

RNA non codificanti

Almeno 2/3 del DNA genomico è trascritto in molecole di RNA, ma solo il 2% corrisponde ad mRNA che vengono poi tradotti in proteine. Tutti gli altri trascritti, inclusi tRNA e rRNA, sono genericamente chiamati RNA non codificanti, o ncRNA.

Le funzioni della maggior parte dei ncRNA sono ancora da chiarire. Si sa però che la maggior parte di questi sono implicati, da soli o in associazione a proteine, nella regolazione della stabilità e della traduzione degli mRNA ed in processi regolativi ad altri livelli.

Le metodiche di sequenziamento NGS risultano molto importanti nello studio di questi ncRNA.

Funzioni degli ncRNAs

I ncRNAs sono coinvolti nella regolazione dell'espressione genica, dalla trascrizione allo splicing alla traduzione, contribuiscono anche all'organizzazione del genoma ed alla sua stabilità. Sono presenti sia negli eucarioti che nei procarioti.

Procarioti

Nei procarioti i ncRNA possono essere raggruppati in due categorie: *piccoli RNA non codificanti* (small non-coding RNA, sncRNA o sRNA) e *riboswitch*.

sncRNA

Riconoscono per appaiamento di basi complementari i loro mRNA bersaglio, regolando in trans la traduzione e la stabilità degli mRNA.

Questi sncRNA sono implicati nell'adattamento della cellula a cambiamenti delle condizioni ambientali. In E.coli ce ne sono più di 50. Non in tutti i casi il meccanismo di funzionamento è stato chiarito. Attraverso il riconoscimento degli mRNA in trans possono determinare una attivazione o una repressione della traduzione.

Riboswitch

Sono sequenze di RNA che subiscono cambiamenti conformazionali in conseguenza al legame di piccole molecole regolatrici. Agiscono in cis. Sono coinvolti nella regolazione di numerosi geni batterici implicati nel metabolismo di aminoacidi e nucleotidi.

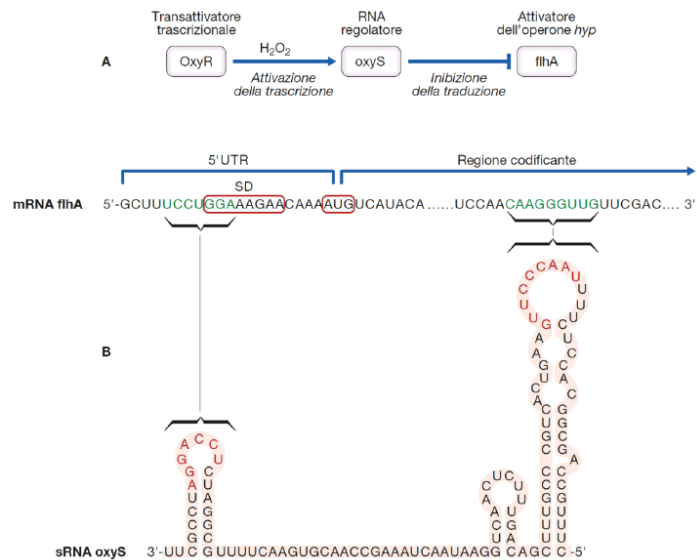
L'azione in cis dei riboswitch vuol dire che questi sono parte degli stessi trascritti che essi regolano, e quindi controllano l'espressione delle adiacenti sequenze codificanti.

sncRNA *oxyS* di *E. coli*

Quando una cellula batterica è sottoposta a stress ossidativo questa risponde attivando la trascrizione di una serie di geni della risposta anti-ossidativa. L'aggiunta di H_2O_2 al terreno di coltura determina l'attivazione di *OxyR*, che è un attivatore trascrizionale che induce l'attivazione di diversi geni tra cui il gene che codifica per lo sncRNA *oxyS*. Questo è un sRNA di un centinaio di nucleotidi. *oxyS* a sua volta regola la traduzione di una decina di mRNA bersaglio, attivandola per alcuni e reprimendola per altri.

Tra i mRNA bersaglio della repressione c'è *flhA*, che è un attivatore dell'operone *hyp*. L'operone *hyp* codifica per geni richiesti nella sintesi di idrogenasi, che sono enzimi coinvolti nella ossidazione di substrati.

oxyS presenta tre strutture a forcina con stelo e ansa. La forcina al 3' e quella al 5' si appaiano mediante complementarità al mRNA di *flhA*. L'ansa al 3' di *oxyS* è complementare ad una sequenza subito a monte del codone AUG (start) e che è parzialmente sovrapposta alla regione di Shine-Delgarno (regione di complementarità alla subunità piccola del ribosoma). L'appaiamento dell'ansa posta al 3' del sncRNA *oxyS* alla regione Shine-Delgarno dell'mRNA *flhA* impedisce il legame di quest'ultima al ribosoma, impedendo la traduzione.



Riboswitch GlmS

La *glucosammina-6-fosfato sintasi GlmS* è un enzima coinvolto nella sintesi di *glucosammina-6-fosfato* a partire da fruttosio-6-fosfato e glutammina. L'RNA che codifica per questo enzima è un riboswitch, quindi un ribozima. Questo RNA presenta una regione 5' con una particolare struttura secondaria con delle strutture a forcina. Questa regione è in grado di legare la *glucosammina-6-fosfato*, ed al legame di questa molecola si ha un cambiamento conformazionale che rende il trascritto un ribozima, con attività endonucleasica che agisce in cis, clivando se stesso in regione 5'. La rimozione della regione 5' causa la degradazione del trascritto, diminuendo la quantità di trascritto che verrà poi tradotto in GlmS, con conseguente diminuzione del suo prodotto, la glucosammina-6-fosfato.

Eucarioti

Il sequenziamento del trascrittoma ha permesso di scoprire una grande quantità di ncRNA. Questi ncRNA sono di lunghezza molto variabile e possono essere suddivisi in due grandi classi.

- Small ncRNA

Hanno una lunghezza compresa tra i 18 ed i 300 nucleotidi. A questa classe appartengono:

- Micro RNA, coinvolti nell'inibizione della traduzione. Lunghi tra i 19 ed i 24 nucleotidi. Sono trascritte come molecole precursori più lunghe, miRNA primari, che vengono poi processati da complessi enzimatici quali Drosha e Dicer. Svolgono varie funzioni, regolando trascrizione e traduzione.
- Small interfering RNA, svolgono una funzione di difesa. ??
- Piwi RNA, lunghi dai 21 ai 35 nucleotidi, sono coinvolti nella gametogenesi e nel silenziamento dei trasposoni.
- Small nuclear RNA, coinvolti nella traduzione degli mRNA. Lunghi circa 150 nucleotidi.
 - Small nucleolar RNA, coinvolti nel processamento dei rRNA. Lunghi dai 60 ai 300 nucleotidi.

- Long ncRNAs

Hanno una lunghezza superiore ai 200 nucleotidi.

- Natural antisense transcript
- Transcribed ultraconserved regions
- Long intergenic ncRNAs
- Pseudogenes
- Circular RNAs

La caratteristica comune di questi RNA è che non sono tradotti in proteine.

miRNA

I miRNA sono localizzati in regioni di DNA non codificante oppure negli introni di geni codificanti per proteine. Circa il 50% dei miRNA sono clusterizzati in regioni cromosomiche e sono trascritti da un promotore comune. In questo modo si creano diverse famiglie di miRNA che condividono un target e processi biologici su cui operano, grazie alla presenza nella loro sequenza di una *seed region*.

La seed region è rappresentata da 5-7 nucleotidi al 5' che si appaiano perfettamente al mRNA target.

Un singolo miRNA è in grado di reprimere più di 100 mRNA in media.

Negli esseri umani oltre il 60% dei geni codificanti proteine sono target di miRNA.

Biogenesi miRNA

I micro RNA vengono generati da precursori più grandi, denominati miRNA primari oppure pri-miRNA.

I pri-miRNA sono trascritti dalla RNA polimerasi II e subiscono processi di capping, slicing poliadenilazione e maturazione. I pri-miRNA sono sottoposti a due processi di maturazione. Il primo prende il nome di

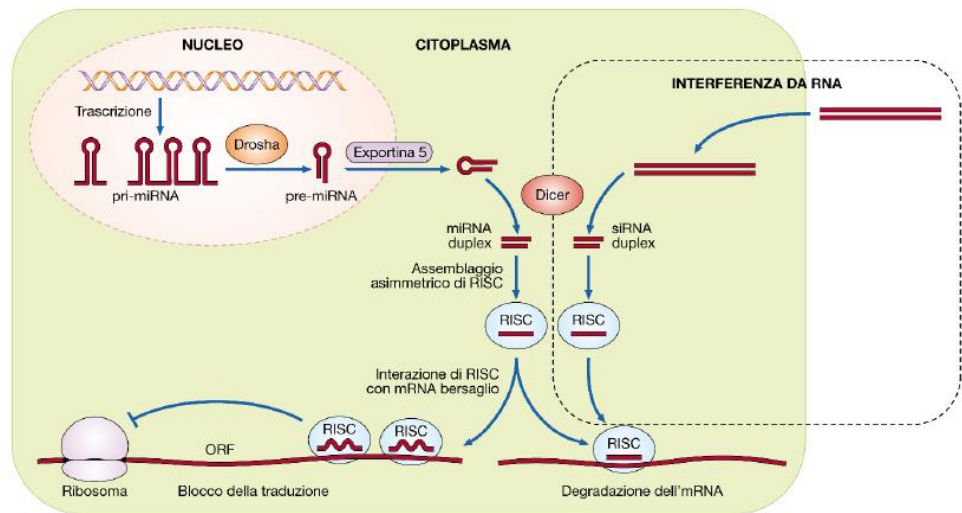
cropping ed avviene al livello del nucleo catalizzato dal complesso Drosha. Il secondo step di maturazione si chiama *Dicing* ed avviene al livello citoplasmatico catalizzato dal complesso Dicer.

Cropping

Nel primo passaggio di maturazione, il *cropping*, grazie all'azione del complesso Drosha vengono rimossi dal pri-miRNA dei tratti al 3' e 5', generando delle strutture *stem-loop* (a forcina) lunga circa 60-100 nucleotidi che prende il nome di *pre-miRNA*. Il pre-

miRNA presenta una sporgenza di una coppia di basi ad una delle estremità. Drosha appartiene alla classe delle RNasi III e taglia le molecole di RNA a doppio filamento generando estremità sfalsate di un paio di nucleotidi.

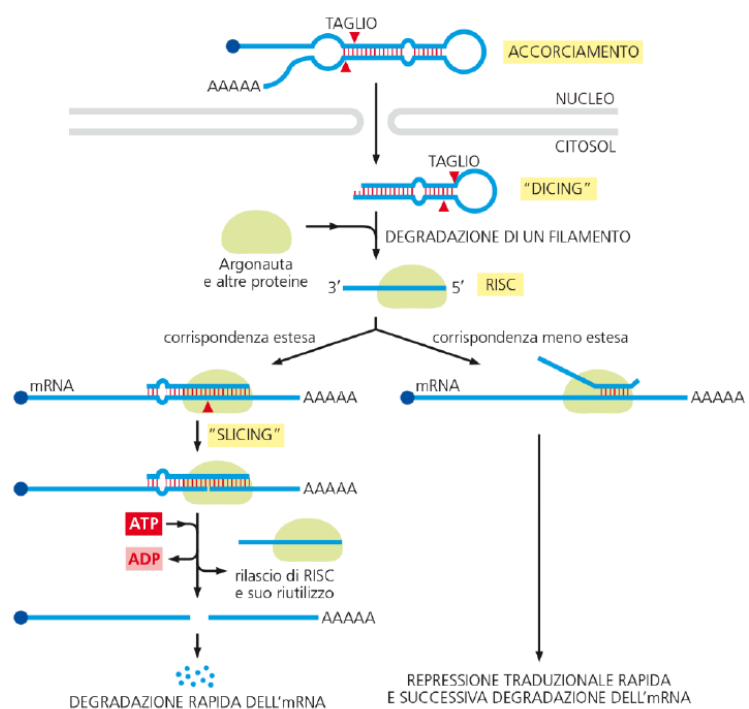
Il pre-miRNA viene trasferito dal nucleo al citoplasma attraverso i pori nucleari grazie al sistema *esportina 5* *ran GTP*.



Dicing

Nel citoplasma il pre-miRNA diventa substrato del complesso Dicer. Anche Dicer è una RNasi III, che rimuove il loop e genera un RNA a doppio filamento (duplex RNA) asimmetrico di lunghezza compresa tra i 19 ed i 24 nucleotidi. Uno dei due filamenti del duplex RNA viene caricato dal complesso RISC e prende il nome di RNA guida, con una lunghezza di circa 21 – 23 nucleotidi; l'altro filamento di RNA viene scartato e prende il nome di RNA passenger.

I miRNA maturi si legano agli mRNA bersaglio. Si legano per appaiamento di basi, per lo più in regione 3' UTR. Più miRNA diversi o più copie dello stesso miRNA possono legare un mRNA bersaglio.



Il riconoscimento specifico del bersaglio è a carico del miRNA guida. I miRNA non funzionano da soli, ma solo se assemblati nel complesso RISC (RNA Induced Silencing Complex).

Il miRNA può appaiarsi più o meno bene al suo mRNA bersaglio. Quando l'appaiamento è perfetto il mRNA bersaglio viene degradato; questo è frequente nelle piante. Quando l'appaiamento è imperfetto si ha il blocco della traduzione, e questo avviene frequentemente nei mammiferi.

Anche se l'appaiamento tra basi tra miRNA ed mRNA non è perfetto, la seed-region deve essere perfettamente complementare.

Nel caso di appaiamento perfetto tra miRNA ed mRNA la conseguente degradazione dell'mRNA è irreversibile. Nel caso di appaiamento imperfetto l'azione di blocco della traduzione può essere reversibile se si ha la separazione del miRNA dall'mRNA.

Lo studio sui miRNA ha portato allo sviluppo di una serie di tecniche utili, quali il silenziamento genico, che prende anche il nome di *interferenza da RNA*. L'interferenza da RNA consiste nell'utilizzare molecole di RNA a doppio filamento prodotte in vitro, con sequenze complementari agli mRNA che si vogliono degradare.

RISC

Il complesso RISC è un meccanismo di silenziamento genico, di cui le proteine *Argonauta* (Ago) sono una componente fondamentale e conservata. Esistono diverse proteine Ago, che vengono utilizzate dal complesso RISC per silenziare moltissimi bersagli. Il riconoscimento specifico del bersaglio resta a carico dell'miRNA.

La proteina Ago è suddivisa in domini funzionali.

Il dominio N è importante per la separazione dei filamenti del miRNA duplex.

Il dominio PAZ lega l'estremità 3' del filamento guida.

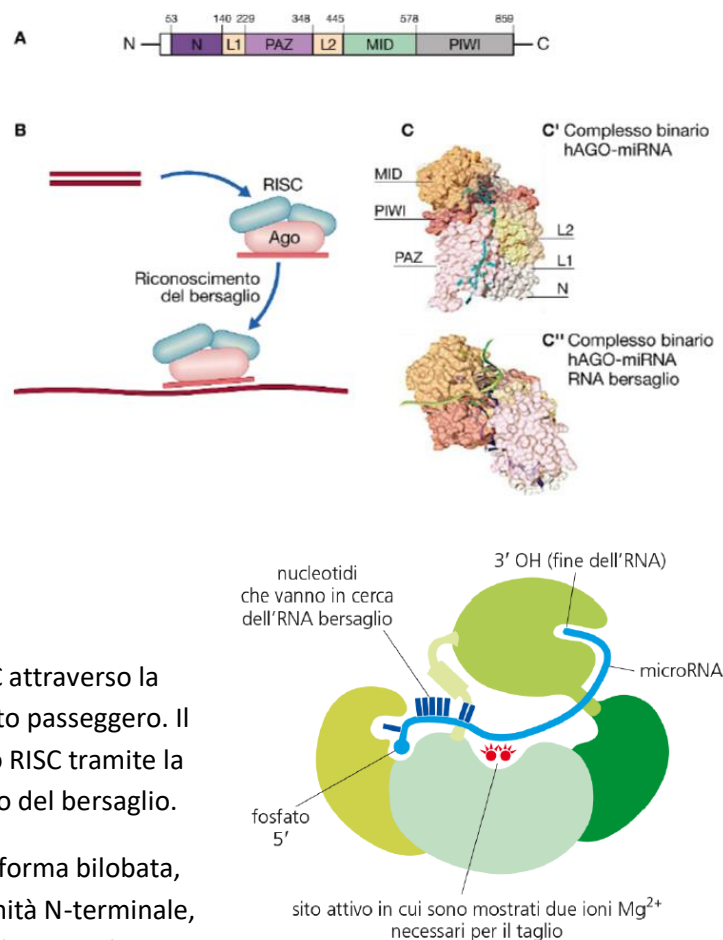
Il dominio MID lega l'estremità 5' del filamento guida, cioè la parte che si appaierà col bersaglio.

Il dominio PIWI contiene il sito catalitico coinvolto nella degradazione del bersaglio.

L1 ed L2 sono i domini di collegamento.

Il miRNA maturo a doppia elica si unisce a RISC attraverso la proteina Ago e si ha l'eliminazione del filamento passeggero. Il filamento guida rimane associato al complesso RISC tramite la proteina Ago ed è coinvolto nel riconoscimento del bersaglio.

Si forma il complesso Ago-miRNA. Ago ha una forma bilobata, in cui i domini N e PAZ sono presenti all'estremità N-terminale, mentre i domini MID e PIWI sono localizzati all'estremità C-terminale.



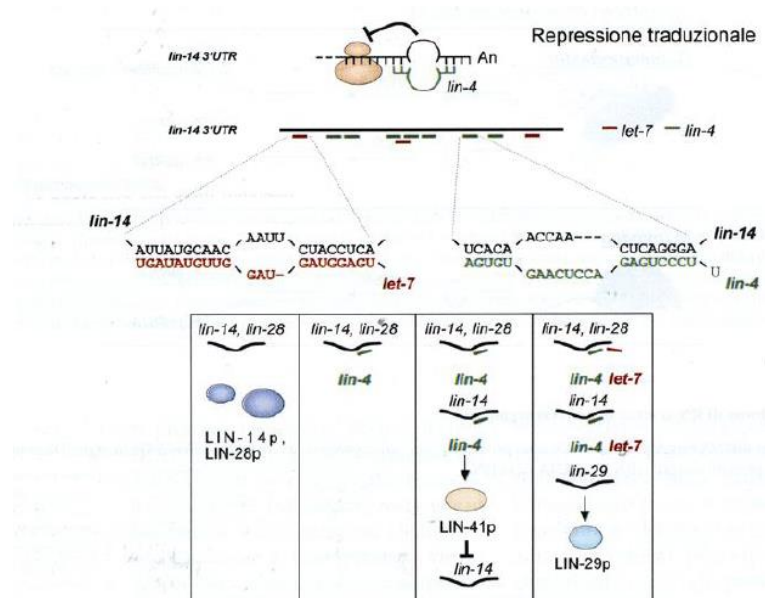
La proteina Argonata associata al miRNA lega l'RNA target. Il sito attivo di Ago taglia l'mRNA bersaglio quando forma un appaiamento esteso con i miRNA. Nel sito attivo sono presenti due ioni Mg^{2+} . Molte proteine Ago sono prive del sito catalitico e legano l'mRNA bersaglio senza degradarlo. Se l'accoppiamento tra miRNA ed mRNA target non è completo si avrà il blocco della traduzione, senza degradazione del target.

miRNA *Lin-4*

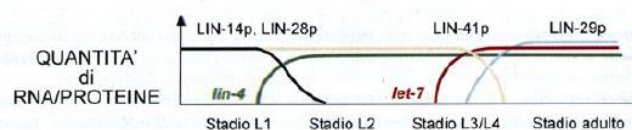
Lin-4 è il primo miRNA ad essere stato scoperto in *C.elegans*. *Lin-4* è un miRNA di 22 nucleotidi che è complementare ad una sequenza di 10 nucleotidi ripetuta sette volte nella regione 3' UTR dell'mRNA bersaglio *lin-14*.

Lin-14 è una proteina coinvolta nello sviluppo larvale di *C.elegans*, ed è regolata negativamente dal prodotto del gene *lin-4*, che è un miRNA. Durante lo stadio larvale 1 il gene *lin-14* viene attivamente trascritto ed il corrispondente mRNA viene tradotto.

Nei seguenti stadi larvali 2 e 3 la trascrizione del gene *lin-4* porta alla sintesi di un miRNA che interferisce, per appaiamento di basi, con il mRNA di *lin-14* e ne blocca la traduzione. Dallo stadio larvale 2 si ha un aumento della trascrizione di *lin-4* ed una diminuzione della proteina *lin-14*.



Nello sviluppo larvale di *C.elegans* sono coinvolti più miRNA, come *let-7*. *Lin-4* e *let-7* non si appaiano perfettamente con *lin-14*, determinando la repressione della traduzione dell'mRNA *lin-14*.



Al termine dello stadio larvale L1, all'aumentare della trascrizione di *lin-4* si ha una diminuzione dei livelli di espressione delle proteine *lin-14* e *lin-28*. Questo permette la progressione agli stadi successivi dello sviluppo. Nelle fasi larvali più tardive la trascrizione del miRNA *lin-7* aumenta e reprime l'espressione di *lin-41*, e ciò permette l'espressione della proteina *lin-29* e la progressione allo stadio adulto.

miRNA 223

miRNA 223 è un micro RNA coinvolto nella differenziazione dei granulociti indotta dall'acido retinoico RA. Nelle cellule indifferenziate la proteina NF1-A si lega al promotore del miRNA 223 e così facendo mantiene bassi i livelli di questo miRNA, effettuando una repressione della trascrizione. miRNA 223 è in grado di legarsi alla regione 3' UTR del mRNA che codifica per la proteina NF1-A stessa. miRNA 223 reprime quindi NF1-A, che a sua volta reprime miRNA 223.

Quando è presente acido retinoico RA, la proteina C/EBP α è in grado di rimpiazzare NF1-A sul promotore di miRNA 223 e di conseguenza promuove l'espressione di miRNA 223, fungendo da fattore di trascrizione. miRNA 223 si lega alla regione 3' UTR del mRNA che codifica per NF1-A, causando la repressione della traduzione del mRNA codificante per NF1-A, con diminuzione della concentrazione della proteina NF1-A. L'effetto di questa repressione traduzionale è la differenziazione dei granulociti.

La differenziazione dei granulociti coinvolge quindi almeno un miRNA e due fattori di trascrizione che competono per il legame al promotore del miRNA 223.

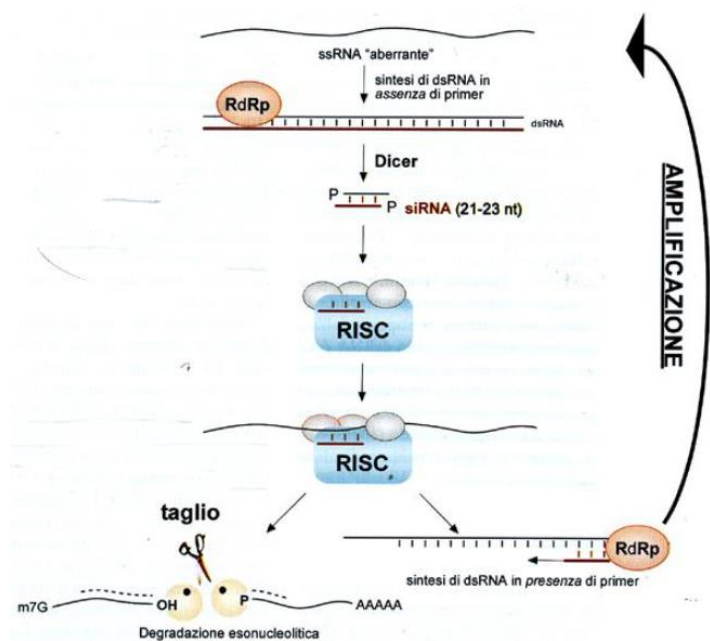
siRNA, small interfering RNA

Gli siRNA sono piccoli RNA lunghi circa 21-23 nucleotidi. Sono molecole di RNA a singolo filamento che derivano da fonti diverse, come virus a doppio o singolo filamento.

Rappresentano un meccanismo di difesa in caso di infezione virale.

Questi siRNA si trovano in piante, funghi e nematodi. Il meccanismo di difesa mediato da siRNA è cruciale per gli organismi che presentano queste molecole. Nei mammiferi il compito di difesa dalle infezioni virali si basa su un sistema mediato da proteine, cioè il sistema immunitario.

Un virus il cui materiale genetico è rappresentato da RNA a singolo filamento può infettare gli organismi che presentano siRNA. Nel caso di infezione l'RNA a singolo filamento del virus viene convertito in RNA a doppio filamento grazie all'azione della *RNA polimerasi RNA dipendente*, in assenza di primer. In questo caso il complesso Dicer converte gli RNA a doppio filamento lunghi in molecole di RNA a doppio filamento più corte, di circa 21-23 nucleotidi. Uno dei filamenti di questo RNA a doppio filamento viene incorporato nel complesso RISC.



L'RNA bersaglio del complesso siRNA/RISC è quello esogeno del virus che ha infettato la cellula. L'RNA bersaglio viene quindi riconosciuto per appaiamento di basi col siRNA, e può andare incontro a due destini. In un caso si può avere taglio dell'RNA virale, che viene poi degradato da delle esonucleasi.

In un altro caso l'RNA bersaglio può fungere da stampo per la RNA polimerasi RNA dipendente, con il siRNA che funge da primer. In quest'ultimo caso viene prodotto un RNA a doppio filamento lungo, che rientra poi nel pathway di taglio di Dicer, determinando l'amplificazione della risposta di silenziamento, e quindi della risposta di difesa.

Le RNA polimerasi RNA dipendente è essenziale per il funzionamento di questo meccanismo di difesa, perché intervengono sia nella fase iniziale che nell'amplificazione del processo di silenziamento. In ogni caso

l'azione della RNA polimerasi RNA dipendente permette la produzione di RNA a doppio filamento, che poi sono substrato di Dicer che permette di generare i siRNA.

Sebbene i miRNA e i siRNA siano generati in modi diversi questi si basano comunque sulle stesse proteine ed hanno un meccanismo di ricerca dei bersagli molto simile.

Nelle piante, ad esempio in *Arabidopsis thaliana* il meccanismo di siRNA è molto importante.

Nell'immagine vediamo come la pianta A sia una pianta sana, che cresce nella maniera corretta.

La pianta B è stata infettata da un virus CMV, e come risultato si ha una crescita ridotta della pianta, che comunque sopravvive grazie al meccanismo di difesa siRNA.

La pianta C ha una mutazione nel gene *sgs*, che è un gene necessario perché avvenga il processo di silenziamento da parte dei siRNA. La pianta non è infettata dal virus e quindi cresce in maniera molto simile alla pianta WT.

La pianta D ha una mutazione sul gene *sgs* ed è stata infettata dal virus. Come si vede la crescita è fortemente ridotta, e spesso non sopravvive.



Silenziamento genico con RNA antisenso

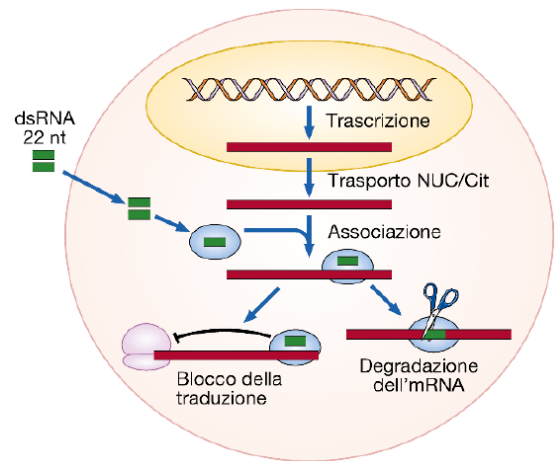
Lo studio dei meccanismi di azione dei miRNA e siRNA hanno permesso lo sviluppo di metodi di silenziamento genico. Uno dei modi per capire come funziona una proteina è deletare il gene che codifica per essa ed osservare il fenotipo risultante.

Per spegnere l'espressione di un gene era stata messa a punto la tecnica dell'RNA antisenso. Questa tecnica consiste nell'introduzione nella cellula di una molecola di RNA che essendo complementare al mRNA bersaglio vi si appaia, inibendone la traduzione.

Questi RNA antisenso presentavano una serie di problemi, come il trasferimento nella cellula, citotossicità dei costrutti ed altri effetti collaterali. Ad oggi si utilizza un'altra tecnica per inibire geni specifici, che prende il nome di *interferenza da RNA* oppure *RNAi*. Questa tecnica sfrutta il principio di funzionamento di miRNA e siRNA. Può essere utilizzata su cellule in coltura e su animali e piante. Questa tecnica ha grande potenzialità anche nel trattamento di patologie umane.

Questa tecnica consiste nella trascrizione all'interno delle cellule di RNA a doppio filamento della lunghezza di 22 nucleotidi. Uno solo dei due filamenti è complementare all'mRNA bersaglio che si vuole inibire. Senza necessità di maturazione i due filamenti vengono separati ed uno dei due viene assemblato nel complesso RISC, che riconosce il bersaglio determinandone il blocco della traduzione o la degradazione dell'mRNA.

Questa tecnica è stata messa a punto grazie all'osservazione che l'introduzione, mediante trascrizione o microiniezione in *Drosophila* o *C.elegans*, di RNA a doppio filamento di cui uno dei filamenti era complementare ad un mRNA bersaglio determinava la degradazione di quest'ultimo. In cellule umane ed in altri mammiferi la tecnica non funzionava, perché in questi organismi esistono almeno tre diversi meccanismi per la difesa dall'invasione da parte di molecole di RNA a doppia elica.



Questi meccanismi di difesa sono il sistema miRNA, una via di difesa mediata da RNA Z1 che determina la degradazione di tutto l'mRNA, ed una via mediata dalla chinasi PKR che fosforila un fattore di inizio della traduzione determinando il blocco della sintesi proteica.

Le ultime due vie, Z1 e chinasi, funzionano solo in risposta a RNA a doppia elica di dimensioni maggiori di 26 nucleotidi. Per questo motivo si usa la trasfezione con RNA lunghi 22 nucleotidi.