

# **Stabilità del genoma!**

- I danni al DNA sono una continua minaccia alla stabilità genomica di una cellula
- Ogni giorno il genoma di una cellula va incontro a circa un milione di lesioni
- Gli organismi viventi hanno evoluto vari meccanismi per proteggere l'integrità del DNA
- Questi meccanismi si basano sulla attività degli enzimi della riparazione del DNA

# La Genomica

- La Genomica è lo studio della struttura, del contenuto e dell'evoluzione dei genomi, ma si estende anche all'analisi dell'espressione e delle funzioni dei geni e delle proteine.
- Le analisi sperimentali sono integrate da due nuove discipline correlate alla genomica, la bioinformatica e la biologia computazionale. La **bioinformatica** è dedicata alla risoluzione di problemi biologici a livello molecolare con metodi informatici. La **Biologia Computazionale** è l'applicazione di strumenti propri delle scienze dell'informazione (es. algoritmi, intelligenza artificiale, databases) a problemi di interesse biologico, biotecnologico e biomedico.

# Le mappe genetiche

- Una mappa genetica è la descrizione dell'ordine relativo dei marcatori genetici nei gruppi di associazione (linkage) nei quali la distanza tra geni di riferimento, o marcatori, è espressa come unità di ricombinazione ( unità di misura centimorgan)

# **Le mappe fisiche**

- Una mappa fisica è un insieme di tratti contigui di DNA cromosomico (contig) nei quali la distanza tra le sequenze usate come marcatori è espressa in kilobasi. La mappa fisica più dettagliata è la sequenza completa del genoma

# **Le mappe citogenetiche**

- Una mappa citologica è costituita dai profili di bandeggio cromosomico osservabili al microscopio.

- Natura della variabilità genetica
- Variabilità interspecie e intraspecie; interpopolazione e intrapopolazione
- Mutazione: mutazione somatica, mutazione germinale; allele ancestrale e allele derivato; polimorfismo genetico per sostituzione di base nell'uomo: 1 ogni 100-200 basi di DNA per un totale di decine di milioni di polimorfismi
- I polimorfismi per piccole delezioni o duplicazioni (polimorfismi indel) sono 10 volte meno comuni. Le indel che coinvolgono una sola base si chiamano SNP (Single Nucleotide Polymorphism)

- Microsatelliti
- Minisatelliti
- Polimorfismi VNTR
- Polimorfismi dovuti a trasposoni
- Polimorfismi strutturali

Polimorfismi dovuti a microsatelliti: ripetizioni in tandem di corte sequenze di 1-6 basi (meno di 30 ripetizioni). **I microsatelliti sono polimorfici e multiallelici.** Si trovano dispersi in tutto il genoma.

Polimorfismi dovuti a minisatelliti: ripetizioni in tandem di una sequenza di circa 100 basi per una lunghezza complessiva di decine di migliaia di basi. Hanno una distribuzione preferenziale nelle regioni subtelomeriche.

Polimorfismi di VNTR, caratterizzati da un numero variabile di ripetizioni in tandem (sequenze dei centromeri, dei telomeri, dell'rRNA)

Polimorfismi dovuti a trasposoni. Nell'uomo in particolare Alu, Line1 e SVA (Sine-VNTR-Alu)

Polimorfismi strutturali, variazioni di struttura del genoma (inversioni, inserzioni, delezioni, duplicazioni) come le CNV (copy number variation)



Il 5% del genoma umano è rappresentato dalle duplicazioni segmentali, duplicazioni di un segmento di DNA la cui grandezza può andare da poche kb fino a centinaia di kb. Le duplicazioni segmentali possono trovarsi in omozigosi, cioè sono state fissate, oppure in eterozigosi. La variabilità per la presenza di duplicazioni o delezioni in una popolazione va sotto il nome di Copy Number Variation (CNV)

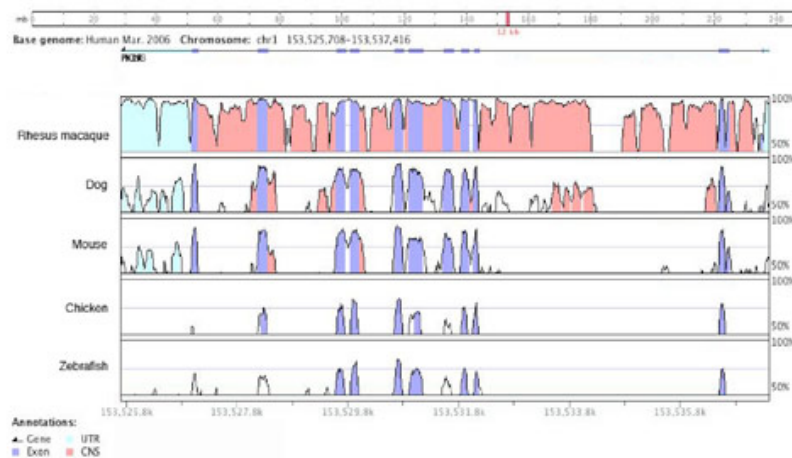
# La genomica comparativa

- La genomica comparativa è un campo della ricerca biologica in cui le sequenze genomiche di differenti specie sono confrontate. Il confronto tra sequenze di interi genomi fornisce un punto di vista dettagliato di come gli organismi sono correlati l'uno all'altro a livello genetico.
- La genomica comparativa fornisce anche un mezzo potente per studiare i cambiamenti evolutivi tra gli organismi, aiutando a identificare i geni che sono conservati o comuni tra le specie, così come i geni che danno a ogni organismo le sue caratteristiche uniche.
- La genomica comparativa può fornire benefici ad un ampio spettro di studi come la biomedicina, l'ecologia, l'agricoltura, la zoologia e le biotecnologie.

La genomica comparativa si basa sulla tecnologia del sequenziamento del DNA.

Punto di entrata: confronto tra specie su dimensione del genoma, numero di geni, numero di cromosomi.

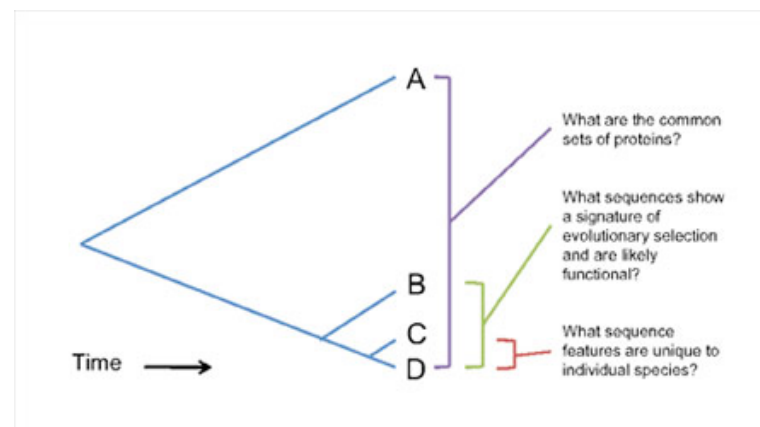
Confronto diretto delle sequenze di DNA:



sintenia

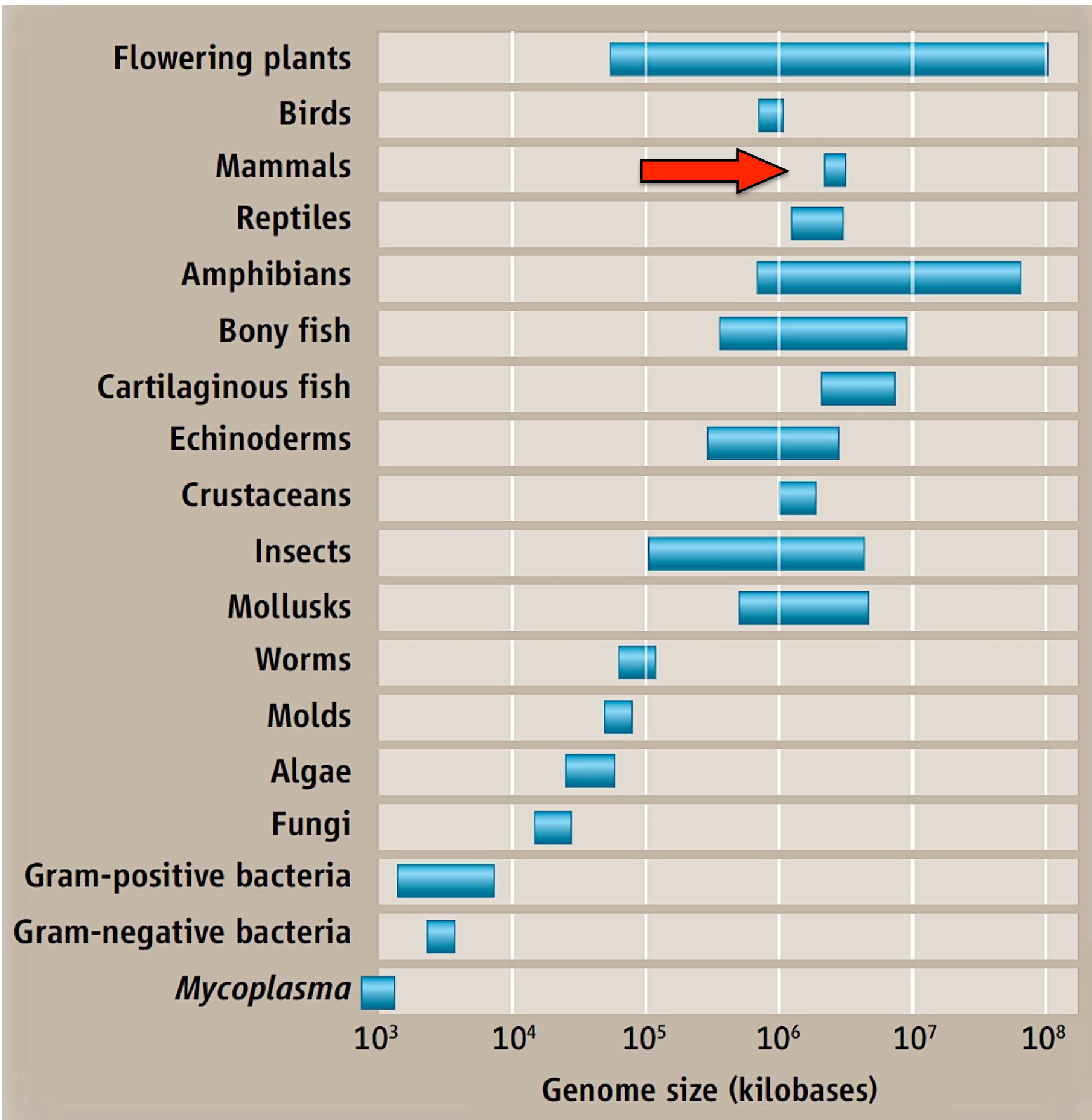
allineamento di DNA  
omologo

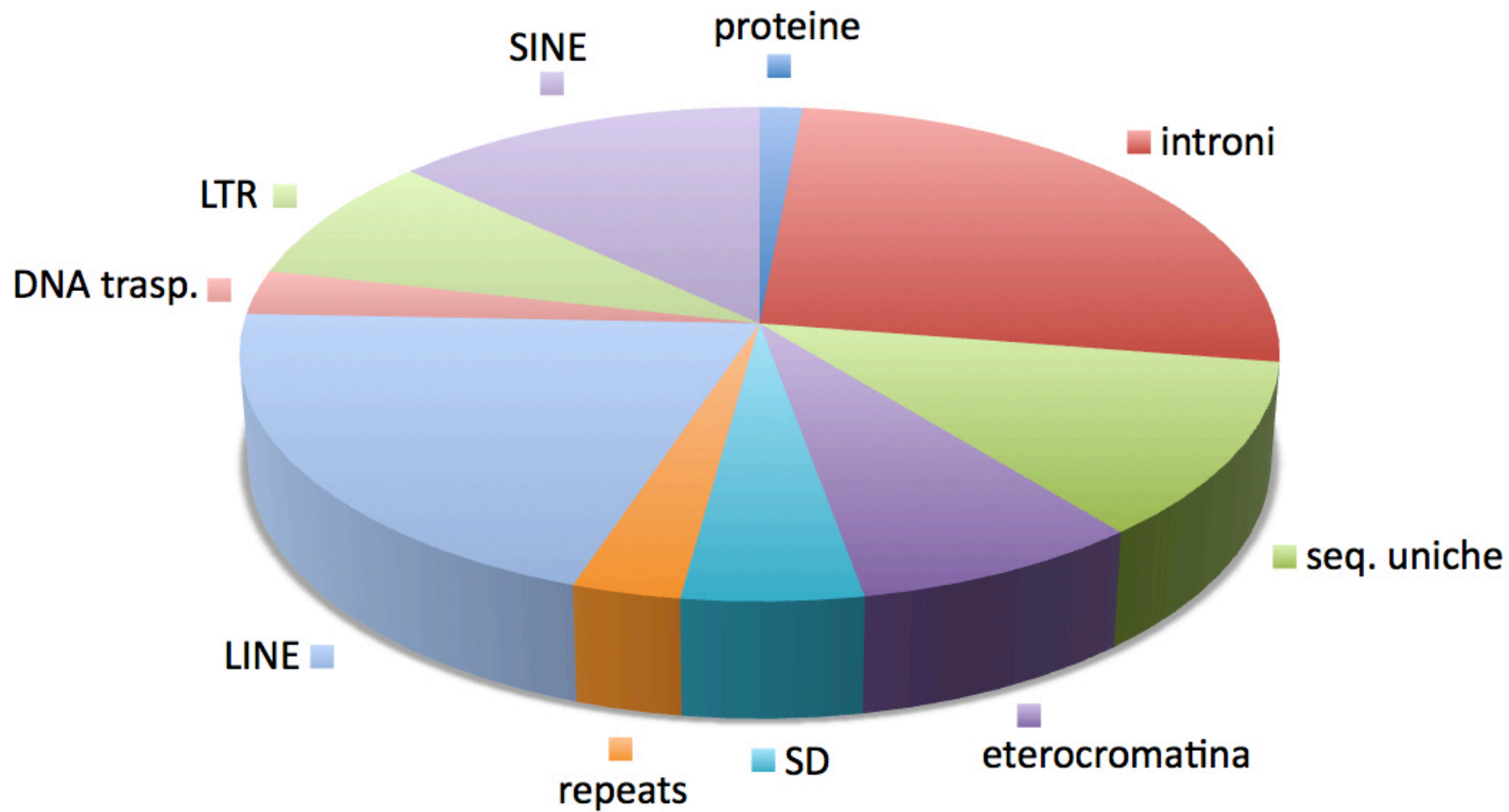
confronti di genomi a  
differenti distanze evolutive



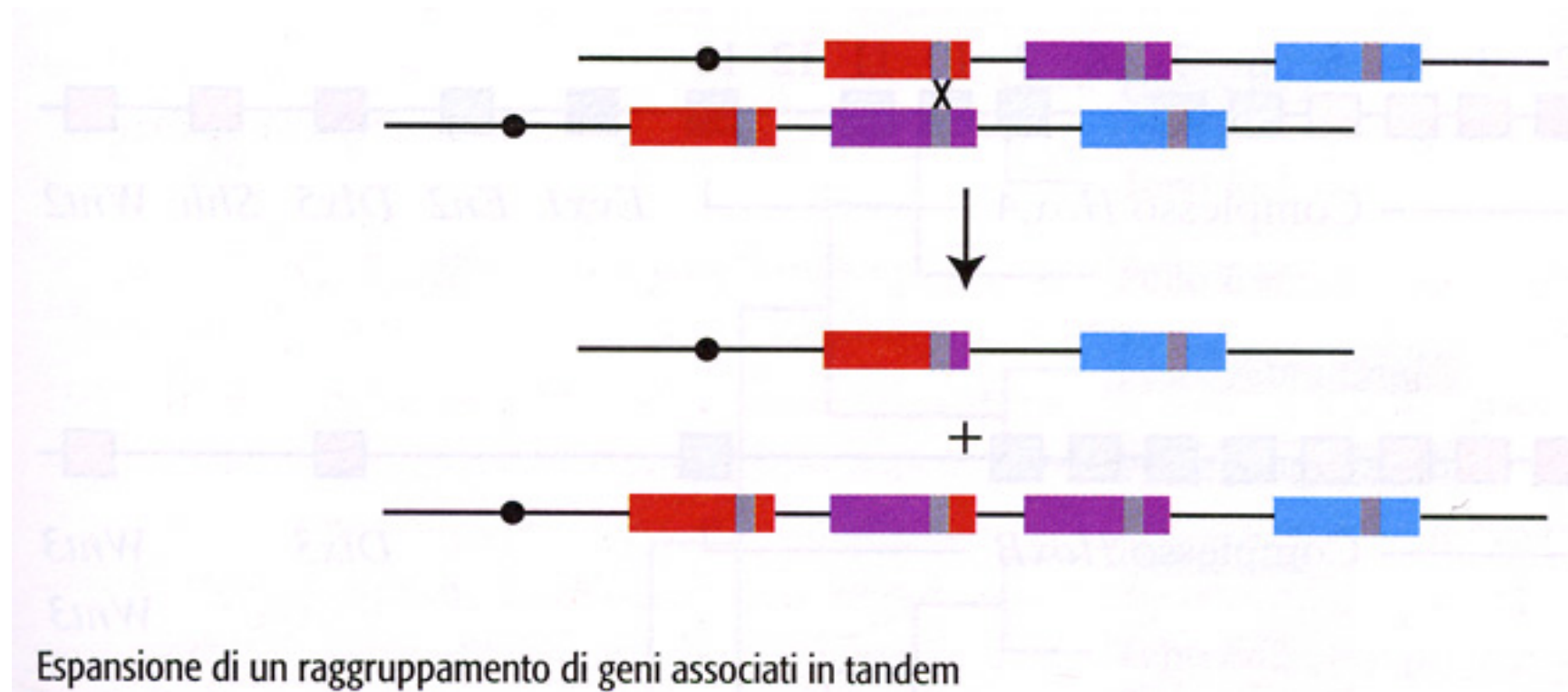
Organism	Genome size
-----	-----
Lilium longiflorum	90,000,000,000
Amphiuma means	84,000,000,000
Necturus maculosus	81,300,000,000
Pinus resinosa	68,000,000,000
Lilium formosanum	36,000,000,000
Coscinodiscus asteromphalus	25,000,000,000
Triturus cristatus	20,600,000,000
Allium cepa	18,000,000,000
<b>Triticum aestivum (wheat)</b>	16,000,000,000
Cercopithecus nigroviridis	5,117,100,000
Zea mays	5,000,000,000
Potorous tridactylus	4,026,600,000
Marmosa pusilla	4,026,600,000
Peromyscus maniculatus	4,006,800,000
Galago crassicaudatus	3,922,900,000
Symphalangus syndactylus	3,705,800,000
Cercocebus aterrimus	3,705,800,000
Cercocebus atys	3,691,000,000

Organism	Genome size
-----	-----
Macaca fascicularis	3,404,800,000
<b>Homo sapiens</b>	3,400,000,000
Macaca nigra	3,399,900,000
<b>Zea mays</b>	<b>3,000,000,000</b>
Callicebus torquatus	2,225,500,000
Eptesicus fuscus	2,220,500,000
Boa constrictor	2,100,000,000
Pipistrellus savii	1,973,800,000
Gallus gallus	1,200,000,000
Musca domestica	900,000,000
Crassostrea virginica	700,000,000
<b>Oryza sativa (riso)</b>	<b>430.000.000</b>
Emericella nidulans	31,000,000
Saccharomyces cerevisiae	12,067,280
<b>Escherichia coli</b>	4,639,221
Mycobacterium tuberculosis	4,397,000
Methanococcus jannaschii	1,664,974
Borrelia garinii	953,000
Borrelia afzelii	948,000
Borrelia burgdorferi	946,000
Mycoplasma pneumoniae	816,394
Mycoplasma genitalium	580,000
Human immunodeficiency virus type 1	9,750

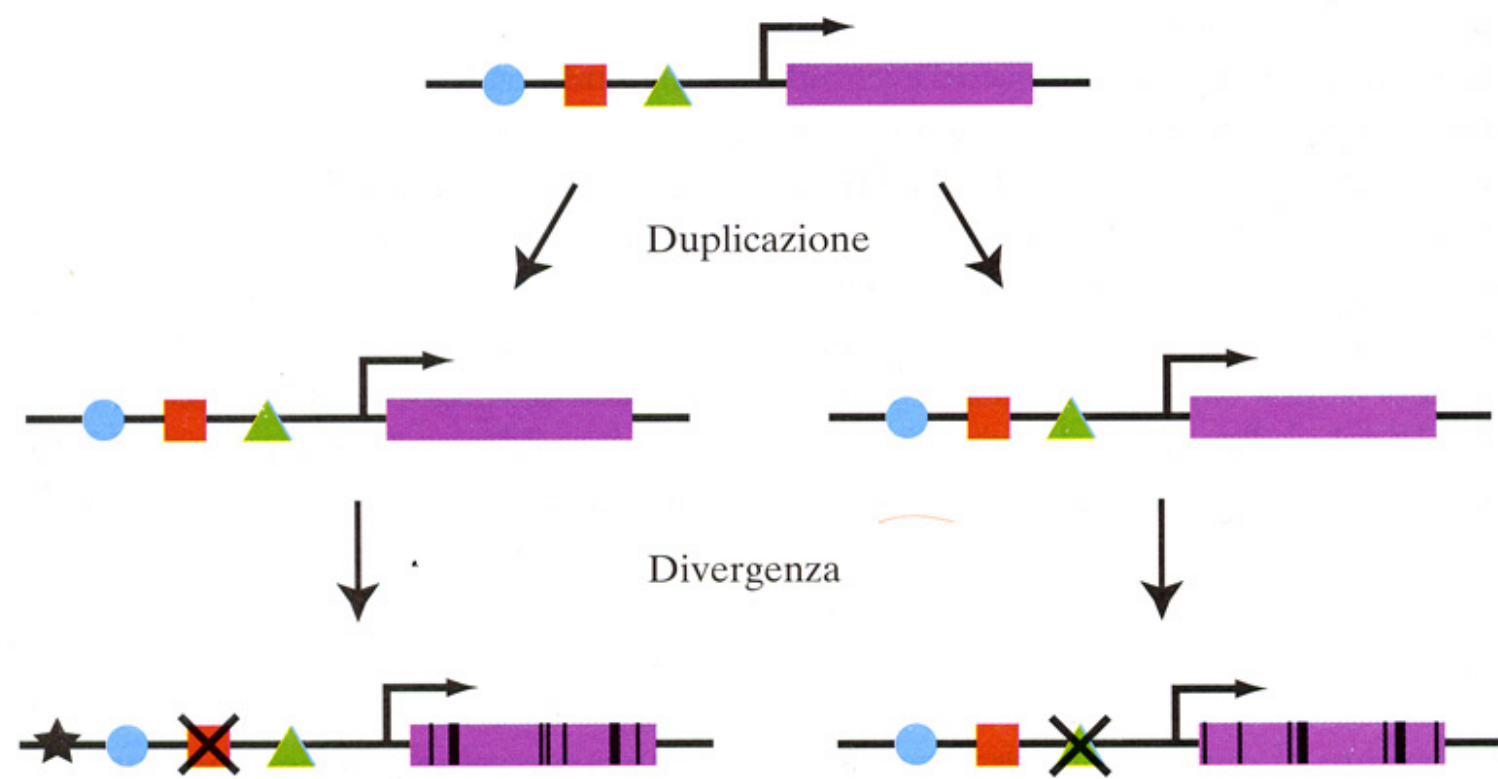




# Duplicazione genica

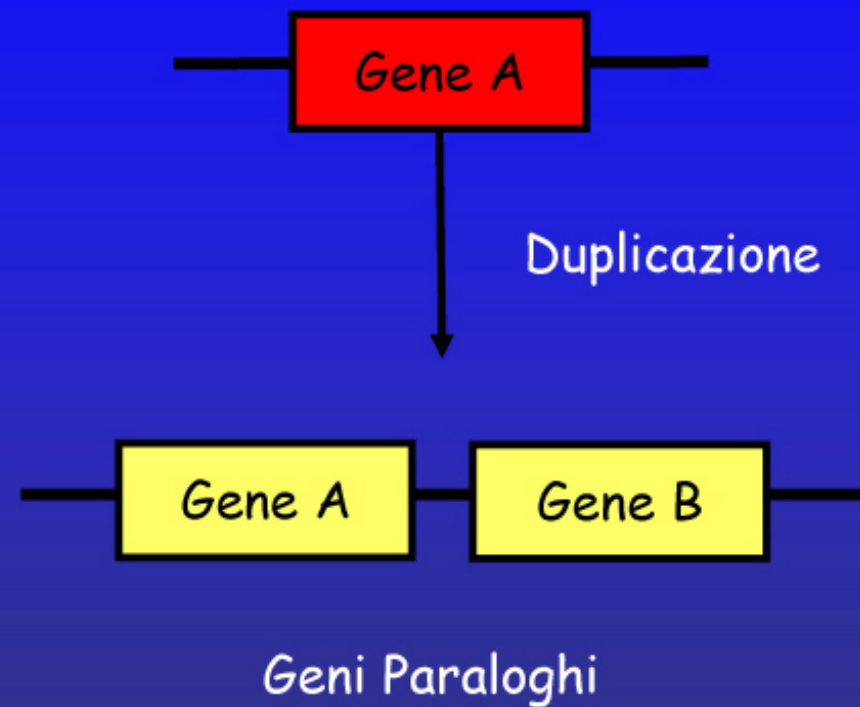
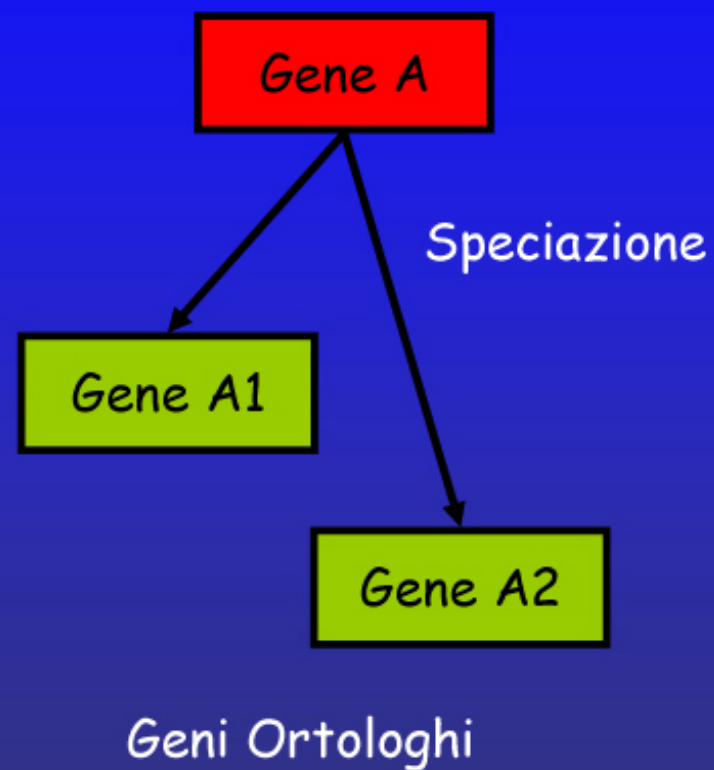






Meccanismi di divergenza genica

# Ortologhi e paraloghi



# Le Famiglie Geniche

- Circa la metà di tutti i geni eucariotici codificanti proteine è presente in singola copia nel genoma. Gli altri geni sono presenti in copie multiple.
- Nel corso dell'evoluzione, le varie copie di uno stesso gene possono subire mutazioni diverse, dando origine a un gruppo di geni strettamente affini, definiti complessivamente **famiglia genica**.
- Le sequenze di DNA di una famiglia genica di solito sono un po' differenti l'una dall'altra. Fintanto che almeno un membro mantiene la primitiva sequenza di DNA, e quindi codifica la proteina giusta, gli altri possono mutare più o meno estesamente. La presenza di copie «supplementari» di un gene è un'importante fonte di variabilità per l'evoluzione a livello molecolare: quando un gene mutato è utile, può superare la selezione e mantenersi nelle generazioni successive; se invece non è funzionale, la copia non mutata salva la situazione. I geni non funzionali diventano pseudogeni e alla fine si possono estinguere
- Esempi di famiglie geniche sono quelle delle globine e quella dei geni Hox

Esempio di vantaggio selettivo per una duplicazione: l'amilasi.

L'amido è importante per la dieta nelle popolazioni di coltivatori-raccoglitori, mentre non lo è nelle popolazioni di cacciatori. L'amilasi è un enzima contenuto nella saliva che inizia la digestione dell'amido nella bocca. Il numero dei geni è maggiore (circa 10) nelle prime popolazioni.

Esempio di delezione non svantaggiosa: la lattasi.

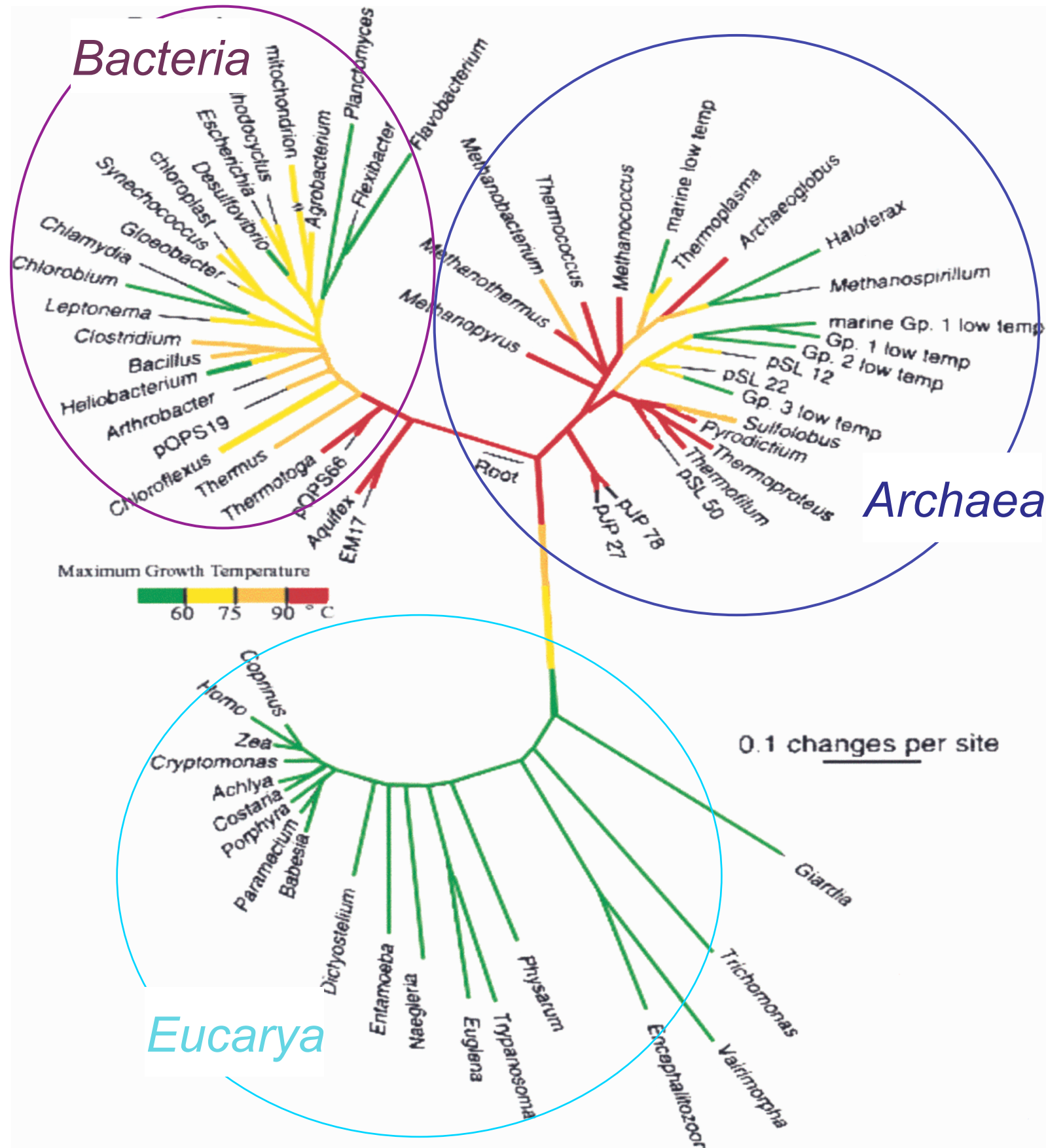
La lattasi idrolizza il lattosio in glucosio e galattosio ed è attiva solo nelle prime fasi della vita. Negli europei una delezione nel primo introne del gene per la lattasi in linkage disequilibrium con tre SNP lo rende attivo per tutta la vita.



# Geni immortali

- Il genoma di un moscerino ha circa il doppio dei geni del lievito di birra
- Il genoma dell'uomo ha circa il doppio dei geni del moscerino
- Il genoma dell'uomo ha circa lo stesso numero di geni del topo
- Il pesce palla ha 7350 geni condivisi con l'uomo
- Le proteine codificate da questi geni sono identiche circa per il 61%
- Ci sono circa 500 geni in comune tra archea, batteri, funghi, piante e animali
- Gli eucarioti esistono da almeno 1,8 miliardi di anni, gli archea e i batteri da più di 2 miliardi di anni

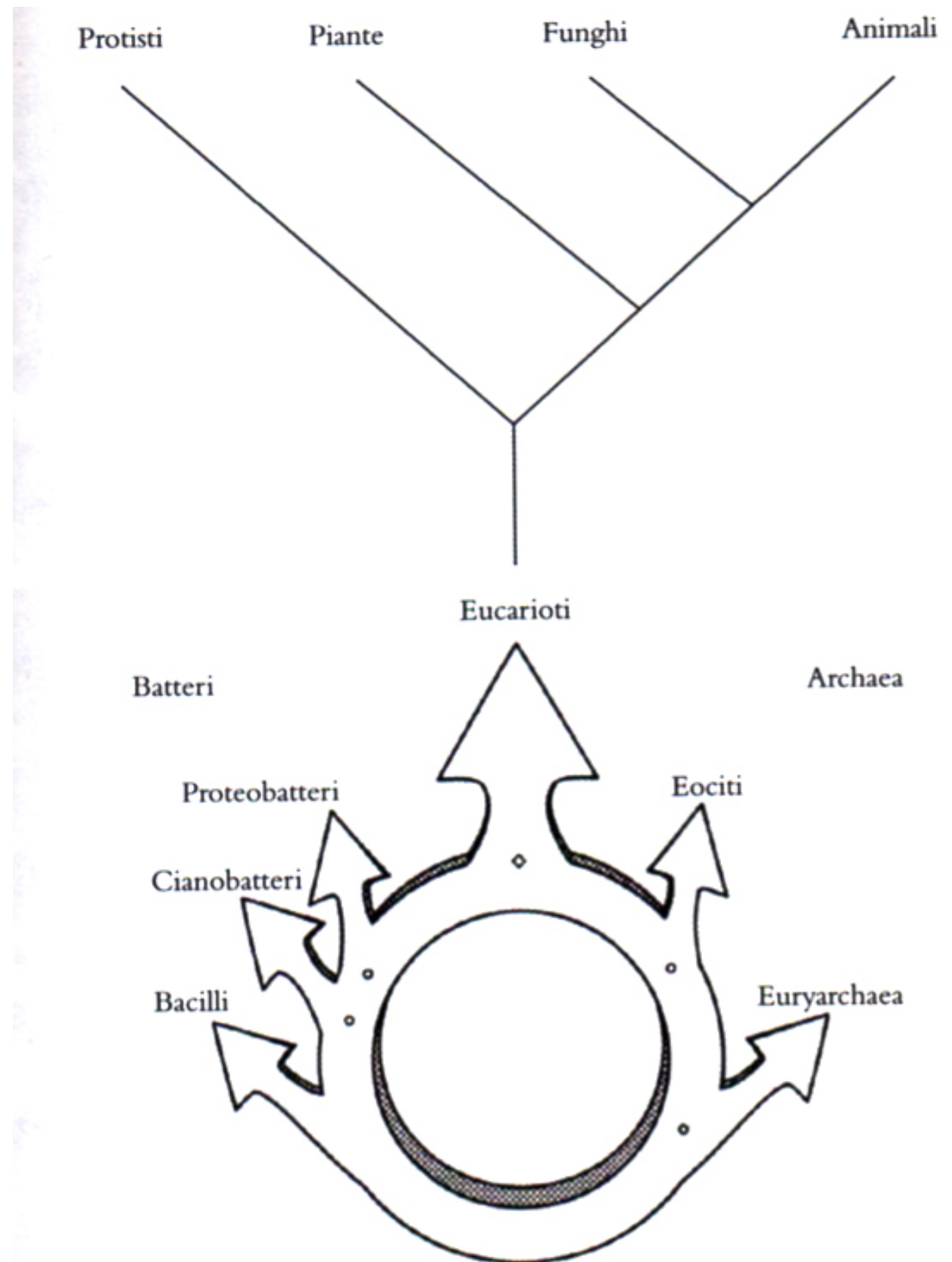
## Bacteria





# Un nuovo albero filogenetico dei viventi

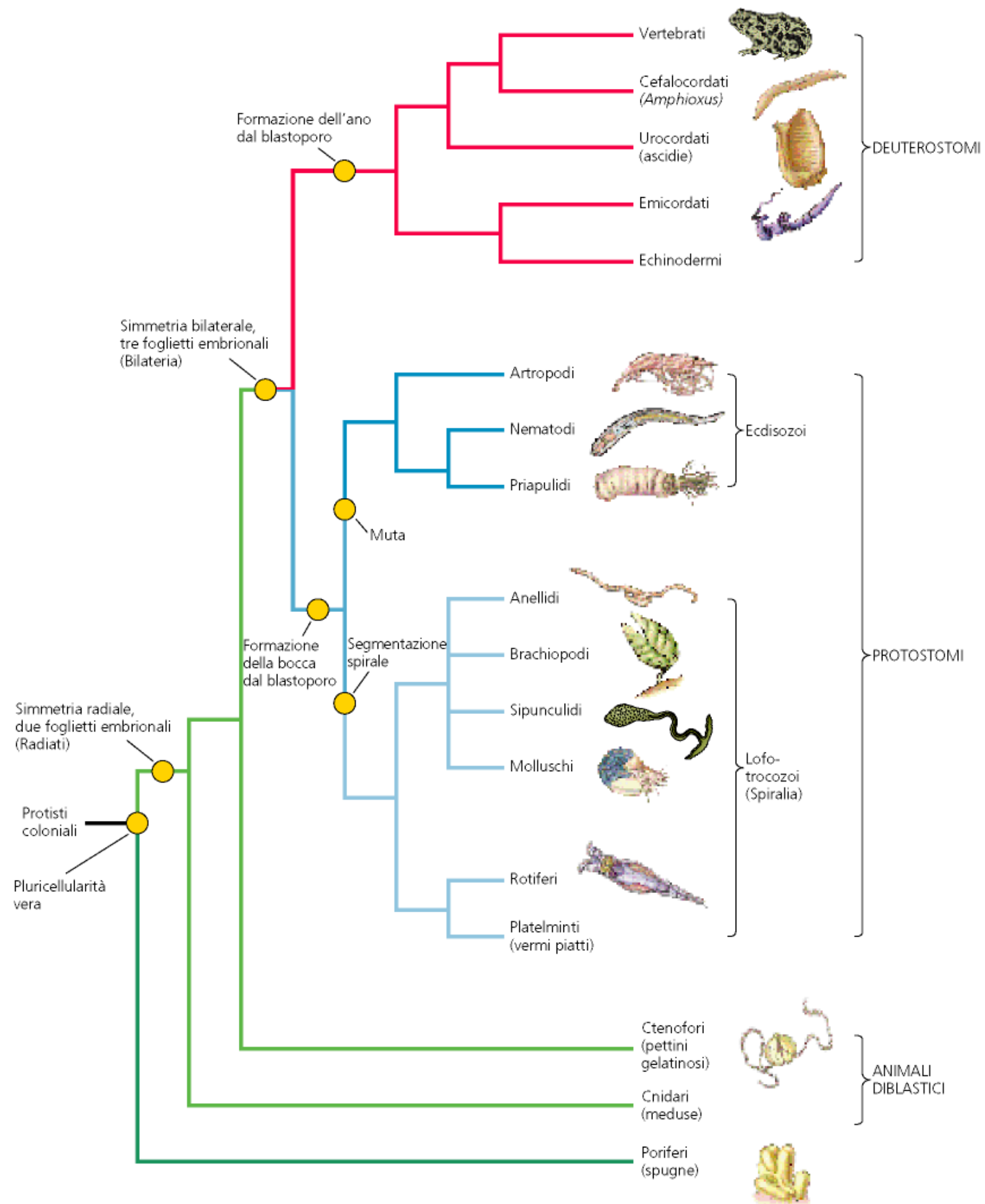
Le sequenze di DNA indicano che un'antica fusione fra qualche tipo di Archaea con qualche tipo di batterio contribuì a dare origine agli eucarioti





# Geni immortali

- La selezione purificatrice mantiene la “purezza” della sequenza di amminoacidi delle proteine eliminando quei cambiamenti che compromettono la loro funzione.
- La selezione purificatrice lascia delle impronte nella maggior parte dei geni, ma è più evidente nei geni immortali che si sono conservati in tutte le divisioni dei viventi
- I geni immortali sono sopravvissuti non perché hanno evitato di subire mutazioni, infatti essi sono vulnerabili come tutti gli altri geni, ma perché sono fondamentali per i processi di base universali in tutte le cellule

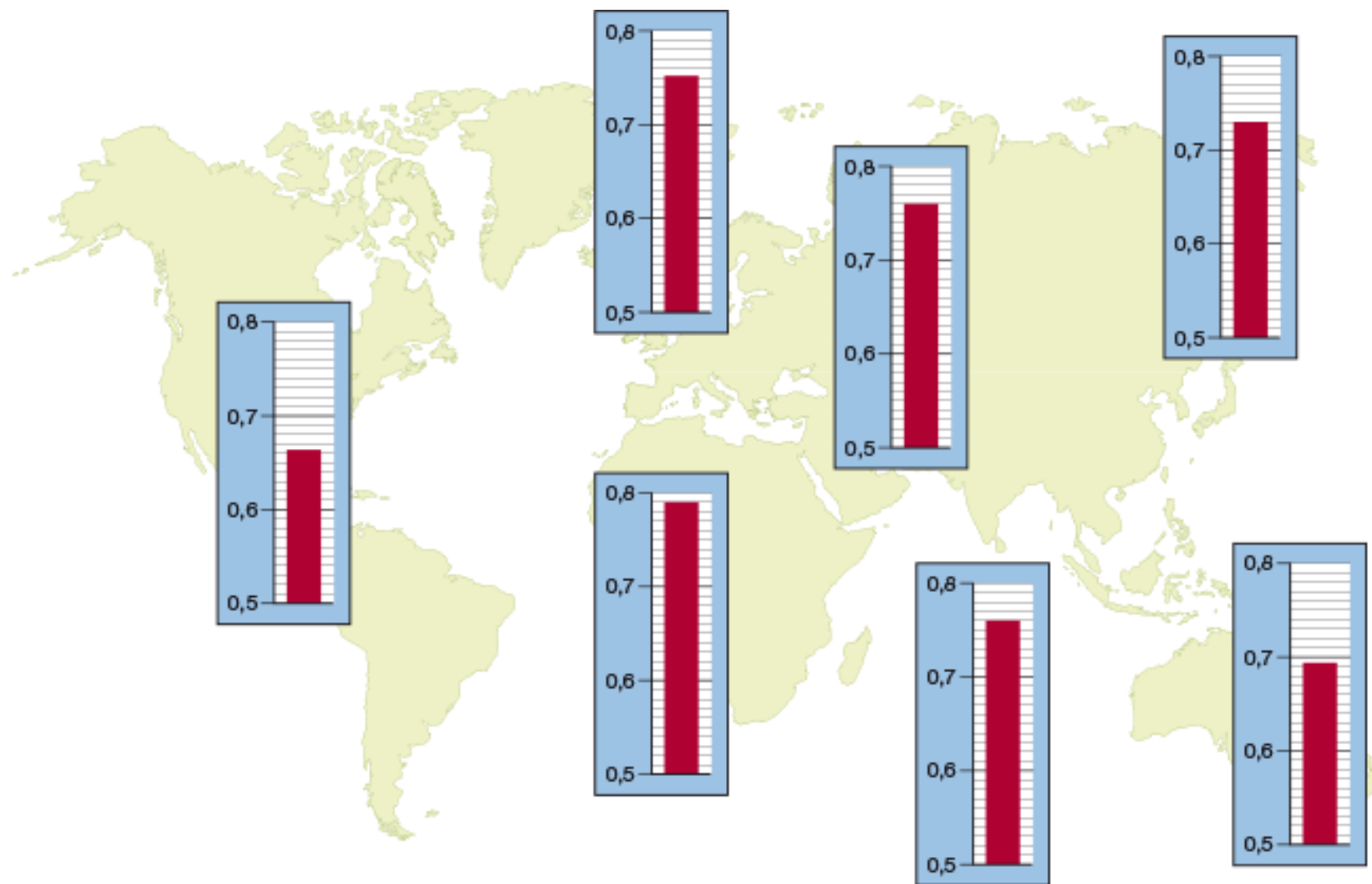


# Origini dell'Homo sapiens

I livelli più elevati di diversità in termini di eterozigosità si trovano nelle popolazioni dell'Africa sub-sahariana che significa maggiore antichità delle popolazioni africane rispetto a quelle degli altri continenti.

I pattern di linkage disequilibrium (LD) tra alleli a siti differenti (aplotipi) sono più bassi in Africa perché la ricombinazione meiotica ha avuto più tempo per separare alleli associati.

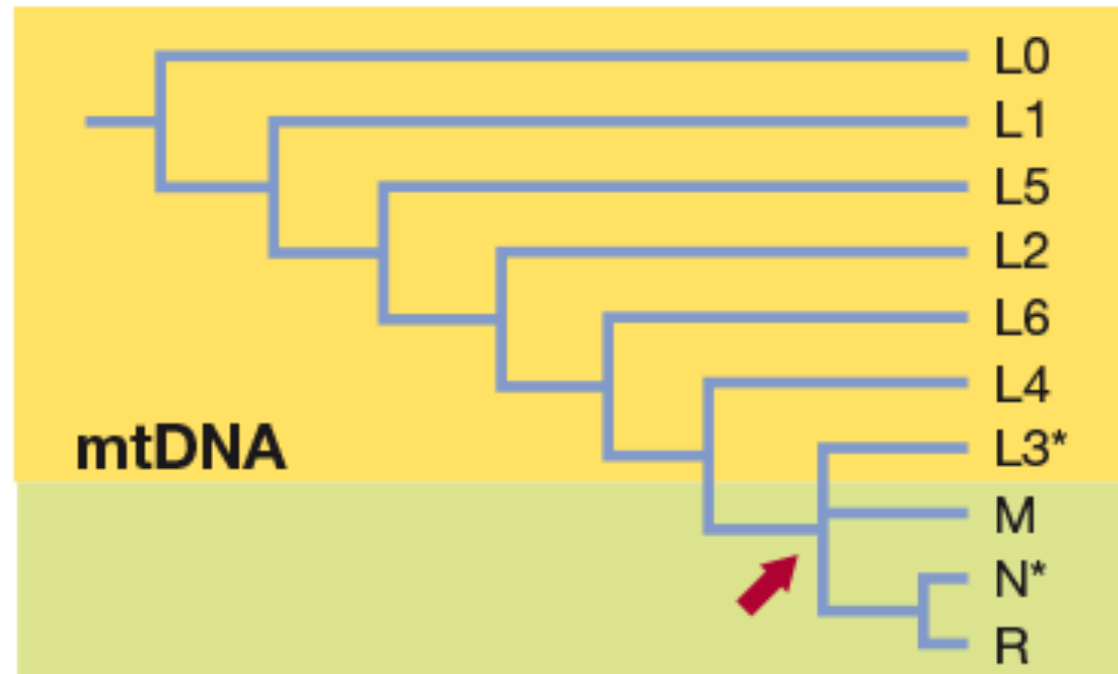
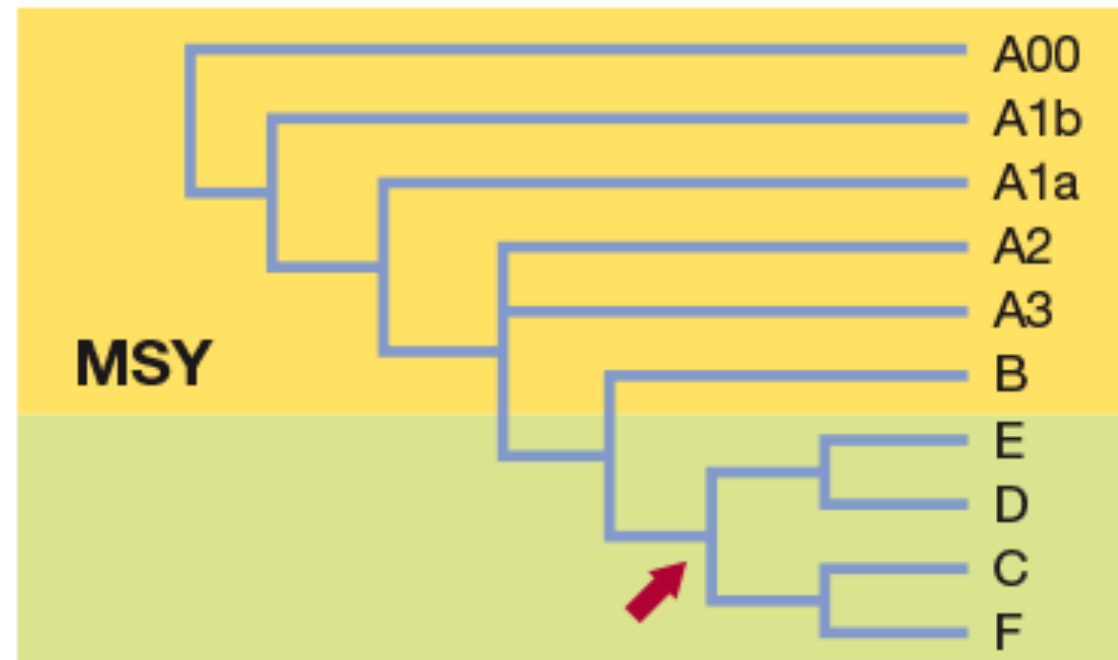
Un'aplotipo è un genotipo aploide per due o più siti polimorfici così vicini in termini di frequenza di ricombinazione da essere quasi sempre ereditati in blocco. Per questo motivo è frequente che gli alleli di questi siti si trovino in **disequilibrio aploide** fra loro (linkage disequilibrium)



DNA mitocondriale e Male-Specific-Y (MSY) sono trasmessi rispettivamente per via materna e paterna.

E' possibile quindi costruire degli alberi filogenetici costituiti da rami che rappresentano le linee di discendenza. I rami terminali rappresentano gli aplotipi oggi esistenti. I diversi rami convergono in nodi che rappresentano le mutazioni e l'insieme dei rami che discende da un nodo prende il nome di clade o aplogruppo.

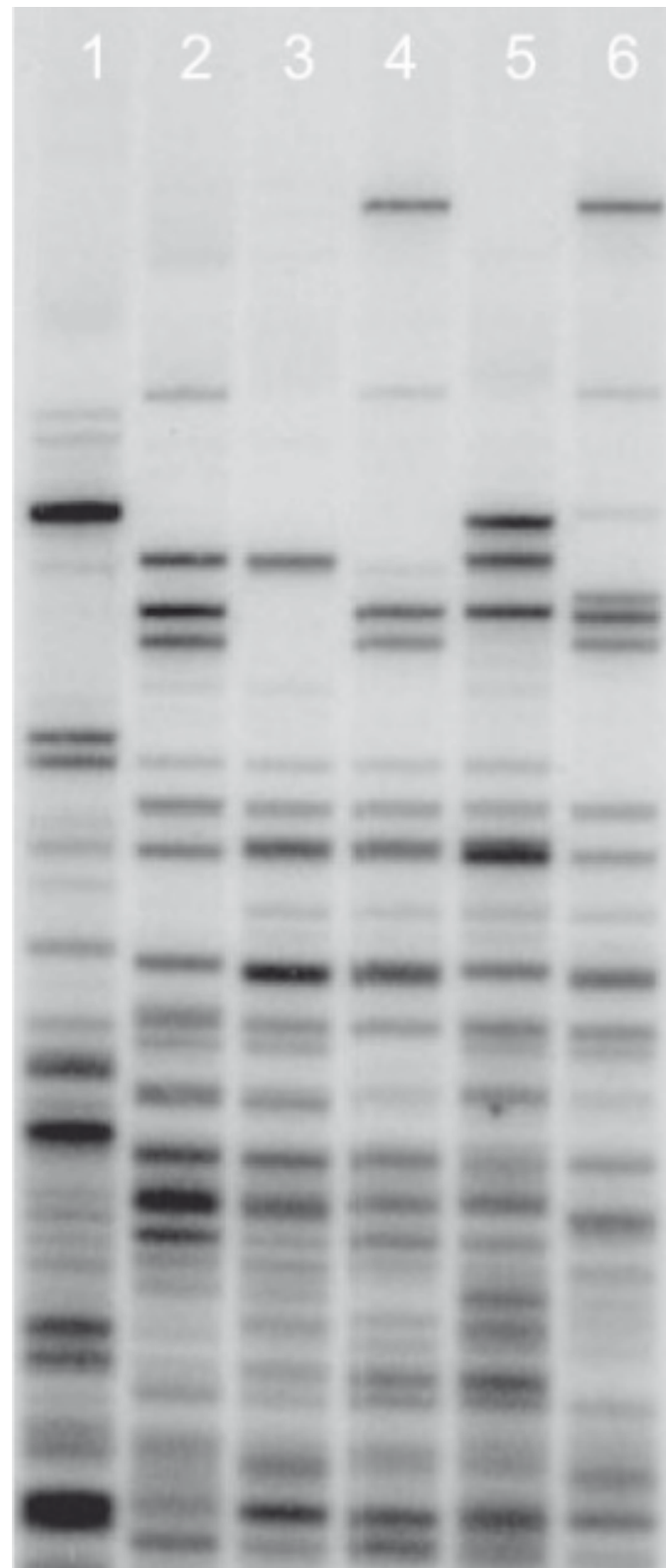
Il nodo più profondo è quello dell'antenato comune.



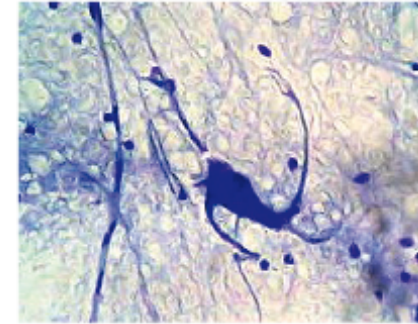
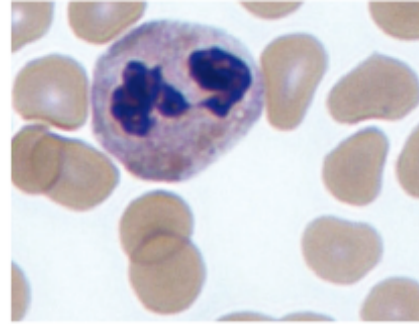
**Tabella E2.1** Lista incompleta di alcuni cataloghi razziali proposti per l'uomo, compresi due cataloghi in uso presso le polizie americana e inglese.

Autore	Numero di razze	Razze o gruppi proposti
Linneo (1735)	6	<i>Europaeus, Asiaticus, Afer, Americanus, Ferus, Monstruosus</i>
Buffon (1749)	6	Europeo, Lappone, Tartaro, Asiatico, Etiopico, Americano
Kant (1775)	4	Bianco, Nero, Unno (o Calmucco, o Mongolo), Industano
Blumenbach (1795)	5	Caucasico, Mongoliaco, Etiopico, Americano, Malese
Cuvier (1828)	3	Caucasoide, Negroide, Mongoloide
Deniker (1900)	29	
Museum of Natural History, Chicago (1933)	109	
Von Eickstedt (1937)	38	
Biasutti (1959)	53	
Coon (1962)	5	Caucasoide, Mongoloide, Capoide, Congoide, Australoide
US Office of Management and Budget (1997)	6	American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, White
Metropolitan Police Service, London (2005)	16	W1 White British, W2 White Irish, W9 Other White background, M1 Mixed White and Black Caribbean, M2 Mixed White and Black African, M3 Mixed White and Asian, M9 Other Mixed background, A1 Asian Indian, A2 Asian Pakistani, A3 Asian Bangladeshi, A9 Any other Asian background, B1 Black Caribbean, B2 Black African, B9 Other Black background, O1 Chinese, O9 Other Ethnic group





Due cellule dello stesso individuo  
(leucocita e neurone)



0/1000

Due gemelli identici



0/1000

Due di noi a caso



1/1000

Uno di noi e uno scimpanzè



10/1000

Uno di noi e un ciclamino



650/1000







# Le tracce della storia evolutiva nel patrimonio genetico

- Nella storia evolutiva della specie umana, il Paleolitico o Età della Pietra antica, occupa gran parte degli ultimi 2 milioni di anni. In questo lungo periodo di tempo le caratteristiche dell'uomo moderno si sono formate nell'interazione tra il genoma e l'ambiente circostante.

# Le tracce della storia evolutiva nel patrimonio genetico

- I grandi cambiamenti ambientali derivanti dall'introduzione 10000 anni fa (Neolitico) dell'agricoltura e dell'allevamento del bestiame, che influenzarono la dieta e lo stile di vita dei nostri progenitori, sono avvenuti troppo recentemente per la capacità di adattamento del genoma umano, considerando i tempi evolutivi.

# Le tracce della storia evolutiva nel patrimonio genetico

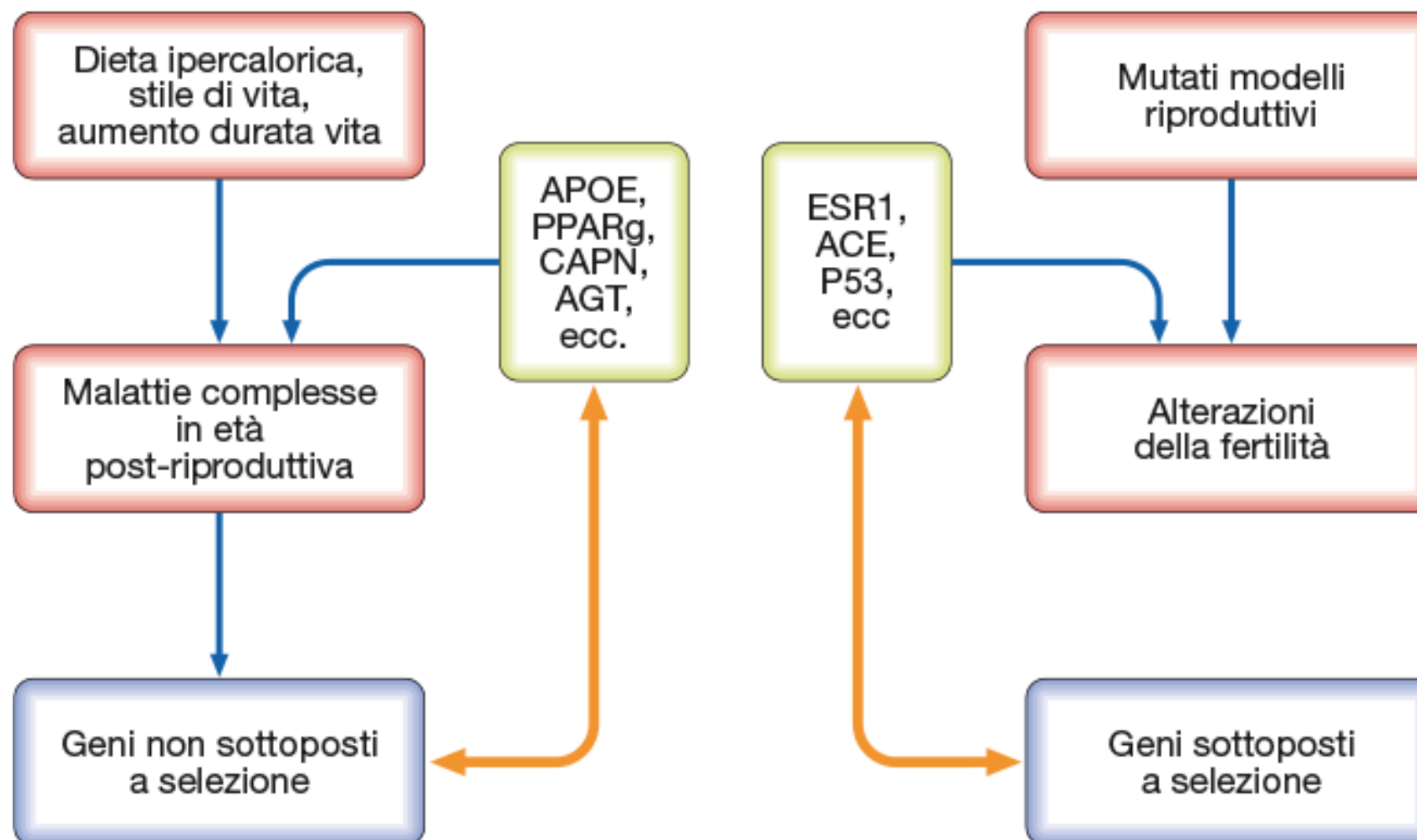
- Nel complesso, quindi, il nostro genoma è quello dell'uomo dell'età della pietra, adattatosi a una dieta e a un livello di esercizio fisico di uno stile di vita pre-rivoluzione agricola.
- L'uomo del Paleolitico era adattato a una dieta varia e flessibile con alternanza di periodi di carestia e di abbondanza. Per massimizzare la capacità di sopravvivenza, l'organismo era adattato a una dieta minimale e a un intenso esercizio fisico.

# Le tracce della storia evolutiva nel patrimonio genetico

- Il primo a proporre una spiegazione evolucionistica per l'aumentata frequenza di malattie complesse fu James Neel nel 1962. Per spiegare l'aumento del diabete nelle popolazioni occidentali, egli suggerì che ci fossero dei geni detti *thrifty*, componenti normali del genoma umano, selezionati nel passato per utilizzare più efficientemente il cibo, ma resi dannosi da progresso.



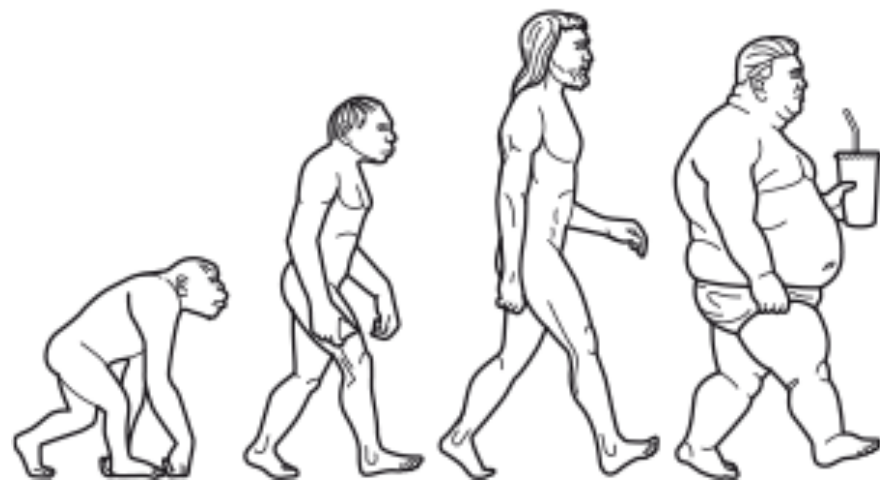
Esempi di geni vecchi in un mondo nuovo: malattie da civilizzazione



**Tabella E1.1** Le più comuni malattie complesse da alterato stile di vita, i fattori ambientali alterati e alcuni dei geni coinvolti.

Malattie	Fattore ambientale	Geni
Diabete di tipo 2	Dieta	CAPN10 PPARG
Obesità	Dieta, esercizio fisico	ADRB2, ADRB3
Ipertensione	Dieta, esercizio fisico, durata della vita	AGT, CYP3AC, ACE
Malattie cardiovascolari	Dieta, esercizio fisico, durata della vita	APOE
Malattia di Alzheimer	Durata della vita	APOE
Subfertilità/infertilità	Modalità riproduttive, posticipazione della riproduzione	ESR1, ACE, TP53

**(a)**



**(b)**

