

# Macroscale patterns of oceanic zooplankton composition and size structure

Manoela C. Brandão<sup>1,2,+,\*</sup>, Fabio Benedetti<sup>3,+,\*</sup>, Séverine Martini<sup>4</sup>, Yawouvi Dodji Soviadan<sup>1</sup>, Jean-Olivier Irisson<sup>1</sup>, Jean-Baptiste Romagnan<sup>5</sup>, Amanda Elineau<sup>1</sup>, Corinne Desnos<sup>1</sup>, Laëtitia Jalabert<sup>1</sup>, Andrea S. Freire<sup>6</sup>, Marc Picheral<sup>1</sup>, Lionel Guidi<sup>1</sup>, Gabriel Gorsky<sup>1</sup>, Chris Bowler<sup>7,8</sup>, Lee Karp-Boss<sup>9</sup>, Nicolas Henry<sup>8,10</sup>, Colomban de Vargas<sup>8,10</sup>, Matthew B. Sullivan<sup>11</sup>, *Tara* Oceans Consortium Coordinators<sup>‡</sup>, Lars Stemmann<sup>1,8</sup>, and Fabien Lombard<sup>1,8,12</sup>

<sup>1</sup>Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, Villefranche-sur-mer, 06230, France

<sup>2</sup>Ifremer, Centre Bretagne, Unité Dynamiques des Ecosystèmes Côtiers, Plouzané, 29280, France

<sup>3</sup>ETH Zürich, Institute of Biogeochemistry and Pollutant Dynamics, Zürich, 8092, Switzerland

<sup>4</sup>Aix Marseille Univ., Université de Toulon, CNRS, IRD, MIO UM 110, Marseille, 13288, France

<sup>5</sup>Ifremer, Centre Atlantique, Unité Ecologie et Modèles pour l'Halieutique, Nantes, 44311, France

<sup>6</sup>Universidade Federal de Santa Catarina, Departamento de Ecologia e Zoologia, Florianópolis, 88010970, Brazil

<sup>7</sup>Institut de biologie de l'École normale supérieure (IBENS), CNRS, INSERM, PSL Université Paris, Paris, 75005, France

<sup>8</sup>Research Federation for the study of Global Ocean Systems Ecology and Evolution, FR2022/*Tara* Oceans GOSEE, Paris, 75016, France

<sup>9</sup>University of Maine, School of Marine Sciences, Orono, 04469, USA

<sup>10</sup>Sorbonne Université, CNRS, Station Biologique de Roscoff, AD2M, UMR 7144, Roscoff, 29680, France

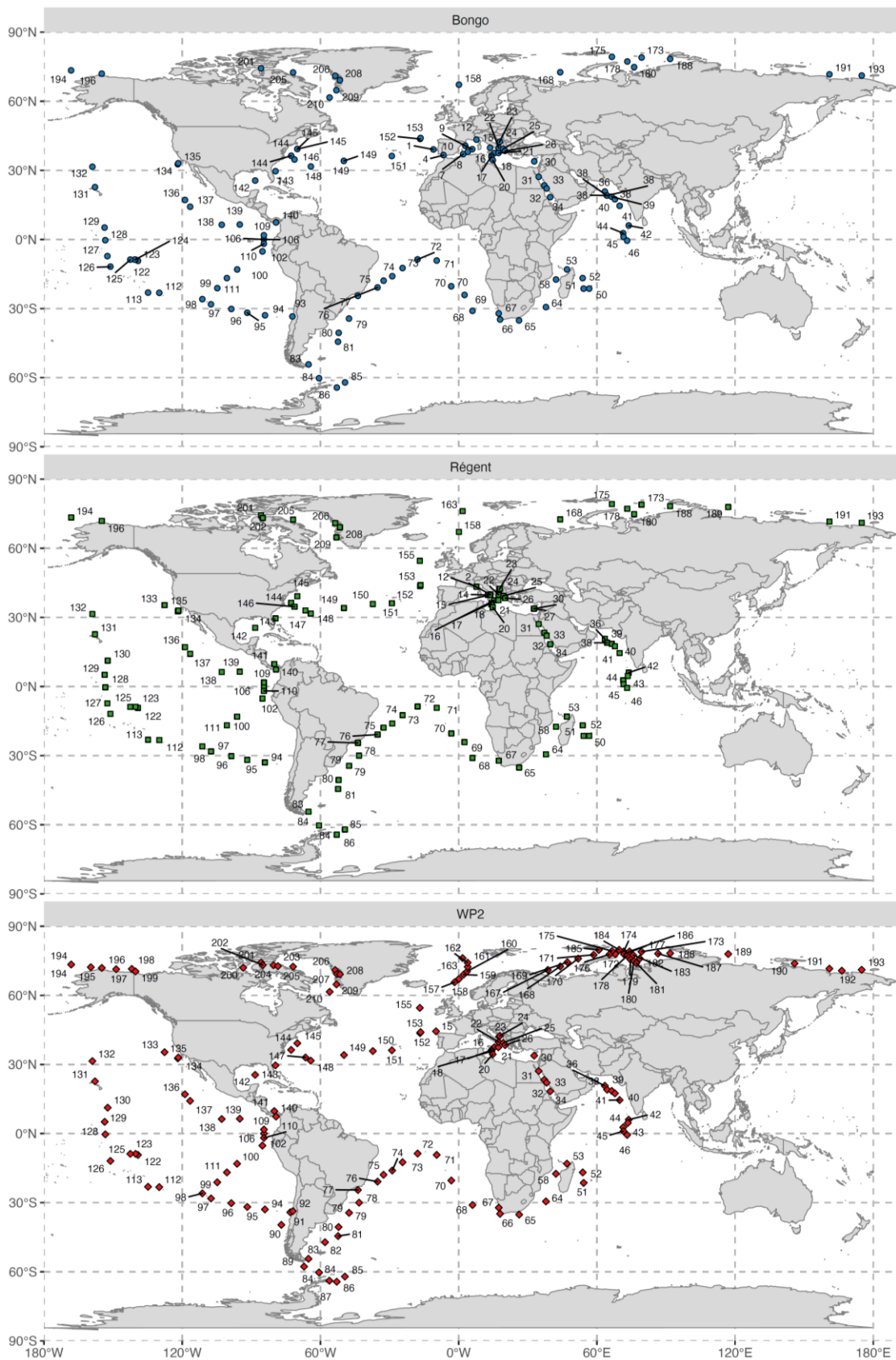
<sup>11</sup>The Ohio State University, Department of Microbiology and Civil, Environmental, and Geodetic Engineering, Columbus, 43214, USA

<sup>12</sup>Institut Universitaire de France, Paris, 75231, France

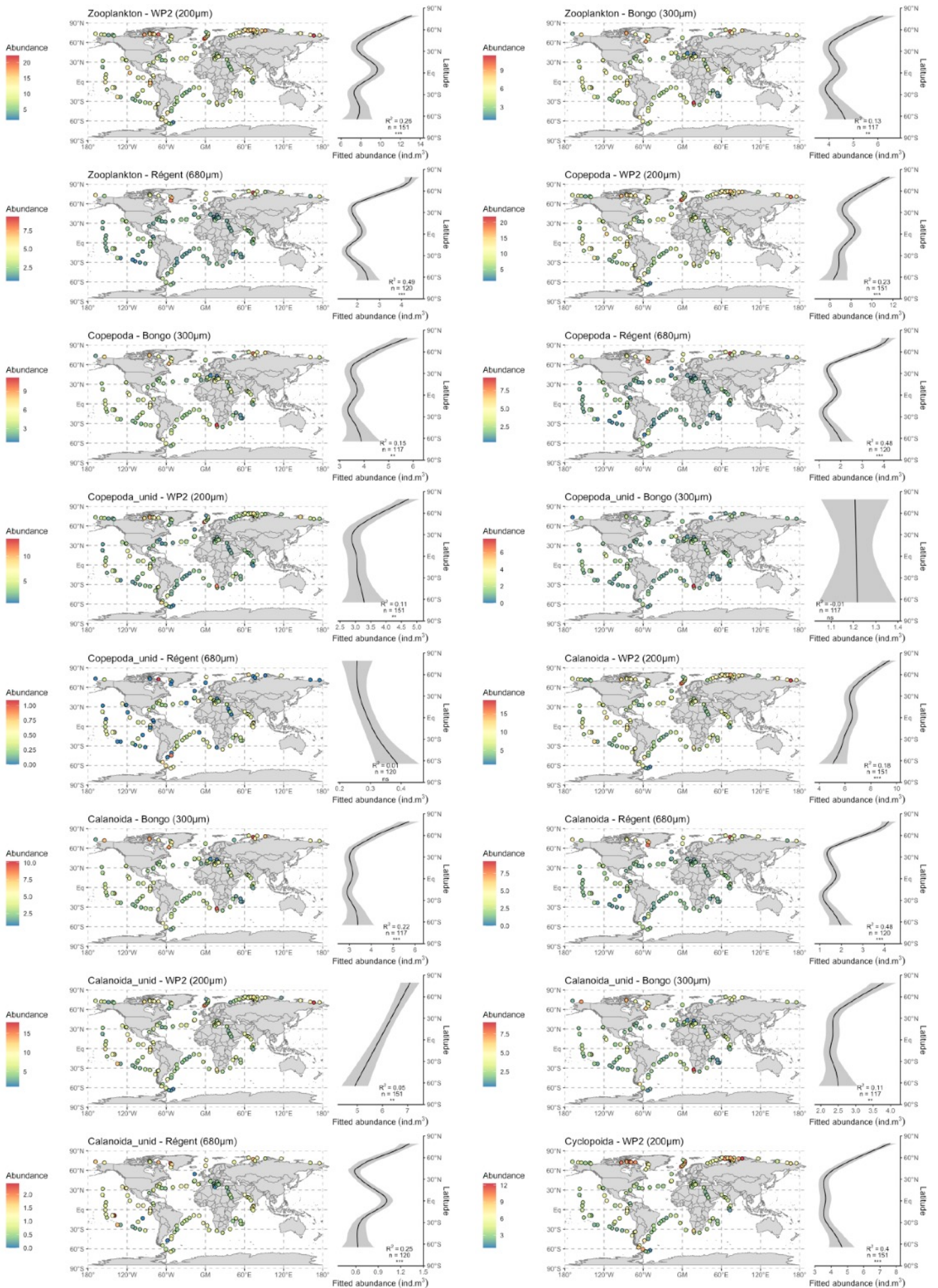
<sup>+</sup>these authors contributed equally to this work

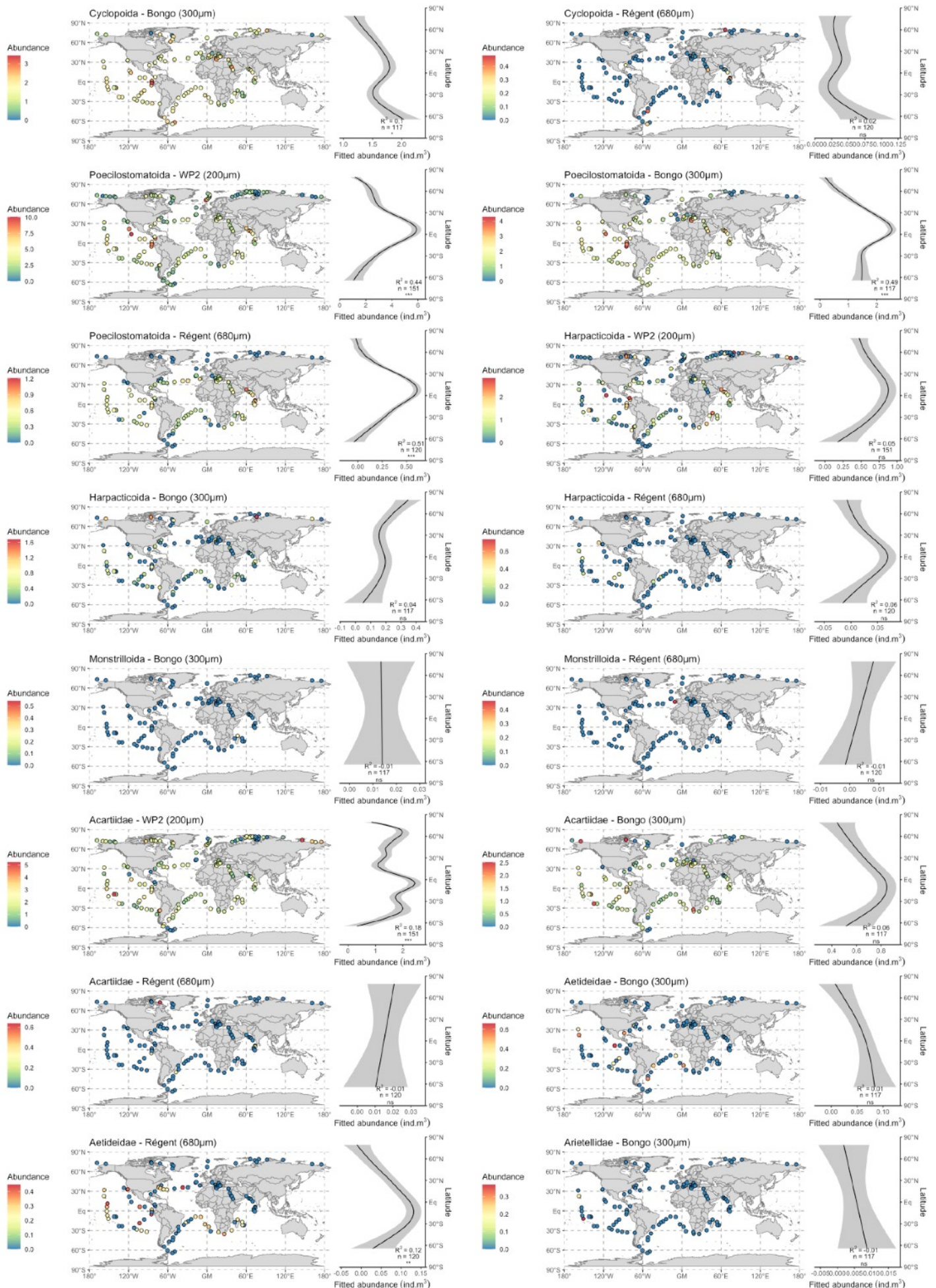
<sup>\*</sup>manoelacbl@gmail.com; fabio.benedetti@usys.ethz.ch

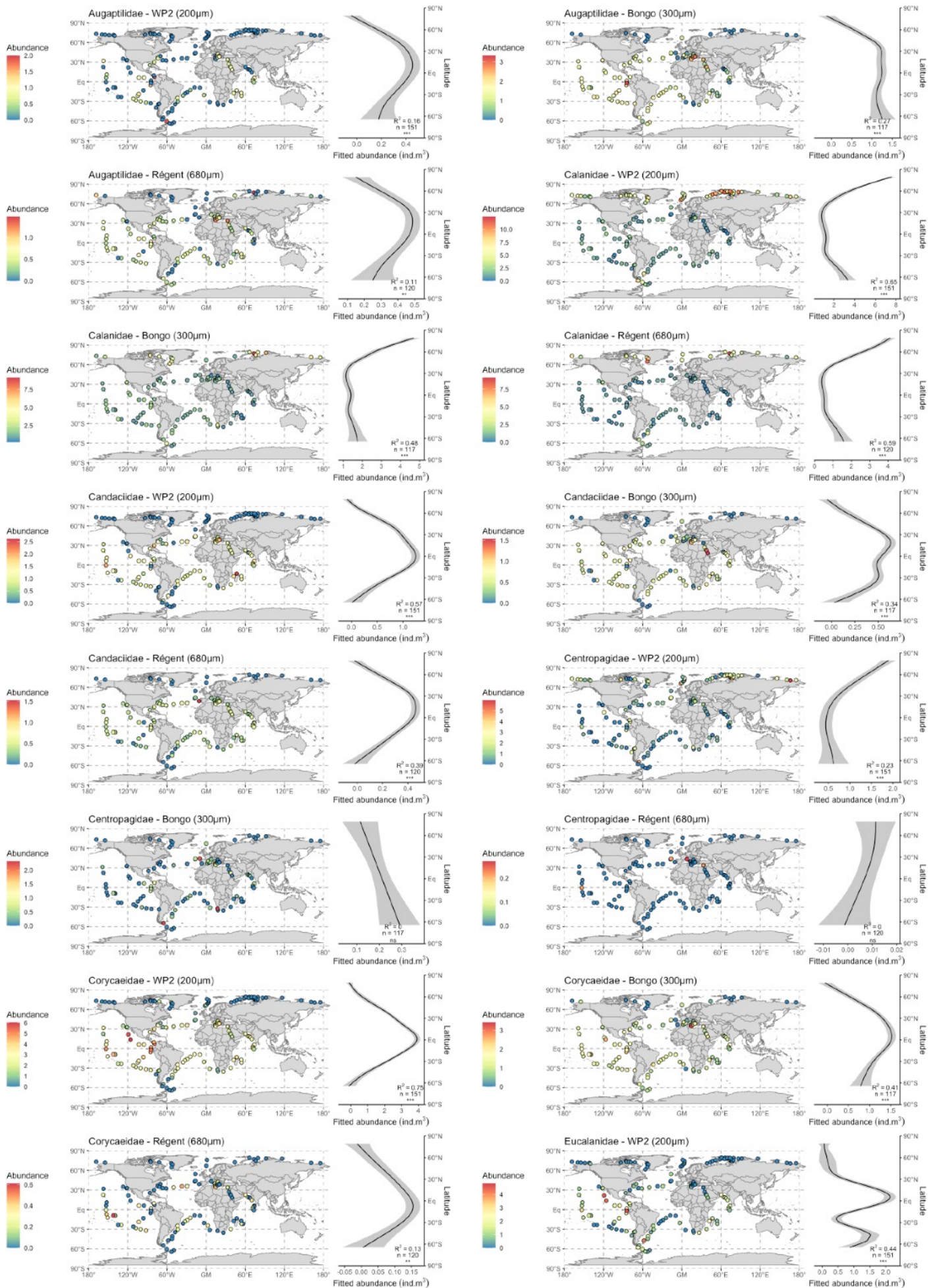
<sup>‡</sup>*Tara* Oceans Consortium Coordinators and affiliations are listed at the end of this manuscript

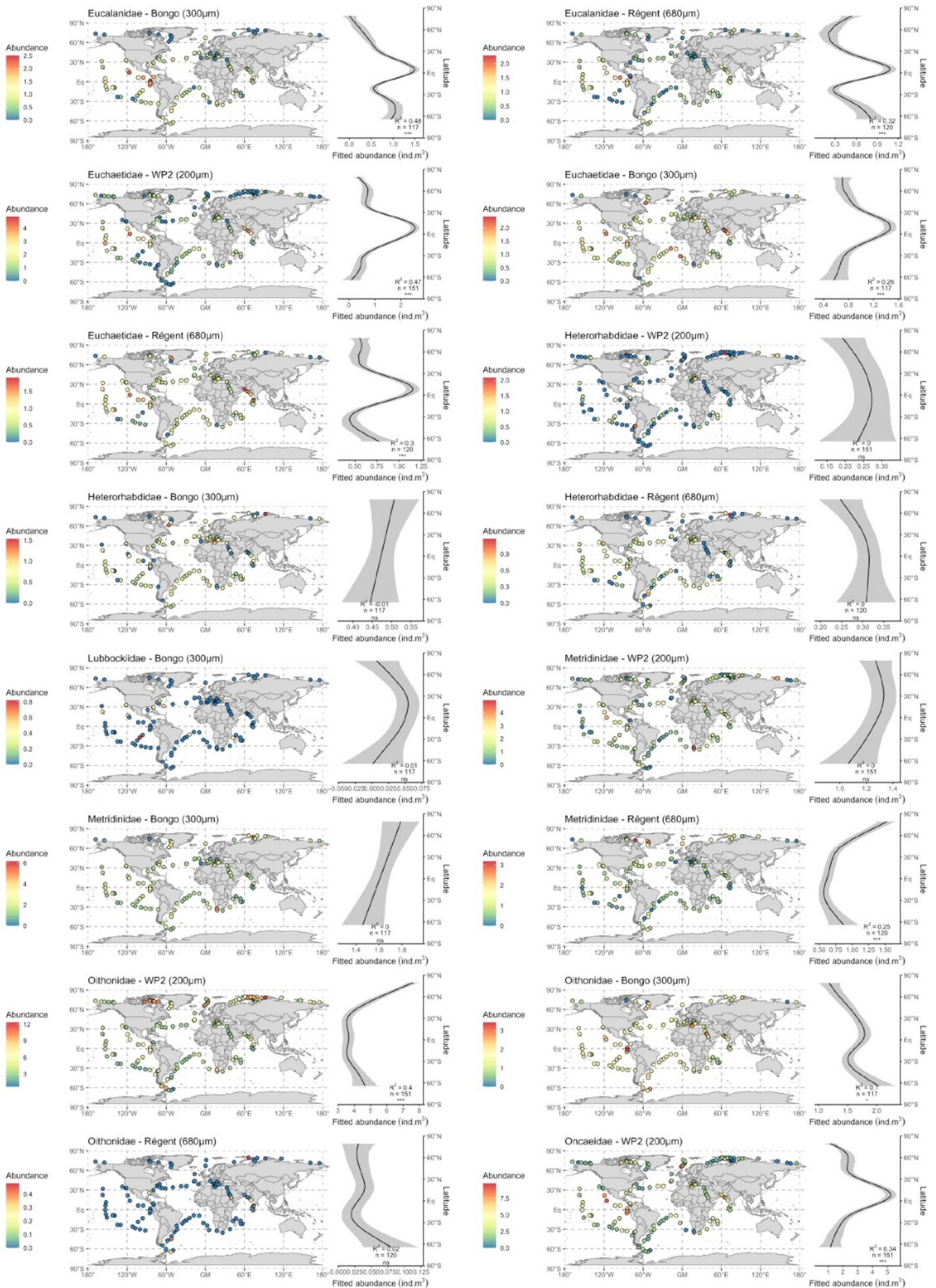


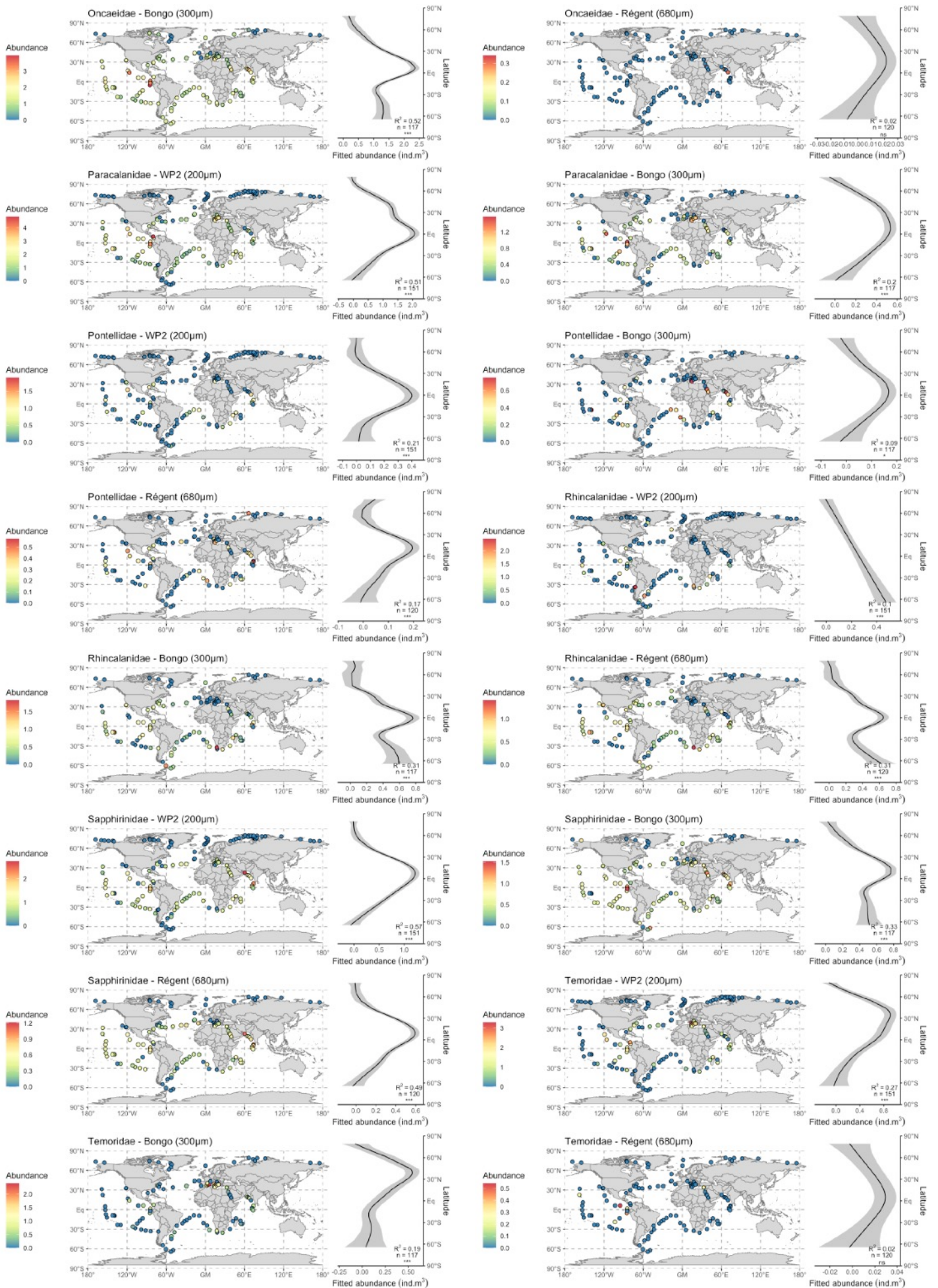
**Supplementary Figure S1:** Location of the sampling stations of the *Tara* Oceans expedition used in the present study discriminated per plankton net.

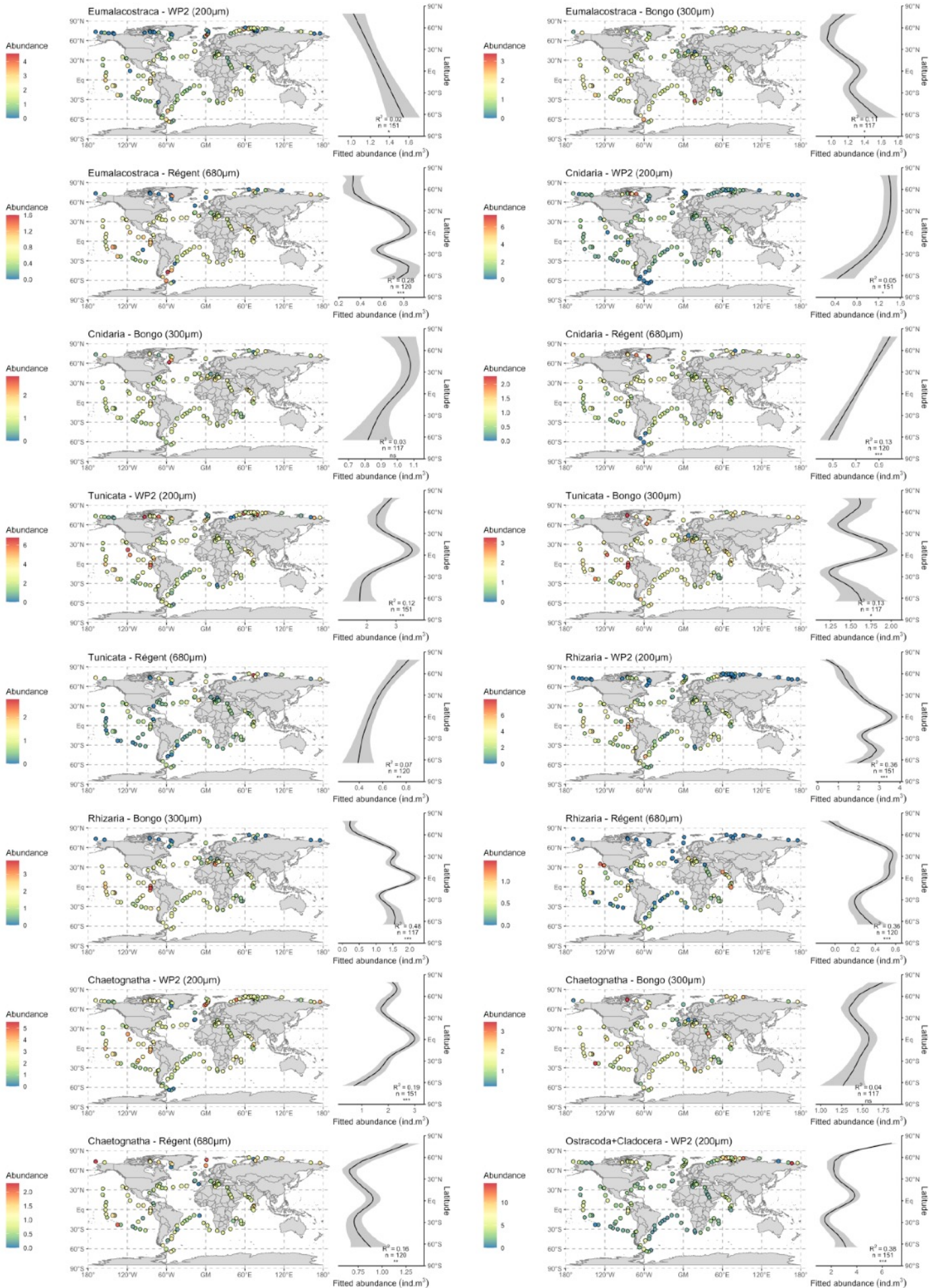




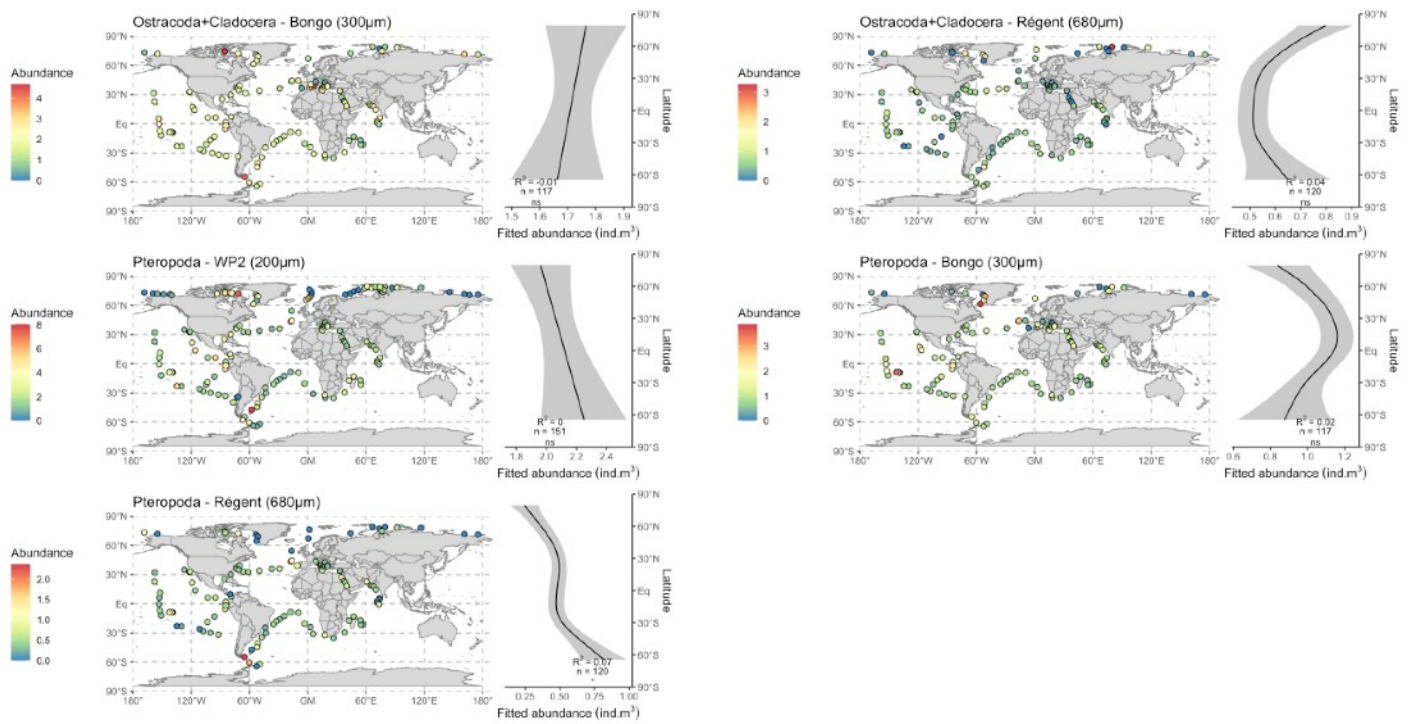




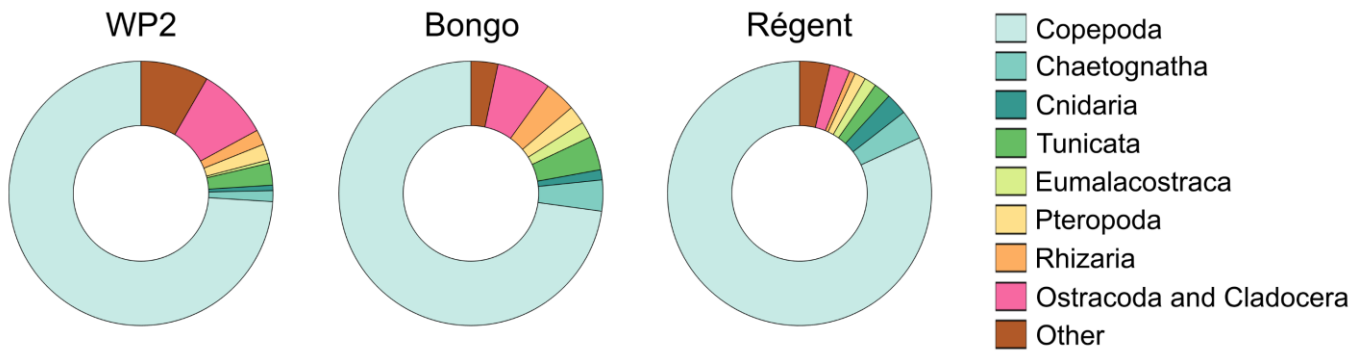








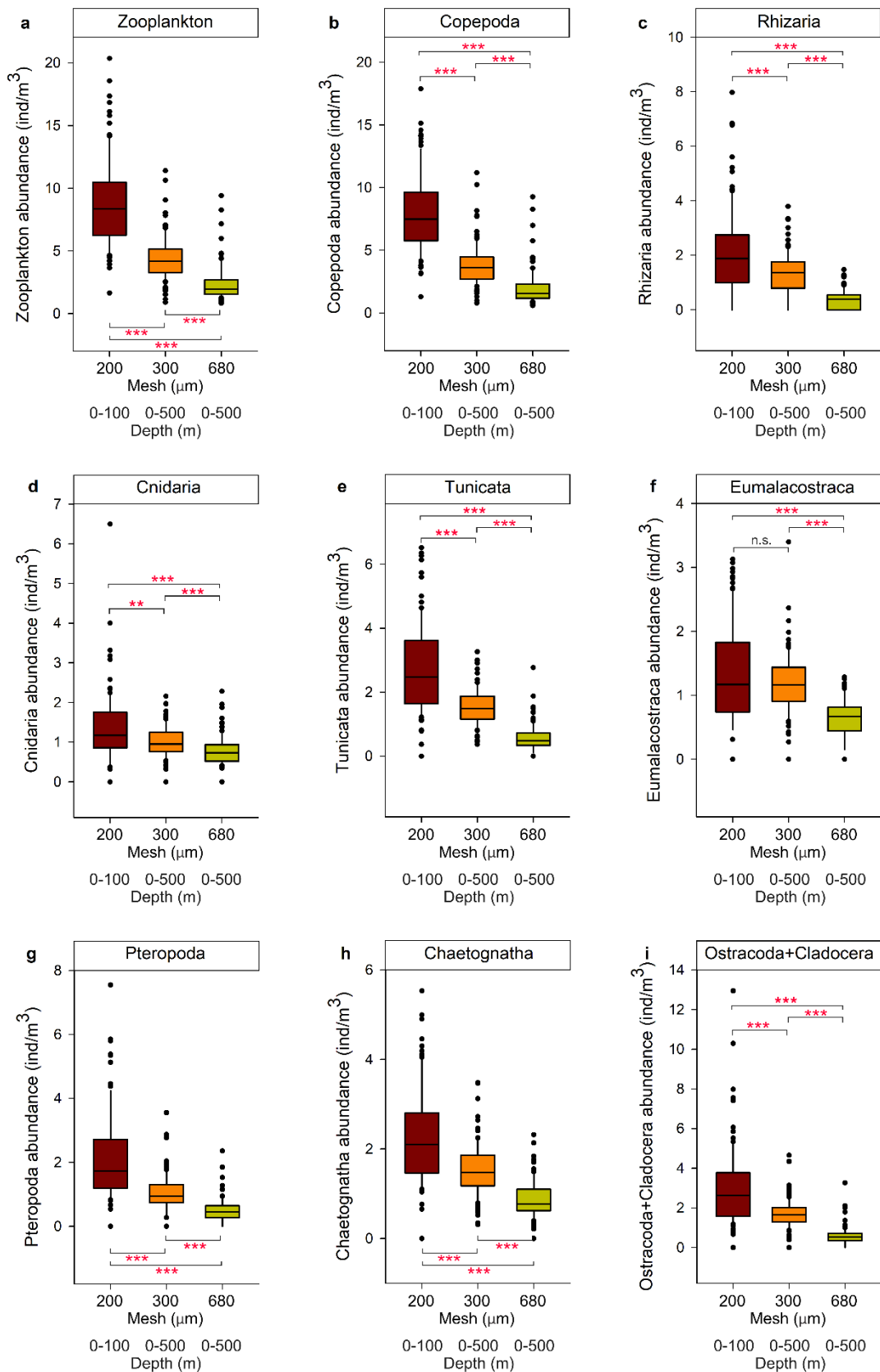
**Supplementary Figure S2:** Maps and latitudinal patterns of the abundance (cubic-transformed ind.m<sup>3</sup>) of the zooplankton groups sampled with the three plankton nets (WP2, Bongo and Régent). The solid curves on the zonal side plots illustrate the prediction from the Generalized Additive Model (GAM) fitting transformed abundance against latitude. The explanatory power of the GAM (adjusted R<sup>2</sup>), 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 prediction.



Net	Copepoda	Chaetognatha	Cnidaria	Tunicata	Eumalacostraca	Pteropoda	Rhizaria	Ostracoda + Cladocera	Other
<b>WP2 (200µm)</b>	73.99	1.34	0.66	2.69	0.40	1.97	1.85	8.77	8.32
<b>Bongo (300µm)</b>	72.85	3.79	1.25	4.14	1.95	2.26	3.84	6.63	3.30
<b>Régent (680µm)</b>	81.90	3.62	2.65	2.09	1.46	1.36	0.72	2.47	3.73

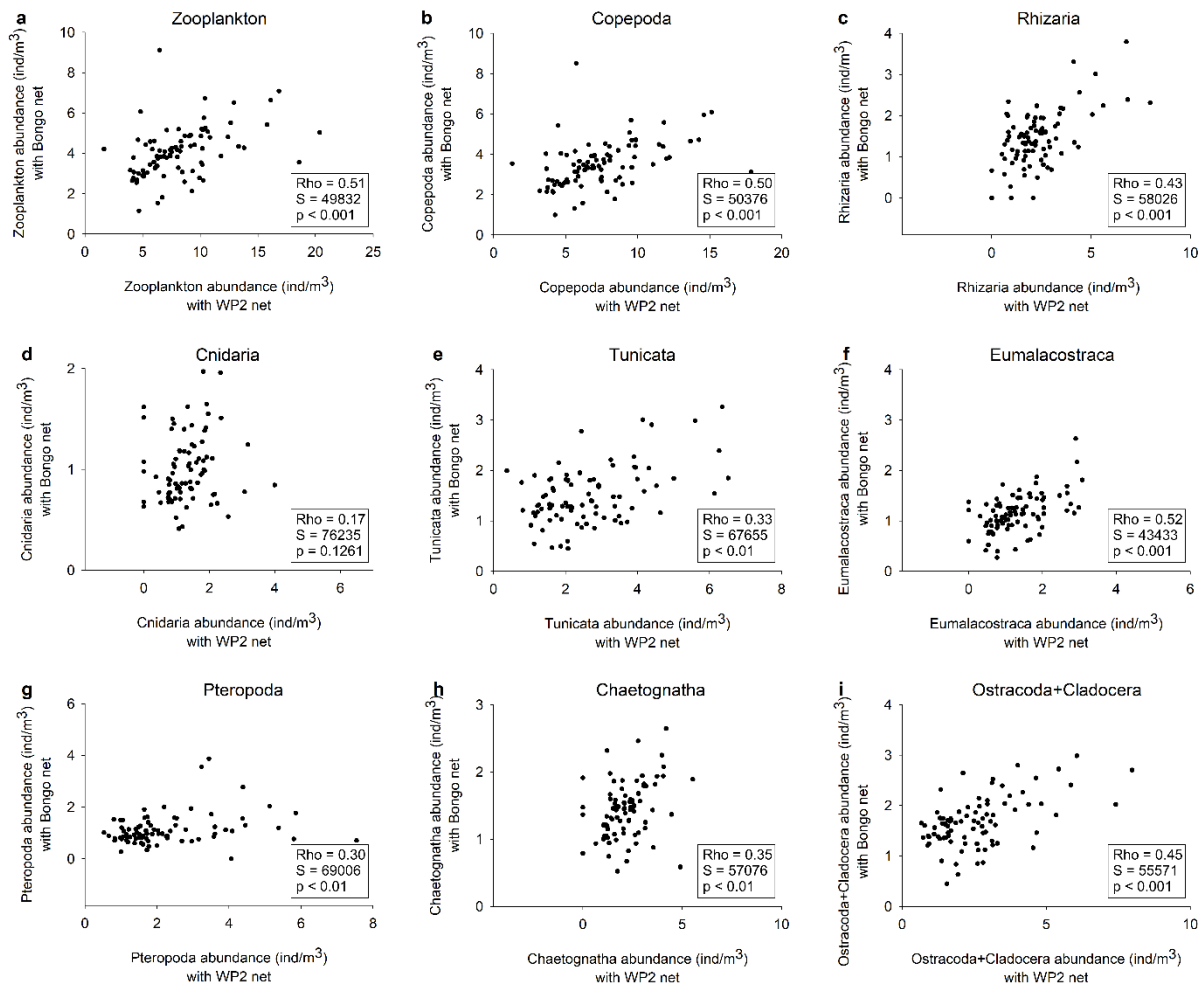
**Supplementary Document S3:** Distribution of the relative contribution (%) of the main mesozooplankton groups to total zooplankton community abundance across the three plankton nets used.

**Supplementary Document S4:** Comparing mesh size and sampling depth effects on the abundance distributions of the main zooplankton groups.

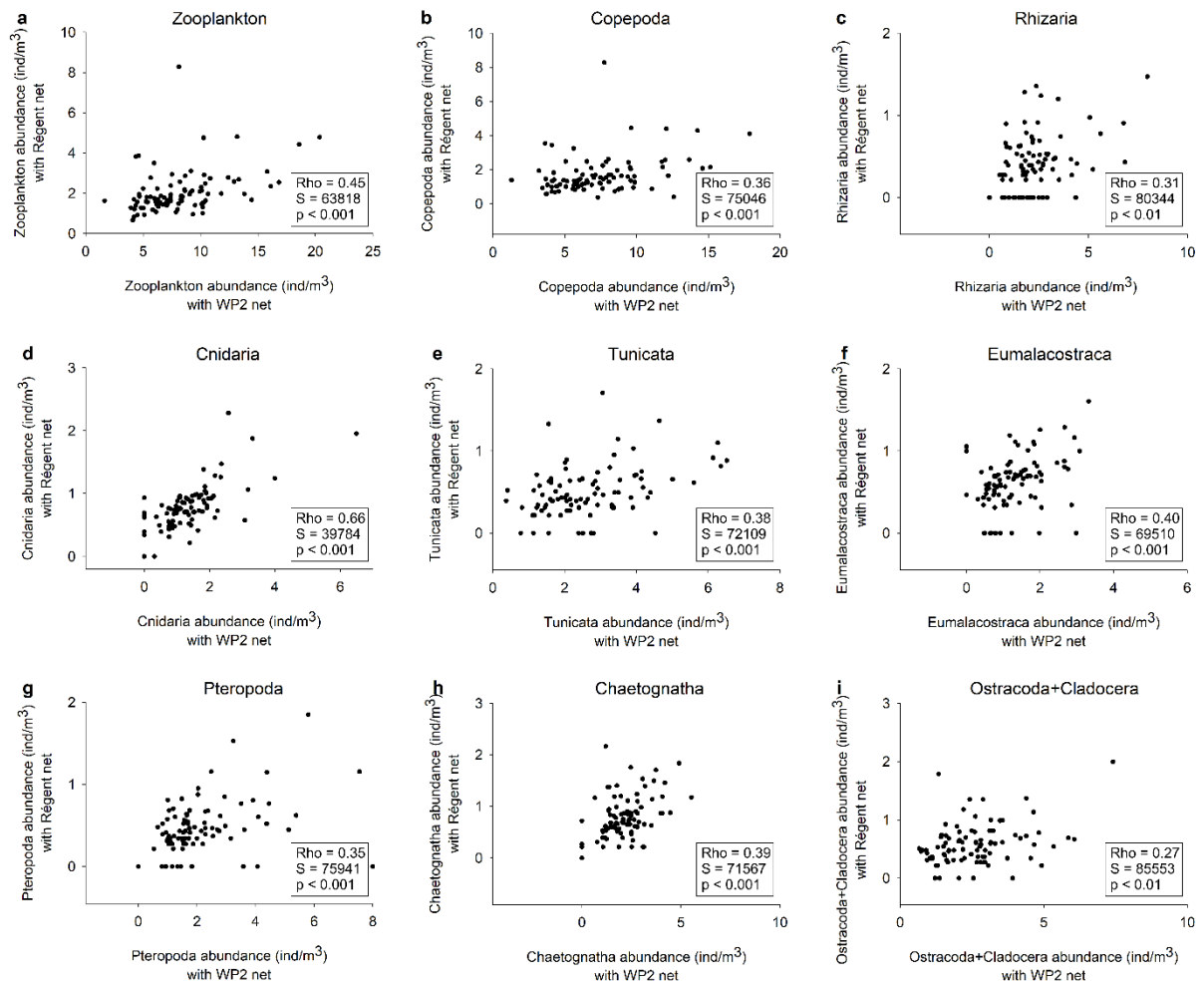


**Figure S4.1:** Comparing the cubic-transformed abundance distributions of the main zooplankton groups studied between the three plankton mesh sizes (200μm, WP2; 300μm, Bongo; 680μm, Régent) and sampling

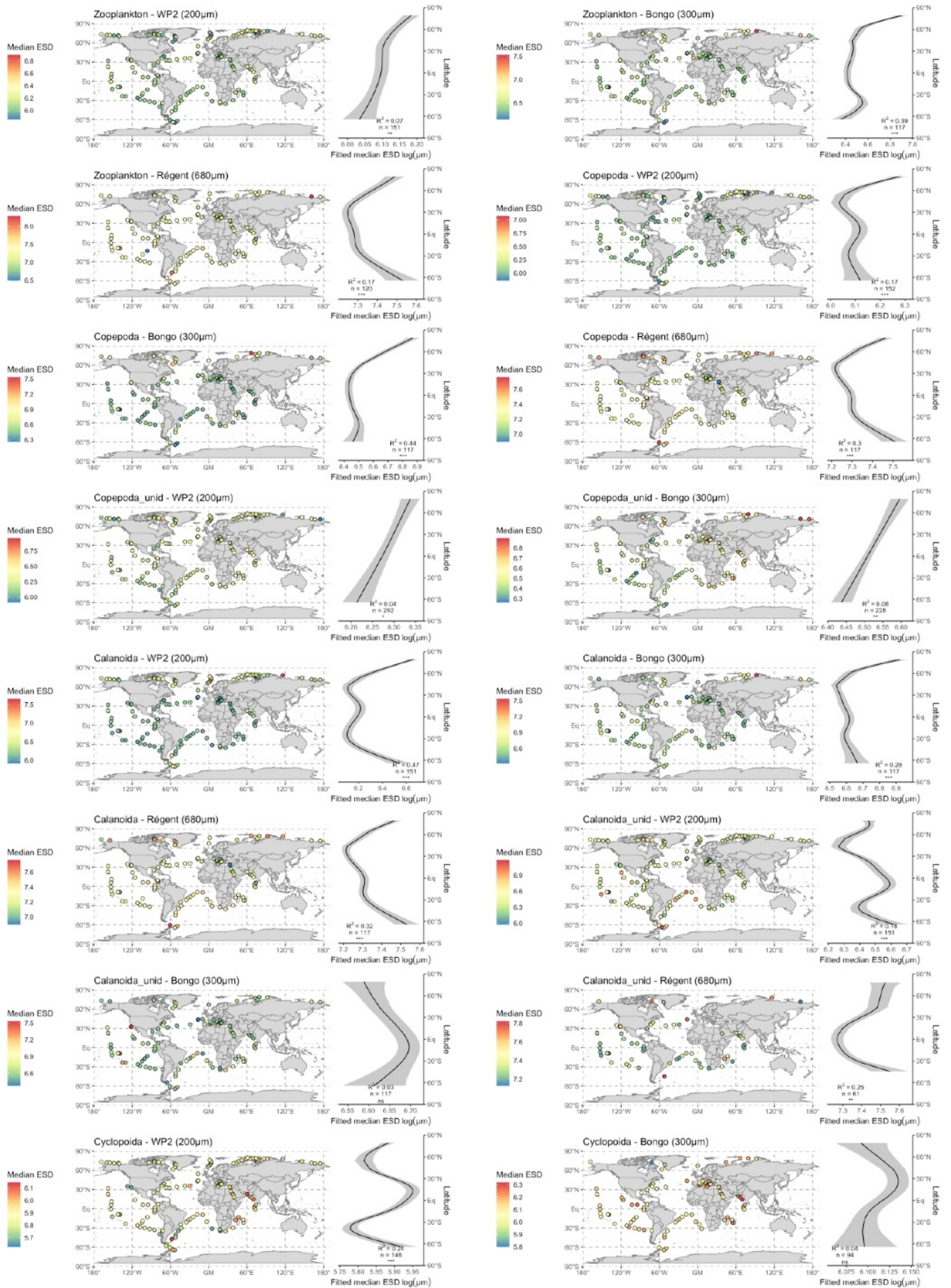
depth of the net hauls (0-100m, WP2; 0-500m, Bongo and Régent). a) Total zooplankton, b) Copepoda, c) Rhizaria, d) Cnidaria, e) Tunicata, f) Eumalacostraca, g) Pteropoda, h) Chaetognatha, and i) Ostracoda + Cladocera. For each group and mesh size, non parametric variance analyses (Kruskal-Wallis tests) were performed to test if the nets displayed significant different zooplankton abundances. Conventional labels illustrate the level of significance of the test: n.s. = not significant; \*\* = p-value < 0.01; \*\*\* = p-value < 0.001.

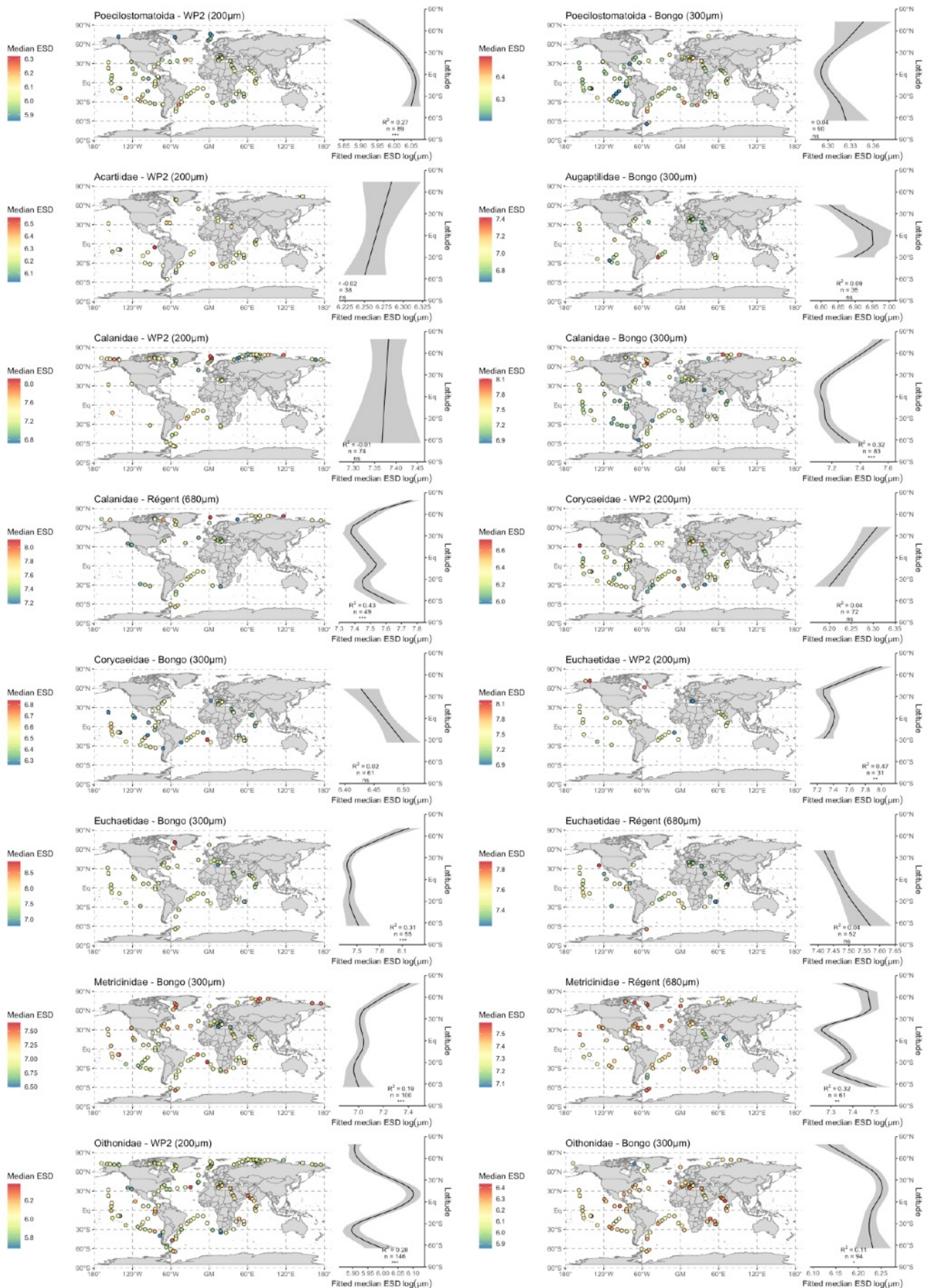


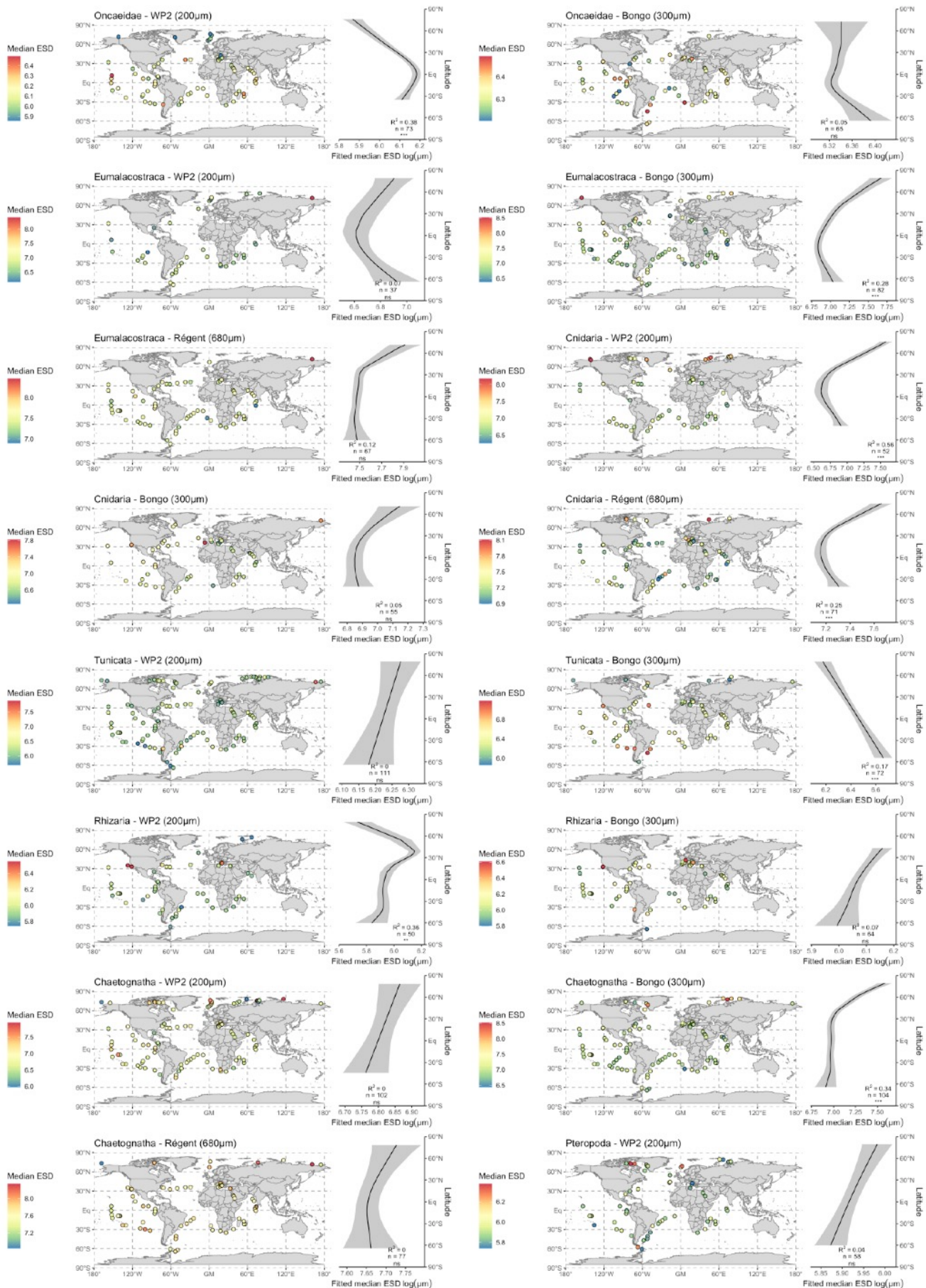
**Figure S4.2:** Pairwise correlations of cubic-transformed abundance distributions of the main zooplankton groups studied between the sampling depth of the net hauls (0-100m, WP2; 0-500m, Bongo). a) Total zooplankton, b) Copepoda, c) Rhizaria, d) Cnidaria, e) Tunicata, f) Eumalacostraca, g) Pteropoda, h) Chaetognatha, and i) Ostracoda + Cladocera. The results from the Spearman's rank correlation tests are shown on the plots.



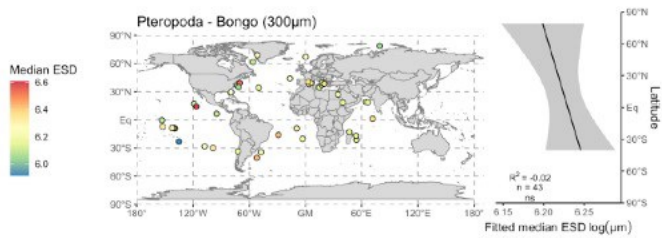
**Figure S4.3:** Pairwise correlations of cubic-transformed abundance distributions of the main zooplankton groups studied between the sampling depth of the net hauls (0-100m, WP2; 0-500m, Régent). a) Total zooplankton, b) Copepoda, c) Rhizaria, d) Cnidaria, e) Tunicata, f) Eumalacostraca, g) Pteropoda, h) Chaetognatha, and i) Ostracoda + Cladocera. The results from the Spearman's rank correlation tests are shown on the plots.





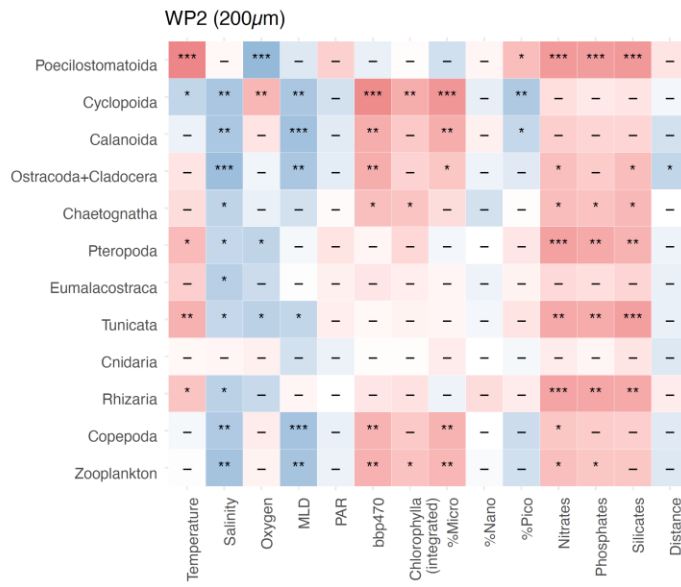




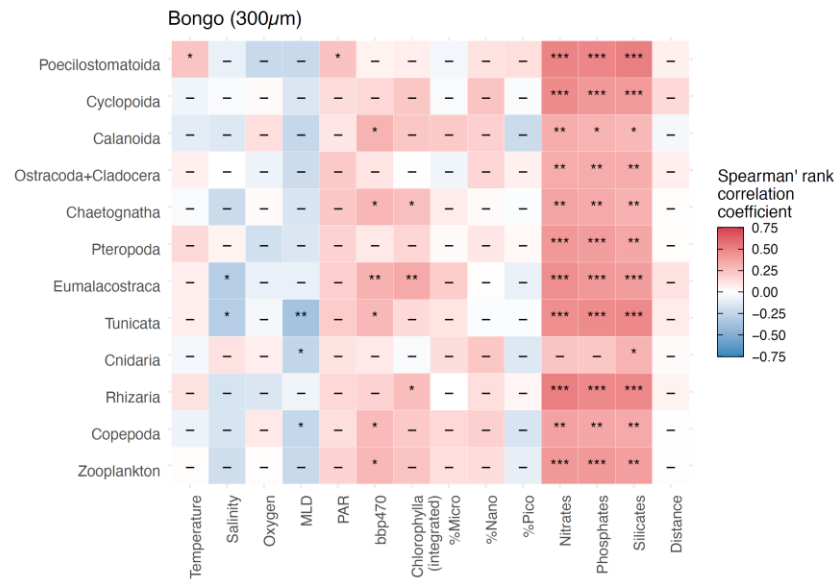


**Supplementary Figure S5:** Maps and latitudinal patterns of the logged median Equivalent Spherical Diameter (ESD,  $\mu\text{m}$ ) observed for the various zooplankton groups sampled with the three plankton nets (WP2, Bongo and Régent). The major and minor axes of the best fitting ellipses were measured for each organism to measure their ESD. Community-level size structure was estimated through the median value of the ESD distribution at individual-level. The solid curve on the zonal 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 of a group and only the groups with at least 30 stations were modelled here.

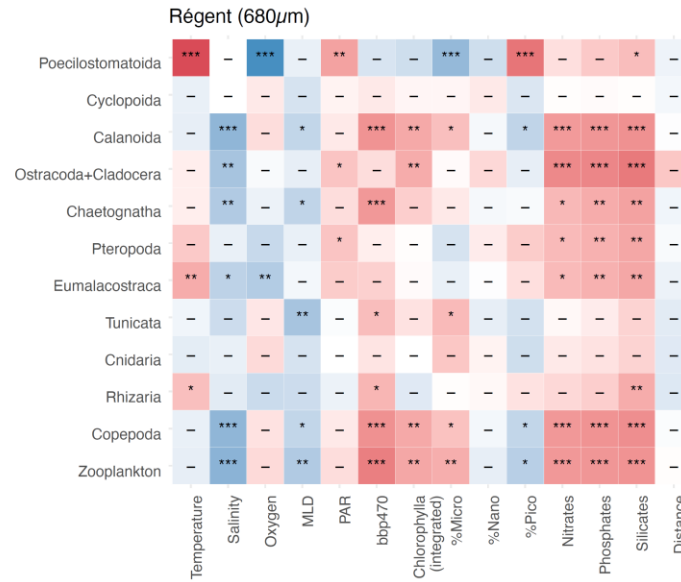
(a)

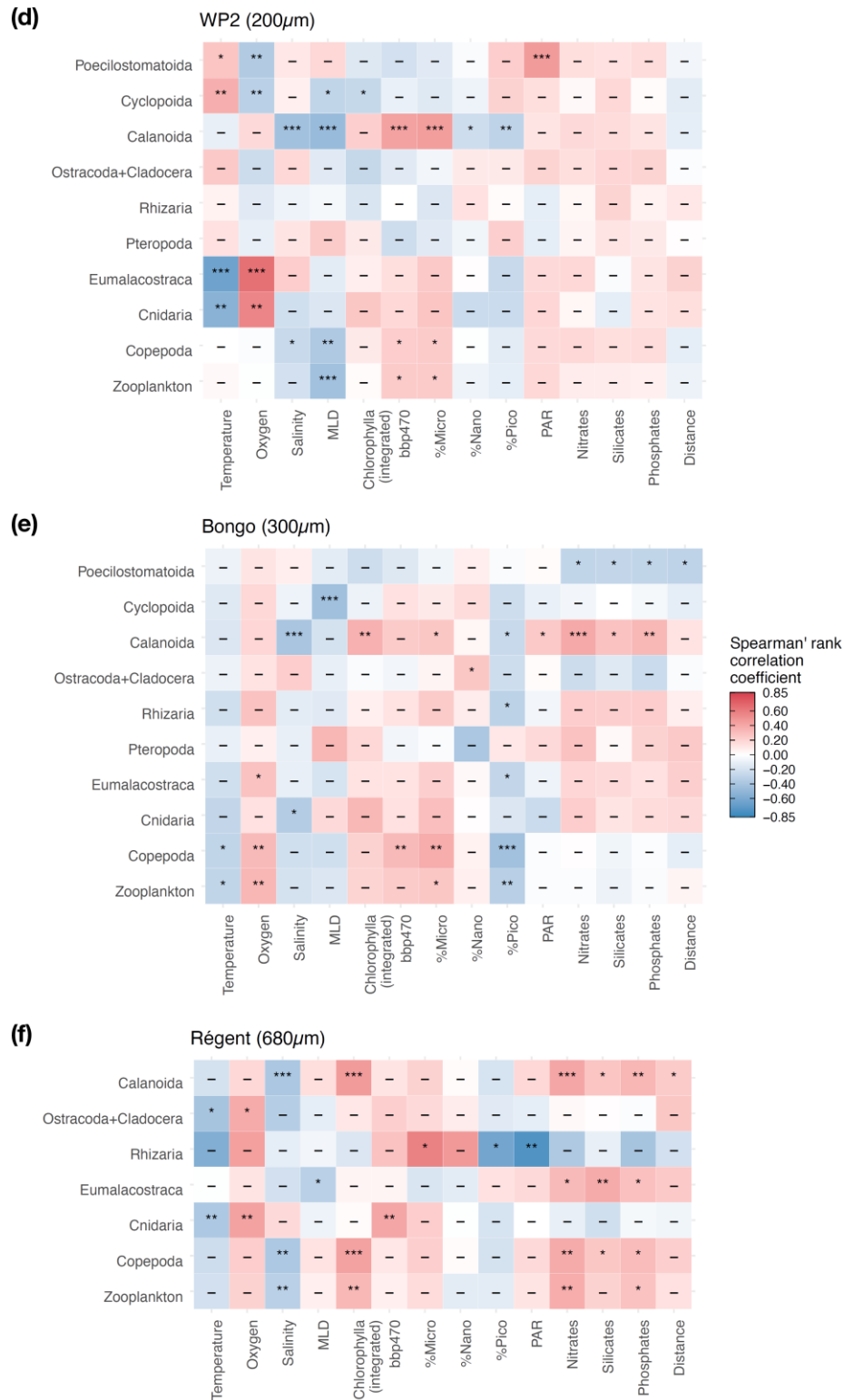


(b)

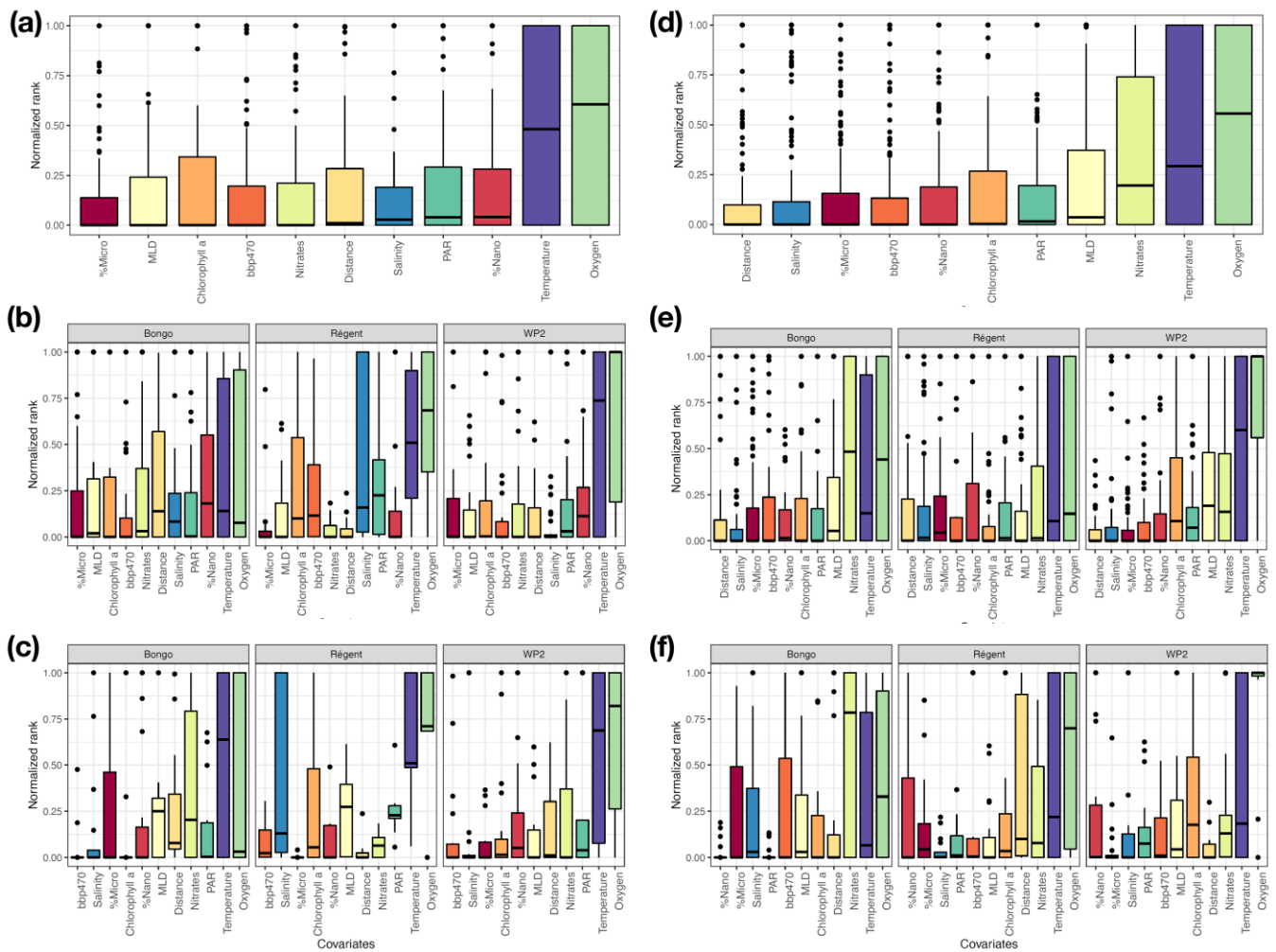


(c)

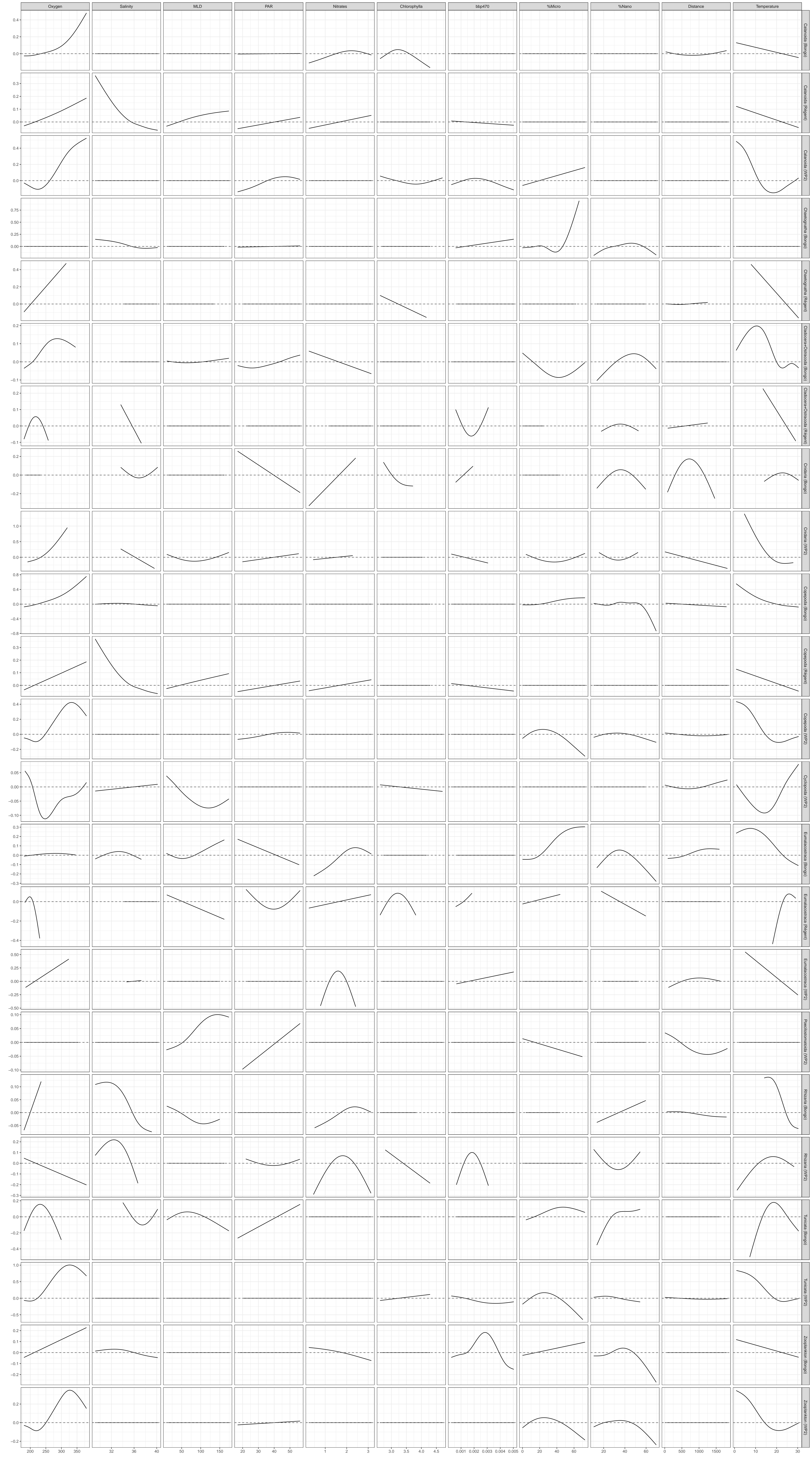




**Supplementary Figure S8:** Heatmaps of the Spearman's rank correlation coefficients computed between the (a)-(b)-(c) abundance (cubic-transformed) and the (d)-(e)-(f) median ESD (logged) of the zooplankton groups and the 13 environmental covariates as measured by the (a)-(d) WP2 net (200 $\mu$ m mesh), (b)-(e) Bongo net (300 $\mu$ m mesh) and (c)-(f) Régent net (680 $\mu$ 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 stations where ESD was measured for at least 20 individuals of a group were considered when computing the correlation coefficients.

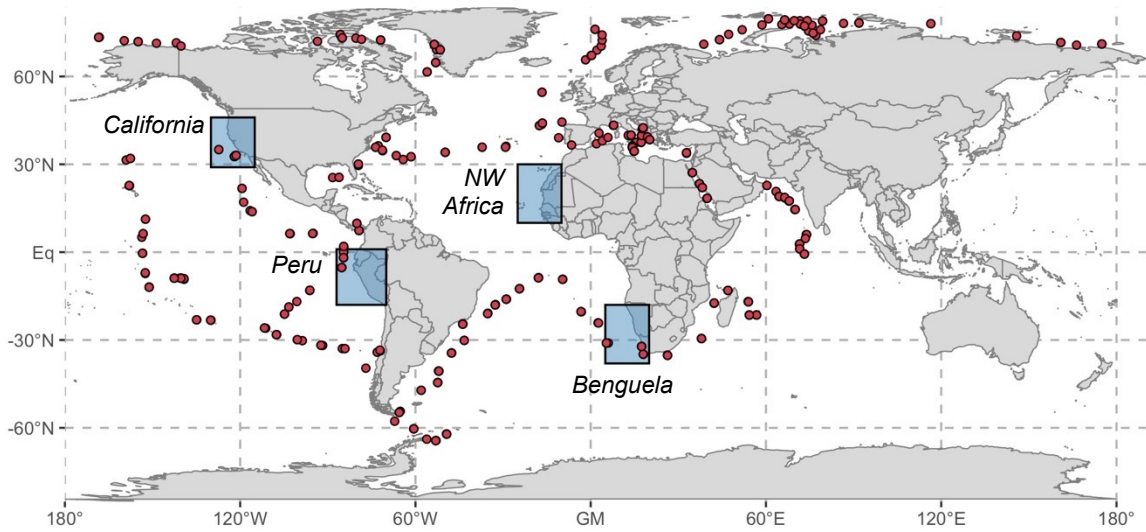


**Supplementary Figure S9:** Distribution of the normalized ranks of variable significance for the environmental covariates used in the the Generalized Additive Models (GAMs) fit to model the observed (a)-(b)-(c) median ESD (logged) and (d)-(e)-(f) abundance (cubic-transformed) measured for all zooplankton groups based on: (a)-(d) all plankton nets together (covariates are displayed in order of increasing significance of median normalized rank); (b)-(e) per net type; and (c)-(f) based on the groups shown in Tables 1 and 2. Ranks range between 0 (least important covariate) and 1 (most important covariate). The F statistic of the smooth term associated to each covariate was extracted and was used to quantitatively rank the covariates according to their relative significance (the higher the F value, the more significant) within their corresponding GAM. For each GAM, ranks were normalized by dividing the F value of every smooth term by the maximum F value across all smooth terms. The complete statistics of the GAMs are given in Supplementary Table S7.

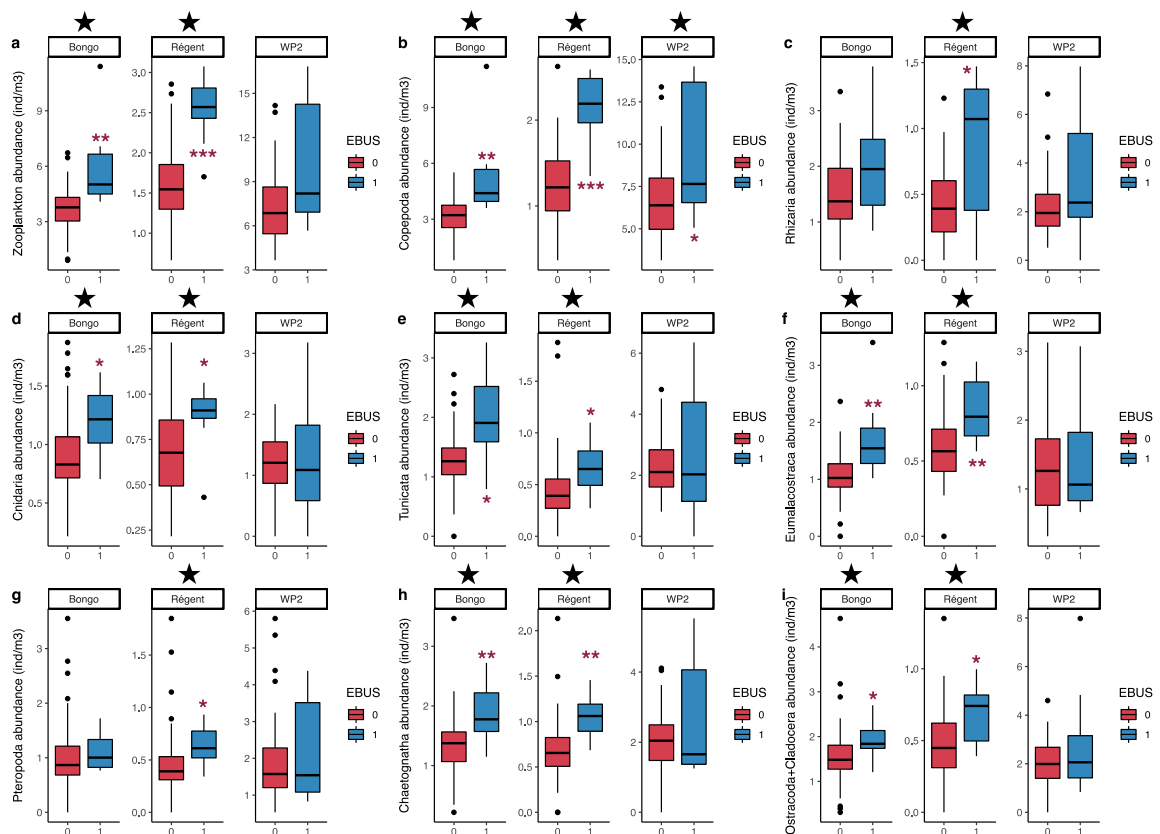




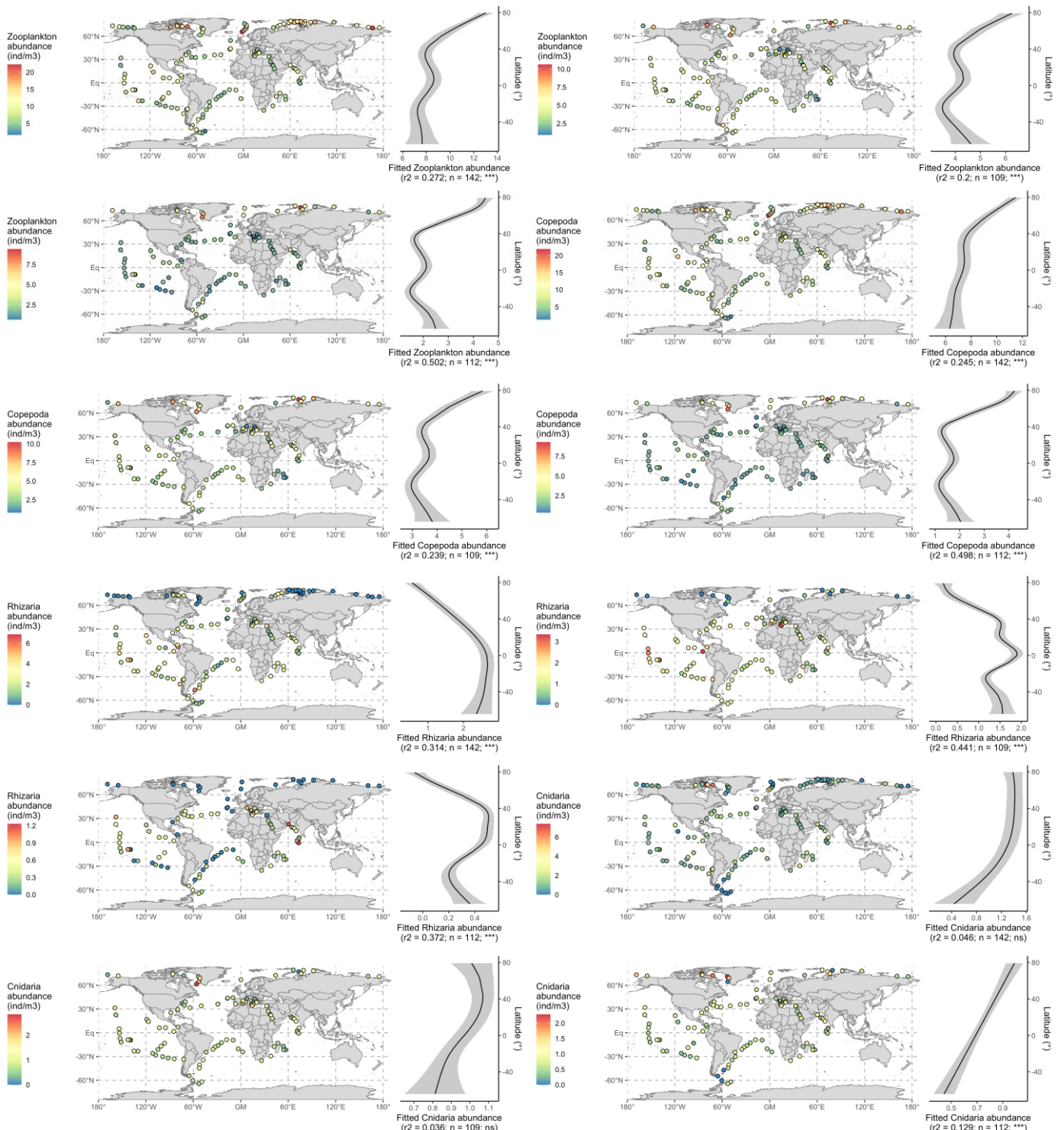
**Supplementary Document S12: Comparing zooplankton groups abundance estimates between stations located within Eastern Boundary Upwelling Systems (EBUS) and outside EBUS.**



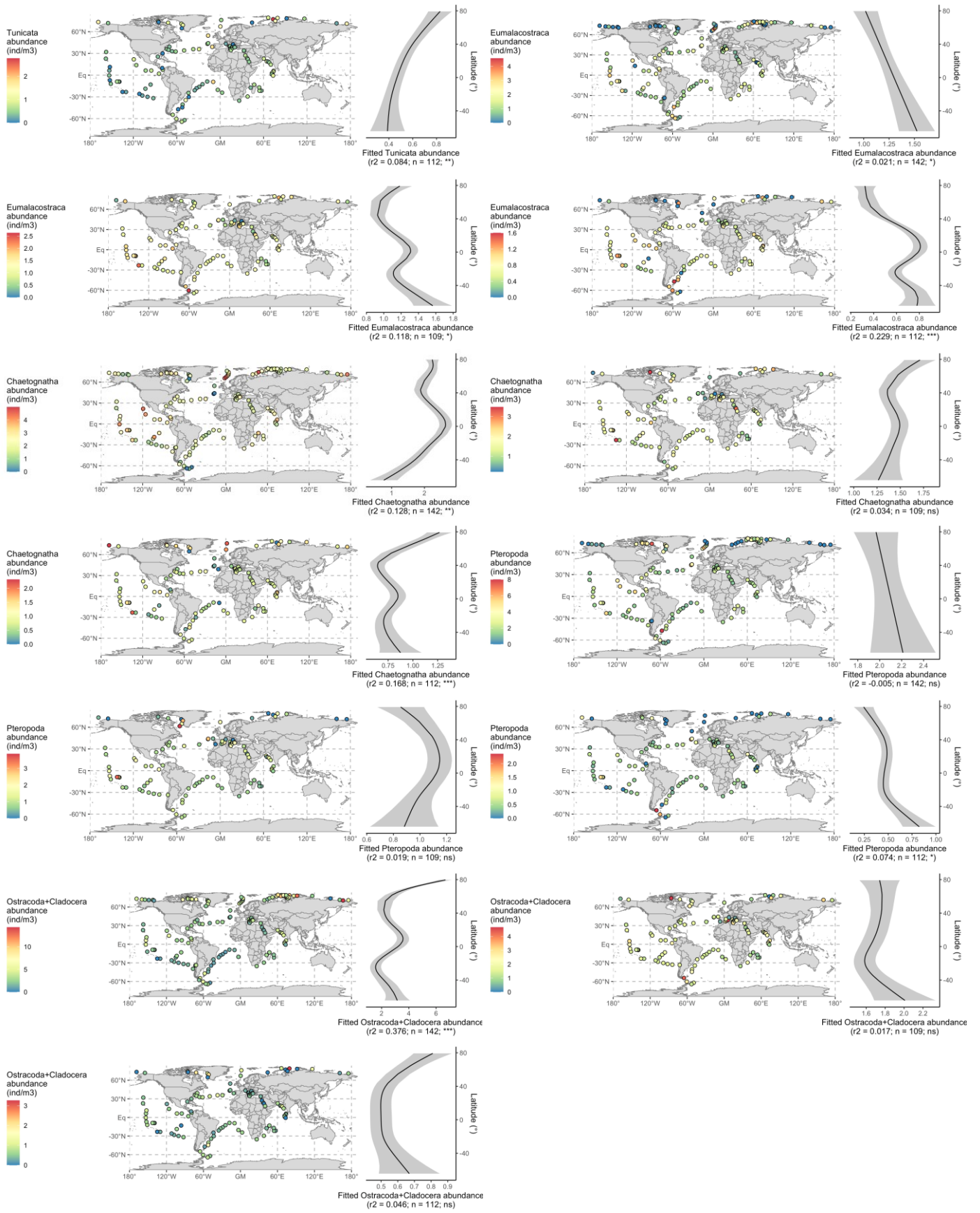
**Figure S12.1:** Geographic position of the Tara Oceans sampling stations (in red) from which the zooplankton abundance estimates used in our study come from and the position of the major Eastern Boundary Upwelling Systems (EBUS). The extent for the spatial boundary boxes are based on those shown in Figure 2 of Chavez & Messié (2009). Any sampling station whose spatial coordinates fell into one of the four EBUS was labelled as a “EBUS station”. The corresponding abundance estimates of the main zooplankton groups were compared to those stemming from the “non EBUS stations” that displayed a similar latitudinal range (i.e. 29°N-46°N and -30°N-1°N) as the EBUS stations (see below). This way, high latitude sampling stations were discarded from this analysis so we can focus on those abundance patterns that mostly reflect effects due to the EBUS conditions.



**Figure S12.2:** Comparing the cubic-transformed abundance distributions of the main zooplankton groups studied across three plankton nets (WP2, 200 $\mu$ m; Bongo, 300 $\mu$ m; Régent, 680 $\mu$ m) between the EBUS stations in blue (see Fig. SXX1) and the comparable non EBUS stations in red (i.e. stations located between 29°N-46°N and -30°N-1°N). a) Total zooplankton, b) Copepoda, c) Rhizaria, d) Cnidaria, e) Tunicata, f) Eumalacostraca, g) Pteropoda, h) Chaetognatha, and i) Ostracod + Cladocera. For each group and net type, non parametric variance analyses (Kruskal-Wallis tests) were performed to test if the EBUS stations displayed significant higher zooplankton abundances. The groups and nets for which a significant ( $p$ -value  $< 0.05$ ) was found between EBUS and non EBUS stations are highlighted with a black star. Conventional labels illustrate the level of significance of the test: \* =  $p$ -value  $< 0.05$ ; \*\* =  $p$ -value  $< 0.01$ ; \*\*\* =  $p$ -value  $< 0.001$ .





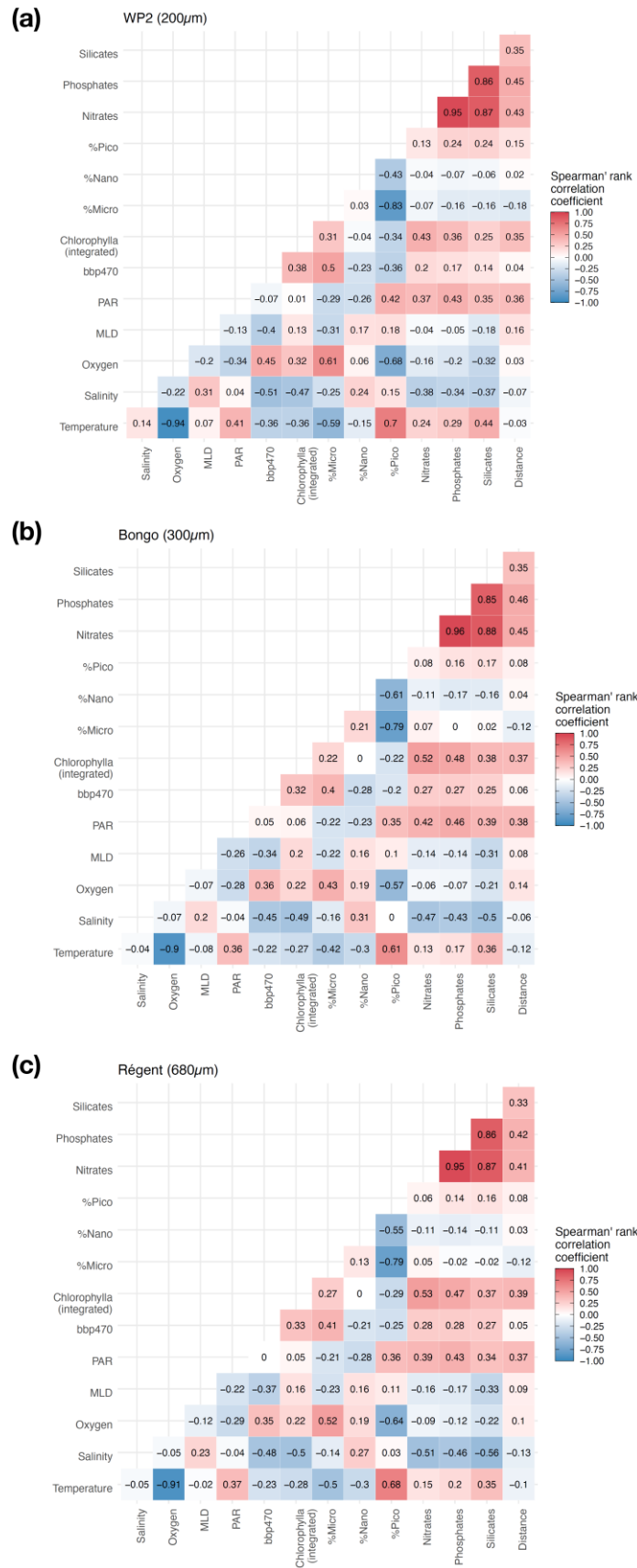


**Figure S12.3:** Maps and latitudinal patterns of the abundance (cubic-transformed ind.m<sup>3</sup>) of the main zooplankton groups sampled with three plankton nets (WP2, 200 $\mu$ m; Bongo, 300 $\mu$ m; Régent, 680 $\mu$ m), excluding those sampling stations that fall within an eastern boundary upwelling system (EBUS). The solid curves on the zonal side plots illustrate the prediction from the Generalized Additive Model (GAM) fitting transformed abundance against latitude. The explanatory power of the GAM (adjusted R<sup>2</sup>), 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 prediction. As in Supplementary Figure S2, the order of appearance of the groups' maps+plots panel is always as follows: WP2 data, Bongo data and then Régent data.

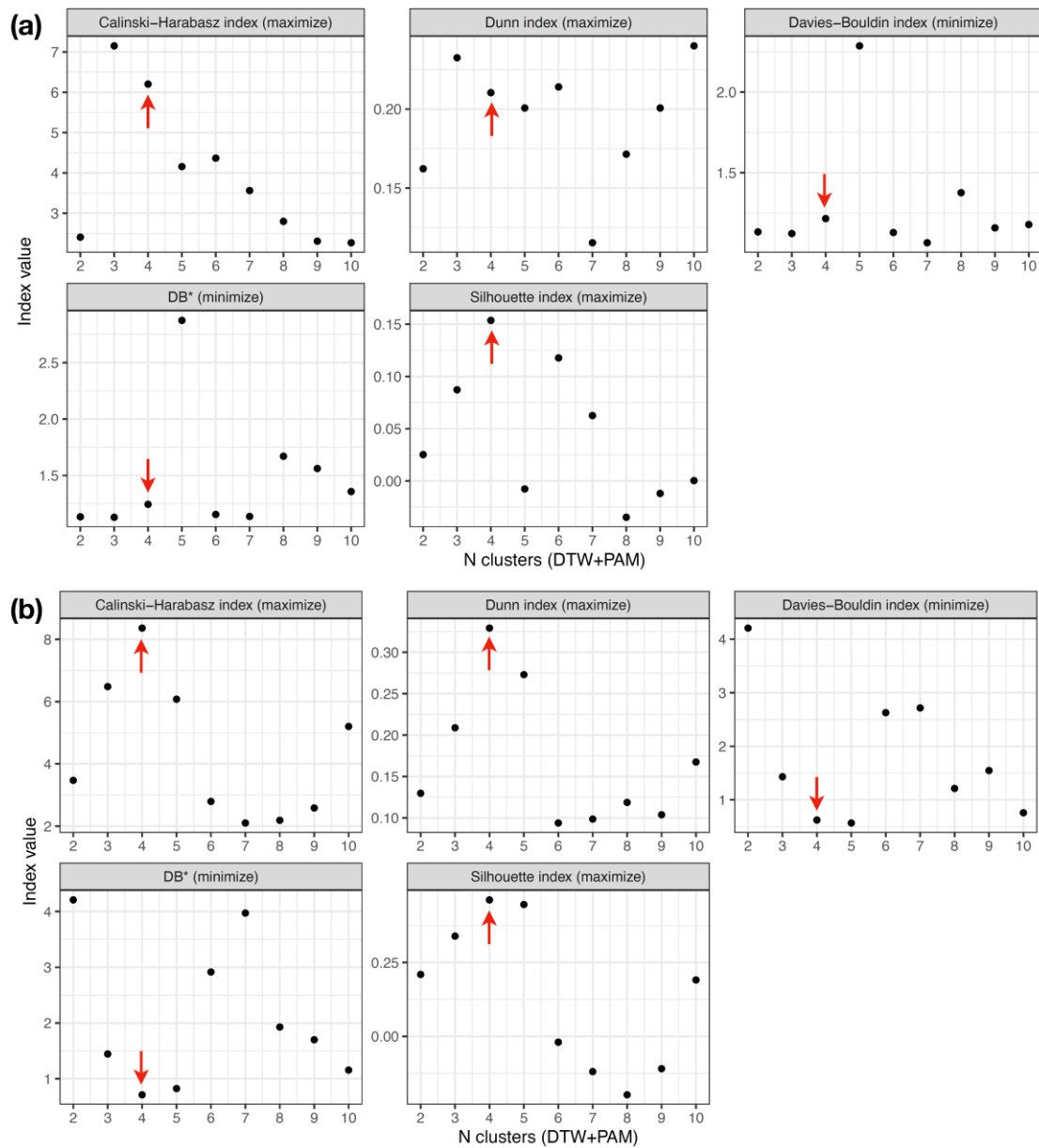
The effect of the EBUS stations on the cubic-transformed abundance latitudinal patterns modelled above can be examined by comparing these maps and zonal latitudinal plots to their counterparts shown in Supplementary Figure S2. Very few differences were found, which suggests that EBUS have a relative weak effect on the macroscale patterns of zooplankton abundances. However, we acknowledge that the sampling design of the *Tara* Oceans cruises does not allow to fully investigate this effect, as no stations sampled the northwestern African EBUS and relatively few stations were taken in the other three EBUS (Fig. S12.1). The most notable differences in latitudinal abundance patterns were found for the:

- **Copepoda (WP2):** the modelled increase in abundances near the equator was way less marked as the relatively high abundances observed in the Peru and Benguela EBUS were discarded.
- **Rhizaria (WP2):** the modelled decrease in abundances near 30°S was removed as the relatively low abundances observed in the Benguela EBUS were discarded.
- **Tunicata (WP2):** the modelled increase in abundances near the equator was less marked as the relatively high abundances observed in the Peru EBUS were discarded.

All the other abundance latitudinal patterns modelled were left unchanged by the removal of the EBUS stations. Therefore, although the EBUS seem to have a positive influence of the abundances of most of the main zooplankton groups studied here (Fig. S12.2), they do not strongly impact the modelled latitudinal gradients in zooplankton abundance.



**Supplementary Figure S14:** Heatmaps of the Spearman's rank correlation coefficients computed between the 13 covariates as measured at the stations sampled with the (a) WP2 net, (b) Bongo net, and (c) Régent net. Macronutrients (Nitrates, Phosphates and Silicates) concentrations were cubic-transformed and integrated Chlorophyll a concentration was log-transformed. See Methods for details regarding to the in situ measurements of these covariates.



**Supplementary Figure S15:** Profiles of the five indices of clustering quality measured for 2 to 10 clusters obtained by computing the *Dynamic Time Warping* (DTW) distance and performing Partitioning Around Medoids (PAM) clustering on the smooth curves resulting from the Generalized Additive Models (GAMs) fit on (a) the median ESD estimates of the zooplankton groups highlighted in Table 1, and (b) the abundance of the groups highlighted in Table 2 (one smooth curve per environmental covariate included in the GAMs; clustering was performed on a the multivariate data series obtained by stacking the smooth curves; only GAMs displaying a deviance explained >40% were considered). The red arrows highlight our final choice of the number of cluster which stems from the trade-off across the five indices examined. DB\* corresponds to the modified version of the Davies-Bouldin index. Indices values were obtained through the *cvi* function of the *dtwclust* R package. Whether maximum or minimum index values indicate higher or lower cluster quality is specified.