Progress in Oceanography 205 (2022) 102819

Contents lists available at ScienceDirect

Progress in Oceanography

journal homepage: www.elsevier.com/locate/pocean

Deep ocean prokaryotes and fluorescent dissolved organic matter reflect the history of the water masses across the Atlantic Ocean

Markel Gómez-Letona [a,*], Javier Arístegui [a,*], Nauzet Hernández-Hernández [a], Xosé Antón Álvarez-Salgado [b], Marta Álvarez [c], Erick Delgadillo [d], María Pérez-Lorenzo [d], Eva Teira [d], Santiago Hernández-León [a], Marta Sebastián [e]

[a] Instituto de Oceanografía y Cambio Global, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain
[b] Instituto de Investigacións Mariñas (IIM), CSIC, Vigo, Spain
[c] Instituto Español de Oceanografía, CSIC, A Coruña 15001, Spain
[d] Centro de Investigación Mariña da Universidade de Vigo (CIM-UVigo), Departamento de Ecología y Biología Animal, Universidade de Vigo, 36310 Vigo, Spain
[e] Department of Marine Biology and Oceanography, Institut de Ciències del Mar (ICM), CSIC, Barcelona, Spain

ARTICLE INFO

Keywords:
Dissolved organic matter
Prokaryotic community
Water mass
Dark ocean
Atlantic Ocean

ABSTRACT

Organic matter is known to influence community composition and metabolism of marine prokaryotes. However, few studies have addressed this linkage in the deep ocean. We studied the relationship between fluorescent dissolved organic matter and prokaryotic community composition in meso- and bathypelagic water masses along a surface productivity gradient crossing the subtropical and tropical Atlantic Ocean. Four fluorescence components were identified, three humic-like and one protein-like. The distributions of the humic-like components were significantly explained by water mass mixing, apparent oxygen utilisation (AOU) and epipelagic productivity proxies in varying degrees, while the protein-like component was explained only by water mass mixing and epipelagic productivity. The diversity and taxonomic composition of the prokaryotic community differed between water masses: the Nitrosopumilales order dominated in water masses with high AOU and humic-like fluorescence (notably, the SubPolar Mode Water), and tended to co-occur with Marine Group II archaea, the SAR324 clade and Thiomicrospirales, while bathypelagic water masses displayed greater abundances of members of Marinimicrobia, SAR202 and SAR324. Water mass mixing regression models suggested that the distribution of some taxa (e.g., Marinimicrobia, SAR202) was dominated by mixing and selection within the water masses during ageing, while others (chiefly, Alteromonadales) were mostly influenced by local processes. Our results suggest a link between the composition of the prokaryotic community, oxygen utilisation and the signal of fluorescent dissolved organic matter, and has implications for our understanding of the processes that shape carbon cycling and prokaryotic communities in the deep ocean.

1. Introduction

Prokaryotic communities (bacteria and archaea) play a major role in the cycling of organic matter in marine ecosystems (Moran et al., 2016). They interact with the continuum of organic matter –from particulate to dissolved organic matter (DOM)– through assimilation, excretion and remineralisation of organic compounds. This interaction sets the stage for their influence on the biogeochemical cycles of carbon and other elements (Falkowski et al., 2008; Offre et al., 2013), the ocean–atmosphere exchange fluxes of CO_2 (Kwon et al., 2009) and the sequestration of organic carbon (Lechtenfeld et al., 2015).

The myriad of prokaryotes that take part in these processes in the global ocean are influenced by a variety of factors, such as temperature (Laufkötter et al., 2017), nutrient availability (Wohlers-Zöllner et al., 2011), oxygen concentration (Beman and Carolan, 2013) and interactions with other organisms (Seymour et al., 2017). Nonetheless, DOM concentration and composition have been shown to play a major role in how and to which extent prokaryotes interact with organic

* Corresponding authors.
E-mail addresses: markel.gomezletona@ulpgc.es (M. Gómez-Letona), javier.aristegui@ulpgc.es (J. Arístegui), nauzet.hernandez@ulpgc.es (N. Hernández-Hernández), xsalgado@iim.csic.es (X.A. Álvarez-Salgado), marta.alvarez@ieo.es (M. Álvarez), erick_delgadillo@ucol.mx (E. Delgadillo), mplorenzo@uvigo.es (M. Pérez-Lorenzo), teira@uvigo.es (E. Teira), shernandezleon@ulpgc.es (S. Hernández-León), msebastian@icm.csic.es (M. Sebastián).

https://doi.org/10.1016/j.pocean.2022.102819
Received 24 November 2021; Received in revised form 19 March 2022; Accepted 16 May 2022
Available online 21 May 2022

matter (Gómez-Consarnau et al., 2012).

In recent years, there has been significant progress towards understanding how microbial activity influences the distribution and processes that shape the DOM pool in the deep ocean, and vice versa (Moran et al., 2016). The composition of DOM has been posed to be one factor determining its bioavailability (Dittmar et al., 2021): from labile, easily degradable molecules –such as amino acids and sugars– to refractory ones that can last for decades to millennia in the deep ocean (Hansell, 2013). Fluorescence spectroscopy of DOM has proven to be a valuable tool to differentiate between the protein-like, and refractory, humic-like, fractions of the DOM pool (Murphy et al., 2008; Stedmon and Bro, 2008). Furthermore, recent ultra-high resolution analytical techniques, such as Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS), have shown that DOM is extremely diverse in composition, with the number of molecular formulae being in the order of 10^5 (Zark et al., 2017).

Chemical characteristics of DOM however are not the sole factor influencing its reactivity, i.e., how readily it is utilised by microbes (Carlson and Hansell, 2015). Environmental conditions (e.g., concentration of organic compounds and availability of nutrients), the metabolic capabilities of the consumers and ecological interactions also seem to influence the ability of microbes to degrade DOM (Baltar et al., 2021; Dittmar et al., 2021). The prokaryotic community hosts a vast array of metabolisms (Sunagawa et al., 2015) that enables the use of organic matter with varying chemical characteristics, both directly (Bergauer et al., 2018) or indirectly via extracellular enzymatic activity (Arnosti, 2011). This yields distinctive imprints in the optical characteristics of DOM (Catalá et al., 2015a). Microbial remineralisation and deep ocean circulation (responsible for DOM transport from subduction zones and around the global ocean, during which microbes transform DOM) are hence the main drivers of DOM distribution patterns in the dark ocean (Catalá et al., 2015b; Hansell et al., 2009).

Some recent studies have shown that the quality of DOM, rather than its quantity, shape the composition and abundance of the prokaryotic community in deep ocean layers (Guerrero-Feijóo et al., 2017; Ruiz-González et al., 2020). In the deep water masses off a coastal system characterized by seasonal upwelling pulses, Guerrero-Feijóo et al. (2017) found that archaeal taxa were linked to fluorescence properties of DOM while bacteria were related to its average molecular weight. Furthermore, Ruiz-González et al. (2020) analysed samples collected in 8 locations from distinct biogeographic provinces within the Pacific, Indian and Atlantic oceans during the Malaspina 2010 circumnavigation expedition and found a link between amino-acid-like FDOM of both surface and bathypelagic (~4000 m) waters, and the activity of bathypelagic prokaryotes. This link was suggested to arise from vertical connectivity through sinking particles, as these would potentially release bioavailable FDOM and fuel prokaryotic activity.

To further elucidate the links between DOM and deep ocean prokaryotic communities, we combined a thorough characterisation of the optical properties of the DOM pool with high-throughput sequencing of the 16S rRNA gene in central, intermediate and deep water masses along a surface productivity gradient in the tropical and subtropical Atlantic Ocean. We hypothesised that water masses could harbour different FDOM pools and prokaryotic communities based on their initial properties in their respective source regions, their ageing history as they circulate through the world ocean, the mixing between different water masses and the remineralisation processes occurring within the study region. On the one hand, the present study crossing distinct oceanographic environments along a productivity gradient bridges the gap between previous local studies (e.g., Guerrero-Feijóo et al. 2017) or a global ocean surveys with low spatial resolution (Ruiz-González et al., 2020). On the other hand, the characterisation of water masses by means of an optimum multiparameter analysis adds relevant information to interpret the link between the origin and fate of DOM, and microbial organisms in the deep ocean. Thus, the present work provides an opportunity to expand our knowledge on how the distributions of DOM and prokaryotic communities in the mesopelagic (200–1000 m) and bathypelagic (1000–4000 m) ocean are mediated by water mass mixing and biogeochemical processes.

2. Materials and methods

2.1. Sample collection

Samples for this study were collected during the MAFIA cruise (*Migrants and Active Flux In the Atlantic ocean*, April 2015) on board the BIO Hespérides. Seawater samples for biogeochemical analyses were collected at 13 stations in the subtropical and tropical Atlantic (Fig. 1) using a General Oceanics oceanographic rosette equipped with 24L PVC Niskin bottles alongside a Seabird 911-plus CTD, a Seapoint Chlorophyll Fluorometer and a Seabird-43 Dissolved Oxygen Sensor. The Chl-a values provided by the fluorometer were based on the factory calibration. Oxygen sensor measurements were calibrated with samples potentiometrically titrated with the Winkler method following the procedure described in Moreno-Ostos (2012). Oxygen solubility was computed using the equation of Benson and Krause (1984). AOU ($\mu mol \cdot Kg^{-1}$) was calculated by subtracting measured oxygen concentration from the oxygen solubility values at saturation, with respect to the atmosphere.

2.2. FDOM samples

Seawater samples (18 depths per station) for the analysis of fluorescent dissolved organic matter (FDOM) were collected directly from the Niskin bottles into 250 mL glass flasks that had previously been

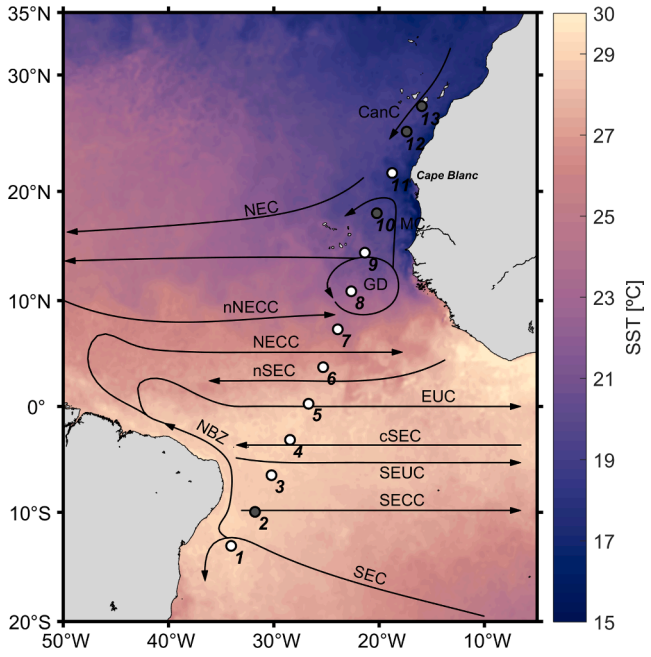


Fig. 1. Sampling stations during the MAFIA cruise. Stations in which samples for 16S amplicon sequencing were analysed are shown in white. The underlying colormap represents the reprocessed Global Ocean OSTIA Sea Surface Temperature (SST) from the CMEMS (product code: SST-GLO-SST-L4-REP-OBSERVATIONS-010-011) for 15/04/2015. Major ocean currents in the study region (based on Stramma and Schott, 1999; Brandt et al., 2008) are also displayed: CanC, Canary Current; NEC, North Equatorial Current; MC, Mauritanian Current; GD, Guinea Dome; nNECC, northern North Equatorial Counter Current; NECC, North Equatorial Counter Current; nSEC, northern South Equatorial Current; EUC, Equatorial Under Current; NBZ, North Brazil Current; cSEC, central South Equatorial Current; SEUC, South Equatorial Under Current; SECC, South Equatorial Counter Current; SEC, South Equatorial Current.

cleaned with solutions of bleach (10%) and HCl (5%), and rinsed with abundant ultrapure water. Samples were left in the dark to allow them reach room temperature prior to their analysis.

2.3. Fluorescence measurements

Fluorescence measurements were performed with a Perkin-Elmer LS55 spectrofluorometer, using excitation (Ex) and emission (Em) slit widths of 10 nm. The excitation range was between 250 and 450 nm and the emission range between 270 and 670 nm. The measurements were collected as synchronic excitation-emission matrices (EEM). A total of 21 synchronic scans were recorded. The first scan started at Ex 250 nm and Em 270 nm (=Ex + 20 nm) and followed with increments of 0.5 nm in both Ex and Em wavelengths up to Ex 450 nm and Em 470 nm. The last scan started at Ex 250 nm and Em 470 nm (=Ex + 220 nm) and followed with increments of 0.5 nm in both Ex and Em wavelengths up to Ex 450 nm and Em 670 nm. Blanks were measured using freshly produced ultrapure water following the same procedure.

Raw measurements were processed using the DOMFluor toolbox (v. 1.7; Stedmon and Bro, 2008) for Matlab (R2017a). First, blank measurements were subtracted from seawater EEMs. Second, EEMs were normalised to the Raman area (RA), which was estimated applying the trapezoidal rule of integration on the emission scan at the 350 nm excitation wavelength in the blank measurements. Recording EEMs in synchronic mode avoids sampling first and second order Rayleigh scatter bands. Inner filter correction was not applied because during the entire cruise the absorption coefficient at 250 nm ($a_{CDOM}(250)$; see Supporting Methods for details) displayed mean values of 0.99 ± 0.25 m^{-1} (max. = 1.95 m^{-1}; n = 250) and this correction is deemed necessary only above 10 m^{-1} (Stedmon and Bro, 2008).

2.4. Parallel factor analysis of fluorescence data

The processed EEMs (see Fig. S1 for examples) were analysed applying a Parallel Factor analysis (PARAFAC) (Stedmon and Bro, 2008) using the DOMFluor toolbox. The PARAFAC model was constructed based on 222 samples (outliers were removed) and it was validated using split-half validation and random initialisation. The resulting model was made up of four components (Fig. S2). For each one, the fluorescence maximum (F_{max}) was recorded in each of the samples. This dataset is available online in the PANGAEA repository (www.pangaea.de) under accession number 942940 (Gómez-Letona et al., 2022b,a).

The optical characteristics of the four components, named according to their fluorescence emission maxima, are summarised in Table S1, along with similar fluorophores found in the literature. The identification of previously described fluorophores was performed using the OpenFluor database (openfluor.lablicate.com, Murphy et al., 2014), based on the combined Tucker Congruence Coefficient of the excitation and emission spectra ($TCC_{ex\cdot em}$). For components $C_{462\text{-}490}$, C_{406} and C_{366}, 6, 7 and 12 matches with high congruence ($TCC_{ex\cdot em} > 0.95$) were found, respectively. $C_{462\text{-}490}$ had an excitation peak at 278 nm, with a secondary one at 370 nm, and a relatively wide emission maxima at 462 nm – 490 nm. This component was found to be related to general humic-like dissolved organic matter. Analogous fluorophores have been described to be positively correlated to AOU in the deep ocean (Catalá et al., 2015b), and were previously found in the study area (Aparicio et al., 2015). C_{454} had an excitation maximum below 250 nm and the emission maximum at 454 nm. No matches with high congruence were found for this component, probably because it was not fully resolved by the PARAFAC model and had a slight contribution from $C_{462\text{-}490}$ (Fig. S2). However, among the matches with lower $TCC_{ex\cdot em}$, C_{454} was found to be similar to fluorophores described as humic/fulvic-like, related to the typical Coble's (1996) peak A (Lapierre and Del Giorgio, 2014; Stedmon and Markager, 2005). C_{406} had excitation and emission maxima at 328 nm and 406 nm, respectively. Similar signals have also been related to humic-like substances, usually of specifically marine origin (Catalá et al., 2015b; Stedmon et al., 2003). Like $C_{462\text{-}490}$, C_{406} has also been linked to AOU and had been previously found in the study area (Catalá et al., 2015b; De La Fuente et al., 2014). Finally, C_{366} was different to the previous components: the excitation peak was located at 286 nm (with a secondary one below 250 nm), while the emission peak was at 366 nm, lower than $C_{462\text{-}490}$ and C_{406}. This protein-like component has been found to be related to material associated with phytoplankton production and has been shown to be partially bioavailable for microbial consumption (Kida et al., 2019; Lønborg et al., 2010; Stedmon et al., 2003).

2.5. DNA sampling, extraction and sequencing

Seawater samples for DNA were collected at stations 1– 9 and 11 from selected meso- and bathypelagic depths: 300 m, 700 m, 1500 m, 2500 m and 3500 m (3200 m for station 5) and kept at 4 °C in the dark (2 h at most) until further processing. Samples were prefiltered with a 20 μm mesh and subsequently filtered through 0.22 μm Sterivex filters (Millipore, SVGP01050) using a peristaltic pump. Upon filtration, filters were sealed with parafilm, flash-frozen with liquid nitrogen and stored at −80 °C.

DNA from biomass retained in the 0.22 μm Sterivex filters was extracted using the PowerWater DNA isolation kit (MoBio Laboratories Inc., CA, USA) according to the manufacturer's instructions. DNA concentration was fluorometrically quantified with a Qubit 3.0 instrument and Qubit dsDNA HS (high sensitivity) Assay Kits (Invitrogen). Prokaryotic community composition was assessed by sequencing the V4 and V5 regions of the 16S rRNA gene (16S rDNA) by using the universal primers "515F" and "926R" (Parada et al., 2016). Amplified regions were sequenced with the Illumina MiSeq platform (paired-end reads; 2 × 250 bp) by Fasteris SA (Geneva, Switzerland).

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI (https://www.ebi.ac.uk/ena) under accession number PRJEB44713.

2.6. 16S amplicon sequence analysis

Bioinformatic analyses were performed at the Marine Bioinformatics platform (Marbits, marbits.icm.csic.es) of the Institut de Ciències del Mar (ICM-CSIC, Barcelona). Primers were removed using *cutadapt* (Martin, 2011). All subsequent analyses were performed in R (3.6; R Core Team, 2019). The *DADA2* package (Callahan et al., 2016) was used to process the amplicon sequence data (*trunclen* = (210, 205), *maxEE* = (2, 4), *minOverlap* = 15). *DADA2* models errors in the Illumina-sequenced amplicon reads to infer exact amplicon sequence variants (ASV) down to one nucleotide difference. Taxonomic assignation of ASVs was carried out using the '*IDTaxa*' function (Murali et al., 2018) from the *DECIPHER* package, with SILVA v. 138 as the training set. Identification of contaminant sequences was performed using the *decontam* package (Davis et al., 2018) applying the frequency method (*threshold* = 0.1). This analysis suggested that sequences of ASVs classified within the orders *Burkholderiales* and *Propionibacteriales* (*Cutibacterium* genus) were contaminants, although not all of them were identified as such. Nonetheless, all sequences assigned to these orders were removed, as they are known reagent/laboratory contaminants (Salter et al., 2014). In the end, samples with at least 1000 reads were preserved for downstream analyses (mean 56028, min. 1092, max. 115430).

2.7. Statistical analyses

All statistical analyses were carried out in R (R Core Team, 2019). Non-metric multidimensional scaling (NMDS; *vegan* package, Oksanen et al., 2019) was performed to visualize the similarity in prokaryotic community composition among samples. Prior to the analysis, the ASV count table was normalised by centred log-ratio (CLR) transformation

(*compositions* package, van den Boogaart et al., 2020) and Euclidean distances were estimated (*vegan* package, Oksanen et al., 2019) based on this normalised table.

Diversity of the prokaryotic community in the different water masses was studied by estimating classical diversity indicators: species richness, the Pielou evenness index (*microbiome* package, Lahti and Shetty, 2019), Faith's phylogenetic diversity index (*picante* package, Kembel et al., 2010) and the Shannon index (*vegan* package, Oksanen et al., 2019). To compute these estimates, amplicon samples with over 30,000 counts (n = 29) were rarefied with permutations using the *EcolUtils* R package (version 0.1, Salazar, 2020).

2.8. Multiparameter water mass analysis

The contribution of each water mass to each sample was objectively quantified applying an optimum multiparameter analysis in Matlab (R2017a). Briefly, based on previous hydrographic studies in the area (e.g. Álvarez et al., 2014; Catalá et al., 2015b), twelve source water types (SWT) were identified in the collected water samples (Fig. 2a, Table S2): Salinity Maximum Water (SMW), Madeira Mode Water (MMW), Equatorial Water (EQ_{13}), Eastern North Atlantic Central Water of 15 °C ($ENACW_{15}$) and 12 °C ($ENACW_{12}$), SubPolar Mode Water (SPMW), Mediterranean Water (MW), Antarctic Intermediate Water of 5 °C ($AAIW_5$) and 3.1 °C ($AAIW_{3.1}$), Circumpolar Deep Water (CDW) and North Atlantic Deep Water of 4.6 °C ($NADW_{4.6}$) and 2 °C ($NADW_2$). The contribution of each SWT to each sample was quantified by solving a set of four conservative linear mixing equations defined by potential temperature (θ), salinity (S), silicate (SiO_4H_4) and the conservative NO tracer ($=O_2 + R_N \cdot NO_3$ with $R_N = 9.3$; Álvarez et al., 2014; Broecker, 1974), plus a fifth equation that constrained the sum of the SWT contributions to 1. These equations were normalised and weighted according to how conservative each variable is, using weights 10, 5, 1, 2 and 100 for potential temperature, salinity, SiO_4H_4, NO and volume, respectively (Álvarez et al., 2014). Therefore, the mixing of a maximum of five of these SWT could be solved simultaneously. Given that we identified 12 SWTs, they were grouped into the following mixing clusters, based on oceanographic criteria of water density and geographic proximity: $AAIW_5$-$AAIW_{3.1}$-CDW-$NADW_2$-$NADW_{4.6}$, $AAIW_5$-$NADW_{4.6}$-EQ_{13}, EQ_{13}-SMW, $NADW_2$-$NADW_{4.6}$-MW, $NADW_{4.6}$-MW-SPMW-$ENACW_{12}$, $NADW_{4.6}$-$AAIW_5$-SPMW-$ENACW_{12}$, $ENACW_{12}$-$ENACW_{15}$, $ENACW_{15}$-MMW, $AAIW_5$-$NADW_{4.6}$-SPMW-EQ_{13}, EQ_{13}-$ENACW_{12}$-SPMW and EQ_{13}-$ENACW_{12}$-$ENACW_{15}$-MMW (Fig. 2a). The initial values of the SWTs were based on previous hydrographic studies (Table S3). Based on the contribution (as %) of SWTs to each sample, archetype values were estimated for each SWT as weighed means of physical and biogeochemical variables, diversity indices, and CLR-transformed abundances of taxonomic orders that contained dominant ASVs (representing > 0.5% of reads at least in a sample) (Álvarez-Salgado et al., 2013; Catalá et al., 2015b). Note that mixed layer samples (here < 100 m) were excluded from the optimum multiparameter analysis because temperature, salinity, SiO_4H_4 and NO do not behave conservatively in this layer. The optimum multiparameter analysis results are available online in the PANGAEA repository (www.pangaea.de) under accession number 942934 (Gómez-Letona et al., 2022b,c).

To assess the degree of variance of FDOM that was explained by conservative water mass mixing and non-conservative biogeochemical processes, multiple linear least squares regressions were performed between FDOM components and water mass proportions alone (mixing models), which account both for the initial values in the source regions plus the large-scale mineralisation processes occurring from the source to the study region. Note that closer to their source regions the proportion of each water type is the highest. This proportion declines as the water type moves away from its source region due to mixing with other water types. At the same time, ageing of the water type occurs in parallel to mixing. Thus, mixing models also retain the large-scale mineralisation processes (Álvarez-Salgado et al., 2013). Individual biogeochemical variables were also included in the regressions (mixing + biogeochemistry models), capturing local biogeochemical processes not considered in the mixing models (Álvarez-Salgado et al., 2013). Among the biogeochemical variables we included the integrated values in the epipelagic layer (as first 200 m) of Chl-a ($Chl\text{-}a_{int}$), the protein-like C_{366} ($C_{366\text{-}int}$) and the chromophoric DOM (CDOM) absorption coefficient at 254 nm ($a_{CDOM}(254)_{int}$, see Supporting Methods for absorption data processing). Integrated values of these variables were estimated by multiplying the discrete measurements by the distance, in meters, between samples. The integrated values of each station were assigned to all samples from that station when estimating the multiple linear regressions. These integrated estimates were included as proxies of epipelagic productivity, as they agreed with satellite-derived net primary production estimates (Fig. S3). They provide insights into potential vertical transport of organic matter, as highly productive areas are known to support high vertical export rates of particles (Fischer et al., 2020), and DOM accumulated in the surface ocean can be eventually transported to the ocean's interior (Baetge et al., 2020). Furthermore,

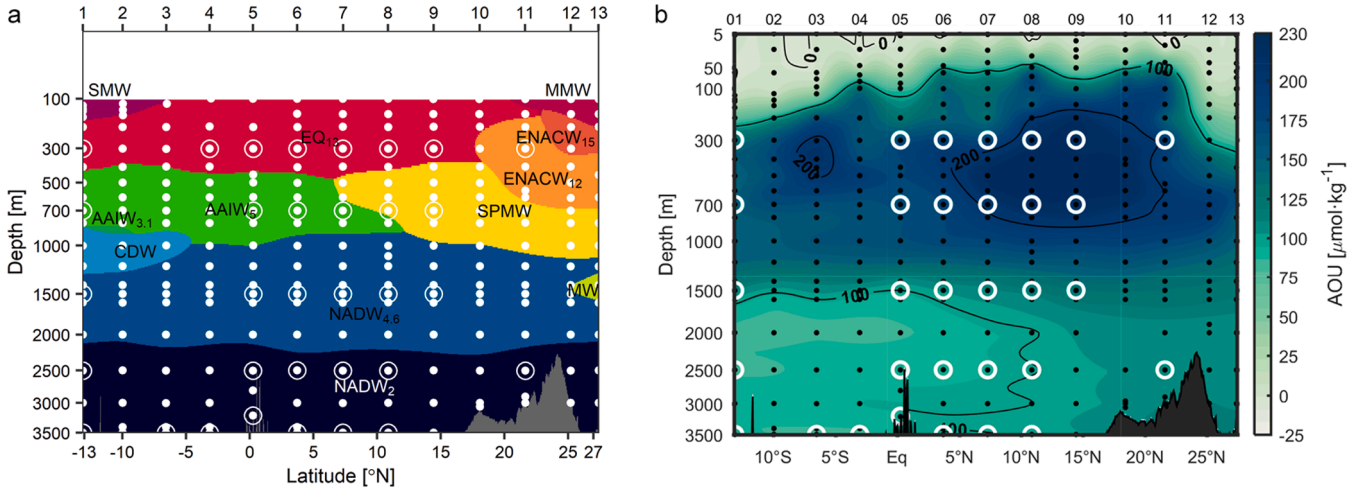


Fig. 2. Distribution of the water masses and Apparent Oxygen Utilisation (AOU) in the study section. (a) Approximate area in which each water mass has its highest contribution is shown. (b) AOU estimates. Dots represent samples of the physical and geochemical variables included in the optimum multiparameter analysis and white circles the locations of amplicon samples. Note that the vertical scale is square root-transformed to allow for a better visualisation of the results. Numbers on top correspond to stations in Fig. 1. Data interpolation was performed with Data Interpolating Variational Analysis (DIVA; Troupin et al., 2012) in Matlab (R2017a).

positive relationships have been found between surface and bathypelagic levels of protein-like FDOM, and also between bathypelagic protein-like FDOM and prokaryotic activity (Ruiz-González et al., 2020), suggesting that surface waters with higher amount of protein-like FDOM lead to higher protein-like FDOM export. The mixing and mixing + biogeochemistry models were also applied to the CLR-transformed abundances of dominant ASVs.

3. Results

3.1. FDOM distribution

The humic-like FDOM components ($C_{462\text{-}490}$, C_{454}, C_{406}) and the protein-like C_{366} component presented remarkably contrasting distributions (Fig. 3). Humic-like components overall had higher fluorescence values below 200 m, although there were latitudinal changes both in the photic and dark layers. $C_{462\text{-}490}$ (Fig. 3a) displayed higher levels north of 5°N, with maximum values (>0.013 RU) at stations 9 and 10 below 300 m, maintaining high values down to the deepest samples. In surface waters $C_{462\text{-}490}$ fluorescence increased from station 1 (~0.001 RU) to station 11 (0.013 RU), off Cape Blanc. Archetype values of $C_{462\text{-}490}$ were higher in water masses located deeper (Table 1): 0.0050 ± 0.0007 RU for the SMW vs 0.0138 ± 0.0003 for the $NADW_2$. Nonetheless, meridional differences were also observed as $AAIW_5$ (core at 0.22 ± 1.86 °N) and SPMW (17.32 ± 1.39 °N), while occupying the same portion of the water column (core at ~700 m), presented markedly different archetype values: 0.0115 ± 0.0003 and 0.0133 ± 0.0003 RU, respectively (Table 1).

C_{454} (Fig. 3b) displayed the highest fluorescence values among all four components, but with similar patterns to $C_{462\text{-}490}$: surface waters showed increased fluorescence from station 1 to 11, and maximum values (>0.027 RU) were found in stations 9 and 10 at depths below 300 m. C_{454} archetype values were highest in the SPMW (0.0261 ± 0.0005 RU; the MW estimate was slightly higher but had a large error, Table 1). C_{454} decreased slightly with depth ($NADW_{4.6}$, 0.0216 ± 0.0008 RU) and notably towards the south ($AAIW_5$, 0.0176 ± 0.0012 RU; SMW, 0.0087 ± 0.0014). There was a marked change in the values of C_{454} throughout the water column close to the equator, from station 5 (<0.019 RU) to station 6 (>0.019 RU). Interestingly, the region between stations 5 and 7 presented a strong SST gradient (Fig. 1), with higher temperatures south of the Equator (station 5), leading to increased stratification of the surface layer (Fig. S4).

Component C_{406} (Fig. 3c) was also widespread in deep waters and presented an increase at stations 8–10, with maximum values of > 0.0105 RU between ~ 150–800 m depth, shallower than the $C_{462\text{-}490}$ and C_{454} maxima. This maximum tended to follow AOU patterns (although not entirely): peaks of AOU exceeding 200 μmol·Kg^{-1} were found in stations 8–10 between depths of ~ 200–800 m (Fig. 2b), where oxygen concentrations were as low as 50 μmol·Kg^{-1} (Fig. S5). Higher fluorescence values at shallower depths were reflected in increased archetype C_{406} values particularly of central and intermediate waters in the tropical North Atlantic. The $ENACW_{12}$ presented the highest C_{406} estimates among water masses (0.0103 ± 0.0004 RU; Table 1), followed by the SPMW (0.0100 ± 0.0003 RU). Again, this contrasted with the archetype values of $AAIW_5$, which were markedly lower (0.0075 ± 0.0003). SMW (0.0063 ± 0.0008) and bathypelagic water masses also displayed low archetype values (Table 1).

The protein-like component C_{366} (Fig. 3d) presented a patchy

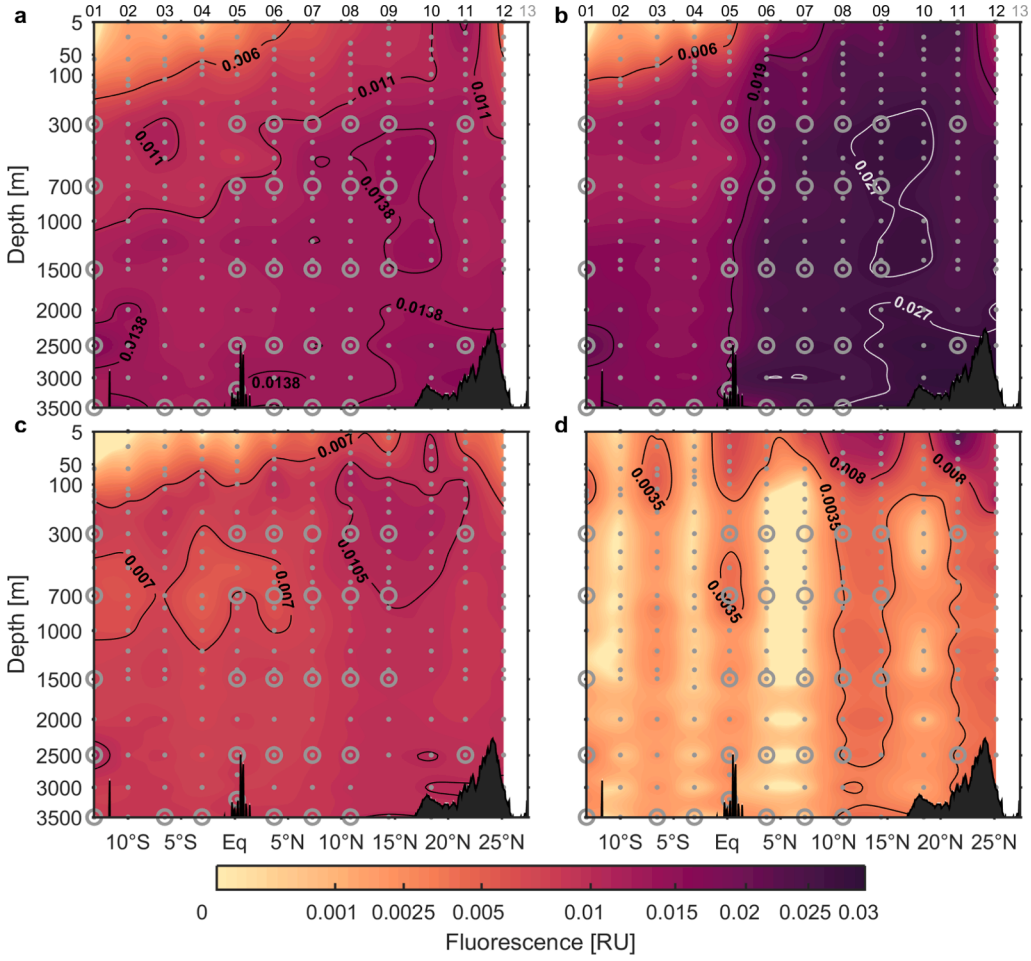


Fig. 3. Distribution of the FDOM components along the cruise section: $C_{462\text{-}490}$ (a), C_{454} (b), C_{406} (c) and C_{366} (d). Grey dots represent FDOM samples and circles indicate samples for which 16S amplicon sequencing was also performed. Fluorescence is represented in Raman Units (RU). Note that both the vertical and colour scales are square root-transformed to allow for a better visualisation of the results. Data interpolation was performed with Data Interpolating Variational Analysis (DIVA; Troupin et al., 2012) in Matlab (R2017a).

Table 1
Water mass characteristics in the study region. Contribution of each water mass to the total sampled volume (as %), along the archetype values of potential temperature (θ), salinity, oxygen concentration (O_2), apparent oxygen utilisation (AOU) and FDOM components. Values are displayed as archetype ± standard error. Water mass acronyms are: Salinity Maximum Water (SMW), Madeira Mode Water (MMW), Equatorial Water (EQ_{13}), Eastern North Atlantic Central Water of 15 °C ($ENACW_{15}$) and 12 °C ($ENACW_{12}$), Subpolar Mode Water (SPMW), Mediterranean Water (MW), Antarctic Intermediate Water of 5 °C ($AAIW_5$) and 3.1 °C ($AAIW_{3.1}$), Circumpolar Deep Water (CDW) and North Atlantic Deep Water of 4.6 °C ($NADW_{4.6}$) and 2 °C ($NADW_2$).

Water mass	Volume [%]	Depth [m]	Latitude [°N]	θ [°C]	Salinity	O_2 [μmol·Kg^{-1}]	AOU [μmol·Kg^{-1}]	$C_{462\text{-}490}$[a] [RU]	C_{454}[a] [RU]	C_{406}[a] [RU]	C_{366}[a] [RU]
SMW	2.88	134 ± 11	−9.68 ± 1.80	21.03 ± 1.11	36.56 ± 0.18	193.0 ± 8.0	26.2 ± 12.7	5.0 ± 0.7	8.7 ± 1.4	6.3 ± 0.8	3.8 ± 1.2
MMW	1.95	130 ± 17	22.13 ± 2.67	16.22 ± 0.63	36.28 ± 0.20	160.7 ± 28.0	79.0 ± 30.9	9.9 ± 1.1	21.5 ± 1.6	8.8 ± 1.4	7.0 ± 2.2
EQ_{13}	15.71	252 ± 23	3.13 ± 1.66	12.70 ± 0.61	35.39 ± 0.09	118.8 ± 6.4	141.3 ± 8.6	10.1 ± 0.4	18.6 ± 1.1	8.9 ± 0.3	2.2 ± 0.4
$ENACW_{15}$	2.09	186 ± 34	25.40 ± 1.07	16.20 ± 0.60	36.40 ± 0.13	188.9 ± 17.1	50.7 ± 19.6	8.8 ± 0.8	20.0 ± 1.2	7.4 ± 0.9	8.3 ± 2.2
$ENACW_{12}$	4.68	413 ± 53	22.40 ± 1.26	11.72 ± 0.51	35.61 ± 0.08	110.5 ± 10.0	153.5 ± 10.5	12.1 ± 0.4	24.9 ± 0.7	10.3 ± 0.4	4.2 ± 0.7
SPMW	8.34	696 ± 56	17.32 ± 1.39	8.20 ± 0.38	35.09 ± 0.04	101.5 ± 7.9	184.5 ± 6.8	13.3 ± 0.2	26.1 ± 0.5	10.0 ± 0.3	3.1 ± 0.5
MW	0.76	1632 ± 386	25.28 ± 1.68	5.17 ± 1.11	35.17 ± 0.09	199.7 ± 16.2	106.9 ± 9.3	13.4 ± 1.4	26.6 ± 2.4	9.6 ± 0.7	4.3 ± 0.8
$AAIW_5$	12.83	713 ± 47	0.22 ± 1.86	6.27 ± 0.33	34.70 ± 0.04	137.8 ± 6.2	162.0 ± 4.7	11.5 ± 0.3	17.6 ± 1.2	7.5 ± 0.3	1.6 ± 0.4
$AAIW_{3.1}$	0.30	1052 ± 315	−9.26 ± 6.56	4.05 ± 0.29	34.60 ± 0.23	188.6 ± 21.4	127.4 ± 21.8	11.2 ± 1.2	15.0 ± 4.9	7.1 ± 1.0	0.5 ± 1.1
CDW	5.89	1917 ± 232	3.92 ± 3.47	3.65 ± 0.25	34.87 ± 0.05	213.1 ± 8.1	105.4 ± 6.6	13.0 ± 0.5	20.9 ± 1.9	8.5 ± 0.4	1.9 ± 0.5
$NADW_{4.6}$	30.00	1485 ± 75	5.99 ± 1.46	4.57 ± 0.17	34.92 ± 0.02	196.2 ± 4.7	115.4 ± 3.7	12.8 ± 0.2	21.6 ± 0.8	8.5 ± 0.2	2.0 ± 0.3
$NADW_2$	14.57	2754 ± 98	4.66 ± 2.12	2.73 ± 0.09	34.94 ± 0.01	240.9 ± 1.9	84.7 ± 1.6	13.8 ± 0.3	22.8 ± 1.1	9.2 ± 0.2	2.0 ± 0.3

[a] $\times 10^{-3}$.

distribution, with high values in surface waters, particularly at stations 8, 9 and 11 (0.025 RU, off Cape Blanc). Deep water masses had lower C_{366} fluorescence but, overall, stations with greater surface C_{366} values also presented higher C_{366} values in deep waters. Archetype values of C_{366} were higher in central and to a lesser extent intermediate water masses, but standard errors were larger than for the other components (Table 1).

3.2. Effects of conservative water mass mixing and biogeochemical processes on FDOM

Multiple linear regressions between the measured variables and the water mass proportions ('mixing models') captured two sources of variability: the effects of 1) conservative water mass mixing, which also retains the effects of temperature and depth on the test variables, and 2) the mineralisation processes that occur from the water mass formation area to the study region (Álvarez-Salgado et al., 2013). We found that depending on the FDOM component, differing degrees of their variability were explained by mixing models (Table 2): $C_{462\text{-}490}$ was explained in 80%, C_{454} in 56%, C_{406} in 57% and C_{366} only in 42%. Including AOU in the models rendered significant results for components $C_{462\text{-}490}$, C_{454}, and C_{406}: the explained variance increased to 87% (+7%), 67% (+11%) and 66% (+9%), respectively (Table 3), while C_{366} was not significantly affected. The positive regression coefficients of AOU suggest a link between prokaryotic oxygen consumption in the water mass since it was formed and the generation of humic-like fluorescence. Including $C_{366\text{-int}}$ and $a_{CDOM}(254)_{int}$ as explanatory variables of FDOM components yielded significant results (but not for Chl-a_{int}). $C_{366\text{-int}}$ increased the explained variance in 1, 2, 10 and 16% for $C_{462\text{-}490}$, C_{454}, C_{406} and C_{366}, respectively, while $a_{CDOM}(254)_{int}$ did it in 4, 15, 14 and 3% (Table 3). These improvements in the explained variance were accompanied by a reduction in the standard error of the residuals.

Table 2
Results of the mixing models of FDOM components and apparent oxygen utilisation (AOU). r^2 is the coefficient of determination, representing the fraction of variance accounted for water mass mixing; SE_{res} the standard error of the residuals; n the number of samples included in the regression; p the p-value of the model.

Parameter	Unit	r^2	SE_{res}	n	p
$C_{462\text{-}490}$[a]	RU	0.80	1.09	178	<0.0001
C_{454}[a]	RU	0.56	4.13	178	<0.0001
C_{406}[a]	RU	0.57	1.09	178	<0.0001
C_{366}[a]	RU	0.42	1.81	178	<0.0001
AOU	μmol·kg^{-1}	0.86	18.47	234	<0.0001

[a] $\times 10^{-3}$.

AOU itself was explained in 86% by the mixing model (Table 2), with significant improvements when adding prokaryotic abundance (+2%), $C_{366\text{-int}}$ (+3%), Chl-a_{int} (+3%) and $a_{CDOM}(254)_{int}$ (+5%) (Table 3). Likewise, Chl-a_{int} and $a_{CDOM}(254)_{int}$ significantly contributed to explain the variance of prokaryotic abundance in the dark ocean (r^2 = 0.94 (+2%) and 0.93 (+1%), respectively; not shown).

Taken together, these results suggest that, while water mass mixing and history, and AOU were the main drivers of humic-like fluorescence variability, there was also a significant link between epipelagic productivity, the production and vertical transport of organic matter (including protein-like DOM), and the generation of humic-like FDOM in the dark ocean.

3.3. Prokaryotic community composition in dark waters of the Atlantic ocean

Prokaryotic community composition was somehow uniform at the phylum level except for three samples which were clearly different from the rest (Fig. 4a): st3-3500 m, st4-3500 m and st9-700 m. Excluding these outlier samples, dark ocean communities were usually dominated in abundance by Crenarchaeota (formerly Thaumarchaeota, comprising 22.5% of reads on average, but ranging between 2.2 and 42.4%), followed by the SAR324 clade (12.6%; 6.3 – 20.6%), Thermoplasmatota (10.8%; 2.3 – 28.1%), Marinimicrobia (10.0%; 4.5 – 15.1%), the

Table 3
Results of the multiple linear regressions combining biochemical variables with the mixing models. *Parameter 1* corresponds to the dependent variable of the regression and *Parameter 2* to the predictor biochemical variable added to the mixing regression model. *Unit 1* and *Unit 2* are their units, respectively. *Coefficient* is the regression coefficient of *Parameter 2*; r^2 the coefficient of determination, representing the fraction of variance accounted for the model; SE_{res} the standard error of the residuals; n the number of samples included in the regression; p_{coef} the p-value of the coefficient of *Parameter 2*. AOU stands for apparent oxygen utilisation. *ns* = not significant ($p > 0.05$).

Parameter 1	Parameter 2	Unit 1	Unit 2	Coefficient[a]	r^2	SE_{res}[a]	n	p_{coef}
$C_{462\text{-}490}$	AOU	RU	μmol kg^{-1}	0.04 ± 0.00	0.87	0.87	178	<0.0001
C_{454}	AOU	RU	μmol kg^{-1}	0.12 ± 0.02	0.67	3.58	178	<0.0001
C_{406}	AOU	RU	μmol kg^{-1}	0.03 ± 0.00	0.66	0.97	178	<0.0001
C_{366}	AOU	RU	μmol kg^{-1}	0.01 ± 0.01	0.42	1.81	178	ns
$C_{462\text{-}490}$	$C_{366\text{-int}}$	RU	RU	0.38 ± 0.18	0.81	1.08	178	0.0345
C_{454}	$C_{366\text{-int}}$	RU	RU	2.34 ± 0.67	0.58	4	178	0.0006
C_{406}	$C_{366\text{-int}}$	RU	RU	1.12 ± 0.16	0.67	0.96	178	<0.0001
C_{366}	$C_{366\text{-int}}$	RU	RU	3.04 ± 0.19	0.78	1.13	178	<0.0001
$C_{462\text{-}490}$	$a_{CDOM}(254)_{int}$	RU	m^{-1}	0.02 ± 0.00	0.84	0.97	178	<0.0001
C_{454}	$a_{CDOM}(254)_{int}$	RU	m^{-1}	0.10 ± 0.01	0.71	3.32	178	<0.0001
C_{406}	$a_{CDOM}(254)_{int}$	RU	m^{-1}	0.03 ± 0.00	0.71	0.89	178	<0.0001
C_{366}	$a_{CDOM}(254)_{int}$	RU	m^{-1}	0.02 ± 0.01	0.45	1.76	178	0.0022
AOU	Prokaryotic abundance	μmol kg^{-1}	cell mL^{-1}	0.29 ± 0.04	0.88	16635.91	219	<0.0001
AOU	$C_{366\text{-int}}$	μmol kg^{-1}	RU	4.92 ± 2.39	0.89	16.13	221	0.0405
AOU	Chl-a_{int}	μmol kg^{-1}	mg Chl-a m^{-2}	0.10 ± 0.04	0.89	16.04	221	0.0115
AOU	$a_{CDOM}(254)_{int}$	μmol kg^{-1}	m^{-1}	0.30 ± 0.04	0.91	14.59	221	<0.0001

[a] $\times 10^{-3}$, if 'Parameter 1' is an FDOM component, or if 'Parameter 1' is AOU and 'Parameter 2' is prokaryotic abundance.

proteobacteria classes Gammaproteobacteria (9.5%; 5.4 – 16.3%) and Alphaproteobacteria (6.3%; 3.7 – 14%), and Chloroflexi (5.0%; 1.0 – 15.8%), respectively. The rest of the taxa made up 12.2% of reads on average, while the mean abundance of unidentified sequences was 7.9%.

The outlier samples had heterogenous community compositions. They presented unusually high abundances of taxa that otherwise were infrequent, especially members of *Actinobacteria* and *Firmicutes*. Nevertheless, Alpha- and Gammaproteobacteria, and Bacteroidota (in st. 9) were also present in high relative abundances (see Supporting results, and Fig. S6 and S7 for a more detailed description).

NMDS ordination of the samples (Fig. 4b) based on their prokaryotic community structure showed that this was clearly associated with the water mass composition, as samples with similar water mass contributions were overall more similar to each other than to those with different water mass proportions.

3.4. Changes in prokaryotic community diversity and composition across water masses

Archetype diversity indices were estimated for water masses that presented a sizeable contribution in the amplicon samples: EQ_{13}, $ENACW_{12}$, SPMW, $AAIW_5$, CDW, $NADW_{4.6}$ and $NADW_2$. In conjunction, these 7 water masses represented 99.87% of the water volume in the amplicon samples (min. $ENACW_{12}$ = 3.12%, max. $NADW_{4.6}$ = 30.14%; Fig. S8). Other water masses were left out as very low contributions would potentially bias archetype estimates. Species (here ASVs) richness (Fig. 5a) was lowest in SPMW (2920 ± 214 ASVs), and highest in bathypelagic water masses (3485–3554). The Pielou index (Fig. 5b), a measure of evenness, was lowest also in the SPMW (0.754 ± 0.026), followed by $AAIW_5$ (0.771 ± 0.016), and highest in $ENACW_{12}$ (0.807 ± 0.004). Similarly, the Shannon alpha diversity index (Fig. 5c) presented minimum values in the SPMW (6.019 ± 0.265) and highest values in $ENACW_{12}$ (6.459 ± 0.091) and $NADW_2$ (6.475 ± 0.116). The phylogenetic diversity to species richness ratio (Fig. 5d) gives an estimate of the phylogenetic relatedness of the ASVs that form the community: low ratios mean that the richness is due to phylogenetically close taxa (microdiversity), and vice versa. This ratio was lowest in EQ_{13} (0.0393 ± 0.0004), relatively stable in the mesopelagic (SPMW, 0.0404 ± 0.0011), and highest in the deep $NADW_2$ (0.0440 ± 0.0009). These results suggest that important differences in prokaryotic community diversity exist between water masses. SPMW, the water mass with the highest archetype AOU value (Table 1), tended to harbour a less diverse and less even prokaryotic community, as opposed to bathypelagic water masses. $ENACW_{12}$, while showing richness and phylogenetic diversity/richness estimates similar to those of SPMW, harboured a more even and diverse community.

An in-depth look into the taxonomic composition of the community in each of the water masses revealed further differences between them. The archetype CLR-transformed abundances of the dominant orders (Fig. 6) showed that Nitrosopumilales, within the Crenarchaeota phylum, displayed clearly superior abundances in the SPMW and were the dominant ASVs in this water mass (Fig. S9 and Supporting results). Nonetheless, there were within-order differences in the abundance patterns of the Nitrosopumilales. A close inspection of the archetype abundances of the dominant Nitrosopumilales ASVs showed two groups with contrasting distributions: one with peak abundances in the SPMW (including most of the ASVs with the highest abundances, thus, dominating the order-level trend) and a second group that presented maximum abundance values in bathypelagic waters (Fig. S10). Fitting mixing models to CLR-transformed abundances of dominant ASVs showed that the variability explained by the mixing models markedly differed within the Nitrosopumilales order (Fig. S11): overall, ASVs associated with the SPMW presented lower r^2 (<0.6), while those with highest archetype abundances in bathypelagic water masses were better explained by mixing ($r^2 > 0.7$). Mixing models account for the mixing between the different water masses, but also for the history of these water masses, i.e., all the remineralization processes that have occurred from the formation area of the water mass to the study region (Álvarez-Salgado et al., 2013). Considering the turnover rates of prokaryotes in the deep North Atlantic (24–55 days, Reinthaler et al, 2013) relative to the age of the studied water masses (few years to over a hundred years, Khatiwala et al. 2012; Catalá et al. 2015a), the ASVs that are highly explained by the mixing models would have experienced a selection over time in those water masses, leading them to sustain the abundances they display. Including biogeochemical variables in the regressions did not always yield significant regression coefficients, but multiple ASVs were positively associated with C_{454} and, to a lesser extent, AOU (Fig. S12).

The SAR324 clade, Marine Group II (maximum in the SPMW, but also in EQ_{13}), SAR11 clade (Alphaproteobacteria), Microtrichales (Acidobacteria), Thiomicrospirales and Steroidobacterales (Gammaproteobacteria) were more abundant in mesopelagic water masses, while the SAR202 clade (Chloroflexi), HOC36 (Gammaproteobacteria), Oceanospirillales (Gammaproteobacteria), Alteromonadales (Gammaproteobacteria), Sphingomonadales (Alphaproteobacteria) and Rhodobacterales (Alphaproteobacteria) presented higher archetype

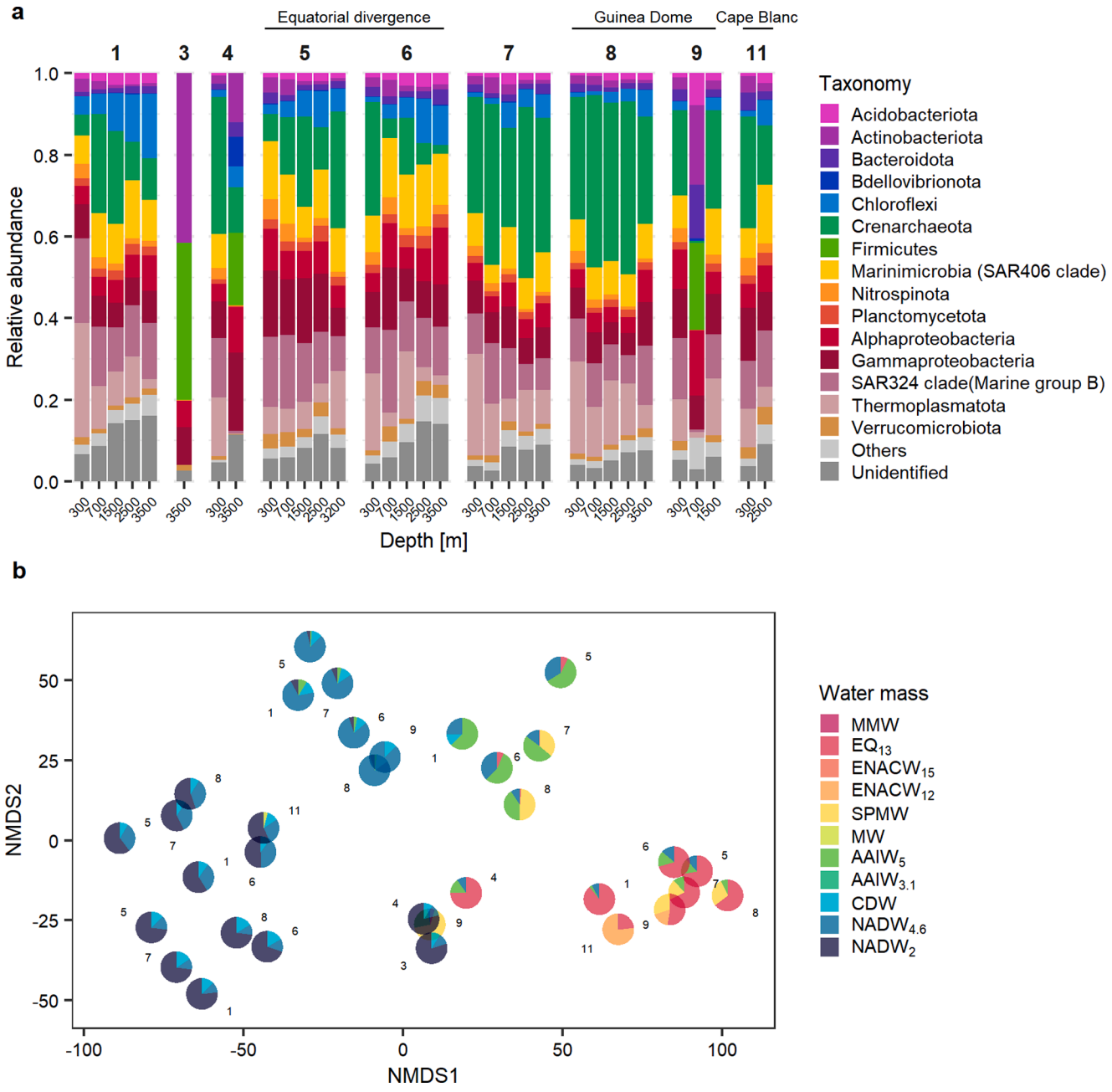


Fig. 4. Prokaryotic community composition. (a) Relative abundances of prokaryotic phyla (or class for Proteobacteria) along the cruise section. Numbers on top of bars represent the stations. (b) Distribution of prokaryotic communities in the NMDS space according to their similarities in taxonomic composition. Each pie chart corresponds to a sample, where colours represent the contribution of the different water masses. See the text for water mass abbreviations. Numbers beside pie charts correspond to stations.

abundances in bathypelagic water masses (Fig. 6, Fig. S10). Thiomicrospirales, Alteromonadales, Sphingomonadales and Steroidobacterales displayed the largest differences between water masses. On the contrary, Marinimicrobia, UBA10353 marine group (Gammaproteobacteria) and the Arctic97B-4 marine group (Verrucomicrobiota) were equally abundant in the different water masses. Marinimicrobia presented differences at the ASV level (Fig. S10), with one group of ASVs enriched in mesopelagic waters and another group in bathypelagic water masses. Mixing models showed greatest explanatory capacity for Marinimicrobia, Marine Group II, SAR202 and HOC36 with r^2 exceeding 0.8–0.9, while Alteromonadales only reached r^2 of 0.35–0.55 (Fig. S11). Among biogeochemical variables, humic-like fluorescence components, AOU and $a_{CDOM}(254)_{int}$ helped explain the distributions of these prokaryotic taxa. SAR202 bacteria showed significant negative coefficients for C_{454}, AOU and $a_{CDOM}(254)_{int}$ (Fig. S12), in agreement with their greater abundances in bathypelagic water masses and, to a lesser extent, in $AAIW_5$ (Fig. 6). Alteromonadales ASVs also displayed negative coefficients for C_{454} and, to a lesser extent, $C_{462\text{-}490}$ and C_{406}. On the contrary, a considerable part of the Marine Group II and HOC36 ASVs presented positive significant coefficients for C_{454} and $a_{CDOM}(254)_{int}$ (Fig. S12).

4. Discussion

4.1. FDOM characteristics and distribution in water masses

Detailed descriptions of FDOM distributions in the deep tropical Atlantic are scarce. Catalá et al. (2015b) reported values for components

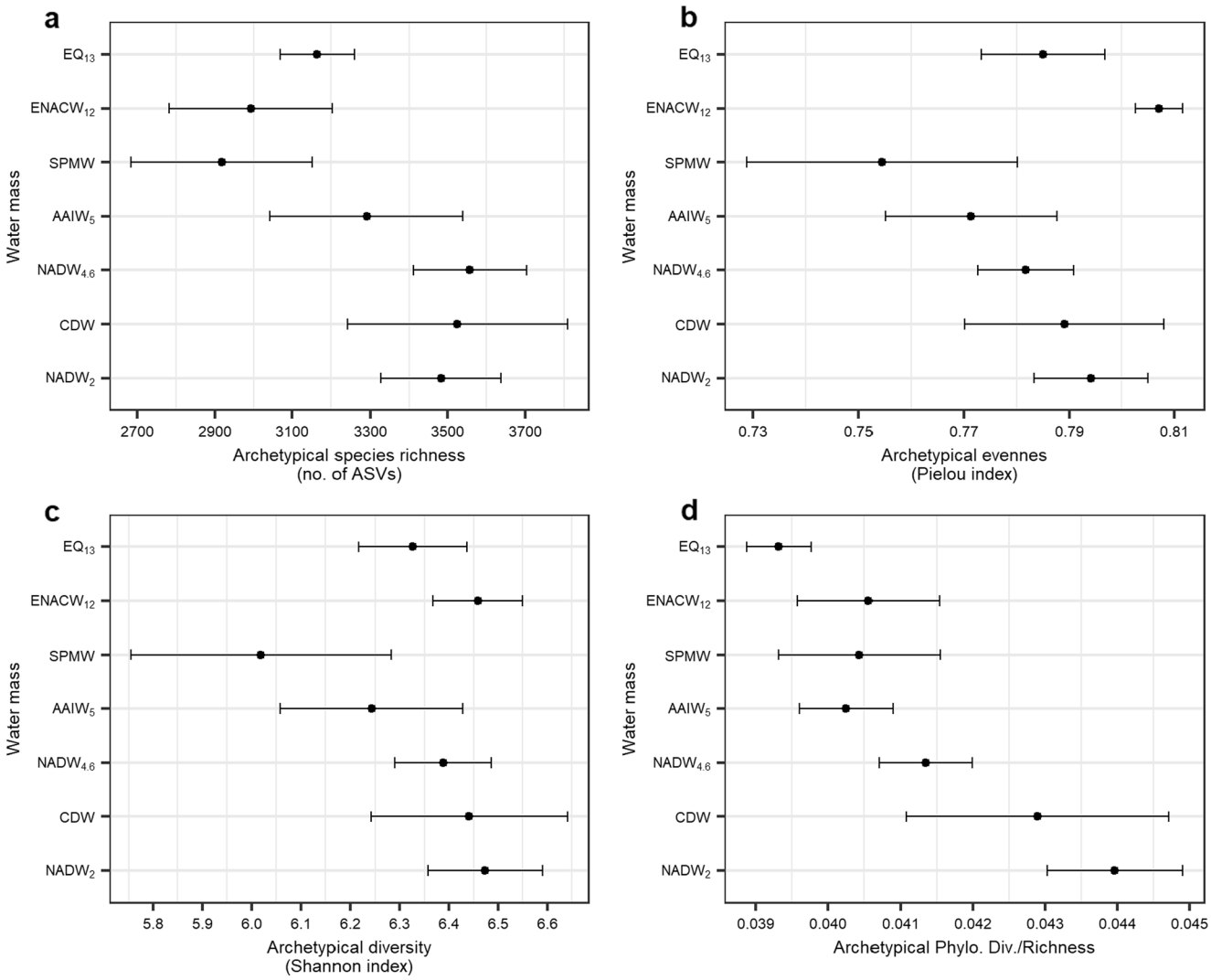


Fig. 5. Prokaryotic community diversity indices. Archetype values of (a) species richness (as no. of amplicon sequence variants, ASVs), (b) Pielou evenness index, (c) Shannon index and (d) Faith's Phylogenetic Diversity index / ASV richness ratio. Error bars represent the standard error of the estimates.

similar to $C_{462\text{-}490}$ and C_{406} within the Equatorial Atlantic Central Water and North Atlantic Deep Water that agree with our findings both qualitatively, with an overall increase in fluorescence from intermediate to deep waters (Fig. 3 and Table 1), and quantitatively, as their fluorescence values fall within the range of those found here. The humic-like components generally displayed higher values below the epipelagic layer, probably due to photobleaching in surface waters (Catalá et al., 2016), while also showing a northward increase in fluorescence in the entire water column, specially marked for component C_{454} (Fig. 3). This change in fluorescence was coupled with an increase in AOU, but not entirely: $NADW_{4.6}$ and $NADW_2$ presented high $C_{462\text{-}490}$ and C_{454} fluorescence relative to their AOU (Table 1, Fig. S13), suggesting a contribution by terrestrial FDOM introduced in the source region of NADW (Benner et al., 2005; Catalá et al., 2015b; Jørgensen et al., 2011). High values of C_{406} were measured within the Oxygen Minimum Zone (OMZ) of the eastern tropical Atlantic (10-18°N, Fig. 2b and S5). This oceanographic feature is generated by the combination of remineralisation by prokaryotes of abundant sinking organic matter from the highly productive surface waters between the Guinea Dome and Cape Blanc (influenced by the Northwest African upwelling zone, Fig. 1 and Fig. S14; Carr and Kearns, 2003) and reduced ventilation by currents (Stramma et al., 2008). Minimum O_2 concentrations in the OMZ were ~50 $\mu mol \cdot Kg^{-1}$, which agree with values reported in the literature that ascribe the eastern tropical Atlantic OMZ as an hypoxic zone (Karstensen et al., 2008; Stramma et al., 2008). The fluorescence-AOU coupling (Fig. S13) suggests that the increase in fluorescence was related to microbial transformation of organic matter, producing humic-like DOM, which has been documented in the literature (Jørgensen et al., 2014). The direct C_{406}-AOU relationship displayed variability (Fig. S13), suggesting that multiple factors, including potential consumption by prokaryotes, might influence its distribution. The transformation of vertically transported organic matter might have enhanced the humic-like FDOM signal, as stations 8, 9, 11 and12 showed high levels of the protein-like C_{366} throughout the water column. Moreover, stations 8–11 also presented widespread, high humic-like fluorescence in the bathypelagic layer which, in addition to the NADW influence, could reflect a considerable remineralisation below 1000 m as a consequence of high export rates in these productive stations. OMZs are known to reduce the remineralisation rates of sinking particles, allowing them to reach greater depths than they otherwise would in environments with greater O_2 concentrations (Rasse and Dall'Olmo, 2019).

As expected from their spatial distributions, humic-like FDOM components ($C_{462\text{-}490}$, C_{454}, C_{406}) were found to be significantly related to AOU (Table 3, Fig. S13), agreeing with previous findings that described fluorophores similar to $C_{462\text{-}490}$ and C_{406} displaying positive correlations with AOU in the deep ocean (Catalá et al., 2015b; De La Fuente et al., 2014). The higher explanatory power of $C_{462\text{-}490}$ by water mass mixing, and slightly lower by AOU relative to C_{454} and C_{406}

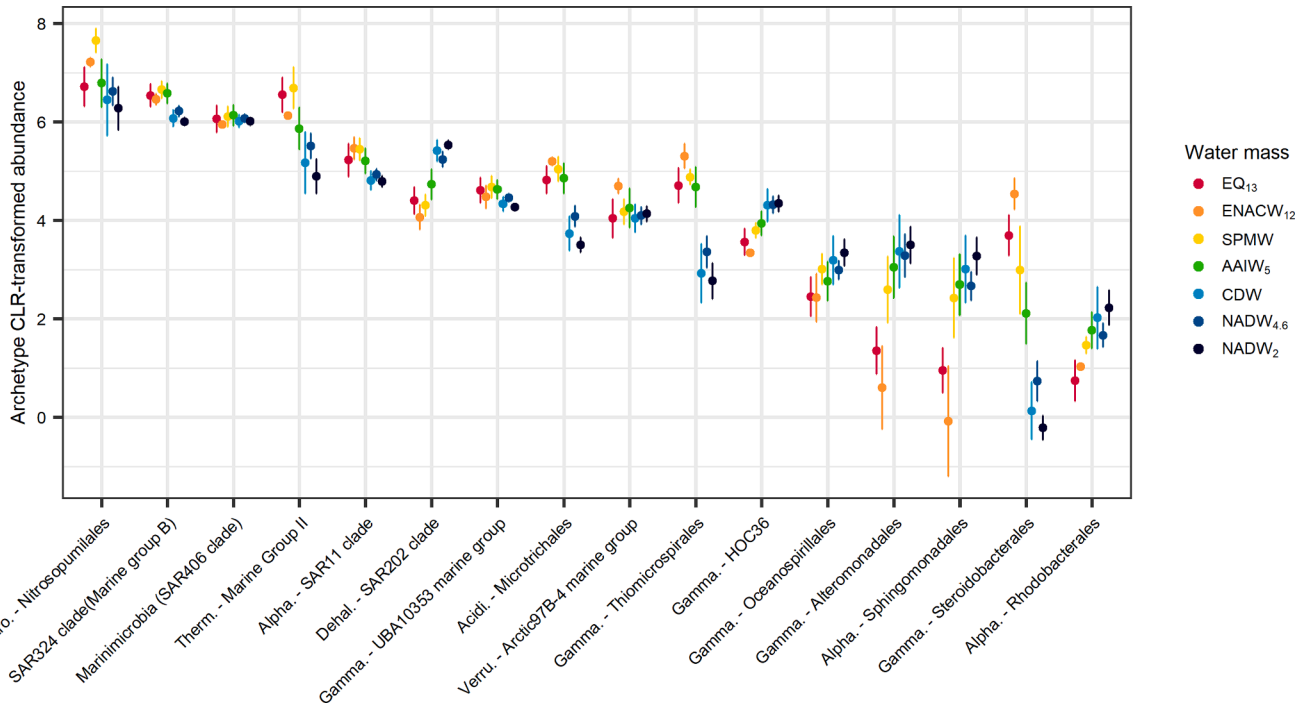


Fig. 6. Taxonomic characterisation of the water masses. CLR-transformed archetype abundances in water masses of orders (phyla, for SAR324 clade and Marinimicrobia) that contain dominant amplicon sequence variants (>0.5% in at least one sample). Bars represent standard errors of archetype abundances. When available, the abbreviated class designation is shown to the left of order label.

(Tables 2 and 3), suggests that the characteristics of compounds comprising this fluorophore hindered biologically mediated degradation processes to a greater degree. Contrary to humic-like components, the protein-like C_{366} was the least explained by the mixing model and showed no improvement when including AOU, suggesting totally different dynamics. The explanatory capacity of $C_{366\text{-int}}$ (C_{366} values integrated over the epipelagic layer) was significant for all FDOM components but highest for C_{366}, further supporting the view that deep ocean protein-like DOM is linked to the downward flux of protein-like material through sinking particles (Ruiz-González et al., 2020). Another factor that could potentially contribute to the protein-like FDOM signal in the dark ocean (especially in the mesopelagic layer) is the excretion of protein-like DOM by zooplankton and micronecton (Morán et al., 2022; Urban-Rich et al., 2006). High biomasses of these organisms were found in stations 8, 9 and 11 (Hernández-León et al., 2019), where the C_{366} signal was strong (Fig. 3d).

Together, these observations would suggest that the increased humic-like fluorescence is partially a consequence of the vertical export of organic matter and subsequent remineralization, which is supported by the fact that several proxies of epipelagic productivity (here as $C_{366\text{-int}}$, $a_{CDOM}(254)_{int}$ proxies) significantly contributed to the explained variance of FDOM components, AOU and prokaryotic abundance (Table 3). This links the surface conditions to the microbial reworking of organic matter in the water masses of the dark tropical and subtropical Atlantic. These results support the view that, while water mass mixing and history play an important role, there is a strong vertical connection between the processes occurring in the water column, as it has been recently suggested (Mestre et al., 2018; Ruiz-González et al., 2020).

4.2. *Water mass-specific differences in prokaryotic community composition*

While the overall composition of the prokaryotic communities followed water mass composition (Fig. 4), this pattern might have been partly influenced by the vertical distribution of communities (DeLong et al., 2006), since bathypelagic water masses occupied much of the section in their depth ranges (Fig. 2a). Nonetheless, differences were identified between water masses both in terms of overall diversity (Fig. 5) and specific taxonomic composition of the prokaryotic communities (Fig. 6 and S10), in agreement with previous studies showing that water masses harbour distinct prokaryotic communities (Agogué et al., 2011; Salazar et al., 2016). Among water masses, the SPMW, directly influenced by the OMZ of the Eastern North Atlantic, presented the less diverse and even community (dominated by several abundant taxa). Decreases in diversity of the prokaryotic community have been observed in OMZs elsewhere (Beman et al., 2020; Bertagnolli and Stewart, 2018) and point to a narrower range of prokaryotes able to thrive in low oxygen conditions, although contrasting results have been reported (Stevens and Ulloa, 2008). The community in deep water masses was in general more diverse and had higher evenness (Fig. 5b), which agrees with previous reports from the deep North Atlantic (Agogué et al., 2011; Frank et al., 2016). This was paired with high phylogenetic diversity/ASV richness ratios (Fig. 5d), indicating that the diversity was due to phylogenetically distant ASVs, which suggests that the prokaryotic communities present in bathypelagic water masses could potentially display a wider range of metabolic capabilities.

Specific orders of prokaryotes showed different distributions across waters masses (Fig. 6), with three main patterns emerging: orders that were 1) more abundant in one or multiple central/intermediate water masses (predominantly, Nitrosopumilales and Marine Group II), 2) more abundant in deep water masses (SAR202 clade, Alteromonadales), and 3) equally abundant across water masses (Marinimicrobia). This association of prokaryotic taxa with specific water masses has been previously reported (Agogué et al., 2011). Nonetheless, within-order variability was observed in multiple instances (Fig. S10) as ASVs belonging to the same order showed differing patterns, suggesting distinct associations with environmental conditions even for closely related prokaryotes.

Nitrosopumilales presented a notable enrichment in the SPMW (Fig. 6 and S10). The fact that ASVs belonging to this order thrived in waters associated to high humic-like FDOM and AOU values suggests that an effective metabolism under suboxic conditions allows this group

to prevail in environments where predominantly recalcitrant DOM and low O_2 concentrations hinder activity of heterotrophic organisms. Nitrosopumilales have an efficient autotrophic metabolism using energy derived from the oxidation of very low concentrations of ammonia via a modified version of the hydroxypropionate/hydroxybutyrate (HP/HB) cycle (Könneke et al., 2014), and are also known to utilise cyanate and urea as sources of energy and nitrogen (Kitzinger et al., 2019). It has been recently shown that they can produce small amounts of O_2 that would aid their ammonia-oxidising metabolism (and potentially other microbes) in low oxygen waters (Kraft et al., 2022). The dominance of this archaeal order in the SPMW agrees with previous reports of organisms closely related to Nitrosopumilales dominating nitrifying processes in suboxic waters (Labrenz et al., 2010; Stewart et al., 2012). Nonetheless, the presence of another group of Nitrosopumilales ASVs with abundance peaks in the bathypelagic (Fig. S10) suggests a functional partitioning within this order adapted to specific conditions within the water column.

ASVs classified as Marine Group II (Thermoplasmatota) were among the most abundant orders in the different water masses and displayed co-occurrence patterns with Nitrosopumilales (Fig. S10). Members of Marine Group II are deemed to play an important role in the cycling of organic matter, as it has been observed that they possess genes for the degradation of high molecular weight compounds (such as proteins, polysaccharides and lipids) and the intermembrane transport of organic molecules (Tully, 2019). Marine Group II ASVs were most abundant in the SPMW (although EQ_{13} presented similar archetype abundance, Fig. 6), where there was a strong signal of FDOM and AOU (Fig. 2b and 3, Table 1) and higher productivity in surface waters (Fig. S14). This is in agreement with the observation in the Iberian upwelling zone of Marine Group II archaea in high abundances in low oxygen conditions, in correlation with humic-like FDOM (Guerrero-Feijóo et al., 2017). Moreover, this order has been reported as an important contributor to both the uptake of glucose and the degradation of complex organic matter in the bathypelagic Mediterranean (Boutrif et al., 2011). Thus, in contrast to Nitrosopumilales, Marine Group II archaea inhabiting the dark ocean seem to have a prevalent heterotrophic life-style (Deschamps et al., 2014). The presence of organisms of this order in low-oxygen conditions has been previously observed (Belmar et al., 2011) and might be favoured by their ability to conduct nitrate respiration (Rinke et al., 2019).

An important number of ASVs classified as members of the SAR324 clade were associated with mesopelagic water masses while others seemed to favour bathypelagic environments (Fig. S10). The widespread presence and high abundances of ASVs belonging to this phylum indicates that they are well adapted to the varying conditions in the dark ocean, including suboxic waters (Pajares et al., 2020), as suggested by their autotrophic and extremely versatile metabolism (Sheik et al., 2014; Swan et al., 2011).

SAR202 showed higher archetype abundances in bathypelagic water masses, increasing from 300 m to 3500 m (Fig. 6 and S10) agreeing with previous reports from the Atlantic (Varela et al., 2008). Members of SAR202 are considered to be heterotrophic, free-living bacterioplankton (Mehrshad et al., 2018), and genomic studies have suggested their ability to utilise recalcitrant DOM (Landry et al., 2017; Saw et al., 2020). For instance, Landry et al. (2017) highlighted the number of genes predicted to encode a variety of monooxygenases and dehydrogenases, which they deemed likely to participate in the activation of alicyclic and long-chain molecules for their degradation. Thus, the ability of SAR202 to utilise diverse recalcitrant compounds would potentially help to explain their high contribution to communities in deep water masses.

Differences were observed in the explanatory capacity of the mixing models on the abundances of dominant taxa (Fig. S11). For instance, ASVs belonging to Marinimicrobia, which were abundant across water masses (Fig. 6 and S10) and known to inhabit a variety of oceanic environments (Hawley et al., 2017), were highly explained by the mixing models, as was the case for SAR202 ASVs. Thus, their distribution seemed to be largely determined by a combination of water mass mixing and the selection they had experienced in the different water masses since their formation. On the contrary, some orders were weakly explained by the mixing models (Fig. S11), suggesting that local biogeochemical processes, such as the vertical connection associated with surface productivity (Frank et al., 2016; Mestre et al., 2018) or other undetermined processes, were more important for their distribution. Alteromonadales was an example of such taxa (Fig. S11), as they are known copiotrophs and opportunists that are usually found associated with particles, quickly reacting to organic matter inputs in 'boom-and-bust' dynamics (Mestre et al., 2018; Reintjes et al., 2020). Nonetheless, while negative relations were observed with humic-like FDOM (Fig. S12), no significant relationship was found between Alteromonadales and protein-like FDOM (Fig. S12). Thus, a definitive explanation about the factors involved in controlling their distribution in accordance with their known copiotrophy cannot be inferred from our results.

5. Conclusions

In this study we jointly characterised the FDOM pool and the prokaryotic community compositions across water masses of the tropical and subtropical Atlantic Ocean. We found that the most humic, aromatic and complex fraction of FDOM ($C_{462\text{-}490}$) was more influenced by mixing and the history of the water masses, while the protein-like component (C_{366}) was the most subjected to intra-regional, non-conservative biogeochemical processes (likely due to vertical transport of sinking organic matter related to surface productivity gradients). Other humic components (C_{454}, C_{406}) laid in between those extremes. AOU was observed to significantly explain the variance of the humic components, linking them to microbial activity. The composition of the prokaryotic community was found to be associated with these processes: diversity decreased with increasing values of AOU, with minimum values in the SPMW, the water mass most influenced by the Eastern North Atlantic OMZ. Likewise, specific taxa displayed differing abundances between water masses: Nitrosopumilales were more abundant in the SPMW, dominating low-oxygen waters, and Marine Group II archaea, SAR324 clade and Thiomicrospirales tended to co-occur with them in this water mass. On the contrary, members of the SAR202 clade were more abundant in bathypelagic water masses, while Marinimicrobia were distributed throughout the water column. Mixing models explained in varying degrees the abundance patterns of different taxa, suggesting that the distribution of some of them (e.g., Marinimicrobia, SAR202) was dominated by mixing and selection within the water masses since their formation, while others (e.g., Alteromonadales) were not. Thus, while water masses showed differences in their DOM characteristics and prokaryotic communities depending on their initial properties, ageing history and mixing, a fraction of the DOM signal and prokaryotic community was deemed to be controlled (in varying degrees) by local processes (such as vertical inputs of organic matter, as suggested by the relationships displayed with the epipelagic productivity proxies). Our observations are relevant to the understanding of how microbial-DOM interactions, including long-term changes within water masses and the vertical connectivity along surface productivity gradients, affect the compositions of both the DOM pool and the prokaryotic community in the dark ocean. Distinct prokaryotic taxonomic groups would display different metabolisms favouring the consumption and transformation of certain types of DOM, affecting the long-term (i.e., centennial scales) storage of C in the deep ocean. Future studies including the characterization of the organic matter with non-targeted ultra-high resolution analytical techniques such as FT-ICR-MS are necessary to provide detailed insights into the role of distinct prokaryotic communities in the degradation and/or generation of specific compound classes.

Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank to the officers and crew of the *BIO Hespérides*, and the staff of the Unit of Marine Technology (UTM) of the Spanish Research Council (CSIC) for their invaluable help at sea.

This work is a contribution to projects MAFIA (CTM2012-39587-C04-01), FLUXES (CTM2015-69392-C3), e-IMPACT (PID2019-109084RB-C21), and INTERES (CTM2017-83362-R) funded by the Spanish "Plan Nacional/Estatal de I + D" and cofounded with FEDER funds, and to project SUMMER (AMD-817806-5) funded by the European Union's Horizon 2020 research and innovation program. MGL is supported by Ministerio de Ciencia, Innovación y Universidades, Gobierno de España (FPU17-01435) during his PhD. MS is supported by the Project MIAU (RTI2018-101025-B-I00) and the 'Severo Ochoa Centre of Excellence' accreditation (CEX2019-000928-S).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pocean.2022.102819.

References

Agogué, H., Lamy, D., Neal, P.R., Sogin, M.L., Herndl, G.J., 2011. Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. Mol. Ecol. 20, 258–274. https://doi.org/10.1111/j.1365-294X.2010.04932.x.

Álvarez-Salgado, X.A., Nieto-Cid, M., Álvarez, M., Pérez, F.F., Morin, P., Mercier, H., 2013. New insights on the mineralization of dissolved organic matter in central, intermediate, and deep water masses of the northeast North Atlantic. Limnol. Oceanogr. 58 (2), 681–696. https://doi.org/10.4319/lo.2013.58.2.0681.

Álvarez, M., Brea, S., Mercier, H., Álvarez-Salgado, X.A., 2014. Mineralization of biogenic materials in the water masses of the South Atlantic Ocean. I: Assessment and results of an optimum multiparameter analysis. Prog. Oceanogr. 123, 1–23. https://doi.org/10.1016/j.pocean.2013.12.007.

Aparicio, F.L., Nieto-Cid, M., Borrull, E., Romero, E., Stedmon, C.A., Sala, M.M., Gasol, J. M., Ríos, A.F., Marrasé, C., 2015. Microbially-mediated fluorescent organic matter transformations in the deep ocean. Do the chemical precursors matter? Front. Mar. Sci. 2 (106) https://doi.org/10.3389/fmars.2015.00106.

Arnosti, C., 2011. Microbial Extracellular Enzymes and the Marine Carbon Cycle. Ann. Rev. Mar. Sci. 3 (1), 401–425. https://doi.org/10.1146/annurev-marine-120709-142731.

Baetge, N., Graff, J.R., Behrenfeld, M.J., Carlson, C.A., 2020. Net Community Production, Dissolved Organic Carbon Accumulation, and Vertical Export in the Western North Atlantic . Front. Mar. Sci. .

Baltar, F., Alvarez-Salgado, X.A., Arístegui, J., Benner, R., Hansell, D.A., Herndl, G.J., Lønborg, C., 2021. What Is Refractory Organic Matter in the Ocean? . Front. Mar. Sci. .

Belmar, L., Molina, V., Ulloa, O., 2011. Abundance and phylogenetic identity of archaeoplankton in the permanent oxygen minimum zone of the eastern tropical South Pacific. FEMS Microbiol. Ecol. 78, 314–326. https://doi.org/10.1111/j.1574-6941.2011.01159.x.

Beman, J.M., Carolan, M.T., 2013. Deoxygenation alters bacterial diversity and community composition in the ocean's largest oxygen minimum zone. Nat. Commun. 4, 2705. https://doi.org/10.1038/ncomms3705.

Beman, J.M., Vargas, S.M., Vazquez, S., Wilson, J.M., Yu, A., Cairo, A., Perez-Coronel, E., 2020. Biogeochemistry and hydrography shape microbial community assembly and activity in the eastern tropical North Pacific Ocean oxygen minimum zone. Environ. Microbiol. 23 (6), 2765–2781. https://doi.org/10.1111/1462-2920.15215.

Benner, R., Louchouarn, P., Amon, R.M.W., 2005. Terrigenous dissolved organic matter in the Arctic Ocean and its transport to surface and deep waters of the North Atlantic. Global Biogeochem. Cycles 19 (2), n/a–n/a.

Benson, B.B., Krause, D., 1984. The concentration and isotopic fractionation of oxygen dissolved in freshwater and seawater in equilibrium with the atmosphere. Limnol. Oceanogr. 29, 620–632. https://doi.org/10.4319/lo.1984.29.3.0620.

Bergauer, K., Fernandez-Guerra, A., Garcia, J.A.L., Sprenger, R.R., Stepanauskas, R., Pachiadaki, M.G., Jensen, O.N., Herndl, G.J., 2018. Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. Proc. Natl. Acad. Sci. USA 115, E400–E408. https://doi.org/10.1073/pnas.1708779115.

Bertagnolli, A.D., Stewart, F.J., 2018. Microbial niches in marine oxygen minimum zones. Nat. Rev. Microbiol. 16 (12), 723–729. https://doi.org/10.1038/s41579-018-0087-z.

Boutrif, M., Garel, M., Cottrell, M.T., Tamburini, C., 2011. Assimilation of marine extracellular polymeric substances by deep-sea prokaryotes in the NW Mediterranean Sea. Environ. Microbiol. Rep. 3, 705–709. https://doi.org/10.1111/j.1758-2229.2011.00285.x.

Brandt, P., Hormann, V., Bourlès, B., Fischer, J., Schott, F.A., Stramma, L., Dengler, M., 2008. Oxygen tongues and zonal currents in the equatorial Atlantic. J. Geophys. Res. Ocean. 113 (C4) https://doi.org/10.1029/2007JC004435.

Broecker, W.S., 1974. "NO", a conservative water-mass tracer. Earth Planet. Sci. Lett. 23 (1), 100–107. https://doi.org/10.1016/0012-821X(74)90036-3.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13 (7), 581–583. https://doi.org/10.1038/nmeth.3869.

Carlson, C.A., Hansell, D.A., 2015. DOM Sources, Sinks, Reactivity, and Budgets, in: Hansell, D.A., Carlson, C.A. (Eds.), Biogeochemistry of Marine Dissolved Organic Matter: Second Edition. Academic Press, Boston, pp. 65–126. https://doi.org/10.1016/B978-0-12-405940-5.00003-0.

Carr, M.-E., Kearns, E.J., 2003. Production regimes in four eastern boundary current systems. Deep-Sea Res Part II Top. Stud. Oceanogr. 50 (22-26), 3199–3221. https://doi.org/10.1016/j.dsr2.2003.07.015.

Catalá, T.S., Álvarez-Salgado, X.A., Otero, J., Iuculano, F., Companys, B., Horstkotte, B., Romera-Castillo, C., Nieto-Cid, M., Latasa, M., Morán, X.A.G., Gasol, J.M., Marrasé, C., Stedmon, C.A., Reche, I., 2016. Drivers of fluorescent dissolved organic matter in the global epipelagic ocean. Limnol. Oceanogr. 61 (3), 1101–1119. https://doi.org/10.1002/lno.10281.

Catalá, T.S., Reche, I., Álvarez, M., Khatiwala, S., Guallart, E.F., Benítez-Barrios, V.M., Fuentes-Lema, A., Romera-Castillo, C., Nieto-Cid, M., Pelejero, C., Fraile-Nuez, E., Ortega-Retuerta, E., Marrasé, C., Álvarez-Salgado, X.A., 2015a. Water mass age and aging driving chromophoric dissolved organic matter in the dark global ocean. Global Biogeochem. Cycles 29 (7), 917–934. https://doi.org/10.1002/2014GB005048.

Catalá, T.S., Reche, I., Fuentes-Lema, A., Romera-Castillo, C., Nieto-Cid, M., Ortega-Retuerta, E., Calvo, E., Álvarez, M., Marrasé, C., Stedmon, C.A., Álvarez-Salgado, X. A., 2015b. Turnover time of fluorescent dissolved organic matter in the dark global ocean. Nat. Commun. 6, 5986. https://doi.org/10.1038/ncomms6986.

Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A., Callahan, B.J., 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6 (1). https://doi.org/10.1186/s40168-018-0605-2.

De La Fuente, P., Marrasé, C., Canepa, A., Antón Álvarez-Salgado, X., Gasser, M., Fajar, N.M., Romera-Castillo, C., Pelegrí, J.L., 2014. Does a general relationship exist between fluorescent dissolved organic matter and microbial respiration?-The case of the dark equatorial Atlantic Ocean. Deep-Sea Res. Part I Oceanogr. Res. Pap. 89, 44–55. https://doi.org/10.1016/j.dsr.2014.03.007.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.-U., Martinez, A., Sullivan, M.B., Edwards, R., Brito, B.R., Chisholm, S.W., Karl, D.M., 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science (80-. 311 (5760), 496–503.

Deschamps, P., Zivanovic, Y., Moreira, D., Rodriguez-Valera, F., Lopez-García, P., 2014. Pangenome evidence for extensive interdomain horizontal transfer affecting lineage coreandshell genes inuncultured planktonic thaumarchaeota and euryarchaeota. Genome Biol. Evol. 6, 1549–1563. https://doi.org/10.1093/gbe/evu127.

Dittmar, T., Lennartz, S.T., Buck-Wiese, H., Hansell, D.A., Santinelli, C., Vanni, C., Blasius, B., Hehemann, J.-H., 2021. Enigmatic persistence of dissolved organic matter in the ocean. Nat. Rev. Earth Environ. 2 (8), 570–583. https://doi.org/10.1038/s43017-021-00183-7.

Falkowski, P.G., Fenchel, T., Delong, E.F., 2008. The microbial engines that drive earth's biogeochemical cycles. Science (80-. 320 (5879), 1034–1039.

Fischer, G., Neuer, S., Ramondenc, S., Müller, T.J., Donner, B., Ruhland, G., Ratmeyer, V., Meinecke, G., Nowald, N., Klann, M., Wefer, G., 2020. Long-Term Changes of Particle Flux in the Canary Basin Between 1991 and 2009 and Comparison to Sediment Trap Records Off Mauritania . Front. Earth Sci.

Frank, A.H., Garcia, J.A.L., Herndl, G.J., Reinthaler, T., 2016. Connectivity between surface and deep waters determines prokaryotic diversity in the North Atlantic Deep Water. Environ. Microbiol. 18, 2052–2063. https://doi.org/10.1111/1462-2920.13237.

Gómez-Consarnau, L., Lindh, M.V., Gasol, J.M., Pinhassi, J., 2012. Structuring of bacterioplankton communities by specific dissolved organic carbon compounds. Environ. Microbiol. 14, 2361–2378. https://doi.org/10.1111/j.1462-2920.2012.02804.x.

Gómez-Letona, M., Arístegui, J., Hernández-Hernández, N., et al., 2022a. Biogeochemical variables including mixed layer samples from the tropical and subtropical Atlantic of the MAFIA cruise in 2015. PANGAEA. https://doi.org/10.1594/PANGAEA.942940.

Gómez-Letona, M., Arístegui, J., Hernández-Hernández, N., Álvarez, M., Álvarez-Salgado, X.A., Delgadillo, E., Pérez-Lorenzo, M., Teira, E., Hernández-León, S., Sebastián, M., 2022b. Fluorescent dissolved organic matter and water mass distribution in the tropical and subtropical Atlantic in 2015. PANGAEA. https://doi.org/10.1594/PANGAEA.942941.

Gómez-Letona, M., Arístegui, J., Hernández-Hernández, N., et al., 2022c. Biogeochemical variables and optimum multiparameter analysis results from the tropical and subtropical Atlantic of the MAFIA cruise in 2015. PANGAEA. https://doi.org/10.1594/PANGAEA.942934.

Guerrero-Feijóo, E., Nieto-Cid, M., Sintes, E., Dobal-Amador, V., Hernando-Morales, V., Álvarez, M., Balagué, V., Varela, M.M., King, G., 2017. Optical properties of dissolved organic matter relate to different depth-specific patterns of archaeal and bacterial community structure in the North Atlantic Ocean. FEMS Mcrobiol. Ecol. 93 (1), fiw224. https://doi.org/10.1093/femsec/fiw224.

Hansell, D.A., 2013. Recalcitrant dissolved organic carbon fractions. Ann. Rev. Mar. Sci. 5 (1), 421–445. https://doi.org/10.1146/annurev-marine-120710-100757.
Hansell, D., Carlson, C., Repeta, D., Schlitzer, R., 2009. Dissolved organic matter in the ocean: a controversy stimulates new insights. Oceanography 22 (4), 202–211.
Hawley, A.K., Nobu, M.K., Wright, J.J., Durno, W.E., Morgan-Lang, C., Sage, B., Schwientek, P., Swan, B.K., Rinke, C., Torres-Beltrán, M., Mewis, K., Liu, W.-T., Stepanauskas, R., Woyke, T., Hallam, S.J., 2017. Diverse Marinimicrobia bacteria may mediate coupled biogeochemical cycles along eco-thermodynamic gradients. Nat. Commun. 8, 1507. https://doi.org/10.1038/s41467-017-01376-9.
Hernández-León, S., Olivar, M.P., Fernández de Puelles, M.L., Bode, A., Castellón, A., López-Pérez, C., Tuset, V.M., González-Gordillo, J.I., 2019. Zooplankton and Micronekton Active Flux Across the Tropical and Subtropical Atlantic Ocean. Front. Mar. Sci.
Jørgensen, L., Stedmon, C.A., Granskog, M.A., Middelboe, M., 2014. Tracing the long-term microbial production of recalcitrant fluorescent dissolved organic matter in seawater. Geophys. Res. Lett. 41 (7), 2481–2488. https://doi.org/10.1002/2014GL059428.
Jørgensen, L., Stedmon, C.A., Kragh, T., Markager, S., Middelboe, M., Søndergaard, M., 2011. Global trends in the fluorescence characteristics and distribution of marine dissolved organic matter. Mar. Chem. 126 (1-4), 139–148. https://doi.org/10.1016/j.marchem.2011.05.002.
Karstensen, J., Stramma, L., Visbeck, M., 2008. Oxygen minimum zones in the eastern tropical Atlantic and Pacific oceans. Prog. Oceanogr. 77 (4), 331–350. https://doi.org/10.1016/j.pocean.2007.05.009.
Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., Webb, C.O., 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26 (11), 1463–1464. https://doi.org/10.1093/bioinformatics/btq166.
Khatiwala, S., Primeau, F., Holzer, M., 2012. Ventilation of the deep ocean constrained with tracer observations and implications for radiocarbon estimates of ideal mean age. Earth Planet. Sci. Lett. 325–326, 116–125. https://doi.org/10.1016/j.epsl.2012.01.038.
Kida, M., Kojima, T., Tanabe, Y., Hayashi, K., Kudoh, S., Maie, N., Fujitake, N., 2019. Origin, distributions, and environmental significance of ubiquitous humic-like fluorophores in Antarctic lakes and streams. Water Res. 163, 114901. https://doi.org/10.1016/j.watres.2019.114901.
Kitzinger, K., Padilla, C.C., Marchant, H.K., Hach, P.F., Herbold, C.W., Kidane, A.T., Könneke, M., Littmann, S., Mooshammer, M., Niggemann, J., Petrov, S., Richter, A., Stewart, F.J., Wagner, M., Kuypers, M.M.M., Bristow, L.A., 2019. Cyanate and urea are substrates for nitrification by Thaumarchaeota in the marine environment. Nat. Microbiol. 4 (2), 234–243. https://doi.org/10.1038/s41564-018-0316-2.
Könneke, M., Schubert, D.M., Brown, P.C., Hügler, M., Standfest, S., Schwander, T., Schada von Borzyskowski, L., Erb, T.J., Stahl, D.A., Berg, I.A., 2014. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. Proc. Natl. Acad. Sci. USA 111 (22), 8239–8244. https://doi.org/10.1073/pnas.1402028111.
Kraft, B., Jehmlich, N., Larsen, M., Bristow, L.A., Könneke, M., Thamdrup, B.o., Canfield, D.E., 2022. Oxygen and nitrogen production by an ammonia-oxidizing archaeon. Science (80-. 375 (6576), 97–100.
Kwon, E.Y., Primeau, F., Sarmiento, J.L., 2009. The impact of remineralization depth on the air-sea carbon balance. Nat. Geosci. 2 (9), 630–635. https://doi.org/10.1038/ngeo612.
Labrenz, M., Sintes, E., Toetzke, F., Zumsteg, A., Herndl, G.J., Seidler, M., Jürgens, K., 2010. Relevance of a crenarchaeotal subcluster related to Candidatus Nitrosopumilus maritimus to ammonia oxidation in the suboxic zone of the central Baltic Sea. ISME J. 4 (12), 1496–1508. https://doi.org/10.1038/ismej.2010.78.
Lahti, L., Shetty, S., 2019. microbiome R package, version 1.8.0 http://microbiome.github.io.
Landry, Z., Swan, B.K., Herndl, G.J., Stepanauskas, R., Giovannoni, S.J., Zhou, J., Aluwihare, L., Santoro, A., Northen, T., 2017. SAR202 genomes from the dark ocean predict pathways for the oxidation of recalcitrant dissolved organic matter. MBio 8 (2). https://doi.org/10.1128/mBio.00413-17.
Lapierre, J.F., Del Giorgio, P.A., 2014. Partial coupling and differential regulation of biologically and photochemically labile dissolved organic carbon across boreal aquatic networks. Biogeosciences 11, 5969–5985. https://doi.org/10.5194/bg-11-5969-2014.
Laufkötter, C., John, J.G., Stock, C.A., Dunne, J.P., 2017. Temperature and oxygen dependence of the remineralization of organic matter. Global Biogeochem. Cycles 31 (7), 1038–1050. https://doi.org/10.1002/2017GB005643.
Lechtenfeld, O.J., Hertkorn, N., Shen, Y., Witt, M., Benner, R., 2015. Marine sequestration of carbon in bacterial metabolites. Nat. Commun. 6, 6711. https://doi.org/10.1038/ncomms7711.
Lønborg, C., Álvarez-Salgado, X.A., Davidson, K., Martínez-García, S., Teira, E., 2010. Assessing the microbial bioavailability and degradation rate constants of dissolved organic matter by fluorescence spectroscopy in the coastal upwelling system of the Ría de Vigo. Mar. Chem. 119, 121–129. https://doi.org/10.1016/j.marchem.2010.02.001.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 17 (1), 10. https://doi.org/10.14806/ej.17.110.14806/ej.17.1.200.
Mehrshad, M., Rodriguez-Valera, F., Amoozegar, M.A., López-García, P., Ghai, R., 2018. The enigmatic SAR202 cluster up close: Shedding light on a globally distributed dark ocean lineage involved in sulfur cycling. ISME J. 12 (3), 655–668. https://doi.org/10.1038/s41396-017-0009-5.
Mestre, M., Ruiz-González, C., Logares, R., Duarte, C.M., Gasol, J.M., Sala, M.M., 2018. Sinking particles promote vertical connectivity in the ocean microbiome. Proc. Natl. Acad. Sci. USA 115, E6799–E6807. https://doi.org/10.1073/pnas.1802470115.
Moran, M.A., Kujawinski, E.B., Stubbins, A., Fatland, R., Aluwihare, L.I., Buchan, A., Crump, B.C., Dorrestein, P.C., Dyhrman, S.T., Hess, N.J., Howe, B., Longnecker, K., Medeiros, P.M., Niggemann, J., Obernosterer, I., Repeta, D.J., Waldbauer, J.R., 2016. Deciphering ocean carbon in a changing world. Proc. Natl. Acad. Sci. USA 113 (12), 3143–3151. https://doi.org/10.1073/pnas.1514645113.
Morán, X.A.G., García, F.C., Røstad, A., Silva, L., Al-Otaibi, N., Irigoien, X., Calleja, M.L., 2022. Diel dynamics of dissolved organic matter and heterotrophic prokaryotes reveal enhanced growth at the ocean's mesopelagic fish layer during daytime. Sci. Total Environ. 804, 150098. https://doi.org/10.1016/j.scitotenv.2021.150098.
Moreno-Ostos, E., 2012. Expedición de circunnavegación Malaspina 2010: cambio global y exploración de la biodiversidad del océano. Libro blanco de métodos y técnicas de trabajo oceanográfico. Consejo Superior de Investigaciones Científicas, Madrid.
Murali, A., Bhargava, A., Wright, E.S., 2018. IDTAXA: A novel approach for accurate taxonomic classification of microbiome sequences. Microbiome 6, 140. https://doi.org/10.1186/s40168-018-0521-5.
Murphy, K.R., Stedmon, C.A., Waite, T.D., Ruiz, G.M., 2008. Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy. Mar. Chem. 108, 40–58. https://doi.org/10.1016/j.marchem.2007.10.003.
Murphy, K.R., Stedmon, C.A., Wenig, P., Bro, R., 2014. OpenFluor- An online spectral library of auto-fluorescence by organic compounds in the environment. Anal. Methods 6 (3), 658–661.
Offre, P., Spang, A., Schleper, C., 2013. Archaea in Biogeochemical Cycles. Annu. Rev. Microbiol. 67 (1), 437–457. https://doi.org/10.1146/annurev-micro-092412-155614.
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2019. vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan.
Pajares, S., Varona-Cordero, F., Hernández-Becerril, D.U., 2020. Spatial Distribution Patterns of Bacterioplankton in the Oxygen Minimum Zone of the Tropical Mexican Pacific. Microb. Ecol. 80 (3), 519–536. https://doi.org/10.1007/s00248-020-01508-7.
Parada, A.E., Needham, D.M., Fuhrman, J.A., 2016. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ. Microbiol. 18 (5), 1403–1414. https://doi.org/10.1111/1462-2920.13023.
R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Rasse, R., Dall'Olmo, G., 2019. Do Oceanic Hypoxic Regions Act as Barriers for Sinking Particles? A Case Study in the Eastern Tropical North Atlantic. Global Biogeochem. Cycles 33, 1611–1630. https://doi.org/10.1029/2019GB006305.
Reinthaler, T., Álvarez Salgado, X.A., Álvarez, M., van Aken, H.M., Herndl, G.J., 2013. Impact of water mass mixing on the biogeochemistry and microbiology of the Northeast Atlantic Deep Water. Global Biogeochem. Cycles 27, 1151–1162. https://doi.org/10.1002/2013GB004634.
Reintjes, G., Fuchs, B.M., Amann, R., Arnosti, C., 2020. Extensive Microbial Processing of Polysaccharides in the South Pacific Gyre via Selfish Uptake and Extracellular Hydrolysis. Front. Microbiol. 11 https://doi.org/10.3389/fmicb.2020.583158.
Rinke, C., Rubino, F., Messer, L.F., Youssef, N., Parks, D.H., Chuvochina, M., Brown, M., Jeffries, T., Tyson, G.W., Seymour, J.R., Hugenholtz, P., 2019. A phylogenomic and ecological analysis of the globally abundant Marine Group II archaea (Ca. Poseidoniales ord. nov.). ISME J. 13 (3), 663–675. https://doi.org/10.1038/s41396-018-0282-y.
Ruiz-González, C., Mestre, M., Estrada, M., Sebastián, M., Salazar, G., Agustí, S., Moreno-Ostos, E., Reche, I., Álvarez-Salgado, X.A., Morán, X.A.G., Duarte, C.M., Sala, M.M., Gasol, J.M., 2020. Major imprint of surface plankton on deep ocean prokaryotic structure and activity. Mol. Ecol. 29 (10), 1820–1838. https://doi.org/10.1111/mec.15454.
Salazar, G., 2020. EcolUtils: Utilities for community ecology analysis. R package version 0.1. https://github.com/GuillemSalazar/EcolUtils.
Salazar, G., Cornejo-Castillo, F.M., Benítez-Barrios, V., Fraile-Nuez, E., Álvarez-Salgado, X.A., Duarte, C.M., Gasol, J.M., Acinas, S.G., 2016. Global diversity and biogeography of deep-sea pelagic prokaryotes. ISME J. 10 (3), 596–608. https://doi.org/10.1038/ismej.2015.137.
Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87. https://doi.org/10.1186/s12915-014-0087-z.
Saw, J.H.W., Nunoura, T., Hirai, M., Takaki, Y., Parsons, R., Michelsen, M., Longnecker, K., Kujawinski, E.B., Stepanauskas, R., Landry, Z., Carlson, C.A., Giovannoni, S.J., Zhou, J., 2020. Pangenomics analysis reveals diversification of enzyme families and niche specialization in globally abundant SAR202 bacteria. MBio 11 (1). https://doi.org/10.1128/mBio.02975-19.
Seymour, J.R., Amin, S.A., Raina, J.-B., Stocker, R., 2017. Zooming in on the phycosphere: The ecological interface for phytoplankton-bacteria relationships. Nat. Microbiol. 2 (7) https://doi.org/10.1038/nmicrobiol.2017.65.
Sheik, C.S., Jain, S., Dick, G.J., 2014. Metabolic flexibility of enigmatic SAR324 revealed through metagenomics and metatranscriptomics. Environ. Microbiol. 16 (1), 304–317. https://doi.org/10.1111/1462-2920.12165.

Stedmon, C.A., Bro, R., 2008. Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. Limnol. Oceanogr. Methods 6 (11), 572–579. https://doi.org/10.4319/lom.2008.6.572.

Stedmon, C.A., Markager, S., 2005. Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. Limnol. Oceanogr. 50 (2), 686–697. https://doi.org/10.4319/lo.2005.50.2.0686.

Stedmon, C.A., Markager, S., Bro, R., 2003. Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. Mar. Chem. 82 (3-4), 239–254. https://doi.org/10.1016/S0304-4203(03)00072-0.

Stevens, H., Ulloa, O., 2008. Bacterial diversity in the oxygen minimum zone of the eastern tropical South Pacific. Environ. Microbiol. 10 (5), 1244–1259. https://doi.org/10.1111/j.1462-2920.2007.01539.x.

Stewart, F.J., Ulloa, O., Delong, E.F., 2012. Microbial metatranscriptomics in a permanent marine oxygen minimum zone. Environ. Microbiol. 14, 23–40. https://doi.org/10.1111/j.1462-2920.2010.02400.x.

Stramma, L., Brandt, P., Schafstall, J., Schott, F., Fischer, J., Körtzinger, A., 2008. Oxygen minimum zone in the North Atlantic south and east of the Cape Verde Islands. J. Geophys. Res. Ocean. 113, C04014. https://doi.org/10.1029/2007JC004369.

Stramma, L., Schott, F., 1999. The mean flow field of the tropical Atlantic Ocean. Deep-Sea Res. Part II Top. Stud. Oceanogr. 46 (1-2), 279–303. https://doi.org/10.1016/S0967-0645(98)00109-X.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D.R., Alberti, A., Cornejo-Castillo, F.M., Costea, P.I., Cruaud, C., d'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J.M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B. T., Royo-Llonch, M., Sarmento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Bowler, C., de Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M.B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S.G., Bork, P., Boss, E., Bowler, C., Follows, M., Karp-Boss, L., Krzic, U., Reynaud, E.G., Sardet, C., Sieracki, M., Velayoudon, D., 2015. Structure and function of the global ocean microbiome. Science (80-. 348 (6237).

Swan, B.K., Martinez-Garcia, M., Preston, C.M., Sczyrba, A., Woyke, T., Lamy, D., Reinthaler, T., Poulton, N.J., Masland, E.D.P., Gomez, M.L., Sieracki, M.E., DeLong, E.F., Herndl, G.J., Stepanauskas, R., 2011. Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. Science (80-. 333 (6047), 1296–1300.

Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J.M., Brasseur, P., Rixen, M., Alvera-Azcárate, A., Belounis, M., Capet, A., Lenartz, F., Toussaint, M.E., Beckers, J.M., 2012. Generation of analysis and consistent error fields using the Data Interpolating Variational Analysis (DIVA). Ocean Model. 52–53, 90–101. https://doi.org/10.1016/j.ocemod.2012.05.002.

Tully, B.J., 2019. Metabolic diversity within the globally abundant Marine Group II Euryarchaea offers insight into ecological patterns. Nat. Commun. 10, 271. https://doi.org/10.1038/s41467-018-07840-4.

Urban-Rich, J., McCarty, J.T., Fernández, D., Acuña, J.L., 2006. Larvaceans and copepods excrete fluorescent dissolved organic matter (FDOM). J. Exp. Mar. Bio. Ecol. 332, 96–105. https://doi.org/10.1016/j.jembe.2005.11.023.

van den Boogaart, K.G., Tolosana-Delgado, R., Bren, M., 2020. compositions: Compositional Data Analysis. R package version 1.40-5. https://CRAN.R-project.org/package=compositions.

Varela, M.M., van Aken, H.M., Herndl, G.J., 2008. Abundance and activity of Chloroflexi-type SAR202 bacterioplankton in the meso- and bathypelagic waters of the (sub) tropical Atlantic. Environ. Microbiol. 10 (7), 1903–1911. https://doi.org/10.1111/j.1462-2920.2008.01627.x.

Wohlers-Zöllner, J., Breithaupt, P., Walther, K., Jürgens, K., Riebesell, U., 2011. Temperature and nutrient stoichiometry interactively modulate organic matter cycling in a pelagic algal-bacterial community. Limnol. Oceanogr. 56 (2), 599–610. https://doi.org/10.4319/lo.2011.56.2.0599.

Zark, M., Christoffers, J., Dittmar, T., 2017. Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. Mar. Chem. 191, 9–15. https://doi.org/10.1016/j.marchem.2017.02.005.

Supporting information

Deep ocean prokaryotes and fluorescent dissolved organic matter reflect the history of the water masses across the Atlantic Ocean

Markel Gómez-Letona, Javier Arístegui, Nauzet Hernández-Hernández, Xosé Antón Álvarez-Salgado, Marta Álvarez, Erick Delgadillo, María Pérez-Lorenzo, Eva Teira, Santiago Hernández-León and Marta Sebastián

This file contains:

Supporting Methods

Supporting Results

Supporting References

Supporting Tables 1-3

Supporting Figures 1-14

Supporting methods

Chromophoric dissolved organic matter

Seawater samples for the analysis of chromophoric dissolved organic matter (CDOM) were collected following the same procedure as for FDOM samples (see Materials and methods section).

Absorption spectra were determined using an Ocean Optics USB2000+UV-VIS-ES Spectrometer alongside a World Precision Instruments liquid waveguide capillary cell (LWCC) with a path length of 0.9982 m. Absorbance was recorded between 200 and 750 nm both for seawater samples and blanks (performed with freshly produced ultrapure water). Raw data was processed by subtracting the blanks to the seawater measurements (blank correction), followed by the subtraction of the average absorbance between 600 and 700 nm to the whole spectrum (dispersion correction).

The absorbance ($A_{CDOM}(\lambda)$) spectra were transformed into absorption coefficient ($a_{CDOM}(\lambda)$) spectra following the definition of the Napierian absorption coefficient:

$$a_{CDOM}(\lambda) = 2.303 \cdot \frac{A_{CDOM}(\lambda)}{L}$$

Where, for each wavelength λ, the absorption coefficient $a_{CDOM}(\lambda)$ is given by the absorbance at wavelength λ (A_λ), the path length of the cuvette (L, in meters) and 2.303, the factor that converts from decadic to natural logarithms.

The absorption coefficient at 250 nm ($a_{CDOM}(250)$) was used during the preprocessing of the excitation-emission matrices of the fluorescence data. The absorption coefficient at 254 nm ($a_{CDOM}(254)$) was used as a proxy for dissolved organic carbon concentration (Catalá et al., 2018; Lønborg and Álvarez-Salgado, 2014).

Prokaryotic cell abundance

Seawater samples for measuring the abundance of prokaryotes were collected in all stations at 22 depths, from surface to 3500 m (or bottom, where it was above this depth). Samples were collected into cryovials and fixed with paraformaldehyde at 1%, left at 4ºC in the dark for 15' and subsequently stored at -80ºC. After 24h they were analysed in a FACSCalibur (Becton-Dickinson) flow cytometer, by staining 1.2 mL of sample with 4 µL of SybrGreen I (Molecular Probes) diluted in DMSO (1:10). Fluorescent beads (Polysciences) were added for internal calibration ($10^5 \cdot mL^{-1}$).

Inorganic nutrients

Inorganic nutrients were sampled from Niskin bottles with polyethylene tubes and stored at −20°C until analysis in the laboratory. The analysis was performed with a QuAAtro 39-SEAL Analytical AutoAnalyzer following Armstrong et al. (1967).

Supporting results

Atypical prokaryotic communities

Three samples presented prokaryotic community compositions which starkly differed from the bulk of samples (Fig. 4 in the main text): samples st3-3500m, st4-3500m and st9-700m displayed rare amplicon sequence variants (ASVs) with unusually high relative abundances. Most of the ASVs present in these samples were also detected in others (Fig. S9), but a number of them (92, 1.05% of total ASVs) were only found in the outliers. The taxonomic classification of the ASVs (Fig. S7) showed important contributions by Actinobacteria (Corynebacterium, Micrococcaceae) and Firmicutes (Exiguobacterium, Anaerococcus, Lactobacillus), followed by Alphaproteobacteria (Sphingomonadales), Gammaproteobacteria (Acinetobacter) and Bacteroidota. Specially interesting is the case of st4-3500m, where these rare prokaryotic community compositions were coupled with maxima of prokaryotic cell abundance (not shown). This may be indicative of a local bloom of these rare prokaryotes. For the other two samples, however, there were no major changes in organic matter or cytometric properties.

Abundance rank curves

Abundance rank curves provided details on the differences in the taxonomy of the prokaryotic communities present in the distinct water masses. The top ASVs consistently belonged to the SAR324 clade, Nitrosopumilales (Crenarchaeota) and Marine Group II (Thermoplasmatota) (Fig. S9). However, clear changes were observed between water masses: in EQ_{13} and $ENACW_{12}$, Marine Group II was prominently found within the top taxa, alongside Nitrosopumilales and, to a lesser extent, Thiomicrospirales (Gammaproteobacteria) and SAR324. SPMW displayed a similar arrange of ASVs, but three Nitrosopumilales ASVs stood out above the rest. In $AAIW_5$ and, specially, CWD, $NADW_{4.6}$ and $NADW_2$, SAR324 ASVs were the most abundant ones and Marinimicrobia ASVs tended to occupy higher ranks. Differences between water masses were further evidenced by the dominance of the top ASVs, which can be assessed by the steepness of the rank curve (Fig. S9). EQ13, $ENACW_{12}$ and $AAIW_5$ contrasted with the bathypelagic water masses, as the former showed a gentle slope, with the top ASVs presenting relatively similar archetype abundances without marked drops. On the contrary, the CDW, $NADW_{4.6}$ and $NADW_2$ presented abrupt decreases in archetype abundances within the top 10 ASVs, especially between the first two (SAR324) and the following ones. A marked decrease was also observed for the first three ASVs in the SPMW, all classified as Nitrosopumilales. The SPMW was the only water mass dominated to such an extent by Nitrosopumilales ASVs.

References

Álvarez-Salgado, X.A., Nieto-Cid, M., Álvarez, M., Pérez, F.F., Morin, P., Mercier, H., 2013. New insights on the mineralization of dissolved organic matter in central, intermediate, and deep water masses of the northeast North Atlantic. Limnol. Oceanogr. 58, 681–696. https://doi.org/10.4319/lo.2013.58.2.0681

Álvarez, M., Álvarez-Salgado, X.A., 2009. Chemical tracer transport in the eastern boundary current system of the North Atlantic. Ciencias Mar. 35, 123–139.

Álvarez, M., Brea, S., Mercier, H., Álvarez-Salgado, X.A., 2014. Mineralization of biogenic materials in the water masses of the South Atlantic Ocean. I: Assessment and results of an optimum multiparameter analysis. Prog. Oceanogr. 123, 1–23. https://doi.org/10.1016/j.pocean.2013.12.007

Ambar, I., Howe, M.R., 1979. Observations of the Mediterranean outflow—I mixing in the Mediterranean outflow. Deep Sea Res. Part A. Oceanogr. Res. Pap. 26, 535–554. https://doi.org/10.1016/0198-0149(79)90095-5

Aparicio, F.L., Nieto-Cid, M., Borrull, E., Romero, E., Stedmon, C.A., Sala, M.M., Gasol, J.M., Ríos, A.F., Marrasé, C., 2015. Microbially-mediated fluorescent organic matter transformations in the deep ocean. Do the chemical precursors matter? Front. Mar. Sci. 2:106. https://doi.org/10.3389/fmars.2015.00106

Armstrong, F.A.J., Stearns, C.R., Strickland, J.D.H., 1967. The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. Deep. Res. Oceanogr. Abstr. 14, 381–389. https://doi.org/10.1016/0011-7471(67)90082-4

Broecker, W.S., Takahashi, Taro, Takahashi, Timothy, 1985. Sources and flow patterns of deep-ocean waters as deduced from potential temperature, salinity, and initial phosphate concentration. J. Geophys. Res. Ocean. 90, 6925–6939. https://doi.org/10.1029/JC090iC04p06925

Castro, C.G., Pérez, F.F., Holley, S.E., Ríos, A.F., 1998. Chemical characterisation and modelling of water masses in the Northeast Atlantic. Prog. Oceanogr. 41, 249–279. https://doi.org/10.1016/S0079-6611(98)00021-4

Catalá, T.S., Martínez-Pérez, A.M., Nieto-Cid, M., Álvarez, M., Otero, J., Emelianov, M., Reche, I., Arístegui, J., Álvarez-Salgado, X.A., 2018. Dissolved Organic Matter (DOM) in the open Mediterranean Sea. I. Basin–wide distribution and drivers of chromophoric DOM. Prog. Oceanogr. 165, 35–51. https://doi.org/10.1016/j.pocean.2018.05.002

Catalá, T.S., Reche, I., Fuentes-Lema, A., Romera-Castillo, C., Nieto-Cid, M., Ortega-Retuerta, E., Calvo, E., Álvarez, M., Marrasé, C., Stedmon, C.A., Álvarez-Salgado, X.A., 2015. Turnover time of fluorescent dissolved organic matter in the dark global ocean. Nat. Commun. 6:5986. https://doi.org/10.1038/ncomms6986

Chen, M., Jung, J., Lee, Y.K., Hur, J., 2018. Surface accumulation of low molecular weight dissolved organic matter in surface waters and horizontal off-shelf spreading of nutrients and humic-like fluorescence in the Chukchi Sea of the Arctic Ocean. Sci. Total Environ. 639, 624–632. https://doi.org/10.1016/j.scitotenv.2018.05.205

Coble, P.G., 1996. Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Mar. Chem. 51, 325–346. https://doi.org/10.1016/0304-4203(95)00062-3

Friedrichs, M.A.M., McCartney, M.S., Hall, M.M., 1994. Hemispheric asymmetry of deep water transport modes in the western Atlantic. J. Geophys. Res. Ocean. 99, 25165–25179. https://doi.org/10.1029/94JC02087

Georgi, D.T., 1981. On the relationship between the large-scale property variations and fine structure in the Circumpolar Deep Water. J. Geophys. Res. 86, 6556–6566. https://doi.org/10.1029/JC086iC07p06556

Harvey, J., 1982. θ-S relationships and water masses in the eastern North Atlantic. Deep Sea Res. A 29, 1021–1033. https://doi.org/10.1016/0198-0149(82)90025-5

Kida, M., Kojima, T., Tanabe, Y., Hayashi, K., Kudoh, S., Maie, N., Fujitake, N., 2019. Origin, distributions, and environmental significance of ubiquitous humic-like fluorophores in Antarctic lakes and streams. Water Res. 163:114901. https://doi.org/10.1016/j.watres.2019.114901

Lapierre, J.F., Del Giorgio, P.A., 2014. Partial coupling and differential regulation of biologically and photochemically labile dissolved organic carbon across boreal aquatic networks. Biogeosciences 11, 5969–5985. https://doi.org/10.5194/bg-11-5969-2014

Lønborg, C., Álvarez-Salgado, X.A., 2014. Tracing dissolved organic matter cycling in the eastern boundary of the temperate North Atlantic using absorption and fluorescence spectroscopy. Deep Sea Res. Part I Oceanogr. Res. Pap. 85, 35–46. https://doi.org/10.1016/j.dsr.2013.11.002

McCartney, M.S., 1982. The subtropical recirculation of mode waters. J. Mar. Res. 40, Supple, 427–464.

Mémery, L., Arhan, M., Alvarez-Salgado, X.A., Messias, M.-J., Mercier, H., Castro, C.G., Rios, A.F., 2000. The water masses along the western boundary of the south and equatorial Atlantic. Prog. Oceanogr. 47, 69–98. https://doi.org/10.1016/S0079-6611(00)00032-X

Montgomery, R.B., 1958. Water characteristics of Atlantic Ocean and of world ocean. Deep Sea Res. 5, 134–148. https://doi.org/10.1016/0146-6313(58)90004-2

Murphy, K.R., Stedmon, C.A., Wenig, P., Bro, R., 2014. OpenFluor- An online spectral library of auto-fluorescence by organic compounds in the environment. Anal. Methods 6, 658–661. https://doi.org/10.1039/c3ay41935e

Pérez, F.F., Mintrop, L., Llinás, O., Glez-Dávila, M., Castro, C.G., Alvarez, M., Körtzinger, A., Santana-Casiano, M., Rueda, M.J., Ríos, A.F., 2001. Mixing analysis of nutrients, oxygen and inorganic carbon in the Canary Islands region. J. Mar. Syst. 28, 183–201. https://doi.org/10.1016/S0924-7963(01)00003-3

Piola, A.R., Gordon, A.L., 1989. Intermediate waters in the southwest South Atlantic. Deep Sea Res. Part A. Oceanogr. Res. Pap. 36, 1–16. https://doi.org/10.1016/0198-0149(89)90015-0

Pollard, R.T., Pu, S., 1985. Structure and circulation of the Upper Atlantic Ocean northeast of the Azores. Prog. Oceanogr. 14, 443–462. https://doi.org/10.1016/0079-6611(85)90022-9

Ríos, A.F., Pérez, F.F., Fraga, F., 1992. Water masses in the upper and middle North Atlantic Ocean east of the Azores. Deep Sea Res. Part A. Oceanogr. Res. Pap. 39, 645–658. https://doi.org/https://doi.org/10.1016/0198-0149(92)90093-9

Siedler, G., Kuhl, A., Zenk, W., 1987. The Madeira Mode Water. J. Phys. Oceanogr. 17, 1561–1570. https://doi.org/10.1175/1520-0485(1987)017<1561:TMMW>2.0.CO;2

Speer, K.G., McCartney, M.S., 1992. Bottom Water Circulation in the Western North Atlantic. J. Phys. Oceanogr. 22, 83–92. https://doi.org/10.1175/1520-0485(1992)022<0083:BWCITW>2.0.CO;2

Stedmon, C.A., Bro, R., 2008. Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial. Limnol. Oceanogr. Methods 6, 572–579. https://doi.org/10.4319/lom.2008.6.572

Stedmon, C.A., Markager, S., 2005. Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. Limnol. Oceanogr. 50, 686–697. https://doi.org/10.4319/lo.2005.50.2.0686

Stedmon, C.A., Markager, S., Bro, R., 2003. Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. Mar. Chem. 82, 239–254. https://doi.org/10.1016/S0304-4203(03)00072-0

Stramma, L., England, M., 1999. On the water masses and mean circulation of the South Atlantic Ocean. J. Geophys. Res. Ocean. 104, 20863–20883. https://doi.org/10.1029/1999JC900139

Talley, L.D., 1996. Antarctic intermediate water in the South Atlantic, in: Wefer, G., Berger, W.H., Siedler, G., Webb, D.J. (Eds.), The South Atlantic. Present and Past Circulation. Springer-Verlag, pp. 219–238.

Tsuchiya, M., 1986. Thermostads and circulation in the upper layer of the Atlantic Ocean. Prog. Oceanogr. 16, 235–267. https://doi.org/10.1016/0079-6611(86)90040-6

Worthington, L. V., 1976. On the North Atlantic circulation, Johns Hopkins Oceanographic Studies, vol. 6. Johns Hopkins University Press, Baltimore.

Wüst, G., 1935. Schichtung und Zirkulation des Atlantischen Ozeans, Die Stratosphäre. Wissenschaftliche Ergebnisse der Deutschen Atlantischen Expedition auf dem Forschungsund Vermessungsschiff “Meteor” 1925–1927. 6, 180pp. English translation edited by W.J. Emery, The stratosphere of the Atlantic Ocean. Scientific Results of the German Atlantic Expedition of the Research Vessel ‘Meteor’ 1925–27. Amerind Publishing Co., 1978.

Zenk, W., 1975. On the Mediterranean outflow west of Gibraltar. Meteor. Forsch.-Ergebnisse A 16, 23–24.

Supporting tables

Supporting table 1. Characteristics of the FDOM components derived from the PARAFAC analysis, along with analogous fluorophores previously described in the literature. FDOM fluorophores from the literature were identified making use of the OpenFluor database (Murphy et al., 2014), applying threshold values of the Tucker Congruence Coefficients (TCC) of 0.95 for both excitation and emission spectra, yielding a minimum $TCC_{ex \cdot em}$ value of 0.9025 (except for C_{454}). Wavelengths between parentheses represent secondary maxima.

This study			Previously identified					
Comp.	Excitation max [nm]	Emission max [nm]	Study	Comp. name	Excitation max [nm]	Emission max [nm]	$TCC_{ex \cdot em}$	Description
$C_{462-490}$	278 (370)	462 - 490	Catalá et al. (2015b)	C1	275 (370)	480	0.9891	Humic-like, positively correlated to apparent oxygen utilisation. Coble's (1996) peak A-C
			Chen et al. (2018)	$C_{<260(365)/476}$	< 265 (365)	476	0.9665	
			Aparicio et al. (2015)	FIC2	275	468	NA	
C_{454}	< 250	454	Stedmon & Markager (2005)	C4	< 250 (360)	440	0.8879	UV humic-like / Fulvic acid fluorophore group. Coble's (1996) Peak A.
			Stedmon et al. (2003)	C1	< 240	436	NA	
			Lapierre & Del Giorgio (2014)	C3	275 (345)	436	0.9184	
C_{406}	328	406	Catalá et al. (2015b)	C2	320	400	0.9720	Marine humic-like, positively correlated to apparent oxygen utilisation. Coble's (1996) Peak M.
			Stedmon et al. (2003)	C4	325 (250)	416	0.9724	
			Aparicio et al. (2015)	FIC1	325	423	NA	
C_{366}	286 (< 250)	366	Stedmon et al. (2003)	C5	280 (< 240)	368	0.9782	Labile, protein-like material generated as a result of biological production in the water column. It is close to Coble's (1996) peak T
			Coble (1996)	Peak T	275	340	NA	
			Kida et al. (2019)	C_{360}	290	362	0.9805	

Supporting Table 2. Water types intercepted during the MAFIA cruise, brief description of source point where they belong to, characteristics, and some references with more details about their origin and circulation.

Name and acronym	Source	Characteristics	SWT	References
Salinity Maximum Water (SMW)	Tropical area (12–22ºS)	Warmest (27ºC) mode water in the SAO, formed by evaporation, transported westward to America with SEC	SMW	Worthington (1976), Stramma and England (1999), Mémery et al. (2000)
Equatorial Atlantic Central Water (13ºC)	Eastern South Atlantic, near Namibia	Formed by mixing of low salinity water outcropped further south with overlying high salinity water. Transported by the South Equatorial current to the Equator and along the Brazilian coast by the North Brazil current.	EQ_{13}	Tsuchiya (1986)
MMW	Madeira Mode Water	Mode water formed near the Madeira Island	MMW	Siedler et al. (1987)
Eastern North Atlantic Central water	Eastern North Atlantic Subtropical gyre	Mode waters defining the upper (15ºC) and lower (12ºC) limits of the subtropical ENACW formed between the area of the Azores and Portugal currents.	$ENACW_{15}$ $ENACW_{12}$	Harvey (1982), Pollard and Pu (1985), Ríos et al. (1992), Álvarez-Salgado et al. (2013)
Mediterranean Water	Gulf of Cadiz	Formed in the Gulf of Cadiz by entrainment of Eastern North Atlantic Central water on the high-salinity outflow from the Mediterranean Sea, spreads at 800–1300 m, $S > 36$ and $\theta \sim 11$–12 ºC.	MW	Zenk (1975), Ambar and Howe (1979), Castro et al. (1998), Álvarez-Salgado et al. (2013)
Antarctic Intermediate Water	Pacific Ocean north of the Sub-Antarctic Front & Malvinas-Brazil Confluence.	Formed north of the Subantarctic Front (SAF) and east of the Drake Passage by ventilation of the Subantarctic Mode Water (SAMW) formed in the Southeast Pacific.	$AAIW_{5.0}$ $AAIW_{3.1}$	McCartney (1982), Piola and Gordon (1989), Talley (1996)
Circumpolar Deep Water	Antarctic Circumpolar Current	Also named Common Water, formed by mixing in the Antarctic Circumpolar current of mid-depth Indian, Pacific and Atlantic deep water with WSDW and NADW.	CDW	Montgomery (1958), Georgi (1981), Broecker et al. (1985)
North Atlantic Deep Water	North Atlantic Ocean	Carried into the South Atlantic by the Deep Western Boundary Current (DWBC). Characterized by salinity maximum and silicate minimum (4.6ºC); and θ–S	$NADW_{4.6}$ $NADW_{2.0}$	Wüst (1935), Speer and McCartney (1992),

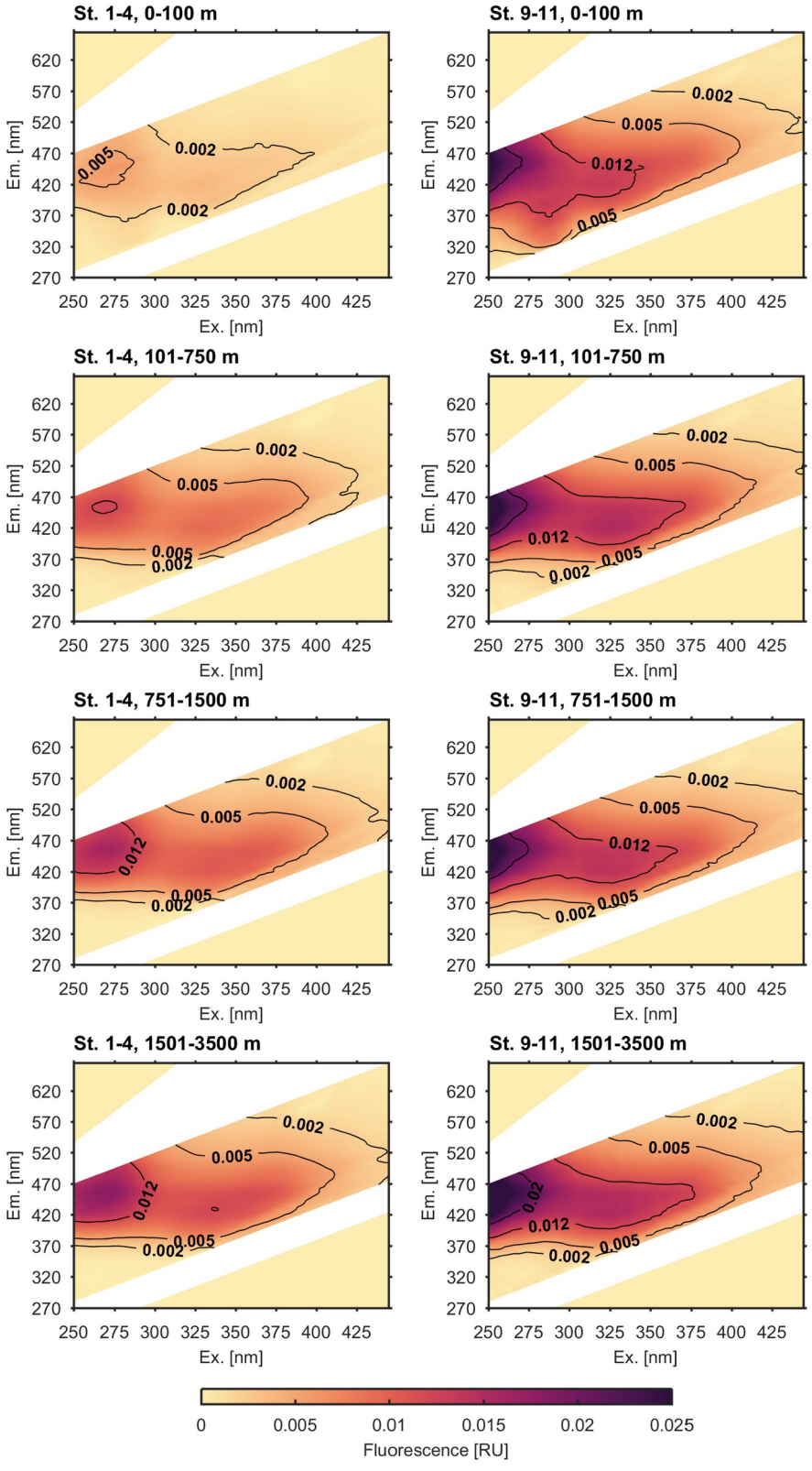
		discontinuity and oxygen maximum (2.0 ºC). Defined at their entry in the South Atlantic Ocean off South America.		Friedrichs et al. (1994)

Supporting Table 3. Thermohaline and chemical characteristics (average value ± uncertainty) of the water types (WT) introduced in the OMP analysis of the water masses intercepted during the MAFIA cruise.

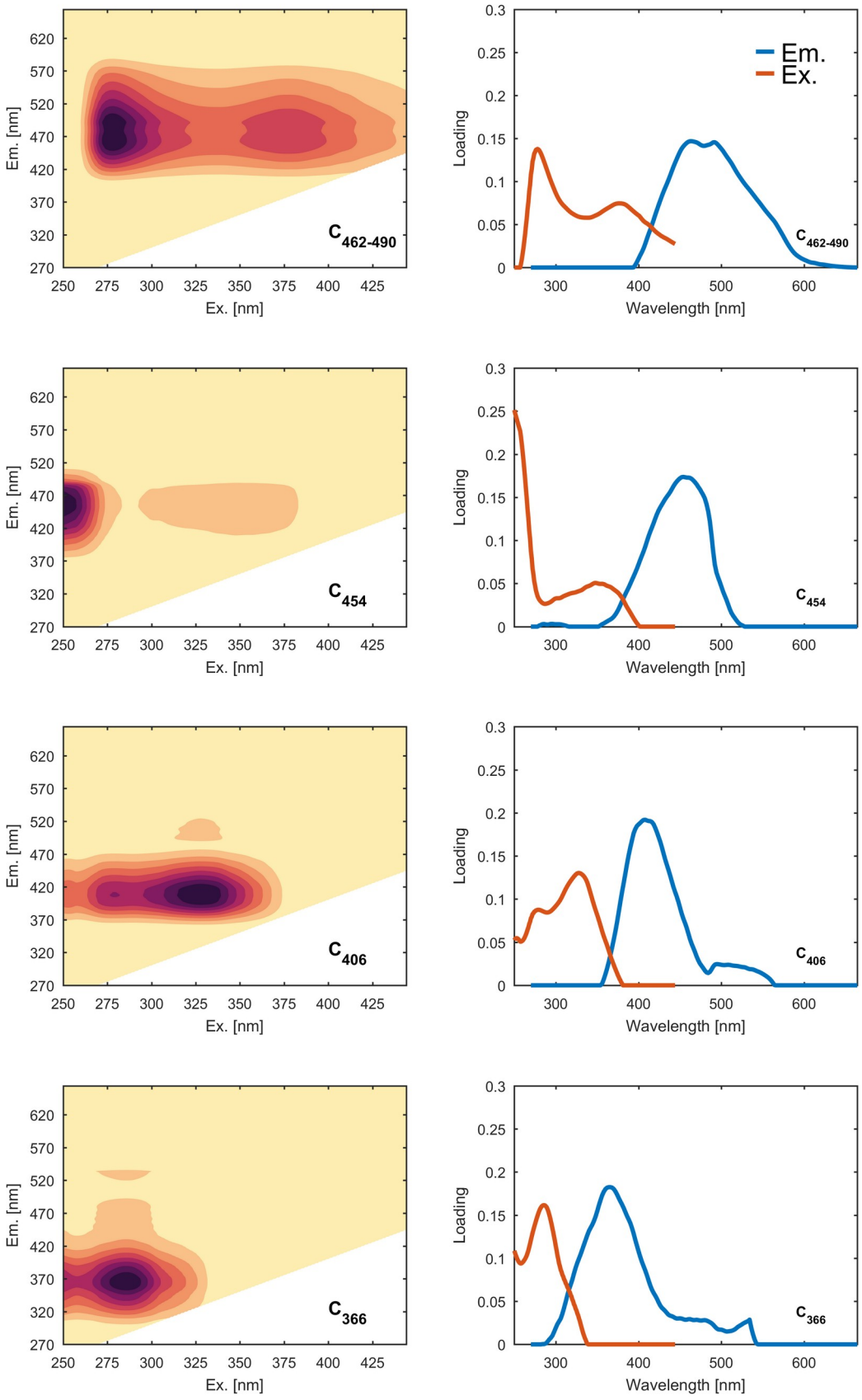
WT	θ_i (ºC)	S_i	SiO_4H_{4i} (μmol kg^{-1})	NO_i (μmol kg^{-1})
SMW [a]	27.0 ± 0.1	37.50 ± 0.01	1.1 ± 0.5	206 ± 3
MMW [b]	20.0 ± 0.5	37.00 ± 0.04	0.4 ± 0.3	225 ± 10
EQ_{13} [a]	13.0 ± 0.1	35.20 ± 0.01	5.3 ± 0.7	315 ± 3
$ENACW_{15}$ [b]	15.3 ± 0.4	36.10 ± 0.02	2.2 ± 1.7	264 ± 8
$ENACW_{12}$ [c]	12.2 ± 0.4	35.66 ± 0.02	4.9 ± 0.2	322 ± 8
SPMW [c]	8.2 ± 0.4	35.23 ± 0.01	14.5 ± 0.4	386 ± 7
MW [c]	11.8 ± 0.1	36.50 ± 0.01	7.2 ± 0.7	304 ± 9
$AAIW_5$ [a]	5.00 ± 0.08	34.14 ± 0.01	7.0 ± 0.7	482 ± 3
$AAIW_{3.1}$ [a]	3.10 ± 0.08	34.12 ± 0.01	16.4 ± 0.7	558 ± 3
CDW [a]	1.60 ± 0.03	34.720 ± 0.003	110.6 ± 0.9	497 ± 1
$NADW_{4.6}$ [a]	4.6 ± 0.1	35.020 ± 0.005	7.3 ± 0.5	426 ± 2
$NADW_2$ [a]	2.02 ± 0.03	34.910 ± 0.003	28.2 ± 0.9	446 ± 1

[a] Álvarez et al. (2014)
[b] Álvarez and Álvarez-Salgado (2009); Lønborg and Álvarez-Salgado (2014)
[c] Pérez et al. (2001); Álvarez and Álvarez-Salgado (2009)

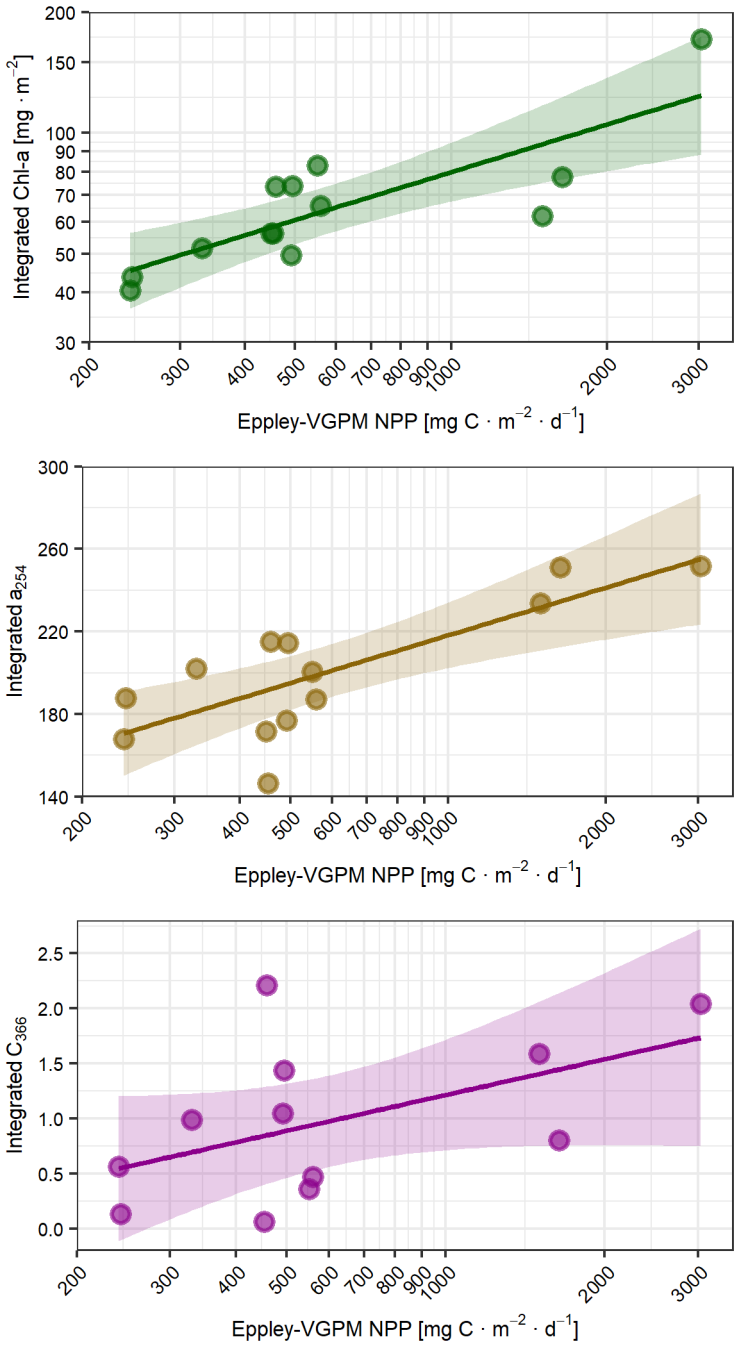
Supporting figures



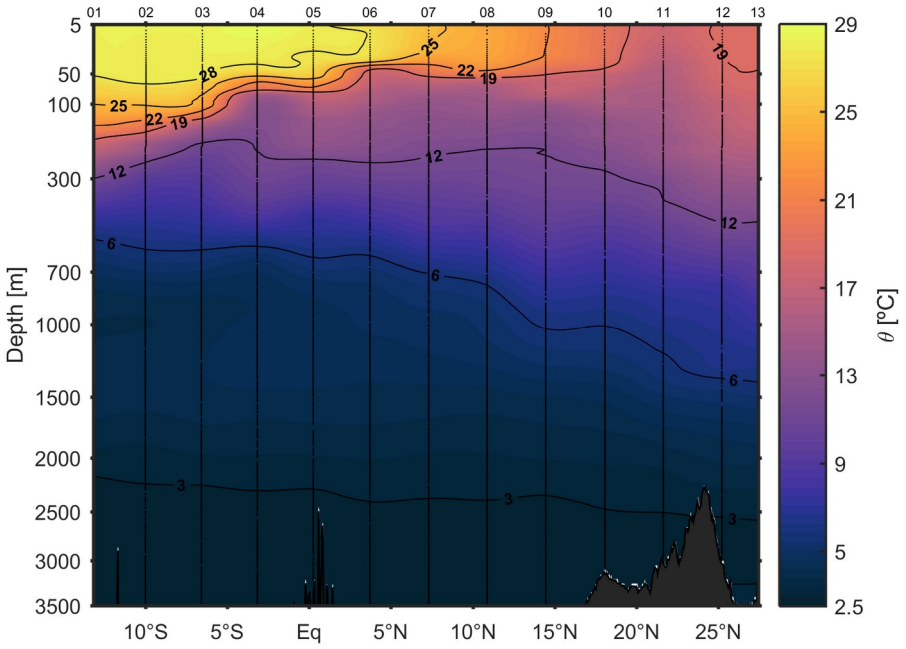
Supporting figure 1. Average excitation-emission matrices of FDOM samples from stations 1-4 (left column) and 9-11 (right column), and depths (from top to bottom) 0-100 m, 101-750 m, 751-1500 m and 1501-3500 m. EEMs were processed using the DOMFluor toolbox (v. 1.7; Stedmon and Bro, 2008) toolbox for Matlab (R2017a), see Methods for details.



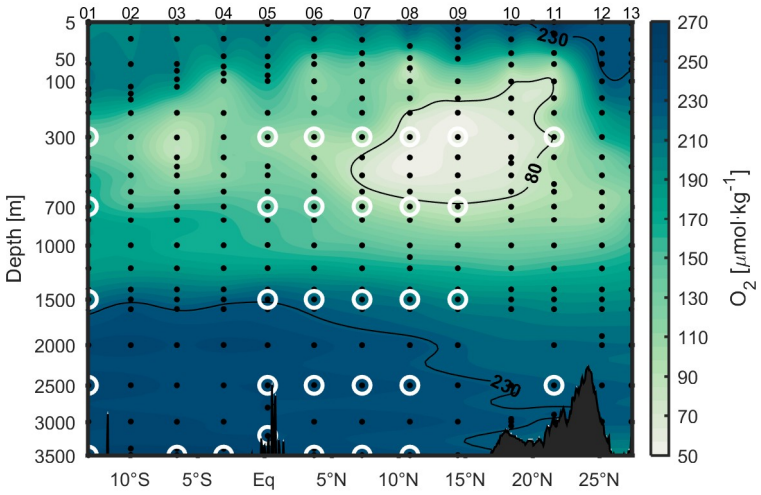
Supporting figure 2. FDOM components derived from the PARAFAC analysis (see Methods for details). In the left column, excitation-emission matrices (EEMs) of each component; in the right column, excitation (red) and emission (blue) spectra. The processed EEMs were analysed using the DOMFluor toolbox (v. 1.7; Stedmon and Bro, 2008) toolbox for Matlab (R2017a).



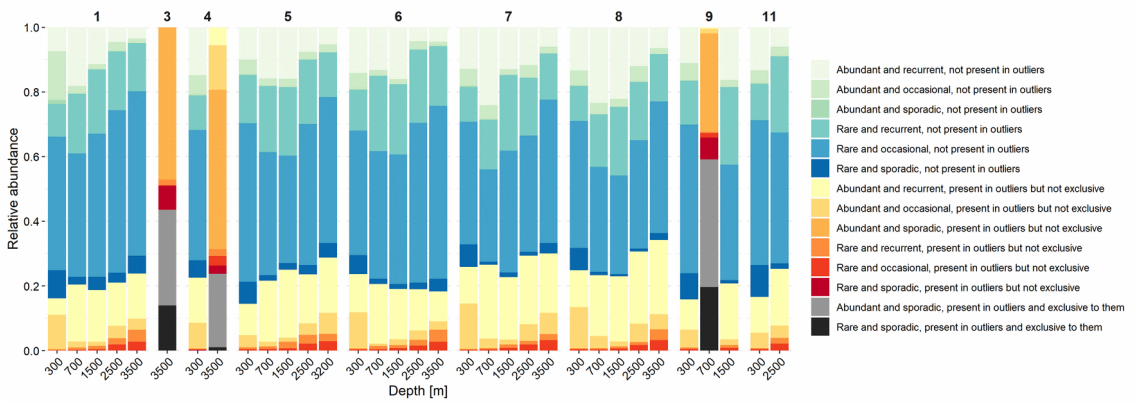
Supporting figure 3. Relationships between satellite-derived NPP (Eppley-VGPM model, MODIS dataset, April 2015; sites.science.oregonstate.edu/ocean.productivity) and integrated values of Chl-a, $a_{CDOM}(254)$ and C_{366} over the upper 200 m.



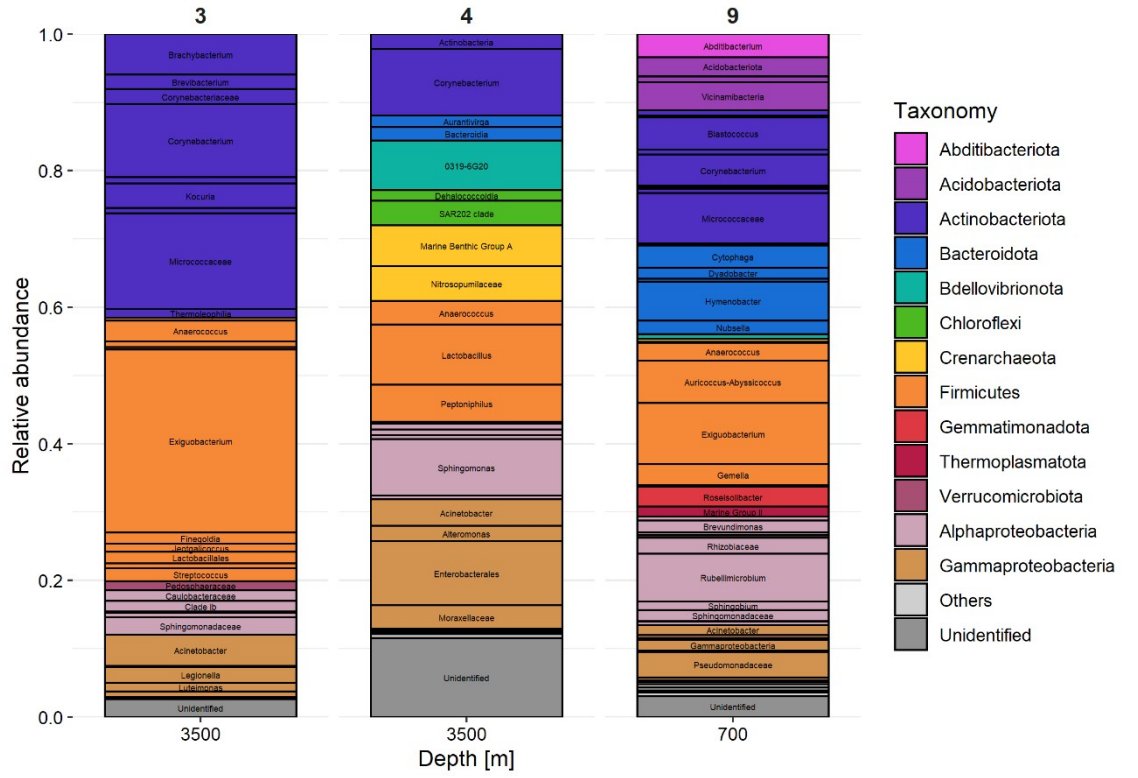
Supporting Figure 4. Potential temperature (θ) along the cruise section.



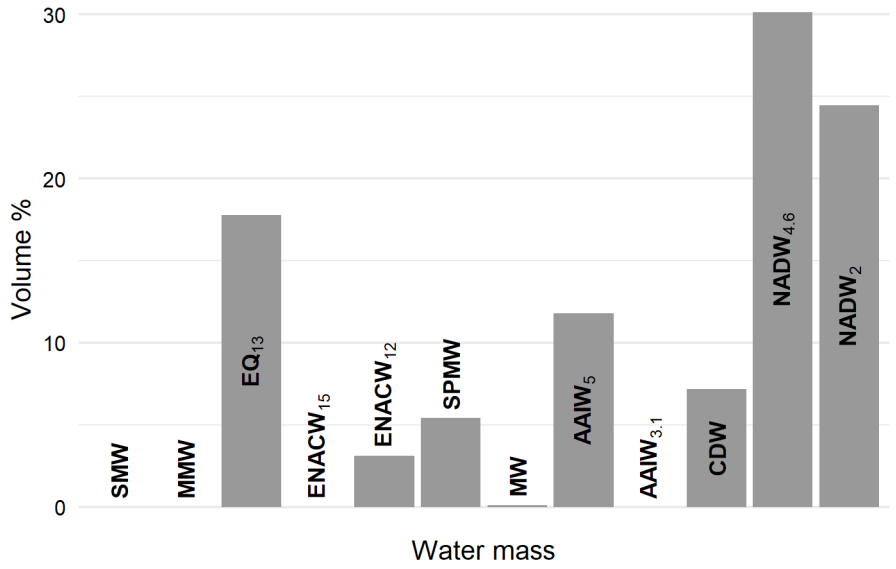
Supporting Figure 5. Latitudinal distribution of dissolved oxygen concentration. Dots represent bottle samples included in the OMP and circles the locations of amplicon samples. Note that the vertical scale is square root-transformed to allow for a better visualisation of the results. Numbers on top correspond to stations in Fig. 1. Data interpolation was performed with DIVA in Matlab (R2017a).



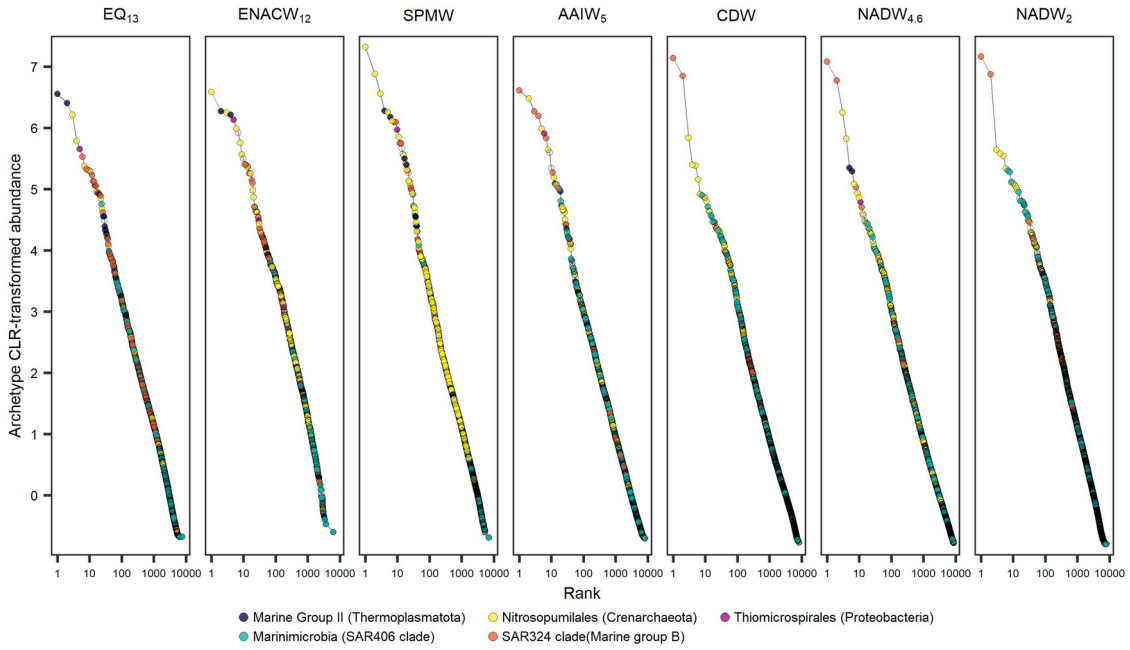
Supporting Figure 6. Amplicon sequence variant (ASV) classification by abundance and appearance frequency. Amplicon sequence variants (ASVs) that represented >1% of the community in at least one sample were defined abundant. The rest were considered rare. As to appearance frequency, ASVs present in >75% of samples were considered recurrent, those present in <75% and >25% of samples were considered occasional and those present only in <25% of samples were considered sporadic. Furthermore, ASVs were also classified according to their relation to outlier samples: ASVs exclusive to outlier samples, ASVs present both in outlier and regular samples, and ASVs absent in outlier samples.



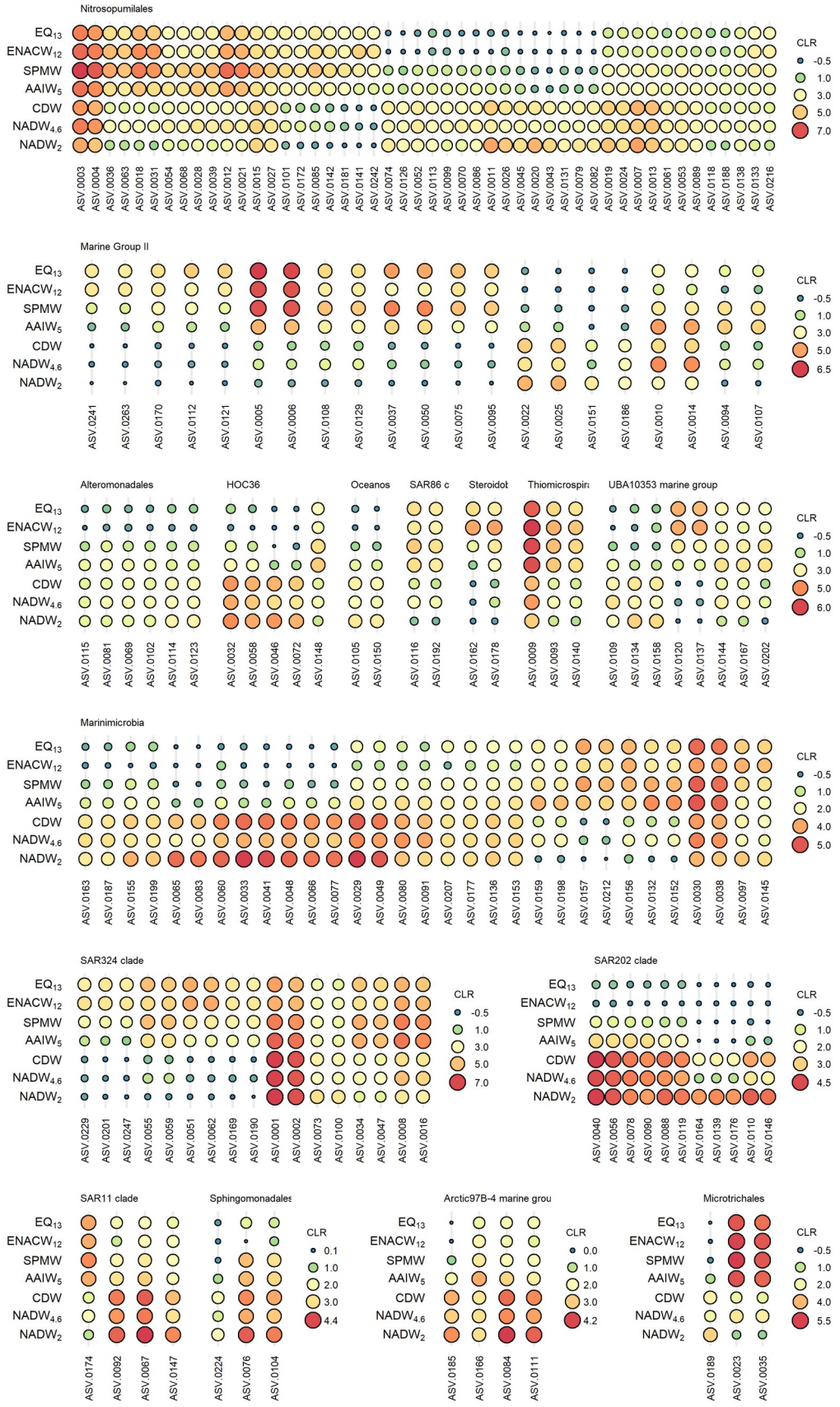
Supporting Figure 7. Detailed taxonomic classification, down to the genus level where possible, of the amplicon sequence variants (ASVs) in the outlying samples (st3-3500m, st4-3500m, st9-700m).



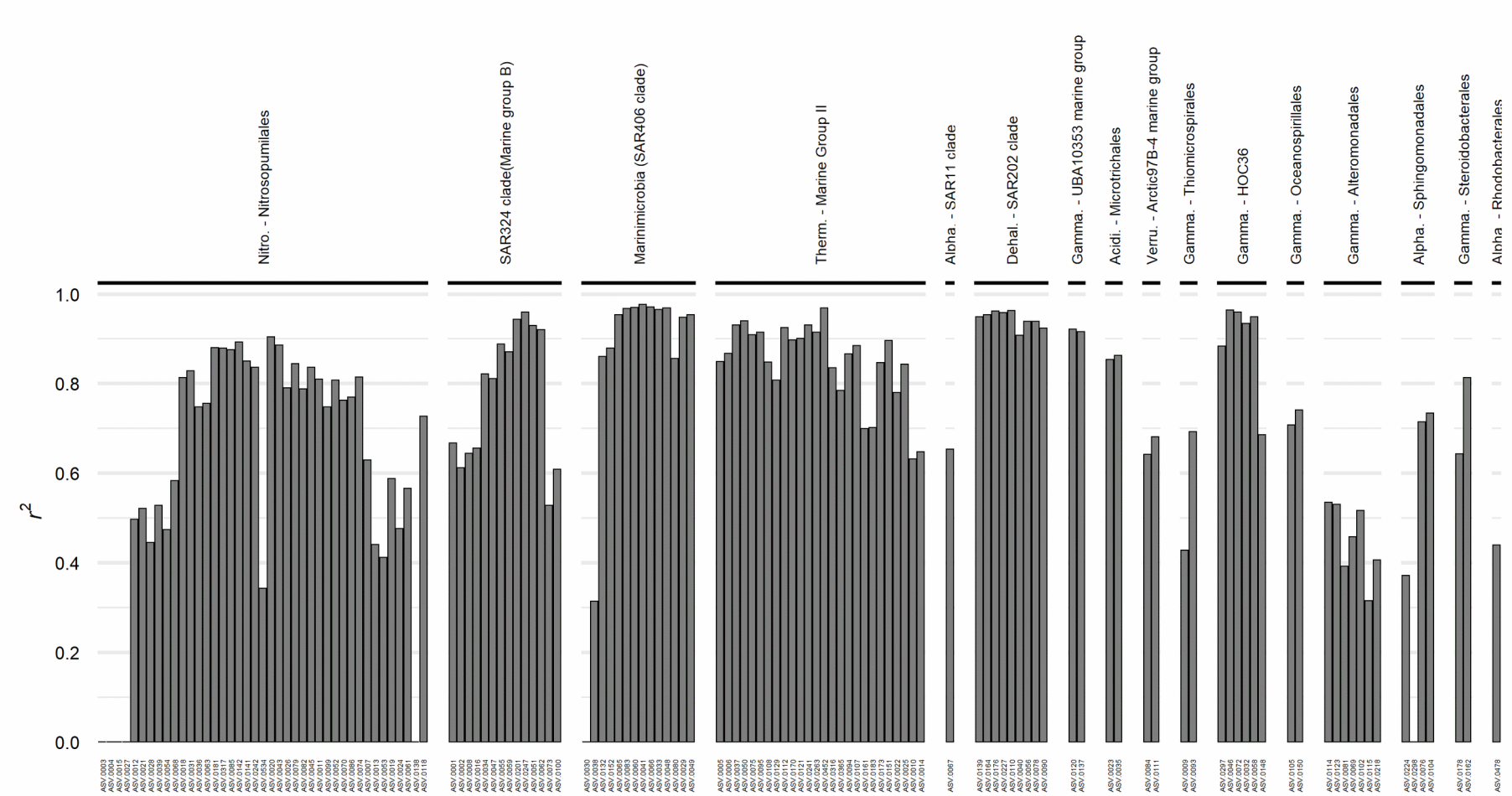
Supporting Figure 8. Accumulated contribution, as % of the total volume sampled, of the water masses in the amplicon samples.



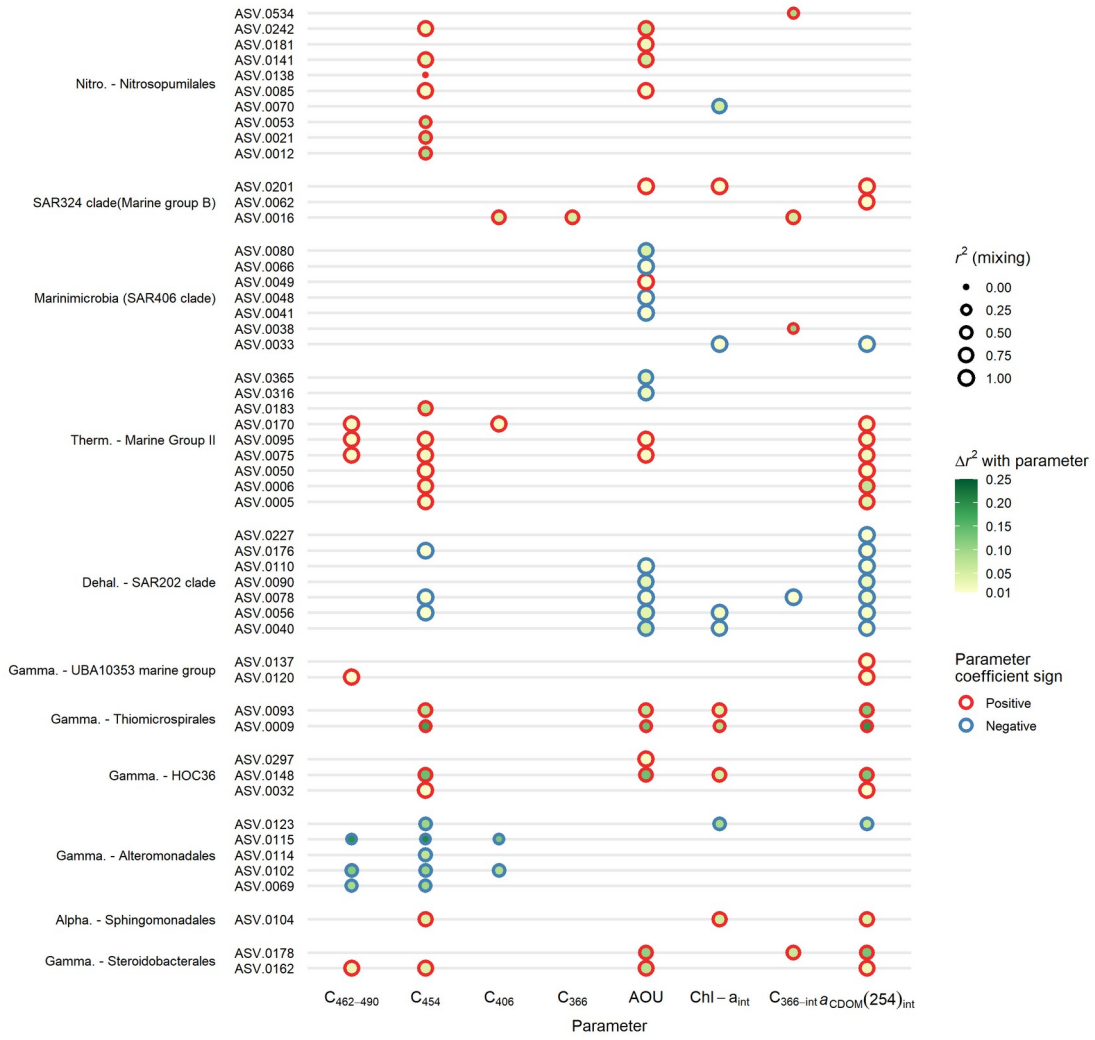
Supporting Figure 9. Taxonomic characterisation of the water masses. Rank abundance curves of amplicon sequence variants (ASVs), per water mass. Only the orders of the most abundant ASVs are colour-coded for simplicity.



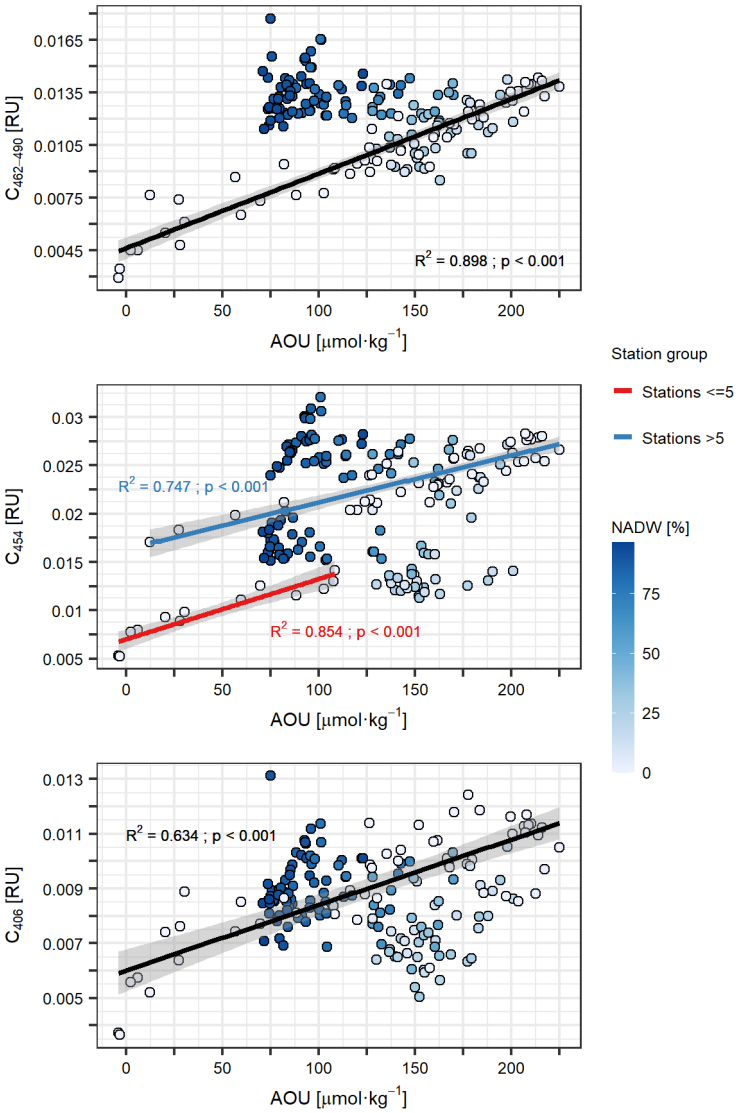
Supporting Figure 10. Archetype CLR-transformed abundances of amplicon sequence variants (ASVs) belonging to the orders/phyla shown in Fig. 6 that are present among the top-200 most abundant ones. ASVs are ordered by hierarchical clustering to group those with similar results in order to aid visualisation.



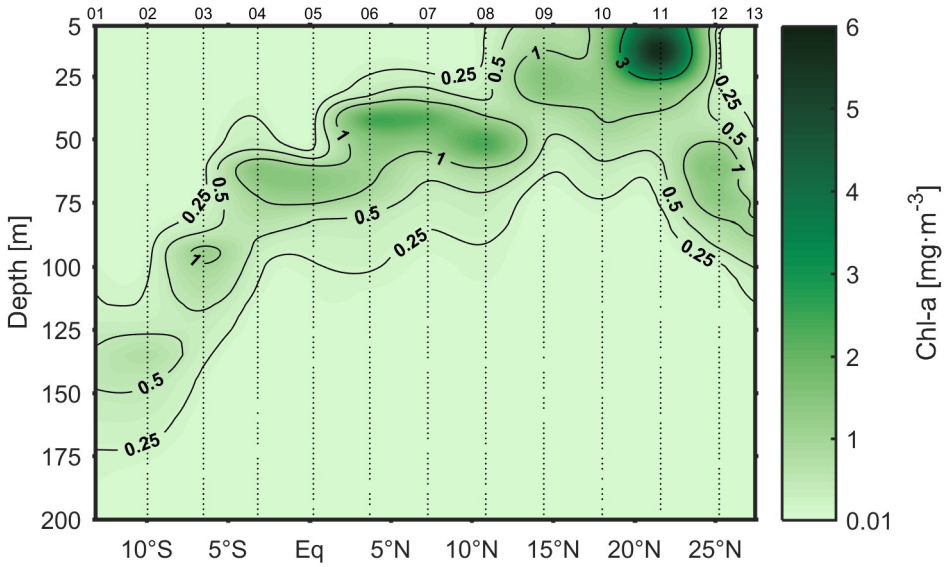
Supporting Figure 11. Results from the mixing regression models for dominant amplicon sequence variants (ASVs) (>0.5% in a sample). The explanatory capacity of the mixing models for each ASV is represented by the coefficient of determination (r^2). ASVs are grouped by order (phylum, for SAR324 clade and Marinimicrobia). When available, the abbreviated class designation is shown to the left of order label.



Supporting Figure 12. Results from the mixing + biogeochemistry regression models for dominant amplicon sequence variants (ASVs) (>0.5% in at least a sample). Only ASVs with significant results are shown. Circle size represents the coefficient of determination (r^2) of the mixing model, circle fill colour represents the r^2 improvement when including each of the biogeochemical parameters into the models, and circle border colour represents the sign of the regression coefficient of each biogeochemical parameter. ASVs are grouped by order (phylum, for SAR324 clade and Marinimicrobia). When available, the abbreviated class designation is shown to the left of order label.



Supporting Figure 13. Influence of NADW in humic-like FDOM signals, and relationship between humic-like FDOM and AOU. Linear regressions were performed only including samples with NADW ($NADW_{4.6}$ + $NADW_2$) contributions of <5%. Colormap represents the contribution of NADW (as %) in each sample. For C_{454}, regressions were divided in two station groups (1-5 and 6-12).



Supporting Figure 14. Fluorescence-based Chl-a estimates derived from the CTD fluorometer within the first 200 m.