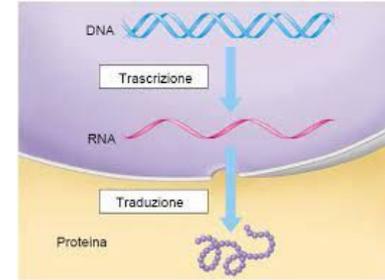


ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

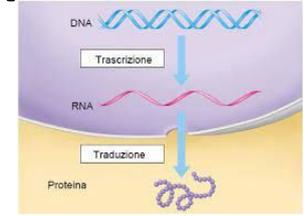
1. introduzione: importanza dell'RNA
2. Metodi di predizione
3. Allineamento globale per predizione di funzione: Dotlet
4. Algoritmo di predizione strutturale
 - MFLOD
 - RNAfold
 - RNAstructure
5. Algoritmo di Nussinov
6. Metodo di previsione basato sull'evoluzione
7. Metodi di previsione basati sulla simulazione
8. Limitazioni
9. Prospettive future



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

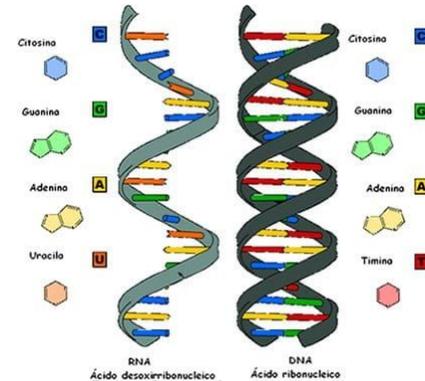
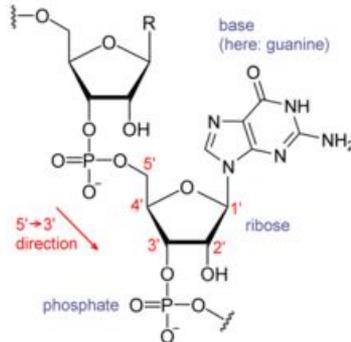
1. Introduzione: importanza dell'RNA

L'RNA (acido ribonucleico) è una molecola a filamento singolo composta da una sequenza di nucleotidi che si legano tramite legami fosfodiesterici.



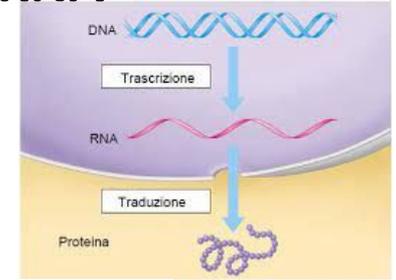
La sua struttura è simile a quella del DNA (acido desossiribonucleico), con la differenza che l'RNA presenta uno zucchero-ribosio al posto del desossiribosio e una base azotata uracile al posto della timina.

L'RNA svolge diverse funzioni all'interno delle cellule, come la sintesi proteica, la regolazione dell'espressione genica e il trasporto dell'informazione genetica.



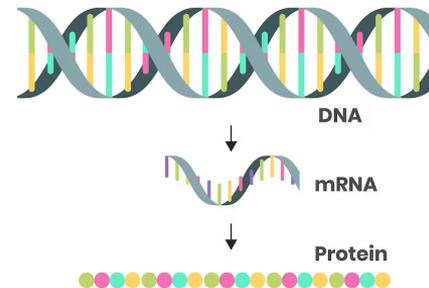
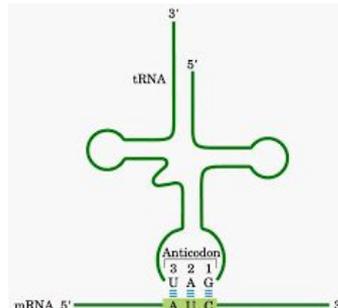
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

1. Introduzione: importanza dell'RNA



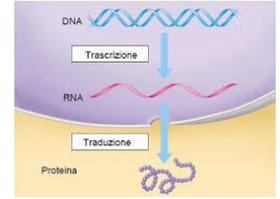
Esistono **quattro classi** particolarmente importanti di RNA che sono:

- mRNA (messenger RNA): è una sequenza che codifica per la formazione di una o più proteine.
- tRNA (transfer RNA): di solito piccole sequenze (circa 80 basi), che trasportano gli amminoacidi fino al ribosoma, nel quale avviene la traduzione dell'mRNA in proteina.
- rRNA (ribosomal RNA): sono le sequenze che costituiscono assieme alle proteine il ribosoma stesso.
- RNA virale: sono i codici di alcuni virus che lo utilizzano per replicarsi



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

1. Introduzione: importanza dell'RNA



È importante poter predire la struttura dell'RNA per diverse ragioni:

1. Comprendere la funzione dell'RNA: la struttura dell'RNA può influenzare la sua **capacità** di **tradurre** l'informazione genetica in proteine e di regolare l'espressione genica.
2. Identificare **nuovi bersagli terapeutici**: la previsione della struttura dell'RNA può aiutare a identificare nuovi bersagli terapeutici per malattie come il cancro e le malattie infettive.
3. Progettare **farmaci**: la conoscenza della struttura dell'RNA può aiutare a progettare farmaci mirati che interagiscono in modo specifico con l'RNA.

Un esempio di applicazione pratica potrebbe essere la ricerca di nuovi farmaci antivirali. La conoscenza della struttura dell'RNA virale può aiutare gli scienziati a sviluppare farmaci che inibiscono specificamente il meccanismo di replicazione virale, impedendo così la diffusione dell'infezione.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

1. Introduzione: importanza dell'RNA

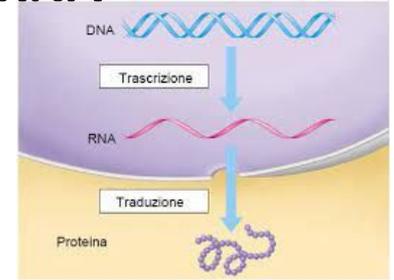
Le sequenze di RNA non sono semplicemente lunghi filamenti ma si accoppiano tra basi con **legami ad idrogeno**, minimizzando la loro **energia libera**.

In questo modo producono strutture tridimensionali. Se consideriamo di visualizzare gli accoppiamenti, e consideriamo che una base può solo accoppiarsi unicamente ad un'altra base avremo la struttura secondaria dell'RNA che possiamo visualizzare come:

Filamento di RNA = AGGAACGATGTCGUCA

Possibile struttura secondaria

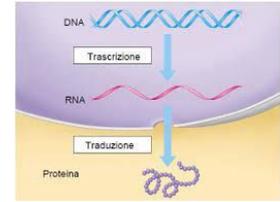
AA CG
G = C T = U
G = G G = C
A AT A



dove abbiamo utilizzato per gli accoppiamenti tra le basi complementari il simbolo =. Gli accoppiamenti possibili sono G=C e T=A (raramente G=C), e i rispettivi simmetrici.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

1. Introduzione: importanza dell'RNA



La stabilità di queste strutture dipende da vari fattori tra cui:

- Il numero di coppie G=C rispetto a quelle T=A. In quanto G=C sono più favorite energeticamente facendo 3 legami idrogeno rispetto a T=A che ne fanno due soltanto.
- Il numero di basi accoppiate rispetto ai loop (singolo filamento).
- La lunghezza e la forma dei loop.

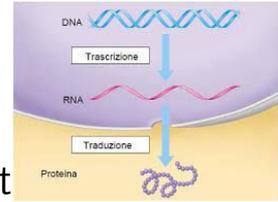
La stabilità di una struttura secondaria è quindi associata a questi fattori che concorrono al computo dell'energia libera del sistema.

Per cui più è **negativa l'energia** di una struttura e maggiormente diviene probabile trovare l'RNA in quella struttura.

L'**obiettivo** degli **algoritmi** che servono per la predizione della struttura secondaria dell'RNA è quello di trovare la configurazione (o le configurazioni) con il minimo dell'energia libera.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

1. Introduzione: importanza dell'RNA



I metodi di predizione della struttura dell'RNA si basano su algoritmi e modelli matematici che analizzano la sequenza di nucleotidi dell'RNA e utilizzano informazioni sui legami elettrostatici, idrogeno e idrofobici che si formano tra le basi per determinare la forma tridimensionale dell'RNA.

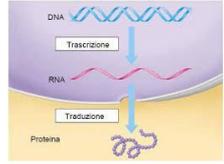
Ci sono diversi metodi di predizione della struttura dell'RNA, tra cui:

1. **Metodi di previsione termodinamica:** questi metodi si basano sul calcolo dell'**energia libera di Gibbs** necessaria per stabilire una particolare struttura dell'RNA. Utilizzando un modello di equilibrio termodinamico, tali metodi determinano la struttura più stabile dell'RNA in condizioni fisiche specifiche.

2. **Metodi di previsione basati sull'evoluzione:** questi metodi confrontano le sequenze di RNA simili attraverso la diversità evolutiva per individuare regioni altamente conservate che sono importanti per la struttura dell'RNA.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

1. Introduzione: importanza dell'RNA

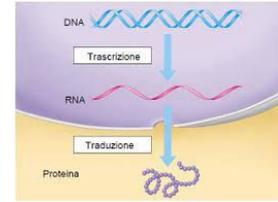


3. Metodi di previsione basati sulla condivisione di informazioni: questi metodi utilizzano **informazioni sperimentali**, come dati cristallografici e di risonanza magnetica nucleare, per determinare la struttura tridimensionale dell'RNA.

4. Metodi di previsione basati **sulla simulazione**: questi metodi utilizzano la modellizzazione molecolare e la simulazione per prevedere la struttura dell'RNA. Tali metodi richiedono l'uso di computer ad alte prestazioni per simulare i movimenti delle basi dell'RNA nel tempo.

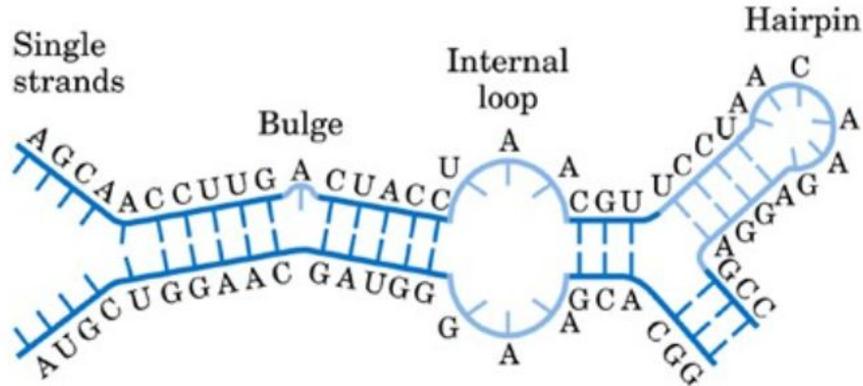
I metodi di predizione della struttura dell'RNA sono importanti per comprendere come l'RNA svolge i suoi vari ruoli biologici, come la regolazione genica e la traduzione proteica. Tuttavia, è importante sottolineare che le previsioni della struttura dell'RNA possono essere influenzate da molti fattori, tra cui la presenza di altri fattori di legame proteici, la flessibilità dell'RNA, le condizioni sperimentali e la popolazione di diverse conformazioni strutturali.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA



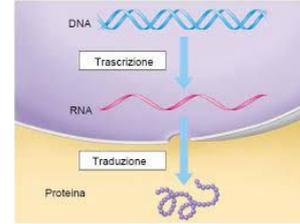
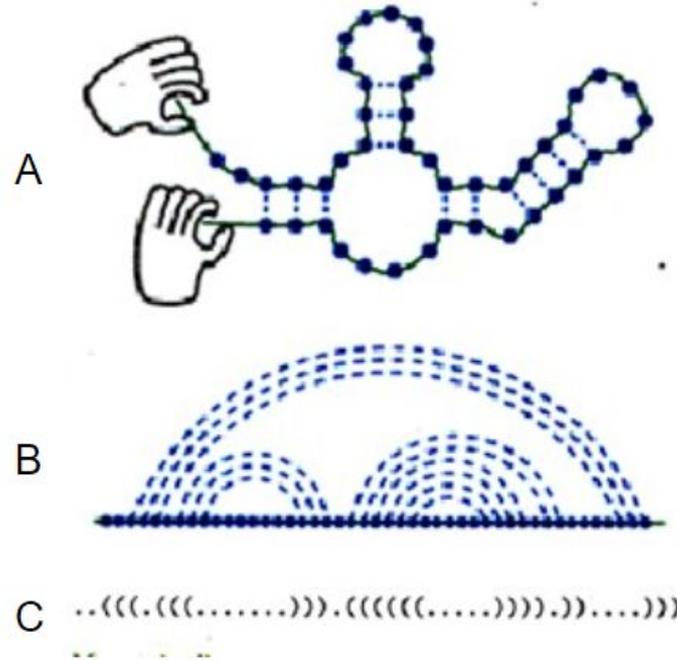
Predizione della struttura secondaria

- ▶ **Regioni a singolo filamento;**
- ▶ **stem:** coppie contigue di nucleotidi complementari;
- ▶ **hairpin loop:** è un loop alla fine di uno stem;
- ▶ **bulge loop:** interruzione di uno stem in un solo lato;
- ▶ **internal loop:** interruzione di uno stem su entrambi i lati;
- ▶ **multi-branched loop:** interruzione fra tre o più stem divergenti;



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- ▶ Gli appaiamenti di basi sono sempre annidati nella struttura secondaria.
- ▶ Questo implica che se si immagina di “aprire” un grafo planare (A) si ottiene un diagramma (B) in cui gli archi non sono mai intrecciati tra di loro.
- ▶ Questo fa sì che la struttura secondaria possa essere scritta con una notazione in parentesi (C) senza che vi sia ambiguità;



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

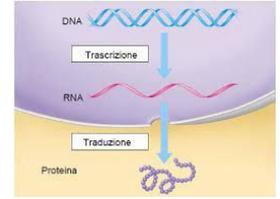
3. Allineamento globale per predizione di funzione: Dotlet

Inoltre la predizione della struttura può derivare anche da un allineamento globale o locale.

Ad esempio DOTLET è un programma di allineamento globale di sequenze di acidi nucleici o proteine, utilizzato per la visualizzazione grafica dell'allineamento. In particolare, DOTLET rappresenta le sequenze allineate come una serie di punti su una griglia bidimensionale, dove l'asse X rappresenta una sequenza e l'asse Y rappresenta l'altra. I punti rappresentano le corrispondenze tra le due sequenze.

DOTLET è spesso utilizzato per confrontare due sequenze simili e identificare le differenze tra di loro.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA



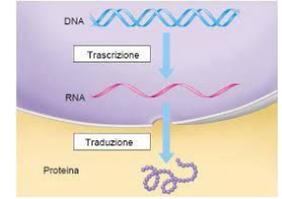
- ▶ **DOTLET** può fornire informazioni su porzioni di sequenza complementari.
- ▶ La sequenza nucleotidica riportata lungo un asse viene confrontata con la stessa sequenza, riportata lungo l'altro asse, invertita e complementata.
- ▶ Nella matrice Dot-Plot **regioni corrispondenti agli stem appaiono come segmenti perpendicolari alla diagonale principale**.
- ▶ Il parametro Windows-Size in questo caso fissa in pratica la lunghezza dello stem minimo.

Vantaggi: consente di mettere in evidenza contemporaneamente tutti i possibili appaiamenti all'interno di una molecola di RNA.

Svantaggi: non consente di determinare una particolare struttura secondaria.

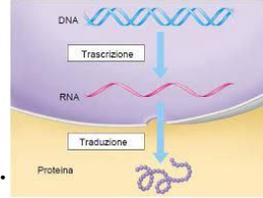
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- ▶ **Predizione:** La tendenza naturale di una molecola di RNA è quello di raggiungere la sua conformazione più stabile. Possiamo plausibilmente assumere di voler:
 - ▶ Massimizzare gli accoppiamenti (Algoritmo di Nussinov). Piuttosto semplicistico;
 - ▶ Minimizzare l'energia (Algoritmo di Zuker);
 - ▶ La stabilità termodinamica di una molecola di RNA ripiegata può essere misurata in termini di variazioni di energia libera tra la molecola a singolo filamento e la molecola ripiegata in una struttura secondaria. Tale variazione dipende dalla sequenza, dalla temperatura e dalla forza ionica;
 - ▶ La variazione di energia libera globale è pari alla somma dei contributi indipendenti dei motivi elementari di struttura;
 - ▶ Tabelle di Freier dei contributi energetici: **Contribuiscono in modo stabilizzante le energie tra basi appaiate e in modo destabilizzante i loops ;**



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

4. Algoritmo di predizione strutturale: MFOLD



Un algoritmo di predizione strutturale basato sulla minimizzazione dell'energia è MFOLD. Utilizzato per prevedere la struttura secondaria di una singola molecola di RNA.

Fa uso di un algoritmo di **minimizzazione dell'energia**, che considera la formazione di legami idrogeno e altri tipi di interazioni fisiche e chimiche che influenzano la formazione della struttura a doppio filamento dell'RNA.

Il Modello di Mfold tiene conto della **stabilità della struttura** dell'RNA, ossia dell'**energia libera di Gibbs** necessaria per rompere le interazioni tra le basi appaiate.

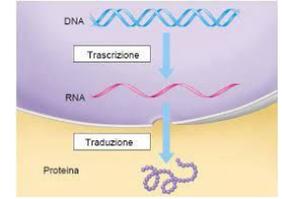
Per prevedere la struttura dell'RNA con il Modello di Mfold, vengono utilizzati diversi parametri termodinamici, come la temperatura, la concentrazione di ioni, il pH e la presenza di molecole di solvente.

Il Modello di Mfold valuta quindi la stabilità di ogni possibile struttura dell'RNA e fornisce quella con l'energia libera di Gibbs minima.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>



- ▶ Ciascuno dei motivi elementari di struttura secondaria contribuisce alla stabilità della struttura secondaria e quindi alla determinazione della sua energia libera.
- ▶ Gli stem sono elementi stabilizzati (contributo negativo dell'energia libera) mentre tutte le regioni a singolo filamento sono destabilizzanti (contributo positivo o nullo dell'energia libera).
- ▶ Data una sequenza di RNA, **Mfold restituisce le strutture a minima energia più probabili**, dato che la struttura biologicamente corretta è di solito sub-ottimale, piuttosto che quella a minima energia (Algoritmo di Zuker);

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

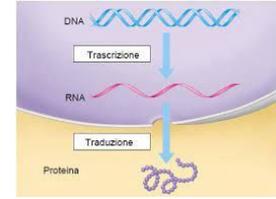
MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>

- Possiamo copiare una sequenza in formato FASTA (e assegnarle un nome);

- Consideriamo la sequenza

>Haemophilus_influenzae_Rd.trna49-
AlaGGC (307354-307279) Ala (GGC) 76
bp Sc: 85.98
GGGGATATAGCTCAGTTGGGAGAGCGC
TTGAATGGCATTCAAGAGGTCGTCCGT
TCGATCCCATTATCTCCACCA



RNA Folding Form

The folding temperature is fixed at 37°. You may still fold with the older version 2.3 RNA parameters.
[DNA mfold server](#) [Quikfold](#). Fold many short RNA or DNA sequences at once.

Enter a name for your sequence:

Enter the sequence to be folded in the box below. All non-alphabet characters will be removed. FASTA

```
GGGGATATAGCTCAGTTGGGAGAGCGCTTGAATGGCATTCAAGAGGTCGTCCGTTCGATC  
CCGATTATCTCCACCA
```

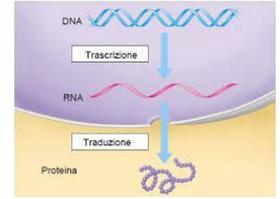
Enter [constraint information](#) in the box at the right. (optional) You may:

1. force bases $i, i+1, \dots, i+k-1$ to be double stranded by entering:
F I O K on 1 line in the constraint box.
2. force consecutive base pairs $i, i+1, j-1, \dots, i+k-1, j-k+1$ by entering:
F I I K on 1 line in the constraint box.
3. force bases $i, i+1, \dots, i+k-1$ to be single stranded by entering:
P I O K on 1 line in the constraint box.
4. prohibit the consecutive base pairs $i, i+1, j-1, \dots, i+k-1, j-k+1$ by entering:
P I I K on 1 line in the constraint box.
5. prohibit bases i to j from pairing with bases k to l by entering:
P H K l on 1 line in the constraint box.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>



Output

The energy dot plot for tRNA. [\(Definition\)](#)
File formats: [Text](#), [PostScript](#), [pdf](#), [png](#), [log](#)

Computed Structures: [\(File Formats\)](#)
The computed foldings contain 28 base pairs out of 28 (100.0%) in the energy dot plot.

Extra files: [sorted](#) ct file; [h-num](#) values; [p-num](#) values; [log](#) file for main computations.

Choose zipped file: or gzipped tar file: [PostScript](#)

View ss-count information

[\(Definition\)](#) [\(ss-count file\)](#) ss value = 0.79 ± 0.85

Averaging window Magnification Base to magnify about Plot format [PostScript](#)

View Individual Structures:

[Circular structure Plots](#)

- [Structure 1](#): Initial dG = -23.40 kcal/mol. [\(Thermodynamic Details\)](#)
Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [ct file](#), [Vienna](#), [RNAML](#), [RnaViz.ct](#), [Mac.ct](#), [RNAdraw](#), [XRNA.ss](#)
- [Structure 2](#): Initial dG = -23.00 kcal/mol. [\(Thermodynamic Details\)](#)
Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [ct file](#), [Vienna](#), [RNAML](#), [RnaViz.ct](#), [Mac.ct](#), [RNAdraw](#), [XRNA.ss](#)

Dot plot folding comparison for tRNA:
Image format: [PostScript](#)

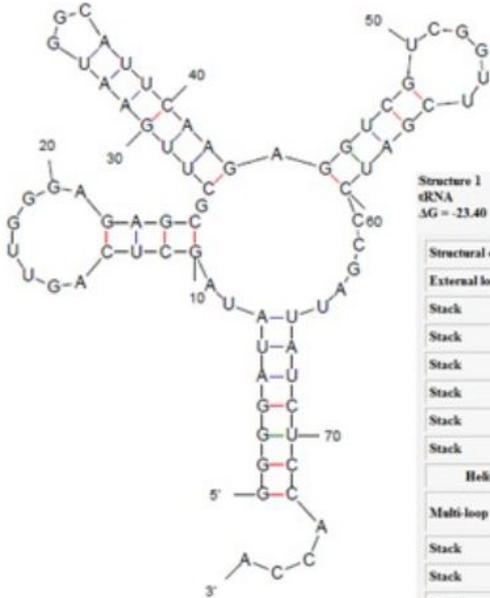
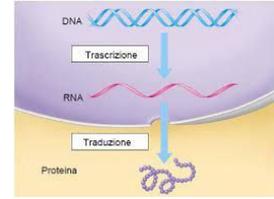
Compare selected foldings: 1 2

Sono state predette due strutture:
Sono ordinate in base alla stabilità.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>



Cliccando ad esempio sul formato JPEG viene visualizzata la struttura predetta

Structure 1
sRNA
 $\Delta G = -23.40$

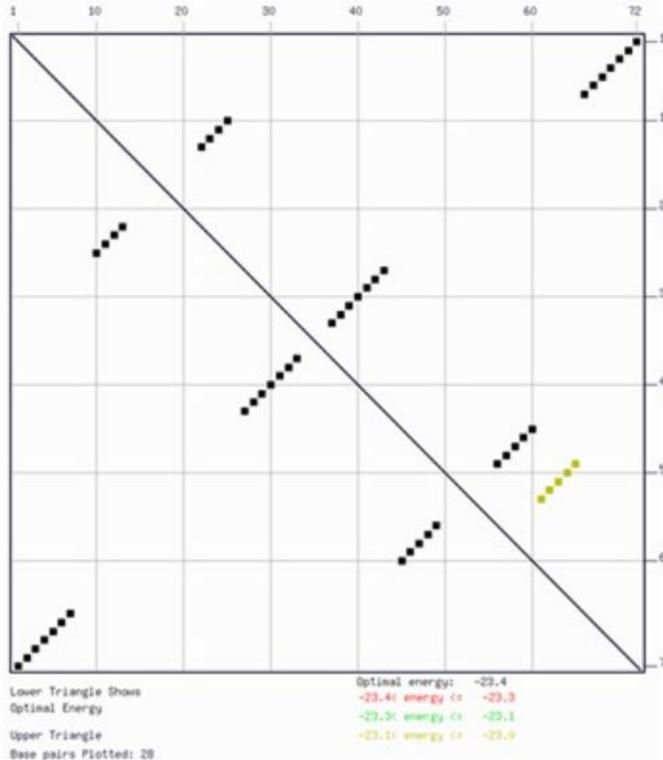
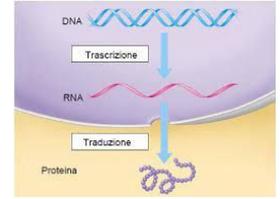
Structural element	ΔG	Information
External loop	-1.70	4 ss bases & 1 closing helices.
Stack	-3.30	External closing pair is G ₁ -C ₇₁
Stack	-1.50	External closing pair is G ₂ -C ₇₁
Stack	-2.10	External closing pair is G ₃ -U ₇₀
Stack	-2.40	External closing pair is G ₄ -C ₆₉
Stack	-1.10	External closing pair is A ₈ -U ₆₈
Stack	-1.30	External closing pair is U ₆ -A ₆₇
Helix	-11.70	7 base pairs.
Multi-loop	0.90	External closing pair is A ₇ -U ₆₆ 9 ss bases & 4 closing helices.
Stack	-1.50	External closing pair is G ₆₂ -C ₆₀
Stack	-1.40	External closing pair is G ₆₁ -U ₅₉
Stack	-2.40	External closing pair is U ₆₇ -A ₆₃
Stack	-2.40	External closing pair is C ₆₃ -G ₆₇
Helix	-7.70	5 base pairs.
Hairpin loop	3.90	Closing pair is G ₆₉ -C ₆₅

Dettagli termodinamici

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>



Energy Dot Plot:

Mostra gli stem che fanno parte di un fold ottimale.

Uno stem è mostrato come una diagonale nera perpendicolare alla diagonale principale

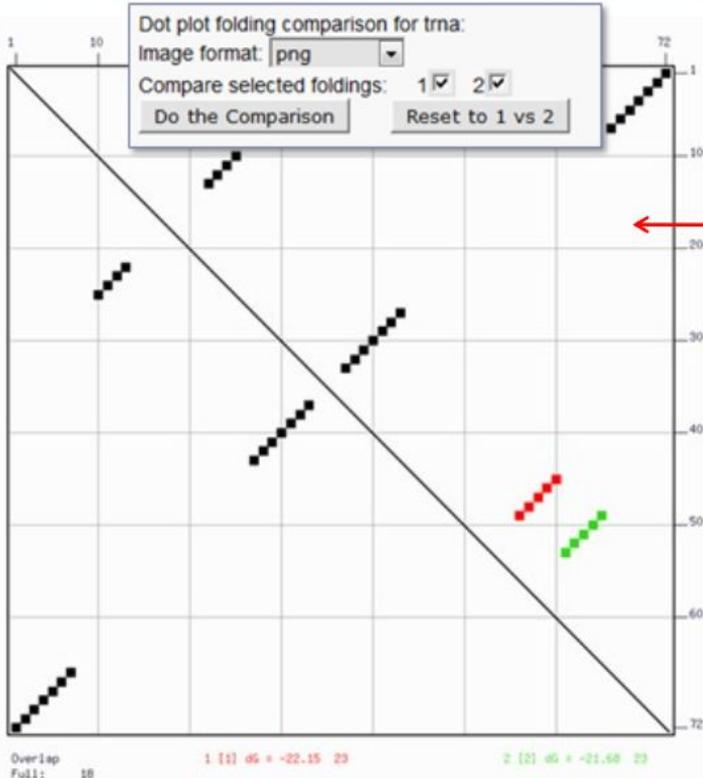
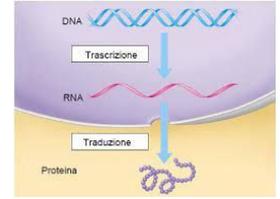
Interazioni distanti sono mostrate come diagonali molto vicine all'angolo in alto a destra o in basso a sinistra.

Diagonali colorate indicano strutture secondarie subottimali. (Red=best, Yellow=worst)

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>



DOT PLOT FOLDING:

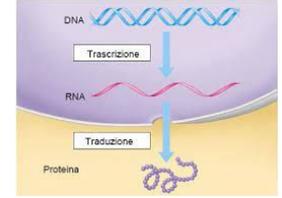
Confronta il fold di tutte le le soluzioni prodette.

Gli elementi colorati corrispondono agli stem che occorrono solo in una predizione.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

MFOLD

<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>



PARAMETRI:

Se si dispone di alcuni dati sperimentali e si sa che delle coppie di basi specifiche si accoppiano è possibile forzare MFOLD ad utilizzare queste informazioni.

Enter [constraint information](#) in the box at the right. (optional) You may:

1. force bases $i, i+1, \dots, i+k-1$ to be double stranded by entering:
F i 0 k on 1 line in the constraint box.
2. force consecutive base pairs $i, j, i+1, j-1, \dots, i+k-1, j-k+1$ by entering:
F i j k on 1 line in the constraint box.
3. force bases $i, i+1, \dots, i+k-1$ to be single stranded by entering:
P i 0 k on 1 line in the constraint box.
4. prohibit the consecutive base pairs $i, j, i+1, j-1, \dots, i+k-1, j-k+1$ by entering:
P i j k on 1 line in the constraint box.
5. prohibit bases i to j from pairing with bases k to l by entering:
P i j k l on 1 line in the constraint box.

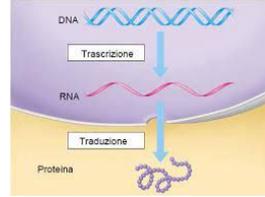


F 10 22 3

Ad esempio se sappiamo che i nucleotidi 10,11,12 interagiscono rispettivamente con i nucleotidi 22,21,20 possiamo indicarlo a MFOLD come in figura (F=Force, P=Prohibit).

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

4. Algoritmo di predizione strutturale: MFOLD



Modello di Mfold: questa tecnica può essere utile per fornire informazioni sulla stabilità e la funzione dell'RNA e per identificare regioni di struttura conservata tra diverse specie.

Tuttavia, va tenuto presente che i metodi di predizione basati su modelli termodinamici possono avere **limitazioni** nella previsione di struttura complessa e altamente dinamica dell'RNA. Questo perché spesso i modelli termodinamici assumono che le molecole di RNA siano rigide e statiche.

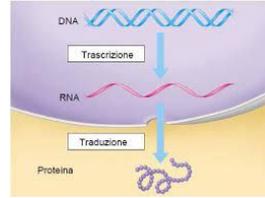
Infatti, molte molecole di RNA hanno **strutture dinamiche** che possono cambiare rapidamente e in modo complesso in risposta a diversi segnali ambientali. Inoltre, la presenza di legami chimici non canonici e di basi modificate può rendere difficile la previsione accurata della struttura dell'RNA utilizzando solo metodi termodinamici.



Pertanto, altre tecniche, come la modellistica di dinamica molecolare e l'analisi di modifiche post-trascrizionali, potrebbero essere necessarie per ottenere previsioni più accurate della struttura dell'RNA.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

4. Algoritmo di predizione strutturale: RNAfold



Un algoritmo di predizione strutturale basato sulla **minimizzazione dell'energia** è RNAfold, un software di predizione della struttura dell'RNA basato sull'algoritmo del minimum free energy (MFE) che richiede la minima energia libera di **Gibbs**, il che significa che è la struttura che richiede la minima quantità di energia per formarsi.

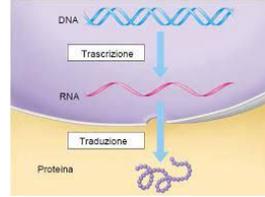
RNAfold è in grado di prevedere la struttura primaria dell'RNA, ovvero la sua successione dei nucleotidi, e la struttura secondaria dell'RNA, ovvero la configurazione complessiva delle basi appaiate e appaiate non appaiate nella molecola di RNA.



<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

4. Algoritmo di predizione strutturale: RNAfold



Esercizio:

Hai a disposizione una sequenza di DNA di 50 nucleotidi e vorresti predire la struttura dell'RNA che sarà prodotto trascurando eventuali interazioni con proteine (utilizzando il programma RNAfold).

Sequenza di DNA: AGTCGAGCTGATCGTAGCCAGTACGTAGCTGCATCGTAGCTAGCTGATC



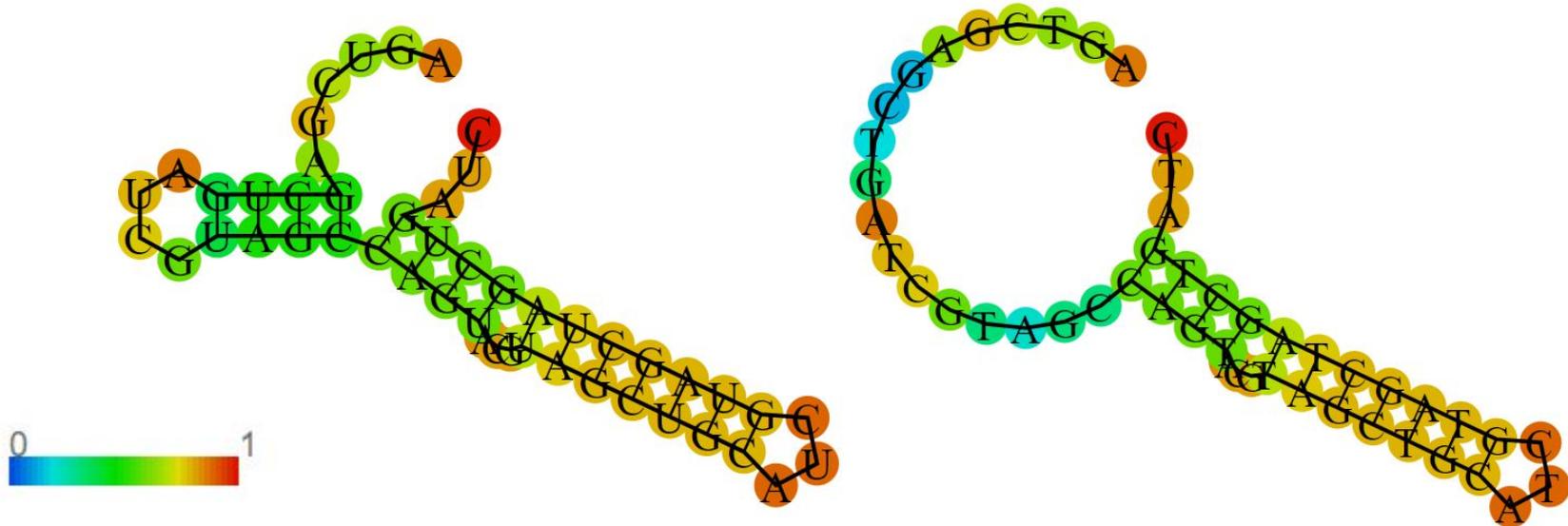
<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

4. Algoritmo di predizione strutturale: RNAfold

Uscita grafica

Puoi guardare il disegno interattivo della struttura MFE qui sotto. Se non vedi il disegno interattivo e stai utilizzando Internet Explorer, installa il [plug-in Adobe SVG](#) . **Una nota sulle probabilità di accoppiamento delle basi:** la struttura sottostante è colorata dalle probabilità di accoppiamento delle basi. Per le regioni non accoppiate il colore indica la probabilità di essere non accoppiate.

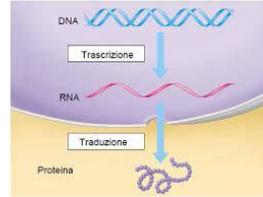


ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

4. Algoritmo di predizione strutturale: RNAstructure

RNAstructure è un software che permette di analizzare la struttura e la stabilità termodinamica delle molecole di RNA.

Il programma utilizza una combinazione di algoritmi per la predizione della struttura e la simulazione delle dinamiche molecolari, e può essere utilizzato per studiare sia RNA singoli che complessi di RNA.



<https://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html>

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

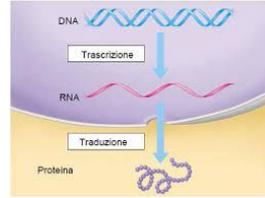
4. Algoritmo di predizione strutturale: RNAstructure

Un esempio pratico di utilizzo di RNAstructure potrebbe riguardare la predizione della struttura di una specifica sequenza di RNA, ad esempio un RNA messaggero (mRNA).

Supponiamo di avere la seguente sequenza di mRNA:
AGGUAUCUGUCUAUUCCUUGCAGUACUUAGUUCGCUCU

Per predire la struttura di questo mRNA utilizzando RNAstructure, è possibile eseguire i seguenti passaggi:

- Aprire il software RNAstructure e caricare la sequenza mRNA
- Selezionare le impostazioni desiderate, ad esempio la temperatura di predizione e il tipo di algoritmo da utilizzare.
- Avviare la predizione della struttura cliccando sul pulsante "Prevedi struttura"
- Dopo alcuni secondi, RNAstructure restituirà una rappresentazione grafica della struttura predetta per l'mRNA.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

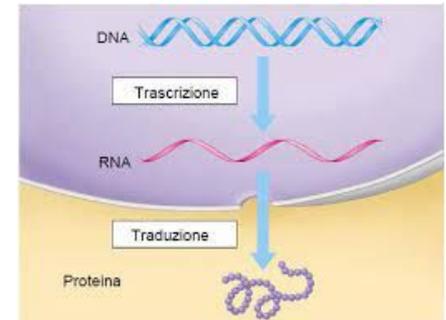
5. Algoritmo di Nussinov

Uno dei primi algoritmi per la predizione della struttura dell'RNA ad essere sviluppato è stato l'algoritmo di Nussinov.

Un algoritmo utilizzato per calcolare la struttura a pieghe delle molecole di RNA (ovvero la forma tridimensionale assunta dall'RNA a causa dell'interazione tra le basi azotate del suo nucleotide).

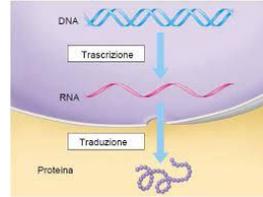
L'algoritmo si basa sulla ricerca delle **coppie di basi complementari** all'interno della molecola di RNA e sulla formazione di una struttura a pieghe di energia minima utilizzando un algoritmo di programmazione dinamica.

La struttura a pieghe è importante perché determina la funzione dell'RNA e il modo in cui interagisce con altre molecole all'interno della cellula.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

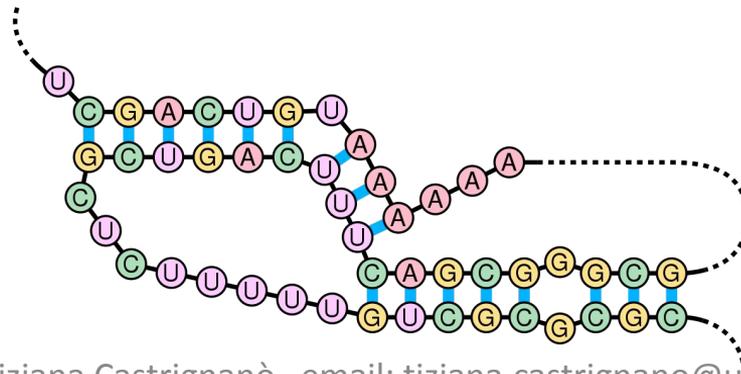
5. Algoritmo di Nussinov



L'algoritmo di **Nussinov**, come la maggior parte degli algoritmi che cercano di predire la struttura secondaria dell'rna, trova strutture secondarie *palindrome*. Per cui gli *pseudoknots* non sono predicibili con questi approcci.

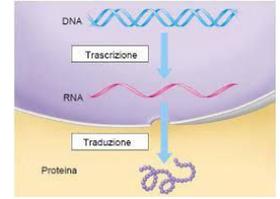
→ **Pseudoknots**: strutture a forma di nodo presenti nell'RNA che si formano quando alcune parti del filamento si accavallano e si legano tra loro tramite ponti di idrogeno.

Gli pseudoknots giocano un ruolo importante nella struttura e nella funzione dell'RNA, influenzando la sua stabilità, la struttura tridimensionale e la capacità di interagire con altre macromolecole cellulari.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

5. Algoritmo di Nussinov



L'algoritmo di Nussinov cerca di massimizzare il numero di coppie, o più in generale un punteggio assegnato a ciascuna possibile coppia base(i),base(j).

La più semplice funzione è ovviamente quella di associare un valore di +1 se le due basi sono complementari e 0 altrimenti.

Se definiamo rna la nostra sequenza di basi e vogliamo valutare il punteggio del possibile accoppiamento tra le basi in posizione i e j , possiamo utilizzare la funzione definita in Figura.

```
def score_bpair(rna, i, j, tol=2):
```

Qui abbiamo anche introdotto tol , che serve ad evitare di calcolare legami tra basi la cui separazione lungo la sequenza ($|i-j| \geq tol$) sia troppo corta (per poter fare un legame richiediamo almeno un loop lungo una base).

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

5. Algoritmo di Nussinov

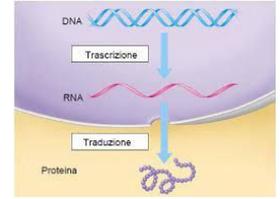
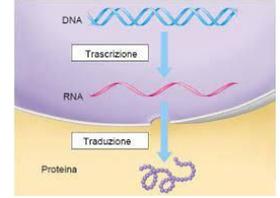


Figura: Funzione per il calcolo del punteggio come basi complementari

```
def score_bpair(rna,i,j,tol=2):  
    ''' score_bpair(rna,i,j,tol=2)  
        compute the score between the elemnt in position i and j  
        in the rna sequence  
        tol is the minimal separation we do not want to count  
    '''  
    score={('A','U'):1.0, ('U','A'):1.0, ('C','G'):1.0, ('G','C'):1.0}  
    a=rna[i]  
    b=rna[j]  
    if(abs(i-j)<tol or ((a,b) not in score.keys())):  
        return 0.0  
    else:  
        return score[(a,b)]
```

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

6. Metodo di predizione basato sull'evoluzione



Vediamo adesso i metodi di predizione basati sull'evoluzione. Un esempio è la **coevoluzione delle coppie di basi**.

Questo metodo si basa sull'idea che alcune coppie di basi nell'RNA sono conservate attraverso l'evoluzione perché sono necessarie per mantenere la struttura e la funzione del RNAs.

Pertanto, se due basi sono coevolute, significa che probabilmente sono fisicamente vicine nella struttura dell'RNA.

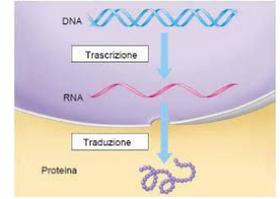
Il metodo consiste nel confrontare le sequenze di RNA di diverse specie che sono imparentate evolutivamente e utilizzano un algoritmo per identificare le coppie di basi che sono coevolutive. Questo processo aiuta a prevedere quale siano le coppie di basi che sono strettamente legate nella struttura dell'RNA. Una volta identificate queste coppie di basi, si possono utilizzare queste informazioni per predire la struttura dell'RNA.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

6. Metodo di previsione basato sull'evoluzione

Un esempio di metodo di previsione basato sulla condivisione di informazioni per la predizione della struttura dell'RNA è il metodo di "predizione di coerenza dell'evoluzione".

Come abbiamo detto, questo metodo si basa sulla condivisione di informazioni riguardanti le sequenze di RNA simili tra specie diverse, utilizzando l'evoluzione come indicatore della conservazione delle caratteristiche funzionali dell'RNA.



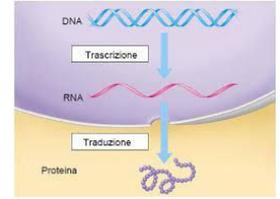
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

6. Metodo di previsione basato sull'evoluzione

In pratica, il metodo confronta le sequenze di RNA attraverso le diverse specie e identifica le regioni che mostrano una forte conservazione, indicando che queste regioni sono cruciali per la funzione dell'RNA.

Successivamente, il metodo si focalizza sui dettagli della struttura dell'RNA, utilizzando tecniche di modellazione molecolare per predire la struttura dell'RNA sulla base delle informazioni di conservazione identificate.

In questo modo, il metodo di predizione di coerenza dell'evoluzione è in grado di fornire previsioni altamente accurate sulla struttura dell'RNA, basate sulla condivisione di informazioni tra diverse specie.

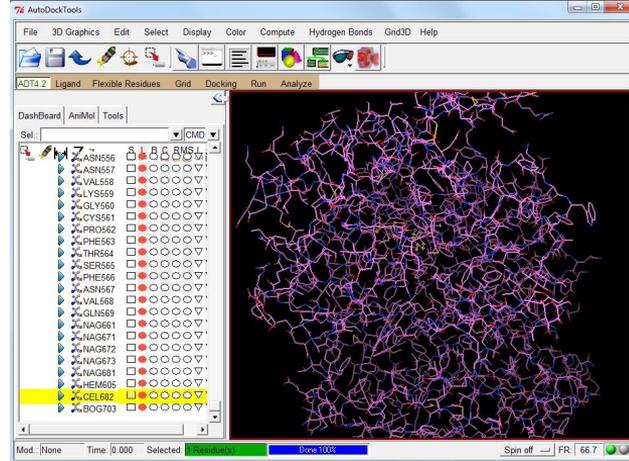
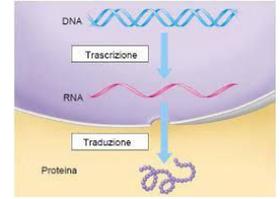


ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

7. Metodi di previsione basati sulla simulazione

I metodi di previsione della struttura dell'RNA basati sulla simulazione sono algoritmi informatici che utilizzano modelli matematici e fisici per simulare il modo in cui le molecole si comportano e interagiscono tra loro.

Questi metodi sono utilizzati per prevedere la struttura tridimensionale dell'RNA, il che è importante perché la struttura determina la funzione dell'RNA.

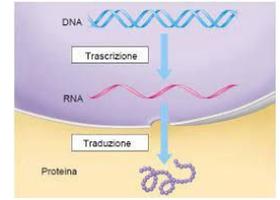


ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

7. Metodi di previsione basati sulla simulazione

Alcuni dei metodi più comuni di previsione della struttura dell'RNA basati sulla simulazione includono:

- la dinamica molecolare,
- la modellizzazione del ribosoma,
- la modellizzazione del folding dell'RNA,
- la modellizzazione del docking dell'RNA,
- la modellizzazione delle interazioni della proteina-RNA.



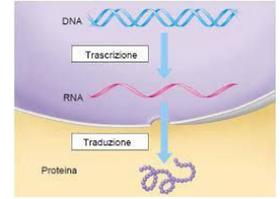
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

7. Metodi di previsione basati sulla simulazione

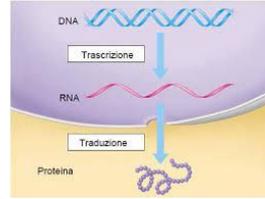
I metodi basati sulla **dinamica molecolare** per la previsione della struttura dell'RNA sono utilizzati per calcolare l'**evoluzione nel tempo** delle **posizioni** e dei **movimenti** degli **atomi** della molecola di RNA.

Questi metodi si basano su equazioni matematiche che descrivono la fisica dei movimenti delle particelle.

- Il primo passo consiste nella generazione di un modello tridimensionale della molecola di RNA, in cui vengono specificate le posizioni degli atomi nella struttura.
- Successivamente, il modello viene salvato in un ambiente in cui la simmetria sferica con un potenziale che simula l'interazione con il solvente viene utilizzata per considerare le interazioni elettrostatiche.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA



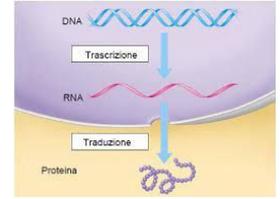
7. Metodi di previsione basati sulla simulazione

- Una volta salvata la molecola di RNA, si avvia la simulazione della dinamica molecolare, in cui la posizione degli atomi viene calcolata in funzione del tempo, utilizzando le leggi della fisica classica.
- I risultati della simulazione possono essere utilizzati per prevedere la struttura tridimensionale dell'RNA attraverso l'analisi delle traiettorie di movimento degli atomi.

L'analisi della dinamica molecolare può anche essere utilizzata per prevedere i movimenti e le interazioni di RNA con altre macromolecole, come proteine e acidi nucleici. Questi metodi sono utilizzati nel campo della biologia molecolare e genetica per comprendere il modo in cui gli RNA svolgono il loro ruolo nella formazione di proteine e nella regolazione dell'espressione genica.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- la modellizzazione del docking dell'RNA



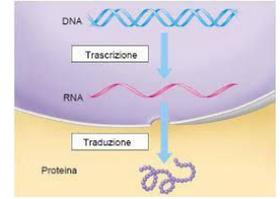
La modellizzazione del docking dell'RNA si riferisce alla creatività di un modello che descrive l'interazione tra un RNA e una proteina. Il docking si riferisce alla posizione e alla conformazione dell'RNA e della proteina quando sono legati.

La modellizzazione del docking dell'RNA può essere utilizzata per comprendere meglio come l'RNA e la proteina interagiscono tra loro e per aiutare a sviluppare nuovi farmaci o terapie basati sull'interazione tra RNA e proteine.

La modellizzazione del docking dell'RNA utilizza tecniche di bioinformatica, simulazione al computer e sperimentazione in vitro e in vivo per creare un modello accurato dell'interazione tra RNA e proteina.

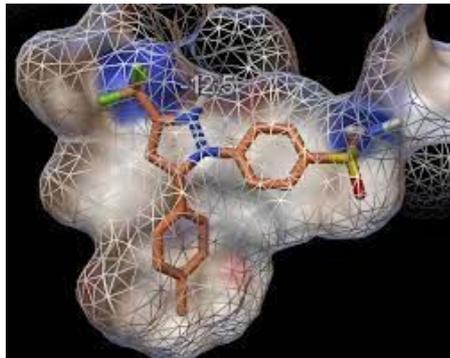
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- la modellizzazione del docking dell'RNA



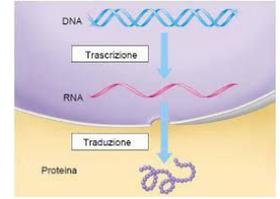
Una possibile modellizzazione del docking dell'RNA potrebbe essere la seguente:

1. Identificazione dei target di docking: vengono scelti i due RNA che si desidera far interagire e si definiscono le zone di interazione, ovvero quelle parti delle molecole che si uniranno.
2. Preparazione dei file di input: si creano i file di input per i software di simulazione molecolare, che includono le informazioni sulla struttura degli RNA e sulle coordinate dei siti di interazione.



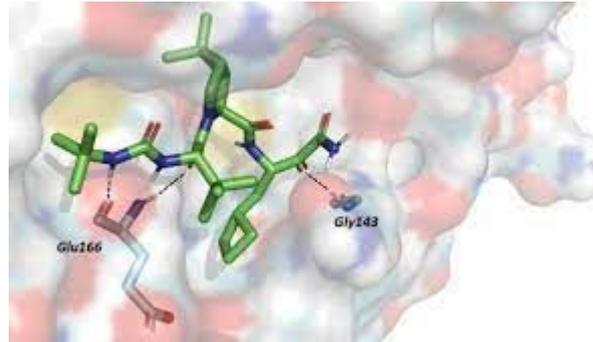
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- la modellizzazione del docking dell'RNA



3. Simulazione del docking: tramite i software di docking, si effettua la simulazione della formazione del complesso tra i due RNA. Vengono esplorate diverse conformazioni fino a trovare quella che presenta l'affinità e la complementarità maggiori.

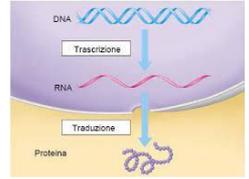
4. Analisi dei risultati: si analizzano i dati ottenuti dalla simulazione per valutare la bontà del docking. Vengono considerati parametri come la distanza tra i siti di interazione, l'energia di legame, la conformazione del complesso e la presenza di interazioni stabilizzanti.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- la modellizzazione del docking dell'RNA

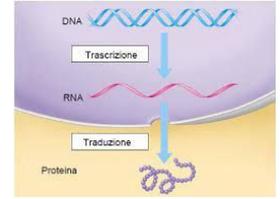
La dinamica molecolare e il docking molecolare sono strettamente correlati. La dinamica molecolare è una tecnica di simulazione che permette di studiare il movimento delle molecole nel tempo e nello spazio, mentre il docking molecolare si riferisce al processo di predizione dell'interazione tra due molecole, come un ligando e una proteina.



Il docking molecolare utilizza le informazioni strutturali per calcolare come due molecole si attireranno e si uniranno tra loro, mentre la dinamica molecolare può essere usata per studiare il complesso formato dalla proteina e dal ligando in varie condizioni, come ad esempio in diversi solventi o temperature.

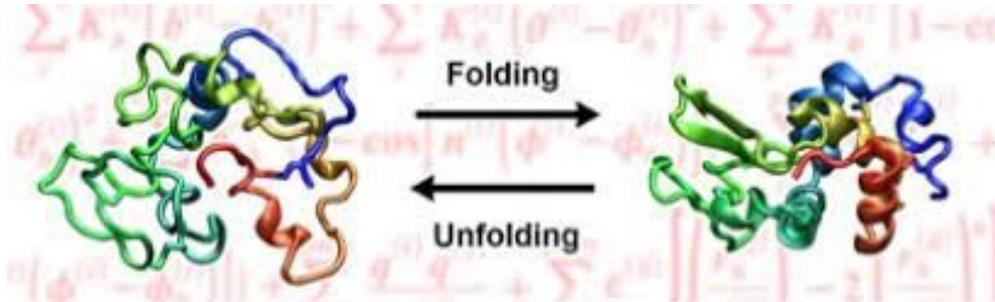
ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

- la modellizzazione del folding dell'RNA,



La modellizzazione del folding dell'RNA si riferisce alla creazione di modelli matematici e bioinformatici per comprendere come le molecole di RNA si piegano su se stesse per formare la loro struttura tridimensionale.

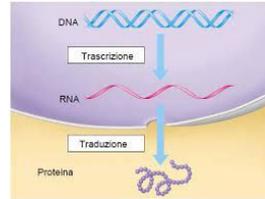
Questo processo è importante perché la struttura tridimensionale dell'RNA influisce sulle sue funzioni biologiche, come la capacità di legarsi a specifici target molecolari. La modellizzazione del folding dell'RNA è un'area di ricerca attiva in bioinformatica e biologia computazionale e utilizza approcci di modellizzazione molecolare per prevedere la struttura tridimensionale dell'RNA e simulazioni al computer per rivelare le dinamiche di folding.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

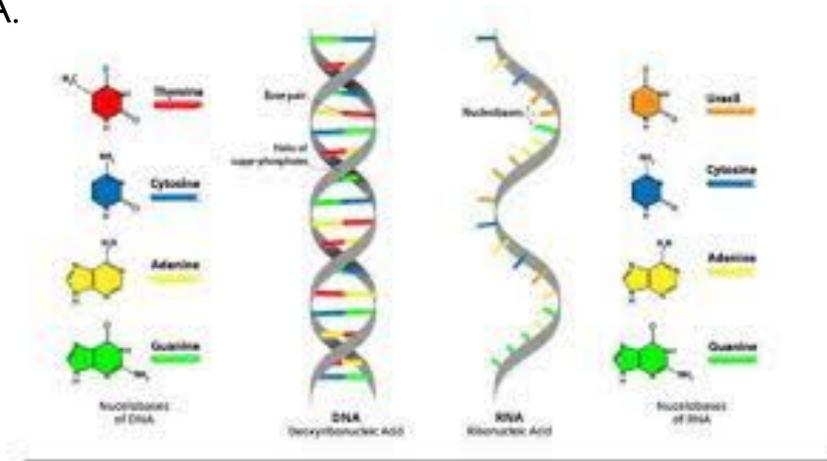
8. Limitazioni

In conclusione, la predizione della struttura dell'RNA è un campo in continua evoluzione che richiede l'uso di algoritmi complessi basati su principi termodinamici e di modellizzazione molecolare.



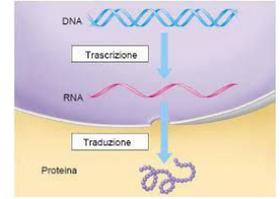
Tuttavia, è importante considerare le limitazioni delle attuali tecnologie di predizione, tra cui:

1) Variabilità nella sequenza delle regioni non codificanti e non conservate che influenzano la stabilità della struttura dell'RNA.



ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA

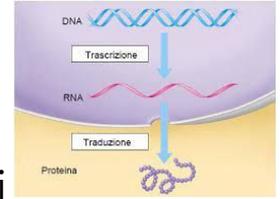
8. Limitazioni



- 2) La mancanza di informazioni sulla struttura tridimensionale dell'RNA, che può influire sulla predizione della struttura secondaria.
- 3) La difficoltà di prevedere legami intramolecolari complessi e interazioni tra diverse regioni dell'RNA.
- 4) L'effetto dell'ambiente cellulare sulla struttura dell'RNA, che può differire dalla struttura prevista in vitro.
- 5) La mancanza di dati sperimentali sulla natura delle modificazioni post-trascrizionali dell'RNA che possono influire sulla sua struttura.

In generale, gli algoritmi di predizione della struttura dell'RNA sono utili strumenti di analisi, ma è importante considerare i loro limiti e integrare i risultati con altre informazioni sperimentali per avere una visione completa della struttura e della funzione dell'RNA.

ALGORITMI DI PREDIZIONE DELLA STRUTTURA DELL'RNA



9. Prospettive future:

Le prospettive future per la predizione della struttura dell'RNA includono lo sviluppo di nuovi algoritmi e tecniche che possono prevedere con maggiore precisione la struttura dell'RNA a livello globale e di dettaglio.

Il successo della predizione della struttura dell'RNA potrebbe essere ampliato mediante l'impiego dell'intelligenza artificiale e dell'apprendimento automatico.

Inoltre, le tecnologie emergenti di imaging ad alta risoluzione consentiranno di ottenere nuovi dati sperimentali sulla struttura dell'RNA in situ, che potranno essere utilizzati per validare e migliorare le previsioni. Tutto ciò consentirà di meglio comprendere le funzioni dell'RNA e di sviluppare nuove terapie per le malattie associate alla disfunzione dell'RNA.