Patterns of eukaryotic diversity from the surface to the deep-ocean sediment

Tristan Cordier1,2,*, Inès Barrechea Angeles1,3, Nicolas Henry4,5, Franck Lejerzowicz6,7, Cédric Berney4,5, Raphaël Morard8, Angelika Brandt9,10, Marie-Anne Cambon-Bonavita11, Lionel Guidi12, Fabien Lombard12,13, Pedro Martinez Arbizu14,15, Ramon Massana16, Covadonga Orejas17, Julie Poulain5,18, Craig R. Smith19, Patrick Wincker18, Sophie Arnould-Haond20, Andrew J. Gooday21,22,*, Colomban de Vargas4,5,*, Jan Pawlowski1,23,24,*

Remote deep-ocean sediment (DOS) ecosystems are among the least explored biomes on Earth. Genomic assessments of their biodiversity have failed to separate indigenous benthic organisms from sinking plankton. Here, we compare global-scale eukaryotic DNA metabarcoding datasets (18S-V9) from abyssal and lower bathyal surficial sediments and euphotic and aphotic ocean pelagic layers to distinguish plankton from benthic diversity in sediment material. Based on 1685 samples collected throughout the world ocean, we show that DOS diversity is at least threefold that in pelagic realms, with nearly two-thirds represented by abundant yet unknown eukaryotes. These benthic communities are spatially structured by ocean basins and particulate organic carbon (POC) flux from the upper ocean. Plankton DNA reaching the DOS originates from abundant species, with maximal deposition at high latitudes. Its seafloor DNA signature predicts variations in POC export from the surface and reveals previously overlooked taxa that may drive the biological carbon pump.

INTRODUCTION

Deep-ocean sediment (DOS) ecosystems cover more than half of Earth’s surface and remain one of the least explored ecosystems on the planet. This vast and heterogeneous environment provides habitats for diverse biological communities that support fundamental ecological processes and services, such as nutrient recycling for the healthy functioning of ocean ecosystems and carbon sequestration for the regulation of Earth’s climate over geological time scales (1). The DOS is exposed to growing anthropogenic pressures, notably from climate change (2, 3), deep-sea mining (4), oil and gas exploitation, and bottom trawling (5), making a scientifically informed protection of its biodiversity a matter of the highest importance (6–8).

For more than 50 years, a considerable effort has been devoted to understanding the diversity and biogeography of benthic organisms thriving in the DOS (9). However, the enormous extent of this habitat and its remote location under several kilometers of water means that only a minute proportion has ever been sampled. Most previous studies have focused on morphological analyses of the macro- and mega-fauna, which typically show high levels of α diversity and small-scale faunal patchiness (9–11), and have recently been proposed as biological indicators for deep-ocean monitoring and conservation (12). Less attention has been paid to the microbial and meiofaunal organisms that numerically dominate DOS communities (13, 14) but can hardly be identified using classical morphotaxonomic approaches. Studying planktonic organisms sinking to DOS is hampered by similar technical limitations related to great depths (resulting in poor spatial coverage of sinking plankton datasets) and limited morphological identification (but for shell-building taxa that can keep distinctive features once in the sediment). Their study is hence often approached indirectly, using sediment traps to capture the sinking flux of taxa over time that contribute most to the biologically driven carbon sequestration in the deep ocean before they reach the sediment (15–17).

The development of high-throughput environmental genomics has begun to fill these gaps in knowledge, revealing substantial unknown diversity among viruses (18) and prokaryotes (19–22) from DOS. Yet, the use of genomics to explore DOS eukaryotes has been limited and focused mostly on particular taxonomic groups (23–25) or geographic regions ([21, 26–28], but see [29]). One major challenge in interpreting molecular data from DOS is to distinguish DNA reads that belong to indigenous benthic eukaryotes from those originating from pelagic
organisms that sink through the water column and leave their DNA traces in the sediments (28, 30–32).

Here, we tackle these problems by comparing a newly generated, global-scale DNA metabarcoding dataset of total eukaryotic diversity from deep oceanic surficial sediments (418 samples collected during 15 oceanographic cruises from 2010 to 2016; table S1) to comparable published datasets from euphotic (1160 samples from the Tara Oceans expeditions) (33, 34) and aphotic (138 samples from the Tara Oceans and Malaspina expeditions) (35) zones across the world ocean. Together, these represent the first consistent molecular meta-dataset spanning the three main open-ocean realms (pelagic euphotic, pelagic aphotic, DOS) at a global scale across 447 sampling sites (Fig. 1A). We assembled ~2.42 billion DNA reads (table S2), produced by polymerase chain reaction (PCR) amplification of the V9 region of the 18S ribosomal RNA gene, and processed them using the DADA2 workflow to infer amplicon sequence variants (ASVs).

On the basis of taxonomic annotations of ASVs using the SILVA and PR database sequences and on the occurrence of a highly conserved DNA sequence motif across eukaryotes, we discarded prokaryotic, plastidic, and mitochondrial ASVs, as well as technical artifacts, allowing us to focus on eukaryotic diversity.

RESULTS AND DISCUSSION
Eukaryotic diversity from the ocean surface to the DOS
We obtained a total of 242,465 eukaryotic ASVs represented by ~1.95 billion DNA reads (Fig. 1B). Only 3806 (1.6%) of these ASVs were detected in all three realms, while 6382 pelagic ASVs were detected in DOS. These ASVs were assumed to correspond to sinking pelagic organisms, mainly plankton, although 29 ASVs (representing 1.29% of the reads of these ASVs) could be ascribed to nekton (e.g., dead vertebrates), which also contribute to the downward flux of organic matter. From the metazoa fraction of sinking pelagic organisms, we curated benthic animals with known meroplanktonic larvae (224 ASVs; see Materials and Methods). The number of ASVs found exclusively in DOS, here assumed to correspond to indigenous deep-sea benthic organisms, was comparable to that found in the pelagic realms, although there were 25 times more pelagic DNA reads in our meta-dataset (Fig. 1B). To account for this variation in sequencing effort, we subsampled each aggregated dataset per realm 1000 times at identical sequencing depths (1 Mio reads) and analyzed the diversity of ASV, together with their distribution and abundances within and across the pelagic (euphotic and aphotic zones) and DOS (sinking pelagic and benthic organisms; fig. S1) realms. This indicated that, although nearly half of eukaryotic DNA reads represent sinking planktonic ASVs, the ASV richness in the pelagic zones (e.g., Dactylopodida amoebae (6.5%), Chromadoria nematodes (5.3%), Monothalamid foraminifera (4.4%), and Oligohymenophorea ciliates (3.7%). Nearly two-thirds (60.1%) of the benthic eukaryotic ASVs (representing 47.8% of the reads) could not be taxonomically annotated using current reference taxonomic databases and a similarity cutoff of 85%, and many of them matched a reference sequence with less than 80% similarity (Fig. 2B). By comparison, the proportion of unassigned ASVs in the pelagic samples is 24.7% (2.6% of the reads) in the euphotic zone and 13.9% (4.1% of the reads) in the aphotic zone (Fig. 2C and table S3).

To better characterize the taxonomic breadth of the unknown eukaryotic diversity in the ocean, we clustered all unassigned eukaryotic ASVs into operational taxonomic units (OTUs) at decreasing similarity thresholds (fig. S6). This revealed that more than 10,000 benthic OTUs are formed with a 90% similarity cutoff, well below the species/genus threshold levels (40). These results indicate that previously unknown high-rank eukaryotic groups with diverse and abundant sublineages likely make up most of the diversity thriving in DOS. A similar number of 90% cutoff OTUs is formed in the pelagic euphotic zone, but their relative abundance (2.6% of the reads) is much lower than for benthic diversity (47.8% of the reads). Many of these unassigned pelagic ASVs may thus correspond to rare unknown eukaryotes or rare intraspecific/intragenomic variants of known eukaryotes with unusually high polymorphism (41). Among the known taxa (39.9% of benthic ASVs and 52.2% of benthic reads), our data for selected typical deep benthic macrofaunal and meiofaunal groups show that some are relatively well represented in the current databases (e.g., polychaetes and nemerteans; fig. S7), while others remain poorly represented (e.g., foraminifera and nematodes).

Biogeography of deep-ocean benthic eukaryotes
Analysis of the strictly benthic eukaryotic diversity revealed global biogeographic patterns among DOS communities. The overall richness...
of benthic ASVs tends to decrease with increasing latitude (fig. S8). Benthic richness follows a bell-shaped trend with increasing export flux of particulate organic carbon (POC) from the surface and particularly with increasing POC reaching the seafloor (the latter explaining up to 10.7% of the variation in overall benthic richness). This pattern is not consistent across benthic groups, with nematodes, foraminifera, and molluscs being notably more diverse at higher latitudes and at sites with higher POC flux reaching the seafloor (fig. S8). The compositional structure of deep benthic communities is in broad agreement with abyssal biogeographic provinces (42) (PERMANOVA $R^2 = 0.136$, $P < 0.001$) and somewhat structured along a gradient of absolute latitude at a global scale ($R^2 = 0.051$, $P < 0.001$), although polar regions are separated on the ordination (Fig. 3A). We used a selection of environmental parameters (see Materials and Methods) in a stepwise model building for constrained ordination to explain the observed pattern. The model explained up to 15.1% of the benthic compositional variation, with seabed nitrate, POC export from the surface, and POC reaching the seafloor together explaining 11.2% (table S4), in line with previous findings on the role of POC export in shaping benthic prokaryotic communities (19).

The average proportion of shared ASVs between pairs of samples as a function of their geographical distance was relatively stable
which the initial similarity is halved (table S5). These results indicate linear models and lower halving distances, i.e., the distance after have a steeper distance-decay, as indicated by the greater slope of (nematodes, foraminifera, and amoebae). The former also generally benthic macrofaunal groups (molluscs and polychaetes) tend to measure of the local presence/absence patchiness) indicate that initial similarities (i.e., similarity at 1 km distance that provides a their community similarity with increasing spatial separation. The environmental drivers. We calculated key distance-decay parameters indicating dispersal limitation among benthic taxa or shifts in envi-
ronmental drivers. We calculated key distance-decay parameters for selected benthic groups (table S5), i.e., the rate of decrease in their common planktonic larval phases (43). We lastly compared spatial structures and distance-decay parameters for whole benthic eukaryotic communities with those for water column communities (table S5). Overall, benthic communities are more spatially structured (mantel: $r = 0.454$, $P < 0.001$) than pelagic communities (euphotic: $r = 0.147$, $P < 0.001$; aphotic: $r = 0.228$, $P < 0.001$) and have lower initial similarity, steeper distance-decay, and smaller halving distances (table S5 and fig. S9), consistent with previous findings for benthic compared to pelagic communities of bacteria in the world ocean (44). This was also shown by fitting neutral community assembly models to each realm dataset, indicating that up to 10 km, after which it decreased steadily (fig. 3B), possibly indicating dispersal limitation among benthic taxa or shifts in environmental drivers. We calculated key distance-decay parameters for selected benthic groups (table S5), i.e., the rate of decrease in their community similarity with increasing spatial separation. The initial similarities (i.e., similarity at 1 km distance that provides a measure of the local presence/absence patchiness) indicate that benthic macrofaunal groups (molluscs and polychaetes) tend to have a stronger local turnover than meiofaunal or protistan groups (nematodes, foraminifera, and amoebae). The former also generally have a steeper distance-decay, as indicated by the greater slope of linear models and lower halving distances, i.e., the distance after which the initial similarity is halved (table S5). These results indicate that dispersal limitation or environmental filtering (or a combination of both) may be stronger for macrofaunal benthic organisms than for meiofaunal or microbial eukaryotes, although macrofaunal taxa are usually thought not to be limited in their dispersion, owing to their common planktonic larval phases (43).
taxonomic groups of the pelagic euphotic zone tend to be more geographically widespread, while benthic groups tend to be less widespread than expected by neutral models (Fig. S10). The importance of benthic community patchiness at local scales is reinforced by the variation of beta diversity as a function of increasing sampling scale (Fig. 3C). The compositional variations within single sediment cores (8 to 10 cm in diameter) and between sediment cores collected from the same multicorer deployment (30 cm to 1 m apart) were comparable to that between deployments at a single station. This confirms the high degree of deep benthic community variation at local scales observed before in the case of selected groups of macro- and meiofauna by both morphological (45) and DNA-based (24) studies.

**Eukaryotic plankton DNA signature on the DOS**

Our study provides the first DNA-based insight into the qualitative and semiquantitative importance of the eukaryotic plankton diversity reaching the DOS at a global scale and thus driving the biological transfer of atmospheric carbon to the seafloor. The taxonomic composition of the 6382 planktonic eukaryotes in the DOS is roughly similar to that in the pelagic euphotic and aphotic zones (Fig. 2). In terms of relative abundance, however, planktonic DNA reads from sediment samples are mainly distributed among diatoms (15.4%) and various groups of rhizarians (26.9%) but include relatively few copepods (1.3%) and dinoflagellates (4.7%), which numerically dominate the plankton in upper ocean layers (Fig. S4 and table S3). Plankton DNA on the DOS also include abundant ASVs assigned to the diplonemids (2.3%) and fungi (4.2%) that are common in the aphotic zone, supporting previous findings that the DOS accumulates DNA from organisms occurring throughout the entire water column (28).

The pelagic ASVs reaching the DOS are generally among the most abundant planktonic eukaryotes in the water column (together representing 75.8% of the reads in the euphotic and 79.3% of the reads in the aphotic), although not all abundant pelagic ASVs are present in the DOS (Fig. 4A). We explored whether their occurrence in the DOS could be explained by their size distribution and by their trophic modes. We found no evidence that larger planktonic taxa are more likely to reach the DOS than smaller taxa (Fig. 4C), reinforcing the idea that most sinking plankton is transferred to the sediment through organic aggregates and not as individual organisms. However, the relative abundance of large planktonic taxa was higher in high-latitude sediments, especially in the Arctic (Fig. 4D), consistent with the trend of increasing sea-surface plankton size with increasing latitude and nutrient content (46, 47). Notably, the proportion of parasitic protists among sinking pelagic ASVs (13.7%) is greater than among nonsinking pelagic ASVs (2.9%; Fig. 4E), indicating that pelagic parasites are more likely to reach the DOS. Their relatively higher abundance in temperate and tropical latitude sediments suggests their ecological importance at these latitudes (Fig. 4F). Greater transfer of parasites to the DOS could reflect their ability to infect and kill larger hosts and/or the massive amounts of resistant and relatively dense propagules that they typically release after host infection and that could persist in sinking aggregates (48–51).

We aggregated our data for each entire realm to investigate whether the most abundant sinking pelagic ASVs in the water column are also the most abundant among sinking pelagic ASVs detected in the sediment, providing insight into their overall taphonomy (Fig. S11). Similar abundance profiles would indicate that the structure of sinking plankton assemblages is overall preserved in surface sediment, whereas dissimilar profiles would indicate that sinking assemblages are consumed or repackaged during their downward transfer in a
nonrandom manner. For instance, the abundance profiles of sinking pelagic ASVs of copepods, dinoflagellates, diatoms, acanthareans, eupelagonemids, hydrozoans, and spumellarians in both the euphotic and aphotic zones are similar to their abundance profiles in the DOS (fig. S11), suggesting that their structure in pelagic ecosystems is preserved on the seafloor. Their profiles in the DOS better mirrored those in the aphotic zone (higher $R^2$ of linear models), suggesting that the transformation of sinking material occurs mostly in the upper oceanic layer. However, this was not the case for diatoms and colloidarians, for which the abundance profiles in the euphotic and aphotic layer were similarly preserved in the DOS, likely because of their higher propensity to sink and form aggregates (52, 53). Although plankton DNA transfer to sediment has yet to be investigated [but see (54–56)], notably by accounting for physical processes (e.g., deep waters currents and vertical mixing) that interact with biological and ecological processes in the deposition of sinking material, our study reinforces the significance of the DOS as a DNA archive of upper-ocean biodiversity and ecology and a source of potential new proxies to document past environmental changes (57).

Last, we attempted to correlate our global DOS plankton biodiversity dataset to yearly average POC export from the surface and the fraction of it reaching the DOS, as estimated, respectively, by thorium-derived export measurements, modeled at a global scale (58), and by evaluating the efficiency of POC transfer through the water column based on sediment trap POC flux data, net primary production estimates, and sea surface temperatures (59). Overall, the pelagic ASVs detected in the DOS represent 21.4% of the DNA reads obtained from sediment samples. The proportion of DNA reads of pelagic origin in the sediment follows an increasing trend from low to high latitudes (Fig. 5A). This proportion also broadly approximates POC export from the surface ($R^2 = 0.23$, $P < 0.001$), despite the higher remineralization rates at productive high latitudes (60). Furthermore, the composition of plankton DNA in the DOS can predict up to 58% of the variation in POC export from the surface and 57% of the POC reaching the seafloor using cross-validated random forest regressions (Fig. 5B). We used a multivariate regression method to identify the sinking pelagic ASVs that best explain the variation of POC export and POC reaching the seafloor (Fig. 5C). Not unexpectedly, diatoms and dinoflagellates (52, 61) were important contributors, but we also identified some previously overlooked taxa that are not usually considered to contribute to POC export, such as alveolate parasites (MALV-II), cercozoans, chrysophytes, and several unknown eukaryotes [see also (51)]. Our time-integrated data from the DOS therefore highlight previously underappreciated taxa that may be keystone drivers of the biological carbon pump.

**Toward a holistic view of ocean biodiversity and ecosystem processes**

Our global molecular meta-dataset from the ocean surface to the DOS provides the first unified vision of eukaryotic biodiversity patterns across the three dimensions of the world ocean (Fig. 1). It shows that the DOS is an extremely rich and unique realm with a strong connection to the water masses above that is reflected in the pelagic DNA signature (Figs. 1, 4, and 5 and fig. S11). Although focused on
smaller-sized organisms (eukaryotic microbes and meiofauna), these DNA-based results are broadly consistent with morphological evidence from larger animals for high deep-sea benthic diversity and small-scale patchiness (high local species turnover) (9–11).

The DOS appears to be much more diverse than oceanic waters (Fig. 1 and fig. S1) and is composed of communities of mostly unknown eukaryotes (Fig. 2) that display clear biogeographic patterns at global scales and considerable patchiness at local scales (Fig. 3). These patterns are likely driven by the flux of sinking organic aggregates and fecal pellets (Fig. 3, table S4, and fig. S8) (62, 63). Our data also show that the DNA-based plankton abundance profiles are broadly preserved in the DOS and that the transformation of sinking material appears to occur mostly between the euphotic and aphotic layers (fig. S11). The deposition of eukaryotic plankton is maximal at productive high latitudes (Fig. 5), and the plankton contribution to time-integrated sedimentary DNA broadly approximates the yearly average POC export from the surface. Moreover, the taxonomic composition of planktonic assemblages in the DOS is an even better predictor of the variation of POC export from the surface and the fraction of it reaching the seafloor (Fig. 5), indicating that biodiversity is key for ocean carbon export and burial. These DOS assemblages comprise not only taxa that are known to be important drivers of the biological carbon pump but also several taxonomic and functional groups that have been overlooked in what is arguably one of the most fundamental ecological processes of the world ocean.

Together, our results highlight the DOS as one of Earth’s richest modern ecosystems and fossil archives. They underline the need for concerted international efforts to further understand DOS biodiversity and its ecological role in planetary biogeochemical cycles. Our study, together with recent evidence that plankton DNA signal can be preserved in subseafloor sediments (54–57), paves the way for using sedimentary planktonic DNA to complement the microfossil-based proxies currently used to reconstruct ancient oceans, including their biological carbon sequestration processes. We hope it will also...
provide the basis for a more informed and effective stewardship strategy for protecting unique and relatively pristine deep-ocean ecosystems as the exploitation of seabed resources gathers pace.

MATERIALS AND METHODS

DOs sample collection

DOs have been collected during two main projects (deep_sea and eDNAbyss). For the deep_sea project, sediments samples were collected at abyssal depths during eight expeditions to the Arctic, Atlantic, Southern, and Pacific Oceans (table S1). We used disposable sterile spoons to subsample the top surface sediment centimeter (ca. 2 g from 0 to 1 cm) following a nested sampling design: up to three pseudo-replicates per core, up to two cores per deployment (multicorer), and up to three deployments per station (detailed list in table S7). Sediment samples were placed in sterile Falcon tubes with (VEMA, SYSTCOII, KuramBio I, MANGAN’16, ABYSSLINE) and without (MSM39, DIVA3, and BIONION) 6 ml of Lifeguard Preservation Solution (QIAGEN) before being frozen onboard at −20°C. The samples were shipped within −20°C containers to the University of Geneva (Switzerland). Upon arrival, sediment samples were stored at −80°C until extraction of nucleic acids.

For the eDNAbyss project, lower bathyal and abyssal sediments were collected in the Mediterranean Sea, North Atlantic, and Arctic Oceans during five cruises (Arctic: MarMine; North Atlantic: MEDWAVES; Mediterranean: PEACETIME, CANHOV, and ESSNAUT; table S1). For each station, triplicate cores (10 cm in diameter) were collected with a multicorer or with a remotely operated vehicle. Surface sediment (0 to 1 cm) was collected using metallic spatulas previously sterilized with bleach or DNA Exitus, rinsed with ethanol 96% and then nanopure water, and transferred into sterile zip-locked bags, homogenized by mixing and flattened to be stored at −80°C until DNA extraction. When possible, other layers (1 to 3, 3 to 5, 5 to 10, and 10 to 15 cm) were also collected from sediment cores. An empty zip lock bag from the stock used served as a blank sampling and extraction control in several stations along each cruise.

Nucleic acid extractions, PCR amplification, and illumina sequencing

For each of the 320 surface sediment samples collected in abyssal plains of the deep_sea project, we extracted the total RNA and DNA contents of c.a. 2 g of material as in (64), and we generated cDNA from deoxyribonuclease-treated RNA as in (65). We controlled that no carried-over DNA molecules remained in the RNA extracts based on the absence of PCR products after 60 cycles. We amplified by PCR the V9 hypervariable region of the ribosomal 18S gene with the following primer pair: the forward 1389F (5′-TTGTACACCCACCGC3′) and the reverse 1510R (5′-CCTTCYGCAGGTTCACCTAC-3′) as designed in (66). Tag-encoded versions of the primers (a unique 8-nt sequence was added in the 5′ end of each primer) were used to multiplex up to 40 samples per sequencing library. Each sample was amplified in duplicate PCR reactions, and each PCR was performed in a total volume of 25 μl as follows: 19.4 μl of H₂O, 2.5 μl buffer (FastStart, Roche), 0.5 μl of bovine serum albumin (20 mg/ml; Invitrogen UltraPure), 0.5 μl of 10 mM dNTPs (deoxyribonucleotide triphosphate) (Roche), 0.1 μl of FastStart DNA Polymerase (5 U/μl; FastStart, Roche), 0.5 μl of forward and reverse primers at 10 μM, and lastly, 1 μl of DNA or RNA template (or 1.5 μl for some samples that did not amplify with 1 μl). All DNA and RNA samples were measured using the double-stranded DNA (dsDNA) High-Sensitivity Assay Kit and the RNA High-Sensitivity Assay Kit on the Qubit 4 fluorometer (Thermo Fisher Scientific) and diluted at 7 ng/μl prior PCR amplification. The PCR reaction conditions were as follows: pre-denaturation step at 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 1 min, extension at 72°C for 1.5 min, and a final extension at 72°C for 2 min. A PCR-negative control for each unique combination of tag-encoded primers was verified by agarose gel electrophoresis. The two PCR replicates for each sample were combined and quantified using high-resolution capillary electrophoresis (QIAxcel System, QIAGEN). The PCR products were pooled in equimolar concentration within each multiplexed library. Each pool of PCR products was purified using a High Pure PCR Product Purification kit (Roche), following the manufacturer’s instructions. The sequencing libraries were prepared using the TruSeq DNA PCR-Free Library Preparation Kit (Illumina), following the manufacturer’s instructions. The libraries were quantified by quantitative PCR (qPCR) using the Kapa Library Quantification Kit for Illumina Platforms (Kapa Biosystems) and sequenced on a MiSeq instrument (Illumina) using paired-end sequencing for 300 cycles with kit v2.

Within the project eDNAbyss, DNA extractions were performed on about 10 g of sediment using the PowerMax Soil DNA Isolation Kit (QIAGEN, Hilden, Germany), following the manufacturer protocol, except for the last step where incubation of the elution buffer was prolonged 10 min on the spin filter membrane to increase the DNA yield. The first solution of the kit was poured into empty field control ziplock bags, before being extracted along with sediment samples, following the exact same protocol. All DNA extracts were then stored at −80°C (and transported to Genoscope on dry ice) until PCR amplifications. The V9 hypervariable region of the 18S ribosomal RNA (rRNA) gene was amplified by PCR using the same primer pair (1389F and 1510R). Each sample was amplified in triplicates, and each PCR reaction was performed in a total volume of 25 μl with the Phusion High-Fidelity PCR Master Mix with GC buffer (Thermo Fisher Scientific), 0.4 μM final concentration of each primer, 3% of dimethyl sulfoxide, 1× Phusion Master Mix, and 2.5 ng of template DNA (less for few extracts with very low DNA concentration). The PCR reaction conditions were as follows: pre-denaturation step at 98°C for 30 s, followed by 25 cycles of denaturation at 98°C for 10 s, annealing at 57°C for 30 s, extension at 72°C for 30 s, and a final extension at 72°C for 10 min. PCR products were purified using 1.8× AMPure XP beads cleanup (Beckmann Coulter Genomics). Aliquots of purified amplicons were then run on an Agilent Bioanalyzer using the DNA High Sensitivity LabChip kit to check their lengths and quantified with a Qubit Fluorometer to check their quality and concentration. Amplicons generated were then used for preparation of sequencing libraries. Amplicons (100 ng) were directly end-repaired, A-tailed, and ligated to Illumina adapters on a Biomek FX Laboratory Automation Workstation. Library amplification was then performed using a Kapa HiFi HotStart NGS library Amplification kit with the same cycling conditions applied for previous steps and cleaned up by AMPure XP purification (1 to 1 volume). All libraries were then quantified first by Quant-it dsDNA HS (high-sensitivity) assay using a Fluoroskan Ascent instrument (Thermo Fisher Scientific) and then by qPCR with the KAPA Library Quantification Kit for Illumina Libraries (Kapa Biosystems) on an MPro instrument (Agilent Technologies). Library profiles were checked using high-throughput microfluidic capillary electrophoresis system (LabChip GX, PerkinElmer, Waltham, MA). Libraries were then normalized to 10 nM by addition of 10 mM
tris-Cl (pH 8.5) and applied to cluster generation according to the Illumina Cbot User Guide (part no. 15006165). PhiX DNA spike-in was adapted for some libraries (20% instead of 1%) to minimize the loss of data due to low nucleotide diversity at the beginning of the sequencing run. Libraries were sequenced on HiSeq4000 or HiSeq2500 instruments (Illumina) on a paired-end mode. The raw sediment sequencing data have been deposited to the European Nucleotide Archive (ENA) under project accessions PRJEB33873 (eDNAbyss) and PRJEB48517 (deep_sea).

Public 18S-V9 rDNA sequencing datasets

We gathered published datasets (table S2) targeting the V9 hyper-variable region of the eukaryotic 18S rRNA gene and using the same PCR primers pair (1389F and 1510R) used here for the DOS samples. These datasets were produced by studies sampling the euphotic/aplothic zones (33–35, 67–69) and the DOS (30, 67).

Environmental variables

Although some environmental variables and sediment descriptors were collected during the oceanic expeditions from which we collected sediment samples, their heterogeneity led us to extract more homogeneous environmental layers from the Global Marine Environmental Datasets (http://gmed.auckland.ac.nz) to standardize our concatenated dataset across multiple studies. These variables included the surface calcite (calcite, in mole per cubic meter), surface nitrate (nitrate, in micromole per liter), surface silicate (silicate, in micromole per liter), surface phosphate (phosphate, in micromole per liter), average photosynthetically active radiation (PAR_mean, in Einstein per square meter per day), surface pH, average sea surface temperature (sst_mean, in celsius), variation in sea surface temperature (sst_range, in celsius), average surface currents strength (srf_current, in meter per second), primary production (primprod, in mgC·m²/day/cell), average stock of particulate inorganic carbon (PIC_mean, in mole per cubic meter), average stock of POC (POC_mean, in mole per cubic meter), total suspended matter (tsm_mean, in grams per cubic meter), seabed slope (slope, degree), seabed nitrate (sb_nitrate, in micromole per liter), seabed silicate (sb_silicate, in micromole per liter), seabed-dissolved oxygen (sb_o2dissolve, in milliliter per liter), seabed-utilized oxygen (sb_o2utilized, in milliliter per liter), seabed temperature (sb_temp, in celsius), seabed salinity [sb_salinity, practical salinity scale (PSS)], and average temperature in the water column (wat_col_temp, in celsius). We also extracted the estimated POC export at 100 m depth below the surface (POC_export, g C m⁻² per year) (58) and the POC fraction reaching the seafloor (POC_seafloor, g C m⁻² per year) (59, 70). The values of each environmental variable for each sample analyzed in this study were extracted from the environmental layers with their Global Positioning System (GPS) coordinates (table S7).

Raw sequencing data processing

For the deep_sea dataset, the sequencing libraries were demultiplexed using Double Tag Demultiplexer (DTD) software (https://github.com/yoann-dufresne/DoubleTagDemultiplexer) to screen the R1 and R2 files of each library and retrieve unique tag-encoded primer combinations associated to each sample (allowing no mismatches). We thus produced pairs of fastq files for each sample. For all other illumina datasets (see table S2), we obtained at least one pair of fastq files (paired-end) per sample (some samples were sequenced several times to obtain enough reads that were subsequently merged before statistical analysis). For the datasets produced with the 454 sequencing technology, we obtained one fastq file per sample. We used two R scripts (for paired-end illumina datasets and for 454 dataset, see the "rds_pipeline_illumina.R" and "rds_pipeline_454.R" scripts) to process all the fastq files by batch of 10 samples per job on a High-Performance Computing cluster (Baobab, University of Geneva). The R scripts implemented the key steps of the DADA2 workflow (71) and additional quality filtering steps (see below). The scripts performed the quality filtering with the filterAndTrim function of the DADA2 v1.12.1 R package with default settings, the trimming of primers using the cutadapt v2.4 software (72), the filtering of any read that still contain traces of primers (fastqFilterPrimersMatches function in the fastqUtils.R script), the filtering of any read below 20 bp (fastqFilterWidth function in the fastqUtils.R script), the training of errors models using the learnErrors function of DADA2 with default settings, the inference of ASVs using the dada function with default settings (but for the 454 data, for which we used the HOMOPOLYMER_GAP_PENALTY = −1, BAND_SIZE = 32 options, as recommended by the DADA2 package developing team), and the merging for the overlapping paired-end reads using the mergePairs function with the option “trimOverhang.” Last, we exported the output of the DADA2 workflow, i.e., the “.rds” files that contain all the ASV sequences and their counts for each sample. We also collected summary statistics on each processing step, the trained errors models, and processing time for each sample.

Combining the datasets into a single ASV-to-sample table, taxonomic and functional annotations, matrix curation

We reimported the rds files into R to build an ASV-to-sample table. We first produced an ASV table per dataset, filtered chimeric ASVs with the option “consensus” within each dataset, and aggregated the replicated libraries per biological sample for the deep_sea, tara, and tara_polar datasets. We also aggregated the reads obtained from DNA and RNA libraries generated for each deep_sea sediment samples, since a comparative analysis revealed that the diversity and biogeographic patterns of eukaryotic communities are mostly similar between DNA and RNA [fig. S12; in line with (29, 73)]. We filtered the ASVs detected in the negative controls of the eDNAbyss dataset across the eDNAbyss ASV table. Last, we concatenated each dataset-based ASV-to-sample table into a single one and exported all the ASVs sequences into a fasta file for taxonomic annotations. We used the “assignment-fasta-vsearch” module of the SLIM v0.6 software (74) that wraps the vsearch v2.2.2 software (75). The ASVs were compared to a custom version of PR2 (33) that focus on the 18S V9 region and that include functional annotations (available at https://doi.org/10.5281/zenodo.3768950) and with the SILVA v138 database (76). Taxonomic annotations were the consensus among up to three candidate reference sequences that are above 85% similarity with the query or directly assigned to the reference sequence if the query had a similarity of at least 99%. We also performed another search without restricting a minimum similarity threshold to match an entry reference sequence in the custom V9 version of PR2, to identify non-18S V9 sequences. We focused our analysis on the eukaryotic diversity by discarding any prokaryotic, plastidic ASVs, or any other artifactual ASV. We used taxonomic annotations obtained with the SILVA database to discard prokaryotic ASVs and the annotations obtained with PR2 to discard organelle-derived ASVs. All ASVs that only loosely match any V9 reference sequence (i.e., <20% similarity) were considered as non-18S V9 sequence and were discarded. We also filtered ASVs that did not contain the “GTCG” motif in the first...
four nucleotides in the 5' end. This motif is widely conserved across
eukaryotes, whereas prokaryotes have a “GTCA” motif highly con-
served in those positions. We lastly used the length distribution of
eukaryotic and prokaryotic ASVs to discard any possible prokaryotic
unassigned ASVs (filter set at 116 bp; fig. S13). For downstream tax-
ononomic analyses, we used PR^2-based annotations using the 85%
minimum similarity threshold. We also inferred the trophic mode of
pelagic ASVs (phototrophic/photosymbiotic/parasitic/heterotrophic
protists and zooplankton/other metazoan) based on the matching
candidate reference sequences in PR^2 (up to three candidates per
queried ASV). We ascribed our ASVs to functional groups only if
the functional attributes across candidates above 95% similarity were
unambiguous, i.e., all the candidates for taxonomic assignment share
similar functional attributes. We also used the different size frac-
tions of the Tara Oceans samples to infer the size of pelagic ASVs,
by using a weighted average of relative abundances across the size
fractions of plankton samples (using the lowest mesh size, e.g., from
the 20- to 180-µm size fraction, we used 20 µm in our calculation for
this size fraction).

Classifying eukaryotic ASVs into pelagic or benthic taxa
We considered the ASVs being detected in pelagic samples as plank-
tonic (or nektic), the ASVs detected exclusively in sediment samples
as benthic, and the ASVs detected in both pelagic and sediment data-
sets as sinking plankton [although 29 vertebrates ASVs, comprising
most of the nektan, represented ~1.29% of the sequences detected
in both pelagic and sediment samples (table S3), we hereafter refer
only to sinking plankton]. However, because multiple benthic groups
have meroplanktonic larvae and hence could be detected in pelagic
samples, we manually curated the ASVs assigned to metazoans within
the sinking plankton fraction, based on their known lifestyles. This
was, for instance, the case for some polychaetes, molluscs, echino-
derms, or harpacticoid copepods that were “forced” into benthic
diversity but not for pteropods that were left in the sinking plank-
ton. Of the 546 metazoan ASVs in the sinking plankton fraction,
224 were curated as benthic (see table S8 for the details of this man-
ual curation).

Eukaryotic community diversity and structural analysis
For α and β diversity analyses, we used functions of the vegan R
package v2.5-3 (77), unless specified differently. Because the size
fractionation of the pelagic samples from the Tara Oceans datasets
has a strong effect on α and β diversity measures, we compared
eukaryotic diversity patterns across pelagic and benthic realms by
considering only the richest nano- (3 to 20 µm) and pico- (0.2 to
5 µm) size fractions of pelagic samples. The eukaryotic ASV accu-
cumulation curves as a function of sampling effort were computed
with the specaccum function with the “random” method. We calcu-
lated the Shannon diversity for each sample and compared the dis-
tribution of sample diversity across both pelagic euphotic and aphytic
with the strictly benthic diversity using the stat_compare_means
function of the ggpubr R package v0.2.5 (78) with default settings.
For β diversity analysis, we removed samples with less than 1000
reads and discarded ASVs represented by less than 100 reads throughout
the dataset. We then normalized the ASV-to-sample matrix with the
cumulative sum scaling (CSS) method (79) and computed a Bray-Curtis
dissimilarity matrix between pairs of samples. The dissimilarity matrix
was used to perform a nonmetric multidimensional scaling (NMDS)
ordination on two axes. Sampling depth and absolute latitude variables
were fit to the NMDS as smooth surfaces using the ordisurf func-
tion. The dissimilarity matrix was also used as input of the adonis
function for PERMANOVA models testing for differences between
eukaryotic compositional structure between realms (pelagic euphotic,
pelagic aphotic, and sediment) and along a gradient of absolute latitude
(nested in type of realm and restricting permutations within type of
realm with the “strata” option), using 999 permutations. Last, we
measured the β diversity dispersion within each realm using the
betadisper function and compared the distances distribution to group
centroids between realms using the stat_compare_means function
of the ggpubr R package.

For α diversity and β diversities of the deep-ocean benthic
communities, we focused on oceanic samples only, i.e., we did not
consider the samples from the Mediterranean Sea nor the ones from
the Gulf of California, to avoid potential effects from coastal ecosys-
tems. We calculated the normalized ASV richness per sample for the
overall benthic communities and for selected benthic groups
(nematodes, foraminifera, platyhelminths, polychaetes, molluscs, and
ciliates) by rarefying each benthic sample at the lowest remaining
sequencing depth (after removing planktonic ASVs and after focus-
ing on a given benthic taxonomic group). We used generalized addi-
tive models (GAMs) to investigate the possible nonlinear variation
of richness and Shannon diversity along gradients of latitude, pri-
mary production, and POC export from the surface and reaching
the seafloor using the gam function of the mgcv R package (https://
cran.r-project.org/web/packages/mgcv/) and the smoothing parameter
set to 3. For β diversity analyses, we used a similar approach than
detailed above (CSS-normalized and Bray-Curtis dissimilarity ma-
trix), although here, we did not filter rare ASVs. We used the pcoa
function of the ape R package (80) to perform a principal coordinate
analysis of the Bray-Curtis dissimilarity matrix and calculate the
structural variation explained by the first two axis of the ordination.
We used the ordisurf and envfit functions to respectively fit the
absolute latitude and a selection of environmental variables (seabed
variables: salinity, temperature, silicate, nitrate, dissolved oxygen, POC
reaching the seafloor, and pelagic variables that connect the surface
to the DOS, namely, the primary productivity and the POC export
from the surface) to the ordination. PERMANOVA models were
used to test for differences in benthic composition between abyssal
postulated biogeographic provinces (42) and along a gradient of ab-
solute latitude using 999 permutations. Then, we used the selected
environmental variables in a stepwise model building for constrained
ordination (distance-based redundancy analysis) using the ordi2step
function in a forward direction and using 999 permutations, to ex-
plain the observed benthic community structure. We calculated the
proportion of shared ASVs between pairs of benthic samples to in-
vestigate the decrease of shared ASV proportion as a function of
increasing spatial distance (calculated from GPS coordinates, see the
“companionFunctions.R” script). We also calculated key distance-
decay parameters as in (81), i.e., the initial similarities (Sørensen
similarities between pairs of samples distant to each other by less
than a kilometer), the slope of distance-decay relationship (here in a
log-linear regression form), and the halving distances, i.e., the spatial
distance after which the initial similarities are halved. We calculated
these parameters on an average Sørensen dissimilarity matrix calcu-
lated over 10 rarefaction draws at the minimum sequencing depth
possible (the sequencing depth of the sample with the lowest num-
ber of reads) and by considering the full benthic community or by
focusing on selected benthic groups only, e.g., polychaetes, molluscs, or
platyhelminths (macrofaunal size classes); nematodes or foraminifera ( meiofauna); and amoebae or ciliates (microbes). We used the *mantel* function to test for correlation between spatial distance and community dissimilarities using 999 permutations. We used the *betadisper* function to calculate the β dispersion of benthic communities at increasing sampling spatial scale (between replicates samples of a sediment core, between cores of the same deployment, between deployments at a given station, and within a given abyssal basin). Last, we aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.

We compared the inferred functional attributes (size and trophic mode) of sinking pelagic ASVs with their nonsinking counterparts to explore whether these traits could explain their transfer to the DOS. We also explored the variation of functional groups and size classes of the sinking planktonic communities in the sediment along the gradient of latitude. We investigated the spatial pattern of planktonic abundance on the seafloor by fitting a GAM on the proportion of planktonic DNA reads in the sediment as function of latitude (with smoothing parameter set to 3). Then, we aggregated the planktonic DNA reads of all sediment samples at the station scale and used random forest models to predict the POC export from the surface and the POC reaching the seafloor in a leave-one-out cross-validation approach. We used the *ranger* function of the ranger R package (83) in a regression mode, growing 300 trees and setting the “mtry” function of the ranger R package (83) as 1. We aggregated all samples at the station scale and fitted neutral community assembly models as in (82) to investigate whether the distribution of ASVs within the pelagic and benthic realms are less or more geographically widespread than expected by neutral models.
31. M. V. Lindh, B. M. Maillot, C. N. Shulze, A. J. Gooday, D. J. Amorn, C. R. Smith, M. J. Church, From the surface to the deep-sea: Bacterial distributions across poly metallic nodule fields in the Clarion-Clipperton Zone of the Pacific Ocean. Front. Microbiol. 8, 1696 (2017).

32. R. Morard, F. Lejezerovicz, K. F. Darling, B. Lecroq-Bennet, M. W. Pedersen, L. Orlando, J. Pawlowski, C. De Vargas, M. Kucera, Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroeconomy. Biogeosciences 14, 2741–2754 (2017).

33. C. de Vargas, S. Audic, N. Henry, J. Decelle, F. Mahe, J. Pomask, M. J. Poulin, S. Romac, C. Colom, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horak, O. Jaillon, S. Kandels-Lewis, M. Picheral, J. Poulain, S. Searson, L. Stemmann, F. Not, H. Ogata, S. Pesant, J. Raes, M. E. Stieracki, S. Speich, L. Stegemann, S. Sunagawa, J. Weissenbach, P. Wincker, E. Karsenti, E. Boss, M. Follows, L. Karp-Boss, E. Kriz, E. G. Reynaud, C. Sardet, M. B. Sullivan, D. Velayoudou, Eukaryotic plankton diversity in the sunlit ocean. Science 348, 1261605 (2015).

34. F. M. Ibarbalz, N. Henry, M. C. Brandão, S. Martini, G. Busseni, H. Byrne, L. P. Coelho, D. Morgan-Smith, M. A. Clouse, G. J. Herndl, A. B. Bochdansky, Diversity and distribution E. Lara, D. Moreira, A. Vereshchaka, P. López-García, Pan-oceanic distribution of new D. Forster, G. Lentendu, S. Filker, E. Dubois, T. A. Wilding, T. Stoeck, Improving the estimate of the strength of the ocean’s biological carbon pump. Geophys. Res. Lett. 38, L04606 (2011).

35. M. J. Lutz, K. Caldeira, B. A. Dunbar, M. J. Behrenfeld, Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. J. Geophys. Res. 112, C10011 (2007).

36. L. Guidi, L. Legendre, G. Reygondeau, J. Utz, L. Stegemann, S. A. Henson, A new look at ocean carbon remineralization for estimating deepwater sequestration. Global Biogeochem. Cycles 29, 1044–1059 (2015).

37. D. M. Billert, R. S. Lampitt, A. L. Rice, R. C. F. Mantoura, Seasonal sedimentation of phytoplankton to the deep-sea benthos, 520–522 (2013).

38. D. K. Steinberg, M. R. Landry, Zooplankton and the Ocean Carbon Cycle. Ann. Rev. Mar. Sci. 9, 413–444 (2017).

39. C. R. Smith, F. C. De Leo, A. F. Bernardino, A. K. Sweetman, P. M. Arbizu, Abyssal food limitation, ecosystem structure and climate change. Trends Ecol. Evol. 23, 518–528 (2008).

40. F. Lejezerovicz, Ch. Lejezerovicz, E. Pilet, C. T. Wilding, K. D. Black, J. Pawlowski, High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Sci. Rep. 5, 13932 (2015).

41. J. Pawlowski, P. M. Arbizu, T. Cedhagen, T. A. Wilding, Environmental monitoring through protist next-generation sequencing metabarcoding: Assessing the impact of fish farming on benthic foraminifera communities. Mol. Ecol. Resour. 14, 1129–1140 (2014).

42. L. A. Amaral-Zettler, E. A. McClintem, H. W. Ducklow, S. M. Huse, A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA Genes. PLOS ONE 4, e26372 (2009).

43. A. A. Y. Lee, Z. Liu, S. K. Hu, A. C. Jones, D. Y. Kim, P. D. Countway, L. A. Amaral-Zettler, S. C. Cary, E. B. Sher, F. B. Sher, R. J. Gast, D. A. Caron, Investigating microbial eukaryotic diversity from a global census: Insights from a comparison of pyrotag and full-length sequences of 185 rRNA genes. Appl. Environ. Microbiol. 80, 4363–4373 (2014).

44. D. Xu, R. Li, H. C. Hu, S. Sun, N. Jiao, A. Warren, Microbial eukaryote diversity and activity in the water column of the South China Sea based on DNA and RNA high throughput sequencing. Front. Microbiol. 8, 1121 (2017).

45. C. de Vargas, T. Pollina, S. Romac, N. Le Besot, N. Henry, C. Berg, S. Colin, N. Haentjens, M. Carmichael, D. Le Guen, J. Decelle, F. Mahe, E. Malpot, C. Beumont, M. Hardy, D. Guillant, J. Probert, D. F. Gruber, A. Allen, G. Gorsky, M. Follows, B. B. Cael, X. Pochon, R. Trouble, F. Lombard, E. Boss, M. Prakash, Plankton Planet: ‘Seatizen’ oceanography to assess open ocean life at the planetary scale. bioRxiv , 2020.08.31.263442 (2020).

46. A. K. Sweetman, A. R. Thuber, C. R. Smith, L. A. Levin, C. Mora, C.-L. Wei, A. J. Gooday, D. O. J. Jones, M. Rex, M. Sasahara, J. Ingels, H. A. Ruih, C. F. Frieder, R. Danovaro, L. Wüstberg, A. Baco, B. M. Grupe, A. Pasulka, S. K. Meyer, K. M. Dunlop, L. A. Henry, J. M. Roberts, Major impacts of climate change on deep-sea benthic ecosystems. Elementa 5, 4 (2017).

47. C. M. Preston, C. A. Durkin, K. M. Yamahara, DNA metabarcoding reveals organisms contributing to particulate matter flux to abyssal depths in the North East Pacific ocean. Deep. Res. II Top. Stud. Oceanogr. 173, 104708 (2020).

48. L. Guidi, S. Chaffron, L. Bittner, D. Esvillard, A. Lahilmi, S. Roux, Y. Darzi, S. Audic, L. Berling, J. R. Burns, L. P. Coelho, C. Berney, Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroeconomy. Biogeosciences 14, 2741–2754 (2017).

49. S. Agusti, J. I. Gonzalez-Gordillo, D. Vaqué, M. Estrada, M. I. Cerezo, G. Salazar, J. M. Gasol, M. C. Durante, Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. Nat. Commun. 6, 7608 (2015).

50. T. Biard, E. Bigeard, S. Audic, J. Poulain, A. Gutierrez-Rodriguez, S. Pesant, L. Stegemann, F. Not, Biogeography and diversity of Collocladia (Radiolalia) in the global ocean. ISME J. 13, 1331–1344 (2019).

51. I. Barrenechea Angeles, F. Lejezerovicz, T. Cordier, J. Scheiplitz, S. Mulitza, M. Kucera, D. Ariztiegui, J. Pawlowski, R. Morard, Planktonic foraminifera eDNA signature deposited on the seafloor remains preserved after burial in marine sediments. Sci. Rep. 20351 (2020).

52. J. Lembrecht, G. Hallegraeff, C. J. S. Bolch, C. Woodward, A. Cooper, Hybridisation capture allows DNA damage analysis of ancient marine eukaryotes. Sci. Rep. 11, 3220 (2021).

53. J. B. Kirkpatrick, E. A. Walsh, D. Honti, Fossil DNA persistence and decay in marine sediment over hundred-thousand-year-to-million-year time scales. Geology 44, 615–618 (2016).

54. D. S. M. Billett, R. S. Lampitt, A. L. Rice, R. C. F. Mantoura, Seasonal sedimentation of phytoplankton to the deep-sea benthos, 520–522 (2013).

55. D. K. Steinberg, M. R. Landry, Zooplankton and the Ocean Carbon Cycle. Ann. Rev. Mar. Sci. 9, 413–444 (2017).

56. C. M. Preston, C. A. Durkin, K. M. Yamahara, DNA metabarcoding reveals organisms contributing to particulate matter flux to abyssal depths in the North East Pacific ocean. Deep. Res. II Top. Stud. Oceanogr. 173, 104708 (2020).

57. L. Guidi, S. Chaffron, L. Bittner, D. Esvillard, A. Lahilmi, S. Roux, Y. Darzi, S. Audic, L. Berling, J. R. Burns, L. P. Coelho, C. Berney, Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroeconomy. Biogeosciences 14, 2741–2754 (2017).

58. L. Guidi, S. Chaffron, L. Bittner, D. Esvillard, A. Lahilmi, S. Roux, Y. Darzi, S. Audic, L. Berling, J. R. Burns, L. P. Coelho, C. Berney, Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroeconomy. Biogeosciences 14, 2741–2754 (2017).
Cordier et al., Sci. Adv. 8, eabj9309 (2022) 4 February 2022

SCIENCE ADVANCES | RESEARCH ARTICLE

M. Guardiola, O. S. Wangensteen, P. Taberlet, E. Coissac, M. J. Uriz, X. Turon, Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. PeerJ. 4, e8207 (2016).

Guido, 10.3897/cph.2012.19226 (2012).

T. Voges, J. Max, V. Meyer, R. Morard, V. Müller, G. Patton, A. Paul, A. Poirier, P. Riesen, T. Schade, http://dx.doi.org/10.2312/cr_msm39. (2015).

B. N., C. A., N. Elsner, F. V., O. Golovan, G. Kamenev, 2012 (ANT-XXVIII/3).

A. Brandt, M. Malyutina, I. Alalykina, B. N., C. A., N. Elsner, F. V., O. Golovan, G. Kamenev, 2012 (ANT-XXVIII/3). Cruise report. (2012); http://dx.doi.org/10.13140/RG.2.1.2473.6401.

E. Paulsen, Ø. Sture, MarMine cruise report - Arctic Mid-Ocean Ridge 15.08.2016–05.09.2016 (2016); http://hdl.handle.net/11250/2427715.

C. Kreuziger, L. A. Henry, M. Hermida, J. A. Jimenez, J. L. Lopez-Jurado, P. Lozano, A. Mateo-Ramirez, G. Mateu, J. L. Matoso, C. Menendez, A. Morillas, J. Movilla, A. Oliariaga, M. Paredes, V. Pelayo, S. Pineiro, M. Raka, T. Ramirez, M. Ramos, J. Reis, J. Rivera, A. Romero, J. L. Rueda, T. Salvador, I. Sampaio, H. Sanchez, R. Santiago, A. Serrano, G. Taranto, J. Urra, P. Velez-Belchi, N. Viladrich, M. Zein, Cruise Summary Report - MEDWaves survey (MEDiterranean outflow WAter and Vulnerable Ecosystem5). (2017); https://doi.org/10.5281/zenodo.556516.

C. Guieu, K. Desboeufs, PEACETIME cruise, RV Pourquoi pas ? (2017); https://doi.org/10.17600/17000300.

M. A. Cambon-Bonavita, ESSNAUT 2017 cruise, RV Pourquoi pas ? (2017); https://doi.org/10.17600/17009100.

E. Raugel, CANHRov cruise, RV L’Europe, (2016); https://doi.org/10.17600/16012300.

Acknowledgments: We thank the captains and crews of all 15 expeditions on RV Polarstern (SYSTCOII, ARK XXIV, and ANDIE II), RV Sonne (KuramBio I, Vema-TRANSIT), RV Maria S. Merian (M539), RV Meteor (DIVA3), RV Kilo Moana (MANGAN-16), RV Thompson (ABYSSLINE2), RV L’Atalante (BIONOD), F. Fradillon (ESSNAUT), RV Pourquoi pas?, C. Guieu and C. Tamburini (PEACETIME), RV Sarmiento de Gamboa (MEDWAVES), RV L’Europe and M.-C. Fabri (CANHRov), RV Polar King and E. Ramirez-Llodra (MarMinC), the Carbon Group at Bjarkeins Centre for Climate Research and the Tara Ocean Foundation and its partners. Samples from the KM16 license area were made available for the investigations by C. Ruhlemann and A. Vink from the Federal Institute for Geosciences and Raw Materials (BGR) in Hanover. The computations were performed at the University of Geneva on the Baobab cluster. This article is contribution 125 of Tara Oceans. Funding: We further thank the following sponsors for support: the Swiss National Science Foundation (grants 31003A_159709, 31003A_179129, and P2GEP3_171829), the Swiss Network for International Studies award (20170024), the European Research Council (grant 818449, AGENTS), CNRS (in particular, FR022), Sorbonne University, the French Government “Investissement d’Avenir” program OCEANOMICS (ANR-11-BTBIR-0008), France Génomique, the Genoscope-CEA (ANR-10-INBS-09) for the project eDNAbyss (2016-228), Ifremer for the project Merlin “Pourquoi pas les Abysses,” the German Research Foundation (grants BR1121/20-1 and BR1121/41-1 and the Center/Cluster of Excellence “The Ocean Floor—Earth’s Uncharted Interface”), the German Ministry for Science and Education (grants 03G0223A and 03G0227A), EU JPIO-Oceans project MinigImpact-2 (German BMBF contract 03F0812E), Spanish Ministry of Economy and Competitiveness (CTM2016-75083-R), European Union’s Horizon 2020 Research and Innovation Program (grant 676780 ATLAS), and the Gordon and Betty Moore Foundation (grant GBMF5257 UniEuk). Samples from the UK and OMS license areas in the Clarion-Clipperton Zone were collected as part of the ABYSSLINE project, funded by UK Seabed Resources Development Ltd. (contract SRLD SRL10100). "Author contributions" T.C., C.d.V., and J.Pa. conceived the study. S.A.-H., A.B., M.-A.C.-B., P.M.A., R.Mo., R.Ma., C.R.S., and J.Pa. led the writing of the manuscript. All authors contributed to the writing of the manuscript and approved the final version.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: The raw sediment sequencing data have been deposited to the EVA under project accessions PRJEB48517 (deep_sea) and PRJEB33873 (eDNAbyss). Accession numbers of additional raw sequencing data have been deposited to the ENA under project accessions PRJEB48517

31003A_159709, 31003A_179129, and P2GEP3_171829, the Swiss Network for International Studies award (20170024), the European Research Council (grant 818449, AGENTS), CNRS (in particular, FR022), Sorbonne University, the French Government “Investissement d’Avenir” program OCEANOMICS (ANR-11-BTBIR-0008), France Génomique, the Genoscope-CEA (ANR-10-INBS-09) for the project eDNAbyss (2016-228), Ifremer for the project Merlin “Pourquoi pas les Abysses,” the German Research Foundation (grants BR1121/20-1 and BR1121/41-1 and the Center/Cluster of Excellence “The Ocean Floor—Earth’s Uncharted Interface”), the German Ministry for Science and Education (grants 03G0223A and 03G0227A), EU JPIO-Oceans project MinigImpact-2 (German BMBF contract 03F0812E), Spanish Ministry of Economy and Competitiveness (CTM2016-75083-R), European Union’s Horizon 2020 Research and Innovation Program (grant 676780 ATLAS), and the Gordon and Betty Moore Foundation (grant GBMF5257 UniEuk). Samples from the UK and OMS license areas in the Clarion-Clipperton Zone were collected as part of the ABYSSLINE project, funded by UK Seabed Resources Development Ltd. (contract SRLD SRL10100). "Author contributions" T.C., C.d.V., and J.Pa. conceived the study. S.A.-H., A.B., M.-A.C.-B., P.M.A., R.Mo., R.Ma., C.R.S., and J.Pa. led the writing of the manuscript. All authors contributed to the writing of the manuscript and approved the final version. Competing interests: The authors declare that they have no competing interests. Data and materials availability: The raw sediment sequencing data have been deposited to the EVA under project accessions PRJEB48517 (deep_sea) and PRJEB33873 (eDNAbyss). Accession numbers of additional raw sequencing data of pelagic and sediment samples analyzed in this study can be found in respective publications listed in table S2. All the metadata and R code to process the raw sequencing data and to reproduce the results and figures are available through a GitHub repository: https://github.com/trtcrd/DOS_V9.

Submitted 10 June 2021 Accepted 13 December 2021 Published 4 February 2022