DR. ORIOL CANALS (Orcid ID : 0000-0003-1016-1302)

MR. ALEIX OBIOL (Orcid ID : 0000-0002-5475-9827)

Article type : Original Article

Ciliate diversity and distribution across horizontal and vertical scales in the open ocean

Oriol Canals[1], Aleix Obiol[1], Imer Muhovic[1], Dolors Vaqué[1], Ramon Massana[1]

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

Corresponding authors: Oriol Canals (ocanalsd@gmail.com) and Ramon Massana (ramonm@icm.csic.es).

Keywords: Ciliophora, global ocean, plankton size-fractioning, Protists, DNA metabarcoding, deep-sea

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/MEC.15528

ABSTRACT

Ciliates are globally distributed eukaryotic organisms inhabiting virtually all environments on Earth. Although ciliates range from 10 μm to few mm in cell size, they are repeatedly reported in the pico-sized fraction (smaller than 2-3 μm) of molecular surveys. Here, we used existing datasets (BioMarKs and Tara Oceans) with different size fractions to demonstrate that the ciliate pico-sized signal, likely derived from cell breakage during filtration, is informative and reliable to study marine ciliate biodiversity and biogeography. Then, we used sequences from the picoeukaryotic fraction of two circumnavigation expeditions, Malaspina-2010 and Tara Oceans, to give insights into the taxonomic composition and horizontal and vertical distribution of ciliates in the global ocean. Results suggested a high homogeneity of ciliate communities along the ocean surface from temperate to tropical waters, with ciliate assemblages dominated by a few abundant and widely distributed taxa. Very few taxa were found in a single oceanic region, therefore suggesting a high level of ciliate cosmopolitanism in the global ocean. In vertical profiles, ciliates were detected up to 4000 m depth, and a clear vertical community structuring was observed. Our results provided evidence supporting ciliates as deeply integrated organisms in the deep-sea trophic web, where they may play a relevant role as symbionts of metazoans and grazers of prokaryotes and small eukaryotes in the water column and in aggregates.

1. INTRODUCTION

Ciliates are unicellular eukaryotic organisms belonging to the Alveolata supergroup, having cell sizes from about 10 µm to a few millimeters (Lynn, 2008). These organisms colonize nearly all environments on Earth, being found in marine and freshwater plankton and benthos, soils (Foissner, Chao, & Katz, 2008; Lynn, 2008), and extreme environments (Hu, 2014). They display a great variety of ecological strategies, from free-living to symbionts and from heterotrophy to mixotrophy, and have a wide range of optimal preys (Lynn, 2008). In marine environments, free-living ciliates are primarily considered to play an important role as trophic links between primary producers and larger consumers (Gifford, 1991; Porter, Sherr, Sherr, Pace, & Sanders, 1985). In addition, mixotrophic ciliates can significantly contribute to primary production (up to 70% in some cases), and many species display symbiotic interactions with fishes (including some parasites causing important economic losses), invertebrates and even prokaryotes, showing that ciliates are deeply integrated in marine ecosystems (Lynn, 2008). The dominant marine planktonic species belong to the Orders Oligotrichida and Choreotrichida (Class Spirotrichea) (Gómez, 2007; Lynn, 2008; Rychert, Nawacka, Majchrowski, & Zapadka, 2014; Santoferrara, Grattepanche, Katz, & McManus, 2014; Vaqué, Blough, & Duarte, 1997; Vaqué et al., 2014), and are members of nanoplankton (2-20 µm) and microplankton (20-200 µm) communities. Consequently, assessments of ciliate biodiversity by molecular methods have been mostly performed on these two size-fractions (Bachy, Moreira, Dolan, & López-García, 2014; Grattepanche, Santoferrara, McManus, & Katz, 2015; Grattepanche, Santoferrara, McManus, & Katz, 2016; Sun et al. 2017) or on the whole community collected in 0.2 to 0.8 µm pore-size filters (Pitsch et al., 2019; Stock et al., 2013; Sun et al., 2019; Zhou et al., 2017). However, ciliate sequences have been systematically reported in the smallest size fraction, the picoplankton (0.2/0.8-2/3 µm) (Cheung, Chu, Li, Kwan, & Wong, 2008; de Vargas et al., 2015; Giner et al., 2019; Not, del Campo, Balagué, de Vargas, & Massana, 2009; Sørensen, Daugbjerg, & Gabrielsen, 2012), but these have been seldom considered, representing a promising additional resource to further study some ciliate ecology issues like biodiversity, community structure and biogeography.

The relatively recent application of HTS (High Throughput Sequencing) technology to assess microbial diversity has contributed to solve the debate about biogeography of protist species (and by extension ciliates). This debate was moved by two alternative beliefs: i) all protist species are cosmopolitan and respond to the aphorism “everything is everywhere, but the environment selects” (Finlay, Esteban, & Fenchel, 2004), and ii) protists have many cosmopolitan species but also a significant number of endemic species, as proposed by the moderate endemicity model (Foissner, 2006; Foissner et al., 2008). Recent microscopic and molecular studies have

confirmed many cases of cosmopolitanism but also biogeographic patterns for benthic (Azovsky & Mazei, 2013) and planktonic marine ciliates (Agatha, 2011; Dolan & Pierce, 2012; Santoferrara, Rubin, & McManus, 2018). However, most studies on marine ciliates have been performed in neritic waters (shallow regions above continental margins), often across short transects, and mainly focusing on Tintinnina species (loricate Choreotrichida), which due to their external lorica or shell are easier to collect in plankton nets, and to be microscopically identified, than aloricate ciliates (ciliates without an external lorica). Indeed, aloricate ciliates remain less studied (Agatha, 2011), even though they account for the majority of ciliate diversity and abundance in marine waters (Pitta, Giannakourou, & Christaki, 2001; Rychert et al., 2014; Vaqué et al., 1997; Vaqué et al., 2004). Therefore, there is still the need for studies focusing on the whole Ciliophora diversity along the global ocean, as the work of Gimmler, Korn, de Vargas, Audic, & Stoeck (2016) in the context of Tara Oceans expedition. These authors aimed to uncover the environmental factors that drive ciliate diversity in the global sunlit ocean, providing insights on ciliate biogeography and diversity at the group level. However, no data of the most abundant ciliates in the open ocean at a species level were provided.

Albeit most studies on marine ciliates have been performed in surface waters, the presence of deep-sea ciliates has been systematically reported by both microscopic (Agusti et al., 2015; Silver, Gowing, Brownlee, & Corliss, 1984; Sohrin, Imazawa, Fukuda, & Suzuki, 2010; Tanaka & Rassoulzadegan, 2002) and molecular approaches (Countway et al., 2007; Giner et al., 2020; Hu et al., 2018; Sauvadet, Gobet, & Guillou, 2010). Few studies have focused on the distribution of ciliate communities across depth, some being limited to the photic layers, i.e. from surface to about 200 m depth (Christaki et al. 2011; Pitta et al., 2001; Santoferrara, Gómez, & Alder, 2011), and others also providing data below the photic zone (Grattepanche et al., 2016; Sun et al., 2019; Wickham, Steinmair, & Kamennaya, 2011; Zhao, Filker, Xu, Huang, & Zheng, 2017). Only Zhao et al. (2017) provided taxonomic insights on open ocean planktonic ciliates below 2000 m depth. Thus, research is needed to uncover the hidden diversity of ciliates in the dark ocean and to explore their role in the deep ocean food web and their involvement in the carbon pump and other biogeochemical cycles.

Here, we used existing sequencing datasets assessing marine eukaryotic diversity to analyze, in the first place, how fractionated filtration determines Ciliophora detection, in order to evaluate the reliability of using the data from the pico-sized fraction to study some aspects of ciliate ecology. For that purpose, we used two molecular surveys containing pico-, nano- and micro-sized fractions: BioMarKs (Massana et al., 2015) and Tara Oceans (de Vargas et al., 2015). Considering the fragility of most ciliate cells and their previous detection in the pico-sized fraction, we expected this fraction to be at least as informative for ciliate ecology as nano- and micro-sized

fractions are. Second, we aimed to study the diversity and horizontal and vertical distribution of open ocean ciliate taxa using the pico-sized fraction from Malaspina-2010 (Giner et al., 2020; Logares et al., 2020) and Tara Oceans (de Vargas et al.; 2015) circumnavigation expeditions. Since diversity, community composition and biogeography of ciliates in the open sea are still understudied as compared to coastal zones, we expect to provide new insights on these ecological aspects both in a wide surface report and across the water column.

2. MATERIAL and METHODS

2.1. Sequencing datasets and environmental data

Data for the size fractionated analysis include previously published eukaryotic surveys of three size fractions from the BioMarKs project (BMK; Massana et al., 2015) and Tara Oceans circumnavigation expedition (from 2009 to 2013) (Tara-Fract; de Vargas et al., 2015). Pico-sized subsamples (0.8 to 3 µm in BMK and 0.8 to 5 µm in Tara), nano-sized subsamples (3 to 20 µm and 5 to 20 µm), and micro-sized subsamples in Tara (20 to 180 µm) were obtained by sequential filtration. Micro-sized subsamples from BMK (20 to 200 µm) were collected by towing for 5–15 min a plankton net of 20 µm mesh size. Extracted DNA from these subsamples was used to amplify the V4 or V9 region of the 18S rDNA gene. In BMK, V4 sequences were obtained in a 454 platform and clustered in OTUs (Operational Taxonomic Units) by similarity at 97% (Massana et al., 2015). In Tara, V9 sequences were obtained in an Illumina platform and clustered by Swarm (de Vargas et al., 2015). Only subsamples with at least 30 Ciliophora OTUs, a threshold established to avoid putative randomness due to low sequencing depth, were considered: 5 samples from BMK (15 subsamples considering the three fractions) and 43 from Tara (129 subsamples).

Data for the distribution of ciliates in the global ocean derive from the Malaspina expedition (from December 2010 to July 2011) (Logares et al. 2020; Giner et al. 2020), targeting the V4 18S rDNA of pico-sized marine plankton (0.2 to 3 µm), and the Tara Ocean expedition, targeting the V9 of the 0.8 to 5 µm size fraction (de Vargas et al. 2015, Pesant et al. 2015). Environmental data for both expeditions included salinity, nitrate, nitrite, phosphate, turbidity, chlorophyll, oxygen and temperature. The horizontal survey consists of surface ocean samples (~3 m depth) collected during Malaspina (subset named MPN-SRF, 113 samples) and Tara Oceans (Tara-SRF, 41 samples). Stations were attributed to a given oceanic region based on their geographic coordinates. Vertical profiles consist on samples from the Malaspina expedition collected at the surface (~3 m), deep chlorophyll maximum (DCM), mesopelagic zone (200 to 1000 m; some of them in the oxygen minimum zone [OMZ] or the deep scattering layer [DSL]) and bathypelagic

zone (1000 to 4000 m) (Giner et al., 2020). For these profiles, community DNA and RNA from the pico-sized fraction (0.2 to 3 µm) was extracted simultaneously, resulting in the MPN-DNA dataset (62 samples from 12 stations) and the MPN-RNA dataset (91 samples from 13 stations). Malaspina V4 sequences from the three datasets were separately processed using DADA2 (Callahan et al., 2016) to get amplicon sequence variants (ASVs) using the following parameters: pool=TRUE, trunclen=220,210 (240,210 in MPN-SRF) and maxEE=6,8. Finally, the extracted DNA from a subset of vertical profiles was used to amplify the V9 region of the 18S rDNA followed by Illumina sequencing (Obiol et al., 2020). The MPN-V9 dataset comprised 30 samples from 6 stations and was processed by DADA2 using the parameters: pool=TRUE, trunclen=110,90 and maxEE=4,6.

A summary of all sequencing datasets is shown in Table 1. Regardless of the clustering method used, the taxonomic units generated (ASVs, swarms and OTU_{97}) are referred to as OTUs throughout the text. Rarefaction was applied to avoid biases due to different sequencing depths by using the *rrarefy* function of the vegan package (v2.5-4; Oksanen et al., 2019) within the R Statistical environment (R version 3.6.2, R Development Core Team, 2008). The rarefaction threshold for each dataset is specified in Table 1. Overall, 99 samples were kept in MPN-SRF, 41 in Tara-SRF, 40 in MPN-DNA (11 in surface, 10 in DCM, 9 in mesopelagic, and 10 in bathypelagic), 75 in MPN-RNA (13, 13, 31, and 18) and 22 in MPN-V9 (5, 4, 6, and 7) after the rarefaction.

2.2. Taxonomic assignment of Ciliophora sequences

Taxonomic assignment of eukaryotic sequences was performed using BLAST (Altschul, Gish, Miller, Myers, & Lipman, 1990) against the PR2 (Guillou et al., 2013) reference database (V4 and V9 OTUs), and against the *eukaryotesV4* (Obiol et al. 2020) database (V4 OTUs). Sequences initially identified as ciliates were retrieved from the original dataset and validated using the EukRef-Ciliophora database (Boscaro et al., 2018). Ciliate sequences were then classified at a Class level (whenever possible) following the current classification (Adl et al., 2019) based on BLAST searches against the PR2 database. For the V4 dataset, classification into ciliate groups was further confirmed with phylogenetic trees with a subset from EukRef-Ciliophora database covering all Ciliophora classes. For tree construction, query and reference sequences were aligned using MAFFT (Katoh & Standley, 2013), and phylogenetic trees were constructed using RAxML 8.2.12 (Stamatakis, 2014) under the model GTRCATI. The Class Spirotrichea was separated into the lower taxonomic ranks Euplotia, Hypotrichia, Licnophoridae, Oligotrichida, Strobilidiina and Tintinnina, in response to the high relevance of spirotrichs in marine ecosystems. Also, some of these groups possess evident morphological traits that allow easy

comparison of molecular and microscopic data. We used Nassophorea and Prostomatea classes although both are nowadays considered polyphyletic (Boscaro et al., 2018).

To identify the V4 and V9 OTUs that putatively derive from the same ciliate species, we selected the 100 most abundant sequences in MPN-SRF (V4 amplicons) and Tara-SRF (V9 amplicons) datasets and performed a BLAST search to get the closest GenBank reference sequence. These sequences were used to construct a reference phylogenetic tree (by RAxML after a MAFFT alignment), which was then used for the placement of V4 and V9 sequences by the evolutionary placement algorithm (EPA; Berger, Krompass, Stamatakis, 2011), designed for phylogenetic placement of short sequences (data available at Zenodo repository, doi: 10.5281/zenodo.3736633). Based on this tree, we considered that a pair of V4 and V9 OTUs correspond to the same phylotype when i) they matched to a same reference sequence at a similarity >98%, and ii) there was no other reference sequence covering both 18S regions displaying a higher similarity with only one sequence of the query pair.

2.3. Statistical analysis

Statistical analyses were performed within the R Statistical environment (R version 3.6.2). Ordination of communities was carried out using non-metric multidimensional scaling (NMDS) analyses based on Bray–Curtis dissimilarity matrices (*vegdist* function, vegan package v2.5-4). ANOSIM (analysis of similarity; Clarke, 1993) was used to test whether there were significant differences between predefined sample groups (*anosim* function, vegan package). Mantel test (*mantel* function, vegan package) was used to assess whether variations of ciliate community composition correlated with environmental parameters. The effect of each environmental parameter on the ciliate community structure was measured by PERMANOVA (permutational analysis of variance) (*adonis* function, vegan package). Venn diagrams were done to visualize the OTU overlap among fractions and water layers (eulerr package v5.1.0; Larsson, 2019). An IndVal (Indicator Value) analysis (Dufrene & Legendre, 1997) was applied to detect the OTUs significantly related to a given depth layer (labdsv package v1.8-0; Roberts, 2006). Following Logares et al. (2013), OTUs displaying a p-value <0.05 and IndVal values >0.5 were considered. Only OTUs appearing in at least 5 samples were considered for the IndVal analysis. Alpha diversity analyses were based on the Shannon index (H') and the richness (number of OTUs).

3. RESULTS

3.1. Ciliophora signal across size fractionated samples

We analyzed the pico-, nano- and micro-sized fractions of BMK (5 water samples) and Tara-Fract (43 water samples) datasets in order to compare Ciliophora diversity retrieved by HTS in the

different plankton fractions. Despite the large difference in the total number of OTUs detected in BMK (109 OTUs) and Tara-Fract (861 OTUs), in both datasets the pico-sized fraction exhibited the highest alpha diversity across fractions (Figure 1A), both in terms of Richness and Shannon indices, while the micro-sized fraction was the least diverse. This lower diversity in the larger size fraction contrasts with a higher proportion of ciliate reads with respect to the total eukaryotic reads (median of 28.2% in BMK and of 2.4% in Tara-Fract; data not shown). Ordination of subsamples by non-metric multidimensional scaling (NMDS) revealed the grouping of Ciliophora communities based on the size fraction analyzed. In BMK, pico- and nano-sized subsamples grouped together (Figure 1B) while micro-sized ones were spread across the two-dimensional space. In Tara-Fract, the three size fractions grouped clearly separated, with higher cohesion observed in pico- and micro-sized fractions as compared to nano-sized ones. The ANOSIM test statistically supported the grouping of samples by fraction in both BMK (R coefficient: 0.416, p-value: 0.0014) and Tara-Fract (R: 0.645, p-value: 0.0001) datasets.

Oligotrichida, Tintininna and to a lesser extent Strobilidiina and Oligohymenophorea were the Ciliophora groups that showed the clearest variation across fractions in both datasets (Figure 1C). Oligotrichida dominated the pico-sized fraction in BMK dataset (40.3% of OTUs and 74.7% of reads) and was also prominent in Tara-Fract (28.5% of reads), while its richness and relative abundance dramatically decreased in the micro-sized fraction, being absent in BMK and accounting for only 2.6% of reads in Tara-Fract. An opposed pattern was observed for Tintinnina, which dominated in the micro-sized fraction and were scarcely detected in the smallest fraction in both datasets. Thus, Tintinnina represented 9.7 and 4.7% of OTUs and 3.5 and 4.1% of reads in the pico-sized fraction of BMK and Tara-Fract, respectively. In the micro-sized subsamples, Tintinnina represented 56.0 and 15.0% of OTUs and 95.4 and 77.3% of reads, respectively. Strobilidiina in BMK and Oligohymenophorea in Tara-Fract showed a distribution pattern across fractions similar to Oligotrichida.

Venn diagrams were used to evaluate the OTU overlap between the different size-fractions (Figure 1D). The BMK dataset showed a moderate overlap between the two smallest size fractions in terms of OTUs (55 and 58% of OTUs in pico- and nano-sized subsamples, respectively) and a remarkable overlap in terms of relative abundance (89 and 87% of reads in pico- and nano-sized fractions corresponded to OTUs shared by both fractions). In contrast, most OTUs from the micro-sized fraction were exclusive of this fraction (64%, representing 58.7% of reads). OTUs shared by the three fractions represented less than 2% of reads for each fraction. The Tara-Fract dataset, on the other hand, showed a remarkably high OTU overlap between all fractions, especially in terms of amplicon abundance: over 94% of reads of each fraction belonged to OTUs shared by the three planktonic fractions.

Overall, with the only exception of Tintinnina taxa, which are less represented in the smallest size fractions, it appears that the pico-sized fraction provides in general a more diverse information about ciliate community than nano- and micro-sized fractions. Therefore, our results support the pico-sized fraction as an informative and reliable fraction to study ciliate diversity and ecology in marine environments, at least to the same extent nano- and micro-sized fractions are, keeping in mind that some taxa (i.e. Tintinnina) may be undersampled.

3.2. Ciliates in the global ocean

Eukaryotic HTS data from Malaspina (0.2 to 3 μm size fraction) and Tara (0.8 to 5 μm) circumnavigation expeditions were analyzed in order to study Ciliophora distribution in offshore waters from the global ocean. The datasets used here, MPN-SRF (99 samples), Tara-SRF (41 samples), MPN-DNA (40 samples), MPN-RNA (75 samples) and MPN-V9 (22 samples), displayed 416, 609, 191, 1531 and 111 OTUs, respectively (Table 1). These belonged to all Ciliophora classes, i.e. Heterotrichea, Karyorelictea, Armophorea, Litostomatea, Spirotrichea (here divided in Euplotia, Hypotrichia, Licnophoridae, Oligotrichida, Strobilidiina and Tintinnina), Colpodea, Oligohymenophorea, Nassophorea, Plagiopylea, Phyllopharyngea and Prostomatea, as well as to the insertae sedis groups *Protocruzia*, Mesodiniidae and Cariacothrix. Many OTUs belonged to environmental clades, i.e. clades exclusively composed by taxa detected by molecular surveys. Nasso_1, Oligo5 and Phyll_4 (based on Boscaro et al., 2018) were the most relevant environmental clades, representing 96.1%, 78.1% and 88.1% of Nassophorea, Oligohymenophorea and Phyllopharyngea sequences, respectively. All datasets displayed a comparable averaged similarity (from 90.8 to 93.1%) to formally described species, regardless of the 18S gene barcode used or whether sequencing derived from DNA or RNA extracts.

3.2.1 Ciliate distribution in surface waters

MPN-SRF and Tara-SRF datasets were used to study ciliate horizontal distribution in surface waters in all major regions of the non-polar global ocean (plus two samples from the Southern Ocean; Fig. 2A). Ordination of samples on NMDS plots (Figure S1), followed by the ANOSIM test, pointed to a weak influence of geographic oceanic regions on ciliate community composition in Malaspina (R: 0.207, p-value: 0.0001), and moderate in Tara (R: 0.266, p-value: 0.0002). Mantel test analyses revealed similar trends, with a weak correlation between abiotic environmental parameters and ciliate community composition in Malaspina (R: 0.101 and p-value: 0.0081) and a moderate correlation in Tara (R: 0.247 and p-value: 0.002). Based on PERMANOVA tests, environmental data, longitude and latitude explained 18.4% of ciliate community variation in MPN-SRF and 49% in Tara-SRF (Table S1). Temperature appeared as

the parameter that mostly explained ciliate community variation in Tara-SRF. The higher influence of geographic regions and environmental parameters in Tara samples was driven by the clear differentiation of Southern Ocean samples in terms of alfa diversity (Figure S2) and community composition (Figure 2B). Indeed, when Southern Ocean samples were excluded from Tara dataset, the effect of oceanic region on ciliate assemblages notably decreased (R: 0.198, p-value: 0.0034, by ANOSIM test), as well as the correlation between biotic and abiotic data (R: 0.136, p-value: 0.04, by Mantel test), yielding results similar to Malaspina data.

Ciliophora community composition in terms of the OTUs present was notably similar along ocean surface in both Malaspina and Tara datasets, either at Class (Figure 2B) and at lower taxonomic ranks (Table S2). However, based on relative abundances, some regions in the Tara survey exhibited clear particularities. First, ciliate assemblages from the Southern Ocean were distinct from the general survey by having very low abundance of Oligohymenophorea and Nassophorea, and high abundance of Strobilidiina. Second, samples from the Red Sea and Indian Ocean displayed an increase in Colpodea, mostly due to two unique taxa: OTU_556 (100% to *Colpoda steinii*, Colpodida) in the Red Sea, and OTU_565 (100% to *Aristerostoma* sp., Cyrtolophosidida) in North and South Indian Ocean. Overall, Strombidiida (Oligotrichida), the environmental clade Oligo5 (Oligohymenophorea), Nassophorea_X (Nassophorea), Cyrtophoria_1 (Phyllopharyngea), Choreotrichida (Strobilidiina), Tintinnida (Tintinnina) and Mesodiniidae (only detected in Tara dataset due to a deficient match of Mesodiniidae sequences with the reverse V4 primer) arose as the most diverse and/or abundant Ciliophora clades in the ocean surface.

Finally, we aimed to study the distribution of specific taxa along the ocean surface. OTUs detected in all oceanic regions (six regions in MPN-SRF and eight in Tara-SRF, Figure 2) accounted for more than half of the Ciliophora abundance (66.4% in MPN-SRF and 63.3% in Tara-SRF, representing 21.8% and 3.7% of OTUs) (Figure 2C). On the other side, ciliates detected only in one region (at least in two samples) accounted for 8.8% of OTUs in Malaspina and 14.1% in Tara, representing only 1.6% and 0.7% of the reads. Oligotrichida displayed the highest number of ubiquitous OTUs (52.4% in MPN-SRF and 21.4% in Tara-SRF), while ubiquity was also found in Oligohymenophorea, Tintinnina, Strobilidiina, Nassophorea, Phyllopharyngea, Prostomatea and Mesodiniidae. We analyzed the 100 most abundant OTUs from MPN-SRF (V4 region) and Tara-SRF (V9 region) and linked the two 18S regions into a single phylotype (when possible) based on BLAST results and on phylogenetic placements. We obtained an unambiguous match for 14 phylotypes (Table S3), which explained 20.1% of Ciliophora abundance in the whole survey (15.0% in Malaspina and 32.2% in Tara). Eight of them belonged to Oligotrichida, three to Oligohymenophorea, two to Tintinnina and one to Prostomatea (Table S3). Only two phylotypes displayed a 100% identity to a formally described species, specifically

to the oligohymenophorean *Pseudocohnilembus persalinus* and *Cardiostomatella vermiformis*. The homogeneous distribution in the global ocean surface of the six more widespread phylotypes within this list is shown in Figure 3.

3.2.2 Ciliates across the vertical scale

Vertical profiles, from surface to bathypelagic waters (up to 4000 m), were taken during the Malaspina expedition, and the distribution of ciliates with depth was analyzed in three sequencing datasets: MPN-DNA, MPN-RNA and MPN-V9 (Table 1). In all datasets, DCM arose as the most diverse layer, and bathypelagic as the less diverse in Ciliophora community composition (Figure 4A). The two DNA approaches, i.e. MPN-DNA and MPN-V9, displayed a larger decrease in alfa diversity in the aphotic zone than the RNA survey, in which the mesopelagic layer showed diversity values slightly higher than at surface. Ciliophora communities from all vertical profiles grouped according to depth layer in the NMDS ordination plot (Figure 4B). This clustering was statistically supported by ANOSIM test in all datasets, both when samples were grouped in photic (surface and DCM) versus aphotic (mesopelagic and bathypelagic) layers (R: 0.613 in MPN-DNA, 0.694 in MPN-RNA and 0.614 in MPN-V9, p-value: 0.0001) or in the four layers (R: 0.578, 0.660 and 0.548, p-value: 0.0001). A more marked differentiation was observed between surface and DCM assemblages (R: 0.736, 0.795 and 0.444, p-value: 0.0001, 0.0001 and 0.0081) than between mesopelagic and bathypelagic assemblages (R: 0.246, 0.482 and 0.237, p-value: 0.0006, 0.0001 and 0.0192). Based on the NMDS plot, photic ciliate communities showed a higher cohesion than aphotic ciliate assemblages. No differences in alpha diversity and taxonomic composition were seen in communities from the oxygen minimum zone or the deep scattering layer as compared with the rest of mesopelagic samples (data not shown).

Oligotrichida and Oligohymenophorea were the dominant groups along the marine water column, followed by Strobilidiina, Nassophorea and to a lesser extent Phyllopharyngea. A marked (and consistent in the three datasets) variation in community composition was observed from surface to deep layers (Figure 4C and Table S4). In the photic zone, Strombidiida (Oligotrichida), Oligo5 (Oligohymenophorea), Nassophorea_X (Nassophorea), Cyrtophoria_1 (Phyllopharyngea), Choreotrichida (Strobilidiina) and Tintinnida (Tintinnina) were the most abundant and/or diverse clades. Even though both photic layers displayed a similar community composition, an evident increase in Oligo5 was seen in the DCM. Further noticeable differences in the photic zone emerged at a Family level, such as an increase in Leegaardiellidae_B (Choreotrichida) and a decrease in Tontoniidae_A (Strombidiida) with depth. The mesopelagic layer was characterized by a marked increase in both diversity and abundance of Apostomatia (Oligohymenophorea) and to a lesser extent of Scuticociliatia (Oligohymenophorea), coupled to a decrease in

Tontoniidae_B (Strombidiida) abundance. Finally, bathypelagic ciliate assemblages displayed higher relevance of Scuticociliatia (mainly Pseudocohnilembidae and Philasterida) and Peritrichia (Zoothamniidae) clades. In addition, a remarkable peak of Euplotia abundance was obtained in the aphotic zone of the MPN-V9 dataset, especially in the bathypelagic layer (reaching up to 47.8% of Ciliophora abundance). This peak, undetected in the V4 surveys, was due to a unique OTU displaying a 97.3% identity to *Euplotes euryhalinus* (Genbank accession number KP297368). We also observed an increase of *Protocruzia* and to a lesser extent of Colpodea richness and relative abundance with depth.

A moderate Ciliophora OTU overlap between layers was observed in all datasets (Figure 4D). In particular, only 4.7, 12.2 and 8.1% of OTUs (in MPN-DNA, MPN-RNA and MPN-V9, respectively) were detected in all layers, while 36.1, 36.0 and 37.8% of OTUs were exclusively found in one layer (percentages refer to OTUs appearing at least twice in the dataset). IndVal analyses revealed 25, 126 and 15 layer-related OTUs in MPN-DNA, MPN-RNA and MPN-V9 datasets, respectively, and those with the highest IndVal values in each layer are shown in Table 2. Photic layers gathered the majority of these depth-related OTUs (96.0% in MPN-DNA, 74.6% in MPN-RNA and 86.7% in MPN-V9), mainly belonging to Oligotrichida (between 53.3 and 58.7% in the three datasets) and to Oligohymenophorea (between 19.0 and 26.7%). The remaining depth-related OTUs belonged to Nassophorea, Tintinnina, Euplotia, Prostomatea and *Protocruzia* (Table S5). DNA and RNA vertical surveys exhibited a large OTU overlap (Figure S3), as only 20 OTUs out of 191 in MPN-DNA dataset (representing 6.6% of abundance) were not detected in MPN-RNA.

4. DISCUSSION

4.1. Methodological considerations

There are several aspects that may influence the reliability of the results exposed here. First, it is well known that the data obtained in metabarcoding assessments may be dependent on the particular assay performed (Santoferrara, 2019). Here we combined surveys using different 18S rDNA regions (V4 and V9), sequencing technologies (454 and Illumina), nucleic acid templates (DNA and RNA) and OTU delineation (similarity, swarms and error correction). Also, sequencing depth significantly differed between datasets (being quite low in some of them) given the broad aim of the initial surveys. As a consequence, some ecological parameters of ciliate communities, e.g. alfa diversity values (richness and H indices), were very different between datasets. However, the general trends on size-fractionation, biodiversity, community structure and biogeography were comparable among datasets, and this consistency indicates robustness of our results on these issues. A second main concern is that biogeographic analyses from

circumnavigation expeditions like Malaspina and Tara are based in a single station across wide oceanographic transects, so overlooking seasonal or temporal variations as well as putative particular conditions of the community sampled. This leads to the justified concern of up to which extent these few samples can provide useful inferences for the global ocean. Our study includes 140 stations (from two different datasets) collected across a huge oceanic area, and the findings that arise from these data provide a set of general patterns that can indeed be applied to shed valuable light on the diversity and biogeography of marine ciliates at a global scale.

4.2. Ciliate detection across fractions

Although pico-sized ciliates have not been described to date, Ciliophora sequences have been repeatedly detected in the pico-sized fraction by HTS surveys (Cheung et al., 2008; de Vargas et al., 2015; López-García, Rodríguez-Valera, Pedrós-Alió, & Moreira, 2001; Not et al., 2009; Sørensen et al., 2012). Our results indicate that these sequences, somehow considered artifactual, are indeed providing useful information about the diversity and community composition of marine ciliates. Several hypotheses have arisen to explain the presence of ciliate sequences in the pico-sized fraction: a) ciliate cell breakage during filtration (Cheung et al., 2008; Not et al., 2009; Terrado, Vincent, & Lovejoy, 2009); b) membrane flexibility in ciliate cells, allowing them to pass through small filter pores (Cheung et al.; 2008); and c) the collection of extracellular DNA derived from non-living ciliates (Not et al., 2009; Schiaffino et al., 2016; Sørensen et al., 2012; Sørensen, Daugbjerg, & Richardson, 2013). The wide across-fractions sharing in Tara-Fract, where over 94% of Ciliophora abundance in each size fraction corresponded to OTUs detected in the three fractions, supports cell breakage as the most plausible hypothesis primarily explaining the Ciliophora signal in pico-sized subsamples. The strong ciliate signal in the MPN-RNA dataset is also favoring this hypothesis, as extracellular RNA is very labile and therefore the RNA signal likely derives from alive cells (broken and collected in the smallest size fraction). A similar high fraction-sharing was not observed in BMK, probably due to a combination of lower sequencing effort (Table 1) and the different setup to sample the microplankton. Thus, the scarcity of non-Tintininna sequences in BMK microplankton samples may be explained by a higher cell flexibility and breakage imposed by the plankton net tow. Extracellular DNA as the source of Ciliophora signal in the three fractions is likely the least supported hypothesis, since this would imply the existence of high amounts of extracellular DNA adhered to differently sized particles or aggregates that would be retained in filters of different pore size (0.2/0.8, 3/5 and 20µm). Thus, our results do not agree with the statements of Sørensen et al. (2013) that considered extracellular DNA as the main source of ciliate signal in the pico-sized fraction.

Tintinnina and Oligotrichida (both belonging to Spirotrichea class) were the main groups responding to fractioning in both coastal (BMK) and offshore (Tara-Fract) samples. In particular, Tintinnina were mainly found in micro-sized subsamples, while Oligotrichida dominated in the smallest fractions. This agrees with previous observations by Grattepanche et al. (2016). This differential distribution across fractions can be explained by specific morphological features. Tintinnina cells possess an external rigid structure (lorica) generally rather large (>100 µm), which may facilitate their retention in 20 µm pore size filters. On the other hand, Oligotrichida are smaller, naked and non-rigid ciliates, which make them more sensitive to cell breakage or more flexible to pass through small filter pores.

In summary, our results indicate that not only cell size but also morphological features may drive protist detection in different planktonic fractions. In ciliates, our findings support the pico-sized fraction as an informative fraction when a sequential filtering process is applied to samples, at least as informative as larger sized fractions. The pico-sized fraction showed a higher tendency to gather diversity compared to micro-sized subsamples (Figure 1A) and a greater community cohesion than nano-sized ones (Figure 1B, Tara-Fract panel). Nevertheless, since the pico-sized fraction does not optimally detect Tintinnina, a complementary analysis of both pico- (instead of nano-) and micro-sized fractions, or a unique filtering on 0.2 to 0.8 µm pore size filters (e.g. Pitsch et al., 2019; Stock et al., 2013; Sun et al., 2019; Zhou et al., 2017) arise as the best choices to fully assess marine Ciliophora diversity by HTS when microbial biomass is collected by filtration.

4.3. Ciliates in the open ocean surface

Analyses of HTS data from global surface waters indicated a limited influence of oceanic region to ciliate community composition. Thus, planktonic ciliates formed rather homogeneous assemblages along the vast tropical and temperate areas studied here (Atlantic, Pacific, and Indian Oceans, Southern Australian Bight, and Mediterranean and Red seas). The largest community differentiation was found in the few samples from the Southern Ocean, as already reported by Gimmler et al. (2016), and in line with other surveys targeting non-loricate Oligotrichea (Oligotrichida and Strobilidiina) (Agatha et al. 2011) or Tintinnina (Dolan & Pierce, 2013; Santoferrara et al., 2018). This could be attributed to the sudden decrease of water temperature at the Polar Front, which likely led to community differentiation by adaptation to lower temperatures, coupled with the strong Antarctic Circumpolar Current (Griffiths, 2010), which may act as a colonization barrier, preventing dispersion from temperate regions. Although no Arctic samples were included in the present study, a similar pattern would be expected for boreal ciliate communities. Indeed, a global trend of lower microbial diversity in the polar regions than in temperate and tropical zones has been recently reported (Ibarbalz et al., 2019), in addition to the

higher ciliate endemism described in boreal marine regions (e.g. Dolan & Pierce, 2013; Santoferrara et al., 2018).

Results exposed here also confirm previous reports indicating that the most abundant taxa tend to be more widespread than the scarcer ones (de Vargas et al., 2015; Santoferrara et al., 2014). Here, up to 66.4% of Ciliophora abundance in MPN-SRF and 63.3% in Tara-SRF (81.1% excluding Southern Ocean samples) corresponded to ciliate taxa present in all oceanic regions (Figure 2C), thus supporting cosmopolitanism of the most abundant taxa in the ocean surface. This fact was also evidenced by the global distribution of a few abundant phylotypes retrieved in the two datasets (Figure 3, Table S3). The existence of a widely distributed core ciliate community therefore suggests a high level of specialization of these core taxa to the typical oligotrophic conditions of open ocean surface waters.

In the context of microbial biogeography, the moderate endemicity model (Foissner, 2006) proposed that about 30% of morphological and molecular suprageneric taxa might be endemic species. Here, regional endemic Ciliophora (those appearing at least twice in a unique oceanic region) were 8.8% of total OTUs in Malaspina and 14.1% in Tara. These results indicate that open ocean surface waters harbor a lower level of endemicity than the 30% proposed by the moderate endemicity model, or the 33% regional endemism found for aloricate Oligotrichea by Agatha (2011). Comparing with the previous report, we found that only 6.6% of aloricate Oligotrichea OTUs in MPN-SRF and 10.2% in Tara-SRF can be considered as regional endemic. The higher regional endemism reported by Agatha (2011) could be attributed to the lower capacity of detection of morphological versus molecular surveys, coupled with the expected higher endemism in coastal environments than in the open ocean. Molecular approaches have also limitations evaluating diversity, such as inconsistencies between clustering methods or, in some cases, the low resolution of the 18S rDNA for species delimitation (Santoferrara et al., 2018). However, the same authors stated that molecular and morphological methods agreed at the genus level, thus supporting this taxonomic rank in biogeographic studies. Our data, based on OTUs with similarity >98% to described species, confirmed the cosmopolitanism of all genera stated as globally distributed by Agatha (2011) except *Paratontonia* and *Lohmaniella* (not found here). Genera considered restricted to the Northern hemisphere, such as *Cyrtostrombidium*, *Novistrombidium*, *Spirotontonia* and *Varistrombidium*, were also found in the Southern hemisphere, while the distribution of *Lynnella* was expanded from the North Pacific, to the Atlantic and South Indian Oceans.

Finally, we want to highlight the high occurrence of *Pseudocohnilembus persalinus* (Oligohymenophorea, Scuticociliatia) along the ocean surface (Table S3). This is one of the many scuticociliate species able to become opportunistic histophagous parasites of fishes that may

cause serious economic losses due to high fish mortality (Xiong et al., 2015). Scuticociliatosis by *Pseudocohnilembus persalinus* has been reported in the freshwater rainbow trout (Jones, Prosperi-Porta, & LaPatra, 2010) and the marine olive flounder (Kim et al., 2004), both commercially relevant fishes.

4.4. From the surface to the deep-sea

Ciliates have been repeatedly reported inhabiting the water column from surface to bathypelagic depths (Countway et al., 2007; Giner et al., 2020; López-García et al., 2001; Not, Gausling, Azam, Heidelberg, & Worden, 2007; Silver et al., 1984), marine sediments (Bik et al., 2012; Hausmann, Hülsmann, Polianski, Schade, & Weitere, 2002) and particular deep habitats like hydrothermal vents (Sauvadet et al., 2010), cold methane seeps (Takishita, Kakizoe, Yoshida, & Maruyama, 2010) and hypersaline basins (Alexander et al., 2009; Stock et al., 2012; Stock et al., 2013). However, only few studies have focused on planktonic deep-sea ciliates and its vertical distribution in the open ocean. Among them, some reported their presence and abundance (Agusti et al., 2015; Silver et al., 1984; Tanaka & Rassoulzadegan, 2002) while others described (to a lesser or greater extent) its community composition (Grattepanche et al., 2015; Grattepanche et al., 2016; Sohrin et al., 2010; Sun et al., 2019; Zhao et al., 2017). Our results agree with previous findings and support the notion that ciliates are not only present in the deep-sea but also play an active role, as indicated by the RNA signal obtained here in mesopelagic and bathypelagic layers and in previous reports (Hu et al., 2016; Hu et al., 2018; Takishita et al., 2010). Indeed, relevant trophic roles have been attributed to ciliates as grazers of bacteria and nano-flagellates, parasites of fishes and invertebrates, and food source for small metazoans (Silver et al., 1984; Sohrin et al., 2010; Sun et al., 2019; Tanaka & Rassoulzadegan, 2002; Zhao et al., 2017).

Consistent with previous reports for the eukaryotic community (Countway et al., 2007; Giner et al., 2020) and for ciliates assemblages (Grattepanche et al., 2015; Silver et al., 1984; Zhou et al., 2017), our data showed an evident differentiation of Ciliophora communities along the water column. Ciliate assemblages grouped according to light availability (photic versus aphotic layers) and also to the main four layers of the water column: surface, DCM, mesopelagic and bathypelagic, as also reported by Zhou et al. (2017). Furthermore, between 13 and 15% of OTUs in each dataset appeared as indicators of a specific layer. With respect to alpha diversity estimates for ciliates, both DNA and RNA surveys showed the DCM as the most diverse layer and the bathypelagic as the least diverse layer. These results are in line with most previous findings (Christaki et al., 2011; Pitta et al. 2011; Sun et al., 2019; Zhao et al. 2017). Grattepanche

et al. (2016) reported higher ciliate diversity in the mesopelagic (850 m depth) than in photic samples, but these findings were based on a single deep sample.

Our results on ciliate community composition along the depth gradient are also in line with recent studies (Sun et al., 2019; Zhao et al., 2017), all supporting Spirotrichea (Oligotrichida, Strobilidiina and Tintinnina), Oligohymenophorea and Nassophorea as the most diverse and abundant groups in the global ocean. It is noteworthy the consistency between our results, based on pico-sized subsamples, and those from studies where samples were collected by a unique filtration on 0.22 μm (Zhou et al., 2017) or 0.8 μm (Sun et al., 2019) pore size filters. These studies also reported an increase of Apostomatia and Scuticociliatia clades in the aphotic layers. Apostomatia are primarily symbionts or parasites of invertebrate groups like shrimps and cephalopods (Lynn, 2008; Soudienne et al., 2016), while Scuticociliatia, typically considered bacterivores, also harbor many species able to parasite fishes (Lynn, 2008). Our results revealed a clear preference of Tontoniidae_A and Tontoniidae_B (Oligotrichida, Strombidiida) to the photic zone, a fact probably explained by the spread mixotrophy within many genera of these clades like *Laboea*, *Pseudotontonia*, or *Spirotontonia* (Lynn, 2008). In addition, the peak of Zoothamniidae in the bathypelagic layer corresponded to the free-floating *Zoothamnium pelagicum* (Gómez, 2017), an oligotrophic bacterivorous species able to actively move throughout the water column.

The peak of Euplotia in the aphotic zone of the MPN-V9 dataset was surprising since no *Euplotes* species has been previously reported as dominant in marine surveys. Also, *Euplotia* OTUs were practically absent in V4 surveys, most likely due to large insertions in the V4 region of most species, which prevents their PCR amplification prior HTS. The single V9 Euplotia OTU was detected in 5 out of 6 stations at all depths, being particularly prominent in the bathypelagic layer, where it reached over 40% of Ciliophora abundance. Euplotia members are free-living, flattened cells possessing cirrus (composite tuft of somatic cilia forming leg-like structures; Lynn, 2008) in the ventral side of the cell, what allows them to crawl and graze over substrates such as marine snow, i.e. sinking particles from surface layers to the deep-sea. As far as we know, our findings are the first suggesting the pelagic deep-sea as a favorable environment for ciliate species adapted to live on substrates, probably feeding on bacteria and small eukaryotes that conform the marine snow (Alldredge & Silver, 1988; Kiørboe, 2003). Overall, the present study provides new data suggesting ciliates as well-integrated organisms in the understudied trophic web of the deep ocean. Our findings point to deep-sea ciliates as symbionts of various marine groups (e.g. fishes, crustaceans, cephalopods), and grazers of prokaryotes and small eukaryotes in the water column and in marine aggregates.

4.5. Highlighting the unknown

A low sequence similarity of marine ciliate OTUs against formally described species was observed, about 90-93% similarity on average in the different datasets, fact that highlights the poor knowledge on ciliate diversity in marine environments (Liu et al., 2017). This is especially relevant in the open ocean, as exemplified by the large number of OTUs belonging to environmental clades, like Nasso_1, Oligo5 and Phyll_4 (Boscaro et al., 2018), among others. Moreover, this sequence novelty was more noticeable in deep waters (Figure S4), indicating that more unknowns exist in the undersampled and understudied deep ocean. The exposed findings highlight not only the need of further taxonomic efforts by the scientific community, but also the need of funding taxonomic efforts by public agencies, as there will be no advances on taxonomy without taxonomists funding.

5. ACKNOWLEDGMENTS

This project was supported by the Spanish Ministry of Economy and Competitiveness through the projects Consolider Malaspina-2010 (CSD2008–00077) and ALLFLAGS (CTM2016-75083-R). We thank all the scientists that contributed samples and data for this study, in particular Caterina R. Giner, as well as the crew of the R/V BIO Hespérides. We also want to thank Ramiro Logares, Aurélie Labarre and David López (ICM-CSIC) for giving their opinion about the manuscript during interesting and productive discussions. Bioinformatics analyses have been performed at the Marbits platform (ICM-CSIC; https://marbits.icm.csic.es).

6. REFERENCES

1. Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A. … Zhang, Q. (2019). Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *Journal of Eukaryotic Microbiology*, 66(1), 4–119. doi: 10.1111/jeu.12691

2. Agatha, S. (2011). Global Diversity of Aloricate Oligotrichea (Protista, Ciliophora, Spirotricha) in Marine and Brackish Sea Water. *PLoS One*, 6(8), e22466. doi:10.1371/journal.pone.0022466

3. Agusti, S., González-Gordillo, J., Vaqué, D., Estrada, M., Cerezo, M. I., Salazar, G., … Duarte, C.M. (2015). Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. *Nature Communications*, 6, 7608. doi: 10.1038/ncomms8608

4. Alexander, E., Stock, A., Breiner, H. W., Behnke, A., Bunge, J., Yakimov, M. M., & Stoeck, T. (2009). Microbial eukaryotes in the hypersaline anoxic L'Atalante deep-sea basin. *Environmental Microbiology*, 11(2), 360–381. doi: 10.1111/j.1462-2920.2008.01777.x

5. Alldredge, A. L., & Silver, M. W. (1988). Characteristics, dynamics and significance of marine snow. *Progress in Oceanography*, 20(1), 41–82. doi: 10.1016/0079-6611(88)90053-5
6. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. doi: 10.1016/S0022-2836(05)80360-2
7. Amaral-Zettler, L., McCliment, E., Ducklow, H., & Huse, S. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One*, 4(7), e6372. doi: 10.1371/journal.pone.0006372
8. Azovsky, A., & Mazei, Y. (2013). Do microbes have macroecology? Large-scale patterns in the diversity and distribution of marine benthic ciliates. *Global Ecology and Biogeography*, 22(2), 163–172. doi: 10.1111/j.1466-8238.2012.00776.x
9. Bachy, C., Moreira, D., Dolan J. R., & López-García, P. (2014). Seasonal dynamics of free-living tintinnid ciliate communities revealed by environmental sequences from the North-West Mediterranean Sea. *FEMS Microbiology Ecology*, 87(2), 330–342. doi: 10.1111/1574-6941.12224
10. Berger, S. A., Krompass, D., & Stamatakis, A. (2011). Performance, accuracy, and web server for evolutionary placement of short sequence reads under Maximum Likelihood. *Systematic Biology*, 60(3), 291–302. doi: 10.1093/sysbio/syr010
11. Bik, H. M., Sung, W., De Ley, P., Baldwin, J.G., Sharma, J., Rocha-Olivares, A., & Thomas W. K. (2012). Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Molecular Ecology*, 21(5), 1048–1059. doi: 10.1111/j.1365-294X.2011.05297.x
12. Boscaro, V., Santoferrara, L. F., Zhang, Q., Gentekaki, E., Syberg-Olsen, M. J., del Campo, J., & Keeling, P. J. (2018). EukRef-Ciliophora: a manually curated, phylogeny-based database of small subunit rRNA gene sequences of ciliates. *Environmental Microbiology*, 20(6), 2218–2230. doi: 10.1111/1462-2920.14264
13. Callahan, B. J., McMurdie, P. J., Rosen, M.J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583. doi: 10.1038/nmeth.3869
14. [dataset]Canals, O., Obiol, A., Muhovic, I., Vaqué, D., & Massana, R. (2020) Ciliate otutable from BioMarKs, Malaspina and Tara datasets, and the metadata file [Dataset]. Zenodo. doi: 10.5281/zenodo.3736633
15. Cheung, M. K., Chu, K. H., Li, C. P., Kwan, H. S., & Wong, C. K. (2008). Genetic diversity of picoeukaryotes in a semienclosed harbour in the subtropical western Pacific Ocean. *Aquatic Microbial Ecology*, 53, 295–305. doi: 10.3354/ame01247
16. Clarke, K. R. (1993). Nonparametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, 18(1), 117–143. doi: 10.1111/j.1442-9993.1993.tb00438.x

17. Christaki, U., Wambeke, F. V., Lefevre, D., Lagaria, A., Prieur, L., Pujo-Pay, M., … Moutin, T. (2011). Microbial food webs and metabolic state across oligotrophic waters of the Mediterranean Sea during summer. *Biogeosciences*, 8, 1839–1852. doi: 10.5194/bg-8-1839-2011

18. Countway, P. D., Gast, R.J., Dennett, M. R., Savai, P., Rose, J. M., & Caron, D. A. (2007). Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environmental Microbiology*, 9(5), 1219–1232. doi: 10.1111/j.1462-2920.2007.01243.x

19. de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., … Karsenti, E. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348, 1261605. doi: 10.1126/science.1261605

20. Dolan, J. R., & Pierce, R. W. (2012). Diversity and distributions of tintinnids. In Dolan, J. R., Montagnes, D. J. S., Agatha, S., Coats, D. W. & Stoecker, D. K. (Eds.). The Biology and Ecology of Tintinnid Ciliates: Models for Marine Plankton. (pp. 214–243).
Oxford: Wiley-Blackwell.

21. Dufrene, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67(3), 345–366. doi: 10.2307/2963459

22. Finlay, B. J., Esteban, G. F., & Fenchel, T. (2004). Protist Diversity is Different? *Protist*, 155(1), 15–22. doi: 10.1078/1434461000160

23. Foissner, W. (2006). Biogeography and dispersal of micro-organisms: A review emphasizing protists. *Acta Protozoologica*, 45, 111–136.

24. Foissner, W., Chao, A., & Katz, L. A. (2008). Diversity and geographic distribution of ciliates (Protista: Ciliophora). *Biodiversity and Conservation*, 17, 345–363. doi: 10.1007/s10531-007-9254-7

25. Gifford, D.J. (1991). The protozoan–metazoan trophic link in pelagic ecosystems. *The Journal of Protozoology*, 38(1), 81–86. doi: 10.1111/j.1550-7408.1991.tb04806.x

26. Gimmler, A., Korn, R., de Vargas, C., Audic, S., & Stoeck, T. (2016). The Tara Oceans voyage reveals global diversity and distribution patterns of marine planktonic ciliates. *Scientific Reports*, 6, 33555. doi: 10.1038/srep33555

27. Giner, C. R., Balagué, V., Krabberød, A. K., Ferrera, I., Reñé, A., Garcés, E., … Massana, R. (2019). Quantifying long-term recurrence in planktonic microbial eukaryotes. *Molecular Ecology*, 28(5), 923–935. doi: 10.1111/mec.14929

28. Giner, C. R., Pernice, M. C., Balagué, V., Duarte, C. M., Gasol, J. M., Logares, R., &
Massana, R. (2020). Marked changes in diversity and relative activity of picoeukaryotes with depth in the world ocean. *ISME J.* 14:437–449. doi.org/10.1038/s41396-019-0506-9

29. Gómez, F. (2007). Trends on the distribution of ciliates in the open Pacific Ocean. *Acta Oecologica*, 32(2), 188–202. doi: 10.1016/j.actao.2007.04.002

30. Gómez, F. (2017). Motile behaviour of the free-living planktonic ciliate *Zoothamnium pelagicum* (Ciliophora, Peritrichia). *European Journal of Protistology*, 59, 65–74. doi: 10.1016/j.ejop.2017.03.004

31. Grattepanche, J. D., Santoferrara, L. F., McManus, G. B., & Katz, L.A. (2015). Distinct assemblage of planktonic ciliates dominates both photic and deep waters on the New England shelf. *Marine Ecology Progress Series*, 526, 1–9. doi: 10.3354/meps11256

32. Grattepanche, J. D., Santoferrara, L. F., McManus, G. B., & Katz, L. A. (2016). Unexpected biodiversity of ciliates in marine samples from below the photic zone. *Molecular Ecology*, 25(16), 3987–4000. doi: 10.1111/mec.13745

33. Griffiths, H. J. (2010). Antarctic marine biodiversity: what do we know about the Distribution of life in the Southern Ocean? *PLoS One*, 5, e11683. doi: 10.1371/journal.pone.0011683

34. Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, … Christen, R. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, 41, D597–D604. doi: 10.1093/nar/gks1160

35. Hausmann, K., Hülsmann, N., Polianski, I., Schade, S., & Weitere, M. (2002). Composition of benthic protozoan communities along a depth transect in the eastern Mediterranean Sea. *Deep-Sea Research Part I,* 49(11), 1959–1970. doi: 10.1016/S0967-0637(02)00095-X

36. Hu, X. (2014). Ciliates in Extreme Environments. *Journal of Eukaryotic Microbiology*, 61(4), 410–418. doi: 10.1111/jeu.12120

37. Hu S. K., Campbell, V., Connell, P., Gellene, A. G., Liu, Z., Terrado, R., & Caron, D. A. (2016). Protistan diversity and activity inferred from RNA and DNA at a coastal ocean site in the eastern North Pacific. *FEMS Microbiology Ecology*, 92(4), fiw050. doi: 10.1093/femsec/fiw050

38. Hu, S. K., Liu, Z., Alexander, H., Campbell, V., Connell, P.E., Dyhrman, S. T., … Caron, D.A. (2018). Shifting metabolic priorities among key protistan taxa within and below the euphotic zone. *Environmental Microbiology*, 20(8), 2865–2879. doi: 10.1111/1462-2920.14259

39. Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byme, H., … Picheral, M. (2019). Global trends in marine plankton diversity across Kingdoms of life. *Cell*, 179, 1084–1097. doi: 10.1016/j.cell.2019.10.008

40. Jones, S. R. M., Prosperi-Porta, G., & LaPatra, S. E. (2010). First Isolation of Pseudocohnilembus persalinus (Ciliophora: Scuticociliatida) from freshwater-reared

rainbow trout, Oncorhynchus mykiss. *The Journal of Parasitology*, 96(5), 1014–1016. doi: 10.1645/GE-2500.1

41. Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. doi: 10.1093/molbev/mst010

42. Kim, S. M., Cho, J. B., Lee, E. H., Kwon, S. R., Kim, S. K., Nam, Y. K., & Kim, K. H. (2004). Pseudocohnilembus persalinus (Ciliophora: Scuticociitida) is an additional species causing scuticociliatosis in olive flounder Paralichthys olivaceus. *Diseases of Aquatic Organisms*, 62(3), 239–244. doi: 10.3354/dao062239

43. Kiørboe, T. (2003). Marine snow microbial communities: scaling of abundances with aggregate size. Aquatic Microbial Ecology, 33: 67–75. doi: 10.3354/ame033067

44. Larsson, J. (2019). eulerr: Area-proportional euler and Venn diagrams with ellipses. R package version 5.1.0. https://cran.r-project.org/package=eulerr.

45. Liu, W., Jiang, J., Xu, Y., Pan, X., Qu, Z., Luo, X., … Pan, H. (2017). Diversity of free-living marine ciliates (Alveolata, Ciliophora): Faunal studies in coastal waters of China during the years 2011-2016. *European Journal of Protistology*, 61(Pt B), 424–438. doi: 10.1016/j.ejop.2017.04.007

46. Logares, R., Lindstrom, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J. … Bertilsson, S. 2013. Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME Journal*, 7(5), 937–948. Doi: 10.1038/ismej.2012.168

47. Logares, R., Deutschmann, I. M., Junger, P. C., Giner, C. R., Krabberød, A. K., Schmidt, T. S. B., … Massana, R. (2020). Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome,* 8, 55. doi: 10.1186/s40168-020-00827-8.

48. López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C. & Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature, 409(6820), 603–607. doi: 10.1038/35054537

49. Lynn D. (2008). The Ciliated Protozoa. Characterization, classification, and guide to the literature. Springer Science + Business Media B.V. Ontario (Canada).

50. Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C. … de Vargas, C. (2015). Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environmental Microbiology*, 17(10), 4035–4049. doi: 10.1111/1462-2920.12955

51. Not, F., Gausling, R., Azam, F., Heidelberg, J. F., & Worden, A. Z. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environmental Microbiology*, 9(5), 1233–1252. doi: 10.1111/j.1462-2920.2007.01247.x

52. Not, F., del Campo, J., Balagué, V., de Vargas, C., & Massana, R. (2009). New Insights into the Diversity of Marine Picoeukaryotes. *PLoS One*, 4(9), e7143. doi: 10.1371/journal.pone.0007143

53. Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D. … Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan.

54. Obiol, A., Giner, C. R., Sánchez, P., Duarte, C. M., Acinas, S. G., & Massana, R. (2020). A metagenomic assessment of microbial eukaryotic diversity in the global ocean. *Molecular Ecology Resources*, 20(3), 718–731. doi: 10.1111/1755-0998.13147.

55. Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., … Tara Oceans Consortium Coordinators. (2015). Open science resources for the discovery and analysis of Tara Oceans data. *Scientific Data*, 2, 150023. doi: 10.1038/sdata.2015.23

56. Pitsch, G., Bruni, E. P., Forster, D., Qu, Z., Sonntag, B., Stoeck, T., & Posch, T. (2019). Seasonality of planktonic freshwater ciliates: Are analyses based on V9 regions of the 18S rRNA gene correlated with morphospecies counts? *Frontiers in Microbiology*, 10, 248. doi: 10.3389/fmicb.2019.00248

57. Pitta, P., Giannakourou, A., & Christaki, U. (2001). Planktonic ciliates in the oligotrophic Mediterranean Sea: longitudinal trends of standing stocks, distributions and analysis of food vacuole contents. *Aquatic Microbial Ecology*, 24(3), 297–311. doi: 10.3354/ame024297

58. Porter, K. G., Sherr, E. B., Sherr, B. F., Pace, M. & Sanders, R. W. (1985). Protozoa in planktonic food webs. *The Journal of Protozoology*, 32, 409–415. doi: 10.1111/j.1550-7408.1985.tb04036.x

59. R Development Core Team. (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

60. Roberts, D. W. (2016). labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.8-0. https://cran.r- project.org/package=labdsv.

61. Rychert, K., Nawacka, B., Majchrowski, R., & Zapadka, T. (2014). Latitudinal pattern of abundance and composition of ciliate communities in the surface waters of the Atlantic Ocean. *Oceanological and Hydrobiological Studies*, 43(4), 436–441. doi: 10.2478/s13545-014-0161-8

62. Santoferrara, L. (2019). Current practice in plankton metabarcoding: optimization and error management. Journal of Plankton Research, 41(5), 571–582. doi: 10.1093/plankt/fbz041

63. Santoferrara, L., Gómez, M. I., & Alder, V. A. (2011). Bathymetric, latitudinal and vertical distribution of protozooplankton in a cold-temperate shelf (southern Patagonian waters) during winter. *Journal of Plankton Research*, 33(3), 457–468. doi: 10.1093/plankt/fbq128

64. Santoferrara, L. F., Grattepanche, J. D., Katz, L. A., & McManus, G. B. (2014). Pyrosequencing for assessing diversity of eukaryotic microbes: analysis of data on marine planktonic ciliates and comparison with traditional methods. *Environmental Microbiology,* 16(9), 2752–2763. doi: 10.1111/1462-2920.12380

65. Santoferrara, L. F., Rubin, E., & Mcmanus, G. B. (2018). Global and local DNA (meta)barcoding reveal new biogeography patterns in tintinnid ciliates. *Journal of Plankton Research*, 40(3), 209–221. doi: 10.1093/plankt/fby011

66. Sauvadet, A. L., Gobet, A., & Guillou L. 2010. Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep hydrothermal habitats. *Environmental Microbiology*, 12(11), 2946–2964. doi: 10.1111/j.1462-2920.2010.02272.x

67. Schiaffino, M. R., Lara, E., Fernández, L. D., Balagué, V., Singer, D., Seppey, C. C. W., … Izaguirre, I. (2016). Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. *Environmental Microbiology*, 18(12), 5249–5264. doi: 10.1111/1462-2920.13566

68. Silver, M. W., Gowing, M. M., Brownlee, D. C., & Corliss, J. O. (1984). Ciliated protozoa associated with oceanic sinking detritus. *Nature*, 309: 246–248. doi: 10.1038/309246a0

69. Sohrin, R., Imazawa, M., Fukuda, H., & Suzuki, Y. (2010). Full-depth profiles of prokaryotes, heterotrophic nanoflagellates, and ciliates along a transect from the equatorial to the subarctic central Pacific Ocean. *Deep-Sea Research Part II*, 57, 1537–1550. doi: 10.1016/j.dsr2.2010.02.020

70. Sørensen, N., Daugbjerg, N., & Gabrielsen, T. M. (2012). Molecular diversity and temporal variation of picoeukaryotes in two Arctic fjords, Svalbard. *Polar Biology*, 35(4), 519–533. doi: 10.1007/s00300-011-1097-8

71. Sørensen, N., Daugbjerg, N., & Richardson, K. (2013). Choice of pore size can introduce artefacts when filtering picoeukaryotes for molecular biodiversity studies. *Microbial Ecology*, 65(4), 964–968. doi: 10.1007/s00248-012-0174-z

72. Soudienne, D., Florent, I., Dellinger, M., Justine, J. L., Romdhane, M. S., Furuya, H., & Grellier, P. (2016). Diversity of apostome ciliates, *Chromidina* spp. (Oligohymenophorea, Opalinopsidae), parasites of cephalopods of the Mediterranean Sea. *Parasite*, 23, 33. doi: 10.1051/parasite/2016033

73. Stamatakis, A. (2014). RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1212–1213. doi: 10.1093/bioinformatics/btu033

74. Stock, A., Breiner, H. W., Pachiadaki, M., Edgcomb, V., Filker, S., La Cono, V., … Stoeck, T. (2012). Microbial eukaryote life in the new hypersaline deep-sea basin Thetis. *Extremophiles*, 16(1), 21–34. doi: 10.1007/s00792-011-0401-4

75. Stock, A., Edgcomb, V., Orsi, W., Filker, S., Breiner, H. W., Yakimov, M. M., & Stoeck, T. (2013). Evidence for isolated evolution of deep-sea ciliate communities through geological separation and environmental selection. *BMC Microbiology*, 13, 150. doi: 10.1186/1471-2180-13-150

76. Sun, P., Huang, L., Xu, D., Huang, B., Chen, N., & Warren, A. (2017). Marked seasonality and high spatial variation in estuarine ciliates are driven by exchanges between the 'abundant' and 'intermediate' biospheres. *Scientific Reports*, 7(1), 9494. doi: 10.1038/s41598-017-10308-y

77. Sun, P., Huang, L., Xu, D., Warren, A., Huang, B., Wang, Y., … Kong, J. (2019). Integrated space-time dataset reveals high diversity and distinct community structure of ciliates in mesopelagic waters of the Northern South China Sea. *Frontiers in Microbiology*, 10, 2178. doi: 10.3389/fmicb.2019.02178

78. Takishita, K., Kakizoe, N., Yoshida, T., & Maruyama, T. (2010). Molecular evidence that phylogenetically diverged ciliates are active in microbial mats of deep-sea cold-seep sediment. *The Journal of Eukaryotic Microbiology*, 57(1), 76–86. doi: 10.1111/j.1550-7408.2009.00457.x

79. Tanaka, T., & Rassoulzadegan, F. (2002). Full-depth profile (0–2000 m) of bacteria, heterotrophic nanoflagellates and ciliates in the NW Mediterranean Sea: Vertical partitioning of microbial trophic structures. *Deep-Sea Research Part II*, 49(11), 2093–2107. doi: 10.1016/S0967-0645(02)00029-2

80. Terrado, R., Vincent, W. F., & Lovejoy, C. (2009). Mesopelagic protists: diversity and succession in a coastal Arctic ecosystem. *Aquatic Microbial Ecology*, 56, 25–39. doi: 10.3354/ame01327

81. Vaqué, D., Blough, H. A., & Duarte, C. M. (1997). Dynamics of ciliate abundance, biomass and community composition in an oligotrophic coastal environment (NW Mediterranean). *Aquatic Microbial Ecology*, 12, 71–83. doi: 10.3354/ame012071

82. Vaqué, D., Alonso-Sáez, L., Arístegui, J., Agustí, S., Duarte, C. M., Sala, M. M., … Gasol, J. M. (2014). Bacterial production and losses to predators along an open ocean productivity gradient in the Subtropical North East Atlantic Ocean. *Journal of Plankton Research*, 36(1), 198–213. doi: 10.1093/plankt/fbt085

83. Wickham, S. A., Steinmair, U., Kamennaya, N. (2011). Ciliate distributions and forcing factors in the Amundsen and Bellingshausen Seas (Antarctic). *Aquatic Microbial Eocology*, 62(3), 215–230. doi: 10.3354/ame01468

84. Xiong, J., Wang, G., Cheng, J., Tian, M., Pan, X., Warren, A., … Miao, W. (2015). Genome of the facultative scuticociliatosis pathogen *Pseudocohnilembus persalinus* provides insight into its virulence through horizontal gene transfer. *Science Reports*, 5, 15470. doi: 10.1038/srep15470

85. Zhao, F., Filker, S., Xu, K., Huang, P., & Zheng, S. (2017). Patterns and drivers of vertical distribution of the ciliate community from the surface to the abyssopelagic zone in the Western Pacific Ocean. *Frontiers in Microbiology*, 8: 2559. doi: 10.3389/fmicb.2017.02559

Data accessibility statement

The data that support the findings of this study are openly available in the open-access repository Zenodo at http://doi.org/10.5281/zenodo.3736633.

Author Contributions

O. Canals and R. Massana conceived and designed the study. O. Canals and A. Obiol analyzed the data. O. Canals wrote the manuscript. I. Muhovic and A. Obiol contributed to the analytical and plotting tools. R. Massana and D. Vaqué reviewed the study and provided relevant conceptual contributions. All authors have seen and approved the manuscript prior to submission.

Figure legends and Table captions

Figure 1. Assessment of ciliate diversity by HTS in three size fractions (P: pico-, from 0.2/0.8 to 3/5 μm; N: nano-, from 3/5 to 20 μm; M: micro-, from 20 to 180/200 μm) from planktonic communities in coastal (BMK, left panels, 5 subsamples for each size fraction) and open ocean samples (Tara-Fract, right panels, 43 subsamples for each size fraction). A: Boxplots displaying the richness and H diversity index values in the three size fractions. B: Placement of ciliate assemblages in each size fraction in a non-metric multidimensional scaling (NMDS) plot. Stress values: 0.032 (BMK) and 0.176 (Tara-Fract). C: Ciliate community composition at a group level in each size-fraction, based on OTU number and the relative abundance of reads. D: Venn diagram displaying the OTU overlap between fractions.

Figure 2. Location of sampling stations and taxonomic composition of surface ciliate communities in Malaspina (left panels) and Tara Oceans (right panels) datasets. A: Geographic locations of sampling stations during the Malaspina cruise (circles with a black border) and the Tara expedition (circles with a white border). NAO: North Atlantic Ocean, MS: Mediterranean Sea, RS: Red Sea, NIO: North Indian Ocean, SIO: South Indian Ocean, SAO: South Atlantic Ocean, SO: Southern Ocean, SPO: South Pacific Ocean, SAB: South Australian Bight, NPO: North Pacific Ocean. B: Percentage of OTUs (and their relative read abundance) of the dominant taxonomic groups in each oceanic region. C: Proportion of ciliate OTUs (together with their relative read abundance) appearing from one to all oceanic regions sampled.

Figure 3. Distribution maps of the six most widespread phylotypes in the ocean surface for which an unambiguous match between the V4 region (in MPN-SRF) and the V9 region (Tara-SRF) was obtained. The text above the maps indicates the GenBank accession number of each phylotype and its closest cultured match with its percentage similarity (in parenthesis). Bubble size indicates the relative abundance of each phylotype in each station with respect to the total ciliate abundance (absences are not shown), scaled from the minimum to maximum in each plot.

Figure 4. Ciliate assemblages in oceanic depth layers, surface (SRF), deep chlorophyll maximum (DCM), mesopelagic (Meso) and bathypelagic (Bathy), as derived from the MPN-DNA (left panels), MPN-RNA (middle panels) and MPN-V9 (right panels) datasets. A: Boxplots displaying the richness and H diversity index values in each layer. B: Placement of ciliate assemblages by layer in a non-metric multidimensional scaling (NMDS) plot. Stress values: 0.170 (MPN-DNA), 0.137 (MPN-RNA) and 0.114 (MPN-V9). One sample was removed from the analysis in MPN-

DNA (D2223) and MPN-RNA (R2223), since it did not allow the correct visualization of the ordination. C: Ciliate community composition at a group level in each layer, based on OTU number and the relative abundance of reads. D: Venn diagram displaying the OTU overlap between layers.

Table 1. Summary of the datasets used in the present study. †: the number of samples includes all depths from a given station, and the number between parentheses refers to the number of fractionated subsamples. SD: standard deviation.

Table 2. List of the five most relevant depth-related OTUs in each layer (chosen according to the IndVal statistic value).

Supplementary Information

Figure S1. Placement of ciliate assemblages from MPN-SRF and Tara-SRF in a non-metric multidimensional scaling (NMDS) plot by oceanic region. Stress values: 0.232 (MPN-SRF) and 0.169 (Tara-SRF). NAO: North Atlantic Ocean, MS: Mediterranean Sea, RS: Red Sea, NIO: North Indian Ocean, SIO: South Indian Ocean, SAO: South Atlantic Ocean, SO: Southern Ocean, SPO: South Pacific Ocean, SAB: South Australian Bight, NPO: North Pacific Ocean.

Figure S2. Boxplots displaying the richness and H diversity index values in each oceanic region from MPN-SRF and Tara-SRF datasets. NAO: North Atlantic Ocean, MS: Mediterranean Sea, RS: Red Sea, NIO: North Indian Ocean, SIO: South Indian Ocean, SAO: South Atlantic Ocean, SO: Southern Ocean, SPO: South Pacific Ocean, SAB: South Australian Bight, NPO: North Pacific Ocean.

Figure S3. OTU overlap between MPN-DNA and MPN-RNA surveys. Number in parenthesis refers to the percentage of ciliate abundance these OTUs represent.

Figure S4. Average of sequence similarity against reference sequences of formally described ciliate species by layer in the MPN-DNA, MPN-RNA and MPN-V9 vertical profiles.

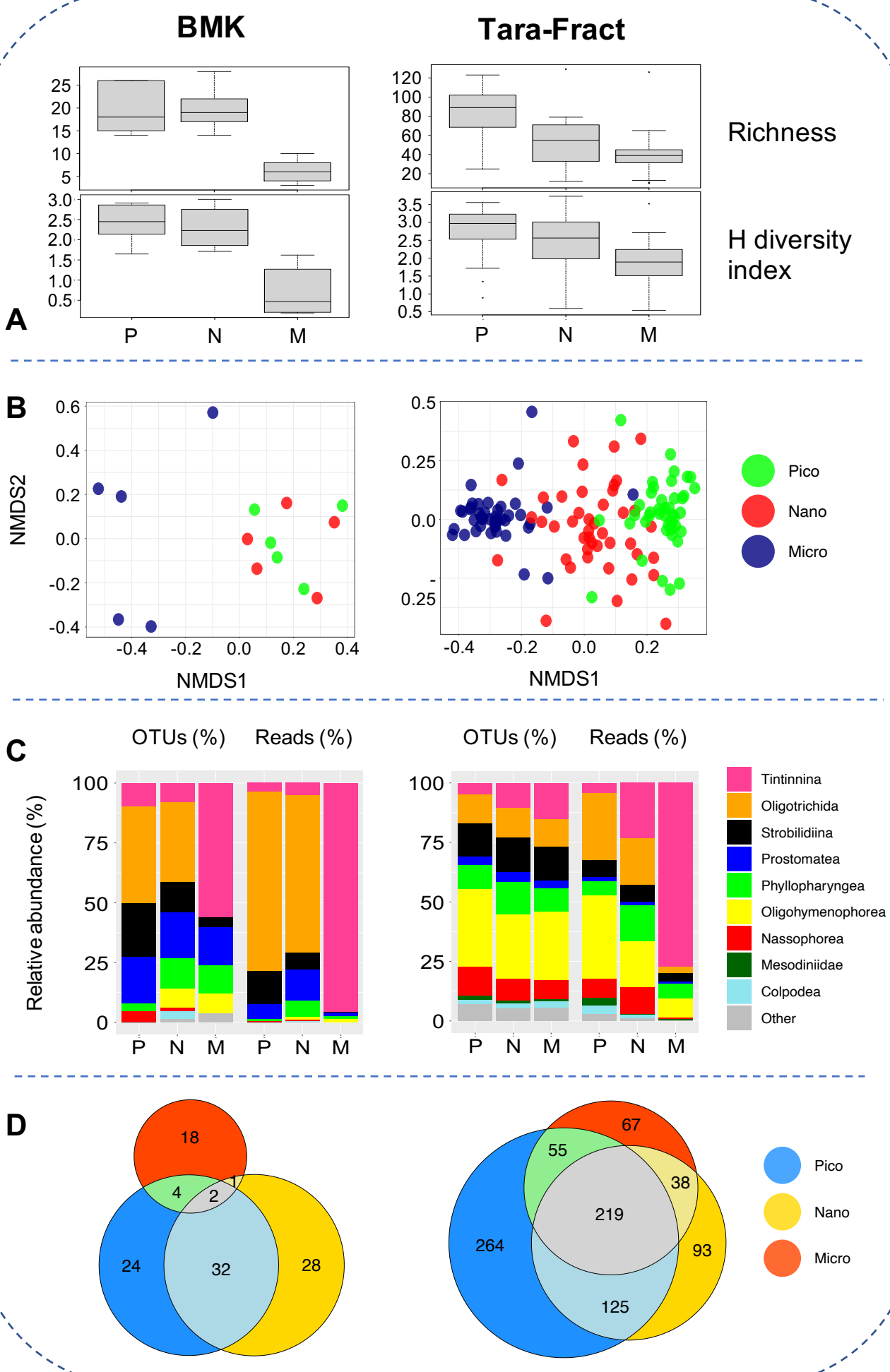
Table S1. Results of PERMANOVA tests. It shows to what extent longitude, latitude and environmental parameters explain the variations in ciliate community composition from MPN-SRF and Tara-SRF. Degree of freedom is 1 in all cases.

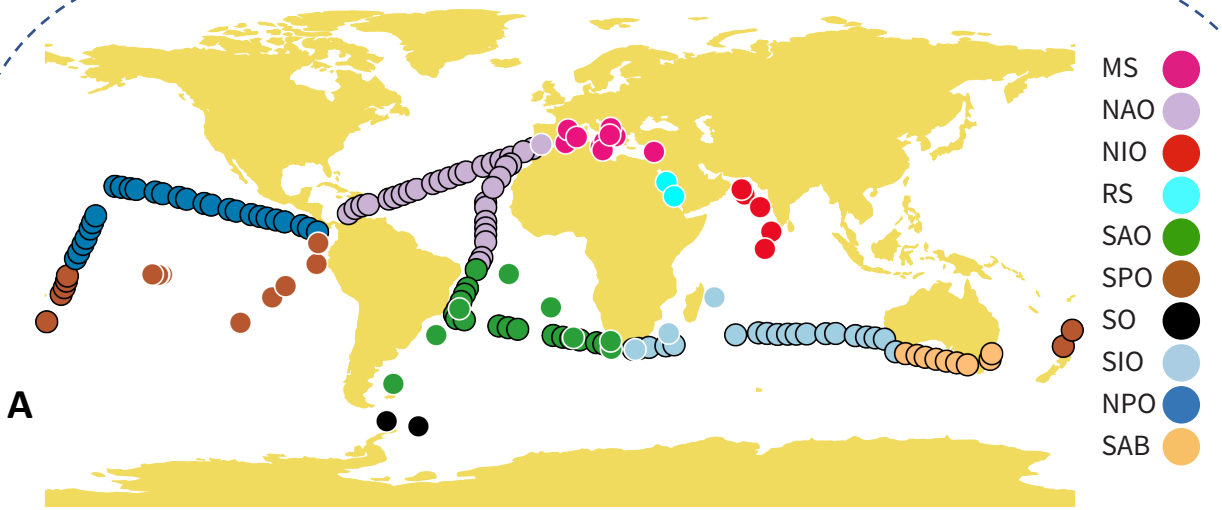
Table S2. A detailed report of the percentage of OTUs and their relative read abundance within Order and Family taxonomic levels for surface samples of each oceanic region of the MPN-SRF and the Tara-SRF datasets. NAO: North Atlantic Ocean, SAO: South Atlantic Ocean, SIO: South Indian Ocean, SAB: South Australian Bight, SPO: South Pacific Ocean, NPO: North Pacific Ocean.

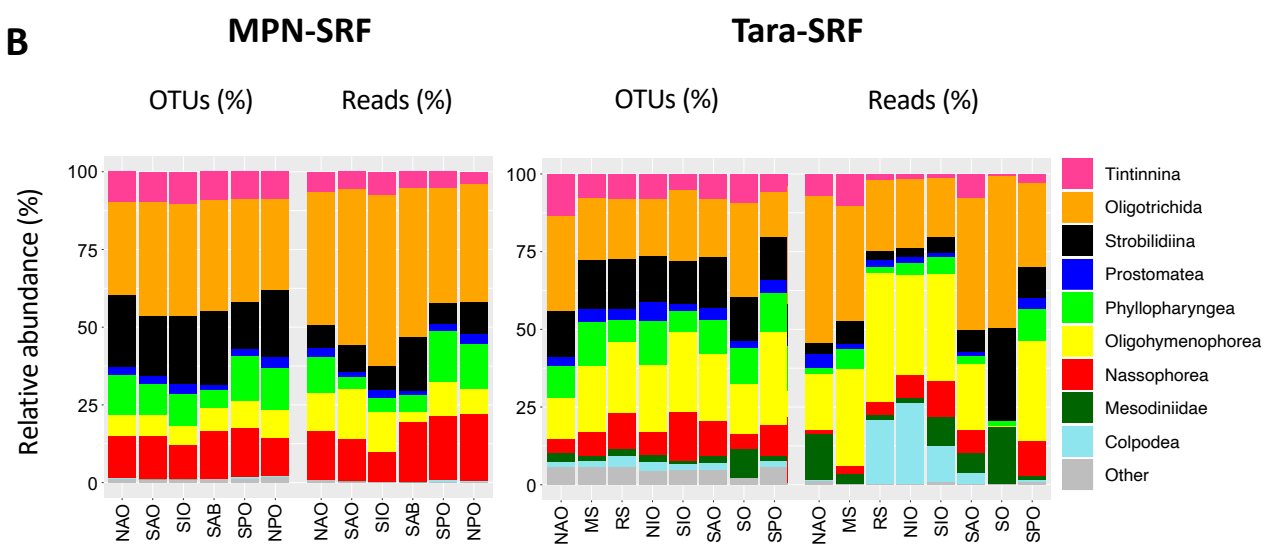
Table S3. List of the 14 most abundant and widely distributed phylotypes in the ocean surface retrieved by linking V4 and V9 sequences from MPN-SRF and Tara-SRF datasets.

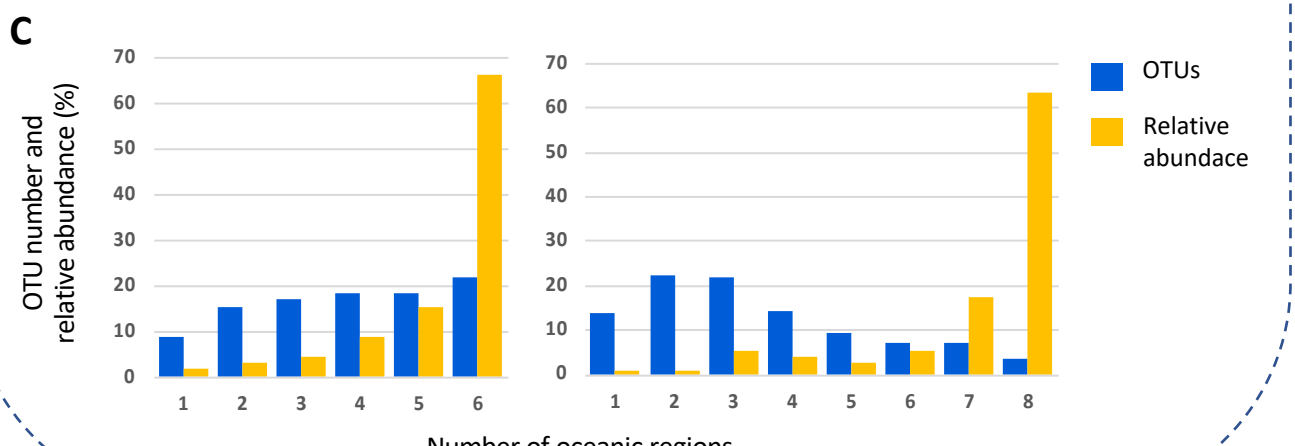
Table S4. A detailed report of the percentage of OTUs and their relative read abundance within Order and Family taxonomic levels for vertical profiles (MPN-DNA, MPN-RNA and MPN-V9) datasets.

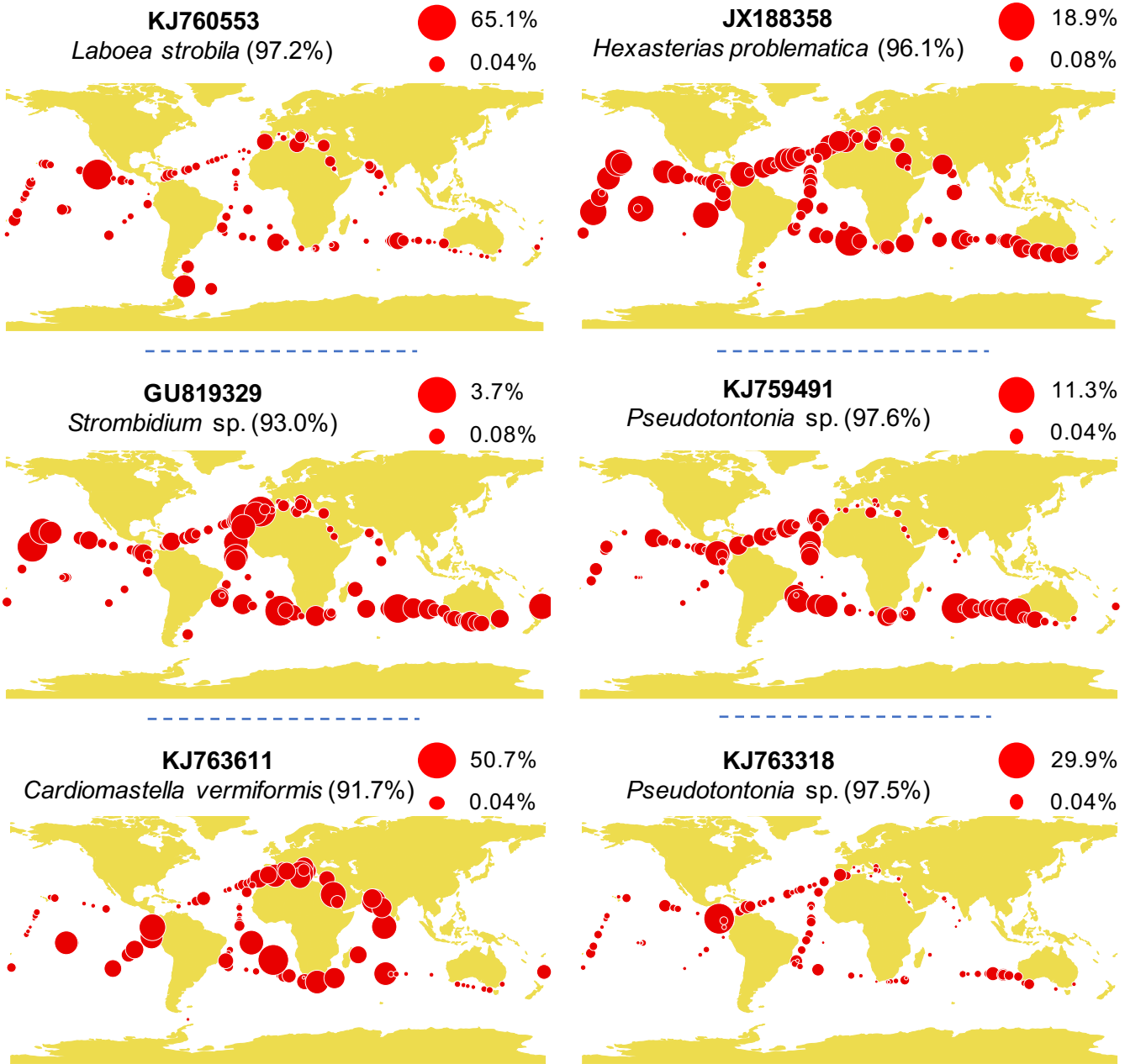
Table S5. Full list of the depth-related OTUs in each layer for the three vertical datasets retrieved by the IndVal analysis. In V4 sequences, † means that the OTU is depth-related in both MPN-DNA and MPN-RNA datasets.

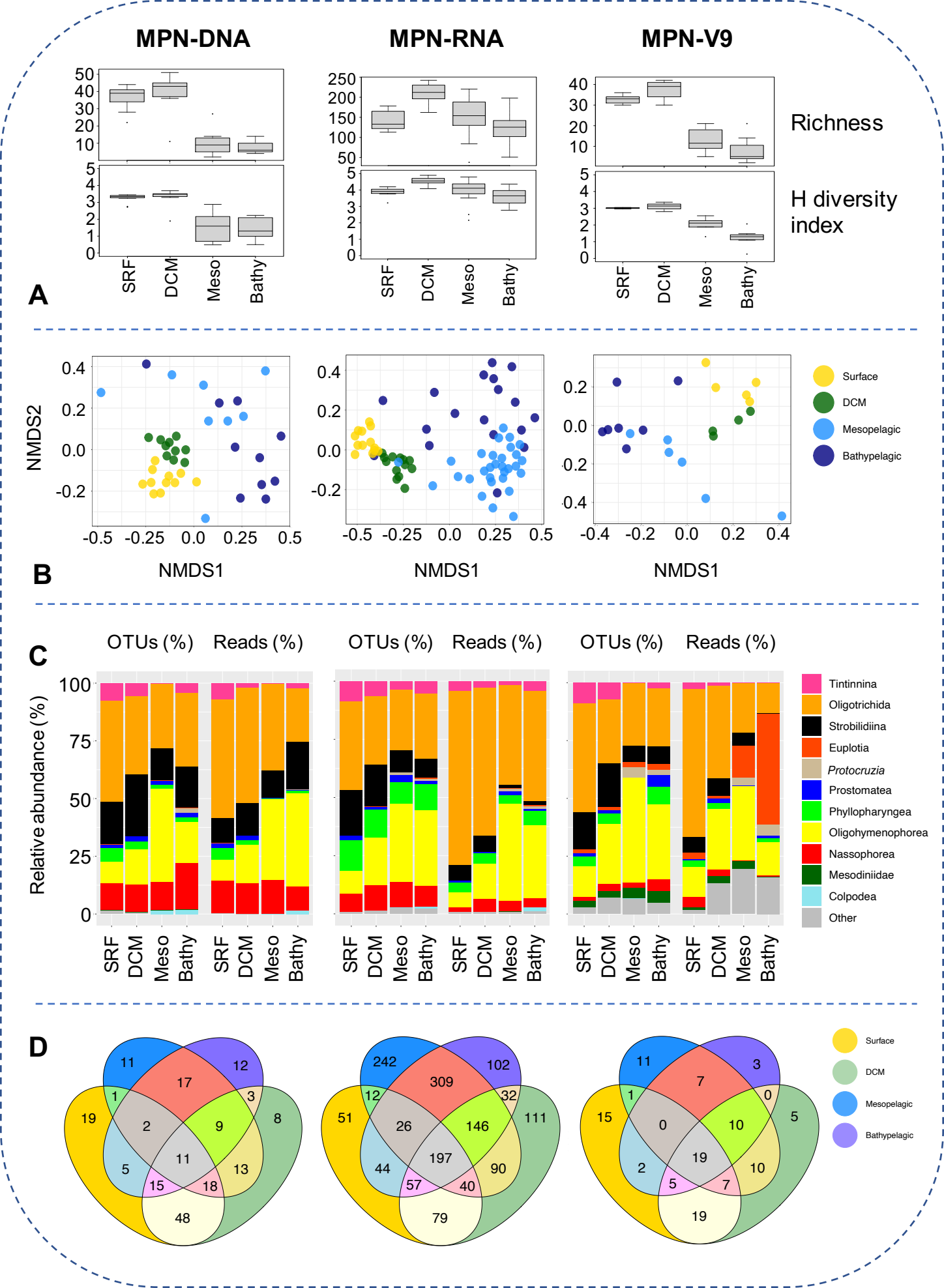

BMK
Tara-Fract
Richness
H diversity index
A
P
N
M
B
NMDS1
NMDS2
Pico
Nano
Micro
C
OTUs (%)
Reads (%)
Relative abundance (%)
Tintinnina
Oligotrichida
Strobilidiina
Prostomatea
Phyllopharyngea
Oligohymenophorea
Nassophorea
Mesodiniidae
Colpodea
Other
D
18
4
2
1
24
32
28
67
55
38
219
264
93
125


A
MS
NAO
NIO
RS
SAO
SPO
SO
SIO
NPO
SAB


B
MPN-SRF
Tara-SRF
OTUs (%)
Reads (%)
OTUs (%)
Reads (%)
Relative abundance (%)
100
75
50
25
0
NAO
SAO
SIO
SAB
SPO
NPO
NAO
SAO
SIO
SAB
SPO
NPO
NAO
MS
RS
NIO
SIO
SAO
SO
SPO
NAO
MS
RS
NIO
SIO
SAO
SO
SPO
Tintinnina
Oligotrichida
Strobilidiina
Prostomatea
Phyllopharyngea
Oligohymenophorea
Nassophorea
Mesodiniidae
Colpodea
Other


C
OTU number and relative abundance (%)
70
60
50
40
30
20
10
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
Number of oceanic regions
OTUs
Relative abundace


KJ760553
Laboea strobila (97.2%)
65.1%
0.04%
JX188358
Hexasterias problematica (96.1%)
18.9%
0.08%
GU819329
Strombidium sp. (93.0%)
3.7%
0.08%
KJ759491
Pseudotontonia sp. (97.6%)
11.3%
0.04%
KJ763611
Cardiomastella vermiformis (91.7%)
50.7%
0.04%
KJ763318
Pseudotontonia sp. (97.5%)
29.9%
0.04%


MPN-DNA
MPN-RNA
MPN-V9
Richness
H diversity index
SRF
DCM
Meso
Bathy
A
NMDS2
NMDS1
B
Surface
DCM
Mesopelagic
Bathypelagic
C
OTUs (%)
Reads (%)
Relative abundance (%)
Tintinnina
Oligotrichida
Strobilidiina
Euplotia
Protocruzia
Prostomatea
Phyllopharyngea
Oligohymenophorea
Nassophorea
Mesodiniidae
Colpodea
Other
D