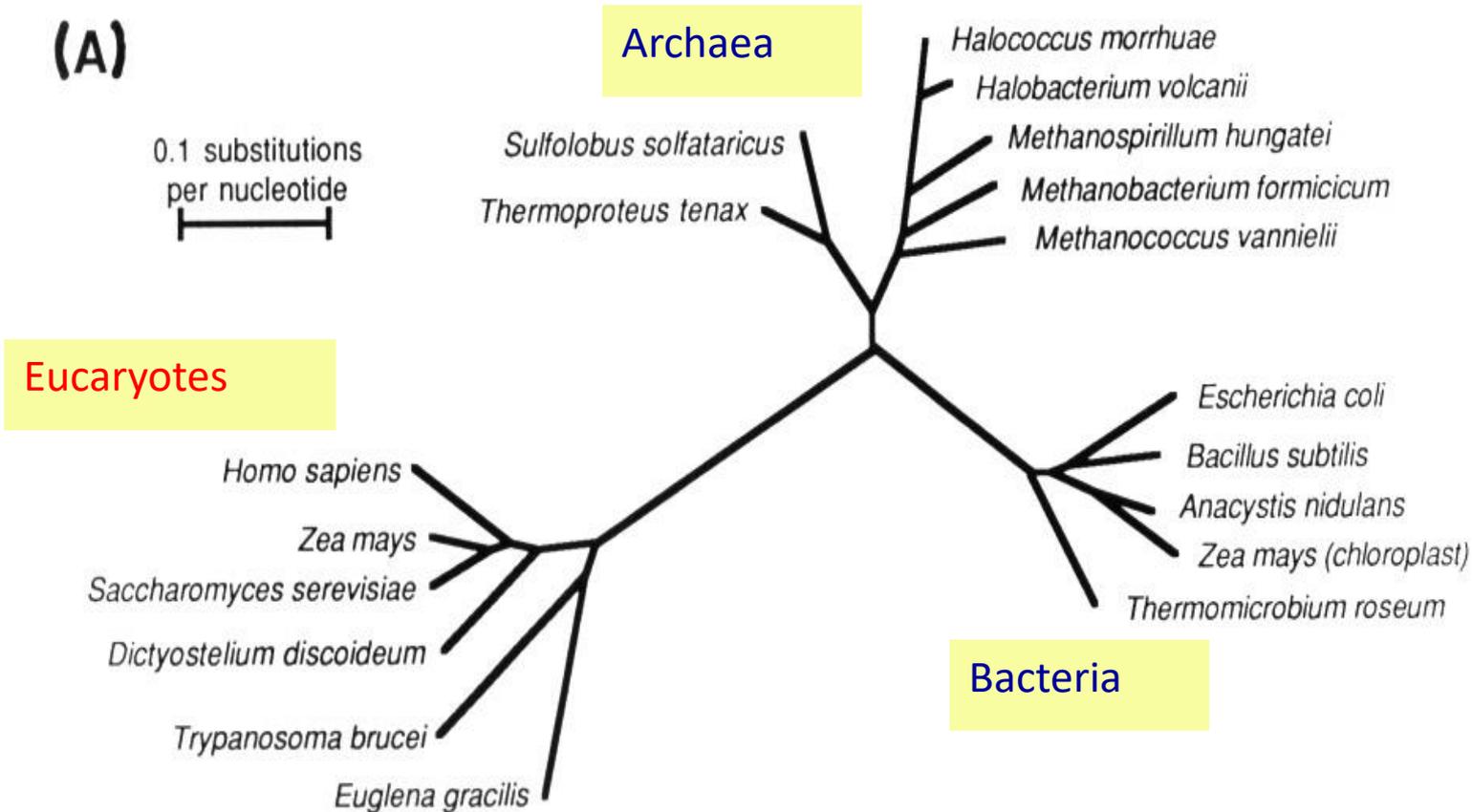


# L'arbre universel du vivant

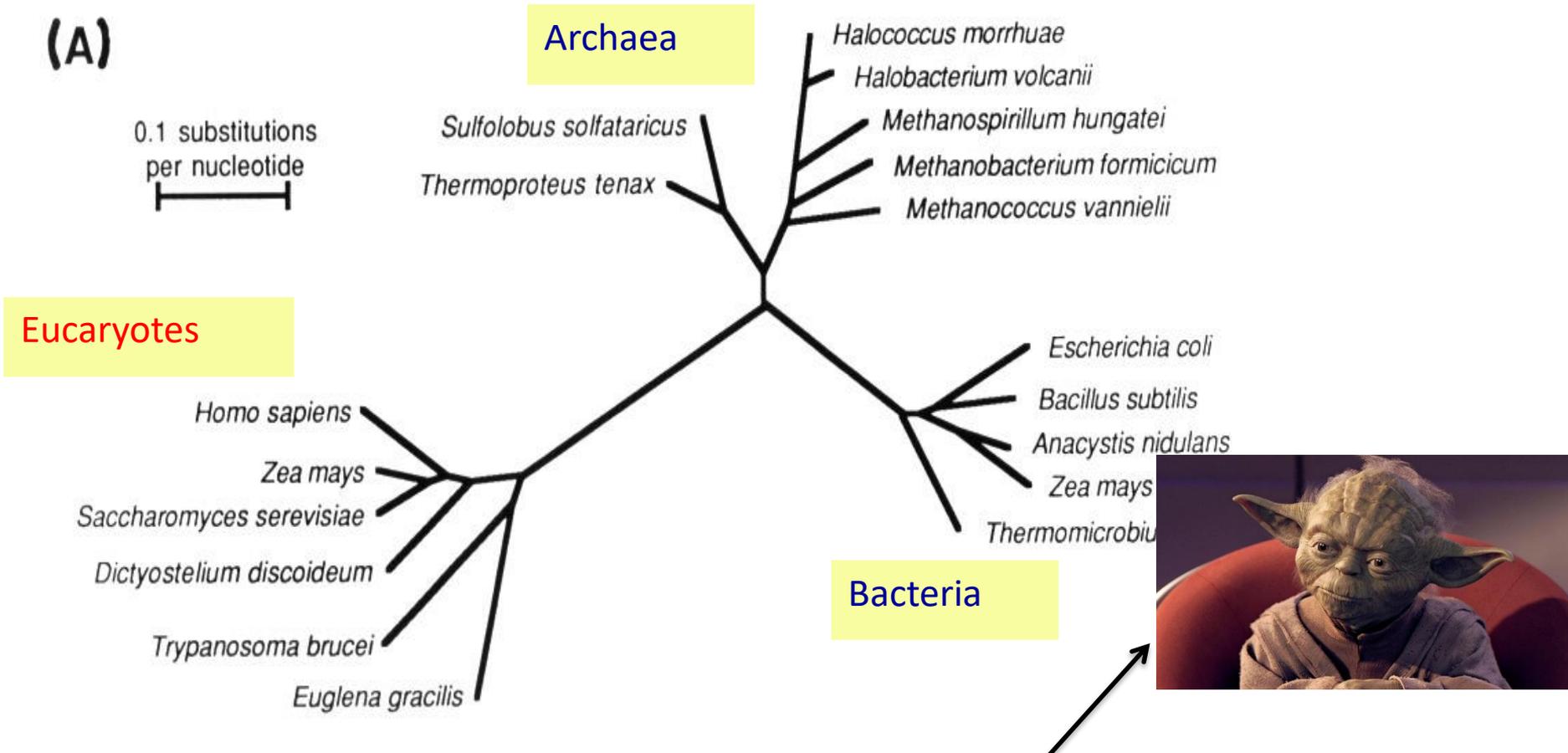
Où est la racine ?



→ pour enracer un arbre il faut un groupe de séquences externe au groupe étudié...

# L'arbre universel du vivant

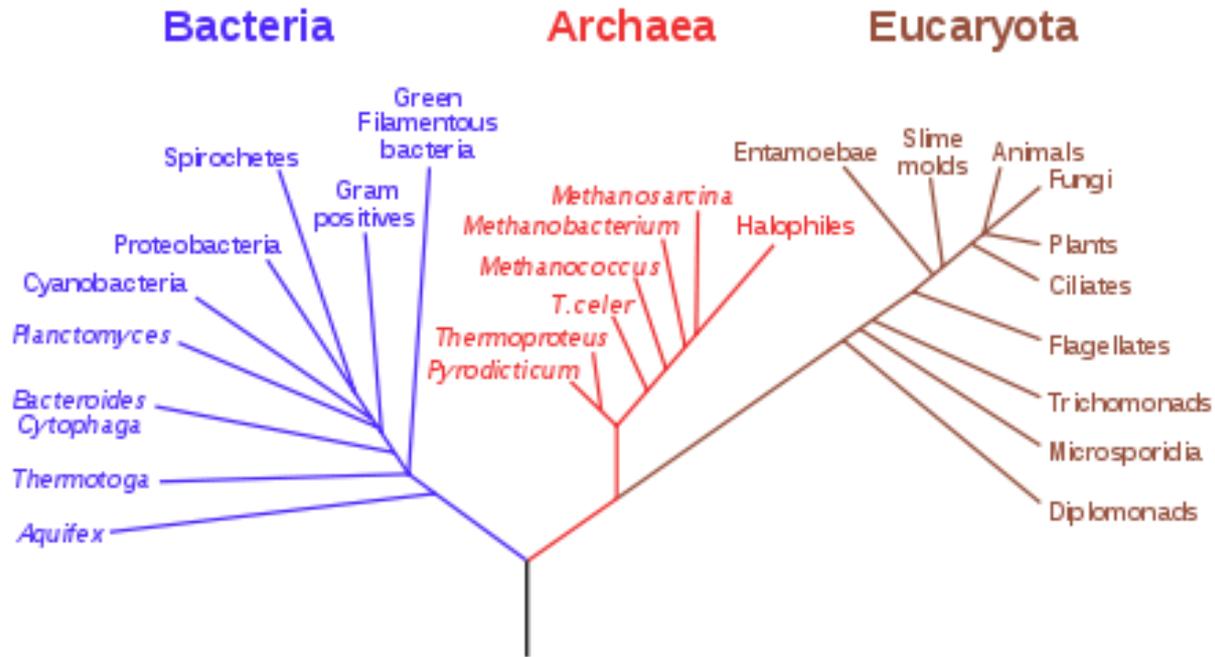
Où est la racine ?



→ pour enracer un arbre il faut un groupe de séquences externe au groupe étudié...

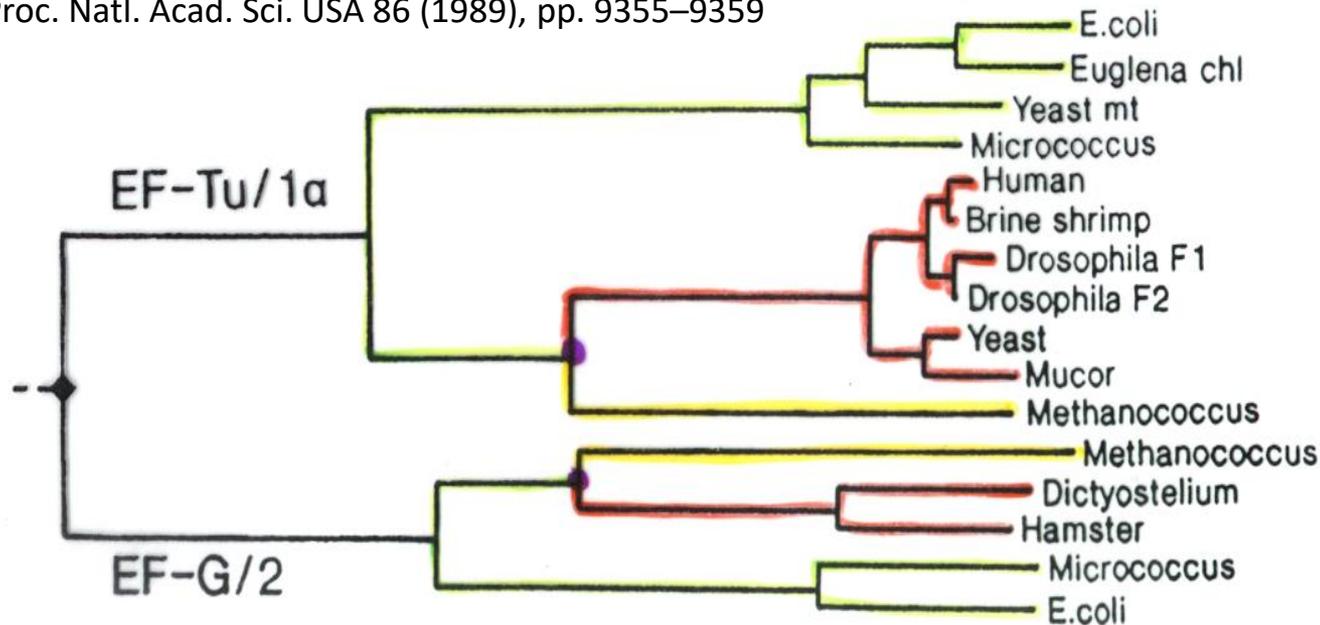
# L'arbre universel du vivant

L'arbre de Woese et Fox, 1990, est enraciné sur la branche des bactéries:



# L'arbre universel du vivant

Iwabe et al., Proc. Natl. Acad. Sci. USA 86 (1989), pp. 9355–9359



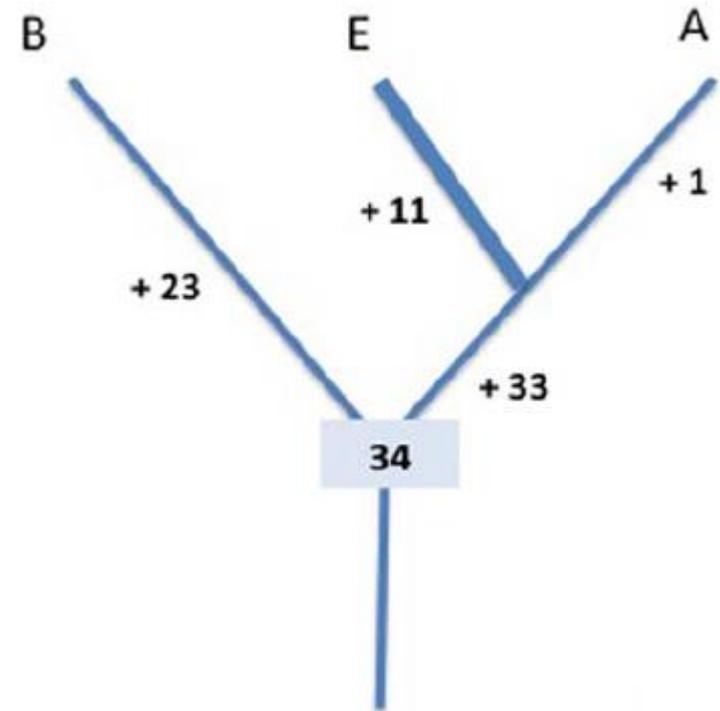
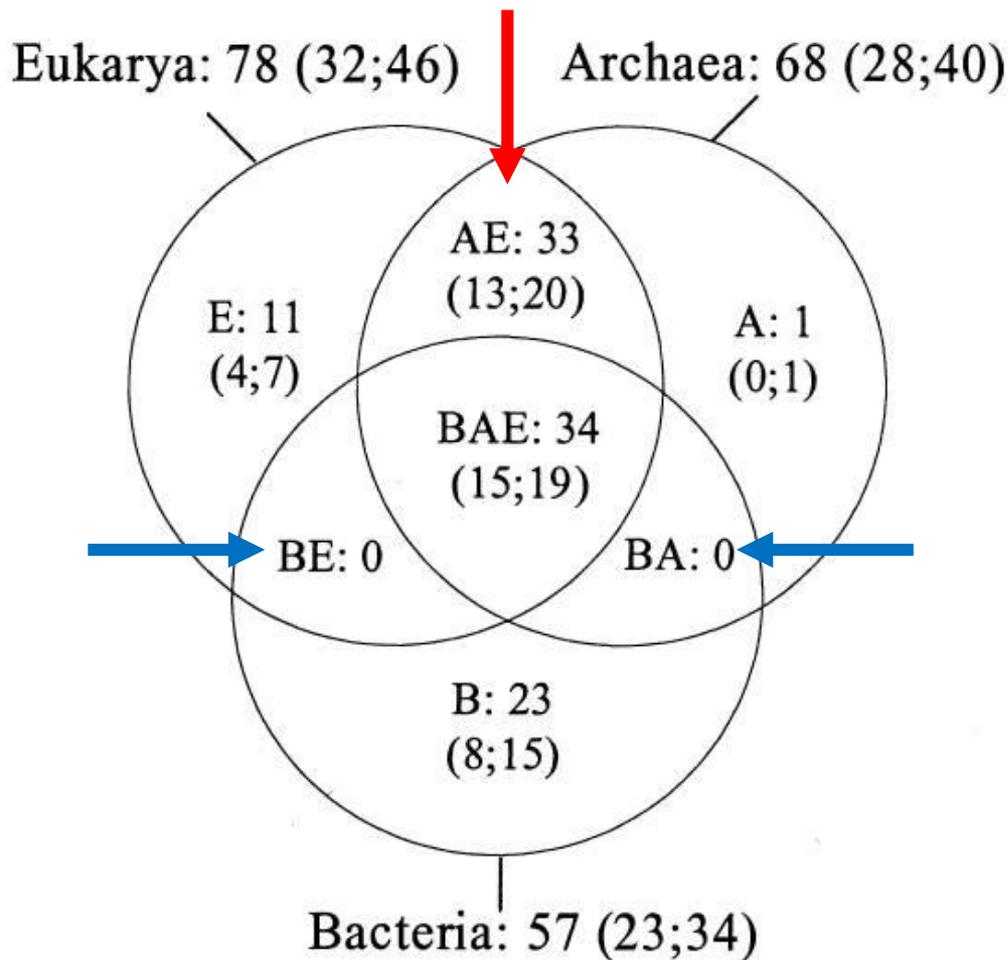
-l'arbre universel peut être enraciné en utilisant une paire des protéines *paralogues* et *universelles*

-exemples: les facteurs d'élongation EF-Tu et EF-G ou les sous-unités  $\alpha$  et  $\beta$  de l'ATPase F1

-la duplication est survenue avant la divergence des trois domaines – par conséquent la racine peut être placée à l'endroit/moment qui correspond à leur duplication

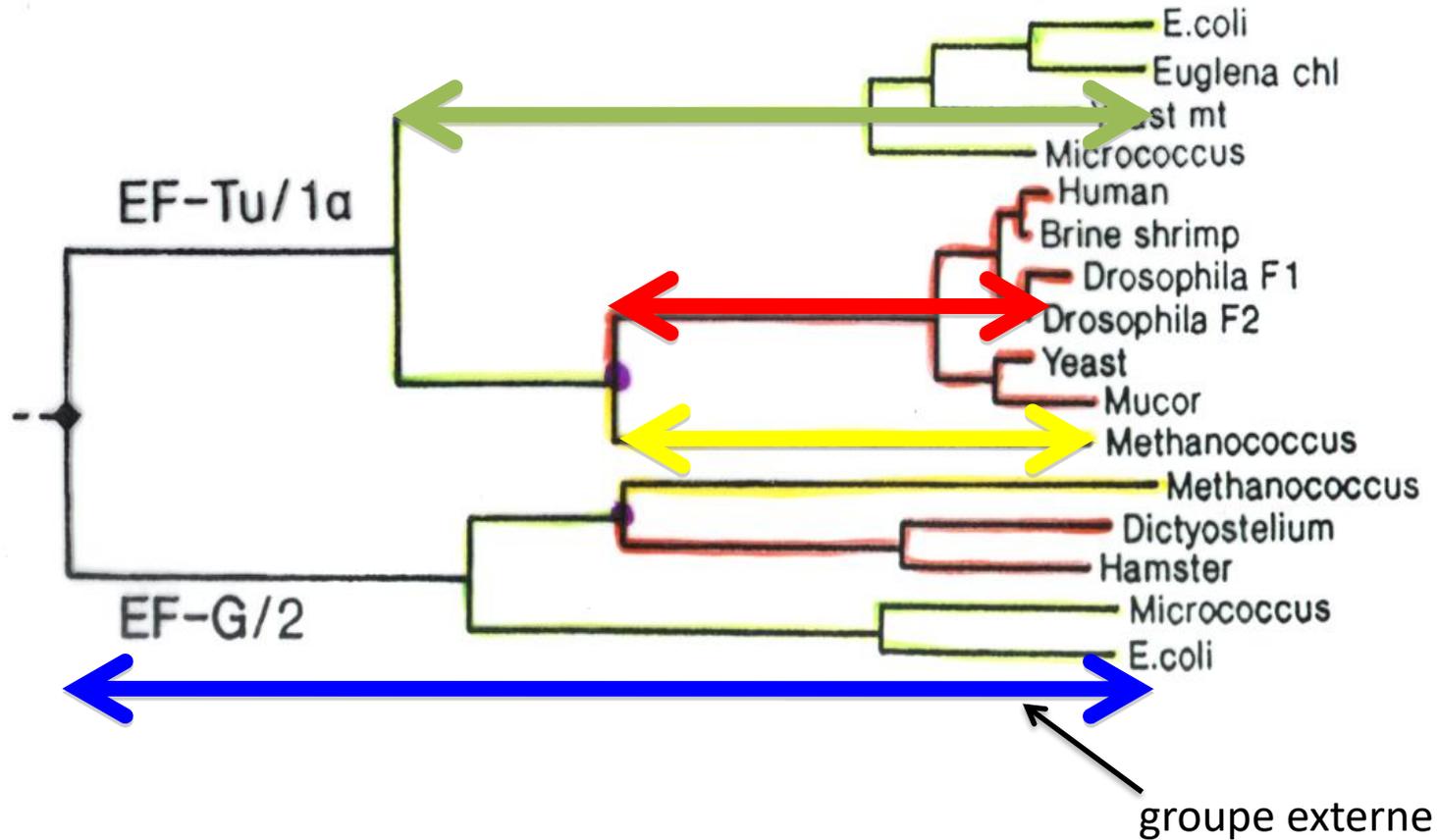
# L'arbre universel du vivant

Distribution des protéines ribosomales entre les trois domaines...



soutien pour la racine sur la branche des bactéries

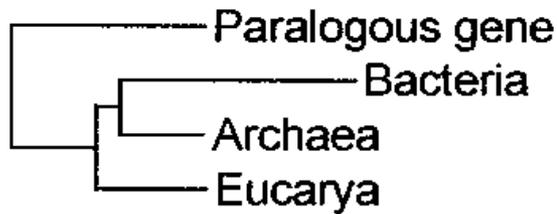
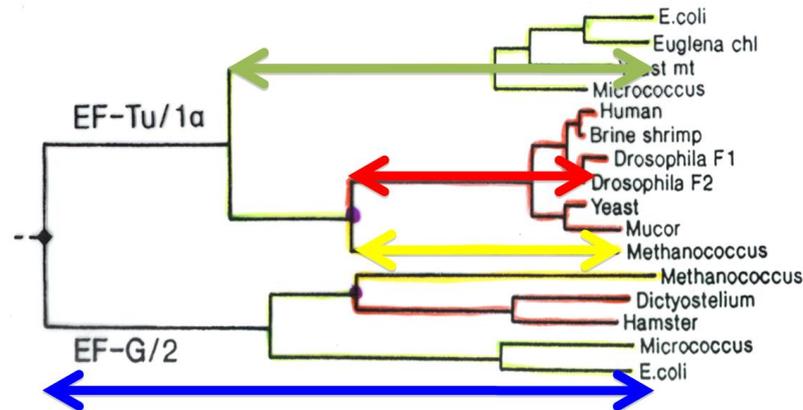
# L'arbre universel du vivant



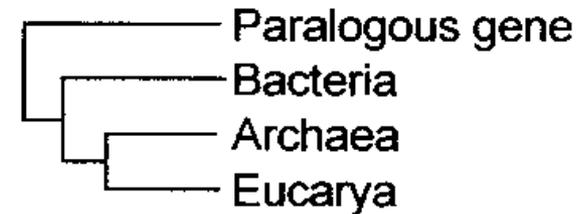
-les branches des bactéries et du groupe externe sont les plus longues →

# L'arbre universel du vivant

→ long branch attraction artefact



⇒  
**LBA**



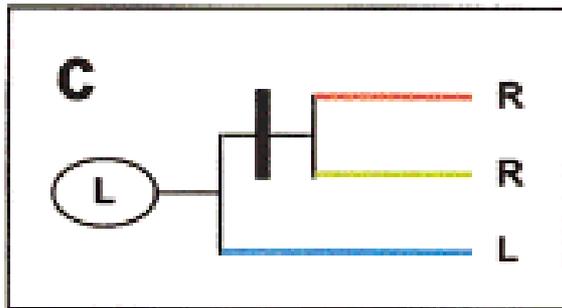
# L'arbre universel du vivant

## Where is the root of the universal tree of life?

Patrick Forterre<sup>1\*</sup> and Hervé Philippe<sup>2</sup>

→ dans les alignements des protéines universelles très peu de positions sont informatives

→ on doit pouvoir polariser un changement d'AA pour que cela soit informatif

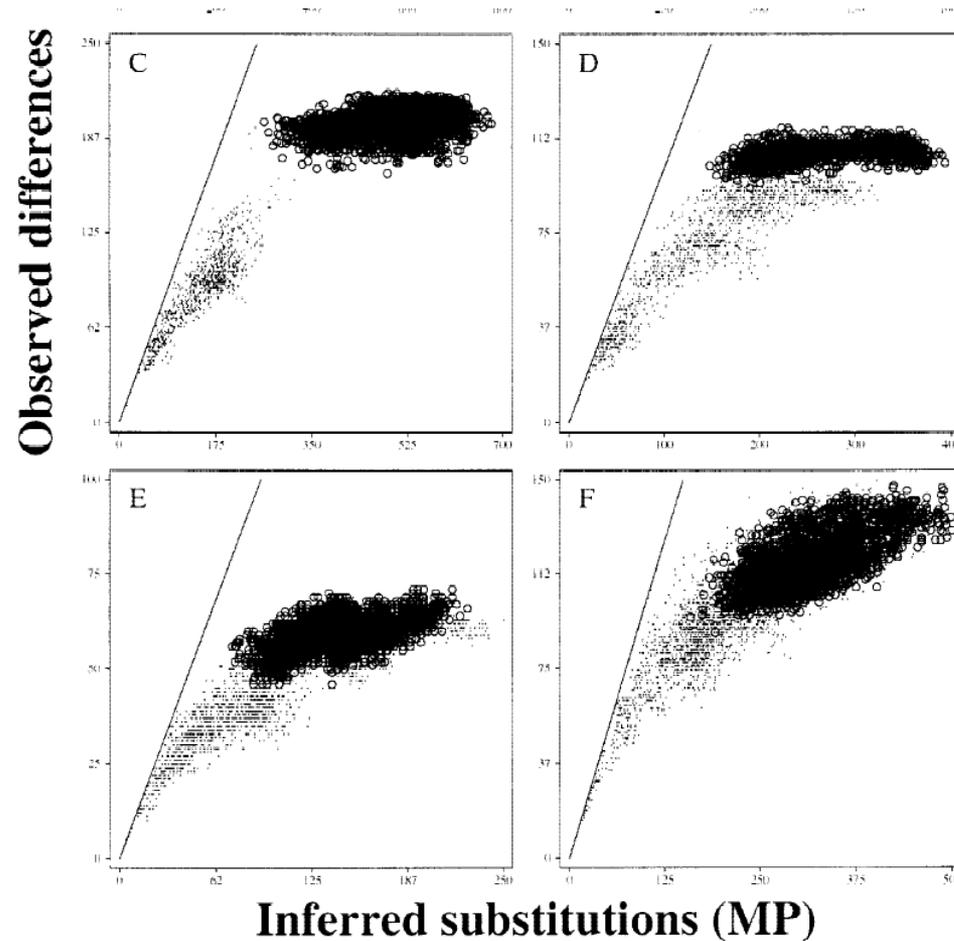


→ le groupe externe contient le caractère L = plésiomorphe  
le caractère R = synapomorphie

→ Pour les protéines universelles les synapomorphies sont très rares (car les positions qui évoluent lentement sont rares)

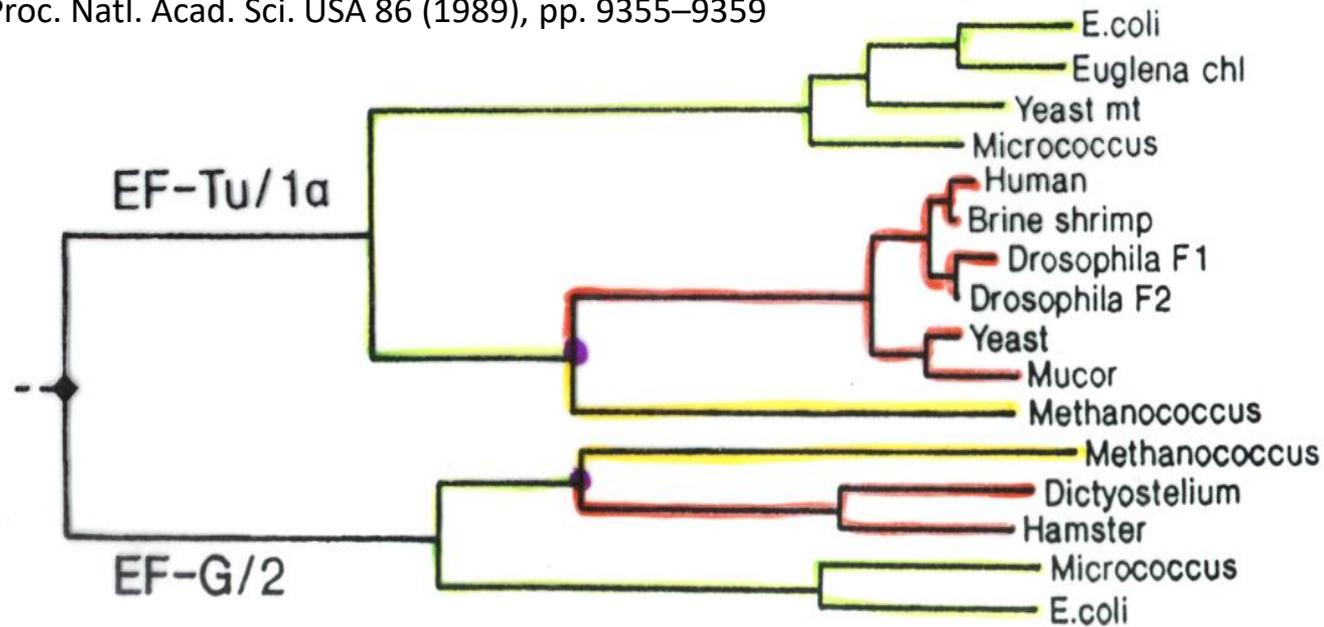
# L'arbre universel du vivant

→ analyse de saturation des gènes marqueurs (comparaison des paires de séquences):



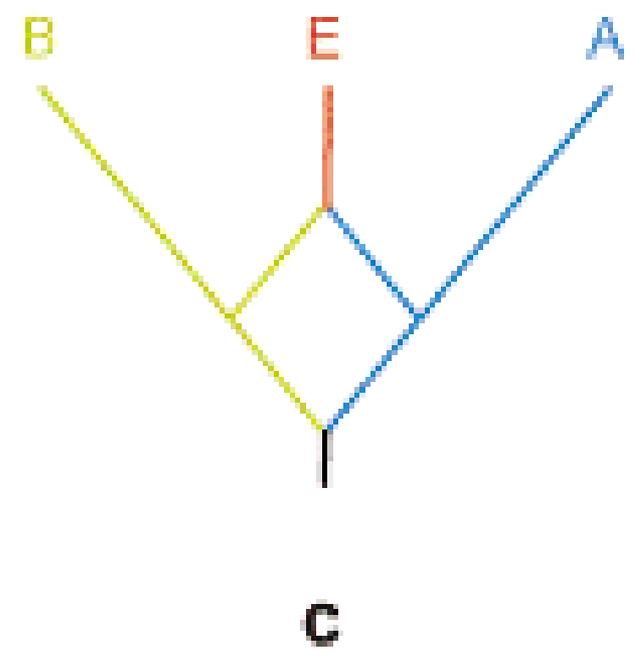
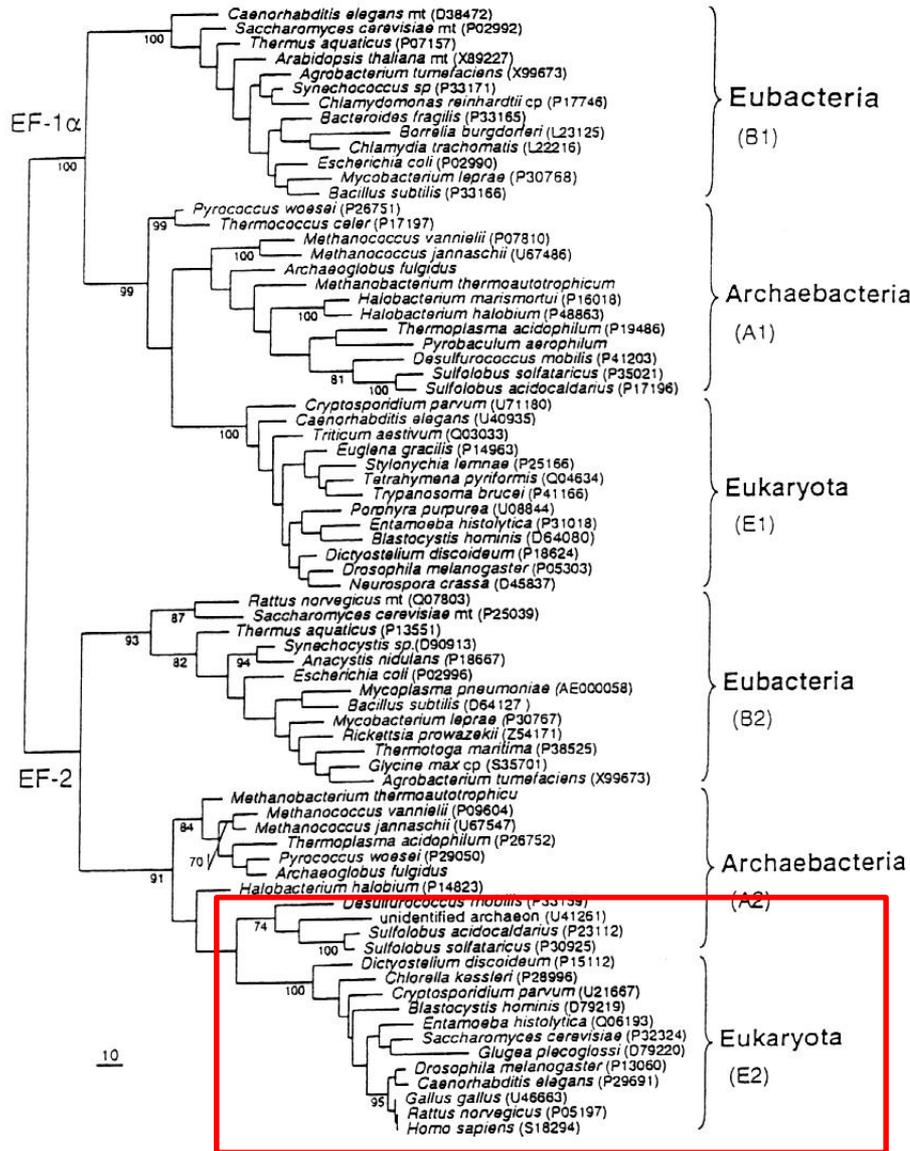
# L'arbre universel du vivant

Iwabe et al., Proc. Natl. Acad. Sci. USA 86 (1989), pp. 9355–9359



-en 1989 peu de séquences sont disponibles

# L'arbre universel du vivant

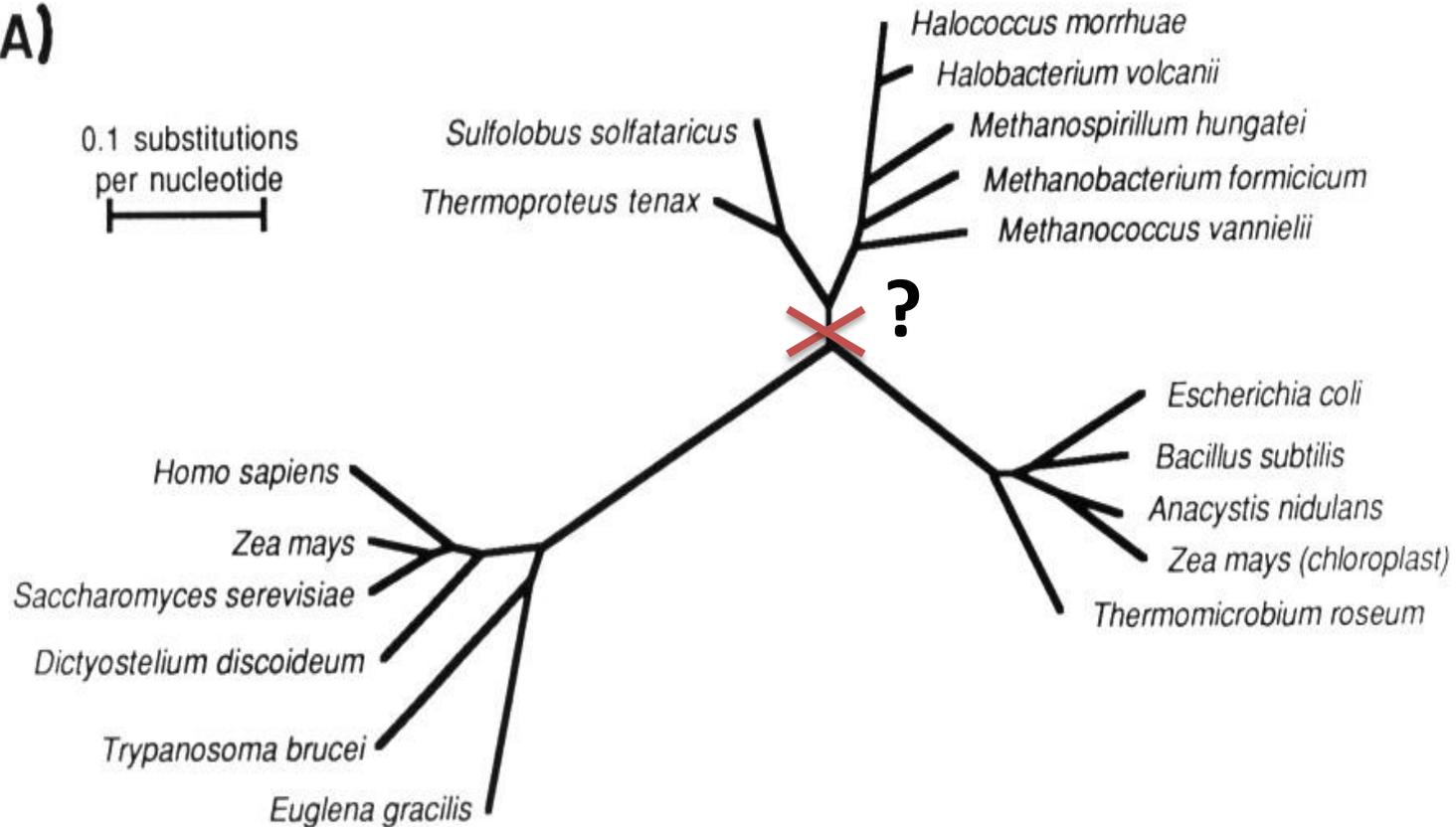


→ le nombre des séquences augmente: les topologies alternatives à celle de C. Woese émergent....

# L'arbre universel du vivant

1985, Olsen and Woese

(A)



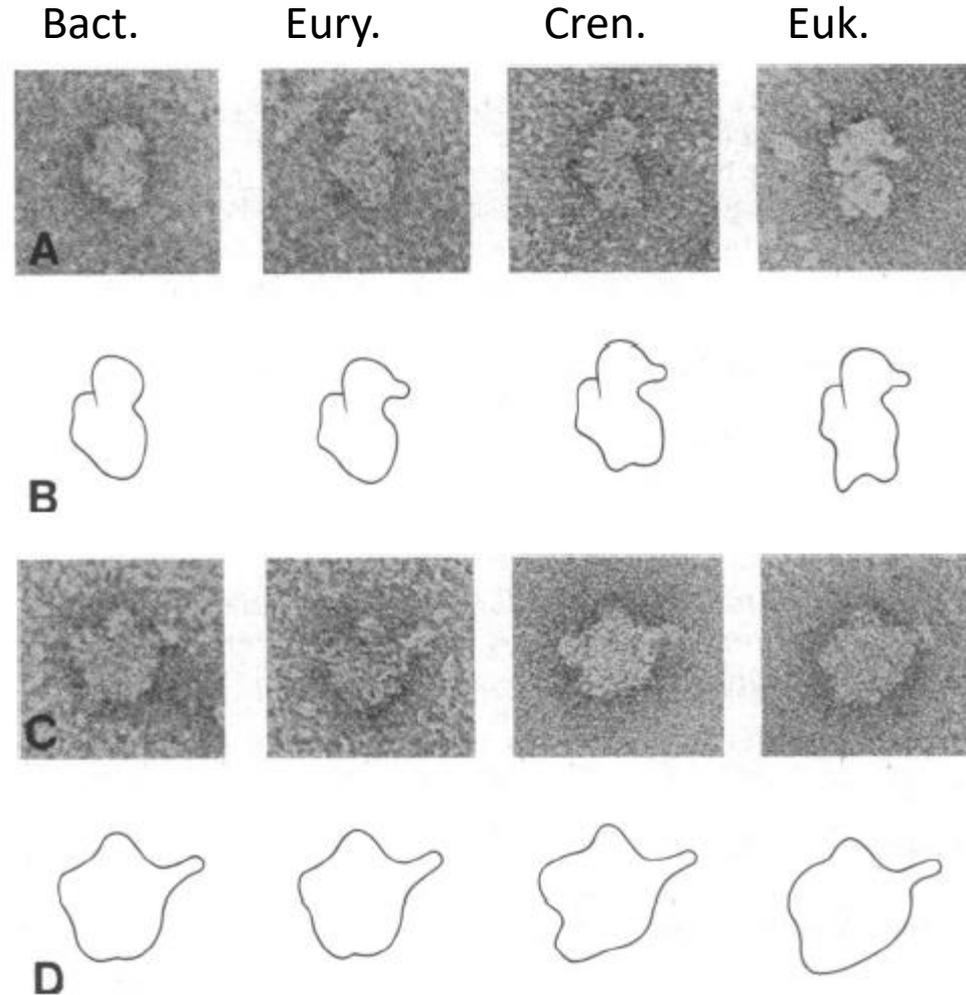
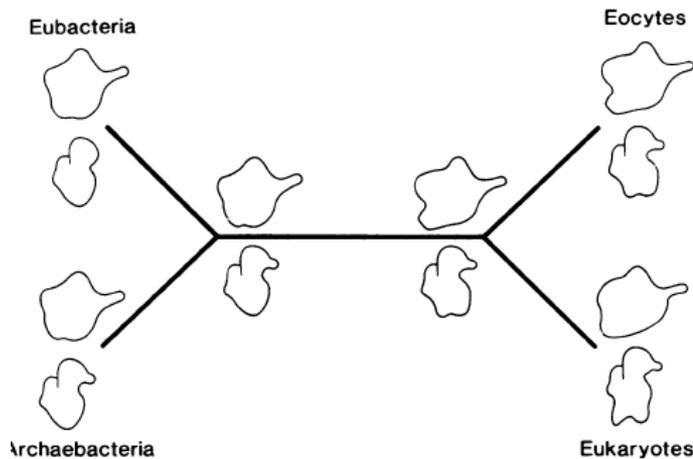
# L'arbre universel du vivant

Jim Lakes, 1984

→ Le scénario « Eocyte »

→ quatre (et non pas trois) types de ribosomes

→ crenarchaeota = eocytes



→ les crenarchaeota sont le groupe sœur des eukaryotes alors que les euryarchaea (halophiles et méthanogènes) sont groupe sœur des bactéries

# L'arbre universel du vivant

## Evidence That Eukaryotes and Eocyte Prokaryotes Are Immediate Relatives 1992, science

Maria C. Rivera and James A. Lake

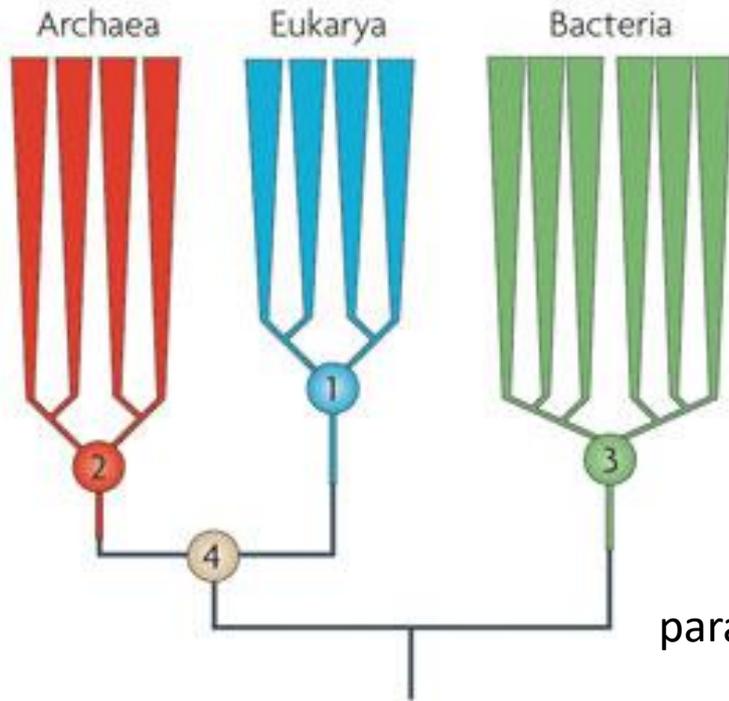
→ facteur d'élongation EF1-alpha (protéine universelle)

Taxon	Organism	Left primer	11-amino acid segment	4-amino acid segment	Right primer
Eukaryotes	Human	KNMITG	TSQADCAVLIVAAGV	GEFEAGISKNG	QTREH
Eukaryotes	Tomato	KNMITG	TSQADCAVLIIDSTT	GGFEAGISKDG	QTREH
Eukaryotes	Yeast	KNMITG	TSQADCAILIIAGGV	GEFEAGISKDG	QTREH
Eocytes	<i>P. occu.</i>	KNMITG	ASQADAAILVVSARK	GEFEAGMSAEG	QTREH
Eocytes	<i>D. muco.</i>	KNMITG	ASQADAAILVVSARK	GEFEAGMSAEG	QTREH
Eocytes	<i>A. infe.</i>	KNMITG	ASQADAAI IAVSAKK	GEFEAGMSEEG	QTREH
Eocytes	<i>Su. acid.</i>	KNMITG	ASQADAAILVVSARK	GEYEAGMSAEG	QTREH
Methanogens and relatives	<i>T. celer</i>	KNMITG	ASQADAAVLVVAATD	---GVMP	QTKEH
	<i>Mc. van.</i>	KNMITG	ASQADAAVLVVNVDD	AKSGIQP	QTREH
Halobacteria	<i>H. maris.</i>	KNMITG	ASQADNAVLVVAADD	---GVQP	QTQEH
Eubacteria	<i>Th. mar.</i>	KNMITG	AAQMDGAILVVAATD	---GPMP	QTREH
Eubacteria	<i>S. plat.</i>	KNMITG	AAQMDGAILVVSAAAD	---GPMP	QTREH
Eubacteria	Mitoch.	KNMITG	AAQMDGAIIVVAATD	---GQMP	QTREH

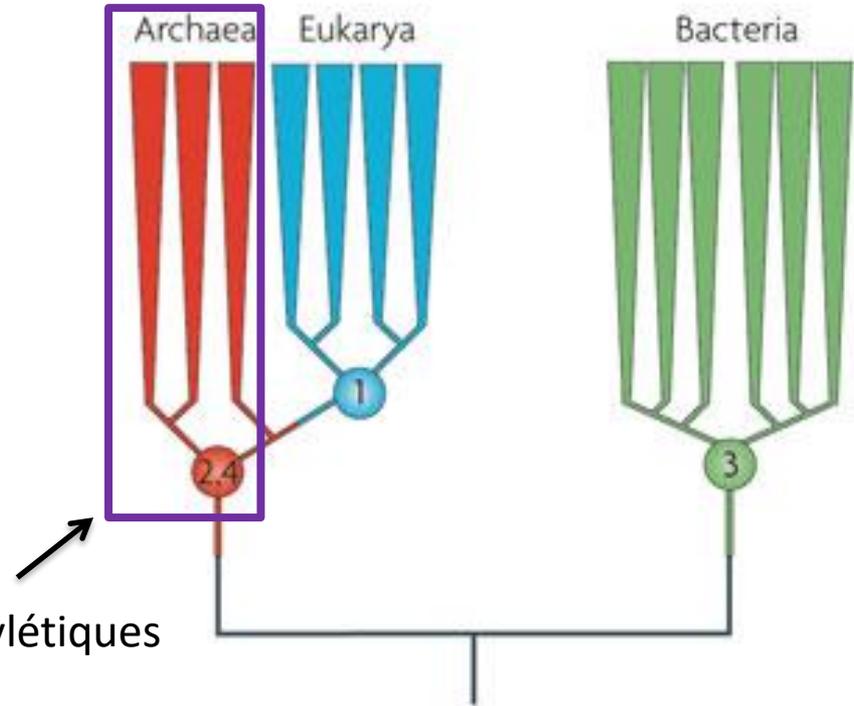
→ l'hypothèse eocyte est soutenue par un indél spécifique trouvé au sein des eucaryotes et des crenarchaeota

# L'arbre universel du vivant

a 'Three primary domains' (3D) scenario



b 'Two primary domains' (2D) scenario



→ fin de la monophylie des archées? Fin des archées tout court?!?

*"modern cells are sufficiently complex, integrated and 'individualized' that further major change in their designs does not appear possible"*

C. Woese

# L'arbre universel du vivant

→ 2008, nouvelle version de l'hypothèse « eocyte »:

→ soutenu par des reconstructions phylogénétiques dans un cadre statistique incorporant les modèles sophistiqués d'évolution des séquences dans leurs hypothèses de base

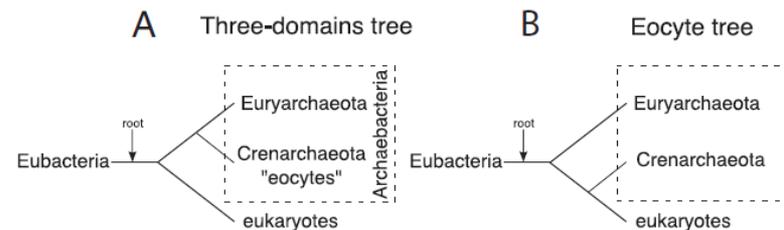
## The archaeobacterial origin of eukaryotes

Cymon J. Cox<sup>a,1</sup>, Peter G. Foster<sup>a,1</sup>, Robert P. Hirt<sup>b</sup>, Simon R. Harris<sup>b</sup>, and T. Martin Embley<sup>b,1</sup>

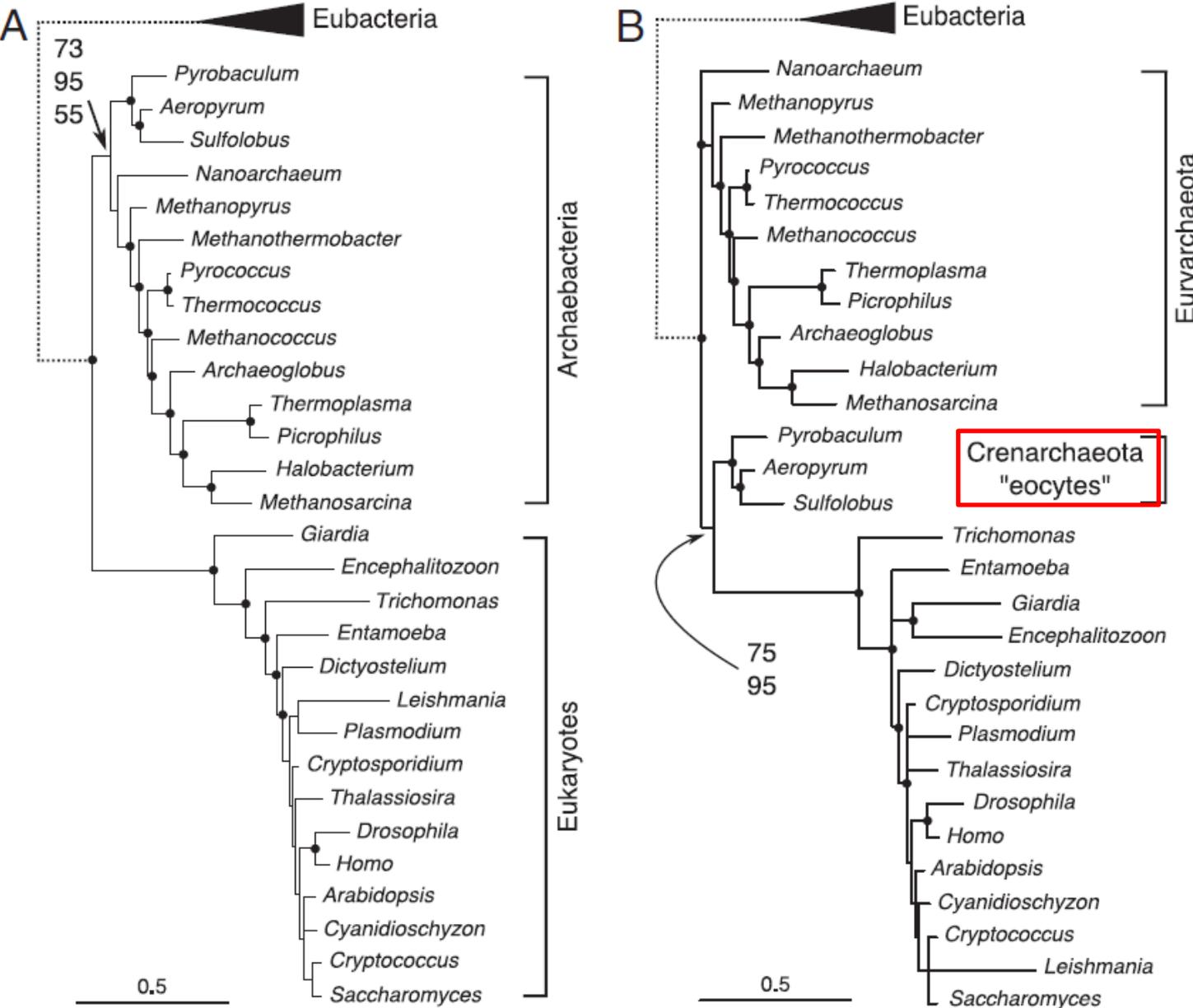
<sup>a</sup>Department of Zoology, Natural History Museum, Cromwell Road, London, SW7 5BD, United Kingdom; and <sup>b</sup>Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle, NE2 4HH, United Kingdom

Communicated by Jeffrey D. Palmer, Indiana University, Bloomington, IN, October 24, 2008 (received for review April 16, 2008)

The origin of the eukaryotic genetic apparatus is thought to be central to understanding the evolution of the eukaryotic cell. Disagreement about the source of the relevant genes has spawned competing hypotheses for the origins of the eukaryote nuclear lineage. The iconic rooted 3-domains tree of life shows eukaryotes and archaeobacteria as separate groups that share a common ancestor to the exclusion of eubacteria. By contrast, the eocyte hypothesis has eukaryotes originating within the archaeobacteria



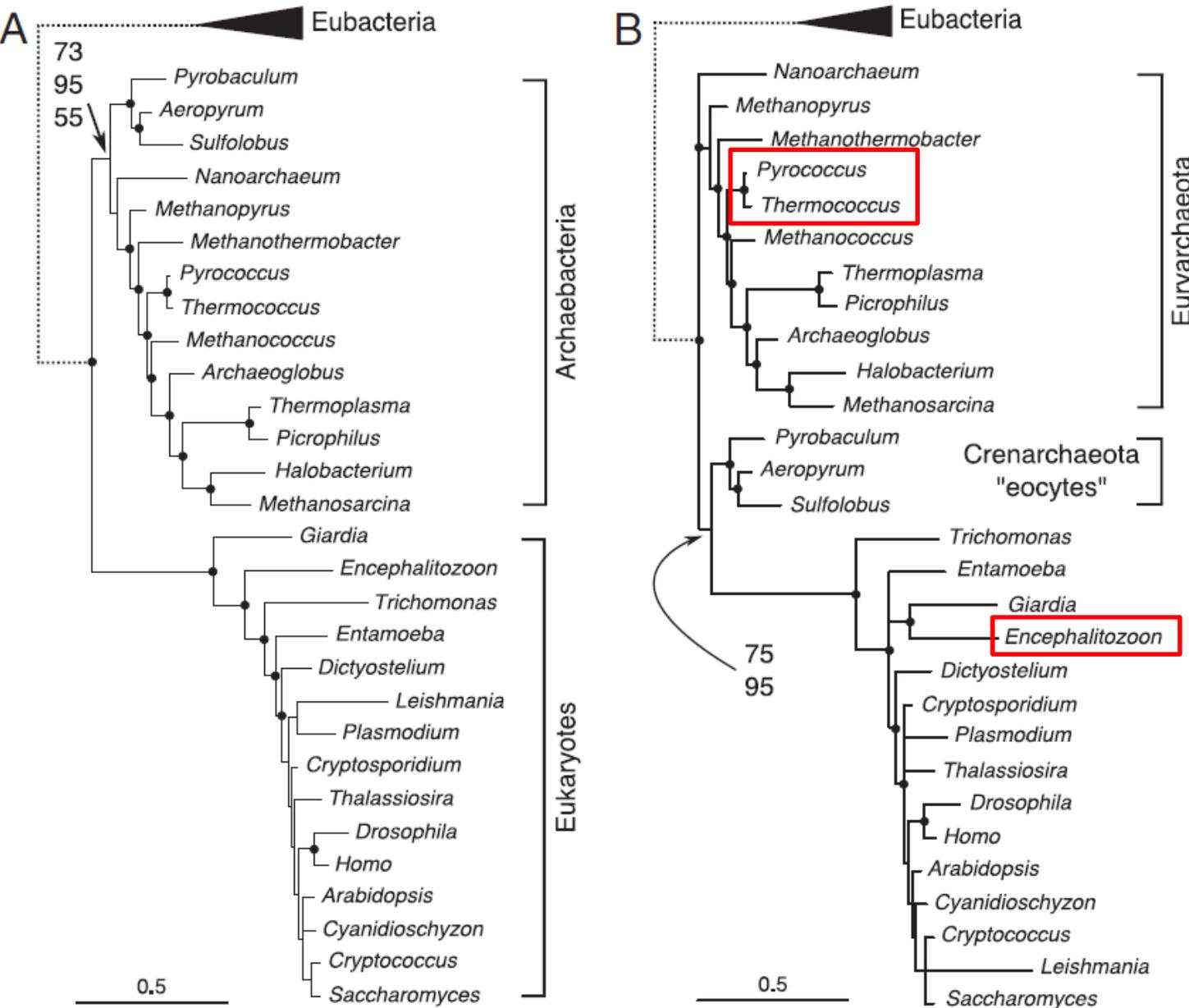
# L'arbre universel du vivant



-SSU + LSU rRNA  
A: vitesse de substitution des positions est homogène chez les différentes lignées

B: vitesse de substitution hétérogène

# L'arbre universel du vivant

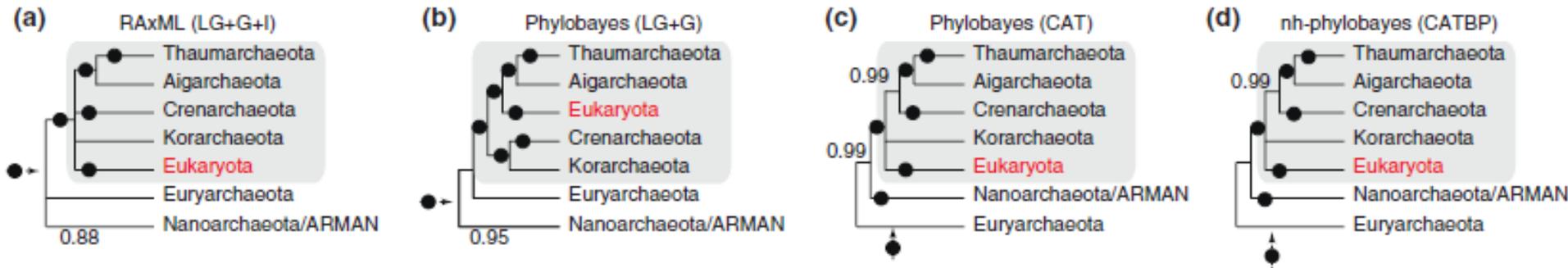


Thermococcales  
parmi les  
methanogènes?

c'est un  
champignon!

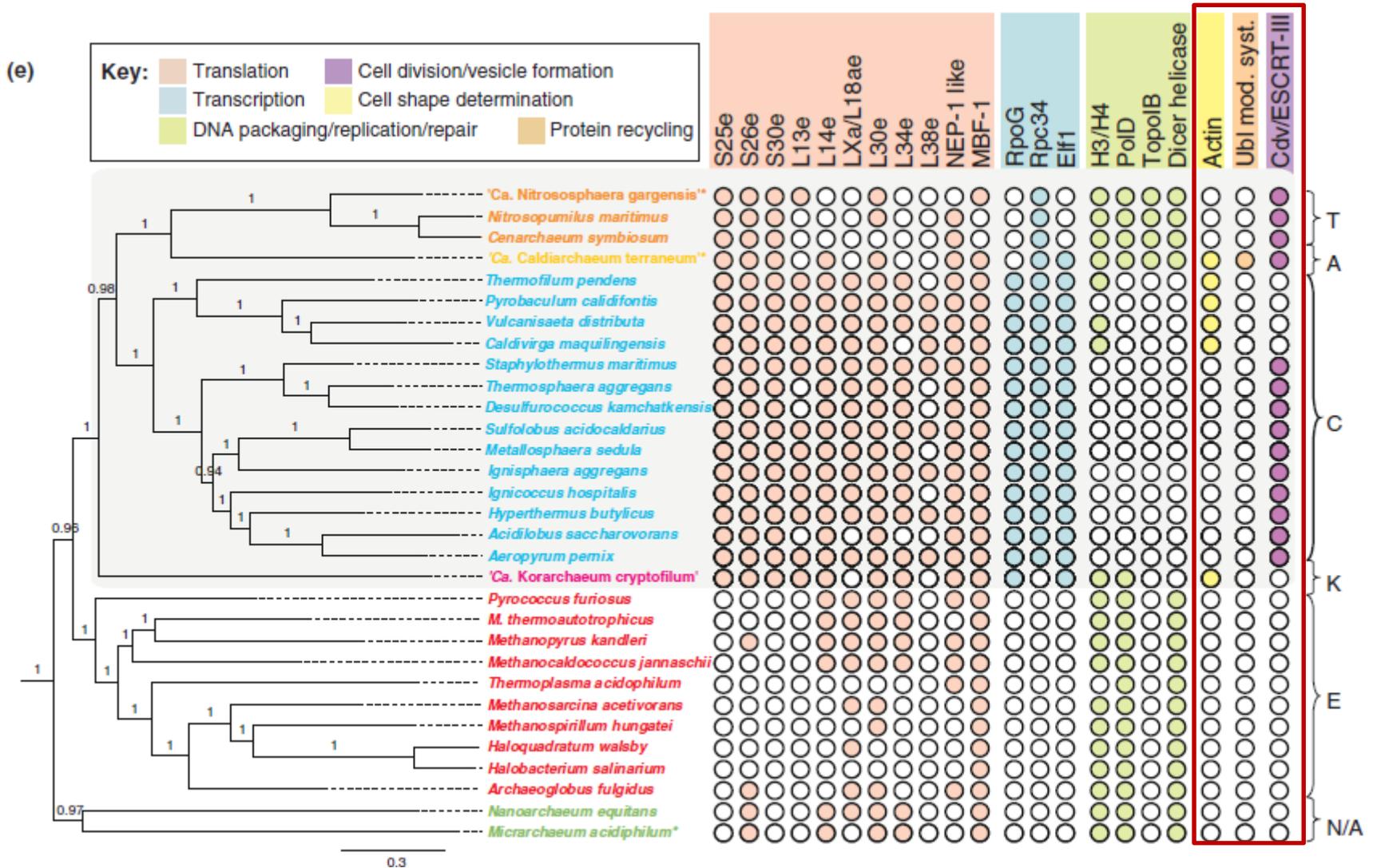
# L'arbre universel du vivant

→ Phylogénie bayésienne, concaténation de 26 protéines universelles



→ Eucaryotes sont placés à l'intérieur des archées → les gènes d'archées présents chez les eucaryotes ont été transmis verticalement depuis une archées parentale apparenté aux TACK

# L'arbre universel du vivant



# L'arbre universel du vivant

Théorie de fusion (symbiogénese): fusion entre une archée et une alpha – proteobactérie est à l'origine des Eucaryotes

Consensus:

1. LECA possédait l'ancêtre d'alpha protéobactéries qui donnera naissance à la mitochondrie
2. Les génomes d'Eucaryotes sont chimériques: les gènes informationnels sont apparentés à ceux d'archées et les gènes opérationnels sont apparentés à ceux des Bactéries
3. Les eucaryotes codent pour les protéines qui leur sont propres (ESP, eukaryotic signature proteins)

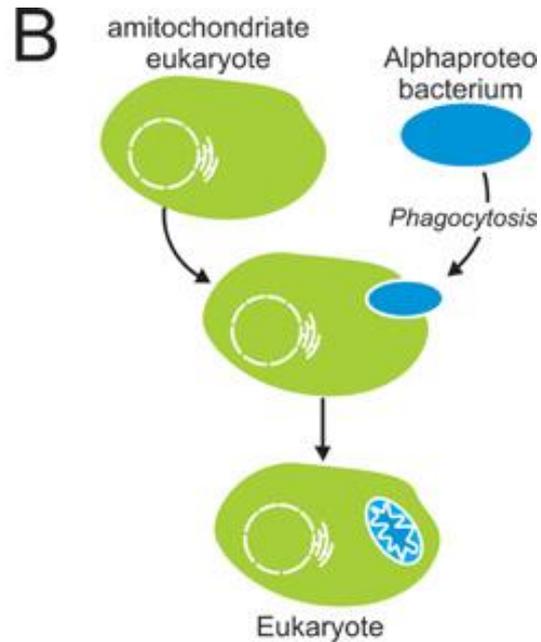
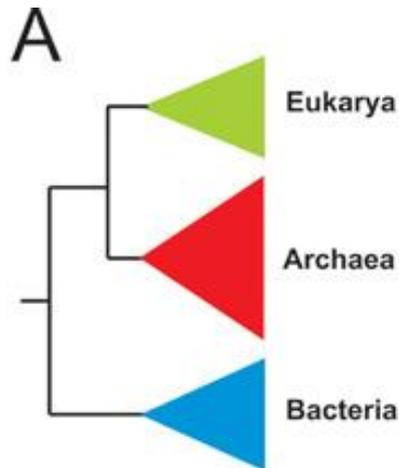
2 questions majeures:

1. Quelle était la nature de l'hôte qui a englobé l'ancêtre de l'alpha-protéobactéries?
2. La complexité eucaryote a-t-elle émergé avant l'événement d'endosymbiose (avant mitochondrie) ou après?

# L'arbre universel du vivant

Modèle « avant mitochondrie » :

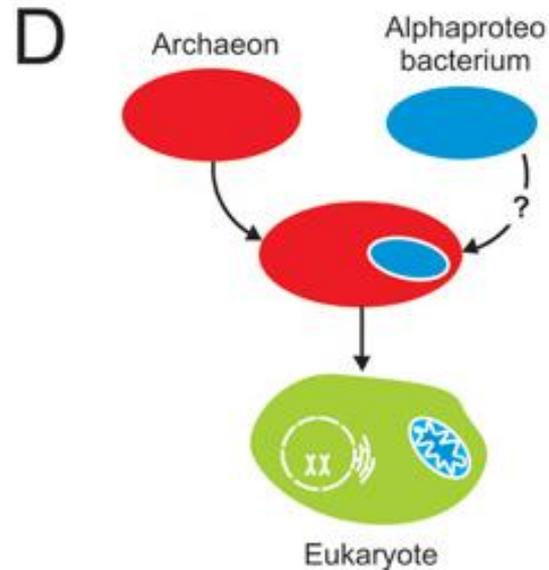
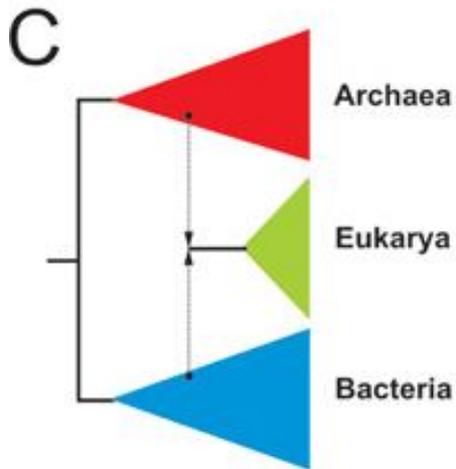
L'ancêtre commun de tous les eucaryotes est une cellule complexe de type eucaryote sans mitochondrie – scénario compatible avec l'arbre classique



# L'arbre universel du vivant

Modèle « après mitochondrie » :

la fusion entre une archée et une bactérie à déclenché l'évolution des caractères eucaryotes



« smoking gun is missing »: les organismes intermédiaires ne sont pas connus!

# L'arbre universel du vivant

Théorie de fusion (symbiogénese): fusion entre une archée (TACK) et une alpha – proteobactérie est à l'origine des Eucaryotes

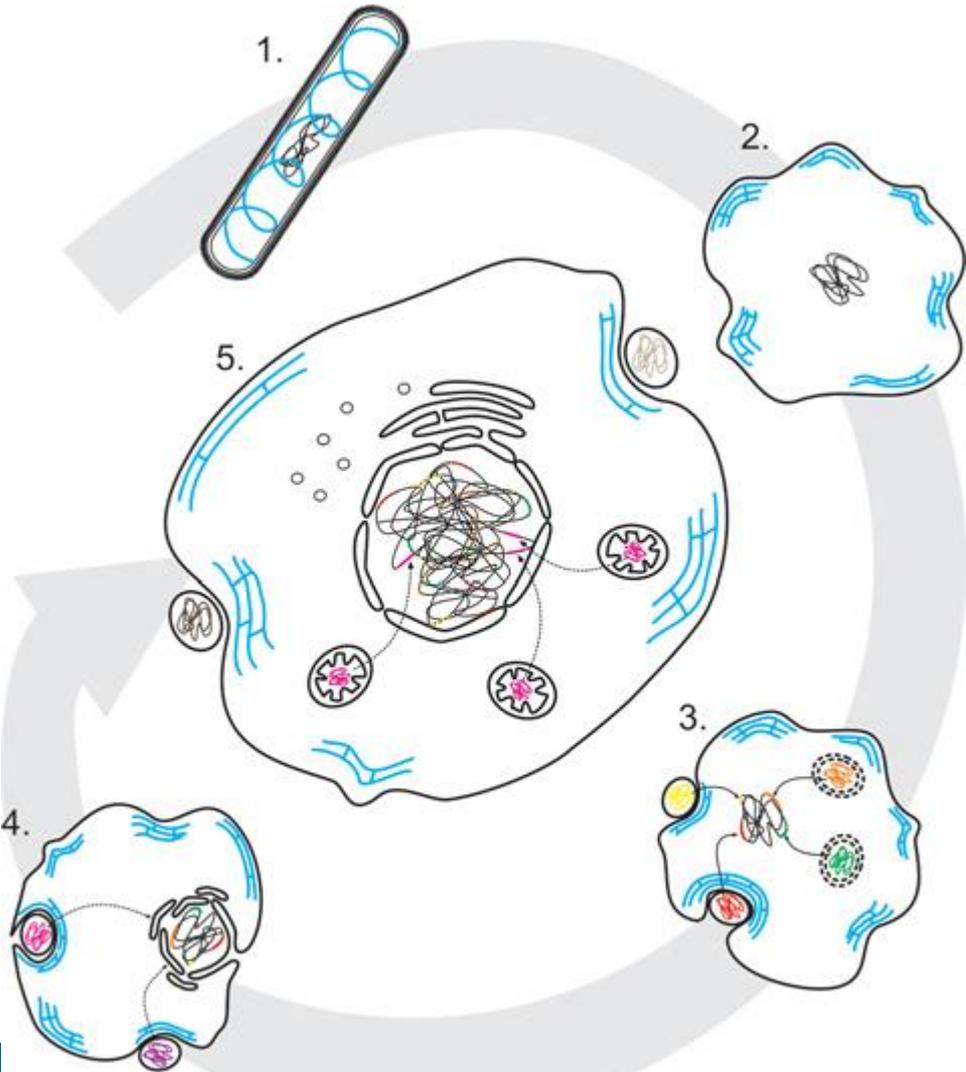
→ argument 1: le partage entre les TACK et les Eucaryotes d'ESP:

actine, tubuline, ESCRT, ubiquitine, plusieurs facteurs de transcription et de traduction

→ argument 2: l'acquisition de la mitochondrie a pu enlever la contrainte énergétique à laquelle sont soumis les archées et les bactéries pour permettre une augmentation drastique du nr. de gènes et de la taille de la cellule

# L'arbre universel du vivant

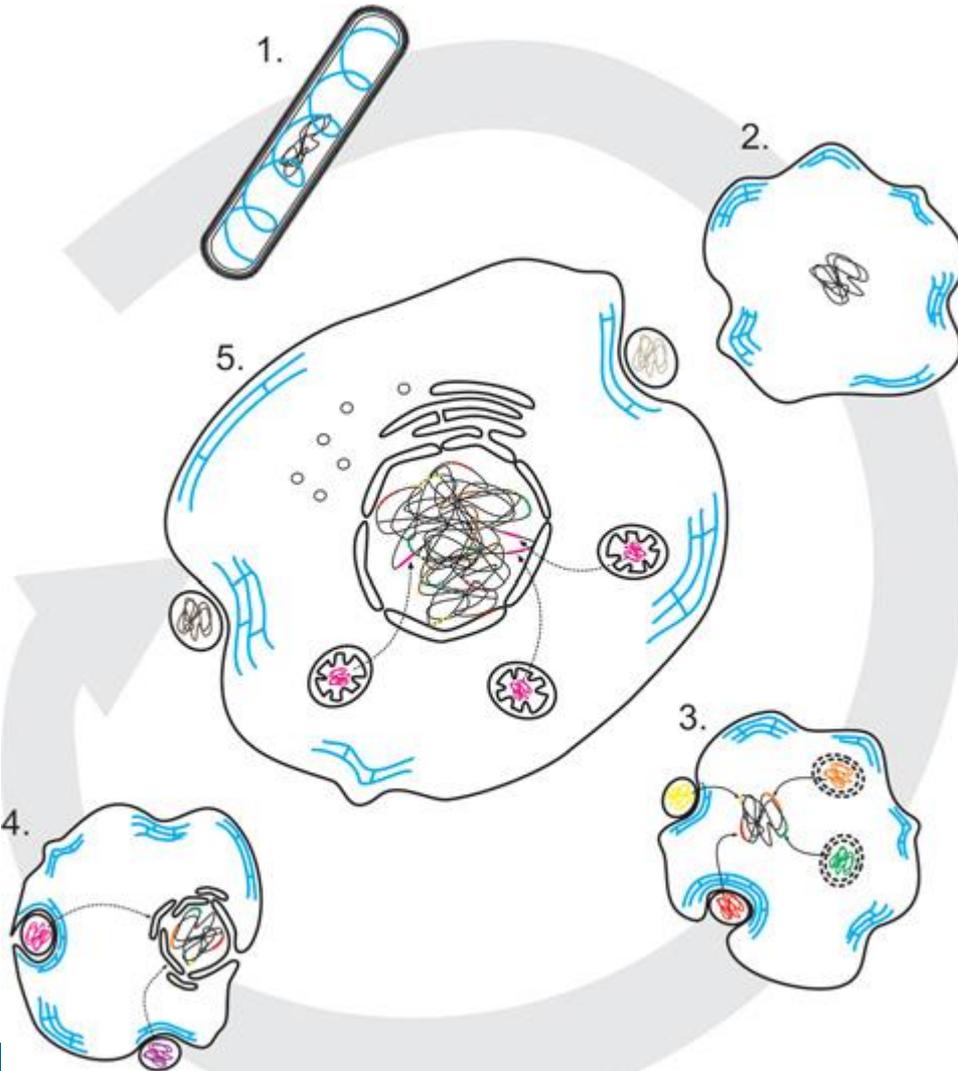
La théorie 'PhAT (phagocytosing archaeon theory)' pour l'émergence de la cellule eucaryote



1. une archée ancestrale appartenant aux TACK – contient tous les ESP connus chez les TACK
2. perte de la paroi protéique – évolution du cytosquelette (actine) - formation des extensions cellulaires

# L'arbre universel du vivant

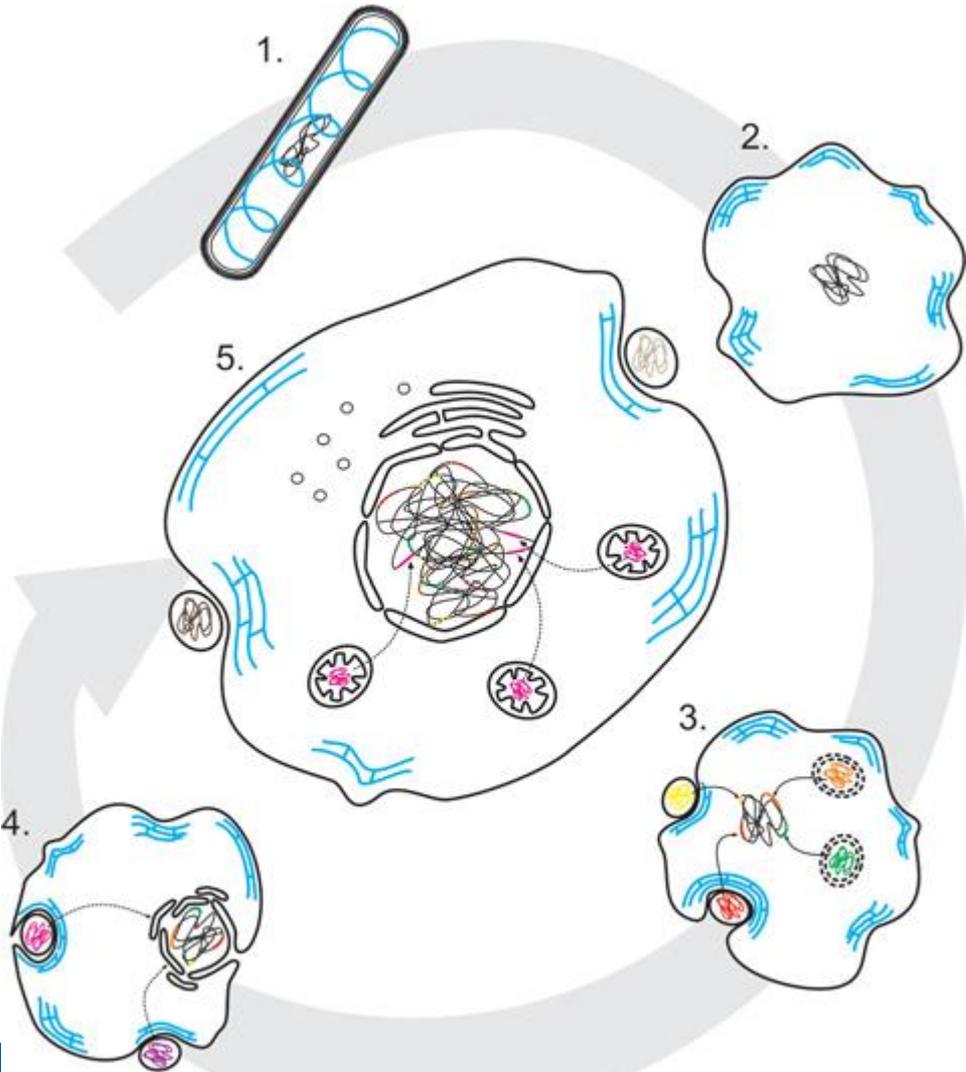
La théorie 'PhAT (phagocytosing archaeon theory)' pour l'émergence de la cellule eukaryote



3. Mise en place de la phagocytose – augmentation drastique du HGT – génome mosaïque – vitesse d'évolution accéléré (déstabilisation due au dosage de gènes inapproprié)

# L'arbre universel du vivant

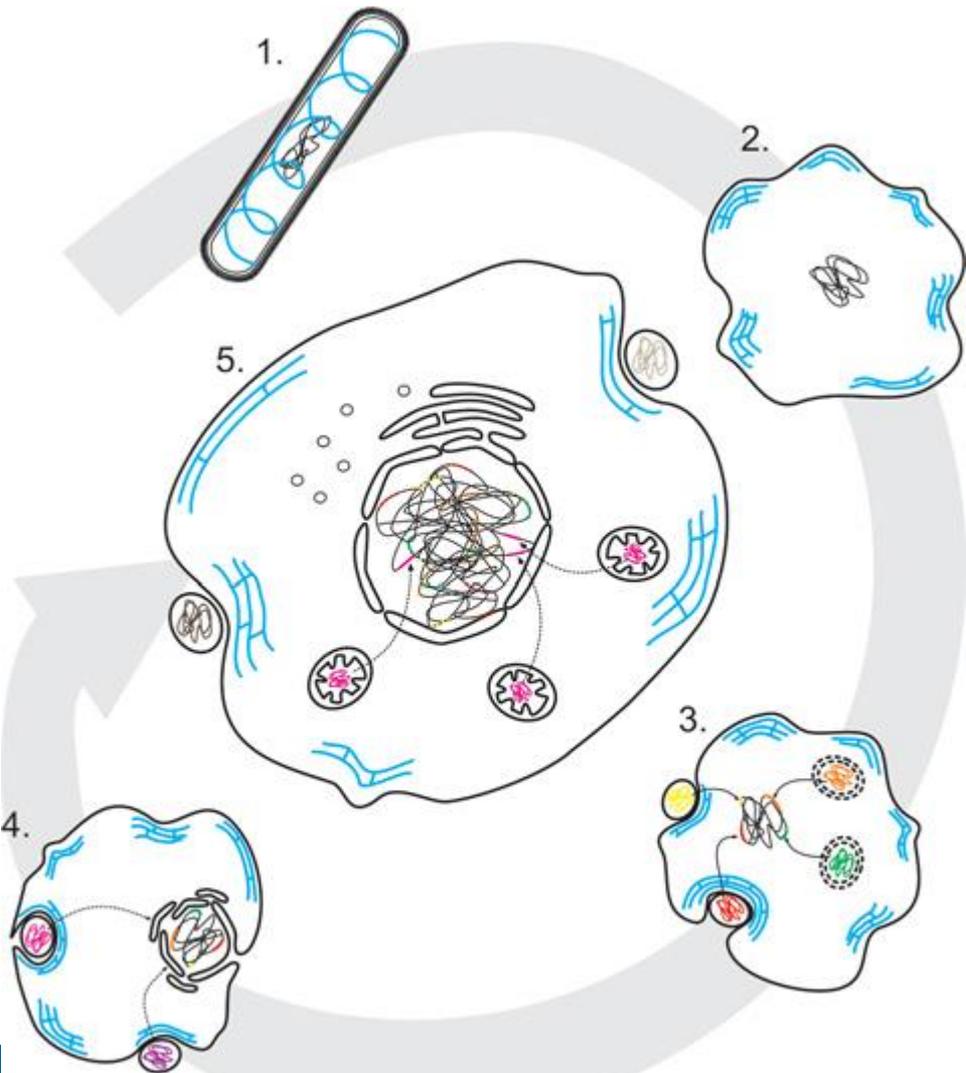
La théorie 'PhAT (phagocytosing archaeon theory)' pour l'émergence de la cellule eukaryote



4. pour protéger l'intégrité génomique  
formation d'une membrane par  
invagination - noyau primitif;  
phagocytose d'une alpha-bactérie -  
symbiose

# L'arbre universel du vivant

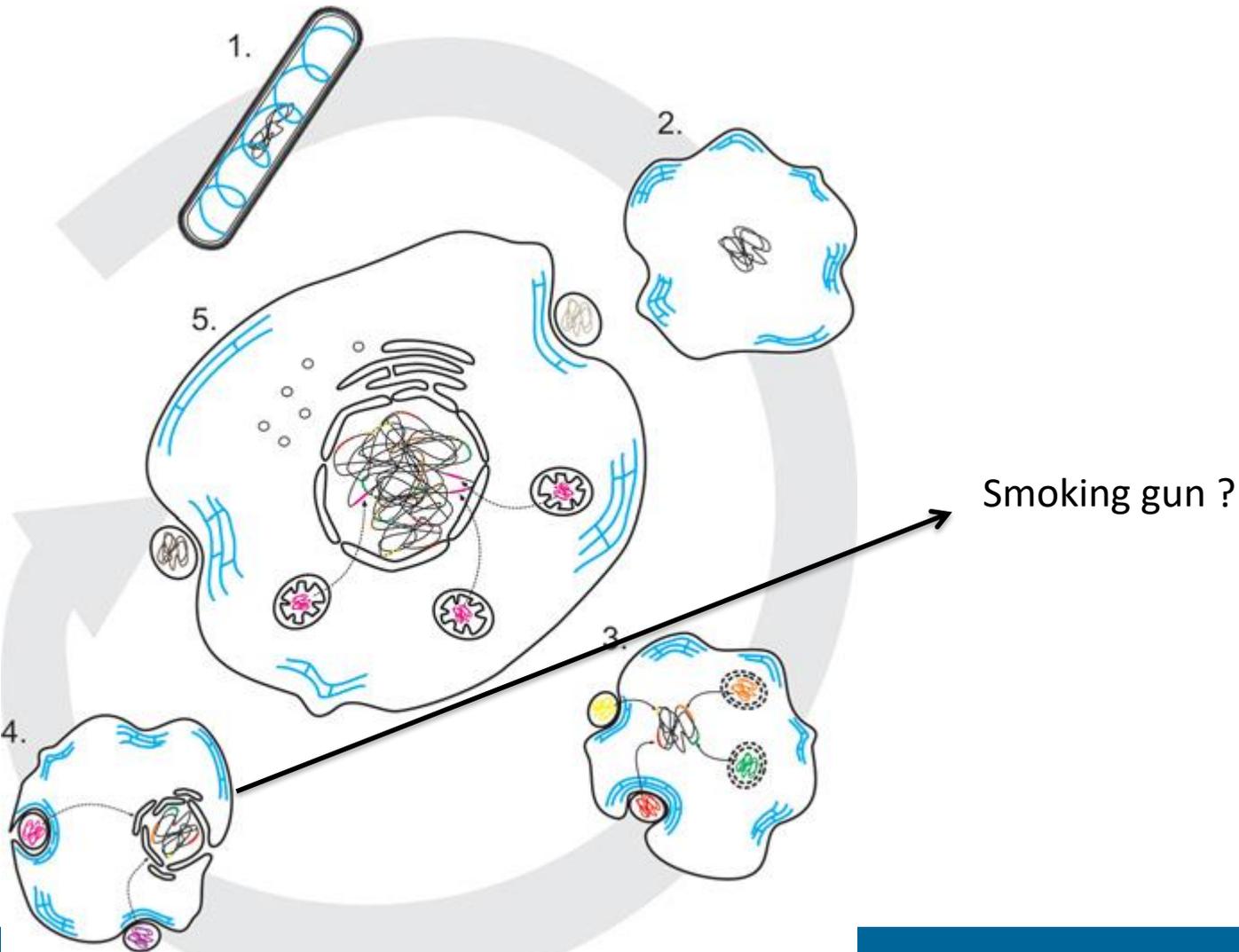
La théorie 'PhAT (phagocytosing archaeon theory)' pour l'émergence de la cellule eukaryote



5. évolution réductive d'alpha-bactérie en organe produisant l'ATP → le surplus énergétique permet complexification → expansion de la taille cellulaire (production d'énergie n'est plus intégré dans la membrane plasmique)

# L'arbre universel du vivant

La théorie 'PhAT (phagocytosing archaeon theory)' pour l'émergence de la cellule eucaryote





Biological: Full-length

# Prokaryote or eukaryote? A unique microorganism from the deep sea

Masashi Yamaguchi<sup>1,\*</sup>, Yuko Mori<sup>2</sup>, Yoshimichi Kozuka<sup>3</sup>, Hitoshi Okada<sup>1,4</sup>, Katsuyuki Uematsu<sup>5</sup>, Akihiro Tame<sup>5</sup>, Hiromitsu Furukawa<sup>2</sup>, Tadashi Maruyama<sup>6</sup>, Cedric O’Driscoll Worman<sup>7</sup> and Koji Yokoyama<sup>1</sup>

## *Parakaryon myojinensis*

- l’organisme avec des caractéristiques intermédiaires entre les procaryotes et eucaryotes
- matériel génétique entouré de membrane (non-intégrale)
- présence de 3 cellules « endosymbiontes » à l’intérieur de la cellule
- pas de mitochondrie

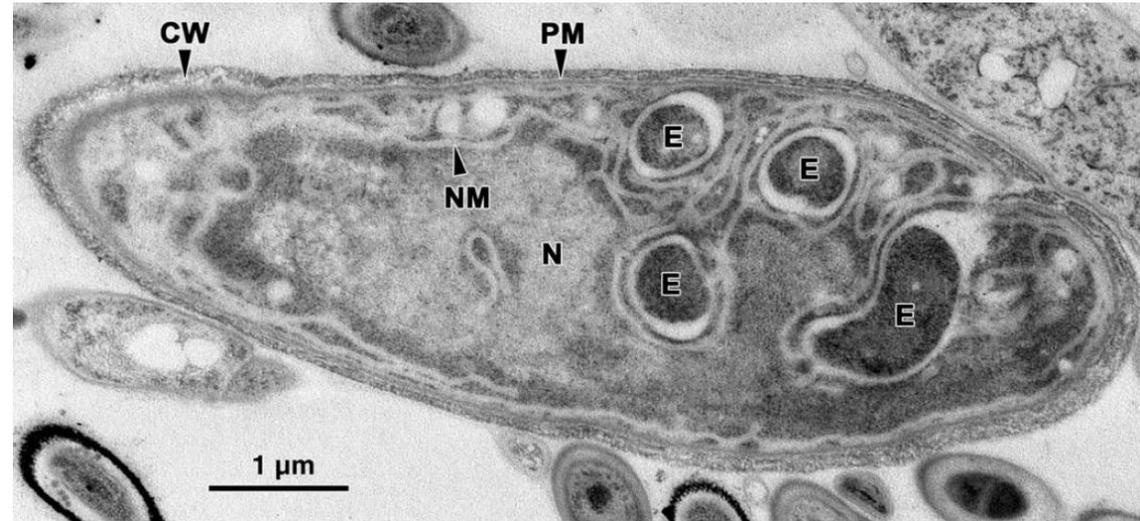
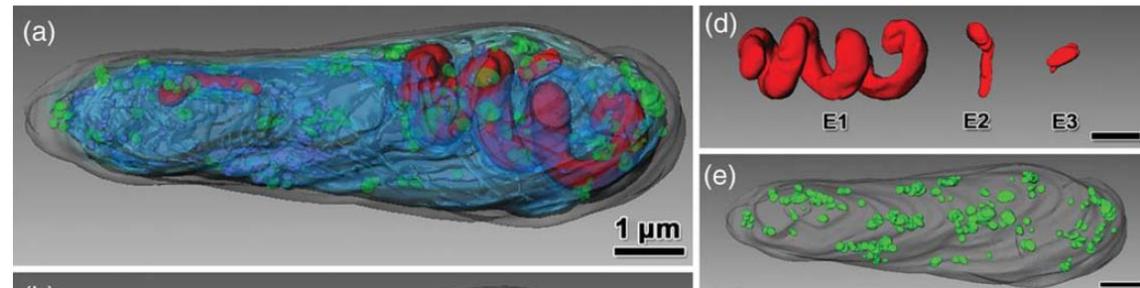


Fig. 1. An ultrathin section of *Parakaryon myojinensis*. Note the large irregular nucleoid (N) with single nucleoid membrane (NM), the presence of endosymbionts (E) and the absence of mitochondria. Also labeled are the cell wall (CW) and plasma membrane (PM).

- une endosymbiose ancienne ou une relation de prédateur – proie récente?





Biological: Full-length

Prokaryote or eukaryote? A unique microorganism from the

deep sea → sédiments – environnement très stable → « fossiles vivants » sont à trouver ici?

Masashi Yamaguchi<sup>1,\*</sup>, Yuko Mori<sup>2</sup>, Yoshimichi Kozuka<sup>3</sup>, Hitoshi Okada<sup>1,4</sup>,  
Katsuyuki Uematsu<sup>5</sup>, Akihiro Tame<sup>5</sup>, Hiromitsu Furukawa<sup>2</sup>,  
Tadashi Maruyama<sup>6</sup>, Cedric O’Driscoll Worman<sup>7</sup> and Koji Yokoyama<sup>1</sup>

*Parakaryon myojinensis*

- l’organisme avec des caractéristiques intermédiaires entre les procaryotes et eucaryotes
- matériel génétique entouré de membrane (non-intégrale)
- présence de 3 cellules « endosymbiontes » à l’intérieur de la cellule

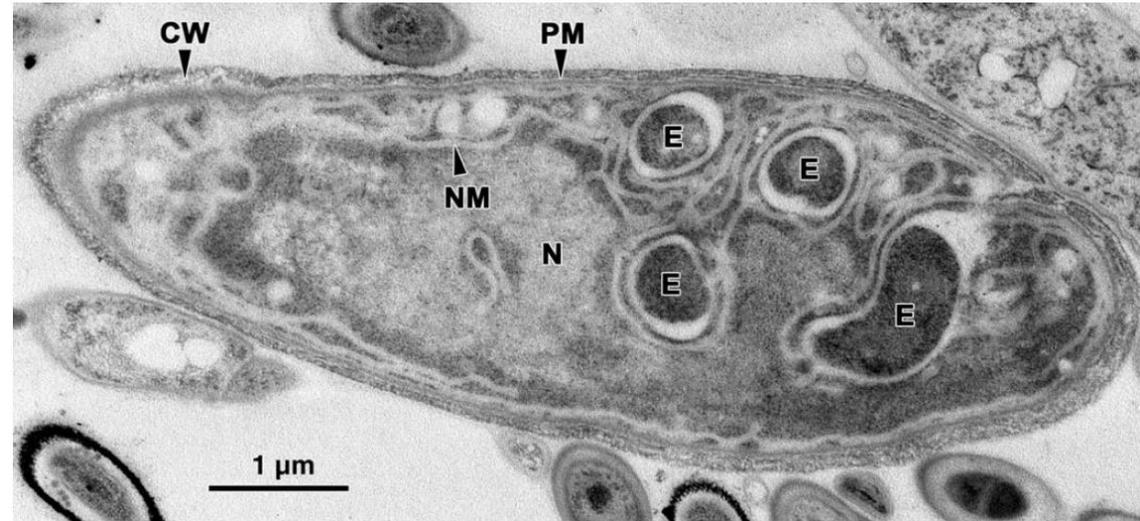
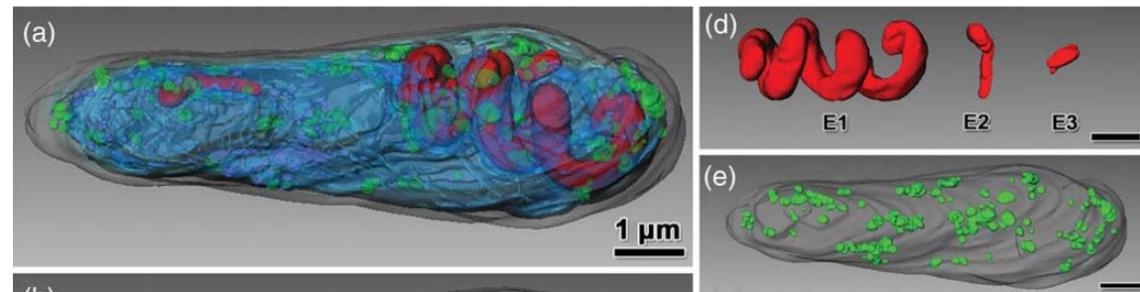


Fig. 1. An ultrathin section of *Parakaryon myojinensis*. Note the large irregular nucleoid (N) with single nucleoid membrane (NM), the presence of endosymbionts (E) and the absence of mitochondria. Also labeled are the cell wall (CW) and plasma membrane (PM).

- une endosymbiose ancienne ou une relation de prédateur – proie récente?





Thijs Ettema

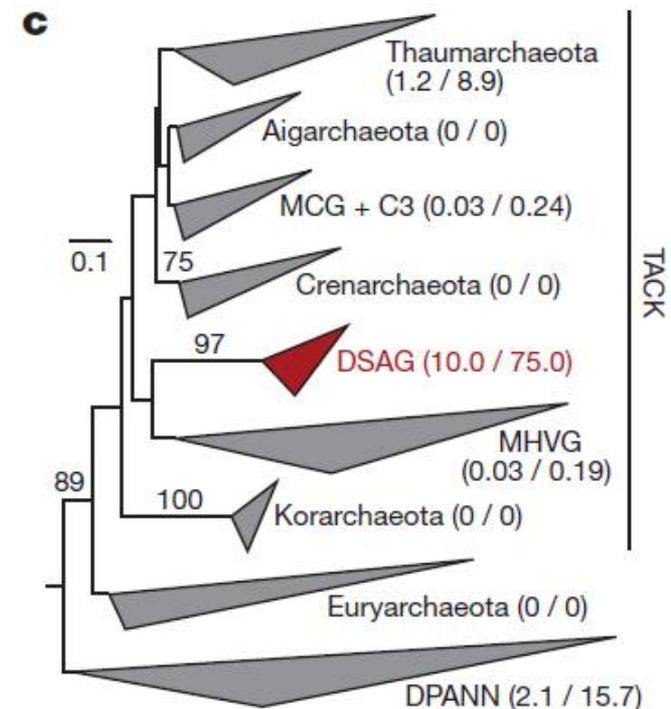
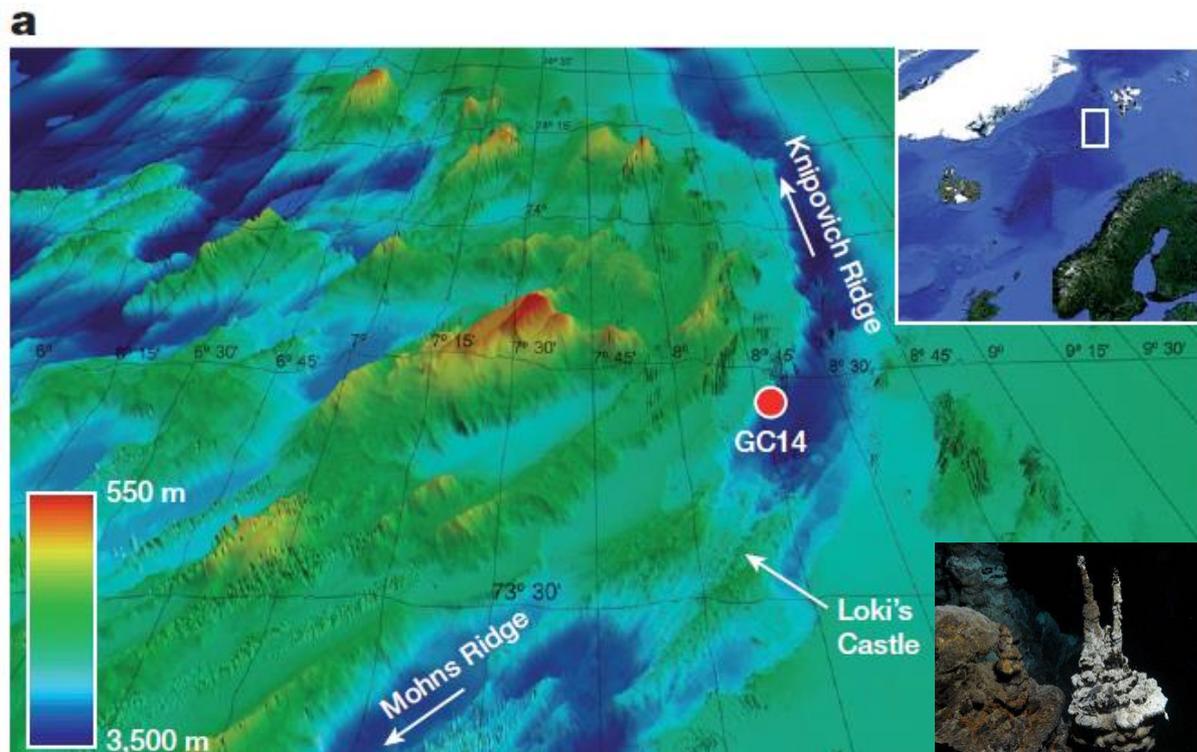
# Complex archaea that bridge the gap between prokaryotes and eukaryotes

Mai, 2015

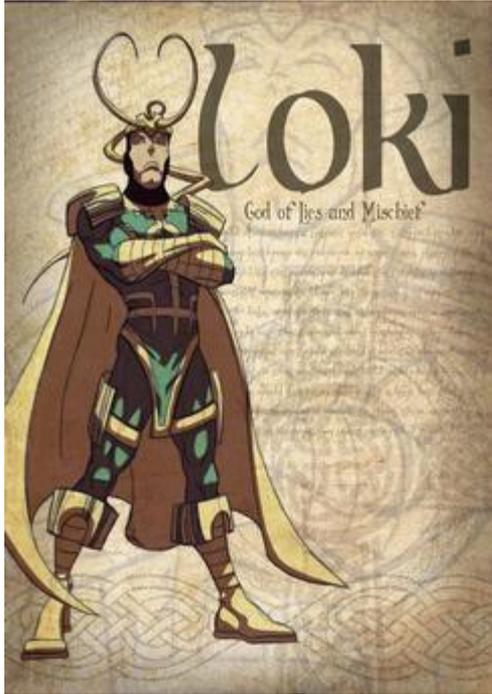
Anja Spang<sup>1\*</sup>, Jimmy H. Saw<sup>1\*</sup>, Steffen L. Jørgensen<sup>2\*</sup>, Katarzyna Zaremba-Niedzwiedzka<sup>1\*</sup>, Joran Martijn<sup>1</sup>, Anders E. Lind<sup>1</sup>, Roel van Eijk<sup>1†</sup>, Christa Schleper<sup>2,3</sup>, Lionel Guy<sup>1,4</sup> & Thijs J. G. Ettema<sup>1</sup>

## Découverte de Lokiarchaeota:

- échantillons de sédiments marins profonds
- amplification et séquençage 16S rDNA → ~10 % sont DSAG (deep sea archaeal group)



# L'arbre universel du vivant

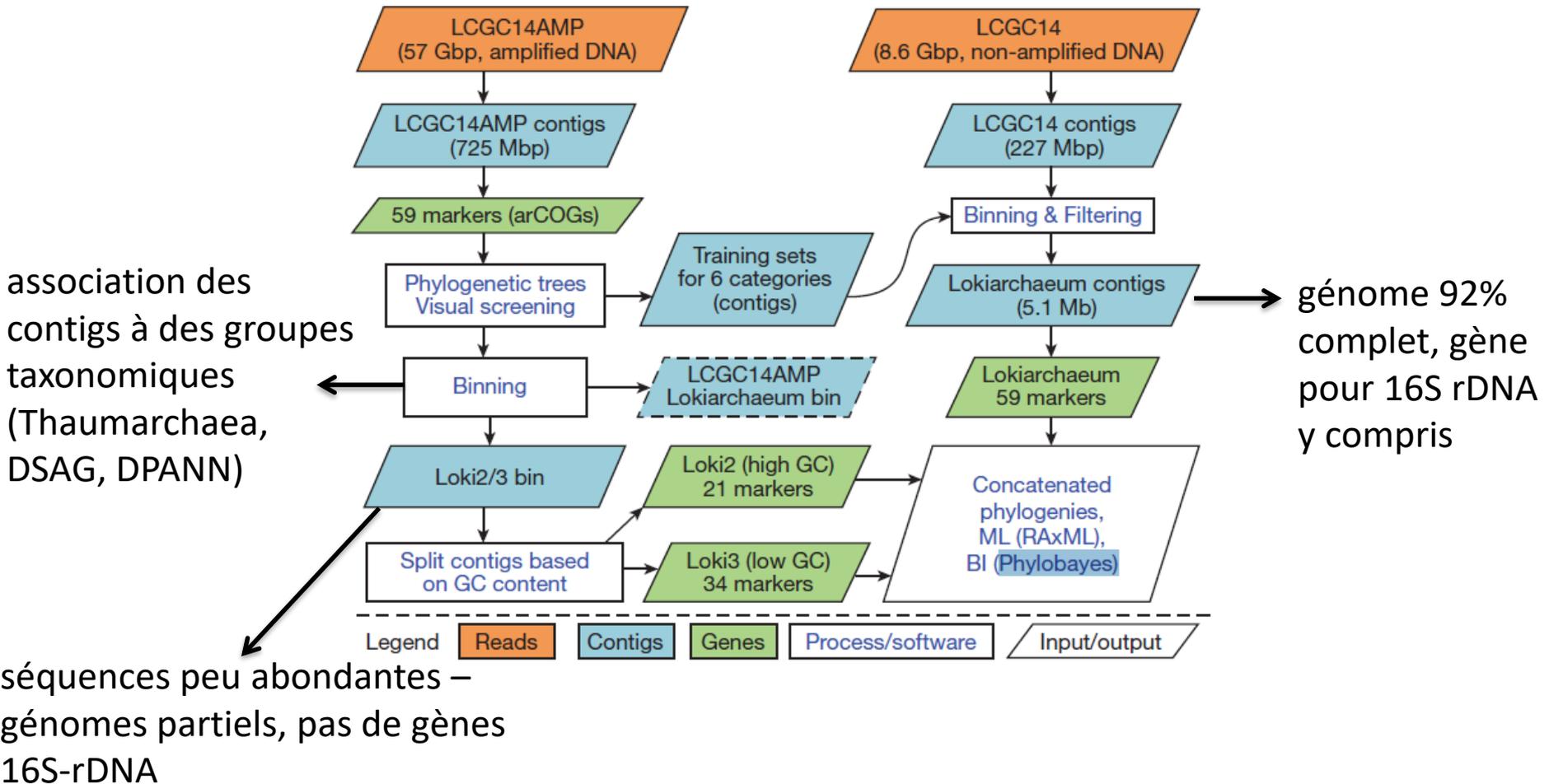


→ dieu dans la mythologie nordique, peut changer de forme

« Loki has been described as “a staggeringly complex, confusing, and ambivalent figure who has been the catalyst of countless unresolved scholarly controversies”, in analogy to the ongoing debates on the origin of eukaryotes.”

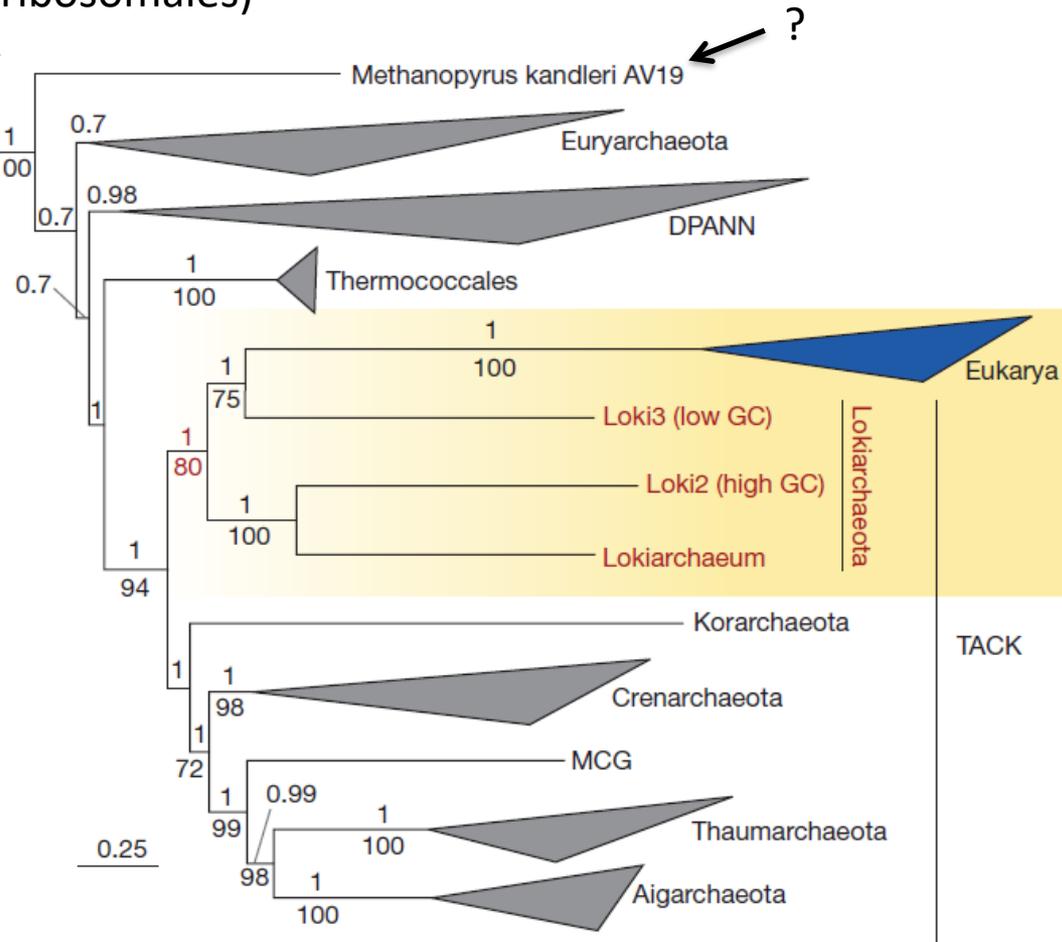


# L'arbre universel du vivant



# L'arbre universel du vivant

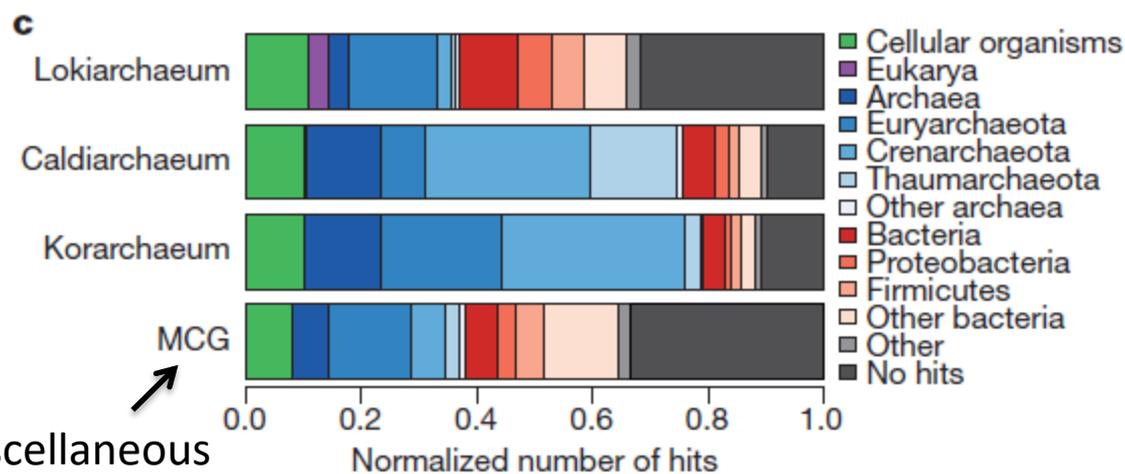
- analyse phylogénétique bayésienne: 36 protéines conservés (majoritairement protéines ribosomales)



Les Lokiarchées ont un ancêtre en commun avec les Eucaryotes!

# L'arbre universel du vivant

→ l'annotation du génome du Lokiarchaeum: 32% des ORFans, 26% protéines d'archées, 29% protéines bactériennes, 175 protéines = 3% → ESP



Miscellaneous  
crenarchaeota  
group – non-cultivés

→ contamination avec des séquences issues des génomes bactériens et eucaryotes (problème d'assemblage des contigs à partir des données brutes de séquençage et/ou problème de « binning » )?

# L'arbre universel du vivant

→ pour les protéines d'origine bactérienne:

-pas de protéines ribosomales parmi les séquences bactériennes, les protéines bactériennes ont les fonctions dans les voies métaboliques, signalisation, les mécanismes de défense, transport...

→ pour les protéines d'origine eucaryote:

-les ESP sont toujours entouré par des protéines procaryotes (aussi des gènes d'origine bactérienne)

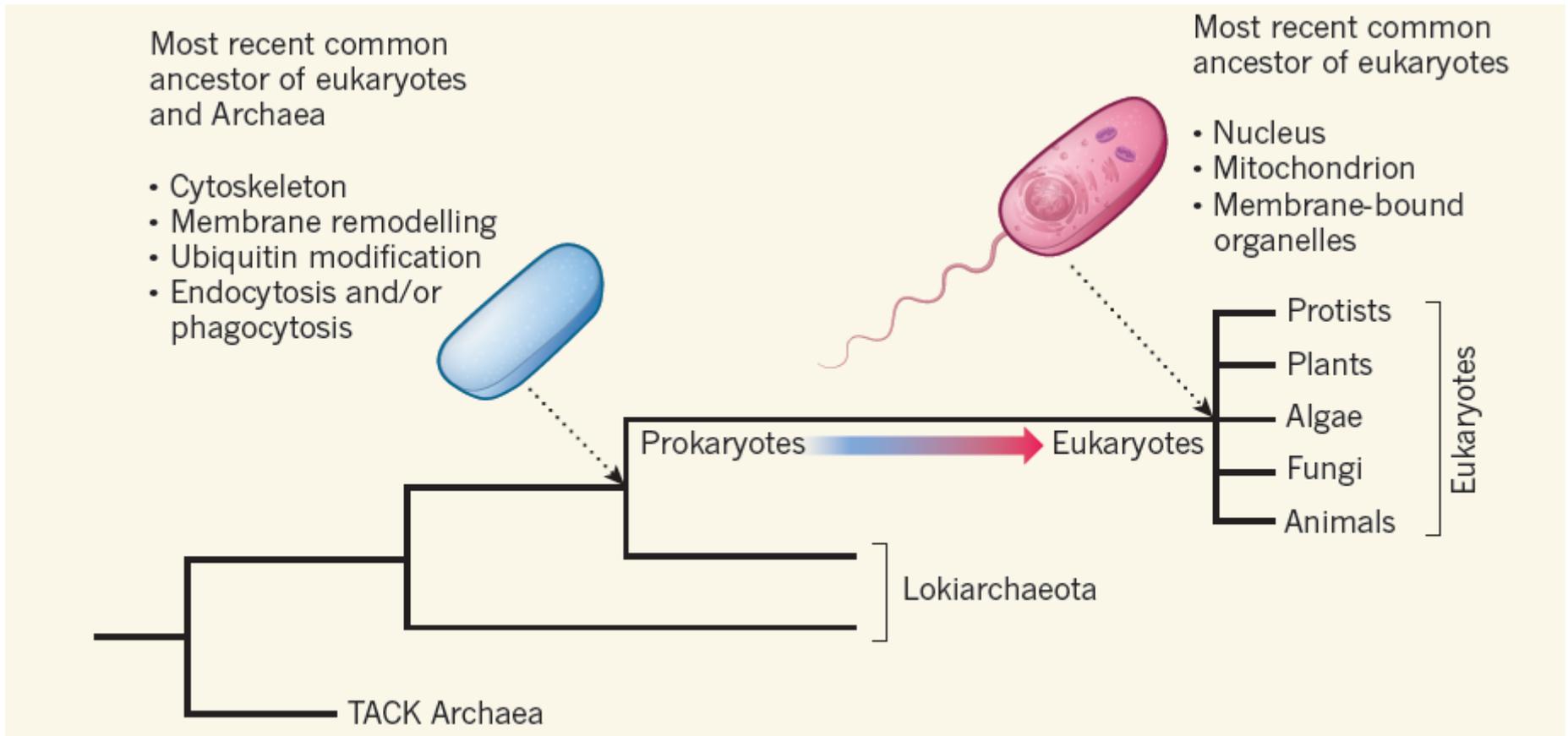
-la plupart contenus dans les contigs contenant des gènes « signature » pour le archées

-les séquences d'ESP sont retrouvés dans les données metagénomiques avec une haute redondance (>20x) mais les séquences eucaryote ne sont pas détectables ou alors à une fréquence très faible

-pas d'amplification de gènes 18S rDNA



# L'arbre universel du vivant

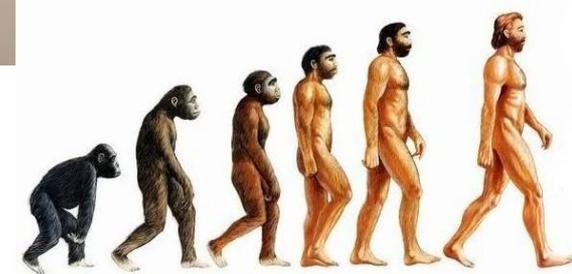
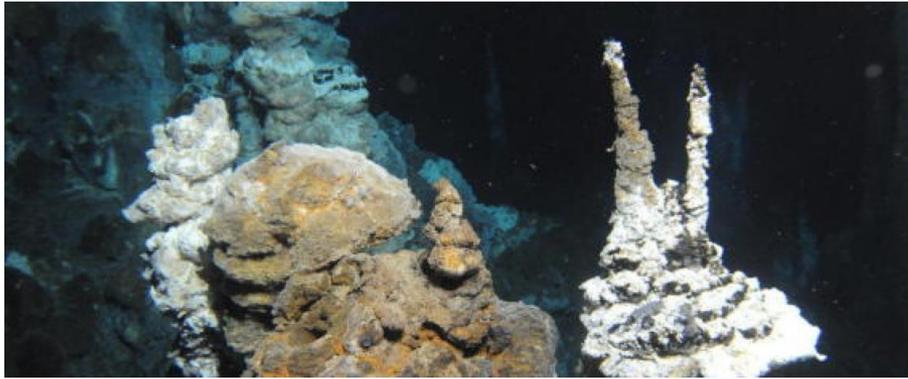


**Figure 1 | Lokiarchaeota are the closest known prokaryotic relatives of eukaryotes.** Phylogenetic trees

# Découverte d'un nouveau microbe, le Lokiarchaeota, qui lève une partie du voile sur l'origine de la vie complexe

Rédaction du HuffPost

Publication: 07/05/2015 18h40 CEST | Mis à jour: 07/05/2015 21h50 CEST



Créa

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)



Wiki Loves Africa: share African cultural fashion and adornment pictures with t

## Lokiarchaeota

From Wikipedia, the free encyclopedia

**Lokiarchaeota** is a proposed **phylum** of the **Archaea**. It was introduced in 2015 after the identification of a candidate **genome** i a **metagenomic** analysis of a mid-oceanic sediment sample. This suggests the existence of a **genus** of unicellular life dubbed **Lokiarchaeum**. The sample was taken near a **hydrothermal vent** at a site known as **Loki's Castle** on **Gakkel Ridge** in the **Arctic Ocean**.<sup>[1]</sup>

Its significance stems from a **phylogenetic** analysis that disclosed a **monophyletic** grouping of the Lokiarchaeota with the **eukaryotes**, and which revealed several **genes** with **cell membrane**-related functions; the presence of such genes support the hypothesis of the **archaeal host** (or **ecocyte-like scenarios**) for the emergence of the eukaryotes.



WIKIPEDIA  
The Free Encyclopedia

- [Main page](#)
- [Contents](#)
- [Featured content](#)
- [Current events](#)
- [Random article](#)
- [Donate to Wikipedia](#)
- [Wikipedia store](#)

Interaction

- [Help](#)
- [About Wikipedia](#)
- [Community portal](#)
- [Recent changes](#)
- [Contact page](#)

## Asgard archaea illuminate the origin of eukaryotic cellular complexity

Katarzyna Zaremba-Niedzwiedzka<sup>1\*</sup>, Eva F. Caceres<sup>1\*</sup>, Jimmy H. Saw<sup>1\*</sup>, Disa Bäckström<sup>1</sup>, Lina Juzokaite<sup>1</sup>, Emmelien Vancaester<sup>1†</sup>, Kiley W. Seitz<sup>2</sup>, Karthik Anantharaman<sup>3</sup>, Piotr Starnawski<sup>4</sup>, Kasper U. Kjeldsen<sup>4</sup>, Matthew B. Stott<sup>5</sup>, Takuro Nunoura<sup>6</sup>, Jillian F. Banfield<sup>3</sup>, Andreas Schramm<sup>4</sup>, Brett J. Baker<sup>2</sup>, Anja Spang<sup>1</sup> & Thijs J. G. Ettema<sup>1</sup>

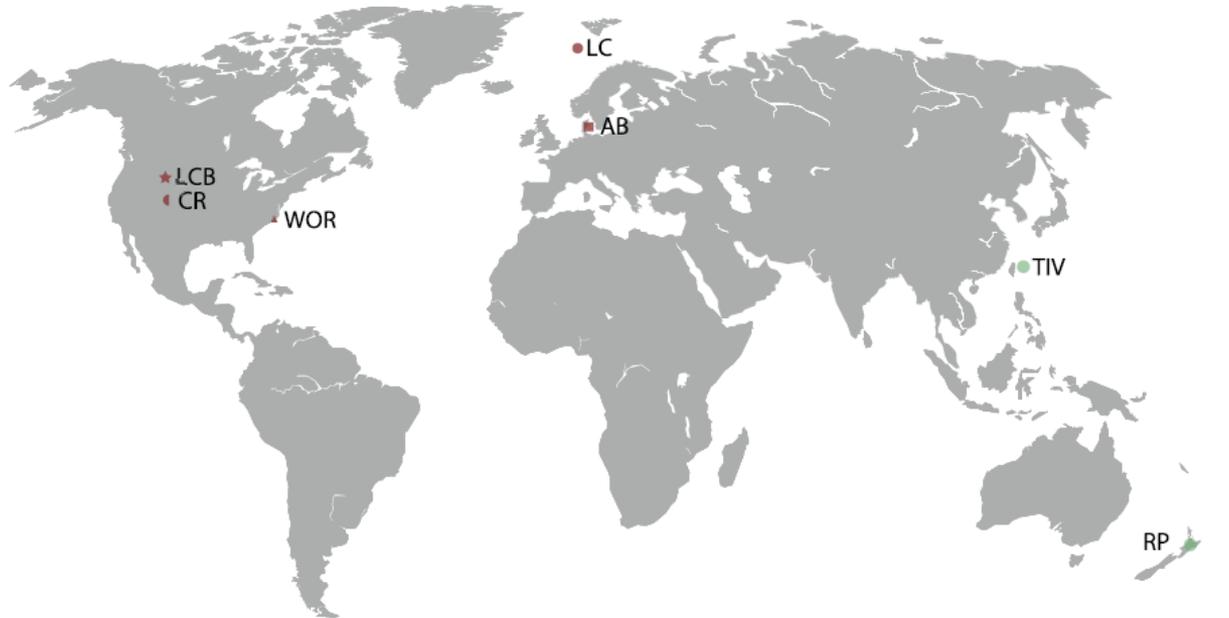
→ Recherche de nouvelles lignées apparentées aux Lokiarchées (métagénomique)

Supplementary Table 1. Overview of sampling sites

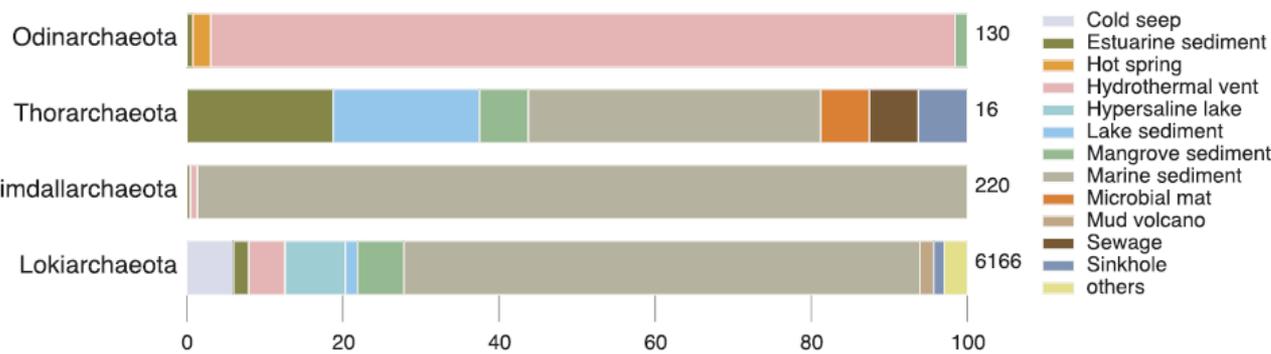
	Sample name	Site	Latitude, Longitude	Sediment Depth (cm)	Additional comments
AB	MM1/PM1	Aarhus Bay, Denmark	56°06'12"N, 10°27'28.2"E	25	
	MM2/PM2			75	
	MM3/PM3			125	
	MM4/PM4			175	
LC	Loki's Castle hydrothermal vent field	73°45'47.4"N, 8°27'50.4"E	75		
LCB	Lower Culex Basin, Yellowstone National Park, USA	44°34'23.0"N, 110°47'40.5"W	n. d.	~70°C; pH 8.6; Saw et al. 2015	
RP	Radiata Pool, Ngatamariki, New Zealand	38°31'48.0"S, 176°10'12.0"E	5	63.3°C; pH 7.5	
TIV	Taketomi Island shallow submarine hydrothermal field, Japan	24°20'54.0"N, 124°06'06.0"E	~10-15	Hirayama <i>et al.</i> 2012	
WOR	White Oak River estuary, USA	34°44'35.5"N, 77°07'26.1"W	26-30	Seitz <i>et al.</i> 2016	
CR	Aquifer near Colorado River, USA	39°31'44.69" N, 107°46'19.71" W	400	~10°C; pG 7.11; Castelle et al. 2015	

# L'arbre universel du vivant

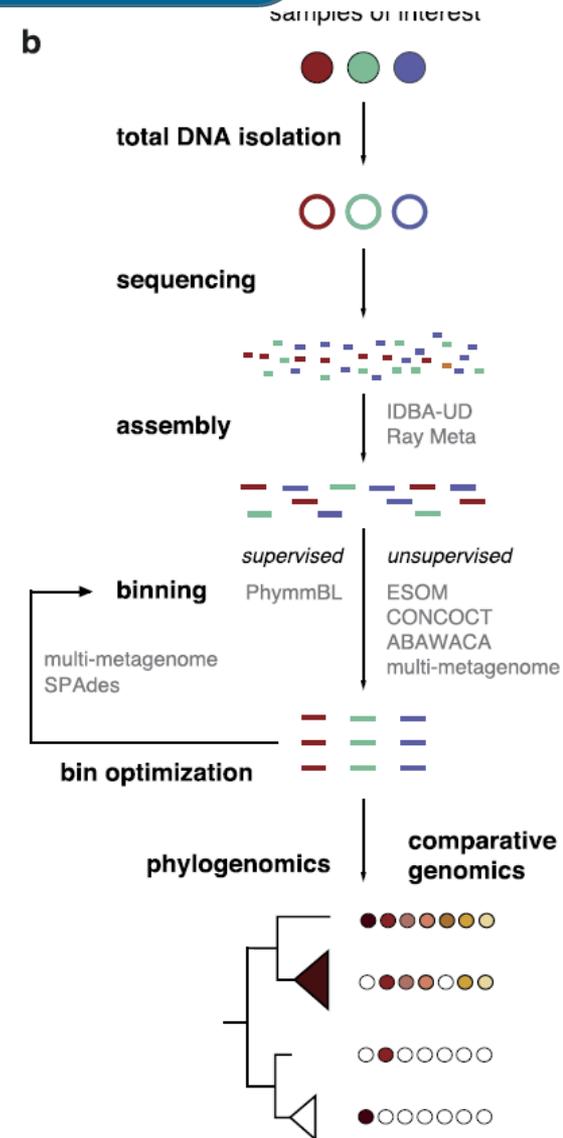
a



c

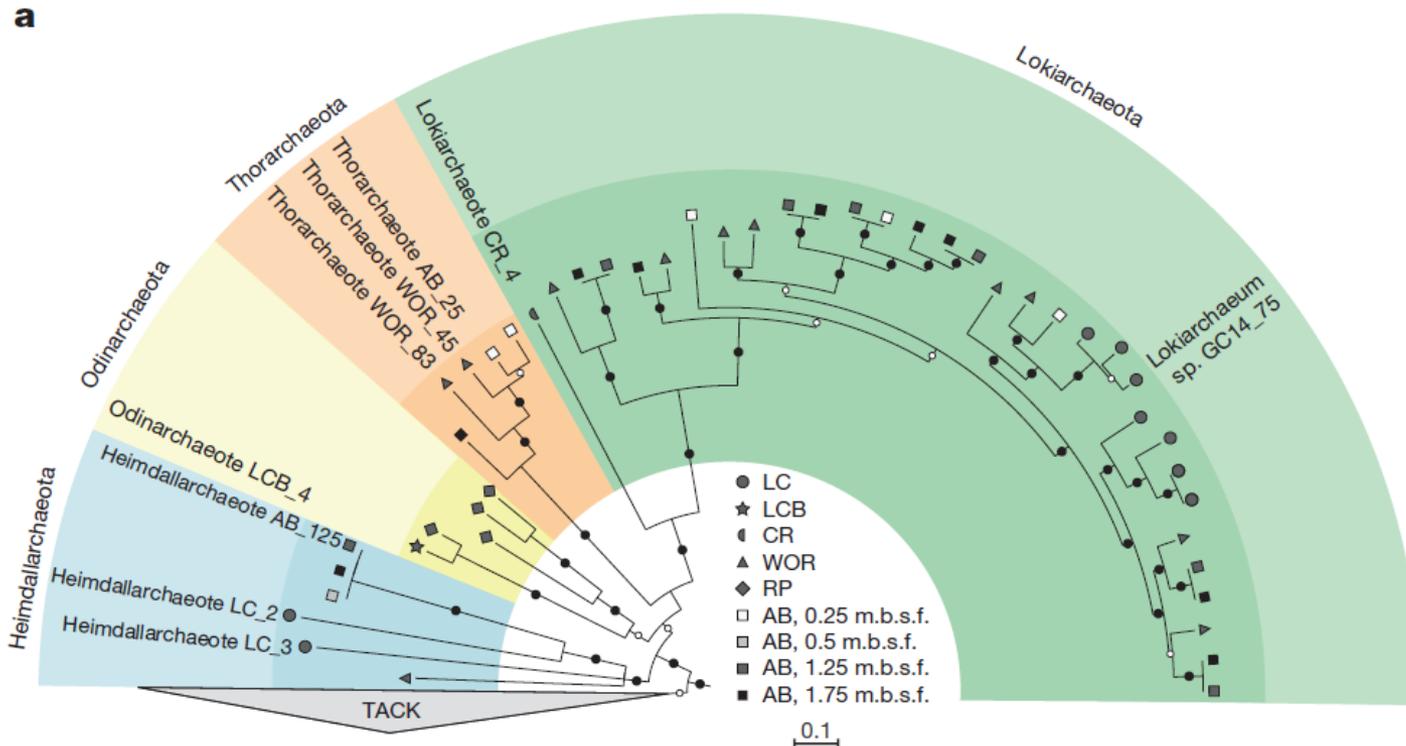


b



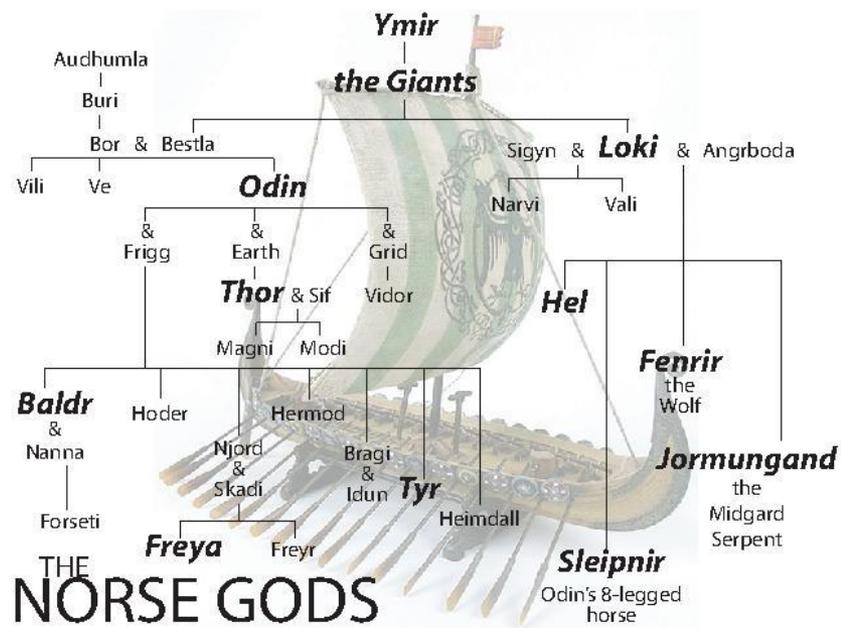
# L'arbre universel du vivant

→ Assemblage de séquences métagénomiques en 3.28 Gpb de contigs > 5 kpb → recherche des gènes ribosomiques conservés pour identifier des organismes proches de Loki → phylogénies



→ Identification de Odin- et Heimdall- archées → Asgard

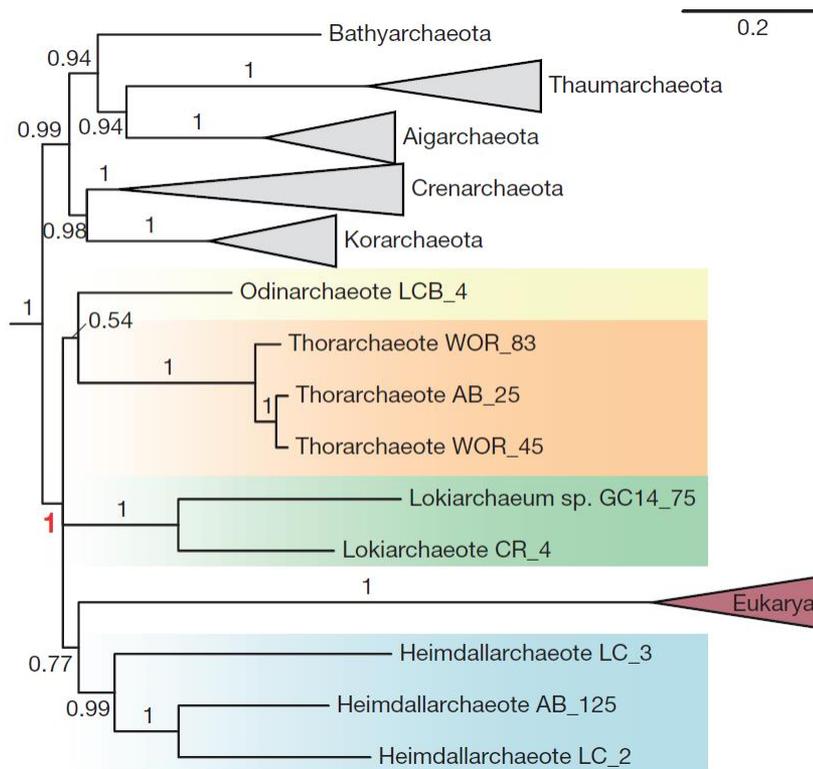
# THE NINE WORLDS



## THE NORSE GODS

# L'arbre universel du vivant

- Binning → association des contigs pour reconstruire des métagénomés
- Phylogénies en utilisant 16S et 23S rRNA, les gènes marqueurs universels, protéines ribosomales universelles et partagées archées/eucaryotes

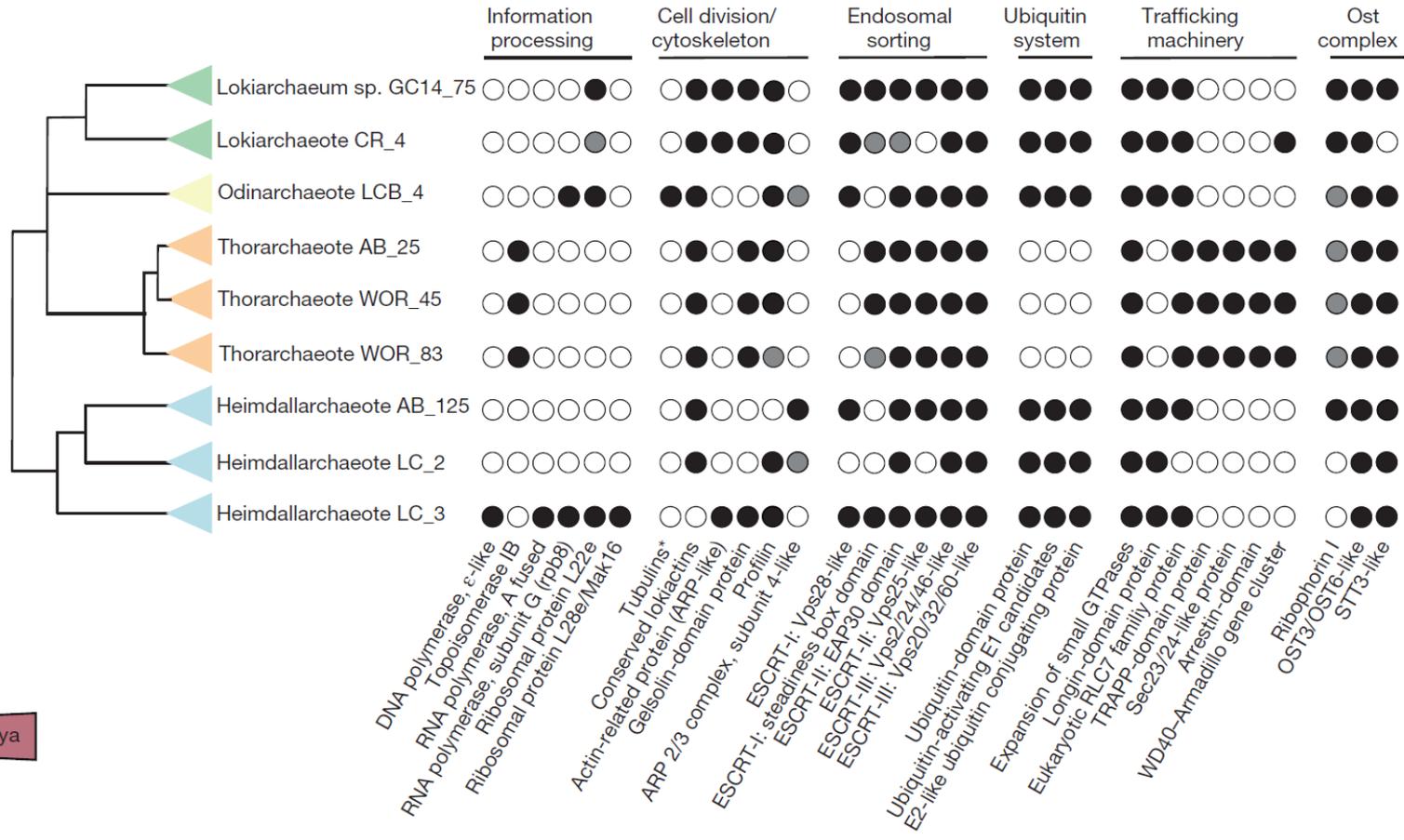


- Les Eucaryotes groupent de manière robuste (PP = 1) avec les Asgard
- les Heimdalarchées forment un groupe sœur avec les eucaryotes (PP = 0.77)
- Les relations de parenté au sein des Asgard ne sont pas résolues

Bayésien, 55 protéines ribosomales A/E, CAT-GTR model d'évolution de séquences

# L'arbre universel du vivant

→ Les métagénomés reconstruits contiennent un répertoire enrichie en ESP (Eukaryotic Signature Proteins)



ya



# L'arbre universel du vivant

Question centrale concernant l'émergence d'Eucaryotes:

Pourquoi seulement la lignée d'Asgardarchées avait acquis une complexité cellulaire caractéristique des Eucaryotes ?

*“Future exploration of novel branches in the tree of life, including novel members of the Asgard, as well as deep-branching eukaryotes, and the detailed characterization of their metabolic repertoires and cell-biological features, will undoubtedly provide new fundamental insights into the process of eukaryogenesis, and ultimately reveal how eukaryotic cells evolved their complex and compartmentalized nature.”*

RESEARCH ARTICLE

Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes

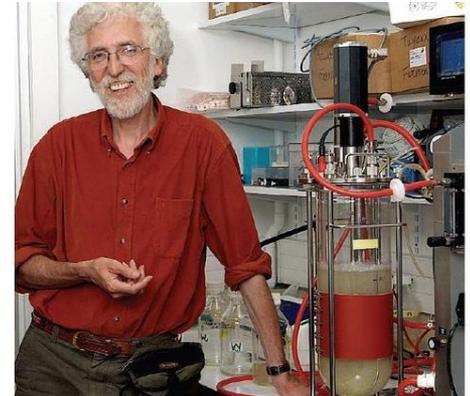
Violette Da Cunha<sup>1,2†</sup>, Morgan Gaia<sup>1‡</sup>, Daniele Gadelle<sup>2</sup>, Arshan Nasir<sup>3</sup>, Patrick Forterre<sup>1,2\*</sup>

- L'affiliation Loki-Eucaryotes a été critiqué par Caetano-Anolles et coll. (Nasir et al., 2015 Trends in Microbiology)
- L'arbre de Loki est déséquilibré (10 B, 10 E, 87 A) → cela peut biaiser la topologie (Nasir et al., 2016, Archaea)

POST-DOC



Violette Da Cunha



Patrick Forterre

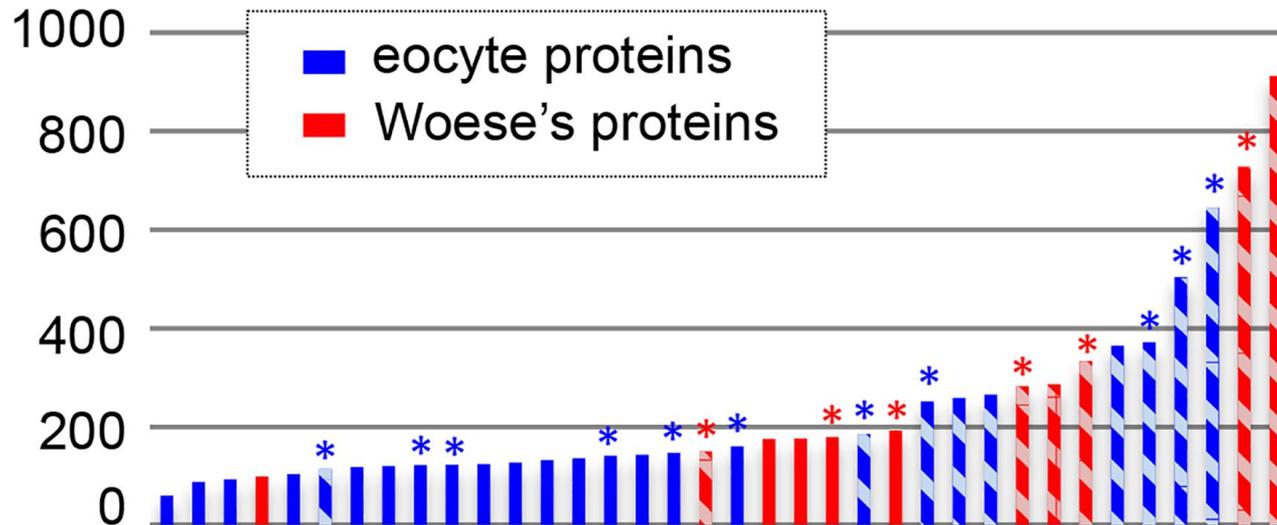
# L'arbre universel du vivant

- Ré-analyse des phylogénies individuelles des 36 protéines universelles utilisées par Spang et coll.
- Identification des protéines universelles (parmi les 36) compatibles avec l'arbre à trois ou deux domaines
- Identification des contamination dans les séquences protéiques utilisés (EF2)
- Analyse phylogénétique de l'ARN polymérase en utilisant un échantillon taxonomique équilibré

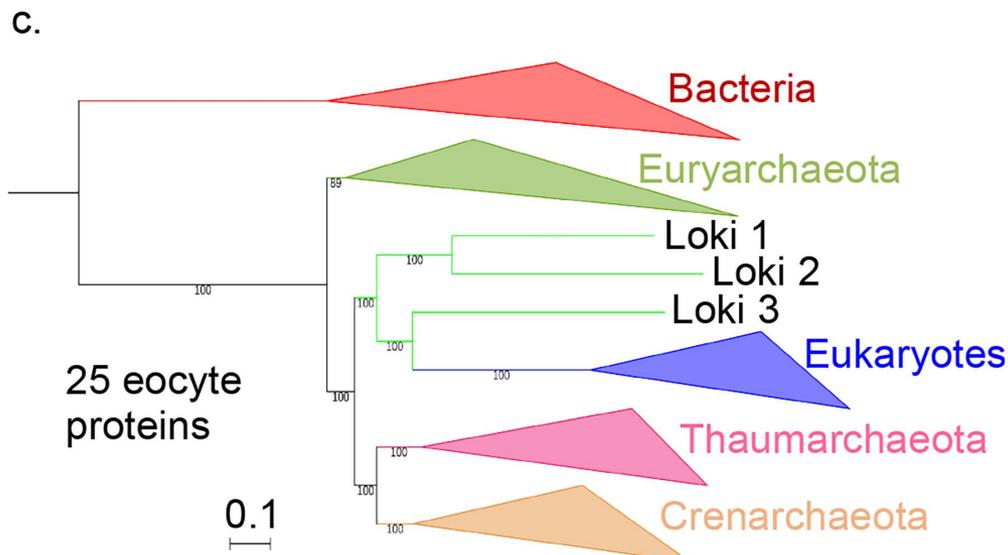
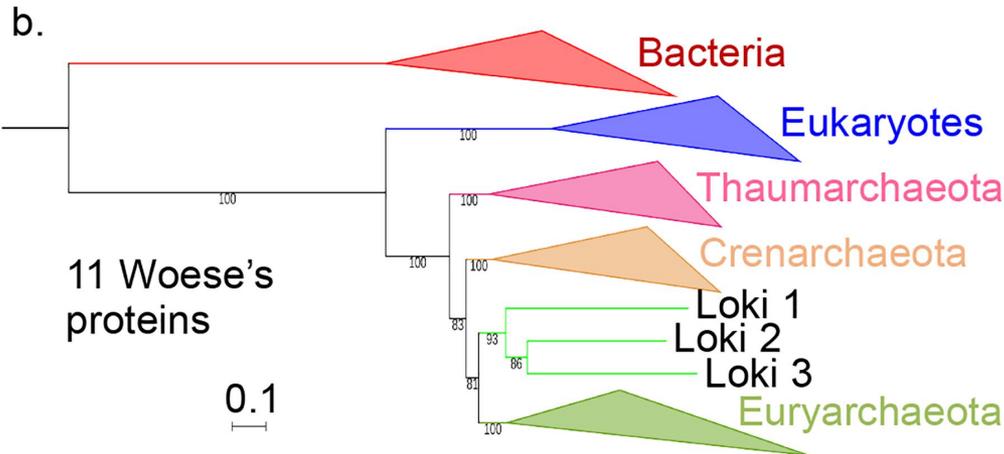
# L'arbre universel du vivant

→ Identification des protéines universelles (parmi les 36) compatibles avec l'arbre à trois ou deux domaines

a.



# L'arbre universel du vivant



→ Elimination des lignées qui évoluent vite (ex. *Methanopyrus kandleri*, Aigarchaeota, DPANN) pour minimiser LBA

# L'arbre universel du vivant

→ l'analyse de la qualité du génome de Loki 1 indique une forte contamination

Software	Number of markers	Completeness	Contamination	Heterogeneity index
CheckM	145	90.29%	45.15 %	78.21
Anvi'o	162	92,6%	56,8%	-

→ Dans le segment qui semble être le plus contaminé se trouve le gène pour EF-2 (#)

Sets									
Color	Name	Contigs	Length	Comp./Conta.	Comp./Conta.	Comp./Conta.	Comp./Conta.	Comp./Conta.	Comp./Conta.
■	Set_6	170	1.60M	21.6%/2.5%					
■	Set_5	54	558K	4.9%/0.6%					
■	Set_4	148	1.86M	71.6%/3.7%					
■	Set_3	7	22.3K	0.6%/0.0%					
■	Set_2	113	1.06M	43.2%/0.6% (#)					
■	Set_1	19	33.7K	0.0%/0.0%					

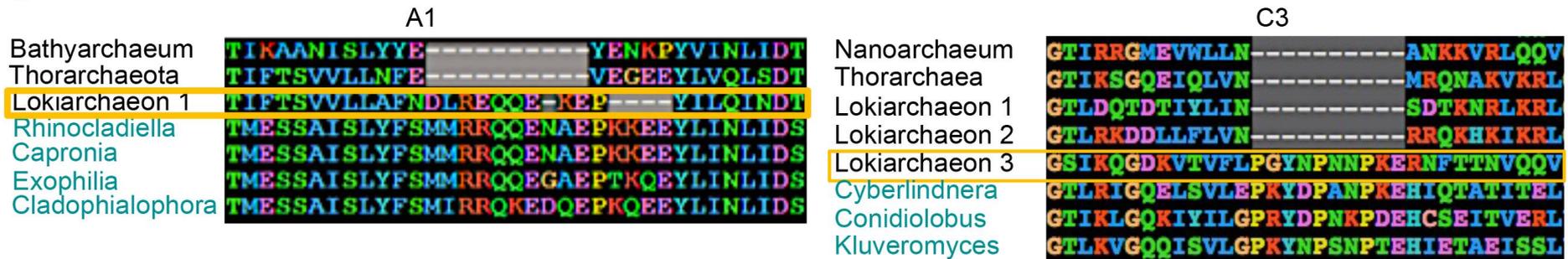
Diagram illustrating the relationship between sets and their completeness/contamination metrics:

- Set\_6 (yellow) and Set\_5 (green) are grouped together with a bracket labeled 90.7%/14.2%.
- Set\_4 (dark green) and Set\_3 (black) are grouped together with a bracket labeled 76.5%/4.3%.
- Set\_2 (pink) and Set\_1 (blue) are grouped together with a bracket labeled 92.6%/56.8%.
- Set\_4 (dark green) and Set\_2 (pink) are grouped together with a bracket labeled 92.6%/56.8%.

# L'arbre universel du vivant

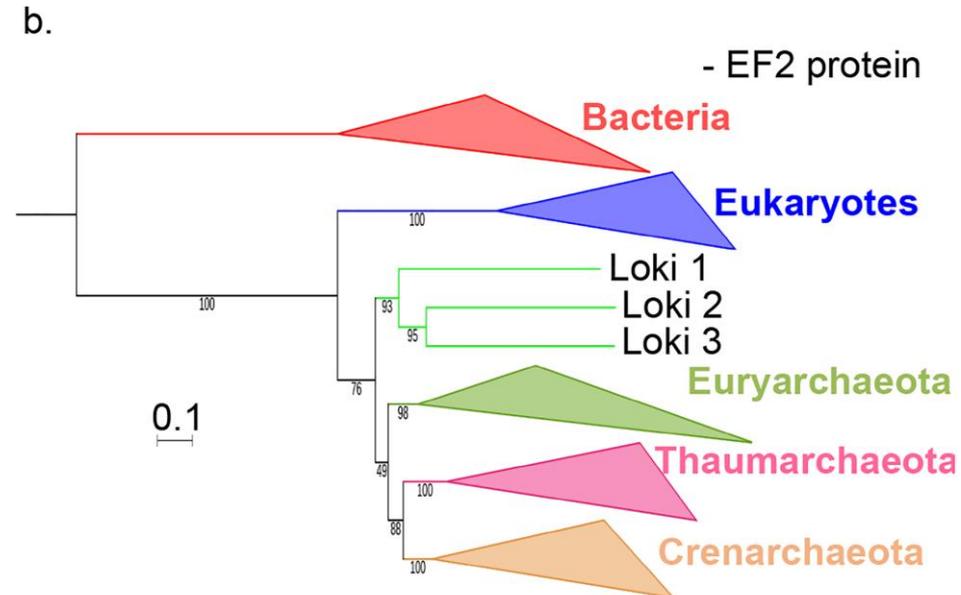
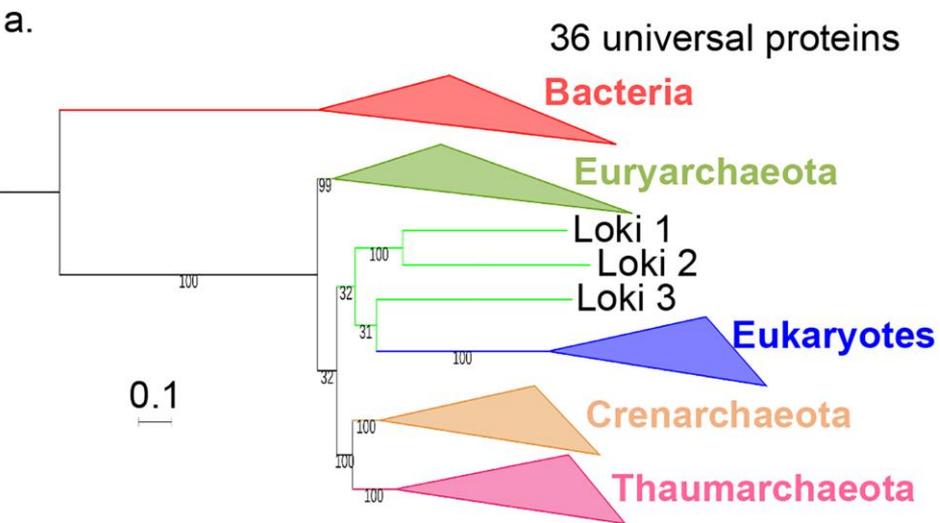
→ Les indel du type eucaryote sont trouvés dans les séquences de EF-2 assignées à Loki

c.



# L'arbre universel du vivant

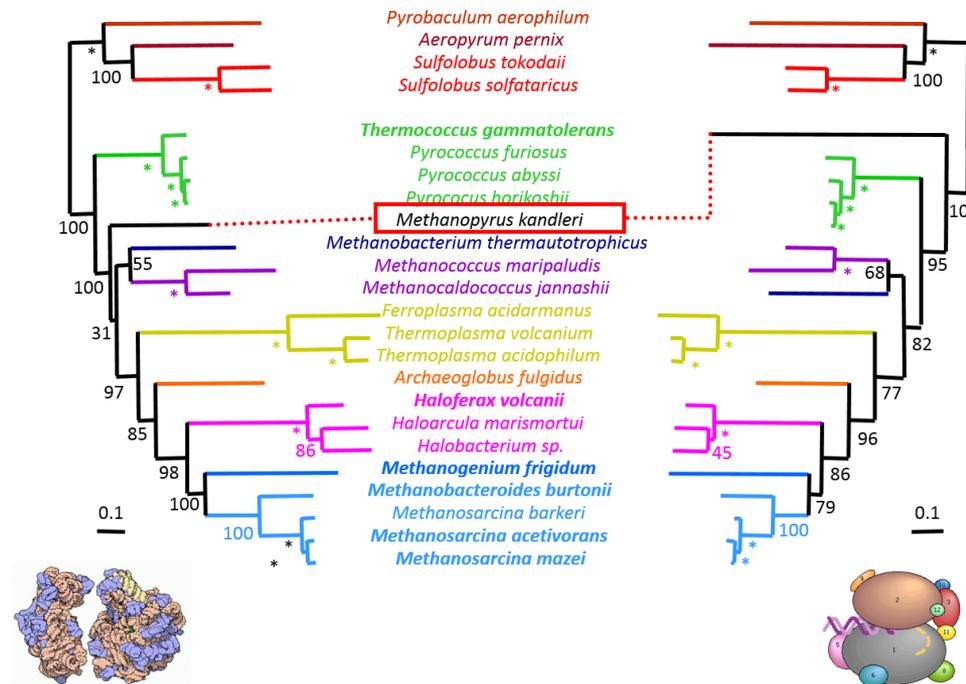
→ L'analyse faite sans les séquences qui évoluent vite et sans EF2 donne l'arbre à trois domaines!



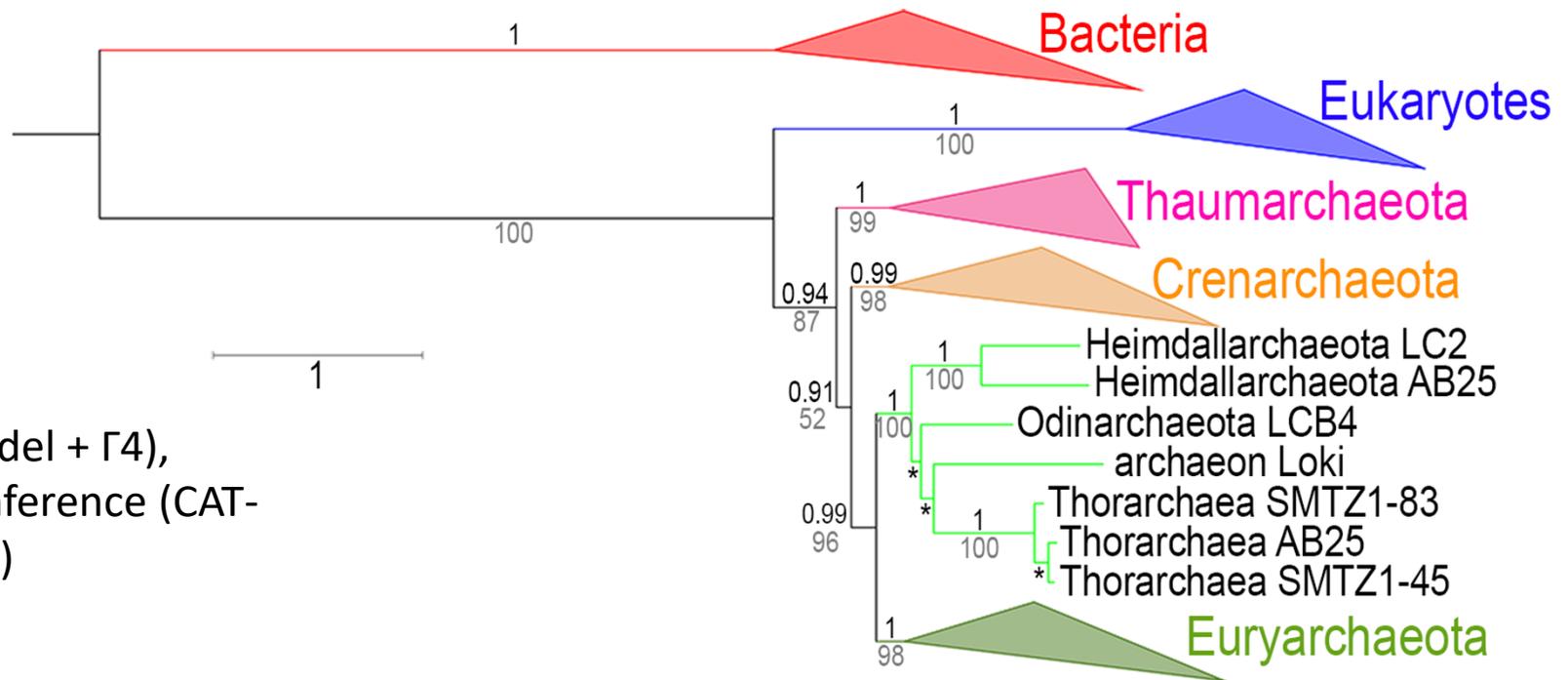
# L'arbre universel du vivant

→ l'arbre de RNA polymérase :

- (i) la protéine universelle la plus longue (1200 AA pour chacune de deux sous-unités principales)
- (ii) Phylogénie d'archées congruente avec la phylogénie des protéines ribosomales



# L'arbre universel du vivant



RNA pol,  
ML (LG model +  $\Gamma$ 4),  
Bayesian inference (CAT-  
GTR model)

→ Asgard archaea forment un groupe monophylétique avec les Euryarchées!

# L'arbre universel du vivant

## Lokiarchaeota: eukaryote-like missing links from microbial dark matter?

Arshan Nasir<sup>1</sup>, Kyung Mo Kim<sup>2</sup>, and Gustavo Caetano-Anollés<sup>3</sup>

<sup>1</sup> Department of Biosciences, COMSATS Institute of Information Technology, Islamabad, Pakistan

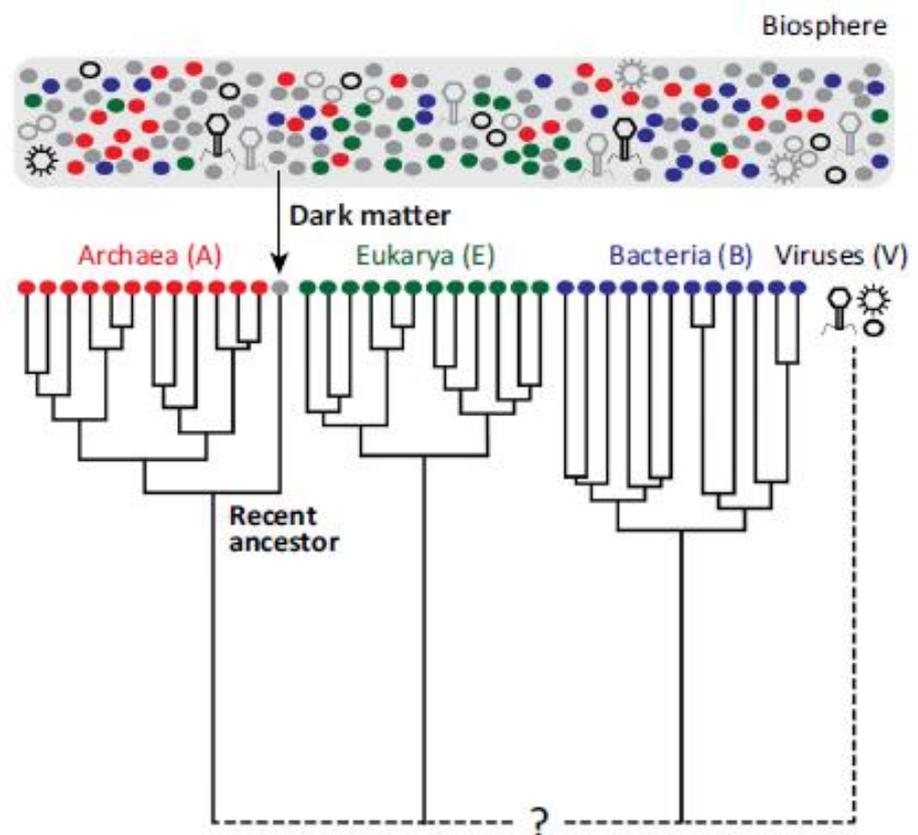
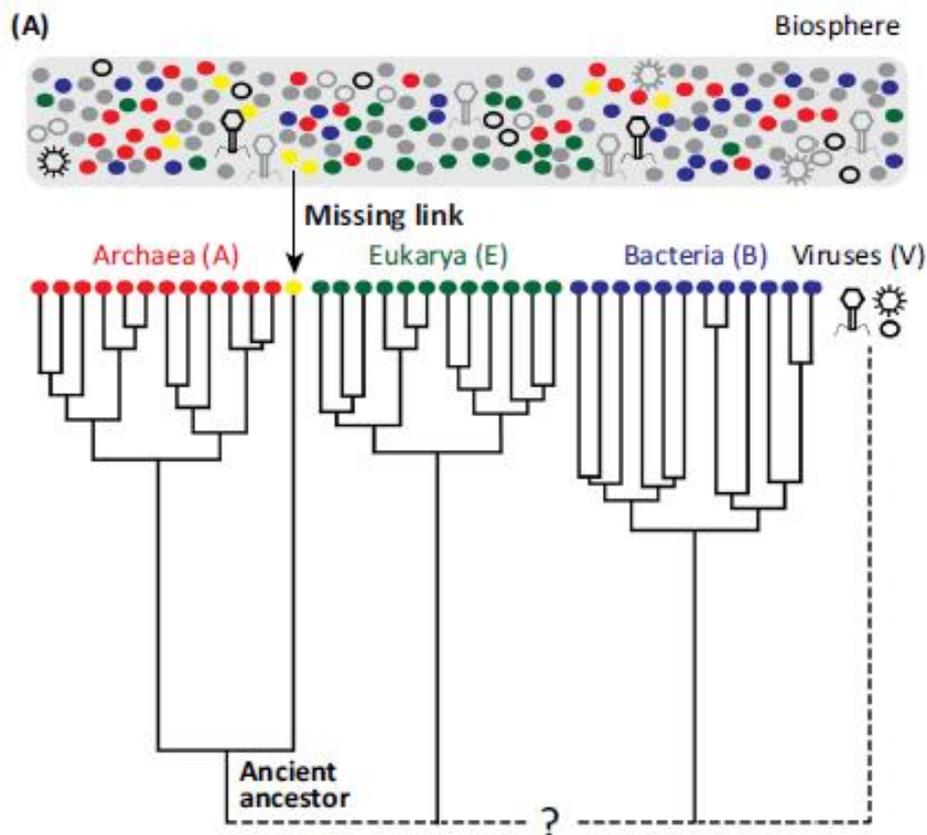
<sup>2</sup> Microbial Resource Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea

<sup>3</sup> Evolutionary Bioinformatics Laboratory, Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA



*Arshan Nasir*

→ Les Loki sont le lien évolutif entre les archées et les eukaryotes ou alors simplement des archées issus de « la matière noire microbienne »?



- la vaste majorité de la diversité microbienne reste à présent inconnue (microbial dark matter)
- la partie connue est utilisé pour établir les arbres phylogénétiques
- «missing link » un organisme qui comporte des caractères anciens, unifiant un grand groupe d'organismes (ex. archées et les eucaryotes ) → car caractères partagés communs

# L'arbre universel du vivant

-comment savoir si un organisme est un « missing link »?

→ Les phylogénies moléculaires: biais dans l'échantillonnage des organismes, les modèles d'évolution de séquences inadéquates, les hypothèses ad hoc pour enraceriner les arbres

→ la distribution des domaines protéiques conservés (DPC) qui ont une structure 3D et fonction en commun (homologues)

**Conserved domain model,**  
based on multiple sequence alignment of related proteins spanning a variety of organisms to reveal sequence regions that contain the same, or similar, patterns of amino acids.

Wide range of organisms represented in multiple sequence alignment

**cd00064: FU** Furin-like repeats.  
Cysteine rich region. Exact function of the domain is not known. Furin is a serine-kinase dependent proprotein processor. Other members of this family include endoproteases and cell surface receptors.

Sequence Alignment

Reformat Format: Hypertext Row Display: up to 10 Color Bits: 2.0 bit Type Selection: the most diverse members

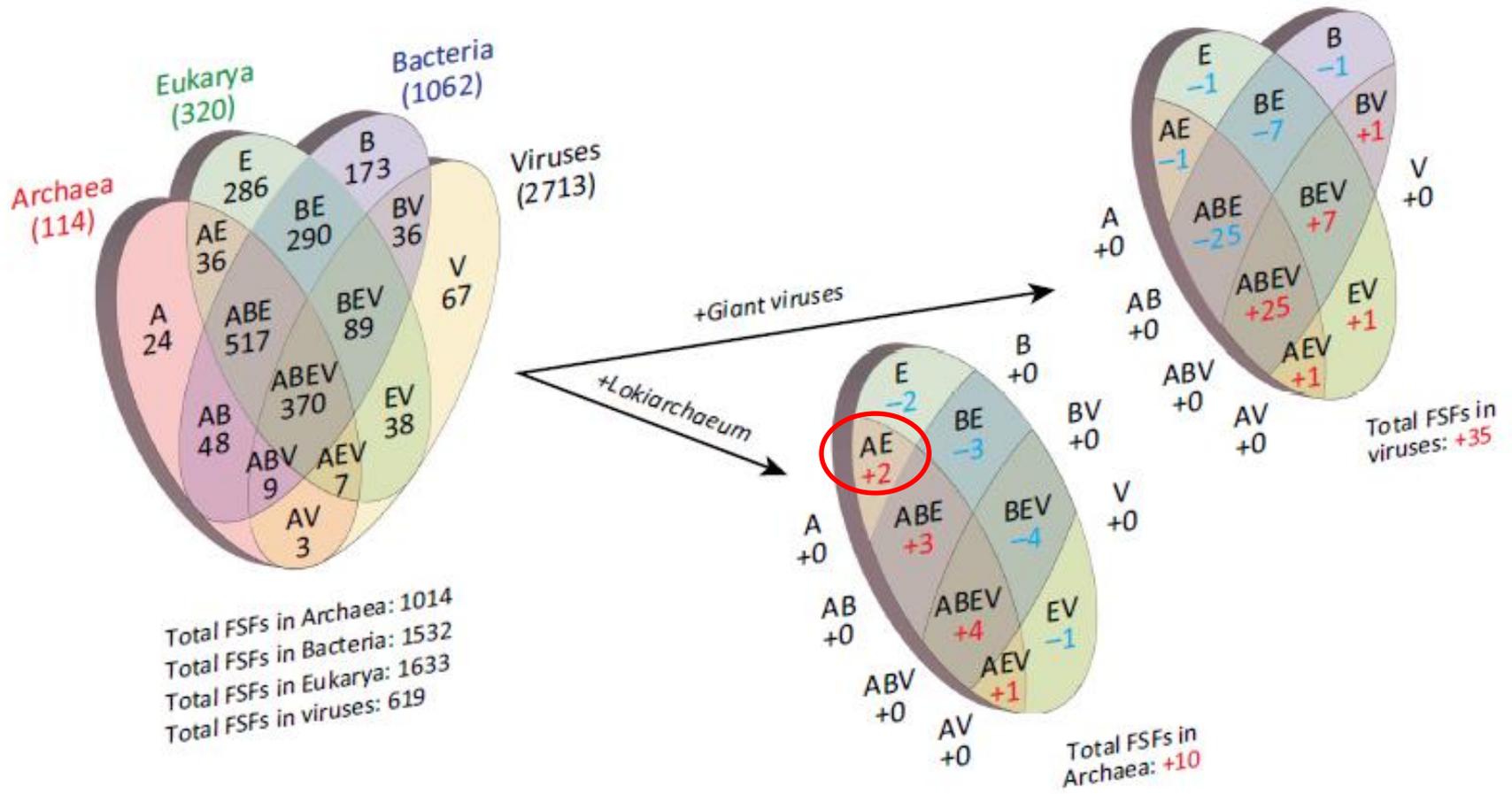
	10	20	30	40	50	60	70	80										
1IGR_A	200	CKMP	ECLGSCS	Apdn	dtAcVACR	hyy	yaqV	VP	ACFpntyz	---Feg	243							
g1 17448068	426	ECKIe	HCE-ACF	Sh	nFcRC	Kegly	lhkgRC	V	ACFpegesaa	---ngt	469							
g1 1749783	614	ACFDe	NCK-SC	Thd	dVCeEC	Adgfs	ldngKCV	sa	tnKSG	lstgaigisva	alvvvgglvq	674						
g1 283446	605	LCSVg	GCK-SC	Vhd	nICgEC	Nsgfs	ldngKCV	saq	---RS	lsagaigisva	vvavvgglva	665						
g1 630804	733	PCHS	---SQ	ECLE	Hea	---dCt	KCP	qdfs	---llng	TCV	qkvkik	---F	Flens	---av	779			
g1 423565	1023	HCPE	---RCQ	DCI	Hea	---kTc	ECM	pdff	---lynd	MCH	r	---SC	---fyp	---dm	1063			
g1 453659	718	PCDV	---GCK	HCE	Ssa	---lHc	tACR	kyys	---lkdd	KCV	sa	avV	S	Vm	---itls	---ivlisv	770	
g1 2499869	579	FIFPe	---IGCK	ILIS	---	gAc	vVC	Eegfs	---lhqk	NCV	q	---HC	fgfa	---vld	hystend	veira	638	
g1 1098514	448	SCSD	athgit	GCK-KC	Alktl	agetest	Vc	EC	Td	krit	---pgn	ACLE	---QC	pagt	ya	---	ingv	503
g1 423565	660	PCFD	---NCE	LCY	np	---nIc	RC	M	sgy	vi	pph	T	CK	---	ECR	qgef	---sey	705

Seulement 2500 DPC dans la nature – hautement conservés → peuvent être utilisés pour inférer des relations de parenté

# L'arbre universel du vivant

-le « missing link » entre les archées et les eucaryotes devrait comporter des nombreux caractères (DPC) en commun avec les Eucaryotes

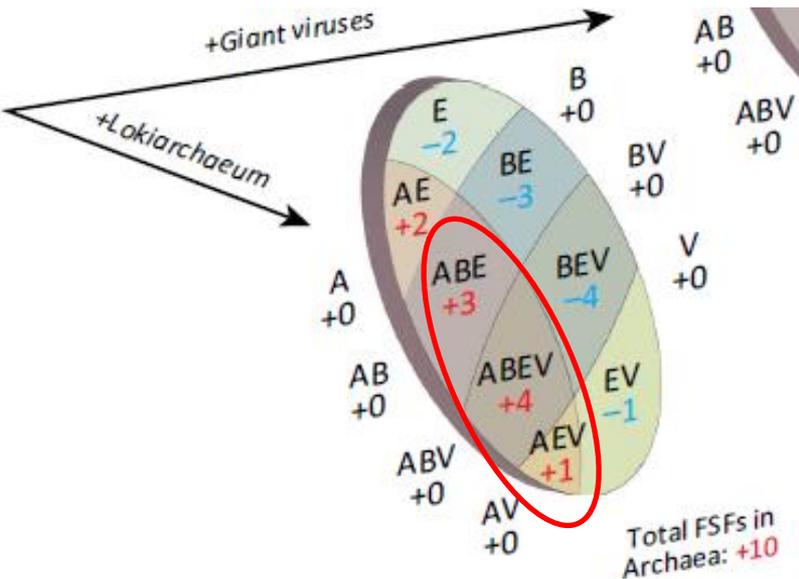
(B)



# L'arbre universel du vivant

-seulement 2 nouveaux DPC en commun avec les eucaryotes!

→ d'autres DPC ont été identifiés: impliqués dans le cytosquelette, ubiquitination des protéines...



...mais on les retrouve chez les bactéries (B) et les virus (V)

-Loki contribuent seulement 10 DPC aux archées ( $10^{14} \rightarrow 10^{24} = 0.1\%$ )

# L'arbre universel du vivant

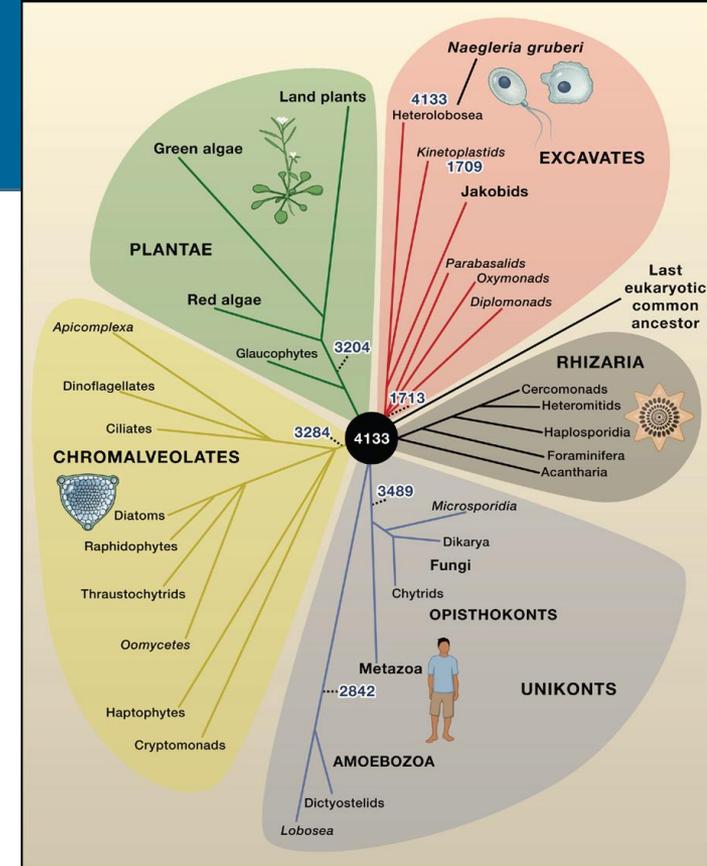
→ l'origine de la complexité chez les Eucaryotes?

-4133 gènes sont partagés entre *N. gruberi* et au moins un supergroupe d'Eucaryotes

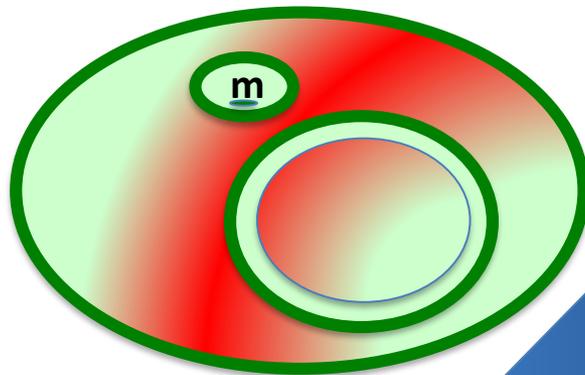
-suggère que LECA était très complexe

**-au moins 40 % des protéines chez LECA sont des ESP!  
(pas d'homologues chez les A et B)**

-Comment expliquer l'existence des nombreux gènes ESP et des structures très complexes par les théories de fusion???



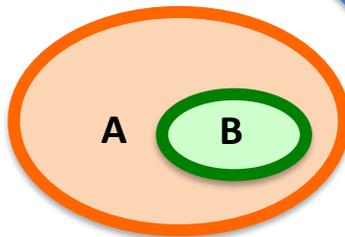
# L'arbre universel du vivant



LECA



FME



All Eukaryotic Specific feature should have appeared after the fusion and before LECA, *all intermediate stages have disappeared*

**Spliceosome,**  
**mRNA capping**

transcriptional machineries with unique components, such as the **mediator**, **endoplasmic reticulum** and derived structures such as **lysosomes**, the **Golgi apparatus** and the nuclear membrane

**nuclear pores**, **nucleolus** and other nucleus-specific structures,

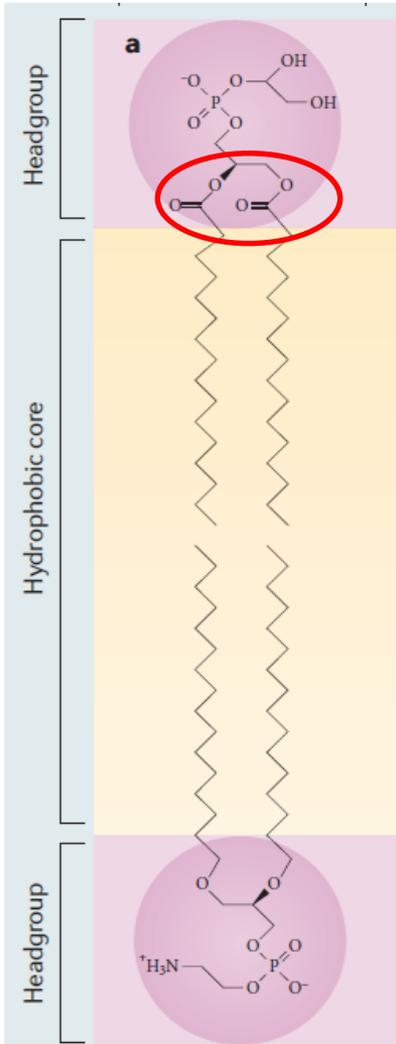
**linear chromosomes with centromeres and telomeres**,

**mitosis** and associated chromosome segregation system

**sex with meiosis** derived from mitosis, **centrioles** and **mid-bodies** for cell division

# Archaea - les lipides membranaires

Bacteria/Eukaryotes



sn-glycerol-3-phosphate

liaison ester

acides gras

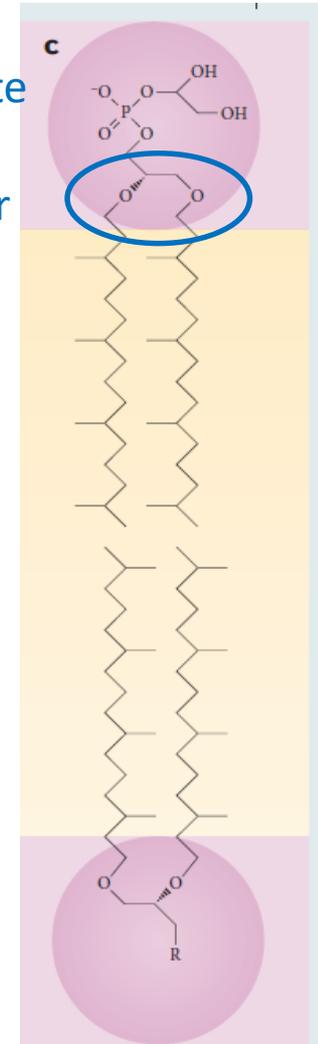
D stéréo-isomère L

sn-glycerol-1-phosphate

liaison éther

chaîne  
latérale  
composé  
d'isoprène  
méthylé

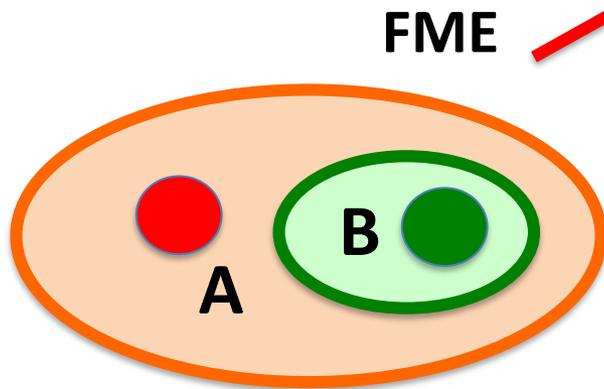
Archaea



→ les lipides d'Archaea sont fondamentalement différents!

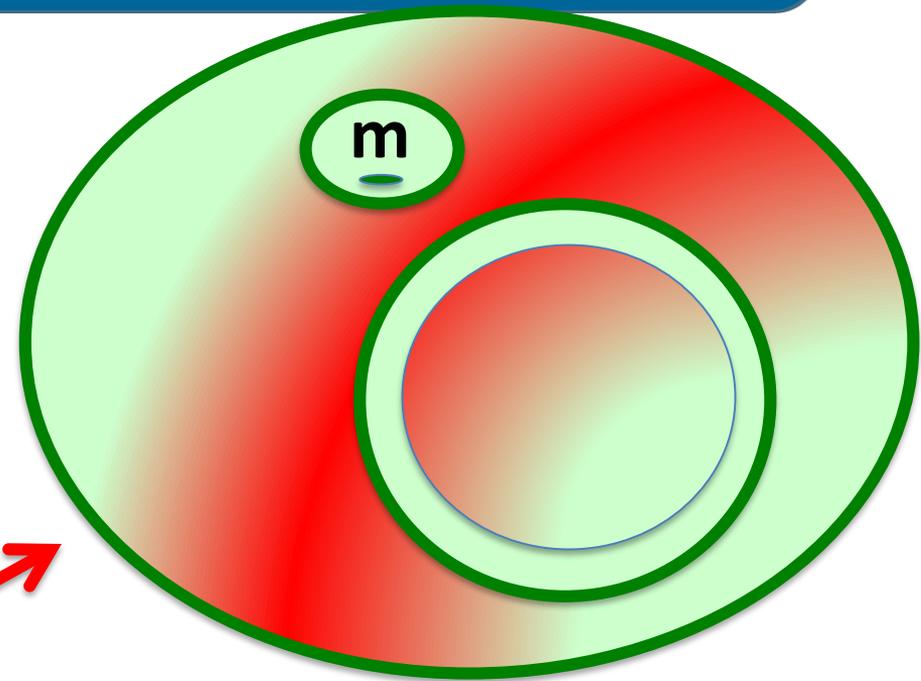
# L'arbre universel du vivant

→ transformation of the host archaeal membrane into the bacterial-like membrane of eukaryotes. Selection pressure?



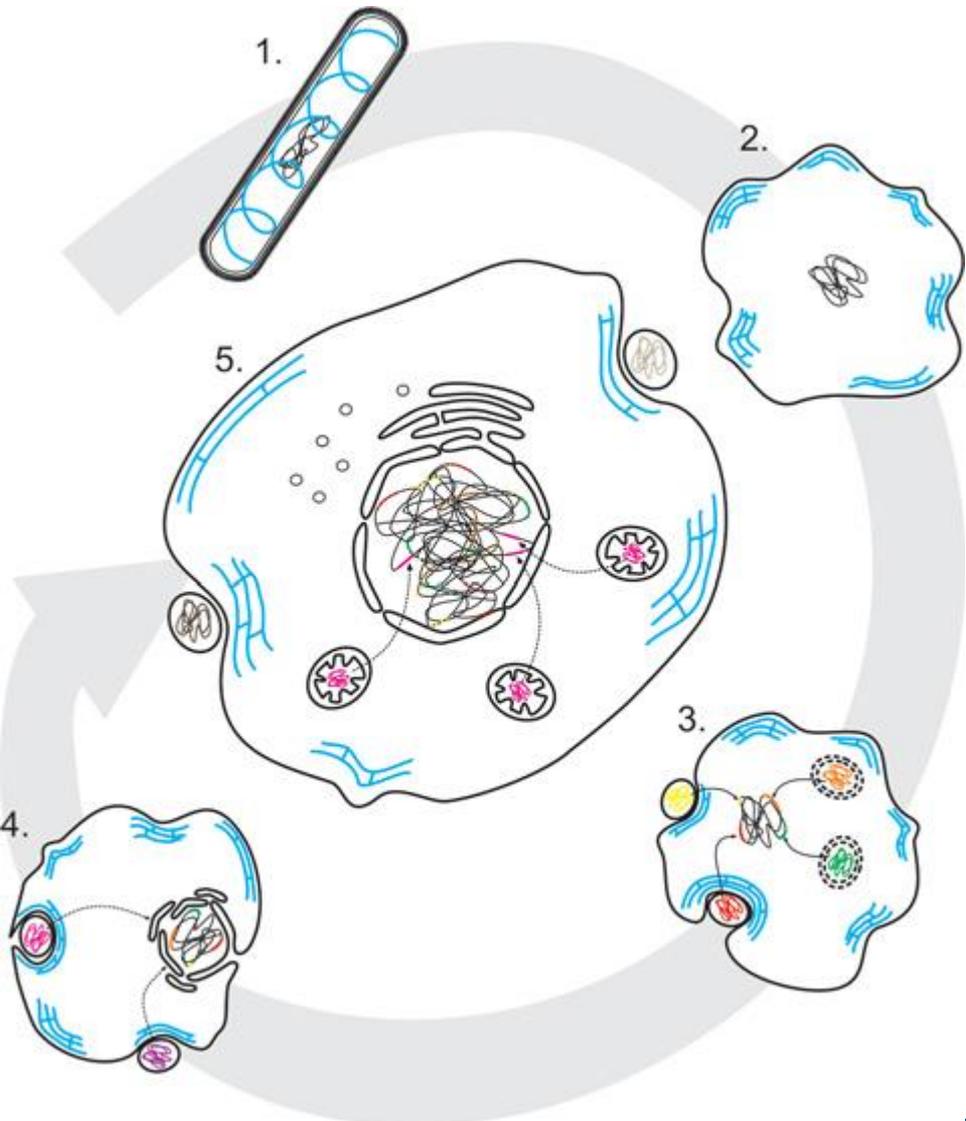
**FME**

**LECA**



- such transformation has never been observed in nature.
- acquisition of a bacterial type of membrane never occurred in any archaeal lineage, despite the massive lateral gene transfer (LGT) of bacterial genes into archaea

# L'arbre universel du vivant



n'explique pas pourquoi l'acquisition des gènes uniquement dans un sens (du symbionte vers le hôte)?

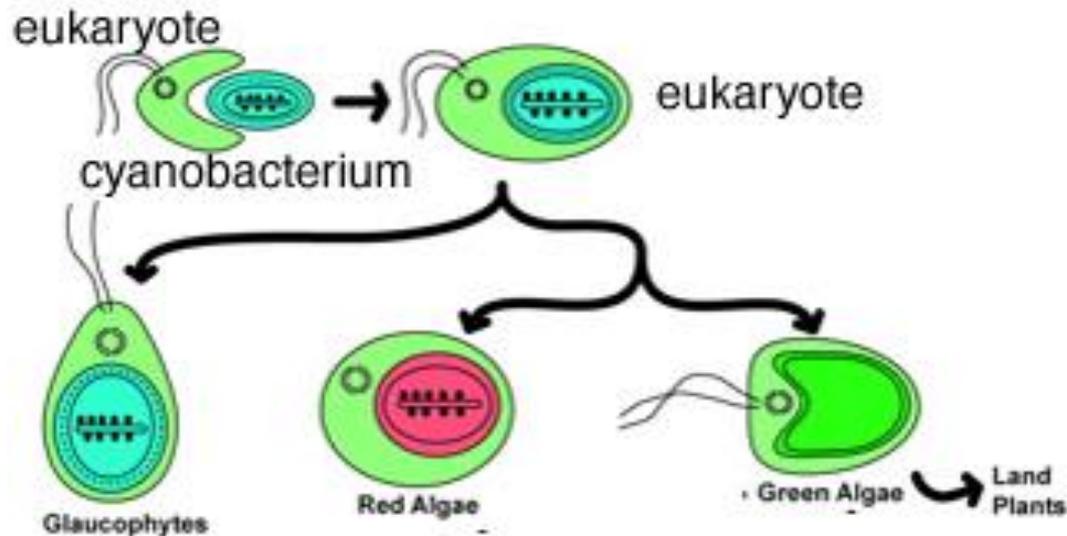
pourquoi la transformation du chromosome circulaire en chromosomes linéaires pourvus de centromères et de télomères ?

pourquoi la machinerie de division simple de l'hôte a été remplacé par le mitose très complexe ?

# L'arbre universel du vivant

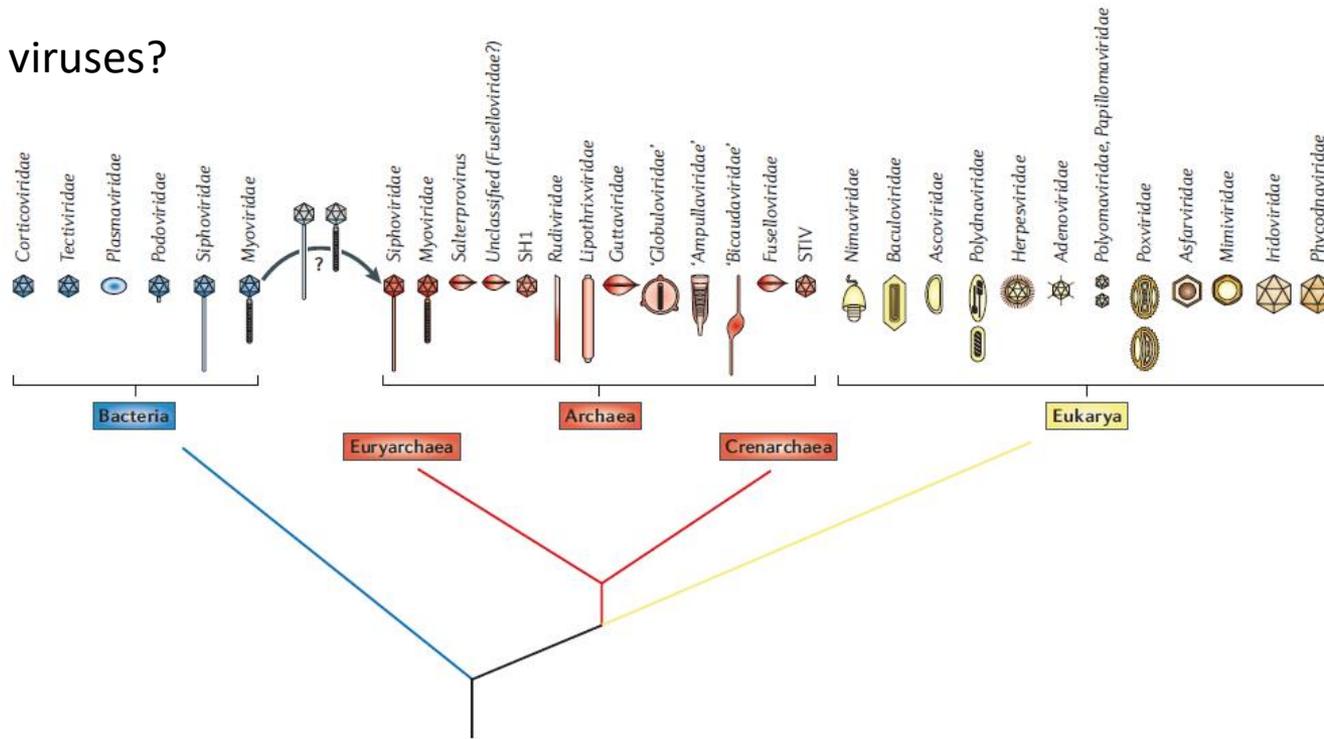
→selon les théories de fusion l'association d'une archée et d'une bactérie a conduit à l'émergence d'un nouveau domaine les Eucaryotes

→On ne connaît pas d'autres exemples d'une telle transformation: l'union d'un eucaryote avec une cyanobactérie n'a pas donné un nouveau domaine; les plantes sont toujours les eucaryotes et les chloroplastes sont reconnaissables en tant que les anciennes bactéries



# L'arbre universel du vivant

How about viruses?

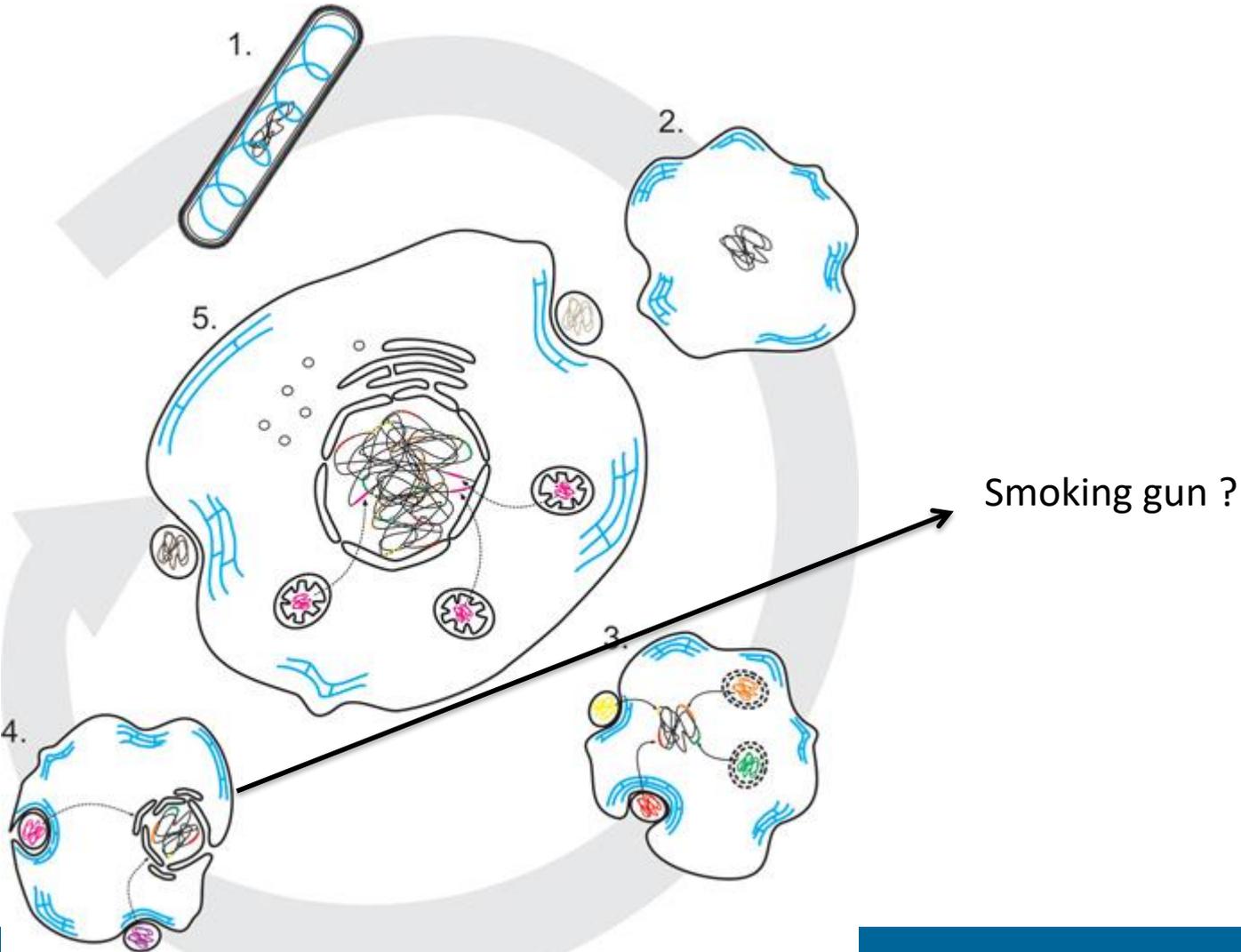


→ each domain has its own « collection » of viruses

→ The fusion theory does not explain how the eukaryotic viruses arose from archaeal ones

# L'arbre universel du vivant

La théorie 'PhAT (phagocytosing archaeon theory)' pour l'émergence de la cellule eucaryote



## Article

# Isolation of an archaeon at the prokaryote–eukaryote interface

<https://doi.org/10.1038/s41586-019-1916-6>

Received: 6 August 2019

Accepted: 5 December 2019

Published online: 15 January 2020

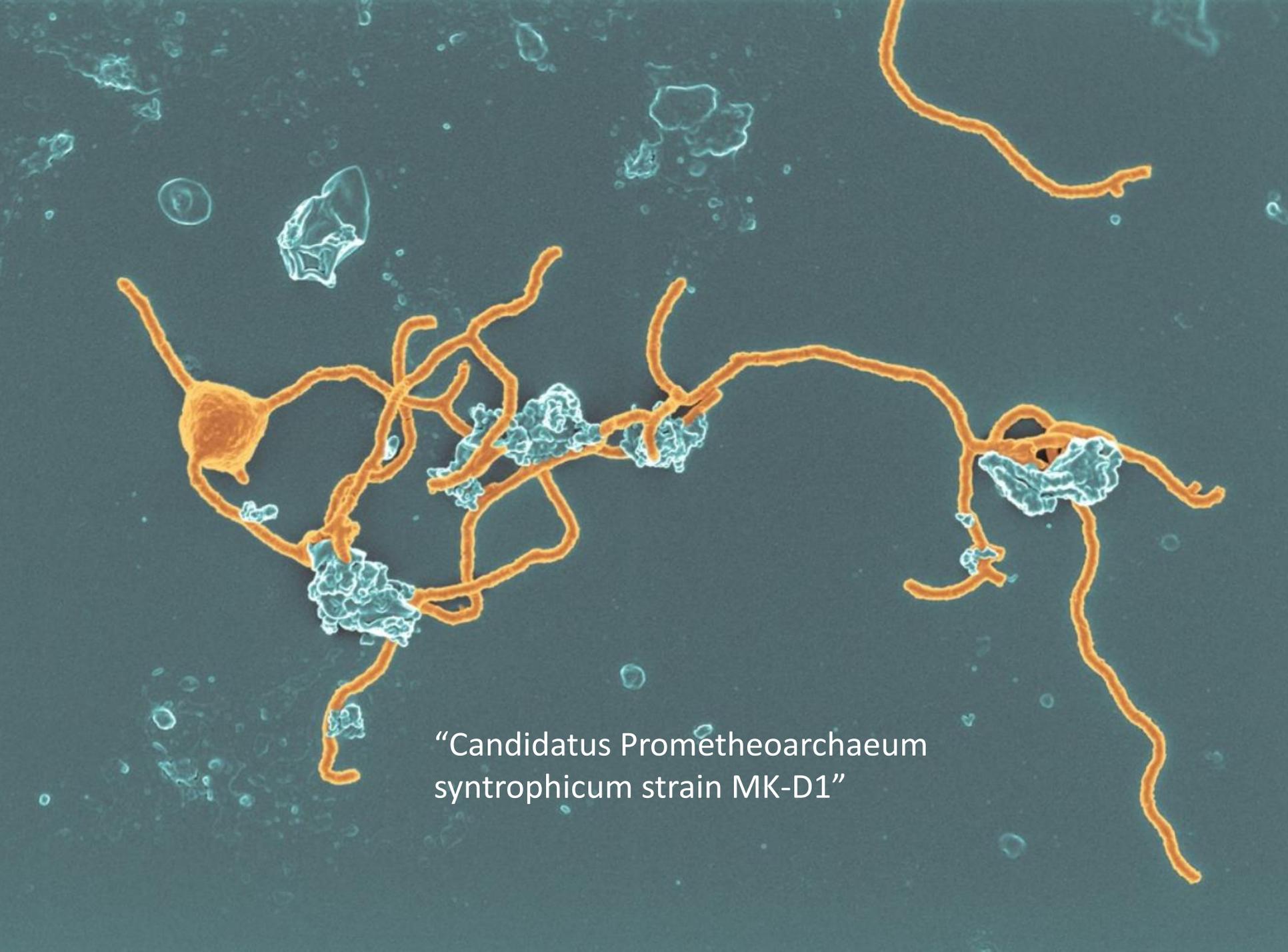
Open access

Hiroyuki Imachi<sup>1,11\*</sup>, Masaru K. Nobu<sup>2,11\*</sup>, Nozomi Nakahara<sup>1,2,3</sup>, Yuki Morono<sup>4</sup>, Miyuki Ogawara<sup>1</sup>, Yoshihiro Takaki<sup>1</sup>, Yoshinori Takano<sup>5</sup>, Katsuyuki Uematsu<sup>6</sup>, Tetsuro Ikuta<sup>7</sup>, Motoo Ito<sup>4</sup>, Yohei Matsui<sup>8</sup>, Masayuki Miyazaki<sup>1</sup>, Kazuyoshi Murata<sup>9</sup>, Yumi Saito<sup>1</sup>, Sanae Sakai<sup>1</sup>, Chihong Song<sup>9</sup>, Eiji Tasumi<sup>1</sup>, Yuko Yamanaka<sup>1</sup>, Takashi Yamaguchi<sup>3</sup>, Yoichi Kamagata<sup>2</sup>, Hideyuki Tamaki<sup>2</sup> & Ken Takai<sup>1,10</sup>

The origin of eukaryotes remains unclear<sup>1</sup>. Current data suggest that eukaryotes

« La chance ne sourit qu'aux esprits bien préparés. »

Louis Pasteur



“Candidatus Prometheoarchaeum  
syntrophicum strain MK-D1”

# L'arbre universel du vivant

Human Occupied Vehicle *SHINKAI 6500*

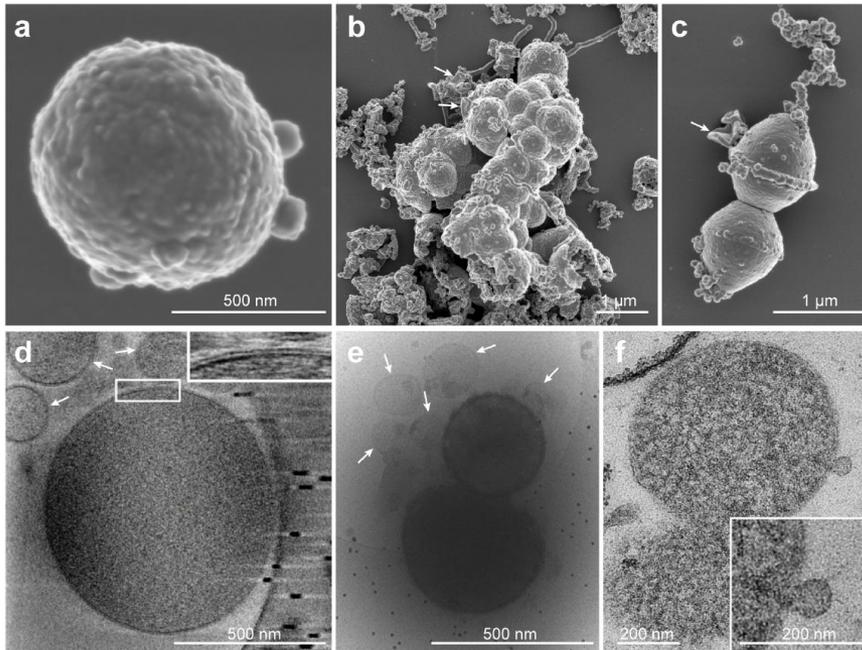


*SHINKAI 6500* is a submersible that can dive to a depth of 6,500 m. Crewed by a maximum of three people within *SHINKAI 6500*'s titanium pressure sphere, three windows allow researchers to observe the deep sea with their own eyes.



down-hanging sponge reactor  
inoculated with deep-sea  
sediments

# L'arbre universel du vivant



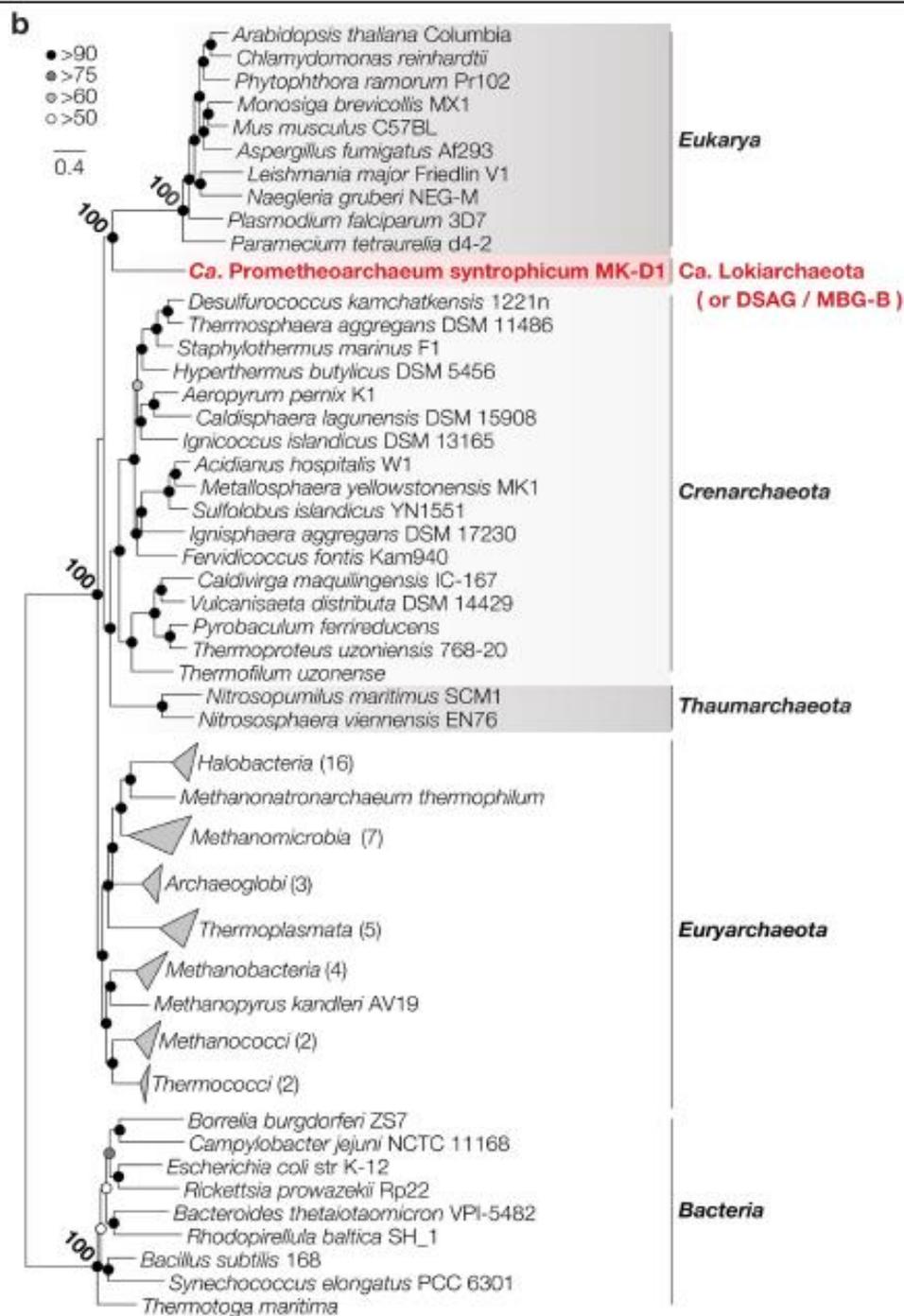
→ 12 years isolation effort!!!

→ 30–60 days of lag phase → required over 3 months to reach full growth with a yield of  $\sim 10^5$  16S rRNA gene copies/ml

→ Anaerobic, AA-oxidizing archaeon, small cocci, ca. 550 nm in diameter, syntrophically grows with hydrogen- and formate-utilizing microorganisms. It produces MVs, chains of blebs, and membrane-based protrusions.

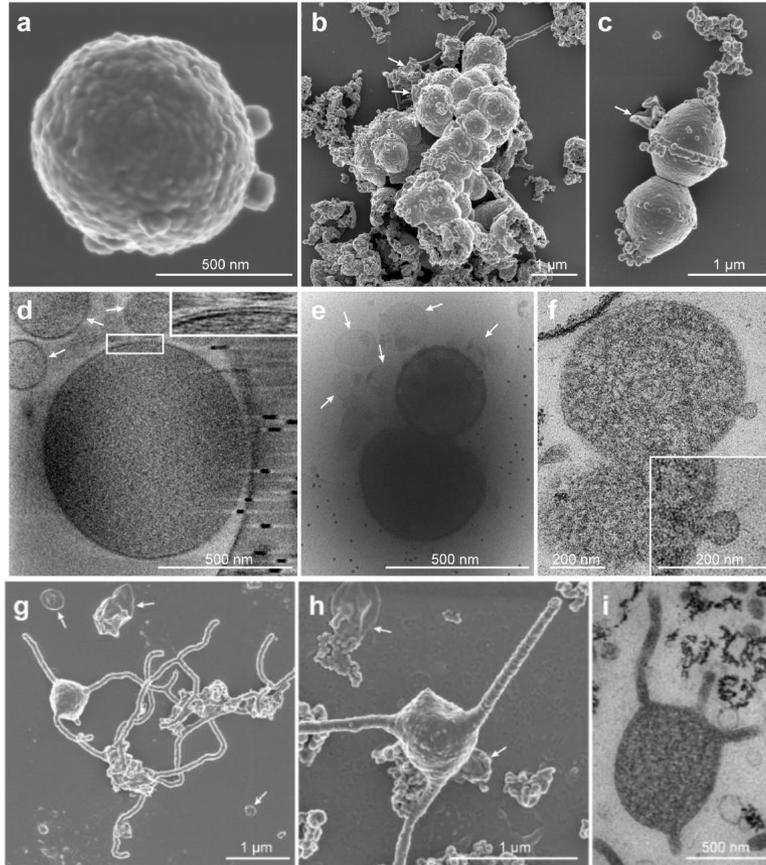
# L'arbre universel

Phylogénie  
bayésienne,  
protéines  
ribosomales



# L'arbre universel du vivant

→ “Candidatus *Prometheoarchaeum syntrophicum* strain MK-D1”

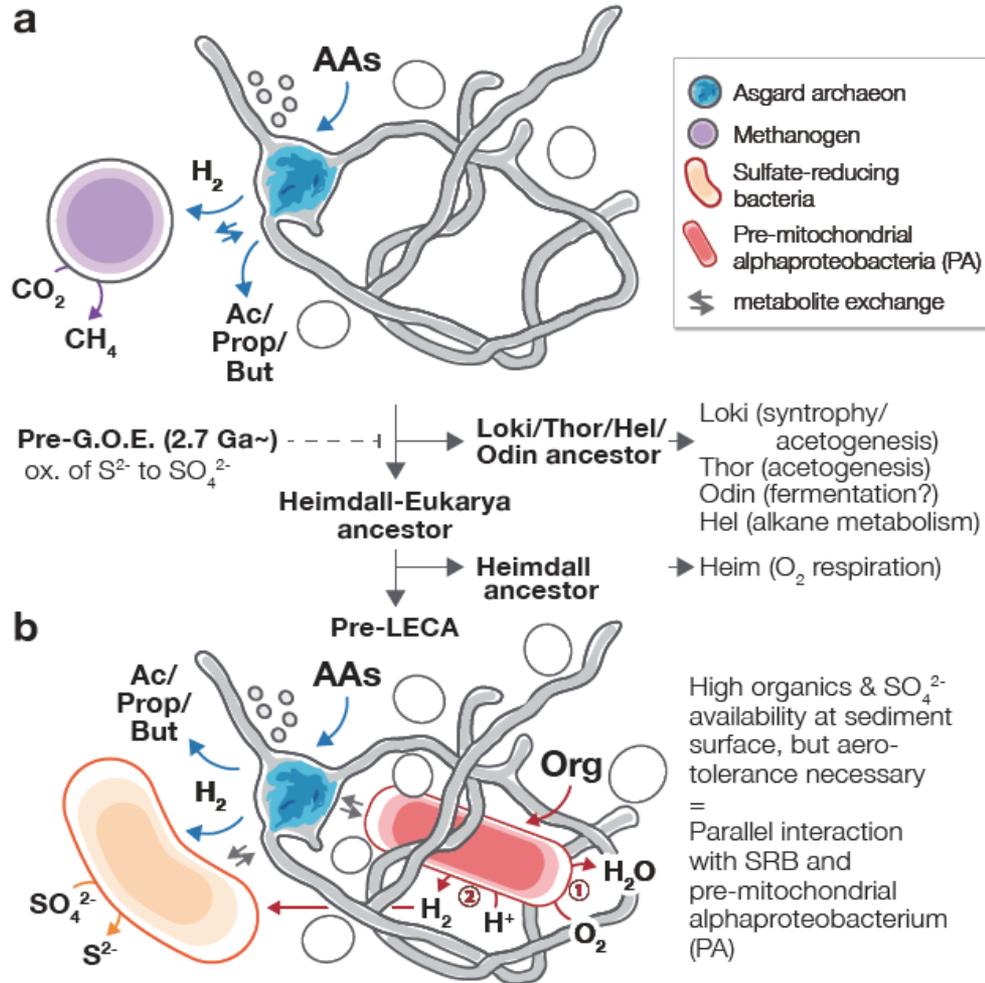


→ Endomembranes ?

→ Phagocytosis ?

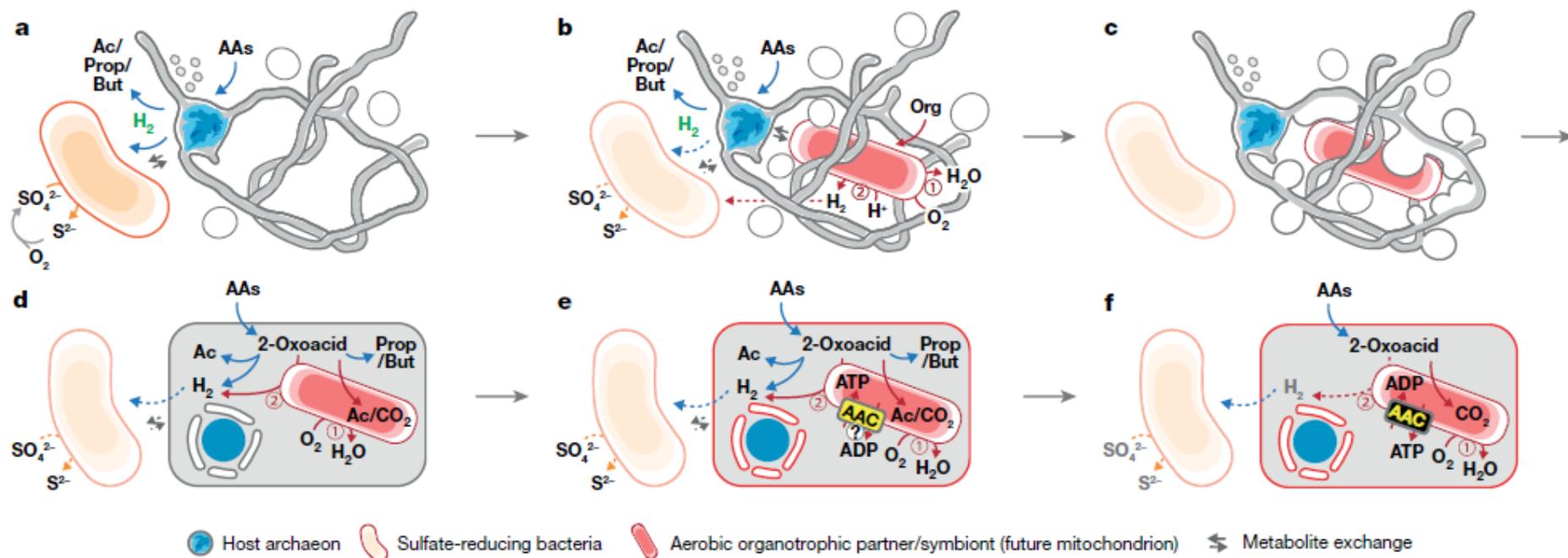
# L'arbre universel du vivant

→ Un « ménage à trois »



# L'arbre universel du vivant

→ Scenario proposé pour l'origine du LECA



# L'arbre universel du vivant

## Anomalous phylogenetic behavior of ribosomal proteins in metagenome assembled asgard archaea

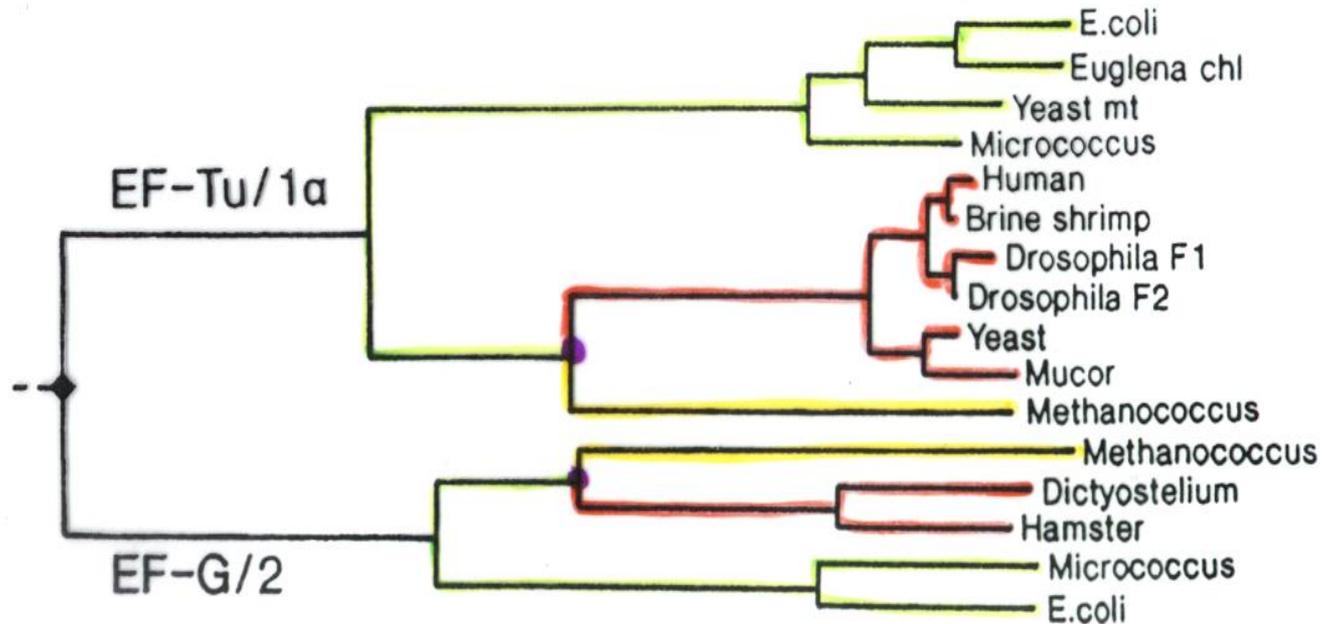
Sriram G. Garg<sup>1\*†</sup>, Nils Kapust<sup>1†</sup>, Weili Lin<sup>2†</sup>, Michael Knopp<sup>1</sup>, Fernando D. K. Triá<sup>1</sup>,  
Shijulal Nelson-Sathi<sup>3</sup>, Sven B. Gould<sup>1</sup>, Lu Fan<sup>4,5,6</sup>, Ruixin Zhu<sup>2</sup>, Chuanlun Zhang<sup>4,5,7</sup>,  
William F. Martin<sup>1\*</sup>

environment. Here we show that Asgard and CPR MAGs fail spectacularly to meet the most basic phylogenetic criterion<sup>8</sup> fulfilled by genome sequences of all cultured prokaryotes investigated to date: the ribosomal proteins of Asgard and CPR MAGs do not share common evolutionary histories. Their phylogenetic behavior is anomalous to a degree never observed with genomes of real organisms. CPR and Asgard MAGs are binning artefacts, assembled from environments where up to 90% of the DNA is from dead cells<sup>9–12</sup>. Asgard and CPR MAGs are unnatural constructs, genome-like patchworks of genes that have been stitched together into computer files by binning.

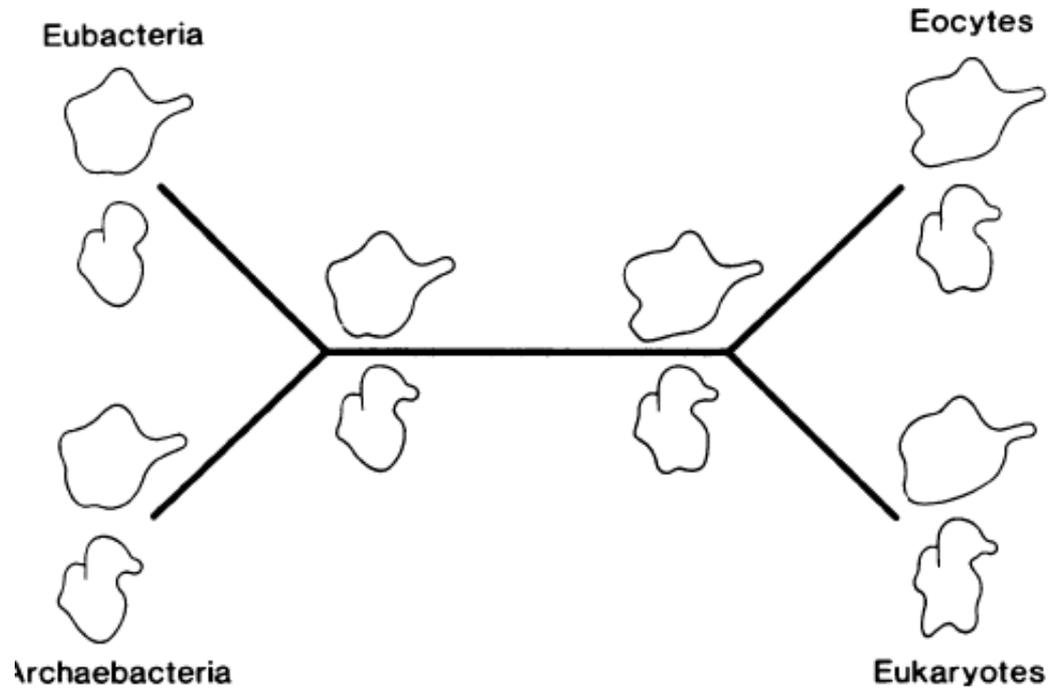
**TO BE CONTINUED...**

# Questions de révision

1. Expliquez comment cet arbre universel a pu être enraciné.

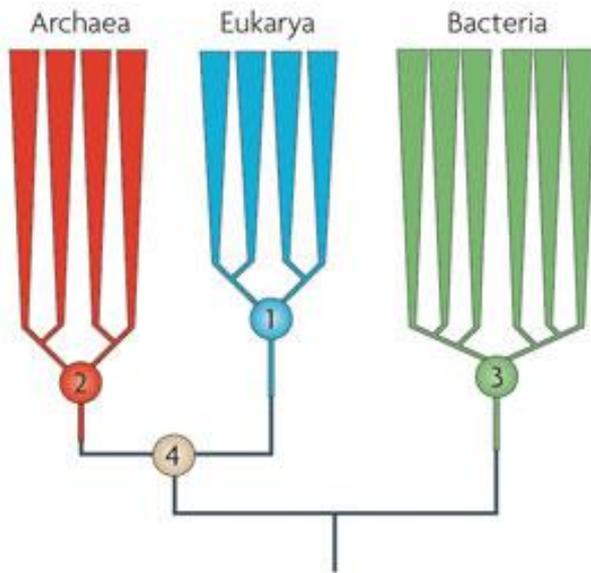


2. Expliquez le scénario « eocyte » proposé par Jim Lake en 1984 en s'appuyant sur le schéma ci-dessous



### 3. Quelle est la différence fondamentale entre la topologie 2D et 3D?

**a** 'Three primary domains' (3D) scenario



**b** 'Two primary domains' (2D) scenario

