

Natural Minimal cells

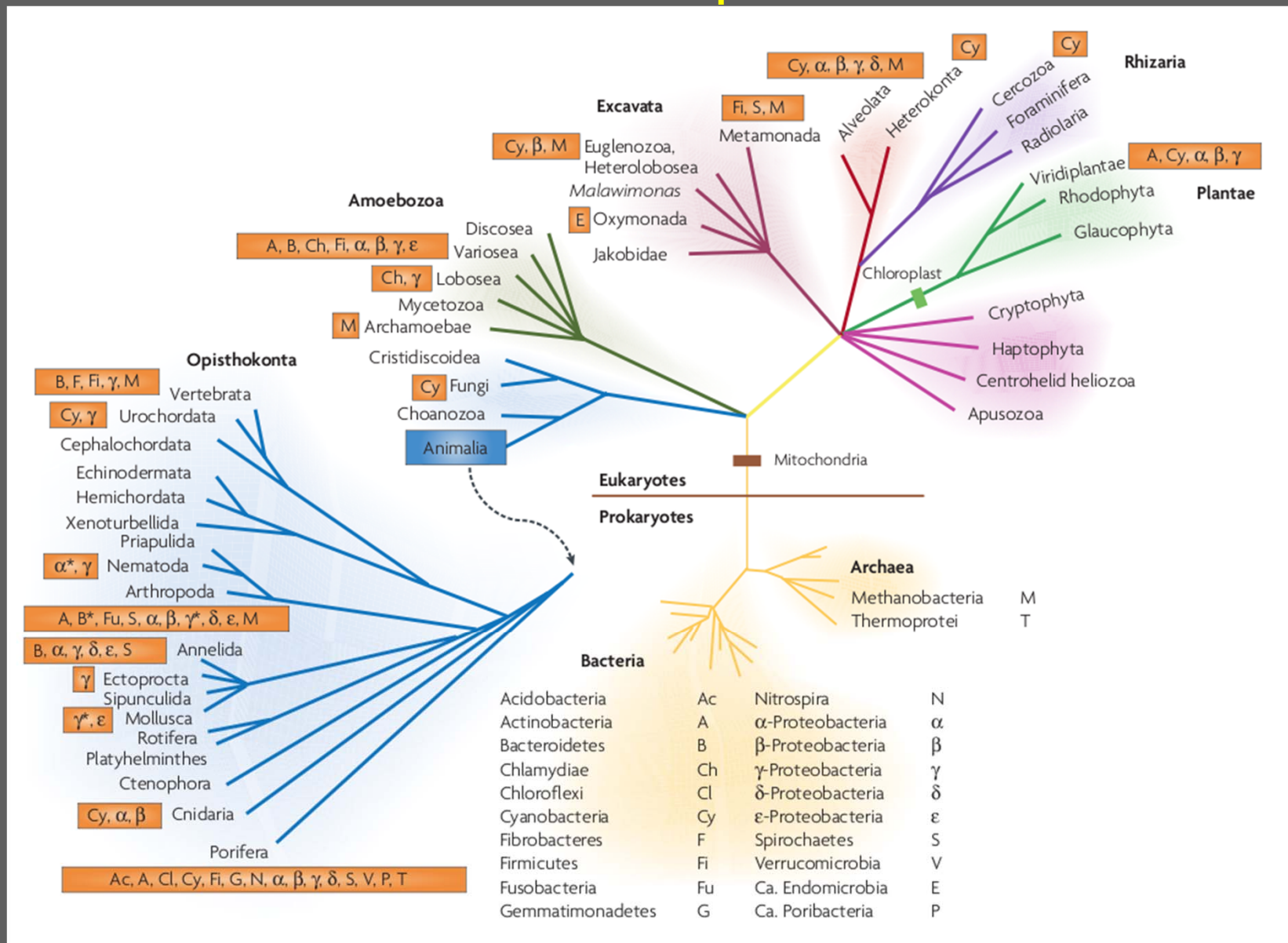
First International School on Biomolecular and Biocellular Computing

Osuna, 5-7 September, 2011

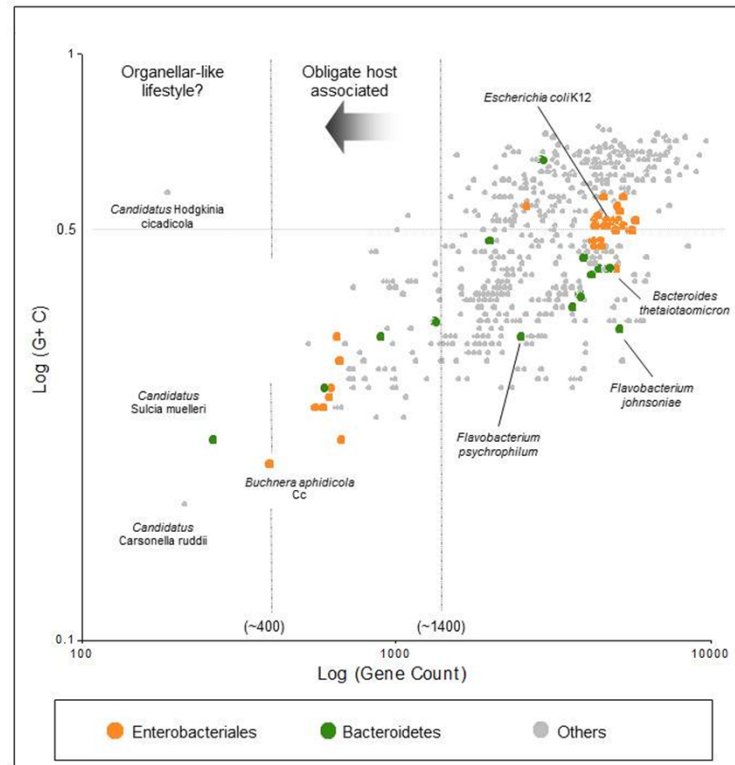
Andrés Moya



Symbiosis is a clearly widespread phenomena in the Biosphere

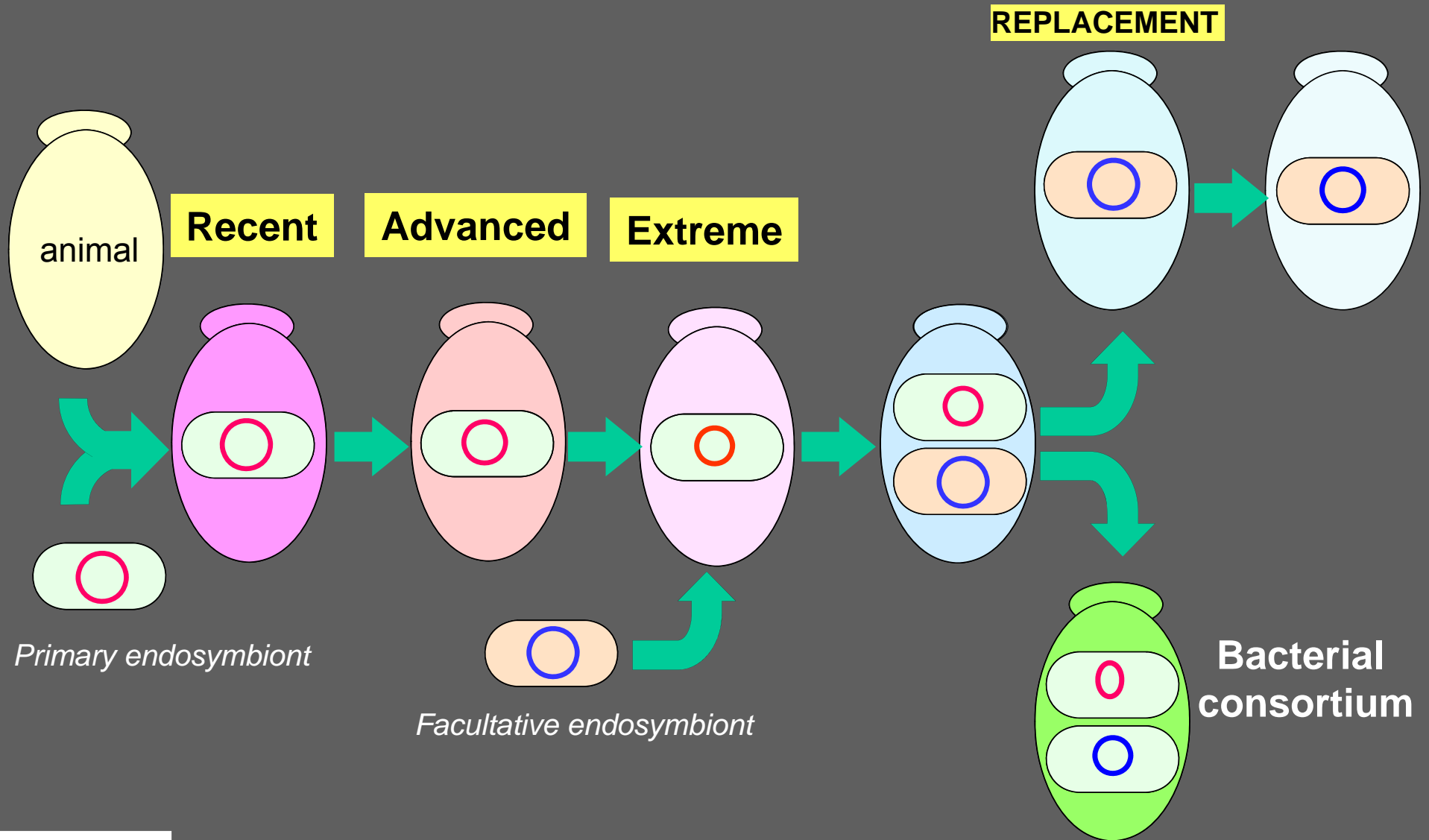


Endosymbiosis produces the smallest genomes



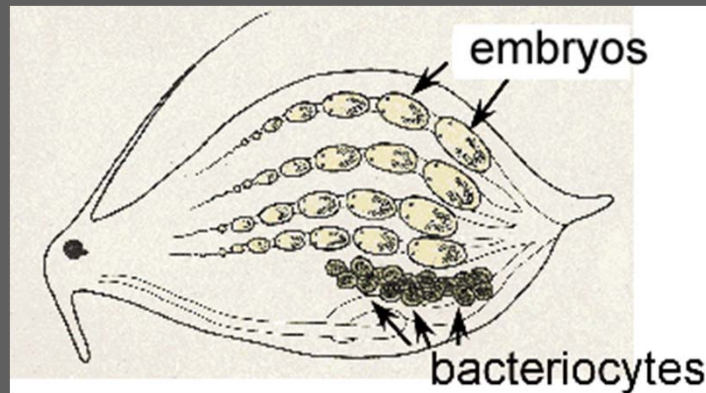
Delaye and Moya. 2010. *Bioessays*.

Establishment, maintenance and evolution of symbiosis

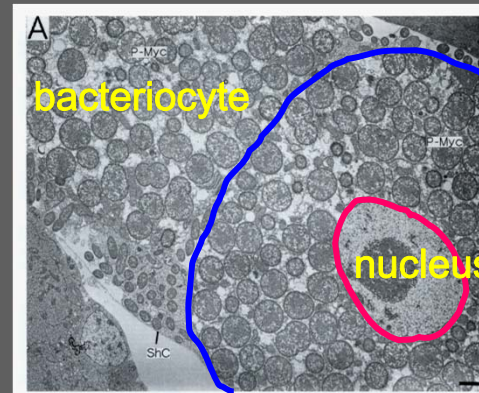


Bacterial symbioses in insects

- More than 20% of insects live in symbiotic relationship with microorganisms, allowing them to exploit a great variety of ecological niches
- In general, such insects feed on unbalanced diets, which are supplemented by **endosymbionts**



longitudinal section of an aphid



- During the adaptation to intracellular life, endosymbionts have undergone **drastic changes in their genomes**



Carpenter ant:
Omnivore



Cockroach:
Omnivore

Insect hosts

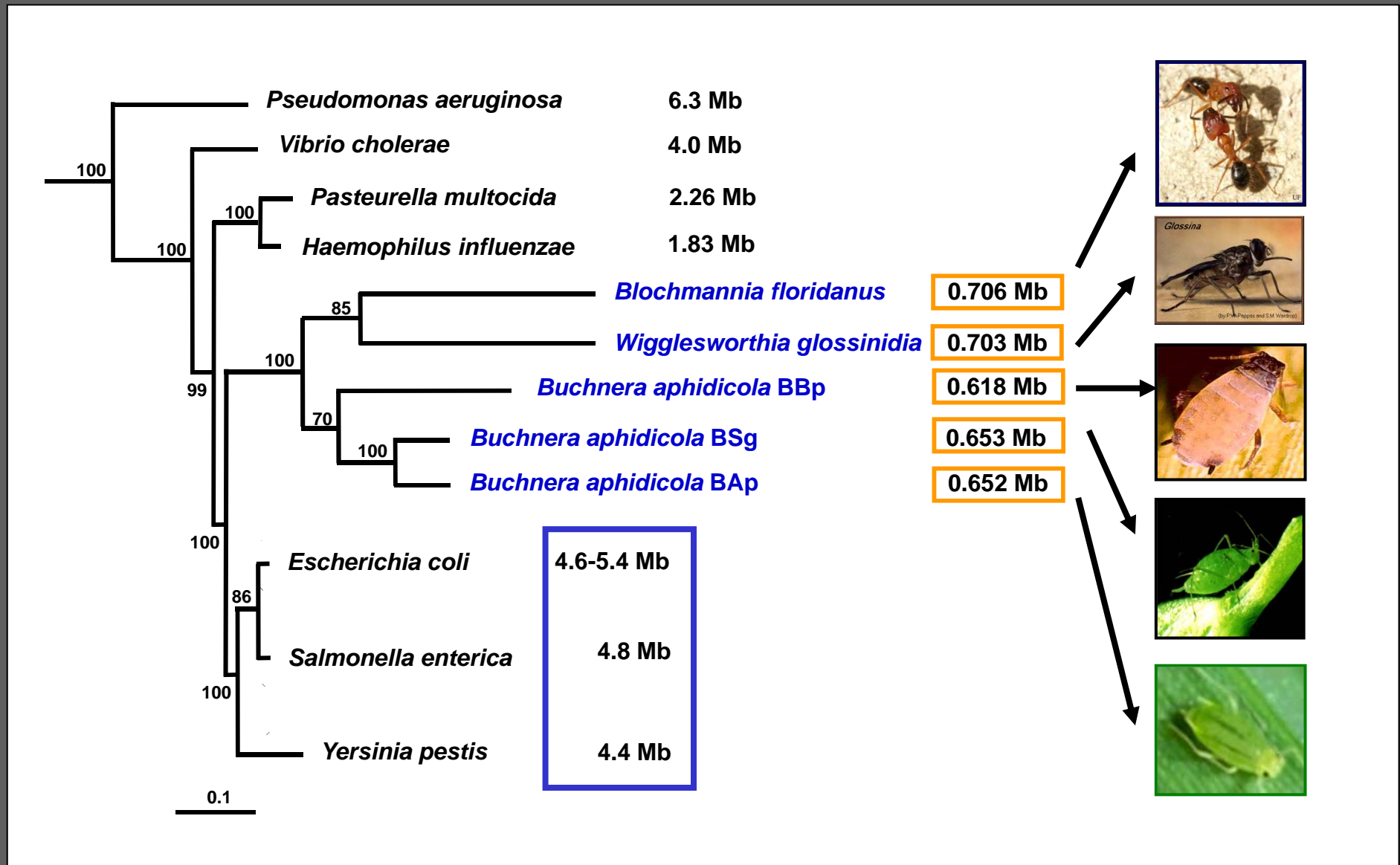


Aphid:
Plant-sap



Tse tse fly:
Blood

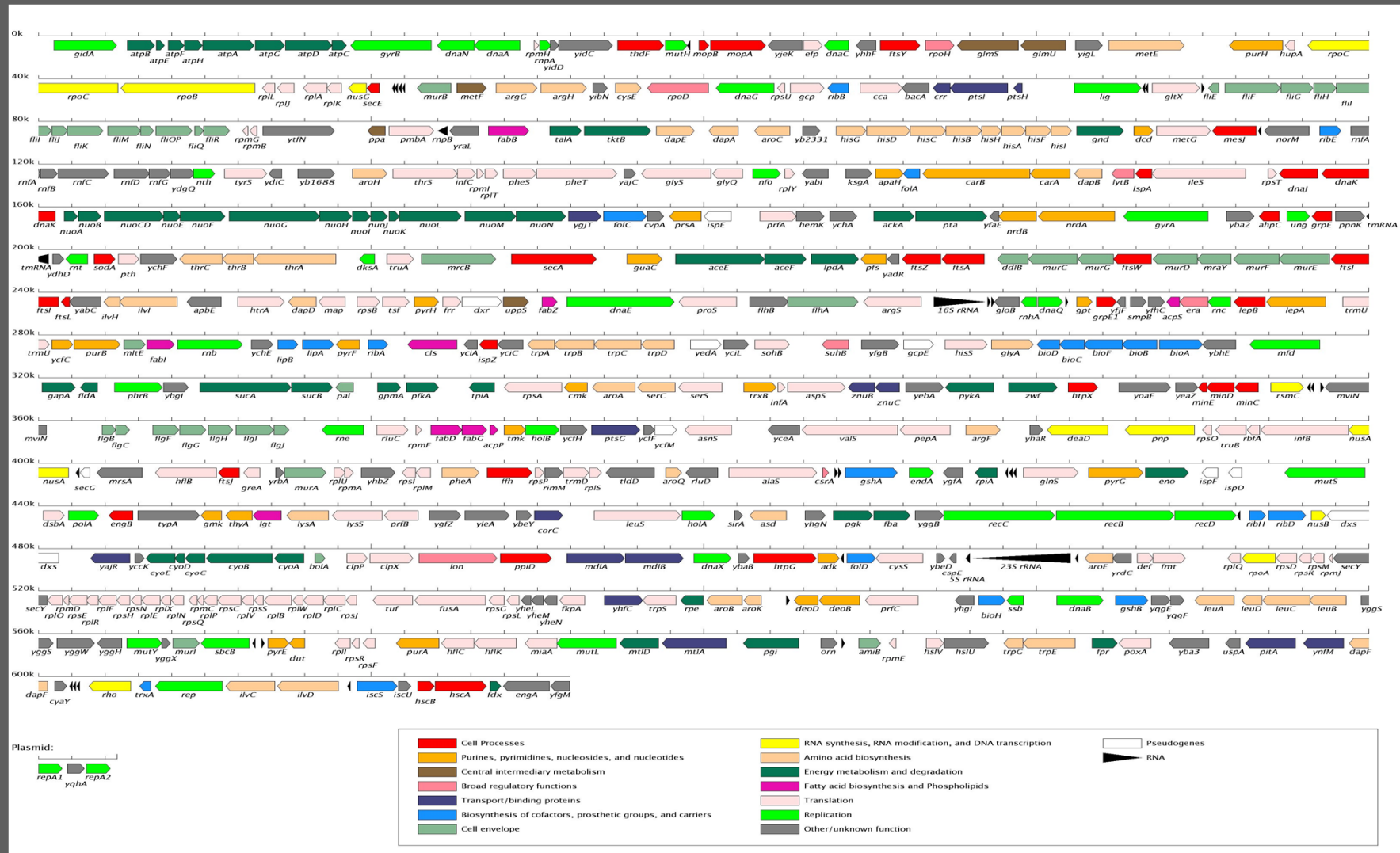
Reduced genome size in Gamma-proteobacteria



Aphid (*Baizongia pistaciae*)



B. aphidicola BPI genome map



van Ham et al. 2003. PNAS.

Carpenter ant (*Camponotus floridanus*)



Blochmannia floridanus genome map

Blochmannia genome

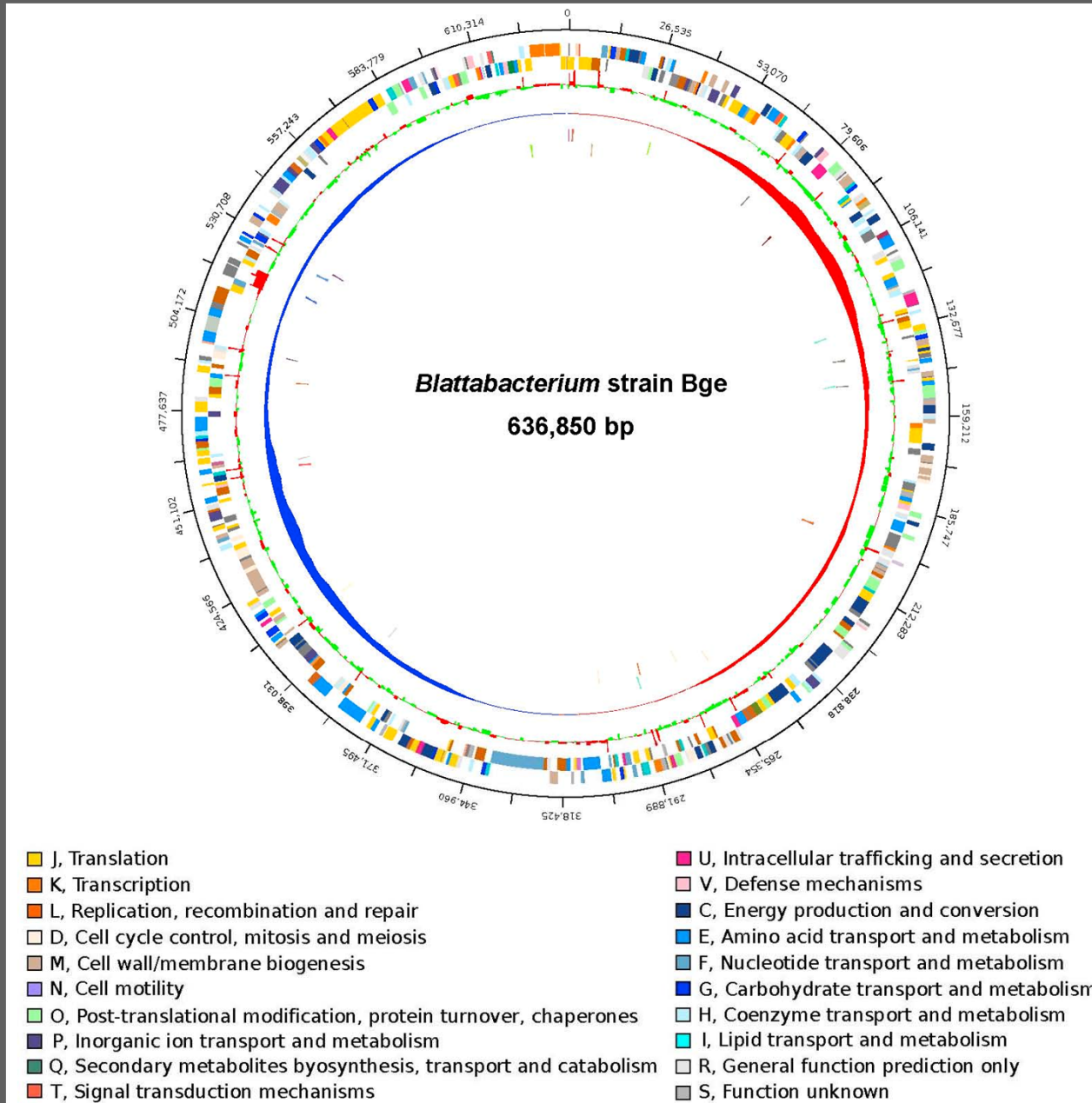


- RNAs
- Pseudogenes
- Energy production and conversion **C**
- Cell division and chromosome partitioning **D**
- Amino acid transport and metabolism **E**
- Nucleotide transport and metabolism **F**
- Carbohydrate transport and metabolism **G**
- Coenzyme metabolism **H**
- Lipid metabolism **I**
- Translation, ribosomal structure and biogenesis **J**
- Transcription **K**
- DNA replication, recombination and repair **L**
- Cell envelope biogenesis, outer membrane **M**
- Posttranslational modification, protein turnover, chaperones **O**
- Inorganic ion transport and metabolism **P**
- Signal transduction mechanisms **T**
- Cell motility and secretion **N**
- General function prediction only **R**
- Function unknown **S**

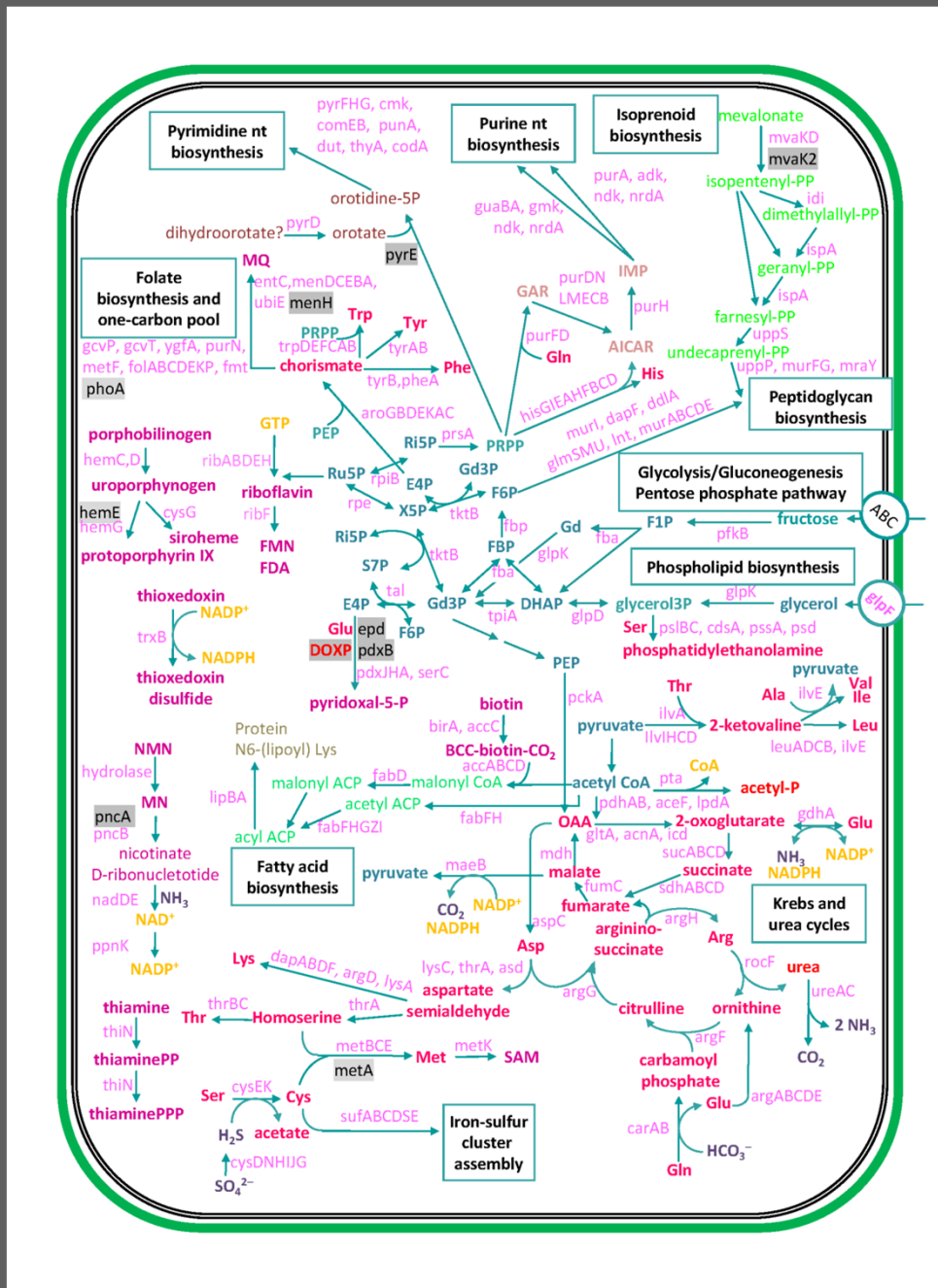
Cockroach (*Blatella germanica*)



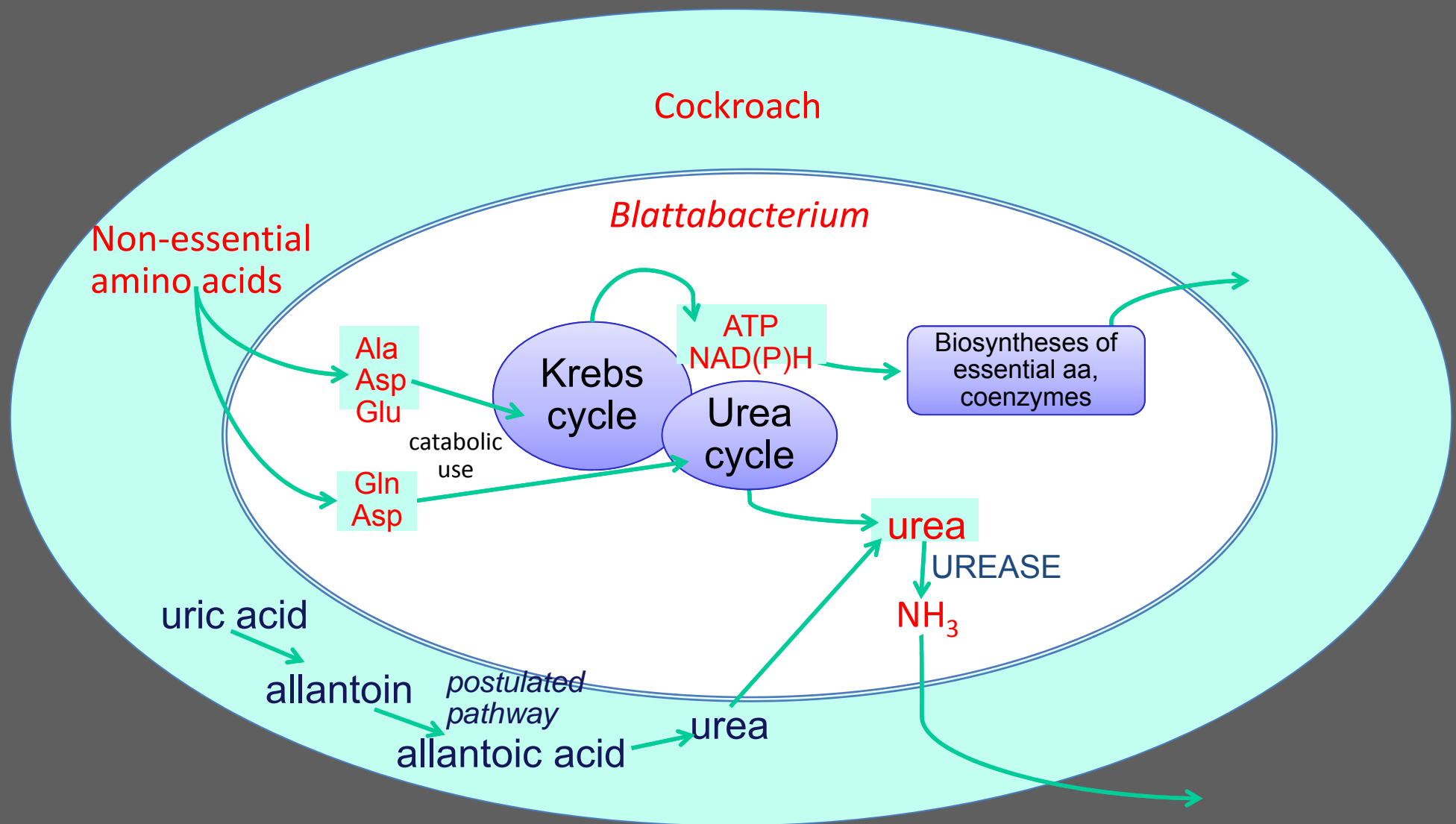
Circular map of *Blattabacterium* sp. Bge genome



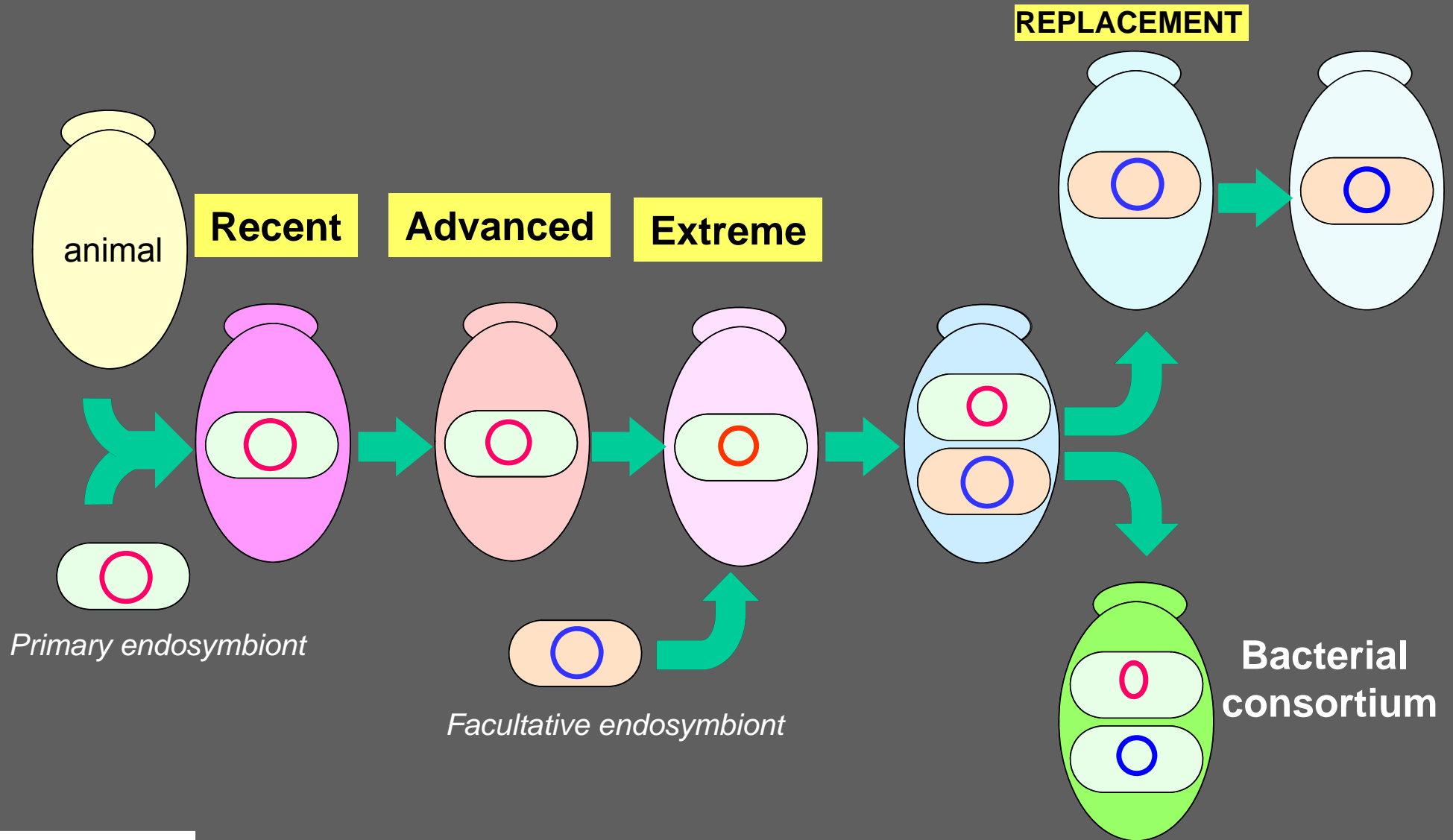
Blattabacterium sp. Bge detailed metabolism



Blattabacterium sp. Bge metabolism



Establishment, maintenance and evolution of symbiosis



Symbiotic consortia



CENTRO SUPERIOR DE
INVESTIGACIÓN EN
SALUD PÚBLICA

VNIVERSITAT ID VALÈNCIA

Institut Cavanilles de Biodiversitat i Biologia Evolutiva

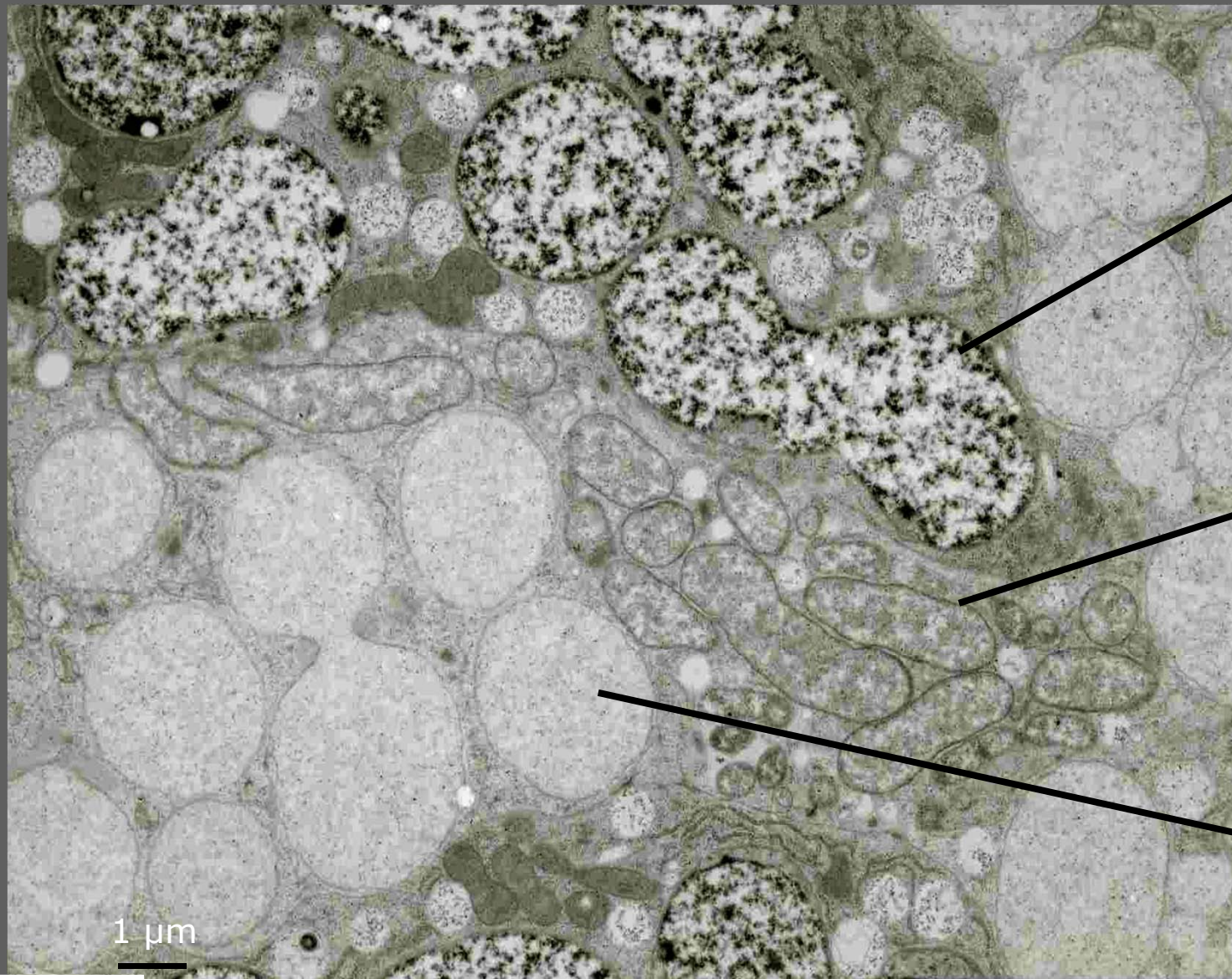
Aphid (*Cinara cedri*)



Buchnera aphidicola: 416 kb



Endosymbionts of *C. cedri*



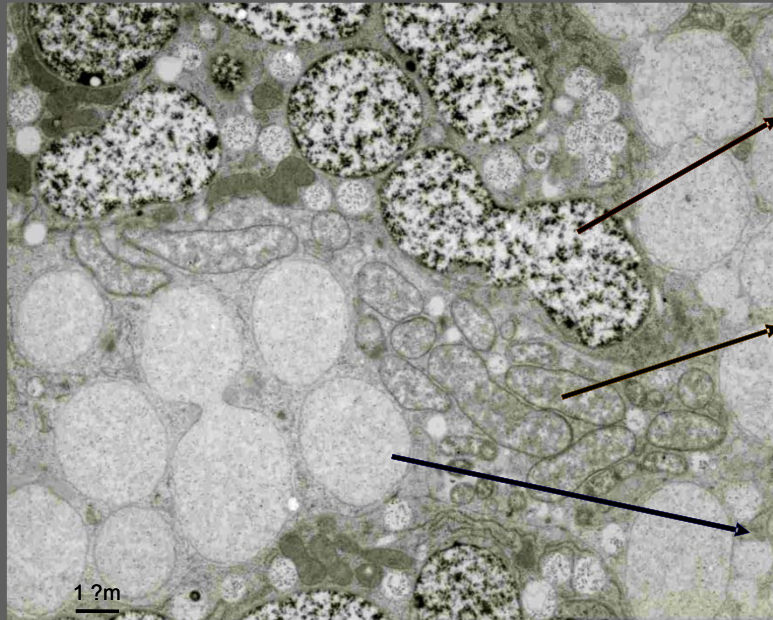
B. aphidicola

Wolbachia

Serratia symbiotica

Morphological analyses

B. aphidicola coexists with *Candidatus Serratia symbiotica*

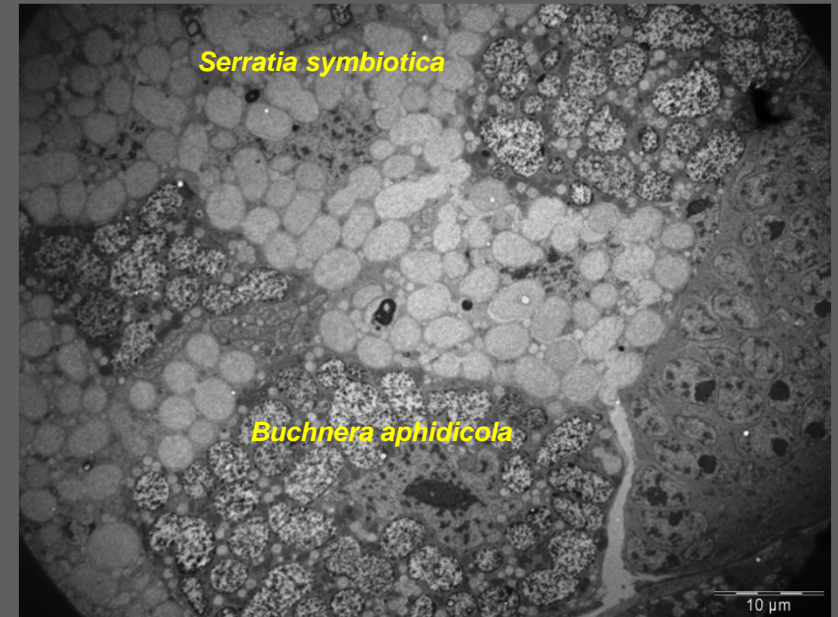


Ultrathin

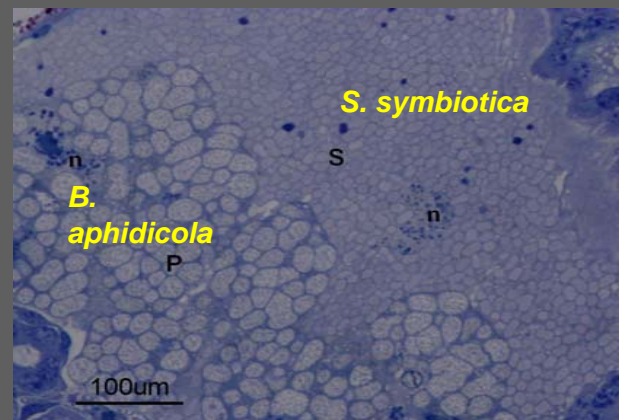
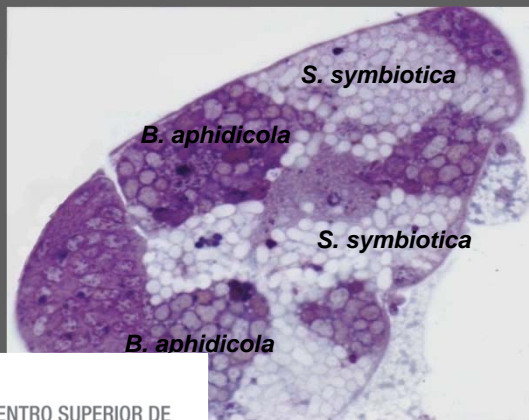
B. aphidicola

Wolbachia

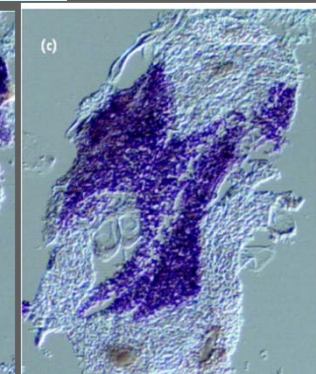
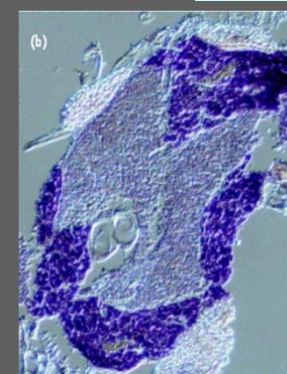
Candidatus
Serratia
symbiotica



Semithin



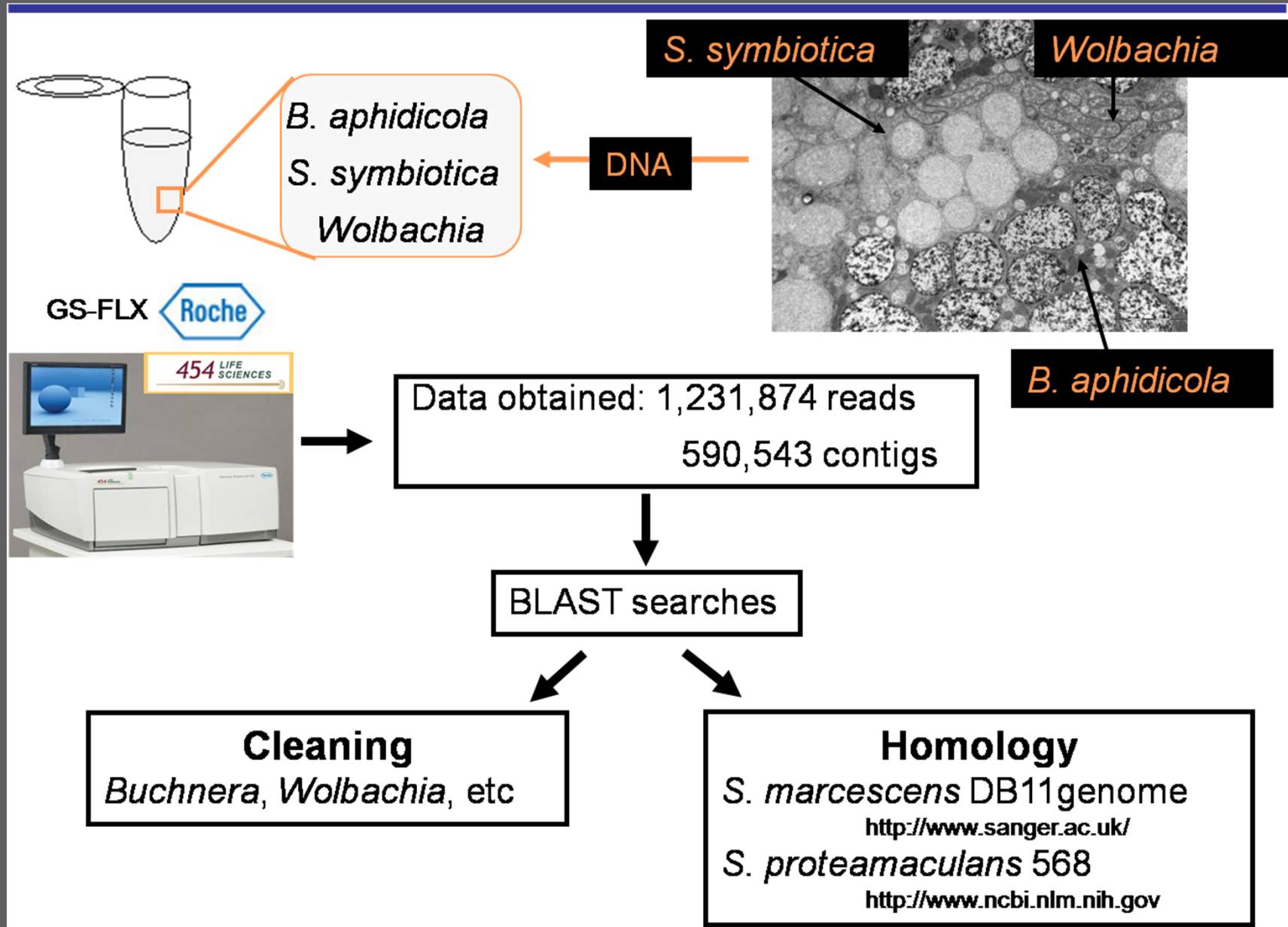
FISH



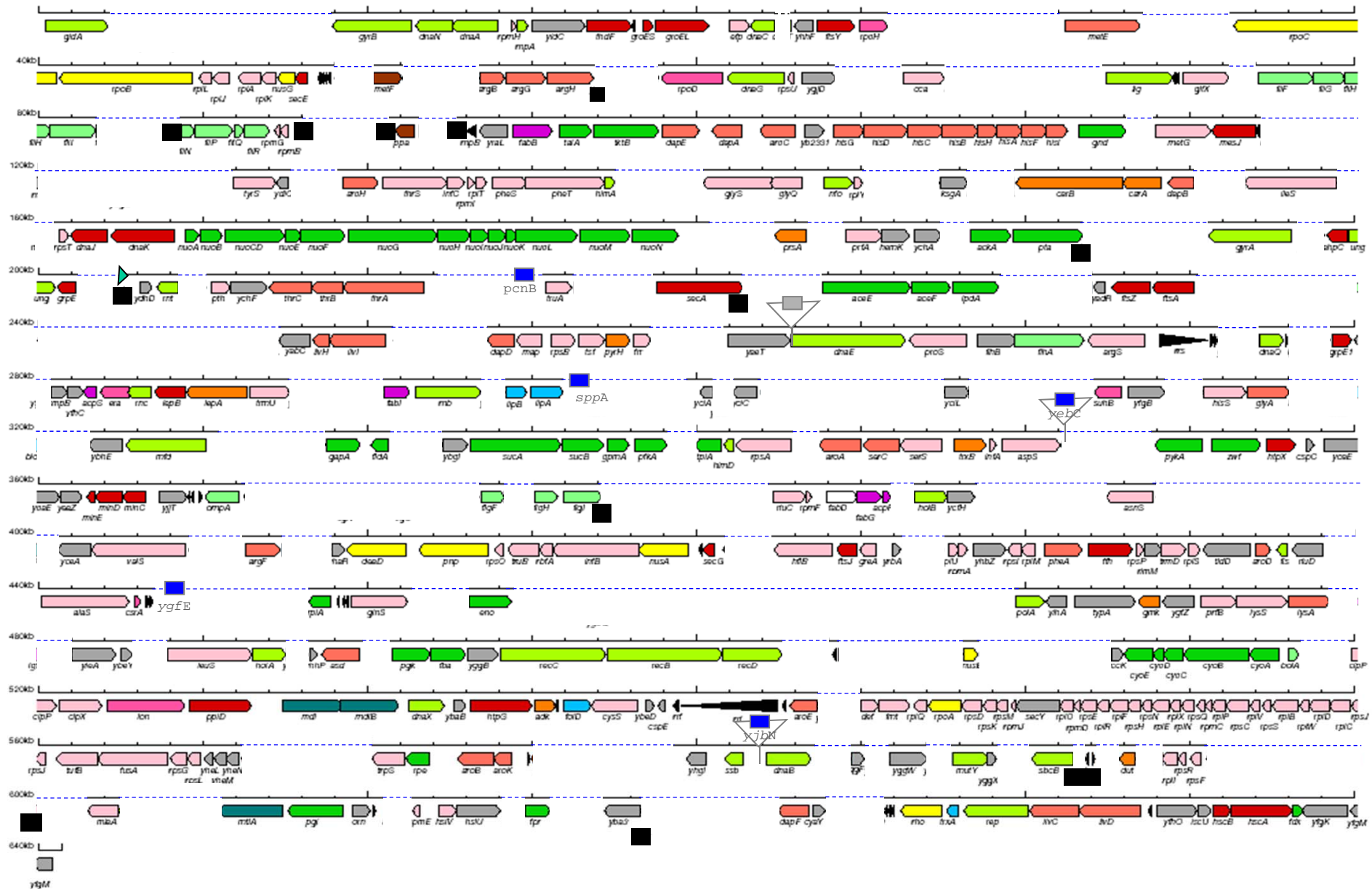
B. aphidicola

S. symbiotica

Metagenomic approach

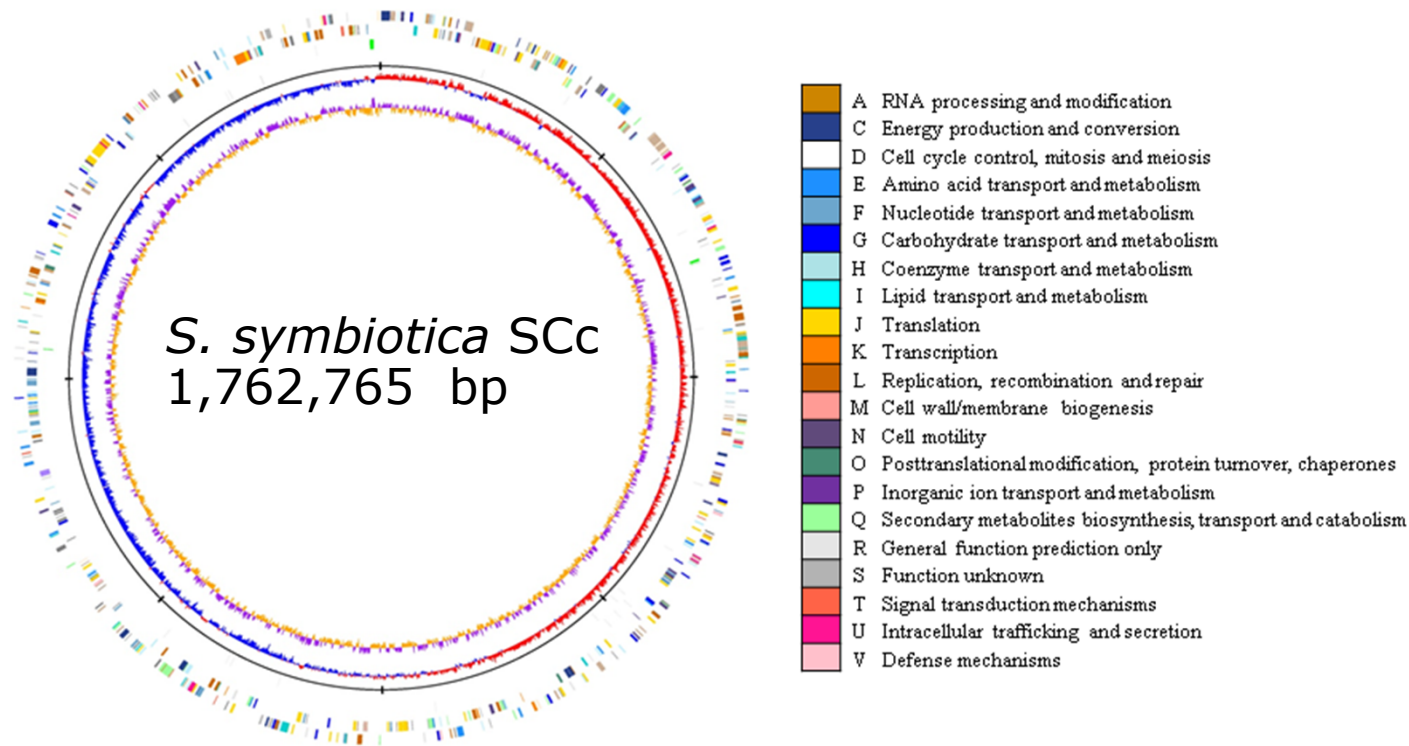


Smallest *Buchnera* genome



Serratia symbiotica SCc genome map

Fig. S1



Genomic data

	<i>S. proteamaculans</i>	<i>S. symbiotica</i> SCc	<i>B. aphidicola</i> BCc
Life style	Free-living	Endosymbiont	Endosymbiont
Genome size (Mbp)	5.495	1.7	0.422
-chromosome	5.448	1.7	0.416
-plasmid	0.047	-	0.006
ORFS	4891 + 51	822	357 + 5
G+C content (%)	55	30	20,2
rDNA copies			
-16S	7	2	1
-23S	7	2	1
-5S	8	1	1
tRNA-aminoacil synthase	86	24	23
RecA	+	-	-
Insertion Sequences (IS)	4 IS type ISSpr1 5 IS type ISSpr2	not found	not present

Functional analyses

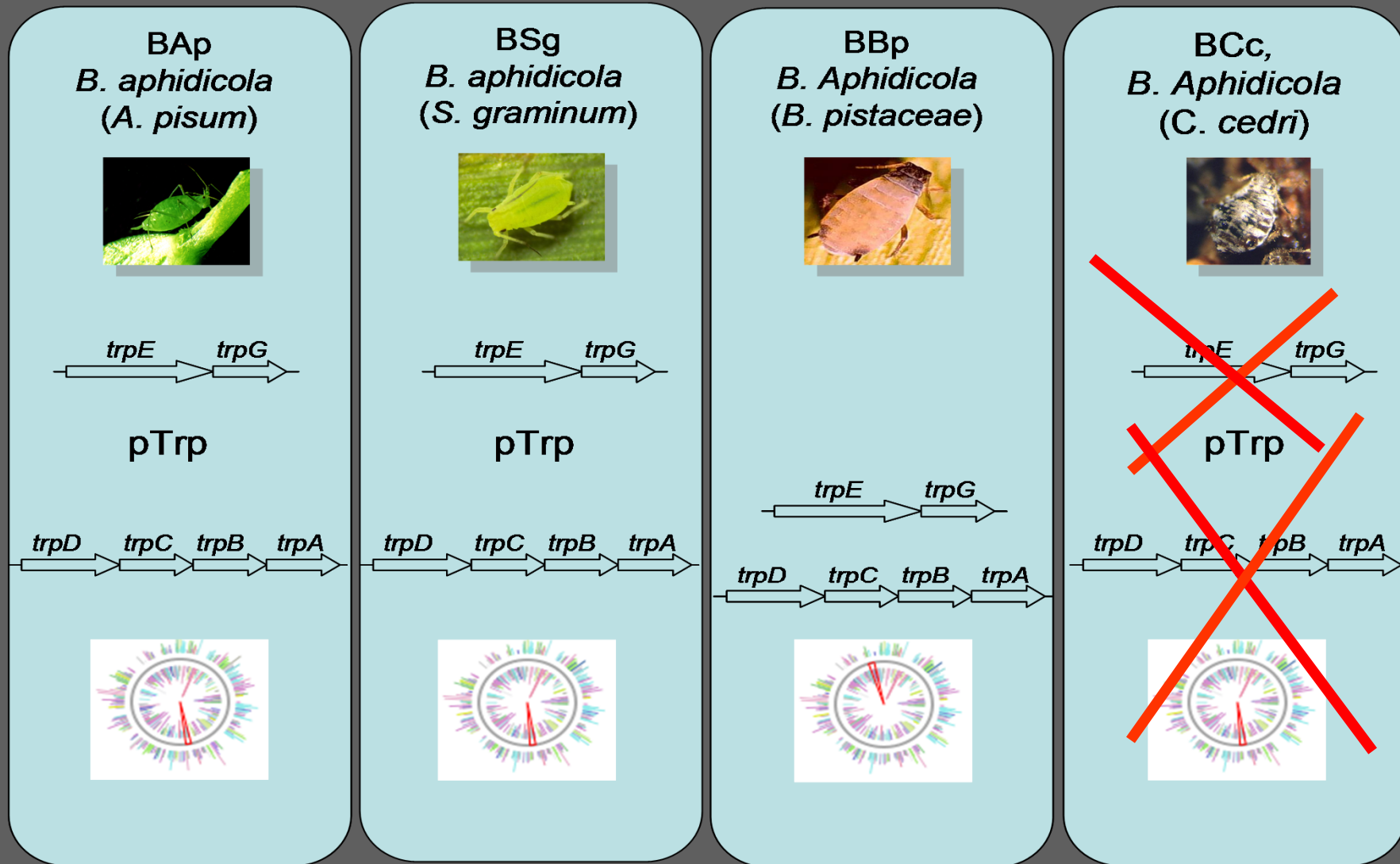
	Cofactors and vitamins biosynthetic capability	
	BCc	SCc
Biotin	-	+
Lipoate	+	+
Folate	-	+
Glutathione	-	+
Heme and porphirin	-	+
Nicotinate and nicotinamide	-	+
Pantothenate and CoA	-	+
Riboflavin	-	+
Thiamine	-	+
Ubiquinone	-	+



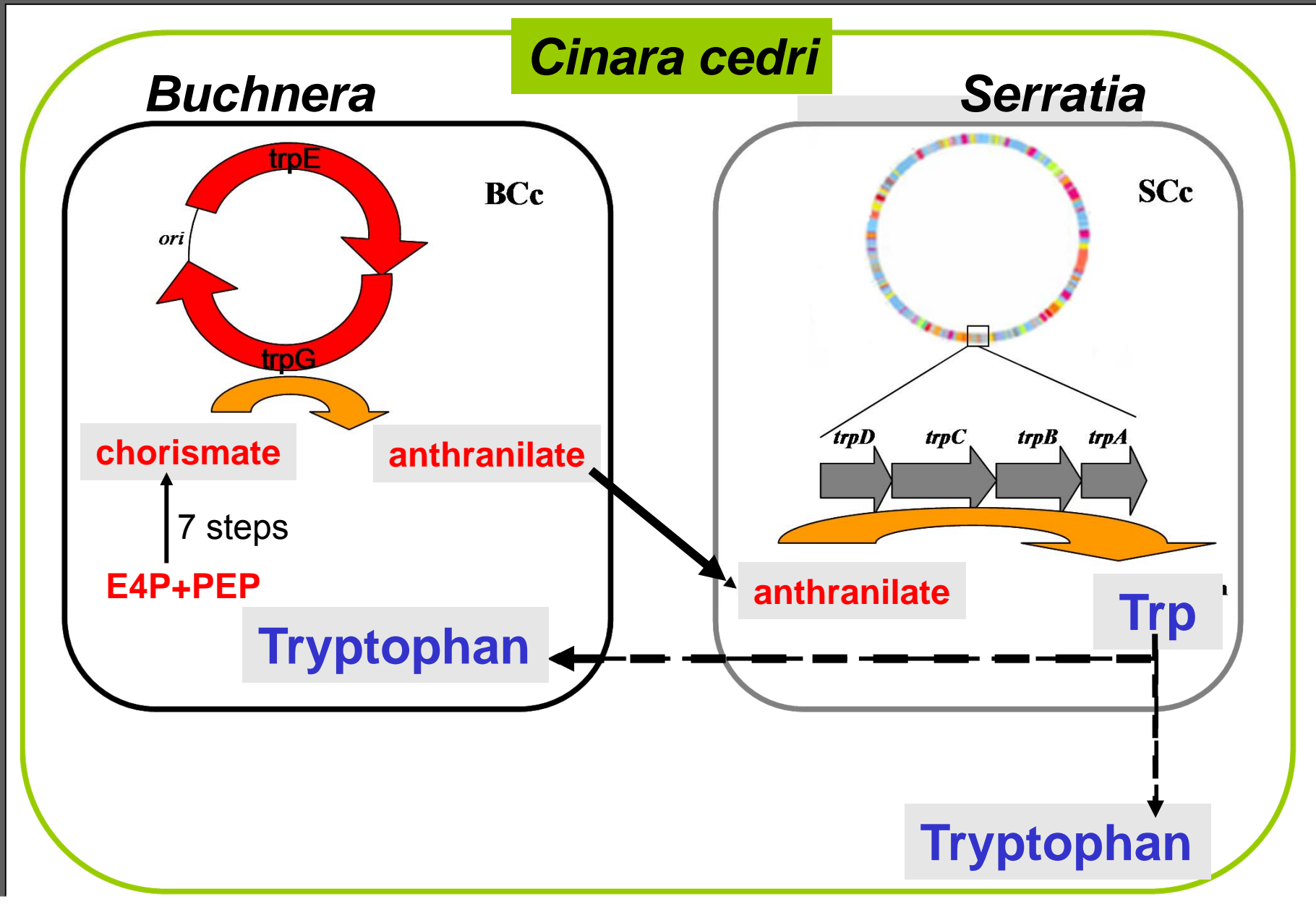
Functional analyses

	Amino acid biosynthetic capability	
	BCc	SCc
Essential		
Arginine	+	-
Valine	+	-
Leucine	+	-
Isoleucine	+	-
Lysine	+	-
Threonine	+	+
Methionine	-	+
Histidine	+	-
Phenylalanine	+	-
Tryptophan	-/+	+/-
Nonessential		
Tyrosine	+	-
Proline, Glutamate, Glutamine, Cysteine, Aspartate, Asparagine, Alanine, Glycine, Serine	-	-

Tryptophan provision in *B. aphidicola*



Metabolic complementation



Tremblaya princeps: nested endosymbiosis

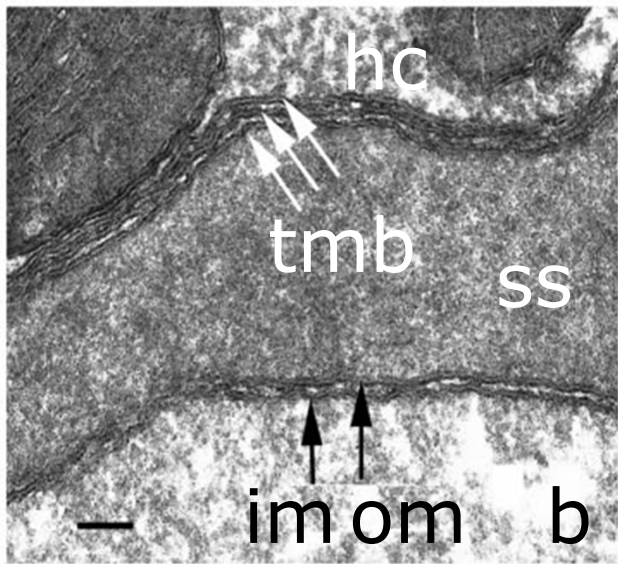
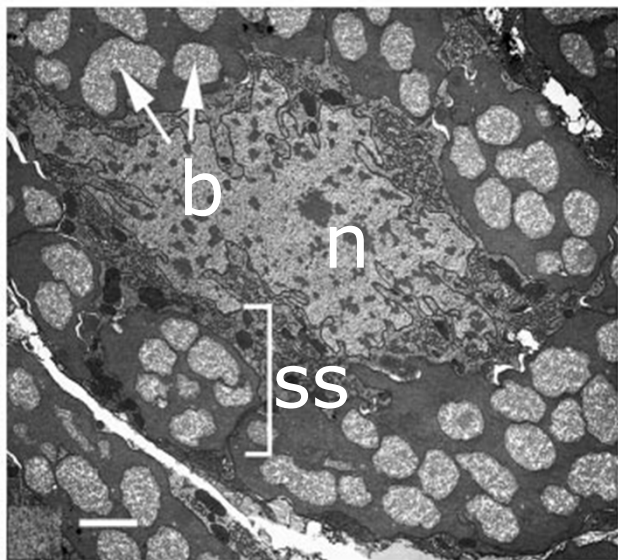


The mealybug *Planococcus citri* (cotonet)

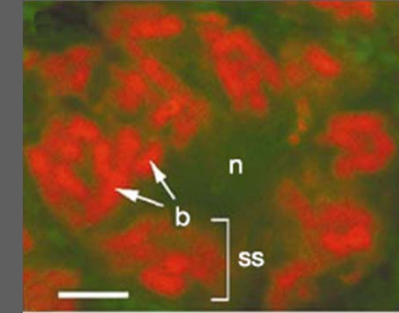


Endosymbionts described in the mealybug *Planococcus citri* (cotonet)

bacteriocyte

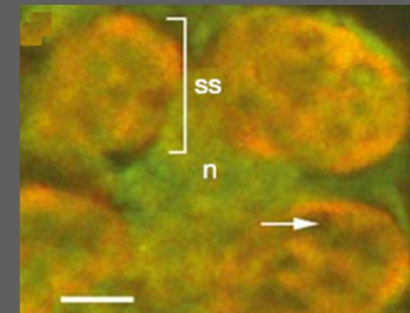


b: bacteria
im: inner membrane
om: outer membrane
hc: host cell cytoplasm
n: nucleus
ss: symbiotic spheres
tmb: three membrane bilayers



γ -proteobacterial probe

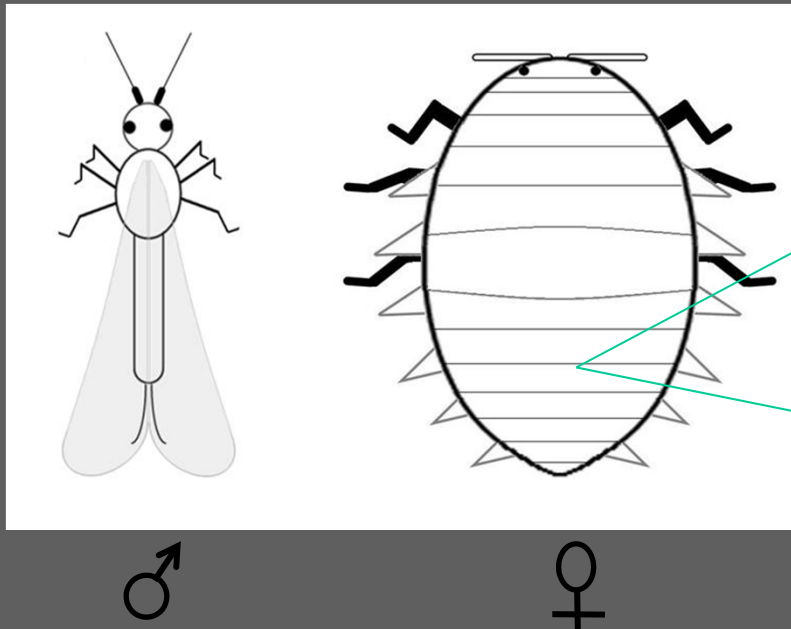
FISH (visualized by
laser-scanning
confocal microscopy)



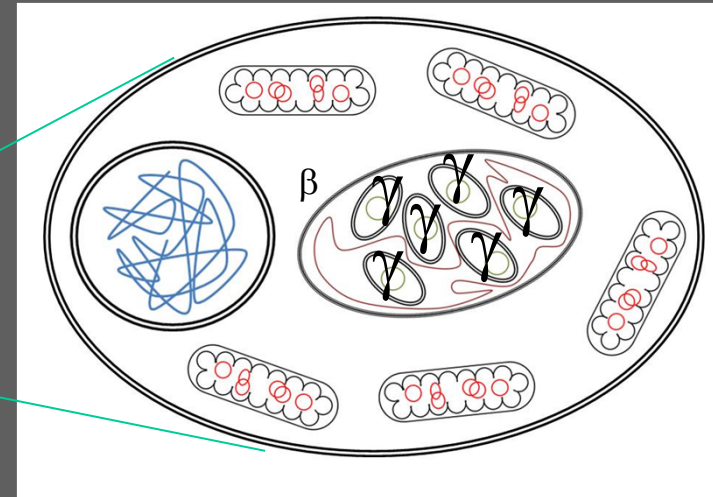
β -proteobacterial probe

von Dohlen et al. 2001. *Nature*.

A bacteria living inside another bacteria



Bacteriocyte



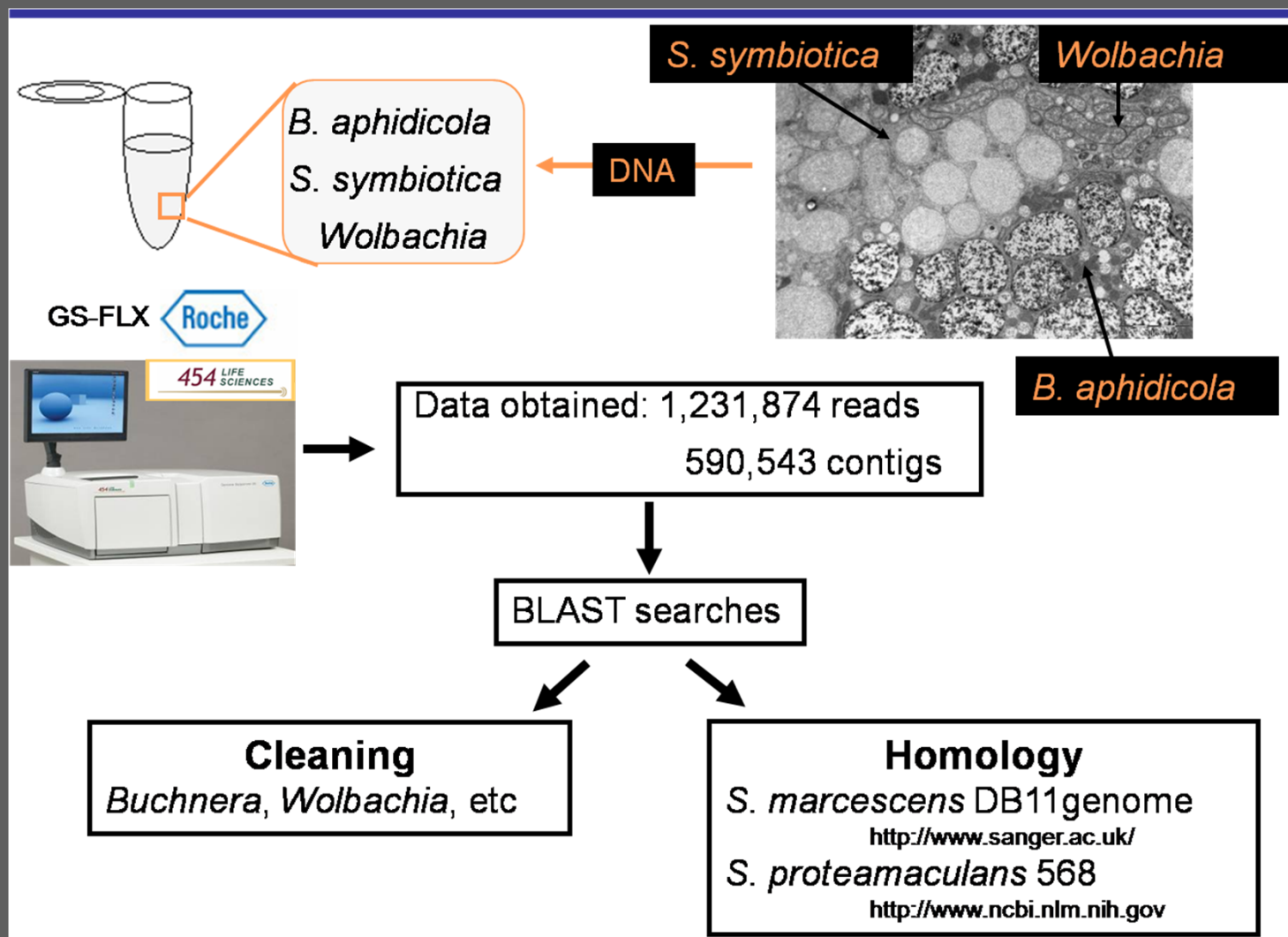
Primary endosymbiont:

β -proteobacteria *Tremblaya princeps*

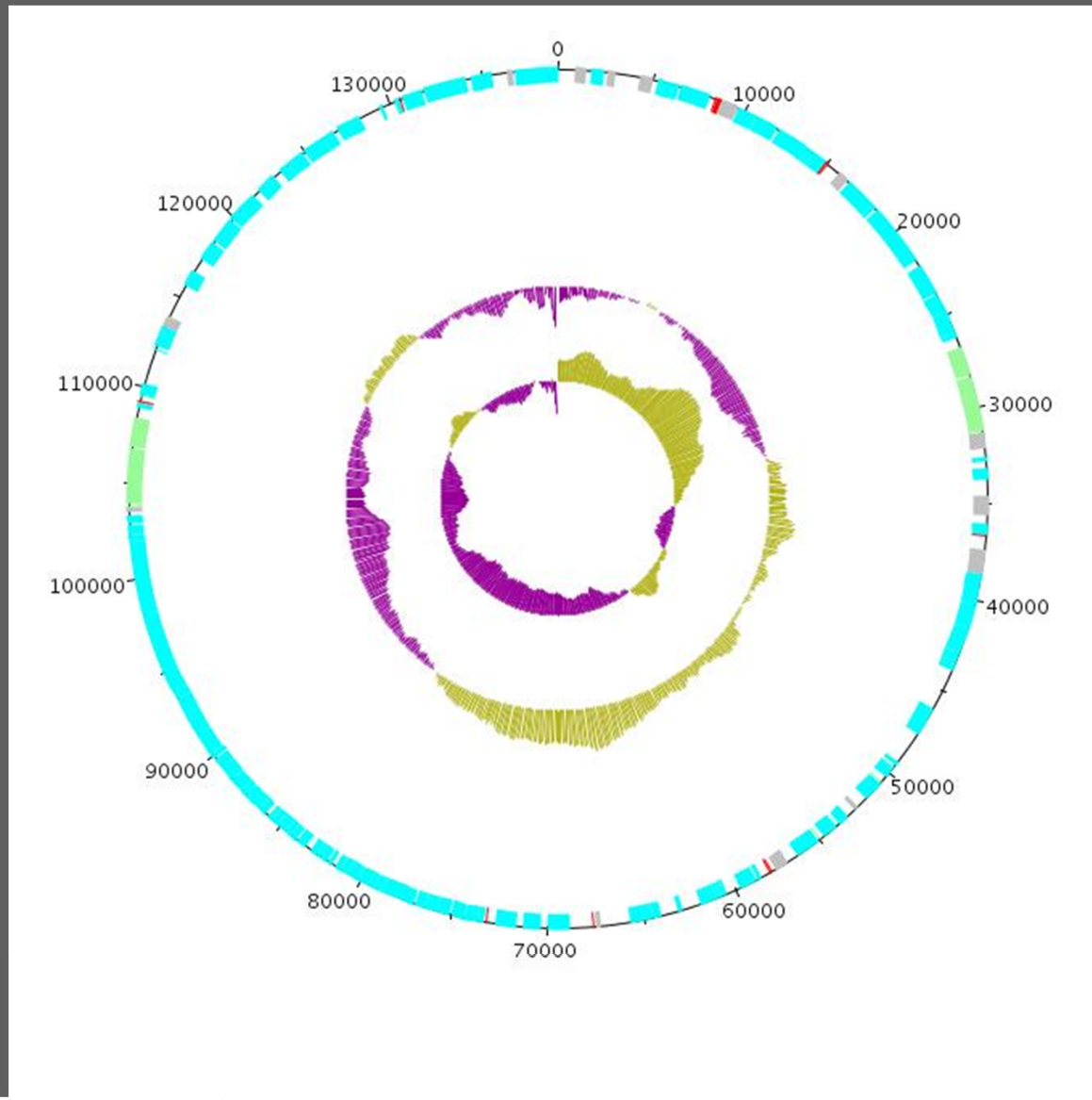
Secondary endosymbiont:

γ -proteobacteria

...and applying the same metagenomic approach we carried out with the aphid *C. cedri*



Tremblaya princeps genome



Outer circle:

- Blue: detected CDS.
- Green: Ribosomal operons
- Red: functional tRNAs and tmRNA
- Grey: pseudogenes

Inner Circles: GCskew

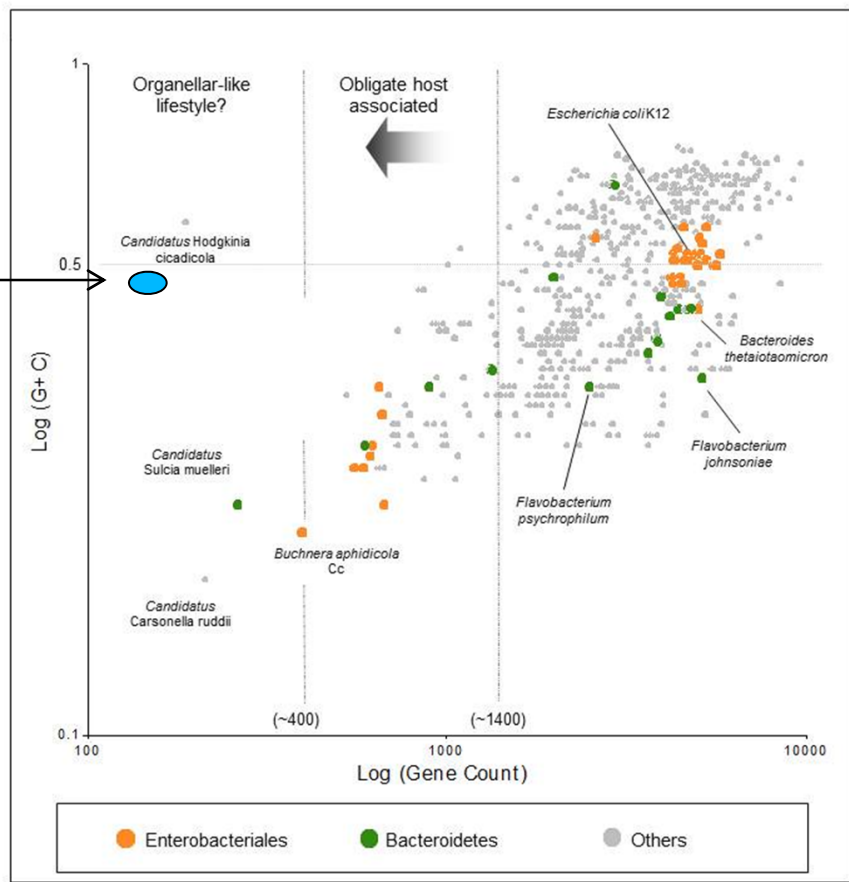
López-Madrugal et al., *J. Bacteriology*, in press.
McCutcheon and von Dohlen, *Current Biology*, 2011.

Some comparative features

	<i>Tremblaya princeps</i>	<i>Candidatus Hodgkinia cicadicola</i> Dsem	<i>P. citri</i> secondary endosymbiont (PCSE)	<i>Buchnera aphidicola</i> BCc
Genome length (kb)	139	144	515	416
Genes	131	189	451	397
Protein coding genes	112	169	413	357
tRNAs	8	16	35	31
Pseudogenes	21	-	33	3
Ribosomal gene set	2	1	1	1
GC content (%)	59	58	44	20

Endosymbiosis produces the smallest genomes

Tremblaya princeps



Minimal Cells and Synthetic Biology

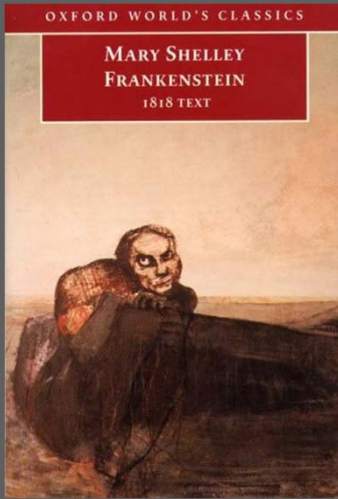


CENTRO SUPERIOR DE
INVESTIGACIÓN EN
SALUD PÚBLICA

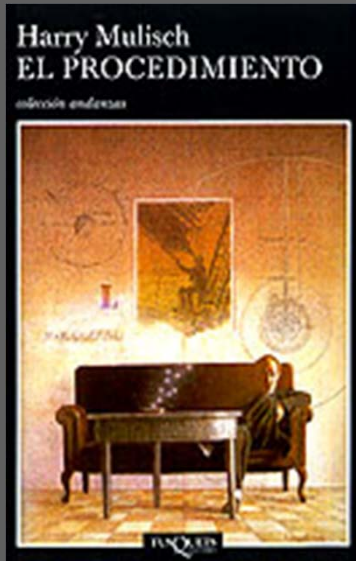
VNIVERSITAT ID VALÈNCIA

Institut Cavanilles de Biodiversitat i Biologia Evolutiva

Synthetic life: two strategies



- Victor Frankenstein's way or top-down



- Victor Werker's way or bottom-up

Synthetic life: two strategies

Theoretical and experimental comparative genomics: minimal genome. Synthetic genome in a genome-free cell. New genetic and metabolic circuits. Interchangeable parts...

à la Frankenstein

Top-down

living
system

Bottom-up

à la Werker

Peretó and Català (2007)
“The Renaissance of
Synthetic Biology” *Biol Theor*
2:128-130

Fundamental concepts
(autonomy/autopoiesis, self-
replication) and their chemical
implementation

From an analytical science to a synthetic one

Total Synthesis of a Tyrosine Suppressor Transfer RNA Gene

XVI. ENZYMATIC JOININGS TO FORM THE TOTAL 207-BASE PAIR-LONG DNA*

(Received for publication, August 17, 1978)

Takao Sekiya,‡ Tatsuo Takeya,§ Eugene L. Brown,¶ Ramamoorthy Belagaje, Roland Contreras,||
Hans-Joachim Fritz,** Michael J. Gait,*** Robert G. Lees,§§ Michael J. Ryan,§§§ and H. Gobind
Khorana

From the Departments of Biology and Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Kjeld E. Norris

From The Danish Institute of Protein Chemistry, DK-2970 Hørsholm, Denmark

4 YEARS 207-bp

JBC 254:5787, 1979

Total Synthesis and Cloning of a Gene Coding for the Ribonuclease S Protein

KRISHNAN P. NAMBIAR

JOSEPH STACKHOUSE

DORA M. STAUFFER

W. POINDEXTER KENNEDY

J. K. ELDREDGE

STEVEN A. BENNER

Science 223:1299, 1984

1.5 YEARS 330-bp

Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template

Jeronimo Cello, Aniko V. Paul, Eckard Wimmer*

Science 297:1016, 2002

MONTHS 7740-bp

Generating a synthetic genome by whole genome assembly: ϕ X174 bacteriophage from synthetic oligonucleotides

Hamilton O. Smith, Clyde A. Hutchison III*, Cynthia Pfannkoch, and J. Craig Venter*

PNAS 100:1299, 2003

14-DAYS 5386-bp

From an analytical science to a synthetic one

Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster

Sarah J. Kodumal, Kedar G. Patel, Ralph Reid, Hugo G. Menzella, Mark Welch, and Daniel V. Santi*

PNAS 101:15573, 2004

~1 MONTH 32,000-bp

24 January 2008 / Page 1 / 10.1126/science.1151721

Scienceexpress

Research Article

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, Jayshree 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*

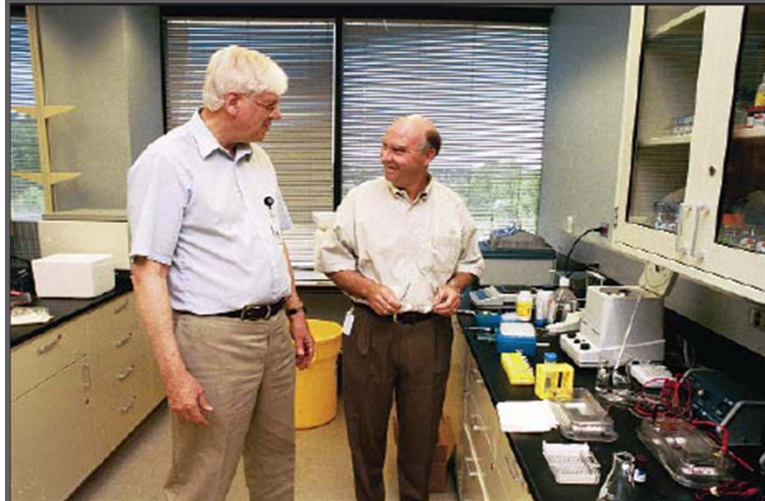
MONTHS? 582,970-bp

SCIENCE VOL 298 29 NOVEMBER 2002

GENETICS

Venter Gets Down to Life's Basics

Never shy about his aims, DNA sequencer J. Craig Venter Jr. announced this week that he has won a government grant to design a novel form of life. The U.S. Department of Energy's science office has awarded his group \$3 million over 3 years to "develop a synthetic chromosome," the first step toward making a self-replicating organism with a completely artificial genome.



Playing God? Craig Venter (right) and Hamilton Smith predict that they will have a synthetic genome up and running in 3 years.

The minimal genome concept

Minimal genome (Koonin, 2000):

“The smallest possible group of genes that would be sufficient to sustain a functioning cellular life form under the most favorable conditions imaginable:

- *in the presence of a full complement of essential nutrients*
- *in the absence of environmental stress”*

Essential functions:

- cell structure maintenance
- reproduction
- evolution

Our approach to the minimal bacterial genome

MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, Sept. 2004, p. 518–537
1092-2172/04/\$08.00+0 DOI: 10.1128/MMBR.68.3.518–537.2004
Copyright © 2004, American Society for Microbiology. All Rights Reserved.

Vol. 68, No. 3

Determination of the Core of a Minimal Bacterial Gene Set†

Rosario Gil,^{1,2*} Francisco J. Silva,^{1,2} Juli Peretó,^{1,3} and Andrés Moya^{1,2}

*Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, Valencia,¹ and
Departament de Genètica² and Departament de Bioquímica i Biologia Molecular,³
Universitat de València, Burjassot (València), Spain*

1. Comparison of reduced genomes

Five complete genomes from insect endosymbionts (500-600 genes/0.6-0.7Mb)
The human parasite *Mycoplasma genitalium* (480 genes/0.54Mb)
Phytoplasma asteris (754 genes/0.86Mb)

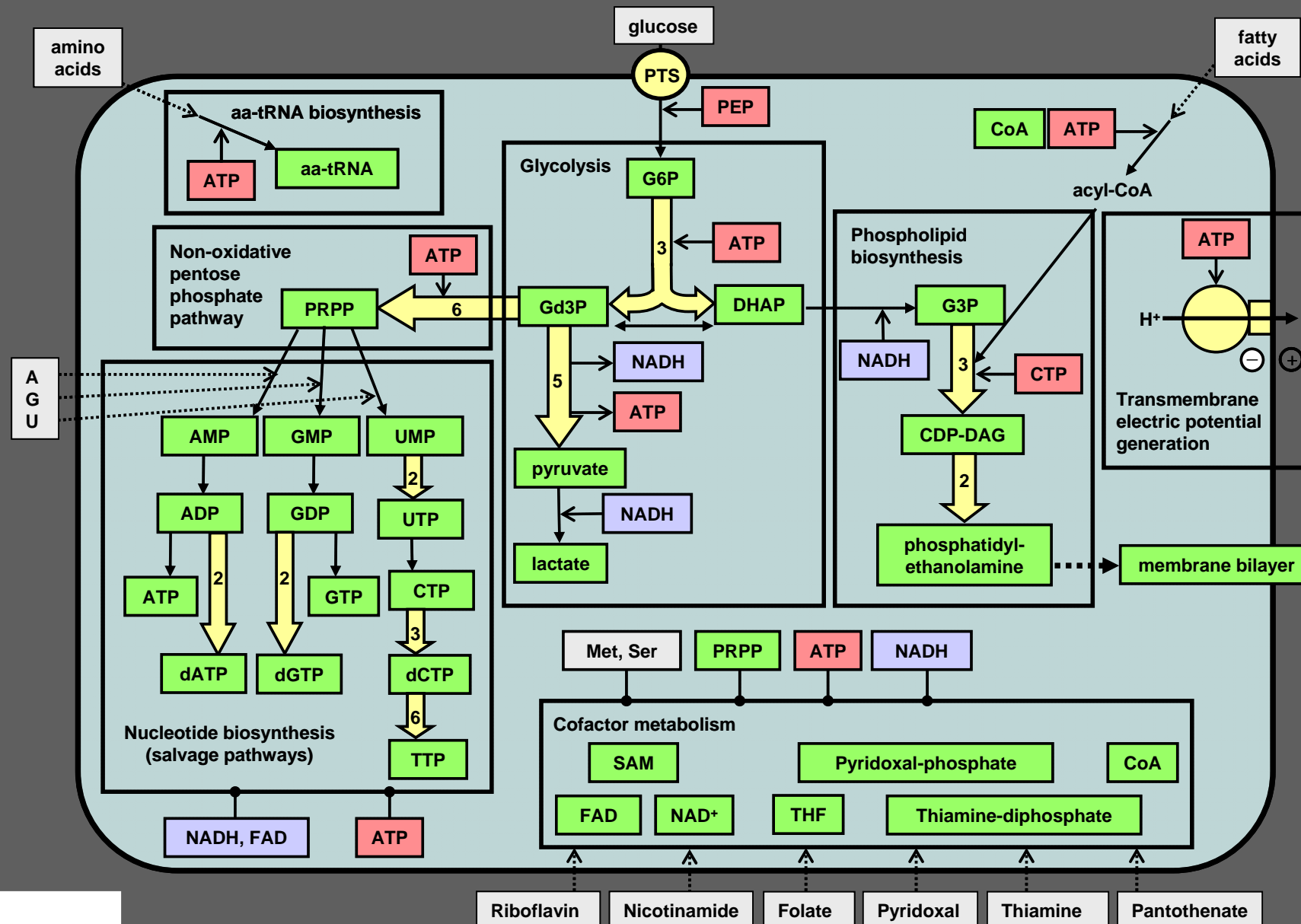
2. Data on experimental genome minimization

M. genitalium, *M. pneumoniae* (global mutagenesis, Hutchison III *et al.* 1999)
Bacillus subtilis (systematic inactivation, Kobayashi *et al.* 2003)
Escherichia coli (global mutagenesis, Gerdes *et al.* 2003)
Staphylococcus aureus (antisense RNA, Forsyth *et al.* 2002)

3. A minimalist metabolic network

Alternative pathways
Incomplete pathways
Metabolic coherence

Metabolic map



Core of a minimal bacterial gene set

DNA metabolism		16
Basic replication machinery	13	
DNA repair, restriction and modification	3	
RNA metabolism		106
Basic transcription machinery	8	
Translation: aminoacyl-tRNA synthesis	21	
Translation: tRNA maturation and modification	6	
Translation: ribosomal proteins	50	
Translation: ribosome function, maturation and modification	7	
Translation factors	12	
RNA degradation	2	
Protein processing, folding and secretion		16
Protein posttranslational modification	2	
Protein folding	5	
Protein translocation and secretion	5	
Protein turnover	3	
Cellular processes		6
Energetic and intermediary metabolism		56
Poorly characterized		8

Core of a minimal bacterial gene set

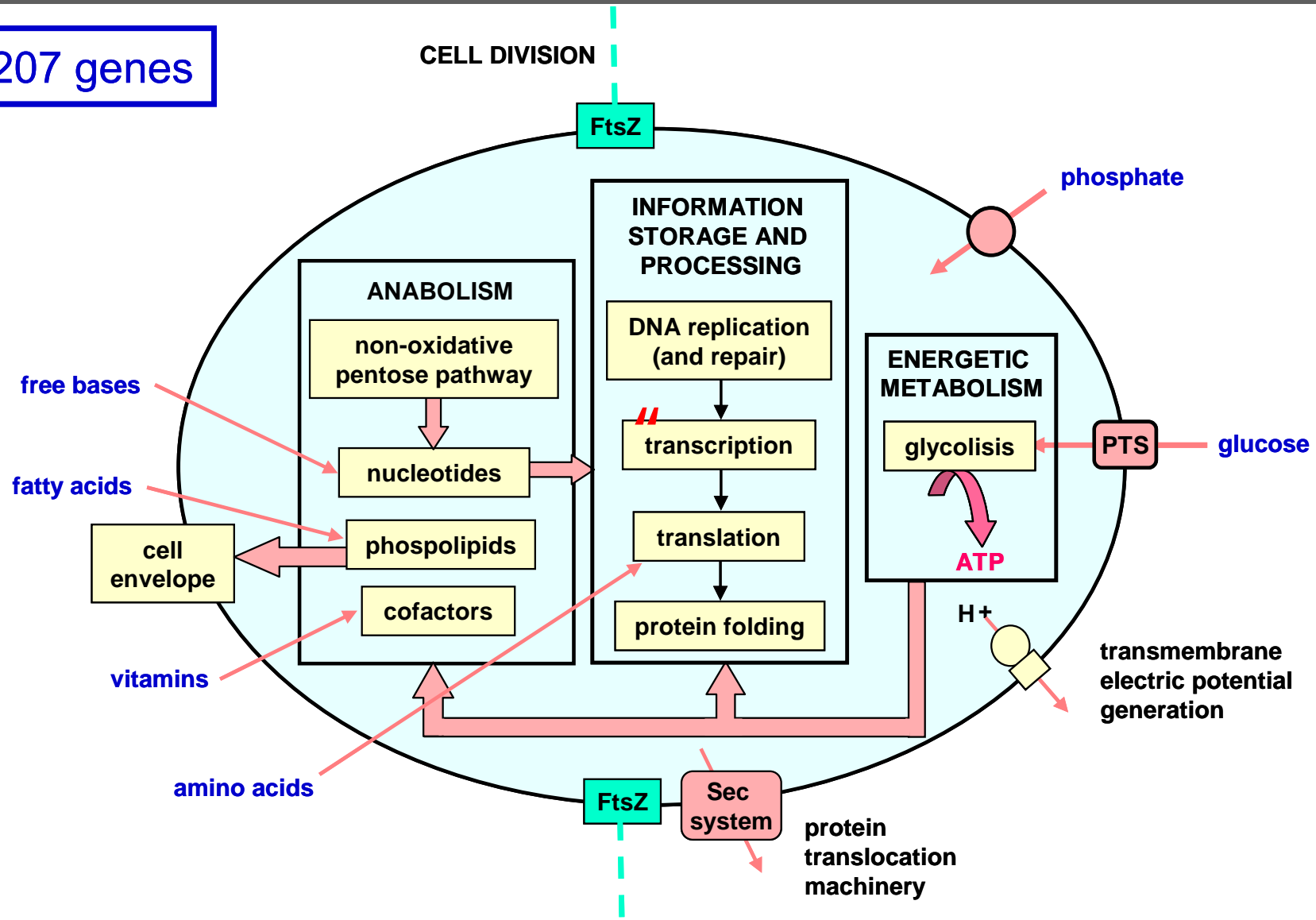
DNA metabolism		16
RNA metabolism		106
Protein processing, folding and secretion		16
Cellular processes		6
Cell division	1	
Transport	4	
Energetic and intermediary metabolism		56
Glycolysis	10	
Proton motive force generation	9	
Pentose phosphate pathway	3	
Lipid metabolism	7	
Biosynthesis of nucleotides	15	
Biosynthesis of cofactors	12	
Poorly characterized		8
Total		207

BASIC INFORMATIONAL MACHINERY	M. genome	<i>T. princeps</i>	PCSE
1. INFORMATION, STORAGE, PROCESSING	122	52	122
1.1. DNA metabolism	16	4	16
1.1.1. Basic replication machinery	13	4	13
1.1.2. DNA repair, restriction, modification	3	0	3
1.2. RNA metabolism	106	48	106
1.2.1. Basic transcription machinery	8	4	8
1.2.2. Translation	96	44	96
1.2.2.1. Aminoacyl-tRNA synthesis	21	0	21
1.2.2.2. tRNA maturation and modification	6	0	6
1.2.2.3. Ribosomal proteins	50	39	50
1.2.2.4. Ribosome function, maturation, modification	7	2	7
1.2.2.5. Translation factors	12	4	12
1.2.3. RNA degradation	2	0	2
2. PROTEIN PROCESSING, FOLDING, SECRETION	15	6	15
2.1. Protein post-translational modification	2	0	2
2.2. Protein folding	5	5	5
2.3. Protein translocation and secretion	5	0	5
2.4. Protein turnover	3	1	3



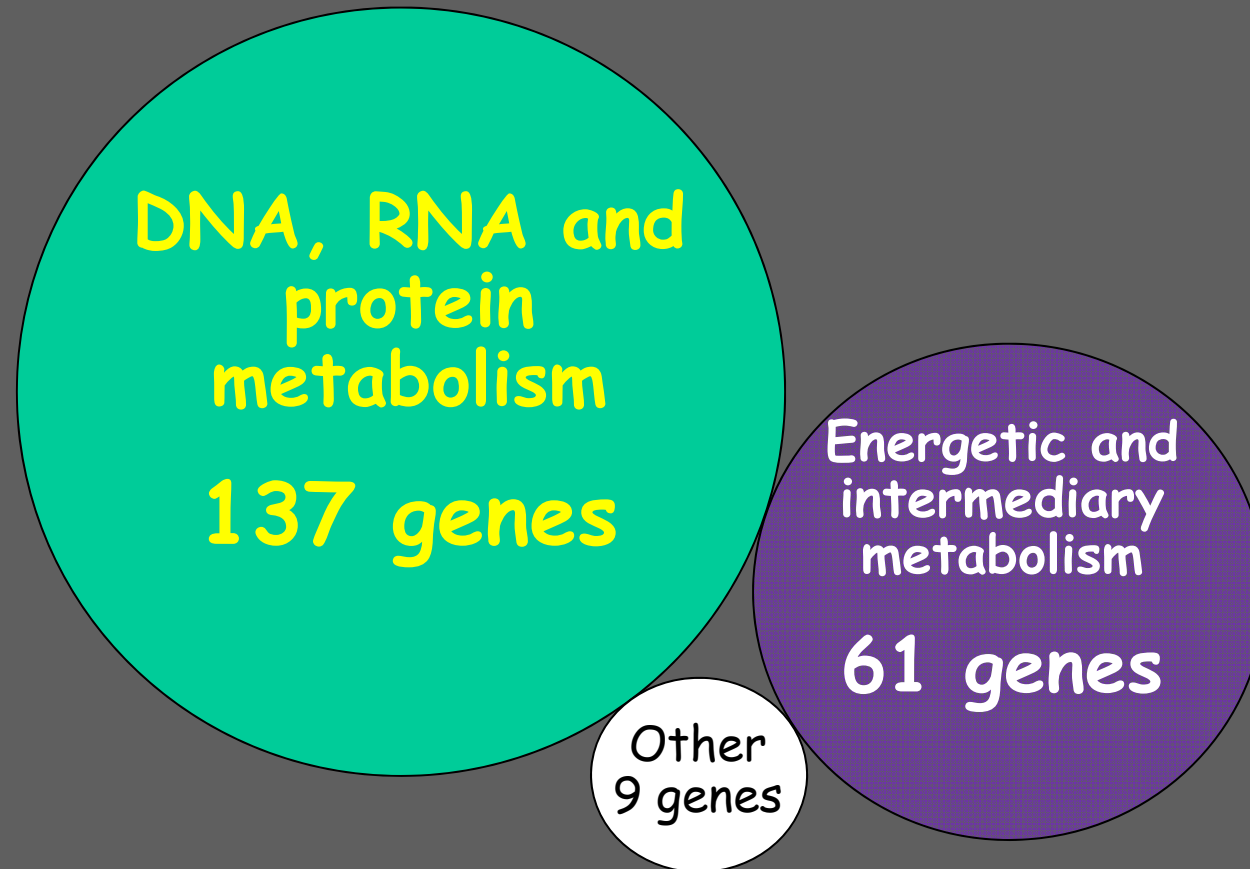
Our minimal genome for a hypothetical heterotrophic cell

207 genes



Gil et al. 2004. MMBR

Our minimal genome for a hypothetical heterotrophic cell



Structural analysis of minimal metabolisms



CENTRO SUPERIOR DE
INVESTIGACIÓN EN
SALUD PÚBLICA

VNIVERSITAT Đ VALÈNCIA

Institut Cavanilles de Biodiversitat i Biologia Evolutiva

Topological parameters of the inferred metabolic networks from the minimal gene set and natural genomes ordered from high to small size

Species	p-c genes	n	L	D	C	Cr
<i>Bradyrhizobium japonicum</i>	8317	1282	10.20	35	0.044	0.00150
<i>Streptomyces coelicolor</i>	8154	1119	10.10	29	0.064	0.00174
<i>Methylobacterium loti</i>	7272	1209	9.71	33	0.055	0.00165
<i>Anabaena sp.</i>	6131	970	9.76	29	0.041	0.00192
<i>Nocardia farcinica</i>	5936	1089	9.79	30	0.047	0.00174
<i>Agrobacterium tumefaciens (w)</i>	5402	1147	9.45	33	0.056	0.00171
<i>Escherichia coli (CFT073)</i>	5379	1120	10.20	34	0.075	0.00201
<i>Escherichia coli (K-12)</i>	4237	1215	10.30	35	0.067	0.00570
<i>Mycobacterium tuberculosis</i>	3991	1139	9.98	31	0.051	0.00167
<i>Clostridium acetobutylicum</i>	3848	784	9.56	25	0.061	0.00246
<i>Synechocystis sp.</i>	3264	918	10.50	30	0.044	0.00192
<i>Brucella melitensis</i>	3198	1197	8.54	31	0.049	0.00161
<i>Lactobacillus plantarum</i>	3059	864	9.64	26	0.067	0.00220
<i>Haemophilus influenzae (d)</i>	1657	775	10.00	30	0.065	0.00250
<i>Prochlorococcus marinus</i>	1760	844	10.50	30	0.045	0.00210
<i>Wolbachia (Bma)</i>	1195	516	8.76	28	0.075	0.00321
<i>Rickettsia prowazekii</i>	886	517	8.41	24	0.042	0.00299
<i>Tropheryma whipplei</i>	839	426	11.60	43	0.027	0.00475
<i>Wigglesworthia brevialpilis</i>	617	561	11.40	35	0.035	0.00308
<i>Blochmannia floridanus</i>	583	634	8.47	26	0.046	0.00273
<i>Buchnera aphidicola</i>	504	443	7.76	25	0.042	0.00395
<i>Mycoplasma genitalium</i>	484	207	7.49	23	0.043	0.00826
Minimal gene set	208	165	5.34	18	0.031	0.00977

p -c, number of protein-coding genes

n , number of nodes

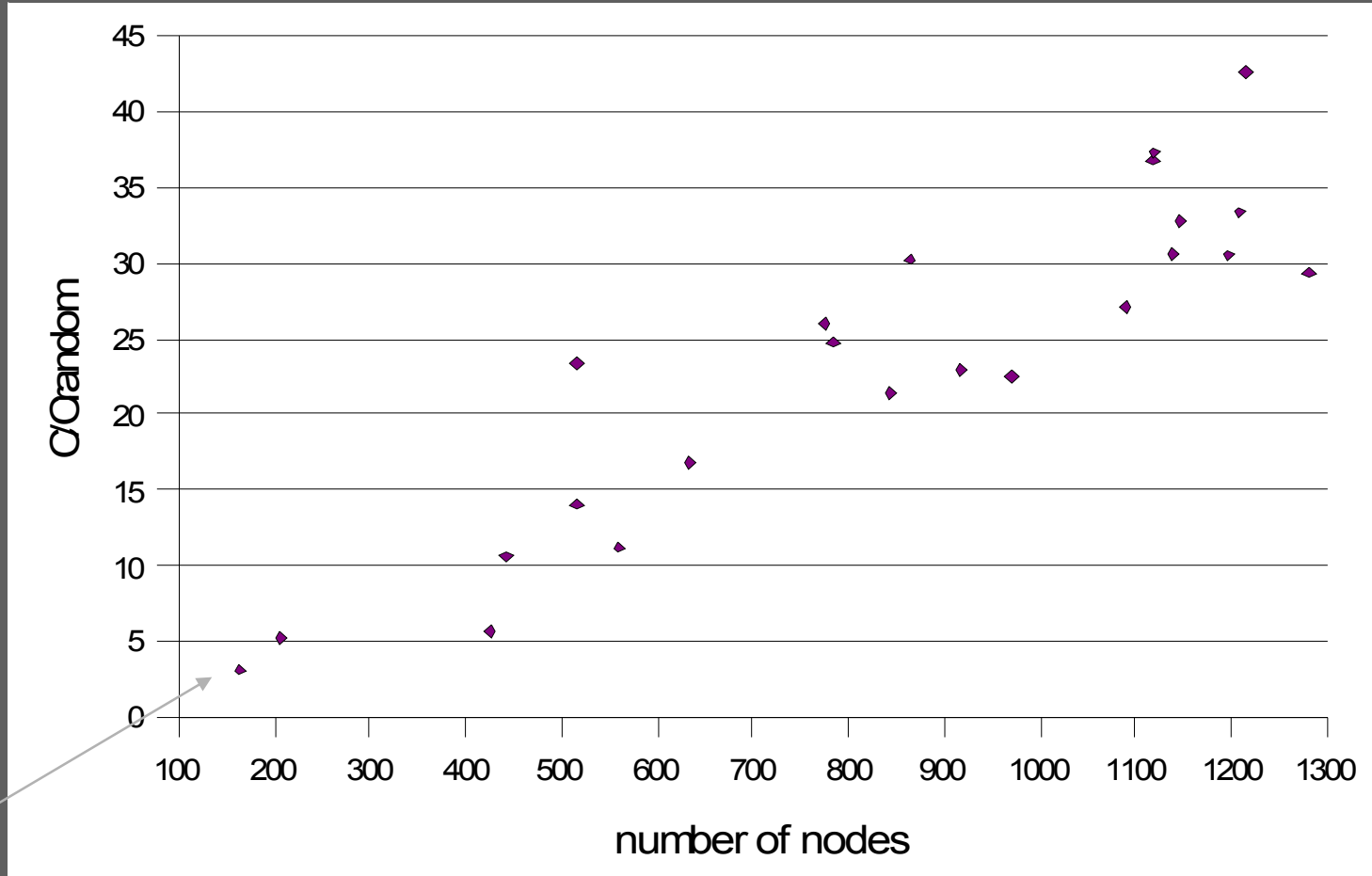
L , average path length;

D , network diameter

C , clustering coefficient;

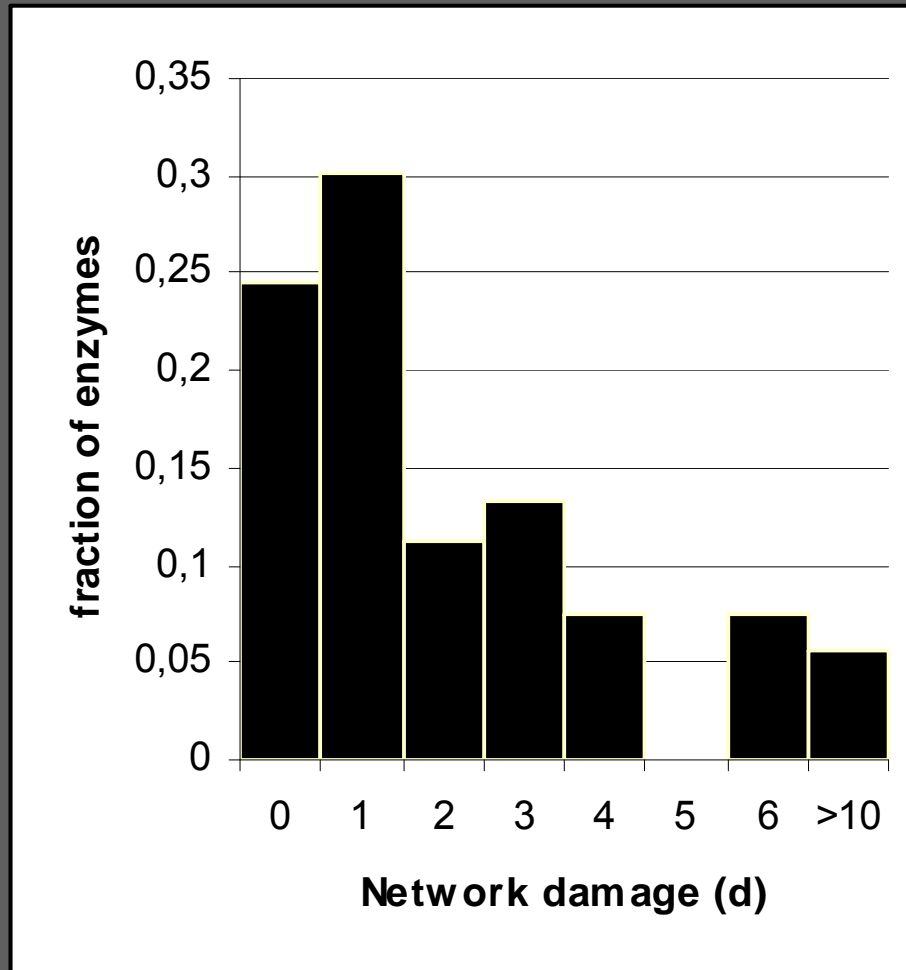
Cr , clustering coefficient for random network.

Effect of network size on the deviation of the clustering coefficient from expected random scenario



Minimal gene set

Network damage analysis for the minimal metabolic network



d , number of metabolites whose synthesis is prevented by a given mutation

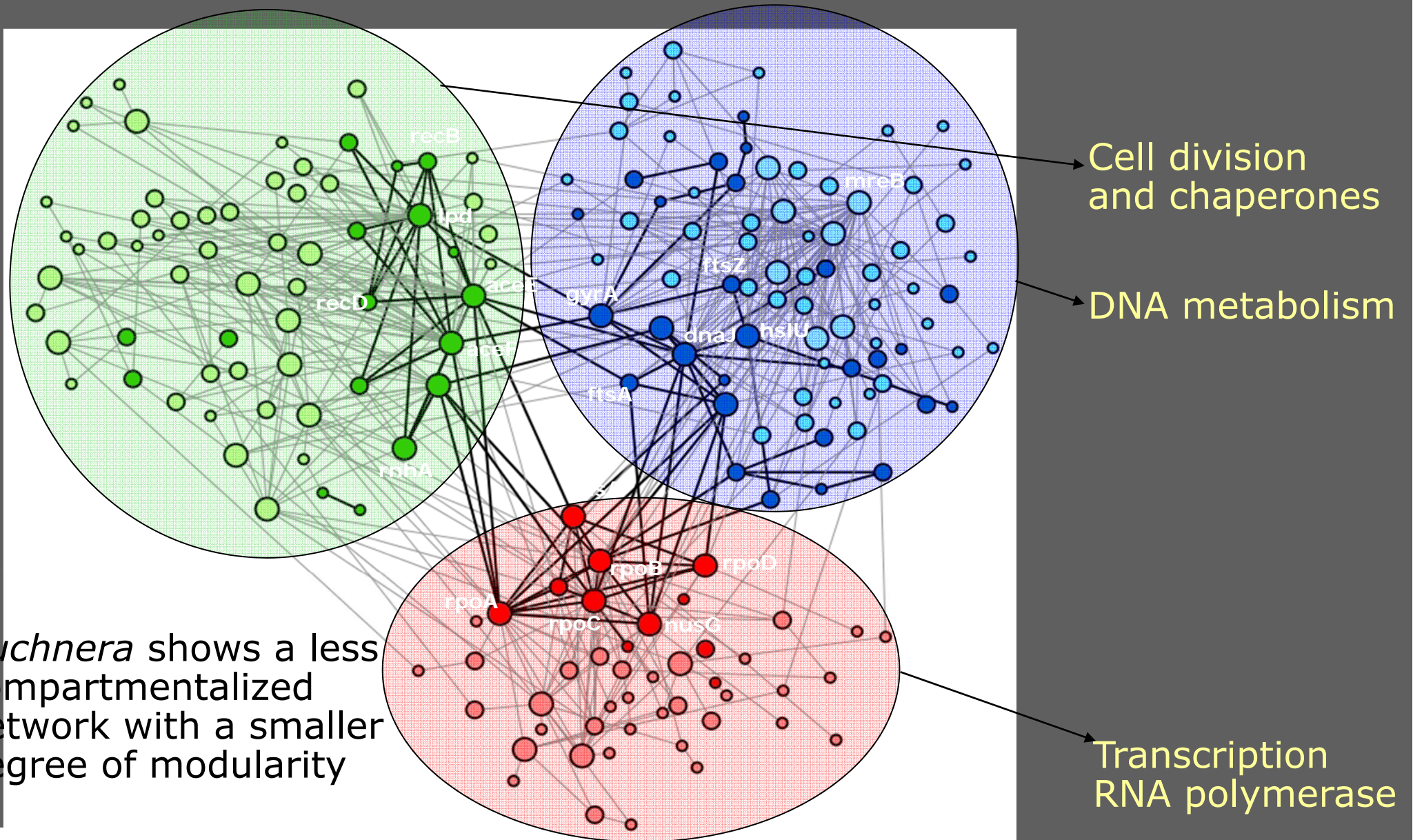
Most mutations (76%) in metabolic enzymes encoded in the minimal genome prevent the synthesis of at least one compound.

In the *E. coli* metabolic network the vast majority of the mutations produced no network damage

Modular organization in the reductive evolution of protein-protein interaction networks

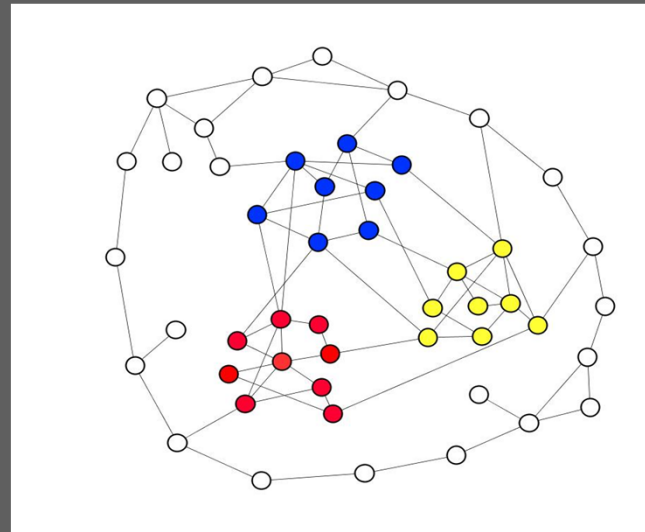


View of the three major modules of *E. coli* and *Buchnera*

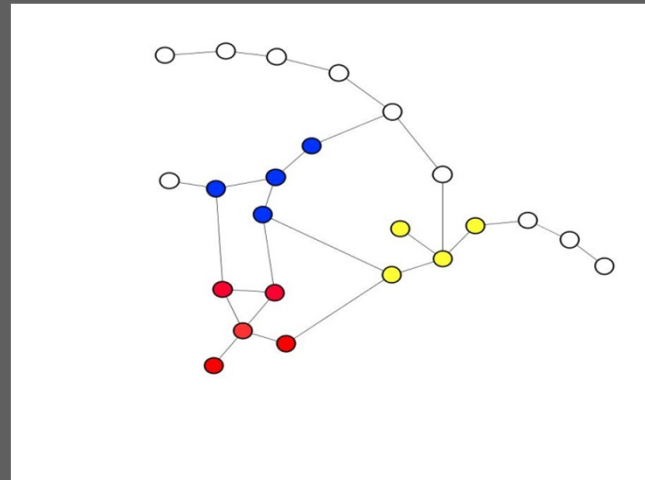


Buchnera shows a less compartmentalized network with a smaller degree of modularity

Deletion of interactions produce reduced modularity



↓
Reductive
evolution



Shared thoughts with this audience

- Under different views and definitions, the history of Biology is a cocktail of analytic and synthetic approaches to living beings.
- Computing probably represents the best language to formalize biological processes.
- Gödel's theorems tell us that there might appear biological features that cannot be fully explained within a particular system.
- Then, we must paid attention to those Synthetic Biology approaches stating that 'every thing is under control'.
- An alternative and positive view to the previous conclusion is that the best Synthetic Biology should be one based on the continuous increase of empirical and computational biological knowledge.
- Synthetic Biology can take advantage from natural minimal cells.

Evolutionary Genetics Group



Genomics and Health



Funding Agencies

Europe

Marie Curie ITN "Symbiomics"
COST on "Arthropod Symbiosis"

Spain

Ministerio de Ciencia e Innovación

Valencian Government

Prometeo Program



FP7/2007-2013, *grant agreement* # 212894



CENTRO SUPERIOR DE
INVESTIGACIÓN EN
SALUD PÚBLICA

VNIVERSITAT ID VALÈNCIA

Institut Cavanilles de Biodiversitat i Biologia Evolutiva