

STRUTTURA E ORGANIZZAZIONE DEL MATERIALE EREDITARIO

- 1928: Griffith scopre il "principio trasformante"

Un esperimento illuminante fu il seguente.
Selezione di Topi: **Linee Pure**.

Ceppo 1 X Ceppo 2 (entrambe Albini): \longrightarrow **Pigmentati**

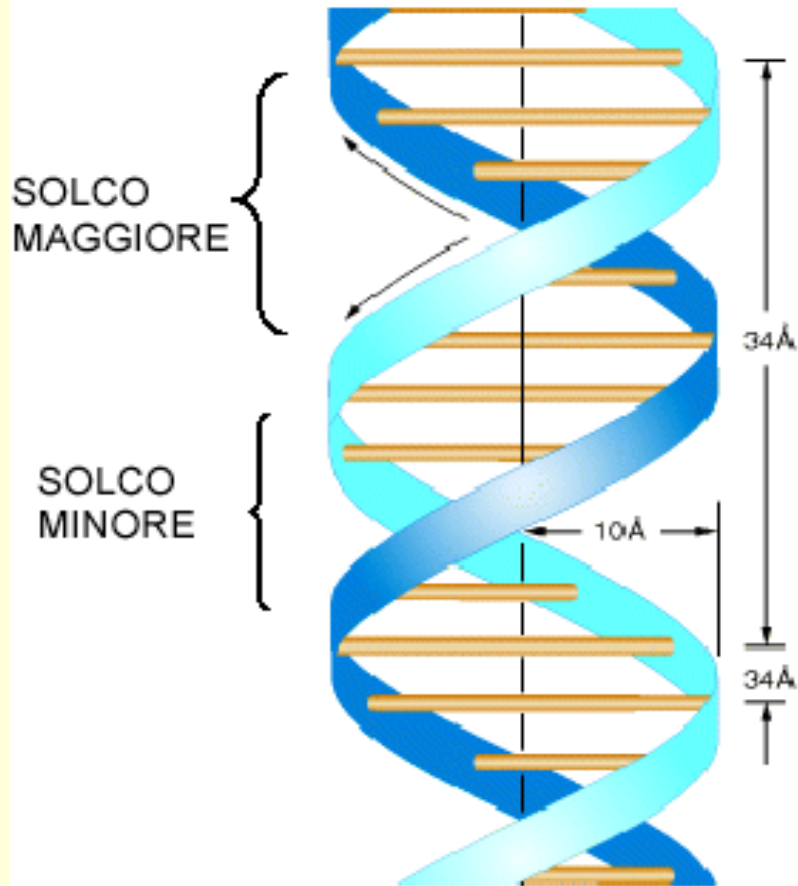
Ciò sembrava in contrasto con le leggi dell'ereditarietà, in quanto l'albinismo è un fenotipo recessivo, e quindi i genitori albini dei topi pigmentati avrebbero dovuto mancare totalmente dei geni necessari per produrre la melanina.

- 1944: Avery, McLeod e McCarthy dimostrano che il "principio trasformante" è DNA
- 1952: Hershey e Chase dimostrano che i fagi quando infettano i batteri iniettano il proprio DNA
- 1953: Watson e Crick determinano la struttura a doppia elica, sulla base dei dati di Wilkins e Franklin

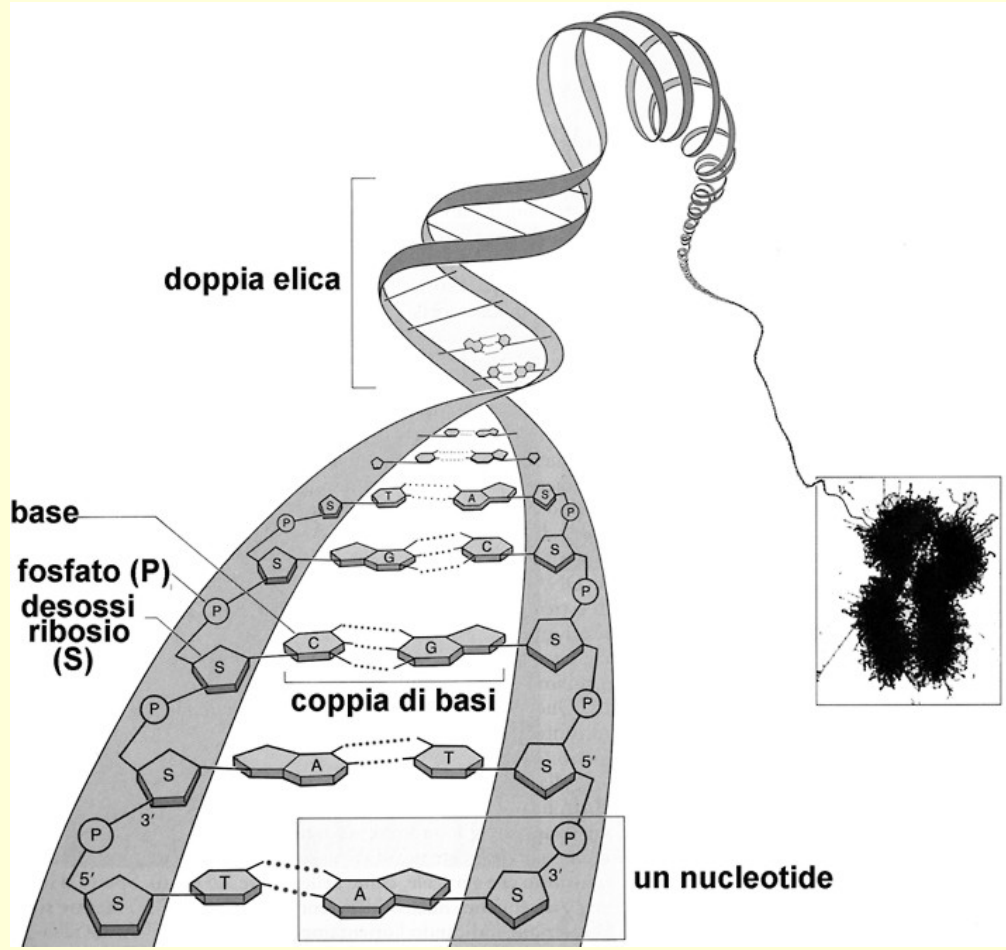
La composizione chimica del DNA

- Il DNA è costituito da unità monomeriche, i **nucleotidi**, ciascuno dei quali contiene:
 - uno zucchero pentoso, il desossiribosio
 - un gruppo fosfato
 - una base azotata
 - purine: adenina e guanina
 - pirimidine: citosina e timina

Caratteristiche della molecola di DNA

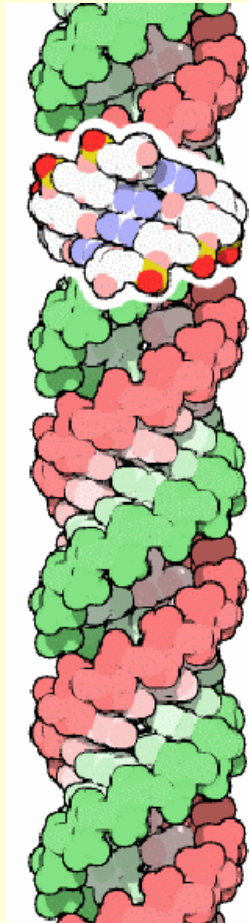


- E' costituita da due catene polinucleotidiche avvolte l'una intorno all'altra a formare una doppia elica Destrorsa
- Le due catene sono Antiparallele
- Per ogni giro dell'elica ci sono 10 coppie di basi
- Si possono distinguere due solchi (maggiore e minore) dovuti al fatto che i due filamenti non hanno sempre la stessa distanza dall'asse dell'elica



LE TRE FORME DEL DNA

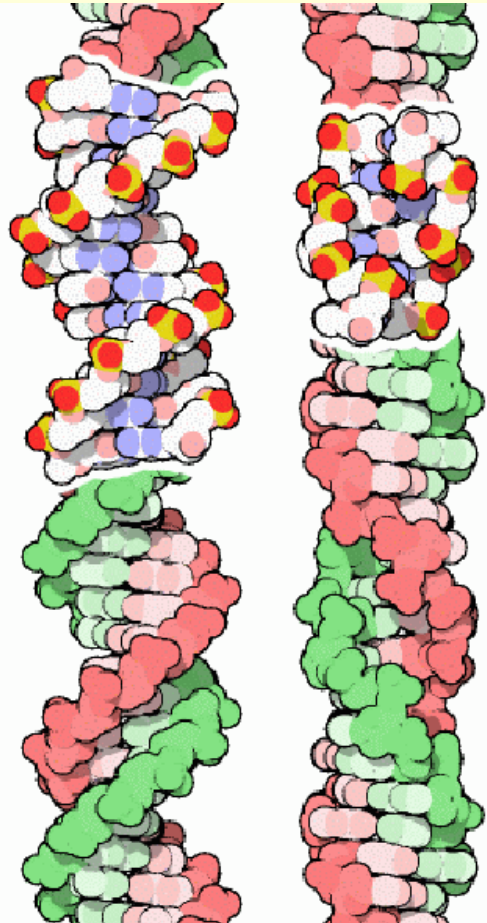
Forma A:



doppia elica destrorsa con
10,9bp/ giro

Si forma in condizioni disidratanti

Forma Z:



doppia elica sinistrorsa con 12bp/
giro

Si forma a concentrazioni elevate di sale
e richiede una particolare sequenza di basi,
con molte coppie alternate citosina-guanina
e guanina-citosina.

Forma B:

(Watson e Crick): doppia elica
destrorsa con 10bp/giro

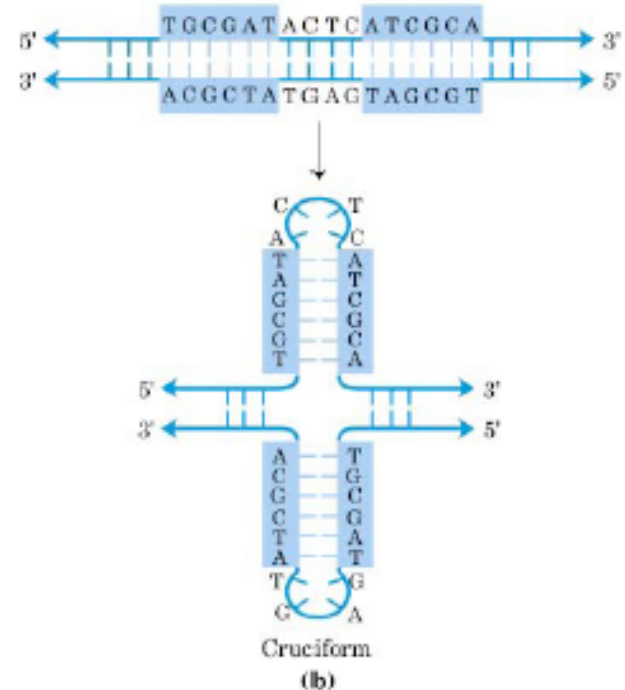
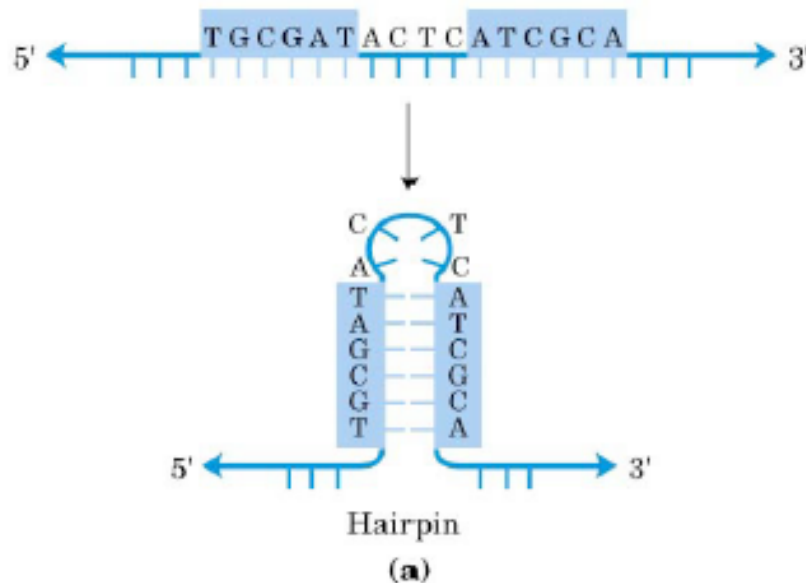
la normale doppia elica, chiamata elica-B,
si trova nelle condizioni tipiche che esistono nelle cellule vive

Il DNA può assumere forme insolite

- Z-DNA
- Forcine (sequenze palindromiche)
- Strutture a croce (sequenze invertite)

gli enzimi di restrizione permettono di tagliare il DNA in frammenti in corrispondenza di specifiche sequenze dette palindromiche, cioè che si possono leggere in entrambi i sensi sui due filamenti complementari; per esempio l'enzima *BamHI* riconosce la seguente sequenza (lo slash / rappresenta il punto di taglio; il colore evidenzia la simmetria nelle sequenze su ciascun filamento e sui due filamenti complementari):

5' ... **G / G A T C** C ... 3'
3' ... C C T A **G / G** ... 5'



La forma Z del DNA ha funzione biologica?

- E' correlata alla trascrizione:
 - Anticorpi anti Z-DNA si legano in corrispondenza di regioni attivamente trascritte dei cromosomi politenici di *Drosophila* e dei nuclei di cellule di mammifero
 - In geni umani c'è un'alta concentrazione di sequenze che favoriscono la formazione di Z-DNA vicino ai siti d'inizio della trascrizione
 - Il movimento della RNA polimerasi determina la formazione di Z-DNA
 - Sono state identificate proteine che si legano specificamente al DNA Z

quanto DNA c'è in una cellula e come è organizzato?

<i>Escherichia coli</i> (1997)	10^6 bp	ca 4.300 geni
Lievito (1996)	1.2×10^7	ca 6.000
Nematode (1998)	9.7×10^7	ca 18.000
<i>Drosophila</i> (2000)	1.2×10^8	ca 14.000
<i>Arabidopsis</i> (2001)	1.25×10^8	ca 25.000
Uomo (2001)	3×10^9	ca 30-40.000
Riso (2002)	4.2×10^8	ca 32-50.000
Topo (2002)	2.7×10^9	ca 30.000

GENOMICA COMPARATIVA

Organismo	Genoma (bp)	n. GENI	% GENI CONDIVISI CON UOMO
E. coli	4.6 milioni	4400	n.d.
Drosophila	165 milioni	13600	50,00%
Topo	3 miliardi	30000	90,00%
Arabidopsis	125 milioni	25000	n.d.
C. elegans	97 milioni	19000	40,00%
Lievito	12 milioni	6000	31,00%
Uomo	3 miliardi	30-40000	

Fonte: Howard Hughes Medical Institute (2001) **The genes we share with yeast, flies, worm and mice: new clues to human health and disease.**

Paradosso del valore C

- Ci possono essere ampie variazioni nel valore C fra specie la cui complessità apparentemente non varia altrettanto es Anfibi
- C'è un eccesso di DNA rispetto a quello che codifica per le proteine

il paradosso del valore C è la mancanza di una diretta relazione tra dimensioni del genoma e complessità dell'organismo

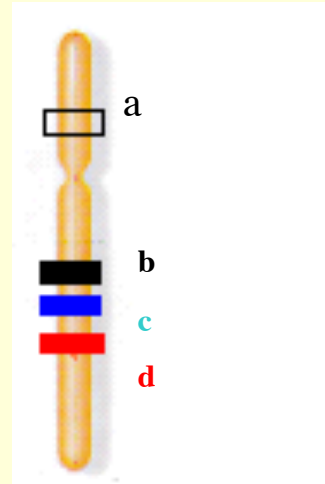
Geni e cromosomi

OGNI CROMOSOMA E' COSTITUITO DA UNA SUCCESSIONE LINEARE DI GENI O LOCI.

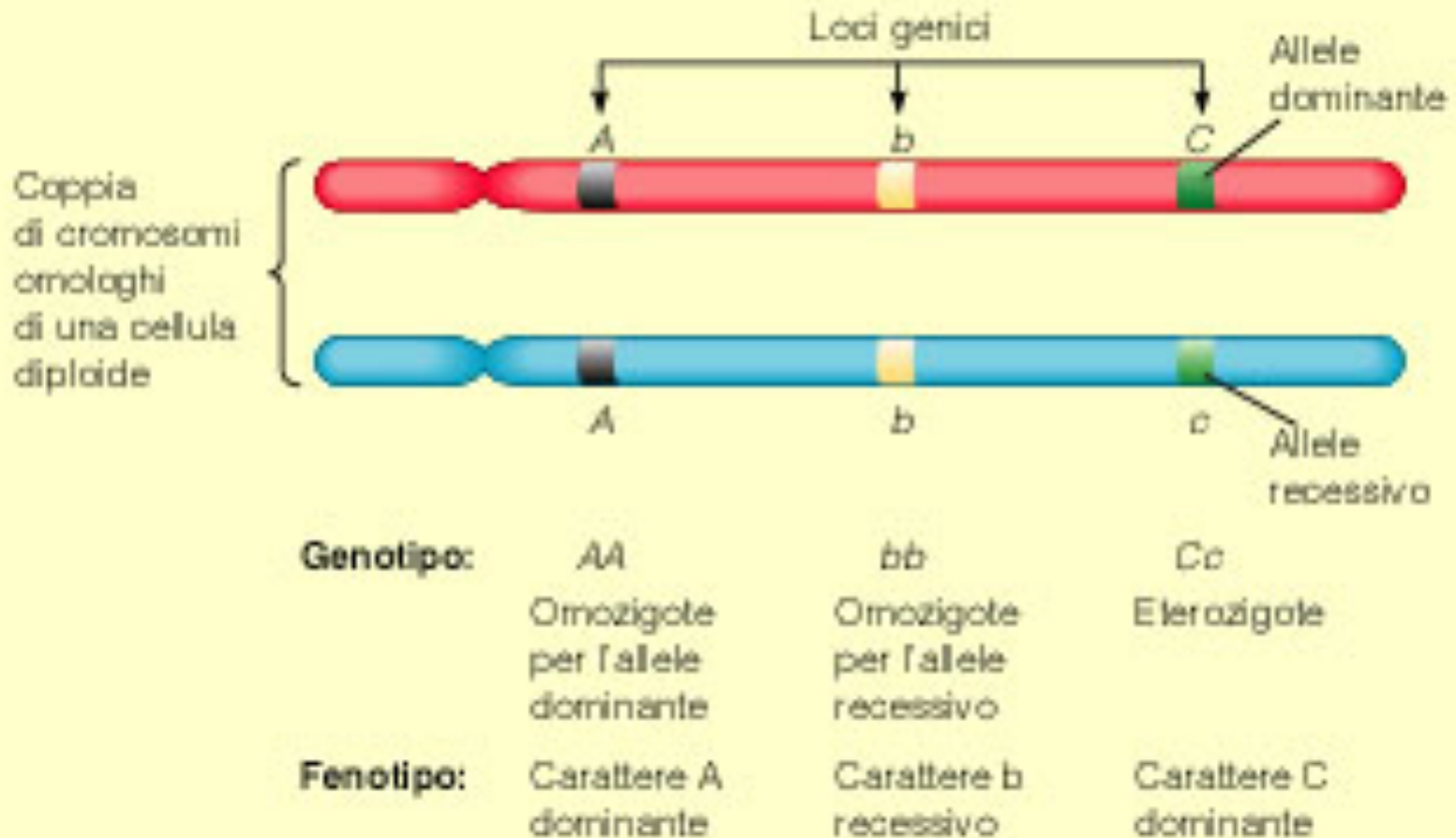
GENE: UNITA' EREDITARIA FONDAMENTALE

LOCUS: POSIZIONE OCCUPATA DA UN GENE SU UN CROMOSOMA.

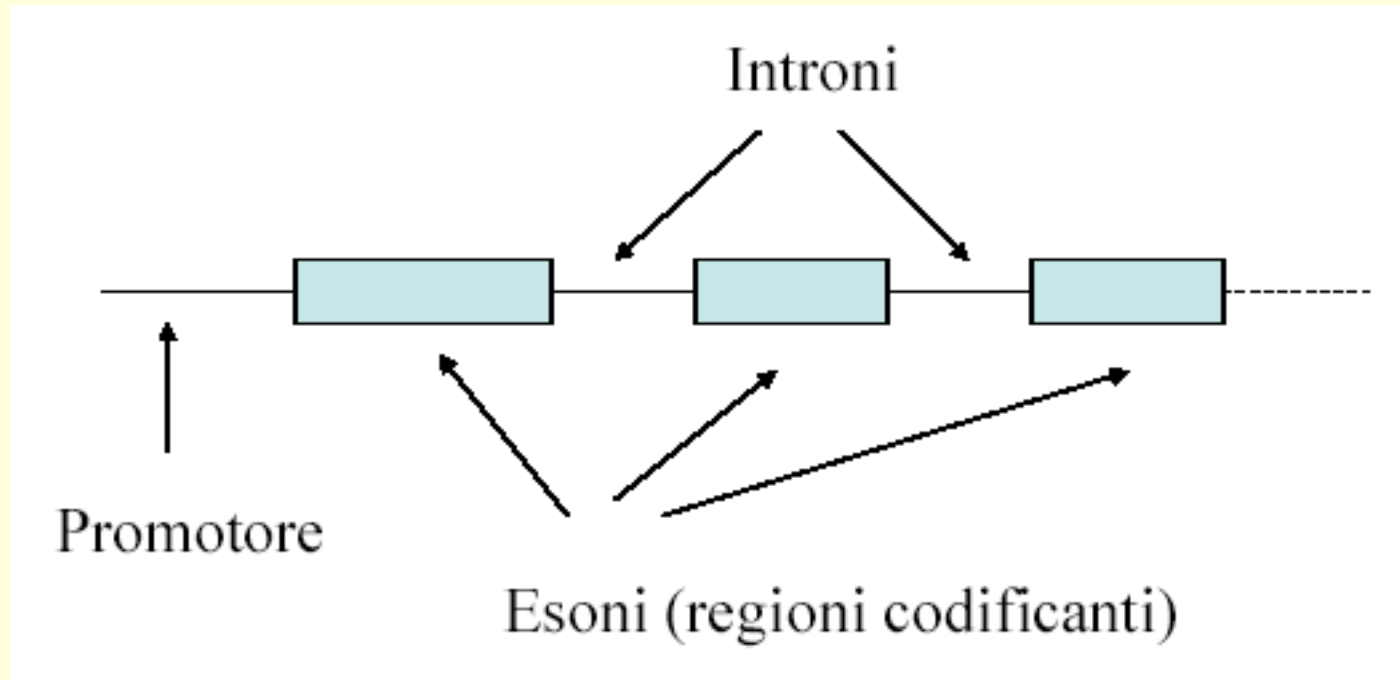
OGNI PAIO DI CROMOSOMI CONTIENE GLI STESSI GENI NELLO STESSO ORDINE MA NON NECESSARIAMENTE IN FORMA IDENTICA.



ALLELI: FORME DIVERSE DI UNO STESSO GENE



Il gene negli Eucarioti è più grande rispetto alla sequenza strettamente codificante:



Nel genoma umano si ha una grande varietà di geni. Alcuni (istoni) sono privi di introni, altri hanno esoni e introni variabili per numero e dimensioni.

Un gene per funzionare correttamente deve esprimersi al momento giusto e nella giusta posizione durante lo sviluppo.

Un gene inteso come unità funzionale deve contenere:

- ➡ una regione regolativa capace di ricevere e rispondere a segnali specifici per l'inizio della trascrizione (**promotore**)
- ➡ una regione che contiene segnali di terminazione per la fine della trascrizione (**terminatore**)

Un gene **eucariotico** inoltre contiene :

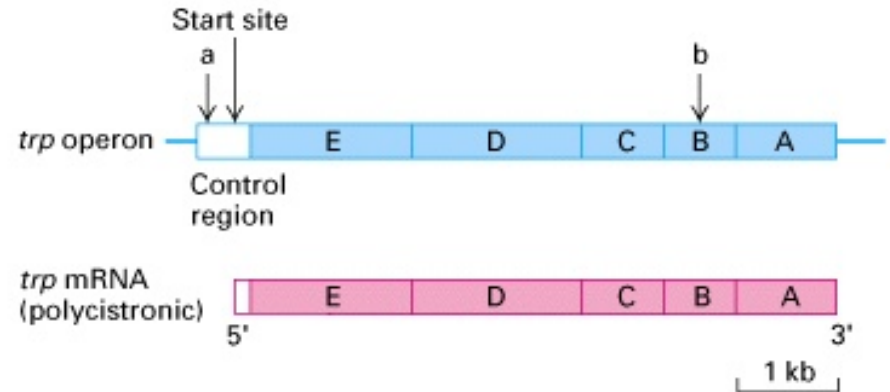
esoni = tratti di DNA codificante (trascritti e tradotti)

introni = tratti di DNA non codificante (trascritti e poi rimossi, non tradotti)

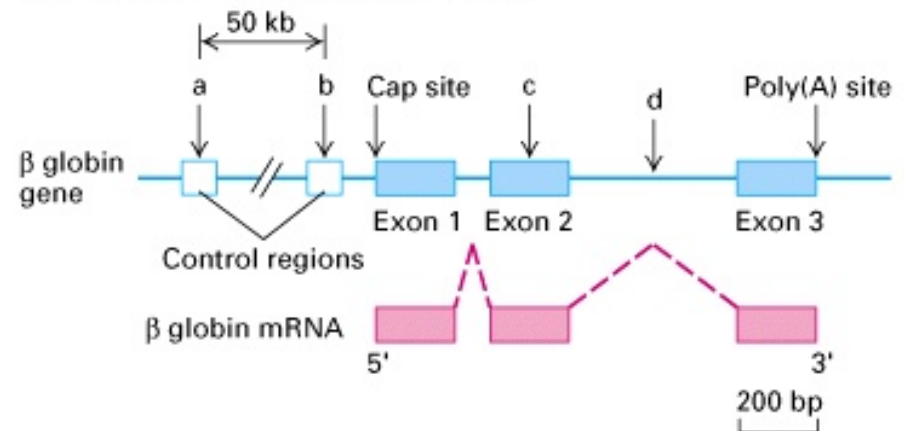
Procarioti vs. eucarioti

- I geni dei procarioti sono organizzati in operoni (policistronici)
- I geni degli eucarioti sono generalmente singole unità trascrizionali (monocistronici)

(a) Prokaryotic polycistronic transcription unit



(b) Eukaryotic simple transcription unit



I genomi degli Eucarioti contengono tre tipi di sequenze

- ✓ altamente ripetute
- ✓ mediamente ripetute
- ✓ uniche

Nell'uomo le sequenze uniche
(codificanti)
costituiscono meno del 5% del genoma

Classes of repetitive DNA

Interspersed (dispersed) repeats (e.g., Alu sequences)



Tandem repeats (e.g., microsatellites)



<http://www3.kumc.edu/jcalvet/bioc801b/>

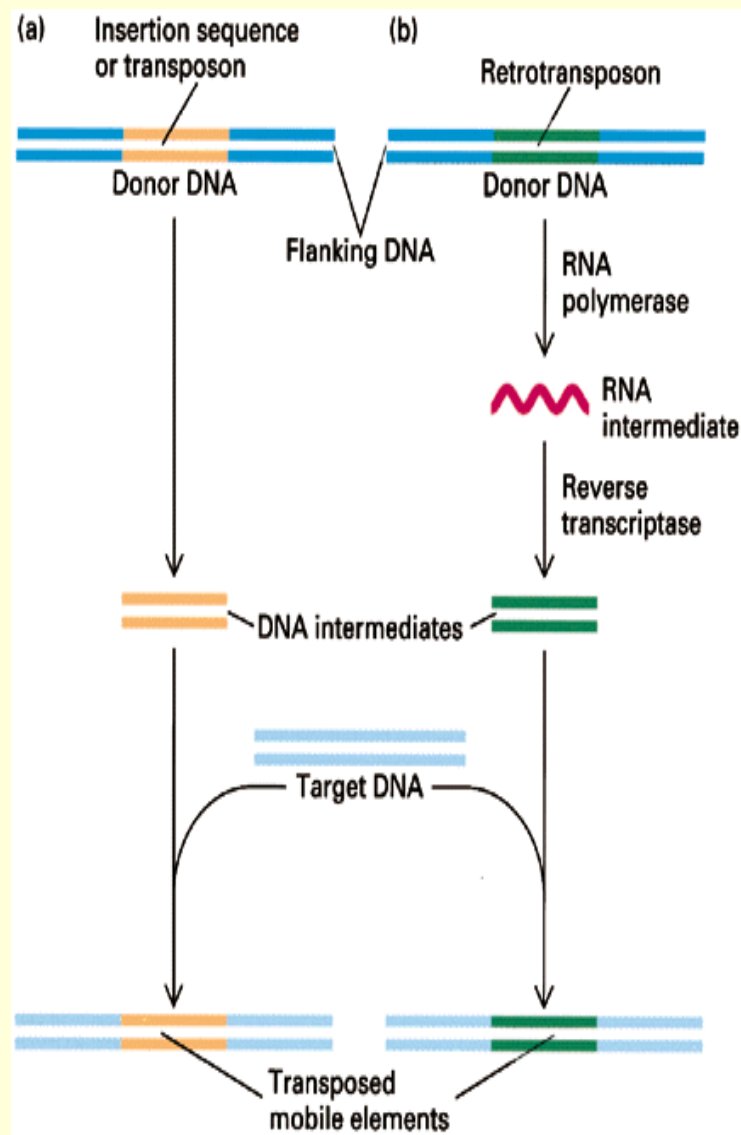
Le sequenze ripetute intersperse sono sparse in tutto il genoma, mentre le ripetute in tandem sono adiacenti le une alle altre e si trovano a livello dei centromeri e dei telomeri.

Nel genoma umano le sequenze ripetute costituiscono almeno il 50% del totale e sono suddivise in 5 classi:

- sequenze ripetute derivate da trasposoni (intersperse)
- copie inattive di geni
- ripetizioni di sequenze semplici (minisatelliti e microsatelliti): $(A)_n$, $(CA)_n$, $(CGG)_n$
 - **macrosatelliti**: regioni ripetute della lunghezza di circa 100 nucleotidi
 - **microsatelliti**: le regioni ripetute sono costituite da 2-6 nucleotidi.
- duplicazioni segmentali
- sequenze ripetute in tandem (centromeri, telomeri, cluster ribosomali)

DNA moderatamente ripetuto

- Famiglie geniche
- Trasposoni
 - DNA mobile - può muoversi da un cromosoma ad un altro
- Retrotrasposoni
 - si muovono attraverso un intermedio ad RNA



Una certa % dei geni umani è costituita da membri di famiglie di sequenze di DNA che condividono tra loro un buon grado di omologia



Ibridazione molecolare
Sequenziamento del DNA
PCR

A seconda del grado di omologia tra i diversi geni di una stessa famiglia, si distinguono:

• **FAM DI GENI RIPETUTI**

ORGANIZZAZIONE IN TANDEM,
IN UNO O PIU' CLUSTERS

(i componenti sono fundamentalmente identici tra loro. es. rRNA e Istoni)

• **FAM GENICHE CLASSICHE**

(i componenti mostrano una certa divergenza tra loro)

ORGANIZZAZIONE IN CLUSTERS
(es. GLOBINE)

ORGANIZZAZIONE INTERSPERSA
(es. actina)

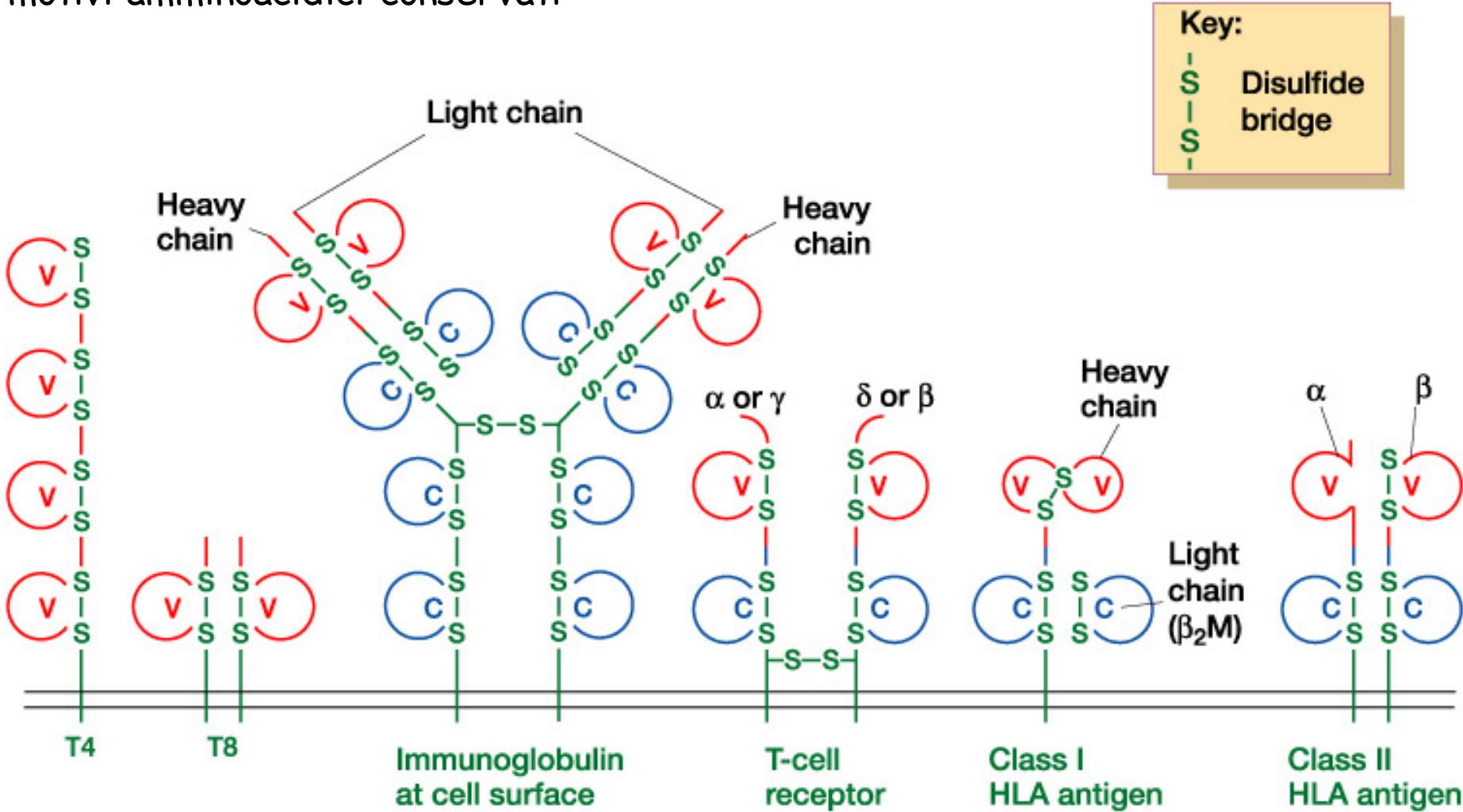
• **SUPERFAMIGLIE GENICHE**

ORGANIZZAZIONE INTERSPERSA
(geni che codificano prodotti funzionalmente correlati, ma con minima omologia di sequenza)

Famiglie geniche classiche con organizzazione interspersa: Non vi è una relazione fisica tra i membri di una famiglia -> localizzazione su più cromosomi. Possono derivare da eventi di duplicazione genica oppure da eventi di trasposizione

famiglia	n° copie	caratteristiche
aldolasi	5	3 geni funzionali e 2 pseudogeni, su 5 cromosomi diversi
NF1	>12	1 gene funzionale (17q), copie difettose, non processate
gliceraldeide 3-fosfato-deidrogenasi	>18	1 gene funzionale
actina	>20	4 geni funzionali

Superfamiglie geniche: i geni codificano prodotti funzionalmente correlati ma che non mostrano elevati gradi di omologia di sequenza né motivi amminoacidici conservati



I membri della superfamiglia delle Ig sono proteine di superficie con strutture e domini simili tra loro

Perché esistono le famiglie geniche?

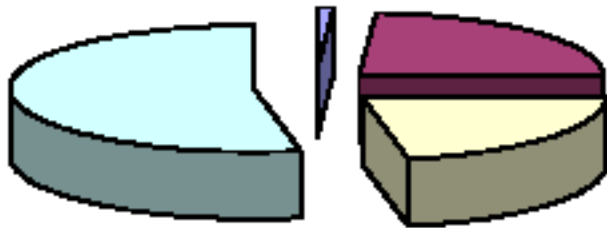
- Richiesta di grandi quantità di prodotto.
- Funzioni specializzate.
- Espressione differenziale.

TABLE 9-2 Protein Families in Vertebrates and Invertebrates

Family	Number of Proteins in Family
COMMON PROTEINS	
Actins	5–30
70 K Heat-shock proteins	3
Keratins	>20
Myosin, heavy chain	5–10
Protein kinases	10–100s
Transcription factors	10–100s (?)
Tubulins, α and β	3–15
INSECT PROTEINS	
Eggshell proteins (silk moth and fruit fly)	50
VERTEBRATE PROTEINS	
Globins (many species)	
α -globin	1–3
β -like globins	5
Immunoglobins, variable regions (many species)	500
Ovalbumin (chicken)	3
Transplantation antigens (mouse and human)	50–100
Visual pigment protein (human)	4
Vitellogenin (frog, chicken)	5

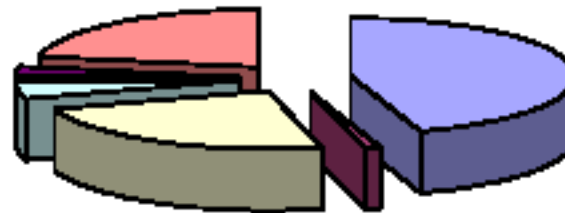
Nel genoma umano ci sono pochi geni

Solo 1% del genoma umano codifica per proteine



- Esoni
- Introni
- DNA intergenico
- DNA ripetuto

Il genoma umano è costituito per la maggior parte da sequenze ripetute

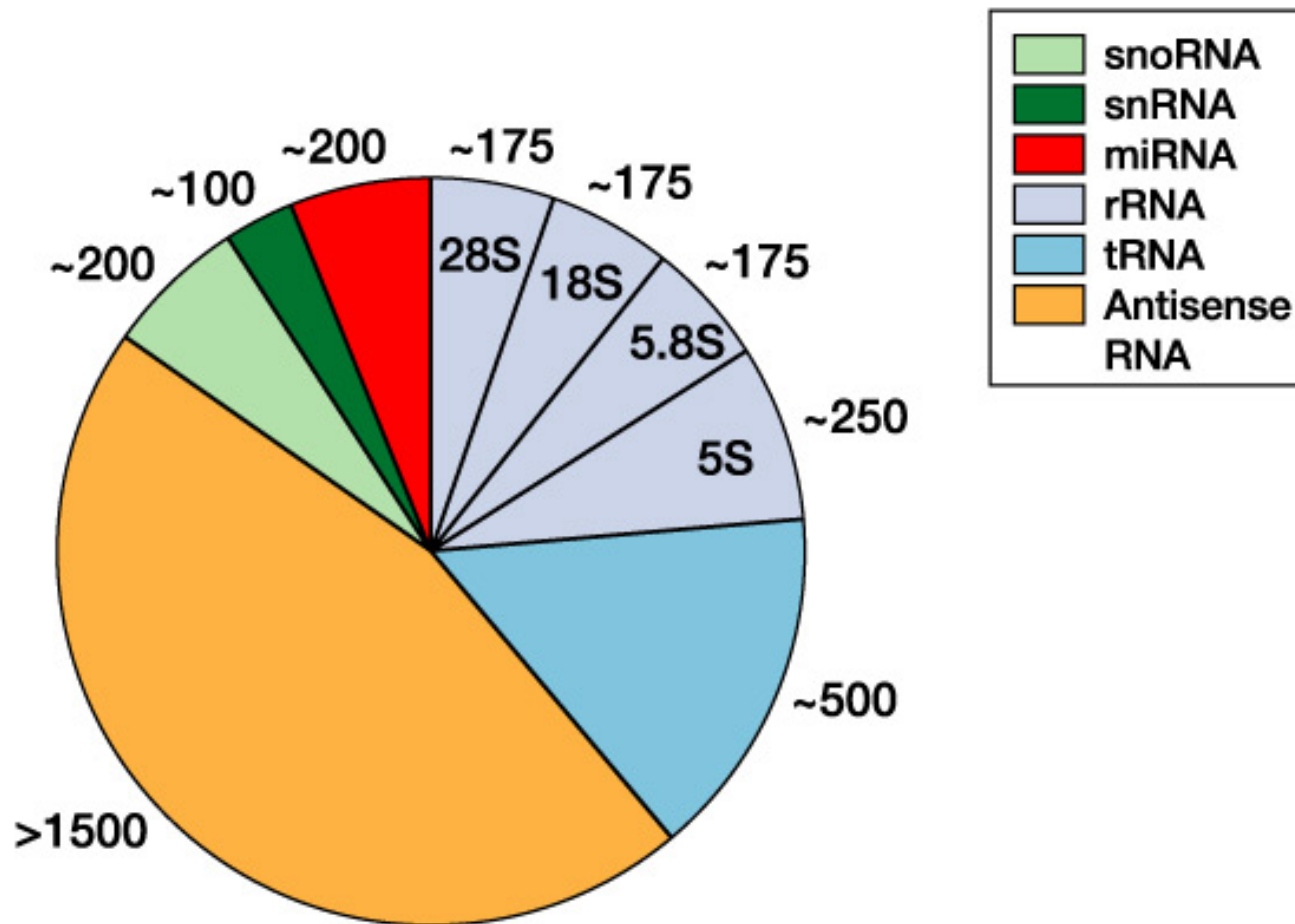


- Trasposoni
- Esoni
- Introni
- Duplicazioni
- Ripetizioni semplici
- DNA intergenico

Funzione dei geni negli eucarioti superiori

- Geni che codificano per prodotti proteici;
- Geni che codificano per RNA non codificanti (RNA-genes).

Il 5-10% dei geni nucleari sono "RNA-genes"



COMPLESSITA' DI UN GENOMA

Lunghezza totale di tutte le differenti sequenze presenti in esso.

La complessità di un genoma si valuta studiando le cinetiche di denaturazione e rinaturazione

COMPLESSITA' DI UN GENOMA

Nonostante la relativa fragilità dei ponti idrogeno che legano ciascuna coppia di basi, ogni molecola di DNA contiene un numero tale di coppie di basi, da impedire che la catena possa separarsi **spontaneamente** in condizioni fisiologiche.

Se tuttavia la doppia elica viene sottoposta alla temperature di 70°C o oltre, o in condizioni di pH *estremo* ($\text{pH} < 3$ o $\text{pH} > 10$), i legami idrogeno tra le basi appaiate (adenina-timina e guanina-citosina) si rompono ed essa si scinde nelle sue due catene complementari: questo processo è chiamato **denaturazione o effetto fusione**.

La denaturazione è un effetto reversibile: abbassando lentamente la temperatura le due catene complementari possono di nuovo appaiarsi formando una doppia elica normale (**rinaturazione**).

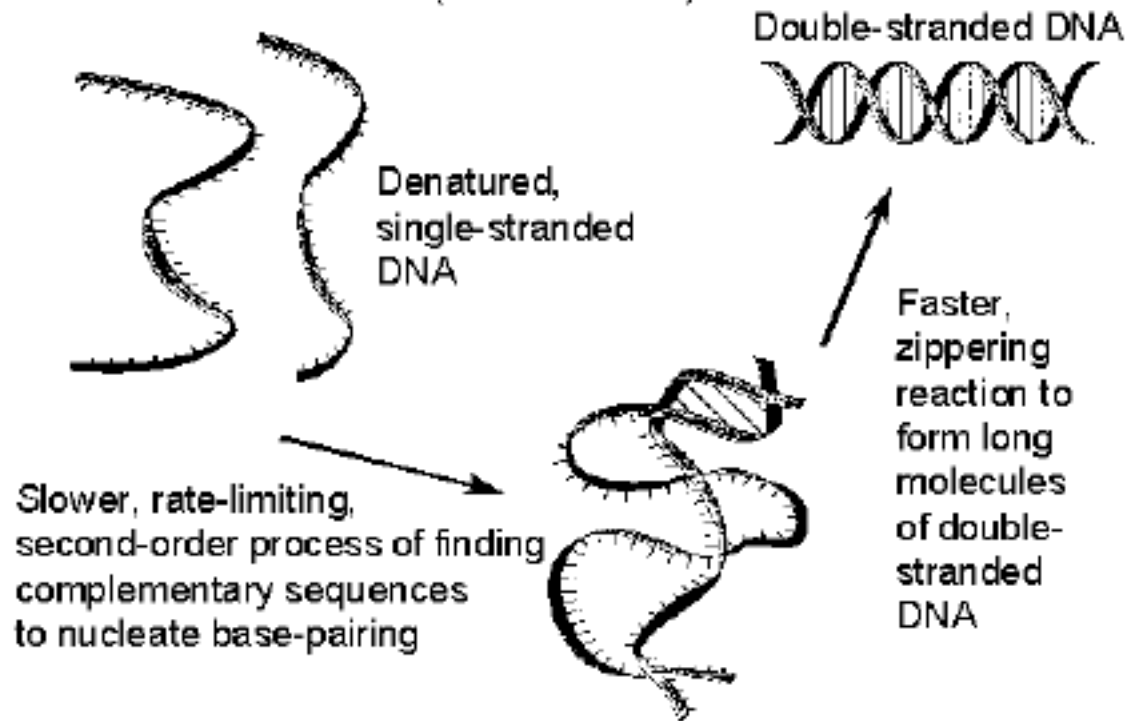
La denaturazione si accompagna ad aumento dell'assorbimento della luce ultravioletta a 2600Å (effetto ipercromico).

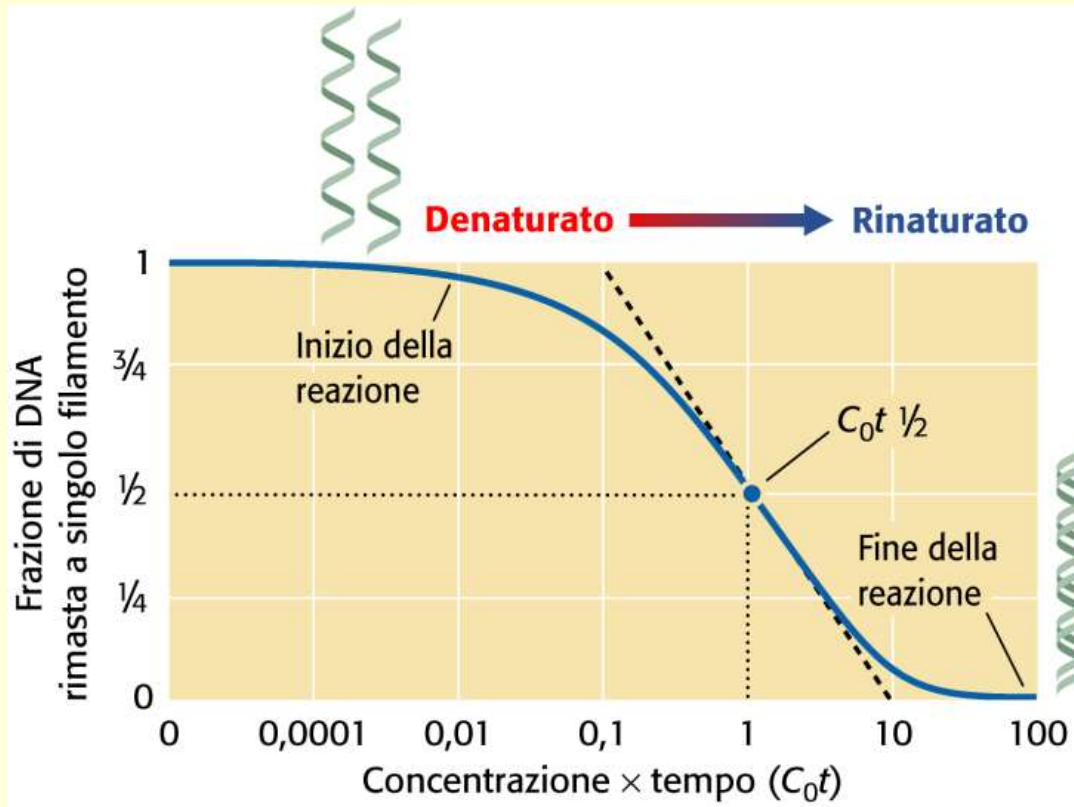
Temperatura di fusione ed effetto ipercromico dipendono dalla composizione in basi del DNA, in quanto la rottura dei tre ponti idrogeno tra le basi G e C richiede una temperatura più elevata di quella che è sufficiente a rompere i due esistenti tra le basi A e T. Le due catene della doppia elica possono essere considerate come una coppia di modelli, positivo e negativo, ognuno dei quali specifica il rispettivo complemento per la sintesi di una catena complementare che quindi sarà identica alla catena opposta. In tal modo le due molecole di DNA *figlie* prodotte risultano identiche alla doppia elica madre. Il meccanismo di duplicazione del DNA risulta così essere **semiconservativo**, in quanto delle due catene originarie una è conservata in ciascuna delle due cellule figlie e l'altra è di nuova sintesi

DENATURAZIONE E RINATURAZIONE DEL DNA

a). Complexity of chromosomal DNA

DNA reassociation (renaturation)

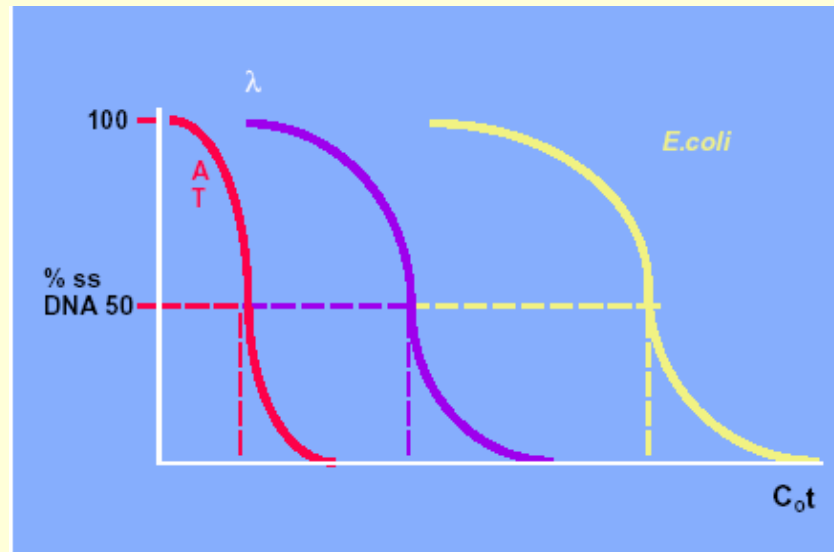




Curva di rinaturazione o curva " C_0t " per un DNA ideale. Sull'asse delle ordinate è riportata la % di DNA a singolo filamento, sulle ascisse il log del prodotto della concentrazione del DNA (C_0) e del tempo.

L'andamento della C_{ot} è funzione di due fattori

- le dimensioni del genoma
- la presenza e quantità di DNA ripetuto



Curve C_{ot} per campioni di DNA a diversa complessità. Il tempo necessario perchè il 50% del DNA si rinaturi è preso come misura della complessità.

La $C_{ot} 1/2$ è direttamente correlata alla quantità di DNA in un genoma: più il genoma è complesso, meno copie ci saranno di una data sequenza in una certa massa di DNA.

Curva C_{ot} di genoma di Mammifero. La cinetica di riassociazione è complessa, perchè si tratta di una miscela di frazioni di DNA.

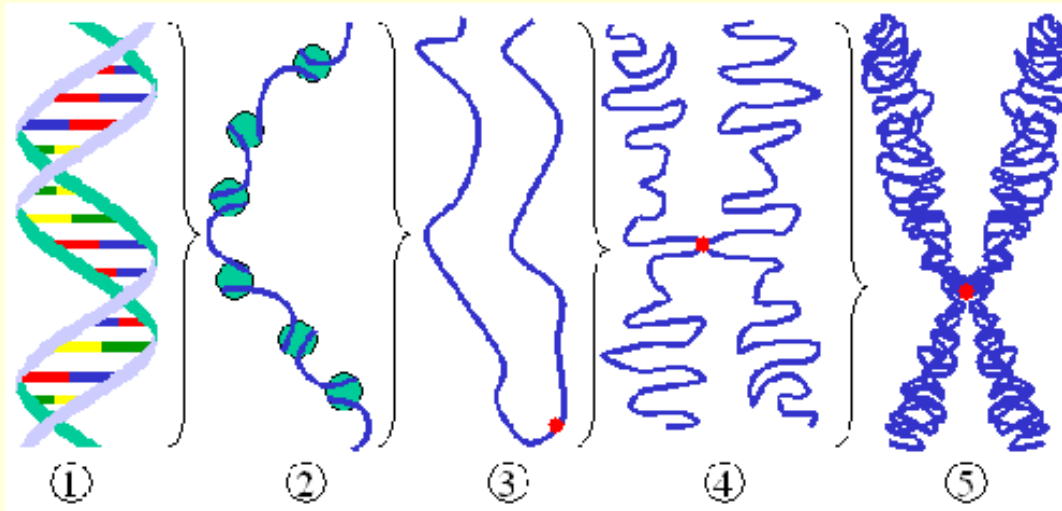
Negli Eucarioti

- Organizzazione della cromatina
 - Cromosomi politenici
 - Cromosomi "lampbrush"

Cromosoma politenico è un tipo speciale di cromosoma costituito da un fascio di parecchi cromatidi derivati da ripetuti cicli di replicazione di singoli cromatidi senza divisione nucleare. Questo cromosoma è caratteristico di vari tessuti dei ditteri.

Cromosomi Lampbrush: si formano durante la meiosi nell'oogenesi di anfibi. Sono grossi cromosomi bivalenti (2 paia di cromatidi fratelli si connettono al chiasma).

La **cromatina** rappresenta la forma in cui gli acidi nucleici si trovano nel nucleo di una cellula eucariota. La cromatina è formata da acido desossiribonucleico, DNA, avvolto su gruppi di proteine dette istoni (proteine basiche), formando un Nucleosoma, e da proteine non-istoniche (proteine neutre o acide); essa è poi ripiegata in vario modo. Esistono infatti diversi livelli di organizzazione della cromatina:



Differenti livelli di condensazione del DNA:

(1) DNA a doppia elica

(2) Cromatina: filo di DNA con istoni

(3) Cromatina condensata durante l'interfase

(4) Cromatina condensata durante la profase (sono presenti due copie di molecole di DNA)

(5) Cromosoma durante la metafase.

Eterocromatina. Regioni che si colorano intensamente

Eucromatina: Regioni che si colorano meno intensamente (perché impacchettata in modo meno denso)

Il nucleosoma

Per nucleosoma s'intende una struttura formata dal cilindretto di istoni, o *core*, alla cui superficie si avvolge il DNA.

Figure 19.3 The nucleosome consists of approximately equal masses of DNA and histones (including H1). The predicted mass of the nucleosome is 262 kD.

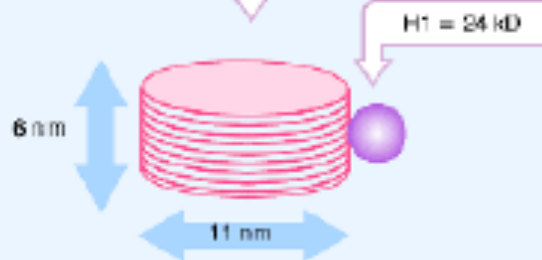
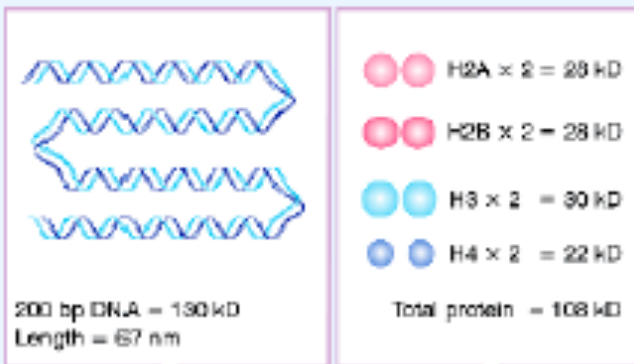
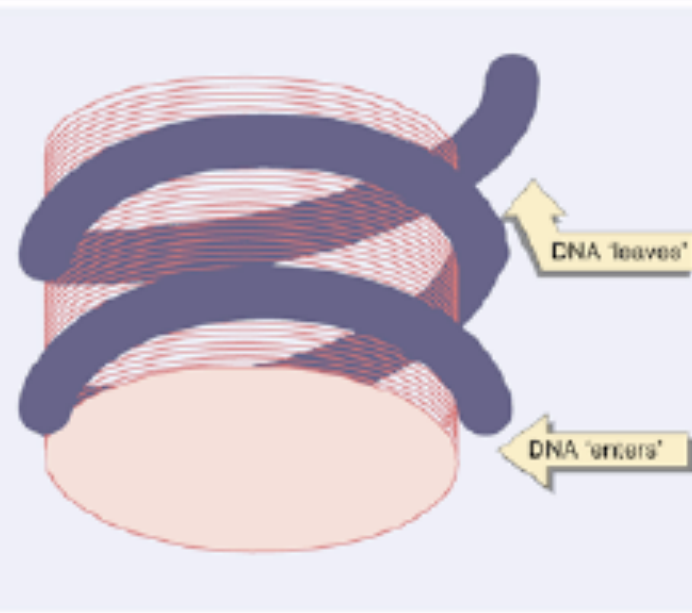
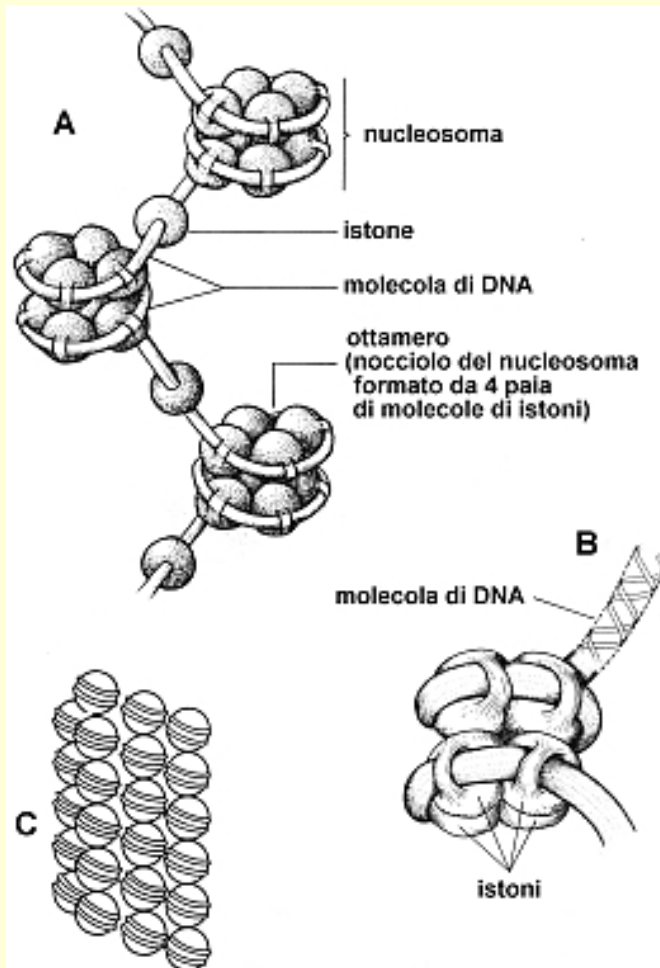


Figure 19.4 The nucleosome may be a cylinder with DNA organized into two turns around the surface.



Struttura dei nucleosomi



Struttura della fibra cromatinica.

A - fibra cromatinica a collana di perle.

B - Nucleosoma.

C - Superspiralizzazione del nucleofilamento, che rappresenta la forma condensata della fibra cromatinica.

L'esame al microscopio elettronico delle nucleoproteine estratte dal nucleo interfaseico ha permesso, in certi casi, di verificare che le fibre cromatiniche hanno una **struttura a filo di perle**. Ogni sferetta, detta nucleosoma, ha un diametro di circa 10 nm, ed è connessa a quelle adiacenti da un filamento lungo 14 nm dello spessore di 3 nm. Il nucleosoma presenta le seguenti caratteristiche:

- possiede un **core proteico** formato da istoni (H2A, H2B, H3, H4) organizzati in ottamero (4 coppie, una per ogni tipo di istone)
- il core è circondato da un **segmento di DNA avvolto ad elica**; la lunghezza di questo DNA corrisponde a 140 coppie di basi complementari.

Un istone H1, situato all'esterno di ogni nucleosoma, controllerebbe il grado di avvolgimento, o *condensazione*, del filamento di cromatina. Il DNA compreso fra due nucleosomi contiene in media 40 coppie di basi. Quando i nucleosomi sono accostati l'uno all'altro, la fibra cromatinica ha un diametro di 10 nm. Questa fibra può spiralizzarsi assumendo la forma di una fibra del diametro di 20 nm.

Genoma nucleare Eucariotico

- Ogni specie ha un numero caratteristico di cromosomi
- I geni sono segmenti di cromosomi nucleari
- La ploidia si riferisce ad un numero di completi sets di cromosomi
 - aploide ($1n$): un completo set di geni
 - diploide ($2n$)
 - poliploide ($=3n$)
- Nei diploidi i cromosomi esistono in paia omologhe (omologhi)
 - strutturalmente simili
 - stessa sequenza di geni
 - possono contenere alleli differenti

PACKING RATIO

$$\frac{\text{lunghezza molecola DNA}}{\text{lunghezza struttura che la contiene}}$$

- Nucleosomi P.R. circa 6
- Fibra 30 nm P.R. circa 40
- Eucromatina P.R. ≥ 1000
- Cromosomi P.R. ≥ 10000