Trait biogeography of marine copepods - an analysis across scales

Brun, Philipp Georg; Payne, Mark R; Kiørboe, Thomas

Published in:
Ecology Letters

Link to article, DOI:
10.1111/ele.12688

Publication date:
2016

Document Version
Peer reviewed version

Trait biogeography of marine copepods – an analysis across scales

Running title (45 char): Trait biogeography of marine copepods

Philipp Brun[1,*] Mark R. Payne[1,a], and Thomas Kiørboe[1,b]

[1] Centre for Ocean Life, National Institute of Aquatic Resources, Technical University of Denmark, Kavalergarden 6, DK-2920 Charlottenlund, Denmark

* Corresponding author:

Email: pgbr@aqua.dtu.dk

Phone: +45 35 88 34 80

Fax +45 35 88 33 33

a mpay@aqua.dtu.dk

b tk@aqua.dtu.dk

Keywords (10): Trait biogeography, copepods, marine zooplankton, global, body size, myelination, offspring size, feeding mode, integrated nested Laplace approximations, Shannon size diversity

Type of article: Letter

Statement of authorship: All authors were involved in the design of the technical set-up of the study. PB compiled and analyzed the data and prepared the manuscript with contributions and support from the other authors.

Number of words in the abstract: (147/150):

Number of words in the main text (excl. acknowledgements, references, legends): (4992/5000)

Number of words in text boxes: (561/750)

Number of references: (55/50)

Number of textboxes/figures/tables (max 6 in total): 6

Abstract

Functional traits, rather than taxonomic identity, determine the fitness of individuals in their environment: traits of marine organisms are therefore expected to vary across the global ocean as a function of the environment. Here, we quantify such spatial and seasonal variations based on extensive empirical data and present the first global biogeography of key traits (body size, feeding mode, relative offspring size and myelination) for pelagic copepods, the major group of marine zooplankton. We identify strong patterns with latitude, season, and between ocean basins that are partially (approximately 50%) explained by key environmental drivers. Body size, for example, decreases with temperature, confirming the temperature-size rule, but surprisingly also with productivity, possibly driven by food-chain length and size-selective predation. Patterns unrelated to environmental predictors may originate from phylogenetic clustering. Our maps can be used as a test-bed for trait-based mechanistic models and to inspire next generation biogeochemical models.

Introduction

Studying the distribution and abundance of organisms is the key task in ecology (Begon *et al.* 2006). In recent decades, the growing availability of observational data and empirical models has increasingly allowed the pursuit of this task on large spatial scales. In particular the distribution patterns of individual species and their links to the physical environment have been studied intensively (Elith & Leathwick 2009). However, a major challenge for such macro-scale studies is the mechanistic linking of the observed patterns to the processes that drive them (Keith *et al.* 2012). One powerful way to identify such links is the trait-based approach, because the functional traits of an organism, rather than its taxonomic identity, determine its fitness in a given environment. The trait-based approach assumes that organism fitness is based on success in the fundamental life missions feeding, survival and reproduction, and that the outcome of each of those missions depends on a few key traits. These key traits are interrelated through trade-offs and their optimal expression is determined by the environmental conditions (Litchman *et al.* 2013).

The trait-based approach in biogeography is well established for primary producers but its potential for animals has rarely been exploited. The trait-based approach has a long tradition in plant ecology (e.g., Westoby *et al.* 2002) and has also been used to describe the distributions of phytoplankton (e.g., Edwards *et al.* 2013). Besides providing ecological insight, trait biogeographies have fostered a more realistic incorporation of primary producers into global vegetation and ocean circulation models and thus have advanced biogeochemistry and climate science research (Scheiter *et al.* 2013; Brix *et al.* 2015). However, trait biogeographies for animals are uncommon, although they may be equally valuable. This is particularly evident for marine zooplankton, and their dominant members, the copepods

(Barton *et al.* 2013b). Marine copepods are ubiquitous, typically dominate the biomass of zooplankton communities, and play a key role in pelagic food webs (Verity & Smetacek 1996). For this group traits and associated trade-offs are relatively well understood (Kiørboe 2011) and comparably rich observational data exists (O'Brien 2010).

Key traits for copepods include body size, feeding mode, relative offspring size, and myelination of the nerves, determining both their fitness and their impact on the ecosystem. Body size governs most vital rates and biotic interactions (Kiørboe & Hirst 2014) and affects marine food webs and carbon fluxes (Turner 2002; García-Comas *et al.* 2016), feeding mode determines feeding efficiency and associated predation risk (Kiørboe 2011), relative offspring size determines the success in recruitment in a given environment (Neuheimer *et al.* 2015), and myelination of the nerves is one aspect of predator defense (Lenz 2012) (Box 1).

The aim of this study is to establish large-scale copepod trait biogeographies, including the first ever global analyses. In addition, we tested two hypotheses: (H1) Between-community trait variation is structured in space and time, i.e., trait distributions can be largely described by assuming that they are more similar to neighboring communities than to distant communities. (H2) These spatiotemporally dependent structures form in response to key environmental drivers including food availability, temperature, water transparency, and seasonality, as suggested in Box 1. We combined information on traits for hundreds of marine pelagic copepod taxa with two of the most extensive sets of observational data for copepods, covering the North Atlantic and the global ocean. We demonstrate distinct spatiotemporal trait biogeographies for most traits that can be partly explained by environmental drivers, and partly, such as in the case of differences between ocean basins, as a result of other structuring processes.

Methods

Overview

The analyses consisted of two steps. Firstly, we combined copepod trait information with field observations of copepod occurrences, defined communities, and summarized those using summary statistics. We combined trait information with two observational datasets with different resolutions in space and time: the North Atlantic with seasonal resolution, and the global ocean without temporal resolution. Secondly, we used statistical models to test our hypotheses, to investigate the spatial/spatiotemporal patterns of trait distributions, and to analyze their relationship with the environment.

Trait data

Trait data originated from a collection of literature information on functional traits for marine copepods (Brun *et al.* 2016). Where multiple measurements were available per species, we took species-specific averages. We used body size measurements from adults irrespective of the life stage of the observed individuals and thus estimated an upper boundary of potential body size. In the global analysis, information on mixed feeding was not sufficient to characterize the communities, and we therefore only distinguished between active feeders and passive feeders, considering mixed feeding taxa as active feeders.

Observational data

North Atlantic

Data from the Continuous Plankton Recorder (CPR) survey was used to estimate the spatiotemporal distributions of North Atlantic copepods. The CPR survey is a large-scale

106 monitoring program of North Atlantic plankton, particularly copepods, diatoms and
107 dinoflagellates (Richardson *et al.*, 2006). The CPR is towed by ships of opportunity at
108 approximately 7 m depth. Each CPR sample corresponds to 10 nautical miles and around 3
109 m^3 of seawater filtered onto a 270 µm-sized silk gauze. We used roughly 49 000 observations
110 of 67 copepod taxa resolved into abundance classes that have been classified by the CPR
111 survey between 1998 and 2008 (Johns 2014, Appendix A).

112 Observations of CPR taxa were matched with taxon-specific trait estimates. Not all
113 taxa sampled in the CPR were resolved to the species level. Traits for higher order taxa were
114 represented by the traits of the most common species in that group, as reported in Richardson
115 *et al.* (2006). Where no information about the most common species was available, we
116 averaged traits of all species in the taxon that have been repeatedly observed in the study
117 area, according to the OBIS database (www.iobis.org, Appendix A). Available trait
118 information largely covered the estimated biomass of observed taxa in the North Atlantic
119 (Table 1).

120 ***Global***

121 For the global analysis we used data from the Coastal and Oceanic Plankton Ecology,
122 Production and Observation Database (COPEPOD), which contains abundance information
123 for various plankton groups (O'Brien 2010). This data is compiled from a global collection of
124 cruises, projects, and institutional holdings. Data for copepods consisted of roughly one
125 million observations distributed across the global ocean. We updated the taxonomic
126 classification of the observations according to the most recent online taxonomy
127 (http://www.marinespecies.org/copepoda/) and utilized only data with abundance information
128 and taxonomic resolution at the genus level or higher. In a few cases, we also included pooled
129 observations for two genera, describing their traits based on the first genus mentioned.

Furthermore, we filtered for observations taken in the top 200 meters of the water column and excluded parasitic taxa. While the absolute number of observations lost through the filtering was minor, observations were removed from most of the Pacific, particularly because of lacking taxonomic resolution of data from this area.

Observations were matched with corresponding trait information. Traits at the genus level were estimated as means of the available estimates for their species. For all traits, match-ups were possible for most of the estimated abundance (Table 1).

COPEPOD data were spatially binned and an expected abundance was estimated for the taxa present. Unlike the CPR data, COPEPOD observations do not have a homogeneous sampling design and no standardized catalogue of taxa was targeted. We therefore split the global ocean into roughly 5000 polygons of similar area, and estimated trait-statistics polygon-wise. For each polygon, we used geometrical means to estimate the relative abundance of each taxon present for which trait information existed.

Summarizing community traits

Community traits were summarized by mass-weighted means and, for body size, also by the Shannon size diversity index. Biomass-weighted means were estimated by using the cubed body length estimates as biomass proxies. In addition, we quantified body-size diversity in copepod communities using the Shannon size diversity index. Body-size diversity characterizes the diversity of size classes within a community, which has been related to food-web properties (García-Comas *et al.* 2016). Furthermore, it indicates whether copepod communities are affected by environmental filtering. The Shannon size diversity index (μ) is analogue to the Shannon diversity index but computed on the probability-density function of a continuous-random variable (Quintana *et al.* 2008). It is estimated as

153 $$\mu = -\int_0^{+\infty} p_x(x) log_2\, p_x(x) dx \quad (1)$$

154 where $p_x(x)$ represents the probability density function of size x.

155 We estimated μ non-parametrically with the Monte Carlo kernel estimation technique
156 (Quintana *et al.* 2008). Shannon size diversity was calculated for all polygons with at least 5
157 observed taxa. The corresponding probability density functions were estimated by weighting
158 the body sizes with the mass fractions of the species present. The Shannon size diversity
159 index is primarily suitable for comparisons between communities.

160 **Environmental data**

161 Environmental variables considered are proxies for the key factors of temperature,
162 available amount of food, prey size, seasonality, and water transparency (Box 1). For
163 temperature, we used the monthly sea surface temperature (SST) data HadISST1 from the
164 Hadley Centre for Climate Prediction and Research, Meteorological Office (Rayner *et al.*
165 2003). Available amount of food was characterized with satellite-derived monthly estimates
166 of net primary productivity (NPP) obtained from
167 http://www.science.oregonstate.edu/ocean.productivity based on the VGPM algorithm
168 (Behrenfeld & Falkowski 1997). Median phytoplankton cell diameter (MD_{50}) was used as
169 proxy for prey size, prey motility, and food quality including lipid content. Flagellates of
170 intermediate size typically have a higher motility and lipid content than large-celled diatoms
171 or small bacterioplankton (Kleppel 1993; McManus & Woodson 2012). Although not all
172 copepods feed solely on phytoplankton, phytoplankton cell size has a strong impact on the
173 entire food web (Barnes *et al.* 2011). MD_{50} was estimated based on empirical relationships
174 with SST and chlorophyll *a* concentration (CHL) (Barnes *et al.* 2011; Boyce *et al.* 2015),
175 where we used the monthly GlobColour CHL1 product (http://www.globcolour.info/) to

represent CHL. Seasonality manifests itself in various ways including photoperiod, temperature, and available diet. For copepods the most immediate impact of seasonality is arguably the food availability. We therefore characterized seasonality by the seasonal variation in chlorophyll *a* concentration, applying the Shannon size diversity index on the CHL data (as this index is suitable to estimate the diversity of any non-negative, continuous variable). Water-column transparency was approximated by Secchi Depth (ZSD), represented by the monthly GlobColour ZSD product. For NPP, data from the period 2003-2008 was considered; for all other predictors, the period considered was 1998-2008.

Environmental variables were aggregated to match the resolution of the copepod communities. For the North Atlantic analysis we produced 1°×1° monthly means for each year for SST, MD_{50}, and ZSD. Since we did not have a complete temporal coverage for NPP, we matched the observations with monthly averages based on the years 2003-2008. CHL seasonality was calculated for each year independently and matched with all months of that year. For the global models, we aggregated the predictors by the polygons used to define the copepod communities, including the entire time-span of data availability. For computational efficiency, and to avoid numerical problems, all environmental variables were discretized to 200 equally-spaced steps, normalized and standardized. Note that particularly on the global scale, some of the predictors showed significant Pearson correlation coefficients (r) up to r=0.86 for SST and MD_{50} (Appendix B). However, the analyses performed here are largely insensitive to collinearity (Dormann *et al.* 2012).

Statistical modelling

The integrated nested Laplace approximation (INLA) approach is a novel and computationally-efficient Bayesian statistical tool that is particularly powerful in handling

spatial and spatiotemporal correlation structures (Rue *et al.* 2009; Blangiardo & Cameletti 2015). We used the INLA approach to model each trait for both observational datasets as a function of i) space (and season), ii) environmental predictors, and iii) as a combination of i) and ii). We modeled the continuous traits (body size, body-size diversity, and relative offspring size) assuming *t*- and normal-distributions for the North Atlantic and the global models, respectively. The categorical traits (feeding modes and myelination) were modeled assuming beta-binomial and binomial distributions, respectively, both of which require a number-of-trials parameter. For the North Atlantic models we defined the numbers of trials by the total counts of individuals per sample and the number of positives was estimated by the weight fraction of these counts showing the trait in question. In the global models, the number of trials was held constant at one. The fitted models were used to map the trait distributions, investigate the relationships between traits and environmental predictors, and to compare the amount of variance explained by the three model set-ups.

Spatial and spatiotemporal models

Spatial and spatiotemporal models were constructed assuming distributions of traits to have a spatially- and temporally-dependent structure. We assumed trait distributions to be isotropic, stationary Gaussian Fields which are approximated with discrete meshes in INLA (Blangiardo & Cameletti 2015). We constructed a spatial mesh for each domain and an additional seasonal mesh for the North Atlantic (Appendix C). Furthermore, we complemented the North Atlantic models with a random effect correcting for variations between the years analyzed.

Environmental models

The environmental modeling approach used is equivalent to ecological niche models, but applied to community properties rather than individual species. For each trait and both observational datasets we fitted models for all possible combinations of the candidate predictors. The predictors were fitted as smooth, non-linear effects using second-order random-walk models (Rue *et al.* 2009), an approach similar to common generalized additive models (GAMs; Wood 2006) where the non-parametric response form of each predictor is determined by the data. Based on these models we assessed the best predictor combination for each trait according to the minimum Watanabe-Akaike information criterion (WAIC), a modified version of the Akaike Information Criteria that is appropriate for use with mixed-effects models (Gelman *et al.* 2014). We further used the univariate environmental models to investigate trait-environment relationships: univariate models were chosen over multivariate models to prevent distortions due to collinear predictors (Dormann *et al.* 2012).

Combined models

"Combined" models were created by adding spatial/spatiotemporal structures to the best environmental models (Blangiardo & Cameletti 2015).

Evaluation of hypotheses

Both of our hypotheses focused on between-community variance of traits. The existence of such variance was confirmed in a preliminary assessment (Appendix D). Hypothesis H1 (community traits are spatially structured) was then tested by quantifying the fraction of variance explained (R^2) by spatial/spatiotemporal models, and hypothesis H2 (spatial structure is explained by key environmental drivers) was evaluated by comparing the R^2 of the best environmental models with the R^2 of the combined models.

243

Results

Evaluation of hypotheses

All traits examined showed distinct structure in space and time, both globally (no temporal resolution) and in the North Atlantic, confirming our hypothesis H1. Our spatial and spatiotemporal models could explain substantial fractions of the between-community trait variance based on the spatial dependency assumption. This was particularly true for global patterns, where R^2 of spatial models ranged from 0.36 for active feeding to 0.75 for body size (Figure 1a). In the North Atlantic, the spatiotemporal models were somewhat less efficient for the more finely-resolved communities of the CPR observations and ranged from R^2=0.32 for body-size diversity to R^2=0.48 for body size (Figure 1b).

Our second hypothesis, that we can explain these spatial patterns with key environmental drivers, proved partially valid. On average, environmental models (green bars in Figure 1c,d) reached approximately half of the R^2 of combined models (yellow bars in Figure 1c,d), indicating that about half the patterns in the investigated traits could be explained by the environmental predictors hypothesized to be important. The ratio between R^2 for environmental models and R^2 for combined models was somewhat higher in the global domain and peaked at 78% for the global myelination model. Similarly, body size and body-size diversity could be explained relatively well by the environment, with corresponding percentages well above the 50% in both domains. For active feeding, on the other hand, environmental models performed relatively poorly and could only explain minor fractions of the identified patterns.

Trait distributions

Seasonal variation in trait distributions in the North Atlantic

All traits examined showed seasonally-varying distribution patterns. Mean community body size varied substantially and mainly ranged between 1 and 5 mm in the North Atlantic (Figure 2a-d), corresponding to a two order-of-magnitude variation in body mass. Communities with the largest mean body size occurred from spring to autumn in the northwestern North Atlantic, in particular in the Labrador Sea (Figure 2b-d). Smallest community-averaged body size was observed in the central and eastern part of the investigated area, mainly during summer (Figure 2c). From spring to autumn, steep spatial gradients in body size existed while the distribution was mostly uniform during winter.

The diversity of body size in copepod communities was estimated to be highest in winter when values were evenly distributed throughout most of the investigated domain (Figure 2e). In spring and autumn, body-size diversity was similarly high in the central North Atlantic, but smaller in the coastal areas in the east and the west (Figure 2f,h). Lowest body-size diversity was found in summer in the entire investigated area, except for the northwestern North Atlantic around the Labrador Sea (Figure 2g).

Active feeding was estimated to be the dominant feeding mode in the North Atlantic. This was particularly true for winter and spring, where, apart from a few exceptions along the coasts, the communities consisted of at least 66% active feeders (Figure 2i,j). In the eastern part of the investigated area, including the northwestern European coasts, this dominance of active feeders was reduced during summer and autumn and often replaced by a co-dominance of mixed and active feeders (Figure 2k,l).

Myelinated copepods dominated the communities in the North Atlantic overall, yet there was considerable spatiotemporal variation. In winter, myelinated and amyelinated

fractions were roughly in balance, except for the northern central part of the investigated area, where the communities were almost exclusively amyelinated (Figure 2m). The patterns changed markedly in spring when the dominance of myelinated copepods was the greatest, foremost in the northern part of the investigated area (Figure 2n). In summer, and particularly in autumn, the fraction of amyelinated copepods increased again, mainly along the coasts and in the southern and eastern part of the investigated area (Figure 2o,p).

On the community level, egg-size varied on average between about 4.5% and 7.5% of the body size of adult females in the North Atlantic. Highest relative offspring size was observed during winter months in the central part of the investigated area (Figure 2q). In spring, relative offspring size was smaller, in particular in the northwestern North Atlantic, while it gradually increased toward the southeastern part of the investigated area (Figure 2r). In summer and autumn relative offspring size showed a patchy distribution with less variation (Figure 2s,t).

Global trait distributions

The traits investigated also showed clear spatial patterns on the global scale. Mean body size mainly ranged between 1.5 and 7 mm for communities observed in the global ocean (polygons in Figure 3). Largest body sizes were found at high latitudes above 50°, except for the North Atlantic where communities with intermediate body size extended somewhat further northward (Figure 3a). According to the best environmental model, the latitudes with the smallest body size were found in the subtropics while around the equator the mean body size was slightly larger. The smallest body sizes were found in the subtropical central Atlantic, 2-3 mm, whereas communities at similar latitudes in the Indian Ocean tended to have larger mean body sizes, around 3-4 mm. Myelination was distributed similarly to body size (pixel to pixel Spearman correlation coefficient, $r_{spearman}$=0.84) but with more

small-scale variation (Figure 3b): at high latitudes myelinated copepods dominated, while at low and intermediate latitudes myelinated and amyelinated taxa were similarly abundant. Again, the central Atlantic differed from the Indian Ocean with a lower fraction of myelinated organisms. Relative offspring size was inversely proportional to body size ($r_{spearman}$=-0.69) and myelination ($r_{spearman}$=-0.65). In the global ocean relative egg sizes varied between about 3% and 8%, with the relatively largest eggs at low latitudes and the relatively smallest eggs at high latitudes (Figure 3c).

Trait-environment relationships

Environmental responses of most traits were comparable between the global ocean and the North Atlantic analyses (Figure 4), although they tended to be weaker in the North Atlantic. Highest body size was found at low NPP, intermediate phytoplankton cell size and low SST (Figure 4a-c). While globally only intermediate chlorophyll seasonality favored copepod communities with large body size, in the North Atlantic these communities were also found at low CHL seasonality (Figure 4d). Communities with high body-size diversity were most common in environments with low NPP, CHL seasonality and phytoplankton cell size (Figure 4e,f,h). Furthermore, high body-size diversity was found at the high and the low end of the temperature spectrum, while temperatures around 10°C were associated with the lowest diversity (Figure 4g). On the global scale, the best model for body-size diversity did not include CHL seasonality. The weight fraction of myelinated copepods was highest in environments with low NPP, and intermediate Secchi Depth (Figure 4i-k). In the global ocean the fraction of myelinated copepods increased with phytoplankton cell size, while in the North Atlantic it peaked at a median cell size of around 6 µm and rapidly decreased with larger phytoplankton. Finally, relative offspring size was smallest for low NPP, intermediate phytoplankton cell size and relatively short Secchi Depths of 5-25 m (Figure 4l-n). The best

global model for relative offspring size did not include Secchi Depth. WAIC values for all model combinations of traits and environmental predictors can be seen in Appendix G.

Discussion

Our analysis of copepod trait distributions revealed a wealth of strong patterns along several spatial and temporal gradients. Most of these patterns were consistent with the literature or comparable to the trait distributions of other organism groups, yet there were some surprising findings too. Several traits showed considerable latitudinal variation. For example, mean body size was clearly larger at high latitudes than at low latitudes, while it was smallest in the subtropics, and slightly larger around the equator. This pattern is equivalent to the distribution of phytoplankton cell size, and, along the Atlantic Meridional Transect, to the distribution of body size of total zooplankton (San Martin *et al.* 2006; Boyce *et al.* 2015). Relative offspring size also changed significantly with latitude and was highest in the subtropics and tropics, paralleling the distribution of seed mass in terrestrial plants (Moles & Westoby 2003). Trait distributions also showed strong seasonal dynamics. For example, body size in the North Atlantic varied considerably throughout the season with largest copepods in March and April. Similar dynamics have been found for diatoms in the same area, with the largest mean cell size between January and March (Barton *et al.* 2013a). More unexpected were the clear differences between the central Atlantic and the Indian Ocean found in all traits investigated. This difference was unrelated to the known environmental parameters and has not been found in phytoplankton trait distributions (Barnes *et al.* 2011).

A substantial fraction of the spatial and temporal patterns could be linked to the environmental predictors investigated. While temperature seemed to affect copepod traits directly, productivity may influence them in more complex ways. It is well established for both terrestrial and aquatic organisms that within species, body size is inversely related to

363 temperature (Forster *et al.* 2012), and this also applies to copepods (Horne *et al.* 2016). Our
364 results demonstrate that this relationship also holds on the community level. However, body
365 size changed relatively little with increasing temperature when compared to its steep decline
366 with increasing productivity. A negative relationship between body size and productivity is
367 surprising: many groups of marine fish and terrestrial mammals grow larger in areas of higher
368 productivity (Huston & Wolverton 2011), and the same was found for copepods in laboratory
369 experiments (Berggreen *et al.* 1988). For copepods in the field this may be different due to
370 size-selective predation by planktivorous fish (Brucet *et al.* 2010), which are particularly
371 abundant in productive ecosystems like upwelling regions (Cury *et al.* 2000). Furthermore, in
372 oligotrophic open ocean areas planktonic food chains tend to be longer (Boyce *et al.* 2015).
373 Thus, although copepods at the same trophic level may be smaller in areas with low
374 productivity, the mean body size of the entire copepod community may be larger.

375 In contrast to body size, relative offspring size was positively correlated with NPP,
376 possibly in response to stronger biotic interactions. Large offspring size is often seen as an
377 adaptation to harsh environments (Segers & Taborsky 2011), and therefore a positive
378 correlation between relative offspring size and productivity may seem surprising at first sight.
379 However, few offspring and comparably high investments in each individual are also
380 characteristics of K-selected species, which live in densely populated communities
381 (MacArthur & Wilson 1967). In this case, relatively larger offspring may be better in
382 competing for resources and avoiding predation, as has been found for fish: fish fry from
383 large eggs are more tolerant to starvation, avoid predation risks more consequently, and have
384 larger reaction distances to potential predators (Miller *et al.* 1988; Segers & Taborsky 2011).
385 Similarly in terrestrial plants, seed mass is positively correlated to NPP (Moles & Westoby
386 2003).

About half of the identified spatiotemporal patterns could not be explained by the environmental predictors, but arose from other structuring processes. Some of these unexplained patterns occurred on large spatial scales, where the most-pronounced and surprising differences occurred between the central Atlantic and the Indian Ocean. On these scales evolutionary history may affect trait distributions. The distribution range of copepod species is limited by their ability to maintain viable populations (Norris 2000), although, in principle, water parcels can travel between any pair of locations in the global ocean within a decade (Jönsson & Watson 2016). Patterns unexplained by the environmental predictors also occurred on smaller spatial scales in the North Atlantic. On these scales other trait-environment interactions, for example, success in overwintering, may play a role, as well as transportation by ocean currents (Melle *et al.* 2014). Finally, sampling bias may have caused some unexplained patterns, in particular in the global dataset, where sampling methods and taxonomic detail may have differed somewhat between sampling efforts in different areas.

Besides identifying potential drivers of trait distributions, our results, particularly the distribution of body size, also provide insight into how copepod communities affect marine ecosystems and carbon fluxes. The distribution of body size in copepod communities has implications for the fate of the primary production, and determines whether it is recycled in the upper ocean, transported to the sea floor via fecal pellets, or channeled toward higher trophic levels. Copepod fecal pellets may contribute a significant but highly variable (0-100 %) fraction to the vertical material fluxes in the ocean (Turner 2002), and body size of copepods appears to be the main determinant of this fraction (Stamieszkin *et al.* 2015): small copepods produce small fecal pellets that are mainly recycled in the upper ocean, while large copepods produce large pellets that rapidly sink to the seafloor. Body-size diversity of mesozooplankton communities, which are typically dominated by copepods (Verity &

411 Smetacek 1996), is furthermore positively correlated with the transfer efficiency of primary
412 production to higher trophic levels (García-Comas *et al.* 2016): the optimal prey size of
413 primary consumers depends on their body size, and therefore communities of primary
414 consumers with diverse body sizes feed efficiently on a range of prey sizes and harvest the
415 phytoplankton communities more exhaustively. Similarly, changes in phyto- and zooplankton
416 community body size composition have been shown to affect the spatial distribution and
417 temporal dynamics of planktivorous fish. In upwelling areas worldwide, spatial distribution
418 and multi-decadal fluctuations of sardine and anchovy stocks have been explained by
419 climate-driven changes in the physical environment and their impact on plankton body size
420 (e.g., Lindegren *et al.* 2013). Smaller-sized plankton promote filter-feeding fish species with
421 fine gill rakes (e.g., sardine) while larger plankton support particulate-feeders with coarse gill
422 rakes (e.g., anchovy) (van der Lingen *et al.* 2006).

423 Focusing on the large-scale spatial and temporal patterns of copepod trait distributions
424 is necessarily crude and ignores conditions specific to certain regions, especially in data-
425 scarce systems like the open ocean. Particularly with our global approach we defined
426 communities in a simplistic way, included some coarse taxonomic groups, and ignored
427 intraspecific variation in continuous traits such as body size. Our observational data were not
428 evenly distributed in the global ocean, and, especially in the Pacific, data with the required
429 quality were largely lacking. Furthermore, our analysis was biased toward large copepods, as
430 it was based on traditional observational data that were mostly taken with mesh sizes of 200
431 µm or coarser (O'Brien 2010). These meshes may not capture one third of the copepod
432 biomass in the small size fractions (Gallienne & Robins 2001), which is particularly rich in
433 passive feeding taxa like *Oithona* - a potential explanation for the small fractions of passive
434 feeders we identified in this study (Figure 2, Appendix E).

Some of these uncertainties could be reduced by employing approaches that measure traits directly in the field rather than indirectly via taxonomic classification and subsequent merging with trait information from the literature. *In-situ* imaging may be one way to do so (Picheral *et al.* 2010). Taking images of plankton communities with cheap, automated devices carried by commercial ships similar to the Continuous Plankton Recorder (Richardson et al. 2006) could greatly speed-up the sampling and improve data coverage. Imaging may be particularly suitable to measure body size compositions (García-Comas *et al.* 2016), but with the rapid development of algorithm-based image recognition, it may soon be possible to also measure other traits such as sac-spawning or swimming behavior.

Nevertheless, our trait biogeographies showed substantial spatial and temporal structure that was consistently linked to environmental predictors for two independent observational datasets, highlighting the relevance of the trait-based approach to describe copepod biogeography. We demonstrated the value of these biogeographies to test and develop new hypotheses about the drivers of the distribution of zooplankton. Furthermore, our results may be used as a test-bed for trait-based mechanistic models. Ultimately we hope our work will contribute to the development of next generation global models of the dynamics of planktonic ecosystems and their reaction to future climate change.

Acknowledgements

We acknowledge the Villum foundation for support to the Centre for Ocean Life and the European Union 7th Framework Programme (FP7 2007–2013) under grant agreement number 308299 (NACLIM). Likewise, we wish to thank the many current and retired scientists at SAHFOS whose efforts over the years helped to establish and maintain the Continuous Plankton Recorder survey.

References

1.Barnes, C., Irigoien, X., De Oliveira, J.A.A., Maxwell, D. & Jennings, S. (2011). Predicting marine phytoplankton community size structure from empirical relationships with remotely sensed variables. *J. Plankton Res.*, 33, 13–24.

2.Barton, A.D., Finkel, Z. V., Ward, B. a., Johns, D.G. & Follows, M.J. (2013a). On the roles of cell size and trophic strategy in North Atlantic diatom and dinoflagellate communities. *Limnol. Oceanogr.*, 58, 254–266.

3.Barton, A.D., Pershing, A.J., Litchman, E., Record, N.R., Edwards, K.F., Finkel, Z. V, *et al.* (2013b). The biogeography of marine plankton traits. *Ecol. Lett.*, 16, 522–534.

4.Begon, M., Townsend, C.R. & Harper, J.L. (2006). *Ecology: From Individuals to Ecosystems*. 4th edn. Blackwell Publishing, Malden, MA.

5.Behrenfeld, M.J. & Falkowski, P.G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.*, 42, 1–20.

6.Berggreen, U., Hansen, B. & Kiørboe, T. (1988). Food size spectra, ingestion and growth of the copepodAcartia tonsa during development: Implications for determination of copepod production. *Mar. Biol.*, 99, 341–352.

7.Blangiardo, M. & Cameletti, M. (2015). *Spatial and Spatio-temporal Bayesian Models with R-INLA*. 1st edn. Wiley, Chichester, West Sussex, United Kingdom.

8.Boyce, D.G., Frank, K.T. & Leggett, W.C. (2015). From mice to elephants: overturning the "one size fits all" paradigm in marine plankton food chains. *Ecol. Lett.*, 18, 504–515.

9.Brix, H., Menemenlis, D., Hill, C., Dutkiewicz, S., Jahn, O., Wang, D., *et al.* (2015). Using Green's Functions to initialize and adjust a global, eddying ocean biogeochemistry general circulation model. *Ocean Model.*, 95, 1–14.

10.Brucet, S., Boix, D., Quintana, X.D., Jensen, E., Nathansen, L.W., Trochine, C., *et al.* (2010). Factors influencing zooplankton size structure at contrasting temperatures in coastal shallow lakes: Implications for effects of climate change. *Limnol. Oceanogr.*, 55, 1697–1711.

11.Brun, P., Payne, M.R. & Kiørboe, T. (2016). A trait database for marine copepods. Earth Syst. Sci. Data Discuss., 1–33.

12.Cury, P., Bakun, A., Crawford, R.J.M., Jarre, A., Quinones, R.A., Shannon, L.J., *et al.* (2000). Small pelagics in upwelling systems: patterns of interaction and structural changes in "wasp-waist" ecosystems. *ICES J. Mar. Sci.*, 57, 603–618.

13.Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., *et al.* (2012). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop.).*, 36, 27–46.

14.Edwards, K.F., Litchman, E. & Klausmeier, C.A. (2013). Functional traits explain phytoplankton community structure and seasonal dynamics in a marine ecosystem. *Ecol. Lett.*, 16, 56–63.

15.Elith, J. & Leathwick, J.R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.*, 40, 677–697.

16.Forster, J., Hirst, A.G. & Atkinson, D. (2012). Warming-induced reductions in body size are greater in aquatic than terrestrial species. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 19310–4.

17.Gallienne, C.P. & Robins, B.D. (2001). Is *Oithona* the most important copepod in the world's oceans? *J. Plankton Res.*, 23, 1421–1432.

18.García-Comas, C., Sastri, A.R., Ye, L., Chang, C., Lin, F., Su, M., *et al.* (2016). Prey size diversity hinders biomass trophic transfer and predator size diversity promotes it in planktonic communities. *Proc. R. Soc. B Biol. Sci.*, 283, 20152129.

19.Gelman, A., Hwang, J. & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Stat. Comput.*, 24, 997–1016.

20.Hansen, B., Bjørnsen, P.K. & Hansen, P.J. (1994). The size ratio between planktonic predators and their prey. *Limnol. Oceanogr.*, 39, 395–403.

21.Hopcroft, R.R., Roff, J.C. & Chavez, F.P. (2001). Size paradigms in copepod communities: a re-examination. *Hydrobiologia*, 453/454, 133–141.

22.Horne, C.R., Hirst, A.G., Atkinson, D., Neves, A. & Kiørboe, T. (2016). A global synthesis of seasonal temperature-size responses in copepods. *Glob. Ecol. Biogeogr.*, 1–12.

23.Huston, M.A. & Wolverton, S. (2011). Regulation of animal size by eNPP, Bergmann's rule, and related phenomena. *Ecol. Monogr.*, 81, 349–405.

24.Johns, D.G. (2014). Raw data for copepods in the North Atlantic (25-73N, 80W-20E) 1998-2008 as recorded by the Continuous Plankton recorder. Doi: 10.7487/2014.344.1.138

25.Jönsson, B.F. & Watson, J.R. (2016). The timescales of global surface-ocean connectivity. *Nat. Commun.*, 7, 11239.

26.Keith, S.A., Webb, T.J., Bohning-Gaese, K., Connolly, S.R., Dulvy, N.K., Eigenbrod, F., *et al.* (2012). What is macroecology? *Biol. Lett.*, 8, 904–906.

27.Kiørboe, T. (2011). How zooplankton feed: mechanisms, traits and trade-offs. *Biol. Rev.*, 86, 311–339.

27.Kiørboe, T. (2011). How zooplankton feed: Mechanisms, traits and trade-offs. *Biol. Rev.*

28.Kiørboe, T. (2013). Attack or Attacked: The Sensory and Fluid Mechanical Constraints of Copepods' Predator-Prey Interactions. *Integr. Comp. Biol.*, 53, 821–831.

29.Kiørboe, T. & Hirst, A.G. (2014). Shifts in Mass Scaling of Respiration, Feeding, and Growth Rates across Life-Form Transitions in Marine Pelagic Organisms. *Am. Nat.*, 183, E118–E130.

30.Kleppel, G. (1993). On the diets of calanoid copepods. *Mar. Ecol. Prog. Ser.*, 99, 183–195.

31.Lenz, P.H. (2012). The biogeography and ecology of myelin in marine copepods. *J. Plankton Res.*, 34, 575–589.

32.Lindegren, M., Checkley, D.M., Rouyer, T., MacCall, A.D. & Stenseth, N.C. (2013). Climate, fishing, and fluctuations of sardine and anchovy in the California Current. *Proc. Natl. Acad. Sci.*, 110, 13672–13677.

572 33.van der Lingen, C., Hutchings, L. & Field, J. (2006). Comparative trophodynamics of
573 anchovy *Engraulis encrasicolus* and sardine *Sardinops sagax* in the southern Benguela: are
574 species alternations between small pelagic fish trophodynamically mediated? *African J. Mar.*
575 *Sci.*, 28, 465–477.

576
577 34.Litchman, E., Ohman, M.D. & Kiørboe, T. (2013). Trait-based approaches to zooplankton
578 communities. *J. Plankton Res.*, 35, 473–484.

579
580 35.MacArthur, R. & Wilson, E.O. (1967). *The Theory of Island Biogeography*. *Theory Isl.*
581 *Biogeogr.* Princeton University Press.

582
583 36.McManus, M.A. & Woodson, C.B. (2012). Plankton distribution and ocean dispersal. *J.*
584 *Exp. Biol.*, 215, 1008–16.

585
586 37.Melle, W., Runge, J., Head, E., Plourde, S., Castellani, C., Licandro, P., *et al.* (2014). The
587 North Atlantic Ocean as habitat for *Calanus finmarchicus*: Environmental factors and life
588 history traits. *Prog. Oceanogr.*, 129, 244–284.

589
590 38.Miller, T.J., Crowder, L.B., Rice, J. a. & Marschall, E. a. (1988). Larval Size and
591 Recruitment Mechanisms in Fishes: Toward a Conceptual Framework. *Can. J. Fish. Aquat.*
592 *Sci.*, 45, 1657–1670.

593
594 39.Moles, A.T. & Westoby, M. (2003). Latitude, seed predation and seed mass. *J. Biogeogr.*,
595 30, 105–128.

596
597 40.Neuheimer, A.B., Hartvig, M., Heuschele, J., Hylander, S., Kiørboe, T., Olsson, K.H., *et*
598 *al.* (2015). Adult and offspring size in the ocean over 17 orders of magnitude follows two life
599 history strategies. *Ecology*, 96, 3303–3311.

600
601 41.Norris, R.D. (2000). Pelagic species diversity, biogeography, and evolution. *Paleobiology*,
602 26, 236–258.

603
604 42.O'Brien, T.D. (2010). *COPEPOD: The Global Plankton Database. An overview of the*
605 *2010 database contents, processing methods, and access interface.* US Dep. Commerce,
606 NOAA Tech. Memo NMFS-F/ST-36, 28 pp.

607
608 43.Picheral, M., Guidi, L., Stemmann, L., Karl, D.M., Iddaoud, G. & Gorsky, G. (2010). The
609 Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of
610 particle size spectra and zooplankton. *Limnol. Oceanogr. Methods*, 8, 462–473.

611
612 44.Quintana, X.D., Brucet, S., Boix, D., López-Flores, R., Gascón, S., Badosa, A., *et al.*
613 (2008). A nonparametric method for the measurement of size diversity with emphasis on data
614 standardization. *Limnol. Oceanogr. Methods*, 6, 75–86.

615
616 45.Rayner, N.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L. V., Rowell, D.P.,
617 *et al.* (2003). Global analyses of sea surface temperature, sea ice, and night marine air
618 temperature since the late nineteenth century. *J. Geophys. Res.*, 108, 4407.

619
620 46.Richardson, A.J., Walne, A.W., John, A.W.G., Jonas, T.D., Lindley, J. a., Sims, D.W., *et*
621 *al.* (2006). Using continuous plankton recorder data. *Prog. Oceanogr.*, 68, 27–74.

622
623 47.Rue, H., Martino, S. & Chopin, N. (2009). Approximate Bayesian inference for latent
624 Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B*
625 *(Statistical Methodol.*, 71, 319–392.

626
627 48.San Martin, E., Harris, R.P. & Irigoien, X. (2006). Latitudinal variation in plankton size
628 spectra in the Atlantic Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, 53, 1560–1572.

629
630 49.Scheiter, S., Langan, L. & Higgins, S.I. (2013). Next-generation dynamic global
631 vegetation models: learning from community ecology. *New Phytol.*, 198, 957–969.

632
633 50.Segers, F.H.I.D. & Taborsky, B. (2011). Egg size and food abundance interactively affect
634 juvenile growth and behaviour. *Funct. Ecol.*, 25, 166–176.

635
636 51.Stamieszkin, K., Pershing, A.J., Record, N.R., Pilskaln, C.H., Dam, H.G. & Feinberg,
637 L.R. (2015). Size as the master trait in modeled copepod fecal pellet carbon flux. *Limnol.*
638 *Oceanogr.*, 60, 2090–2107.

639
640 52.Turner, J. (2002). Zooplankton fecal pellets, marine snow and sinking phytoplankton
641 blooms. *Aquat. Microb. Ecol.*, 27, 57–102.

642
643 53.Verity, P. & Smetacek, V. (1996). Organism life cycles, predation, and the structure of
644 marine pelagic ecosystems. *Mar. Ecol. Prog. Ser.*, 130, 277–293.

645
646 54.Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A. & Wright, I.J. (2002). Plant
647 Ecological Strategies: Some Leading Dimensions of Variation Between Species. *Annu. Rev.*
648 *Ecol. Syst.*, 33, 125–159.

55.Wood, S. (2006). *Generalized Additive Models: An Introduction with R*. CRC Press, Boca Raton, Florida.

Tables

Table 1: Trait data coverage for taxa included in observational datasets: covered fractions of taxonomic diversity and biomass/abundance are shown for the North Atlantic and the global ocean. Biomass fractions could be estimated for the North Atlantic using cubed total length as mass proxies, since data on total length was available for all taxa. For the global ocean this was not the case and we therefore report percentages of abundance (number of individuals). North Atlantic data stems from the Continuous Plankton Recorder; global data stems from the Coastal and Oceanic Plankton Ecology, Production and Observation Database.

Trait	North Atlantic (67 taxa)		Global (607 taxa)	
	Diversity	Biomass	Diversity	Abundance
Body size	100%	100%	95%	99%
Feeding mode	99%	100%	78%	96%
Myelination	100%	100%	100%	100%
Relative offspring size	55%	99%	23%	70%

Figure captions

Figure 1: Fraction of variance explained by INLA models for each trait based on spatial/spatiotemporal predictors (red), environmental predictors (green), and both types of predictors (yellow). Results are shown for global models (left panels) and North Atlantic models (right panels). Combined and environmental models for the North Atlantic were run on a subset of the observations used for the spatiotemporal models due to missing environmental data (satellite observations during winter months). R^2 of spatiotemporal models can thus be slightly higher than corresponding R^2 combined models.

Figure 2: Seasonal succession of community traits in the North Atlantic 1998-2008. Estimated trait distributions are shown for the beginning of January, April, July, and October (columns) for body size, body-size diversity, feeding modes, myelination and relative offspring size (columns). Displayed are only pixels with a maximum distance of 400 kilometers from observations in every season. Estimates of spatial and temporal autocorrelation of trait distributions in the North Atlantic are shown in Appendix F.

Figure 3: Global distributions of community mean traits for body size (a), myelination (b), and relative offspring size (c). Polygons on the maps represent simulated communities. Colored polygons are data-based estimates; polygons in gray scales are predictions with the best environmental models. The panels on the right show trait distributions per latitude. Median model predictions (lines) and 90% confidence intervals (polygons) are shown in grey. Data-based trait patterns are superimposed in orange, including median (circles), inter quartile range (thick lines), and 90% confidence intervals (thin lines). Global maps for further traits can be seen in Appendix E. Estimates of spatial autocorrelation lengths of global trait distributions are shown in Appendix F.

Figure 4: Responses of trait distributions to environmental predictors of hypothetical importance based on single-predictor models. Traits include body size, body-size diversity, myelinated fraction, and relative offspring size (rows). Responses for fractional traits are shown on the logit scale. Environmental predictors are net primary production (left row), phytoplankton cell diameter (second row from left), sea surface temperature (second row from right), seasonality of chlorophyll *a* concentration (right row top), and Secchi Depth (right row bottom). Lines in dark blue represent global models, lines in cyan represent North Atlantic models. Shaded areas surrounding the lines illustrate 95% confidence intervals. Dashed lines represent predictors not included in the best environmental models of the corresponding trait and domain. Responses for active feeding are shown in Appendix H.

Figures

Figure 1

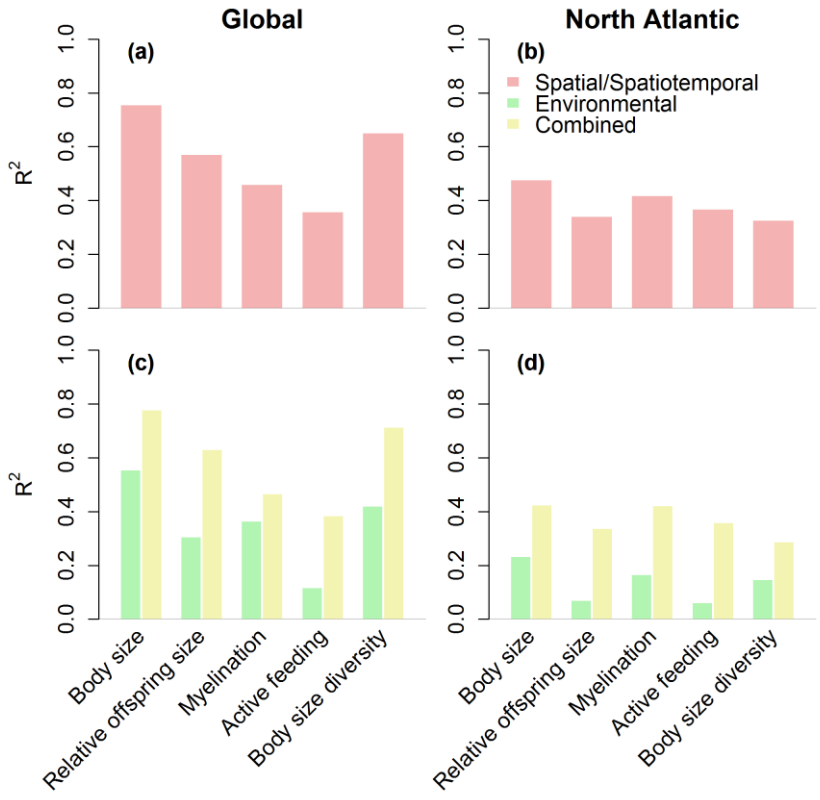


Figure 2

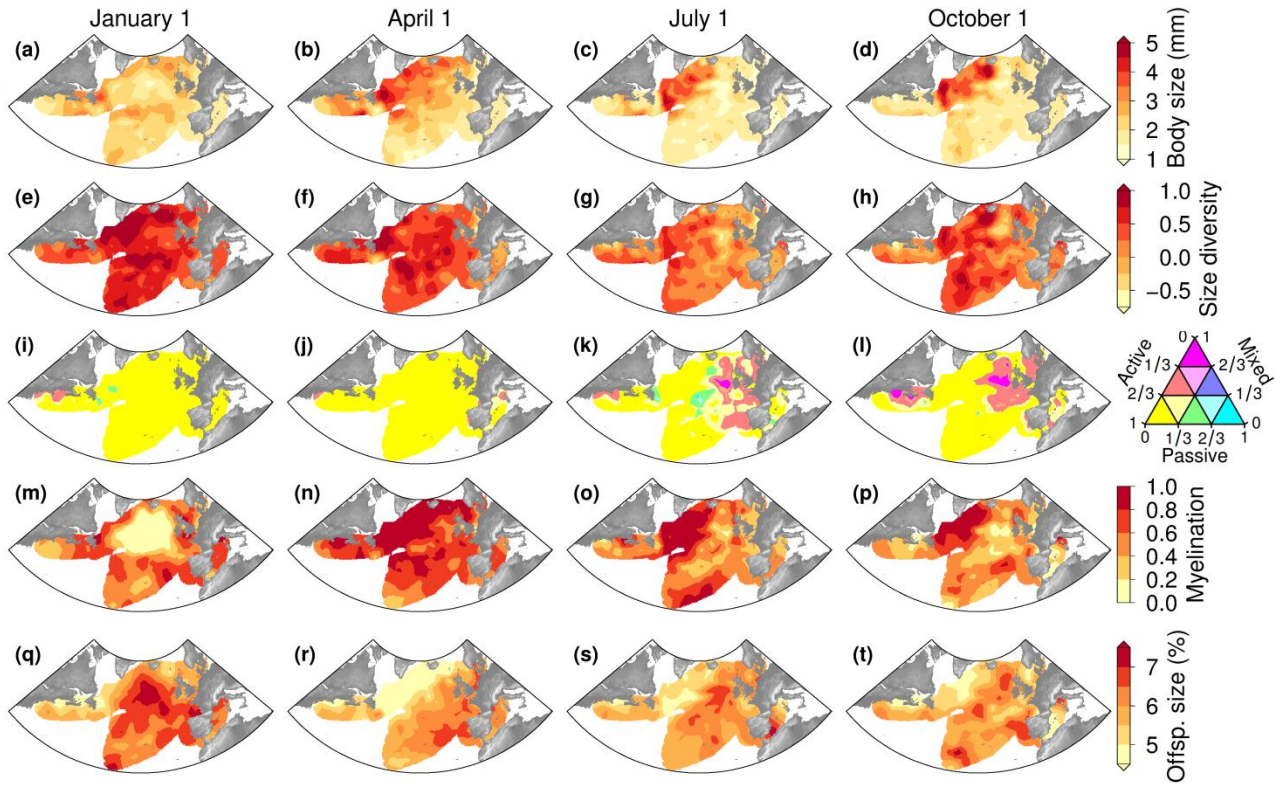

January 1
April 1
July 1
October 1
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
(p)
(q)
(r)
(s)
(t)
Body size (mm)
5
4
3
2
1
Size diversity
1.0
0.5
0.0
−0.5
Active
Mixed
Passive
Myelination
1.0
0.8
0.6
0.4
0.2
0.0
Offsp. size (%)
7
6
5

Figure 3

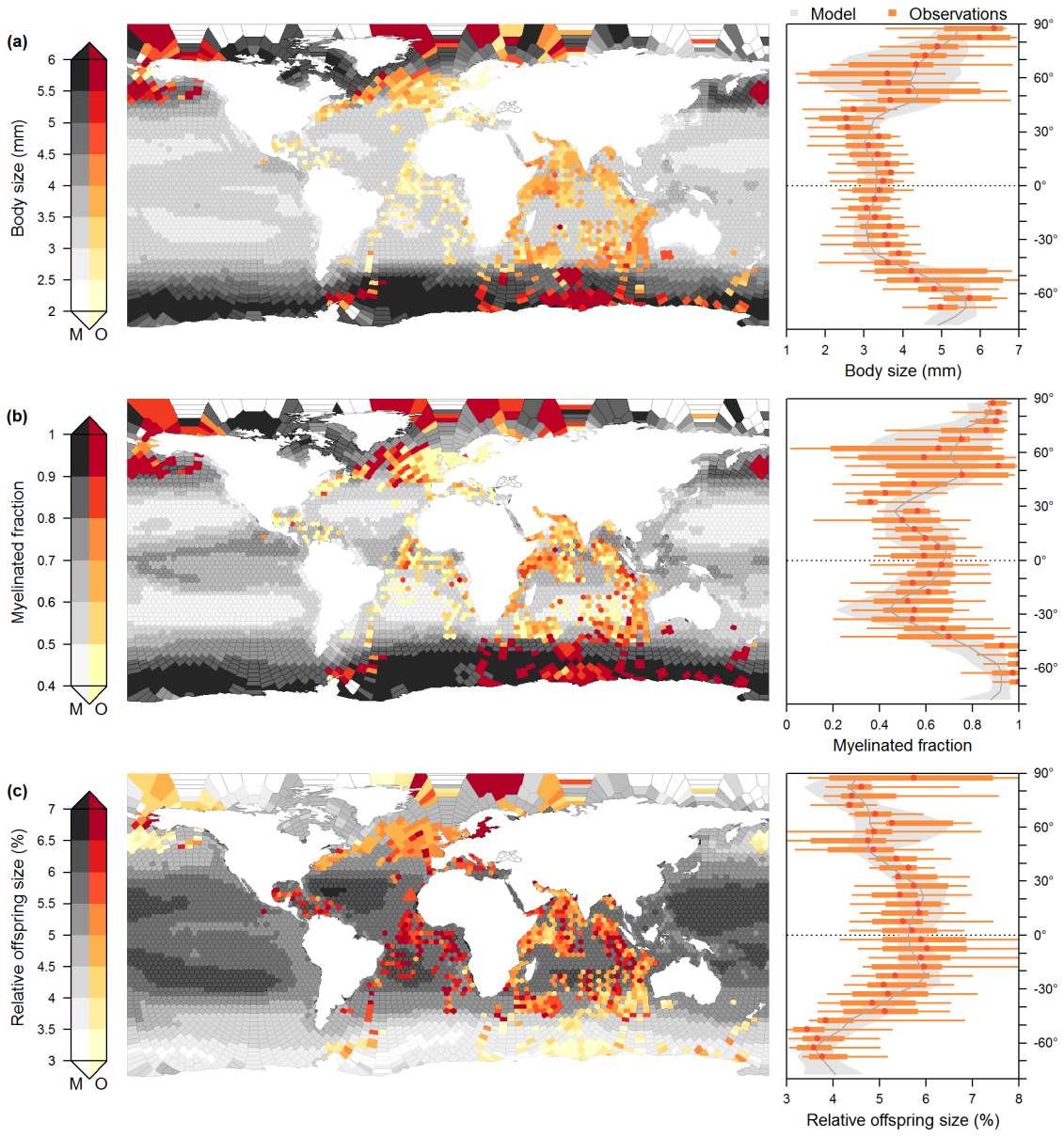
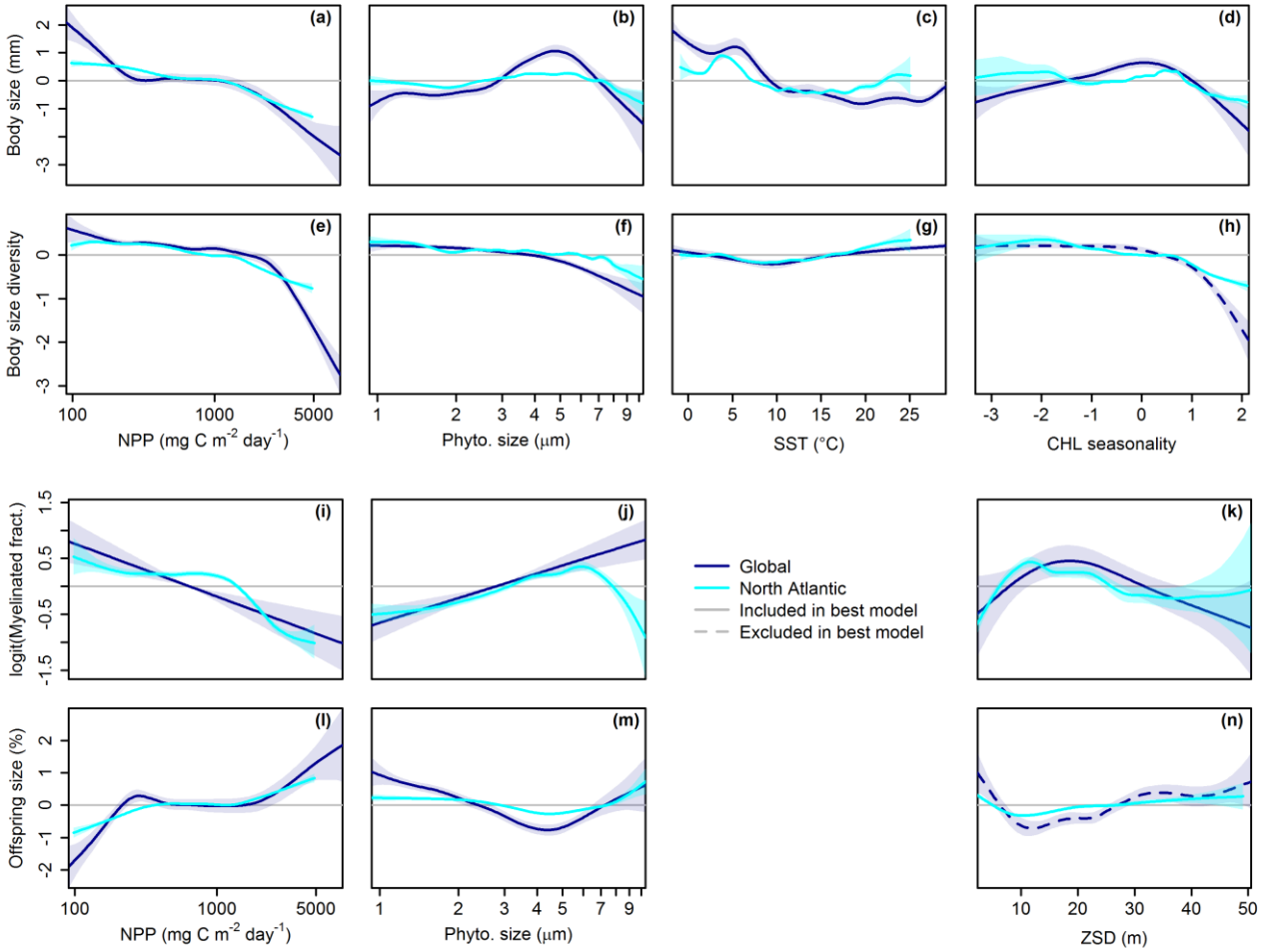

(a)
Body size (mm)
6
5.5
5
4.5
4
3.5
3
2.5
2
M O
Model
Observations
90°
60°
30°
0°
-30°
-60°
1 2 3 4 5 6 7
Body size (mm)
(b)
Myelinated fraction
1
0.9
0.8
0.7
0.6
0.5
0.4
M O
0 0.2 0.4 0.6 0.8 1
Myelinated fraction
(c)
Relative offspring size (%)
7
6.5
6
5.5
5
4.5
4
3.5
3
M O
3 4 5 6 7 8
Relative offspring size (%)

Figure 4


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
Body size (mm)
Body size diversity
logit(Myelinated fract.)
Offspring size (%)
NPP (mg C m-2 day-1)
Phyto. size (µm)
SST (°C)
CHL seasonality
ZSD (m)
Global
North Atlantic
Included in best model
Excluded in best model

Text boxes

Box1: Traits considered and their hypothesized dependence on the environment

Body size

Body size is a master trait affecting all major life missions of an organism, i.e., feeding, survival, and reproduction (Litchman *et al.* 2013). It can be considered a proxy for several other essential properties such as most vital rates, mobility, and prey size. Here, body size is represented by the total length of adults. We hypothesize that mean body size in copepod communities decreases with increasing temperatures. Such a relationship is known to occur within copepod species, potentially due to oxygen limitation of large organisms at warm temperatures (Forster *et al.* 2012). Furthermore, we expect copepod body size to be positively correlated to productivity, as has been shown for many animal groups (Huston & Wolverton 2011). Larger body size has also been shown to be beneficial for copepods to cope with seasonal environments (Maps *et al.* 2014), and we thus expect body size to be positively related to the intensity of the seasonal cycle. Finally, we hypothesize that copepod body size is positively related to the size of the local prey, as feeding efficiency in copepods is a function of the predator to prey size ratio (Hansen *et al.* 1994).

Feeding mode

We distinguish between three different feeding modes: passive feeding, active feeding, and mixed feeding (Kiørboe 2011). Passive feeding includes mainly ambush feeding but also particle feeding copepods. The former copepods wait for prey to pass within their perceptive range, while the latter feed on large particles of marine snow. Active strategies comprise cruise feeding and feeding current feeding, where the copepod either moves

732 through the water or generates a feeding current. Most taxa exclusively use either an active or
733 a passive feeding behavior, but some taxa are able to alternate (called mixed feeders in this
734 paper). Ambush feeders rely on motile prey for feeding and therefore we hypothesize that
735 passive feeders are more common in areas with more motile phytoplankton like flagellates.
736 Furthermore, we expect passive feeders to be less common in unproductive areas as they
737 have lower feeding rates (Kiørboe 2013) and may struggle more with low prey
738 concentrations. Lastly, we hypothesize mixed feeding to be a trait that is beneficial in
739 seasonal environments with varying prey types and concentrations.

740 Relative offspring size

741 Some copepod species have relatively larger (and fewer) eggs than others, suggesting
742 differences in the investment made per offspring. We estimate these differences as relative
743 offspring size, the ratio between egg diameter and the length of the adult female. We do not
744 study absolute egg diameters here, as they scale positively with body size (Neuheimer *et al.*
745 2015): according to our data the corresponding Pearson correlation coefficient is r=0.84
746 (n=166), while r for relative offspring size versus body size is -0.19 (n=164). We expect large
747 relative offspring size to be beneficial in harsh environments (Segers & Taborsky 2011) with
748 low productivity, low quality of food but also low predation pressure.

749 Myelination

750 Copepods can be grouped into myelinated and amyelinated taxa (Lenz 2012). Myelin
751 is a membranous sheath that surrounds the axons of neurons and greatly enhances the speed
752 of signal transmission. Myelinated copepods are more efficient in escaping predators and
753 need less energy to maintain their nervous systems, but they rely on a more lipid-rich diet

754 (Lenz 2012). We hypothesize that myelination to common in areas where predation pressure
755 is high, where productivity is low, and where food quality is high (Lenz 2012).

Appendix A: CPR taxa considered

CPR taxa considered in the North Atlantic copepod community and species, based on which traits were estimated.

CPR taxon	Species considered for trait estimate
Acartia spp. (unidentified)[a]	*A. clausi*
Acartia danae	*A. danae*
Acartia longiremis	*A. longiremis*
Aetideus armatus	*A. armatus*
Anomalocera patersoni	*A. patersoni*
Calanoides carinatus	*C. carinatus*
Calanus finmarchicus	*C. finmarchicus*
Calanus glacialis	*C. glacialis*
Calanus helgolandicus	*C. helgolandicus*
Calanus hyperboreus	*C. hyperboreus*
Calocalanus spp.[b]	*C. contractus, C. pavo, C. plumulosus, C. styliremis, C. tenuis*
Candacia armata	*C. armata*
Candacia ethiopica	*C. ethiopica*
Candacia pachydactyla	*C. pachydactyla*
Paracandacia simplex	*C. simplex*
Centropages bradyi	*C. bradyi*
Centropages chierchiae eyecount	*C. chierchiae*
Centropages hamatus	*C. hamatus*
Centropages typicus	*C. typicus*
Centropages violaceus	*C. violaceus*
Clausocalanus spp.[b]	*C. arcuicornis, C. furcatus, C. paululus, C. pergens*
Corycaeus spp.[a,b]	*C. speciosus, Ditrichocorycaeus anglicus*
Ctenocalanus vanus	*C. vanus*
Eucalanus spp.[b] (Unidentified)	*E. elongatus, Pareucalanus attenuatus*
Eucalanus hyalinus	*E. hyalinus*
Euchaeta acuta	*E. acuta*
Euchaeta marina	*E. marina*
Euchirella rostrata	*E. rostrata*
Heterorhabdus norvegicus	*H. norvegicus*
Heterorhabdus papilliger	*H. papilliger*
Isias clavipes	*I. clavipes*
Labidocera spp.[b] (Unidentified)	*L. acutifrons, L. aestiva, L. wollastoni*
Lucicutia spp.[a]	*L. flavicornis*
Mecynocera clausi	*M. clausi*
Mesocalanus tenuicornis	*M. tenuicornis*

Metridia longa	*M. longa*
Metridia lucens	*M. lucens*
Harpacticoida Total Traverse[a,b]	*Microsetella norvegica, Microsetella rosea*
Nannocalanus minor	*N. minor*
Neocalanus gracilis	*N. gracilis*
Oithona spp.[b]	*O. atlantica, O. linearis, O. nana, O. plumifera, O. robusta, O. setigera, O. similis*
Oncaea spp.[b]	*O. media, O. mediterranea, O. ornata, O. venusta*
Para-Pseudocalanus spp.[b]	*Paracalanus parvus, Pseudocalanus elongatus, Pseudocalanus minutus*
Paracandacia bispinosa	*P. bispinosa*
Paraeuchaeta gracilis	*P. gracilis*
Paraeuchaeta hebes	*P. hebes*
Paraeuchaeta norvegica	*P. norvegica*
Parapontella brevicornis	*P. brevicornis*
Pleuromamma abdominalis	*P. abdominalis, P. indica*
Pleuromamma borealis	*P. borealis*
Pleuromamma gracilis	*P. gracilis*
Pleuromamma piseki	*P. piseki*
Pleuromamma robusta	*P. robusta*
Pleuromamma xiphias	*P. xiphias*
Pontellina plumata	*P. plumata*
Scolecithricella spp.[b]	*P. ovata, S. dentata, S. minor, S. vittata*
Rhincalanus nasutus	*R. nasutus*
Scolecithrix danae	*S. danae*
Subeucalanus crassus	*S. crassus*
Subeucalanus monachus	*S. monachus*
Temora longicornis	*T. longicornis*
Temora stylifera	*T. stylifera*
Tortanus discaudatus	*T. discaudatus*
Undeuchaeta major	*U. major*
Undeuchaeta plumosa	*U. plumosa*
Undinula vulgaris	*U.vulgaris*
Urocorycaeus spp.[b]	*U. furcifer, U. lautus, U. longistylis*

759 [a]Most common species in taxon according to (Richardson *et al.* 2006) was considered for trait information.

760 [b]Trait estimates for genus based on arithmetic mean of species common in the North Atlantic according to

761 www.iobis.org.

762

Appendix B: Correlation analysis of environmental variables

Pearson correlation coefficients between all pairs of environmental predictors used: values in italic indicate correlation coefficients for observations in the North Atlantic; non-italic values indicate values on the global scale. Grey color represents variable combinations which are never used in the models (ZSD and CHL seasonality). Fields highlighted in yellow represent combinations used in the models with correlation coefficients higher than 0.7.

	SST[a]	**ZSD**[b]	**NPP**[c]	**CHL seasonality**[d]	**MD_{50}**[e]
SST	1	0.47	-0.06	-0.52	-0.86
	1	*0.48*	*-0.15*	*-0.49*	*-0.58*
ZSD	0.47	1	-0.78	-0.92	-0.82
	0.48	*1*	*-061*	*-0.6*	*-0.79*
NPP	-0.06	-0.78	1	0.77	0.5
	-0.15	*-0.61*	*1*	*0.37*	*0.4*
CHL seasonality	-0.52	-0.92	0.77	1	0.86
	-0.49	*-0.6*	*0.37*	*1*	*0.59*
MD_{50}	-0.86	-0.82	0.5	0.86	1
	-0.58	*-0.79*	*0.42*	*0.59*	*1*

[a]Sea surface temperature; [b]Secchi Depth; [c]net primary productivity; [d]seasonality in chlorophyll *a* concentrations; [e]median diameter of phytoplankton cells

Appendix C: Spatial and temporal meshes for INLA

North Atlantic

Models for the North Atlantic were constructed including both, a spatial and a seasonal mesh. The spatial mesh covered the North Atlantic and was constrained by the coastlines (islands with an area smaller than 100 000 km^2 were ignored). The maximum distance between mesh points was chosen to be about 300 km (Figure C1). The seasonal mesh had nodes at the beginning of January, April, July, and October and was cyclic at its boundaries.

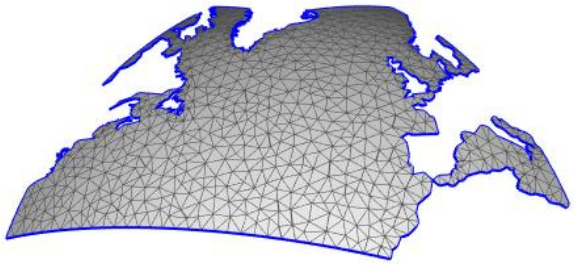


Figure C1: Delaunay triangulation of the North Atlantic domain. Points (intersections) of the field are used to estimate the spatial dependencies in INLA models. We projected the coordinates onto a sphere in order to realistically represent the spatial distances.

Global

Spatial models of global trait distributions were modeled based on a spherical, global mesh defined with a maximum distance of about 500 km between the points and constrained by coarse continental borders (again, islands with an area smaller than 100 000 km^2 were ignored) (Figure C2).

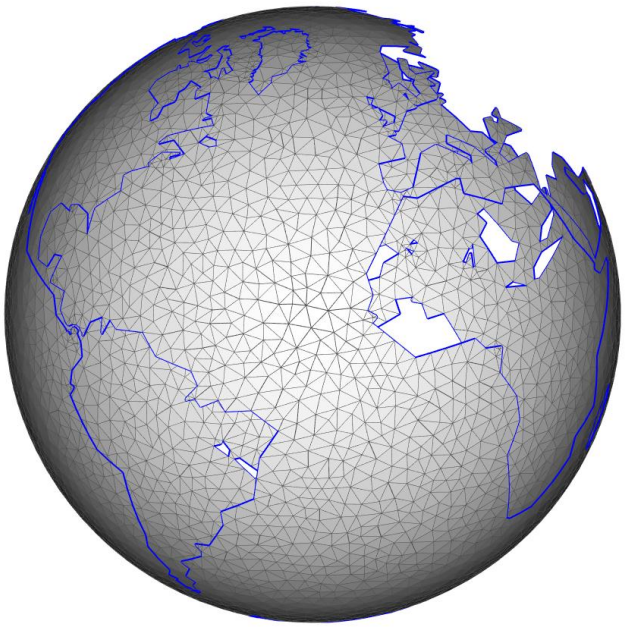


Figure C2: Delaunay triangulation of the global domain. Points (intersections) of the field are used to estimate the spatial dependencies in INLA models. We projected the coordinates onto a sphere in order to realistically represent the spatial distances.

Appendix D: Verification of the existence of between-community trait variance

We found clear variation between communities in all traits of both the North Atlantic and the global domain. The existence of variation was assessed using a bootstrapping approach on the variance of the summary statistics (see Methods). We tested whether the variance among communities of these summary statistics differed from zero. To this end we resampled each summary statistic in of both domains 1000 times with replacement. For each of these 1000 pseudo-samples of communities we then calculated the variance. The histograms for these variances are shown in Figure D1. For all traits and both domains we could clearly confirm our hypothesis that a significant variation of traits exists between copepod communities.

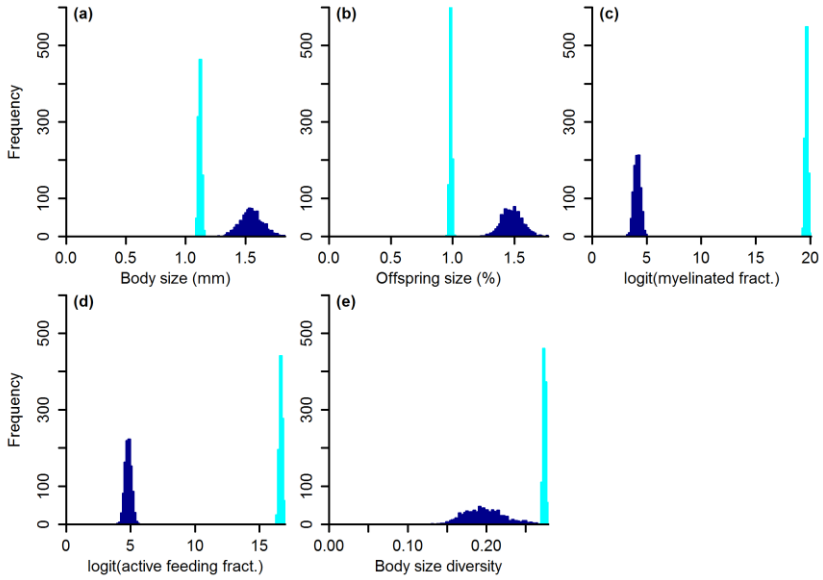
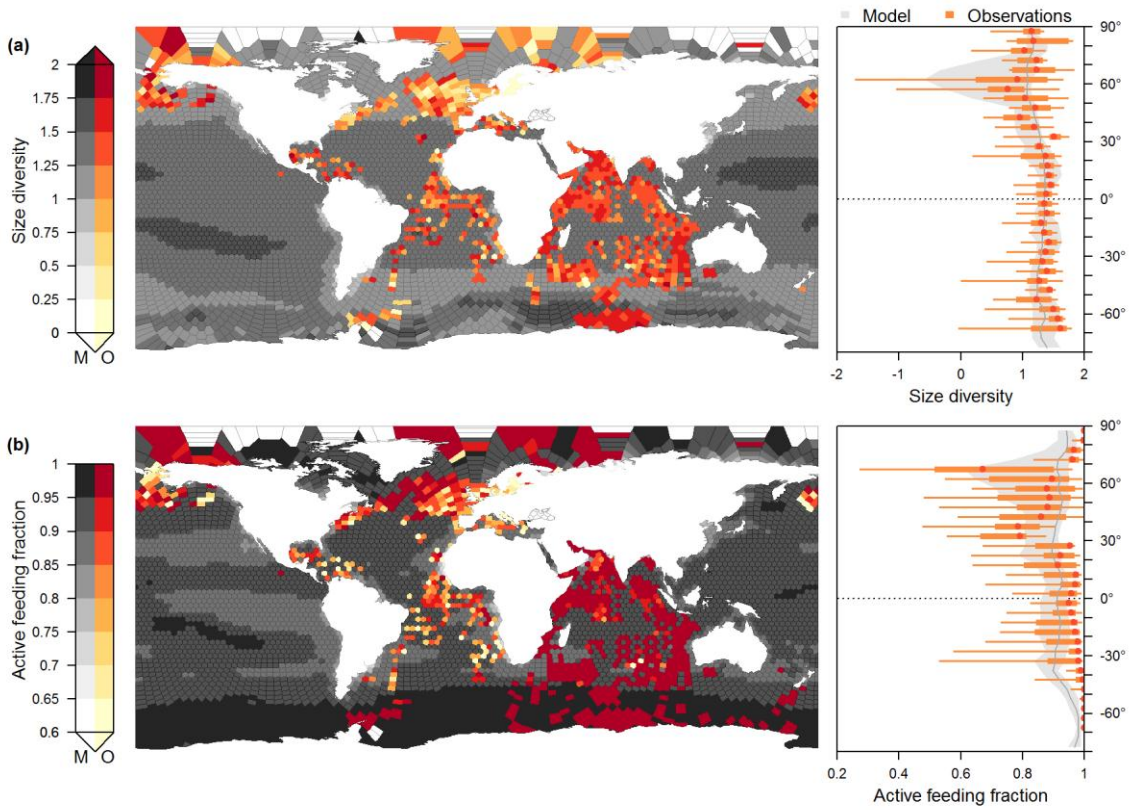


Figure D1: Histograms of standard deviations for body size (a), relative offspring size (b), the logit transformed fraction of myelinated copepods (c), the logit transformed fraction of active feeding copepods (d), and body-size diversity (e). Variance estimates for the North

812 Atlantic domain are shown in cyan and variance estimates for global domain are shown in
813 dark blue.

814

Appendix E: Further global traits



Global distributions of community mean traits for body-size diversity (a) and active feeding (b). Polygons on the maps represent simulated communities. Colored polygons are data-based estimates; polygons in gray scales are predictions with the best environmental models. The panels on the right show latitudinal trait variation. Median model predictions (lines) and 90% confidence intervals (polygons) are shown in grey. Data-based trait patterns are superimposed in orange, including median (circles), inter quartile range (thick lines), and 90% confidence intervals (thin lines).

Appendix F: Spatial and temporal correlations

Table F1: Spatial and temporal autocorrelation of trait distributions in the North Atlantic obtained from spatiotemporal models. Depicted are means and standard deviations. Temporal autocorrelation is defined as Pearson correlation coefficients between subsequent seasons; spatial autocorrelation length is defined as the distance at which the Pearson correlation coefficients between points fall below about 0.13.

Trait	Temporal autocorrelation (between seasons)	Spatial autocorrelation length (km)
Body size	0.511 ± 0.054	810 ± 87
Relative offspring size	0.277 ± 0.082	1017 ± 85
Myelination	0.243 ± 0.073	998 ± 90
Active feeding	0.406 ± 0.069	1074± 127
Mixed feeding	0.522 ± 0.066	970 ± 88
Passive feeding	0.153 ± 0.085	675 ± 83
Body-size diversity	0.250 ± 0.074	634 ± 6

Table F2: Spatial autocorrelation length of trait distributions in the global ocean obtained from spatial models. Depicted are means and standard deviations. Spatial autocorrelation length is defined as the distance at which the Pearson correlation coefficients between points fall below about 0.13.

Trait	Spatial autocorrelation length (km)
Body size	5575 ± 1286
Relative offspring size	4117 ± 787
Myelination	30 745 ± 22 955

Active feeding	2549± 5
Body-size diversity	1721 ± 316

836

837

Appendix G: Skill of environmental models with all predictor combinations

Table G1: Model skill in terms of deviance information criterion (DIC), Wanatabe-Akaike information criterion (WAIC), and explained variance (R^2) of global environmental models. Best models for each trait are highlighted in yellow.

Response	Predictors	DIC	WAIC	R^2	Best model
Feeding_mode.Active		521.80	521.01		0
Feeding_mode.Active	diverCHL	520.73	519.18	0.02	0
Feeding_mode.Active	meanNPP	507.63	505.99	0.11	0
Feeding_mode.Active	medianPhyto	523.12	521.56	0.00	0
Feeding_mode.Active	diverCHL & medianPhyto	521.52	519.13	0.03	0
Feeding_mode.Active	meanNPP & diverCHL	502.49	500.07	0.13	1
Feeding_mode.Active	meanNPP & medianPhyto	507.36	504.93	0.10	0
Feeding_mode.Active	meanNPP & diverCHL & medianPhyto	503.62	500.35	0.14	0
Myelination		1103.57	1102.82		0
Myelination	meanNPP	1088.48	1086.95	0.08	0
Myelination	meanZSD	1087.71	1084.27	0.12	0
Myelination	medianPhyto	1083.23	1081.79	0.11	0
Myelination	meanNPP & medianPhyto	1029.80	1027.42	0.31	0
Myelination	meanZSD & meanNPP	1024.59	1022.14	0.34	0
Myelination	meanZSD & medianPhyto	1048.60	1044.45	0.26	0
Myelination	meanZSD & meanNPP & medianPhyto	1019.67	1016.37	0.36	1

OffspringSize		2652.67	2655.54		0
OffspringSize	meanNPP	2575.61	2574.39	0.11	0
OffspringSize	meanZSD	2563.92	2563.02	0.12	0
OffspringSize	medianPhyto	2450.52	2452.46	0.22	0
OffspringSize	meanNPP & medianPhyto	2325.52	2328.54	0.33	1
OffspringSize	meanZSD & meanNPP	2380.24	2380.92	0.29	0
OffspringSize	meanZSD & medianPhyto	2347.13	2349.12	0.32	0
OffspringSize	meanZSD & meanNPP & medianPhyto	2331.31	2331.70	0.33	0
Size		2748.86	2749.15		0
Size	diverCHL	2663.16	2667.00	0.10	0
Size	meanNPP	2621.78	2621.75	0.15	0
Size	meanSST	2316.70	2324.12	0.41	0
Size	medianPhyto	2530.59	2533.88	0.24	0
Size	diverCHL & medianPhyto	2363.88	2367.20	0.38	0
Size	meanNPP & diverCHL	2294.15	2295.89	0.42	0
Size	meanNPP & medianPhyto	2265.79	2266.23	0.44	0
Size	meanSST & diverCHL	2197.55	2203.25	0.50	0
Size	meanSST & meanNPP	2160.57	2168.47	0.52	0
Size	meanSST & medianPhyto	2174.24	2182.39	0.51	0
Size	meanNPP & diverCHL & medianPhyto	2241.91	2242.00	0.46	0
Size	meanSST & diverCHL & medianPhyto	2134.15	2145.48	0.53	0
Size	meanSST & meanNPP & diverCHL	2147.14	2156.90	0.52	0
Size	meanSST & meanNPP & medianPhyto	2130.55	2142.20	0.54	0

Size	meanSST & meanNPP & diverCHL & medianPhyto	2089.48	2106.09	0.56	1
Size_diversity		988.22	995.21		0
Size_diversity	diverCHL	756.29	770.96	0.27	0
Size_diversity	meanNPP	624.68	631.16	0.38	0
Size_diversity	meanSST	911.16	923.45	0.11	0
Size_diversity	medianPhyto	855.45	867.05	0.16	0
Size_diversity	diverCHL & medianPhyto	751.58	761.19	0.27	0
Size_diversity	meanNPP & diverCHL	623.02	630.48	0.39	0
Size_diversity	meanNPP & medianPhyto	596.43	610.23	0.41	0
Size_diversity	meanSST & diverCHL	721.89	736.67	0.31	0
Size_diversity	meanSST & meanNPP	594.31	602.39	0.41	0
Size_diversity	meanSST & medianPhyto	721.33	732.50	0.31	0
Size_diversity	meanNPP & diverCHL & medianPhyto	588.82	599.09	0.42	0
Size_diversity	meanSST & diverCHL & medianPhyto	680.14	697.85	0.35	0
Size_diversity	meanSST & meanNPP & diverCHL	597.90	605.54	0.41	0
Size_diversity	meanSST & meanNPP & medianPhyto	581.59	595.75	0.43	1
Size_diversity	meanSST & meanNPP & diverCHL & medianPhyto	582.21	596.36	0.43	0

843

844 Table G2: Model skill in terms of deviance information criterion (DIC), Wanatabe-
845 Akaike information criterion (WAIC), and explained variance (R^2) of North Atlantic
846 environmental models. Best models for each trait are highlighted in yellow.

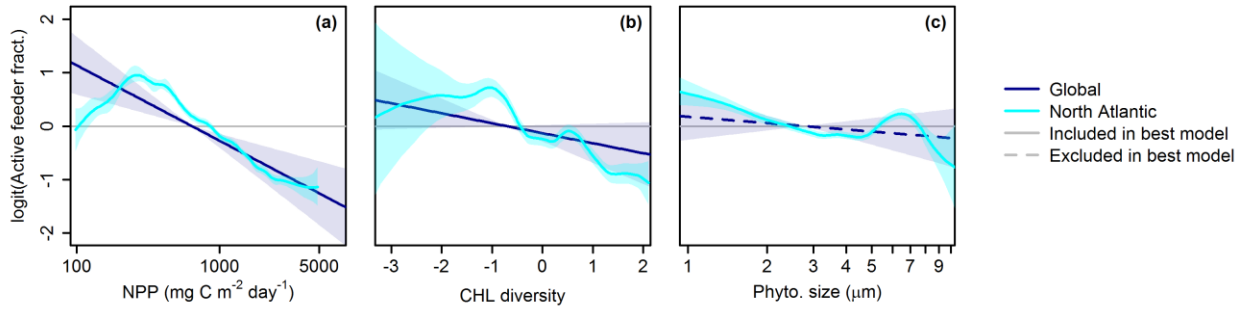
Response	Predictors	DIC	WAIC	R^2	Best model
Feeding_mode.Active		215857	215863	0.00	0
Feeding_mode.Active	Diver_CHL	210778	210784	0.01	0
Feeding_mode.Active	NPP	208409	208410	0.02	0
Feeding_mode.Active	Phyto_size	211310	211312	0.01	0
Feeding_mode.Active	Diver_CHL & Phyto_size	210529	210536	0.04	0
Feeding_mode.Active	NPP & Diver_CHL	208143	208149	0.02	0
Feeding_mode.Active	NPP & Phyto_size	207843	207845	0.04	0
Feeding_mode.Active	NPP & Diver_CHL & Phyto_size	207459	207469	0.06	1
Myelination		242754	242757	0.00	0
Myelination	NPP	241690	241692	0.07	0
Myelination	Phyto_size	242291	242294	0.01	0
Myelination	ZSD	242179	242183	0.04	0
Myelination	NPP & Phyto_size	240331	240334	0.11	0
Myelination	NPP & ZSD	241302	241306	0.08	0
Myelination	ZSD & Phyto_size	240022	240027	0.14	0
Myelination	NPP & ZSD & Phyto_size	239348	239353	0.16	1
OffspringSize		86733	86734	0.00	0
OffspringSize	NPP	85972	85972	0.03	0
OffspringSize	Phyto_size	86061	86062	0.02	0

OffspringSize	ZSD	86157	86159	0.02	0
OffspringSize	NPP & Phyto_size	84842	84841	0.06	0
OffspringSize	NPP & ZSD	85256	85257	0.05	0
OffspringSize	ZSD & Phyto_size	85196	85197	0.05	0
OffspringSize	NPP & ZSD & Phyto_size	84145	84147	0.09	1
Size		97476	97478	0.00	0
Size	Diver_CHL	92815	92823	0.04	0
Size	NPP	94444	94444	0.08	0
Size	Phyto_size	93403	93409	0.03	0
Size	SST	90243	90251	0.11	0
Size	Diver_CHL & Phyto_size	95434	95435	0.06	0
Size	NPP & Diver_CHL	92736	92735	0.12	0
Size	NPP & Phyto_size	91645	91645	0.15	0
Size	NPP & SST	89445	89444	0.21	0
Size	SST & Diver_CHL	92424	92424	0.13	0
Size	SST & Phyto_size	89597	89612	0.13	0
Size	NPP & Diver_CHL & Phyto_size	91088	91086	0.17	0
Size	NPP & SST & Diver_CHL	89219	89216	0.21	0
Size	NPP & SST & Phyto_size	84696	84736	0.23	0
Size	SST & Diver_CHL & Phyto_size	92156	92155	0.14	0
Size	NPP & SST & Diver_CHL & Phyto_size	84477	84485	0.23	1
Size_diversity		49562	49559	0.01	0
Size_diversity	Diver_CHL	48154	48157	0.05	0
Size_diversity	NPP	45518	45513	0.13	0

Size_diversity	Phyto_size	49191	49188	0.02	0
Size_diversity	SST	48973	48974	0.03	0
Size_diversity	Diver_CHL & Phyto_size	48086	48086	0.05	0
Size_diversity	NPP & Diver_CHL	45267	45263	0.13	0
Size_diversity	NPP & Phyto_size	45295	45291	0.13	0
Size_diversity	NPP & SST	45379	45375	0.13	0
Size_diversity	SST & Diver_CHL	47922	47921	0.06	0
Size_diversity	SST & Phyto_size	48662	48671	0.04	0
Size_diversity	NPP & Diver_CHL & Phyto_size	44943	44943	0.14	0
Size_diversity	NPP & SST & Diver_CHL	45147	45144	0.14	0
Size_diversity	NPP & SST & Phyto_size	45171	45168	0.14	0
Size_diversity	SST & Diver_CHL & Phyto_size	47851	47846	0.06	0
Size_diversity	NPP & SST & Diver_CHL & Phyto_size	44855	44857	0.15	1

847

Appendix H: Environmental responses of active feeding



Responses of active feeding to environmental predictors of hypothetical importance, based on single-predictor models. Responses are shown on the logit scale. Environmental predictors are net primary production, seasonality of chlorophyll *a* concentration, and phytoplankton cell diameter (columns). Lines in dark blue represent global models, lines in cyan represent North Atlantic models. Shaded areas surrounding the lines illustrate 95% confidence intervals. Dashed lines represent predictors not included in the best models of the corresponding trait and domain.