Journal of Biogeography
January 2012, Volume 39, Issue 1, pages 114–129
http://dx.doi.org/10.1111/j.1365-2699.2011.02582.x

Biogeography of tuna and billfish communities

Gabriel Reygondeau[1, *], Olivier Maury[1], Gregory Beaugrand[2], Jean Marc Fromentin[3], Alain Fonteneau[1], Philippe Cury[1]

[1] Institut de Recherche pour le Développement (IRD), UMR EME 212, Centre de Recherches Halieutiques Méditerranéennes et Tropicales, Av. Jean Monnet, BP 171, 34203 Sète Cedex, France
[2] Centre National de la Recherche Scientifique (CNRS), Laboratoire d'Océanologie et de Géosciences, UMR LOG CNRS 8187, Station Marine, Université des Sciences et Technologies de Lille – Lille 1, BP 80, 62930 Wimereux, France
[3] Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), UMR EME 212, Centre de Recherches Halieutiques Méditerranéennes et Tropicales, Av. Jean Monnet, BP 171, 34203 Sète Cedex, France

*: Corresponding author : Gabriel Reygondeau, email address : gabriel.reygondeau@hotmail.fr

Abstract :

Aim The aims of this study were: (1) to identify global communities of tuna and billfish species through quantitative statistical analyses of global fisheries data; (2) to describe the spatial distribution, main environmental drivers and species composition of each community detected; and (3) to determine whether the spatial distribution of each community could be linked to the environmental conditions that affect lower trophic levels by comparing the partitions identified in this study with Longhurst's biogeochemical provinces.

Location The global ocean from 60° S to 65° N.

Methods We implemented a new numerical procedure based on a hierarchical clustering method and a nonparametric probabilistic test to divide the oceanic biosphere into biomes and ecoregions. This procedure was applied to a database that comprised standardized data on commercial longline catches for 15 different species of tuna and billfish over a period of more than 50 years (i.e. 1953–2007). For each ecoregion identified (i.e. characteristic tuna and billfish community), we analysed the relationships between species composition and environmental factors. Finally, we compared the biogeochemical provinces of Longhurst with the ecoregions that we identified.

Results Tuna and billfish species form nine well-defined communities across the global ocean. Each community occurs in regions with specific environmental conditions and shows a distinctive species composition. High similarity (68.8% homogeneity) between the spatial distribution of the communities of tuna and billfish and the biogeochemical provinces suggests a strong relationship between these species and the physical and chemical characteristics of the global ocean.

Main conclusions Despite their high tolerance for a wide range of environmental conditions, these highly migratory species are partitioned into clear geographical communities in the ocean at a global scale. The similarity between biogeochemical and biotic divisions in the ocean suggests that the global ocean is a mosaic of large biogeographical ecosystems, each characterized by specific environmental conditions that have a strong effect on the composition of the trophic web.

Keywords : Biogeochemical provinces ; global ocean ; *Istiophorus* ; *Katsuwonus* ; Macroecology ; *Makaira* ; marine biogeography ; *Tetrapturus* ; *Thunnus* ; *Xiphias*

INTRODUCTION

At the macroscale, one of the main goals of marine biogeography is to identify the spatial distribution of marine organisms and biodiversity and explain it by elucidating the relationship of abundance or species diversity with the environment (Lomolino *et al.*, 2006). Of the many types of classifications that have been proposed, almost all are based on either the physical structure of the global ocean, with respect to such parameters as temperature, stratification and circulation (Emery & Meincke, 1986; Cushing, 1989), or the spatial distribution of marine organisms (Beklemishev, 1961; McGowan, 1971). A distinctive approach was developed by Platt *et al.* (1991) who proposed classifying ecosystems in both the open ocean and continental shelves on the basis of the concentration of surface chlorophyll *a*, which is determined using the Continental Zone Colour Scanner (CZCS). The last approach was developed and refined by Longhurst (1998) at the global scale by the addition of remote sensing and *in situ* measurements of environmental parameters that affect phytoplankton growth and production (e.g., currents, nutrient concentrations, and stratification index). He used these data to partition the global ocean into what he called biogeochemical provinces (BGCPs) (Longhurst, 1998). The Longhurst scheme has been used as the basis for an atlas of oceanic ecosystems and has provided data on the specific conditions and fluctuations of environmental parameters that drive biogeochemical processes and affect the dynamics of marine species in each province. BGCPs have been used by the oceanographic community in ecological studies as a source of geographical data on the different types of environmental conditions found in oceans.

At a regional or basin scale, several studies have investigated the relevance of BGCPs by analysing the spatial distribution of marine species at different low trophic levels, from bacteria (Li *et al.*, 2004) to plankton (Gibbons, 1997; Beaugrand *et al.*, 2002; Woodd-Walker *et al.*, 2002; Alvain *et al.*, 2005). All of these studies indicated a significant match between

the BGCPs and the abundance of species, associations among species, or biodiversity. It was concluded that BGCPs represent specific environmental conditions that directly affect the abundance of species at lower trophic levels, due to their low physiological tolerance to variations in abiotic parameters (Richardson & Schoeman, 2004; IOCCG, 2009). Tuna and billfish species are oceanic top predators that are important for both ecological and economic reasons. They migrate over long distances during their biological cycle, with the result that they are widely distributed over the global ocean. However, previous analyses of data from commercial fisheries have consistently revealed differences in the spatial distribution between each tuna and billfish species (Fonteneau, 1998; Worm *et al.*, 2005). These results suggest that characteristic tuna and billfish species communities may be detected over the global ocean and relied to distinct oceanic biotopes. Nevertheless, no studies have investigated the appropriateness of using BGCPs to study the spatial distribution of these species and communities at higher trophic levels, which have physiologies that allow them to survive in a wider range of environmental conditions than lower trophic species.

The aims of this study were as follows: (1) to identify global communities of tuna and billfish species through quantitative statistical analyses of global fisheries data, (2) to describe the spatial distribution, main environmental drivers, and composition of species of each community detected, and (3) to determine whether the spatial distribution of each community detected could be linked to the environmental conditions that affect lower trophic levels. To achieve these goals, we investigated the biogeography of tuna and billfish species by applying a recently developed nonparametric methodology, using a dataset that comprised fisheries data on 15 species of tuna of the genera *Thunnus* and *Katsuwonus* and billfish of the genera *Xiphias*, *Makaira, Tetrapturus* and *Istiophorus,* obtained at a global scale (from 60° S to 65° N) over a period of 50 years. Using the results, we discuss the latitudinal division of the ecoregions that were detected by matching to specific environmental conditions the

physiological and behavioural characteristics of each tuna and billfish species used in the study. Furthermore, the model for partitioning the global ocean implemented using tuna and billfish was compared to the BGCPs of Longhurst (1998) to determine whether environmental conditions affect not only planktonic communities (Beaugrand *et al.*, 2002), but also the spatial distribution of communities at higher trophic levels.

MATERIALS AND METHODS

Biological data

Fisheries data were used to determine the spatial distribution of 15 species of tuna and billfish (Table 1). The data used were obtained from 180°W to 180°E and from 60°S to 65°N at a spatial resolution of 5° by 5°, and from 1953 to 2007, at monthly intervals. For each geographical cell and for each month, data on longline catches (number of fish) and fishing effort (number of hooks) were gathered for Taiwanese and Japanese fleets. These fleets were selected because each has a long fishing history and because they fish over a wider area than the fleets of other countries. The data on Japanese and Taiwanese fleets represent nearly 70% of the total catch of longline fisheries world-wide over the period from 1953 to 2007 (Fonteneau, 1998). Data were obtained from four Regional Fisheries Management Organizations (RFMOs; namely, the Indian Ocean Tuna Commission, International Commission for the Conservation of Atlantic Tunas, Inter-American Tropical Tuna Commission, and Western and Central Pacific Fisheries Commission) in a standardized form and are available on the Climate Impact on Oceanic Top Predators (CLIOTOP) website (http://vmmdst-proto.mpl.ird.fr/MDST/).

Given that only data on fish catches by the longline fishing technique (in which hooks are deployed on lines over a wide vertical range) were used in the study, not all

developmental stages of each species were sampled. Longline fisheries catch mostly adult fish (potential spawners). On the other hand, small tuna are less able to migrate than adults (Gunn & Block, 2001), so it is likely that any bias that exists is minor. Furthermore, due to confusion in the naming of some species in the RFMO database during the 1950s and 1960s, several comparable species (bluefin tuna, sailfish and blue marlin) were grouped together; these species are characterized by having similar levels of tolerance to environmental conditions, similar prey, and a rather low phylogenetic differentiation (Block & Stevens, 2001). This grouping reduced the number of species from 15 species (Table 1, name) to 11 grouped species (Table 1, code) and so also reduced the variance of the species matrix and anomalies in the spatial distribution of species, which could influence the numerical procedure described below (i.e. clustering analysis).

To infer the biogeography of the tuna and billfish community, the average abundance index of each of these species at a global scale has to be estimated from fisheries data. As a consequence, variations in the monthly average catches were not taken into account. The catch rate, measured by catch per unit of effort (CPUE), was calculated by dividing the sum of the annual catches by the sum of the annual efforts for each geographical area cell. The above-mentioned CPUE index is more accurate than the mean value of the standard CPUE index, which is calculated as the annual average catch divided by the associated effort and can be biased markedly by large changes in the catch of different species during the period studied. Due to the dense distribution of tuna populations (Sibert & Hampton, 2003) and the high mobility of longline fleets (Fonteneau, 1998), fleets tend to concentrate their efforts on patches with a high concentration of biomass until the yield decreases. Thereafter, fleets move to another region with a high concentration of biomass. As a consequence, the CPUE tends to remain spatially constant in different regions (Maury & Gascuel, 1999). This effect creates bias when the standard CPUE index (Polacheck, 2006) is used to estimate the spatial

distribution of fish by leading to an overestimation of the number of fish in regions with low densities and underestimation in regions with high densities, which in turn results in flat estimates of species distribution that vary little from place to place (Maury & Gascuel, 1999; Walters, 2003).

To avoid this bias and maintain a contrast among the spatially distributed mean estimates for the catch, we calculated the nominal CPUE (catch for 1000 hooks) over the entire study period as follows:

$$CPUE_{i,j,f,s} = \frac{\sum_t C_{i,j,f,s,t}}{\sum_t E_{i,j,f,s,t}}, \quad \textbf{(1)}$$

where C is the catch data accumulated for each geographical cell of longitude i and latitude j, for each time t, each species s, and each fleet f, and E is the number of hooks deployed in each geographical cell of longitude i and latitude j for each time t, species s, and fleet f. Visual inspection of the data and reference to the literature (Rouyer *et al.*, 2008) revealed that the patterns of variation in time series for tuna and billfish catches and CPUEs are affected strongly by the fishing method and fisheries, because of differences in the spatial movement of the fisheries and in the targeting of species. To avoid bias in the methodology, we considered the catches of the 11 grouped species separately from the Taiwanese and Japanese fleets, as suggested by Rouyer *et al.* (2008).

Environmental data

On the basis of the literature on the biogeography and ecology of tuna and billfish species (Longhurst, 1995; Block & Stevens, 2001), 12 environmental variables were selected and used in the study to characterize each region (see Appendix S1 in the Supporting Information).

The average annual data for sea surface temperature and selected biogeochemical variables (nitrate, silicate, phosphorus and salinity) were retrieved from the World Ocean Atlas 2005 (WOA05) (Antonov *et al.*, 2005; Boyer *et al.*, 2006; Rombouts *et al.*, 2009). The level of dissolved oxygen at the sea surface and at a depth of 100 metres was used as a proxy for the vertical distribution of this parameter (Prince & Goodyear, 2006). Mean annual values and standard deviations for the concentration of chlorophyll *a* were calculated using yearly averages derived from remote sensing data (SeaWiFS) for the period 1997 to 2007, inclusive. Data on sea surface currents were obtained from the OSCAR data access system of the National Oceanic and Atmospheric Administration (NOAA), from 60° S to 65° N (Rombouts *et al.*, 2009). The mean annual mixed layer depth was obtained from de Boyer Montégut *et al.* (2004) and bathymetry data from the General Bathymetric Chart of the Oceans (GEBCO) (Smith & Sandwell, 1997). The intensity and depth of thermoclines were taken from Reygondeau & Beaugrand (2011). The entire environmental dataset was aggregated on a 5° longitude by 5° latitude grid that extended from 180°W to 180°E and from 60°S to 65°N.

Analysis 1: Division of the oceanic biosphere into ecoregions

We identified ecoregions, defined here as subdivisions of the global ocean, on the basis of tuna and billfish CPUE using a procedure based on the methodology developed by Souissi *et al.* (2001) and Beaugrand *et al.* (2002) (Fig. 1). The procedure comprises five main steps:

Step 1: *Mathematical transformation of the data and numerical criteria for species selection*

CPUE data were log-transformed (i.e. $\log_{10}$ (CPUE+1)) to account for the heteroscedasticity of the data (Legendre & Legendre, 1998). Species that are harvested commonly by longline vessels [yellowfin tuna (YFT), bluefin tuna (BFT), bigeye tuna (BET), albacore tuna (ALB), swordfish (SWO), and striped marlin (MLS); see Table 1] were caught more frequently than

species that can be considered as longliner by-catch [skipjack tuna (SKJ), sailfish (SFA), black marlin (BLM), blue marlin (BUM), and short-billed spearfish (SSP); see Table 1]. With regard to species that were sampled less frequently, the spatial distribution derived from the longline CPUE might be biased and might affect the next step of the numerical procedure (i.e. the clustering method). In light of this possible bias, the spatial coverage of each species and its contribution to the total CPUE were computed (Appendix S2) for each of the 11 grouped species, taking into account the fact that Japanese and Taiwanese catches were analysed separately. The species caught were ranked on the basis of their number as a proportion of the total from matrix X (Fig. 1). A level of 0.5% of the total CPUE (as in Soussi *et al.,* 2001) was used to separate the species into two groups (Appendix S2): (1) the dominant species, whose relative contribution was greater than 0.5% and were found in more than 50% of the cells (Fig. 1, step 1, matrix I, 1183 geographic cells, seven species caught by the Japanese fleet and six by the Taiwanese fleet), and (2) secondary species, whose relative contribution was less than 0.5% and were found in less than 50% of the cells (Fig. 1, step 1, matrix II, 1183 geographic cells, four species caught by the Japanese fleet and five by the Taiwanese fleet).

Step 2: *Clustering of ecoregions that are based on tuna and billfish data*

The general hierarchical agglomerative clustering model of Lance & Williams (1967; β=-0.25, see Legendre & Legendre, 1998) was used on matrix I, which was composed of the data for the 13 dominant tuna and billfish species (Fig. 1, matrix I; 1183 geographic cells), to identify ecoregions on the basis of the CPUE within each geographical cell (Q mode) (Fig. 1, step 2). The clustering model was applied by computing a distance matrix from matrix I using the Bray–Curtis coefficient (Bray & Curtis, 1957). Hence, the geographical cells were agglomerated according to their Bray–Curtis distance computed over matrix I. The resulting dendrogram is presented in Fig. 2.

None of the numerical indices that have been proposed previously to determine an optimal cut-off level on such a dendrogram (Hardman-Mountford *et al.*, 2008; Guidi *et al.*, 2009) were appropriate for our methodology. Hence, different cut-off levels (Fig. 2, cut-off levels I to VI) were tested by a nonparametric methodology and examined visually as recommended by Legendre & Legendre (1998). After careful examination, we decided to use six cut-off levels at a Bray–Curtis distance of 9.5, 8.5, 5.7, 4.3, 3.8 and 3.2, respectively (Fig. 2), because the resulting maps of the spatial distribution of ecoregions detected at each cut-off levels provided a good compromise between global and local biogeochemical features.

Step 3: *Probabilities that a geographical cell belongs to a given ecoregion*

The probability that a given geographical cell (5° longitude × 5° latitude) belonged to a particular ecoregion was computed using a simplified version of the multiple response permutation procedure (MRPP, Mielke *et al.*, 1981) that was implemented recently in the nonparametric probabilistic ecological niche model (NPPEN; Beaugrand & Helaouët, 2008; Beaugrand *et al.*, 2011; Lenoir *et al.*, 2011). Mathematically, the NPPEN determines the probability that an observation that is composed of p variables (p, CPUE of the 13 dominant species of tuna and billfish in matrix I) belongs to a group $\boldsymbol{G_{m,p}}$ detected on the dendrogram at a given cut-off level (m, the number of geographical cells that vary between groups; p, the associated CPUE of the dominant species in matrix I), using the generalized Mahalanobis distance (Mahalanobis, 1936). The generalized Mahalanobis distance enables the correlation between variables (here the abundance of each species) to be taken into account (Ibañez, 1981):

$$D^2_{x,G} = \left(x - \overline{G}\right)' R^{-1} \left(x - \overline{G}\right), \quad (2)$$

where x is the vector of length p and represents the CPUE of the dominant species, $\boldsymbol{R}_{p,p}$ is the correlation matrix of the group $G_{m,p}$ (where m varies between groups), and $\overline{G}$ is the average cluster condition inferred from $\boldsymbol{G}_{m,p}$ (with $m < n$). The probability that a given geographical cell belongs to each group $G_{m,p}$, detected at each of the six cut-off levels according to the spatial distribution of the CPUE of the matrix I, was calculated for each geographical cell (n = 1188) (see Fig. 3, for cut-off level VI). Then, for each of the six cut-off levels, each geographical cell was assigned to the group, or ecoregion, to which it has the greatest likelihood of belonging at a given cut-off level (Fig. 1, step 4). The results for each cut-off level are mapped in Appendix S3, and summarized in Fig. 2 (cut-off levels II, V and VI) and Fig. 4 (cut-off level VI).

Step 4: *Calculation of the indicator value of each species and each group*

Indicator species that characterized each ecoregion were determined using the indicator value of Dufrêne & Legendre (1997) (Fig. 1, step 5). The indicator value is calculated by combining measures of specificity and fidelity. The specificity $A_{i,j}$ is the ratio of the mean abundance of species i in the geographical cells of group j ($N_{i,j}$) to the sum of the mean abundance of species i in all the groups (N_i):

$$A_{i,j} = \frac{N_{i,j}}{N_i} \qquad (3)$$

The fidelity $B_{i,j}$ is the ratio of the number of geographical cells in group j where species i is present ($S_{i,j}$) to the total number of pixels in this group (S_j):

$$B_{i,j} = \frac{S_{i,j}}{S_j} \qquad (4)$$

The indicator value ($V_{i,j}$) is calculated by multiplying the specificity and fidelity indices, because these two quantities represent independent information:

$$V_{i,j} = A_{i,j} \times B_{i,j} \times 100 \quad (5)$$

According to Rouyer *et al.* (2008), the clustering of ecoregions must take into account the differences in the behaviour of each fleet. In light of this recommendation, the species were divided into two groups to account for differences in fishing techniques between the Japanese and Taiwanese fleets. However, differences between fleets with respect to the distribution of species are not consistent in the case of the analysis of species composition (Rouyer *et al.*, 2008); hence, in our analysis, we considered information on both dominant and secondary species for both Japanese and Taiwanese fleets (see Table 1). As a consequence, we calculated the indicator value of Dufrêne & Legendre (1998) for 11 grouped species (see Table 1, code) and nine groups detected at cut-off level VI of the dendrogram. The results are presented in Fig. 3 as radar plots.

Analysis 2: Characterization of the environment in each ecoregion

We used principal components analysis (PCA; Jolliffe, 1986) to characterize the environmental conditions in all the ecoregions that were identified at the cut-off level VI of the dendrogram (see Step 2). The values of the 12 selected environmental factors (see Environmental data) were assigned to every geographical cell in each ecoregion and PCA were performed on these 12 variables (see Appendix S1) for each ecoregion separately. The environmental factors that contributed most to the first three principal components (PC) were identified for each ecoregion and are shown in Fig. 4.

Analysis 3: Comparison between the identified ecoregions and BGCP

Due to the fact that the separation of the oceans into BGCPs or ecoregions is semiquantitative and given the differences in the spatial resolution of the two methods of partitioning,

inferential or exploratory statistical tests were not used. Instead, a homogeneity analysis was conducted to compare the partitioning of the global ocean at cut-off level VI (Fig. 4) with the BGCP described by Longhurst (1995). The analysis quantifies the average homogeneity of a referential partition (i.e. BGCPs) with respect to that of another one (i.e. detected ecoregions). Here, the BGCPs determined by Longhurst were selected as reference partitions. Then, the average homogeneity of each ecoregion was quantified relative to each province.

For each BGCP, the total number of geographical cells in the BGCP (at a resolution of 5° × 5°) was calculated first (Table 2, number of cells). Then, the total number of geographical cells for each ecoregion in each BGCP was calculated. The percentage of each ecoregion that corresponded spatially to each BGCP was then determined [Table 2, ecoregion (%)]. To quantify the global similarity between the two types of biogeographical partition, the geographical cells in each ecoregion that were dominant in each BGCP were summed, then divided by the total number of geographical cells studied (n = 1188). The resulting number represents the average homogeneity of all the BGCPs represented in Fig. 4.

RESULTS

Ecological partitioning of the world's oceans on the basis of tuna and billfish catches

The ecoregions defined at each of the six cut-off levels are presented in Fig. 2 (cut-off levels II, V and VI) and Appendix S3. At cut-off levels I and II (distances of 9.5 and 8.5, respectively), three biomes (*sensu* Longhurst, 1998) were identified: those of subpolar and temperate, trade winds (or tropical), and westerly winds (or westerlies). At the next level (III, 5.7), a cluster that comprised the coastal group of geographical cells was detected. At the fourth level (at a distance of 4.3), the tropical ocean was shown to be composed of three ecoregions (tropical I and II and coastal tropics). Tropical I was the tropical open ocean, whereas tropical II was closer to the continental shelf. The coastal tropical ecoregion included

the Mexican, Indo-Pacific and Arabian seas, and areas located in the North Pacific from the Kurushio Current to the North Pacific transitional area. The Mexican, Indo-Pacific and Arabian seas were distinguished separately from the North Pacific areas at cut-off level V (3.8) and the Mexican seas (in the Tropical eastern Pacific) were detected separately from Arabian and Indo-Pacific seas at cut-off level VI (at a distance of 3.2). Cut-off level VI also showed that the oceanic subtropical gyres comprised two groups: one representing the seasonal extension of the gyres and the other their core areas. The temperate ecoregion was also identified at this level and comprised a cluster of cells that matched a small area in the Western Australian continental shelf in addition to the temperate ecoregions of the ocean.

Probability of ecoregions identified at cut-off level VI and associated indicator species

Ecoregions 1 (Mexican coast), 5 (Temperate), 6 (Western Australian continental shelf), and 7 (Transition zone) showed monospecific dominance (Fig. 3). In ecoregion 1 (Mexican coast), the sailfish was the predominant species. This region was identified most clearly on the Central Pacific-American continental shelf (Fig. 3). Bluefin species dominated ecoregion 5 (Temperate), which was located from the temperate to the subpolar areas of the open ocean and in some specific seas (the Java Sea, the Mediterranean Sea, and the Gulf of Mexico) (Fig. 3). Bluefin species also dominated ecoregion 6, which was located mainly over the Western Australian continental shelf. Swordfish dominated ecoregion 7. This region extended over a large area; it included the North Pacific transition zone and specific upwelling systems (Humboldt, California).

The ecoregions that represented oceanic gyres and tropical oceans were characterized by a more diverse group of species in which one or two species dominated (Fig. 3). Ecoregion 3 corresponded to the extension zones of oceanic gyres and was dominated by albacore tuna and swordfish. In contrast, ecoregion 4 was located in the core of the gyres, where albacore

tuna, together with striped marlin and swordfish, dominated. Ecoregion 8 (Tropical I) corresponded to cells in which bigeye tuna and striped marlin were present, whereas ecoregion 9 was dominated by yellowfin tuna, blue marlin, and, to a lesser extent, bigeye tuna. The spatial distribution of ecoregion 2 (Indo-Pacific and Arabian Seas) was located mainly in the Arabian Sea and Indo-Pacific seas (Fig. 3). A highly diversified fish community characterized this group and no single species was clearly predominant (Fig. 3).

Environmental factors that characterized the identified regions

PCA was used to identify the main environmental factors (see Appendix S1) that characterized each ecoregion at cut-off level VI (distance = 3.2). These factors are summarized in Fig. 4.

In the temperate ecoregion 5, the main characteristic environmental conditions were a low sea surface temperature and a high concentration of oxygen at the surface, and, to a lesser extent, a high concentration of chlorophyll *a*. The low sea surface temperature and high concentration of oxygen at the surface contributed to PC1, which explained 45.65% of the variance of the environmental matrix computed on the spatial distribution of ecoregion 5. The mean chlorophyll *a* concentration and standard deviation contributed to PC2, which explained 13.92% of the total variance of the environmental matrix. These same main environmental conditions were also detected in the transition zone ecoregion 7, which was located between the temperate and gyre biomes (PC1: 46.45% and PC2: 16.33% of the total variance). However, PC3 (12.86% of the total variance) showed a notable difference in that the standard deviation of the chlorophyll concentration contributed more to PC3, due to there being a greater degree of seasonal variation of the chlorophyll *a* concentration in this ecoregion than in the temperate one.

The environmental conditions of the two ecoregions that were associated with tropical oceanic gyres (ecoregions 3 and 4; Fig. 4) were characterized by: (1) intermediate values of sea surface temperature, which contributed to PC1 for both ecoregions; (2) high salinity, which contributed to PC1 for ecoregion 3 and PC3 for ecoregion 4; and (3) weak and deep stratification of the water column, which contributed to PC2 for both ecoregions. In ecoregions 3 and 4, PC1 contributed 32% and 33.7 % of the total variance, respectively; PC2 contributed 21.06% and 15.07% of the variance, respectively; and PC3 contributed 15.6% and 12.5% of the total variance, respectively. The analysis revealed that these two ecoregions were differentiated by the fact that the seasonal extension regions of the gyres contained higher concentrations of chlorophyll *a* than the core regions (chlorophyll *a* concentration contributed to PC2 for both ecoregions).

The tropical ecoregions (ecoregions 8 and 9, Fig. 4) were both characterized by warm sea surface temperatures, which contributed to PC3. PC3 accounted for 15.03% and 15.32% of the total variance in ecoregions 8 and 9, respectively. However, ecoregion 8 (Tropical I) was characterized mainly by strong stratification and a high velocity of oceanic currents (PC1: 30.52% of the total variance), and a low concentration of dissolved oxygen at 100 m (PC2: 22% of the total variance). In contrast, ecoregion 9 (Tropical II) was characterized strongly by a high mean concentration of chlorophyll *a* and shallow thermocline (PC1: 36.56% of the total variance).

The three coastal ecoregions (i.e. the Mexican coast, the Indonesian and Arabian seas, and the Western Australian continental shelf; ecoregions 1, 2, and 6, respectively) were characterized mostly by sea surface temperature, nutrient concentration, and bathymetry. Ecoregion 1 was characterized by shallow stratification (in terms of mixed layer depth and depth of the thermocline; PC1: 74.92% of the total variance) whereas the concentrations of silicate and nitrate, respectively, contributed more to PC2 and PC3 (11.9% and 9.72% of the

total variance, respectively). In ecoregion 2, a warm sea surface temperature and a high level of surface oxygen contributed to PC1 (33.42% of the total variance), whereas a high concentration of oxygen at 100 m and high mean chlorophyll *a* contributed to PC2 (25.2% of the total variance). In ecoregion 6, PC1 was explained mainly by the nutrient concentration (51.42% of the total variance) and PC2 by the bathymetry and the intensity of the thermocline (16.12% of the total variance).

DISCUSSION

One of the main goals of marine biogeography is to identify how the environment influences the spatial distribution of marine organisms and thus patterns of biodiversity. Achieving such a grand goal requires the integration of many very large datasets. The commercial fisheries dataset studied herein is one of the most exhaustive sets of observations on marine species at a global scale. Nevertheless, the fact that certain species are targeted and sampled non-homogeneously, as a result of the commercial interests of fleets, makes it difficult to interpret such a dataset for ecological purposes. Thus, we investigated the average biogeographical distribution of tuna and billfish species over the period from 1953 to 2007 using a newly developed methodology that enables the global ocean to be partitioned on the basis of the distribution of top predators. Specific environmental conditions and the community composition of each identified ecoregion were estimated. The results revealed latitudinal divisions among communities related to the physiological and behavioural capacities of species to cope with regional environmental conditions. Comparison of these biogeographical regions with BGCPs identified previously on the basis of environmental parameters revealed spatial overlap of 68.8% on average, which suggested that environmental conditions affect both the spatial distribution and the species composition of marine ecosystems from low to high trophic levels.

Issues related to fisheries data

The longline catches that we analysed in the study cover an area from 60° S to 65° N and yield the most exhaustive dataset available with respect to the spatial distribution of species of tuna and billfish that are exploited commercially (Fonteneau, 1998). However, as a result of the specificity of the longline method, only large individuals are caught; hence, the dataset does not cover the entire population of a given species, because individuals at the earlier stages of biological development will not be caught by the large hook. Indeed, the appropriateness and utility of using fisheries statistics for scientific purposes is debatable, for various reasons.

First, the data for some species may introduce bias, in that the species are considered to be by-catch or reported in the category of 'others'. In the present study, to address this problem, the contribution to the total CPUE and the spatial distribution of each species were examined (Appendix S2). Only dominant species, which showed the highest spatial coverage and contributed most to the CPUE, were used in the numerical procedure.

Second, fisheries statistics are often biased due to changes in the spatial distribution of the fishing, as well as temporal variation. Given that fishing fleets usually explore the regions with the highest abundance of targeted fish and, once a given region has been exploited sufficiently, move quickly to another region to maintain optimal productivity, CPUE series often show little variation between different areas and do not reflect the heterogeneity of the resource (Fonteneau & Richard, 2003). Furthermore, studies that focused on global or local variations in catch rates over time (Myers & Worm, 2003) have been criticized strongly due to the large amount of variation in commercial fisheries statistics (Polacheck, 2006). Indeed, Polacheck (2006) stressed the difficulty in using CPUE indices due to the fluctuations in the relationship between fishing effort and abundance of the targeted fish populations over time,

especially for longline fisheries. Furthermore, Rouyer *et al.* (2008) showed that the patterns of variation in fisheries time series for catches of tuna and billfish varied primarily in terms of the spatial distribution of the species and secondarily in relation to the fishing equipment used. The authors concluded that the patterns of variation in fisheries time series fail to represent accurately the underlying dynamics of populations of large pelagic fish. Consequently, the fisheries data used in the present study cannot be used readily for ecological purposes.

To characterize the biogeography of large pelagic species, we have proposed a new CPUE index that captures the spatial distribution of the species studied over the entire period of study. Over the period studied, these species were under- or fully exploited, but not overexploited, and so their spatial distributions arguably were not affected by the increase in pressure from fisheries. Only bluefin tuna is classified as having been overexploited: this is the case for the southern bluefin tuna and the West Atlantic bluefin tuna since the late 1970s and for the East Atlantic and possibly the Pacific bluefin tuna since the 1990s. Furthermore, Rouyer *et al.* (2008) observed that the effect of changes in environmental conditions on the spatial distribution of large pelagic species is detected predominantly on a time scale of decades, rather than a shorter time scale; hence, long time series are required to detect long-term variations. The combined use of a long time period and the CPUE index allowed us to estimate the complete average spatial distribution of each species over the period studied in accordance with previous findings (Fonteneau, 1998).

Third, commercial longline fleets from different countries do not always target the same species and do not cover the same spatial area and period. To reduce this bias, only data from Taiwanese and Japanese fleets were used and they were analysed separately, as recommended by Rouyer *et al.* (2008). In addition, both fleets were observed to change the depth at which longline hooks were deployed to catch different species at the same location

(Maunder *et al.*, 2006). These changes in fishing strategy might bias the dataset by altering the spatial distribution of some species and thereby affecting the outputs of the clustering analysis. To minimize the effects of this potential bias, only data for the cumulative catch per geographical cell for well-sampled species (dominant species, see Fig. S1) were considered for the cluster analysis.

Finally, the longline dataset used in this study has a low spatial resolution (5° × 5° geographical cells) and provides data on tuna and billfish which are found only rarely on continental shelves. As a consequence, the results for ecoregions near coasts need to be interpreted with caution.

Biogeography of tuna and billfish

On the basis of the results, the tropical ocean can be divided into two ecoregions (Figs 4 & 5, tropical I and II). Tropical I corresponded to areas that are dominated by bigeye tuna, which is a warm-water fish but has a much better ability to thermoregulate (both physiologically and behaviourally) than yellowfin tuna, especially with respect to the large adults targeted by longline fleets (Brill *et al.*, 2005). This capacity, combined with the ability to hunt in dark waters (Nicol & Somiya, 1989; Somiya *et al.*, 2000) and a tolerance to relatively anoxic waters, extends their feeding grounds well below the thermocline (from 300 to 500 m), where abundant mesopelagic organisms are concentrated by advection away from upwelling regions (Musyl *et al.*, 2003). The second ecoregion (Tropical II) corresponded to the reproduction and nursery areas of yellowfin tuna in three oceans (Maury *et al.*, 2001). This species, which uses its eyesight to detect prey, is also in general distributed in warm tropical waters where a shallow thermocline enhances phytoplankton production and concentrates epipelagic prey in euphotic surface waters (Kirby, 2001) (Fig. 5).

Albacore tuna prefer temperate waters (Laurs *et al.*, 1984), which restricts their range to oceanic gyres and adjacent regions (Penney *et al.*, 1998). As adults, this smaller species generally feeds in gyres (Bertrand *et al.*, 2002) where the mixed layer depth is deep, and where the temperature falls gradually as the water depth increases (Tomczak & Godfrey, 2003). Our classification split the gyres into two regions: (1) ecoregion 3, which corresponded to the feeding areas of juvenile albacore tuna (seasonal extent of gyres), and (2) ecoregion 4, which corresponded to the core of the gyres and was inhabited by adult albacore tuna (Bard *et al.*, 1998). The swordfish occurred in more temperate regions than bigeye and yellowfin tuna (transition zone in Figs 4 & 5). Although it cannot thermoregulate well, the swordfish has high thermal inertia and can see well in dark waters. These characteristics allow this species to dominate productive areas (i.e. areas of high primary and secondary production), such as the transition zone ecoregion (Longhurst, 1995) that exists between subpolar regions and the margins of oceanic gyres (Kurushio and Gulf Stream currents).

The temperate ecoregion (ecoregion 5) was dominated by bluefin tuna. Bluefin tuna have the highest thermoregulatory capacity of all tuna and billfish species and inhabit temperate (and even subpolar) waters where they can find plentiful food resources. Atlantic bluefin tuna (*Thunnus thynnus*) also migrate to the Mediterranean Sea and the Gulf of Mexico to reproduce, which explains the extension of the temperate ecoregion (Fig. 4).

The last three proposed ecoregions may be found in coastal areas. The putative Western Australian continental shelf ecoregion matched the Japanese fishing grounds during the feeding migrations of southern bluefin tuna from the area in which they reproduce in the Java Sea to the south subpolar convergence in the Indian Ocean where they feed (Proctor *et al.*, 1995). The shelves of the putative Indo-Pacific and Arabian Seas ecoregion show high phytoplankton productivity, which results in a high availability of food for forage species and hence many tuna and billfish species feed in these regions (Price, 2002). Moreover, it is easier

to catch fish in these regions, thanks to a shallow oxycline that pushes predatory fish near the surface (Prince & Goodyear, 2006). The last proposed ecoregion, located along the Pacific Mexican coast, was dominated by sailfish, which are known to reside in this area, close to the site at which they reproduce on the continental shelf. Particular caution is warranted when considering the validity of this ecoregion due to the low quality of the fisheries statistics describing sailfish.

Tuna and billfish communities

Nine major ecoregions, characterized by specific assemblages of species, were detected (Figs 3 & 4, radar plot). Analysis of the species composition of each ecoregion (Fig. 3) revealed three types of pattern: (1) single-species dominance, (2) multispecies dominance, and (3) a diversified community without clear dominance. The last of these, with no apparent competitive exclusion, can be related to the ecological characteristics of tuna and billfish species that may share the same area by occupying different depths (Young *et al.*, 2010). These specific characteristics are exploited by different longline fleets, which target different species in the same region by changing the depth to which the hooks are lowered (Yang & Gong, 1987; Nakano *et al.*, 1997). As a consequence, each ecoregion that represents a specific community needs to be considered in three dimensions to understand fully the assemblage of species. For this purpose, we propose a conceptual scheme that is based on the ecoregion that predominates at each latitude (Fig. 4) and is refined using the available literature on the vertical division and movement of tuna and billfish species (Block & Stevens, 2001; Fig. 5).

Furthermore, the migratory behaviour of large pelagic fish can affect the composition of each community directly. In fact, some of the species studied show a clear pattern of migration between feeding grounds and breeding grounds (Walli *et al.*, 2009). As a consequence, the species assemblages in a given area might vary in a seasonal manner,

depending on whether migrant reproductive adults are present (Fromentin & Powers, 2005). The biogeographical partitioning proposed herein (Fig. 4) does not fully take into account the seasonal migration of the species because some data are limited due to restrictions on the access of longliners to certain areas, such as the western spawning ground of the Atlantic bluefin tuna since the early 1980s. Therefore, the biogeography of large pelagic fish proposed here has to be considered as an average condition over the period studied.

Comparison of the described ecoregions with the BGCPs

The importance of BGCPs (Longhurst, 1995) for research in ecology has already been supported by several studies (Beaugrand *et al.*, 2002; Woodd-Walker *et al.*, 2002; Li *et al.*, 2004). However, all these studies were performed on stenoecious species, which are sensitive to small changes in the environment and for which fluctuations in abundance can be attributed directly to such changes (Reygondeau & Beaugrand, 2011). The partitioning of the global ocean that we report herein, which is based on the distribution of large pelagic fish, were compared with the BGCPs (Longhurst, 1995) to determine the ability of the BGCPs to discriminate between the spatial distributions of species at high trophic levels at the global oceanic scale. The homogeneity between the spatial distribution of the BGCPs and the ecoregions that we identified reached 68.8% on average (Table 2). This result indicates that specific environmental conditions, captured by the BGCPs, partially control the spatial distribution and co-presence among species of large pelagic fish.

The level of homogeneity between our proposed ecoregions and Longhurst's BGCPs was lower in coastal provinces and higher in open ocean provinces (Table 2). Furthermore, the mean level of mismatch of 30% detected between these two schemes might be related to the fact that the BGCPs have a geographical resolution of 1° × 1° whereas our ecoregions are resolved at only 5° × 5°. As a consequence, many geographical cells of the ecoregions

identified herein might belong to two or three different BGCPs, which would lead to some biogeographical boundaries being missed, moved, or diluted. The numerous biases that affect commercial fisheries data might also contribute to imprecise boundaries between ecoregions, for example because the longline data may lead to the spatial distribution of some species being underestimated. In addition, even if Longhurst's BGCPs remain the most accepted scheme of partitioning in marine biogeography, some of the boundaries that it draws may be open to revision, in light of new observations made since the BGCPs model of partitioning was developed (Longhurst, 1995). More recent observations might explain some of the differences between our partitioning results for top predators and the BGCP, for example, in the Caribbean Sea, the South Pacific subtropical gyre, the North Atlantic subtropical gyre, and the monsoon band. Finally, the patterns of migration of these top predators and their physiology (specifically, their capacity to thermoregulate) make them less dependent on, or sensitive to, environmental variations than species at lower trophic levels. Given these considerations, the strong overlap between BGCPs and the spatial distributions of large pelagic fish is rather surprising.

According to the general definition of the BGCPs, each partition represents specific environmental conditions that differ from those of the adjacent province, and within a given province the pattern of variation of key biogeochemical parameters can be predicted (Longhurst, 2005). As a consequence, each BGCP can be considered to reflect a characteristic habitat to which species have adapted and the environment in which the species can develop and maintain populations (cf. Hutchinson, 1957). This assumption is used generally to explain the appropriateness of using the BGCPs for ecological studies of species at lower trophic levels and stenoecious species (from bacteria to mesozooplankton; see Beaugrand *et al.*, 2002; Woodd-Walker *et al.*, 2002; Li *et al.*, 2004). Our results revealed a high degree of similarity between Longhurst's BGCPs and the spatial distribution of species at high trophic levels (i.e.

large pelagic species), even though these organisms are thought in general to be affected less by variation in environmental conditions than species at lower levels (Block & Stevens, 2001). Thus, we infer that this relative match between BGCPs and the spatial distribution of communities of large pelagic species may result from bottom-up processes (Cury et *al.*, 2008). First, the main biological processes (e.g. survival of larvae) and life history traits of large pelagic fish are affected directly by environmental parameters, such as temperature (Block & Stevens, 2001). Second, the abundance of top predators depends strongly on the presence of foraging species, which in turn are linked closely to environmental conditions (Lehodey *et al.* 1998; Beaugrand *et al.*, 2002). As a consequence, each BGCP may represent specific environmental envelopes (i.e. characteristic intervals of variation in environmental conditions that affect the dynamics of marine species) in which adapted species at lower trophic levels are present. Thus, the BGCPs can also be related to the spatial distribution of top predator species that in turn reflect their varied feeding preferences and differences in tolerance to their environment.

CONCLUSIONS

The use of an extensive commercial fisheries dataset on tuna and billfish characterized by a wide range of environmental tolerance has revealed clear spatial partitioning of the global ocean into well-defined communities (i.e. ecoregions). We suggest that this spatial division of the oceans results from the spatial distribution of the species on the basis of their different physiological and behavioural adaptations to the environment. Although previous studies have already demonstrated a match between the distribution of other taxa and the BGCPs of Longhurst (1995), the species studied were generally ectothermal and thus were affected more directly by the local environment. In our study, the similarities identified between the partitioning of the ocean with respect to top predators and the BGCP shows the strong

influence of the environment on the species composition and spatial distribution of apical species. Therefore, BGCPs correspond to certain physical conditions (biotopes) in which marine species can maintain their populations and constitute specific trophic webs. Thus, the BGCPs seem to provide a geographical framework that is relevant when studying the possible effects of climate change on the abundance of top predators and which will enable better management and improved conservation of marine resources.

ACKNOWLEDGEMENTS

The authors are grateful to the referees and the editors who helped to improve the paper. F. Menard, S. Bonhommeau, E. Chassot, I. Rombouts and especially F. Ibañez improved our understanding of the methodology used in this study and the interpretation of the results. We are grateful to A. Longhurst for his crucial help in the interpretation of the results. The authors are also grateful to A.C. Gandrillon for her help in the editing of this paper and to P. Lopez for his help in the development of the figures. To obtain the complete biological dataset, please contact Laurent.floch@ird.fr or Emmanuel.chassot@ird.fr. For any other information or requests about the dataset used in the study, please contact the authors of the article.

REFERENCES

Alvain, S., Moulin, C., Dandonneau, Y. & Bréon, F.M. (2005) Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Research Part I*, **52,** 1989-2004.

Antonov, J.I., Levitus, S. & Boyer, T.P. (2005) Thermosteric sea level rise, 1955-2003. *Geophysical Research Letters*, **32,** L12602.

Bard, F.X., Bach, P. & Josse, E. (1998) Habitat, écophysiologie des thons: quoi de neuf depuis 15 ans? *Collective volume of scientific papers: international commission for the conservation of Atlantic tunas,* **50,** 319-342.

Beaugrand, G. & Helaouët, P. (2008) Simple procedures to assess and compare the ecological niche of species. *Marine Ecology Progress Series*, **363,** 29-37.

Beaugrand, G., Ibañez, F., Lindley, J.A. & Reid, P.C. (2002) Diversity of calanoid copepods in the North Atlantic and adjacent seas: species associations and biogeography. *Marine Ecology Progress Series,* **232,** 179-195.

Beaugrand, G., Lenoir, S., Ibañez, F. & Manté, C. (2011) A new model to assess the probability of occurrence of a species based on presence-only data. *Marine Ecology Progress Series,* **424,** 175-190.

Beklemishev, C.W. (1961) On the spatial structure of plankton communities in dependence of oceanic circulation. Boundaries of ranges of oceanic plankton animals in the North Pacific. *Okeanologia*, **5,** 1059-1072.

Bertrand, A., Josse, E., Bach, P., Gros, P. & Dagorn, L. (2002) Hydrological and trophic characteristics of tuna habitat: consequences on tuna distribution and longline catchability. *Canadian Journal of Fisheries and Aquatic Sciences*, **59,** 1002-1013.

Block, B.A. & Stevens, E.D. (2001) *Tuna: physiology, ecology, and evolution*. Academic Press, San Diego.

de Boyer Montégut, C., Madec, G., Fischer, A.S., Lazar, A. & Ludicone, D. (2004) Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, **109,** C12003.

Boyer, T.P., Antonov, J.I., Garcia, H., Johnson, D.R., Locarnini, R.A., Mishonov, A.V., Pitcher, M.T., Baranova, O.K. & Smolyar, I. (2006) *World Ocean Database 2005, Chapter 1: Introduction, NOAA atlas NESDIS 60*. US Government Printing Office, Washington, DC.

Bray, R.J. & Curtis, J.T. (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, **27,** 325-347.

Brill, R.W., Bigelow, K.A., Musyl, M.K., Fritsches, K.A. & Warrant, E.J. (2005) Bigeye tuna (*Thunnus obesus*) behavior and physiology and their relevance to stock assessments and

fishery biology. *Collective volume of scientific papers: international commission for the conservation of Atlantic tunas,* **57,** 142-161.

Cushing, D.H. (1989) A difference in structure between ecosystems in strongly stratified waters and in those that are only weakly stratified. *Journal of Plankton Research*, **11,** 1-13.

Cury, P.M., Shin, Y.J., Planque, B., Durant, J.M., Fromentin, J.M., Kramer-Schadt, S., Stenseth, N.C., Travers, M. & Grimm, V. (2008) Ecosystem oceanography for global change in fisheries. *Trends in Ecology and Evolution*, **23**, 338-346.

Dufrêne, M. & Legendre, P. (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*, **67,** 345-366.

Emery, W.J. & Meincke, J. (1986) Global water masses: summary and review. *Oceanologica Acta*, **9,** 383-391.

Fonteneau, A. (1998) *Atlas of tropical tuna fisheries*. Edition ORSTOM, Paris.

Fonteneau, A. & Richard, N. (2003) Relationship between catch, effort, CPUE and local abundance for non-target species, such as billfishes, caught by Indian Ocean longline fisheries. *Marine and Freshwater Research*, **54,** 383-392.

Fromentin, J.M. & Powers, J.E. (2005) Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. *Fish and Fisheries*, **6,** 281-306.

Gibbons, M.J. (1997) Pelagic biogeography of the south Atlantic Ocean. *Marine Biology*, **129,** 757-768.

Graham, J.B. & Dickson, K.A. (2004) Tuna comparative physiology. *Journal of Experimental Biology*, **207,** 4015-4024.

Guidi, L., Ibañez, F., Calcagno, V. & Beaugrand, G. (2009) A new procedure to optimize the selection of groups in a classification tree: applications for ecological data. *Ecological Modelling*, **220,** 451-461.

Gunn, J. & Block, B. (2001) Advances in acoustic, archival, and satellite tagging of tunas. *Fish Physiology*, **19,** 167-224.

Hardman-Mountford, N.J., Hirata, T., Richardson, K.A. & Aiken, J. (2008) An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Remote Sensing of Environment,* **112,** 3341-3352.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22,** 415-427.

Ibañez, F. (1981) Immediate detection of heterogeneities in continuous multivariate, oceanographic recordings. Application to time series analysis of changes in the Bay of Villefranche sur Mer. *Limnology and Oceanography*, **26,** 336- 349.

IOCCG (2009) Remote sensing in fisheries and aquaculture. *Reports of the International Ocean-Colour Coordinating Group* (ed. by M.H. Forget, V. Stuart and T. Platt), pp. 0-95. International Ocean-Colour Coordinating Group, Dartmouth.

Jolliffe, I.T. (1986) *Principal component analysis*. Springer-Verlag, New York.

Kirby, D.S. (2001*) On the integrated study of tuna behaviour and spatial dynamics: tagging and modelling as complementary tools. Electronic tagging and tracking in marine fisheries* (ed. by Sibert J. R. and J.L. Nielson), pp. 407-420. Kluwer Academic Publishers, Boston, MA.

Lance, G.N. & Williams, W.T. (1967) A general theory of classificatory sorting strategies. I. Hierarchical systems. *Computer Journal*, **9,** 373-380.

Laurs, R.M., Fiedler, P.C. & Montgomery, D.R. (1984) Albacore tuna catch distributions relative to environmental features observed from satellites. *Deep-Sea Research Part I*, **31,** 1085-1099.

Legendre, P. & Legendre, L. (1998) *Numerical ecology*, 2nd edn. Elsevier Science BV, Amsterdam.

Lehodey, P., Andre, J.M., Bertignac, M., Hampton, J., Stoens, A., Menkes, C., Memery, L. & Grima, N. (1998) Predicting skipjack tuna forage distributions in the equatorial Pacific using a coupled dynamical bio-geochemical model. *Fisheries Oceanography*, **7,** 317-325.

Lenoir, S., Beaugrand, G. & Lécuyer, E. (2011) Modelled spatial distribution of marine fish and projected modifications in the north Atlantic Ocean. *Global Change Biology*, **17,** 115-129.

Li, W.K.W., Head, E.J.H. & Glen Harrison, W. (2004) Macroecological limits of heterotrophic bacterial abundance in the ocean. *Deep-Sea Research Part I*, **51,** 1529-1540.

Lomolino, M.V., Riddle, B.R. & Brown, J.H. (2006) *Biogeography*, 3rd edn, Sunderland, MA.

Longhurst, A. (1995) Seasonal cycles of pelagic production and consumption. *Progress in Oceanography*, **36,** 77-167.

Longhurst, A. (1998) *Ecological geography of the sea*. Academic Press, London.

Mahalanobis, P.C. (1936) On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2,** 49-55.

Maunder, M.N., Sibert, J.R., Fonteneau, A., Hampton, J., Kleiber, P. & Harley, S.J. (2006) Interpreting catch per unit effort data to assess the status of individual stocks and communities. *ICES Journal of Marine Science*, **63,** 1373-1385.

Maury, O. & Gascuel, D. (1999) SHADYS ('Simulateur Halieutique de DYnamiques Spatiales), a GIS based numerical model of fisheries. Example application: the study of a marine protected area. *Aquatic Living Resources*, **12,** 77-88.

Maury, O., Gascuel, D., Marsac, F., Fonteneau, A. & Rosa, A.L.D. (2001) Hierarchical interpretation of nonlinear relationships linking yellowfin tuna (*Thunnus albacares*) distribution to the environment in the Atlantic Ocean. *Canadian Journal of Fisheries and Aquatic Sciences*, **58,** 458-469.

McGowan, J.A. (1971) *Ocean biogeography of the Pacific. The micropaleontology of the oceans* (ed. by F.A. Riedl), pp. 3-74. Cambridge University Press, Cambridge.

Mielke, P.W., Berry, K.J. & Brier, G.W. (1981) Application of multi response permutation procedures for examining seasonal changes in monthly mean sea-level pressure patterns. *Monthly Weather Review*, **109,** 120-126.

Musyl, M.K., Brill, R.W., Boggs, C.H., Curran, D.S., Kazama, T.K. & Seki, M.P. (2003) Vertical movements of bigeye tuna (*Thunnus obesus*) associated with islands, buoys, and seamounts near the main Hawaiian islands from archival tagging data. *Fisheries Oceanography*, **12,** 152-169.

Myers, R.A. & Worm, B. (2003) Rapid worldwide depletion of predatory fish communities. *Nature*, **423,** 280-283.

Nakano, H., Okazaki, M. & Okamoto, H. (1997) Analysis of catch depth by species for tuna longline fishery based on catch by branch lines. *Bulletin National Research Institute of Far Seas Fisheries,* **24,** 43-62.

Nicol, J.A.C. & Somiya, H. (1989) *The eyes of fishes.* Clarendon Press, Oxford.

Penney, A.J., Yeh, S.Y., Kuo, C.L. & Leslie, R.W. (1998) Relationships between albacore (*Thunnus alalunga*) stocks in the southern Atlantic and Indian oceans. *Collective volume of scientific papers: international commission for the conservation of Atlantic tunas*, **50,** 261-272.

Platt, T., Caverhill, C. & Sathyendranath, S. (1991) Basin-scale estimates of oceanic primary production by remote sensing: the north Atlantic. *Journal of Geophysical Research*, **96,** C8.

Polacheck, T. (2006) Tuna longline catch rates in the Indian Ocean: did industrial fishing result in a 90% rapid decline in the abundance of large predatory species? *Marine policy*, **30,** 470-482.

Price, A.R. (2002) Simultaneous 'hotspots' and 'coldspots' of marine biodiversity and implications for global conservation. *Marine Ecology Progress Series*, **241,** 23-270.

Prince, E.D. & Goodyear, C.P. (2006) Hypoxia-based habitat compression of tropical pelagic fishes. *Fisheries Oceanography*, **15,** 451-464.

Proctor, C.H., Thresher, R.E., Gunn, J.S., Mills, D.J., Harrowfield, I.R. & Sie, S.H. (1995) Stock structure of the southern bluefin tuna (*Thunnus maccoyii*): an investigation based on probe microanalysis of otolith composition. *Marine Biology*, **122,** 511-526.

Reygondeau, G. & Beaugrand, G. (2011) Water column stability and *Calanus finmarchicus*. *Journal of Plankton Research*, **33,** 119-136.

Richardson, A.J. & Schoeman, D.S. (2004) Climate impact on plankton ecosystems in the northeast Atlantic. *Science*, **305,** 1609-1612.

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S. & Legendre, L. (2009) Global latitudinal variations in marine copepod diversity and environmental factors. *Proceedings of the Royal Society B: Biological Sciences*, **276,** 3053–3062.

Rouyer, T., Fromentin, J.M., Stenseth, N.C. & Cazelles, B. (2008) Analysing multiple time series and extending significance testing in wavelet analysis. *Marine Ecology Progress Series*, **359,** 11-23.

Sibert, J. & Hampton, J. (2003) Mobility of tropical tunas and the implications for fisheries management. *Marine Policy*, **27,** 87-95.

Smith, W.H.F. & Sandwell, D.T. (1997) Global sea floor topography from satellite altimetry and ship depth soundings. *Science*, **277,** 1956–1962.

Somiya, H., Takei, S. & Mitani, I. (2000) Guanine and its retinal distribution in the tapetum of the bigeye tuna, *Thunnus obesus. Ichthyological Research*, **47,** 367-372.

Souissi, S., Ibañez, F., Ben Hamadou, R., Boucher, J., Cathelineau, A.C., Blanchard, F. & Poulard, J.C. (2001) A new multivariate mapping method for studying species assemblages and their habitats: example using bottom trawl surveys in the bay of Biscay (France). *Sarsia*, **86,** 527-542.

Tomczak, M. & Godfrey, J.S. (2003) *Regional oceanography: an introduction*, 2nd edn. Daya Publishing House, Delhi.

Walli, A., Teo, S.L.H., Boustany, A., Farwell, C.J., Williams, T., Dewar, H., Prince, E.D. & Block, B.A. (2009) Seasonal movements, aggregations and diving behavior of Atlantic bluefin tuna (*Thunnus thynnus*) revealed with archival tags. *PLoS ONE*, **4,** e6151.

Walters, C. (2003) Folly and fantasy in the analysis of spatial catch rate data. *Canadian Journal of Fisheries and Aquatic Sciences*, **60,** 1433-1436.

Woodd-Walker, R.S., Ward, P. & Clarke, A. (2002) Large-scale patterns in diversity and community structure of surface water copepods from the Atlantic Ocean. *Marine Ecology Progress Series*, **236,** 189-203.

Worm, B., Sandow, M., Oschlies, A., Lotze, H.K. & Myers, R.A. (2005) Global patterns of predator diversity in the open oceans. *Science*, **309,** 1365-1369.

Yang, W.S. & Gong, Y. (1987) The vertical distribution of tunas and billfishes, and fishing efficiency between Korean regular and deep longlines in the Atlantic Ocean. *Collective volume of scientific papers: international commission for the conservation of Atlantic tunas,* **24,** 184-187.

Young, J., Lansdell, M., Campbell, R., Cooper, S., Juanes, F. & Guest, M. (2010) Feeding ecology and niche segregation in oceanic top predators off eastern Australia. *Marine Biology*, **157,** 2347-2358.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Characteristics of the twelve environmental factors selected.

Appendix S2 Numerical criteria of selection of dominant and secondary species.

Appendix S3 Maps of the spatial distribution of each group for the six cut-off levels selected by the NPPEN procedure, and map of the BGCPs proposed by Longhurst (1995).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

Gabriel Reygondeau is a PhD student at the Centre of Fisheries Research (Centre de Recherche Halieuthique) in the research team 'Exploited marine ecosystems' (EME 212 Sète, France). His research aims to identify and characterize the effects of climate change on a global scale and its consequences for the biogeography of marine ecosystems.
The objectives of the research team (http://www.umr-eme.org) are to study the effects of global change on marine ecosystems, to investigate how such systems are and should be governed and exploited, and to predict changes in global ecosystems on the basis of IPCC scenarios.

Author contributions: G.R., O.M., A.F. and P.C. conceived the idea; G.R., O.M. and A.F. collected the data; and G.R., G.B. and J.M.F. analysed the data. G.R. led the writing with every co-author contributing.

Editor: Michael Dawson

Figure Captions

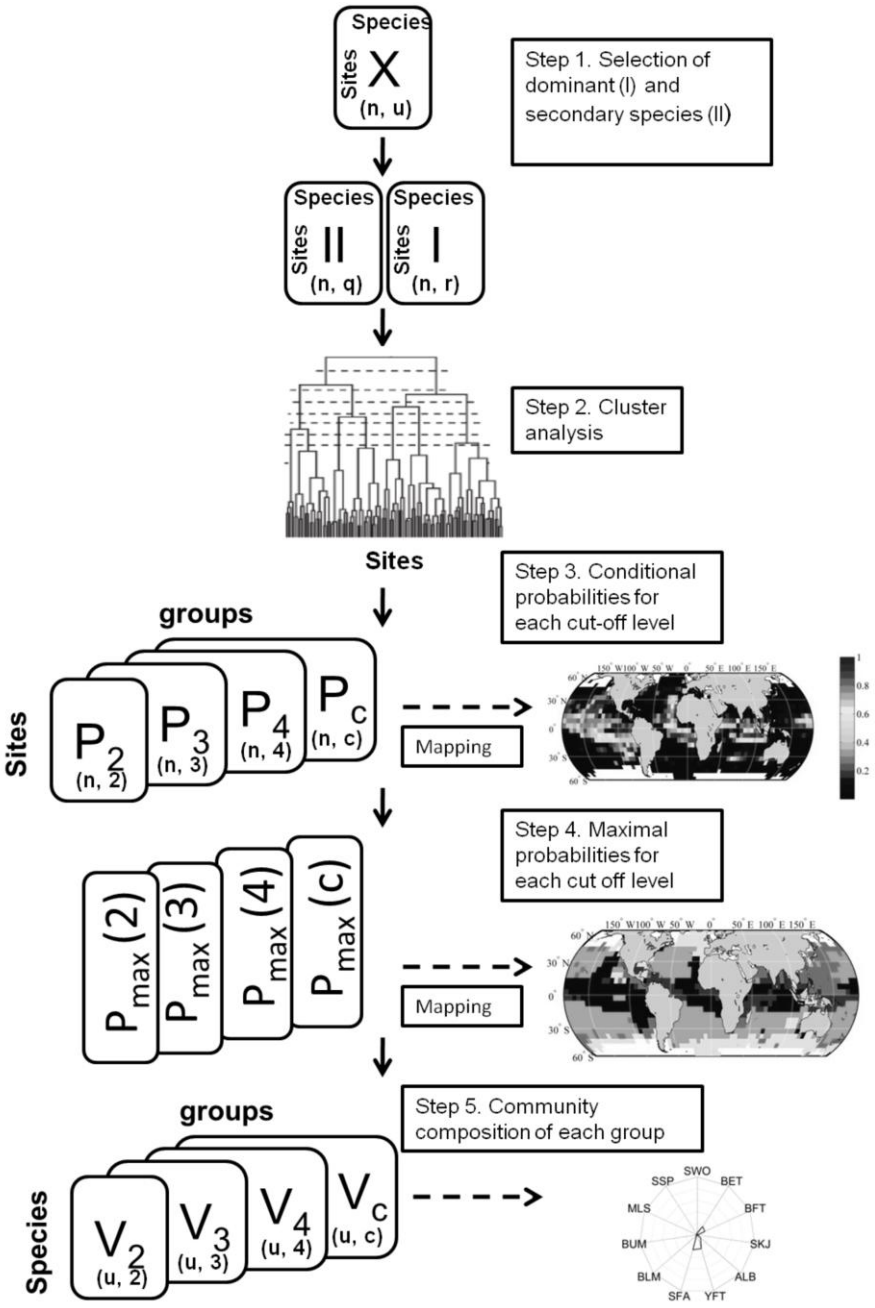

Figure 1 Schematic representation of the numerical procedures used in this study (see Materials and Methods for details). The clustering analysis (step 2) is shown in detail in Fig. 2. A map of the probability of occurrence of each group cited in step 3 is shown in Fig. 3 for all groups at cut-off level VI. A map of the ecoregions obtained with the maximal probability (step 4) is shown in Fig. 2 (cut-off levels II, V, and VI), Fig. 4 (cut-off level VI), and Appendix S3 (all cut-off levels). The radar plots that indicate the species composition of each group are shown in Fig. 3 and Fig. 4.

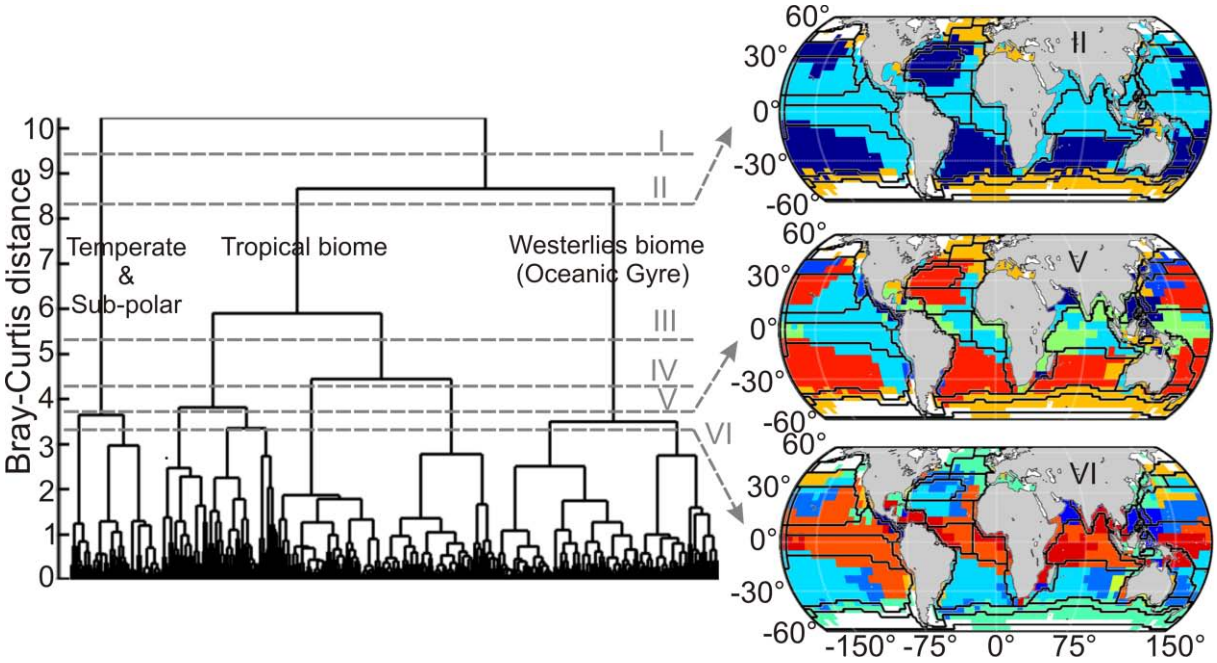

Figure 2 Identification of ecoregions on the basis of tuna and billfish data. The Dendrogram derived from the cluster analysis performed on the matrix (I, 1189 geographical cells and 13 species) showing the cut-offs at the six different levels that were tested (dashed lines). The names of each cut-off level are only qualitative and do not refer to the number of groups detected in the resulting partitioning. The projection used is Eckert IV.

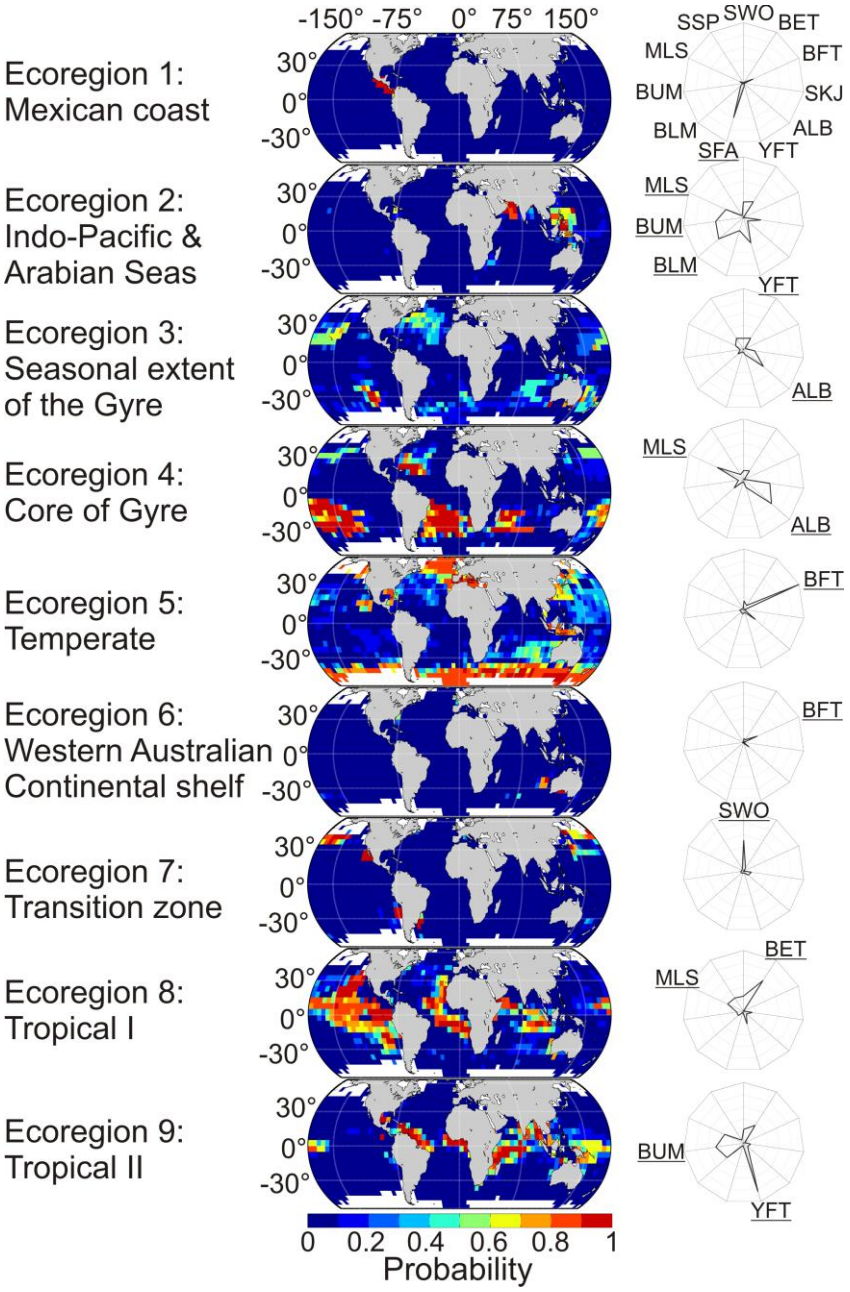

Figure 3 Maps of the probabilities that each group of geographical cells identified for cut-off level VI is in a given ecoregion. The corresponding radar plot shows the indicator species of tuna and billfish for the ecoregion. All radar plots are constricted between 0 and 50% with lines at 10% intervals. Each species code corresponds to a species name that can be retrieved from Table 1. The projection used for each map is Eckert IV.

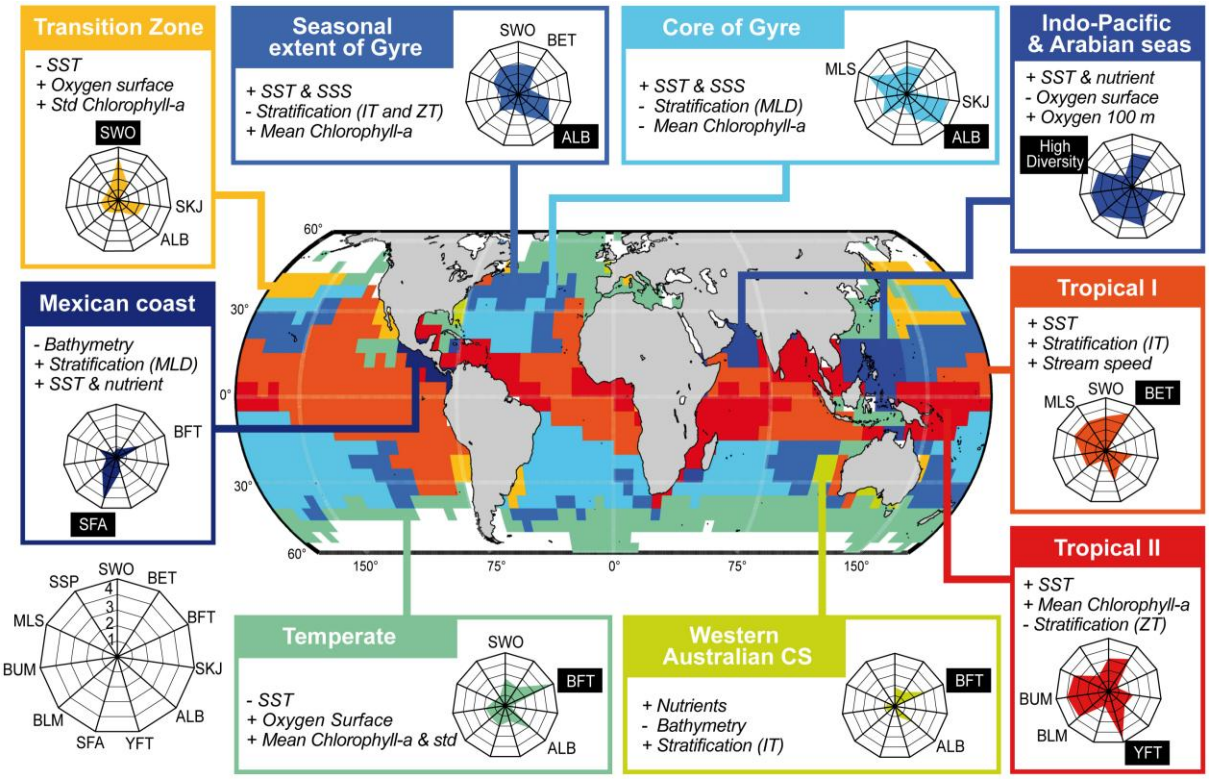

Figure 4 Proposed ecological partitioning of the global ocean based on the distribution of tuna and billfish species (with an Eckert IV projection). Each community is associated with a specific colour on the map and a descriptive box. For each ecoregion, the name, species association, and main environmental driver factors are provided in the corresponding box. The annotation '+' denotes a high value and '-' a low value of the environmental parameter. The radar plots are on a $\log_{10}$ scale. The dominant species and secondary species of each community detected are annotated on the plot. Each species code corresponds to a species name in Table 1. The environmental factors are annotated on the figure using acronyms (see Appendix S1): SST= sea surface temperature; SSS = sea surface salinity; MLD= mixed layer depth; IT= intensity of the thermocline; ZT= depth of the thermocline; std= standard deviation; Oxygen 100m = Oxygen at 100m.

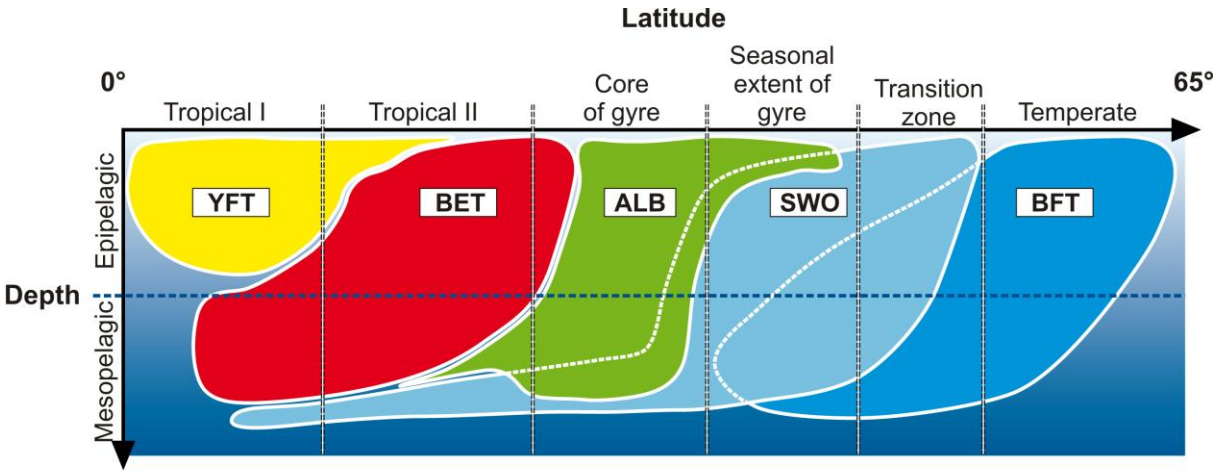

Figure 5 Conceptual scheme of the spatial distribution of each ecoregion as a function of latitude and depth. The figure was produced using the ecoregion that predominates at each latitude (Fig. 4) and refined (vertically) using established knowledge on the spatial distribution of the species studied (see e.g. Block & Stevens, 2001). Oceanic ecoregions are represented by dashed black lines. The spatial distributions of the five dominant species are each represented by a specific colour (YFT, yellowfin tuna; BET, bigeye tuna; ALB, albacore tuna; SWO, swordfish; BFT, all bluefin tuna species). Depth is divided into epipelagic (water mass from the surface to the mixed layer depth) and mesopelagic (from the mixed layer depth to 1000 m) zones.

Table 1 The common name, Latin name, spatial distribution, and code name for each species of tuna and billfish in the study.

Name	Latin name	Taxonomic authority	Ocean	Code
Northern bluefin	*Thunnus thynnus*	Linnaeus, 1758	Atlantic	BFT
Southern bluefin	*Thunnus maccoyii*	Castelnau, 1872	Atlantic, Pacific, Indian	
Pacific bluefin	*Thunnus orientalis*	Temminck & Schlegel, 1844	Pacific	
Bigeye tuna	*Thunnus obesus*	Lowe, 1839	Atlantic, Pacific, Indian	BET
Yellowfin tuna	*Thunnus albacares*	Bonnaterre, 1788	Atlantic, Pacific, Indian	YFT
Albacore tuna	*Thunnus alalunga*	Bonnaterre, 1788	Atlantic, Pacific, Indian	ALB
Skipjack tuna	*Katsuwonus pelamis*	Linnaeus, 1758	Atlantic, Pacific, Indian	SKJ
Swordfish	*Xiphias gladius*	Linnaeus, 1758	Atlantic, Pacific, Indian	SWO
Indo-Pacific blue marlin	*Makaira mazara*	Lacepedein 1802	Pacific, Indian	BUM
Atlantic blue marlin	*Makaira nigricans*	Lacepède, 1801	Atlantic	
Black marlin	*Makaira indica*	Cuvier, 1832	Atlantic, Pacific, Indian	BLM
Striped marlin	*Tetrapturus audax*	Philippi, 1887	Atlantic, Pacific, Indian	MLS
Atlantic sailfish	*Istiophorus albicans*	Latreille, 1804	Atlantic	SFA
Pacific sailfish	*Istiophorus platypterus*	Shaw, 1792	Pacific, Indian	
Short-billed spearfish	*Tetrapturus angustirostris*	Tanaka, 1914	Pacific, Indian	SSP

Table 2 Homogeneity index between each biogeochemical province (BGCP; Longhurst, 1998) and the nine ecoregions identified using the CPUEs of 13 dominant tuna and billfish species. The percentage of geographical cells represented by each ecoregion was calculated. The total percentage homogeneity was computed from the sum of the geographical cells of the dominant ecoregion in each province divided by the total geographical cells considered in the study.

Province name	Code	Biome	Ocean	Number of cells	Ecoregion (%)								
					1	2	3	4	5	6	7	8	9
Brazilian current coast	BRAZ	Coastal	Atlantic	8	0.0	0.0	0.0	50.0	0.0	0.0	50.0	0.0	0.0
Benguela current coast	BENG	Coastal	Atlantic	5	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0
Guinea current coast	GUIN	Coastal	Atlantic	9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.1	88.9
Canary current coast	CNRY	Coastal	Atlantic	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0
Guianas coast	GUIA	Coastal	Atlantic	9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100
Northeast Atlantic shelves	NECS	Coastal	Atlantic	2	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
Northwest Atlantic shelves	NWCS	Coastal	Atlantic	8	0.0	0.0	12.5	12.5	25.0	37.5	0.0	0.0	12.5
Eastern India coast	EAFR	Coastal	Indian	15	0.0	13.3	0.0	26.7	0.0	0.0	0.0	0.0	60.0
Western Australian and Indonesian coast	AUSW	Coastal	Indian	25	0.0	13.3	0.0	0.0	6.7	26.7	0.0	46.7	6.7
Eastern India coast	INDE	Coastal	Indian	8	0.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0	87.5
Western India coast	INDW	Coastal	Indian	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100
Humboldt current coast	HUMB	Coastal	Pacific	8	0.0	0.0	0.0	0.0	25.0	0.0	25.0	50.0	0.0
East Australian coast	AUSE	Coastal	Pacific	6	0.0	0.0	66.7	0.0	33.3	0.0	33.3	0.0	0.0
Sunda-Arafura shelves	SUND	Coastal	Pacific	17	0.0	5.9	0.0	0.0	41.2	0.0	0.0	17.6	35.3
China Sea	CHIN	Coastal	Pacific	4	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Central American coast	CAMR	Coastal	Pacific	9	77.8	0.0	0.0	0.0	0.0	0.0	22.2	0.0	0.0
Alaska Coastal downwelling	ALSK	Coastal	Pacific	1	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
New Zealand coast	NEWZ	Coastal	Pacific	7	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Coastal Californian current	CCAL	Coastal	Pacific	14	0.0	0.0	0.0	0.0	27.3	0.0	27.3	45.5	0.0
Antarctic	ANTA	Polar	Antarctic	16	0.0	0.0	10.0	0.0	90.0	0.0	0.0	0.0	0.0
Atlantic Arctic	ARCT	Polar	Atlantic	8	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Atlantic subArctic	SARC	Polar	Atlantic	9	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
South Atlantic gyral	SATL	Trade wind	Atlantic	68	0.0	0.0	10.3	76.5	1.5	0.0	0.0	11.8	0.0
Eastern tropical Atlantic	ETRA	Trade wind	Atlantic	18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	69.2	30.8
Western tropical Atlantic	WTRA	Trade wind	Atlantic	16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	62.5	37.5
Caribbean	CARB	Trade wind	Atlantic	16	0.0	0.0	6.3	0.0	0.0	37.5	0.0	0.0	56.3
North Atlantic tropical gyral	NATR	Trade wind	Atlantic	27	0.0	0.0	18.5	40.7	0.0	0.0	0.0	29.6	11.1
Indian south subtropical gyre	ISSG	Trade wind	Indian	66	0.0	0.0	30.3	51.5	12.1	6.1	0.0	0.0	0.0
Indian monsoon gyre	MONS	Trade wind	Indian	46	0.0	0.0	0.0	0.0	0.0	0.0	0.0	47.8	52.2
Archipelagic deep basins	ARCH	Trade wind	Pacific	31	0.0	32.3	32.3	16.1	0.0	0.0	0.0	0.0	19.4
Pacific equatorial divergence	PEQD	Trade wind	Pacific	68	0.0	0.0	0.0	0.0	0.0	0.0	0.0	91.4	8.6
North Pacific equatorial countercurrent	PNEC	Trade wind	Pacific	25	0.0	0.0	0.0	0.0	12.0	0.0	0.0	84.0	4.0
North Pacific Tropical gyre	NPTG	Trade wind	Pacific	135	0.0	4.3	39.1	11.3	6.1	0.0	9.6	29.6	0.0
South Pacific gyre	SPSG	Trade wind	Pacific	151	0.0	0.0	13.2	80.2	0.0	0.0	1.7	18.2	0.0
Western Pacific warm pool	WARM	Trade wind	Pacific	42	0.0	4.8	9.5	16.7	0.0	0.0	0.0	16.7	52.4
South subtropical convergence	SSTC	Westerly wind	Antarctic	65	0.0	0.0	15.4	4.6	80.0	0.0	0.0	0.0	0.0
SubAntarctic water ring	SANT	Westerly wind	Antarctic	72	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Northeast Atlantic subtropical gyral	NAST E	Westerly wind	Atlantic	21	0.0	0.0	38.1	61.9	0.0	0.0	0.0	0.0	0.0
Mediterranean Sea	MEDI	Westerly wind	Atlantic	13	0.0	0.0	0.0	0.0	92.3	0.0	7.7	0.0	0.0
Northwest Atlantic subtropical gyral	NAST W	Westerly wind	Atlantic	19	0.0	0.0	47.4	0.0	26.3	0.0	0.0	26.3	0.0
Gulf Stream	GFST	Westerly wind	Atlantic	8	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0

North Atlantic Drift	NADR	Westerly wind	Atlantic	14	0.0	0.0	7.1	0.0	92.9	0.0	0.0	0.0	0.0
Northwest Arabian Sea upwelling	ARAB	Westerly wind	Indian	12	0.0	0.0	0.0	0.0	0.0	0.0	58.3	25.0	16.7
Tasman Sea	TASM	Westerly wind	Pacific	6	0.0	0.0	83.3	0.0	16.7	0.0	0.0	0.0	0.0
Kuroshio Current	KURO	Westerly wind	Pacific	18	0.0	0.0	0.0	14.3	64.3	0.0	21.4	0.0	0.0
Eastern Pacific subArctic gyres	PSAE	Westerly wind	Pacific	2	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Western Pacific subArctic gyres	PSAW	Westerly wind	Pacific	9	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
North Pacific polar front	NPPF	Westerly wind	Pacific	26	0.0	0.0	26.9	0.0	19.2	0.0	53.8	0.0	0.0
Total				**1188**	**68.8**								

1