



**HAL**  
open science

## Cross-ocean patterns and processes in fish biodiversity on coral reefs through the lens of eDNA metabarcoding

Laëtitia Mathon, Virginie Marques, David Mouillot, Camille Albouy, Marco Andrello, Florian Baletaud, Giomar H. Borrero-Pérez, Tony Dejean, Graham J. Edgar, Jonathan Grondin, et al.

### ► To cite this version:

Laëtitia Mathon, Virginie Marques, David Mouillot, Camille Albouy, Marco Andrello, et al.. Cross-ocean patterns and processes in fish biodiversity on coral reefs through the lens of eDNA metabarcoding. *Proceedings of the Royal Society B: Biological Sciences*, 2022, 289 (1973), pp.20220162. 10.1098/rspb.2022.0162 . hal-03669624

**HAL Id: hal-03669624**

<https://hal.umontpellier.fr/hal-03669624v1>

Submitted on 23 Nov 2023

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

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

1 **Cross-ocean patterns and processes in fish biodiversity on coral reefs through the lens**  
2 **of eDNA metabarcoding**

3  
4 Laetitia Mathon<sup>1,2\*†‡</sup>, Virginie Marques<sup>1,3†</sup>, David Mouillot<sup>3,4</sup>, Camille Albouy<sup>5</sup>, Marco  
5 Andrello<sup>3,17</sup>, Florian Baletaud<sup>2,3,6</sup>, Giomar H. Borrero-Pérez<sup>7</sup>, Tony Dejean<sup>8</sup>, Graham J. Edgar<sup>9</sup>,  
6 Jonathan Grondin<sup>8</sup>, Pierre-Edouard Guerin<sup>1</sup>, Régis Hocdé<sup>3</sup>, Jean-Baptiste Juhel<sup>3</sup>, Kadarusman<sup>10</sup>,  
7 Eva Maire<sup>3,11</sup>, Gael Mariani<sup>3</sup>, Matthew McLean<sup>12</sup>, Andrea Polanco F.<sup>7</sup>, Laurent Pouyaud<sup>13</sup>, Rick  
8 D. Stuart-Smith<sup>9</sup>, Hagi Yulia Sugeha<sup>14</sup>, Alice Valentini<sup>8</sup>, Laurent Vigliola<sup>2</sup>, Indra B Vimono<sup>14</sup>,  
9 Loïc Pellissier<sup>15,16‡</sup>, Stéphanie Manel<sup>1‡</sup>

10  
11  
12 <sup>1</sup> CEFE, Univ. Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France

13 <sup>2</sup> ENTROPIE, Institut de Recherche pour le Développement (IRD), Univ. Réunion, UNC, CNRS, IFREMER, Centre  
14 IRD de Nouméa, Nouméa, New Caledonia, France

15 <sup>3</sup> MARBEC, Univ Montpellier, CNRS, IFREMER, IRD, Montpellier, France

16 <sup>4</sup> Institut Universitaire de France

17 <sup>5</sup> DECOD (Ecosystem Dynamics and Sustainability), IFREMER, INRAE, Institut Agro - Agrocampus Ouest ,  
18 Nantes, France.

19 <sup>6</sup> SOPRONER, groupe GINGER, 98000 Noumea, New Caledonia, France

20 <sup>7</sup> Instituto de Investigaciones Marinas y Costeras-INVEMAR, Colombia. Museo de Historia Natural Marina de  
21 Colombia (MHNMC), Programa de Biodiversidad y Ecosistemas Marinos. Calle 25 No. 2 – 55 Playa Salguero, Santa  
22 Marta, Colombia.

23 <sup>8</sup> SPYGEN, Le Bourget-du-Lac, France

24 <sup>9</sup> Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

25 <sup>10</sup> Politeknik Kelautan dan Perikanan Sorong, KKD BP Sumberdaya Genetik, Konservasi dan Domestikasi, Papua  
26 Barat, Indonesia

27 <sup>11</sup> Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK

28 <sup>12</sup> Department of Biology, Dalhousie University, Halifax, NSB3H4R2, Canada

29 <sup>13</sup> ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

30 <sup>14</sup> Research Center for Oceanography, National Research and Innovation Agency, Jl. Pasir Putih 1, Ancol Timur,  
31 Jakarta Utara 14430, Indonesia

32 <sup>15</sup> Landscape Ecology, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH  
33 Zürich, Zürich, Switzerland

34 <sup>16</sup> Unit of Land Change Science, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

35 <sup>17</sup> Institute for the study of Anthropic Impacts and Sustainability in the marine environment, National Research Council  
36 (CNR-IAS), Rome, Italy

37  
38 \* Corresponding authors: laetitia.mathon@gmail.com

39 † These authors contributed equally as first author to this work

40 ‡ These authors contributed equally as senior author to this work

## 41 **Abstract**

42 Increasing speed and magnitude of global change threaten the world's biodiversity and particularly  
43 coral reef fishes. A better understanding of large-scale patterns and processes on coral reefs is  
44 essential to prevent fish biodiversity decline but it requires new monitoring approaches. Here, we  
45 use environmental DNA metabarcoding to reconstruct well-known patterns of fish biodiversity on  
46 coral reefs and uncover hidden patterns on these highly diverse and threatened ecosystems. We  
47 analyzed 226 eDNA seawater samples from 100 stations in 5 tropical regions (Caribbean, Central  
48 and Southwest Pacific, Coral Triangle and Western Indian ocean) and compared those to 2,047  
49 underwater visual censuses from the Reef Life Survey (RLS) in 1,224 stations. Environmental  
50 DNA reveals a higher (16%) fish biodiversity, with 2,650 taxa, and 25% more families than  
51 underwater visual surveys. By identifying more pelagic, reef-associated and crypto-benthic  
52 species, eDNA offers a fresh view on assembly rules across spatial scales. Nevertheless, RLS  
53 identified more species than eDNA in 47 shared families, which can be due to incomplete sequence  
54 assignment, possibly combined with incomplete detection in the environment, for some species.  
55 Combining eDNA metabarcoding and extensive visual census offers novel insights on the spatial  
56 organization of the richest marine ecosystems.

## 1. Introduction

57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80

Coral reefs host the highest fish diversity on earth despite covering less than 0.1% of the ocean's surface (1,2). They are also severely threatened (3), with near future outlooks predominantly pessimistic (4). Data syntheses over decades of surveys estimate the total number of coral reef fishes from 2,400 to 8,000 species (5,6), distributed among approximately 100 families (7). Typically, coral reef biodiversity displays clear spatial patterns, including longitudinal and latitudinal gradients outwards the Indo-Australian Archipelago (8,9), also known as the 'Coral Triangle', hosting the world's highest level of marine biodiversity (10). The exceptional biodiversity in the Coral Triangle has recently been suggested to strongly relate to higher diversity among fish families that feed on plankton (11). Other trophic groups are also very important on coral reefs but are often undetected because they are transient or hidden (12,13). Intriguingly, the proportions of fish species among families are shown to be strongly conserved across the Indo-Pacific (8). The spatial patterns of coral reef fishes are also marked by strong variations in taxonomic composition (species turnover or  $\beta$  diversity), often due to isolation (14). Many species on coral reefs are geographically localized, but can sometimes be locally abundant, while others are widespread (15).

Coral reef fishes have evolved in a physically complex environment and present a wide range of forms and functions (16). Small cryptic species, hereafter called crypto-benthic, that live inside the reef structure, can be very difficult to sample or survey using non-destructive methods (17), yet represent half of the fish diversity on coral reefs (13). Even though fishes are among the best-studied taxa inhabiting coral reefs (18), our knowledge of their biodiversity is only partial (19), the taxonomy is complex, uncertain for many species (5), and countless species remain undescribed.

81 Environmental DNA (eDNA) metabarcoding, a method retrieving and analyzing DNA naturally  
82 released by organisms in their environment (20), provides an opportunity to not only better  
83 understand classical biodiversity patterns, but also uncover novel ones hidden by our incomplete  
84 taxonomic and biogeographic coverage (21). Environmental DNA is particularly powerful in  
85 aquatic ecosystems (22) and is now well established for marine microorganisms (23,24). By  
86 contrast, its potential to provide an integrated biodiversity assessment of macroorganisms,  
87 including vertebrates of all trophic levels (from crypto-benthic to large pelagic fish species), is  
88 only shown at local (25) and regional (26–30) scales but not yet at spatial scales including more  
89 than one biogeographic region or multiple ocean basins.

90 Here, we investigate how a cross-ocean basin snapshot of eDNA sampling could describe the  
91 distribution of fish biodiversity on coral reefs, reveal unknown patterns, and challenge well-  
92 established assembly rules. From 226 eDNA seawater samples (2,712 PCR replicates) collected  
93 in 100 stations at 26 sites covering five tropical regions (Southeast Polynesia, Tropical  
94 Northwestern Atlantic, Tropical Southwestern Pacific, Western Indian Ocean and Western Coral  
95 Triangle) across the Indian, Pacific and Atlantic Oceans (figure S1-S2), we produced a final dataset  
96 of 189,350,273 mitochondrial 12S rRNA gene sequence reads (see Methods), clustered into 2,023  
97 molecular operational taxonomic units (MOTUs), and assigned to Actinopterygii (bony fishes)  
98 and Chondrichthyes (cartilaginous fishes) taxa (tables S1- S2). We then compared fish biodiversity  
99 patterns obtained from eDNA to those observed from 2,047 standardized visual surveys of reef  
100 fishes in 1224 stations at 219 sites within 24 tropical regions (31).

101

## 102 **2. Results**

### 103 **(a) Global estimates of fish biodiversity on coral reefs**

104 We estimated total fish diversity on coral reefs using the asymptote of a multi-model accumulation  
105 curve for both eDNA MOTUs (32) and visual census species (Methods). The asymptote estimated  
106 from 100 eDNA stations distributed in five regions sampled over a 28-month period reaches 2,650  
107 MOTUs (figure 1a). This detectable fish MOTUs diversity, including also MOTUs unassigned at  
108 the species-level, is 16% higher than the estimate from visual census data, which reaches an  
109 asymptote at 2,268 fish species from 2,047 tropical transects surveyed during 13 years (figure 1b).  
110 The asymptotic estimation of family richness obtained with eDNA reaches 147 families, 25% more  
111 than the asymptotic number of families estimated with visual census data (118 families, figure 1c-  
112 d). Among the 71 families shared between both datasets, 24 have a higher number of MOTUs from  
113 eDNA survey than species from visual survey while 47 have more species from visual survey than  
114 MOTUs from eDNA survey (figure 1e). Families with more taxa identified using eDNA include  
115 those often associated with reef-adjacent habitats such as mangroves or soft sediments like  
116 Mugilidae (e.g. *Mugil rubrioculus*), Elopidae and Gerreidae (33, e.g. *Gerres oyena*), and crypto-  
117 benthic species that live hidden in crevices (e.g. Gobiidae) or nocturnal fish species (34, e.g.  
118 Congridae). Families with more taxa with visual census include Acanthuridae, Chaetodontidae,  
119 Blenniidae, Labridae, Pomacentridae and Scaridae. Fifty-five families are detected only with  
120 eDNA, including Myctophidae, Engraulidae, Atherinidae and Exocoetidae, while 24 families are  
121 detected only by visual census, including Caesionidae, Chaenopsidae, Labrisomidae and  
122 Microdesmidae. Environmental DNA estimates a diversity of crypto-benthic species 13% higher  
123 than with visual census, and, among many others, includes species such as the elegant firefish  
124 (*Nemateleotris decora*), which lives on the outer reef slope between 25 to 70 m (figure 2a). Yet,  
125 the difference in fish diversity assessment between the two methods is the strongest for pelagic  
126 and wide-ranging species, for which eDNA reveals more than 7 times higher richness than with

127 visual census. These species mainly belong to Scombridae (e.g. *Katsuwonus pelamis*), Clupeidae,  
128 Carcharhinidae (e.g. *Carcharhinus leucas*, *Sphyrna lewini*) and Belonidae (figure 2b).

129

130 MOTU richness per fish family retrieved with eDNA is strongly correlated with fish species  
131 richness within families recorded in visual census data (Pearson correlation = 0.84,  $p < 0.001$ ,  $n =$   
132 71, figure 1e). Highly diverse families seen on coral reefs are also well represented in eDNA  
133 samples, with Gobiidae, Labridae and Pomacentridae containing more than 100 MOTUs each,  
134 together representing about 20% of MOTUs (figure 1f, figures S3- S4). The slope of the log-log  
135 relationship between MOTUs richness per family and species richness per family is equal to 0.8  
136 showing that the relationship is not proportional but saturating. The richest fish families contain  
137 more MOTUs detected with eDNA than species detected with visual surveys.

138

### 139 (b) Biogeography of eDNA sequences

140 The spatial distribution of MOTUs follows clear biogeographic patterns, with a peak in the coral  
141 triangle and lower values of MOTU richness toward Southeast Polynesia (figure S5). The richest  
142 region (West Papua, Indonesia, Western Coral Triangle) contains ~50% of the global pool of fish  
143 MOTUs while the poorest region (Fakarava, French Polynesia, Southeast Polynesia) contains only  
144 9% of the global pool (figures S6-S7 and table S2). Distance-based Redundancy Analysis  
145 (dbRDA) was performed on fish family proportions at each site (i.e. number of MOTUs or species  
146 assigned to each family in each site, see Methods) for eDNA and visual surveys with the region  
147 and the site MOTU/species richness as explanatory variables, including their interaction (figure 3,  
148 table S3). For eDNA, the dbRDA explains up to 42% of variation in family proportions between  
149 pairs of sites with region and MOTU/species richness both having significant effects ( $F = 4.1$  and  
150 5.7, respectively,  $p < 0.001$ ), but no significant interaction ( $F = 1.99$ ,  $p > 0.05$ ). The partial dbRDA

151 on eDNA showed a significant effect of region while controlling for MOTU richness ( $F = 2.79$ ,  $p$   
152  $< 0.001$ ). The first axis explains 17.2% of variation in family proportions and separates the Western  
153 Coral Triangle from other regions (figure 3*a-b*). The first axis shows a higher proportion of  
154 Lutjanidae but lower proportions of Labridae and Gobiidae in sites of the Western Coral Triangle.  
155 It also confirms the longitudinal diversity gradient from the Coral Triangle. The second axis  
156 explains 11.2% of variation and discriminates the Tropical Northwestern Atlantic from the  
157 Western Indian Ocean, due to a higher proportion of Clupeidae and Carangidae in the Atlantic  
158 Ocean and a higher proportion of Acanthuridae in the Indian Ocean. The dbRDA performed on  
159 visual census data explained greater variation ( $R^2 = 0.5$ ,  $p < 0.001$ ) and the region also had a  
160 significant, albeit weaker than for MOTUs, effect on fish family proportions ( $F = 17.7$ ,  $p < 0.01$ ),  
161 while species richness and interaction between the two variables also had significant effects ( $F =$   
162  $6.28$  and  $2$ ,  $p < 0.01$  respectively). The first axis explains 41.6% of variance in family proportions  
163 and separates the Tropical Northwestern Atlantic from the other regions with a higher proportion  
164 of Gobiidae and Serranidae. The second axis explains 5.7% of variance in family proportions and  
165 separates the Southeast Polynesia from Indo-Pacific regions, and is mostly driven by the higher  
166 proportion of Pomacentridae in the Indo-Pacific (figure 3*c-d*).

167

### 168 (c) **Global patterns of fish turnover and rarity**

169 Our eDNA survey shows that a majority of MOTUs are geographically restricted, with 85% of the  
170 MOTUs detected in only one region (figure 4*a*), and 35% in only one site (figure S8). Geographic  
171 restriction is one aspect of species rarity but is shown to play a primary role in determining  
172 extinction risk while local abundance and habitat specialization have secondary roles (35). We  
173 hierarchically partitioned the global MOTU diversity ( $\gamma_{global}$ ) into additive diversity components



174 (i.e. dissimilarity) due to difference between regions ( $\beta_{inter-region}$ ), mean difference between  
175 sites within regions ( $\bar{\beta}_{inter-site}$ ), mean difference between stations within sites ( $\bar{\beta}_{inter-station}$ )  
176 and mean station diversity ( $\bar{\alpha}_{station}$ ) (36). As a consequence of the geographic restriction of most  
177 MOTUs to one region, the total fish MOTU ( $\gamma$ ) diversity is mainly due to inter-region  $\beta$ -diversity  
178 (~74%) followed by inter-site (14.8%) and inter-station (5.9%)  $\beta$ -diversity (figure 4b). The same  
179 partitioning using different site delineations (10 and 20 km) provides similar results (table S4).  
180 Diversity partitioning of crypto-benthic fish MOTUs only or pelagic fish MOTUs only reveals  
181 similar patterns (table S5). The partitioning diversity of species detected by visual census also  
182 revealed similar patterns but with a stronger effect of  $\beta_{inter-region}$  (84%) and lower (3x)  
183  $\bar{\beta}_{inter-site}$  and  $\bar{\beta}_{inter-station}$  (table S5, figure S9).

184

185 Beyond the hierarchical partitioning of diversity, we compared the distribution of fish MOTUs  
186 and species visual occurrences independently of the survey method and sampling effort using  
187 global species abundance distributions (gSAD) (37). We fitted the fish MOTU and species visual  
188 occurrences to three distributions (log-series, Pareto and Pareto with exponential finite adjustment,  
189 *i.e.* Pareto Bended, see Methods) and estimated the parameters by maximum likelihood. For the  
190 visual census gSAD, the best fit was obtained with the log-series and Pareto distributions (table  
191 S6) with a slope of -0.95 (confidence interval at 95% [-0.98;-0.92]) (figure S10). This suggests a  
192 distribution of geographically restricted or rare species close to the neutral theory ( $\beta$  close to -1).  
193 By contrast, the best fit for fish MOTUs was obtained with the Pareto Bended distribution with a  
194 slope  $\beta = -0.76$  (confidence interval at 95% [-0.85;-0.65]) and then with the log-series distribution,  
195 suggesting a lower prevalence of rarity than under the neutral theory, in agreement with previous  
196 tests based on species distributions on coral reefs (38).

197

### 3. Discussion

199 Environmental DNA allows the detection and identification of more taxa than traditional  
200 techniques (26,39), but further offers novel insights on the spatial organization of the richest  
201 marine ecosystem at large scale. Over a timespan of 2.3 years, in major tropical ocean basins,  
202 eDNA metabarcoding reveals a higher proportion of crypto-benthic, pelagic and soft-sediment-  
203 associated fishes on coral reefs than detected in the most extensive visual census over 13 years.  
204 We found a high local MOTU turnover, but we were not able to conclude if it is due to an  
205 insufficient sampling at the station level, or if it suggests that differences in fish species  
206 composition may exist between adjacent reefs that are not detected by visual surveys (40), so that  
207 fish biodiversity is more patchy than previously thought on coral reefs.

208

209 We were also able to retrieve well-known patterns of fish diversity on coral reefs such as the  
210 biogeographic boundaries between the Atlantic and Pacific oceans, the longitudinal diversity  
211 gradient from the center of the Coral Triangle, with Southeast Polynesia being the least diverse  
212 region and Western Coral Triangle the richest, and that Gobiidae, Labridae, Pomacentridae and  
213 Apogonidae are the most diverse fish families on coral reefs (8). We found a lower proportion of  
214 rare MOTUs than expected under the neutral theory with eDNA, which is in agreement with the  
215 findings of a previous study from coral reefs in the Indo-Pacific (38), while visual census data  
216 suggests higher rarity close to that predicted from the neutral theory. More surprising, our study  
217 calls into question the pattern of fish family stability composition across the Indo-Pacific that was  
218 revealed more than 20 years ago (8), and the recent finding that planktivore families drive fish  
219 biodiversity patterns on coral reefs (11). We found significant effects of species richness and  
220 region on family composition, which appears less stable than previously thought.

221  
222 Environmental DNA identified many pelagic, deepwater and crypto-benthic species not seen by  
223 divers. Among the pelagic species identified with eDNA, many belong to the Scombridae and  
224 Carcharhinidae families, which likely avoid divers or are not permanent residents on coral reefs  
225 so can be missed in visual surveys (41). Some crypto-benthic or reef-associated species, hidden in  
226 the reef, can also be missed by divers so were also more represented in eDNA than in visual  
227 surveys. Crypto-benthic species also have a crucial role for coral reef functioning, by promoting  
228 biomass production and fueling the reef trophodynamics (42), but their diversity has been  
229 underestimated so far (13). Transient, pelagic and deep-water species may be very important for  
230 reef functioning, through pelagic larval stages or nocturnal migration up the reef slope (12,43,44),  
231 but their presence and role need further investigation. In contrast, visual census also detected many  
232 families not detected, or not identified, by eDNA, such as Acanthuridae, Blenniidae, Caesionidae,  
233 Chaenopsidae, Chaetodontidae, Labrisomidae, Labridae or Microdesmidae. This limited  
234 identification by eDNA can be due to the very low representation of these families in 12S reference  
235 databases (between 0 and 12%), or to the low resolution of the teleo marker for species of these  
236 families, so several species can share the same sequence and be grouped under the same MOTU.  
237 Environmental DNA may also be inappropriate to detect these species in the environment.

238  
239 The finding of a strong regional effect on both species composition (figure 3) and species  
240 differentiation (figure 4) at a large scale is in agreement with visual surveys and previous  
241 knowledge (45), while the suggestion of a strong turnover at the local scale may be an unexpected  
242 result for coral reef fishes. This predominant role of large-scale bioregional differentiation explains  
243 the exceptional fish diversity on coral reefs, probably associated with long-term geological  
244 isolation (2). Overall, the Tropical Northwestern Atlantic region has a very distinct MOTU

245 composition compared to the four other regions (figure 3) with only 1.2% of MOTUs being shared  
246 between the Tropical Northwestern Atlantic and any other region, while 20% of MOTUs are  
247 shared between at least two Indo-Pacific regions (figure 4a). The isolation of the Tropical  
248 Northwestern Atlantic region can be explained by the hard vicariant barrier of the Isthmus of  
249 Panama (14,46), and a limited suitable area for coral reefs during the past quaternary glaciation.  
250 By contrast, the Indo-Pacific maintained extensive coral reef refuges that have served as centers  
251 of survival during ice-age periods (9).

252 The greater local compositional dissimilarity of reef fishes among adjacent stations with eDNA  
253 than with visual census may correspond to local environmental or habitat differences, to stochastic  
254 or random processes (47), or may be due to an insufficient sampling at the station level  
255 (Supplementary Analyses Fig. 6). A higher number of replicates per station would be necessary to  
256 characterize exhaustively the diversity at the station level and more confidently conclude on the  
257 local turnover hypothesis.

258

259 While our results confirm the potential of eDNA to monitor biodiversity in marine ecosystems,  
260 some limitations should be addressed in the future to fully exploit this potential. Completing public  
261 reference databases would improve the accuracy of taxonomic assignment, which is essential for  
262 a better estimation of biodiversity patterns. At such a large spatial scale, reference databases are  
263 far from exhaustive with only up to 13% of fish species sequenced on our marker (52), preventing  
264 assignment to the species level for 81% of our eDNA sequences. Using multiple markers is an  
265 alternative to the database limitation (53,54), but it is much more expensive. For these reasons, we  
266 used MOTUs curated by a combination of a clustering algorithm and conservative abundance-  
267 based post-clustering filters. While un-curated MOTUs are prone to overestimate real diversity  
268 (55) and a given MOTU can represent several species within one cluster or several MOTUs

269 belonging to one species, MOTUs with conservative curation have been shown to reflect the true  
270 level of fish diversity across scales in streams (56,57). Additionally, some species share the same  
271 barcode sequence due to insufficient genetic differentiation on such a small mitochondrial marker  
272 (54). This lack of taxonomic resolution combined with a conservative curated MOTUs pipeline  
273 can underestimate MOTUs richness. Moreover, some crypto-benthic or rare fish families are still  
274 underrepresented in public databases, and their diversity is potentially underestimated with eDNA  
275 (i.e. Blenniidae, Gobiesocidae, Chaenopsidae, Aploactinidae).

276 Differences in sampling method and in sample size might influence the detected biodiversity with  
277 eDNA. The lower volume of water sampled in the Western Coral Triangle region (2L per sample,  
278 so 4L per station using point-sampling instead of 2-km transect with 30L elsewhere), could  
279 underestimate fish biodiversity. However, previous studies show that MOTU accumulation curves  
280 based on this dataset were close to the total fish diversity reported in this region (32). Furthermore,  
281  $\beta$ -diversity between samples within stations in each region indicates that dissimilarity between  
282 samples is not greater in the Western Coral Triangle than in other regions (figure S11). To account  
283 for differences in sample size and obtain a balanced design, we performed sensitivity analyses by  
284 rarefying our complete dataset to i) 4 stations for all sites and ii) 4 sites per region after removing  
285 the lowest sampled region (Southeast Polynesia) (Supplementary Analyses Fig. 1-4). We obtained  
286 similar patterns even after subsampling stations or sites. However, our site-based and station-based  
287 accumulation curves do not reach plateaus suggesting that our sampling effort was not sufficient  
288 to exhaustively estimate fish biodiversity for each site (Supplementary Analyses Fig. 5) and station  
289 (Supplementary Analyses Fig. 6). Twenty-five replicates (so, 12 stations in case of field  
290 duplicates) could accurately estimate biodiversity regionally due to high local turnover (58). A  
291 higher number of eDNA samples would be necessary here to reach MOTU accumulation per site  
292 and station.

293 The transport and degradation of eDNA can also impact species detection. As some evidence  
294 suggests that eDNA from pelagic fishes degrades slower than from inshore species (59), we cannot  
295 exclude that eDNA from pelagic and deep-water families (*e.g.* Myctophidae) might disperse  
296 sufficiently with sea currents such that species living close to reef habitats are detected.  
297 Environmental DNA transport could also explain the detection of some freshwater fish families  
298 (*i.e.* Centrarchidae, Osphronemidae or Channidae) in a few samples located near an estuary or in  
299 an enclosed bay with freshwater inputs.

300

301 Better understanding and anticipating the effects of multiple threats to the marine environment  
302 depends on the temporal and spatial extent of our monitoring capacity in the vast ocean.  
303 Environmental DNA is a powerful tool to investigate biodiversity patterns at large scale and  
304 monitor biodiversity, but still benefits from the combination with complementary approaches as  
305 visual methods for an exhaustive biodiversity survey across space and time to keep pace with  
306 ongoing changes.

307

## 308 **4. Methods**

### 309 **(a) Environmental DNA collection and sample processing**

310 Environmental DNA seawater samples were collected between 2017 and 2019, following a  
311 hierarchical pattern. A total of 226 eDNA samples (filters) were collected in 100 stations  
312 (gathering of replicates at the same location) located in 26 sites (groups of stations separated by at  
313 least 35 km) distributed across five tropical regions (figure S1-S2). Three different sampling  
314 methods were used comprising a 2km-long sampling transect of 30L (surface or bottom depth) or  
315 point samples of 2L (table S7 and Methods S1), and between 12 and 64 samples were collected by

316 region. Filtration was performed with Polyethersulfone (PES) filters, 0.2 µm pore size. For each  
317 sampling campaign, a strict contamination control protocol was followed in both field and  
318 laboratory stages (39). Negative field controls were performed in multiple sites, and revealed no  
319 contamination from the boat or samplers.

320

### 321 **(b) eDNA extraction, amplification and sequencing**

322 DNA extraction was performed in a dedicated DNA laboratory (SPYGEN, [www.spygen.com](http://www.spygen.com))  
323 equipped with positive air pressure, UV treatment and frequent air renewal. Decontamination  
324 procedures were conducted before and after all manipulations. Detailed protocols of DNA  
325 extraction, amplification and sequencing can be found in Method S2 and in (32,39). A teleost-  
326 specific 12S mitochondrial rRNA primer pair (teleo, forward primer -  
327 ACACCGCCCGTCACTCT, reverse primer – CTTCCGGTACTTACCATG (39)) was used  
328 for the amplification of metabarcoding sequences. As we analysed our data using MOTUs as a proxy  
329 for species to overcome genetic database limitations, we chose to amplify only one marker. Teleo  
330 marker has been shown to be the most appropriate for fish, owing to its high interspecific  
331 variability, and its short size allowing us to detect rare and degraded DNA reliably (39,54,60,61).  
332 Twelve DNA amplifications PCR per sample were performed.

333

### 334 **(c) Bioinformatic analysis**

335 Following sequencing, reads were processed using clustering and post-clustering cleaning to  
336 remove errors and estimate the number of species using Molecular Operational Taxonomic Units  
337 (MOTUs) (56). First, reads were assembled using *VSEARCH* (62), then demultiplexed and  
338 trimmed using *CUTADAPT* (63) and clustering was performed using *SWARM* v.2 (64) with a

339 minimum distance of 1 mismatch between clusters. Taxonomic assignment of MOTUs was carried  
340 out using the Lower Common Ancestor (LCA) algorithm *ecotag* implemented in the OBITOOLS  
341 toolkit (65) and the European Nucleotide Archive (ENA) as a reference database (release 143,  
342 March 2020). Details on the bioinformatics analysis can be found in Methods S3. Taxonomic  
343 assignments obtained from the LCA algorithm at the species level were accepted if the percentage  
344 of similarity with the reference sequence was 100%, at the genus level if the similarity was between  
345 90 and 99%, and at the family level if the similarity was > 85% following previous studies (32,66).  
346 If these criteria were not met, the MOTU was left unassigned. Only 21% of assigned MOTUs are  
347 assigned to the family level with a similarity between 85 and 90% (Table S8).

348

#### 349 **(d) Visual census data**

350 The visual census survey data used here is a subset (2047 transects, in 219 sites, figure S1) of the  
351 complete visual census data (3027 transects) provided by the RLS (31), and comprises all species  
352 observed on standardized 50 m surveys at sites in tropical biogeographic realms between 2006 and  
353 2017 (Methods S4) (67). We selected only the most recent survey for each station and only  
354 transects with more than five percent of coral cover. Two different sampling protocols were  
355 adapted to detect both reef and crypto-benthic fishes.

356

#### 357 **(e) Statistical analysis**

358 More details on the statistical analysis are available in Methods S5.

359 Accumulation curves were calculated for species per 500 m<sup>2</sup> transect, MOTUs per eDNA sample,  
360 and families per transect and sample. We used the functions “specaccum” and “fitspecaccum”  
361 from the R package “vegan” which calculates the expected species accumulation curve using a



362 sample-based rarefaction method and fit a nonlinear accumulation model. In order to assess the  
363 impact of the irregular sampling on the estimates measured with accumulation curves, we subset  
364 randomly half of the transects in the 3 most sampled regions in Australia, and calculated again the  
365 accumulation curves for species and families (figure S12). The results were unchanged.

366 Linear regression models were fitted between the number of MOTUs per family in the eDNA  
367 dataset and the number of species per family in the visual census dataset, after  $\log(x+1)$   
368 transformation (figure 1e).

369 Accumulation curves were also calculated by sub-setting MOTUs belonging to crypto-benthic  
370 orders, or to pelagic families, for both datasets (figure 2). The asymptote was calculated as  
371 described above.

372 We performed distance-based Redundancy Analysis (dbRDA) on family proportions, with *region*  
373 and *site richness* as explanatory variables, using the function *capscale* from the *vegan* package.  
374 We subset the Visual Census to select only the 68 sites that fell into the 5 regions in common with  
375 the eDNA dataset. Total dbRDA provided the effects of each of the variables and their interaction.  
376 We then calculated partial dbRDA to measure the effect of the Region while correcting for the  
377 effect of site richness (figure 3, table S3).

378 We applied an additive partitioning framework (68) to separate the total MOTUs diversity at the  
379 global scale ( $\gamma$  global) into contributions at smaller scales from regions to local richness :  $\gamma_{global} =$   
380  $\beta_{inter-region} + \text{mean } \beta_{inter-site} + \text{mean } \beta_{inter-station} + \text{mean } \bar{\alpha}_{station}$ . In this additive framework, the three  
381 levels of biodiversity (69) (i.e.  $\alpha$ ,  $\beta$  and  $\gamma$ ) are expressed with the same unit and consequently the  
382 contribution of  $\alpha$  and  $\beta$  diversity to total diversity ( $\gamma$ ) can be directly compared (70).

383 We analyzed the distribution of fish MOTU and species occurrences using global species  
384 abundance distribution (gSAD) which plots, on a log-log scale, the number of species as a function  
385 of the number of observations (37).

386

387

## 388 **Funding**

389 The sampling in the Caribbean, Indian Ocean and Polynesia and the sequencing were funded by  
390 Monaco Explorations. Fieldwork in Indonesia and laboratory activities were supported by the  
391 Lengguru 2017 Project ([www.lengguru.org](http://www.lengguru.org)), conducted by the French National Research Institute  
392 for Sustainable Development (IRD), National Research and Innovation Agency (BRIN)  
393 with the Research Center for Oceanography (RCO), the Politeknik Kelautan dan Perikanan  
394 Sorong), the University of Papua (UNIPA) with the help of the Institut Français in Indonesia (IFI),  
395 funding from Monaco Explorations, and corporate sponsorship from the Total Foundation and  
396 TIPCO company. Fieldwork and laboratory activities in New-Caledonia were supported by the  
397 projects ANR SEAMOUNTS and CIFRE REEF 3.0 conducted by the French National Institute  
398 for Sustainable Development (IRD) and GINGER-BURGEAP-SOPRONER company with  
399 funding from Monaco Explorations. Fieldwork and laboratory activities in Colombia were  
400 supported by Monaco Explorations, ETH Global grant and the project Reefish, conducted in  
401 collaboration with the Instituto de Investigaciones Marinas y Costeras – INVEMAR. Monaco  
402 Explorations supported also sampling and sequencing in the Caribbean. Fieldwork in the French  
403 Scattered islands was supported by the Terres Australes et Antartiques Françaises (TAAF).

404

## 405 **Acknowledgements**

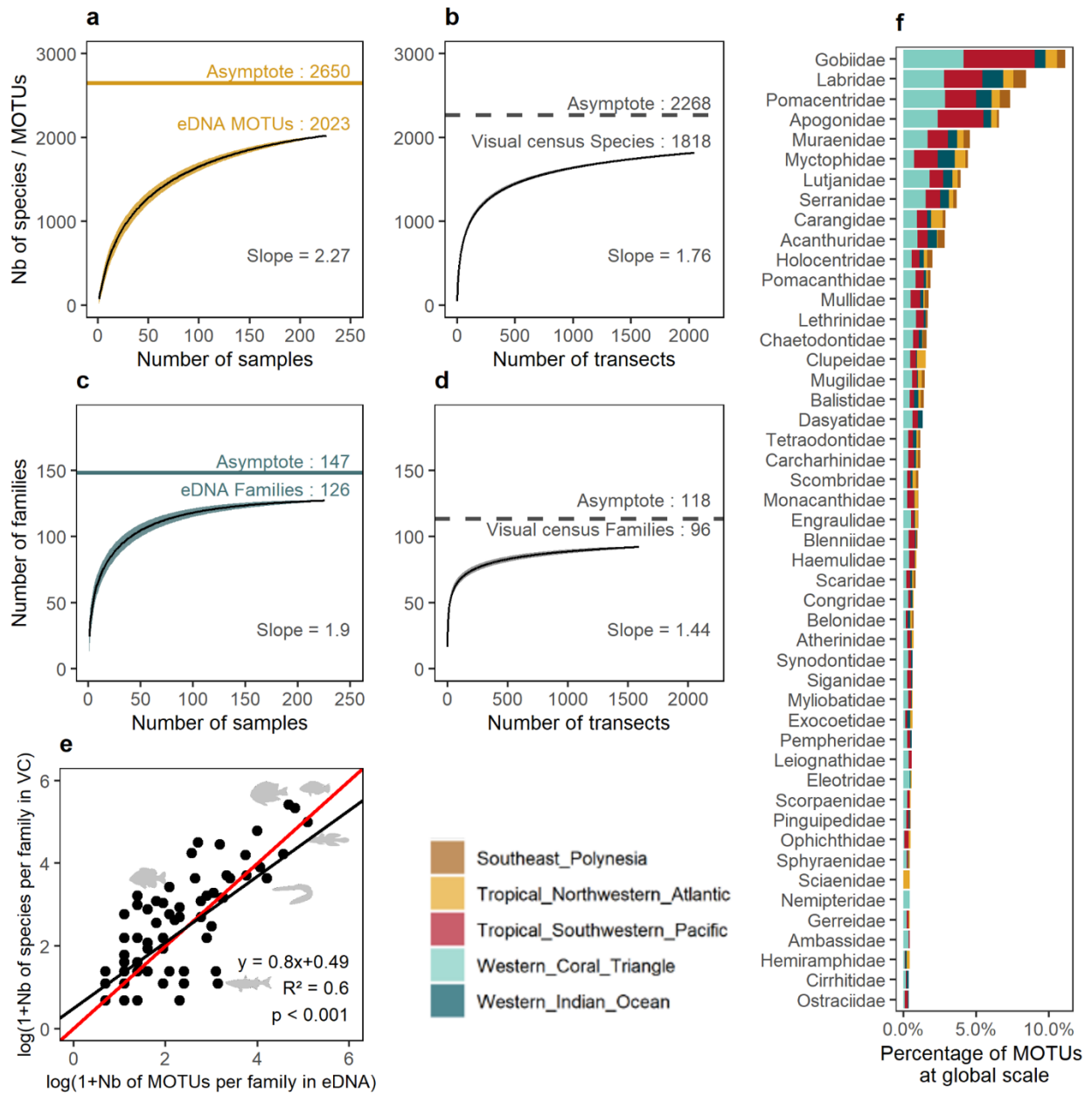
406 The authors thank all staff and students involved in fieldwork and acknowledge SPYGEN staff for  
407 the technical support in the eDNA laboratory.

408

- 410 1. Parravicini V, Kulbicki M, Bellwood DR, Friedlander AM, Arias-Gonzalez JE, Chabanet P, et al. Global  
411 patterns and predictors of tropical reef fish species richness. *Ecography (Cop)*. 2013;36(12):1254–62.
- 412 2. Cowman PF, Bellwood DR. The historical biogeography of coral reef fishes: Global patterns of origination  
413 and dispersal. *J Biogeogr*. 2013;40(2):209–24.
- 414 3. Cinner JE, Zamborain-mason J, Gurney GG, Graham NAJ, Macneil MA, Hoey AS, et al. Meeting fisheries,  
415 ecosystem function, and biodiversity goals in a human-dominated world. *Science (80- )*.  
416 2020;368(April):307–11.
- 417 4. Hoegh-Guldberg O, Jacob D, Taylor M, Guillén Bolaños T, Bindi M, Brown S, et al. The human imperative  
418 of stabilizing global climate change at 1.5°C. *Science (80- )*. 2019;365(6459).
- 419 5. Victor BC. How many coral reef fish species are there? Cryptic diversity and the new molecular taxonomy.  
420 In: *Ecology of fishes on coral reefs* Cambridge University Press, Cambridge. 2015. p. 76–88.
- 421 6. Siqueira AC, Morais RA, Bellwood DR, Cowman PF. Trophic innovations fuel reef fish diversification. *Nat*  
422 *Commun*. 2020;11(1):1–11.
- 423 7. Bellwood D, Wainwright P. The history and biogeography of fishes on coral reefs. In: Sale P, editor. *Coral*  
424 *Reef Fishes: Dynamics and Diversity in a Complex Ecosystem*. Academic P. San Diego; 2002.
- 425 8. Bellwood DR, Hughes TP. Regional-scale assembly rules and biodiversity of coral reefs. *Science (80- )*.  
426 2001;292(5521):1532–4.
- 427 9. Pellissier L, Leprieur F, Parravicini V, Cowman PF, Kulbicki M, Litsios G, et al. Quaternary coral reef  
428 refugia preserved fish diversity. *Science (80- )*. 2014;344(6187):1016–20.
- 429 10. Veron J, Devantier LM, Turak E, Green AL, Kininmonth S, Stafford-Smith M, et al. Delineating the Coral  
430 Triangle. *Galaxea, J Coral Reef Stud*. 2009;11(2):91–100.
- 431 11. Siqueira AC, Morais RA, Bellwood DR, Cowman PF. Planktivores as trophic drivers of global coral reef  
432 fish diversity patterns. *PNAS*. 2021;118(9).
- 433 12. Morais RA, Bellwood DR. Pelagic Subsidies Underpin Fish Productivity on a Degraded Coral Reef. *Curr*  
434 *Biol*. 2019;29(9):1521–7.
- 435 13. Brandl SJ, Goatley CHR, Bellwood DR, Tornabene L. The hidden half: ecology and evolution of  
436 cryptobenthic fishes on coral reefs. *Biol Rev*. 2018;93:1846–73.
- 437 14. Bender MG, Leprieur F, Mouillot D, Kulbicki M, Parravicini V, Pie MR, et al. Isolation drives taxonomic  
438 and functional nestedness in tropical reef fish faunas. *Ecography (Cop)*. 2017;40(3):425–35.
- 439 15. Hughes TP, Bellwood DR, Connolly SR, Cornell H V., Karlson RH. Double jeopardy and global extinction  
440 risk in corals and reef fishes. *Curr Biol*. 2014;24(24):2946–51.
- 441 16. Mouillot D, Villéger S, Parravicini V, Kulbicki M, Arias-gonzález JE. Functional over-redundancy and high  
442 functional vulnerability in global fish faunas on tropical reefs. *PNAS*. 2014;111(38):13757–62.
- 443 17. Alzate A, Zapata FA, Giraldo A. A comparison of visual and collection-based methods for assessing  
444 community structure of coral reef fishes in the Tropical Eastern Pacific. *Rev Biol Trop*.  
445 2014;62(February):359–71.
- 446 18. Bellwood D, Renema W, Rosen BB. Biodiversity hotspots, evolution and coral reef biogeography: a review.  
447 In: *Biotic Evolution and Environmental Change in Southeast Asia* Cambridge: Cambridge University Press.  
448 2012. p. 216–45.
- 449 19. Mora C. Ecology of fishes on coral reefs. *Ecology of Fishes on Coral Reefs*. 2015. 1–374 p.
- 450 20. Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH. Environmental DNA. *Mol Ecol*. 2012;21(8):1789–93.
- 451 21. Boulanger E, Loiseau N, Valentini A, Arnal V, Boissery P, Dejean T, et al. Environmental DNA  
452 metabarcoding reveals and unpacks a biodiversity conservation paradox in Mediterranean marine reserves,  
453 Dryad, Dataset. *Proc R Soc B*. 2021;288(20210112).
- 454 22. Harrison JB, Sunday JM, Rogers SM. Predicting the fate of eDNA in the environment and implications for  
455 studying biodiversity. *Proc R Soc B Biol Sci*. 2019;286(1915):1–9.
- 456 23. Cordier T, Alonso-Saez L, Apothéloz-Perret-Gentil L, Aylagas E, Bohan DA, Bouchez A, et al. Ecosystem  
457 monitoring powered by environmental genomics: A review of current strategies with an implementation  
458 roadmap. *Mol Biol Evol*. 2020;may:1–22.
- 459 24. De Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, et al. Eukaryotic plankton diversity in the  
460 sunlit ocean. *Science (80- )*. 2015;348(6237):1–11.
- 461 25. Valdivia-Carrillo T, Rocha-Olivares A, Reyes-Bonilla H, Domínguez-Contreras JF, Munguia-Vega A.  
462 Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish  
463 communities in a marine biodiversity hotspot. *Mol Ecol Resour*. 2021;(March):1558–74.

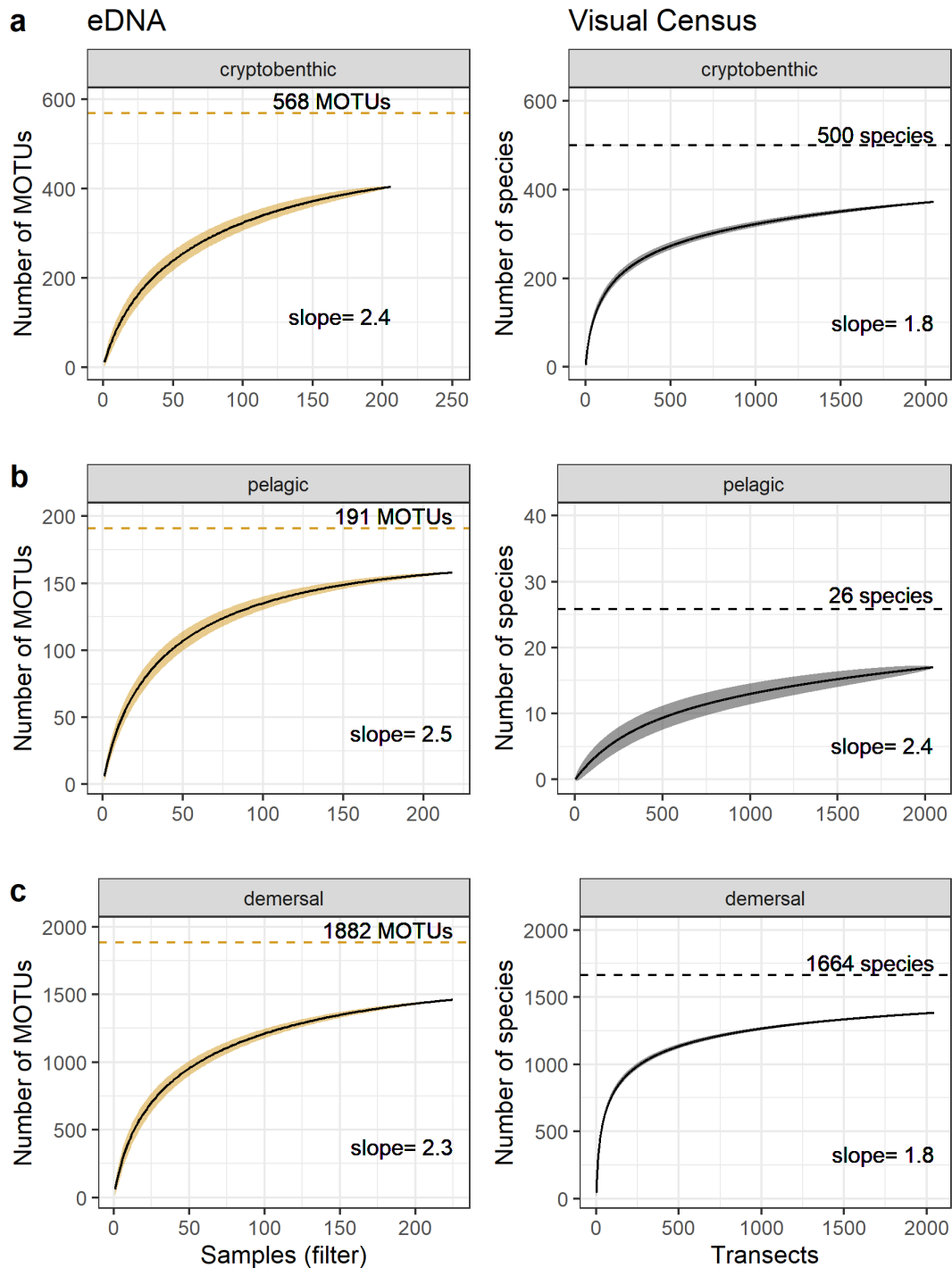
- 464 26. West K, Travers MJ, Stat M, Harvey ES, Richards ZT, DiBattista JD, Newman SJ, Harry A, Skepper CL,  
465 Heydenrych M, Bunce M. Large-scale eDNA metabarcoding survey reveals marine biogeographic break and  
466 transitions over tropical north-western Australia. *Diversity and Distributions*. 2021 Oct;27(10):1942-57.  
467 27. Kume M, Lavergne E, Ahn H, Terashima Y, Kadowaki K, Ye F, et al. Factors structuring estuarine and  
468 coastal fish communities across Japan using environmental DNA metabarcoding. *Ecol Indic* [Internet].  
469 2021;121(November 2020):107216. Available from: <https://doi.org/10.1016/j.ecolind.2020.107216>  
470 28. Fraija-Fernández N, Bouquieaux MC, Rey A, Mendibil I, Cotano U, Irigoien X, et al. Marine water  
471 environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial  
472 patterns in a large oceanic area. *Ecol Evol*. 2020;10(14):7560–84.  
473 29. Aglieri G, Baillie C, Mariani S, Cattano C, Calò A, Turco G, et al. Environmental DNA effectively captures  
474 functional diversity of coastal fish communities. *Mol Ecol*. 2020;(August):1–13.  
475 30. DiBattista JD, Berumen ML, Priest MA, De Brauwer M, Coker DJ, Sinclair-Taylor TH, et al.  
476 Environmental DNA reveals a multi-taxa biogeographic break across the Arabian Sea and Sea of Oman.  
477 *Environ DNA*. 2021;(December 2020):1–16.  
478 31. Edgar GJ, Stuart-Smith RD. Systematic global assessment of reef fish communities by the Reef Life Survey  
479 program. *Sci Data*. 2014;1:1–8.  
480 32. Juhel JB, Utama RS, Marques V, Vimono IB, Sugeha HY, Kadarusman, et al. Accumulation curves of  
481 environmental DNA sequences predict coastal fish diversity in the coral triangle. *Proceedings Biol Sci*.  
482 2020;287(1930):1–10.  
483 33. Castellanos-Galindo GA, Krumme U, Rubio EA, Saint-Paul U. Spatial variability of mangrove fish  
484 assemblage composition in the tropical eastern Pacific Ocean. *Rev Fish Biol Fish*. 2013;23(1):69–86.  
485 34. Willis TJ, Anderson MJ. Structure of cryptic reef fish assemblages: Relationships with habitat  
486 characteristics and predator density. *Mar Ecol Prog Ser*. 2003;257:209–21.  
487 35. Harnik PG, Simpson C, Payne JL. Long-term differences in extinction risk among the seven forms of rarity.  
488 *Proc R Soc B Biol Sci*. 2012;279(1749):4969–76.  
489 36. Crist TO, Veech JA. Additive partitioning of rarefaction curves and species-area relationships: Unifying  $\alpha$ -,  
490  $\beta$ - and  $\gamma$ -diversity with sample size and habitat area. *Ecol Lett*. 2006;9(8):923–32.  
491 37. Enquist BJ, Feng X, Boyle B, Maitner B, Newman EA, Jørgensen PM, et al. The commonness of rarity:  
492 Global and future distribution of rarity across land plants. *Sci Adv*. 2019;5(11):1–14.  
493 38. Dornelas M, Connolly SR, Hughes TP. Coral reef diversity refutes the neutral theory of biodiversity. *Nature*.  
494 2006;440(7080):80–2.  
495 39. Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, et al. Next-generation monitoring of  
496 aquatic biodiversity using environmental DNA metabarcoding. *Mol Ecol*. 2016;25(4):929–42.  
497 40. West K, Travers MJ, Stat M, Harvey ES, Richards ZT, Dibattista JD, et al. Large-scale eDNA  
498 metabarcoding survey reveals marine biogeographic break and transitions over tropical north- western  
499 Australia. *Divers Distrib*. 2021;(00):1–16.  
500 41. Boussarie G, Kiszka JJ, Mouillot D, Bonnin L, Manel S, Kulbicki M, et al. Environmental DNA illuminates  
501 the dark diversity of sharks. *Sci Adv*. 2018;4(5):1–8.  
502 42. Brandl SJ, Tornabene L, Goatley CHR, Casey JM, Morais RA, Côté IM, et al. Demographic dynamics of  
503 the smallest marine vertebrates fuel coral-reef ecosystem functioning. *Science* (80- ). 2019;(May):799–802.  
504 43. Kimmerling N, Zuqert O, Amitai G, Gurevich T, Armoza-Zvuloni R, Kolesnikov I, et al. Quantitative  
505 species-level ecology of reef fish larvae via metabarcoding. *Nat Ecol Evol*. 2018;2(2):306–16.  
506 44. Beckley LE, Holliday D, Sutton AL, Weller E, Olivar MP, Thompson PA. Structuring of larval fish  
507 assemblages along a coastal-oceanic gradient in the macro-tidal, tropical Eastern Indian Ocean. *Deep Res*  
508 *Part II*. 2019;161(March 2018):105–19.  
509 45. McLean M, Stuart-Smith RD, Villéger S, Auber A, Edgar GJ, MacNeil MA, et al. Trait similarity in reef  
510 fish faunas across the world's oceans. *Proc Natl Acad Sci*. 2021;118(12):e2012318118.  
511 46. Gaboriau T, Leprieur F, Mouillot D, Hubert N. Influence of the geography of speciation on current patterns  
512 of coral reef fish biodiversity across the Indo-Pacific. *Ecography* (Cop). 2018;41(8):1295–306.  
513 47. Ahmadi GN, Tornabene L, Smith DJ, Pezold FL. The relative importance of regional, local, and  
514 evolutionary factors structuring cryptobenthic coral-reef assemblages. *Coral Reefs*. 2018;37(1):279–93.  
515 48. Coker DJ, DiBattista JD, Sinclair-Taylor TH, Berumen ML. Spatial patterns of cryptobenthic coral-reef  
516 fishes in the Red Sea. *Coral Reefs*. 2018;37(1):193–9.  
517 49. Goatley CHR, González-Cabello A, Bellwood DR. Reef-scale partitioning of cryptobenthic fish  
518 assemblages across the Great Barrier Reef, Australia. *Mar Ecol Prog Ser*. 2016;544(February):271–80.  
519 50. Wang S, Loreau M. Biodiversity and ecosystem stability across scales in metacommunities. *Ecol Lett*.  
520 2016;19:510–8.

- 521 51. Benkwitt C, Wilson S, Graham NAJ. Biodiversity increases ecosystem functions despite multiple stressors  
522 on coral reefs. *Nat Ecol Evol.* 2020;4:916–26.
- 523 52. Marques V, Milhau T, Albouy C, Dejean T, Manel S, Mouillot D, et al. GAPeDNA : Assessing and  
524 mapping global species gaps in genetic databases for eDNA metabarcoding. 2020;(April):1–13.
- 525 53. Ruppert KM, Kline RJ, Rahman MS. Past, present, and future perspectives of environmental DNA (eDNA)  
526 metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob Ecol*  
527 *Conserv.* 2019;17:e00547.
- 528 54. Polanco A, Richards FE, Flück B, Valentini A, Altermatt F, Jean- SB, et al. Comparing the performance of  
529 12S mitochondrial primers for fish environmental DNA across ecosystems. *Environ DNA.* 2021;(June):1–  
530 15.
- 531 55. Brandt MI, Trouche B, Quintric L, Günther B, Wincker P, Poulain J, et al. Bioinformatic pipelines  
532 combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic  
533 metabarcoding. *Mol Ecol Resour.* 2021;21(6):1904–21.
- 534 56. Marques V, Guérin PÉ, Rocle M, Valentini A, Manel S, Mouillot D, et al. Blind assessment of vertebrate  
535 taxonomic diversity across spatial scales by clustering environmental DNA metabarcoding sequences.  
536 *Ecography (Cop).* 2020;43:1–12.
- 537 57. Sales NG, Wangensteen OS, Carvalho DC, Deiner K, Praebel I, McDevitt A, et al. Space-time dynamics in  
538 monitoring neotropical fish communities using eDNA metabarcoding. *bioRxiv.* 2020;1–38.
- 539 58. Stauffer S, Jucker M, Keggin T, Marques V, Andrello M, Bessudo S, et al. How many replicates to  
540 accurately estimate fish biodiversity using environmental DNA on coral reefs ? Authors : *bioRxiv.* 2021;
- 541 59. Collins RA, Wangensteen OS, O’Gorman EJ, Mariani S, Sims DW, Genner MJ. Persistence of  
542 environmental DNA in marine systems. *Commun Biol.* 2018;1(185):1–12.
- 543 60. Collins RA, Bakker J, Wangensteen OS, Soto AZ, Corrigan L, Sims DW, et al. Non-specific amplification  
544 compromises environmental DNA metabarcoding with COI. *Methods Ecol Evol.* 2019;10(11):1985–2001.
- 545 61. Zhang S, Zhao J, Yao M. A comprehensive and comparative evaluation of primers for metabarcoding eDNA  
546 from fish. *Methods Ecol Evol.* 2020;2020(January):1609–25.
- 547 62. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for  
548 metagenomics. *PeerJ.* 2016;4:1–22.
- 549 63. Martin. M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.*  
550 1994;17(1):10–2.
- 551 64. Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm v2: highly-scalable and high-resolution  
552 amplicon clustering. *PeerJ.* 2015;3:1–12.
- 553 65. Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E. obitools: A unix-inspired software package  
554 for DNA metabarcoding. *Mol Ecol Resour.* 2016;16(1):176–82.
- 555 66. Polanco Fernández A, Marques V, Fopp F, Juhel J, Borrero-Pérez GH, Cheutin M, et al. Comparing  
556 environmental DNA metabarcoding and underwater visual census to monitor tropical reef fishes. *Environ*  
557 *DNA.* 2020 Oct 2;1–15.
- 558 67. Spalding MD, Fox HE, Allen GR, Davidson N, Ferdaña ZA, Finlayson M, et al. Marine Ecoregions of the  
559 World: A Bioregionalization of Coastal and Shelf Areas. *Bioscience.* 2007;57(7):573–83.
- 560 68. Escalas A, Troussellier M, Yuan T, Bouvier T, Bouvier C, Mouchet MA, et al. Functional diversity and  
561 redundancy across fish gut, sediment and water bacterial communities. *Environ Microbiol.*  
562 2017;19(8):3268–82.
- 563 69. Whittaker RH. Evolution and measurement of species diversity. *Taxon.* 1972;21:213–51.
- 564 70. Veech JA, Summerville KS, Crist TO, Gering JC. The additive partitioning of species diversity: Recent  
565 revival of an old idea. *Oikos.* 2002;99(1):3–9.
- 566
- 567



569

570 **Figure 1. Estimates of overall fish richness from environmental DNA (eDNA) and visual**  
 571 **census.** (a) accumulation curve of molecular operational taxonomic units from eDNA (eDNA  
 572 MOTUs), (b) accumulation curve of species from the visual census database, (c) accumulation  
 573 curve of eDNA families, (d) accumulation curve of visual census families. For (a-d), Species  
 574 accumulation model is fitted according to Lomolino method (see methods). (e) linear regression  
 575 (black line) between the number of species per family in visual census data and the number of  
 576 MOTUs per family in eDNA (log(x+1) transformation) over  $n = 77$  families. Each point is a  
 577 family. Red line is  $x=y$ . (f) percentage of MOTUs assigned to each family at global scale, and  
 578 proportion in each region.

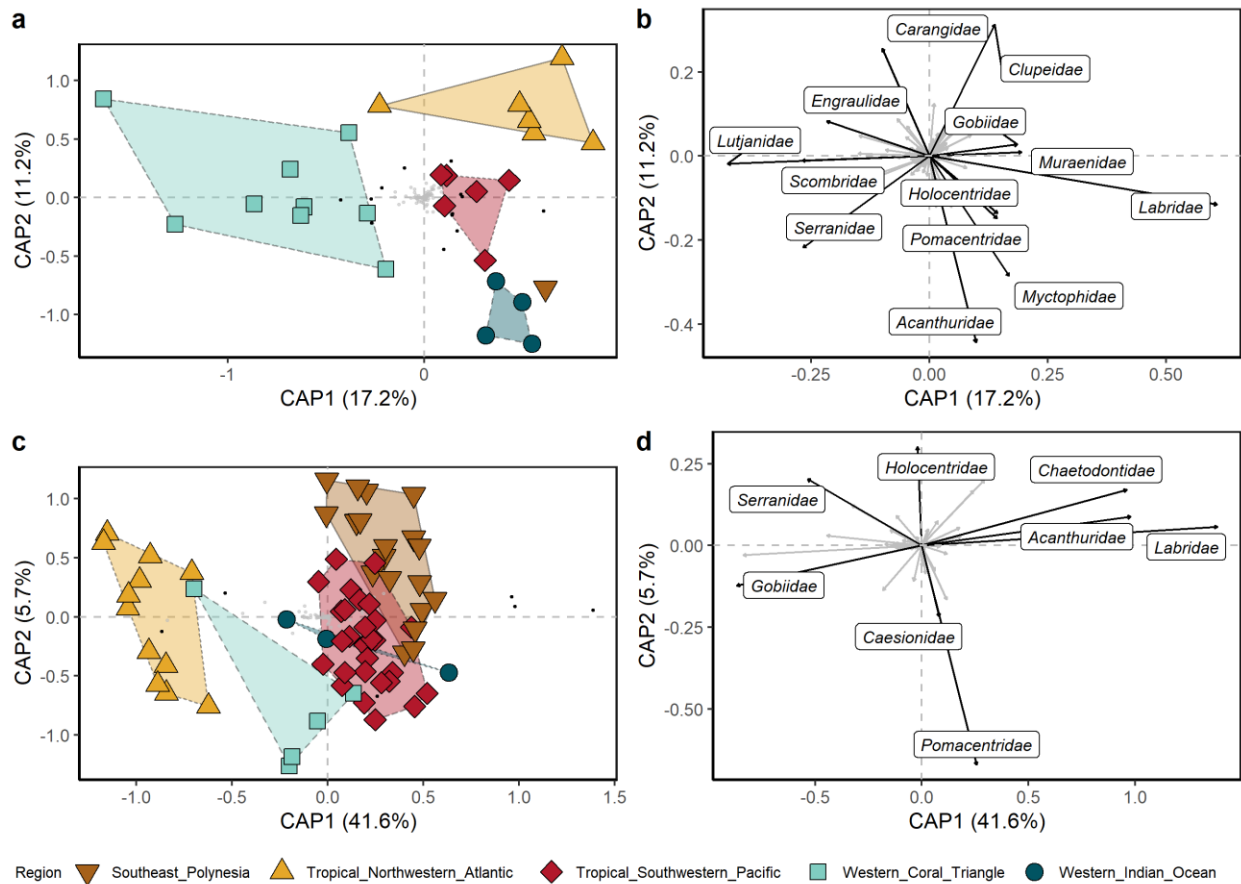


579

580 **Figure 2. Estimates of overall fish richness from eDNA and visual census across habitat**  
 581 **categories.** (a) accumulation curve of crypto-benthic eDNA MOTUs (left) and visual census  
 582 species (right), (b) accumulation curve of pelagic MOTUs (left) and visual census species (right),  
 583 c, accumulation curve of demersal MOTUs (left) and visual census species (right). Accumulation  
 584 model is fitted with a nonlinear Lomolino model (see Methods).

585

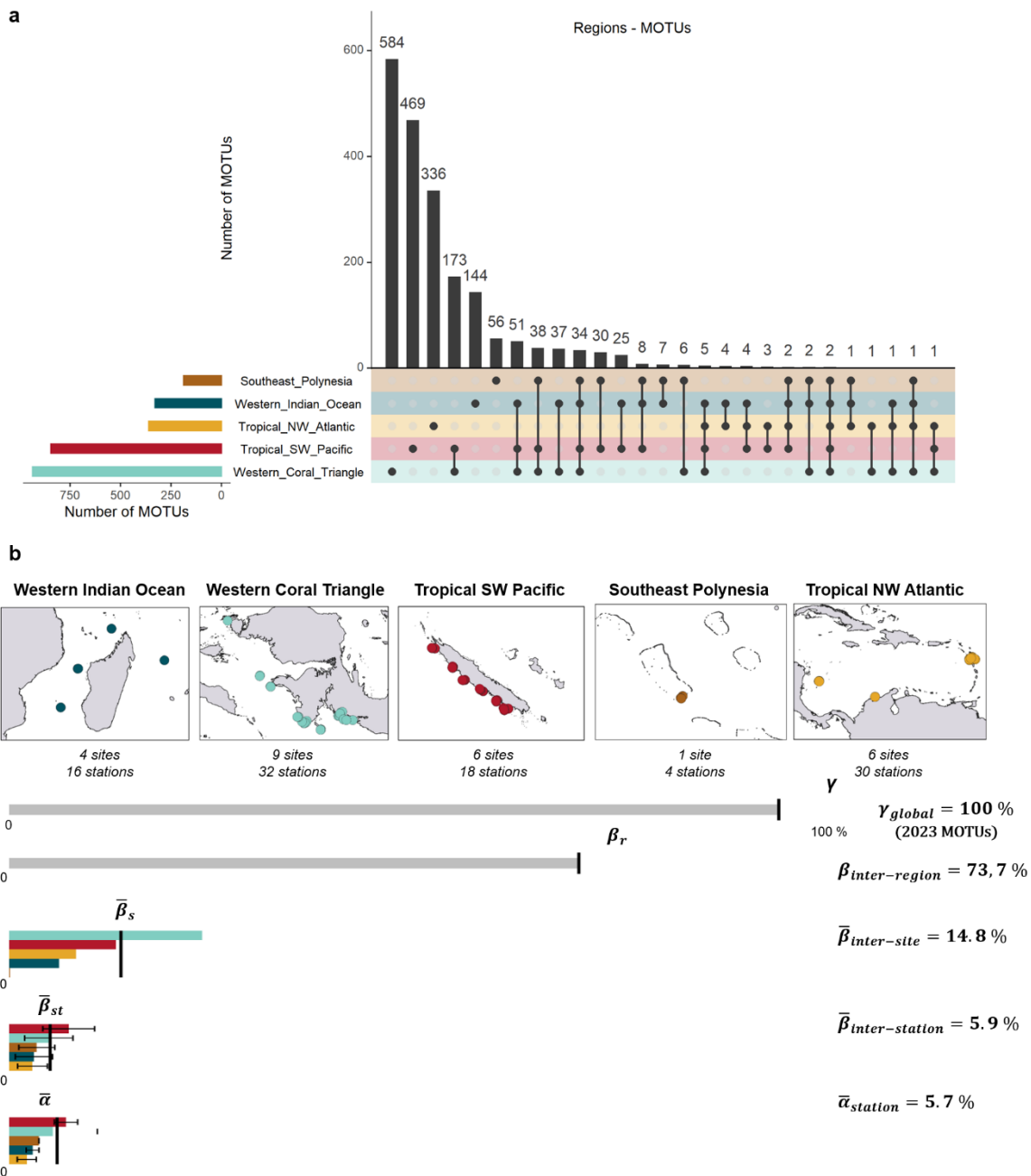




586

587 **Figure 3. Partial Distance-based Redundancy analysis of MOTU proportions of each family**  
 588 **in each site.** (a) dbRDA on eDNA dataset, with 133 families in 26 sites ( $R^2=0.21$ ,  $F=3.11$ ,  
 589  $p=0.001$ ), (b) families with scores > 95% of scores distribution on each axis for eDNA, (c) dbRDA  
 590 on a subset of Visual Census dataset to select only the sites in the same regions as in the eDNA  
 591 dataset, with 76 families in 68 sites ( $R^2=0.5$ ,  $F=15.8$ ,  $p=0.001$ ), (d) families with scores > 95% of  
 592 scores distribution on each axis for Visual Census. Axis labels indicate the percentage of variance  
 593 explained by the 2 first dbRDA dimensions (CAP1 and CAP2).

594



595

596 **Figure 4. Hierarchical partitioning of MOTU occurrences across spatial scales.** (a) Number  
 597 of MOTUs found in only one region, or shared between 2, 3, 4 or all 5 regions. Histograms indicate  
 598 the number of MOTUs present in all the regions identified by the dots in the lower part. (b) Global  
 599 fish diversity ( $\gamma_{global}$ ) is partitioned into  $\beta_{inter-region}$  + mean  $\beta_{inter-site}$  + mean  $\beta_{inter-station}$  + mean  $\bar{\alpha}_{station}$ .  
 600 Mean values at global scales are indicated with the black vertical segments. For  $\beta_{inter-site}$ ,  $\beta_{inter-station}$   
 601 and  $\bar{\alpha}_{station}$ , mean values are given for each region (colored bars) with the standard errors.  $\beta_{inter-region}$   
 602 contributes the highest to gamma global (73.7%).  
 603