



Plankton assessment across the distribution of West African hake and tuna based on eDNA metabarcoding

Sara Fernandez^a, Alba Ardura^a, Jose L. Martinez^b, Johannes Rick^c, Gonzalo Machado-Schiaffino^a, Eva Garcia-Vazquez^{a,*}

^a Department of Functional Biology, University of Oviedo, Spain

^b Unit of DNA Analysis, University of Oviedo, Spain

^c Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, AWI - Department of Coastal Ecology, Germany

ARTICLE INFO

Keywords:

Diatom-based indices
Environmental DNA
Fish preys
Plankton community inventory

ABSTRACT

The richness of plankton communities determines the fish productivity in the ocean, including important resources that rely on extractive fisheries, such as hakes (genus *Merluccius*) and tunas (genus *Thunnus*). Their preys forage on zooplankton, and the latter feed on phytoplankton. Inventories of plankton communities for scientific advice to sustainable fishing are essential in this moment of climate change. Plankton is generally inventoried using conventional methodologies based on large water volumes and visual morphological analyses of samples. In this study, we have employed metabarcoding on environmental DNA (eDNA) samples extracted from small water volumes for plankton inventory from twelve distant sampling stations in the East Atlantic Ocean. Zones rich in hake and tuna prey were detected from eDNA, and multivariate multiple regression analysis was able to predict those zones from diatom-based indices and planktonic diversity based on functional groups. Salinity was negatively correlated with the proportion of diatoms in phytoplankton, highlighting expected impacts of current global change on marine plankton communities. The results emphasise the importance of the plankton richness for fish productivity and support the utility of environmental DNA as a tool to monitor plankton composition changes.

1. Introduction

Overfishing has led to an important erosion of wild fish stocks; sustainable fishing supported by scientific advice is an urgent need in regions that rely on extractive fisheries for livelihoods and food such as West African countries (Sumaila et al., 2016). In this moment of climate change such scientific advice is especially important because species are shifting due to changing oceanographic conditions, and productivity may sharply oscillate over years (Karp et al., 2019). Hakes of the genus *Merluccius* and tunas of the genus *Thunnus* are fishing targets of enormous economic value. Although the bluefin tuna *Thunnus thynnus* aquaculture is taking place including the full cycle, attempts to farm *Merluccius* species in Norway (Bjelland and Skiftesvik, 2006), Spain (Iglesias et al., 2010) or Chile (Bustos et al., 2007) showed high mortality in aquaculture conditions. Therefore, global production relies on extractive fisheries that are very important from the economic point of view.

These hakes and tunas are carnivorous. They prey upon a variety of

zooplankton taxa and/or ichthyoplankton depending on the life stage (e.g., Bozzano et al., 2005; Ménard et al., 2006; Mahe et al., 2007). Many of their direct preys, such as Euphausiacea krill for fish juveniles (e.g., Bozzano et al., 2005), feed on phytoplankton. Therefore, both phytoplankton and zooplankton are important components in the trophic chain(s) where hakes and tunas occupy high positions.

Satellite data of sea-surface chlorophyll-a has been routinely employed as an indicator of fish productivity (Sachoemar, 2015). However, not all the phytoplankton species have the same value for zooplankton to graze, and there are differences between species. Diatoms provide a large part of carbon to zooplankton consumers (Fry and Wainright, 1991). Although some macrozooplankton like Cladocerans may select dinoflagellates and green algae over diatoms and cryptophytes (Levine et al., 1999), diatoms contribute substantially to the growth of shrimp (Fernandes et al., 2019), and are key in determining the abundance of large copepods (Benedetti et al., 2019). On the other hand, diatom dominance over dinoflagellates indicates good environmental status (Wasmund et al., 2017). The composition of the

* Corresponding author.

E-mail address: egv@uniovi.es (E. Garcia-Vazquez).

<https://doi.org/10.1016/j.marenvres.2023.106312>

Received 20 October 2023; Received in revised form 28 November 2023; Accepted 18 December 2023

Available online 22 December 2023

0141-1136/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

phytoplankton can be determined from conventional plankton surveys that comprise plankton sampling using nets of different types and mesh, visual sorting and counting under the microscope and taxonomic identification with the help of specialized guides (e.g., Ferronato et al., 2021; Kasyan et al., 2022). To shorten this process, some methodologies such as pigment signatures, using the photosynthetic pigment to estimate the phytoplankton abundance in field populations by HPLC method (e.g., Llarío et al., 2020), have been applied.

Environmental DNA (eDNA) has been employed to inventory plankton, detect changes in the ocean community structure with depth, and study the spatial distribution of species (Chavez et al., 2021). The results obtained using eDNA approaches reflect the plankton community composition: DNA concentration is significantly correlated with chlorophyll-a, the number of 18S rDNA reads corresponding to diatoms is correlated with diatom biomass, and the same happens roughly with the Cytochrome c oxidase subunit I (COI) reads and the biomass of copepods (Chavez et al., 2021). Diversity estimations from metabarcoding reads are robust and can be considered reliable in aquatic environments (Nguyen et al., 2020; Drummond et al., 2021), being able to identify microalgae at a species level (e.g. Ardura et al., 2021). The preys that form part of carnivorous fish diet can be identified from eDNA metabarcoding on plankton samples; even small water volumes can be sufficient for this purpose (Garcia-Vazquez et al., 2021).

In this study, we aim to disentangle the elements that contribute to the landscape of the planktonic resources using eDNA as a data source. We will focus on the East Atlantic Ocean, where *Merluccius* and *Thunnus* species are fished by the European fleet in European waters, and in West African waters under Sustainable Fisheries Partnership Agreements (SFPA; https://oceans-and-fisheries.ec.europa.eu/fisheries/international-agreements/sustainable-fisheries-partnership-agreements-sfpas_en, accessed June 2023). Results obtained from eDNA metabarcoding will be employed to explore on its use to identify plankton and other marine organisms, estimating the richness of phytoplankton and zooplankton as well as grazing indices. Departure hypotheses were:

- i) From Chavez et al. (2021) results of eDNA, we expected a correlation between chlorophyll-a and eDNA concentration.
- ii) From the variety of zooplankton preys and the importance of diatoms for zooplankton to graze (e.g., Benedetti et al., 2019; Fernandes et al., 2019), we expected that planktonic diversity and diatom-based indices would predict the richness of *Merluccius* and *Thunnus* preys.
- iii) Since diatoms are indicators of good environmental status (Wasmund et al., 2017), the relative abundance of diatoms would correlate with Molecular Operational Taxonomic Units (MOTU)-based biodiversity indices.

2. Material and methods

2.1. Oceanic transect studied, physic-chemical parameters and sampling locations

Water samples were taken from twelve points along a latitudinal transect on board RV Polarstern in the research cruise PS116 (Supplementary Table 1). The cruise across the East Atlantic was from Bremerhaven (North Sea) to South African waters (Fig. 1) between November and early December 2018. The cruise details are in Hanfland and König (2019). Water temperature and salinity were continuously recorded during the cruise (Hanfland et al., 2019); their values during the water sample uptake and the ship position (coordinates) in each sampling point are in Supplementary Table 1. Chlorophyll concentration data were taken from the NASA Earth Observation (NEO - <https://neo.gsfc.nasa.gov/>). A map was generated considering the period of sampling, from 12/11/2018 to 5/12/2018 and the exact coordinates of the sampling points in the latitudinal transect (Fig. 1, Supplementary Table 1). The chlorophyll values as $\text{mg}\cdot\text{m}^{-3}$ are given in Supplementary

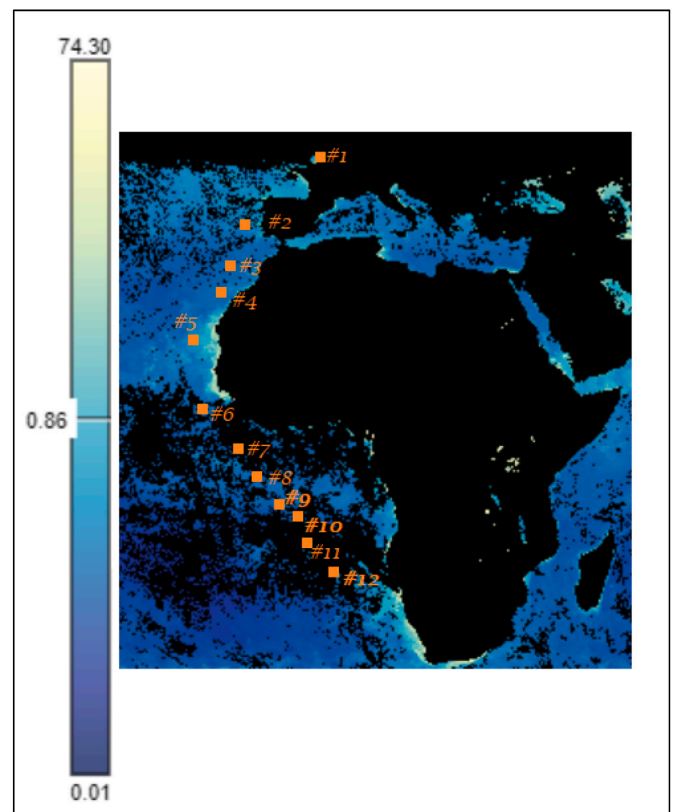


Fig. 1. Map showing the sampling points (#1 to #12) and concentrations of satellite-derived chlorophyll in $\text{mg}\cdot\text{m}^{-3}$ in the sampling dates. The map was produced through the Nasa Earth Observations (NEO) tool (<https://neo.gsfc.nasa.gov/>). The black areas show where the satellite could not measure chlorophyll.

Table 1. The chlorophyll concentration detected varied between $1.8 \text{ mg}\cdot\text{m}^{-3}$ in Bremerhaven port (location #1) to $0 \text{ mg}\cdot\text{m}^{-3}$ in the sampling locations #6 to #9 coinciding with the known oligotrophic area at the northwest of the Gulf of Guinea due to coastally trapped eddies in this area (McGillicuddy et al., 2007; Djakourke et al., 2014).

The research cruise crossed three FAO major fishing areas (<https://www.fao.org/fishery/en/area/search>): 27 (Atlantic European waters), 34 (north central west African waters) and 47 (central south west African waters). These include territorial waters of African countries that have SFPA with Europe. Going southwards, there are three mixed agreements that include explicitly tuna and hake (with Morocco, Senegal and The Gambia), four tuna agreements (Cabo Verde, Cote d'Ivoire, Gabon, Sao Tomé e Príncipe), and other two mixed agreements that contain tuna and a variety of cephalopods, shrimp, small pelagic and demersal fish thus hake may be caught (Mauritania, Guinea-Bissau). The agreement with Morocco includes other pelagic and demersal species in addition to tuna and hake (https://oceans-and-fisheries.ec.europa.eu/fisheries/international-agreements/sustainable-fisheries-partnership-agreements-sfpas_en).

2.2. Sampling procedure

Sampling details can be found in Garcia-Vazquez et al. (2021). Briefly, three replicates of 2L water each were taken from twelve sampling points using an underwater pump system integrated in the Research vessel Polarstern. Water samples were filtered on board using PES Supor 200 Membrane Filters (Pall Corporation, Life Sciences) with $0.2 \mu\text{m}$ pore size and 47 mm diameter. The filters were preserved in 100% ethanol at room temperature.

Some individuals retained in the filters were carefully taken with

tweezers and put on separate glass slides. They were classified *de visu* under the microscope with the help of taxonomic guides, for validation of the metabarcoding methodology.

All the material was sterilized with 10% bleach after and before samplings to avoid cross contamination. Laboratory coat and disposable gloves were worn all time. Onboard, an independent wet laboratory with working surfaces cleaned with 10% at the beginning and at the end of each working day, was employed for these filtrations to prevent contamination.

2.3. DNA analysis

Once in the laboratory facilities at the University of Oviedo, eDNA was extracted from filters using PowerWater DNA Isolation Kit (Qiagen) following manufacturer's instructions. A previous step to pellet the ethanol content and add it in the extraction process was included. All extractions were performed in a pre-PCR laboratory under a flow laminar hood equipped with UV light. Negative controls were added to each extraction process.

Fragments from two genes were amplified with polymerase chain reaction (PCR) from the extracted eDNA using universal primers. A fragment of the cytochrome oxidase subunit 1 gene (COI), using primers from Leray et al. (2013) that were modified with a PGM sequencing adaptor, the barcodes (one per sample) needed to differentiate the reads belonging to each water sample, and a "GAT" spacer. Amplification was carried out in a total volume of 20 μ L including Green GoTaq Buffer 1X, 2.5 mM MgCl₂, 0.25 mM dNTPs, 20 pmol of each primer, 4 μ L of template DNA, 200 ng/ μ L of bovine serum albumin (BSA), and 0.65 U of DNA Taq polymerase (Promega). PCR conditions in a Veriti Thermal Cycler (Applied Biosystems, Foster City, California) were 95 °C for 1 min, followed by 35 cycles of 95 °C for 15 s, 46 °C for 15 s, 72 °C for 10 s, and a final extension of 72 °C for 3 min.

For the second gene, a fragment of the eukaryotic V4 region of the nuclear small subunit ribosomal DNA (18S rRNA gene) was amplified using the universal primers Uni18SF and Uni18SR from Zhan et al. (2013), properly modified with the PGM sequencing adaptor and the barcodes to differentiate between samples, as well as the "GAT" spacer. Amplification was carried out in a total volume of 20 μ L including Green GoTaq Buffer 1X, 2.5 mM MgCl₂, 0.25 mM dNTPs, 20 pmol of each primer, 4 μ L of template DNA, 200 ng/ μ L of bovine serum albumin (BSA), and 0.65 U of DNA Taq polymerase (Promega). PCR conditions in a Veriti Thermal Cycler (Applied Biosystems, Foster City, California) were 95 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 90 s, and a final extension of 72 °C for 8 min. Extraction of negative controls were also included in the PCRs and PCR blanks were also performed.

The amplification success was visually assessed on 2% agarose gel. PCR amplicons were purified from agarose gel using the Montage DNA Gel Extraction Kit (Millipore) and quantified using the Qubit BR dsDNA Kit (Thermo Fisher Scientific). A Bioanalyzer 2100 (Agilent Technologies) was employed to confirm the fragment size, the absence of by-products, and to accurately quantify DNA before sequencing. Amplicon samples were diluted down to 26 pmol for the pooling in an equimolar concentration. The pool was processed by liquid emulsion PCR in the One Touch System using the Ion PGM™ OT2 Supplies Kit (Life Technologies) following the manufacturer's instructions. Then, it was loaded in the Ion Chip (Life Technologies) and sequenced employing the Ion Torrent Personal Genome Machine (Life technologies) following the specifications in the protocol Ion PGM™ Sequencing Kit.

2.4. Bioinformatics and taxonomic assignment

Adaptors from Ion Torrent platform were trimmed within the platform software and fastq sequences files were used to filter per quality. Qiime2 2020.2 (Caporaso et al., 2011) was used to trim the primers and to filter by length (amplicon size 200–400 bp reads were retained in COI

and 100–600 bp reads in 18S) using *Cutadapt* software (Martin, 2011). Denoising option was employed where sequences are dereplicated and denoised using the *unoise3* algorithm from *vsearch* 11.0.6 (Rognes et al., 2016). For taxonomic classification, filtered sequences were compared against a public COI and 18S reference database (NCBI, accessed on June 16, 2020) and stored locally. The database was downloaded using the *esearch* queries "COI; NOT Bacteria; NOT environmental; NOT viruses; NOT unclassified" and "18S; NOT uncultured; NOT bacteria; NOT unclassified" and constructed with the respective taxonomic information using the script "Entrez_qiime.py" by Baker (2017). Finally, "qiime feature-classifier" command was employed to assign the taxonomy, using a 97% and 80% for COI and 18S reads respectively as identity percentage and an e-value of 10⁻⁵⁰. Resulting MOTUs from taxonomic assignment were employed in further steps, choosing genus as the assignment level to be considered a MOTU. Taxonomic information was checked in World Register of Marine Species (WORMS; <http://www.marinespecies.org/>).

2.5. Biota classification, functional groups and biological indices considered

Plankton indices were calculated based on the MOTU table. There are many plankton indices of community functioning (e.g. Pomerleau et al. (2015) for zooplankton, and Weithoff and Beisner (2019) for phytoplankton), but most of them require the individuals of each species to be counted while in metabarcoding the number of sequence counts are correlated with the biomass (Ershova et al., 2021). In this study, the taxonomic diversity was estimated calculating Shannon index based on the number of MOTUs per class as a unit. Diversity based on functional groups was also estimated in each sampling location from the number of MOTUs in each of the following: phytoplankton, zooplankton, ichthyoplankton and large predators (species with big eggs that cannot be a part of the plankton). Shannon index was calculated from these data. The rest of indices were based on the number of species as in Moncheva et al. (2002) but using MOTU instead of morphologically described taxa. These were: phytoplankton taxonomic dominance (Bacillariophyceae: Dinophyceae species ratio = Diatoms/Dinoflagellates), grazing pressure (phytoplankton: zooplankton species; and Bacillariophyceae: Copepoda = Diatoms/Copepods) where the higher value corresponds to the lower grazing pressure).

The proportion of tuna and hake preys over the total number of zooplankton MOTU were calculated from the MOTU table considering the sampling depth. Being the samples collected in the upper 6 m of the water layer, we focused on prey of tuna of any age – tunas are pelagic and prey of juvenile hakes, because adult hakes are demersal and will rarely eat at that shallow depth. Preys occurring in Atlantic *Thunnus* diet were based on Ortiz de Zárate (1987), Ménard and Marchal (2003), Pusineri et al. (2005), Ménard et al. (2006), Rudershausen et al. (2010), Logan et al. (2011), Olafsdottir et al. (2016) and Valls et al. (2022). Those occurring in the diet of juvenile *Merluccius* species distributed in the East Atlantic (Old Continent species *Merluccius merluccius*, *M. senegalensis*, *M. polli*, *M. capensis* and *M. paradoxus*, from North to South) were taken from Maurin (1954), Alheit and Pitcher (1995), Pillar and Wilkinson (1995), Roel and Macpherson (1988), Garrison and Link (2000), Lloris et al. (2005), Mahe et al. (2007) and Wilhelm et al. (2015).

2.6. Statistics

Exploratory data analysis was run with correlations between variables first, then a principal component analysis (PCA). For the PCA the correlation option was employed. We created a scatter plot with diagonals proportional to the weight of each variable. The threshold considered sufficient for Eigenvalues was 0.7.

Pairwise correlations between variables were estimated from Pearson's *r* after checking normality. Normality was tested using Shapiro-Wilk parameter. Comparison between sampling locations for the

distribution of MOTUs in taxonomic groups was done with contingency statistics, using chi-square and Cramer's V as a proxy of effect size.

Multivariate multiple regression analysis was run to identify the independent variables (plankton indices) that predict significantly the dependent variables of interest, i.e. the proportion of hake and tuna preys.

The significance threshold applied was 0.05. PAST software was employed (Hammer et al., 2001).

3. Results

3.1. NGS results

DNA was extracted and successfully PCR-amplified from the water samples obtained in the twelve sampling points considered (Supplementary Table 2). DNA concentration varied between 0.170 ng/ μ L in the location #3 to 1.469 ng/ μ L in the location #12, and the total number of quality reads (summing COI and 18S metabarcodes) ranged from 846 to 383,422 in locations #6 and #7, respectively (Supplementary Table 2). For a higher reliability we used a threshold of >30,000 quality reads for downstream functional and multivariate analysis; thus we excluded location #6 from those analyses.

The NGS results are available in the Bioproject number PRJNA675458 and Biosample accession number SAMN16708468. After bioinformatics assignment, all reads assigned to non-marine species (human DNA, birds, terrestrial plants in locations near the coast) were discarded from the analysis, number of considered reads per sampling point is showed in Supplementary Table 2 (varying from 0.4 to 0.94 of the total reads recovered). The results of COI metabarcode for the sampling sites #5 to #8 were published in Garcia-Vazquez et al. (2021), where they were named WA#1 to WA#4 (WA for West Africa); the rest of data are novel.

Merging the results obtained from the two primer sets, 77 putative marine genera (MOTU) were identified (Supplementary Table 3). Of those, the diatom *Odontella sinensis*, the copepods *Nitokra spinipes* and *Delibus* sp., and the krill *Euphausia* sp., were identified *in situ* confirming the molecular result; some were also confirmed individually from DNA (Garcia-Vazquez et al., 2021). The taxonomic profile was significantly different between sampling locations (Fig. 2) with contingency $\chi^2 = 77.65$, d.f. = 54, $p = 0.019$, Cramer's V = 0.29. In the two ports the majority of MOTUs were of marine mammals (four genera of Delphinidae), plus a few cartilaginous fish in Bremerhaven (two genera of Carcharhiniformes, one genus of Rajiformes), and a dinoflagellate (genus *Alexandrium*) in Las Palmas. In contrast, locations such as #2, #5 or #12 contained a variety of different animals and algae classes (Supplementary Table 3, Fig. 2). Location #6 had only three MOTUs, as expected from small number of quality reads.

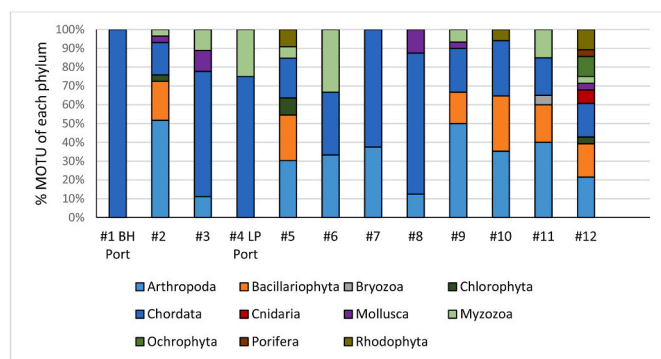


Fig. 2. Taxonomic variation found from environmental DNA (eDNA), presented as the proportion of MOTU (= putative genera) per phylum in the sampling locations considered. BH and LP stand for Bremerhaven and Las Palmas ports, respectively.

3.2. Functional groups and plankton indices

For small number of quality reads, the location #6 was excluded from this and further analysis. By functional groups, the sampling locations were also different ($\chi^2 = 70.37$, d.f. = 30, $p < 0.001$, Cramer's V = 0.34) (Fig. 3). DNA of large predators (marine mammals, large fish) occurred in most sites, but in the two ports they were clearly dominant (Fig. 3). Phytoplankton was dominant in locations #5 and #12, zooplankton in #2, #9 and #11, and large predators in the rest except in #10 where phytoplankton and zooplankton were balanced and more abundant (in MOTUs) than the rest. The number of phytoplankton genera per location (or phytoplankton richness) varied greatly among sampling sites. In the locations #1 (Bremerhaven), #7 and #8 (both in front of the Gulf of Guinea, see Fig. 1) phytoplankton DNA was not found in this study (Fig. 3). Ichthyoplankton DNA occurred in the majority of locations but corresponding only to one or a few MOTU per location (Fig. 3).

The molecular ecological indices calculated differed among locations, according to different functional groups and taxa (Table 1). The highest taxonomic diversity corresponded to sampling point #5, where the ratio phytoplankton/zooplankton was also the highest, and the lowest to. The highest diversity based on functional groups occurred in location #10 and the lowest in Bremerhaven port (#1) where we found DNA of large predators only. Diatoms dominated over dinoflagellates in the sites with a high taxonomic diversity (Table 1).

3.3. Hake and tuna prey richness and influencing variables

Potential preys of hake identified in this study from DNA were: one genus of amphipod, three genera of Euphausiids (krill), two of decapods, and three fish from different orders (Blenniiformes, Clupeiformes and Gobiiformes) (Table 2). In addition to those, one isopod and two genera of Pleuronectiformes can be preys of tuna. DNA of at least one tuna prey was found from all water samples except those taken from ports, while DNA of hake preys occurred from eight sampling sites, being absent from the sampling point #5 in front of Mauritania (Table 2).

The landscape of the eleven locations regarding the analysed variables is visualized in the PCA scatter plot (Fig. 4), and the PCA results, including Eigenvalues, proportion of the variance explained by each PC and loadings of the different factors, in the Supplementary Table 4. PC1, PC2, PC3 and PC4 had Eigenvalues >0.7. The two measures of diversity had the higher load in PC1; the proportion of tuna prey MOTUs and the grazing index Diatoms/Copepods in PC2; the proportion of hake prey MOTUs and salinity in PC3; and temperature and tuna preys in PC4 (Supplementary Table 4). In the scatter plot of PC1 over PC2 (the two components summed 66.4% of the total variance), the two grazing indices were located close to each other, and the Phytoplankton dominance index (Diatoms/Dinoflagellates) was next to eDNA concentration (Fig. 4). The proportion of hake preys was between the two diversity

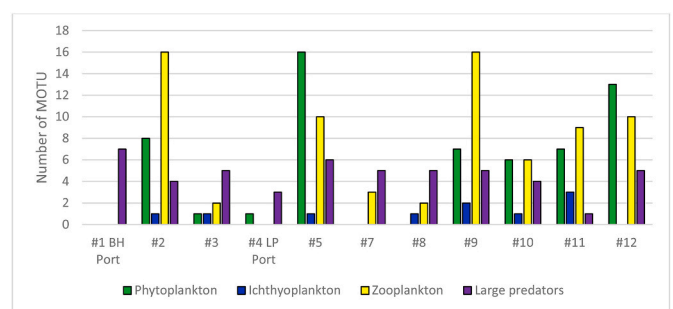


Fig. 3. Number of MOTUs (= putative genera) of each functional group found in the 12 sampling locations analysed across a latitudinal gradient in the East Atlantic Ocean. BH and LP represent Bremerhaven and Las Palmas ports, respectively.

Table 1

Indices calculated for each sampling location based on eDNA metabarcoding. Number of MOTUs are shown in parenthesis after each sample code. BH and LP are Bremerhaven and Las Palmas ports, respectively. Plankton indices as in Moncheva et al. (2002), Diatoms/Dinoflagellates being Bacillariophyceae/Dinophyceae = Phytoplankton taxon dominance.

Sample	Grazing pressure			Taxonomic diversity	Diversity based on functional groups
	Phytoplankton/Zooplankton	Diatoms/Dinoflagellates	Diatoms/Copepods		
#1 BH Port (7)	–	–	–	1.95	0
#2 (29)	0.5	6	0.46	3.37	1.07
#3 (9)	0.5	0	0	2.20	1.15
#4 LP Port (4)	–	0	–	1.39	0.56
#5 (33)	1.6	4	0.8	3.50	1.13
#7 (8)	0	–	0	2.08	0.66
#8 (8)	0	–	0	2.08	0.90
#9 (30)	0.44	2.5	0.45	3.40	1.15
#10 (17)	1	–	1	2.83	1.24
#11 (20)	0.78	1.33	0.5	3.00	1.16
#12 (28)	1.3	5	1.25	3.33	1.03

Table 2

Genera identified as preys of *Merluccius* (hake) and/or *Thunnus* (tuna) species from relevant literature. Presence/absence in each sampling site as 1/0. Port samples are excluded.

Prey type	Order	Genus	In the diet of	Sampling site									
				#2	#3	#5	#7	#8	#9	#10	#11	#12	
Invertebrates	Amphipoda	<i>Sarothrogammarus</i>	Hake, tuna	0	0	0	0	0	0	0	0	0	1
	Decapoda	<i>Cambarellus</i>	Hake, tuna	0	0	0	1	0	0	0	0	0	0
	Decapoda	<i>Pugettia</i>	Hake, tuna	0	0	0	0	0	0	0	0	0	1
	Euphausiacea	<i>Euphausia</i>	Hake, tuna	0	0	0	0	0	1	0	0	0	0
	Euphausiacea	<i>Thysanoessa</i>	Hake, tuna	1	0	0	0	0	1	1	0	0	0
	Euphausiacea	<i>Thysanopoda</i>	Hake, tuna	1	0	0	0	0	1	0	0	0	0
	Isopoda	<i>Ligia</i>	Tuna	0	0	0	1	0	1	0	0	0	0
Fish	Blenniiformes	<i>Salarias</i>	Hake, tuna	0	0	0	0	0	0	0	0	1	0
	Clupeiformes	<i>Engraulis</i>	Hake, tuna	0	0	0	0	1	1	0	0	0	0
	Gobiiformes	<i>Eleotris</i>	Hake, tuna	0	1	0	0	0	0	0	0	0	0
	Pleuronectiformes	<i>Solea</i>	Tuna	1	0	1	0	0	1	1	1	1	0
	Pleuronectiformes	<i>Cynoglossus</i>	Tuna	0	0	0	0	0	0	0	0	1	0
		Hake prey richness		2	1	0	1	1	4	1	1	1	2
		Tuna prey richness		3	1	1	2	1	6	2	3	2	

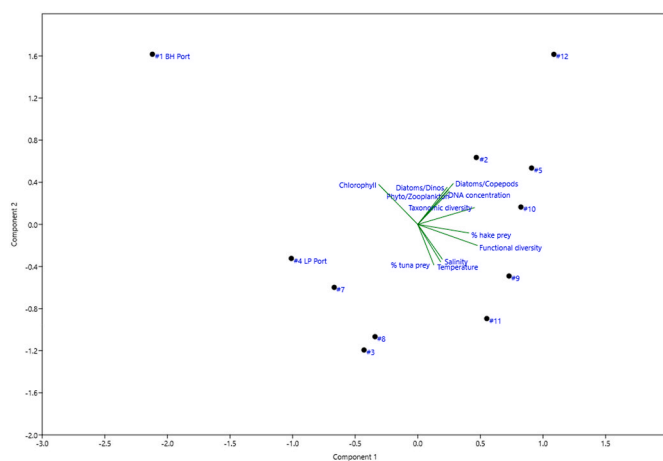


Fig. 4. Eigenvalue-scaled scatter plot of the Principal Component Analysis constructed from the indices considered. The length of diagonals is proportional to the weight of each factor in the PCA. BH and LP are Bremerhaven and Las Palmas ports, respectively. Diversity of functional groups is abbreviated here as “Functional diversity”.

indices, closer to that calculated from functional groups than to the taxonomic diversity. The proportion of tuna preys was located also in the lower part of the scatter plot near temperature and not very far from the diagonal representing hake preys, which could be expected

considering that the two fish are carnivorous and share crustaceans and fish in their diet. The locations with low phytoplankton richness (the two ports, #3, #6, #7, #8) were in the quadrants opposite to the locations richer in phytoplankton (Fig. 4). The factor Chlorophyll was relatively separated of the rest of factors in this scatter plot, but closer to the phytoplankton-based indices and the DNA concentration than to the diversity indices, see Fig. 4.

Pairwise correlations between the considered variables are presented in Supplementary Table 5. Correlations were generally in accordance with the PCA results. Regarding the physical variables (displayed in Supplementary Table 1), temperature (higher in the tropical zone where #7 to #9 sampling points are located) was highly and negatively correlated with chlorophyll: $r = -0.788$, 9 d.f., $p = 0.004$. Chlorophyll was below detection levels in those locations over the sampling days in November (see Supplementary table 1). This is normal because upper chlorophyll levels generally decline with increased temperature in tropical oceans (Feng et al., 2015), and low chlorophyll in the study area in November is common, from the consistent seasonal variation of chlorophyll-a occurring in tropical East Atlantic Ocean (e.g., Dunstan et al., 2018). Another interesting result was the highly significant negative correlation between salinity and the dominance index Diatoms/Dinoflagellates ($r = -0.882$, $p = 0.009$), which is consistent with the known decrease of the density of large diatoms under low salinity conditions while the effect is much smaller on phytoflagellates (e.g., Hernando et al., 2015).

Opposite to expectations, no significant correlation was found between the DNA concentration and the chlorophyll concentration in the analysed samples ($r = -0.035$, d.f. = 9, $p = 0.92$ n.s.). However, there

was a significant and positive correlation between eDNA concentration and the grazing index Diatoms/Copepods ($r = 0.822$, $p = 0.007$) (Supplementary Table 5). This result reveals the importance of diatoms in the community detected from eDNA, and would connect, although indirectly, primary productivity measures with eDNA. Logically, chlorophyll measures were positively correlated with the grazing index Phytoplankton/Zooplankton ($r = 0.671$, $p = 0.04$). The proportion of tuna preys was not significantly correlated with any other variable, and the proportion of hake prey was positively correlated with the two measures of diversity: $r = 0.79$ with $p = 0.004$ with taxonomic diversity, and $r = 0.73$ with $p = 0.03$ with functional diversity (Supplementary Table 5).

Only independent planktonic indices not significantly correlated to each other were employed for multivariate multiple regression analysis. For this reason, we excluded taxonomic diversity (correlated with the grazing indices and the diversity of functional groups), and the grazing index Phytoplankton/Zooplankton that was significantly correlated with the other grazing index Diatoms/Copepods (Supplementary Table 5). The analysis run with hake and tuna preys as dependent and the remaining indices as independent variables had a significant overall MANOVA, with $F = 3.921$, $df_1 = 6$, $df_2 = 12$, Wilks' $\lambda = 0.114$ and $p = 0.021$. The multiple regression on the dependent variables (Table 3) showed that the proportion of tuna preys was significantly predicted by the diversity based on functional groups ($t = 2.53$, $p = 0.039$). The proportion of hake preys was significantly predicted by the dominance index Diatoms/Dinoflagellates ($t = 2.54$, $p = 0.038$, $r^2 = 0.26$), that is, a phytoplankton richer in diatoms would predict hake preys. It was also significantly predicted by the functional groups diversity index with $t = 3.676$ and $p = 0.008$, $r^2 = 0.52$ (Table 3), this meaning that both tuna and hake preys would be more abundant in functionally richer communities.

4. Discussion

4.1. Insights into eDNA-based analysis of plankton-based indices

This study provides eDNA signals that support the importance of some phytoplankton elements, such as diatoms, in the richness of fish' preys. Here, the proportion of diatoms (over dinoflagellates) significantly predicted the richness of hake preys, supporting our Hypothesis (ii), at least for hakes. eDNA data would support the importance of diatoms for zooplankton grazing, as proposed by Benedetti et al. (2019) or Fernandes et al. (2019). Unlike the results by Sachoemar (2015), where chlorophyll-a was an indicator of fish productivity, in this study satellite-based data of chlorophyll were not directly related with the

Table 3

Multiple regression analysis showing the predictive value of independent variables for the proportion of hake and tuna preys. Regression coefficients and statistics are presented. SE, standard error. Significant p-values are marked in italics.

		β coefficient	SE	t	p	R ²
% tuna prey	Constant	0.014	0.039	0.346	0.739	
	Diatoms/ Dinoflagellates	-0.009	0.008	-1.051	0.328	0.131
	Diatoms/ Copepods	-0.051	0.038	-1.329	0.225	0.134
	Functional groups diversity	0.096	0.038	2.533	<i>0.039</i>	0.27
% hake prey	Constant	-0.089	0.099	-0.899	0.398	
	Diatoms/ Dinoflagellates	0.053	0.021	2.539	<i>0.038</i>	0.259
	Diatoms/ Copepods	-0.13	0.096	-1.351	0.219	0.01
	Functional groups diversity	0.349	0.095	3.676	<i>0.008</i>	0.524

richness of hake and tuna preys. Remote-sensed patterns of primary productivity have some limitations (Sigman and Hain, 2012; Chen et al., 2013; Babin et al., 2015); satellite-based productivity patterns are not sufficient to interpret the foraging behaviour of marine predators, as seen in the Benguela upwelling zone (Grémillet et al., 2008).

The results obtained from eDNA are generally according to known characteristics of the species detected and oceanographic features. The negative effect of low salinity on the density (in this case the relative abundance) of large diatoms described for example in Antarctica (Herando et al., 2015) was confirmed here from a negative correlation between the dominance index Diatoms/Dinoflagellates and the salinity. This would support expected deep changes in phytoplankton communities along current climate change, with a decrease of diatom diversity following an increase of temperature and the subsequent decrease of salinity for ice sheet melting (Sugie et al., 2020).

Two areas rich in hake and tuna prey were found. One in the sub-equatorial zone around #9 (Fig. 1) is influenced by the cold Benguela current flowing northwards; for its upwelling it is rich in phytoplankton and ichthyoplankton (e.g., Hutchings et al., 2006), like *Engraulis* which is a prey of both hake and tuna. DNA of African sardines (*Sardinops sagax*) that are important prey of hakes and tuna were not found there, which is not surprising because the species collapsed due to overfishing and environmental change at the beginning of this century, with a cascade of changes in its predators (Erasmus et al., 2021). The zone of site #2, in south Portugal waters, contains rich fish assemblages that include the hake *M. merluccius* and its prey in both shallow and deep waters (e.g., Gomes et al., 2001). Prediction of hake and tuna prey from plankton diversity is logical, because diversity based on the functional groups ensures availability of food for such preys, especially for opportunistic feeders like *Solea* (Fanelli et al., 2022) here detected from eDNA. On the other hand, the predictive value of the dominance index Diatoms/Dinoflagellates for the proportion of hake preys can be also expected; for example, *Euphausia* and other Euphausiacea (krill) that are hake preys feed primarily on diatoms (e.g., Cavan et al., 2019), thus the dominance of diatoms would favour these species.

4.2. Environmental DNA for the inventory of planktonic communities

For the application of eDNA method of community inventory our results would agree with previous studies, at least partially. As in other studies, metabarcoding provides a complete view of the planktonic community including primary producers, zooplankton, ichthyoplankton and large predators (e.g., Ardura et al., 2021; Chavez et al., 2021). However, a direct correlation between chlorophyll-a and eDNA quantity (Chavez et al., 2021), expected in our Hypothesis (i), was not found. Instead, the importance of phytoplankton in the eDNA amount –which is essential for the application of eDNA-based community inventories– was here revealed from the significant correlation between the grazing index Phytoplankton/Zooplankton and the concentration of eDNA. Another interesting result was the significant correlation between the satellite-measured chlorophyll and the relative abundance of diatoms. The chlorophyll and the phytoplankton biomass can be differentially related depending on the type and composition of phytoplankton (Geider et al., 1997; Johnson et al., 2010), and in this case diatoms would be especially important.

Supporting our expectations in Hypothesis iii) based on Wasmund et al. (2017), diatom-based indices correlated positively and significantly with MOTU-based Shannon's diversity index. At the same time, the diversity based on functional groups significantly predicted the richness of hake and tuna preys. Altogether, the significant relationships between functional components of the plankton found in our study based on eDNA are consistent with the importance to diatoms as a key phytoplankton component indicator of a good environmental status (Wasmund et al., 2017). Despite being obtained from a relatively small number of locations at a large spatial dimension, these results support the validity of eDNA metabarcoding as a source of data with predictive

value in fisheries sciences.

This study indeed confirms the utility of eDNA for plankton inventories (Chavez et al., 2021), and opens a new application in the calculation of diatom-based indices where diatom richness is the main variable. Environmental DNA has been applied in phytoplankton studies with different purposes, like the identification of harmful algae (Ardura et al., 2020) and invasive species (Zaiko et al., 2020). Extending its application to the calculation of indices based on relative richness of plankton elements, as we did in this study, will contribute to a better understanding of ecosystem functioning, especially in areas difficult to sample like in the open ocean. Moreover, it will help to give scientific advice to sustainable fisheries.

It must be noted that eDNA of *Merluccius* and *Thunnus* species was not found in this survey. Not finding hake larvae in the majority of sampling points is not strange because many sampling stations are far off the continental shelves, while hake larvae are generally found on the continental shelf or at the shelf break (García-Fernández et al., 2021). However, older individuals –especially pelagic tuna– could be there. A failure of the primers to amplify these fish is not a likely explanation (although cannot be totally ruled out), because we found sequences of other fish and of large predators – that logically correspond to floating eDNA. Bluefin tuna adults undertake long displacements between foraging and spawning areas and could be around the northern sampling points at the sampling time, but they exhibit a diel pattern of depth use, occupying shallower depths at night and deeper depths in the day (Horton et al., 2020). Since the samples were taken during the day, eDNA from adult Bluefin tuna would not be expected at the sampling depth of 6m. On the other hand, foraging grounds of young and adolescent *Thunnus thunnus* aggregations are mainly visited in spring and summer in East Atlantic waters (Logan et al., 2011), and the same happens for *T. alalunga* that is in turn associated with chlorophyll-a distribution (e.g., Sagarmínaga and Aguirrezabalaga, 2014); thus, November would not be a good timing for the detection of young tunas in the studied zone.

4.3. Limitations of this study

This study has some limitations. One is that, likely for low eDNA concentration in open ocean waters, the replicates had to be merged for obtaining sufficient eDNA for metabarcoding. This could be solved sampling larger water volumes (50L or higher volumes), although filtering high volumes can be challenging to afford.

On the other hand, the study could be enriched taking samples at different depths; here all the samples were taken at the same depth, 6m below sea level. Surface sampling would likely allow to detect other community components, principally phytoplankton. Another improvement could be additional sampling at night, not only during the day. Nocturnal sampling would allow to detect organisms with diel vertical migration, like many copepods and squid (e.g., Rosa and Seibel, 2010; Holliland et al., 2012), and even tunas that prefer shallower waters at night (Horton et al., 2020).

Finally, samples were obtained in North Hemisphere winter and South Hemisphere summer. In temperate regions a seasonal variation in plankton composition is expected (e.g., Benedetti et al., 2019); but the results obtained in this study correspond to a single season in each hemisphere. Seasonal sampling in the future would enrich the results, reflecting better the plankton communities and consequently the fish species that feed on those plankton in different stages of their life cycle. Altogether, these improvements could provide a much better view of the plankton diversity in the study area.

5. Conclusions

In this study plankton communities were inventoried from environmental DNA, and planktonic indices calculated.

The proportion of prey of hake and tuna was significantly predicted

from planktonic diversity, and hake prey from the relative abundance of diatoms that, although not direct prey of hake, are grazed by hake prey.

Significant negative correlation between an abiotic factor, salinity, and an index calculated from eDNA, Diatoms/Dinoflagellates, would confirm the potential impact of current global change on planktonic communities.

These results support the utility of eDNA for the inventory of plankton communities and scientific advice for fisheries sustainability.

CRediT authorship contribution statement

Sara Fernandez: Data curation, Investigation, Methodology, Software, Writing – review & editing. **Alba Ardura:** Investigation, Methodology, Visualization, Writing – review & editing. **Jose L. Martinez:** Investigation, Methodology, Software. **Johannes Rick:** Investigation, Methodology, Validation. **Gonzalo Machado-Schiaffino:** Investigation, Writing – review & editing. **Eva Garcia-Vazquez:** Conceptualization, Investigation, Supervision, Writing – original draft.

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.

Data availability

Data will be made available on request.

Acknowledgments

This study was funded from the Spanish Ministry of Science and Innovation Grant GLOBALHAKE PID2019-108347RB-I00, and from the Government of Asturias, Grant AYUD/2021/50967. The work on board RV Polarstern was funded from the Alfred Wegener Institute Grant No. AWI_PS116_00. We are grateful to Polarstern crew for their help with sampling and tasks on board. Sara Fernandez holds Margarita Salas Grant AYUD/2021/58385.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.marenvres.2023.106312>.

References

- Alheit, J., Pitcher, T.J., 1995. Hake. Fisheries, Ecology and Markets, Fish and Fisheries Series 15. Chapman & Hall, London, p. 478.
- Ardura, A., Borrell, Y.J., Fernandez, S., Gonzalez-Arenales, M., Martinez, J.L., Garcia-Vazquez, E., 2020. Nuisance algae in ballast water facing international conventions. Insights from DNA metabarcoding in ships arriving in bay of biscay. Water 12, 2168. <https://doi.org/10.3390/w12082168>.
- Ardura, A., Rick, J., Martinez, J.L., Zaiko, A., Garcia-Vazquez, E., 2021. Stress resistance for unraveling potential biopollutants. Insights from ballast water community analysis through DNA. Mar. Pollut. Bull. 163, 111935 <https://doi.org/10.1016/j.marpolbul.2020.111935>.
- Babin, M., Bélanger, S., Ellingsen, I., Forest, A., Le Fouest, V., Lacour, T., Ardyna, M., Slagstad, D., 2015. Estimation of primary production in the Arctic Ocean using ocean colour remote sensing and coupled physical–biological models: strengths, limitations and how they compare. Prog. Oceanogr. 197–220. <https://doi.org/10.1016/j.pcean.2015.08.008>.
- Baker, C., 2017. Workflow for Generating a Qiime-Compatible Blast Database from an Entrez Search, pp. 1–4. Issue February.
- Benedetti, F., Jalabert, L., Sourisseau, M., Becker, B., Cailliau, C., Desnos, C., Elineau, A., Irisson, J.-O., Lombard, F., Picheral, M., Stemmann, L., Poulain, P., 2019. The seasonal and inter-annual fluctuations of plankton abundance and community structure in a north atlantic marine protected area. Front. Mar. Sci. 6, 214. <https://doi.org/10.3389/fmars.2019.00214>.
- Bjelland, R.M., Skiftesvik, A.B., 2006. Larval development in European hake (*Merluccius merluccius* L.) reared in a semi-intensive culture system. Aquacult. Res. 37, 1117–1129. <https://doi.org/10.1111/j.1365-2109.2006.01536.x>.

- Moncheva, S., Dontcheva, V., Shtereva, G., Kamburska, L., Malej, A., Gorinstein, S., 2002. Application of eutrophication indices for assessment of the Bulgarian Black Sea coastal ecosystem ecological quality. *Water Sci. Technol.* 46 (8), 19–28. <https://doi.org/10.2166/wst.2002.0136>.
- Nguyen, B.N., Shen, E.W., Seemann, J., et al., 2020. Environmental DNA survey captures patterns of fish and invertebrate diversity across a tropical seascape. *Sci. Rep.* 10, 6729. <https://doi.org/10.1038/s41598-020-63565-9>.
- Olafsdottir, D., MacKenzie, B.R., Chosson, P.V., Ingimundardottir, T., 2016. Dietary evidence of mesopelagic and pelagic foraging by atlantic bluefin tuna (*Thunnus thynnus* L.) during autumn migrations to the Iceland basin. *Front. Mar. Sci.* 3, 108. <https://doi.org/10.3389/fmars.2016.00108>.
- Ortiz de Zárate, V., 1987. Datos sobre la alimentación del Atún blanco (*Thunnus alalunga*) juvenil capturado en el Golfo de Vizcaya. *Col. Vol. Sci. Pap. ICCAT* 26 (2), 243–247.
- Pillar, S.C., Wilkinson, I.S., 1995. The diet of Cape hake *Merluccius capensis* on the south coast of South Africa. *S. Afr. J. Mar. Sci.* 15 (1), 225–239. <https://doi.org/10.2989/02577619509504845>.
- Pomerleau, C., Sastri, A.R., Beisner, B.E., 2015. Evaluation of functional trait diversity for marine zooplankton communities in the Northeast subarctic Pacific Ocean. *J. Plankton Res.* 37 (4), 712–726. <https://doi.org/10.1093/plankt/fbv045>.
- Pusineri, C., Vasseur, Y., Hassani, S., Meynier, L., Spitz, J., Ridoux, V., 2005. Food and feeding ecology of juvenile albacore, *Thunnus alalunga*, off the Bay of Biscay: a case study. *ICES (Int. Counc. Explor. Sea) J. Mar. Sci.* 62 (1), 116–122. <https://doi.org/10.1016/j.icesjms.2004.09.004>.
- Roel, B.A., Macpherson, E., 1988. Feeding of *Merluccius capensis* and *M. paradoxus* off Namibia. *S. Afr. J. Mar. Sci.* 6 (1), 227–243. <https://doi.org/10.2989/025776188784480546>.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584. <https://doi.org/10.7717/peerj.2584>.
- Rosa, R., Seibel, B.A., 2010. Metabolic physiology of the Humboldt squid, *Dosidicus gigas*: implications for vertical migration in a pronounced oxygen minimum zone. *Prog. Oceanogr.* 86 (1–2), 72–80. <https://doi.org/10.1016/j.pocean.2010.04.004>.
- Rudershausen, P.J., Buckel, J.A., Edwards, J., 2010. Feeding ecology of blue marlins, dolphinfish, yellowfin tuna, and wahoos from the North Atlantic Ocean and comparisons with other oceans. *Trans. Am. Fish. Soc.* 139, 1335–1359. <https://doi.org/10.1577/T09-105.1>.
- Sachoemar, S., 2015. Variability of sea-surface chlorophyll-a, temperature and fish catch within Indonesian region revealed by satellite data. *Mar. Res. Indones.* 37 (2), 75–87. <https://doi.org/10.14203/mri.v37i2.25>.
- Sagarminaga, Y., Arribabalaga, H., 2014. Relationship of Northeast Atlantic albacore juveniles with surface thermal and chlorophyll-a fronts. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 107, 54–63. <https://doi.org/10.1016/j.dsr2.2013.11.006>.
- Sigman, D.M., Hain, M.P., 2012. The biological productivity of the ocean. *Nature Education Knowledge* 3 (10), 21.
- Sugie, K., Fujiwara, A., Nishino, S., Kameyama, S., Harada, N., 2020. Impacts of temperature, CO₂, and salinity on phytoplankton community composition in the Western Arctic Ocean. *Front. Mar. Sci.* 6, 821. <https://doi.org/10.3389/fmars.2019.00821>.
- Sumaila, U.R., Bellmann, C., Tipping, A., 2016. Fishing for the future: an overview of challenges and opportunities. *Mar. Pol.* 69, 173–180. <https://doi.org/10.1016/j.marpol.2016.01.003>.
- Valls, M., Saber, S., Gómez, M.J., Reglero, P., Mele, S., 2022. Diet and feeding behaviour of albacore *Thunnus alalunga* in the western Mediterranean Sea during the spawning period. *J. Fish. Biol.* 100 (1), 203–217. <https://doi.org/10.1111/jfb.14935>.
- Wasmund, N., Kownacka, J., Göbel, J., Jaanus, A., Johansen, M., Jurgensone, I., Lehtinen, S., Powilleit, M., 2017. The diatom/dinoflagellate index as an indicator of ecosystem changes in the baltic sea 1. Principle and handling instruction. *Front. Mar. Sci.* 4, 22. <https://doi.org/10.3389/fmars.2017.00022>.
- Weithoff, G., Beisner, B.E., 2019. Measures and approaches in trait-based phytoplankton community ecology – from freshwater to marine Ecosystems. *Front. Mar. Sci.* 6, 40. <https://doi.org/10.3389/fmars.2019.00040>.
- Wilhelm, et al., 2015. Biology and fisheries of the shallow-water hake (*Merluccius capensis*) and the deep-water hake (*M. paradoxus*) in Namibia. In: Arancibia, H. (Ed.), *Hakes: Biology and Exploitation*. John Wiley & Sons, Ltd, Chichester, UK, pp. 70–100. <https://doi.org/10.1002/9781118568262.ch> (Chapter 3).
- Zaiko, A., Wood, S., Pochon, X., Biessy, L., Laroche, O., Croot, P.L., Garcia-Vazquez, E., 2020. Elucidating biodiversity shifts in ballast water tanks during a cross-latitude transfer: complementary insights from molecular analyses. *Environ. Sci. Technol.* 54 (13), 8443–8454. <https://doi.org/10.1021/acs.est.0c01931>.
- Zhan, A., Hulák, M., Sylvester, F., Huang, X., Adebayo, A.A., Abbott, C.L., Adamowicz, S. J., Heath, D.D., Cristescu, M.E., Macisaac, H.J., 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods Ecol. Evol.* 4 (6), 558–565. <https://doi.org/10.1111/2041-210X.12037>.