



HAL
open science

Insights into biotic and abiotic modulation of ocean mesopelagic communities

Janaina Rigonato, Marko Budinich, Alejandro Murillo, Manoela Brandão, Juan Karlusich, Yawouvi Dodji Soviadan, Ann Gregory, Hisashi Endo, Florian Kokoszka, Dean Vik, et al.

► **To cite this version:**

Janaina Rigonato, Marko Budinich, Alejandro Murillo, Manoela Brandão, Juan Karlusich, et al.. Insights into biotic and abiotic modulation of ocean mesopelagic communities. 2022. hal-03166899

HAL Id: hal-03166899

<https://hal.science/hal-03166899>

Preprint submitted on 16 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

1 **Insights into biotic and abiotic modulation of ocean mesopelagic communities**

2
3 Janaina Rigonato^{1,17*#}, Marko Budinich^{2,3*}, Alejandro A. Murillo⁴, Manoela C. Brandão⁵, Juan
4 J. Pierella Karlusich⁶, Yawouvi Dodji Soviadan⁵, Ann C. Gregory⁷, Hisashi Endo⁸, Florian
5 Kokoszka^{6,9}, Dean Vik⁷, Nicolas Henry², Paul Frémont¹, Karine Labadie¹, Ahmed A. Zayed⁷,
6 Céline Dimier^{2,6,17}, Marc Picheral^{5,17}, Sarah Searson⁵, Julie Poulain^{1,17}, Stefanie Kandels^{4,10},
7 Stéphane Pesant^{11,12}, Eric Karsenti^{6,10}, The Tara Oceans coordinators[§], Peer Bork^{6,13}, Chris
8 Bowler^{6,17}, Samuel Chaffron^{3,17}, Colombaro de Vargas^{2,17}, Damien Eveillard^{3,17}, Marion
9 Gehlen¹⁴, Daniele Iudicone⁹, Fabien Lombard^{5,17}, Hiroyuki Ogata⁸, Lars Stemmann^{5,17},
10 Matthew B. Sullivan^{7,15}, Shinichi Sunagawa^{4,16}, Patrick Wincker^{1,17} and Olivier Jaillon^{1,17#}

11
12
13 1 Génomique Métabolique, Genoscope, Institut de Biologie François Jacob, Commissariat à
14 l'Energie Atomique (CEA), CNRS, Université Evry, Université Paris-Saclay, 91000 Evry,
15 France

16 2 Sorbonne Université, CNRS, Station Biologique de Roscoff, AD2M, UMR 7144, 29680
17 Roscoff, France

18 3 Université de Nantes, CNRS UMR 6004, LS2N, F-44000 Nantes, France

19 4 Structural and Computational Biology, European Molecular Biology Laboratory, Meyerhofstr.
20 1, 69117 Heidelberg, Germany

21 5 Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche (LOV), 06230
22 Villefranche-sur-Mer, France

23 6 Institut de Biologie de l'ENS (IBENS), Département de biologie, Ecole normale supérieure,
24 CNRS, INSERM, Université PSL, 75005 Paris, France

25 7 Department of Microbiology, The Ohio State University, Columbus, OH 43214, USA

26 8 Bioinformatics Center, Institute for Chemical Research Kyoto University, Gokasho, Uji,
27 Kyoto, 611-0011, Japan

28 9 Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy

29 10 Directors' Research European Molecular Biology Laboratory Meyerhofstr. 1 69117
30 Heidelberg Germany

31 11 MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen,
32 Germany

33 12 PANGAEA, Data Publisher for Earth and Environmental Science, University of Bremen,
34 Bremen, Germany

35 13 Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany

36 14 Institut Pierre Simon Laplace, Laboratoire des Sciences du Climat et de l'Environnement,
37 CEA, CNRS, Université Paris-Saclay, 91191 Gif-sur-Yvette cedex, France

38 15 Department of Civil, Environmental and Geodetic Engineering, The Ohio State University,
39 Columbus OH 43214 USA

40 16 Department of Biology; Institute of Microbiology and Swiss Institute of Bioinformatics, ETH
41 Zurich; Zurich, 8093; Switzerland

42 17 Research Federation for the study of Global Ocean Systems Ecology and Evolution,
43 FR2022/Tara Oceans GOSEE, 3 rue Michel-Ange, 75016 Paris, France

44
45 *These authors share the first authorship

46 # Corresponding authors: jana_rigonato@protonmail.com (JR);

47 ojaillon@genoscope.cns.fr (OJ)

48
49
50 Running title: Mesopelagic community ecology

51
52
53
54

55 **Abstract**

56

57 Marine plankton mitigate anthropogenic greenhouse gases, modulate biogeochemical cycles,
58 and provide fishery resources. Plankton is distributed across a stratified ecosystem of sunlit
59 surface waters and a vast, though understudied, mesopelagic 'dark ocean'. In this study, we
60 mapped viruses, prokaryotes, and pico-eukaryotes across 32 globally-distributed cross-depth
61 samples collected during the *Tara* Oceans Expedition, and assessed their ecologies. Based
62 on depth and O₂ measurements, we divided the marine habitat into epipelagic, oxic
63 mesopelagic, and oxygen minimum zone (OMZ) eco-regions. We identified specific
64 communities associated with each marine habitat, and pinpoint environmental drivers of dark
65 ocean communities. Our results indicate that water masses primarily control mesopelagic
66 community composition. Through co-occurrence network inference and analysis, we identified
67 signature communities strongly associated with OMZ eco-regions. Mesopelagic communities
68 appear to be constrained by a combination of factors compared to epipelagic communities.
69 Thus, variations in a given abiotic factor may cause different responses in sunlit and dark
70 ocean communities. This study expands our knowledge about the ecology of planktonic
71 organisms inhabiting the mesopelagic zone.

72

73 **Keywords:** mesopelagic community, metabarcoding, plankton, pan-oceanic expedition,
74 oxygen minimum zone

75

76 **Introduction**

77
78 Below the ocean's sunlit layer lies the mesopelagic zone that occupies around 20% of the
79 global ocean volume [1]. The mesopelagic zone is biologically defined as starting where
80 photosynthesis no longer occurs (<1% irradiance around 200m depth), down to its lower
81 boundary where there is no detectable sunlight (around 1000m depth) [2]. This twilight
82 ecosystem cannot rely on photoautotrophy, but sustains its energetic requirements by the
83 combination of heterotrophic, chemoautotrophic, and chemo-mixotrophic metabolisms,
84 together with physicochemical processes. Among the latter, the fraction of upper ocean
85 productivity that escapes epipelagic recycling and sinks by gravity or is delivered by the daily
86 migration of zooplankton constitutes an essential energy source in deep waters and is a vector
87 for attached organisms [3].

88 The biodiversity and biomass in mesopelagic communities have been underestimated in the
89 past [2, 4], and previous work showed these communities hold an enormous unexploited
90 biological resource [5, 6, 1]. Mesopelagic organisms are considered to be a vast source of fat
91 and protein, potentially becoming the primary source of global bioeconomy [7]. So far, efforts
92 have been made to increase knowledge of mesopelagic mega/macrofauna by studying the
93 abundance and diversity of nekton. These efforts are of great importance, given the rapid
94 increase in the exploitation of this zone by nutraceutical and fisheries industries [6]. However,
95 less attention has been devoted to the mesopelagic community's microscopic fraction, despite
96 the pivotal role of the marine microbiome in biogeochemical cycles. The marine microbiome
97 makes crucial links in the food web between primary production and dark ocean specialized
98 consumers. Previous reports have shown stratification of planktonic communities with depth.
99 In this regard, the mesopelagic zone displays a distinct assemblage of dsDNA viruses [8], giant
100 viruses [9], prokaryotes [10, 11], and eukaryotes [12]. However, unlike the epipelagic layer,
101 mesopelagic plankton diversity does not show the latitudinal diversity gradient trends from
102 pole-to-pole, peaking at lower latitudes [13].

103 Among the studies conducted in mesopelagic zones, we highlight the efforts to explore regions
104 of extreme conditions, such as oxygen minimum zones (OMZs). These zones are formed by

105 relatively old slowly upwelling waters, often lying below highly productive surface zones [14],
106 and are currently increasing in volume in the oceans [15]. OMZ prokaryotic communities are
107 well documented and predominated by taxa such as Nitrospira, Marinimicrobia, and anammox
108 bacteria from the phylum Planctomycetes, while Thaumarchaeota abundance is frequently
109 lower in these zones [16, 17, 18, 19, 20]. In contrast, knowledge of eukaryotic diversity in OMZs
110 is still rudimentary, but a prevalence of specific taxa such as Ciliophora, Dinoflagellata, MALV,
111 and Acantharia has been reported, together with a higher metabolic activity of these taxa [21,
112 22, 23]. Understanding plankton community structure and dynamics is fundamental to
113 anticipate the impacts of global warming and acidification in these regions.

114 The last decades have seen a significant increase in large-scale oceanic surveys [24, 25, 26].
115 However, most mesopelagic community studies have been limited to geographically or
116 ecologically fragmented regions, or to specific taxonomic groups, mainly because of the
117 inherent difficulties of accessing this zone on a global scale [5]. Hence, these studies have
118 given us a limited picture of community composition. Moreover, the factors influencing
119 community structure, presumably a combination of biotic and abiotic factors [27, 28], have
120 been little explored in the mesopelagic zone.

121 The present study takes advantage of the *Tara* Oceans large-scale survey conducted in
122 different water layers using a systematic sampling protocol, spanning viruses to small
123 eukaryote size fractions, to investigate the mesopelagic biome [29]. We capitalized on genomic
124 data together with extensive contextual data and ocean geography to explore the particularities
125 of mesopelagic communities compared to communities found in the euphotic zone. We also
126 investigated potential water deoxygenation effects on these communities by comparing OMZ
127 communities with those from well-oxygenated waters. This work expands our knowledge of
128 the web of relationships underpinning mesopelagic plankton ecosystems on a broad
129 geographic scale.

130

131 **Materials and Methods**

132 **• Sample collection and pre-processing**

133 The environmental and biological data were obtained during the *Tara* Oceans expedition
134 (2009- 2012) in 32 oceanographic stations located in the Indian Ocean (IO - 037, 038, 039),
135 Pacific Ocean (PO - 097, 098, 100, 102, 106, 109, 110, 111, 112, 122, 131, 132, 133, 135,
136 137, 138), South Atlantic Ocean (SAO - 068, 070, 072, 076, 078) and North Atlantic Ocean
137 (NAO - 142, 143, 144, 145, 146, 148, 149, 152) comprising tropical and subtropical regions
138 (Figure 1). Physico-chemical environmental data were obtained along a vertical profile at each
139 station. Temperature, salinity, and oxygen were measured using a CTD-rosette system with a
140 coupled dissolved oxygen sensor. Chlorophyll-a concentrations were measured using high-
141 performance liquid chromatography. Nutrient concentrations were determined using
142 segmented flow analysis. Metadata are available at PANGAEA [30, 31, 32, 33, 34, 35] -
143 (<https://doi.pangaea.de/10.1594/PANGAEA.875582>).

144 The vertical distribution of marine particles was investigated with an Underwater Vision Profiler
145 (UVP, [36, 37]) mounted on the CTD-Rosette. The UVP acquires images in a coherent water
146 volume (1 L) delimited by a light sheet issued from red light-emitting diodes. Automatic
147 identification of objects was made using Ecotaxa, based on a learning set of visually identified,
148 manually classified objects and associated features. Images were classified to distinguish
149 mesozooplankton from non-living objects and artifacts (e.g., detritic particles, fibers, and out-
150 of-focus objects).

151 Water vertical profiles of temperature and salinity generated from the CTD were used to identify
152 the water masses by plotting a temperature x salinity (T/S) diagram using the Ocean Data View
153 V 5.0 (ODV) software package [38].

154 Three different water layers were sampled: surface (SRF, 3-7 m), deep chlorophyll maximum
155 (DCM - depth identified according to the peak of chlorophyll-a fluorescence obtained *in situ*),
156 and mesopelagic (ranging from 200-1000 m) [39]. The planktonic community was sampled by
157 partitioning the seawater by filtering each sampled depth with different filter sizes [34]. Among
158 the mesopelagic zones, 13 of them were identified as deficient in oxygen and classified as

159 oxygen minimum zone (OMZ, stations IO - 037, 038, 039 / PO - 100, 102, 106, 109, 110, 111,
160 133, 135, 137, 138). The OMZ were categorized as suboxic: $<10 \mu\text{M O}_2/\text{kg}$ seawater and
161 anoxic: ($<0.003 \mu\text{M}/\text{kg}$ seawater or undetectable with most sensitive techniques, e.g., STOX
162 sensors) [Units of O_2 concentration: $1 \text{ mL}\cdot\text{L}^{-1}=1.43 \text{ mg}\cdot\text{L}^{-1}=44.64 \mu\text{M}$] [22].

163 Our dataset comprises different organismal size-fractions from viruses (two dsDNA-virus
164 families *Podoviridae* and *Myoviridae* - hereafter named as phages and NCDLV giant viruses -
165 hereafter named as giruses) to pico-eukaryotes. Phage libraries were constructed from
166 seawater samples filtered at $0.22 \mu\text{m}$, concentrated using iron chloride flocculation, and treated
167 with deoxyribonuclease (DNase). Girus *polB* and prokaryotic 16S rDNA sequences were
168 extracted from plankton metagenomes sequenced from $0.22\text{--}1.6$ or $0.22\text{--}3 \mu\text{m}$ filters, and the
169 pico-eukaryote dataset was obtained by V9-18S rDNA marker amplification from $0.8\text{--}3$ or 0.8--
170 $5 \mu\text{m}$ filters. Details of sample preparation and sequencing procedures are described in Alberti
171 et al. [40].

172 Phage relative abundance was accessed through the search for the marker genes *gp23*
173 (*Myoviridae*) and *polA* (*Podoviridae*) in the protein collection GOV2.0 derived from
174 metagenomic sequencing described in Gregory et al. [8]. The girus abundance profile was
175 obtained from *polB* marker gene gathered from the OM-RGC.v2 catalog [11] as described in
176 Endo et al. [9]. The Prokaryote 16S rDNA marker derived from the metagenome assembly,
177 named 16S Mitag, is described in Sunagawa et al. [10]. Sequences matching “Eukaryota”,
178 “chloroplasts”, and “mitochondria” were removed from the final table. Clustering and annotation
179 of pico-eukaryote V9-18S rDNA amplicons are described in de Vargas et al. [41], and functional
180 annotation of taxonomically assigned V9-18S rDNA metabarcodes was improved afterward; in
181 this case, we conserved in the final data only sequences assigned to the “Eukaryota” domain.

182 We concatenated SRF and DCM samples for each taxonomic group to obtain an epipelagic
183 dataset (EPI). Counts of OTUs shared in SRF and DCM samples were summed. OTU
184 abundance was normalized by the total counts for each taxonomic group within each sample.

185

186 • **Ecological Analysis**

187 **Epipelagic and Mesopelagic Community and Environmental differences**

188 We applied an NMDS analysis based on the Bray-Curtis dissimilarity matrix on relative
189 abundances using the ‘metaNMDS’ function from the vegan R package [42] to confirm
190 community differences between epipelagic and mesopelagic layers. Homogeneity of the
191 sampled environmental parameters was checked using the ‘betadisper’ function (Homogeneity
192 of multivariate dispersions in the vegan package). The analysis was conducted using the
193 Euclidean distance matrix of the environmental variables using the depths (epipelagic,
194 mesopelagic) as group factor. A permutation test statistically confirmed the results.

195

196 **Ecological model selection**

197 We used the Species Abundance (SAD) and Rank Abundance Distribution curves (RAD: log
198 abundances vs. rank order) to fit some of the most popular ecological models assessing how
199 communities are assembled [43, 44]. Niche/deterministic models presume that the community
200 is under selection due to biotic and abiotic interactions, while neutral/stochastic models
201 assume that random processes structure the community such as drift, migration, birth, and
202 death [45, 46]. SAD is essential for describing and understanding community assemblage and
203 its management [44]. Further, by showing logarithmic species abundance against a rank order,
204 RAD is commonly used to investigate a community’s structure from observations made at one
205 point in space and time [43, 44]. All the abundance distributions were fitted using ‘fitsad’ (sads
206 R package - [47]), and ‘radfit’ (vegan) functions using maximum likelihood estimation. A set of
207 candidate models was selected *a priori* to be tested by SAD: Log-series, Poisson-lognormal,
208 Broken Stick, Power Law, and the neutral ecological model Zero-sum multinomial distribution.
209 The neutral ecological model describes the SAD of a sample taken from a neutral
210 metacommunity under random drift. The models selected to be tested by RAD were niche pre-
211 emption models (geometric series or Motomura model), lognormal, Zipf, and Zipf–Mandelbrot,

212 and the null model that infers that the individuals are randomly distributed among observed
213 species. For a complete description of each model see Magurran and McGill [48].

214 The Akaike's Information Criterion (AIC) was used to evaluate the fitted model's quality based
215 on log-likelihood penalized by the number of estimated parameters. AIC estimates the loss of
216 information if the model is assumed for a given dataset. In this manner, models with lower AIC
217 values are selected as better fit conditions [49]. The AIC values produced for each model were
218 compared using the delta AIC (Δ_i). A Δ_i value <2 indicates equally likely models, values $3 < \Delta_i$
219 <9 indicate less likely models, and $\Delta_i > 10$ for no likely models [49].

220

221 **Ecological inferences and statistics**

222 Ecological patterns were inferred using environmental variables to constrain the variation
223 observed in biological data for planktonic samples using Canonical Correspondence Analysis
224 (CCA) in the vegan R package. A set of physico-chemical variables for the discrete depths
225 were selected for the ecological inferences, such as nitrate (NO_3^-), oxygen, temperature,
226 salinity, density, and particles using particle flux UVP data. In order to avoid collinearity among
227 factors, the selected variables were checked for variance inflation factor using the `vif.cca`
228 function and tested for significance by 'anova' implemented in vegan with 999 permutations.
229 Each variable effect significance was tested individually using all the others as covariables
230 (independently from their order in the model) by applying the option 'margin' to the 'anova'
231 function in vegan.

232 Permutational multivariate analysis of variance (PERMANOVA) was performed with the
233 function 'adonis' in vegan to determine the relationship between mesopelagic community
234 composition and predefined water masses based on 999 permutations.

235

236 **Organism Eco-region classification**

237 In order to detect organisms specific to epipelagic (EPI), oxic mesopelagic (Oxic MES), and
238 OMZ eco-regions, only data containing both epipelagic and mesopelagic information were

239 considered: in total, 25 stations for giruses, prokaryotes, and pico-eukaryotes, and 13 for
240 phages. First, we discarded OTUs with fewer than 100 reads to remove biases due to rare
241 species' presence, and then recalculated the relative abundances for each dataset. Next, we
242 ran a Kruskal-Wallis test ('kruskal.test' from stats R package [50]) to detect differential OTU
243 relative abundances between eco-regions, followed by a Bonferroni correction to avoid Type I
244 error. Organisms with a p-value <0.05, indicating a difference within groups, were subject to a
245 post-hoc Dunn test ('dunn.test' from dunn.test R package [51]) to identify preferential Eco-
246 regions for each OTU. From these results, OTUs statistically equally abundant in all Eco-
247 regions or non-significant Kruskal-Wallis tests were assigned to the "ubiquitous" group. In
248 contrast, those with significant p-values were classified as EPI, Oxidic MES, or OMZ if only the
249 corresponding Eco-region was elected according to the Dunn test. Organisms with no
250 significant differences between Oxidic MES and OMZ were assigned to Core MES.

251

252 **Co-occurrence Network**

253 For investigation of associations between organisms across Eco-regions, a co-occurrence
254 network was inferred. In this analysis, phage samples were not included due to the lower
255 number of stations sampled. Therefore, samples for giruses, prokaryotes and pico-eukaryotes
256 from stations 038, 039, 068, 070, 072, 076, 078, 098, 100, 102, 109, 110, 111, 112, 122, 132,
257 133, 137, 138, 142, 145, 146, 148, 149 and 152 were retained. OTUs with a relative abundance
258 $<10^{-4}$ and counting fewer than 5 observations were discarded. Next, each sample was
259 normalized by applying a centered-log-ratio (CLR) transformation. Network inferences were
260 performed using Flashweave version 0.18 developed in Julia version 1.2 [52], using the
261 sensitive and heterogeneous mode.

262 We analyzed this global co-occurrence network by delineating communities (or modules) using
263 the Clauset-Newman-Moore algorithm [53]. These modules are subsets of OTUs, obtained by
264 maximizing the co-occurrences within the module and minimizing connections between them.
265 Next, we investigated modules enriched in OTUs from specific Eco-regions using Fisher's
266 exact test using the "fisher.test" function from the stats R package.

267

268 **Results and Discussion**

269

270 Leveraging the resources produced by the *Tara* Oceans project, we deciphered differences
271 between epipelagic and mesopelagic beta-diversity stratification, with a particular emphasis
272 on the role of environmental variables such as temperature, oxygen, salinity, NO_3^- , chlorophyll-
273 a, and particle flux (see Methods). As previously reported, we observed a stratification by depth
274 of the epipelagic and the mesopelagic communities (i.e., phages, giruses, prokaryotes, and
275 pico-eukaryotes) (Supplementary Figure 1). Consequently, we investigated differences among
276 epipelagic and mesopelagic sampling sites based on Euclidean distance of physicochemical
277 characteristics from each site. We observed a high dissimilarity gradient among sites for both
278 layers (Supplementary Figure 2a, b). Mesopelagic samples were spread in the plot, with most
279 of the points placed distant from the group centroid (located in the center of the cloud of points
280 identified for each group) (Supplementary Figure 2a). In contrast, epipelagic points displayed
281 a large variance due to a few samples positioned apart from the main cluster (Supplementary
282 Figure 2a). These results probably underlie the heterogeneity of environmental conditions
283 encountered in both sampled layers, and this environmental variation may be an important
284 factor that can directly influence community composition.

285 Community composition variations can also be shaped by four main eco-evolutionary
286 processes: selection, dispersal, drift, and speciation [54]. Mathematical models based on these
287 processes and applied to species abundance (SAD) and rank abundance (RAD) distributions
288 are historically used to infer the ecological or evolutionary mechanisms that structure a given
289 community [55]. Here, we addressed these biological processes to infer the abundance
290 distributions as observed in such natural communities [56, 57] in the most diverse
291 environments (terrestrial and aquatic) [45, 46, 43]. We used both SAD and RAD as evidence
292 of ecological processes related to variations in plankton community composition. When applied
293 at the local scale, we excluded random/neutral evolutive effects because all epipelagic and

294 mesopelagic communities fit the niche/deterministic ecological models. In fact, in all cases, we
295 found a Δ_i value <2 for both methods evaluated, which allows us to select one best-fitting model
296 without a subjective judgment (Supplementary Material Table 1 and 2) [49]. In these cases,
297 the community assemblages are mediated by a combination of environmental conditions,
298 interspecies interactions (competition, predation, or mutualisms), and species traits (for
299 instance, phototrophy, parasitism). Specifically, 93% and 65% of epipelagic and mesopelagic
300 samples better fit the Lognormal model (Poisson Lognormal SAD [58] and Lognormal RAD
301 [59]), respectively). These results suggest that EPI and MESO plankton populations are
302 affected by the combination of many independent variables, including competitive biotic
303 interactions and abiotic factors [43]. According to this model, the community has a broad and
304 elementary form of organization [57]. Ser-Giacomi et al. [60] stated that the eukaryotic rare-
305 biosphere (non-dominant OTU's) composed of "transient" or occasional taxa in the ocean sunlit
306 layer are mainly governed by dispersal/neutral events. In contrast, and in the case of our
307 findings, the dominant fraction of the assemblage follows idiosyncratic environmental
308 conditions.

309 Next, to quantify how much of the differences in the assemblages' variance can be explained
310 by environmental conditions, we employed canonical correspondence analysis (CCA) using
311 the environmental variables measured at discrete depths as constraint variables. The results
312 showed that the environment explains only a small fraction of community variance for both
313 layers (32% on average) (Figure 2). The phage assemblage was the exception, for which about
314 55% of the epipelagic variation and 65% of the mesopelagic variation could be explained by
315 the variables investigated (Figure 2). We tested the variance explained by single explanatory
316 variables individually. In contrast with epipelagic communities mainly governed by temperature
317 and oxygen, as observed here and elsewhere [10, 61, 8, 13, 12, 62], we could not identify a
318 single environmental predictor structuring entire mesopelagic assemblages. However, a few
319 different variables appeared to be significant for each group (Table 1, complete analysis in
320 Supplementary Material Table 3). Notably, our analysis identified oxygen as the main

321 mesopelagic driver for phage and prokaryote assemblages, confirming previous reports [20,
322 63, 64].

323 Even though we observed in the ordination-plots the distinction of OMZ and oxic mesopelagic
324 (Oxic MES) stations for viruses and pico-eukaryote assemblages (Figure 2, diamond and down
325 triangle), we could not disentangle the effect of oxygen from the other variables included in the
326 analyses (Table 1, complete analysis in Supplementary Material Table 3). This result shows
327 that these assemblages are probably affected by a combination of the predictors evaluated,
328 reflecting their need to cope with a broader environmental gradient that maximizes their niche-
329 space partitioning. Previous studies have identified oxygen as one of the main drivers of the
330 eukaryotic community structure in OMZ regions [65, 22, 23]. These studies mainly compared
331 community composition along the oxygen gradient within the water column depth, from the
332 surface downwards. However, depth stratification of plankton communities is evident even in
333 regions with high oxygen concentrations, so distinct parameters co-varying with depth must be
334 taken into account in addition to the oxygen gradients [66].

335 In addition to the physicochemical parameters, our results show that particle flux derived from
336 UVP measurements was also a significant variable structuring the phage assemblages in both
337 epipelagic and mesopelagic layers (Table 1, complete analysis in Supplementary Material
338 Table 3). This data supports previous reports about the high correlation of this environmental
339 factor with phages, finding possible relevance for the carbon pump's functioning in epipelagic
340 layers [67]. This observation may also reflect the association with virus inputs from overlaying
341 water layers via sinking particles [68].

342 *In situ* physico-chemical measurements have revealed the dynamics and fluctuating nature of
343 the ocean, even at short time scale [69]. The heterogeneity in mesopelagic layers given by
344 deep currents, and by the impact of the surface production, together with the low mixing levels,
345 may favor a diversification in the mesopelagic community living in different water masses,
346 leading to species adaptation-acclimation. The *Tara* Oceans expedition route included

347 samples from common or distinct water masses defined by temperature/salinity profiles - T/S,
348 comprising regionally connected or unconnected stations. We identified nine different water
349 masses in the mesopelagic sampled locations (Figure 3). We could confirm significant
350 differences in mesopelagic communities sampled in different water masses based on the
351 PERMANOVA test (Table 2). This result indicates that the oceanic patchiness created by
352 distinct water masses can favor beta-diversity diversification, indicating it to be a critical
353 component for mesopelagic community variation for all the assemblages studied (phages,
354 giruses, prokaryotes, and pico-eukaryotes). Thus, we hypothesize that this result may be
355 explained by two non-exclusive causes related to water masses: (i) past common origin among
356 water masses that have drifted or (ii) constant connectivity by ocean circulation between
357 sampled sites belonging to the same water mass.

358 Following the biotic control suggested by the deterministic model previously identified, we
359 addressed another lingering question, resolving planktonic community signatures of Oxic MES
360 and OMZ regions for epipelagic communities. For this, we classified OTUs based on their
361 relative abundance into three eco-regions: 1) EPI, 2) Oxic MES, and 3) OMZ. OTUs were
362 classified as Core MES when commonly present in Oxic MES and OMZ samples. The taxa
363 that were either equally abundant in all three eco-regions or not statistically confirmed to a
364 single eco-region were classified as ubiquitous (Supplementary Figure 3). Using this approach,
365 we could identify ubiquitous taxa, that are likely to thrive in a wide range of environmental
366 conditions, or that may be detected in mesopelagic samples due to the simple vertical
367 movement of sinking particles. This classification should help avoid putative biases inherent of
368 the metabarcoding methodology.

369 More specifically, we were able to identify Oxic MES and OMZ signatures mainly at the infra-
370 taxonomic level (OTU-species) for all biotic groups investigated (Figure 4, Supplementary
371 Figure 3, 4, 5, 6, 7). This reflected the wide ecological niche occupied by the different species
372 at a higher taxonomic level (i.e. family). At the species level, we observed large taxonomic
373 plasticity of OTUs that occurred equally in both Oxic MES and OMZ samples, called Core MES,

374 principally for girus assemblages. However, most OTUs are not yet classified at the infra-
375 taxonomic level (data not shown). This observation reflects the knowledge gap about the
376 biodiversity and functional plasticity of species thriving in this ecosystem.

377 The great majority of phage taxa occurred at similar abundance in all regions (ubiquitous)
378 (Supplementary Figure 3, 4). Surprisingly, this ubiquity is vertically linked at each independent
379 station (Supplementary Figure 4), supporting the seed-bank hypothesis raised by Brum et al.
380 [60], and the correlation to the sinking particles observed here. We observed taxa specific to
381 the mesopelagic layer in second place, mostly related to the OMZ eco-region (Figure 4a,
382 Supplementary Figures 3, 8). This mesopelagic specificity agrees with the sharp increase in
383 marine phage microdiversity following depth, as previously shown by Gregory et al. [8]. Our
384 results emphasize that one cause for phage stratification in the water column might be the
385 adaptation to the mesopelagic environment. Two hypotheses arise here, 1) the environment
386 acts as a strong driver, directly selecting phages independently of their hosts, and 2) there is
387 higher phage-host specificity in the mesopelagic layer, promoting phage selection. Following
388 the first hypothesis, we can posit that the environment can directly impact phage assemblage
389 composition. The direct contact with the environment of free phage particles (released from
390 their hosts) may reduce infectivity, degrade, or remove virus particles, and adversely affect
391 adsorption to the host [70]. This direct environmental effect over marine phages was reported
392 for different ionic gradients [71], daylight conditions, and temperature [72]. However, the
393 enrichment of prokaryotic OTUs specific to mesopelagic regions (Supplementary Figures 3,
394 8), especially in OMZs, does not exclude the phage-host indirect selection relationship.

395 We found fewer but abundant mesopelagic-specific girus OTUs in both Oxic and OMZ eco-
396 regions, indicating that giruses can be less diverse in the mesopelagic layer (Figure 4b,
397 Supplementary Figures 5, 3, 8). Also, giruses can encode genes such as transporters for
398 ammonium, magnesium, and phosphate that are important in marine oligotrophic areas [73].
399 This characteristic can improve the host's fitness in the short-term but ultimately favor girus
400 fecundity and endurance. This property is named NCLDV-mediated host reprogramming [73].

401 The great majority of mesopelagic girus OTUs were assigned to the Core MES group,
402 indicating that these entities may infect a wide range of hosts adapted to diverse environmental
403 conditions.

404 We could better distinguish the prokaryotic mesopelagic signatures between Oxic MES and
405 OMZ, confirming the influence of oxygen reported here and in previous studies [18, 19, 20]
406 (Figure 4c, Supplementary Figures 6, 3, 8). Among the planktonic microorganisms,
407 prokaryotes have been, so far, the most investigated group in OMZ regions, especially in the
408 Pacific Ocean [60, 19, 18]. We observed similar occurrence and abundance for the OMZ
409 signature taxa in the Indian Ocean stations (IO - 037, 038, 039) and in stations PO - 100, 137,
410 and 138 from the Pacific Ocean (Figure 4c). These Pacific stations are located in the open
411 ocean (PO - 137 and 138 located in the Equatorial upwelling zone and station PO - 100 in the
412 South Pacific Subtropical Gyre). They present a strong upwelling signature, disclosing an
413 intense decrease in oxygen concentration almost reaching shallow waters. Likewise, the
414 sampling stations in the Indian Ocean are located in well-stratified waters, markedly
415 characterized by the abrupt decrease of oxygen concentration below the thermohaline at 100-
416 120m depth. In all these stations, the oxygen concentration ranges from 0.83 to 3 $\mu\text{mol/kg}$,
417 characterizing functionally anoxic waters since this oxygen level cannot sustain aerobic
418 metabolisms [74]. The other OMZ stations in the Pacific Ocean (PO - 102, 109, 110, 111) are
419 located in coastal areas. Although they are also under the influence of upwellings presenting
420 an oxygen depletion, the oxygen level does not achieve anoxic conditions, and thus are
421 classified as suboxic waters. This microoxic environment is enough to completely change the
422 microbial metabolism delineating the community composition in those sites. Also, differences
423 in offshore and coastal upwelling formation, for instance, or the influence of river runoffs,
424 transporting anthropogenic nutrient enrichment from the continent to coastal zones [74], could
425 be crucial to support the differences in OMZ communities we observed.

426 The same combination of OMZ anoxic and suboxic samples was observed for the pico-
427 eukaryotic groups MALV-II and Diplonemida, suggesting these OTUs as the true OMZ

428 eukaryotic signatures (Figure 4d). Some OTUs of these groups exhibited similar occurrences
429 in the anoxic Indian and Pacific Oceans but not in suboxic samples from the Pacific Ocean.
430 However, we observed a lower number of pico-eukaryotic taxa in the OMZ eco-region, the
431 prevailing OTUs being specific to Oxidic MES locations in most cases. Although the CCA
432 analysis did not disentangle the oxygen from the other variables to explain the pico-eukaryote
433 assemblage variations, here we can verify that OMZ conditions do act negatively on selection
434 of pico-eukaryotes in marine environments.

435 Another step to better understand mesopelagic community dynamics is to dissect the
436 ecological relationships among species that thrive in this layer. Co-occurrence networks can
437 indicate how the environment may structure the community acting as a filter for resident
438 species [76]. They can also give us glimpses of organisms' ecological interactions based on
439 species connectivity [76, 77]. Combining the viroplasm, prokaryote, and pico-eukaryote data, we
440 inferred a network containing 6,154 nodes and 12,935 edges (Figure 5a, Table 3). Due to the
441 lower number of stations sampled for phages, we excluded this group from the analysis. We
442 found mainly positive relationships (94%), suggesting a predominance of putative biotic
443 interactions (e.g. competition, symbiosis) rather than taxa avoidance or exclusion. This
444 dominance of positive relations was also reported for epipelagic plankton communities [27,78].
445 The global network had a modularity value greater than 0.4 (Table 3), indicating that the
446 network has a modular structure [74]. Using a module detection algorithm, we were able to
447 identify 36 distinct modules in the global network. Three of them were mainly composed of
448 OTUs significantly enriched in mesopelagic OTUs (Oxidic MES enriched module 1 and OMZ
449 enriched modules 4 and 17; Figure 5). Together, these three modules cover almost the total
450 richness found in the mesopelagic zone (Figure 5b), and present similar values for the average
451 degree, clustering coefficient, and average path length (Table 3). These parameters indicate
452 a network complexity [76], hinting at distinct ecological niches within the mesopelagic at the
453 level of investigated organismal fractions.

454 In more detail, the OMZ signature modules were composed of a few connected nodes (323
455 and 175 nodes for modules 4 and 17, respectively), potentially indicating two distinct OMZ
456 community niches. The Oxic MES module 1 counted more nodes composing the network
457 associations (731 nodes), and both modules presented a variation in taxonomic composition
458 and proportions. OMZ module 4 contained mainly prokaryotic (23%) and girus (55%) OTUs
459 (Figure 5c, d). Among the prokaryotes, we detected taxa previously determined as OMZ
460 signatures (Nitrospinae, Marinimicrobia SAR 406, and Planctomycetes). Module 17 is mainly
461 comprised of pico-eukaryotes (82%), notably MALV-II (14%) and Diplonemida (17%)
462 previously indicated as OMZ signatures. Module 1 is taxonomically more diverse but is mainly
463 comprised of giruses and pico-eukaryotes. These groups accounted for 36% and 55%,
464 respectively, of OTUs in this module (Figure 5c, d). Giruses contributed to 598 associations
465 (edges) in mesopelagic module 1, of which 177 occurred between giruses and pico-
466 eukaryotes. Giruses from the *Mimiviridae* family are the most numerous taxa in all three
467 mesopelagic modules. *Mimiviridae* is a very abundant family in the ocean, present in various
468 size ranges from piconanoplankton (0.8-5 μm) up to mesoplankton (180- 2,000 μm) [9, 79].
469 This observation supports our findings that giruses are a prosperous group in mesopelagic
470 waters, undertaking different strategies to endure in such environmental conditions. In all three
471 modules, we observed the presence of Foraminifera, of which some species can use nitrate
472 over oxygen as an electron acceptor, favoring their survival in OMZ regions [80].

473 Sugihara [57] affirms that a hierarchical niche structure can explain the lognormal abundance
474 pattern in communities, and this assertion is valid for small assemblages and for large
475 ensembles. Consistently, the observation of three distinct community modules supports the
476 evidence of the lognormal ecological model empirically defined locally for each assemblage.
477 Our results converge and suggest that the mesopelagic presents at least three well-defined
478 ecological niches (Oxic MES, OMZ-4 and OMZ-17), with established conditions and resources
479 (abiotic and biotic) that allow the survival of a given species in these environments. Differences
480 between OMZ and Oxic MES networks suggest a potential loss of connections and interactions

481 among mesopelagic community members, directly affecting ecosystem stability due to habitat
482 change.

483

484 **Conclusions**

485

486 In this study, we explored mesopelagic pico-plankton ecological structuring, and concluded
487 that this component of oceanic plankton is heterogeneous regarding its environmental
488 parameters. The ecological parameters drive mesopelagic community assemblages, as they
489 fit niche ecological models. Although we could not identify a single driver of community
490 composition (such as temperature for sunlit ocean) for all organisms, we could pinpoint the
491 relevance of oxygen for phages and prokaryotic fractions and the relation of the former with
492 particle flux. Also, we show that water masses defined by their T/S profiles can explain the
493 differences in the observed pico-plankton structure, pointing to the role of a set of
494 environmental parameters rather than single drivers for community composition.

495 By establishing Eco-regions (Epipelagic, Oxic MES, and OMZ), we were able to discriminate
496 specific OTUs for all fractions studied. While we recovered known markers for Oxic MES and
497 OMZ regions at high taxonomic levels, we also found that most of these OTU signatures are
498 observed at low taxonomic levels, which sometimes cannot be resolved using current
499 databases. Crossing these specific OTUs with co-occurrence networks, we identified three
500 niches with biotic and abiotic conditions that characterize mesopelagic waters.

501 The limiting access to data is usually the bottleneck for knowledge about mesopelagic
502 dynamics. Our study benefits from a more significant number of organism samples and distinct
503 oceanic provinces than previous ones, allowing us to combine data to derive an expanded
504 vision of mesopelagic composition. Our results emphasize the need for better understanding
505 of mesopelagic life, in particular by improving our knowledge about oxic and oxygen-depleted
506 mesopelagic-dwelling communities, especially as climate change can be expected to expand
507 marine OMZs shortly.

508 **Acknowledgments**

509
510 This study is part of the “Ocean Plankton, Climate and Development” project conducted by the
511 Tara Ocean Foundation with the support of the French Facility for the Global Environment
512 (FFEM). Rigonato J., Budinich M., Murillo A.A., Brandão M.C., Pierella Karlusich J.J. and
513 Soviadan Y.D. received financial support from FFEM to execute the project. Brandão, M.C.
514 also received financial support from Coordination for the Improvement of Higher Education
515 Personnel of Brazil (CAPES 99999.000487/2016-03). Tara Oceans (which includes both the
516 Tara Oceans and Tara Oceans Polar Circle expeditions) would not exist without the leadership
517 of the *Tara* Ocean Foundation and the continuous support of 23 institutes
518 (<https://fondationtaraocean.org>). We further thank the commitment of the following sponsors:
519 CNRS (in particular Groupement de Recherche GDR3280 and the Research Federation for
520 the study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Oceans-GOSEE),
521 European Molecular Biology Laboratory (EMBL), Genoscope/CEA, The French Ministry of
522 Research, and the French Government ‘Investissements d’Avenir’ programs OCEANOMICS
523 (ANR- 11-BTBR-0008), FRANCE GENOMIQUE (ANR-10-INBS-09-08), MEMO LIFE (ANR-
524 10-LABX-54), and PSL* Research University (ANR-11-IDEX-0001-02). We also thank the
525 support and commitment of agnès b. and Etienne Bourgois, the Prince Albert II de Monaco
526 Foundation, the Veolia Foundation, Region Bretagne, Lorient Agglomeration, Serge Ferrari,
527 World Courier. The global sampling effort was enabled by countless scientists and crew who
528 sampled aboard the schooner *Tara* from 2009-2013, and we thank MERCATOR-CORIOLIS
529 and ACRIST for providing daily satellite data during the expeditions. We are also grateful to
530 the countries who graciously granted sampling permission. The authors declare that all data
531 reported herein are fully and freely available from the date of publication, with no restrictions,
532 and that all of the analyses, publications, and ownership of data are free from legal
533 entanglement or restriction by the various nations whose waters the Tara Oceans expeditions
534 sampled in. This article is contribution number XX of Tara Oceans.

535

536

537 Tara Oceans Coordinators (alphabetical order) Silvia G. Acinas, Marcel Babin, Peer Bork,
538 Emmanuel Boss, Chris Bowler, Guy Cochrane, Colomban de Vargas, Michael Follows, Gabriel
539 Gorsky, Nigel Grimsley, Lionel Guidi, Pascal Hingamp, Daniele Iudicone, Olivier Jaillon,
540 Stefanie Kandels, Lee Karp-Boss, Eric Karsenti, Fabrice Not, Hiroyuki Ogata, Stéphane
541 Pesant, Nicole Poulton, Jeroen Raes, Christian Sardet, Sabrina Speich, Lars Stemmann,
542 Matthew B. Sullivan, Shinichi Sunagawa and Patrick Wincker.

543 **Conflict of Interest**

544

545 The authors declare that they have no conflict of interest.

546

547 **References**

548

- 549 1. Proud R, Cox MJ, and Brierley AS. Biogeography of the Global Ocean's Mesopelagic
550 Zone. en. *Current Biology* 2017 Jan; 27:113–9. doi: [10.1016/j.cub.2016.11.003](https://doi.org/10.1016/j.cub.2016.11.003)
- 551 2. Robinson C, Steinberg DK, Anderson TR, Arístegui J, Carlson CA, Frost JR, et al.
552 Mesopelagic zone ecology and biogeochemistry – a synthesis. en. *Deep Sea Research*
553 *Part II: Topical Studies in Oceanography* 2010 Aug; 57:1504–18. doi:
554 [10.1016/j.dsr2.2010.02.018](https://doi.org/10.1016/j.dsr2.2010.02.018)
- 555 3. Mestre M, Ruiz-González C, Logares R, Duarte CM, Gasol JM, and Sala MM. Sinking
556 particles promote vertical connectivity in the ocean microbiome. en. *Proceedings of the*
557 *National Academy of Sciences* 2018 Jul; 115:E6799–E6807. doi:
558 [10.1073/pnas.1802470115](https://doi.org/10.1073/pnas.1802470115)
- 559 4. Whitman WB, Coleman DC, and Wiebe WJ. Prokaryotes: The unseen majority. en.
560 *Proceedings of the National Academy of Sciences* 1998 Jun; 95:6578–83. doi:
561 [10.1073/pnas.95.12.6578](https://doi.org/10.1073/pnas.95.12.6578)
- 562 5. Hidalgo M and Browman HI. Developing the knowledge base needed to sustainably
563 manage mesopelagic resources. en. *ICES Journal of Marine Science* 2019 May;
564 76:609–15. doi: [10.1093/icesjms/fsz067](https://doi.org/10.1093/icesjms/fsz067)
- 565 6. St. John MA, Borja A, Chust G, Heath M, Grigorov I, Mariani P, et al. A Dark Hole in Our
566 Under- standing of Marine Ecosystems and Their Services: Perspectives from the
567 Mesopelagic Community. en. *Frontiers in Marine Science* 2016 Mar; 3. doi:
568 [10.3389/fmars.2016.00031](https://doi.org/10.3389/fmars.2016.00031)
- 569 7. Glover AG, Wiklund H, Chen C, and Dahlgren TG. Managing a sustainable deep-sea
570 'blue economy' requires knowledge of what actually lives there. *eLife* 2018 Nov; 7. Ed.
571 by Groll H and Rodgers PA:e41319. doi: [10.7554/eLife.41319](https://doi.org/10.7554/eLife.41319)
- 572 8. GregoryAC, ZayedAA, Conceição-NetoN, TempertonB, BolducB, AlbertiA,et al.
573 *MarineDNA Viral Macro- and Microdiversity from Pole to Pole*. en. *Cell* 2019 Apr
574 :S0092867419303411. doi: [10.1016/j.cell.2019.03.040](https://doi.org/10.1016/j.cell.2019.03.040)

- 575 9. Endo H, Blanc-Mathieu R, Li Y, Salazar G, Henry N, Labadie K, et al. Biogeography of
576 marine giant viruses reveals their interplay with eukaryotes and ecological functions. en.
577 Nature Ecology & Evolution 2020 Sep. doi: [10.1038/s41559-020-01288-w](https://doi.org/10.1038/s41559-020-01288-w)
- 578 10. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, et al. Structure
579 and function of the global ocean microbiome. en. Science 2015 May; 348:1261359–
580 1261359. doi: [10.1126/science.1261359](https://doi.org/10.1126/science.1261359)
- 581 11. Salazar G, Paoli L, Alberti A, Huerta-Cepas J, Ruscheweyh HJ, Cuenca M, et al. Gene
582 Expression Changes and Community Turnover Differentially Shape the Global Ocean
583 Metatranscriptome. en. Cell 2019 Nov; 179:1068–1083.e21. doi:
584 [10.1016/j.cell.2019.10.014](https://doi.org/10.1016/j.cell.2019.10.014)
- 585 12. Giner CR, Pernice MC, Balagué V, Duarte CM, Gasol JM, Logares R, et al. Marked
586 changes in diversity and relative activity of picoeukaryotes with depth in the world ocean.
587 en. The ISME Journal 2020 Feb; 14:437–49. doi: [10.1038/s41396-019-0506-9](https://doi.org/10.1038/s41396-019-0506-9)
- 588 13. Ibarbalz FM, Henry N, Brandão MC, Martini S, Busseni G, Byrne H, et al. Global Trends
589 in Marine Plankton Diversity across Kingdoms of Life. en. Cell 2019 Nov; 179:1084–
590 1097.e21. doi: [10.1016/j.cell.2019.10.008](https://doi.org/10.1016/j.cell.2019.10.008)
- 591 14. Capone DG and Hutchins DA. Microbial biogeochemistry of coastal upwelling regimes
592 in a changing ocean. en. Nature Geoscience 2013 Sep; 6:711–7. doi:
593 [10.1038/ngeo1916](https://doi.org/10.1038/ngeo1916)
- 594 15. Breitburg D, Levin LA, Oschlies A, Grégoire M, Chavez FP, Conley DJ, et al. Declining
595 oxygen in the global ocean and coastal waters. en. 2018 :13
- 596 16. Stevens H and Ulloa O. Bacterial diversity in the oxygen minimum zone of the eastern
597 tropical South Pacific. en. Environmental Microbiology 2008 May; 10:1244–59. doi:
598 [10.1111/j.1462-2920.2007.01539.x](https://doi.org/10.1111/j.1462-2920.2007.01539.x)
- 599 17. Stewart FJ, Ulloa O, and DeLong EF. Microbial metatranscriptomics in a permanent
600 marine oxygen minimum zone: OMZ community gene expression. en. Environmental
601 Microbiology 2012 Jan; 14:23–40. doi: [10.1111/j.1462-2920.2010.02400.x](https://doi.org/10.1111/j.1462-2920.2010.02400.x)
- 602 18. Ulloa O, Wright JJ, Belmar L, and Hallam SJ. Pelagic Oxygen Minimum Zone Microbial
603 Communities. en. The Prokaryotes. Ed. by Rosenberg E, DeLong EF, Lory S,
604 Stackebrandt E, and Thompson F. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013
605 :113–22. doi: [10.1007/978-3-642-30123-0_45](https://doi.org/10.1007/978-3-642-30123-0_45)
- 606 19. Ulloa O and Pantoja S. The oxygen minimum zone of the eastern South Pacific. en.
607 Deep Sea Research Part II: Topical Studies in Oceanography 2009 Jul; 56:987–91. doi:
608 [10.1016/j.dsr2.2008.12.004](https://doi.org/10.1016/j.dsr2.2008.12.004)
- 609 20. Wright JJ, Konwar KM, and Hallam SJ. Microbial ecology of expanding oxygen minimum
610 zones. en. Nature Reviews Microbiology 2012 Jun; 10:381–94. doi:
611 [10.1038/nrmicro2778](https://doi.org/10.1038/nrmicro2778)
- 612 21. Duret MT, Pachiadaki MG, Stewart FJ, Sarode N, Christaki U, Monchy S, et al. Size-
613 fractionated diversity of eukaryotic microbial communities in the Eastern Tropical North
614 Pacific oxygen minimum zone. en. FEMS Microbiology Ecology 2015 May; 91. doi:
615 [10.1093/femsec/fiv037](https://doi.org/10.1093/femsec/fiv037)
- 616 22. Orsi W, Song YC, Hallam S, and Edgcomb V. Effect of oxygen minimum zone formation
617 on communities of marine protists. en. The ISME Journal 2012 Aug; 6:1586–601. doi:
618 [10.1038/ismej.2012.7](https://doi.org/10.1038/ismej.2012.7)
- 619 23. Parris DJ, Ganesh S, Edgcomb VP, DeLong EF, and Stewart FJ. Microbial eukaryote
620 diversity in the marine oxygen minimum zone off northern Chile. en. Frontiers in
621 Microbiology 2014 Oct; 5. doi: [10.3389/fmicb.2014.00543](https://doi.org/10.3389/fmicb.2014.00543)

- 622 24. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, et al. The
623 Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern
624 Tropical Pacific. en. PLOS Biology 2007 Mar; 5:e77. doi: [10.1371/journal.pbio.0050077](https://doi.org/10.1371/journal.pbio.0050077)
- 625 25. Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, Raes J, et al. A Holistic
626 Approach to Marine Eco-Systems Biology. en. PLoS Biology 2011 Oct; 9:e1001177. doi:
627 [10.1371/journal.pbio.1001177](https://doi.org/10.1371/journal.pbio.1001177)
- 628 26. Pernice MC, Forn I, Gomes A, Lara E, Alonso-S´aez L, Arrieta JM, et al. Global
629 abundance of planktonic heterotrophic protists in the deep ocean. en. The ISME Journal
630 2015 Mar; 9:782–92. doi: [10.1038/ismej.2014.168](https://doi.org/10.1038/ismej.2014.168)
- 631 27. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, et al. Determinants of
632 community structure in the global plankton interactome. en. Science 2015 May;
633 348:1262073–1262073. doi: [10.1126/science.1262073](https://doi.org/10.1126/science.1262073)
- 634 28. Louca S, Hawley AK, Katsev S, Torres-Beltran M, Bhatia MP, Kheirandish S, et al.
635 Integrating bio- geochemistry with multiomic sequence information in a model oxygen
636 minimum zone. en. Proceedings of the National Academy of Sciences 2016 Oct;
637 113:E5925–E5933. doi: [10.1073/pnas.1602897113](https://doi.org/10.1073/pnas.1602897113)
- 638 29. Ardyna M, Tara Oceans Consortium C, and Tara Oceans Expedition P. Environmental
639 context of all station from the Tara Oceans Expedition (2009-2013), about the annual
640 cycle of key parameters estimated daily from remote sensing products at a spatial
641 resolution of 9km. en. type: dataset. 2017. doi: [10.1594/PANGAEA.883613](https://doi.org/10.1594/PANGAEA.883613)
- 642 30. Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Acinas, S. G., Babin, M., ... de
643 Vargas, C. (2020). Tara Oceans: towards global ocean ecosystems biology. *Nature*
644 *Reviews Microbiology*, 18(8), 428–445. doi: [10.1038/s41579-020-0364-5](https://doi.org/10.1038/s41579-020-0364-5)
- 645 31. Guidi L, Morin P, Coppola L, Tremblay JE, Pesant S, Tara Oceans Consortium C, et al.
646 Environmental context of all samples from the Tara Oceans Expedition (2009-2013),
647 about nutrients in the targeted environmental feature. en. type: dataset. 2017. doi:
648 [10.1594/PANGAEA.875575](https://doi.org/10.1594/PANGAEA.875575)
- 649 32. Guidi L, Picheral M, Pesant S, Tara Oceans Consortium C, and Tara Oceans Expedition
650 P. Environ- mental context of all samples from the Tara Oceans Expedition (2009-2013),
651 about sensor data in the targeted environmental feature. en. type: dataset. 2017. doi:
652 [10.1594/PANGAEA.875576](https://doi.org/10.1594/PANGAEA.875576)
- 653 33. Guidi L, Ras J, Claustre H, Pesant S, Tara Oceans Consortium C, and Tara Oceans
654 Expedition P. Environmental context of all samples from the Tara Oceans Expedition
655 (2009-2013), about pigment concentrations (HPLC) in the targeted environmental
656 feature. en. type: dataset. 2017. doi: [10.1594/PANGAEA.875569](https://doi.org/10.1594/PANGAEA.875569)
- 657 34. Pesant S, Tara Oceans Consortium C, and Tara Oceans Expedition P. Methodology
658 used on board to prepare samples from the Tara Oceans Expedition (2009-2013). en.
659 type: dataset. 2017. doi: [10.1594/PANGAEA.875580](https://doi.org/10.1594/PANGAEA.875580)
- 660 35. Speich S, Chaffron S, Ardyna M, Pesant S, Tara Oceans Consortium C, and Tara
661 Oceans Expedition P. Environmental context of all samples from the Tara Oceans
662 Expedition (2009-2013), about the water column features at the sampling location. en.
663 type: dataset. 2017. doi: [10.1594/PANGAEA.875579](https://doi.org/10.1594/PANGAEA.875579)
- 664 36. Picheral M, Guidi L, Stemmann L, Karl DM, Iddaoud G, and Gorsky G. The Underwater
665 Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle
666 size spectra and zooplankton. *Limnology and Oceanography: Methods* 2010; 8:462–73
- 667 37. Picheral M, Searson S, Taillandier V, Bricaud A, Boss E, Ras J, et al. Vertical profiles of
668 environmental parameters measured on discrete water samples collected with Niskin
669 bottles during the Tara Oceans expedition 2009-2013. en. type: dataset. 2014. doi:
670 [10.1594/PANGAEA.836319](https://doi.org/10.1594/PANGAEA.836319)

- 671 38. Schlitzer R. Interactive analysis and visualization of geoscience data with Ocean Data
672 View. en. *Computers & Geosciences* 2002 Dec; 28:1211–8. doi: [10.1016/S0098-
673 3004\(02\)00040-7](https://doi.org/10.1016/S0098-3004(02)00040-7)
- 674 39. Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., ...
675 Searson, S. (2015). Open science resources for the discovery and analysis of Tara
676 Oceans data. *Scientific Data*, 2(Lmd), 1–16. doi: [10.1038/sdata.2015.23](https://doi.org/10.1038/sdata.2015.23)
- 677 40. Alberti A, Poulain J, Engelen S, Labadie K, Romac S, Ferrera I, et al. Viral to metazoan
678 marine plankton nucleotide sequences from the Tara Oceans expedition. en. *Scientific
679 Data* 2017 Aug; 4:170093. doi: [10.1038/sdata.2017.93](https://doi.org/10.1038/sdata.2017.93)
- 680 41. Vargas C de, Audic S, Henry N, Decelle J, Mahe F, Logares R, et al. Eukaryotic plankton
681 diversity in the sunlit ocean. en. *Science* 2015 May; 348:1261605–1261605. doi:
682 [10.1126/science.1261605](https://doi.org/10.1126/science.1261605)
- 683 42. Jari Oksanen FGB, Friendly M, Kindt R, Legendre P, McGlenn D, Minchin PR, et al.
684 vegan: Community Ecology Package. R package version 2.5-6. 2019
- 685 43. Wilson JB. Methods for fitting dominance/diversity curves. en. *Journal of Vegetation
686 Science* 1991 Feb; 2:35–46. doi: [10.2307/3235896](https://doi.org/10.2307/3235896)
- 687 44. Matthews TJ and Whittaker RJ. Fitting and comparing competing models of the species
688 abundance distribution: assessment and prospect. en. *Frontiers of Biogeography* 2014
689 Jun; 6. doi: [10.21425/F56220607](https://doi.org/10.21425/F56220607)
- 690 45. Chave J. Neutral theory and community ecology: Neutral theory and community
691 ecology. en. *Ecology Letters* 2004 Feb; 7:241–53. doi: [10.1111/j.1461-
692 0248.2003.00566.x](https://doi.org/10.1111/j.1461-0248.2003.00566.x)
- 693 46. McGill BJ, Maurer BA, and Weiser MD. EMPIRICAL EVALUATION OF NEUTRAL
694 THEORY. en. *Ecology* 2006 Jun; 87:1411–23. doi: [10.1890/0012-
695 9658\(2006\)87\[1411:EEONT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1411:EEONT]2.0.CO;2)
- 696 47. Paulo I. Prado MDM and Chalom A. sads:Maximum Likelihood Models for Species
697 Abundance Distributions. R package version 0.4.2. 2018
- 698 48. Magurran AE and McGill BJ. Biological Diversity frontiers in measurement and
699 assessment. en. 24. New York: Oxford University press, 2011
- 700 49. Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC
701 and BIC in model selection. *Sociological Methods and Research*, 33(2), 261–304. doi:
702 [10.1177/0049124104268644](https://doi.org/10.1177/0049124104268644)
- 703 50. R Core Team. R: A Language and Environment for Statistical Computing. ISBN 3-
704 900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013
- 705 51. Dinno A. Dunn's Test of Multiple Comparisons Using Rank Sums. R package version
706 1.3.5. 2017
- 707 52. Tackmann J, Rodrigues JFM, and Mering C von. Rapid inference of direct interactions
708 in large-scale ecological networks from heterogeneous microbial sequencing data. *Cell
709 Systems* 2019. doi: [10.1016/j.cels.2019.08.002](https://doi.org/10.1016/j.cels.2019.08.002)
- 710 53. Clauset, Newman, and Moore, "Finding Community Structure in Very Large Networks."
711 *Physical Review E*. 2004; 70:066111 doi: [10.1103/PhysRevE.70.066111](https://doi.org/10.1103/PhysRevE.70.066111)
- 712 54. Vellend M. Conceptual Synthesis in Community Ecology. en. *The Quarterly Review of
713 Biology* 2010 Jun; 85:183–206. doi: [10.1086/652373](https://doi.org/10.1086/652373)
- 714 55. Filho RC, Martins FR, and Gneri MA. Fitting abundance distribution models in tropical
715 arboreal communities of SE Brazil. en. *Community Ecology* 2002; 3:169–80
- 716 56. Magurran AE and Henderson PA. Explaining the excess of rare species in natural
717 species abundance distributions. en. 2003; 422:3

- 718 57. Sugihara G. Minimal Community Structure: An Explanation of Species Abundance
719 Patterns. en. *The American Naturalist* 1980; 116:770–87
- 720 58. Bulmer MG. On Fitting the Poisson Lognormal Distribution to Species-Abundance Data.
721 en. *Biometrics* 1974 Mar; 30:101. doi: [10.2307/2529621](https://doi.org/10.2307/2529621)
- 722 59. Preston FW. The Commonness, And Rarity, of Species. en. *Ecology* 1948 Jul; 29:254–
723 83. doi: [10.2307/1930989](https://doi.org/10.2307/1930989)
- 724 60. Ser-Giacomi, E., Zinger, L., Malviya, S., De Vargas, C., Karsenti, E., Bowler, C., & De
725 Monte, S. (2018). Ubiquitous abundance distribution of non-dominant plankton across
726 the global ocean. *Nature Ecology and Evolution*, 2(8), 1243–1249. doi: [10.1038/s41559-](https://doi.org/10.1038/s41559-018-0587-2)
727 [018-0587-2](https://doi.org/10.1038/s41559-018-0587-2)
- 728 61. Brum JR, Ignacio-Espinoza JC, Roux S, Doucier G, Acinas SG, Alberti A, et al. Patterns
729 and ecological drivers of ocean viral communities. en. *Science* 2015 May;
730 348:1261498–1261498. doi: [10.1126/science.1261498](https://doi.org/10.1126/science.1261498)
- 731 62. Ghiglione JF, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K, et al. Pole-
732 to-pole bio- geography of surface and deep marine bacterial communities. en.
733 *Proceedings of the National Academy of Sciences* 2012 Oct; 109:17633–8. doi:
734 [10.1073/pnas.1208160109](https://doi.org/10.1073/pnas.1208160109)
- 735 63. Ulloa O, Canfield DE, DeLong EF, Letelier RM, and Stewart FJ. Microbial oceanography
736 of anoxic oxygen minimum zones. en. *Proceedings of the National Academy of Sciences*
737 2012 Oct; 109:15996– 6003. doi: [10.1073/pnas.1205009109](https://doi.org/10.1073/pnas.1205009109)
- 738 64. Aldunate M, De la Iglesia R, Bertagnolli AD, and Ulloa O. Oxygen modulates bacterial
739 community composition in the coastal upwelling waters off central Chile. en. *Deep Sea*
740 *Research Part II: Topical Studies in Oceanography* 2018 Oct; 156:68–79. doi:
741 [10.1016/j.dsr2.2018.02.001](https://doi.org/10.1016/j.dsr2.2018.02.001)
- 742 65. De la Iglesia R, Echenique-Subiabre I, Rodríguez-Marconi S, Espinoza JP, Dassow P
743 von, Ulloa O, et al. Distinct oxygen environments shape picoeukaryote assemblages
744 thriving oxygen minimum zone waters off central Chile. en. *Journal of Plankton*
745 *Research* 2020 Sep; 42. Ed. by Dolan J:514–29. doi: [10.1093/plankt/fbaa036](https://doi.org/10.1093/plankt/fbaa036)
- 746 66. Schnetzer A, Moorthi SD, Countway PD, Gast RJ, Gilg IC, and Caron DA. Depth
747 matters: Microbial eukaryote diversity and community structure in the eastern North
748 Pacific revealed through environmen- tal gene libraries. en. *Deep Sea Research Part I:*
749 *Oceanographic Research Papers* 2011 Jan; 58:16–26. doi: [10.1016/j.dsr.2010.10.003](https://doi.org/10.1016/j.dsr.2010.10.003)
- 750 67. Guidi L, Chaffron S, Bittner L, Eveillard D, Larhlimi A, Roux S, et al. Plankton networks
751 driving carbon export in the oligotrophic ocean. en. *Nature* 2016 Feb; 532:465–70. doi:
752 [10.1038/nature16942](https://doi.org/10.1038/nature16942)
- 753 68. Parada V, Sintés E, Aken HM van, Weinbauer MG, and Herndl GJ. Viral Abundance,
754 Decay, and Diversity in the Meso- and Bathypelagic Waters of the North Atlantic. en.
755 *Applied and Environmental Microbiology* 2007 Jul; 73:4429–38. doi:
756 [10.1128/AEM.00029-07](https://doi.org/10.1128/AEM.00029-07)
- 757 69. Bates AE, Helmuth B, Burrows MT, Duncan MI, Garrabou J, Guy-Haim T, et al.
758 *Biologists ignore ocean weather at their peril.* en. *Nature* 2018 Aug; 560:299–301
- 759 70. Mojica KDA and Brussaard CPD. Factors affecting virus dynamics and microbial host–
760 virus interactions in marine environments. en. *FEMS Microbiol Ecol* 2014 :21
- 761 71. Kukkaro P and Bamford DH. Virus–host interactions in environments with a wide range
762 of ionic strengths. en. *Environmental Microbiology Reports* 2009 :7
- 763 72. Bettarel Y, Bouvier T, and Bouvy M. Viral persistence in water as evaluated from a
764 tropical/temperate cross-incubation. en. *Journal of Plankton Research* 2009 Aug;
765 31:909–16. doi: [10.1093/plankt/fbp041](https://doi.org/10.1093/plankt/fbp041)

- 766 73. Schulz F. Giant virus diversity and host interactions through global metagenomics. en.
767 Nature 2020; 578:432–6. doi: <https://doi.org/10.1038/s41586-020-1957-x>
- 768 74. Thamdrup B, Dalsgaard T, and Revsbech NP. Widespread functional anoxia in the
769 oxygen minimum zone of the Eastern South Pacific. en. Deep Sea Research Part I:
770 Oceanographic Research Papers 2012 Jul; 65:36–45. doi: [10.1016/j.dsr.2012.03.001](https://doi.org/10.1016/j.dsr.2012.03.001)
- 771 75. Helly JJ and Levin LA. Global distribution of naturally occurring marine hypoxia on
772 continental margins. en. Deep Sea Research Part I: Oceanographic Research Papers
773 2004 Sep; 51:1159–68. doi: [10.1016/j.dsr.2004.03.009](https://doi.org/10.1016/j.dsr.2004.03.009)
- 774 76. Berry D and Widder S. Deciphering microbial interactions and detecting keystone
775 species with co- occurrence networks. en. Frontiers in Microbiology 2014 May; 5. doi:
776 [10.3389/fmicb.2014.00219](https://doi.org/10.3389/fmicb.2014.00219)
- 777 77. Zhu W, Qin C, Ma H, Xi S, Zuo T, Pan W, et al. Response of protist community dynamics
778 and co- occurrence patterns to the construction of artificial reefs: A case study in Daya
779 Bay, China. en. Science of The Total Environment 2020 Nov; 742:140575. doi:
780 [10.1016/j.scitotenv.2020.140575](https://doi.org/10.1016/j.scitotenv.2020.140575)
- 781 78. Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., ... Eveillard, D.
782 (2020). Environmental vulnerability of the global ocean plankton community
783 interactome. *BioRxiv*. doi: 10.1101/2020.11.09.375295
- 784 79. Mihara T, Koyano H, Hingamp P, Grimsley N, Goto S, and Ogata H. Taxon Richness of
785 “Megaviridae” Exceeds those of Bacteria and Archaea in the Ocean. en. Microbes and
786 Environments 2018; 33:162–71. doi: [10.1264/jsme2.ME17203](https://doi.org/10.1264/jsme2.ME17203)
- 787 80. Glock N, Roy AS, Romero D, Wein T, Weissenbach J, Revsbech NP, et al. Metabolic
788 preference of nitrate over oxygen as an electron acceptor in foraminifera from the
789 Peruvian oxygen minimum zone. en. Proceedings of the National Academy of Sciences
790 2019 Feb; 116:2860–5. doi: [10.1073/pnas.1813887116](https://doi.org/10.1073/pnas.1813887116)

791

792 Figure legends

793 **Figure 1:** Geographical locations of *Tara* Oceans epipelagic and mesopelagic sampling sites
794 included in this study. Symbols: ▲ refers to epipelagic, ▼ Oxidic MES and ◆ OMZ eco-regions.
795 Symbol colors represent organism groups evaluated in the present study.

796 **Figure 2:** Ordination plot of epipelagic (left) and mesopelagic (right) communities based on
797 OTU composition based on canonical correspondence analysis (CCA). Percentages in
798 parentheses are the amount of variation constrained, in titles represent the total in each
799 analysis, and in the axis the correspondent value for each dimension. Arrows represent
800 environmental quantitative explanatory variables with arrowheads indicating their direction of
801 increase. Shapes represent sampling sites. Shape formats represent eco-regions, epi:
802 epipelagic, Oxidic MES: oxidic mesopelagic, OMZ: oxygen minimum zone mesopelagic. IO: Indian
803 Ocean, NAO: North Atlantic Ocean, NPO: North Pacific Ocean, SAO: South Atlantic Ocean,
804 SPO: South Pacific Ocean.

805 **Figure 3:** Temperature and salinity plot indicating water mass designation for all mesopelagic
806 samples. Formats represent the different oceanic basins (■ - North Atlantic Ocean, ● - South
807 Atlantic Ocean, ▲ - Pacific Ocean, ★ - Indian Ocean). Colors indicate the oxygen concentration
808 at the sampling depth. LSW - Labrador Sea Water; AAIW - Antarctic Intermediate Water;
809 tNPIW transitional North Pacific Intermediate Water; SAMW - Subantarctic Mode Water;
810 SPSTMW - South Pacific Subtropical Mode Water; modAAIW - modified Antarctic Intermediate
811 Water; PGW - Persian Gulf Water mass; RSW - Red Sea Water mass; NASTMW - North
812 Atlantic Subtropical Mode Water.

813 **Figure 4:** Relative abundance of OTUs assigned to mesopelagic eco-regions. A) Phages, B)
814 Giruses, C) Prokaryotes and D) pico-Eukaryotes.

815 **Figure 5:** Co-occurrence network in epipelagic and mesopelagic communities. A) Global
816 network, with connected modules for OMZ (purple and orange) and MES (green) highlighted.
817 B) Relative taxa abundance in each module in each station and depth. C) Relative number of
818 OTUs classified in taxonomic groups. D) Network representation of modules.

819 **Supplementary Figure 1:** Non-metric multidimensional scaling (NMDS) showing epipelagic
820 and mesopelagic community stratification for each organism group.

821 **Supplementary Figure 2:** Epipelagic and mesopelagic group dispersion based on physical-
822 chemical oceanic properties (Euclidian method). A) First two axes of PCoA. B) Dispersion of
823 distances from samples to centroids.

824 **Supplementary Figure 3:** Relative abundance of OTUs classified into different eco-regions
825 by ocean layers.

826 **Supplementary Figure 4:** Normalized relative abundance of phages and their preferred eco-
827 region.

828 **Supplementary Figure 5:** Normalized relative abundance of giruses and their preferred eco-
829 region.

830 **Supplementary Figure 6:** Normalized relative abundance of prokaryotes and their preferred
831 eco-region.

832 **Supplementary Figure 7:** Normalized relative abundance of pico-eukaryotes and their
833 preferred eco-region.

834 **Supplementary Figure 8:** Relative abundance of OTUs from taxonomic groups for epipelagic
835 and mesopelagic (Oxic MES and OMZ) samples enriched in each eco-region (UBI: ubiquitous,
836 EPI: epipelagic, Core MES: core mesopelagic, Oxic MES: oxic mesopelagic, OMZ: oxygen
837 minimum zone. A) Phages B) Giruses C) Prokaryotes D) pico-eukaryotes.

611 **Figures:**

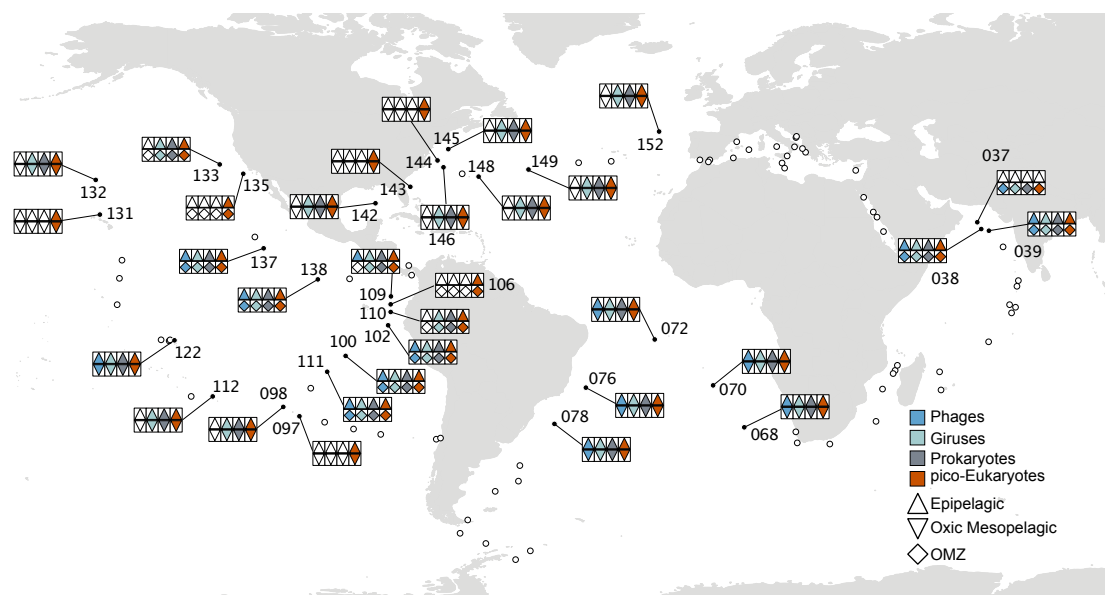


Figure 1: Geographical locations of Tara Oceans epipelagic and mesopelagic sampling sites included in this study. Symbols up-triangle refers to epipelagic, down-triangles Oxidic MES and diamond OMZ eco-regions. Symbols colors represent organism groups evaluated in the present study

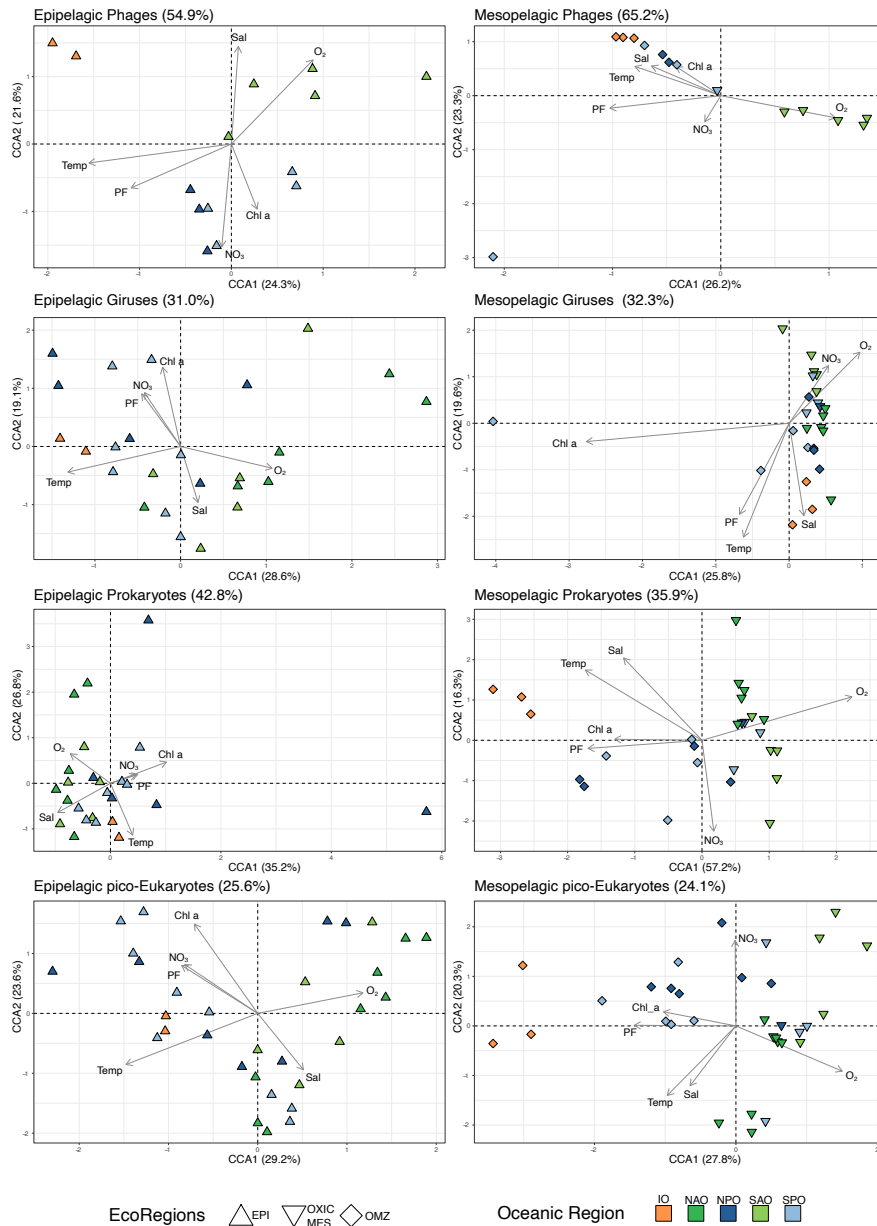


Figure 2: Ordination plot of epipelagic (left) and mesopelagic (right) communities based on OTU's composition based on canonical correspondence analysis (CCA). Percentages in parenthesis are the amount of variation constrained, in titles represents the total in each analysis, and in the axis the correspondent value for each dimension. Arrows represent environmental quantitative explanatory variables with arrowheads indicating their direction of increase. Shapes represent sampling sites. Shape formats represent eco-regions, epi: epipelagic, Oxic_mes: oxic mesopelagic, OMZ: OMZ mesopelagic. IO: Indian Ocean, NAO: North Atlantic Ocean, NPO: North Pacific Ocean, SAO: South Atlantic Ocean, SPO: South Pacific Ocean

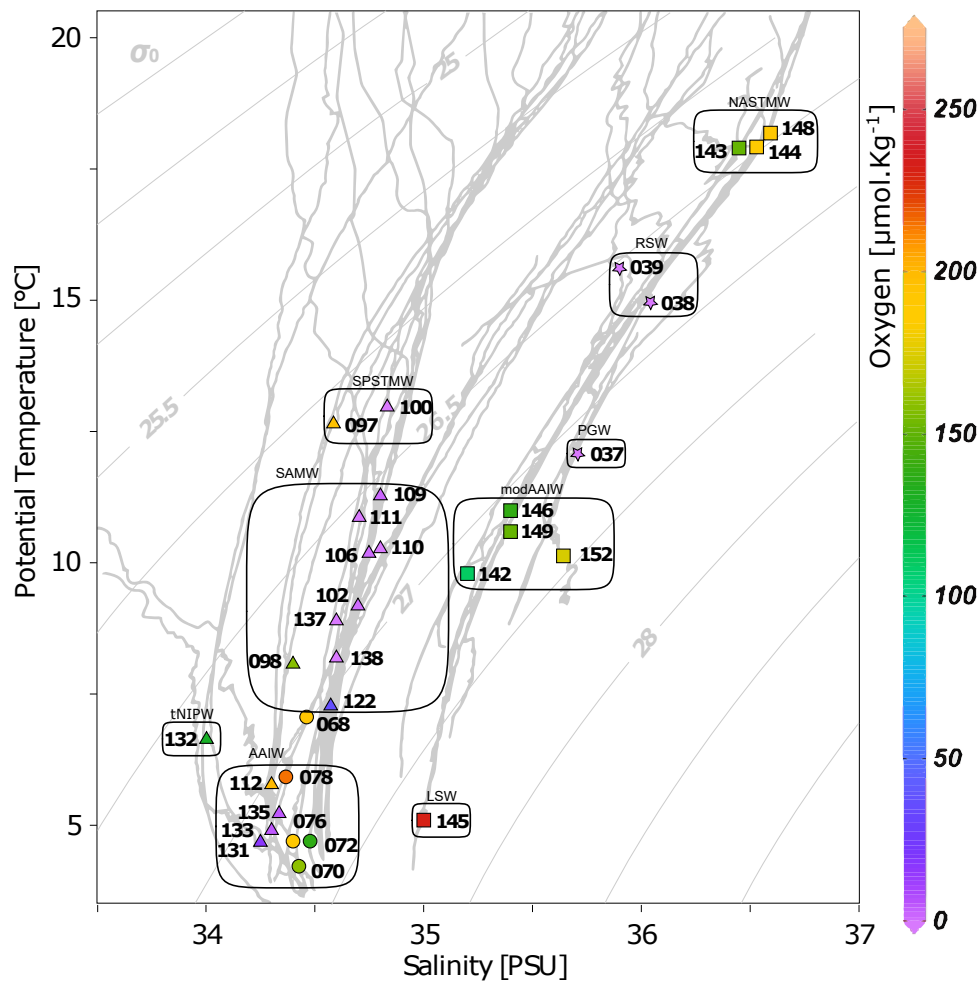


Figure 3: Temperature and salinity plot indicating water masses designation for all mesopelagic samples. Formats represent the different oceanic basins (□ - North Atlantic Ocean, ○ - South Atlantic Ocean, △ - Pacific Ocean, ★ - Indian Ocean). Colours indicate the oxygen concentration at the sampling depth. LSW - Labrador Sea Water; AAIW - Antarctic Intermediate Water; tNIPW ? transitional North Pacific Intermediate Water; SAMW - Subantarctic Mode Water; SPSTMW - South Pacific Subtropical Mode Water; modAAIW - modified Antarctic Intermediate Water; PGW - Persian Gulf Water mass; RSW - Red Sea Water mass; NASTMW - North Atlantic Subtropical Mode Water.

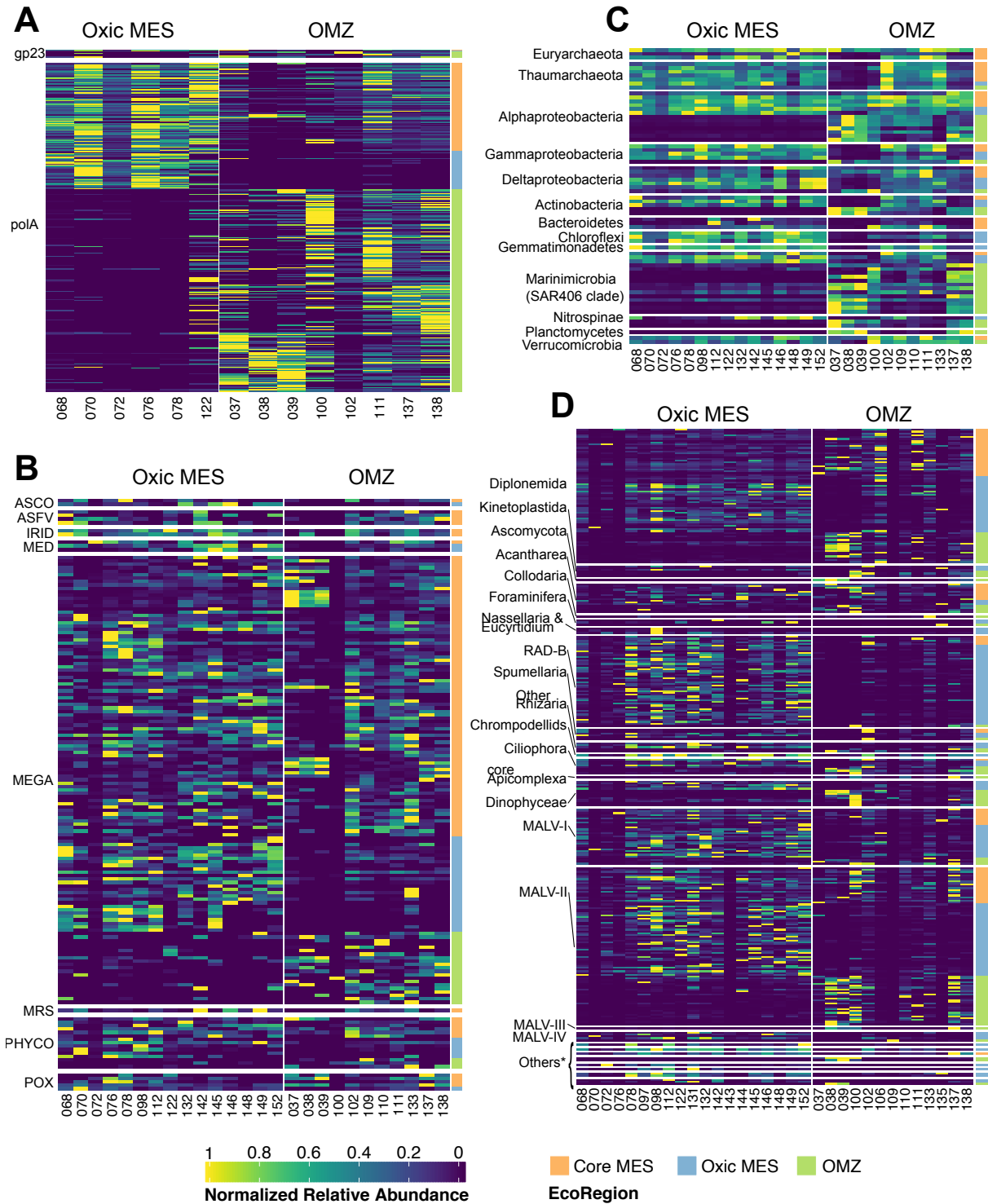


Figure 4: Relative abundance of OTUs assigned to Mesopelagic Eco-regions. A) Phages, B) Giruses, C) Prokaryotes and D) Eukaryotes

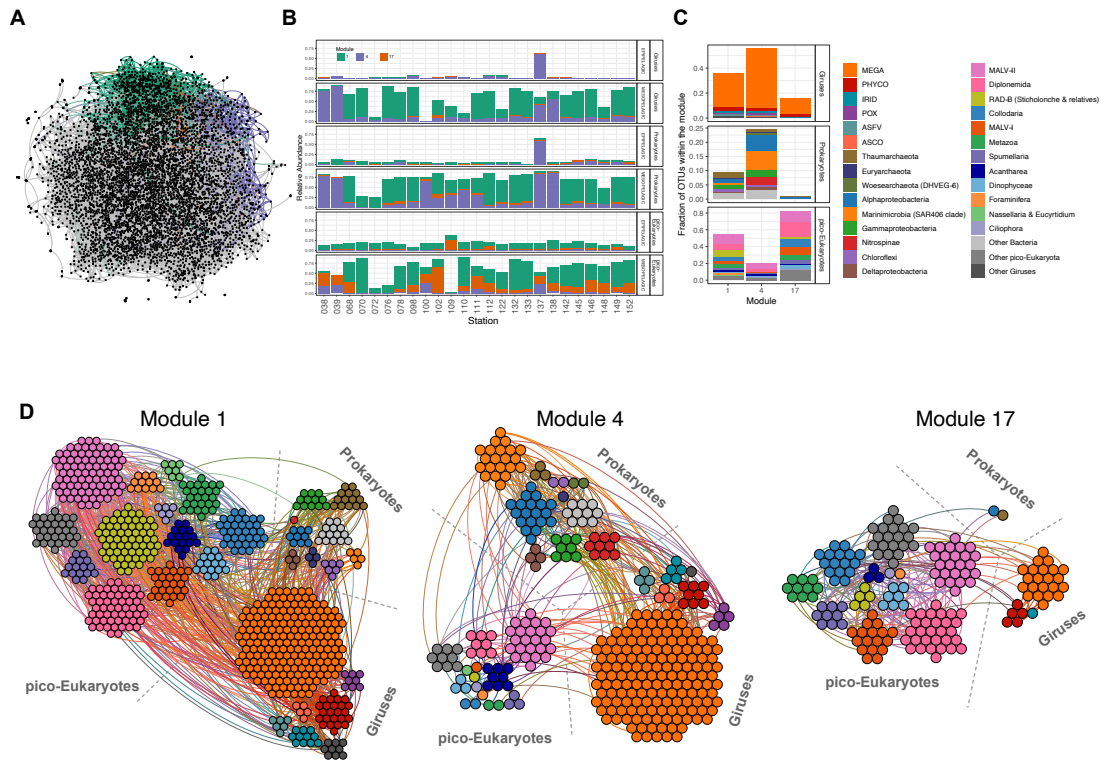


Figure 5: Co-Occurrence Network in Epipelagic and Mesopelagic communities. A) Total Network, with connected modules for OMZ (purple and orange) and MES (green) highlighted. B) Relative taxa abundance in each Module in each station and depth. C) Relative number of OTUs classified in taxonomic groups. D) Network representation of modules

612 **Tables**

Table 1: ANOVA p-value for each environmental factor used as explanatory variable. Global correspond to the p-value for the whole set of variables, the following columns to p-values for each of the environmental variables considering the others as covariable.

assemblage	Depth	Global	Temp. °C	Salinity	O ₂ [μ mol/kg]	[NO ₃] ⁻ [μ mol/l]	Chl-a [μ g/m ³]	Particle flux
Phages	EPI	0.001	0.057	0.144	0.043	0.031	0.420	0.017
Phages	MES	0.001	0.059	0.004	0.027	0.124	0.110	0.002
Giruses	EPI	0.001	0.023	0.074	0.051	0.152	0.053	0.221
Giruses	MES	0.002	0.032	0.211	0.145	0.134	0.008	0.211
Prokaryotes	EPI	0.035	0.048	0.141	0.002	0.006	0.044	0.568
Prokaryotes	MES	0.006	0.501	0.304	0.039	0.444	0.966	0.486
pico-Eukaryotes	EPI	0.025	0.027	0.166	0.412	0.292	0.216	0.659
pico-Eukaryotes	MES	0.001	0.606	0.191	0.243	0.468	0.271	0.477

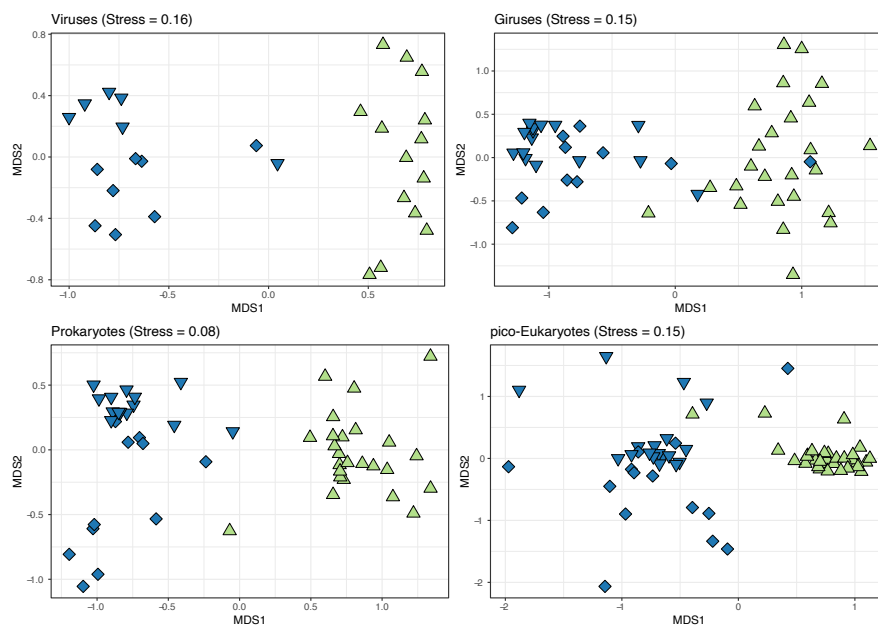
Table 2: Proportion of the variation in community composition that explained by water masses using the Permutation multivariated analysis of variance (PERMANOVA)

assemblage	Df	Sum Of Squares	Mean Squares	F. Model	R ²	Pr($\geq F$)
Phages	4	1.87	0.47	2.32	0.51	0.002
Giruses	8	3.19	0.40	1.81	0.46	0.001
Prokaryotes	8	1.30	0.16	3.29	0.60	0.001
pico-Eukaryotes	8	2.60	0.32	1.62	0.36	0.001

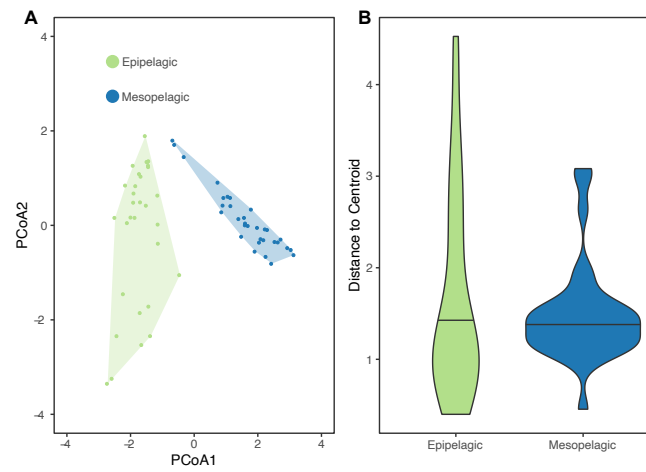
Table 3: Network topological features derived from global analysis including giruses, prokaryotes and pico-eukaryotes samples in epipelagic and mesopelagic depths

Name	Global	Mod 1	Mod 4	Mod 17
Nodes	6154	731	323	175
Positive Edges	12193	1236	480	223
Negative Edges	742	70	49	9
Avg. degree	4.20	3.57	3.28	2.65
Clustering	0.03	0.03	0.09	0.05
Density	0.00	0.00	0.01	0.02
Average.path.length	7.28	6.01	6.27	6.30
Betweenness	0.01	0.05	0.10	0.22
Degree Centralization	0.00	0.01	0.02	0.04
Modularity	0.47	0.60	0.67	0.66

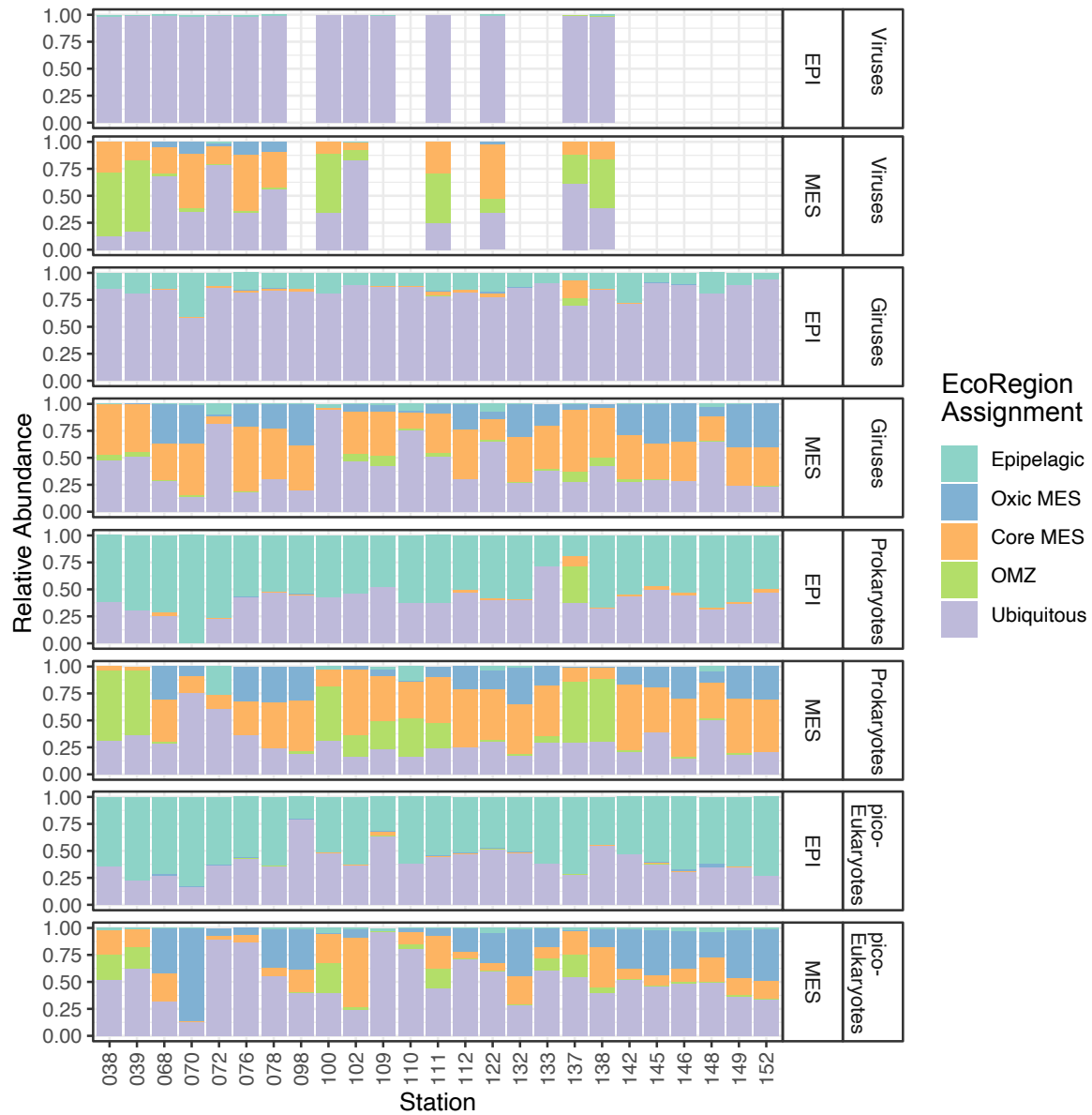
613 **Supplementary Figures:**



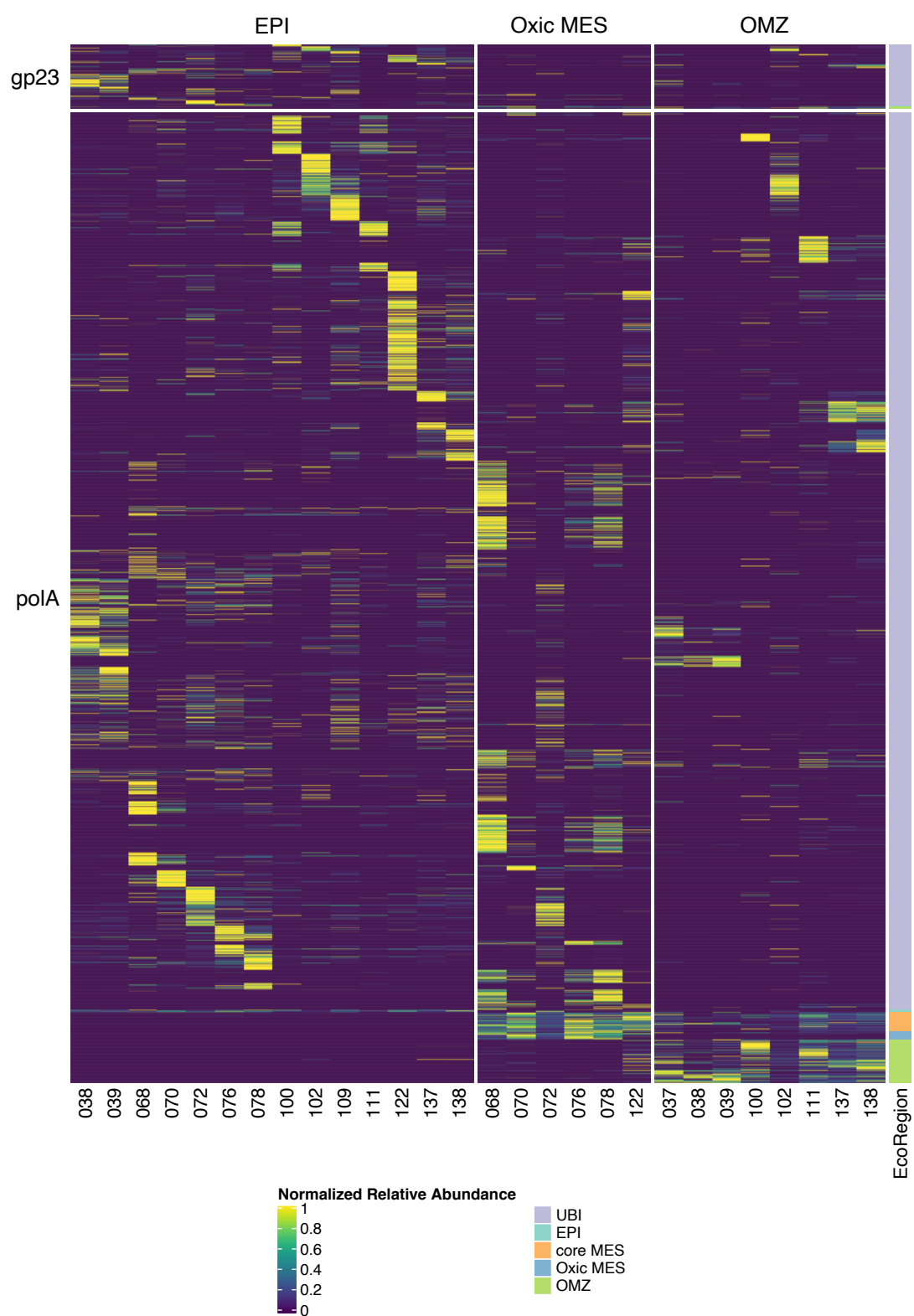
Supplementary Figure 1: Non-metric multidimensional scaling (NMDS) showing epipelagic and mesopelagic communities stratification for each organism group



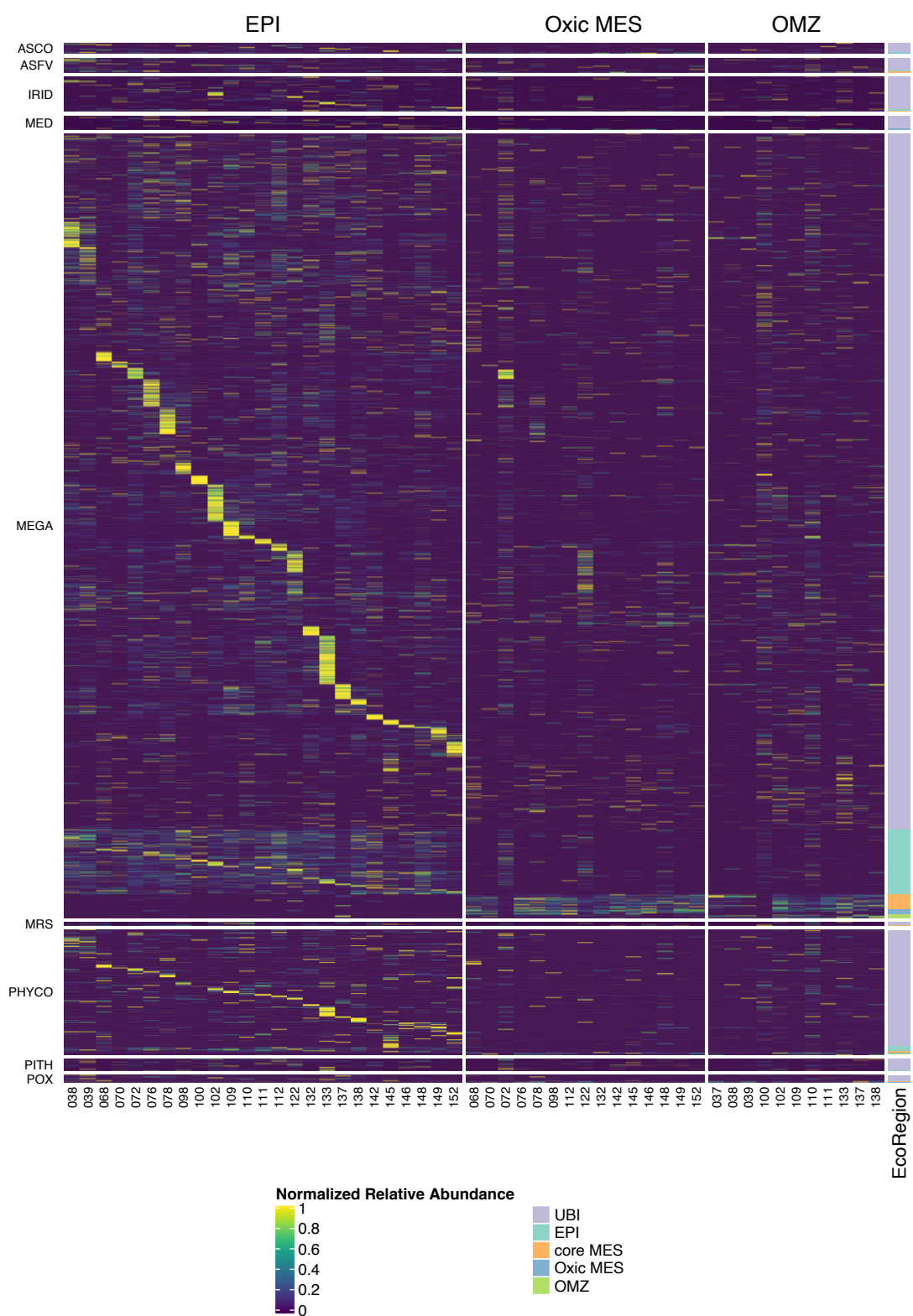
Supplementary Figure 2: Epipelagic and mesopelagic group dispersion based on physical-chemical oceanic properties (Euclidian method). A) First two axes of PCoA. B) Dispersion of distances from samples to centroids.



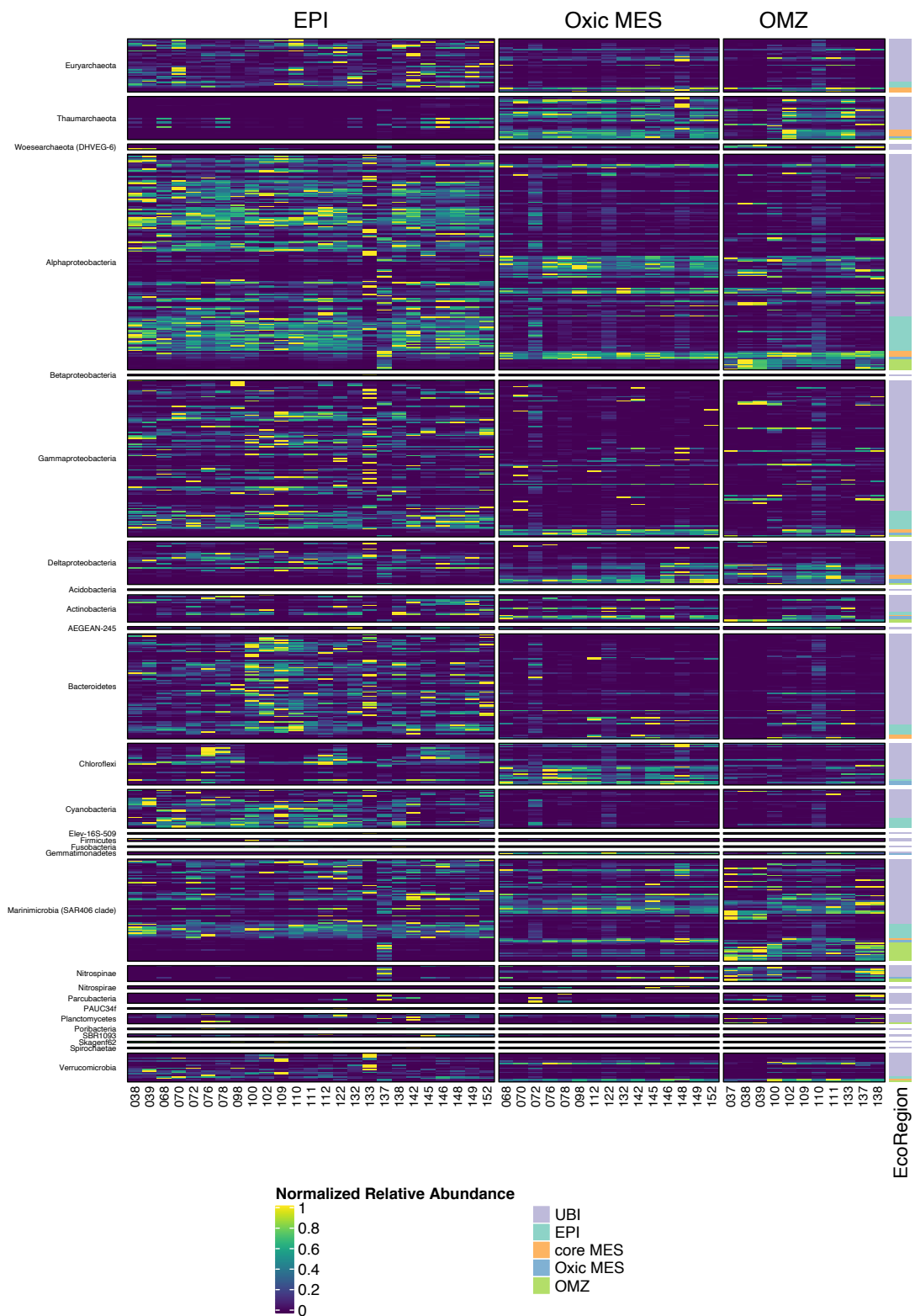
Supplementary Figure 3: Relative abundance of OTUs classified into different eco-regions in to ocean layers



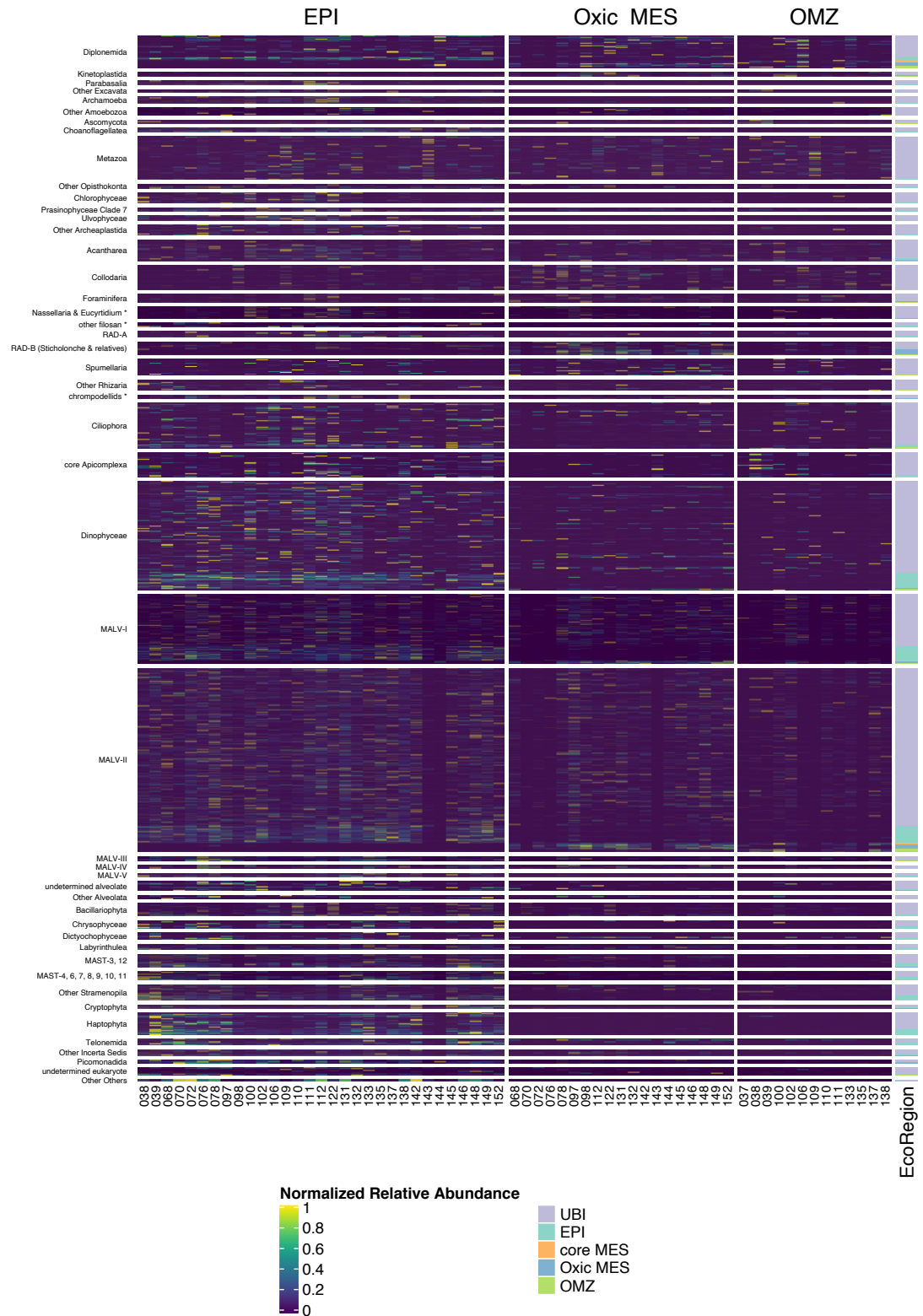
Supplementary Figure 4: Normalized Relative abundance of Phages and their preferred eco-region



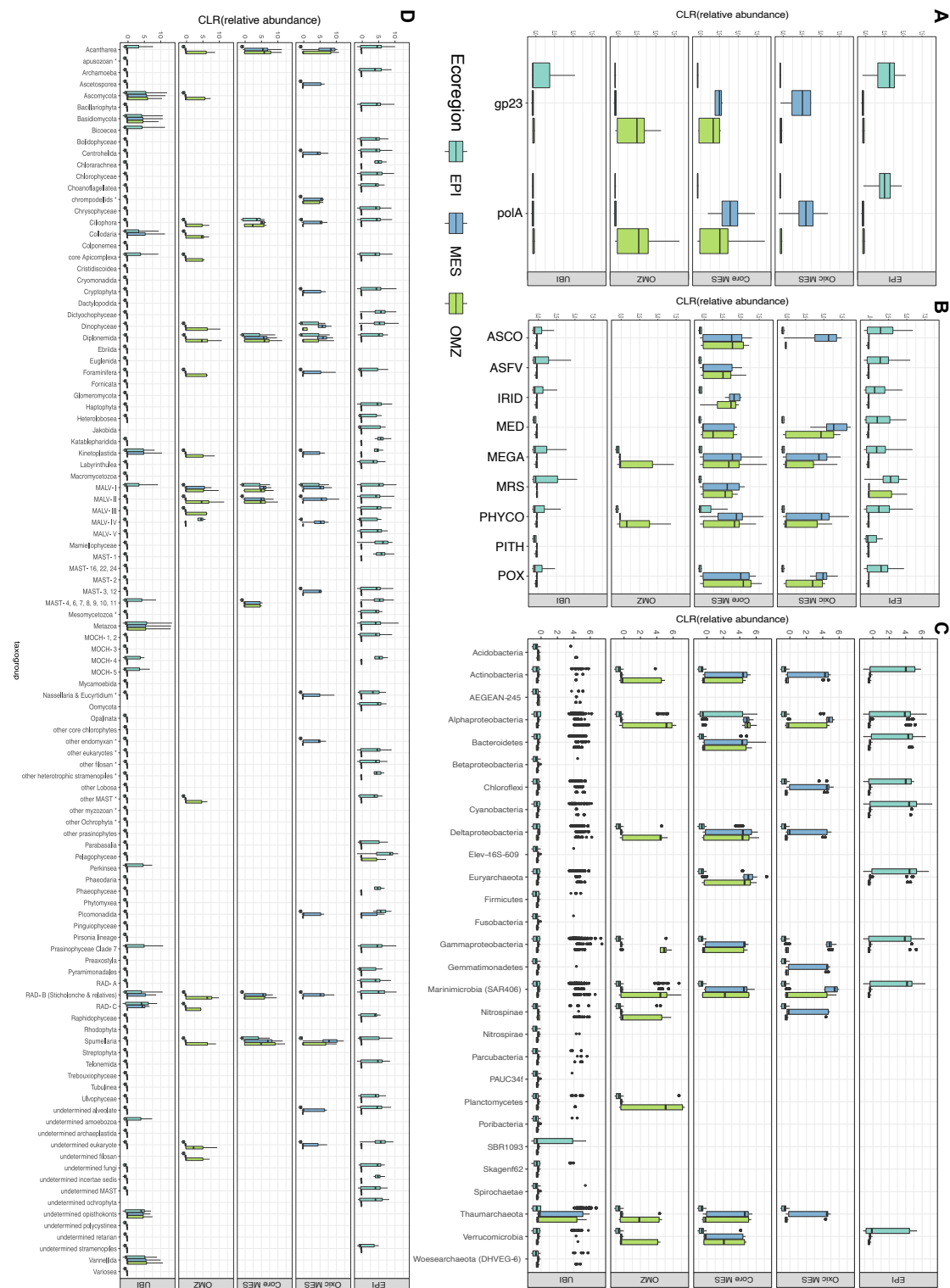
Supplementary Figure 5: Normalized Relative abundance of giruses and their preferred eco-region



Supplementary Figure 6: Normalized Relative abundance of Prokaryotes and their preferred eco-region



Supplementary Figure 7: Normalized Relative abundance of pico-Eukaryotes and their preferred eco-region



Supplementary Figure 8: Relative abundances of OTUs from taxonomic groups for epipelagic and mesopelagic (Oxic MES and OMZ) samples enriched in each eco-region (UBI: ubiquitous, EPI: epipelagic, Core MES: core mesopelagic, Oxic MES: oxenic mesopelagic, OMZ: oxygen minimum zone). A) Phages B) Gyruses C) Prokaryotes D) pico-Eukaryotes