

Novel diversity within marine Mamiellophyceae (Chlorophyta) unveiled by metabarcoding

Supplementary material

Margot Tragin

Daniel Vaultot

Sorbonne Université, CNRS, UMR 7144

Station Biologique, Place Georges Teissier, 29680 Roscoff, France

Submitted to **Scientific Report**

Date: October 21, 2018

All supplementary data have been deposited to

Figshare: <https://figshare.com/s/6e074ed24b01d3d1f8ea>

GitHub : https://github.com/vaultot/Paper-2018_Tragin_Mamiellophyceae-R-scripts

1 **Supplementary data**

2 Supplementary Data S1: (OSD_mamiello_mothur.sh) Mothur script for sequence analysis.

3 Supplementary Data S2: (OSD_Mamiello.html). R script for generating distribution maps and
4 heatmaps with output.

5 Supplementary Data S3: (Excel file OSD_Mamiello_Tables Supplementary.xlsx). Description of
6 the different sheets:

- 7 • "metadata": list of OSD stations with coordinates and metadata.
- 8 • "samples": list of OSD samples.
- 9 • " otus LGC Mamiello": Mamiellophyceae LGC ASV table with taxonomy and number of
10 reads assigned to each ASV per OSD 2014 samples.
- 11 • " otus LW Mamiello": Mamiellophyceae LW ASV table with taxonomy and number of reads
12 assigned to each ASV per OSD 2014 samples.
- 13 • "sequences GenBank": Genbank accession numbers for sequences used to build the
14 trees.
- 15 • "simil_Ostreococcus": Matrix of the pairwise percent identity between *Ostreococcus* clades.
16 The calculation was done based on the alignment available as Supplementary data.
- 17 • "simil_Micromonas": idem for *Micromonas*.
- 18 • "simil_Mantoniella": idem for *Mantoniella*.

19 Supplementary Data S4: (OSD_Mamiello_ASV_LGC.fasta) OSD 2014 Mamiellophyceae ASV
20 representative sequences (LGC sequencing).

21 Supplementary Data S5: (OSD_Mamiello_ASV_LW.fasta) OSD 2014 Mamiellophyceae ASV
22 representative sequences (LW sequencing).

23 Supplementary Data S6: (OSD_Mamiello_ASV_LGC_Ostreococcus_align.fasta) Alignment of
24 major LGC and LW *Ostreococcus* ASV with related GenBank sequences from clones and
25 cultures.

26 Supplementary Data S7: (OSD_Mamiello_ASV_LGC_Micromonas_align.fasta) Idem for
27 *Micromonas*.

28 Supplementary Data S8: (OSD_Mamiello_ASV_LGC_Mantoniella_align.fasta) Idem for
29 *Mantoniella*.

30 **List of supplementary tables**

31 Table S1 List of OSD stations.

32 Table S2 Matrix of the pairwise percent identity between *Ostreococcus* clades. The calcula-
33 tion was done based on the alignment available as Supplementary data S6.

34 Table S3 Matrix of the pairwise percent identity between *Micromonas* clades. The calculation
35 was done based on the alignment available as Supplementary data S7.

36 Table S4 Matrix of the pairwise percent identity between *Mantoniella* clades. The calculation
37 was done based on the alignment available as Supplementary data S8.

38 **List of supplementary figures**

39 Fig. S1 A. Percentage of OSD samples where a given ASV was detected (at least one read).
40 Numbers at right of bars correspond to number of samples. B. Idem but for samples
41 where the ASV contributed more than 1% of the Mamiellophyceae reads. Stations
42 with less than 100 Mamiellophyceae reads were not considered. The three major *O.*
43 *tauri* ASVs have been pooled together.

44 Fig. S2 Heatmap of the Mamiellophyceae communities in OSD samples. The number
45 of read corresponding to the major ASVs were normalized by the number of
46 Mamiellophyceae reads per station. Stations with less than 100 Mamiellophyceae
47 reads were not considered.

Fig. S3 Distribution of major *Ostreococcus* ASVs for OSD2014 (LGC) in European waters. The three major *O. tauri* ASVs have been pooled together. Circle surface corresponds to the percentage of ASV reads relative to the total number of Mamielophyceae reads. Samples for which the ASV contribution was lower than 1 % are represented by blue crosses.

Fig. S4 Alignments of *Ostreococcus* V4 regions and localization of the qPCR primers and probes (Demir-Hilton et al. 2011) used to quantify the OI (A) and OII (B) clades. Sequences with no mismatches to each probe set are in bold. Blue and green sequences correspond to LGC and LW datasets, respectively.

Fig. S5 Same as Fig. S3 for *Micromonas*.

Fig. S6 Same as Fig. S3 for *Bathycoccus*.

Fig. S7 Same as Fig. S3 for *Mantoniella*.

Table S1. List of OSD stations.

Station	Name	Ocean	Country
OSD1	Plymouth - L4	Atlantic Ocean	UK
OSD2	Roscoff - SOMLIT	Atlantic Ocean	France
OSD3	Helgoland	Atlantic Ocean	Germany
OSD4	LTER-MC	Mediterranean Sea	Italy
OSD5	Crete	Mediterranean Sea	Greece
OSD6	Blanes	Mediterranean Sea	Spain
OSD7	Moorea - Tiahura	Pacific Ocean	Tahiti
OSD8	BATS	Atlantic Ocean	Bermuda
OSD10	Lake Erie W4		USA
OSD13	Varna Bay	Black Sea	Bulgaria
OSD14	Banyuls	Mediterranean Sea	France
OSD15	Villefranche - SOMLIT	Mediterranean Sea	France
OSD17	VLIZ	Atlantic Ocean	Belgium
OSD18	Kyrenia	Mediterranean Sea	Turkish Republic of Northern
OSD19	Famagusta	Mediterranean Sea	Turkish Republic of Northern
OSD20	Faxaflói	Atlantic Ocean	Iceland
OSD21	Croatia	Mediterranean Sea	Croatia
OSD22	Marseille Solemio SOMLIT	Mediterranean Sea	France
OSD24	Marchica	Mediterranean Sea	Morocco
OSD25	Saidia Rocher	Mediterranean Sea	Morocco
OSD26	Tangier	Atlantic Ocean	Morocco
OSD28	Belize	Atlantic Ocean	Belize
OSD29	Florida	Atlantic Ocean	USA
OSD30	Tvärminne	Atlantic Ocean	Finland
OSD34	Alexandria	Mediterranean Sea	Egypt

OSD35	Cheasapeake Bay	Atlantic Ocean	USA
OSD36	Delaware	Atlantic Ocean	USA
OSD37	Port Everglades	Atlantic Ocean	USA
OSD38	Long Key	Atlantic Ocean	USA
OSD39	Charleston Harbor	Atlantic Ocean	USA
OSD41	Sequim Bay Park	Pacific Ocean	Alaska
OSD42	Faro Lake	Mediterranean Sea	Italy
OSD43	SIO Pier	Pacific Ocean	USA
OSD45	Tampa Bay	Atlantic Ocean	USA
OSD46	Horn Island	Atlantic Ocean	USA
OSD47	Venice Lagoon	Mediterranean Sea	Italy
OSD48	Venice Gulf	Mediterranean Sea	Italy
OSD49	Vida	Mediterranean Sea	Slovenia
OSD50	Pasaia	Atlantic Ocean	Spain
OSD51	Bocas del Toro	Atlantic Ocean	USA
OSD52	Abu Hashish	Mediterranean Sea	Egypt
OSD53	Ras Disha	Mediterranean Sea	Egypt
OSD54	Maine Booth Bay	Atlantic Ocean	USA
OSD55	Maine Damariscotta River	Atlantic Ocean	USA
OSD56	Hawaii Kakaako	Pacific Ocean	Hawaii
OSD57	Hawaii Oahu	Pacific Ocean	Hawaii
OSD58	PICO	Atlantic Ocean	USA
OSD60	South Carolina 2 - North Inlet	Atlantic Ocean	USA
OSD61	Vineyard Sound	Atlantic Ocean	USA
OSD62	Manai Straits	Atlantic Ocean	UK
OSD63	Venice Acqua Alta	Mediterranean Sea	Italy
OSD64	Odessa	Black Sea	Ukraine
OSD65	Leigh Marine Laboratory	Pacific Ocean	New Zealand

OSD69	Marghera	Mediterranean Sea	Italy
OSD70	Lido	Mediterranean Sea	Italy
OSD71	Otago	Pacific Ocean	New Zealand
OSD72	Boknis Eck	Atlantic Ocean	Germany
OSD73	Lima Estuary	Atlantic Ocean	Portugal
OSD74	Douro Estuary	Atlantic Ocean	Portugal
OSD76	Foglia	Mediterranean Sea	Italy
OSD77	Metauro	Mediterranean Sea	Italy
OSD78	CONISMA	Mediterranean Sea	Italy
OSD80	Young Sound	Arctic Ocean	Greenland
OSD81	Ria Formosa Lagoon	Atlantic Ocean	Portugal
OSD90	Etoliko Lagoon	Mediterranean Sea	Greece
OSD91	Oualidiya	Atlantic Ocean	Morocco
OSD92	Casablanca	Atlantic Ocean	Morocco
OSD93	Eljadida	Atlantic Ocean	Morocco
OSD94	Saidia Marina	Mediterranean Sea	Morocco
OSD95	Singapore Indigo_V	Pacific Ocean	Singapore
OSD96	Sao Miguel Azores I	Atlantic Ocean	Portugal
OSD97	Faial Azores	Atlantic Ocean	Portugal
OSD98	Sao Jorge Azores	Atlantic Ocean	Portugal
OSD99	C1	Mediterranean Sea	Italy
OSD100	Crete - GOS	Mediterranean Sea	Greece
OSD101	Quinta do Lorde	Atlantic Ocean	Portugal
OSD102	Marina do Funchal	Atlantic Ocean	Portugal
OSD103	Porto da Cruz	Atlantic Ocean	Portugal
OSD105	Cambridge Bay, Nunavut	Arctic Ocean	USA
OSD106	REYKIS	Atlantic Ocean	Iceland
OSD107	Lisboa	Atlantic Ocean	Portugal

OSD108	Alcochete	Atlantic Ocean	Portugal
OSD109	Rosario	Atlantic Ocean	Portugal
OSD110	Figueira da Foz	Atlantic Ocean	Portugal
OSD111	Ria de Aveiro_1	Atlantic Ocean	Portugal
OSD113	CascaisWatch	North Atlantic Ocean	Portugal
OSD114	BerlengasWatch	Atlantic Ocean	Portugal
OSD115	Santa Cruz	Atlantic Ocean	Portugal
OSD116	Lagoa de Obidos	Atlantic Ocean	Portugal
OSD117	Tavira Beach	Atlantic Ocean	Portugal
OSD118	Lough Hyne	Atlantic Ocean	Ireland
OSD122	Station A Gulf Of Eilat	Red Sea	Israel
OSD123	Shikmona	Mediterranean Sea	Israel
OSD124	Osaka Bay	Pacific Ocean	Japan
OSD125	Cullercoats Beach	Arctic Ocean	UK
OSD126	Eyafjordur_1	Arctic Ocean	Iceland
OSD128	Eyafjordur_3	Arctic Ocean	Iceland
OSD129	Eyafjordur_4	Arctic Ocean	Iceland
OSD130	Eyafjordur_5	Arctic Ocean	Iceland
OSD131	Zlatna ribka	Black Sea	Bulgaria
OSD132	Sdot YAM	Mediterranean Sea	Israel
OSD133	Robben Island	Atlantic Ocean	South Africa
OSD141	Raunefjorden	Atlantic Ocean	Norway
OSD142	Gray's Reef National Marine Sanctuary	Atlantic Ocean	USA
OSD143	Skidaway Institute of Oceanography	Atlantic Ocean	USA
OSD144	Maunalua Bay O'ahu	Pacific Ocean	USA
OSD145	Blankenberge	Atlantic Ocean	Belgium
OSD146	Fram Strait	Arctic Ocean	Germany
OSD147	Rajarata	Indian Ocean	Sri Lanka

OSD148	Wadden Sea	Atlantic Ocean	Germany
OSD149	Laguna Rocha Norte	Atlantic Ocean	Uruguay
OSD150	Laguna Rocha Sur	Atlantic Ocean	Uruguay
OSD151	South Atlantic Microbial Observatory	Atlantic Ocean	Uruguay
OSD152	Compass Buoy Station	Atlantic Ocean	Canada
OSD153	Faro Island	Atlantic Ocean	Portugal
OSD154	Arcachon-SOMLIT	Atlantic Ocean	France
OSD155	Steilene Oslofjord	Atlantic Ocean	Norway
OSD156	Hvaler Tisler Site	Atlantic Ocean	Norway
OSD157	ELLEIm2	Atlantic Ocean	Norway
OSD158	Sao Miguel Azores II	Atlantic Ocean	Portugal
OSD159	Brest-SOMLIT	Atlantic Ocean	France
OSD162	Stonehaven	Atlantic Ocean	UK
OSD163	Scapa	Atlantic Ocean	UK
OSD164	Scalloway	Atlantic Ocean	UK
OSD165	Loch Ewe	Atlantic Ocean	UK
OSD166	Armintza	Atlantic Ocean	Spain
OSD167	Eyafjordur_6	Arctic Ocean	Iceland
OSD168	IMST_izmir	Mediterranean Sea	Turkey
OSD169	Brightlingsea Creek, Essex	Atlantic Ocean	UK
OSD170	Belgium - 130	Atlantic Ocean	Belgium
OSD171	Belgium - 230	Atlantic Ocean	Belgium
OSD173	Belgium - 710	Atlantic Ocean	Belgium
OSD174	Belgium - 780	Atlantic Ocean	Belgium
OSD175	ZG02	Atlantic Ocean	Belgium
OSD176	Belgium - 215	Atlantic Ocean	Belgium
OSD177	Belgium - 120	Atlantic Ocean	Belgium
OSD178	Belgium - 435	Atlantic Ocean	Belgium

OSD182	W08	Atlantic Ocean	Belgium
OSD183	W09	Atlantic Ocean	Belgium
OSD184	W10	Atlantic Ocean	Belgium
OSD185	Belgium - 421	Atlantic Ocean	Belgium
OSD186	SERC Rhode River Maryland	Atlantic Ocean	USA
OSD187	Palmer station	Southern Ocean	Antarctica

Table S2. Matrix of the pairwise percent identity between *Ostreococcus* clades. The calculation was done based on the alignment available as Supplementary data S6.

Ostreococcus	"O. lucimarinus"	O. clade B	O. tauri	O.mediterraneus	clade E
"O. lucimarinus"	99.70	98.80	99.40	98.00	99.20
O. clade B	98.80	99.70	99.10	97.70	99.40
O. tauri	99.40	99.10	99.70	98.30	99.50
O.mediterraneus	98.00	97.70	98.30	99.80	98.00
clade E	99.20	99.40	99.50	98.00	99.80

Table S3. Matrix of the pairwise percent identity between *Micromonas* clades. The calculation was done based on the alignment available as Supplementary data S7.

Micromonas	M. polaris	B sub-arctic	M. bravo II	M. bravo I	B._4	M. pusilla	B warm I	B warm II	M. commoda(C)	M. commoda(AB)
M. polaris	99.90	99.10	98.90	98.30	97.40	97.30	97.70	98.20	97.20	97.40
B sub-arctic	99.10	100.00	98.40	97.80	96.80	96.80	97.20	97.50	97.10	97.40
M. bravo II	98.90	98.40	100.00	99.40	97.00	97.60	98.10	98.20	97.50	97.50
M. bravo I	98.30	97.80	99.40	99.50	97.00	97.10	97.70	97.90	97.10	97.30
B._4	97.40	96.80	97.00	97.00	100.00	97.10	97.70	97.70	97.50	97.40
M. pusilla	97.30	96.80	97.60	97.10	97.10	99.80	98.30	98.60	97.40	96.90
B warm I	97.70	97.20	98.10	97.70	97.70	98.30	100.00	99.80	98.30	97.90
B warm II	98.20	97.50	98.20	97.90	97.70	98.60	99.80	100.00	98.30	97.20
M. commoda(C)	97.20	97.10	97.50	97.10	97.50	97.40	98.30	98.30	99.80	99.20
M. commoda(AB)	97.40	97.40	97.50	97.30	97.40	96.90	97.90	97.20	99.20	100.00

Table S4. Matrix of the pairwise percent identity between *Mantoniella* clades. The calculation was done based on the alignment available as Supplementary data S8.

Mantoniella	M. antarctica	M. squamata	M. clade A	M. clade B
M. antarctica	99.70	99.00	96.70	99.30
M. squamata	99.00	100.00	97.50	99.80
M. clade A	96.70	97.50	97.60	97.30
M. clade B	99.30	99.80	97.30	99.90

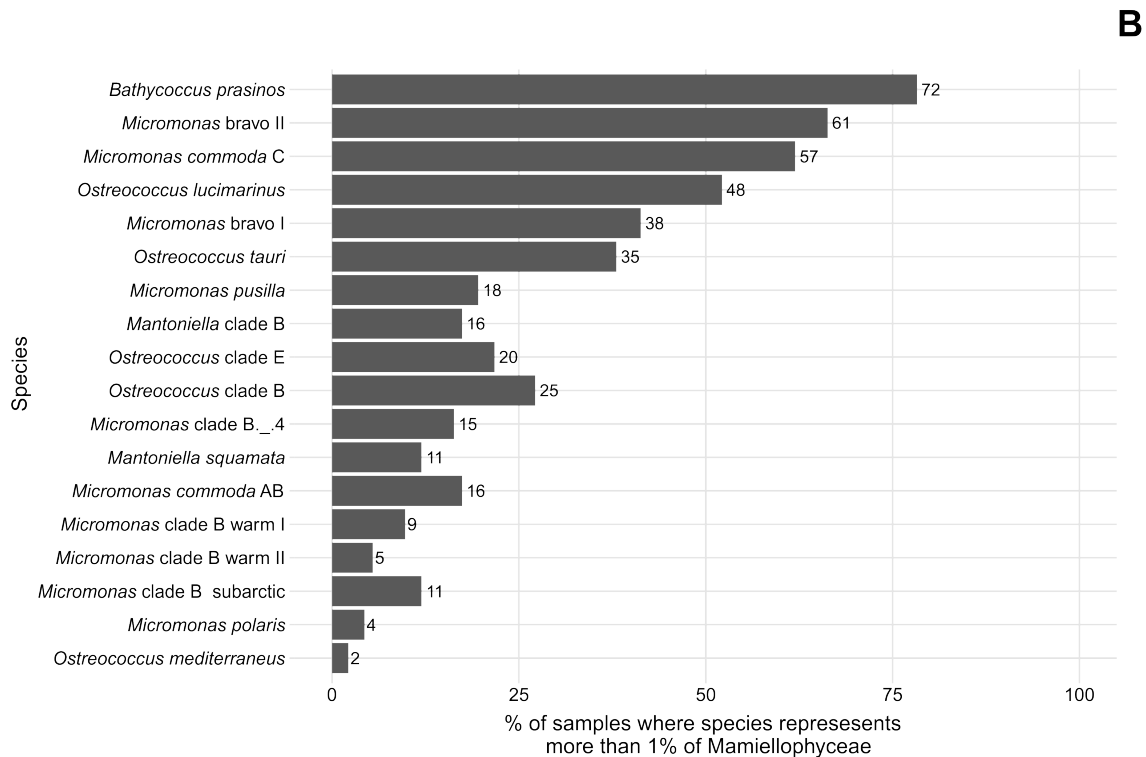
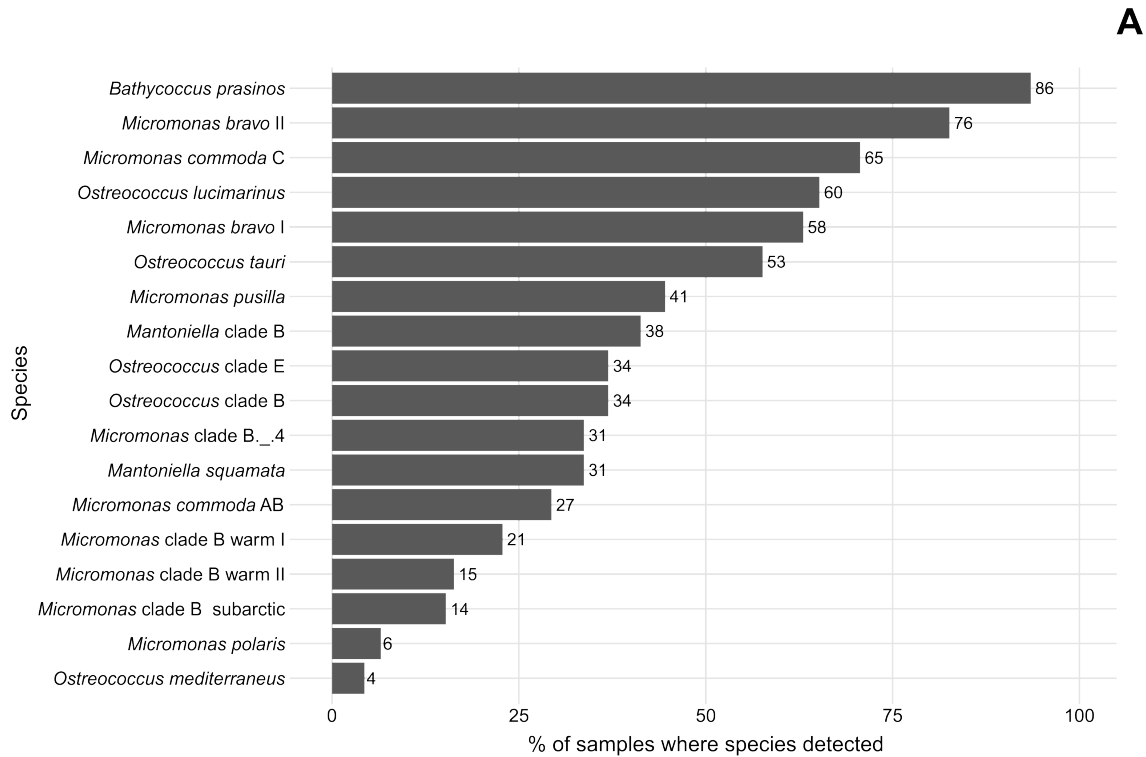


Figure S1. A. Percentage of OSD samples where a given ASV was detected (at least one read). Numbers at right of bars correspond to number of samples. B. Idem but for samples where the ASV contributed more than 1% of the Mamiellophyceae reads. Stations with less than 100 Mamiellophyceae reads were not considered. The three major *O. tauri* ASVs have been pooled together.

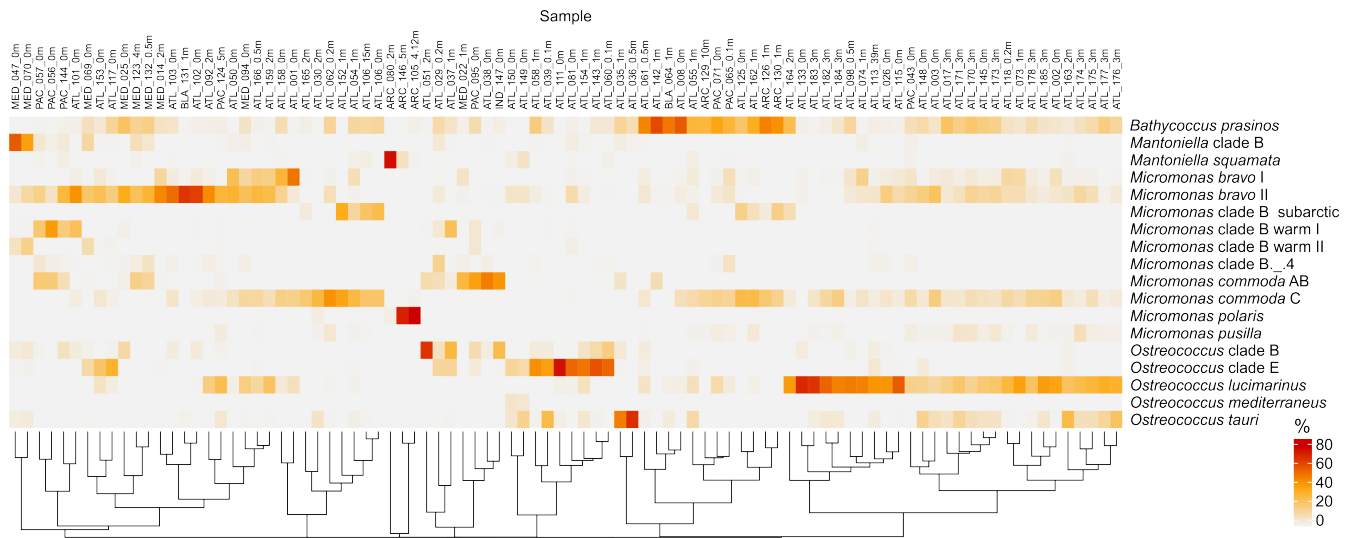


Figure S2. Heatmap of the Mamiellophyceae communities in OSD samples. The number of read corresponding to the major ASVs were normalized by the number of Mamiellophyceae reads per station. Stations with less than 100 Mamiellophyceae reads were not considered.

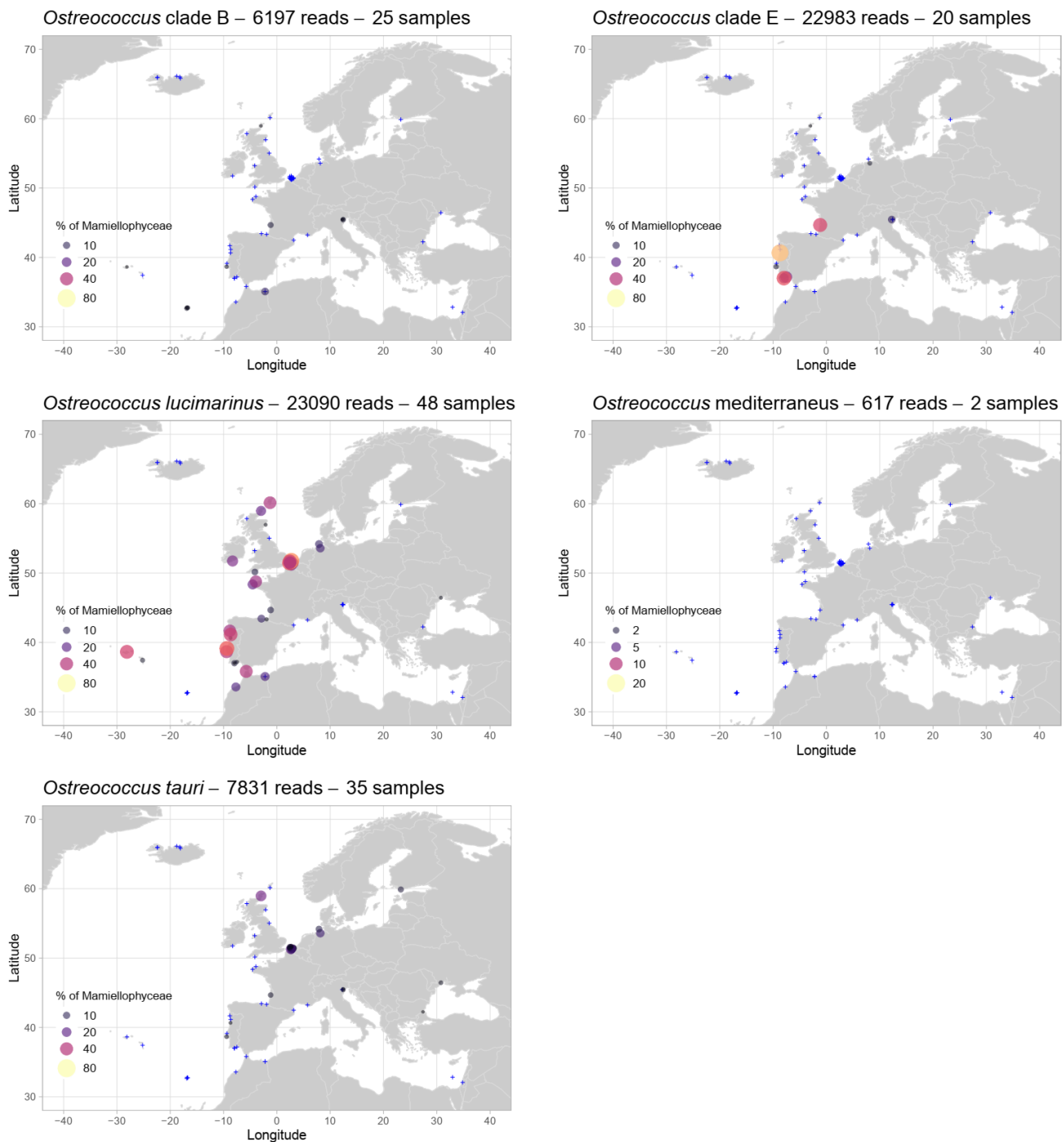


Figure S3. Distribution of major *Ostreococcus* ASVs for OSD2014 (LGC) in European waters. The three major *O. tauri* ASVs have been pooled together. Circle surface corresponds to the percentage of ASV reads relative to the total number of Mamiellophyceae reads. Samples for which the ASV contribution was lower than 1 % are represented by blue crosses.

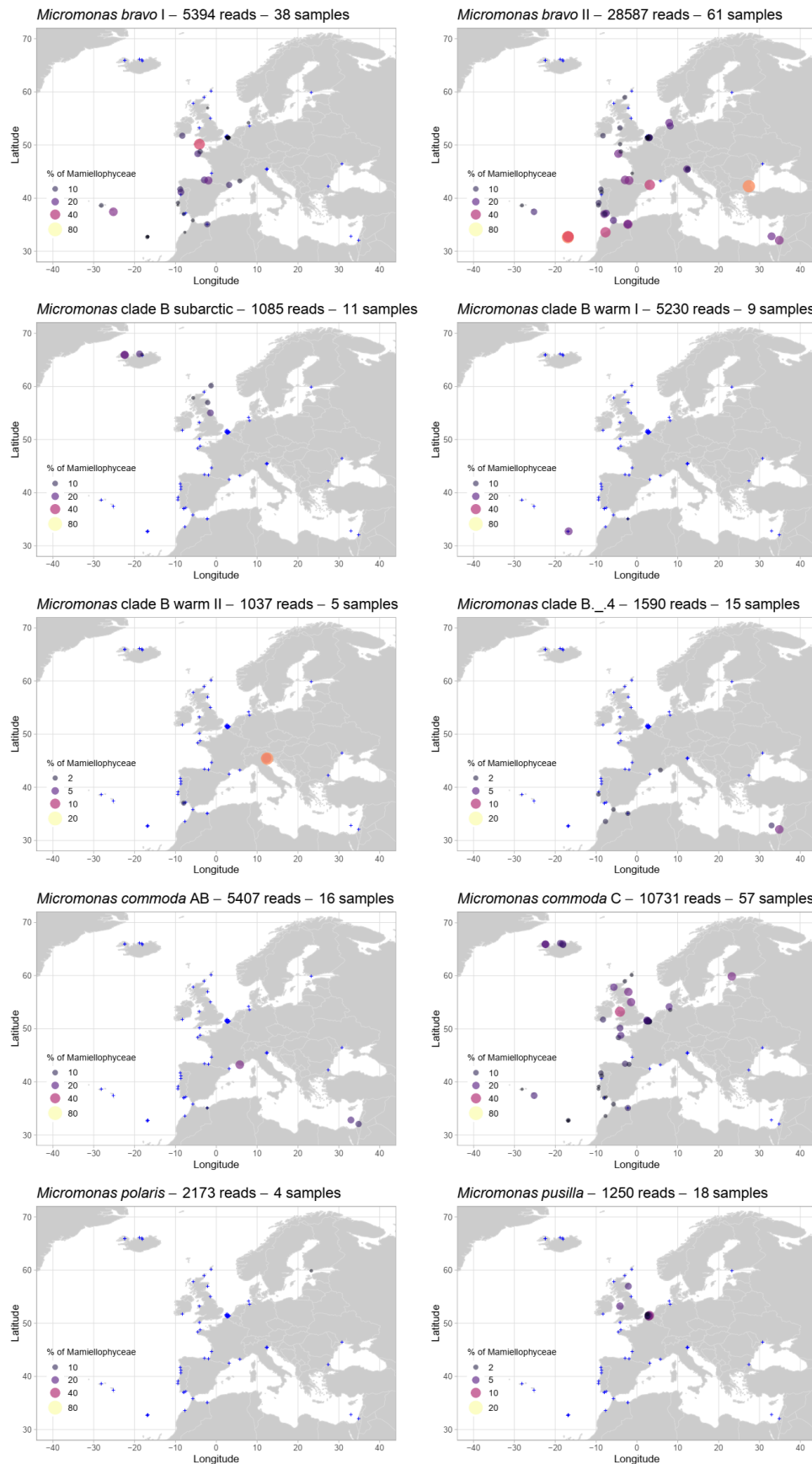


Figure S5. Same as Fig. S3 for *Micromonas*.

Bathycoccus prasinus – 16651 reads – 72 samples

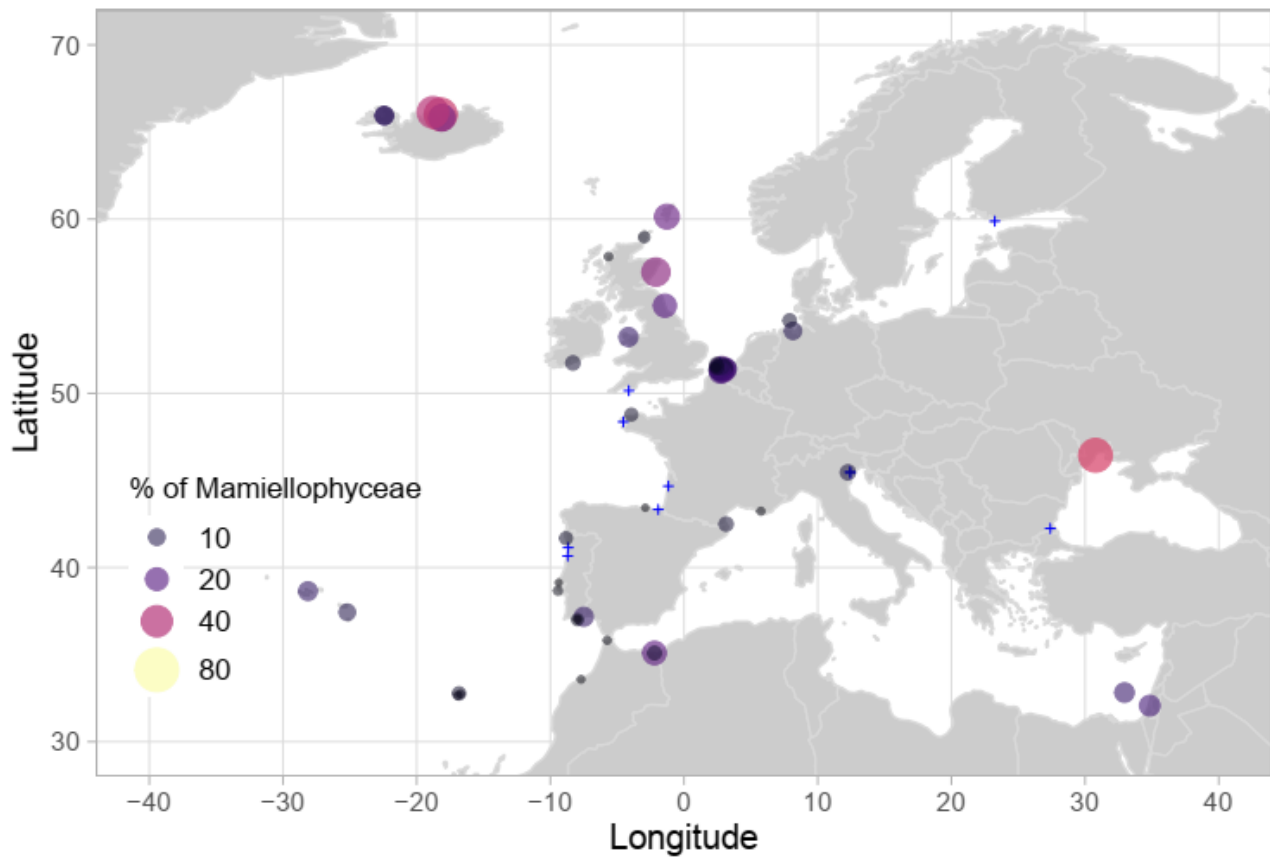
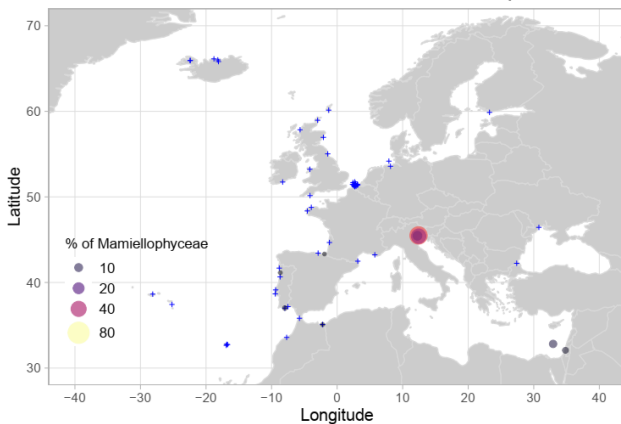


Figure S6. Same as Fig. S3 for *Bathycoccus*.

Mantoniella clade B – 4037 reads – 16 samples



Mantoniella squamata – 4498 reads – 11 samples

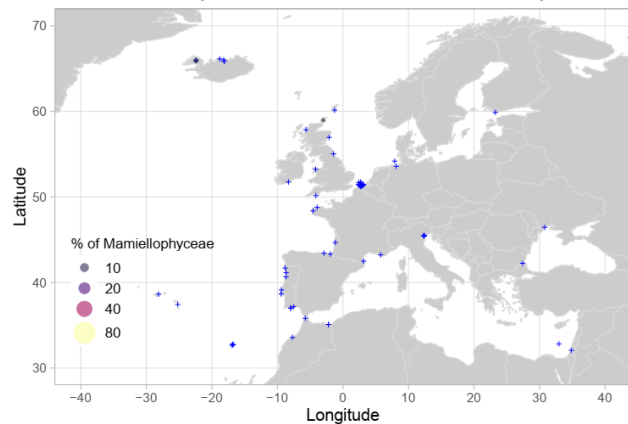


Figure S7. Same as Fig. S3 for *Mantoniella*.