

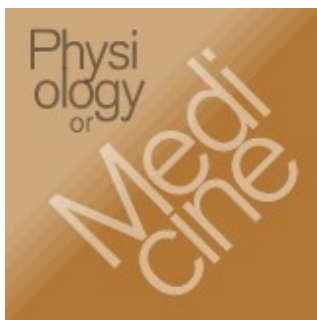


IL PREMIO NOBEL PER LA FISIOLOGIA O LA MEDICINA 2024

ASSEGNATO A VICTOR AMBROS E GARY RUVKUN

La scoperta, negli anni Novanta ad opera dei due scienziati statunitensi, di un nuovo principio di regolazione genica tramite microRNA ha aperto enormi prospettive. Oggi sappiamo che ci sono più di mille geni che codificano diversi microRNA negli esseri umani e che la regolazione genica da parte dei microRNA è universale tra gli organismi multicellulari. Le possibilità applicative riguardano la diagnosi e la prognosi di diverse patologie.

* Già Ordinario di Biochimica presso il Dipartimento di Biotecnologie e Bioscienze dell'Università degli Studi di Milano Bicocca docente di Scienze



Il premio Nobel per la Fisiologia o la Medicina del 2024 è stato assegnato congiuntamente agli statunitensi Victor Ambros e Gary Ruvkun "per la scoperta del microRNA e del suo ruolo nella regolazione genica post-trascrizionale".

Per comprendere il significato e la portata di tale scoperta è indispensabile premettere brevemente alcune elementari conoscenze di biologia molecolare.

L'informazione genetica

Nelle cellule di tutti gli organismi esistono molteplici componenti molecolari, e tra di essi l'acido deossiribonucleico (*DNA*) e le *proteine* rivestono una fondamentale importanza. In termini semplificati, il DNA contiene le informazioni per la "costruzione" dell'organismo, mentre ciascuna proteina attua una specifica funzione. Di conseguenza, nel loro insieme le proteine formano il macchinario che consente a un organismo di vivere e riprodursi. Specifici tratti di DNA che contengono l'informazione per la sintesi di una determinata proteina sono denominati *geni*. Il DNA è composto di quattro diversi tipi di nucleotidi che rappresentano l'alfabeto con cui l'informazione viene immagazzinata. Ciascun nucleotide contiene una delle



Victor Ambros (a sinistra) e Gary Ruvkun (a destra)

seguenti quattro basi azotate: adenina (A), timina (T), guanina (G) e citosina (C). In ciascun organismo, il DNA possiede una specifica sequenza di basi azotate, diversa da quella di tutti gli altri. Inoltre, come è noto il DNA consiste di due filamenti che si avvolgono elicoidalmente uno sull'altro e che contengono sequenze complementari di basi, in quanto una T su un filamento si appaia con una A sull'altro, e una G con una C. Ne deriva che ciascun filamento contiene la stessa informazione, sia in forma complementare, e di conseguenza nella duplicazione del DNA che accompagna la duplicazione cellulare, ciascun filamento funge da stampo per la sintesi di due molecole identiche tra loro e al DNA progenitore (Figura 1).

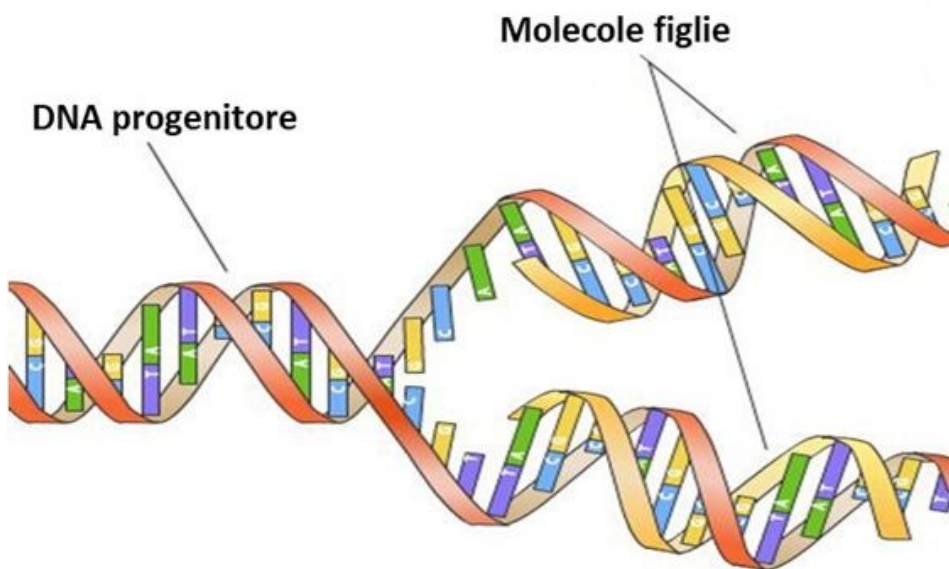


Figura 1. Una rappresentazione semplificata di una molecola di DNA. I due filamenti si avvolgono elicoidalmente uno sull'altro e al loro interno sono presenti le basi azotate (A, T, G, C) la cui sequenza codifica l'informazione genetica contenuta nella molecola. Ciascun filamento funge da stampo per la sintesi di uno nuovo avente una sequenza complementare allo stampo.

Le proteine derivano invece dall'assemblaggio di piccole molecole, gli amminoacidi, dei quali esistono 20 tipi diversi. In particolare, essi possiedono la stessa struttura base, di cui ciascuno rappresenta una "variazione sul tema". Una proteina è dunque un lungo polimero lineare, la cui individualità e funzione è stabilita da una determinata sequenza in amminoacidi (da poche decine a parecchie centinaia). Dopo la sintesi, il polimero si ripiega formando una precisa struttura, di norma globulare, che a sua volta possiede una specifica funzione. L'informazione genetica codificata nel DNA sotto forma di sequenza di basi viene decodificata ed espressa come proteine attraverso un complesso sistema in cui interviene un altro acido nucleico, l'acido ribonucleico (RNA), molto simile ma non identico al DNA. In particolare, esso si trova in forma di singolo filamento e codifica l'informazione secondo le stesse regole con cui lo fa il DNA (salvo che al posto della timina presenta una base simile, l'uracile, che come la timina si appaia con l'adenina). In un processo detto *trascrizione*, un filamento di DNA di un gene viene copiato generando una molecola di RNA complementare, chiamata RNA messaggero (mRNA). Esso viene utilizzato per sintetizzare una proteina in un processo che richiede un complesso apparato molecolare. Tale processo è detto *traduzione* in quanto consiste nella decodificazione dell'informazione contenuta nell'mRNA e quindi ultimamente nel gene. È evidente che un tale sistema di codificazione, detto *codice genetico*, prevede una corrispondenza tra sequenza di basi nel DNA (o nell'RNA) e sequenza di amminoacidi nella corrispondente proteina, in quanto ogni sequenza di tre basi codifica un

determinato amminoacido. In sintesi, si verifica un flusso di informazione nella direzione DNA → RNA → proteine (Figura 2).

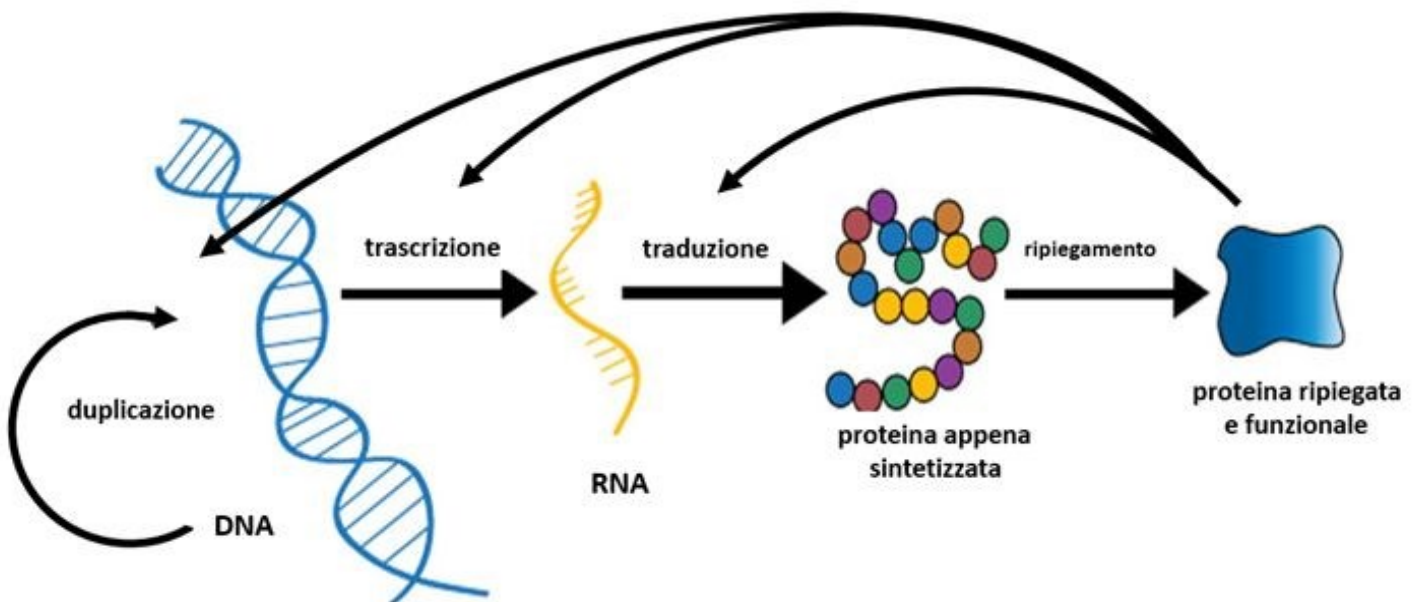


Figura 2. Il flusso dell'informazione genetica. L'informazione contenuta in sequenze specifiche di DNA codificanti proteine (i geni) viene trascritta in molecole di messaggero (mRNA), le quali a loro volta sono utilizzate per la sintesi di una specifica proteina (nelle cellule eucariotiche al di fuori del nucleo). L'informazione è codificata in modo tale per cui a ogni sequenza di tre basi del DNA e dell'mRNA corrisponda un determinato amminoacido nella sequenza amminoacidica della proteina sintetizzata. Lo schema mette in evidenza che traduzione, trascrizione e duplicazione del DNA sono esse stesse funzioni sostenute da specifiche proteine.

Nelle specie animali e vegetali più complesse, pluricellulari, le cellule sono *eucariotiche*; ciò sta a indicare che esse presentano molteplici compartimenti interni, tra i quali uno noto a tutti è il nucleo. Questa è un'osservazione attinente alla materia che stiamo trattando in quanto la trascrizione avviene nel nucleo ma l'mRNA prodotto nella trascrizione viene usato per la sintesi delle proteine (tradotto) al di fuori del nucleo.

Le risposte da un nematode “elegante”

Questo preambolo iniziale ci sembra indispensabile per contestualizzare opportunamente la scoperta che è valsa il premio Nobel ad Ambros e Ruvkun. Oltre a ciò, è anche opportuno premettere che già dagli anni Sessanta del secolo scorso era stato assodato che la sintesi delle diverse proteine è regolata, vale a dire che ciascuna di esse non viene prodotta a un tasso costante; al contrario, esso varia in misura anche cospicua a seconda delle esigenze cellulari (come per esempio il fabbisogno energetico o l'avvio della duplicazione cellulare). Tale meccanismo di controllo è basato sull'azione di proteine specializzate, note come *fattori di trascrizione* che, legandosi a regioni specifiche del DNA decidono quali mRNA debbano essere prodotti, quando e in quale misura. All'inizio degli anni Novanta del secolo scorso, migliaia di fattori di trascrizione erano già stati identificati e ciò indusse a credere che i principi fondamentali della regolazione genica fossero stati pienamente chiariti.

E in questa fase dello sviluppo delle conoscenze che si inserisce la storia della scoperta di Ambros e Ruvkun. Alla fine degli anni Ottanta del secolo scorso i due erano giovani dottori di ricerca nel laboratorio di Robert Horvitz (al MIT di Cambridge negli Stati Uniti), a sua volta premio Nobel per la Fisiologia o Medicina nel 2002 insieme a Sydney Brenner e John Sulston. Horvitz investigava un piccolo organismo, un verme nematode lungo circa 1 mm denominato *Caenorhabditis elegans*, nel tentativo di chiarire i meccanismi che governano la differenziazione dei vari tipi cellulari durante lo sviluppo.

Può apparire sorprendente il fatto che un organismo così apparentemente semplice possa fornire indicazioni valide anche al riguardo dei meccanismi in atto nelle specie più complesse, uomo compreso. Eppure sappiamo bene che molte informazioni basilari in merito alla biologia umana sono state ottenute attraverso studi su specie molto più semplici, grazie alla comune origine evolutiva e alla conseguente adozione di strategie simili a livello molecolare in organismi anche molto diversi. Nello specifico, anche il *C. elegans*, pur nella sua semplicità, consiste di diversi tipi cellulari come per esempio neuroni, cellule muscolari ed epiteliali; dunque presenta anch'esso fenomeni di differenziazione cellulare nel corso del suo sviluppo. In effetti, la storia che stiamo raccontando documenta in modo eloquente l'utilità di tale modello animale nell'illuminare meccanismi molecolari di validità generale, come apparirà chiaro col procedere della presente narrazione.

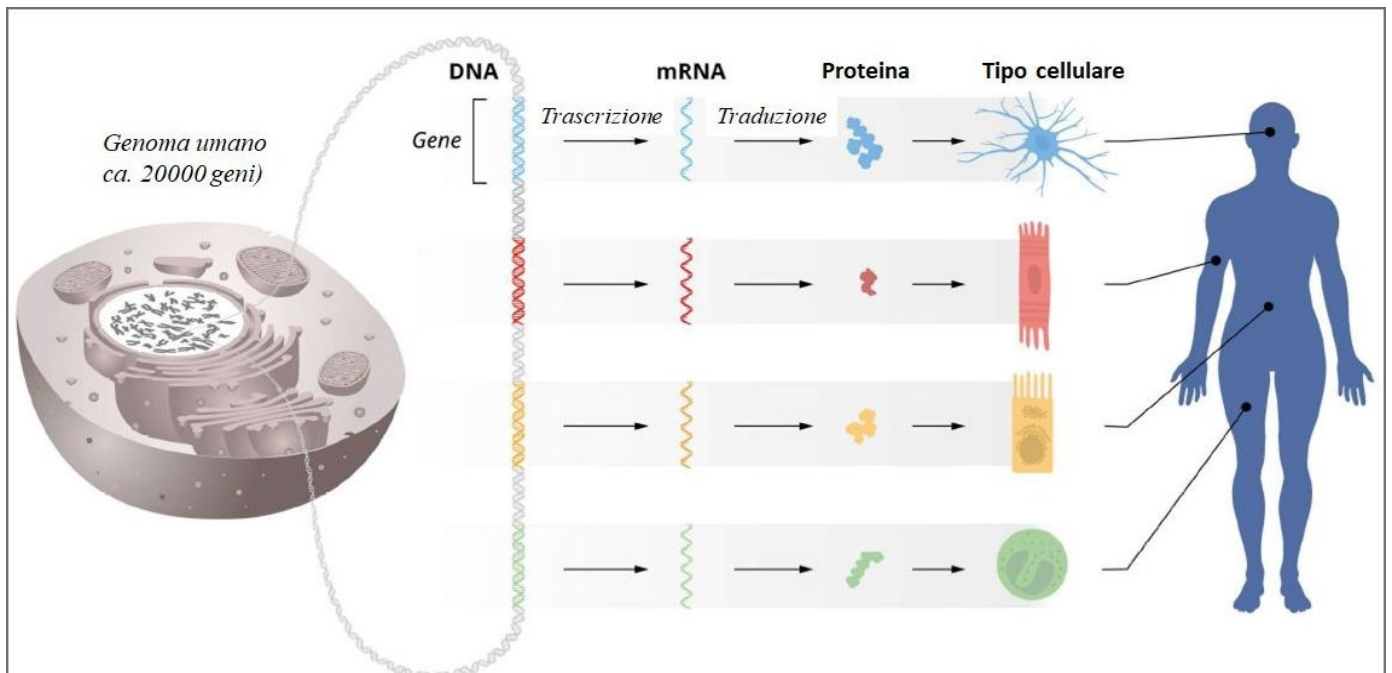


Figura 3. Il problema della differenziazione cellulare. Dato che in tutte le cellule di un organismo è depositata una identica informazione genetica, come fanno a svilupparsi diversi tipi cellulari nel corso dello sviluppo? Ciò richiede un'accurata regolazione dell'attività genica, in modo tale che in ogni tipo cellulare sia attivo uno specifico repertorio

In effetti, il problema della differenziazione cellulare ha rappresentato un interrogativo di enorme portata fin dai primordi della biologia cellulare e molecolare. Infatti il *genoma*, cioè l'insieme del DNA di un organismo, è lo stesso in tutte le cellule in quanto tutte derivano da un'unica cellula uovo fecondata: come dunque possono originarsi i diversi tessuti se tutte le cellule sono geneticamente identiche? (Figura 3).

Come abbiamo visto, la scoperta dei fattori di trascrizione di parecchi decenni fa aveva in parte risposto a tale interrogativo; tuttavia Ambros e Ruvkun dimostrarono che questa era solo una parte della storia. Alla ricerca di altri fattori responsabili della differenziazione cellulare, essi si misero alla ricerca di geni del *C. elegans* coinvolti nel controllo della tempistica di attivazione al momento opportuno di diversi programmi genetici e di conseguenza dello sviluppo dei diversi tipi cellulari. In particolare poterono analizzare due ceppi di questo nematode, che presentavano mutazioni in due geni, denominati *lin-4* e *lin-14*. Oggi sappiamo che *lin-14* è un fattore di trascrizione avente un ruolo fondamentale nello sviluppo, in quanto capace di controllare molteplici proteine coinvolte nel processo. In effetti, una mutazione inattivante di *lin-14* si traduce nel blocco dello sviluppo larvale. Inoltre Ambros dimostrò già nei suoi primi studi che *lin-4* era un regolatore negativo di *lin-14*; in altre parole, l'azione di *lin-4* inibiva l'espressione

del fattore di trascrizione codificato da *lin-14*, e, in tal modo, poteva modularne la funzione, determinando quando e in che misura *lin-14* dovesse funzionare e temporizzando così gli eventi necessari per il normale sviluppo dell'animale. Di conseguenza, mutazioni del gene *lin-4* non bloccavano lo sviluppo ma portavano alla formazione di un adulto con profonde alterazioni anatomiche (Figura 4). Naturalmente queste osservazioni suscitarono subito l'interrogativo circa le modalità con cui *lin-4* regolava *lin-14*.

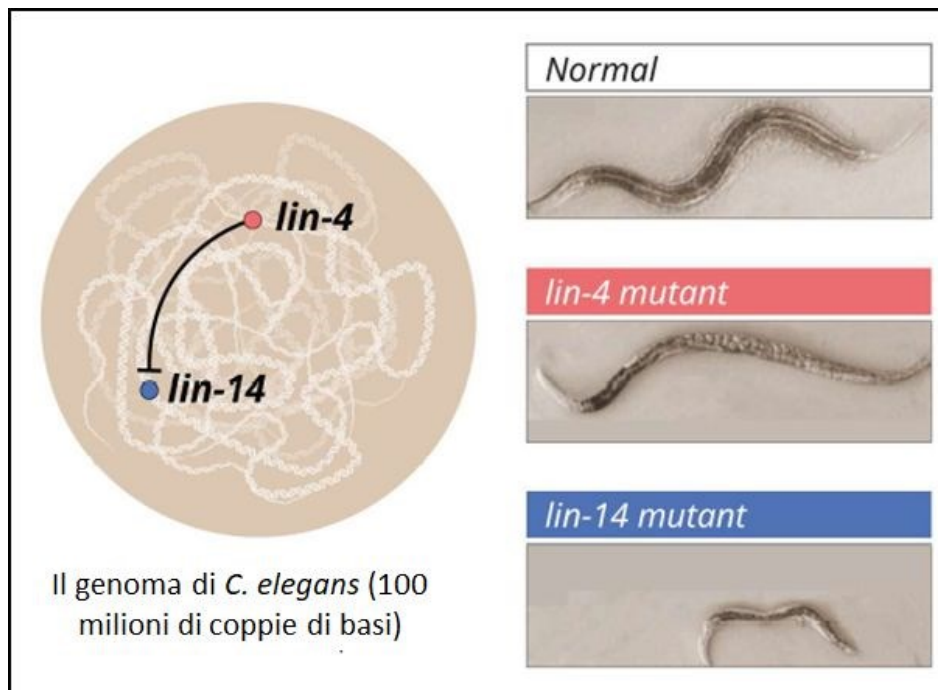


Figura 4. In C. elegans, il gene lin-4 inibisce l'espressione del gene lin-14, il quale a sua volta codifica un fattore di trascrizione, responsabile dell'espressione di molti geni coinvolti nello sviluppo dell'organismo. Mutazioni inattivanti di lin-14 bloccano lo sviluppo allo stadio larvale, mentre la mutazione lin-4, abolendo i meccanismi di controllo di lin-14, porta allo sviluppo di un adulto con sostanziali difetti anatomici.

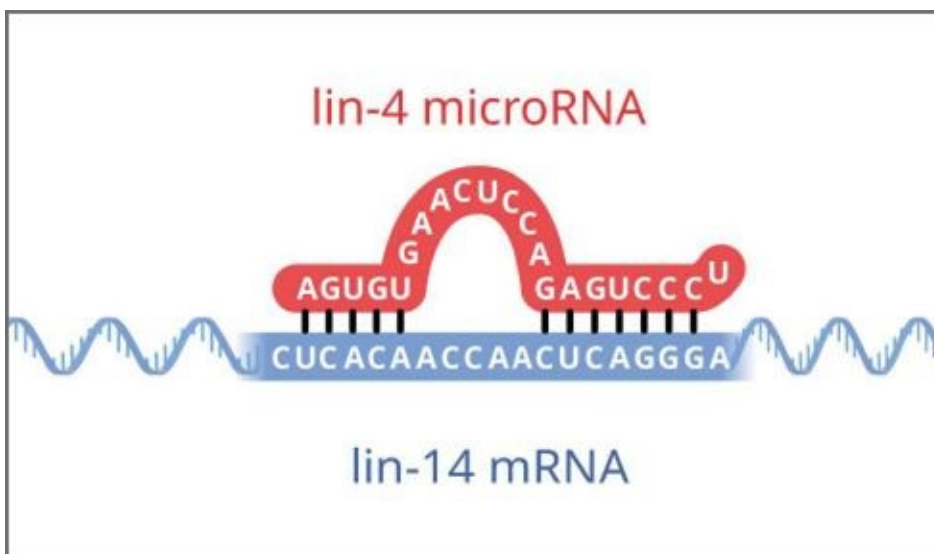
Al fine di meglio comprendere la natura del prodotto del gene *lin-4* e i suoi meccanismi di azione, dopo il suo periodo post-dottorato Victor Ambros proseguì il lavoro nel suo laboratorio appena fondato presso l'Università di Harvard. Riuscì così a clonare il gene e ciò portò a una scoperta sorprendente: esso codificava una molecola di RNA insolitamente corta (poche decine di nucleotidi). Dunque non poteva essere un mRNA, cioè un acido nucleico portante l'informazione per la sintesi di una proteina, sia perché i veri mRNA sono in ogni caso molto più lunghi, sia il piccolo RNA perché mancava di quelle sequenze di basi necessarie per poter essere riconosciuto e destinato a questo utilizzo.

Il microRNA alla base della regolazione genica

Al contempo, Gary Ruvkun, proseguì le sue ricerche nel suo laboratorio di recente istituzione presso il *Massachusetts General Hospital* e la *Harvard Medical School*, dove realizzò una scoperta che chiariva le osservazioni di Ambros circa il possibile meccanismo di azione del prodotto del gene *lin-4*. A differenza di come si sapeva allora che funzionasse la regolazione genica, Ruvkun dimostrò che non era la produzione di mRNA da *lin-14* a essere inibita da *lin-4*. La regolazione sembrava invece verificarsi in una fase successiva del processo di espressione genica, attraverso l'arresto della produzione della proteina codifica-

ta nel gene *lin-14*. I due scienziati confrontarono quindi le loro scoperte, dimostrando che la breve sequenza di *lin-4* era complementare a uno specifico segmento dell'mRNA di *lin-14* e che vi si legava bloccando la produzione della proteina *lin-14*. Era stato così scoperto un nuovo principio di regolazione genica, mediato da un tipo di RNA precedentemente sconosciuto, che fu denominato microRNA (Figura 5).

I risultati furono pubblicati nel 1993 in due articoli sulla rivista scientifica *Cell*. Tuttavia, come spesso accade nel caso di scoperte di fenomeni completamente nuovi, la comunità scientifica non ne colse immediatamente l'importanza. In particolare, si sospettava che l'inusuale meccanismo di regolazione genica fosse una peculiarità di *C. elegans*, probabilmente irrilevante per gli esseri umani e altri animali più complessi. Tale percezione cambiò nel 2000 quando il gruppo di ricerca di Ruvkun pubblicò la scoperta di un altro microRNA, codificato dal gene *let-7*. A differenza di *lin-4*, il gene *let-7* era altamente conservato e presen-



te in tutto il regno animale.

*Figura 5. Una rappresentazione schematica dell'appaiamento del piccolo RNA codificato nel gene *lin-4* con uno specifico tratto dell'mRNA prodotto dal gene *lin-14*. La figura mostra che in realtà la complementarità tra le due molecole è parziale e ciò porta alla formazione di un'ansa non appaiata. Ciò non di meno il legame è stabile grazie all'appaiamento delle due sequenze complementari.*

Questo articolo scientifico, contrariamente a quello pubblicato nel 1993, suscitò un enorme interesse e negli anni successivi molti altri scienziati si dedicarono a queste ricerche, così che a oggi sono stati identificati centinaia di diversi microRNA. Oggi sappiamo che ci sono più di mille geni che codificano diversi microRNA negli esseri umani e che la regolazione genica da parte dei microRNA è universale tra gli organismi multicellulari.

Oltre alla mappatura di nuovi microRNA, esperimenti successivi condotti da diversi gruppi di ricerca hanno chiarito i meccanismi che governano la produzione dei microRNA e il loro indirizzamento alle sequenze bersaglio complementari sui mRNA, con la conseguente regolazione della sintesi proteica. In particolare, il legame dei microRNA porta all'inibizione di quest'ultima e alla degradazione degli mRNA. Curiosamente, un singolo microRNA può regolare l'espressione di molti geni diversi e, al contrario, un singolo gene può essere regolato da molteplici microRNA, con la possibilità di coordinare finemente l'espressione di intere reti di geni.

La regolazione genica tramite microRNA, rivelata per la prima volta da Ambros e Ruvkun, è in atto da centinaia di milioni di anni. Questo meccanismo ha consentito l'evoluzione di organismi sempre più complessi. Sappiamo infatti dalla ricerca genetica che cellule e tessuti non si sviluppano normalmente senza microRNA, come nel caso dei geni *lin-4* e *lin-14*.

Il ruolo critico dei fenomeni regolativi operati dai microRNA può essere messo in evidenza dal fatto che difetti in tali processi (dovuti a mutazioni nei geni che codificano microRNA) possono contribuire a diverse patologie, come il cancro, alcuni disturbi congeniti dell'udito, della vista e dello sviluppo scheletrico.

Dunque, questa inaspettata e fondamentale scoperta nel piccolo verme *C. elegans* ha rivelato una nuova dimensione della regolazione genica, essenziale per tutte le forme di vita complesse.

Come suggerito da queste ultime osservazioni, la scoperta del mondo dei microRNA rivela vaste possibilità applicative in ambito clinico, basate su tecnologie già disponibili, cioè quelle capaci di determinare la presenza e la quantità di queste molecole nei fluidi biologici, come pure di realizzarne la sintesi. Al momento, le prospettive più promettenti riguardano soprattutto la messa a punto di test diagnostici e prognostici che valutano le alterate concentrazioni di specifici microRNA. Per esempio, approcci di questo tipo possono essere applicati a diversi tipi di tumori, a malattie ereditarie e congenite, a patologie epatiche e al diabete.

Paolo Tortora

(già Ordinario di Biochimica presso il Dipartimento di Biotecnologie e Bioscienze dell'Università degli Studi di Milano Bicocca)