




## Supplementary Material for

# metaPR<sup>2</sup>: a database of eukaryotic 18S rRNA metabarcodes with an emphasis on protists.

Daniel Vaultot<sup>1</sup> ✉ , Clarence Wei Hung Sim<sup>2</sup>, Denise Ong<sup>2</sup> , Bryan Teo<sup>2</sup>, Charlie Biver<sup>3</sup>, Mahwash Jamy<sup>3</sup>, Adriana Lopes dos Santos<sup>2</sup> 

<sup>1</sup> UMR 7144, ECOMAP, CNRS, Sorbonne Université, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>2</sup> Asian School of the Environment, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

<sup>3</sup> Department of Organismal Biology (Systematic Biology), Uppsala University, Uppsala, Sweden

### ORCID

- Daniel Vaultot: 0000-0002-0717-5685
- Adriana Lopes dos Santos: 0000-0002-0736-4937
- Denise Ong: 0000-0001-6053-6948
- Clarence Wei Hung Sim: 0000-0003-2190-7261
- Mahwash Jamy: 0000-0002-2930-9226

✉ Corresponding author: vaultot@gmail.com

Submitted to **Molecular Ecology Resources**

Date: May 26, 2022

**Keywords:** 18S rRNA, metabarcodes, database, R, shiny, PCR, protists

**Short title:** metaPR<sup>2</sup> - a database of eukaryotic metabarcodes

**Table S1:** Ecological function of taxa according to Table S1 of Sommeria-Klein et al. (2021). Taxa present in the PR2 database for which ecological function was not present in Table S1 were assigned an ecological function based on the literature. Ecological function was propagated to all taxa below the taxon for which it was defined using an R script.

Taxon	Taxonomic level	Function	Reference
Acantharea	class	phagotrophs	Sommeria-Klein et al. 2021
Annelida	class	metazoans	Sommeria-Klein et al. 2021
Apicomplexa	class	parasites	Sommeria-Klein et al. 2021
Arthropoda	class	metazoans	Sommeria-Klein et al. 2021
Endomyxa-Ascetosporea	class	parasites	Sommeria-Klein et al. 2021
Ascomycota	class	phagotrophs	Sommeria-Klein et al. 2021
Bacillariophyta	class	phototrophs	Sommeria-Klein et al. 2021
Basidiomycota	class	phagotrophs	Sommeria-Klein et al. 2021
Bicoecea	class	phagotrophs	Sommeria-Klein et al. 2021
Bolidophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Bryozoa	class	metazoans	Sommeria-Klein et al. 2021
Chaetognatha	class	metazoans	Sommeria-Klein et al. 2021
Chlorarachniophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Chlorophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Chloropicophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Choanoflagellata	class	phagotrophs	Sommeria-Klein et al. 2021
Chromodellids	division	phagotrophs	Sommeria-Klein et al. 2021
Chrysophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Chytridiomycota	class	parasites	Sommeria-Klein et al. 2021
Ciliophora	division	phagotrophs	Sommeria-Klein et al. 2021
Cnidaria	class	metazoans	Sommeria-Klein et al. 2021
Cryomonadida	class	phagotrophs	Sommeria-Klein et al. 2021
Cryptophyta	division	phototrophs	Sommeria-Klein et al. 2021
Ctenophora	class	metazoans	Sommeria-Klein et al. 2021
Dactylopodida	order	parasites	Sommeria-Klein et al. 2021
Dictyochophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Dinophyceae	class	dinoflagellates	Sommeria-Klein et al. 2021
Diplonemida	order	phagotrophs	Sommeria-Klein et al. 2021
Ebriida	order	phagotrophs	Sommeria-Klein et al. 2021
Echinodermata	class	metazoans	Sommeria-Klein et al. 2021
Eucyrtidium	genus	phagotrophs	Sommeria-Klein et al. 2021
Euglenida	class	phagotrophs	Sommeria-Klein et al. 2021

**Table S1:** *(continued)*

Taxon	Taxonomic level	Function	Reference
Foraminifera	division	phagotrophs	Sommeria-Klein et al. 2021
Haptophyta	division	phototrophs	Sommeria-Klein et al. 2021
Katablepharidophyta	division	phagotrophs	Sommeria-Klein et al. 2021
Kinetoplastea	class	parasites	Sommeria-Klein et al. 2021
Labyrinthulomycetes	class	phagotrophs	Sommeria-Klein et al. 2021
Dino-Group-I	order	parasites	Sommeria-Klein et al. 2021
Dino-Group-II	order	parasites	Sommeria-Klein et al. 2021
Dino-Group-III	order	parasites	Sommeria-Klein et al. 2021
Dino-Group-IV	order	parasites	Sommeria-Klein et al. 2021
Dino-Group-V	order	parasites	Sommeria-Klein et al. 2021
Mamiellophyceae	class	phototrophs	Sommeria-Klein et al. 2021
MAST-1	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-10	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-11	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-12	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-3	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-4	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-6	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-7	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-8	class	phagotrophs	Sommeria-Klein et al. 2021
MAST-9	class	phagotrophs	Sommeria-Klein et al. 2021
Mesomycetozoa	division	parasites	Sommeria-Klein et al. 2021
MOCH-1	class	phototrophs	Sommeria-Klein et al. 2021
MOCH-2	class	phototrophs	Sommeria-Klein et al. 2021
Mollusca	class	metazoans	Sommeria-Klein et al. 2021
Nassellaria	order	phagotrophs	Sommeria-Klein et al. 2021
Nemertea	class	metazoans	Sommeria-Klein et al. 2021
Oomycota	class	parasites	Sommeria-Klein et al. 2021
Pelagophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Platyhelminthes	class	metazoans	Sommeria-Klein et al. 2021
Porifera	class	metazoans	Sommeria-Klein et al. 2021
Pyramimonadophyceae	class	phototrophs	Sommeria-Klein et al. 2021
RAD-A	class	phagotrophs	Sommeria-Klein et al. 2021
RAD-B	class	phagotrophs	Sommeria-Klein et al. 2021
RAD-C	class	phagotrophs	Sommeria-Klein et al. 2021

**Table S1:** *(continued)*

Taxon	Taxonomic level	Function	Reference
Rhodophyta	division	phototrophs	Sommeria-Klein et al. 2021
Spumellaria	order	phagotrophs	Sommeria-Klein et al. 2021
Streptophyta	division	phototrophs	Sommeria-Klein et al. 2021
Telonemia	division	phagotrophs	Sommeria-Klein et al. 2021
Trebouxiophyceae	class	phototrophs	Sommeria-Klein et al. 2021
Vannellida	order	phagotrophs	Sommeria-Klein et al. 2021
Amoebozoa	supergroup	parasites	Literature
Perkinsea	division	parasites	Literature
Alveolata_X	division	parasites	Literature
Dinoflagellata	division	phagotrophs	Literature
Apusozoa	supergroup	phagotrophs	Literature
Chlorophyta	division	phototrophs	Literature
Glaucophyta	division	phototrophs	Literature
Prasinodermophyta	division	phototrophs	Literature
Centrohelioczoa	division	phagotrophs	Literature
Metamonada	division	parasites	Literature
Picozoa	division	phagotrophs	Literature
Choanoflagellida	division	phagotrophs	Literature
Fungi	division	parasites	Literature
Metazoa	division	metazoans	Literature
Cercozoa	division	phagotrophs	Literature
Aurearenophyceae	class	phototrophs	Literature
Chrysomerothyceae	class	phototrophs	Literature
Eustigmatophyceae	class	phototrophs	Literature
MOCH-3	class	phototrophs	Literature
MOCH-4	class	phototrophs	Literature
MOCH-5	class	phototrophs	Literature
Phaeophyceae	class	phototrophs	Literature
Ochrophyta	division	phototrophs	Literature
Pinguicophyceae	class	phototrophs	Literature
Raphidophyceae	class	phototrophs	Literature
Synchromophyceae	class	phototrophs	Literature
Synurophyceae	class	phototrophs	Literature
Xanthophyceae	class	phototrophs	Literature
MAST-16	class	phagotrophs	Literature

**Table S1:** *(continued)*

Taxon	Taxonomic level	Function	Reference
MAST-22	class	phagotrophs	Literature
MAST-24	class	phagotrophs	Literature
Opalozoa	division	parasites	Literature
Pseudofungi	division	parasites	Literature
Sagenista	division	phagotrophs	Literature
Stramenopiles	supergroup	phagotrophs	Literature
Discoba	division	parasites	Literature
Archaeplastida	supergroup	phototrophs	Literature
Eukaryota_X	supergroup	parasites	Literature
Malawimonadidae	division	parasites	Literature
Radiolaria	division	phagotrophs	Literature
Protalveolata	supergroup	phagotrophs	Literature

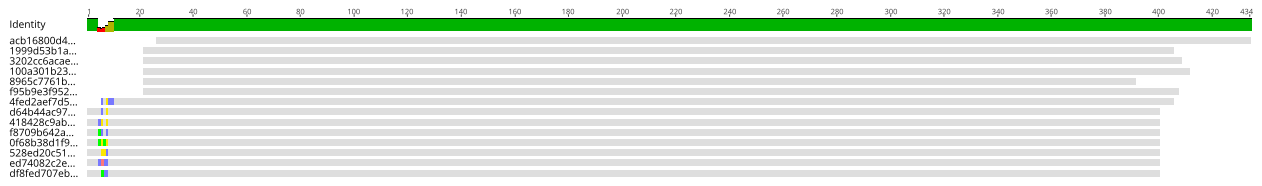
**Table S2:** Eukaryotic 18S rRNA primers used for metaPR2 datasets with the number of datasets (N) where used (Table 1).

Name	Sequence	Region	Direction	Reference	DOI	N
TAReuk454FWD1	CCAGCASCYCGCGTAATTCC	V4	fwd	Stoeck et al (2010)	10.1111/j.1365-294X.2009.04480.x	21
E572F	CYCGGTAATTCCAGCTC	V4	fwd	Comeau et al. (2011)	10.1371/journal.pone.0027492	7
3NDF	GGCAAGTCTGGTGCCAG	V4	fwd	Cavalier-Smith et al. (2009)	10.1016/j.protis.2009.03.003	2
528F	GCGGTAATTCAGTCCAA	V4	fwd	Cheung et al. (2010)	10.1038/ismej.2010.26	2
NSF573	CGCGTAATTCAGTCCA	V4	fwd	Mangot et al. (2013)	10.1111/1462-2920.12065	2
1380F	CCCTGCCHTTGTACACAC	V9	fwd	Amaral Zettler et al (2009)	10.1371/journal.pone.0006372	1
1389F	TTGTACACACCGCCC	V9	fwd	Amaral Zettler et al (2009)	10.1371/journal.pone.0006372	1
515F	GTGCCAGCMGCCGCGTAA	V4	fwd	Parfrey et al. (2014)	10.3389/fmicb.2014.00298	1
528F	CCGCGTAATTCAGCTC	V4	fwd	Zhu et al. (2005)	10.1016/j.femsec.2004.10.006	1
EK-565F	GCAGTAAAAAGCTCGTAGT	V4	fwd	Simon et al. (2015)	10.1111/1462-2920.12591	1
EuF-V4	CCAGCASCYCGGTAATWCC	V4	fwd	Boscaro et al. (2017)	10.1007/s00248-016-0912-8	1
EuF-V4	CCAGCASCYCGGTAATWCC	V4	fwd	Belevich et al. (2017)	10.1007/s00248-017-1076-x	1
TAReukFWD1	CCAGCASCYCGGTAAT	V4	fwd	Annenkova et al. (2020)	10.3390/microorganisms8040543	1
TAReukREV3	ACTTTCGTTCTTGATYRA	V4	rev	Stoeck et al (2010)	10.1111/j.1365-294X.2009.04480.x	15
V4 18S Next.Rev	ACTTTCGTTCTTGATYRATGA	V4	rev	Piredda et al. (2017)	10.1093/femsec/fiw200	7
E1009R	AYGGTATCTRATCRTCTTYG	V4	rev	Comeau et al. (2011)	10.1371/journal.pone.0027492	6
1055R	ACGGCCATGCACCACCCAT	V4	rev	Alves-de-Souza et al (2011)	10.5194/bg-8-2125-2011	2
1510R	CCTTCYGCAGTTCCACCTAC	V9	rev	Lopez-Garcia et al. (2003)	10.1073/pnas.0235779100	2
NSR951	TTGGYRAATGCTTTCGC	V4	rev	Mangot et al. (2013)	10.1111/1462-2920.12065	2
V4_euk_R2	ACGGTATCTRATCRTCTTCG	V4	rev	Brate et al. (2010)	10.1038/ismej.2010.39	2
1119r	GGTGCCCTCCGTC	V4	rev	Parfrey et al. (2014)	10.3389/fmicb.2014.00298	1
897R	TCYDAGAATTYCACCTCT	V4	rev	Hugerth et al. (2014)	10.1371/journal.pone.0095567	1
EUK1134-R	TTTAAGTTTCAGCCTTGCG	V4	rev	Carnegie et al. (2003)	10.3354/dao054219	1
Nex_18S_0964_R	GATCCCYAACTTTCGTTCTTGA	V4	rev	Kim et al. (2016)	10.1111/1462-2920.13523	1
picoR2	AKCCCCYAACTTTCGTTCTTGA	V4	rev	Belevich et al. (2017)	10.1007/s00248-017-1076-x	1

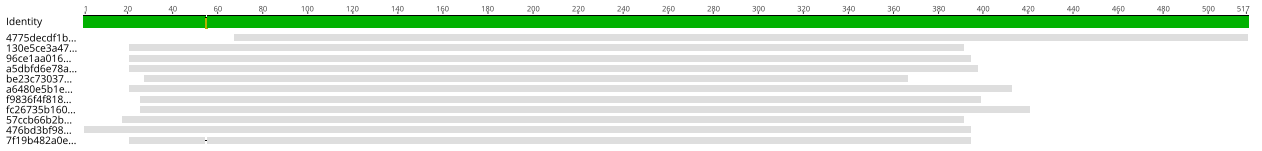
**Table S3:** 18S rRNA primer sets used for metaPR2 datasets with the number of datasets (N) where used (Table 1). Refer to Table S2 for sequence and reference of primers.

Primer fwd	Primer rev	Region	N
TAReuk454FWD1	TAReukREV3	V4	14
TAReuk454FWD1	V4 18S Next.Rev	V4	7
E572F	E1009R	V4	6
3NDF	V4_euk_R2	V4	2
528F	1055R	V4	2
NSF573	NSR951	V4	2
1380F	1510R	V9	1
1389F	1510R	V9	1
E572F	897R	V4	1
EK-565F	UNonMet	V4	1
EuF-V4	picoR2	V4	1
F515	R119	V4	1
Nex_18S_0587_F	Nex_18S_0964_R	V4	1
TAReukFWD1	TAReukREV3	V4	1

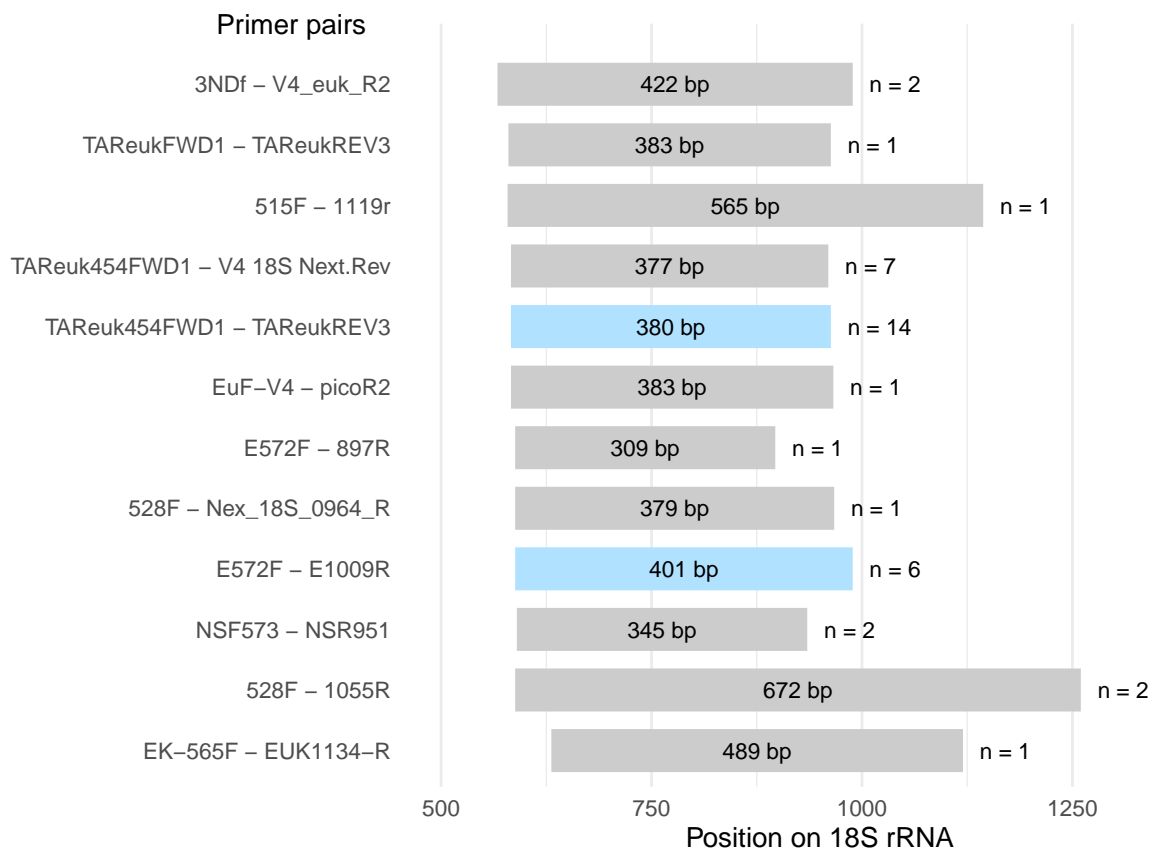
## A - MAST8



## B - Ciliate



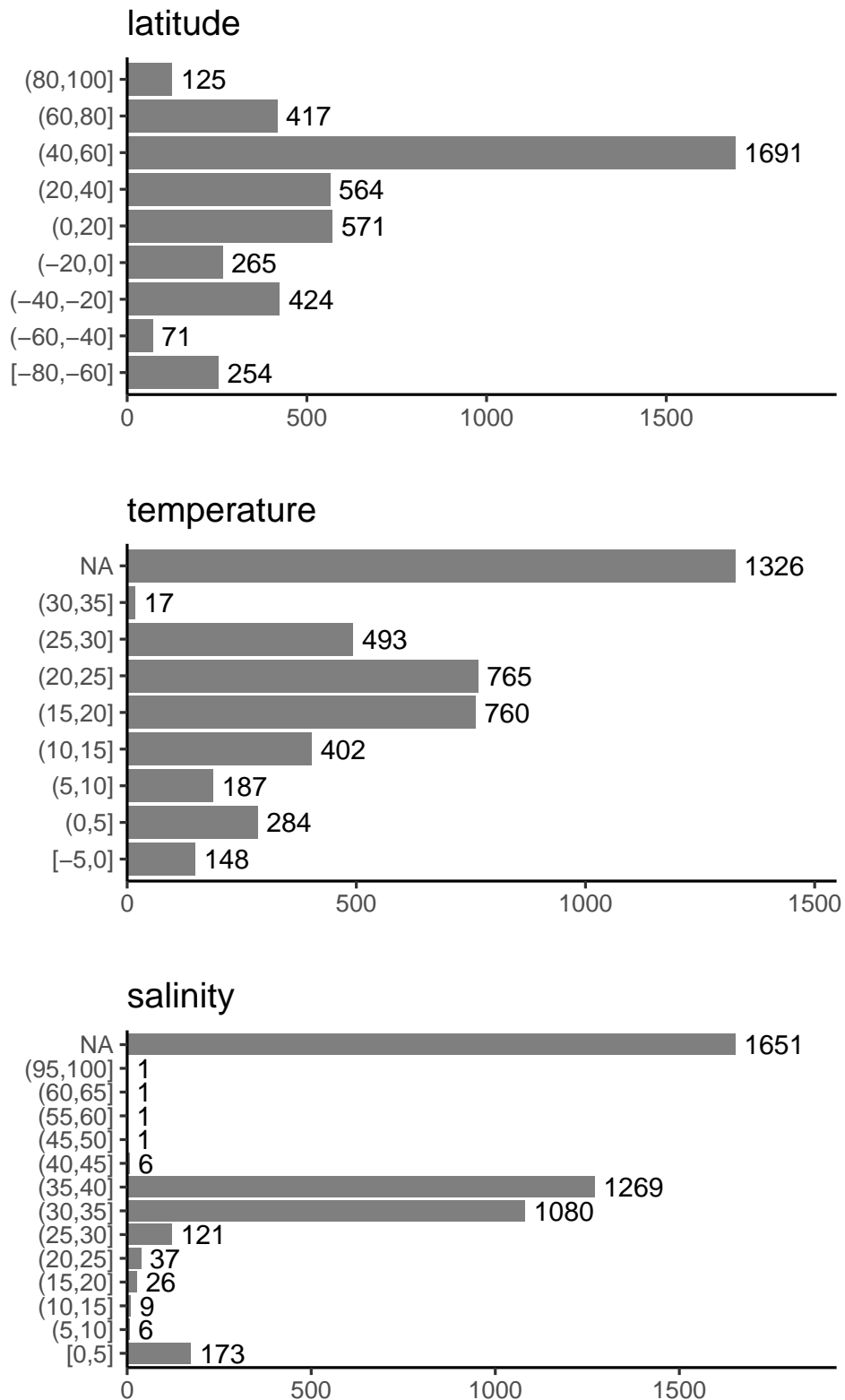
**Figure S1:** Two examples of V4 sequence clusters ASV (cASV) for Stramenopiles MAST 8 (A) and ciliates (B).



**Figure S2:** Amplicon size and position on the 18S rRNA gene (yeast), with the number of datasets for each V4 primer pair on the right side.

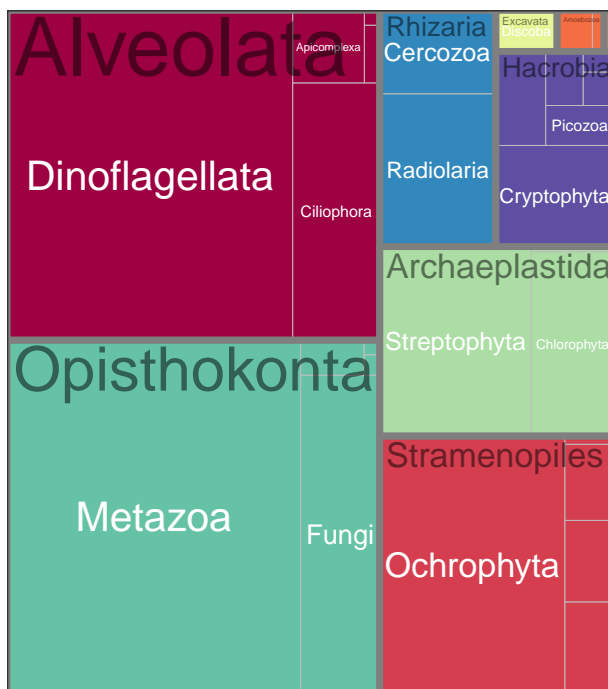


## Number of samples

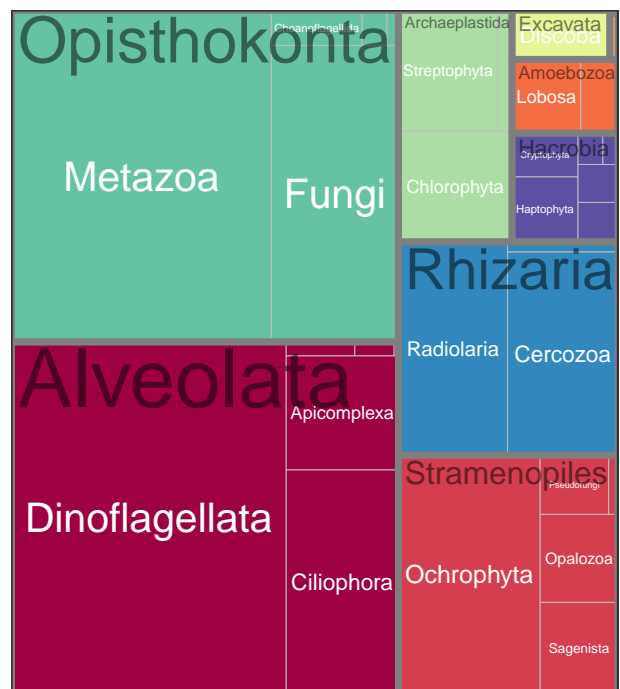


**Figure S3:** Distribution of samples by latitude, temperature and salinity ranges. NA corresponds to samples for which the data are not available.

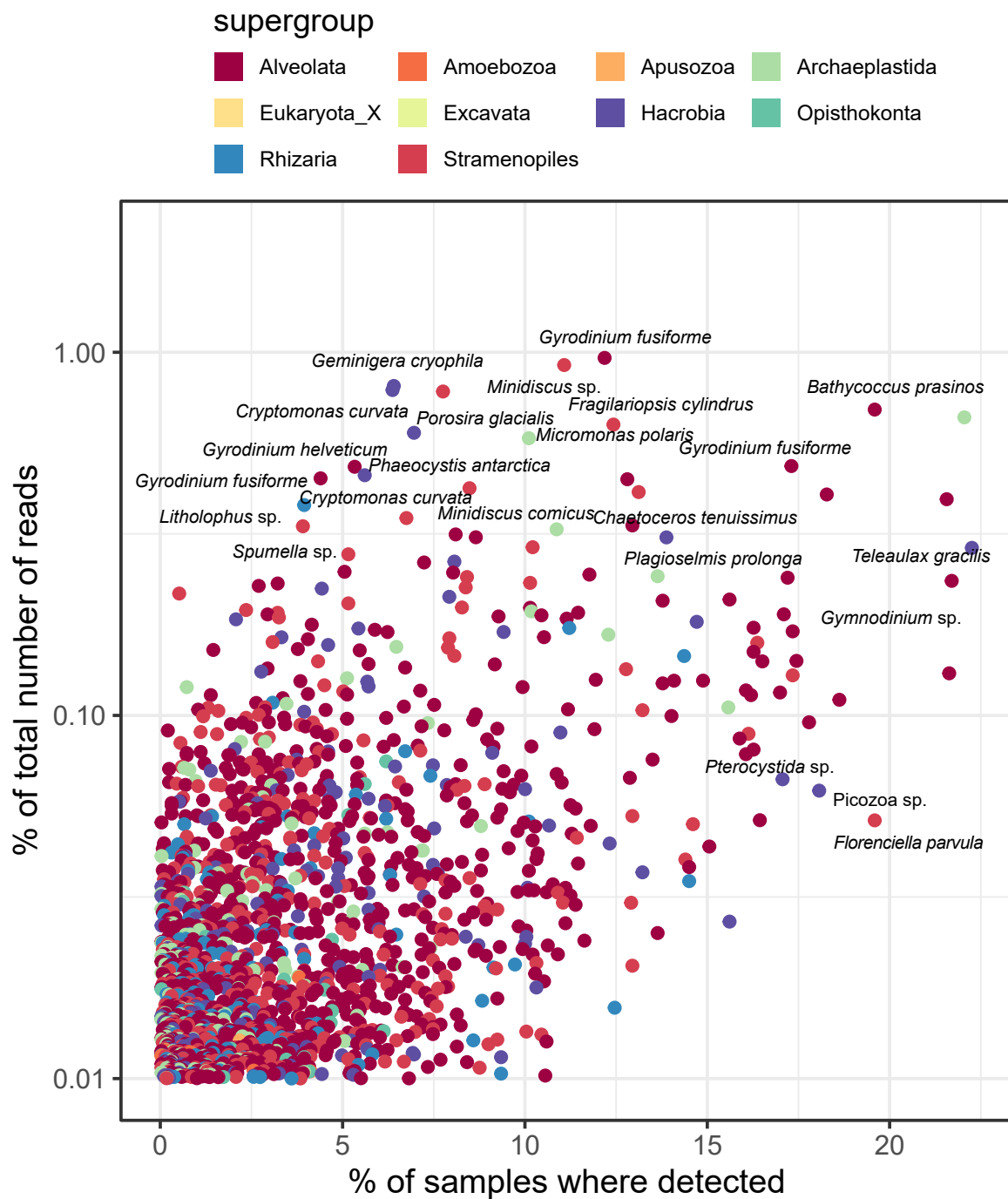
Reads



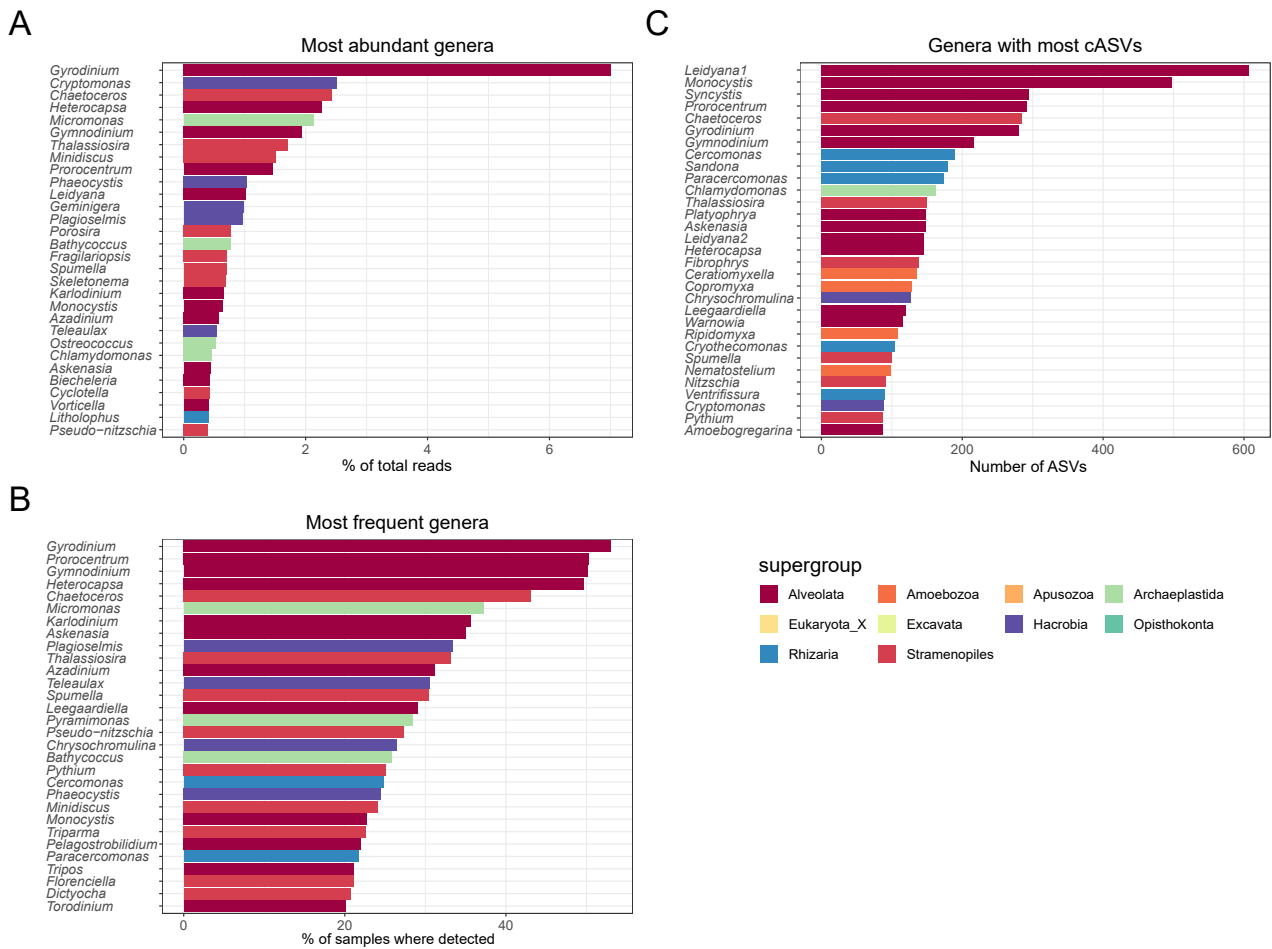
cASVs



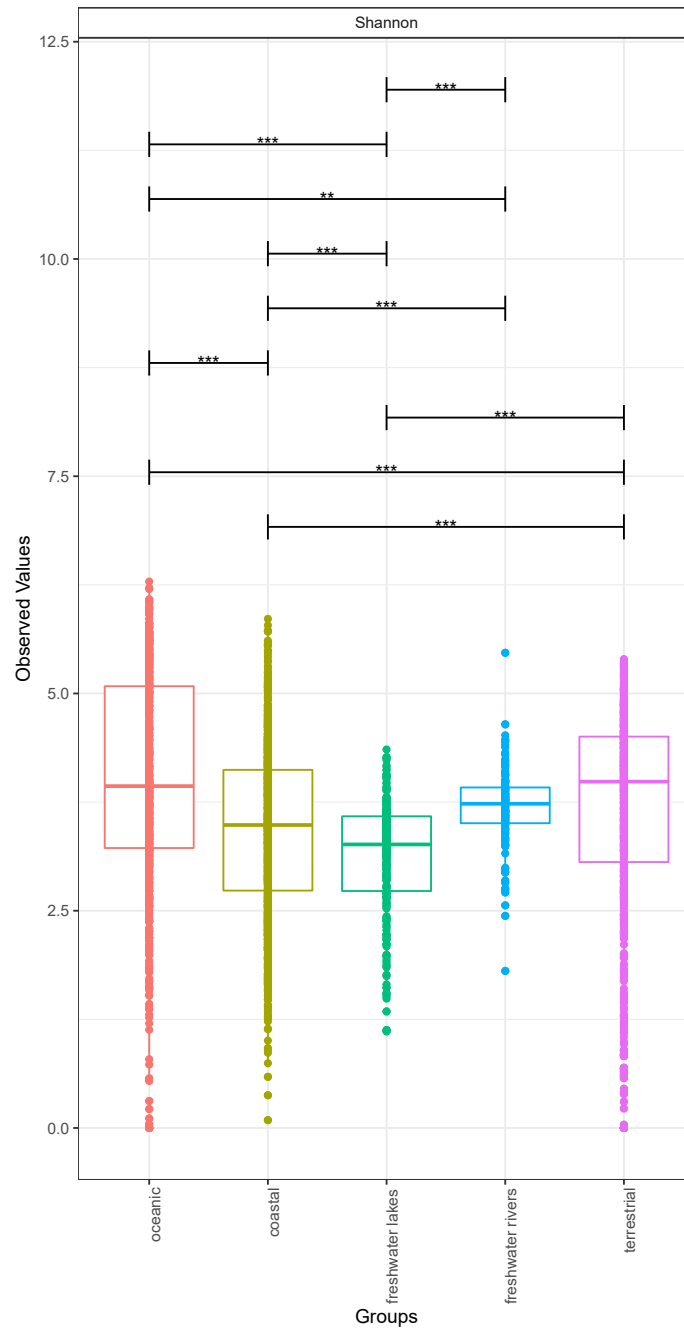
**Figure S4:** Treemaps of most abundant taxa (supergroup and division) for all datasets (V4 and V9) based on number of reads after normalization (left) or number of cASVs (right).



**Figure S5:** Protist V4 cASVs. Relationships between frequency and abundance of cASVs. Colours correspond to supergroups.



**Figure S6:** Protist genus analysis for the V4 dataset after normalization. A. Most abundant genera. B. Most frequent genera. C. Genera with most cASVs.



**Figure S7:** Protist V4 ASVs. Shannon's diversity index as a function of the environment with significance (\*\* p-value < 0.01, \*\*\* p-value < 0.001).

Datasets Treemap Map Barplot Diversity Query Download About

# metaPR2

DOI [10.5281/zenodo.6507135](https://doi.org/10.5281/zenodo.6507135) release **1.0.3** date 30 April 2022



## A database of 18S rRNA metabarcodes

Database version: 1.1 - 41 datasets

Shiny application version: 1.0.3

### Presentation

MetaPR2 is a database of published 18S rRNA metabarcodes that have been reprocessed and assigned using PR2.

### Accessing the database

Access to the database to map, search and download the barcodes can be done in three different ways:

- Using a [web interface](#).
- Download the R package and launch the shiny application.

#### 1 - Web interface

- Launch in your browser: <http://shiny.metapr2.org>
- Help : <https://pr2database.github.io/metapr2-shiny/articles/>

#### 2 - metaPR2 shiny R package

- Install the package from GitHub and launch function `metapr2App()`

```
install.packages(devtools)

devtools::install_github("pr2database/metapr2-shiny")
metapr2::metapr2App()
```

Note: You may have to install some packages required by metapr2 if they are not installed on your machine

### Help

Extensive help is provided [here](#).

### Errors

Please report errors in the [Issues page](#) of the metaPR2 database.

### Citation

Vaulot, D., Sim, C.W.H., Ong, D., Teo, B., Biver, C., Jamy, M., Lopes dos Santos, A., 2022. metaPR2: a database of eukaryotic 18S rRNA metabarcodes with an emphasis on protists. *BioRxiv*. <https://doi.org/10.1101/2022.02.04.479133>

### Resources

- Website: <https://shiny.metapr2.org/>
- Docker: <https://hub.docker.com/repository/docker/vaulot/metapr2>
- Source code: <https://github.com/pr2database/metapr2-shiny>

### Maintainer

- Daniel Vaulot: [vaulot@gmail.com](mailto:vaulot@gmail.com)

### Contributors

- Daniel Vaulot, CNRS Roscoff, NTU-ASE Geek lab
- Adriana Lopes dos Santos, NTU-ASE Geek lab
- Clarence Sim, NTU-ASE Geek lab
- Denise Ong, NTU-ASE Geek lab
- Bryan Teo, NTU-ASE Geek lab
- Charlie Biver, Uppsala University Sweden
- Mahwash Jamy, Uppsala University Sweden

Figure S8: Shiny panel "About".



The metaPR2 database

Select datasets

41 items selected

Select All Deselect All

- Antarctic\_Fieldees\_Bay\_2013 ✓
- Antarctic\_Fieldees\_Bay\_2015\_18S\_V4 ✓
- Antarctic\_Fieldees\_Bay\_2015\_18S\_V4\_sorted ✓
- Arctic\_Baffin\_Bay\_2013 ✓
- Arctic\_Beaufort\_Sea\_MALINA\_2014 ✓
- Arctic\_Nansen\_Basin\_2012 ✓
- Arctic\_Nares\_Strait\_2014 ✓
- Arctic\_Ocean\_Central\_2012 ✓
- Arctic\_Ocean\_PS80\_2012 ✓
- Arctic\_Ocean\_Survey\_2005\_2011 ✓
- Arctic\_White\_Sea\_2013\_2015 ✓
- Baltic\_Sea\_2012\_2013 ✓
- Baltic\_Sea\_Gdansk\_2012 ✓
- Chukchi\_Sea\_IGESCAPE\_2010 ✓
- European\_coast\_Biomarkers\_2009 ✓
- Italy\_Naples\_2011 ✓
- Lake\_Baikal\_2013 ✓
- Lake\_Chaohu\_2014\_2015 ✓
- Lake\_Chevreuse\_2012 ✓
- Lake\_Fuxian\_2015 ✓
- Lake\_Garda ✓
- Lakes\_Argentina ✓
- Lakes\_mountain\_2013 ✓
- Lakes\_Scandinavia ✓
- Malaspina\_surface\_2010\_2011 ✓
- Malaspina\_vertical\_2010\_2011 ✓
- Mariana\_Trench\_2016\_1 ✓
- Mariana\_Trench\_2016\_2 ✓
- Norway\_Oslo\_fjord\_2009\_2011 ✓
- OSD\_2014\_V4\_LGC ✓
- OSD\_2014\_V4\_LW ✓
- OSD\_2015\_V4 ✓
- River\_Parana ✓
- River\_Saint\_Charles\_2016\_2017 ✓
- Soils\_Global\_2012 ✓
- Soils\_Neotropical ✓
- Soils\_Swiss ✓
- Spain\_Blanes\_2004\_2013 ✓
- Tara\_Arctic\_V4 ✓
- Tara\_Ocean\_V4 ✓
- Tara\_Oceans\_V9 ✓

All

About Datasets Treemap Map Barplot Diversity Query Download Help

Quick dataset selection.

Dataset groups

marine\_global\_V4  oceanic  coastal  rivers  lakes  soils  arctic  antarctic  temperate  tropical  time series


Show 10 entries Search:

dataset_id	dataset_name	region	paper_reference	sample_number	asv_number	n_reads_mean	selected
11	Antarctic Fieldees Bay- 2013	Southern Ocean	Luo, W. et al. Molecular diversity of microbial eukaryotes in sea water from Fildes Peninsula, King George Island, Antarctica. <i>Polar Biol.</i> (2015)	10	69	13631	true
16	Antarctic Fieldees Bay 2015 18S V4	Southern Ocean	Trefault, N., De la Iglesia, R., Moreno-Pino, M., Lopes dos Santos, A., G-e9-rikas Ribeiro, C., Parada-Pozo, G., Cristi, A., Marie, D., & Vaulot, D. (2021). Annual phytoplankton dynamics in coastal waters from Fildes Bay, Western Antarctic Peninsula. <i>Scientific Reports</i> , 11(1), 1368.	123	685	48261	true
18	Antarctic Fieldees Bay 2015 18S V4 sorted	Southern Ocean	Trefault, N., De la Iglesia, R., Moreno-Pino, M., Lopes dos Santos, A., G-e9-rikas Ribeiro, C., Parada-Pozo, G., Cristi, A., Marie, D., & Vaulot, D. (2021). Annual phytoplankton dynamics in coastal waters from Fildes Bay, Western Antarctic Peninsula. <i>Scientific Reports</i> , 11(1), 1368.	60	280	31615	true
9	Arctic Nansen Basin - 2012	Arctic Ocean	Metties, K., von Appen, W.-J., Killas, E., Nicolaus, A. & N-16>thg, E.-M. Biogeography and Photosynthetic Biomass of Arctic Marine Pico-Eukaryotes during Summer of the Record Sea Ice Minimum 2012. <i>PLoS One</i> 11, 20 pp. (2016)	17	328	13700	true
42	Arctic Nares Strait - 2014	Arctic Ocean	Kalenitchenko D., Joli N., Potvin M., Tremblay J.-c&e9-, Lovejoy C. 2019. Biodiversity and Species Change in the Arctic Ocean: A View Through the Lens of Nares Strait. <i>Frontiers in Marine Science</i> 6:1-96>17.	247	1510	38626	true
6	Arctic Ocean Central - 2012	Arctic Ocean	Stecher, A., Neuhaus, S., Lange, B., Frickenhaus, S., Beszteri, B., Kroth, P.G. & Valentin, K. 2015. rRNA and rDNA based assessment of sea ice protist biodiversity from the central Arctic Ocean. <i>Eur. J. Phycol.</i> 1-96>16.	8	182	36628	true
40	Arctic Ocean Survey - 2005- 2011	Arctic Ocean	Thaler M., Lovejoy C., Sea B. 2015. Biogeography of Heterotrophic Flagellate Populations Indicates the Presence of Generalist and Specialist Taxa in the Arctic Ocean. <i>Applied and Environmental Microbiology</i> 81.2137-96>2148	36	467	7136	true
5	Arctic Ocean, Beaufort Sea, MALINA cruise - 2009	Arctic Ocean	Monier, A., Terrado, R., Thaler, M., Comeau, A., Medrinal, E. & Lovejoy, C. 2013. Upper Arctic Ocean water masses harbor distinct communities of heterotrophic flagellates. <i>Biogeosciences</i> . 10:4273-96>88. Monier, A., Comte, J., Babin, M., Forest, A., Matsuoka, A. & Lovejoy, C. 2014. Oceanographic structure drives the assembly processes of microbial eukaryotic communities. <i>ISME J.</i> 9:990-96>1002	24	270	6704	true
39	Arctic Polarstern expedition ARK-XXVII/3 - 2012	Arctic Ocean	Rapp J.Z., Fern-s1>ndez, M-e9>ndez M., Bienhold C., Boetius A. 2018. Effects of Ice-Algal Aggregate Export on the Connectivity of Bacterial Communities in the Central Arctic Ocean. <i>Frontiers in Microbiology</i> 9:1035	45	978	73833	true
38	Arctic White Sea - 2013-2015	Arctic Ocean	Belevich TA., Ilyash L.V., Milyutina IA., Logacheva MD., Goryunov D.V., Troitsky A.V. 2017. Photosynthetic Picoeukaryotes in the Land-Fast Ice of the White Sea, Russia. <i>Microbial Ecology</i> . 1-96>16.	17	380	23990	true

Showing 1 to 10 of 41 entries

Previous 1 2 3 4 5 Next

Figure S9: Shiny panel "Datasets".



The metaPR2 database

Select datasets

41 items selected

Select Samples

Gene regions

V4

DNA or RNA

DNA

Ecosystems

oceanic, coastal, freshwater lakes, freshwater riv

Substrates

water

Size fractions

pico, total

Depth levels

surface

Select ASVs

Minimum number of total reads per ASV

1000

Select Taxa

Supergroup

All

Division

All

Class

All

Order

All

Family

All

Genus

All

Species

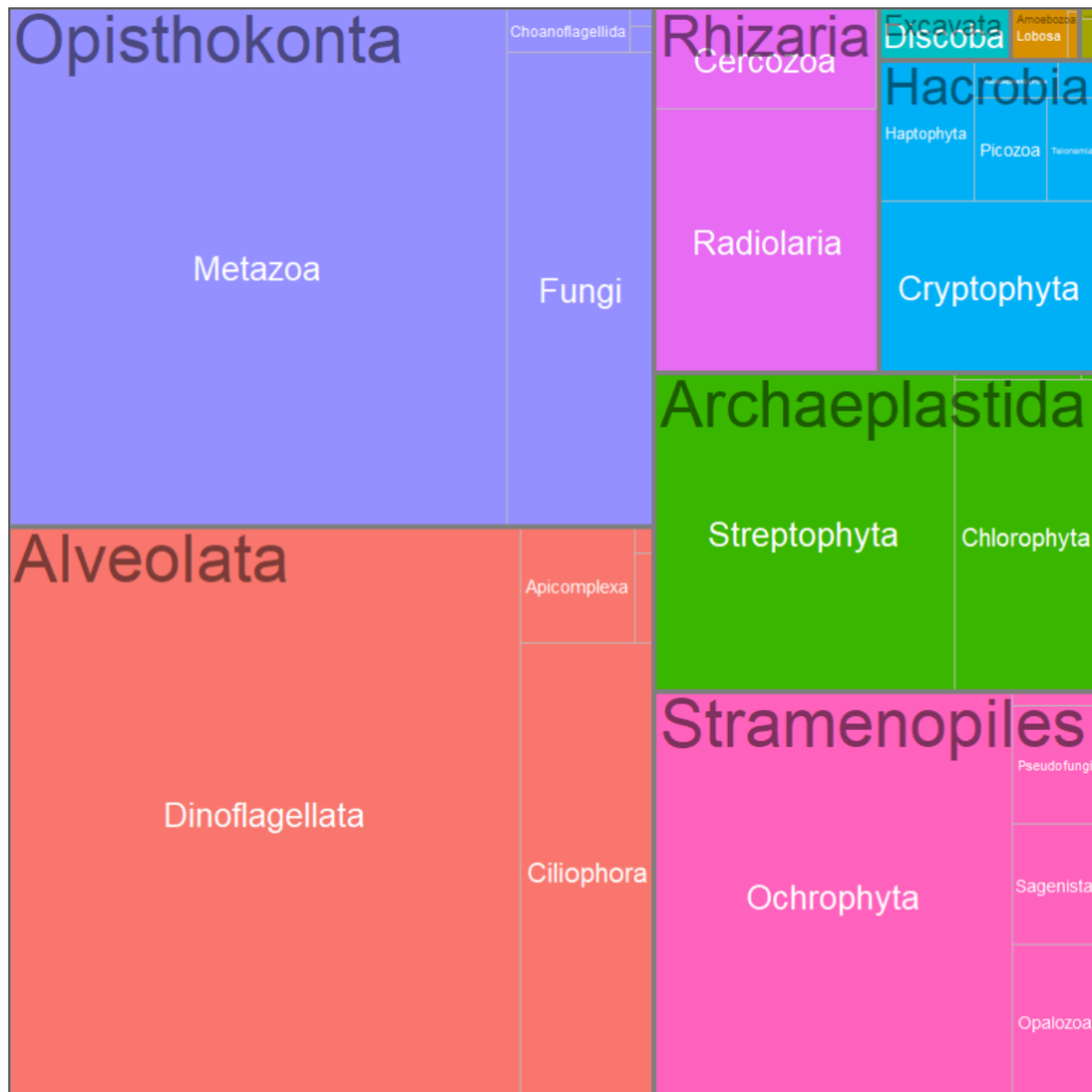
All

Figure S10: Shiny sample and taxonomy selection sidebar.



[About](#) [Datasets](#) [Treemap](#) [Map](#) [Barplot](#) [Alpha diversity](#) [Beta diversity](#) [Query](#) [Download](#)

Number of reads have been normalized (not rarefield) to 100 with 3 decimals.



**Figure S11:** Shiny panel "Treemap".



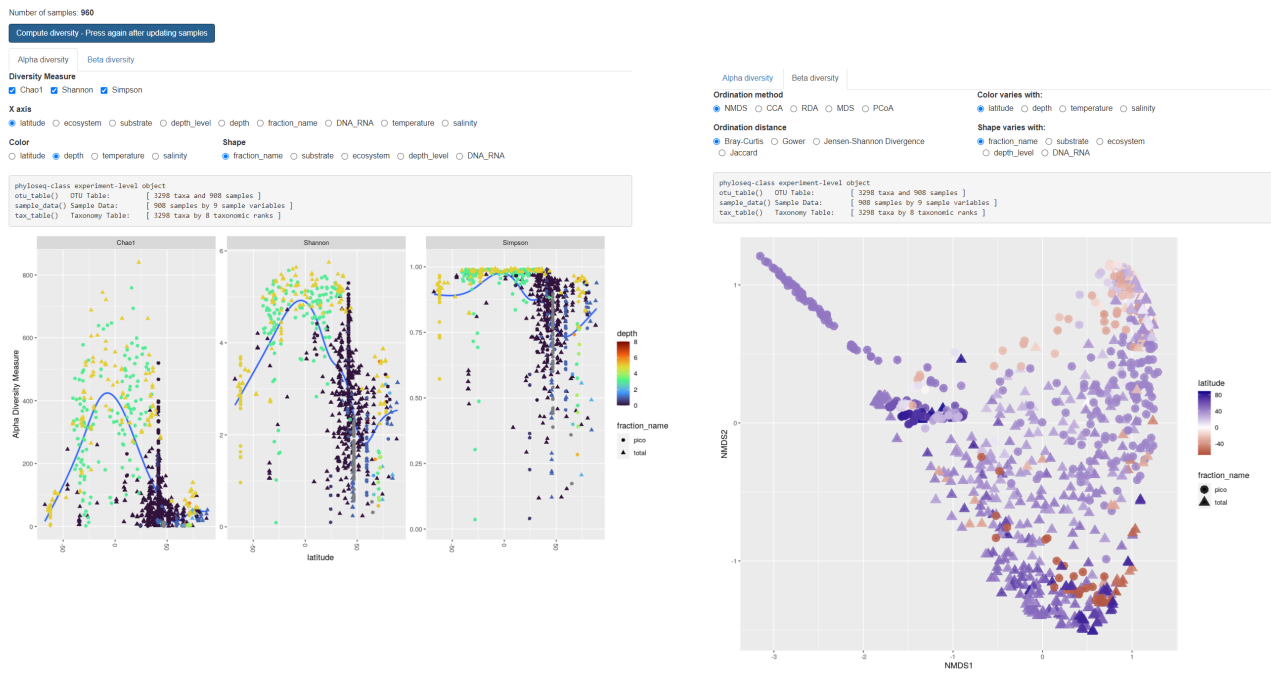


Figure S14: Shiny panel "Diversity".

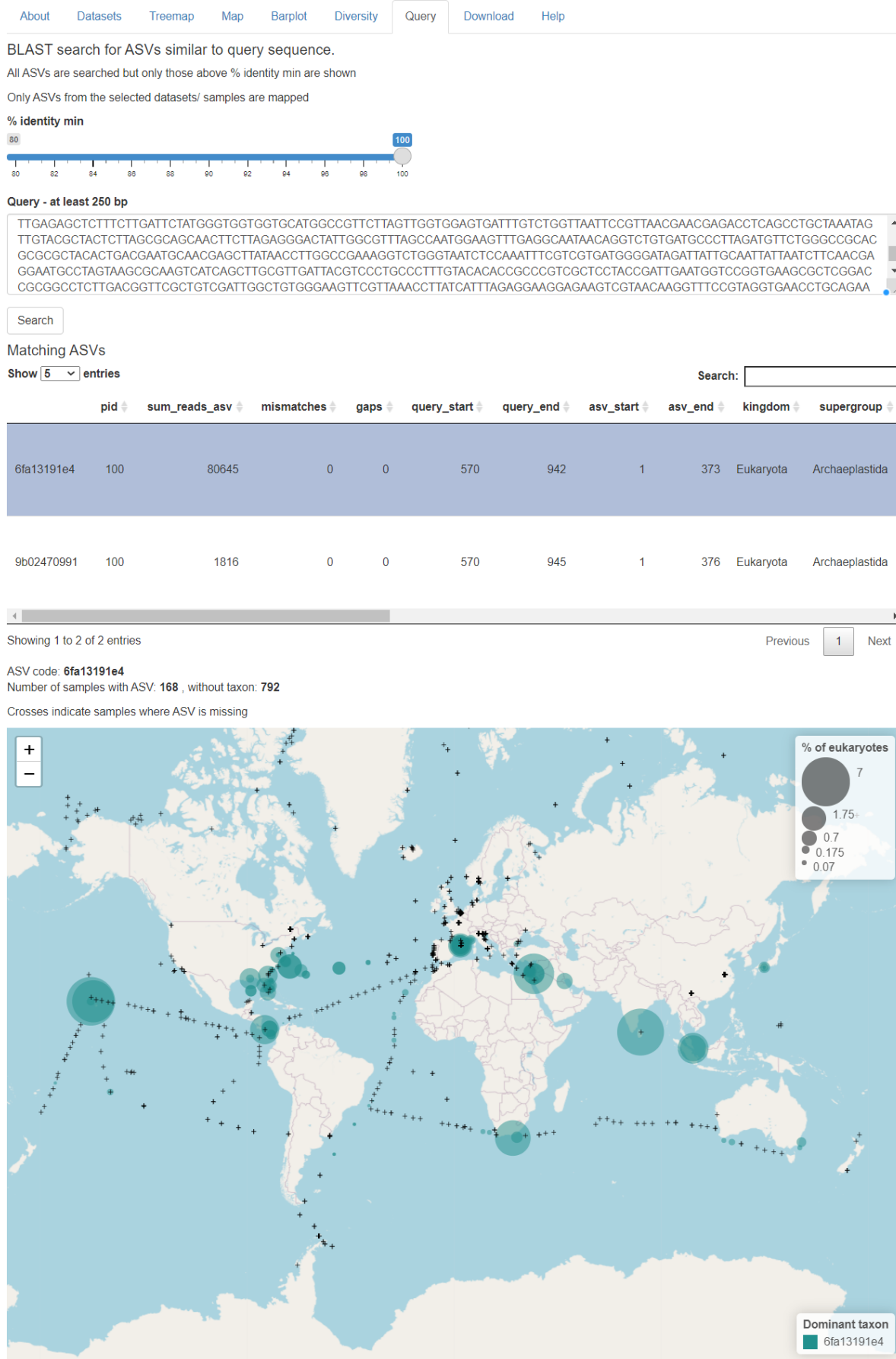


Figure S15: Shiny panel "Query".

[About](#)
[Datasets](#)
[Treemap](#)
[Map](#)
[Barplot](#)
[Diversity](#)
[Query](#)
[Download](#)
[Help](#)

### Selected damples, datasets and taxa

The following files are provided. They can be linked by key fields. They only contain the selected datasets, sample type and taxa. **The asv\_reads and phyloseq files can be very big** if you download all datasets and all taxa. Phyloseq files will only be created for less than 1000 samples selected.

file	content	key fields
datasets.xlsx	Information on the different datasets selected including reference and GenBank id	dataset_id
samples.xlsx	List of samples selected with medadata	file_code
asv.xlsx	ASV selected with taxonomy and sequence	asv_code
asv_reads.tsv.gz	Percent of reads (normalized to total number of eukaryotic reads in the sample), for each ASV and each sample (long form).	asv_code, file_code
phyloseq.rds	File to use with phyloseq R package ( <a href="https://joey711.github.io/phyloseq/">https://joey711.github.io/phyloseq/</a> ). Use readRDS() function to read	<b>5000 samples max</b>

Number of samples: **960**

[Download datasets, samples and ASVs \(zip\)](#)
[Download ASVs abundance \(tsv.gz\)](#)

```

Make phyloseq done
phyloseq-class experiment-level object
otu_table() OTU Table: [ 3298 taxa and 908 samples ]
sample_data() Sample Data: [ 908 samples by 9 sample variables ]
tax_table() Taxonomy Table: [ 3298 taxa by 8 taxonomic ranks ]
phyloseq-class experiment-level object
otu_table() OTU Table: [ 3298 taxa and 908 samples ]
sample_data() Sample Data: [ 908 samples by 9 sample variables ]
tax_table() Taxonomy Table: [ 3298 taxa by 8 taxonomic ranks ]

```

[Download phyloseq file \(rds\)](#)

**Figure S16:** Shiny panel "Download".