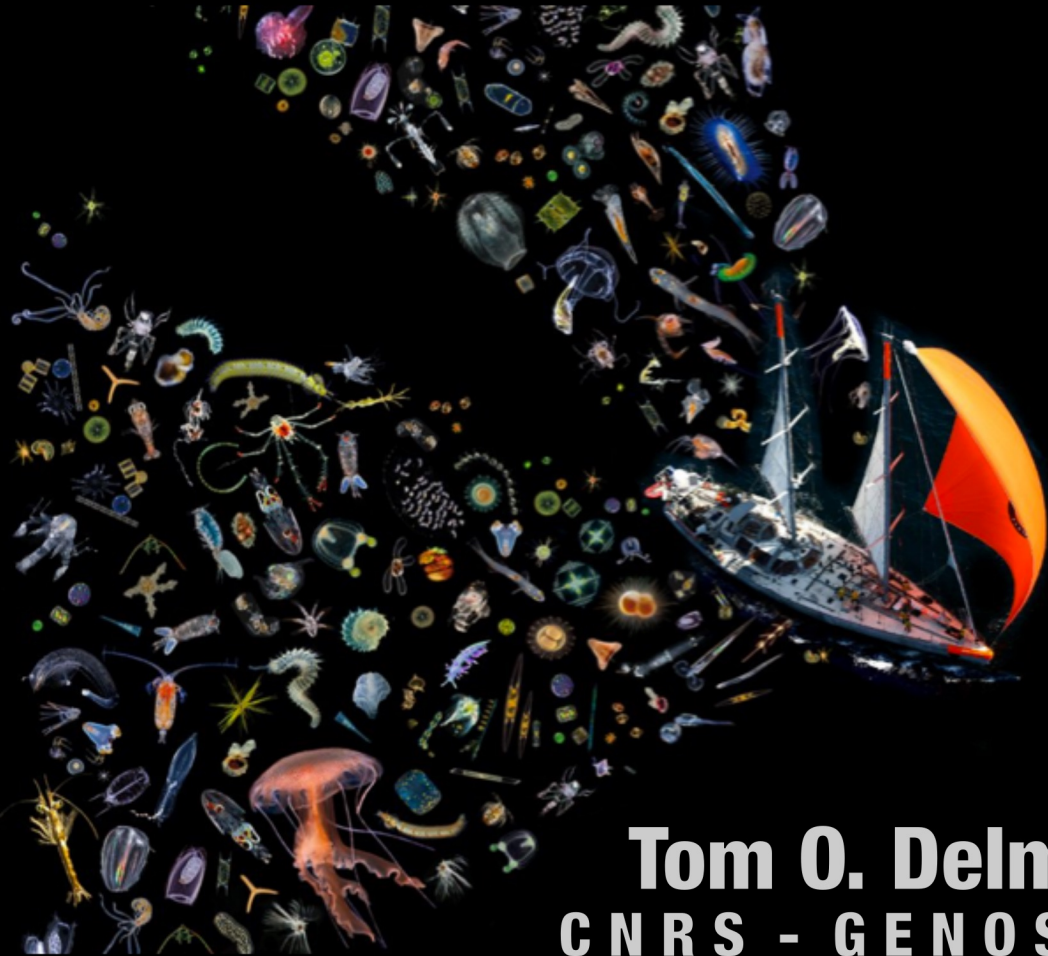


Plankton genomics with Tara Oceans



Tom O. Delmont
CNRS - GENOSCOPE

Plankton:

Millions of lineages

viruses

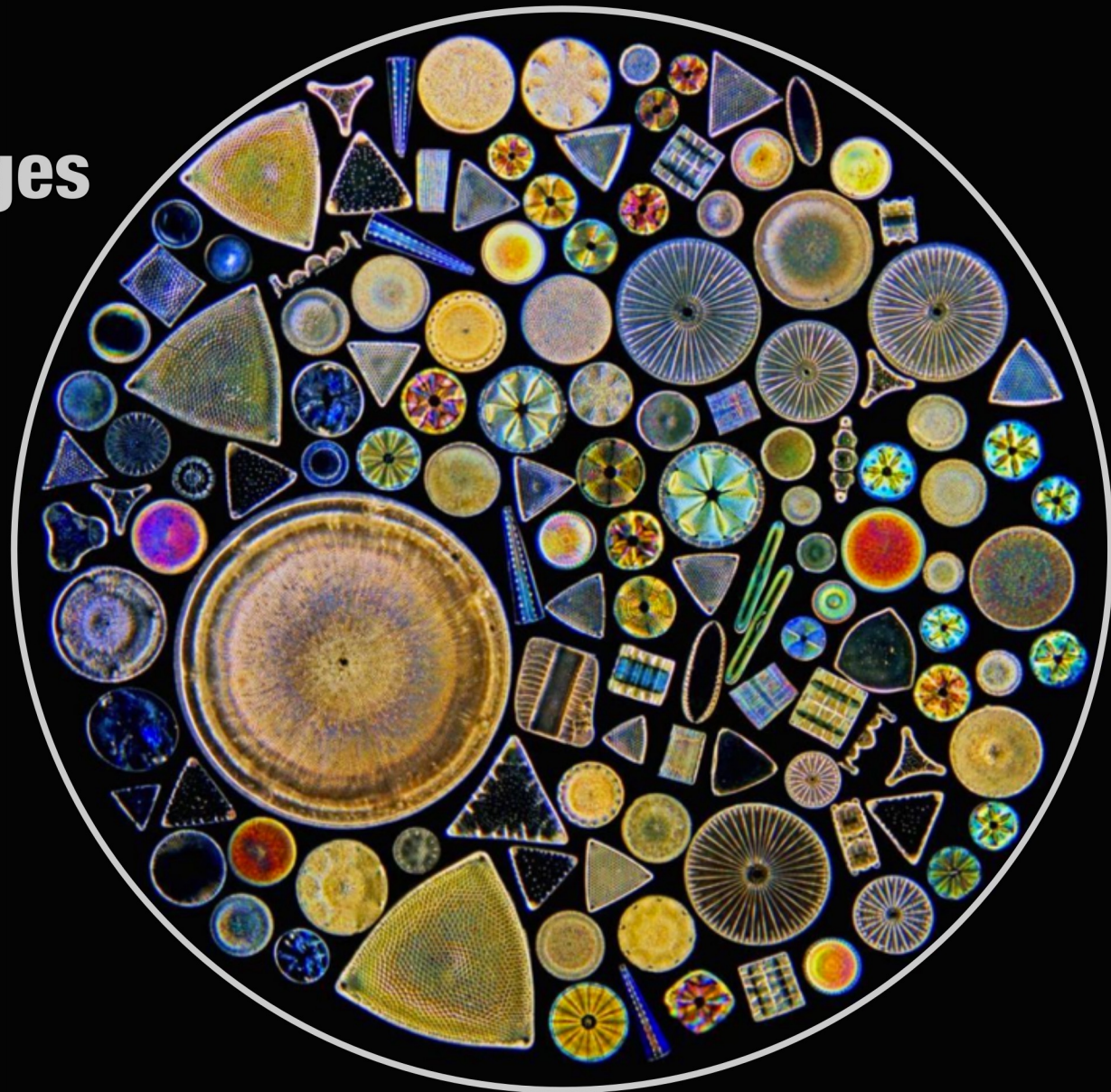
Archaea

Bacteria

Eukarya

Complex
ecological
networks

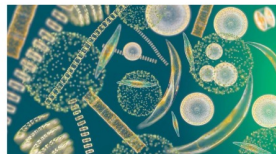
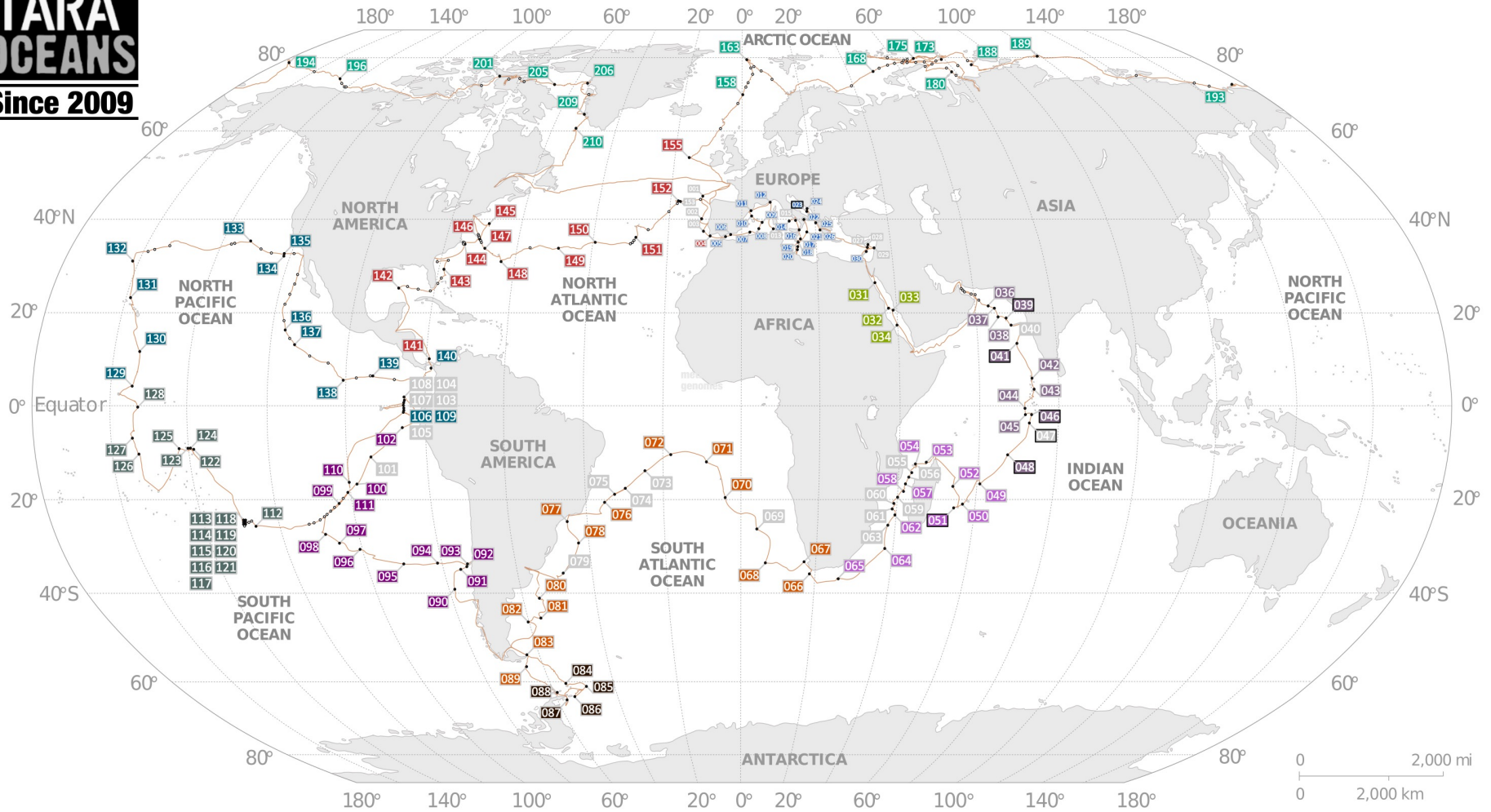
~50% of oxygen
production
on Earth



Diatoms represent
a tiny part of this diversity

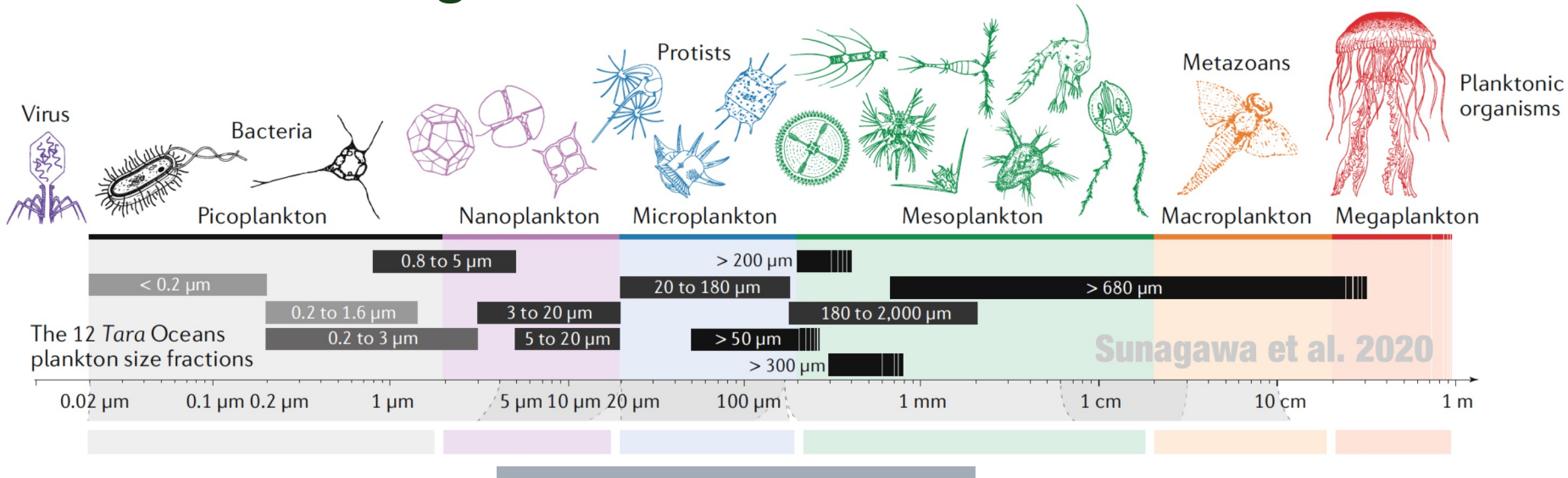
Sampling the oceans to understand plankton

TARA OCEANS
Since 2009



Equivalent of 10,000 human genomes sequenced

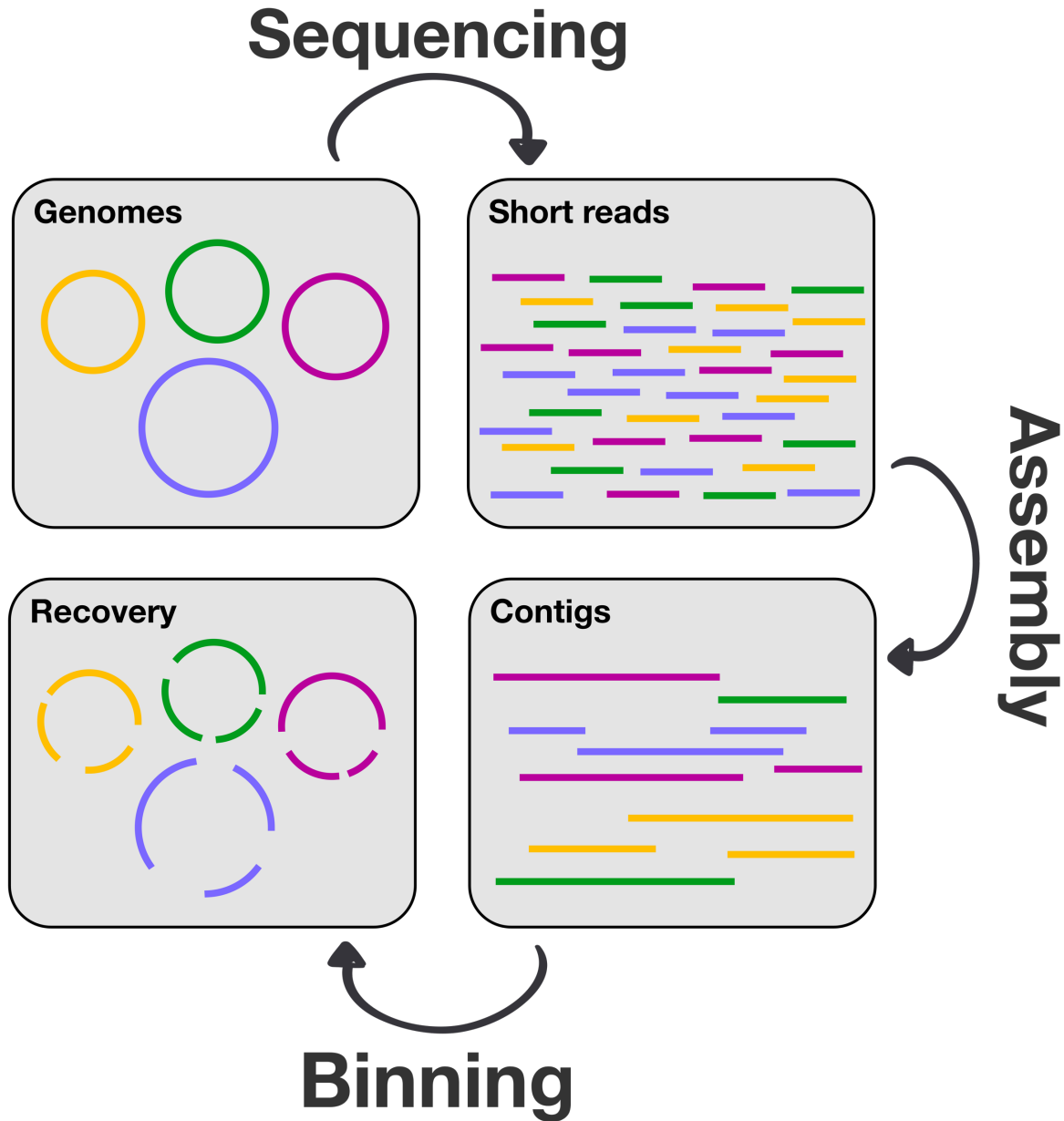
~300 billion metagenomic reads



A considerable opportunity to better understand the genomic diversity of plankton

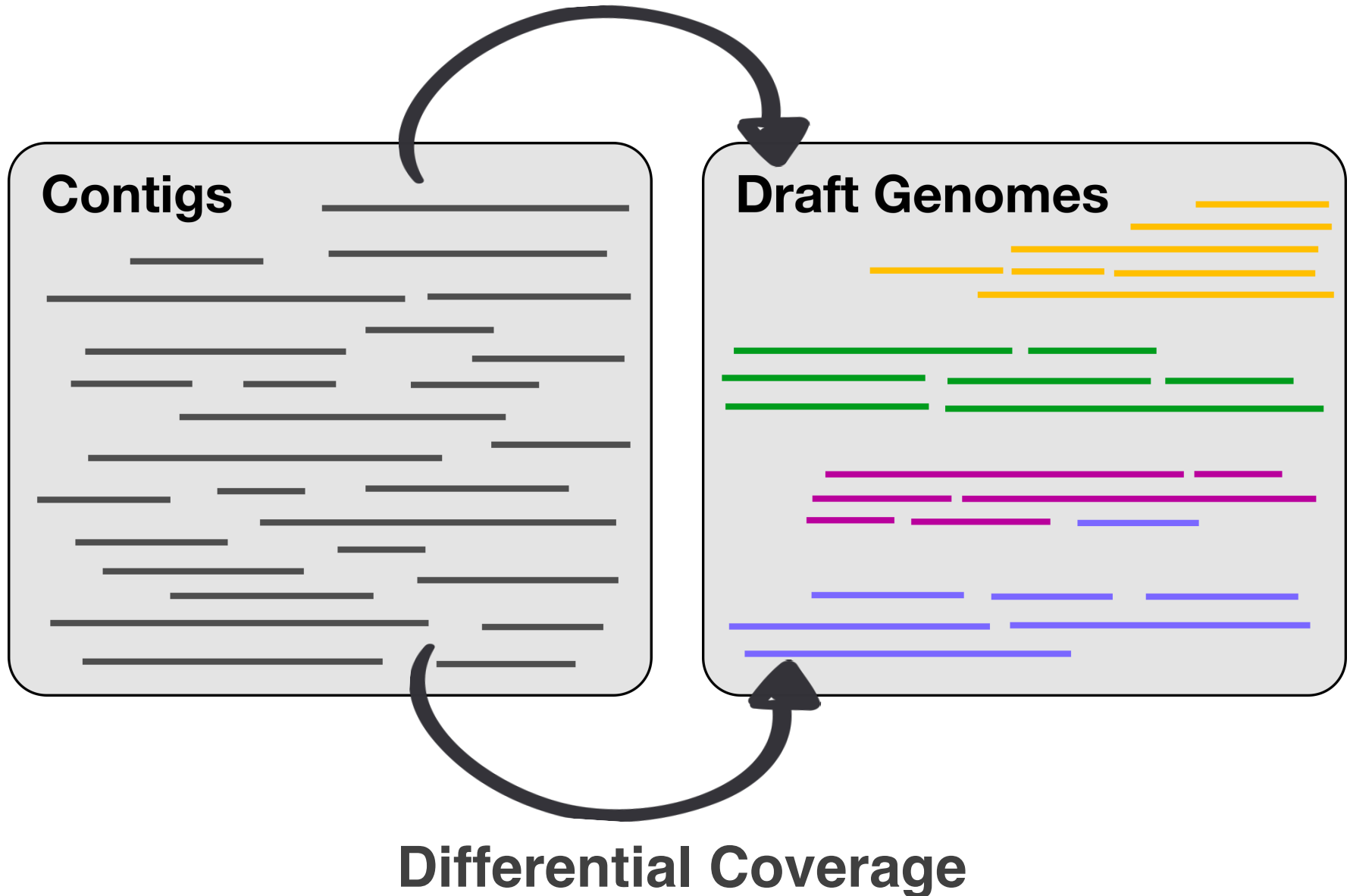
We need to solve a complex metagenomic puzzle

Genome-resolved metagenomics



Genome-resolved metagenomics

Genomic signatures





**Genome-resolved
metagenomics**

**Population
genetics**

**Comparative
genomics**

Phylogenomics

Eren et al., PeerJ 2015

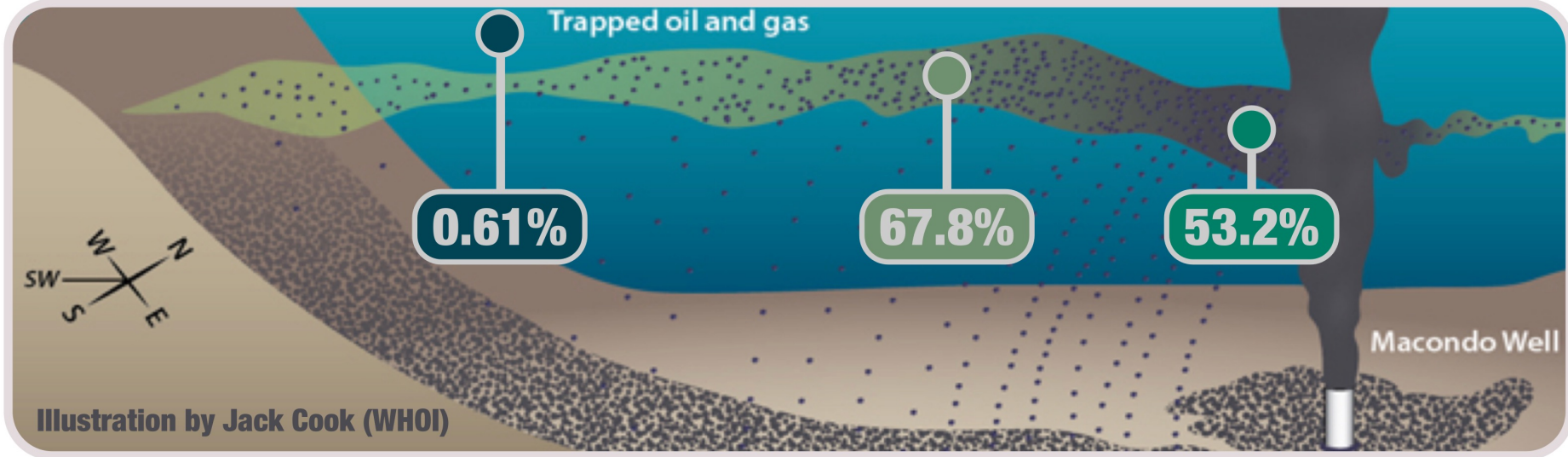
Eren et al., Nature Microbiology 2021

Integrated multi-omics at scale

An open-source, community-driven analysis and viualization platform for microbial omic.

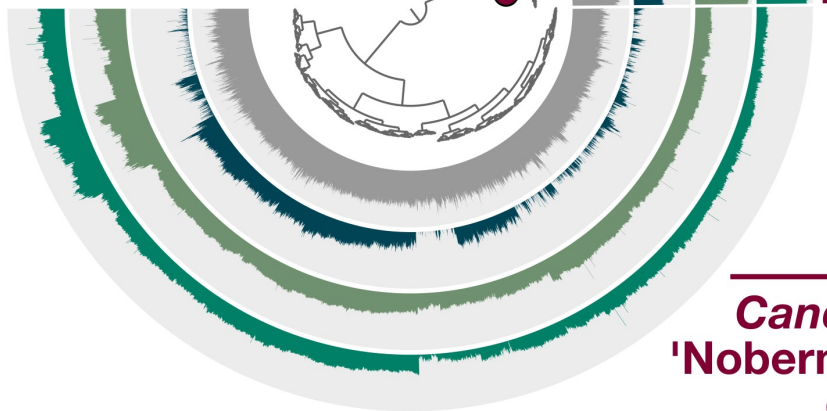
[IN A NUTSHELL](#)[INSTALL](#)[GET HELP](#)[LEARN](#)[PEOPLE](#)[COMMUNITY BLOG](#)['OMICS VOCABULARY](#)[HOW TO CITE](#)

anvio.org



PROXIMAL MGs
DISTAL MGs
CONTROL MGs
GC-content

6,123 contigs
 Binning métagénomique
 Composition de séquence
 et distribution



Delmont and Eren
 2017, PNAS

2,579x
3,530x
30x
GC-content

526 contigs
 Curation génomique
 Composition de séquence

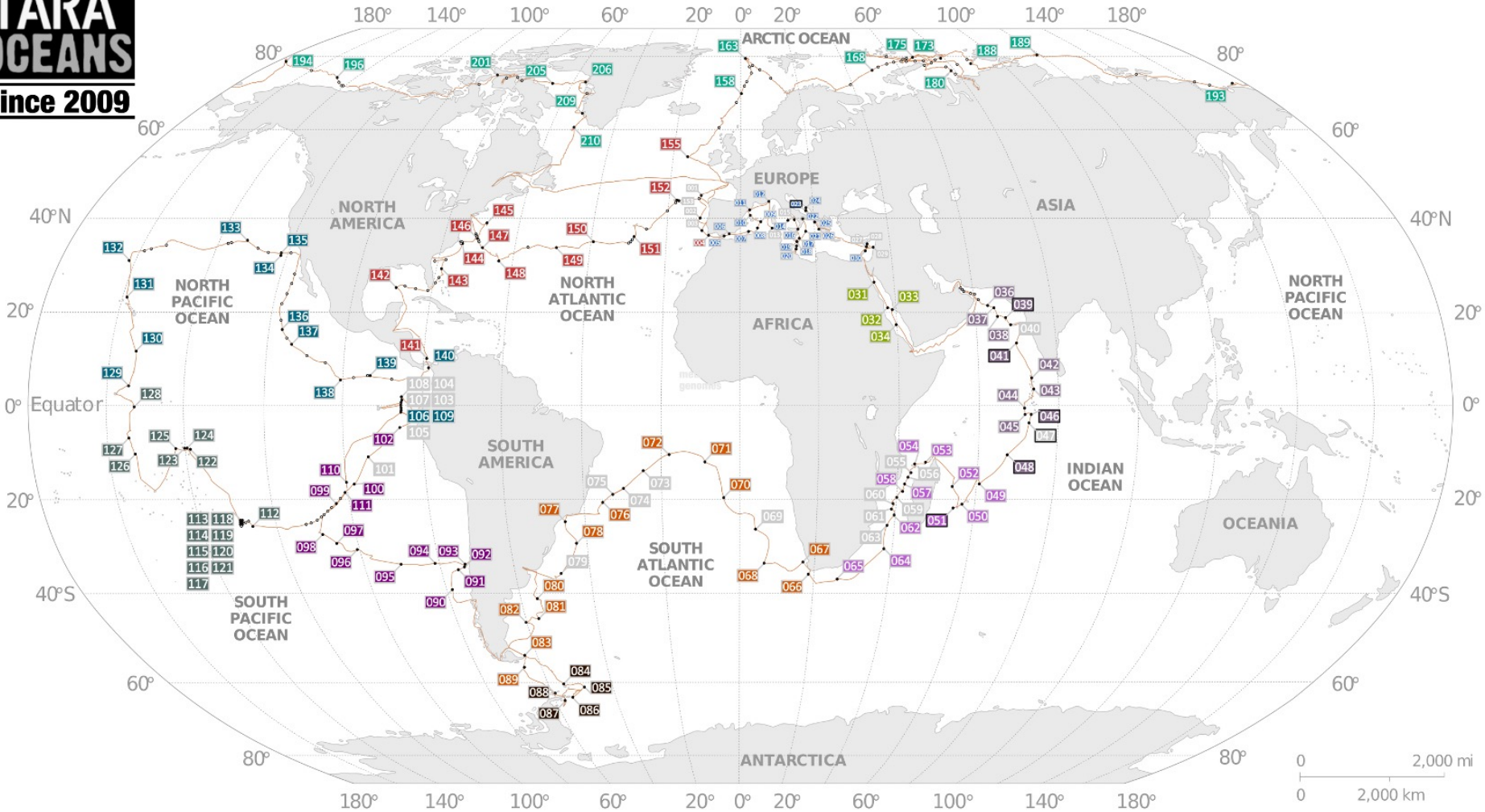


Candidatus
'*Nobermanella*
***desum*'**
 2.81 Mbp
 97% complet
 0% redondant

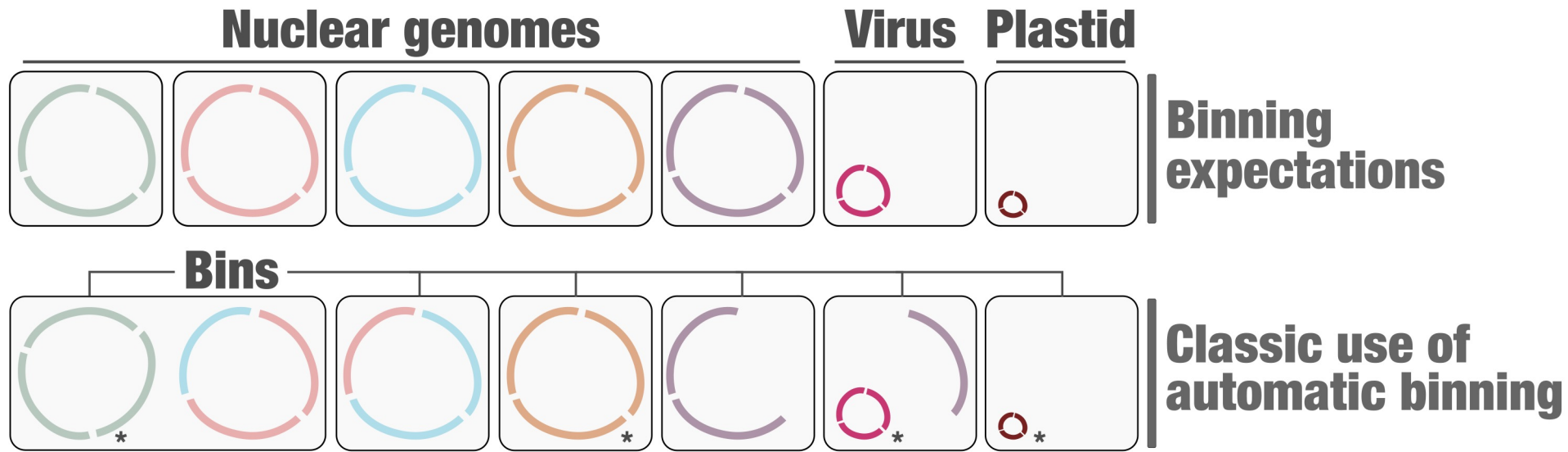


- Nearly 1,000 metagenomes
- 280 billion metagenomic reads
- 11 large co-assemblies by region
- 12 million contigs > 2.5 kbp

TARA OCEANS
Since 2009



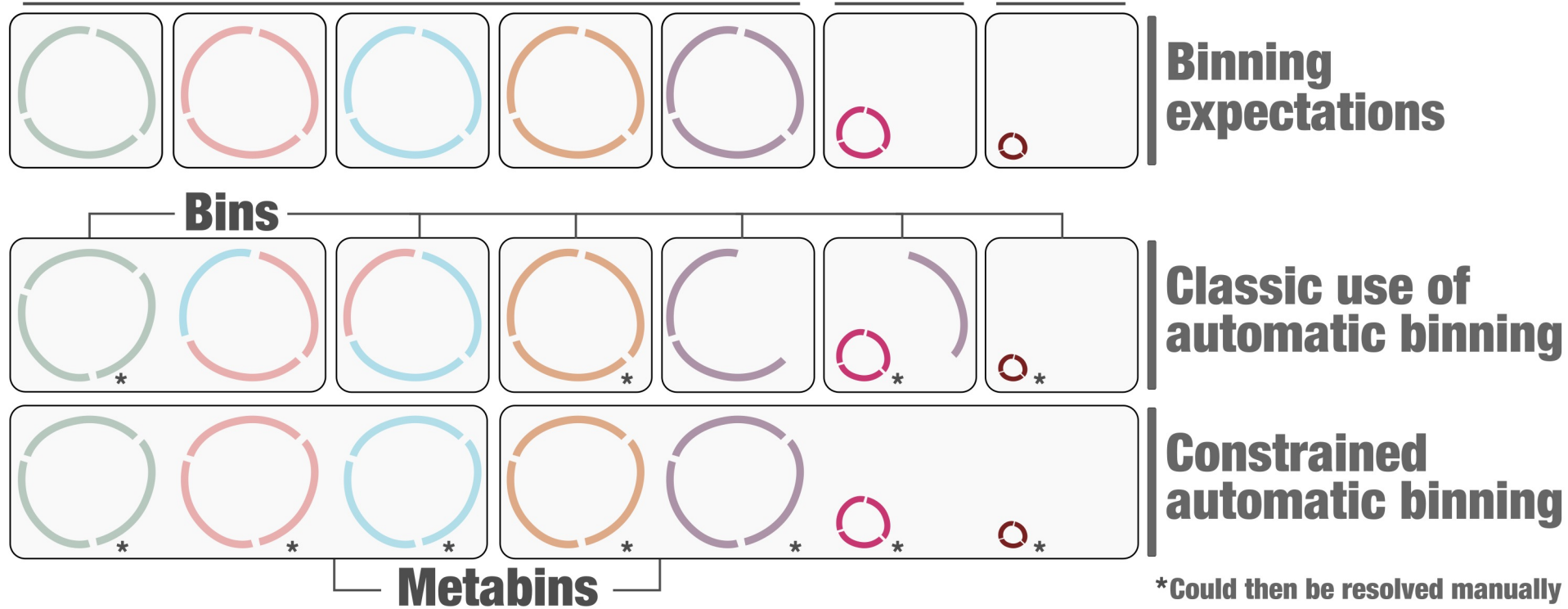
The “classic” automatic binning approach



“Constrained binning” and the metabins

Nuclear genomes

Virus **Plastid**

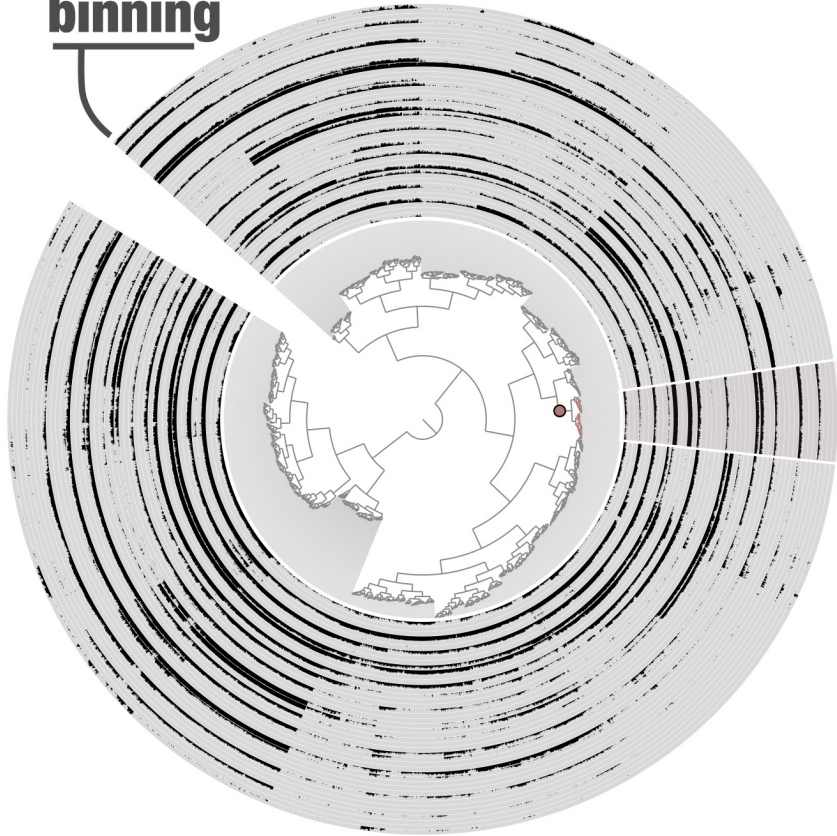




Total of 2.550 metabins

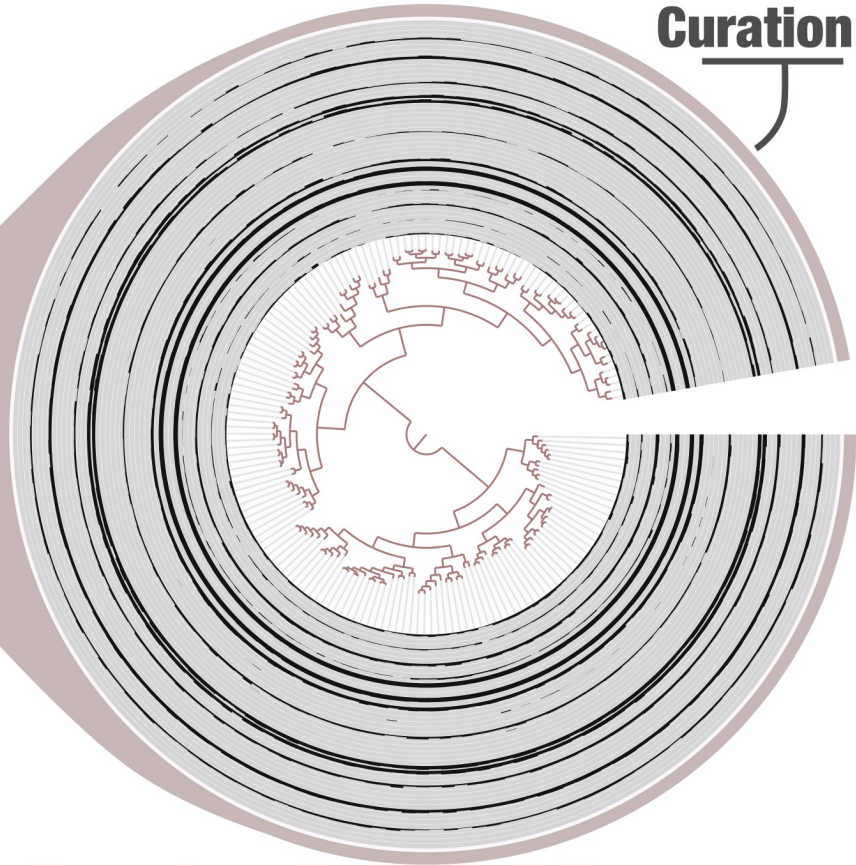
(large bins containing multiple genomes)

Manual binning



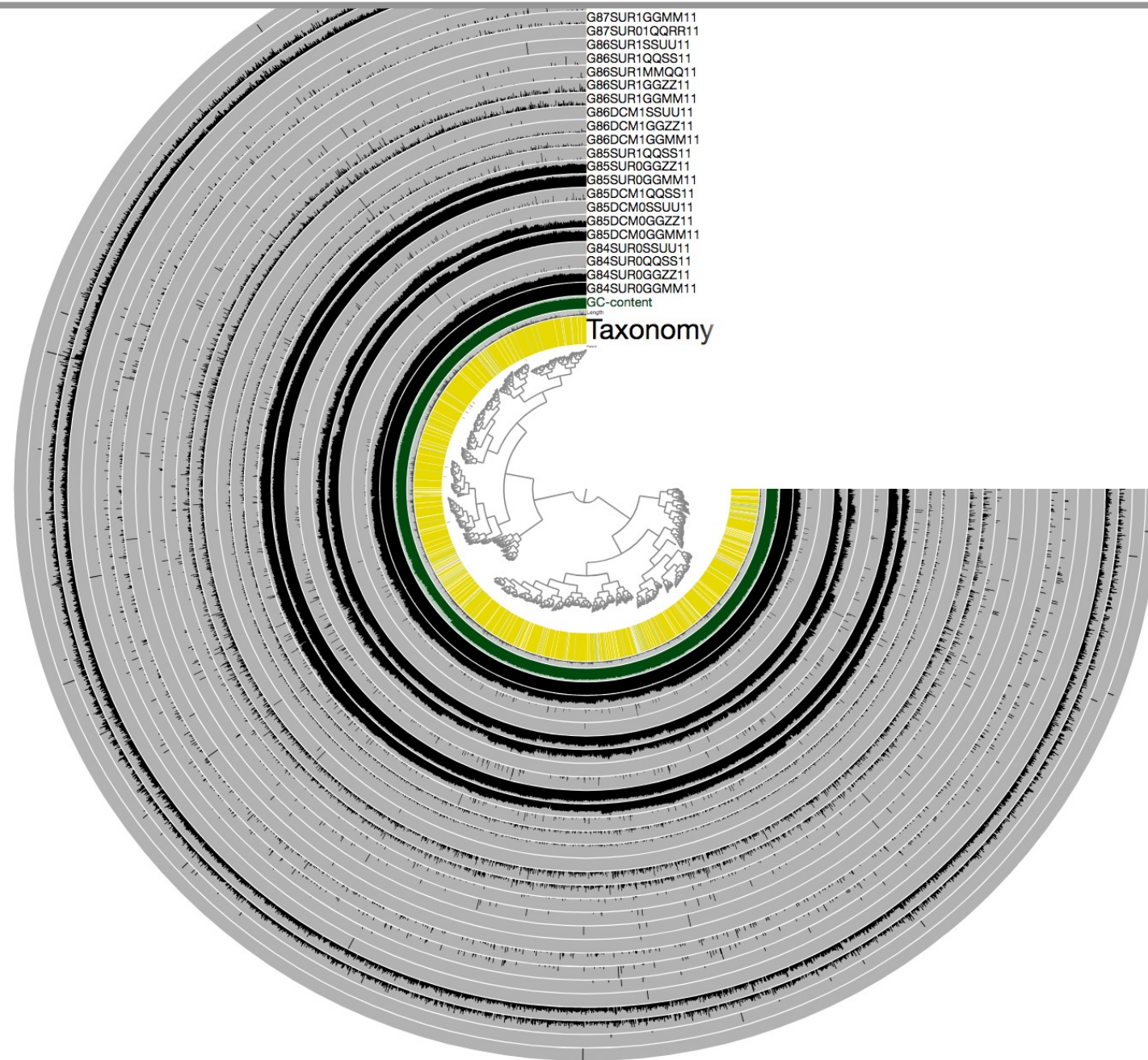
One Tara Oceans metabin
4,111 contigs - 19 Mbp

Visualization & Curation

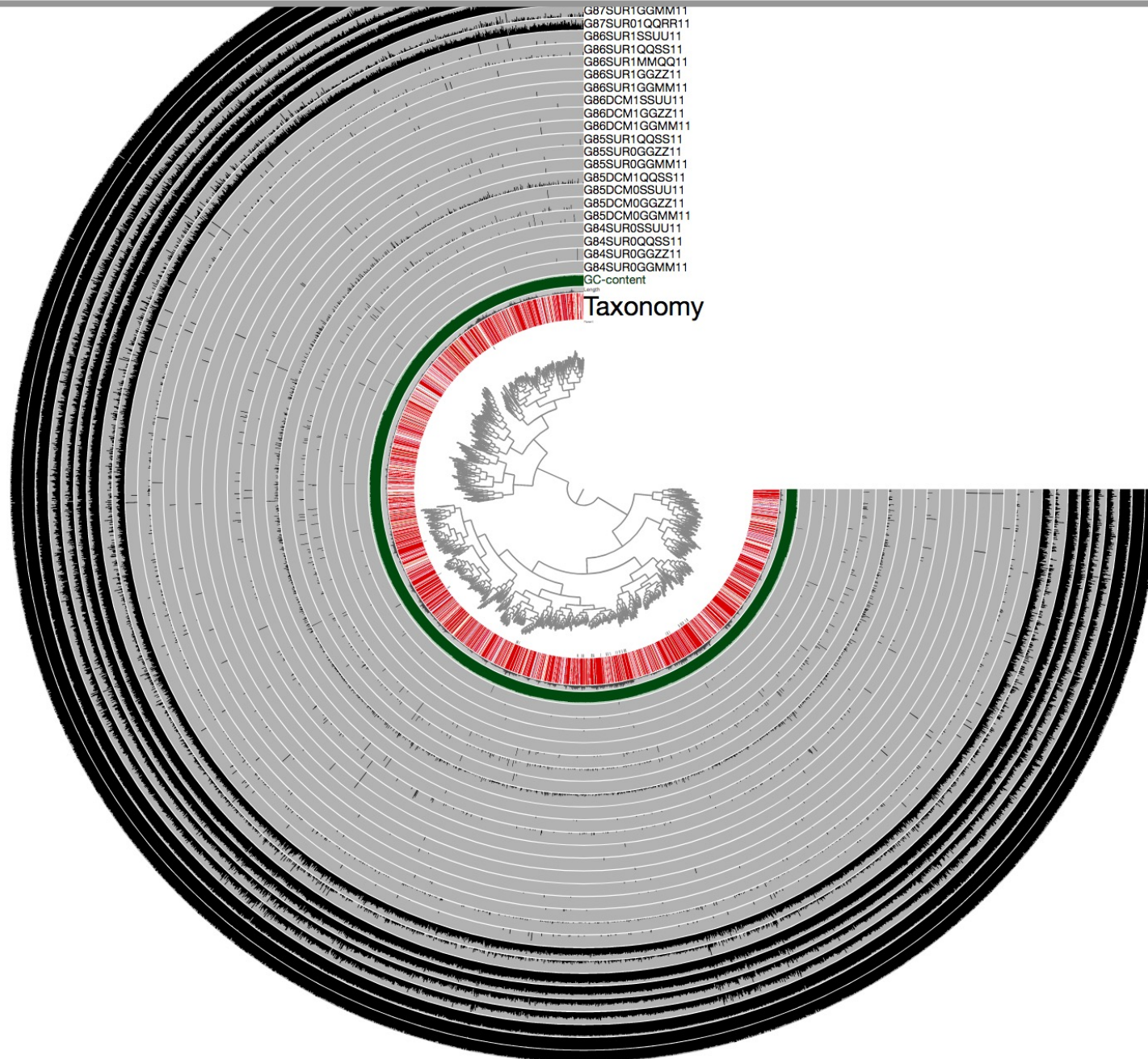


Asgard marine population genome
1.46 Mbp - completion of 82%

A curated eukaryotic MAG (example 01)



A curated eukaryotic MAG (example 02)



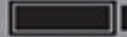


Total of 2.550 metabins

A long journey of manual binning & curation ...

Metabin
#1




100% 

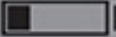


75% 



50% 



25% 



0% 

Metabin
#2.550



Total of **2.550 metabins**

~2,000 bacterial & archaeal MAGs
(all >70% complete)

✓ **40 heterotrophic bacterial diazotrophs:**
more abundant compared to the cyanobacterial
diazotrophs ! (ISMEj - 2021)



Total of **2.550 metabins**

~**2,000 bacterial & archaeal MAGs**
(all >70% complete)

- ✓ **40 heterotrophic bacterial diazotrophs:**
more abundant compared to the cyanobacterial diazotrophs ! (ISMEj - 2021)
- ✓ **2 non-diazotrophic *Trichodesmium* species:**
not all *Trichodesmium* colonies & filaments can fix nitrogen ! (PNAS - 2021)



Total of 2.550 metabins

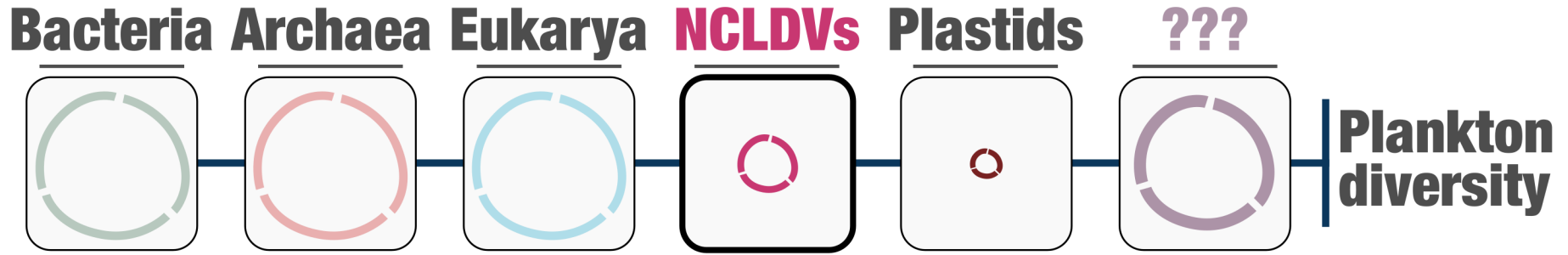
~2,000 bacterial & archaeal MAGs

~700 eukaryotic MAGs (all > 10 Mbp)

✓ 40 heterotrophic bacterial diazotrophs:
more abundant compared to the cyanobacterial
diazotrophs ! (ISMEj - 2021)

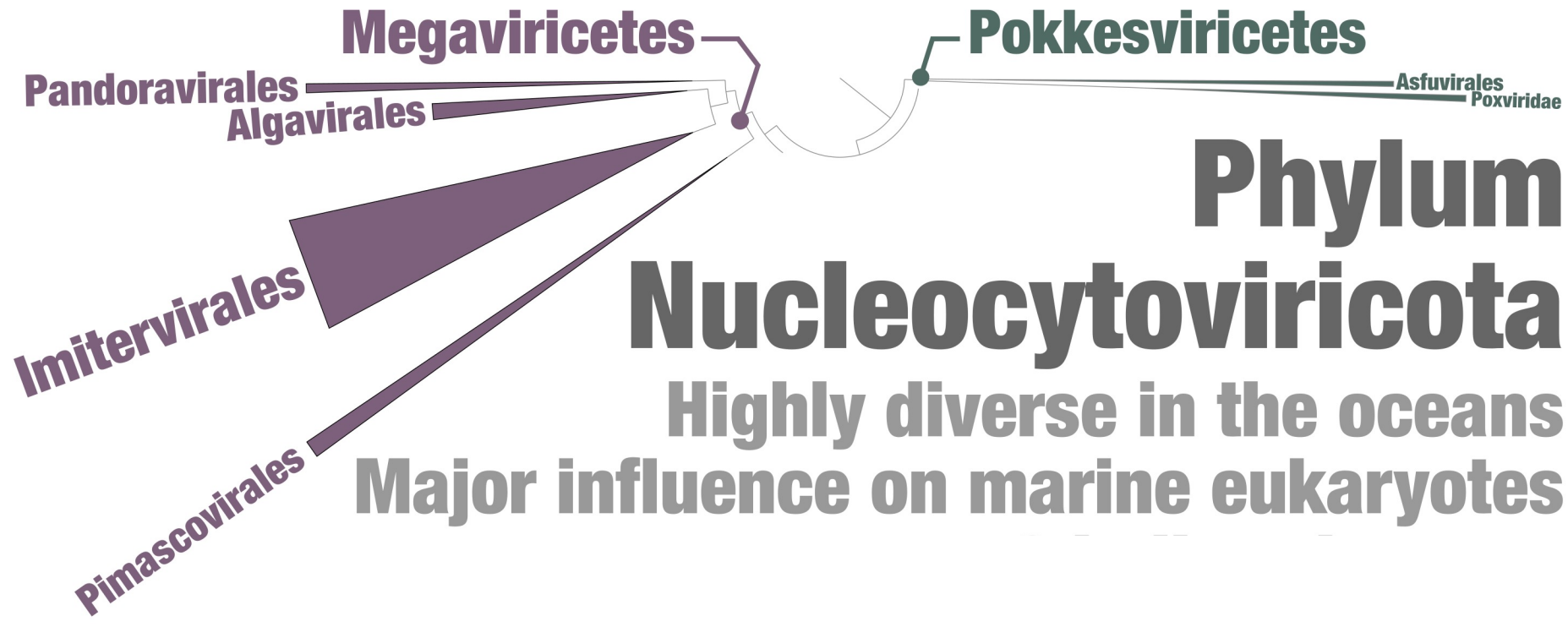
✓ 2 non-diazotrophic *Trichodesmium* species:
not all *Trichodesmium* colonies & filaments can
fix nitrogen ! (PNAS - 2021)

✓ Recovery of eukaryotic MAGs en masse:
functional convergence of distantly related
eukaryotic lineages ! (Cell Genomics - 2022)



**Single copy core gene
collections for the
three domains
of life**

The genomics of giant viruses & their close relatives



Copyright © 2017

FANTASTIC MAGs

AND WHERE TO FIND THEM



TARA Oceans

Copyright © W. G. B. R. Y

FANTASTIC MAGs

AND WHERE TO FIND THEM



TARA Oceans

- **Large co-assemblies**
- **Mapping reads**
- **Data processing (with anvi'o)**
- **Classic binning**

FANTASTIC MAGs

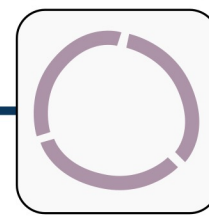
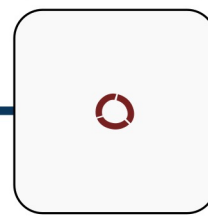
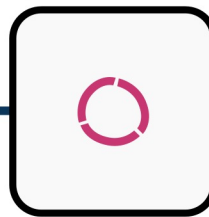
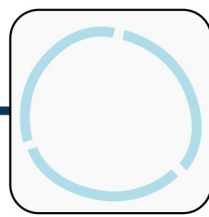
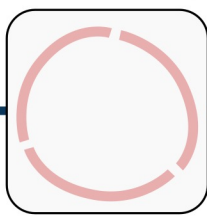
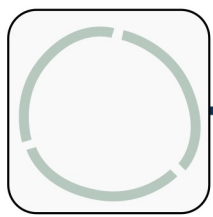
AND WHERE TO FIND THEM



TARA Oceans

- Large co-assemblies
- Mapping reads
- Data processing (with anvi'o)
- Classic binning
- Targeted binning

Bacteria **Archaea** **Eukarya** **NCLDVs** **Plastids** **???**

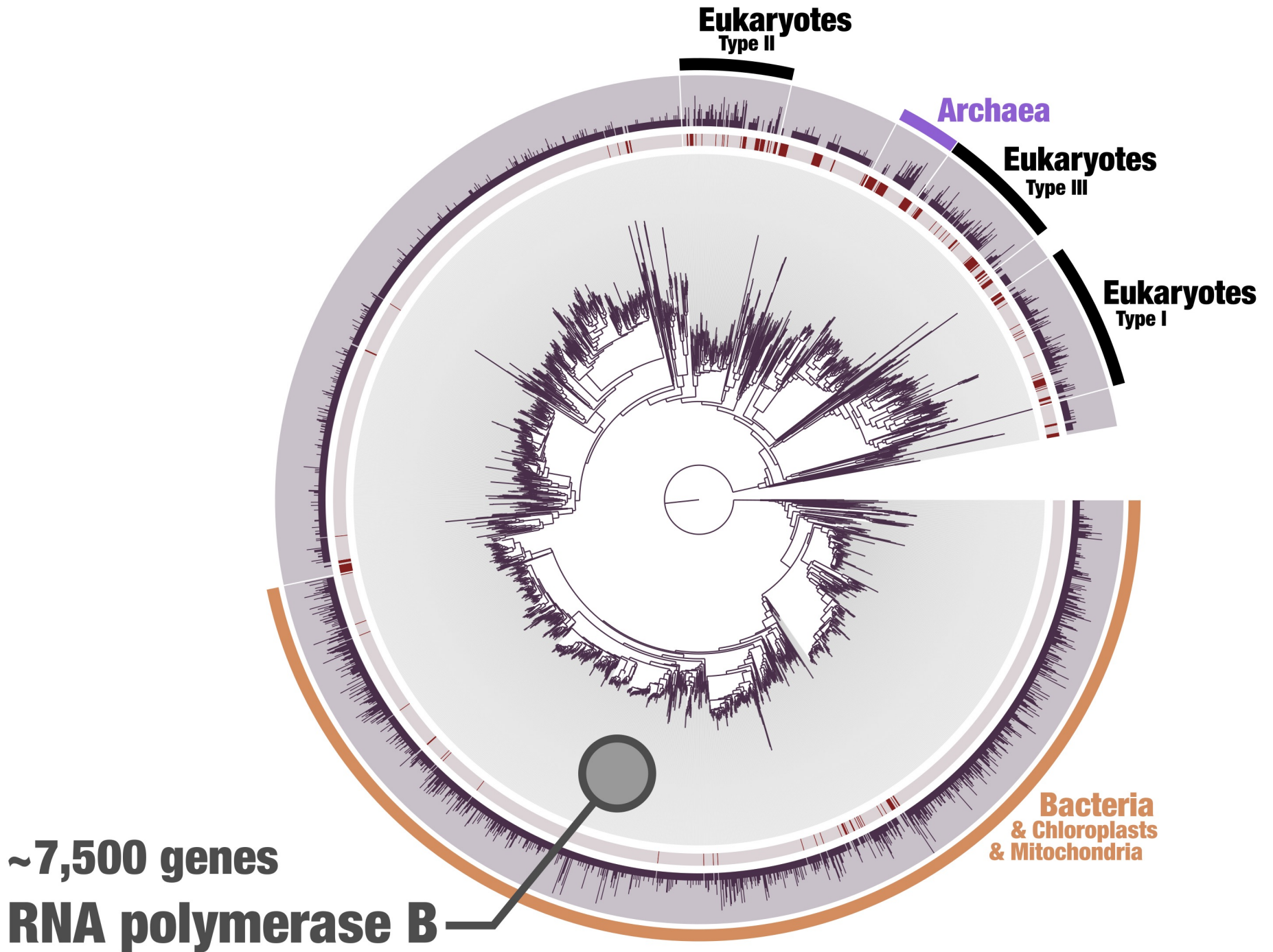


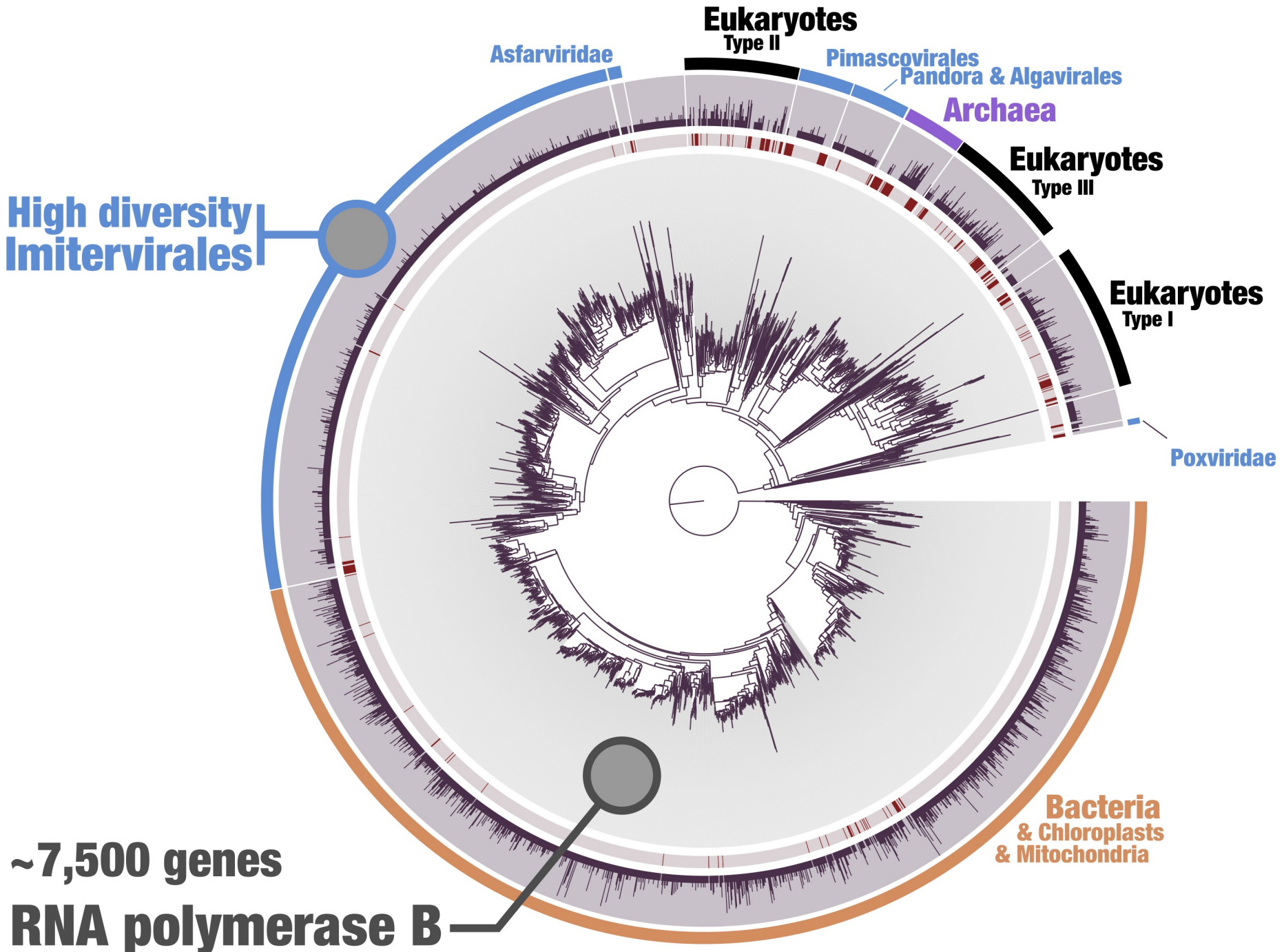
**Plankton
diversity**



The remarkable RNA polymerase genes

**A compass for
targeted binning**





Deep-branching
with no references

Eukaryotes
Type II

NEW 01
Asfarviridae

Pimascovirales
Pandora & Algavirales

Archaea

Eukaryotes
Type III

NEW 02 & 03

Eukaryotes
Type I

NEW 04

NEW 05 &
06 & 07
Poxviridae

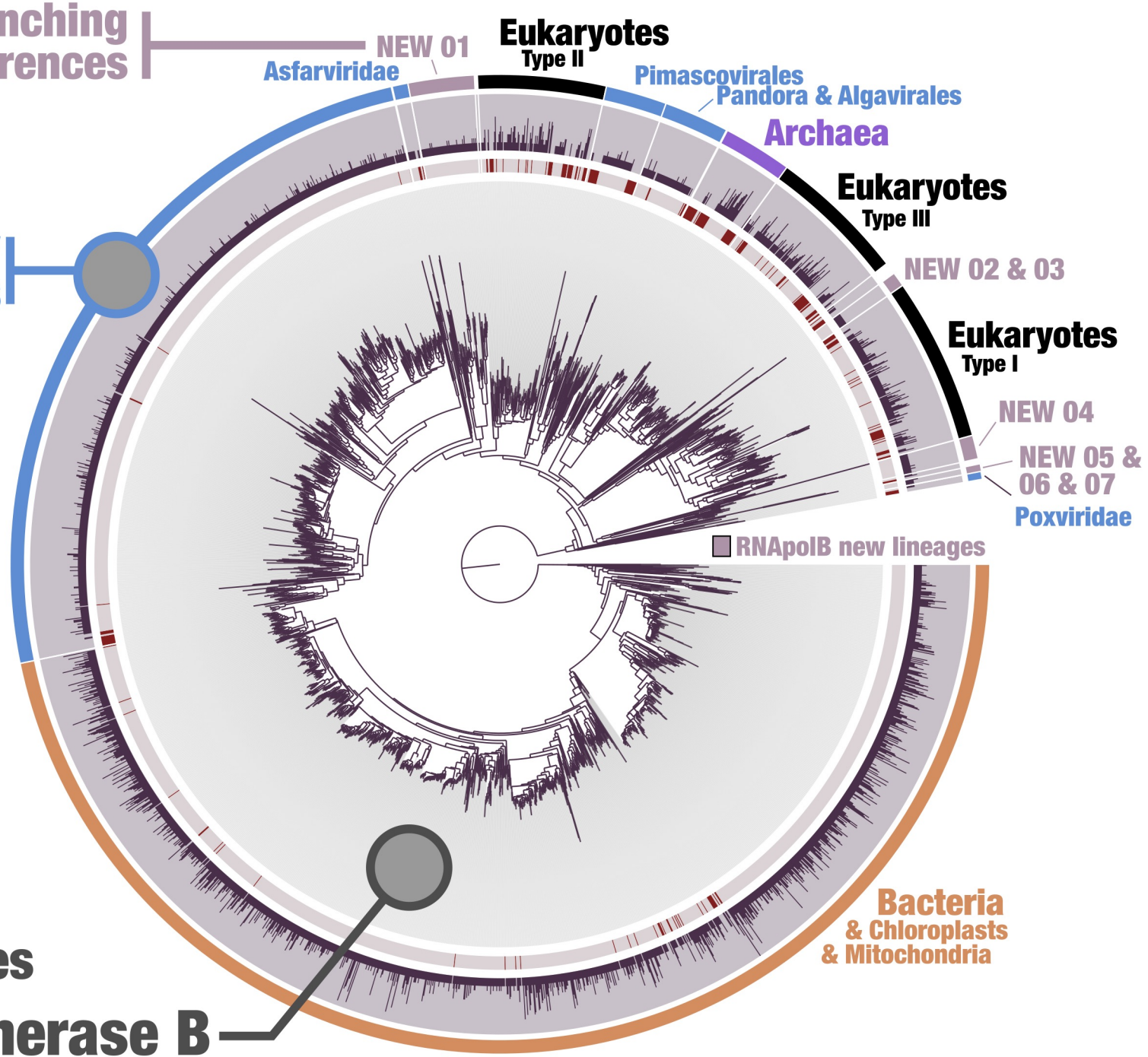
■ RNAPolB new lineages

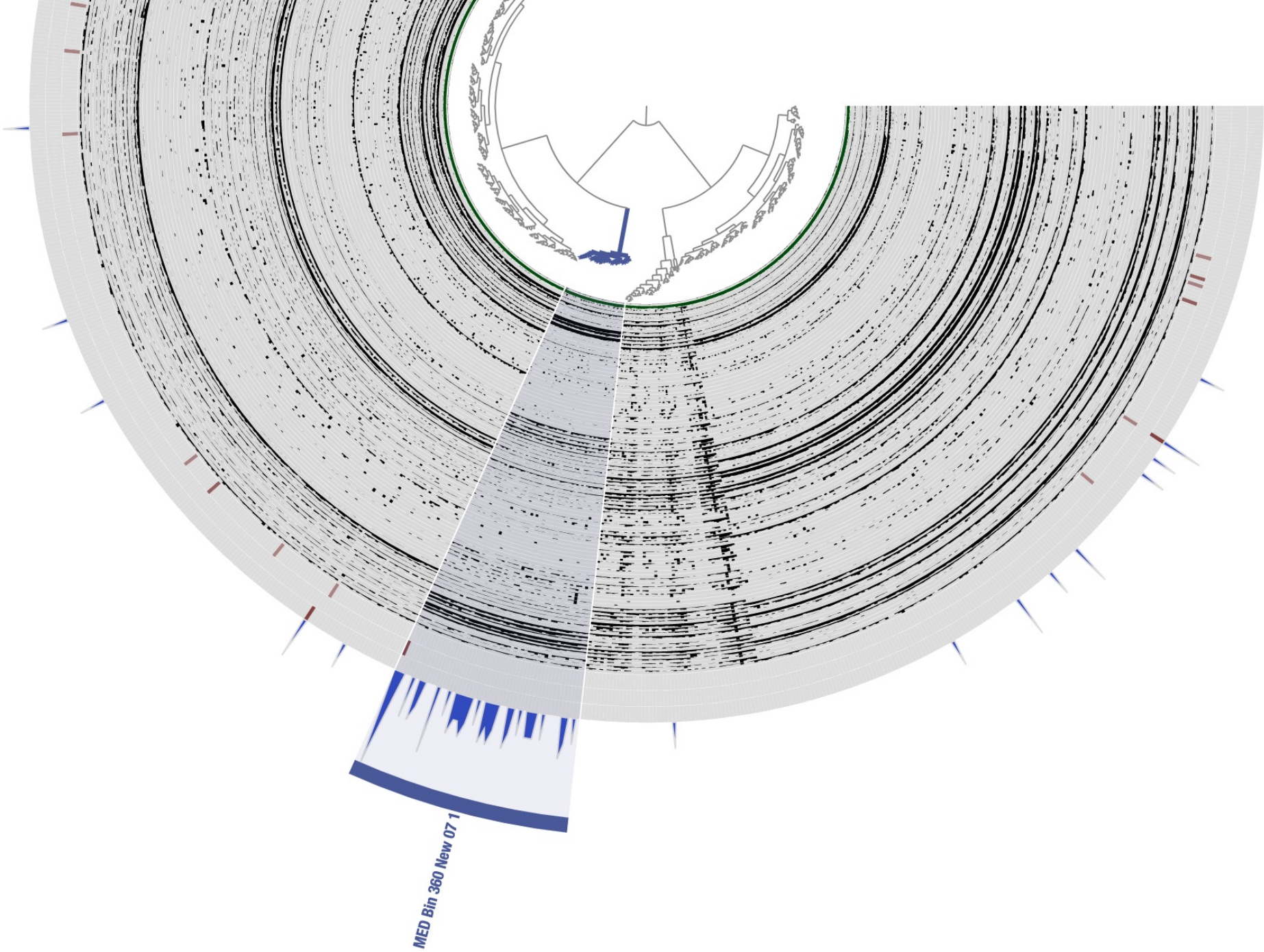
Bacteria
& Chloroplasts
& Mitochondria

High diversity
Imitervirales

~7,500 genes

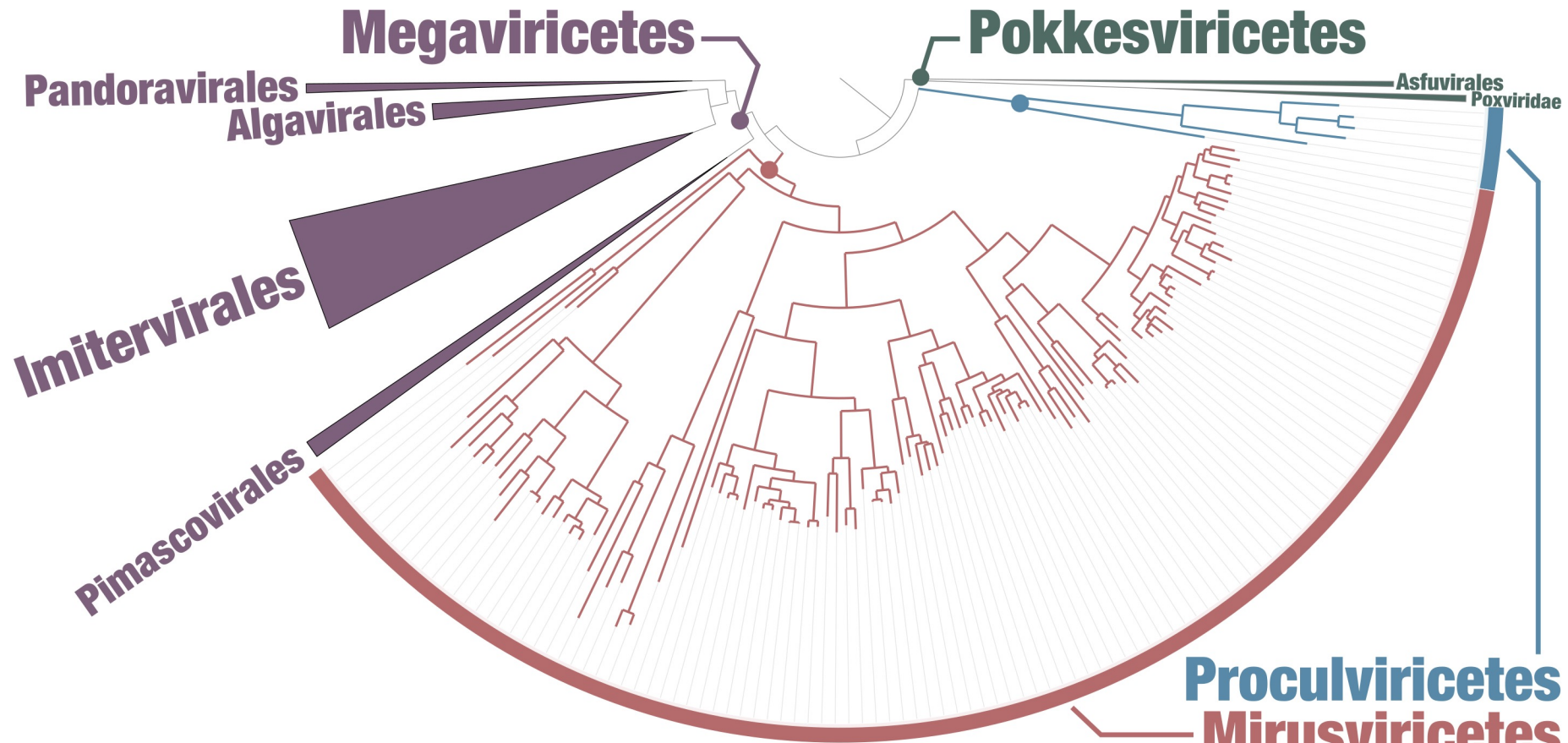
RNA polymerase B



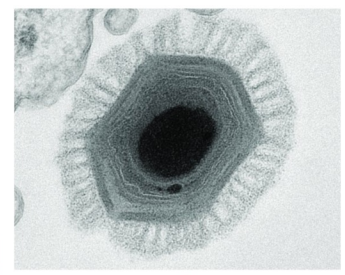


MED Bin 360 New 07 1

The genomics of giant viruses & their close relatives



	RNAp0IA	RNAp0IB	DNAp0IB	TFIIS	MCP	pATPase	VLTF3	Primase
Megaviricetes	○	○	●	●	●	●	●	●
Pokkesviricetes	●	●	●	●	●	●	●	●
Proculviricetes	●	●	●	●	●	●	●	●
Mirusviricetes	●	●	○	●	○	○	○	○



Virion morphogenesis

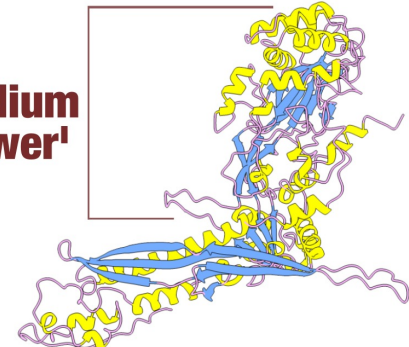
-- Major capsid protein --

No 'tower'



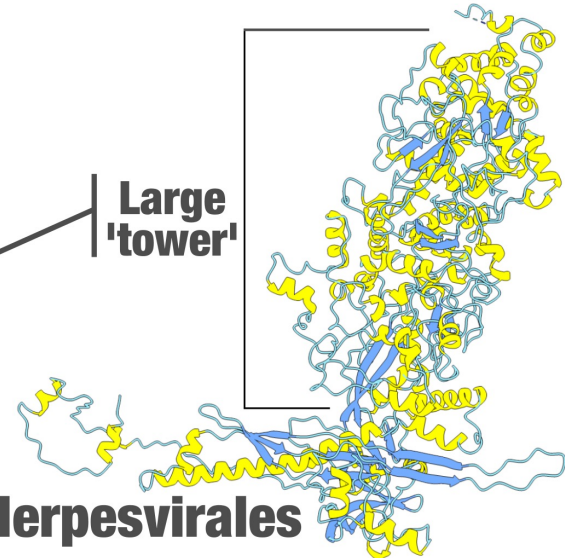
Caudoviricetes

**Medium
'tower'**



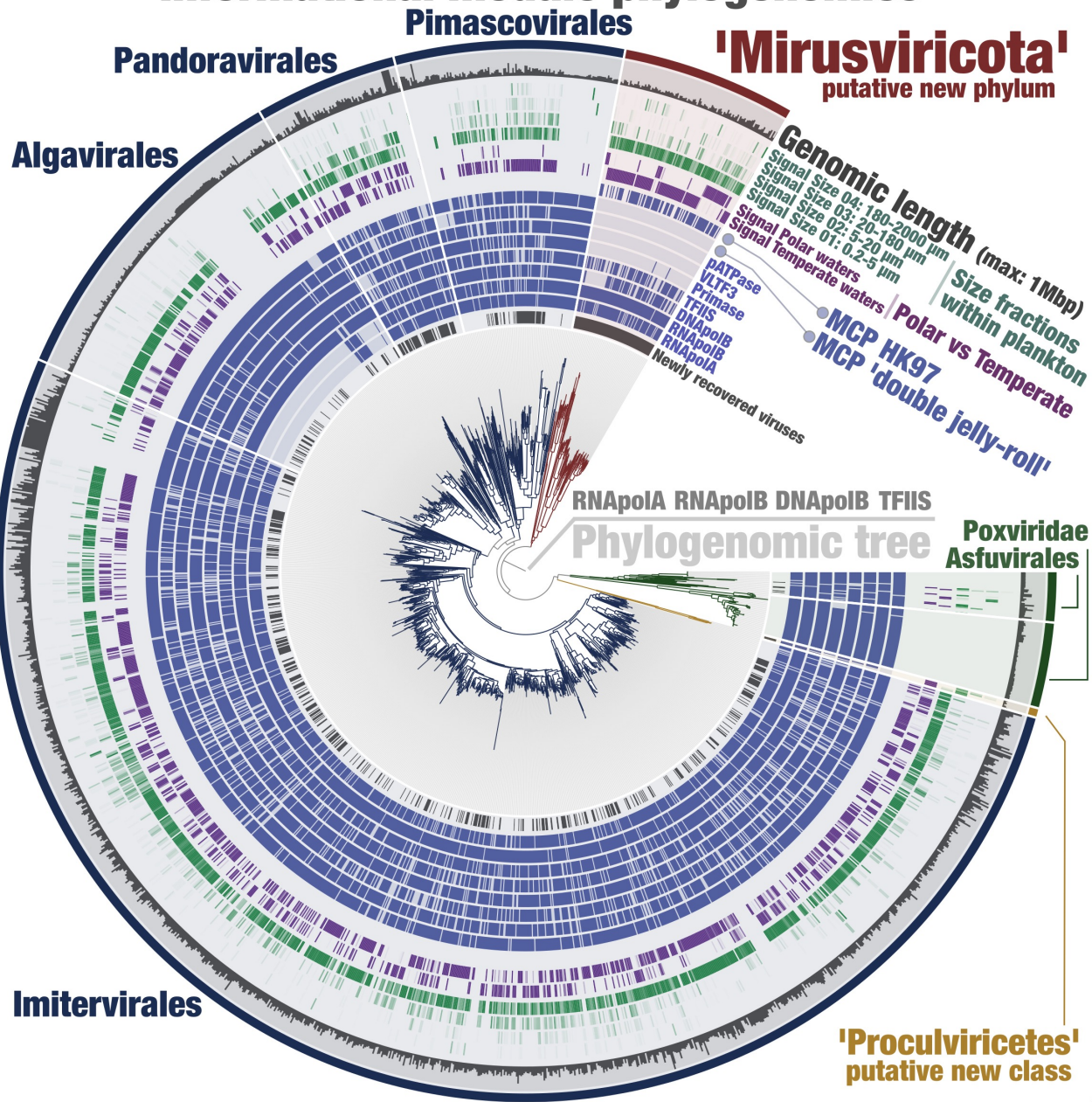
'Mirusviricota' MCP HK97

**Large
'tower'**

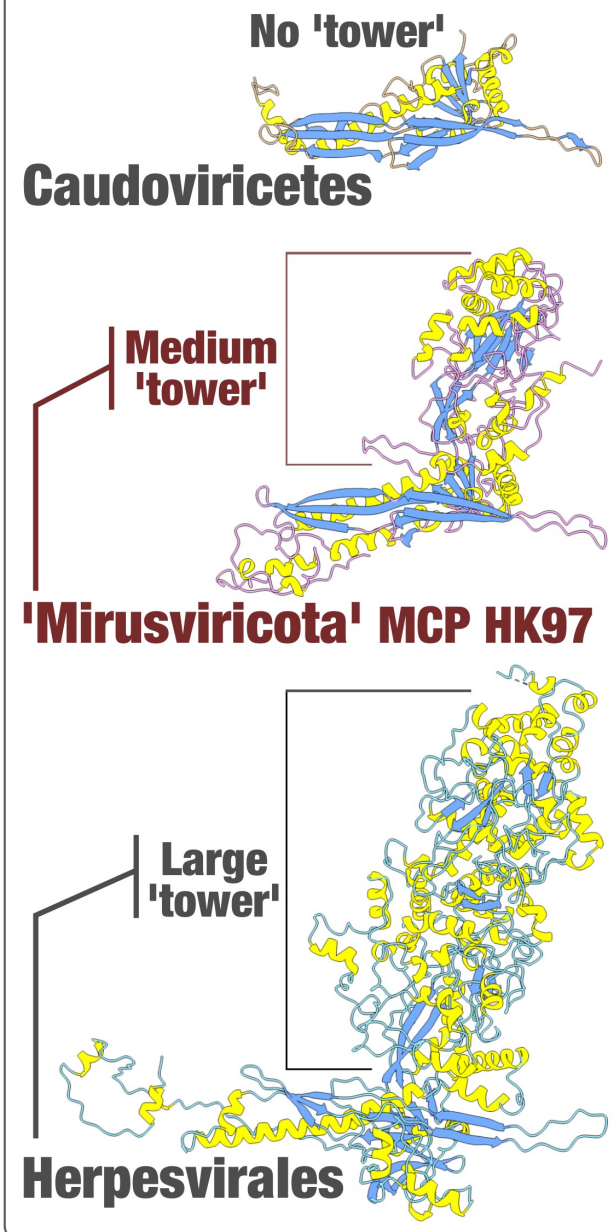


Herpesvirales

-- Informational module phylogenomics --



-- Major capsid protein --



Plankton-infecting relatives of herpesviruses clarify the evolutionary trajectory of giant viruses

Morgan Gaïa^{1,2*}, Lingjie Meng^{3*}, Eric Pelletier^{1,2}, Patrick Forterre^{4,5}, Chiara Vanni⁶, Antonio Fernandez-Guerra⁷, Olivier Jaillon^{1,2}, Patrick Wincker^{1,2}, Hiroyuki Ogata³, Mart Krupovic⁸, and Tom O. Delmont^{1,2#}

¹ Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France.

² Research Federation for the study of Global Ocean systems ecology and evolution, FR2022/Tara G0see, Paris, France.

³ Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, 611-0011, Japan

⁴ Institut de Biologie Intégrative de la Cellule (I2BC), CNRS, Université Paris-Saclay, 91198 Gif sur Yvette, France

⁵ Institut Pasteur, Département de Microbiologie, 25 rue du Docteur Roux, 75017, Paris, France

⁶ MARUM center for marine environmental sciences, University of Bremen, Germany

⁷ Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

⁸ Institut Pasteur, Université Paris Cité, CNRS UMR6047, Archaeal Virology Unit, Paris, France

* Co-first authors

Corresponding author

FANTASTIC MAGs

AND WHERE TO FIND THEM



TARA Oceans

- Large co-assemblies
- Mapping reads
- Data processing (with anvi'o)
- Classic binning
- Targeted binning

Mirus

fantastic MAGs
connecting
2 viral realms

FANTASTIC MAGs

AND WHERE TO FIND THEM



TARA Oceans

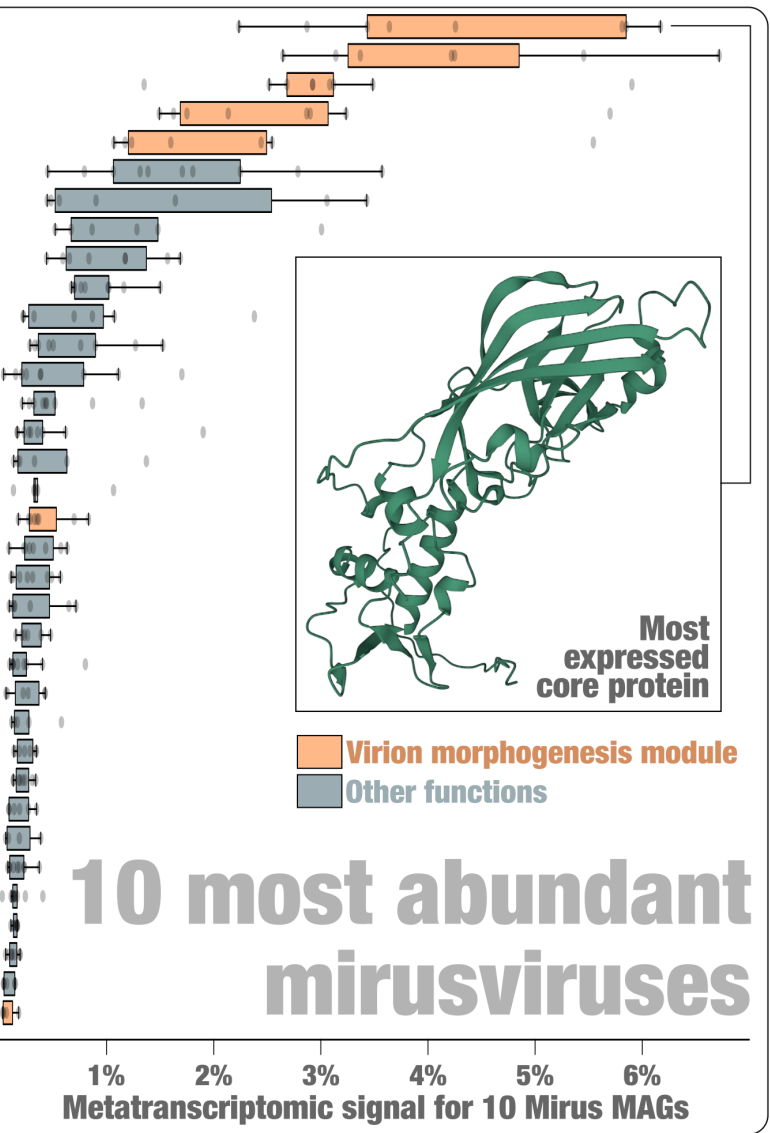
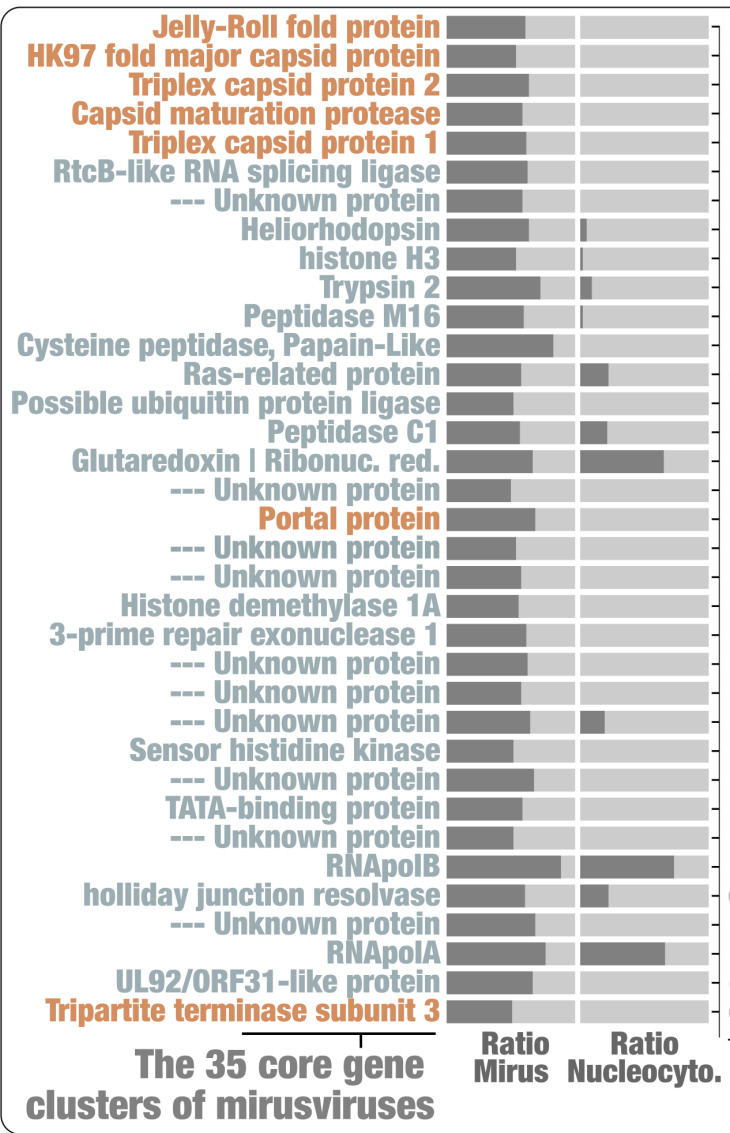
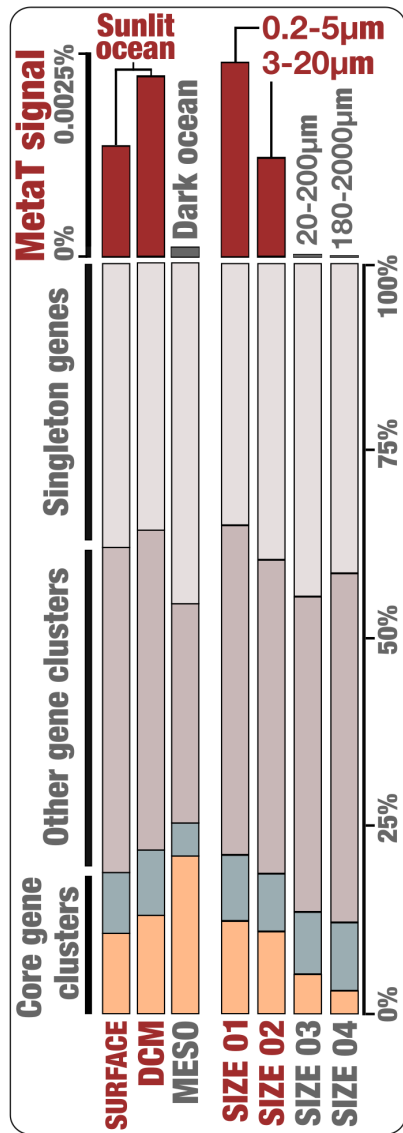
- **Large co-assemblies**
- **Mapping reads**
- **Data processing (with anvi'o)**
- **Classic binning**
- **Targeted binning**
- **An opportunity for many microbiologists to contribute:**

**Decentralized
genome-resolved
metagenomics**

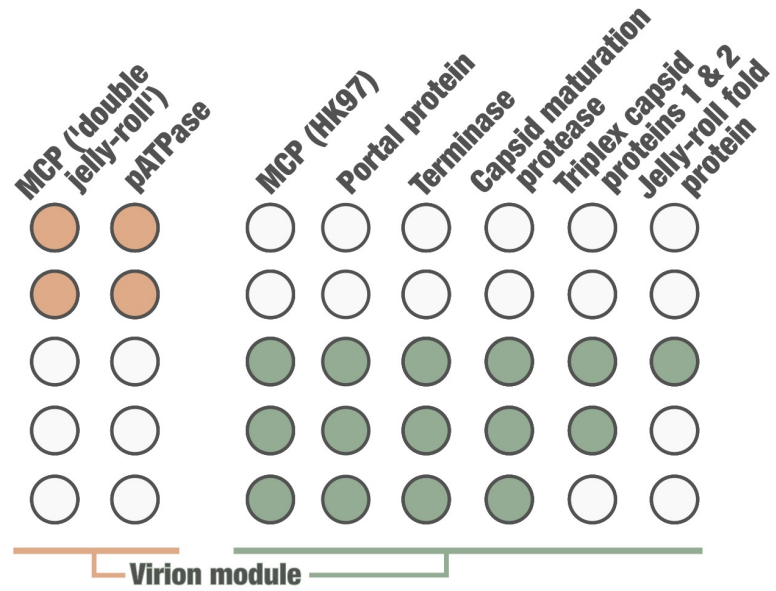
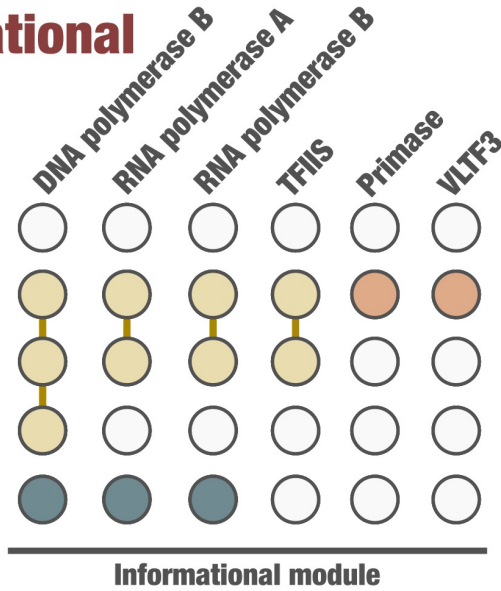
Plankton genomics with Tara Oceans



Together!



A Occurrence of informational and virion modules



● Eukaryotic informational module connecting two realms

● Sporadic and highly divergent from the eukaryotic informational module

A Occurrence of informational and virion modules

