



OPEN

Macroscale patterns of oceanic zooplankton composition and size structure

Manoela C. Brandão^{1,2,33}✉, Fabio Benedetti^{3,33}✉, Séverine Martini⁴, Yawouvi Dodji Soviadan¹, Jean-Olivier Irisson¹, Jean-Baptiste Romagnan⁵, Amanda Elineau¹, Corinne Desnos¹, Laëtitia Jalabert¹, Andrea S. Freire⁶, Marc Picheral¹, Lionel Guidi¹, Gabriel Gorsky¹, Chris Bowler^{7,8}, Lee Karp-Boss⁹, Nicolas Henry^{8,10}, Colomban de Vargas^{8,10}, Matthew B. Sullivan¹¹, Tara Oceans Consortium Coordinators*, Lars Stemmann^{1,8} & Fabien Lombard^{1,8,12}

Ocean plankton comprise organisms from viruses to fish larvae that are fundamental to ecosystem functioning and the provision of marine services such as fisheries and CO₂ sequestration. The latter services are partly governed by variations in plankton community composition and the expression of traits such as body size at community-level. While community assembly has been thoroughly studied for the smaller end of the plankton size spectrum, the larger end comprises ectotherms that are often studied at the species, or group-level, rather than as communities. The body size of marine ectotherms decreases with temperature, but controls on community-level traits remain elusive, hindering the predictability of marine services provision. Here, we leverage *Tara Oceans* datasets to determine how zooplankton community composition and size structure varies with latitude, temperature and productivity-related covariates in the global surface ocean. Zooplankton abundance and median size decreased towards warmer and less productive environments, as a result of changes in copepod composition. However, some clades displayed the opposite relationships, which may be ascribed to alternative feeding strategies. Given that climate models predict increasingly warmed and stratified oceans, our findings suggest that zooplankton communities will shift towards smaller organisms which might weaken their contribution to the biological carbon pump.

Body size has been defined as a “master trait” for plankton as it is a morphological characteristic shared by organisms across taxonomy and that characterizes the functions performed by organisms in ecosystems^{1,2}. It has a paramount effect on growth, reproduction, feeding strategies and mortality³. One of the oldest manifestations of the biogeography of traits was proposed over 170 years ago, namely Bergmann’s rule, in which field observations showed that larger species tend to be found at higher, colder latitudes⁴.

In the oceans, size is critical in determining trophic links in planktonic ecosystems and is thus a critical factor in regulating the efficiency of the biological carbon pump⁵. Body size is sensitive to changes in temperature due to the thermal dependence of physiological processes⁶. The plankton is mainly composed of ectotherms which are organisms that do not generate sufficient metabolic heat to elevate their body temperature, so their metabolic processes depends on external temperature⁷. Consequently, ectotherms grow more slowly and reach

¹Sorbonne Université, CNRS, Laboratoire d’Océanographie de Villefranche, 06230 Villefranche-sur-mer, France. ²Ifremer, Centre Bretagne, Unité Dynamiques des Ecosystèmes Côtiers, 29280 Plouzané, France. ³ETH Zürich, Institute of Biogeochemistry and Pollutant Dynamics, 8092 Zürich, Switzerland. ⁴Aix Marseille Univ., Université de Toulon, CNRS, IRD, MIO UM 110, 13288 Marseille, France. ⁵Ifremer, Centre Atlantique, Unité Ecologie et Modèles Pour l’Halieutique, 44311 Nantes, France. ⁶Departamento de Ecologia e Zoologia, Universidade Federal de Santa Catarina, Florianópolis 88010970, Brazil. ⁷Institut de Biologie de l’École Normale Supérieure (IBENS), CNRS, INSERM, PSL Université Paris, 75005 Paris, France. ⁸Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Oceans GOSEE, 75016 Paris, France. ⁹School of Marine Sciences, University of Maine, Orono 04469, USA. ¹⁰Sorbonne Université, CNRS, Station Biologique de Roscoff, AD2M, UMR 7144, 29680 Roscoff, France. ¹¹Department of Microbiology and Civil, Environmental, and Geodetic Engineering, The Ohio State University, Columbus 43214, USA. ¹²Institut Universitaire de France, 75231 Paris, France. ³³These authors contributed equally: Manoela C. Brandão and Fabio Benedetti. *List of authors and their affiliations appear at the end of the paper. ✉email: manoelacb1@gmail.com; fabio.benedetti@usys.ethz.ch

maturity at a larger body size in colder environments, which has long puzzled biologists because classic theories of life-history evolution predict smaller adult sizes in environments delaying growth⁸. This pattern of body size variation, known as the temperature-size rule (TSR⁹), has been observed for a wide range of ectotherms, including single-celled and multicellular species, invertebrates and vertebrates^{8,10}.

The processes underlying the inverse relationship between body size and temperature remain to be identified⁸. Despite temperature playing a major role in shaping latitudinal variations in organism size, these patterns may also rely on complex interactions between physical, chemical and biological factors. For instance, oxygen supply plays a central role in determining the magnitude of ectothermic temperature-size responses, but it is hard to disentangle the relative effects of oxygen and temperature from field data because these two variables are often strongly inter-related in the surface ocean^{11,12}.

The major drivers of community-level plankton size structure (i.e. distribution of individual body size in a given community) must be identified to effectively perform the ecological predictions that are progressively requested in a context of climate change¹³. Global patterns of phytoplankton biomass, size and community composition have been extensively studied thanks to satellite sensors that can detect phytoplankton pigments from space. Satellite observations showed that larger phytoplankton dominate in upwelling regions and at high latitudes where seasonal mixing regimes elicit higher macronutrients availability^{14,15}. In contrast, zooplankton size structure and composition remain challenging to study in situ and remain poorly constrained by observations. Body size variations of planktonic copepods have been derived from literature-based relationships and have been found to display latitudinal patterns driven by variations in temperature and primary production¹⁶. Previous studies showed that temperature, rather than food availability, is the dominant variable in explaining variations in copepod body size¹⁷. Body size can be altered experimentally in the laboratory^{18,19}. However, how these species-specific-based and/or laboratory-based observations can be transferred to the size structure of natural communities remains unclear. Knowing how size structure and abundance scale with changing abiotic conditions at the community level is critical because these factors determine the production and the functioning of the entire ecosystem²⁰.

Here, we use plankton samples homogeneously collected at a macroscale during the *Tara* Oceans expeditions (2009–2013) that were analyzed with the ZooScan imaging system²¹ to document how zooplankton composition (i.e., the abundance of different groups) and size structure at the community level varies with latitude, temperature, oxygen, macronutrient concentrations, phytoplankton biomass and other ecosystem properties. We develop multivariate regression models to identify the underlying drivers of the global gradients of abundance and size structure for more than 30 zooplankton clades.

Results

Latitudinal patterns of zooplankton abundance and composition. Based on the Zooscan analysis of the WP2 (200 µm mesh), Bongo (300 µm mesh) and Régent (680 µm mesh) net samples, we found that most of the 36 zooplankton groups retained displayed significant latitudinal patterns of abundance (Fig. 1). Here, we focused on the significant patterns observed for total zooplankton and those broad groups displaying the highest contributions to total abundance (i.e., Copepoda, Rhizaria, Cnidaria, Tunicata, Chaetognatha and Ostracoda plus Cladocera) based on WP2 net samples, as this net showed the broadest spatial coverage (Supplementary Fig. S1). The spatial patterns were nonetheless consistent across all three nets and for all other groups (Supplementary Fig. S2). Total zooplankton and its main constituting groups displayed non-monotonic gradients of abundance with peaks in the Arctic and/or near the equator, and depressions in the tropical gyres. Zooplankton abundance was highest in the Arctic (Fig. 1a,b), north of 60°N, and decreased progressively towards the equator. A secondary peak was visible near the equator because of the relatively higher abundance in the eastern tropical Pacific Ocean. Zooplankton abundance decreased towards the Southern Ocean, whose few sampled stations displayed the lowest abundances.

Gradients in zooplankton abundance were clearly driven by copepods (Fig. 1c,d) as those dominated community composition (74% of total abundance in WP2 samples, 73% and 82% in the Bongo and Régent samples, respectively; Supplementary Doc. S3). Copepods displayed the same abundance pattern as total zooplankton but showed a slightly weaker tropical peak. The latter was actually more marked for other groups, especially the Rhizaria (Fig. 1e,f) that showed very low abundances towards the poles. Gelatinous groups displayed contrasted patterns. Latitudinal gradients were more marked for Tunicata (Fig. 1i,j) and Chaetognatha (Fig. 1k,l) than for total zooplankton as their abundance levels observed in tropical upwelling regions compete with those observed in the Arctic Ocean. Carnivorous jellyfishes (Cnidaria; Fig. 1g,h) displayed a weakly significant latitudinal pattern that was driven by higher abundances in the western Arctic Ocean. Eumalacostraca (i.e., macrozooplankton such as euphausiids, amphipods and decapods) also showed strong bimodal gradient but only in the Régent data, and pteropods showed no distinguishable latitudinal abundance pattern (Supplementary Fig. S2).

Considering the dominance of copepods in terms of abundances in the communities sampled, we examined the underlying latitudinal gradients in copepod order and family composition (Fig. 2). All nets (Fig. 2a–c) showed an increase in the relative contribution of calanoid families, and especially the large-bodied Calanidae, to the detriment of Cyclopoida (Oithonidae) and Poecilostomatoida (Oncaeiidae, Corycaeiidae and Sapphirinidae). The relative abundances of copepod families were more evenly distributed in the tropics than in the poles, reflecting gradients of decreasing copepod diversity with latitude (already documented by Ibarbalz et al.²²). The variations in copepod abundance were driven by the increase in calanoids (mainly Calanidae) and oithonids towards the Arctic Ocean (Supplementary Fig. S2). Conversely, the following families showed clear abundance peaks in tropical regions (gyres or upwelling): Augaptilidae, Candaciidae, Corycaeiidae, Eucalanidae, Euchaetidae, Oncaeiidae, Paracalanidae, Sapphirinidae and Temoridae. The WP2 data showed less marked variations (Fig. 2a) as this net better samples the smaller Poecilostomatoida and Cyclopoida. The Régent net captured a lower quantity of

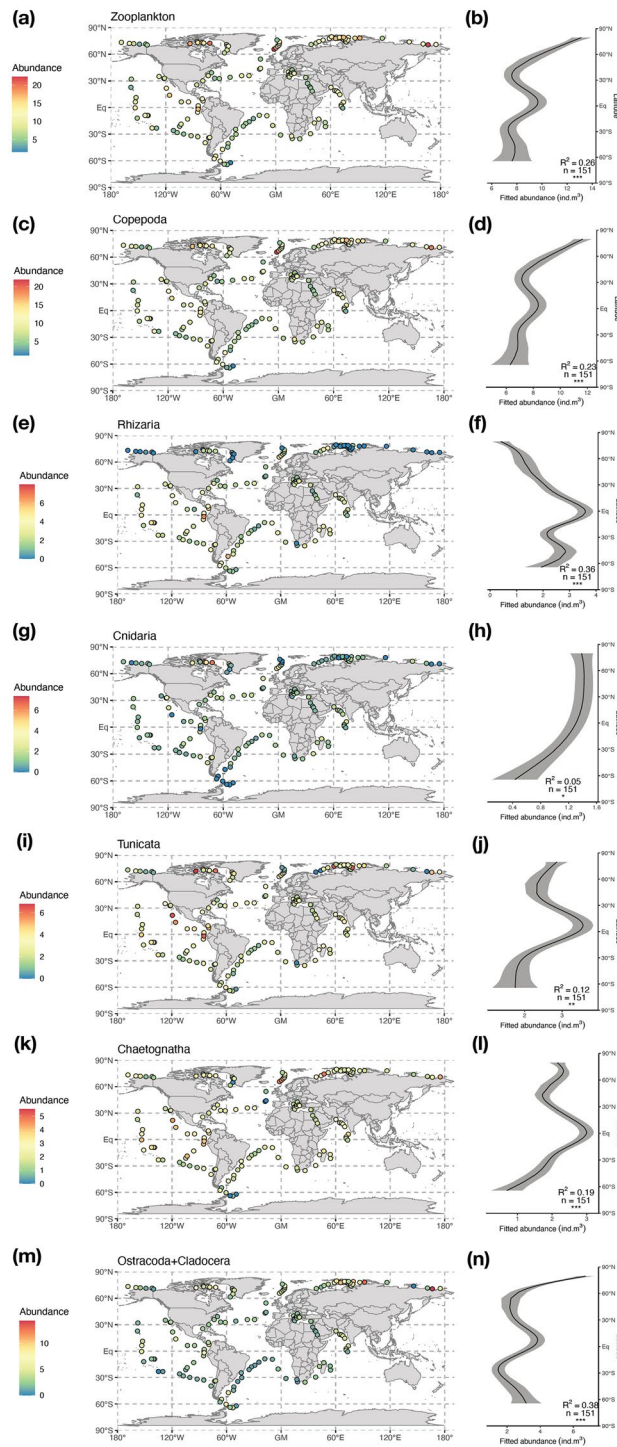


Figure 1. Maps and latitudinal patterns of the abundance (cubic-transformed ind m³) of **(a,b)** Total zooplankton, **(c,d)** Copepoda, **(e,f)** Rhizaria, **(g,h)** Cnidaria, **(i,j)** Tunicata, **(k,l)** Chaetognatha, and **(m,n)** Ostracoda + Cladocera observed in samples collected by the WP2 net. The solid curves on the right-hand side plots illustrate the prediction from the Generalized Additive Model (GAM) fitting abundance against latitude. The explanatory power of the GAM (adjusted R²), the number of samples used and the significance of the smooth term (p < 0.001 = ***, p < 0.01 = **, p < 0.05 = *, p > 0.05 = ns) are reported on the plots. The grey ribbon illustrates the standard error of the GAM prediction.

unidentified Calanoida (Fig. 2c) as the WP2 and Bongo nets (Fig. 2b) as the relatively coarse mesh of this net is not able to retain smaller organisms, a pattern that was found across all zooplankton groups (Supplementary

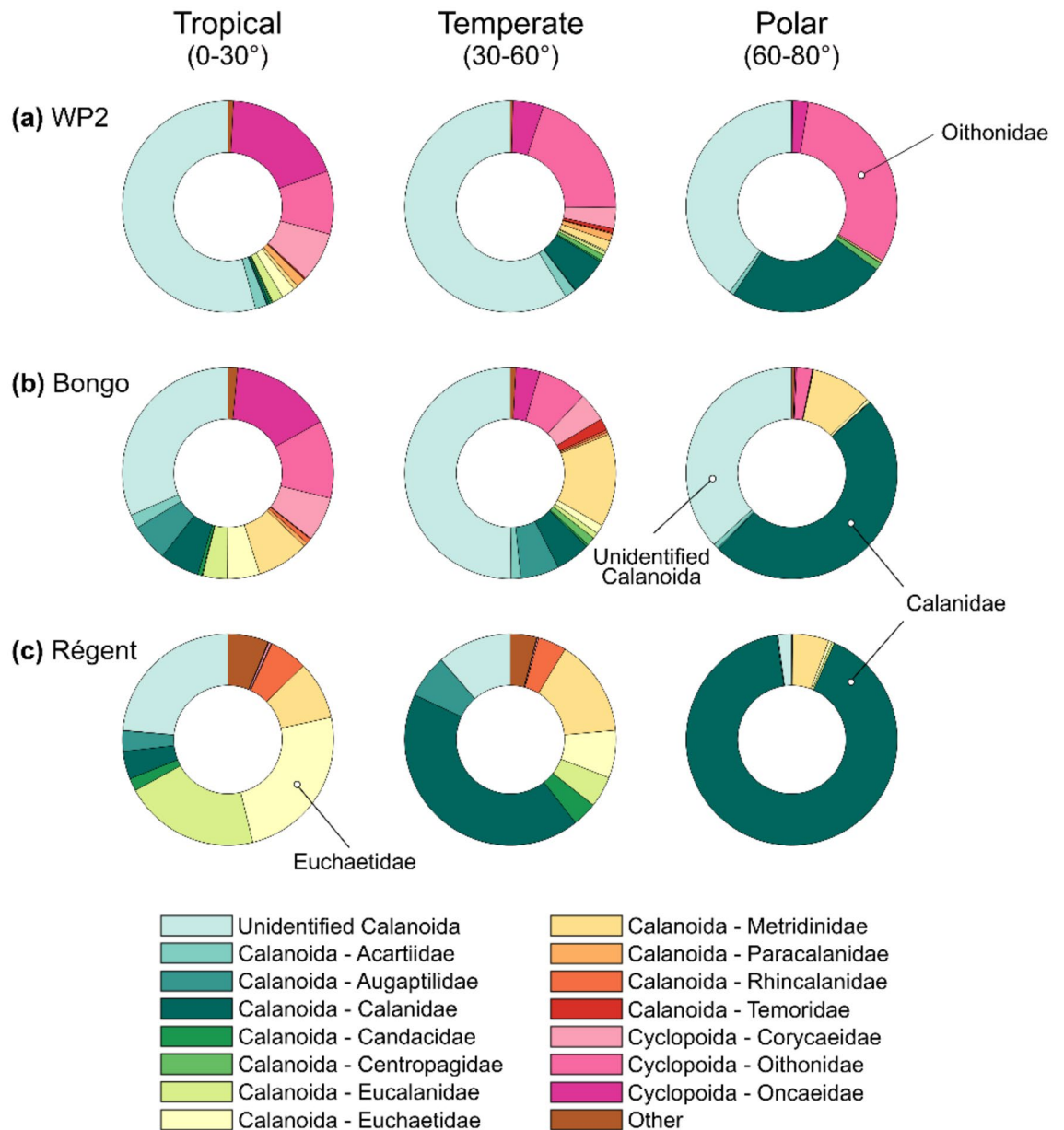


Figure 2. Variations in Copepoda community composition across the tropical (0–30°), temperate (30°–60°) and polar (>60°) latitudinal bands, depicted through the changes in relative abundances of the copepod Orders (Calanoida, Cyclopoida and Poecilostomatoida) and Families sampled by the (a) Bongo net, (b) WP2 net, and (c) Régent net. Taxa with lower than 1% are not shown. Unidentified categories correspond to those organisms that could be assigned to an Order but not to a Family because of the limited resolution of the imaging system.

Doc. S4). The Bongo net showed lower zooplankton abundances than the WP2 net because of its coarser mesh (1.5 times coarser), yet the global patterns in abundances between these two nets showed relatively high positive correlations ($\rho > 0.4$) for several of the main zooplankton groups (e.g. Total zooplankton, Copepoda, Rhizaria, Eumalacostraca, and Ostracoda + Cladocera; Supplementary Doc. S4). Differences between the WP2 samples and the Régent samples were more marked as the latter was equipped with a mesh 3.4 times larger than the former. Only the abundances of total zooplankton, Cnidaria and Eumalacostraca showed relatively high correlations to the WP2 data.

Latitudinal patterns of zooplankton size structure. Variations in median Equivalent Spherical Diameter (ESD) were explored to examine latitudinal patterns in zooplankton size structure. The most consistent cross-net patterns of median ESD were found for the total zooplankton community, which was driven by the median ESD of calanoid copepods (Fig. 3; see Supplementary Fig. S5 for the other groups). The most prominent feature of the copepod median ESD pattern was a sharp decline from the Arctic to the equator, which was more marked in the Bongo (Fig. 3b) and Régent data (Fig. 3c) than in the WP2 (Fig. 3a). In the southern hemisphere, patterns differed across nets: copepod median ESD sharply increased towards the Southern Ocean

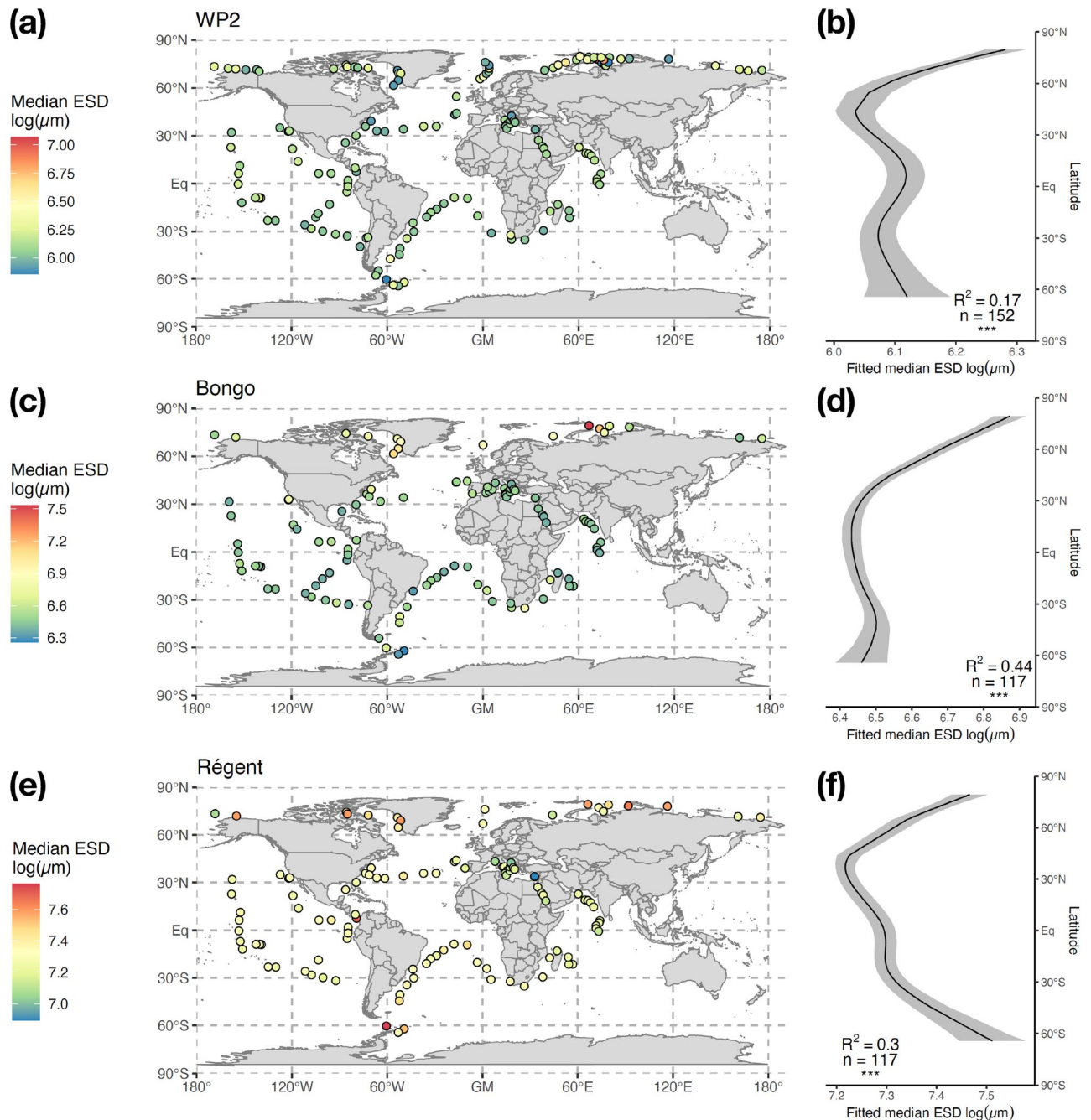


Figure 3. Maps and latitudinal patterns of the logged median Equivalent Spherical Diameter (ESD, μm) observed for Copepoda based on (a,b) WP2 samples (200 μm mesh), (c,d) Bongo samples (300 μm mesh) and (e,f) Régent samples (680 μm mesh). The major and minor axes of the best fitting ellipses were measured for each organism to estimate their ESD. Community-level size structure was determined through the median value of the ESD distribution at individual-level. The solid curves in the right-hand side plots illustrate the prediction from the Generalized Additive Model (GAM) fitting median ESD as a function of latitude. The explanatory power of the GAM (adjusted R^2), the number of samples used and the significance of the smooth term ($p < 0.001 = \text{***}$, $p < 0.01 = \text{**}$, $p < 0.05 = *$, $p > 0.05 = \text{ns}$) are reported on the plots. The grey ribbon illustrates the standard error of the prediction. Only the stations where ESD was measured for at least 20 individuals were considered.

according to the Régent net, whereas it showed no variations or a slight decrease according to the WP2 and Bongo samples, respectively. Considering the relatively poor coverage of the Southern Ocean by *Tara* Oceans, these latter patterns should be interpreted with caution. According to the WP2, the net that best sampled the smaller Poecilostomatoida, the latter showed median ESD patterns that were opposite to the Calanoida: their

median ESD clearly increased from the poles to the tropics and peaked in the southern hemisphere around 30°S (Supplementary Fig. S5).

Contrary to abundances, a secondary tropical peak in median ESD was not observed for zooplankton (Supplementary Fig. S5). Abundance and median ESD were significantly positively correlated for total zooplankton in the WP2 and Régent data, and for the Copepoda and Calanoida in all nets (Supplementary Table S6).

Among non-copepod groups, the Cnidaria also showed a sharp decrease in median ESD from the Arctic Ocean to the equator in both WP2 and Régent samples (Supplementary Fig. S5). The median ESD of Rhizaria followed the opposite pattern according to the WP2 and Bongo samples as it peaked around 40°N and decreased towards lower latitudes. Our approach did not detect clear latitudinal gradients in median ESD for most of the other zooplankton groups (Supplementary Table S7), either because of insufficient observations or because median ESD is not controlled by factors that vary latitudinally. Therefore, we examined the potential environmental drivers of median ESD variations to help us explain why size structure estimates display less marked latitudinal patterns.

Relationships with environmental covariates. The strength of the linear covariance between the groups' abundance, median ESD and environmental covariates was examined through non parametric correlation coefficients (Fig. 4; Supplementary Fig. S8). The median ESD of most zooplankton groups displayed similar significant correlation patterns across nets: the median ESD of total zooplankton, Copepoda, Calanoida, Cnidaria and Eumalacostraca decreased with temperature, salinity and picophytoplankton (%Pico), but increased weakly with oxygen, chlorophyll a, macronutrient concentrations, microphytoplankton (%Micro) and the intensity of particles backscattering (bbp470). Total zooplankton median ESD decreased significantly with Mixed Layer Depth (MLD) only in the WP2 samples (Fig. 4a), a pattern driven by the Calanoida. The median ESD of the Poecilostomatoida increased with temperature, salinity, %Pico and Photosynthetically Active Radiation (PAR). We also found PAR to be the main covariate associated with a lower median ESD of Rhizaria according to the Régent data (Fig. 4c). The Rhizaria showed less significant correlations but differed from the main pattern as their median ESD slightly increased with %Micro and decreases with %Pico and PAR. The only groups displaying a similar pattern were the Tunicata and to a lesser extent the Chaetognatha (Supplementary Fig. S8).

Zooplankton abundances displayed stronger correlation patterns than median ESD (Supplementary Fig. S8) and seem to be more strongly linked to productivity-related covariates (i.e. chlorophyll a, bbp470, %Micro, %Pico and macronutrient concentrations) than physical ones (i.e. temperature and oxygen). The abundance of most groups increased significantly with chlorophyll a, macronutrient concentrations, %Micro and bbp47, but decreased with %Pico. The abundance of some groups presented correlation patterns that departed from the abovementioned trend as they increased with temperature and decreased with oxygen (Supplementary Fig. S8): Rhizaria (WP2 and Régent), Eumalacostraca (Régent mainly), Chaetognatha and Poecilostomatoida (WP2 only).

Nonlinear relationships between median ESD estimates and a subset of environmental covariates were explored through Generalized Additive Models (GAMs, see "Methods") to identify and rank the drivers of size structure of zooplankton groups (Table 1). In total, 102 GAMs were fitted to median ESD estimates ($n = 40$ for the WP2 and Bongo data, $n = 22$ for the Régent; Supplementary Table S7). These GAMs showed reasonable to good fit as the median (\pm IQR) %Dev was 53.6% (\pm 33.4%). The GAMs based on the Régent observations displayed significantly higher %Dev ($57.9\% \pm 24.7\%$) than those based on the WP2 ($55.4\% \pm 31.3\%$) and Bongo ($48.7\% \pm 34.9\%$) (Kruskal–Wallis test, $\text{Chi}^2 = 143.6$, $p < 2.2 \times 10^{-16}$). The GAMs including temperature did not show higher %Dev than those including oxygen except with the Régent data but the difference was found to be marginal ($\text{Chi}^2 = 19.1$, $p = 1.3 \times 10^{-5}$). Substantial variations in smooth term rankings were visible across nets and zooplankton groups (Table 1; Supplementary Fig. S9). Oxygen and temperature were the two top-ranking significant covariates, while the remaining eight covariates displayed lower median ranks (Supplementary Fig. S9) though some (e.g., salinity, MLD, chlorophyll a or %Micro) emerged as key covariates for modelling the median ESD of some groups (Table 1).

The smoothing curves of the GAMs displaying a %Dev > 50% were extracted to cluster the groups based on the shape of these curves along with each covariate (see Methods). This way, we were able to identify clusters of zooplankton groups displaying similar functional responses to the covariates selected (i.e. zooplankton groups sharing similar drivers of global median ESD), and we could project their similarities in a two dimensional metric dimensional scaling (MDS) space to summarize the main trends. Four clusters were identified (Table 1 and Fig. 5). Cluster 1 comprised six models with a mix of Bongo and Régent observations: the median ESD of total zooplankton and the Calanoida (Bongo), Copepoda and Calanoida (Régent) and Cladocera + Ostracoda (Bongo). This cluster gathered groups whose median ESD showed linear increases with oxygen and PAR and no response to temperature (Supplementary Fig. S10). The smoothing curves modelled for the other covariates were either non-significant or highly variable between groups (Fig. 5; Supplementary Fig. S10). The smoothing curves of the zooplankton WP2 data and its main driving group (i.e. calanoid copepods) were clustered with the Tunicata and Eumalacostraca (WP2), the Copepoda (Bongo) and the Chaetognatha (Régent). Contrary to cluster 1, these groups displayed non linear decreases in median ESD with temperature and relatively strong non linear increases with oxygen. Cluster 3 was the largest as it comprised nine models from various groups and nets: Cyclopoida, Poecilostomatoida and the Rhizaria (all WP2), the Eumalacostraca (Bongo and Régent), the Cladocera + Ostracoda (Régent) and the gelatinous zooplankton (Cnidaria, Tunicata and Chaetognatha) sampled with the Bongo net. Because of the cluster's larger size, the response curves modelled for these groups were diverse. The main trend was an overall non linear decrease in median ESD with oxygen concentration. Finally, Cluster 4 gathered a single model (Cnidaria, WP2) meaning it displayed an original combination of modelled response curves. The median ESD of the Cnidaria (WP2) decreased linearly with temperature and increased non

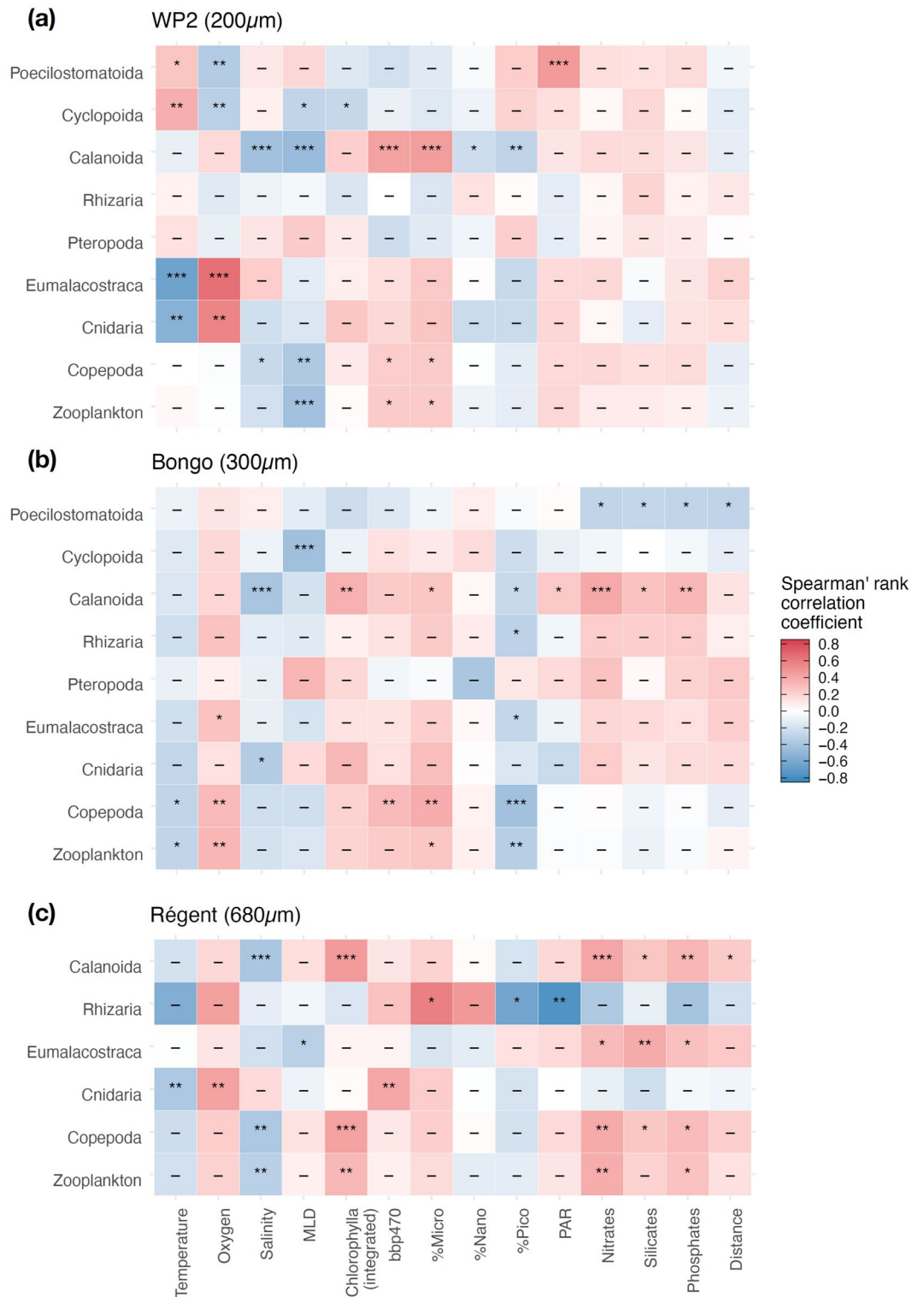


Figure 4. Heatmaps of the Spearman's rank correlation coefficients computed between the size structure (i.e., logged median Equivalent Spherical Diameter; ESD) of the main zooplankton groups and the selected 14 covariates depicting the environmental conditions in the global surface ocean as sampled by (a) WP2 net (200 μ m mesh), (b) Bongo net (300 μ m mesh) and (c) Régent net (680 μ m). The significance of the Spearman's rank correlation tests are reported in the tiles ($p < 0.001 = ***$, $p < 0.01 = **$, $p < 0.05 = *$, $p > 0.05 = ns$). Only the zooplankton groups displaying significant correlation coefficients for more than one environmental covariate in at least one net parameter are shown (see Supplementary Fig. S8 for all groups). Only the stations where ESD was measured for at least 20 individuals of a group were considered when computing the correlation coefficients. Distance stands for distance to coast (in km).

Group (median ESD)	Net	First term	Deviance explained (%)	Significant smooth terms ($p < 0.05$)	Cluster (PAM based on DTW)—only for models with Deviance > 40%
Zooplankton	WP2	Oxygen	0.59	Oxygen, %Micro, %Nano	2
		Temperature	0.54	Temperature, %Micro, %Nano	
	Bongo	Oxygen	0.71	Oxygen, Nitrates, bbp470, %Nano	1
		Temperature	0.72	Temperature, Salinity, Nitrates, bbp470, %Micro, %Nano	
	Régent	Oxygen	0.23	Salinity, bbp470	–
		Temperature	0.23	Salinity, bbp470, %Micro	–
Copepoda	WP2	Oxygen	0.68	Oxygen, PAR, %Micro	2
		Temperature	0.64	Temperature %Micro	
	Bongo	Oxygen	0.89	Oxygen, %Micro, %Nano, Distance to coast	2
		Temperature	0.88	Temperature, %Micro, %Nano, Distance to coast	
	Régent	Oxygen	0.59	Oxygen, Salinity, MLD, PAR	1
		Temperature	0.57	Temperature, Salinity, MLD, PAR	
Rhizaria	WP2	Oxygen	0.76	Oxygen, Salinity, MLD, PAR, Nitrates, bbp470, Chlorophylla, %Nano	3
		Temperature	0.77	Temperature, Salinity, MLD, PAR, Nitrates, bbp470, Chlorophylla, %Nano	
	Bongo	Oxygen	0.47	Oxygen, Salinity	1
		Temperature	0.51	Temperature, Salinity	
Cnidaria	WP2	Oxygen	0.98	Oxygen, Salinity, MLD, PAR, bbp470, %Micro, %Nano, Distance to coast	4
		Temperature	0.96	Temperature, MLD, Nitrates, %Micro, Distance to coast	
	Bongo	Oxygen	0.77	PAR, Nitrates, Chlorophylla, %Nano, Distance to coast	3
		Temperature	0.78	PAR, Nitrates, Chlorophylla, %Nano, Distance to coast	
	Régent	Oxygen	0.22	Oxygen	–
		Temperature	0.20	–	–
Tunicata	WP2	Oxygen	0.64	Oxygen, %Micro, %Nano	2
		Temperature	0.59	Temperature, %Micro, %Nano	
	Bongo	Oxygen	0.71	Oxygen, Salinity, MLD, PAR, %Nano	3
		Temperature	0.70	Temperature, PAR, %Nano	
Eumalacostraca	WP2	Oxygen	0.79	Oxygen, Nitrates	2
		Temperature	0.87	Temperature, Nitrates, bbp470, Chlorophylla, %Nano, Distance to coast	
	Bongo	Oxygen	0.54	PAR, Nitrates, %Micro, %Nano	3
		Temperature	0.53	Temperature, PAR, Nitrates, %Nano	
	Régent	Oxygen	0.71	Oxygen, MLD, PAR, Chlorophylla, %Nano	3
		Temperature	0.75	Temperature, MLD, PAR, bbp470, Chlorophylla, %Nano	
Pteropoda	WP2	Oxygen	0.33	Nitrates, bbp470	–
		Temperature	0.33	Nitrates, bbp470	–
Chaetognatha	WP2	Oxygen	0.32	PAR, Chlorophylla	–
		Temperature	0.32	PAR, Chlorophylla	–
	Bongo	Oxygen	0.68	Salinity, bbp470, %Micro, %Nano	3
		Temperature	0.68	Salinity, bbp470, %Micro, %Nano	
	Régent	Oxygen	0.54	Oxygen, Chlorophylla	2
		Temperature	0.60	Temperature, Chlorophylla	
Calanoida	WP2	Oxygen	0.73	Oxygen, PAR, Chlorophylla, %Micro	2
		Temperature	0.73	Temperature, %Micro	
	Bongo	Oxygen	0.50	Oxygen, Nitrates, Chlorophylla	1
		Temperature	0.39	Temperature, Salinity, PAR	
	Régent	Oxygen	0.60	Oxygen, Salinity, MLD, PAR, Nitrates	1
		Temperature	0.58	Temperature, Salinity, MLD, PAR	
Poecilostomatoida	WP2	Oxygen	0.51	MLD, PAR, Distance to coast	3
		Temperature	0.51	MLD, PAR, Distance to coast	
	Bongo	Oxygen	0.39	Oxygen, %Micro, Distance to coast	–
		Temperature	0.35	Temperature, %Micro, Distance to coast	–
Continued					

Group (median ESD)	Net	First term	Deviance explained (%)	Significant smooth terms ($p < 0.05$)	Cluster (PAM based on DTW)—only for models with Deviance > 40%
Cyclopoida	WP2	Oxygen	0.41	Oxygen, MLD	3
		Temperature	0.41	Temperature, MLD	
	Bongo	Oxygen	0.32	MLD	–
		Temperature	0.32	MLD	–

Table 1. Summary of the explanatory power (i.e. % of deviance explained) of the Generalized Additive Models (GAMs) fit to model the global gradients log-transformed Equivalent Spherical Diameter (ESD, μm) as a function of the ten environmental covariates selected, measured for the zooplankton groups and for the plankton nets that sampled enough stations (> 30) and enough individuals (> 20) per group. The significant covariates ($p < 0.05$) were ranked based on their relative F statistic and are shown. The GAMs displaying a % of deviance explained > 40% were clustered into four groups based on the shape of the smoothing curves of each covariate (Fig. S10).

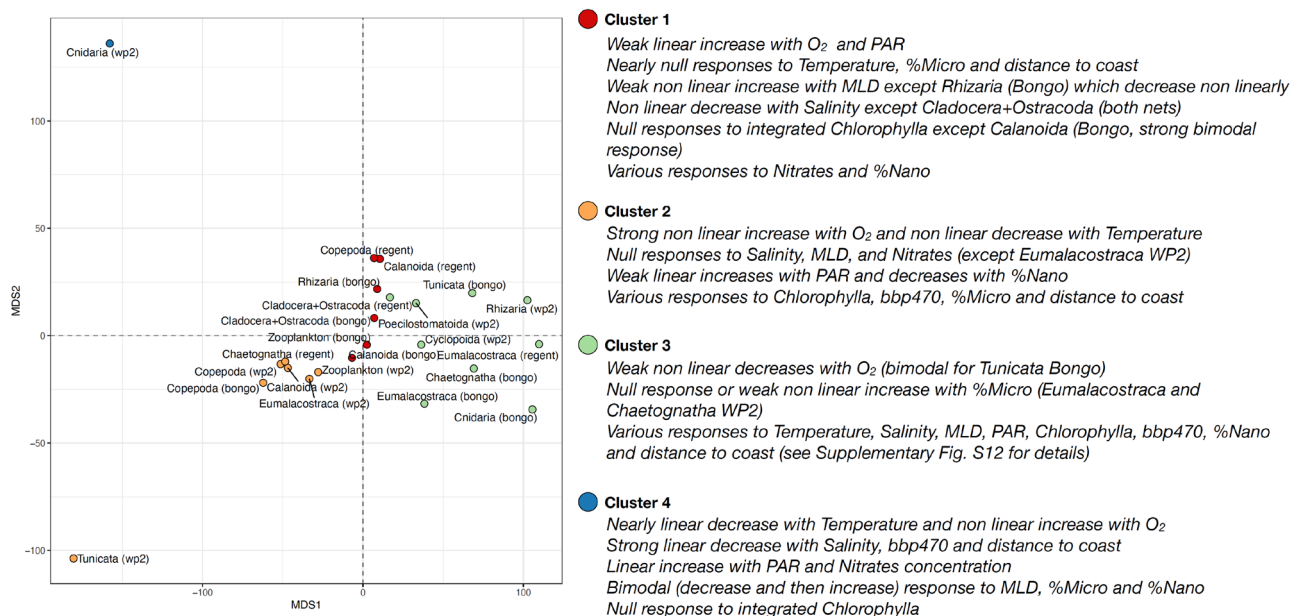


Figure 5. Two dimensional metric dimensional scaling (MDS) plot illustrating the similarity between the responses of the groups' median ESD to the environmental covariates selected. The smoothing curves from the Generalized Additive Models (GAMs) modelling the global gradients in log-transformed median Equivalent Spherical Diameter (ESD, μm) of the zooplankton groups (estimated for various plankton nets) as a function of ten environmental covariates and displaying a deviance explained > 40%. The smoothing curves were combined into a multivariate data series to compute Dynamic Time Warping (DTW) distances and perform partitioning around medoids (PAM) clustering. This way the GAMs were clustered into four clusters representing combinations of zooplankton groups and plankton nets that exhibit similar median ESD-covariate relationships.

linearly with oxygen, and it departed from the other groups because of its strong linear decreases with salinity, particles backscattering and distance to coast (Supplementary Fig. S10).

The same approach was applied to investigate the drivers of global abundance patterns (Table 2; Supplementary Table S7). The median ESD-based GAMs display higher %Dev than the abundance-based GAMs ($40.7\% \pm 27.3\%$; $\text{Chi}^2 = 1697.3.6$, $p < 2.2 \times 10^{-16}$) whatever the net, despite the higher number of observations available for modelling abundances (i.e. 200 GAMs were fitted based on the transformed abundance data). The WP2-based GAMs presented slightly higher %Dev ($\text{Chi}^2 = 62.9$, $p = 2.2 \times 10^{-14}$) than the ones based on the Régent and Bongo observations. The GAMs including temperature displayed a lower %Dev than those including oxygen (37.9 ± 25.8 versus 42.5 ± 27.2 ; $\text{Chi}^2 = 144.1$, $p < 2.2 \times 10^{-16}$). Contrary to median ESD-GAMs, the inclusion of oxygen instead of temperature substantially increased the %Dev for total zooplankton and Calanoida, and Cyclopoida (Table 2), implying that oxygen could be a stronger driver than temperature for zooplankton abundances. Again, the smooth terms associated with temperature and oxygen emerged as the two most significant terms (Table 2; Supplementary Fig. S9). Substantial variations in smooth terms rankings were observed across nets and groups again (Table 2). However, NO_2NO_3 concentrations, chlorophyll a and MLD showed higher significance rankings than in the ESD-based GAMs (Table 2), implying these covariates were more critical to include when modelling zooplankton abundance than size structure.

Again, the smooth curves of the GAMs displaying a %Dev > 40% were extracted to cluster the zooplankton groups based on the similarity of their responses to the covariates. Four clusters could be identified and these are

Group (abundance)	Net	First term	Deviance explained (%)	Significant smooth terms ($p < 0.05$)	Cluster (PAM based on DTW)—only for models with Deviance > 40%
Zooplankton	WP2	Oxygen	0.67	Oxygen, MLD, Nitrates, Chlorophylla, %Nano	4
		Temperature	0.26	Salinity, MLD, Chlorophylla	
	Bongo	Oxygen	0.40	Oxygen, bbp470, %Micro	2
		Temperature	0.38	Temperature, bbp470, %Micro	
	Régent	Oxygen	0.87	Oxygen, PAR, Nitrates, Chlorophylla, %Nano, Distance to coast	1
		Temperature	0.81	Temperature, Nitrates, %Micro, %Nano	
Copepoda	WP2	Oxygen	0.67	Oxygen, MLD, Nitrates, Chlorophylla, %Nano	4
		Temperature	0.30	MLD, Chlorophylla	
	Bongo	Oxygen	0.41	Oxygen, bbp470, %Micro	2
		Temperature	0.38	Temperature, bbp470, %Micro	
	Régent	Oxygen	0.88	Oxygen, Nitrates, Chlorophylla, %Micro, %Nano, Distance to coast	1
		Temperature	0.86	Temperature, Nitrates, %Micro, %Nano, Distance to coast	
Rhizaria	WP2	Oxygen	0.32	Oxygen, Salinity, Nitrates	–
		Temperature	0.32	Temperature, Salinity, Nitrates	
	Bongo	Oxygen	0.49	Oxygen, Nitrates	2
		Temperature	0.48	Nitrates	
	Régent	Oxygen	0.44	Oxygen, MLD, PAR, bbp470, %Micro	2
		Temperature	0.44	Temperature, PAR, bbp470, %Micro	
Cnidaria	WP2	Oxygen	0.35	Nitrates, Chlorophylla	–
		Temperature	0.35	Nitrates, Chlorophylla	
	Bongo	Oxygen	0.41	Salinity, MLD, Nitrates	2
		Temperature	0.39	Salinity, MLD, Nitrates	
	Régent	Oxygen	0.21	Distance to coast	–
		Temperature	0.18	Distance to coast	
Tunicata	WP2	Oxygen	0.40	Oxygen, MLD, Nitrates	3
		Temperature	0.38	Temperature, Nitrates	
	Bongo	Oxygen	0.57	MLD, Nitrates, Chlorophylla	2
		Temperature	0.57	MLD, Nitrates, Chlorophylla	
	Régent	Oxygen	0.22	–	–
		Temperature	0.22	–	
Eumalacostraca	WP2	Oxygen	0.16	Salinity	–
		Temperature	0.16	Salinity	
	Bongo	Oxygen	0.24	Nitrates	–
		Temperature	0.24	Nitrates	
	Régent	Oxygen	0.42	Chlorophylla, %Micro, Distance to coast	2
		Temperature	0.41	Temperature, Chlorophylla	
Pteropoda	WP2	Oxygen	0.33	Oxygen, Nitrates, Chlorophylla, Distance to coast	–
		Temperature	0.32	Temperature, Nitrates, Chlorophylla, Distance to coast	
	Bongo	Oxygen	0.29	Oxygen, Chlorophylla	–
		Temperature	0.28	Temperature, Chlorophylla	
	Régent	Oxygen	0.29	%Micro	–
		Temperature	0.27	%Micro	
Chaetognatha	WP2	Oxygen	0.52	Oxygen, PAR, Chlorophylla, %Nano	2
		Temperature	0.19	Nitrates, %Nano	
	Bongo	Oxygen	0.46	Oxygen, Nitrates, bbp470	2
		Temperature	0.27	bbp470, Chlorophylla	
	Régent	Oxygen	0.43	Nitrates, bbp470, Chlorophylla, %Micro, Distance to coast	2
		Temperature	0.44	Nitrates, bbp470, Chlorophylla, %Micro, Distance to coast	
Continued					

Group (abundance)	Net	First term	Deviance explained (%)	Significant smooth terms ($p < 0.05$)	Cluster (PAM based on DTW)—only for models with Deviance > 40%
Calanoida	WP2	Oxygen	0.69	Oxygen, MLD, Chlorophylla, %Micro, %Nano	4
		Temperature	0.35	MLD, Chlorophylla, %Micro	
	Bongo	Oxygen	0.38	Oxygen, bbp470, %Micro	–
		Temperature	0.37	Temperature, bbp470, %Micro	
	Régent	Oxygen	0.88	Oxygen, Nitrates, Chlorophylla, %Micro, %Nano, Distance to coast	1
		Temperature	0.86	Temperature, Nitrates, %Micro, %Nano	
Poecilostomatoida	WP2	Oxygen	0.53	Oxygen, MLD, Nitrates, Chlorophylla	3
		Temperature	0.51	Temperature, MLD, PAR, Nitrates, Chlorophylla	
	Bongo	Oxygen	0.65	Oxygen, Salinity, MLD, Nitrates	3
		Temperature	0.66	Temperature, Salinity, MLD, Nitrates	
	Régent	Oxygen	0.60	Oxygen, PAR, Distance to coast	2
		Temperature	0.65	Temperature, Salinity, PAR, Distance to coast	
Cyclopoida	WP2	Oxygen	0.64	Oxygen, MLD, PAR, bbp470, Chlorophylla, %Nano	2
		Temperature	0.49	MLD, PAR	
	Bongo	Oxygen	0.45	Salinity, Nitrates, %Nano	2
		Temperature	0.64	Salinity, Nitrates, bbp470, %Nano	
	Régent	Oxygen	0.62	Oxygen, PAR, %Micro	2
		Temperature	0.31	bbp470	

Table 2. Summary of the explanatory power (i.e. % of deviance explained) of the Generalized Additive Models (GAMs) fit to model the global gradients cubic-transformed abundance (ind m^3) as a function of the ten environmental covariates selected, measured for the zooplankton groups and for the plankton nets that sampled enough stations (> 30) and enough individuals (> 20) per group. The significant covariates ($p < 0.05$) were ranked based on their relative F statistic and are shown. The GAMs displaying a % of deviance explained > 40% were clustered into five groups based on the shape of the smoothing curves of each covariate (Fig. S11).

more clearly delineated than those based on the median ESD response curves as evidenced by the relatively more scattered positions of the groups in MDS space (Fig. 6; Supplementary Fig. S11). Cluster 1 gathered the smooth curves modelled for total zooplankton, Copepoda and Calanoida based on the Régent data. Their abundances showed: (i) a strong nonlinear decrease with temperature and increase with oxygen concentration, (ii) non linear decreases with salinity and distance to coast, and (iii) slight increases with PAR and NO_2NO_3 concentrations. Cluster 2 was the largest clusters as it gathered the responses of diverse range of 17 different models based on various groups and nets. This implies that a relative broad ranges of abundances responses within this cluster (Supplementary Fig. S11), which is why it holds a relatively neutral central position in the MDS space (Fig. 6). Yet, nearly all groups showed null responses in abundances to temperature, except total zooplankton and Copepoda (Bongo data) which showed non linear decreases. Cluster 3 was also a smaller cluster composed of three models only: the Poecilostomatoida (both WP2 and Bongo) and the Tunicata (WP2 only). Contrary to clusters 1 and 2, these were characterized by non linear increases in abundance with temperature and NO_2NO_3 concentrations but decreases with oxygen and MLD. This is why these groups are positioned on the negative side of MDS2. Finally, cluster 4 also comprised the same groups as cluster 1 but based on the WP2 abundance estimates instead of the Régent ones. Contrary to the latter, total zooplankton and calanoid copepods here showed null response to temperature. Yet, similar to cluster 1, they also showed strong abundances increase with oxygen concentrations, which explains why both clusters are positioned on the positive side of MDS2 (Fig. 6). This cluster also displayed original strong gaussian responses to chlorophyll a and particles backscattering.

Discussion

Here, we provide a homogeneous dataset of zooplankton composition and size structure based on individual measurements of body size and document the shape of the relationships between community-level size structure and key environmental drivers on a macroecological scale. We find that zooplankton communities exhibit larger median size and abundance towards the poles and towards the tropical upwelling regions sampled (Supplementary Doc. S12), a pattern that is largely driven by copepods. The higher contributions of the large-bodied grazing Calanidae relative to the smaller omnivorous-carnivorous Cyclopoida and Poecilostomatoida (i.e. Oithonidae, Oncaeidae and Corycaeidae) drives the latitudinal increase in median size towards the poles, in addition to the observed negative scaling of body length with temperature which is in line with the TSR^{9,17}. Indeed, our inspection of size structure-environment relationships show that zooplankton size decreases with temperature, salinity, MLD and the contribution of the smallest phytoplankton cells to phytoplankton biomass. Conversely, it increases with concentrations of oxygen, macronutrients, phytoplankton biomass and the contribution of large phytoplankton (e.g. diatoms) to said biomass. Using species body size estimates from the literature, Brun et al.¹⁶ also found copepod mean body size to increase towards the poles, a pattern driven by a negative temperature-size

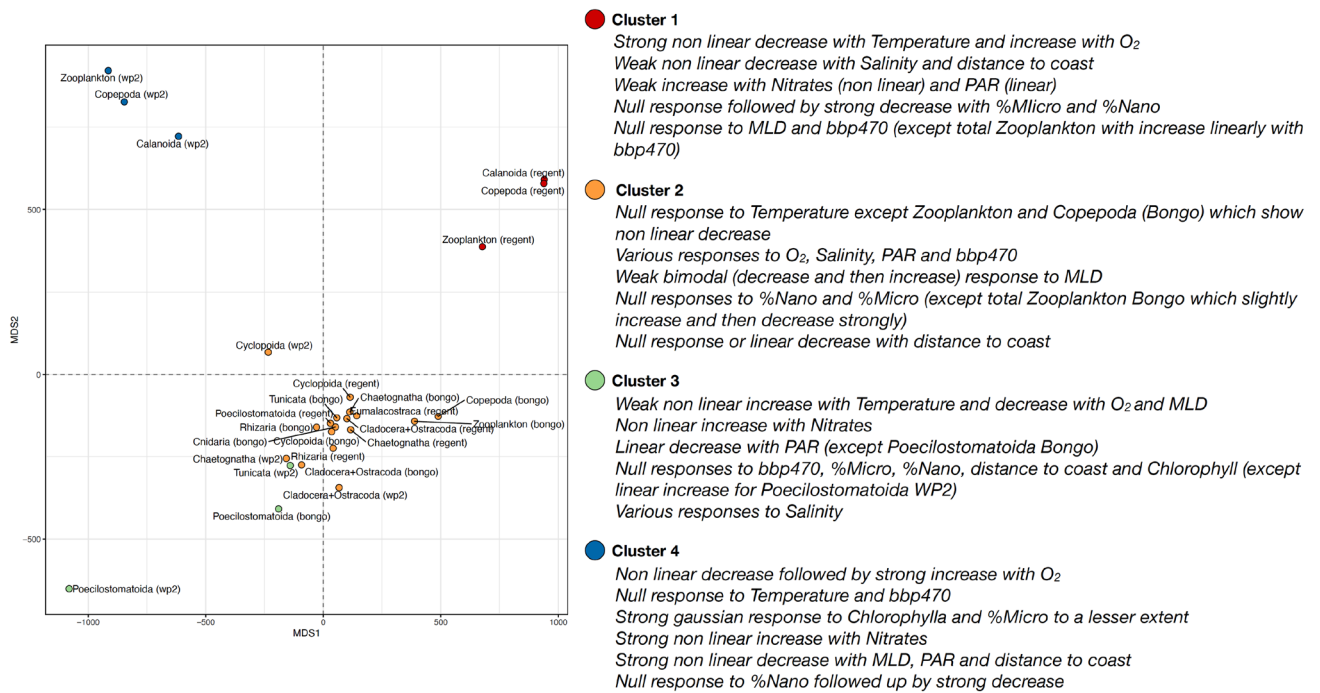


Figure 6. Two dimensional metric dimensional scaling (MDS) plot illustrating the similarity between the responses of the groups' abundances to the environmental covariates selected. The smoothing curves from the Generalized Additive Models (GAMs) modelling the global gradients in cubic-transformed abundances (ind m^3) of the zooplankton groups (estimated for various plankton nets) as a function of ten environmental covariates and displaying a deviance explained $>40\%$. Smoothing curves span a 1–100 scale spanning the range of the covariates measured values. The smoothing curves were combined into a multivariate data series to compute Dynamic Time Warping (DTW) distances and perform partitioning around medoids (PAM) clustering. This way the GAMs were clustered into four clusters that represent combinations of zooplankton groups and plankton nets that exhibit similar abundance-covariate relationships.

relationship and a positive relationship between phytoplankton size and zooplankton size. Several explanations for increased body size towards the poles have been proposed, varying from the stimulating effects of temperature on ectotherm metabolism, the synergetic effects of the presence of larger prey, and the availability of oxygen as a function of temperature^{2,23}. For metazoan ectotherms, the effects of temperature on somatic and gonad growth seem to be the most robust explanation²⁴. The negative correlation between community-level size and temperature might stem from the positive effect of temperature on growth rates. At low latitudes, metabolic rates are higher and life cycles become shorter for the various species composing the community. Consequently community-level median size decreases because of warmer temperatures, and the body surface area to body volume ratio increases²⁵. Despite decades of research, it is still uncertain whether the temperature-size rule is an adaptive response to temperature-related physiological processes (i.e. enzyme activity) or ecological constraints (e.g. food availability, predation and other mortality causes), or a response to biological constraints operating at cellular level such as oxygen supply¹². Arthropods and rotifers have been shown to reach smaller body sizes in poorly oxygenated waters^{23,26}. The potential role of oxygen concentration on the onset of maturation and on size variations remains unclear and is mostly masked by its strong collinearity with surface temperature¹².

In contrast to the decrease in zooplankton median size and abundance observed towards oligotrophic subtropical gyres, an increase was observed near the equatorial regions where the upwelling regime creates colder and more productive conditions. We found the main groups of the zooplankton communities sampled in the eastern boundary upwellings (EBUS) to display significantly higher abundances relative to communities sampled at comparable latitudes (Supplementary Doc. S12). However, the EBUS do not strongly affect the modelled latitudinal patterns of zooplankton abundance (Supplementary Doc. S12). Yet, the effects of the upwelling regime are more marked for abundances than for size structure. This could be linked to the way we estimated median ESD (e.g. aggregated distributions of body size estimated from particles images) compared to the more direct and less uncertain counting of abundance, or to the fact that fewer stations are available when studying size structure gradients (see "Methods"). Overall, abundances showed correlation patterns with the environmental covariates that are quite similar to median size for the total zooplankton community and its major constituting groups (Calanoida, but also Tunicata, Chaetognatha and Cnidaria). This suggests that zooplankton size structure and abundance respond similarly to environmental drivers. Temperature and/or oxygen concentration were found to be the two main covariates in explaining the quasi-global variations of both size structure and abundance. However, we found productivity-related covariates (i.e. Chlorophyll a, NO_2NO_3 concentration, bbp470 and %Micro) to be of higher importance for modelling zooplankton groups abundance. This is an important factor to consider when defining the key parameters to model either zooplankton size or biomass. Our results support the view

that temperature and oxygen are more important parameters than the available biomass of photoautotrophs in driving zooplankton community-level and individual-level body size variations^{12,17} and therefore in controlling the expression of physiological traits that scale allometrically (e.g., growth, respiration).

Yet, the abundance of some zooplankton groups (Poecilostomatoida, Rhizaria and Chaetognatha, and Pteropoda to a lesser extent) show correlation patterns that are opposite to the general copepod-driven trend: their abundance actually increases with temperature, PAR and the contribution of small phytoplankton. These groups rely on feeding strategies that are very different from the filter-feeding Calanoida^{16,27,28}. For instance, the Poecilostomatoida are cruise-feeding and ambush-feeding copepods displaying a broad omnivorous-carnivorous diet^{27–29}. Similarly, chaetognaths are carnivorous ambush-feeders and many pteropods deploy mucus nets for feeding passively on particles fluxes²⁷. Therefore, these groups are able to thrive in large phytoplankton-depleted conditions where mortality-risks and competition for food are more pronounced than in phytoplankton-replete conditions thanks to their alternative feeding strategies. If their growth and reproduction are less dependent on phytoplankton biomass while still promoted in warmer conditions, then spatial patterns driven by positive temperature-abundance relationships can emerge. Our results further support the view that zooplankton is not a homogeneous category whose size structure and biomass dynamics can be adequately modelled through a few size classes^{1,30}.

We found the median ESD of large protists (i.e. Rhizaria, which mainly comprise Foraminifera and Radiolaria) to increase linearly with %Micro but to decrease with %Pico, PAR, and chlorophyll a to a lesser extent. Contrary to Copepoda, temperature and oxygen did not show clear effect on the size structure of those large protists as their median ESD shows contrasted responses to these two covariates across nets. Large protists abundance increased significantly with temperature, macronutrients concentrations, bbp470 and decreased significantly with oxygen. Therefore, the drivers underlying the patterns of Rhizaria abundance and size structure seem distinct, or even opposite, to those that govern copepod size structure and abundance patterns. Again, this could be ascribed to their notable difference in life strategies. Numerous species of Rhizaria are large single-celled mixotrophic protists that host obligate intracellular microalgal symbionts (photosymbionts³¹). Spinose foraminifera show higher contents of chlorophyll a than the shorter non-spinose species³². The efficient photosynthesis performed by photosymbionts, promoted in conditions of higher irradiance and macronutrient concentrations, can lead to oxygen concentrations reaching nearly 200% of the oxygen saturation levels^{33,34}, and potentially even more within their cytoplasm. Such high oxygen availability in the protist cells may weaken the usual temperature- and oxygen-driven constraints on their body size. High oxygen concentrations promote the formation of reactive oxygen species (ROS), which could significantly damage cell structures through the oxidation of DNA, cell membranes or proteins. Overproduction of ROS driven by temperature increase is suspected to trigger coral bleaching, either by symbiont expulsion or digestion³⁵. Similar reactions may occur within protists³⁶. We hypothesize that large protists attempt to prevent ROS accumulation by optimizing the distance between the photosymbionts and themselves. Indeed, most symbiont-bearing Foraminifera tend to display large spinose formation, as a support for the symbiont swarms located further away from the central shell^{33,34,37}, but also enhance prey encounter rates³⁸. Keeping larger sizes to enhance prey capture and avoid ROS could explain the observed stability in median size and abundance of these organisms in the warmer tropical conditions.

The heterogeneity of sampling strategies between surveys usually hinders global scale plankton studies that require the combination of data from multiple oceanographic cruises. The data collected from the *Tara* Oceans expeditions allow us to examine the in situ properties of plankton communities at a very large spatial scale, thanks to the uniform sampling strategy. However, it should be reminded that the one-time nature of such sampling impedes us from addressing the temporal variations of plankton community size structure across the different provinces studied. In addition, it is also worth to point out that the distribution of the sampling stations are unequal across latitudes (Supplementary Fig. S1). Notwithstanding, the latitudinal patterns we observe for copepod size structure are consistent with those of previous studies that resolved seasonal variations^{16,17}, therefore providing some support for the temporal consistency of our results. The correlations we report between abundance, size structure and the environmental variables do not ascertain the ecological and biological processes through which the observed latitudinal patterns emerge. Nonetheless, correlative studies such as ours are key for identifying the major drivers of biological changes and pinpoint further studies to be performed under more controlled conditions that will seek to identify and test the precise biological processes underlying the patterns.

While the level of taxonomic identification of the ZooScan imaging system remains suitable for a size-based community-level study, it does not enable us to depict finer variations in species composition that could be important to further understand the assembly of plankton communities in response to environmental gradients. However, it allowed us to observe large scale patterns and to identify the shape of the relationships between environmental drivers and size structure that would have taken years to depict through non-automated methods. The observed latitudinal patterns in abundance and size structure are relatively consistent across the three nets used but some discrepancies were found (e.g., unidentified Copepoda, Cyclopoida, or Pteropoda; Supplementary Fig. S2). These were likely due to the relative coarse mesh of the Bongo and Régent nets, which underestimated the abundances of most groups (Supplementary Doc. S4). Therefore, these nets could have underestimated the strength of some latitudinal abundance and size structure patterns and their relationships to environmental covariates. Discrepancies between the WP2 data and the two other nets could also stem from differences in sampling depth and net tow, which are known to affect plankton community estimates. The potential effects of these sampling parameters remain difficult to describe here, as only the effects of the mesh size could be evaluated (Supplementary Doc. S4). While the WP2 net was towed vertically from 100 m depth to the surface while the Bongo and Régent nets were towed obliquely from 500 m to the surface. Although these two nets were equipped with coarser meshes, they were towed deeper so they could have captured the deeper living community better³⁹. Nonetheless, considering that most of the zooplankton organisms are concentrated in the 0–200 m layer³⁹, we are

confident that the sampling design of the present *Tara* expeditions adequately captured the macroscale patterns of zooplankton community composition.

Our study follows a trait-based approach to examine the distribution of a “master trait” (i.e. body size) to better investigate how community composition relates to ecosystem functioning¹. We report quasi-global size-latitude relationships in the size structure of major marine zooplankton groups, as well as their scaling with environmental covariates at the community-level. Larger zooplankton are known to enhance energy fluxes to higher trophic levels and to promote carbon export towards deeper layers^{40,41}. Therefore, our observations bring further support to the view that ongoing global climate warming will elicit a decrease in zooplankton size and lower their contribution to the biological carbon pump⁴¹ as well as to overall metabolic rates³. However, fully understanding and predicting such anticipated changes requires a precise parameterization of how environmental conditions impact marine organisms in marine ecosystem models. The representation of plankton diversity in mechanistic marine ecosystem models is improving as the latter may now include from ten⁴² to hundreds of plankton functional types in the case of self-assembling traits-based models^{43,44}. Yet vast inter-model discrepancies exist in terms of their parameterization⁴⁵. Models often aim to validate their parameterization using emergent constraints^{46,47}. The relationships observed between zooplankton community size structure and environmental covariates, or community biomass per size classes and environmental covariates provide such constraints for model validation and evaluation^{45–47} but also shows that one single parameterization is not sufficient to fully capture the variety of the responses observed among plankton organisms. Therefore, our study allows a more precise parameterization of such models, and thus a more precise estimation of future climatic impact on zooplankton organisms abundance, size and by extension effect on the biological carbon pump. We call for closer collaborations between the fields of macroecology, biology, experimental physiology and adaptation to disentangle the roles of multiple drivers in shaping individual traits and the community-level response of marine ecosystems to current and future cumulative effects of stressors, through cell-to-ecosystem studies⁴⁸.

Methods

Sample collection. Zooplankton samples and environmental data were collected at 168 stations across all major oceanic provinces during the *Tara* Oceans expeditions (2009–2013) (Supplementary Fig. S1). Zooplankton was collected with three different types of nets to cover the 200–680 μm size range, encompassing most of the organisms constituting the mesozooplankton. A WP2 net of 200 μm mesh size and 0.57 m^2 opening was towed vertically or obliquely from 100 m depth to the surface. A Bongo net and a Régent net, of 300 and 680 μm mesh size (0.57 and 1.12 m^2 opening), respectively, were towed obliquely from 500 m depth to the surface. Samples were preserved with buffered formaldehyde (4%) for later digitization and morphological analyses. The *Tara* Oceans expeditions sampling strategy and methodologies are fully described in Pesant et al.⁴⁹.

Measurements of environmental covariates. To describe the abiotic habitat associated with each plankton sample, vertical profiles of physical and biogeochemical variables (hereinafter called environmental covariates) were measured by a conductivity temperature depth sensor/rosette (CTD) and Niskin bottles following a published sampling package⁵⁰. A detailed description of each method used as well as all metadata used are available on PANGAEA^{51–54}.

Temperature ($^{\circ}\text{C}$), salinity (psu) and oxygen concentration ($\mu\text{mol kg}^{-1}$) were measured at 10 m depth. Mixed Layer Depth (MLD, m) was estimated based on the 0.03 kg m^{-3} sigma differential density relative to the density at 10 m depth⁵⁵. Chlorophyll a concentration was estimated from vertical CTD casts. The values derived from the fluorescence composite profiles were integrated from 0 to 200 m (or 100 m depending on seafloor depth), using the trapezoidal method. Nutrients concentrations [nitrite/nitrate (NO_2/NO_3 , $\mu\text{mol l}^{-1}$), phosphate (PO_4 , $\mu\text{mol l}^{-1}$) and silicate (SiO_2 , $\mu\text{mol l}^{-1}$)] were determined using segmented flow analysis⁵⁶. For nutrient concentrations, the average of the median values corresponding to each integrated nets samples⁵³ was used as it is a better indicator of the overall conditions over the course of a sampling station.

The contribution of the three main phytoplankton size classes to total phytoplankton biomass, %Pico ($< 2 \mu\text{m}$), %Nano (2–20 μm), and %Micro ($> 20 \mu\text{m}$) were estimated based on HPLC analysis⁵⁷. The measurements were integrated over the 0–200 water column.

Surface Photosynthetically Active Radiation (PAR, $\text{mol quanta m}^{-2} \text{day}^{-1}$) was calculated from in situ sensor data, calibrated using factory settings. Surface backscattering coefficient of particles at 470 nm (bbp470, m^{-1}) was calculated from in situ sensor data, corrected with in situ measurements in dark conditions. For both PAR and bbp470, we used the median value around the sampling date and location^{51–54}.

Among all the contextual metadata provided by the TARA consortium^{51–54}, the above-mentioned covariates were selected because: (i) they were the most complete across most sampling stations; (ii) presented the most normal-like distribution and because they were collinear with their alternative versions. Finally, distance to coast (km) was added a posteriori to the suite of covariates to help disentangling coastal samples from the open ocean ones and include this geographical effects in our statistical models. Distance to coast was computed as the shortest Haversine distance to 0 m isobath, on a 15 min resolution. The bathymetric data from the ETOPO1 database (<https://ngdc.noaa.gov/mgg/global/global.html>) were used and obtained through the *marmap* R package⁵⁸.

Zooplankton abundance and size estimates. Zooplankton samples were analyzed using the ZooScan imaging system²¹. Zooplankton images classification was performed using an automatic recognition algorithm and validated into taxonomic groups by a posteriori expert inspection. Organisms were classified into coarse taxonomic groups on Ecotaxa⁵⁹, generally at the class or order-level except for copepods which were identified down to the family level whenever possible. For our spatial analyses (see below), 36 taxonomic groups were retained, including total zooplankton, Copepoda, Chaetognatha, Cnidaria, Tunicata (mainly appendicularians,

salps and doliolids, which present a range in size, due to change in clade), Eumalacostraca (mainly amphipods, decapods and euphausiids), Rhizaria (mainly foraminifers and radiolarians), Pteropoda and small crustacean grazers (Cladocera, Ostracoda and nauplii larvae). The Copepoda class was then broken down into its main five orders (Calanoida, Cyclopoida, Poecilostomatoida, Harpacticoida and Monstrilloida) which were also broken down to families whenever possible (i.e. to attain $n > 20$ individuals per clade per station; e.g. Oithonidae, Calanidae, Oncaeiidae etc.). The last two groups gathered the unidentified Copepoda and the small unidentified Calanoida for which the resolution of the ZooScan did not permit a more precise classification. The list of living organisms identified on Ecotaxa as well as their final taxonomic classification is summarized in Supplementary Table S13. Abundance values were standardized to the number of individuals per m^3 according to the volume of water filtered. The final mean abundances are the sum of all individuals divided by the volume of water filtered by each net and sample.

The major and minor axes of the best fitting ellipses were measured for each living organism to derive their equivalent spherical diameter (ESD) which is here used as a proxy of body size at the individual-level. Then, the community-level size structure of each of the 36 abovementioned groups was estimated through the median value of the ESD distribution at the individual level. When estimating both abundances and median ESD estimates of a group, all the individuals from the smaller nested groups were accounted for.

Numerical analyses. Three main steps were carried out: (i) spatial patterns of abundance and size structure were explored; (ii) the strength of their linear relationship with various environmental covariates was examined; and (iii) nonlinear statistical models were fitted based on the groups' abundances and median ESD and the selected environmental covariates to examine the underlying drivers of global abundance and size structure.

First, the distribution of the groups' abundance and median ESD were visually inspected for each of the 36 groups and four transformations (square-root, natural log, log to base 10 and cubic) were applied to examine which would provide the distribution closest to a normal distribution (based on the p-value of Shapiro–Wilk normality tests). As a result, the groups' abundances were cubic-transformed and the median ESD estimates were log-transformed.

To identify the groups displaying the most compelling global patterns of abundance and size structure, a first set of spatially-explicit generalized additive models (GAMs)⁶⁰ were fitted to the latitude of the sampling stations for each net separately. GAMs are generalized linear models that allow to incorporate flexible nonlinear responses through smoothing functions. GAMs allow us to identify and model nonlinear latitudinal patterns that might emerge because of species–environment relationships. Thin plate regression splines were applied and the smoothing parameters were determined through restricted maximum likelihood (REML) and a Gaussian link function. We insured that at least 30 stations, with each at least 20 individuals were available for fitting the GAMs and thus avoid focusing on groups presenting few observations. Variability in the groups' abundance may bias the associated median ESD estimates. Some groups of finer taxonomic resolution (i.e. families) may display very low abundances and thus an insufficient number of individuals to derive a robust median ESD estimate from. Therefore, we ensured to consider only on those sampling stations that displayed a sufficient amount of individuals ($n \geq 20$) for examining size structure patterns. The adjusted R^2 of the GAMs, as well as the p-values of the latitude smooth terms, were examined to identify the groups displaying significant latitudinal patterns.

Prior to examining the correlation between the groups' abundance and median ESD and the selected environmental covariates, the latter were transformed (i.e. square-root, natural log, log to base 10 and cubic) and the normality of the distributions was tested. Macronutrients (NO_2NO_3 , PO_4 and SiO_2) concentrations were cubic-transformed and chlorophyll a concentration was log-transformed. The values of the other covariates were kept as is. Spearman's rank correlation coefficients (ρ) were computed between the groups' abundance/median ESD and all covariates to examine the strength of their linear relationships and identify the main drivers of the spatial patterns. Then, the shape of specific abundance–environment and size–environment relationships were modelled through GAMs. A prior selection of covariates was carried out to discard those that would be too collinear⁶¹. Covariates collinearity is a sensitive issue when modelling biotic–abiotic relationships through regressive models as it may inflate parameters and errors estimates⁶¹. For each net data, pairwise Spearman's correlation coefficients (ρ) were computed between covariates (Supplementary Fig. S14). When a pair of covariates displayed a $|\rho| \geq 0.7$, the covariate displaying the distribution closest to normality and the least amount of missing values was retained. As a result, %Pico was discarded to the advantage of %Micro ($\rho = 0.85$), and PO_4 and SiO_2 concentrations were discarded to the advantage of NO_2NO_3 concentration ($\rho = 0.95$ and $\rho = 0.87$, respectively). NO_2NO_3 was thus used to represent gradients in macronutrients concentration. Temperature and dissolved oxygen concentration also displayed high collinearity ($\rho = -0.95$). Yet, we wanted to assess how important both these covariates could be in explaining abundance and size structure patterns. Therefore we ran the subsequent analyses by accounting for temperature and O_2 but separately.

Using the groups transformed abundance and median ESD as response variables, we fitted GAMs for each net data using the following ten covariates: temperature, salinity, MLD, PAR, NO_2NO_3 , chlorophyll a, bbp470, %Micro, %Nano and distance to coast. A second set of GAMs was trained by replacing temperature by oxygen. Again, thin plate regression splines were applied and the parameters were determined through REML and a Gaussian link function. The dimensions of the basis of the smooth terms were adjusted by dividing the number of available observations by the number of covariates. Extra penalties were added to each smoothing term so the parameter estimation can completely remove terms estimated as insignificant. Model terms were then selected by backwards removal of insignificant variables. The percentage of explained deviance (%Dev) and the adjusted R^2 of the GAMs were retrieved to evaluate their performance. For each GAM, covariates significance was ranked according to their relative F statistic (Supplementary Fig. S9). We tested for significant variations in %Dev across net data or covariates sets through non-parametric variance analyses (Kruskal–Wallis test) followed by posthoc

pairwise comparisons using Dunn's test and Bonferroni's method for adjusting p-values. Only those stations displaying more than 20 individuals when modelling the groups' size structure estimates were considered. The significant smooth terms (p-values < 0.05) were identified and their smoothing function and standard error estimates were plotted on the scale of the corresponding covariate (covariate-specific smooth curve).

The covariate-specific smooth curves of the GAMs fitted on median ESD estimates displaying a %Dev \geq 50% were then used to cluster the zooplankton groups based on the shapes of the modelled smooth curves. This way, we identified groups that respond similarly to environmental gradients. In short, the smooth curves were treated as independent data series and we used Dynamic Time Warping (DTW)⁶² to compare them to each other. DTW is an algorithm that tries to find the optimum warping path between two univariate or multivariate data series. DTW stretches the data series locally to have one match the other(s) as much as possible. Then, the Euclidean distance between the data series is computed by summing the distances of the aligned data points. The modelled smooth curves were projected on a scale of 1 to 100 values scaling the range observed for each covariate. All the covariate-specific smooth curves were out together in a list that can be assimilated to a multivariate data series even in lengths for each retained GAM. DTW distances were then computed and partitioning around medoids (PAM) clustering⁶³ was performed to cluster the GAMs (i.e. a zooplankton group + a net type) into two to ten clusters. Five different indices (Calinski-Harabasz, Dunn's, Silhouettes, classic and modified Davies Bouldin indices) were examined to choose the optimal number of cluster. Four clusters were retained for the median-ESD based GAMs, and four for the abundance-based ones (Supplementary Fig. S15). Hierarchical clustering approaches were also examined but they yielded unclear performance indices so PAM was preferred. Every modelled covariate-specific smooth curves used for the DTW clustering are reported in Supplementary Figs. S10 and S11 for the median ESD responses and the abundance models, respectively. To summarize such large amount of information and illustrate the similarities between the zooplankton groups, the inter-group distance matrix issued from the DTW algorithm was projected onto to a two dimensional space based on classical multidimensional scaling.

The analyses were performed with the R v3.5.2⁶⁴ environment and with MATLAB R2017a. All maps presented were plotted in R v3.5.2. The main packages used for data analyses and plotting were *tidyverse*⁶⁵ and *HH*⁶⁶ and *FactoMineR*⁶⁷. GAMs were built using the *mgcv*⁶⁰ package. The partitioning of the zooplankton groups based on the shape of the smoothing curves issued from the GAMs was performed using the *dtwclus*⁶⁸ package.

Data availability

Median ESD and abundance values by zooplankton groups are available at <https://doi.org/10.17632/nwvfwccgvh.1>. Zooplankton imaging datasets from the *Tara* Oceans expeditions are available through the collaborative web Ecotaxa application and repository under the addresses: <https://ecotaxa.obs-vlfr.fr/prj/377>, <https://ecotaxa.obs-vlfr.fr/prj/2245>, <https://ecotaxa.obs-vlfr.fr/prj/378> for the WP2 net; <https://ecotaxa.obs-vlfr.fr/prj/397>, <https://ecotaxa.obs-vlfr.fr/prj/398>, <https://ecotaxa.obs-vlfr.fr/prj/395> for the Bongo net; <https://ecotaxa.obs-vlfr.fr/prj/415>, <https://ecotaxa.obs-vlfr.fr/prj/409>, <https://ecotaxa.obs-vlfr.fr/prj/408>, <https://ecotaxa.obs-vlfr.fr/prj/411>, <https://ecotaxa.obs-vlfr.fr/prj/412> for the Régent net. Contextual data from the *Tara* Oceans expedition, including those that are newly released from the Arctic Ocean, are available at <https://doi.org/10.1594/PANGAEA.875582>.

Received: 29 September 2020; Accepted: 13 July 2021

Published online: 03 August 2021

References

- Litchman, E., Ohman, M. D. & Kiørboe, T. Trait-based approaches to zooplankton communities. *J. Plankton Res.* **35**, 473–484. <https://doi.org/10.1093/plankt/fbt019> (2013).
- Kiørboe, T. & Hirst, A. G. Shifts in mass scaling of respiration, feeding, and growth rates across life-form transitions in marine pelagic organisms. *Am. Nat.* **183**, E118–E130. <https://doi.org/10.1086/675241> (2014).
- Andersen, K. H. *et al.* Characteristic sizes of life in the oceans, from bacteria to whales. *Annu. Rev. Mar. Sci.* **8**, 1–25. <https://doi.org/10.1146/annurev-marine-122414-034144> (2015).
- Bergmann, C. Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien* **3**, 595–708 (1847).
- Woodson, C., Schramski, J. R. & Joye, S. B. A unifying theory for top-heavy ecosystem structure in the ocean. *Nat. Commun.* **9**, 1–8. <https://doi.org/10.1038/s41467-017-02450-y> (2018).
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789. <https://doi.org/10.1890/03-9000> (2004).
- Gardner, J. L., Peters, A., Kearney, M. R., Joseph, L. & Heinsohn, R. Declining body size: A third universal response to warming?. *Trends Ecol. Evol.* **26**, 285–291. <https://doi.org/10.1016/j.tree.2011.03.005> (2011).
- Angilletta, M. J., Steury, T. D. & Sears, M. W. Temperature, growth rate, and body size in ectotherms: Fitting pieces of a life-history puzzle. *Integr. Comp. Biol.* **44**, 498–509. <https://doi.org/10.1093/icb/44.6.498> (2004).
- Atkinson, D. Temperature and organism size: A biological law for ectotherms?. *Adv. Ecol. Res.* **25**, 1–58. [https://doi.org/10.1016/S0065-2504\(08\)60212-3](https://doi.org/10.1016/S0065-2504(08)60212-3) (1994).
- Atkinson, D. & Sibly, R. M. Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends Ecol. Evol.* **12**, 235–239. [https://doi.org/10.1016/S0169-5347\(97\)01058-6](https://doi.org/10.1016/S0169-5347(97)01058-6) (1997).
- Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science* **348**, 1261359. <https://doi.org/10.1126/science.1261359> (2015).
- Audzijonyte, A. *et al.* Is oxygen limitation in warming waters a valid mechanism to explain decreased body sizes in aquatic ectotherms?. *Glob. Ecol. Biogeogr.* **28**, 64–77. <https://doi.org/10.1111/geb.12847> (2018).
- Begon, M., Townsend, C. R. & Harper, J. L. *Ecology: From Individuals to Ecosystems* 4th edn. (Blackwell Publishing, New York, 2006).
- Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J. & Barlow, R. G. An absorption model to determine phytoplankton size classes from satellite ocean colour. *Remote Sens. Environ.* **112**, 3153–3159. <https://doi.org/10.1016/j.rse.2008.03.011> (2008).
- Kostadinov, T., Siegel, D. & Maritorea, S. Global variability of phytoplankton functional types from space: Assessment via the particle size distribution. *Biogeosciences* **7**, 3239–3257. <https://doi.org/10.5194/bg-7-3239-2010> (2010).

16. Brun, P., Payne, M. R. & Kjørboe, T. Trait biogeography of marine copepods: An analysis across scales. *Ecol. Lett.* **19**, 1403–1413. <https://doi.org/10.1111/ele.12688> (2016).
17. Horne, C. R., Hirst, A. G., Atkinson, D., Neves, A. & Kjørboe, T. A global synthesis of seasonal temperature–size responses in copepods. *Glob. Ecol. Biogeogr.* **25**, 988–999. <https://doi.org/10.1111/geb.12460> (2016).
18. Garzke, J., Hansen, T., Ismar, S. M. H. & Sommer, U. Combined effects of ocean warming and acidification on copepod abundance, body size and fatty acid content. *PLoS ONE* **11**, e0155952. <https://doi.org/10.1371/journal.pone.0155952> (2006).
19. Stelzer, C. P. Phenotypic plasticity of body size at different temperatures in a planktonic rotifer: Mechanisms and adaptive significance. *Funct. Ecol.* **16**, 835–841. <https://doi.org/10.1046/j.1365-2435.2002.00693.x> (2002).
20. Riemer, K., Anderson-Teixeira, K. J., Smith, F. A., Harris, D. J. & Ernest, S. K. M. Body size shifts influence effects of increasing temperatures on ectotherm metabolism. *Global Ecol. Biogeogr.* **27**, 958–967. <https://doi.org/10.1111/geb.12757> (2018).
21. Gorsky, G. *et al.* Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.* **32**, 285–303. <https://doi.org/10.1093/plankt/fbp124> (2010).
22. Ibarbalz, F. M. *et al.* Global trends in marine plankton diversity across kingdoms of life. *Cell* **179**, 1084–1097. <https://doi.org/10.1016/j.cell.2019.10.008> (2019).
23. Hoefnagel, K. N. & Verberk, W. C. Is the temperature–size rule mediated by oxygen in aquatic ectotherms?. *J. Therm. Biol.* **54**, 56–65. <https://doi.org/10.1016/j.jtherbio.2014.12.003> (2015).
24. Wojewodzic, M. W., Kyle, M., Elser, J. J., Hessen, D. O. & Andersen, T. Joint effect of phosphorus limitation and temperature on alkaline phosphatase activity and somatic growth in *Daphnia magna*. *Oecologia* **165**, 837–846. <https://doi.org/10.1007/s00442-010-1863-2> (2011).
25. Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M. & Charnov, E. L. Effects of size and temperature on metabolic rate. *Science* **293**, 2248–2251. <https://doi.org/10.1126/science.1061967> (2001).
26. Czarnoleski, M., Ejsmont-Karabin, J., Angilletta, M. K. & Kozłowski, J. Colder rotifers grow larger but only in oxygenated waters. *Ecosphere* **6**, 1–5. <https://doi.org/10.1890/ES15-00024.1> (2015).
27. Kjørboe, T. How zooplankton feed: Mechanisms, traits and trade-offs. *Biol. Rev.* **86**, 311–339. <https://doi.org/10.1111/j.1469-185X.2010.00148.x> (2011).
28. Benedetti, F., Gasparini, S. & Ayata, S.-D. Identifying copepod functional groups from species functional traits. *J. Plankton Res.* **38**, 159–166. <https://doi.org/10.1093/plankt/fbv096> (2016).
29. Brun, P., Payne, M. R. & Kjørboe, T. A trait database for marine copepods. *Earth Syst. Sci. Data* **9**, 99–113. <https://doi.org/10.5194/essd-9-99-2017> (2017).
30. Anderson, T. R. Plankton functional type modelling: Running before we can walk?. *J. Plankton Res.* **27**, 1073–1081. <https://doi.org/10.1093/plankt/fbi076> (2005).
31. Biard, T. *et al.* *In situ* observations unveil an unexpectedly large biomass of Radiolaria and Phaeodaria (Rhizaria) in the oceans. *Nature* **532**, 504–507. <https://doi.org/10.1038/nature17652> (2016).
32. Takagi, H. *et al.* Characterizing photosymbiosis in modern planktonic foraminifera. *Biogeosciences* **16**, 3377–3396. <https://doi.org/10.5194/bg-16-3377-2019> (2019).
33. Rink, S., Kühl, M., Bijma, J. & Spero, H. J. Microsensor studies of photosynthesis and respiration in the symbiotic foraminifer *Orbulina universa*. *Mar. Biol.* **131**, 583–595. <https://doi.org/10.1007/s002270050350> (1998).
34. Lombard, F., Erez, J., Michel, E. & Labeyrie, L. Temperature effect on respiration and photosynthesis of the symbiont-bearing planktonic foraminifera *Globigerinoides ruber*, *Orbulina universa*, and *Globigerinella siphonifera*. *Limnol. Oceanogr.* **54**, 210–218. <https://doi.org/10.4319/lo.2009.54.1.0210> (2009).
35. Lesser, M. P. Coral Bleaching: Causes and Mechanisms. In *Coral Reefs: An Ecosystem in Transition* (eds Dubinsky, Z. & Stambler, N.) (Springer, 2011).
36. Villar, E. *et al.* Symbiont chloroplasts remain active during bleaching-like response induced by thermal stress in *Collozoum pelagicum* (Collozoum, Retaria). *Front. Mar. Sci.* **5**, 387. <https://doi.org/10.3389/fmars.2018.00387> (2018).
37. Hemleben, C., Spindler, M. & Anderson, O. R. *Modern Planktonic Foraminifera* (Springer-Verlag, 1989).
38. Suzuki, N. & Not, F. Biology and ecology of radiolaria. In *Marine Protists: Diversity and Dynamics* (eds Ohtsuka, S. *et al.*) (Springer, 2015).
39. de Puelles, F. *et al.* Zooplankton abundance and diversity in the tropical and subtropical ocean. *Diversity* **11**, 203. <https://doi.org/10.3390/d11110203> (2019).
40. Beauprand, G., Edwards, M. & Legendre, L. Marine biodiversity, ecosystem functioning, and carbon cycles. *PNAS* **107**, 10120–10124. <https://doi.org/10.1073/pnas.0913855107> (2010).
41. Brun, P. *et al.* Climate change has altered zooplankton-fuelled carbon export in the North Atlantic. *Nat. Ecol. Evol.* **3**, 416–423. <https://doi.org/10.1038/s41598-018-0780-3> (2019).
42. Buitenhuis, E. T., Le Quéré, C., Bednaršek, N. & Schiebel, R. Large contribution of pteropods to shallow CaCO₃ export. *Glob. Biogeochem. Cycles* **33**, 458–468. <https://doi.org/10.1029/2018GB006110> (2019).
43. Follows, M. J., Dutkiewicz, J., Grant, S. & Chisholm, S. W. Emergent biogeography of microbial communities in a model ocean. *Science* **315**, 1843–1846. <https://doi.org/10.1126/science.1138544> (2007).
44. Ward, B. A., Dutkiewicz, S., Jahn, O. & Follows, J. F. A size-structured food-web model for the global ocean. *Limnol. Oceanogr.* **57**, 1877–1891. <https://doi.org/10.4319/lo.2012.57.6.1877> (2012).
45. Saille, S. F. *et al.* Comparing food web structures and dynamics across a suite of global marine ecosystem models. *Ecol. Model.* **261**, 43–57. <https://doi.org/10.1016/j.ecolmodel.2013.04.006> (2013).
46. Le Quéré, C. *et al.* Role of zooplankton dynamics for Southern Ocean phytoplankton biomass and global biogeochemical cycles. *Biogeosciences* **13**, 4111–4133. <https://doi.org/10.5194/bg-13-4111-2016> (2016).
47. Kwiatkowski, L. *et al.* Emergent constraints on projections of declining primary production in the tropical oceans. *Nat. Clim. Change* **7**, 355–358. <https://doi.org/10.1038/nclimate3265> (2017).
48. Sunagawa, S. *et al.* Tara Oceans: Towards global ocean ecosystems biology. *Nat. Rev. Microbiol.* **18**, 428–445. <https://doi.org/10.1038/s41579-020-0364-5> (2020).
49. Pesant, S. *et al.* Open science resources for the discovery and analysis of Tara Oceans data. *Sci. Data* **2**, 150023. <https://doi.org/10.1038/sdata.2015.23> (2015).
50. Picheral, M. *et al.* Vertical profiles of environmental parameters measured on discrete water samples collected with Niskin bottles at station TARA_147 during the Tara Oceans expedition 2009–2013. PANGAEA <https://doi.org/10.1594/PANGAEA.839235> (2014).
51. Guidi, L. *et al.* Environmental context of all samples from the Tara Oceans Expedition (2009–2013), about sensor data in the targeted environmental feature. PANGAEA <https://doi.org/10.1594/PANGAEA.875576> (2017).
52. Guidi, L. *et al.* Environmental context of all samples from the Tara Oceans Expedition (2009–2013), about pigment concentrations (HPLC) in the targeted environmental feature. PANGAEA <https://doi.org/10.1594/PANGAEA.875569> (2017).
53. Guidi, L. *et al.* Environmental context of all samples from the Tara Oceans Expedition (2009–2013), about nutrients in the targeted environmental feature. PANGAEA <https://doi.org/10.1594/PANGAEA.875575> (2017).
54. Speich, S. *et al.* Environmental context of all samples from the Tara Oceans Expedition (2009–2013), about the water column features at the sampling location. PANGAEA <https://doi.org/10.1594/PANGAEA.875579> (2017).

55. de Boyer-Montegut, C., Madec, G., Fischer, A. S., Lazar, A. & Iudicone, D. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.* **109**, C12003. <https://doi.org/10.1029/2004JC002378> (2004).
56. Aminot, A., K erouel, R. & Coverly, S. C. Nutrients in seawater using segmented flow analysis. In *Practical Guidelines for the Analysis of Seawater* (ed. Wurl, O.) (CRC Press, 2009).
57. Uitz, J., Claustre, H., Morel, A. & Hooker, S. B. Vertical distribution of phytoplankton communities in Open Ocean: An assessment based on surface chlorophyll. *J. Geophys. Res.* **111**, C08005. <https://doi.org/10.1029/2005JC003207> (2006).
58. Pante, E. & Simon-Bouhet, B. marmap: A Package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* **8**(9), e73051. <https://doi.org/10.1371/journal.pone.0073051> (2013).
59. Picheral, M., Colin, S. & Irisson J.-O. *EcoTaxa, A Tool for the Taxonomic Classification of Images*. <http://ecotaxa.obs-vlfr.fr> (2017).
60. Wood, S. N. *Generalized Additive Models: An Introduction with R* 2nd edn. (Chapman and Hall/CRC, 2017).
61. Dormann, C. F. *et al.* Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x> (2013).
62. Giorgino, T. Computing and visualizing dynamic time warping alignments in R: The dtw package. *J. Stat. Softw.* **31**, 1–24. <https://doi.org/10.18637/jss.v031.i07> (2009).
63. Park, H.-S. & Jun, C.-H. A simple and fast algorithm for K-medoids clustering. *Expert Syst. Appl.* **36**, 3336–3341. <https://doi.org/10.1016/j.eswa.2008.01.039> (2009).
64. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2018). <https://www.R-project.org/>.
65. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686. <https://doi.org/10.21105/joss.01686> (2019).
66. Heiberger, R. M. *HH: Statistical Analysis and Data Display: Heiberger and Holland. R package version 3.1–40*, <https://CRAN.R-project.org/package=HH> (2020).
67. L e, S., Josse, J. & Husson, F. FactoMineR: An R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18. <https://doi.org/10.18637/jss.v025.i01> (2008).
68. Sarda-Espinosa, A. *dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance. R package version 5.5.6*, <https://CRAN.R-project.org/package=dtwclust> (2019).

Acknowledgements

Tara Oceans (which includes both the *Tara* Oceans and *Tara* Oceans Polar Circle expeditions) would not exist without the leadership of the *Tara* Expeditions Foundation and the continuous support of 23 institutes (<http://oceans.taraexpeditions.org>). We further thank the commitment of the following sponsors: CNRS (in particular Groupement de Recherche GDR3280 and the Research Federation for the study of Global Ocean Systems Ecology and Evolution, FR2022/*Tara* Oceans-GOSEE), European Molecular Biology Laboratory (EMBL), Genoscope/CEA, The French Ministry of Research, and the French Government ‘Investissements d’Avenir’ programmes OCEANOMICS (ANR-11-BTBR-0008), FRANCE GENOMIQUE (ANR-10-INBS-09-08), MEMO LIFE (ANR-10-LABX-54), and PSL Research University (ANR-11-IDEX-0001-02). M.C.B. acknowledges post-doc fellowships from the Coordination for the Improvement of Higher Education Personnel of Brazil (CAPES) (99999.000487/2016-03) and the Fonds Franais pour l’Environnement Mondial (FFEM). F.B. received support from ETH Z rich and from the European Union’s Horizon 2020 research and innovation program under grant agreement n SEP-210591007. We also thank the support and commitment of Agn s b. and Etienne Bourgeois, the Prince Albert II de Monaco Foundation, the Veolia Foundation, Region Bretagne, Lorient Agglomeration, Serge Ferrari, World Courier, and KAUST. The global sampling effort was enabled by countless scientists and crew who sampled aboard the *Tara* from 2009–2013, and we thank MERCATOR-CORIOLIS and ACRI-ST for providing daily satellite data during the expeditions. We are also grateful to the countries who graciously granted sampling permissions. The authors declare that all data reported herein are fully and freely available from the date of publication, with no restrictions, and that all of the analyses, publications, and ownership of data are free from legal entanglement or restriction by the various nations whose waters the *Tara* Oceans expeditions sampled. The following people were involved in plankton image sorting: B. Serranito, N. Monferrer, C. Merland and F. Roullier. We thank the EMBRC platform PIQv for image analysis. This work was supported by EMBRC-France, whose French state funds are managed by the ANR within the Investments of the Future program under reference ANR-10-INBS-02. We are grateful to Meike Vogt for sharing her thoughts regarding the potential use of the present observations for zooplankton modelling. This article is contribution number 121 of *Tara* Oceans.

Author contributions

M.C.B., F.B., L.S. and F.L. conceptualized the study and wrote the original draft. M.C.B., F.B., S.M., Y.D.S., J.-B.R., A.E., C.D., L.J., M.P., L.G., G.G., L.S. and F.L. participated in data curation. M.C.B., F.B. and F.L. lead the preparation of the data and performed the numerical analyses. M.C.B., F.B., S.M., Y.D.S., J.-O.I., L.S. and F.L. helped interpret the data. L.S. and F.L. supervised the study. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94615-5>.

Correspondence and requests for materials should be addressed to M.C.B. or F.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Tara Oceans Consortium Coordinators

Silvia G. Acinas¹³, Marcel Babin¹⁴, Peer Bork¹⁵, Emmanuel Boss⁹, Chris Bowler^{7,8}, Guy Cochrane¹⁶, Colombaro de Vargas^{8,10}, Gabriel Gorsky¹, Lionel Guidi¹, Nigel Grimsley^{17,18}, Pascal Hingamp⁴, Daniele Iudicone¹⁹, Olivier Jaillon²⁰, Stefanie Kandels^{15,21}, Lee Karp-Boss⁹, Eric Karsenti^{7,21}, Fabrice Not¹⁰, Hiroyuki Ogata²², Nicole Poulton²³, Stephane Pesant^{24,25}, Jeroen Raes^{26,27,28}, Christian Sardet²⁹, Sabrina Speich^{30,31}, Lars Stemmann¹, Matthew B. Sullivan¹¹, Shinichi Sunagawa³² & Patrick Wincker²⁰

¹³Institute of Marine Sciences (ICM) – CSIC, Pg. Marítim de la Barceloneta, 37-49, 08003 Barcelona, Spain.

¹⁴Takuvik Joint International Laboratory (UMI3376), Université Laval (Canada) – CNRS (France), Université Laval, Québec, QC G1V 0A6, Canada. ¹⁵Structural and Computational Biology, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. ¹⁶European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. ¹⁷CNRS Biologie Intégrative Des Organismes Marins (BIOM), UMR7232, 1 avenue Pierre Fabre, 66650 Banyuls-sur-Mer, France. ¹⁸Sorbonne Université, Observatoire Océanologique de Banyuls-Sur-Mer, 1 avenue Pierre Fabre, 66650 Banyuls-sur-Mer, France. ¹⁹Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy. ²⁰Génomique Métabolique, Genoscope, Institut de Biologie François Jacob, Commissariat à l'Énergie Atomique (CEA), CNRS, Université ÉvrY, Université Paris-Saclay, ÉvrY, France. ²¹Directors' Research, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. ²²Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto 611-001, Japan. ²³Bigelow Laboratory for Ocean Sciences, East Boothbay, ME 04544, USA. ²⁴PANGAEA, Data Publisher for Earth and Environmental Science, University of Bremen, Bremen, Germany. ²⁵MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany. ²⁶Department of Microbiology and Immunology, Rega Institute, KU Leuven, Herestraat 49, 3000 Leuven, Belgium. ²⁷Center for the Biology of Disease, VIB, Herestraat 49, 3000 Leuven, Belgium. ²⁸Department of Applied Biological Sciences, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium. ²⁹Sorbonne Université, CNRS, UMR 7009 Biodev, Observatoire Océanologique, 06230 Villefranche-sur-mer, France. ³⁰Department of Geosciences, Laboratoire de Météorologie Dynamique (LMD), Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France. ³¹Laboratoire de Physique des Océans, UBO-IUEM, Place Copernic, 29820 Plouzané, France. ³²Department of Biology, Institute of Microbiology and Swiss Institute of Bioinformatics, ETH Zürich, Vladimir-Prelog-Weg 4, 8093 Zürich, Switzerland.