

La corsa al sequenziamento dei genomi batterici

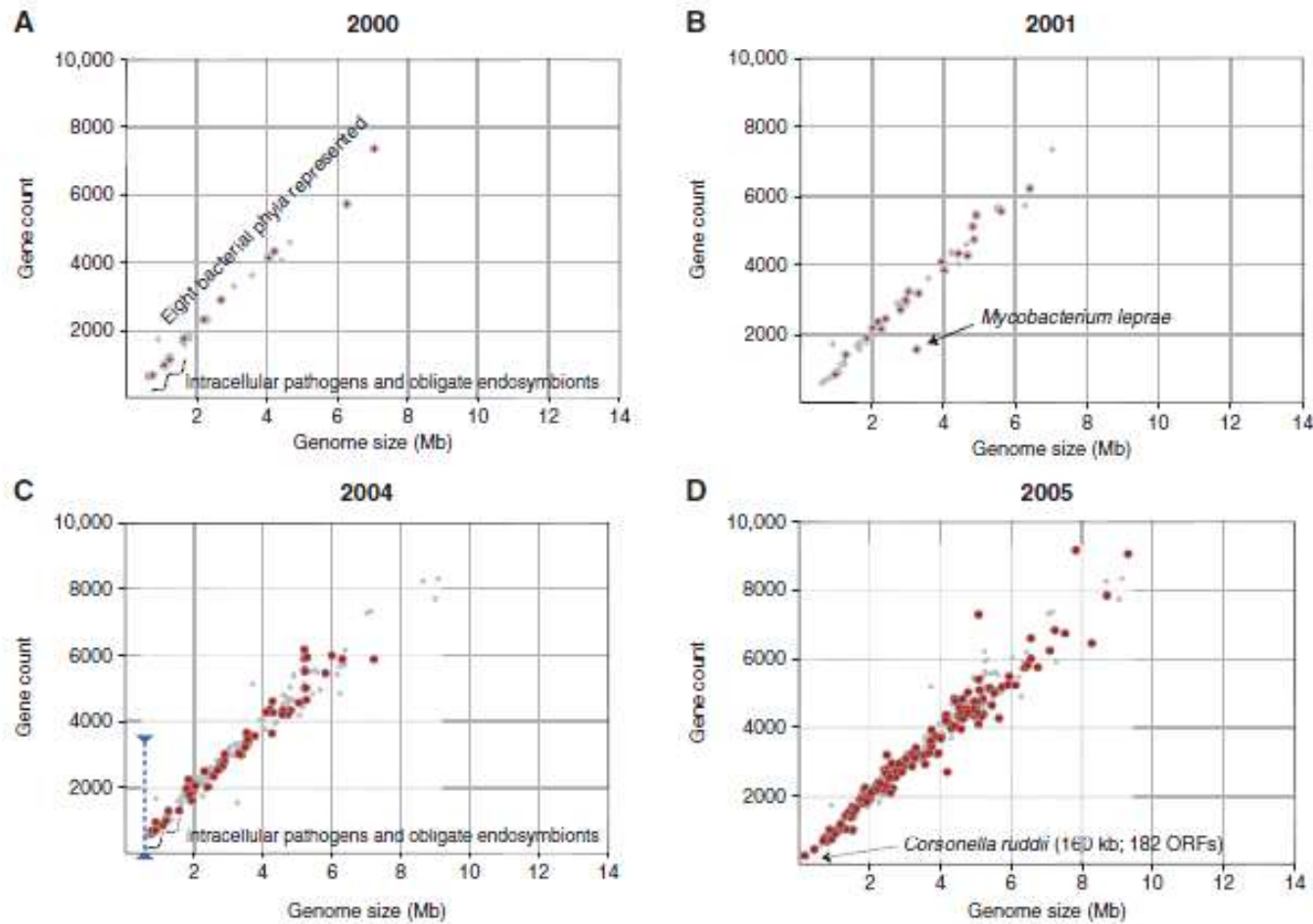
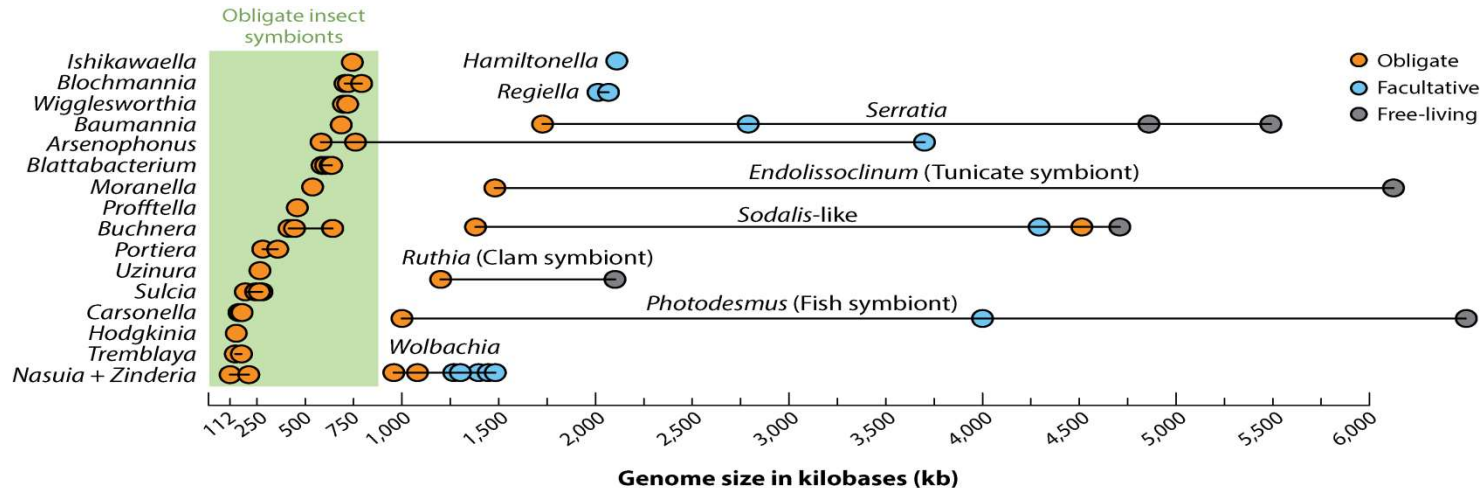


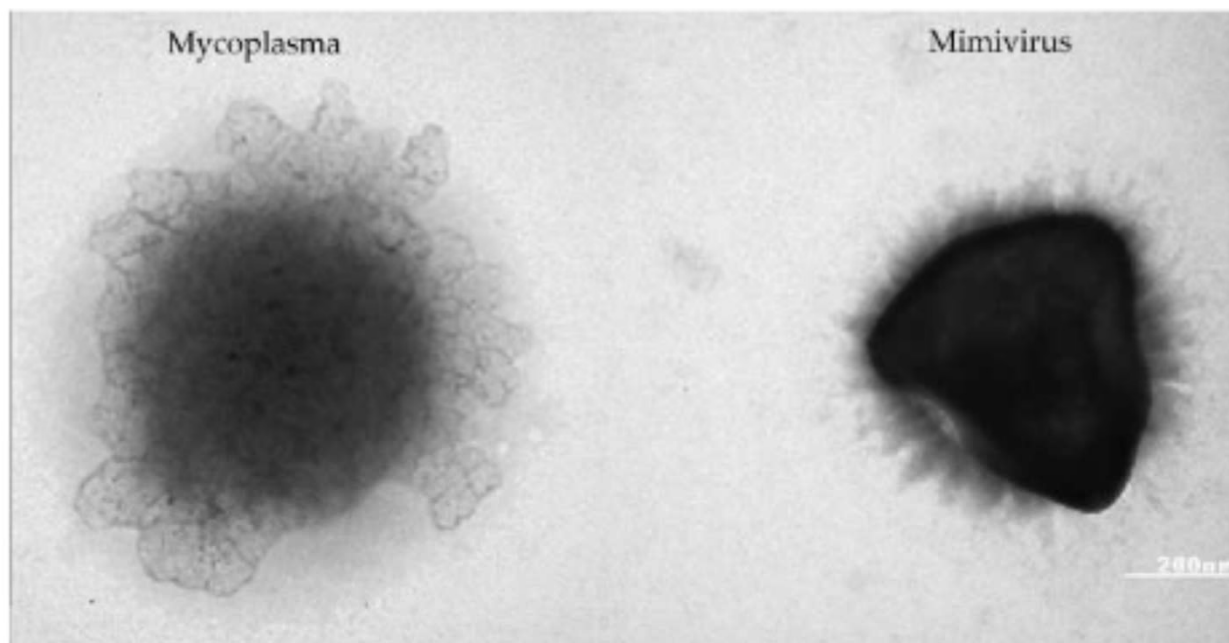
Figure 2. Genome size and gene count in bacterial genomes sequenced from 2001 to 2005. Red dots indicate genomes that were published in the designated year and smaller gray dots represent genomes published in all prior years. Panels A–D show results for the year indicated. ORE, Open reading frame.

Genomi al di sotto della cellula minima!! 112 Kb



AR Moran NA, Bennett GM. 2014.
Annu. Rev. Microbiol. 68:195–215

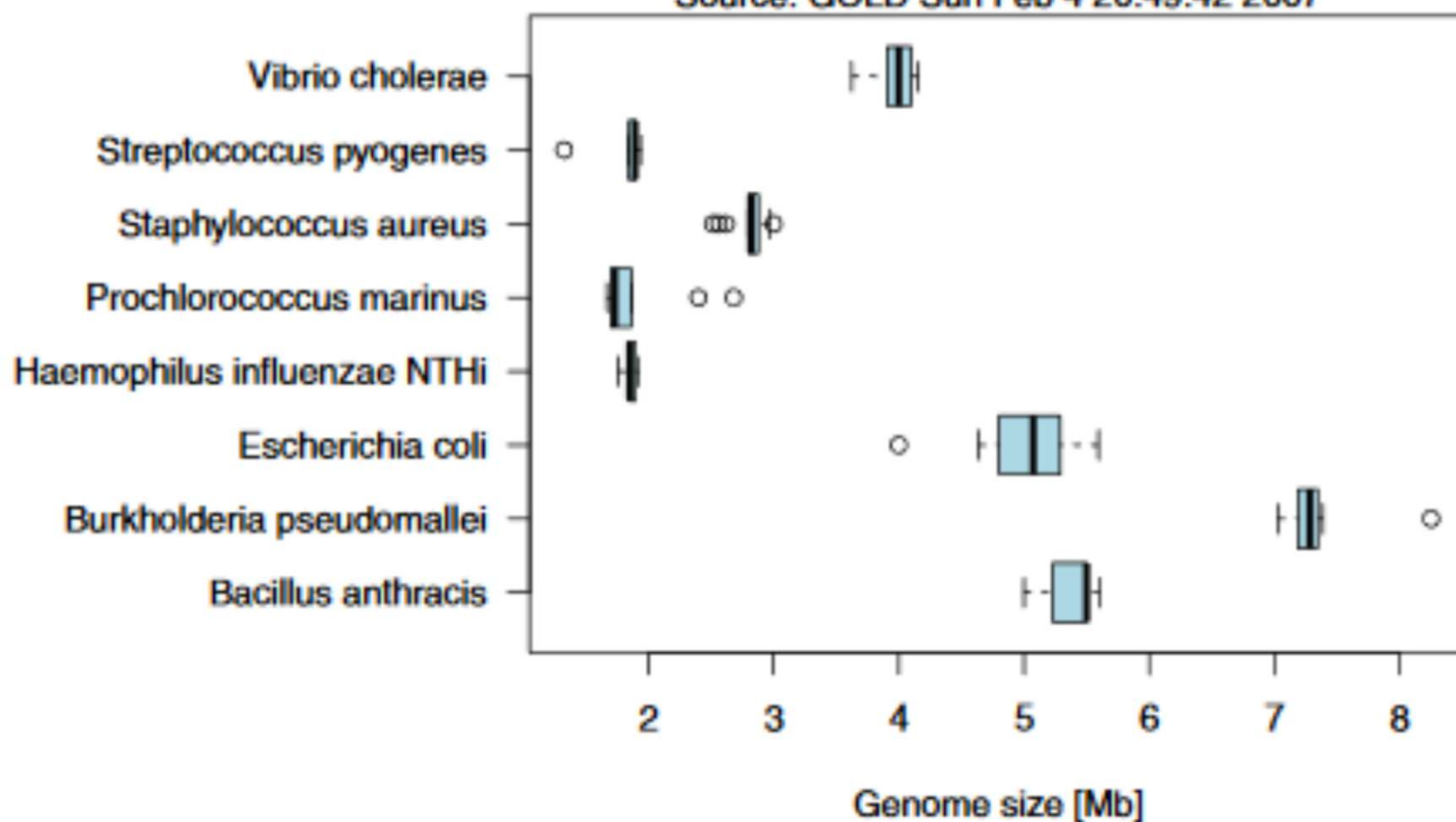
Despite their size, each retains some genes that enable provisioning of limiting nutrients or other capabilities required by hosts. Genome sequence analyses show that genome reduction is an ongoing process, resulting in a continuum of sizes, with the smallest genome currently known at **112 kilobases**. Genome reduction is typical in host-restricted symbionts and pathogens, but the tiniest genomes are restricted to symbionts required by hosts and restricted to specialized host cells, resulting from long coevolution with hosts. Genes are lost in all functional categories, **but core genes for central informational processes, including genes encoding ribosomal proteins, are mostly retained, whereas genes underlying production of cell envelope components are especially depleted**. Thus, these entities retain **cell-like properties but are heavily dependent on coadaptation of hosts**, which continuously evolve to support the symbionts upon which they depend.



Electronic microscopy of a "bacteria" on the left (*Ureaplasma urealyticum (parvum)*) with a genome size of 0.751 Mb and mimivirus on the right with a genome size of 1.181 Mb. Credit: the Mimivirus picture gallery from <http://giantvirus.org/>. Copyright: Prof. Didier Raoult, Rickettsia Laboratory, La Timone, Marseille, France.

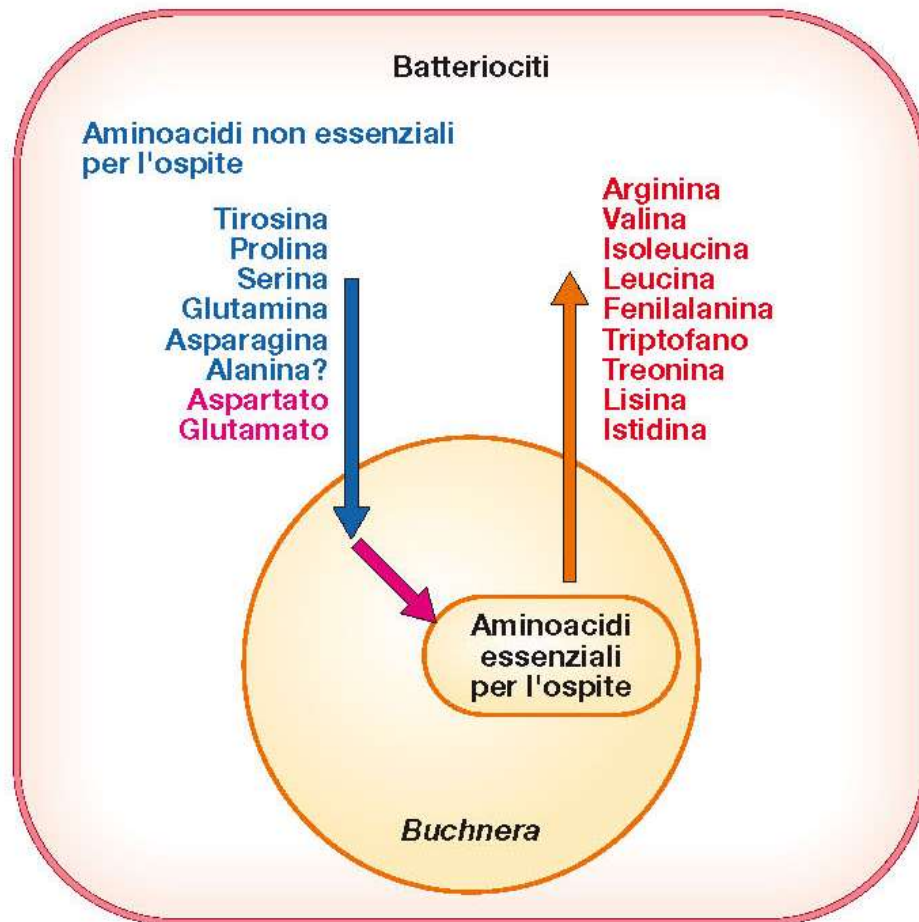
Genome size within species with at least 10 strains

Source: GOLD Sun Feb 4 20:49:42 2007



Genomi di piccole dimensioni e adattamento all'ambiente: parassitismo obbligato

Il caso di *Buchnera*, un batterio con omologie con a *E. coli*, parassita obbligato degli afidi. Ne gli afidi ne il batterio possono vivere senza uno dell'altro : non si possono trattare gli afidi con antibiotici che uccidono il batterio ne si possono coltivare in provetta i batteri senza afidi !!



Batterio → Afide

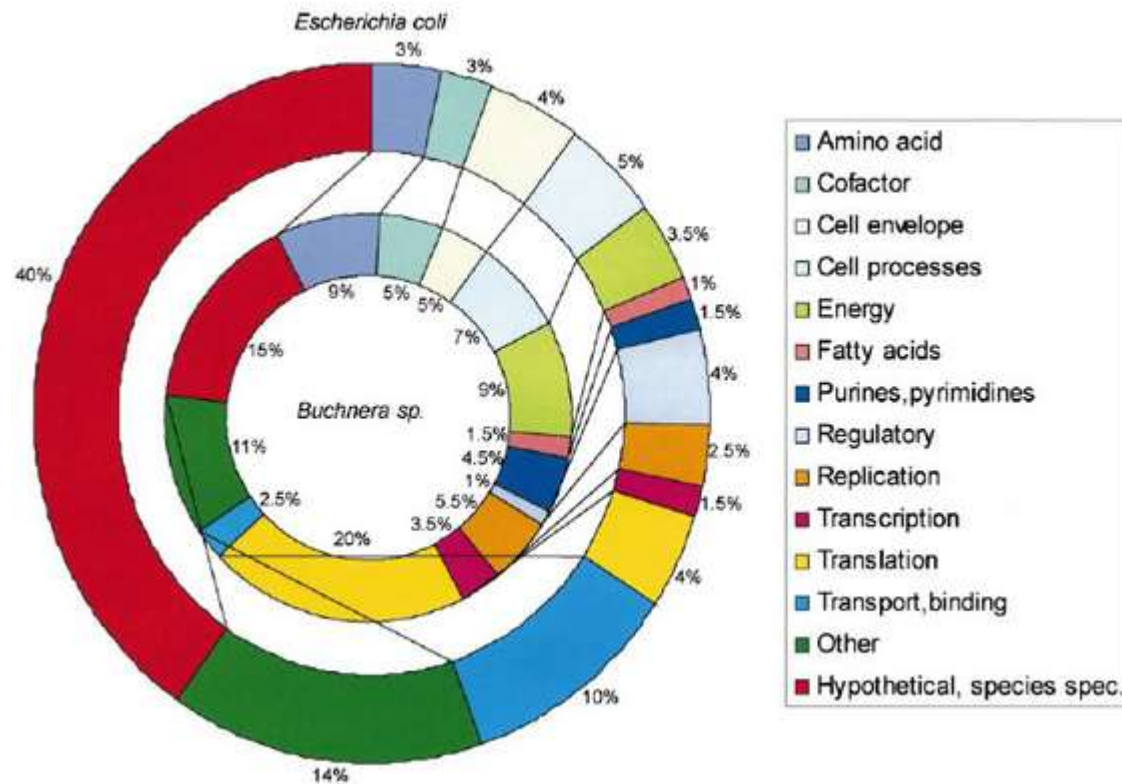
Aminoacidi essenziali che l'insetto non trova nella pianta di cui si nutre ,

Batterio ← Afide

Nutrienti ed energia per il metabolismo

Gli afidi si nutrono della linfa delle piante un liquido povero di proteine e richiedono 10 AA che non sanno sintetizzare . I batteri sintetizzano alcuni AA ma utilizzano i metaboliti dell'ospite per dare l'avvio alla sintesi

Buchnera vive all'interno di cellule degli afidi dette batteriociti.



La riduzione del genoma di *Buchnera* è considerevole 450Kb -670 Kb

In *Buchnera* molti geni coinvolti nella sintesi di aminoacidi sono portati da plasmidi multicopie che ne aumentano il dosaggio genico

L'adattamento di *Buchnera* alla vita intracellulare ha portato ad una riduzione del suo genoma con perdita dei geni per la sintesi degli aminoacidi forniti dall'ospite e conservazione di quei geni per la biosintesi degli aminoacidi essenziali.

In *Buchnera burgdoferi* sono stati trovati plasmid che contengono pseudogeni

Riduzione dei genomi

Molti dei geni presenti in *Buchnera* hanno un'elevata omologia con i geni di enterobatteri si pensa quindi che possano essere derivati per genome reduction.

I geni che codificano per la biosintesi degli aminoacidi essenziali sono stati trovati disposti in tandem con uno dei due geni inattivato da mutazioni.

Il genoma di *Buchnera* di tre isolati diversi in tre specie diverse di afidi varia da 670 a 40 kb .

Il più piccolo codifica solo per 396 proteine contro i 470 geni di *M.genitalium*

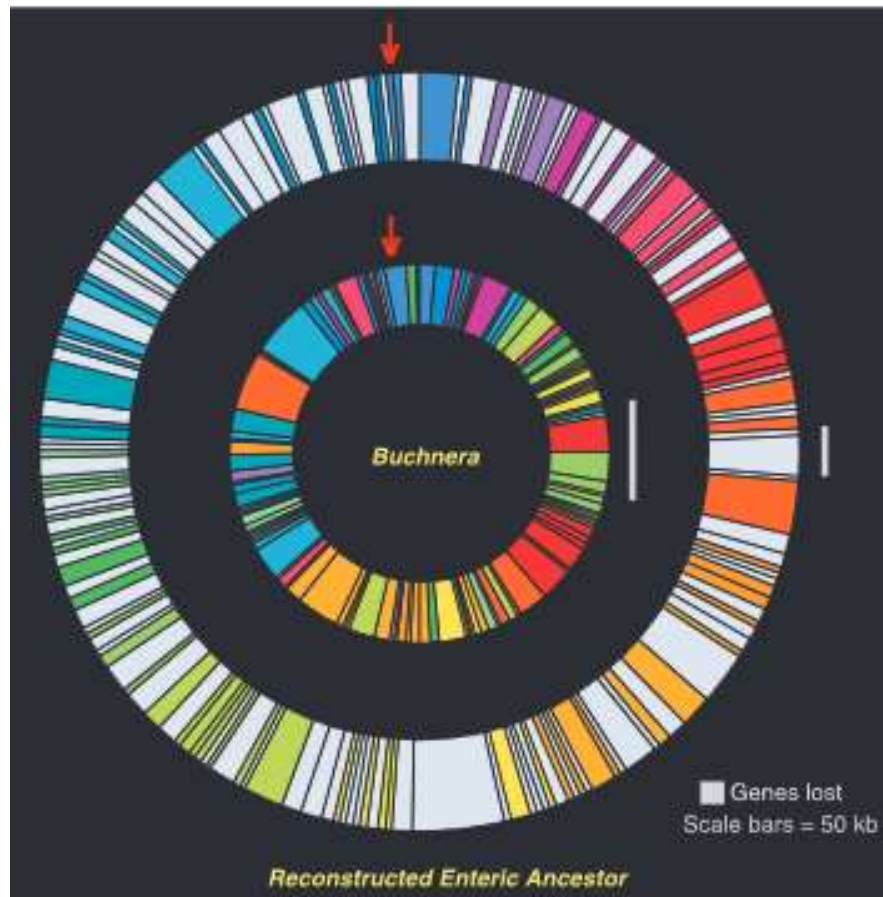
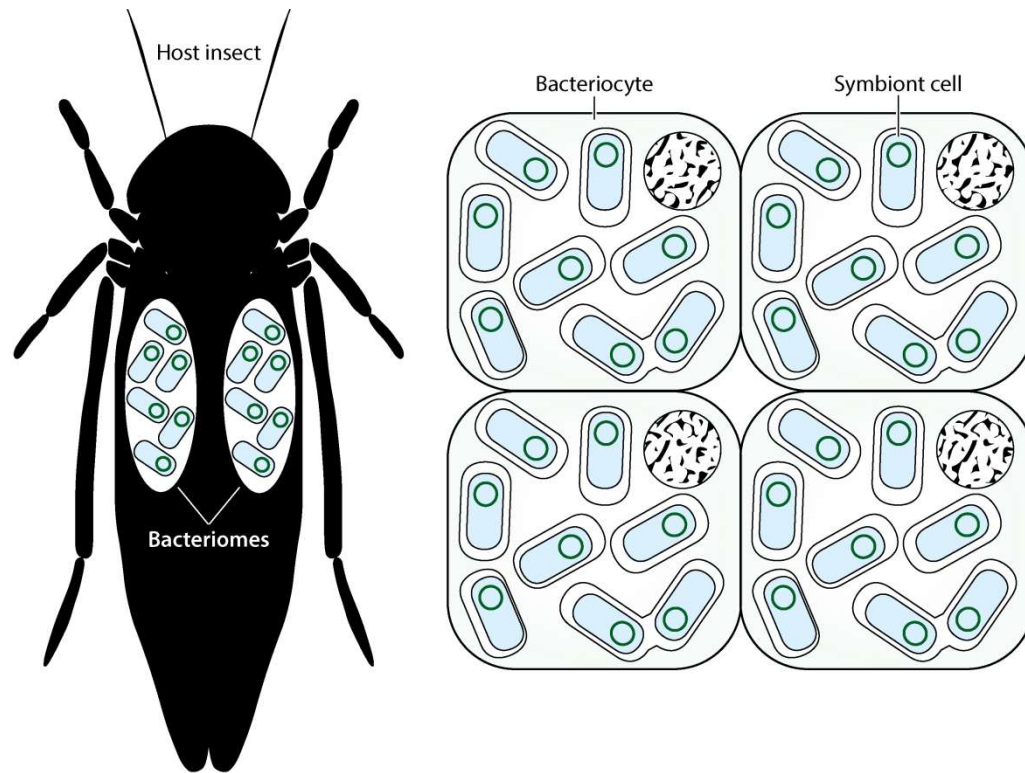


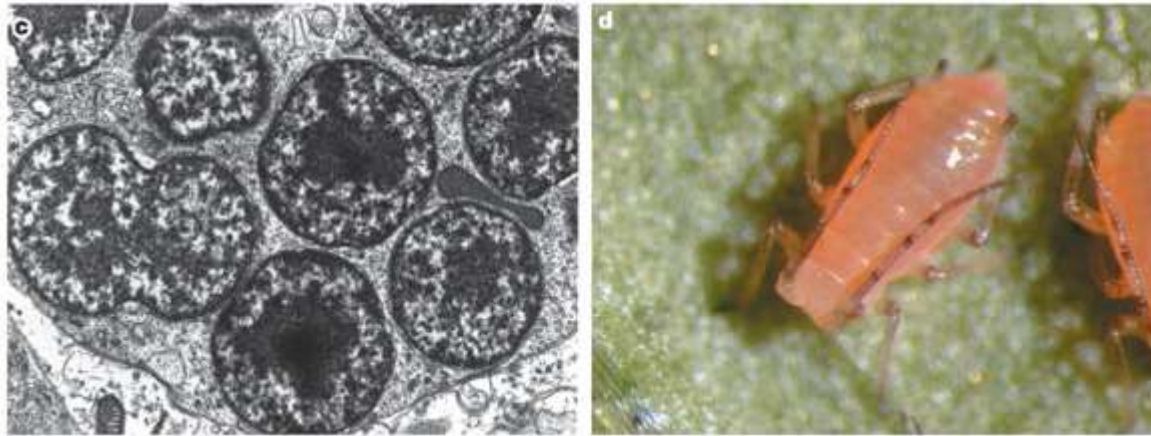
Fig. 1. Genome reduction in *Buchnera*-APS, the bacterial endosymbiont of aphids. *Buchnera* has undergone massive genome reduction (to 0.64 Mb) [34]. Virtually all of its 590 genes have close homologs in the genomes of the enteric bacteria, which, along with phylogenetic evidence, indicates that its genome was derived from a much larger genome resembling modern enteric bacteria such as *E. coli*. In this depiction, the outer ring represents the hypothetical ancestral genome for enteric bacteria, produced by removing horizontally acquired regions [identified on the basis of phylogenetic distribution among enteric bacteria (51)] from the genome of *E. coli* MG1655. Gray bands denote ancestral sections of the genome that have been eliminated during the evolution of the *Buchnera* lineage. Colored bands represent regions within which ancestral gene arrangements persist in *Buchnera*, although many individual genes within these regions have been lost. *Buchnera* retains 21% of ancestral genes. The *Buchnera* genome is drawn at a larger scale than the ancestor (scale bars beside each genome are 50 kb), and the bands are colored to match corresponding ancestral regions. For each genome, the arrow indicates origin of replication.



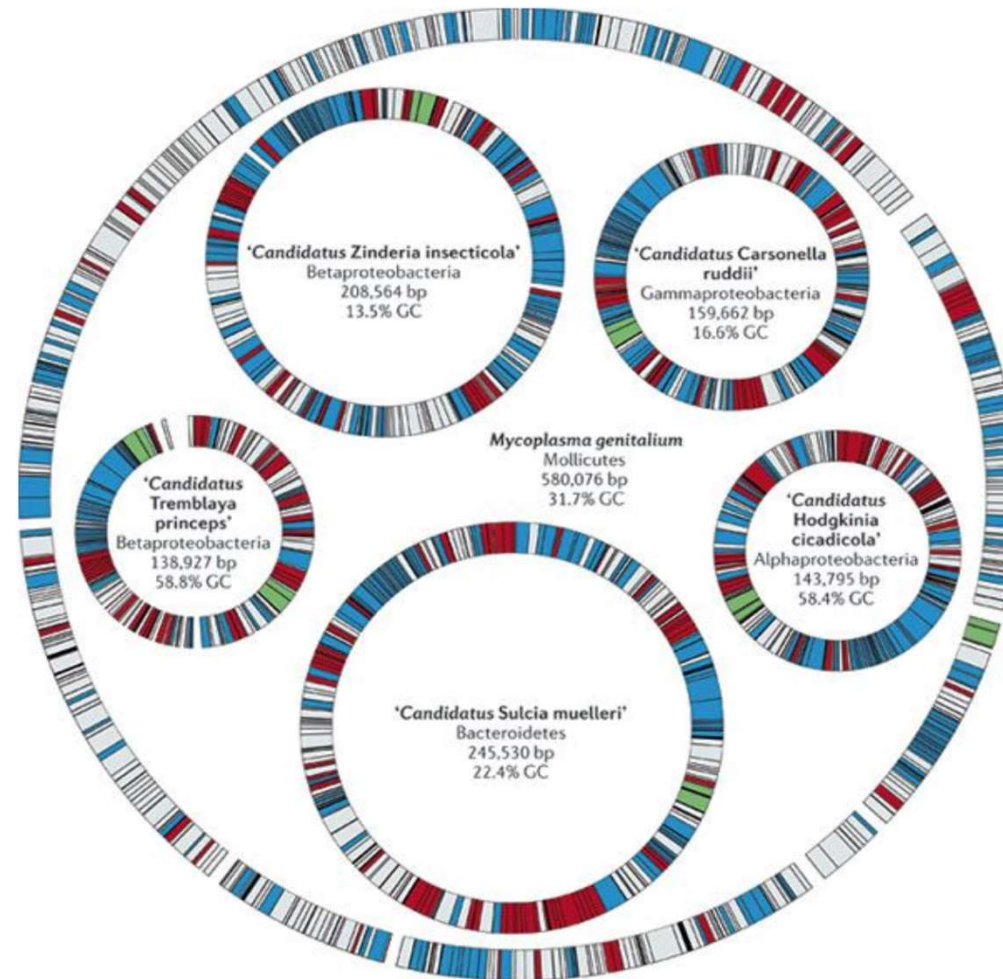
Ma dove vivono i batteri con i genomi super minimi??

AR Moran NA, Bennett GM. 2014.
Annu. Rev. Microbiol. 68:195–215

Diagram showing location of obligate symbionts within bacteriocytes in an insect. Individual symbiont cells are typically surrounded by a host-derived membrane within the bacteriocyte cytosol. Bacteriocytes are often clustered into a bacteriome, usually located in the insect abdomen



P-endosymbionts of insects are characterized by their occurrence in specialized host cells called **bacteriocytes**, which are located at various positions in the insect body depending on the host group¹. Bacteriocytes are sometimes grouped into organ-like structures called bacteriomes that occur in the body cavity of aphids and in the anterior gut region of tsetse flies.



Nature Reviews | Microbiology

The genome of *Mycoplasma genitalium*, the free-living organism with the smallest genome, is two to four times as large as the genomes of five symbionts recently shown to have tiny genomes (that is, smaller than 300 kb): '*Candidatus Sulcia muelleri*', '*Candidatus Zinderia insecticola*', '*Candidatus Carsonella ruddii*', '*Candidatus Hodgkinia cicadicola*' and '*Candidatus Tremblaya princeps*'. Genes involved in informational processes are in blue, those involved in vitamin or amino acid biosynthesis are in maroon, ribosomal RNA genes are in green, other genes are light grey and breaks are non-coding regions.

Ma quali geni hanno conservato i minigenomi ??

Gene	Product	Present in ' <i>Candidatus Tremblaya princeps</i> '?
Replication		
<i>dnaE</i>	DNAP III α -subunit	Yes
<i>dnaQ</i>	DNAP III ϵ -subunit	Yes
Transcription		
<i>rpoA</i>	RNAP α -subunit	Yes
<i>rpoB</i>	RNAP β -subunit	Yes
<i>rpoC</i>	RNAP β' -subunit	Yes
<i>rpoD</i>	RNAP factor σ^{70}	No (gene present is a pseudogene)
Protein folding or stability		
<i>groL</i>	GroEL (chaperone Hsp60 family member)	Yes
<i>groS</i>	GroES (chaperone Hsp60 regulator)	Yes
<i>dnaK</i>	The main component of the chaperone Hsp70	Yes
tRNA modification		
<i>mnmA</i>	tRNA-specific 2-thiouridylase	No (gene present is a pseudogene)
<i>mnmE</i>	A GTP-binding protein with a role in tRNA modification	No
<i>mnmG</i>	A protein involved in tRNA modification	No (gene present is a pseudogene)
Sulphur metabolism		
<i>sufS</i> or <i>iscS</i>	Cysteine desulfurase	Yes (<i>iscS</i>)
<i>sufBC</i> or <i>iscAU</i>	Cysteine desulfurase accessory proteins	Yes (<i>iscU</i> , but not <i>iscA</i>)
RNA modification		
<i>rlu</i> genes	Ribosomal large-subunit pseudo-uridine synthase genes	No

Translation

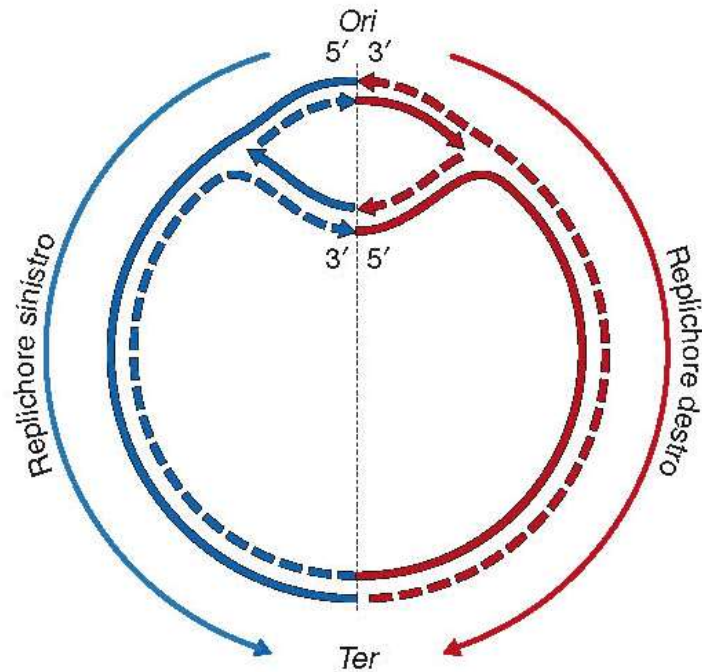
<i>infA</i>	IF-1	Yes
<i>infB</i>	IF-2	Yes
<i>infC</i>	IF-3	Yes
<i>fusA</i>	EF-G	Yes
<i>tsf</i>	EF-Ts	No
<i>prfA</i>	RF-1	No
<i>prfB</i>	RF-2 [†]	No
<i>frr</i>	Ribosome-recycling factor	No
<i>def</i>	Peptide deformylase	No
<i>alaS</i>	Alanyl-tRNA synthetase	No
<i>gltX</i>	Glutamyl-tRNA synthetase	No
<i>glyQ</i>	Glycyl-tRNA synthetase α -subunit	No
<i>ileS</i>	Isoleucyl-tRNA synthetase	No
<i>metG</i>	Methionyl-tRNA synthetase	No
<i>pheS</i>	Phenylalanyl-tRNA synthetase, α -subunit	No
<i>trpS</i>	Tryptophanyl-tRNA synthetase	No
<i>valS</i>	Valyl-tRNA synthetase	No
<i>rps</i> genes	30S ribosomal subunit proteins S1, S2, S3, S4, S5, S7, S8, S9, S10, S11, S12, S13, S14, S16, S17, S18 and S19 (17 of the 21 possible subunits)	Yes
<i>rpl</i> and <i>rpm</i> genes	50S ribosomal subunit proteins L2, L3, L4, L5, L6, L11, L13, L14, L15, L16, L20, L22, L27, L28, L33 and L36 (16 of the 32 possible subunits)	Yes (except <i>rplE</i> and <i>rplV</i> , encoding L5 and L22, for which the genes present are pseudogenes)
<i>rrsA</i>	16S rRNA	Yes
<i>rrlA</i>	23S rRNA	Yes
<i>rrfA</i>	5S rRNA	Yes
tRNA genes	tRNAs recognizing codons for Met (three), Gly (two), Cys, Phe, Lys, Ala, Glu, Pro, Gln and Ile	No (except for those for Met, Lys and Ala)

DNAP, DNA polymerase; EF, elongation factor; IF, translation initiation factor; RF, peptide chain release factor; RNAP, RNA polymerase; rRNA, ribosomal RNA.

^{*}Genes conserved in '*Candidatus* Sulcia muelleri', '*Candidatus* Zinderia insecticola', '*Candidatus* Carsonella ruddii' and '*Candidatus* Hodgkinia cicadicola'.

[†]RF-2 is missing in '*Ca.* Hodgkinia cicadicola' and '*Ca.* Zinderia insecticola'; they have reassigned UGA from a stop codon to a tryptophan codon, so the protein is no longer needed.

Posizionamento dei geni sul cromosoma e trascrizione



Il posizionamento dei geni sul cromosoma è in parte condizionato da come si muove l'apparato di replicazione.

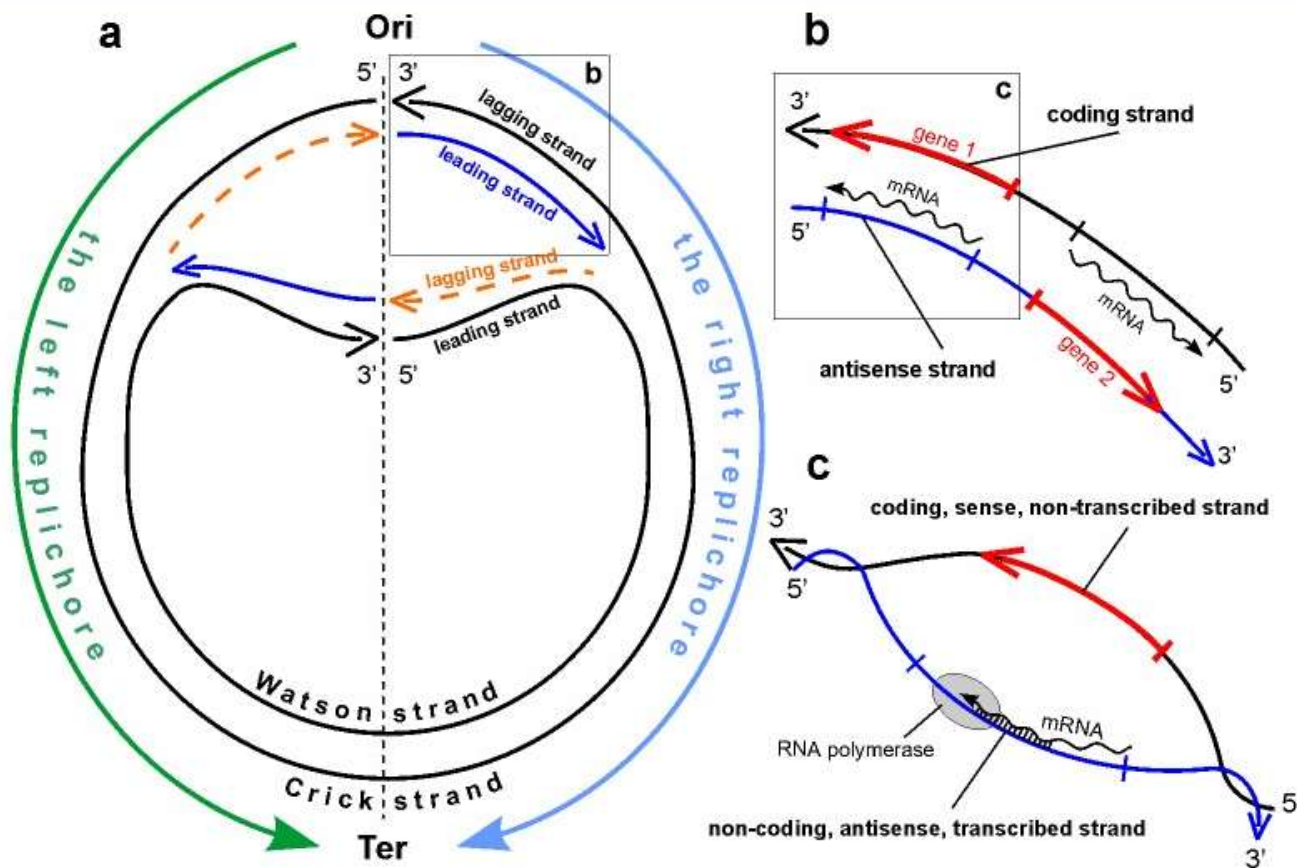
Possiamo definire due regioni definite replichores che dividono il genoma in due metà corrispondenti al movimento delle due forche. In ognuno di questi replichores ci sarà un'elica leading ed una lagging strand che saranno opposte.

E' stato osservato che:

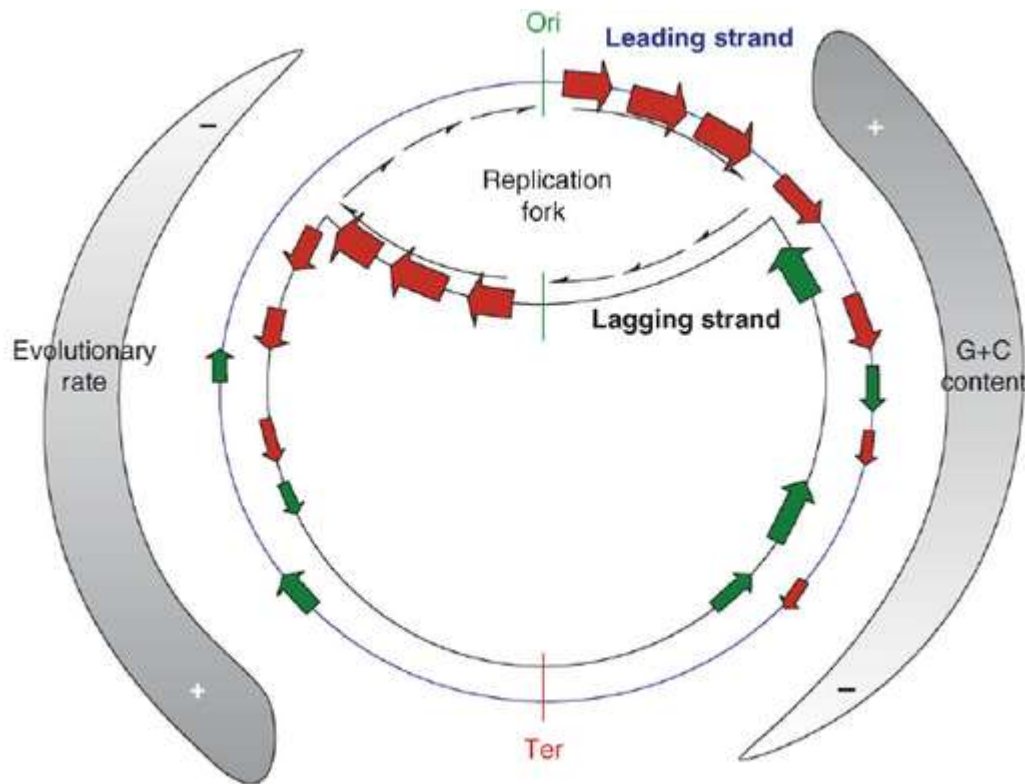
- Il numero di geni trascritti nella stessa direzione in cui si muove la forca replicativa è maggiore rispetto ai geni trascritti in direzione opposta ovvero filamento leading e filamento codificante corrispondono.
- Il cromosoma nelle due opposte metà può presentare asimmetria nel contenuto in G+C, nell'orientamento dei geni nell'accumulo di mutazioni, nella biosintesi di aminoacidi essenziali.

L'assimmetria delle due porzioni indica

1- che grossi apparati macromolecolari che replicano (DNA polimerasi) e trascrivono(RNA polimerasi) il DNA sono coordinati . Infatti i due apparati entrano in collisione quando vi sono eventi di trascrizione in senso opposto alla replicazione.



In *E. coli* i geni espressi ad alta frequenza come i 7 operoni per gli RNA ribosomiali e 53 degli 86 geni per i tRNA sono nella leading strand



Un altro aspetto dell'asimmetria riguarda la composizione in basi delle eliche C/G e A/T dei singoli filamenti che si osserva in 185 genomi (appartenenti a 11 Phyla) e non si trova in lievito, *Drosophila* e uomo. Perché si origina? Essendo replicati in modo asimmetrico e si potrebbero accumulare mutazioni silenti sui frammenti codificanti

Distribuzione dei geni, orientamento e posizione sono fattori non casuali

L'organizzazione del genoma batterico è caratterizzata dalla presenza di macrodomini funzionali , ampie regioni di DNA nel quale ogni singolo gene ha un corretto livello di espressione che dipende dal suo orientamento dalla sua posizione rispetto all'origine.

Se alcune regioni vengono invertite come orientamento sul genoma (pur rimanendo presenti i geni) si può avere la non espressione o espressione molto ridotta: queste grandi inversioni possono indurre anche morte nella cellula .

Elemento importante è anche il posizionamento rispetto all'origine: geni posizionati in posizione prossimale all'origine sono espressi di più in modo statisticamente significativo rispetto a quelli in posizione distale

Le sequenze Chi(χ)

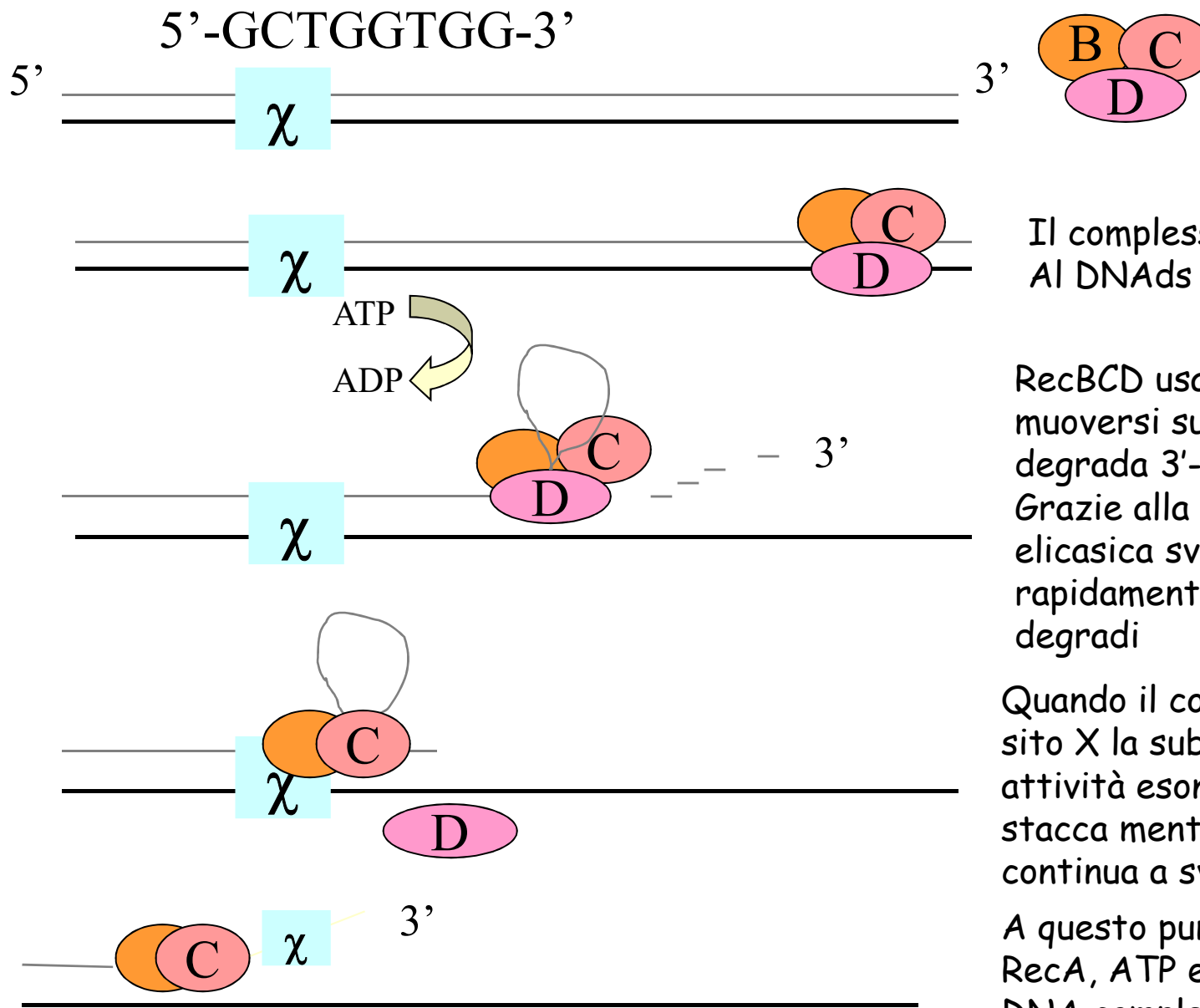
E. coli contiene un numero molto elevato di sequenze chi(χ): queste sequenze promuovono la ricombinazione.

Sequenza specifica di basi (5' GCTGGTGG 3')

Nel 75% dei casi sono presenti nel medesimo orientamento sulla leading strand

Nel cromosoma di *E. coli* è presente

una sequenza Chi circa ogni 5 kb

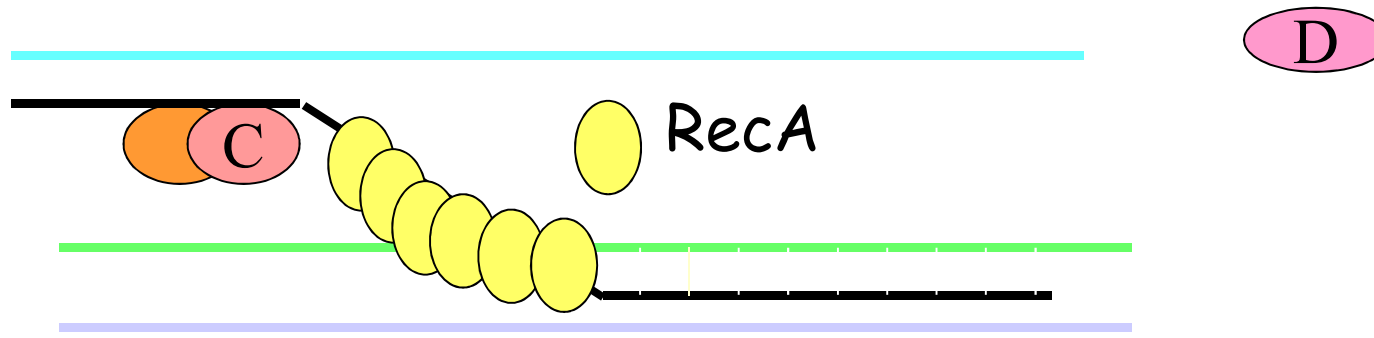


Il complesso RecBCD si lega Al DNAd

RecBCD usa ATP per muoversi sul DNAd e degrada 3'-5' (ExoV) Grazie alla sua attività elicastica svolge il DNA più rapidamente di quanto lo degradi

Quando il complesso arriva al sito X la subunità D dotata di attività esonucleasica si stacca mentre l'elicasi continua a svolgere il DNA

A questo punto se è presente RecA, ATP e una molecola di DNA complementare può avvenire ricombinazione



Grazie a questo processamento mediato dal complesso RecBCD si può creare una regione di DNA a singolo filamento.

RecA sarà in grado di riconoscere e legarsi a questo filamento di ssDNA

Si potrà avere il processo di ricombinazione per invasione di una regione omologa di dsDNA da parte del complesso RecA-ssDNA

SEQUENZE GENOMICHE ED EVOLUZIONE ; quale contributo?

La disponibilità di intere sequenze genomiche da molti microrganismi costituisce un'opportunità unica per ottenere una descrizione più obiettiva delle relazioni evolutive tra i diversi microrganismi

La scoperta che i processi di

- **trasferimento genico orizzontale (HGT)**
- **perdita di geni**

svolgono un ruolo cruciale nell'evoluzione dei microrganismi ha profondamente influito sugli studi dei genomi batterici.

In particolare è cambiato il concetto che **mutazioni all'interno dei singoli geni fossero la fonte principale** delle variazioni fenotipiche in funzione della selezione naturale.

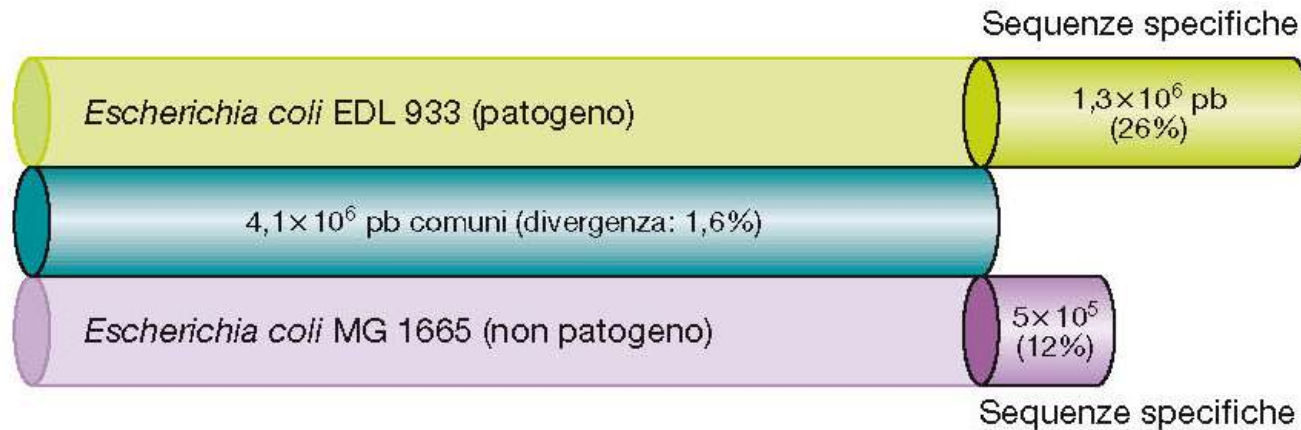
Meccanismi di plasticità genomica

- Trasferimento genico orizzontale (HGT)
- Riduzione dei genomi (Genome reduction)
- Duplicazioni geniche e riarrangiamenti genomici
- Generazione di mutazioni puntiformi

si intende il processo con il quale l'informazione genetica viene trasferita tra batteri diversi

Plasticità genomica un esempio importante: E. coli

Un pool di geni comuni di 4.1 Mb e poi possiamo trovare sequenze specifiche dei E. coli patogeni in questo caso 1.3Mb in più) o nel ceppo di laboratorio (0.5 Mb)



Genoma di *E.coli* : esempio di grande variabilità

La sequenza dell'intero genoma di *E.coli* ha rivelato una variabilità intraspecie estremamente elevata

Sono disponibili 4 sequenze genomiche di *E.coli*

Analisi genomica comparativa ha rivelato che *E.coli* 0157 ha un genoma di 1 Mb più grande di quella di *E.coli* K12 e circa 25% dei geni non sono conservati nel genoma di *E.coli* K12.

Molti dei geni presenti in 0157 si pensa siano stati acquisiti tramite eventi di trasferimento orizzontale e tramite elementi genetici mobili quali fagi, profagi e sequenze IS

Soltanto 3.000 geni sono in comune tra i 4 genomi di *E.coli* mentre erano 4.000 tra *E.coli* K12 e 0157

I 3000 geni comuni presentano SINTENIA suggerendo una base di trasmissione verticale

Meccanismi che contribuiscono alla plasticità del genoma

Elemento genetico o meccanismo
Acquisizione di proprietà

Conseguenze

• Mutazioni puntiformi

• Alterazione dell'espressione genica

• Ricombinazione omologa

• Riarrangiamenti DNA, inversioni, duplicazioni, delezioni, integrazione di DNA per HGT

• Trasformazione

• Acquisizione di informazione genetica

• IS, elementi compositi

• Inserzioni, delezioni, inversioni, alterazione espressione genica

• Integroni

Trasferimento di geni, riarrangiamenti del DNA

• Trasposoni coniugativi e plasmidi

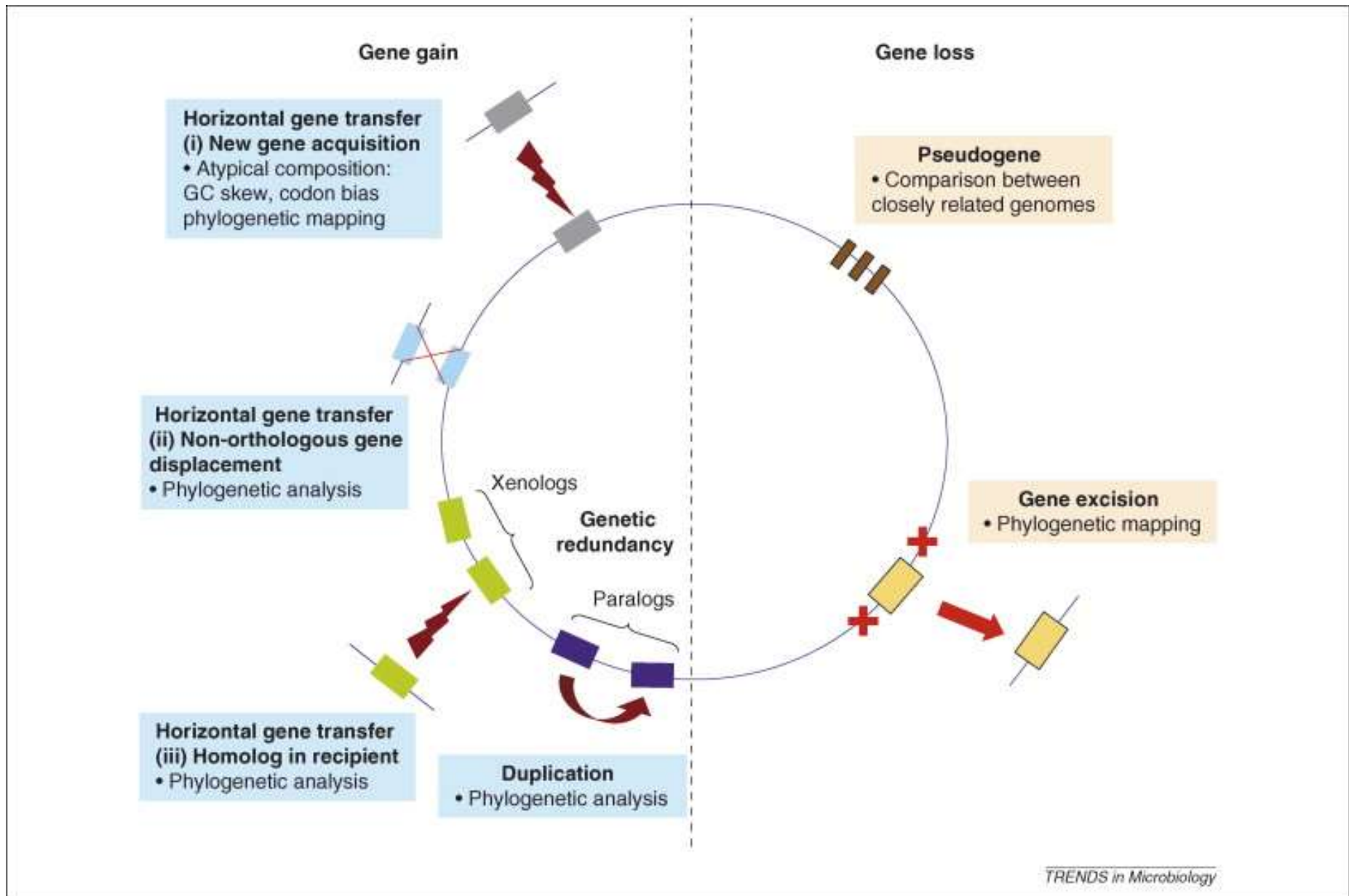
Coniugazione, HGT, Mobilizzazione

• Batteriofagi

• Trasduzione generalizzata e specializzata HGT

• GEI, PAI, isolotti

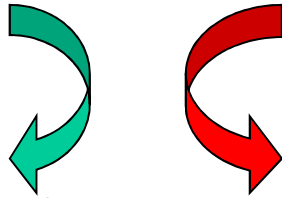
• HGT, Integrazione delezione di grandi regioni di DNA



The dynamics of genome repertoire. Bacterial genomes are dynamic entities that constantly gain (left; blue boxes) and lose genes (right; beige boxes). These modifications of gene repertoires arise by different mechanisms. First, bacterial genomes can acquire genetic material from other organisms, even distantly related ones. Horizontal gene transfers are evidenced by different types of approaches that generally identify distinct sets of genes. (i) Analysis of gene composition (GC%, codons) identifies mostly genes that are rarely found in other species. This generally precludes a confirmation of their foreign status by a phylogenetic analysis. However, a mapping of gene presence on a phylogenetic tree of complete genomes can confirm that they have been recently acquired in the genome. By contrast, phylogenetic analysis can reveal HGT for genes that have wider phylogenetic distribution, and these genes only rarely show a striking difference in composition. In this case, HGT can result in the addition of a completely new gene (i), the replacement of an existing gene (ii) or genetic redundancy (iii) if a homologous gene is already present in the recipient genome. Genetic redundancy can also arise from gene duplication and only phylogenetic analysis can distinguish between these two origins. Recent analyses have demonstrated that HGT participates significantly in the degree of redundancy in a bacterial genome. Gene excision and formation of pseudogenes are the mechanisms for gene loss. Excision occurs when a gene is completely deleted from the genome, and pseudogene formation occurs when mutations (point mutation and/or insertion/deletion) accumulate, resulting in function loss. The loss of a gene is evidenced by the absence of the gene in the analyzed genome whereas it is present in related species (phylogenetic mapping). Pseudogenes can be identified by comparisons of closely related genomes.

Disponibilità di un'ampia gamma di genomi di batteri ha fornito delle informazioni notevoli sull'evoluzione dei genomi dei microrganismi

Due fenomeni contrastanti non spiegabili con un meccanismo di evoluzione VERTICALE:



Drammatiche differenze nel genoma di microrganismi simili quali *E.coli* ed *Haemophilus* o tra *E.coil* 0157 ed *E.coil* K12.

Ampie regioni di omologie tra genomi di batteri filogeneticamente distanti o tra ARCHEA e di BATTERI

HGT è uno dei principali meccanismi per spiegare l'evoluzione dei microrganismi

Horizontal gene transfer HGT

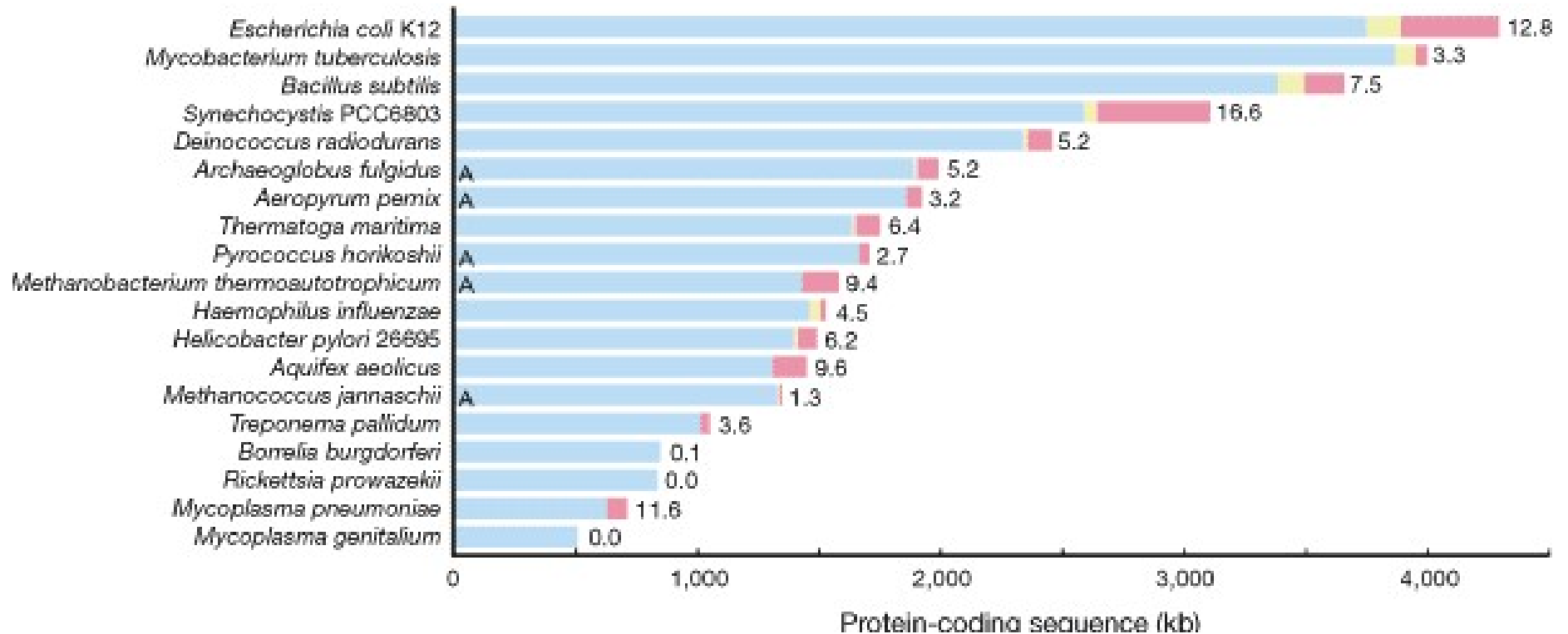
Trasferimento genico orizzontale o trasferimento genico laterale è un fenomeno evolutivo che coinvolge scambi genetici tra differenti linee evolutive

Per lungo tempo si è dibattuto sul ruolo del HGT nell'evoluzione perché variava completamente la tradizionale visione dell'evoluzione della vita ed era difficile da provare

Nel 1991 si riuscì a dimostrare che in E.coli 10-16% del genoma risultava da HGT:

Il fenomeno quindi aveva un ruolo essenziale nell'evoluzione

Ma che percentuale rappresenta il DNA acquisito per HGT?



La lunghezza delle barre rappresenta la quantità di DNA codificante suddiviso in DNA "nativo" (in blu), DNA presente in elementi genetici mobili (in giallo) e DNA proveniente da HGT o straniero (in rosa). I numeri si riferiscono alle percentuali di DNA "straniero"

HGT e la filogenesi

HGT (Horizontal Gene Transfer) è

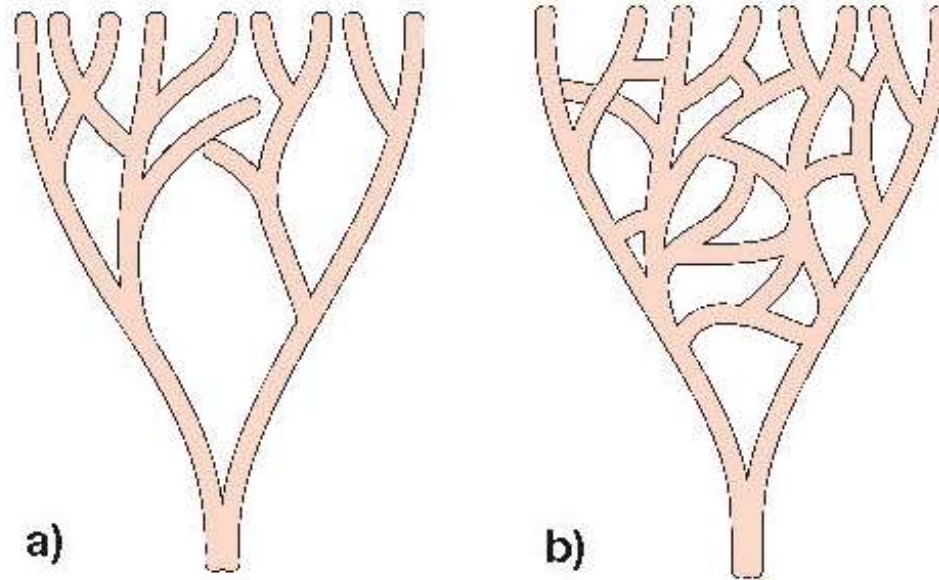
- un evento molto frequente e coinvolge ceppi appartenenti a gruppi tassonomici distanti.
- ha profondamente modificato la filogenesi di tutti i procarioti .

Nella visione darwiniana la filogenesi è il risultato di modificazioni e diversificazioni che avvengono nella discendenza verticale attraverso le generazioni

HGT implica la trasmissione di caratteri non alla discendenza ma **ORIZZONTALMENTE**.

Geni diversi possono quindi aver subito strade evolutive diverse all'interno dello stesso organismo.

Effetti del HGT sulla filogenesi



Dalle sequenze dei genomi emerge che i segmenti trasferiti orizzontalmente sono molto numerosi e coinvolgono ampie regioni .
L'albero filogenetico considerando HGT assume quindi un aspetto ramificato

Horizontal Gene Transfer (HGT)

Definito anche Trasferimento Genico Laterale (HGT and LGT)

Come avviene : le principali strategie sono :

Trasformazione :

- frammenti di DNA nudo penetrano o sono captati dalle cellule

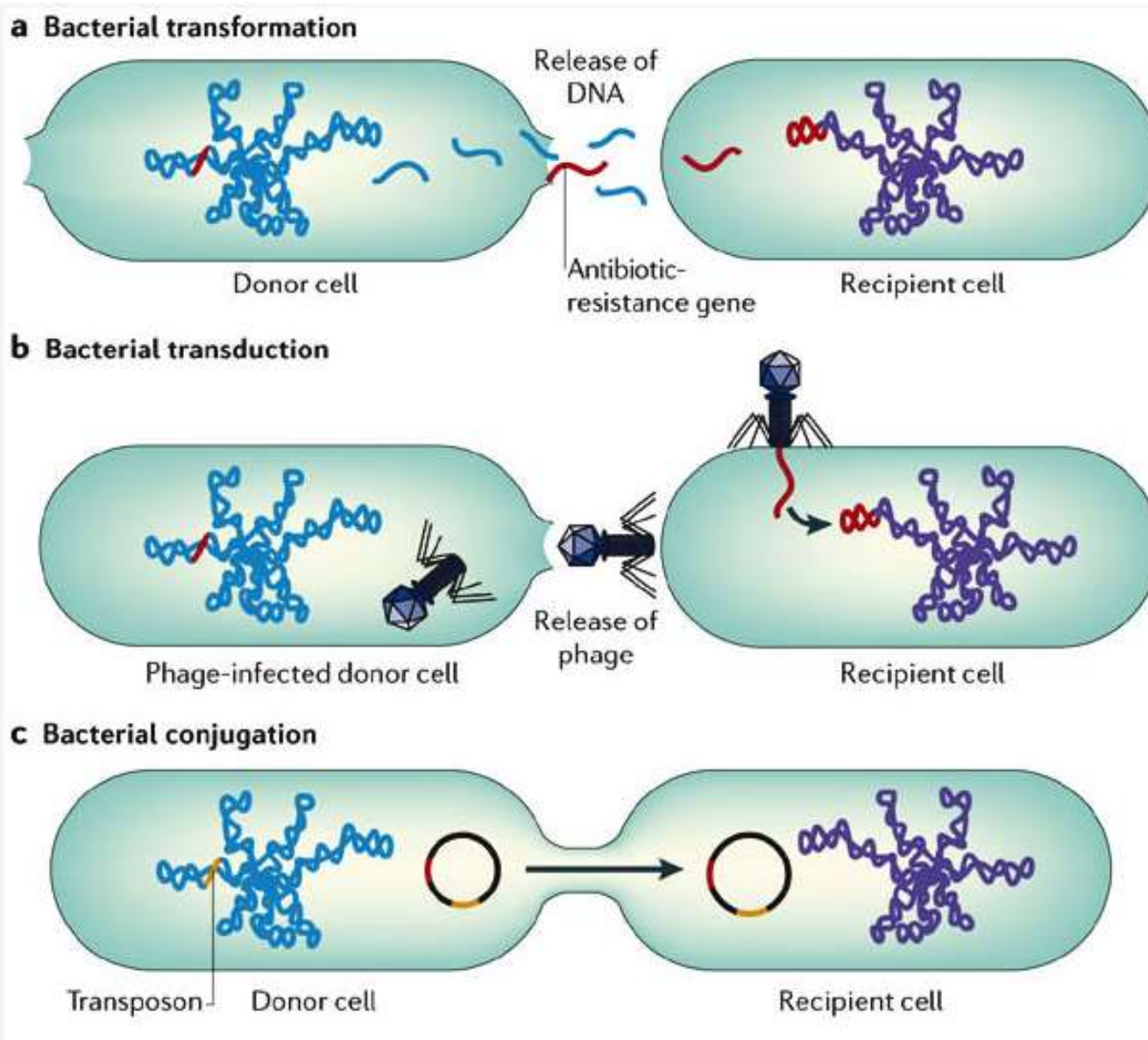
Trasduzione

- I batteriofagi fanno transitare DNA batterico da una cellula donatrice ad una ricevente :
- c'è un limite nella quantità di DNA trasferito determinata dalle dimensioni della testa fagica

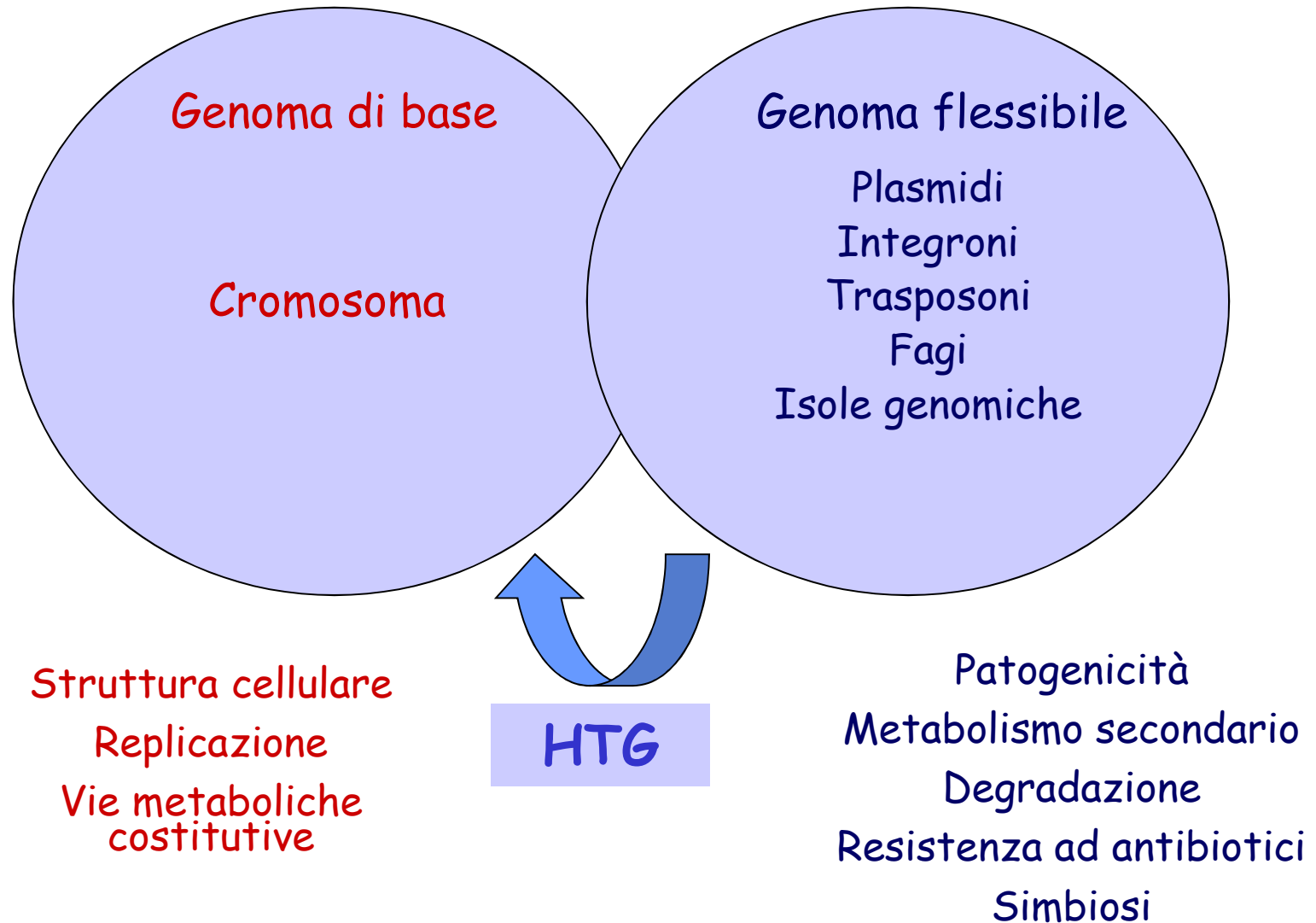
Coniugazione :

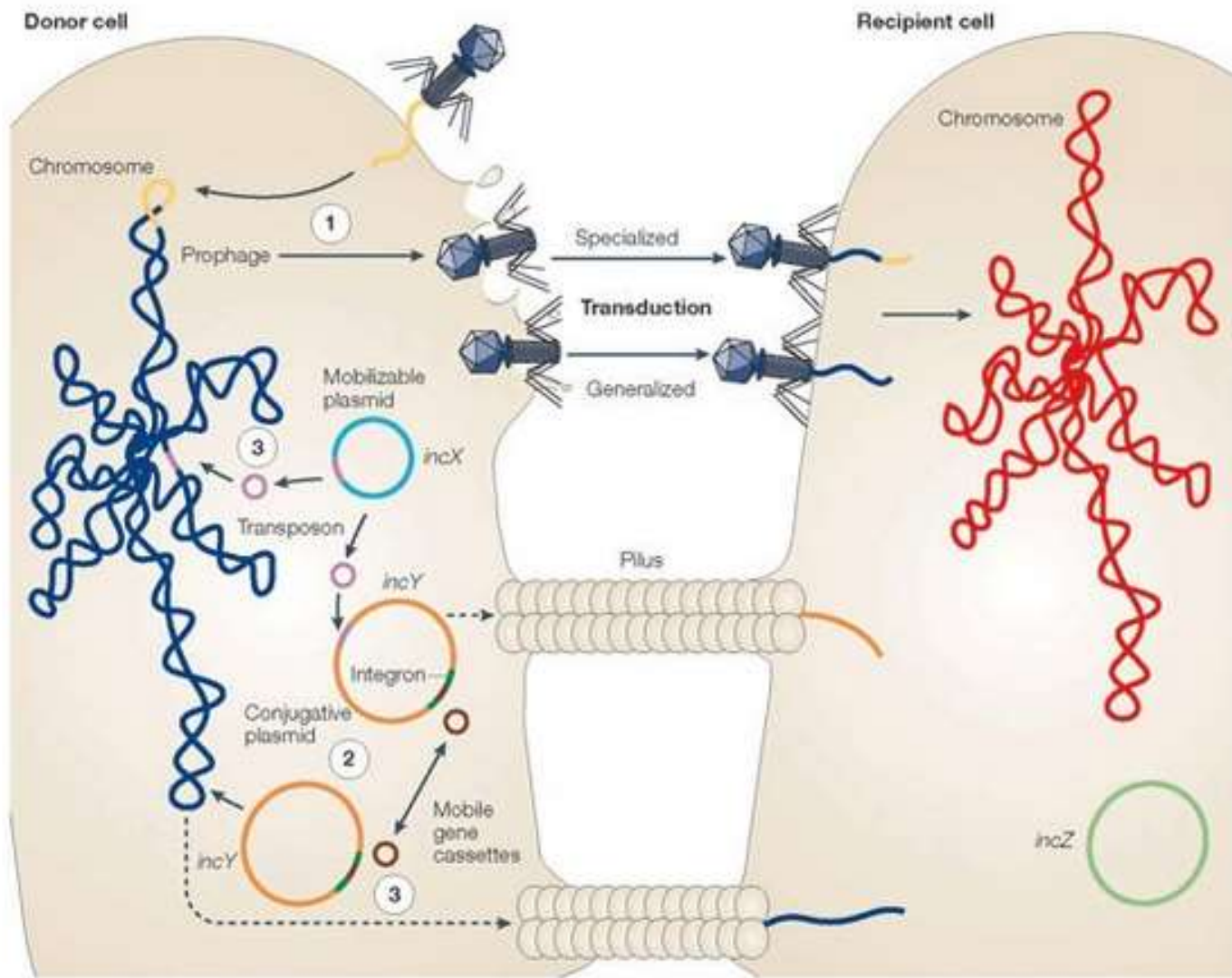
necessita di elementi genetici mobili quali i plasmidi , gli ICE,
richiede il contatto cellula-cellula
possono passare frammenti di DNA di dimensioni maggiori

Le 3 strategie a confronto nella trasmissione di geni di antibiotico resistenza



EVOLUZIONE DEL GENOMA BATTERICO





Classici esempi di HGT

- Geni di antibiotico resistenza su plasmidi
- Le sequenze di Inserzione o IS
- Le Isole di Patogenicità
- I geni per le tossine sui plasmidi
- Il plasmid eTi di *Agrobacterium*

Che cosa serve perchè avvenga il trasferimento

- Disponibilità del DNA donatore
- Stabilità del DNA nell'ambiente
- Un vettore per la trasmissione (fago/plasmide/ICE)
- Un sistema di captazione e inserimento
- Un sistema di mantenimento
- Un sistema per la stabilizzazione
- Un vantaggio selettivo

Quali sono i limiti e che cosa può prevenire il trasferimento

- Instabilità nel nuovo ospite
- I sistemi di restrizione
- L'incompatibilità nell'uso di codoni
- Mancanza di geni appropriati per l'espressione
- Silenziamento dei geni trasferiti

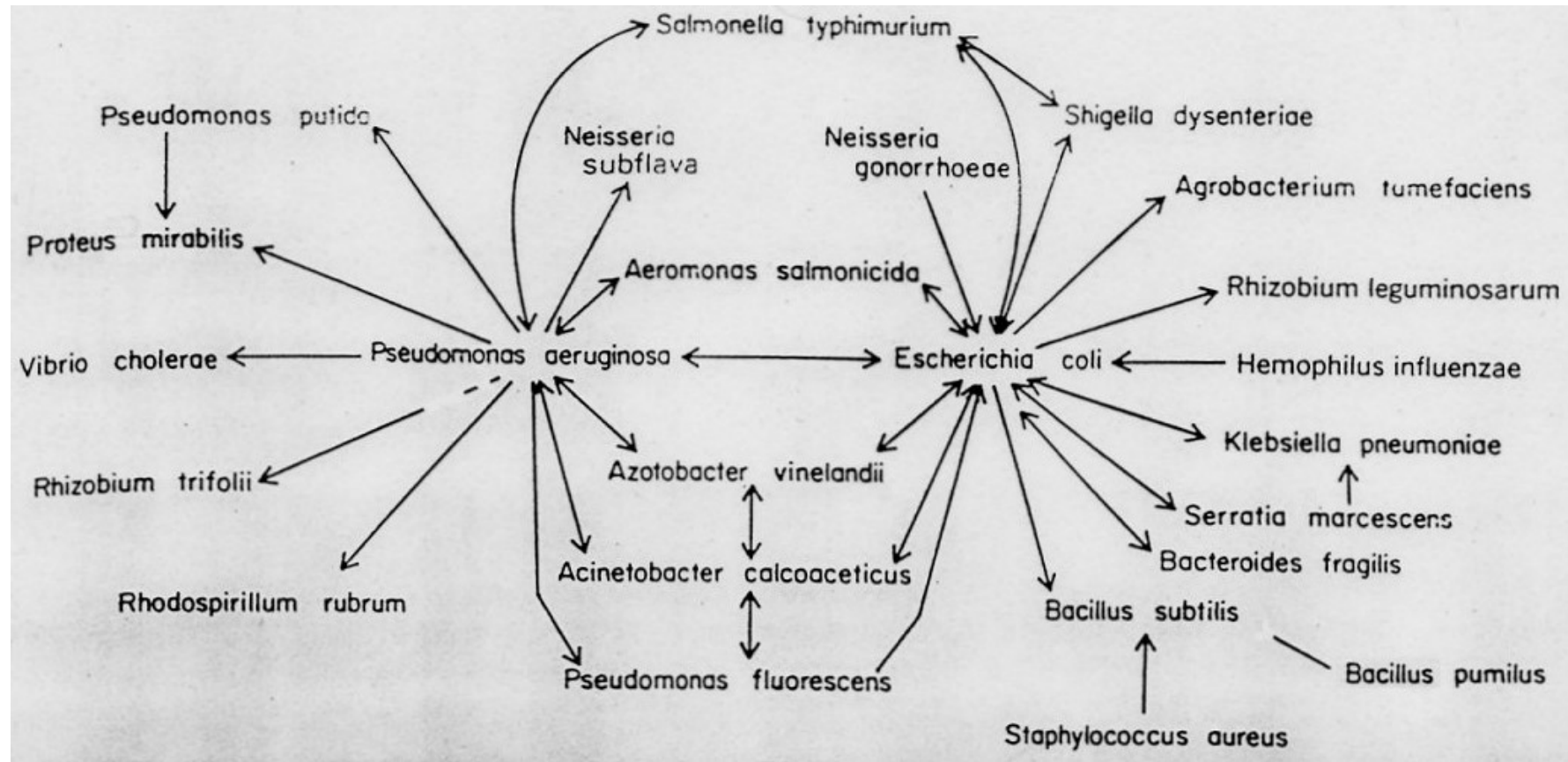
Tre principali caratteristiche nell'evoluzione di un patogeno sono trasmesse da HGT:

ANTIBIOTICO RESISTENZA

DETERMINANTI DI PATOGENICITA'

DETERMINANTI COINVOLTI
NELL'ADATTAMENTO METABOLICO o
MIGLIORAMENTO DELLA FITNESS

I plasmidi permettono lo scambio genetico tra le diverse specie e sono tra i principali responsabili del **Trasferimento Genico Orizzontale (HGT)**



ANTIBIOTICO-RESISTENZA

I determinanti di antibiotico resistenza sono spesso associati con elementi genetici mobili quali:

- plasmidi
- integroni
- superintegroni
- trasposoni complessi
- isole genomiche di antibiotico resistenza

INTEGRONI sono dei sistemi naturali :

- per il clonaggio ed espressione genica
 - in grado di incorporare ORF e convertirle in geni funzionali
- Costituiscono sistema primario per 'evoluzione della antibiotico resistenza e cattura dei geni di virulenza (specialmente nei batteri Gram-

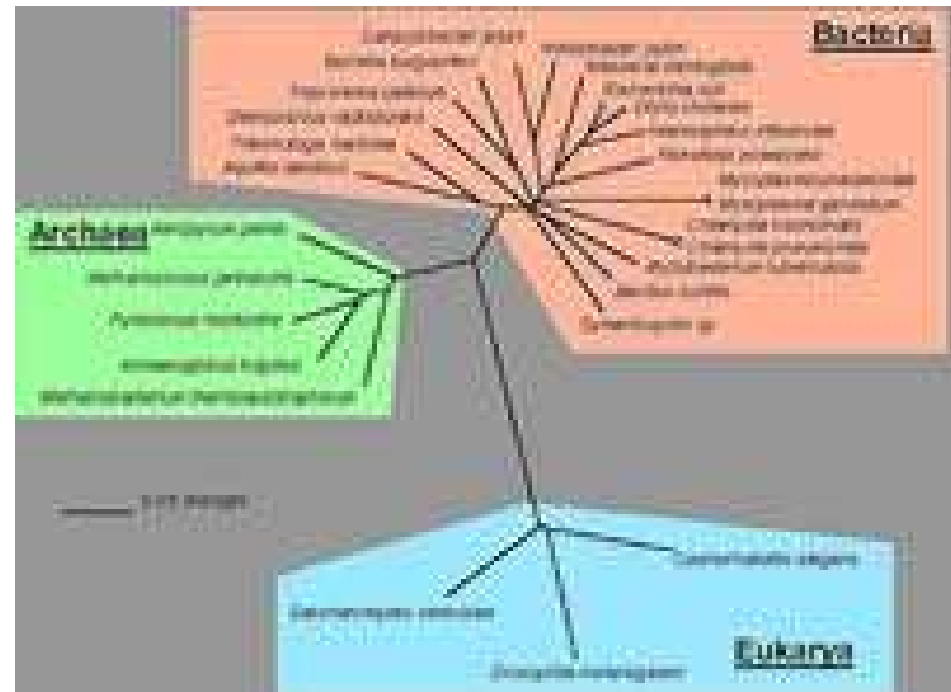
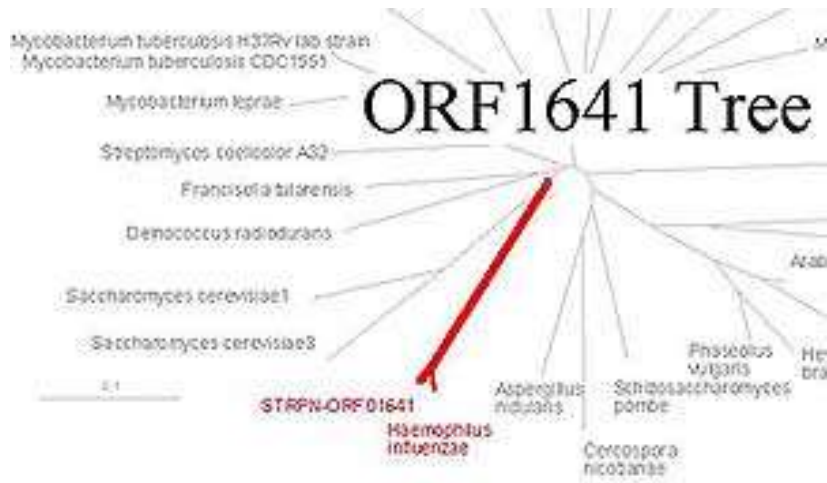
SUPERINTEGRONI sono dei sistemi di stoccaggio delle cassette geniche con funzioni diverse

Come si può rilevare la presenza di sequenze arrivate per HGT

Le sequenze arrivate in un nuovo genoma per HGT mantengono per un periodo più o meno lungo le caratteristiche del genoma d'origine distinte dal genoma ricevente

- Analisi della sequenza di DNA
 - Alberi filogenetici
 - Contenuto in G+C
 - Utilizzazione dei codoni

Confrontando le omologie di un gene con l'albero filogenetico ottenuto con quello relativo rRNA 16S



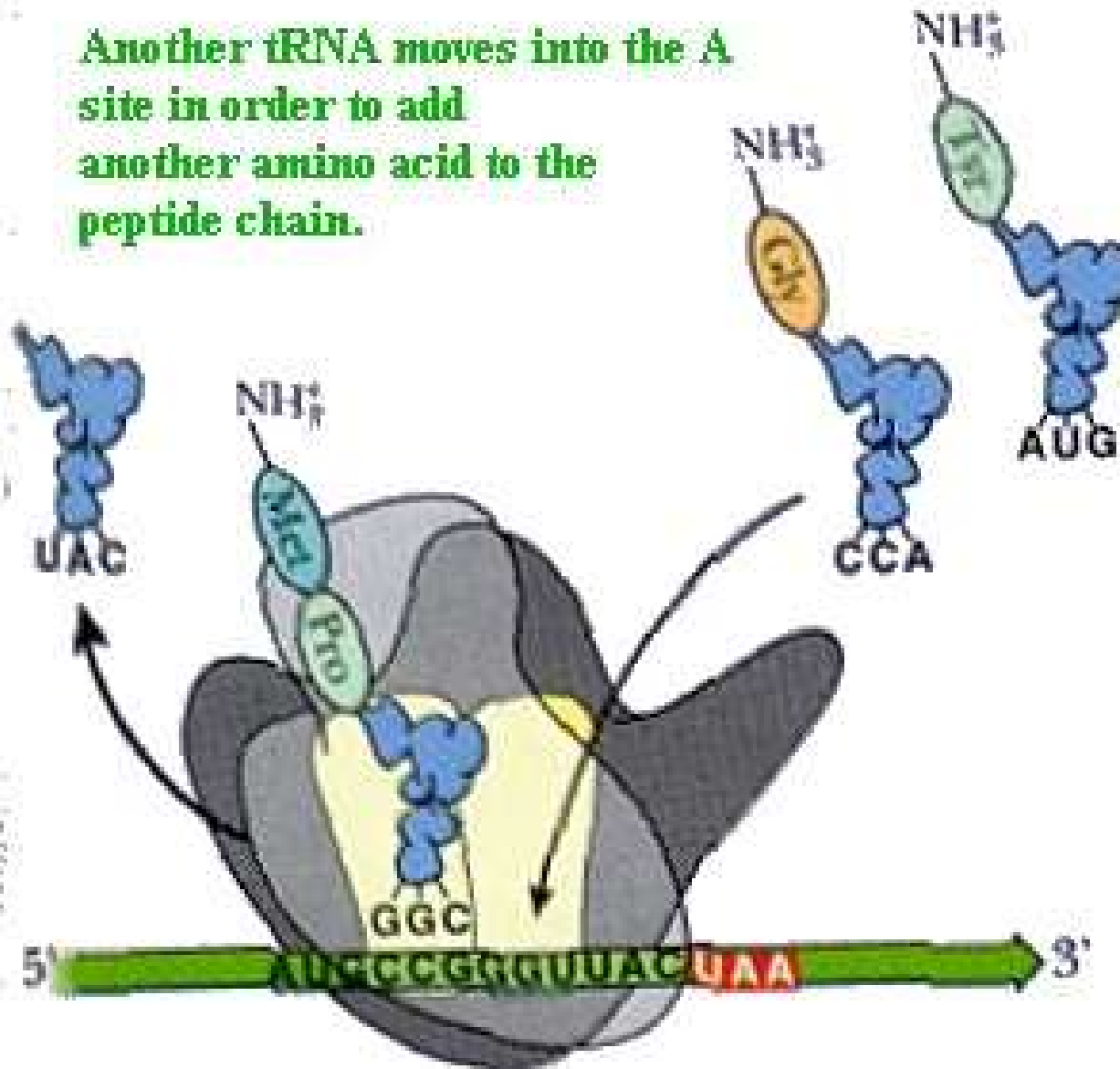
Confrontando il contenuto in G+C

- Nel DNA G (Guanina) sono accoppiate con le C (Citosina)
- Misurando la quantità in G+C di una determinata regione
- Se una regione ha un contenuto in G+C molto diverso dalla gran parte del genoma molto probabilmente deriva da HGT



L'utilizzazione dei codoni è tipica di ogni specie

Another tRNA moves into the A site in order to add another amino acid to the peptide chain.

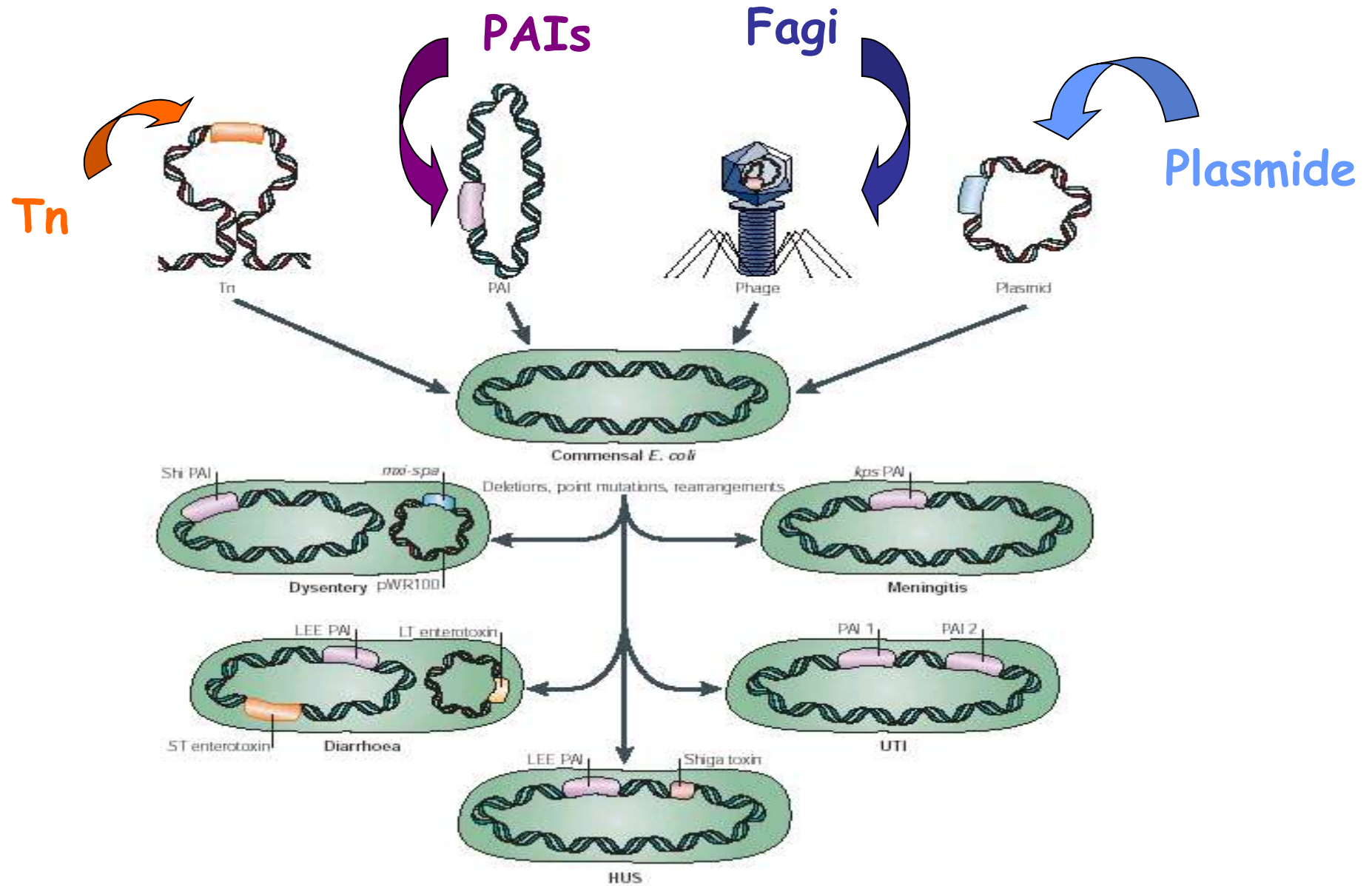


Second letter

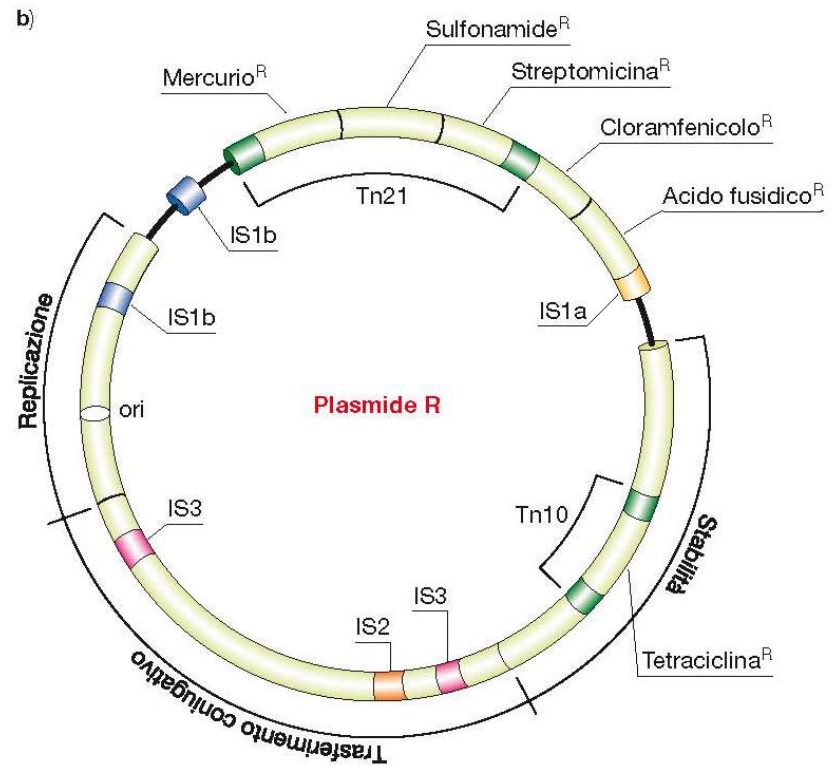
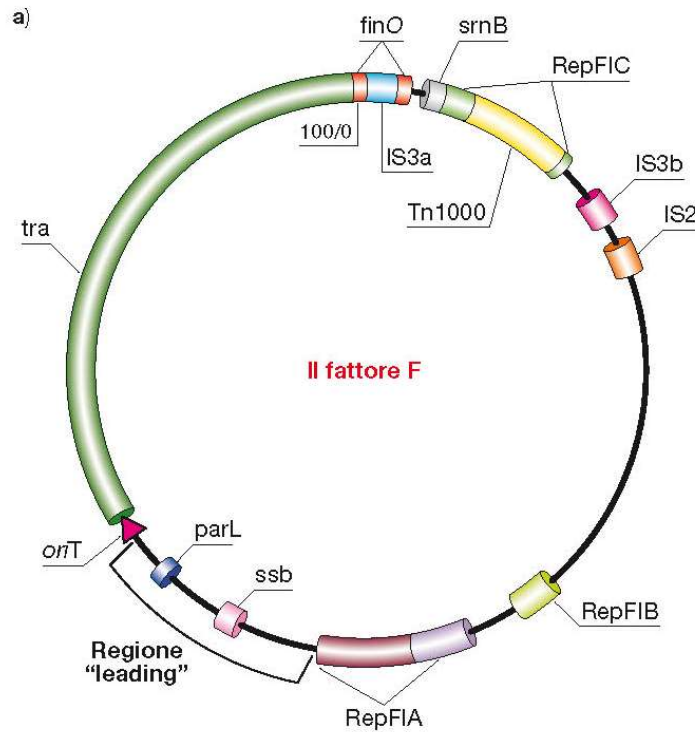
First letter

	U	C	A	G	
U	UUU Phenylalanine UUC UUA Leucine UUG	UCU Serine UCC UCA UCG	UAU Tyrosine UAC UAA Stop codon UAG Stop codon	UGU Cysteine UGC UGA Stop codon UGG Tryptophan	U C A G
C	CUU Leucine CUC CUA CUG	CCU Proline CCC CCA CCG	CAU Histidine CAC CAA Glutamine CAG	CGU Arginine CGC CGA CGG	U C A G
A	AUU Isoleucine AUC AUA AUG Methionine; initiation codon	ACU Threonine ACC ACA ACG	AAU Asparagine AAC AAA Lysine AAG	AGU Serine AGC AGA Arginine AGG	U C A G
G	GUU Valine GUC GUA GUG	GCU Alanine GCC GCA GCG	GAU Aspartic acid GAC GAA Glutamic acid GAG	GGU Glycine GGC GGA GGG	U C A G

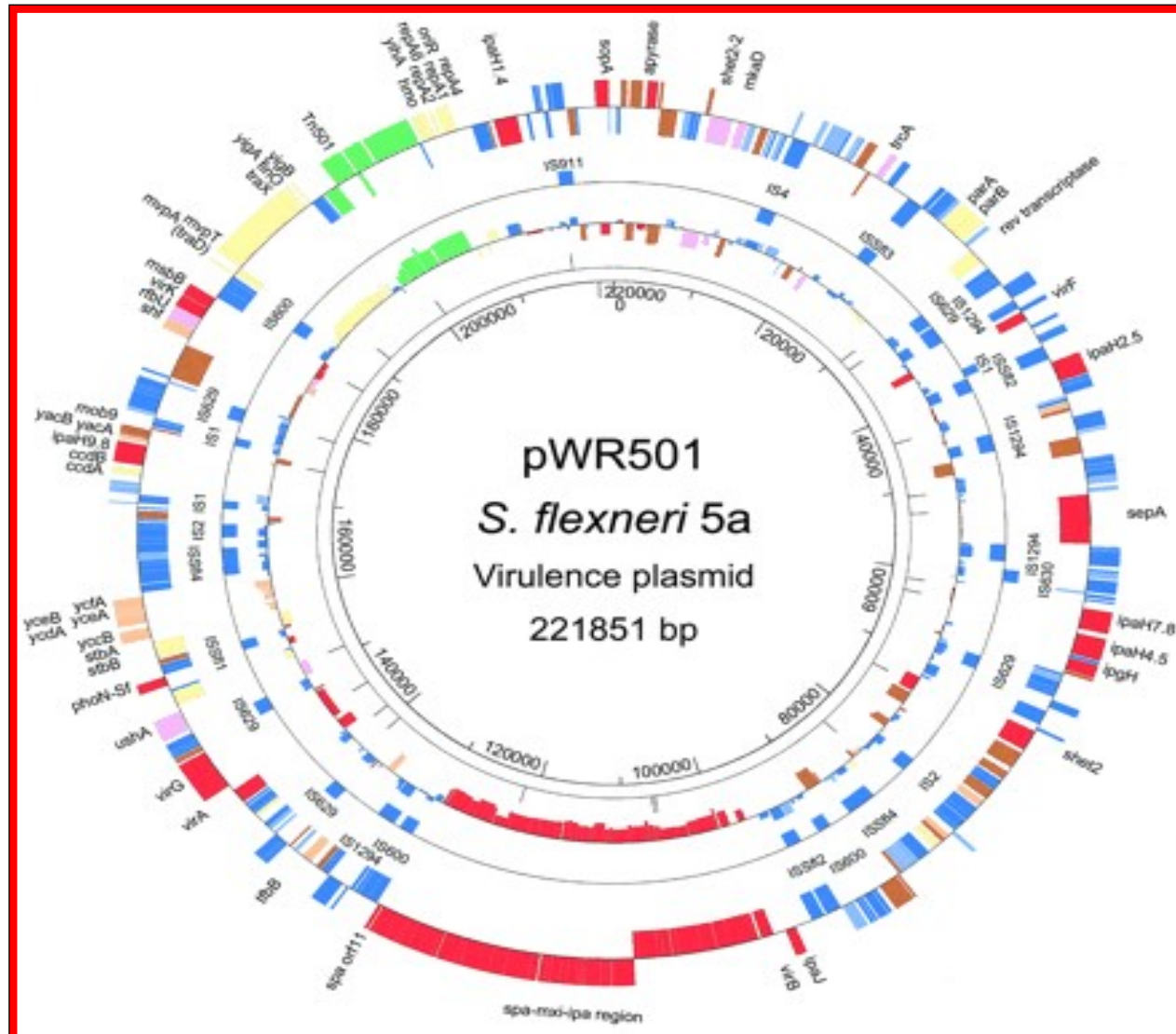
Le principali strutture coinvolte nel trasferimento genico orizzontale



Struttura di F e di un classico plasmide R1 contenente numerosi geni di antibiotico resistenza

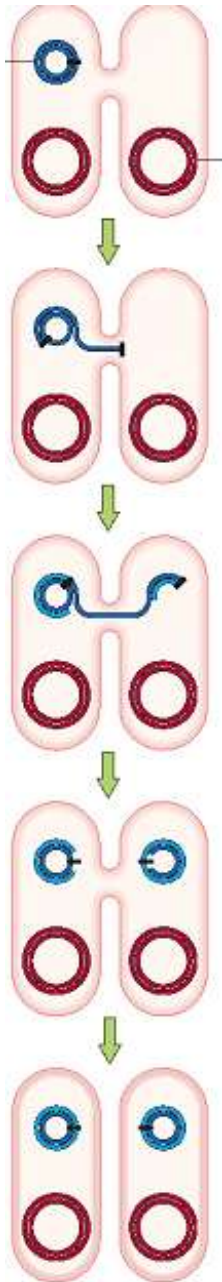


Plasmide di virulenza di *Shigella* contiene numerose IS



rosso → geni di virulenza

blu → IS

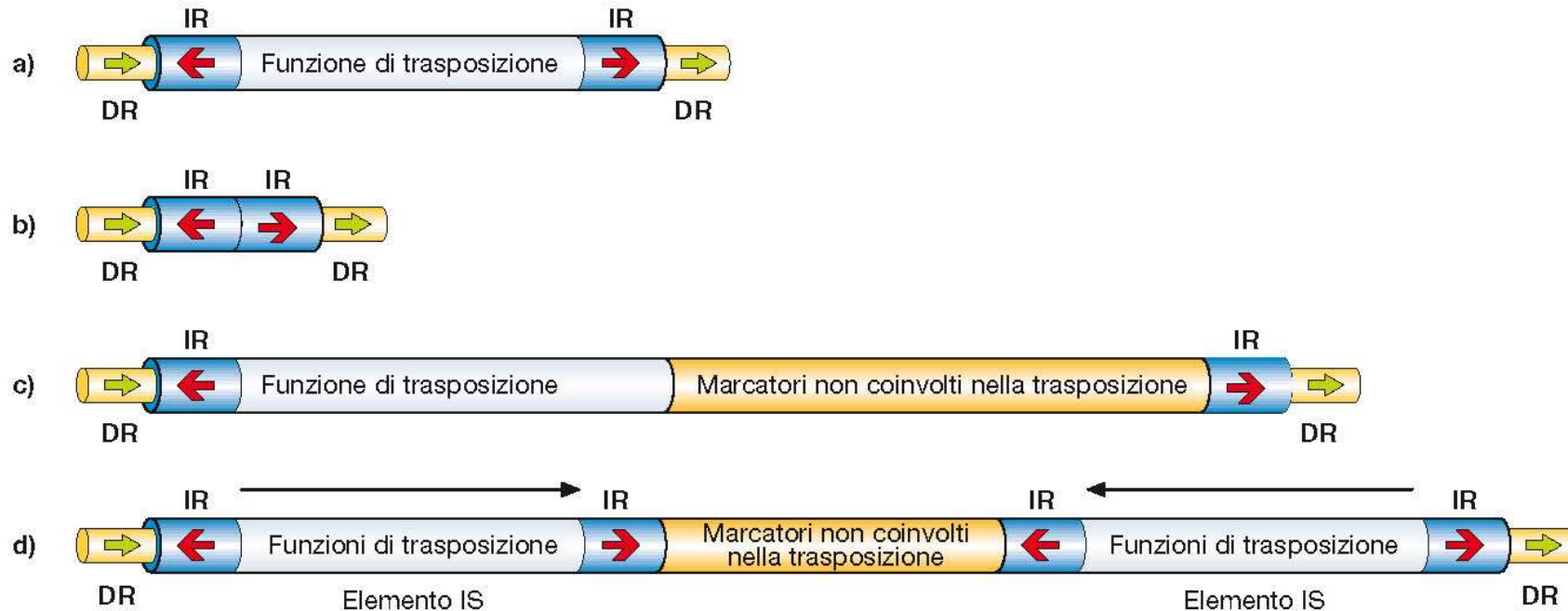


I plasmidi sono i responsabili del processo di coniugazione: potente sistema di trasferimento genico intra ed inter specie

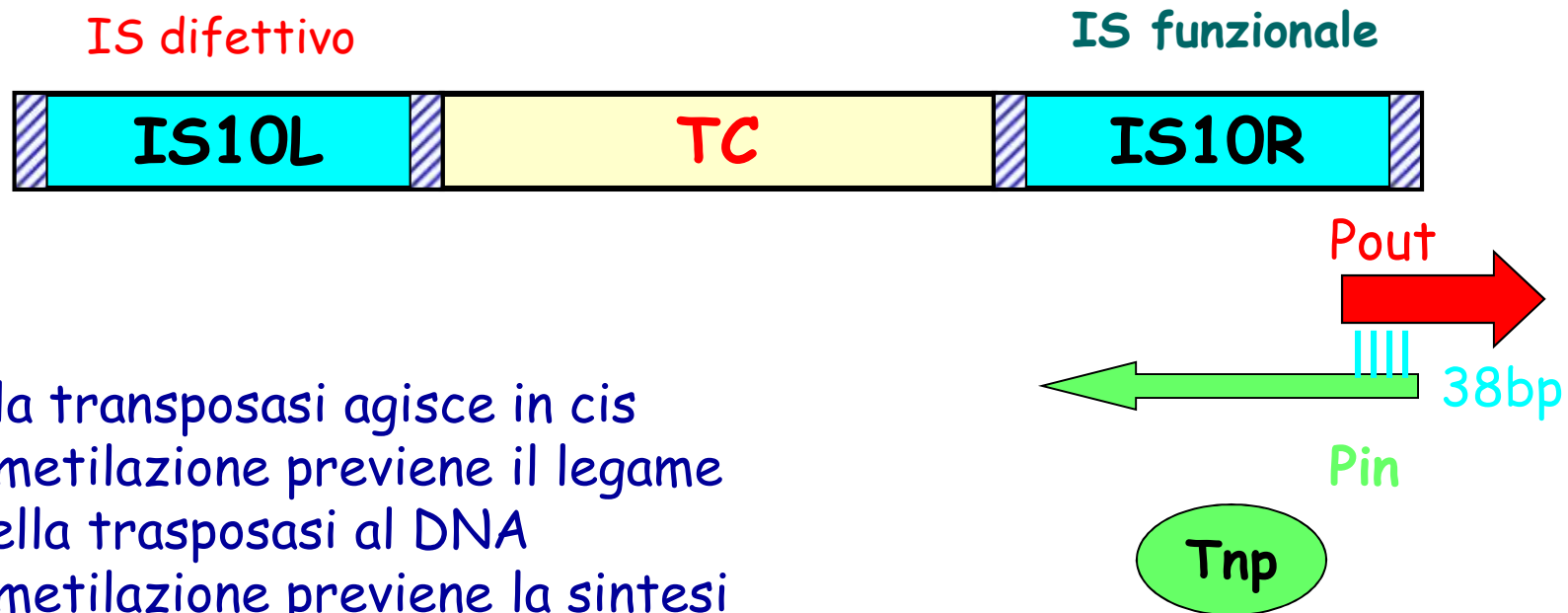
Gli elementi trasponibili dai più semplici ai più complessi

Caratteristiche comuni

1. **Contengono i geni per la trasposasi**
2. **Contengono delle sequenze inversamente ripetute all'estremità**
3. **Generano duplicazioni dirette nel sito d'inserzione**

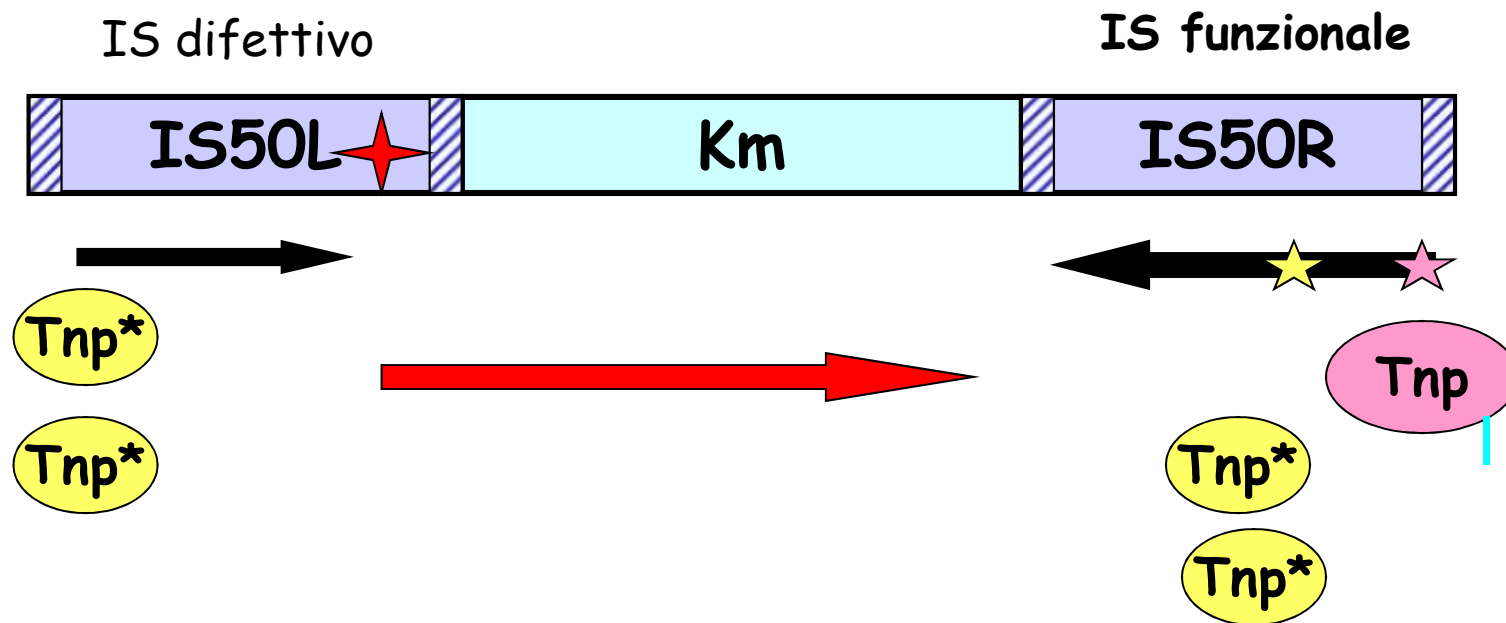


Struttura di Tn10



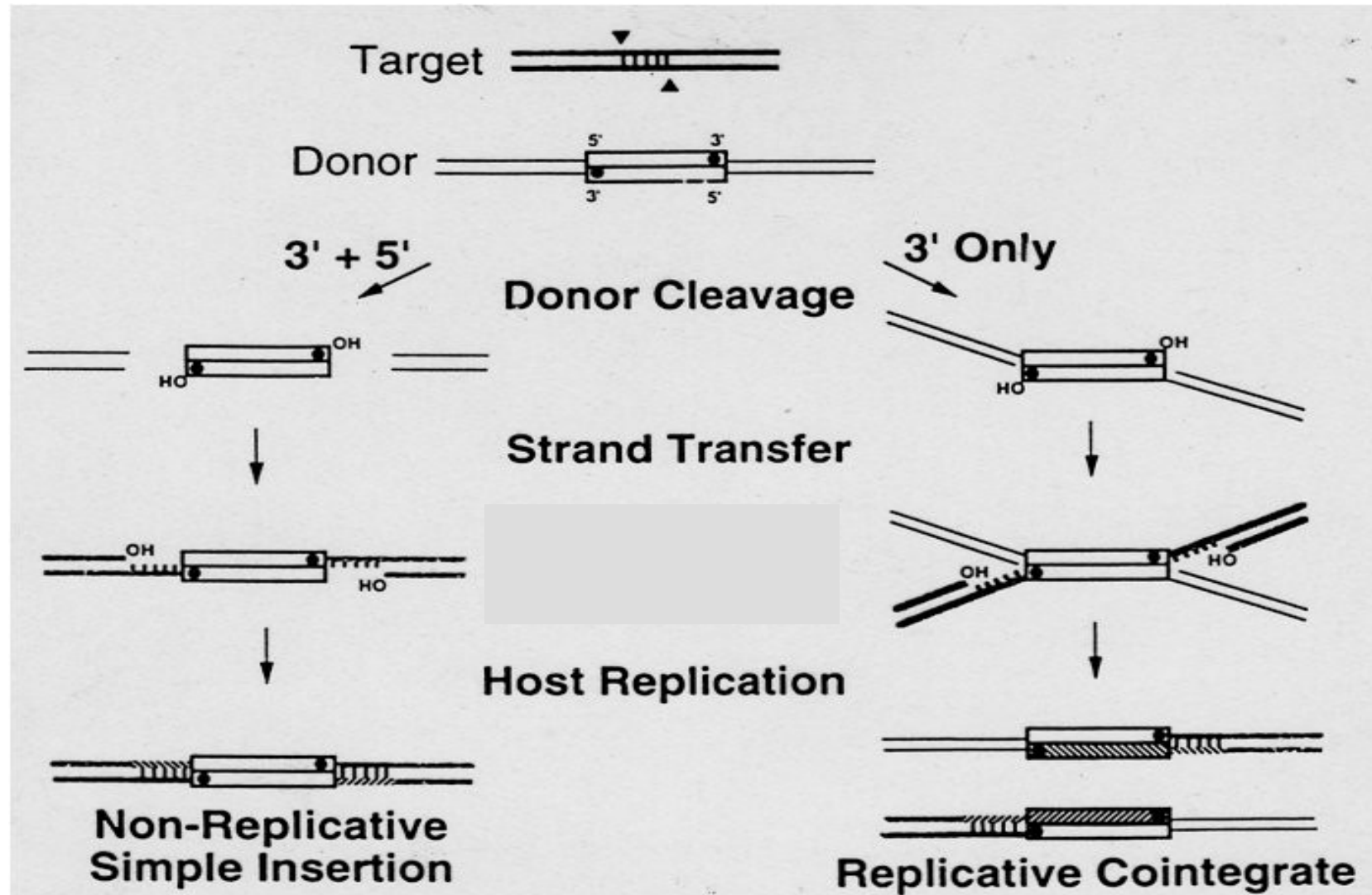
- la transposasi agisce in cis
- metilazione previene il legame della trasposasi al DNA
- metilazione previene la sintesi della trasposasi

Struttura di Tn5



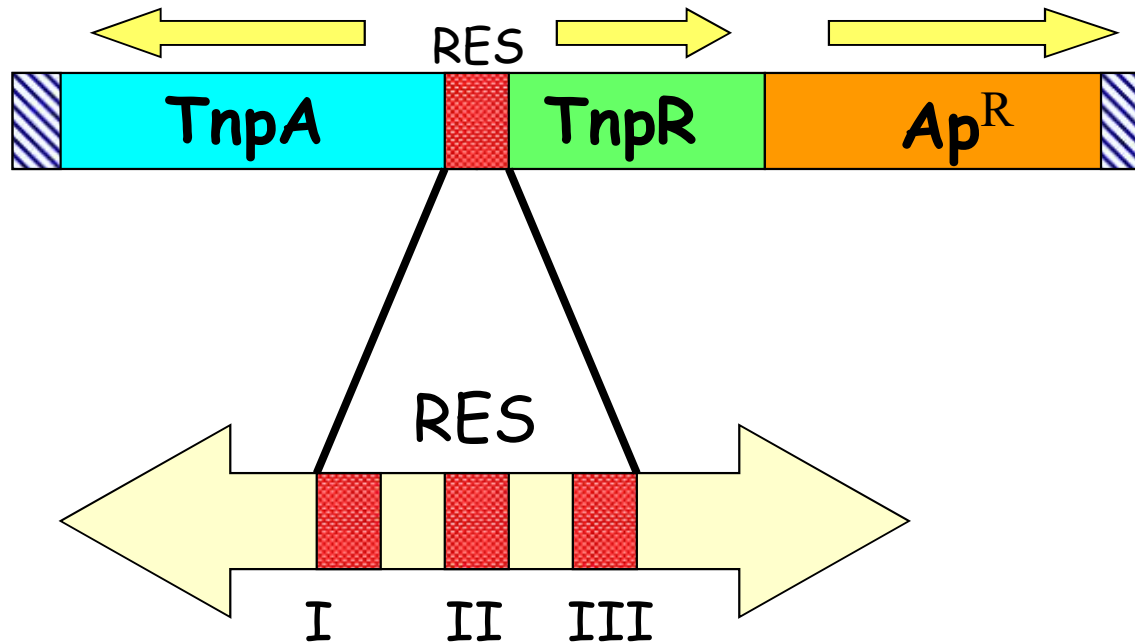
Trasposizione conservativa e replicativa a confronto:

si noti il taglio solo al 3' (replicativa) e
al 3' e 5' (conservativa) nella molecola donatore



Struttura dei trasposoni di tipo Tn3 (TnAp)

Modalità di trasposizione di tipo replicativo



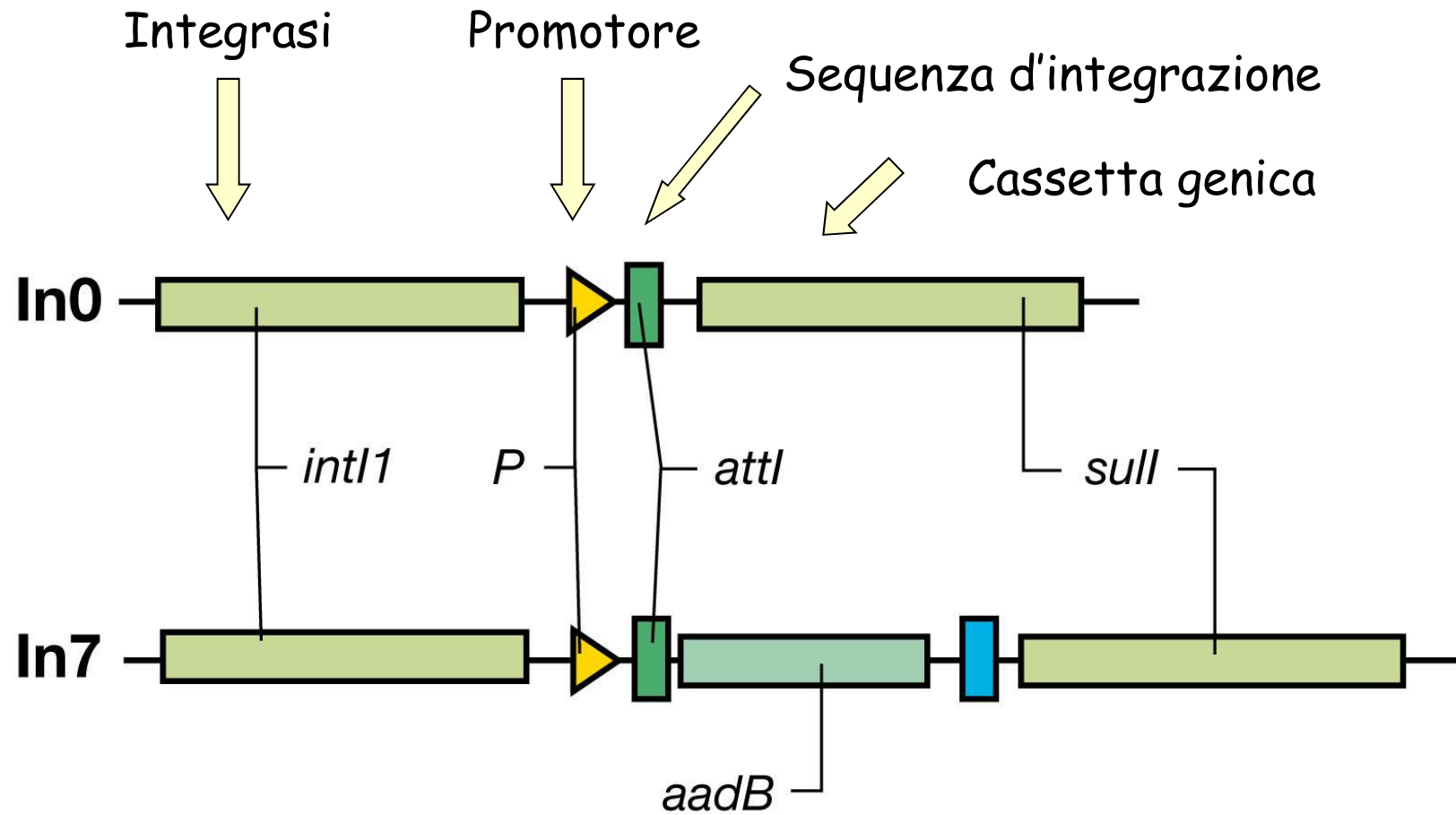
TnpR resolvasi e repressore della trasposasi

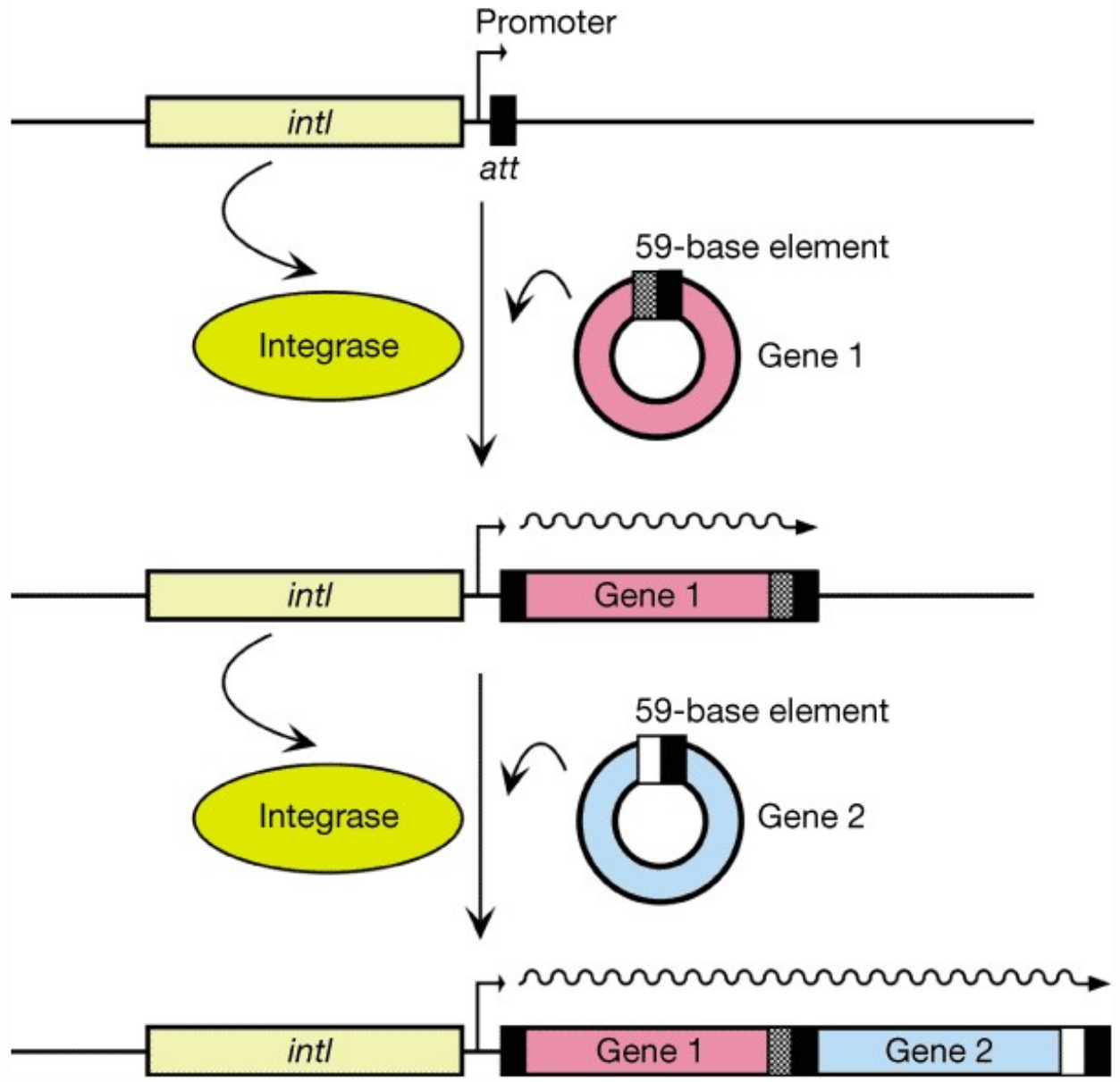
INTEGRONI sono elementi genetici che hanno la capacità di recuperare geni e di farli esprimere

Gli integroni sono caratterizzati da 3 elementi chiave:

- un gene (**intI**) che codifica per l'integrasi, un enzima in grado di indurre ricombinazione sito specifica;
- una sequenza specifica (**attI**) che viene utilizzata per l'integrazione integrasi-mediata;
- un promotore orientato verso l'esterno che permette l'espressione del gene integrato.

Struttura degli integroni

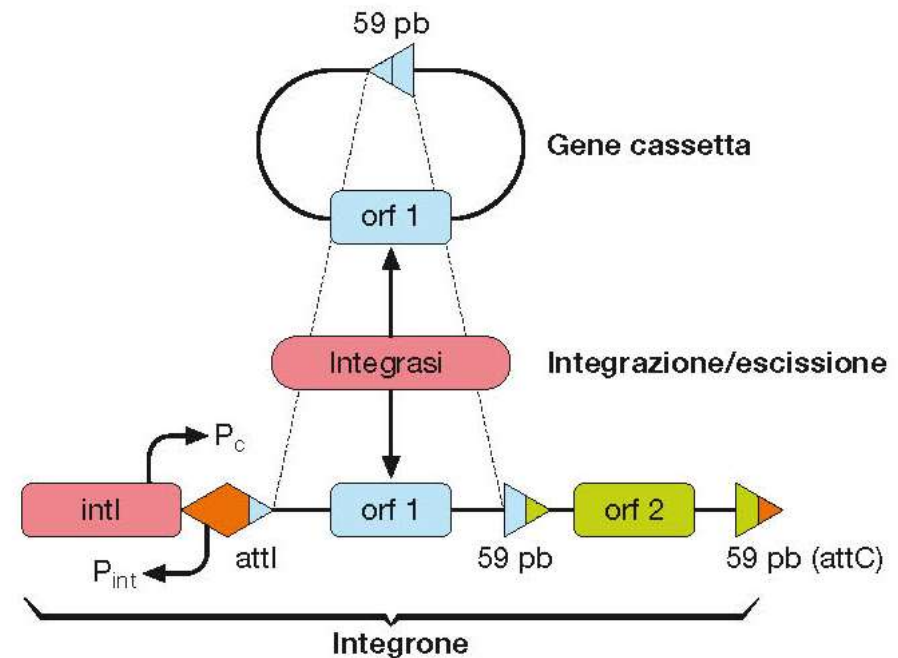




Gli **INTEGRONI** integrano al loro interno non geni casuali ma specifiche **cassette geniche** caratterizzate:

- geni fiancheggiati da sequenze specifiche di DNA riconosciute dall'integrasi;

- geni che non si esprimono finchè non vengono integrati nell'integrone e possono essere trascritti dal promotore dell'integrone

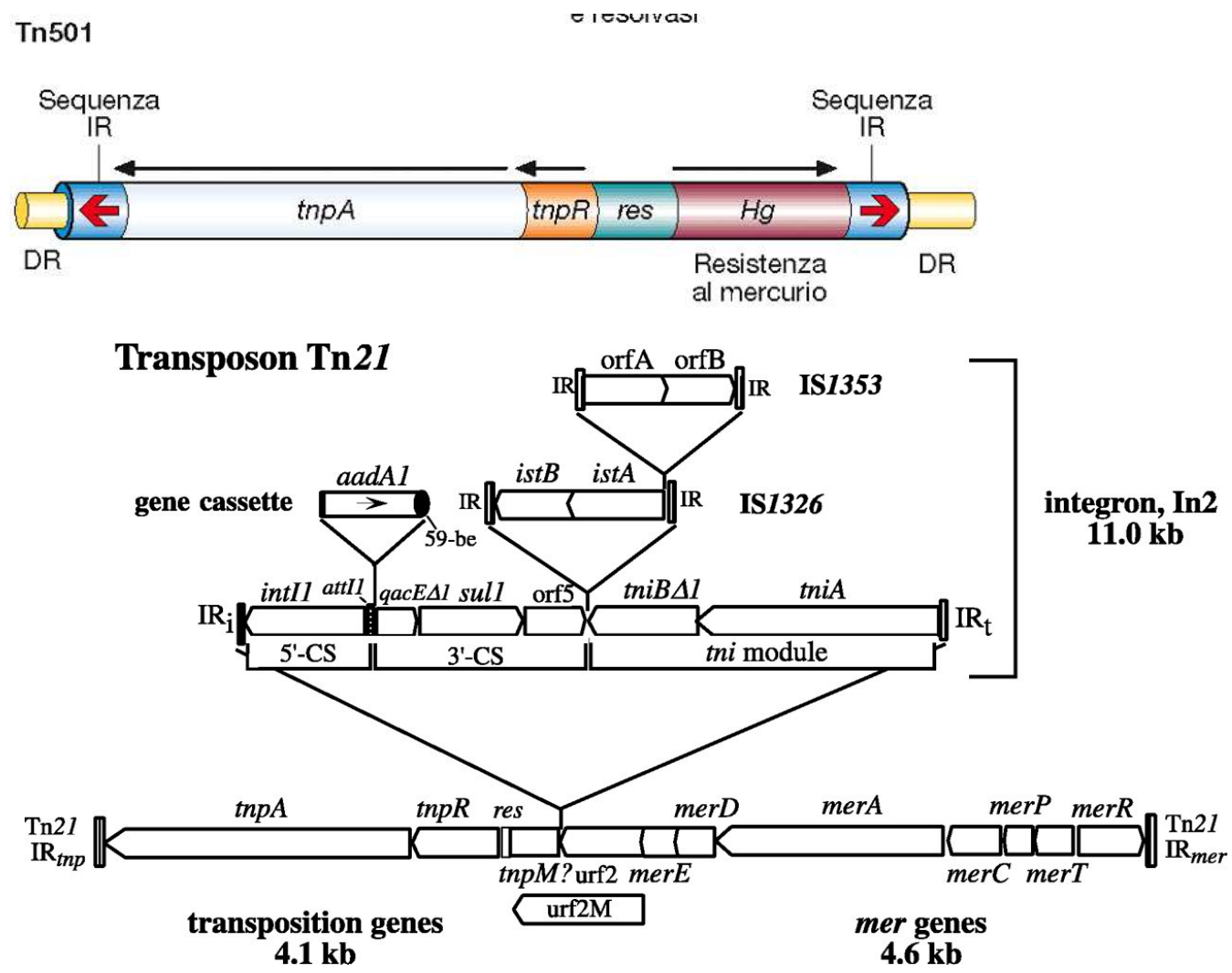


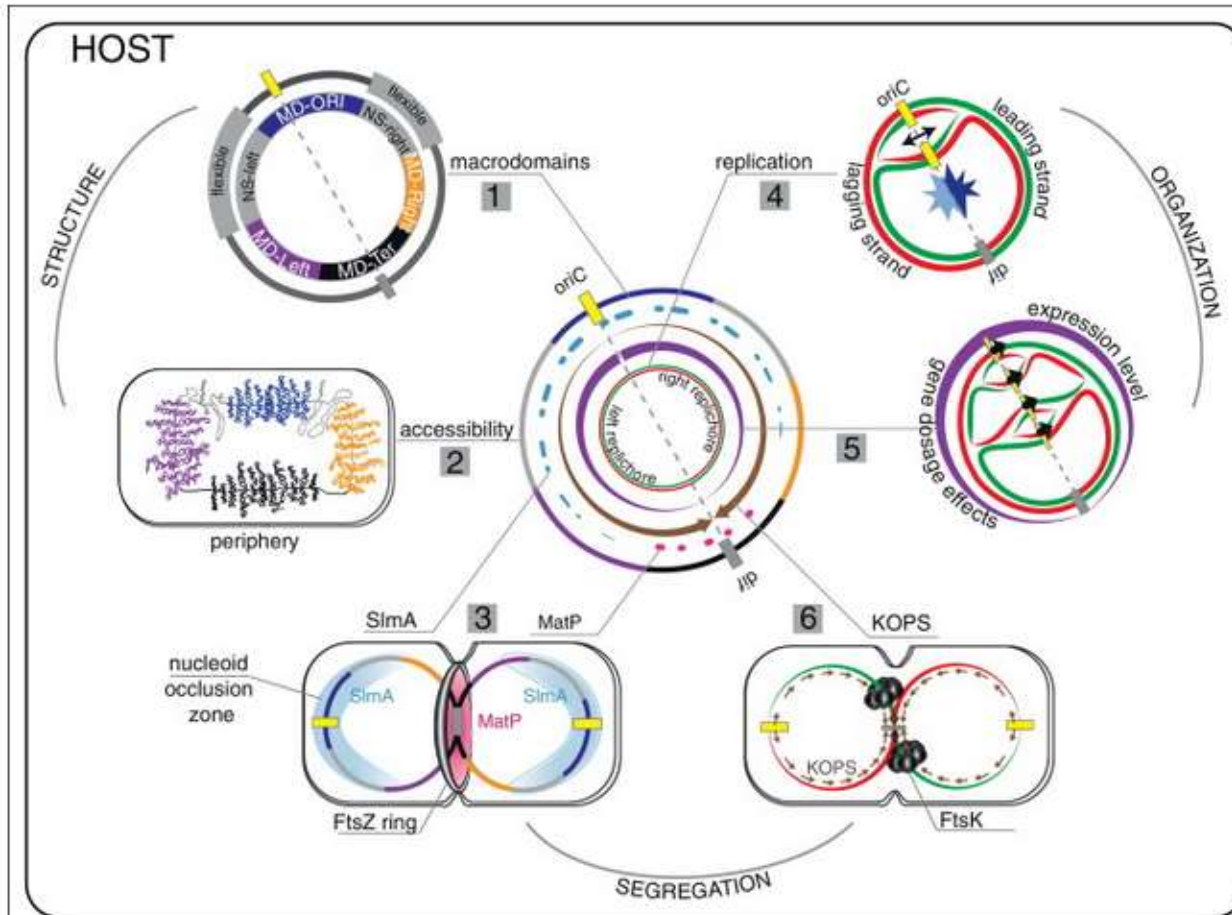
Le cassette geniche

- condividono delle caratteristiche strutturali comuni
- generalmente contengono un singolo gene
- hanno delle sequenze invertite ripetute all'estremità 3' definito sito *attC* o sequenza 59 bp

Gli integroni non sono elementi mobili di per se come le sequenze IS o i Tn ma sono associati ad elementi trasponibili o a plasmidi coniugativi che li possono veicolare all'interno della stessa specie o tra specie diverse.

Da una a molte antibiotico-resistenze: evoluzione di un trasposone





Organizzazione del cromosoma: i domini funzionali



Current Opinion in Microbiology
Volume 22, December 2014, Pages 22-29



The chromosomal accommodation and domestication of mobile genetic elements

Marie Touchon^{1,2}, Louis-Marie Bobay^{1,2,3}, Eduardo PC Rocha^{1,2}

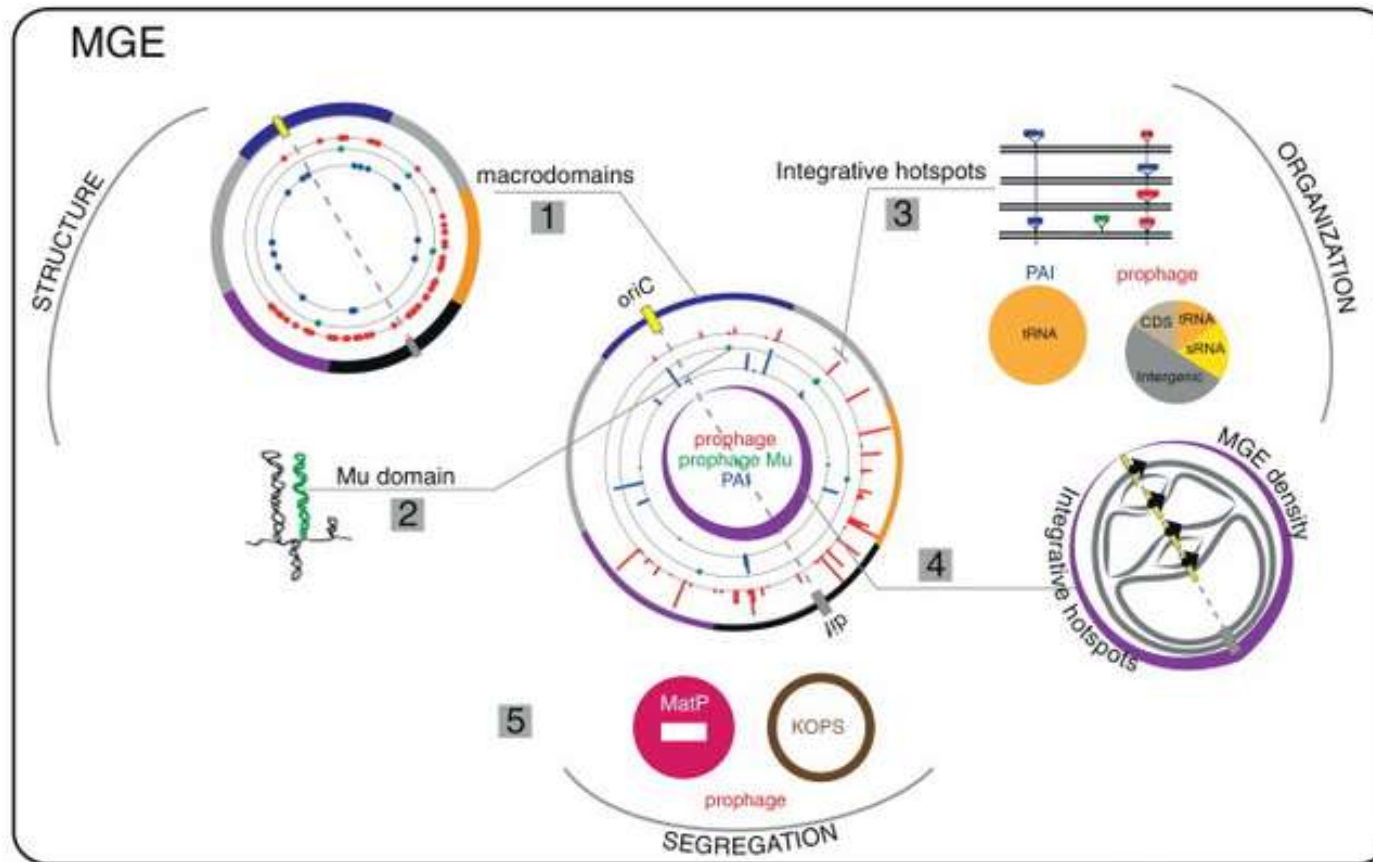
Show more

<https://doi.org/10.1016/j.mib.2014.09.010>

Get rights and content

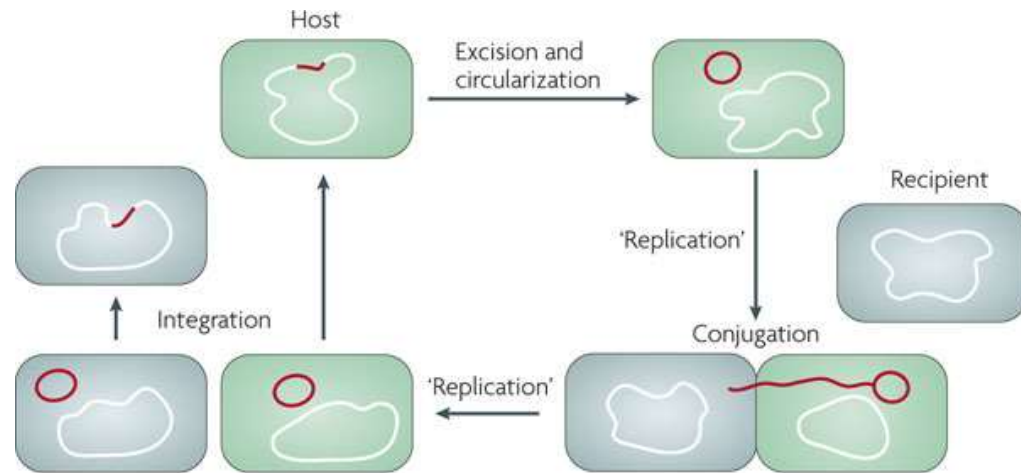
The *E. coli* circular chromosome is organized into 4 structured macrodomains (MD) and 2 unstructured flexible chromosomal regions (NS) [17]. (2) The macrodomain Ter (MD-Ter, black) is preferentially located at the periphery of the nucleoid [25]. (3) MatP defines the MD-Ter by binding to *matS* sequence motifs [34]. SlmA binds to DNA motifs frequent in MD-Ori and the flanking NS-regions to prevent chromosome fragmentation during septum formation [35]. (4) Chromosomal replication proceeds bi-directionally (black arrow) from a single origin (*oriC*, yellow) to the opposite termination site (*dif*, gray). The leading (resp. lagging) strands are represented in green (resp. red) and are different in terms of composition (GC skew), number of genes, and Chi and KOPS motifs [12]. (5) Presence of multiple replication forks in fast-growing bacteria produces a transient replication-associated gene dosage effect that leads to selection of highly expressed genes (violet) near the origin of replication [23]. (6) KOPS (FtsK-orienting polar sequence, brown) are polarized motifs very frequent in the Ter-proximal regions. They orient the last stages of chromosome segregation

Ma dove si inseriscono gli elementi genetici mobili (MGE)



Current Opinion in Microbiology

MGE accommodation . (1) Transposable prophage Mu (green) integrates randomly in the genome, whereas most other prophages (red) integrate more frequently in the MD-Ter and the flanking MD-left and MD-right. Pathogenicity islands (PAI, blue) are more frequent in the other half of chromosome (i.e. MD-Ori and NS-regions) [15]. (2) Mu forms a stable chromosomal domain. (3) Mobile elements using site-specific recombinases are concentrated in a few integrative hotspots. PAI integrated at tRNA genes, whereas prophages also use other targets. (4) The occupancy rate and the number of integrative hotspots increase with the distance from the origin, in inverse relation with replication-associated gene dosage effects. (5) Prophages show avoidance of MatP and over-representation of KOPS that are relevant only at the prophage state in the context of the biology of the host.



Gli ICE : Integrative conjugative elements

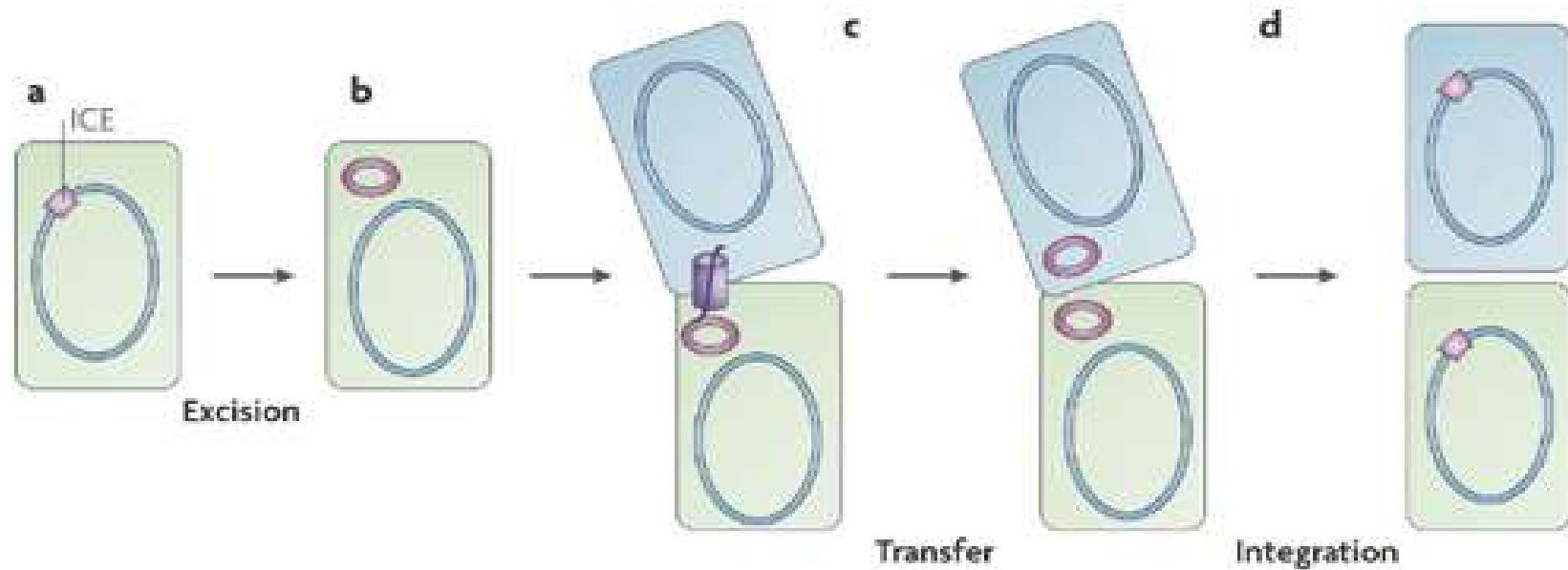
fagi , plasmidi o trasposoni?

Gli ICE si ritrovano integrati nel cromosoma dell'ospite in un sito specifico (attB) e sono fiancheggiati da sequenze specifiche definite attL and attR. In seguito ad excisione precisa tramite ricombinazione sito specifica tra le sequenze attL e attR si ottiene la molecola di ICE circolare che conterà il sito attP mentre sul cromosoma della cellula l'ospite si troverà il sito att B.

Durante il processo di coniugazione con una cellula ICE-free, il donatore ed il recipiente vengono in contatto e una singola elica di ICE viene trasferita tramite il meccanismo a cerchio rotante.

Dopo il trasferimento la DNA polimerasi del recipiente riformerà una molecola circolare a doppia elica.

Tramite ricombinazione sito specifica tra i siti attP dell'ICE e attB del cromosoma l'ICE si integrerà nel cromosoma del recipiente.



Gli ICE sono normalmente integrati nel cromosoma della cellula ospite e sono replicati e trasmessi alle cellule figlie allo stato integrato come fossero regioni del cromosoma del batterico.

In alcune condizioni gli ICE si possono excidere dal cromosoma batterico, formando un intermedio circolare. L'intermedio circolare si replica ed un filamento viene trasferito alle cellula recipiente tramite un poro di coniugazione. I componenti del sistema di trasferimento sono codificati dall'ICE.

PATOGENICITA'

I geni di virulenza sono frequentemente localizzati su elementi genetici mobili o ex mobili quali le isole di patogenicità PAI

I geni acquisiti devono poi inserirsi in sistemi di regolazione che permettano di ottimizzare l'espressione e di coordinarla con gli altri geni di virulenza

Spesso i sistemi a due componenti o le proteine del nucleotide annettono sotto il loro controllo i nuovi geni in modo che vengano espressi in risposta a stimoli ambientali ed in modo coordinato

FITNESS

Determinanti coinvolti nel miglioramento dell'adattamento all'ambiente o nel potenziamento delle strategie di sopravvivenza possono essere trasferiti per HGT e si ritrovano sia in batteri patogeni che in commensali

Come si può arrivare all'identificazione che una data sequenza provenga da HGT?

Prova diretta sperimentale (raramente possibile)

Prova indiretta

Analisi filogenetica :

generare alberi filogenetici per ogni gene analisi particolare per i geni sono molto distanti rispetto al restante genoma

Miglior appaiamento (Best Sequence Match Detection) :

ricerca in data base della omologia di sequenza con altri genomi ed osservazione della distribuzione di questa regione

Analisi della composizione nucleotidica (Nucleotide Compositional Analysis)

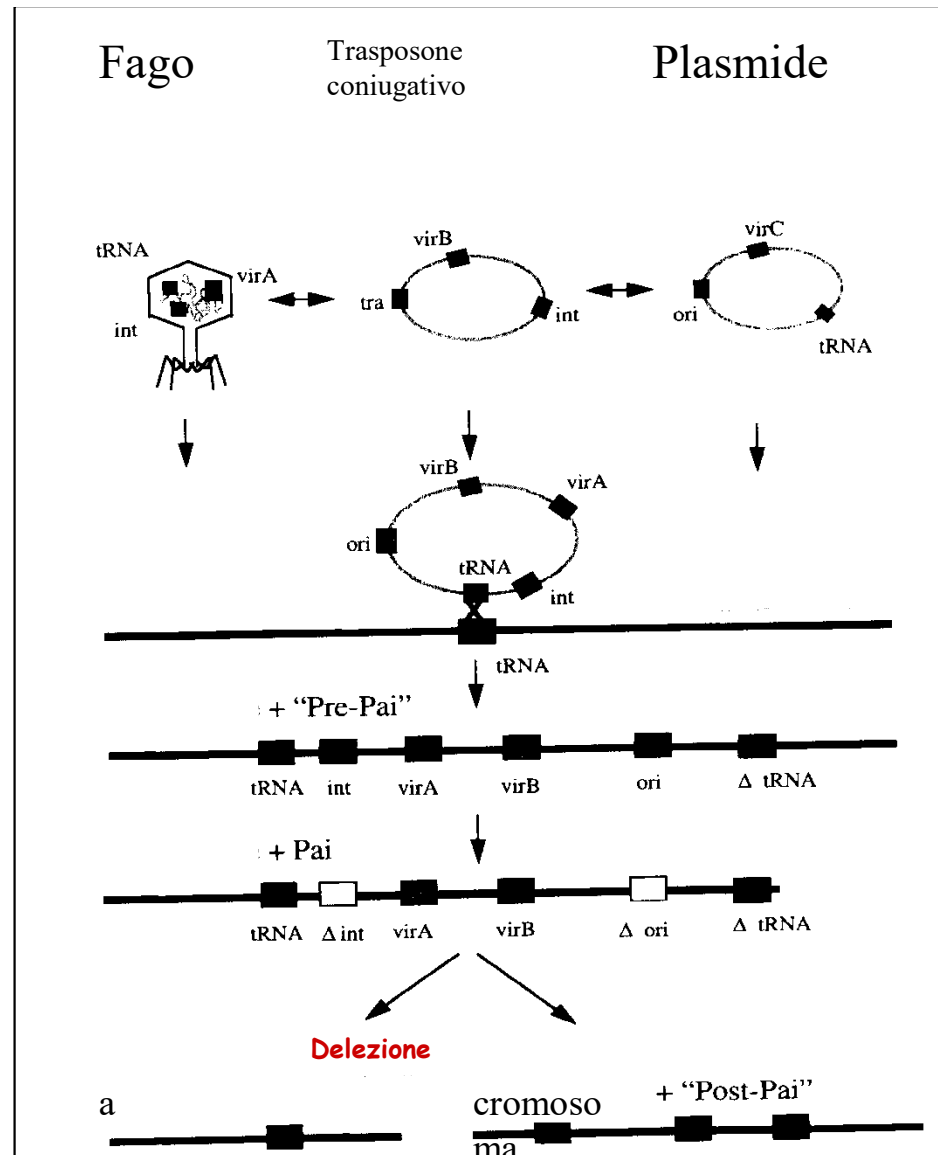
batteri esibiscono una grande variabilità nel contenuto in GC ma i geni all'interno di un genoma sono molto simili in termini di composizione di basi, uso dei codoni. I frammenti di DNA ottenuti tramite HGT mantengono le caratteristiche di sequenza del genoma del donatore

ISOLE DI PATOGENICITA'



- Contengono uno o più geni di virulenza
- Sono presenti solo nei ceppi patogeni
- Sono di grandi dimensioni 10-200 kb
- Hanno un diverso contenuto in G+C (recente HTG)
- Sono spesso inserite in geni per tRNA
- Sono fiancheggiate da sequenze di DNA direttamente ripetute (DR)
- Sono associate a elementi genetici mobili
- Sono instabili
- Rappresentano strutture a mosaico

Percorso evolutivo nell'origine di una PAI



Formazione di un cointegrato

Trasferimento orizzontale

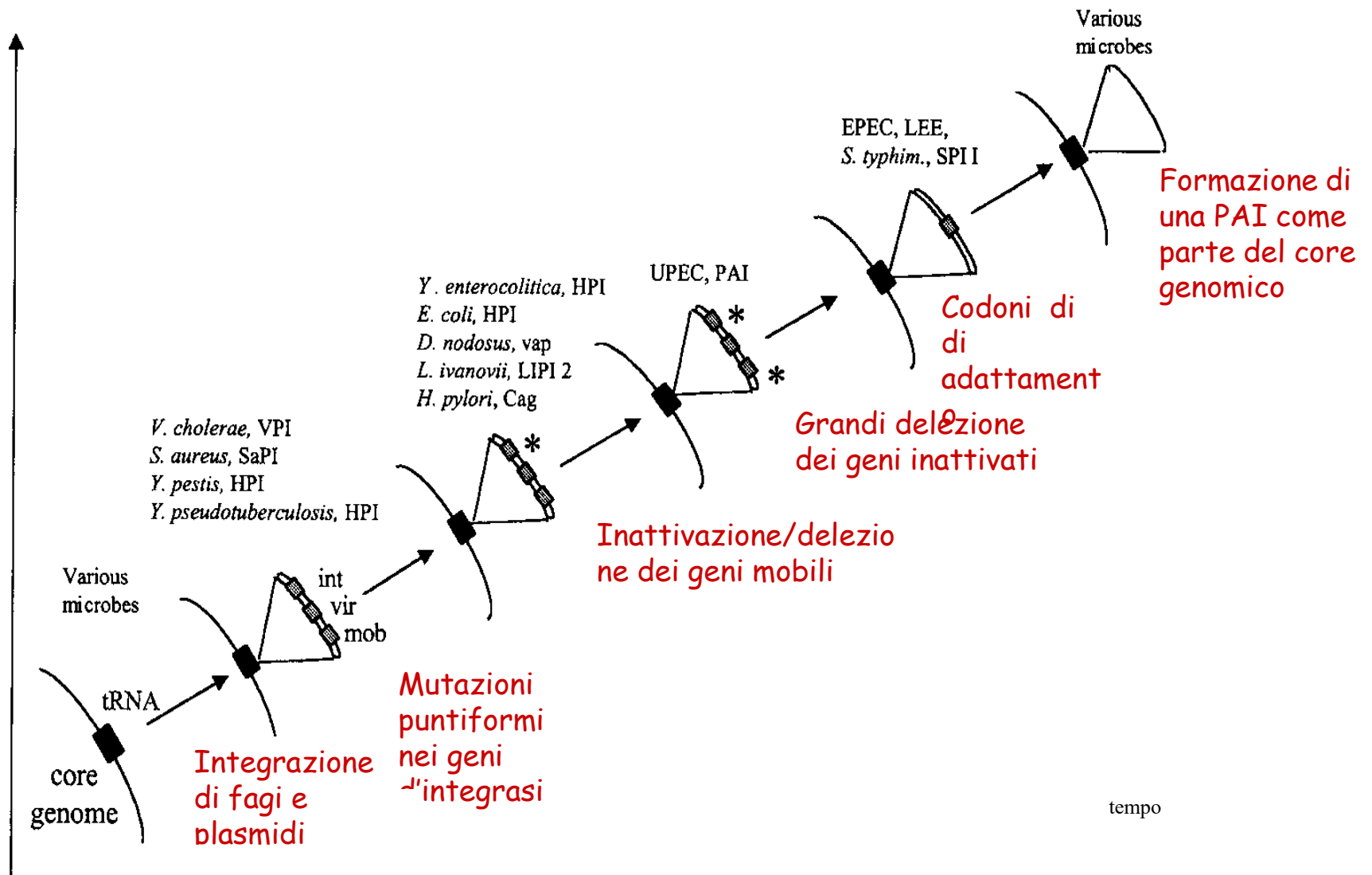
Ricombinazione

Integrazione

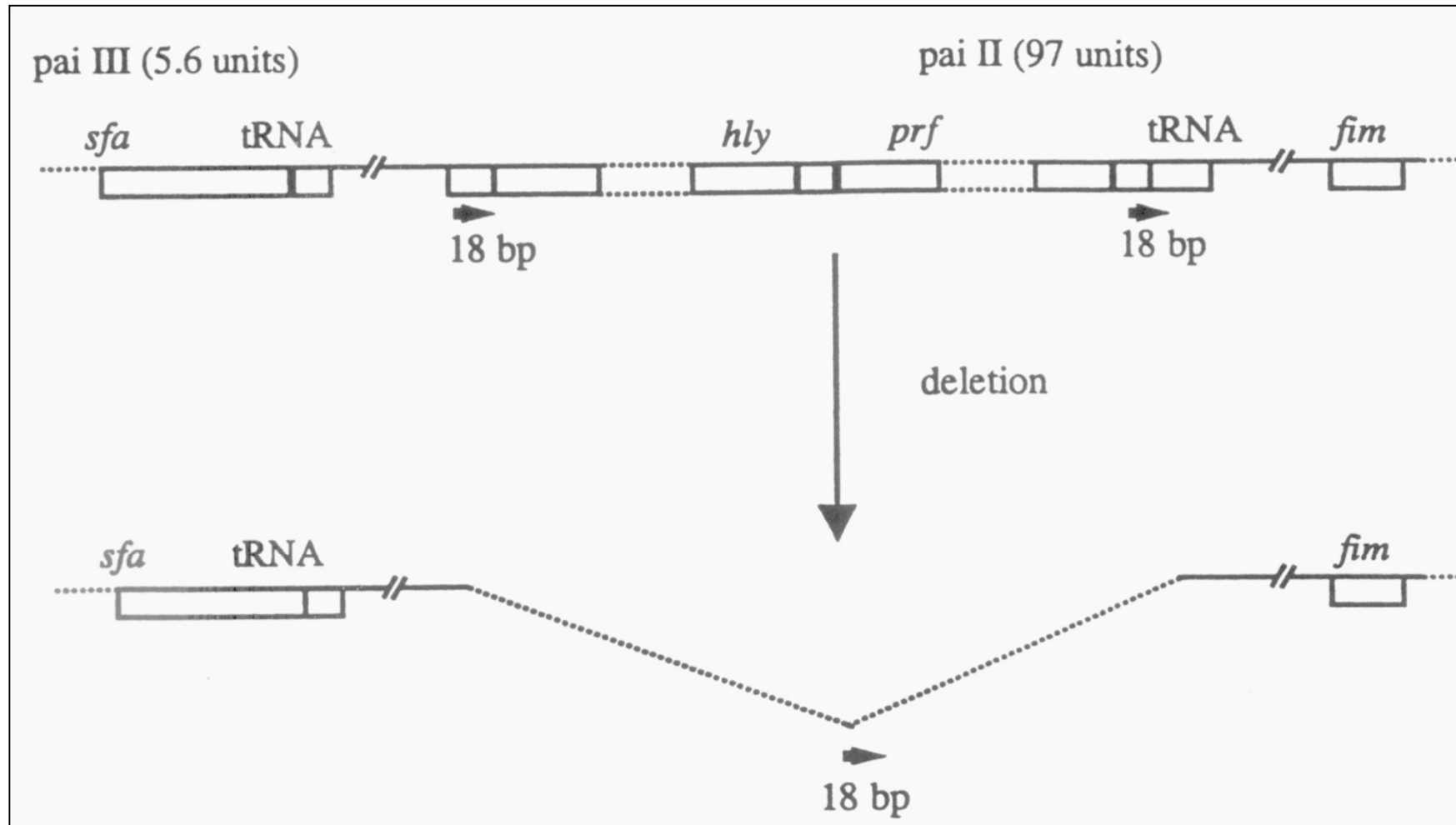
Mutazione, HOMING

Adattamento

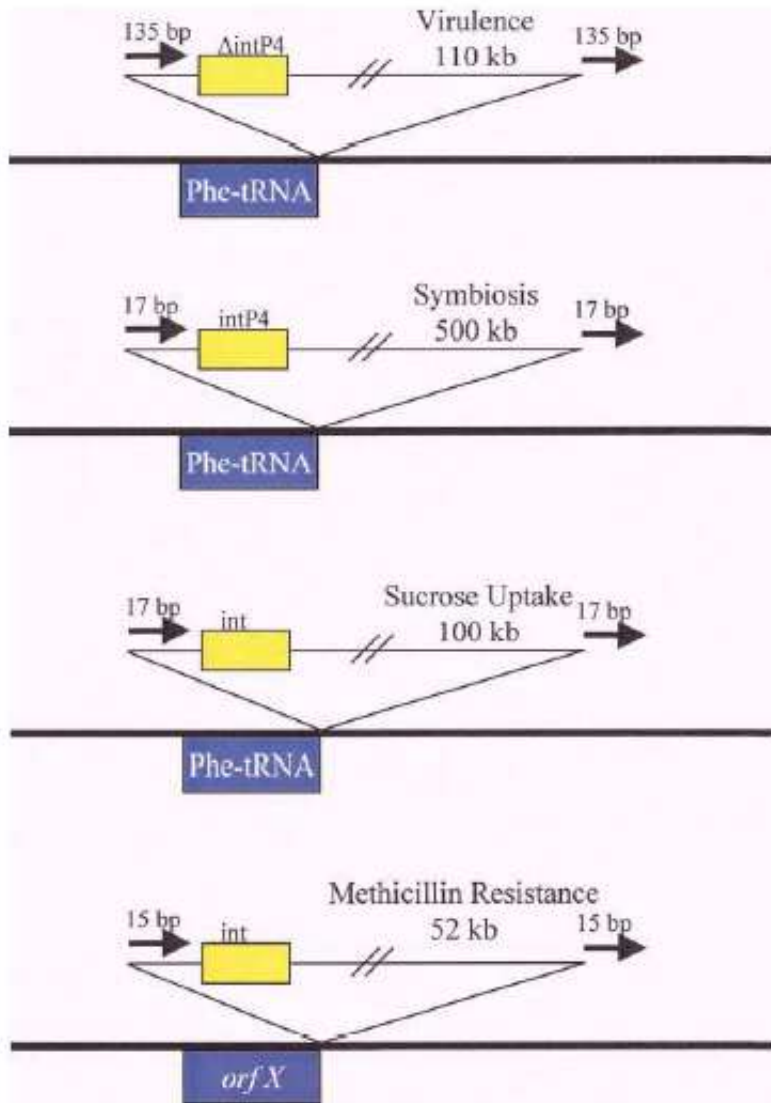
Stadi evolutivi della formazione dell'isola di patogenicità



Le Isole di Patogenicità si possono excidere dal genoma attraverso ricombinazione tra le sequenze presenti nelle sequenze direttamente ripetute (DR)



Non soltanto Pathogenicity Island : il fenomeno coinvolge altri gruppi di geni → Genomic Island



Geni di virulenza

Pathogenicity island
Uropathogenic E.coli

Geni per la simbiosi

Symbiosis Island
Mesorhizobium loti

Geni per l'utilizzazione di zuccheri

Metabolic Island
Salmonella senftenberg

Geni per la resistenza ad uno o più antibiotici

Resistance Island
Staphylococcus aureus

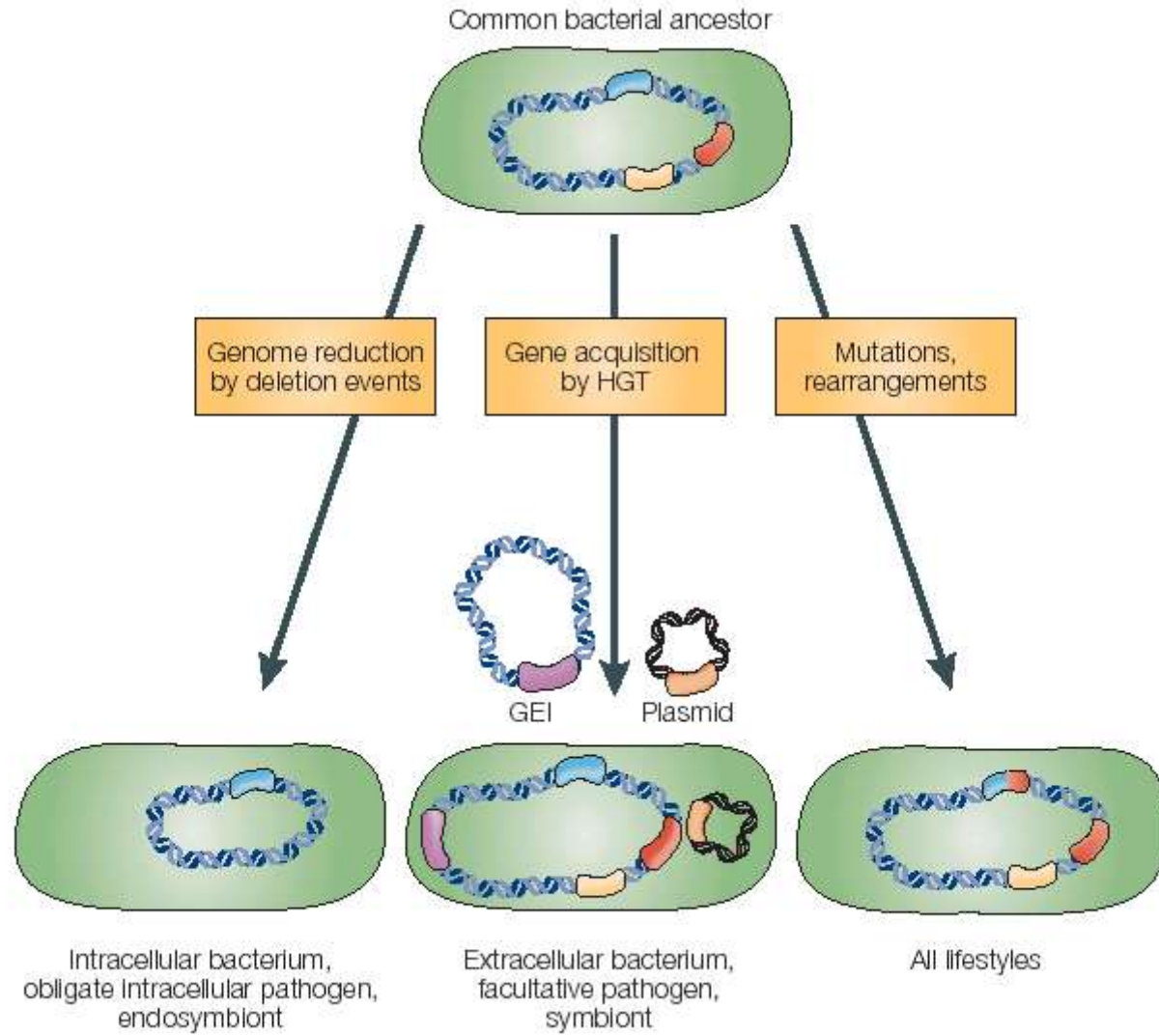
Caratteristiche generali delle Isole genomiche

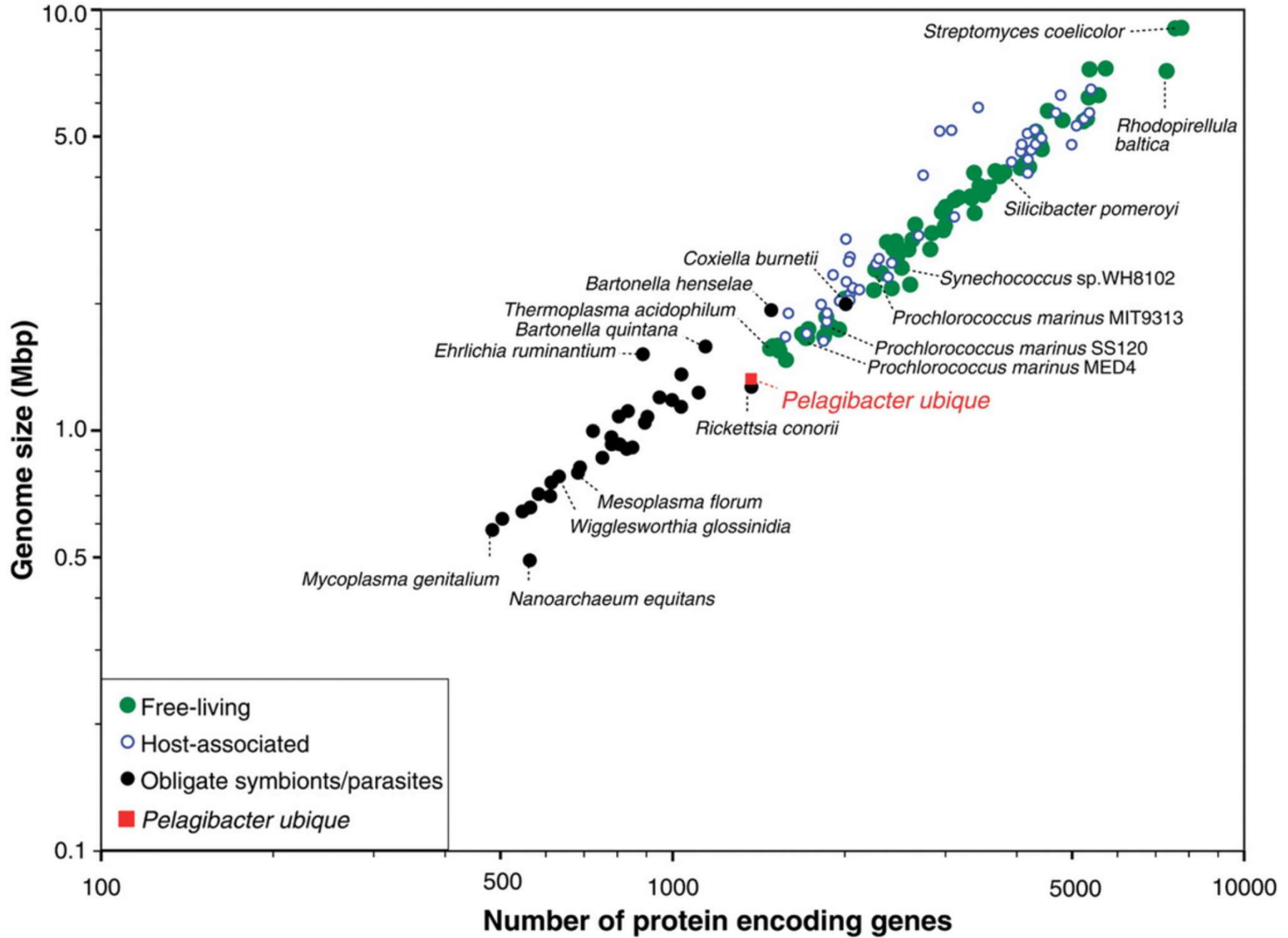
- Sono regioni di DNA acquisite tramite HGT inserite nelle vicinanze dei **tRNA**.
- Sono fiancheggiate da **DR**
- Contengono diversi geni coinvolti nel processo di adattamento
- Contengono elementi IS funzionali o difettivi
- contiene geni legati alla motilità
- Ruolo fondamentale svolto dal gene **INT** che codifica per l'**INTEGRASI** coinvolta nell'inserzione e delezione di regioni di DNA fiancheggiate da DR

Come si identifica un 'isola genomica?

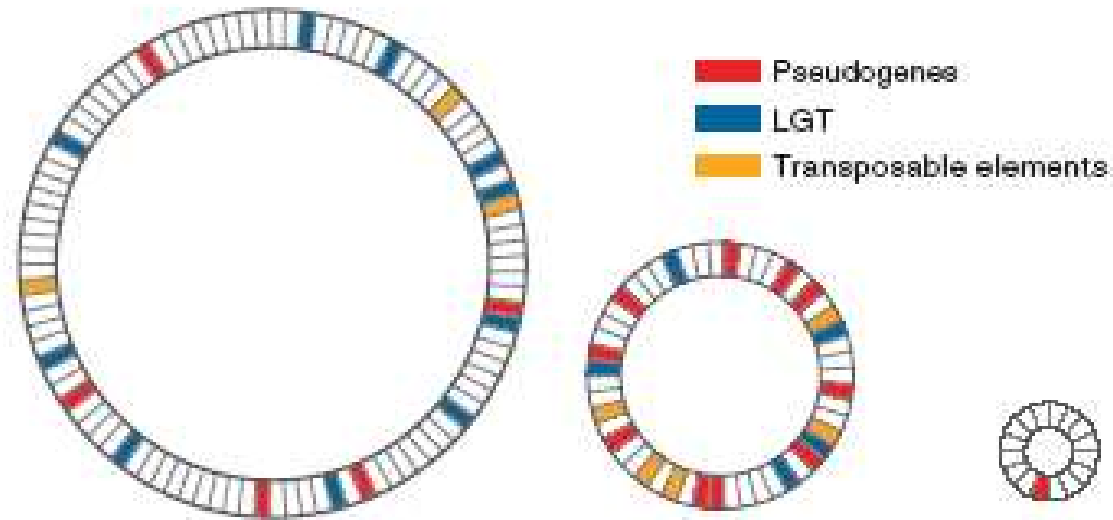
- Analisi comparativa del genoma per ibridazione sottrattiva :
identificazione delle regioni tipiche di un genoma
- tRNA screening: analisi delle regioni fiancheggianti i tRNA
- Rilevazione delle informazioni genetiche acquisite per
HGT: strutture ripetute, IS e trasposoni, analisi della composizione
in base , uso dei codoni etc
- Island probing: analisi della stabilità di una regione usnd
controselezioni idonee

Evoluzione di un microrganismo





Pseudogeni e dimensioni dei genomi



	Free-living	Recent or facultative pathogen	Obligate symbiont or pathogen
Genome size	Large (5-10 MB)	Intermediate (2-5 MB)	Small (0.5-1.5 MB)
Number of pseudogenes	Few	Many	Rare
Incidence of LGT	Frequent	Frequent to rare	Rare to none
Selfish genetic elements	Few	Common	Rare
Genome organization	Stable or unstable	Unstable	Stable
Effective population size	Large	Small	Small

La selezione sembra operare più efficacemente sulle specie che vivono anche nell'ambiente e non solo nell'ospite.

Dal momento che la selezione permette di eliminare i geni deleteri queste specie contengono ampi genomi con un numero relativamente piccolo di pseudogeni o elementi genetici mobili.

In patogeni recenti abbiamo invece un accumulo di pseudogeni a dimostrazione della grande plasticità di questo genomi

In specie che sono da lungo tempo dipendenti dall'ospite invece la tendenza verso l'insorgenza di delezioni ha eliminato tutte le sequenze superflue determinando genomi molto ridotti con pochi pseudogeni o Tn

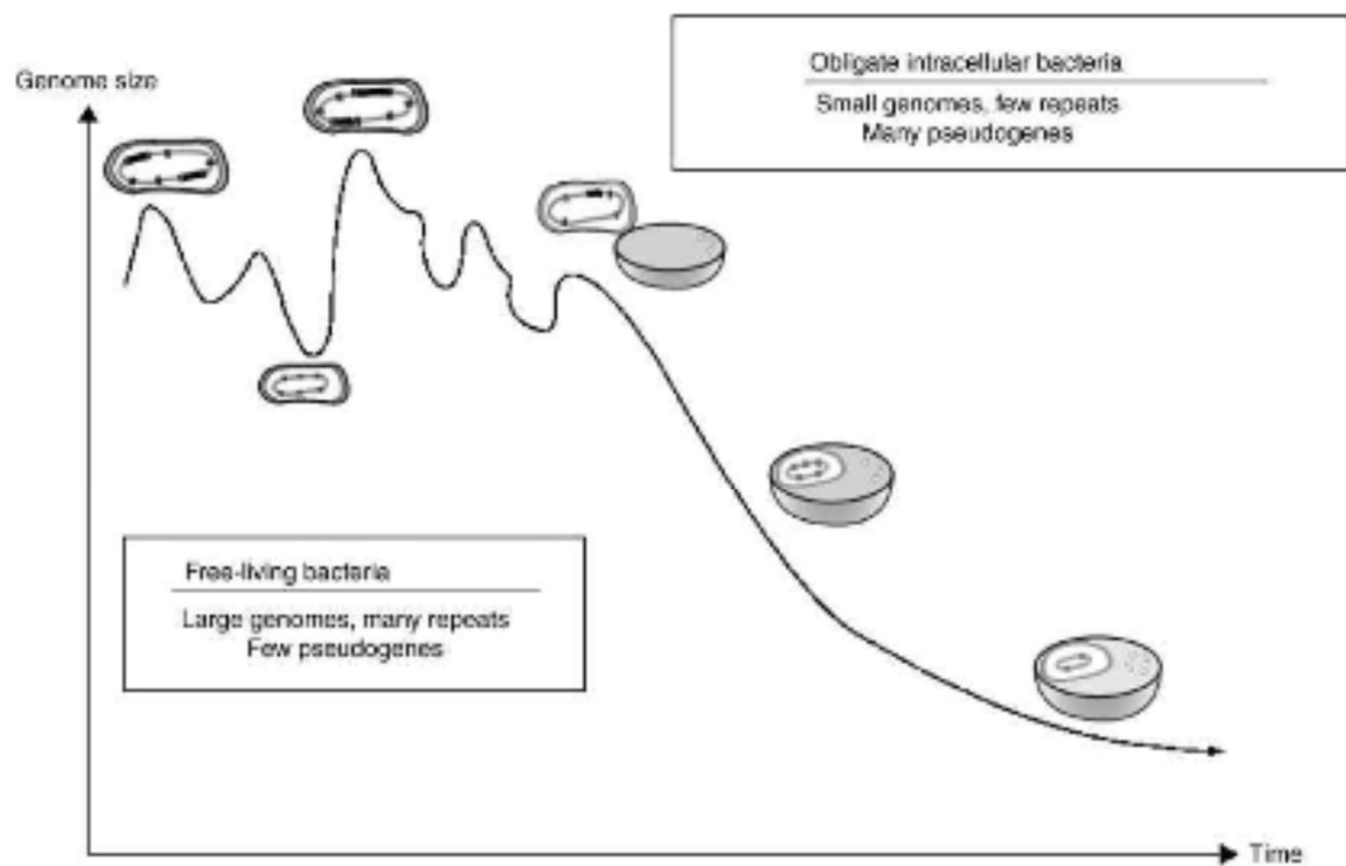
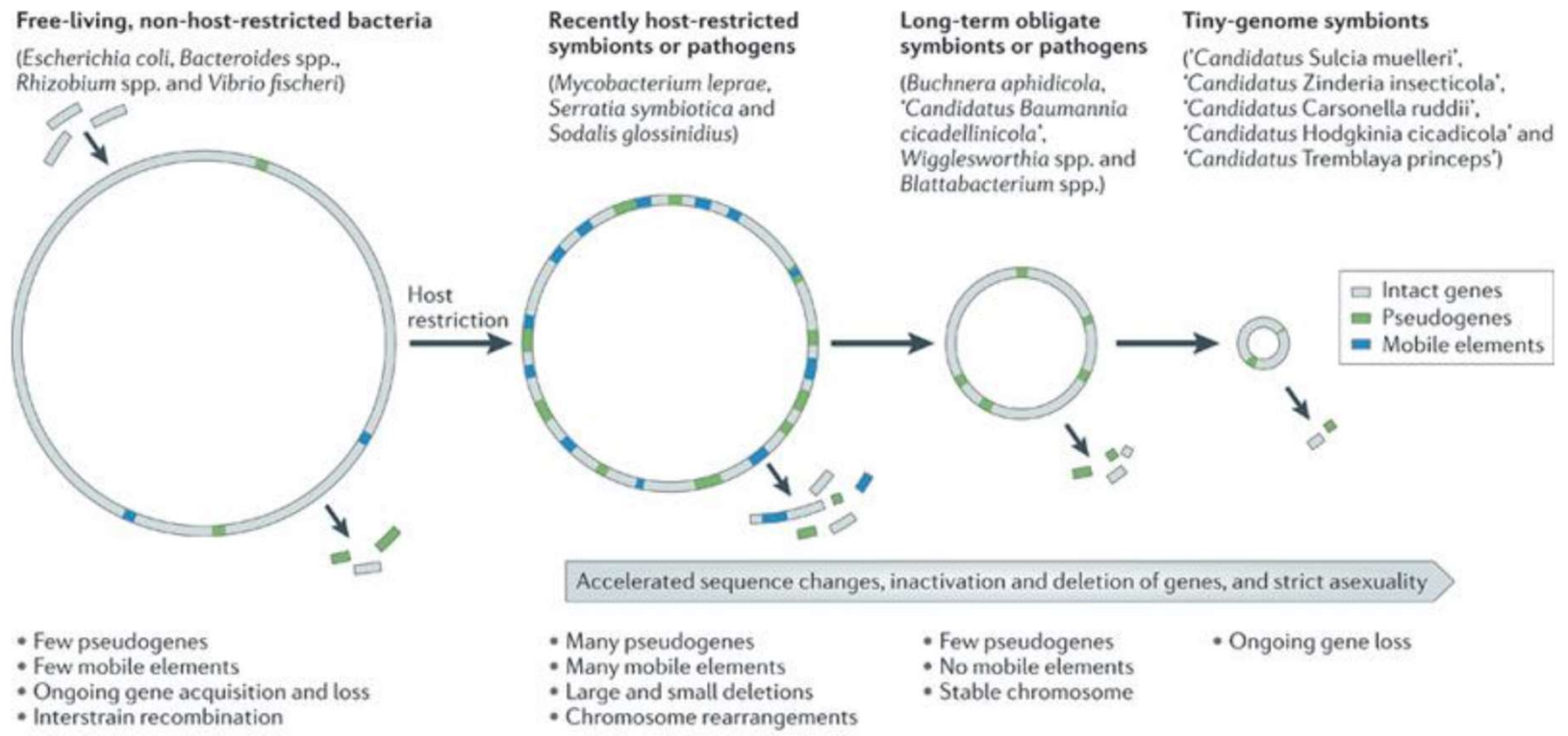


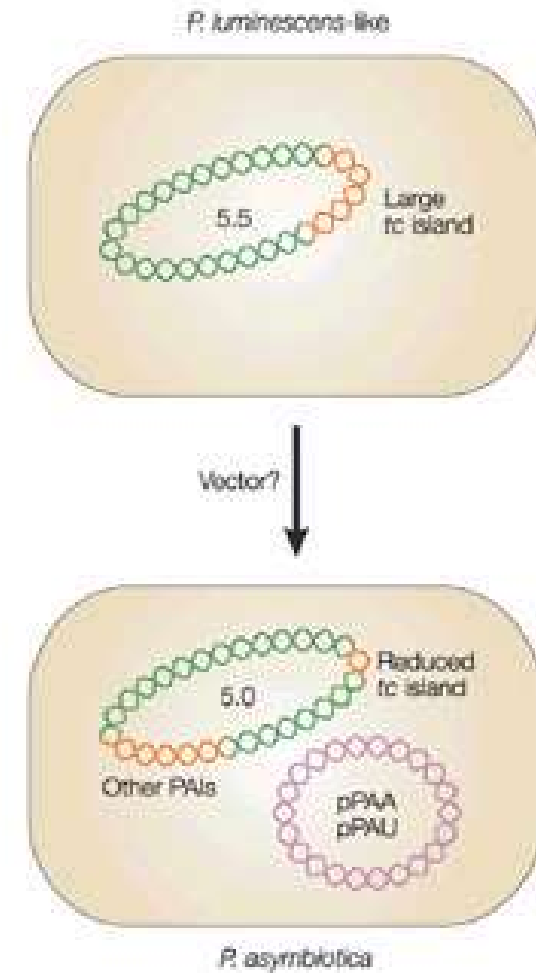
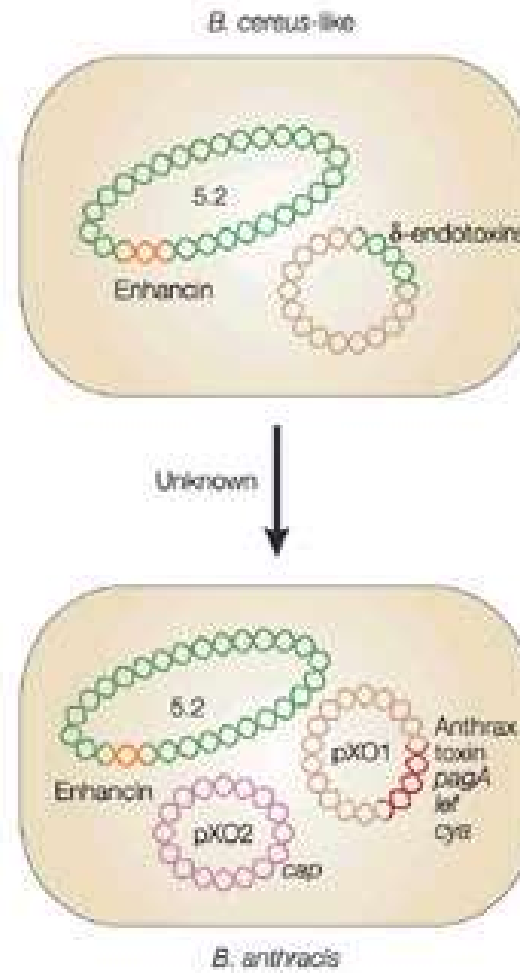
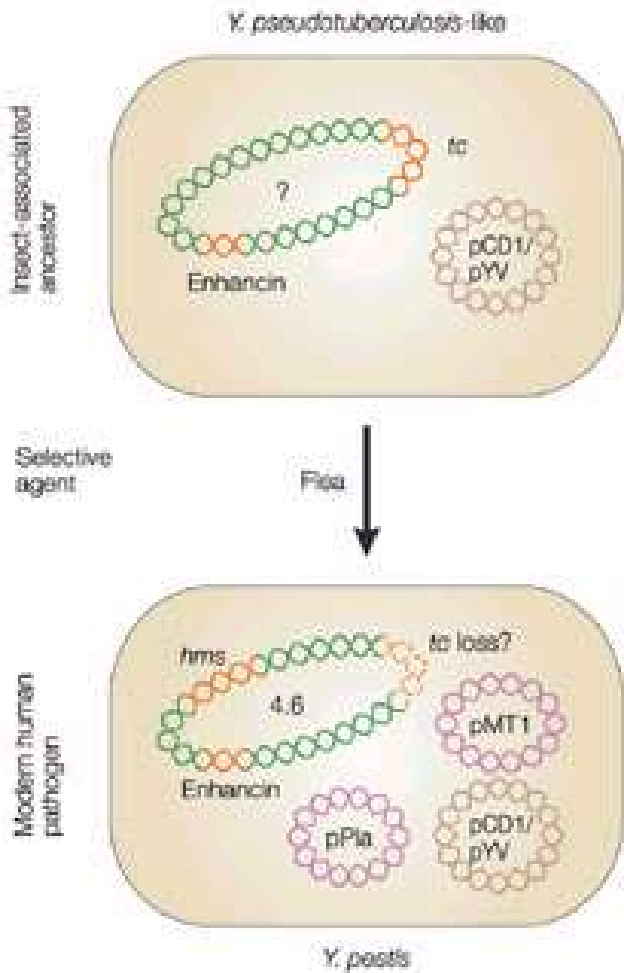
Figure 3. Schematic illustration of genome size variations as a function of time during transitions to intracellular growth habitats. Filled boxes represent mobile genetic elements. Genomes of obligate intracellular bacteria are smaller and have a lower content of repeated sequences (//) and a higher content of pseudogenes (x) than genomes of free-living bacteria and facultative intracellular parasites.



Y. pseudotuberculosis → *Y. pestis*
 Acquisizione di 2 plasmidi necessari per la sopravvivenza nell'uomo e per la trasmissione tramite mosche

Bacillus cereus → *B. anthracis*
 Acquisizione del plasmide codificante la tossina

P. luminescens like → *P. asymbiotica*
 acquisizione di un plasmide



Riduzione del genoma

L'acquisizione di geni è controbilanciata dalla perdita di geni nativi: l'insorgenza di delezioni costituisce una delle principali forze per il modellamento del genoma

Perdita di funzioni geniche  Vantaggio selettivo

Molti pseudogeni non espressi in *Y.pestis* sono funzionali in altre specie di *Yersinia* la perdita dei geni contribuisce all'adattamento di *Y.pestis* al vettore (insetto) che è un prerequisito per la trasmissione del patogeno all'uomo tramite le mosche

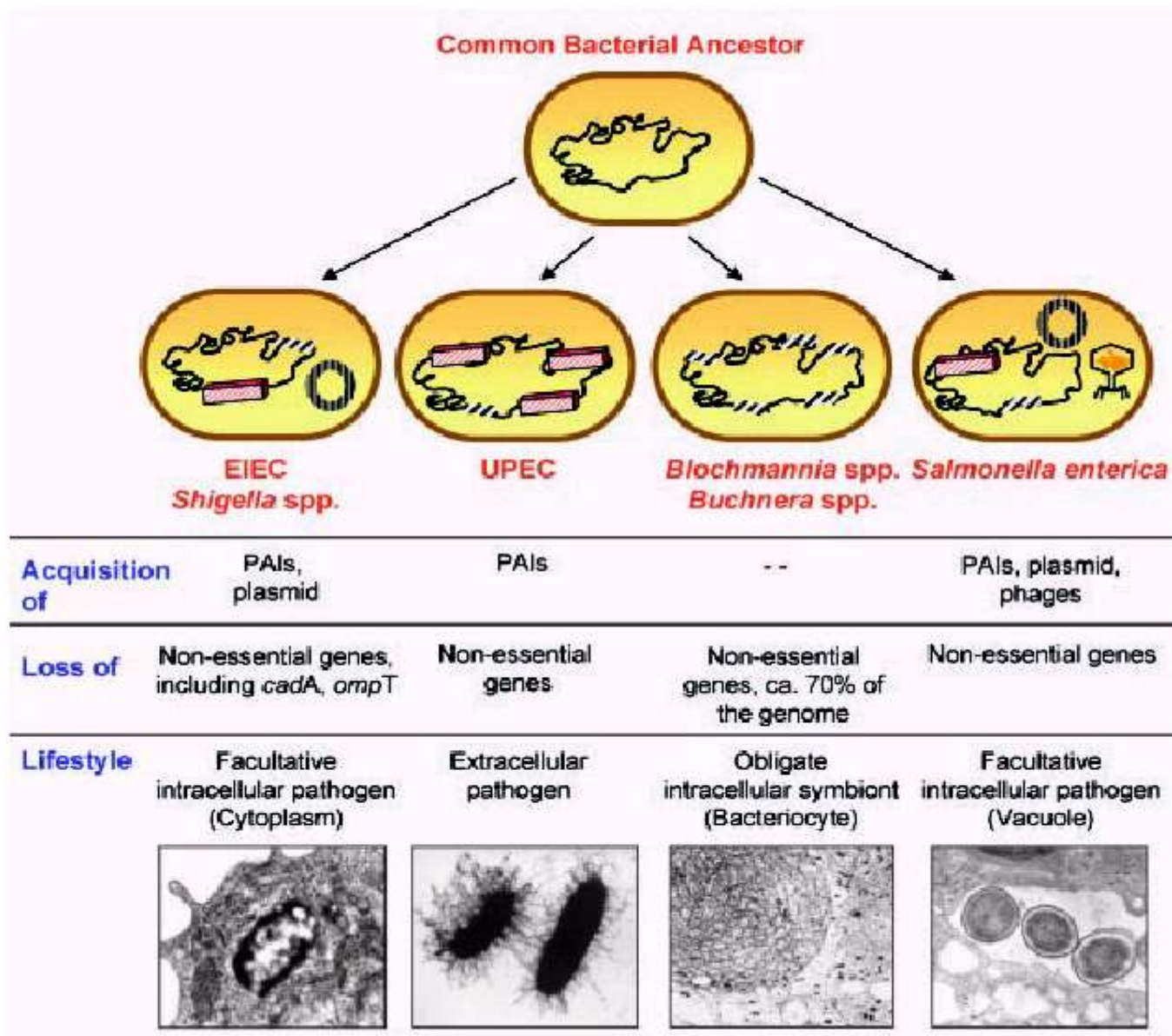
PATOGENI INTRACELLULARI OBBLIGATI

Clamydia
Rickettsia
Buchnera aphidicola
Blochmania



Perdita di notevoli quantità di materiale genetico, programmata in modo da assicurare la persistenza a lungo termine nell'ospite

Modellamento del genoma : le strategie



Genoma di *E.coli* : esempio di grande variabilità

La sequenza dell'intero genoma di *E.coli* ha rivelato una variabilità intraspecie estremamente elevata

Sono disponibili 4 sequenze genomiche di *E.coli*

Analisi genomica comparativa ha rivelato che *E.coli* 0157 ha un genoma di 1 Mb più grande di quella di *E.coli* K12 e circa 25% dei geni non sono conservati nel genoma di *E.coli* K12.

Molti dei geni presenti in 0157 si pensa siano stati acquisiti tramite eventi di trasferimento orizzontale e tramite elementi genetici mobili quali fagi, profagi e sequenze IS

Soltanto 3.000 geni sono in comune tra i 4 genomi di *E.coli* mentre erano 4.000 tra *E.coli* K12 e 0157

I 3000 geni comuni presentano SINTENIA suggerendo una base di trasmissione verticale

L'intuizione di Carl Woese di utilizzare il gene che codifica per RNA16S per la costruzione degli alberi filogenetici è ancora valida e resta un'analisi necessaria per stabilire la filogenesi tra gruppi tassonomici .

La struttura dell'RNA ribosomiale è tale che alcune regioni implicate nella formazione delle strutture tridimensionali complesse sono estremamente conservate e le loro sequenze si sono evolute lentamente differenziandosi molto poco tra gruppi filogeneticamente distanti