

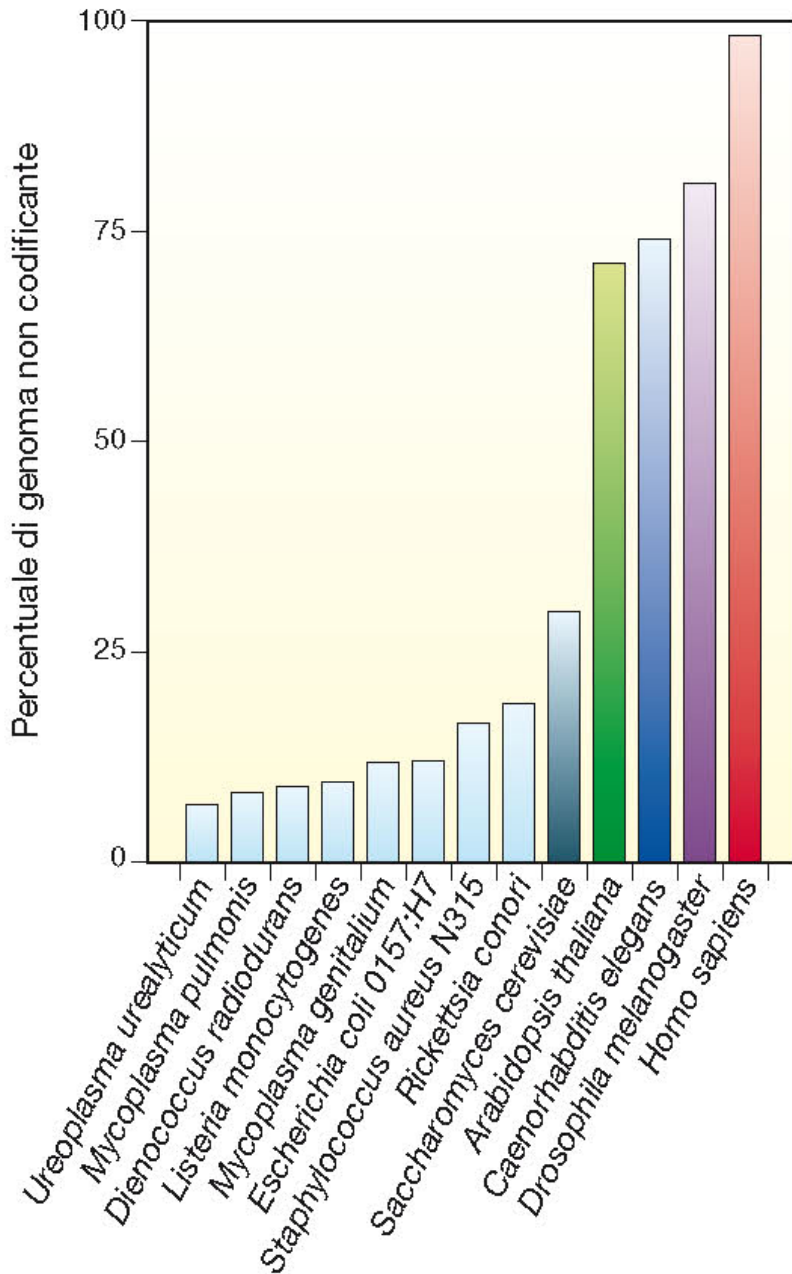
**Organismo****Dimensioni  
(in paia di basi)****ORF<sup>b</sup>****Commenti****Genomi di Batteri sequenziati****Bacteria**

<i>Mycoplasma genitalium</i>	580 070	470	Il più piccolo genoma cellulare noto (➤ vol. 2A, cap. 16.21)
<i>Mycoplasma pneumoniae</i>	816 394	677	Causa la polmonite (➤ vol. 2A, cap. 16.21)
<i>Borrelia burgdorferi</i>	910 725	853	È uno spirochete, possiede un genoma lineare e causa la malattia di Lyme (➤ vol. 2A, cap. 16.33 e vol. 2B, cap. 31.4)
<i>Chlamydia trachomatis</i>	1 042 519	894	Parassita intracellulare obbligato, patogeno umano comune (➤ vol. 2A, cap. 16.27 e vol. 2B, cap. 30.13)
<i>Rickettsia prowazekii</i>	1 111 523	834	Parassita intracellulare obbligato, causa il tifo epidemico (➤ vol. 2A, cap. 16.13 e vol. 2B, cap. 31.3)
<i>Treponema pallidum</i>	1 138 006	1041	Spirochete, causa la sifilide (➤ vol. 2A, cap. 16.33 e vol. 2B, cap. 30.12)
<i>Aquifex aeolicus</i>	1 551 335	1544	Ipertermofilo, autotrofo (➤ vol. 2A, cap. 16.37)
<i>Prochlorococcus marinus</i>	1 657 990	1716	Il più abbondante fototrofo nell'oceano (➤ vol. 2A, capp. 16.26 e 20.6)
<i>Helicobacter pylori</i>	1 667 867	1590	Causa l'ulcera peptica (➤ vol. 2B, cap. 30.10)
<i>Streptococcus pyogenes</i>	1 852 442	1752	Patogeno, causa la faringite e la scarlattina (➤ vol. 2B, cap. 30.2)
<i>Thermotoga maritima</i>	1 860 725	1877	Ipertermofilo (➤ vol. 2A, cap. 16.36) <i>vedi</i> anche fig. 13.7
<i>Chlorobium tepidum</i>	2 154 946	2288	Batterio fototrofico modello (➤ vol. 2A, cap. 16.32)
<i>Staphylococcus aureus</i>	2 814 816	2593	Principale responsabile delle infezioni nosocomiali (➤ vol. 2A, cap. 16.19 e vol. 2B, cap. 29.7)
<i>Deinococcus radiodurans</i>	3 284 156	2185	Resistente alle radiazioni, contiene cromosomi multipli (➤ vol. 2A, cap. 16.34)
<i>Synechocystis</i> sp.	3 573 470	3168	Cianobatterio (➤ vol. 2A, cap. 16.25)
<i>Bdellovibrio bacteriovorus</i>	3 782 950	3584	Parassita di altri procarioti (➤ vol. 2A, cap. 16.14)
<i>Geobacter sulfurreducens</i>	3 814 139	3467	Modello per il biorisanamento (➤ vol. 2A, cap. 18.18)
<i>Caulobacter crescentus</i>	4 016 942	3767	Ciclo di vita complesso (➤ vol. 2A, cap. 16.16)
<i>Bacillus subtilis</i>	4 214 810	4100	Modello genetico per i Gram-positivi (➤ vol. 2A, cap. 16.20)
<i>Mycobacterium tuberculosis</i>	4 411 529	3924	Causa la tubercolosi (➤ vol. 2A, cap. 16.23 e vol. 2B, cap. 30.5)
<i>Escherichia coli</i>	4 639 221	4288	Modello genetico per i Gram-negativi (➤ vol. 2A, cap. 16.11)
<i>Bacillus anthracis</i>	5 227 293	5738	Patogeno, agente per la guerra biologica (➤ vol. 2A, cap. 16.20)
<i>Rhodospseudomonas palustris</i>	5 459 213	4836	Fototrofo anossigenico metabolicamente versatile (➤ vol. 2A, cap. 16.2)
<i>Pseudomonas aeruginosa</i>	6 264 403	5570	Patogeno opportunistico metabolicamente versatile (➤ vol. 2A, cap. 16.7)
<i>Streptomyces coelicolor</i>	8 667 507	7825	Possiede un cromosoma lineare, produce antibiotici (➤ vol. 2A, cap. 16.24)
<i>Bradyrhizobium japonicum</i>	9 105 828	8317	Fissatore di azoto, forma noduli nelle piante di soia (➤ vol. 2A, cap. 20.22)

# Genomi degli Archea sequenziati

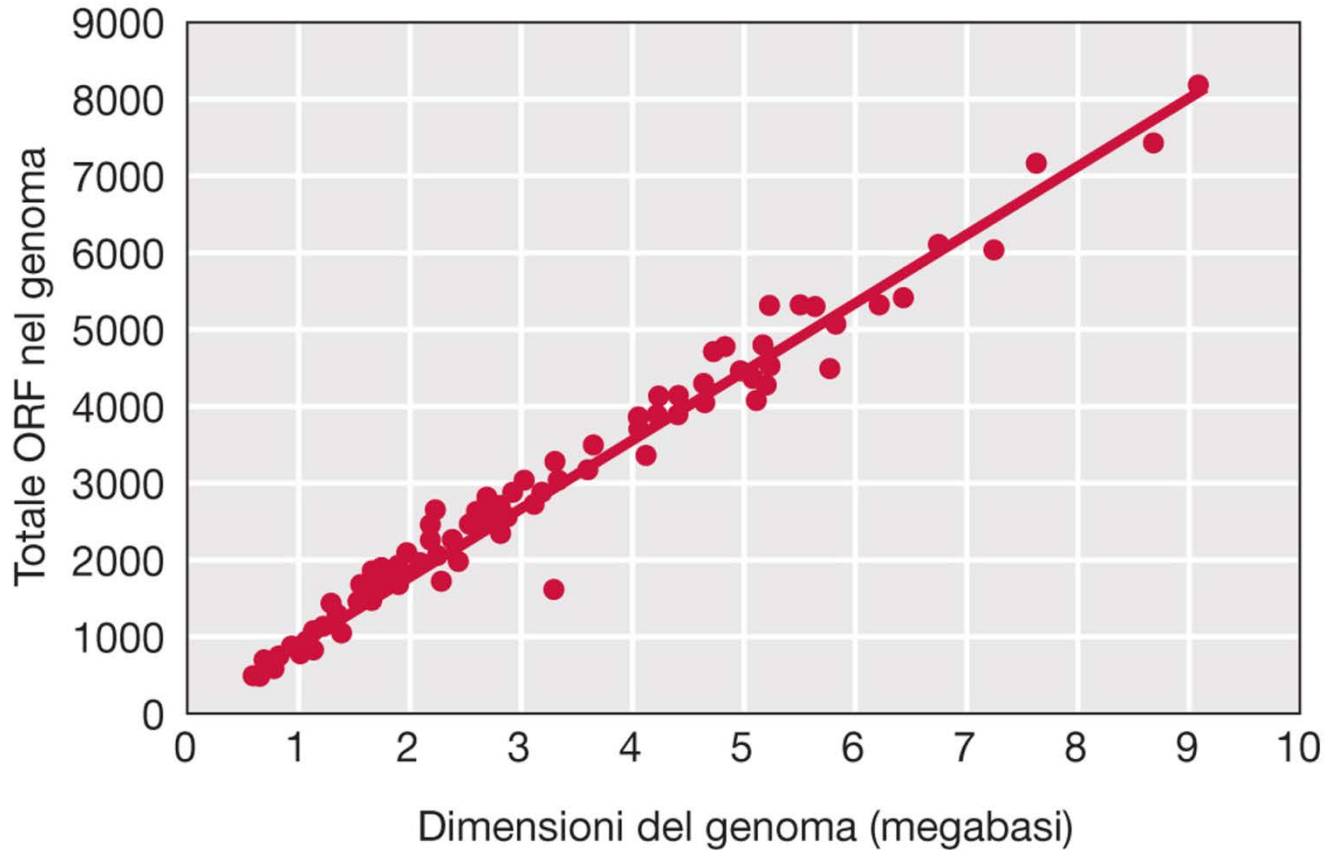
## Archaea

<i>Nanoarchaeum equitans</i>	490 885	552	Il più piccolo genoma conosciuto (▶ vol. 2A, cap. 17.11)
<i>Thermoplasma acidophilum</i>	1 564 905	1509	Termofilico, acidofilo (▶ vol. 2A, cap. 17.5)
<i>Methanocaldococcus jannaschii</i>	1 664 976	1738	Metanogeno (▶ vol. 2A, cap. 17.4)
<i>Aeropyrum pernix</i>	1 669 695	1841	Ipertermofilo (▶ vol. 2A, cap. 17.9)
<i>Pyrococcus horikoshii</i>	1 738 505	2061	Ipertermofilo (▶ vol. 2A, cap. 17.6)
<i>Methanothermobacter thermoautotrophicus</i>	1 751 377	1855	Metanogeno (▶ vol. 2A, cap. 17.4)
<i>Archeoglobus fulgidus</i>	2 178 400	2436	Ipertermofilo (▶ vol. 2A, cap. 17.7)
<i>Halobacterium salinarum</i>	2 571 010	2630	Alofilo estremo, contiene bacteriorodopsina (▶ vol. 2A, cap. 17.3)
<i>Sulfolobus solfataricus</i>	2 992 245	2977	Ipertermofilo, chemiolitotrofo sulfureo (▶ vol. 2A, cap. 17.9)



Vi è quindi un enorme disparità nelle capacità codificanti dei genomi dei procarioti ( l'intero genoma) rispetto a quello del genoma umano ( 2%)

## Correlazione tra dimensioni del genoma e numero di ORF nei procarioti



Dati derivanti dall'analisi di 115 genomi procariotici ( Batteri ed Archea)

*Mycoplasma genitalium* e *Mycoplasma pneumoniae* sono microrganismi con un genoma estremamente piccolo.

Le 470 ORF di *M. genitalium* sono presenti anche in *M. pneumoniae*

Attraverso mutagenesi e analisi comparativa con altri genomi si è visto che sono necessari

**300 geni codificanti per stabilire la minima funzionalità di una cellula.**

Non si conoscono finora genomi con un numero inferiore di ORF.

Il più piccolo genoma finora identificato è quello di un Archea *Nanoarchaeum equitans* il cui genoma è di 90 kb più piccolo di quello di *M.genitalium*.

*N.equitans* pur essendo privo di geni per la sintesi di proteine coinvolte nei processi di catabolismo ed anabolismo ha un genoma interamente codificante ed un numero di ORF superiore a *M.genitalium*

### Genomi molto grandi

*Bradyrhizobium japonicum* ( responsabile della fissazione dell'azoto) contiene 2800 ORF in più rispetto al genoma di *Saccharomyces cerevisiae*.

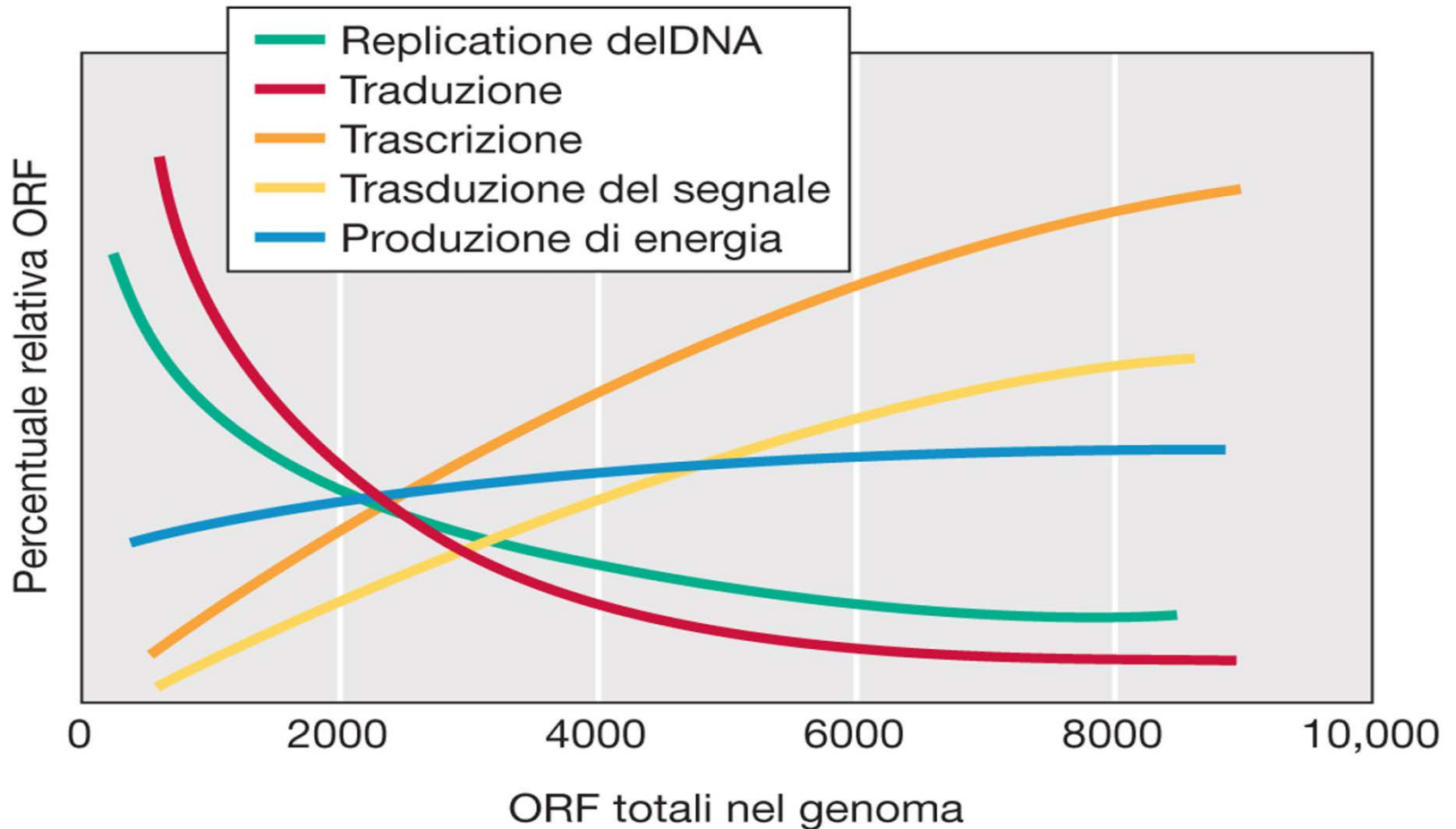
*Streptomyces coelicolor* ha un genoma di 8 Mb, 7846 ORF circa 1800 in più rispetto a *S. cerevisiae*.

## Funzione genica nei genomi batterici : rapporto funzione dimensioni

<b>Categorie funzionali</b>	<b>Percentuale di geni sul cromosoma nella categoria di riferimento</b>		
	<b><i>Escherichia coli</i> (4,64 Mbp)<sup>a</sup></b>	<b><i>Haemophilus influenzae</i> (1,83 Mbp)<sup>a</sup></b>	<b><i>Mycoplasma genitalium</i> (0,58 Mbp)<sup>a</sup></b>
Metabolismo	21,0	19,0	14,6
Strutturali	5,5	4,7	3,6
Trasporto	10,0	7,0	7,3
Regolazione	8,5	6,6	6,0
Traduzione	4,5	8,0	21,6
Trascrizione	1,3	1,5	2,6
Replicazione	2,7	4,9	6,8
Altri noti	8,5	5,2	5,8
Sconosciuti	38,1	43,0	32,0

<sup>a</sup> Dimensioni dei cromosomi. Ognuno dei microrganismi elencati contiene un solo cromosoma circolare.

Geni coinvolti nella sintesi proteica sono essenziali : più i genomi sono piccoli maggiore è la componente percentuale dei geni dedicati ai processi di traduzione.





Organismi di dimensioni maggiori mostrano un numero più elevato di geni coinvolti nei processi di trascrizione o di regolazione rispetto ai microrganismi con genomi più piccoli.

Questi meccanismi regolativi permettono alla cellula di rispondere in maniera migliore alla disponibilità di substrati diversi attraverso l'espressione di geni specifici.

Gli organismi di piccole dimensioni fanno a meno di questi processi regolativi e sono in genere parassiti

## *Microrganismi con genomi di grandi dimensioni*

Gli organismi con genomi di grandi dimensioni possiedono la capacità di codificare per molti geni coinvolti sia nel metabolismo che nei processi regolativi.

I meccanismi di regolazione permettono alla cellula di rispondere in modo migliore alla disponibilità di substrati differenti attraverso l'attivazione di geni specifici.

*Uno degli habitat principali dei microrganismi è il suolo.*

Tutti i microrganismi con genomi di dimensioni maggiori alle 6 Mbp sono microrganismi del suolo.

Il suolo è un habitat nel quale le fonti di carbonio e di energia sono spesso scarse o disponibili in forme diverse e spesso fruibili in maniera intermittente.

## I genomi dei Batteri sono caratterizzati

- un numero elevato di geni per il metabolismo dei carboidrati
- un numero significativo di geni per funzioni correlate alla membrana
- un elevato numero di geni ancora a funzione sconosciuta
- un elevato numero di proteine ipotetiche

### Rispetto agli Archea

un numero minore di geni per il metabolismo dei coenzimi

un numero minore di geni per la produzione di energia

Oltre alle dimensioni del genoma anche l'appartenenza al Dominio ( Batteri o Archea ) sembra influenzare la categorizzazione funzionale dei geni nei procarioti

**Tabella 8.3** La perdita di funzioni metaboliche nei batteri patogeni.

SPECIE (NUMERO DI GENI)	GLICOLISI	CICLO ACIDI TRICARBOSSILICI	BIOSINTESI AMINOACIDI	BIOSINTESI PURINE	BIOSINTESI PIRIMIDINE
<i>Mycoplasma genitalium</i> (470)	+	-	-	-	-
<i>Buchnera</i> spp. (588)	+	-	+	+	+
<i>Rickettsia prowazekii</i> (834)	-	+	-	-	-
<i>Chlamidia trachomatis</i> (894)	+	-	+	-	-
<i>Treponema pallidum</i> (1014)	+	-	-	-	-
<i>Mycobacterium leprae</i> (1604)	parziale	in decadimento	+	+	+

Batteri con genomi piccoli dipendono dall'ospite per numerose funzioni: sono state infatti perse funzioni metaboliche importanti come la glicolisi, biosintesi di aminoacidi e purine e pirimidine che vengono fornite dall'ospite.

Molti di questi genomi possiedono ancora geni per queste funzioni ma non funzionali ovvero pseudogeni.

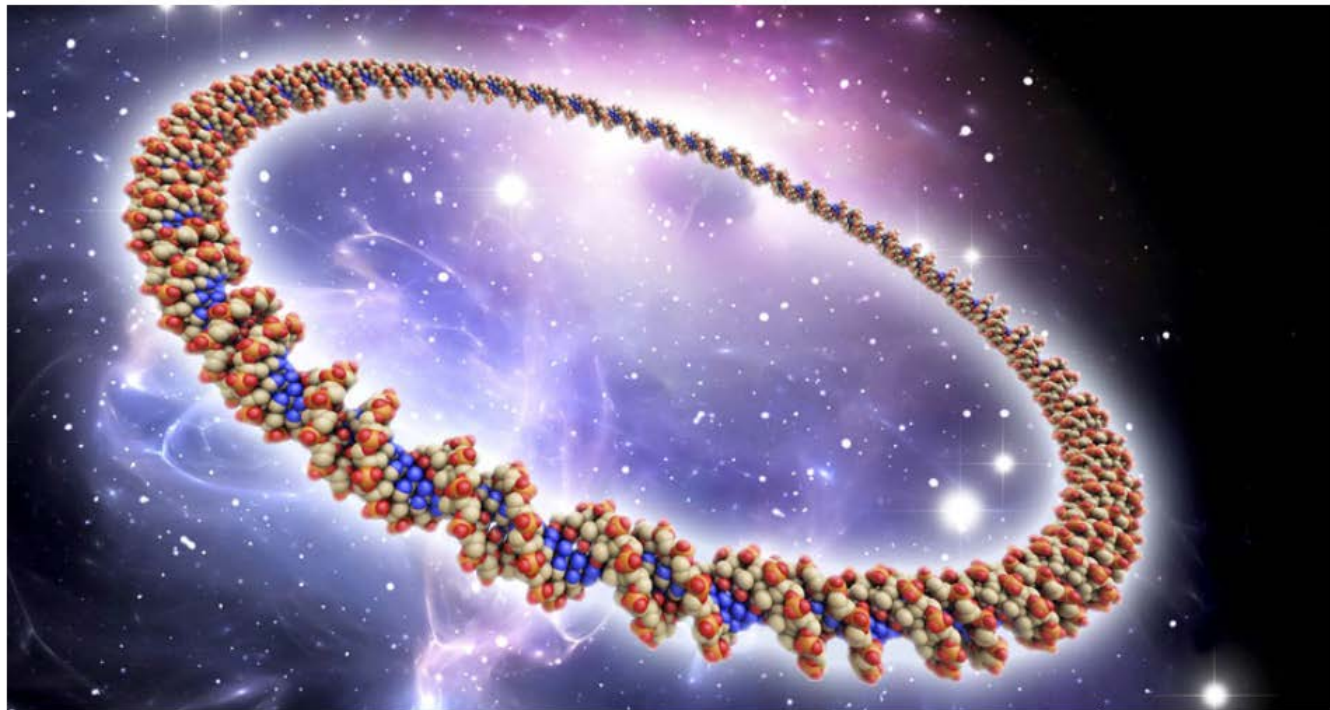
In *Mycobacterium leprae* circa 50 % dei geni sono pseudogeni che mantengono omologia con *Mycobacterium tuberculosis*

GENETICS MICROBES

# Genes: How few needed for life?

People need more than 20,000, but one bacterium was rebuilt — and lived — with a mere 473

BY TINA HESMAN SAEY APR 5, 2016 — 7:15 AM EST



Una domanda che scaturisce dal confronto dei genomi riguarda la dimensione minima sufficiente perchè vi sia VITA.

E' possibile che esistano specie con un genoma più piccolo di 0.58Mb ovvero quello di *M.genitalium* che per ora sembra la forma di batterio con minore geni.

E' stata fatta quindi in silico un 'analisi genomica comparativa tra i genomi più piccoli per capire quali fossero i geni conservati

Mutagenesi di un gene alla volta per verificare se questo fosse essenziale alla vita della cellula

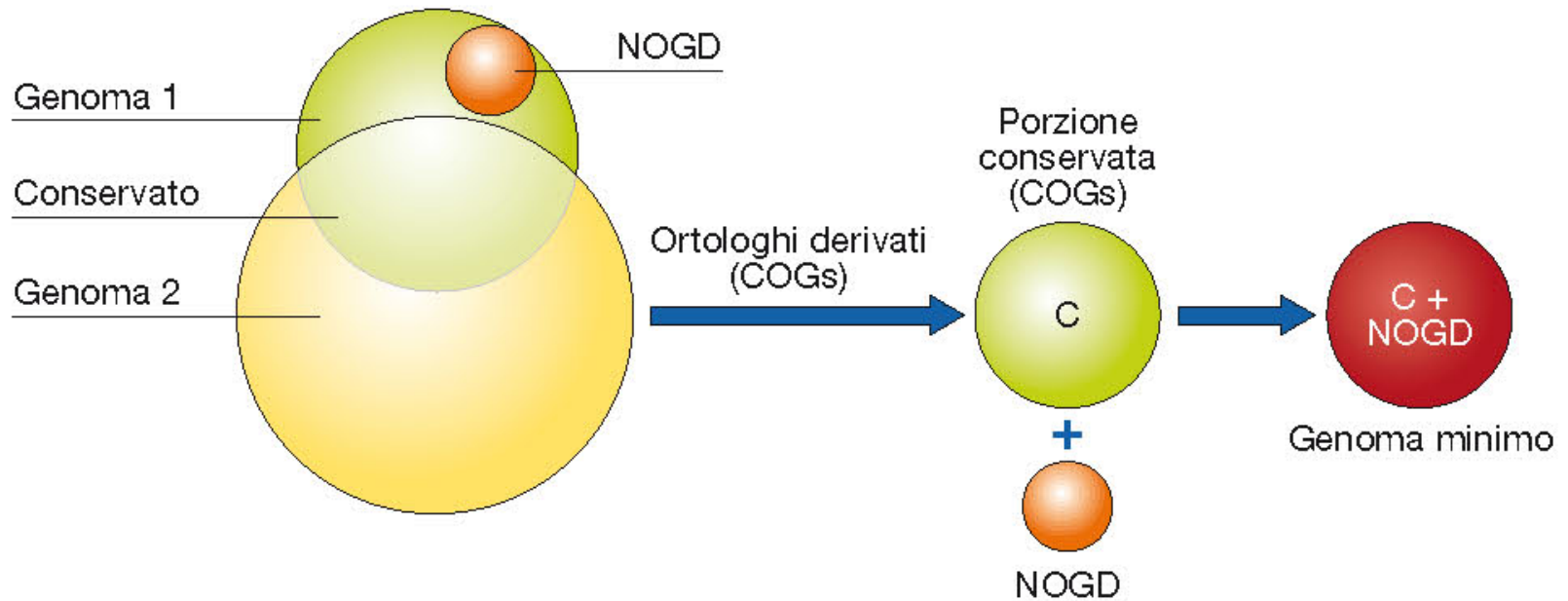
La prima proposta di genoma minimo viene dall'analisi dei 2 genomi sequenziati per primi *Haemophilus influenzae* ( 1815 geni) e *Mycoplasma genitalium* 525 geni) che rivela un gruppo comune di **geni 256** molto più piccolo del numero di geni presenti nel genoma di *Mycoplasma*.

Da qui l'idea di introdurre il metodo di global transposon mutagenesis che ha permesso, effettuando uno studio sulla vitalità delle cellule portatrici di mutazioni in ciascuno dei 525 geni di *Mycoplasma* di stabilire che

150 geni sono NON ESSENZIALI

375 geni sono essenziali

Nasce la possibilità di produrre un genoma minimo che fosse più piccolo di quelli esistenti in natura ma che sarebbe stato in ogni caso più grande di questi 256 geni identificati come CORE

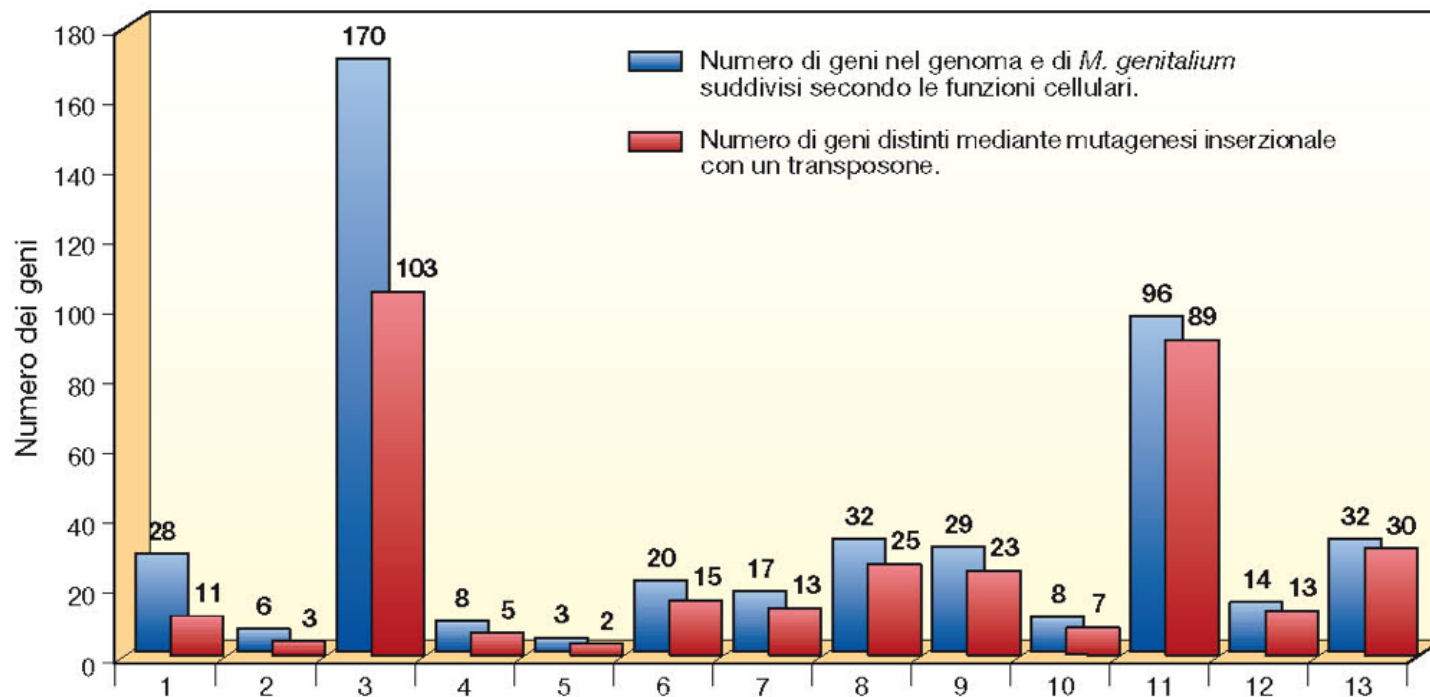


L'analisi comparativa delle sequenze dei genomi di due organismi filogeneticamente distanti mostra l'esistenza di geni conservati nelle due specie (geni ortologi o COGs Cluster of Orthologous Genes) e di geni non ortologi specifici dell'uno o dell'altro organismo (NOGD Non Orthologous Gene Displacement). Il genoma minimo è il numero dei geni essenziali per la vita di uno specifico organismo ed è la somma dei geni COGs più i NOGS che saranno più o meno ampi.



Nel caso di *M. genitalium* , l'analisi di 1354 inserzioni casuali di trasposoni nel genoma mostra che dei 480 geni codificanti , 265-350 sono essenziali nelle condizioni di laboratorio, tra questi circa 100 geni codificano per funzioni non note.

Il primo ipotetico genoma minimo è stato costruito sulla base di dati derivanti dalla mutagenesi e delezioni e da informazioni di letteratura sui pathway più importanti.



#### Funzioni cellulari delle classi dei geni

- |  |   |
|--|---|
| 1. Membrana                              | 8. Trasporto                                    |
| 2. Regolazione                           | 9. Replicazione/ricombinazione/riparazione      |
| 3. Sconosciuto                           | 10. Metabolismo degli acidi grassi e dei lipidi |
| 4. Metabolismo principale                | 11. Traduzione                                  |
| 5. Biosintesi di cofattori               | 12. Processi cellulari                          |
| 6. Metabolismo delle purine e pirimidine | 13. Produzione di energia                       |
| 7. Trascrizione                          |   |

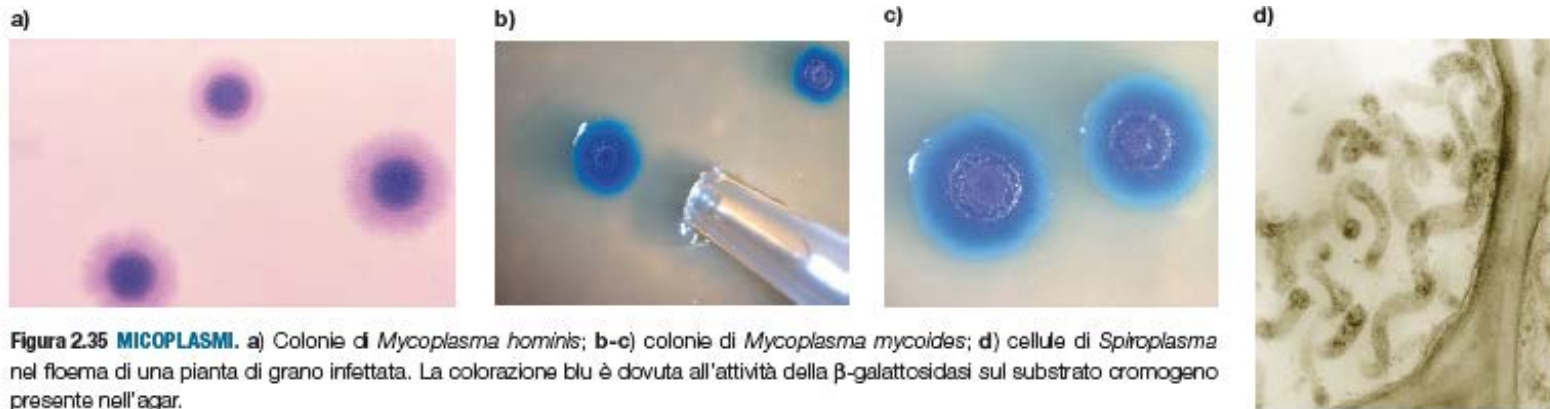


Figura 2.35 MICOPLASMI. a) Colonie di *Mycoplasma hominis*; b-c) colonie di *Mycoplasma mycoides*; d) cellule di *Spiroplasma* nel floema di una pianta di grano infettata. La colorazione blu è dovuta all'attività della  $\beta$ -galattosidasi sul substrato cromogeno presente nell'agar.

I micoplasmi batteri senza parete caratterizzati da un genoma molto piccolo (meno di 1 megabase), molte richieste nutrizionali.

Dal RNA16S appartengono ai Gram+, sono tutti commensali parassiti o patogeni, infezioni difficili da trattare con antibiotici

## Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome

Daniel G. Gibson, Gwynedd A. Benders, Cynthia Andrews-Pfannkoch, Evgeniya A. Denisova, Holly Baden-Tillson, Jaysree Zaveri, Timothy B. Stockwell, Anushka Brownley, David W. Thomas, Mikkel A. Algire, Chuck Merryman, Lei Young, Vladimir N. Noskov, John I. Glass, J. Craig Venter, Clyde A. Hutchison III, Hamilton O. Smith\*

We have synthesized a 582,970–base pair *Mycoplasma genitalium* genome. This synthetic genome, named *M. genitalium* JCVI-1.0, contains all the genes of wild-type *M. genitalium* G37 except MG408, which was disrupted by an antibiotic marker to block pathogenicity and to allow for selection. To identify the genome as synthetic, we inserted “watermarks” at intergenic sites known to tolerate transposon insertions. Overlapping “cassettes” of 5 to 7 kilobases (kb), assembled from chemically synthesized oligonucleotides, were joined by *in vitro* recombination to produce intermediate assemblies of approximately 24 kb, 72 kb (“1/8 genome”), and 144 kb (“1/4 genome”), which were all cloned as bacterial artificial chromosomes in *Escherichia coli*. Most of these intermediate clones were sequenced, and clones of all four 1/4 genomes with the correct sequence were identified. The complete synthetic genome was assembled by transformation-

genome, we needed to reliable methods for the much larger synthetic I.

Strategy for synthesizing 580,076-bp *M. gen* (*Mycoplasma genitalium* genomic sequence; acc was partitioned into 10 roughly 5 to 7 kb in 10 individually synthesized and then joined together cassette boundaries we so that each cassette c complete genes. This deletion or manipulation ual cassettes. Most ca adjacent neighbors by segments overlapped t Cassette 101 overlappi ngs. This deletion or manipulation ual cassettes. Most ca adjacent neighbors by segments overlapped t Cassette 101 overlappi ngs. This deletion or manipulation ual cassettes. Most ca adjacent neighbors by segments overlapped t Cassette 101 overlappi ngs.

Short “watermark” in cassettes 14, 29, 39, 5 inserted or substituted sa or encode information in can be either in noncod

*Science* (2008) :319, 1215

## Global Transposon Mutagenesis and a Minimal *Mycoplasma* Genome

Clyde A. Hutchison III,<sup>1,2\*</sup> Scott N. Peterson,<sup>1\*†</sup> Steven R. Gill,<sup>1</sup> Robin T. Cline,<sup>1</sup> Owen White,<sup>1</sup> Claire M. Fraser,<sup>1</sup> Hamilton O. Smith,<sup>1‡</sup> J. Craig Venter<sup>1‡§</sup>

*Mycoplasma genitalium* with 517 genes has the smallest gene complement of any independently replicating cell so far identified. Global transposon mutagenesis was used to identify nonessential genes in an effort to learn whether the naturally occurring gene complement is a true minimal genome under laboratory growth conditions. The positions of 2209 transposon insertions in the completely sequenced genomes of *M. genitalium* and its close relative *M. pneumoniae* were determined by sequencing across the junction of the transposon and the genomic DNA. These junctions defined 1354 distinct sites of insertion that were not lethal. The analysis suggests that 265 to 350 of the 480 protein-coding genes of *M. genitalium* are essential under laboratory growth conditions, including about 100 genes of unknown function.

*Science* (1999) :286, 2165

## RESEARCH ARTICLE SUMMARY

### SYNTHETIC BIOLOGY

## Design and synthesis of a minimal bacterial genome

Clyde A. Hutchison III,<sup>†</sup> Ray-Yuan Chuang,<sup>‡</sup> Vladimir N. Noskov, Nacaya Assad-Garcia, Thomas J. Deerinck, Mark H. Ellisman, John Gill, Krishna Kannan, Bogamil J. Karas, Li Ma, James F. Peletier, Zhi-Qing Qi, R. Alexander Richter, Elizabeth A. Strychalski, Lijie Sun, Yo Suzuki, Bilyana Tsvetanova, Kim S. Wise, Hamilton O. Smith, John I. Glass, Chuck Merryman, Daniel G. Gibson, J. Craig Venter\*

**INTRODUCTION:** In 1984, the simplest cells capable of autonomous growth, the mycoplasmas, were proposed as models for understanding the basic principles of life. In 1995, we reported the first complete cellular genome sequences (*Haemophilus influenzae*, 1815 genes, and *Mycoplasma genitalium*, 525 genes). Comparison of these sequences revealed a conserved core of about 250 essential genes, much smaller than either genome. In 1999, we introduced the method of global transposon mutagenesis and experimentally demonstrated that *M. genitalium* contains many genes that are nonessential for growth in the laboratory, even though it has the

smallest genome known for an autonomously replicating cell found in nature. This implied that it should be possible to produce a minimal cell that is simpler than any natural one. Whole genomes can now be built from chemically synthesized oligonucleotides and brought to life by installation into a receptive cellular environment. We have applied whole-genome design and synthesis to the problem of minimizing a cellular genome.

**RATIONALE:** Since the first genome sequences, there has been much work in many bacterial models to identify nonessential genes and

define core sets of conserved genetic functions, using the methods of comparative genomics. Often, more than one gene product can perform a particular essential function. In such cases, neither gene will be essential, and neither will necessarily be conserved. Consequently, these approaches cannot, by themselves, identify a set of genes that is sufficient to constitute a viable genome. We set out to define a minimal cellular genome experimentally by designing and building one, then testing it for viability. Our goal is a cell so simple that we can determine the molecular and biological function of every gene.

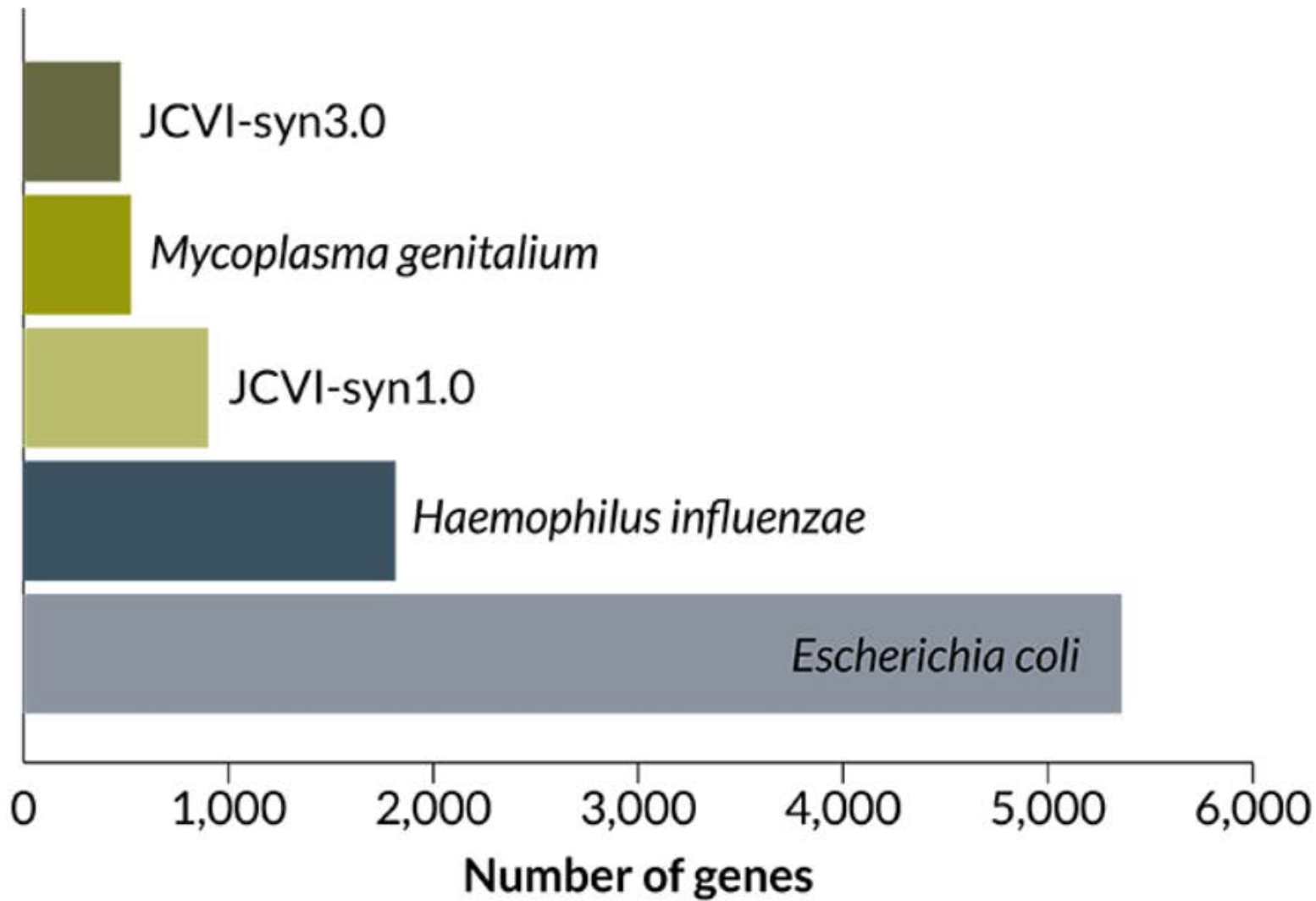
**RESULTS:** Whole-genome design and synthesis were used to minimize the 1079-kilobase pair (kbp) synthetic genome of *M. mycoides* JCVI-syn1.0. An initial design, based on collective knowledge of molecular biology in combination

#### ON OUR WEB SITE

Read the full article at <http://dx.doi.org/10.1126/science.aad6253>

with limited transposon mutagenesis data, failed to produce a viable cell. Improved transposon mutagenesis methods revealed a class of quasi-essential genes that are needed for robust growth, explaining the failure of our initial design. Three more cycles of design, synthesis, and testing, with retention of quasi-essential genes, produced JCVI-syn3.0 (531 kbp, 473 genes). Its genome is smaller than that of any autonomously replicating cell found in nature. JCVI-syn3.0 has a doubling time of

Corrected 28 June 2016; see full text.



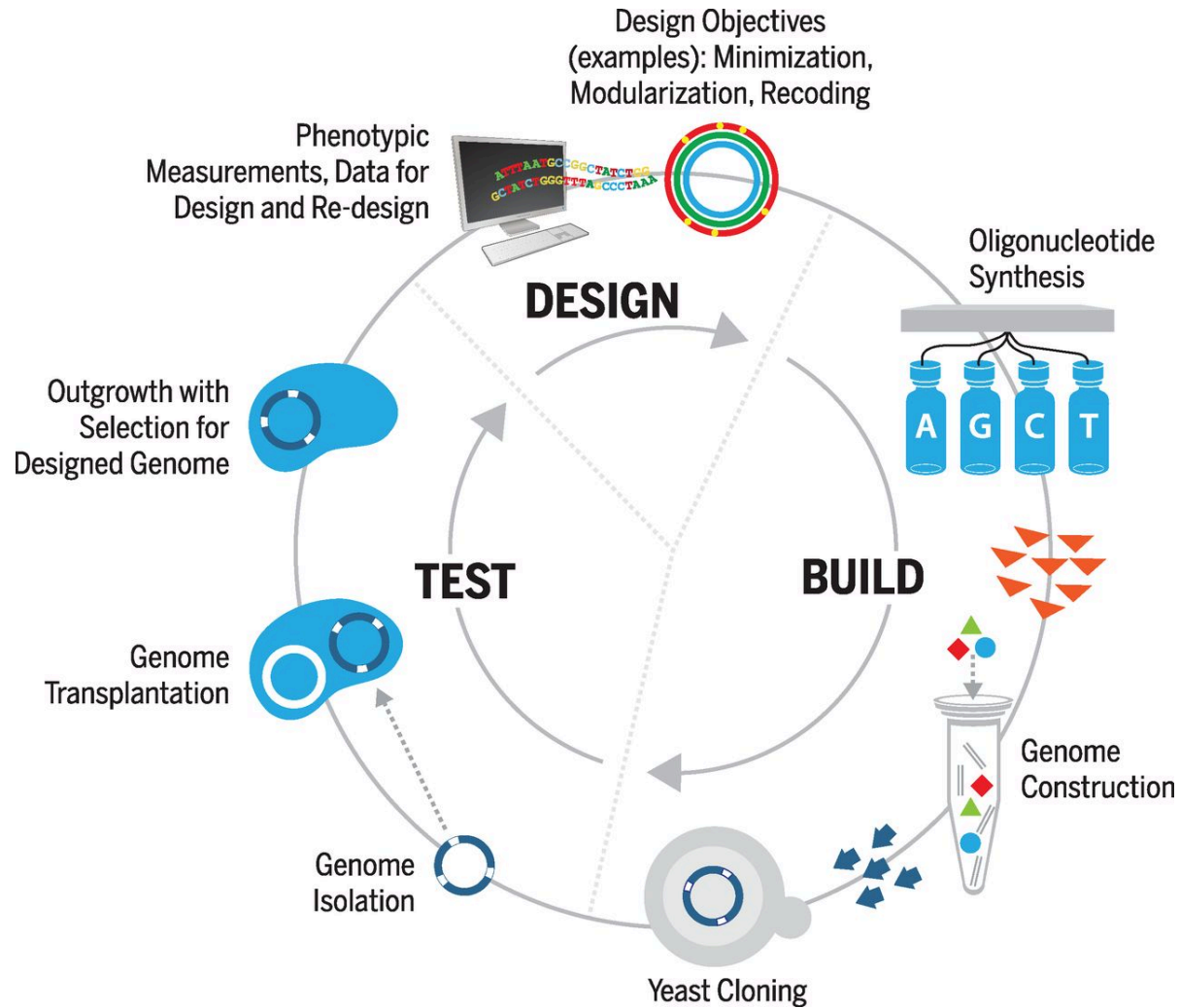
Il primo genoma sintetizzato artificialmente riguarda quello di *M. mycoides*, un *Mycoplasma* in grado di crescere più rapidamente.

Questo genoma definito syn 1.0 è stato trapiantato all'interno di una specie diversa *M. capricolum* e corrisponde esattamente al genoma di *M. mycoides* con aggiunta di alcuni marcatori e di sequenze del vettore.

A partire da questo genoma è stato creato syn 3.0 costituito da 531 kb che codifica 438 proteine e 35 sRNA.

Il suo genoma è più piccolo di *M. genitalium* ma si duplica 5 volte più velocemente.

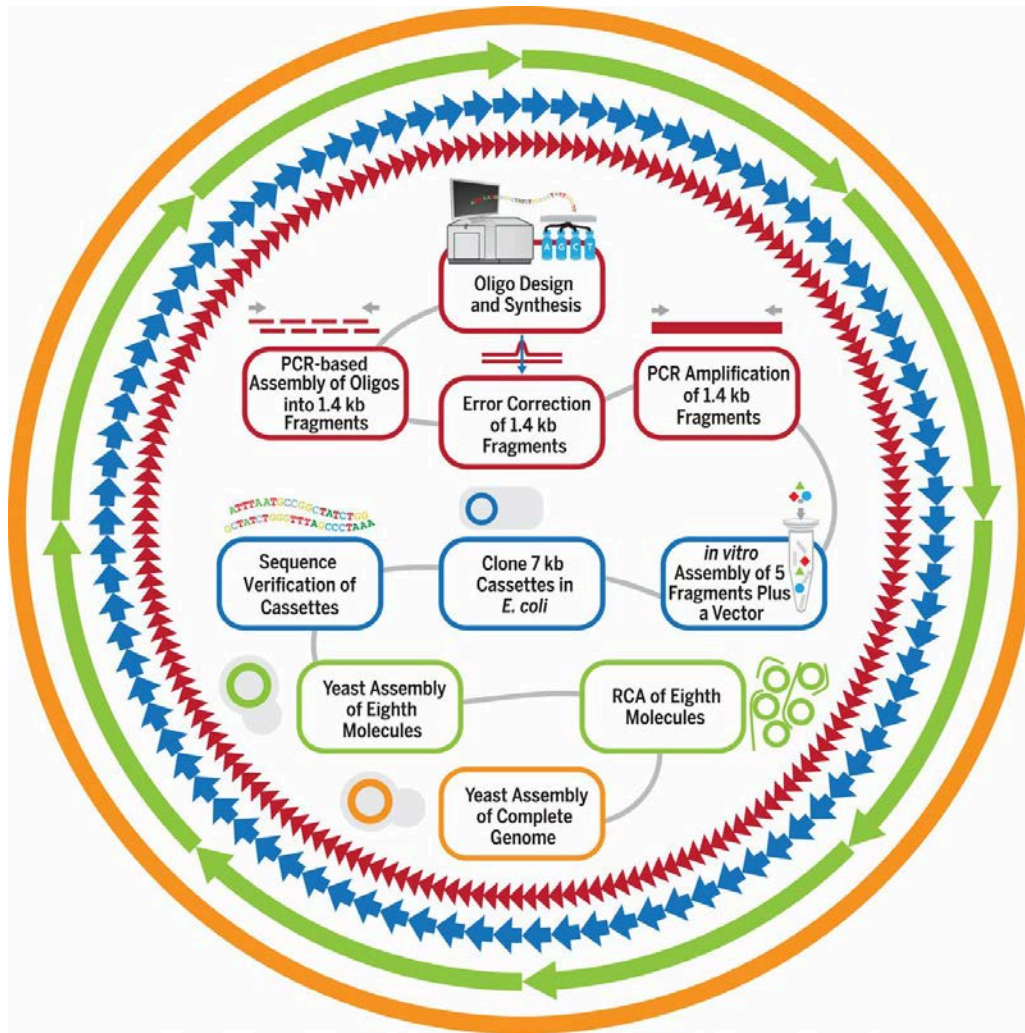
**Fig. 1 The JCVI DBT cycle for bacterial genomes.**



Clyde A. Hutchison III et al. *Science* 2016;351:aad6253

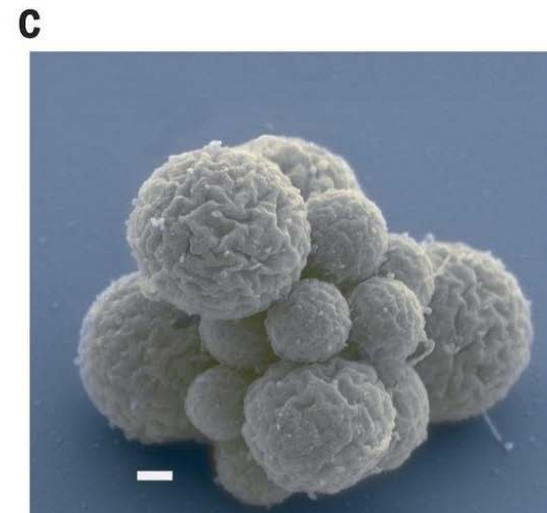
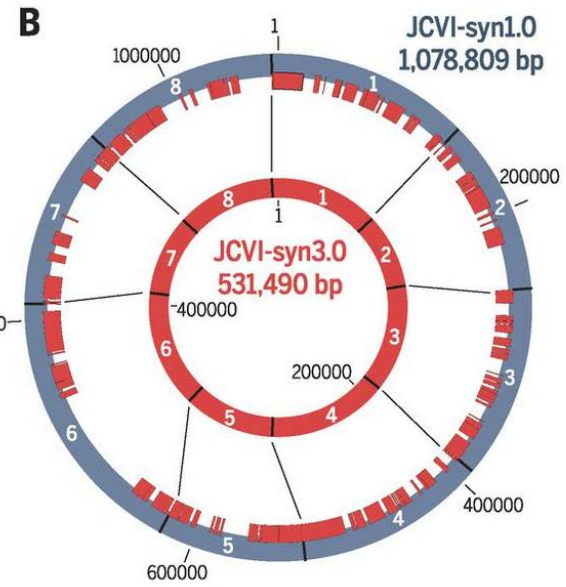
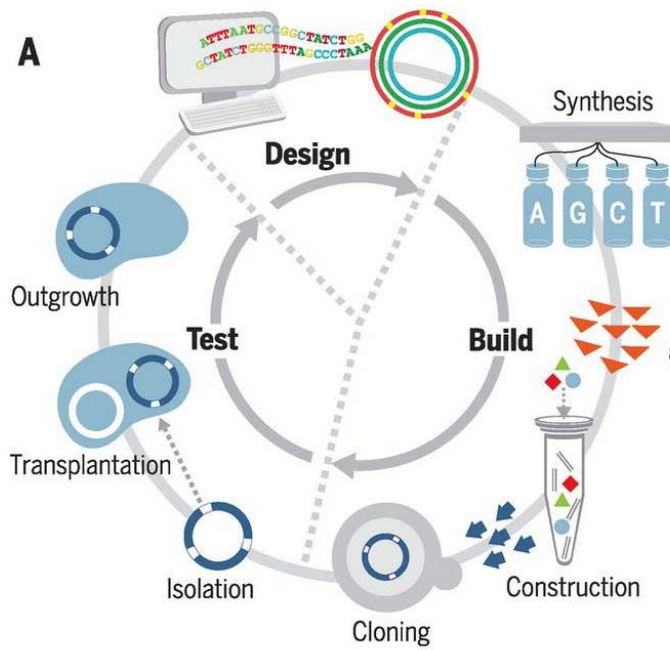


## Strategy for whole-genome synthesis.



Overlapping oligonucleotides (oligos) were designed, chemically synthesized, and assembled into 1.4-kbp fragments (red). After error correction and PCR amplification, five fragments were assembled into 7-kbp cassettes (blue). Cassettes were sequence-verified and then assembled *in vitro* in yeast to generate one-eighth molecules (green). The eight molecules were amplified by RCA and then assembled in yeast to generate the complete genome (orange).



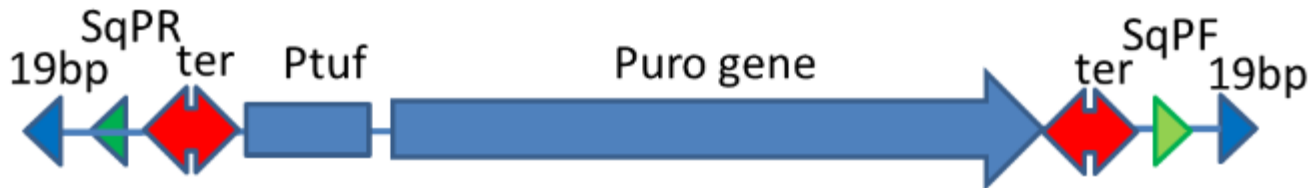


## Procedura

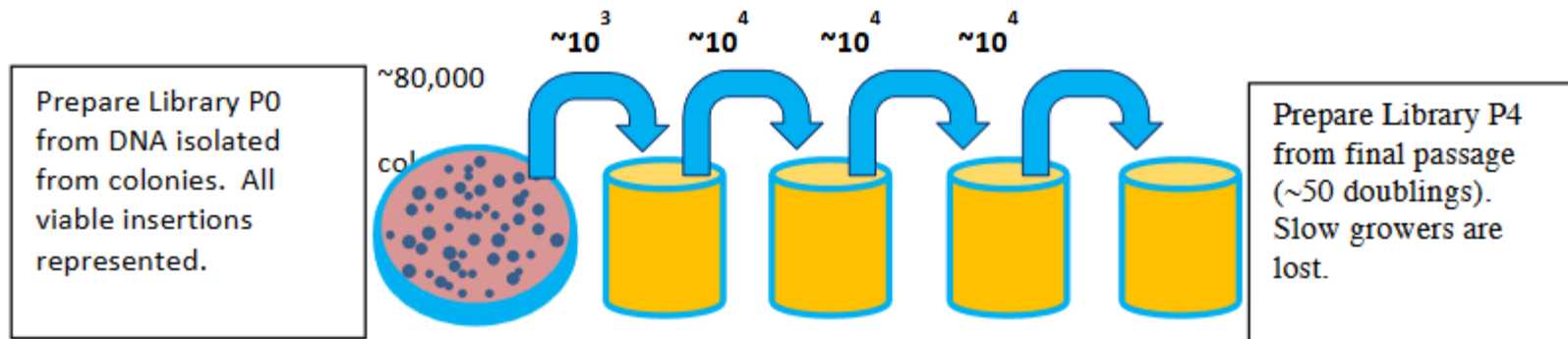
Partendo dal genoma syn 1 , effettuata mutagenesi con miniTn5 purR. Dal Pull di colonie contenenti Tn inserito in siti diversi (PO) si sono fatte una serie di passaggi seriali ( circa 40 passaggi )per controselezionare i ceppi che crescevano lentamente (P4) .

- I geni che non venivano mai colpiti da mutagenesi ESSENZIALI
- Quelli colpiti da mutagenesi e presenti sia nel PO che nel P4 definiti NON ESSENZIALI
- I geni colpiti nel PO ma non nel P4 definiti QUASI ESSENZIALI : se la delezione di questi geni generava crescita ridotta definiti i (I genes impairment growth)

**Step 1.** Construct Tn5 transposon containing 19 bp mosaic ends, sequencing primer sites, terminator sequences, and a selectable marker (puromycin-resistance gene). Bind Tn5 transposase (Epicentre) to the 19bp termini to form the active transposome.



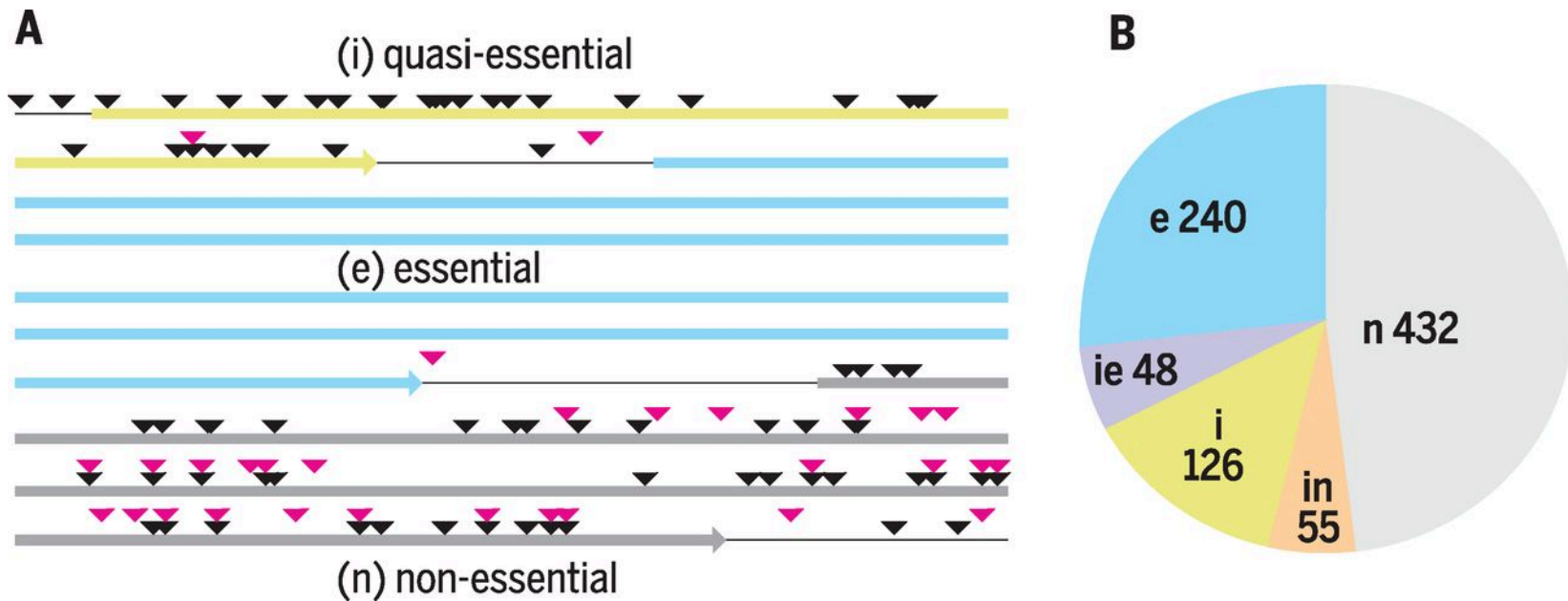
**Step 2.** Introduce transposome into *Mycoplasma mycoides* JCVI-syn1.0 R-M (minus) strain by polyethylene glycol (PEG) transformation method. Collect puromycin-resistant colonies and serially propagate to eliminate slow growers.



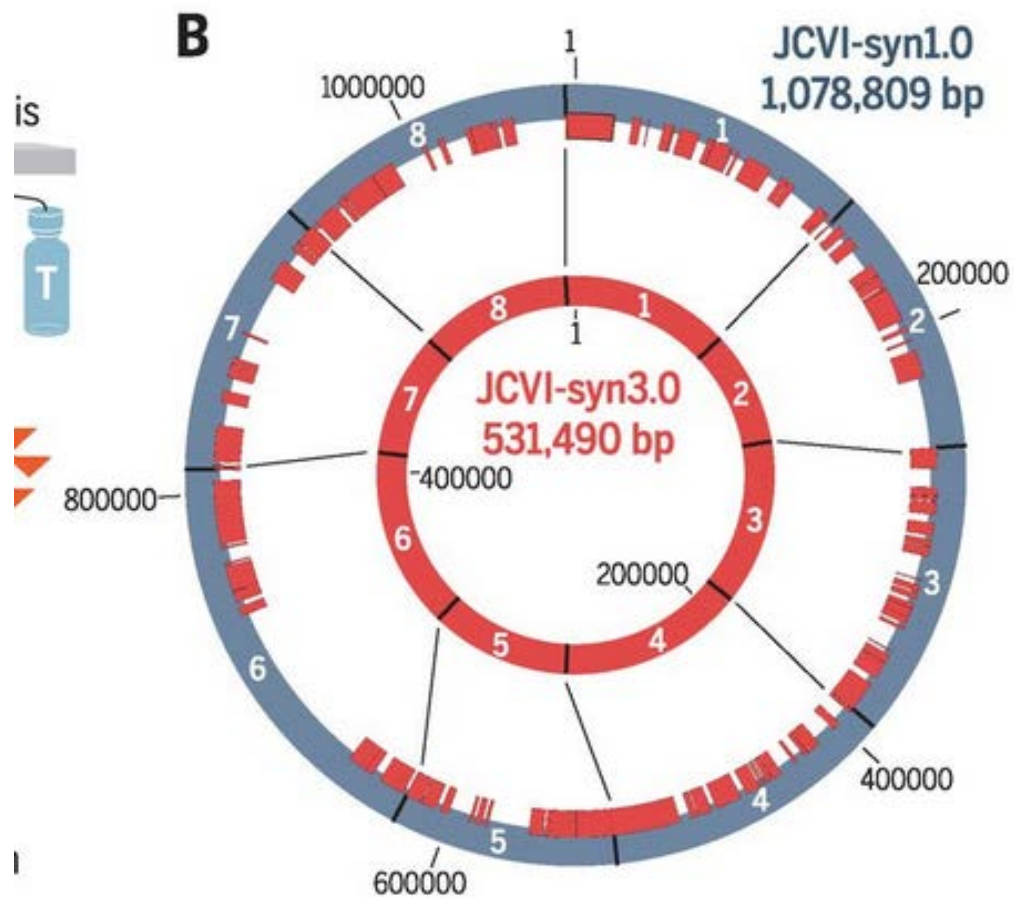
## Procedura

Partendo dal genoma syn 1 , effettuata mutagenesi con miniTn5 purR.  
 Dal Pull di colonie contenenti Tn inserito in siti diversi (P0) si sono fatte una serie di passaggi seriali ( circa 40 passaggi )per controselezionare i ceppi che crescevano lentamente (P4) .

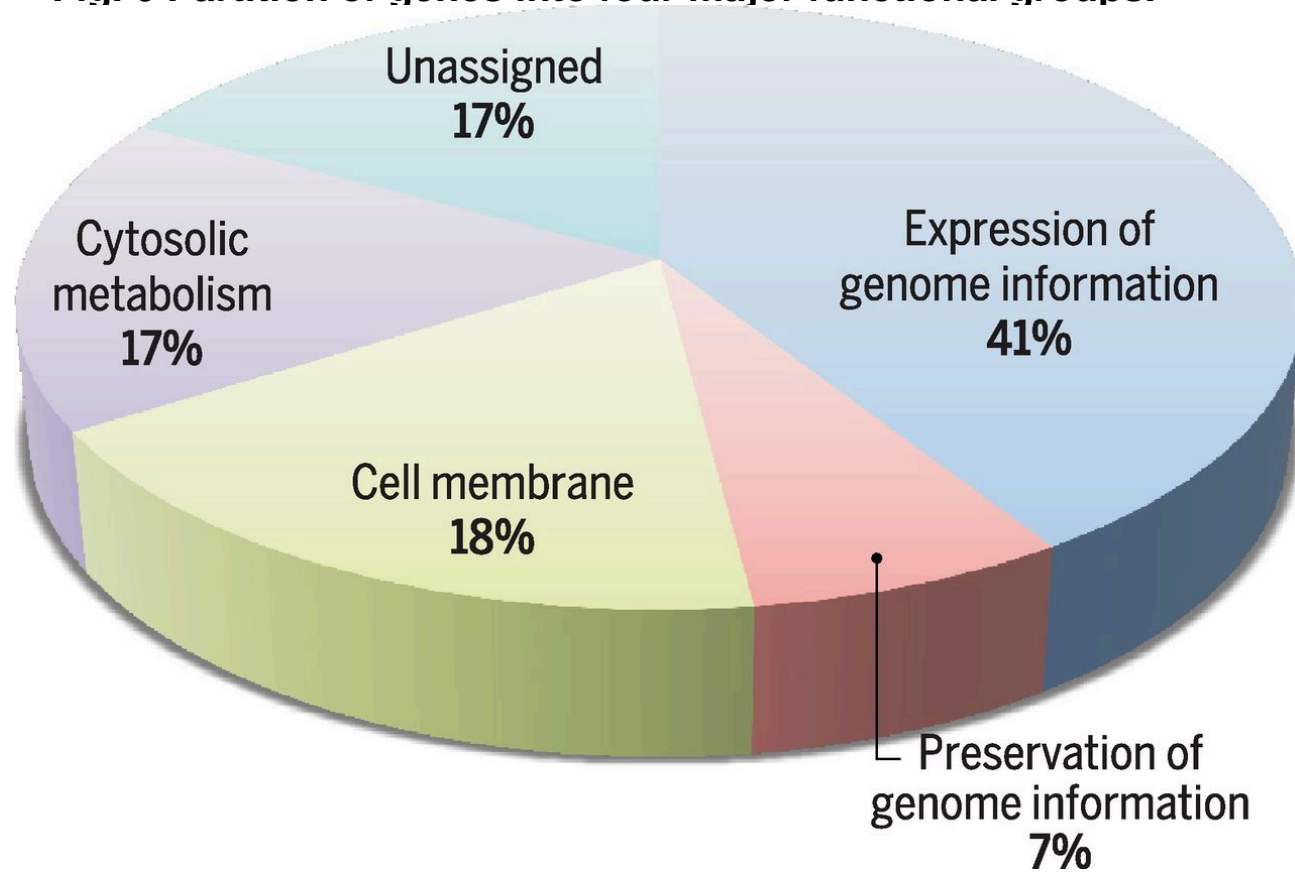
**Fig. 3 Classification of gene essentiality by transposon mutagenesis.**



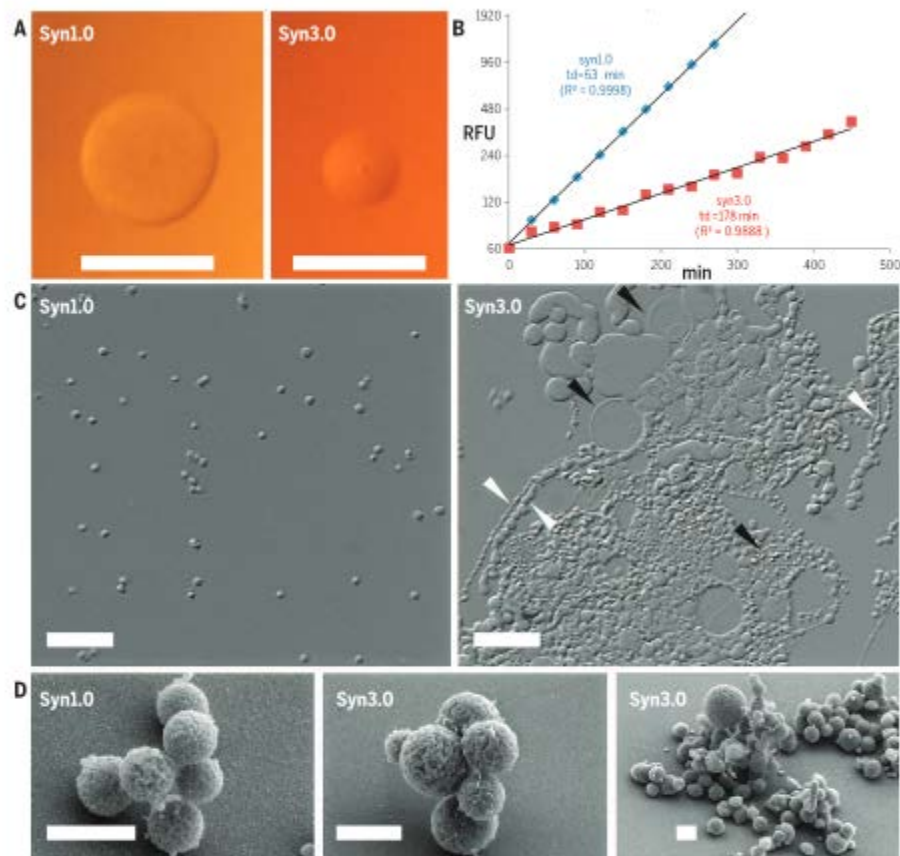
Classification of gene essentiality by transposon mutagenesis. **(A)** Examples of the three gene classifications, based on Tn5 mutagenesis data. The region of syn1.0 from sequence coordinates 166,735 to 170,077 is shown. The gene *MMSYN1\_0128* (lime arrow) has many P0 Tn5 inserts (black triangles) and is an i-gene (quasi-essential). The next gene, *MMSYN1\_0129* (light blue arrow), has no inserts and is an e-gene (essential). The last gene, *MMSYN1\_0130* (gray arrow), has both P0 (black triangles) and P4 (magenta triangles) inserts and is an n-gene (nonessential). Intergenic regions are indicated by black lines. **(B)** The number of syn1.0 genes in each Tn5-mutagenesis classification group. The n- and in-genes are candidates for deletion in reduced genome designs.



**Fig. 6 Partition of genes into four major functional groups.**



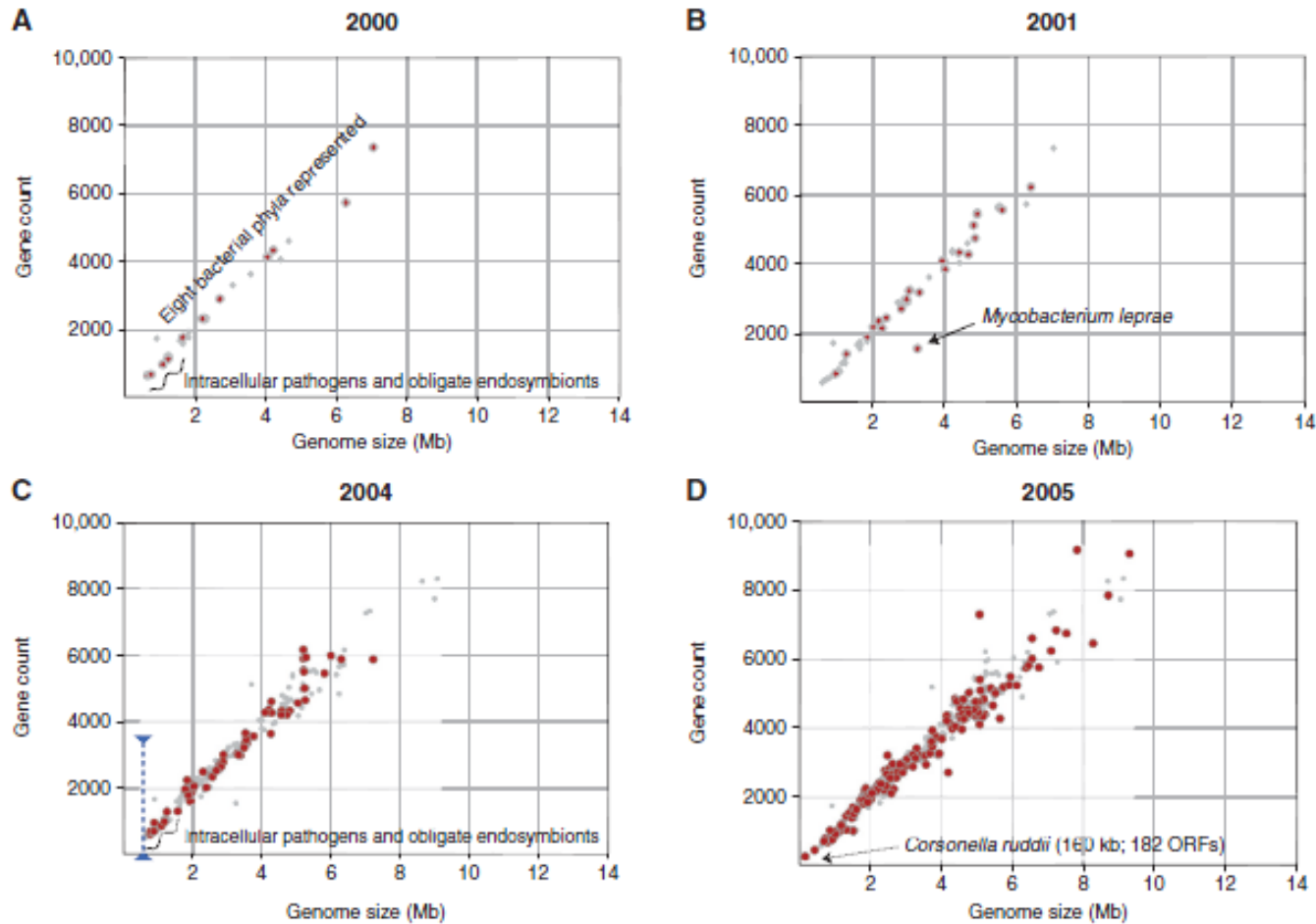
Partition of genes into four major functional groups. Syn3.0 has 473 genes. Of these, 79 have no assigned functional category (Table 1). The remainder can be assigned to four major functional groups: (i) expression of genome information (195 genes); (ii) preservation of genome information (34 genes); (iii) cell membrane structure and function (84 genes); and (iv) cytosolic metabolism (81 genes). The percentage of genes in each group is indicated.



**Fig. 7. Comparison of *syn1.0* and *syn3.0* growth features.** (A) Cells derived from 0.2  $\mu\text{m}$ -filtered liquid cultures were diluted and plated on agar medium to compare colony size and morphology after 96 hours (scale bars, 1.0 mm). (B) Growth rates in liquid static culture were determined using a fluorescent measure (relative fluorescent units, RFU) of double-stranded DNA accumulation over time (minutes) to calculate doubling times (td). Coefficients

of determination ( $R^2$ ) are shown. (C) Native cell morphology in liquid culture was imaged in wet mount preparations by means of differential interference contrast microscopy (scale bars, 10  $\mu\text{m}$ ). Arrowheads indicate assorted forms of segmented filaments (white) or large vesicles (black). (D) Scanning electron microscopy of *syn1.0* and *syn3.0* (scale bars, 1  $\mu\text{m}$ ). The picture on the right shows a variety of the structures observed in *syn3.0* cultures.

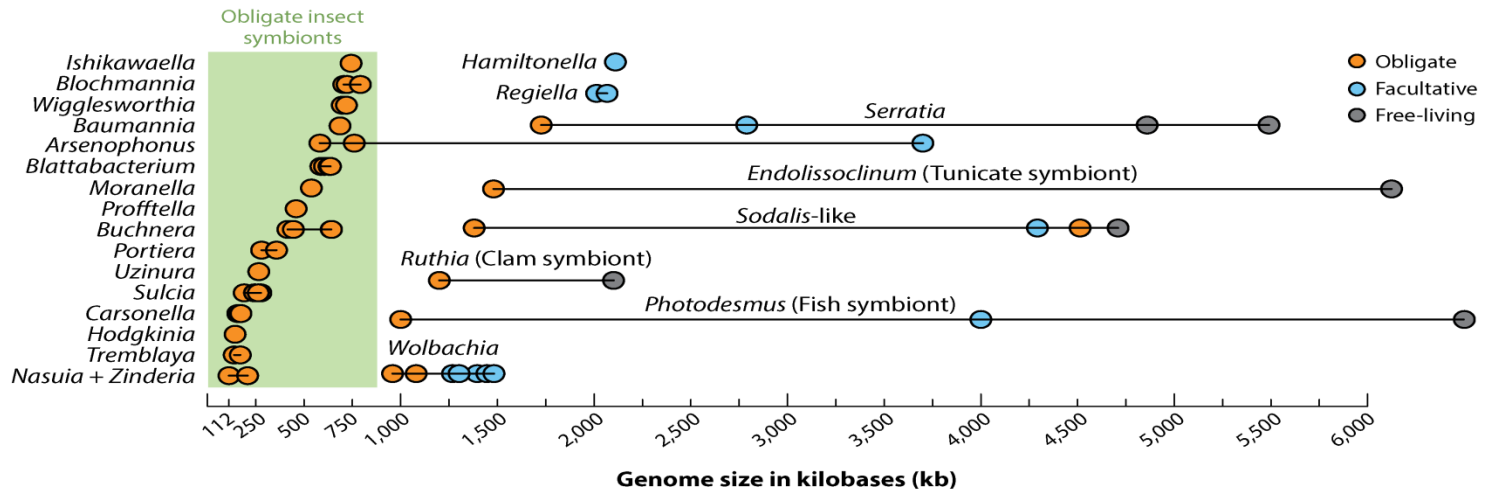
# La corsa al sequenziamento dei genomi batterici



**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.

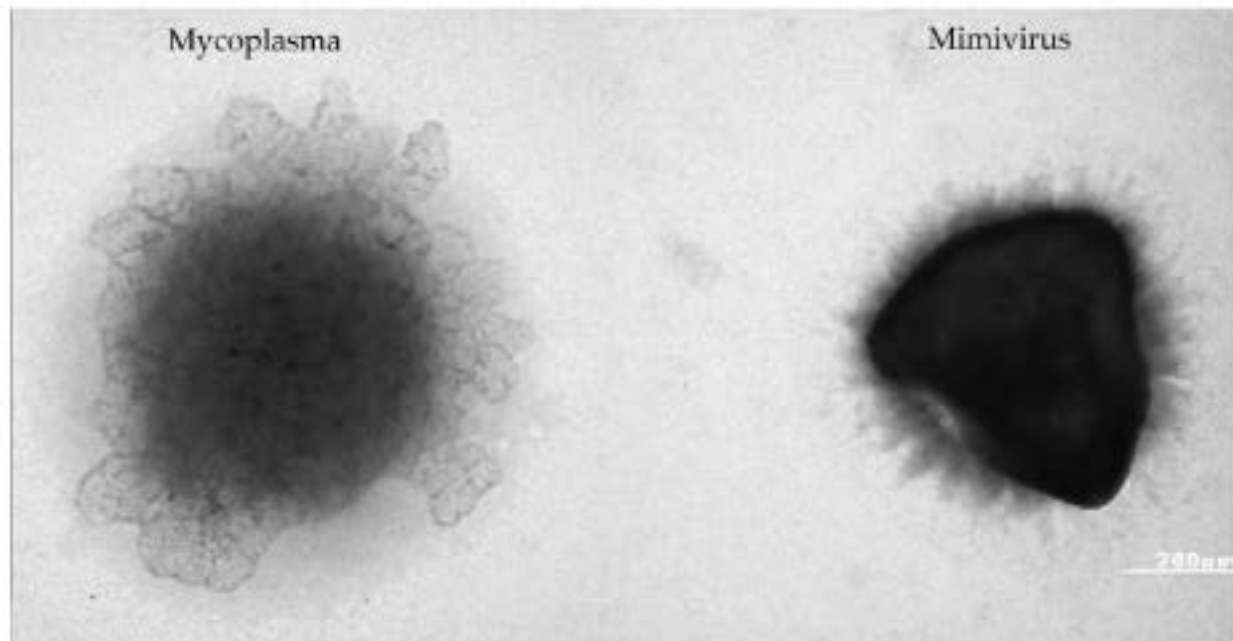


# Genomi al di sotto della cellula minima! 112 Kb



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

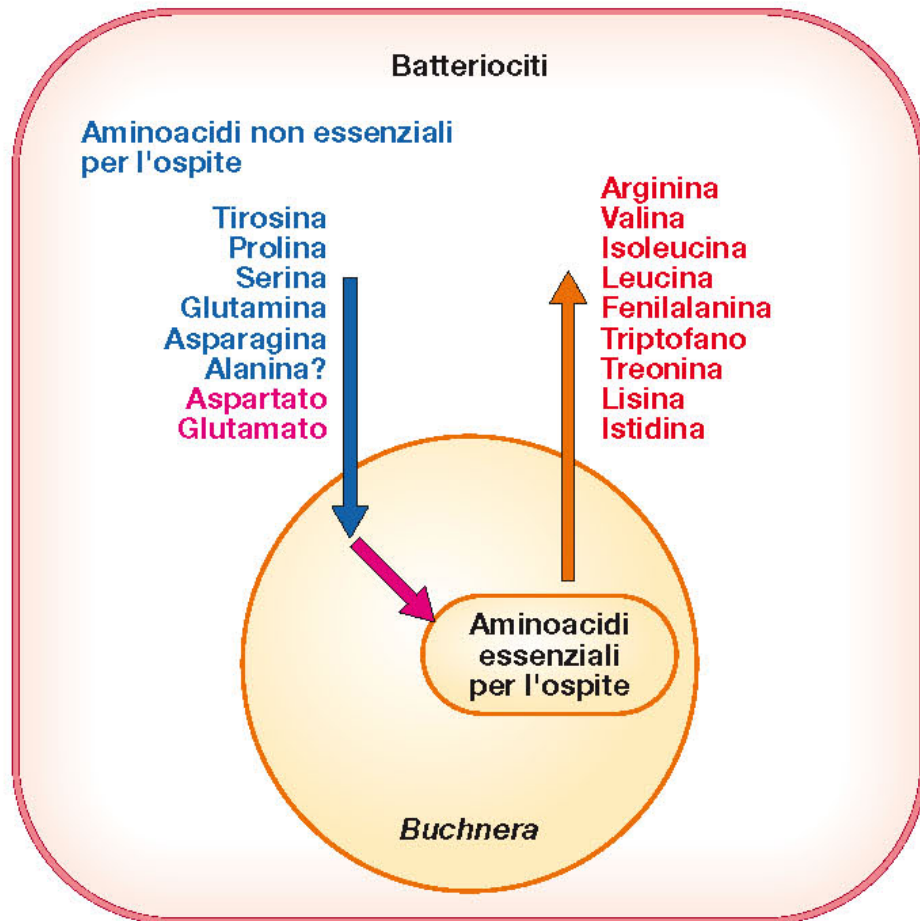
Despite their size, each retains some genes that enable provisioning of limiting nutrients or other capabilities required by hosts. Genome sequence analyses show that genome reduction is an ongoing process, resulting in a continuum of sizes, with the smallest genome currently known at **112 kilobases**. 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.

# 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 !!

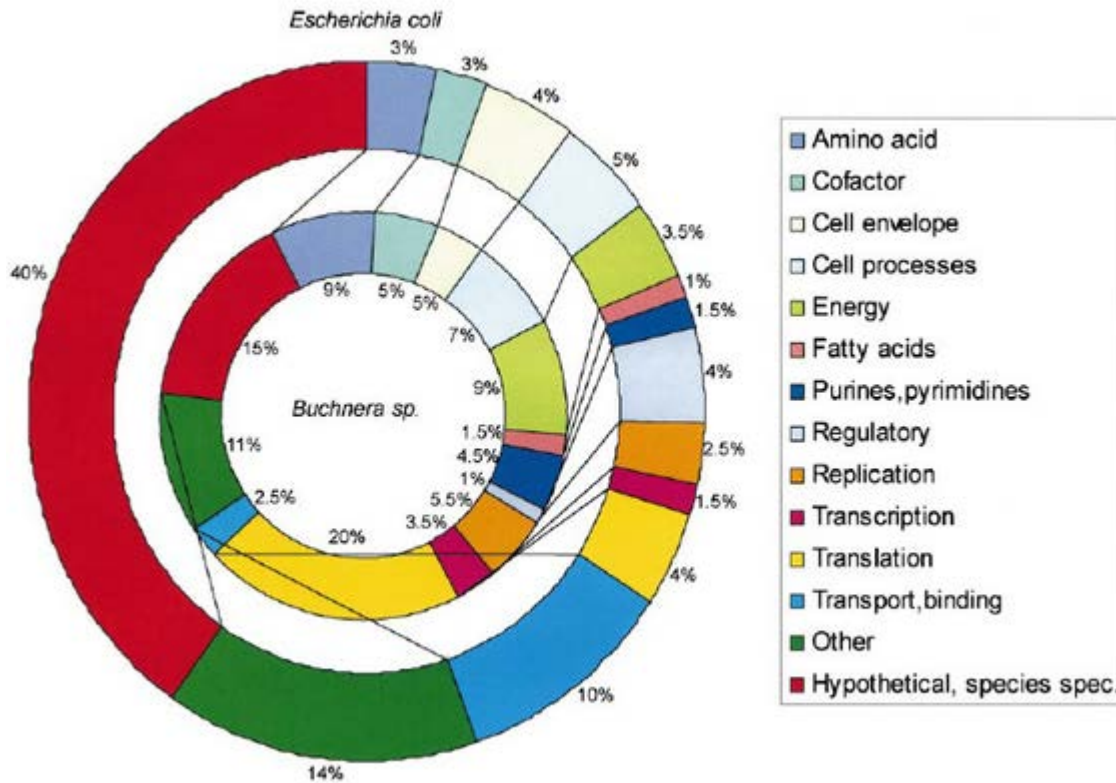


Batterio → Afide  
Aminoacidi essenziali che l'insetto non trova nella pianta di cui si nutre ,

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 essenziali 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.

# 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 400 kb .

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

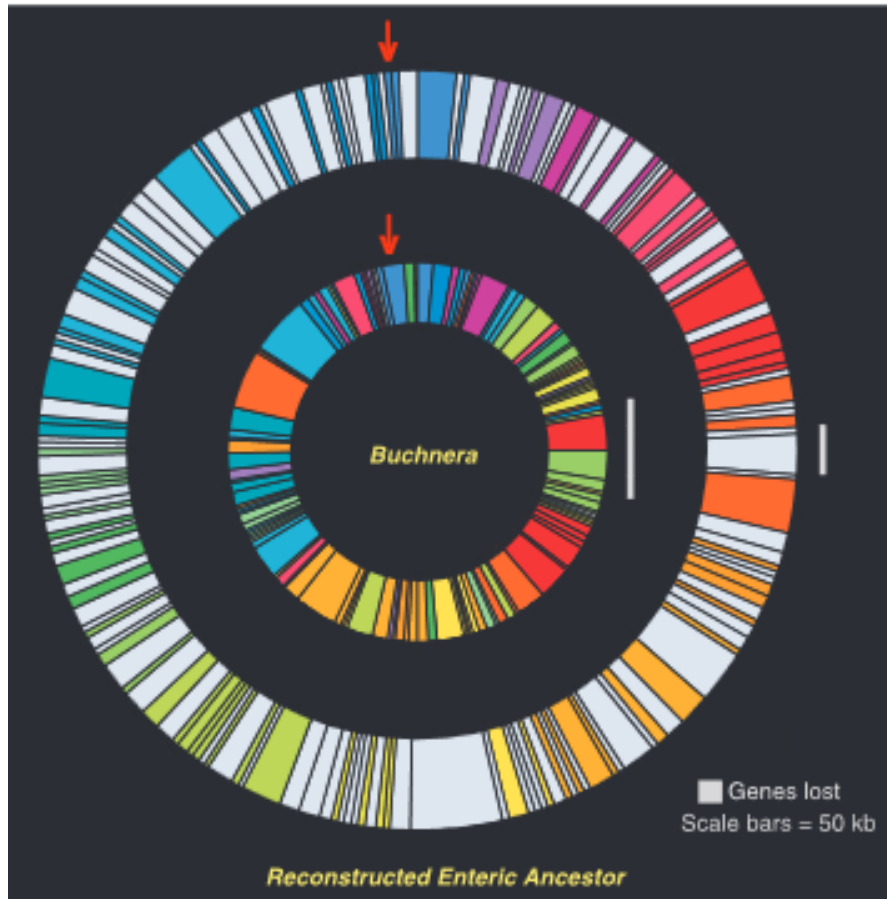
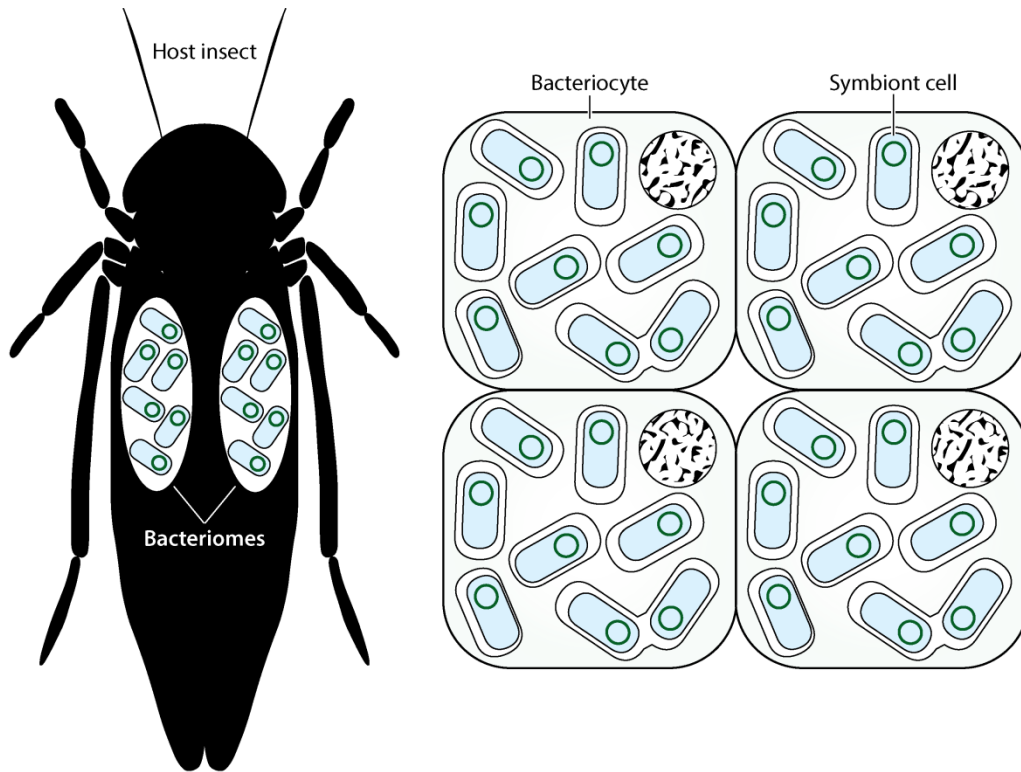


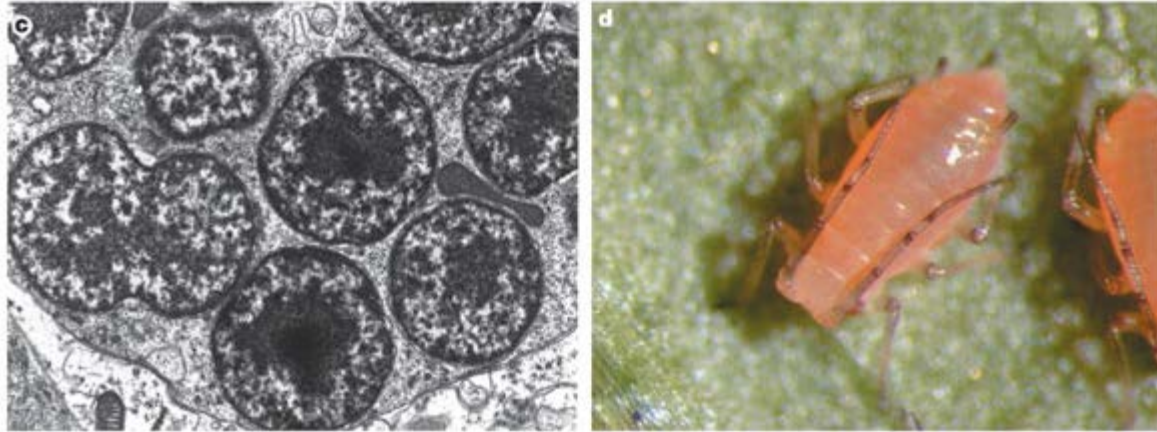
Fig. 1. Genome reduction in *Buchnera*-APS, the bacterial endosymbiont of aphids. *Buchnera* has undergone massive genome reduction (to 0.64 Mb) (34). Virtually all of its 590 genes have close homologs in the genomes of the enteric bacteria, which, along with phylogenetic evidence, indicates that its 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 (57)] 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.



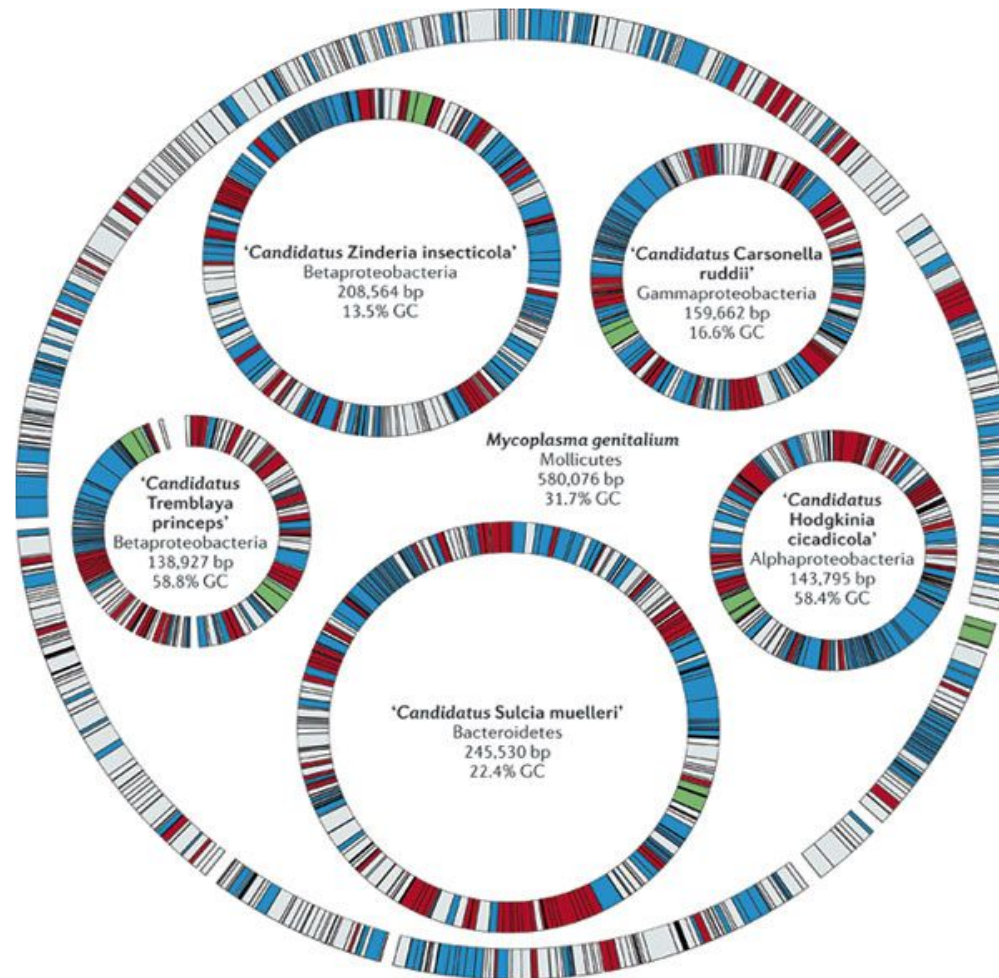
Ma dove vivono i batteri con i genomi super minimi??

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 group<sup>1</sup>. 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.



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> '?
<b>Replication</b>		
<i>dnaE</i>	DNAP III $\alpha$ -subunit	Yes
<i>dnaQ</i>	DNAP III $\epsilon$ -subunit	Yes
<b>Transcription</b>		
<i>rpoA</i>	RNAP $\alpha$ -subunit	Yes
<i>rpoB</i>	RNAP $\beta$ -subunit	Yes
<i>rpoC</i>	RNAP $\beta'$ -subunit	Yes
<i>rpoD</i>	RNAP factor $\sigma^{70}$	No (gene present is a pseudogene)
<b>Protein folding or stability</b>		
<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
<b>tRNA modification</b>		
<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)
<b>Sulphur metabolism</b>		
<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> )
<b>RNA modification</b>		
<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 <sup>‡</sup>	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 $\alpha$ -subunit	No
<i>ileS</i>	Isoleucyl-tRNA synthetase	No
<i>metG</i>	Methionyl-tRNA synthetase	No
<i>pheS</i>	Phenylalanyl-tRNA synthetase, $\alpha$ -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.

<sup>\*</sup>Genes conserved in '*Candidatus* Sulcia muelleri', '*Candidatus* Zinderia insecticola', '*Candidatus* Carsonella ruddii' and '*Candidatus* Hodgkinia cicadicola'.

<sup>‡</sup>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.