

Diversity Matters
Ramon Rosselló-Móra
Marine Microbiology Group
Institut Mediterrani d'Estudis Avançats (CSIC-UIB)



Acknowledgements:

Josefa Antón (Universidad de Alicante, Spain)
Rudolf Amann (Max Planck Institute MM, Germany)
Edward Moore (CCUG/Sweden)
Philippe Schmitt-Kopplin (Helmholz Zentrum, Germany)
Silvia Marqués (EEZ, Spain)
Karl-Heinz Schleifer (TU Munich, Germany)
Frank Oliver Glöckner (Max Planck Institute, Germany)
William B. Whitman (University of Georgia, USA)
Jean Euzéby (SBSV, France)
Javier Tamames de la Huerta (CNB, Spain)
Roberto Vidal (Universidad de Chile, Chile)
Bernardo González (Universidad Adolfo Ibañez, Chile)
Joan Rita (Universitat Illes Balears)
Daniel Ginard (Hospital Son Espases)
Sam Khorrami (Hospital Son Espases)

BOS-2000-1123-C02-01
BOS-2003-05198-C02-01
CGL2006-12714-C02-02
CE-CSD2007-0005

CGL2009-12651-C02-02
CGL2012-39627-C03-03
INFRA-2012-2.2.5/312251
MEBICROHN

At the MMG – IMEDEA:

Mercedes Urdiain (technical assistant)
Raúl Muñoz (Ph-D student)
Tomeu Viver (Ph-D student)
Merit Mora (Ph-D student)
Xisca Font (Master)
Sara Díaz (technical assistant)
Carlota Alejandre (Master)

Away of the IMEDEA:

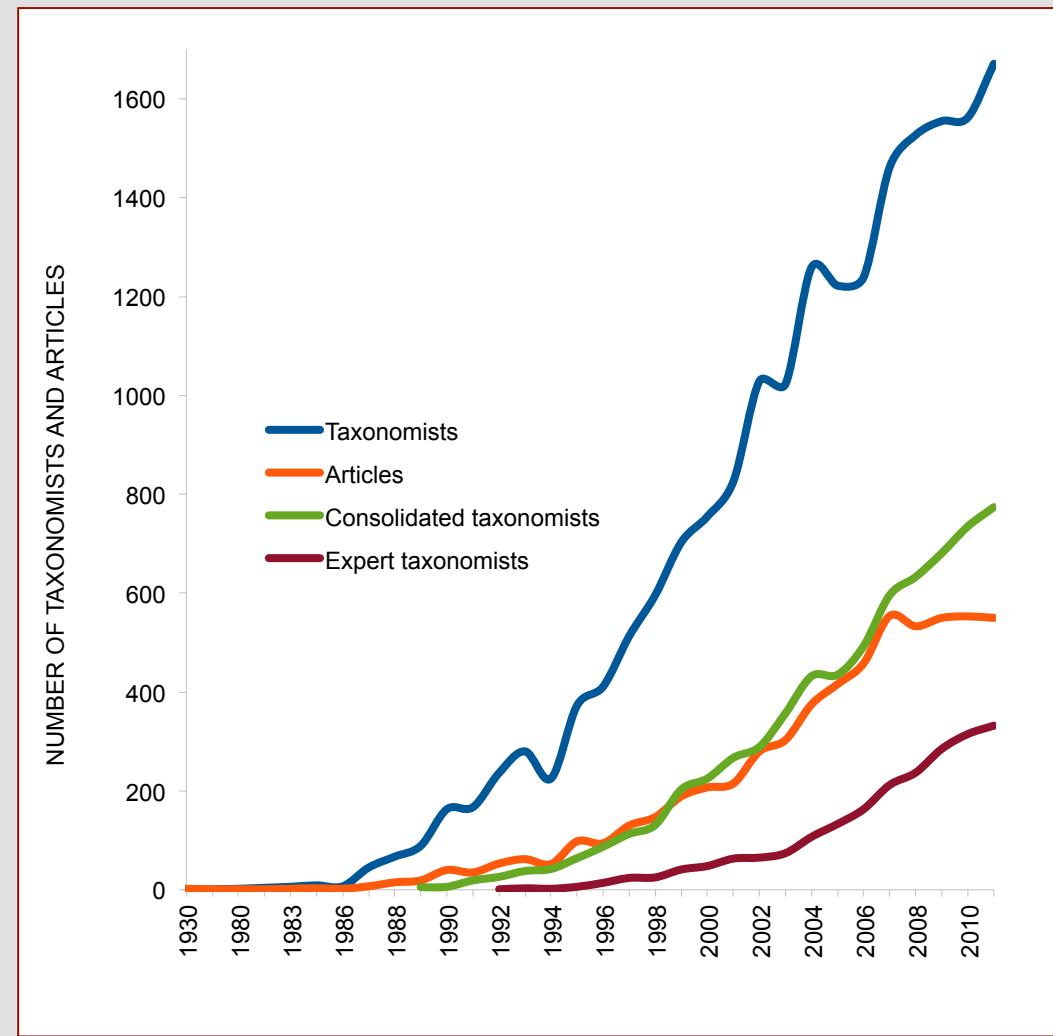
Arantxa López (postdoc)
Michael Richter (postdoc)
Ana Cifuentes (postdoc)
Pablo Yarza (Ph-D student)
Jocelyn Brito (Ph-D student)
Ana Belén Suárez (Ph-D student)
Sara Cortès (Master)

- Salines de Levante
- Salines de S'Avall
- Salines de Ibiza
- Salines de Formentera
- Salinas de Santa Pola
- Salinas de Fuerteventura
- Salinas de Lanzarote
- Salines de la Trinidad
- Salinas de Añana
- Aquàrium de Barcelona
- Marineland Catalunya
- Marineland Mallorca
- Parc Natural del Delta
- Cabildo de Gran Canaria
- Salinas de Lo Valdivia

⇒ www.imedea.uib.es/jspecies
⇒ www.arb-silva.de/living_tree
⇒ www.elsevier.de/syapm

Growth of expertise in taxonomy of prokaryotes

There are a lot of newcomers in the field, but that publishing >5 (green) or >10 descriptions are minority



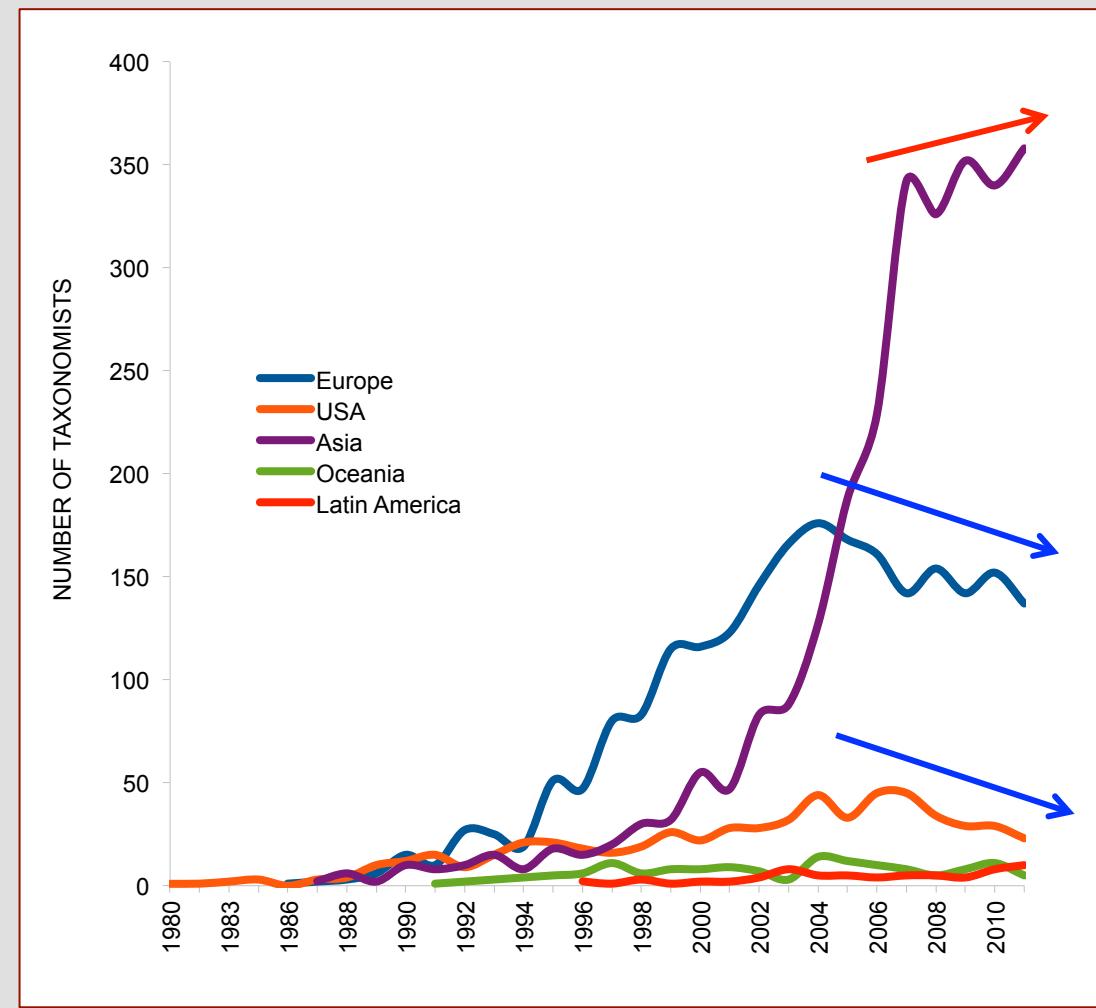
Tamames & Rosselló-Móra 2012 Trends in Microbiol 20:514-516

1- Loss of expertise in occidental countries

Loss of Know-How in Occidental Countries

Oriental taxonomists are at least doubling those of western countries

In Europe and USA it seems that a decline in the numbers of experts occurs



Tamames & Rosselló-Móra 2012 Trends in Microbiol 20:514-516

1- Loss of expertise in occidental countries



<http://www.mirri.org/>

Current Census of European Taxonomists

Bibliographic list:

583 taxonomists
>5 papers

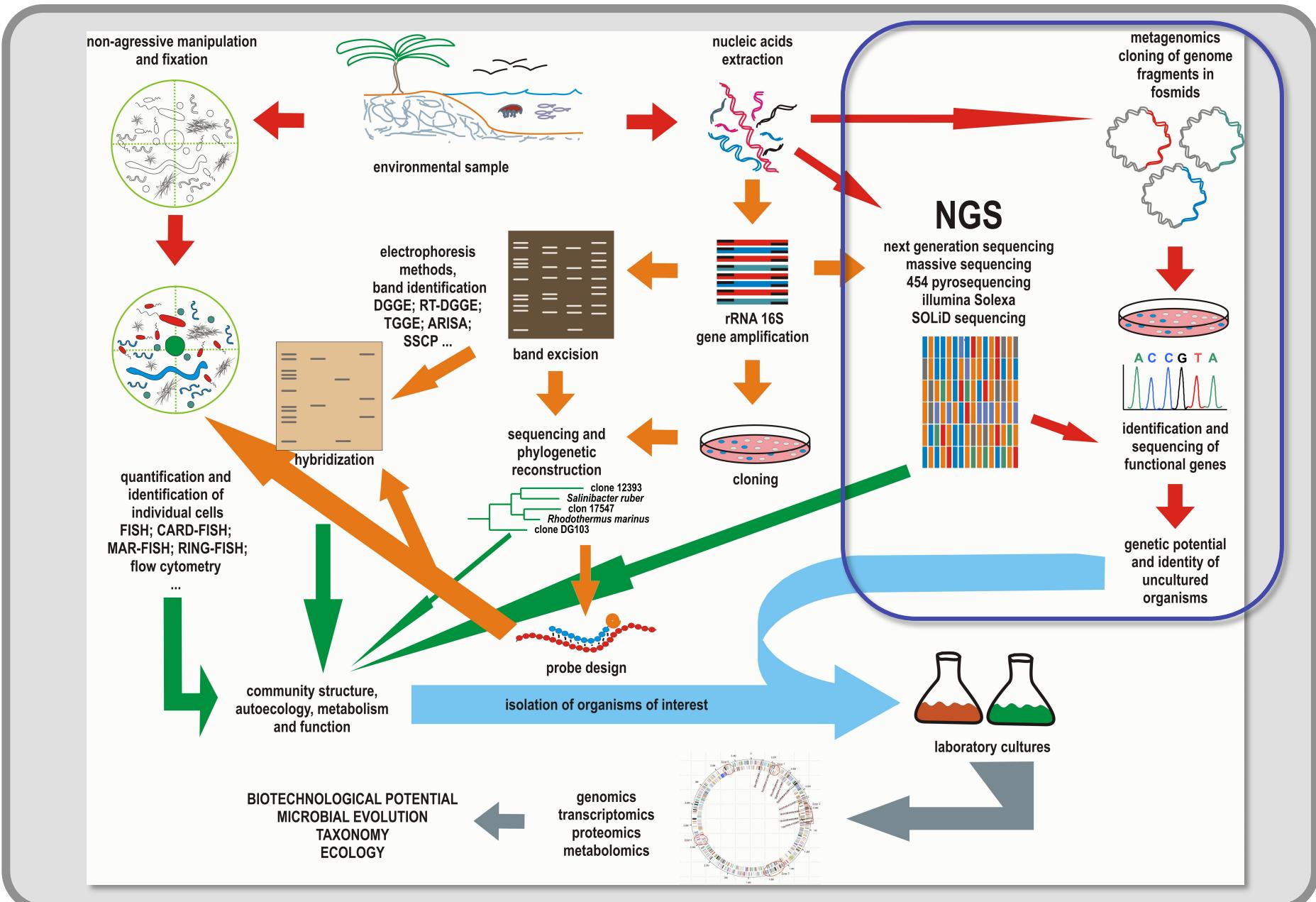
Email exchanges

65 retired/deceased
88 never answered
82 wrongly identified

Current census

293 taxonomists

we overestimated
the census in about
50%



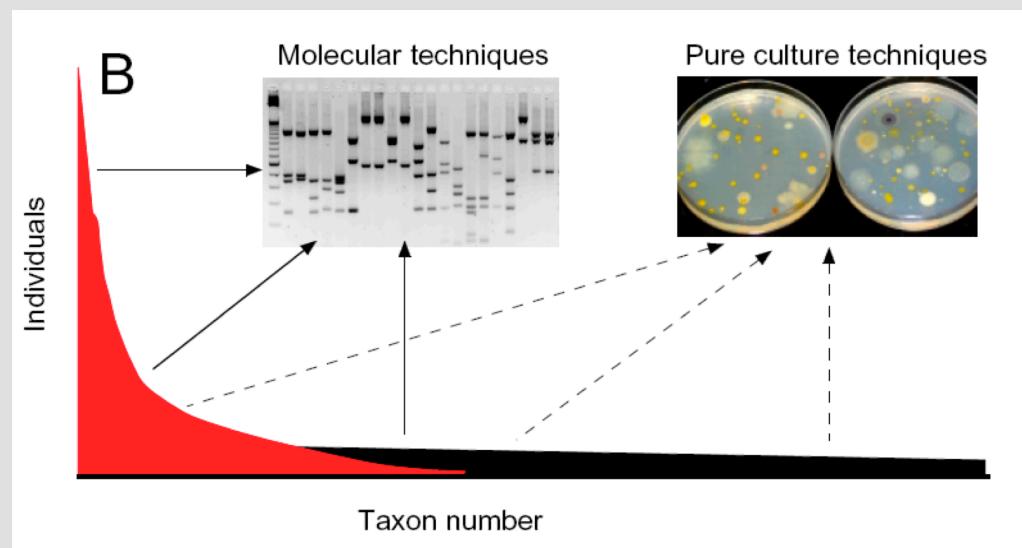
2- Diversity approaches

Species ↔ microbial molecular ecology

MOLECULAR TECHNIQUES

- ➔ generally identification of units (species) by means of 16S rRNA genes
- ➔ generally inform about the highly abundant organisms
- ➔ it is not clear where to set a threshold of what is a species

Red => knowable diversity / black => seed bank, unknown, difficult to know



Pedrós-Alió, 2006 Trends Microbiol 14:257-263

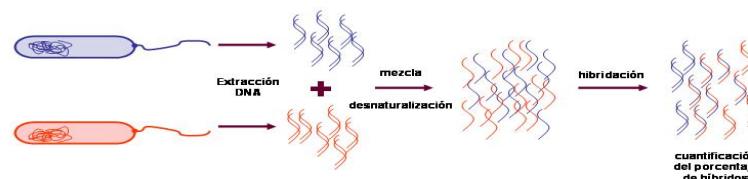
Culturing is important to understand uncultured

**For metagenomics, it is important to understand the meaning of
Species at the genomic level**

Species at the genomic level

DDH (DNA-DNA hybridization):

- 70% similarity (50-70%)



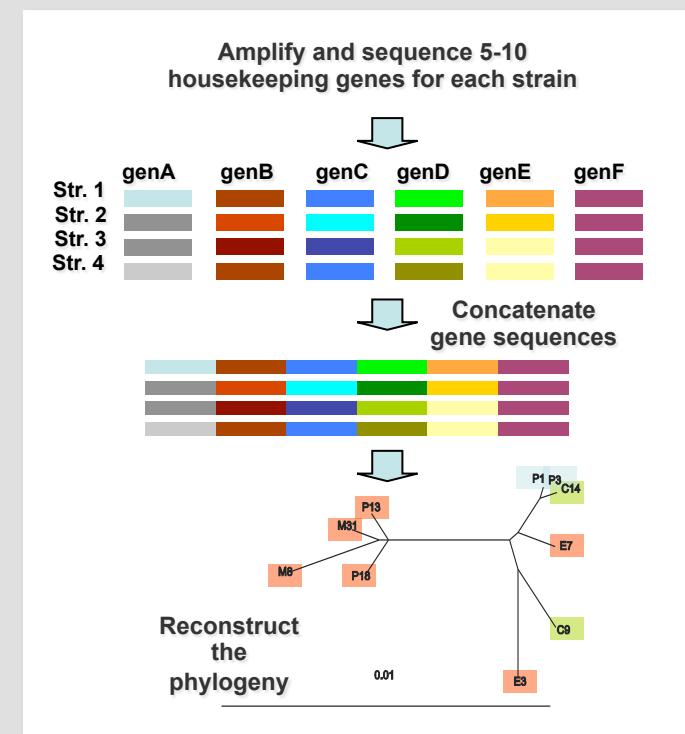
- used since the 60's ⇔ strong influence
- non cumulative DB

Rosselló-Mora & Amann 2001, FEMS Rev. 25:39-67

DDH has been the gold standard ⇔
the “sex” for higher eukaryotes

MLSA (multilocus sequence analysis):

- 5-10 full/partial sequences
- house keeping genes
 - primer design difficulties
 - biases in the selection of genes
 - time consuming
 - ↓↓ number for stable topology

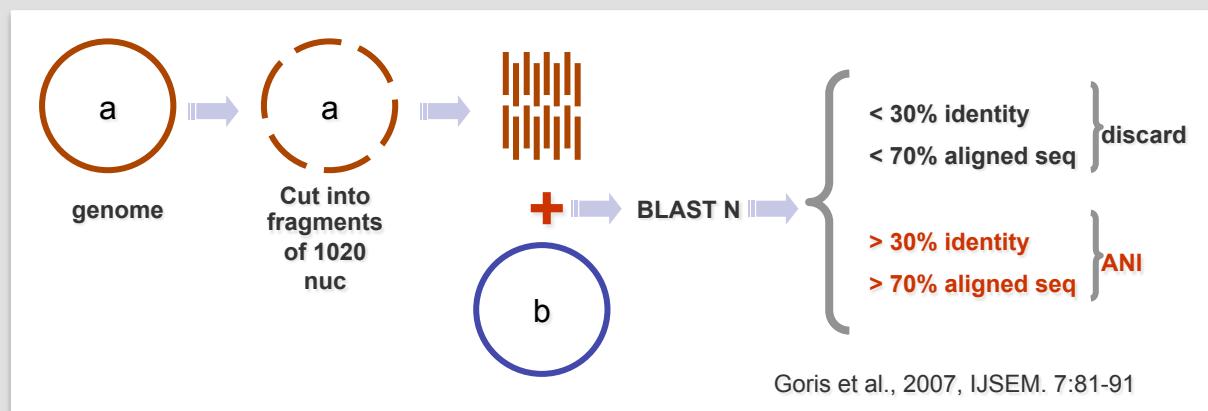
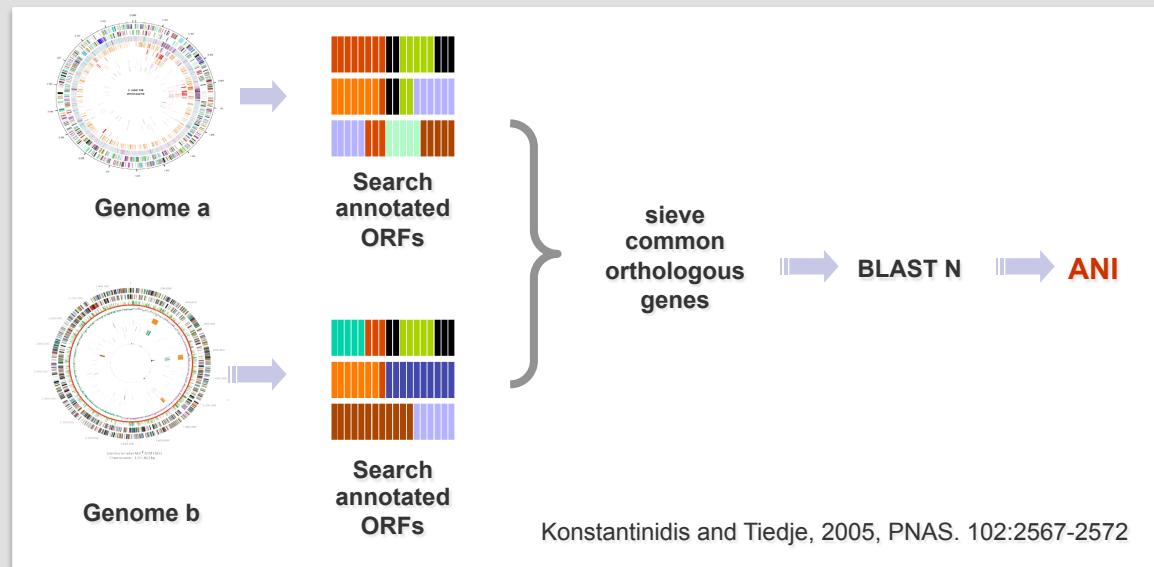


Stackebrandt et al., 2002, Int J Syst Evol Microbiol. 52:846-849

Gevers et al., 2005, Nature Rev. Microbiol. 3:733-739

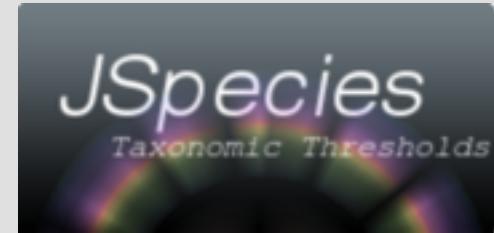
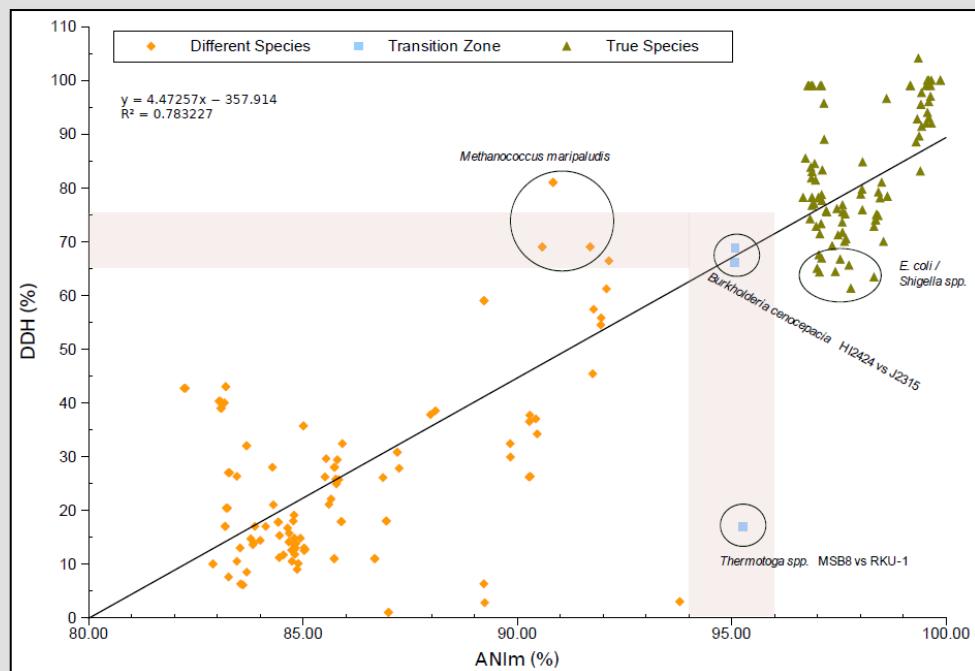
3- Studies on cultured prokaryotes

Alternative approaches ⇔ ANI



3- Studies on cultured prokaryotes

ANI can be calculated with JSpecies:
www.imedea.uib.es/jspecies (downloadable program)
<http://jspecies.ribohost.com/jspeciesws/> (online service)

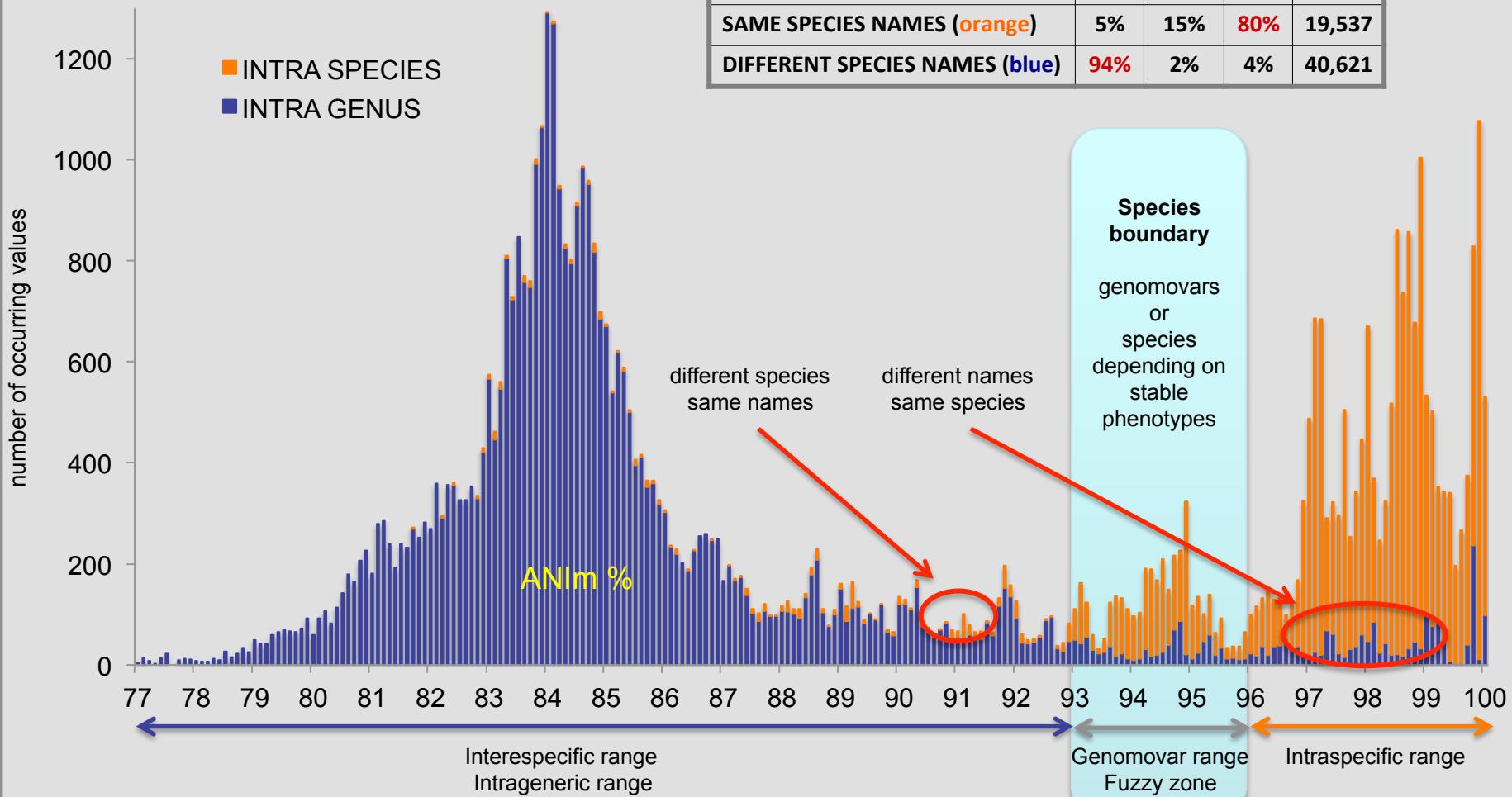


ANIm vs DDH:

- 85 genospecies evaluated
- **94-96%** a plausible borderline
- inconsistent results most probably due to wrong DDH values

- **ANI** thresholds of 94-96% ⇔ genomospecies
- **20%** random sequences (i.e., 250 nuc) of two genomes is enough
- Complete catalogue of type strain genomes ⇔ only **4%** random genome sequence is enough

Species boundaries



Rosselló-Móra & Amann (2015) Syst Appl. Microbiol 38:209-216

3- Studies on cultured prokaryotes

Tetranucleotide variation: $4^4 = 256$

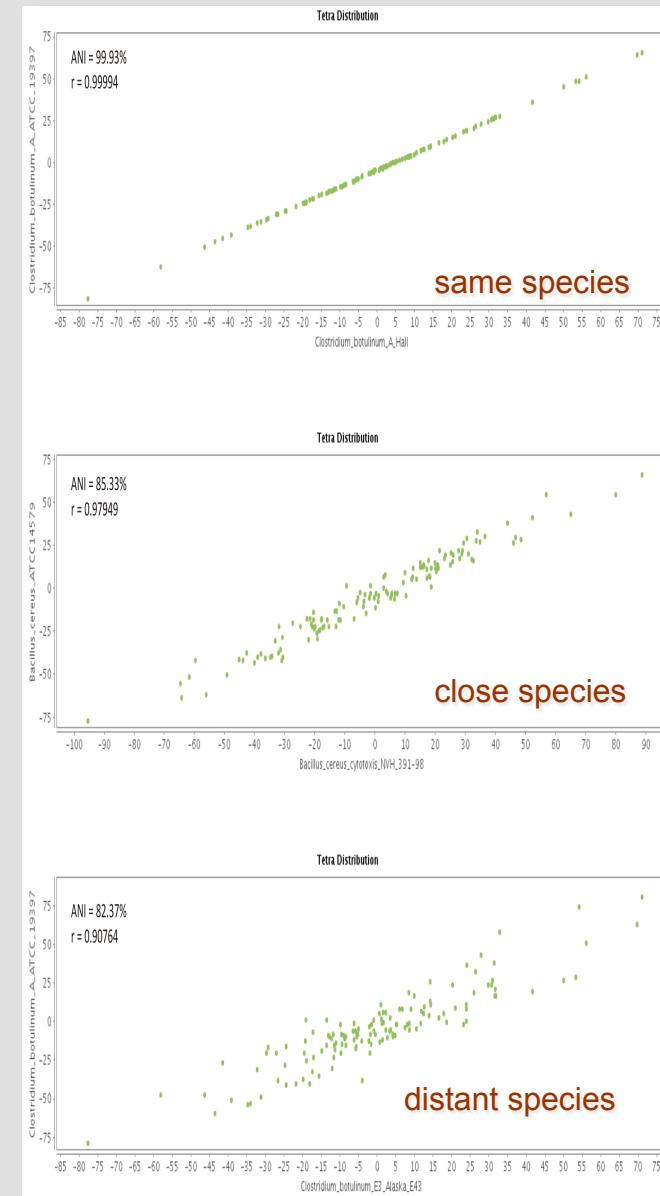
```

aaaa aaag aaac aaat aaga aagg aagc aagt aaca aacg aacc aact aata aatg aatc aatt
agaa agag agac agat agga aggg aggc aggt agca agcg agcc agct agta agtg agtc agtt
acaa acag acac acat acga acgg acgc acgt acca accg accc acct acta actg actc actt
ataa atag atac atat atga atgg atgc atgt atca atcg atcc atct atta attg attc attt
gaaa gaag gaac gaat gaga gagg gagc gagt gaca gacg gacc gact gata gatg gatc gatt
ggaa ggag ggac ggat ggga gggg gggt ggca ggcg ggcc ggct ggta ggtg ggtc ggtt
gcaa gcag gcac gcat gcga gcgg gcgc gcgt gcca gcgg gccc gcct gcta gctg gctc gctt
gtaa gtag gtac gtat gtga gtgg gtgc gtgt gtca gtcg gtcc gtct gtta gttg gttc gttt
caaa caag caac caat caga cagg cagc cagt caca cacg cacc cact cata catg catc catt
cgaa cgag cgac cgat cgga cggg cgcc cggt cgca cgcg cgcc cgct cgta cgta cgtc cgtt
ccaa ccag ccac ccat ccga ccgg ccgc ccgt ccca cccg cccc ccct ccta cctg cctc cctt
ctaa ctat ctac ctat ctga ctgg ctgc ctgt ctca ctcg ctcc ctct ctta ctgg cttc cttt
taaa taag taac taat taga tagg tagc tagt taca tacg tacc tact tata tatg tate tatt
tgaa tgag tgac tgat tgga tggg tggc tggt tgca tgcg tgcc tgct tgta tgtg tgtc tgtt
tcgg tcac tcat tcga tcgg tcgc tcgt tcca tccg tccc tcct tcta tctg tctc tctt
ttaa ttat ttac ttat ttga ttgg ttgc ttgt ttca ttcc ttct ttta ttgg tttc tttt

```

TETRA:

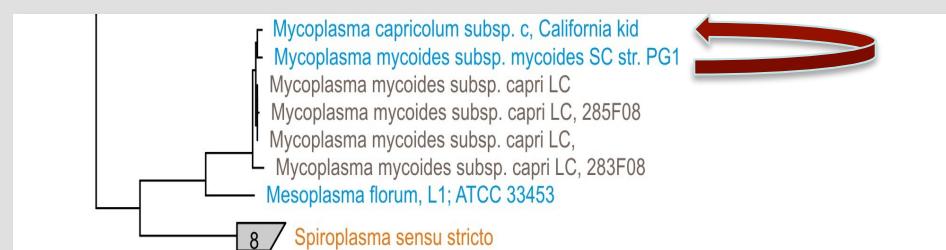
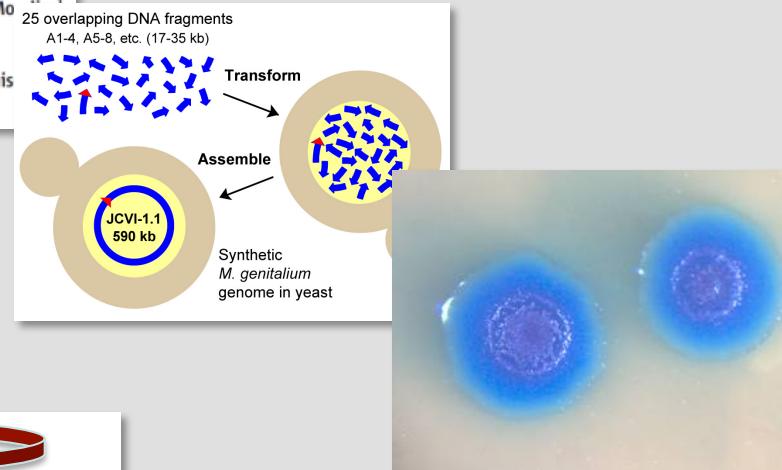
- Genomes have an oligonucleotide usage (not yet understood, related to codon usage)
- Similar genomes might have similar usage
- ALIGNMENT FREE PARAMETER
- may be useful in deciding whether a group of strains deserve a species status
- Same species >0.999



- The case of the synthetic genome of *M. mycoides* strain GM12 transplanted to *M. capricolum*
(Science (2010) 329: 52)

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

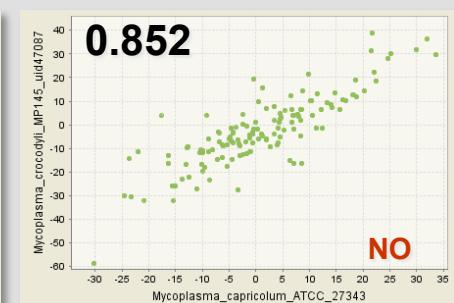
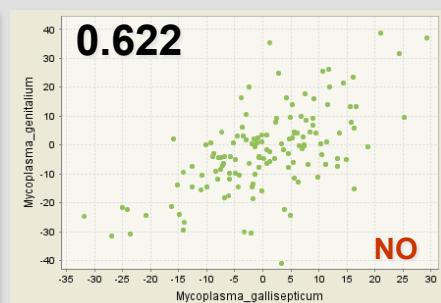
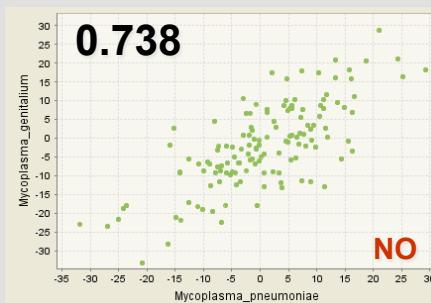
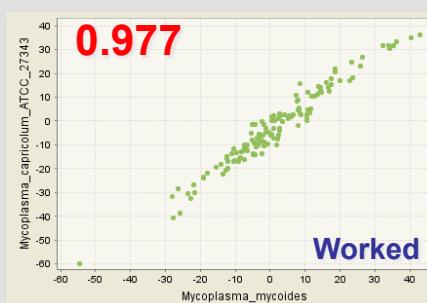
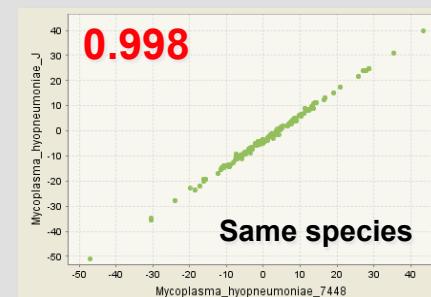
Daniel G. Gibson,¹ John I. Glass,¹ Carole Lartigue,¹ Vladimir N. Noskov,¹ Ray-Yuan Chuang,¹ Mikkel A. Algire,¹ Gwynedd A. Benders,² Michael G. Montague,¹ Li Ma,¹ Monzia M. Mo^{1,3}, Chuck Merryman,¹ Sanjay Vashee,¹ Radha Krishnakumar,¹ Nacyra Assad-Garcia,¹ Cynthia Andrews-Pfankoch,¹ Evgeniya A. Denisova,¹ Lei Young,¹ Zhi-Qing Qi,¹ Thomas H. Segall-Shapiro,¹ Christopher H. Calvey,¹ Prashanth P. Parmar,¹ Clyde A. Hutchison III,¹ Hamilton O. Smith,² J. Craig Venter^{1,2,*}



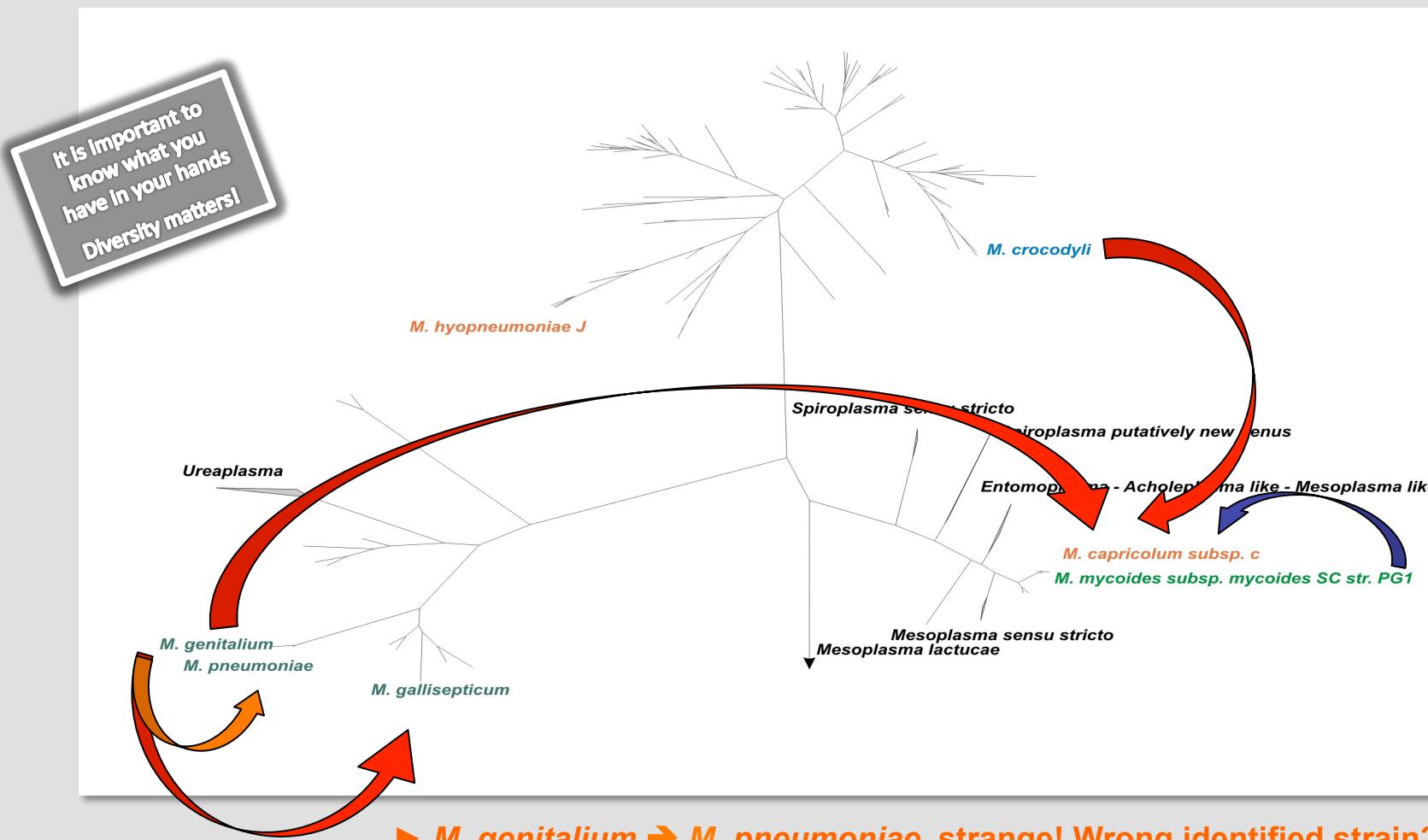
Genome transplantation experiments of Venter

- ▶ Only one of the several transplants worked out!
- ▶ Different ways of reading the genome?

organism	target	ANI	TETRA (r)
<i>M. hyopneumoniae</i> 7448	<i>M. hyopneumoniae</i> J	98.2	0.999
<i>M. mycoides</i> LC	<i>M. capricolum</i>	87.8	0.977
<i>M. genitalium</i>	<i>M. capricolum</i>	63.4	0.620
<i>M. genitalium</i>	<i>M. pneumoniae</i>	68.9	0.738
<i>M. genitalium</i>	<i>M. gallisepticum</i>	64.1	0.622
<i>M. aligatoris</i> (crocodyli)	<i>M. capricolum</i>	65.2	0.852



- The phylogenetic (evolutionary) distance plays an important role in the recognition of how the genetic information is coded



A wrong name may cause important consequences:

- ➔ *Bacillus cereus* var. *toyoi* ➔ important probiotic, in use since >30 years
- ➔ Misidentification promoted its banning in the European Union
- ➔ After reclassification still under study by the EFSA



Species

poultry
swine
cattle
rabbits
aquaculture

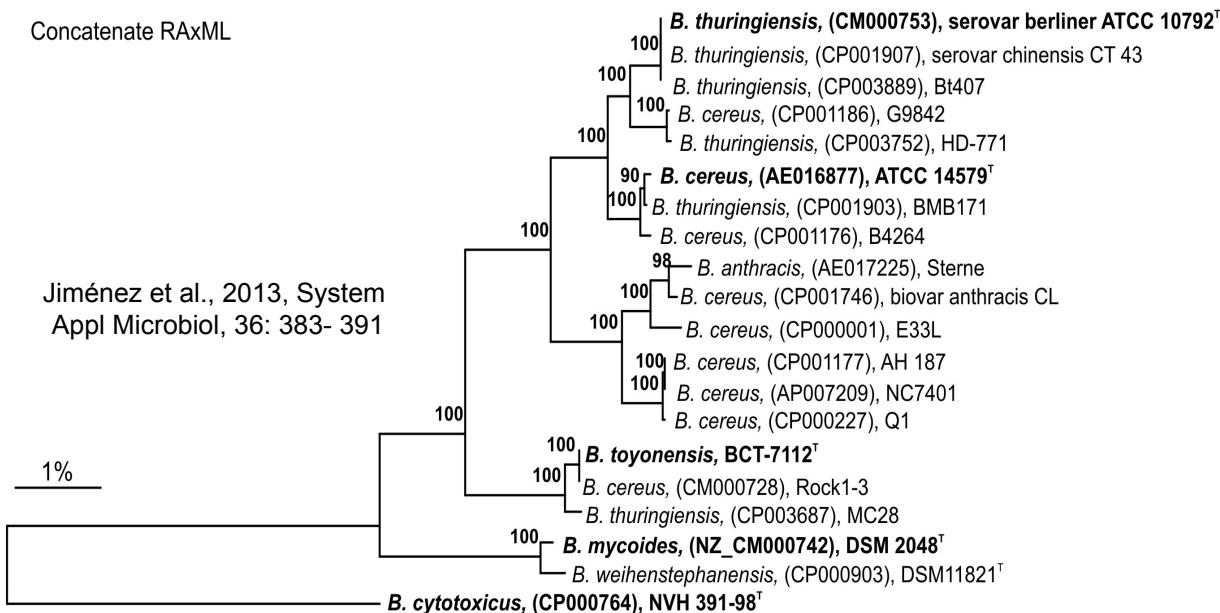
Toyocerin, Naturally profitable

TOYOCERIN® is a natural micro-organism (*Bacillus cereus* var. *toyoi*), with the function of microflora stabilizer in the gut of the animal which improves performance and enables a better digestion of the nutrients provided by the feed.

TOYOCERIN® is the first probiotic to receive authorisation in the European Union to be used as an additive in animal nutrition and is one of the most studied at scientific level. *Bacillus toyoi* is a strain originally isolated from soil in Japan, and which is not genetically modified. It is a microorganism naturally ubiquitous in the soil and is therefore likely to be ingested by wild animals that root in the soil, such as wild pigs or poultry. Besides its extensive market history - not only in Europe but worldwide (for example Japan, South East countries, South America) - *Bacillus toyoi* has demonstrated in vitro and in vivo its capacity to inhibit the growth of pathogenic bacteria of human concern such as *Salmonella* and *E. coli*. These results have been corroborated with studies that have demonstrated a reduction in *Salmonella* prevalence in broilers and fattening pigs fed with *Bacillus toyoi*.

Concatenate RAxML

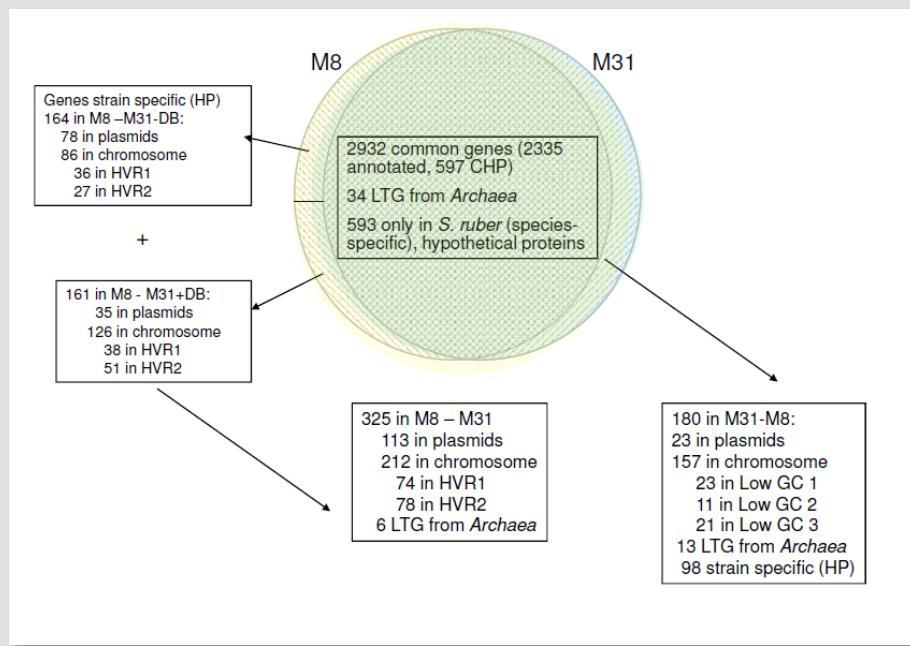
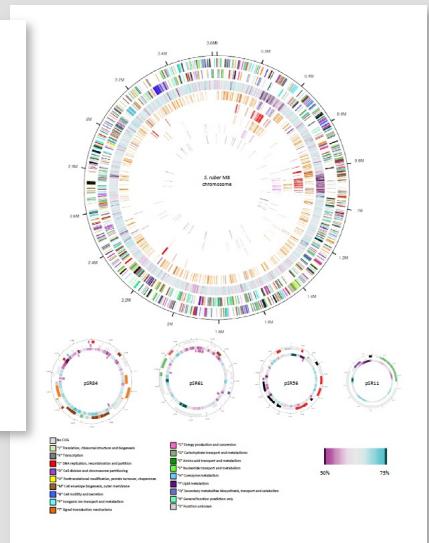
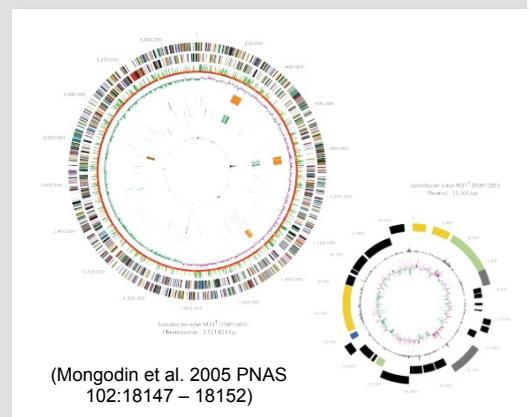
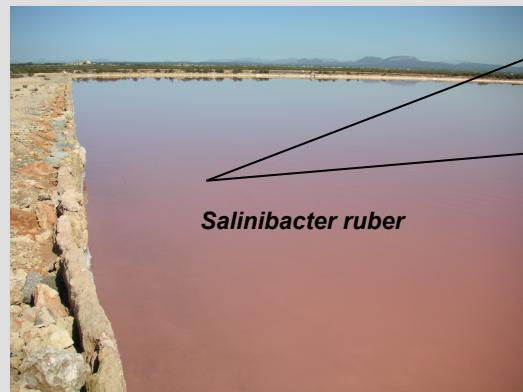
Jiménez et al., 2013, System Appl Microbiol, 36: 383- 391



It is important to
know who do you
have in your hands
Diversity matters!

3- Studies on cultured prokaryotes

One species in the environment is composed by different populations



- M8 (3,6Mb) vs M31 (3,5Mb) (3,086 ⇔ 2,934 orfs)
- 98% ANI; 80.3% DDH
- 4 plasmids vs 1 plasmid
- 2932 genes => core genome
- 593 unique species specific (Hyp Prot)
- 325 + 180 (505) strain specific genes (auxiliary)

(Peña et al. (2010) ISMEJ 4:882-895)

3- Studies on cultured prokaryotes

(Peña et al. (2010) ISMEJ 4:882-895)

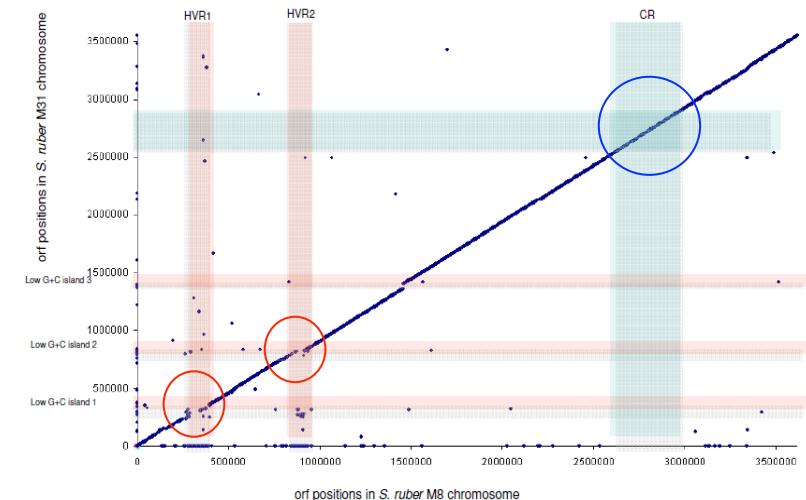
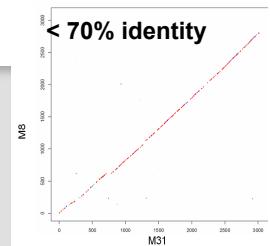
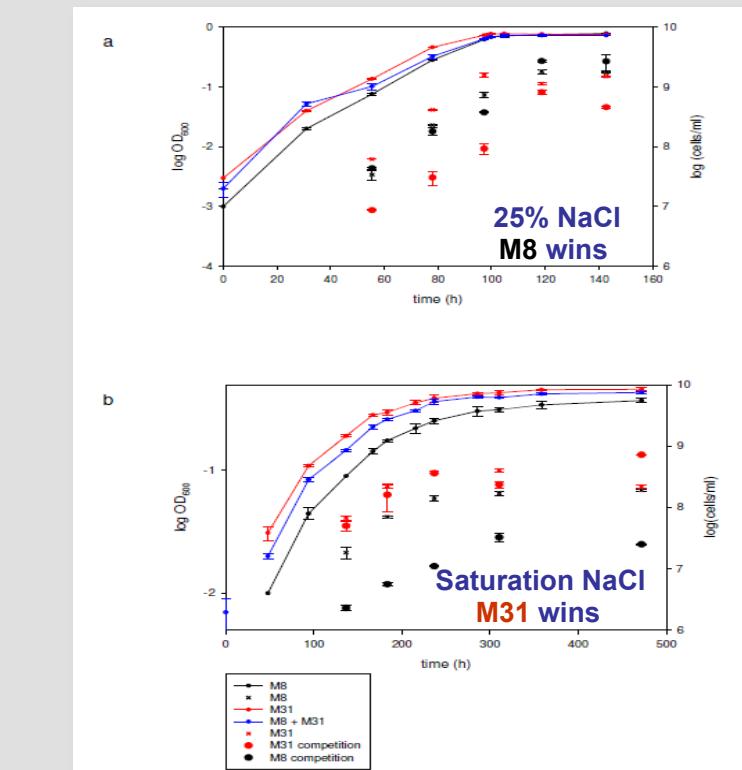
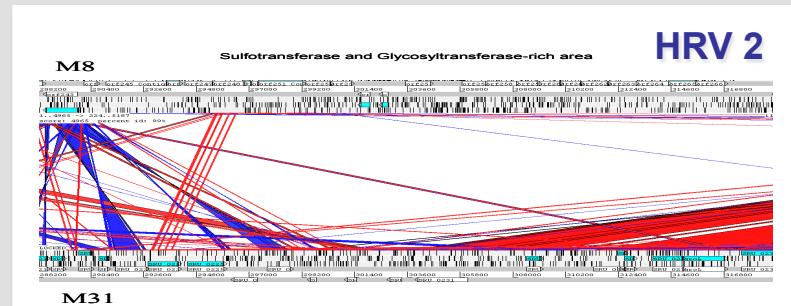


Figure sup. 2

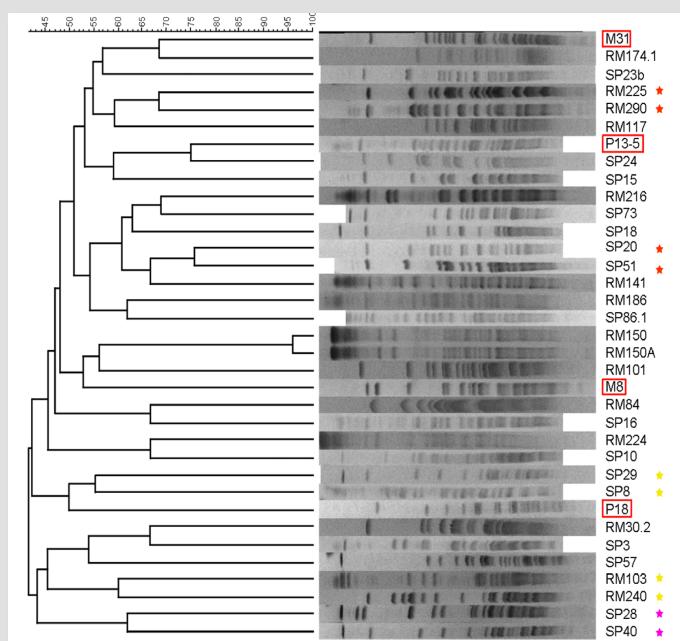
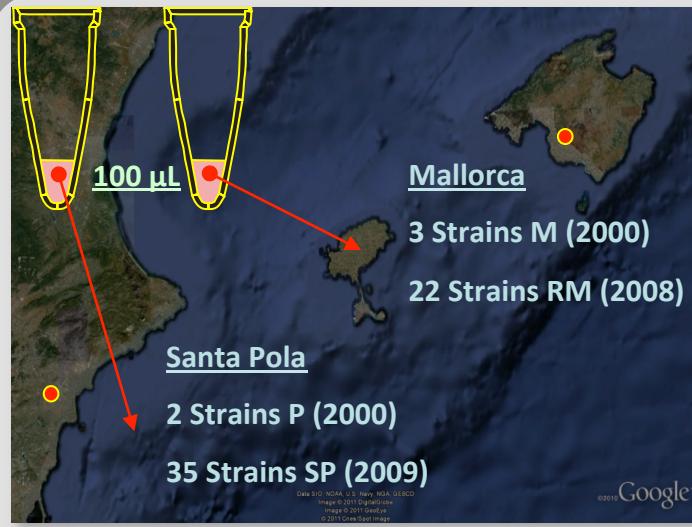


- good synteny of the genes
- divergent genes appear equally spread through the genome
- 1 conserved region (CR, >99.5% identity; ~400 Kb; inorganic transporters (K), aminoacid transporters)
- 2 hypervariable islands (HVR1 & HVR 2; ≠ GC/TETRA; cell envelope/outer membrane, transposases, glycosyltransferases)



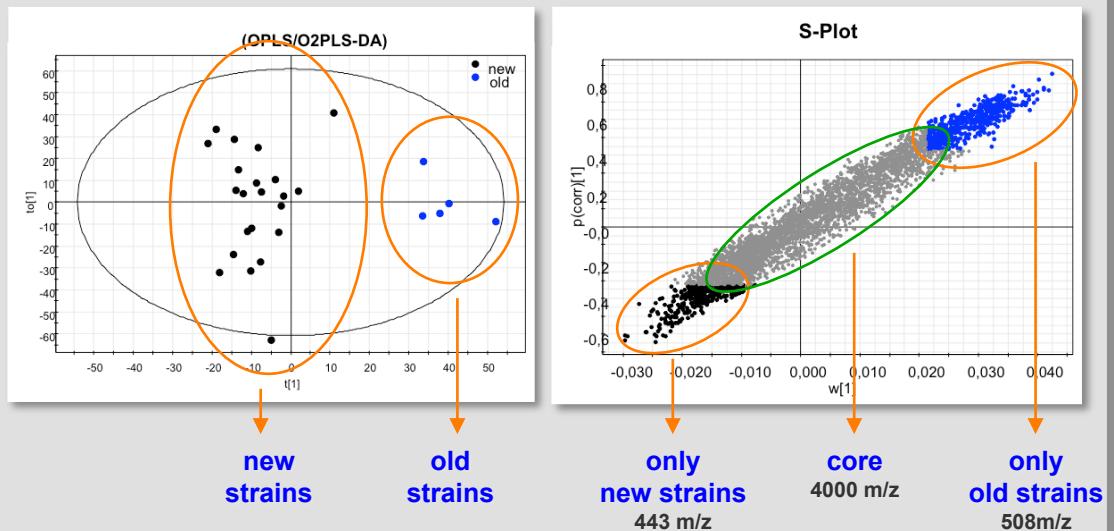
- M8 seems to be more competitive in 25% NaCl
- M31 is more competitive under NaCl saturation

3- Studies on cultured prokaryotes



Microdiversity study (isolates from 100 µL of brines)

Metabolomics



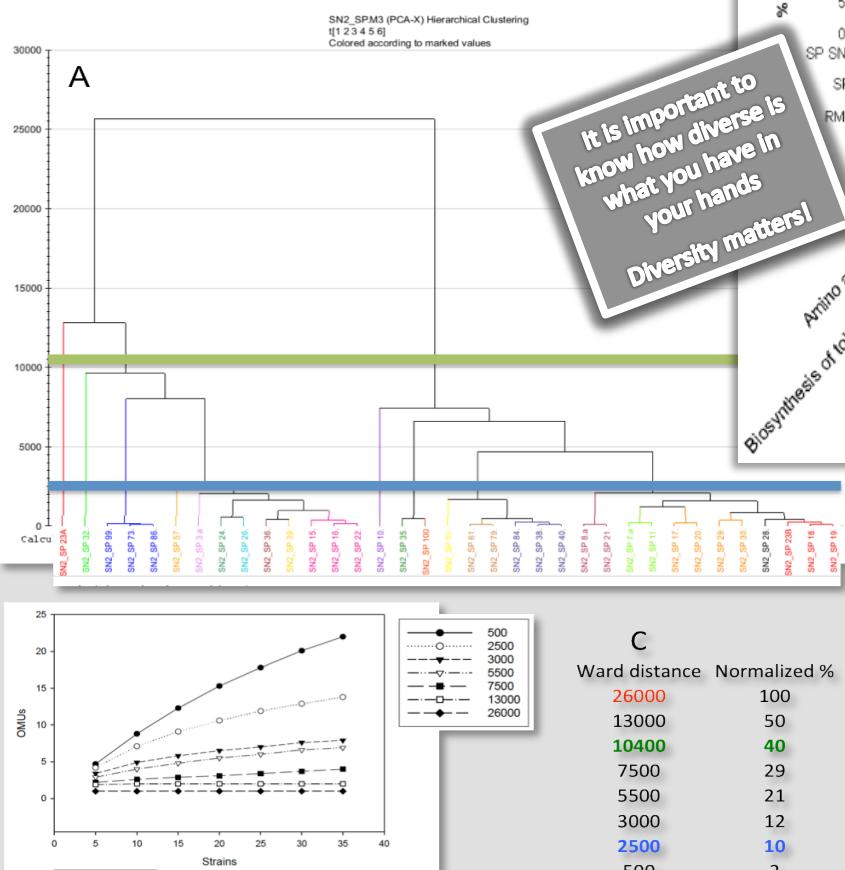
- No clonal varieties co-isolated
- clear distinction between old – new strains
- old strains resemble due to laboratory adaptation (only old)
- new strains synthesize metabolites only as wild type (only new)
- the core – metabolome is expressed in all strains

Antón et al. 2013 PLoS ONE 8(5): e64701

3- Studies on cultured prokaryotes

Metabolomic microdiversity

- metabolomes were very similar (one species)
- metabolomes were NOT identical



Discriminative metabolites among isolates:

- amino acid metabolism
- lipid metabolism
- secondary metabolism (antibiotics; polyketides...)

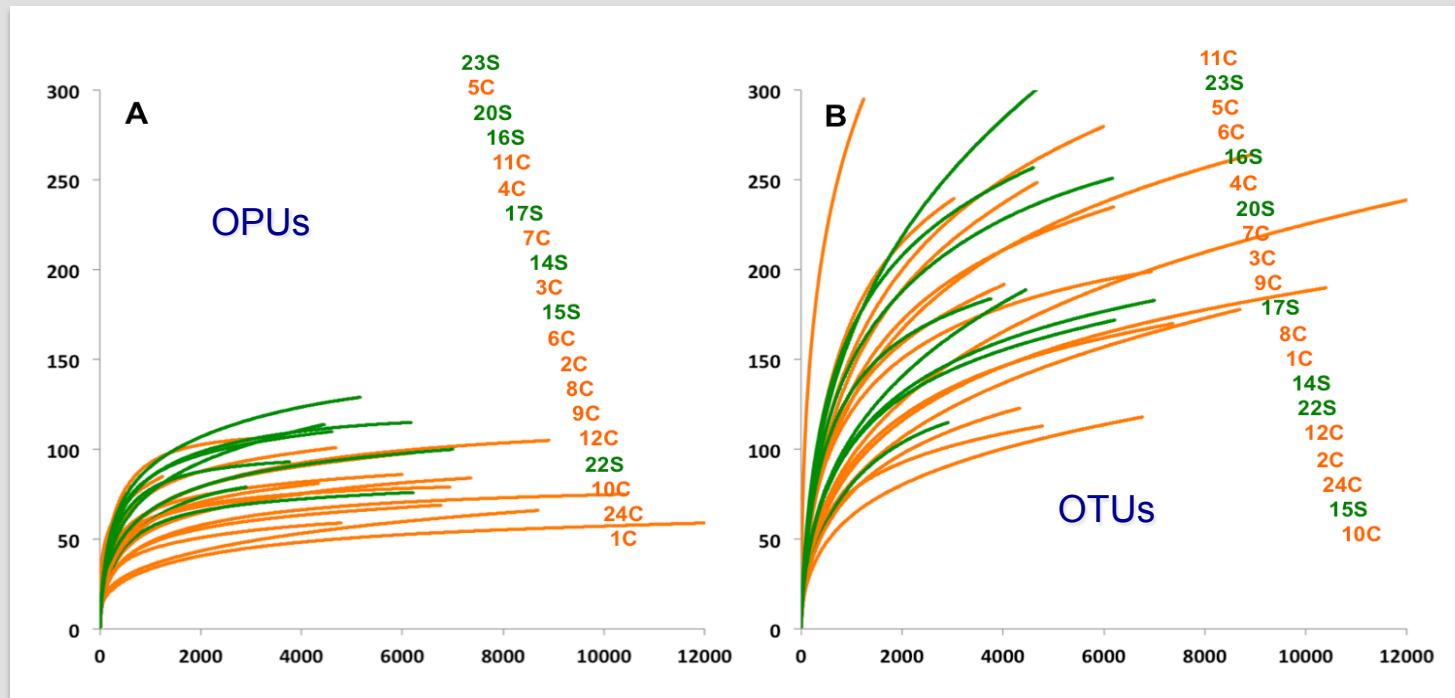
Antón et al. 2013 PLoS ONE 8(5): e64701

3- Studies on cultured prokaryotes

Culture-free studies help in classifying uncultured diversity

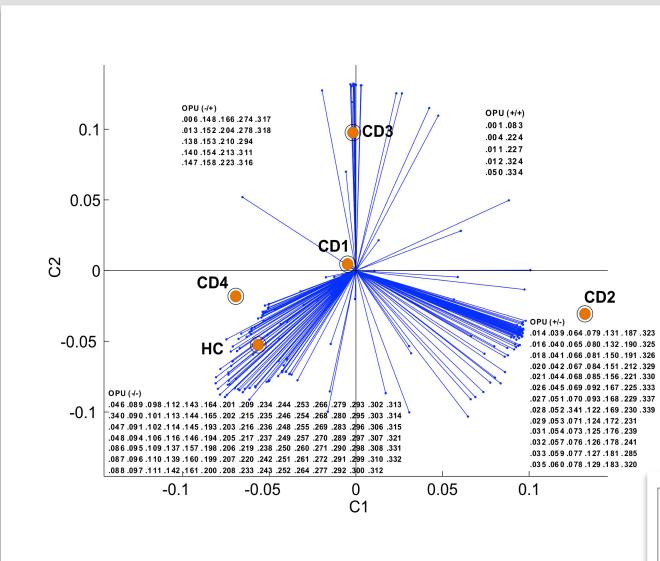
Example case ⇔ biopsies associated to Crohn microbiomes

- 20 samples (7 Healthy ; 13 Crohn)
- 150,000 sequences (mean 550 nuc; <300 nuc discarded)
- Mean of $6,592 \pm 2,622$ in each sample
- 73% with known taxa (up to species identity)

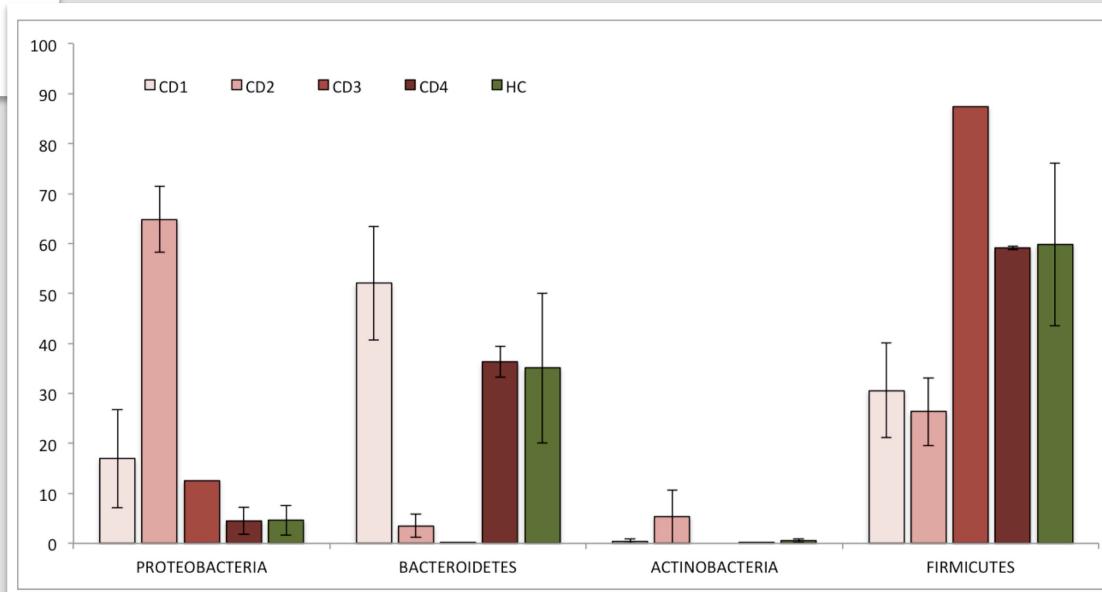


Vidal et al. Syst Appl Microbiol (2015) 38, 442-452

Example case ⇔ biopsies associated to Crohn microbiomes



- PCA allowed the discrimination of 4 CD groups and the responsible OPUs
- HC (balanced *Firmicutes / Bacteroidetes*)
- CD1 (depletion of *Firmicutes*, increase of *Bacteroidetes*)
- CD2 (depletion of *Firmicutes & Bacteroidetes*; increase of *Proteobacteria*)

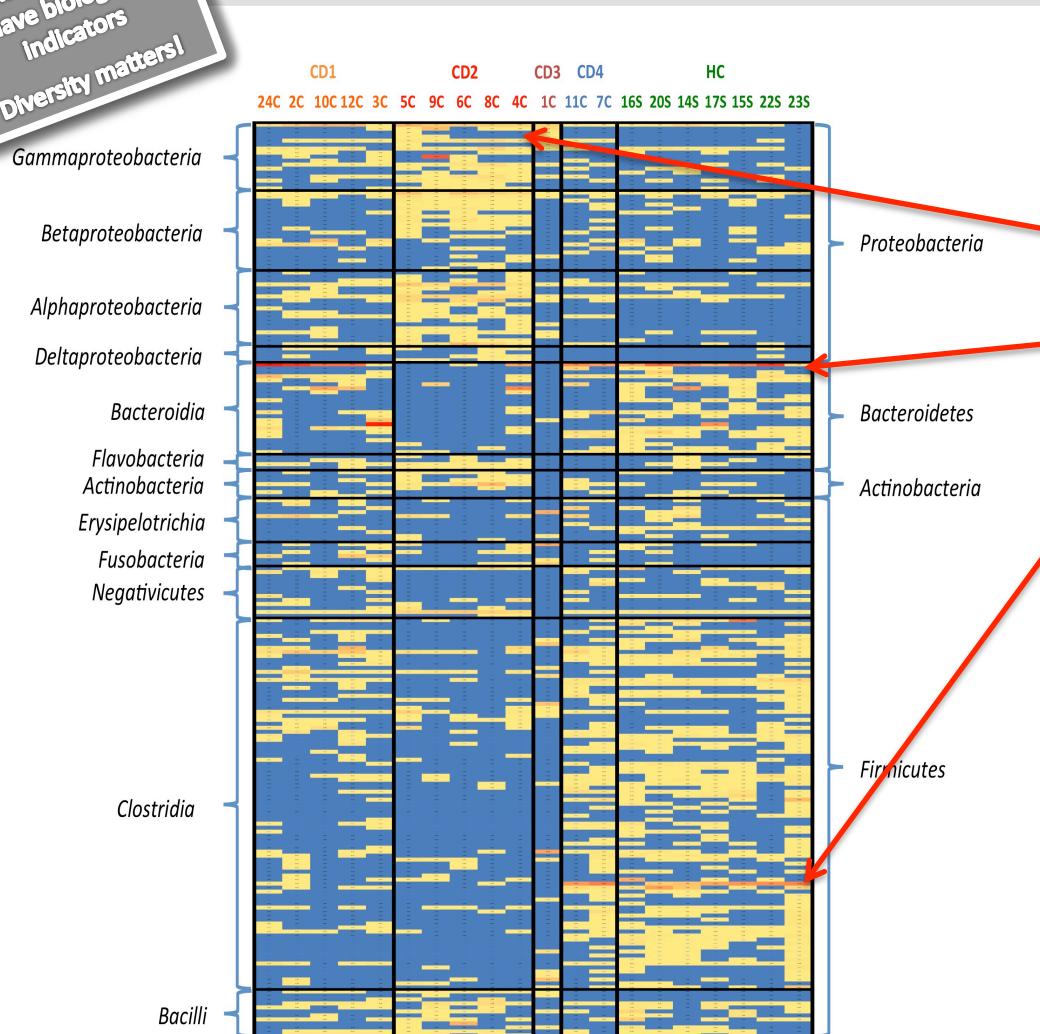


Vidal et al. Syst Appl Microbiol (2015) 38, 442-452

4- Studies on uncultured prokaryotes

It is important to
know whether you
have biological
indicators
Diversity matters!

Example case ⇔ biopsies associated to Crohn microbiomes



- CD 4 was indistinguishable from healthy patients
- *E. coli* the most representative proteobacterium in CD
- *Bacteroides volutans* / *B. dorei* the most representative organisms in the mucose
- *Faecalibacterium prausnitzii* disappears in CD
- CD may correspond to different disorders with same inflammatory response
- Or
- CD1 is an intermediate stage between Healthy and CD2

Vidal et al. Syst Appl Microbiol (2015) 38, 442-452

Example case ⇔ epi- and endophytic microbiomes of halophytes

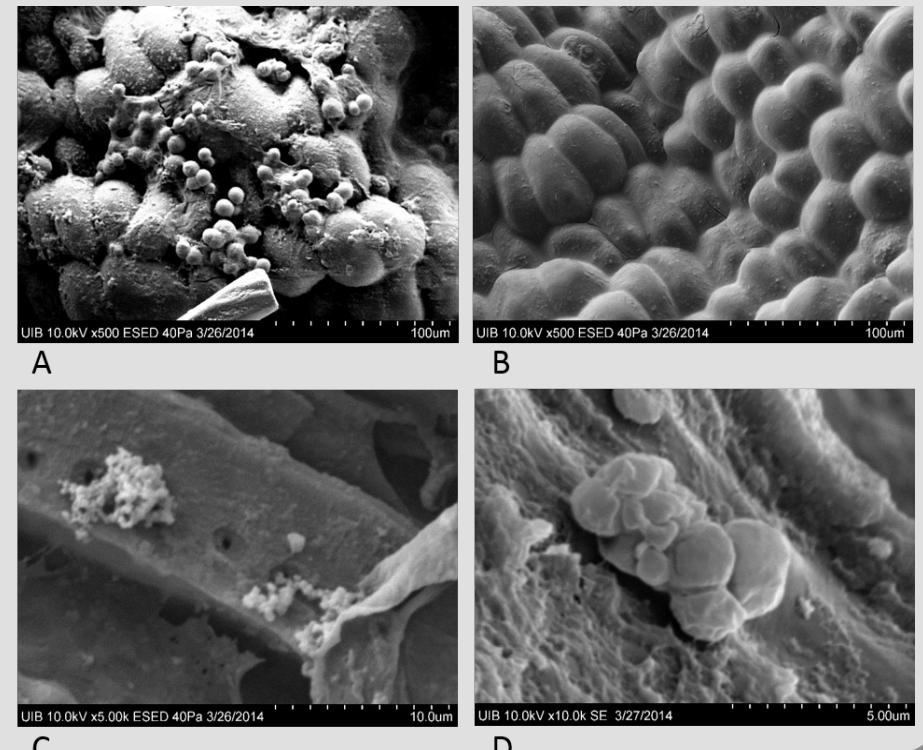


Study on halophytes (*Arthrocnemum macrostachyum*)

- ➔ 5 locations in Spain (Mallorca and Alicante) and Chile (Lo Valdivia and Pichidangui)
- ➔ Collaborative project between IMEDEA and Univ. Adolfo Ibañez
- ➔ **Pyrotagging + culturing** (culture dependent and independent methods)
- ➔ 14,000 sequences (mean 550 nuc)
- ➔ 823 isolates
- ➔ Surface sterilization

Endosymbionts may
be growth
promoters with
interest in
biotechnology
Diversity matters!

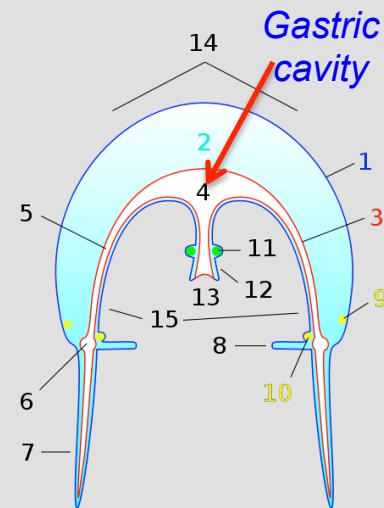
A. macrostachyum grows in
salinized soils and accumulates
high amounts of salts in their
tissues



Mora-Ruiz et al. (2015) Syst Appl Microbiol 38: 406-416

4- Studies on uncultured prokaryotes

Example case ⇔ microbiomes of jellyfishes (*Cotylorhiza tuberculata*)

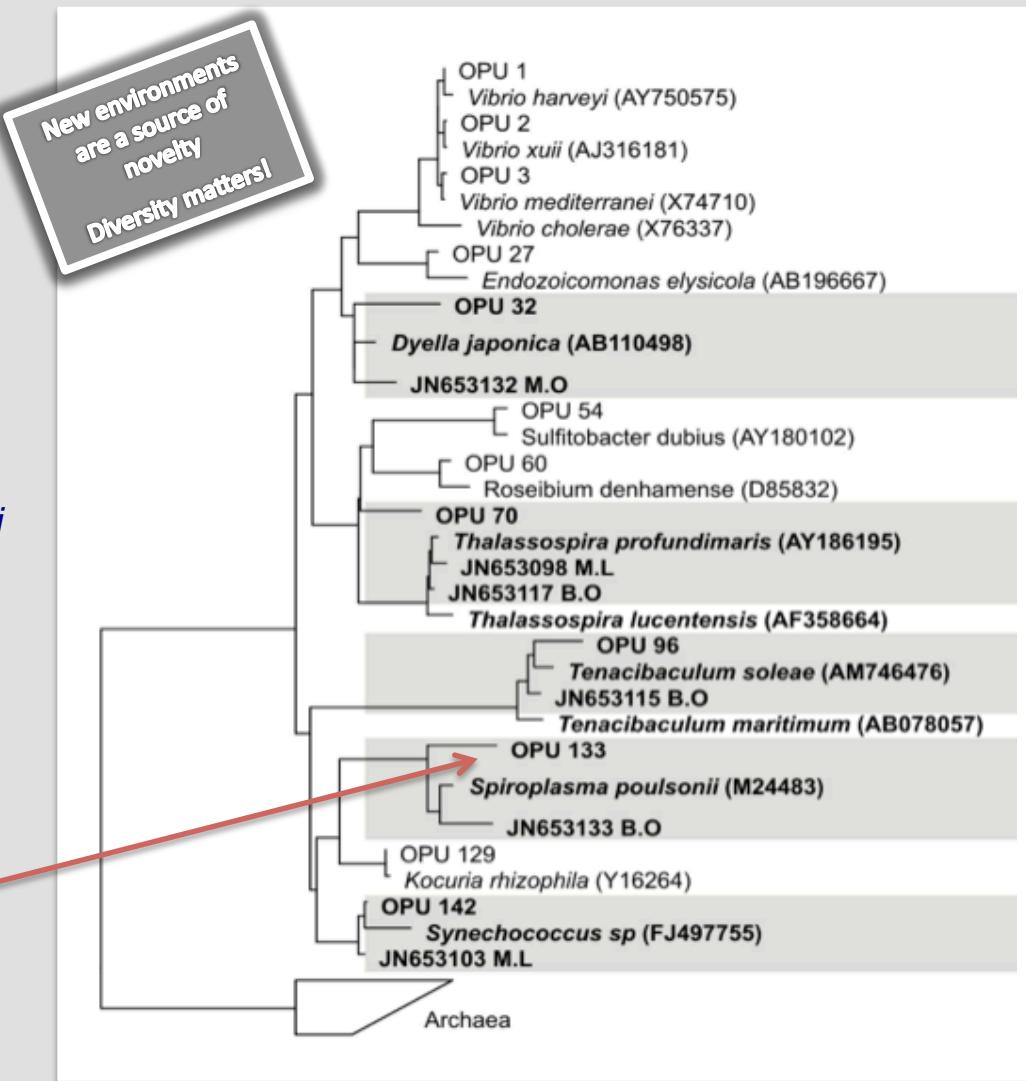


- 4 exemplars from the shore of Mallorca
- *C. tuberculata* generates large outbreaks in autumn (tons of dead biomass at the end of November)
- Pyrotagging + culturing (culture dependent and independent methods)
- 59,117 sequences (mean 550 nuc)
- 190 isolates
- Only in the gastric cavity (digestive organ tightly connected with the seawater by mouth arm openings)

Cortés et al. (2015) Syst Appl Microbiol 38: in press

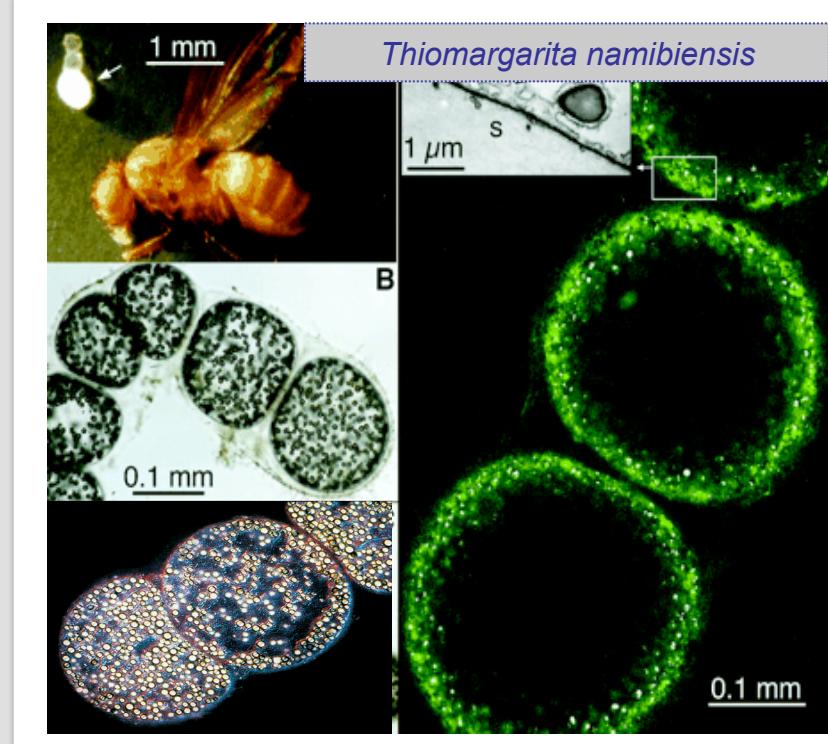
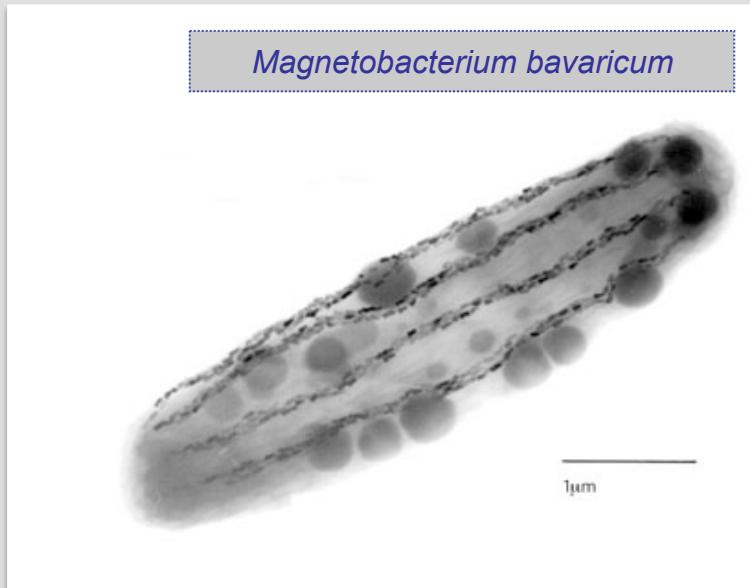
Example case ⇔ microbiomes of jellyfishes (*Cotylorhiza tuberculata*)

- 1st report in *Scyphozoa*
- Very low diversity 3 organisms making 99% diversity
- 3 key players (*Spiroplasma* 82%; *Thalassospira* 11% and *Tenacibaculum* 6%)
- 2 major cultured organisms *Vibrio harveyi* & *V. xuii* (both detected in pyrotagging)
- *Tenacibaculum* and *V. harveyi* potential pathogens (especially fishes) ⇔ jellyfish as carrier
- *Spiroplasma* potential mutualist ⇔ new lineage ⇔ **new candidatus?**



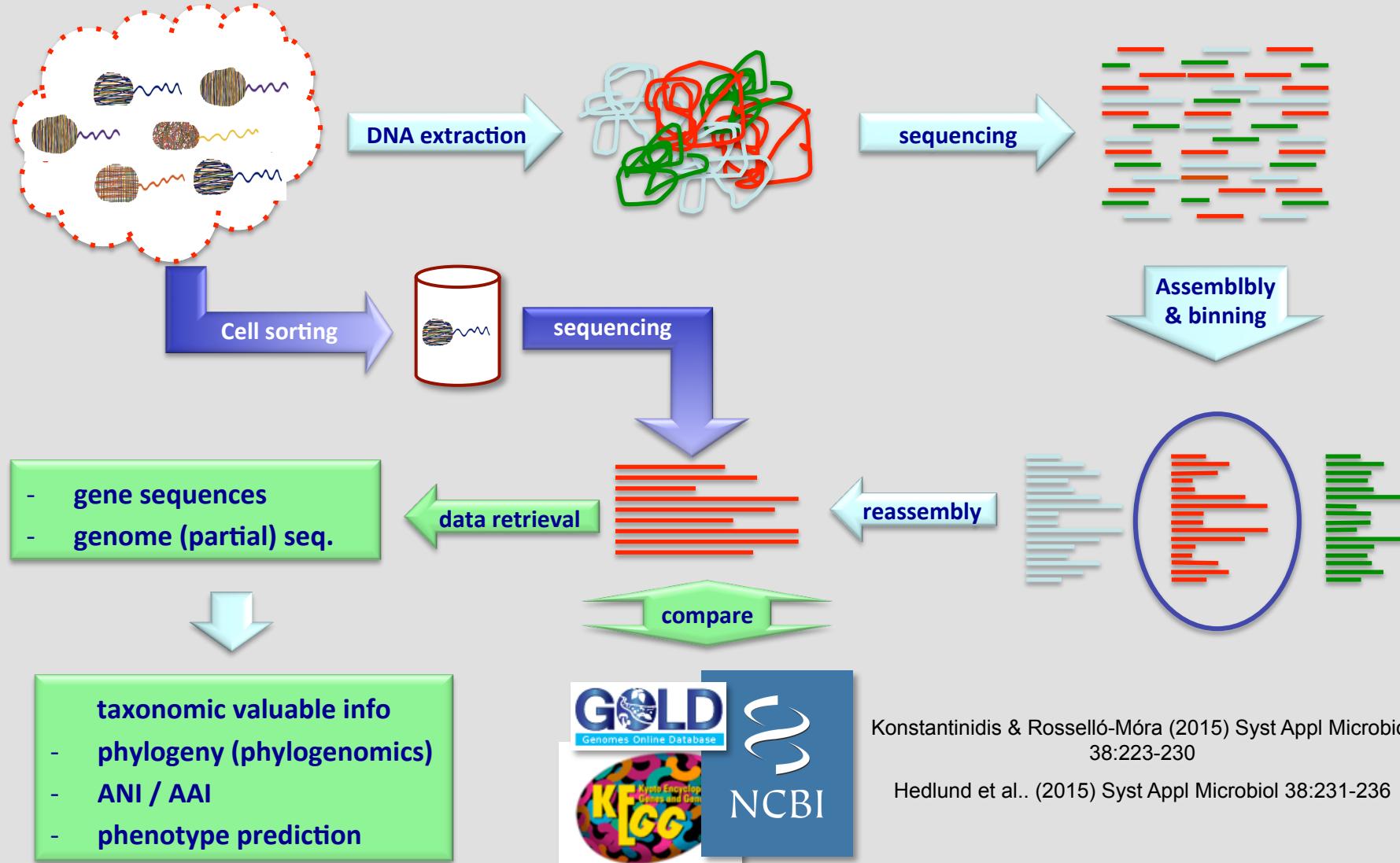
Cortés et al. (2015) Syst Appl Microbiol 38: in press

Candidatus status for uncultured taxa (only for conspicuous organisms)



- ➔ PROVISIONAL STATUS
- ➔ Clear phylogenetic reconstruction (know their genealogy)
- ➔ quantify & explain ecological traits (FISH)
- ➔ Explain some metabolic traits (MAR-FISH, conspicuous metabolism, cellular features, ...)
- ➔ Up to now only 360 *Candidatus* species recognized

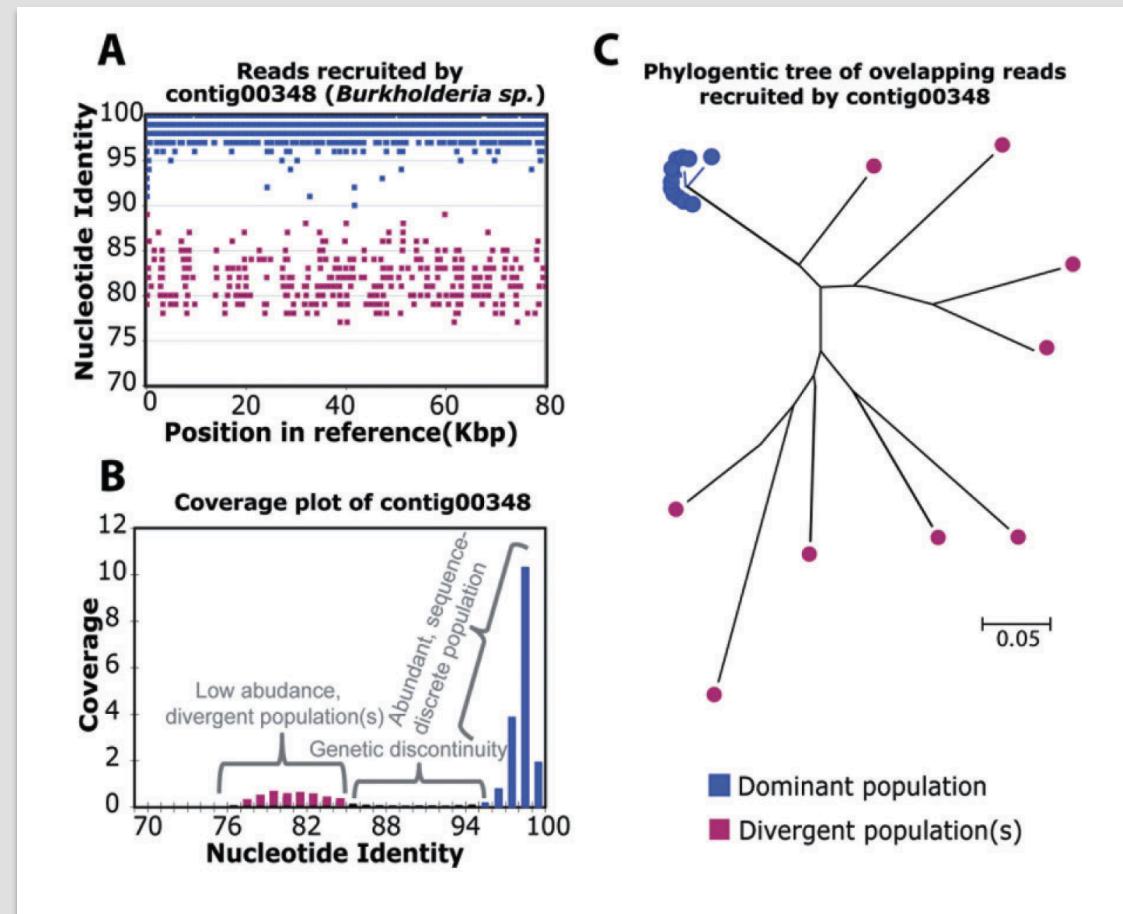
Metagenomics & Single Cell Genomics (a source of very useful data)



4- Studies on uncultured prokaryotes

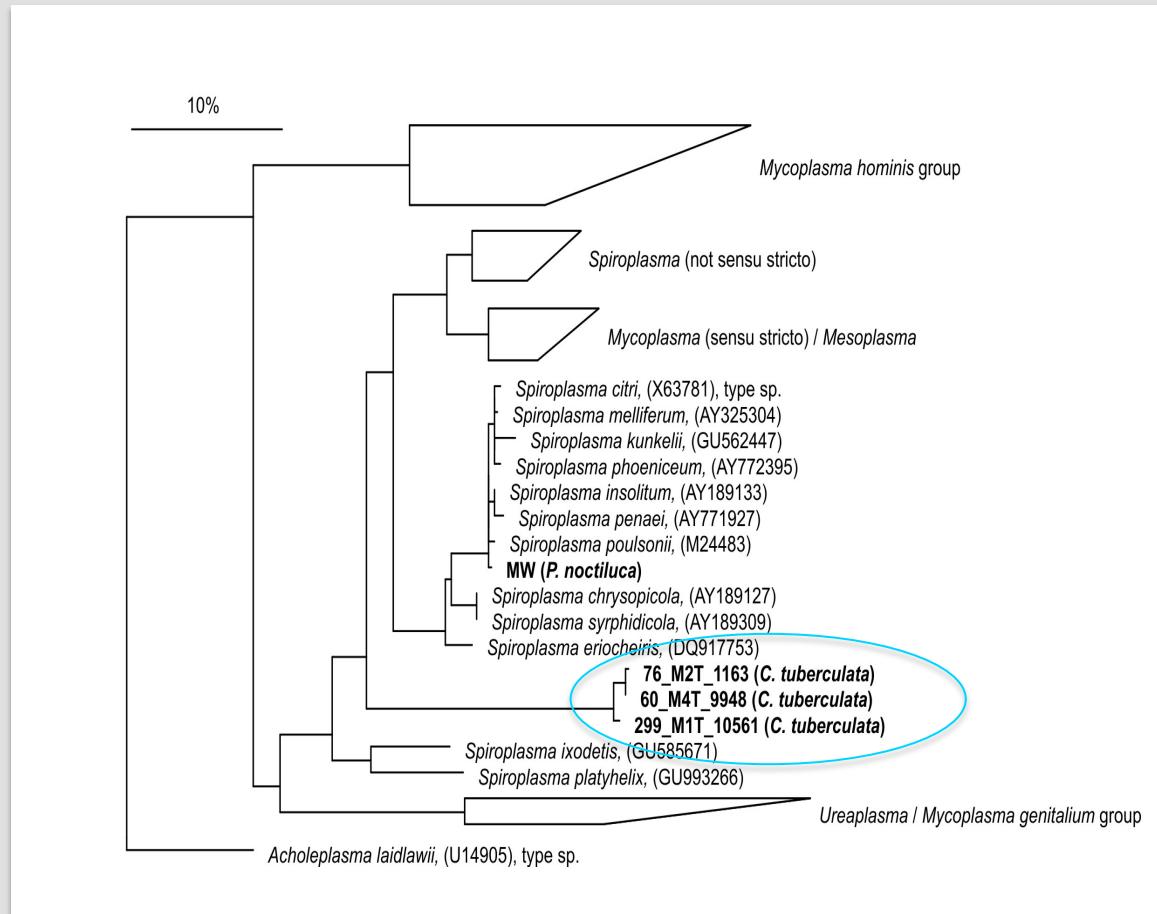
Candidatus based on metagenomes (new approach)

- reference ⇔ large contig
- Sequences aligned at 70% identity
- two populations:
 - blue => highly related
 - red => divergent populations
- genetic discontinuity in the same environment



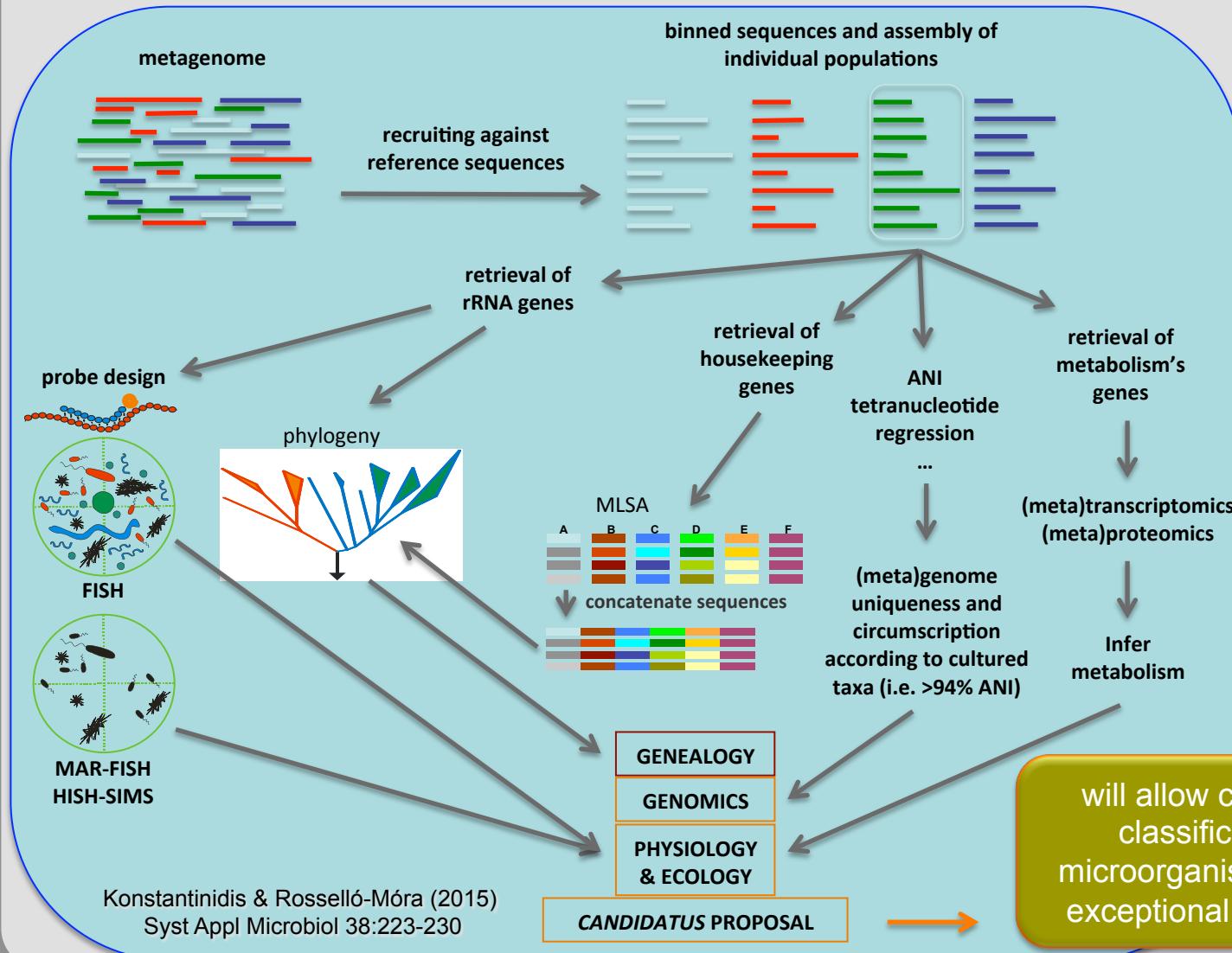
Caro-Quintero & Konstantinidis 2012, Environ Microbiol 42: 347-355

Candidatus based on metagenomes (new approach)

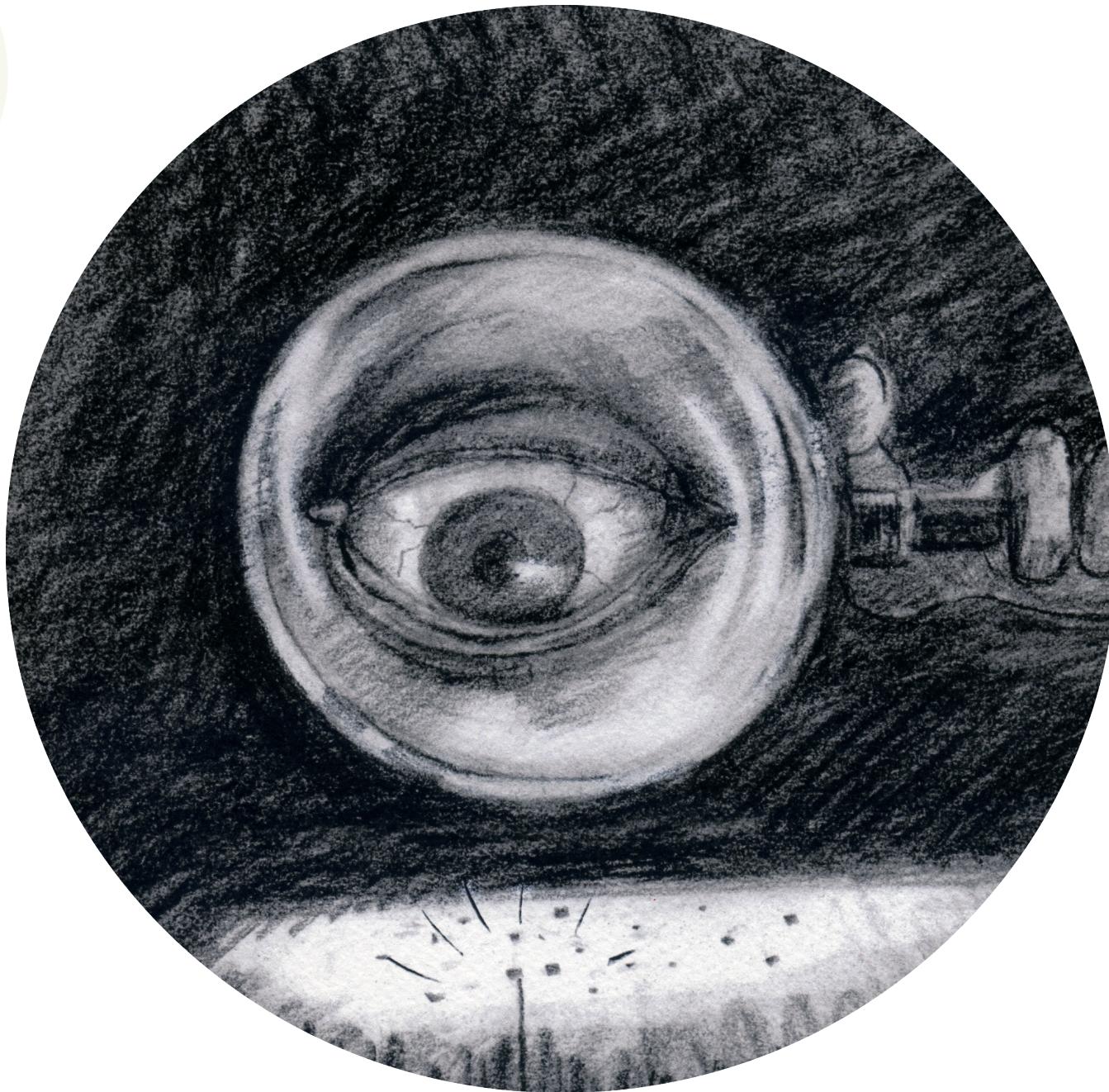


➔ *Spiroplasma* potential mutualist ⇔ new lineage ⇔ new candidatus? ⇔ metagenomics

Candidatus based on metagenomes (new approach)



- ▶ ANI >94% may reflect species (mosaic ⇔ distinct populations)
- ▶ ANI <94% may reflect different coexisting species
- ▶ *Candidatus* proposal of inconspicuous organisms (most) will be possible



OTHER PARAMETERS

Maximal Unique and Exact Matches (MUM)

De Loger et al., 2009, J. Bacteriol. 191: 91-99

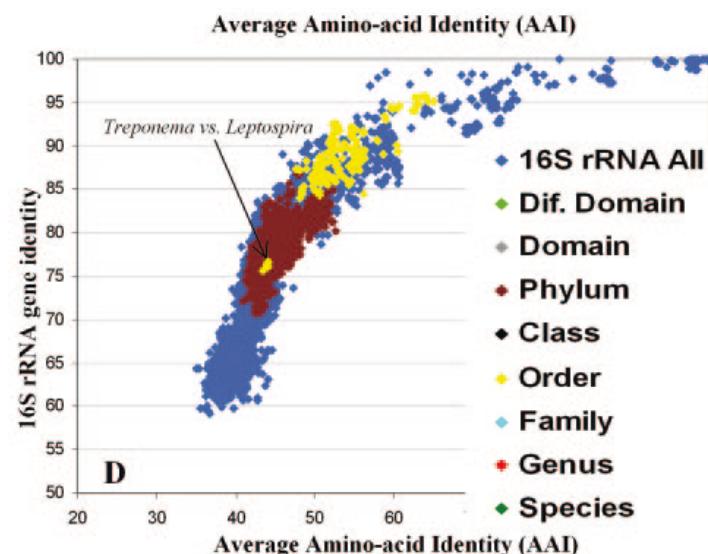
High Scoring Segment Pairs (HSP)

Auch et al., 2010. Std Gen Sci 2:117-134

And more to come

Need full genome sequences
The easiest is the best

Average Aminoacid Identity (AAI)



Kostantinidis & Tiejde, 2005, J. Bacteriol. 187:6258-6264