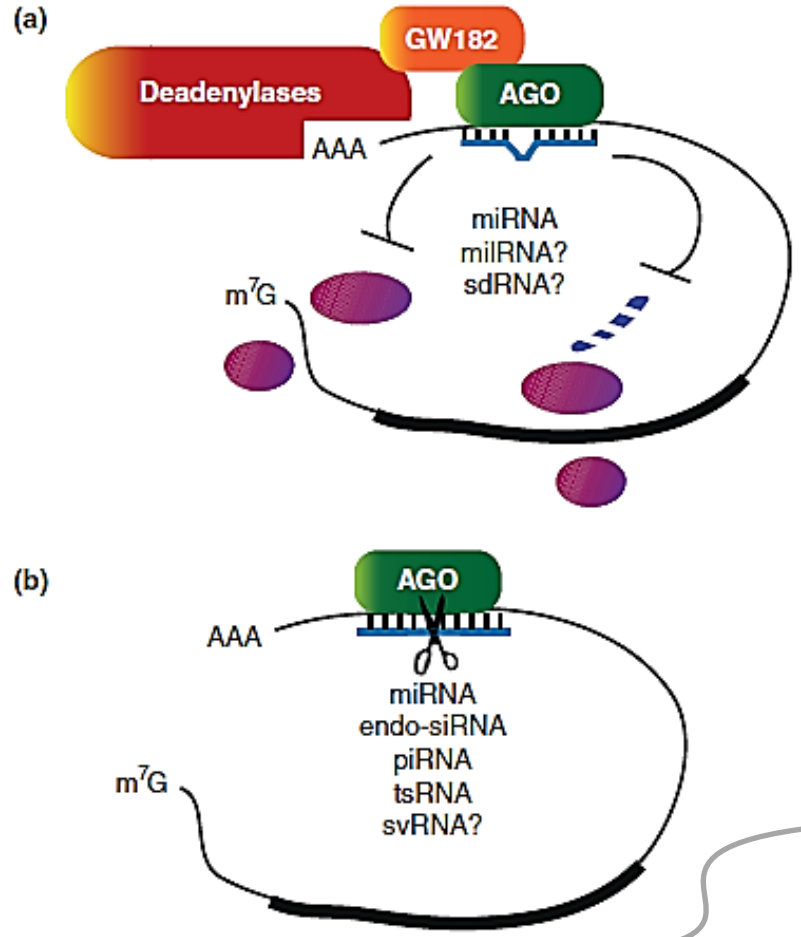
siRISC assembly
Insects:
R2D2/Dicer heterodimer binds siRNA duplex
+ other factors→ RISC loading compex
+ Ago2 → pre-RISC
Ago2 cleaves the passenger strand (eliminated) → RISC

Humans:
Dicer, TRBP, Ago2 → assembling
Binds dsRNA, processes to siRNA, eliminates passenger → RISC

Which siRNA strand is selected ?
thermodynamic stabilities of the two duplex ends: Guide strand is the one that has its 5' terminus at the less stably base-paired end.
Selection graded: siRNAs with equal base-pairing stabilities at their ends → either strand used (Opposite strand miRNA is called with a «∗»).

How do small interfering RNA exert their silencing effects?

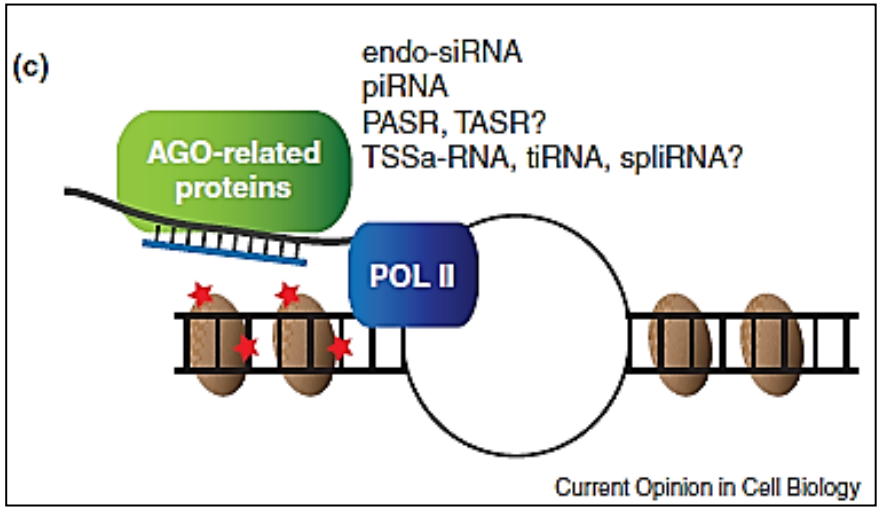

Heterochromatin
Histone methylation
HMT
RNA polymerase
DMT
DNA methylation (plants)
Heterochromatin
Translation inhibition
AAAA
Partially complementary mRNA
Exonucleolytic degradation
Nascent transcripts
Ago
Complementary mRNA
(Dicer)
2° siRNA/ dsRNA synthesis
(*)
2° siRNA/ dsRNA synthesis
AAAA
Slicer
OH
P
AAAA
PPP
PPP
PPP



Many roads to silencing.
Mechanisms used by sRNAs to silence gene expression.

(a) Typically, miRNAs, and possibly milRNAs and sdRNAs, partially pair to mRNA target sequences and promote mRNA degradation through the recruitment of deadenylases by GW182 proteins bound to Argonaute (AGO). Alternatively, miRISC inhibits translation initiation, stalls translation elongation or stimulates proteolysis of nascent peptides encoded by the target mRNA.

(b) Perfect pairing of many types of sRNAs results in cleavage of the target mRNA by Ago2.



(c) 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

From Aalto 2012

Textbook

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Cell Biology

Small non-coding RNAs mount a silent revolution in gene expression

Antti P Aalto[1,a] and Amy E Pasquinelli[2]

During the past decade, it has become evident that small non-coding RNAs (ncRNAs) participate in widespread and essential regulatory mechanisms in most eukaryotic cells. Novel classes of small RNAs, their biogenesis pathways and cellular effects are continuously being described, and new properties of already established ncRNAs are still being discovered. As the list of small RNA molecules and their roles becomes more and more extensive, one can get lost in the midst of new information. In this review, we attempt to bring order to the small ncRNA transcriptome by covering some of the major milestones of recent years. We go through many of the new properties that have been attributed to already familiar RNA molecules, and introduce some of the more recent novel classes of tiny ncRNAs.

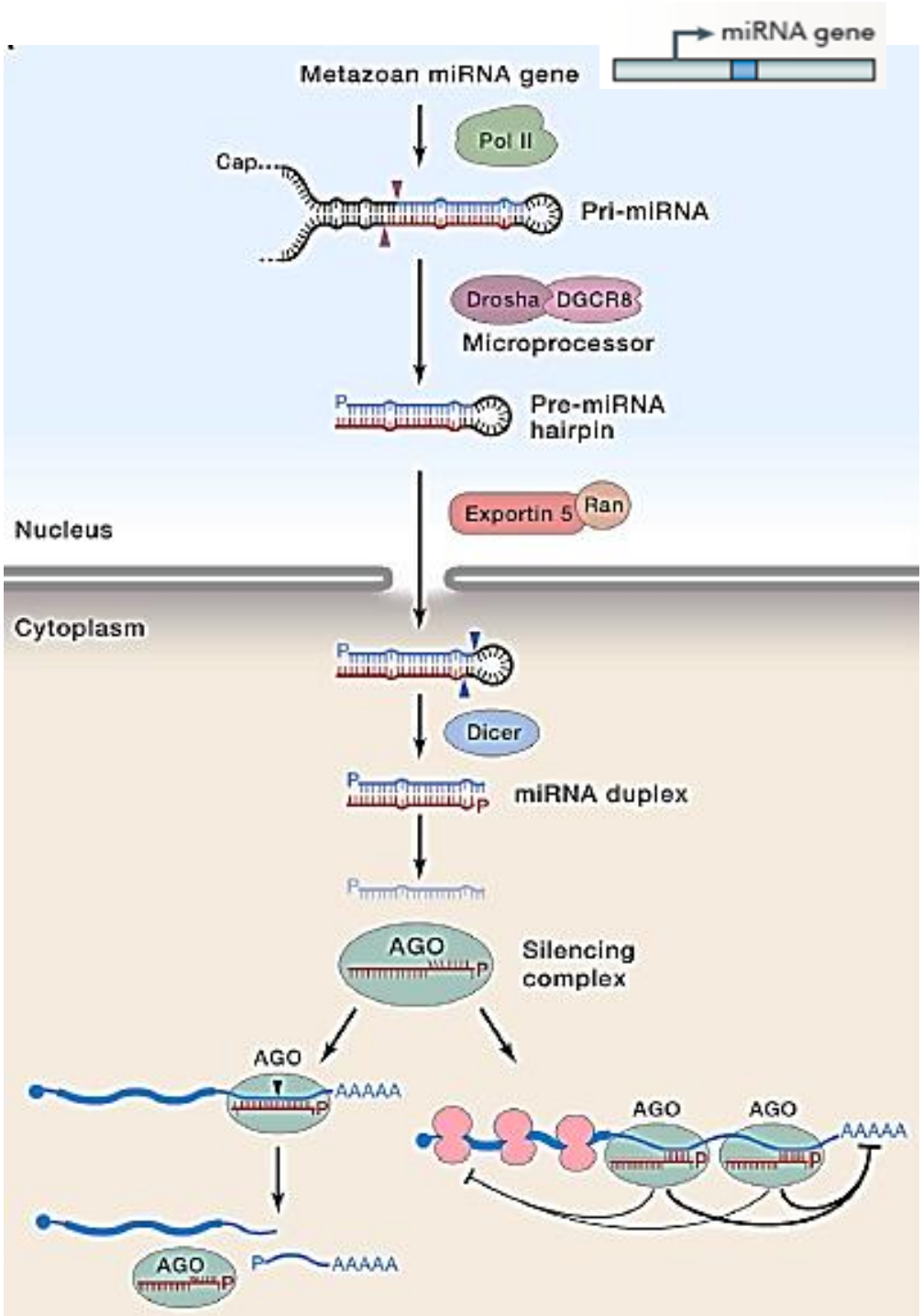
Current Opinion in Cell Biology 2012, 24:333–340

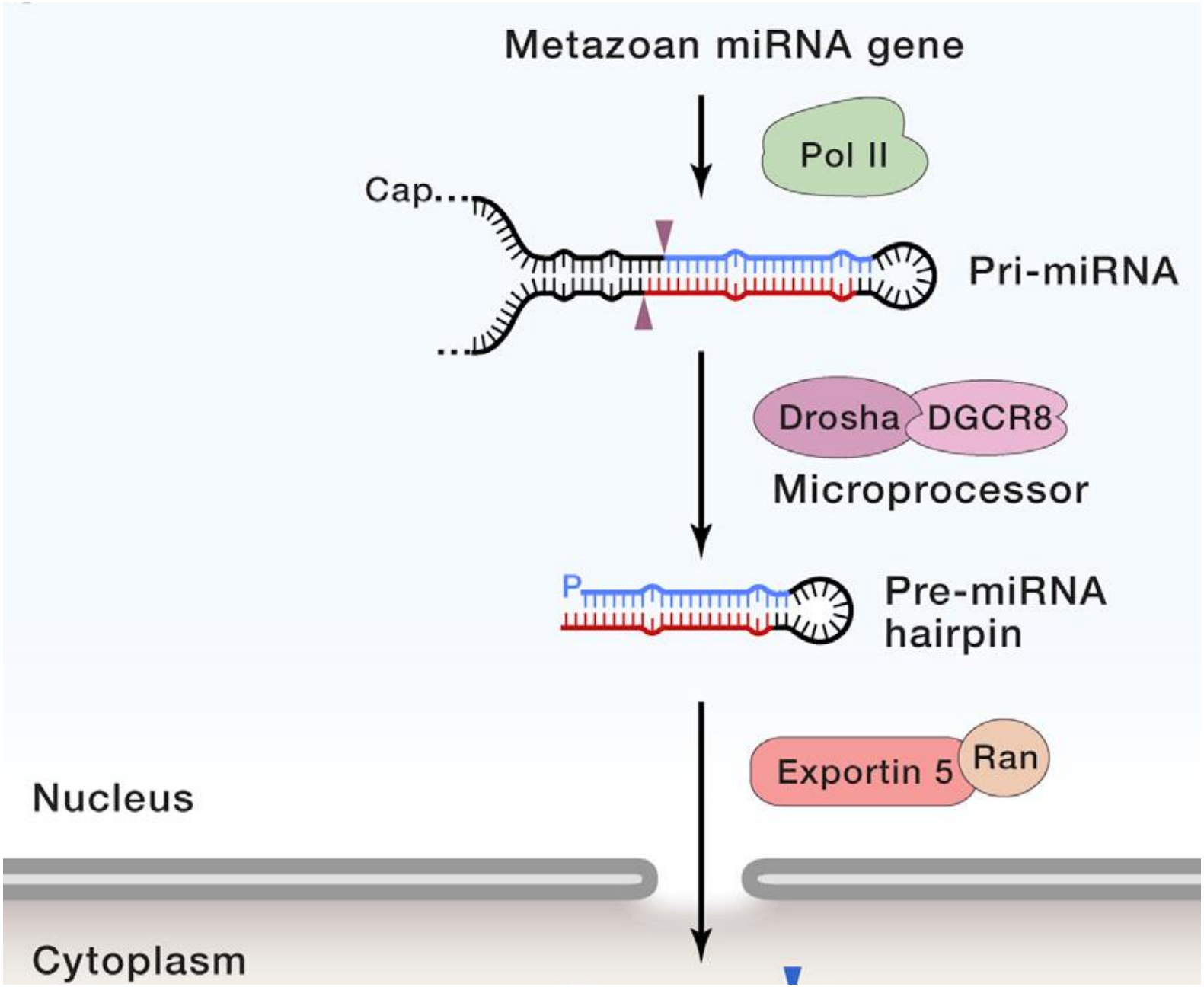
miRNA

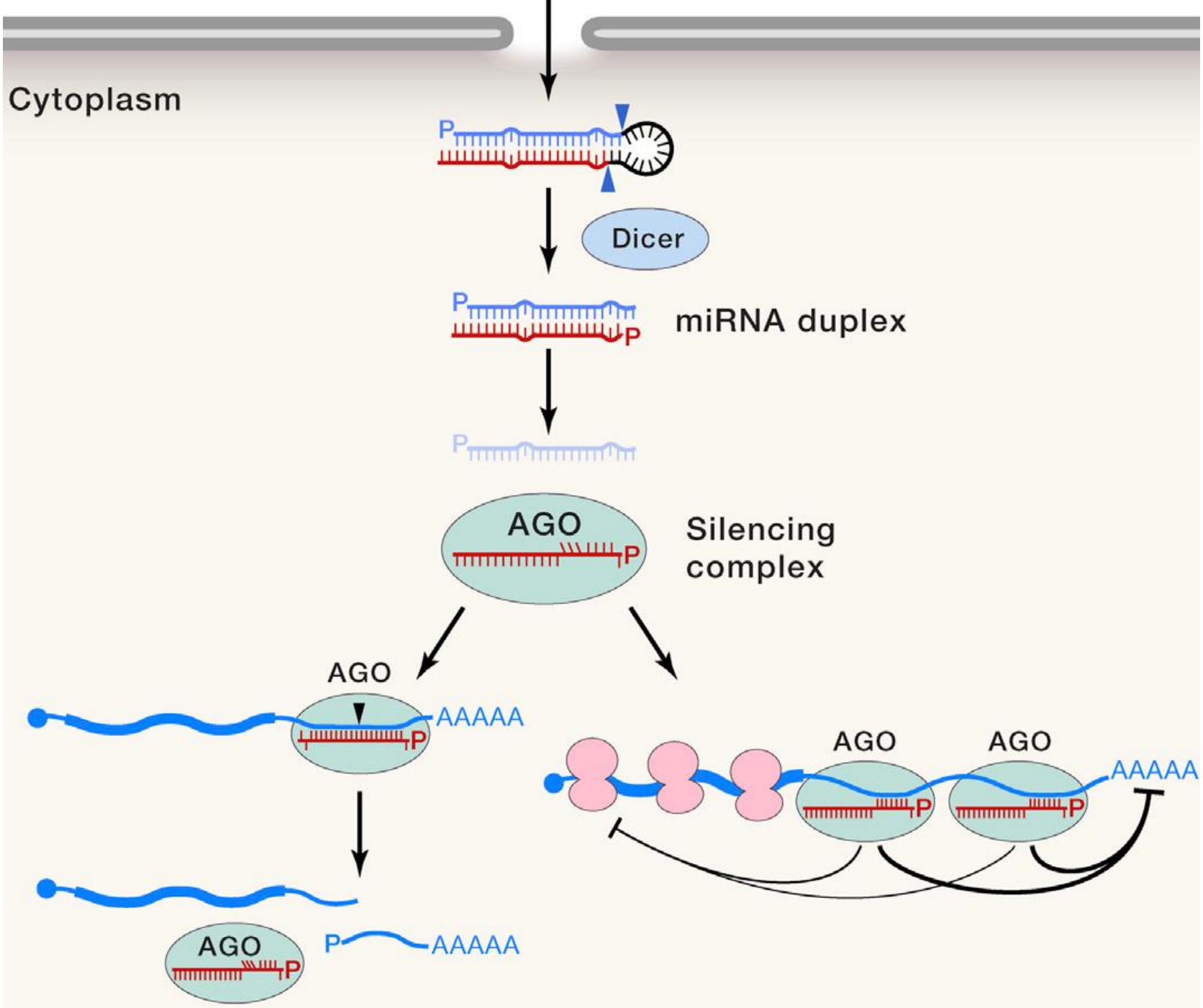
Micro RNA are a family of small RNA that are transcribed from several locations in genomes.

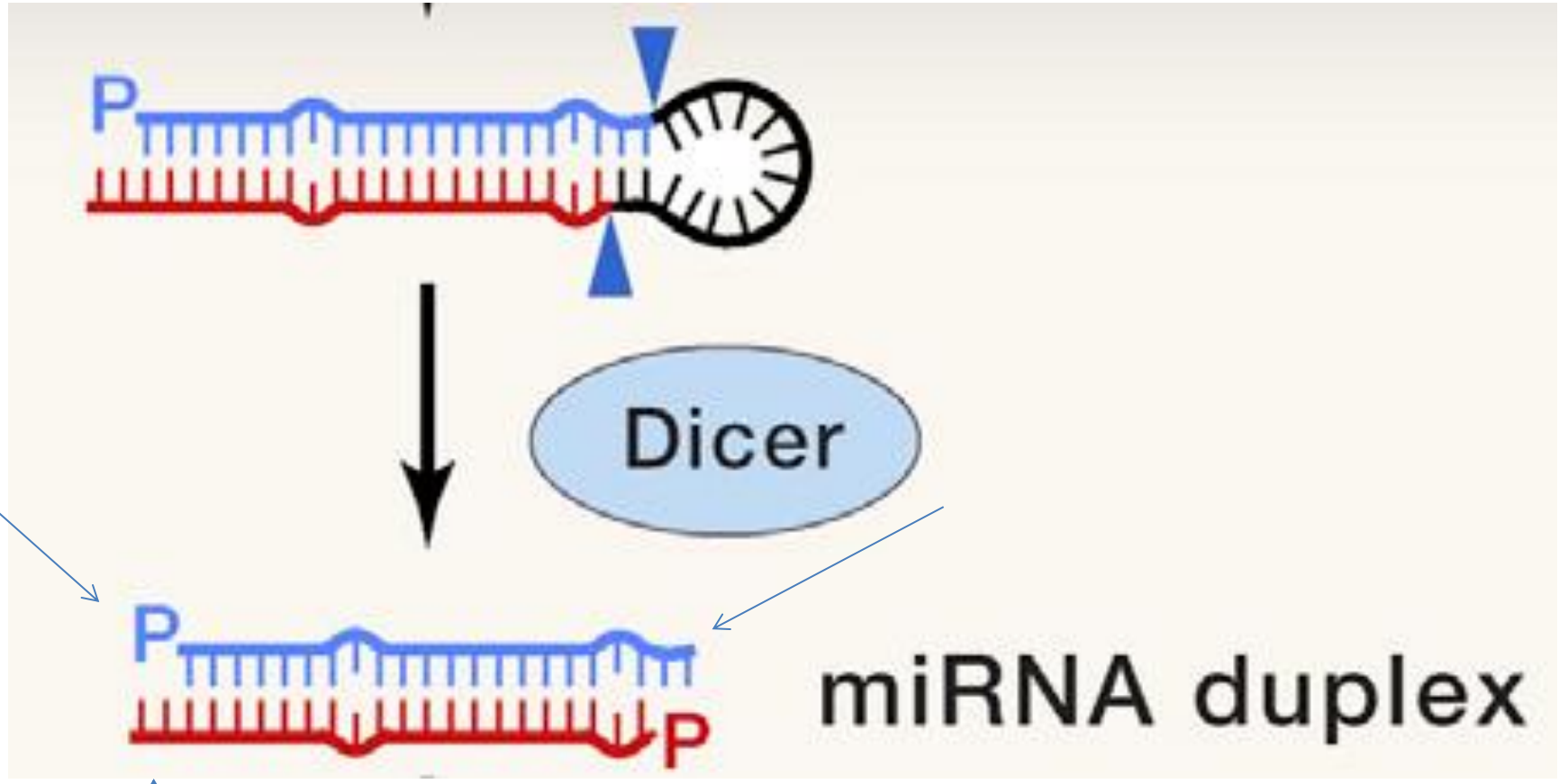
They have a typical structure, making a stem-loop structure with some mismatches in the stem

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


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


Metazoan miRNA gene
Pol II
Cap
Pri-miRNA
Drosha
DGCR8
Microprocessor
P
Pre-miRNA
hairpin
Exportin 5
Ran
Nucleus
Cytoplasm


Cytoplasm
P
Dicer
P
P
miRNA duplex
P
AGO
P
Silencing
complex
AGO
P
AAAAA
AGO
AGO
AAAAA
P
P
AGO
P
P
AAAAA


P
Dicer
P
P
miRNA duplex

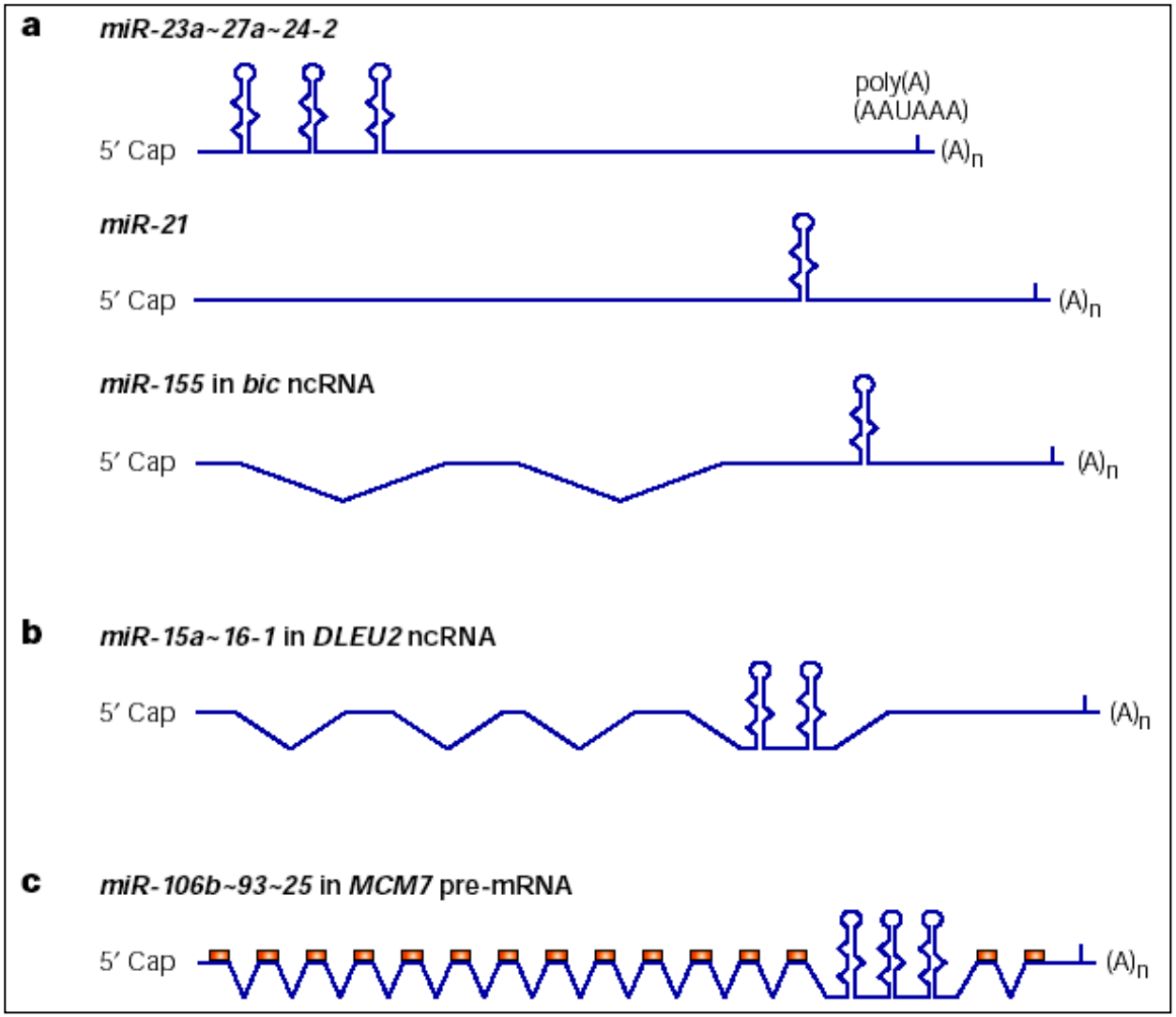
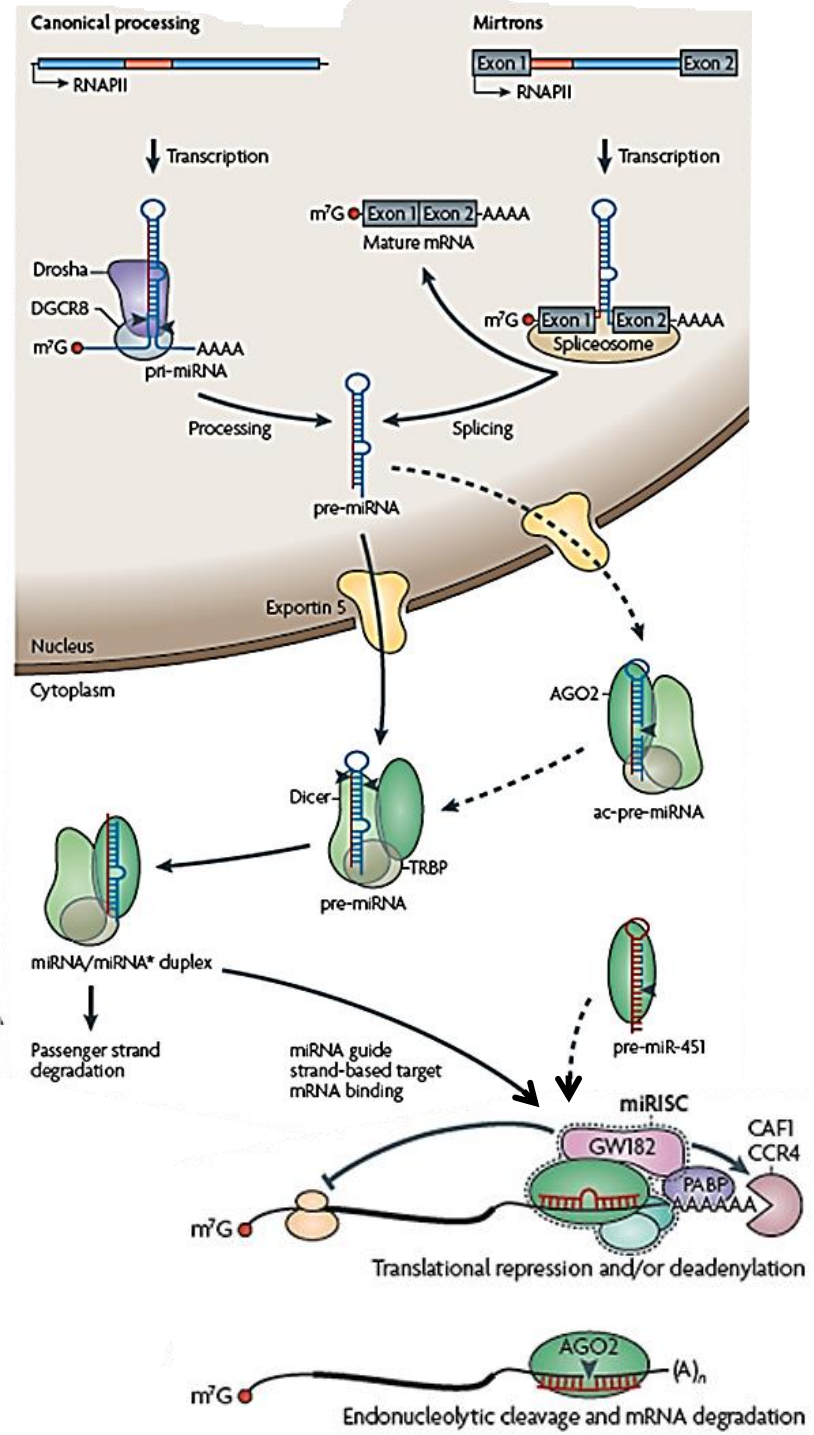


Figure 1 | **The structure of five pri-miRNAs.**

Primary transcripts that encode miRNAs, primiRNAs, contain 5′ cap structures as well as 3′ poly(A) tails. miRNAs can be categorized into three groups according to their genomic locations relative to their positions in an exon or intron.

a | Exonic miRNAs in non-coding transcripts such as an *miR-23a~27a~24-2* cluster, *miR-21* and *miR-155*. *miR-155* was found in a previously defined non-coding RNA (ncRNA) gene, *bic*17.

b | Intronic miRNAs in non-coding transcripts. For example, an *miR-15a~16-1* cluster was found in the fourth intron of a previously defined non-coding RNA gene, *DLEU2* (REF. 126). **c** | Intronic miRNAs in protein-coding transcripts. For example, an *miR-106b~93~25* cluster is embedded in the thirteenth intron of DNA replication licensing factor *MCM7* transcript (variant 1, which encodes isoform 1). The mouse *miR-06b~93~25* homologue is also found in the thirteenth intron of the mouse *MCM7* homologue gene15. The hairpins indicate the miRNA stem-loops. Orange boxes indicate the protein-coding region. This figure is not to scale.



MicroRNAs (miRNAs) are processed from RNA polymerase II (RNAPII)-specific transcripts of independent genes or from introns of protein-coding genes. In the canonical pathway, primary precursor (pri-miRNA) processing occurs in two steps, catalysed by two members of the RNase III family of enzymes, Drosha and Dicer, operating in complexes with dsRNA-binding proteins (dsRBPs), for example DGCR8 and transactivation-responsive (TAR) RNA-binding protein (TRBP) in mammals. In the first nuclear step, the Drosha–DGCR8 complex processes pri-miRNA into an ~70-nucleotide precursor hairpin (pre-miRNA), which is exported to the cytoplasm. Some pre-miRNAs are produced from very short introns (mirtrons) as a result of splicing and debranching, thereby bypassing the Drosha–DGCR8 step. In either case, cleavage by Dicer, assisted by TRBP, in the cytoplasm yields an ~20-bp miRNA/miRNA* duplex. In mammals, argonaute 2 (AGO2), which has robust RNaseH-like endonuclease activity, can support Dicer processing by cleaving the 3′ arm of some pre-miRNAs, thus forming an additional processing intermediate called AGO2-cleaved precursor miRNA (ac-pre-miRNA)70. Processing of pre-miR-451 also requires cleavage by AGO2, but is independent of Dicer and the 3′ end is generated by exonucleolytic trimming. Following processing, one strand of the miRNA/miRNA* duplex (the guide strand) is preferentially incorporated into an miRNA-induced silencing complex (miRISC), whereas the other strand (passenger or miRNA*) is released and degraded (not shown). Generally, the retained strand is the one that has the less stably base-paired 5′ end in the miRNA/miRNA* duplex. miRNA* strands are not always by-products of miRNA biogenesis and can also be loaded into miRISC to function as miRNAs. See BOX 2 for details of miRISC function. GW182, glycine-tryptophan protein of 182 kDa; m7G, 7-methylguanosine-cap; PABP, poly(A) binding protein.
(*From Krol 2010 NRG, modified)*