

lunedì 5 ottobre 2020

## Geni e genomi

---

### Elementi aggiuntivi rispetto agli appunti

- I virus hanno genomi estremamente semplici: molto più piccolo rispetto anche ai procarioti (NON hanno gli apparati di sintesi proteica: NON sono organismi viventi)
- Genoma procariotico -> genoma compatto (ovvero sequenze TUTTE codificanti: NON ci sono introni)
- Dimensione genomi eucariotico -> genoma molto più grande ma NON tutti i geni codificano per proteine

### Genoma eucariotico

**Paradosso del valore C** (n totale di nt di un genoma -> ovvero il contenuto del DNA): il contenuto di DNA NON è correlato in maniera lineare con la complessità del genoma di un organismo

**Paradosso del valore G:** il numero di geni presenti in un organismo NON è correlato in maniera lineare con la complessità del suo genoma

**Esperimento di re-annaeling** (ovvero esperimento di ibridazione degli acidi nucleici): la complessità del genoma è correlata linearmente al tempo necessario al suo DNA per rinaturarsi (ovvero alla *velocità di rinaturazione*)

Tecnica: purificazione DNA dei vari organismi (virus, procarioti ed eucarioti), frammentazione del DNA, denaturazione DNA -> seguire allo spettrofotometro la rinaturazione

La curva di rinaturazione del genoma eucariotico è diversa dalle altre: si osservano 3 diverse velocità di rinaturazione -> 3 frazioni differenti (corrispondenti a sequenze diverse) nel genoma degli eucarioti:

- Frazione **veloce** -> DNA altamente ripetuto (elementi parassitici le cui informazioni sono rimasti nel nostro genoma)
- Frazione **intermedia** -> sequenze presenti in un numero di copie abbastanza elevato (per es. geni che codificano rRNA: c'è bisogno di tanti ribosomi quindi c'è bisogno di molte copie in modo tale che la trascrizione sia molto più efficiente)

lunedì 5 ottobre 2020

- Frazione **lenta** -> sequenze *uniche* (la maggior parte dei geni codificanti proteine)

La densità genica degli eucarioti è bassissima

## Frazione lenta: sequenze *uniche*

Le sequenze *uniche* rappresentano il **50%** del nostro genoma, di cui solo il **1,5% codifica per proteine** -> quindi la frazione del genoma che codifica per proteine è minima!

Rappresenta la porzione del genoma in copia singola o presente in un numero di copie molto basso

Contiene esoni e introni -> infatti il **genoma eucariote è discontinuo** (NON continuo come quello di Coli) e **NON compatto** (gli introni NON hanno una funzione attualmente ben chiara)

Si parla di genoma interrotto, però quei geni che sono più espressi solitamente sono più compatti

**Pseudogeni** -> considerati sequenze *uniche*: durante l'evoluzione si sono formate 2 o più copie di uno stesso gene (per *duplicazione*), 1 dei due geni duplicati NON viene espresso e accumula mutazioni -> quindi sono geni spenti, NON funzionali (a causa di un accumulo di mutazioni o di mancanza di sequenze regolative)

Talvolta gli pseudogeni possono avere funzione regolativa -> piccoli RNA non codificanti (snRNA) che regolano l'espressione genica

Dimensione geni eucariotici: da 100 nt (tRNA) fino ad 1 milione di nt

## Sequenze mediamente ripetute

Sequenze presenti in centinaia o migliaia di copie

Rappresentano il **25%** del genoma

Sequenze NON funzionanti: sequenze ripetute interposte, **trasposoni** (sequenze di DNA capaci di saltare da una posizione del genoma all'altro) e **retrotrasposoni** attivi o non più attivi => si tratta di resti di infezioni

Geni funzionanti delle sequenze mediamente ripetute:

- rRNA
- tRNA

lunedì 5 ottobre 2020

## - Istoni

Sono elementi molto importanti: devono esserci tante copie in modo che vengano trascritti e tradotti nel tempo necessario

Delle sequenze mediamente ripetute fanno parte i macro, mini e microsattelliti (mediamente ripetute in tandem)

## Sequenze altamente ripetute: **DNA satellite**

Sequenze molto corte presente in moltissimi copie

Sono spesso definite **DNA satellite**

Possono essere ripetute in:

- **Tandem** (una attaccata all'altra: blocchi di sequenze ripetute tutte vicine) -> si trovano in telomeri e centromeri (la frazione più veloce nella curva di rinaturazione)

Hanno la funzione di garantire il mantenimento del nostro genoma

1. Centromeri: la componente che rinatura più velocemente (varia abbastanza da specie a specie) -> sequenze che contano circa 171 bp
  2. Telomeri: brevi ripetizioni in tandem -> sequenze ben conservate nell'evoluzione diminuisce con l'aumentare dei cicli di replicazione
- **Intersperse** (ripetute in posti differenti): sequenze LINE e SINE (discendenti dei trasposoni)

Delle sequenze altamente ripetute fanno parte le **sequenze consensus**: sequenza più probabile di una particolare sequenza

Visualizzazione del DNA satellite utilizzando **gradienti di CsCl** (*cloruro di cesio*) -> da questo deriva proprio il nome DNA satellite

CsCl determina la formazione di un gradiente di densità: permette di separare molecole con densità differenti (la densità del DNA dipende dalla composizione in basi)

Il DNA satellite presenta un composizione in basi differenti rispetto al genoma ripetuto meno -> rende possibile la sua visualizzazione su CsCl => deposizione dei vari strati di densità differenti

lunedì 5 ottobre 2020

Il DNA satellite si disporrà in una frazione differente della provetta contenente CsCl rispetto al resto del DNA (quando si effettua un gradiente di densità si centrifuga il DNA in una provetta contenente CsCl)

**Gradienti di cloruro di cesio:** gradiente di densità -> separazione sulla base della densità

**Gradienti di saccarosio:** separazione sulla base della dimensione (e alla conformazione) -> separazione in base alla capacità di muoversi nel gradiente (pian piano diventa più difficile)

Genoma umano rispetto agli altri genomi -> tantissimi introni

[Coli NON presenta DNA satellite!]

La densità genomica cresce con la complessità dell'organismo -> sono anche presenti anche delle sequenze *intergeniche* (sequenze che spaziano tra loro i geni)