Science Advances
AAAS

Supplementary Materials for

Global biogeography of the smallest plankton across ocean depths

Pedro C. Junger *et al.*

Corresponding author: Pedro C. Junger, pedro.junger@gmail.com; Ramiro Logares, ramiro.logares@icm.csic.es

Sci. Adv. **9**, eadg9763 (2023)
DOI: 10.1126/sciadv.adg9763

This PDF file includes:

Supplementary Text

Methods

Dataset, sampling, and analytical methods

We compiled a dataset (Fig. 1) composed of 451 samples from surface (3 m depth) to deep waters (up to 4,800 m), covering three depth zones of the ocean: epi- (0-200 m – including DCM), meso- (200-1,000 m), and bathypelagic (1,000-4,000 m). This dataset combines samples obtained during two oceanographic expeditions with similar sampling strategies: *i)* the *Malaspina-2010* circumglobal expedition (*40*) from which we included 263 samples collected between December 2010 and July 2011 in 120 stations distributed along the tropical and subtropical portions (latitudes between 35° N and 40° S) of the Pacific, Atlantic, and Indian oceans (Fig. 1); and *ii)* the *HotMix* trans-Mediterranean cruise (*10*, *55*) from which we considered 188 samples collected between April and May 2014 in 29 stations distributed along the whole Mediterranean Sea (from -5° W to 33° E) and the adjacent Northeast Atlantic Ocean (Fig. 1A). This dataset therefore allows the comparison of the tropical and subtropical ocean (samples hereafter called "open ocean") to a semi-enclosed basin such as the Mediterranean Sea, which displays unique features such as higher temperature and salinity as well as lower nutrient concentration than the open ocean, particularly in the meso- and bathypelagic (Fig. 1B). The *Malaspina-2010* dataset contains 13 stations where the whole vertical profile was sampled (VP stations in Fig. 1). A detailed vertical distribution of the samples is available in Fig. S1. Due to the differences in the sampling size between depth zones, we also generated a dataset with a standardized number of samples (n=39) evenly distributed across space (Fig. S5 and Fig. S6).

The used ASV datasets include contextual data that considers 6 standardized environmental variables (temperature, salinity, fluorescence, PO_4^{3-}, NO_3^-, and SiO_2) as well as prokaryotic and picoeukaryotic abundances determined by flow cytometry and bacterial activity measurements. Water samples were obtained with 20L (in *Malaspina*) or 12L (in *HotMix*) Niskin bottles attached to a rosette sampler equipped with a conductivity–temperature–depth (CTD) profiler (except surface samples in *Malaspina*, that were obtained with individual 30 L bottles, not attached to the rosette). Vertical profiles of temperature, conductivity, and fluorescence were continuously recorded throughout the water column with the CTD sensors. Conductivity measurements were converted into practical salinity scale values. Inorganic nutrients (NO_3^-, PO_4^{3-}, SiO_2) were measured from the Niskin bottle samples with standard spectrophotometric protocols (*106*), using a Skalar autoanalyzer SAN++, as described in (*40*, *53*). Missing nutrient concentration values were extracted from the World Ocean Database (*107*). Prokaryotic populations and phototrophic picoeukaryotes abundances were enumerated using a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA, USA) as detailed elsewhere (*108*). Prokaryotic heterotrophic activity was estimated using the centrifugation method and measuring ^{3}H-leucine incorporation (*109*). For deep water samples, we used the filtration method with a larger volume and undiluted hot leucine. Significant differences in microbial abundances and bacterial activity between depth

zones were tested with an analysis of variance (ANOVA), followed by a Tukey post-hoc test.

To obtain picoplankton biomass, between ~4–12 L of seawater was pre-filtered with a 200-µm net mesh (to remove large organisms and particles). *Malaspina* samples were then sequentially filtered through a 20 µm nylon mesh followed by 3-µm and 0.2-µm polycarbonate filters (47-mm for surface and 142-mm diameter for vertical profiles; Isopore, Merck Millipore, Burlington, MA, USA) using a peristaltic pump. *HotMix* samples were sequentially filtered through 47-mm polycarbonate filters with 3-µm mesh size (Isopore, Merck Millipore) and finally filtered through 0.2-µm Sterivex units. Filters were flash-frozen in liquid N_2 and stored at −80 °C until DNA extraction. Here, only the free-living 'picoplankton' size-fraction (0.2–3 µm) was used in downstream analyses.

Nucleic acid extraction, sequencing, and bioinformatics

DNA extraction was conducted with a standard phenol-chloroform protocol (*87*) for the *Malaspina* surface samples. DNA from the *Malaspina* vertical profiles samples was extracted using the Nucleospin RNAkit (Macherey-Nagel) plus the Nucleospin RNA/DNA Buffer Set (Macherey-Nagel) procedures. *HotMix* DNA samples were extracted using the PowerWater Sterivex™ DNA isolation Kit (MO BIO Laboratories). DNA extracts were quantified with Qubit 1.0 (Thermo Fisher Scientific) and preserved at −80 °C. The same extracts were used for both the 16S and 18S rRNA-gene amplification, and all samples were sequenced with the same prokaryotic and eukaryotic primers. The hypervariable V4–V5 (≈ 400 bp) region of the 16S rRNA gene was PCR amplified with the primers 515F-Y (5'-GTGYCAGCMGCCGCGGTAA) - 926R (5'-CCGYCAATTYMTTTRAGTTT) to target prokaryotes – both Bacteria and Archaea (*11*). The hypervariable V4 region of the 18S rRNA gene (≈ 380 bp) was PCR amplified with the primers TAReukFWD1 (5'-CCAGCASCYGCGGTAATTCC-3') and TAReukREV3 (5'-ACTTTCGTTCTTGATYRA-3') to target eukaryotes (*89*). PCR amplification was carried out with a QIAGEN HotStar Taq master mix (Qiagen Inc., Valencia, CA, USA). Amplicon libraries were then paired-end sequenced on an Illumina (San Diego, CA, USA) MiSeq platform (2 × 250 bp or 2 × 300 bp) at the Research and Testing Laboratory facility, Texas, USA (https://rtlgenomics.com/), or Genoscope (France), for the *Malaspina* deep sea dataset. See details about gene amplification and sequencing in (*10*, *15*).

Raw Illumina MiSeq reads (2x250 or 2x300) were processed using DADA2 (*90*) to determine amplicon sequence variants (ASVs). For the 16S rRNA gene, forward reads were trimmed at 220 bp and reverse reads at 200 bp, whilst for the 18S rRNA gene, we trimmed the forward reads at 240 bp and the reverse reads at 180 bp. Then, for the 16S, the maximum number of expected errors (maxEE) was set to 2 for the forward reads and to 4 for the reverse reads, while for the 18S, the maxEE was set to 7 and 8 for the forward and reverse reads respectively. Error rates for each possible nucleotide substitution type were estimated using a machine learning approach implemented in DADA2 for both the 16S and 18S. Finally, DADA2 was used to estimate error rates for both the 16S and 18S genes in order to delineate the ASVs

Prokaryotic ASVs were assigned taxonomy using the naïve Bayesian classifier method (*91*) alongside the SILVA v.132 database (*92*) as implemented in DADA2, while Eukaryotic ASVs were BLASTed (*93*) against the Protist Ribosomal Reference database [PR^2, version 4.11.1; (*94*)]. Eukaryotes, chloroplasts, and mitochondria were removed from the 16S ASVs table, while Streptophyta, Metazoa, and nucleomorphs were removed from the 18S ASVs table. Both, the 16S and 18S ASVs tables were rarefied to 20,000 reads per sample with the function *rrarefy* from the Vegan R package. To be consistent with our previous study (*15*), for the calculation of ecological processes and associated analysis, ASVs with total abundances < 100 reads across all samples were removed to avoid biases. This filtering procedure removed ~5% of the total reads and ~90% of the total ASVs from both the 16S and the 18S rRNA gene datasets.

Computing analyses were conducted at the MARBITS bioinformatics platform of the Institut de Ciències del Mar (ICM; http://marbits.icm.csic.es). Sequences are publicly available at the European Nucleotide Archive (http://www.ebi.ac.uk/ena) under accession numbers PRJEB23913 [18S rRNA genes] & PRJEB25224 [16S rRNA genes] for the *Malaspina* surface dataset; PRJEB23771 [18S rRNA genes] & PRJEB45015 [16S rRNA genes] for the *Malaspina* vertical profiles dataset; PRJEB45011 [16S rRNA genes] for the *Malaspina* deep sea dataset, and PRJEB44683 [18S rRNA genes] & PRJEB44474 [16S rRNA genes] for the *HotMix* dataset.

Phylogenetics

Phylogenetic trees were built for both the 16S and 18S rRNA gene datasets using the ASVs' sequences. ASV sequences were aligned using mothur (*110*) and aligned templates: SILVA template – for the 16S rRNA gene – and PR^2 template – for the 18S rRNA gene. Poorly aligned regions or sequences were removed using trimAl (parameters: -gt 0.3 -st 0.001) (*111*). Aligned sequences were also visually curated with seaview v4 (*112*) and sequences with >=40% of gaps were removed. Finally, phylogenetic trees were inferred from the aligned quality-filtered sequences using FastTree v2.1.9 (*113*). Additional phylogenetic analyses were carried out with the R package *picante* (*114*).

Environmental heterogeneity, water masses characterization, and least-cost distance calculations

We calculated the average pairwise dissimilarity (*EnvHt*) as an index of environmental heterogeneity based on the main standardized environmental variables that we measured: temperature, salinity, fluorescence, PO_4^{3-}, NO_3^-, and SiO_2. We firstly computed an Euclidean distance matrix for each depth zone using the *vegan* R package and then determined the dissimilarity among samples by dividing the Euclidean distance matrix (*Euc*) by the maximum Euclidean distance (Euc_{max}) of a given depth zone as described in (*30*) and summarized here: $EnvHt=(Euc/Euc_{max})+0.001$. Finally, the mean *EnvHt* was calculated as an estimation of environmental heterogeneity in each depth zone. Significant differences in environmental heterogeneity between depth zones were tested with a Kruskal-Wallis test, followed by a Wicoxon post-hoc test.

The presence of different water masses is an important feature to properly describe the deep dark ocean (> 200 m depth). Water masses are well-established features with unique properties characterized by their thermohaline and chemical characteristics. A water mass is composed of different proportions of one or more water types of a given origin (*115*). Here, the percentage of different water types contributing to the water mass composition of each sample (from 200 m to the bottom) was calculated using an optimum multiparameter water mass analysis (*95*). This method characterizes water types using conservative variables such as salinity and potential temperature (see (*55*) for details). We have identified 22 and 19 water types in the open ocean and the Mediterranean Sea, respectively. We computed the dissimilarity (Euclidean distance) between pairwise samples based on their water mass composition (% of each water type) to use in our downstream analysis. A nonmetric multidimensional scaling (NMDS) analysis based on these Euclidean distances was conducted to determine the differences among samples.

Least-cost geographical distances were calculated using the 'lc.dist()' function of the *marmap* R package (*96*). We first computed three transition matrices (using the 'trans.mat()' function) with different minimum depths, corresponding to the epi- (surface), meso- (200 m), and bathypelagic (1,000 m). Each generated transition matrix contained the probability of transition from one cell to adjacent cells of a given bathymetric grid. We used the high-resolution (15 arc-second) GEBCO bathymetric database hosted on the British Oceanographic Data Centre server (https://www.gebco.net/). Since the Mediterranean Sea's deep waters (>400 m) are completely separated by the Strait of Sicily, the *marmap* algorithm could not calculate the horizontal distance between bathypelagic samples situated in the western and eastern Mediterranean. To deal with this issue, we simulated the vertical trajectory needed to overcome the Strait of Sicily by simply summing each sample's depth to the geographical distances between 'isolated' stations. To calculate the least-cost distances, 'marmap' sets a depth limit for geographic barriers to compute the transition matrices (*96*). For example, if the limit is set to 0, the program calculates the distance by turning around the continents. However, in the case of the Mediterranean Sea, the western and eastern basins are completely isolated (at least horizontally) in depths below 400m, so the program generates unrealistically long distances between western and eastern samples from the deep ocean. To deal with this issue, for these isolated samples, we computed the least-cost distances by calculating the geographic distances (geodesic) between samples (not considering geographic barriers) and then summed the vertical distances to theoretically overcome the Strait of Sicily. For example, a western 1,400 m depth sample (1 km deeper than the top of the Strait of Sicily) located 200 km from an eastern 1,400 m depth sample had a final least-cost distance of 200 km + 2x 1 km = 202 km.

General analysis

Distance-based redundancy analyses (dbRDA) were performed on community differentiation (based on Bray-Curtis dissimilarities) of both prokaryotic (16S rRNA gene) and picoeukaryotic (18S rRNA gene) samples using the 'capscale()' function of the *vegan* R package (*103*). Analyses of dissimilarities were conducted using the 'adonis2()' function of the *vegan* package to investigate the percentage of variance in community

composition explained by environmental or geographic variables (*104*). Classic biogeographic provinces classifications (e.g.: Longhurst provinces; (*116*)) are only applied to the upper sunlit ocean (above 200 m), while deep-oceanic basins classifications (based on isolated water masses) are only applied to the deep (bellow 3,500 m) (*36*). Therefore, we used the classic oceanic basins (South Atlantic Ocean, North Atlantic Ocean, North Pacific Ocean, South Pacific Ocean, and Indian Ocean) as standard categorical explanatory variables to test the effect of geography between depth zones of the open ocean. For the Mediterranean Sea, we used the sub-basin classification (Levantine Sea, Ionan Sea, Sicily Strait, Tirrenyan Sea, Sardinian Sea, Alborean Sea, and Gibraltar Strait), based on Mediterranean internal circulation patterns (*117*) as well as physico-chemical, and biological features (*118*).

Spearman correlations were computed between β-diversity (Bray-Curtis and βNTI) and environmental Euclidean distances matrices using the 'cor.test()' function of the *stats* R package. Spearman correlations were also carried out to test the association between community (Bray-Curtis dissimilarity) and water masses composition (Euclidean distances) in the meso- and bathypelagic. Mantel correlograms were carried out with the 'mantel.correlog()' function in *Vegan* to test for the decrease in picoplankton community similarity (β-diversity) with increasing geographic distances (distance-decay). We used distance classes of 1,000 km for the open ocean, while for the Mediterranean Sea, we used distance classes of 350 km. Sequential differences in picoplankton β-diversity (Bray-Curtis dissimilarity) were computed in the sampling sequence of each cruise (see arrow directions in Fig. S13). Statistical differences between zones in sequential Bray-Curtis values were tested using analysis of variance (ANOVA) followed by a Tukey post-hoc test.

Pearson correlation matrices between diversity metrics and environmental variables were computed using the 'cor()' function and plotted with the *ggcorrplot* R package. Nonmetric multidimensional scaling (NMDS) based on Euclidean distances was used to analyze similarities and differences in water mass composition among ocean depth zones and basins, followed by an analysis of similarities (ANOSIM) to test for differences among delineated groups. The NMDS and ANOSIM were implemented using the 'metaMDS()' and 'anosim()' *vegan* functions, respectively. Analysis of variance (ANOVA), followed by a Tukey post-hoc test, was used to test for statistical differences in β-diversity metrics (Bray-Curtis, βNTI, and RC_{bray}). Differences in environmental heterogeneity between zones were tested using Kruskal-Wallis, followed by a Wicoxon post-hoc test. Linear regression models were carried out to investigate the influence of water masses (Euclidean distance) on community composition (Bray-Curtis dissimilarity) in each vertical profile. Spearman correlations were used to test the correlation between the ecological processes results obtained with the entire (unbalanced) dataset and the results found with a standardized sampling size dataset. All statistical analyses were conducted in the R statistical environment (*105*) and all plots were generated using the R package *ggplot2* (*119*).

Results

In general, we observed an inverted diversity pattern between the two main components of the picoplankton community: while prokaryotic diversity (richness, Shannon index, and phylogenetic diversity) increased with depth, picoeukaryotic diversity decreased towards the deep ocean (Fig. S2). The Pielou evenness index increased for prokaryotes and decreased for picoeukaryotes from the surface to the deep ocean (Fig. S2). The gamma diversity of prokaryotes was 26776 ASVs in the open ocean and 11795 ASVs in the Mediterranean Sea. The gamma diversity of picoeukaryotes was 35165 ASVs in the open ocean and 12367 ASVs in the Mediterranean Sea. The prokaryotic gamma diversity (adjusted to a standard sample set; n=38) increased from the epi- (4203 in the open ocean and 2723 in the Mediterranean Sea) to the meso- (8898 in the open ocean and 4303 in the Mediterranean Sea) and bathypelagic (8898 in the open ocean and 4303 in the Mediterranean Sea). The picoeukaryotic adjusted gamma diversity (n=38) decreased from the epi- (12290 in the open ocean and 5699 in the Mediterranean Sea) to the meso- (6496 in the open ocean and 4217 in the Mediterranean Sea) and bathypelagic (4668 in the open ocean and 2330 in the Mediterranean Sea).

Prokaryotes and picoeukaryotes displayed significant differences ($p<0.05$) between depth zones for bNTI and RC_{Bray} metrics (Fig S3). We also found significant differences ($p<0.05$) between ocean zones with regard to additional β-diversity metrics (i.e. Bray-Curtis, Jaccard, and Sorensen) and their partitioning (Fig. S4). In the surface open ocean (~3m), the role of selection was higher for prokaryotic (~27%) than for picoeukaryotic community turnover (~11%) (Fig. S7). *Heterogeneous selection* had a higher importance in structuring picoeukaryotes as compared to prokaryotes (~7% vs. ~4%, respectively). Conversely, *homogeneous selection* was more important for prokaryotes (~23%) than for picoeukaryotes (~4%) (Fig. S7). *Dispersal limitation* explained ~67% of picoeukaryotic and ~25% of prokaryotic community turnover in the surface open ocean (~3m). *Drift* was a relevant process structuring prokaryotic communities (~31%) but not picoeukaryotic counterparts (~6%) (Fig. S7). Similarly, *selection* explained ~26% of prokaryotic and ~10% of picoeukaryotic turnover in the surface Mediterranean Sea (Fig. S7). Unlike the open ocean, *heterogeneous selection* explained a higher percentage of prokaryotic pairwise comparisons (~17%) than *homogeneous selection* (~9%). An equal proportion of *heterogeneous* (~5%) and *homogeneous selection* (~5%) explained the structure of picoeukaryotes in the surface Mediterranean (Fig. S7). *Dispersal limitation* was the most important process (~63%) shaping picoeukaryotes and explained ~23% of the turnover of prokaryotes (Fig. S7), whereas *drift* explained most of the turnover of prokaryotes (~46%) and ~25% of picoeukaryotic assembly in the Mediterranean Sea's surface waters (Fig. S7).

In the open ocean DCM, *selection* explained ~21% and ~46% of the turnover of prokaryotic and picoeukaryotic communities, respectively (Fig. S7). While *heterogeneous selection* was relatively more important for picoeukaryotes (~45%) than prokaryotes (~10%), *homogeneous selection* was much smaller for picoeukaryotes (0.8%) than prokaryotes (11%) in the DCM of the open ocean (Fig. S7). *Dispersal limitation* was relatively more important for prokaryotes (52%) than picoeukaryotes

(37%) in this zone. *Homogenizing dispersal* was very low for both prokaryotes (~4%) and picoeukaryotes (~0.6%) in the DCM. *Drift* explained 26% and 13% of the turnover of prokaryotes and picoeukaryotes, respectively, in the DCM of the open ocean. In the Mediterranean Sea, *selection* was also relatively more important for picoeukaryotes (~23%) than for prokaryotes (~16%) (Fig. S7). *Heterogeneous selection* explained a larger proportion of the turnover of picoeukaryotes (~22%) than of prokaryotes (~9%), whereas *homogeneous selection* was more important in structuring prokaryotes (~7%) than picoeukaryotes (~0.4%) in the DCM of the Mediterranean Sea (Fig. S7).

Since there is also vertical dispersal between ocean surface and deep waters, we have also estimated the ecological processes integrating all depths (from 3 to 4,000 m) in each of the 13 vertical profile stations (VP stations in Fig. 1A). We found that *selection* was the most important process vertically shaping free-living picoplankton communities in the *Malaspina* vertical profiles, explaining ~52-81% of the prokaryotic community turnover (Fig. 2C) and ~24-52% of the picoeukaryotic community turnover (Fig. 2C). The role of vertical *heterogeneous selection* ranged from 29% to 52% in prokaryotes and from 10% to 52% in picoeukaryotes (Fig. 2C). *Homogeneous selection* was much higher for prokaryotes than for picoeukaryotes in most of the *Malaspina vertical profiles* (Fig. 2C). The role of vertical *dispersal limitation* ranged from 10% to 43% in prokaryotes and from 5% to 43% in picoeukaryotes (Fig. 2C). Vertical *homogenizing dispersal* was negligible (~0-5%) for both domains (Fig. 2C). The role of *drift* was relatively more important in picoeukaryotes (~15-43%) than in prokaryotes (~5-24%) across vertical profiles (Fig. 2C).

Environmental heterogeneity was significantly higher in the epipelagic than in the meso- and bathypelagic of the open ocean and the Mediterranean Sea (Fig. S8). The bathypelagic displayed slightly higher environmental heterogeneity than the mesopelagic in the open ocean (Fig. S8). On the contrary, the Mediterranean bathypelagic had the lowest level of environmental heterogeneity. (Fig. S8). Water mass composition was vertically structured in the open ocean (stress = 0.19521) with two separated clusters, meso- and bathypelagic, along the NDMS1 axis (Fig. S11). There was also a secondary separation between the Atlantic Ocean and the Pacific+Indian oceans along the NMDS2 axis (Fig. S11). The ANOSIM analysis confirmed that this separation in water mass composition (Euclidean distance) between zones ($r=0.48$, $p<0.05$) was stronger than that between ocean basins ($r=0.39$, $p<0.05$). In the Mediterranean Sea, there was a strong (stress = 0.1828) horizontal segregation in water masses between the Western and Eastern basins along the NMDS1 axis (Fig. S11). This segregation between basins was much stronger ($r=0.52$, $p<0.05$) than between zones ($r=0.18$, $p<0.05$) in the Mediterranean Sea.

Supplementary Figures

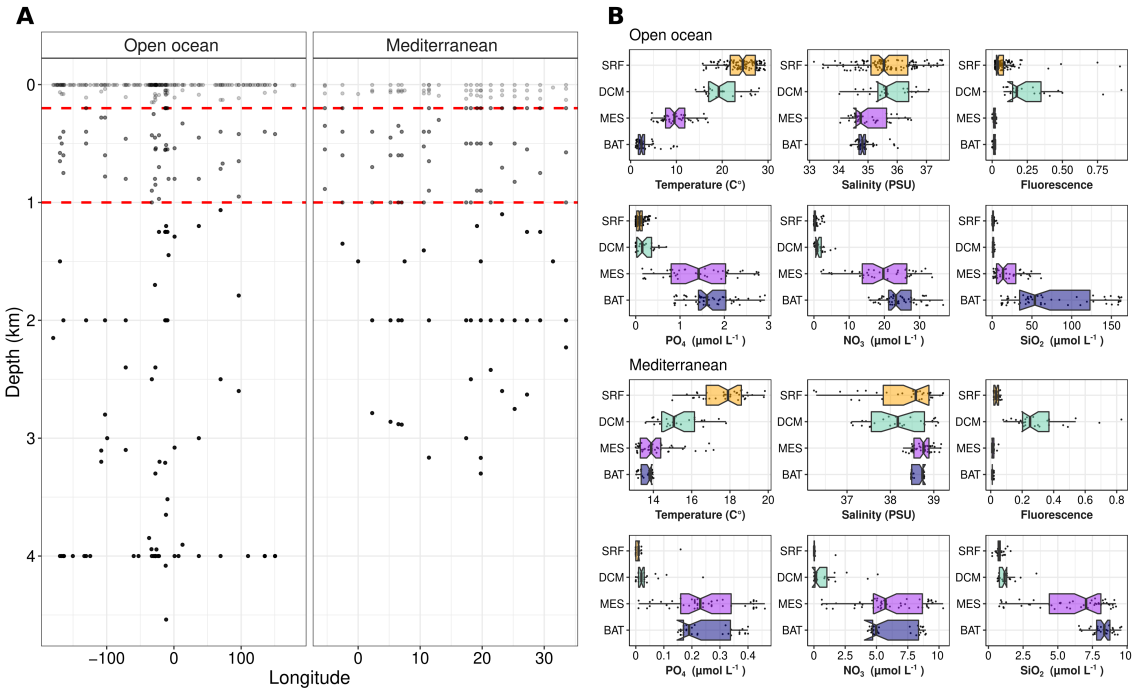

Fig. S1. The analyzed dataset covers environmentally contrasting depths zones of the ocean (A) Distribution of the samples across depths and longitude. The dashed red lines depict the division between zones: epi- (0-200 m), meso- (200-1,000 m), and bathypelagic (>1,000 m) **(B)** Boxplots showing the variability, by depth zones, of the environmental variables used in this study. Note the difference in scales between the open ocean and the Mediterranean Sea. Means were significantly different (ANOVA, Tukey post-hoc test; $p<0.001$) between upper (Surface [SRF] and DCM) and deep (Mesopelagic [MES] and Bathypelagic [BAT]) zones.

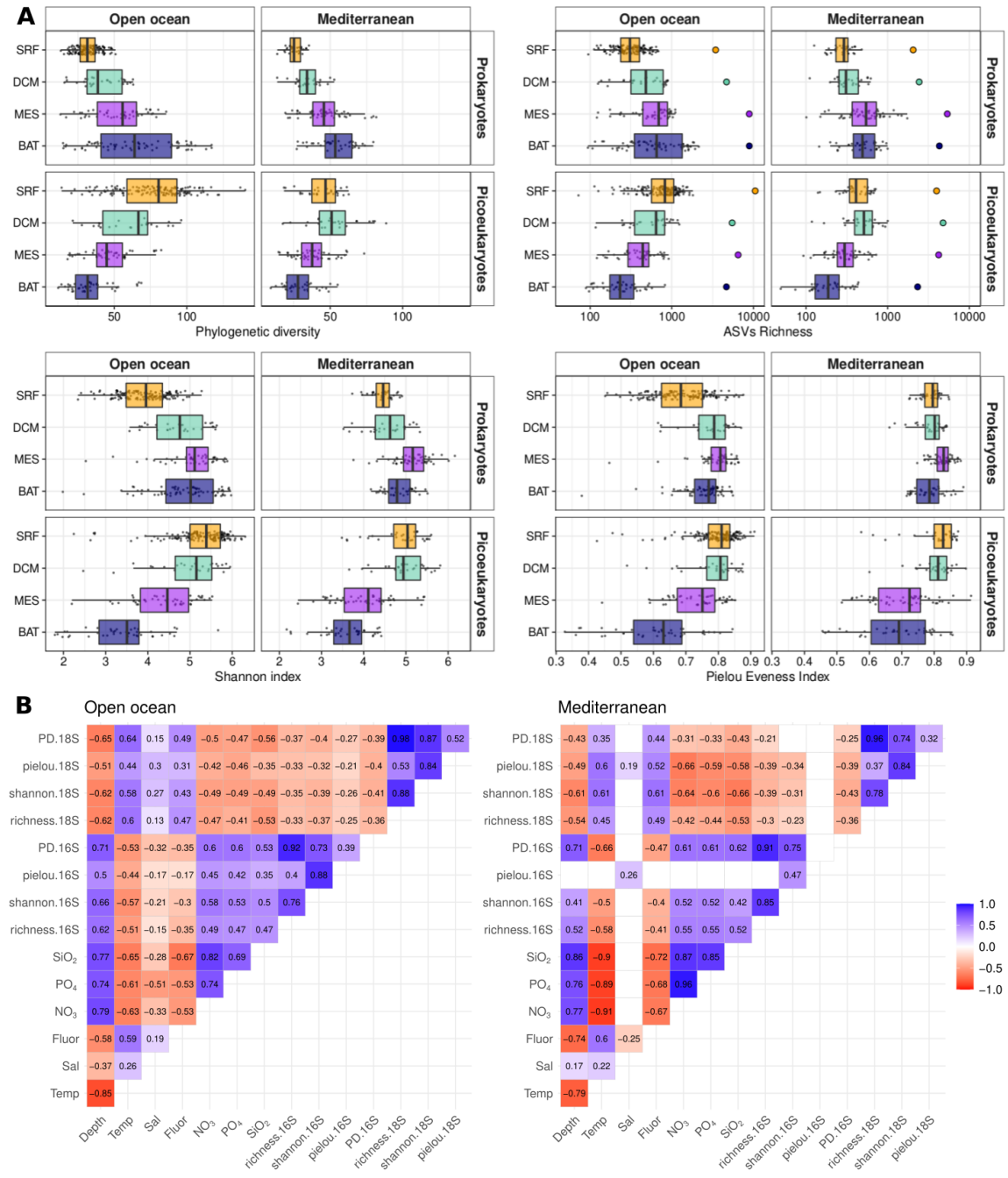

Fig. S2. Diversity indexes and their correlations with environmental variables (A) Picoplankton diversity depicted as phylogenetic diversity, ASVs richness, Shannon and Pielou's evenness index by depth zones (SRF, surface; DCM, deep chlorophyll maxima; MES, Mesopelagic; BAT, Bathypelagic). The circles in the ASVs richness boxplots stand for gamma diversity adjusted by sampling size. **(B)** Correlation matrix of *Pearson* (R) correlation values between picoplankton diversity metrics and environmental variables in the open ocean and the Mediterranean Sea. The empty boxes represent non-significant correlations (p>0.05). The '16S' tags in the metrics depict prokaryotic communities, while '18S' tags represent picoeukaryotic

communities. PD = phylogenetic diversity; Temp = Temperature; Sal = Salinity; Fluor = Fluorescence.

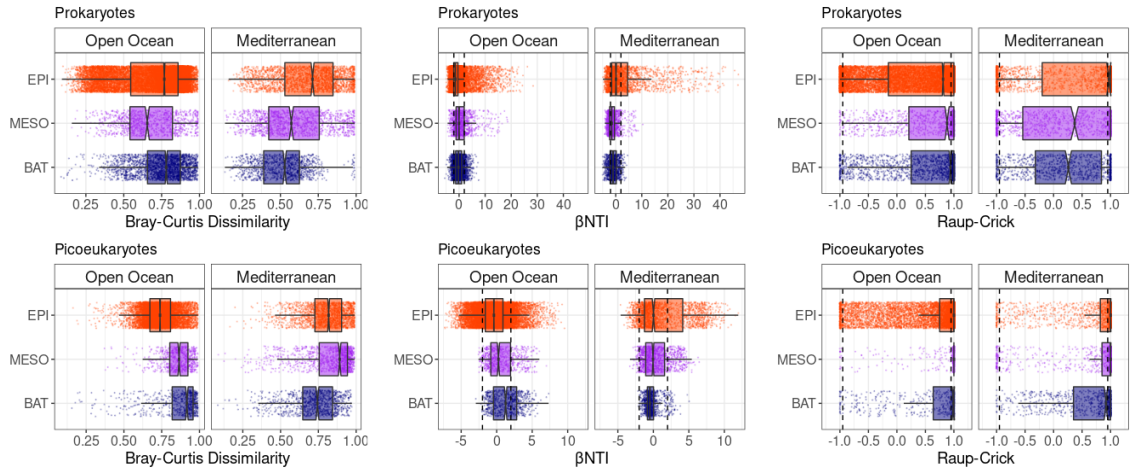

Fig. S3. Bray-Curtis, βNTI, and RC_{bray} metrics by depth zones for prokaryotes and picoeukaryotes. Means were significantly different (ANOVA, Tukey post-hoc test; $p<0.001$) between depth zones for both prokaryotes and picoeukaryotes. See Fig. S4. for β-diversity partitioning plots. EPI = epipelagic, MESO = mesopelagic, BAT = bathypelagic.

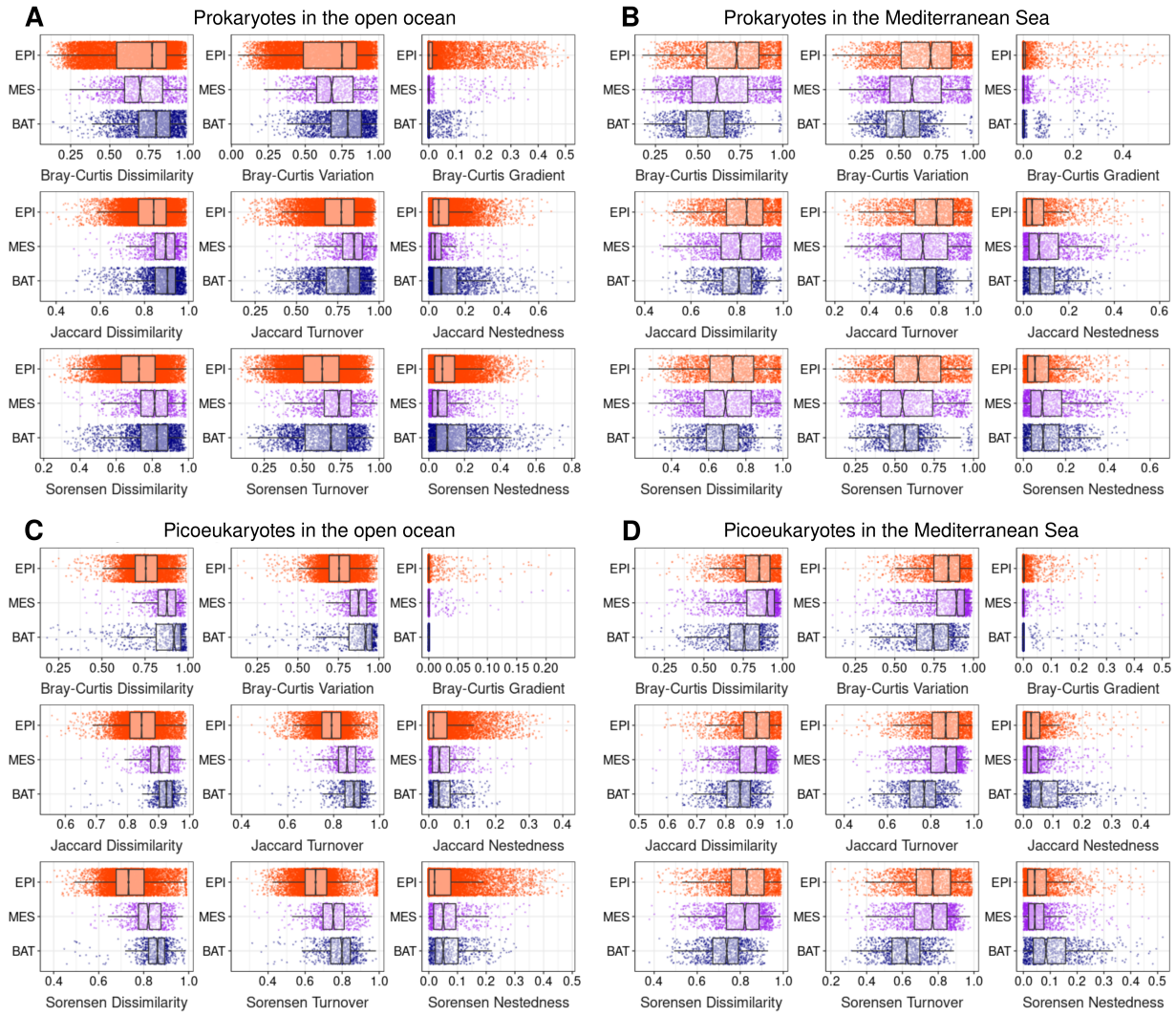

Fig. S4. Picoplankton β-diversity partitioning in the different ocean depth zones. Bray-Curtis dissimilarity (variation and gradient), Jaccard dissimilarity (turnover and nestedness), and Sorensen (turnover and nestedness) for prokaryotes and picoeukaryotes in the open ocean **(A, C)** and Mediterranean Sea **(B, D)**. Means were significantly different (ANOVA, Tukey post-hoc test; $p<0.001$) between depth zones for both prokaryotes and picoeukaryotes. EPI = epipelagic, MESO = mesopelagic, BAT = bathypelagic.

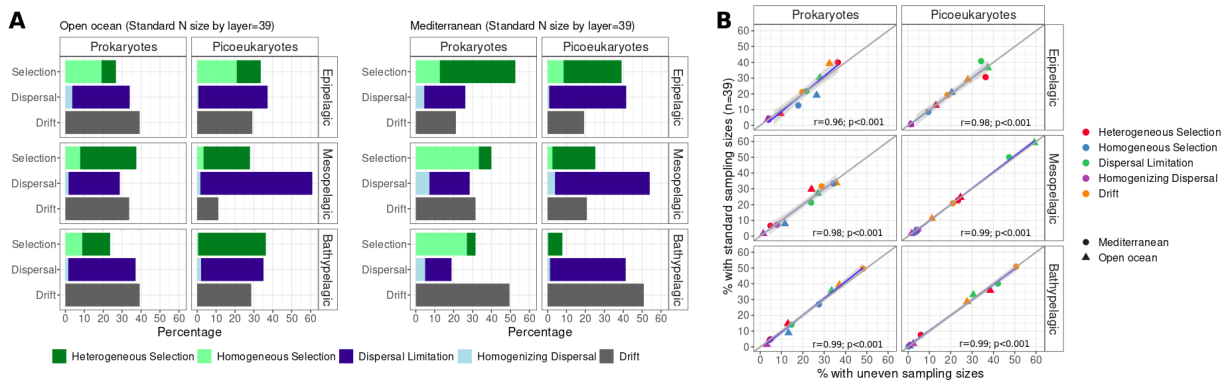

Fig. S5. Community assembly processes in picoplankton across ocean depth zones using standardized sampling sizes (n=39). (A) Relative importance of the ecological processes structuring the picoplankton communities in different depth zones of the open ocean and the Mediterranean Sea: Epi- (n=39), Meso- (n=39), and Bathypelagic (n=39). **(B)** Linear regression between the ecological processes' results obtained with the entire (unbalanced) and the standardized (n=39) datasets. The latter samples were evenly distributed across space, as shown in Fig S6.

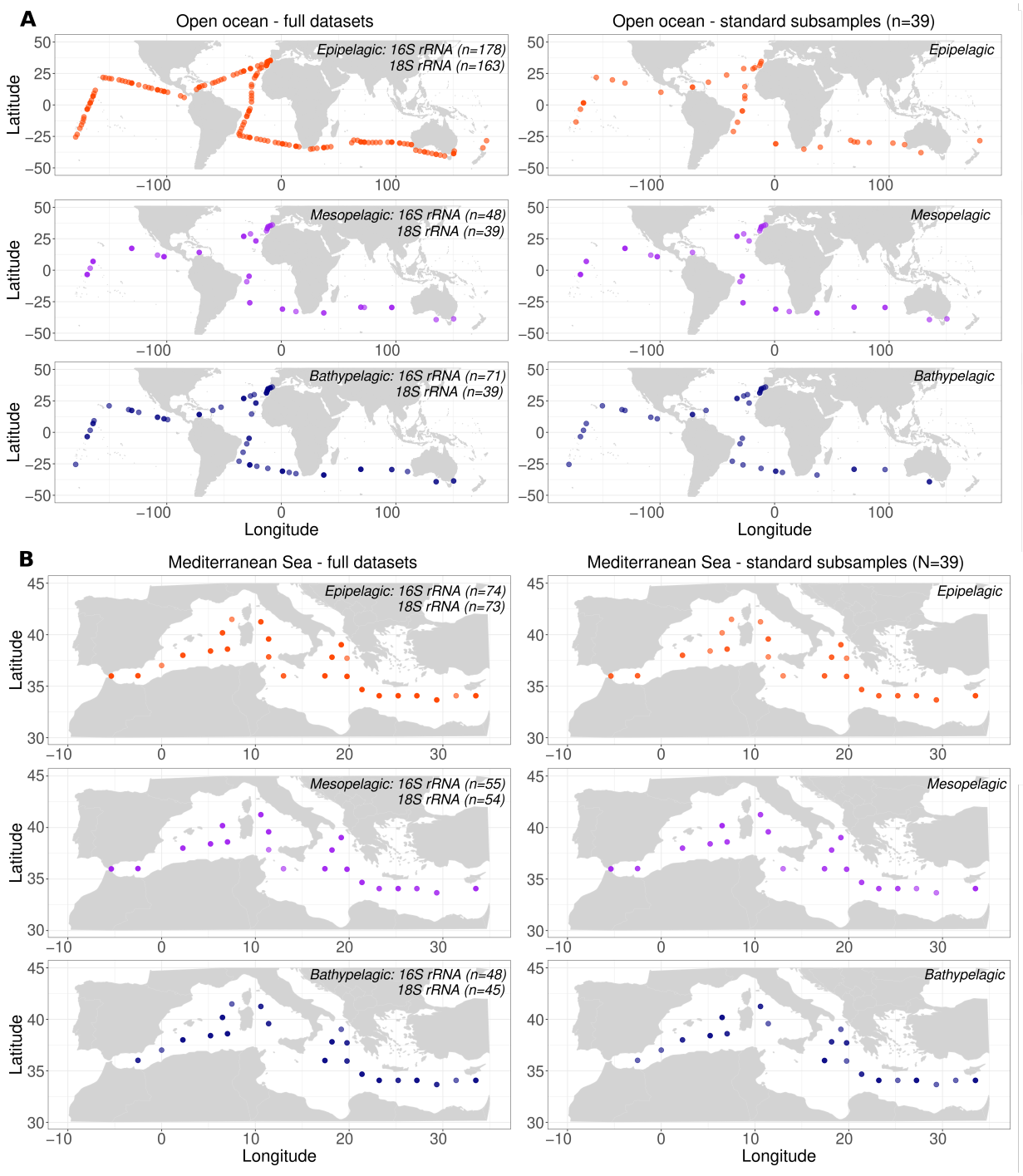

Fig. S6. Original and standardized samples distributions in each depth zone. Geographic distribution of sampling stations in each depth zone for the entire dataset and for those with standardized sampling size (n=39) in the **(A)** open ocean and **(B)** Mediterranean Sea. Samples were evenly distributed across space.

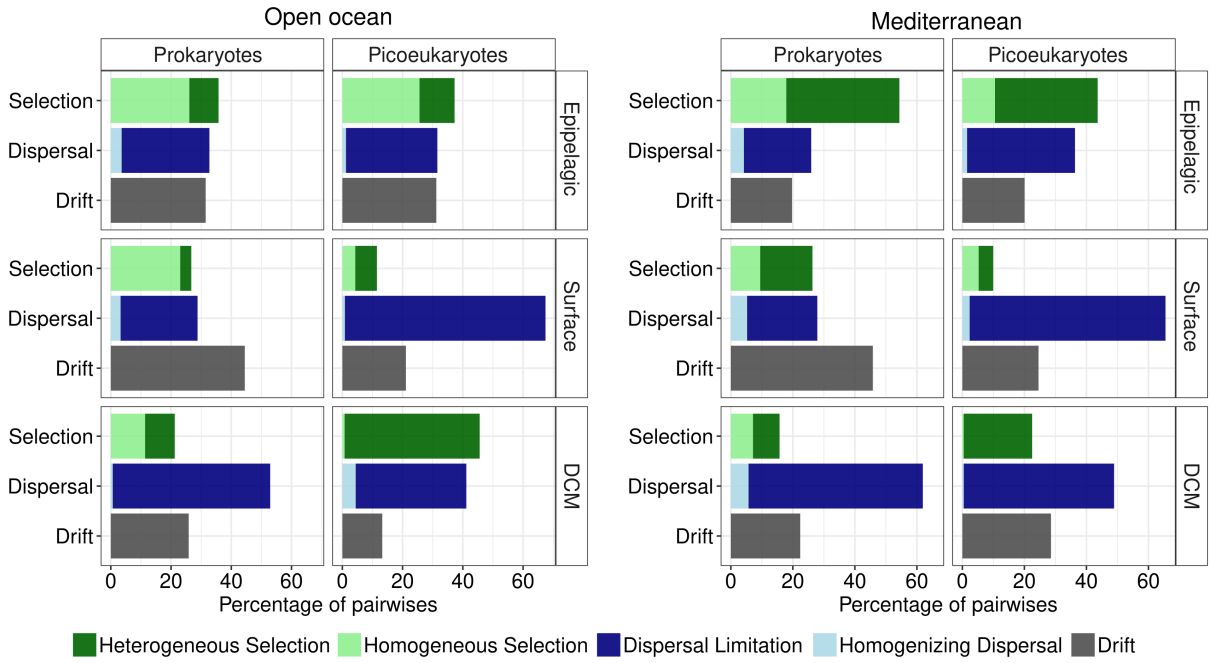

Fig. S7. Picoplankton community assembly processes in different depth zones of the epipelagic. Relative importance of the ecological processes structuring the picoplanktonic community in the epipelagic and its subdivisions, surface, and DCM zones of the open ocean and the Mediterranean Sea.

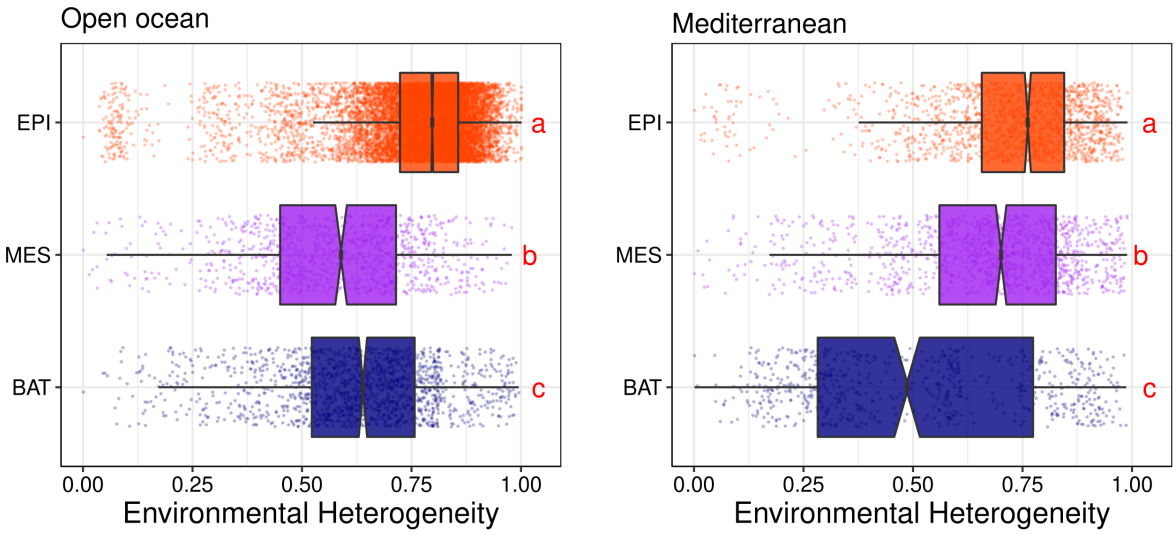

Fig. S8. Differences in environmental heterogeneity between ocean depth zones. Environmental heterogeneity computed as the mean environmental dissimilarity between samples considering the main variables that we measured (Temperature, Salinity, Fluorescence, NO_3, PO_4, SiO_2) in the open ocean and the Mediterranean Sea. Different red letters represent significantly different means [Kruskal-Wallis, Wilcoxon post-hoc test, $p<0.05$] between depth zones.

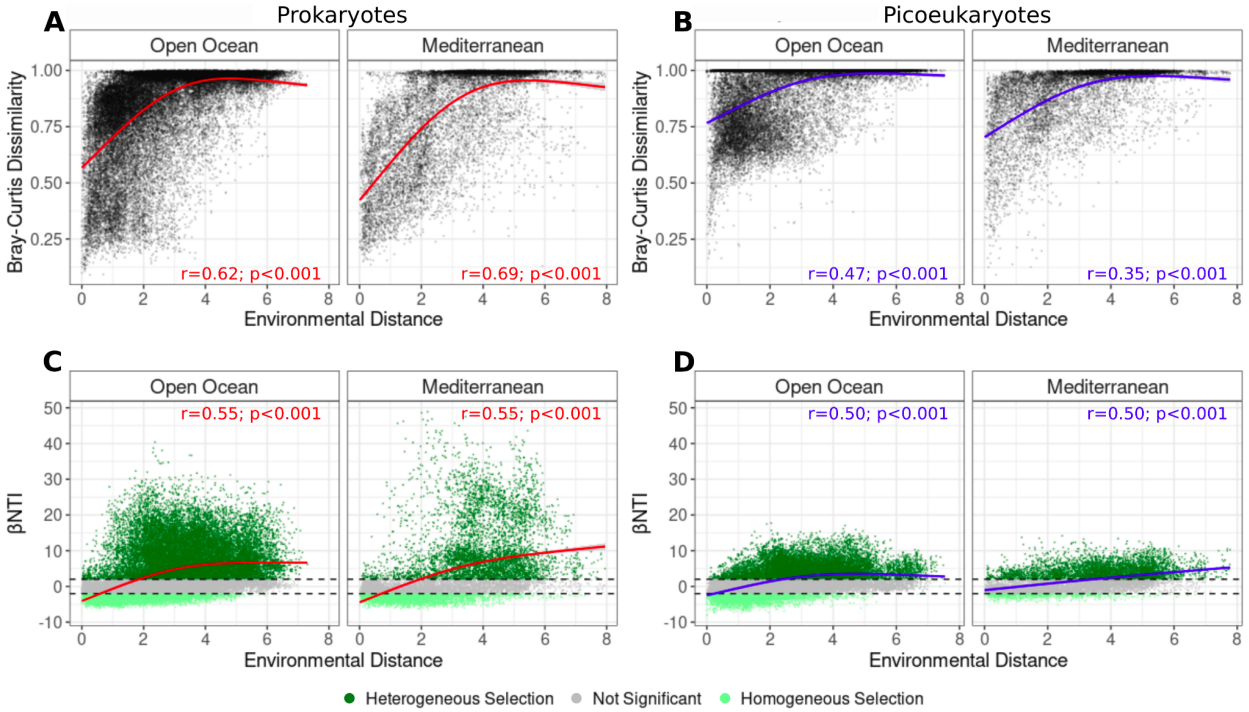

Fig. S9. Picoplankton community composition and phylogeny were positively related to environmental heterogeneity. Difference in taxonomic (Bray-Curtis dissimilarity) and phylogenetic (βNTI) composition for all pairwise picoplankton community comparisons as a function of environmental distance for both prokaryotes **(A, C)** and picoeukaryotes **(B, D)** in the open ocean and the Mediterranean Sea. The solid curves illustrate the nonlinear regressions. Spearman's rank correlation coefficients are depicted on the panel. Outliers with high environmental distances (>10) corresponding to pairwise comparisons with epipelagic samples from the Costa Rica Dome upwelling system were removed from the open ocean plot (see also Fig S13).

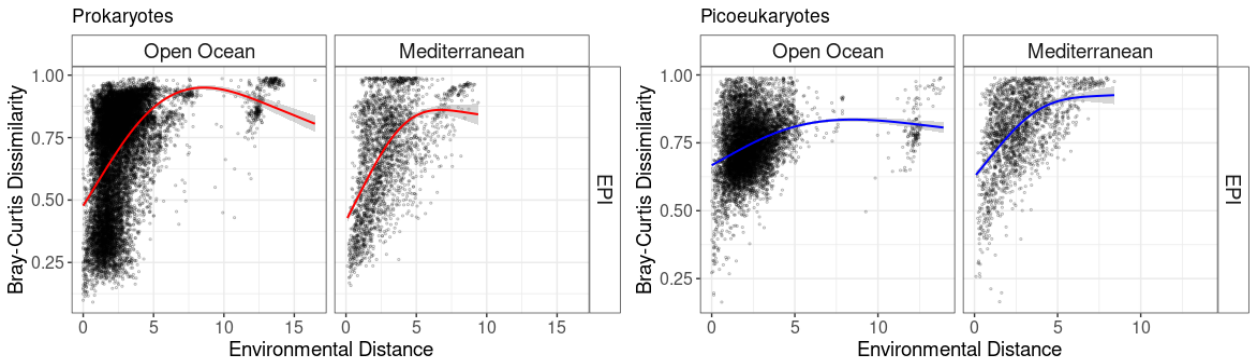

Fig. S10. Picoplankton community composition is positively correlated with environmental heterogeneity in the epipelagic. Difference in composition (Bray-Curtis dissimilarity) for all pairwise picoplankton community comparisons as a function of environmental distance for both prokaryotes and picoeukaryotes in the epipelagic of the open ocean and the Mediterranean Sea. The points with high environmental distances (>10) correspond to the pairwise comparisons with epipelagic samples from the Costa Rica Dome. The solid curves illustrate the nonlinear regressions. Spearman's rank correlation coefficients are depicted on the panel. EPI = epipelagic

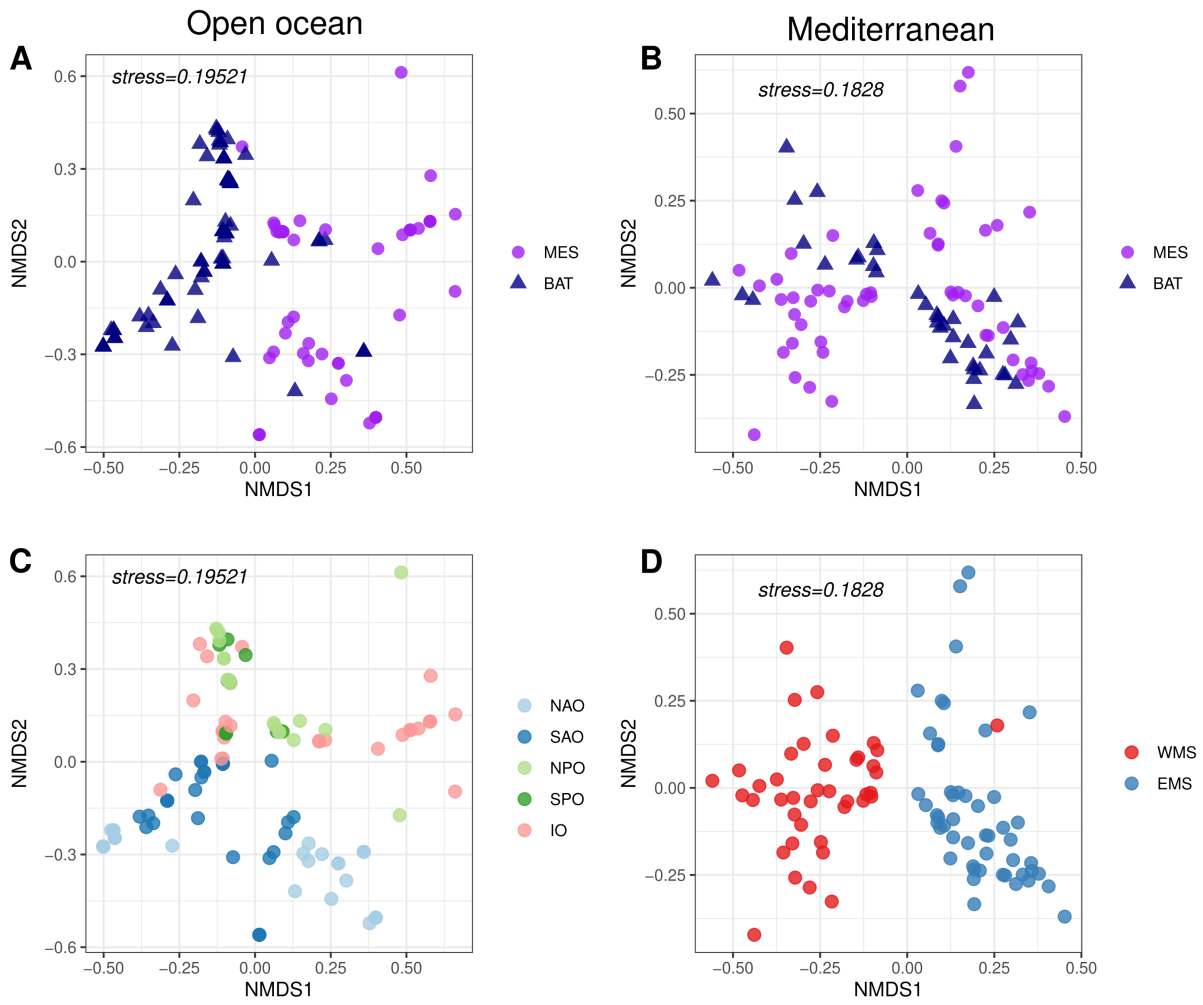

Fig. S11. Differences in water mass composition are segregated by depth zones and ocean basins. Nonmetric multidimensional scaling (NMDS) based on the Euclidean distances of the samples' water mass composition – labeled by zones and basin – in the open ocean **(A, C)** and the Mediterranean Sea **(B, D)**. MES = Mesopelagic; BAT = Bathypelagic. NAO = North Atlantic Ocean, SAO = South Atlantic Ocean, NPO = North Pacific Ocean, SPO = South Pacific Ocean, IO = Indian Ocean, WMS = Western Mediterranean Sea, EMS = Eastern Mediterranean Sea.

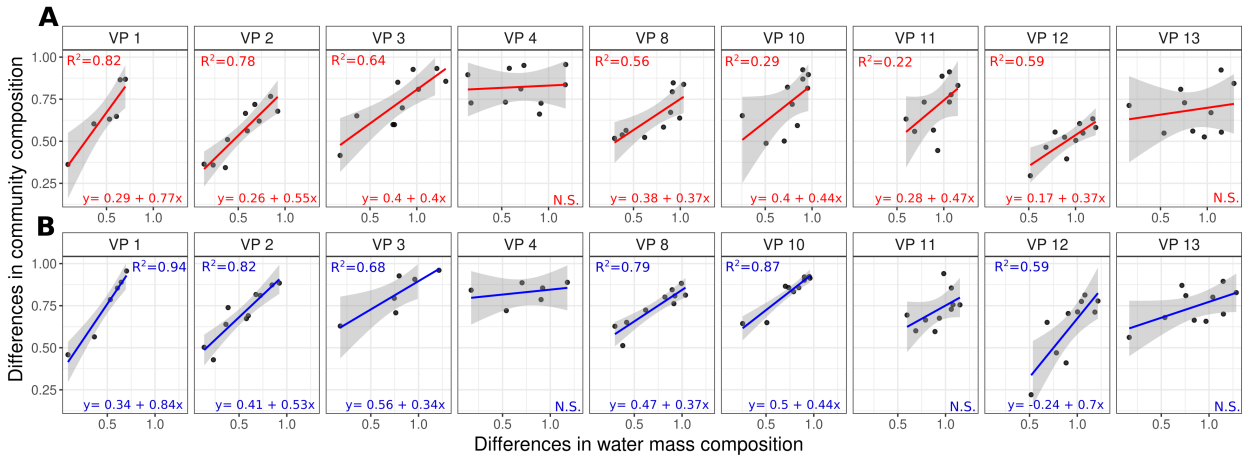

Fig. S12. Picoplankton community composition and potential dispersal are vertically correlated to differences in water mass composition. Difference in community composition (Bray-Curtis dissimilarity) as a function of water mass composition dissimilarity (Euclidean distances) for prokaryotes (in red) **(A)** and picoeukaryotes (in blue) **(B)** in *Malaspina* vertical profiles. Note that only meso- and bathypelagic samples were used in this analysis. The equation, the explanatory power of the linear regression models (adjusted R^2), and the significance of the smooth terms ($p<0.001$) are shown in the plots.

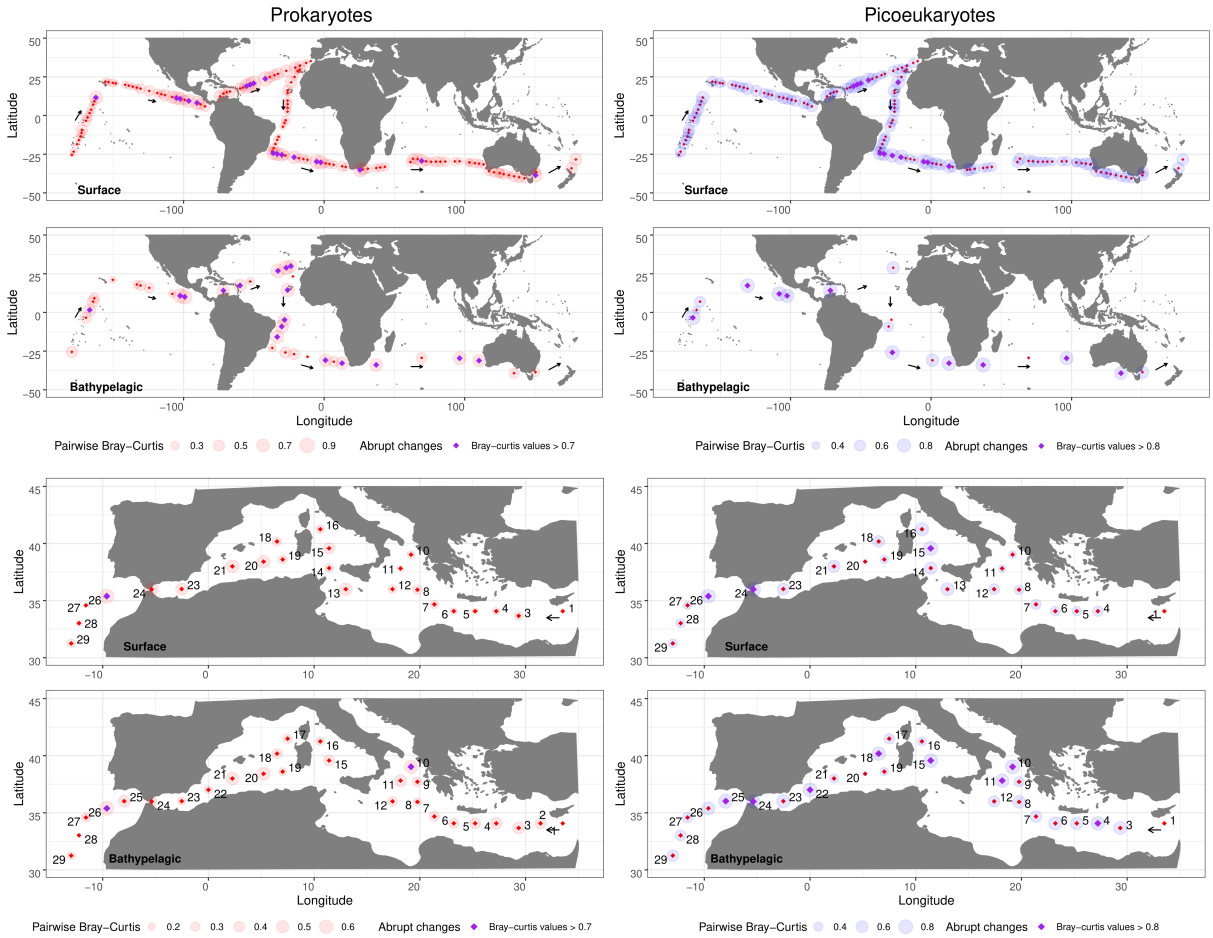

Fig. S13. Sequential change in community composition across space (sequential β-diversity). Communities were sampled along the *Malaspina* and *Hotmix* expeditions (black arrows), and the composition of each community was compared against its immediate predecessor. Note that only surface and bathypelagic samples are considered. The size of each bubble represents the Bray-Curtis dissimilarity between a given community and the community sampled previously.

REFERENCES AND NOTES

1. E. Sherr, B. Sherr, Understanding roles of microbes in marine pelagic food webs: A brief history, in *Microbial Ecology of the Oceans* (John Wiley & Sons Ltd., 2008), pp. 27–44.

2. L. Guidi, S. Chaffron, L. Bittner, D. Eveillard, A. Larhlimi, S. Roux, Y. Darzi, S. Audic, L. Berline, J. R. Brum, L. P. Coelho, J. C. I. Espinoza, S. Malviya, S. Sunagawa, C. Dimier, S. Kandels-Lewis, M. Picheral, J. Poulain, S. Searson, L. Stemmann, F. Not, P. Hingamp, S. Speich, M. Follows, L. Karp-Boss, E. Boss, H. Ogata, S. Pesant, J. Weissenbach, P. Wincker, S. G. Acinas, P. Bork, C. de Vargas, D. Iudicone, M. B. Sullivan, J. Raes, E. Karsenti, C. Bowler, G. Gorsky, Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532**, 465–470 (2016).

3. P. G. Falkowski, T. Fenchel, E. F. Delong, The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**, 1034–1039 (2008).

4. Y. M. Bar-On, R. Milo, The biomass composition of the oceans: A blueprint of our blue planet. *Cell* **179**, 1451–1454 (2019).

5. S. Sunagawa, L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. d'Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis; Tara Oceans coordinators, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, P. Bork, E. Boss, C. Bowler, M. Follows, L. Karp-Boss, U. Krzic, E. G. Reynaud, C. Sardet, M. Sieracki, D. Velayoudon, Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).

6. C. de Vargas, S. Audic, N. Henry, J. Decelle, F. Mahé, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horák, O. Jaillon, G. Lima-

Mendez, J. Lukeš, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis; Tara Oceans Coordinators, S. G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M. E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, E. Karsenti, Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**, 1261605 (2015).

7. R. Massana, Eukaryotic picoplankton in surface oceans. *Annu. Rev. Microbiol.* **65**, 91–110 (2011).

8. J. Arístegui, J. M. Gasol, C. M. Duarte, G. J. Herndld, Microbial oceanography of the dark ocean's pelagic realm. *Limnol. Oceanogr.* **54**, 1501–1529 (2009).

9. M. V. Brown, G. K. Philip, J. A. Bunge, M. C. Smith, A. Bissett, F. M. Lauro, J. A. Fuhrman, S. P. Donachie, Microbial community structure in the North Pacific ocean. *ISME J.* **3**, 1374–1386 (2009).

10. M. Sebastián, E. Ortega-Retuerta, L. Gómez-Consarnau, M. Zamanillo, M. Álvarez, J. Arístegui, J. M. Gasol, Environmental gradients and physical barriers drive the basin-wide spatial structuring of Mediterranean Sea and adjacent eastern Atlantic Ocean prokaryotic communities. *Limnol. Oceanogr.* **66**, 4077–4095 (2021).

11. C. R. Giner, M. C. Pernice, V. Balagué, C. M. Duarte, J. M. Gasol, R. Logares, R. Massana, Marked changes in diversity and relative activity of picoeukaryotes with depth in the world ocean. *ISME J.* **14**, 437–449 (2020).

12. P. E. Galand, M. Potvin, E. O. Casamayor, C. Lovejoy, Hydrography shapes bacterial biogeography of the deep Arctic Ocean. *ISME J.* **4**, 564–576 (2010).

13. S. E. Morales, M. Meyer, K. Currie, F. Baltar, Are oceanic fronts ecotones? Seasonal changes along the subtropical front show fronts as bacterioplankton transition zones but not diversity hotspots. *Environ. Microbiol. Rep.* **10**, 184–189 (2018).

14. E. J. Raes, L. Bodrossy, J. van de Kamp, A. Bissett, M. Ostrowski, M. V. Brown, S. L. S. Sow, B. Sloyan, A. M. Waite, Oceanographic boundaries constrain microbial diversity gradients in the South Pacific Ocean. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8266–E8275 (2018).

15. R. Logares, I. M. Deutschmann, P. C. Junger, C. R. Giner, A. K. Krabberød, T. S. B. Schmidt, L. Rubinat-Ripoll, M. Mestre, G. Salazar, C. Ruiz-González, M. Sebastián, C. de Vargas, S. G. Acinas, C. M. Duarte, J. M. Gasol, R. Massana, Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome* **8**, 55 (2020).

16. E. Villarino, J. R. Watson, G. Chust, A. J. Woodill, B. Klempay, B. Jonsson, J. M. Gasol, R. Logares, R. Massana, C. R. Giner, G. Salazar, X. A. Alvarez-Salgado, T. S. Catala, C. M. Duarte, S. Agusti, F. Mauro, X. Irigoien, A. D. Barton, Global beta diversity patterns of microbial communities in the surface and deep ocean. *Glob. Ecol. Biogeogr.* **31**, 2323–2336 (2022).

17. S. G. Acinas, P. Sánchez, G. Salazar, F. M. Cornejo-Castillo, M. Sebastián, R. Logares, M. Royo-Llonch, L. Paoli, S. Sunagawa, P. Hingamp, H. Ogata, G. Lima-Mendez, S. Roux, J. M. González, J. M. Arrieta, I. S. Alam, A. Kamau, C. Bowler, J. Raes, S. Pesant, P. Bork, S. Agustí, T. Gojobori, D. Vaqué, M. B. Sullivan, C. Pedrós-Alió, R. Massana, C. M. Duarte, J. M. Gasol, Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities. *Commun. Biol.* **4**, 604 (2021).

18. M. Vellend, The theory of ecological communities. *Monogr. Popul. Biol.* **57**, 229 (2016).

19. J.-F. Ghiglione, P. E. Galand, T. Pommier, C. Pedrós-Alió, E. W. Maas, K. Bakker, S. Bertilson, D. L. Kirchman, C. Lovejoy, P. L. Yager, A. E. Murray, Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc. Natl. Acad. Sci.* **109**, 17633–17638 (2012).

20. H. Sarmento, C. Morana, J. M. Gasol, Bacterioplankton niche partitioning in the use of phytoplankton-derived dissolved organic carbon: Quantity is more important than quality. *ISME J.* **10**, 2582–2592 (2016).

21. A. Auladell, A. Barberán, R. Logares, E. Garcés, J. M. Gasol, I. Ferrera, Seasonal niche differentiation among closely related marine bacteria. *ISME J.* **16**, 178–189 (2022).

22. J. Zhou, D. Ning, Stochastic community assembly: Does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* **81**, e00002–e000017 (2017).

23. S. Louca, The rates of global bacterial and archaeal dispersal. *ISME J.* **16**, 159–167 (2022).

24. S. Fodelianakis, A. Valenzuela-Cuevas, A. Barozzi, D. Daffonchio, Direct quantification of ecological drift at the population level in synthetic bacterial communities. *ISME J.* **15**, 55–66 (2021).

25. J. Heino, A. S. Melo, T. Siqueira, J. Soininen, S. Valanko, L. M. Bini, Metacommunity organisation, spatial extent and dispersal in aquatic systems: Patterns, processes and prospects. *Freshw. Biol.***60**, 845–869 (2015).

26. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).

27. D. R. Nemergut, S. K. Schmidt, T. Fukami, S. P. O'Neill, T. M. Bilinski, L. F. Stanish, J. E. Knelman, J. L. Darcy, R. C. Lynch, P. Wickey, S. Ferrenberg, Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* **77**, 342–356 (2013).

28. J. C. Stegen, X. Lin, J. K. Fredrickson, X. Chen, D. W. Kennedy, C. J. Murray, M. L. Rockhold, A. Konopka, Quantifying community assembly processes and identifying features that impose them. *ISME J.* **7**, 2069–2079 (2013).

29. C. R. Woese, Bacterial evolution. *Microbiol. Rev.* **51**, 221–271 (1987).

30. P. Huber, S. Metz, F. Unrein, G. Mayora, H. Sarmento, M. Devercelli, Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. *ISME J.* **14**, 2951–2966 (2020).

31. C. Ruiz-González, R. Logares, M. Sebastián, M. Mestre, R. Rodríguez-Martínez, M. Galí, M. M. Sala, S. G. Acinas, C. M. Duarte, J. M. Gasol, Higher contribution of globally rare bacterial taxa reflects environmental transitions across the surface ocean. *Mol. Ecol.* **28**, 1930–1945 (2019).

32. J. L. Reid, On the mid-depth circulation of the world ocean. *Evol. Phys. Oceanogr.* **623**, 70–111 (1981).

33. E. Mayol, J. M. Arrieta, M. A. Jiménez, A. Martínez-Asensio, N. Garcias-Bonet, J. Dachs, B. González-Gaya, S.-J. Royer, V. M. Benítez-Barrios, E. Fraile-Nuez, C. M. Duarte, Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nat. Commun.* **8**, 201 (2017).

34. D. J. Richter, R. Watteaux, T. Vannier, J. Leconte, P. Frémont, G. Reygondeau, N. Maillet, N. Henry, G. Benoit, O. Da Silva, T. O. Delmont, A. Fernàndez-Guerra, S. Suweis, R. Narci, C. Berney, D. Eveillard, F. Gavory, L. Guidi, K. Labadie, E. Mahieu, J. Poulain, S. Romac, S. Roux, C. Dimier, S. Kandels, M. Picheral, S. Searson; Tara Oceans Coordinators, S. Pesant, J.-M. Aury, J. R. Brum, C. Lemaitre, E. Pelletier, P. Bork, S. Sunagawa, F. Lombard, L. Karp-Boss, C. Bowler, M. B. Sullivan, E. Karsenti, M. Mariadassou, I. Probert, P. Peterlongo, P. Wincker, C. de Vargas, M. Ribera d'Alcalà, D. Iudicone, O. Jaillon, Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems. *eLife* **11**, e78129 (2022).

35. E. Villarino, J. R. Watson, B. Jönsson, J. M. Gasol, G. Salazar, S. G. Acinas, M. Estrada, R. Massana, R. Logares, C. R. Giner, M. C. Pernice, M. P. Olivar, L. Citores, J. Corell, N. Rodríguez-Ezpeleta, J. L. Acuña, A. Molina-Ramírez, J. I. González-Gordillo, A. Cózar, E. Martí, J. A. Cuesta, S. Agustí, E. Fraile-Nuez, C. M. Duarte, X. Irigoien, G. Chust, Large-scale ocean connectivity and planktonic body size. *Nat. Commun.* **9**, 142 (2018).

36. G. Salazar, F. M. Cornejo-Castillo, V. Benítez-Barrios, E. Fraile-Nuez, X. A. Álvarez-Salgado, C. M. Duarte, J. M. Gasol, S. G. Acinas, Global diversity and biogeography of deep-sea pelagic prokaryotes. *ISME J.* **10**, 596–608 (2016).

37. J. P. Bethoux, B. Gentili, P. Morin, E. Nicolas, C. Pierre, D. Ruiz-Pino, The Mediterranean Sea: A miniature ocean for climatic and environmental studies and a key for the climatic functioning of the North Atlantic. *Prog. Oceanogr.* **44**, 131–146 (1999).

38. S. Sammartino, J. García Lafuente, C. Naranjo, J. C. Sánchez Garrido, R. Sánchez Leal, A. Sánchez Román, Ten years of marine current measurements in Espartel Sill, Strait of Gibraltar. *J. Geophys. Res.* **120**, 6309–6328 (2015).

39. M. D. Krom, N. Kress, S. Brenner, L. I. Gordon, Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. *Limnol. Oceanogr.* **36**, 424–432 (1991).

40. M. Estrada, M. Delgado, D. Blasco, M. Latasa, A. M. Cabello, V. Benítez-Barrios, E. Fraile-Nuez, P. Mozetič, M. Vidal, Phytoplankton across tropical and subtropical regions of the Atlantic, Indian and Pacific Oceans, *PLoS One* **11**, e0151699 (2016).

41. M. Cornec, H. Claustre, A. Mignot, L. Guidi, L. Lacour, A. Poteau, F. D'Ortenzio, B. Gentili, C. Schmechtig, Deep chlorophyll maxima in the global ocean: Occurrences, drivers and characteristics. *Global Biogeochem. Cycles* **35**, e2020GB006759 (2021)

42. E. Villar, G. K. Farrant, M. Follows, L. Garczarek, S. Speich, S. Audic, L. Bittner, B. Blanke, J. R. Brum, C. Brunet, R. Casotti, A. Chase, J. R. Dolan, F. d'Ortenzio, J.-P. Gattuso, N. Grima, L. Guidi, C. N. Hill, O. Jahn, J.-L. Jamet, H. Le Goff, C. Lepoivre, S. Malviya, E. Pelletier, J.-B. Romagnan, S. Roux, S. Santini, E. Scalco, S. M. Schwenck, A. Tanaka, P. Testor, T. Vannier, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis; Tara Oceans Coordinators, S. G. Acinas, P. Bork, E. Boss, C. De Vargas, G. Gorsky, H. Ogata, S. Pesant, M. B. Sullivan, S. Sunagawa, P. Wincker, E. Karsenti, C. Bowler, F. Not, P. Hingamp, D. Iudicone, Environmental characteristics of Agulhas rings affect interocean plankton transport. *Science* **348**, 1261447 (2015).

43. T. Soukissian, D. Denaxa, F. Karathanasi, A. Prospathopoulos, K. Sarantakos, A. Iona, K. Georgantas, S. Mavrakos, Marine renewable energy in the Mediterranean Sea: Status and perspectives. *Energies* **10**, 1512 (2017).

44. F. Milke, I. Wagner-Doebler, G. Wienhausen, M. Simon, Selection, drift and community interactions shape microbial biogeographic patterns in the Pacific Ocean. *ISME J.* **16**, 2653–2665 (2022).

45. J. Kong, L. Wang, C. Lin, F. Kuang, X. Zhou, E. A. Laws, P. Sun, H. Huang, B. Huang, Contrasting community assembly mechanisms underlie similar biogeographic patterns of surface microbiota in the tropical North Pacific Ocean. *Microbiol. Spectr.* **10**, e0079821 (2022).

46. H. K. Mod, M. Chevalier, M. Luoto, A. Guisan, Scale dependence of ecological assembly rules: Insights from empirical datasets and joint species distribution modelling. *J. Ecol.* **108**, 1967–1977 (2020).

47. F. Dini-Andreote, J. C. Stegen, J. D. van Elsas, J. F. Salles, Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci.* **112**, E1326–E1332 (2015).

48. R. Logares, S. V. M. Tesson, B. Canbäck, M. Pontarp, K. Hedlund, K. Rengefors, Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environ. Microbiol.* **20**, 2231–2240 (2018).

49. M. Vass, A. J. Székely, E. S. Lindström, S. Langenheder, Using null models to compare bacterial and microeukaryotic metacommunity assembly under shifting environmental conditions. *Sci. Rep.* **10**, 2455 (2020).

50. C. Yesson, M. R. Clark, M. L. Taylor, A. D. Rogers, The global distribution of seamounts based on 30 arc seconds bathymetry data. *Deep Res. Part I Oceanogr. Res. Pap.* **58**, 442–453 (2011).

51. P. Sun, Y. Wang, X. Huang, B. Huang, L. Wang, Water masses and their associated temperature and cross-domain biotic factors co-shape upwelling microbial communities. *Water Res.* **215**, 118274 (2022).

52. H. Agogué, D. Lamy, P. R. Neal, M. L. Sogin, G. J. Herndl, Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Mol. Ecol.* **20**, 258–274 (2011).

53. A. M. Martínez-Pérez, T. S. Catalá, M. Nieto-Cid, J. Otero, M. Álvarez, M. Emelianov, I. Reche, X. A. Álvarez-Salgado, J. Arístegui, Dissolved organic matter (DOM) in the open Mediterranean Sea. II: Basin-wide distribution and drivers of fluorescent DOM. *Prog. Oceanogr.* **170**, 93–106 (2019).

54. M. Gómez-Letona, J. Arístegui, N. Hernández-Hernández, X. A. Álvarez-Salgado, M. Álvarez, E. Delgadillo, M. Pérez-Lorenzo, E. Teira, S. Hernández-León, M. Sebastián, Deep ocean prokaryotes and fluorescent dissolved organic matter reflect the history of the water masses across the Atlantic Ocean. *Prog. Oceanogr.* **205**, 102819 (2022).

55. T. S. Catalá, I. Reche, M. Álvarez, S. Khatiwala, E. F. Guallart, V. M. Benítez-Barrios, A. Fuentes-Lema, C. Romera-Castillo, M. Nieto-Cid, C. Pelejero, E. Fraile-Nuez, E. Ortega-Retuerta, C. Marrasé, X. A. Álvarez-Salgado, Water mass age and aging driving chromophoric dissolved organic matter in the dark global ocean. *Global Biogeochem. Cycles* **29**, 917–934 (2015).

56. R. Massana, R. Logares, Eukaryotic versus prokaryotic marine picoplankton ecology. *Environ. Microbiol.* **15**, 1254–1261 (2013).

57. A. B. Zouari, M. B. Hassen, V. Balagué, E. Sahli, M. Y. B. Kacem, F. Akrout, A. Hamza, R. Massana, Picoeukaryotic diversity in the Gulf of Gabès: Variability patterns and relationships to nutrients and water masses. *Aquat. Microb. Ecol.* **81**, 37–53 (2018).

58. E. F. Neave, H. Seim, S. M. Gifford, O. Torano, Z. I. Johnson, D. Páez-Rosas, A. Marchetti, Protistan plankton communities in the Galápagos Archipelago respond to changes in deep water masses resulting from the 2015/16 El Niño. *Environ. Microbiol.* **24**, 1746–1759 (2022).

59. T. Severin, C. Sauret, M. Boutrif, T. Duhaut, F. Kessouri, L. Oriol, J. Caparros, M. Pujo-Pay, X. Durrieu de Madron, M. Garel, C. Tamburini, P. Conan, J.-F. Ghiglione, Impact of an intense water column mixing (0–1500 m) on prokaryotic diversity and activities during an open-ocean convection event in the NW Mediterranean Sea. *Environ. Microbiol.* **18**, 4378–4390 (2016).

60. O. Flegontova, P. Flegontov, N. Jachníková, J. Lukeš, A. Horák, Water masses shape pico-nano eukaryotic communities of the Weddell Sea. *Commun. Biol.* **6**, 64 (2023).

61. T. Cordier, I. B. Angeles, N. Henry, F. Lejzerowicz, C. Berney, R. Morard, A. Brandt, M.-A. Cambon-Bonavita, L. Guidi, F. Lombard, P. M. Arbizu, R. Massana, C. Orejas, J. Poulain, C. R. Smith, P. Wincker, S. Arnaud-Haond, A. J. Gooday, C. de Vargas, J. Pawlowski, Patterns of eukaryotic diversity from the surface to the deep-ocean sediment. *Sci. Adv.* **8**, eabj9309 (2022).

62. L. Zinger, A. Boetius, A. Ramette, Bacterial taxa–area and distance–decay relationships in marine environments. *Mol. Ecol.* **23**, 954–964 (2014).

63. C. A. Hanson, J. A. Fuhrman, M. C. Horner-Devine, J. B. H. Martiny, Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **10**, 497–506 (2012).

64. K. J. Locey, M. E. Muscarella, M. L. Larsen, S. R. Bray, S. E. Jones, J. T. Lennon, Dormancy dampens the microbial distance–decay relationship. *Philos. Trans. R Soc. B Biol. Sci.* **375**, 20190243 (2020).

65. K. J. Gaston, T. M. Blackburn, J. J. D. Greenwood, R. D. Gregory, R. M. Quinn, J. H. Lawton, Abundance–occupancy relationships *J. Appl. Ecology* **37**, 39–59 (2000).

66. C. J. Brislawn, E. B. Graham, K. Dana, P. Ihardt, S. J. Fansler, W. B. Chrisler, J. B. Cliff, J. C. Stegen, J. J. Moran, H. C. Bernstein, Forfeiting the priority effect: Turnover defines biofilm community succession. *ISME J.* **13**, 1865–1877 (2019).

67. T. Bie, L. Meester, L. Brendonck, K. Martens, B. Goddeeris, D. Ercken, H. Hampel, L. Denys, L. Vanhecke, K. Gucht, J. Wichelen, W. Vyverman, S. A. J. Declerck, Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecol. Lett.* **15**, 740–747 (2012).

68. C. Pedrós-Alió, Time travel in microorganisms. *Syst. Appl. Microbiol.* **44**, 126227 (2021).

69. M. Mestre, C. Ruiz-González, R. Logares, C. M. Duarte, J. M. Gasol, M. M. Sala, Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6799–E6807 (2018).

70. D. A. Gittins, P.-A. Desiage, N. Morrison, J. E. Rattray, S. Bhatnagar, A. Chakraborty, J. Zorz, C. Li, O. Horanszky, M. A. Cramm, F. Bisiach, R. Bennett, J. Webb, A. MacDonald, M. Fowler, D. C. Campbell, C. R. J. Hubert, Geological processes mediate a microbial dispersal loop in the deep biosphere. *Sci. Adv.* **8**, eabn3485 (2022).

71. N. Arandia-Gorostidi, A. E. Parada, A. E. Dekas, Single-cell view of deep-sea microbial activity and intracommunity heterogeneity. *ISME J.* **17**, 59–69 (2023).

72. G. J. Herndl, B. Bayer, F. Baltar, T. Reinthaler, Prokaryotic life in the deep ocean's water column. *Ann. Rev. Mar. Sci.* **15**, 461–483 (2023).

73. L. Kwiatkowski, O. Torres, L. Bopp, O. Aumont, M. Chamberlain, J. R. Christian, J. P. Dunne, M. Gehlen, T. Ilyina, J. G. John, A. Lenton, H. Li, N. S. Lovenduski, J. C. Orr, J. Palmieri, Y. Santana-Falcón, J. Schwinger, R. Séférian, C. A. Stock, A. Tagliabue, Y. Takano, J. Tjiputra, K. Toyama, H. Tsujino, M. Watanabe, A. Yamamoto, A. Yool, T. Ziehn, Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences* **17**, 3439–3470 (2020).

74. S. Chaffron, E. Delage, M. Budinich, D. Vintache, N. Henry, C. Nef, M. Ardyna, A. A. Zayed, P. C. Junger, P. E. Galand, C. Lovejoy, A. E. Murray, H. Sarmento; Tara Oceans coordinators, S. G. Acinas, M. Babin, D. Iudicone, O. Jaillon, E. Karsenti, P. Wincker, L.

Karp-Boss, M. B. Sullivan, C. Bowler, C. de Vargas, D. Eveillard, Environmental vulnerability of the global ocean epipelagic plankton community interactome. *Sci. Adv.* **7**, eabg1921 (2021).

75. G. C. Hays, Ocean currents and marine life. *Curr. Biol.* **27**, R470–R473 (2017).

76. B. A. Ward, B. B. Cael, S. Collins, C. R. Young, Selective constraints on global plankton dispersal. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2007388118 (2021).

77. Y. Silvy, E. Guilyardi, J.-B. Sallée, P. J. Durack, Human-induced changes to the global ocean water masses and their time of emergence. *Nat. Clim. Change* **10**, 1030–1036 (2020).

78. J. D. Zika, J. M. Gregory, E. L. McDonagh, A. Marzocchi, L. Clément, Recent water mass changes reveal mechanisms of ocean warming. *J. Climate* **34**, 3461–3479 (2021).

79. P. Cermeño, S. Dutkiewicz, R. P. Harris, M. Follows, O. Schofield, P. G. Falkowski, The role of nutricline depth in regulating the ocean carbon cycle. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20344–20349 (2008).

80. E. van Sebille, C. Wilcox, L. Lebreton, N. Maximenko, B. D. Hardesty, J. A. van Franeker, M. Eriksen, D. Siegel, F. Galgani, K. L. Law, A global inventory of small floating plastic debris. *Environ. Res. Lett.* **10**, 124006 (2015).

81. L. A. Amaral-Zettler, E. R. Zettler, T. J. Mincer, Ecology of the plastisphere. *Nat. Rev. Microbiol.* **18**, 139–151 (2020).

82. F. L. Hellweger, E. van Sebille, N. D. Fredrick, Biogeographic patterns in ocean microbes emerge in a neutral agent-based model. *Science* **345**, 1346–1349 (2014).

83. E. van Sebille, P. Scussolini, J. V. Durgadoo, F. J. C. Peeters, A. Biastoch, W. Weijer, C. Turney, C. B. Paris, R. Zahn, Ocean currents generate large footprints in marine palaeoclimate proxies. *Nat. Commun.* **6**, 6521 (2015).

84. B. F. Jönsson, J. R. Watson, The timescales of global surface-ocean connectivity. *Nat. Commun.* **7**, 11239 (2016).

85. L. Cheng, K. von Schuckmann, J. P. Abraham, K. E. Trenberth, M. E. Mann, L. Zanna, M. H. England, J. D. Zika, J. T. Fasullo, Y. Yu, Y. Pan, J. Zhu, E. R. Newsom, B. Bronselaer, X. Lin, Past and future ocean warming. *Nat. Rev. Earth Environ.* **3**, 776–794 (2022).

86. G. Salazar, F. M. Cornejo-Castillo, E. Borrull, C. Díez-Vives, E. Lara, D. Vaqué, J. M. Arrieta, C. M. Duarte, J. M. Gasol, S. G. Acinas, Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Mol. Ecol.* **24**, 5692–5706 (2015).

87. R. Massana, A. E. Murray, C. M. Preston, E. F. DeLong, Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Appl. Environ. Microbiol.* **63**, 50–56 (1997).

88. A. E. Parada, D. M. Needham, J. A. Fuhrman, Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).

89. T. Stoeck, D. Bass, M. Nebel, R. Christen, M. D. M. Jones, H.-W. Breiner, T. A. Richards, Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**, 21–31 (2010).

90. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, S. P. Holmes, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

91. Q. Wang, G. M. Garrity, J. M. Tiedje, J. R. Cole, Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).

92. C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, F. O. Glöckner, The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

93. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

94. L. Guillou, D. Bachar, S. Audic, D. Bass, C. Berney, L. Bittner, C. Boutte, G. Burgaud, C. de Vargas, J. Decelle, J. del Campo, J. R. Dolan, M. Dunthorn, B. Edvardsen, M. Holzmann, W. H. C. F. Kooistra, E. Lara, N. Le Bescot, R. Logares, F. Mahé, R. Massana, M. Montresor, R. Morard, F. Not, J. Pawlowski, I. Probert, A.-L. Sauvadet, R. Siano, T. Stoeck, D. Vaulot, P. Zimmermann, R. Christen, The Protist Ribosomal Reference database (PR^2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, D597–D604 (2013).

95. J. Karstensen, M. Tomczak, Age determination of mixed water masses using CFC and oxygen data. *J. Geophys. Res.* **103**, 18599–18609 (1998).

96. E. Pante, B. Simon-Bouhet, Marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLOS ONE* **8**, e73051 (2013).

97. C. R. Gazulla, A. Auladell, C. Ruiz-González, P. C. Junger, M. Royo-Llonch, C. M. Duarte, J. M. Gasol, O. Sánchez, I. Ferrera, Global diversity and distribution of aerobic anoxygenic phototrophs in the tropical and subtropical oceans. *Environ. Microbiol.* **24**, 2222–2238 (2022).

98. J. Cavender-Bares, K. H. Kozak, P. V. A. Fine, S. W. Kembel, The merging of community ecology and phylogenetic biology. *Ecol. Lett.* **12**, 693–715 (2009).

99. J. M. Chase, J. A. Myers, Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R Soc. B Biol. Sci.* **366**, 2351–2363 (2011).

100. A. Baselga, C. D. L. Orme, Betapart: An R package for the study of beta diversity. *Methods Ecol. Evol.* **3**, 808–812 (2012).

101. A. Baselga, Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* **19**, 134–143 (2010).

102. E. B. Graham, A. R. Crump, C. T. Resch, S. Fansler, E. Arntzen, D. W. Kennedy, J. K. Fredrickson, J. C. Stegen, Coupling spatiotemporal community assembly processes to changes in microbial metabolism. *Front. Microbiol.* **7**, 1949 (2016).

103. P. Legendre, M. J. Anderson, Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* **69**, 1–24 (1999).

104. B. H. McArdle, M. J. Anderson, Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology* **82**, 290–297 (2001).

105. R Core Team, R: A Language and Environment for Statistical Computing. R Found. Statistical Computing (2014); http://r-project.org/.

106. K. Grasshoff, K. Kremling, M. Ehrhardt, *Methods of Seawater Analysis* (John Wiley & Sons, 2009).

107. T. P. Boyer, J. I. Antonov, O. K. Baranova, H. E. Garcia, D. R. Johnson, A. V. Mishonov, T. D. O'Brien, D. Seidov, I. Smolyar, M. M. Zweng, C. R. Paver, R. A. Locarnini, J. R. Reagan, C. Coleman, A. Grodsky, *World Ocean Database 2013* (2013).

108. J. M. Gasol, X. A. G. Morán, Flow cytometric determination of microbial abundances and its use to obtain indices of community structure and relative activity, in *Hydrocarbon and Lipid Microbiology Protocols: Single-Cell and Single-Molecule Methods*, T. J. McGenity, K. N. Timmis, B. Nogales, Eds. (Springer, 2015), Springer Protocols Handbooks, pp. 159–187.

109. D. C. Smith, F. Azam, A simple, economical method for measuring bacterial protein synthesis rates in seawater using 3H-leucine. *Mar. Microb. Food Webs* **6**, 107–114 (1992).

110. P. D. Schloss, S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G.

Thallinger, D. J. Van Horn, C. F. Weber, Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).

120. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

121. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).

122. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).

123. S. W. Kembel, P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, C. O. Webb, Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).

124. M. Tomczak, Some historical, theoretical and applied aspects of quantitative water mass analysis. *J. Mar. Res.* **57**, 275–303 (1999).

125. A. R. Longhurst, *Ecological Geography of the Sea* (Academic Press, 2007).

126. A. Bergamasco, P. Malanotte-Rizzoli, The circulation of the Mediterranean Sea: A historical review of experimental investigations. *Adv. Oceanogr. Limnol.* **1**, 11–28 (2010).

127. S.-D. Ayata, J.-O. Irisson, A. Aubert, L. Berline, J.-C. Dutay, N. Mayot, A.-E. Nieblas, F. D'Ortenzio, J. Palmiéri, G. Reygondeau, V. Rossi, C. Guieu, Regionalisation of the Mediterranean basin, a MERMEX synthesis. *Prog. Oceanogr.* **163**, 7–20 (2018).

128. H. Wickham, *Ggplot2: Elegant Graphics for Data Analysis* (Springer Science+Business Media LLC, ed. 2, 2016).