

# Genoma e geni

## Espressione ed utilizzo della informazione genetica I

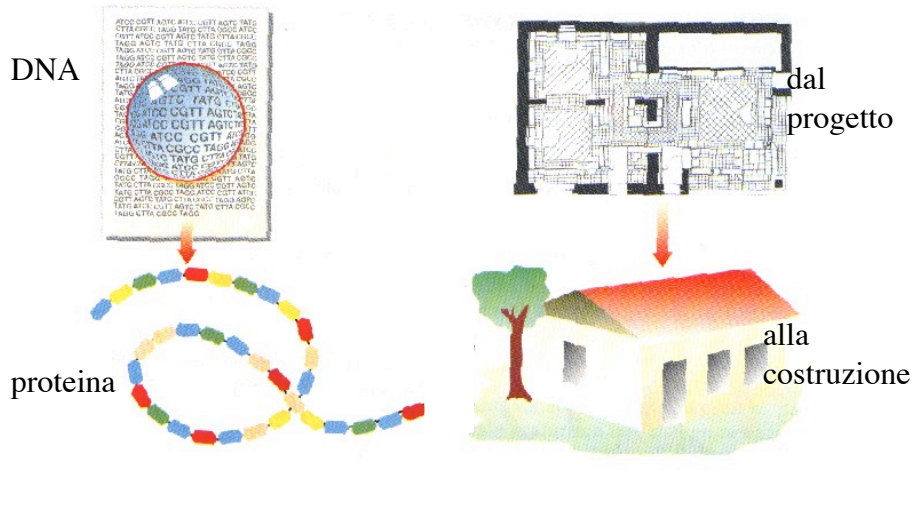
CdL Tecnici di Lab Biomedico  
AA. 2011-12 - Prof.ssa Frabetti

**L'informazione genetica**, contenuta nel DNA, ha lo scopo di:

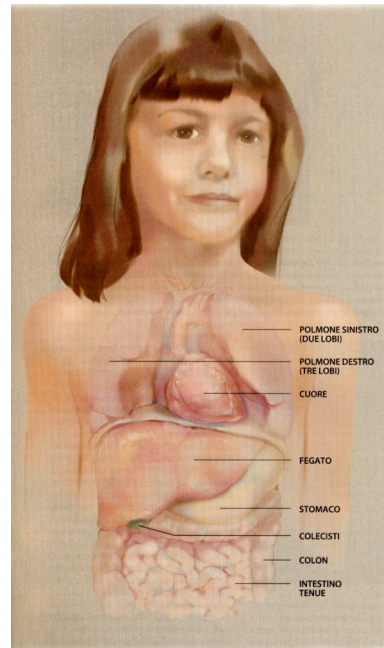
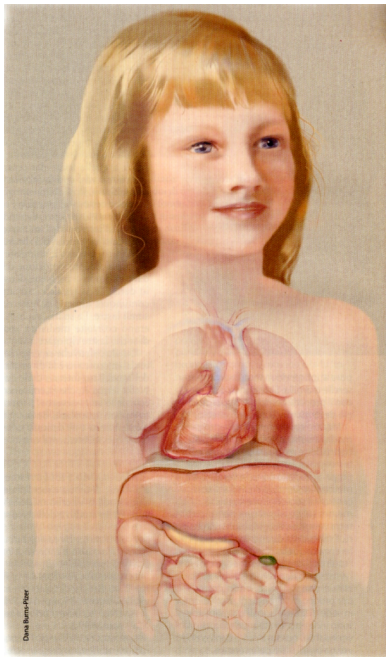
- 1) mantenere lo stato vitale (strutture e funzioni)
- 2) realizzare l'adattamento (risposta ad un cambiamento)
- 3) determinare il differenziamento (specializzazione di strutture e funzioni)

**L'informazione genetica** ereditata determina il manifestarsi di specifici caratteri inducendo la sintesi di certe proteine: **le proteine sono l'anello di congiunzione tra informazione e la realizzazione/effettuazione di strutture ed attività.**

**ESPRESSIONE DELLA INFORMAZIONE BIOLOGICA**  
**non solo le strutture, ma anche l'architettura e la distribuzione**  
**spaziale di queste, nonché i cambiamenti temporali**



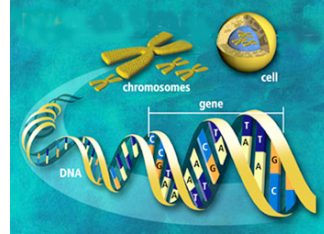
*Situs inversus (< 1:10000)*



Cosa studiamo?

Dove è scritta l'informazione genica:

**DNA**  
**Cromatina / cromosomi**  
**Nucleo eucariotico**



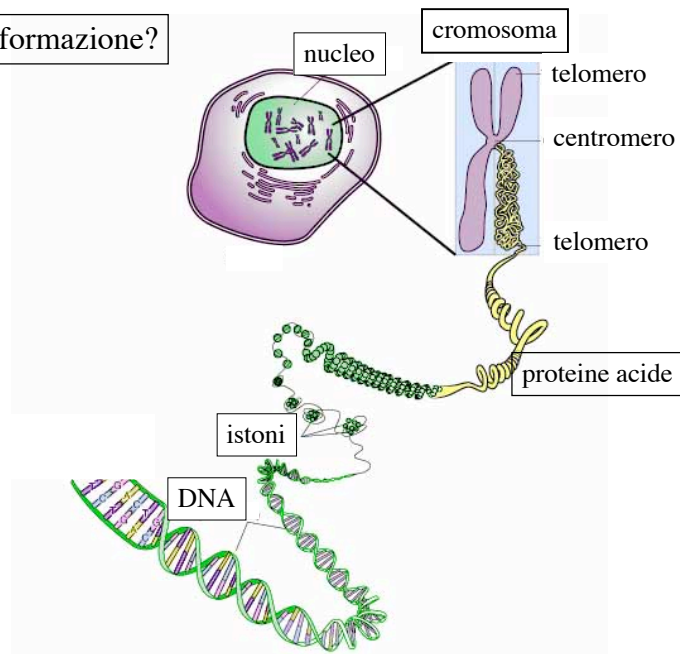
Come è organizzata l'informazione genica:

**Genoma**  
**Geni - come concetto**  
**- "struttura del gene eucariota tipo"**

Come si esprime l'informazione genica:

**Trascrizione e Traduzione**  
**Codice genetico**  
**Meccanismi di CONTROLLO della espressione genica**

Dove è scritta la informazione?

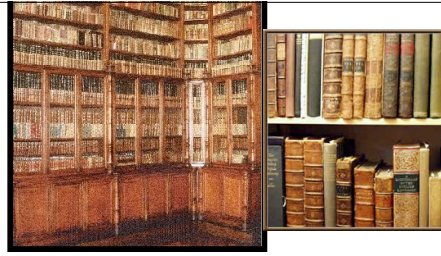


Come è organizzata l'informazione?

Una enorme quantità di informazione  
dentro ad ogni singola cellula:  
il GENOMA

Aprendo ogni libro si trovano brani così scritti:

.....ATCCGAGCTTTACGTACGGTTACCGGATCGAGCATACT.....



Vedremo:

- ✓ **Definizione** di genoma
- ✓ **Complessità** del genoma
- ✓ **Organizzazione** del genoma umano
- ✓ **Variabilità** del genoma umano

## Genoma - definizioni

Il genoma è:

- il complesso dell'informazione genetica di una cellula
- la massa totale del DNA cellulare
- il patrimonio ereditario dell'organismo a cui appartiene

## Complessità del genoma

Servono informazioni proporzionali alla complessità dell'organismo ?

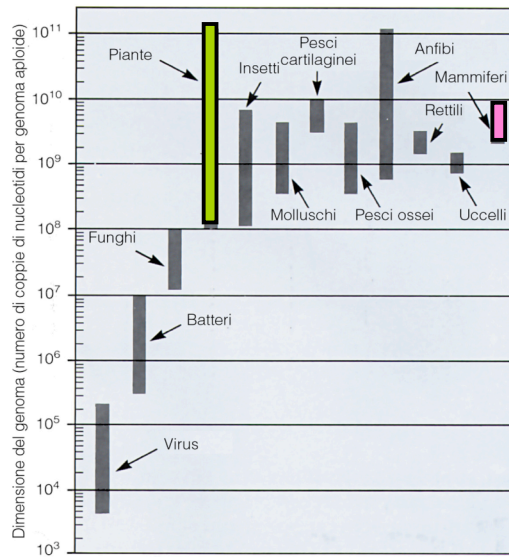
La complessità dipende dalle dimensioni del genoma o dal numero di geni o unità informative?

## DIMENSIONI

**PARADOSSO DELLA  
NON PROPORZIONALITA'  
TRA DIMENSIONI GENOMA E  
COMPLESSITA' ORGANISMO**

**Paradosso  
del valore "C"**  
*C-value paradox*

**C = contenuto in  
DNA di un  
genoma aploide  
misurato in bp**

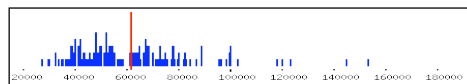


**Figura 16-11** Correlazione fra dimensione del genoma e tipo di organismo. Per ogni gruppo di organismi la barra rappresenta la variabilità approssimativa della dimensione del genoma, misurata come quantità di coppie di nucleotidi per genoma aploide.

## NUMERO DEI GENI codificanti per proteine



**Media: 61.710**  
**Minimo: 27.462**  
**Massimo: 153.478**



*Drosophila melanogaster*  
**16.321**



*Caenorhabditis elegans*  
**21.210**



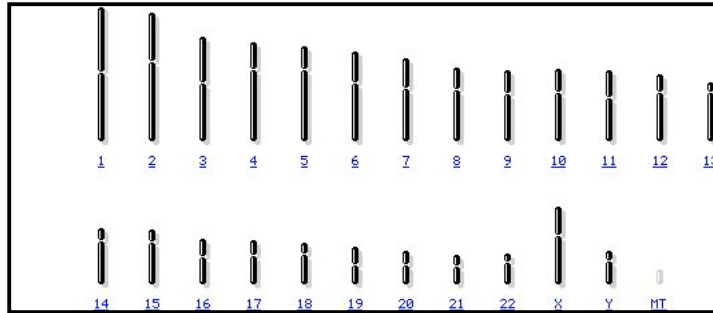
*Arabidopsis thaliana*  
**26.000**



*Homo sapiens*  
**STIMA: 23.000**  
**(analisi computer)**

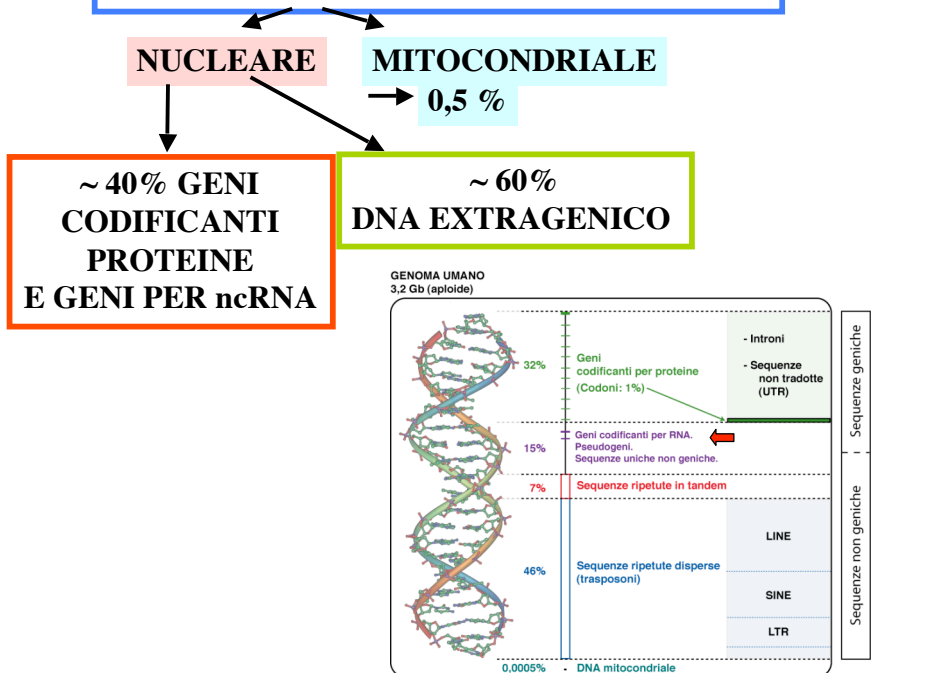
## Organizzazione generale - genoma umano

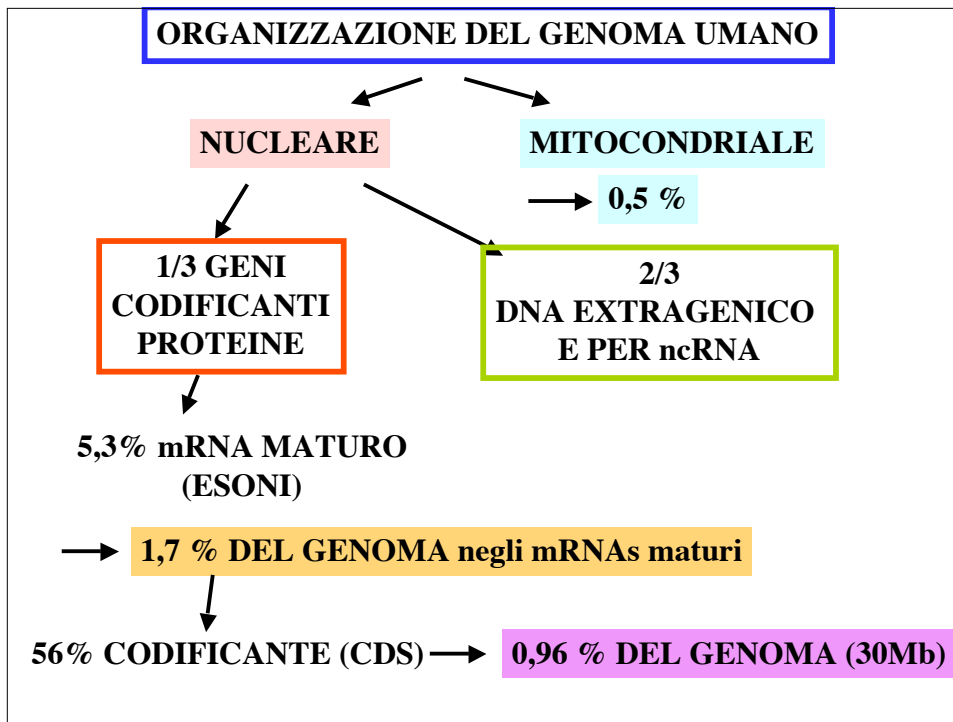
Il **genoma umano** è distribuito in molecole di DNA che costituiscono i cromosomi: 22 tipi di AUTOSOMI, 2 tipi di ETEROCROMOSOMI (X e Y)



Corredo aploide 3.200.000.000 bp ovvero 3,2 Gbp  
 (Cr. 1 246.200.000 bp, Cr. 21 45.000.000 bp)  
 Genoma mitocondriale 16.569 bp  
**Peso** (genoma diploide) 7 pg  
**Lunghezza** (genoma diploide) 2 m


## ORGANIZZAZIONE DEL GENOMA UMANO





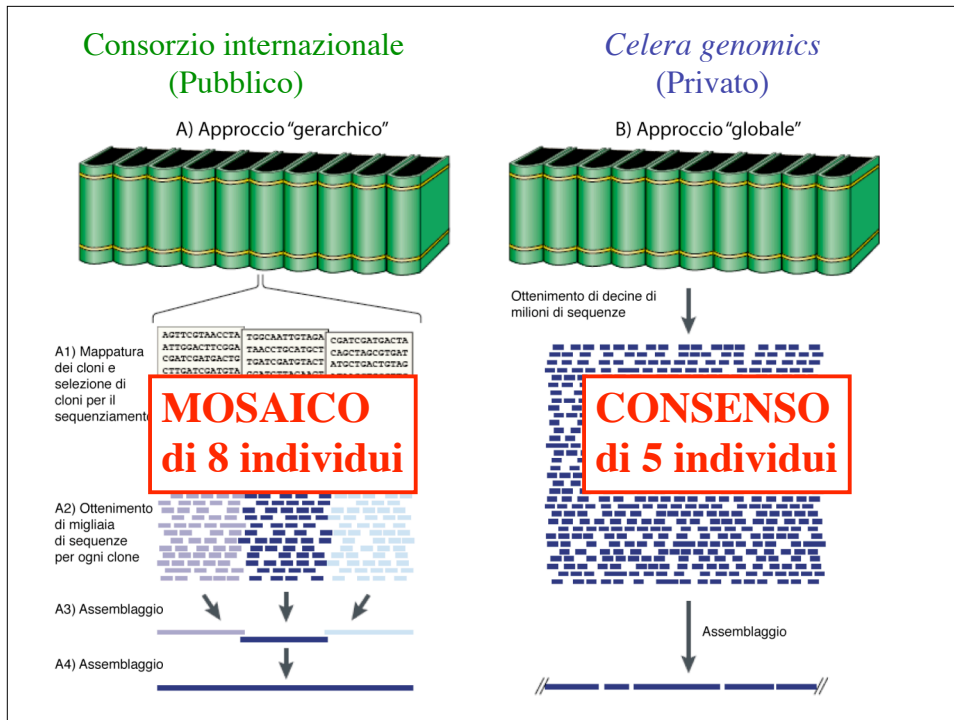
**Variabilità del genoma umano**

**Genoma umano =  
genoma di  
*un individuo*  
umano**



**Ciascuno ha la propria sequenza!  
Si può sequenziare il genoma di uno o più individui**





## VARIABILITA' ALL'INTERNO DELLA SPECIE



**99,9% IDENTITA'**

**0,1%**

**Differenze dell'1 per 1000**

**6x10<sup>6</sup> bp DIVERSE!!**

## VARIABILITA' TRA LE SPECIE



**98,8% IDENTITA'**

**1,2%**

**Differenze dell'1 per 100**



La complessità del genoma si riflette sulla **espressione** dell'informazione genica e sulla **definizione stessa di gene**

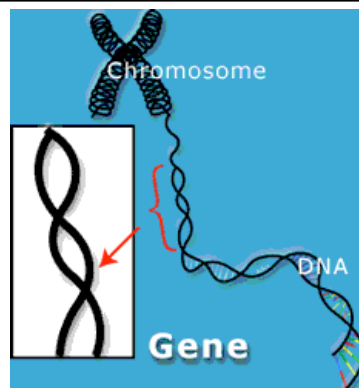


Come una sorta di *matryoska* il **genoma** si organizza nei **cromosomi**, che contengono i **geni** che altro non sono che tratti di DNA

I geni sono le **unità responsabili delle caratteristiche ereditarie**



Il gene è **una regione di DNA trascritta che contiene istruzioni per la sintesi di una proteina, di un RNA o ignota**



Il gene è una delle tante istruzioni contenute in ogni cellula

## FUNZIONI DEI GENI

<b>METABOLISMO</b>	<b>22%</b>
<b>INFORMAZIONE GENETICA</b>	<b>25%</b>
<b>STRUTTURA CELLULARE</b>	<b>21%</b>
<b>SEGNALI</b>	<b>12%</b>
<b>FUNZIONI TESSUTO-SPECIFICHE</b>	<b>20%</b>

**GENI PER RNA NON TRADOTTI**  
(rRNA **85%**, tRNA **10%**, RNA non-codificanti)

### Alcuni geni sono:

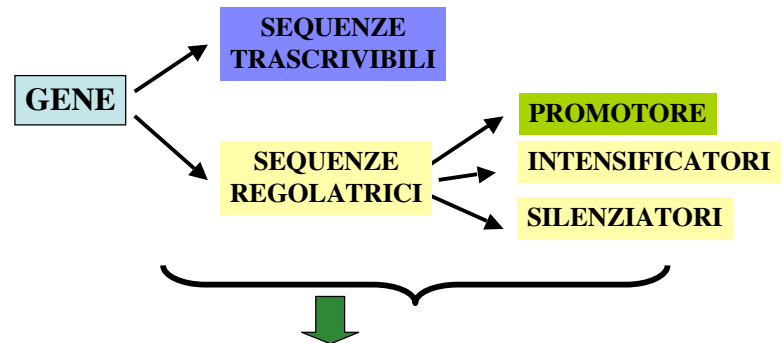
#### 1) Geni essenziali o geni *housekeeping*

Geni che devono essere espressi praticamente in tutti i tipi cellulari poiché codificano prodotti utili al funzionamento generale delle cellule, per es. alla sintesi proteica o alla produzione di energia

#### 2) Geni con limitazioni spatio-temporali nella espressione:

SPAZIO	TEMPO
Espressione diversa in diversi organi e tessuti	Stadio del ciclo cellulare
Specificità per tessuto, tipo cellulare	Stadio dello sviluppo
Singole cellule (specificità di cellula es. Ab)	Stadio del differenziamento
Distribuzione intracellulare	Espressione <i>inducibile</i>

## Gene e sua funzione: Cosa fa accendere il gene?



### UNITA' DI TRASCRIZIONE

Tratto di DNA che codifica per una molecola di RNA e delle sequenze necessarie per la sua trascrizione

### PROMOTORE

Sequenza nel DNA tra i 40 e 200 nucleotidi, con affinità più o meno elevata per la **RNA polimerasi** e di solito posta *a monte* (cioè prima) del gene.

Ogni promotore fornisce all'enzima:

- sequenze di riconoscimento (es. **TATA box**)
- sequenze di legame stabile
- sito di inizio della trascrizione (il nucleotide "+1")
- sequenze di regolazione** (che possono intensificare e/o silenziare l'attività della polimerasi)

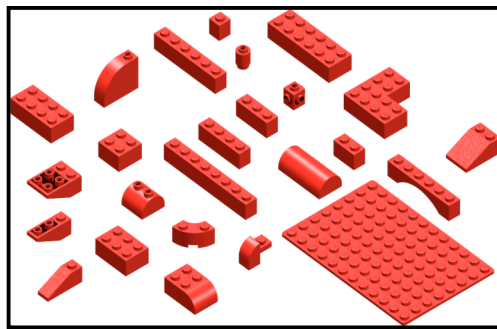


Si tratta come di una **bandierina** in mezzo al mare di DNA: che segna quale informazione leggere e trascrivere

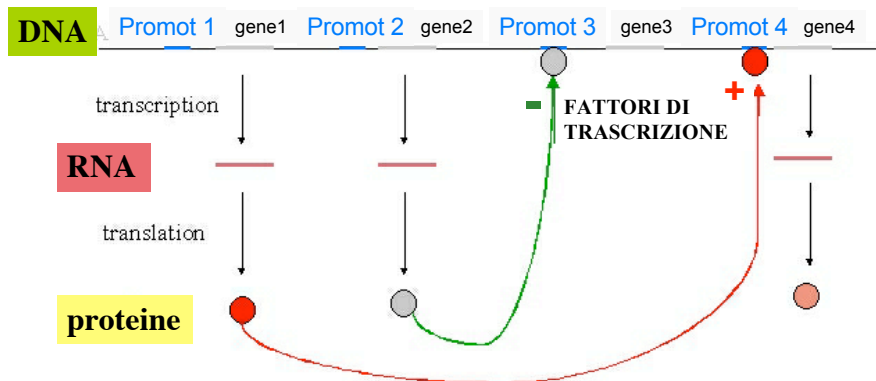
## II PROMOTORE

I PROMOTORI HANNO NATURA MODULARE!

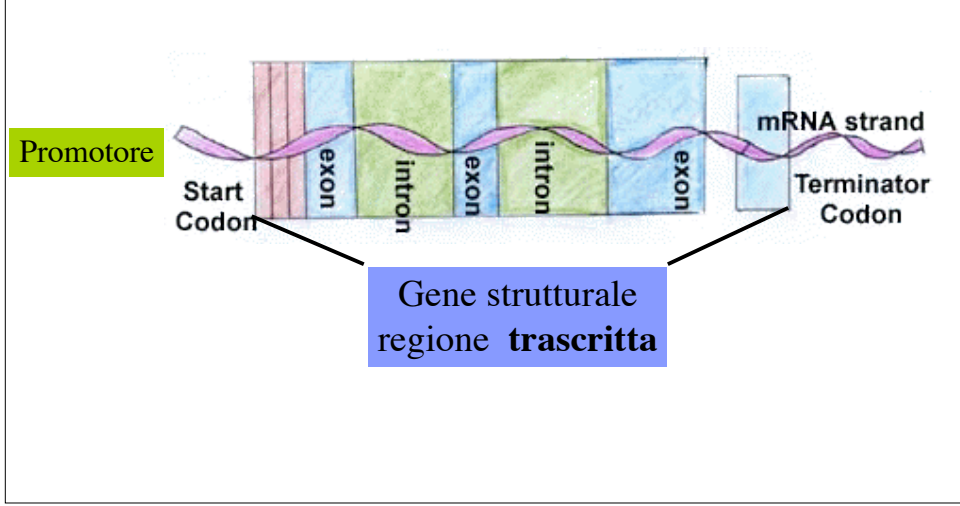
Sequenza data dalla combinazioni di corti elementi utili al legame con proteine che sono **FATTORI DI TRASCRIZIONE** : *piccolo interattoma!*



**Gene networks** - un gene può esprimere le proprie informazioni solo nel contesto dell'intero genoma della cellula



## Gene eucariota - la struttura nasconde complessità



TTAGCACTACCGTATTTGCGCATTACCAGATTAGAGAA  
 ATGCTAGTCGATCTATCGATCGGCTATTCGCAAAGCTG  
 CGCGACTGCGATGCGCTAGCATGCGATTTCGCGATCGCC  
 GAGCGCTCGGAGCGCGCTAGCGGAATACTATATAGC  
 GCGGATCAGTCTAGATCTATGAGATCGATAGCGATCTA  
 GAGATAGGATCGAGATCGAGGCGAGATCATATGAGCG  
 CGGCTATTTAGGCTTAGAGGATTCGGAGATTCGGAGCT  
 TAGGATTACAGAGAGCTTCTTAGGCGCTCCCGGTATCG  
 CTCATCCCATATTTAAATCTATCGATCGAGCTCTCCA  
 ATGCGATCGATAGGACTAGTAGCT  
 TGATAGGCTCGATGAGCATGAGAT  
 CATAGGCATGACTGATCGACTGCA  
 TGCATGATGACTGCATATGAGGACTCGCATTAGCAC  
 TACCGTATTTGCGCATTACGATTAGAGAAATGCTAG  
 TCGATCTATCGATCGGCTATTCGCAAAGCTGCGCGACT  
 GCGATGCGCTAGCATGCGATTTCGCGATCGCCGAGCGC  
 TCGGAGCGCGCTAGCGGAATACTATATAGCGGGAT  
 CAGTCTAGATCTATGAGATCGATAGCGATCTAGAGATA  
 GGATCGAGATCGAGGCGAGATCATATGAGCGCGGCTA  
 TTTAGGCTTAGAGGATTCGGAGATTCGGAGCTTAGGAT  
 TACAGAGAGCTTCTTAGGCGCTCCCGGTATCGCTCCA  
 TCCCATATTTAAATCTATCGATCGAGCTCTCCAATGCG  
 ATCGATAGGACTAGTAGCTAGCTGAGCATGATA  
 GGCTCGATGAGCATGAGATGCATGTACGACTGCATAG  
 GCATGACTGATCGACTGCATCATGACGCATGACTGCAT  
 GCATGACTGCATATGACGGACTCGCATTAGCACTACCG  
 TATTTGCGCATTACCAGATTAGAGAAATGCTAGTCGAT  
 CTATCGATCGGCTATTCGCAAAGCTGCGCGACTGCGAT  
 GCGCTAGCATGCGATTTCGCGATCGCCGAGCGCTCGG  
 AGCGCGCTAGCGGAATACTATATAGCGGGATCAGTC  
 TAGATCTATGAGATCGATAGCGATCTAGAGATAGGATC  
 GAGATCGAGGCGAGATCATATGAGCGCGGCTATTTAG  
 GCTTAGAGGATTCGGAGATTCGGAGCTTAGGATTACA  
 GAGAGCTTCTTAGGCGCTCCCGGTATCGCTCCATCCC  
 ATATTTAAATCTATCGATCGAGCTCTCCAATGCGATCG  
 ATAGGACTAGTAGCTAGCTAGCTGAGCATGATAGGCT  
 CGATGAGCATGAGATGCATGTACGACTGCATAGGCAT  
 GACTGATCGACTGCATCATGACGCATGACTGCATGCA  
 GACTGCATATGACGGACTCGCA

Nel "libro"  
**DNA,**  
 le pagine  
 sono scritte in  
 codice/i

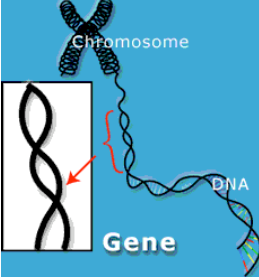
TTAGCACTACCGTATTTGCGCATTACCAGATTAGAGAA  
 ATGCTAGTCGATCTATCGATCGGCTATTCGCAAAGCTG  
 CGCGACTGCGATCGCTAGCATGCGATTGCGGATCGCC  
 GAGCGCTCGGAGCGCGCTAGCGGAATACTATATAGC  
 GCGGATCAGTCTAGATCTATGAGATCGATAGCGATCTA  
 GAGATAGGATCGAGATCGAGGCGAGATCATATGAGCG  
 CGGCTATTTAGGCTTAGAGGATTCGGAGATTCGGAGCT  
 TAGGATTACAGAGAGCTTCTTAGCGCTCCCGGTATCG  
 CTCCATCCATATTTAAATCTATCGATCGAGCTTCCA  
 ATGCGATCGATAGGACTAGTAGCT  
 TGATAGGCTCGATGAGCATGAGAT **GENE**  
 CATAGGCATGACTGATCGACTGCA  
 TGCATGCATGACTGCATATGAGGACTCGCATTAGCAC  
 TACCGTATTTGCGCATTACCAGATTAGAGAAATGCTAG  
 TCGATCTATCGATCGGCTATTCGCAAAGCTCGCGGACT  
 GCGATGCGCTAGCATGCGATTTCGCGATCGCCGAGCGC  
 TCGCGAGCGGCTAGCGGAATACTATATAGCGCGGAT  
 CAGTCTAGATCTATGAGATCGATAGCGATCTAGAGATA  
 GGATCGAGATCGAGGCGAGATCATATGAGCGCGGCTA  
 TTTAGGCTTAGAGGATTCGGAGATTCGGAGCTTAGGAT  
 TACAGAGAGCTTCTTAGGCGCTCCCGGTATCGCTCCCA  
 TCCCATATTTAAATCTATCGATCGAGCTTCCAATGCG  
 ATCGATAGGACTAGTAGCTAGCTAGCTGAGCATGATA  
 GGCTCGATGAGCATGAGATGATGATGACGACTGCATAG  
 GCATGACTGATCGACTGCATCATGACGCATGACTGCAT  
 GCATGACTGCATATGACGGACTCGCATTAGCACTACCG  
 TATTTGCGCATTACCAGATTAGAGAAATGCTAGTCGAT  
 CTATCGATCGGCTATTCGCAAAGCTCGCGGACTCGGAT  
 GCGCTAGCATGCGATTTCGCGATCGCCGAGCGCTCGCG  
 AGCGCGTAGCGGAATACTATATAGCGGGATCAGTC  
 TAGATCTATGAGATCGATAGCGATCTAGAGATAGGATC  
 GAGATCGAGGCGAGATCATATGAGCGCGCTATTTAG  
 GCTTAGAGGATTCGGAGATTCGGAGCTTAGGATTACA  
 GAGAGCTTCTTAGGCGCTCCCGGTATCGCTCCCATCCC  
 ATATTTAAATCTATCGATCGAGCTTCCAATGCGATCG  
 ATAGGACTAGTAGCTAGCTAGCTGAGCATGATAGGCT  
 CGATGAGCATGAGATGATGATGACGACTGCATAGGCA  
 TACTGATCGACTGCATCATGACGCATGACTGCATGCA  
 GACTGCATATGACGGACTCGCA

1 cctgtggag ccacacctta gggttggcca atctactccc aggaagcagg agggcaggag  
 61 ccaggcttg gca<sup>aaa</sup>agt caggcgagag ccatctattg cttaacattg etctgacac  
 121 aactgtgttc actagcaacc tcaaacagac acc<sup>ttg</sup>gtgc acctgactcc tgaggagaag  
 181 tctgcccgtta ctgcccctgt gggcaaggtg aactgtgatg aagttggtg tgaggccctg  
 241 ggc<sup>ag</sup>gttg tatcaaggtt acaagacag ttaaggaga ccaatagaaa ctgggcatg  
 301 ggagacagag aagactcttg gttttctgat aggcactgac tctctctgcc tattgtctta  
 361 ttttcccaacc ct<sup>ag</sup>gtgc tgggtgtcta ccttggacc cagaggttct ttgagtctct  
 421 tggggatctg tccactctct atgctgttat gggcaacct aagtggaag ctcatggcaa  
 481 gaaagtgtc ggtgccc<sup>tta</sup> gtgatggcct ggctcaactg gacaacctca agggcaacct  
 541 tgcccactg agtgagctg actgtgaca gctgcaactg gatcctgaga acctcagggt  
 601 gagtctatg gacccttgat gttttcttc cctctcttt ctatggttaa gttcatgtca  
 661 taggaaggg agaagtaaca ggttacagtt tagaatgga aacagacgaa tgattgcatc  
 721 agtgggaag tctcaggatc gtttttagtt cttttattg ctgtctcaaa caattgtttt  
 781 cttttg<sup>tta</sup> attctgtct tcttttttt tctctctcc aatttttact attatactta  
 841 atgccc<sup>taac</sup> attgtgtata caaaaaggaa atatctctga gatacattaa gtaacttaaa  
 901 aaaaaactt acacagctg cctagta<sup>cat</sup> tactatttgg aatatatgtg tgcttattg  
 961 catattcata atctccc<sup>taac</sup> tttattttct tttattttta attgatacat aatcattata  
 1021 catatttat ggttaaagt taatgtttta atatgtgac acatattgac caaatcaggg  
 1081 taattttgca ttgttaatt taaaaaatgc tttctcttt taatataact tttgtttat  
 1141 ctatttteta atactttccc taatctcttt ctctcaggcc aataatgata caatgtatca  
 1201 tgcctctttg caccattcta aagaataaca gtgataatt ctgggttaag gcaatagcaa  
 1261 tatttctgca tataaatatt tctgcatata aattgtaact gatgtaagag gtttcaatt  
 1321 gctaatagca gctacaatcc agtcaactt ctgcttttat tttatggtg ggataaagct  
 1381 ggattattct gagtccaagc taggccc<sup>ttt</sup> tgtcaatcat gttcacaact cttatcttcc  
 1441 tcccac<sup>agt</sup> cctgggcaac gtgctggtct gtgtgctg<sup>cc</sup> ccaacttt ggcaagaat  
 1501 tca<sup>cccc</sup>caacc agtgcaggct gcctatcaaga aagtggtg<sup>cc</sup> tgggtggtc aatgccc<sup>ttg</sup>  
 1561 cccacaagta tca<sup>ct</sup>agct gcctttctgt ctgtccaatt tctat<sup>taaa</sup>g gctctttgt  
 1621 tccctaagt<sup>c</sup> caactactaa actgggggat attatgaag gctctgaca tctggattc  
 1681 gct<sup>taataa</sup>aaa aaacatttat tttcattgca atgatgtatt taaattatt ctgaatttt  
 1741 tactaaaaa ggaatgtgg aggtcagtgc attttaaaa taaga<sup>aaat</sup> atgagctgtt  
 1801 caaaccttg gaaatacac tatatcttaa actccatgaa agaaggtgag gctgcaacca  
 1861 gctaagtca atgtgcaaca gccctgatg cctatgctt atccatccc cagaaaagg  
 1921 tcttgtgaga gcttgattt gcagg<sup>ttaa</sup> gttt<sup>gt</sup>ctat gctgtattt acattactta  
 1981 ttgttttaga tgcctcatg aatgtcttt cactaccat ttgcttatcc tgcactctc  
 2041 tca<sup>g</sup>ccctgca ct

**Gene della beta-globina (Homo sapiens)**  
 Sequenza GenBank n. [NM\\_000518](#)  
 DNA a filamento doppio (l'altro si ricava per complementarietà)

1 cccgtggag coacacccta gggttggcca atctactccc aggagcagg agggcaggag  
61 coaggctgg gca~~aaa~~agt cagggcagag coactctatg cttacatttg cttctgacac  
121 aactgtgttc actagcaacc tcaaacagac acc~~ATG~~gtgc acctgactcc tgaggagaag  
181 tctgccgtta ctgccctgtg gggcaagtg aactggatg aagttggtg tgaggccctg  
241 ggca~~g~~gttg tatcaaggtt acaaga~~gg~~ ttaaggaga coaatagaaa ctggccatgt  
301 ggagacagag aagactctg g~~act~~ctgat aggcactgac tctctctgcc tattgtctca  
361 ttttcccacc ctt~~ag~~gtgc tgggtgtca cccctggacc cagaggttct tggagtcctt  
421 tgggtgtctg tcaactctg atgctgttat gggcaacct aaggtgaag ctcatggcaa  
481 gaaagtgtc ggtgccctta gtgatggcct ggctcactg gacaacctca agggcactt  
541 tgcccaactg agtgagctgc actgtgaaa gctgcactg gatcctgaga acttcaagg~~t~~  
601 gagtctatgg gacccttgat gtttctttc ccttctttt ctatg~~aaa~~aa gttcatgtca  
661 taggaaggg agaagtaaca ggttacagtt tagaatgga aag~~g~~acgaa tgattgcatc  
721 agtggtgaag tctcaggatc gtttagtctt cttttatctg cgttcataa caattgtttt  
781 cttttgtta attctgtctt tttttttt tttcttgc aattttact attatactta  
841 atgcttaac attgtgata acaaaaggaa at~~g~~cttga gatacattaa gtaacttaaa  
901 aaaaaacttt acacagctg cctagtacat actatttgg aatatatgtg tgcttattg  
961 catattcata atctccctac tttatttct tttatttta attgatacat aatcattata  
1021 catatttatg ggttaaaatg ta~~g~~gttta atatgtgac acatattgac caaatcaggg  
1081 taattttgca ttgtaat~~t~~ aaaaaatgc tttctcttt taatatact tttgtttat  
1141 cttatttcta ata~~ctt~~ta taatctctt ctttcaggc aataatgata caatgtatca  
1201 tgctcttgg cac~~ct~~tcta aagaataaca gtgataatt ctgggttaag gcaatagcaa  
1261 tatttctgca ct~~ta~~aatatt tctgcatata aattgtaact gatgtaagag gttcattat  
1321 gctaataca gctacaatcc agctaccat ctgctttat tttatggtg ggataaggt  
1381 gga~~ct~~tctc gagtccaagc taggcccctt tgctaatcat gttcaatcc cttatctcc  
1441 tcccag~~ct~~ cctgggcaac gtgctggtct gtgtgtgccc coataccttt ggcaagaat  
1501 tca~~cc~~caacc agtgcaggct gctatcaga aagtgtgccc tgggtgctc aatgccctgg  
1561 cccacaagta tca~~ct~~agct cgtttctgt ctgccaatt tctattaag gtcctttgt  
1621 tccctaagtc caactactaa actggggat attatgaag gccctgagca tctggattct  
1681 gct~~ta~~ataa aaacatttat tttcattgca atgatgtatt taaatt~~ct~~ ctgaaatatt  
1741 tactaaaaag ggaatgtggg aggtcagtc atttaaaaca ta~~g~~aaaatg atgagctgt  
1801 caaaccttgg gaaaatacac tatatcttaa actccatgag agaaaggtgag gctgcaacca  
1861 gctaagtcac attggcaaca gccctctgat cctat~~ct~~tt attcatcctc cagaaaagga  
1921 ttctgtaga ggcttgatt gcaggttaaa g~~ct~~ctgat gctgtattt acattaacta  
1981 ttgttttagc tgtcctatg aatgct~~tt~~ caactaccat ttgcttatcc tgcattcttc  
2041 tcagccttga ct

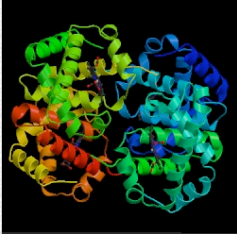
**Cromosoma 12**



**Gene della beta-globina (Homo sapiens)**  
da 1 a 2052 nucleotidi  
Sequenza GenBank n. [NM\\_000518](#)  
DNA a filamento doppio (l'altro si ricava per complementarità)

acattg cttctgacac  
aactgtgttc actagcaacc tcaaacagac acc~~ATG~~gtgc acctgactcc tgaggagaag  
tctgccgtta ctgccctgtg gggcaagtg aactggatg aagttggtg tgaggccctg  
ggcaggtctg tgggtgtca cccctggacc cagaggttct tggagtcctt tggggtctg  
tcaactctg atgctgttat gggcaacct aaggtgaag ctcatggcaa gaaagtgtc  
ggtgccctta gtgatggcct ggctcactg gacaacctca agggcactt tgcccaactg  
agtgagctg actgtgaaa gctgcactg gatcctgaga cctctgggca agtgctggt  
ctgtgtctg gccatacct ttggaaaaga attcaacca coagtgcaag ctgcctatca  
gaaagtgtg gctggtgtg ctaatgccct ggcccacaag tctaac~~TAA~~g ctgcctttct  
tgctgtcaa tttctattaa aggttccttt gttccctaag tcaactact aaactggggg  
atattatgaa ggccttgag catctggatt ctgct~~ta~~ata aaaaacttt at

**mRNA maturo per la beta-globina**



**proteina beta-globina, catena beta della emoglobina e mioglobina**

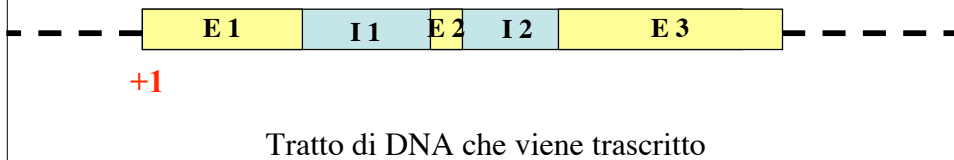


## Struttura del gene eucariota

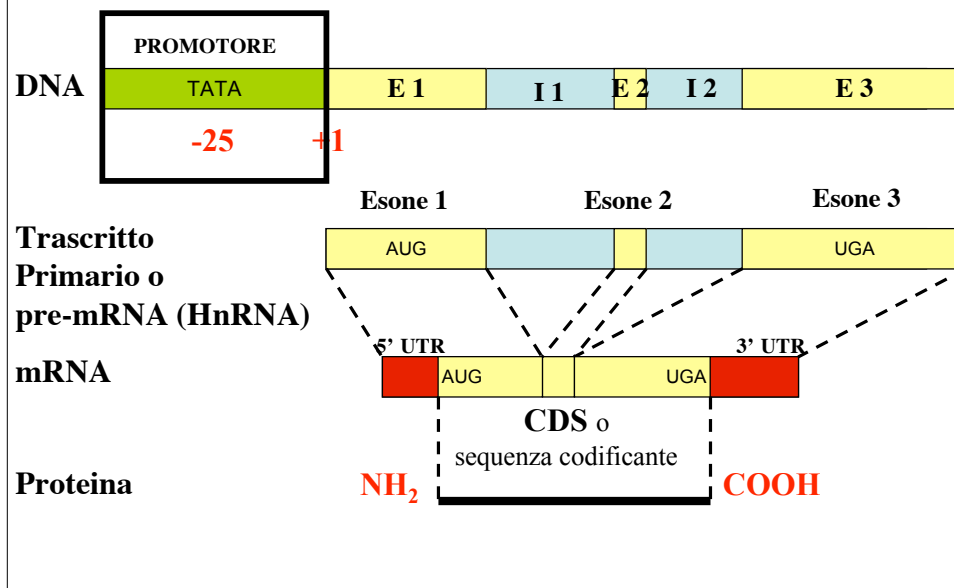
**Il gene ha una natura discontinua:**

**esoni** indicati con E

**introni** indicati con I



**ESONI** (in giallo), intervallati da lunghi INTRONI (in grigio)



Dimensione media geni  
 57.000 bp (57 kb)  
 istoni 100-400 bp (0,1-0,4 kb)  
 distrofina 2.220.000 bp (2,22 Mb)

**Nella realtà gli INTRONI sono molto più lunghi degli esoni**

ESONI  
 Numero medio 11  
 (minimo)-(massimo) 1 - 363  
**Dimensione media 280 bp**  
 INTRONI  
**Dimensione media ca. 6.000 bp**

**mRNA maturo** (in media)  
 Dimensione 3.000 basi  
 Sequenza codificante 1.600 basi (56%)  
**Proteina** 540 AA

