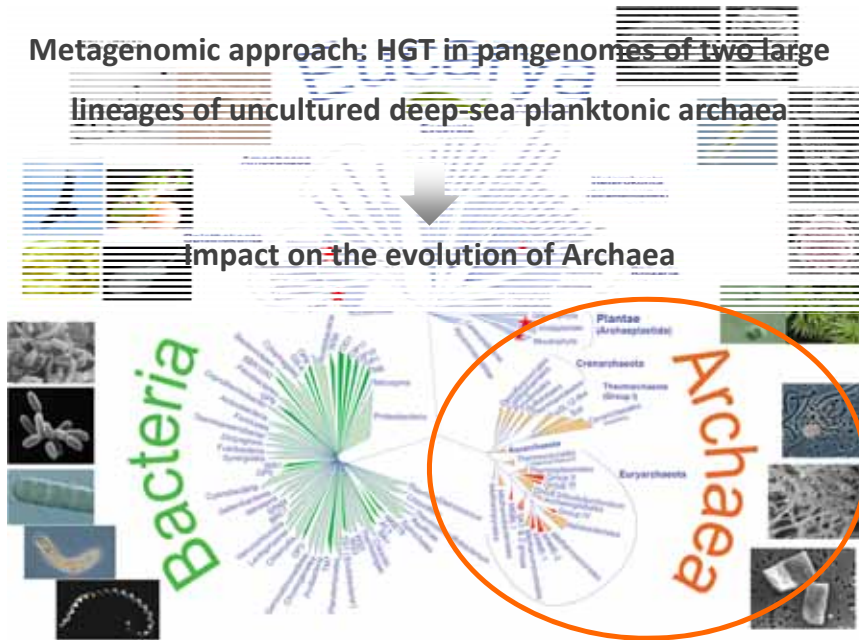


Inter-domain horizontal gene transfer in the evolution of Archaea

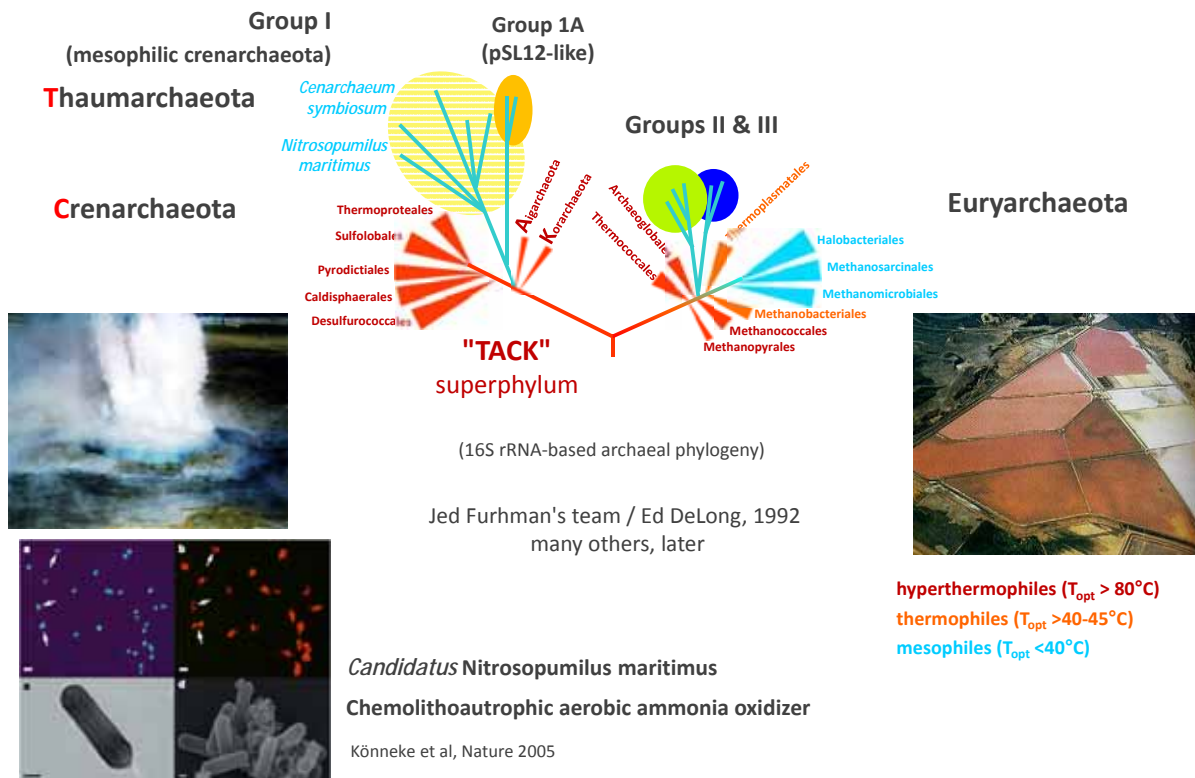
Purificación López-García

Unité d'Ecologie, Systématique et Evolution, CNRS & Université Paris-Sud, France

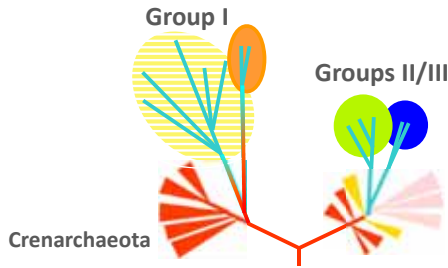
Metagenomic approach: HGT in pangenomes of two large lineages of uncultured deep-sea planktonic archaea



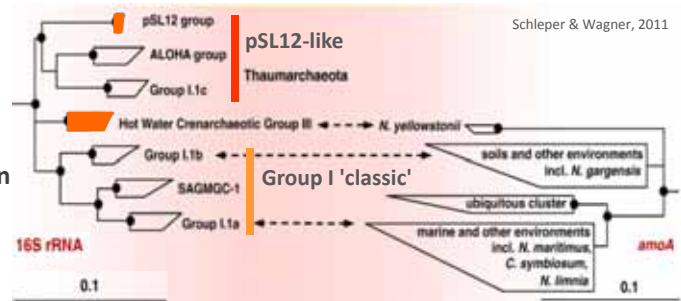
Discovery of marine planktonic archaea



Group I archaea: Thaumarchaeota



Euryarchaeota



Schleper & Wagner, 2011

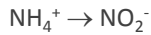
- High diversity in sea, soil, freshwater
- Non cultivated but abundant in **deep ocean** waters (up to 30% picoplankton)

- Novel phylum **Thaumarchaeota**

Brochier et al., Nat. Rev. Microbiol. 2008

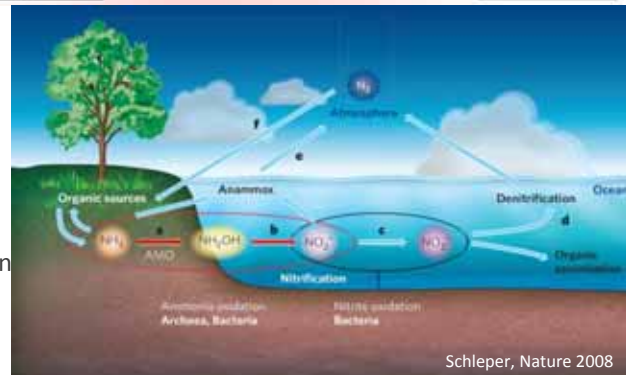
- **Energy metabolism:**

aerobic oxidation of ammonia



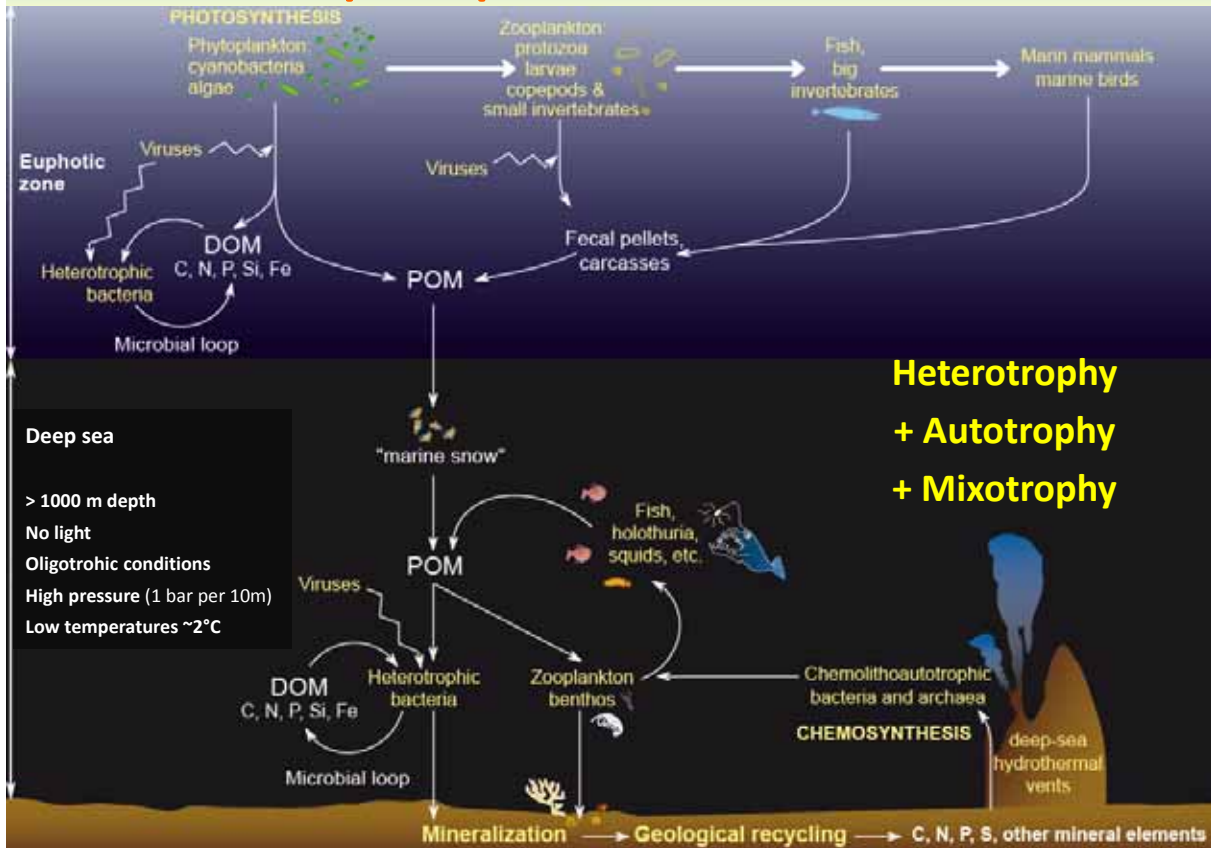
- **C metabolism:** autotrophic and/or mixotrophic?

3-hydroxypropionate/4-hydroxybutyrate genes in genomes of *C. symbiosum* & *N. maritimus*



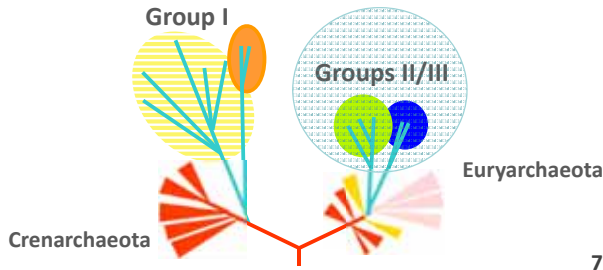
Schleper, Nature 2008

Deep-sea planktonic communities



**Heterotrophy
+ Autotrophy
+ Mixotrophy**

The enigmatic marine Groups II & III archaea



- Exclusively marine
- Non cultured
- Metabolism & lifestyle unknown

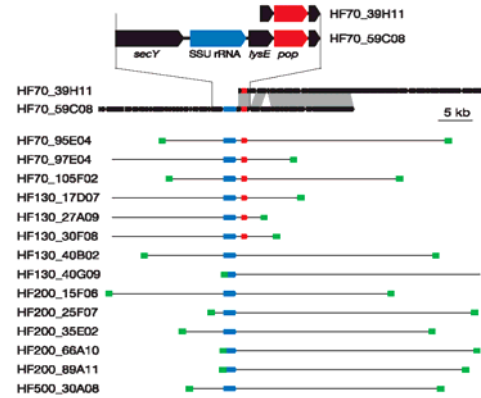
Group II

- Usually higher abundance in surface waters (DCM)

Group III

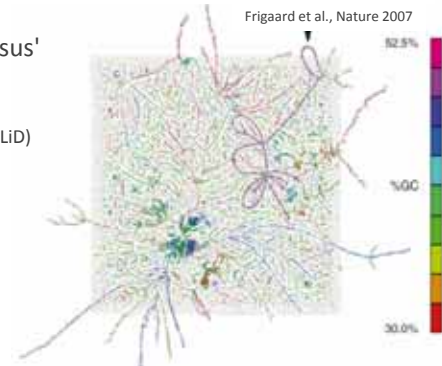
- Deep-sea waters

70-130 m



- 'Majority consensus' MGII genome (de novo assembly, SOLiD)
- Surface waters
- Proteorhodopsin

Iverson et al., Science 2012



Frigaard et al., Nature 2007

Metagenomics: getting access to the genomes of uncultured microbes



-----> [Single cell genomics]

DNA

Massive sequencing

(Shotgun libraries-Sanger sequencing)

Direct NGS sequencing: pyrosequencing 454 / Illumina

Gene content of given environment

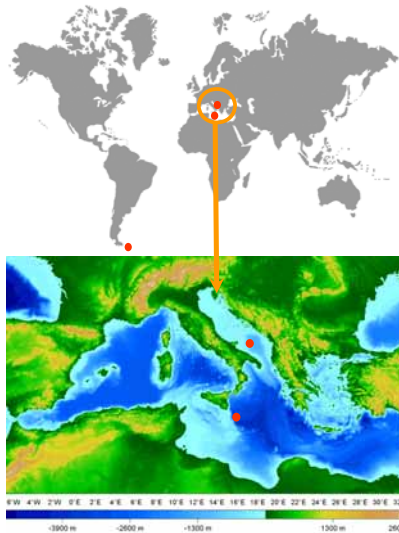
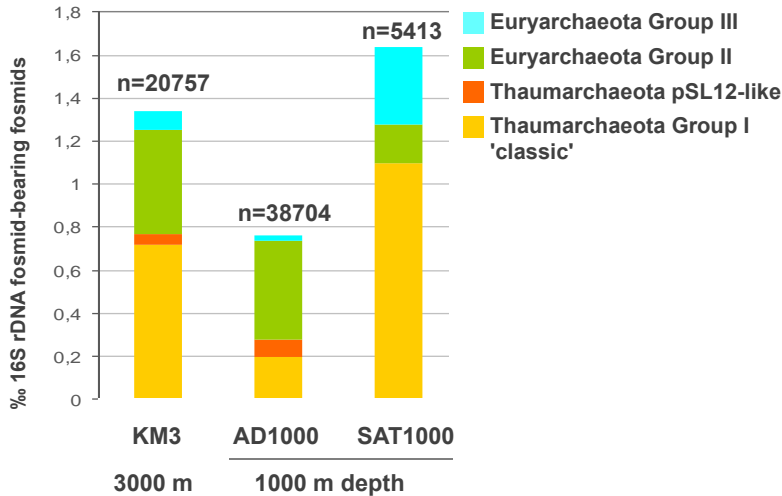
(Genome scaffolds in low diversity environments or of dominant organisms)

Metagenomic libraries

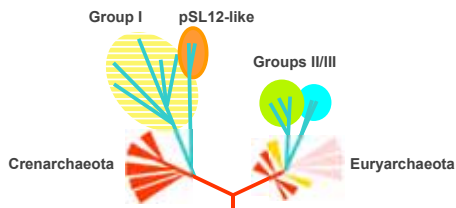
fosmids, BACs

Gene content & genome structure of phylogenetically identifiable organisms

Archaea in deep-sea metagenomic libraries



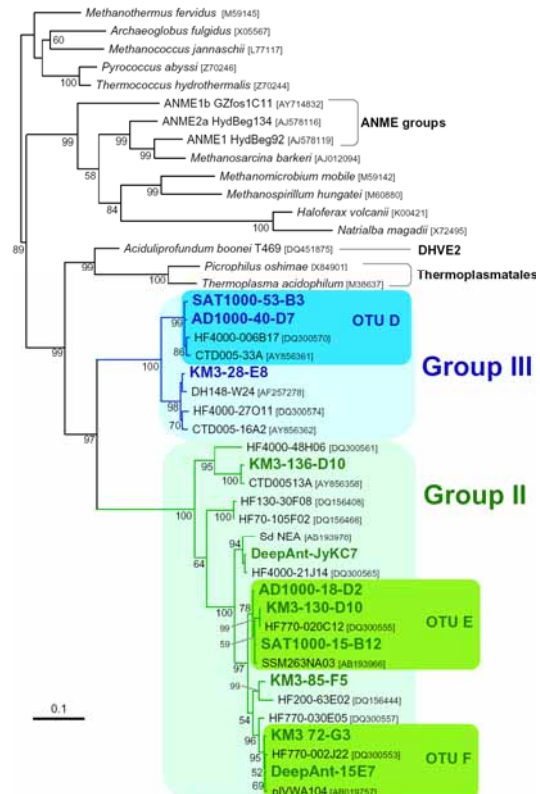
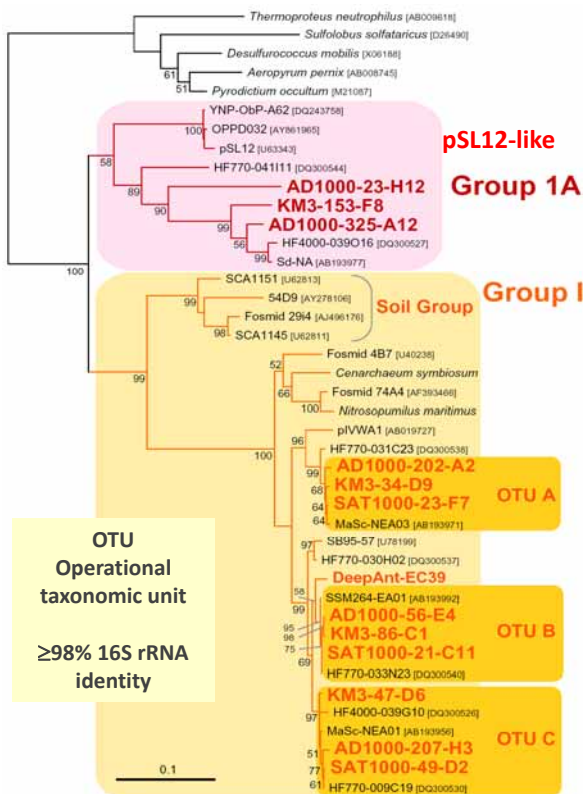
Average depth 2000 m
Warm at depth, ca. 13-14°C



Thaumarchaeota

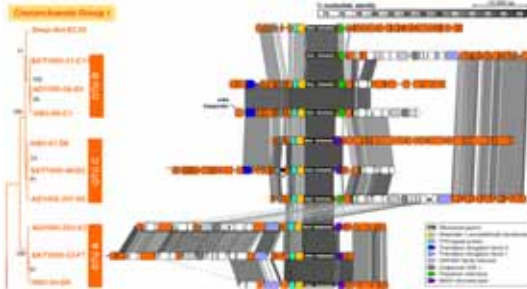
16S rRNA genes in fosmids

Euryarchaeota



HGT from in deep-sea archaea?

Thaumarchaeota (low GC 33.8%)

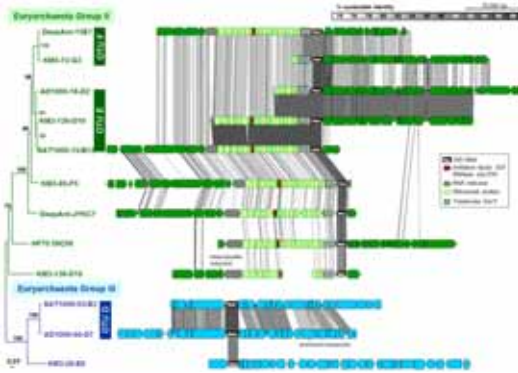


Molecular phylogenetic analyses (ML)

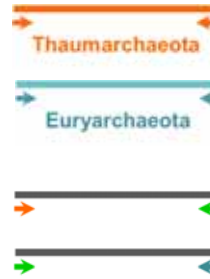
Genes in 21 fosmids
(200 ORFs Thaumarchaeota
90 ORFs GII/III Euryarchaeota)

Genes in fosmid ends
(12,774 fosmids from 3 metagenomic libraries)

Euryarchaeota (high GC 58.7%)



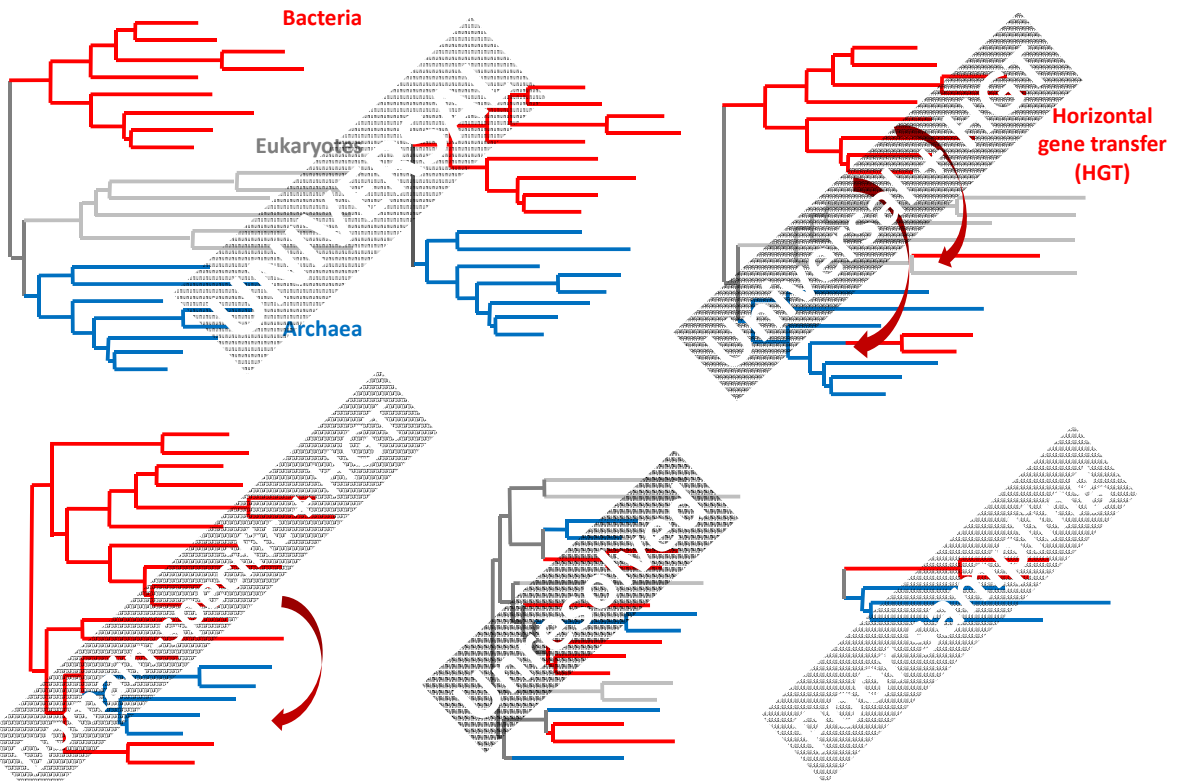
Martin-Cuadrado et al., ISME J (2008)



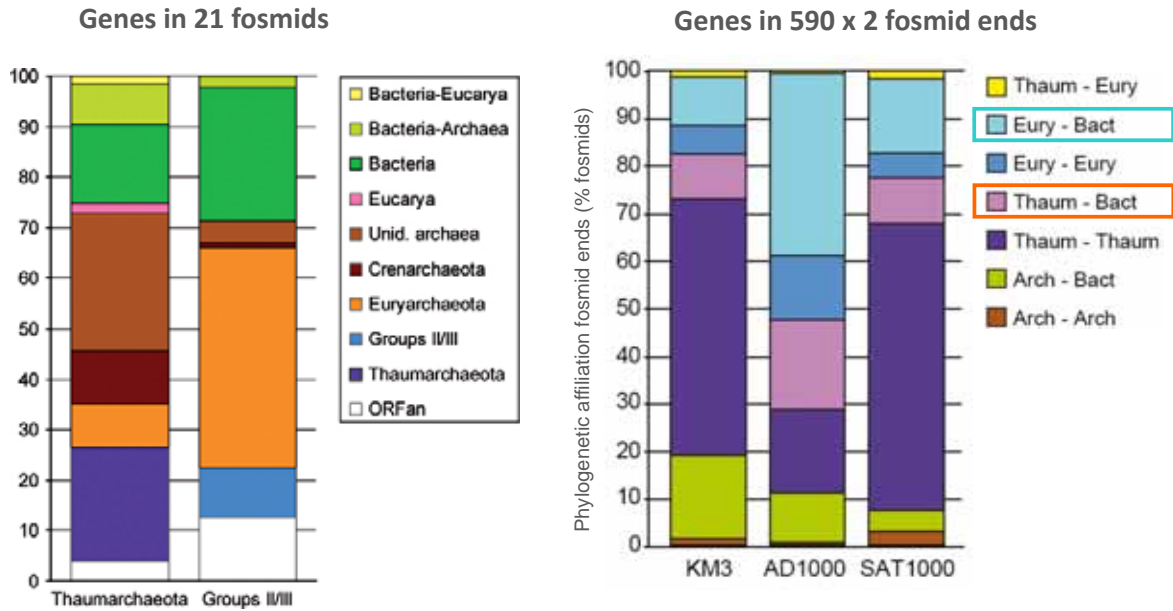
590 fosmids
dual affinity

e.g. bacterial-like
putative HGT case

Interpreting phylogenetic tree topology



HGT from distant donors in deep-sea archaea?



Between **20 and 30%** HGT events from distant donors

10-20 % Thaumarchaeota - Bacteria
10-38% Euryarchaeota - Bacteria

Brochier-Armanet, et al. ISME J 2011

Horizontal gene transfer from distant donors in marine archaea

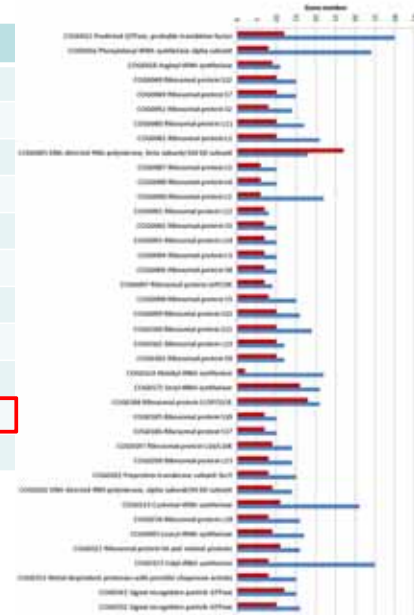
- Real?
- Genome-wide?
- Ancient, recent, ongoing?
- Which genes are transferred?

Getting acces to genomes of uncultured archaea

997 deep Mediterranean archaeal fosmid sequenced & analyzed

40 genes usually in single copy

	Thaumarchaeota	Euryarchaeota GII/III
Number of fosmids	545	452
Total sequence (bp)	19,717,229	16,310,525
Mean fosmid length (bp)	36,178	36,085
GC content (%)	47.13	54.82
rRNAs (5S, 16S, 23S)	42	28
tRNAs	610	489
Mean of 40 single-copy genes	16.5	9.3
ORFs ≥ 90 nt	150,170	164,605
Number of annotated genes	23,665	13,227
Orthologous gene clusters (OGs)	2,098	3,527
Orphan genes	416	1,293



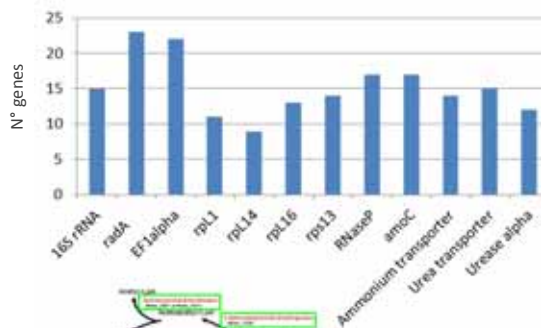
~ 16 Thaumarchaeota genome equivalents

~ 9 GII/III Euryarchaeota genome equivalents

■ Thaumarchaeota
■ Euryarchaeota

Functional potential of deep-sea archaea

Thaumarchaeota



→ Ammonia oxidizing

→ Urea metabolizing

→ Potential autotrophic C fixation



Proposed 3-hydroxypropionate/4-hydroxybutyrate cycle for autotrophic carbon fixation by *N. maritimus*

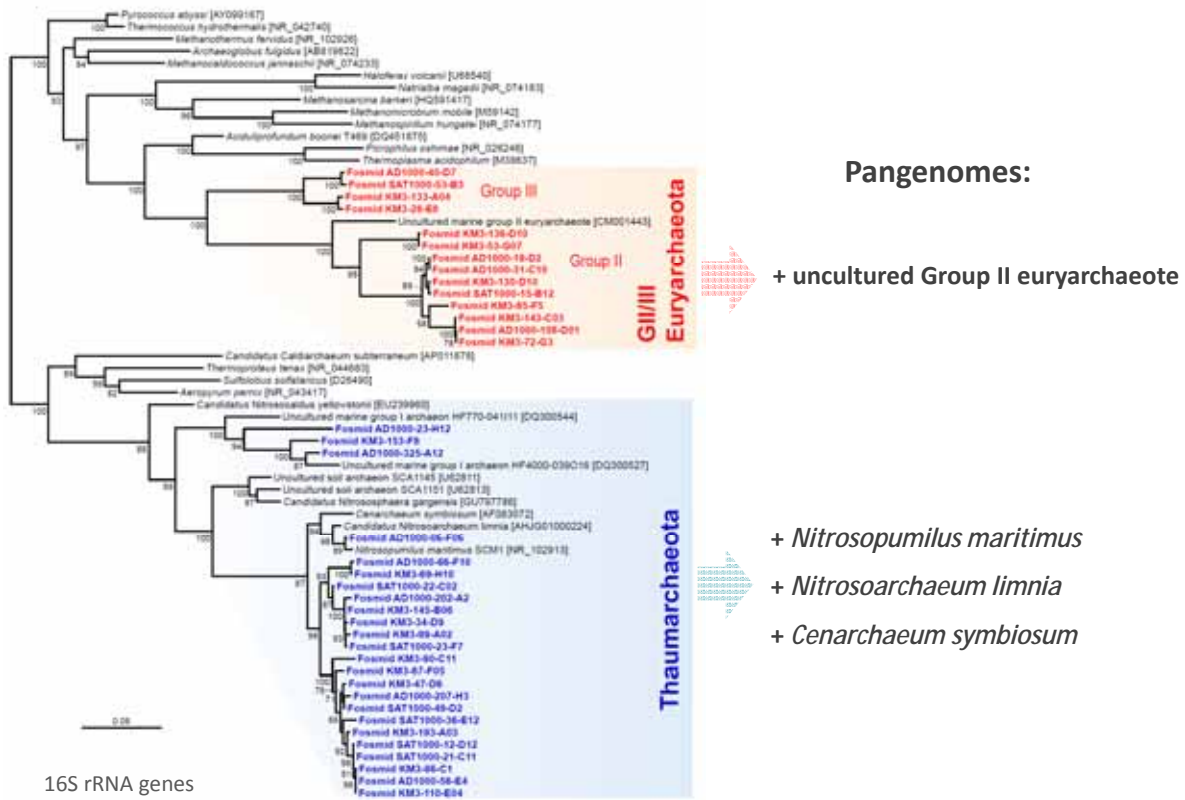
GII/III Euryarchaeota

No proteorhodopsin homologs

→ genuine deep-sea dwellers

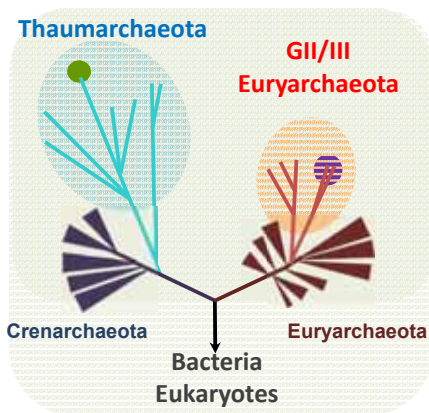
→ likely versatile heterotrophs

Thaumarchaeota and GII/III Euryarchaeota pangenomes



Gene classes ↔ Phylogenetic analyses

* Orthologous gene clusters (OGs)



2,098 genes (OGs) - Thaumarchaeota

3,527 genes (OGs) GII/III Euryarchaeota

+ local database – genomes:
 120 archaea, 297 bacteria, 120 eukaryotes

Shell

Lineage-specific core

Core

Automated

ML phylogenetic trees

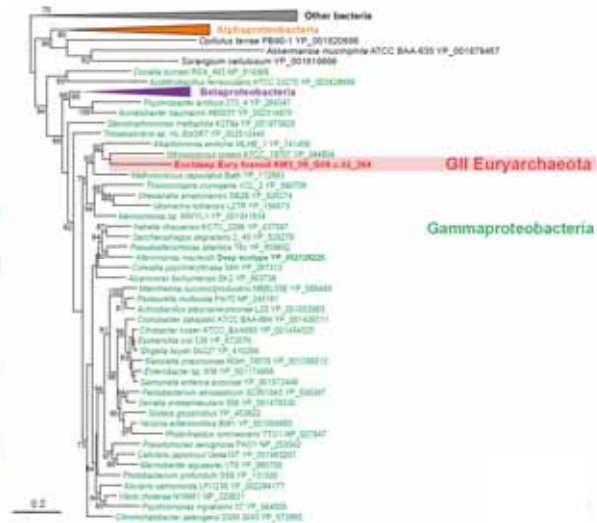
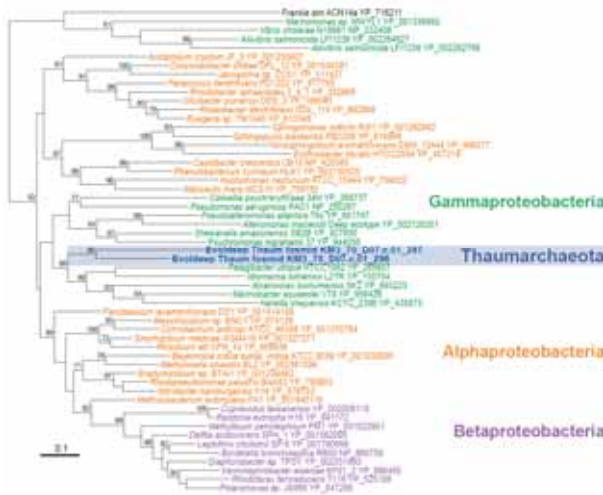
Manual !!!

Phylogenetic origin

Horizontally transferred genes (HT-genes) in 'shell' class

To Thaumarchaeota

To GII/III Euryarchaeota

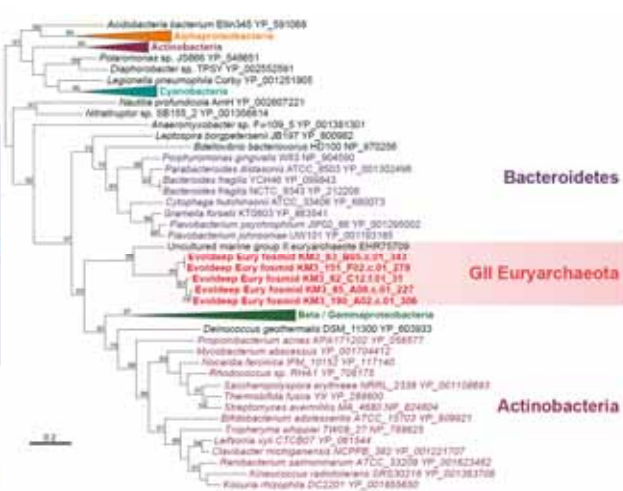
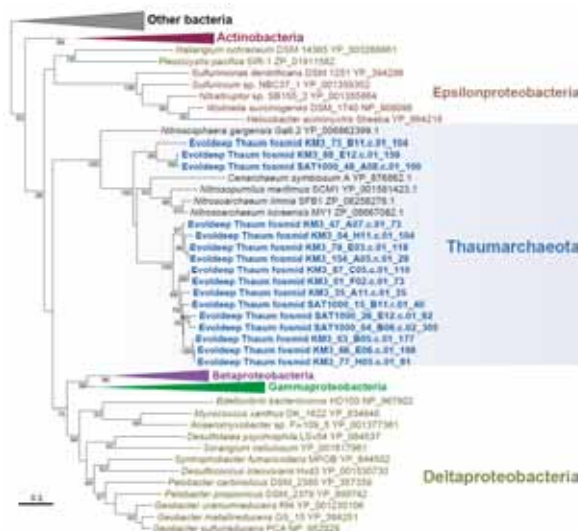


→ Late HGT events (recent import)

HT-genes in 'lineage-specific core' class

To Thaumarchaeota

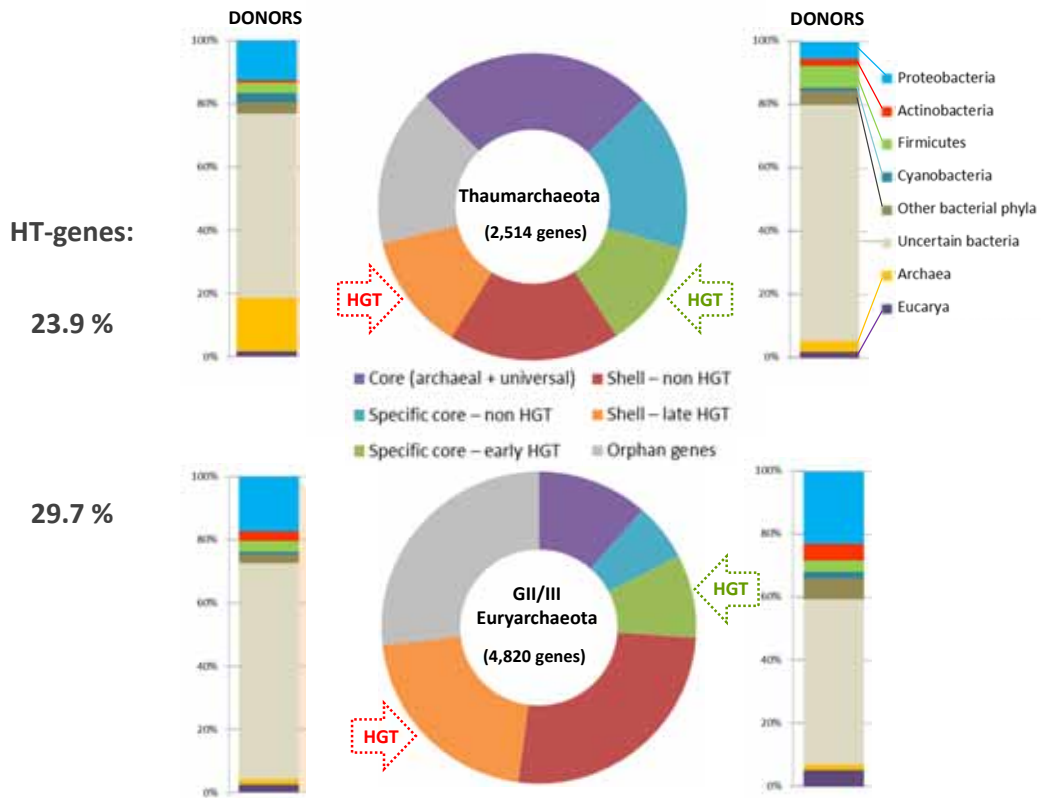
To GII/III-Euryarchaeota



→ Ancient HGT events

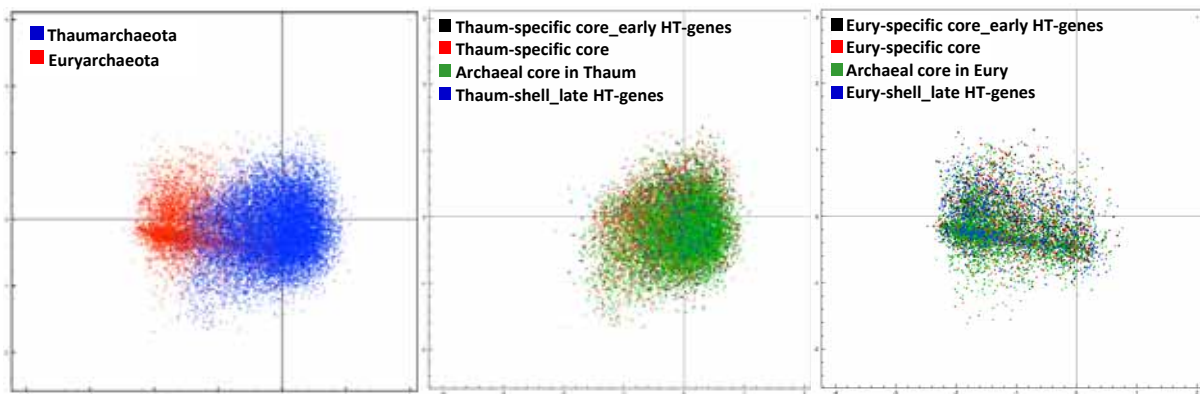
196 out of 290 (67%) early HT-genes also present in soil Thaumarchaeota

Archaeal pangenomes: gene classes and HGT donors



Codon usage in different gene classes

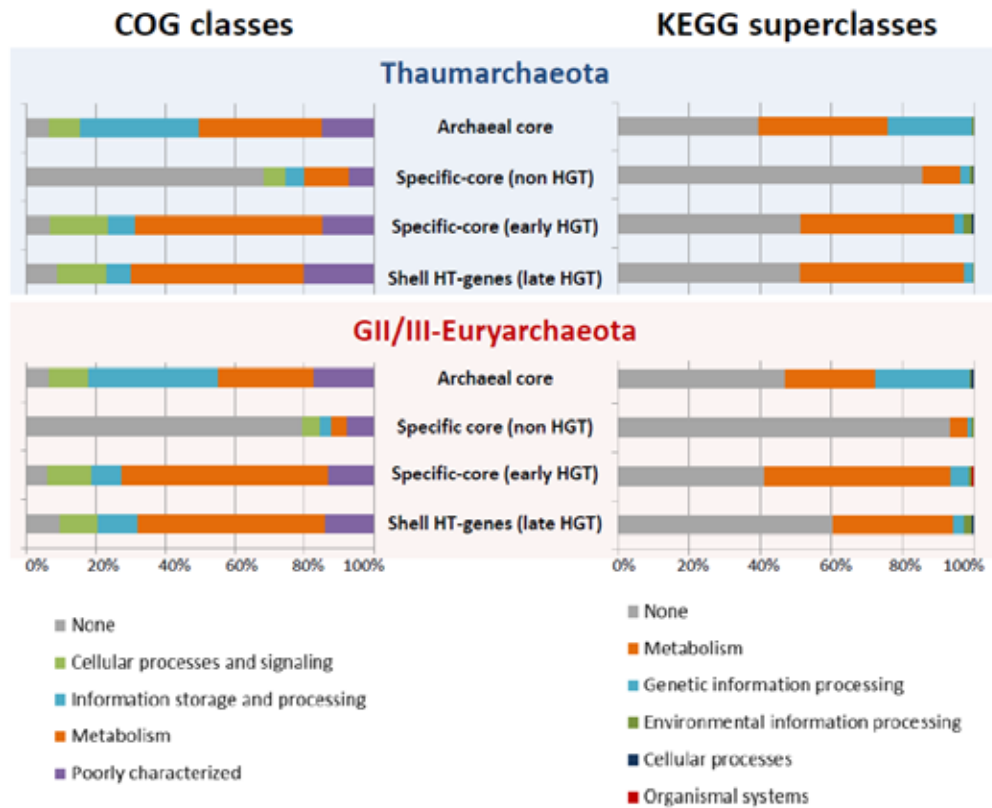
PCA



(GC content bias)

No difference of codon usage among gene classes:
genes adapted to their genomic environment

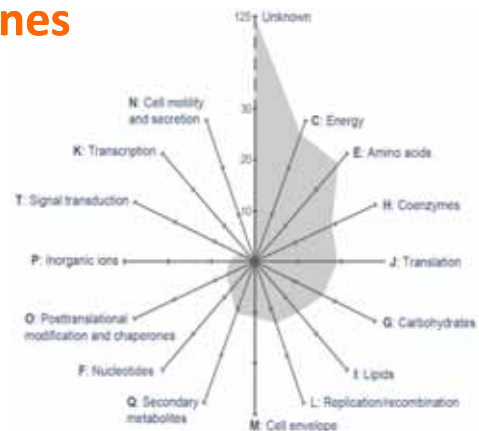
Functional categories



Functions of imported genes

Functional categories (COGs)

- Similarities Thaum/Eury: Nucleotide, coenzyme, carbohydrate, lipid and amino acid **transport** and metabolism, inorganic ion transport, **energy production and conversion**, and **cell wall/membrane biogenesis**
- Differences Thaum/Eury: GII/III more HT-genes affecting transcription/signal transduction and posttranslational modifications



Pathways (KEGGs)

- Similarities Thaum/Eury: Benzoate degradation, phenylalanine metabolism, folate biosynthesis, fatty acid biosynthesis, ABC transporters, oxidative phosphorylation or the metabolism of glyoxylate and dicarboxylate, cysteine, methionine and cofactors
- Differences: Thaum: more HT-genes involved in **sugar metabolism**; GII/III-Eury: more genes in **amino acid and nucleotide metabolism**, streptomycin biosynthesis

→ Role in adaptation to novel environment

At the 'Microbial diversity and evolution' team:



David Moreira

(ESE, CNRS/Univ. Paris-Sud)



Philippe Deschamps

Collaborators:



Yvan Zivanovic

(IGM, CNRS/Univ. Paris-Sud)



**Francisco
Rodriguez-Valera**

(UMH, Spain)



**Céline Brochier-
Armanet**

(Université Lyon1)

<http://www.ese.u-psud.fr/rubrique7.html?lang=en>