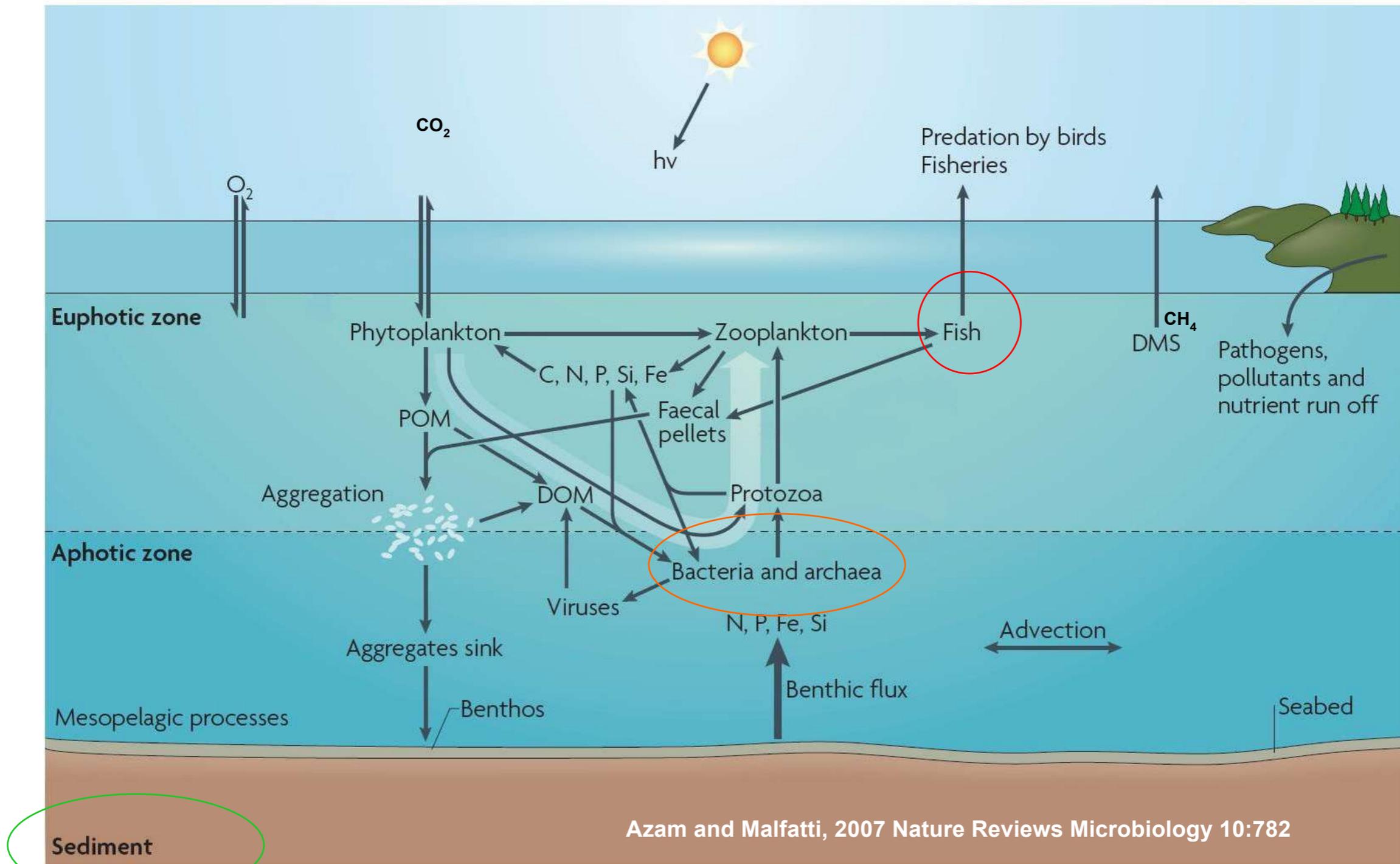


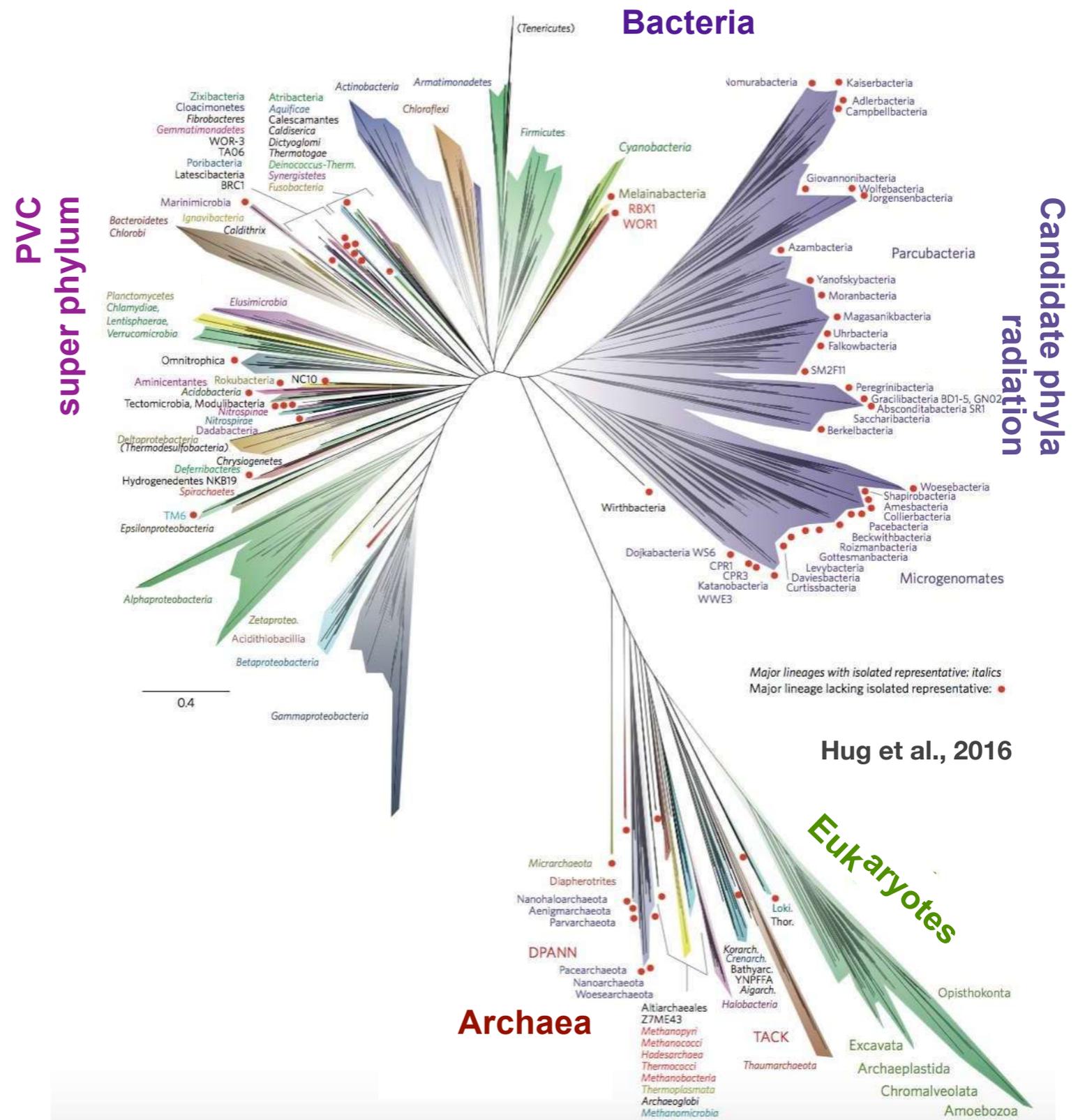
# **L03a: Diversity of Marine Bacteria, Archaea, Eukarya and Viruses**

# Marine microbes are highly diverse

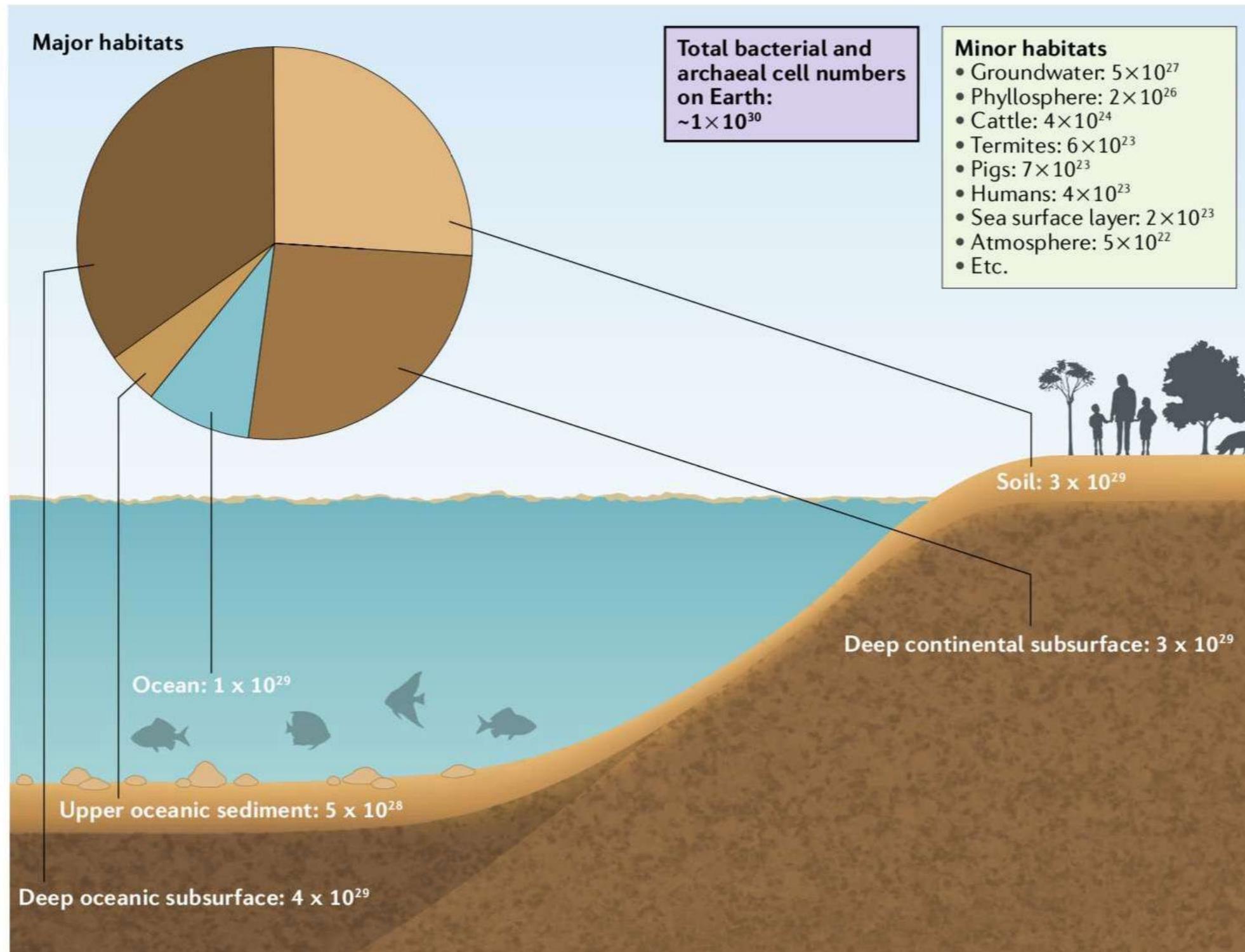


# Microbial diversity on Earth

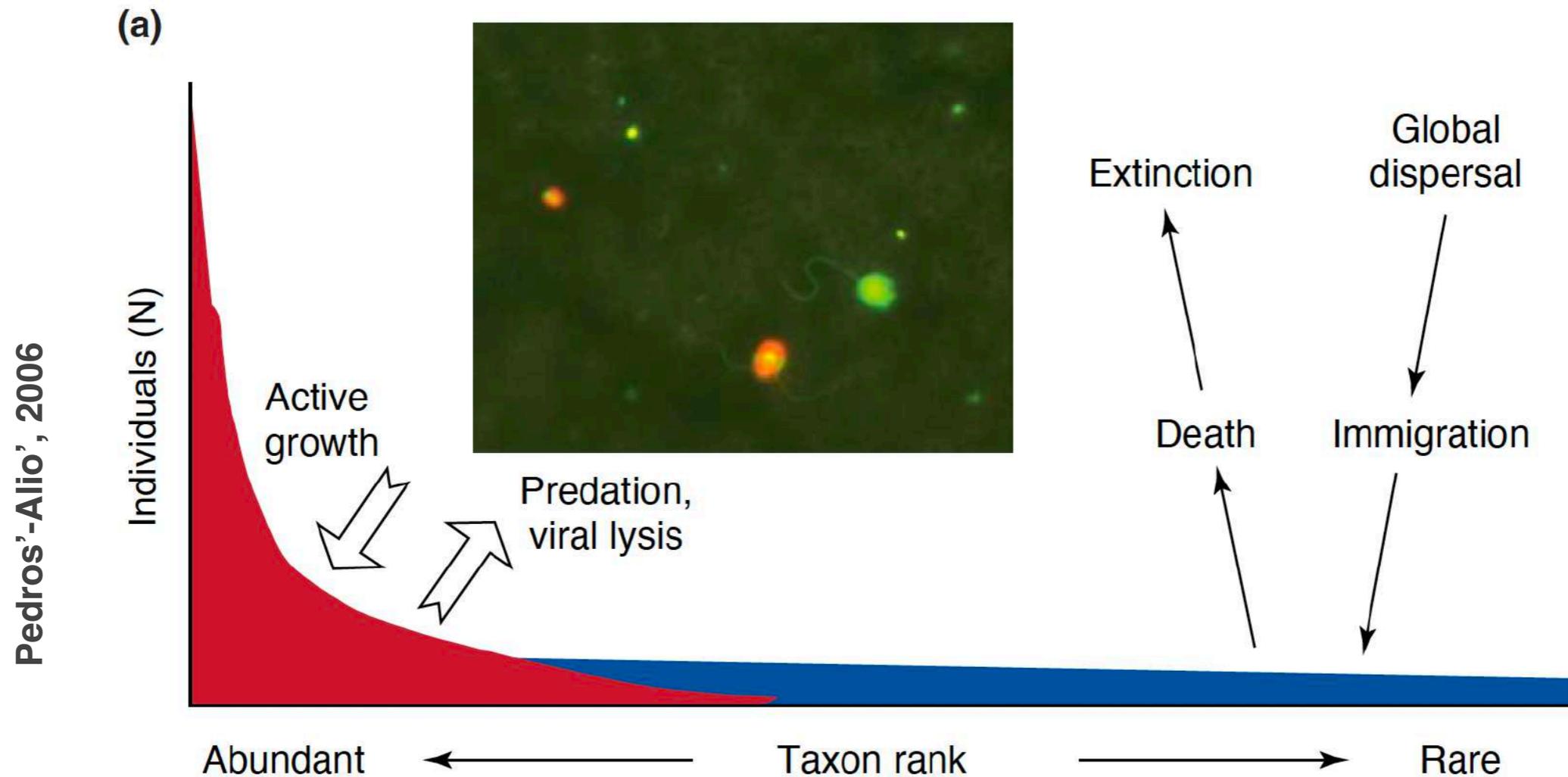
- Still finding new microbes with massive sequencing approaches
- Deeper branching → early in evolution



# Abundance of Bacteria and Archaea on Earth

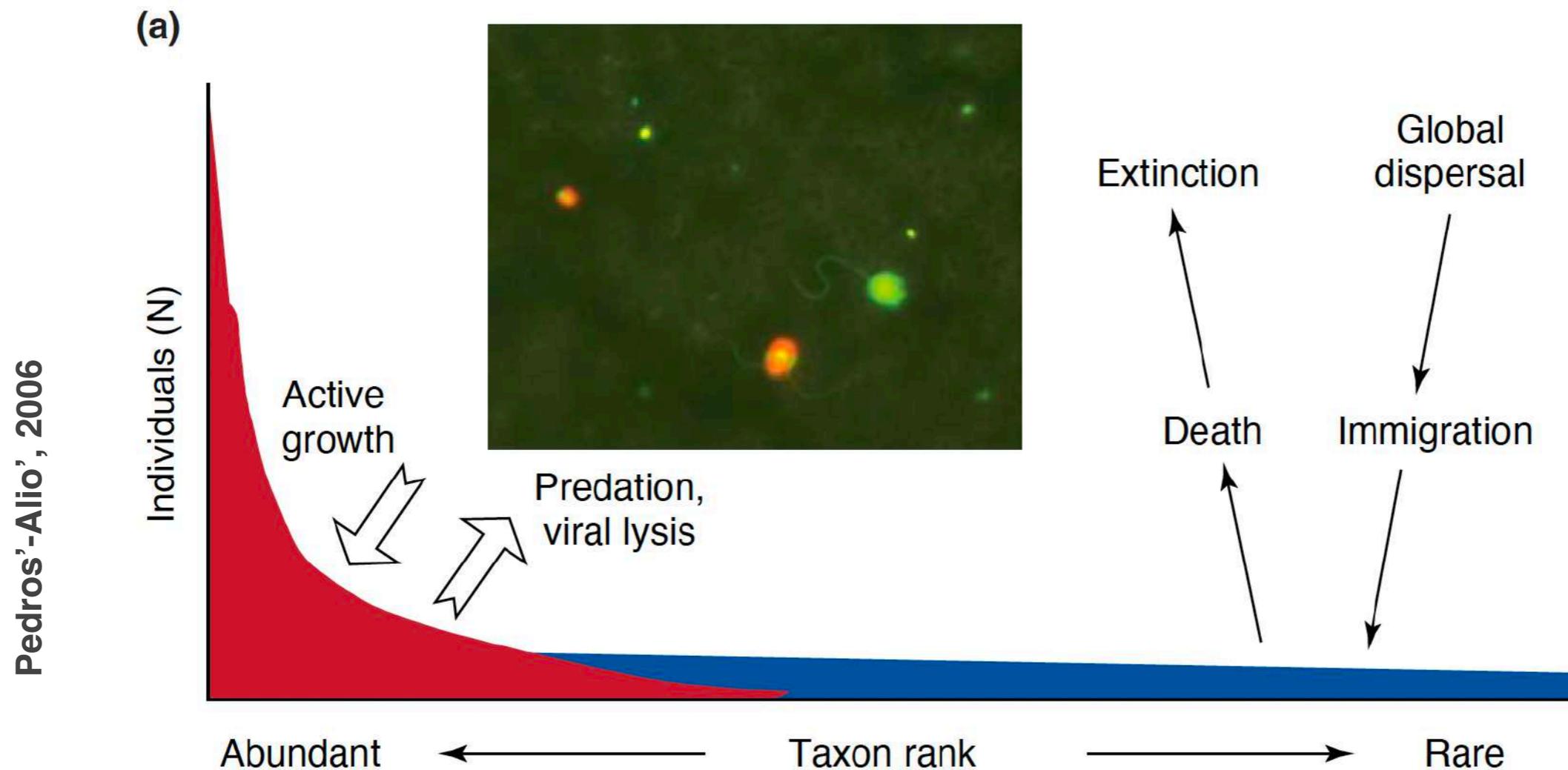


# Microbial diversity: Individual vs. Taxon



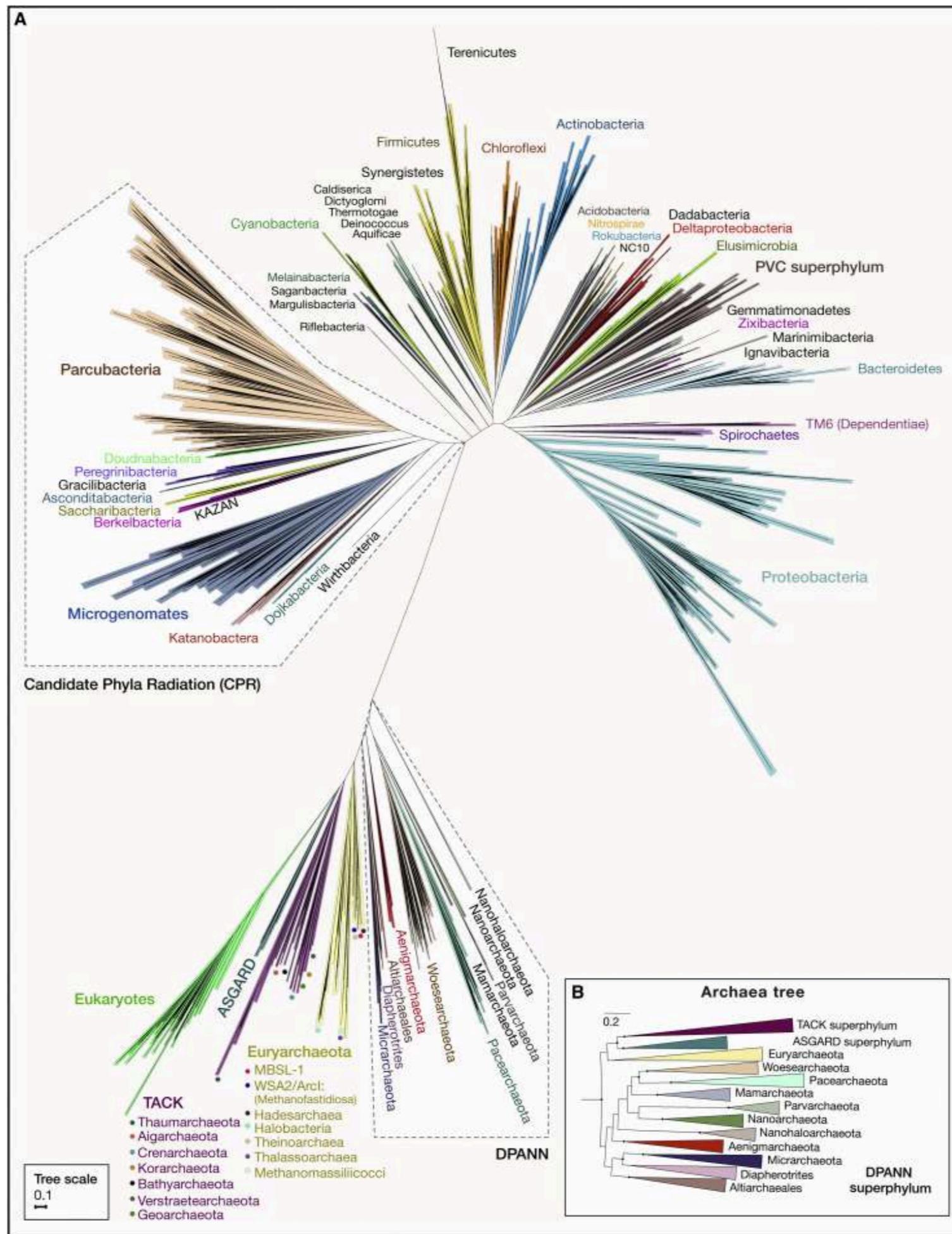
- (i) The red section curve: abundant taxa with active role in carbon and energy flow in ecosystem → active growth and suffer losses because of predation and viral lysis
- (ii) The blue section curve: rare taxa, which survive in ecosystem at low abundance (? with resting stages or spores) → This 'seed bank' recruits new species through immigration

# Microbial diversity: Individual vs. Taxon



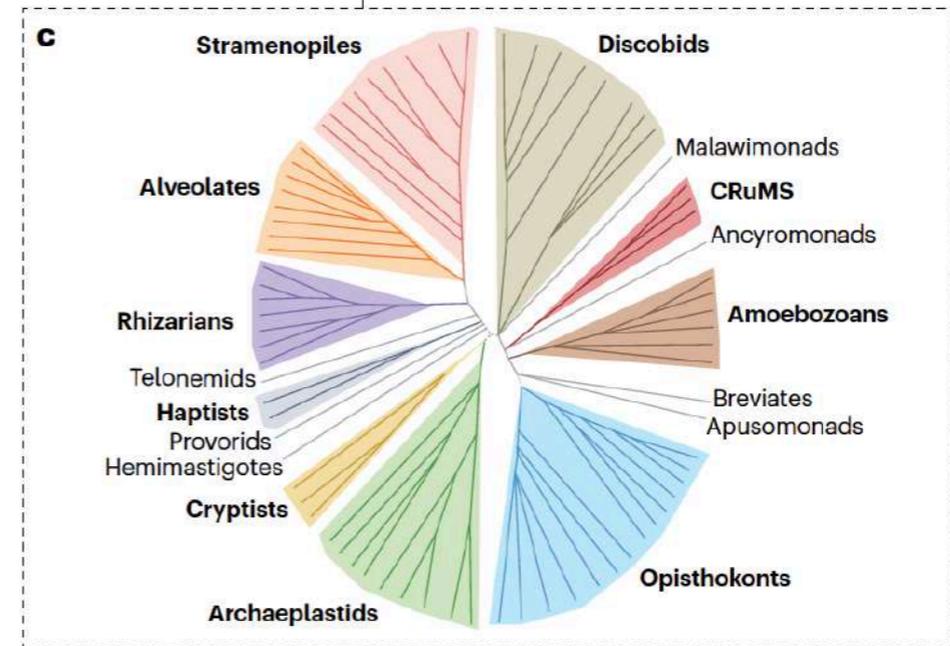
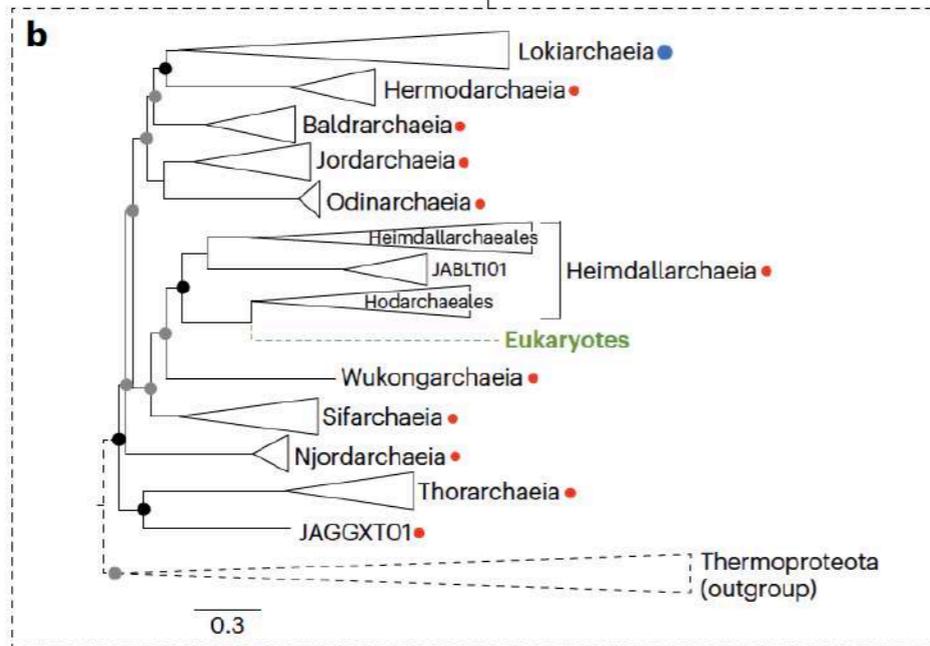
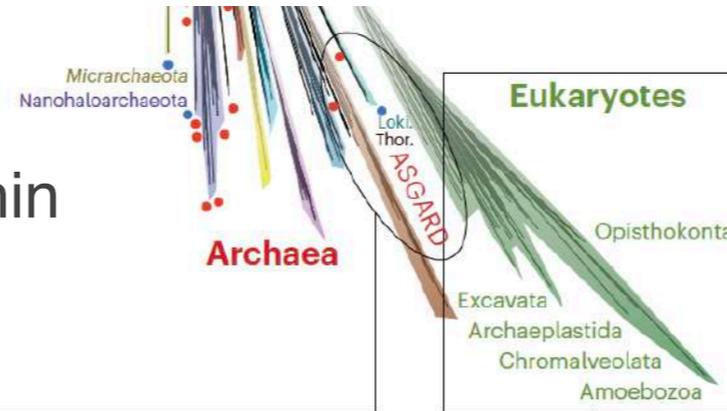
- Because of easy dispersal of microbes, it is likely that essentially all microorganisms reach this ecosystem periodically
- By being rare, these microbes do not experience viral attack and predation is minimized; thus, extinction can only be a result of cell death
- Because extinction can be expected to be low while immigration is probably high, this part of the curve might be long and perhaps include the whole biodiversity of microbes

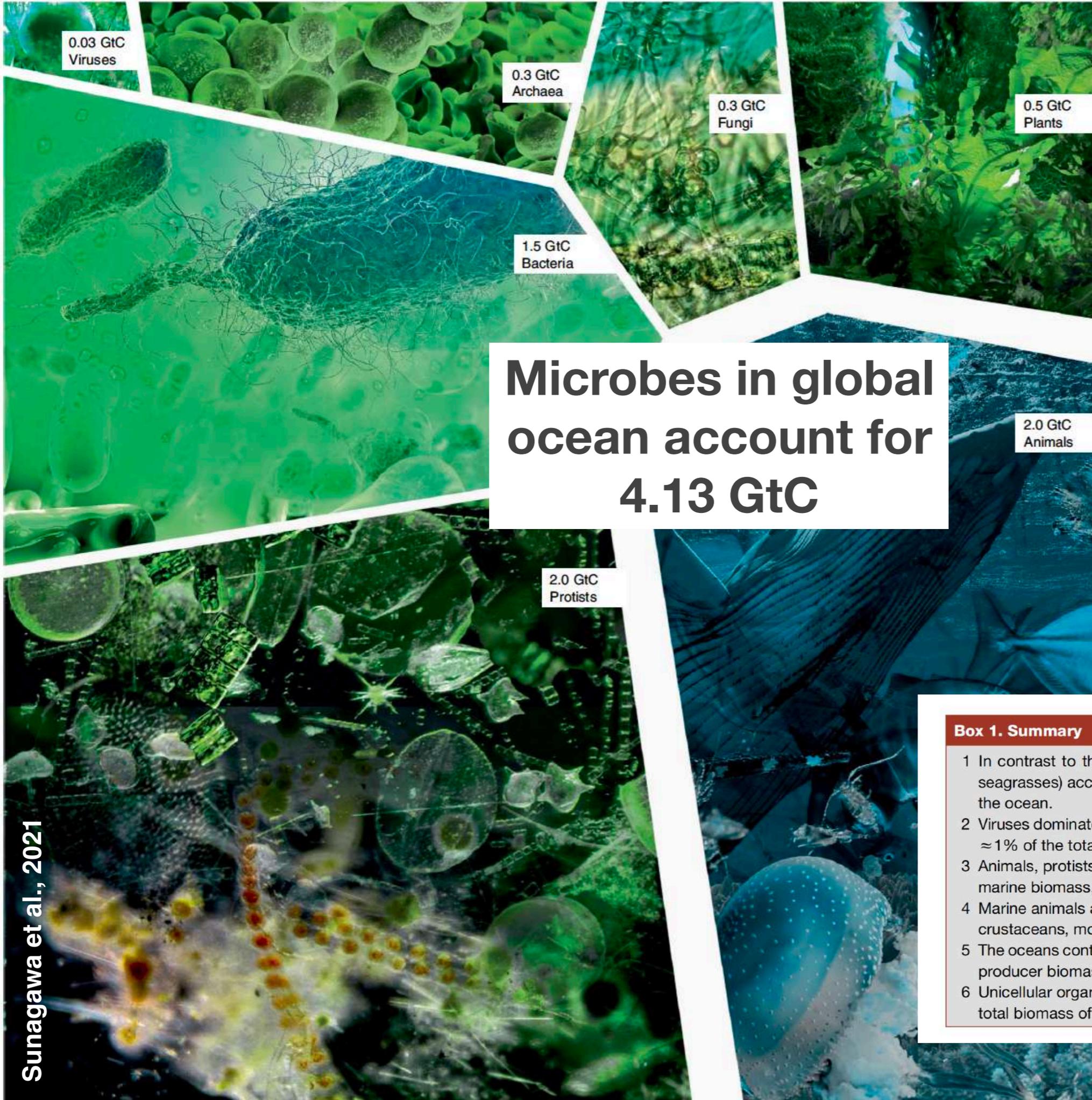
# Diversity of Marine Bacteria



# Archaea and Eukaryotes

A two-domain tree with Eukarya branching from within the Archaea, sister to the Asgardarchaeota



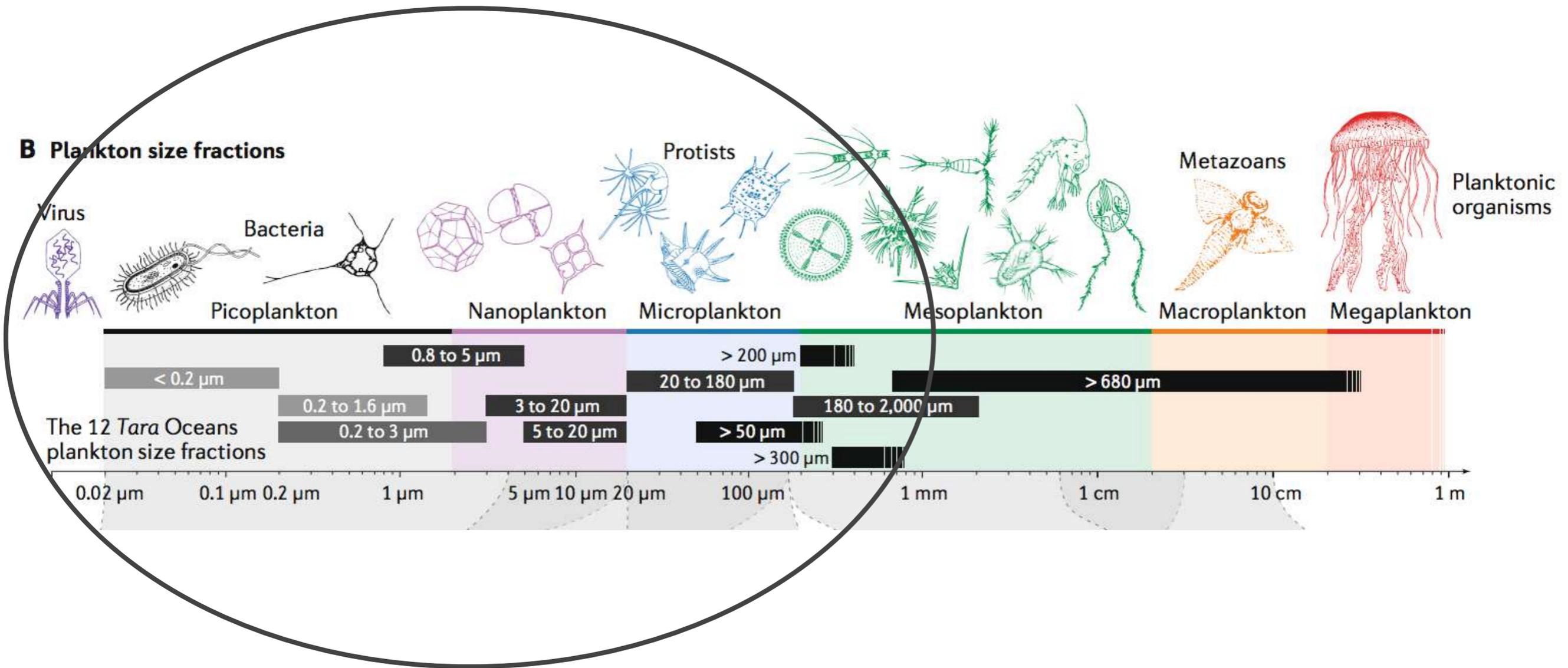


Bar-On & Milo, 2019

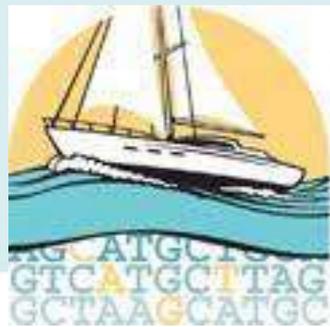
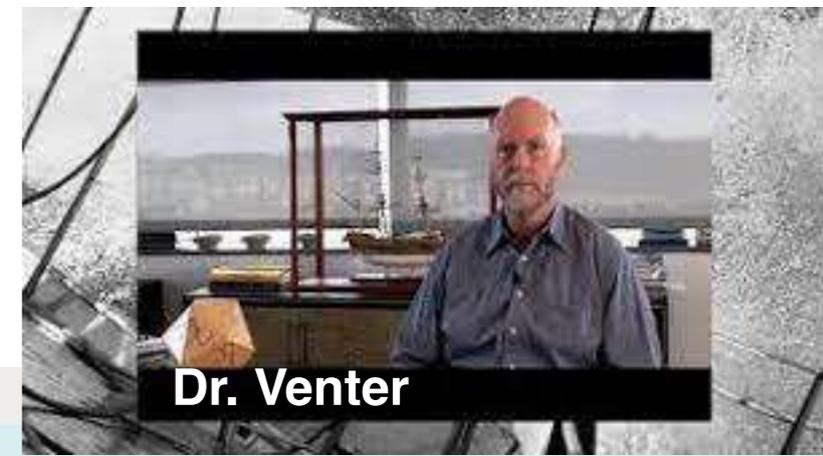
**Box 1. Summary**

- 1 In contrast to their domination on land, plants (green algae and seagrasses) account for less than 10% of the total biomass in the ocean.
- 2 Viruses dominate the ocean in terms of number but constitute only  $\approx 1\%$  of the total biomass.
- 3 Animals, protists, and bacteria together account for  $\approx 80\%$  of the marine biomass, whereas on land they comprise only  $\approx 2\%$ .
- 4 Marine animals are dominated by small mesopelagic fish and crustaceans, mostly copepods, shrimp, and krill.
- 5 The oceans contain much more consumer biomass ( $\approx 5$  Gt C) than producer biomass ( $\approx 1$  Gt C).
- 6 Unicellular organisms contribute approximately two-thirds of the total biomass of marine organisms.

# Size-range of marine microbes



# GLOBAL OCEAN SURVEY, GOS



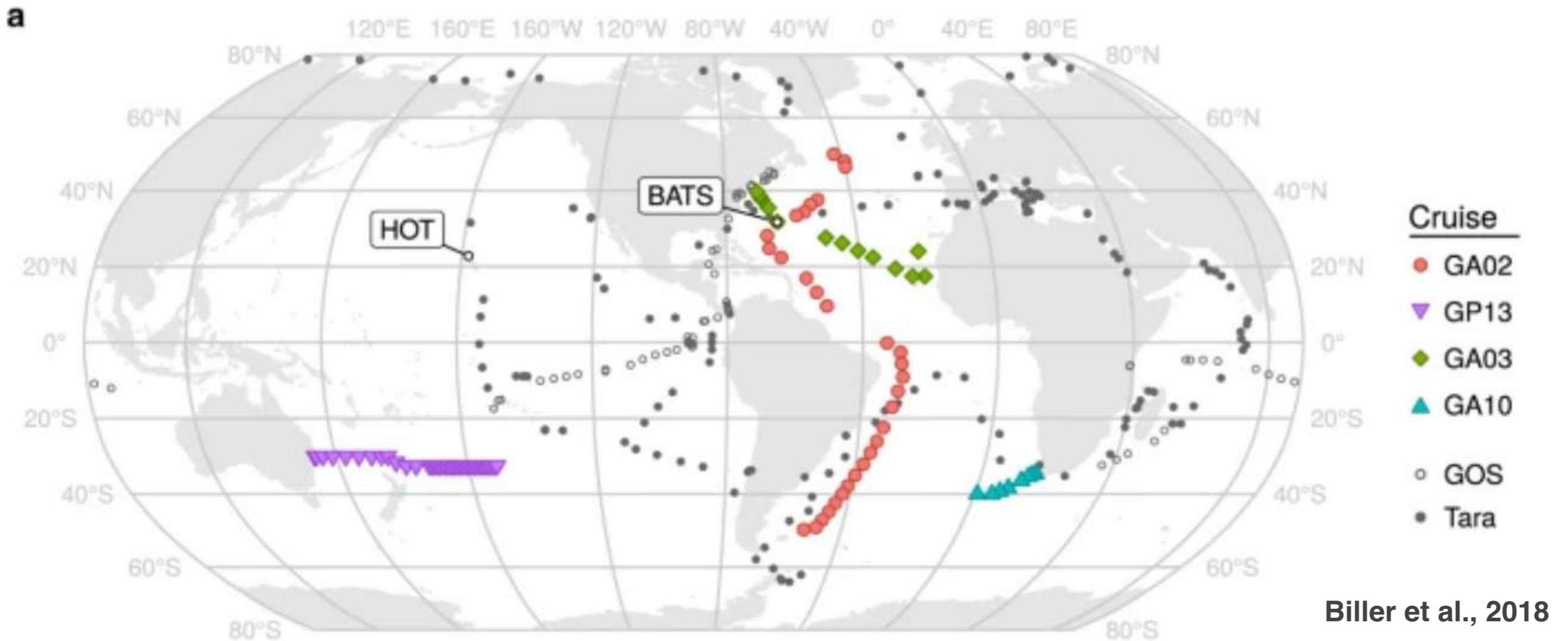
This article is part of the Oceanic Metagenomics collection.

● Past Routes: 2003–2008

● Europe Expedition: 2009–2010

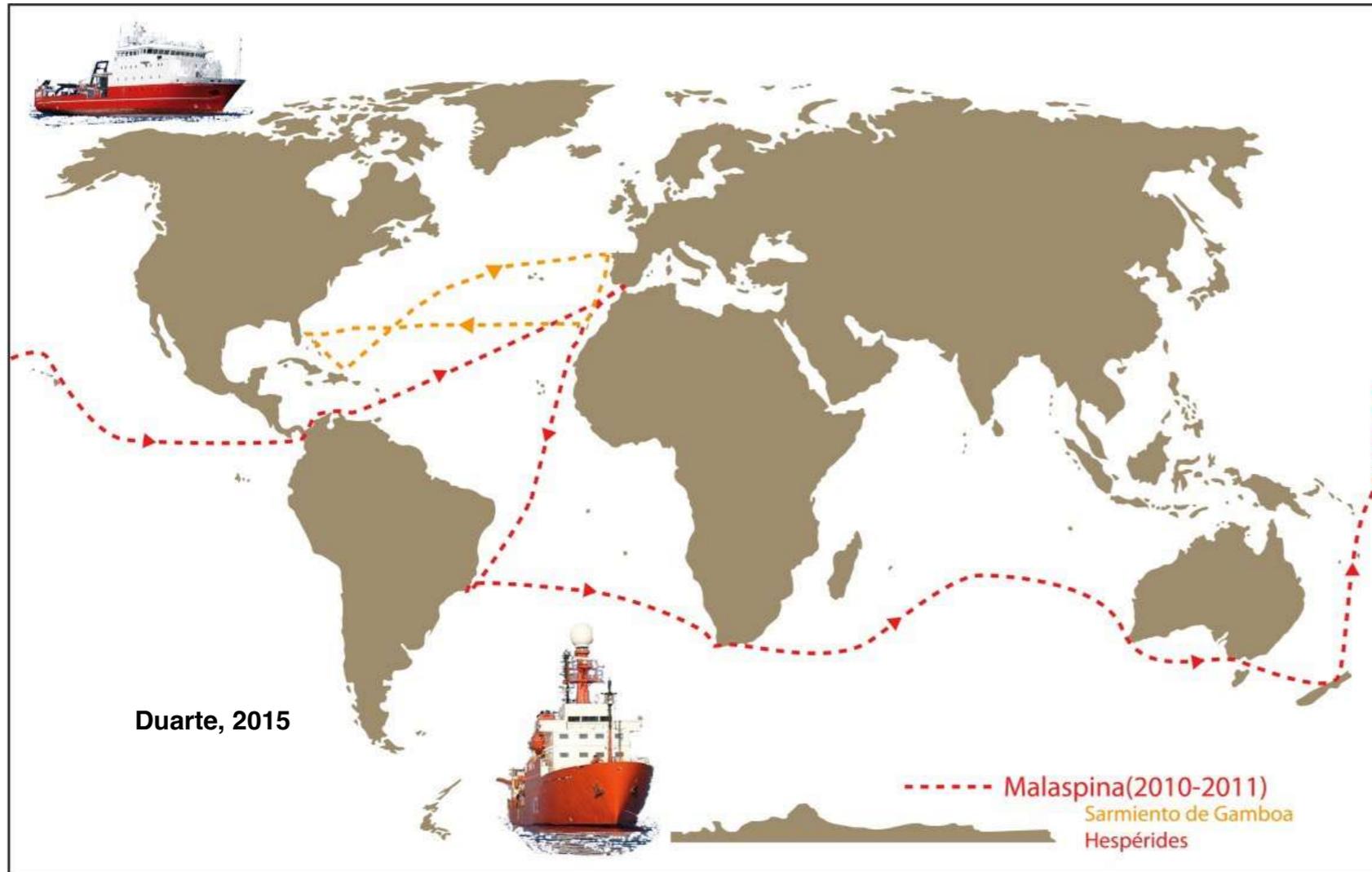
# GEOTRACES and microbial exploration

Figure 1: GEOTRACES metagenomic sampling locations.

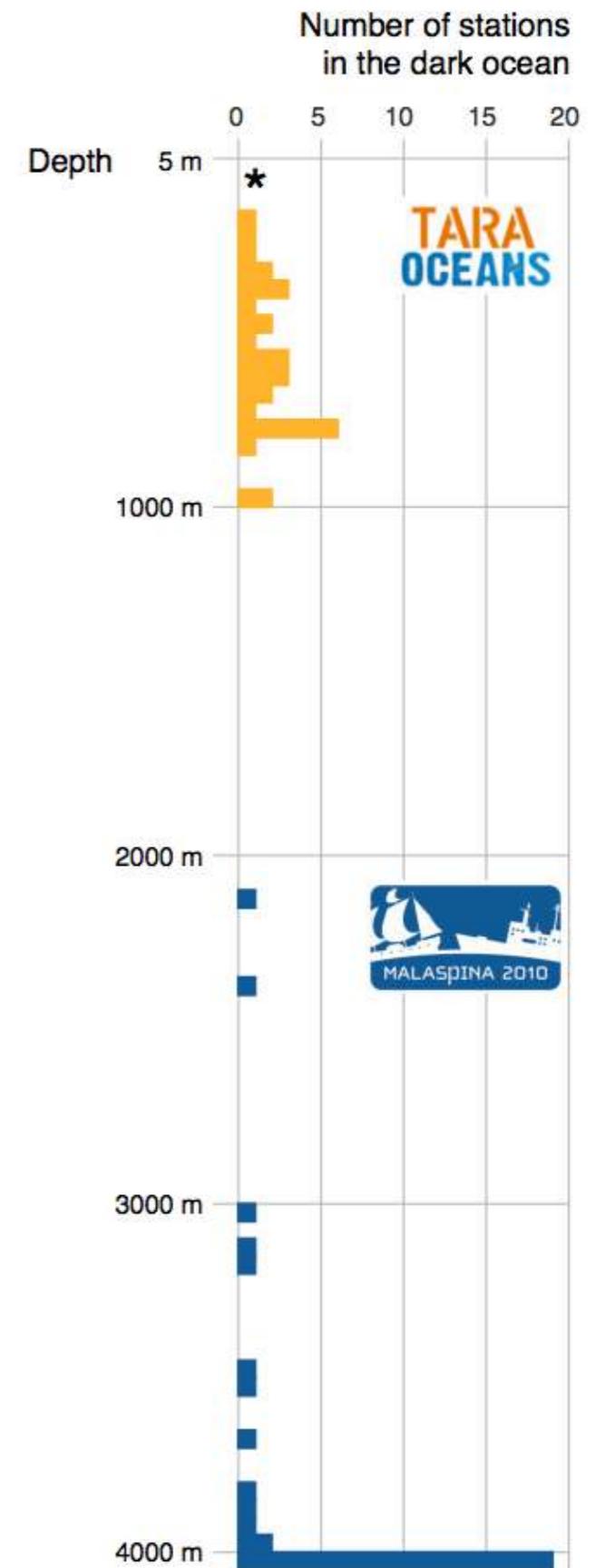


GEOTRACES is an international study of the marine biogeochemical cycles of trace elements and their isotopes, to accelerate research under a global program (since 2008, Anderson et al., 2014)

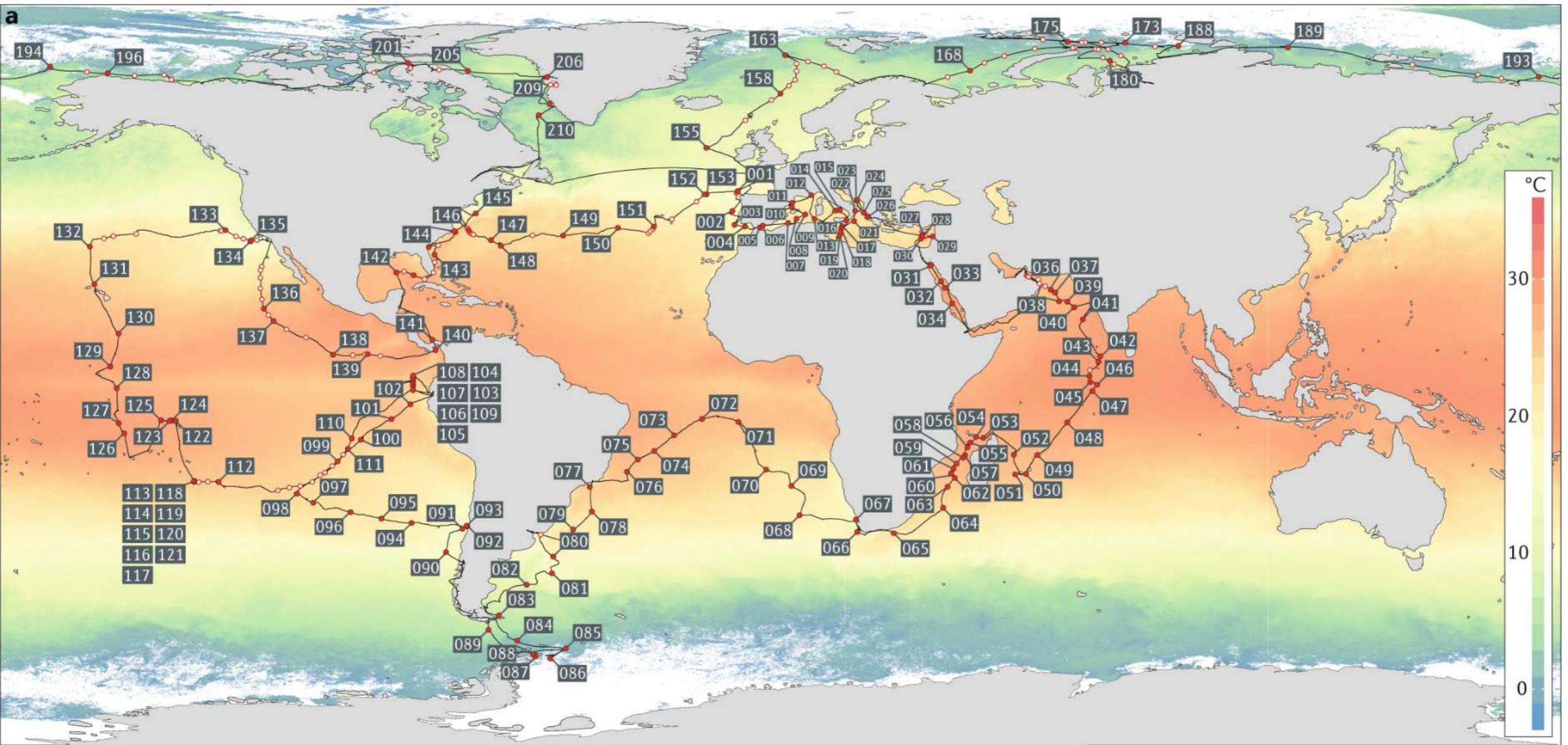
# MALASPINA



Acinas et al., 2019

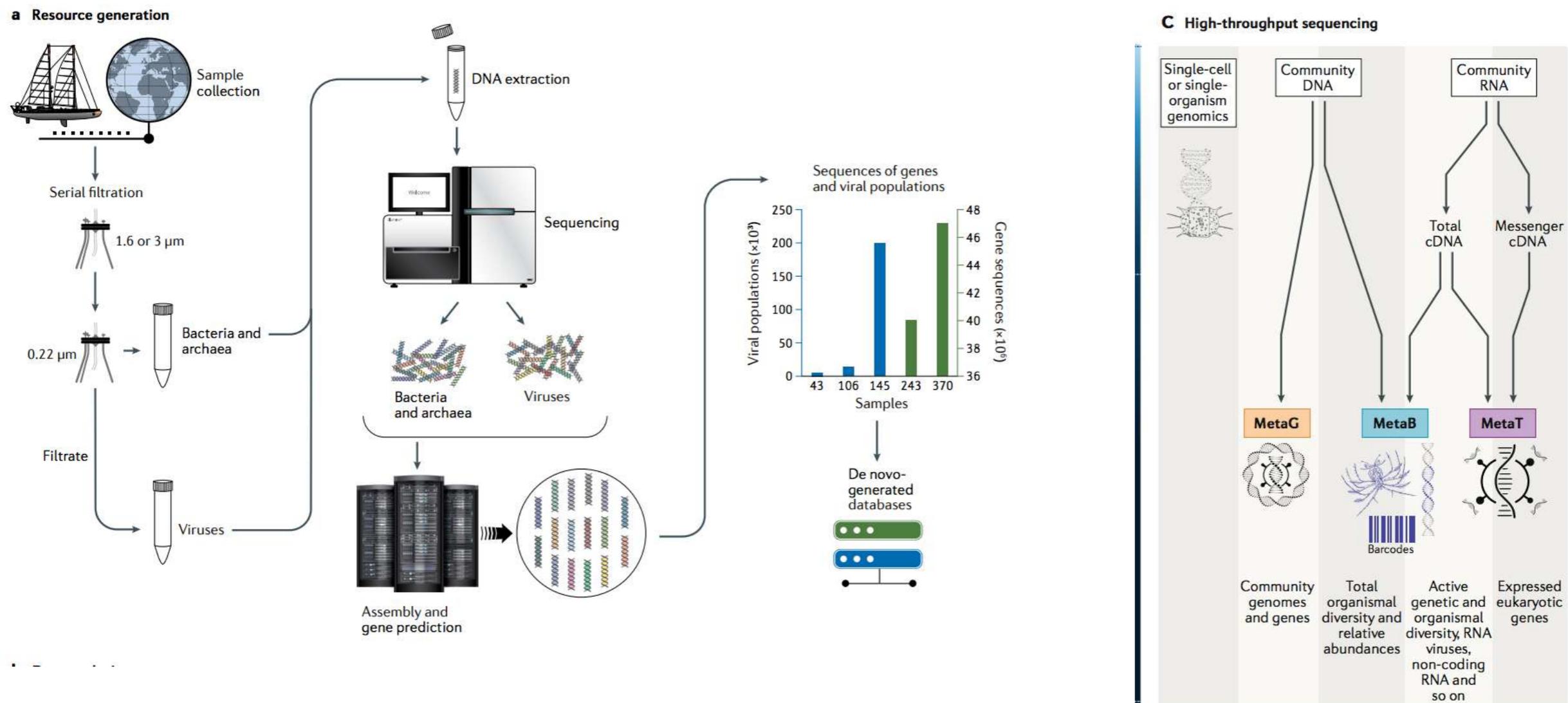


# TARA OCEANS



Sunagawa et al., 2020

# DNA technology as a tool to characterize microbial ecosystems



# Sequencing effort dictates gene discovery in marine microbial metagenomes

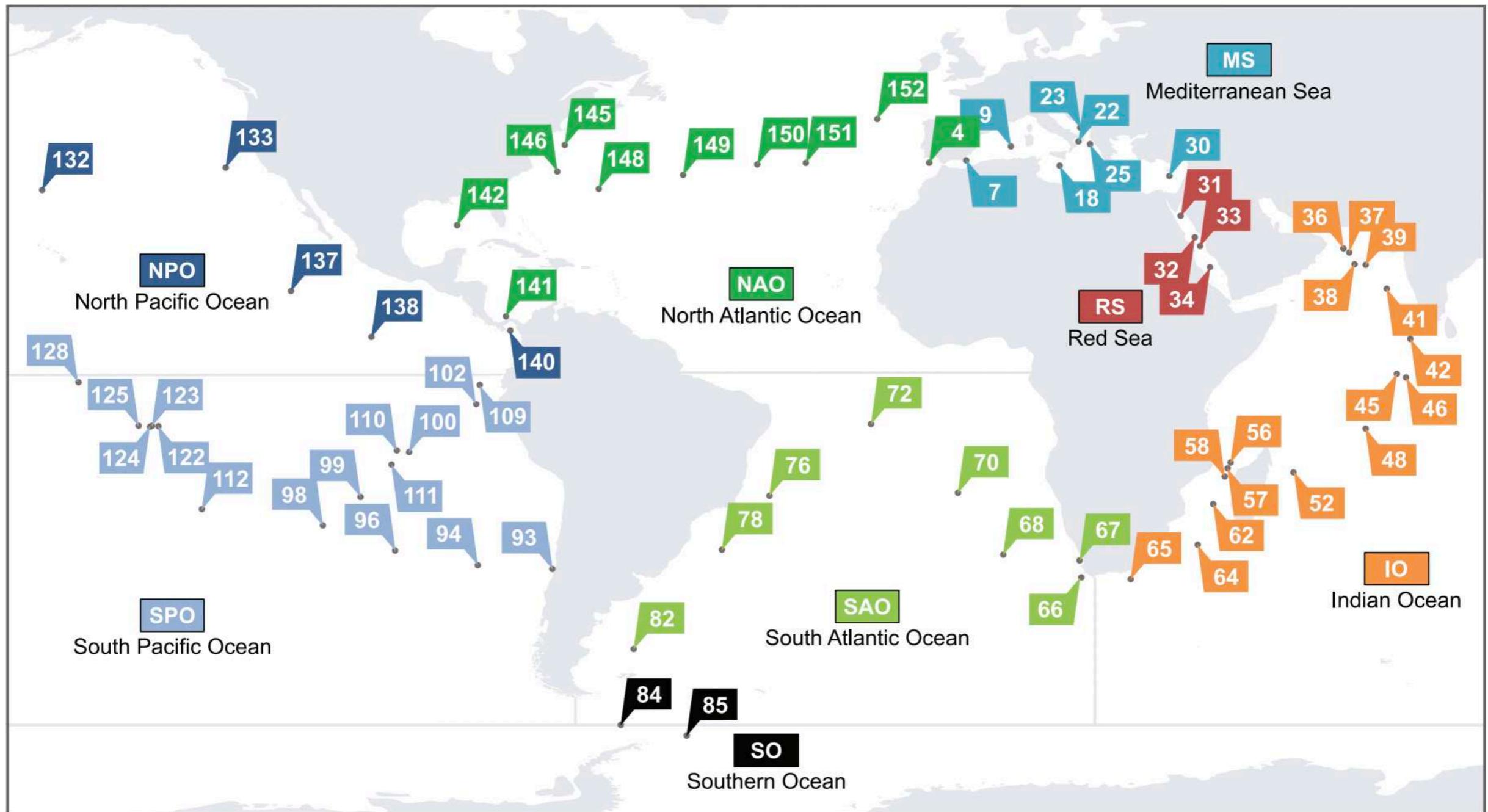
**Table 1.** The number of non-redundant gene sequence clusters predicted in various metagenome projects exploring marine pelagic microbial communities and mammalian enteric microbiomes and corresponding yield relative to the sequencing depth applied.

Project (gene catalogue)	Analytical procedures <sup>b</sup>	Samples	Sequenced depth (Tbp) <sup>c</sup>	Gene/protein sequences ( $\times 10^6$ ) <sup>a</sup>			Yield ( $10^6$ per Tbp)		Original data source <sup>f</sup>
				Redundant	Non-redundant		Genes	Proteins	
					Genes <sup>d</sup>	Proteins <sup>e</sup>			
<b>Marine</b>									
Global Ocean Sampling (GOS)	GP + GC	44	0.00625	13.6	4.5	3.9	720.0	624.0	Rusch <i>et al.</i> (2007)
Baltic Sea reference metagenomes	GP + GC	81	0.586	8.7	8.6	8.3	14.7	14.2	Hugerth <i>et al.</i> (2015)
Tara Ocean (OM-RGC)	GP + GC	243	5.821	61.3	33.3	27.7	5.7	4.8	Sunagawa <i>et al.</i> (2015)
RSCK2011 <sup>g,h</sup>	AM + GP + GC	45	0.0483	2.0	1.3	1.2	26.9	24.8	Thompson <i>et al.</i> (2017)
Station ALOHA (HOTGC) <sup>g,i</sup>	RAM + GP + GC	103	0.638	47.3	29.6	26.1	46.4	40.9	Mende <i>et al.</i> (2017)
GEOTRACES program <sup>g,j</sup>	GP + GC	610	5.024	72.9	29.1	24.1	5.8	4.8	Biller <i>et al.</i> (2018)
MALASPINA-Deep (MDSGC) <sup>g</sup>	RAM + GP + GC	60	0.121	11.8	6.7	6.3	55.4	52.1	Acinas <i>et al.</i> (2019)
MALASPINA-profiles (MRGC) <sup>g</sup>	AM + GP + GC	116	1.714	61.6	32.7	29.0	19.1	16.9	P. Sanchez <i>et al.</i> in prep.
MESOCOSM <sup>g</sup>	AM + GP + GC	65	0.163	5.1	2.8	2.6	17.2	16.0	This study
<b>Non-marine</b>									
Human gut microbiome I <sup>g</sup>	GP + GC	124	0.577	9.7	4.1	3.8	7.1	6.6	Qin <i>et al.</i> (2010)
Human gut microbiome II	GP + GC	1267	6.298	121.3	18.5	16.0	2.9	2.5	Li <i>et al.</i> (2014)
Mouse gut microbiome <sup>k</sup>		184	0.781	22.2	2.6	n.d.	3.3	n.d.	Xiao <i>et al.</i> (2015)
Rat gut microbiome	GP + GC	98	0.222	26.8	7.6	6.9	34.2	31.0	Pan <i>et al.</i> (2018)
Pig gut microbiome <sup>k</sup>		287	1.761	62.9	7.7	n.d.	4.4	n.d.	Xiao <i>et al.</i> (2016)

Abbreviations: OM-RGC, Ocean Microbial Reference Gene Catalogue; MDGC, Malaspina Deep-Sea Gene Collection; MPRGC, Malaspina Reference Gene Catalogue; Red Sea Centre Cruise 2011; HOTGC, Hawaii Ocean Time-series Gene Catalogue, AM, assembled; RAM, re-assembled; GP, gene prediction; GC, gene cluster with mmseq2; n.d. not determined.

# TARA OCEANS (0-1000 m)

## A *Tara* Oceans sampling stations



# Proteobacteria

- **Most abundant sequences in ocean environment**
- **Metabolic functions: autotrophy, heterotrophy, an oxygenic phototrophy, methylotrophy, sulfate reduction, nitrogen fixation**
- **O<sub>2</sub> requirement**
- **CO<sub>2</sub> fixation**
- **Alpha-**
- **Beta-**
- **Gamma-**
- **Delta-**
- **Epsilon**
- **Zeta-**

# Alphaproteobacteria

- From oligotrophic genome to copiotrophic lifestyle including intracellular symbiont
- From 20-50 % of bacterial abundance via 16SrRNA gene (amplicon sequencing)
- **Caulobacterales**: cellular differentiation with stalk, extrusion → epigenetic control → important for biofilm formation and larval settlement (e.g. corals, sponges, oysters..)
- **Rhodospirillales** and **Magnetococcales**: magnetotaxis → have magnetic particles → oxic/anoxic interface → 50MYA speciation (in salt marshes and sediment)
- SAR11, most abundant bacterium

# Betaproteobacteria

- Copiotrophic heterotrophic lifestyle associated with plants and animals
- Chemolithoautotrophic: sulfur oxidizing bacteria
- From ~5-10 % of bacterial abundance via 16SrRNA gene (amplicon sequencing)
- Include also rare biosphere

# Gammaproteobacteria

- Present everywhere with the exception of hot or alkaline environment
- Alteromonadales, Oceanospirillales, Vibrionales → almost entirely marine
- From 25-30% of bacterial abundance via 16SrRNA gene (amplicon sequencing)
- From oligotrophic genome to copiotrophic lifestyle
- Many are uncultivated of sulfide-oxidizing bacteria as intracellular symbiont of invertebrate (e.g. tube worm) and mat and sediment
- Vibrionaceae: pathogens and symbiont, large genome, 2 chromosomes
- Oceanospirillales: break down complex organic compounds → with Osedax worm and sediment by whale falls, hydrocarbon spill (i.e. Alcanivorax); degradation of DMSP (=dimethylpropionosulfonate)
- Thiotrichaceae: filamentous forms, chemotaxis, sulfur-oxidizing chemolithotrophs (using nitrate in anoxic conditions)

# Deltaproteobacteria

- **Anaerobic sulfate bacteria and sulfur-reducing bacteria (SRB)**
- **Sediments, anoxic mud, rich of organic matter, particle associated**
- **With methanogenic archaea → organic matter remineralisation in coastal and shelf-sediment**
- **SRB: organic compounds and  $H_2$  → dissimilative sulfate reduction →  $H_2S$  is toxic but chemotrophic and phototrophic bacteria can use it**
- **Cable bacteria (multi cells): coupling  $O_2$  reduction to water to sulfide oxidation to sulfate via long distance e- transport**

# Epsilonproteobacteria

- Anaerobic respiration with organic or inorganic electron donor and acceptors
- Hydrothermal vents → important primary producers chemolithoautotrophy
- Important in colonizing new vents
- Mycobacteria: large genome, social behaviours, gliding motility, make spores, prey on bacteria
- Bdellovibrionales: prey on Gram-Negative bacteria, top-down population control → new probiotic

# Zetaproteobacteria

- Hydrothermal vents, brine, Fe-rich waters
- Steel-corrosion
- Microaerophilic Fe(II)-oxidizing chemolithoautothrops

# Cyanobacteria

- Evolution 2.5 MYA
- Oxygenic photosynthesis, membranes to anchor photosystems
- Chl a, accessory pigments: phycobilins and phycoerythrin
- Surface ocean, sea ice, shallow sediments, microbial mats, on algae and animal surfaces
- Some anaerobically growth using H<sub>2</sub>, H<sub>2</sub>S, or organic compounds as electron donor
- Single cells to filaments, to endosymbionts
- Nitrogen fixation evolved about 3 BYA in anoxic conditions
  - UCYN-A in symbiosis with a microalga Prymnesiophyte, N<sub>2</sub> fixing
  - *Crocospaera*, filamentous, N<sub>2</sub> fixing
  - Microbial mats: secondary products, toxin production

# Cyanobacteria

✱ *Prochlorococcus* → 20% O<sub>2</sub> production to the atmosphere

- Streamlined genome
- Different ecotype, according to light preferences
- Assimilate to glucose and sugars
- Produce outer membrane vesicles

✱ *Synechococcus* → very abundant

- Different ecotypes, open ocean vs coastal
- Circadian clock regulation

✱ *Trichodesmium* → colony-former

- Nitrogen fixation
- Tropical waters

# Planctomycetes

- **Floating fungi**
- **Marine sediments, surfaces and marine waters**
- **Expandable periplasmic space**
- **Belong to superphylum: Planctomycetes-Verrucomicrobia-Chlamydiae, PVC**
- **Aerobic chemoheterotrophs**
- **From 6% of bacterial abundance via 16SrRNA gene (amplicon sequencing) and < 1% in open ocean**
- **Associated with marine snow and algae bloom**
- **Break down complex heteropolysaccharides**
- **Annamoxisome, anaerobic oxidation of ammonia**

# Bacteroidetes

- From 20% of bacterial abundance via 16SrRNA gene (amplicon sequencing) in coastal and open ocean
- Attached to phytoplankton particulate aggregates, sediment, hydrothermal vents, polare region and associated with animals
- Pigments, carotenoid
- Degrade agar and simple and complex carbohydrates (e.g. chitin)
- Present in human gut—> pollution indicator

# Chloroflexi

- From 40% of bacterial abundance via 16SrRNA gene (amplicon sequencing) in deep ocean
- Wide habitat
- Diverse metabolisms: anoxygenic photosynthesis, aerobic or anaerobic heterotrophy
- Degradation of hydrocarbons
- DOM to refractory DOM
- Still unknown

# Aquificae & Thermotogae

- Primitive branching
- Aquificae:
  - \* Thermophilic, microaerophilic oxidizing  $H_2$  → using  $H_2$ , tiosulfate or sulfur as e-donor and  $O_2$  as e- acceptor, fixing C via reductive TCA cycle
  - \* Primary production in hydrothermal vent system
- Thermotogae:
  - \* Anaerobic, thermophilic, as well as mesophilic
  - \* Unique outer membrane
  - \* Shallow and deep hydrothermal vents
  - \* Fermentative and anaerobic chemoorganotrophs → carbohydrates
  - \* Fix  $N_2$  and reduce sulfur to  $H_2S$

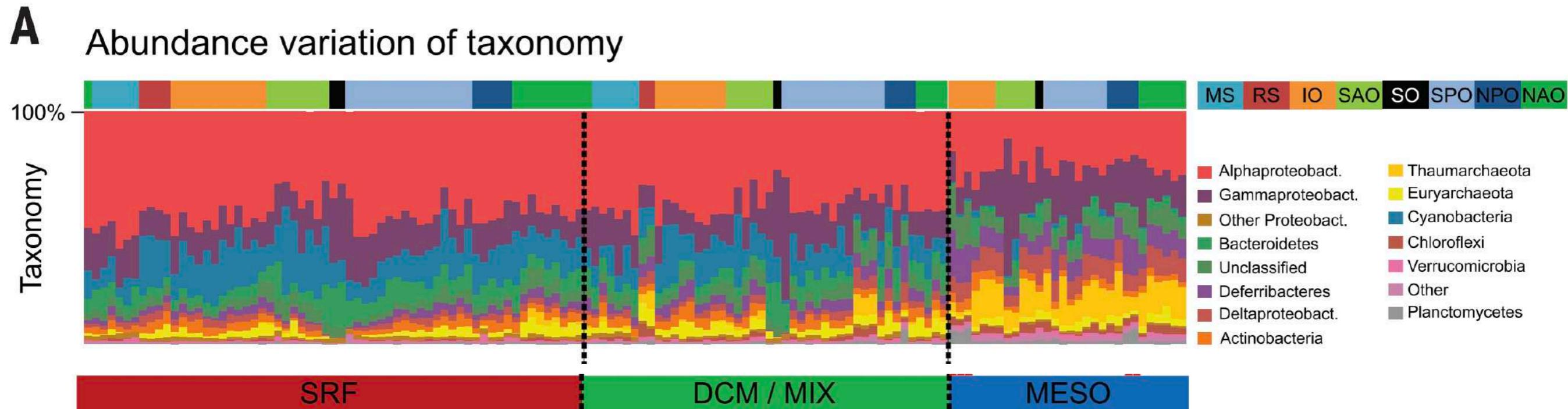
# Firmicutes

- Gram-positive, thick peptidoglycan
- Low G-C content
- Free-living or filamentous
- Bacilli and Clostridia
- ~4 % in water column and ~ 8% in sediment bacterial abundance via 16SrRNA gene (amplicon sequencing)
- Produce spores
- Degrade complex polysaccharides → gut associated

# Actinobacteria

- Gram-positive
- High G-C content
- Free-living or filamentous
- Sediments, marine particles in water column
- Decompose organic matter
- Produce secondary metabolites and antibiotics

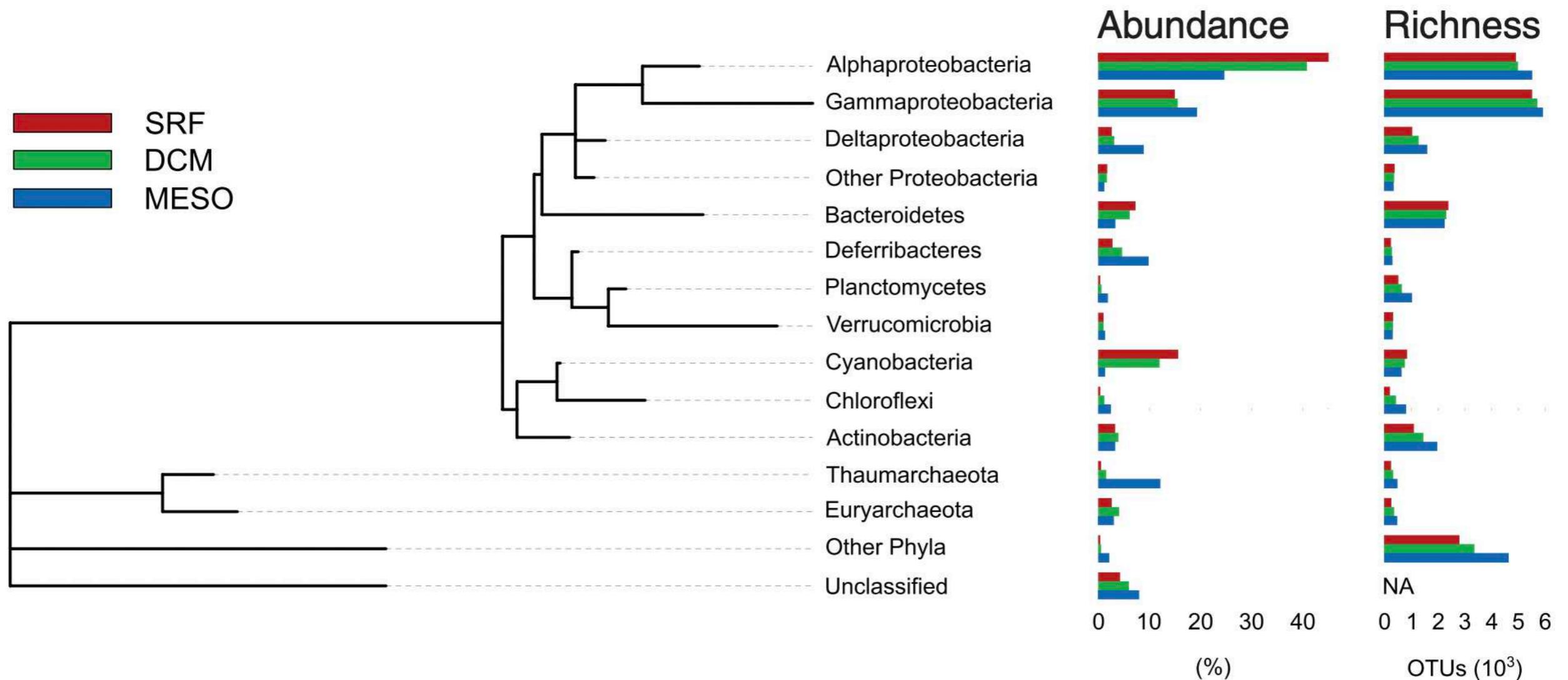
# Tara: 0-1000 m water column, I



- Alphaproteobacteria the most abundant in surface and DCM layers
- Alpha- and Gammaproteobacteria and Bacteroidetes most abundant in surface and DCM layers
- In Mesopelagic, Deltaproteobacteria, Actinobacteria and Archaea are dominant with Alpha- and Gammaproteobacteria
- Cyanobacteria dominant in surface waters



# Tara: 0-1000 m water column, II

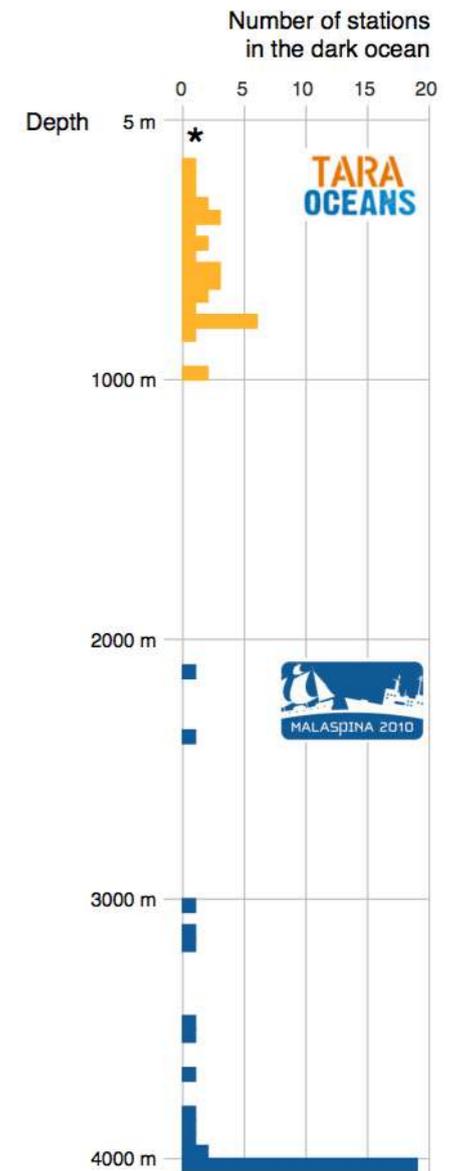


- Highest abundance of Alphaproteobacteria, Gammaproteobacteria and Cyanobacteria
- Highest richness of Alpha-, Gammaproteobacteria and Bacteroidetes

# MALASPINA: 2000-4000 m water column I

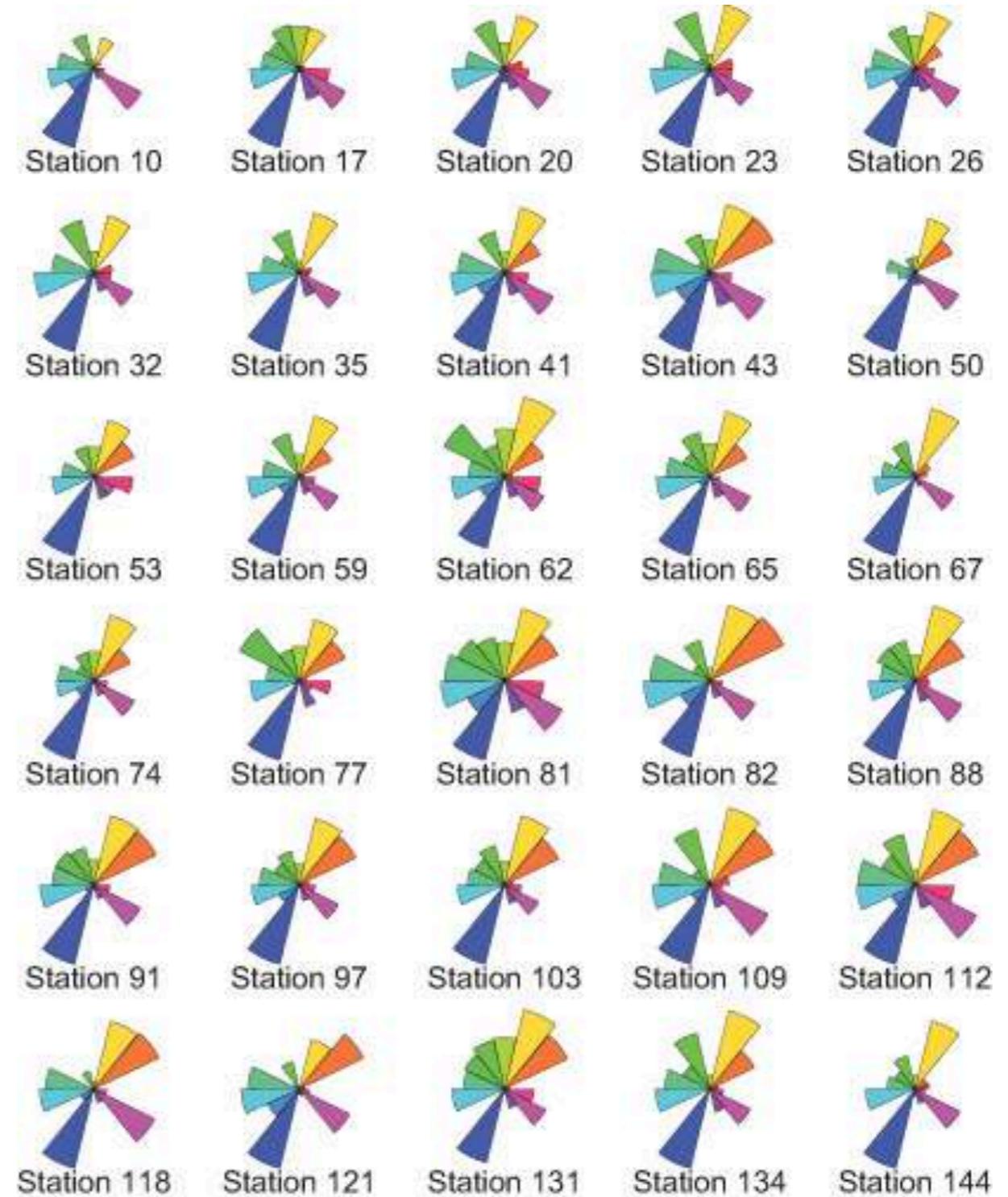
Fig. 1: Malaspina Deep Ocean Genetic Resources.

a

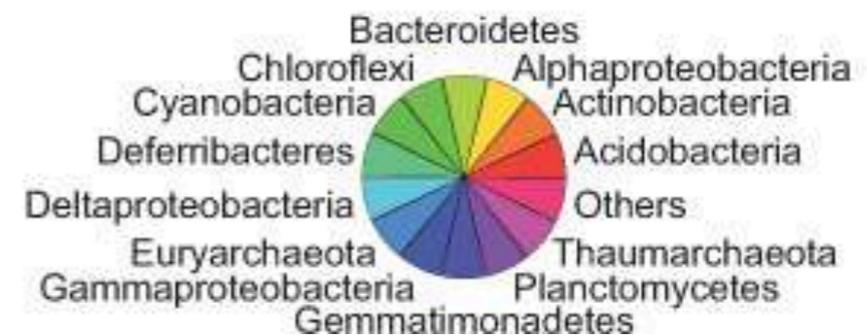


Acinas et al., 2019 and 2021

# MALASPINA: 2000-4000 m water column II



- Alpha- and Gammaproteobacteria the most abundant
- Thaumarcheota and Actinobacteria abundant



Salazar et al., 2015