

Metagenomics is for protists too

Mini-workshop on metagenomics
Protistology Nordics Meeting (May 2nd, 2024)

Jennah Dharamshi (Uppsala University)

Why now? ■

Preview

Metagenome assembled genomes are for eukaryotes too

Ramon Massana^{1,*} and David López-Escardó¹

¹Institut de Ciències del Mar (ICM-CSIC), Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Catalonia, Spain

*Correspondence: ramonm@icm.csic.es

<https://doi.org/10.1016/j.xgen.2022.100130>

Environmental community sequencing is suitable for producing metagenome assembled genomes (MAGs) of prokaryotes, but there is the perception that it cannot work for eukaryotes. In this issue of *Cell Genomics*,

“We envision a future when eukaryotic MAGs will be better extracted from metagenomic datasets, perhaps by importing new sequencing approaches or better bioinformatic tools, providing exhaustive genome characterizations of dominant species in nature, and saturating new phylum discovery as is occurring with prokaryotes.”

Who are we?



Jennah Dharamshi

Microbial evolution and symbiosis

Metagenomics, Genome-Resolved Metagenomics, Metabarcoding, Comparative Genomics, Phylogenomics, Transcriptomics



Mercè Montoliu Nerín

Evolutionary genomics and symbiosis

Single-cell genomics, Genomics, Comparative Genomics, Phylogenomics



Mahwash Jamy

Protist diversity and evolution

Metabarcoding, Phylogenomics, Transcriptomics, Metagenomics



Nina Pohl

Protist diversity and plastid evolution

Metagenomics, Single-cell genomics, Phylogenomics



Plan for the mini-workshop

Overview: The current state of metagenomics for protists

Tutorial Part 1: Searching metagenomes

Tutorial Part 2: Binning metagenomes

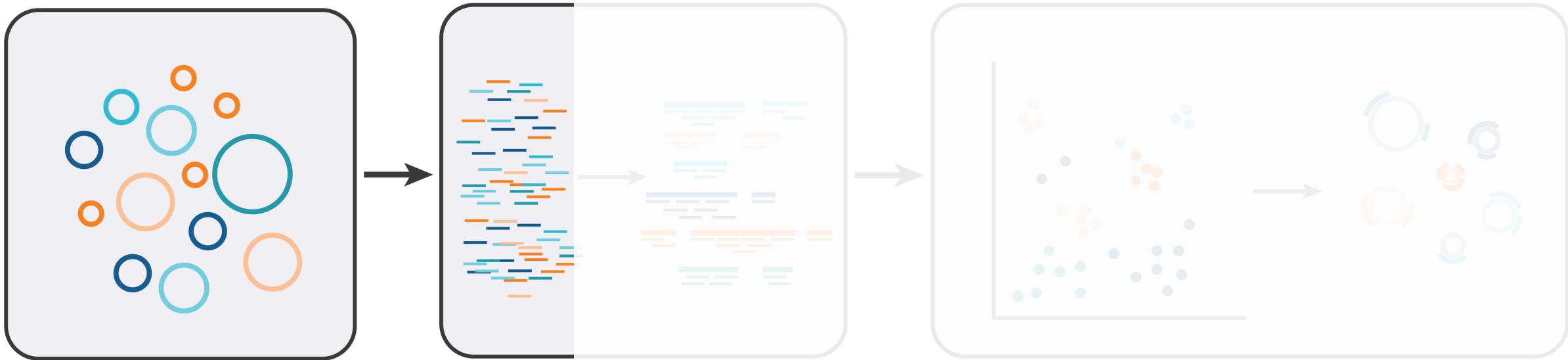
Overview: The current state of metagenomics for protists

What do we mean by metagenomics?



Shotgun metagenomics: “the untargeted ('shotgun') sequencing of all ('meta-') microbial genomes 'genomics' present in a sample” – Quince et al., 2017 (*Nature Biotechnology*, doi: 10.1038/nbt.3935)

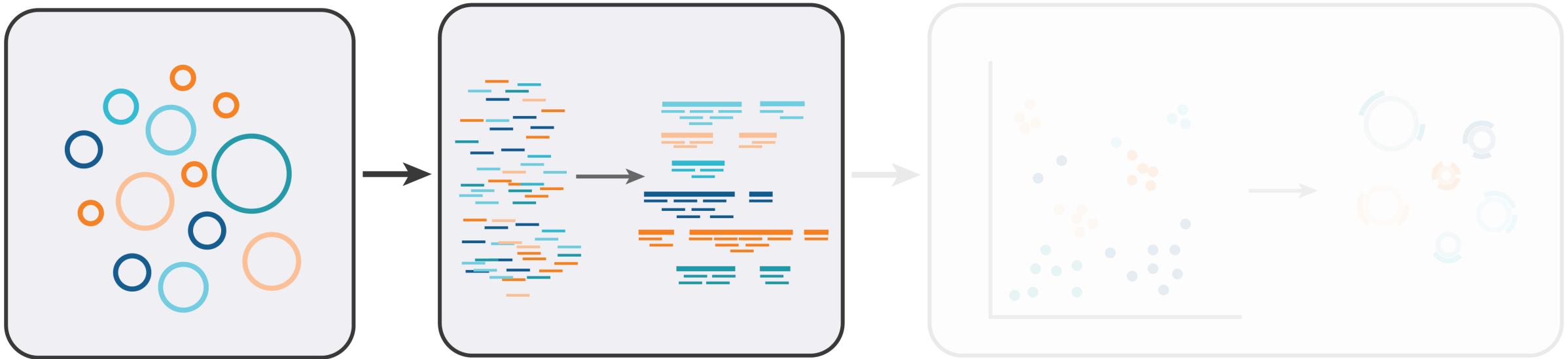
What do we mean by metagenomics?



Shotgun metagenomics: “the untargeted ('shotgun') sequencing of all ('meta-') microbial genomes 'genomics' present in a sample” – Quince et al., 2017 (*Nature Biotechnology*, <https://doi.org/10.1038/nbt.3935>)

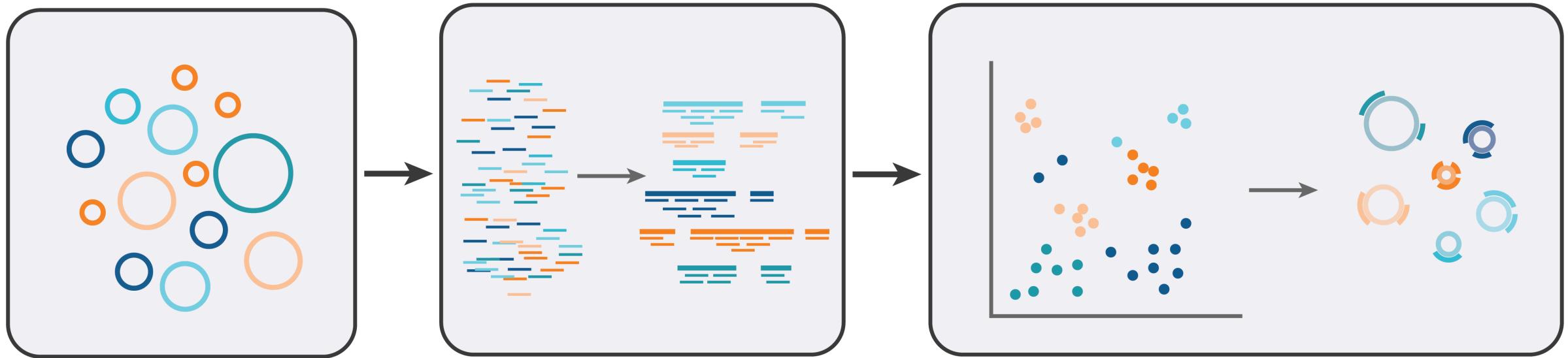
mOTUs or mTAGs: Clustering and classification of marker gene metagenomic reads into operational taxonomic units (OTUs) or sequence variants at different taxonomic levels

What do we mean by metagenomics?



Metagenome assembly: Reconstructed longer contiguous sequences (contigs) obtained from shorter fragments that originate from short or long shotgun metagenomic reads

What do we mean by metagenomics?



Genome resolved metagenomics: The approach of obtaining metagenome-assembled genomes from metagenomic data. For example, by clustering contigs into **bins** (i.e., **binning**) using tetranucleotide frequency, GC content, taxonomic hits, reference genomes, coverage, and/or differential coverage.

Metagenome assembled genome (MAG): Genomes reconstructed from metagenome assemblies.

Metagenomics vs. metabarcoding: pros and cons

Pros

- Primer-free so no primer bias
- Often no amplification so fewer errors and chimeric sequences introduced
- Captures the whole community
- Can also retrieve functional information and look at the diversity of various genes
- Can get the entire rRNA gene operon
- Potential to retrieve MAGs and perform phylogenomic and comparative genomic analyses

Cons

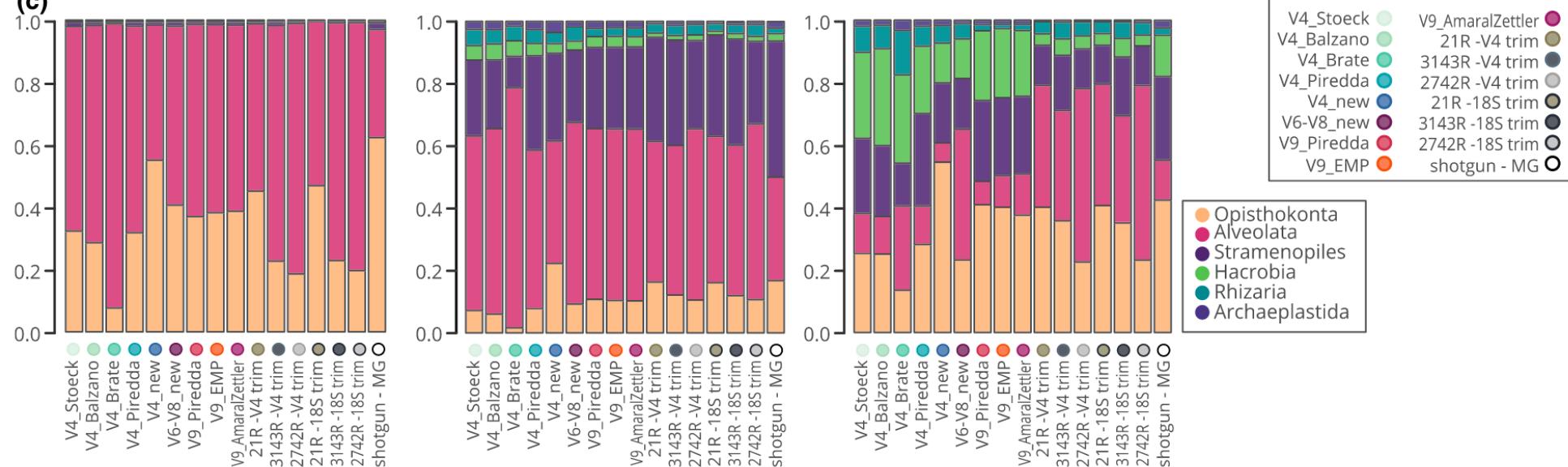
- Doesn't capture rare taxa or taxa with low relative abundance
- Currently fewer metagenomes available than metabarcoding datasets
- More data generated per sample and higher sequencing costs
- Genome-resolved metagenomics often performs poorly for eukaryotes (large and difficult to assemble genomes)

Metagenomics vs. metabarcoding: diversity

(a)



(c)



Sequencing shotgun metagenomes

Short reads

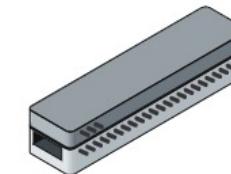


Illumina

Long reads



Pacific Biosciences



Oxford Nanopore

mOTUs or mTAGs: tools



taxonomic meta-omics profiling
using universal marker genes

<https://github.com/motu-tool/mOTUs>

Ruscheweyh et al., *Microbiome*, 2022
(<https://doi.org/10.1186/s40168-022-01410-z>)

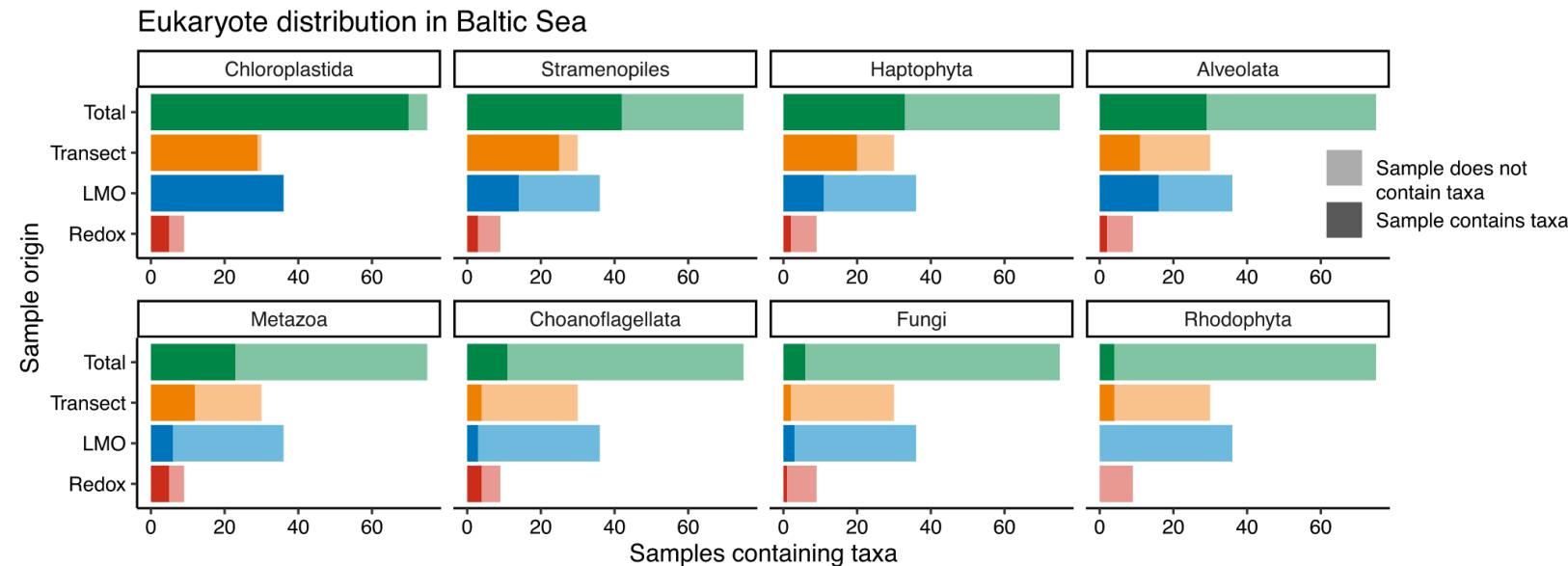


mTAGs: taxonomic profiling using degenerate consensus reference sequences of ribosomal RNA genes

<https://github.com/SushiLab/mTAGs>

Salazar et al., *Bioinformatics*, 2022
(<https://doi.org/10.1093/bioinformatics/btab465>)

EukDetect



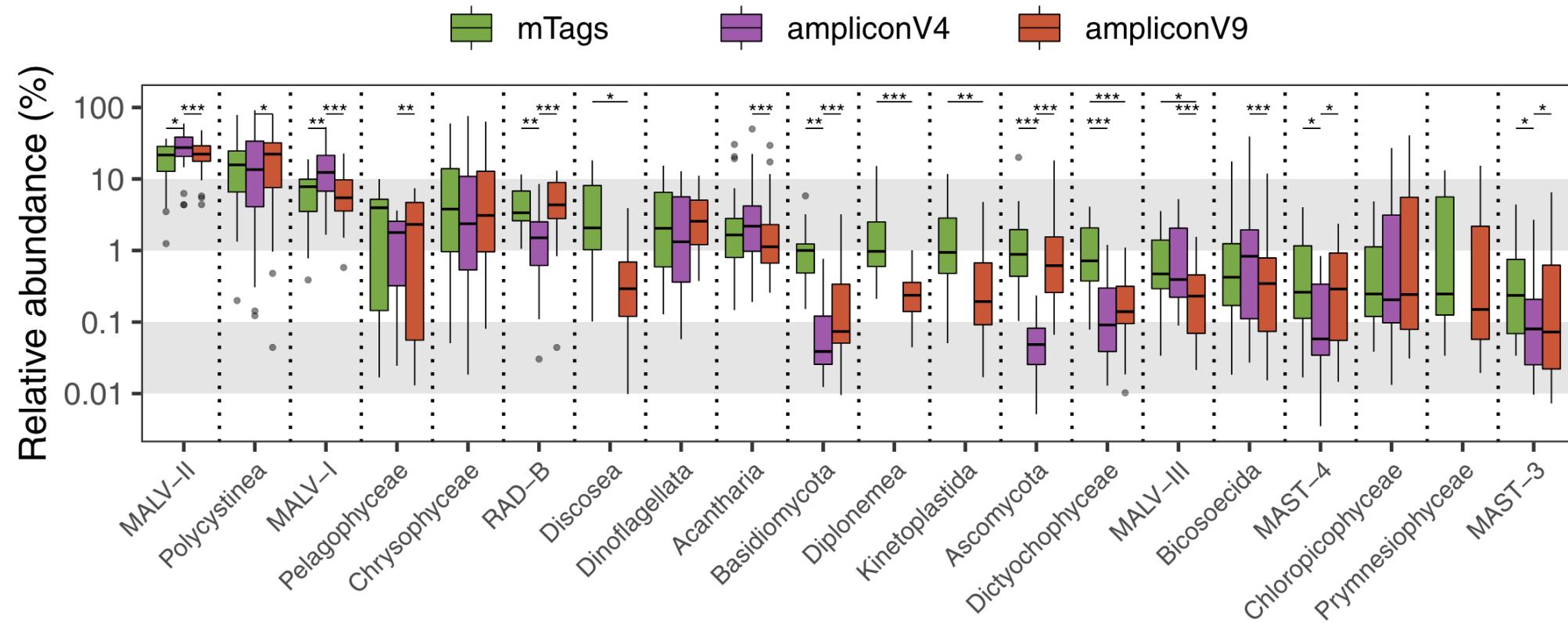
<https://github.com/allind/EukDetect>

For identifying eukaryotes in shotgun metagenomic data

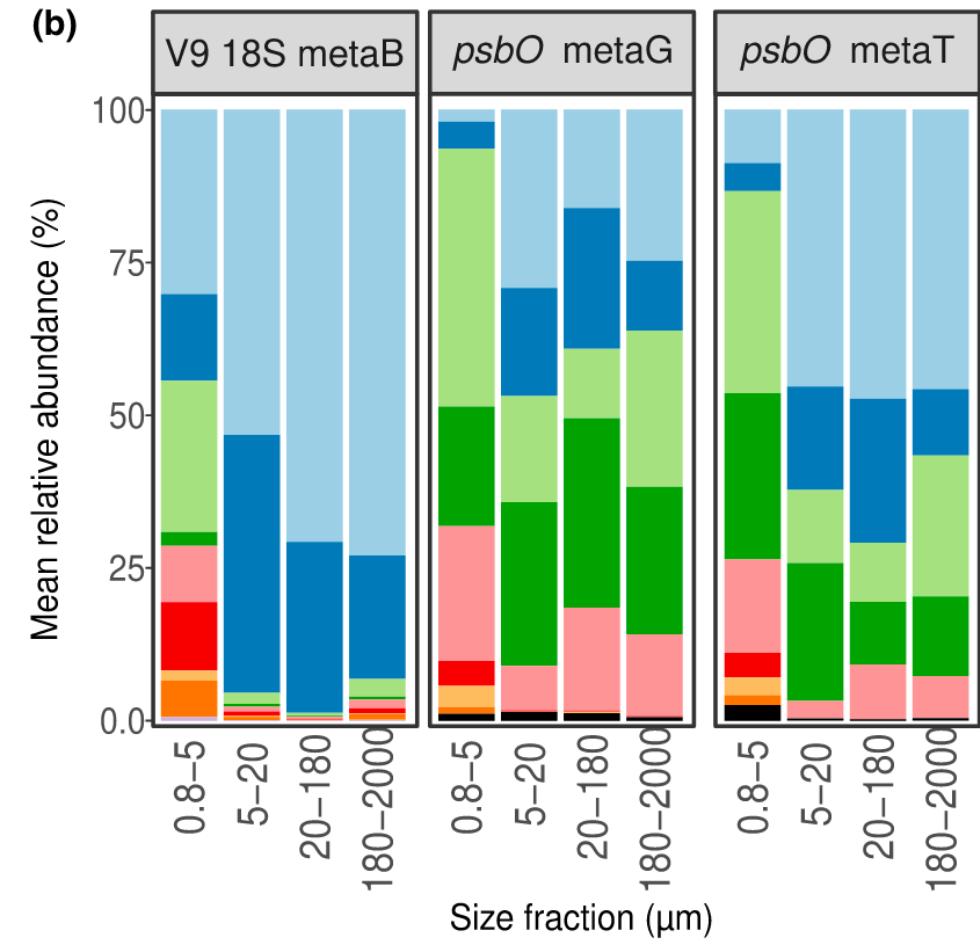
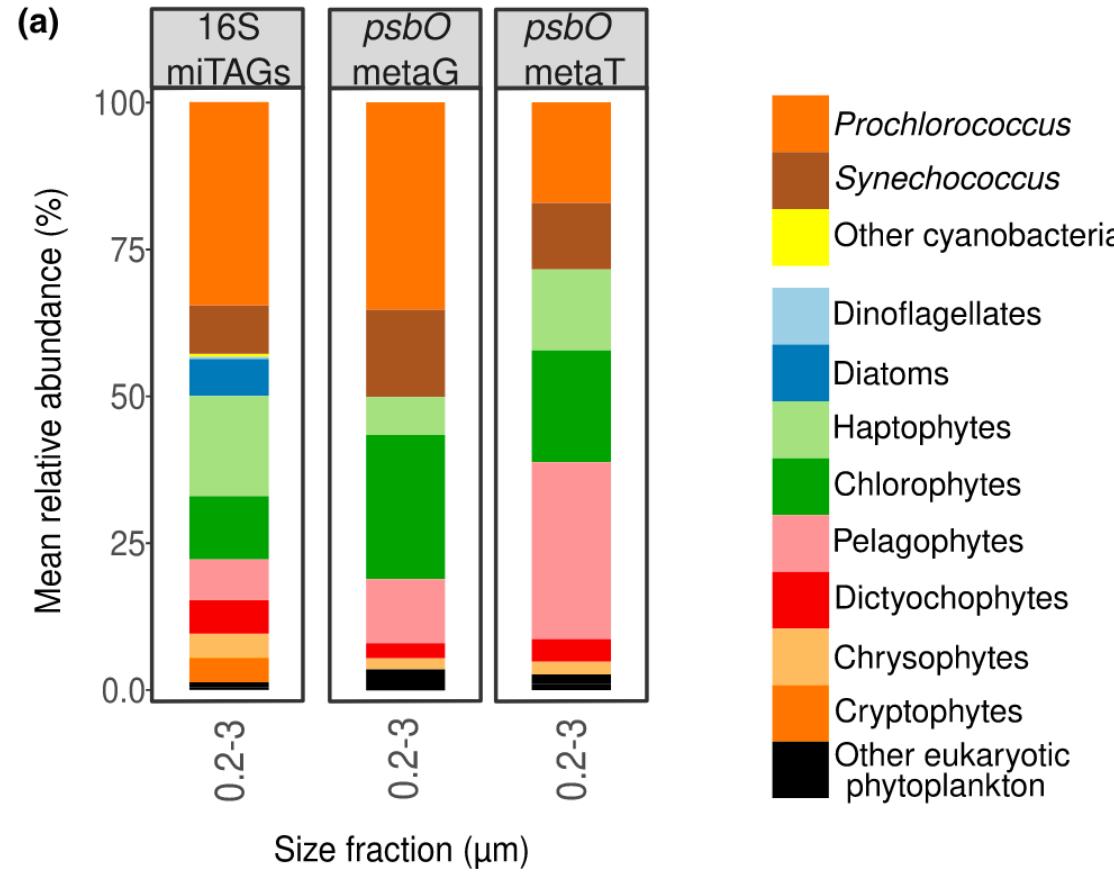
Aligns metagenomic reads to a database of 521,824 universal marker genes from 241 conserved gene families

Includes broad taxonomic coverage of microbial eukaryotes and performs well on low-abundance and closely related species

mTAGs or mOTUs: Diversity and relative abundance

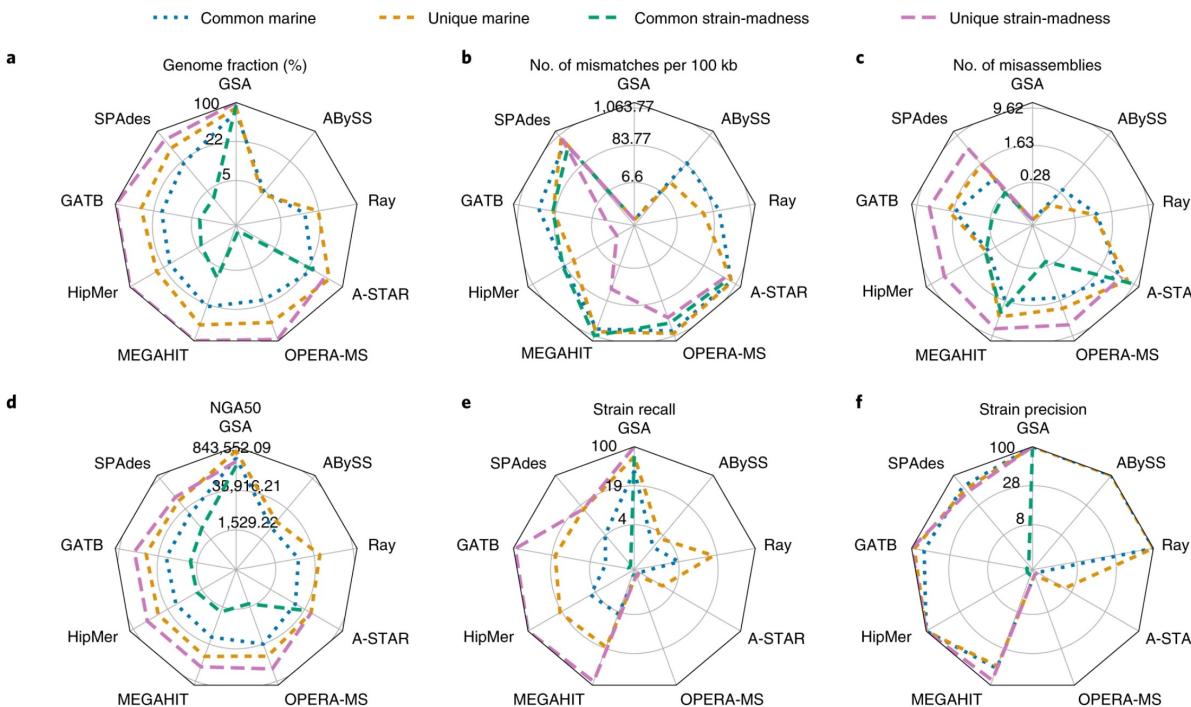


mTAGs or mOTUs: Diversity and relative abundance



Assembling metagenomes

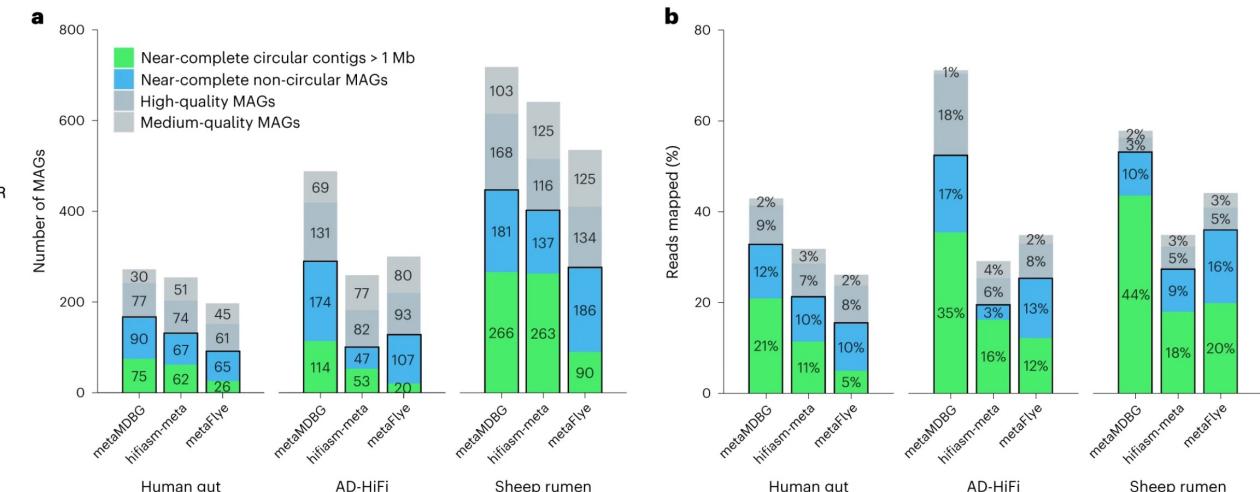
Short read & hybrid assemblers



Meyer et al., *Nature Methods*, 2022
<https://doi.org/10.1038/s41592-022-01431-4>

CAMI - Critical Assessment of Metagenome Interpretation

Long read assemblers (metaMDBG)



Benoit et al., *Nature Biotechnology*, 2024
<https://doi.org/10.1038/s41587-023-01983-6>

metaMDBG – made for high-quality long reads from PacBio

Assembled metagenomes: Uncovering diversity

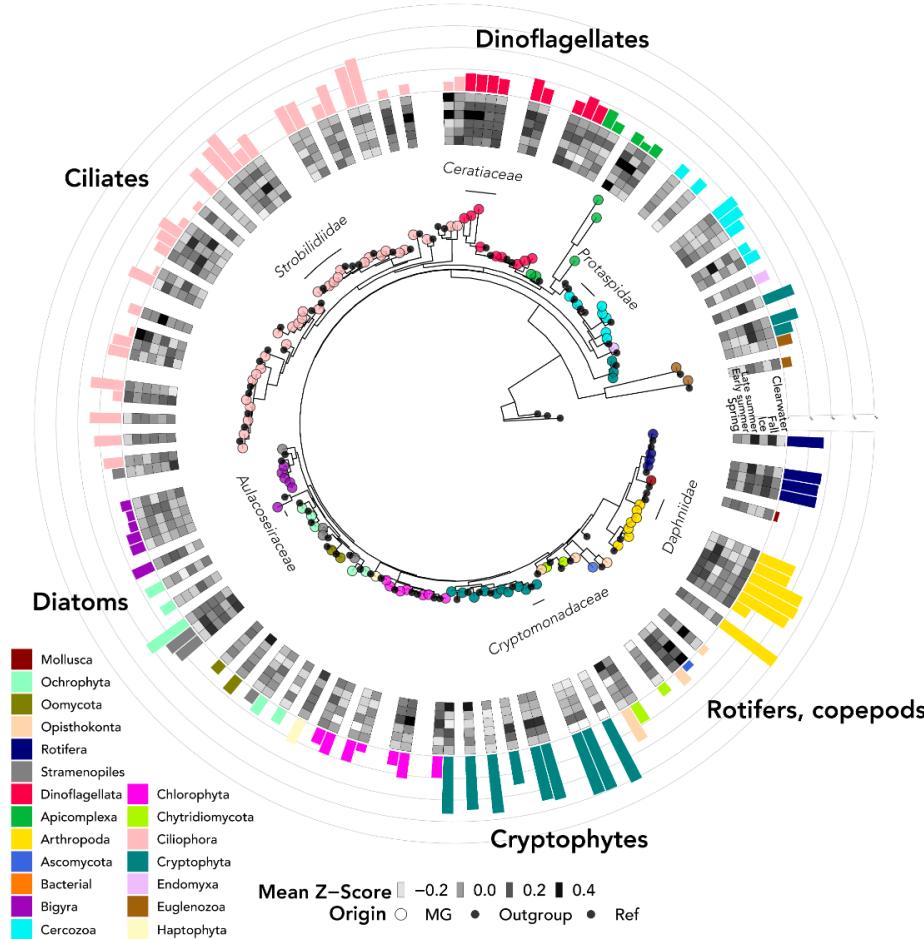
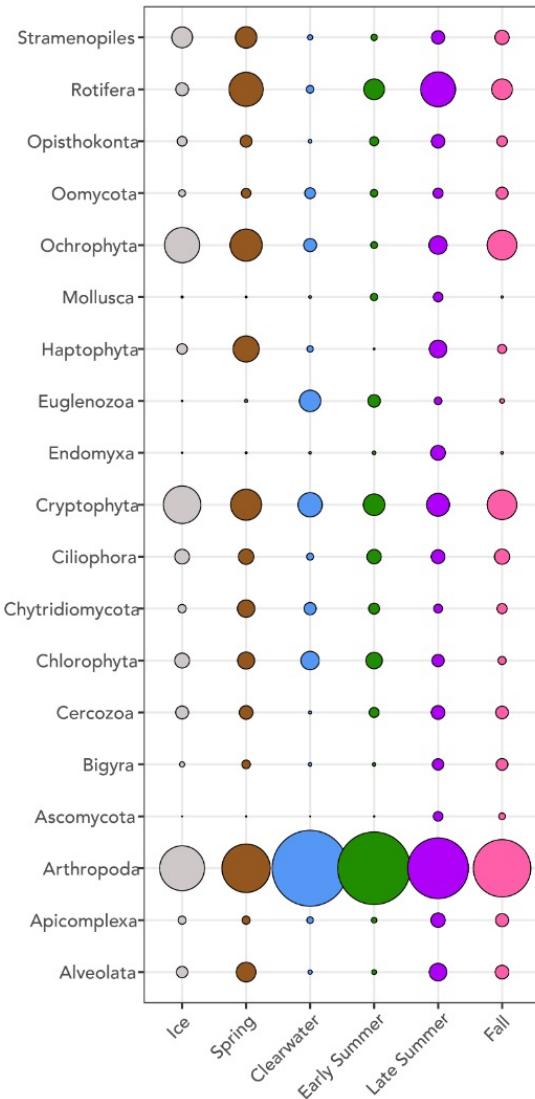
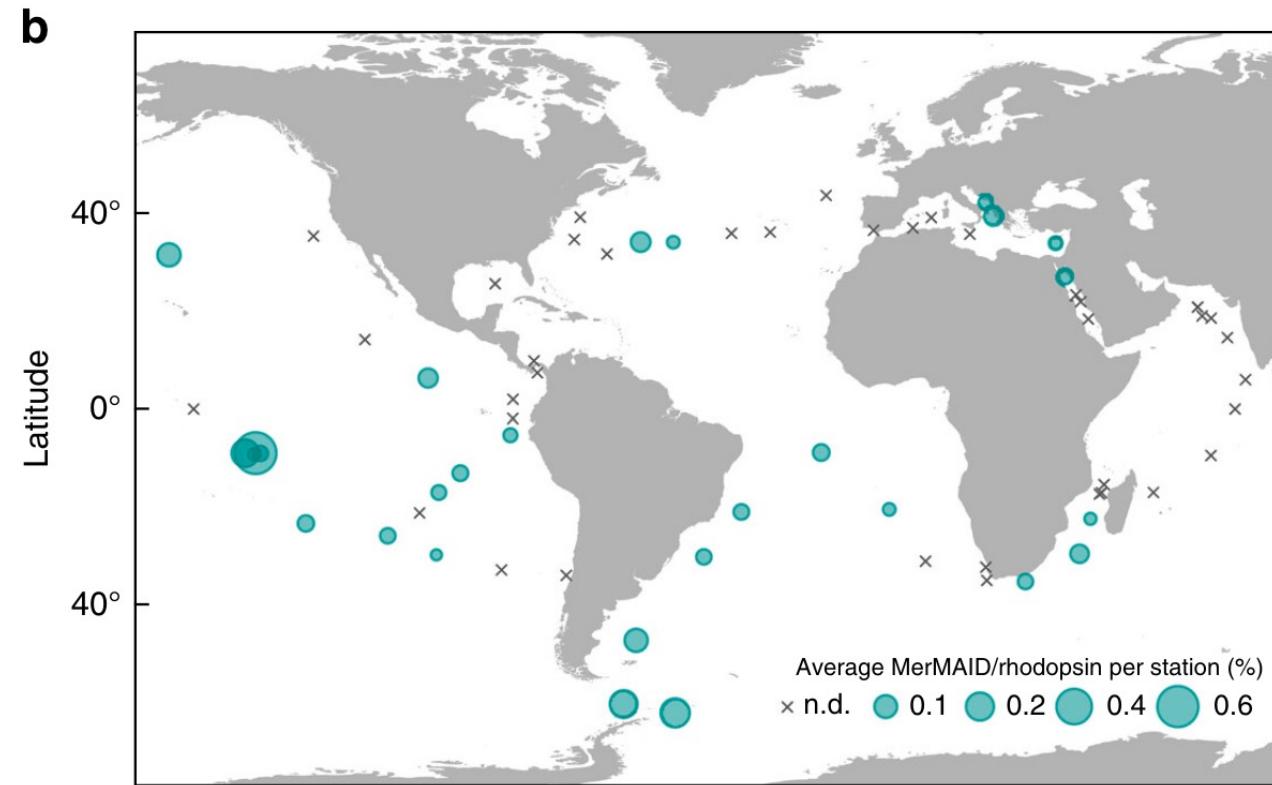
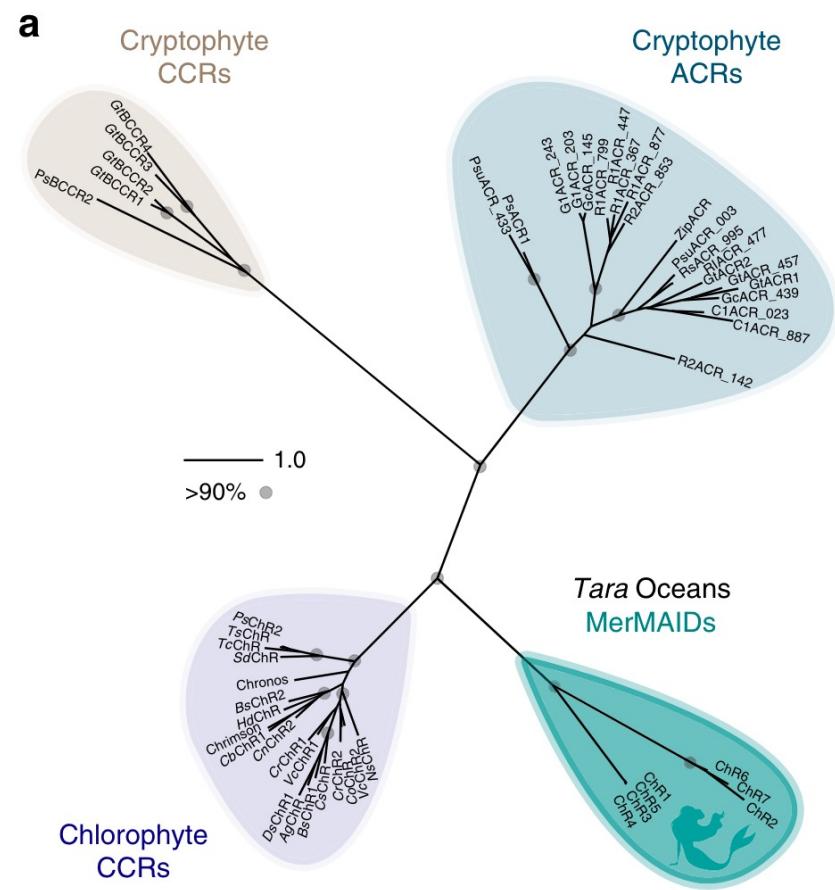


Figure 1: Phylogenetic tree of 18S rRNA gene sequences extracted from Lake Mendota metagenomes over the twenty-year time series. Colored, filled circular points on the tree indicate novel sequences extracted from the Mendota metagenomes, while smaller black circles denote a representative set of previously published sequences. The rings of the heatmap represent the mean Z-score of extracted sequence during each of the seasons of the year, and bars show the number of samples the 18S rRNA gene was found in.

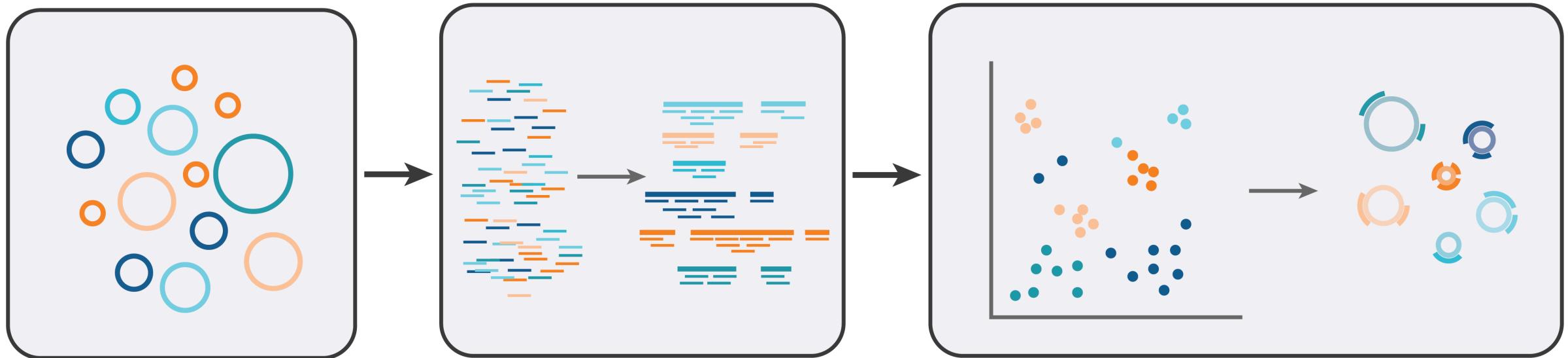


Assembled metagenomes: Gene discovery



Metagenomically identified family of phylogenetically distinct anion-conducting channelrhodopsins (designated MerMAIDs)

Metagenomic binning: How do I retrieve a MAG?



Identifying eukaryotic metagenome contigs: tools

Whokaryote

Whokaryote

Whokaryote uses a random forest classifier that uses gene-structure based features and optionally Tiara (<https://github.com/ibe-uw/tiara>) predictions to predict whether a contig is from a eukaryote or from a prokaryote.

You can use Whokaryote to determine which contigs need eukaryotic gene prediction and which need prokaryotic gene prediction.

<https://github.com/LottePronk/whokaryote>

Pronk and Medema, Microbial Genomics, 2022 (<https://doi.org/10.1099/mgen.0.000823>)

TIARA

Tiara

Deep-learning-based approach for identification of eukaryotic sequences in the metagenomic data powered by PyTorch.

The sequences are classified in two stages:

- In the first stage, the sequences are classified to classes: archaea, bacteria, prokarya, eukarya, organelle and unknown.
- In the second stage, the sequences labeled as organelle in the first stage are classified to either mitochondria, plastid or unknown.

<https://github.com/ibe-uw/tiara>

Karlicki, Antonowicz and Karnkowska, Bioinformatics, 2022
(<https://doi.org/10.1093/bioinformatics/btab672>)

EukREP

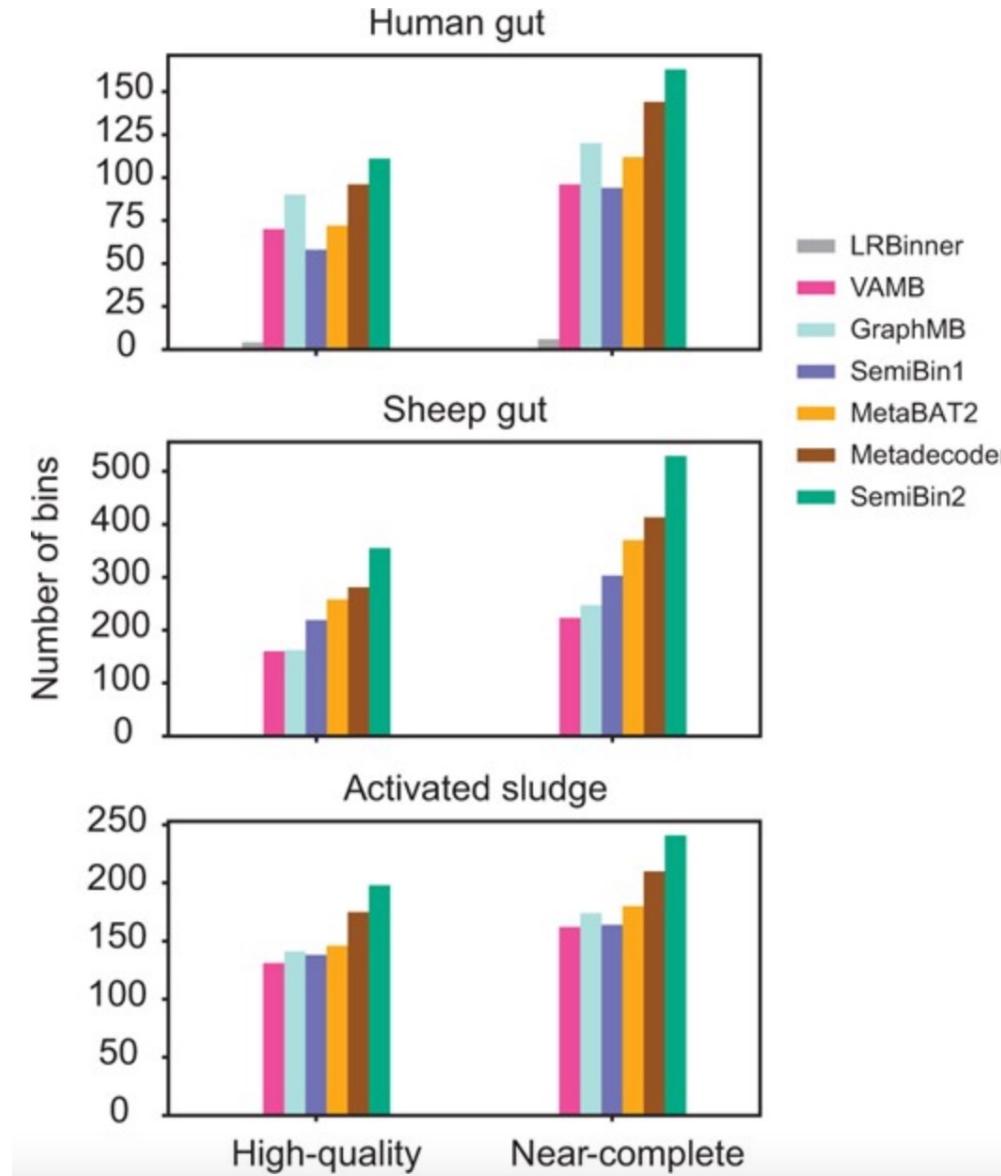
EukRep

Classification of Eukaryotic and Prokaryotic sequences from metagenomic datasets

<https://github.com/patrickwest/EukRep>

West et al., Genome Research, 2018
(<https://doi.org/10.1101/gr.228429.117>)

MAG binning tools: comparison



ACR (Additional Clustering Refiner)

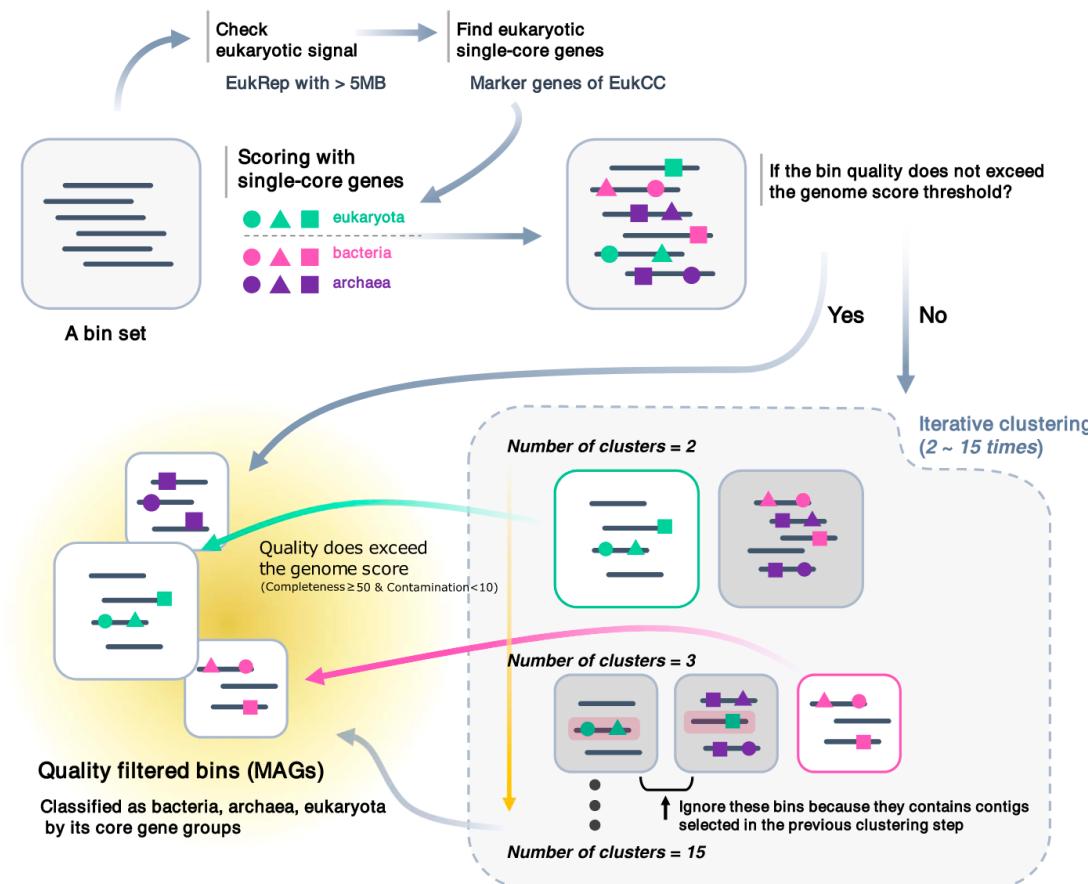
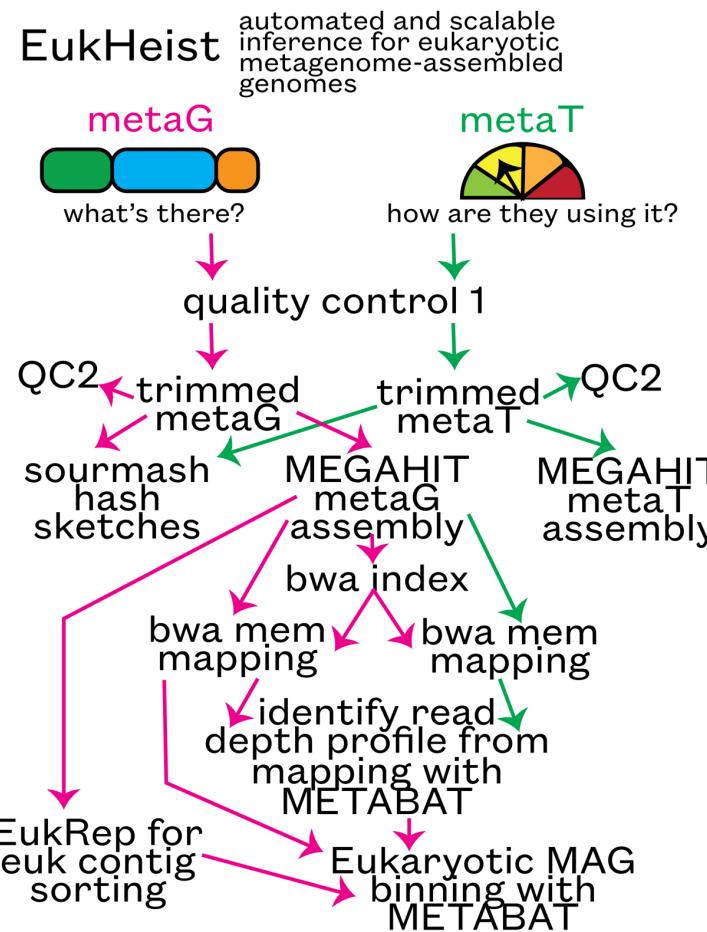


Figure 1. Overall ACR workflow.

<https://github.com/hoonjeseong/acr>

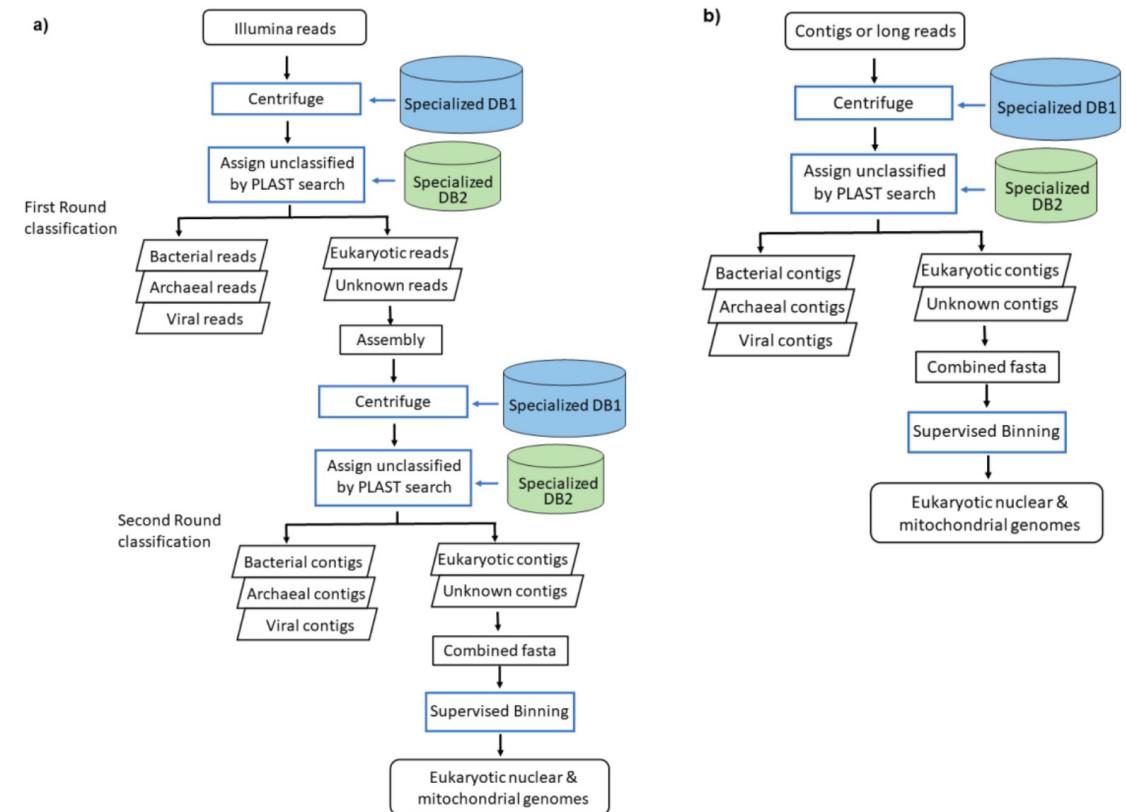
MAG binning pipelines: Eukaryote-specific



<https://github.com/AlexanderLabWHOI/EukHeist>

Alexander et al., *mBio*, 2023
(<https://doi.org/10.1128/mbio.01676-23>)

EukFinder

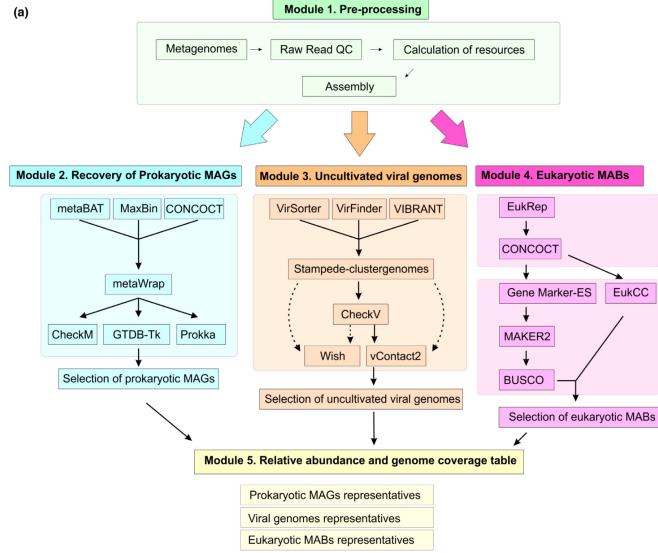


<https://github.com/RogerLab/Eukfinder>

Zhao et al., *bioRxiv*, 2024
(<https://doi.org/10.1101/2023.12.28.573569>)

MAG binning pipelines: General

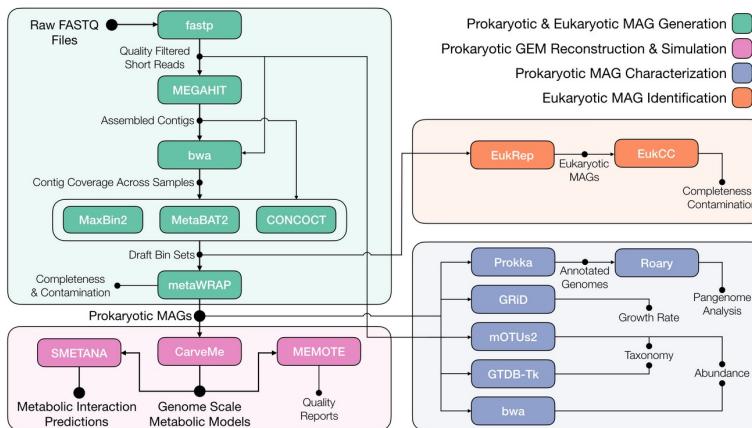
MuDoGeR



<https://github.com/mdsufz/MuDoGeR>

Rocha et al., *Molecular Ecology Resources*,
2022 (<https://doi.org/10.1111/1755-0998.13904>)

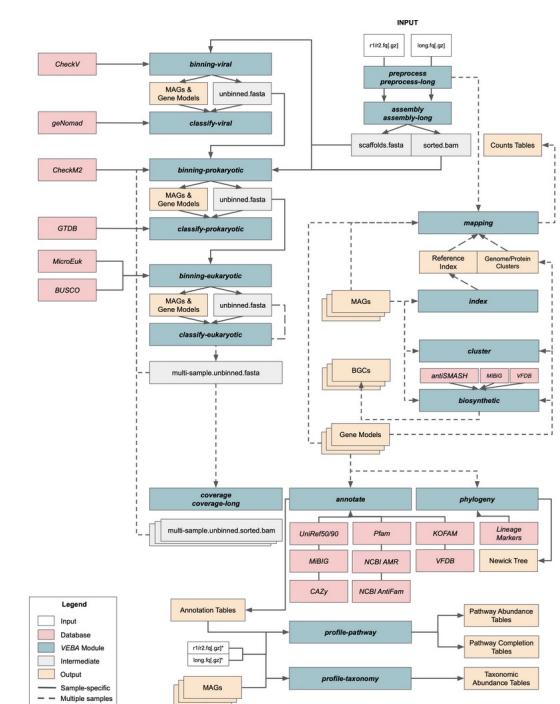
metaGEM



<https://github.com/franciscozorrilla/metaGEM>

Zorrilla et al., *Nucleic Acids Research*, 2021
(<https://doi.org/10.1093/nar/gkab815>)

VEBA



<https://github.com/jolespin/veba>

Espinoza and Dupont, *BMC Bioinformatics*, 2022
(<https://doi.org/10.1186/s12859-022-04973-8>)

Annotating eukaryotic metagenome contigs

Levy Karin et al. *Microbiome* (2020) 8:48
https://doi.org/10.1186/s40168-020-00808-x

RESEARCH

Microbiome

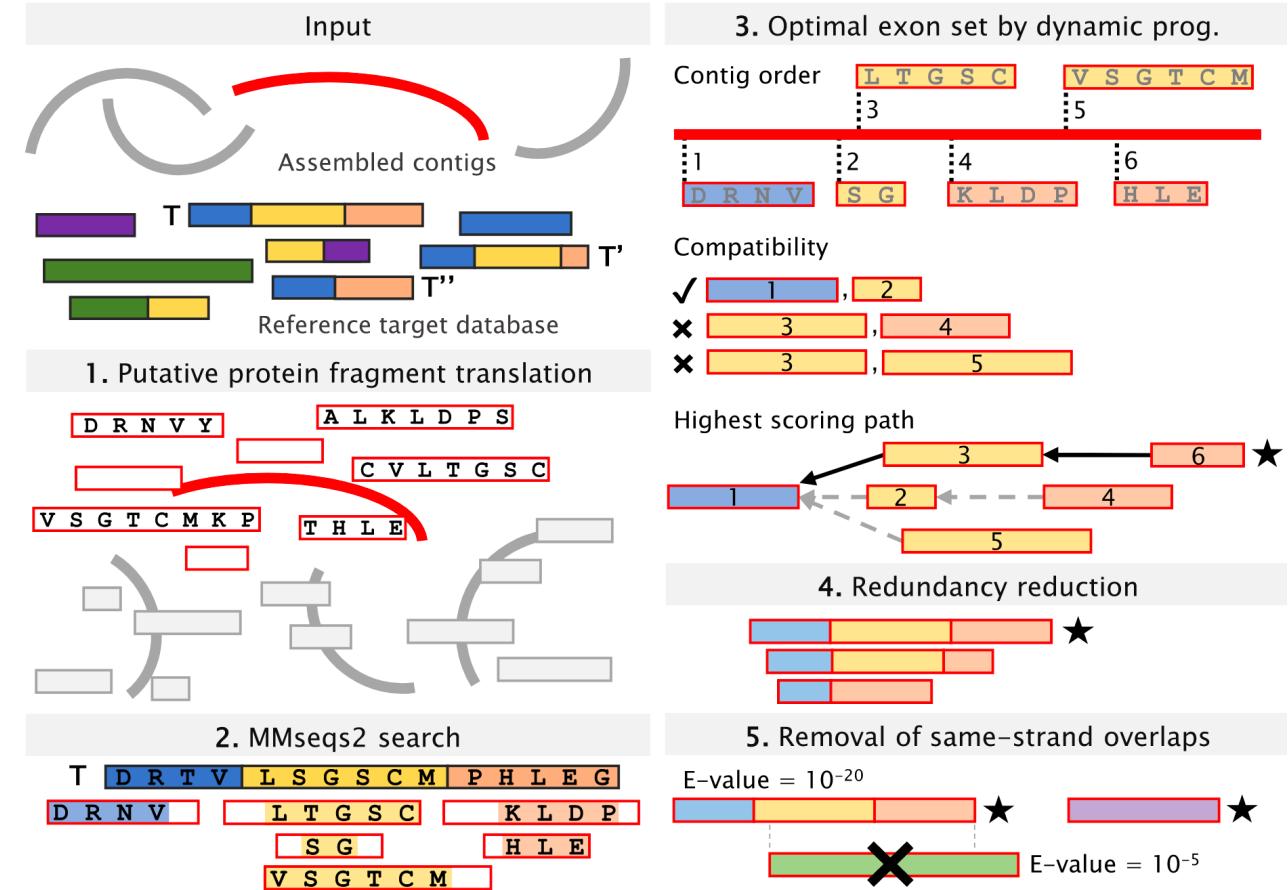
Open Access



MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics

Eli Levy Karin*, Milot Mirdita and Johannes Söding*

<https://github.com/soedinglab/metaeuk>



Estimating MAG completeness and redundancy

Saary et al. *Genome Biology* (2020) 21:244
https://doi.org/10.1186/s13059-020-02155-4

Genome Biology

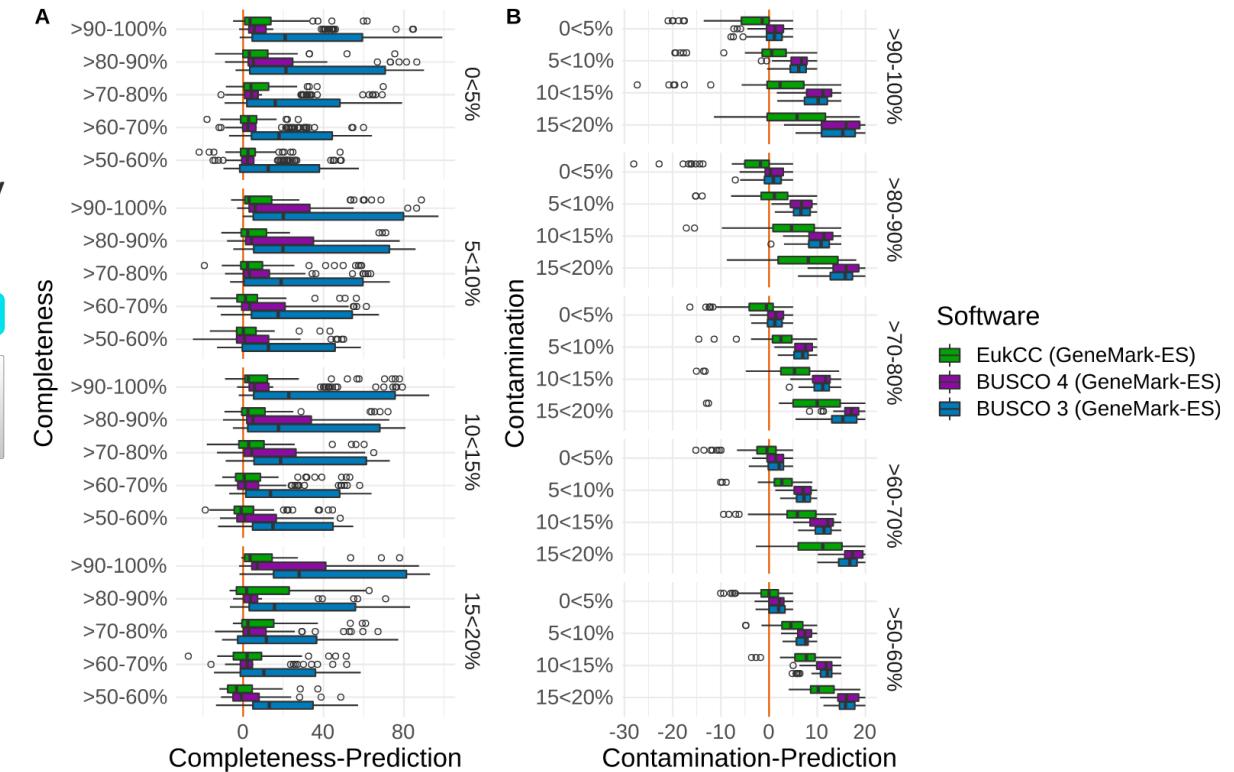
SOFTWARE

Open Access

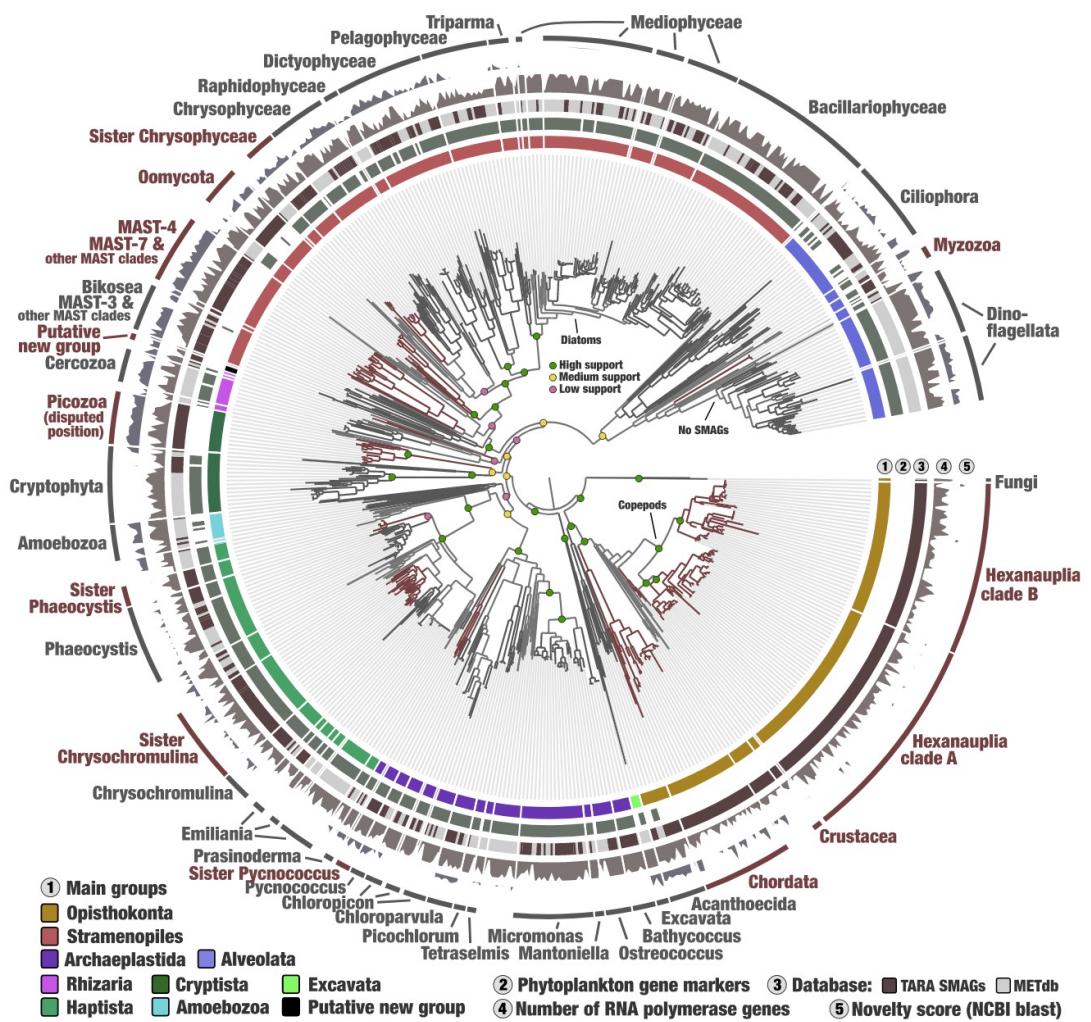
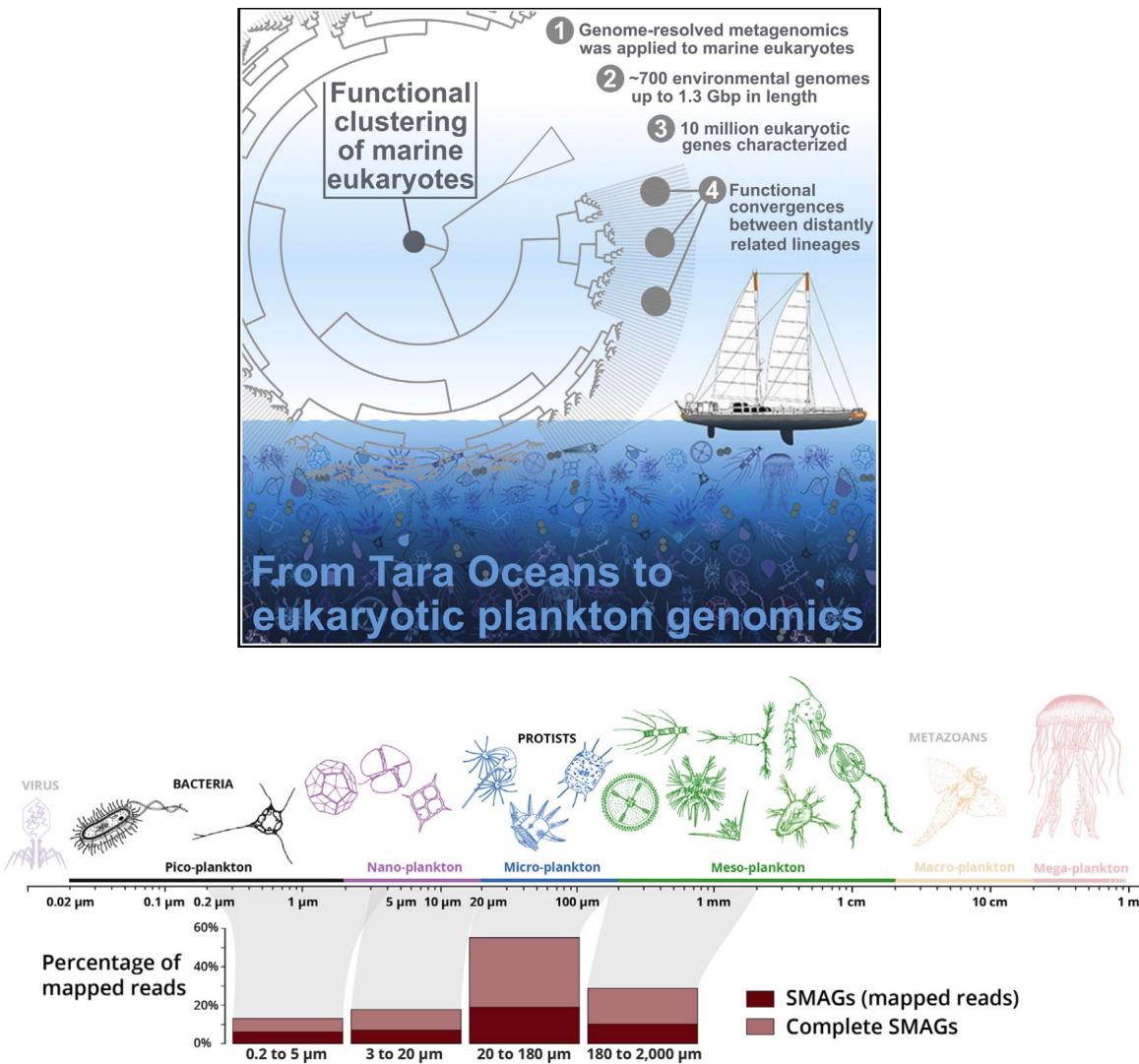
Estimating the quality of eukaryotic genomes recovered from metagenomic analysis with EukCC

Paul Saary , Alex L. Mitchell and Robert D. Finn*

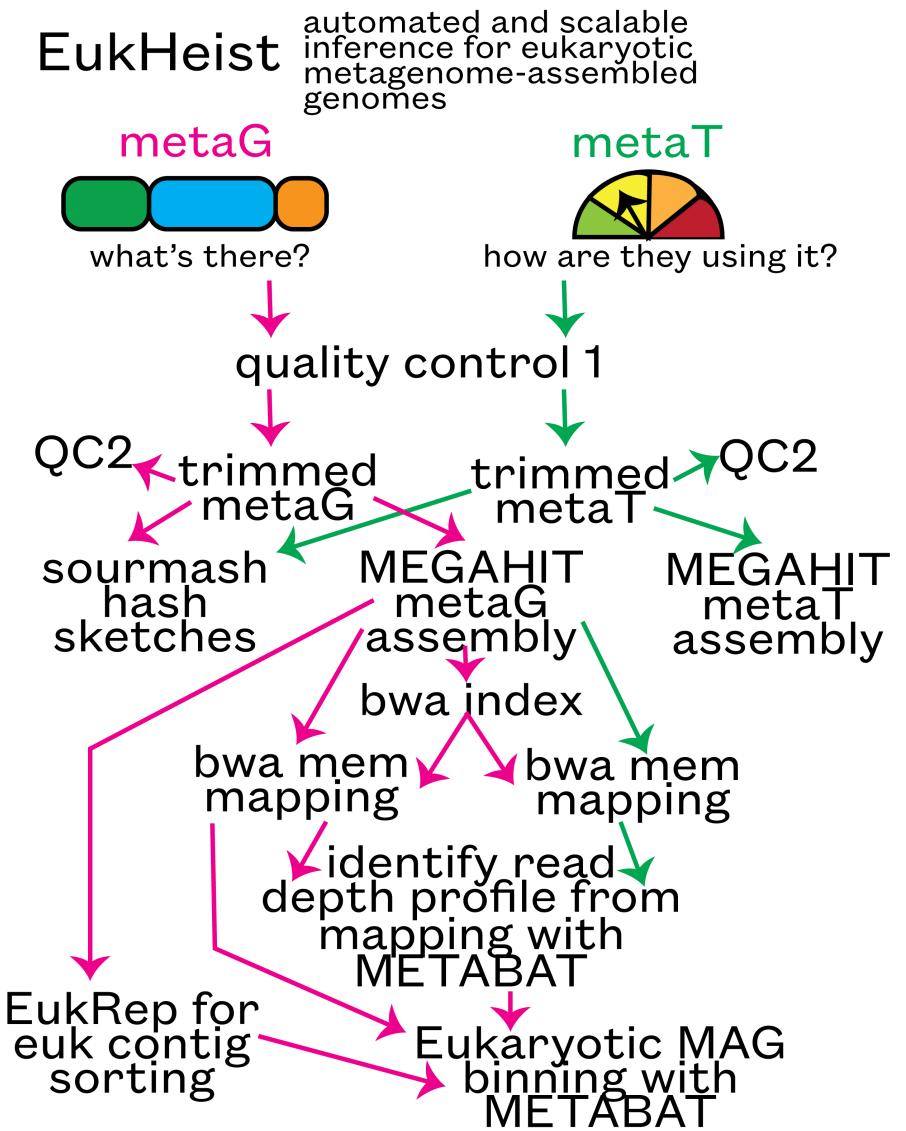
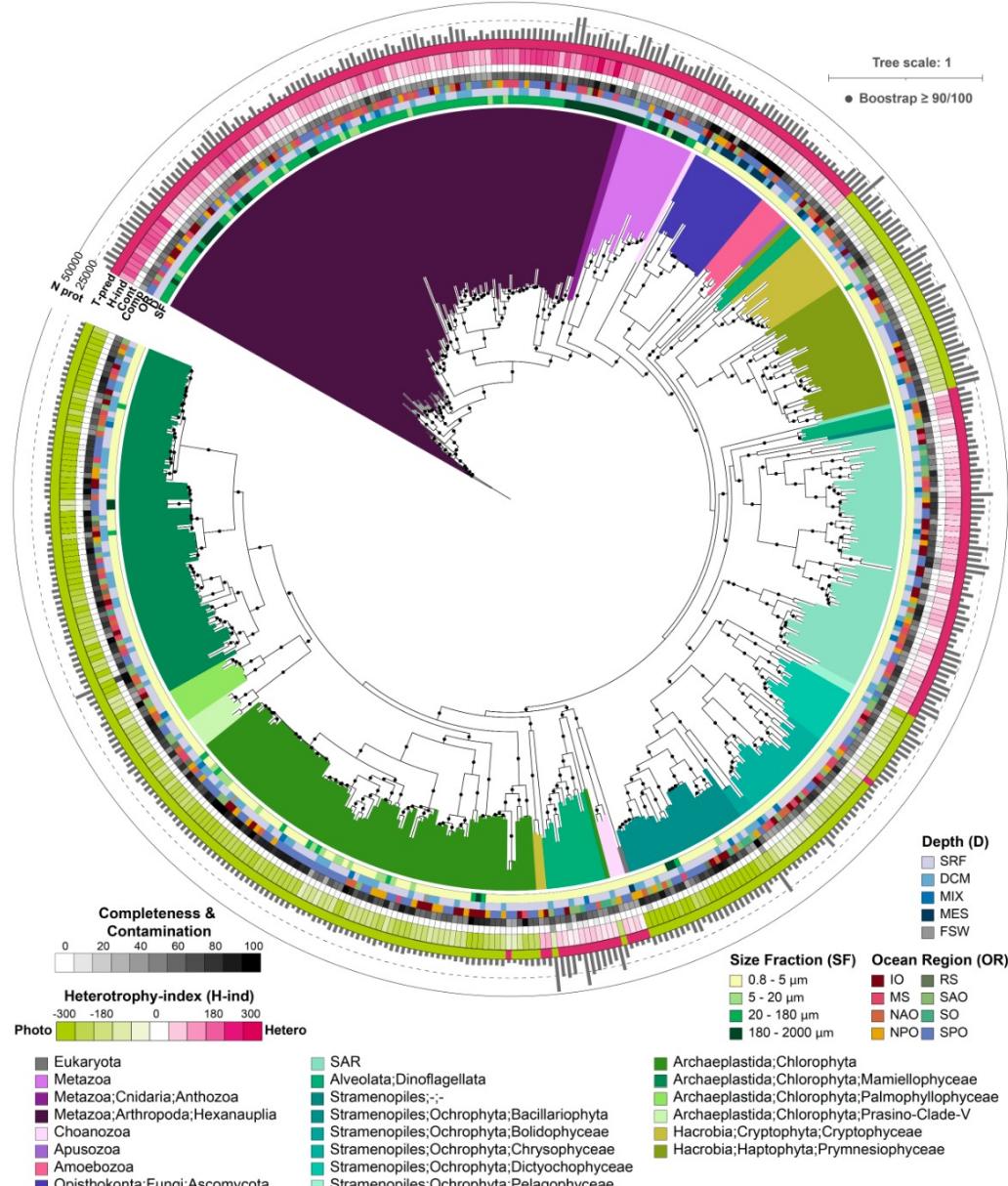
<https://github.com/EBI-Metagenomics/EukCC>



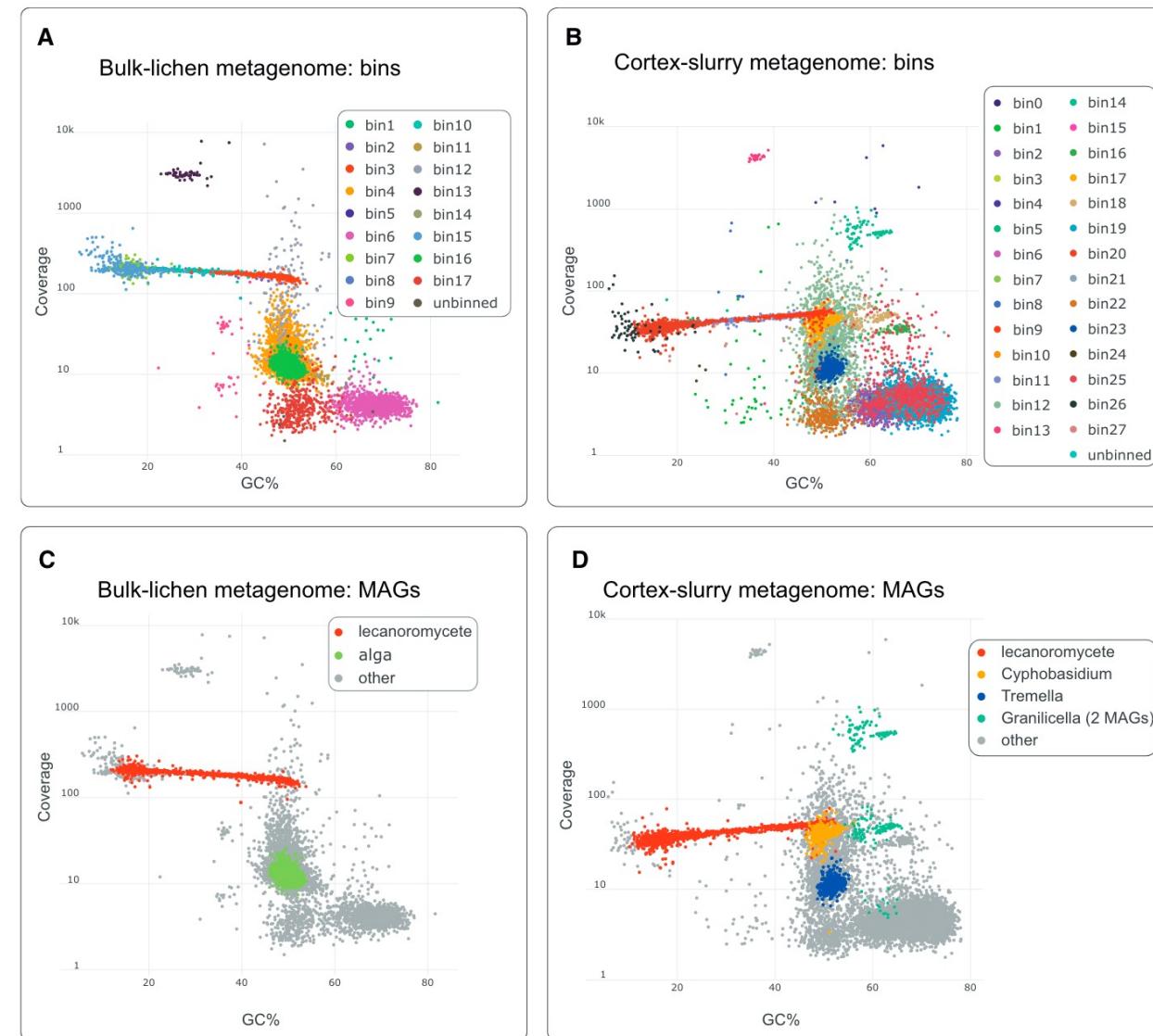
Eukaryotic MAGs: Environmental samples



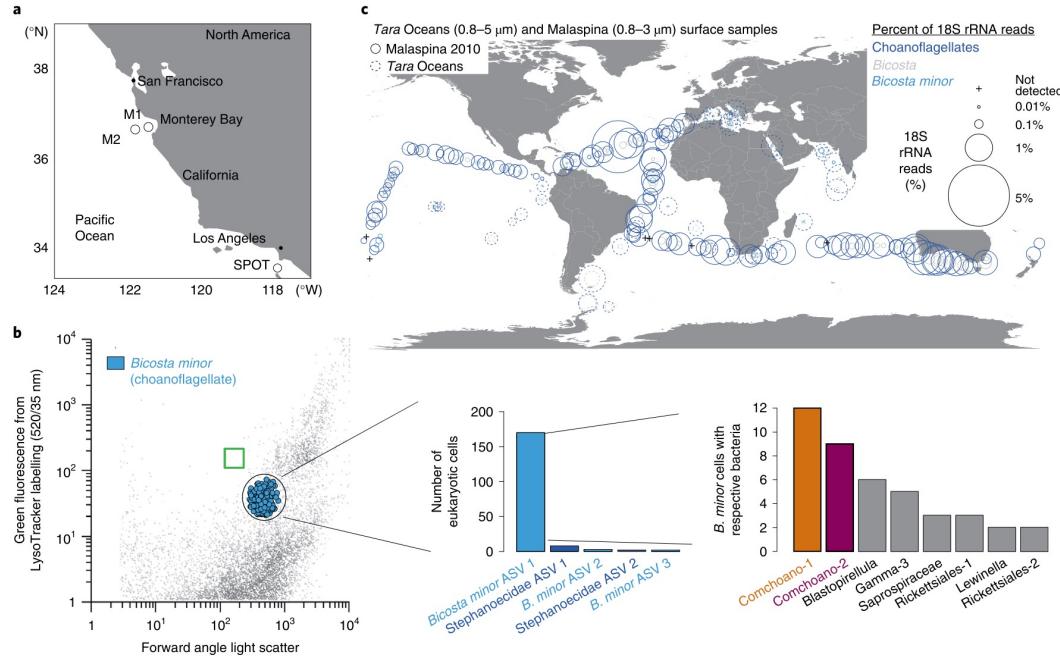
Eukaryotic MAGs: Environmental samples



Eukaryotic MAGs: Low-complexity samples and mixed cultures



Eukaryotic MAGs: Single-cell metagenomes



Needham et al., *Nature Microbiology*, 2022
<https://doi.org/10.1038/s41564-022-01174-0>

Pacific Ocean
 choanoflagellate
Bicosta Minor
 and
 Gammaproteobacteria
 endosymbiont

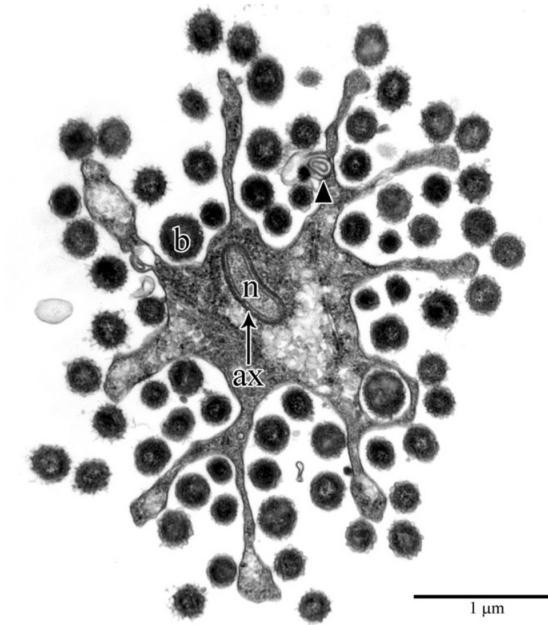
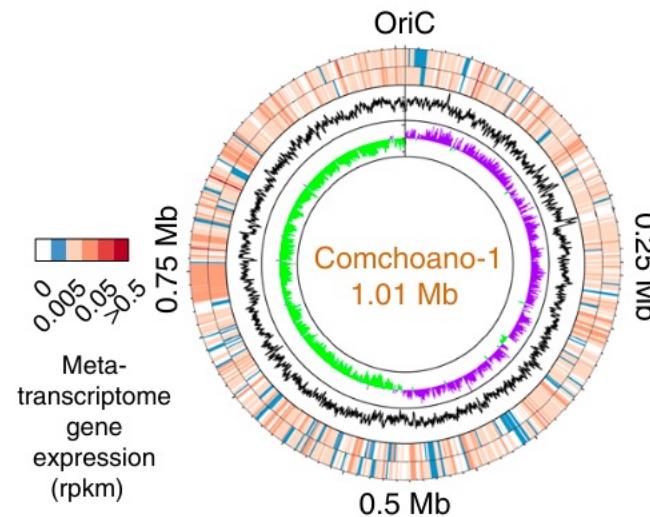
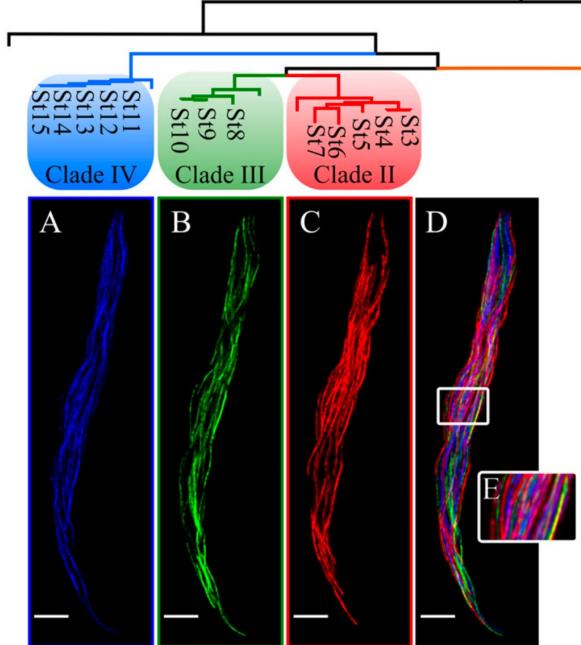


Fig. 1. Transmission electron micrograph of *S. strix* from the gut of *Z. angusticollis* showing the stellar-shaped eukaryotic cell with bacteria (b) attached to the surface. The cell nucleus (n) surrounded by the axostyle (ax) is clearly visible. The arrowhead indicates the putative process of phagocytosis of a bacterium by *S. strix*. (Scale bar, 1 µm.)



Treitli, PNAS, 2019 (
<https://doi.org/10.1073/pnas.1910793116>)

Termite gut oxymonad
Streblomastix strix
 and
 Bacteroidetes symbionts

Associated prokaryote and organellar MAGs

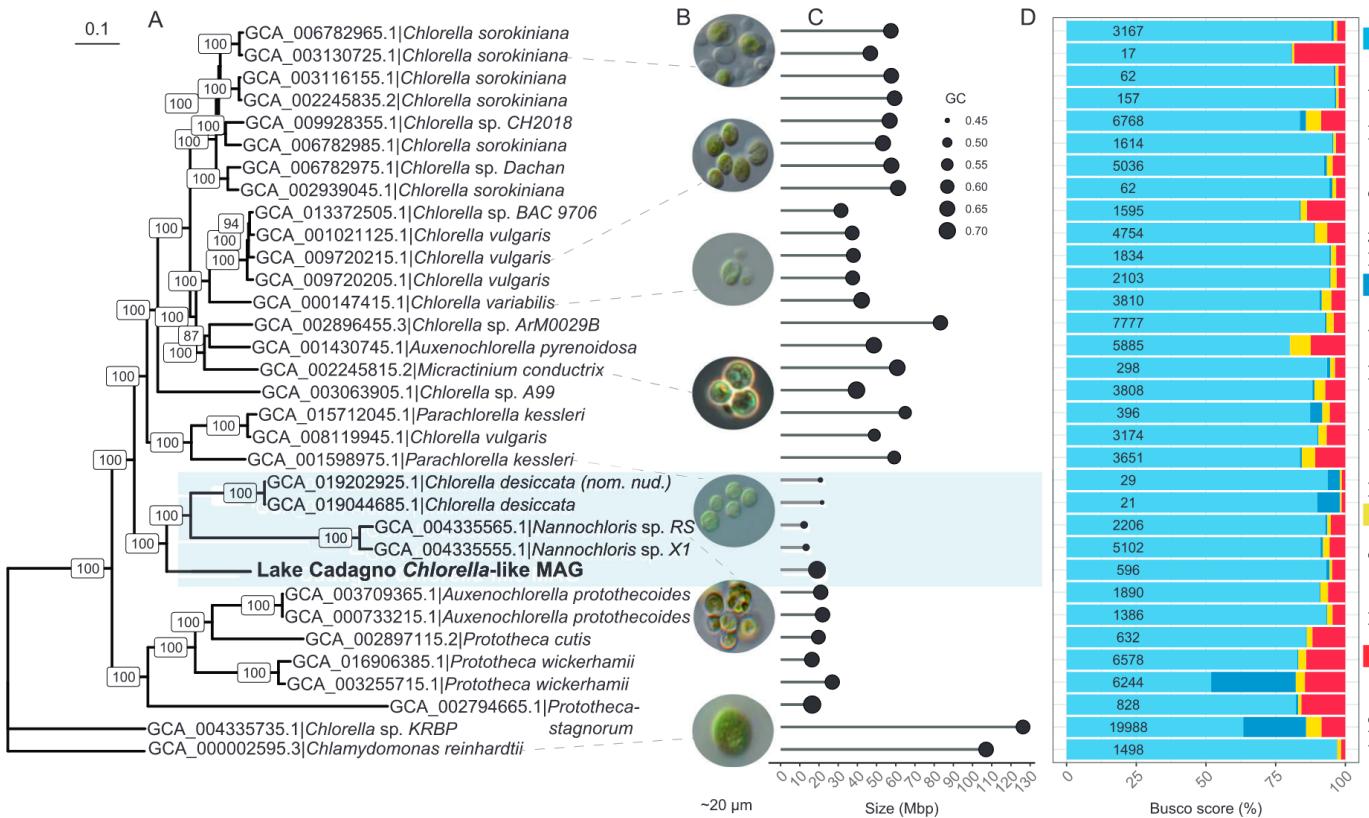
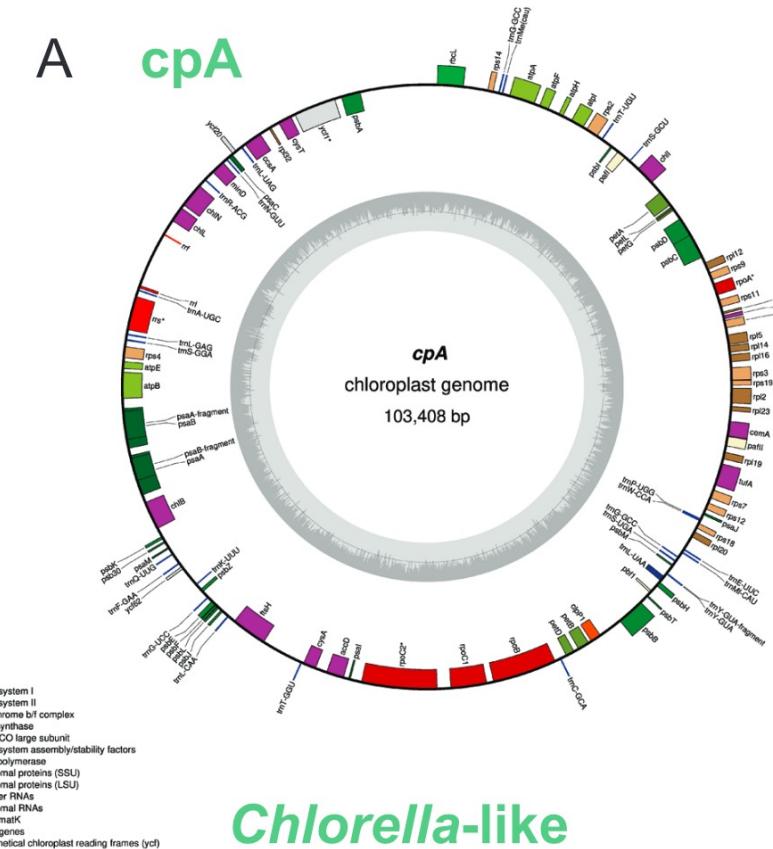
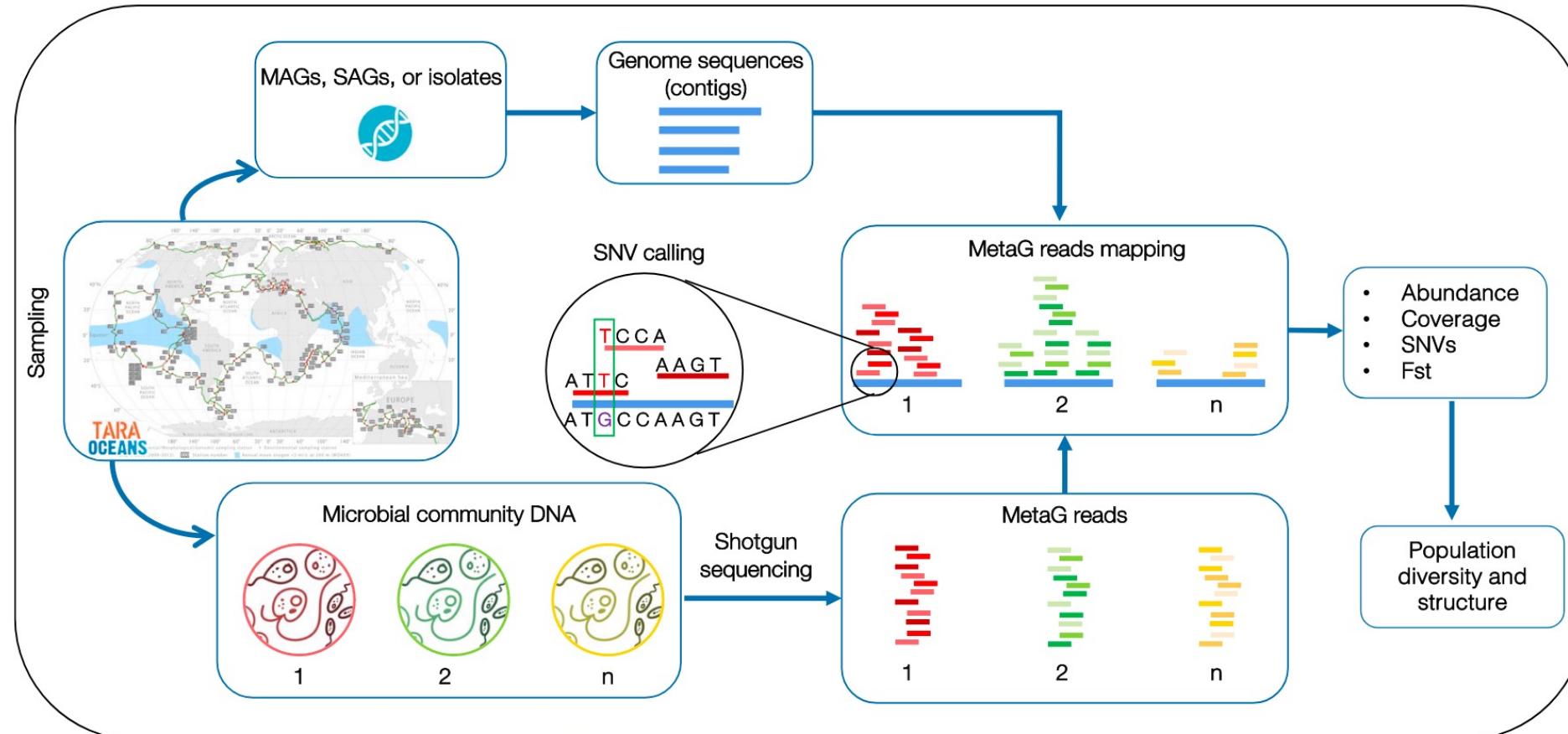


Fig. 3 Comparative analyses of Lake Cadagno *Chlorella*-like MAG with relatives from the Chlorellaceae family (NCBI ID 35461).



Chlorella-like

Metagenome-based population genomics



Caveats of metagenomics for eukaryotes...

Received: 7 April 2022 | Revised: 23 January 2023 | Accepted: 21 February 2023

DOI: 10.1111/1755-0998.13776

RESOURCE ARTICLE

MOLECULAR ECOLOGY
RESOURCES WILEY

Recovery of 197 eukaryotic bins reveals major challenges for eukaryote genome reconstruction from terrestrial metagenomes

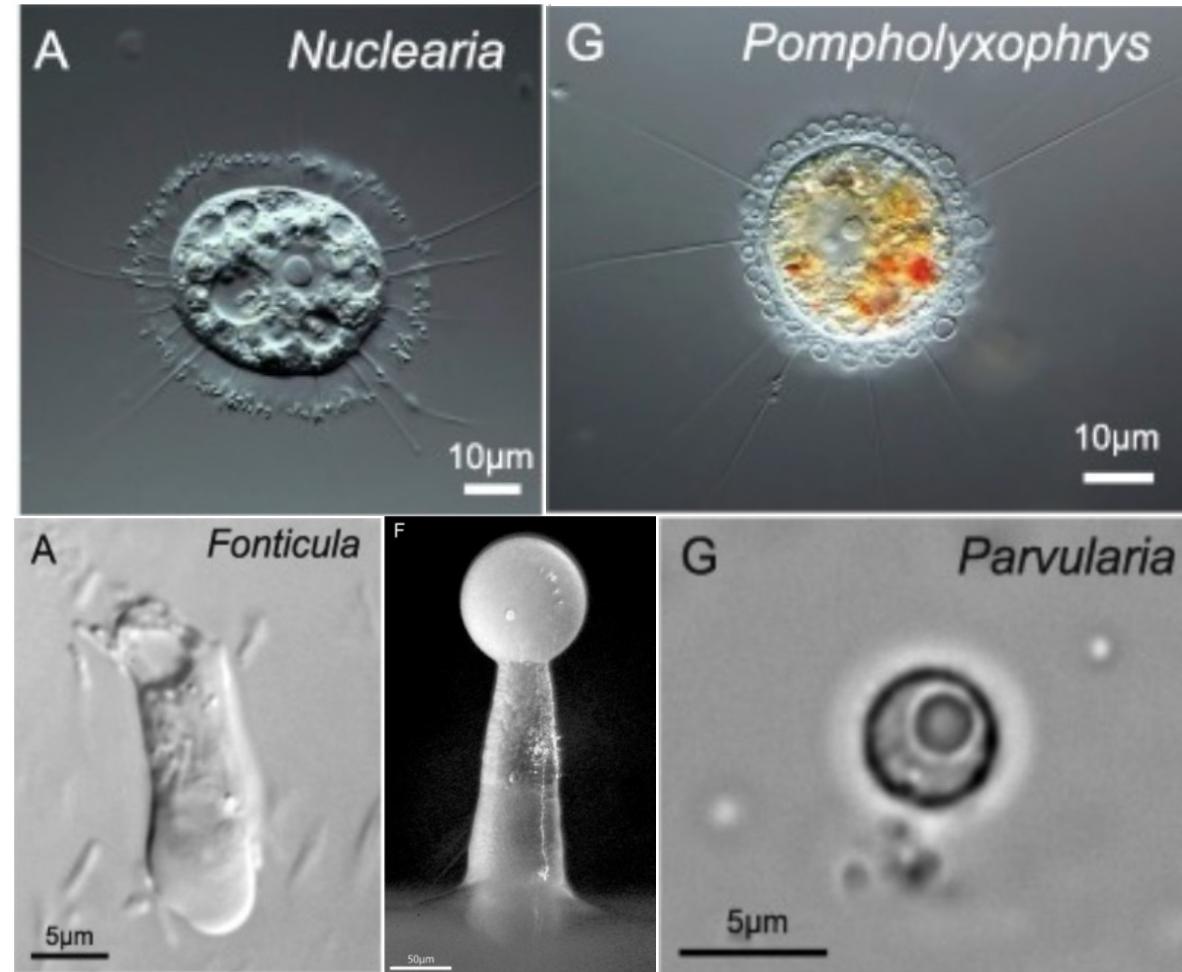
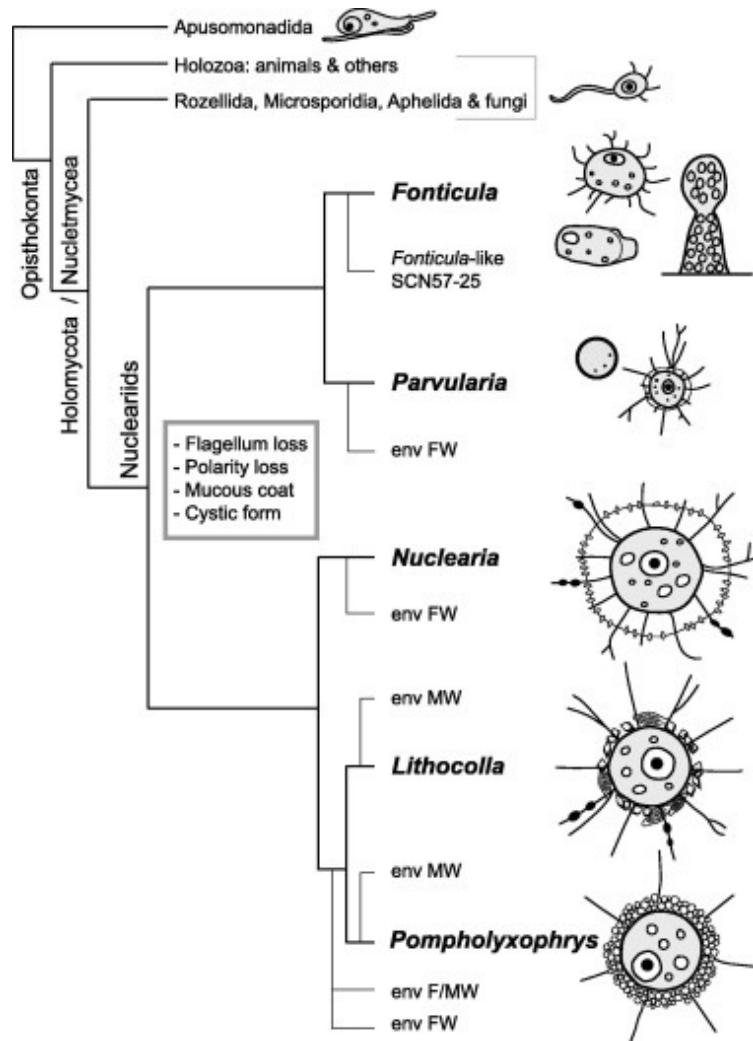
Joao Pedro Saraiva¹ | Alexander Bartholomäus²  | Rodolfo Brizola Toscan¹ |
Petr Baldrian³  | Ulisses Nunes da Rocha¹ 

- Attempted to obtain eukaryotic MAGs using the EukRep pipeline from 6000 terrestrial metagenomes
- Only retrieved eukaryotic bins from 215 metagenomes and 121 MAGs with at least 30% completeness
- Higher sequencing depth needed to obtain eukaryotic genomes
- Multiploidy can cause issues for read assemblers
- Easier to recover eukaryotes with smaller genomes
- The use of **long-read metagenomes** could help with complex eukaryotic genomes
- **Methods development for eukaryotes is needed!** (e.g., to identify introns to improve gene calling)

If I missed any tools
please send them my way ☺

Tutorial Overview

Our example protist group: Nucleariida and *Parvularia atlantis*



Parvularia atlantis genome: ESOM map and manual curation

Article | [Open access](#) | Published: 24 August 2022

Divergent genomic trajectories predate the origin of animals and fungi

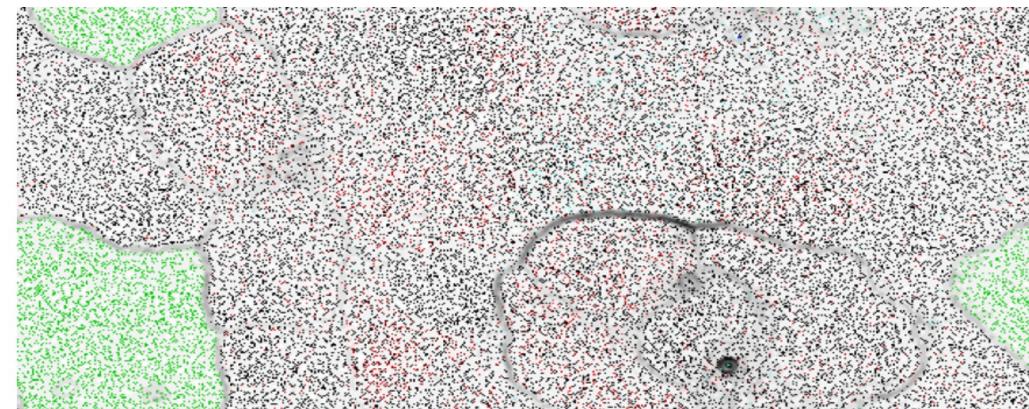
[Eduard Ocaña-Pallarès](#)  , [Tom A. Williams](#), [David López-Escardó](#), [Alicia S. Arroyo](#), [Jananan S. Pathmanathan](#), [Eric Baptiste](#), [Denis V. Tikhonenkov](#), [Patrick J. Keeling](#), [Gergely J. Szöllősi](#) & [Iñaki Ruiz-Trillo](#) 

[Nature](#) **609**, 747–753 (2022) | [Cite this article](#)

29k Accesses | **25** Citations | **245** Altmetric | [Metrics](#)



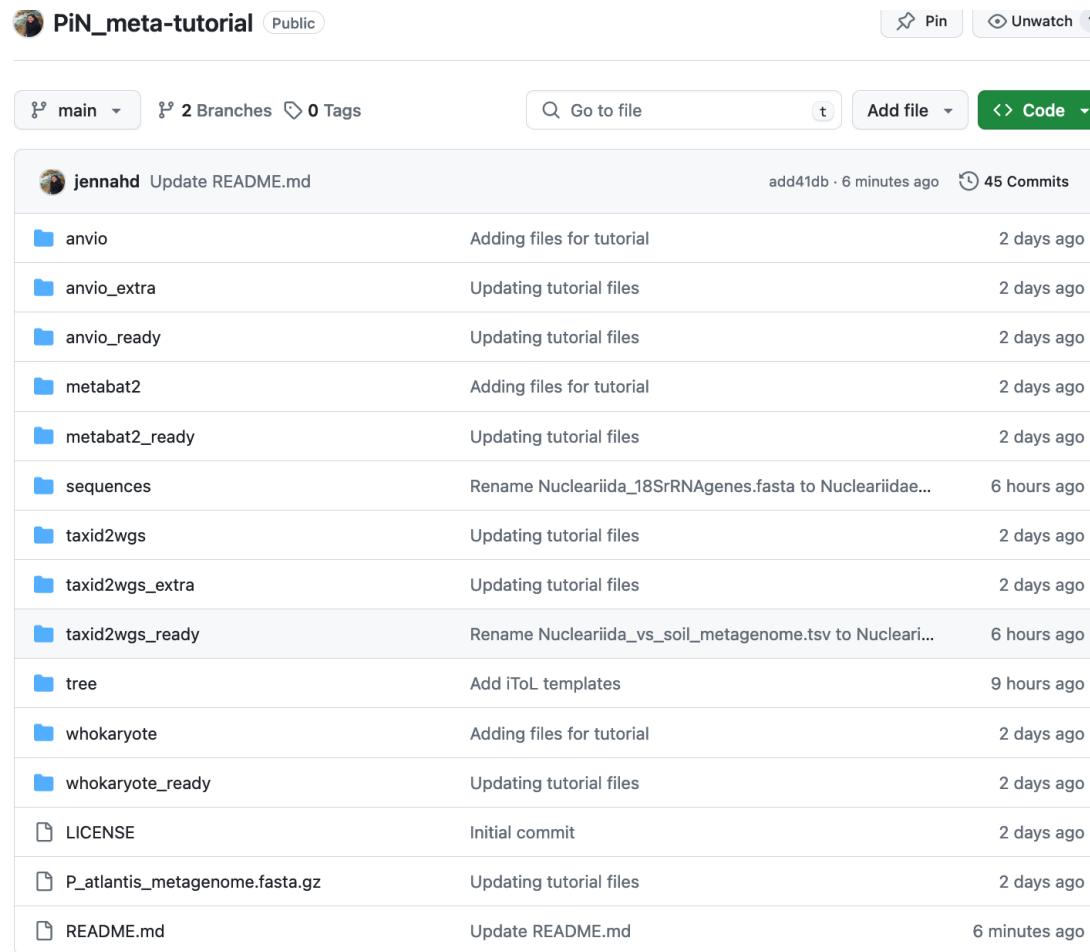
Supplementary Information 1-Fig. 10. ESOM map of the *P. atlantis* contigs from the first assembly, as in Supplementary Information 1-Fig. 9, but colored in green those contig/contig windows included in the *P. atlantis* dataset, and colored in yellow those contig/contig windows that further inspected to determine whether they should be included in the *P. atlantis* dataset, or in the 'Contaminant' set. Other contig/contig windows were included in the 'Contaminant' set.



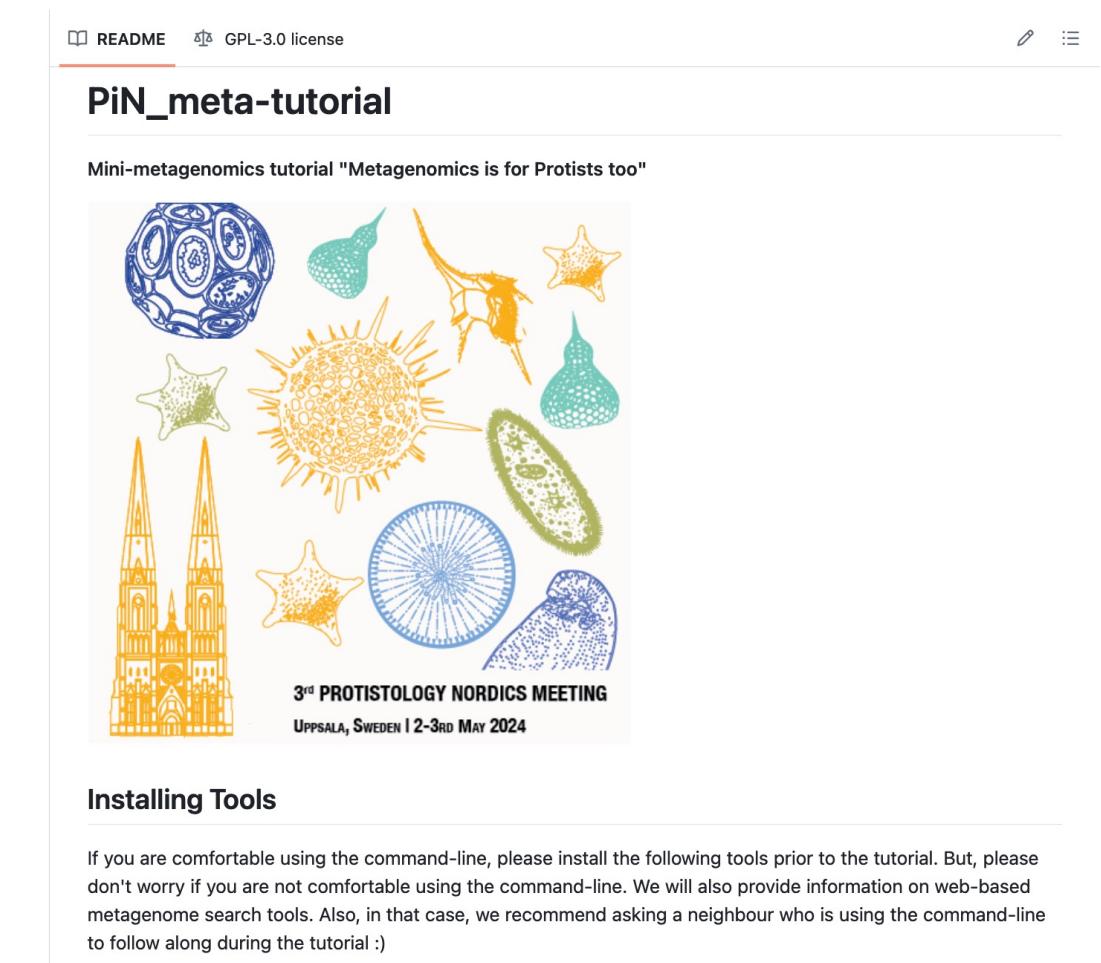
Supplementary Information 1-Fig. 12. ESOM map of the *P. atlantis* scaffolds from the second assembly, as in Supplementary Information 1-Fig. 11, but colored in light green the scaffold windows within the region of the map considered to correspond to *P. atlantis* genome.

Files and tutorial vignette available on Github

https://github.com/jennahd/PiN_meta-tutorial

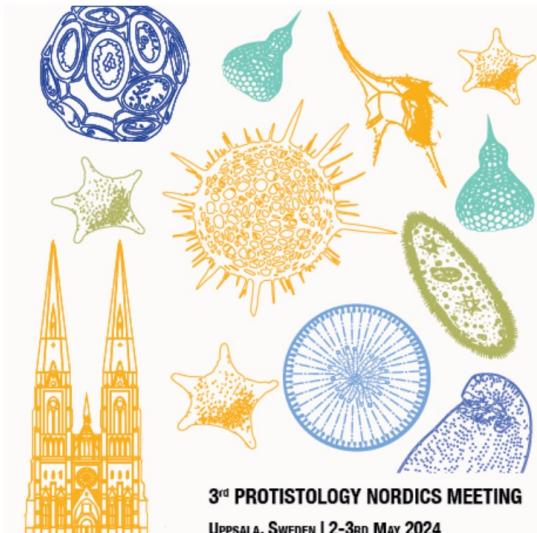


File / Commit Message	Time Ago
jennahd Update README.md	6 minutes ago
anvio Adding files for tutorial	2 days ago
anvio_extra Updating tutorial files	2 days ago
anvio_ready Updating tutorial files	2 days ago
metabat2 Adding files for tutorial	2 days ago
metabat2_ready Updating tutorial files	2 days ago
sequences Rename Nucleariida_18SrRNAGenes.fasta to Nucleariida...	6 hours ago
taxid2wgs Updating tutorial files	2 days ago
taxid2wgs_extra Updating tutorial files	2 days ago
taxid2wgs_ready Rename Nucleariida_vs_soil_metagenome.tsv to Nucleariida...	6 hours ago
tree Add iTOL templates	9 hours ago
whokaryote Adding files for tutorial	2 days ago
whokaryote_ready Updating tutorial files	2 days ago
LICENSE Initial commit	2 days ago
P_atlantis_metagenome.fasta.gz Updating tutorial files	2 days ago
README.md Update README.md	6 minutes ago



PiN_meta-tutorial

Mini-metagenomics tutorial "Metagenomics is for Protists too"

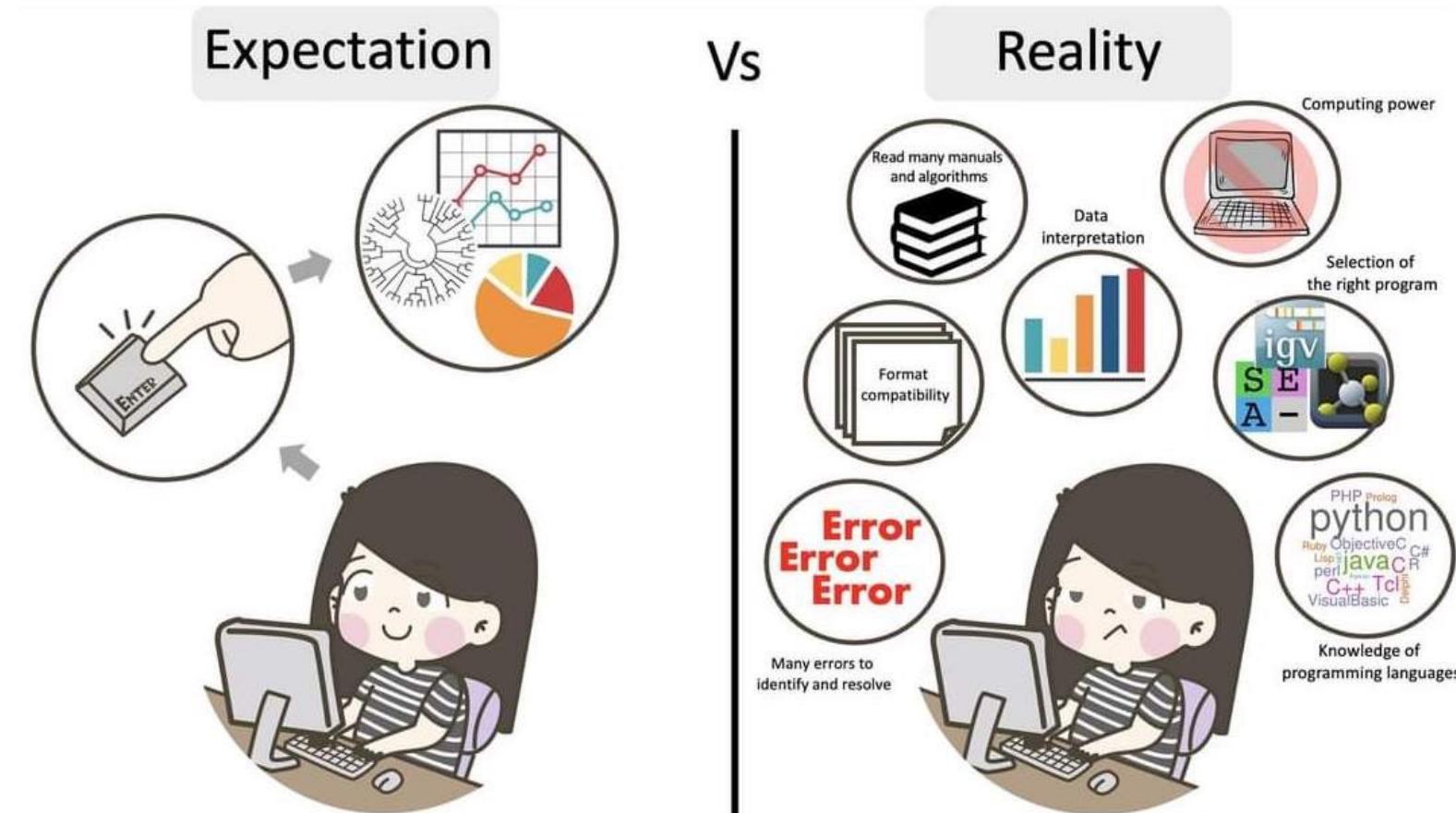


3rd PROTISTOLOGY NORDICS MEETING
UPPSALA, SWEDEN | 2-3RD MAY 2024

Installing Tools

If you are comfortable using the command-line, please install the following tools prior to the tutorial. But, please don't worry if you are not comfortable using the command-line. We will also provide information on web-based metagenome search tools. Also, in that case, we recommend asking a neighbour who is using the command-line to follow along during the tutorial :)

Don't get discouraged if things don't work right away 😊



Tutorial Part 1: Searching Metagenomes

Searching publicly available assembled metagenomes (WGS)

NCBI Taxonomy Browser

Search for **metagenomes** as complete name lock

Display 3 levels using filter: none

Nucleotide Protein Structure Genome Popset SNP
 Gene HomoloGene SRA Experiments LinkOut BLAST GEO Profiles
 BioSample Assembly dbVar Genetic Testing Registry Host Viral Host

[Lineage](#) (full): unclassified entries; unclassified sequences

- o [metagenomes](#) LinkOut Click on organism name to get more information.
 - o [ecological metagenomes](#) LinkOut
 - [abattoir metagenome](#)
 - [activated carbon metagenome](#)
 - [activated sludge metagenome](#)
 - [adhesive metagenome](#)
 - [aerosol metagenome](#)
 - [air conditioner metagenome](#)
 - [air metagenome](#)
 - [alkali sediment metagenome](#)
 - [alkali spring metagenome](#)
 - [anaerobic digester metagenome](#)
 - [anchialine metagenome](#)
 - [ant fungus garden metagenome](#)
 - [aquaculture metagenome](#)
 - [aquarium metagenome](#)
 - [aquatic metagenome](#)
 - [aquifer metagenome](#)
 - [ballast water metagenome](#)
 - [beach sand metagenome](#)
 - [bentonite metagenome](#)
 - [bioanode metagenome](#)
 - [biocathode metagenome](#)
 - [biochar metagenome](#)
 - [biocrust metagenome](#)
 - [biofilm metagenome](#)
 - [biofilter metagenome](#)

blastn blastp blastx tblastn tblastx

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) Clear

Query subrange From To

Or, upload file No file selected. ?

Job Title
Enter a descriptive title for your BLAST search ?

Align two or more sequences ?

Choose Search Set

Database Standard databases (nr etc.): rRNA/ITS databases Genomic + transcriptomes New Experimental databases

Limit by Whole-genome shotgun contigs (wgs) ?

Limit to Organism BioProjectID WGS Project

eukaryotes ?

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown ?

Sequences from type material

Try experimental taxonomic nomenclature For more info see What are taxonomic names?

Things to consider

- *Which species do we find closer hits to in moss metagenomes?*
- *How many would you say are "close" hits to each species? Are they the same species?*
- *Do we find the same top hits for the different searched species?*
- *Do you expect that the top hits will cluster together with the species in a tree?*

Searching publicly available metagenomic reads (SRA)

Branchwater Metagenome Query

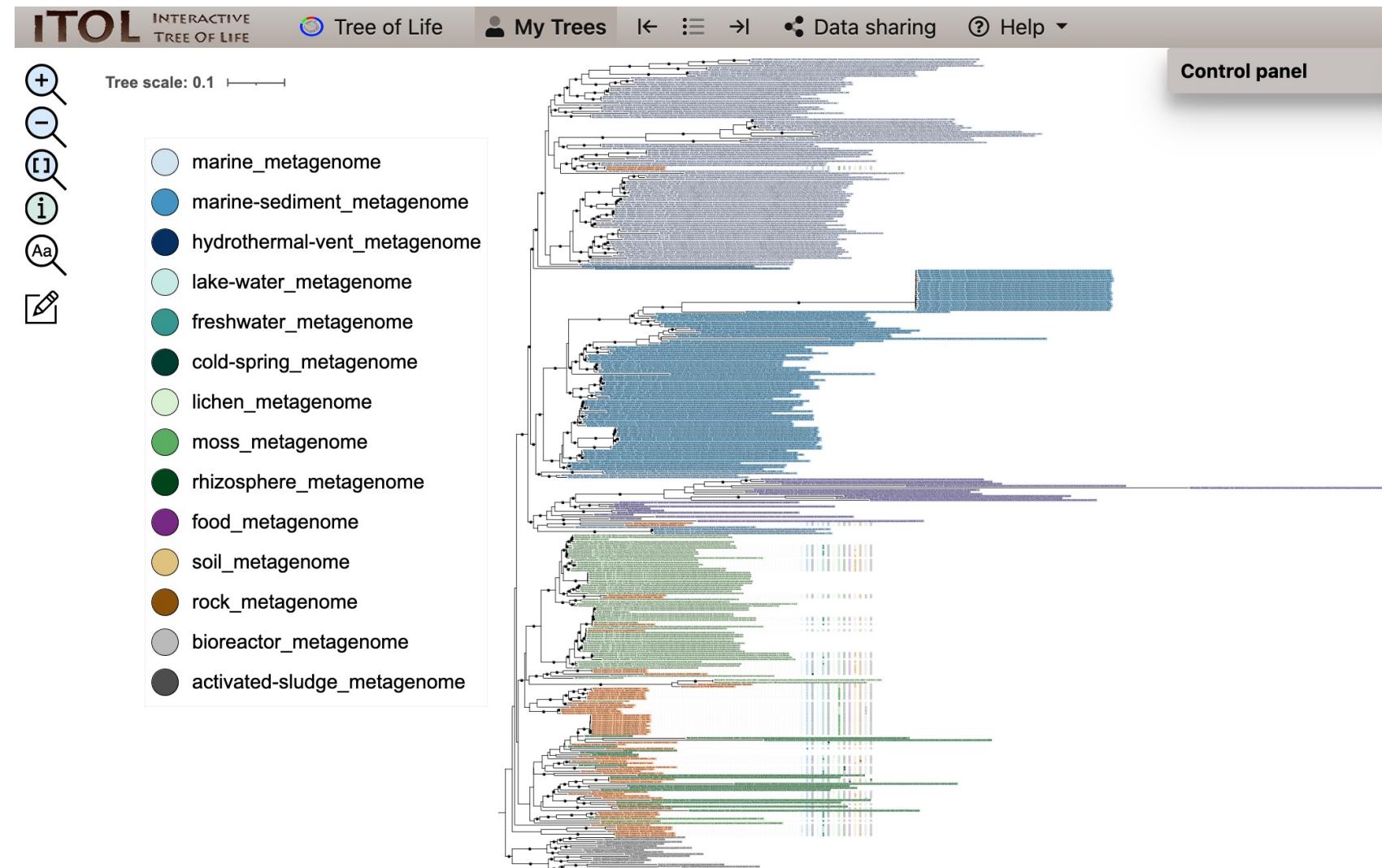
Real-time search for a genome within metagenomes in the SRA.

[Home](#) [Advanced](#) [About](#) [Examples](#) [Contact](#)

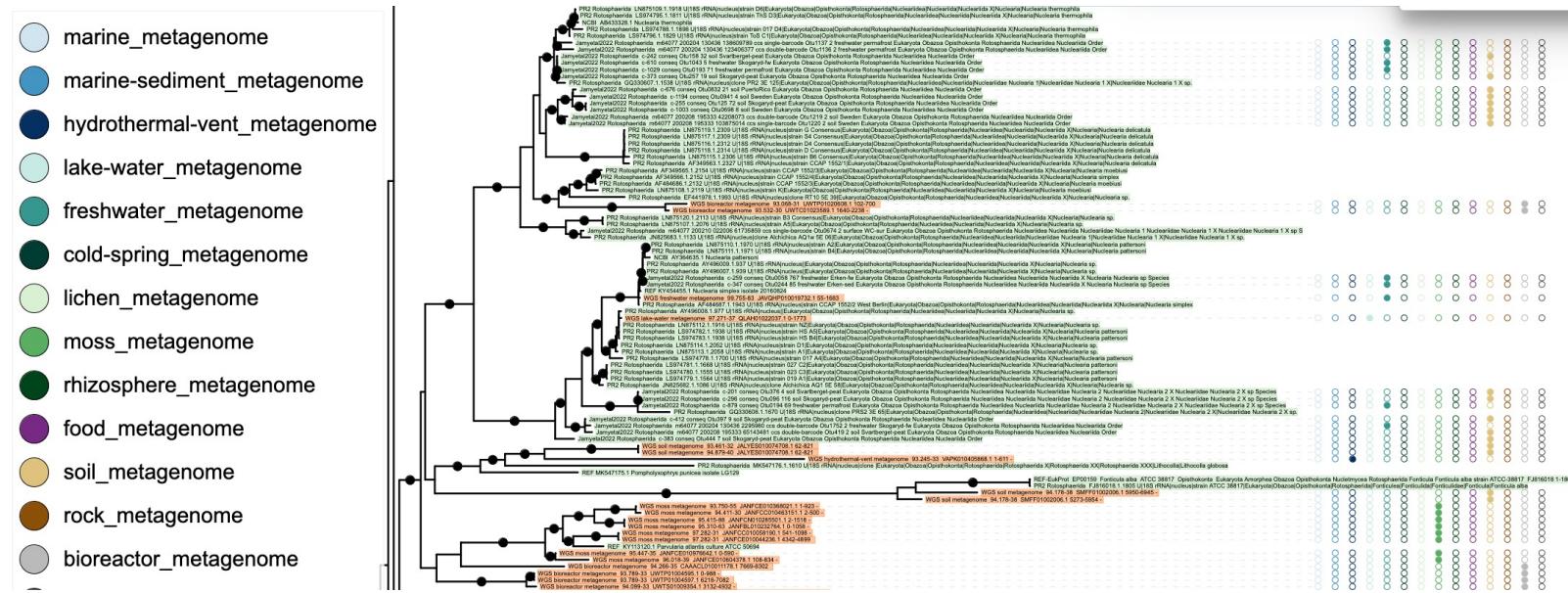
Things to consider

- *What does the location data tell us about the global distribution of hits?*
- *Which kinds of environments are most hits from? Are there any patterns?*
- *Which specific environments are the closest relatives of each species found in?*

Phylogenetic tree including metagenomic sequences



Things to consider



- Where do our metagenomic sequences go in the tree? Are they all affiliated with Nucleariidae (Rotosphaerida)? Perhaps you can see why it is important to always check hits with a phylogeny.
- Which environments do the metagenomic and long-read metabarcoding sequences that affiliate with the different Nucleariidae species come from?
- Do some species have more metagenomic sequences than long-read metabarcoding sequences affiliated with them? What could this tell us about biases with either type of data?

Tutorial Part 2: Binning metagenomes

Separating prokaryotic and eukaryotic contigs and binning

> *Microb Genom.* 2022 May;8(5):mgen000823. doi: 10.1099/mgen.0.000823.

Whokaryote: distinguishing eukaryotic and prokaryotic contigs in metagenomes based on gene structure

Lotte J U Pronk¹, Marnix H Medema¹

Affiliations + expand

PMID: 35503723 PMCID: [PMC9465069](#) DOI: [10.1099/mgen.0.000823](#)

[PeerJ.](#) 2019; 7: e7359.

Published online 2019 Jul 26. doi: [10.7717/peerj.7359](#)

PMCID: [PMC6662567](#)

PMID: [31388474](#)

MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies

[Dongwan D. Kang](#),¹ [Feng Li](#),² [Edward Kirton](#),¹ [Ashleigh Thomas](#),¹ [Rob Egan](#),¹ [Hong An](#),² and [Zhong Wang](#)^{✉1,3,4}

Academic Editor: Joseph Gillespie

Interactive metagenome visualization and manual binning



nature microbiology

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature microbiology](#) > [comment](#) > [article](#)

Comment | Published: 21 December 2020

Community-led, integrated, reproducible multi-omics with anvi'o

A. Murat Eren , Evan Kiefl, Alon Shaiber, Iva Veseli, Samuel E. Miller, Matthew S. Schechter, Isaac Fink, Jessica N. Pan, Mahmoud Yousef, Emily C. Fogarty, Florian Trigodet, Andrea R. Watson, Özcan C. Esen, Ryan M. Moore, Quentin Clayssen, Michael D. Lee, Veronika Kivenson, Elaina D. Graham, Bryan D. Merrill, Antti Karkman, Daniel Blankenberg, John M. Eppley, Andreas Sjödin, Jarrod J. Scott, ... Amy D. Willis

+ Show authors

[Nature Microbiology](#) 6, 3–6 (2021) | [Cite this article](#)

9384 Accesses | 294 Citations | 97 Altmetric | [Metrics](#)

Anvi'o has many tutorials and workflows available

An introduction to Unix

 Mike Lee

Basics of the command line environment, the best friend of any scientist who wishes to work with data.

hands-on beginner

Basics of metagenomic read recruitment

 A. Murat Eren (Meren)

A very simple introduction to metagenomic read recruitment.

metagenomics theoretical beginner

A metagenomic read recruitment exercise

 A. Murat Eren (Meren)

A mini hands-on exercise is to conduct a simple read recruitment experiment with a mock dataset.

metagenomics hands-on beginner

The anvi'o pangenomics workflow

 A. Murat Eren (Meren)  Alon Shaiber  Mahmoud Yousef  Özcan C. Esen

A comprehensive description of the anvi'o pangenomics workflow.

pangenomics theoretical beginner

A pangenomics exercise

 A. Murat Eren (Meren)

A reproducible pangenomic analysis of seven *Vibrio jasicida* genomes isolated in Woods Hole by the participants of Microbial Diversity Course.

pangenomics hands-on beginner

An exercise on single-cell genomics

 Florian Trigodet

A hands-on tutorial to learn how to integrate single-cell genomics data. It covers functional and taxonomic assignment, pangenomic and phylogenomic analyses, and more.

pangenomics phylogenomics hands-on beginner

Anvi'o phylogenomics workflow

 A. Murat Eren (Meren)

A comprehensive description of the anvi'o phylogenomics workflow.

phylogenomics theoretical beginner

An exercise on metabolic reconstruction in anvi'o

 Iva Veseli

A hands-on exercise for the metabolism suite of programs in anvi'o. It starts with a section on how to estimate metabolism for a single bacterial isolate, and goes into larger, real-world dataset of metagenome-assembled genomes to demonstrate how to estimate metabolism compute enrichment scores for metabolic modules.

metabolism metabolic-enrichment genomics metagenomics hands-on beginner

Anvi'o tools for microbial population genetics

 Evan Kiefl  Tom O. Delmont  A. Murat Eren (Meren)

A theoretical tutorial describing how to characterize SNVs, SCVs and SAAVs with anvi'o, and how to interpret the output.

microbial-population-genetics theoretical intermediate

A primer on anvi'o with the Infant Gut Dataset

 A. Murat Eren (Meren)  Iva Veseli  Evan Kiefl  Tom O. Delmont

A tutorial that walks you through some of the most fundamental aspects of anvi'o and its application to a real-world dataset organized in multiple interconnected chapters, which all use the same dataset.

metagenomics binning pangenomics phylogenomics microbial-population-genetics hands-on beginner

A metagenomics workshop featuring anvi'o

 Iva Veseli  Matthew Schechter

Materials from a 2022 workshop on metagenomics in anvi'o, including read recruitment and metabolism estimation.

metagenomics metabolism hands-on beginner

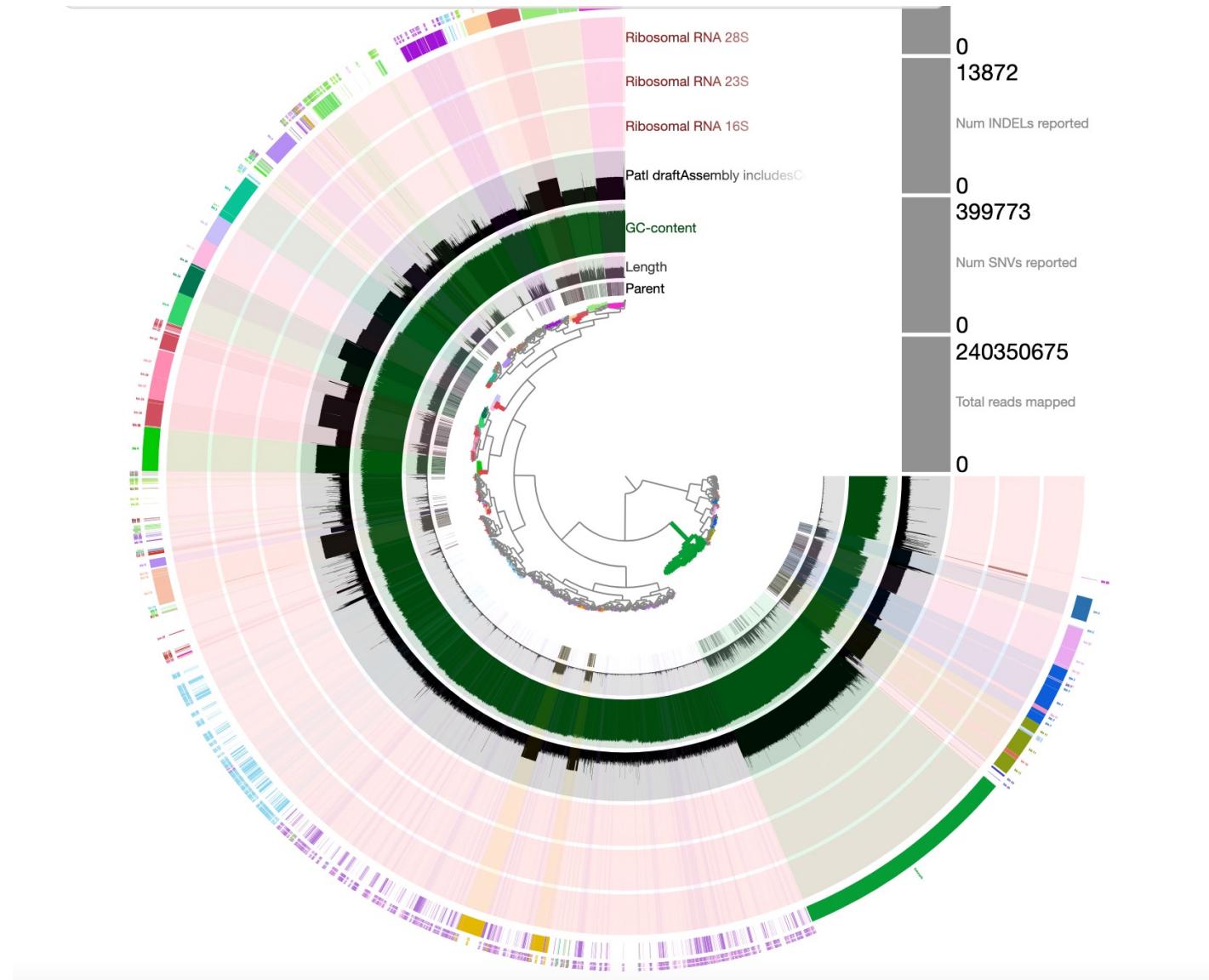
Another anvi'o workshop on metagenomics

 Iva Veseli  Matthew Schechter

Materials and a tutorial from the 2023 iteration of our metagenomics workshop.

metagenomics metabolism hands-on beginner

Using Anvi'o for metagenomic binning and manual curation



Things to consider

- *Which cluster do you think most likely corresponds to *Parvularia atlantis*? Try manually binning it and check its completeness and redundancy if you haven't.*
- *How well did Whokaryote perform at distinguishing prokaryotic and eukaryotic contigs?*
- *How well did MetaBAT2 perform at binning a eukaryotic MAG?*
- *How do the binned bacterial MAGs look? Do they have high quality? Which bacterial groups are present in the mixed culture with *Parvularia atlantis*?*

Thank you! Questions?

Jennah Dharamshi (UU)

A big thank you to the TAs!

Mahwash Jamy (SLU)

Mercè Montoliu Nerín (CBGP-UPM)

Nina Pohl (UU)

