



**HAL**  
open science

## Absence of a core metabolic network common to symbiotic bacteria

Cecilia Coimbra Klein, Ludovic Cottret, Hubert Charles, Christian Gautier, Ana Tereza Ribeiro de Vasconcelos, Vincent Lacroix, Marie-France Sagot

► **To cite this version:**

Cecilia Coimbra Klein, Ludovic Cottret, Hubert Charles, Christian Gautier, Ana Tereza Ribeiro de Vasconcelos, et al.. Absence of a core metabolic network common to symbiotic bacteria. 6. Meeting GDRE-RA Comparative genomics, Nov 2011, Lyon, France. 33 diapositives. hal-02806316

**HAL Id: hal-02806316**

**<https://hal.inrae.fr/hal-02806316>**

Submitted on 6 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Absence of a core metabolic network common to symbiotic bacteria

**Cecilia C. Klein**<sup>1,2,5</sup>, Ludovic Cottret<sup>3</sup>, Hubert Charles<sup>1,4</sup>,  
Christian Gautier<sup>1,2</sup>, Ana Tereza Ribeiro de Vasconcelos<sup>2,4,5</sup>,  
Vincent Lacroix<sup>1,2</sup>, Marie-France Sagot<sup>1,2</sup>

<sup>1</sup> BAMBOO Team, INRIA Grenoble-Rhône-Alpes, France

<sup>2</sup> CNRS, UMR5558, Laboratoire de Biométrie et Biologie Évolutive, UCBL, France

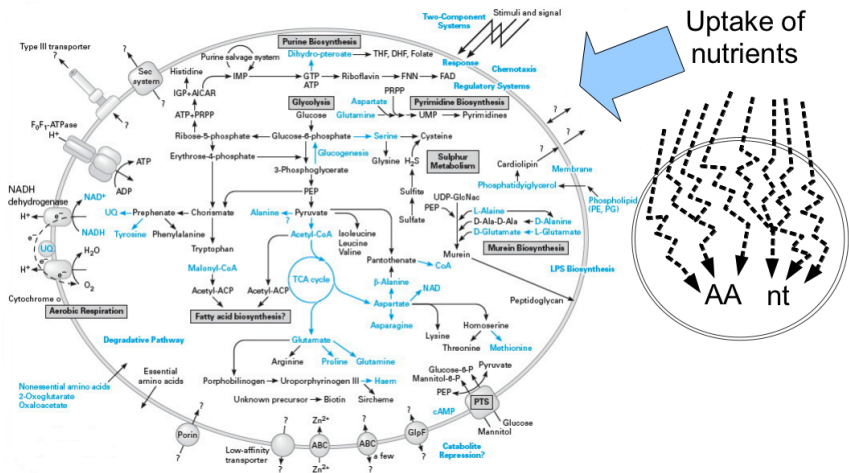
<sup>3</sup> UMR1089 Xénobiotiques INRA-ENVT, Toulouse, France

<sup>4</sup> UMR203 Biologie Fonctionnelle Insectes et Interactions, INRA, INSA-Lyon, France

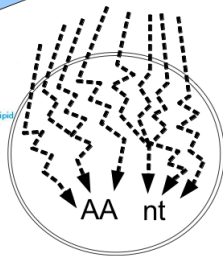
<sup>5</sup> Laboratório Nacional de Computação Científica (LNCC), Brazil

GDRE Comparative Genomics  
November 2011

# Metabolism - basic building blocks

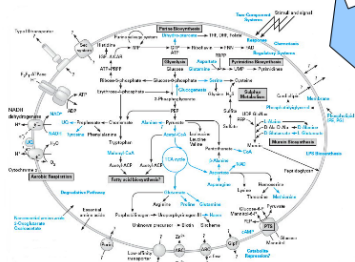


Uptake of nutrients



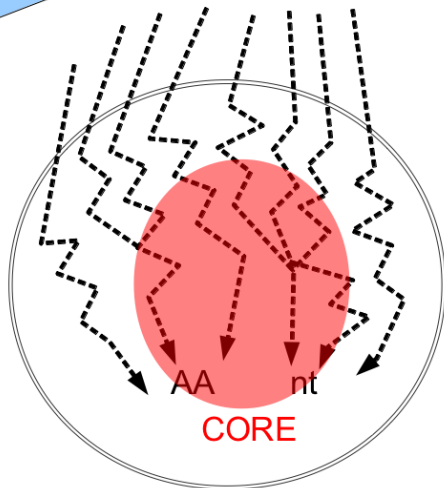
Modified from <http://www.foodmate.net/4images/3/5/21525.html>

# A core of metabolic reactions ??

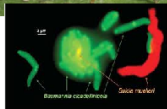
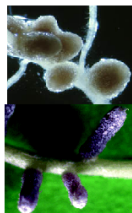
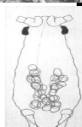
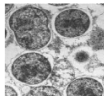


Small Molecule  
Metabolism  
of bacteria

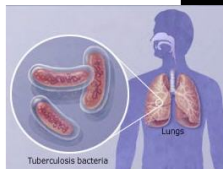
Uptake of  
nutrients



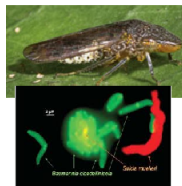
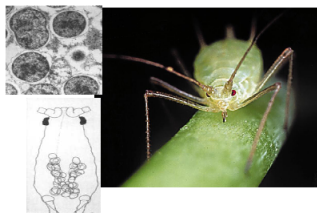
# A core of metabolic reactions ??



Lifestyle



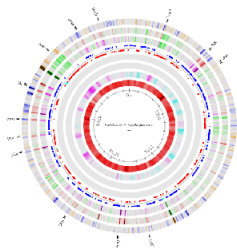
# A core of metabolic reactions ??



Environment is  
the host cell

Type of interaction  
Reproductive manipulators  
Nutritional mutualism  
Host defense

...



Genome reduction  
Reduced metabolism

# How to do it?

- By comparing several metabolic networks of bacteria with different lifestyles
- By comparing the reaction and compound content

# Dataset

## Intracellular

### Intracellular mutualists vertically transmitted (MIV)

*Candidatus Hodgkinia cicadicola* Dsem (HODCCD)  
*Candidatus Carsonella ruddii* PV (CARRP)  
*Candidatus Sulcia muelleri* GWSS (SULMW)  
*Buchnera aphidicola* Cc (BUCCC)  
*Buchnera aphidicola* APS (BUCAI)  
*Blattabacterium* sp., Bge (BLASB)  
*Baumannia cicadellinicola* str. Hc (BAUCH)  
*Candidatus Blochmannia floridanus* (BLOFL)  
*Wigglesworthia glossinidia* (WIGBR)  
*Candidatus Blochmannia pennsylvanicus* (BLOPB)  
*Wolbachia pipentis* wBm (WOLTR)

### Intracellular parasites vertically transmitted (PIV)

*Rickettsia typhi* Wilmington (RICTY)  
*Wolbachia pipentis* wMeI (WOLPM)  
*Orientia tsutsugamushi* Boryong (ORITB)

### Intracellular parasites horizontally transmitted (PIH)

*Chlamydia trachomatis* D/UW-3/CX (CHLTR)  
*Lawsonia intracellularis* PHE/MN1-00 (LAWIP)

## Cell associated (CA)

### Cell associated mutualists vertically transmitted (MCAV)

*Candidatus Hamiltonella defensa* T5A (HAMD5)  
*Sodalis glossinidius morsitans* (SODGM)

### Cell associated parasites horizontally transmitted (PCAH)

*Mycoplasma genitalium* G37 (MYCGE)  
*Mycoplasma hyopneumoniae* J (MYCHJ)  
*Bartonella quintana* Toulouse (BARQU)  
*Helicobacter pylori* 26695 (HELPHY)  
*Neisseria gonorrhoeae* NCCP11945 (NEIG2)  
*Listeria monocytogenes* EGD-e (LISMO)  
*Brucella melitensis* bv 1 16M (BRUME)  
*Yersinia pestis* CO92 (YERPE)  
*Yersinia pseudotuberculosis* YPIII (YERPY)  
*Mycobacterium tuberculosis* H37Rv (MYCTU)  
*Salmonella enterica* serovar Typhi (SALTY)  
*Shigella flexneri* 2a str. 301 (SHIFL)  
*Photobacterium luminescens* TTO1 (PHOLL)  
*Burkholderia mallei* ATCC 23344 (BURMA)  
*Bacillus anthracis* Ames Ancestor (BACAN)

## Extracellular

### Extracellular mutualists horizontally transmitted (MEH)

*Cupriavidus taiwanensis* LMG19424 (CUPTR)  
*Frankia alni* ACN14a (FRAAA)  
*Sinorhizobium meliloti* 1021 (RHIME)

### Extracellular commensals horizontally transmitted (CEH)

*Streptococcus thermophilus* LMD-9 (STRTD)  
*Wolfinella succinogenes* DSM 1740 (WOLSU)  
*Lactobacillus casei* ATCC 334 (LACC3)  
*Bacillus amyloliquefaciens* FZB42 (BACA2)  
*Escherichia coli* K-12 (ECOLI)  
*Pseudomonas fluorescens* SBW25 (PSEFS)  
*Mycobacterium smegmatis* MC2 155 (MYCS2)

### Extracellular parasites horizontally transmitted (PEH)

*Streptococcus agalactiae* 2803V/R (STRA5)  
*Xylella fastidiosa* 9a5c (XYLFA)  
*Vibrio cholerae* O1 biovar El Tor str. N16961 (VIBCH)  
*Erwinia carotovora* subsp. atroseptica SCRI1043 (ERWCT)  
*Pseudomonas entomophila* L48 (PSEE4)  
*Escherichia coli* O157:H7 (ECO57)  
*Bacillus thuringiensis* serovar konkukian str. 97-27 (BACHK)  
*Agrobacterium tumefaciens* C58 (AGRT5)  
*Pseudomonas aeruginosa* UCBPP-PA14 (PSEAB)

## Free-living (FL)

*Thiomicrospira crunogena* XCL-2 (THICR)  
*Desulfotalea psychrophila* Lsv54 (DESPS)  
*Pseudoalteromonas haloplanktis* TAC125 (PSEHT)  
*Bacillus subtilis* 168 (BACSU)  
*Rhodobacter sphaeroides* 2.4.1 (RHOSA)  
*Ralstonia eutropha* H16 (RALEH)



# Data acquisition & methods

- MicroCyc/MicroScope<sup>1</sup>:
  - ▶ genomic data and metabolic network reconstructions
  - ▶ to avoid as much as possible introducing biases related to different annotation techniques
  
- HAMAP classification<sup>2</sup>:
  - ▶ guide the grouping of the organisms

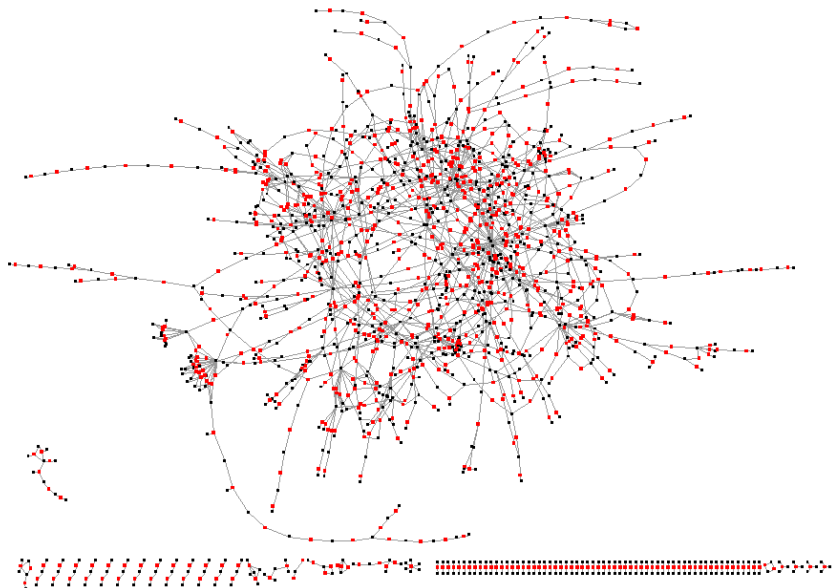
<sup>1</sup>Vallenet *et al.* 2009 Database (Oxford).

<sup>2</sup>Lima *et al.* 2009 Nucleic Acids Res.

# How to do it?

- Comparison of several metabolic networks of bacteria with different lifestyles
- By comparing the reaction and compound content

# Metabolic network



# Data acquisition & methods

- MetExplore: network modelling and filtering<sup>1</sup>
  - ▶ Small molecule metabolism filtering
- Cytoscape: visualisation of the networks<sup>2</sup>
- Analyses were performed using R<sup>3</sup>:
  - ▶ *ade4* package for statistics<sup>4</sup>
  - ▶ *Igraph* package for graph manipulation<sup>5</sup>

<sup>1</sup>Cottret *et al.* 2010 Nucleic Acids Res.

<sup>2</sup>Shannon *et al.* 2003 Genome Res.

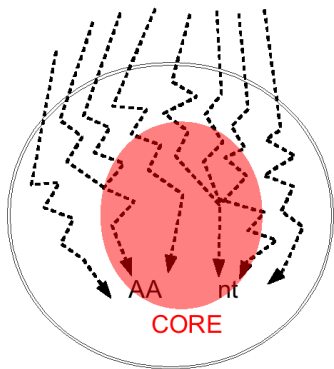
<sup>3</sup><http://cran.r-project.org/>

<sup>4</sup>Dray & Dufour *et al.* 2007 Journal of Statistical Software

<sup>5</sup>Csardi *et al.* 2006 InterJournal Complex Systems

# Metabolic core

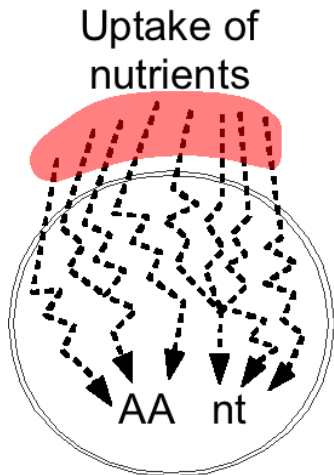
- Metabolic core: a conserved set of elements in these bacteria
- It was obtained by computing the intersection of the sets of reactions (resp. compounds) for each species



# Environment: inputs to the network

Identification of potential inputs to the networks using the Borenstein method<sup>1</sup>

- ▶ metabolites each bacterium potentially acquires from its environment

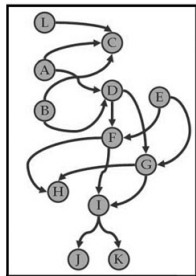


<sup>1</sup>Borenstein *et al.* 2008 Proc Natl Acad Sci U S A

# Environment: inputs to the network

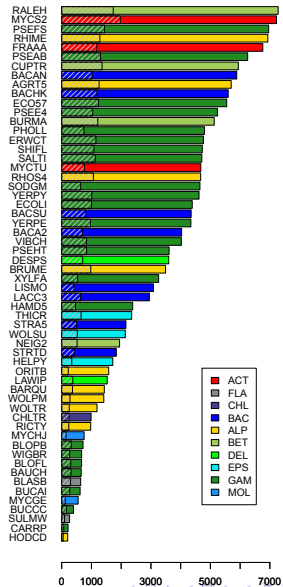
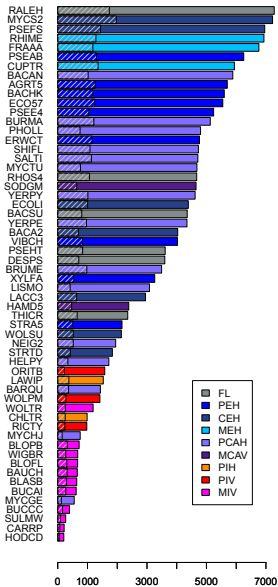
Identification of potential inputs to the networks using the Borenstein method<sup>1</sup>

- ▶ metabolites each bacterium potentially acquires from its environment
- ▶ directed compound graph
- ▶ cofactors and current compounds filtered
- ▶ based on the detection of the strongly connected components (SCC) in the compound graph
- ▶ a version of the Borenstein method using the Igraph package in R



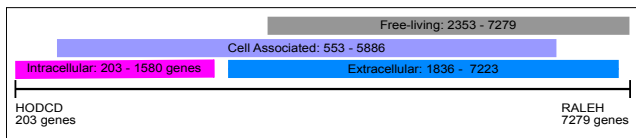
<sup>1</sup>Borenstein *et al.* 2008 Proc Natl Acad Sci U S A

# Data overview: number of genes



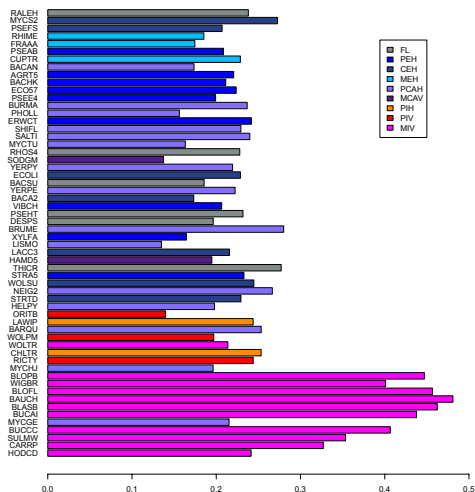


# Data overview: number of genes

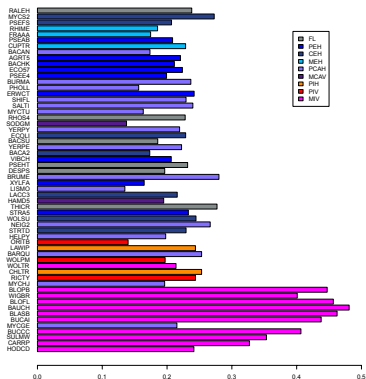


The intracellular life is paired with an extreme reduction of the genome and of the metabolism

# Ratio metabolic genes over the total number of genes

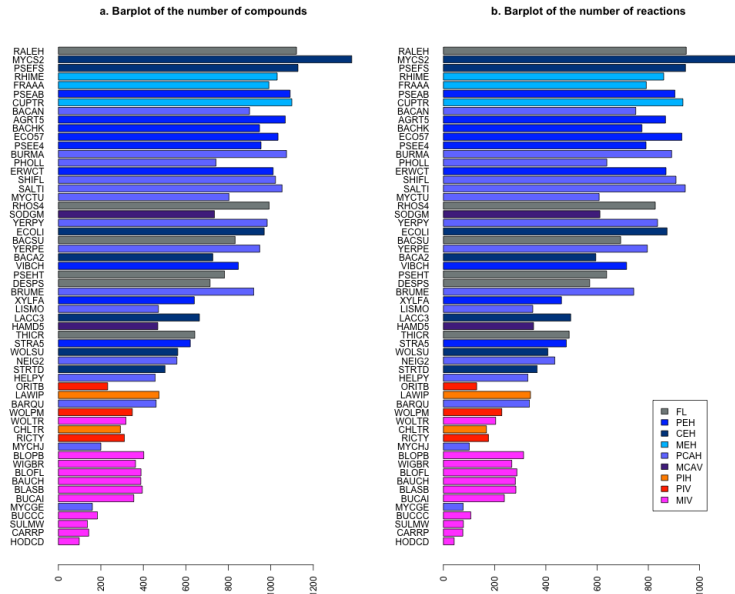


# Ratio metabolic genes over the total number of genes

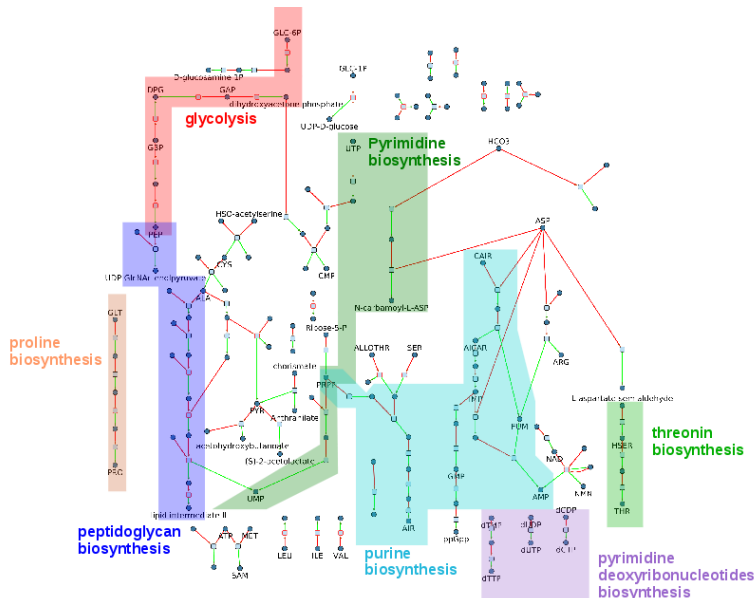


The portion dedicated to metabolism is greater in the genomes of the mutualistic endosymbionts

# Number of compounds and reactions

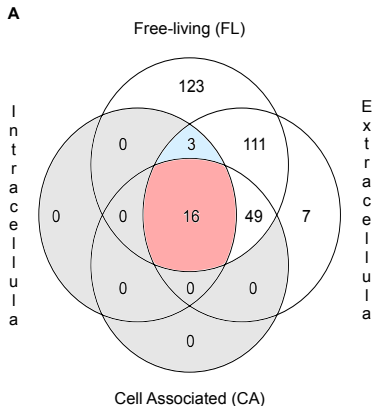


# Core metabolism - extracellular bacteria



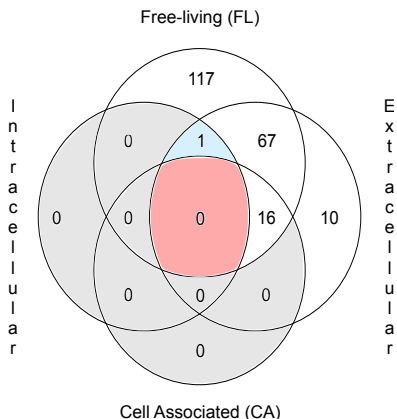


# Core metabolism - compounds



Compounds	Classification
pyruvate	Acids
L-alanine	Amino Acids
L-cysteine	Amino Acids
L-glutamate	Amino Acids
L-glutamine	Amino Acids
L-serine	Amino Acids
D-ribose-5-phosphate	Carbohydrates
coenzyme A	Cofactors
diphosphate	Ions
phosphate	Ions
H <sup>+</sup> (PROTON)	Ions
AMP	Nucleosides
ADP	Nucleosides
ATP	Nucleosides
formate	Unclassified
H <sub>2</sub> O (WATER)	Unclassified

# Core metabolism - reactions



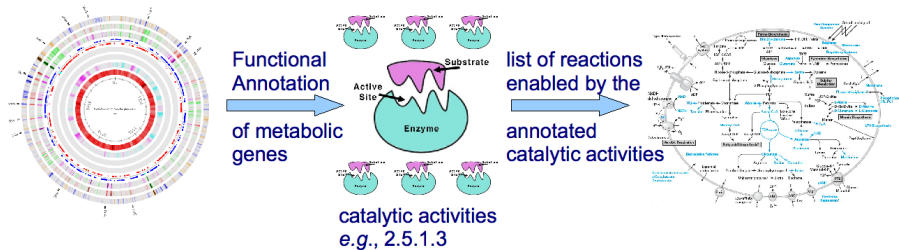
- ▶ There is not a single enzymatic reaction common to all 58 bacteria
- ▶ No metabolic core is preserved among the endosymbionts
- ▶ The reduction of the metabolism affects different parts of the network among the endosymbionts



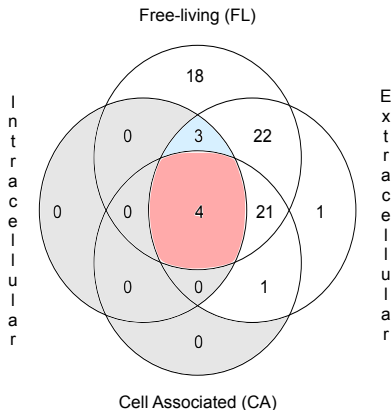
# EC number sets

- The analysis of partial EC number sets, e.g. 2.5.1.-, may allow some flexibility, possibly increasing the size of the common sets.

## Metabolic network reconstruction from genomic data



# EC number sets



EC number	Classification
<b>2</b>	<b>Transferases</b>
2.3	Acytransferases
2.3.1	Transferring groups other than aminoacyl groups
2.5	Transferring alkyl or aryl groups, other than methyl groups
2.5.1	Transferring alkyl or aryl groups, other than methyl groups (only subclass identified to date)
<b>3</b>	<b>Hydrolases</b>
3.5	Acting on carbon-nitrogen bonds, other than peptide bonds
3.5.1	In linear amides
<b>4</b>	<b>Lyases</b>
4.2	Carbon-oxygen lyases
4.2.1	Hydro-lyases

# EC number sets - simulation

The small intersection could be explained by differential random loss of enzymes during genome reduction of intracellular symbionts?



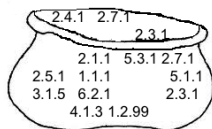
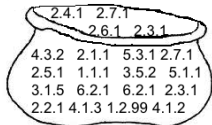
For each MIV  
 $\gamma$ -proteobacteria (7)



Randomly picked  
one extracellular  
 $\gamma$ -proteobacteria

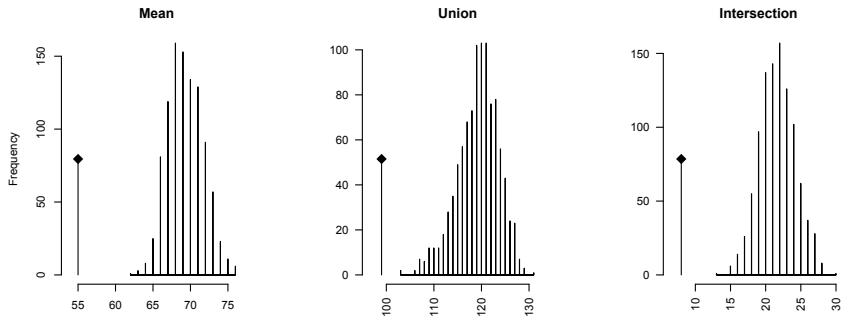
Randomly removed  
elements, until we reached  
the size of the MIV  
metabolic network

*Buchnera*: 246 partial EC numbers



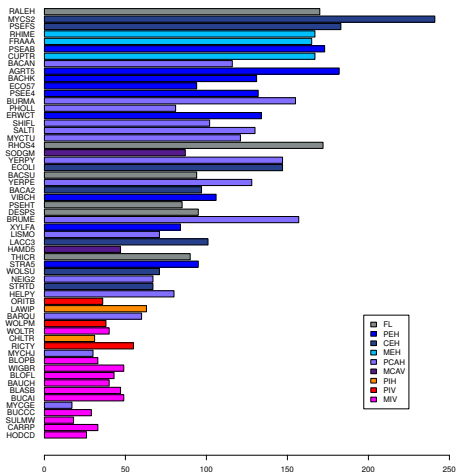
246 partial EC numbers

# EC number sets - simulation



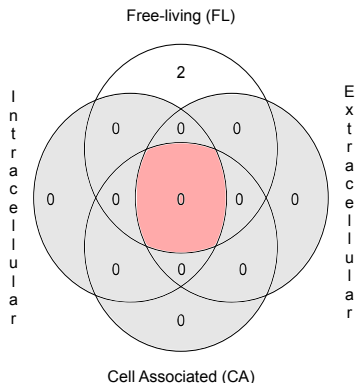
	Mean	Union	Intersection
MIV	55	99	8
Simulation	$69 \pm 4.7$	$119 \pm 8.5$	$22 \pm 5.0$
Estimated $p$ - value	$\leq 0.001$	$\leq 0.001$	$\leq 0.001$

# Environment - inputs in the network



Is the absence of a metabolic core in the obligate intracellular bacteria linked to the differences in their environment?

# Environment - inputs in the network



The absence of common potential inputs to the bacteria

It indicates that the diversity of the environment alone does not explain the loss of a metabolic core

# Conclusions

- Comparison of the small molecule metabolism of 58 bacteria representing a wide range of lifestyles
- No enzymatic reaction common to all of them
- It is essentially due to intracellular symbionts

# Conclusions

- Extracellular symbionts:
  - ▶ a core restricted to 94 reactions
  - ▶ this core is however non connected
  
  - ▶ it involves mainly biosynthetic reactions (88%)
  - ▶ Vieira *et al.* 2011<sup>1</sup> found that the core-metabolism of *Escherichia coli* is enriched in biosynthetic pathways
  
  - ▶ they have similar needs in terms of the building blocks of their cells
  - ▶ each one of them has to adapt to a distinct environment

<sup>1</sup>Vieira *et al.* 2011 J Bacteriol.



# Conclusions

- Obligate intracellular symbionts:
  - ▶ the host provides an interface between the external and the internal environments
  - ▶ most degradation pathways have disappeared and have been outsourced to the host
  - ▶ synthetic routes have not all disappeared but they have been selected for, depending on the nature of the symbiosis
  
- The fact that there is no reaction common to all symbionts implies that virtually any reaction can be outsourced to the host

# Thanks

- Funding:
  - ▶ French project ANR MIRI BLAN08-1335497
  - ▶ ERC Advanced Grant SISYPHE
  - ▶ FAPERJ/INRIA

