

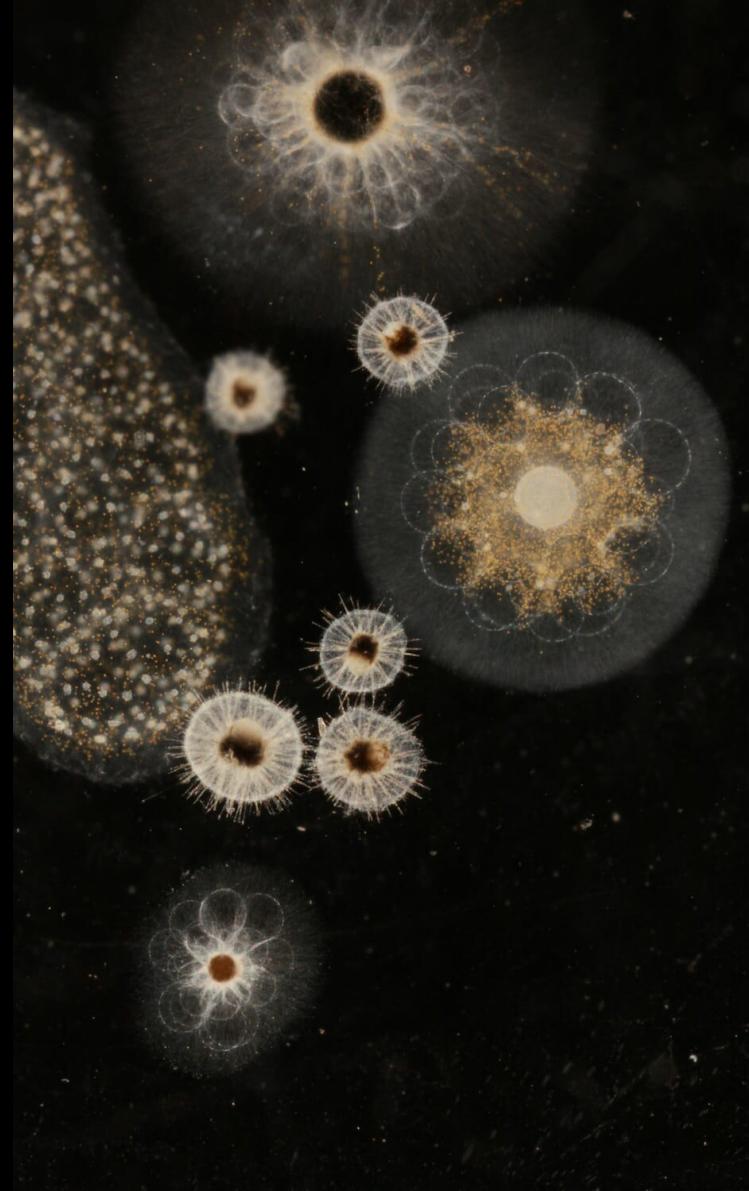


Qu'est ce que le plancton ?

Le plancton

**Ensemble des organismes portés
par le courant**

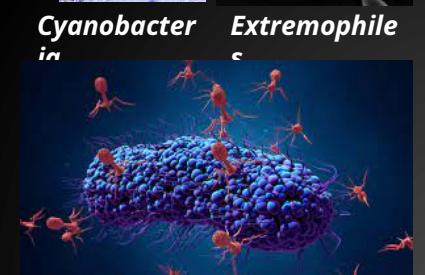
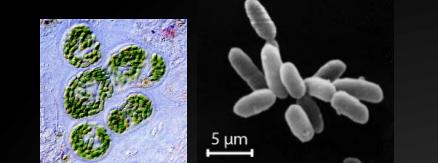
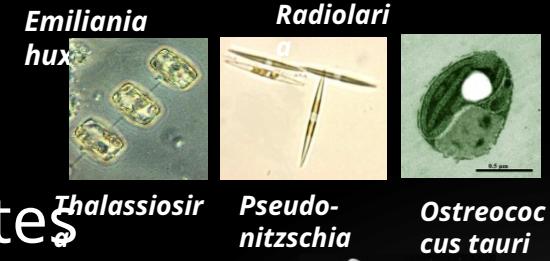
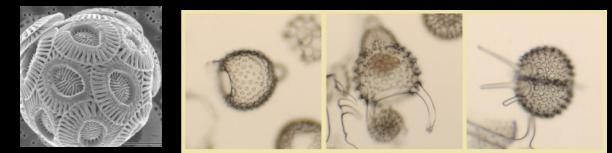
- Rôle écologique important
- Forte biomasse
- Immense diversité



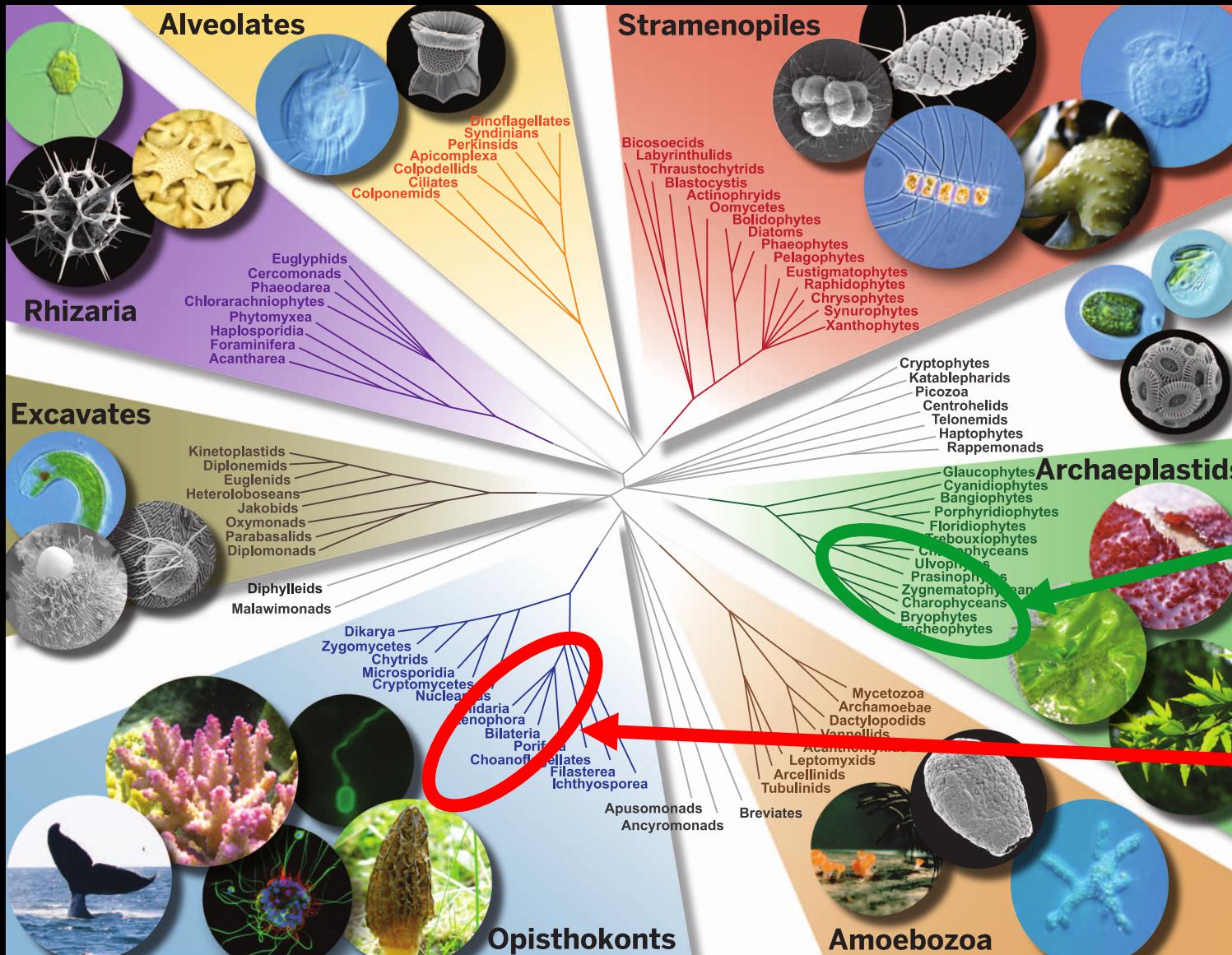


Dans 1 litre d'eau de mer

- 10.000 – 100.000 metazoaires
- 1 to 100 million protistes
- 1 to 10 milliards procaryotes
 - (bacteria and archaeabacteria)
- 1 to 10 milliards de viruses

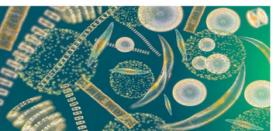
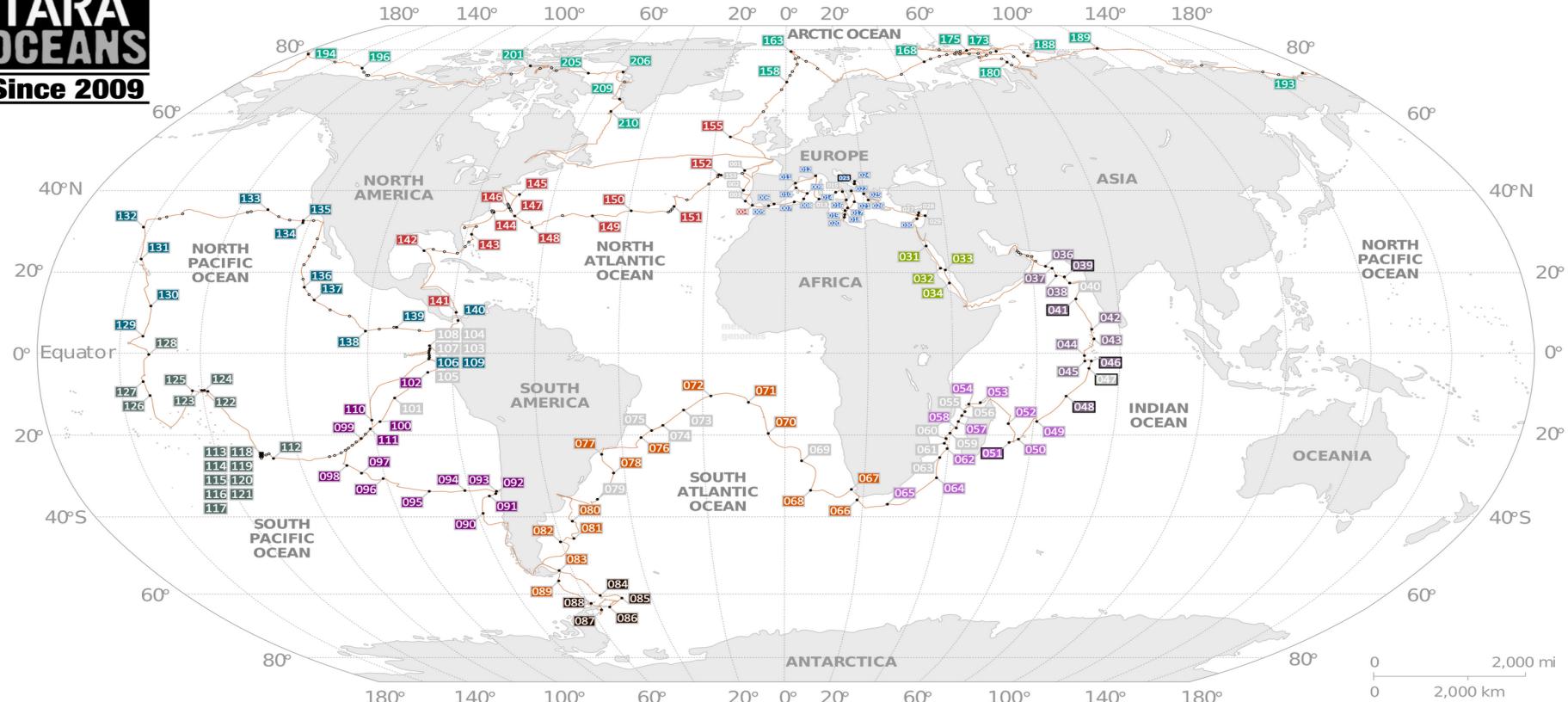


A huge diversity

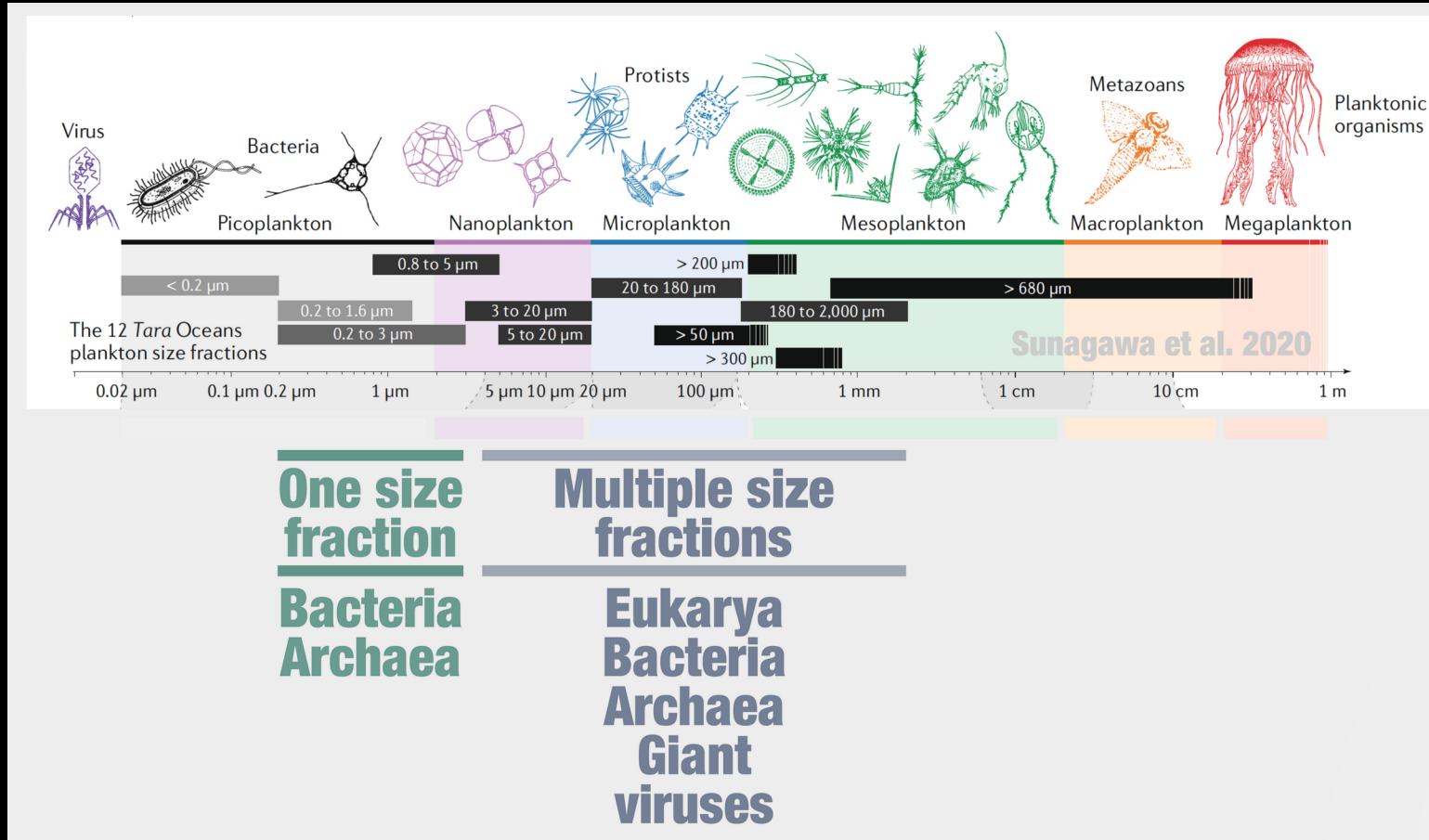


Sampling the oceans to understand plankton

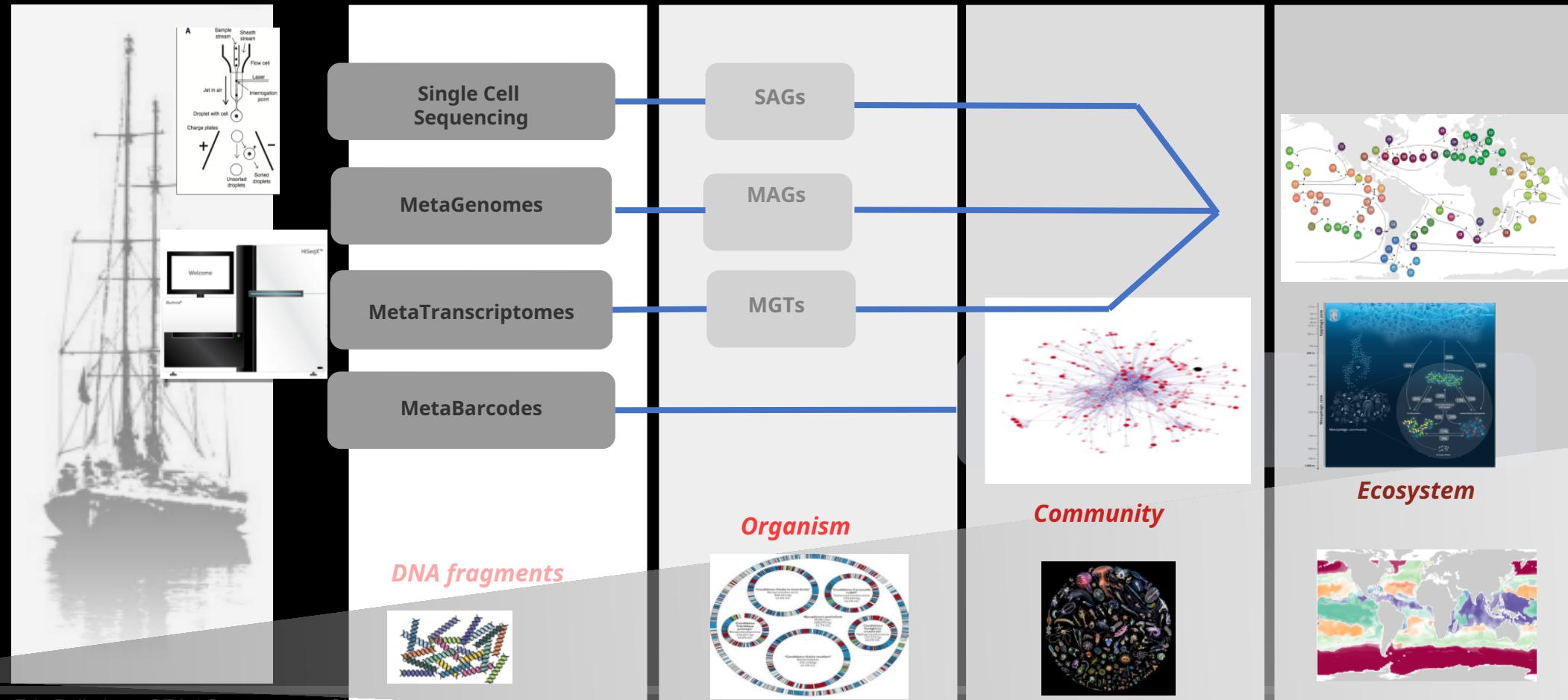
**TARA
OCEANS**
Since 2009



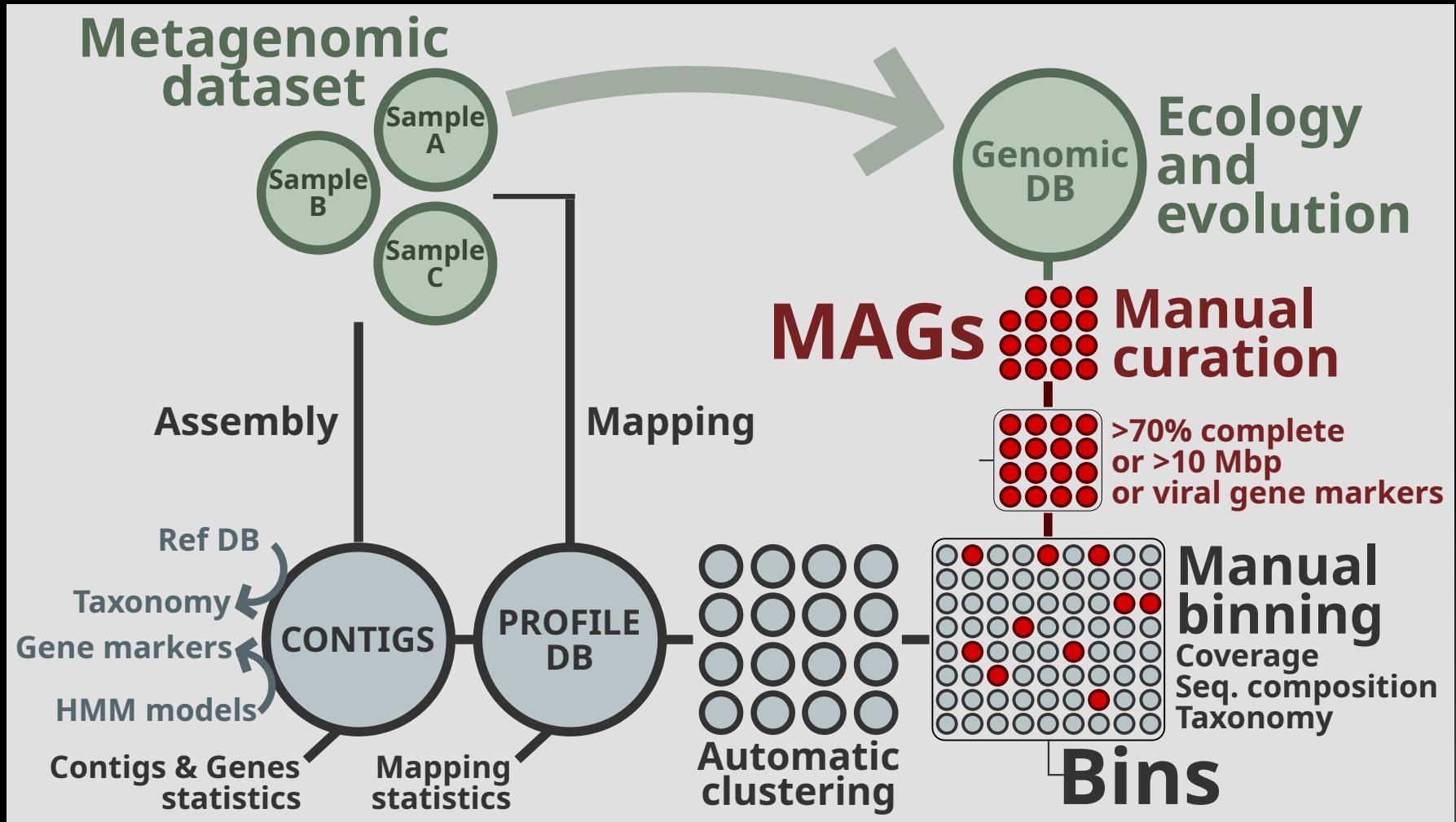
Fractionating sea water in different sizes to study different classes of organisms



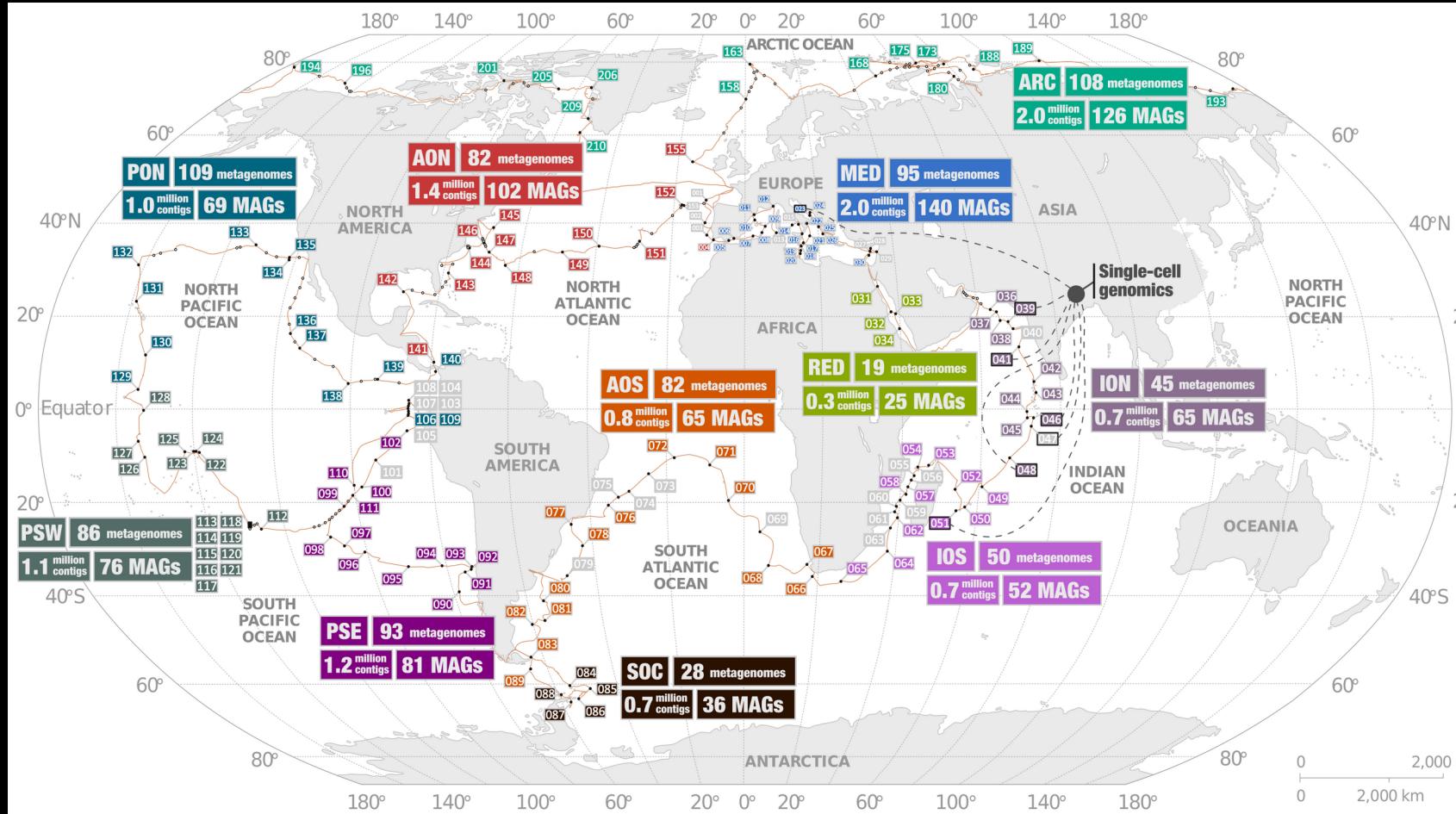
Tara Oceans Eukaryote Plankton From Meta Reads to Genomes ... to Ecosystems



Metagenome-based Assembled Genomes (MAGs)

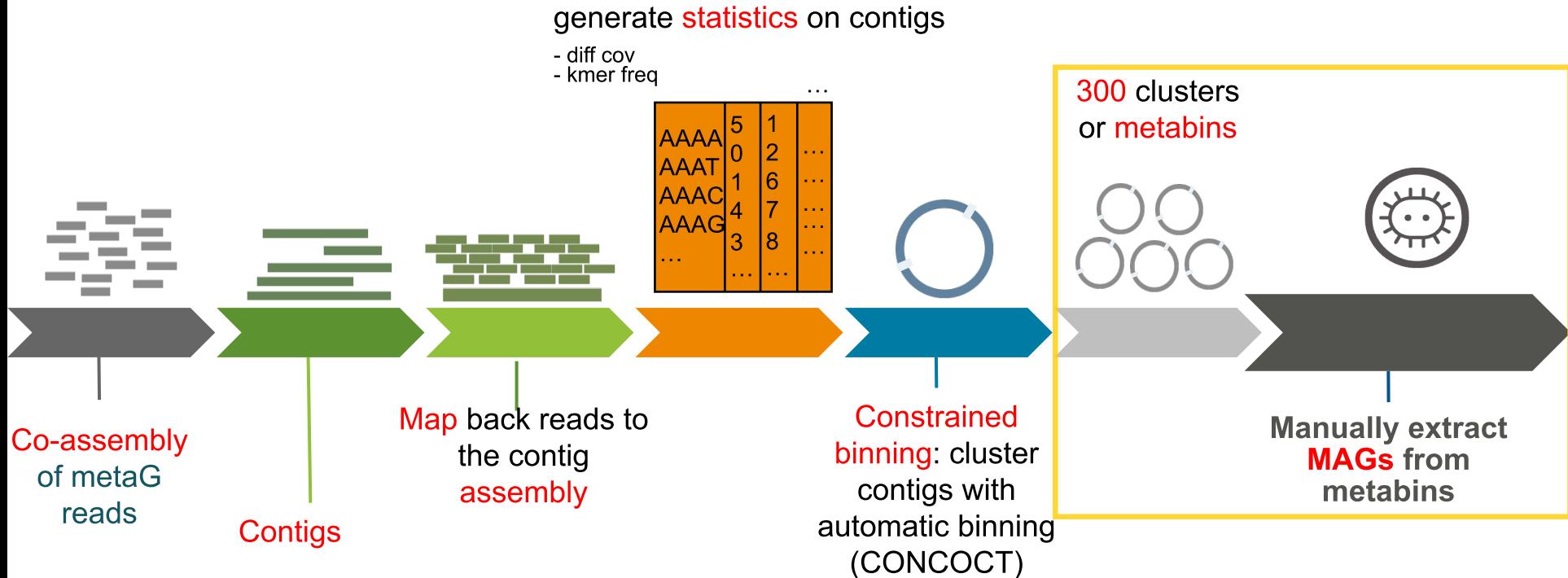


Metagenome Assembled Genomes (MAGs)



713 eukaryote SAGs and MAGs from *Tara Oceans* samples

Workflow for eukaryotic MAGs reconstruction



Un meta-bin est composé d'un ensemble de n contigs composant potentiellement m MAGs

Sur les contigs certains gènes sont prédis, dont certains nous servent de marqueurs.

Nous disposons des informations suivantes

- la liste des gènes marqueurs cibles pour déterminer la complétion/redondance des génomes (71 pour les bactéries et archées, 82 pour les eucaryotes)
- le fichier qui liste les gènes identifiés dans les différents contigs, avec l'affiliation taxonomique associée
- la matrice de distance entre contigs en terme de composition en tétranucléotides
- la matrice de distance entre contigs en terme de variation d'abondance dans les différents échantillons.

On vise à regrouper les contigs en MAGs de sorte que

- on minimise la somme des distances intra-MAG (fréquences tetramères et couverture dans les échantillons)
- on maximise les même distances inter-MAGs
- on maximise la completion par MAG (le ratio de gènes marqueurs trouvé sur le nombre attendu)
- on minimise la redondance (la présence de plusieurs fois le même marqueur par MAG)
- on minimise l'hétérogénéité taxonomique (le nombre de classes taxonomiques différentes vues par MAG).
- on maximise le nombre de contigs classifiés dans des MAGs