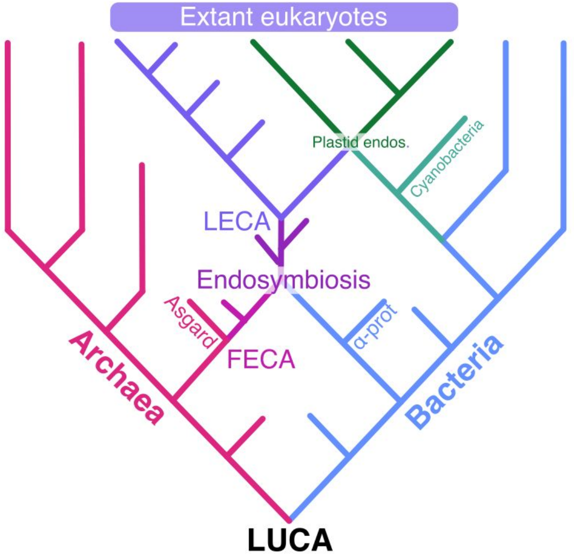
Congress of the European Society for Evolutionary Biology 2025

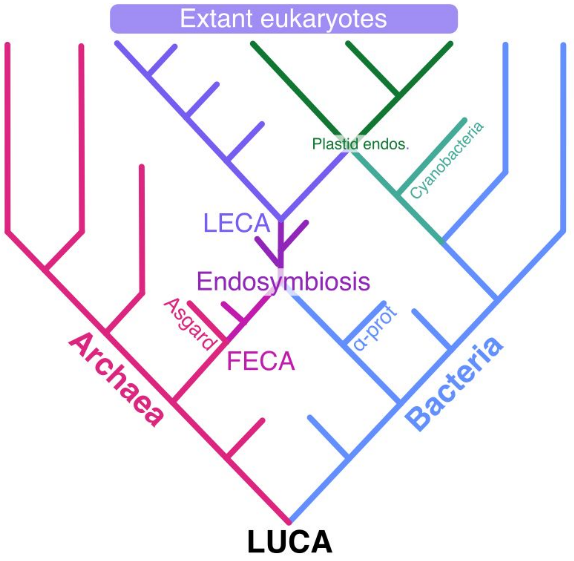
Diverse ancestries reveal complex symbiotic interactions during eukaryogenesis

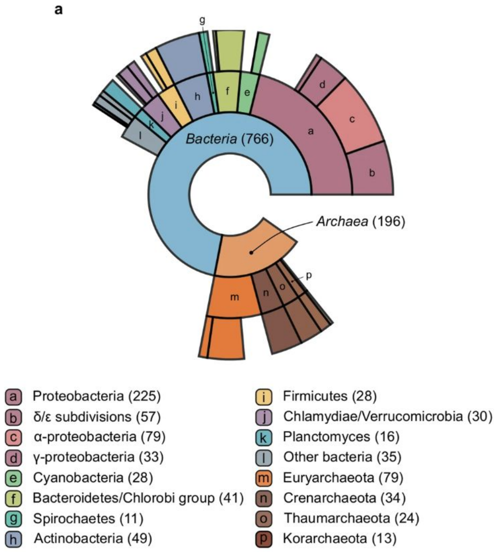
Moisès Bernabeu*, **Saioa Manzano-Morales***, Marina Marcet-Houben*, Toni Gabaldón

Eukaryotes stem from a symbiosis event



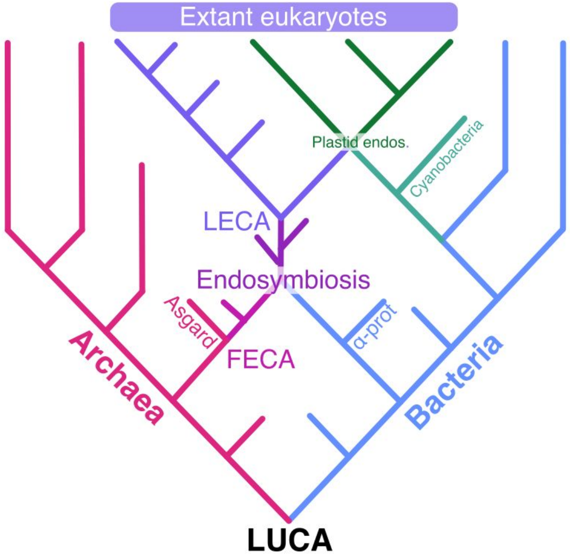
Eukaryotes stem from a symbiosis event, but a binary scenario does not encompass all data

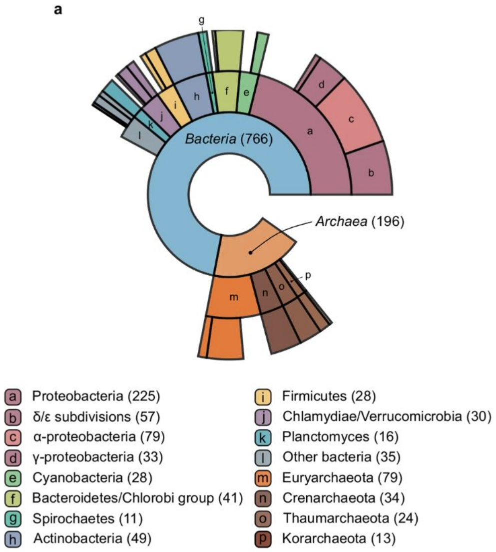




(Pittis & Gabaldón, 2016)

Eukaryotes stem from a symbiosis event, but a binary scenario does not encompass all data - open questions remain

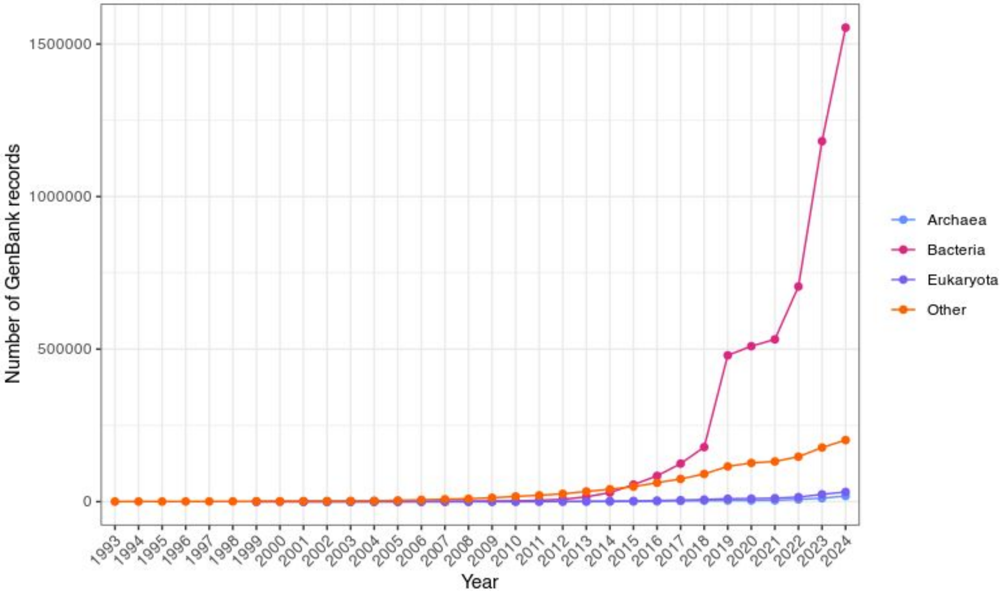




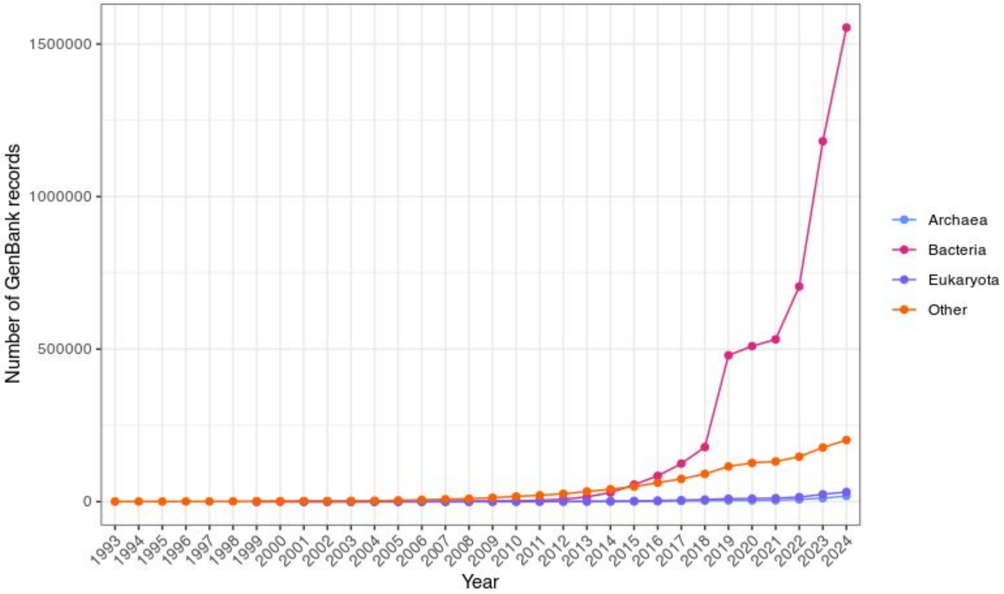
(Pittis & Gabaldón, 2016)

- **Nature** of the additional partners
- **Tempo** of acquisitions
- **Function** in proto-eukaryote

Why now? New (mostly) eukaryotic lineages reshape our understanding of LECA

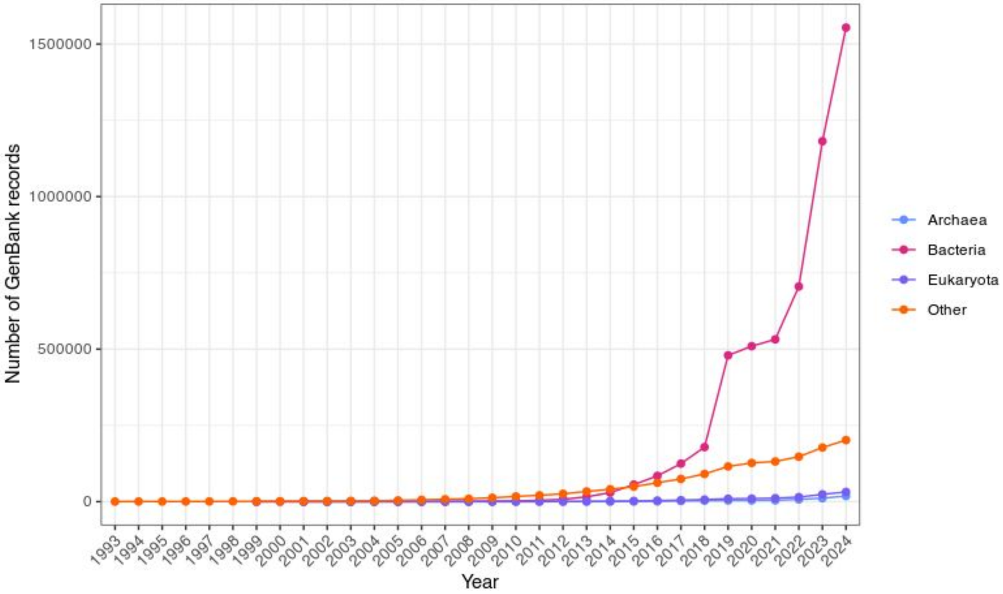


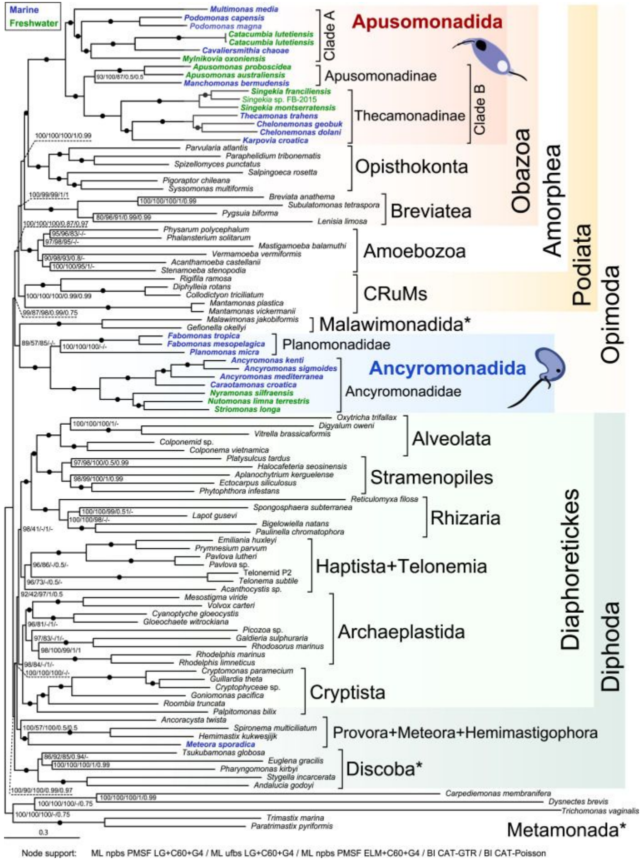
Why now? New (mostly) eukaryotic lineages reshape our understanding of LECA



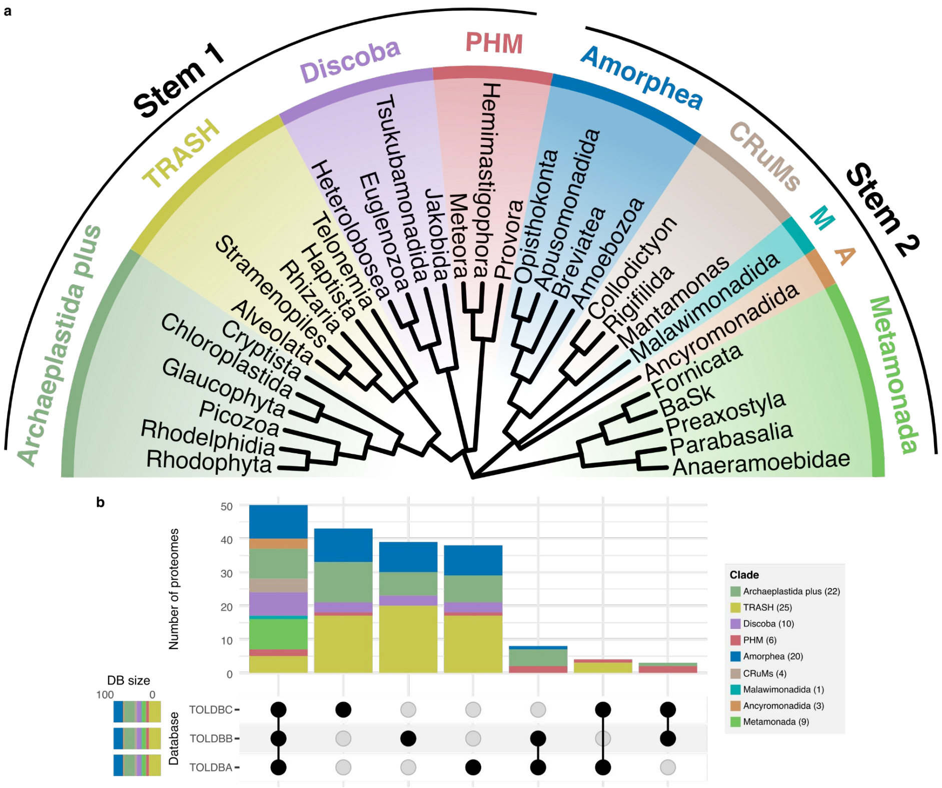
P03.082 “Widespread intra- and inter-domain horizontal transfer and bursts of gene duplication shape the size and content of Asgard archaeal genomes”

Why now? New (mostly) eukaryotic lineages reshape our understanding of LECA

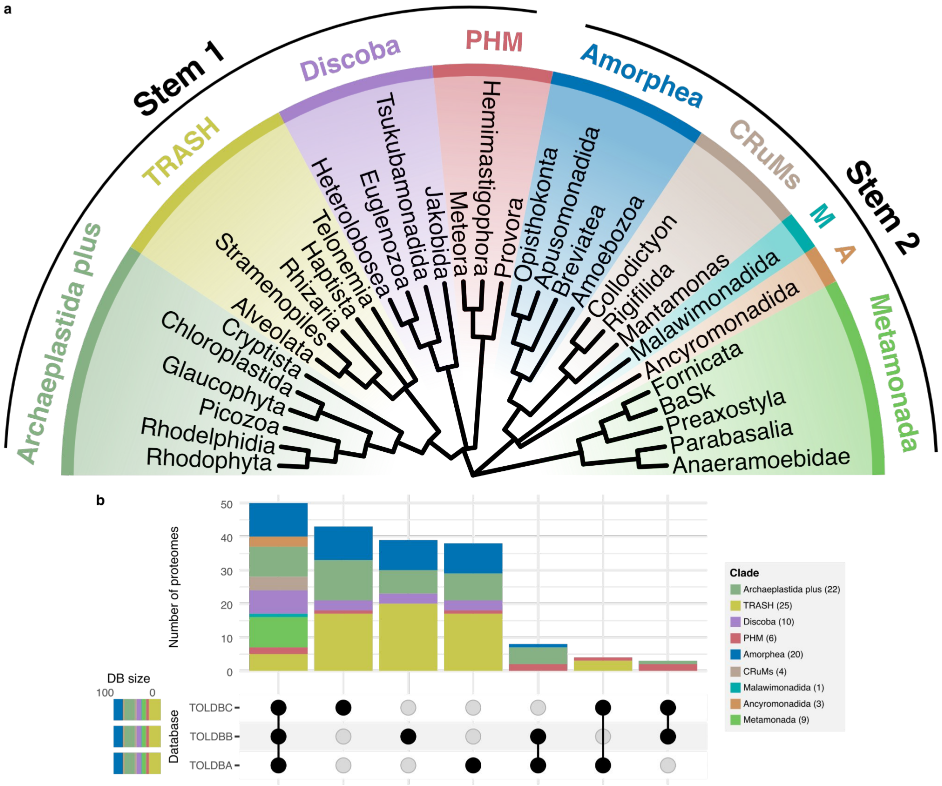




(Torruella *et al.*, 2025)

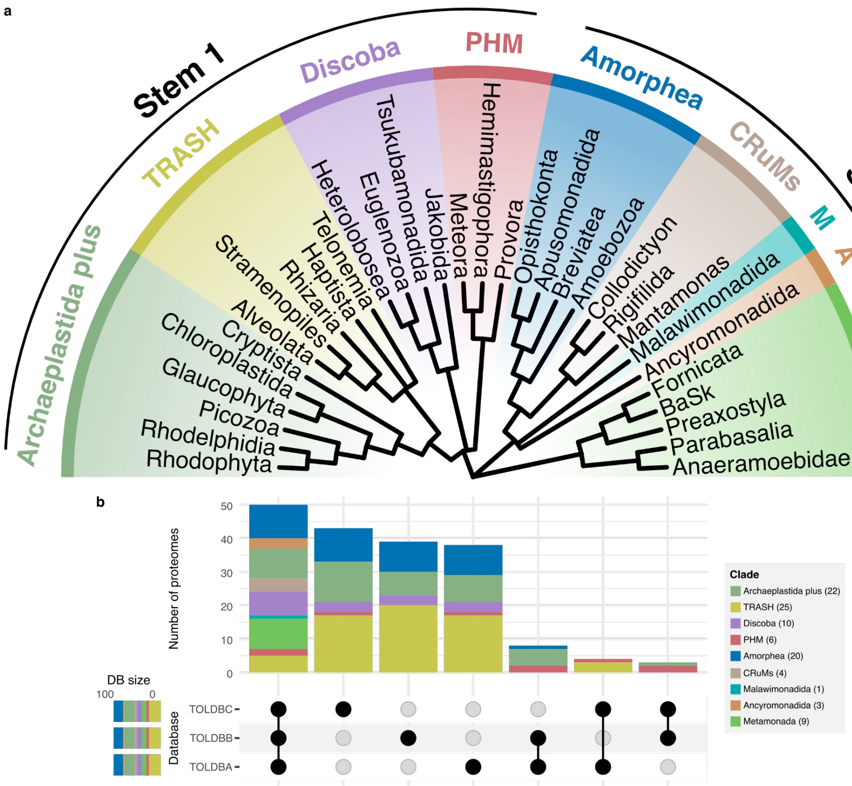


(Bernabeu *et al*., 2024) (in rev.)

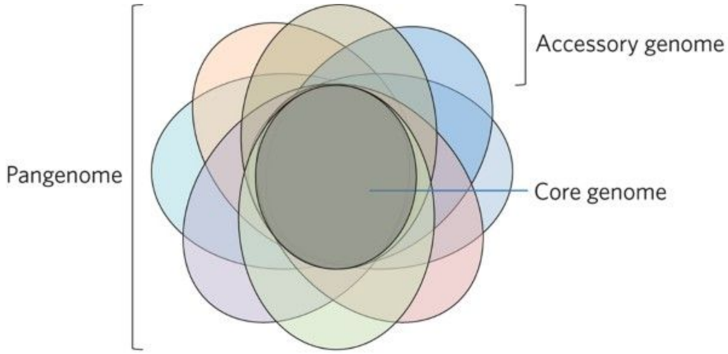


- Reference Viral DataBase (RVDB)
- Genome Taxonomy Database (GTDB)
    - Genus representatives clustered at order level

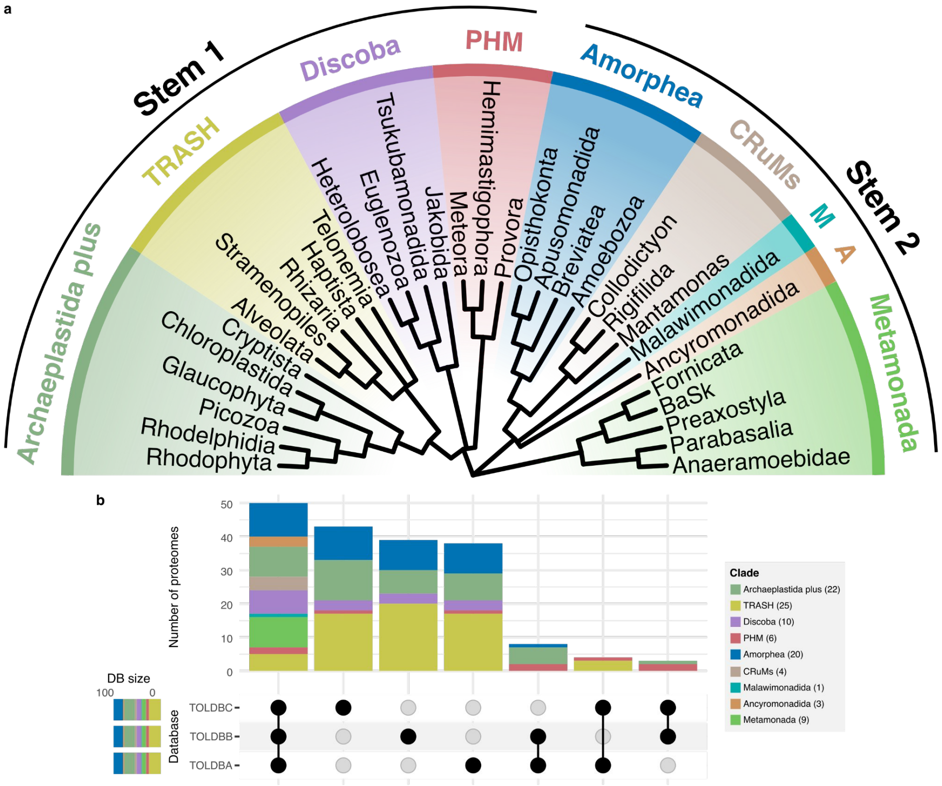
(Bernabeu *et al*., 2024) (in rev.)



- Reference Viral DataBase (RVDB)
- Genome Taxonomy Database (GTDB)
    - Genus representatives clustered at order level
    - Pangenome approach



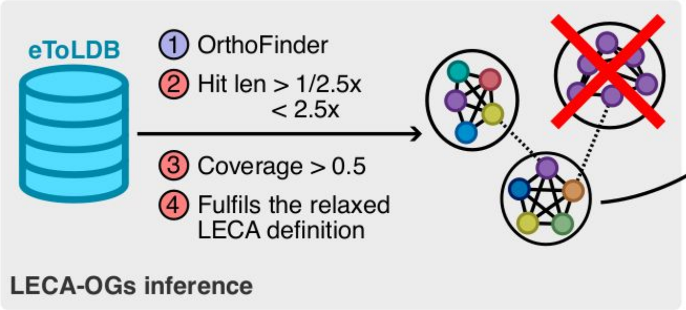
(Bernabeu *et al*., 2024) (in rev.)



(Bernabeu *et al*., 2024) (in rev.)

- Reference Viral DataBase (RVDB)
- Genome Taxonomy Database (GTDB)
  - Genus representatives clustered at order level
  - Pangenome approach

Clustered broad non-eukaryotic db



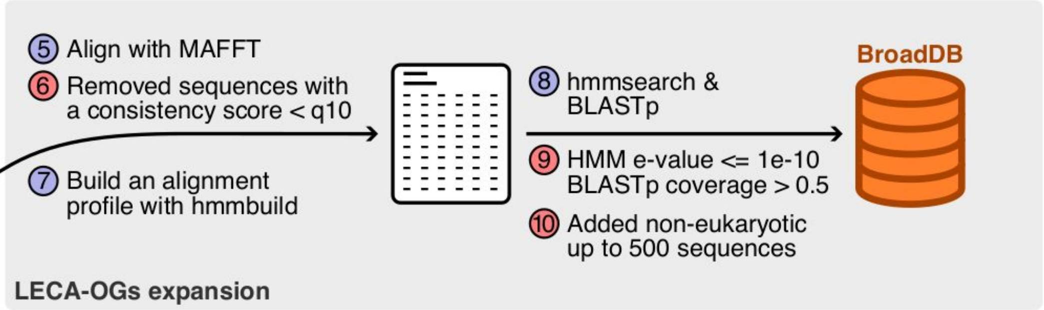
① OrthoFinder

② Hit len > 1/2.5x < 2.5x

③ Coverage > 0.5

④ Fulfils the relaxed LECA definition

LECA-OGs inference



⑤ Align with MAFFT

⑥ Removed sequences with a consistency score < q10

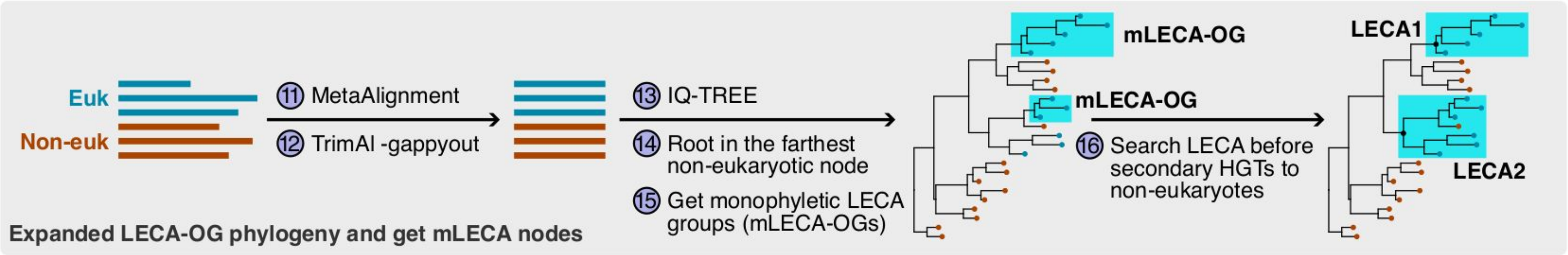
⑦ Build an alignment profile with hmmbuild

⑧ hmmsearch & BLASTp

⑨ HMM e-value <= 1e-10 BLASTp coverage > 0.5

⑩ Added non-eukaryotic up to 500 sequences

LECA-OGs expansion



⑪ MetaAlignment

⑫ TrimAl -gappyout

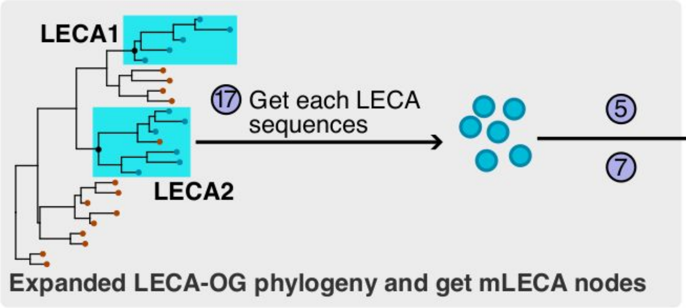
⑬ IQ-TREE

⑭ Root in the farthest non-eukaryotic node

⑮ Get monophyletic LECA groups (mLECA-OGs)

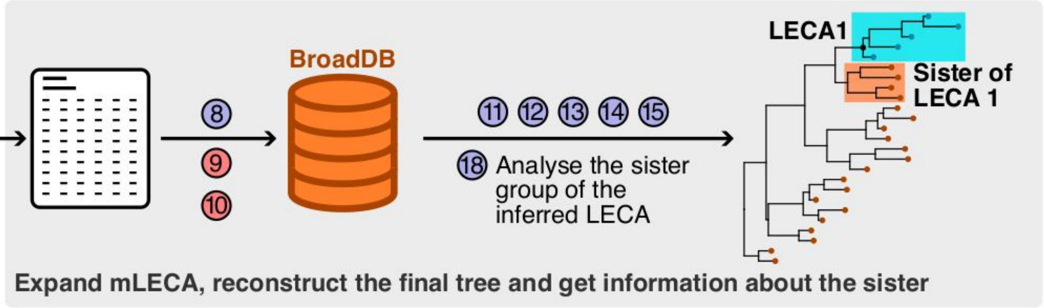
⑯ Search LECA before secondary HGTs to non-eukaryotes

Expanded LECA-OG phylogeny and get mLECA nodes



⑰ Get each LECA sequences

Expanded LECA-OG phylogeny and get mLECA nodes

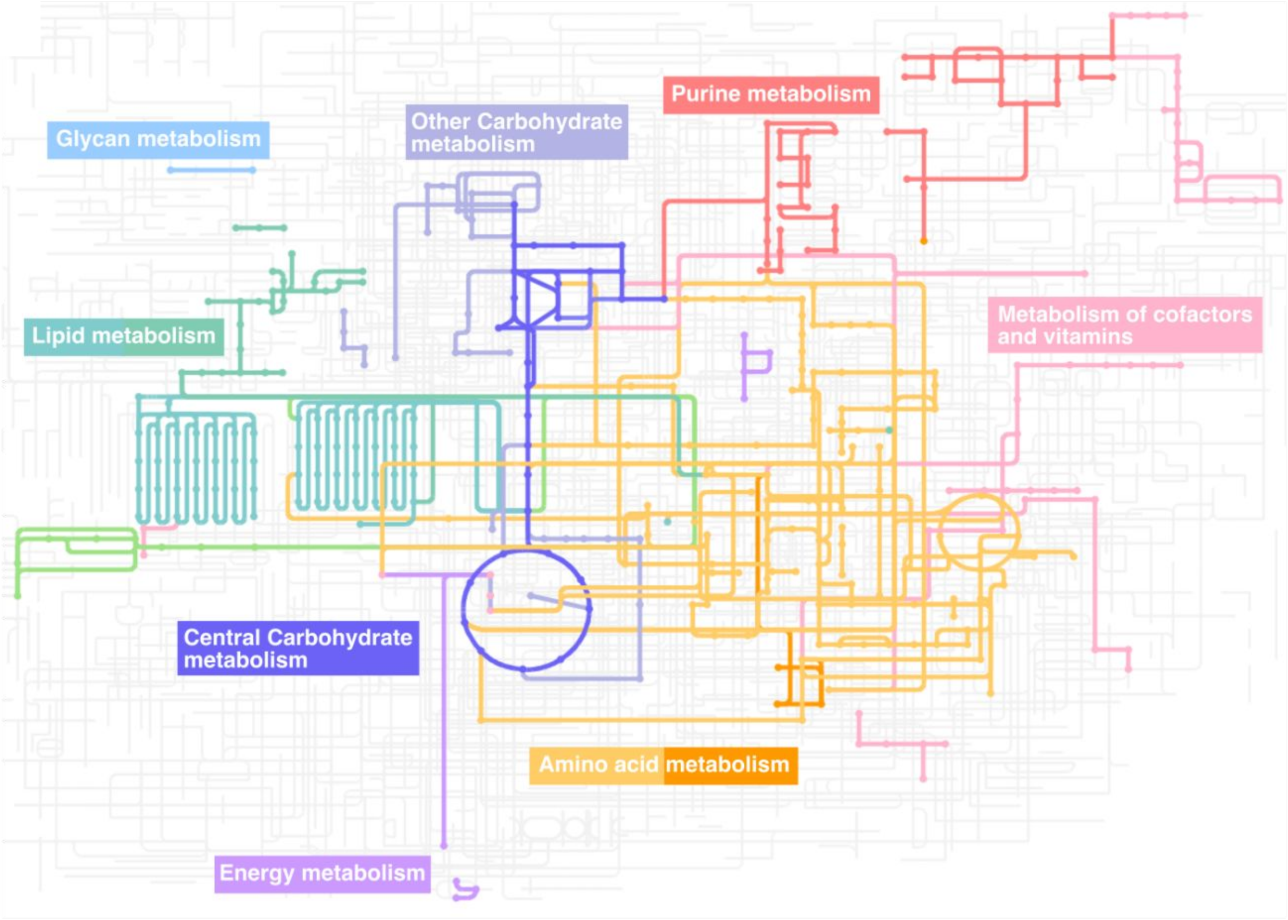


⑱ Analyse the sister group of the inferred LECA

Expand mLECA, reconstruct the final tree and get information about the sister

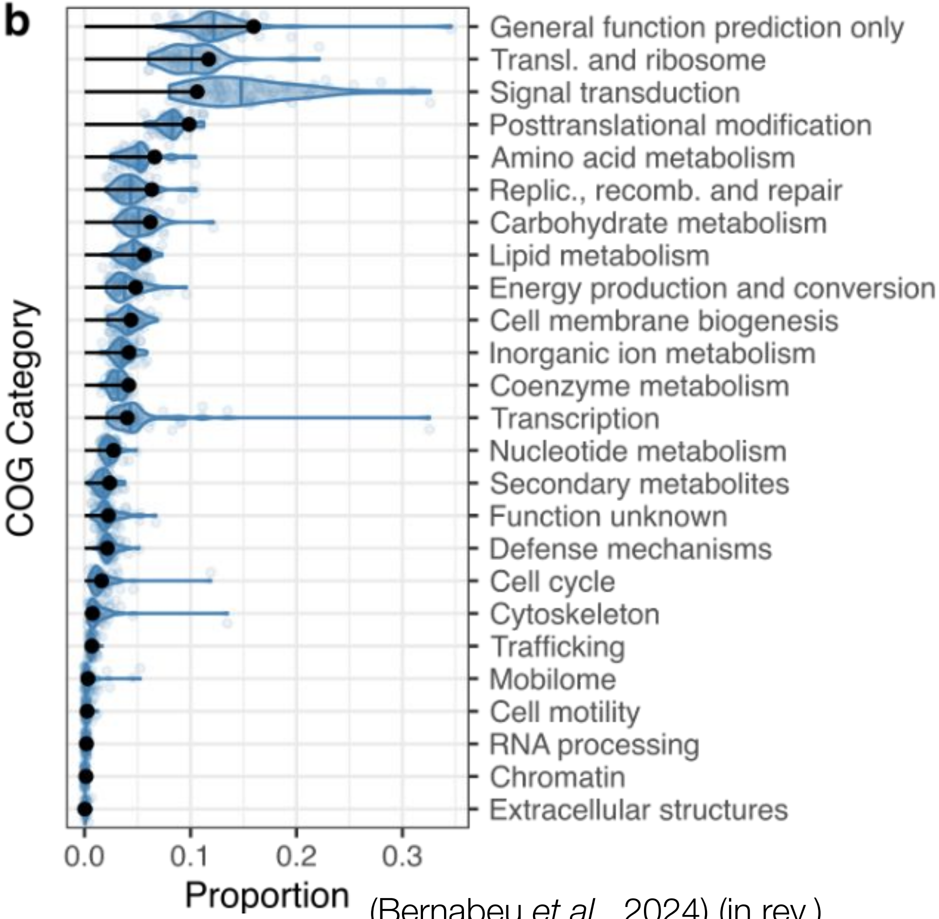
(Bernabeu *et al.*, 2024) (in rev.)

What did LECA look like?



(Bernabeu *et al.*, 2024) (in rev.)

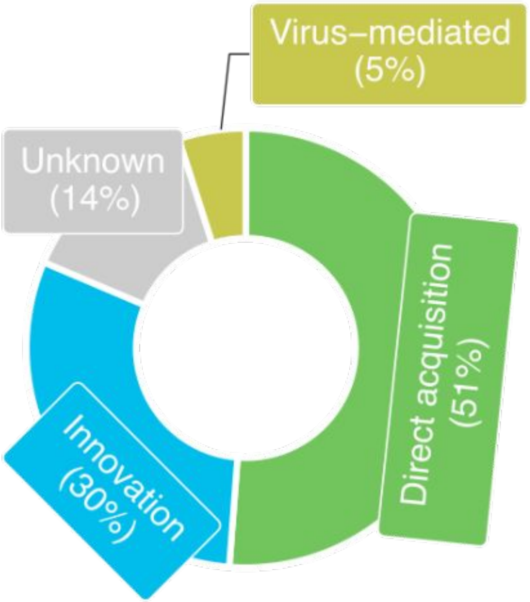
LECA: not (all that) unlike modern eukaryotes



(Bernabeu *et al*., 2024) (in rev.)

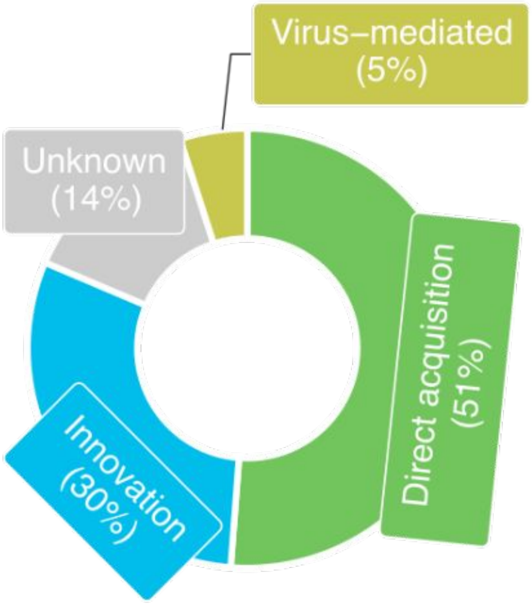
What processes shaped LECA?

A mosaic of acquisitions and innovations generated the LECA proteome



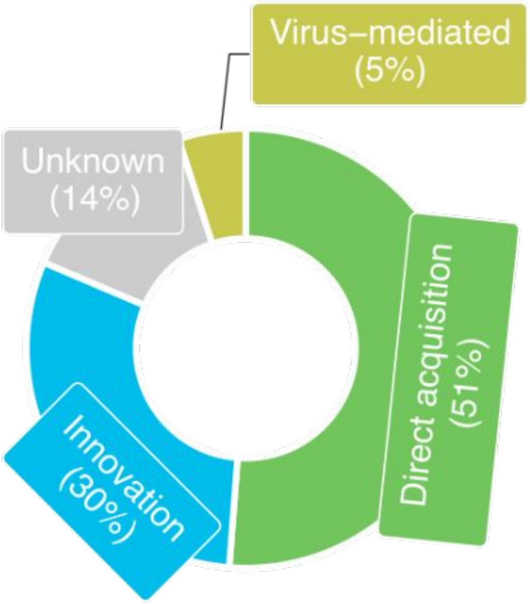
(Bernabeu *et al*., 2024) (in rev.)

A mosaic of acquisitions and innovations generated the LECA proteome



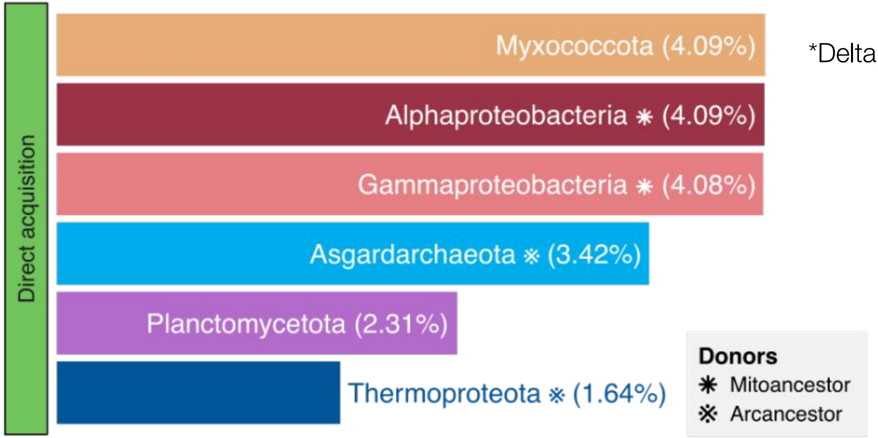
- Impact of innovations on LECA proteome

(Bernabeu *et al*., 2024) (in rev.)



A mosaic of acquisitions and innovations generated the LECA proteome

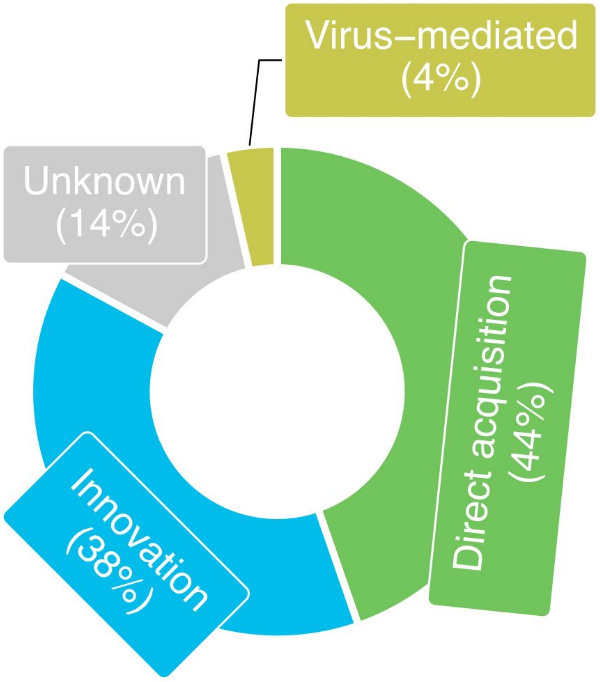
- Significant contributions from outside Asgard archaea and Alphaproteobacteria



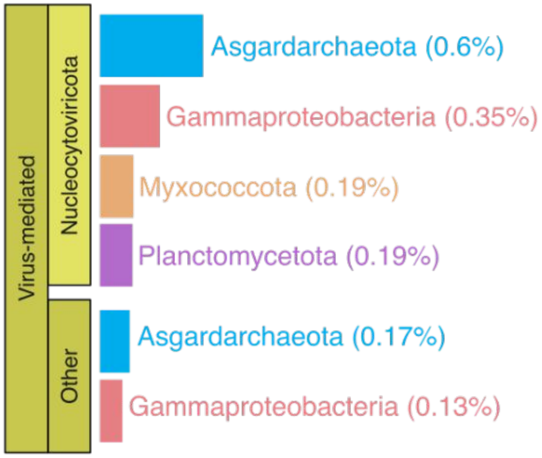
*Deltaproteobacteria

(Bernabeu *et al*., 2024) (in rev.)

A mosaic of acquisitions and innovations generated the LECA proteome

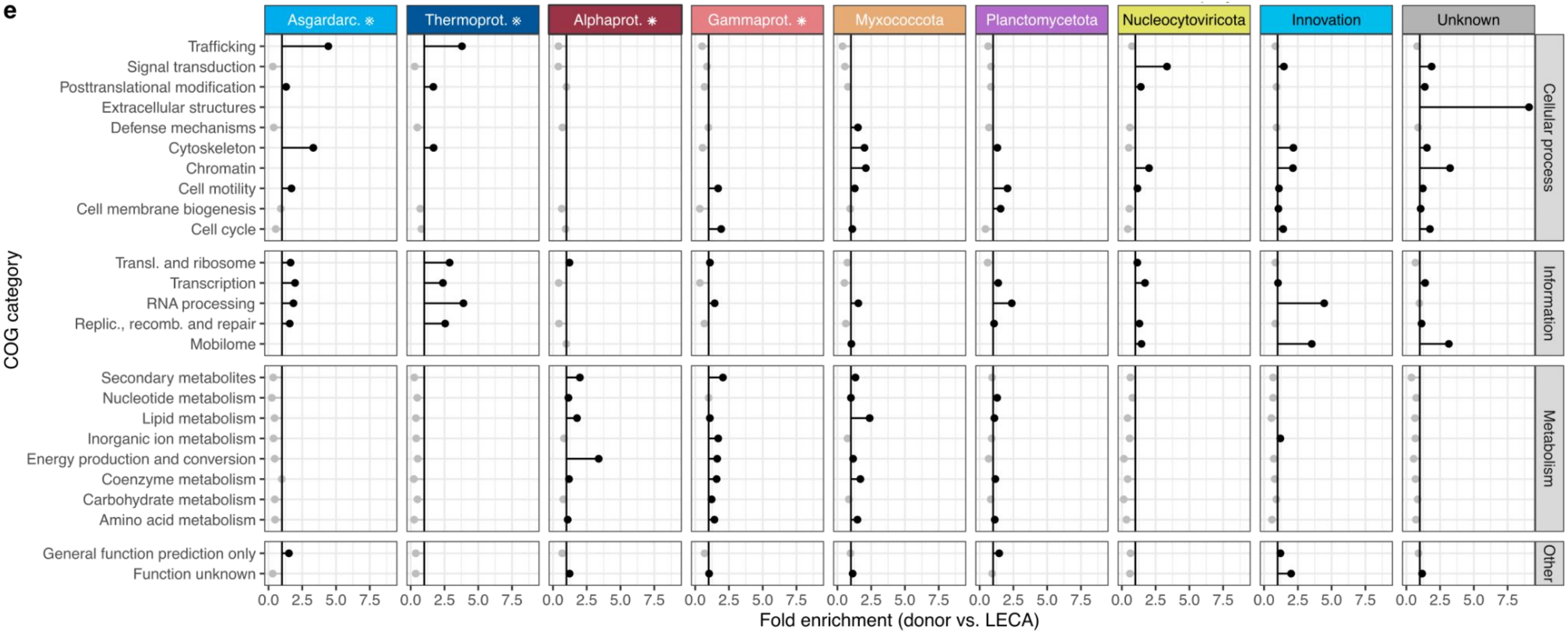


- Viral contribution



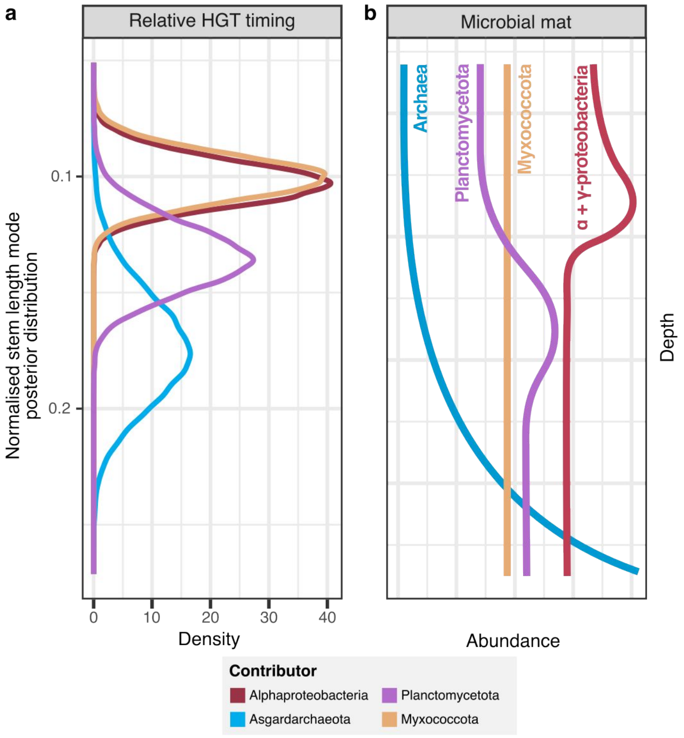
(Bernabeu *et al.*, 2024) (in rev.)

Informational/metabolic divide



How did this proteome evolve?

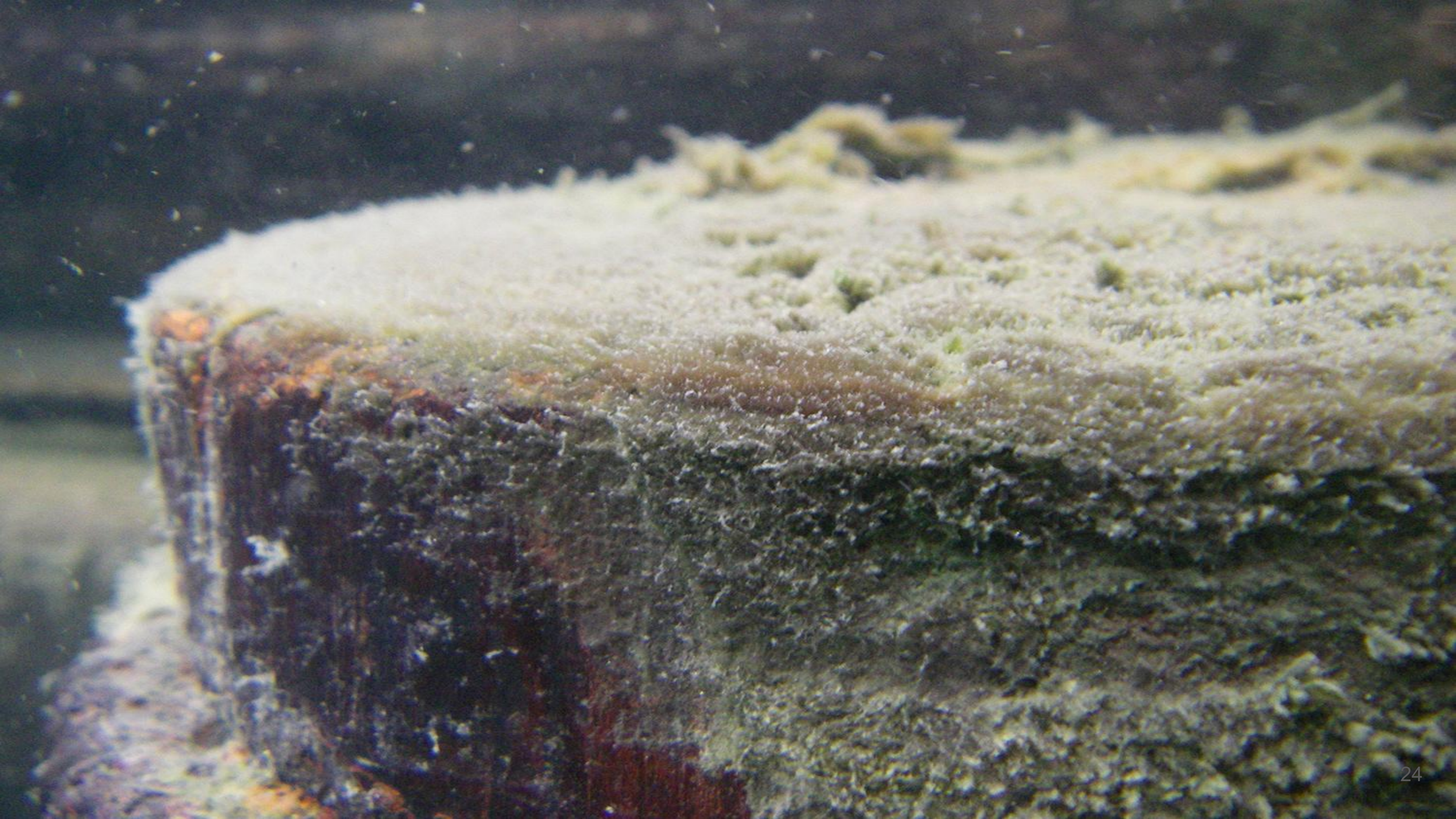
Relative timing of gene acquisitions reveals stepwise evolution in a bacteria-rich environment



“Probabilistic modelling improves relative dating from gene phylogenies”

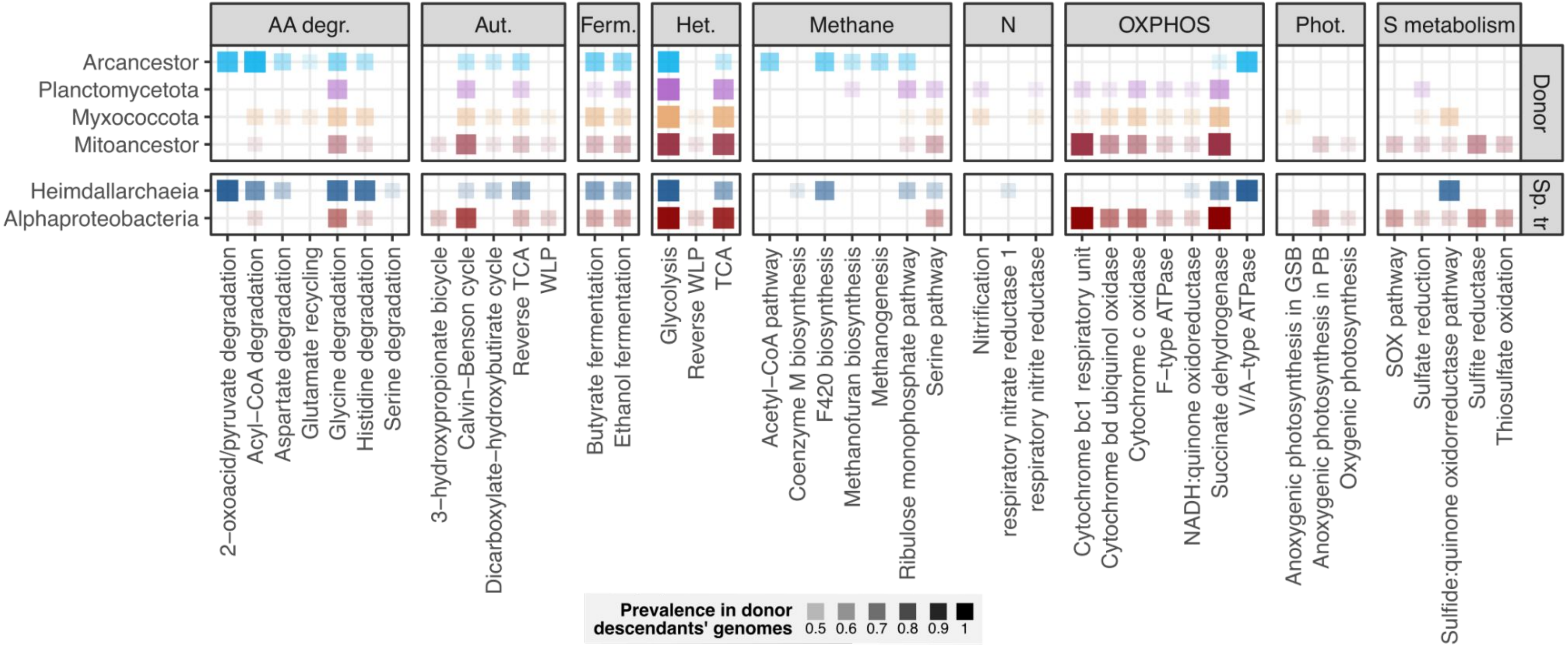
bioRχiv
THE PREPRINT SERVER FOR BIOLOGY

(Bernabeu *et al.*, 2024) (in rev.)



The footprints left in LECA can aid our understanding of the putative donors

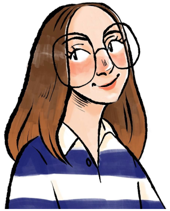
(Gabellone, 2015)



(Bernabeu *et al.*, 2024) (in rev.)

Take-home messages

1. Diverse gene flow from prokaryotes to the (pre-mitochondrial) proto-eukaryote
2. Non-negligible role of viruses as vectors
3. Gradual and complex prokaryote-to-eukaryote transition
4. Likely ecological interactions - bacteria-rich environment



Saioa Manzano-Morales

✉ saioa.manzano@bsc.es
🦋 microbiobits.bsky.social

The LECA team

Marina
Marcet-Houben

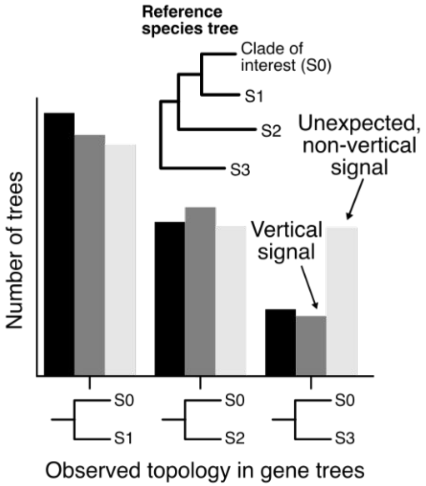
Moisès
Bernabeu

Comparative Genomics
Toni Gabaldón

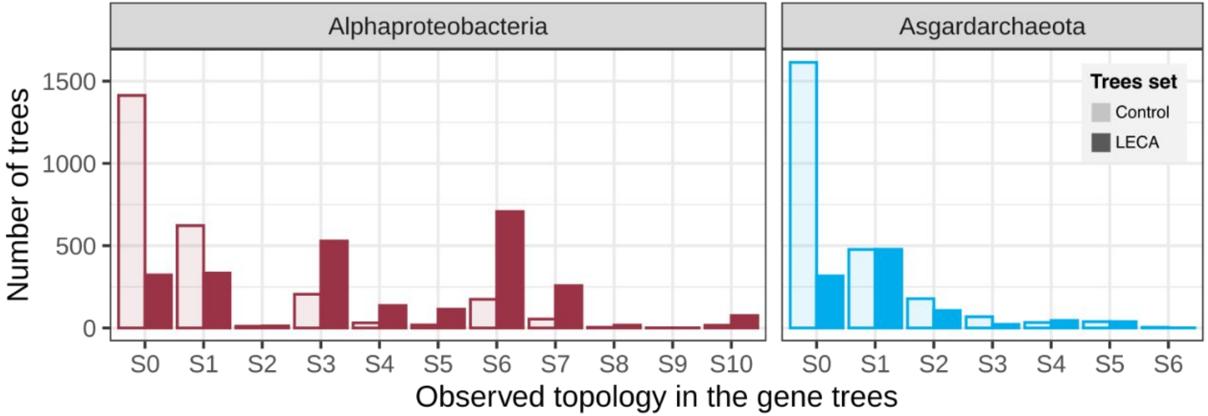
GBMF9742

GA: 724173

b



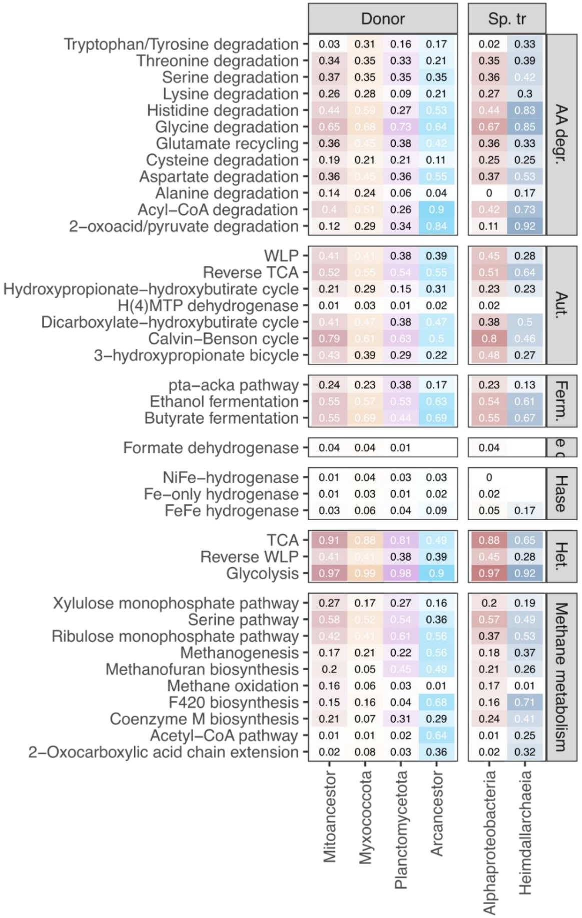
c



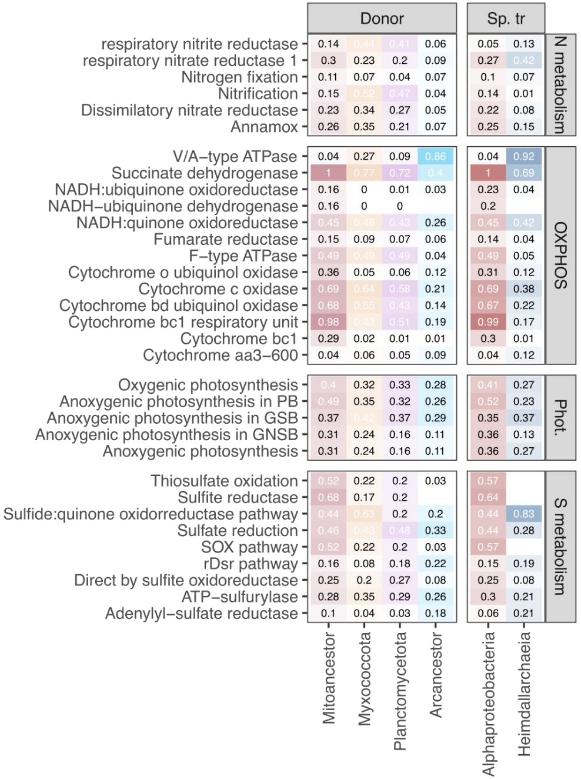
Contributor
🟥 Alpthaproteobacteria
🟦 Asgardarchaeota

Variable
D Donor proportion in S1
T Taxonomic bootstap
U Ultrafast bootstrap

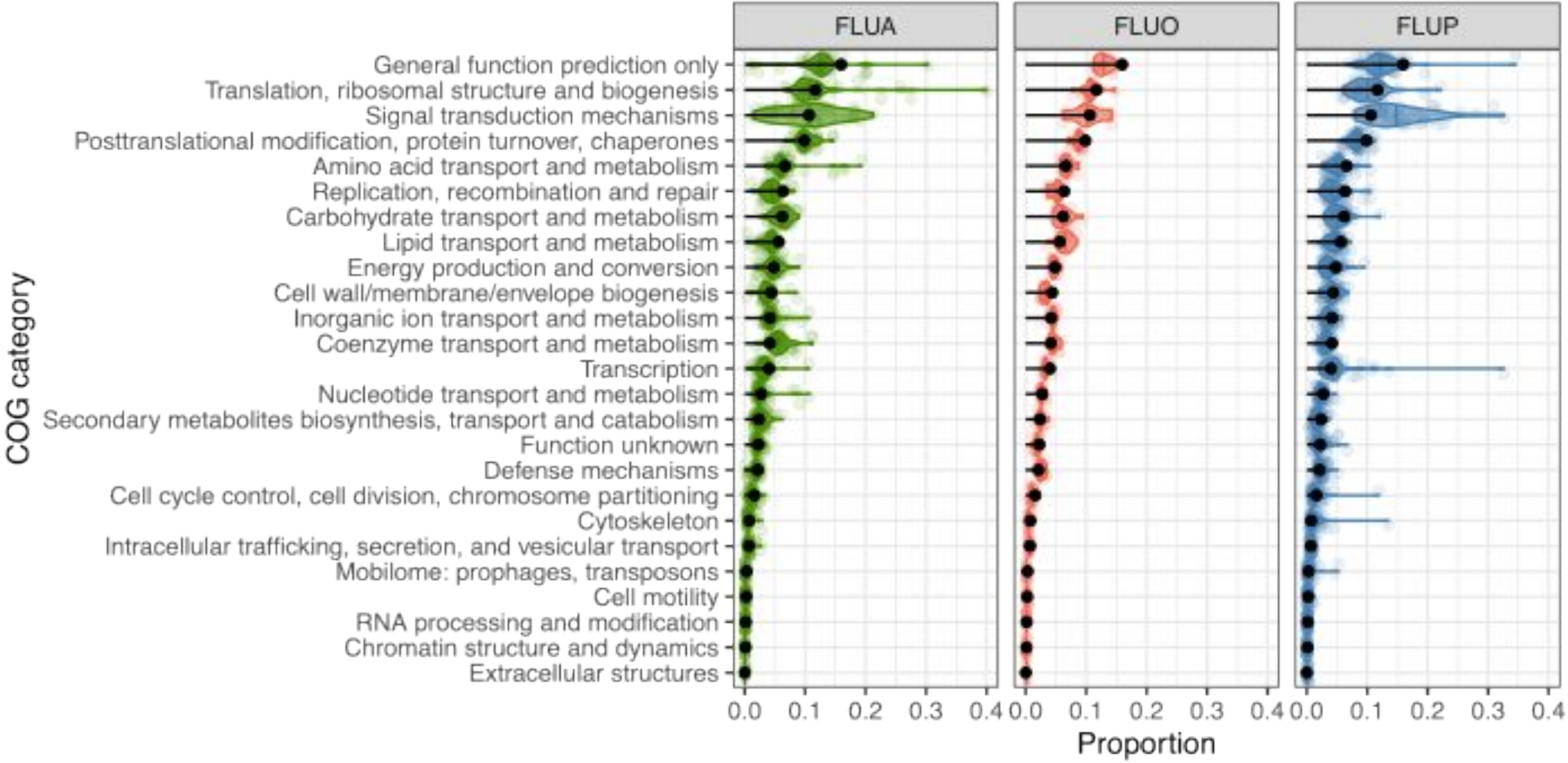
Stringency category
L Low
M Medium
H High

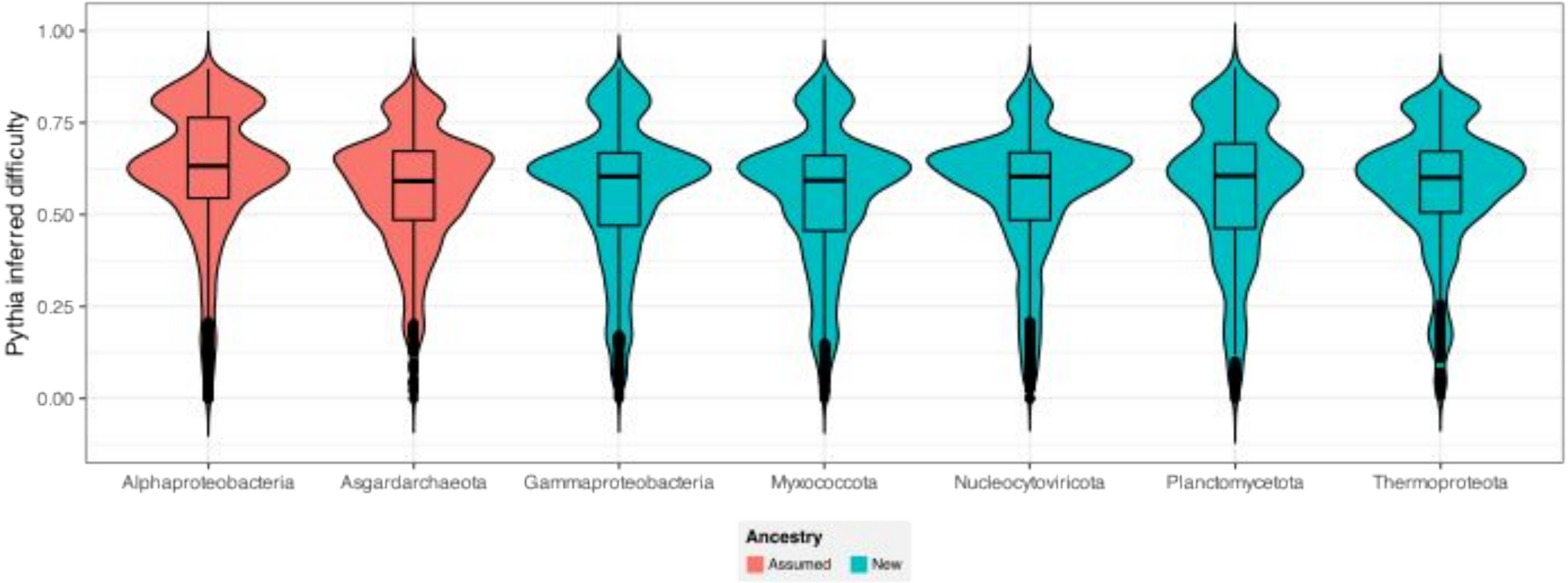


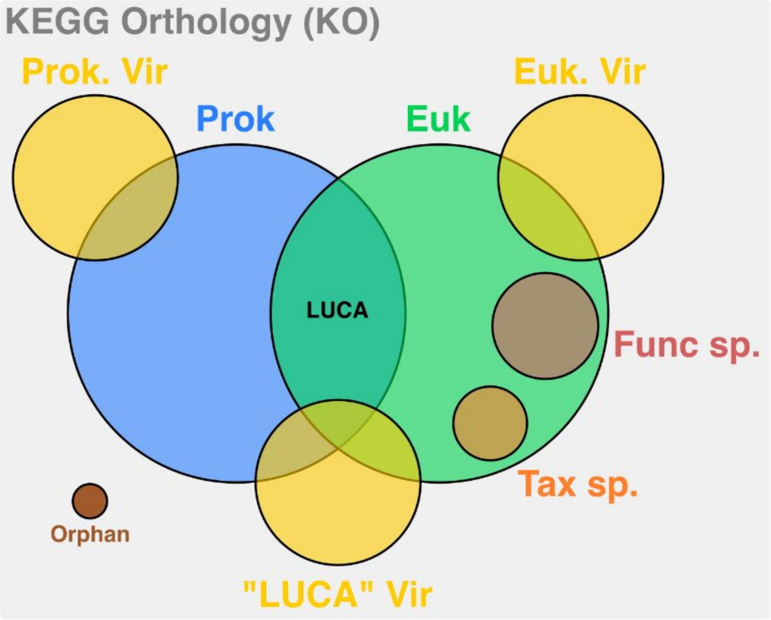
		Donor				Sp. tr	
		Mitoancestor	Myxococcota	Planctomycetota	Arcancestor	Alphaproteobacteria	Heimdallarchaeia
AA degr.	Tryptophan/Tyrosine degradation	0.03	0.31	0.16	0.17	0.02	0.33
	Threonine degradation	0.34	0.35	0.33	0.21	0.35	0.39
	Serine degradation	0.37	0.35	0.35	0.35	0.36	0.42
	Lysine degradation	0.26	0.28	0.09	0.21	0.27	0.3
	Histidine degradation	0.44	0.59	0.27	0.53	0.44	0.83
	Glycine degradation	0.65	0.68	0.73	0.64	0.67	0.85
	Glutamate recycling	0.36	0.45	0.38	0.42	0.36	0.33
	Cysteine degradation	0.19	0.21	0.21	0.11	0.25	0.25
	Aspartate degradation	0.36	0.45	0.36	0.55	0.37	0.53
	Alanine degradation	0.14	0.24	0.06	0.04	0	0.17
	Acyl–CoA degradation	0.4	0.51	0.26	0.9	0.42	0.73
	2–oxoacid/pyruvate degradation	0.12	0.29	0.34	0.84	0.11	0.92
Aut.	WLP	0.41	0.41	0.38	0.39	0.45	0.28
	Reverse TCA	0.52	0.55	0.54	0.55	0.51	0.64
	Hydroxypropionate–hydroxybutirate cycle	0.21	0.29	0.15	0.31	0.23	0.23
	H(4)MTP dehydrogenase	0.01	0.03	0.01	0.02	0.02	
	Dicarboxylate–hydroxybutirate cycle	0.41	0.47	0.38	0.47	0.38	0.5
	Calvin–Benson cycle	0.79	0.61	0.63	0.5	0.8	0.46
	3–hydroxypropionate bicycle	0.43	0.39	0.29	0.22	0.48	0.27
Ferm.	pta–acka pathway	0.24	0.23	0.38	0.17	0.23	0.13
	Ethanol fermentation	0.55	0.57	0.53	0.63	0.54	0.61
	Butyrate fermentation	0.55	0.69	0.44	0.69	0.55	0.67
e c	Formate dehydrogenase	0.04	0.04	0.01		0.04	
Hase	NiFe–hydrogenase	0.01	0.04	0.03	0.03	0	
	Fe–only hydrogenase	0.01	0.03	0.01	0.02	0.02	
	FeFe hydrogenase	0.03	0.06	0.04	0.09	0.05	0.17
Het.	TCA	0.91	0.88	0.81	0.49	0.88	0.65
	Reverse WLP	0.41	0.41	0.38	0.39	0.45	0.28
	Glycolysis	0.97	0.99	0.98	0.9	0.97	0.92
Methane metabolism	Xylulose monophosphate pathway	0.27	0.17	0.27	0.16	0.2	0.19
	Serine pathway	0.58	0.52	0.54	0.36	0.57	0.49
	Ribulose monophosphate pathway	0.42	0.41	0.61	0.56	0.37	0.53
	Methanogenesis	0.17	0.21	0.22	0.56	0.18	0.37
	Methanofuran biosynthesis	0.2	0.05	0.45	0.49	0.21	0.26
	Methane oxidation	0.16	0.06	0.03	0.01	0.17	0.01
	F420 biosynthesis	0.15	0.16	0.04	0.68	0.16	0.71
	Coenzyme M biosynthesis	0.21	0.07	0.31	0.29	0.24	0.41
	Acetyl–CoA pathway	0.01	0.01	0.02	0.64	0.01	0.25
	2–Oxocarboxylic acid chain extension	0.02	0.08	0.03	0.36	0.02	0.32



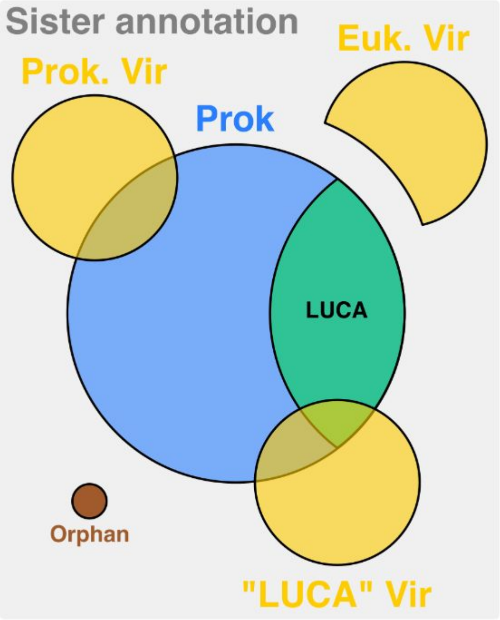
		Donor				Sp. tr	
		Mitoancestor	Myxococcota	Planctomycetota	Arcancestor	Alphaproteobacteria	Heimdallarchaeia
N metabolism	respiratory nitrite reductase	0.14	0.44	0.41	0.06	0.05	0.13
	respiratory nitrate reductase 1	0.3	0.23	0.2	0.09	0.27	0.42
	Nitrogen fixation	0.11	0.07	0.04	0.07	0.1	0.07
	Nitrification	0.15	0.52	0.47	0.04	0.14	0.01
	Dissimilatory nitrate reductase	0.23	0.34	0.27	0.05	0.22	0.08
	Annamox	0.26	0.35	0.21	0.07	0.25	0.15
OXPHOS	V/A–type ATPase	0.04	0.27	0.09	0.86	0.04	0.92
	Succinate dehydrogenase	1	0.77	0.72	0.4	1	0.69
	NADH:ubiquinone oxidoreductase	0.16	0	0.01	0.03	0.23	0.04
	NADH–ubiquinone dehydrogenase	0.16	0	0		0.2	
	NADH:quinone oxidoreductase	0.45	0.48	0.43	0.26	0.45	0.42
	Fumarate reductase	0.15	0.09	0.07	0.06	0.14	0.04
	F–type ATPase	0.49	0.49	0.49	0.04	0.49	0.05
	Cytochrome o ubiquinol oxidase	0.36	0.05	0.06	0.12	0.31	0.12
	Cytochrome c oxidase	0.69	0.64	0.58	0.21	0.69	0.38
	Cytochrome bd ubiquinol oxidase	0.68	0.55	0.43	0.14	0.67	0.22
	Cytochrome bc1 respiratory unit	0.98	0.43	0.51	0.19	0.99	0.17
	Cytochrome bc1	0.29	0.02	0.01	0.01	0.3	0.01
	Cytochrome aa3–600	0.04	0.06	0.05	0.09	0.04	0.12
Phot.	Oxygenic photosynthesis	0.4	0.32	0.33	0.28	0.41	0.27
	Anoxygenic photosynthesis in PB	0.49	0.35	0.32	0.26	0.52	0.23
	Anoxygenic photosynthesis in GSB	0.37	0.42	0.37	0.29	0.35	0.37
	Anoxygenic photosynthesis in GNSB	0.31	0.24	0.16	0.11	0.36	0.13
	Anoxygenic photosynthesis	0.31	0.24	0.16	0.11	0.36	0.27
S metabolism	Thiosulfate oxidation	0.52	0.22	0.2	0.03	0.57	
	Sulfite reductase	0.68	0.17	0.2		0.64	
	Sulfide:quinone oxidorreductase pathway	0.44	0.63	0.2	0.2	0.44	0.83
	Sulfate reduction	0.46	0.43	0.48	0.33	0.44	0.28
	SOX pathway	0.52	0.22	0.2	0.03	0.57	
	rDsr pathway	0.16	0.08	0.18	0.22	0.15	0.19
	Direct by sulfite oxidoreductase	0.25	0.2	0.27	0.08	0.25	0.08
	ATP–sulfurylase	0.28	0.35	0.29	0.26	0.3	0.21
	Adenylyl–sulfate reductase	0.1	0.04	0.03	0.18	0.06	0.21


FLUA
FLUO
FLUP
COG category
General function prediction only
Translation, ribosomal structure and biogenesis
Signal transduction mechanisms
Posttranslational modification, protein turnover, chaperones
Amino acid transport and metabolism
Replication, recombination and repair
Carbohydrate transport and metabolism
Lipid transport and metabolism
Energy production and conversion
Cell wall/membrane/envelope biogenesis
Inorganic ion transport and metabolism
Coenzyme transport and metabolism
Transcription
Nucleotide transport and metabolism
Secondary metabolites biosynthesis, transport and catabolism
Function unknown
Defense mechanisms
Cell cycle control, cell division, chromosome partitioning
Cytoskeleton
Intracellular trafficking, secretion, and vesicular transport
Mobilome: prophages, transposons
Cell motility
RNA processing and modification
Chromatin structure and dynamics
Extracellular structures
0.0
0.1
0.2
0.3
0.4
Proportion

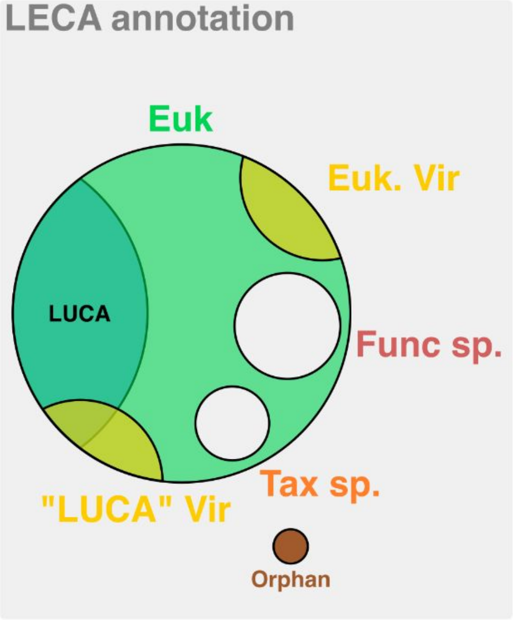

Pythia inferred difficulty
1.00
0.75
0.50
0.25
0.00
Alphaproteobacteria
Asgardarchaeota
Gammaproteobacteria
Myxococcota
Nucleocytoviricota
Planctomycetota
Thermoproteota
Ancestry
Assumed
New



KEGG Orthology (KO)



Sister annotation



LECA annotation

Vir: viruses
Prok: prokaryotes
Euk: eukarypotes

Tax. sp.: taxonomically specific
Func. sp.: functionally specific