MicroRNAs (I)

Regulación de la expresión génica: Niveles de regulación

La regulación génica puede tomar lugar en cualquiera de los pasos que llevan del DNA (gen) a la proteína

Eucariotas tienen mas posibilidades de regulación que procariotas en los que se regula principalmente el inicio:

- La estructura de la cromatina puede y tiene que ser modificada de diferentes maneras para iniciar la transcripción
- El inicio y los niveles de transcripción se regulan mediante factores de transcripción que se unen a elementos regulatorios en *cis*
- *Splicing*, (adición del *cap* 5') y cola poli(A)
- Proceso del transporte al citoplasma
- Degradación del mRNA: mayor rango de vida media en eucariotas (en procariotas se degradan rápidamente)
- **Regulación post-transcripcional (microRNAs)**
- Se regula la iniciación de la traducción
- Modificaciones postraduccionales pueden cambiar la conformación o llevar a la degradación

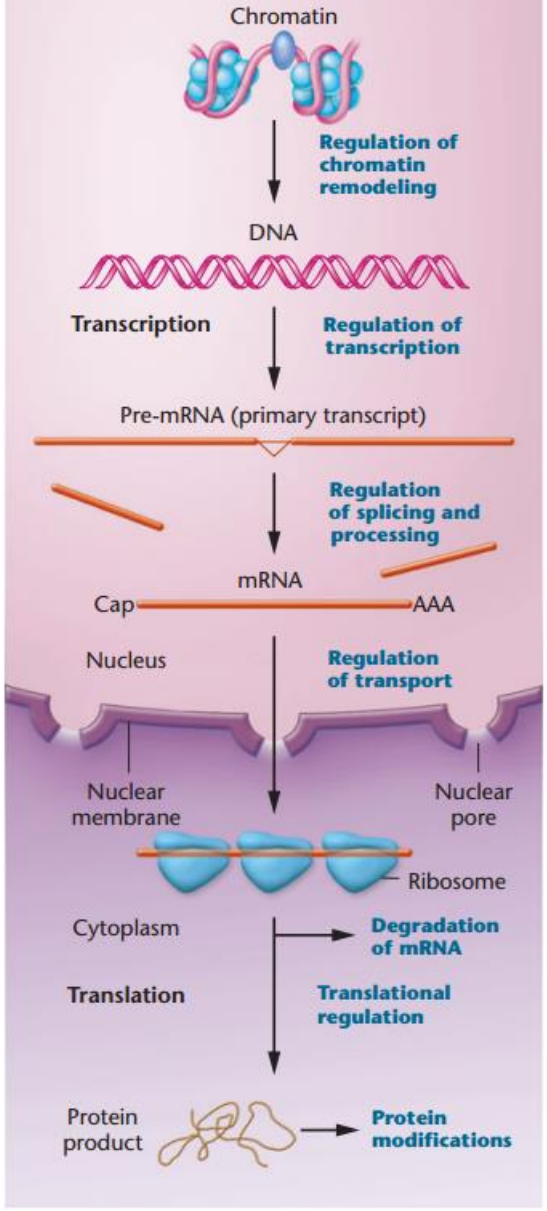
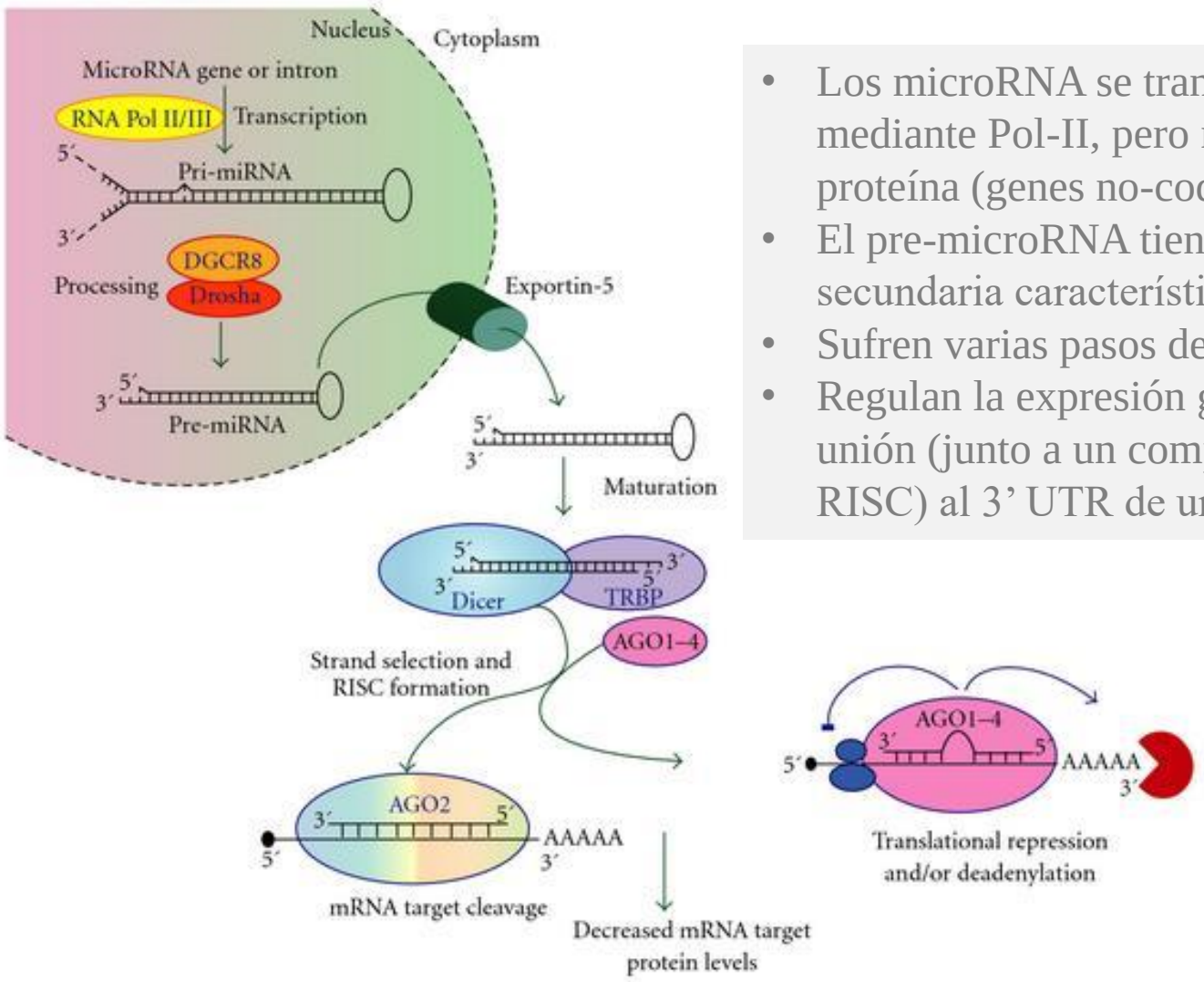


FIGURE 17–1 Regulation can occur at any stage in the expression of genetic material in eukaryotes. All these forms of regulation affect the degree to which a gene is expressed.

Un microRNA es un RNA corto de entre 20 y 23 nt de longitud

1. Regulación génica post-transcripcional
2. Metilación del ADN en plantas.



- Los microRNA se transcriben por general mediante Pol-II, pero no se traducen a proteína (genes no-codificantes)
- El pre-microRNA tiene una estructura secundaria característica de ‘horquilla’
- Sufren varias pasos de maduración
- Regulan la expresión génica mediante la unión (junto a un complejo proteico llamado RISC) al 3’ UTR de un gen

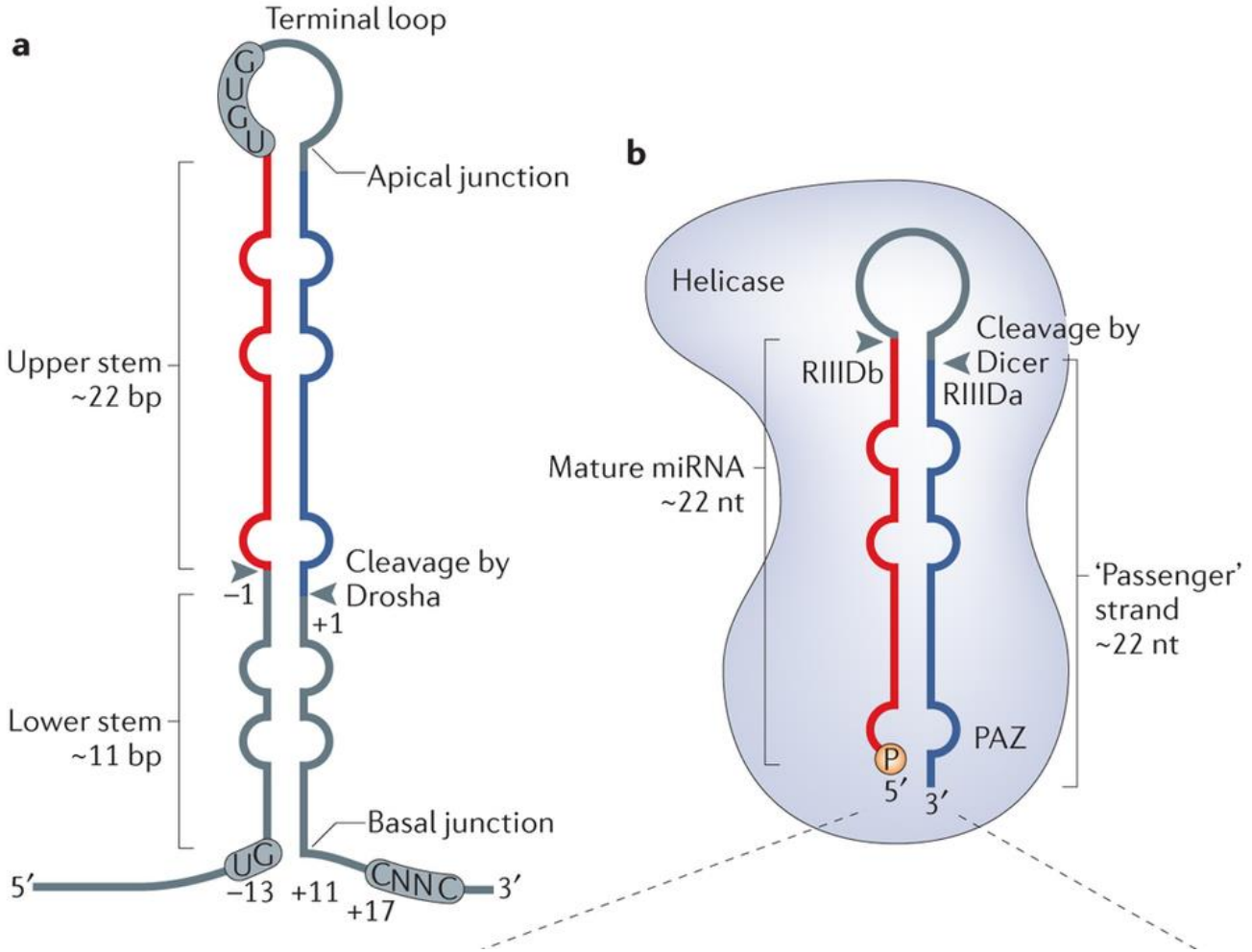
Biogénesis en animales

From

Regulation of microRNA biogenesis

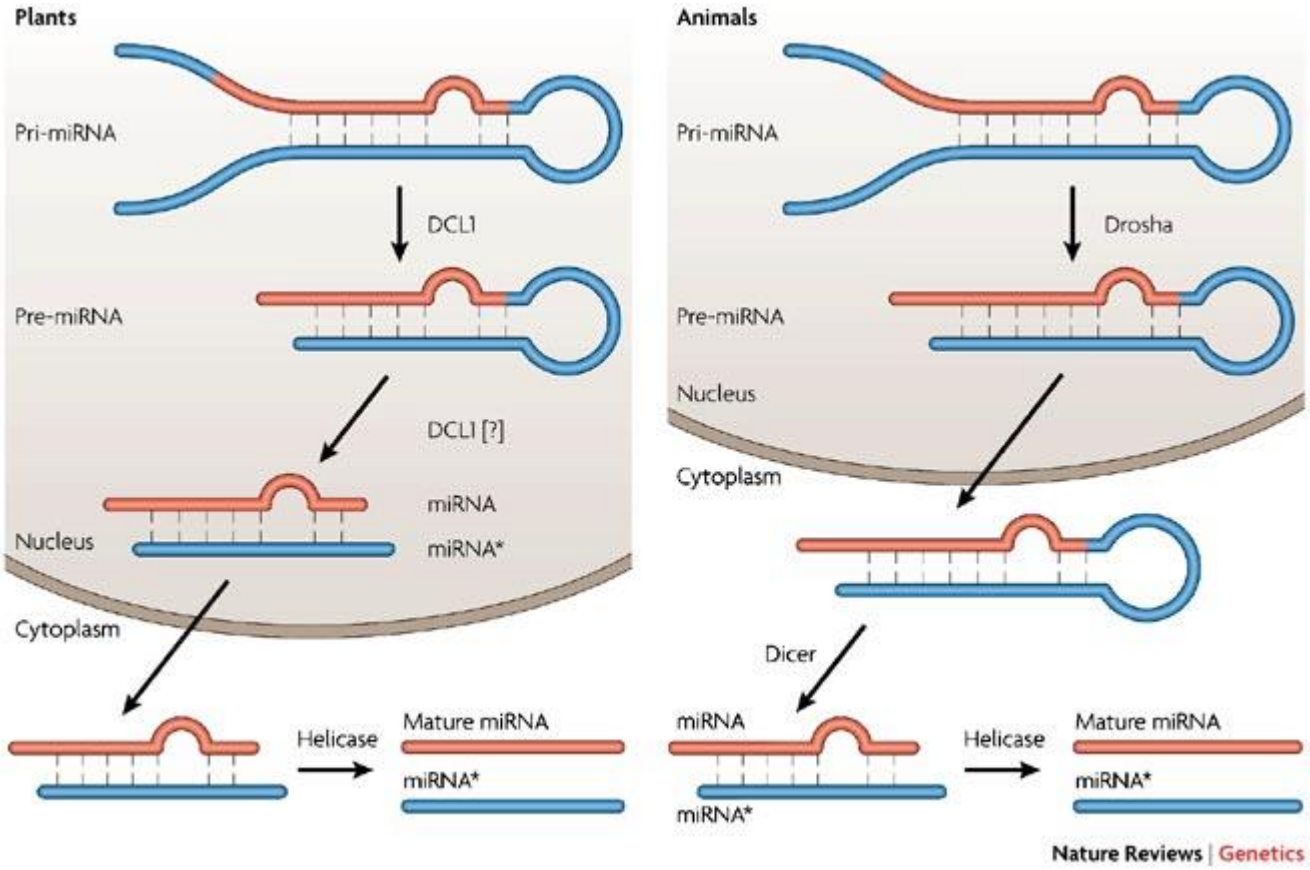
Minju Ha & V. Narry Kim

Nature Reviews Molecular Cell Biology **15**, 509–524 (2014) | doi:10.1038/nrm3838



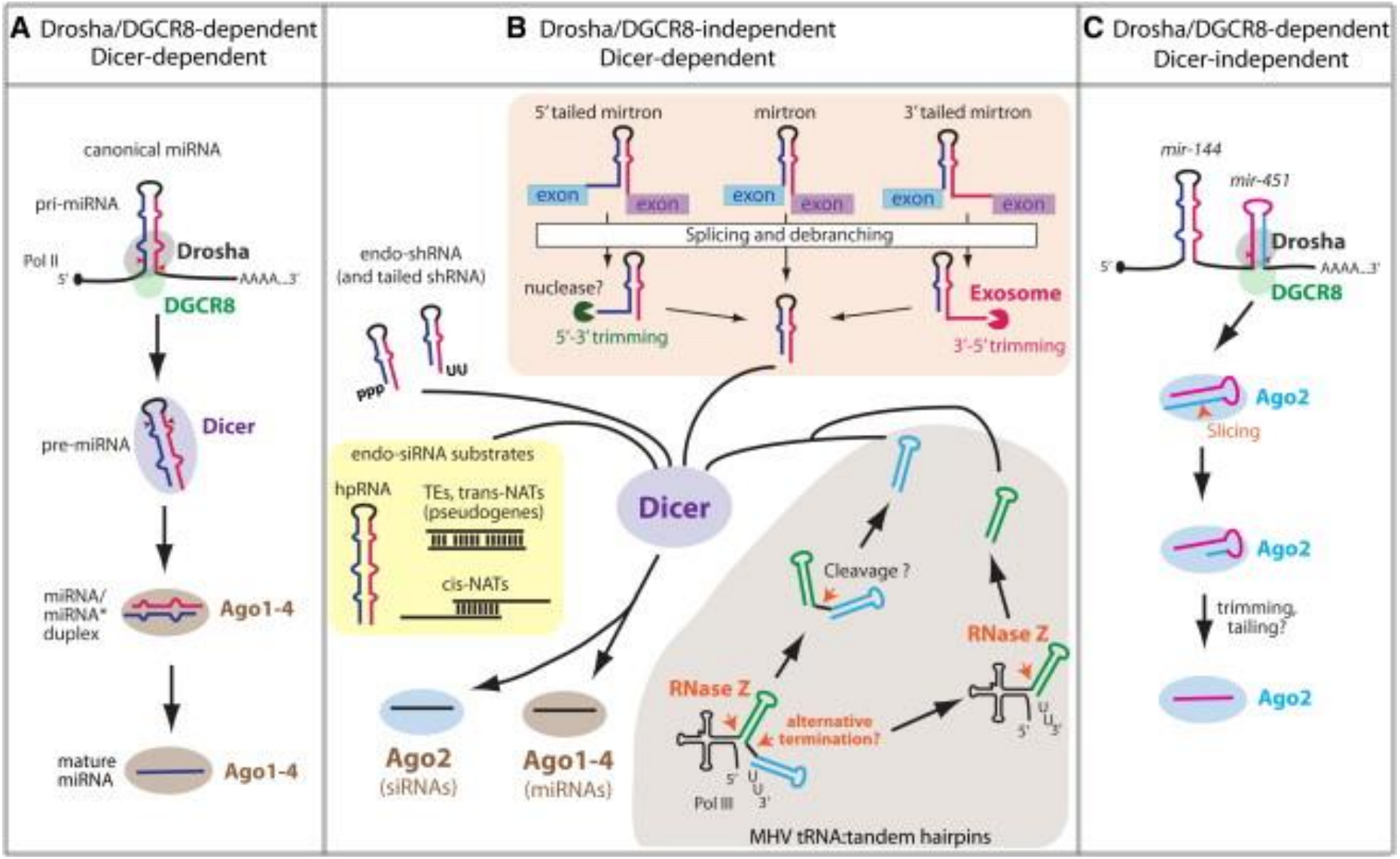
Comparación con plantas

- **Se desconoce microRNAs homólogos entre animales y plantas**
- **La estructura secundaria es similar – la longitud es mas heterogénea en plantas**
- **En plantas se finaliza la biogénesis en el núcleo**



FROM THE FOLLOWING ARTICLE:
The evolution of gene regulation by transcription factors and microRNAs
Kevin Chen & Nikolaus Rajewsky
Nature Reviews Genetics **8**, 93-103 (February 2007)
doi:10.1038/nrg1990

Biogénesis non-canónica

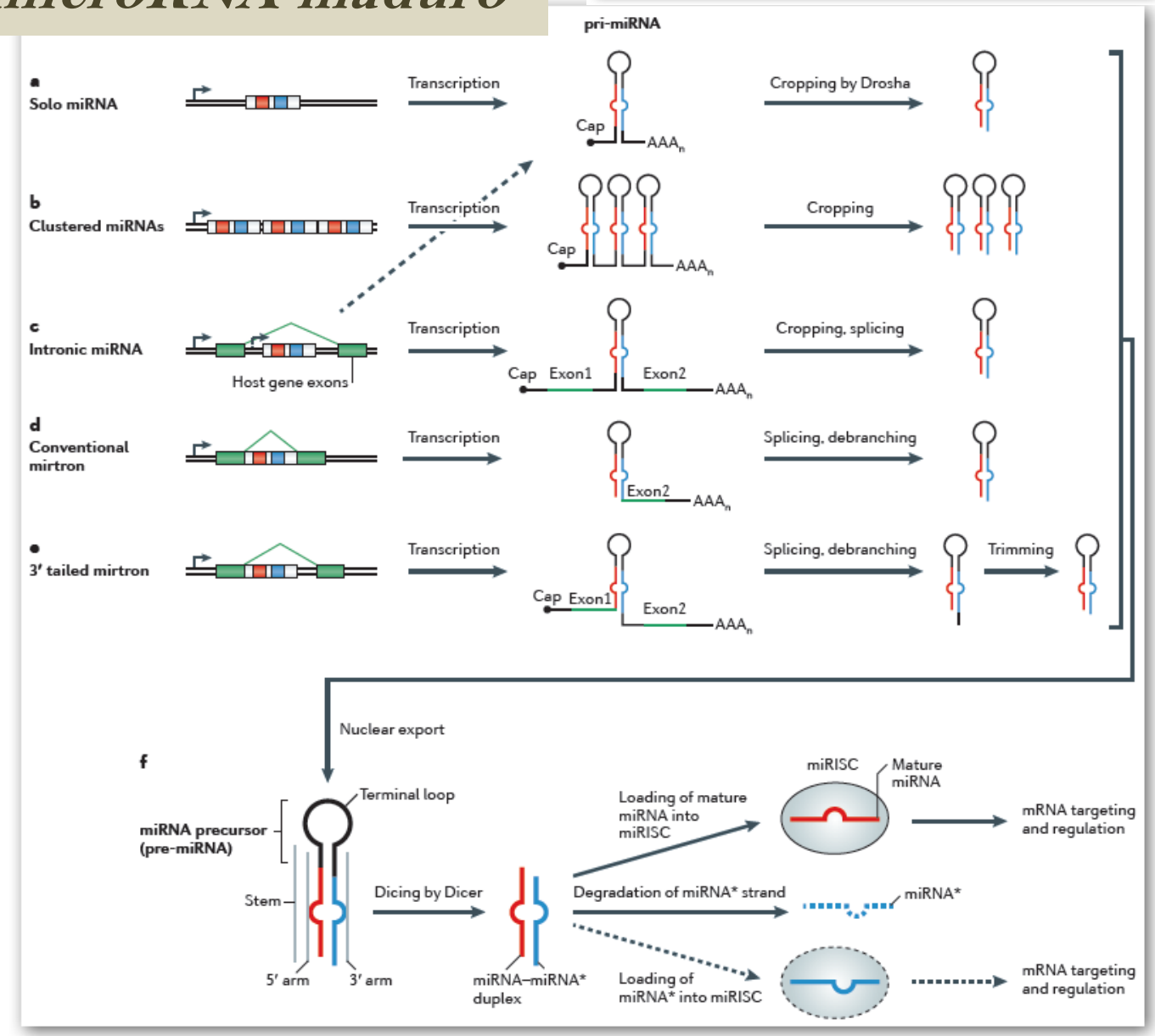


Distribución cromosómica y selección del microRNA maduro

NON-CODING RNA

Evolution of microRNA diversity and regulation in animals

Eugene Berezikov

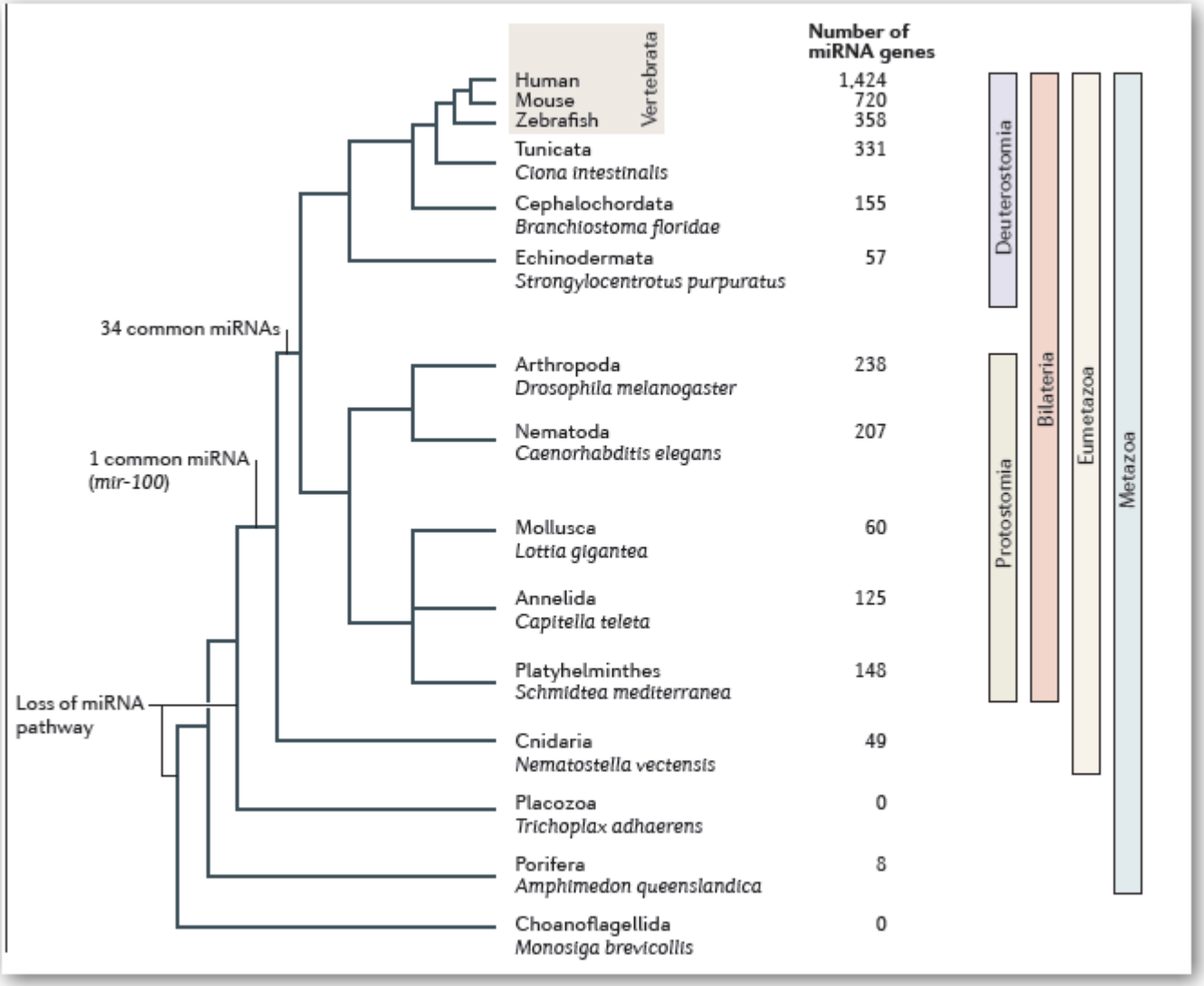


Evolución y conservación

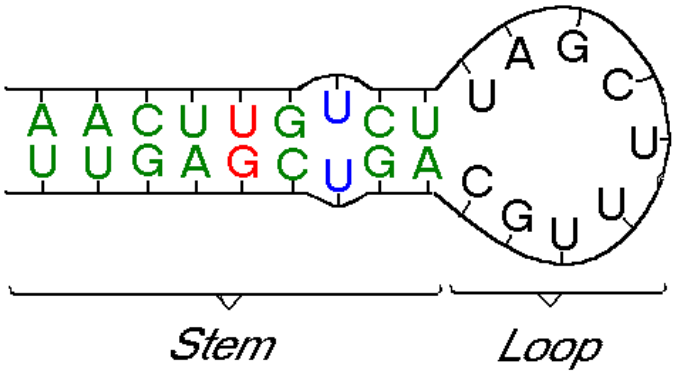
NON-CODING RNA

Evolution of microRNA diversity and regulation in animals

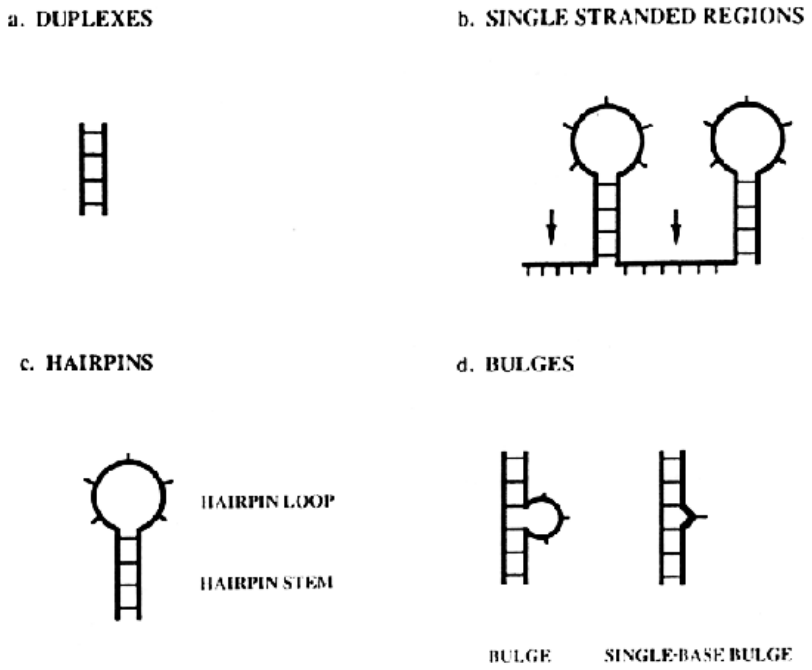
Eugene Berezikov



Estructura secundaria

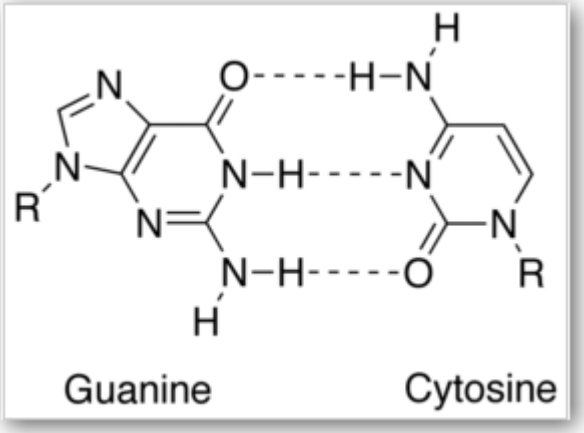


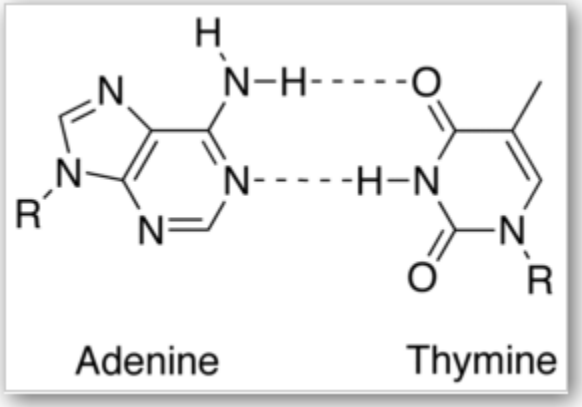
Watson-Crick pairs
UG pairs
Mismatch

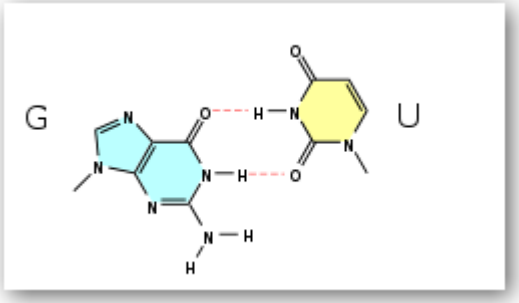


1) Dúplex: región con complementariedad perfecta – todas las bases se emparejan
2) Bulge: en una de las dos hebras hay bases que no tienen pareja
3) Bucles internos: en las dos hebras hay bases que no tienen parejas
4) Hairpin: la estructura secundaria forma una horquilla

Energía de enlace

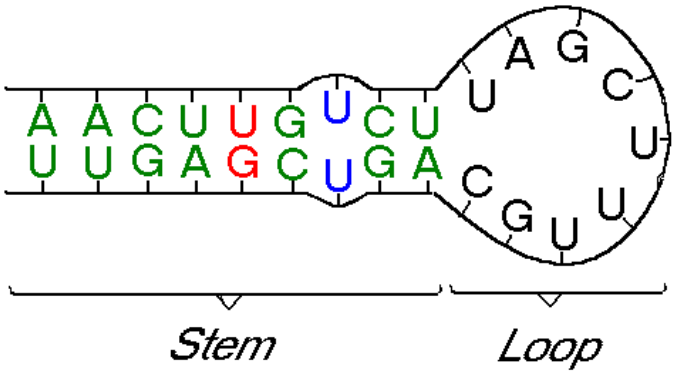




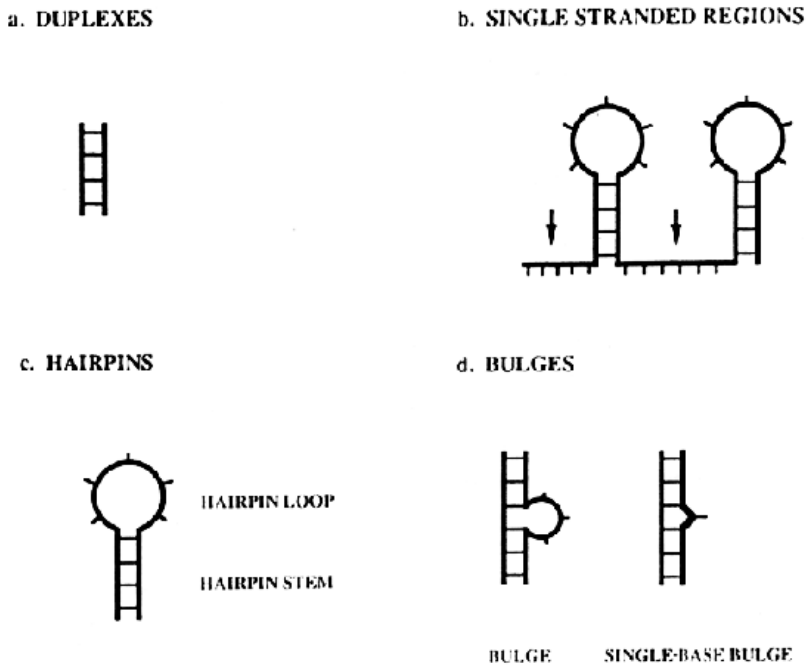


	Stacking Energies for base pairs					
	A/U	**C/G**	**G/C**	**U/A**	**G/U**	**U/G**
A/U	-0.9	-1.8	-2.3	-1.1	-1.1	-0.8
C/G	-1.7	-2.9	-3.4	-2.3	-2.1	-1.4
G/C	-2.1	-2.0	-2.9	-1.8	-1.9	-1.2
U/A	-0.9	-1.7	-2.1	-0.9	-1.0	-0.5
G/U	-0.5	-1.2	-1.4	-0.8	-0.4	-0.2
U/G	-1.0	-1.9	-2.1	-1.1	-1.5	-0.4

Estructura secundaria



Watson-Crick pairs
UG pairs
Mismatch

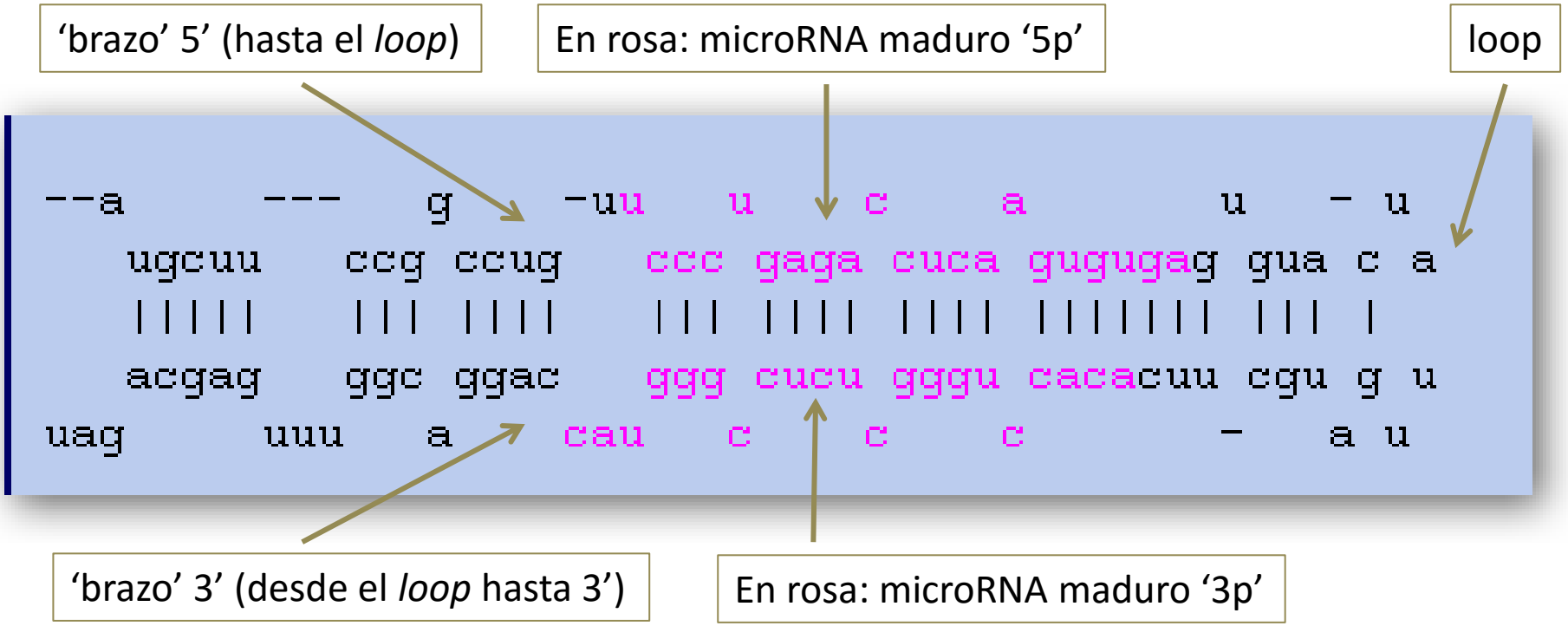


1) Dúplex: región con complementariedad perfecta – todas las bases se emparejan
2) Bulge: en una de las dos hebras hay bases que no tienen pareja
3) Bucles internos: en las dos hebras hay bases que no tienen parejas
4) Hairpin: la estructura secundaria forma una horquilla

Estructura secundaria

El primer microRNA: lin-4

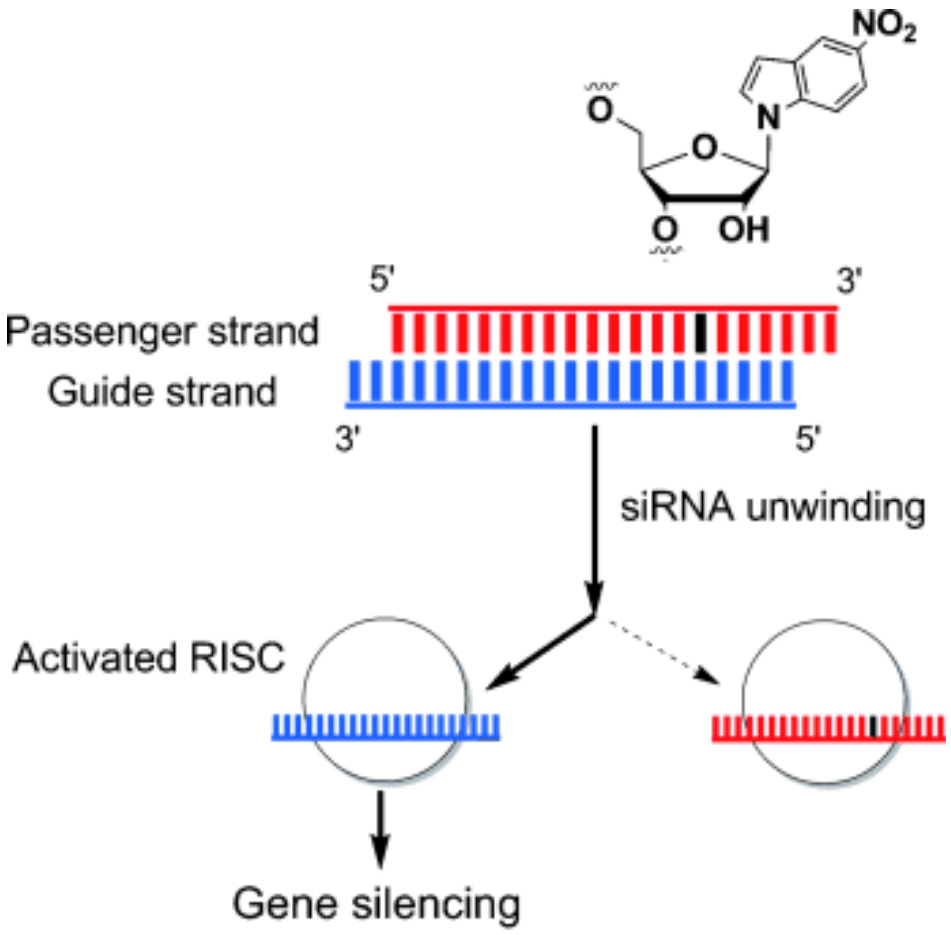
- El pre-microRNA tiene aproximadamente entre 70-100 nt
- La estructura secundaria del pre-microRNA está caracterizada por una 'horquilla'



Nomenclatura

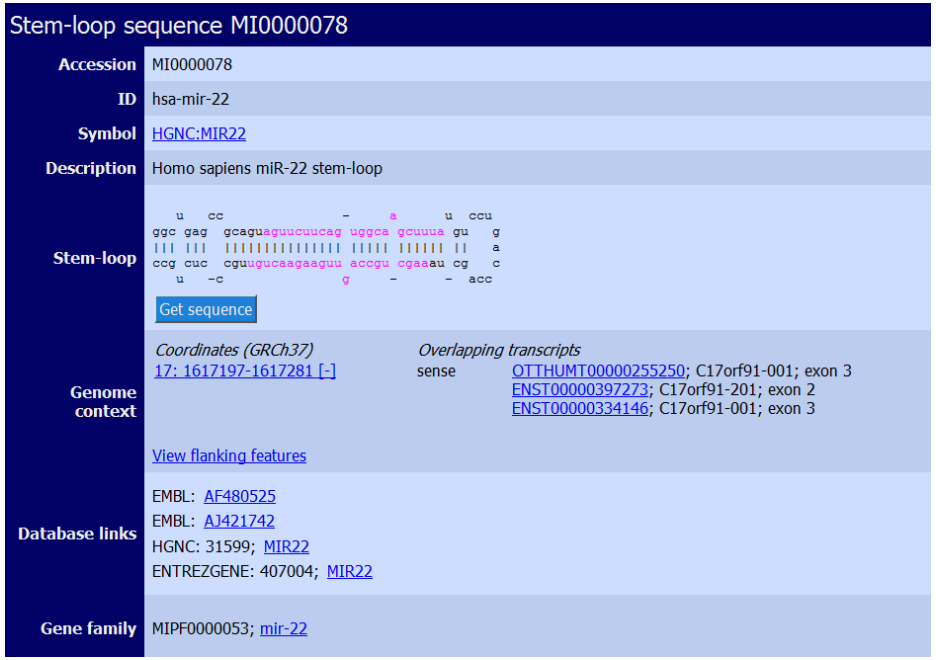
Guide strand: el microRNA maduro funcional (la secuencia que se va a incorporar en RISC)

Passenger strand: la secuencia que normalmente se degrada



miRBase

- miRBase es la base de referencia de los microRNAs: 193 especies, mas de 20,000 secuencias (microRNA maduro y pre-microRNA)
- descarga de datos de secuencia, de las familias conservadas y las coordenadas genómicas
- Permite la búsqueda de una secuencia anónima mediante BLAST



miRBase: nomenclatura

- Los microRNA en miRBase suelen tener nombres compuestos de 4 partes como: mmu-miR-375-5p.
  - Las primeras tres letras indican la especie. Por ejemplo, hsa para humano, mmu para ratón, rno para rata, etc.
  - Un nombre de microRNA en minúscula (hsa-**mir**-22) hace referencia al gen de microRNA o al pre-microRNA. Al microRNA maduro se refiere con '**miR**' (hsa-miR-22-5p)
  - El número que lleva el nombre del microRNA se asigna de forma secuencial.
  - 5p hace referencia al brazo (5p o 3p), 5' en este caso
  - ¡OJO! Antiguamente al microRNA maduro menos frecuente se asignaba un asterisco, por ejemplo hsa-miR-19 (microRNA predominante) y hsa-miR-19*

- Excepciones a estas reglas se mantienen por motivos históricos en las familias let-7 y lin-4
- Mas información acerca de la nomenclatura se puede encontrar en el siguiente articulo: Victor Ambros, Bonnie Bartel, David P. Bartel, Christopher B. Burge, James C. Carrington, Xuemei Chen, Gideon Dreyfuss, Sean R. Eddy, Sam Griffiths-Jones, Mhairi Marshall, Marjori Matzke, Gary Ruvkun, and Thomas Tuschl. **A uniform system for microRNA annotation.** *RNA* 2003 9(3):277-279.

Diversificación de un microRNA maduro: los isomiRs

IsomiRs – the overlooked repertoire in the dynamic microRNAome

Corine T. Neilsen[1,2], Gregory J. Goodall[1,2,3] and Cameron P. Bracken[1,3]

[1] Centre for Cancer Biology, SA Pathology, Frome Road, Adelaide, South Australia, 5000, Australia
[2] School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, South Australia, 5005, Australia
[3] School of Medicine, The University of Adelaide, Adelaide, South Australia, 5005, Australia

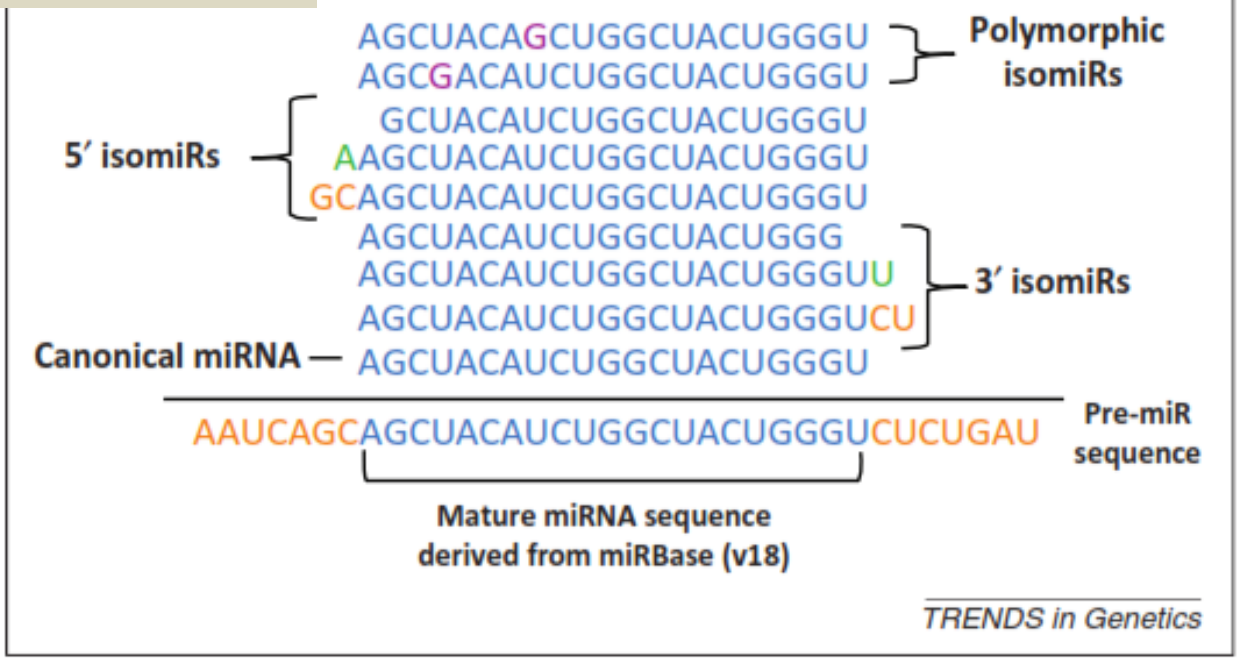
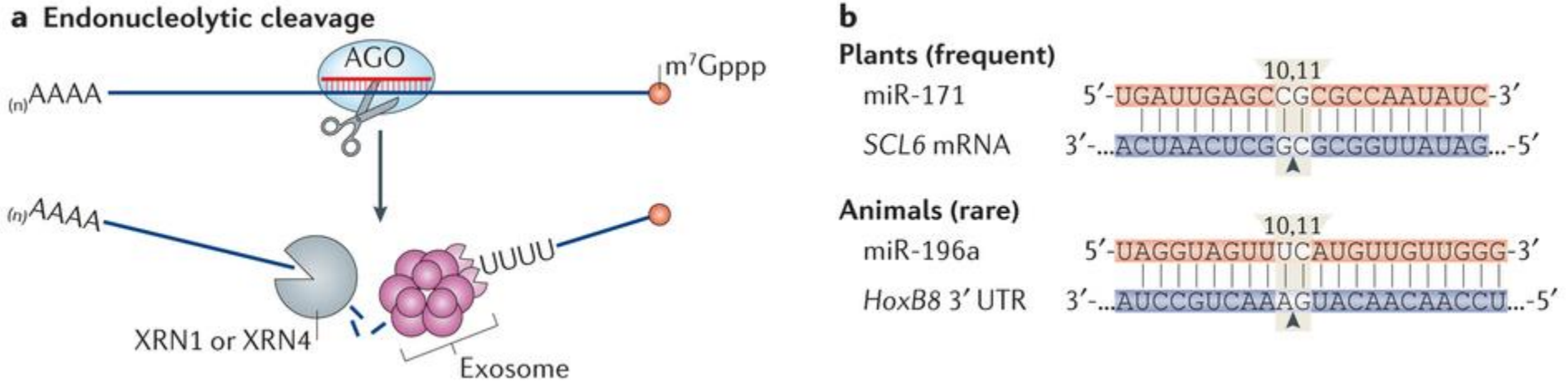


Figure 1. Schematic representation of isomiR species.
The isoforms of human microRNA (miR)-222 are shown as an example of isomiR heterogeneity. The presented sequences are derived from miRBase, the authoritative sequence annotation database for miRNAs. The canonical miR-222 sequence is denoted in blue. The 5′ and 3′ isomiRs are miRNA variants with different sequences at their 5′ and 3′ ends, respectively. The sequence modifications can be either templated (orange) or nontemplated (green). The polymorphic isomiRs harbor distinct nucleotide compositions (purple) within the miRNA sequences.

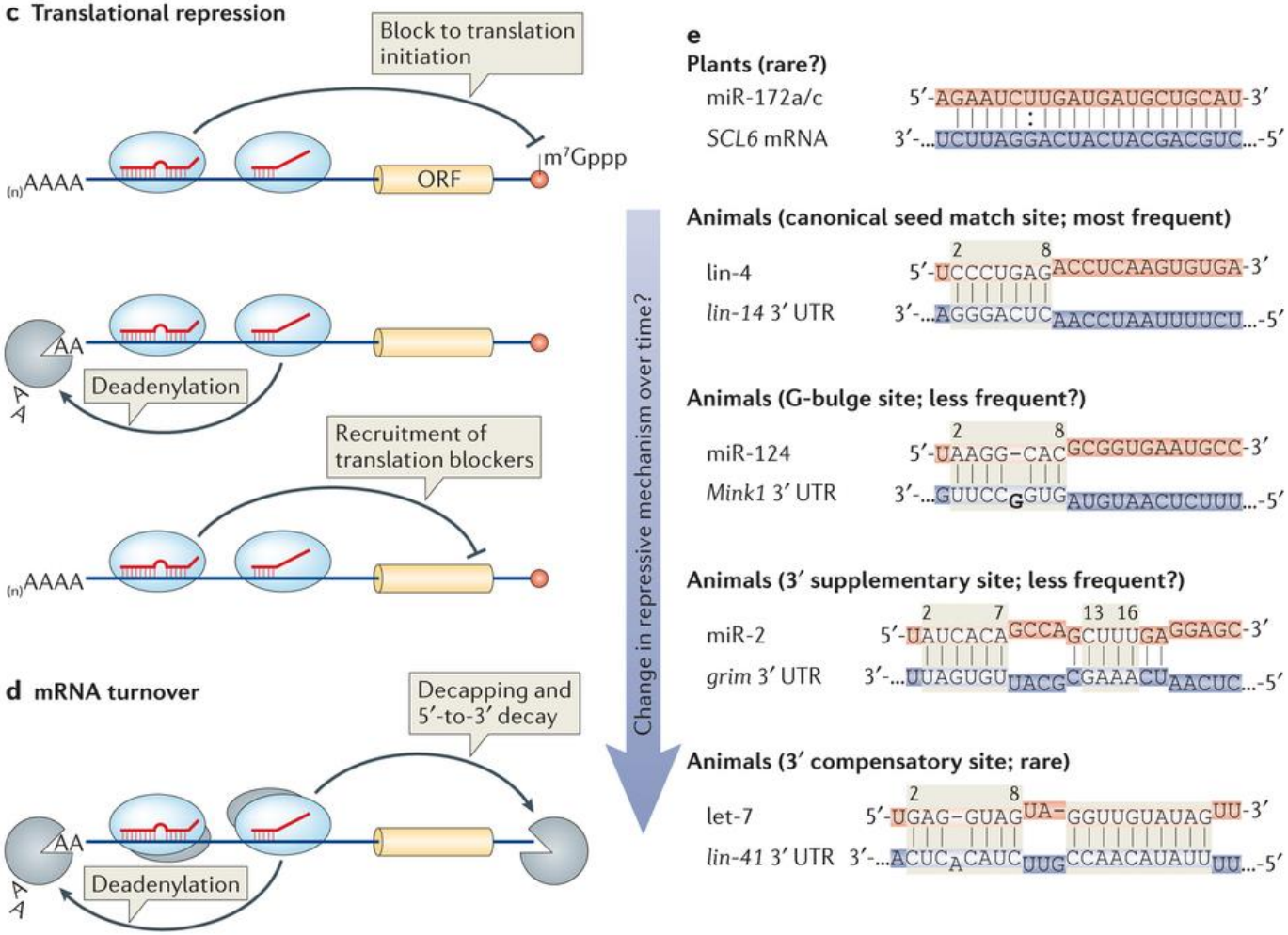
Target sites / dianas

From
Diversifying microRNA sequence and function
Stefan L. Ameres & Phillip D. Zamore
Nature Reviews Molecular Cell Biology **14**, 475–488 (2013) | doi:10.1038/nrm3611

- El microRNA reconoce su diana mediante la complementariedad de secuencia
- En plantas la complementariedad suele ser casi 'perfecta'
- En animales, las primeras 8 bases se llaman *seed* (semilla) y son especialmente importante para reconocer la diana



a | Most plant microRNAs (miRNAs) and a few animal miRNAs direct endonucleolytic cleavage (slicing) of their mRNA targets. The 5'-to-3' exoribonuclease XRN4 in plants and XRN1 in animals, together with the major cellular 3'-to-5' exonucleolytic complex, the exosome, subsequently degrade the sliced mRNA fragments. The 3' end of the 5' cleavage product is frequently uridylated, perhaps to mark these fragments for decay. **b** | miRNA-directed endonucleolytic cleavage of mRNAs requires extensive complementarity between the miRNA and its target site (representative examples in plants and mammals are shown). Endonucleolytic cleavage occurs at the phosphodiester bonds across from nucleotides 10 and 11 of the miRNA, counted from the miRNA 5' end. Although slicing seems to be the dominant mode of action for miRNAs in plants, highly complementary sites in the transcriptome of animals are rare. **c** | In animals, miRNAs were originally proposed to

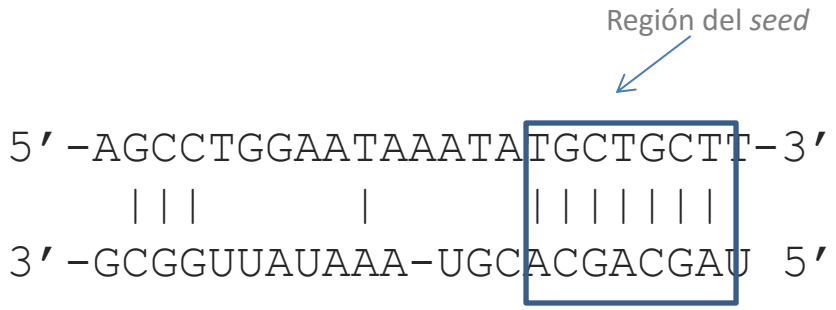


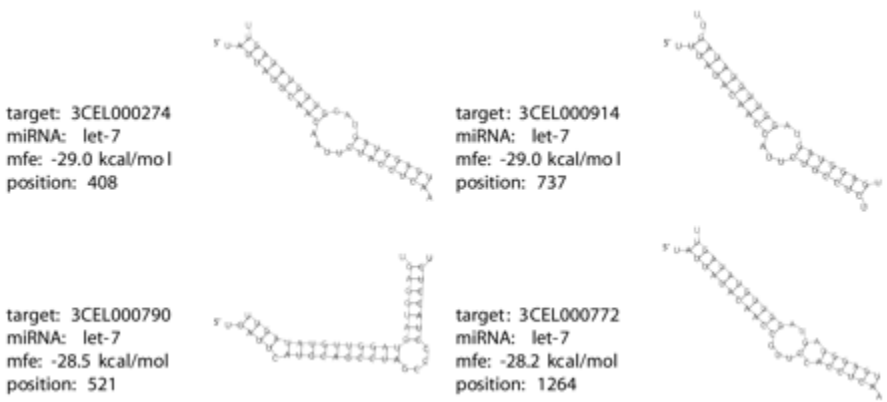
mode of action for miRNAs in plants, highly complementary sites in the transcriptome of animals are rare. **c** | In animals, miRNAs were originally proposed to repress translation of an open reading frame (ORF). Biochemical studies have suggested that miRNAs have a role in blocking translational initiation, in poly(A) tail shortening or in the recruitment of protein cofactors that can interfere with translation. **d** | In many cells and tissues, miRNA-directed translational repression is indistinguishable from mRNA destruction via decapping and 5'-to-3' decay. This has led to the suggestion that miRNAs directly target mRNAs for decay. Another possibility is that the inhibition of translation by miRNAs (part **c**) triggers subsequent mRNA decay, and the temporal delay between these two effects can vary depending on the surveillance mechanisms in place in particular cellular contexts (depicted by blue arrow on the right). **e** | Like animal miRNAs, plant miRNAs may regulate target mRNAs via mechanisms other than endonucleolytic cleavage. But the underlying miRNA–target RNA interaction (for example, miR-172–*SCL6*) is indistinguishable from sites that trigger cleavage (part **b**). In contrast to plants, animal miRNAs are almost always imperfectly complementary to the target mRNAs they regulate. The seed sequence (miRNA nucleotides 2 to 8; sometimes interrupted by a specific G-bulge) is the major determinant for target binding and often suffices to trigger mRNA repression. Additional pairing of miRNA nucleotides 12 to 16 can sometimes bolster seed binding (3' supplementary sites). Only in rare cases can an imperfectly matching seed sequence be compensated for by extensive complementarity between the miRNA 3' region and the target site (3' compensatory sites). Binding of animal miRNAs to partially complementary target sites generally results in translational inhibition and/or mRNA decay via the mechanisms shown in part **c** and part **d**. *HoxB8*, homeobox B8; UTR, untranslated region.

From
Diversifying microRNA sequence and function
Stefan L. Ameres & Phillip D. Zamore
Nature Reviews Molecular Cell Biology **14**, 475–488 (2013) | doi:10.1038/nrm3611

Predicción de dianas

- La predicción se basa en los híbridos que se forman entre el microRNA y el mRNA
- Los híbridos se calculan mediante modelos termodinámicos – *minimum free energy algorithm* (RNAfold, RNAhybrid, etc.)
- Alunas propiedades importantes: energía libre, existencia de una región semilla (*seed region*), numero de desemparejamientos, numero de bucles
- Otras propiedades importantes pero menos entendidos son la "accesibilidad" de la estructura secundaria y la interacción entre varias dianas en la misma 3' UTR
- La predicción de dianas de microRNA se basa fuertemente en la presencia de un "seed" (emparejamiento perfecto de los 7 primeros nucleótidos entre el extremo 5' del microRNA y la región 3' UTR) y la señal filogenética
- Se estima que aprox. El 40% de todas las dianas no tienen "seed" → tienen regiones compensatorias , es decir muchos emparejamientos entre el extremo 3' del microRNA y la UTR 3'
- Algunos de los algoritmos mas usados son: TargetScanS, PicTar, miRanda, RNAhybrid, TargetSpy

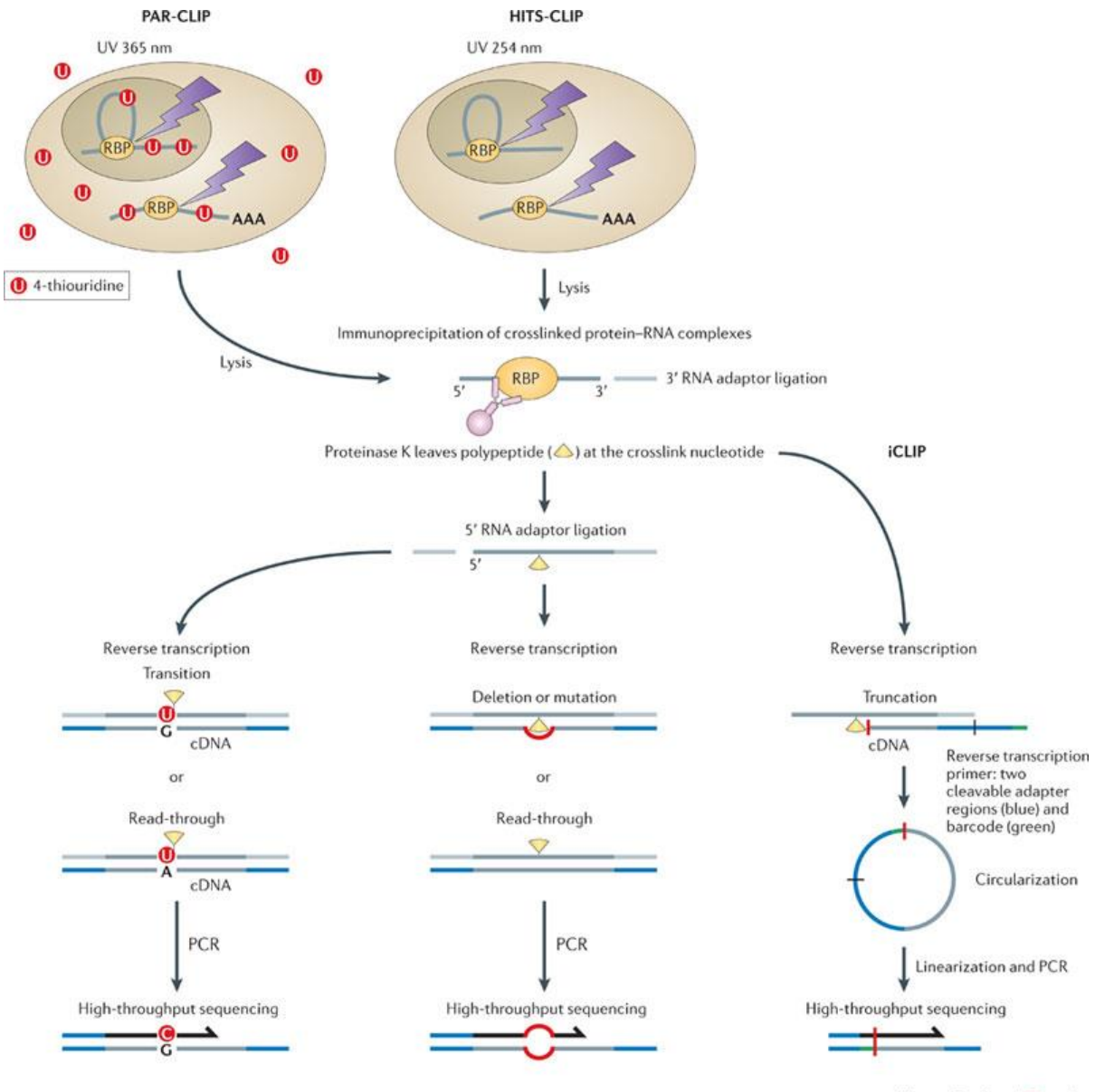




Hibrido entre hsa-miR-16 (abajo) & NM_004178 (arriba) - Posición 249-269 en la 3' UTR

Detectar la interacción miRNA/mRNA: CLIP

CLIP:
crosslinking and immunoprecipitation

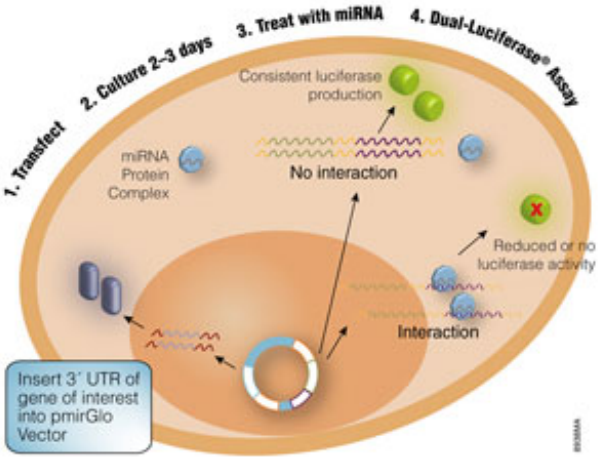


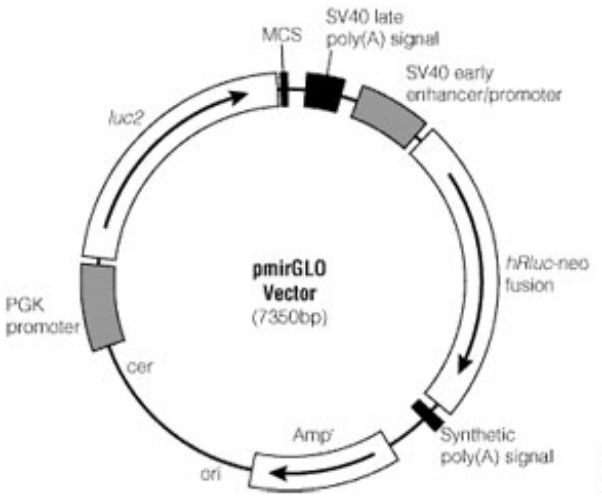
Nature Reviews | Genetics

Reporter assay

http://www.promega.es/products/pm/applications-of-reporter-gene-assays/analyzing-mirna-targets/

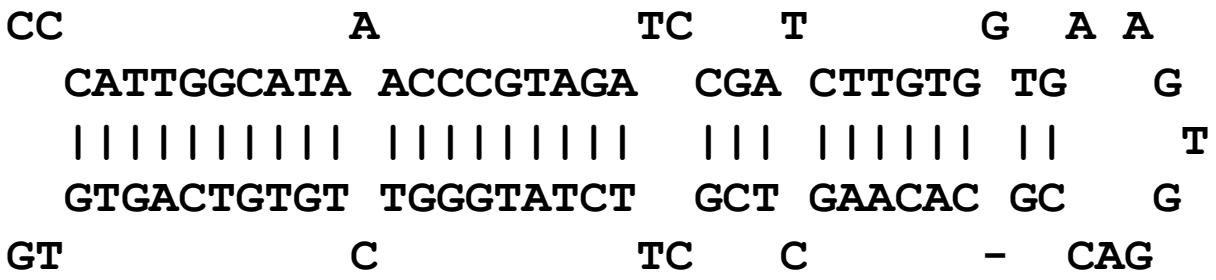
- MicroRNAs (miRNAs) are short RNAs that interact with targets in the 3' UTR of transcripts and result in either mRNA degradation or inhibition of translation.
- Observing miRNA-mediated effects requires a reporter under the control of a weaker promoter so that subtle changes in gene expression can be observed.
- The pmirGLO Dual-Luciferase miRNA Target Expression Vector contains a 3' UTR cloned downstream of the *luc2* firefly luciferase gene under the control of the human phosphoglycerate kinase (PGK) promoter.
  - The PGK is a nonviral universal promoter; can be expressed in yeast, rat, mouse and human cells.
  - The pmirGLO Vector can be used to create Neomycin-resistant stable cell lines.
  - Either the Dual-Luciferase® Reporter Assay System or the Dual-Glo® Luciferase Assay can be used to obtain data.
- The psiCHECK™-2 Vector, designed for target knockdown experiments, has been used to analyze miRNA targets.
  - 3' UTR of interest is cloned into the MCS of psiCHECK™-2 Vector.
  - *Renilla* luciferase under control of the strong SV40 promoter is the reporter.
  - Strong promoter may mask subtle changes in gene expression.





The pmirGLO Vector (Cat. # E1330) has a multiple cloning site downstream of the *luc2* stop codon for insertion of the 3'UTR of interest.

Novel microRNAs: pre-NGS



Prediction based on secondary structure properties

- Minimum Free Energy
- Number of bindings
- Loop Length
- Number of bulges, bulbs
- Compositional features (G+C content, dinucleotide frequencies)

In the human genome exist approx. 11 million hairpin structures
→ High rate of false positive predictions

Predicción de genes

Salida del programa CID-miRNA:

Con los parámetros por defecto, el programa predice 42 genes de microRNA en una secuencia de aprox. 100 kb

(Computational Identification of micro RNA)

An miRNA scanning Tool by:
Center for Computational Biology and Bioinformatics
Jawaharlal Nehru University
New Delhi 110067
E-mails: alok0200@mail.jnu.ac.in, st0645@students.jnu.ac.in
Telephone: +91-11-26717630

CIDmiRNA scan is through.

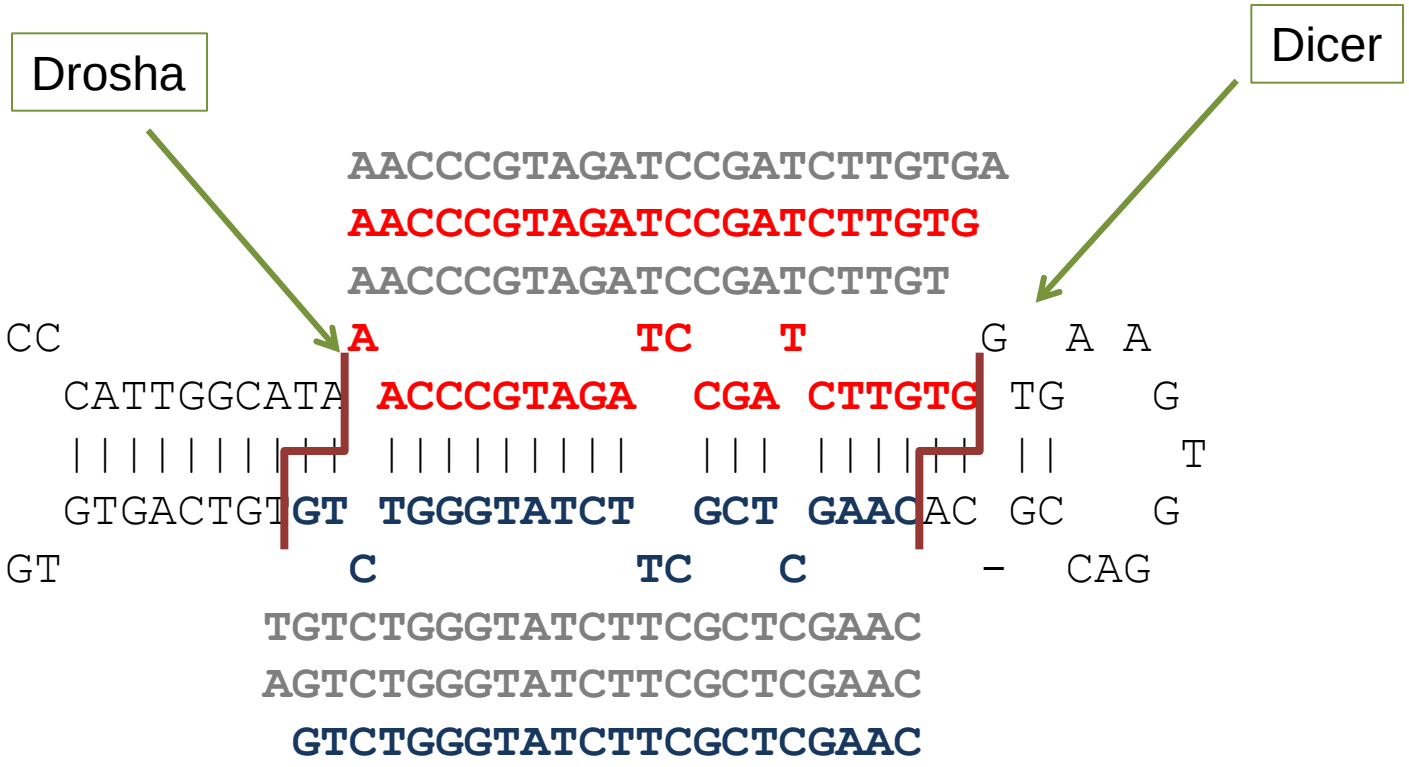
Total precursor miRNAs found: **42**.

FASTA format (sequences only) FASTA format (structures only) FASTA format (both sequences and structures)

S.No.	Position	Length	Normalised Grammar Score	Structural Score
1	1919	76	-0.607686	25
	GAGAATGCAAGGCATTTGGGAGCTGTGTAGCCAGAATCTGGACTTTTAATAGGTCACCAAATGCTCTTGCTTCCTC			
	aau - - g gu-- - g gag gcaag gcauuugg ga cugu ag cca a ||| |||| ||||||| || |||| || ||| a cuc cguuc cguaaacc cu gaua uc ggu u cuu u a g auuu a c			
2	3037	62	-0.569458	41
	GCATGCATGTGTGTGTGTGTGTGTGTGTGTGTGCATGCATGCATACATGTTGGGATGC			
	---- ug gu gcau gcaugugugugugugug u g |||| |||||||||||||| | | cgua uguacauacguacguacgu g u gggu gu ug			
3	3684	94	-0.603886	25
	GTTATTGACTTACATTTCCTGGAGAATAAGAGCACTATAGGCATGAAGACATTTTCTTTTCTCTTGTGCATTTTGCCCAGGAAAGAATAATAAC			
	ac aca a aau a uau cau c guuauug uu uuuccugg g aag gcac agg gaaga a |||||| || ||||||| | ||| |||| ||| |||| u caauaau aa aaaggacc c uuu cgug ucu uuucu u -- g-- - gu- a u-- cu- u			
4	4953	62	-0.564125	53
	GTATATGCACACATGTGTGTGTATCTATGTGTATCTATGTGTATGCATGTGTGTGTGTGT			
	g ug uc u uauaugcacacaugugug ugua ua g |||||||||||||||| ||| || gugugugugugguacguau guau au u u gu cu g			
5	6095	98	-0.580363	82
	GTGTATGCGTGTACTCATGTGTGTGTGTGCACACGCTTACGTGCACGTGTGTGTGTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTACAC			
	ug u ug ug ug u ug u guguaugcg uac caug ug ug cacacgc uacg cacg g ||||||||| ||| |||| || || ||||||| ||| |||| cacaugugu gug gugu gu gu gugugug gugc gugu u gu u gu gu gu u gu g			

La secuenciación masiva brinda nuevas posibilidades a la predicción de microRNAs ya que reduce drásticamente el número de candidatos (secuencias transcritas)

With NGS



- Drosha/Dicer (DCL in plants) processing patterns can be detected
- Both mature microRNAs (both arms) are represented in the sample?
- 5’ end of the mature microRNA shows less fluctuation
- Virtually all reads are organized in one of two clusters

Detect a microRNA gene

NON-CODING RNA

Evolution of microRNA diversity and regulation in animals

Eugene Berezikov

- **Hairpin structure**
- **2nt 3' *overhang* between 5p (red) and 3p (blue) arm**
- **'Little' fluctuation at 5'**

Box 1 | **Identification of miRNA genes**