



HAL
open science

A multivariate approach to large-scale variation in copepod diversity and environmental descriptors

I. Rombouts, Gregory Beaugrand, F. Ibanez, S. Chiba, L. Legendre

► To cite this version:

I. Rombouts, Gregory Beaugrand, F. Ibanez, S. Chiba, L. Legendre. A multivariate approach to large-scale variation in copepod diversity and environmental descriptors. *Limnology and Oceanography*, 2010, 55 (5), pp.2219-2229. hal-00767796

HAL Id: hal-00767796

<https://hal.science/hal-00767796v1>

Submitted on 26 Sep 2024

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.



Distributed under a Creative Commons Attribution 4.0 International License

A multivariate approach to large-scale variation in marine planktonic copepod diversity and its environmental correlates

Isabelle Rombouts,^{a,b,c,*} Grégory Beaugrand,^c Frédéric Ibañez,^{a,b} Stéphane Gasparini,^{a,b} Sanae Chiba,^d and Louis Legendre^{a,b}

^a Université Pierre et Marie Curie, UMR 7093, Laboratoire d'Océanographie de Villefranche, Villefranche-sur-Mer, France

^b Centre National de la Recherche Scientifique (CNRS), UMR 7093, Laboratoire d'Océanographie de Villefranche, Villefranche-sur-Mer, France

^c Centre National de la Recherche Scientifique (CNRS), Université des Sciences et Technologies de Lille—Lille 1, Laboratoire d'Océanologie et de Géosciences (LOG), Wimereux, France

^d Frontier Research Center for Global Change/Japan Agency for Marine–Earth Science and Technology (JAMSTEC), Kanazawa-ku, Yokohama, Kanagawa, Japan

Abstract

We have investigated the relationships between covariations in environmental variables and variations in distributions of marine copepod diversity over an extensive latitudinal range from 86.5°N to 46.5°S. For this purpose, 7 data sets (representing 13,713 samples) of copepod species composition data and 11 environmental data sets were assembled. Principal components analysis was applied to investigate the relationships among the mean and seasonal variations in environmental descriptors (ocean temperature, chlorophyll *a* [Chl *a*], net primary production, and other physical and chemical properties of the ocean) and their relationships with spatial variations in copepod diversity. High copepod diversity corresponded to a combination of high ocean temperature and salinity and low Chl *a* and nutrient concentrations (nitrate, silicate, phosphate). To a lesser extent, high-diversity regions were also correlated to low seasonal variability in oxygen, ocean temperature, and mixed-layer depth. Regression on principal components provided a robust prediction of global copepod diversity ($R^2 = 0.45$, $p < 0.001$) as our subset of environmental data was representative of the full range of environmental variability that occurs globally.

The study of latitudinal patterns in species diversity can help to resolve the underlying mechanisms that regulate the distributions of organisms (Roy et al. 1998). Strong evidence suggests that contemporary environmental descriptors drive broadscale gradients of species richness in terrestrial (Hawkins et al. 2003; Field et al. 2008) and marine (Roy et al. 1998) environments. In particular, energy inputs in terms of photosynthetically active solar radiation and temperature appear to be important for species diversity (Hessen et al. 2007). In the context of climate change, it is timely to identify the drivers of diversity in key marine trophic groups to help anticipate the future responses of marine ecosystems to the changing environment. Pelagic copepods are a key trophic group in the marine plankton, where they play important roles in both the transfer of energy from primary producers to higher trophic levels and biogeochemical cycles (Roemmich and McGowan 1995; Beaugrand 2009). These ectotherms are also particularly suitable for studies of biogeographical distributions in relation to environmental conditions. Indeed, because of the relatively short life spans and therefore rapid population overturn of copepods, tight coupling exists between environmental change and plankton community dynamics, e.g., in the North Atlantic (Reid and Beaugrand 2002; Hays et al. 2005). In this study, we

address the hypothesis that large-scale patterns in copepod diversity are related to variations in climatic factors and energy availability.

There have been only few studies of open ocean diversity at large scales mainly because of the low availability and comparability of biological data at these scales. In the case of copepods, high-quality and fine taxonomic resolution information has recently become available in web-based databases (e.g., Copepod Database <http://www.st.nmfs.noaa.gov/plankton/>), enabling large-scale syntheses of geographical diversity patterns. Using these large data sets, Rombouts et al. (2009) found a polar–tropical contrast in copepod diversity with maximum values around 20°N and a tropical–subtropical plateau in the Southern Hemisphere. Diversity maxima away from the equator have also been observed for other planktonic groups such as foraminifers (Rutherford et al. 1999) and tintinnid ciliates (Dolan et al. 2006), and for higher trophic levels, e.g., fish (Worm et al. 2005). This recurrent spatial pattern suggests the existence of a common underlying mechanism.

Although several physical factors may influence copepod diversity, there have been no investigations as yet of the relationships between covariations in environmental variables and variations in distributions of marine copepod species over large latitudinal ranges. At the regional scale, environmental drivers such as sea surface temperature have been found to be strongly correlated with geographic variations in zooplankton species (Beaugrand et al. 2002; Mackas et al. 2007). Other factors at that scale are water mass distributions (White 1994; Fernández-Álamo and

* Corresponding author: isabelle.rombouts@univ-lille1.fr

¹ Present address: Station Marine de Wimereux (LOG), Wimereux, France

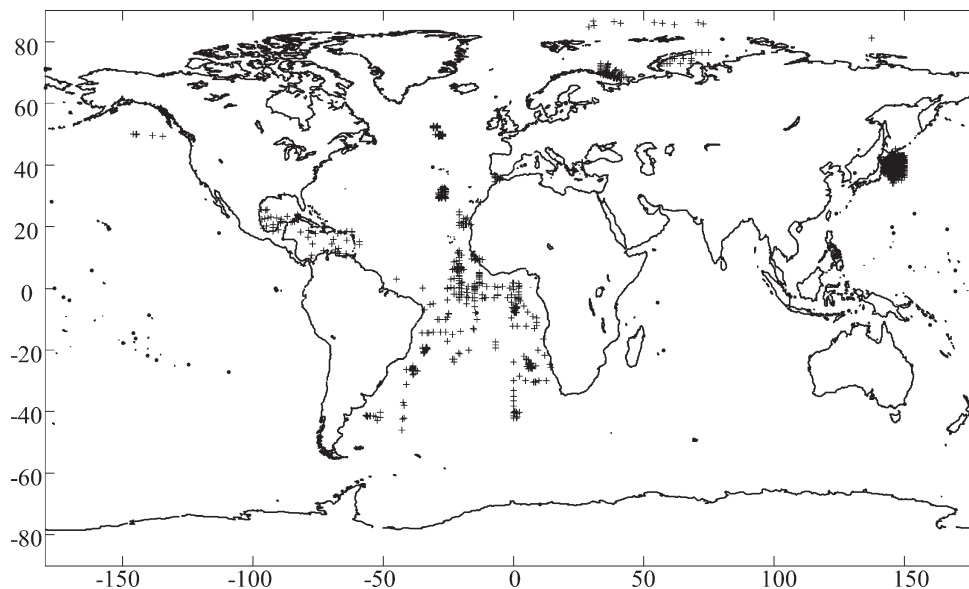


Fig. 1. Locations of the sampling stations used for the analysis. There were 13,713 samples from seven large data sets of copepod taxonomic composition (Table 1).

Färber-Lorda 2006) and variations in nutrient availability and productivity (Turner 2004). Our world-ocean analysis will include variables that are likely to affect large-scale patterns of copepod diversity: ocean temperature (Rutherford et al. 1999), primary production, and chlorophyll *a* (Chl *a*; Connell and Orias 1964), seasonality in these environmental variables (Begon et al. 1996; Rex et al. 2000), and other physical and chemical properties of the ocean (Ruddiman 1969; Stevens 1996). All these variables are likely to be influenced by climate change (Sarmiento et al. 2004). Therefore, the identification of large-scale links between variations in diversity and the environment is important to predict possible large-scale consequences of climate change on the species distributions.

Rombouts et al. (2009) described large-scale geographic variations in marine copepod diversity and single-factor correlations with environmental variables. In the present study, we have extended this first study to a multivariate approach since it is likely that, in the context of climate change, not only these environmental correlates of diversity but also their interrelationships will be affected (Stenseth et al. 2002). Therefore, we considered the relationships among environmental descriptors in addition to their relationship with diversity. Our approach was in two steps. We used a multivariate technique, i.e., principal components analysis (PCA), to investigate the relationships among environmental descriptors (ocean temperature, net primary production, Chl *a*, and other physical and chemical properties of the ocean, and seasonality in these environmental variables), and their relationships with spatial variations in copepod diversity. We applied regression on principal components (PCs) to forecast global copepod diversity.

Methods

Copepod taxonomic composition data—To relate copepod diversity to environmental parameters, we constructed

a database of copepod species composition data covering a large latitudinal extent (86.5°N to 46.5°S). We included only data sets having a temporal range of more than 5 yr, to reduce noise due to patchiness of zooplankton spatial distributions and seasonality in calculating diversity (Rombouts et al. 2009). Although several long-term time series of zooplankton diversity are available, we considered only data that offered high taxonomic resolution, i.e., identification generally to the species level. Another selection criterion was to include only data from samples collected by vertical hauls to eliminate effects of vertical layering and migrations on the calculation of diversity (0–200 m). Our final data set was extracted from the Copepod Database, the World Ocean Database, individual sampling programs (e.g., Sta. L4), and personal communications (e.g., Odate collection, Sta. P, and adjacent sites), which gave a total of 13,713 samples from seven data sets (Fig. 1; Table 1). The compiled database covered several geographic regions (Atlantic Ocean and its adjacent seas, and two regions of the Pacific Ocean). Unfortunately, and to the best of our knowledge, no information is available on copepods in the Indian Ocean and the Southern Ocean that satisfied our above selection criteria. Methodological differences, e.g., mesh sizes used for sampling, may introduce biases in the calculation of taxonomic richness. Visual inspection of some data sets (e.g., California Cooperative Fisheries Investigations time series; Rebstock 2001) showed that when mesh sizes > 330 μm were used in surveys, there was skewness toward larger copepod taxa. Since there is no uniform mathematical solution to standardize taxonomic composition data across data sets, we decided to include in our database only samples collected with a mesh size < 330 μm to avoid possible biases (Rombouts et al. 2009).

The copepod taxonomic composition data were reorganized to create a global-scale grid (1° latitude \times 1° longitude). Even if most individuals had been identified

Table 1. Data sources of copepod composition data.

Institute or data set	Geographic area	Data source	Time period	Sampling protocol
Plymouth Marine Laboratory; Sta. L4	Western English Channel	www.pml.ac.uk/L4	1988–2005	Working party 2 (WP2) net; 200 μm
Murmansk Marine Biological Institute	Barents Sea, Kara Sea	Copepod database http://www.st.nmfs.noaa.gov/plankton/index.html	1938–1981	Juday net; 168, 170 μm
White Sea Biological Station, Zoological Institute, Russian Academy of Sciences	White Sea	http://www.nodc.noaa.gov/OC5/WH_SEA/index1.html Berger et al. (2003)	1963–1998	Juday net; 168 μm
Pelagic ecosystems of the Tropical Atlantic	Tropical and South Atlantic	Copepod database http://www.st.nmfs.noaa.gov/plankton/index.html	1963–1989	Juday and Bogorov Rass net; 125, 300, 330 μm
Odate collection, Tohoku National Fisheries Research Institute	East of Japan	S. Chiba (pers. comm.)	1960–2002	Norpac net; 330 μm
Ocean Station Weather Papa and adjacent sites	North East Pacific	D. L. Mackas (pers. comm.)	1997–2008	Scientific Committee on Oceanic Research (SCOR) net; 220 μm
Laboratoire d'Océanographie de Villefranche; Point B	Ligurian Sea	S. Gasparini (pers. comm.)	1972–2003	WP2 net; 200 μm

to the species level, the taxonomic resolution was not always uniform across data sets. To resolve this problem, individuals that were not consistently identified to the species level in all data sets were reduced to the genus level before determining diversity. This loss in taxonomic resolution for some species is unlikely to have affected the calculations of diversity since genera included in the database tended to have low genus: species ratios, i.e., each genus was represented by relatively few species. Because the uneven taxonomic resolution of the data in our study would have biased the calculation of indices that are based on relative frequencies of taxa, we chose taxonomic richness, i.e., the number of taxa present per geographical cell, as diversity measure (Rombouts et al. 2009). It was calculated as follows. The number of taxa in each grid cell was integrated over the epipelagic layer (0–200 m) and monthly means were pooled to obtain yearly taxonomic richness for each cell. The resulting values were time-averaged over all observation years. Before the calculation of linear correlation coefficients between richness and PCs, diversity data (i.e., taxonomic richness) were log-transformed (\log_{10}) to linearize the relationships.

The values of diversity we calculated should not be interpreted as reflecting the true epipelagic diversity of copepods, but they are appropriate to study diversity in relation to its environment along geographical gradients. Simple richness estimators underestimate diversity at all sample sizes, but the methods to reduce underestimation, e.g., jackknife and bootstrap estimators, could not be used here since we did not have sufficient numbers of replicates to correctly apply these techniques (Beaugrand and Edwards 2001). Furthermore, inconsistencies in the calculation of species richness caused by differences in sample sizes could be resolved by rarefaction methods (Sanders 1968; Hurlbert 1971). However, these methods could not be applied to our data because of small sample sizes in certain regions. For example, the Tropical Atlantic database consisted of a collation of research cruises that covered an extensive geographical region but had relatively low

temporal coverage and therefore contained grid cells with occasional small samples compared with areas with point sampling data (e.g., Sta. L4, Sta. P). Furthermore, rarefaction methods could not be applied because the necessary information on the number of individuals for each species was not always known, i.e., our data set was a mixture of species and genus numbers. A detailed description of the data sources and database homogenization is given in Rombouts et al. (2009).

Environmental data—We created a complementary database of 11 potential environmental correlates of diversity, i.e., ocean temperature, salinity, oxygen, concentrations of nitrate, phosphate, and silicate, net primary production, Chl *a*, water depth, ocean surface currents, and mixed-layer depth (MLD). Data sources are given in Table 2. Most data sets were already formatted as monthly standardized data per unit area on a global-scale grid (1° latitude \times 1° longitude), which facilitated matching with the gridded copepod diversity data. For each environmental variable except water depth, we calculated the annual mean and the coefficient of variation, used here as an index of seasonal variation. This resulted in four matrices of environmental variables: two global grids, one of annual mean values, and one of coefficients of variation and two local matrices, i.e., only those cells (total of 433) where biological data were available, that contained mean annual values and coefficients of variation, respectively. The two matrices of annual mean values and the two matrices of coefficients of variation included 11 and 10 variables, respectively, as water depth did not change with time.

Statistical analyses—Standardized PCA on the environmental data sets: Standardized PCA is a multivariate technique that can be viewed as an orthogonal decomposition of the variance of a data set (Jolliffe 1986). This technique was applied to the potential environmental correlates of copepod diversity to determine: (1) the environmental variables that explained the largest variation

Table 2. Data sources of the environmental variables used in principal components analyses (na, not applicable).

Environmental variable	Data source	Time period	Geographic area	Spatial resolution (°)
Ocean temperature (°C)	World Ocean Atlas 2005; http://www.nodc.noaa.gov/OC5/WOA05/wao05data.html	Long-term monthly mean (1950–2004)	Global	1
Salinity				
Oxygen concentration (mL L ⁻¹)				
Nitrate concentration (μmol L ⁻¹)				
Silicate (μmol L ⁻¹)				
Phosphate (μmol L ⁻¹)	Ocean Surface Current Analyses—Real time (OSCAR); Bonjean and Lagerloef (2002); http://www.oscar.noaa.gov/index.html	Monthly mean (1992–2006)	60°N–60°S	1
Ocean surface currents (m s ⁻¹)				
Chlorophyll <i>a</i> (mg m ⁻³)	Sea-viewing Wide Field-of-view Sensor (SeaWiFS); http://oceancolor.gsfc.nasa.gov	Monthly mean (1997–2005)	Global	1
Net primary production (mg m ⁻² d ⁻¹)	Standard VGPM from MODIS 1.1 data; Behrenfeld and Falkowski (1997); http://www.science.oregonstate.edu/ocean.productivity/index.php	Monthly mean (2002–2007)	Global	1/6
Bathymetry (m)	Smith and Sandwell (1997)	na	Global	1
Mixed-layer depth (m)	Institute Pierre Simon Laplace; http://www.locean-ipsl.upmc.fr/~cdblod/mld.html	Long-term monthly mean (1941–2002)	Global	2

in the data set, (2) the relationships among these potential environmental predictors, and (3) in the case of the local data set, how the scores of the PCs were related to copepod diversity. The procedure was applied to the four environmental data sets separately.

The environmental data were normalized using the omnibus procedure (Legendre and Legendre 1998). The correlation matrix of all standardized variables was used to calculate the eigenvectors and the PCs. The PCs were then ranked in order of significance and the contribution of each variable to each PC was calculated. To check for nonlinearity among environmental descriptors, the multilinearity of the PCs was tested.

For the PCA on the local data sets, taxonomic diversity, latitude, and longitude were added by passive ordination (i.e., as supplementary variables). The outcome of the PCA was used to investigate the relationships of copepod diversity with a combination of environmental factors instead of computing a suite of correlation coefficients of diversity with single factors. Linear Bravais–Pearson's correlations were calculated to assess the relationship between each PC and log-transformed copepod diversity.

Correlations between the eigenvectors of the PCAs on the global and local environmental data sets—We used correlation analyses to investigate if the variation in the local subset of environmental data for which copepod data were available captured the full range of variation that existed in environmental correlates that occurred globally. If this were not so, the relationships found between diversity and environmental variables at local scale would not be representative of the relationships at global scale. Linear correlations (Bravais–Pearson) were calculated between the eigenvectors from PCA on the global environmental data set and the eigenvectors from the local data set for the mean annual values.

Regression on PCs—Multicollinearity among variables can often distort the results of a multiple linear regression, in particular when a large number of environmental variables is used. To avoid the problem of collinearity among predictor variables X_j s, we used regression on PCs (Jolliffe 1986; Legendre and Legendre 1998), which corresponds to a multiple regression between the response variable Y and a subset of x PCs C_j of matrix X computed on the matrix of correlations among the X_j s. Regression on PCs has been used in previous studies for reconstructing paleoclimates (Buckley et al. 2004) and meteorological forecasting (Kung and Sharif 1980).

A regression model on PCs was fitted using diversity (taxonomic richness) as the response variable and a subset of PCs resulting from the PCA that had been performed on the global data set of mean environmental variables as explanatory variables (see above). The PCs included in the subset were chosen with the specific purpose of predicting global copepod diversity. To do so, we used the selection procedure proposed by Guidi et al. (2009), which allows selecting those PCs that contribute significantly to fitting the regression model ($p < 0.05$). (1) Each PC was regressed separately against diversity (taxonomic richness) and the coefficients of determination were calculated (r^2) for each regression. (2) These coefficients were ranked and the cumulated sum was calculated. (3) The sums were plotted in addition to the approximate derivatives of the cumulated sum of the ranked coefficients of determination to decide which PCs to keep for prediction. (4) Because the residuals of the regression model must not be spatially autocorrelated when predicting global copepod diversity, we checked for spatial autocorrelation by calculating a semivariogram of the residuals and regressing the values of semivariance against the distance classes; the absence of a significant slope, $b = 0$, would indicate the absence of spatial autocorrelation. (5) We tested the normality of the

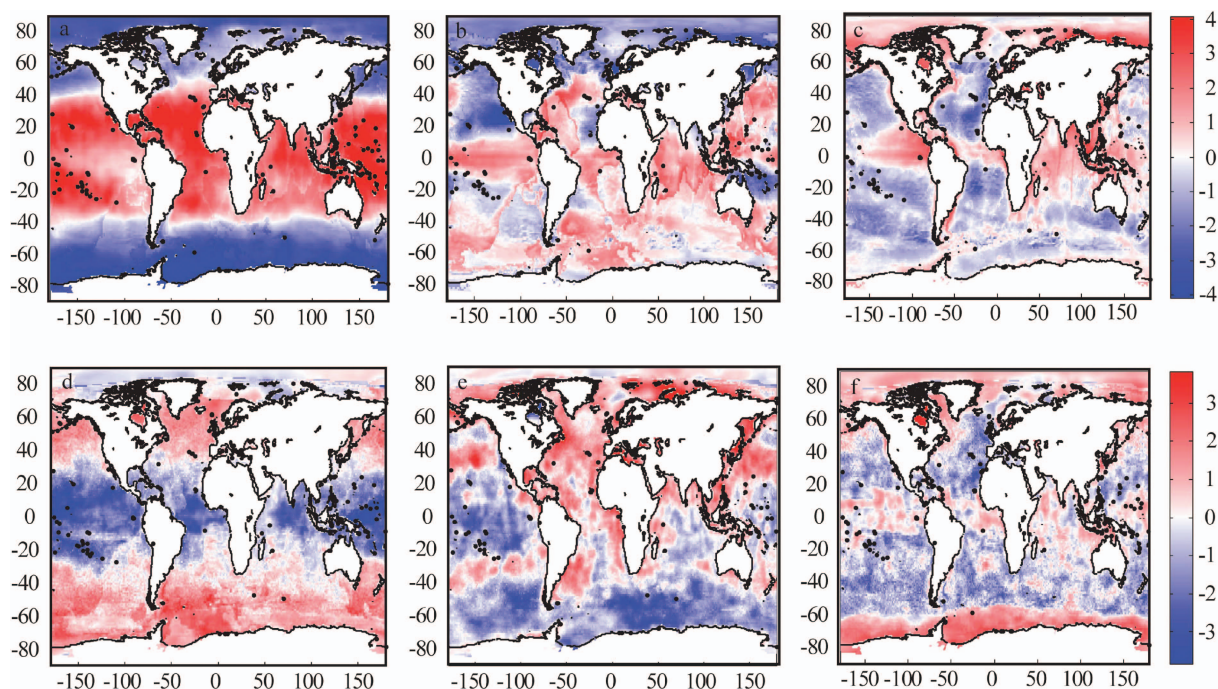


Fig. 2. (a–c) Representation of the principal components from the PCA on the global environmental data set of annual means and (d–f) coefficients of seasonal variation: (a, d) first principal components, (b, e) second principal components, and (c, f) third principal components.

residuals from the skewness and kurtosis of their distribution.

Results

Variations in the global environmental data sets—Mean annual data: Maps of the PC scores show the main patterns of spatial variability of the mean annual values of environmental factors (Fig. 2a–c). The first three PCs accounted for 74% of the variance in the global data set. The first PC (PC1) (Fig. 2a) mostly represented ocean temperature, oxygen concentration, and nutrients. We found a steep gradient in the scores around 40°N and 40°S and low scores in upwelling regions (e.g., Peru and Benguela) (Fig. 2a). The second PC (PC2) (Fig. 2b) mostly represented net primary production. Bathymetry and sea surface currents contributed most to the third PC (PC3) (Fig. 2c). The equatorial currents, the Gulf Stream, and the Antarctic circumpolar currents are clearly visible on Fig. 2c. The multinormality of the PCs was tested and confirmed that the relationships among environmental descriptors were linear.

Seasonal variation data: Most of the variance of the seasonal variation data set was explained by the first three PCs (67%). PC1 mostly represented the seasonal variation in ocean temperature, MLD, and primary production (positive scores). The scores were negative in the equatorial regions, which indicated low seasonal variation in ocean temperature, MLD, and net primary production in these areas (Fig. 2d). PC2 mostly represented the seasonal variation in phosphate and nitrate, which were positively

related to the PC. Relatively high scores and therefore high seasonal variation in these variables were visible in coastal regions and areas where surface currents were strong (Fig. 2e). PC3, which is mapped in Fig. 2f, mostly represented the seasonal variability in salinity (positive scores) with high scores in equatorial and polar regions.

Variations in local environmental data sets and their relationships with diversity—Copepod diversity and mean environmental data: The first three PCs explained 73% of the total variance in the local data set of annual means. The contributions of the 11 environmental variables (annual means) to PC1 and PC2 are presented in Fig. 3. The relationships among variables in the reduced space are indicated by the angles among the vectors and the lengths of the vectors indicate their contributions to the reduced space. Most descriptors contributed highly to the first two principal axes except ocean surface currents, which contributed to the second axis, and depth and net primary production, which contributed to the third axis. Ocean temperature and salinity were on the positive side of PC1, and Chl *a*, silicate, phosphate, nitrate, and oxygen on the negative side. This representation also shows that high copepod diversity (supplementary variable) generally corresponded to high temperature and salinity and low concentrations of nutrients, oxygen, and Chl *a*. Sea surface currents had high negative loadings on PC2, which accounted for 16% of the total variance in the data set. The opposition of diversity against latitude and longitude in the reduced space indicated high diversity at low latitudes. High latitudes corresponded to high concentrations of nutrients, oxygen, and Chl *a*.

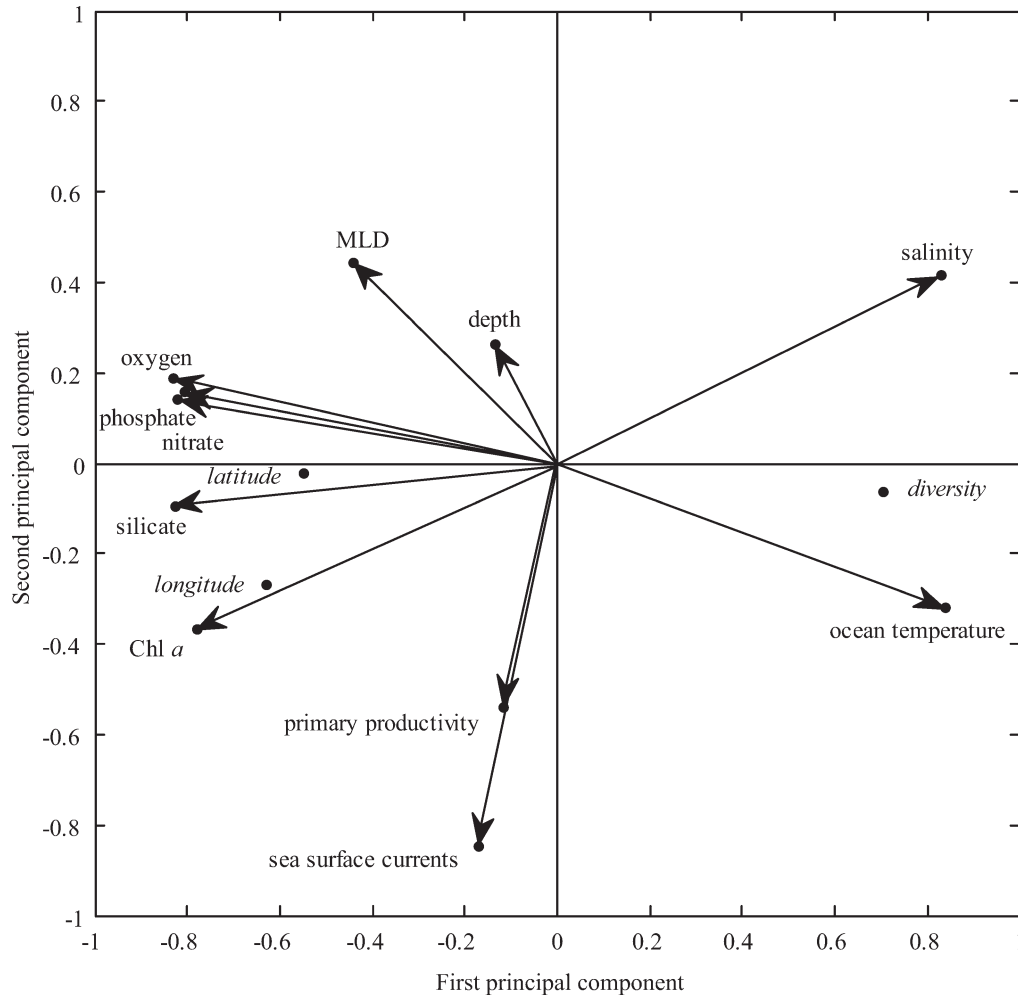


Fig. 3. Contributions of environmental descriptors (annual means) to the space of the first two principal components. Diversity, latitude, and longitude were added as supplementary variables (in italics). The circle of equilibrium descriptor contribution was drawn at $\sqrt{2/p} = 0.42$, where $p = 11$ descriptors. MLD = mixed-layer depth.

Correlation analyses confirmed a strong positive relationship ($r = 0.70$, $p < 0.001$) of copepod diversity with PC1 (Fig. 4). Furthermore, there was a significant negative relationship ($r = -0.25$, $p < 0.001$) between diversity and PC5 (Fig. 4), which mainly represented net primary productivity (not shown here).

Copepod diversity and seasonal variability in the environment: The first PC (PC1) of seasonal variability explained 31% of the total variance in the data set, which was considerably lower than the variance explained by PC1 of the local annual mean data. On the first PC axis (Fig. 5), seasonal variability of oxygen, temperature, and MLD all showed positive scores and were therefore positively interrelated, whereas sea surface currents showed a negative score. Hence, in contrast to the outcome of the PCA on mean values (Fig. 3), the largest variation in the data set (PC1) was mainly accounted for by the seasonal variation in the physical structure of the environment. The opposition of diversity to positive scores on PC1 in the reduced space indicated that high-diversity areas corre-

sponded to low seasonal variability in oxygen, ocean temperature, and MLD (Fig. 5).

In general, correlations of diversity with PCs representing seasonal environmental variability were weaker than with the PCs based on mean values. Nevertheless, there were significant correlations of diversity with PC1 ($r = -0.32$, $p < 0.001$), PC2 ($r = 0.33$, $p < 0.001$), and PC4 ($r = -0.33$, $p < 0.001$) of the seasonal variability data (Fig. 6). These negative correlations indicated that high diversity corresponded to low seasonal variability of the descriptors contributing to the first and fourth PCs. The latter PC mainly represented the seasonal variation in net primary production.

Comparison between global and local variations in environmental data—To determine if our local subsets of environmental data were representative of the variation in environmental variables at the global scale, we plotted the correlations between the values (mean annual averages) of the cumulated number of eigenvectors of the global and local data sets (not shown here). The correlation was

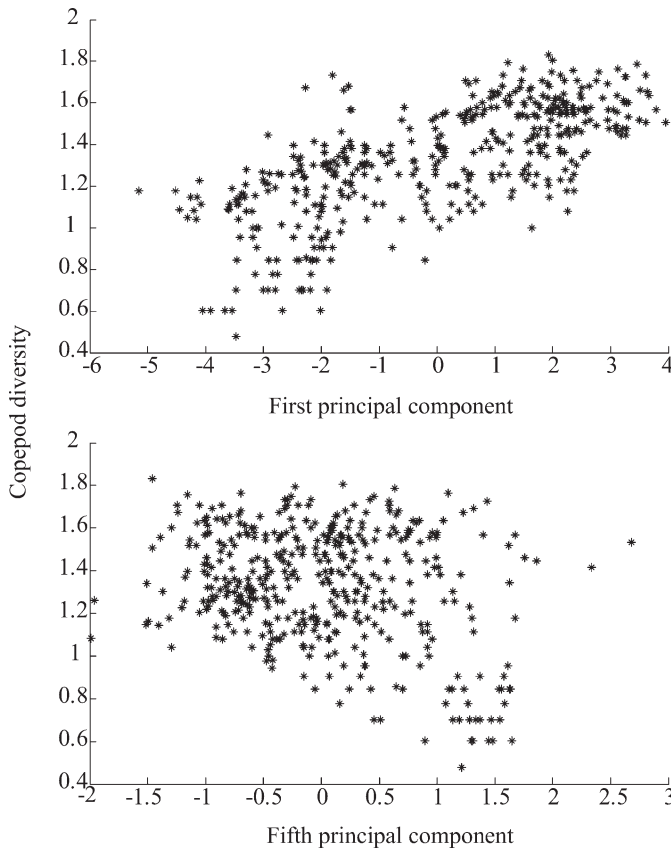


Fig. 4. Relationships between copepod diversity (calculated as \log_{10} of taxonomic richness) and the first ($r = 0.70$, $p < 0.001$) and fifth ($r = -0.25$, $p < 0.001$) principal components from the PCA performed on the mean environmental variables of the local data set. Only these two PCs were significantly correlated with diversity.

strongest when only the first eigenvector of the two matrices ($r = 0.95$) was used, and decreased rapidly thereafter ($r = 0.19$ when the first two eigenvectors were included). The statistical significance of the correlation coefficients could not be tested since the assumption of independence between matrices was not met.

Predicting global copepod diversity using PCs—Copepod diversity (taxonomic richness) was regressed against the geographically corresponding PCs from the global data set of mean environmental data. For selecting the PCs to be used in the regression model, we looked at the slope of the approximate derivatives of the cumulated sums of ranked coefficients of determination (not shown here) plotted against their ranks. A sharp point of inflection after the third derivative indicated that adding a fourth PC to the regression model would not provide additional information. Therefore, only the first three ranked PCs were retained in the regression model ($R^2 = 0.45$, $p < 0.001$), where PC1 ($\beta = 4.8$, $p < 0.05$), PC4 ($\beta = 2.4$, $p < 0.05$), and PC2 ($\beta = 2.2$, $p < 0.05$) were used to predict copepod diversity at the global scale (Fig. 7). The PCs retained in the model were selected solely on their ability to predict diversity, and not on the total variance they explained in

the environmental data set. Indeed, low variance explanation by a component does not necessarily imply that it has low predictive power (Jolliffe 1986). The residuals of the regression were normally distributed, and their spatial distribution did not indicate any distinct areas of over- or underprediction. However, in areas of low copepod diversity (mainly polar regions), some predicted diversity values were negative and were replaced by zero values (null diversity). Finally, the regression of the values of semivariance on distance classes indicated the absence of a significant slope, $b = 0.02$. Hence, the residuals corresponded to white noise, and there was no significant spatial autocorrelation.

Discussion

Our results confirm the previous result of Rombouts et al. (2009) that mean ocean temperature is the most important correlate of large-scale variations in copepod diversity. The multivariate approach used in this study has an advantage over single-factor correlations since the relationships among environmental variables were examined simultaneously with their contributions to the geographic variations in diversity. When taking into account the contributions of all environmental variables, high copepod diversity areas (Fig. 3) corresponded to a combination of high ocean temperature and salinity and low Chl *a* and nutrient concentrations (nitrate, silicate, phosphate). These interrelationships were consistent with previous findings on environmental correlates of large-scale diversity of planktonic bacteria (Fuhrman et al. 2008).

Our regression on PCs model predicted diversity reasonably well using three PCs ($R^2 = 0.45$, $p < 0.001$), and thus proved to be a robust tool to predict global copepod diversity as a function of interrelated environmental variables. The first PC from the global data set contributed most to the model, and represented mainly ocean temperature, Chl *a*, and nutrients. This is not surprising since diversity was highly correlated to the first PC of the local data set. Because the interrelationships among temperature, salinity, Chl *a*, and nutrients were the same at the regional and global scales, we conclude that these environmental descriptors were important correlates of diversity at the two scales. The subtropical gyres in the Atlantic and Pacific showed higher diversity values, and upwelling regions (e.g., Arabian Sea, Benguela, and Humboldt currents) lower diversity (Fig. 7). Even more important is the asymmetry in the distribution of diversity predicted by the model between the Northern and the Southern hemispheres in the Atlantic, which can be explained by the hemispheric asymmetry of climatic conditions in the equatorial region. This asymmetry is consistent with a previous analysis of the field diversity data used in the present study (Rombouts et al. 2009).

We used the relationships between environmental variations and copepod diversity to build a regression model, but there are at least two shortcomings to our approach. Although our model did explain almost half the variance of our heterogeneous data set, it is possible that inclusion in the model of some additional, unavailable

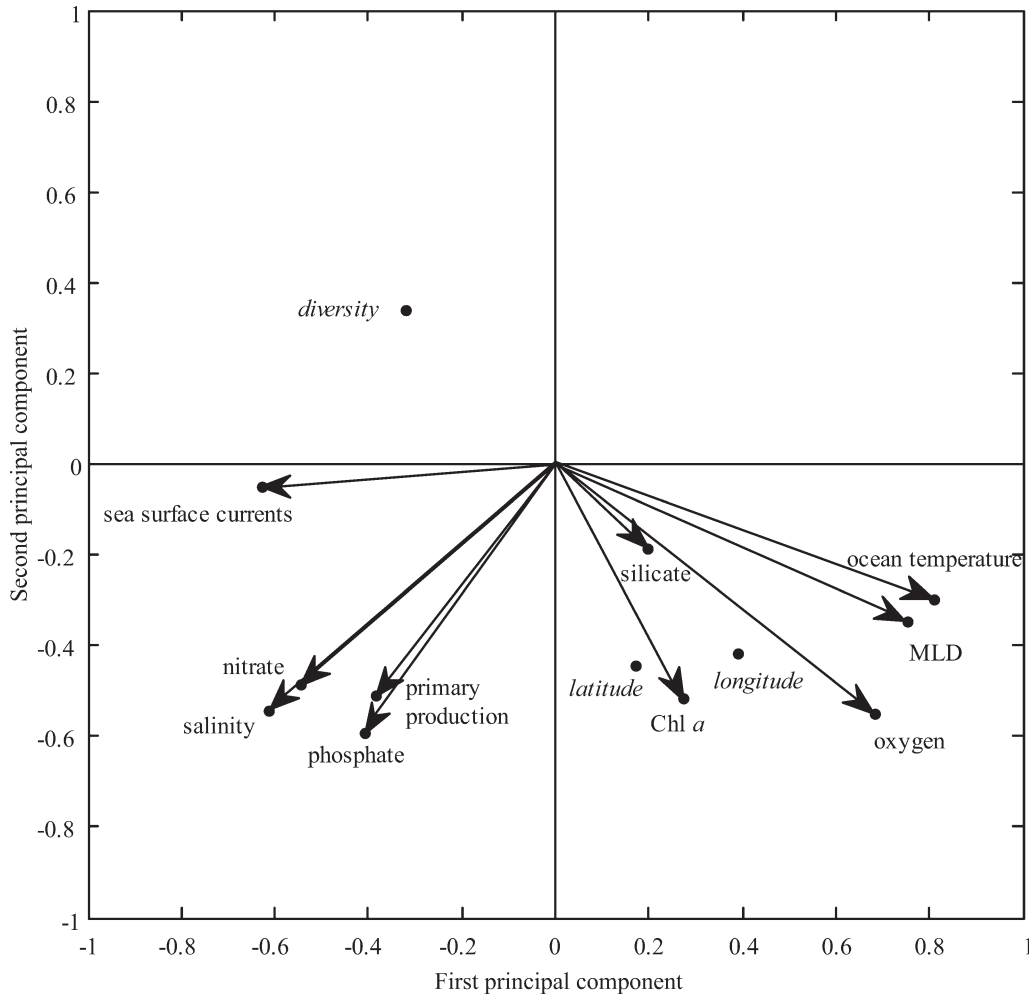


Fig. 5. Contributions of environmental descriptors (seasonal variability) to the space of the first two principal components. Diversity, latitude, and longitude were added as supplementary variables (in italics). The circle of equilibrium descriptor contribution was drawn at $\sqrt{2/p} = 0.44$, where $p = 10$ descriptors. MLD = mixed-layer depth.

variables could have improved its explanatory power. Furthermore, some of the PCs included in the model did not explain a large variation in the environmental data set (e.g., PC4), although they were important predictors of diversity. The lack of correspondence between the explanatory and predictive power of some PCs impedes the ecological interpretation of the regression equation.

The negative relationship between diversity and both nutrients and Chl *a* suggests that copepod diversity was high in oligotrophic regions, in particular the subtropical gyres in the two hemispheres, which are characterized not only by low nutrients but also high temperatures and salinity (Robinson and Henderson-Sellers 1999). The latitudinal distribution of modeled copepod diversity reflected the high diversity values already reported around 20–30°N globally (Rombouts et al. 2009) and in the South Atlantic anticyclonic gyre (Piontkovski et al. 2003). High diversity in oligotrophic regions has been related to a complex network of biological interactions (McGowan and Walker 1985; Piontkovski et al. 2003). Low food resource availability for copepods in oligotrophic regions could result in competition for resources and hence, if species are

to coexist in a community, efficient division of resources would be essential. In this respect, resource partitioning could regulate the number of species in a community (Schoener 1974), as is the case for some terrestrial communities (Tilman et al. 1996). In coastal marine ecosystems, an inverse relationship between zooplankton diversity and resource availability has also been observed (Badosa et al. 2006).

The strong negative relationship we observed between diversity and MLD (Fig. 4) indicates that diversity was high in stratified waters. A similar observation was made by Rutherford et al. (1999) who proposed that a stable physical structure in the near-surface ocean provides high vertical niche availability than can support a high number of species.

Copepod diversity was higher in areas with stable seasonal ocean temperature, MLD, and oxygen. Furthermore, copepod diversity was negatively related to seasonal variations in net primary production. This result corresponds with the observation that oligotrophic regions, with low seasonal variability, have high diversity, although seasonal variability has lower correlations with diversity

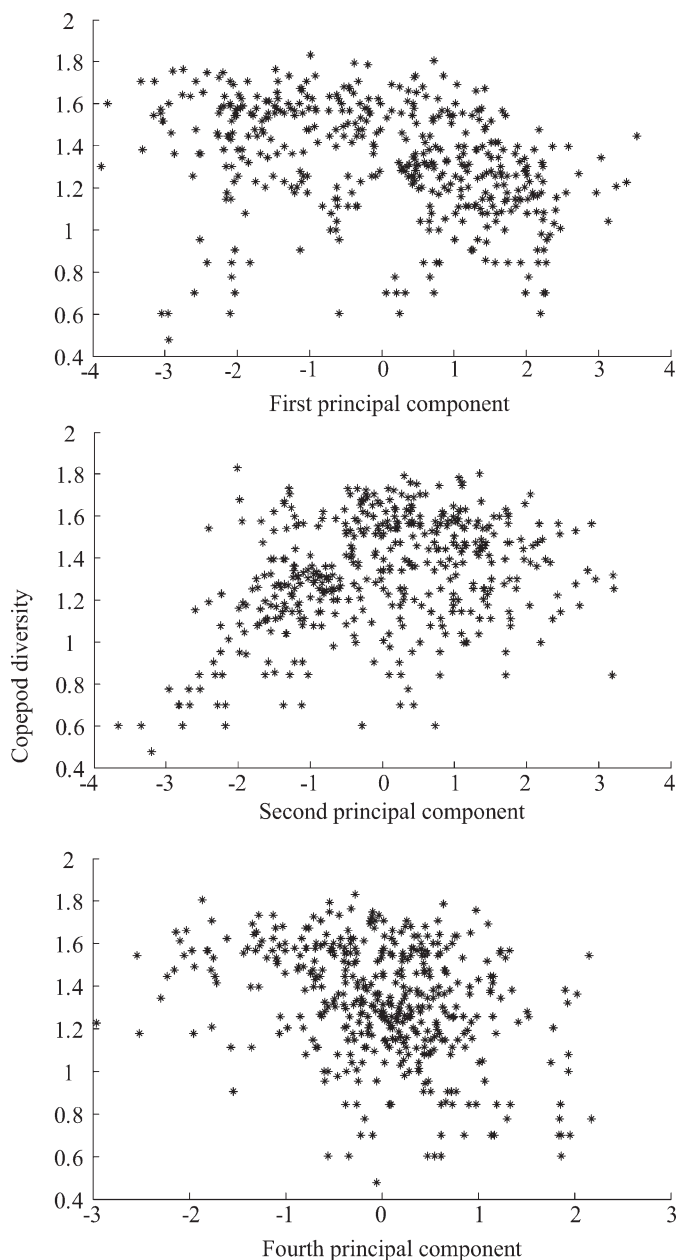


Fig. 6. Relationships between copepod diversity (calculated as \log_{10} of taxonomic richness) and the first ($r = -0.32$, $p < 0.001$), second ($r = 0.33$, $p < 0.001$), and fourth ($r = -0.33$, $p < 0.001$) principal components from the PCA performed on the seasonal variability in environmental variables of the local data set. Only these three PCs were significantly correlated with diversity.

than the mean environmental variables described above. These results are consistent with previous ideas that latitudinal trends in diversity are a positive function of the available energy rather than an inverse function of environmental variability or harshness (Roy et al. 2000; Turner 2004).

Hypotheses related to climate and energy availability may be among the most plausible explanations for the regulation and maintenance of diversity at large spatial

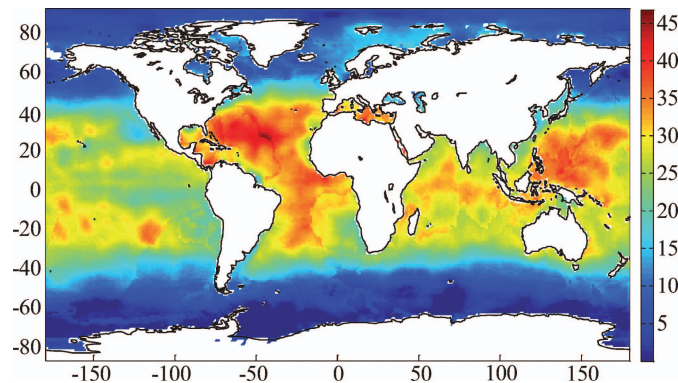


Fig. 7. Predicted global copepod diversity (D) with the following regression equation: $D = 17.85 + 4.8PC1 + 2.4PC4 + 2.2PC2$, where the three PCs are principal components combining mean environmental variables.

scales (Currie 1991; Roy et al. 1998; Field et al. 2008). However, these hypotheses frequently fail to distinguish between fundamentally different forms of energy that influence diversity in different ways (Clarke and Gaston 2006). Species–energy relationships can be influenced by several mechanisms that include the effects of temperature on biochemical reaction rates (Allen et al. 2002), the ability of organisms to maintain homeostasis (Currie 1991; Hawkins et al. 2003), and the effects of ecosystem productivity on total community abundance (Wright 1983; Allen et al. 2002; Clarke and Gaston 2006). Allen et al. (2007) developed a theoretical framework combining different energetic pathways that influence diversity. More specifically, they proposed that thermal kinetic energy (temperature) influences the evolutionary rate of each population and chemical potential energy (net primary production) influences total number of populations. However, temperature is also likely to influence diversity indirectly, for example, through the effects on the rate of utilization of chemical energy by organisms (Clarke and Gaston 2006). Even though the mechanisms that could influence diversity were not explicitly addressed in this study, the strong positive relationship between copepod diversity and ocean temperature suggests that the available kinetic energy was an important factor in shaping the geographic variations in diversity, whereas net primary production as a measure of chemical energy was only weakly correlated to copepod diversity.

Even though climate and energy availability have an overarching influence on species richness, smaller-scale processes are likely to fine-tune community interactions (Gaston and Blackburn 2000), e.g., biotic interactions could act in concert with other mechanisms to enhance and strengthen climatic diversity gradients (Currie et al. 2004). For example, as mentioned before, copepod diversity in oligotrophic regions may be ultimately regulated by climate and energy availability, but may be enhanced locally by species interactions. The ecological importance of historical processes and other abiotic factors (area, spatial heterogeneity, etc.) cannot be ignored at the global scale, but it was beyond the scope of this study to investigate their potential influence.

Incorporating diversity into already complex ecosystem and biochemical models is of pressing importance in the face of accelerating global environmental change (Duffy and Stachowicz 2006). Climate change is likely to affect environmental correlates of diversity such as ocean temperature (Sarmiento et al. 2004), salinity (Toledo et al. 2007), currents (Greene and Pershing 2007), and their interactions (Stenseth et al. 2002). These changing environmental conditions will affect the patterns of marine biodiversity mostly through changes in distributions of taxa as a result of shifts in latitudinal ranges (Beaugrand et al. 2002; Worm et al. 2006). Bioclimatic models have predicted that pelagic taxa in particular will be affected by climate-induced global warming, and showed correspondence between temperature warming close to the surface and a high rate of shift in the geographical ranges of pelagic taxa due to their high dispersal abilities (Cheung et al. 2009). The possible disruption of the structure and functioning of established ecosystems induced by changes in patterns of taxonomic richness (Worm et al. 2006) can lead to changes in major biogeochemical cycles (Beaugrand et al. 2010), which would in turn affect climate (IPCC 2007).

Acknowledgments

We thank the institutes and research programs that shared their valuable data with us, in particular, the Tohoku National Fisheries Institute, which provided us with the Odate collection. We also thank David L. Mackas and Todd O'Brien for helping us with the acquisition of additional data sets and two anonymous reviewers for helpful comments on this manuscript.

This research was funded by the European Marie Curie Early-Stage Training project Meta-Oceans (MEST-CT-2005-019678).

References

- ALLEN, A. P., J. H. BROWN, AND J. F. GILLOOLY. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* **297**: 1545–1548, doi:10.1126/science.1072380
- , J. F. GILLOOLY, AND J. H. BROWN. 2007. Recasting the species–energy hypothesis: The different roles of kinetic and potential energy in regulating biodiversity, p. 283–299. *In* D. Storch, P. A. Marquet, and J. H. Brown [eds.], *Scaling biodiversity*. Cambridge Univ. Press.
- BADOSA, A., D. BOIX, S. BRUCET, R. LOPEZ-FLORES, S. GASCO, AND X. D. QUINTANA. 2006. Zooplankton taxonomic and size diversity in Mediterranean coastal lagoons (NE Iberian Peninsula): Influence of hydrology, nutrient composition, food resource availability and predation. *Estuar. Coast. Shelf. S.* **71**: 335–346, doi:10.1016/j.ecss.2006.08.005
- BEAUGRAND, G. 2009. Decadal changes in climate and ecosystems in the North Atlantic Ocean and adjacent seas. *Deep Sea Res. II* **56**: 656–673, doi:10.1016/j.dsr2.2008.12.022
- , AND M. EDWARDS. 2001. Differences in performance among four indices used to evaluate diversity in planktonic ecosystems. *Oceanol. Acta* **24**: 467–477, doi:10.1016/S0399-1784(01)01157-4
- , ———, AND L. LEGENDRE. 2010. Marine biodiversity, ecosystem functioning, and carbon cycles. *Proc. Natl. Acad. Sci. USA* **107**: 10120–10124, doi:10.1073/pnas.0913855107
- , P. C. REID, F. IBANEZ, J. A. LINDLEY, AND M. EDWARDS. 2002. Reorganization of North Atlantic marine copepod biodiversity and climate. *Science* **296**: 1692–1694, doi:10.1126/science.1071329
- BEGON, M., J. L. HARPER, AND C. R. TOWNSEND. 1996. *Ecology: Individuals, populations and communities*. Blackwell.
- BEHRENFELD, M. J., AND P. G. FALKOWSKI. 1997. Photosynthetic rates derived from satellite-based chlorophyll concentrations. *Limnol. Oceanogr.* **42**: 1–20, doi:10.4319/lo.1997.42.1.0001
- BERGER, V. J., A. D. NAUMOV, N. USOV, M. ZUBAHA, I. SMOLYAR, R. TATUSKO, AND S. LEVITUS. 2003. 36-year time series (1963–1998) of zooplankton, temperature, and salinity in the White Sea. *NOAA Atlas NESDIS 57 IV*.
- BONJEAN, F., AND G. S. E. LAGERLOEF. 2002. Diagnostic model and analysis of the surface currents in the Tropical Pacific Ocean. *J. Phys. Oceanogr.* **32**: 2938–2954, doi:10.1175/1520-0485(2002)032<2938:DMAAOT>2.0.CO;2
- BUCKLEY, B. M., R. J. S. WILSON, P. E. KELLY, D. W. LARSON, AND E. R. COOK. 2004. Inferred summer precipitation for southern Ontario back to AD 610 as constructed from ring widths of *Thuja occidentalis*. *Can. J. For. Res.* **34**: 2541–2553, doi:10.1139/x04-129
- CHEUNG, W. W. L., V. W. Y. LAM, J. L. SARMIENTO, K. KEARNEY, R. WATSON, AND D. PAULY. 2009. Projecting global marine biodiversity impacts under climate change scenarios. *Fish. Fish.* **10**: 235–251.
- CLARKE, A., AND K. J. GASTON. 2006. Climate, energy and diversity. *Proc. R. Soc. Lond. B Biol. Sci.* **273**: 2257–2266, doi:10.1098/rspb.2006.3545
- CONNELL, J. H., AND E. ORIAS. 1964. The ecological regulation of species diversity. *Am. Nat.* **98**: 399–414, doi:10.1086/282335
- CURRIE, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. *Am. Nat.* **137**: 27–49, doi:10.1086/285144
- , AND OTHERS. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* **7**: 1121–1134, doi:10.1111/j.1461-0248.2004.00671.x
- DOLAN, J. R., S. GASPARINI, L. MOUSSEAU, AND C. HEYNDRIKX. 2006. Probing diversity in the plankton: Using patterns in Tintinnids (planktonic marine ciliates) to identify mechanisms. *Hydrobiologia* **555**: 143–157, doi:10.1007/s10750-005-1112-6
- DUFFY, J. E., AND J. J. STACHOWICZ. 2006. Why biodiversity is important to oceanography: Potential roles of genetic, species, and trophic diversity in pelagic ecosystem processes. *Mar. Ecol. Prog. Ser.* **311**: 179–189, doi:10.3354/meps311179
- FERNÁNDEZ-ÁLAMO, M. A., AND J. FÄRBER-LORDA. 2006. Zooplankton and the oceanography of the eastern tropical Pacific: A review. *Prog. Oceanogr.* **69**: 318–359, doi:10.1016/j.pocean.2006.03.003
- FIELD, R., AND OTHERS. 2008. Spatial species-richness gradients across scales: A meta-analysis. *J. Biogeogr.* **36**: 132–147, doi:10.1111/j.1365-2699.2008.01963.x
- FUHRMAN, J. A., J. A. STEELE, I. HEWSON, M. S. SCHWALBACH, M. V. BROWN, J. L. GREEN, AND J. BROWN. 2008. A latitudinal gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci. USA* **105**: 7774–7778, doi:10.1073/pnas.0803070105
- GASTON, J. K., AND T. M. BLACKBURN. 2000. *Pattern and process in macroecology*. Blackwell.
- GREENE, C. H., AND A. J. PERSHING. 2007. Climate drives sea change. *Science* **315**: 1084–1085, doi:10.1126/science.1136495
- GUIDI, I., L. STEMMANN, G. A. JACKSON, F. IBÁÑEZ, H. CLAUSTRE, L. LEGENDRE, M. PICHERAL, AND G. GORSKY. 2009. Effects of phytoplankton community on production, size and export of large aggregates: A world-ocean analysis. *Limnol. Oceanogr.* **54**: 1951–1963.
- HAWKINS, B. A., R. FIELD, H. V. CORNELL, D. J. CURRIE, J.-F. GUÉGAN, AND D. M. KAUFMAN. 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* **84**: 3105–3117, doi:10.1890/03-8006

- HAYS, G. C., A. J. RICHARDSON, AND C. ROBINSON. 2005. Climate change and marine plankton. *Trends Ecol. Evol.* **20**: 337–344, doi:10.1016/j.tree.2005.03.004
- HESSEN, D. O., V. BAKKESTUEN, AND B. WALSENG. 2007. Energy input and zooplankton species richness. *Ecography* **30**: 749–758, doi:10.1111/j.2007.0906-7590.05259.x
- HURLBERT, S. H. 1971. The non-concept of species diversity: A critique and alternative parameters. *Ecology* **52**: 577–586, doi:10.2307/1934145
- INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC) WORKING GROUP II. 2007. Climate change 2007: Impacts, adaptation and vulnerability. Cambridge Univ. Press.
- JOLLIFFE, I. T. 1986. Principal components analysis. Springer-Verlag.
- KUNG, E. C., AND T. A. SHARIF. 1980. Regression forecasting of the Indian summer monsoon with the antecedent upper air conditions. *J. Appl. Meteorol.* **19**: 370–380, doi:10.1175/1520-0450(1980)019<0370:RFOTOO>2.0.CO;2
- LEGENDRE, P., AND L. LEGENDRE. 1998. Numerical ecology. Elsevier.
- MACKAS, D. L., S. BATTEN, AND M. TRUDEL. 2007. Effects on zooplankton of a warmer ocean: Recent evidence from the Northeast Pacific. *Prog. Oceanogr.* **75**: 223–252, doi:10.1016/j.pocean.2007.08.010
- MCGOWAN, J. A., AND P. W. WALKER. 1985. Dominance and diversity maintenance in an oceanic ecosystem. *Ecol. Monogr.* **55**: 103–108, doi:10.2307/1942527
- PIONTKOVSKI, S. A., AND OTHERS. 2003. Plankton communities of the South Atlantic anticyclonic gyre. *Oceanol. Acta.* **26**: 255–268, doi:10.1016/S0399-1784(03)00014-8
- REBSTOCK, G. A. 2001. Long-term stability of species composition in calanoid copepods off southern California. *Mar. Ecol. Prog. Ser.* **215**: 213–224, doi:10.3354/meps215213
- REID, P. C., AND G. BEAUGRAND. 2002. Interregional biological responses in the North Atlantic to hydrometeorological forcing, p. 27–48. *In* K. Sherman and H. R. Skjoldal [eds.], Large marine ecosystems of the North Atlantic: Changing states and sustainability. Elsevier.
- REX, M. A., C. T. STUART, AND G. COYNE. 2000. Latitudinal gradients of species richness in the deep sea benthos of the North Atlantic. *Proc. Natl. Acad. Sci. USA* **97**: 4082–4085, doi:10.1073/pnas.050589497
- ROBINSON, P. J., AND A. HENDERSON-SELLERS. 1999. Contemporary climatology. Prentice Hall.
- ROEMMICH, D., AND J. MCGOWAN. 1995. Climatic warming and the decline of zooplankton in the California current. *Science* **267**: 1324–1326, doi:10.1126/science.267.5202.1324
- ROMBOUTS, I., G. BEAUGRAND, F. IBAÑEZ, S. GASPARINI, S. CHIBA, AND L. LEGENDRE. 2009. Global latitudinal variations in marine copepod diversity and environmental factors. *Proc. R. Soc. Lond. B Biol. Sci.* **276**: 3053–3062, doi:10.1098/rspb.2009.0742
- ROY, K., D. JABLONSKI, AND J. W. VALENTINE. 2000. Dissecting latitudinal diversity gradients: Functional groups and clades of marine bivalves. *Proc. R. Soc. Lond. B Biol. Sci.* **267**: 293–299, doi:10.1098/rspb.2000.0999
- , ———, ———, AND G. ROSENBERG. 1998. Marine latitudinal diversity gradients: Tests of causal hypotheses. *Proc. Natl. Acad. Sci. USA* **95**: 3699–3702, doi:10.1073/pnas.95.7.3699
- RUDDIMAN, W. F. 1969. Recent planktonic foraminifera: Dominance and diversity in North Atlantic surface sediments. *Science* **164**: 1164–1167, doi:10.1126/science.164.3884.1164
- RUTHERFORD, S., S. D'HONDT, AND W. PRELL. 1999. Environmental controls on the geographic distribution of zooplankton diversity. *Nature* **400**: 749–753, doi:10.1038/23449
- SANDERS, H. L. 1968. Marine benthic diversity: A comparative study. *Am. Nat.* **102**: 243–282, doi:10.1086/282541
- SARMIENTO, J. L., AND OTHERS. 2004. Response of ocean ecosystems to climate warming. *Global Biogeochem. Cycles* **18**: GB3003, doi:10.1029/2003GB002134
- SCHOENER, T. W. 1974. Resource partitioning in ecological communities. *Science* **185**: 27–39, doi:10.1126/science.185.4145.27
- SMITH, W. H. F., AND D. T. SANDWELL. 1997. Global seafloor topography from satellite altimetry and ship depth surroundings. *Science* **277**: 1956–1962, doi:10.1126/science.277.5334.1956
- STENSETH, N. C., A. MYSTERUD, G. OTTERSEN, J. W. HURRELL, K.-S. CHAN, AND M. LIMA. 2002. Ecological effects of climate fluctuations. *Science* **297**: 1292–1296, doi:10.1126/science.1071281
- STEVENS, G. C. 1996. Extending Rapoport's rule to Pacific marine fishes. *J. Biogeogr.* **23**: 149–154, doi:10.1046/j.1365-2699.1996.00977.x
- TILMAN, D., D. WEDIN, AND J. KNOPS. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* **379**: 718–720, doi:10.1038/379718a0
- TOLEDO, F. A. L., K. B. COSTA, AND M. A. G. PIVEL. 2007. Salinity changes in the western tropical South Atlantic during the last 30 kyr. *Global Planet. Change* **57**: 383–395, doi:10.1016/j.gloplacha.2007.01.001
- TURNER, J. R. G. 2004. Explaining the global biodiversity gradient: Energy, area, history and natural selection. *Basic Appl. Ecol.* **5**: 435–448, doi:10.1016/j.baae.2004.08.004
- WHITE, B. N. 1994. Vicariance biogeography of the open-ocean Pacific. *Prog. Oceanogr.* **34**: 257–284, doi:10.1016/0079-6611(94)90012-4
- WORM, B., AND OTHERS. 2006. Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**: 787–790, doi:10.1126/science.1132294
- , M. SANDOW, A. OSCHLIES, H. K. LOTZE, AND R. A. MYERS. 2005. Global patterns of predator diversity in the open oceans. *Science* **309**: 1365–1369, doi:10.1126/science.1113399
- WRIGHT, D. H. 1983. Species–energy theory: An extension of species–area theory. *Oikos* **41**: 496–506, doi:10.2307/3544109

Associate editor: Michael R. Landry

Received: 10 November 2009

Accepted: 15 June 2010

Amended: 09 July 2010