Received: 27 March 2023 | Revised: 17 October 2023 | Accepted: 21 November 2023

DOI: 10.1111/1755-0998.13907

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Morphological and taxonomic diversity of mesozooplankton is an important driver of carbon export fluxes in the ocean

Margaux Perhirin[1] | Hannah Gossner[2] | Jessica Godfrey[2] | Rodney Johnson[2] | Leocadio Blanco-Bercial[2] | Sakina-Dorothée Ayata[1,3]

[1]Sorbonne Université, UMR 7159 CNRS-IRD-MNHN, LOCEAN-IPSL, Paris, France

[2]Bermuda Institute of Ocean Sciences, Arizona State University, St. Georges, Bermuda

[3]Institut universitaire de France (IUF), Paris, France

Correspondence
Margaux Perhirin, Sorbonne Université, UMR 7159 CNRS-IRD-MNHN, LOCEAN-IPSL, Paris, France.
Email: margaux.perhirin@locean.ipsl.fr

Funding information
Agence Nationale de la Recherche; Institut des Sciences du Calcul et des Données; National Science Foundation, Grant/Award Number: NSF 1522206, NSF 1756105 and NSF 1829318; Simons Foundation

Handling Editor: Simon Creer

Abstract

Mesozooplankton is a very diverse group of small animals ranging in size from 0.2 to 20 mm not able to swim against ocean currents. It is a key component of pelagic ecosystems through its roles in the trophic networks and the biological carbon pump. Traditionally studied through microscopes, recent methods have been however developed to rapidly acquire large amounts of data (morphological, molecular) at the individual scale, making it possible to study mesozooplankton using a trait-based approach. Here, combining quantitative imaging with metabarcoding time-series data obtained in the Sargasso Sea at the Bermuda Atlantic Time-series Study (BATS) site, we showed that organisms' transparency might be an important trait to also consider regarding mesozooplankton impact on carbon export, contrary to the common assumption that just size is the master trait directing most mesozooplankton-linked processes. Three distinct communities were defined based on taxonomic composition, and succeeded one another throughout the study period, with changing levels of transparency among the community. A co-occurrences' network was built from metabarcoding data revealing six groups of taxa. These were related to changes in the functioning of the ecosystem and/or in the community's morphology. The importance of Diel Vertical Migration at BATS was confirmed by the existence of a group made of taxa known to be strong migrators. Finally, we assessed if metabarcoding can provide a quantitative approach to biomass and/or abundance of certain taxa. Knowing more about mesozooplankton diversity and its impact on ecosystem functioning would allow to better represent them in biogeochemical models.

KEYWORDS
carbon export, metabarcoding, quantitative imaging, trait-based approaches, transparency, zooplankton

1 | INTRODUCTION

Mesozooplankton is a group comprised of small metazoans of 0.2–20 mm in size drifting with ocean currents. They constitute a critical link between primary producers and higher trophic levels and play a major role in the biological carbon pump. They contribute directly to particulate organic carbon (POC) export through the production of sinking faecal pellets, moults and carcasses (Lomas et al., 2010; Steinberg & Landry, 2017). Species in this group can also repackage particles into larger and faster-sinking faecal pellets or can fragment sinking POC, particularly by sloppy feeding, into smaller and slower-sinking particles. Additionally, Diel Vertical Migrations (DVM) realized by some mesozooplankton species between the surface and the deep ocean increase the export of dissolved carbon and the remineralisation of nutrients at depth (Kelly et al., 2019; Tarrant et al., 2021). Hence, mesozooplankton is a highly diverse group in terms of species, ecological strategies, size and among others (Litchman et al., 2013). Its biomass and community composition can impact both the productivity and the biogeochemistry of the environment (Schnetzer & Steinberg, 2002). Better understanding of how variations in mesozooplankton diversity impact ecosystem functionalities will be a fundamental step to improve the general knowledge of ocean ecosystems. This would also contribute to improving the representation of mesozooplankton diversity in biogeochemical models since this key component of the marine ecosystem is usually represented by only one variable (Quéré et al., 2005).

The Sargasso Sea has been identified as one of the highest zooplanktonic diversity locations in the world (Rombouts et al., 2010). The Bermuda Atlantic Time-series Study (BATS, Lomas et al., 2013) site is located in the vicinity of Bermuda Islands (31°40′ N, 64°10′ W) in the North Atlantic subtropical oligotrophic gyre and has been studied since 1988. Long-term sampling and research have shown that environmental variables and zooplanktonic biomass follow a classical seasonal cycle. The phytoplankton bloom, stimulated by optimal nutrients and light conditions, is followed by a peak in zooplankton biomass, usually in March or April. In summer, thermal stratification is strong and leads to nutrient depleted upper waters and thus, a decrease in zooplankton biomass (Deevey, 1971; Madin et al., 2001). Secondary peaks can occur in late summer or early fall, when the mixed layer deepens again. Zooplankton biomass is minimal during winter (Ivory et al., 2019) when nutrient flux fertilizes the surface layer thanks to deep mixing and mode water formation (Steinberg et al., 2001). Distinct taxa assemblages, based exclusively on the community composition, are associated with each of the four seasonal hydrographic periods (Blanco-Bercial, 2020).

Historically, zooplankton have been investigated by collecting samples using plankton nets and identifying specimens at the species (or genus) level under a microscope. However, this technique of visual identification is time-consuming and complex, partly due to the presence of cryptic species, species with multiple life stages or with distinct sexual dimorphism (Huo et al., 2020; Rey et al., 2020). Moreover, taxonomic identification can be strongly biased by expert's specialization (Harvey et al., 2017) and more generally, biased towards the largest and most known plankton species. Molecular studies have revealed that plankton knowledge is biased towards sampled or cultivable species, representing less than 30% of the total estimated plankton diversity, leading to a 'Mare Incognitum' (Chust et al., 2017).

Next-generation sequencing technologies, such as DNA metabarcoding, could therefore be an alternative to overcome these microscopy-based monitoring limitations (Abad et al., 2016). DNA metabarcoding allows for the identification of taxa by identifying specific genes (Hebert et al., 2003) from a bulk sample containing DNA from dozens of taxa. This inexpensive and rapid method can detect rare and cryptic species (Huo et al., 2020), and be used to explore the diversity and structure of marine eukaryotic communities (Chen et al., 2021) as well as their ecosystem dynamics (Matthews et al., 2021). However, it is not free of several technical biases: barcode and primers selection, DNA amplification's efficiency, bioinformatics pipelines' parameters choice, or reference database quality are known sources of biases (Harvey et al., 2017). This method produces compositional data, meaning that values inferred from counts of reads for any taxon will be influenced by the ones from other taxa (Brisbin et al., 2020; MacNeil et al., 2022). Hence, the quantitative property of these data is still controversial and hard to evaluate (Ershova et al., 2021), necessitating verification by imaging methods (Harvey et al., 2017).

Over the past few decades, imaging systems (e.g., Underwater Vision Profiler, Picheral et al., 2010; In Situ Ichthyoplankton Imaging System, Cowen & Guigand, 2008 or Zooscan, Gorsky et al., 2010) have been developed, along with machine learning algorithms, to acquire and analyse large numbers of individual organisms' images in a quantitative way (Irisson et al., 2022). These systems are less or non (if in situ) invasive compared with classical sampling methods (Martini et al., 2021), allowing for relatively rapid acquisition of a large quantity of data, and offering new opportunities to study some previously neglected organisms (Biard et al., 2016).

These images give consistent-rich morphological information of individual organisms, making it possible to use a trait-based approach and thus a more ecosystem-focussed point of view than a species or taxonomic one (Martini et al., 2021). Trait-based approaches refer to functional traits, that is, any individual's characteristic (morphological, physiological) impacting its fitness (Violle et al., 2007), and are based on the assumption that the fitness of an individual is based on its success regarding feeding, reproduction and survival which would depend on a few key traits (Brun et al., 2017). Size is easily measured, and due to its influence on many physiological and ecological aspects of an organism's life, it has been defined, notably in zooplankton, as a 'master trait' (Barton et al., 2013; Litchman et al., 2013). For example, larger organisms will produce larger faecal pellets (Uye & Kaname, 1994) and/or have a stronger migration behaviour, hence playing an important role in export flux out of the euphotic zone (Stamieszkin et al., 2015; Steinberg et al., 2001). Organisms' transparency, although an important zooplankton morphological characteristic (Orenstein et al., 2022), is however understudied (Johnsen &

Widder, 1998). Quantifying transparency might give information about the organism's ecology (reproduction with gonads or eggs, feeding with guts content, Orenstein et al., 2022). Transparency and colour can also be related to physiological functions (Vilgrain et al., 2022), for example, carotenoids effects against oxidative stress (Brüsin et al., 2016), or possibly the utilization of pigmentation and DVM in order to escape visual predators (Hays et al., 1994).

While both metabarcoding and imaging techniques can provide information on the temporal and spatial distribution of zooplankton communities and their environmental dynamics (Vilgrain et al., 2021), imaging data is also able to quantify taxa's abundance and biomass (e.g., Monferrer et al., 2022). Hence, it would be ideal to combine high throughput image acquisition methods with metabarcoding to improve zooplanktonic community studies (Bucklin et al., 2019). Each of these approaches has uncertainties (Monferrer et al., 2022) but when combined, many resolution issues are resolved. For example, metabarcoding fills in some of the limited size spectrum resolution of imaging techniques (MacNeil et al., 2022) and allows a joint study of zooplankton communities taxonomic and functional aspects. Previous studies have tried to assess the quantitative property of metabarcoding by comparing counts of sequence reads to (relative) biomass (e.g., Djurhuus et al., 2018; Durkin et al., 2022 on particles) or to (relative) abundance (e.g., Ibarbalz et al., 2019; Rey et al., 2020) for various zooplankton taxonomic groups but results from each approach seldom agree (Harvey et al., 2017), requiring further studies.

This study, therefore, seeks to use the complimentary methods of metabarcoding and high throughput imaging in order to link variations in mesozooplankton community composition throughout the year to morphological changes in the highly diverse ecosystem of the Sargasso Sea. Additionally, we evaluate how modifications in community composition will impact the ecosystem's export intensity. Finally, utilizing the potential of our 'twin' morphological and molecular data sets, we aim to test the correlations between abundance and biomass, and relative read counts, for wide taxonomic categories.

2 | MATERIALS AND METHODS

2.1 | Environmental data acquisition

Data were collected at BATS site (31°40′ N; 64°10′ W) on a monthly basis, from March 2016 to May 2017, except in March 2017 due to rough weather. Environmental samples were taken during the day, except in February 2017 when they were taken at night. Environmental data were collected vertically using a SeaBird 911 CTD rosette equipped with temperature, conductivity, fluorescence, oxygen and photosynthetic active radiation (PAR) sensors. Density (σ_θ) was calculated from temperature and conductivity. Depth of the Deep Chlorophyll Maximum (DCM, in metre) was defined as the maximum of the fluorescence profile. Mixed Layer Depth (MLD, in metre) was calculated as the depth for which σ_θ is larger or equal to $\sigma_{\theta surface} + 0.125\,kg\,m^{-3}$ (Levitus, 1982). Sediment traps were allowed to drift for 72 h at 200 m (Steinberg et al., 2001). The total organic carbon (TOC) measured at 200 m (in mg m^{-2} d^{-1}) was used as a proxy of carbon export to the mesopelagic layer. The total organic carbon data were averaged from three replicates and blank corrected, following BATS protocol (Steinberg et al., 2001).

2.2 | Zooplankton data acquisition

Zooplankton sampling was conducted by employing a 1 m^2 rectangular net with 202 µm mesh, towed at a speed of about 1 knot, paying out 250 m of wire at 15 m per minute both ways (maximum depth sampled: 168.68 m ± 37.60 m, Table S1). Three replicate day and night tows were made for each cruise, except in October 2016 for which the day sample for molecular analysis was not obtained due to rough weather. Samples from two tows were preserved in formaldehyde at 4% for Zooscan analyses (see following sections). Zooplankton from the last net was fixed in 95% undenatured ethanol then stored at −20°C until returning to BIOS (1–3 days later) for molecular analyses. Ethanol was changed right after arrival to BIOS.

2.3 | Zooscan process and analyses

2.3.1 | Zooscan process

A fraction of each half-split sample from two net replicates (obtained with a Motoda splitter; Motoda, 1959) was scanned using the Zooscan system (Gorsky et al., 2010). All individuals larger than 2 cm were selected by eye and scanned separately. Remaining samples were sieved through a 1000 µm sieve and at least 1500 particles from two size classes (202–1000 µm and >1000 µm) were scanned. Morphological variables (as in Vilgrain et al., 2021, Table S2) were calculated for each individual image of zooplankton organism using the ZooProcess software (Gorsky et al., 2010). Individual images, their metadata, and their morphological descriptors were stored in the EcoTaxa web application (https://ecotaxa.obs-vlfr.fr/ Picheral et al., 2017); project 'BATS_timeseries [4236]'). Automatic taxonomic classification was done on these images at a relatively high level using a random forest algorithm (Irisson et al., 2022; 42 categories, Table S3). Taxonomic identification of each individual image was then validated by a human expert. In total, the data set gathered 133,389 images. Only the images from 29 taxonomic categories (in bold in Table S3) and with a major axis length higher than 500 µm were studied, to focus on mesozooplankton. Moreover, Actinopterygii images with a major axis length longer than 2 cm were removed to consider only larvae that cannot swim against the flow, leading to a total of 69,949 mesozooplankton images. "Oikopleura" category was renamed Larvacea to take into account all the Appendicularia's orders but not be mistaken with the genus name *Appendicularia*. For each

image, a density factor d was computed to account for the water volume sampled as follows:

$$d = \frac{\text{Motoda fraction scanned}}{\text{Total volume sampled}}.$$

2.3.2 | Biovolume and dry weight computations from images

Area, major and minor axes from Ecotaxa were converted from pixels to millimetres (1 pixel = 0.0053 mm). The biovolume was computed for all the 69,949 mesozooplankton images using the Equivalent Spherical Diameter (ESD).

$$\text{ESD} = 2 \times \sqrt{\frac{\text{area}}{\pi}}$$

$$\text{Biovolume} = \frac{4}{3} \times \pi \times \left(\frac{\text{ESD}}{2}\right)^3$$

This computation was chosen to prevent overestimation of the individual's biovolume, due to unfolded copepods' antennae for example (as done in Monferrer et al., 2022). Then, each individual biovolume was multiplied by its density factor d to account for the water volume sampled, and a taxonomic-specific factor t to get the individual dry weight from individual biovolume value. These taxonomic-specific values were obtained and adapted from Maas et al. (2021) who defined linear relationships between biomass and biovolume for various taxa present at BATS (in bold, Table S3).

$$\text{Biomass} = \text{Biovolume} \times d \times t$$

2.3.3 | Relative biomasses and abundances

For comparison with relative read counts from metabarcoding (see below), normalized biomasses were transformed into relative biomasses for each taxonomic category i, going from 1 to $a = 29$, and per sample j, the sample rank going from 1 to $b = 27$, as follows:

$$\text{Relative Biomass}_{\text{tax}_{i\,j}} = \frac{\sum_{i=1}^{i=a} \text{Biomass}_{\text{tax}_{i\,j}}}{\sum_{j=1}^{j=b} \sum_{i=1}^{i=a} \text{Biomass}_{\text{tax}_{i\,j}}}$$

For the same purpose, relative abundances for each taxonomic category were computed per sample, using the same formula with $\text{Abundance} = \#\text{ind} \times d$ instead of Biomass. Data are presented in Table S4 for relative biomasses, expressed as a biomass per volume of water, normalized by the volume of water sampled (mg m^{-3}); and in Table S5 for relative abundances (ind m^{-3}).

2.3.4 | Construction of a multivariate morphological space from Zooscan images

To summarize the morphological characteristics of each organism in the mesozooplankton community and identify the main criteria of morphological variations, a morphological space was defined as in Vilgrain et al. (2021) by performing a weighted PCA (Legendre & Legendre, 2012) on a selection of 18 morphological variables (Table S2) for the 69,949 mesozooplankton images, weighted by their density factor d to give more weight to the most abundant organisms. The 18 morphological descriptors could be divided into four main categories, describing the shape, the size, the transparency, and the complexity of the organism itself, at the individual scale. The weighted PCA then allows the definition of a reduced morphological space (i.e., a morphological space with reduced dimensions) in which each image is located depending on its main morphological characteristics. By doing so, we can identify a few significant axes that summarize the main morphological traits that vary among images. 0.5% of the extreme variables were transformed into NAs to avoid being influenced by outlying points. They were all then normalized by the Yeo-Johnson transformation to satisfy the conditions for PCA application. Axes were considered significant only if their associated eigenvalue was greater than the average of all the eigenvalues (Kaiser-Guttman criterion; Cattell, 1966). The position of each mesozooplankton image in the PCA space is thus a function of its morphology: this PCA space will then be named morphospace. To ease the visualization of the morphospace, example images were mapped to their position in the morphological space as follows: for each pair of principal components (i.e., PC1 vs. PC2 or PC3 vs. PC4) a 15-by-15 grid was defined and five images were randomly chosen and superposed at each node of the grid, to represent a morphotype for each node. Individuals, meaning each of the 69,949 images, can also be plotted with dots in the morphospace. Supplementary variables were added to this PCA and represented in the morphological space as follows. Carbon export at 200 m was added as a quantitative variable and represented by arrows that pointed towards the morphospace area positively correlated to their variations. Pearson's correlation coefficients and their associated p-values were computed based on the coordinates of each individual in the morphospace and the carbon export value obtained during the matching sample. The three molecular-based clusters (see below), the sample time (Day or Night) and taxonomic categories were also added as qualitative supplementary variables.

2.3.5 | Linear relationship between size and transparency traits and carbon export

Pearson's correlation coefficients were computed between carbon export (ln-transformed to meet the normality assumption), opacity and size variables. Each individual's trait value was multiplied by its density factor to account for the water volume sampled. Data were averaged by sample dates (no Day/Night distinction) since there was only one carbon export measurement per cruise. For significant correlations ($p < .05$), linear relationships were represented in the graphs.

2.4 | Metabarcoding process and analyses

2.4.1 | DNA extraction and 18S amplification

Half of each ethanol-preserved sample (obtained with a Motoda splitter; Motoda, 1959) was used for DNA extraction. The other half was kept as archival material at BIOS. Since we focussed the study on mesozooplankton, macroplankton (>2 cm) were removed from the sample (mostly fish larvae). Samples were filtered through a 53 µm mesh, and gently washed with milliQ water to remove as much ethanol as possible. They were then transferred to a 50-mL falcon tube and centrifuged at 3500 *g* for 10 min to remove as much water and ethanol as possible. The pellet was then weighed and transferred to a new 50-mL falcon tube and 15 or 25 mL SDS buffer (10 mM Tris–HCl, 100 mM EDTA pH 8.0, 200 mM NaCl, 1% SDS) depending on the total mass of the pellet. Samples were then vortexed to ensure rapid mixing of the plankton with the buffer and thus degradation of the DNA. They were then ground using a Fisher Scientific FSH125 homogenizer for 5 min, at 1/3 of maximum speed to avoid excessive DNA fragmentation and rapid heating of the sample. After the grinding step, the homogenizer was rinsed with 10 mL SDS buffer (or 15 mL depending on sample mass) to minimize sample loss. Between samples, the homogenizer was washed and sterilized. Proteinase K (Sigma-Aldrich; St. Louis, MO) was added to the buffer (0.2 mg mL^{-1} final concentration), with 5 mL of sterilized stainless steel Shot, 3/32″ Ellipse (Rio Grande, CA). Samples were then placed in an oven at 60°C for 4 h, vortexing at maximum speed for 30 s every 30 min. Tubes were centrifuged at 3500 *g* for 15 min and three 400 µL replicates of supernatant were transferred to 1.5-mL Eppendorf tubes. DNA was extracted using the E.Z.N.A. Mollusc DNA Kit (Omega Bio-Tek, Norcross, GA), with some modifications explained in Blanco-Bercial (2020). Quality control of extracted DNA was performed by fluorimetry using a NanoDropTM One instrument. Extracted DNA was stored at −20°C until amplification.

Amplification of the V1V2 region of the 18S (about 350 bp) was done following Fonseca et al. (2010). Negatives were run during the extraction and amplification steps to ensure the lack of contamination. Negatives produce neither DNA nor bands during amplification. Samples were then sent to the Genomics Research Center of Rochester University (NY, USA) for sequencing. Sequencing was performed in an Illumina MiSeq using the MiSeq Reagent Kit v3 (600-cycles; 2×300) V3 chemistry. To test for replicability within the procedure, two samples were submitted using different indexes in separate batches. No major difference was found between the samples, the main taxa stayed within each sample in similar proportion.

2.4.2 | Bioinformatic pipeline

Demultiplexed samples were analysed in MOTHUR ver. 1.39.5. (Schloss et al., 2009). The entire annotated pipeline is accessible at the following link: https://github.com/margauxperhirin/mesozoo_bats. Unique sequences were aligned against a V1V2 region of the SILVA 138 database (Quast et al., 2013). Sequences were trimmed to the length of the V1V2 region, and only complete ones (starting at the first base and ending at the last base of the alignment) were kept in order to avoid bias due to unfinished amplifications. After gap removal, chimaeras were removed using UCHIME (Edgar et al., 2011) as implemented in MOTHUR, and unique sequences, called thereafter Amplicon Sequence Variant (ASV) were obtained after denoising using UNOISE2 (Edgar, 2016) as implemented in MOTHUR (diffs = 1). ASVs taxonomy was retrieved from a database developed from the complete SILVA 138 release database for SSU, modified to define clade levels (kingdom, phylum class, order, family, genus and species) to fixed columns in the taxonomy file (Bucklin et al., 2021). Using this classification and to focus only on mesozooplankton, only metazoans were kept for the downstream analyses (44,441 ASVs removed).

2.4.3 | Metabarcoding standardized analyses

ASVs belonging to Vertebrata taxa were eliminated, meaning that fish ASVs were removed to not be affected by cells from fish larvae able to swim against currents and thus not belonging to mesozooplankton (17,479 ASVs removed). Moreover, ASVs with less than 2 reads were removed. To control for uneven sequencing effort between samples, all samples were standardized randomly to a minimum number of reads, i.e., 92,150 reads. Percentages of relative abundance in the entire sampling period were computed per ASV, as follows, with j the sample rank going from 1 to $b = 27$ and k the ASV rank going from 1 to $c = 92{,}150$.

$$\%\,\mathrm{RA}_{\mathrm{ASV}_{\mathrm{TOT}k}} = \frac{\sum_{j=1}^{j=b} \mathrm{reads}_{\mathrm{ASV}_{jk}}}{\sum_{j=1}^{j=b} \sum_{k=1}^{k=c} \mathrm{reads}_{\mathrm{ASV}_{jk}}} \times 100$$

A minimum percentage of relative abundance threshold of 0.05% was then applied for all the following analyses. Indeed, metabarcoding reads are a viable proxy of taxon relative abundances for the major taxonomic groups, but it should be interpreted cautiously for the less abundant ones (Matthews et al., 2021). Hence, only 225 ASVs were considered in this study out of 22,814 ASVs. Finally, relative abundances (in percentage) estimated per ASV and per sample were computed as follows, with j the sample rank going from 1 to $b = 27$ and k the ASV rank going from 1 to $c = 225$.

$$\%\,\mathrm{RA}_{\mathrm{ASV}_{jk}} = \frac{\mathrm{reads}_{\mathrm{ASV}_{jk}}}{\sum_{j=1}^{j=b} \sum_{k=1}^{k=c} \mathrm{reads}_{\mathrm{ASV}_{jk}}} \times 100$$

These were then Hellinger-transformed (square-root transformation on relative abundances) to reduce the weight of very abundant ASVs (Legendre & Gallagher, 2001).

2.4.4 | ASVs relative counts of read per Zooscan categories and correlation with abundance and biomass

For direct comparison with image analyses, 22,814 ASVs were grouped into the taxonomic levels used in Zooscan/EcoTaxa systems. For each sample, the relative number of reads per coarse taxonomic category (in bold, Table S3) was computed. Pearson's correlation coefficients were then computed between relative counts of read and relative normalized abundances and biomasses from Zooscan data.

2.4.5 | Succession of ASVs assemblies throughout the sampling period

To explore the differences in mesozooplankton community throughout the sampling period, a Principal COordinates Analysis (PCoA, Legendre & Legendre, 2012) based on Bray–Curtis dissimilarity index (Bray & Curtis, 1957) was performed. Sample dates were clustered by hierarchical classification carried out on their coordinates along the first seven axes of the PCoA to explain more than 70% of the variance with at least 4% in each axis (72.59% of variance explained), using a Euclidean distance and a synoptic Ward linkage (Murtagh & Legendre, 2014). A threshold (−0.55, Figure 3) was set to minimize the total within-cluster variations, defining three meaningful groups of stations with homogeneous values for the variables used. The significance of the PCoA clustering pattern was tested using an ANalysis Of SIMilarity (ANOSIM, 1000 permutations, Clarke, 1993) and permutational multivariate analysis of variance (ADONIS, 1000 permutations, Anderson, 2001) Holm corrected (Holm, 1979).

2.4.6 | Co-occurrences network and 'modules' identification

A network of co-occurring ASVs was built on the relative Hellinger-transformed abundances. Nodes represented ASVs and edges were the positive correlations between those, computed with Spearman's correlation coefficient (rho, Spearman, 1904). Cut-offs of rho and p-values were respectively 0.6 and 0.01 to consider the most statistically robust correlations, meaning only links between pairs of ASVs that often appeared at the same time in the samples and reducing the risks to consider by-chance co-occurrences (Barberán et al., 2012). Modularity estimates the tendency of a network to subdivide in densely connected modules (or clusters) (Bellisario et al., 2019). Networks with high modularity have dense connections between the nodes within modules but limited connections between nodes in different modules. If modularity is higher than 0.40, then we consider the network can be divided into modules, that is, groups of ASVs there are more densely connected between each other than with the rest of the network (Newman, 2006). The Leading Eigenvector community detection algorithm was used (Newman, 2006) to identify the modules. In a few words, the network is splitted into two parts that would be kept if the separation leads to a significant increase in the modularity, and each part would be divided again into two parts at the next step of the algorithm. If the separation does not lead to a significant increase in the modularity, the network is splitted into two other parts, and the process starts again. The algorithm stop when no further separation leads to a significant increase in the modularity.

2.4.7 | Identification of taxa composing the modules

A more precise and trustable taxonomy of ASVs belonging to modules was assigned via BLASTN from the NCBI website (https://www.ncbi.nlm.nih.gov/, Altschul et al., 1990) followed by Batch Entrez analyses to retrieve species identities from Accession Hit numbers (https://www.ncbi.nlm.nih.gov/sites/batchentrez). Then, obtained files were formatted in R. Taxonomy was obtained from LifeWatch Belgium services (https://www.lifewatch.be/data-services/). Unique best hits, meaning that an ASV matches with a unique sequence from the NCBI database at 100% of similitude, were firstly chosen. When multiple species had equal identity (and when one was not a clear misidentification), the lowest common taxonomic rank was kept, e.g. if an ASV matches at 95% of similitude with two sequences in NCBI database belonging to *Candacia truncata* and *Candacia elongata*, then it was characterized as *Candacia* sp. The complete taxonomy of the 225 most abundant ASVs analysed in this study can be found in Table S7.

For each module m (between 1 and $e=6$), their relative abundances were computed per sample as follows, with j the sample rank going from 1 to $b=27$.

$$\%\,\mathrm{RA}_{\mathrm{Mod}_{jm}} = \frac{\sum_{m=1}^{m=e} \mathrm{reads}_{\mathrm{ASV}_{jm}}}{\sum_{j=1}^{j=b} \sum_{m=1}^{m=e} \mathrm{reads}_{\mathrm{ASV}_{jm}}} \times 100$$

To test the Day/Night effect on the modules' composition, a Wilcoxon–Mann–Whitney test (Mann & Whitney, 1947) was conducted on the percentages of relative abundance per sample, for all the modules. Finally, a Kruskal–Wallis analysis (Kruskal & Wallis, 1952) helped to test the impact of the three clusters obtained from the PCoA on the modules' percentages of relative abundance, followed by Dunn pairwise test (Dunn, 1961) with Holm p-values correction for multiple comparisons to identify which cluster affected the most the modules' percentages of relatives abundance. In both cases, nonparametric statistical tests were realized because Shapiro (Shapiro & Wilk, 1965) and Levene's (Brown & Forsythe, 1974) tests showed that not all modules' percentages of relative abundance were normally distributed with similar variances among groups. Moreover, Pearson's correlation coefficients were computed between percentages of relative abundance per module and per sample, and their coordinates along each morphological dimension per sample to have information about the overall morphology of the module; and carbon export per cruise to see if some modules were related to higher or lower carbon export.

2.5 | Numerical tools

All statistical analyses were conducted in the programming environment R 4.1.2 (R Core Team, 2021). The package tidyverse (Wickham et al., 2019) was used for data manipulations, lubridate (James and Hornik, 2020) to deal with temporal variables, vegan (Oksanen et al., 2009) and FactoMineR (Lê et al., 2008) for multivariate analysis, especially the function *dimdesc* was used to compute correlations between carbon export and morphospace axes. pairwiseAdonis (https://github.com/pmartinezarbizu/pairwiseAdonis) was used for permutational multivariate analyses of variance, and igraph (Csardi & Nepusz, 2006) for network and modularity analyses. rstatix (Kassambara, 2021), NbClust (Charrad et al., 2014) and Hmisc (Harrell Jr & Dupon, 2022) were used for statistical analyses. The packages RColorBrewer (Neuwirth, 2022), ggplot2 (Wickham et al., 2021), ggpubr (Kassambara, 2020), ggrepel (Slowikowski et al., 2021) and cowplot (Wilke, 2020) were used to produce general graphics, morphr (https://github.com/jiho/morphr) was used to represent individual images in the morphological space.

3 | RESULTS

3.1 | Seasonal dynamics of the BATS ecosystem

Temporal changes in environmental variables followed a classical seasonal cycle. As expected, temperatures reflected a deep winter mixing, from December 2016 to April 2017. Summer stratification was strong and deep (about 40 m) from July 2016 to September 2016 (Figure S1). MLD increased in autumn and allowed the reinjection of nutrients in the surface layer, creating ideal conditions for an autumnal bloom. Fluorescence was low in surface waters except during the deep mixing in January 2017 and April 2017 (Figure S1). DCM was around 100 m. Carbon export was the highest in April 2017 and in May 2016 and nearly absent in September 2016 ($0.75\,mgC\,m^{-2}\,day^{-1}$). Otherwise, it fluctuated between $6\,mgC\,m^{-2}\,day^{-1}$ and $25\,mgC\,m^{-2}\,day^{-1}$.

3.2 | Carbon export is more transparency-dependant than size-dependant

The four first axes of the PCA based on the morphological variables were significant. They explained about 87.68% of the morphological variance observed in mesozooplankton samples (Figure 1). PC1 (42.87%) was mostly influenced by size and shape related variables (e.g., Perimeter, Circularity, FerretDiameter or Symmetry) with smaller and more circular organisms for PC1 < 0 and larger and more elongated ones for PC1 > 0. PC2 (23.89%) characterized the grey level of mesozooplankton and their size, with PC2 < 0 corresponding to larger, less transparent organisms and PC2 > 0 corresponding to smaller, more transparent ones. PC3 (14.53%) referred to the elongation of the organisms and how their appendages were positioned when organisms were scanned. Negative PC3 < 0 corresponded to organisms with visible appendages while PC3 > 0 represented elongated organisms with a well-defined body (folded or no appendages). Finally, PC4 (6.59%) described the heterogeneity of grey colour within each organism, with organisms with more heterogeneous grey colour for PC4 < 0 and organisms with homogeneous grey colour for PC4 > 0. Morphology of mesozooplankton was quite similar between Day and Night samples, except along PC4 (homogeneity in grey colour)

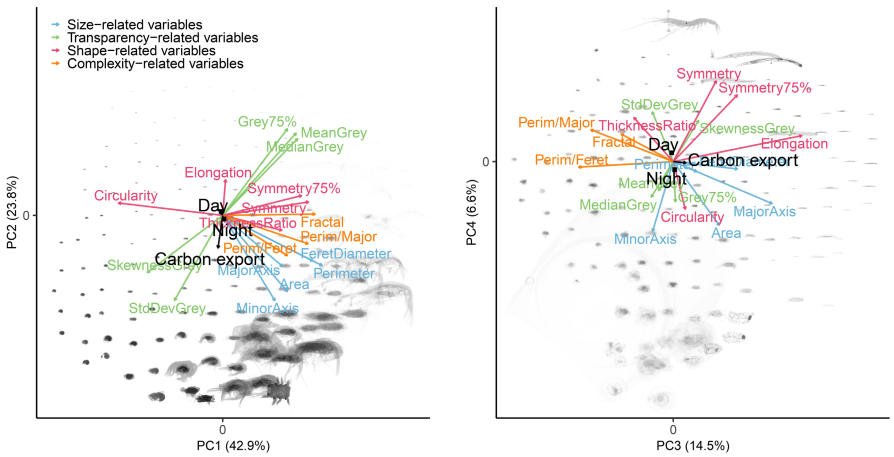

FIGURE 1 Morphological space of mesozooplankton Zooscan images. The four significant axes of the PCA performed on morphological variables are represented: PC1 and PC2 on the left, PC3 and PC4 on the right. Morphological variables (see details in Table S2) are coloured according to their category (size-, transparency-, shape-, or complexity-related). PC1 is mostly related to circularity and size, PC2 to transparency, PC3 to shape complexity and appendages' visibility and PC4 to homogeneity in grey colour of the organisms. For each factorial plane, morphotypes are represented according to their coordinates in the morphological space: 15-by-15 grid was de fined and 5 images close to the factorial plane considered were randomly selected, aligned and superimposed at each node of the grid for visualization. Carbon export is represented on both panels as a supplementary quantitative variable (black arrow) and the sample time (Day or Night) is also represented on both panels as a supplementary qualitative variable (dark square).

with organisms presenting less homogeneous grey levels during the Day (Figure 1).

Correlations between carbon export and principal components were significant for PC1, PC2, and PC3 and reached −0.04, −0.30 and +0.08, respectively ($p < .01$, Figure 1). Thus, according to the visible signal in morphological space (Figure 1), confirmed by the correlations just mentioned, carbon export seems to be more strongly linked to PC2, the transparency of organisms, than to the other axes. To validate this signal, linear relationships were computed between carbon export values and size-related morphological descriptors (to test the assumption that size is a master-trait), and also transparency-related morphological descriptors (correlated with PC2). The results are visible in Figure 2, no significant correlations (p-values $> .05$) were found between carbon export and size-related morphological descriptors. However, significant correlations (p-values $< .05$) were found between morphological descriptors related to transparency (correlated to PC2) and carbon export. From these correlations, it is possible to infer the relationships between imaged organism's transparency and carbon export rates. First, lower organisms' transparency, signified by low values of the three grey intensity variables (Figure 2f,g,h) was associated with higher carbon export. Second, carbon export was higher when StdDevGrey values were high (Figure 2i), that is, when mesozooplankton individuals presented larger ranges of grey level values. Third, carbon export was higher when SkewnessGrey values were higher (Figure 2j). This descriptor indicates how skewed the normal distribution of the grey level values is per image. Therefore, a more normal distribution of grey values is associated with higher carbon export.

3.3 | Three communities with different levels of transparency succeed one another at BATS

A PCoA was conducted to analyse the temporal evolution of the taxonomic composition from metabarcoding data (Figure 3). The first PCoA axis explained about 22.99% of the variance in the taxonomic composition, the 2nd one explained 14.84%. Hierarchical clustering from the first seven PCoA axes (72% of variance explained) coordinates revealed that three communities could be distinguished from their ASV composition (Figure 3). Molecular-based cluster 1 corresponds to all the samplings done between June 2016 and October 2016. The second contains day and night samplings from May 2016 and April 2017. Molecular-based cluster 3 is composed of all the other samplings. Based on their timeline, these groups represent communities present during the summer stratification period (Cluster #1), the zooplankton bloom period (Cluster #2) and the deep winter mixing season, except samplings done in April 2016 and May 2017 (Cluster #3). These patterns were confirmed by the identification of ASVs specific to each of the molecular-based clusters whose identity and functional traits match with environmental conditions (Analysis S1). Clustering pattern is significant ($p = .001$) and communities are all significantly different from each other (Holm p-adjusted $= .003$ between all pairs of clusters).

Fitting the three molecular-based communities in the morphospace, it appears that they were quite similar along PC1 (size and circularity) (Figure 4). However, individuals belonging to the bloom community tended to be less transparent (lowest coordinate along PC2 axis) than winter mixing cluster ones. Individuals from the molecular-based summer cluster were the most transparent ones (lowest coordinate along PC2 axis). This signal of transparency changes, along PC2, matches for carbon export intensity, as explained in the previous paragraph. A similar trend is observed along PC3 (Figure S2) with taxa from the bloom community tending to be more elongated with fewer or no visible appendages. On the opposite end, taxa from the summer stratification community were less elongated with more visible appendages. The winter mixing community was between both. This succession of communities along PC2 and PC3 also matches for carbon export intensity.

3.4 | Metabarcoding read counts and abundance and biomass values from Zooscan are positively correlated, for some taxonomic groups

Correlations between relative number of reads and relative abundance or relative biomass derived from Zooscan images were computed for each taxonomic category (Table 1). Cyclopoida, Doliolida, Harpacticoida and Larvacea relative abundance and biomass were positively correlated with relative read count. Only relative biomass was positively correlated with relative read count for Cephalopoda (Table 1), while relative abundance was positively correlated with relative read count for Chaetognatha and Cladocera. It was negatively correlated for Ostracoda.

3.5 | Migrating and seasonal of groups' effects on carbon export

To better identify the degree to which specific taxa drove morphological variations and carbon export changes described above, a co-occurrence network was built on the 225 ASVs. It contained 413 edges (average of 3.67 co-occurrences per ASV). It had a modularity coefficient of 0.72, meaning that the network can be divided into modules. A total of 15 modules were found. Six modules were composed of at least 10 ASVs, named hereafter A–F. They contained between 11 (module F) and 36 (module D) ASVs.

Taxa were quite different from one module to another. Module A contained many Euphausiacea, Ostracoda from three genus, and some gelatinous taxa (Table S6). Module B was exclusively composed of Arthropoda (Table S6). Module D included some Appendicularia and Thaliacea, among other organisms (Table S6). Modules' entire taxonomic composition can be found in Table S7.

Relative abundances trends were highly variable among modules (Figure 5). They were very similar between day and night samples, except for Module A (Wilcoxon–Mann–Whitney test; $p < .001$, Figure 6 & Table 2), especially during the winter period (from

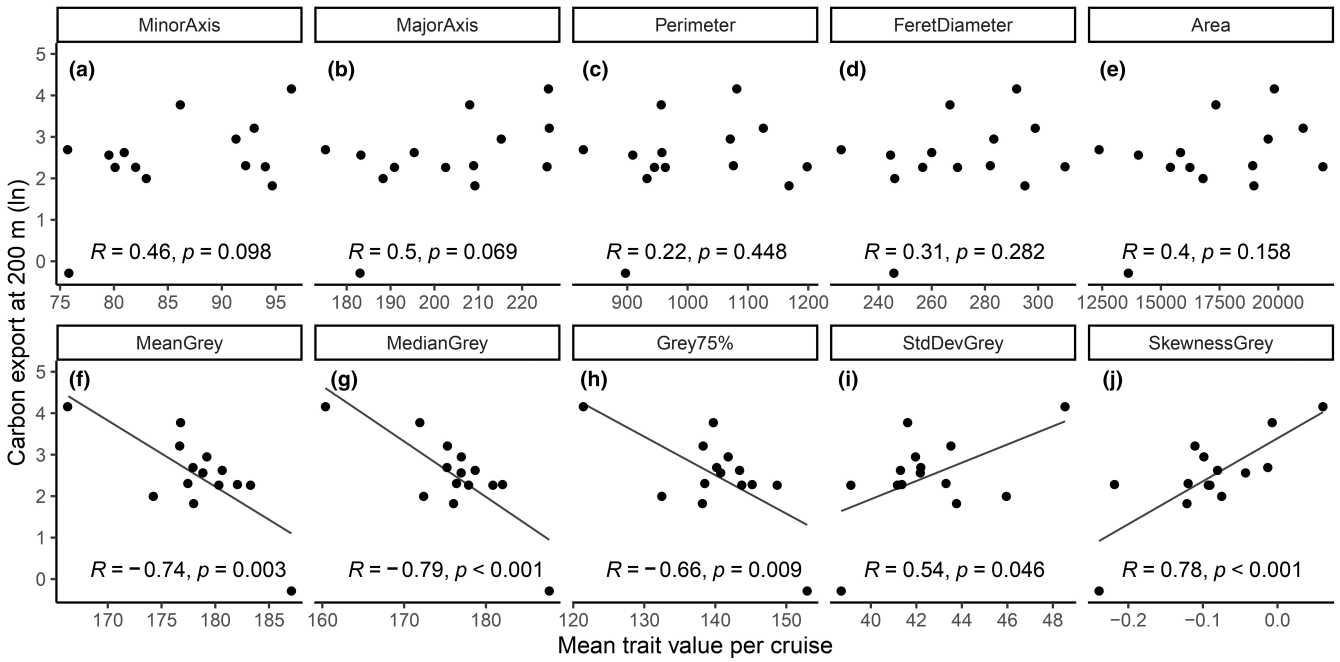

FIGURE 2 Carbon export (mg m^{-2} d^{-1}, ln-transformed) in function of the mean values of the main morphological variables, per cruise. Size-related variables are on the top (a–e), transparency-related variables are on the bottom (f–j). Correlations are represented when significant. Pearson correlation coefficients and p-values are displayed. See Table S2 for details about morphological variables' meaning.

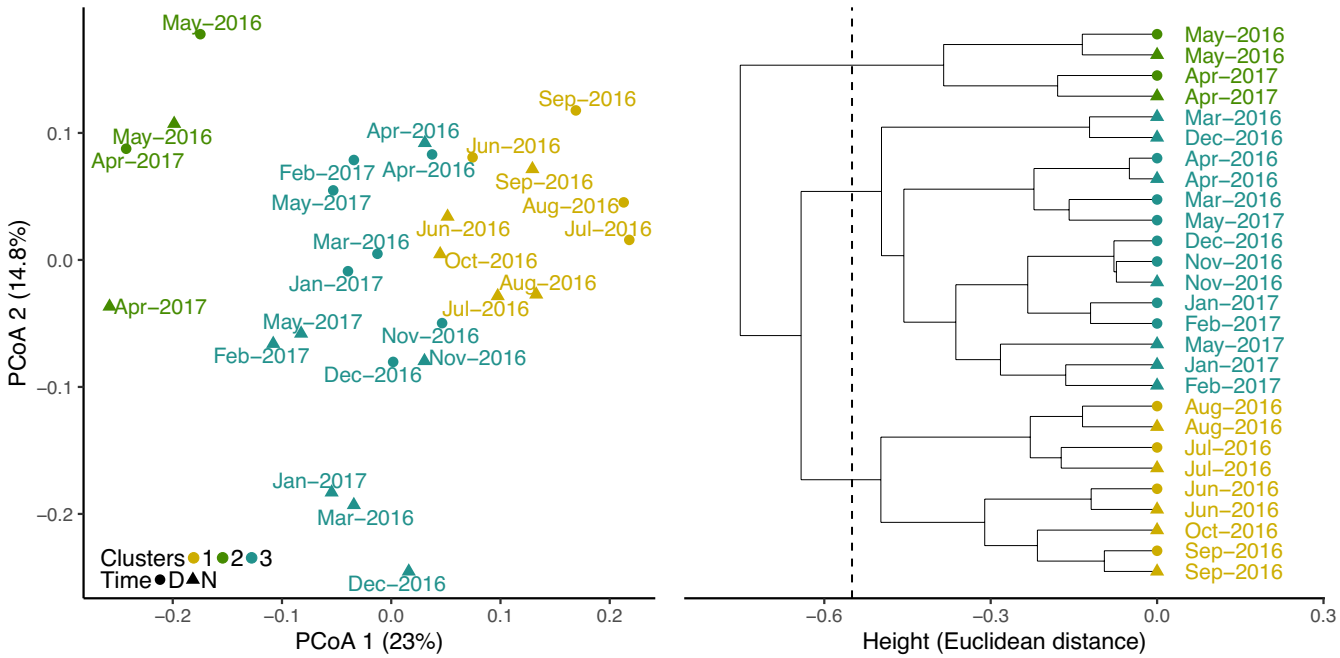

FIGURE 3 Molecular-based clustering. PCoA based on Bray–Curtis dissimilarity matrix computed on Hellinger-transformed ASVs relative abundances per sampling (PCoA1PCoA2 panel on the left) and Ward.D2 clustering on Euclidean distances between samplings' coordinates along the first seven axes (on the right). Colours refer to clusters in both panels. The vertical dashed line represents the threshold used to define the three clusters (−0.55).

November 2016 to January 2017). For all modules except A, there were significant changes in their relative abundances between molecular-based cluster samplings (Kruskal–Wallis test, $p < .05$ except for module A: $p = .182$; Figure 6). Compared with the overall study period, module B was relatively more abundant during the bloom period (molecular-based cluster 2) while module D was relatively less abundant during the same period (Dunn pairwise comparison test; Figure 6 & Table 2). Taxa included in module C were relatively

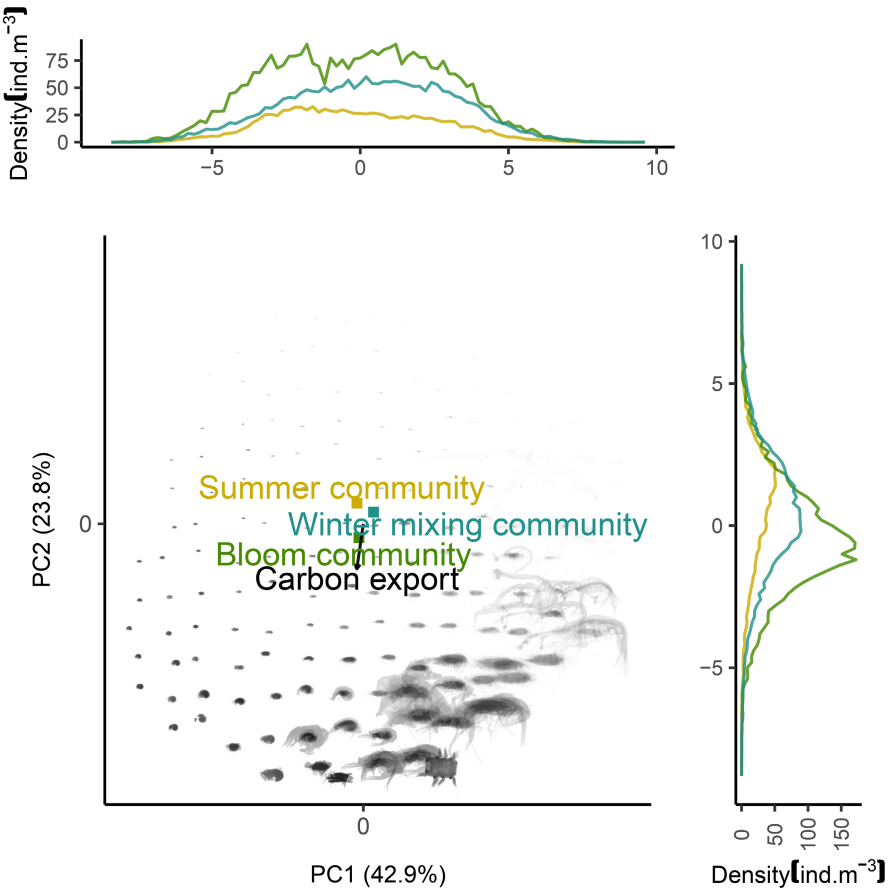

FIGURE 4 Seasonal variation in mesozooplankton morphology. The first two significant components of the morphological space and corresponding morphotypes. As a reminder, PC1 represents organisms' circularity and size (the more positive, the less circular and the larger organisms are) while PC2 represents organisms' transparency (the more positive, the more transparent organisms are). Each organism is associated to a molecular-based cluster, via its sampling date (Figure 3). This cluster membership was added in the morphological space as a supplementary qualitative variable and centroids of the distribution of each cluster (i.e., Summer community, Bloom community, Winter mixing period) are represented. Carbon export is also represented as a supplementary variable (black arrow). Side plots illustrate the distribution of mesozooplankton individuals' density (ind m^{-3}) in the morphospace, discriminated per molecular-based cluster. For example, individuals belonging to the bloom community are more abundant than individuals belonging to other molecular-based clusters, and they tend to appear less transparent since distribution's mode corresponds to a lower PC2 value.

Taxonomic category	Relative read counts/relative abundance, Pearson ρ (*p*)	Relative read counts/relative biomass, Pearson ρ (*p*)
Cephalopoda	n.s.	+0.91 ($p = .030$)
Chaetognatha	+0.60 ($p < .001$)	n.s.
Cladocera	+0.95 ($p < .001$)	n.s.
Cyclopoida	+0.55 ($p = .003$)	+0.51 ($p = .007$)
Doliolida	+0.58 ($p = .001$)	+0.38 ($p = .047$)
Harpacticoida	+0.64 ($p < .001$)	+0.65 ($p < .001$)
Larvacea	+0.80 ($p < .001$)	+0.87 ($p < .001$)
Ostracoda	−0.43 ($p = .025$)	n.s.

Note: No significant correlations were found for either abundance or biomass for the following categories: Annelida, Bivalvia, Bryozoa, Calanoida, Cavoliniidae, Cephalochordata, Cirripedia, Cnidaria, Creseidae,Decapoda, Echinodermata, Euphausiacea, Gymnosomata, Limacinidae, and Salpida.

TABLE 1 Pearson correlation coefficients ρ and their *p*-values, between relative numbers of reads, obtained by metabarcoding and relative abundances, or relative biomasses taken into account the total volume sampled, obtained from the Zooscan, for large taxonomic categories. Correlations compute on variables' means per category and sampling.

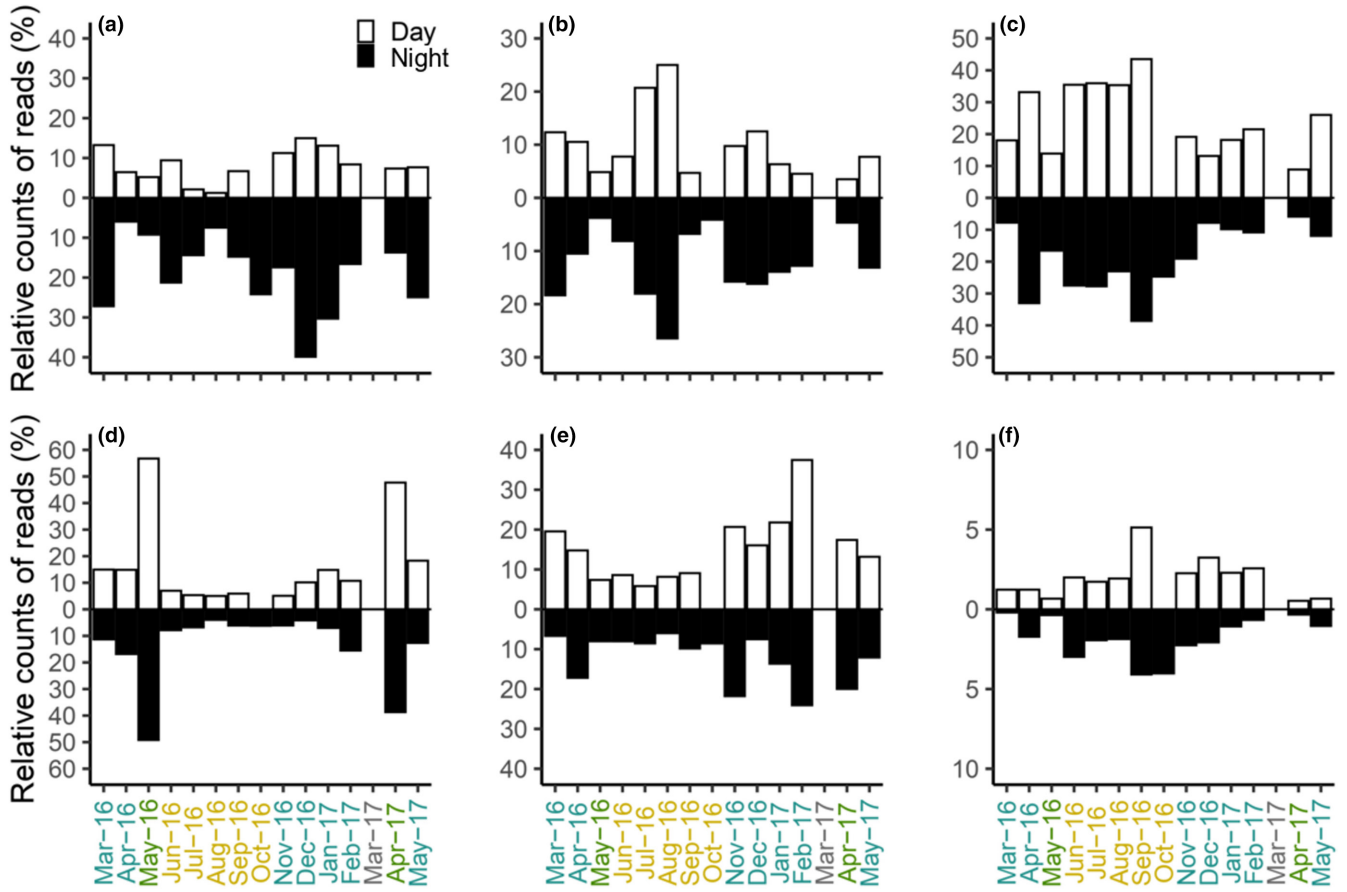

FIGURE 5 Relative counts of reads per module, throughout the sampling period, with a Day/Night distinction. Letters in the top left corner of each plot indicated the represented module. The colours of the x-axis labels represent the 3 molecular-based cluster's periods (Figure 3). There is no data on October-16 Day due to bad weather and on March-17 due to their bad quality. Note that the y-axis scale changes throughout panels.

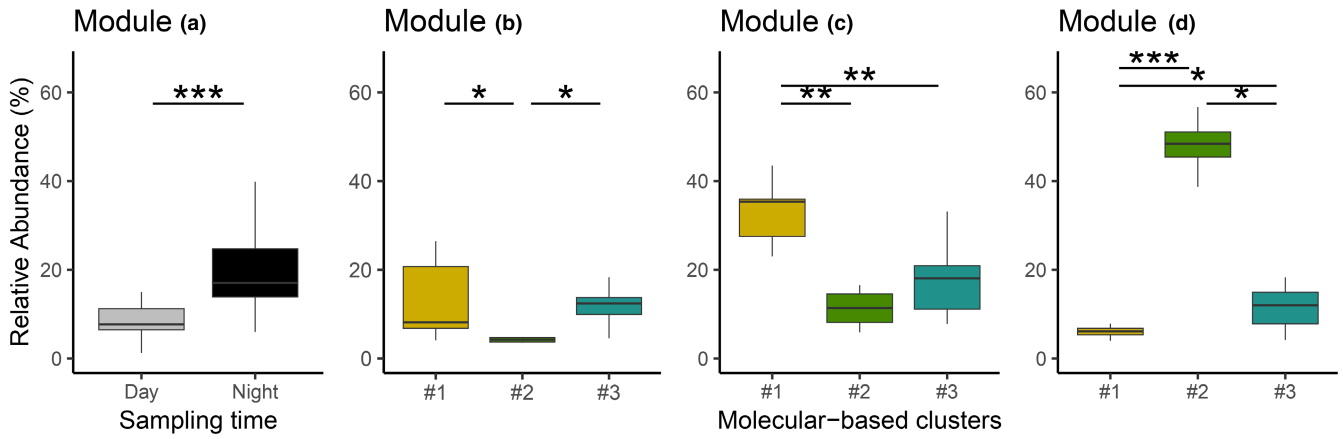

FIGURE 6 Temporal variations in modules' relative abundances. Boxplots of the percentages of relative abundances of ASVs belonging to module (a), sampled during the Day or at Night; and of the percentages of relative abundances of ASVs belonging to modules (b, c) or (d), merged according to the samples' affiliation to a molecular-based cluster. Asterisks represent the strength of the significant difference between distributions (Kruskal-Wallis test followed by Dunn pairwise test; * for $0.01 < p\text{-value} < .05$; ** for $0.001 < p\text{-value} < .01$ and *** for $p\text{-value} < .001$). Boxes are coloured by molecular-based clusters for module (b, c) and (d).

more abundant in BATS mesozooplankton community during the summer stratification period (molecular-based cluster 1) ($p < .001$, Figure 6 & Table 2). Percentages of relative abundance from module F were steady and low throughout the study period. Morphology differed also between modules (Table 2). When ASVs from module E were more present, they tended to increase the size and reduce

TABLE 2 Significant ($p < .01$) correlations between percentages of relative abundance per sampling for each module (from A to F) and diverse variables.

Module	PC 1	PC 2	PC 3	PC 4	Carbon export	Other relationships
A	n.s.	n.s.	n.s.	n.s.	n.s.	Night ($p < .001$)
B	n.s.	n.s.	n.s.	n.s.	n.s.	Not bloom community ($p = .019$)
C	n.s.	+0.74	n.s.	n.s.	−0.51	Summer community ($p < .001$)
D	n.s.	n.s.	0.66	n.s.	0.87	Bloom community ($p < .001$)
E	0.55	n.s.	n.s.	n.s.	n.s.	n.s.
F	n.s.	+0.55	−0.54	n.s.	−0.6	n.s.

Note: PC X represents the coordinates of the sampling points considered as supplementary qualitative variables in the morphospace along axis number X. Carbon export has no Day/Night distinction. Last column contains pieces of information from non-parametric statistical tests.

the circularity (PC1 > 0) of the overall mesozooplankton community. Sampling positions along the second axis were significantly and positively correlated to percentages of relative abundances of modules C and F (respectively +0.74 and +0.55). Hence, when ASVs from these modules are more abundant, the overall mesozooplankton community tended to be more transparent (PC2 > 0) than usual. Finally, Pearson's correlation coefficient was significantly positive between relative abundances of module D and sampling positions along the third axis of the morphospace (+0.67). It was negative between relative abundances of module F, and sampling positions along the third axis of the morphospace (−0.54). Thus, the overall mesozooplankton community appeared with folded appendices, or was more composed of taxa without appendices (PC3 > 0) when ASVs from module D are relatively more abundant, and inversely when module F is relatively more present. Relative abundances of module E and sampling positions along the first axis of the morphospace were positively correlated (+0.56, Table 2). Finally, carbon export was positively correlated to relative abundance of module D (+0.87) and negatively correlated to modules C (−0.51) and F (−0.60).

4 | DISCUSSION

Despite the common assumption that size is the master trait that controls most planktonic processes, here transparency was the key trait defining one of the main biogeochemical processes influenced by plankton, carbon export. Higher export was found to be linked to less transparent animals. Potential explanations for these findings could include opacity due to full guts or the presence of pigments particularly during the bloom period. It could also be due to the presence of strong swimmers participating in the DVM, as those participating require stronger muscles and thicker carapace walls (where muscles attach), increasing the opacity of the community when they become dominating, especially during the spring bloom. This study also shows morphological and molecular data from the same samples can be analysed jointly to better infer how various facets of mesozooplankton diversity (morphological, taxonomical) vary throughout seasons and impact biogeochemical processes. The image-based functional approach allowed for the assignment of morphological properties to individual plankton and added taxa that were underrepresented in the metabarcoding data due to miss amplification. Moreover, benefiting from the "twin" imaging and molecular data sets, we studied the quantitative property of metabarcoding for wide taxonomic categories. Metabarcoding proved to be a good proxy of abundance or biomass for only a few taxa (e.g., Cyclopoida, Doliolida, Larvacea), but could not be related to most of the considered taxa (e.g., Calanoida, Euphausiacea, Salpida).

4.1 | The importance of transparency in mesozooplankton communities

Transparency is a widespread, however largely understudied, characteristic in oceanic zooplankton (Johnsen & Widder, 1998; Orenstein et al., 2022). While most of the research on this morphological characteristic have been done in relation to its protective role from UV radiations or from visual predators (e.g., Brüsin et al., 2016; Hays et al., 1994; Sha et al., 2021), here we showed that transparency might be a good parameter for assessing carbon export intensity at the scale of the mesozooplankton community. Indeed, while all transparency-related variables used were linearly linked to carbon export intensity, none of the size-related parameters were significantly correlated. This might be due to a constant size variation in the BATS mesozooplankton community despite seasonal dynamics among mesozooplankton size classes (Steinberg et al., 2012). Moreover, differences between the three communities could be defined by variations in overall transparency levels. Transparency gives information about body structures and pigmentation links to physiological functions (Orenstein et al., 2022). Hence, these variations might reveal gut content or pigmentation as developed in Vilgrain et al. (2021). Changes in community transparency also reflect taxonomic variations in the community. For example, during the bloom period, imaging data showed more individuals with folded appendages or more elongated organisms. This might be explained by the presence of chaetognaths and larvaceans, as represented on the positive side of PC3 axis (Figure S3).

With copepods representing about 70% of the total scanned community in this study well representing the community's proportions (Deevey, 1971), and knowing the high diversity of this group (probably around 400 species in the Sargasso Sea; Deevey &

Brooks, 1977), inter-group variations are likely to affect the overall community morphological characteristics. For example, transparency variations can be linked to the presence of migrating copepods, such as *Pleuromamma* sp., during the bloom period (Steinberg et al., 2000). These copepods migrate more than a hundred thousand times their body length twice a day (Goetze, 2011), and more generally diel vertical migrations vary between dozens of metres to 2000 metres (Deevey & Brooks, 1977). To be sufficiently efficient, they develop two systems for steady propulsion and intensive escapes. The last one relies on a particular arrangement of copepods' muscles forming a lever system by being attached to organisms' exoskeleton (Kiørboe et al., 2010). Hence, it could be that the tougher or quicker a copepod needs to be, the stronger its exoskeleton must be to resist their muscles. This may explain why organisms appeared more opaque during the bloom period to escape more predators. Finally, a transparency-productivity relationship was found for copepods in the Bay of Baffin, with a proportionally inverse relationship between transparency and chlorophyll-*a* concentrations (Vilgrain et al., 2021), which is quite similar to the transparency-stratification and opacity-bloom dynamics found in this study. Hence, transparency is related to the survival of an organism at the individual scale (Martini et al., 2021) but it might also be a carbon export intensity indicator at the entire community scale, as it was suggested by Da Silva (2021).

4.2 | Different taxonomic assemblages and ecosystem dynamics

Among the six modules identified in the co-occurrences network, two confirmed the link between transparency and low export. Module C was related to the summer stratification community. At this time of the year, new nutrient inputs in the ecosystem were scarce and taxa were typical of a recycling community. It means that nutrients are recycled by coprophagy, particle-feeding and throughout the trophic network which helps to maintain zooplankton biomass in the upper mixed layer during periods of intense stratification of the water column (Kiørboe, 1997). Images analyses further revealed that the community tended to be more transparent when module C was dominant which could be explained by the presence of low carbon content organisms (Siphonophorae, Salpida). The taxa in module C fit in these ecological niches. *Oithona* sp. biomass is negatively correlated to the vertical flux of faecal pellets carbon (Svensen & Nejstgaard, 2003) and they also are also known to reduce the vertical flux of faecal pellets (González & Smetacek, 1994). *Paracalanus aculeatus*, a small copepod typical of low production ecosystems which helps in nutrients recycling, was also present (Paffenhöfer et al., 2006). The abundance of the recycling community and the presence of organisms known to have a low carbon content explain the reduced carbon export during summer stratification. Module C contained five Gastropoda taxa, three of which do not have any shell reducing their dry weight content (volume:C ratios 20–100 times those of non-gelatinous organisms; Kiørboe, 2013) and contribute to the overall transparency of the community (Alldredge, 1984). Found commonly in the samples, *Atlanta selvagensis* has a transparent shell (de Vera & Seapy, 2006). Along with *Thliptodon* sp., they are strong recyclers which feed on other mesozooplankton, including *Cymbulia* sp. (personal communication of Amy Maas; Mifsud, 2001). Further strengthening the link between opacity and low carbon export, *Cymbulia* sp. and *Peracle reticulata* are particle-feeders associated with recycling communities. Module F, although relatively less abundant than module C, was also negatively correlated with carbon export. It included Doliolida, Larvacea or Bryozoa taxa whose centroids are on the transparent and low export side of the morphospace ($PC2 > 0$, Figure S3), further linking transparency and recycling community. Again, it contained taxa indicating a recycling community. For example, *Halocypris* sp. are predominantly detritivores (Angel, 1983) and play a role in the recycling of organic matter (Castellani & Edwards, 2017). In addition to having the same transparency signal as module C, module F was also linked to a community with more visible appendages ($PC3 < 0$, Figure S3), for example, Bryozoa larvae can use their cilia to capture particles (Strathmann, 2006).

Despite the above, an opacity-export signal could have been expected for two other modules but was not measured in this study. Module A contained almost exclusively taxa known to migrate, and using the co-occurrences network, we identified these taxa potentially interacting (Berry & Widder, 2014) that we would have missed otherwise. Euphausidae, except for the genus *Stylocheiron*, are in general strong migrators (Brinton, 1962; Roe et al., 1984) and known to contribute to the diel community changes at BATS station (Blanco-Bercial, 2020). Siphonophorae and Salpidae also realize diel vertical migrations (Andersen et al., 1992; Lüskow et al., 2019). DVM was shown for several genus of ostracods (Kaeriyama & Ikeda, 2002; McHardy & Bary, 1965; Ursella et al., 2018). Even though DVM is a process known to significantly contribute to the biological pump (Steinberg et al., 2008), no significant positive correlation was found between carbon export and module A relative abundances. There are several possible explanations for this discrepancy. Euphausidae migration ranges are usually larger than 200 m (Andersen & Sardou, 1992), meaning that sediment traps might have missed their contribution to the overall export. A possible explanation for DVM is the predator-evasion hypothesis, in which large or pigmented organisms can hide at depth during the day to be less visible while transparent organisms can stay in surface waters (Johnsen, 2014; Sha et al., 2021). However, no correlation was found between migrators and opacity, which might be explained by the non-amplification of certain taxa (see Quantitative property of metabarcoding). For example, Amphipoda are common vertical migrators in the Sargasso Sea (Steinberg et al., 2000). Individuals were part of the morphological data set and very opaque in a large majority ($PC2 < 0$, Figure S3), however they were not present at all in the molecular data set, leading to a potential DVM signal discrepancy. Finally, Module D matched with the zooplankton bloom community when the peak of biomass usually occurs (Madin et al., 2001; Steinberg et al., 2001). This community was composed

of organisms more elongated without visible appendices, e.g., Larvacea, Chaetognatha (PC3 > 0, Figure S3). Module D relative abundances were positively and strongly correlated to carbon export; however, no opacity signal was found. It might be due to the gelatinous plankton taxa it contained (five *Oikopleura* sp., four Doliolidae and one Salpidae). Indeed, by their ability to bloom and/or to produce high quantity of faecal pellets, they can play an important role in carbon export (10%–79% of the 200-m POC flux for salps, Stone & Steinberg, 2016), but by being highly transparent (PC2 > 0, Figure S3; Johnsen & Widder, 1998), their abundance might have skewed the expected opacity-export signal.

This link between transparency or mesozooplankton and carbon export, observed at BATS within a 12-month time-series, could be true for oligotrophic ecosystems like BATS one. However, other studies should test it for longer time-series and for other pelagic ecosystems, taking advantage of the transferability of the method used here. Imaging data were obtained from formalin-fixed organisms through the Zooscan process. Grey colour levels and shape might then not exactly represent how organisms looked when alive, however, it is known that the position of copepods preserved in formalin remain stable after an initial shrinkage period (Sampei et al., 2009) which is much shorter (in the order of days) than the time gap between sampling and scanning in our study. Because more than 1500 organisms were scanned per size-class and sample, we believe that shape-related morphological descriptors, such as Size or Elongation (PC1 and PC3, Figure 1), were robust at the scale of the community. Similarly, all the samples were stored under exactly the same conditions, following the same protocols, and therefore underwent the same potential degradation. There is still the fact that fixation affects each taxonomic group differently (e.g. gelatinous vs. arthropods). The morphological descriptors linked to the transparency of the organisms are however inter-comparable since we are analysing trends and changes in the community, therefore the patterns found would then be reliable.

4.3 | Quantitative property of metabarcoding

Metabarcoding on the 18S rRNA region of DNA is a common strategy to study zooplankton. Various sections of this region have been used as barcodes (e.g., V9 or V4 the standard markers for planktonic eukaryotes since Tara Oceans studies). In this study, we choose to amplify and sequence the V1V2 region, originally used for meiofaunal zoobenthos (Fonseca et al., 2014) and specific to Metazoan (Lejzerowicz et al., 2021). It appears to be one of the best markers to assess marine community changes (Cordier et al., 2018). However, it is important to consider that the use of different molecular markers could have produced slightly different results and that some taxa are not well characterized, for example, *Pleuromamma* sp. or Chaetognatha. This may be due to a secondary structure on their ribosomal region which might make their DNA harder to amplify as the primer regions show high fidelity in the available sequences in the NCBI. This could impact our results since *Pleuromamma* sp., which is highly abundant at BATS, is known to rapidly react to the new primary production and to contribute significantly to the ecosystem carbon flux (between 4% and 70%, Steinberg et al., 2000). Chaetognaths are also common in this study region. They are thought to be linked to the bloom community, since their abundance usually follows copepod density maxima and decreases with increasing stratification of the water column in the Sargasso Sea (Ivory et al., 2019) and their morphology matches this cluster's characteristics. Therefore, due to chaetognath's sequences low amplification we might have missed potential interspecific interactions. Generally, correlations are found between the number of reads per ASV and the corresponding taxon biomass or abundance (Bucklin et al., 2019). Significant relationship also exists between the proportion of input material and the proportion of sequences obtained per species from metabarcoding, however large uncertainties remain (Lamb et al., 2019). Here, positive correlations were found for some coarse taxonomic groups. A negative correlation was found for ostracods abundances versus sequence reads proportions. It was not the case in Matthews et al. (2021) in which they found a positive correlation between ostracods' relative abundances and proportion of sequences for both 18S V4 and COI primers. It might be due to a technical bias from the metabarcoding process. On the contrary, some groups appeared to be well-suited for a quantitative metabarcoding use. Positive correlations were found for abundance and biomass of Cyclopoida, Doliolida, Harpacticoida and Larvacea. For biomass, similar trends were observed for Cyclopoida, Doliolida and Larvacea (Matthews et al., 2021). The proportion of reads of Chaetognatha and Cladocera well correlated with their abundance but not their biomass. Chaetognatha are not well identified by 18S V1V2 (Matthews et al., 2021) which might explain the absence of correlation between relative number of sequences and biomass. On the contrary, relative number of sequences and biomass was positively correlated for Cephalopoda despite their low abundance. For the remaining taxa, the absence of significant correlations might be due to not-specific or inaccurate conversion factors from biovolume to dry weight. Hence, more studies such as the one by Maas et al. (2021) should be done before assuming a quantitative nature.

4.4 | Conclusion

Morphological changes in the mesozooplankton community were correlated with the ecosystem carbon export intensity, but, contrary to our hypothesis, opacity was the parameter most strongly correlated export, likely due to taxonomical changes in the community composition in response to environmental forcing. Transparency variations throughout time and with respect to carbon export should be investigated further. Additionally, this study demonstrates the benefits of combining a trait-based strategy from quantitative imaging of individual organisms with a molecular approach. Finally, significant correlations were found for some taxonomic groups between proportions of sequences from metabarcoding data and biomass or

abundance, derived from Zooscan images showing potential for this method to be refined in a way that allows for abundance estimation. Numerous efforts should be done in this direction to study the potential of a quantitative molecular approach and develop possible support methods.

AUTHOR CONTRIBUTIONS

SDA, LBB and MP designed research. MP performed research and analysed data. HG and JG carried out laboratory work. MP wrote the paper with the help of SDA and LBB. All co-authors reviewed and approved the manuscript.

ACKNOWLEDGEMENTS

French co-authors wish to thank public taxpayers who fund their salaries. SDA and MP acknowledge funding from the Institut des Sciences du Calcul et des Données (ISCD) of Sorbonne Université through the support of the sponsored project-team FORMAL (From ObseRving to Modeling oceAn Life), the mobility of MP to BIOS, and the PhD funding of MP. SDA, LBB and MP also acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-22-CE02-0023-01 (project TRAITZOO). BIOS/ASU authors acknowledge the funding by the Simons Foundation International as part of the BIOS-SCOPE (https://scope.bios.edu/), the NSF funded OCE-1829318, and the Bermuda Atlantic Time-series Study program (BATS; NSF OCE-1756105). Special thanks to the BATS personnel and to the Captains and crew of the RV Atlantic Explorer for their logistical support at sea. This research made use of the FCMMF laboratory equipment at BIOS (NSF DBI-1522206).

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY AND BENEFIT-SHARING STATEMENT

Raw sequence reads and metadata are deposited in the SRA (BioProject PRJNA928675). Morphological, metabarcoding and environmental data are archived on Pangea (https://doi.org/10.1594/PANGAEA.960038; https://doi.org/10.1594/PANGAEA.960033). R scripts and the bio-informatic pipeline used in this manuscript are available at https://github.com/margauxperhirin/mesozoo_bats. Individual mesozooplankton images, their metadata and morphological descriptors are stored in EcoTaxa web application https://ecotaxa.obs-vlfr.fr/ (project 'BATS_timeseries [4236]'). Benefits Generated: Benefits from this research accrue from the sharing of our data and results on public databases as described above.

ORCID

Margaux Perhirin https://orcid.org/0000-0002-0467-7794
Leocadio Blanco-Bercial https://orcid.org/0000-0003-0658-7183
Sakina-Dorothée Ayata https://orcid.org/0000-0003-3226-9779

REFERENCES

Abad, D., Albaina, A., Aguirre, M., Laza-Martínez, A., Uriarte, I., Iriarte, A., Villate, F., & Estonba, A. (2016). Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. *Marine Biology*, *163*(7), 149. https://doi.org/10.1007/s00227-016-2920-0

Alldredge, A. L. (1984). The quantitative significance of gelatinous zooplankton as pelagic consumers. In M. J. R. Fasham (Ed.), *Flows of energy and materials in marine ecosystems* (pp. 407–433). Springer US. https://doi.org/10.1007/978-1-4757-0387-0_16

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andersen, V., & Sardou, J. (1992). The diel migrations and vertical distributions of zooplankton and micronekton in the northwestern Mediterranean Sea. 1. Euphausiids, mysids, decapods and fishes. *Journal of Plankton Research*, *14*(8), 1129–1154. https://doi.org/10.1093/plankt/14.8.1129

Andersen, V., Sardou, J., & Nival, P. (1992). The diel migrations and vertical distributions of zooplankton and micronekton in the northwestern Mediterranean Sea. 2. Siphonophores, hydromedusae and pyrosomids. *Journal of Plankton Research*, *14*(8), 1155–1169. https://doi.org/10.1093/plankt/14.8.1155

Anderson, M. J. (2001). Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian Journal of Fisheries and Aquatic Sciences*, *58*(3), 626–639. https://doi.org/10.1139/f01-004

Angel, M. V. (1983). A review of the progress of research on halocyprid and other oceanic planktonic ostracods 1972–1982. In R. F. Maddocks (Ed.), *Applications of Ostracoda* (pp. 529–548). Univ. Houston Geosciences.

Barberán, A., Bates, S. T., Casamayor, E. O., & Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal*, *6*(2), 343–351. https://doi.org/10.1038/ismej.2011.119

Barton, A. D., Pershing, A. J., Litchman, E., Record, N. R., Edwards, K. F., Finkel, Z. V., Kiørboe, T., & Ward, B. A. (2013). The biogeography of marine plankton traits. *Ecology Letters*, *16*(4), 522–534. https://doi.org/10.1111/ele.12063

Bellisario, B., Camisa, F., Abbattista, C., & Cimmaruta, R. (2019). A network approach to identify bioregions in the distribution of Mediterranean amphipods associated with Posidonia oceanica meadows. *PeerJ*, *7*, e6786. https://doi.org/10.7717/peerj.6786

Berry, D., & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology*, *5*, 219. https://doi.org/10.3389/fmicb.2014.00219

Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Hauss, H., Gorsky, G., Guidi, L., Kiko, R., & Not, F. (2016). In situ imaging reveals the biomass of giant protists in the global ocean. *Nature*, *532*(7600), 504–507. https://doi.org/10.1038/nature17652

Blanco-Bercial, L. (2020). Metabarcoding analyses and seasonality of the zooplankton community at BATS. *Frontiers in Marine Science*, *7*, 173. https://doi.org/10.3389/fmars.2020.00173

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland Forest communities of southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349. https://doi.org/10.2307/1942268

Brinton, E. (1962). The distribution of Pacific euphausiids. https://escholarship.org/uc/item/6db5n157

Brisbin, M., Brunner, O. D., Grossmann, M. M., & Mitarai, S. (2020). Paired high-throughput, in situ imaging and high-throughput sequencing illuminate acantharian abundance and vertical distribution. *Limnology and Oceanography*, *65*(12), 2953–2965. https://doi.org/10.1002/lno.11567

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, *69*(346), 364–367. https://doi.org/10.1080/01621459.1974.10482955

Brun, P., Payne, M. R., & Kiørboe, T. (2017). A trait database for marine copepods. *Earth System Science Data*, *9*(1), 99–113. https://doi.org/10.5194/essd-9-99-2017

Brüsin, M., Svensson, P. A., & Hylander, S. (2016). Individual changes in zooplankton pigmentation in relation to ultraviolet radiation and predator cues. *Limnology and Oceanography*, *61*(4), 1337–1344. https://doi.org/10.1002/lno.10303

Bucklin, A., Peijnenburg, K., Kosobokova, K., O'Brien, T., Blanco-Bercial, L., Cornils, A., Falkenhaug, T., Hopcroft, R., Hosia, A., Laakmann, S., Li, C., Martell, L., Questel, J., Wall-Palmer, D., Minxiao, W., Wiebe, P., & Weydmann, A. (2021). Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology*, *168*, 78. https://doi.org/10.1007/s00227-021-03887-y

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., & Wiebe, P. H. (2019). Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. *ICES Journal of Marine Science*, *76*(4), 1162–1176. https://doi.org/10.1093/icesjms/fsz021

Castellani, C., & Edwards, M. (2017). *Marine plankton: A practical guide to ecology, methodology, and taxonomy*. Oxford University Press.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*, 1–36. https://doi.org/10.18637/jss.v061.i06

Chen, T., Zhang, Y., Song, S., Liu, Y., Sun, X., & Li, C. (2021). Diversity and seasonal variation of marine phytoplankton in Jiaozhou Bay, China revealed by morphological observation and metabarcoding. *Journal of Oceanology and Limnology*, *40*, 577–591. https://doi.org/10.1007/s00343-021-0457-7

Chust, G., Vogt, M., Benedetti, F., Nakov, T., Villéger, S., Aubert, A., Vallina, S. M., Righetti, D., Not, F., Biard, T., Bittner, L., Benoiston, A.-S., Guidi, L., Villarino, E., Gaborit, C., Cornils, A., Buttay, L., Irisson, J.-O., Chiarello, M., … Ayata, S.-D. (2017). Mare Incognitum: A glimpse into future plankton diversity and ecology research. *Frontiers in Marine Science*, *4*, 68. https://doi.org/10.3389/fmars.2017.00068

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, *18*(1), 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T., & Pawlowski, J. (2018). Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Molecular Ecology Resources*, *18*(6), 1381–1391. https://doi.org/10.1111/1755-0998.12926

Cowen, R., & Guigand, C. (2008). In situ Ichthyoplankton imaging system(ISIIS): System design and preliminary results. *Limnology and Oceanography: Methods*, *6*, 126–132. https://doi.org/10.4319/lom.2008.6.126

Csardi & Nepusz. (2006). *The igraph software package for complex network research* [R]. https://igraph.org

Da Silva, O. (2021). Structure de l'écosystème planctonique: Apport des données à haut débit de séquençage et d'imagerie [Thesis, Sorbonne université]. In http://www.theses.fr. http://www.theses.fr/2021SORUS183

de Vera, A., & Seapy, R. R. (2006). Atlanta selvagensis, a new species of heteropod molluscfrom the northeastern Atlantic Ocean(Gastropoda: Carinarioidea). *Vieraea Folia Scientiarum Biologicarum Canariensium*, *34*(Vieraea 34), 45–54. https://doi.org/10.31939/vieraea.2006.34.06

Deevey, G. B. (1971). The annual cycle in quantity and composition of the zooplankton of the Sargasso Sea off Bermuda. I. The upper 500 M1. *Limnology and Oceanography*, *16*(2), 219–240. https://doi.org/10.4319/lo.1971.16.2.0219

Deevey, G. B., & Brooks, A. L. (1977). Copepods of the Sargasso Sea off Bermuda: Species composition, and vertical and seasonal distribution between the surface and 2000 M. *Bulletin of Marine Science*, *27*(2), 256–291.

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., Muller-Karger, F., & Breitbart, M. (2018). Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. *Limnology and Oceanography: Methods*, *16*(4), 209–221. https://doi.org/10.1002/lom3.10237

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, *56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Durkin, C. A., Cetinić, I., Estapa, M., Ljubešić, Z., Mucko, M., Neeley, A., & Omand, M. (2022). Tracing the path of carbon export in the ocean though DNA sequencing of individual sinking particles. *The ISME Journal*, *1-11*, 1896–1906. https://doi.org/10.1038/s41396-022-01239-2

Edgar, R. C. (2016). UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing (p. 081257). bioRxiv. https://doi.org/10.1101/081257

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, *27*(16), 2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Ershova, E. A., Wangensteen, O. S., Descoteaux, R., Barth-Jensen, C., & Præbel, K. (2021). Metabarcoding as a quantitative tool for estimating biodiversity and relative biomass of marine zooplankton. *ICES Journal of Marine Science*, *78*(9), 3342–3355. https://doi.org/10.1093/icesjms/fsab171

Fonseca, V. G., Carvalho, G. R., Nichols, B., Quince, C., Johnson, H. F., Neill, S. P., Lambshead, J. D., Thomas, W. K., Power, D. M., & Creer, S. (2014). Metagenetic analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes. *Global Ecology and Biogeography*, *23*(11), 1293–1302. https://doi.org/10.1111/geb.12223

Fonseca, V. G., Carvalho, G. R., Sung, W., Johnson, H. F., Power, D. M., Neill, S. P., Packer, M., Blaxter, M. L., Lambshead, P. J. D., Thomas, W. K., & Creer, S. (2010). Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature Communications*, *1*(1), Art. 1. https://doi.org/10.1038/ncomms1095

Goetze, E. (2011). Population differentiation in the Open Sea: Insights from the pelagic copepod Pleuromamma xiphias. *Integrative and Comparative Biology*, *51*(4), 580–597. https://doi.org/10.1093/icb/icr104

González, H. E., & Smetacek, V. (1994). The possible role of the cyclopoid copepod Oithona in retarding vertical flux of zooplankton faecal material. *Marine Ecology-Progress Series - MAR ECOL-PROGR SER*, *113*, 233–246. https://doi.org/10.3354/meps113233

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., Cawood, A., Pesant, S., García-Comas, C., & Prejger, F. (2010). Digital zooplankton image analysis using the ZooScan integrated system. *Journal of Plankton Research*, *32*(3), 285–303. https://doi.org/10.1093/plankt/fbp124

Harrell, F. E., Jr., & Dupon, C. (contributed several functions and maintains latex). (2022). *Hmisc: Harrell Miscellaneous* (4.7–0). https://CRAN.R-project.org/package=Hmisc

Harvey, J. B. J., Johnson, S. B., Fisher, J. L., Peterson, W. T., & Vrijenhoek, R. C. (2017). Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. *Journal of Experimental Marine Biology and Ecology*, *487*, 113–126. https://doi.org/10.1016/j.jembe.2016.12.002

Hays, G. C., Proctor, C. A., John, A. W. G., & Warner, A. J. (1994). Interspecific differences in the diel vertical migration of marine copepods: The implications of size, color, and morphology. *Limnology and Oceanography*, *39*(7), 1621–1629. https://doi.org/10.4319/lo.1994.39.7.1621

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(suppl_1), S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Huo, S., Li, X., Xi, B., Zhang, H., Ma, C., & He, Z. (2020). Combining morphological and metabarcoding approaches reveals the freshwater eukaryotic phytoplankton community. *Environmental Sciences Europe*, *32*(1), 37. https://doi.org/10.1186/s12302-020-00321-w

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P., Endo, H., Gasol, J. M., Gregory, A. C., Mahé, F., Rigonato, J., Royo-Llonch, M., Salazar, G., Sanz-Sáez, I., Scalco, E., Soviadan, D., Zayed, A. A., Zingone, A., … Zinger, L. (2019). Global trends in marine plankton diversity across kingdoms of life. *Cell*, *179*(5), 1084–1097.e21. https://doi.org/10.1016/j.cell.2019.10.008

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L., & Stemmann, L. (2022). Machine learning for the study of plankton and marine snow from images. *Annual Review of Marine Science*, *14*(1), 277–301. https://doi.org/10.1146/annurev-marine-041921-013023

Ivory, J. A., Steinberg, D. K., & Latour, R. J. (2019). Diel, seasonal, and interannual patterns in mesozooplankton abundance in the Sargasso Sea. *ICES Journal of Marine Science*, *76*(1), 217–231. https://doi.org/10.1093/icesjms/fsy117

James and Hornik. (2020). Chron: Chronological objects which can handle dates and times. https://CRAN.R-project.org/package=chron

Johnsen, S. (2014). Hide and seek in the open sea: Pelagic camouflage and visual countermeasures. *Annual Review of Marine Science*, *6*, 369–392. https://doi.org/10.1146/annurev-marine-010213-135018

Johnsen, S., & Widder, E. A. (1998). Transparency and visibility of gelatinous zooplankton from the northwestern Atlantic and Gulf of Mexico. *Biological Bulletin*, *195*(3), 337–348. https://doi.org/10.2307/1543145

Kaeriyama, H., & Ikeda, T. (2002). Vertical distribution and population structure of the three dominant planktonic ostracods (Discoconchoecia pseudodiscophora, Orthoconchoecia haddoni and Metaconchoecia skogsbergi) in the Oyashio region, western North Pacific. 9.

Kassambara, A. (2020). Ggpubr: « ggplot2 » based publication ready plots (0.4.0). https://CRAN.R-project.org/package=ggpubr

Kassambara, A. (2021). Rstatix: Pipe-friendly framework for basic statistical tests (0.7.0). https://CRAN.R-project.org/package=rstatix

Kelly, T. B., Davison, P. C., Goericke, R., Landry, M. R., Ohman, M. D., & Stukel, M. R. (2019). The importance of Mesozooplankton diel vertical migration for sustaining a mesopelagic food web. *Frontiers in Marine Science*, *6*, 508. https://doi.org/10.3389/fmars.2019.00508

Kiørboe, T. (1997). Population regulation and role of mesozooplankton in shaping marine pelagic food webs. *Hydrobiologia*, *363*(1), 13–27. https://doi.org/10.1023/A:1003173721751

Kiørboe, T. (2013). Zooplankton body composition. *Limnology and Oceanography*, *58*(5), 1843–1850. https://doi.org/10.4319/lo.2013.58.5.1843

Kiørboe, T., Andersen, A., Langlois, V. J., & Jakobsen, H. H. (2010). Unsteady motion: Escape jumps in planktonic copepods, their kinematics and energetics. *Journal of the Royal Society Interface*, *7*(52), 1591–1602. https://doi.org/10.1098/rsif.2010.0176

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology*, *28*(2), 420–430. https://doi.org/10.1111/mec.14920

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*, 1–18. https://doi.org/10.18637/jss.v025.i01

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271–280. https://doi.org/10.1007/s004420100716

Legendre, P., & Legendre, L. (2012). Chapter 9—Ordination in reduced space. In P. Legendre & L. Legendre (Eds.), *Developments in environmental modelling* (Vol. *24*, pp. 425–520). Elsevier. https://doi.org/10.1016/B978-0-444-53868-0.50009-5

Lejzerowicz, F., Gooday, A. J., Barrenechea Angeles, I., Cordier, T., Morard, R., Apothéloz-Perret-Gentil, L., Lins, L., Menot, L., Brandt, A., Levin, L. A., Martinez Arbizu, P., Smith, C. R., & Pawlowski, J. (2021). Eukaryotic biodiversity and spatial patterns in the clarion-Clipperton zone and other abyssal regions: Insights from sediment DNA and RNA metabarcoding. *Frontiers in Marine Science*, *8*, 671033. https://doi.org/10.3389/fmars.2021.671033

Levitus, S. (1982). *Climatological atlas of the World Ocean*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

Litchman, E., Ohman, M. D., & Kiørboe, T. (2013). Trait-based approaches to zooplankton communities. *Journal of Plankton Research*, *35*(3), 473–484. https://doi.org/10.1093/plankt/fbt019

Lomas, M., Steinberg, D. K. T. D., Carlson, C., Nelson, N., Condon, R., & Bates, N. (2010). Increased ocean carbon export in the Sargasso Sea linked to climate variability is countered by its enhanced mesopelagic attenuation. *Biogeosciences*, *7*, 57–70.

Lomas, M. W., Bates, N. R., Johnson, R. J., Knap, A. H., Steinberg, D. K., & Carlson, C. A. (2013). Two decades and counting: 24-years of sustained open ocean biogeochemical measurements in the Sargasso Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, *93*, 16–32. https://doi.org/10.1016/j.dsr2.2013.01.008

Lüskow, F., Neitzel, P., Miller, M. J., Marohn, L., Wysujack, K., Freese, M., Pohlmann, J.-D., & Hanel, R. (2019). Distribution and abundance of net-captured calycophoran siphonophores and other gelatinous zooplankton in the Sargasso Sea European eel spawning area. *Marine Biodiversity*, *49*(5), 2333–2349. https://doi.org/10.1007/s12526-019-00971-x

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Maas, A. E., Gossner, H., Smith, M. J., & Blanco-Bercial, L. (2021). Use of optical imaging datasets to assess biogeochemical contributions of the mesozooplankton. *Journal of Plankton Research*, *43*(3), 475–491. https://doi.org/10.1093/plankt/fbab037

MacNeil, L., Desai, D. K., Costa, M., & LaRoche, J. (2022). Combining multi-marker metabarcoding and digital holography to describe eukaryotic plankton across the Newfoundland shelf. *Scientific Reports*, *12*(1), Art. 1. https://doi.org/10.1038/s41598-022-17313-w

Madin, L. P., Horgan, E. F., & Steinberg, D. K. (2001). Zooplankton at the Bermuda Atlantic time-series study (BATS) station: Diel, seasonal and interannual variation in biomass, 1994–1998. *Deep Sea Research Part II: Topical Studies in Oceanography*, *48*(8), 2063–2082. https://doi.org/10.1016/S0967-0645(00)00171-5

Martini, S., Larras, F., Boyé, A., Faure, E., Aberle, N., Archambault, P., Bacouillard, L., Beisner, B. E., Bittner, L., Castella, E., Danger, M., Gauthier, O., Karp-Boss, L., Lombard, F., Maps, F., Stemmann, L., Thiébaut, E., Usseglio-Polatera, P., Vogt, M., … Ayata, S.-D. (2021). Functional trait-based approaches as a common framework for aquatic ecologists. *Limnology and Oceanography*, *66*(3), 965–994. https://doi.org/10.1002/lno.11655

Matthews, S. A., Goetze, E., & Ohman, M. D. (2021). Recommendations for interpreting zooplankton metabarcoding and integrating molecular methods with morphological analyses. *ICES Journal of Marine Science*, *78*(9), 3387–3396. https://doi.org/10.1093/icesjms/fsab107

McHardy, R. A., & Bary, B. M. K. (1965). Diurnal and seasonal changes in distribution of two planktonic ostracods, Conchoecia elegans and Conchoecia alata minor. *Journal of the Fisheries Research Board of Canada*, *22*(3), 823–840. https://doi.org/10.1139/f65-072

Mifsud, C. (2001). A mysterious, living, 'giant' Gymnosomata species near the Maltese Islands (Gastropoda, Opisthobranchia). *Basteria*, *65*, 57–60.

Monferrer, N., Biard, T., Sandin, M. M., Lombard, F., Picheral, M., Elineau, A., Guidi, L., Leynaert, A., Tréguer, P. J., & Not, F. (2022). Siliceous Rhizaria abundances and diversity in the Mediterranean Sea assessed by combined imaging and metabarcoding approaches. *Frontiers in Marine Science*, *9*, 895995. https://doi.org/10.3389/fmars.2022.895995

Motoda, S. (1959). DEVICES OF SIMPLE PLANKTON APPARATUS. 40.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? | SpringerLink. https://doi.org/10.1007/s00357-014-9161-z

Neuwirth, E. (2022). RColorBrewer: ColorBrewer palettes (1.1-3). https://CRAN.R-project.org/package=RColorBrewer

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Oksanen, J., Kindt, R., Legendre, P., Hara, B., Simpson, G., Solymos, P., Henry, M., Stevens, H., Maintainer, H., & Oksanen@oulu, jari. (2009). *The vegan Package*.

Orenstein, E. C., Ayata, S.-D., Maps, F., Becker, É. C., Benedetti, F., Biard, T., de Garidel-Thoron, T., Ellen, J. S., Ferrario, F., Giering, S. L. C., Guy-Haim, T., Hoebeke, L., Iversen, M. H., Kiørboe, T., Lalonde, J.-F., Lana, A., Laviale, M., Lombard, F., Lorimer, T., … Irisson, J.-O. (2022). Machine learning techniques to characterize functional traits of plankton from image data. *Limnology and Oceanography*, *67*(8), 1647–1669. https://doi.org/10.1002/lno.12101

Paffenhöfer, G.-A., Mazzocchi, M. G., & Tzeng, M. W. (2006). Living on the edge: Feeding of subtropical open ocean copepods. *Marine Ecology*, *27*(2), 99–108. https://doi.org/10.1111/j.1439-0485.2006.00086.x

Picheral, M., Colin, S., & Irisson, J. O. (2017). EcoTaxa, a tool for the taxonomic classification of images. URL Httpecotaxa Obs-Vlfr Fr.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., & Gorsky, G. (2010). The underwater vision profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnology and Oceanography: Methods*, *8*, 462–473. https://doi.org/10.4319/lom.2010.8.462

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Quéré, C. L., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., & Wolf-Gladrow, D. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, *11*(11), 2016–2040. https://doi.org/10.1111/j.1365-2486.2005.1004.x

R Core Team. (2021). *R: A language and environment for statistical computing [Logiciel]*. R Foundation for Statistical Computing. https://www.R-project.org/

Rey, A., Corell, J., & Rodriguez-Ezpeleta, N. (2020). Metabarcoding to study zooplankton diversity (p. 252–263). https://doi.org/10.1201/9781351021821-14

Roe, H. S. J., James, P. T., & Thurston, M. H. (1984). The diel migrations and distributions within a mesopelagic community in the north East Atlantic. 6. Medusae, ctenophores, amphipods and euphausiids. *Progress in Oceanography*, *13*(3), 425–460. https://doi.org/10.1016/0079-6611(84)90015-6

Rombouts, I., Beaugrand, G., Ibanez, F., Gasparini, S., Chiba, S., & Legendre, L. (2010). A multivariate approach to large-scale variation in marine planktonic copepod diversity and its environmental correlates. *Limnology and Oceanography*, *55*(5), 2219–2229. https://doi.org/10.4319/lo.2010.55.5.2219

Sampei, M., Forest, A., Sasaki, H., Hattori, H., Makabe, R., Fukuchi, M., & Fortier, L. (2009). Attenuation of the vertical flux of copepod fecal pellets under Arctic Sea ice: Evidence for an active detrital food web in winter. *Polar Biology*, *32*, 225–232. https://doi.org/10.1007/s00300-008-0523-z

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schnetzer, A., & Steinberg, D. (2002). Active transport of particulate organic carbon and nitrogen by vertically migrating zooplankton in the Sargasso Sea. https://doi.org/10.3354/MEPS234071

Sha, Y., Zhang, H., Lee, M., Björnerås, C., Škerlep, M., Gollnisch, R., Herzog, S. D., Ekelund Ugge, G., Vinterstare, J., Hu, N., Pärssinen, V., Hulthén, K., Nilsson, P. A., Rengefors, K., Brönmark, C., Langerhans, R. B., & Hansson, L.-A. (2021). Diel vertical migration of copepods and its environmental drivers in subtropical Bahamian blue holes. *Aquatic Ecology*, *55*(4), 1157–1169. https://doi.org/10.1007/s10452-020-09807-4

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Slowikowski, K., Schep, A., Hughes, S., Dang, T. K., Lukauskas, S., Irisson, J.-O., Kamvar, Z. N., Ryan, T., Christophe, D., Hiroaki, Y., Gramme, P., Abdol, A. M., Barrett, M., Cannoodt, R., Krassowski, M., Chirico, M., & Aphalo, P. (2021). Ggrepel: Automatically position non-overlapping text labels with « ggplot2 » (0.9.1). https://CRAN.R-project.org/package=ggrepel

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. https://doi.org/10.2307/1412159

Stamieszkin, K., Pershing, A. J., Record, N. R., Pilskaln, C. H., Dam, H. G., & Feinberg, L. R. (2015). Size as the master trait in modeled copepod fecal pellet carbon flux. *Limnology and Oceanography*, *60*(6), 2090–2107. https://doi.org/10.1002/lno.10156

Steinberg, D. K., Carlson, C. A., Bates, N. R., Goldthwait, S. A., Madin, L. P., & Michaels, A. F. (2000). Zooplankton vertical migration and the active transport of dissolved organic and inorganic carbon in the Sargasso Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, *47*(1), 137–158. https://doi.org/10.1016/S0967-0637(99)00052-7

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., & Knap, A. H. (2001). Overview of the US JGOFS Bermuda Atlantic time-series study (BATS): A decade-scale look at ocean biology and biogeochemistry. *Deep Sea Research Part II: Topical Studies in Oceanography*, *48*(8), 1405–1447. https://doi.org/10.1016/S0967-0645(00)00148-X

Steinberg, D. K., & Landry, M. R. (2017). Zooplankton and the ocean carbon cycle. *Annual Review of Marine Science*, *9*(1), 413–444. https://doi.org/10.1146/annurev-marine-010814-015924

Steinberg, D. K., Lomas, M. W., & Cope, J. S. (2012). Long-term increase in mesozooplankton biomass in the Sargasso Sea: Linkage to climate and implications for food web dynamics and biogeochemical cycling. *Global Biogeochemical Cycles*, *26*(1). https://doi.org/10.1029/2010GB004026

Steinberg, D. K., Van Mooy, B. A. S., Buesseler, K. O., Boyd, P. W., Kobari, T., & Karl, D. M. (2008). Bacterial vs. zooplankton control of sinking particle flux in the ocean's twilight zone. *Limnology and Oceanography*, *53*(4), 1327–1338. https://doi.org/10.4319/lo.2008.53.4.1327

Stone, J. P., & Steinberg, D. K. (2016). Salp contributions to vertical carbon flux in the Sargasso Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, *113*, 90–100. https://doi.org/10.1016/j.dsr.2016.04.007

Strathmann, R. R. (2006). Versatile ciliary behaviour in capture of particles by the bryozoan cyphonautes larva. *Acta Zoologica*, *87*(1), 83–89. https://doi.org/10.1111/j.1463-6395.2006.00224.x

Svensen, C., & Nejstgaard, J. C. (2003). Is sedimentation of copepod faecal pellets determined by cyclopoids? Evidence from enclosed ecosystems. *Journal of Plankton Research*, *25*(8), 917–926. https://doi.org/10.1093/plankt/25.8.917

Tarrant, A. M., McNamara-Bordewick, N., Blanco-Bercial, L., Miccoli, A., & Maas, A. E. (2021). Diel metabolic patterns in a migratory oceanic copepod. *Journal of Experimental Marine Biology and Ecology*, *545*, 151643. https://doi.org/10.1016/j.jembe.2021.151643

Ursella, L., Cardin, V., Batistić, M., Garić, R., & Gačić, M. (2018). Evidence of zooplankton vertical migration from continuous southern Adriatic buoy current-meter records. *Progress in Oceanography*, *167*, 78–96. https://doi.org/10.1016/j.pocean.2018.07.004

Uye, S., & Kaname, K. (1994). Relations between fecal pellet volume and body size for major zooplankters of the Inland Sea of Japan. *Journal of Oceanography*, *50*, 43–49. https://doi.org/10.1007/BF02233855

Vilgrain, L., Maps, F., Basedow, S., Trudnowska, E., Madoui, M.-A., Niehoff, B., & Ayata, S.-D. (2022). Copepods' true colors: Astaxanthin pigmentation as an indicator of fitness. *Ecosphere*, *14*, e4489.

Vilgrain, L., Maps, F., Picheral, M., Babin, M., Aubry, C., Irisson, J.-O., & Ayata, S.-D. (2021). Trait-based approach using in situ copepod images reveals contrasting ecological patterns across an Arctic ice melt zone. *Limnology and Oceanography*, *66*(4), 1155–1167. https://doi.org/10.1002/lno.11672

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, *116*(5), 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H, Chang, W., Henry, L., Pedersen, T. L., Kohske, T., Claus, W., Kara, W., Hiroaki, Y., Dewey, D., & R Studio Team. (2021). ggplot2: Create elegant data visualisations using the grammar of graphics (3.3.5). https://CRAN.R-project.org/package=ggplot2

Wilke, C. O. (2020). Cowplot: Streamlined plot theme and plot annotations for « ggplot2 » (1.1.1). https://CRAN.R-project.org/package=cowplot

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Perhirin, M., Gossner, H., Godfrey, J., Johnson, R., Blanco-Bercial, L., & Ayata, S.-D. (2024). Morphological and taxonomic diversity of mesozooplankton is an important driver of carbon export fluxes in the ocean. *Molecular Ecology Resources*, *24*, e13907. https://doi.org/10.1111/1755-0998.13907