

## ARTICLE

<https://doi.org/10.1038/s41467-022-28488-1>

OPEN

# Global patterns in functional rarity of marine fish

Isaac Trindade-Santos <sup>1</sup>✉, Faye Moyes <sup>1</sup> & Anne E. Magurran <sup>1</sup>

Rare species, which represent a large fraction of the taxa in ecological assemblages, account for much of the biological diversity on Earth. These species make substantial contributions to ecosystem functioning, and are targets of conservation policy. Here we adopt an integrated approach, combining information on the rarity of species trait combinations, and their spatial restrictedness, to quantify the biogeography of rare fish (a taxon with almost 13,000 species) in the world's oceans. We find concentrations of rarity, in excess of what is predicted by a null expectation, near the coasts and at higher latitudes. We also observe mismatches between these rarity hotspots and marine protected areas. This pattern is repeated for both major groupings of fish, the Actinopterygii (bony fish) and Elasmobranchii (sharks, skates and rays). These results uncover global patterns of rarity that were not apparent from earlier work, and highlight the importance of using metrics that incorporate information on functional traits in the conservation and management of global marine fishes.

<sup>1</sup>Centre for Biological Diversity, School of Biology, University of St Andrews, St Andrews, Scotland, UK. ✉email: [its5@st-andrews.ac.uk](mailto:its5@st-andrews.ac.uk)

Pioneering ecologists, including Charles Darwin<sup>1</sup> and Henry Walter Bates<sup>2</sup>, realised the importance of documenting geographic patterns in numbers of rare species when explaining variation in biological diversity over space and time. Rabinowitz's influential paper<sup>3</sup> on seven forms of rarity showed that it is possible to quantify these rare taxa objectively based on information about their geographic extent, local population size, and habitat specificity. Rare taxa play key roles in underpinning ecosystem function<sup>4–6</sup>, and have priority in conservation planning<sup>7,8</sup>. Progress towards the goal of identifying which parts of the Earth support concentrations of rare species is greatest in terrestrial systems<sup>9,10</sup>, with the geography of rarity in the oceans less well charted (but see Connolly, et al.<sup>11</sup> and Grenié, et al.<sup>7</sup> (for the tropics only)). To date, most evaluations of rarity have, *sensu* Rabinowitz, focussed on taxonomic diversity. Yet, as is now evident, biodiversity is a multi-faceted concept, and it is vital to incorporate information on ecological function in biodiversity assessments<sup>12,13</sup> and in conservation<sup>14,15</sup>. New rarity metrics<sup>16,17</sup> make possible an integrated assessment of rarity that takes account of the extent to which species exhibit rare combinations of traits, as well as being geographically restricted while a null model approach can be used to gauge whether such estimates of rarity levels are in excess of an expectation based on species richness. Here we identify concentrations, towards higher latitudes and in coastal systems, of bony and cartilaginous fish species that are rare in both taxonomic and functional dimensions, and document some mismatches between these rarity hotspots and the Marine Protected Areas of the world's oceans.

Our approach to measuring rarity draws on elements of taxonomic and functional biodiversity. First, we use the distribution data of all marine fish species (11,961 Actinopterygii and 866 Elasmobranchii) in AquaMaps<sup>18</sup> and allocate these to 2° grid cells in order to calculate a key facet of rarity, namely species geographic restrictedness<sup>3</sup> (see Supplementary Fig. 1 for workflow and Supplementary Fig. 2 for geographic restrictedness calculation). We then examine eleven ecological and biological relevant traits (7 continuous and 4 categorical traits, drawn from FishBase<sup>19</sup>, representing key environmental, morphological, life history and reproductive variables) for each species, and use these data to quantify how functionally distinctive and unique each species is<sup>16</sup> (see Supplementary Fig. 2 for functional distinctiveness and uniqueness calculations). We define rare taxa as those that occur in the top (i.e. most rare) quartiles<sup>20</sup> of both restrictedness and distinctiveness distributions. Our measure of rarity thus reflects both taxonomic rarity (geographic restrictedness) and functional rarity (global functional distinctiveness and functional uniqueness), i.e. facets of rarity conceptualized respectively by Rabinowitz<sup>3</sup> and Grenié, et al.<sup>16</sup>. Based on previous findings that the tropics are the hotspots for marine species richness (taxonomic diversity) but have high levels of functional redundancy, e.g. species playing similar ecological roles<sup>21,22</sup> we expect to find low functional rarity in the tropics. Our analyses were undertaken independently for bony fish (Actinopterygii) and cartilaginous fish (Elasmobranchii). Moreover, our investigation was conducted separately for each Coastal System and High Seas in seven Oceanic Regions (14 different systems in total) (see Supplementary Fig. 3). That is, for each system, we assessed species rarity based on the distribution and functional traits of species present in that system. Finally, we repeated all the analyses using rarity defined by functional uniqueness<sup>9</sup> (instead of functional distinctiveness) to assess the generality of our findings.

## Results and discussion

As Fig. 1 shows, we detect substantial heterogeneity in the distribution of rarity in the world's Coastal Systems. Rarity is most

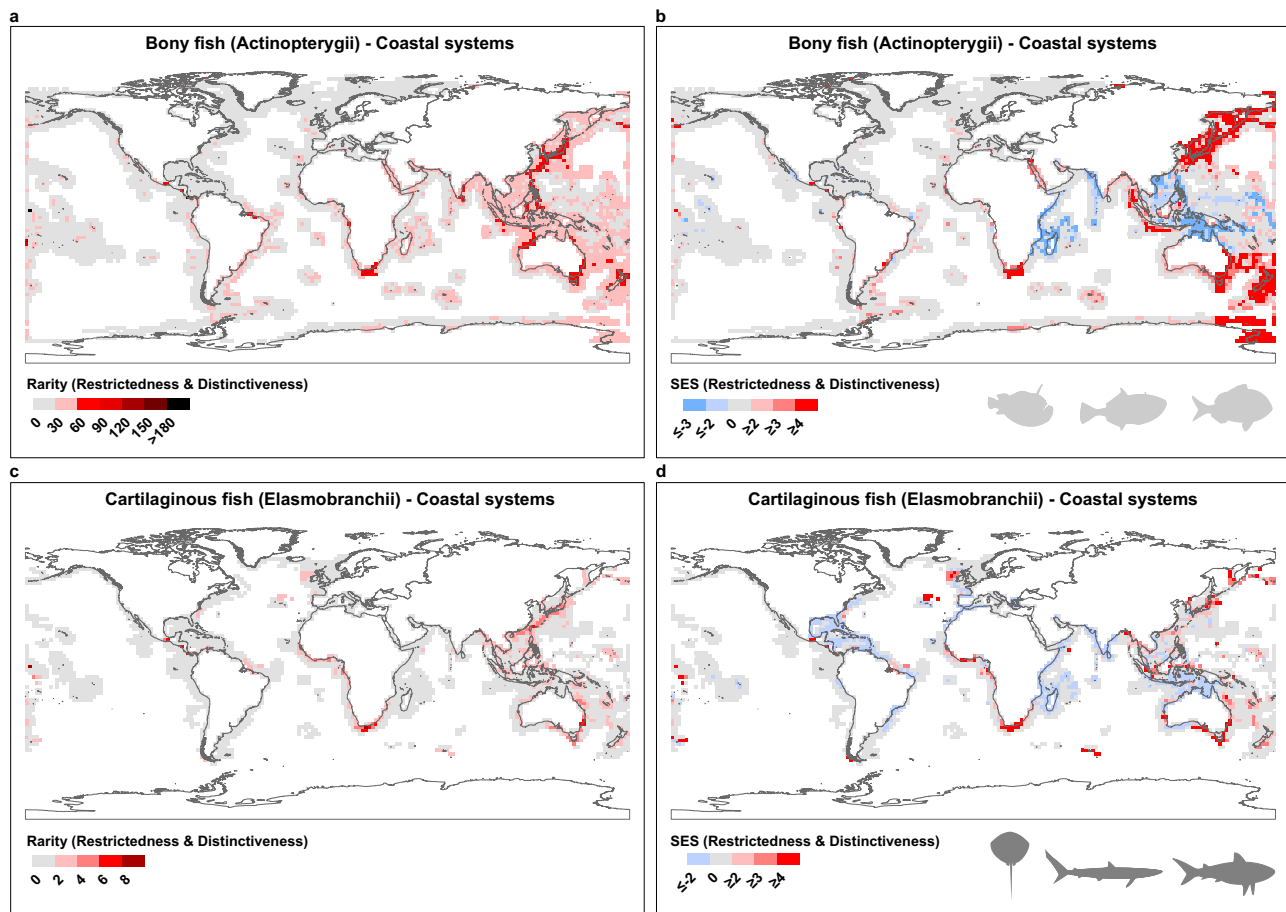
striking in higher latitudes and coastal ecosystems. Hotspots of rarity occur on the East coast of Asia, Northeast and Southwest Australia and Southern Africa, with some concentration at the estuary of the Amazon River in Brazil and Central America for the bony fishes (Actinopterygii), (Fig. 1a), and on the East coast of Asia, the Southwest coast of Australia, South and West coasts of Africa, Central America and the UK for the cartilaginous fishes (sharks, skates and rays—Elasmobranchii) (Fig. 1c). When analysing the results for the High Seas we found more concentrations of rare species for both groups of species in the Atlantic and Eastern Pacific Oceans (Fig. 2a for bony fish and c for cartilaginous). These patterns are likely, to some extent, to reflect variation in species richness, which in line with other taxa<sup>20,23</sup> is highest in the lower latitudes (see Fig. 3a, b which depict species richness variation across latitude for bony and cartilaginous species respectively). However, when we control for species richness by comparing the observed rare species with the null expectation (see *Null Model* section at the Methods), our finding that high latitudes and coastal regions are associated with levels of rarity higher than expected by chance are reinforced (see Figs. 1 and 2b, d excesses (red) and deficits (blue) in rarity, relative to the null, and Fig. 3c–f for latitudinal variation in rarity, again as evaluated against the null). The hotspots identified in Figs. 1 and 2a, c remain, and are supplemented by a much larger sweep of coastline (Fig. 1), particularly in the Asia and Australasia, Africa, Red Sea (mainly bony fish), Amazon Estuary, Central America and UK (mainly cartilaginous fish) (Fig. 1b, d). Moreover, there is a marked concentration of rarity, in excess of expectations, at higher latitudes (Fig. 3, from c to f).

When we compared these results (Figs. 1 and 2, which employ the functional distinctiveness metric) with an alternative method of calculating functional rarity, namely functional uniqueness (Supplementary Fig. 4 (Coastal Systems) and Supplementary Fig. 5 (High Seas), see Supplementary Fig. 2 for indices differences) the conclusions were consistent. Interestingly, as the Supplementary Fig. 4 (Coastal Systems) and Supplementary Fig. 5 (High Seas) show, many of the same hotspots of rarity were identified, albeit with some differences along the South-western African Coast, Southern Indian Ocean, Red Sea (for bony fishes (Supplementary Fig. 4b)), and in Central America and South-eastern North America (Supplementary Fig. 4d) for cartilaginous fish.

All analyses presented here used the probability of occurrences provided by the AquaMaps database >0.9 (see *Occupancy Data (Assemblage Matrix)* section in Methods for more details). As a further check, we repeated the investigations using probabilities >0.7 and >0.5. In all scenarios, our findings were consistent (see Supplementary Fig. 6a to d for probability >0.7 and Supplementary Fig. 6e to h for probability >0.5).

Distance decay plots, using the Northwest Pacific Ocean as an example (Supplementary Fig. 7) provide reassurance that spatial autocorrelation is not obscuring the results we report. In addition, we performed a sensitivity analysis to test if the environmental traits “mean temperature preference” and “maximum depth”, along with the categorical traits (reproductive guild, body shape, swimming mode and position in water column) influenced our results. We found that even when we remove these traits from the analysis our results are consistent (Supplementary Fig. 8a to d). This supports our argument that the global patterns of rarity we uncover are not simply a response to traits influenced by local features. In conjunction with a trait correlation analysis (Supplementary Fig. 9a and b), this shows that no single trait was unduly influential on the patterns we found.

Do these patterns of rarity reflect heterogeneity in habitat use by the rare taxa? To explore this we compared the frequencies of “habitat specialization” categories (i.e. bathydemersal, bathypelagic, benthopelagic, demersal, pelagic neritic, pelagic oceanic and reef



**Fig. 1 Global biogeography of rarity for bony fishes (a-Actinopterygii) and cartilaginous fishes (c-Elasmobranchii) across Coastal Systems.** The functional index used here was distinctiveness (see Supplementary Fig. 1 for workflow and Supplementary Fig. 2 for index details). Plots **a** and **c** illustrate the numbers of rare species found in each 2° grid cell (species that are both rare **taxonomically** and **functionally (distinct)**). Plots **b**, **d** shows the Standardized Effect Sizes (SES), where red shaded cells represent an excess of rare species higher than expected by chance and blue cells represent fewer rare species than expected. For results using functional uniqueness, see Supplementary Fig. 4.

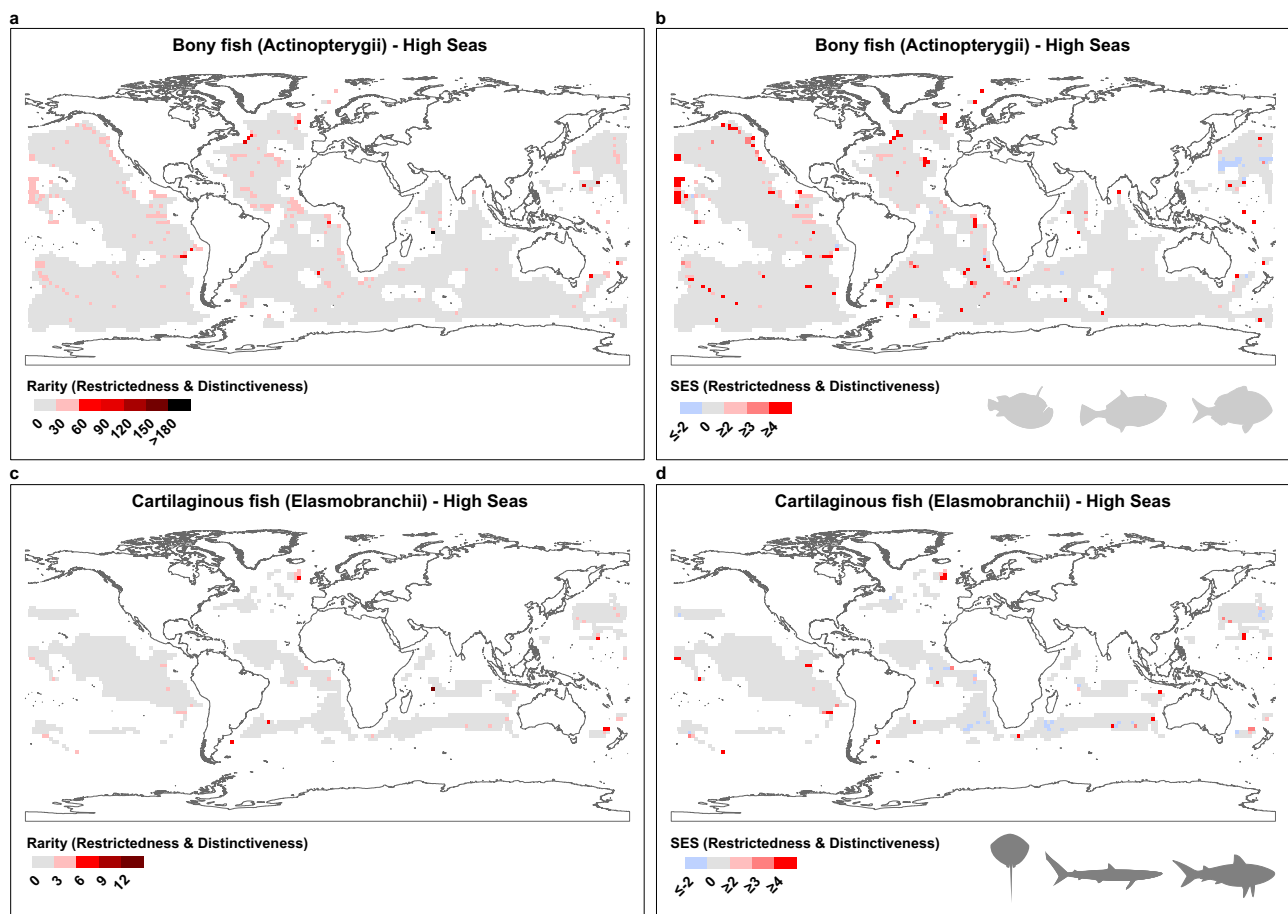
associated) in the rare and non-rare species and found evidence of differences in their association with different habitat types (Supplementary Fig. 10). For example, in the Northwest Pacific Coast most non-rare species are classified as reef associated (Supplementary Fig. 10a (pale green)), while rare species are mostly classified as demersal (Supplementary Fig. 10a (dark green)).

It is possible that a sampling bias towards records in coastal waters could partly explain our findings. However, we note that regions of the earth that are typically under-recorded in biodiversity surveys<sup>24</sup>, such as South America and Africa (but see Floeter et al.<sup>25,26</sup>, and Pinheiro et al.<sup>27,28</sup>), are identified in our analysis as important repositories of rarity, while marine areas that have been intensively monitored, such as the coasts of NW Europe<sup>29–31</sup> are not. Additional analyses showed that sampling effect has no influence on the identification of rarity hotspots (see Supplementary Analysis and Supplementary Fig. 11).

The striking concentrations of rarity uncovered by our analysis are incompletely represented by existing Marine Protected Areas (MPAs). Here we defined cells exhibiting rarity as those where  $SES > 2$ . We found mismatches between the location of 2° grid cells with these high levels of rarity and MPAs (see Fig. 4a and b, bony and cartilaginous fish respectively). This was true for both bony fish (matches in 47% of MPAs (Fig. 4a)) and cartilaginous fish (matches in 63% (Fig. 4b)). Thus, on average, only around half of the localities in the oceans with notable levels of rare fish are protected by MPAs.

AquaMaps<sup>18,32</sup> provided us with a unique opportunity to use occurrence data for measuring rarity based on occupancy data. We treated species occurrences in each of the two-degree grid cells as a proxy for rarity and used this information to distinguish between rare species (highly restricted species) and common species (widely occurring species). The global coverage of AquaMaps is an advantage here. There are, however, some caveats associated with our approach. For example, the data provided by AquaMaps include not only actual observations but also probabilities of occupancy based on combinations of sightings and internal algorithms. Moreover, AquaMaps is based on summed information rather than being a representation of the species present at a given point in time. Additionally, the nature of our data means we cannot calculate shifts in rarity on a temporal scale or with any weighting by abundances. We also appreciate that when abundance and/or biomass data are available, another taxonomic rarity metric, namely Taxonomic Scarcity<sup>16</sup>, could be used. Our decision to focus on restrictedness was shaped by the fact that we are using geographical information from all described fish available in AquaMaps, and that this database provides occupancy patterns based on different probabilities of occurrences.

Our analysis highlights the new insights gained from an integrative approach to quantifying rarity. It shows that using other proxies of biodiversity importance, such as hotspots of richness, and MPAs do not effectively capture the regions of the ocean that



**Fig. 2** Global biogeography of rarity for bony fishes (a-Actinopterygii) and cartilaginous fishes (c-Elasmobranchii) across High Seas. The functional index used here was distinctiveness. Plots **a** and **c** illustrate the numbers of rare species found in each 2° grid cell (species that are rare both **taxonomically** and **functionally** (distinct)). Plots **b** and **d** show the Standardized Effect Sizes (SES). For results using functional uniqueness, see Supplementary Fig. 5.

support concentrations of rare taxa. Indeed, it is higher latitudes, and in particular coastal areas, where rare taxa are most evident. May<sup>33</sup> observed that the oceans are richer in phyla, but poorer in species, than terrestrial ecosystems. He attributed this to both the long evolutionary history of marine life, and the relative lack of spatial heterogeneity in the oceans. Both factors are likely to contribute to the patterns we uncovered. On the one hand the ancient phylogenetic roots of marine fish provide opportunities for trait innovations to arise<sup>34</sup>. On the other, the spatial structure of coastlines and inshore waters is likely to foster local adaptation<sup>35,36</sup>. Similarly, Rabosky<sup>34</sup> found faster speciation rates outside the tropics, particularly at higher latitudes; it is possible that those regions acts as centres for evolutionary and functional novelty (as shown in our study).

The results we report demonstrate the importance of an integrative approach in conservation to protect rare species that exhibit distinct functional traits<sup>37</sup> and make important contributions to ecosystem functioning<sup>4,10,38</sup>. The advantages of such an approach are already clear. For example, Hidasi-Neto et al.<sup>14</sup> used the combination of different facets of biodiversity (species distribution, functional and phylogenetic distinctiveness) and the IUCN species threat status of terrestrial mammals to highlight the importance of the effectiveness of conservation efforts by focusing on regions with high ecologically and evolutionarily distinct mammal species. However, rare taxa are not necessarily conservation priorities at present. For instance, rare terrestrial mammals and bird species are not covered by protected areas at a global scale<sup>10</sup> even though elevated extinction risk is linked to high ecological distinctiveness<sup>39</sup>.

Concentrations of rare coral reef species also have mismatches with MPAs<sup>7</sup>.

Human exploitation, combined with environmental change, is threatening rare taxa with important functional roles, including marine megafauna<sup>8</sup>. In particular, we note that the ongoing global decline of sharks, which display the highest functional distinctiveness amongst the fishes<sup>8</sup>, will contribute to a substantial loss of functional diversity in marine megafauna<sup>8,40,41</sup>. Fishing pressure and climate change are known threats to fish functional diversity<sup>42</sup> and rarity, as a recent study of the North Sea shows<sup>43</sup>.

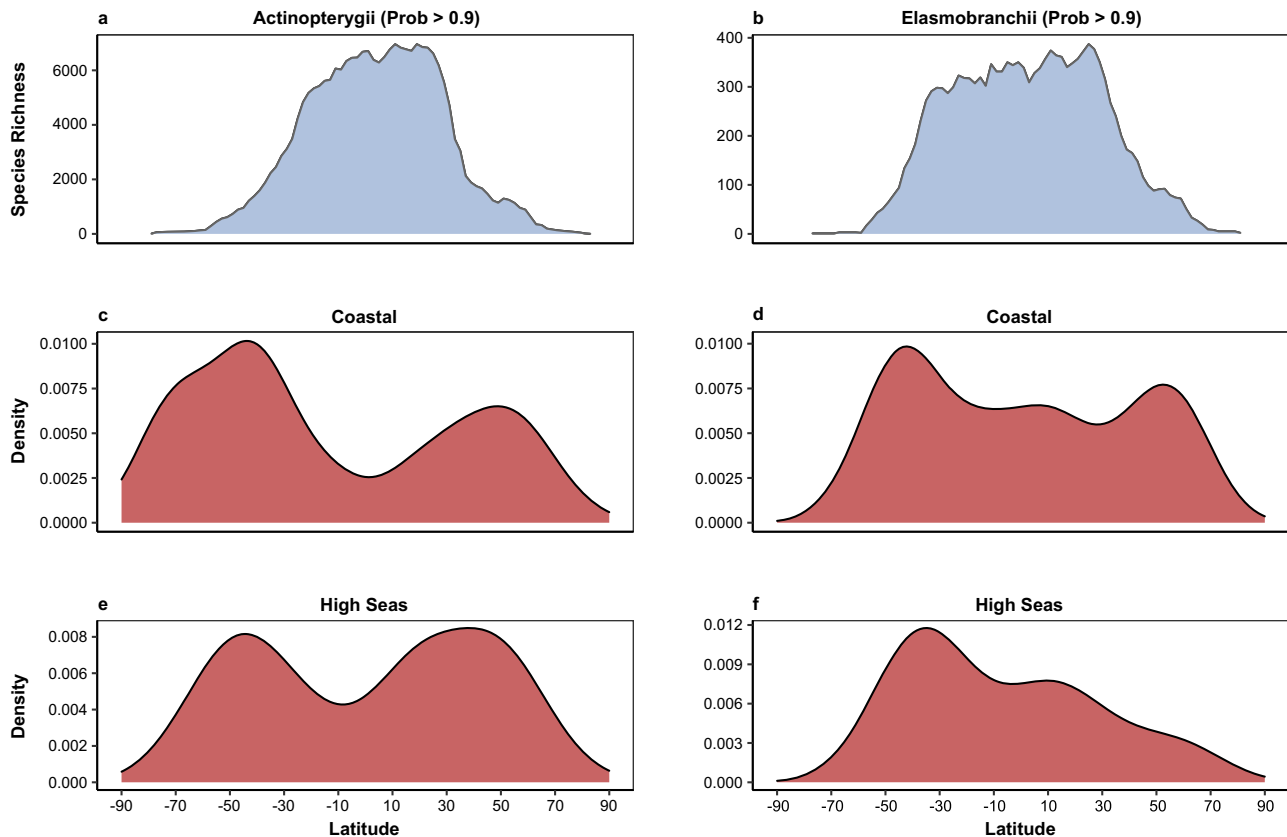
As the vision for the United Nations Decade of Ocean Science for Sustainable Development makes clear, there are still substantial knowledge gaps concerning the status and trends of marine biodiversity. Our paper helps fill one of these knowledge gaps and emphasises the importance of adopting an integrated approach to evaluating, and protecting, the biodiversity on which we depend.

## Methods

We used AquaMaps<sup>18,32</sup> to obtain species occurrence data, and extracted eleven ecologically and biologically relevant traits (seven continuous and four categorical) from FishBase<sup>19</sup>. All calculations were done separately for the two group of fishes (bony fishes and cartilaginous fishes). Our workflow is in Supplementary Fig. 1.

### Step 1- Input

**Occupancy Data (assemblage matrix).** We extracted occupancy data from AquaMaps<sup>18,32</sup>. This online database provides half-degree grid cell species occurrences based on data from GBIF<sup>44</sup> and OBIS<sup>45</sup> complemented with information from FishBase<sup>19</sup> and SeaLifeBase<sup>46</sup>. AquaMaps gives probabilities of the occurrence



**Fig. 3** Illustration of the latitudinal biogeographic distribution of different metrics: species richness (a and b), positive SES ( $\geq 2$ ) based on distinctiveness (from c to f). a, b plot the distribution of species richness in relation to latitude for Actinopterygii (left plots) and Elasmobranchii (right plots) respectively. Plots from c and d represent **Coastal Systems**, while e and f show the SES distribution for the **High Seas**. Here we use the **probability of species occurrences higher than 0.9** (see Supplementary Fig. 6 for the latitudinal results found using probability  $> 0.7$  and  $> 0.5$ ).

of a given species between 0 (based on the environmental envelope indicating there is no chance of finding the species in that grid cell) to 1 (highest probability to find the species in that grid cell)<sup>32</sup>.

These occurrence data are then combined with an algorithm implemented by AquaMaps using “estimates of environmental preferences with respect to depth, water temperature, salinity, primary productivity, and association with sea ice or coastal areas”. For more information see Kesner-Reyes, et al.<sup>32</sup>.

In our analysis, we selected occurrence data (the presence of a given species in a certain half degree grid cell) with a probability  $> 0.9$ . To ensure our results were not simply an artefact of using a high probability of occurrence we also examined probabilities higher than 0.7 and 0.5. Further analyses were repeated independently for each of these probabilities. After this initial step, we allocated each half-degree grid cell to a  $2^\circ$  grid cell. We created the  $2^\circ$  grid cells using features available in ArcGIS<sup>47</sup>. These occupancy data also allowed us to compute the species richness of each  $2^\circ$  grid cell.

We applied all analyses described here (see Supplementary Fig. 1) separately for the Coastal Systems and High Seas of each of the Seven Oceanic Regions (see Supplementary Fig. 3). This gives us seven Coastal Systems and seven High Seas, a total of 14 assemblage matrices for each of the probabilities (probability  $> 0.9$ , 0.7 and 0.5).

**Trait Data Compilation (trait matrix).** Our final occurrence data (each of the 14 systems) were split between marine bony fishes (11,961 Actinopterygii species) and cartilaginous fishes (866 Elasmobranchii species). For each species in both groups of fish we assembled biologically and ecologically relevant traits from the most recent version of the FishBase database<sup>19</sup>. This gave us seven continuous and four categorical traits that are largely uncorrelated with one another, (see Supplementary Fig. 9a and b):

**Environmental Traits:** (1) *Position in Water Column*—The vertical position of a species indicates its feeding habitat<sup>19</sup> and its influence on the process of transferring nutrients through the water column<sup>48,49</sup>. This trait has 8 categories. (2) *Maximum Depth (m)*—This trait reflects the environmental conditions that each species occur<sup>27</sup>. (3) *Mean Temperature Preference (°C)*—Thermal variations in preferences indicates the species tolerances to changes in temperature<sup>50–52</sup>.

**Life History Traits:** (4) *Growth (k)*—This coefficient parameter ( $k = 1 \text{ year}^{-1}$ ) is derived from the von Bertalanffy growth function ( $L_t = L_\infty (1 - \exp(-K(t-t_0)))$ ).

Faster growth rates are associated with higher  $k$  values<sup>19</sup>. (5) *(Q/B)*—this trait represents the proportional ratio between food consumption ( $Q$ ) and biomass ( $B$ ) and can be used as a proxy for trophic interactions, evidencing the flow of energy in the system<sup>53,54</sup>. (6) *Trophic Level*—the position of a given species in the food chain is expressed as its trophic level and as discussed by Froese and Pauly<sup>19</sup> can be assessed as the amount of “energy-transfer steps to that level”. The trophic level also gives information on interactions between species, for example predator-prey and trophic cascades<sup>55</sup>.

**Morphological Traits:** (7) *Body Shape*—The body shape of a species relates to ecological behaviours, such as migration patterns<sup>56</sup>. We divided this trait into 38 categories (see Supplementary Table 1). (8) *Swimming Mode*—the mobile strategy adopted by fish species has a direct relationship with ecology and behaviour<sup>57</sup>. Following FishBase we used 12 different swimming modes (see Supplementary Table 1) based on anatomical and morphological features; these traits provide information on the functional role that each species plays.

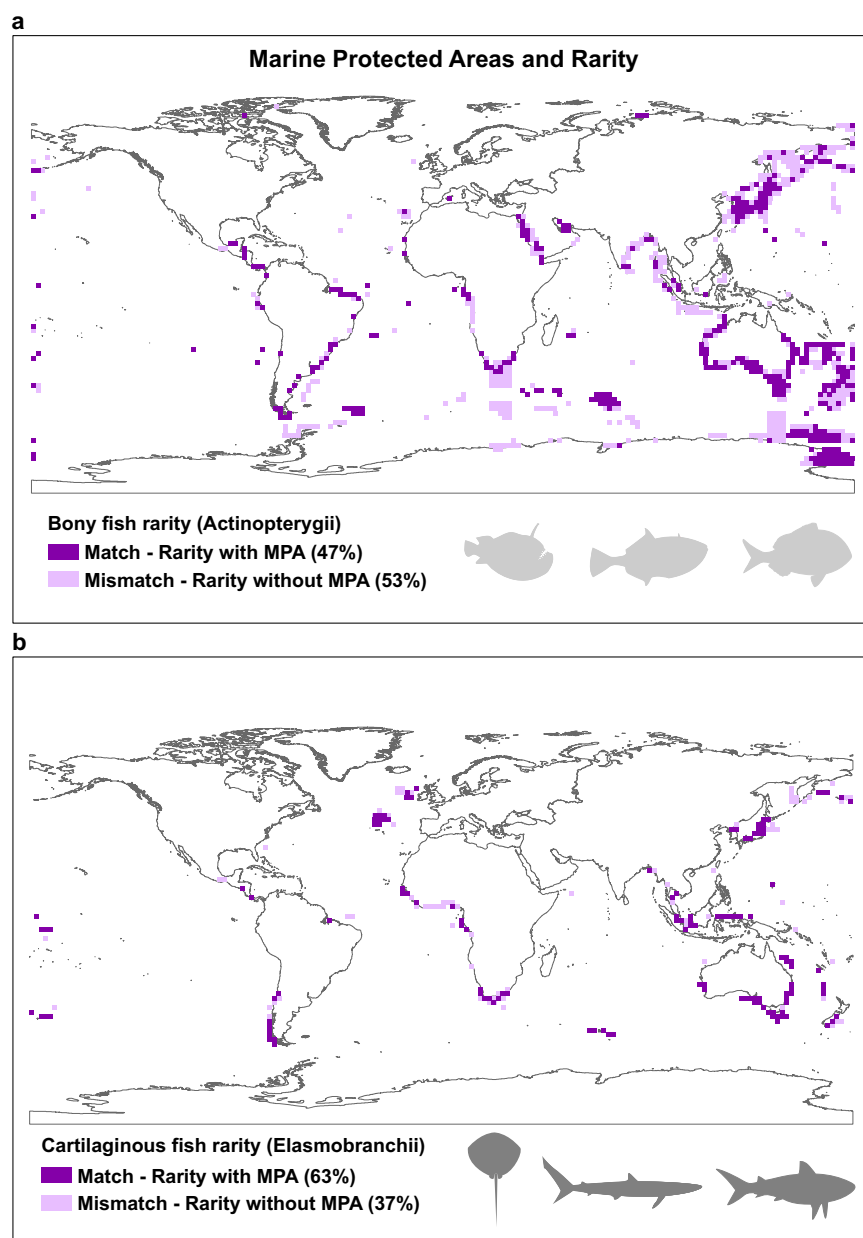
**Reproductive Traits:** (9) *Generation Time*—defined by FishBase as “the time period from birth to average age of reproduction”. (10) *Length of First Maturity (mm)*—body length when around 50% of a given species becomes mature<sup>58,59</sup>. (11) *Reproductive Guild*—15 categories of reproductive guild as defined by FishBase (see Supplementary Table 1). Reproductive traits have a direct influence on population dynamics and species resilience<sup>60,61</sup>, and are therefore commonly used in fisheries management<sup>62</sup>.

These traits were selected for their ecological and biological relevance as described above. We tested the correlations of the traits to ensure complementarity, and as shown in the (Supplementary Fig. 9a and b), these traits are largely uncorrelated. We also took into consideration the data gaps inevitable in a large data set such as FishBase (traits were selected if a maximum of 30% of data were missing). To overcome this limitation we applied random forest algorithms to fill the missing traits<sup>63</sup>, by using the package “missForest”<sup>64</sup>.

## Step 2—Rarity indices

**Species restrictedness.** We calculated species restrictedness ( $Res_i$ ) by dividing the species geographical extent ( $GE = \text{number of } 2^\circ \text{ grid cells that a species occurs in based on the assemblage matrix compiled from AquaMaps}$ ) by the total number of grid cells (TOT), minus one (see Supplementary Fig. 2). We scaled the values





**Fig. 4 Congruence and mismatches between MPAs and rarity hotspots.** We illustrate the results for the Actinopterygii **a** and Elasmobranchii **b** for the Seven Oceanic Regions. In this example, the grid cells presenting Standardized Effect Sizes higher than 2 with the presence of a Marine Protected Area are represented as dark purple, meanwhile the soft purple are the grid cells with SES > 2 and no MPA. The figure shows that 47% of grid cells identified with high concentrations of rarity coincide with MPAs for bony fish **a** and 63% for cartilaginous fish **b**.

between 0 (species occurring in all 2° grid cells) and 1 (the most restricted species). We used the function “restrictedness” in the “funrar” package to do this calculation<sup>16</sup>.

**Functional distinctiveness.** Functional distinctiveness ( $Dis_i$ ) quantifies the level of dissimilarity in trait combination between species<sup>16,22</sup> (see Supplementary Fig. 2). This index is the average of functional distance of a given species compared with all other species in the assemblage<sup>16</sup>.

We calculated how distinct or common each species is by using the function “distinctiveness\_global” available in the “funrar” package<sup>16</sup>. We then scaled the values found between 0 (species with common combinations of traits) and 1 (species with the most dissimilar combination of traits). This analysis was conducted using presence/absence data.

**Functional uniqueness.** Functional uniqueness quantifies the level of species isolation in the multidimensional functional space<sup>16,17</sup>. This index is calculated by quantifying the distance of each species in relation to its nearest neighbour<sup>16</sup>. This mathematical approach applied to multidimensional functional space was adapted

from the mean nearest neighbour distance developed initially to calculate the phylogenetic distance between species<sup>65</sup> (see equation descriptions at the Supplementary Fig. 2).

### Step 3—Selecting rare species & Step 4—Rarity biogeography

**Quartile analysis.** We examined the distribution of values for species restrictedness and functional distinctiveness (or species restrictedness and functional uniqueness) and used the quartile criterion (performed using the base R function “quantile” from the package “stats” in R Core Team<sup>66</sup>) as proposed by Gaston<sup>20</sup> to identify the rare species. By this definition the species considered rare lie in the top quartile of both metrics (i.e. values between 0 (less restricted) and 1 (more restricted)). We next assigned the observed number of rare species (Step 4), as defined above, to each 2° grid cell. The analysis was undertaken separately for Actinopterygii and Elasmobranchii.

**Step 5—Null model.** Does the number of rare species in a given grid cell differ from the null expectation? To answer this question we applied a null model

approach based on the curveball algorithm<sup>67</sup>. This algorithm keeps constant the total number of species (rare + non-rare) and the number of grid cells that each species occurs. It then randomizes the presence and absence of all species following these thresholds. We ran the model for 2000 iterations; in each loop it randomizes the occurrences of all species, identifies where the rare species are falling and then counts the total number of rare species in each grid cell.

To quantify how the observed number of rare species differ from the null expectation we then use Standardized Effect Sizes (SES) as follows:

$$SES = (X - Y)/Z$$

X as the number of rare species observed in each grid cell,

Y as the average of rare species found from the null model after 2000 interactions and

Z as the standard deviation from Y.

A positive SES indicates more rare species than would be expected by chance and a negative SES fewer than expected.

We are using 14 different systems (7 Coastal Systems and 7 High Seas (see Supplementary Fig. 3)) for 2 groups of organisms (bony and cartilaginous fish), 2 functional rarity indices (distinctiveness and uniqueness), and using 3 different probabilities of occurrences (prob. >0.9, >0.7 and >0.5).

We then have the following “roadmap”:

- i. Scales—7 Coastal Systems and 7 High Seas.
- ii. Groups—bony and cartilaginous fish.
- iii. Indices—distinctiveness and uniqueness.
- iv. Probability of occurrences—>0.9, >0.7 and >0.5.
- v. Total—168 independent cases analysis (each having its own assemblage and trait matrices).

Therefore, as mentioned above, the null model ran for 2000 iterations in each of those independent cases. The final matrices from these initial steps contain grid cells as rows and as columns we have the raw number of rare species along with the SES values for each. These matrices were important to map our results.

**Step 6—Mapping the results.** After the above steps (and using the matrices with the results), to visualise the results for Coastal Systems we plotted the geographic distribution of rarity, measured using the observed number of rare species, based on species Restrictedness and functional Distinctiveness using Fig. 1a, c, and the results from the SES using Fig. 1b, d. Meanwhile in Fig. 2 we constructed the same plots for the High Seas. The complementary results of the alternative approach using species Restrictedness and functional Uniqueness are shown in Supplementary Fig. 4 (for Coastal Systems) and Supplementary Fig. 5 (for High Seas).

The flow chart in Supplementary Fig. 1 provides step by step details of what was done for each of the 168 independent cases explained at the “roadmap” above. The comprehensive list of all rare fish species found for each system is available in Supplementary Table 2.

## Further analyses

**Latitudinal rarity biogeography.** We then produced the density plots of the positive SES values (using the function *geom\_density* from the package *ggplot2*<sup>68</sup>) to further understand these patterns in relation to latitude (Fig. 3, from c to j). These were compared with the latitudinal gradient of species richness (Fig. 3a, b). The main text discusses results focused on rarity measured using the probability of occurrence higher than 0.9 (Fig. 3). We also examined density of positive SES values across the latitudinal distribution using the probability of occurrences higher than 0.7 and 0.5 (see Supplementary Fig. 6a–d for probability >0.7 and Supplementary Fig. 6e–f for probability >0.5).

**Spatial autocorrelation.** We constructed distance decay plots to examine spatial autocorrelation, and fitted a quantile regression to these relationships. The results are illustrated in Supplementary Fig. 7, which shows the distance decay calculated by pairwise differences (Supplementary Fig. 7a—Coastal Systems and b—High Seas for bony fish, and Supplementary Fig. 7c—Coastal Systems and d—High Seas for cartilaginous fish) between a given grid cell and all other grid cells present in the Northwest Pacific Ocean. These plots provide reassurance that spatial autocorrelation is not obscuring the results we report.

**Sensitivity analysis.** We performed a sensitivity analysis to ensure that the environmental traits “Depth” and “Mean Temperature Preference” had no major influence on determining the level of distinctiveness and uniqueness of the species. We did this by excluding each trait in turn from the analysis (each of those were removed individually and a third time without both together) and compared the results with the full analysis. We found strong correlations in rarity estimates in all cases (see Supplementary Fig. 8a, b (for bony fish (distinctiveness and uniqueness) respectively, c and d (for cartilaginous fish (distinctiveness and uniqueness)).

**Trait correlation analysis.** We tested the correlation between traits to ensure that those were largely uncorrelated, as shown in Supplementary Fig. 9a, b.

**Supplementary analysis.** We tested the possible influence of sampling effect on the rarity hotspots observed by creating random fill matrices and comparing those with the observed matrices from four scenarios: Northwest Pacific Coast (bony and cartilaginous fish species) and Southwest Pacific Coast (bony and cartilaginous species). The subsequent results showed no evidence of sampling effect (see Supplementary Fig. 11).

**Mapping marine protected areas (MPAs).** We used the MPAs shapefiles provided by the UNEP-WCMC and IUCN<sup>69</sup> to measure the level of congruence between marine protected areas and hotspots of rarity. The distances between each MPA and the centroid of each grid cell were calculated using the spatial analysis tool in ArcGIS (the unit of the distance calculated is in decimal degrees). We then assigned each MPA to its nearest 2° degree grid cell centroid (the distance cut point used was < 0.75 decimal degrees (the distance from a given MPA to a grid cell centroid)). We plotted these global spatial patterns from the 2° grid cells indicating either congruence or mismatches between Marine Protected Areas (MPAs) and Rarity Hotspots (species rare in both dimensions of biodiversity; taxonomically—highest restrictedness and functionally—highest distinctiveness) (Fig. 4a and b, bony and cartilaginous fish respectively).

**Habitat specialization.** All species were classified according to their position in the water column (bathypelagic, bathydemersal, benthopelagic, demersal, pelagic neritic, pelagic oceanic and reef associated (as categorized in FishBase<sup>13</sup>)); here we are using this trait as a proxy for “habitat specialization”. We then used a G test, and Cramer’s V (using the functions *GTest* and *CramerV* from the package *DescTools*<sup>70</sup>) to compare the frequency distribution in habitat specialization between rare and non-rare species (see Supplementary Fig. 10 for all frequency distributions and statistical results).

**Forms of functional rarity classification scheme.** In their 2017 paper Violle et al.<sup>17</sup>, suggested 12 forms of functional rarity. We believe that the approach applied here is similar to the classification scheme they described as: “Rare traits irrespective of the scale and the species pool”. The authors pointed to two possible extremes: rare traits (exhibited by range-restricted species) and common traits (supported by many widespread species). In this case, our approach identifies species that are both geographically restricted within each of the 14 systems (coastal and high seas systems) and present a distinct (or unique) combination of traits. Our approach to the classification of rarity differs slightly, however, in that we follow Gaston’s approach<sup>20</sup> of quantile distribution as illustrated in Supplementary Fig. 1, step 3 QUANTILES.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The datasets generated during and/or analysed in this study are available in the Research Portal of the University of St Andrews repository, <https://doi.org/10.17630/397bc872-f7de-4ded-9ed8-4f734c11b14a><sup>71</sup>.

## Code availability

Code is provided in the Supplementary Methods section of the Supplementary Information file.

Received: 2 March 2021; Accepted: 28 January 2022;

Published online: 15 February 2022

## References

- Darwin, C. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. Vol. 11 (London: John Murray [1st edition]. Harvard University Press, 1859).
- Bates, H. W. The naturalist on the river Amazons. (London; Toronto: JM Dent; New York: EP Dutton, [1910, reprinted 1921], 1921).
- Rabinowitz, D. (John Wiley and Sons: Chichester, UK, 1981).
- Mouillot, D. et al. Rare species support vulnerable functions in high-diversity ecosystems. *PLoS Biol.* **11**, e1001569 (2013).
- Bracken, M. E. & Low, N. H. Realistic losses of rare species disproportionately impact higher trophic levels. *Ecol. Lett.* **15**, 461–467 (2012).
- Dee, L. E. et al. When do ecosystem services depend on rare species? *Trends Ecol. Evolution* **34**, 746–758 (2019).
- Grenié, M. et al. Functional rarity of coral reef fishes at the global scale: hotspots and challenges for conservation. *Biol. Conserv.* **226**, 288–299 (2018).
- Pimiento, C. et al. Functional diversity of marine megafauna in the Anthropocene. *Sci. Adv.* **6**, eaay7650 (2020).
- Enquist, B. J. et al. The commonness of rarity: global and future distribution of rarity across land plants. *Sci. Adv.* **5**, eaaz0414 (2019).

10. Loiseau, N. et al. Global distribution and conservation status of ecologically rare mammal and bird species. *Nat. Commun.* **11**, 1–11 (2020).
11. Connolly, S. R. et al. Commonness and rarity in the marine biosphere. *Proc. Natl Acad. Sci. USA* **111**, 8524–8529 (2014).
12. Díaz, S. et al. Incorporating plant functional diversity effects in ecosystem service assessments. *Proc. Natl Acad. Sci. USA* **104**, 20684–20689 (2007).
13. Díaz, S. & Cabido, M. Vive la différence: plant functional diversity matters to ecosystem processes. *Trends Ecol. Evol.* **16**, 646–655 (2001).
14. Hidasi-Neto, J., Loyola, R. & Cianciaruso, M. V. Global and local evolutionary and ecological distinctiveness of terrestrial mammals: identifying priorities across scales. *Diversity Distrib.* **21**, 548–559 (2015).
15. Trindade-Santos, I. et al. Simulating shifts in taxonomic and functional  $\beta$ -diversity of ray-finned fishes: probing the Mariana disaster. *Perspect. Ecol. Conserv.* **16**, 186–192 (2018).
16. Grenié, M., Denelle, P., Tucker, C. M., Munoz, F. & Violle, C. funrar: An R package to characterize functional rarity. *Diversity Distrib.* **23**, 1365–1371 (2017).
17. Violle, C. et al. Functional rarity: the ecology of outliers. *Trends Ecol. Evolution* **32**, 356–367 (2017).
18. Kaschner, K. et al. AquaMaps: Predicted range maps for aquatic species. Retrieved from [www.aquamaps.org](http://www.aquamaps.org), (Version October, 2019).
19. Froese, R. & Pauly, D. Editors. FishBase. World Wide Web electronic publication, [www.fishbase.org](http://www.fishbase.org), (Version February, 2021).
20. Gaston, K. J. in *Rarity. Population and Community Biology Series Vol. 13* 1–21 (Springer, 1994).
21. Stuart-Smith, R. D. et al. Integrating abundance and functional traits reveals new global hotspots of fish diversity. *Nature* **501**, 539–542 (2013).
22. Ricotta, C. et al. Measuring the functional redundancy of biological communities: a quantitative guide. *Methods Ecol. Evolution* **7**, 1386–1395 (2016).
23. Gaston, K. J. Global patterns in biodiversity. *Nature* **405**, 220–227 (2000).
24. Dornelas, M. et al. BioTIME: a database of biodiversity time series for the Anthropocene. *Glob. Ecol. Biogeogr.* **27**, 760–786 (2018).
25. Floeter, S. R. et al. Atlantic reef fish biogeography and evolution. *J. Biogeogr.* **35**, 22–47 (2008).
26. Floeter, S. R., Krohling, W., Gasparini, J. L., Ferreira, C. E. & Zalmon, I. R. Reef fish community structure on coastal islands of the southeastern Brazil: the influence of exposure and benthic cover. *Environ. Biol. Fishes* **78**, 147–160 (2007).
27. Pinheiro, H. T. et al. Deep reef fishes in the world's epicenter of marine biodiversity. *Coral Reefs* **38**, 985–995 (2019).
28. Pinheiro, H. T. et al. South-western Atlantic reef fishes: zoogeographical patterns and ecological drivers reveal a secondary biodiversity centre in the Atlantic Ocean. *Diversity Distrib.* **24**, 951–965 (2018).
29. Smith, T. D. *Scaling fisheries: the science of measuring the effects of fishing, 1855–1955*. (Cambridge University Press, 1994).
30. ICES. Database of Trawl Surveys (DATRAS). International Council for the Exploration of the Sea (ICES) <https://www.ices.dk/data/data-portals/Pages/DATRAS.aspx> (2021).
31. Moyes, F. & Magurran, A. E. Change in the dominance structure of two marine-fish assemblages over three decades. *J. Fish. Biol.* **94**, 96–102 (2019).
32. Kesner-Reyes, K., C. Garilao, K. Kaschner, J. Barile & Froese, R. AquaMaps: concept, algorithms and data sources for modelling the distribution of marine organisms. In: R. Froese & D. Pauly (Eds). [www.fishbase.org](http://www.fishbase.org), (Version October, 2019), [https://www.aquamaps.org/main/AquaMaps\\_Algorithm\\_and\\_Data\\_Sources.pdf#page=1](https://www.aquamaps.org/main/AquaMaps_Algorithm_and_Data_Sources.pdf#page=1) (2019).
33. May, R. M. Biological diversity: differences between land and sea. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **343**, 105–111, (1994).
34. Rabosky, D. L. et al. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).
35. Stein, A., Gerstner, K. & Kreft, H. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecol. Lett.* **17**, 866–880 (2014).
36. Massicotte, P., Proulx, R., Cabana, G. & Rodriguez, M. A. Testing the influence of environmental heterogeneity on fish species richness in two biogeographic provinces. *PeerJ* **3**, e760 (2015).
37. Lavergne, S., Garnier, E. & Debussche, M. Do rock endemic and widespread plant species differ under the Leaf-Height-Seed plant ecology strategy scheme? *Ecol. Lett.* **6**, 398–404 (2003).
38. Leitão, R. P. et al. Rare species contribute disproportionately to the functional structure of species assemblages. *Proc. R. Soc. B: Biol. Sci.* **283**, 20160084 (2016).
39. Cooke, R. S., Eigenbrod, F. & Bates, A. E. Ecological distinctiveness of birds and mammals at the global scale. *Glob. Ecol. Conserv.* **22**, e00970 (2020).
40. Dulvy, N. K. et al. Challenges and priorities in shark and ray conservation. *Curr. Biol.* **27**, R565–R572 (2017).
41. Dulvy, N. K. et al. Extinction risk and conservation of the world's sharks and rays. *elife* **3**, 34 (2014).
42. Trindade-Santos, I., Moyes, F. & Magurran, A. E. Global change in the functional diversity of marine fisheries exploitation over the past 65 years. *Proc. R. Soc. B* **287**, 20200889 (2020).
43. Murgier, J. et al. Rebound in functional distinctiveness following warming and reduced fishing in the North Sea. *Proc. R. Soc. B* **288**, 20201600 (2021).
44. GBIF. Global Biodiversity Information Facility (GBIF), Free and open access to biodiversity data. Home Page available from: <https://www.gbif.org>, (2020).
45. OBIS. Ocean Biodiversity Information System. Intergovernmental Oceanographic Commission of UNESCO. Available from: [www.iobis.org](http://www.iobis.org), (2020).
46. Palomares, M. L. D., Pauly, D. & Editors. SeaLifeBase. World Wide Web electronic publication. [www.sealifebase.org](http://www.sealifebase.org), (Version February, 2019).
47. Environmental Systems Research Institute (ESRI). ArcGIS Release 10.7. Redlands, CA. (2021).
48. Schaus, M. H. & Vanni, M. J. Effects of gizzard shad on phytoplankton and nutrient dynamics: role of sediment feeding and fish size. *Ecology* **81**, 1701–1719 (2000).
49. Mouillot, D. et al. Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proc. Natl Acad. Sci. USA* **111**, 13757–13762 (2014).
50. Givan, O., Edelist, D., Sonin, O. & Belmaker, J. Thermal affinity as the dominant factor changing Mediterranean fish abundances. *Glob. Change Biol.* **24**, e80–e89 (2018).
51. Pauly, D. & Cheung, W. W. Sound physiological knowledge and principles in modeling shrinking of fishes under climate change. *Global Change Biology*, <https://doi.org/10.1111/gcb.13831> (2017).
52. Cheung, W. W. L., Watson, R. & Pauly, D. Signature of ocean warming in global fisheries catch. *Nature* **497**, 365–368 (2013).
53. Palomares, M. L. D. & Pauly, D. Predicting food consumption of fish populations as functions of mortality, food type, morphometrics, temperature and salinity. *Mar. Freshw. Res.* **49**, 447–453 (1998).
54. Pauly, D. Food consumption by tropical and temperate fish populations: some generalizations. *J. Fish. Biol.* **35**, 11–20 (1989).
55. Pauly, D., Christensen, V., Dalsgaard, J., Froese, R. & Torres, F. Jr. Fishing down marine food webs. *Science* **279**, 860–863 (1998).
56. Wiedmann, M. A. et al. Functional diversity of the Barents Sea fish community. *Mar. Ecol. Prog. Ser.* **495**, 205–218 (2014).
57. Webb, P. W. Form and function in fish swimming. *Sci. Am.* **251**, 72–83 (1984).
58. Vazzoler, A. E. A. M. *Biologia da reprodução de peixes teleosteos: teoria e prática*. (Nupelia, 1996).
59. Fonteles-Filho, A. A. *Recursos pesqueiros: biologia e dinâmica populacional*. (Imprensa Oficial do Ceará, 1989).
60. Winemiller, K. O. Life history strategies, population regulation, and implications for fisheries management. *Can. J. Fish. Aquat. Sci.* **62**, 872–885 (2005).
61. Jakobsen, T., Fogarty, M. J., Megrey, B. A. & Moksness, E. Fish reproductive biology: implications for assessment and management. (John Wiley & Sons, 2016).
62. Trindade-Santos, I. & Freire, K. D. M. F. Analysis of reproductive patterns of fishes from three Large Marine Ecosystems. *Frontiers in Marine Science* **2**, <https://doi.org/10.3389/fmars.2015.00038> (2015).
63. Penone, C. et al. Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods Ecol. Evolution* **5**, 961–970 (2014).
64. Stekhoven, D. J. & Bühlmann, P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2011).
65. Webb, C. O., Ackerly, D. D., McPeck, M. A. & Donoghue, M. J. Phylogenies and community ecology. *Annual review of ecology and systematics*, 475–505, (2002).
66. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, <http://www.R-project.org/>, Vienna, Austria, 2018).
67. Strona, G., Nappo, D., Boccacci, F., Fattorini, S. & San-Miguel-Ayanz, J. A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nat. Commun.* **5**, 1–9 (2014).
68. ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York, 2016).
69. UNEP-WCMC & IUCN. Protected Planet: The World Database on Protected Areas (WDPA) [Online], September 2020, Cambridge, UK: UNEP-WCMC and IUCN. Available at: [www.protectedplanet.net](http://www.protectedplanet.net), (2020).
70. Signorell, A. et al. DescTools: Tools for descriptive statistics. R package version 0.99 18 (2016).
71. Trindade-Santos, I., Moyes, F. & Magurran, A. Data Underpinning Global patterns in functional rarity of marine fish. University of St Andrews, <https://doi.org/10.17630/397bc872-f7de-4ded-9ed8-4f734c11b14a> (2022).



## Acknowledgements

I.T.S. thanks CAPES (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* -Coordination for the Improvement of Higher Education Personnel), process number: #88881.129579/2016-01 (Finance Code 001) for a PhD scholarship. A.E.M. thanks the Leverhulme Trust (RPG-2019-402) for support. We thank the FishBase team (The Philippines and Nicolas Bailly) and the AquaMaps team (Cristina Garilao) for making their data available, the “Fish Lunch” group at the Centre for Biological Diversity (University of St Andrews), Daniel Pauly for their helpful comments, Ylenia Cascone for implementing the null model in Python and Piter Nascimento West for fish drawings. I.T.S. thanks Kátia de Meirelles Felizola Freire and Sidney Feitosa Gouveia for earlier guidance, research opportunities and inspiration.

## Author contributions

I.T.S. developed the concept of the paper in conjunction with A.E.M and F.M. All authors contributed to the analyses and to the writing of the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-28488-1>.

**Correspondence** and requests for materials should be addressed to Isaac Trindade-Santos.

**Peer review information** *Nature Communications* thanks Sergio Floeter, Juan Quimbayo, and Giovanni Strona for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022