Mixotrophic protists display contrasted biogeographies in the global ocean

Emile Faure^{1,2} · Fabrice Not³ · Anne-Sophie Benoiston² · Karine Labadie⁴ · Lucie Bittner² · Sakina-Dorothee Ayata^{1,2}

Received: 2 July 2018 / Revised: 13 December 2018 / Accepted: 13 December 2018 / Published online: 14 January 2019
© International Society for Microbial Ecology 2019

Abstract

Mixotrophy, or the ability to acquire carbon from both auto- and heterotrophy, is a widespread ecological trait in marine protists. Using a metabarcoding dataset of marine plankton from the global ocean, 318,054 mixotrophic metabarcodes represented by 89,951,866 sequences and belonging to 133 taxonomic lineages were identified and classified into four mixotrophic functional types: constitutive mixotrophs (CM), generalist non-constitutive mixotrophs (GNCM), endo-symbiotic specialist non-constitutive mixotrophs (eSNCM), and plastidic specialist non-constitutive mixotrophs (pSNCM). Mixotrophy appeared ubiquitous, and the distributions of the four mixotypes were analyzed to identify the abiotic factors shaping their biogeographies. Kleptoplastidic mixotrophs (GNCM and pSNCM) were detected in new zones compared to previous morphological studies. Constitutive and non-constitutive mixotrophs had similar ranges of distributions. Most lineages were evenly found in the samples, yet some of them displayed strongly contrasted distributions, both across and within mixotypes. Particularly divergent biogeographies were found within endo-symbiotic mixotrophs, depending on the ability to form colonies or the mode of symbiosis. We showed how metabarcoding can be used in a complementary way with previous morphological observations to study the biogeography of mixotrophic protists and to identify key drivers of their biogeography.

Introduction

Marine unicellular eukaryotes, or protists, have a tremendous range of life styles, sizes and forms [1], showing a

taxonomic and functional diversity that remains hard to define [2, 3]. This variety of organisms is having an impact on major biogeochemical cycles such as carbon, oxygen, nitrogen, sulfur, silica, or iron, while being at the base of marine trophic networks [4–8]. Hence, they are key actors of the global functioning of the ocean.

Historically, marine protists have been classified into two groups depending on their trophic strategy: the photo-synthetic plankton (phytoplankton) and the heterotrophic plankton (zooplankton). It is now clear that mixotrophy, i.e., the ability to combine autotrophy and heterotrophy, has been largely underestimated and is commonly found in planktonic protists [6, 9–13]. Instead of a dichotomy between two trophic types, their trophic regime should be regarded as a continuum between full phototrophy and full heterotrophy, with species from many planktonic lineages lying between these two extremes [10]. Mitra et al. [11] have proposed a classification of marine mixotrophic protists into four functional groups, or mixotypes. The constitutive mixotrophs, or CM, are photosynthetic organisms that are capable of phagotrophy, also called “phytoplankton that eat” [11]. They include most mixotrophic nano-flagellates (e.g., *Prymnesium parvum*, *Karlodinium micrum*). On the opposite, the non-constitutive mixotrophs,

These authors contributed equally: Lucie Bittner, Sakina-Dorothee Ayata

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-018-0340-5>) contains supplementary material, which is available to authorized users.

✉ Emile Faure
emile.faure@etu.upmc.fr

¹ Sorbonne Université, CNRS, Laboratoire d’océanographie de Villefranche, LOV, 06230 Villefranche-sur-Mer, France

² Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d’Histoire naturelle, CNRS, Sorbonne Université, EPHE, CP 50, 57 rue Cuvier, 75005 Paris, France

³ Sorbonne Université, CNRS, UMR7144 Adaptation and Diversity in Marine Environment (AD2M) Laboratory, Ecology of Marine Plankton team, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France

⁴ Genoscope, Institut de biologie François-Jacob, Commissariat à l’Energie Atomique (CEA), Evry, France, 91057 Evry, France

or “photosynthetic zooplankton”, are heterotrophic organisms that have developed the ability to acquire energy through photosynthesis [9]. This ability can be acquired in three different ways: the generalist non-constitutive mixotrophs (GNCM) steal the chloroplasts of their prey, such as most plastid-retaining oligotrich ciliates (e.g., *Laboea strobila*), the plastidic specialist non-constitutive mixotrophs (pSNCM) steal the chloroplasts of a specific type of prey (e.g., *Mesodinium rubrum* or *Dinophysis* spp.), and finally the endo-symbiotic specialist non-constitutive mixotrophs (eSNCM) are bearing photosynthetically active endo-symbionts (most mixotrophic Rhizaria from Collozaria, Acantharea, Polycystinea, and Foraminifera, as well as dinoflagellates like *Noctiluca scintillans*).

As drivers of biogeochemical cycles in the global ocean, and particularly of the biological carbon pump [5, 14, 15], marine protists are a key part of ocean biogeochemical models [7, 16–18]. However, physiological details of mixotrophic energy acquisition strategies have only been studied in a restricted number of lineages [9, 19, 20]. They appear to be quite complex and greatly differ across mixotypes, which makes mixotrophy hard to include in a simple model structure [21–25]. Hence at this time, mixotrophy is not included in most biogeochemical models, neglecting the amount of carbon fixed by non-constitutive mixotrophs through photosynthesis, and missing the population dynamics of photosynthetically active constitutive mixotrophs that can still grow under nutrient limitation [23, 26]. This is most probably skewing climatic models predictions [11, 26], as well as our ability to understand and prevent future effects of global change.

A better understanding of the environmental diversity of marine mixotrophic protists, as well as a description of the abiotic factors driving their biogeography at global scale are still needed, in particular to integrate them in biogeochemical models. Leles et al. [27] attempted to tackle this problem by reviewing about 110,000 morphological identification records of a set of more than 60 mixotrophic protists species in the ocean, taken from the Ocean Biogeographic Information System (OBIS) database. They found distinctive patterns in the biogeography of the three different non-constitutive mixotypes (GNCM, pSNCM, and eSNCM), highlighting the need to better understand such diverging distributions [27]. Environmental molecular biodiversity surveys through metabarcoding have been widely used in the past fifteen years to decipher planktonic taxonomic diversity [2, 28–30]. Here, we exploited the global *Tara* Oceans datasets [31–33], and identified 133 mixotrophic lineages, that we classified into the four mixotypes defined by Mitra et al. [11]. This first ever set of mixotrophic metabarcodes allowed us to investigate the global biogeography of both constitutive and non-constitutive mixotrophs, in relation with in-situ abiotic measurements.

We tested (i) if new information on marine mixotrophic protists distribution can be gained in comparison with previous morphological identifications [27]; (ii) if the constitutive mixotrophs, which are not addressed in Leles et al. [27], and the non-constitutive mixotrophs diverge in terms of biogeography; (iii) if the study of diversity and abundance of environmental metabarcodes could lead to the definition of key environmental factors shaping mixotrophic communities.

Materials and methods

Samples collection and dataset creation

Metabarcoding datasets from the worldwide *Tara* Oceans sampling campaigns that took place between 2009 and 2013 [31, 33] (data published in open access at the European Nucleotide Archive under project accession number PRJEB6610) were investigated. We analyzed 659 samples from 122 distinct stations, and for each sample, the V9-18S ribosomal DNA region was sequenced through Illumina HiSeq [32]. Assembled and filtered V9 metabarcodes (cf. details in de Vargas et al. [2]) were assigned to the lowest taxonomic rank possible via the Protist Ribosomal Reference (PR2) database [34]. To limit false positives, we chose to only analyze the metabarcodes (i.e., unique versions of V9 sequences) for which the assignment to a reference sequence had been achieved with a similarity of 95% or higher. This represents 65% of the total dataset in terms of metabarcodes and 84% in terms of total sequences. Our dataset involved 1,492,912,215 sequences, distributed into 4,099,567 metabarcodes assigned to 5071 different taxonomic assignments, going from species to kingdom level precision.

Defining a set of mixotrophic organisms

Among these 5071 taxonomic assignments, we searched for mixotrophic protist lineages, taking into account the four mixotypes described by Mitra et al. [11]: constitutive mixotrophs (CM), generalist non-constitutive mixotrophs (GNCM), endo-symbiotic specialist non-constitutive mixotrophs (eSNCM), and plastidic specialist non-constitutive mixotrophs (pSNCM). We used the table S2 from Leles et al. [27], which is referencing 71 species or genera belonging to three non-constitutive mixotypes (GNCM, pSNCM, and eSNCM), as well as multiple other sources coming from the recent literature on mixotrophy [6, 9–12, 35–47], and inputs from mixotrophic protists’ taxonomy specialists (cf. Acknowledgments section). Within the 5071 taxonomic assignments of variable precisions, we identified 5 GNCM, 9 pSNCM, 77 eSNCM, and 42 CM lineages (detailed list available publicly under the <https://doi.org/10.>

6084/m9.figshare.6715754, and all metabarcodes were tagged with their mixotypes in the PR2 database). Among these 133 taxonomic assignments that we will call “lineages”, 92 were defined at the species level, 119 at the genus level, and the last 14 at higher taxonomic levels where mixotrophy is always present (mostly eSNCM groups like Collodaria). In the Chrysophyceae family, metabarcodes assigned to clades B2, E, G, H, and I were included even though we couldn't find a general proof that all species included in these clades have mixotrophic capabilities. However, if we exclude the photolithophilic Synurophyceae and genera like *Paraphysomonas* and *Spumella*, which we did, a vast majority of Chrysophyceae are considered mixotrophic [10]. The final dataset included 318 054 metabarcodes assigned to the 133 mixotrophic lineages selected, as well as their sequence abundance in 659 samples (table available publicly under the <https://doi.org/10.6084/m9.figshare.6715754>).

Environmental dataset

We built a corresponding contextual dataset using the environmental variables available in the PANGAEA repository from the *Tara* Oceans expeditions [33, 48]. The set of 235 environmental variables was reduced to 57 due to several selection steps (Data available publicly under the <https://doi.org/10.6084/m9.figshare.6715754>; see the details of variable selection in section 1 of Supp. Mat.).

Distribution and diversity of mixotrophic protists

For each mixotype, the number of metabarcodes, the total sequence abundance and the mean sequence abundance by metabarcode was computed (Table 1). Also, we measured each metabarcode's station occupancy, i.e., the number of stations in which it was found, and station evenness, i.e., the homogeneity of its distribution among the stations in which

it was detected (Fig. 2). Diversity of mixotrophic protists was investigated through mixotype-specific metabarcode richness per station (Table 1). As the number of samples taken per station can impact the abundance and diversity of detected metabarcodes, richness was computed only at stations for which the maximum number of eight samples were available (40 stations over 122).

Global biogeography of mixotrophic protists

Two statistical analyses were performed to investigate mixotrophic protists biogeography. One at the metabarcode level, and one at the lineage level, i.e., merging the sequence abundance of metabarcodes sharing the same taxonomical assignment. The metabarcodes abundance table was composed of 318,054 rows/metabarcodes, and 659 columns/samples, whereas the lineage abundance table was composed of 133 rows/lineages and 659 columns/samples (both datasets are available publicly under the <https://doi.org/10.6084/m9.figshare.6715754>). The two analyses led to very similar conclusions, but the biogeography of lineages appeared easier to visually represent and interpret than the one of metabarcodes. Hence, we only present here the results of the lineage-based analysis (See section 3 of Sup. Mat. for metabarcode-level analysis results and discussion).

Our statistical model was designed to identify lineages (or metabarcodes) with contrasted biogeographies, and relate their presence to the environmental context. We normalized the sequence counts from the lineage abundance matrix using a Hellinger transformation [49]. We used the environmental dataset and the mixotrophic lineages' abundance matrix as explanatory and response matrices, respectively, to conduct a redundancy analysis (RDA) [49]. For that, we made a species pre-selection using Escoufier's vectors [50], which allowed to keep only the 62 most significant mixotrophic lineages. This method selects lineages

Table 1 Detailed number of lineages found for each mixotype, as well as the number of metabarcodes, the corresponding total sequence counts over all stations, the mean sequence abundance by metabarcode, and mean metabarcode richness

Mixotypes	CM	eSNCM	pSNCM	GNCM
Number of lineages used in this study	42	77	9	5
Number of V9 metabarcodes	26,015	288,536	2143	1360
Total sequence abundance	3,581,751	86,098,397	208.096	63.622
Mean sequence counts per metabarcode	137.7	298.4	97.1	46.8
Mean metabarcode richness per station ^a (std dev)	2162 (1115)	18502 (9238)	67 (102)	84 (111)
Number of absences/station	0/122	0/122	5/122	3/122

The richness was computed as the number of different metabarcodes present at each station. It was calculated for each mixotype and means are indicated in the fifth column. Absences correspond to the number of stations in which no sequences were detected for the corresponding mixotype

CM constitutive mixotrophs, GNCM generalist non-constitutive mixotrophs, eSNCM endo-symbiotic specialist non-constitutive mixotrophs, pSNCM plastidic specialist non-constitutive mixotrophs

^aThe mean indicated here was calculated using only stations having the maximum number of samples (see main text)

according to a principal component analysis (PCA), sorting them based on their correlation to the principal axes. We then used a maximum model ($Y \sim X$) and a null model ($Y \sim 1$) to conduct a two directional stepwise model selection based on the Akaike information criterion (AIC) [51]. The resulting model contained 28 environmental response variables. More details about statistical analyses are available in section 2 and 3 of the Supplementary Materials. analyses and graphs were realized with the R software version 3.4.3 [52]. All scripts are available on GitHub platform (<https://github.com/upmcgenomics/MixoBioGeo>).

Results

Global distribution and diversity of marine mixotrophic protists

Mixotrophic protists metabarcodes were detected in all the 659 samples with a total sequence abundance of 89,951,866, representing 12.56% of the total sequence abundance in the 659 samples studied. They represented a

mean of 12.64% of the total sequence abundance per sample, with a maximum of 96.96% and a minimum of 0.01%. To avoid any potential overestimation of mixotrophic lineages presence in the following results, we marked all records of less than a hundred sequences as questionable. We found both eSNCM and CM in each of the 122 stations studied (Table 1 and Fig. 1). In only two occasions the number of sequences belonging to CM was questionable, at stations for which only one sample was sequenced. GNCM were found absent in only two stations and their presence was questionable in 39 stations (Fig. 1). pSNCM were absent at five stations (three in the Indian Ocean, and two in the Pacific Ocean) and detected with questionable presence in 54 additional stations, which were mostly located in the central Pacific and the Indian Ocean (Fig. 1). We found significant amounts of sequences corresponding to GNCM in the Central Pacific, Southern subtropical Atlantic, and Indian Ocean. The presence of GNCM in these areas has not yet been recorded through morphological identifications during field expeditions [27]. Also, we detected more than 100 sequences of pSNCM metabarcodes at 11 stations belonging to biogeographical

Fig. 1 Global distribution of mixotypes from metabarcoding data. Maps showing for each station the proportion of sequences (in %) belonging to each mixotype over the total number of mixotrophic sequences. Stations in which no sequence was found were marked as absent, ones with less than 100 sequences marked as questionable.

Each Longhurst biogeographical provinces [53] is colored in the background if more than 100 sequences are detected in at least one of its stations. A coloured version of this figure can be downloaded at <https://doi.org/10.6084/m9.figshare.6715754>

Fig. 2 Sequence abundance, occupancy, and spatial evenness of each mixotrophic metabarcode across sampled stations. Each metabarcode is plotted as a bubble, with its station occupancy, i.e., the number of stations in which it was found, and its station evenness, i.e., the

homogeneity of its distribution among the stations in which it was detected, as coordinates. Violin plots were drawn for each mixotype on both the x and y axes. The size of each bubble is scaled to the sequence abundance found globally for the corresponding metabarcode

provinces in which no morphological identifications had been published [27, 53], mostly in offshore areas of the Atlantic and Pacific Ocean (Fig. 1).

The mean evenness of mixotrophic metabarcodes across stations was of 0.87, and 82.3% of the metabarcodes had a station evenness above 0.5 (Fig. 2). Station occupancy varied a lot depending on the metabarcodes, with a high density of rare metabarcodes leading to a mean of 5.14 stations over a maximum of 122, and a standard deviation of 7.7. However, three eSNCM metabarcodes were found in all the 122 stations, and three CM metabarcodes were detected in 121 stations. The maximum occupancy for a GNCM metabarcode was of 111 stations, while 92 stations was the maximum for a pSNCM metabarcode. CM and GNCM metabarcodes showed a strong tendency towards high evenness values (Fig. 2, means of 0.90 and 0.95, respectively), even for the most sequence abundant metabarcodes. Many eSNCM metabarcodes had high evenness values, but below average values were detected for the most abundant ones (Fig. 2, global mean of 0.87). pSNCM metabarcodes had a similar mean of

evenness values (0.87), but a different distribution compared to other mixotypes (Fig. 2). Among the 50 most abundant metabarcodes, 43 corresponded to *Collodaria* lineages, 47 were eSNCM and 3 were CM, all three assigned to *Gonyaulax polygramma*. GNCM and pSNCM metabarcodes had homogeneously low sequence abundances (Fig. 2 and Table 1).

Main factors affecting the biogeography of mixotrophic protists

The redundancy analysis helped to investigate further the environmental variables responsible for the mixotrophic protists' biogeography. The 62 lineages selected with the Escoufier's vector method corresponded to 20 CM, 34 eSNCM, 3 GNCM, and 5 pSNCM. Even after selection, a significant part of the lineages did not show any response to environmental data in their distribution (Fig. 3, e.g., 19 of the 62 lineages were found between -0.01 and 0.01 on both RDA1 and RDA2). The adjusted R-squared of the RDA was of 34.89% (41.43% unadjusted), with 24.01% of

Fig. 3 Impact of environmental variables on the distribution of marine mixotrophs. Triplot of the redundancy analysis (RDA) computed on the 62 Escoufier-selected lineages, after model selection. The adjusted R-squared of the analysis is of 34.89% (41.43% unadjusted). Each gray dot corresponds to a sample, i.e., one filter at one depth at one station. The blue dashed arrows correspond to the quantitative environmental variables. Abbreviations: MLD mixed layer depth, O2MaxD O2 maximum depth, EuphzoneD euphotic zone depth, PAR photosynthetically active radiations, Calcite Sat. St. Calcite Saturation State,

c_660 optical beam attenuation coefficient at 660 nm, c_part beam attenuation coefficient of particles, acCDOM absorption coefficient of colored dissolved organic matter. Plain arrows correspond to mixotrophic lineages, colors indicating mixotypes. For more readability, we do not represent all qualitative variables included in the model. That is why only the filter centroids are appearing, even though the sampling depth, season, season moment, i.e., early, middle or late, and biogeographical province were used in the RDA calculation

variance explained on the two first axes (Fig. 3). The first RDA axis (14.96%) marks an opposition between samples from oligotrophic waters with low productivity (RDA1 > 0) and samples from eutrophic and productive water masses (RDA1 < 0). This axis is negatively correlated to chlorophyll concentration, particles density, ammonium concentration, absorption coefficient of colored dissolved organic matter (acCDOM), duration of daylight, silica, CO3, oxygen, and PO4 concentration, as well as longitude. It is positively correlated to bathymetry, deep euphotic zone, deep oxygen maximum, deep mixed layer, as well as to the distance to coast. The second RDA axis (9.05%) is opposing offshore and subpolar samples (RDA2 > 0) to coastal and subtropical ones (RDA2 < 0). The axis is positively correlated to the depth of the mixed layer, the distance to coast, the bathymetry, high maximum Lyapunov exponents as well as high concentrations of PO4, oxygen,

CO3 and silica. It is negatively correlated to temperature, salinity, and photosynthetically active radiations (PAR).

Among the 20 CM lineages, seven clearly emerged from the redundancy analysis (Fig. 3) and showed distinct biogeographies related to environmental variables. *Gonyaulax polygramma*, *Alexandrium tamarense*, and *Fragilidium mexicanum*, three Dinophyceae belonging to the Gonyaulacales order, were mainly found in oligotrophic waters with a deep euphotic zone, warm temperature, high salinity, and PAR (RDA1 > 0, RDA2 < 0). The four other CMs (involving all the Chrysophyceae included in the analysis as well as one Dinophyceae from the Kareniaceae family, *Karlodinium micrum*) were found mostly in productive water masses (RDA1 < 0).

eSNCMs can be divided in three groups in the RDA space. The first group (RDA1 < 0) corresponds to eSNCM species dominating rich and productive environments. It

Fig. 4 Contrasted global distributions of metabarcodes corresponding to two eSNCM lineages. Maps of Hellinger-transformed sequence count abundances for metabarcodes assigned to the Collodaria *Siphonosphaera cyathina* **a** and the Acantharia Acanthrometridae F3 spp. **b** These two lineages are opposed on the first RDA axis (Fig. 3

and S1). Size and color both illustrate abundance for better readability. Ellipses were drawn to highlight high abundance zones, and reveal the differences in lineages distribution. A coloured version of this figure can be downloaded at <https://doi.org/10.6084/m9.figshare.6715754>

includes mainly Acantharia and Spumellaria species. The second group ($\text{RDA1} > 0$) dominates oligotrophic environments, and includes multiple Collodaria as well as one Dinophyceae genus (*Ornithocercus*). Within this group, *Ornithocercus* spp. is found mainly in coastal subtropical environments ($\text{RDA2} < 0$), as opposed to *Sphaerozoum punctatum* that is found mainly in offshore subpolar regions ($\text{RDA2} > 0$). *Siphonosphaera cyathina* lies between these two trends as it is found only in oligotrophic samples, but is not influenced by temperature or bathymetry (Figs. 3 and 4). The third group corresponds to the eSNCM lineages that can be interpreted as distributed homogeneously in regards of the environmental data we are using (e.g., lineages with the shortest arrows in Fig. 3). These notably include the 12 Foraminifera lineages present in the RDA. Looking at filters centroids in the RDA space (Fig. 3), we can suppose that

eSNCM lineages dominating eutrophic systems ($\text{RDA1} < 0$) are smaller in size than those dominating oligotrophic ones ($\text{RDA1} > 0$).

Out of the five pSNCM included in the RDA, only *Mesodinium rubrum*, the most abundant one, is distinctively represented in the RDA space. This suggests that the other pSNCM have homogeneous distributions in response to our environmental variables. *Mesodinium rubrum* dominates eutrophic environments, independently from the bathymetry or the temperature ($\text{RDA1} < 0$, $\text{RDA2} \approx 0$). We find a similar pattern for GNCM, with only *Pseudotontonia simplicidens* well represented in the RDA space out of the three species included in the analysis. Like *M. rubrum*, *Pseudotontonia simplicidens* is the most abundant species in its group and it is mainly found in eutrophic waters ($\text{RDA1} < 0$).

Discussion

Mixotrophy occurs everywhere in the global ocean

Our metabarcoding survey confirms that marine mixotrophic protists are ubiquitous in the global ocean [27], possibly extending the known range of distribution of two mixotypes (Figs. 1 and 2). Mixotrophic organisms represented more than 12% of the sequences in the complete *Tara* Oceans metabarcoding dataset, showing that they should not be understated. We found contrasted biogeographies among metabarcodes and their corresponding lineages, both within and across mixotypes (Figs. 2–4 and S1, Sup. Mat. section 3). We found constitutive mixotrophs (CM) and endo-symbiotic specialist non-constitutive mixotrophs (eSNCM) metabarcodes at all the 122 stations included in this global study (Table 1 and Fig. 2), verifying that these two mixotypes are the most abundant in the ocean [27, 47, 54, 55]. This dominance of eSNCM and CM in our data is also linked to the relatively high number of metabarcodes available for these two mixotypes in databases. Using 1360 generalist non-constitutive mixotrophs (GNCM) metabarcodes corresponding to only five lineages, we detected them in ten biogeographical provinces [53] where no morphological identification had been recorded before [27]. GNCM metabarcodes had consistently high evenness values, and some had station occupancy records comparable to the most abundant eSNCM and CM metabarcodes (Fig. 2). These results support the hypothesis of a globally ubiquitous distribution of GNCM. Plastidic specialist non-constitutive mixotrophs (pSNCM) were found in five provinces in which no record existed so far from morphological identification field studies [27]. However, these observations were often in a questionable range in terms of sequence abundance (Fig. 1), and the overall distribution of pSNCM in our data appears as very concordant with morphological observations [27]. pSNCM metabarcodes had dominantly low station evenness values, which again supports the conclusions of Leles et al. [27] that identified pSNCM as highly seasonal and spatially restricted in their distribution.

While building our set of mixotrophic lineages, some widespread and potentially mixotrophic genera did not appear, such as *Ceratium* spp., *Tontonia* spp., *Amphisolenia* spp., *Triposolenia* spp., or *Citharistes* spp., mainly because of a poor representation in the PR2 database. Also, we decided to only consider metabarcodes with more than 95% similarity to a reference sequence. This threshold could be too selective for some species and not enough for some others, as single similarity threshold are hardly efficient when studying whole eukaryotic populations [56, 57]. For example, some species appeared with low sequence abundance in the data even though they could not have been

sampled, such as three lacustrine species, e.g., *Poteroisopumella lacustris*. Considering these biases and the sometimes relatively low sequence counts (marked as questionable in Fig. 1), some of the new GNCM and pSNCM records we observed should be considered with care, as they could be over-estimated or even sometimes artefactual. However, the low number of lineages found for these two mixotypes in PR2 and in our dataset are leading us to think that we were unable to capture the whole GNCM and pSNCM communities. This supposes a global under-estimation of GNCM and pSNCM abundances in our results.

Tara Oceans metabarcoding dataset is built on snapshot samples taken irregularly during a 3-year cruise, hence allowing no proper seasonal variations investigations. However, morphological identifications of mixotrophic protists revealed seasonal variations in their abundance, with *Mesodinium* biomass blooming in spring in coastal seas for example [27]. As metabarcoding datasets have been successfully applied on time series to detect species successions across gradients of time and space [58–60], it would be interesting to similarly investigate seasonal trends in mixotrophic communities. Our set of mixotrophic lineages and metabarcodes being publicly available, our method will be applicable to any other metabarcoding dataset, including time series. It will also be open to inputs and updates from the global scientific community.

The contrasted biogeographies of marine mixotypes

Constitutive mixotrophs

Constitutive mixotrophs (CM) have very diverse feeding behaviors, with some species requiring phototrophy to grow, others phagotrophy, and some being obligate mixotrophs [9]. They were described in all waters of the global ocean [61–65]. We found them distributed in a range of conditions almost as wide as non-constitutive mixotrophs (Figs. 1 and 3). Among highly abundant lineages, most were dominantly found in eutrophic and shallow habitats. However, a few dinoflagellates were found to be highly dominant in oligotrophic, subtropical waters, showing how wide of a range of conditions constitutive mixotrophs can grow in (Fig. 3). This illustrates how mixotrophy can allow organisms to dominate ecosystems even when environmental conditions are poorly adapted to purely phototrophic or heterotrophic organisms. When taken explicitly into account in biogeochemical models, marine mixotrophs increase carbon export by up to 30% [22]. Hence, their global ubiquity supposes that the carbon export of the biological carbon pump could be underestimated in both oligotrophic and eutrophic areas [26].

Plastidic specialist and generalist non-constitutive mixotrophs (pSNCM and GNCM)

Like Leles et al. [27], we found pSNCM and GNCM to have quite similar biogeographies (Fig. 3, section 3 of Sup. Mat.). Sequence abundance of most of the metabarcodes for these two mixotypes was homogeneously low (Table 1), but the two most abundant species, *Mesodinium rubrum* (pSNCM) and *Pseudotontonia simplicidens* (GNCM), were found mostly in coastal and eutrophic waters, consistently with Leles et al.'s [27] morphological observations (Fig. 3, section 3 of Sup. Mat.). No species-level barcode is available in the PR2 database for the *Tontonia* genus, and only one can be found for *Pseudotontonia* and *Laboea* genera, even though morphological records of these GNCM are numerous [27]. Experiments using meso- and microcosms combined with individual counts and morphological identification have found that GNCM ciliates can represent up to half of the individuals in ciliate communities of the photic zone [11, 66, 67]. A proportion we would have trouble to reach with the five lineages we were able to consider, knowing that there are 8686 different ciliate lineages available in PR2. This highlights the urgent need for supplementing 18S reference databases for mixotrophic ciliates.

Endo-symbiotic specialist non-constitutive mixotrophs (eSNCM)

Endo-symbiotic specialist non-constitutive mixotrophs (eSNCM) is by far the most widespread and abundant non-constitutive mixotype in the global ocean (Figs. 1 and 2) [27, 47, 54]. Their biogeography stands out, with a lot of highly abundant ubiquitous lineages, and some other specialized towards certain types of ecosystems (Fig. 3). They represent 95.7% of the sequence counts in our study and correspond to 90.7% of the metabarcodes (Table 1), which highlights their abundance and diversity. The very high number of rDNA copies present in Rhizaria orders such as Collodaria [47] might lead the eSNCM to appear more abundant in metabarcoding datasets than they ecologically are. However, in oligotrophic open oceans the Rhizaria biomass is estimated to be equivalent to that of all other mesozooplankton [68], and positively correlated to the carbon export [15], showing how ecologically important they can be.

Investigating the divergent biogeographies of Collodaria and Acantharia

Collodaria are living either as solitary large cells or as colonies [47], which explains why they are predominantly found in macro-sized (180–2000 µm) filter samples (Fig. 3).

All described Collodaria species so far harbor photosynthetic endo-symbionts, mostly identified as the dinoflagellate species *Brandtodinium nutricula* [47, 69]. These dinoflagellates are able to get in and out of their symbiotic state, which implies a light and/or reversible effect of the Collodarian host on its symbiont metabolism [69]. Based on the same metabarcoding dataset, Collodaria were described as particularly abundant and diverse in the oligotrophic open ocean [47]. In our results, Collodaria dominate oligotrophic, relatively deep waters (Figs. 3 and 4a). These Collodaria appear opposed to another set of Rhizaria (Acantharia and Spumellaria) linked to eutrophic and shallow waters (Figs. 3 and 4b, section 3 of Sup. Mat.). Acantharia are found ubiquitously in the global ocean, but display particularly high sequence abundances in some specific regions [54]. Mixotrophic Acantharia live in symbiosis with the cosmopolitan haptophyte *Phaeocystis*, which is highly abundant and ecologically active in its free-living phase [54]. Unlike the one of Collodaria, this symbiosis is irreversible: an algal symbiont can not go back to its free-living phase [54]. Our results suppose that these specific symbiotic modes could enable Acantharia and Collodaria to dominate different ecosystems (Figs. 3 and 4). Moreover, living in colonies as Collodaria could help to dominate oligotrophic systems, e.g., by accumulating more food and nutrients through their gelatinous extra-cellular matrix [47]. Experiments and modeling studies should help to investigate the contribution of this assumption, comparing food acquisition capacity and growth rates of free individuals versus in colony.

Towards an integration of mixotrophic diversity into marine ecosystem models

The future of marine communities' modeling lies in the integration of omics datasets into modeling frameworks [18, 70–73]. The use of metabolic networks and gene-centric methods has already shown very promising results in modeling prokaryotic ecological dynamics [18, 73]. However, eukaryotic metabolic complexity makes these methods hard to apply on protists for now, and we still lack a universal theoretical framework on how to integrate such methods into concrete modeling [70]. Mixotrophic protists are physiologically complex, and their feeding behavior can vary drastically on short time scales [9]. It will then take a few more years of comparative genomics and transcriptomics studies before being able to model their physiology with purely gene-based approaches. Still, mechanistic models of mixotrophy exist and are quite complex [21, 23], even if the one from Ghyyoot et al. [23] could be implemented in a global biogeochemical model [74]. Most models make the choice to represent either one or two (NCM and CM) types of organisms able to play the

role of all mixotypes depending on parameterization. However, no global agreement has been reached on to what extent the different mixotypes should be modeled. This is mainly due to a lack of quantitative and comparative data on the global impact of grazing and carbon fixation by the different mixotypes [75]. With our study, we show how meta-omics data can be used to identify groups of organisms distributed differently in response to the environment. It also allows the identification of ecological traits and environmental factors potentially responsible for these divergences. This information can be used to identify key species or lineages, and design controlled experiments with variations of targeted environmental factors to produce the quantitative data needed by modelers. Considering our results, we propose that host-symbiont dynamics of eSNCM should be investigated as a trait playing a potential role on Rhizaria ability to thrive in oligotrophic conditions. Particularly, the mechanisms behind holobiont formation and its potential reversibility could play major roles on eSNCM carbon fixation in various nutrient conditions. Future experiments comparing responses of Collodaria and Acantharia holobionts to different stresses in terms of grazing and carbon fixation could lead to a better understanding of the physiological differences between their two modes of symbiosis. Also, our results suggest that the metabolic flexibility of CM should allow this mixotype to grow in almost any conditions, with individual species probably spanning continuously between complete autotrophy and complete heterotrophy. The risk is then to create a “perfect beast” mixotroph dominating all systems [21]. To avoid that, we need more comparative data on grazing and carbon fixation of obligate phototrophs versus obligate heterotrophs in response to nutrient depletion and environmental fluctuation. Here again, meta-omics data could help to identify candidates for efficient experiment designs. Finally, the small number of lineages of GNCM and pSNCM in our study makes it hard to come up with strongly supported conclusions on whether they should be differentiated in models or not. They seem to share similar biogeographies using snapshot data (Fig. 3, section 3 of Sup. Mat.), but considering that they have different abilities for conserving stolen chloroplasts over time, it might not be the case when looking at a time series analysis [20, 76, 77].

Our study uses meta-omics data to investigate the global distribution and biogeography of mixotrophic protists in the ocean. Our results, currently based on metabarcoding data, complement morphological records and will be complemented in the near future by metagenomics and meta-transcriptomics studies. The latter will allow to distinguish the protists with mixotrophic capabilities from the ones with ongoing mixotrophic activity. This could lead to quantitative estimations of mixotrophic rates in environmental samples, allowing a sharpened study of mixotrophic protists

ecology in the global ocean. It could also lead to a metabolic description of complex processes like kleptoplasty and endo-symbiosis, hence facilitating the modeling of mixotrophic behaviors and its incorporation in ocean biogeochemical models.

Acknowledgements We would like to particularly thank Stéphane Pesant and Stéphane Audic for their work on making *Tara Oceans* datasets available. We also thank John Dolan (CNRS, LOV, Villefranche-sur-mer, France), Miguel Mendez-Sandin (Sorbonne Université, Station Biologique de Roscoff, France), and Wei-Ting Chen (National Taiwan Ocean University, Taiwan) for their essential help during the construction of the mixotrophic lineages set. We also thank Florentin Constancias for his help on the metabarcodes clustering tests conducted. Finally, we thank the three anonymous reviewers for their very constructive comments. This article is contribution number #84 of *Tara Oceans*. For the *Tara Oceans* expedition, we thank the commitment of the CNRS (in particular, Groupement de Recherche GDR3280), European Molecular Biology Laboratory (EMBL), Genoscope/CEA, VIB, Stazione Zoologica Anton Dohrn, UNIMIB, Fund for Scientific Research—Flanders, Rega Institute, KU Leuven, The French Ministry of Research. We also thank the support and commitment of Agnès b. and Etienne Bourgois, the Veolia Environment Foundation, Région Bretagne, Lorient Agglomération, World Courier, Illumina, the EDF Foundation, FRB, the Prince Albert II de Monaco Foundation, the *Tara* schooner and its captains and crew. We are also grateful to the French Ministry of Foreign Affairs for supporting the expedition and to the countries who graciously granted sampling permissions. *Tara Oceans* would not exist without continuous support from 23 institutes (<http://oceans.taraexpeditions.org>).

Funding This work was funded by the FunOmics project of the French national program EC2CO-LEFE of CNRS and by the ModelOmics project of the Émergence program of Sorbonne Université, and partly supported by the project MEGALADOM, part of the MASTODON program from the MITI, CNRS France. Emile Faure acknowledges a 3-year Ph.D. grant from the “Interface Pour le Vivant” (IPV) doctoral program of Sorbonne Université. SD Ayata acknowledges the CNRS for her sabbatical year as visiting researcher at ISYEB.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Caron DA, Countway PD, Jones AC, Kim DY, Schnetzer A. Marine protistan diversity. *Annu Rev Mar Sci*. 2012;4:467–93.
2. de Vargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R, et al. Eukaryotic plankton diversity in the sunlit ocean. *Science*. 2015;348:1261605–605.
3. Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, et al. CBOL Protist working group: Barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLOS Biol*. 2012;10:e1001419.
4. Caron DA, Alexander H, Allen AE, Archibald JM, Armbrust EV, Bachy C, et al. Probing the evolution, ecology and physiology of marine protists using transcriptomics. *Nat Rev Microbiol*. 2017;15:6–20.

5. Keeling PJ, Campo J del. Marine protists are not just big bacteria. *Curr Biol*. 2017;27:R541–49.
6. Caron DA. Mixotrophy stirs up our understanding of marine food webs. *Proc Natl Acad Sci*. 2016;113:2806–08.
7. Le Quéré C, Harrison SP, Colin Prentice I, Buitenhuis ET, Aumont O, Bopp L, et al. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob Change Biol*. 2005;11:2016–40.
8. Amacher J, Neuer S, Anderson I, Massana R. Molecular approach to determine contributions of the protist community to particle flux. *Deep Sea Res Part Oceanogr Res Pap*. 2009;56:2206–15.
9. Stoecker DK, Hansen PJ, Caron DA, Mitra A. Mixotrophy in the marine plankton. *Annu Rev Mar Sci*. 2017;9:311–5.
10. Flynn KJ, Stoecker DK, Mitra A, Raven JA, Glibert PM, Hansen PJ, et al. Misuse of the phytoplankton-zooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types. *J Plankton Res*. 2013;35:3–11.
11. Mitra A, Flynn KJ, Tillmann U, Raven JA, Caron D, Stoecker DK, et al. Defining planktonic protist functional groups on mechanisms for energy and nutrient acquisition: Incorporation of diverse mixotrophic strategies. *Protist*. 2016;167:106–120.
12. Esteban GF, Fenchel T, Finlay BJ. Mixotrophy in ciliates. *Protist*. 2010;161:621–41.
13. Selosse M-A, Charpin M, Not F, Jeyasingh P. Mixotrophy everywhere on land and in water: the grand écart hypothesis. *Ecol Lett*. 2017;20:246–63.
14. Ducklow HW, Steinberg DK, Buesseler KO. Upper ocean carbon export and the biological pump. *Oceanogr-Wash DC-Oceanogr Soc*. 2001;14:50–8.
15. Guidi L, Chaffron S, Bittner L, Eveillard D, Larhlimi A, Roux S, et al. Plankton networks driving carbon export in the oligotrophic ocean. *Nature*. 2016;532:465–70.
16. Aumont O, Ethé C, Tagliabue A, Bopp L, Gehlen M. PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geosci Model Dev*. 2015;8:2465–513.
17. Follows MJ, Dutkiewicz S, Grant S, Chisholm SW. Emergent biogeography of microbial communities in a model. *Ocean Sci*. 2007;315:1843–46.
18. Reed DC, Algar CK, Huber JA, Dick GJ. Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proc Natl Acad Sci*. 2014;111:1879–84.
19. Johnson MD. Acquired phototrophy in ciliates: A review of cellular interactions and structural adaptations. *J Eukaryot Microbiol*. 2011;58:185–195.
20. Stoecker DK, Johnson MD, de Vargas C, Not F. Acquired phototrophy in aquatic protists. *Aquat Microb Ecol*. 2009;57:279–310.
21. Flynn KJ, Mitra A. Building the ‘perfect beast’: modelling mixotrophic plankton. *J Plankton Res*. 2009;31:965–92.
22. Ward BA, Follows MJ. Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux. *Proc Natl Acad Sci*. 2016;113:2958–63.
23. Ghyoot C, Flynn KJ, Mitra A, Lancelot C, Gypens N. Modeling plankton mixotrophy: A mechanistic model consistent with the shutter-type biochemical approach. *Front Ecol Evol*. 2017;5:78.
24. Ward BA, Dutkiewicz S, Barton AD, Follows MJ. Biophysical aspects of resource acquisition and competition in algal mixotrophs. *Am Nat*. 2011;178:98–112.
25. Berge T, Chakraborty S, Hansen PJ, Andersen KH. Modeling succession of key resource-harvesting traits of mixotrophic plankton. *ISME J*. 2017;11:212–23.
26. Mitra A, Flynn KJ, Burkholder JM, Berge T, Calbet A, Raven JA, et al. The role of mixotrophic protists in the biological carbon pump. *Biogeosciences*. 2014;11:995–1005.
27. Leles SG, Mitra A, Flynn KJ, Stoecker DK, Hansen PJ, Calbet A, et al. Oceanic protists with different forms of acquired phototrophy display contrasting biogeographies and abundance. *Proc R Soc B Biol Sci*. 2017;284:20170664.
28. Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner H-W, et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol*. 2010;19:21–31.
29. Bik HM, Porazinska DL, Creer S, Caporaso JG, Knight R, Thomas WK. Sequencing our way towards understanding global eukaryotic biodiversity. *Trends Ecol Evol*. 2012;27:233–43.
30. Bittner L, Gobet A, Audic S, Romac S, Egge ES, Santini S, et al. Diversity patterns of uncultured Haptophytes unravelled by pyrosequencing in Naples Bay. *Mol Ecol*. 2013;22:87–101.
31. Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, Raes J, et al. A holistic approach to marine eco-systems biology. *PLoS Biol*. 2011;9:e1001177.
32. Alberti A, Poulain J, Engelen S, Labadie K, Romac S, Ferrera I, et al. Viral to metazoan marine plankton nucleotide sequences from the Tara Oceans expedition. *Sci Data*. 2017;4:170093.
33. Pesant S, Not F, Picheral M, Kandels-Lewis S, Bescot NL, Gorsky G, et al. Open science resources for the discovery and analysis of Tara Oceans data. *Sci Data*. 2015;2:150023.
34. Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, et al. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucl Acids Res*. 2013;41:D597–604.
35. Granéli E, Edvardsen B, Roelke DL, Hagström JA. The ecophysiology and bloom dynamics of *Prymnesium* spp. *Harmful Algae*. 2012;14:260–70.
36. Liu H, Aris-Brosou S, Probert I, de Vargas C. A time line of the environmental genetics of the haptophytes. *Mol Biol Evol*. 2010;27:161–176.
37. Hansen P, Moldrup M, Tarangkoon W, Garcia-Cuetos L, Moestrup Ø. Direct evidence for symbiont sequestration in the marine red tide ciliate *Mesodinium rubrum*. *Aquat Microb Ecol*. 2012;66:63–75.
38. Agatha S, Strüder-Kypke MC, Beran A, Lynn DH. *Pelagostrobilidium neptuni* (Montagnes and Taylor, 1994) and *Strombidium biarmatum* nov. spec. (Ciliophora, Oligotrichea): phylogenetic position inferred from morphology, ontogenesis, and gene sequence data. *Eur J Protistol*. 2005;41:65–83.
39. Jones HLJ, Leadbeater BSC, Green JC. Mixotrophy in marine species of *Chrysochromulina* (Prymnesiophyceae): ingestion and digestion of a small green flagellate. *J Mar Biol Assoc U K*. 1993;73:283.
40. Johnsen G, Dalløkken R, Eikrem W, Legrand C, Aure J, Skjoldal HR. Eco-physiology bio-optics and toxicity of the ichthyotoxic *Chrysochromulina leadbeateri* (Prymnesiophyceae). *J Phycol*. 1999;35:1465–76.
41. Rhodes L, Burke B. Morphology and growth characteristics of *Chrysochromulina* species (Haptophyceae=Prymnesiophyceae) isolated from New Zealand coastal waters. *N Z J Mar Freshw Res*. 1996;30:91–103.
42. Hemleben C, Be AWH, Anderson OR, Tuntivate S. Test morphology, organic layers and chamber formation of the planktonic foraminifer *Globorotalia menardii* (d’Orbigny). *J Foraminifer Res*. 1977;7:1–25.
43. Fehrenbacher JS, Spero HJ, Russell AD. Observations of living non-spinose planktic foraminifers *Neoglobobulimina dutertrei* and *N. pachyderma* from specimens grown in culture. *AGU Fall Meet Abstr*. 2011;41:PP41A-1724.
44. Spero HJ, Parker SL. Photosynthesis in the symbiotic planktonic foraminifer *Orbulina universa*, and its potential contribution to oceanic primary productivity. *J Foraminifer Res*. 1985;15:273–81.
45. Faber WW, Anderson OR, Caron DA. Algal-foraminiferal symbiosis in the planktonic foraminifer *Globigerinella aequilitalis*;

- II, Effects of two symbiont species on foraminiferal growth and longevity. *J Foraminif Res.* 1989;19:185–93.
46. Kuile Bter, Erez J. In situ growth rate experiments on the symbiont-bearing foraminifera *Amphistegina lobifera* and *Amphisorus hemprichii*. *J Foraminif Res.* 1984;14:262–76.
47. Biard T, Bigeard E, Audic S, Poulain J, Gutierrez-Rodriguez A, Pesant S, et al. Biogeography and diversity of Collodaria (Radiolalia) in the global ocean. *ISME J.* 2017;11:1331–44.
48. Ardyna M, Ovidio F, Speich S, Leconte J, Chaffron S, Audic S, et al. Environmental context of all samples from the Tara Oceans Expedition (2009–2013), about mesoscale features at the sampling location. 2017. PANGAEA.
49. Legendre P, Legendre LFJ. Numerical ecology. Elsevier Science, Amsterdam; 1998;197:333.
50. Escouffier Y. Le traitement des variables vectorielles. *Biometrics.* 1973;29:751.
51. Borcard D, Gillet F, Legendre P. Numerical ecology with R. Springer, New York; 2011;176:177.
52. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
53. Longhurst AR. Ecological geography of the sea. Academic Press, San Diego; 1998.
54. Decelle J, Probert I, Bittner L, Desdevises Y, Colin S, de Vargas C, et al. An original mode of symbiosis in open ocean plankton. *Proc Natl Acad Sci.* 2012;109:18000–5.
55. Le Bescot N, Mahé F, Audic S, Dimier C, Garet M-J, Poulain J, et al. Global patterns of pelagic dinoflagellate diversity across protist size classes unveiled by metabarcoding. *Environ Microbiol.* 2016;18:609–26.
56. Wu S, Xiong J, Yu Y. Taxonomic resolutions based on 18S rRNA genes: A case study of subclass Copepoda. *PLoS ONE.* 2015;10:e0131498.
57. Brown EA, Chain FJJ, Crease TJ, MacIsaac HJ, Cristescu ME. Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? *Ecol Evol.* 2015;5:2234–51.
58. Egge E, Bittner L, Andersen T, Audic S, de Vargas C, Edvardsen B. 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. *PLoS ONE.* 2013;8:e74371.
59. Gilbert JA, Field D, Swift P, Thomas S, Cummings D, Temperton B, et al. The taxonomic and functional diversity of microbes at a temperate coastal site: A ‘multi-omic’ study of seasonal and diel temporal variation. *PLoS ONE.* 2010;5:e15545.
60. DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard N-U, et al. Community genomics among stratified microbial assemblages in the ocean’s interior. *Science.* 2006;311:496–503.
61. Arenovski AL, Lim EL, Caron DA. Mixotrophic nanoplankton in oligotrophic surface waters of the Sargasso Sea may employ phagotrophy to obtain major nutrients. *J Plankton Res.* 1995;17:801–20.
62. Safi KA, Hall JA. Mixotrophic and heterotrophic nanoflagellate grazing in the convergence zone east of New Zealand. *Aquat Microb Ecol.* 1999;20:83–93.
63. Moorthi S, Caron DA, Gast RJ, Sanders RW. Mixotrophy: a widespread and important ecological strategy for planktonic and sea-ice nanoflagellates in the Ross Sea, Antarctica. *Aquat Microb Ecol.* 2009;54:269–77.
64. Unrein F, Gasol JM, Massana R. Dinobryon *faculiferum* (Chrysophyta) in coastal Mediterranean seawater: presence and grazing impact on bacteria. *J Plankton Res.* 2010;32:559–64.
65. Sanders RW, Gast RJ. Bacterivory by phototrophic picoplankton and nanoplankton in Arctic waters. *FEMS Microbiol Ecol.* 2012;82:242–53.
66. Calbet A, Martínez RA, Isari S, Zervoudaki S, Nejstgaard JC, Pitta P, et al. Effects of light availability on mixotrophy and microzooplankton grazing in an oligotrophic plankton food web: Evidences from a mesocosm study in Eastern Mediterranean waters. *J Exp Mar Biol Ecol.* 2012;424–425:66–77.
67. Dolan JR, Pérez MT. Costs benefits and characteristics of mixotrophy in marine oligotrichs. *Freshw Biol.* 2000;45:227–38.
68. Biard T, Stemmann L, Picheral M, Mayot N, Vandromme P, Hauss H, et al. In situ imaging reveals the biomass of giant protists in the global ocean. *Nature.* 2016;532:504–7.
69. Probert I, Siano R, Poirier C, Decelle J, Biard T, Tuji A, et al. *Brandtodinium* gen. nov. and *B. nutricula* comb. Nov. (Dinophyceae), a dinoflagellate commonly found in symbiosis with polycystine radiolarians. *J Phycol.* 2014;50:388–99.
70. Stec KF, Caputi L, Buttigieg PL, D’Alelio D, Ibarbalz FM, Sullivan MB, et al. Modelling plankton ecosystems in the meta-omics era. Are we ready? *Mar Genom.* 2017;32:1–17.
71. Dick GJ. Embracing the mantra of modellers and synthesizing omics, experiments and models. *Environ Microbiol Rep.* 2017;9:18–20.
72. Mock T, Daines SJ, Geider R, Collins S, Metodiev M, Millar AJ, et al. Bridging the gap between omics and earth system science to better understand how environmental change impacts marine microbes. *Glob Change Biol.* 2016;22:61–75.
73. Coles VJ, Stukel MR, Brooks MT, Burd A, Crump BC, Moran MA, et al. Ocean biogeochemistry modeled with emergent trait-based genomics. *Science.* 2017;358:1149–54.
74. Shuter B. A model of physiological adaptation in unicellular algae. *J Theor Biol.* 1979;78:519–52.
75. Millette NC, Grosse J, Johnson WM, Jungbluth MJ, Suter EA. Hidden in plain sight: The importance of cryptic interactions in marine plankton. *Limnol Oceanogr Lett.* 2018;3:341–56.
76. Johnson MD, Oldach D, Delwiche CF, Stoecker DK. Retention of transcriptionally active cryptophyte nuclei by the ciliate *Myrionecta rubra*. *Nature.* 2007;445:426–8.
77. Schoener DM, McManus GB. Plastid retention, use, and replacement in a kleptoplastidic ciliate. *Aquat Microb Ecol.* 2012;67:177–87.