

Incontro con bioinformatici

Giuseppe Macino

Universita' di Roma "La Sapienza"

Quanto DNA e' contenuto nei genomi di

Amoeba dubia 670 miliardi c.b



Zea mays 4 miliardi c.b.



Homo sapiens 2,9 miliardi c.b



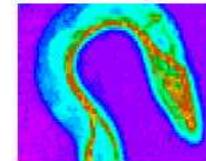
Arabidopsis thaliana 125 milioni c.b.



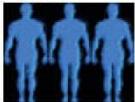
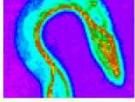
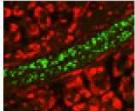
Drosophila melanogaster 120 milioni c.b.



Caenorhabditis elegans 97 milioni c.b.



Quanti geni ci sono nelle varie specie?

Specie	Dimensioni del genoma	Numero di geni
Uomo 	2,9 miliardi paia di basi	30.000
Drosophila melanogaster 	120 milioni paia di basi	13.601
Saccaromyces cerevisiae 	12 milioni paia di basi	6.275
Caenorhabditis elegans 	97 milioni paia di basi	19.000
E.coli 	4.1 milioni paia di basi	4.800
Arabidopsis thaliana 	125 milioni paia di basi	25.000

Percentuali di geni sul genoma totale

Amoeba	0.001%
Homo sapiens	2%
Zea maize	1%
Arabidopsis	80%
Drosophila	50%
Nematode C.elegans	85%
Lievito del pane	70%
Echerichia coli	85%

Da cosa sono composte le
sequenze che costituiscono i
genomi?

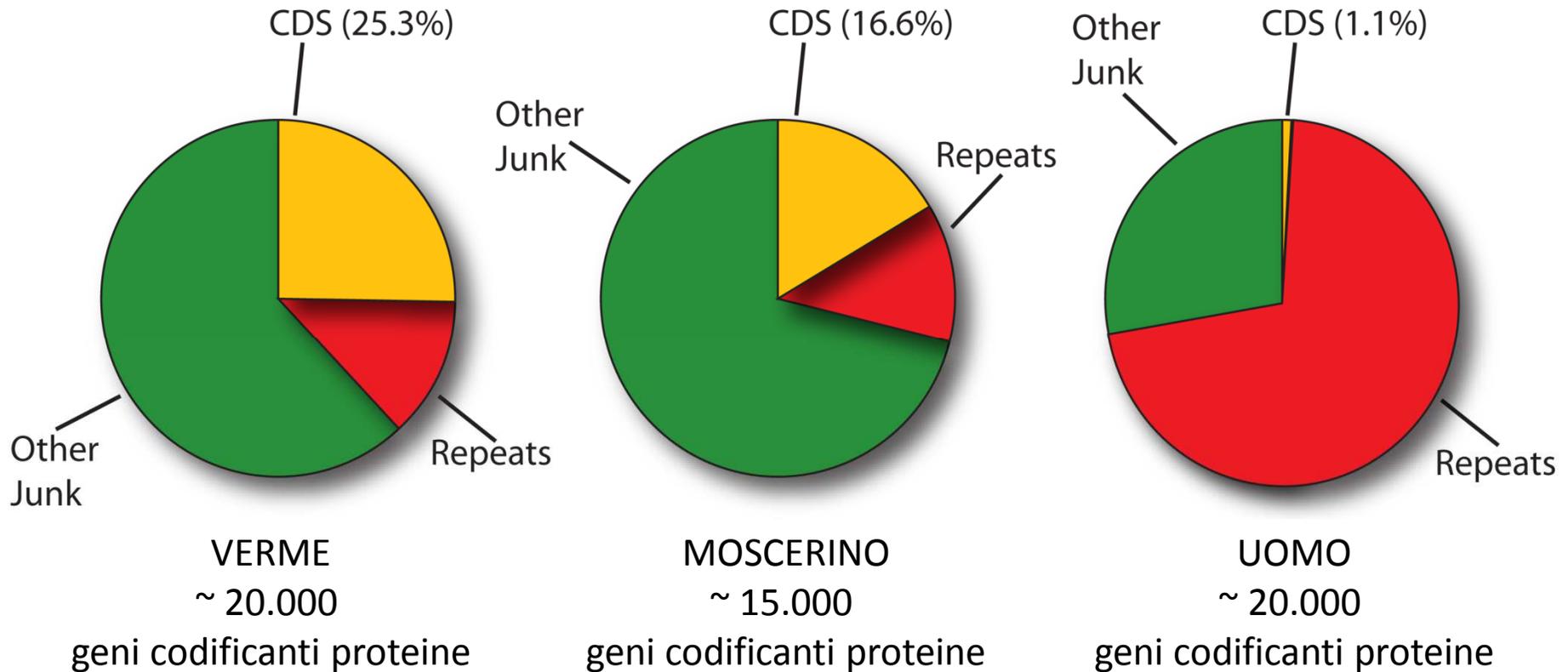
GENI, duplicazioni geniche

SEQUENZE RIPETUTE fino a milioni di volte

TRASPOSONI a DNA in migliaia di copie

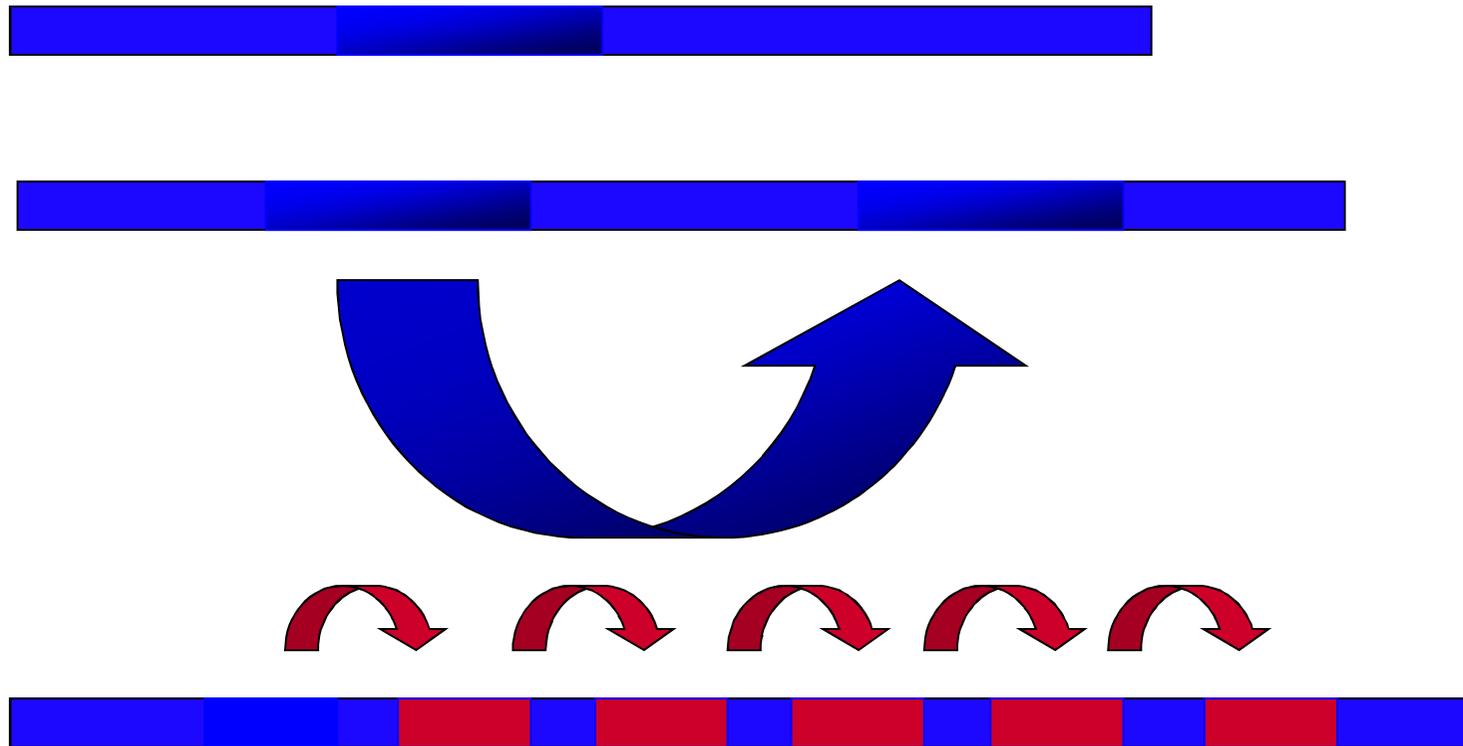
RETRO TRASPOSONI AD RNA in centinaia di
migliaia di copie

Siamo pieni di (DNA) spazzatura

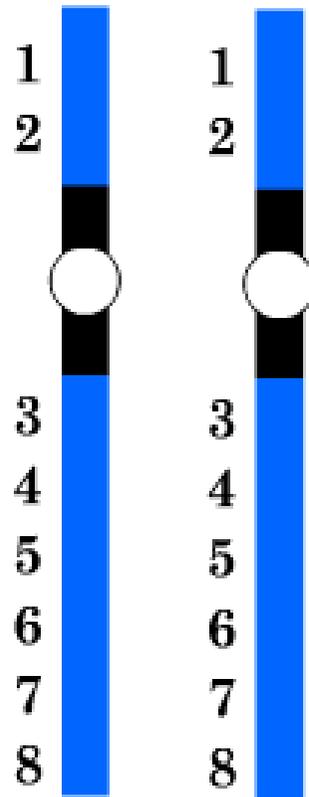


Con quali meccanismi avviene
l'amplificazione?

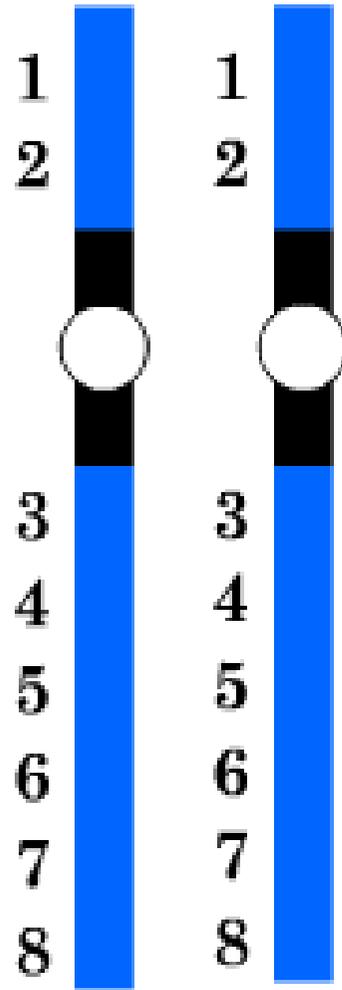
Come gli elementi mobili (trasposoni)
si integrano ed espandono

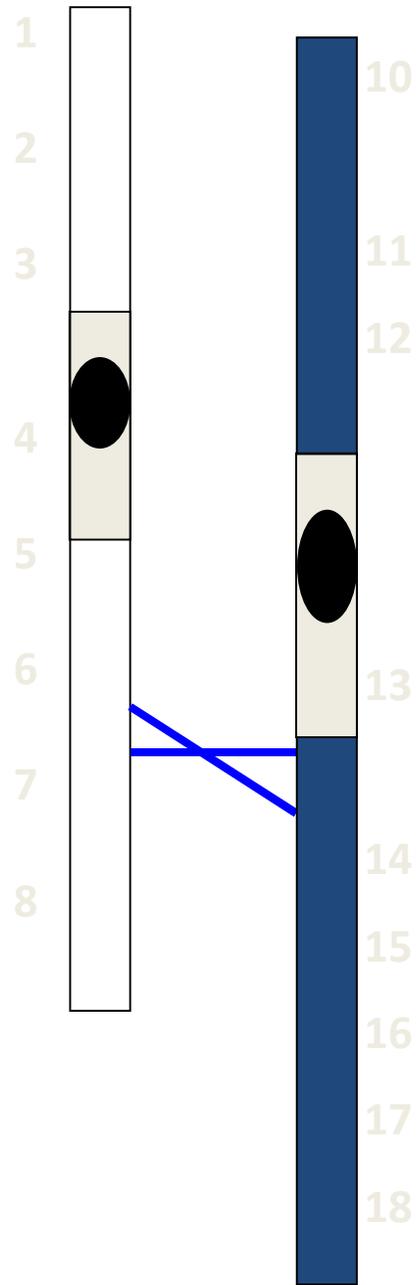


Duplicazione di tratti di cromosomi

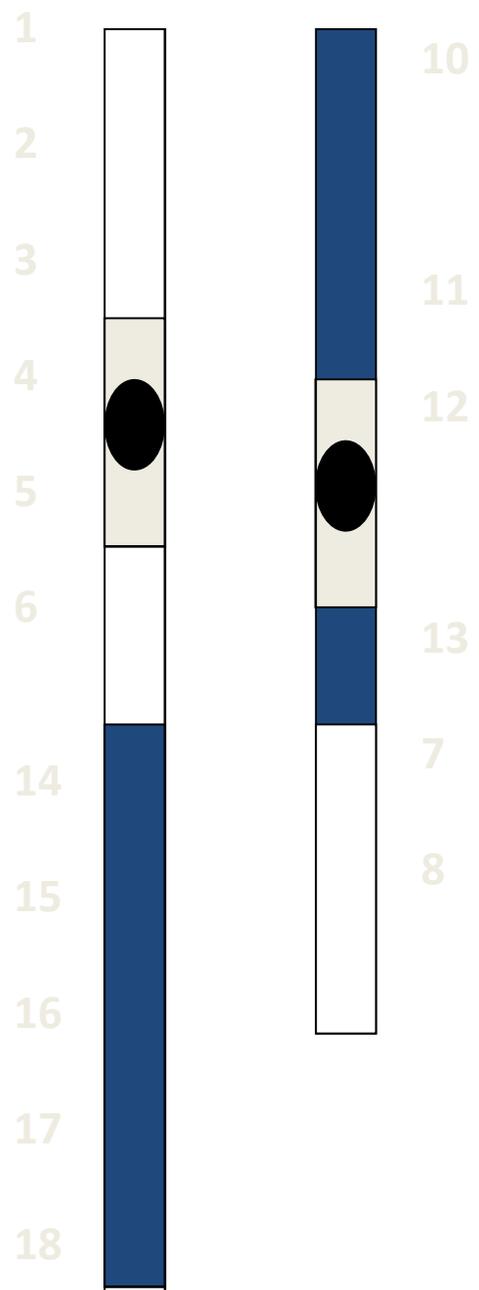


Inversioni cromosomiche

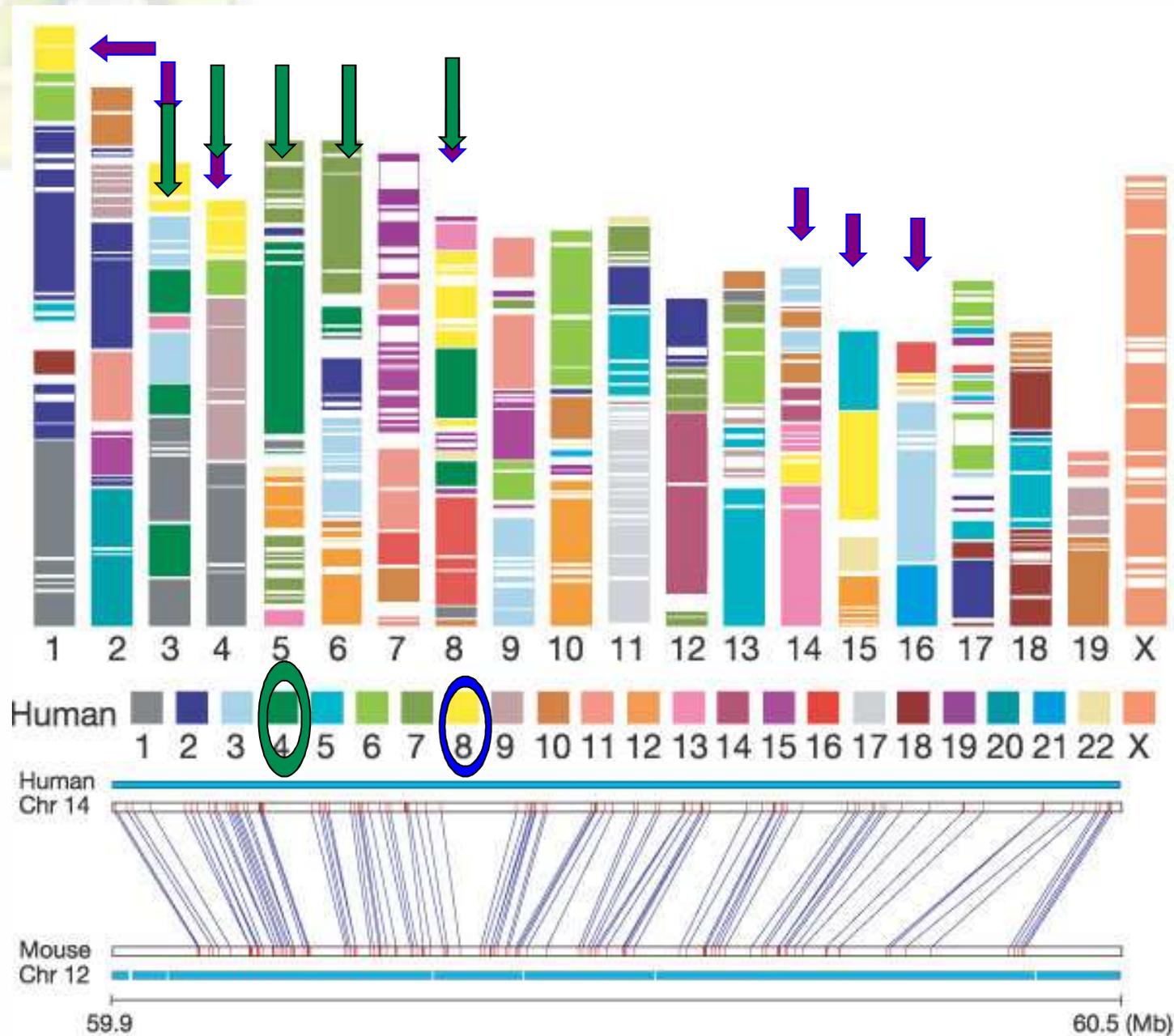




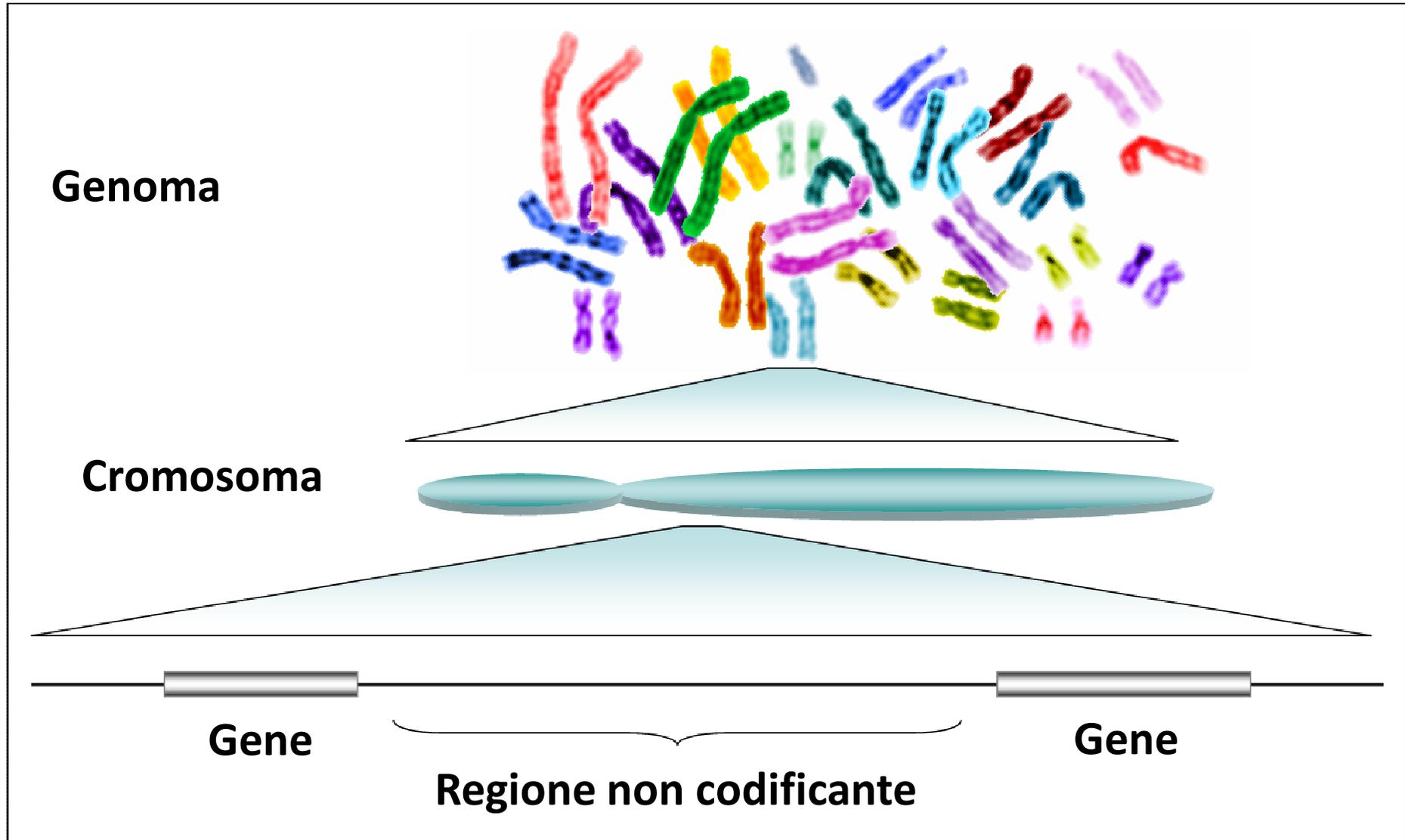
Crossing over ineguale



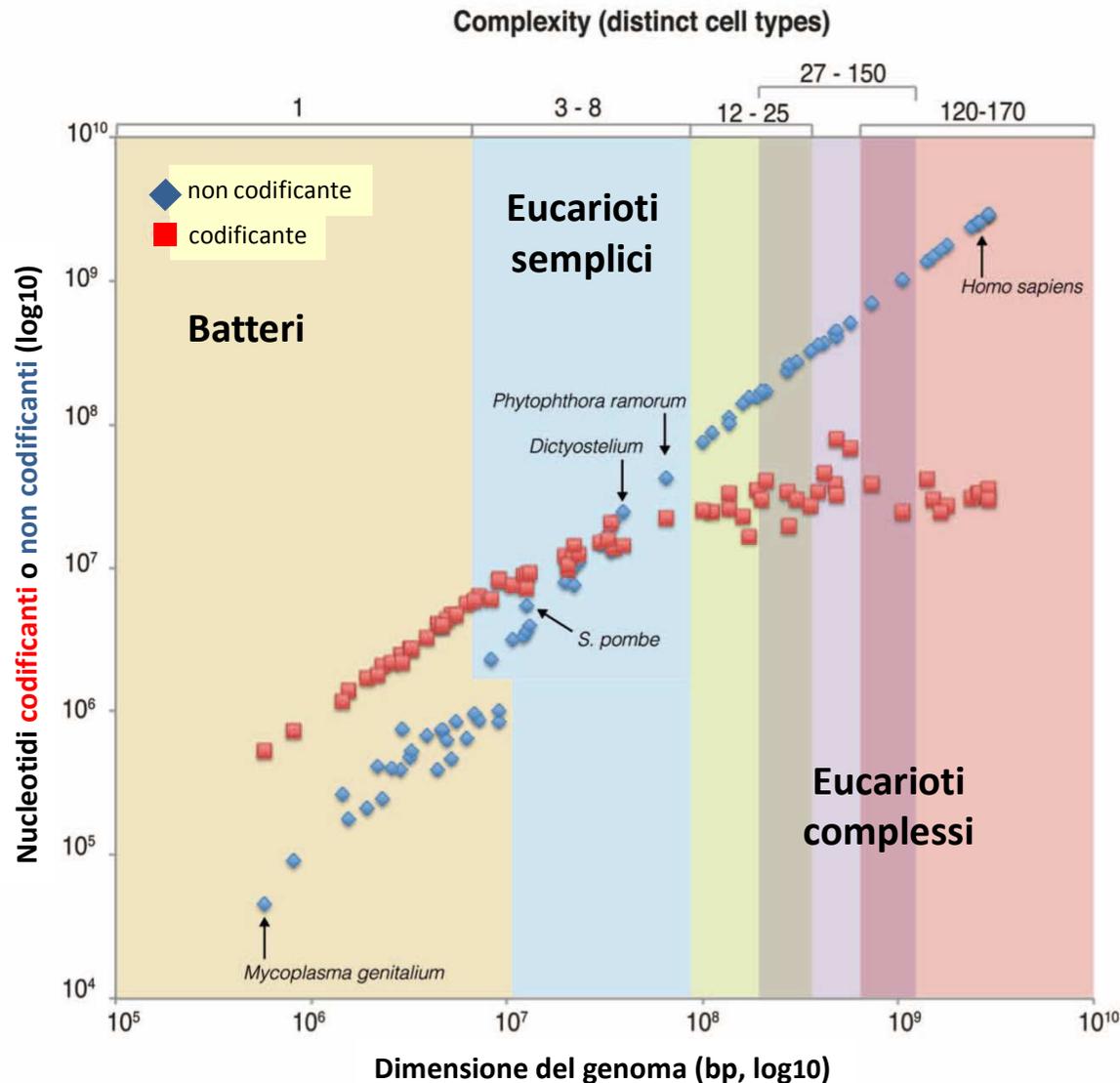
Sintenia tra uomo e topo



Il genoma comprende sia i geni che codificano per le proteine che le sequenze non codificanti.



La componente non codificante aumenta esponenzialmente con la complessità



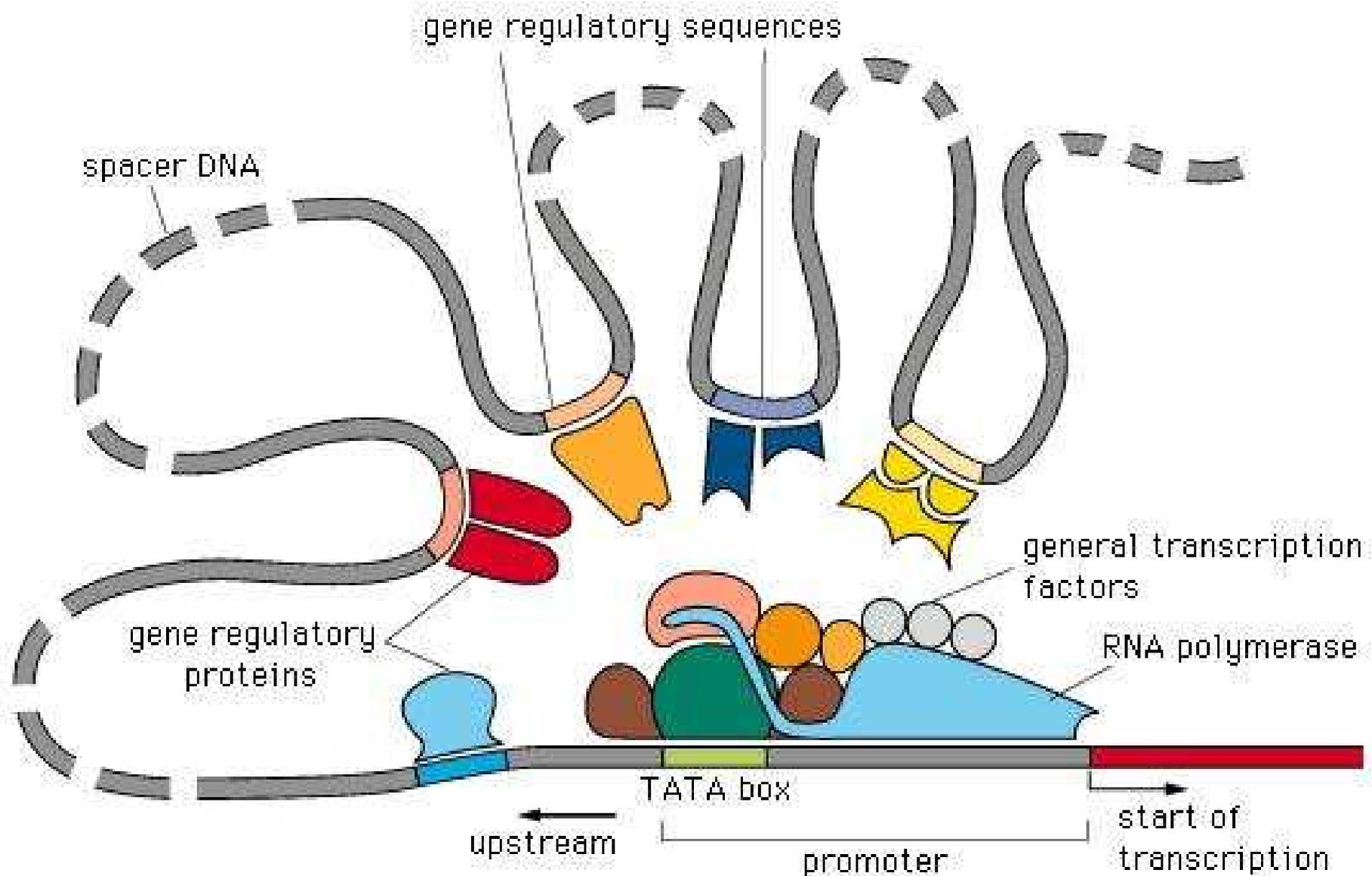
Ruolo fisiologico del genoma ripetuto

- ★ principale sede delle modificazioni epigenetiche
- ★ integrità strutturale del genoma (centromeri, telomeri)
- ★ fonte di informazioni regolatrici
(insulators, promotori alt, esoni alt, siti polyA alt)
- ★ altamente trascritto (più di rRNA, tessuto-specifico)
- ★ impalcatura di RNA della cromatina interfascica
- ★ lncRNAs regolatori generati da specifici-loci

Ruolo patologico del genoma ripetuto

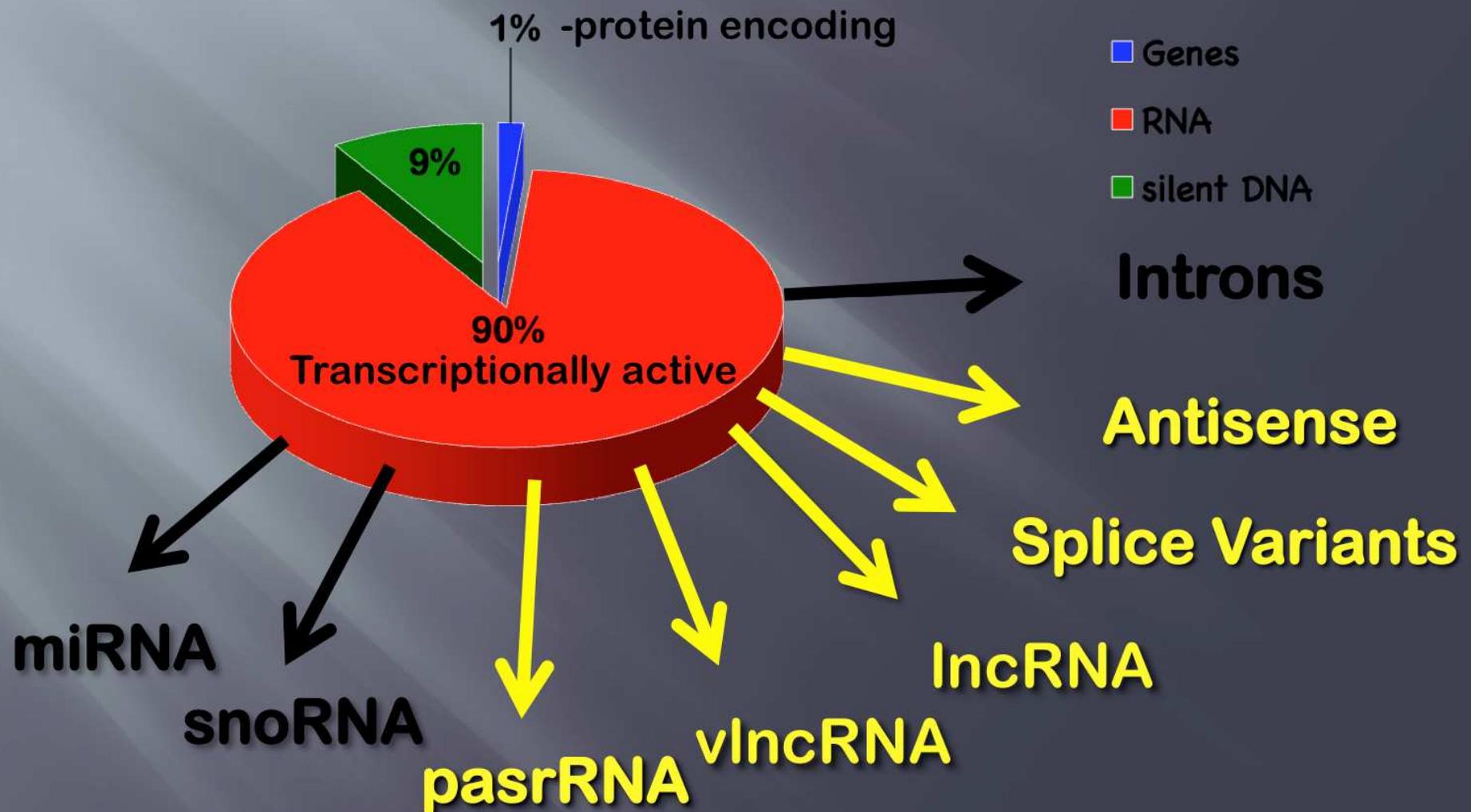
- ✦ mutagenesi inserzionale dovuta ad attiva trasposizione
- ✦ instabilità meiotica associata a malattie genetiche
- ✦ espressione aberrante legata a riarrangiamenti cromosomici nel cancro

Quello che sapevamo sulla trascrizione



Il genoma è quasi completamente trascritto su tutte e due le eliche

Non-coding Genes Account for Most Transcription from the Genome:



Gli RNA sono implicati nelle regolazioni

- Gli RNA che conoscevamo: rRNA, mRNA, tRNA
- I nuovi RNA: Short non coding, Long non coding
- Short: micro RNA (20-24 nucl.), siRNA, piRNA, swiRNA etc.
- Long: circa 25000 diversi nell'uomo (da 200 a 20.000 nucl.)
- Circular RNA: circa 8000 nell'uomo

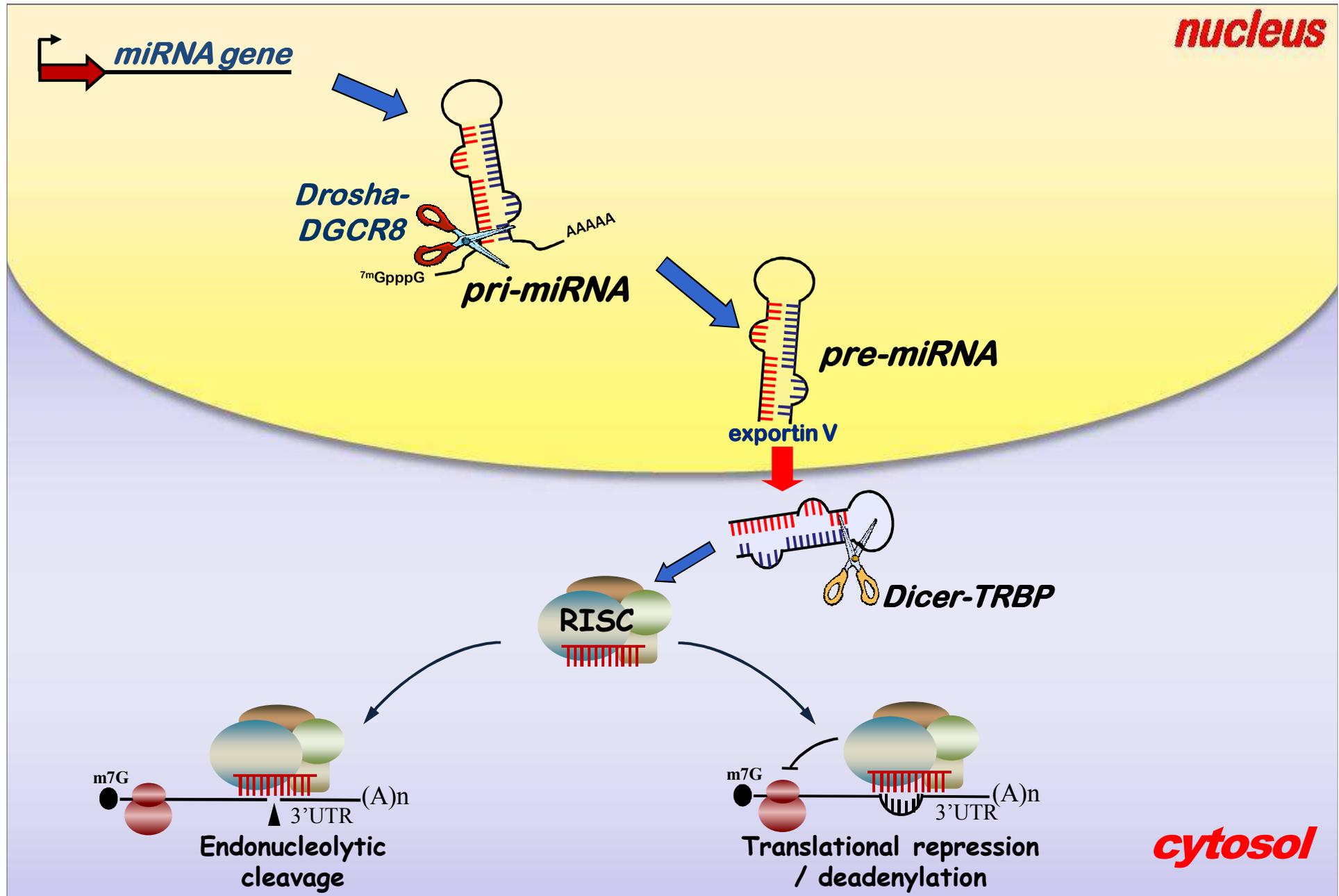
Gli RNA sono implicati nelle regolazioni

- Regolano la struttura della cromatina
- Regolano la trascrizione dei geni
- Regolano la maturazione (splicing) dei trascritti
- Regolano la stabilita' dei trascritti
- Regolano la traduzione
- Regolano la localizzazione spaziale delle proteine e degli RNA
- etc.

miRNAs

- miRNA lin4 e let7 sono stati scoperti la prima volta nei nematodi
- I miRNA sono conservati evolutivamente dai vermi all'uomo
- Il loro ruolo è quello di regolare la stabilità dei trascritti e di regolare la loro traduzione
Circa 200 mRNA ogni miRNA
- Funzionano quindi come repressori (modulatori)

MicroRNA biogenesis



I geni per MicroRNA sono la famiglia genica
piu' abbondante

Da 700 nei vermi a 3000 nell'uomo

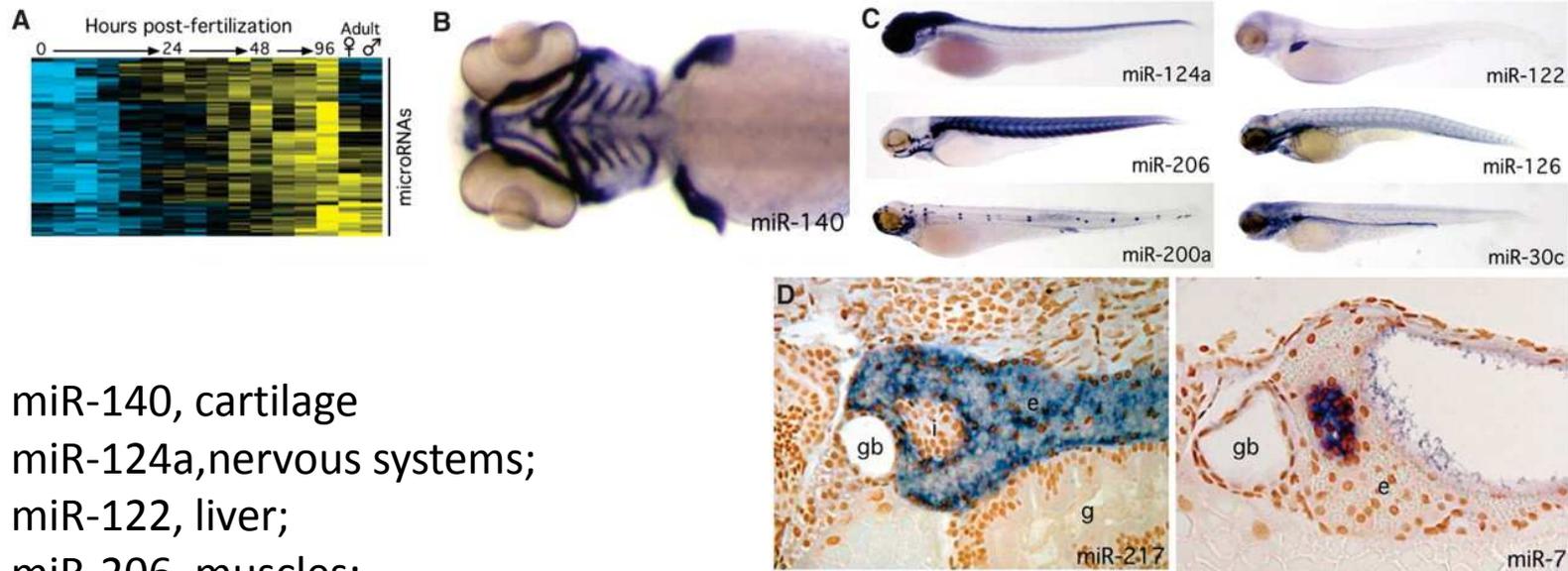
Quanti geni sono regolati dai miRNA?

- Ci sono mediamente 200 mRNA bersaglio per ogni miRNA con punte fino a 1500 bersagli (let 7)
- Almeno il 80% dei geni e' sotto regolazione post trascrizionale

Funzione dei miRNA

- Controllo della proliferazione cellulare
- Controllo della apoptosi
- Sviluppo dei Neuroni
- Differenziamento ematopoietico
- Controllo dello sviluppo negli animali
- Controllo dello sviluppo dei fiori e delle foglie nelle piante

miR expression during development in zebrafish



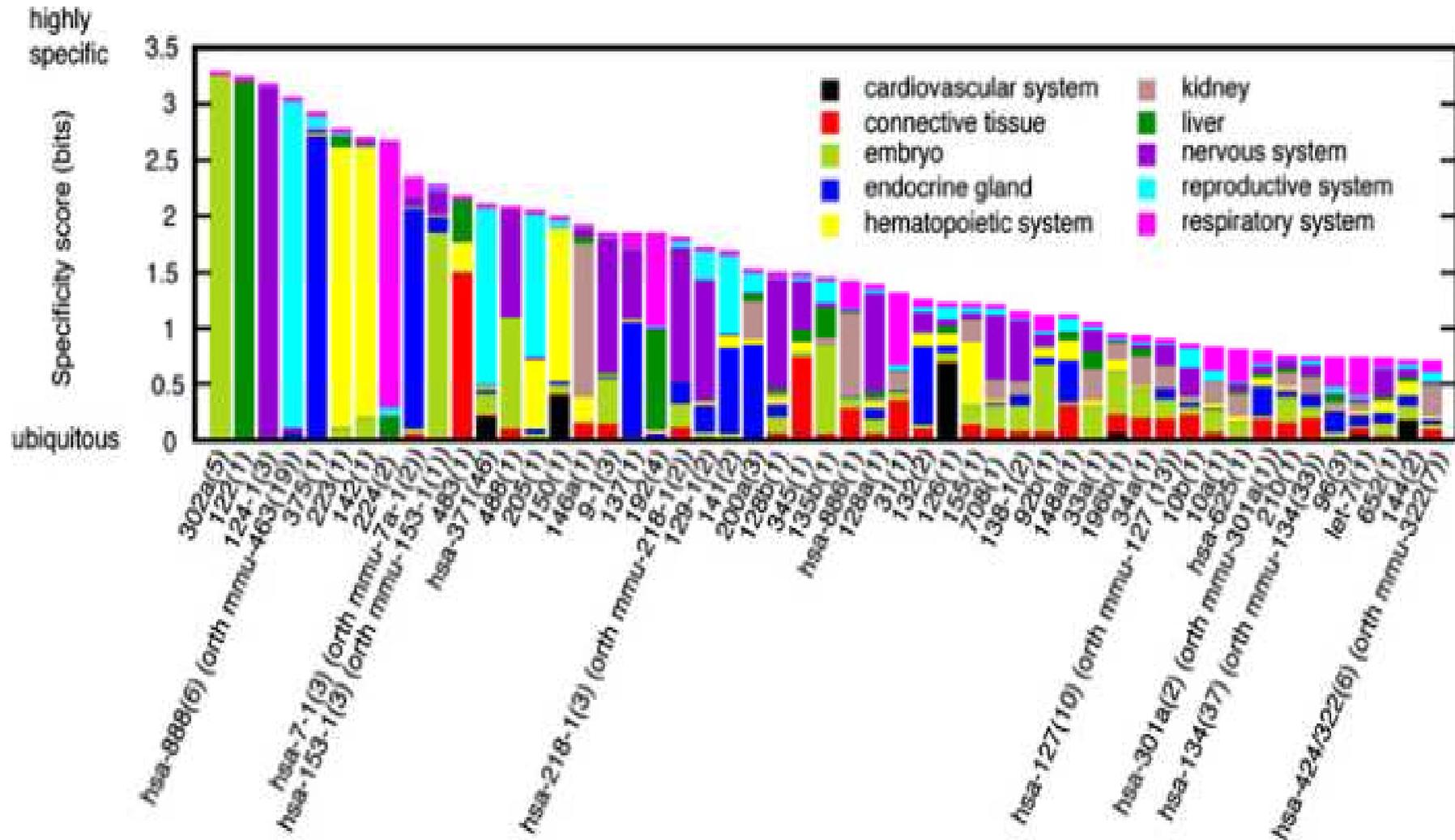
miR-140, cartilage
miR-124a, nervous systems;
miR-122, liver;
miR-206, muscles;
miR-126, blood vessels and heart;
miR-200a, lateral line system and sensory organs;
miR-217, miR-7 in pancreas

Localizzazione della espressione dei miRNA

- Stadi specifici dello sviluppo embrionale
- Mir-1 e' principalmente presente nel cuore di mammiferi
- Mir-122 principalmente nel fegato
- Mir-223 principalmente nei granulociti e nei macrofagi
- Mir-90 Mir-295 preferenzialmente nelle cellule staminali e non nelle cellule differenziate
- Diversi Mir sono espressi in regioni diverse del cervello durante lo sviluppo e nell'adulto

Where are miRNA expressed?

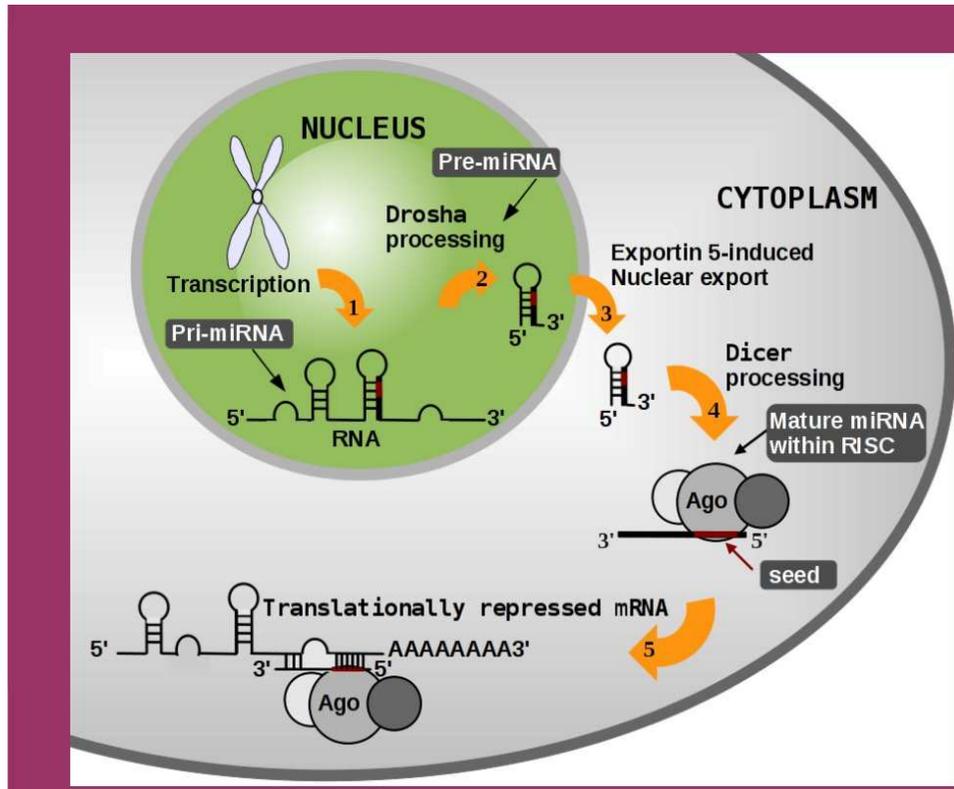
A



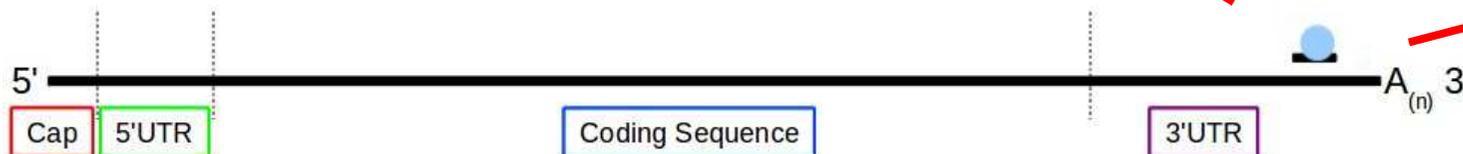
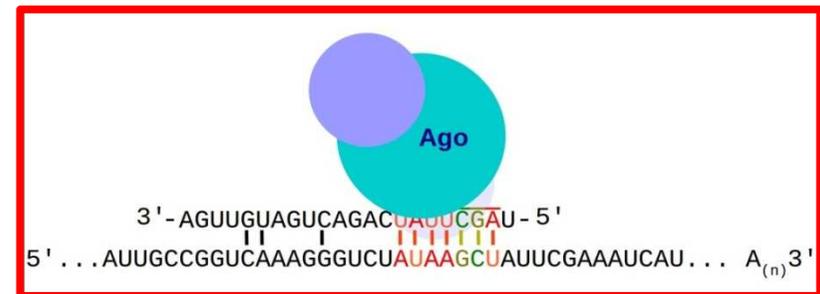
miRNAs: Definition and Biogenesis

5' - UAGCUUAU CAGACUGAUGUUGA - 3'

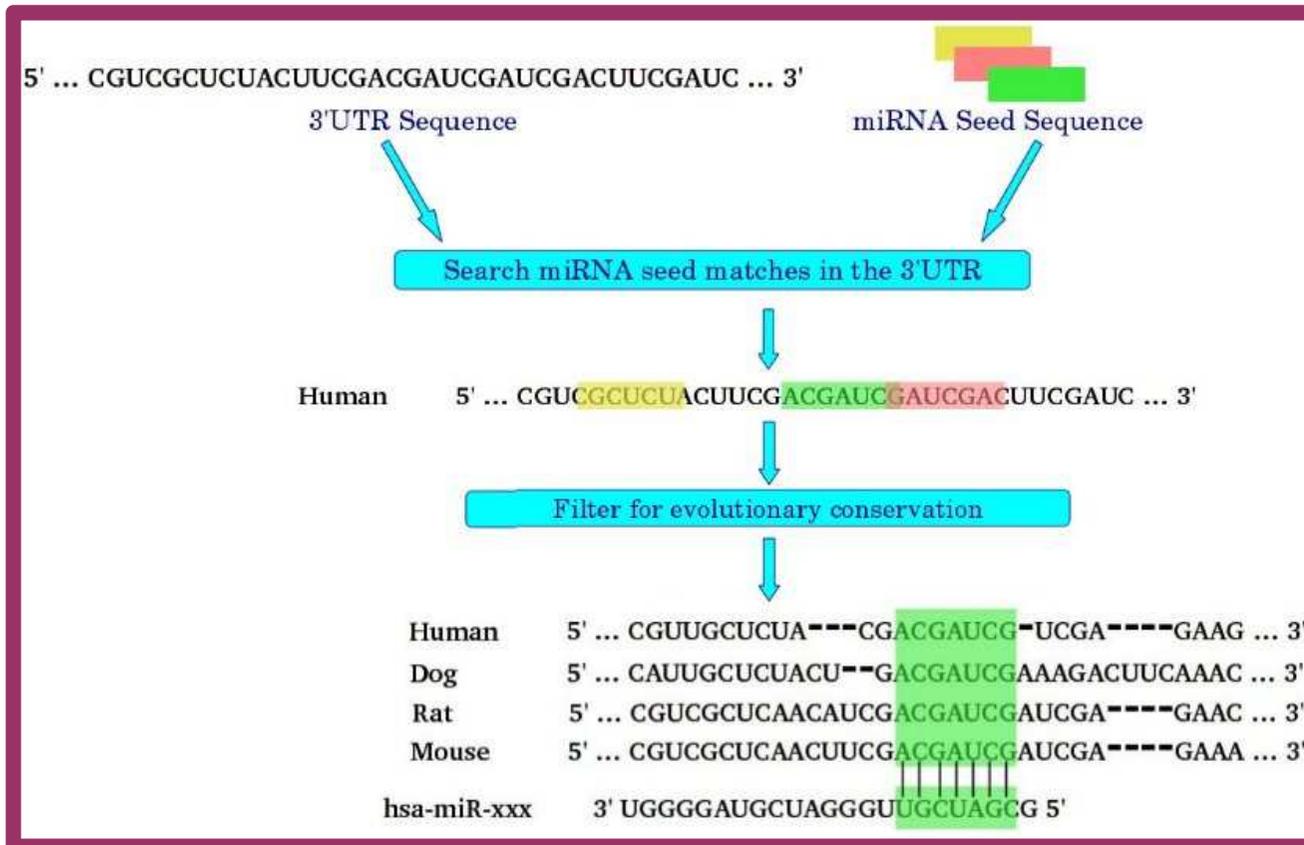
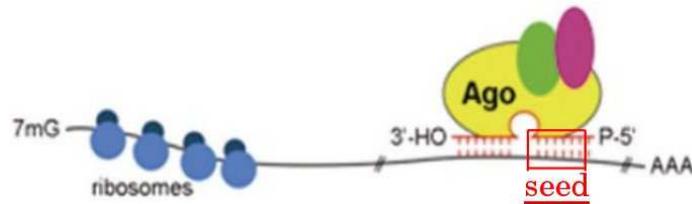
seed



miRNAs are ~22nt single-stranded RNAs that negatively regulate gene expression at the post-transcriptional level



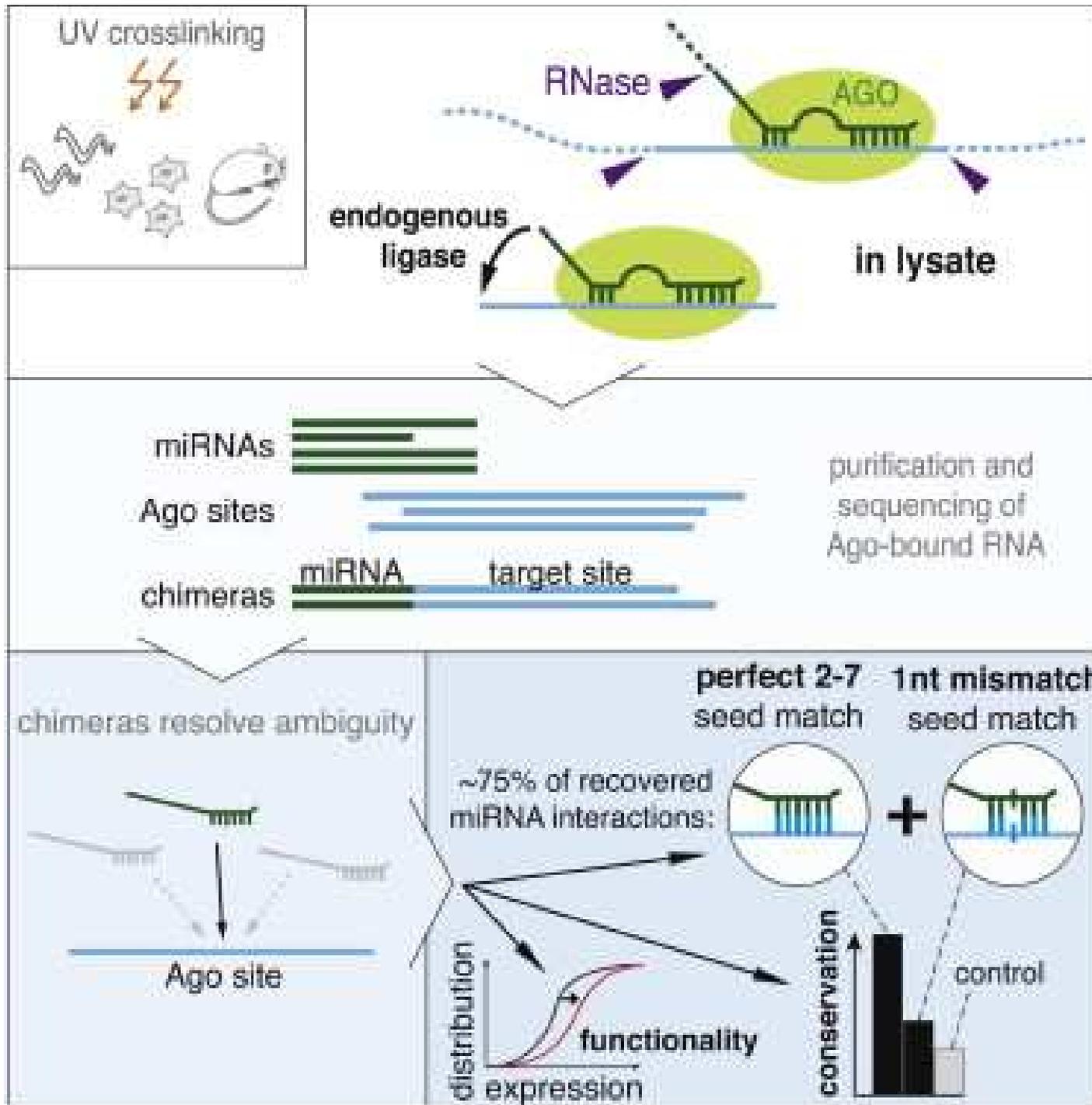
miRNA Target Prediction: standard approach



Limits:

Overwhelming number of predictions, exceeding by far range suitable for experimental testing

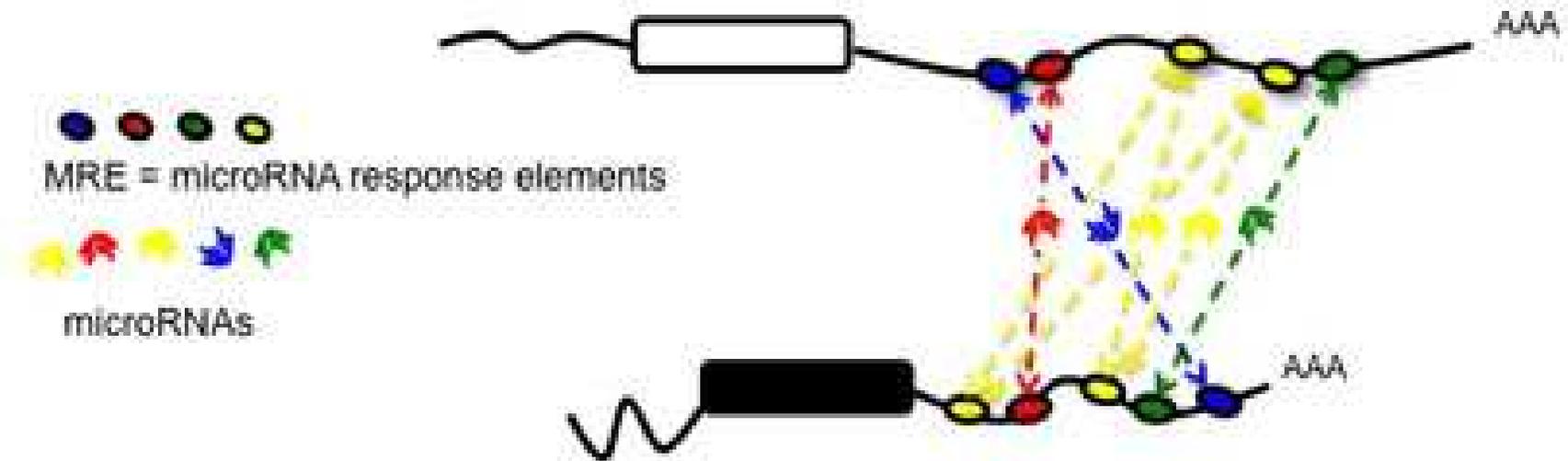
High-rate of False Positives



Coinvolti in tutte le patologie



Competing endogenous RNAs (ceRNA)



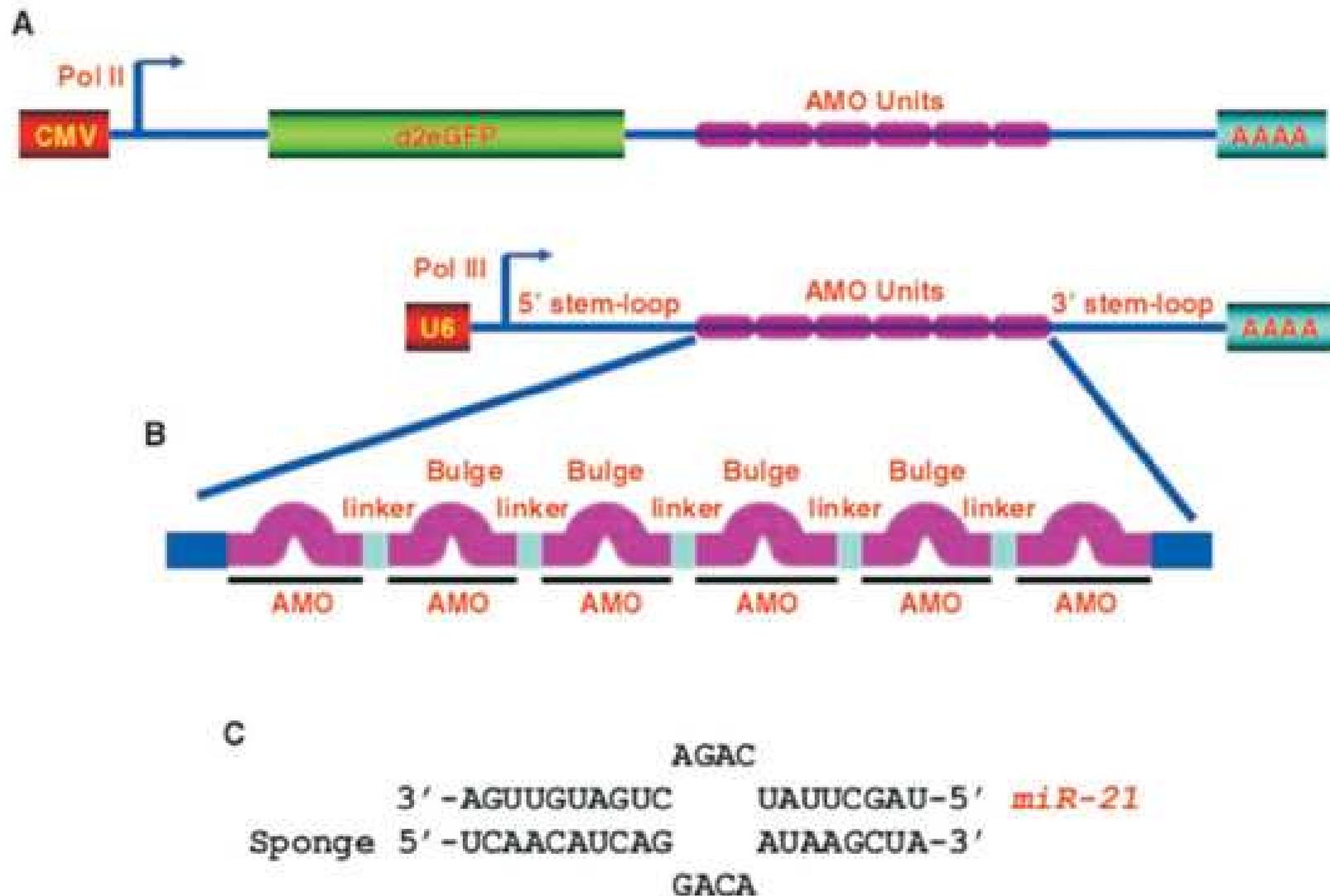
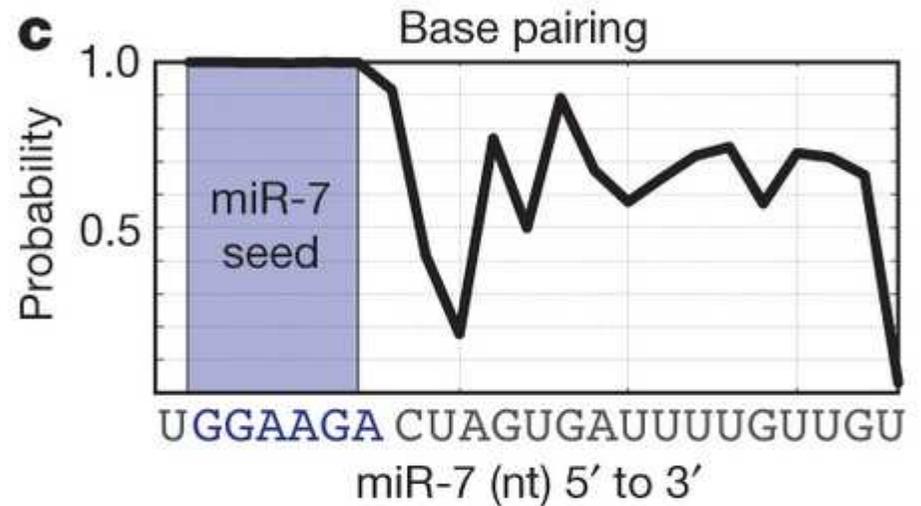
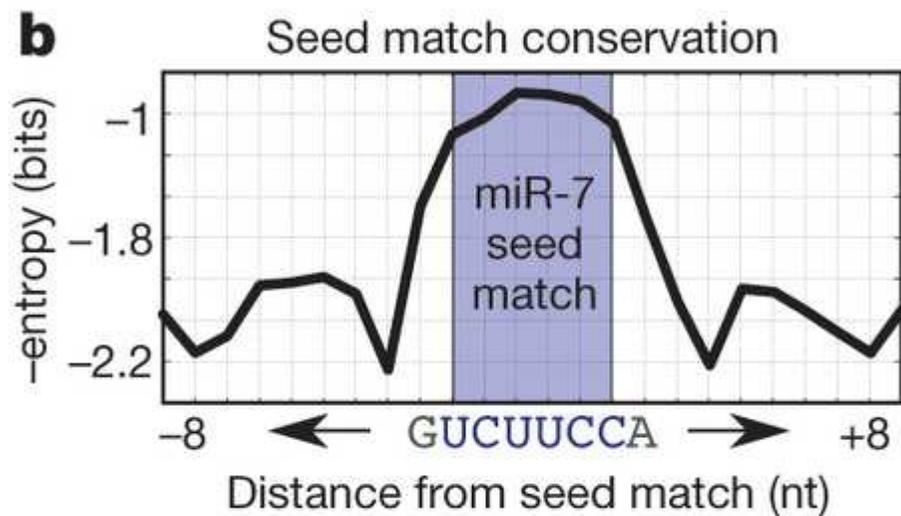
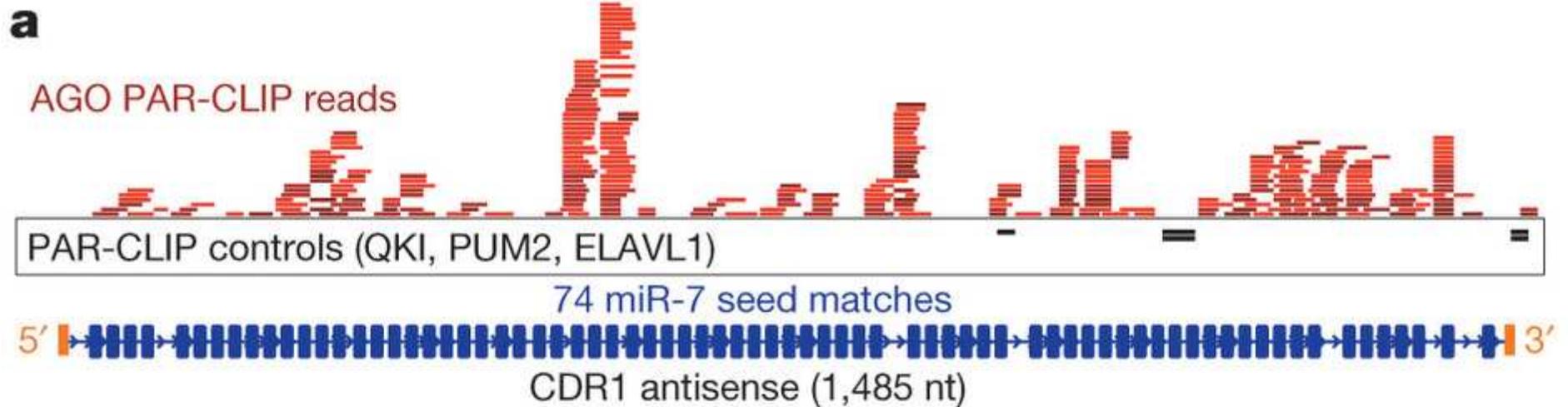


图2.7 miRNA sponge原理示意图

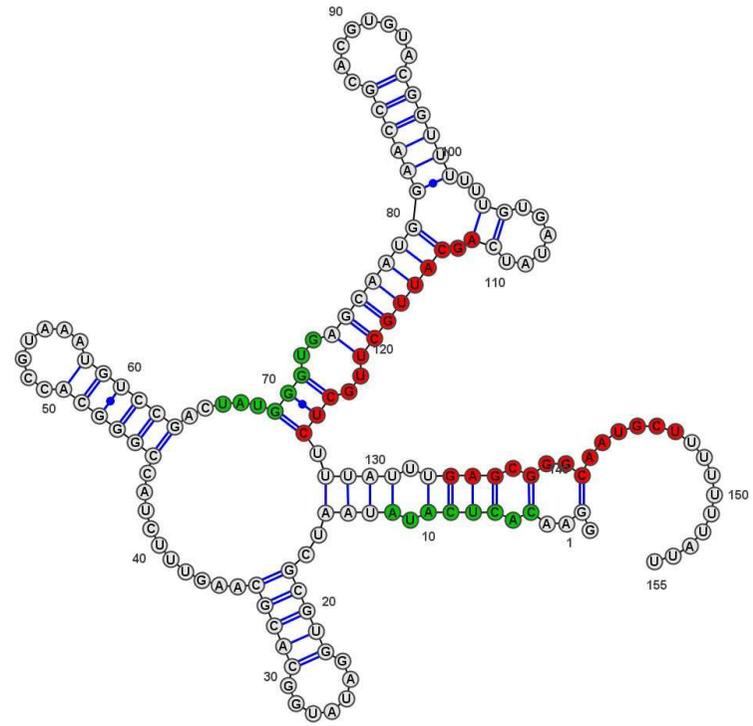
Sequestrano e localizzano



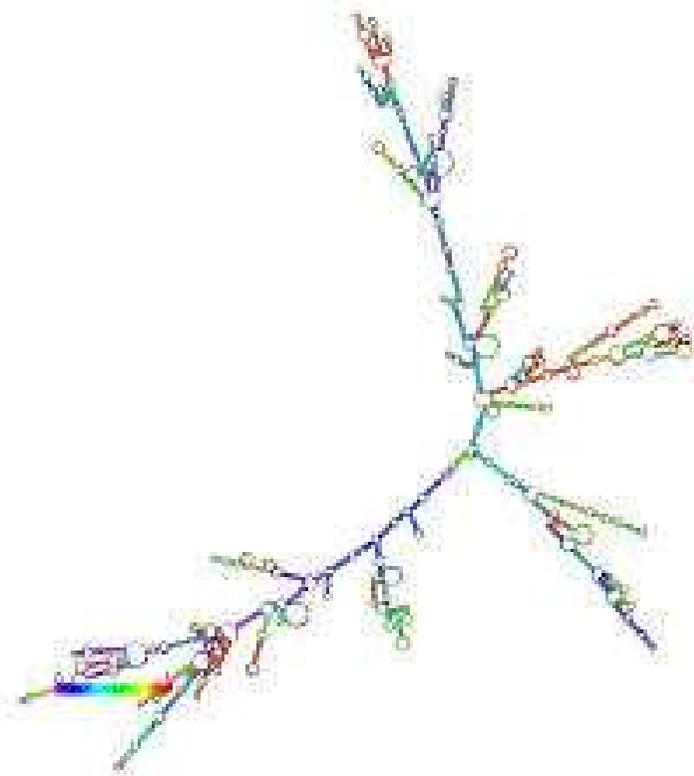
Long non coding RNA

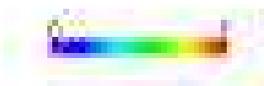
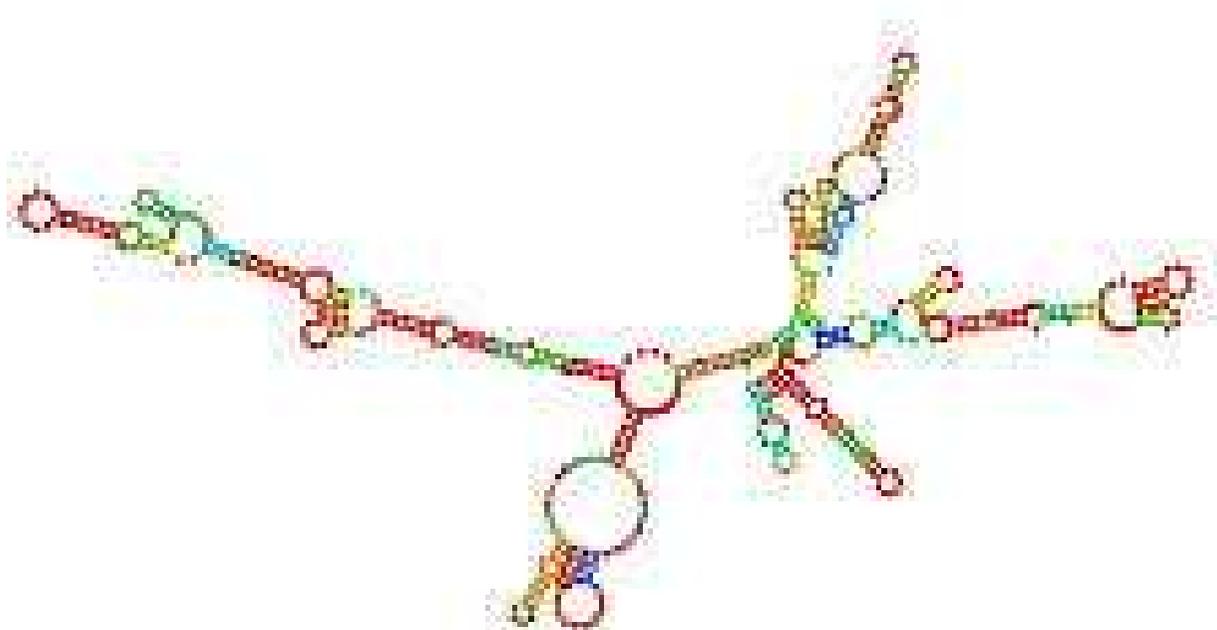
Circa 25000 nell'uomo

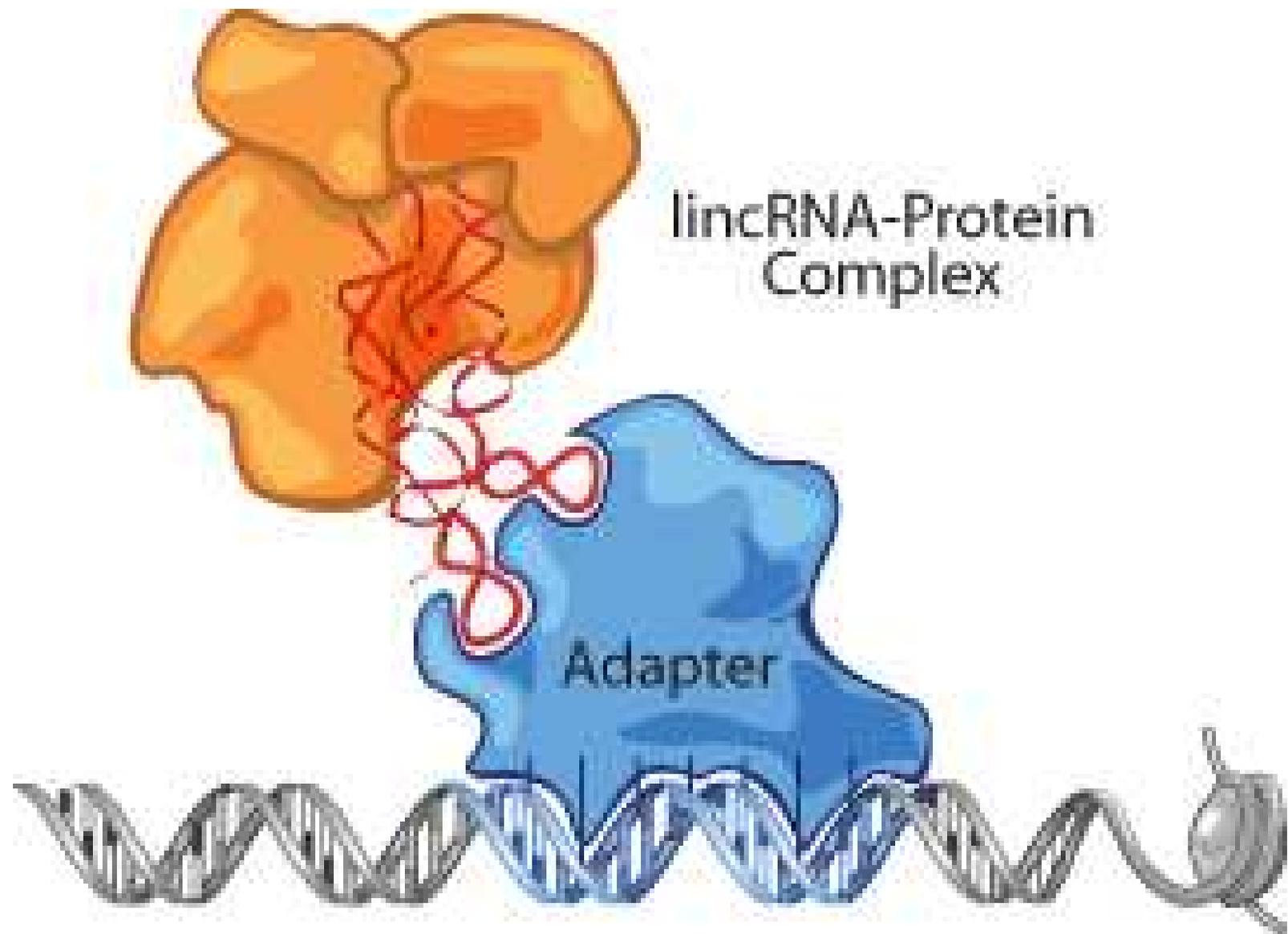
Mondo ad RNA



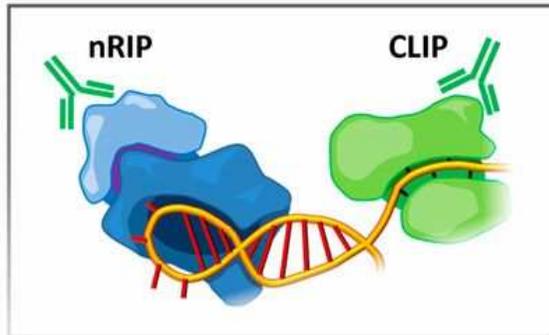
xpt-on



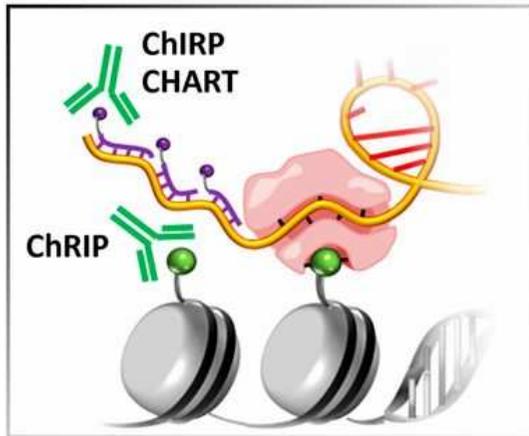




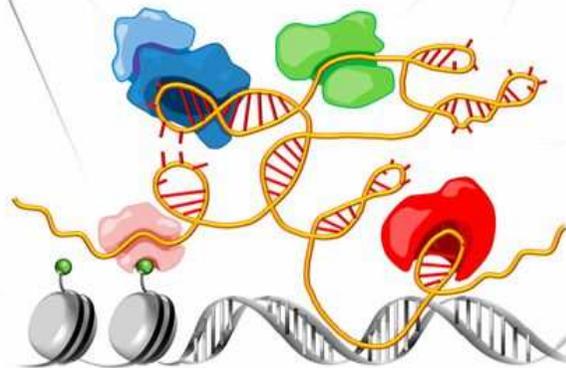
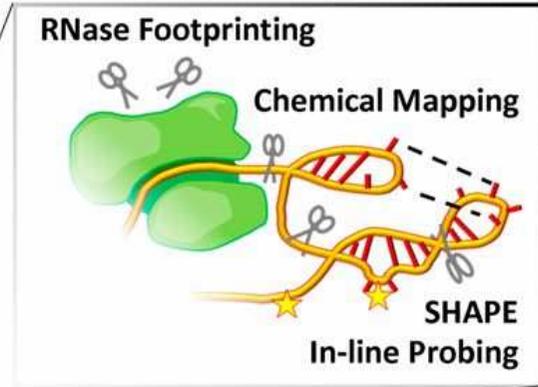
A PROTEIN INTERACTIONS



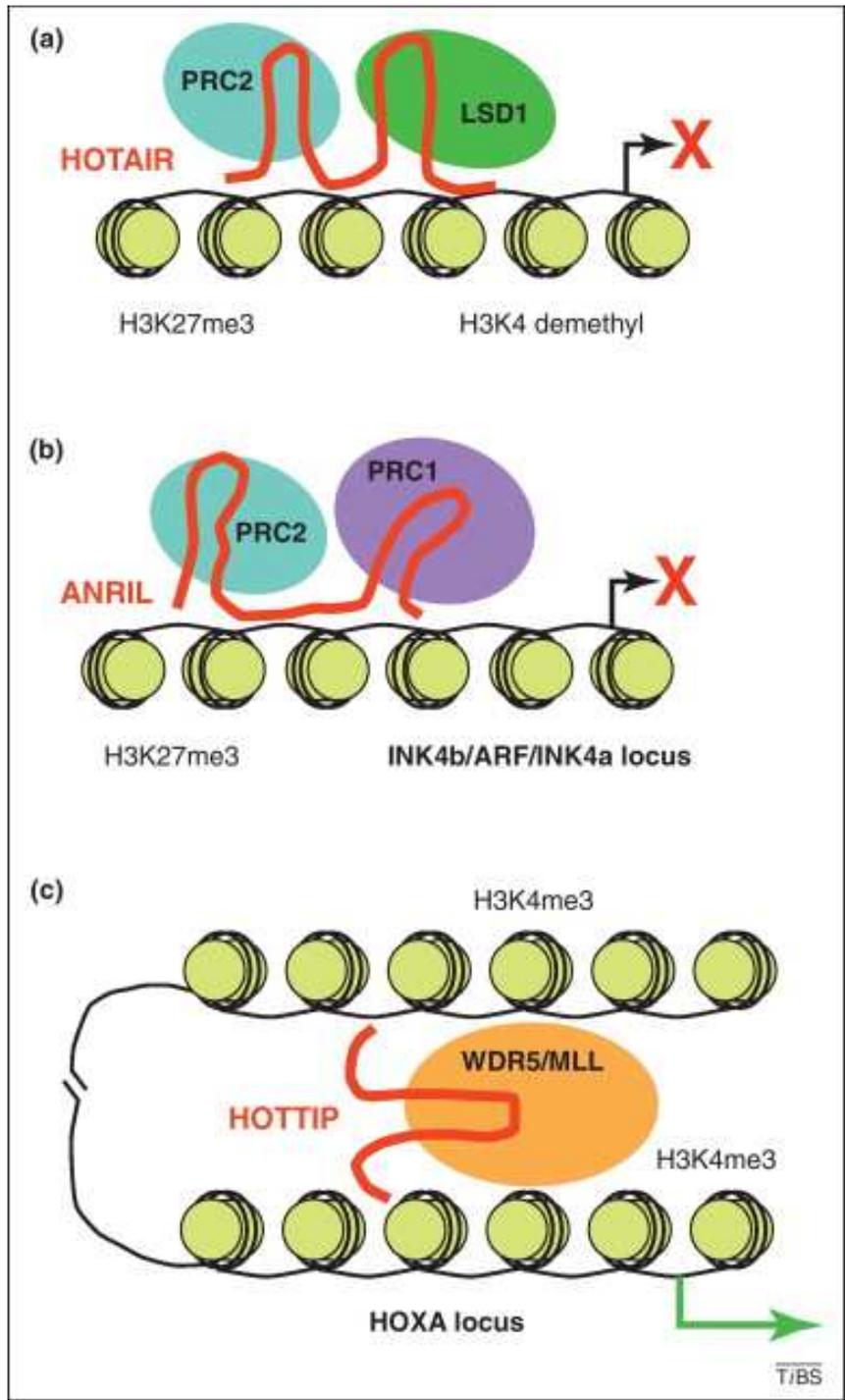
B DNA INTERACTIONS

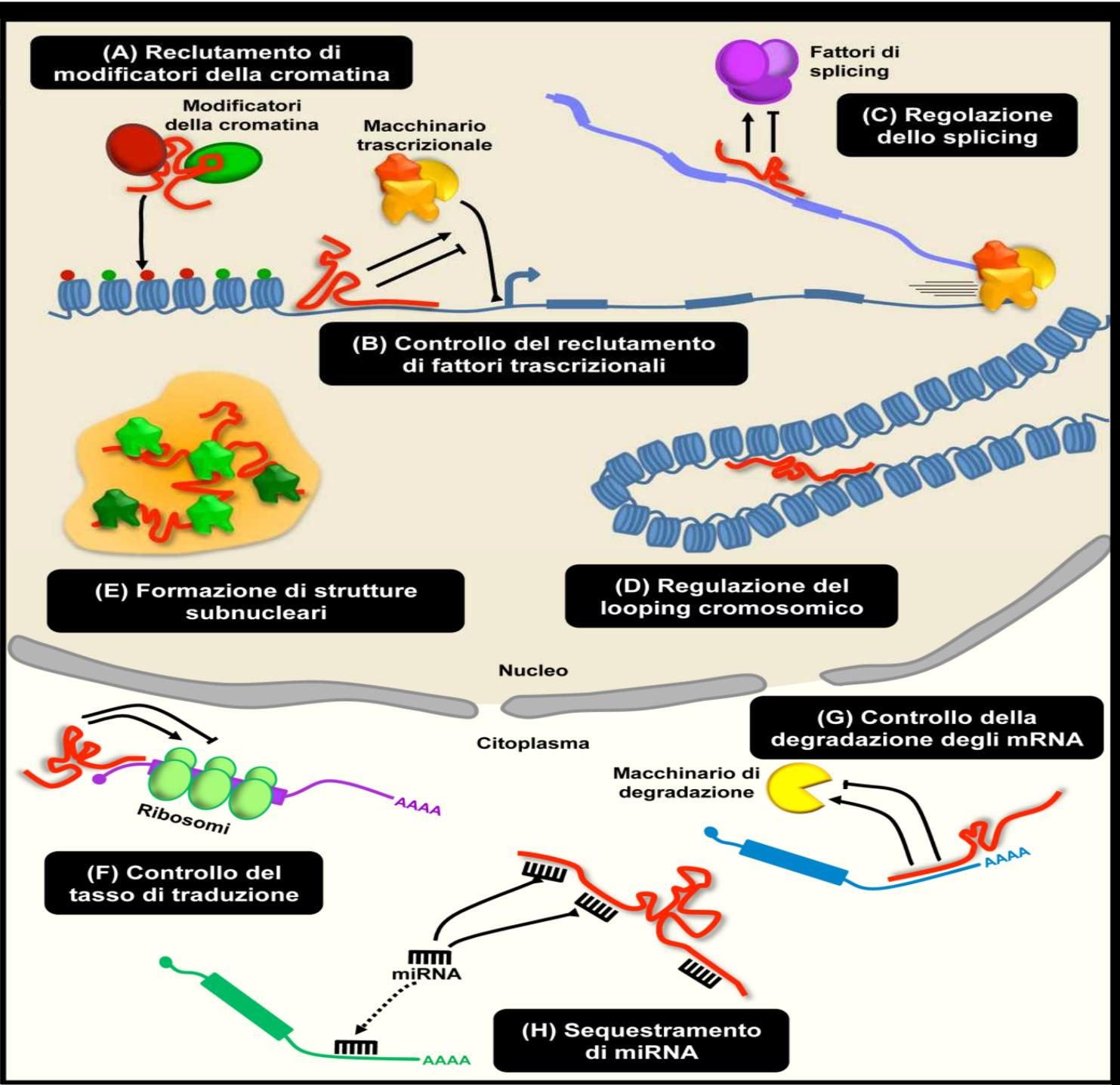


C STRUCTURAL FEATURES



- Antibody
- Histone modification
- Crosslink
- Tagged oligonucleotide
- RNase/chemical cleavage
- Flexible RNA nucleotide

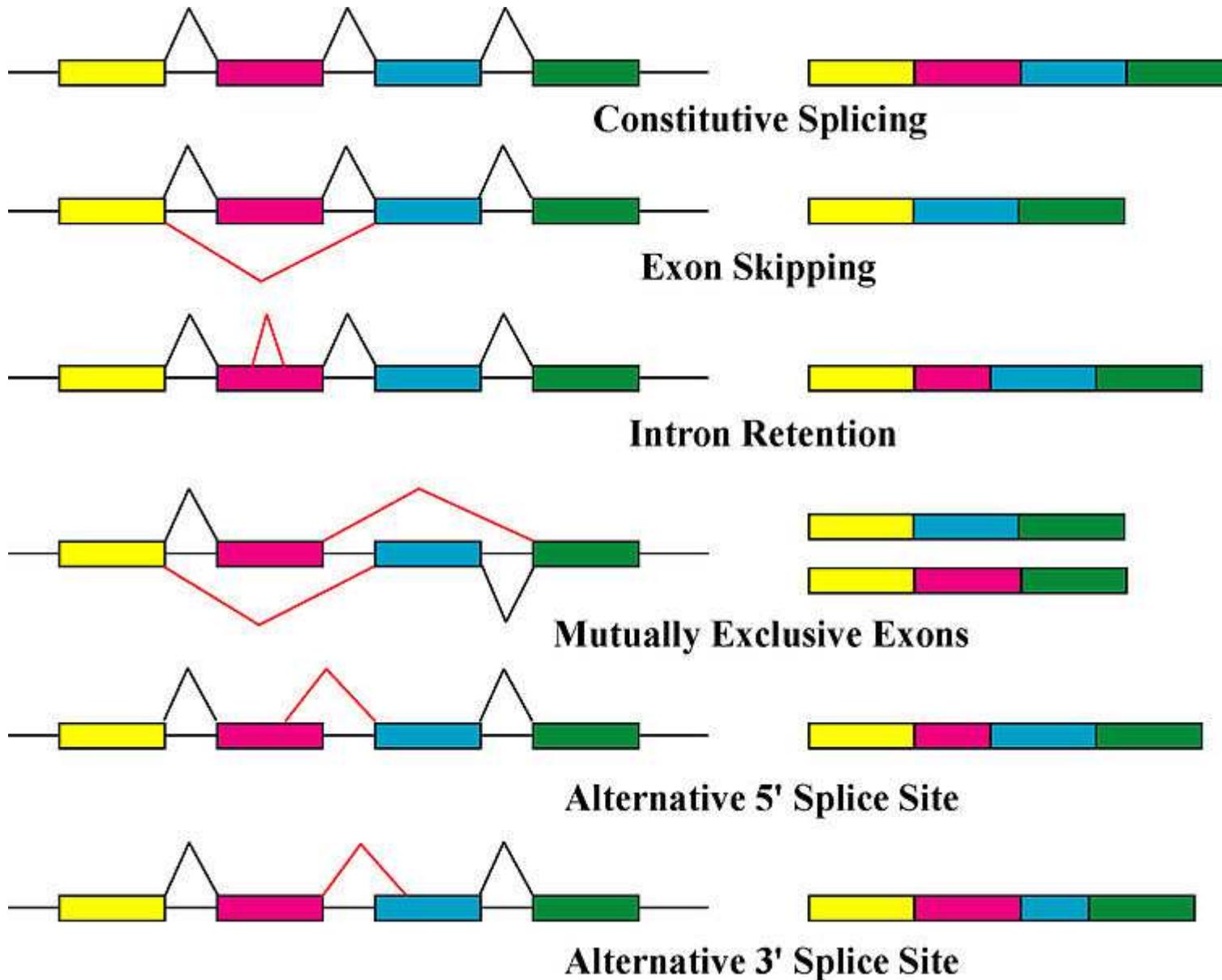




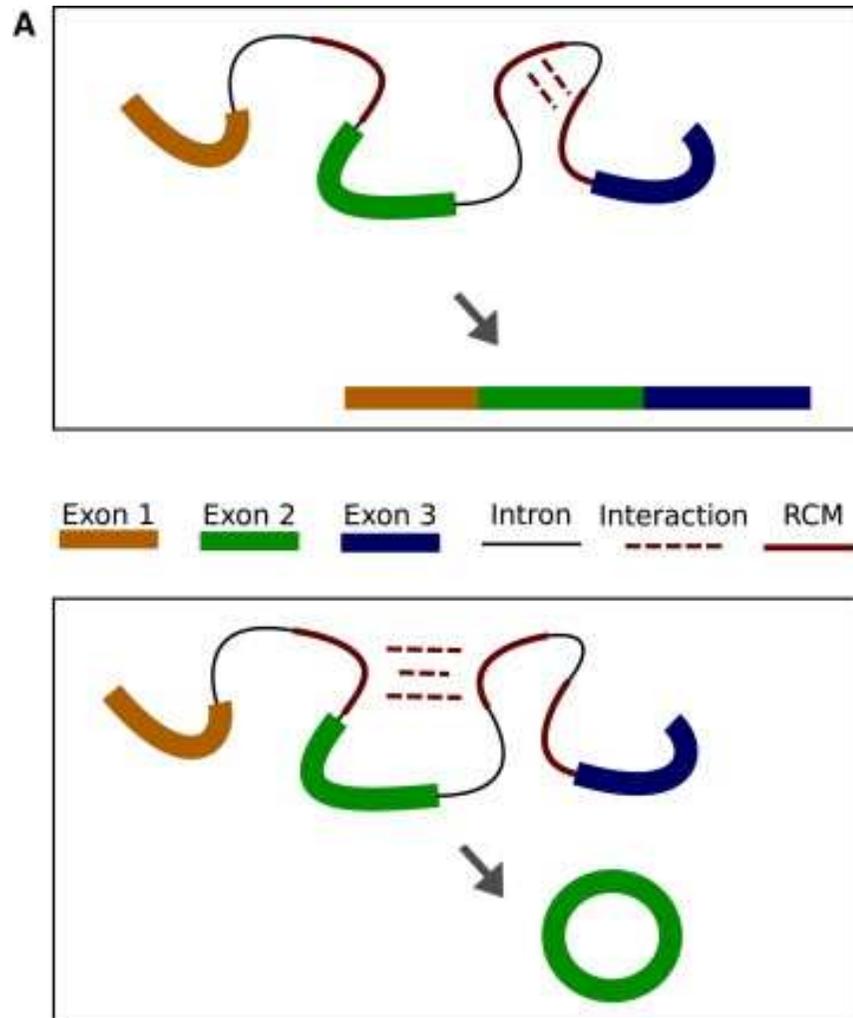
splicing

- Lo splicing moltiplica il numero di funzioni geniche (nell'uomo di quattro volte)

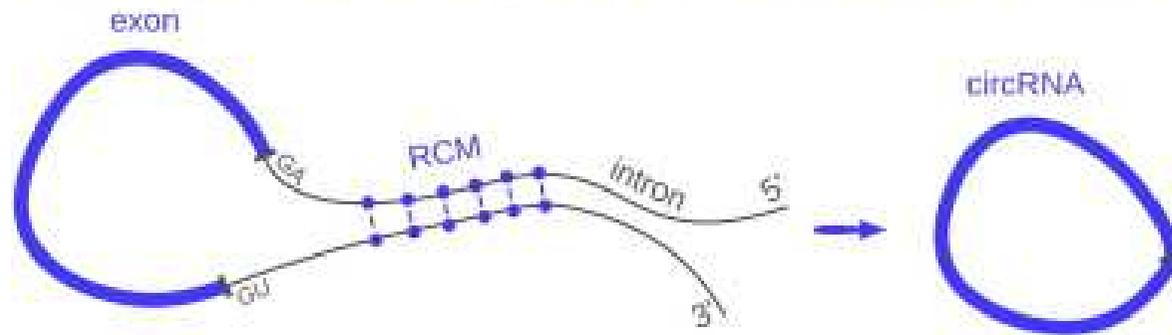
Splicing alternativo di RNA



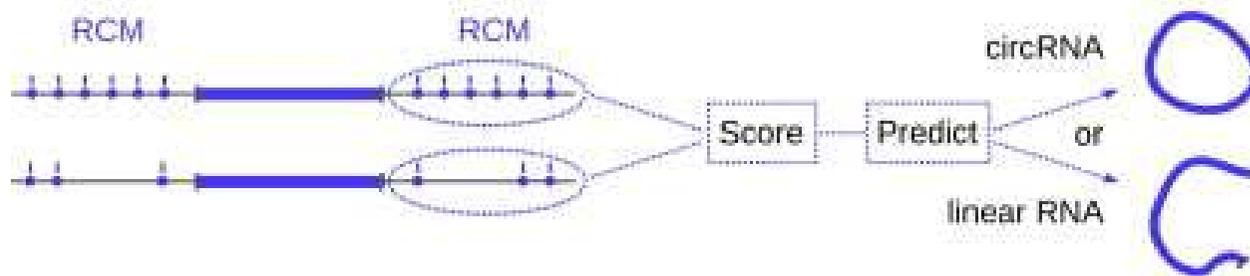
RNA circolari (circa 8000 nell'uomo)



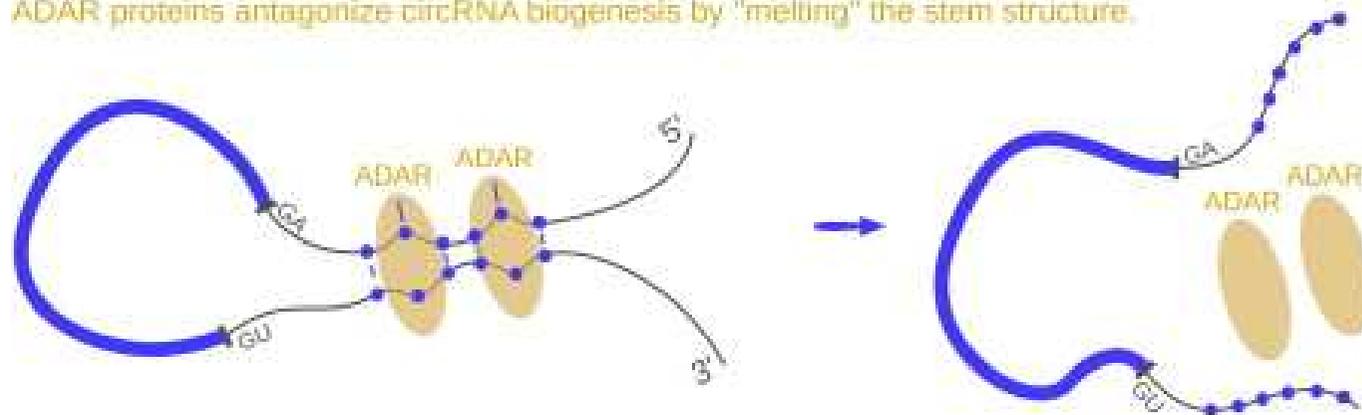
Reverse complementary matches (RCMs) are a conserved feature of circRNA biogenesis



Computational analysis of Intron sequences allows circRNA prediction.



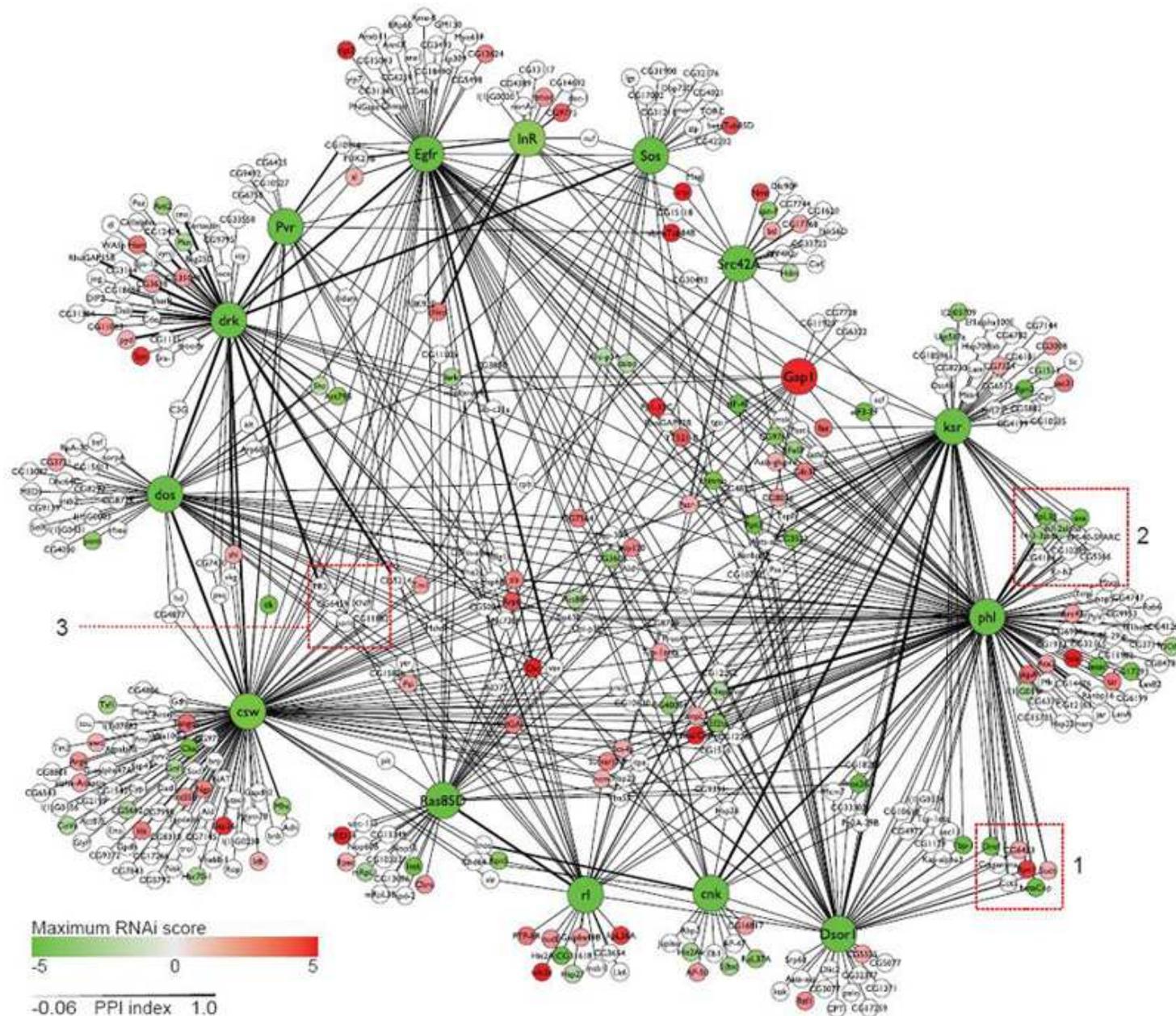
ADAR proteins antagonize circRNA biogenesis by "melting" the stem structure.



Quale e' la funzione degli RNA circolari?

- Per ora ne sono state scoperte solamente due
- 1- Effetto spugna
- 2- Vengono tradotti in proteina che agisce (forse) da dominante negativo
- Sono estremamente stabili e tendono a concentrarsi (nel cervello in particolare nelle sinapsi)
- Moltissimi sono localizzati nel citoplasma

Generano reti di controllo

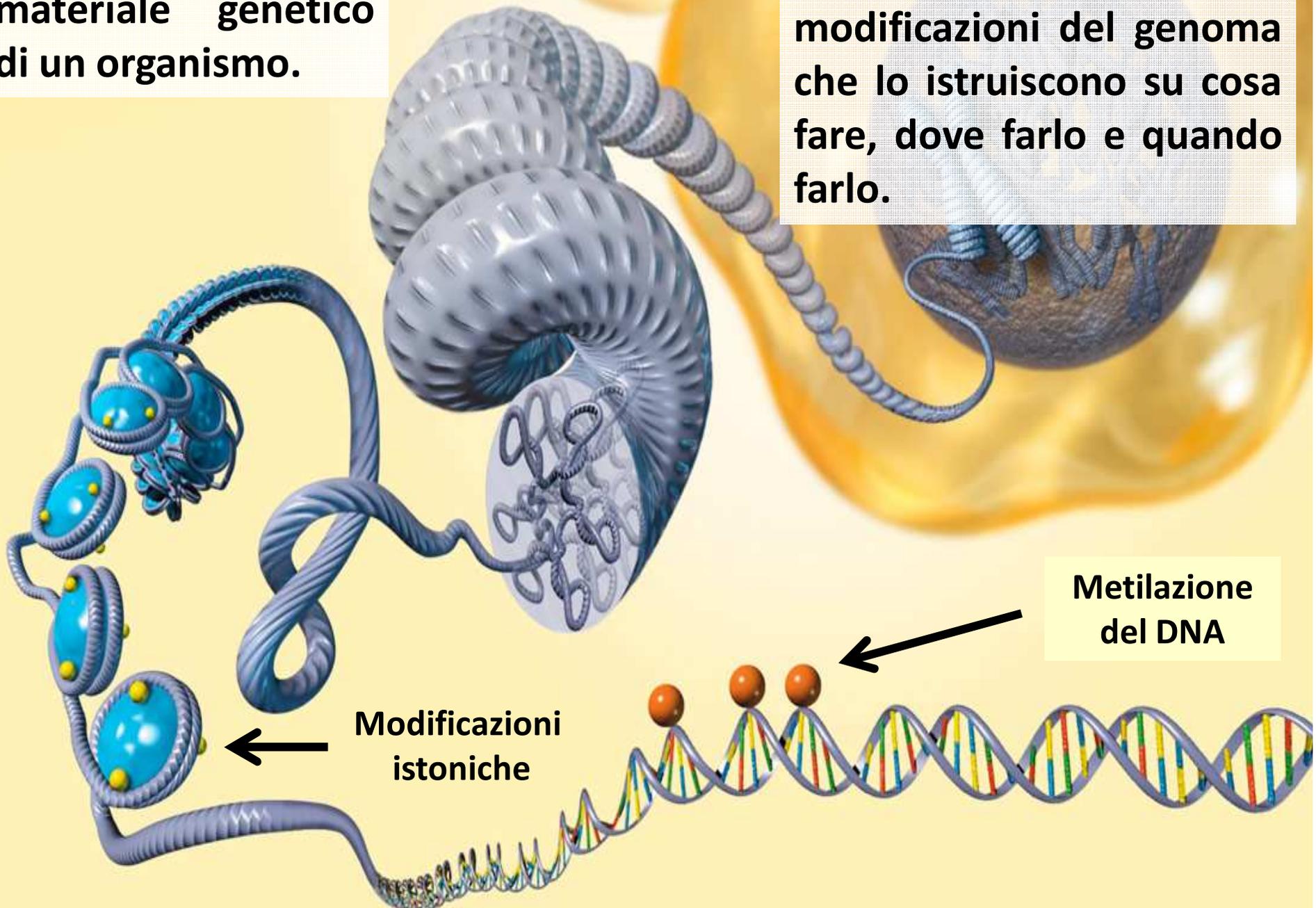


GENOMA

materiale genetico di un organismo.

EPIGENOMA

modificazioni del genoma che lo istruiscono su cosa fare, dove farlo e quando farlo.



Modificazioni istoniche

Metilazione del DNA

Poiché cellule diverse possono avere modificazioni diverse, un singolo genoma può dare origine a centinaia di epigenomi.

