

Struttura della cromatina

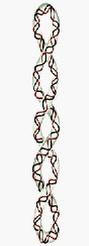
•Direzione dell'avvolgimento del DNA:



(nelle forme A e B del duplex di DNA)

DNA destrorso

•Direzione del superavvolgimento del DNA



Right-handed (negative) superhelix

superavvolgimenti negativi (sinistrorsi)



Normal circular helix

DNA non superavvolto



Left-handed (positive) superhelix

superavvolgimenti positivi (destrorsi)

figura 2

Compattamento del DNA nei cromosomi

Il DNA cellulare e virale è associato a proteine in complessi detti cromosomi.

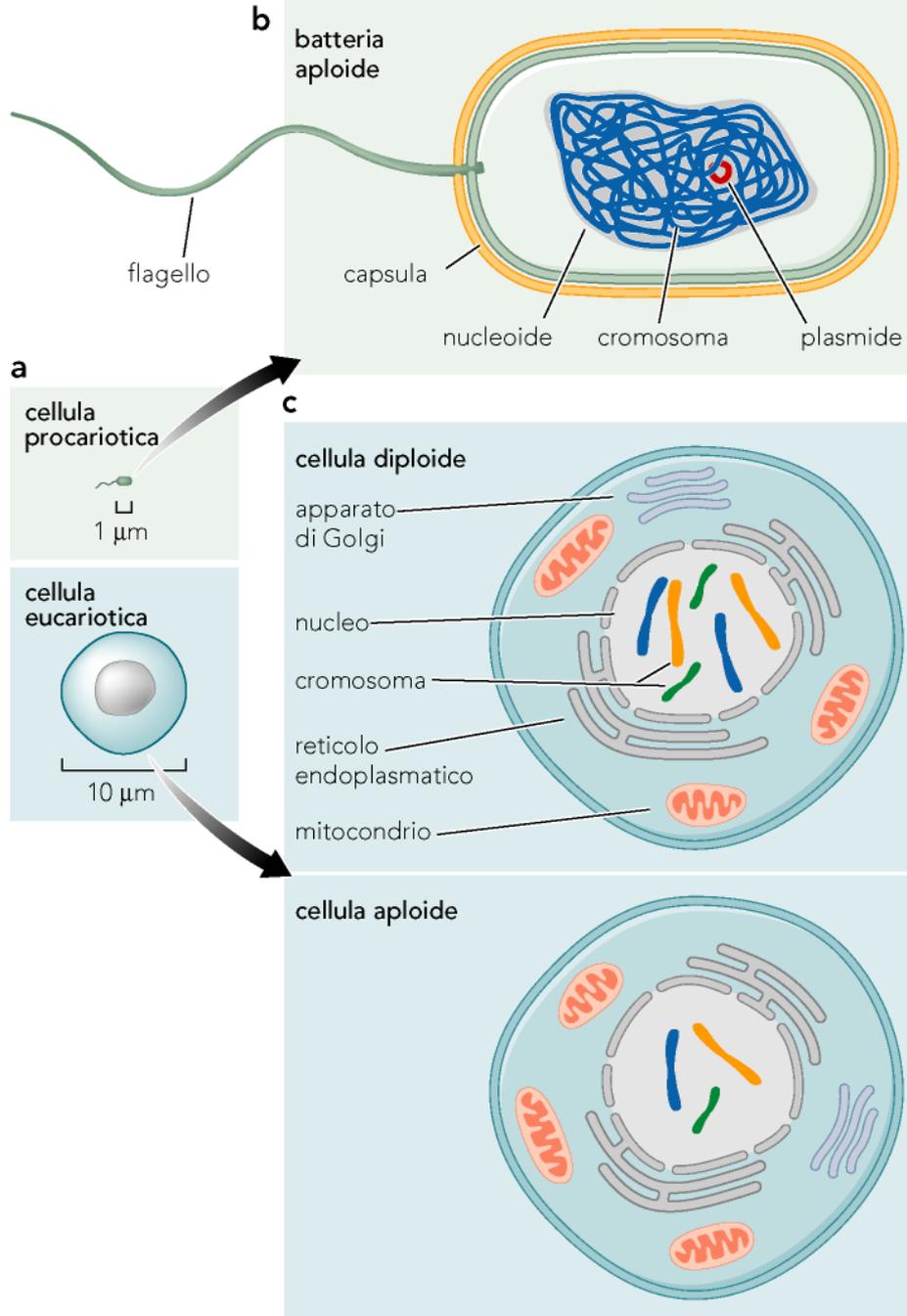
Il compattamento è essenziale per vari motivi:

- a) contenimento nel ristretto spazio cellulare (o virale)
- b) protezione da possibili danni
- c) maggiore stabilità per una corretta espressione dell'informazione
- d) efficiente trasmissione alle cellule figlie
- e) organizzazione strutturale di ordine superiore, che facilita l'espressione genica e la ricombinazione omologa, che genera la diversità individuale.

Cromosoma eucariotico: metà DNA e metà proteine, la maggior parte sono istoni (**cromatina**). Le proteine non istoniche regolano replicazione, trascrizione, riparo, ricombinazione. Le proteine della cromatina possono compattare DNA ≈ 10.000 volte (da cm a μm). Compattamento avviene in stadi successivi: 1) associazione DNA-istoni => nucleosomi, 2) associazione fra nucleosomi vicini, e così via.

La competizione spazio-temporale fra compattamento e accesso di proteine a diverse regioni del cromosoma è alla base della regolazione dell'espressione genica, che in ultima analisi dipende da modificazioni enzimatiche \pm reversibili dei nucleosomi.

Il compattamento del DNA nei procarioti è mediata da proteine istone-simili, che non formano nucleosomi, ma il processo è ancora poco chiaro.



Confronto fra cellula eucariotica e procariotica

I procarioti hanno una copia unica del/dei cromosoma/i, organizzata in una struttura detta **nucleotide**.

Spesso contengono una o più molecole di DNA piccole e circolari, dette **plasmidi**, che portano geni non essenziali.

La maggior parte delle cellule eucariotiche è diploide (coppie di cromosomi omologhi, ognuno da un genitore, nel nucleo).

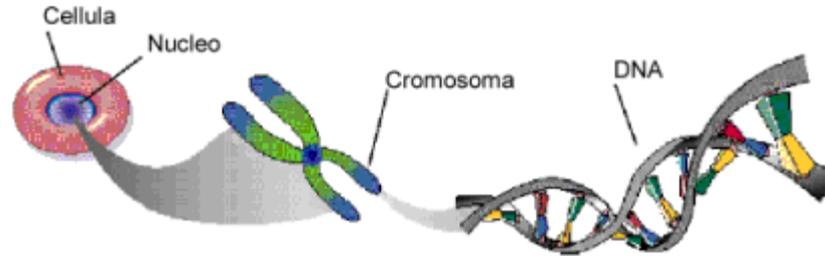
Le cellule coinvolte nella riproduzione sessuale (gameti) sono aploidi (una sola copia di ciascun cromosoma), altre sono poliploidi (più copie, fino a migliaia).

Nell'uomo sono tetraploidi (92 cromosomi) le cellule del fegato

I megacariociti hanno in media 128 copie di ciascun cromosoma, producono migliaia di piastrine, prive di cromosomi, componenti essenziali del sangue (fino a 200.000/ml).

Il DNA, insieme a diverse proteine si organizza in una struttura che è detta **CROMOSOMA**.

I cromosomi sono costituiti da **cromatina**, che consiste di fibre contenenti DNA e proteine. Quando una cellula non è in divisione, la cromatina si trova sotto forma di lunghi filamenti. Durante la divisione le fibre di cromatina si condensano e si rendono visibili come strutture distinte.

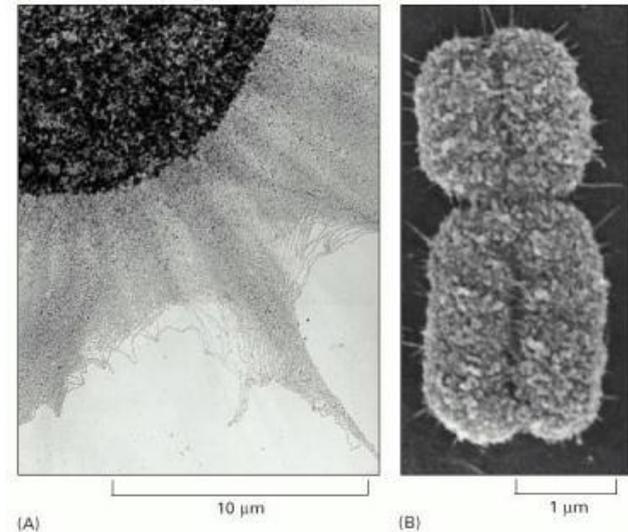


Nelle cellule umane i cromosomi si possono osservare durante la divisione cellulare e distinguere per forma e dimensioni.

La cromatina è un complesso sovramolecolare di acidi nucleici (DNA e RNA) e proteine

Caratteristiche:

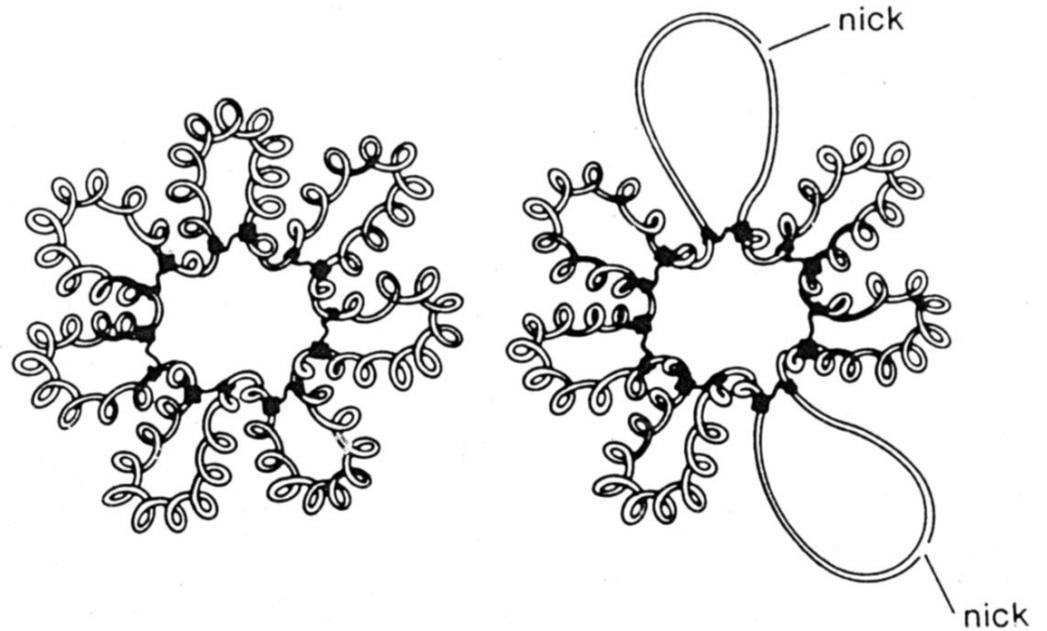
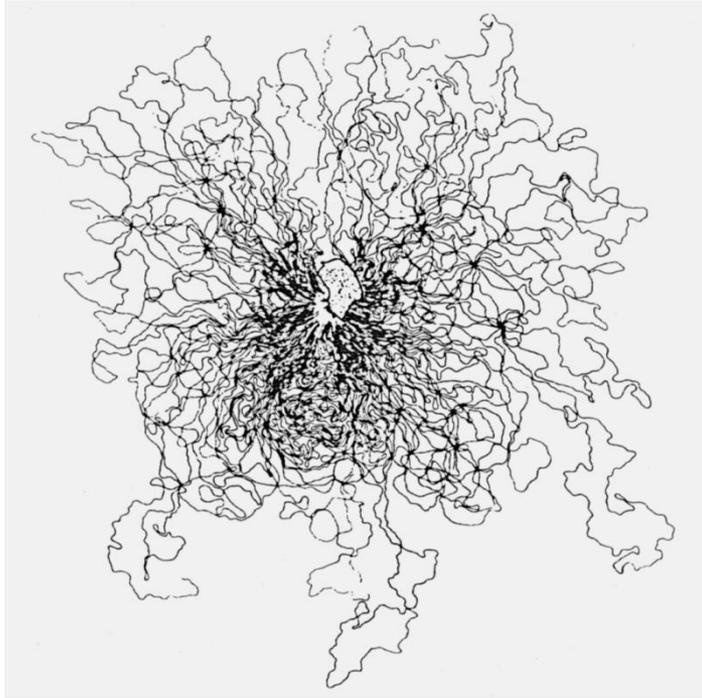
struttura altamente dinamica il cui **grado di organizzazione o compattazione o “condensazione”** varia da cellula a cellula, durante il ciclo vitale della cellula, tra zone diverse dei nuclei interfascici



Impacchettamento del DNA nei cromosomi

- **E. coli ha un cromosoma circolare di circa 4×10^6 b.p. che sarebbe lungo 1,35 mm (1000 volte più dell'intero batterio) se non fosse impacchettato con le proteine.**
- **Una cellula eucariotica (spermatozoo) ha 3×10^9 b.p. e questo DNA sarebbe lungo circa 1m e deve stare in 5-10 μm**
- **Il DNA è compattato grazie a proteine basiche (istoni e istoni-simili e proteine non istoniche)**

Modello dell'organizzazione ad anse nel nucleotide batterico

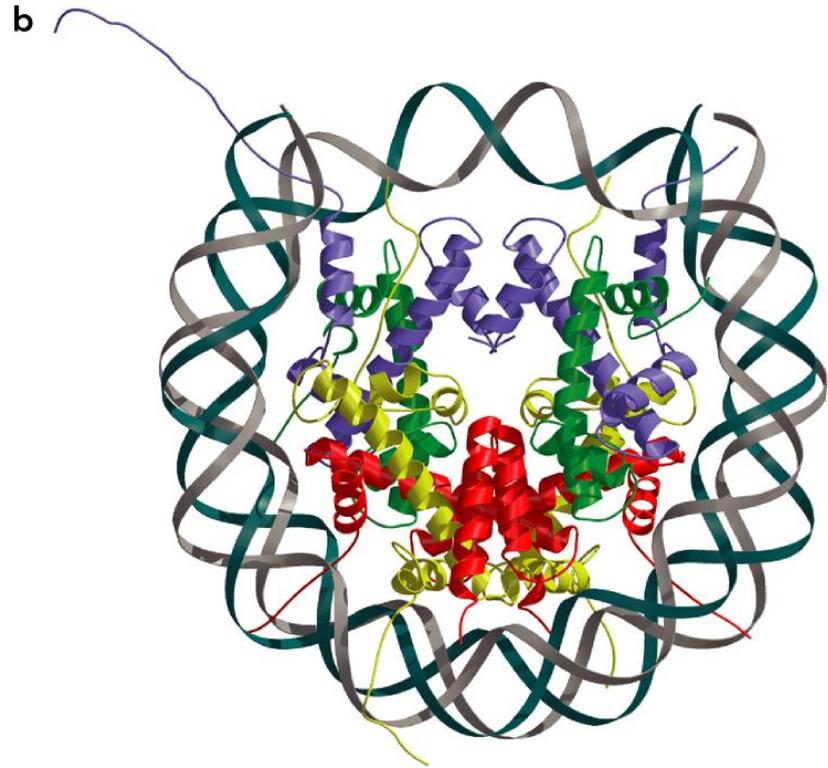
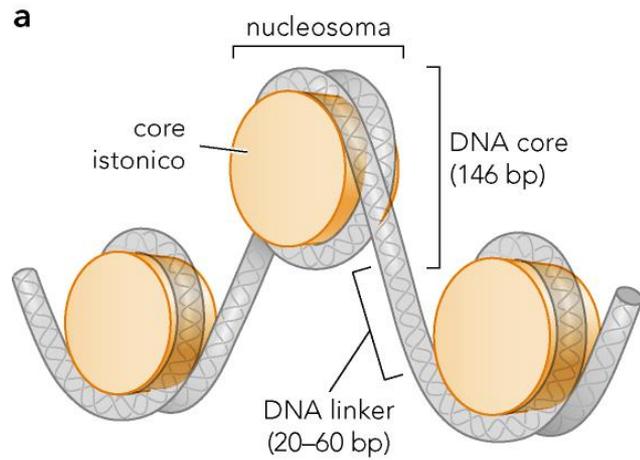


Impacchettamento del DNA

- Gli **istoni** sono Proteine cariche positivamente (ricche in Arg, Lys, His) e si legano ai gruppi fosfato del DNA formando i nucleosomi
- Il ruolo degli istoni è strutturale e funzionale (influenzano anche l'attività del DNA a cui sono associati)
- Il **nucleosoma**: è una “perla” costituita da 147 b.p avvolte su un nucleo di 8 istoni (2 x 4 tipi).
- I nucleosomi sono uniti da **DNA linker**

Impacchettamento del DNA

- Il **nucleosoma impacchettato** si ha quando gli istoni **H1** (un quinto tipo) si associano al DNA linker avvicinando i nucleosomi adiacenti
- Nella **cromatina decondensata** i nucleosomi impacchettati formano anse spiralate tenute insieme dalle **Proteine dell'impalcatura** (non istoniche)
- Nella **cromatina condensata** le anse interagiscono fra di loro per costituire i **cromosomi metafasici**



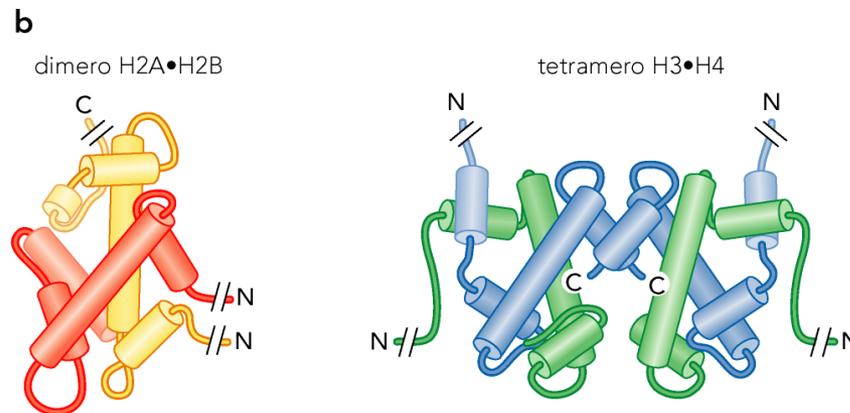
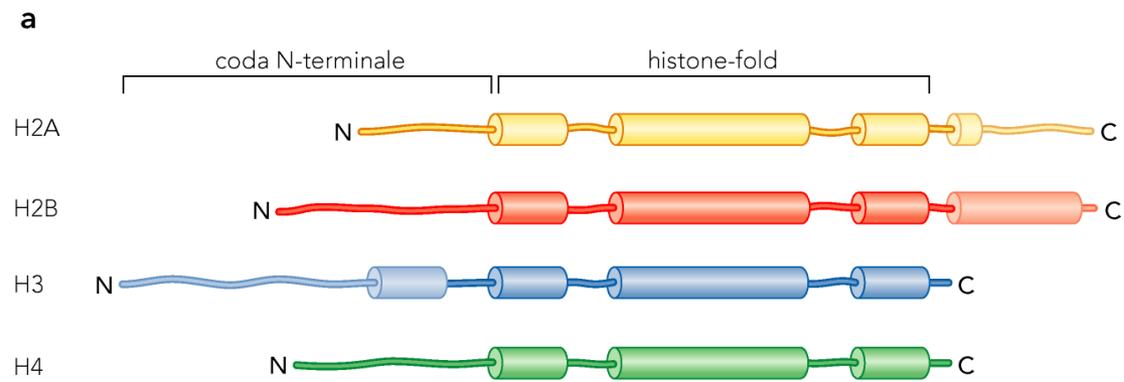
Impacchettamento del DNA eucariotico nei nucleosomi

Il nucleosoma è composto da un nucleo di 8 proteine istoniche attorno a cui è avvolto il DNA. Il DNA fra due nucleosomi è detto DNA linker. Il DNA viene compattato ~ 6 volte nei nucleosomi (I stadio del compattamento).

Il DNA associato ai nucleosomi (DNA core) è avvolto ~ 1.7 volte attorno all'ottamero (147 bp in tutte le cellule eucariotiche).

Il DNA linker invece è variabile (20-60 bp), con una lunghezza tipica per ogni organismo, che si riflette in una diversità di strutture di ordine superiore della cromatina.

Le zone di DNA non organizzate in nucleosomi sono impegnate nell'espressione genica, replicazione o ricombinazione.



Gli istoni

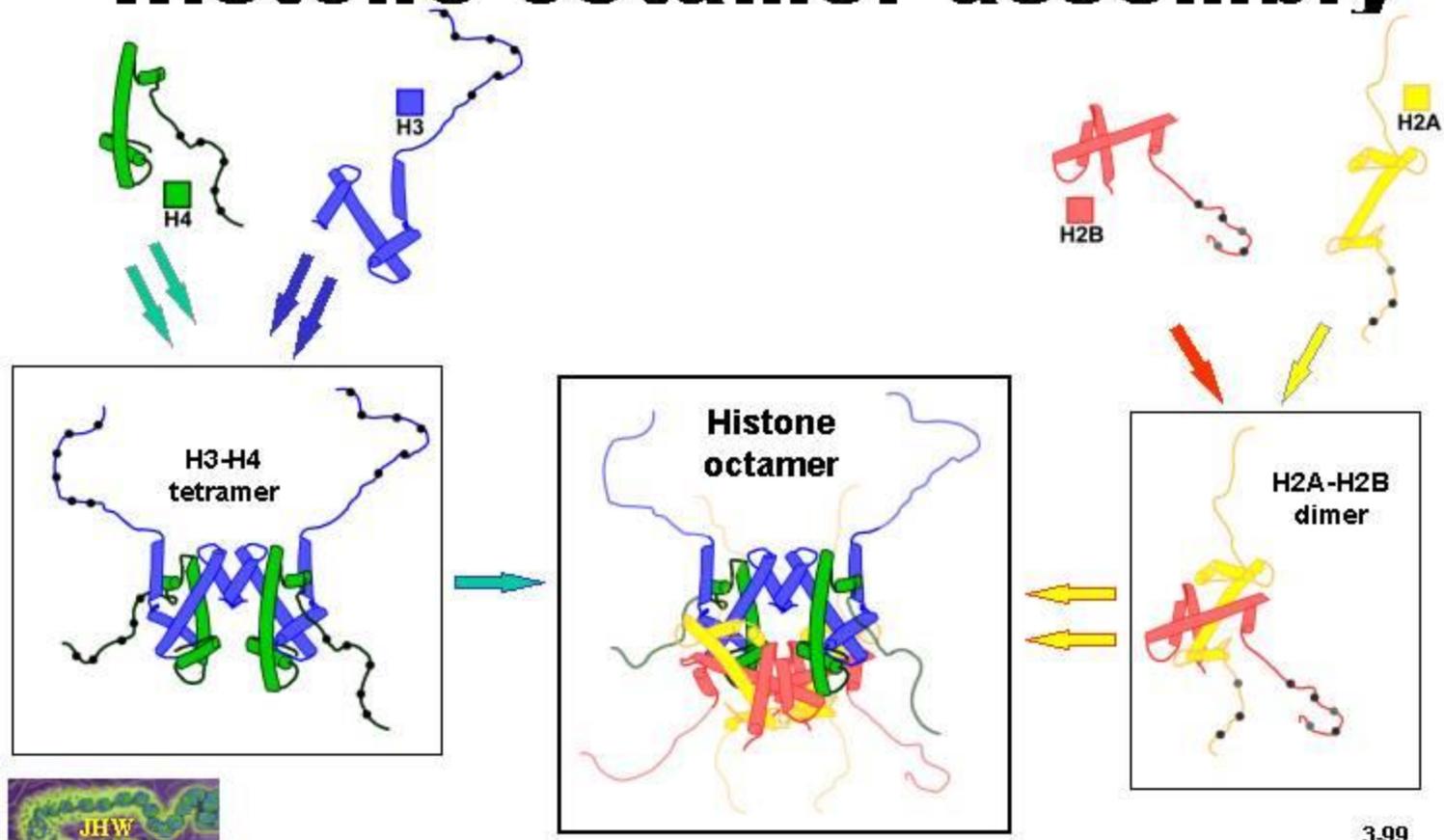
Sono le proteine più abbondanti associate al DNA eucariotico. Sono di 5 tipi: H2A, H2B, H3, H4 (formano il core) e H1 (lega il DNA linker fra 2 nucleosomi). I primi 4 sono presenti in quantità uguali nella cellula, mentre H1 è metà (1 nucleosoma).

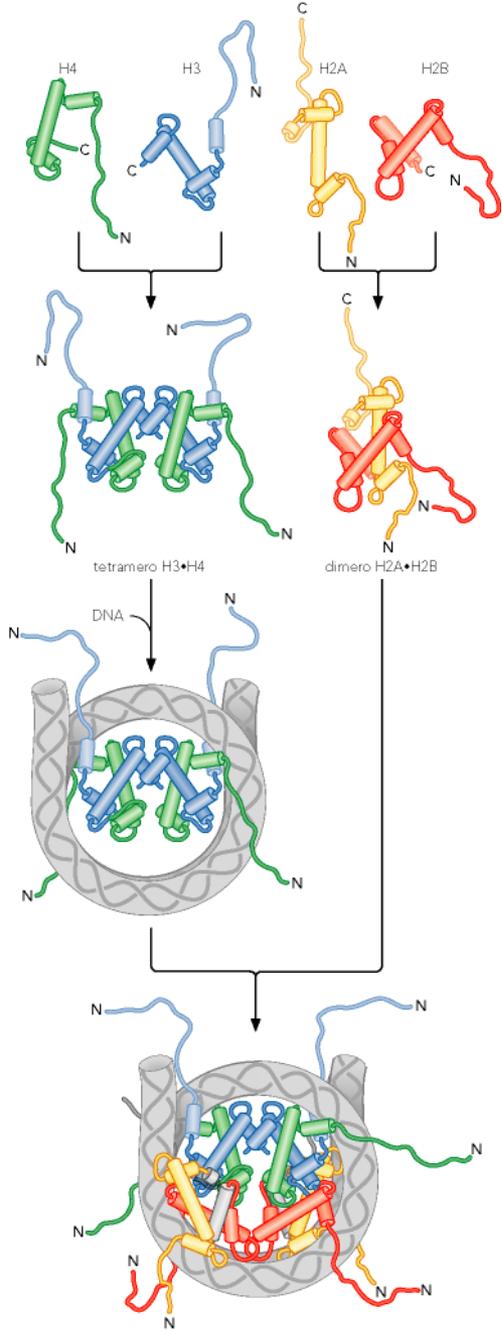
Gli istoni sono proteine basiche, cioè cariche positivamente (> 20% di lisina e arginina), di PM 11-15 kDa (H1 è 20 kDa).

Il core si assembla solo in presenza di DNA, mentre da soli in soluzione formano complessi intermedi.

Sono formati da un dominio ben conservato, detto histone-fold, formato da 3 α -eliche, coinvolte nella formazione di dimeri H2A-H2B e H3-H4, questi ultimi poi formano tetrameri.

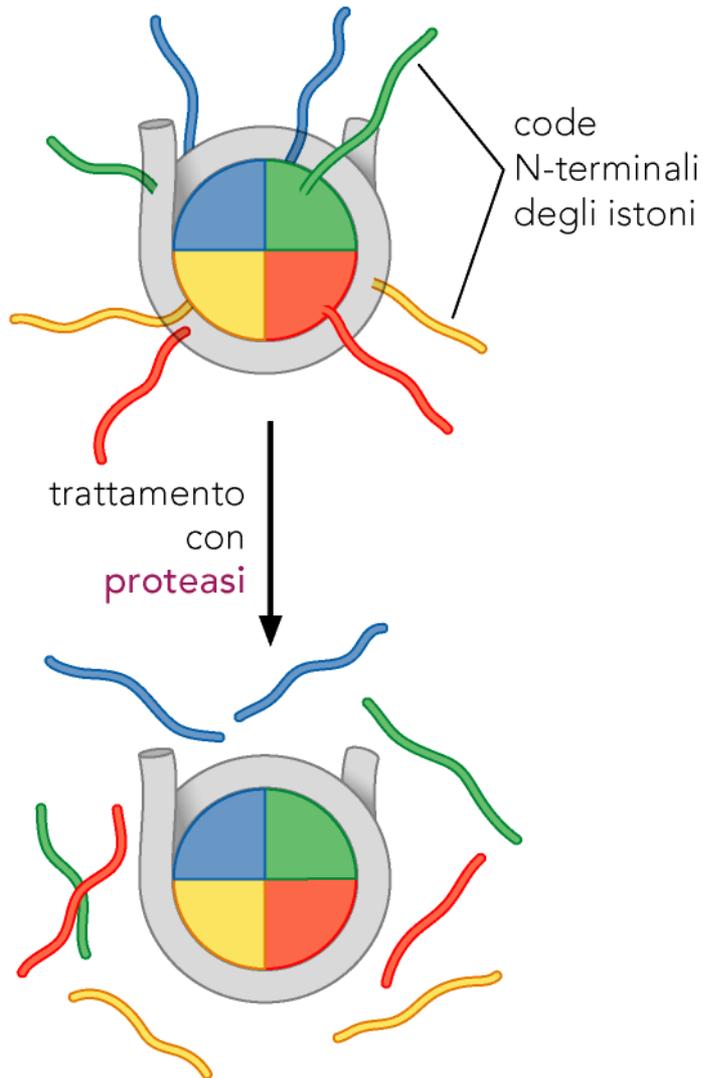
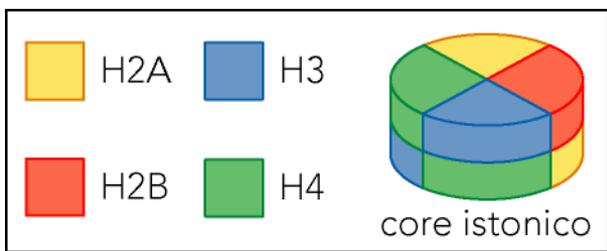
Histone octamer assembly





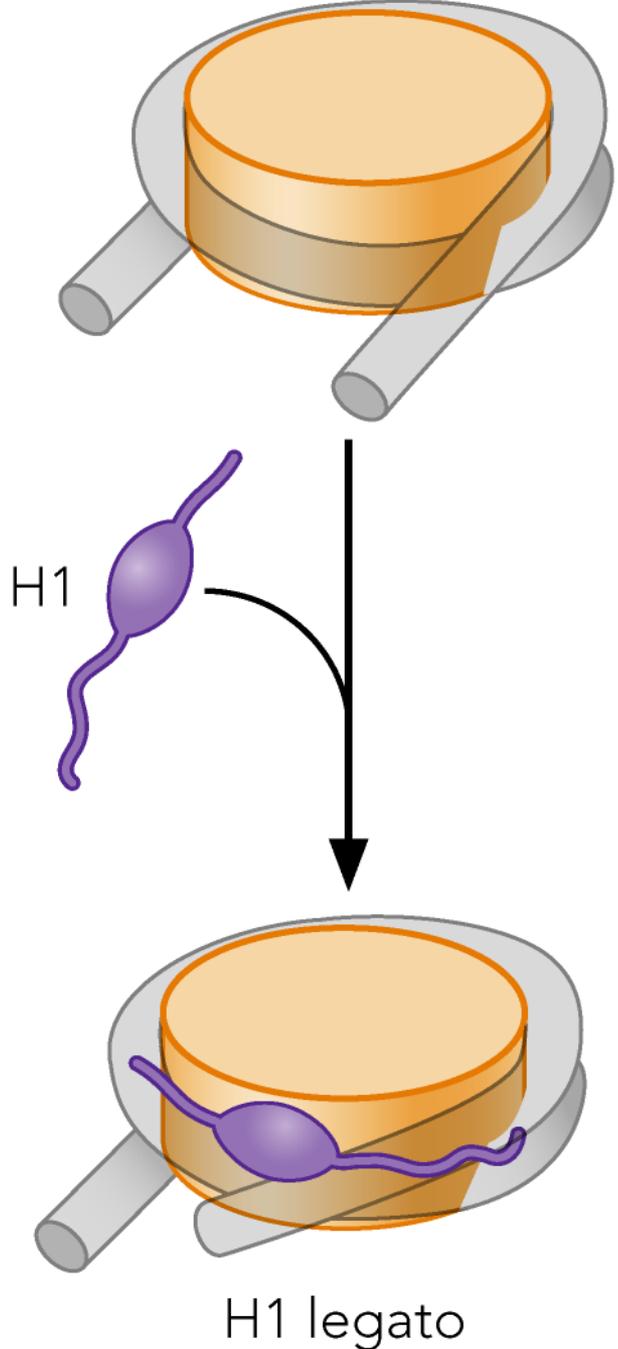
Assemblaggio di un nucleosoma

Al DNA si lega prima un tetramero (H3-H4)₂ e poi due dimeri H2A-H2B.



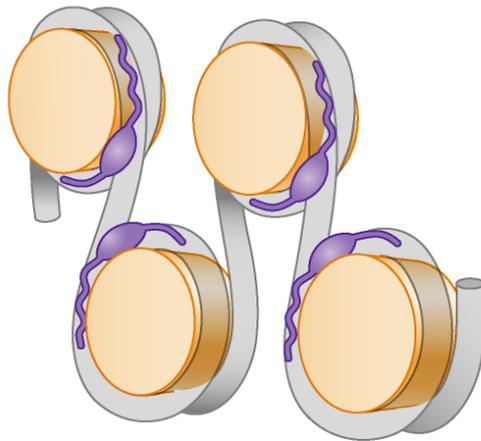
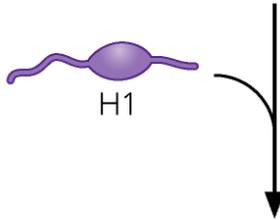
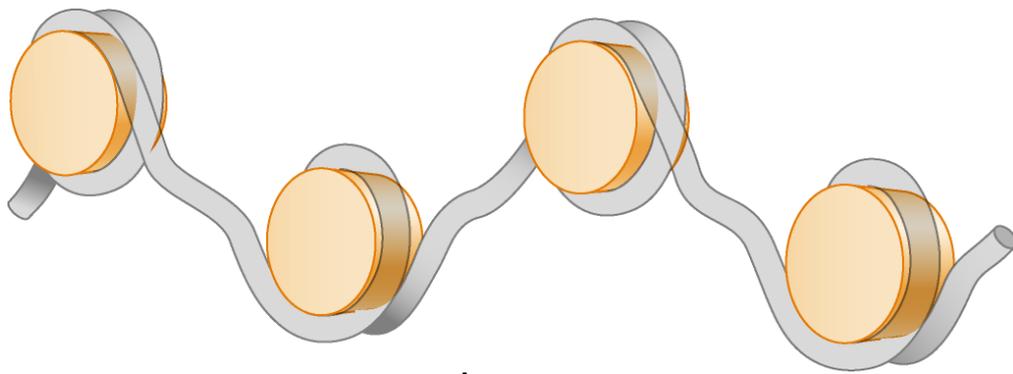
Accessibilità delle code N-terminali degli istoni core alle proteasi

Per trattamento con tripsina le code vengono digerite, mentre il core resta intatto, quindi non necessarie per l'assemblaggio del nucleosoma. Sono però sede di molte modificazioni enzimatiche, che fanno cambiare funzione al nucleosoma (fosforilazioni, acetilazioni e metilazioni di residui di lisina e serina).



Interazione dell'istone H1 col DNA nucleosomico

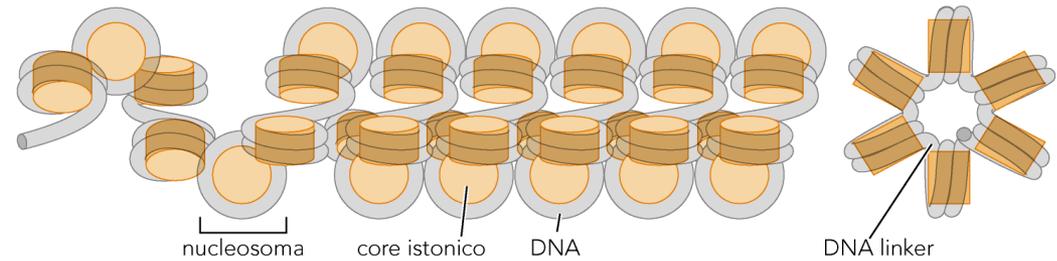
Il passaggio successivo dell'impacchettamento dei nucleosomi è mediato dall'istone H1, che interagisce sia con una porzione centrale del DNA core che con ~20 bp di uno dei due DNA linker, facendo aumentare la lunghezza di DNA arrotolato.



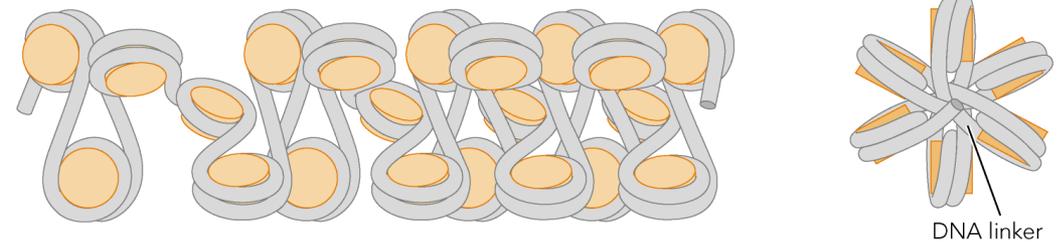
La presenza di H1 fa aumentare la quantità di DNA avvolto attorno al core istonico

Il ripiegamento del DNA linker, provocato dall'istone H1, di circa 20° rispetto all'asse diade del nucleosoma, favorisce la formazione di una struttura a zig-zag della cromatina, aumentando l'impacchettamento del DNA.

a solenoide



b zig-zag



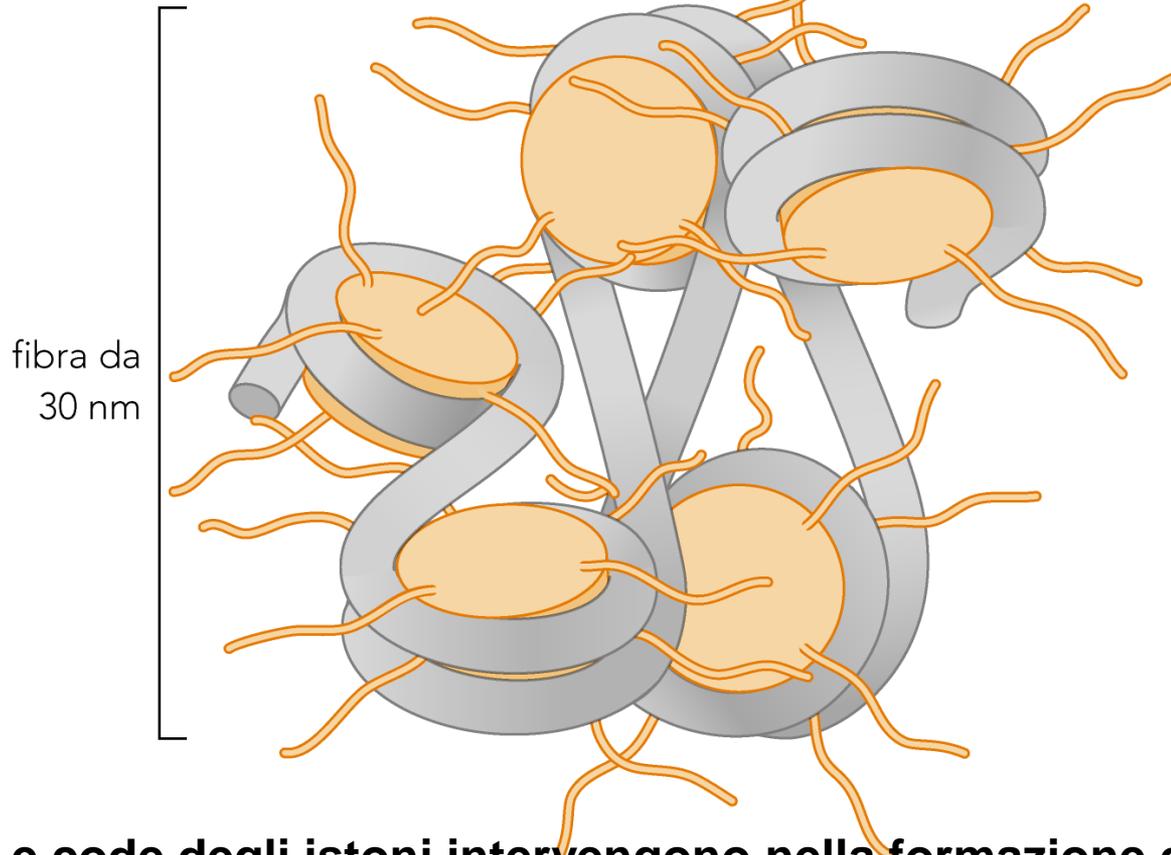
Modelli di formazione della fibra da 30 nm

In vitro ad alta conc. salina si forma una fibra di diametro 30 nm, che riduce notevolmente l'accessibilità del DNA.

La fibra 30 nm è spiegabile in base a due modelli:

a) modello solenoide: 6 nucleosomi/giro, con un buco centrale di 11 nm (visto ai raggi X e in ME) e le superfici superiore e inferiore di ciascun nucleosoma adiacenti. Il DNA linker è parte della superelica e non attraversa mai l'asse del solenoide.

b) modello a zig-zag: differisce dal precedente per la diversa posizione del DNA linker, che attraversa l'asse longitudinale della fibra. E' favorito dalla presenza di DNA linker più lunghi, variabili da specie a specie.



Le code degli istoni intervengono nella formazione della fibra 30 nm

Istoni del core, privati di coda N-terminale, sono incapaci di formare la fibra 30 nm. Nel cristallo sono state osservate legami a H fra la coda di H4 e parti di H2A-H2B di nucleosomi vicini ben conservate in tutti gli eucarioti, ma non coinvolte col legame al DNA e nella formazione dell'ottamero.

E' probabile che le code, in quanto frequentemente modificate nella cellula, influenzino la formazione della fibra da 30 nm e di altre strutture d'ordine superiore.

Caratteristiche degli istoni

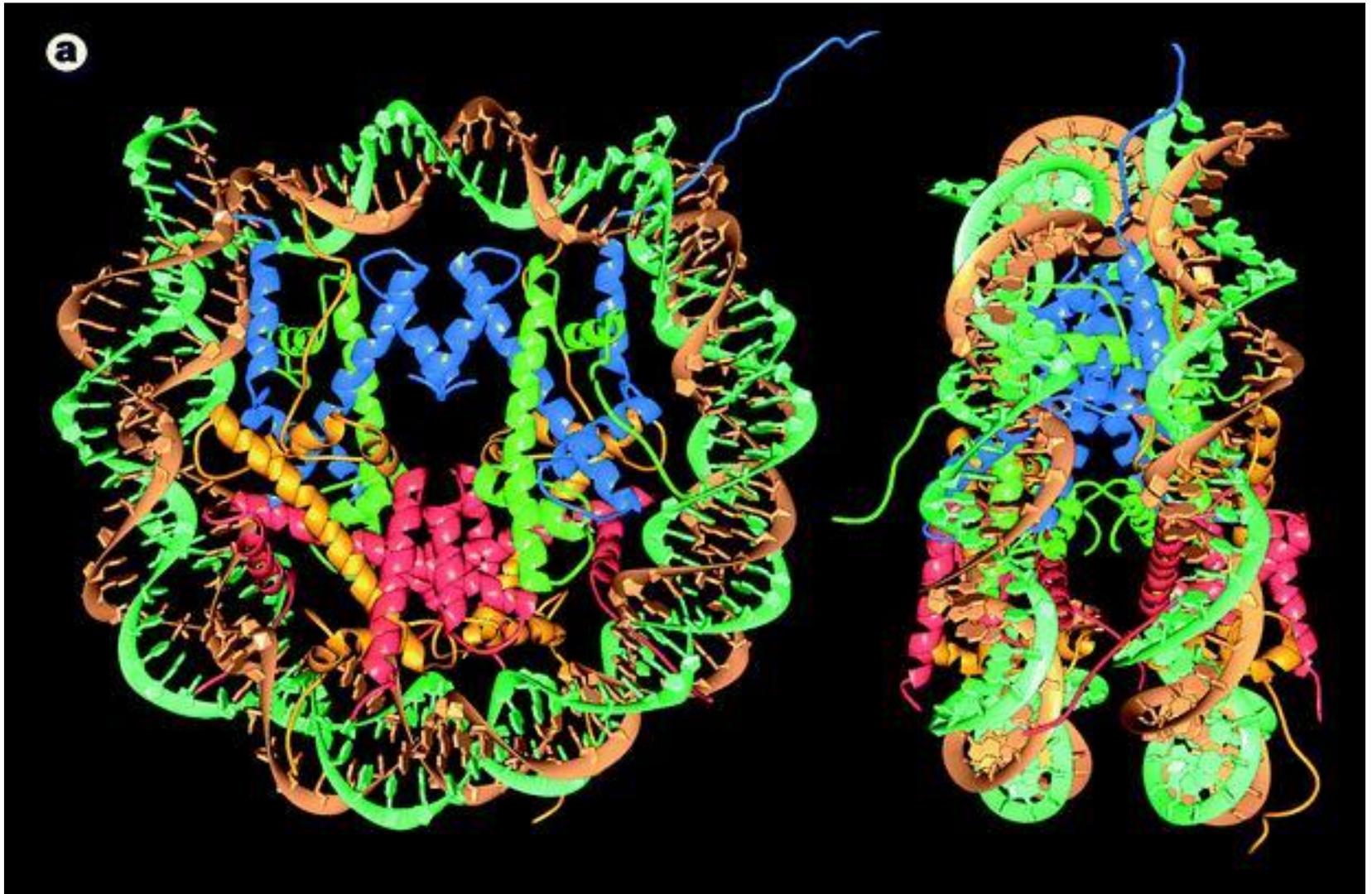
Tipo	Massa mol. N.mol./ kDa nucleosoma	N.mol./ nucleosoma	Caratteristiche
H1	21,5	1	Esterno al “core”, più variabile
H2A	14,0	2	Parte del “core”, molto conservato
H2B	13,8	2	Parte del “core”, molto conservato
H3	15,4	2	Parte più interna del “core”, altamente conservato
H4	11,3	2	Parte più interna del “core”, altamente conservato

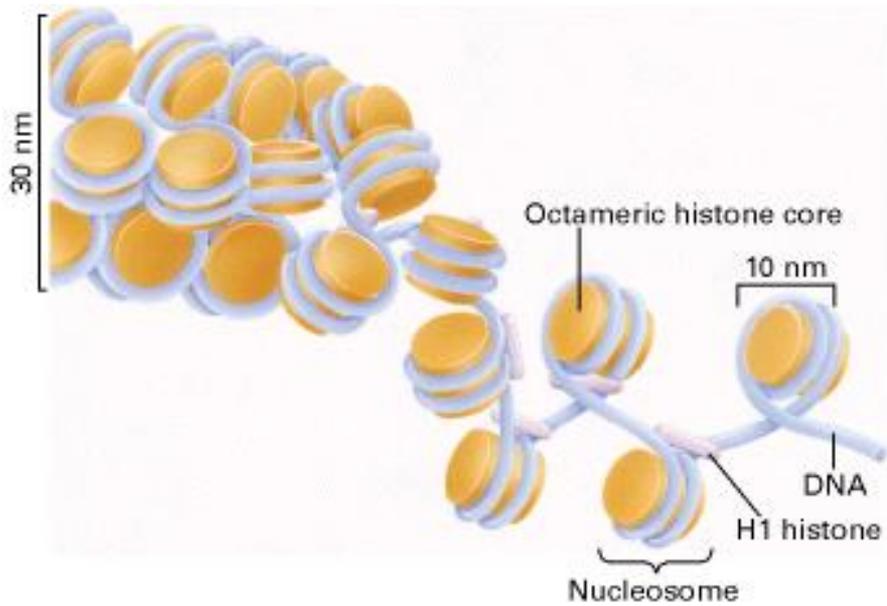
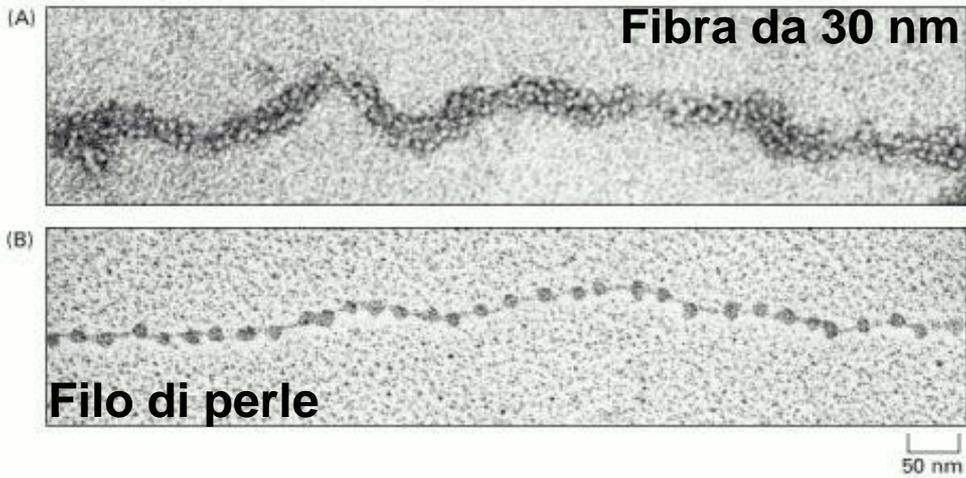
Massa molecolare degli istoni per nucleosoma 130 kDa circa

**DNA per nucleosoma: 200 paia di basi circa (pari a 140 kDa circa),
di cui 146 strettamente associate, il resto “linker”**

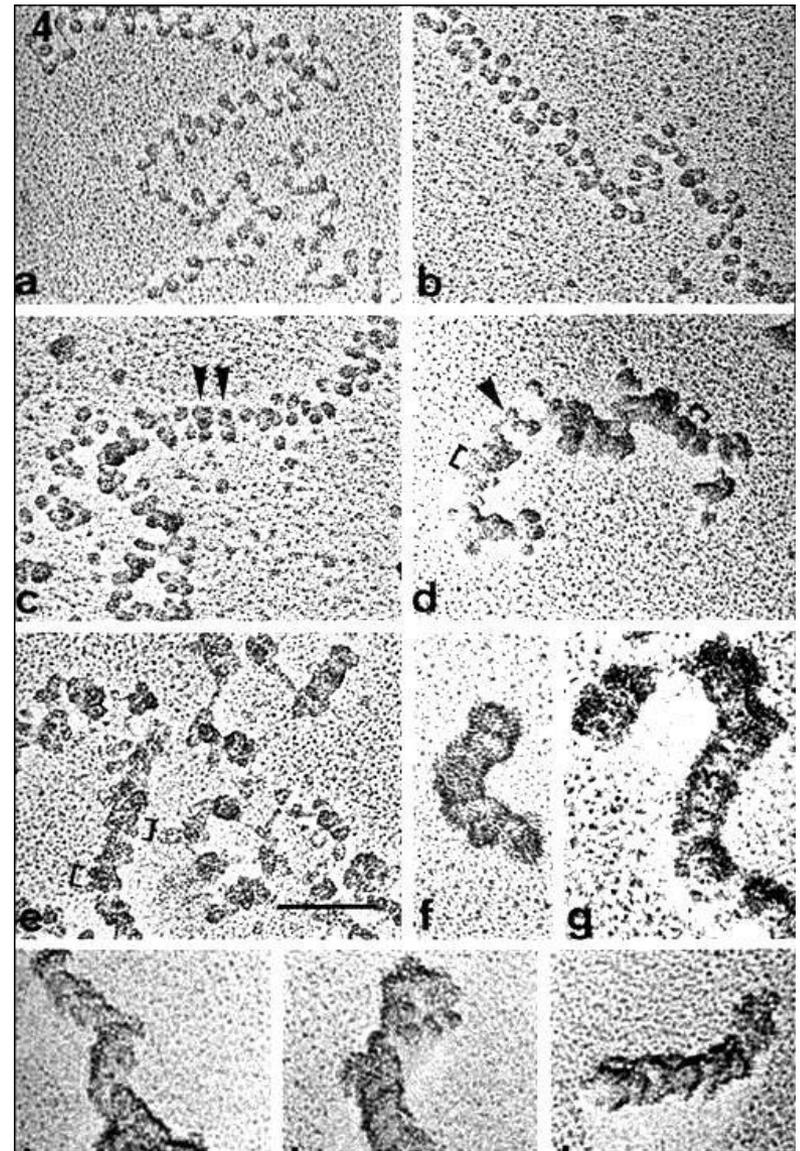
Massa molecolare totale per nucleosoma: 270 kDa circa

IL NUCLEOSOMA

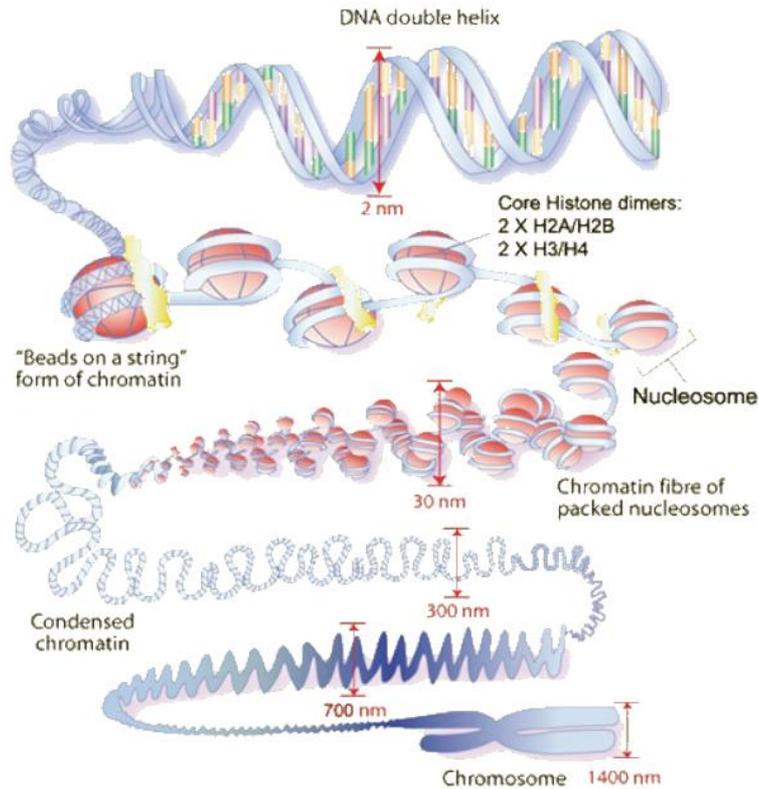




La cromatina al microscopio elettronico

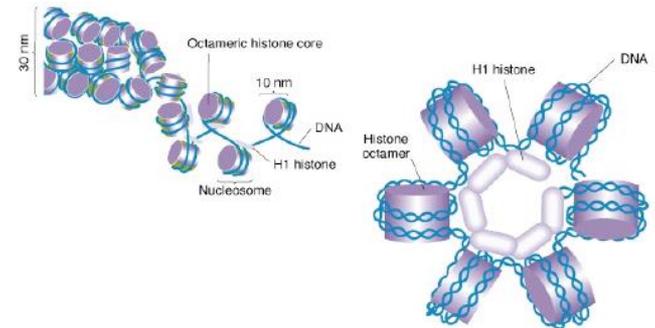


Vari livelli di compattazione

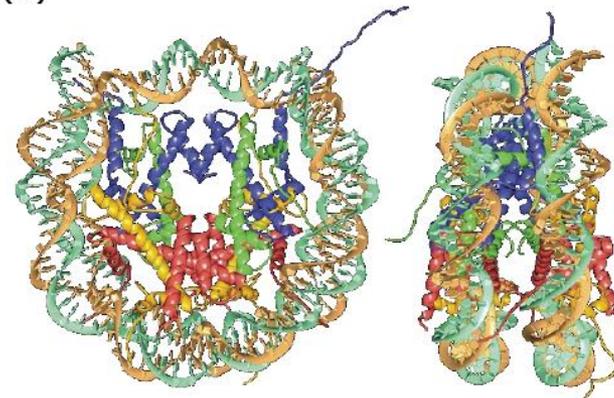


(a)

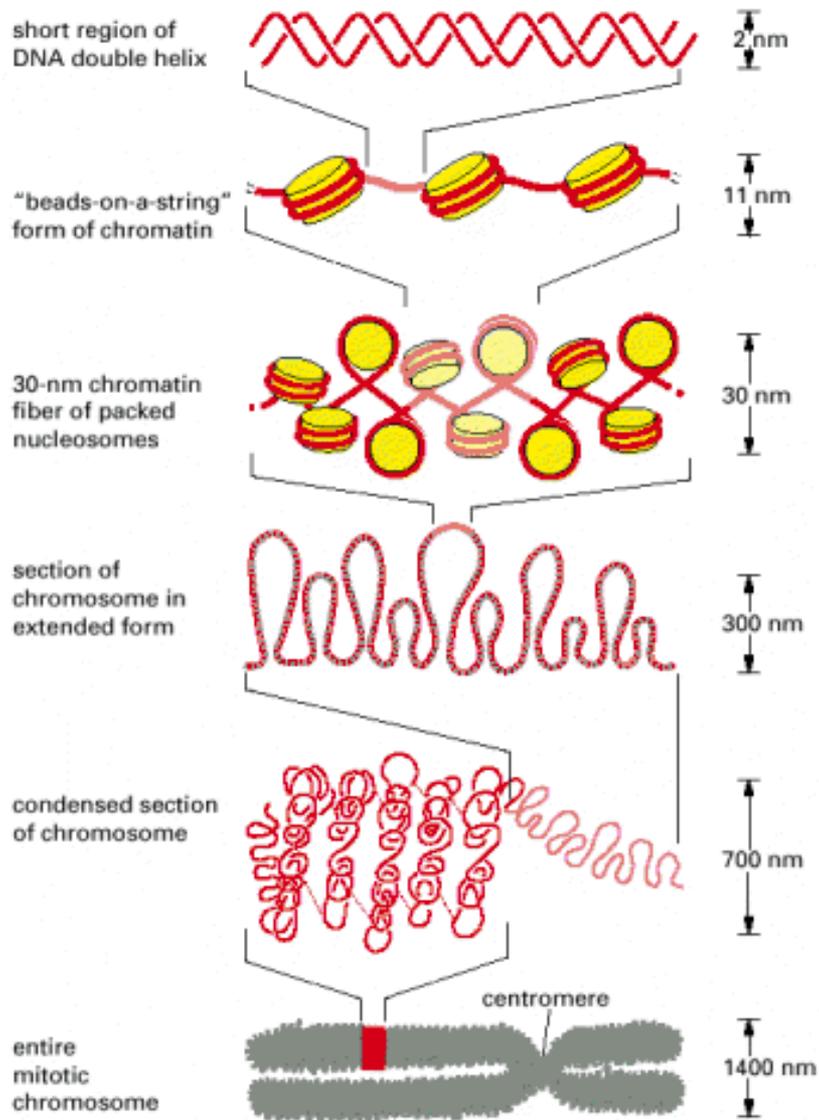
Fig 5.1 (a) Organizzazione gerarchica della struttura compatta del DNA. (b) Livello intermedio di compattazione: l'esamero di nucleosomi. (c) il nucleosoma: DNA in verde-ocra, proteine istoniche in blu-verde-giallo-rosso. È visibile la coda degli istoni.



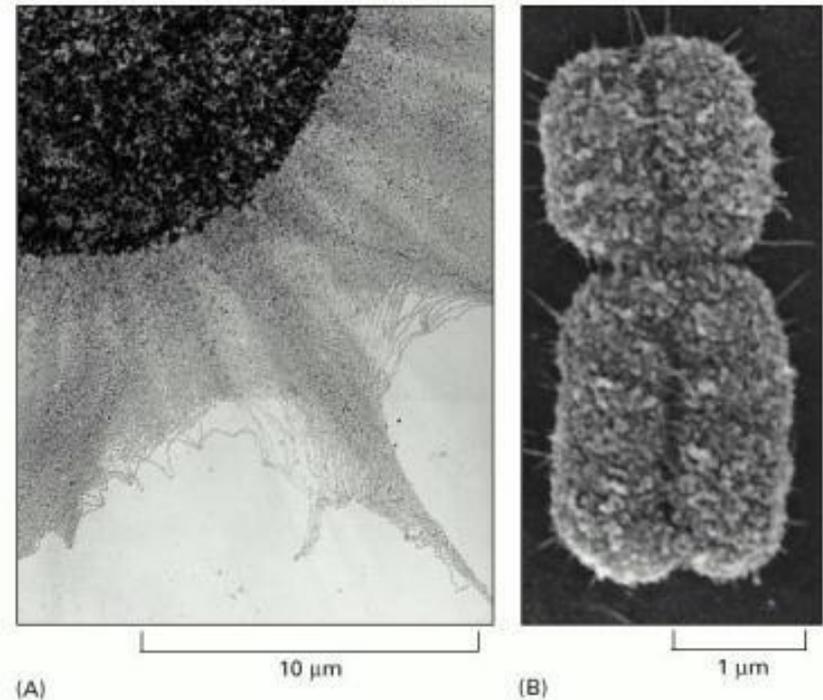
(b)



(c)



NET RESULT: EACH DNA MOLECULE HAS BEEN PACKAGED INTO A MITOTIC CHROMOSOME THAT IS 10,000-FOLD SHORTER THAN ITS EXTENDED LENGTH



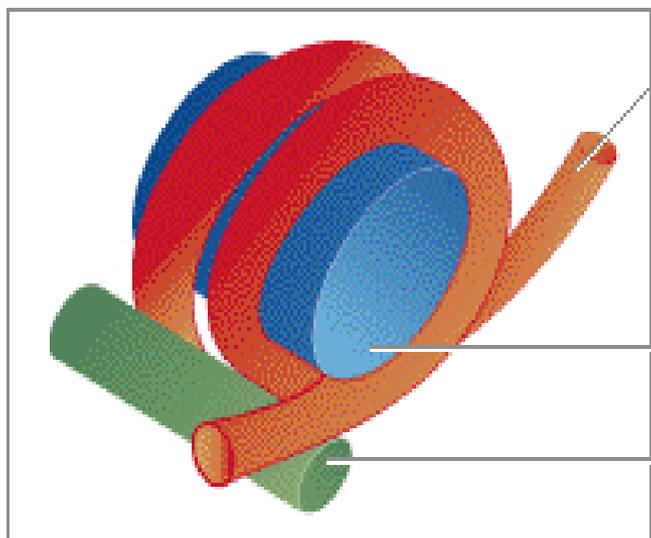
A comparison of extended interphase chromatin with the chromatin in a mitotic chromosome.

(A) An electron micrograph showing an enormous tangle of chromatin spilling out of a lysed interphase nucleus.

(B) A scanning electron micrograph of a mitotic chromosome: a condensed duplicated chromosome in which the two new chromosomes are still linked together. The constricted region indicates the position of the centromere.

Note the difference in scales.

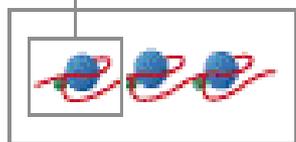
Avvolgimento del DNA su un nucleo di istoni e formazione di un nucleosoma



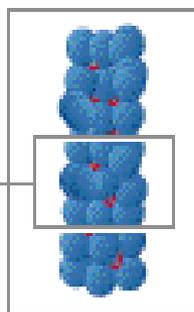
Super-elica di DNA
(l'asse della doppia elica del DNA si avvolge a sua volta nello spazio in modo elicoidale)

Nucleo di 8 molecole di istoni

Molecola di istone che delimita il nucleosoma

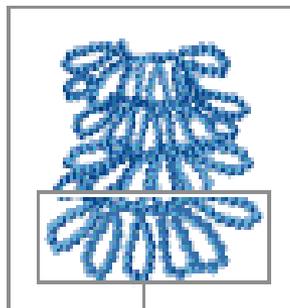


Formazione di un filamento di nucleosomi (diametro 10 nm), che costituisce i filamenti di cromatina visibili nel nucleo quando la cellula non è in fase di divisione (filamento cromatinico)

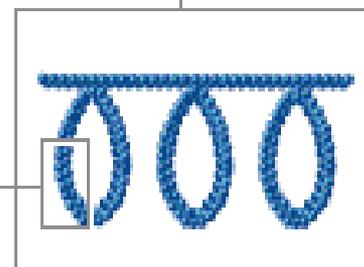


Avvolgimento del filamento e formazione di una fibra (diametro 25 nm)

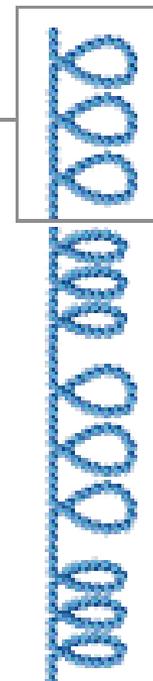
Avvolgimento del filamento ad anse e formazione del cromosoma (diametro 1 μm)



Formazione di un filamento ad anse (diametro 0,5 μm)

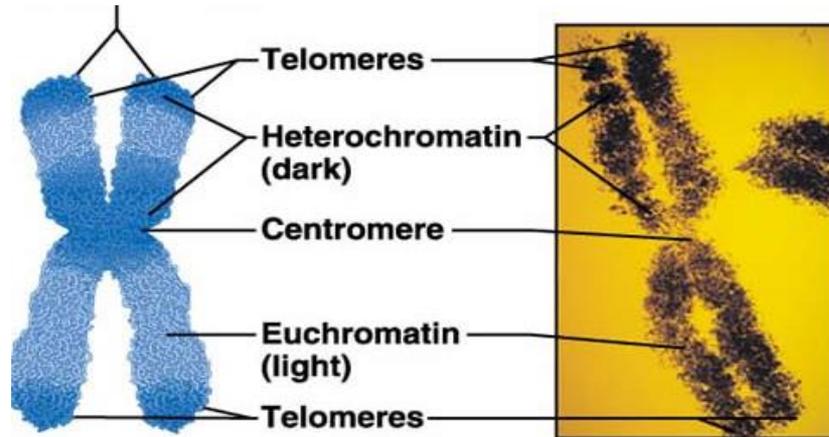


Formazione di bande orizzontali più o meno scure, dovute alla maggiore o minore distanza tra le anse del filamento (bande G)

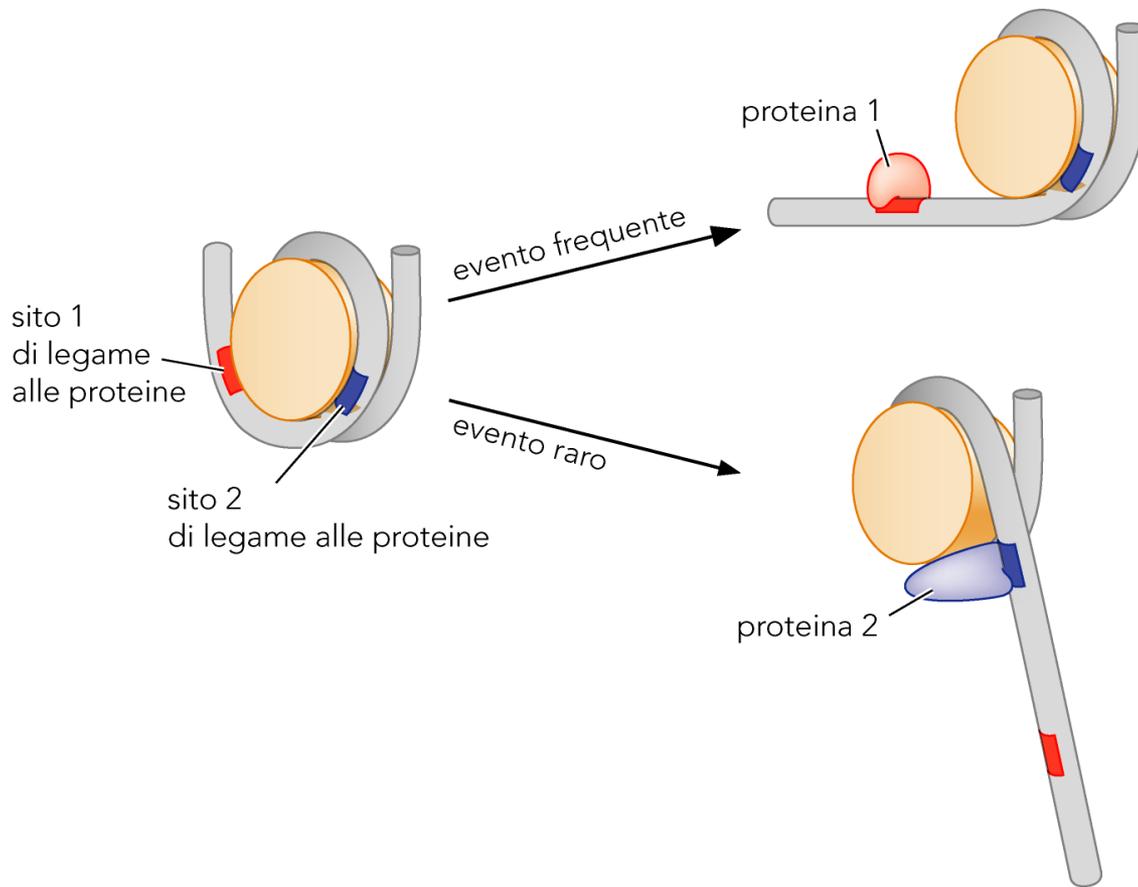


Cromatina

- La cromatina è formata da DNA, avvolto su gruppi di istoni (proteine basiche), formando un nucleosoma, e da proteine non-istoniche (proteine neutre o acide); essa è poi ripiegata in vario modo
- *eucromatina*: meno condensata e corrisponde a zone in cui vi è un'intensa attività di trascrizione per la sintesi proteica
- *eterocromatina*: più condensata, non sembra presentare attività di trascrizione.

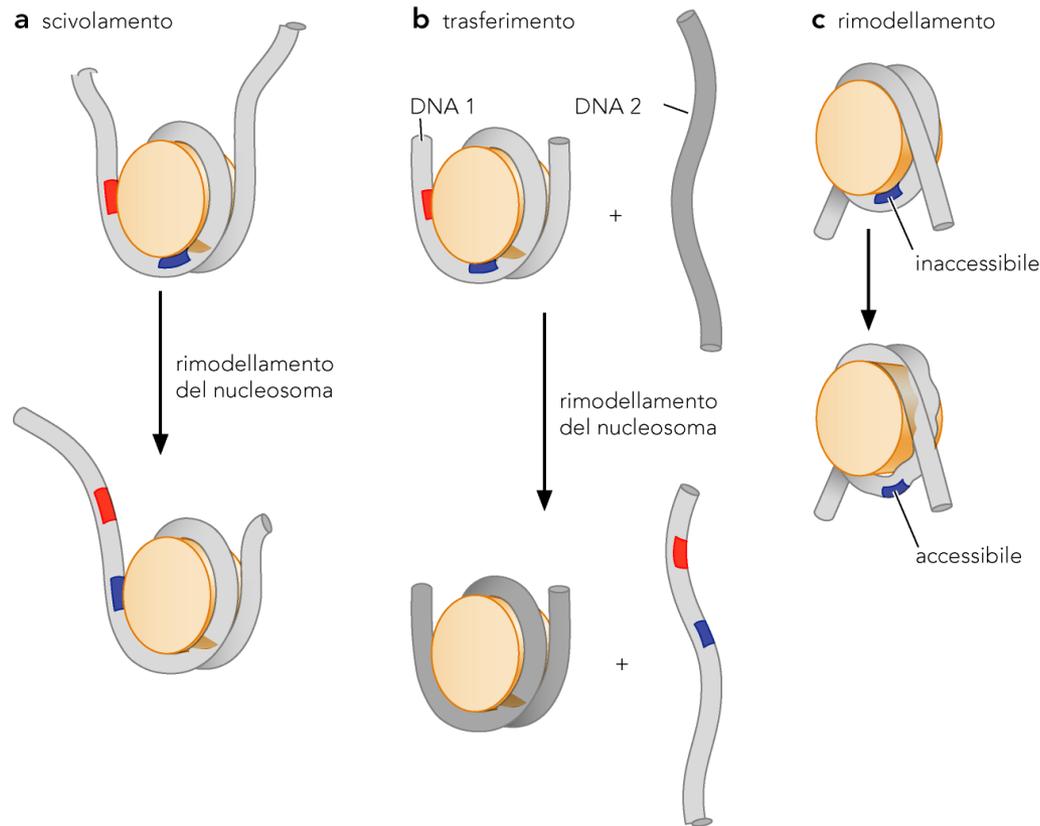


Regolazione della struttura della cromatina



Natura dinamica dell'associazione DNA – istoni

L'accessibilità del DNA nei nucleosomi cambia con le diverse esigenze della cellula, data la natura non covalente delle interazioni fra istoni e DNA. Uno srotolamento più o meno rilevante del DNA, senza distacco, può liberare siti di legame specifici per certe proteine. In generale, più i siti sono interni al nucleosoma, meno sono accessibili.

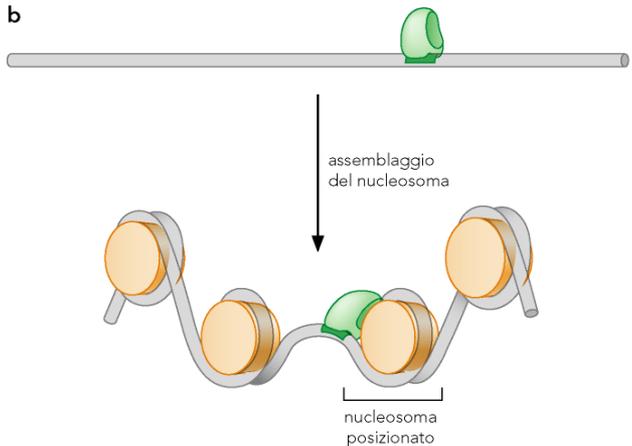
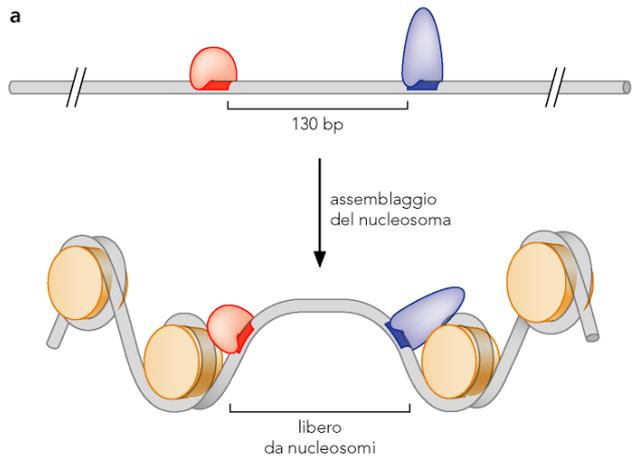


Complessi di rimodellamento nucleosomico

Molti complessi proteici influenzano il posizionamento dei nucleosomi e le interazioni DNA-istoni, per idrolisi di ATP, tramite:

- **scivolamento** dell'ottamero sul DNA
- **trasferimento** da un tratto di DNA ad un altro
- **rimodellamento** del nucleosoma, tale da permettere un maggior accesso al DNA.

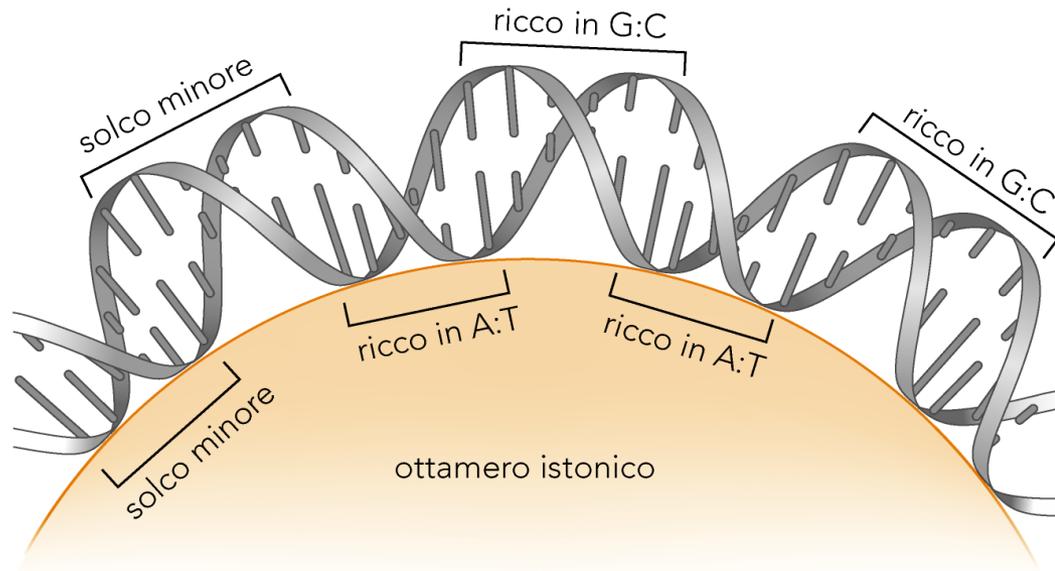
Esistono molti tipi di complessi di rimodellamento nei diversi tessuti.



IL posizionamento dei nucleosomi può dipendere da proteine che legano il DNA

Se due proteine si legano sul DNA a distanza < 150 bp, non si può formare un nucleosoma intermedio

Se alcune proteine legano fortemente il nucleosoma, lo posizionano preferenzialmente nelle loro vicinanze.

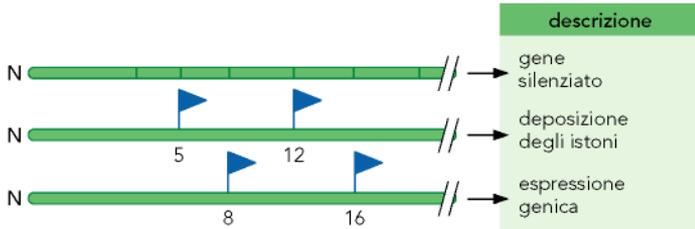
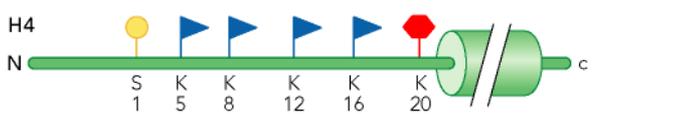
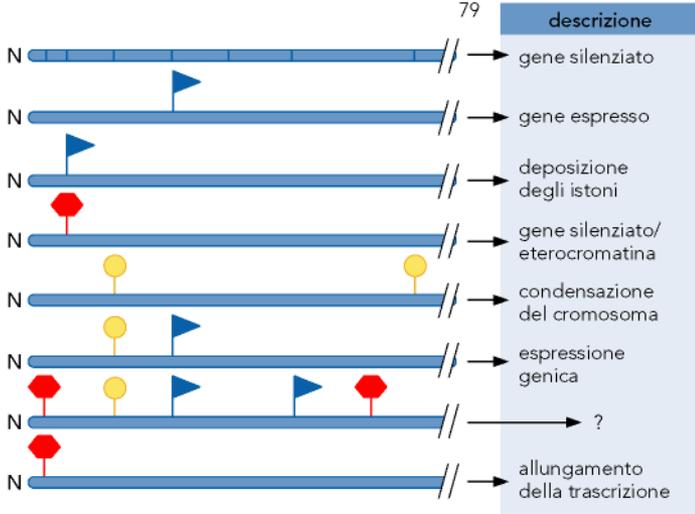
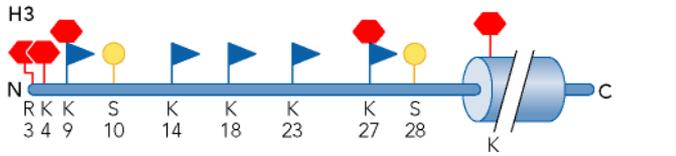
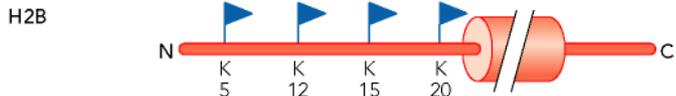


I nucleosomi si formano in modo preferenziale sul DNA ricurvo

Sequenze di DNA ricche in AT, che ripiegano il DNA dalla parte del solco minore, favoriscono il posizionamento del nucleosoma. Viceversa, sequenze ricche in GC, avendo tendenza inversa, posizionano il solco minore esternamente.

Tratti di 5 AT + 5 CG sono siti di posizionamento preferenziale del nucleosoma, anche se non richiesti per l'assemblaggio.

Il posizionamento può avere sia un effetto positivo che negativo sull'accessibilità di siti specifici del DNA.



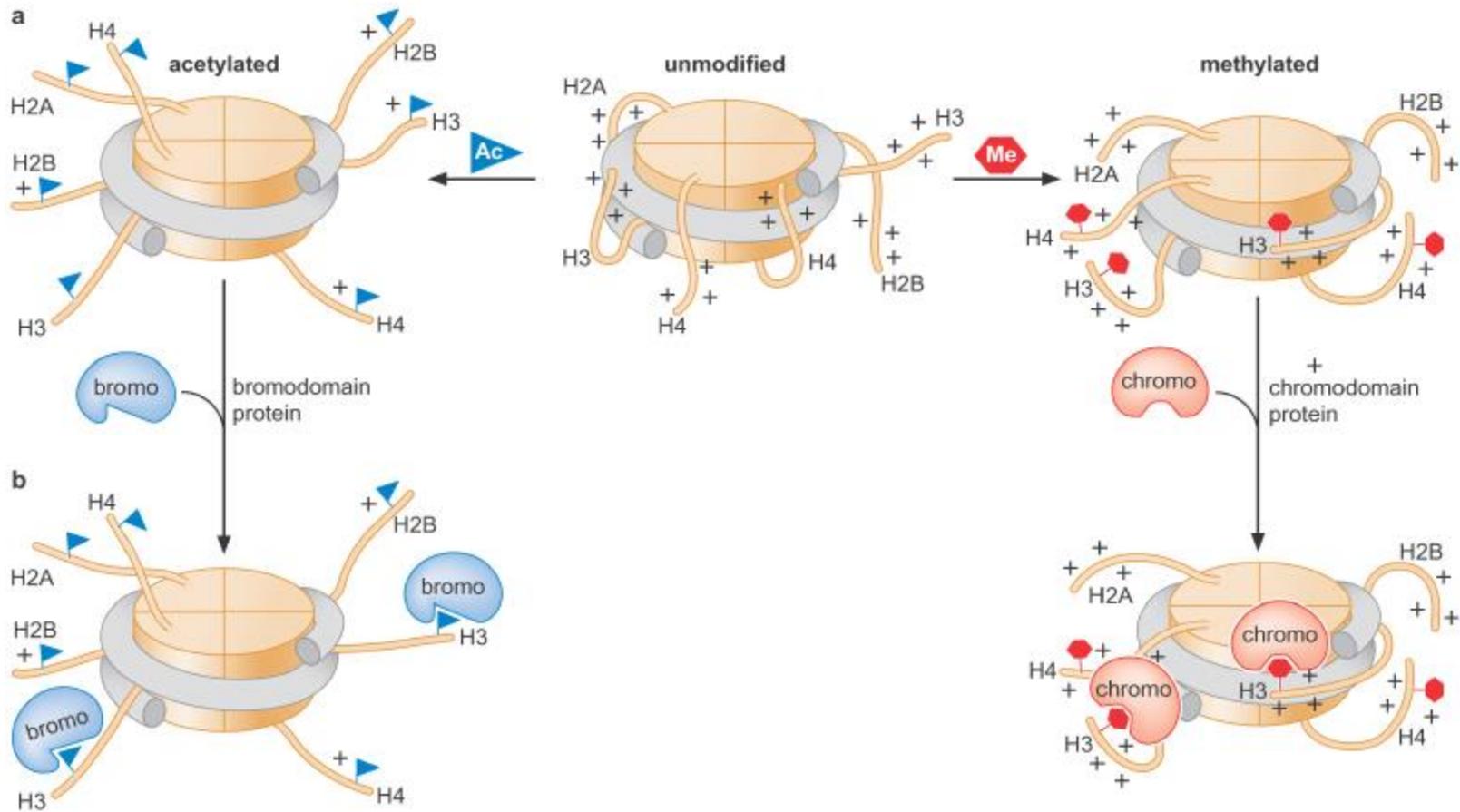
Modificazioni degli istoni alterano la funzione della cromatina

Istoni isolati dalle cellule risultano fortemente modificati nella zona N-terminale: **acetile o metile su lisina (K)**, **fosfato su serina (S)**.

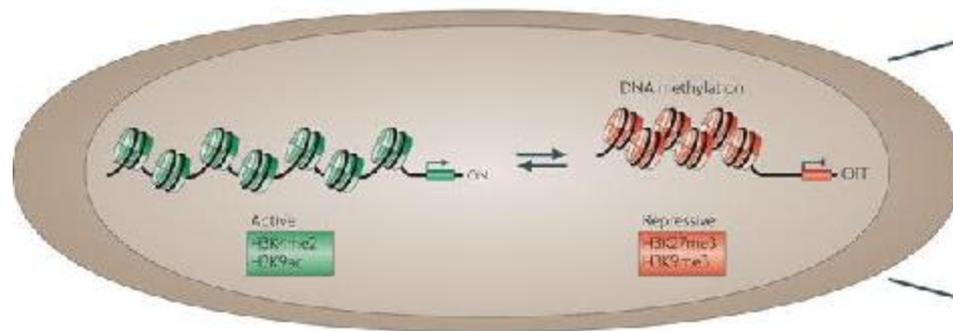
- Acetilazione associata con attivazione della trascrizione,
- Deacetilazione associata con repressione,
- Metilazione con entrambe (dipende dalla posizione dell'amminoacido modificato).
- Fosforilazione di H3 in cromatina altamente condensata (cromosomi mitotici).

Acetilazione e fosforilazione riducono la carica positiva netta degli istoni, quindi l'affinità per il DNA.

Si pensa che esista un codice di riconoscimento delle modifiche da parte delle proteine non istoniche (**codice istonico**).



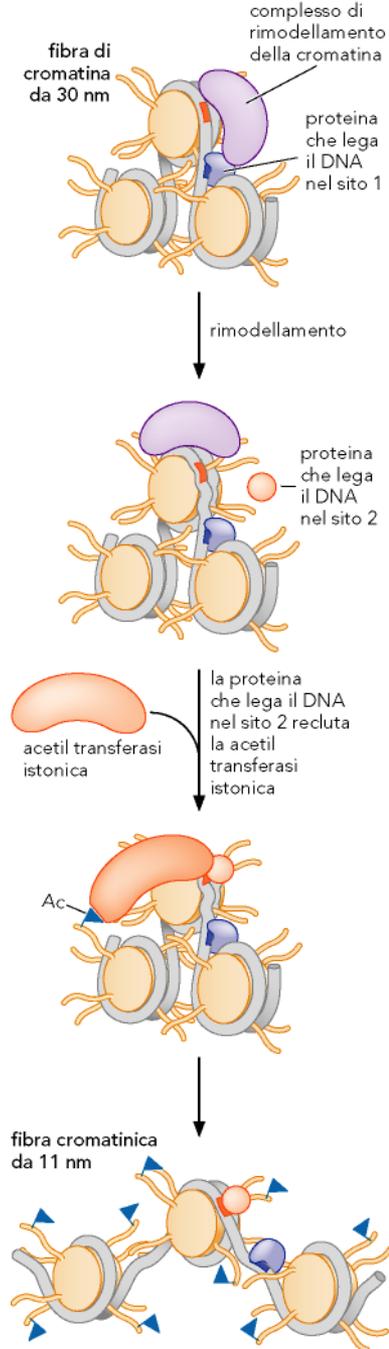
Histone Modifications Affect Chromatin Structure



H3K4 methylation and H3K9 acetylation are hallmarks of active chromatin

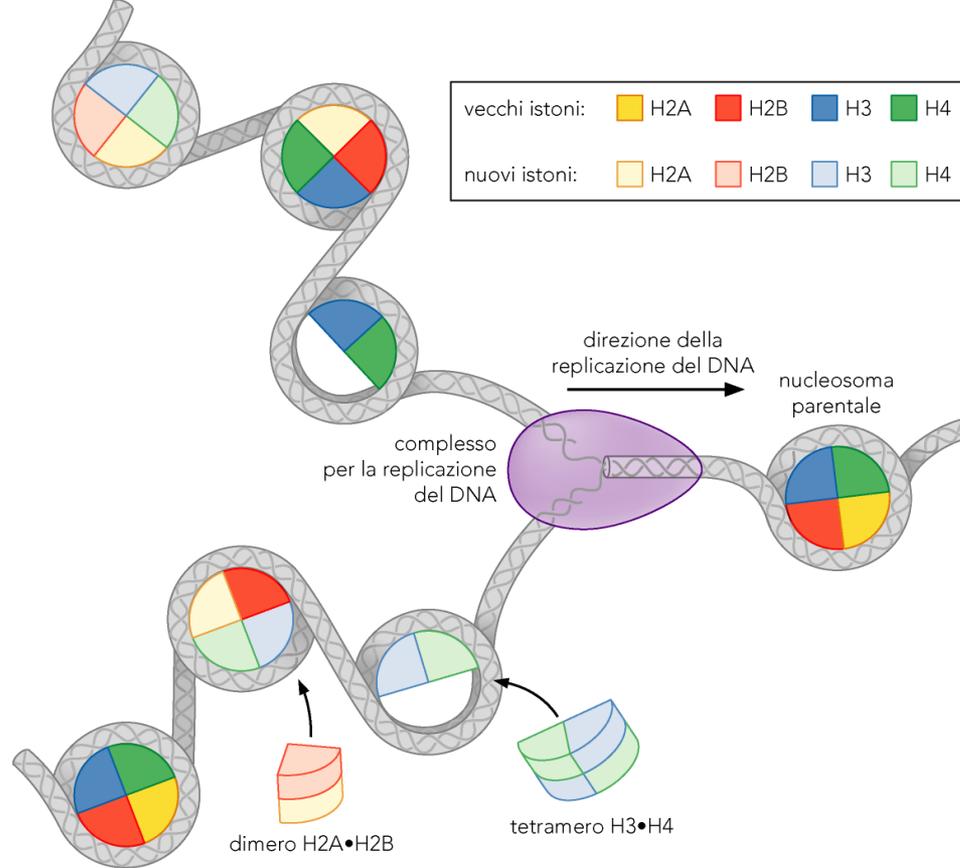
H3K27 methylation and H3K9 methylation are hallmarks of silent chromatin

from Johnstone and Baylin, *Nature Rev. Genet.* 11, 806 (2010)



Complessi di rimodellamento e enzimi che modificano gli istoni agiscono di concerto sulla cromatina

Proteine specifiche reclutano gli enzimi responsabili delle modificazioni istoniche. Acetilasi, deacetilasi e metilasi istoniche, che fanno parte di complessi multiproteici, agiscono su particolari residui istonici. Il reclutamento di questi enzimi su particolari regioni del DNA, indotta da queste proteine non istoniche, è responsabile del posizionamento dei nucleosomi sulla cromatina e modulazione dell'espressione genica, modificando l'accessibilità del DNA.

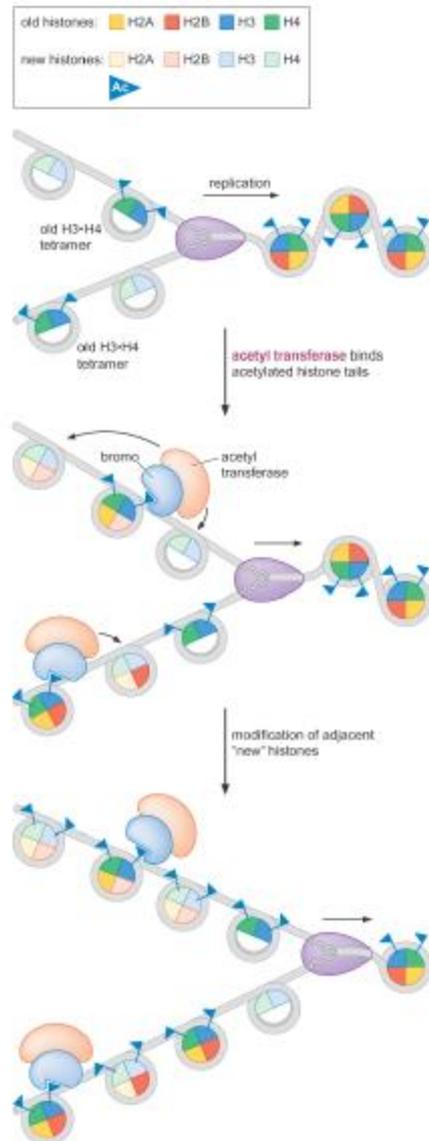


Assemblaggio dei nucleosomi dopo la replicazione

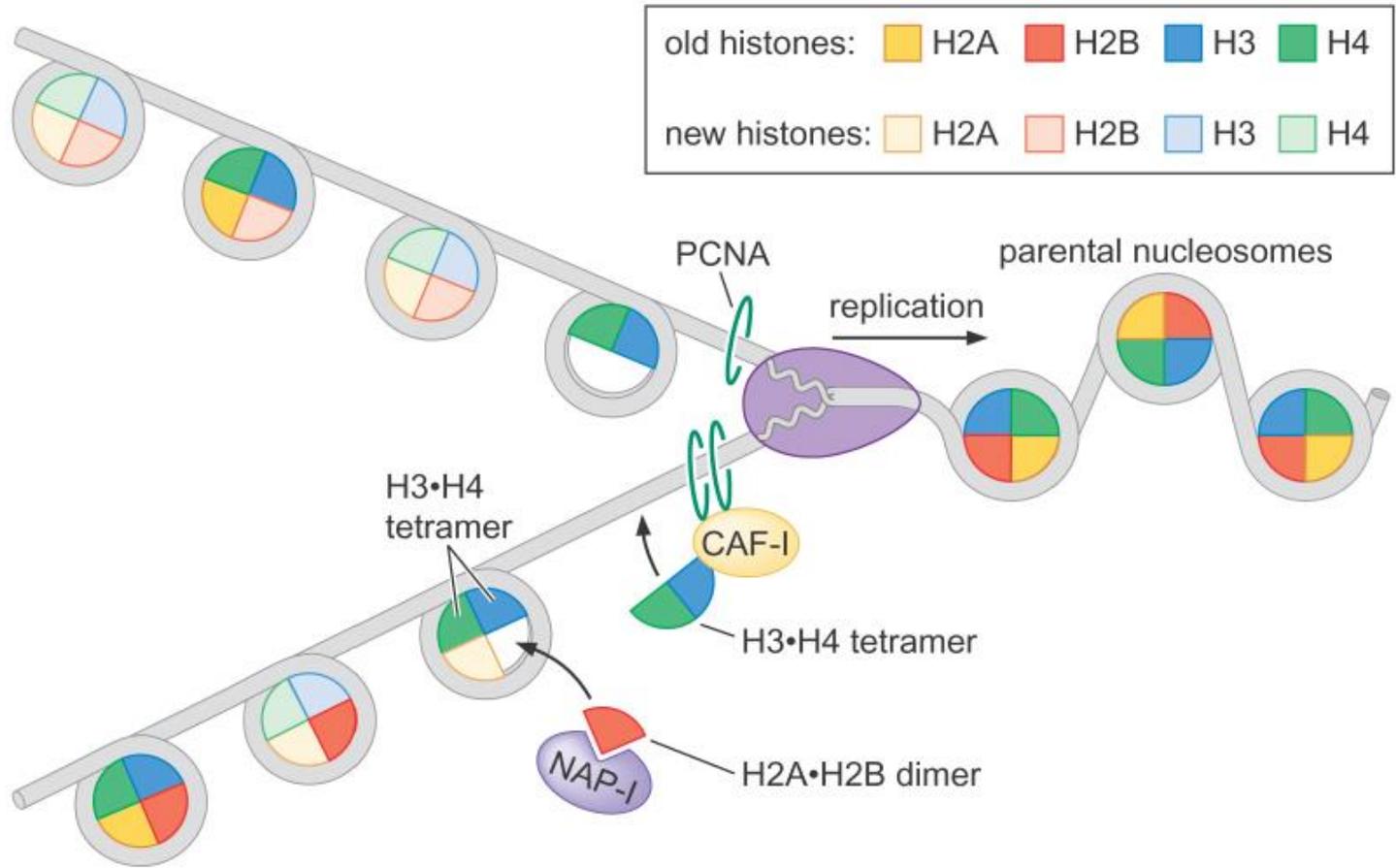
La replicazione del DNA richiede un parziale disassemblaggio dei nucleosomi in tetrameri H3-H4 e dimeri H2A-H2B, ma il DNA viene rapidamente riorganizzato in cromatina per riassociazione del tetramero H3-H4, dei dimeri H2A-H2B e di H1.

Metà dei nucleosomi sui cromatidi fratelli deve necessariamente essere di neosintesi, mentre i vecchi istoni si ridistribuiscono su entrambi i cromosomi: **il tetramero H3-H4 resta legato a caso ad una delle due nuove doppie eliche**, mentre i dimeri H2A-H2B rilasciati entrano a far parte del pool cellulare dei dimeri H2A-H2B, per poi riassociarsi a caso. Ciò permette di propagare le modifiche istoniche parentali e di posizionare i vecchi istoni in zone simili a quelle nel cromosoma parentale.

Ereditarietà dei nucleosomi



Chaperonine e riassettaggio dei nucleosomi



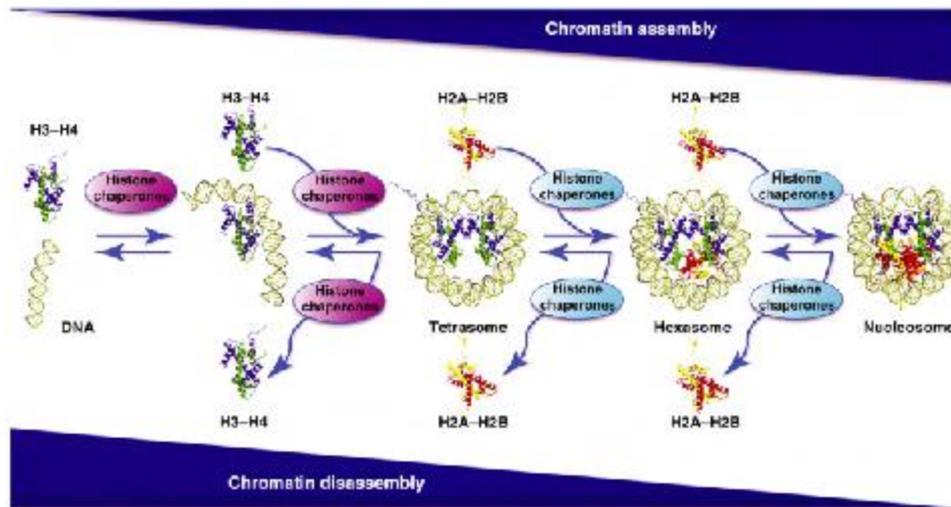
C

o

p

y

Assembly of Nucleosomes



Histone chaperones assemble histones into nucleosomes

Histone chaperones prevent non-specific associations of histones with DNA

Histone chaperones prevent formation of deleterious off-pathway intermediates

from Das *et al.*, *Trends Biochem.Sci.* 35, 476 (2010)

Struttura e complessità di genomi procariotici ed eucariotici

Struttura e dimensioni di genomi cellulari

Procarioti (Bacteria e Archaea):

Tipicamente c'è **un solo cromosoma** nel citoplasma, contenente una molecola di **DNA circolare** lungo da **0,5 a 10 milioni** di paia di basi circa. Spesso possono essere presenti anche uno o più **plasmidi**, piccoli DNA circolari accessori, costituiti da alcune migliaia di paia di basi. Sono note eccezioni: alcuni procarioti con due o più cromosomi circolari, o un cromosoma lineare.

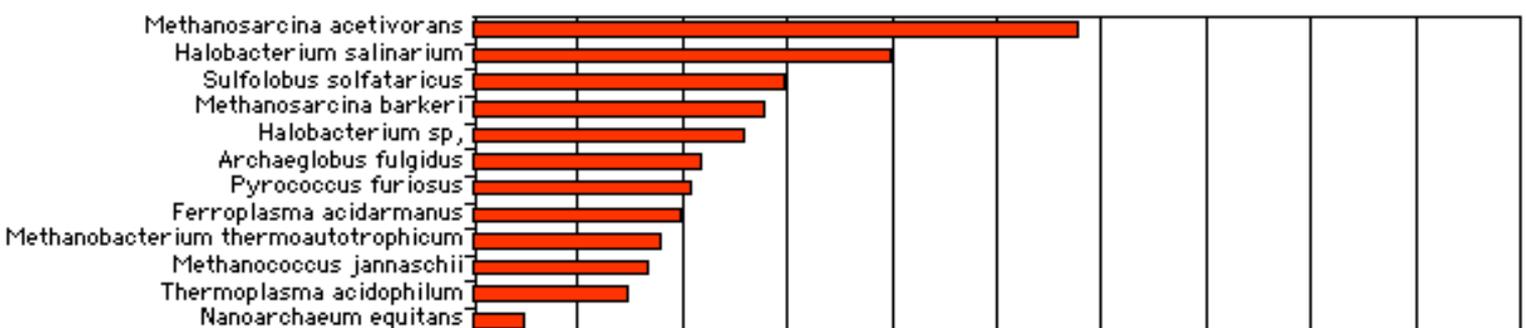
Eucarioti (unicellulari, piante, funghi, animali):

Di regola c'è **più di un cromosoma nel nucleo**, mediamente poche decine, anche se sono noti casi di parecchie centinaia e, all'altro limite, di uno solo. Ciascun cromosoma contiene una molecola di DNA lineare. Nel nucleo può essere presente un solo corredo di cromosomi diversi (la cellula si dice **aploide**), due serie (**diploide**), e sono i casi più frequenti, o più (tri-, tetra-, esa-, ... poli-ploide). Le **dimensioni di un corredo aploide** variano da **una decina di milioni** di paia di basi circa (ad esempio i lieviti o qualche alga unicellulare) a un **centinaio di miliardi** (alcune felci, i pesci polmonati e gli anfibi urodela) o anche più in alcune amebe. **Non si riscontra relazione tra numero di cromosomi, lunghezza del loro DNA, e dimensioni del genoma (il paradosso C).**

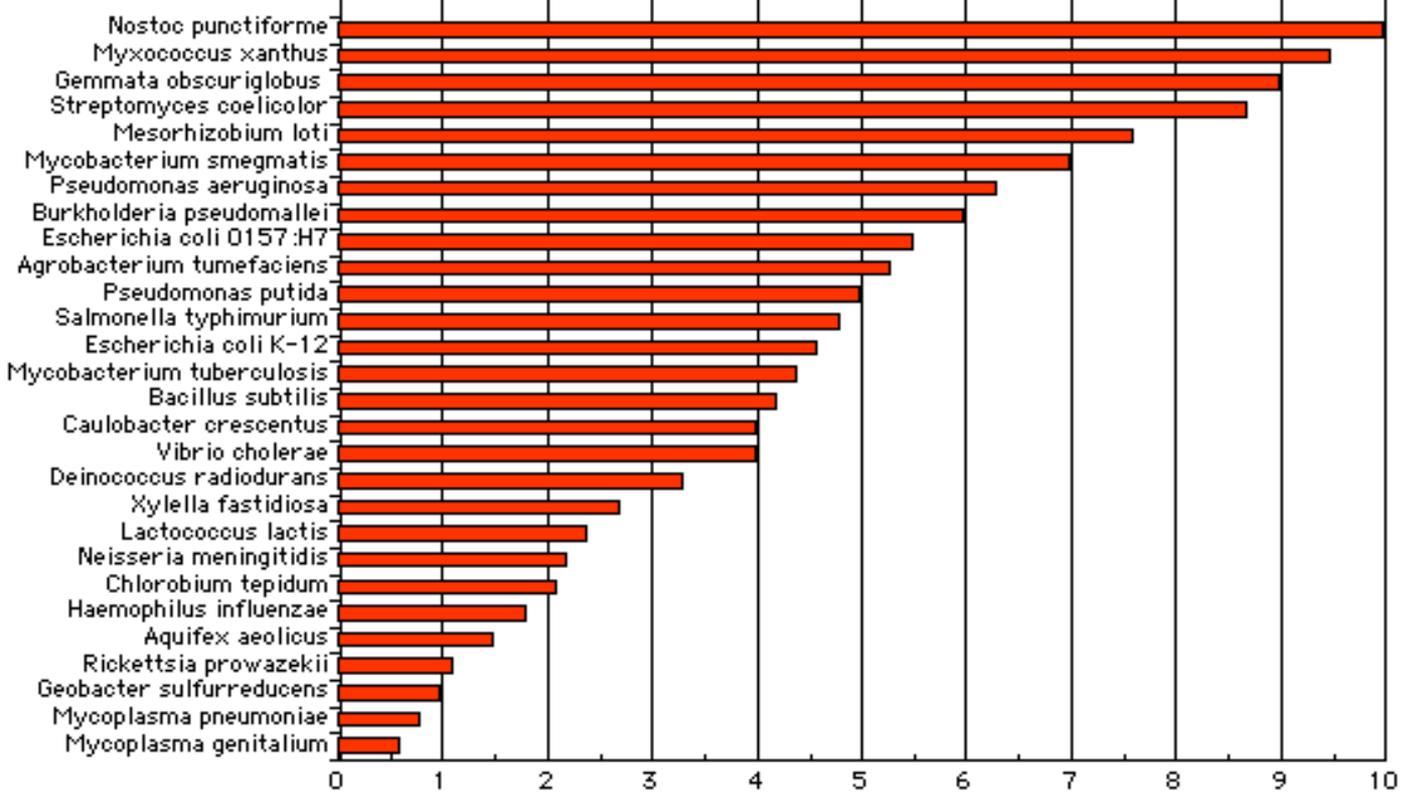
Al di fuori del nucleo è generalmente presente un certo numero di **mitocondri**, dotati di un proprio minigenoma, costituito da una molecola di **DNA circolare** di circa **15.000 – 100.000** paia di basi. I vegetali contengono anche un certo numero di **cloroplasti**, pure dotati di un proprio minigenoma, costituito da una molecola di **DNA circolare** di circa **100.000** paia di basi.

Al di là di questo schema base si riscontrano **casi particolari**. Alcuni unicellulari sono privi di mitocondri (es.: *Giardia lamblia*), alcune alghe unicellulari (es. *Guillardia theta*) presentano due nuclei distinti ciascuno con più cromosomi (il minore è detto **nucleomorfo**). I Ciliati (es. *Tetrahymena*), accanto al nucleo normale, detto germinale o **micronucleo**, sviluppano secondariamente un ulteriore nucleo più grande, detto somatico o **macronucleo**, contenente molte migliaia di minicromosomi a DNA lineare, copie numerosissime di segmenti del DNA dei cromosomi del micronucleo. Le cellule salivari di *Drosophila* replicano il proprio DNA a cascata una decina di volte senza disgiungerlo dando origine a **cromosomi giganti** di oltre 1000 copie, detti **politenuci**. Il batterio gigante *Epulopiscium* (batterio gigante simbiote del pesce chirurgo) contiene un numero molto elevato di copie del proprio genoma (poliploide).

Archaea:



Bacteria:



Genome size (Mbp)

I 3 paradossi del genoma

K

N

C

Tra uomo e scimpanzè



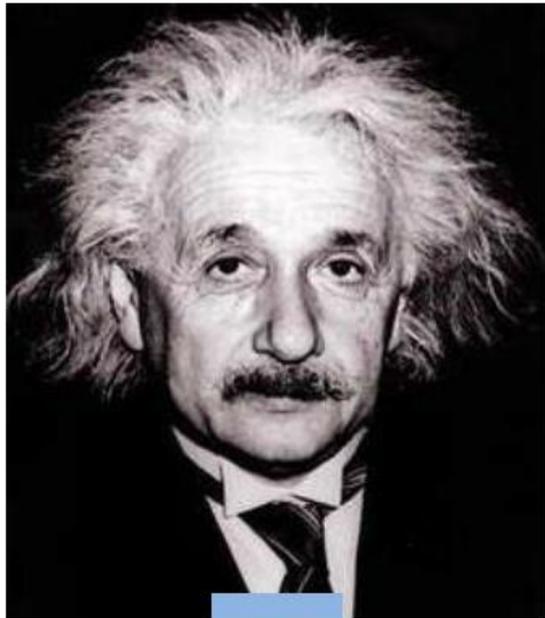
Differenza del 4% nel DNA non codificante

Paradosso del valore **K**: la complessità non correla con il numero di cromosomi.

Homo sapiens

Lysandra atlantica

Ophioglossum reticulatum



46

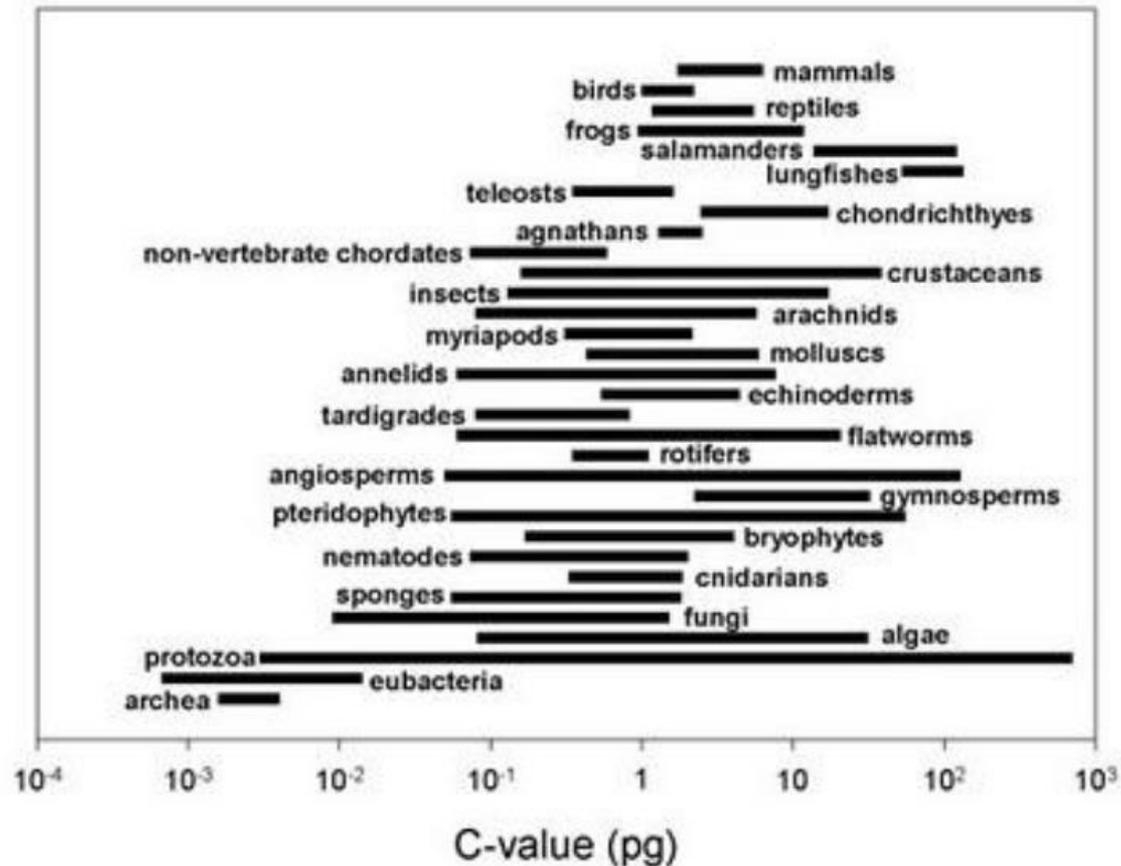


250

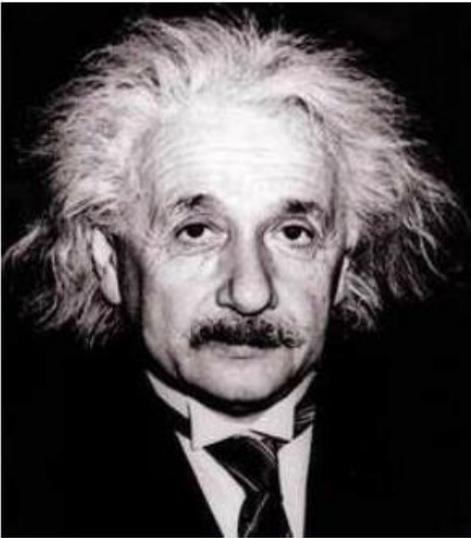


~1260

Il valore C (quantità di DNA contenuto nel nucleo di una cellula aploide di un organismo)



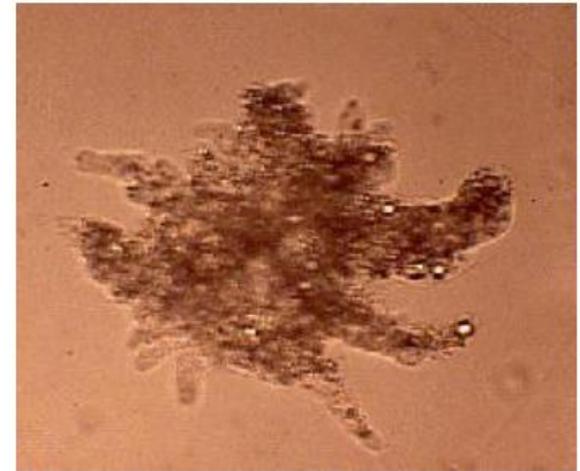
Paradosso del valore **C**: la complessità non correla con la **grandezza del genoma**.



3.4×10^9 bp
Homo sapiens



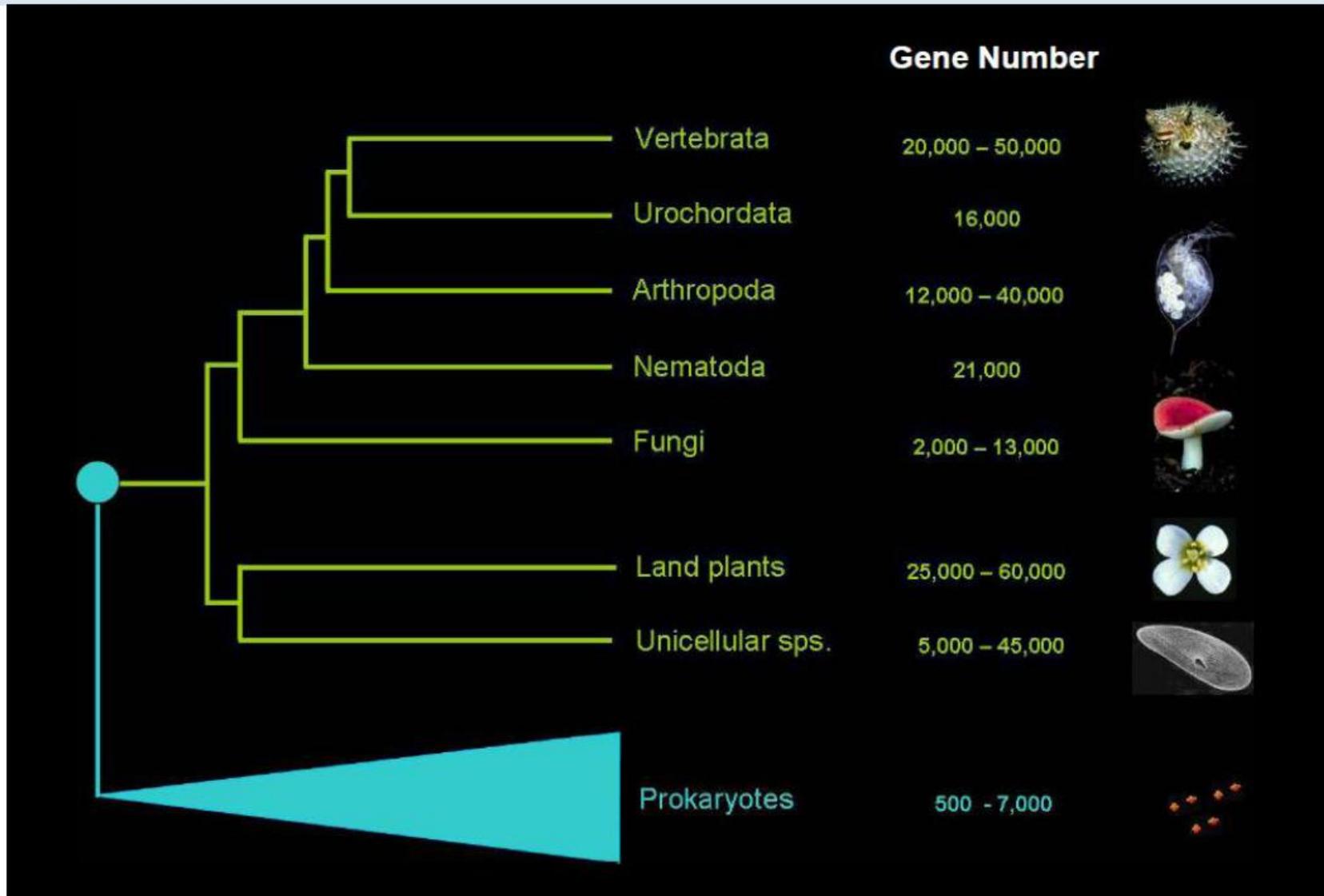
1.5×10^{10} bp
Allium cepa



6.8×10^{11} bp
Amoeba dubia

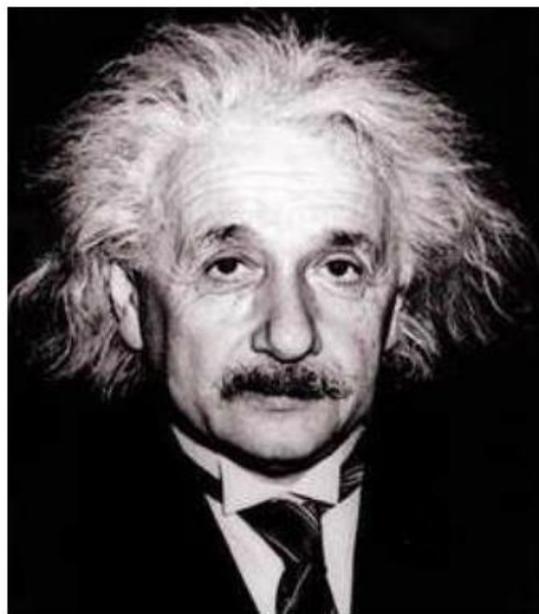
Paradosso del valore N:

Il numero di geni e la complessità degli organismi non sono correlati



Paradosso del valore N:

Il numero di geni e la complessità degli organismi non sono correlati



~21000 geni



~25000 geni



~60000 geni

Il paradosso C dei genomi eucariotici

Nei **procarioti** numero di geni e dimensioni genomiche delle varie specie sono approssimativamente proporzionali, in ragione di circa **1000-1200 pb/gene**.
Appare ragionevole.

Le dimensioni genomiche (corredo aploide) nelle **specie eucariotiche** per contro **variano enormemente** (da circa 10^7 a più di 10^{11}) **senza alcuna relazione con la complessità dell'organismo**. Ad esempio, negli unicellulari alcune amebe hanno le massime dimensioni genomiche mentre alcune alghe unicellulari e alcuni lieviti le minime. Nelle piante superiori si va da circa 10^8 a circa 10^{11} , e lo stesso succede con gli animali.

Questa osservazione è stata chiamata il **Paradosso C** (Complessità).

Il paradosso (in parte) rientra quando si scopre che la variazione nel **numero di geni** (che nella casistica attuale, non vasta ma certamente significativa, vanno da circa **5000** a circa **30000**) risulta molto più contenuta e approssimativamente connessa alla complessità dell'organismo.

Il grosso della differenza nelle dimensioni genomiche è dovuto alla **enorme variazione** nella quantità di **DNA non codificante** per proteine e in buona parte costituito da **sequenze ripetitive**. La ragione di tale variabilità non è ancora chiara.

Dimensioni genomiche di alcuni organismi rappresentativi:

	N° cromosomi (corredo aploide)	M bp tot	N° geni approx	<u>k bp</u> gene
PROCARIOTI				
<i>Nanoarchaeum equitans</i>	1	0.5	420	1,2
<i>Mycoplasma genitalium</i>	1	0.6	480	1,2
<i>Haemophilus influenzae</i>	1	1.8	1750	1,0
<i>Escherichia coli</i>	1	4.7	4300	1,1
<i>Bradyrhizobium japonicum</i>	1	9.1	8300	1,1
EUCARIOTI				
Protisti				
<i>Plasmodium falciparum</i>	14	23	5300	4,3
Funghi				
<i>Saccharomyces cerevisiae</i>	16	12	6000	2,0
<i>Aspergillus nidulans</i>	8	31	..	
Piante				
<i>Ophioglossum petiolatum</i>	510	160000	..	
<i>Arabidopsis thaliana</i>	5	120	25000 ca.	5
<i>Oryza sativa</i>	12	450	30000 ca.	15
<i>Triticum aestivum</i>	7/21	17000	..	
Animali				
<i>Caenorhabditis elegans</i>	6	100	19000	5
<i>Drosophila melanogaster</i>	4	180	13500	13
<i>Takifugu rubripes</i>	22	380	25000	13
<i>Danio rerio</i>	25	1900	25000	60
<i>Protopterus aethiopicus</i>	19	140000	..	
<i>Gallus gallus</i>	39	1200	25000	50
<i>Mus musculus</i>	20	3450	25000	140
<i>Homo sapiens</i>	23	3200	25000*	130

* Secondo una stima di Craig Venter (nel 2007) i geni sarebbero 23.224, mentre secondo Jim Kent (2007) sarebbero 20.433 codificanti e 5.871 non codificanti.

Quanto DNA c'è in una cellula di *Escherichia coli*?

Massa molecolare delle basi:

-Adenina 135 Da

-Citosina 111 Da

-Guanina 151 Da

-Timina 126 Da

Massa molecolare media delle basi 131 Da

Massa molecolare media di un deossinucleoside 247 Da

Massa molecolare media di un deossinucleotide libero 327 Da

Massa molecolare media di un deossinucleotide in catena 309 Da

Massa molecolare media di un paio di deossinucleotidi in catena 618 Da

Il cromosoma di *E. coli* contiene circa 4.7×10^6 paia di basi, quindi ha una massa di circa 2.9×10^9 Da, corrispondenti a 4.8×10^{-15} g
(1Da= 1.66×10^{-24} g)

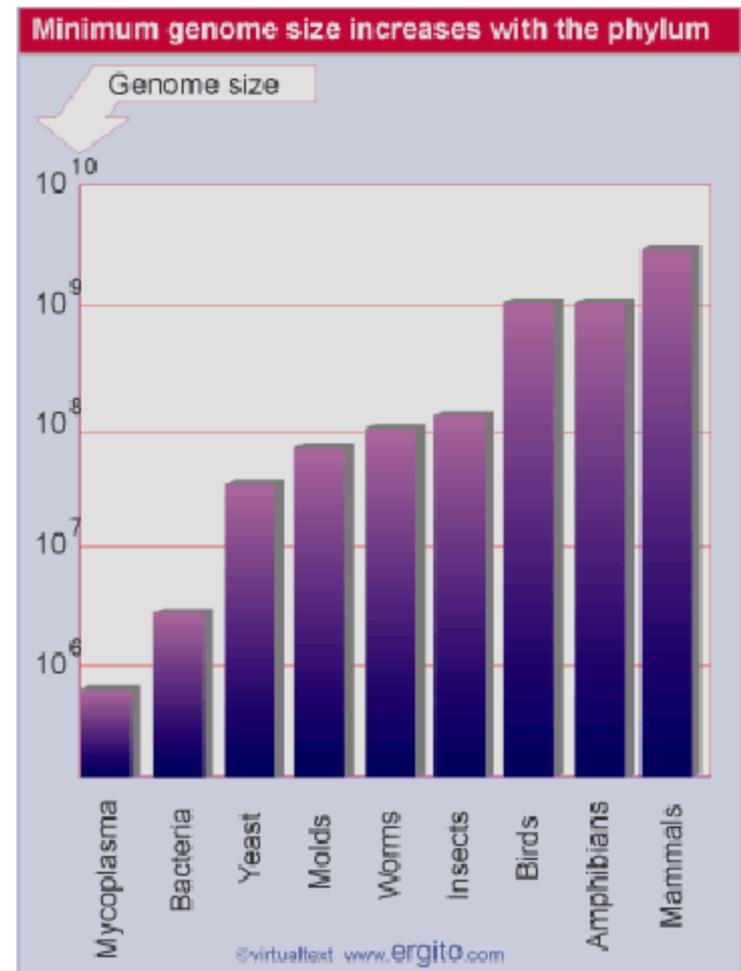
E in una cellula umana?

Un corredo aploide contiene circa 3.3×10^9 paia di basi, una cellula diploide il doppio, quindi una massa di circa 2.1×10^{12} Da, corrispondenti a 3.5×10^{-12} g, a cui andrebbe aggiunta una piccola quantità di DNA mitocondriale.

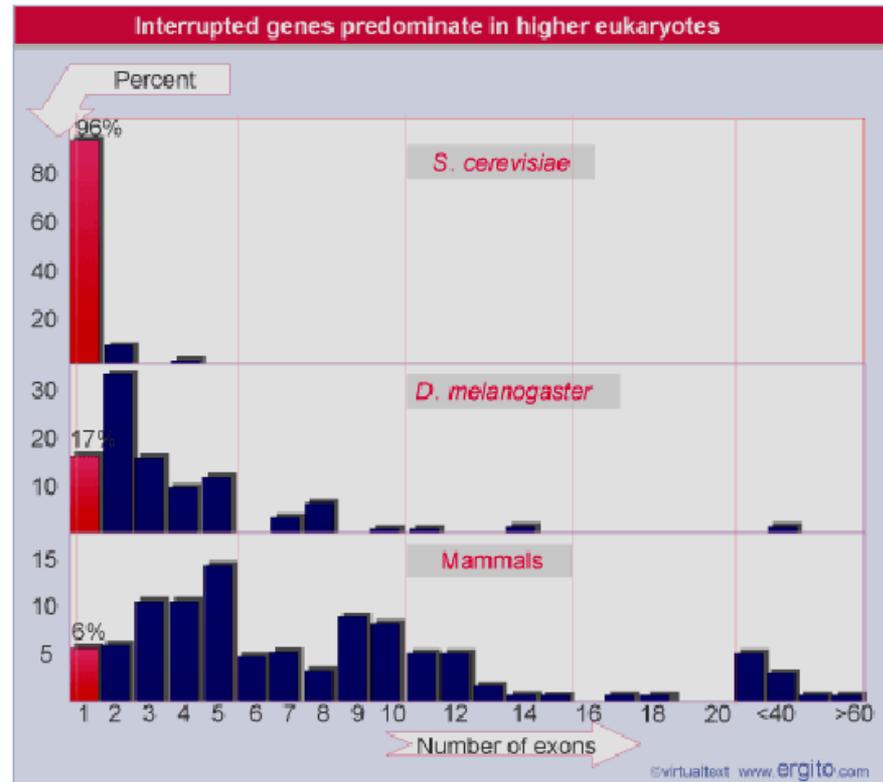
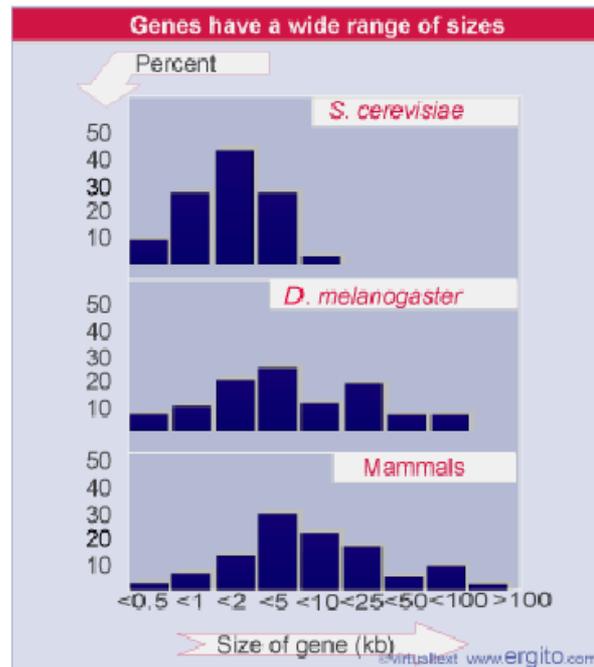
Essendo il numero stimato di cellule in un adulto dell'ordine di 10^{14} , per la gran parte diploidi, ne consegue un contenuto complessivo di DNA dell'ordine di qualche ettogrammo .

Quantità minima di DNA

- Il grafico indica che un aumento della grandezza del genoma è necessaria per la complessità dei procarioti ed eucarioti inferiori

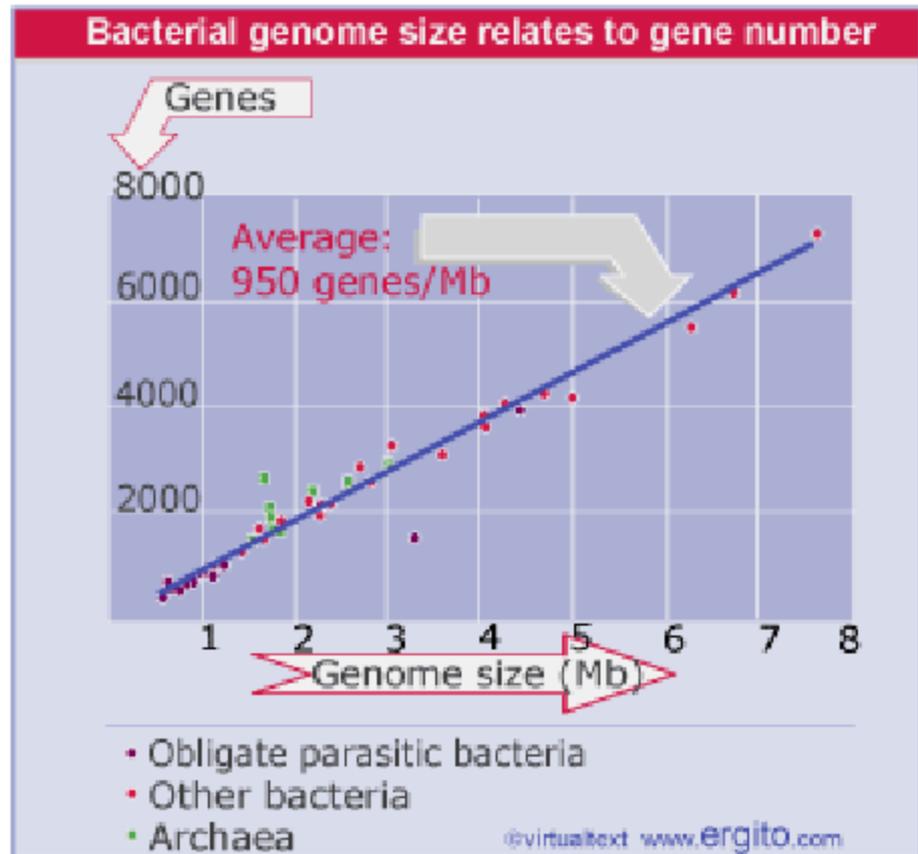


La dimensione dei geni varia grandemente



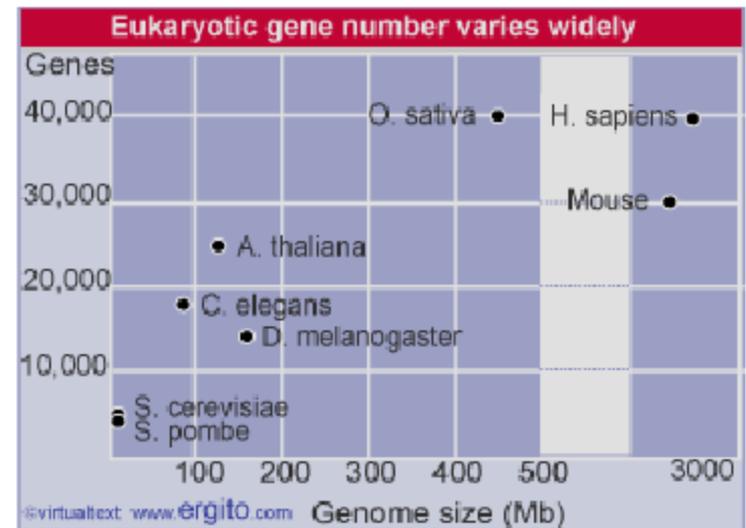
Geni batterici

- Nei batteri quasi tutto il DNA codifica proteine o RNA
- Il genoma varia entro un ordine di grandezza ed è proporzionale al numero di geni



Geni negli eucarioti

- Il genoma degli **eucarioti unicellulari** è simile a quello dei **procarioti**
- Gli eucarioti superiori hanno più geni, ma il loro numero non correla con la dimensione del genoma



Tre classi di geni

- I geni non sono sequenze casuali ma hanno caratteristiche ben precise.
- Buona parte dell'informazione contenuta in un gene viene "copiata" in una molecola di RNA; il resto del gene è coinvolto comunque nel processo di "copia" (trascrizione).
- Alcuni tipi di RNA vengono utilizzati per la sintesi delle proteine, altri svolgono svariati tipi di funzioni.
- Esistono tre classi di geni, che differiscono in base al tipo di RNA che viene prodotto con la loro espressione:
 - Geni della I classe
 - RNA ribosomiale (rRNA)
 - Geni della II classe
 - RNA messaggero (mRNA)
 - Piccoli RNA nucleari (snRNA)
 - Micro RNA (miRNA)
 - Geni della III classe
 - RNA transfer (tRNA)
 - Piccoli RNA nucleolari (snoRNA)
 - Piccoli RNA citoplasmatici (scRNA)
 - Micro RNA (miRNA)

Gli Pseudogeni

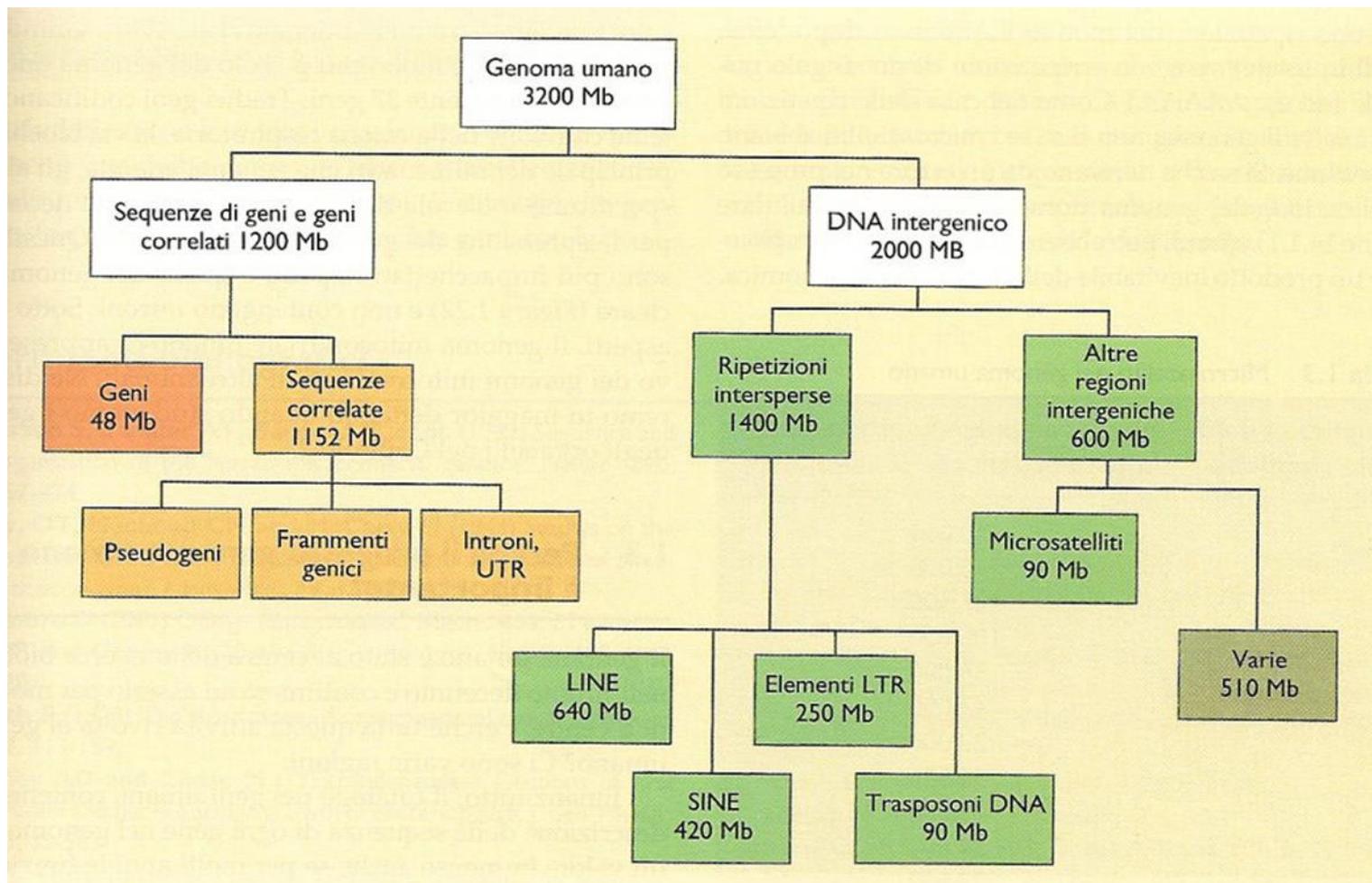
- Gli pseudogeni sono copie non funzionali di geni.
- Sono una sorta di relitti evolutivi.
- Gli pseudogeni convenzionali sono geni inattivati in seguito ad una o più mutazioni nella loro sequenza nucleotidica.
- Una volta che uno pseudogene è diventato completamente non funzionale si degraderà per accumulazione di ulteriori mutazioni e potrebbe addirittura non essere più riconosciuto come relitto genico.

Pseudogeni

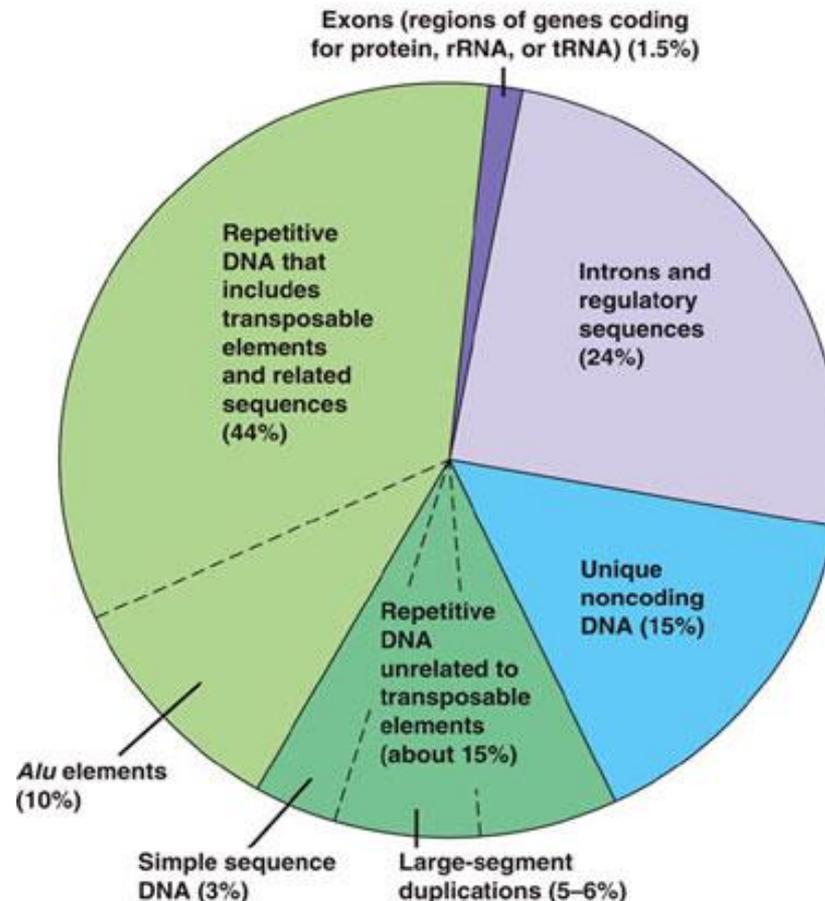
- Sono copie difettose di geni funzionali, prima ritenute un residuo evolutivo
- Contribuiscono alla formazione di qualche RNA attivo
- Si è notato che la loro compromissione determina la morte dell'organismo soggetto

Ripetizioni disperse e microsatelliti

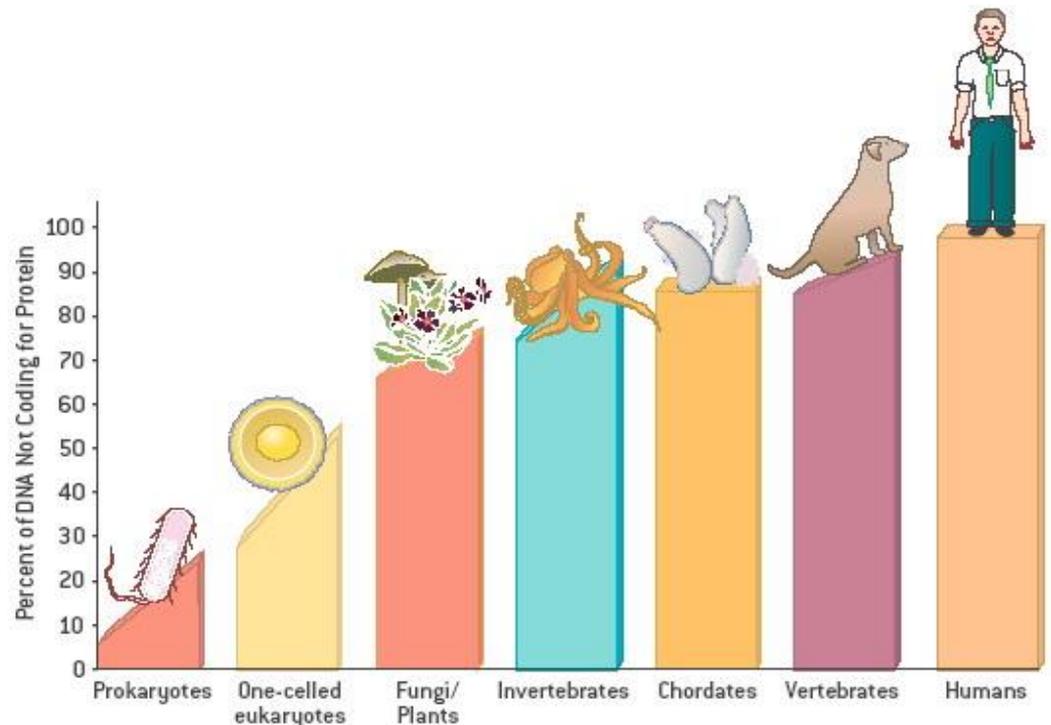
- La grande maggioranza del DNA intergenico è rappresentata da sequenze ripetute di vario tipo.
- Il **DNA ripetitivo** può essere diviso in due categorie:
 - Ripetizioni intersperse
 - DNA ripetuto in tandem



DNA codificante e non codificante nel genoma umano



Problema della complessità

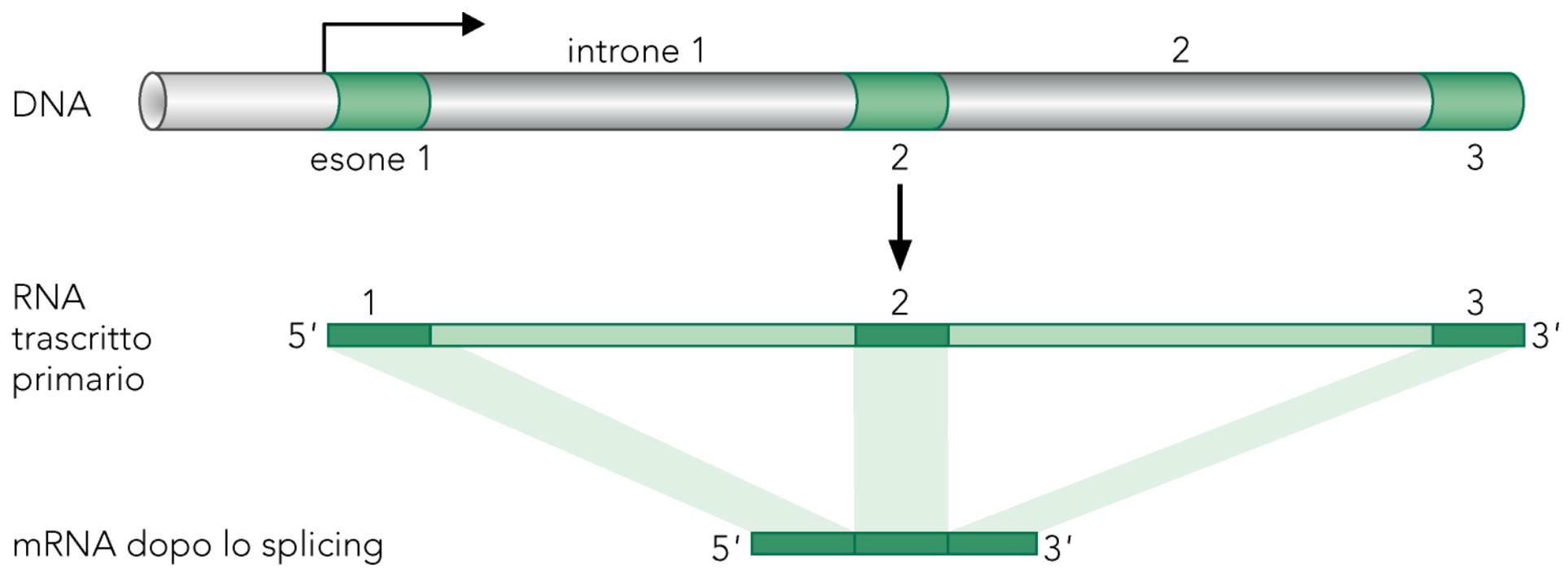


NONPROTEIN-CODING SEQUENCES make up only a small fraction of the DNA of prokaryotes. Among eukaryotes, as their complexity increases, generally so, too, does the proportion of their DNA that does not code for protein. The noncoding sequences have been considered junk, but perhaps it actually helps to explain organisms' complexity.

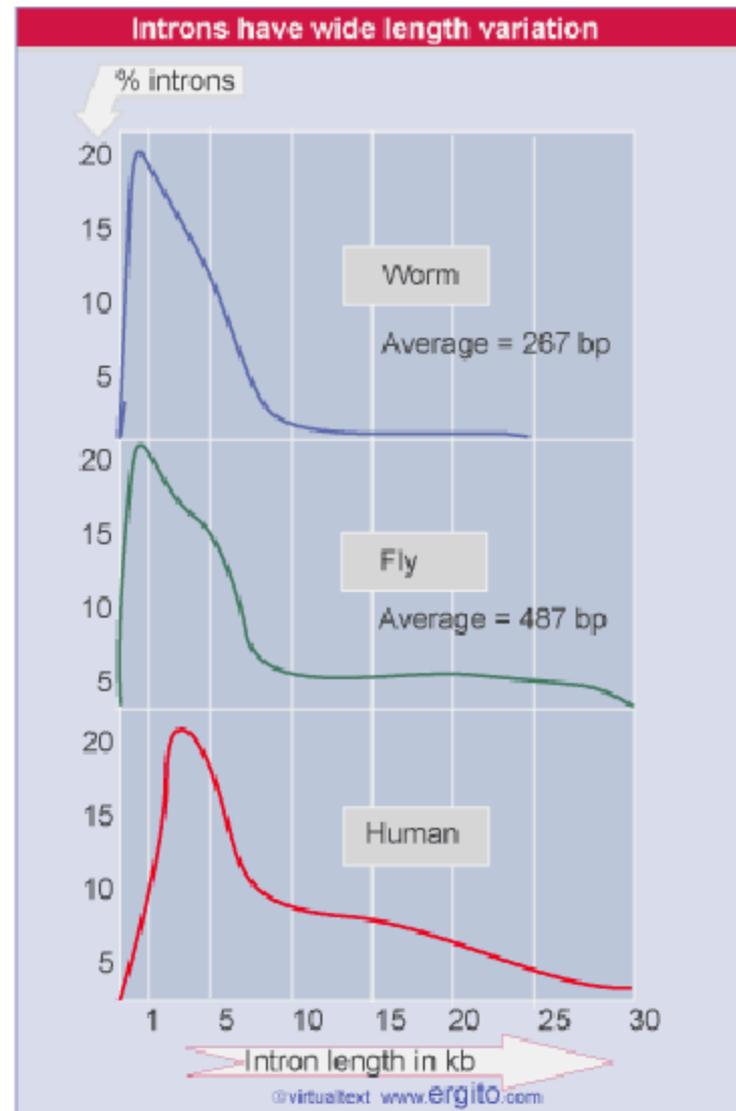
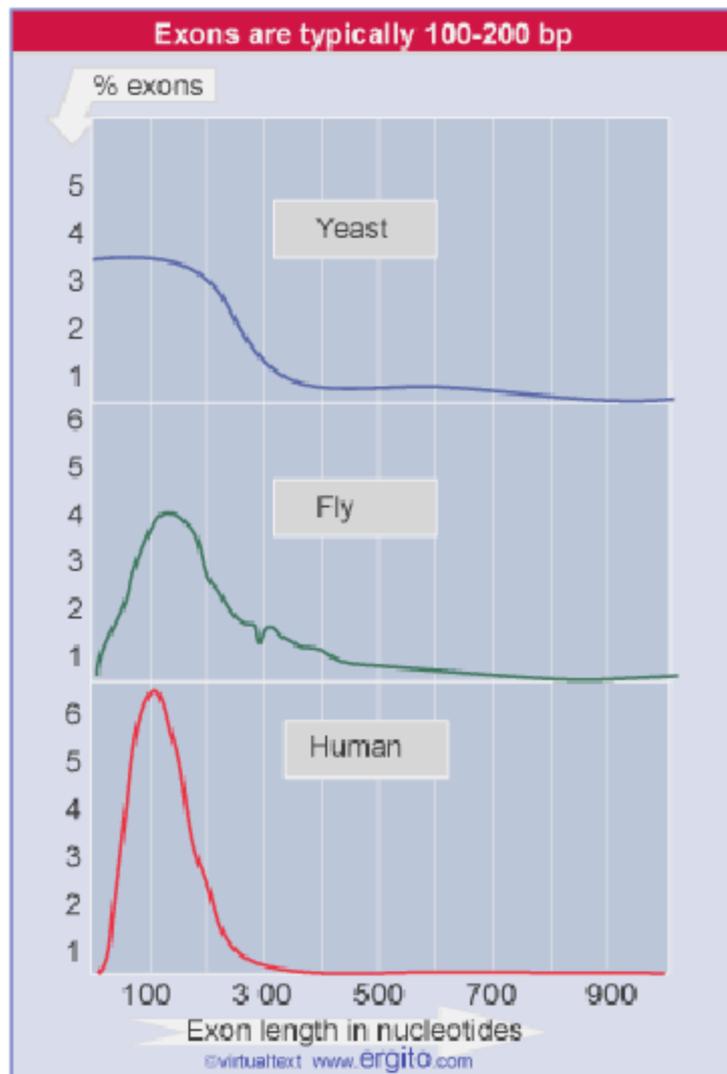
La complessità degli organismi è correlata alle dimensioni del genoma non codificante

Struttura dei Geni eucariotici codificanti

- I geni codificanti sono quelli che vengono trascritti in mRNA.
- Contengono una parte realmente codificante, che specifica la sequenza degli aminoacidi che costituiranno la proteina, ed una parte non codificante.
- A monte della sequenza che verrà trascritta in mRNA vi sono le **sequenze regolatrici**.
- La sequenza trascritta è costituita da due tipi di elementi, detti **esoni** ed **introni**.
- Solo gli esoni contengono informazioni per la sintesi della proteina.



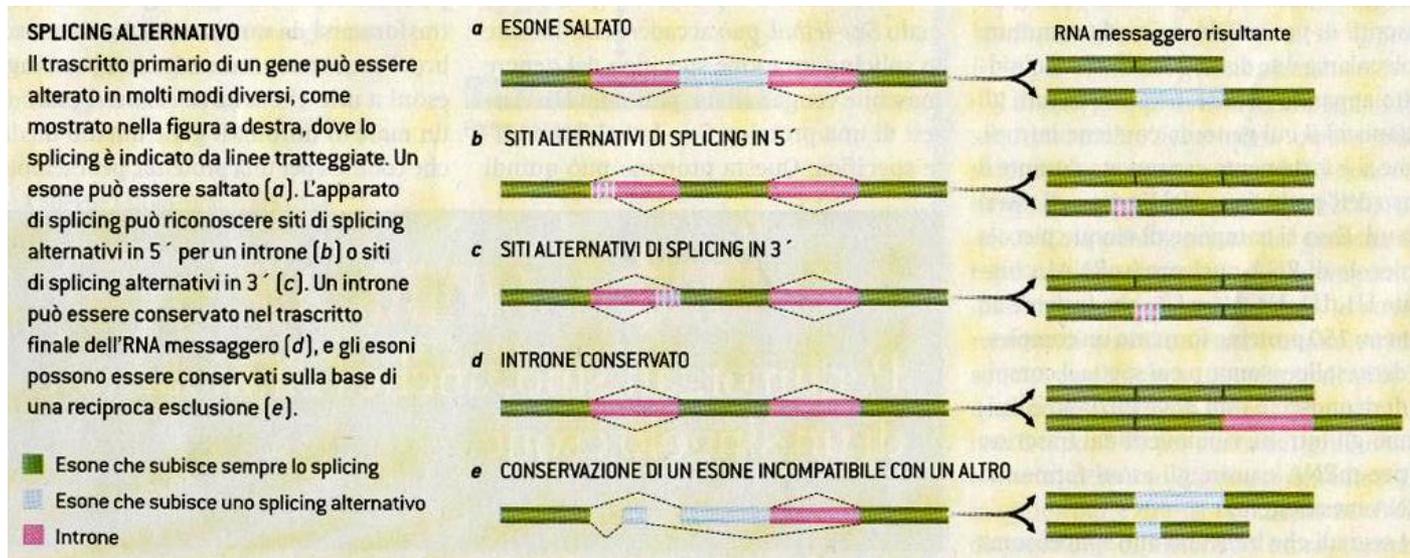
Dimensione esoni e introni



- La lunghezza media degli esoni non varia, quella degli introni aumenta con la filogenesi

Utilità degli introni

- Attraverso lo **splicing alternativo**, nel quale sequenze introniche sono conservate o eliminate, aumenta considerevolmente il numero di proteine codificabili per ogni gene



Utilità degli introni

- Le famiglie Alu, brevi sequenze trasposoniche di circa 300 basi terminanti con la caratteristica coda poli-A, sembrano avere giocato un importante ruolo per l'evoluzione proteica
- Inserendosi casualmente all'interno degli introni e subendo una mutazione (anche solo di una base), possono creare un sito di splicing alternativo, rendendo un pezzo di introne un nuovo esone.

Utilità degli introni

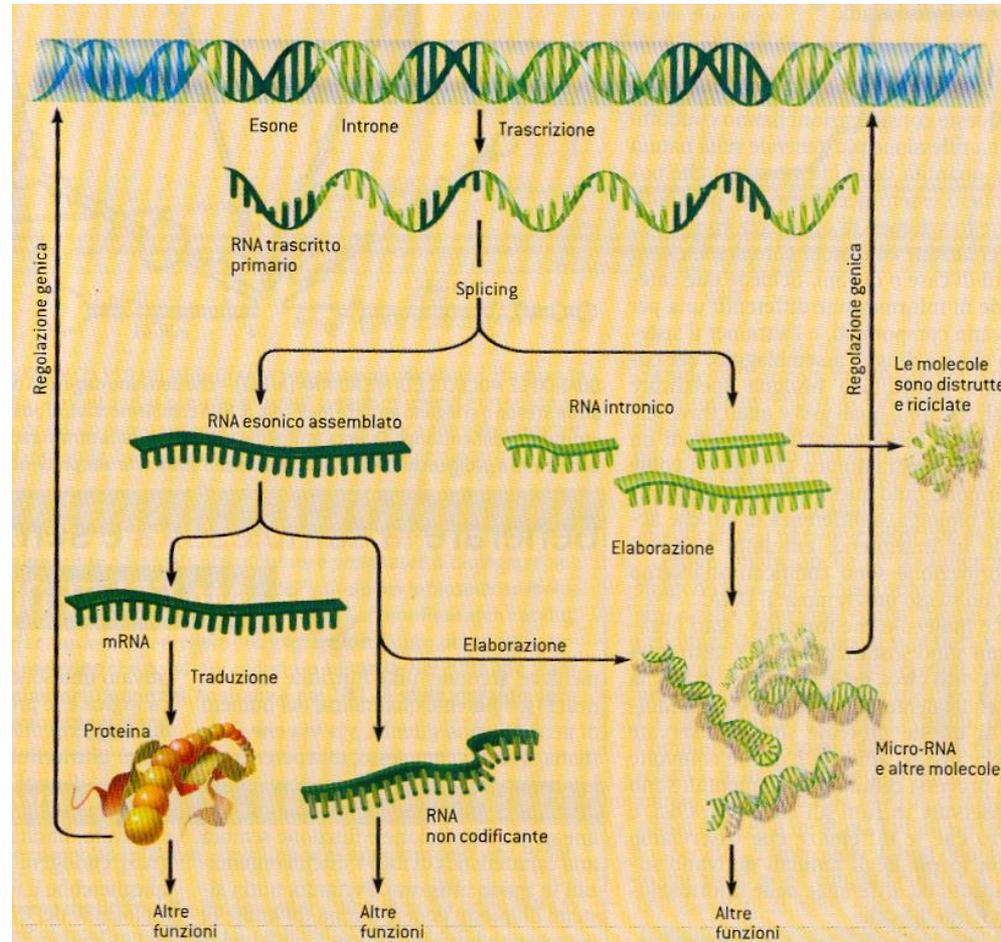
- Questa mutazione non può che portare vantaggi per l'organismo ospitante: la vecchia proteina è sempre conservata ma c'è la possibilità di crearne una nuova, che, se utile, può essere conservata.

Introni = Laboratori di sperimentazione per nuove proteine

Utilità degli introni

- Recenti studi hanno mostrato che gli introni non vengono immediatamente riciclati dopo essere stati rimossi dall'RNA
- Importantissima funzione di regolazione tramite RNA attivi (miRNA e riboswitch)

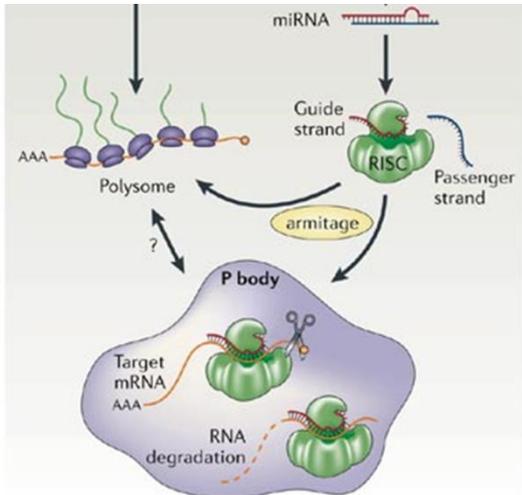
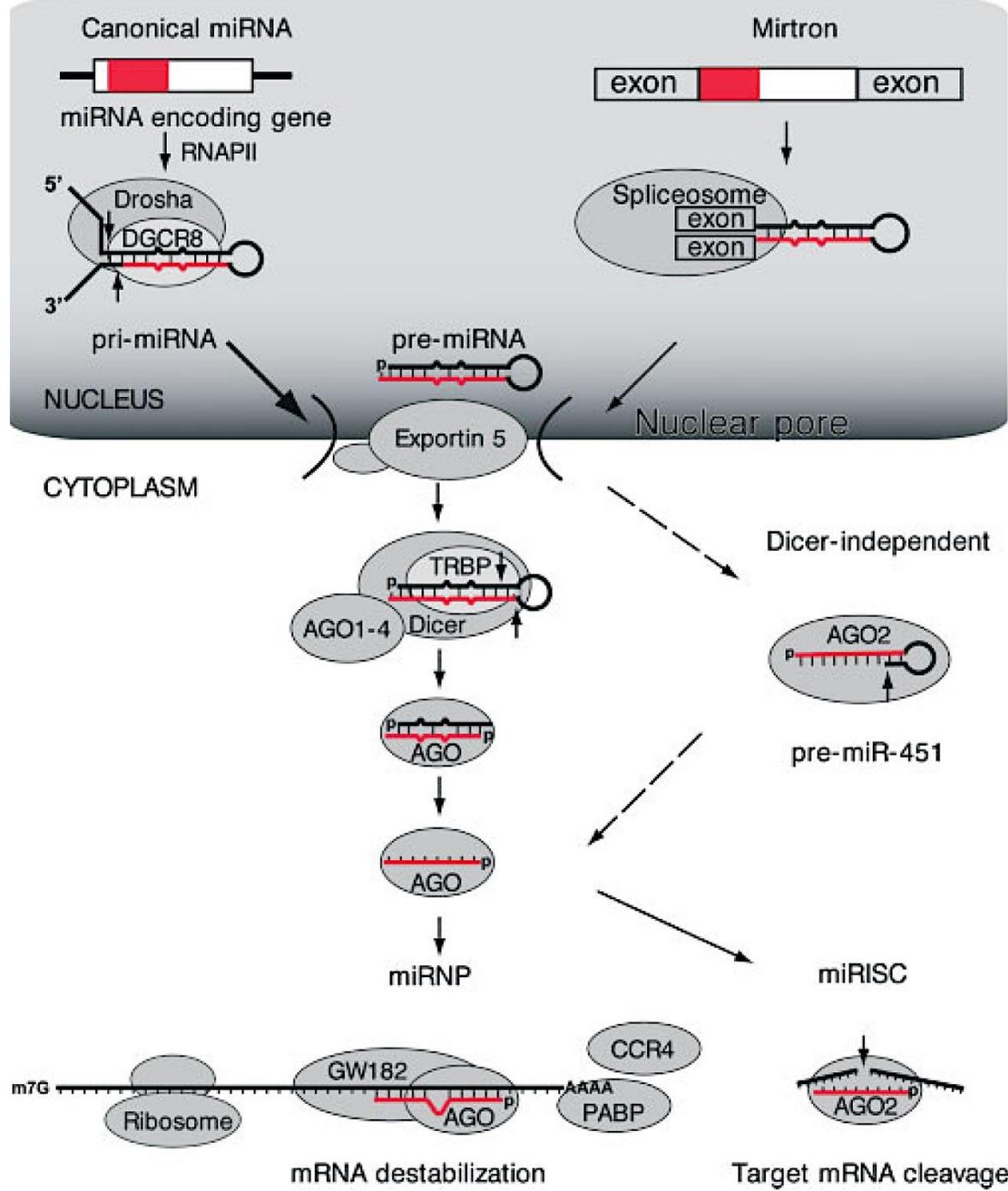
Recente teoria sull'attività genica negli eucarioti



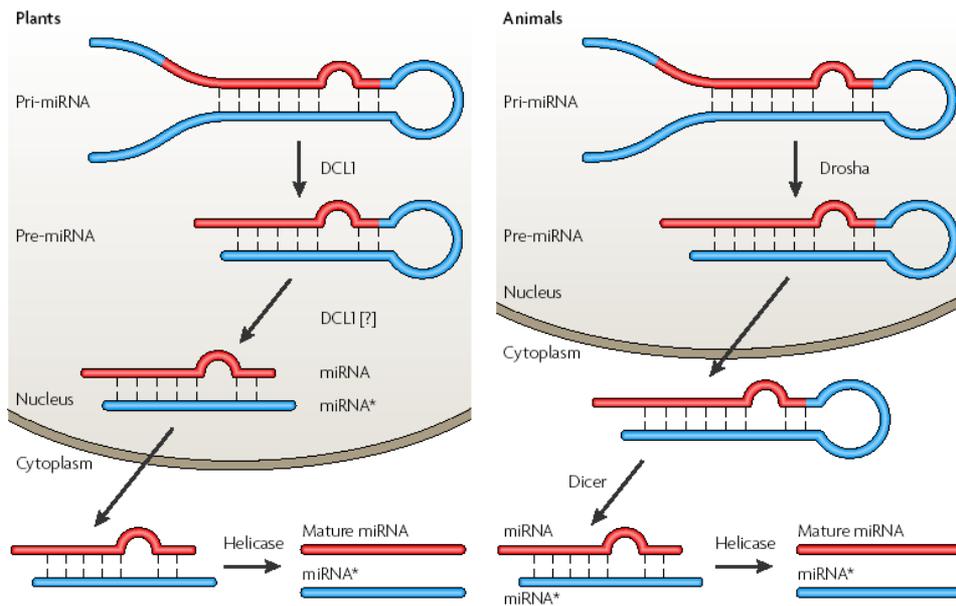
microRNA

- RNA non codificante di 21-22 nt
- Il trascritto è un pre-miRNA di circa 70 nt che forma una struttura stem-loop
- Il pre-miRNA è processato in una molecola di 21-22 nt dall'enzima Dicer.
- La maggior parte dei miRNA regola l'espressione genica dei loro bersagli mRNA.
- La complementarità perfetta con i loro bersagli porta a degradazione dell'mRNA.
- La complementarità imperfetta con mRNA bersaglio porta all'inibizione della traduzione.
- **PRESENTI IN TUTTI GLI EUCARIOTI**

Biogenesi dei miRNA



Generation of miRNAs in Plants and Animals



In plants, miRNA maturation occurs in the nucleus

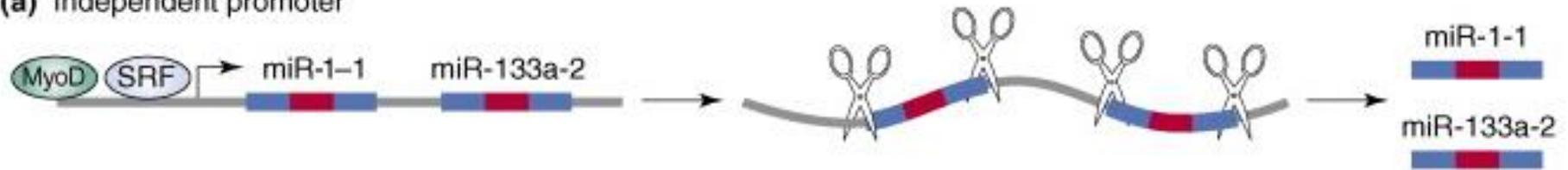
In animals, pre-miRNA is formed in the nucleus and mature miRNA occurs in the cytoplasm

miRNAs regulate ~50% of the human transcriptome

from Chen and Rajewsky, *Nature Rev.Genet.* 8, 93 (2007)

Genomic Organization of miRNA Genes

(a) Independent promoter



(b) Intronic



(c) Exonic



- Intronic miRNAs often in antisense direction, made from own promoter
- Exonic miRNAs - non-coding (or in alternatively spliced exons)

Classification of MiRNAs

- MicroRNAs are classified in two group depending on their origin-
 - ❑ **Intergenic or Exonic miRNAs:** located between the introns of genes & transcribed by RNA pol II or pol III as a stem loop structure called pri-miRNA
 - ❑ **Interagenic or Intronic miRNAs:** miRNAs located within an intron of a protein coding gene & transcribed by RNA pol II as part of pre-mRNA

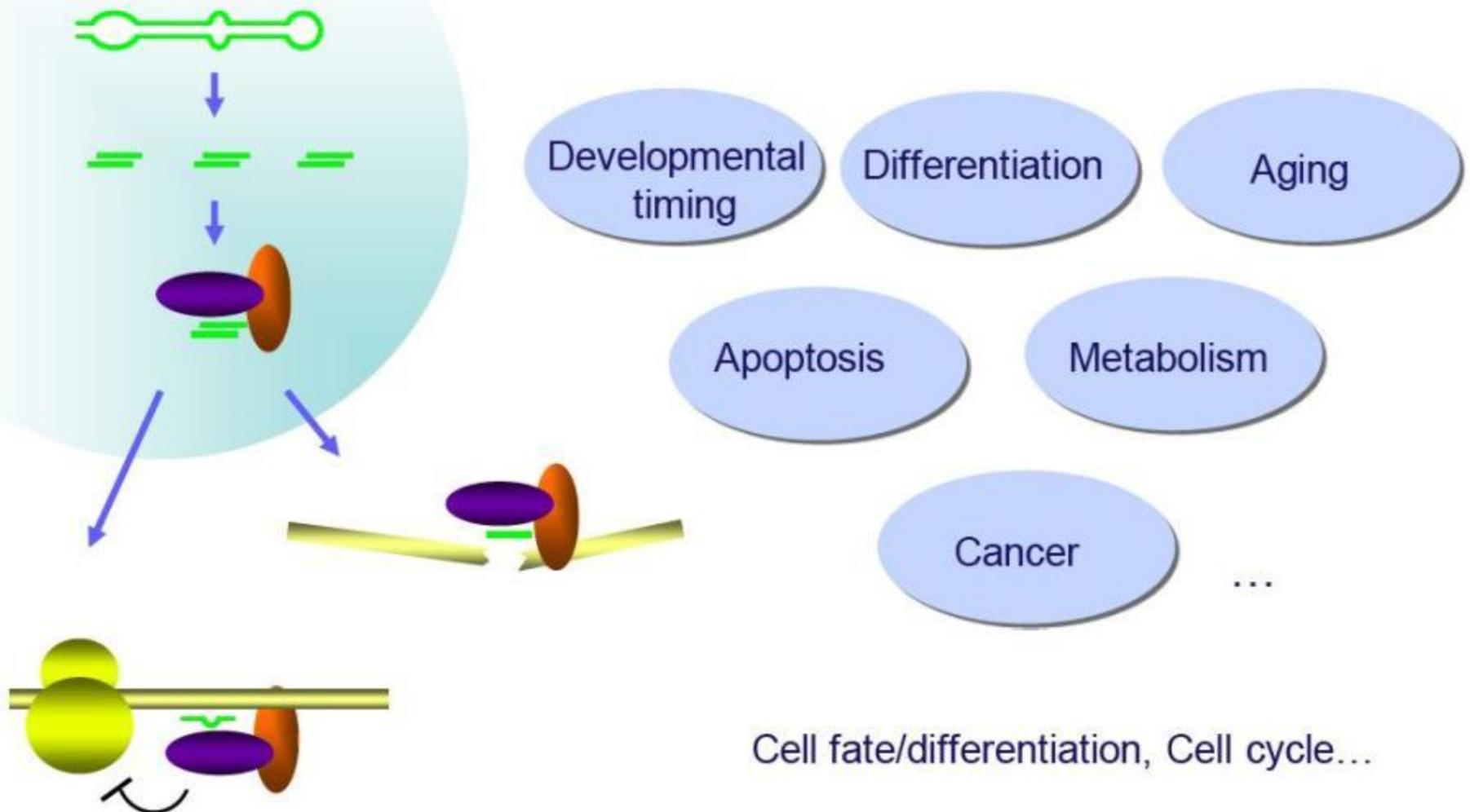
Human miRNAs (February 22, 2010)

- Total miRNA genes in 115 species 10,882
- Total number of miRNAs known 1,580
- Number human miRNAs identified 851
(actualmente 1900)
- Number of human mRNA targets 34,788
- miRNAs can have multiple targets
- Target mRNAs can have multiple miRNA binding sites

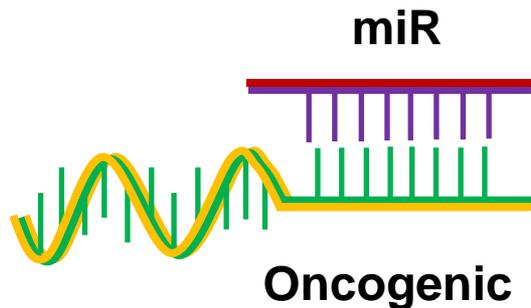
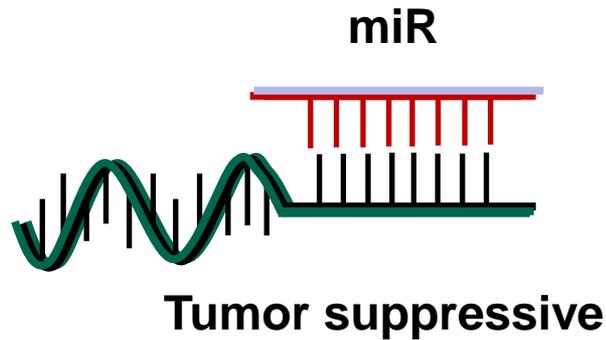
miRBase @ <http://www.mirbase.org/>

MicroCosm @ <http://www.ebi.ac.uk/enright-srv/microcosm/>

Thousands of microRNAs act in multiple biological events

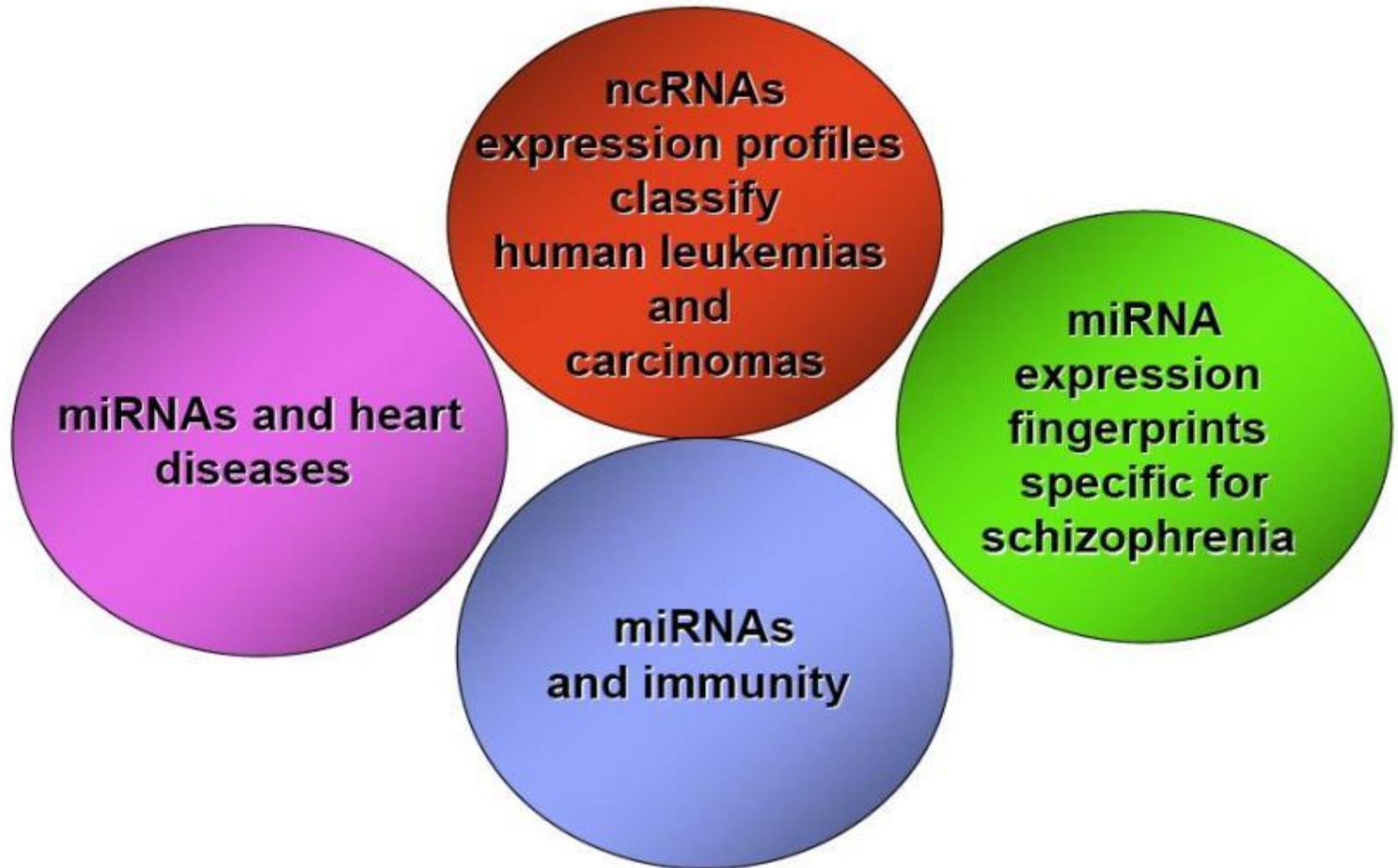


MicroRNA ACTIVITY IN CANCER: TUMOR SUPPRESSIVE OR ONCOGENIC



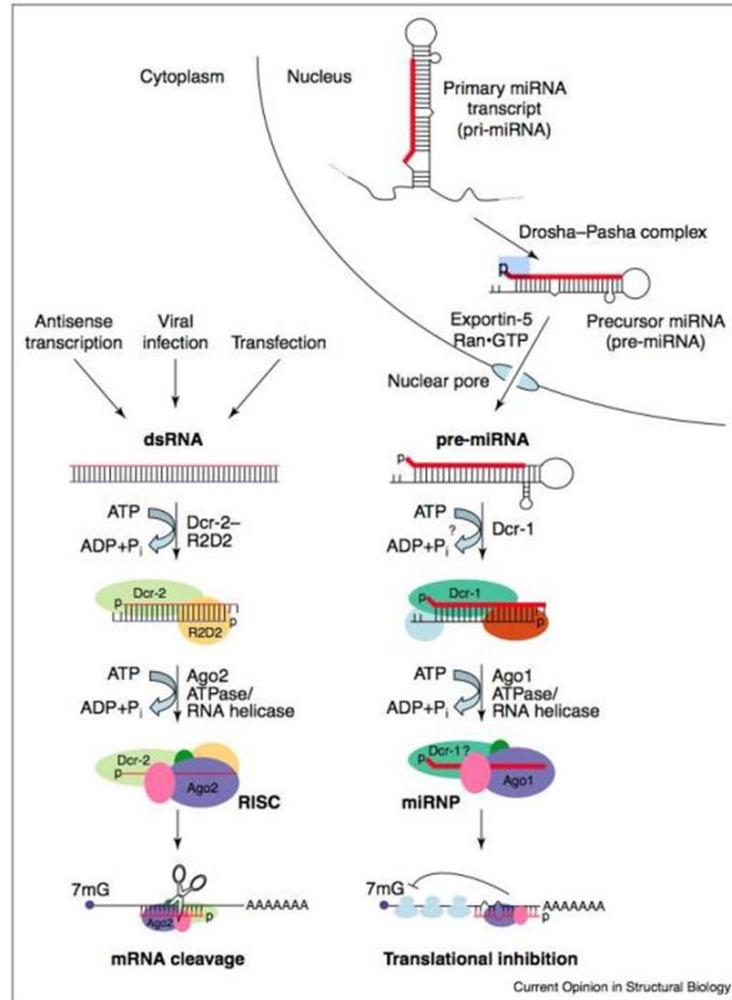
- Suppress expression of oncogenes, growth promoting, survival and angiogenic genes (low in tumors)
- Suppress expression of tumor suppressor, growth inhibitory, proapoptotic genes (high in tumors)

ALTERATIONS OF NONCODING RNAS ARE FOUND IN EVERY TYPE OF HUMAN DISEASE



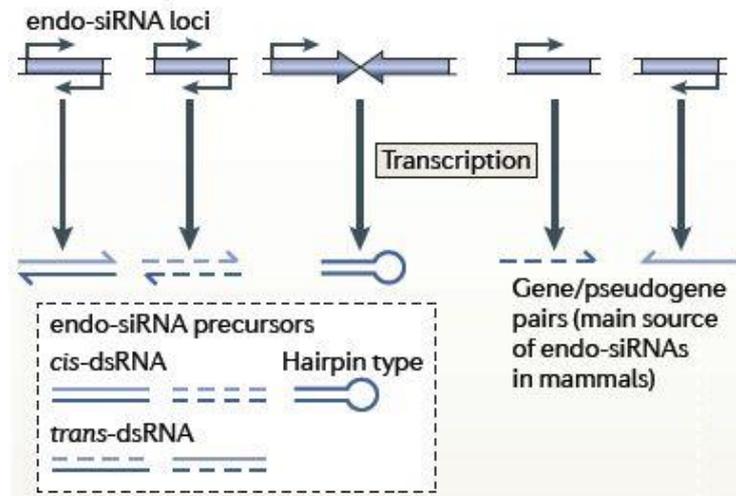
(Calin et al, PNAS 2002; Lu et al, Nature, 2005; Landgraf et al, Cell 2007; Perkins et al Genome Biol 2007; Hansen et al PLoS ONE, 2007; Beveridge et al, Hum Molec Genet 2008, Baltimore D, Nat Immunol 2008; van Rooij, Trends Genet, 2008)

miRNA vs siRNA

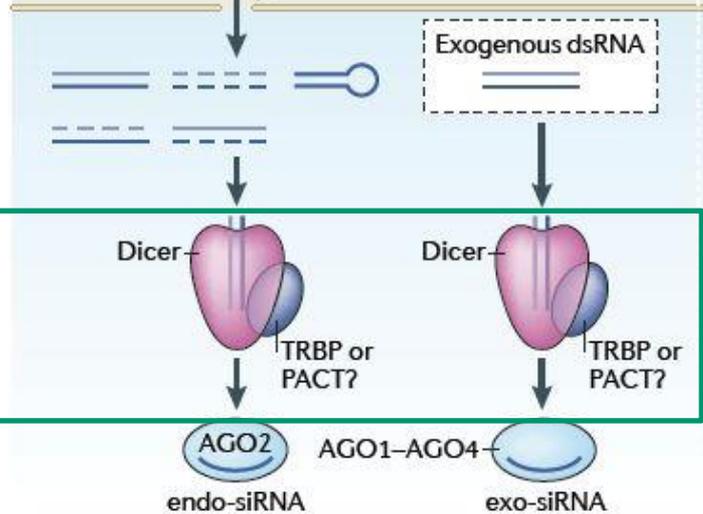


Biogenesis of small RNAs

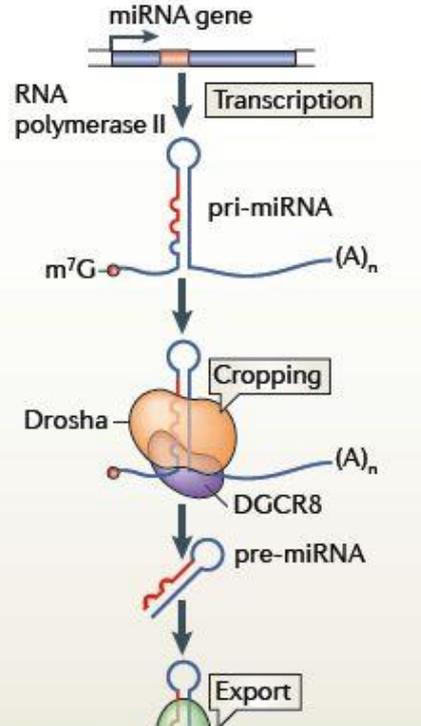
a siRNA biogenesis



Export

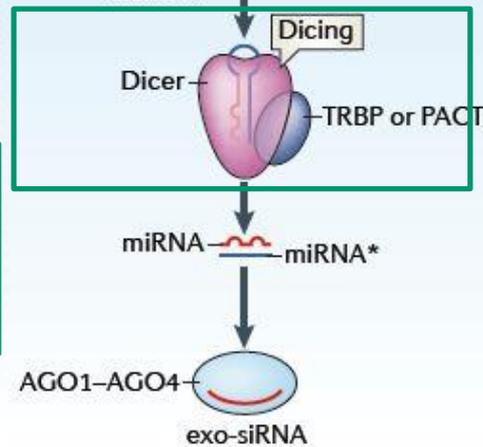


b miRNA biogenesis

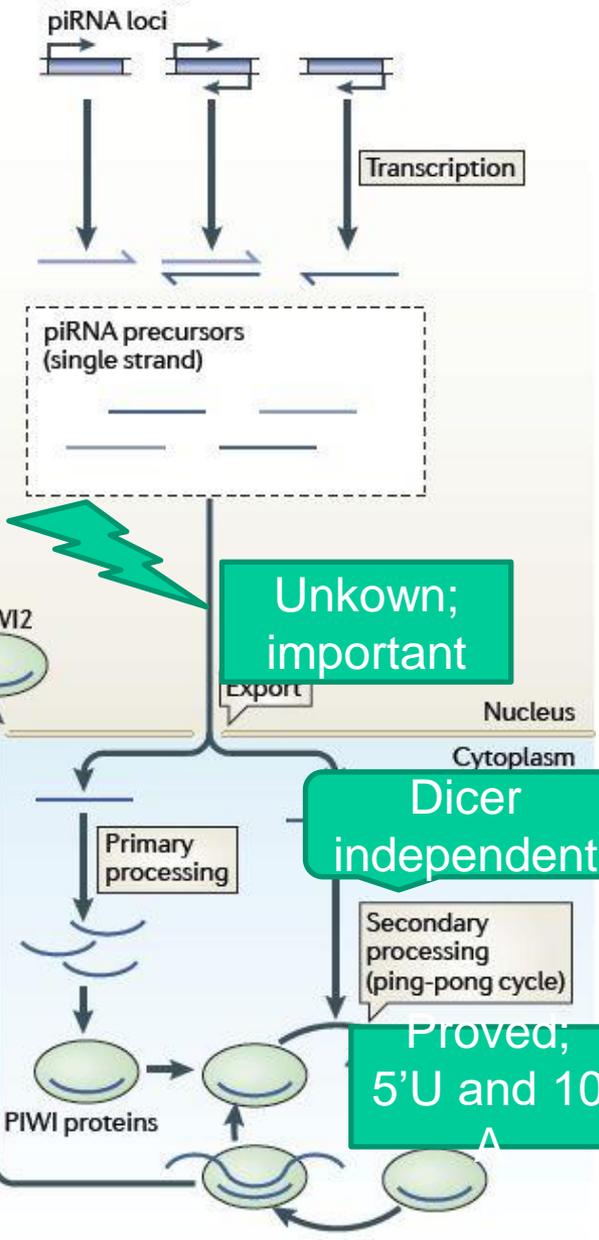


Exportin 5

RanGTP

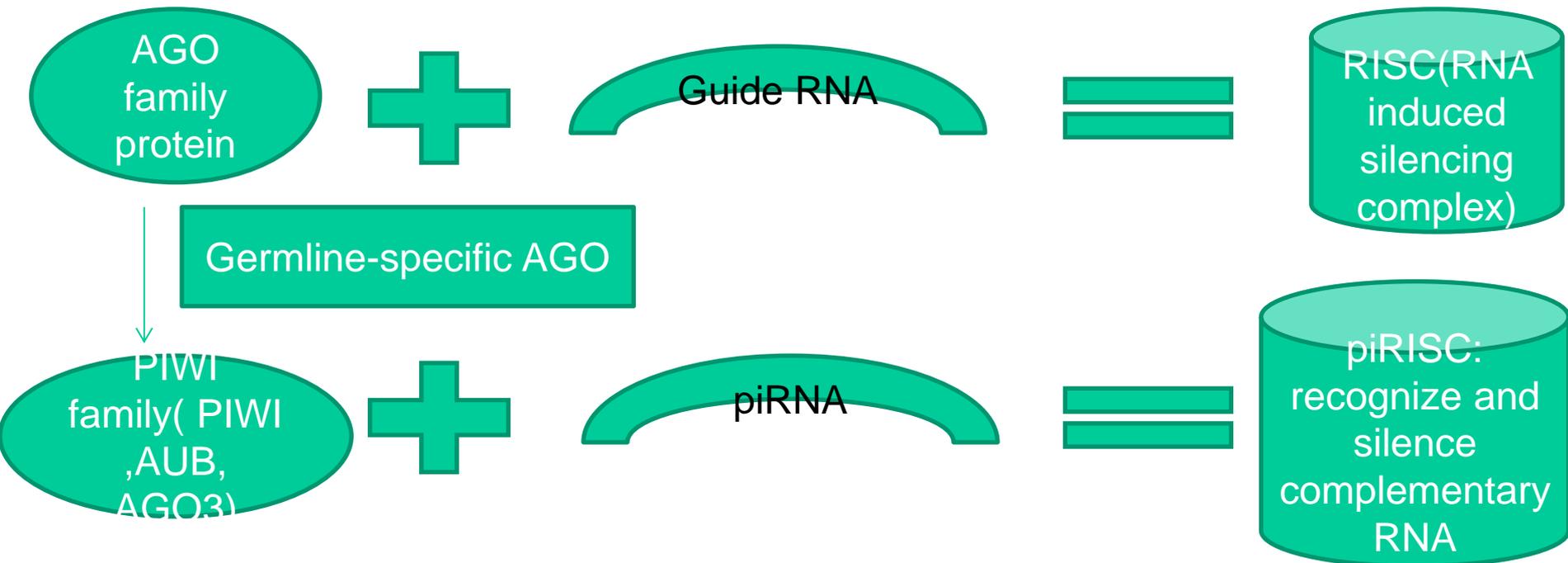


c piRNA biogenesis



What is piRNA

- Small RNAs: piRNA, siRNA, miRNA.
- piRNA: derive from repetitive genomic element, interact with PIWI

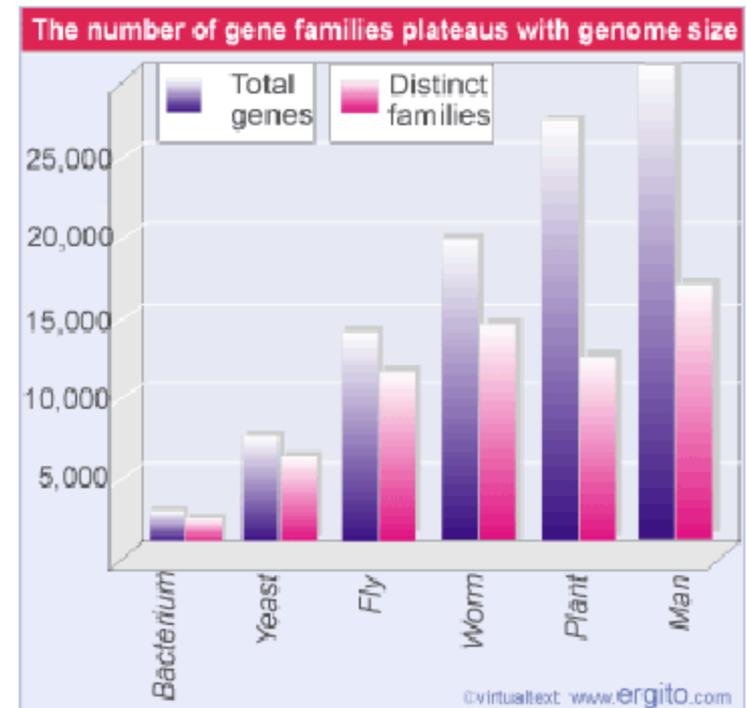


Current Evolutionary studies demonstrate

- piRNA strongly repress retrotransposons
- piRNA evolves very fast among species
- piRNA loci locates in low recombination region
- piRNA show a signature of selective constraint in African populations

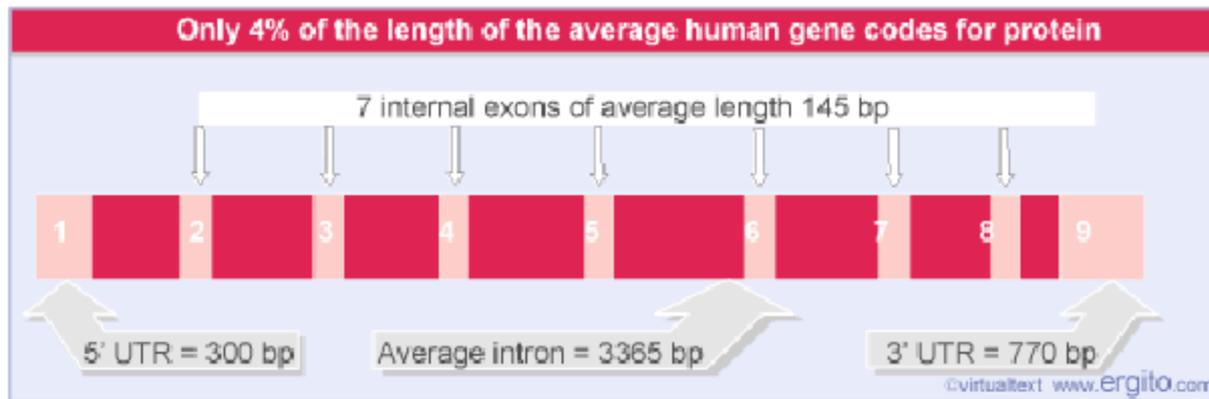
Famiglie di geni

- Alcuni geni sono unici, altri appartengono a famiglie
- Il n° di geni aumenta con la complessità, ma negli eucarioti superiori il n° di famiglie rimane costante



Il gene umano medio

- In media un gene umano ha 9 esoni e 7 introni ed è lungo 27 kb. Gli esoni terminali sono più lunghi. Solo 5% è codificante



I geni occupano il 25% del genoma umano ma solo 1% sono esoni!

Compattezza di alcuni genomi eucariotici

Proprietà del genoma	<i>S.cerevisiae</i>	<i>D.melanogaster</i>	<i>H. sapiens</i>
Densità genica (numero medio di geni per Mb)	479	79	9
Introni per gene (media)	0,04	3	7
% del genoma occupata dalle ripetizioni intersperse	3,4%	12%	44%

Densità genica: quantità di geni contenuti in una megabase di dna genomico

Procarioti: 850-1000geni /Mb

Table 1–2 The Numbers of Gene Families, Classified by Function, That Are Common to All Three Domains of the Living World

GENE FAMILY FUNCTION	NUMBER OF “UNIVERSAL” FAMILIES
Information processing	
Translation	63
Transcription	7
Replication, recombination, and repair	13
Cellular processes and signaling	
Cell cycle control, mitosis, and meiosis	2
Defense mechanisms	3
Signal transduction mechanisms	1
Cell wall/membrane biogenesis	2
Intracellular trafficking and secretion	4
Post-translational modification, protein turnover, chaperones	8
Metabolism	
Energy production and conversion	19
Carbohydrate transport and metabolism	16
Amino acid transport and metabolism	43
Nucleotide transport and metabolism	15
Coenzyme transport and metabolism	22
Lipid transport and metabolism	9
Inorganic ion transport and metabolism	8
Secondary metabolite biosynthesis, transport, and catabolism	5
Poorly characterized	
General biochemical function predicted; specific biological role unknown	24

For the purpose of this analysis, gene families are defined as “universal” if they are represented in the genomes of at least two diverse archaea (*Archaeoglobus fulgidus* and *Aeropyrum pernix*), two evolutionarily distant bacteria (*Escherichia coli* and *Bacillus subtilis*) and one eucaryote (yeast, *Saccharomyces cerevisiae*). (Data from R.L. Tatusov, E.V. Koonin and D.J. Lipman, *Science* 278:631–637, 1997; R.L. Tatusov et al., *BMC Bioinformatics* 4:41, 2003; and the COGs database at the US National Library of Medicine.)

