Environmental DNA metabarcoding reveals and unpacks a biodiversity conservation paradox in Mediterranean marine reserves

Emilie Boulanger1,2, Nicolas Loiseau2, Alice Valentini3, Véronique Arnal1, Pierre Boissery4, Tony Dejean5, Julie Deter5, Nacim Guellati2, Florian Holon2, Jean-Baptiste Juhel2, Philippe Lenfant6, Stéphanie Manel1 and David Mouillot2,7

1CEFE, Univ Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France
2MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Montpellier, France
3SPYGEN, Le Bourget-du-Lac, France
4Agence de l’Eau Rhône-Méditerranée-Corse, Délégation de Marseille, Marseille, France
5Andromède Océanologie, Carnon, France
6CEFREM, CNRS, UPVD, Université de Perpignan, Perpignan, France
7Institut Universitaire de France, Paris, France

Although we are currently experiencing worldwide biodiversity loss, local species richness does not always decline under anthropogenic pressure. This conservation paradox may also apply in protected areas but has not yet received conclusive evidence in marine ecosystems. Here, we survey fish assemblages in six Mediterranean no-take reserves and their adjacent fishing grounds using environmental DNA (eDNA) while controlling for environmental conditions. We detect less fish species in marine reserves than in nearby fished areas. The paradoxical gradient in species richness is accompanied by a marked change in fish species composition under different managements. This dissimilarity is mainly driven by species that are often overlooked by classical visual surveys but detected with eDNA: cryptobenthic, pelagic, and rare fishes. These results do not negate the importance of reserves in protecting biodiversity but shed new light on how under-represented species groups can positively react to fishing pressure and how conservation efforts can shape regional biodiversity patterns.

1. Introduction

Marine ecosystems and their resources are severely threatened by multiple pressures including climate change [1], over-exploitation [2], and habitat degradation [3]. However, despite the prevailing trend of biodiversity loss at the global scale [4,5], the number of species does not necessarily decline at the local scale [6,7]. Long-term time series show that only 3% of coastal marine ecosystems are experiencing a local decline in species richness while a positive trend was reported in 16% of the studied cases [8]. This biodiversity conservation paradox, i.e. species richness can increase under disturbance, can be explained by a balance between extinction and colonization rates for a given location and a high species turnover [7,9]. Such turnover can occur when endemic species are replaced by exotic species [10] or by the range expansion of species from adjacent regions under climate change [8]. Biodiversity can also increase in disturbed areas if the removal of certain vulnerable species allows the establishment and coexistence of more resistant species under the intermediate disturbance hypothesis [11].

This conservation paradox has received little attention in the spatial context of protected areas. Protected areas are expanding worldwide following the new commitment to protect at least 30% of the global ocean and land by 2030, to achieve both biodiversity and climate goals [12–14]. Since some
conservation-dependent species can be rapidly extirpated by human activities outside protected areas [15,16], we can expect more species within protected areas than their non-protected counterparts. Yet, this assumption is supported by scarce evidence [17,18], while other studies fail to show any marked difference in species richness as a result of protection [19–23] or even report higher local species richness in human-modified natural habitats [24]. Here, we suggest that this lack of consensus may come from incomplete species detection and uncontrolled habitat or environmental covariates, at least in coastal marine ecosystems.

Marine protected areas (MPAs), and in particular marine reserves which are strictly no-take MPAs [25], offer a unique opportunity to test this conservation paradox and some underlying hypotheses. Marine reserves are widely recognized as effective conservation tools supporting greater density and biomass of exploited species within their boundaries than nearby fished areas [21,26–28]. Comparatively, the extent to which marine reserves and nearby fished areas support different levels of species richness or different species compositions remains unclear. On one side, large-bodied and predator species are often overexploited by fisheries and extirpated outside marine reserves increasing species richness within reserves [15]. On the other side, marine reserves can restore predator populations and thus strengthen ‘top-down’ trophic cascades thereby affecting biodiversity at lower trophic levels [29]. Elucidating a potential marine conservation paradox would thus require the detection of a broad range of fish species constituting assemblages from large predators to small prey. Yet, many fish species are missed by most capture- or visual-based surveys because they are cryptobenthic, rare, or elusive [30–34]. Moreover, mobile species may not be recorded as they only occur for short amounts of time in a given location [30].

As an alternative, the environmental DNA (eDNA) metabarcoding approach overcomes some shortcomings of classical surveys to characterize marine fish assemblages [35,36] including small, cryptic, and elusive species [33,37]. eDNA is made of small fragments of intra- and extracellular DNA generated by organisms in their proximate environment, and can be sampled to infer the presence of species, monitor coastal ecosystems, and unveil ecological processes [33,38–40]. Yet, the gaps in public genetic databases can limit the breadth of species inventories based on eDNA metabarcoding [41]. In this study, we take advantage of eDNA detectability potential, using a regionally augmented genetic reference database, to uncover a previously hidden conservation paradox in coastal fish assemblages: marine reserves host less species than nearby fished areas after controlling for environmental conditions. We also show that changes in fish species composition along a gradient of human impact are mainly driven by species groups typically overlooked in most MPA studies, namely cryptobenthic, pelagic, and rare fishes.

2. Results

A stratified sampling design was carried out to survey six Mediterranean no-take reserves and their surrounding sites, hereafter referred to as regions (see Methods, figure 1a; electronic supplementary material, table S1). Three sites were considered for each of the six regions: one within the reserve boundaries, one outside at 5 km from the reserve boundaries, and one outside at 10 km, hereafter referred to as protection levels (figure 1b). eDNA was filtered along 2 km transects using a protocol optimized for monitoring coastal species with four replicates per site. We assembled a new reference genetic database (115 species sequenced to reach 75% coverage) for North-Western Mediterranean coastal fishes to assign more eDNA sequences to known species using a stringent bioinformatic pipeline.

(a) Metabarcoding and taxonomic assignment

The 72 eDNA samples (6 regions × 3 sites × 4 replicates per site) yielded a total of 51 506 234 reads, with on average 715 364 reads per sample (±s.d. = 293 934). Assigning reads to the reference database detected 122 unique fish taxa with an average 35 taxa per sample (±10), among which 104 were identified at species level whereas 16 could only be assigned to the genus level and two to the family level. After removal of foreign species, uncertain assignments, and freshwater fishes (electronic supplementary material, Methods), 46 034 170 reads were assigned to a known marine fish species. The mean number of reads per sample dropped to 639 364 (±293 380), and read abundances were transformed to presence/absence for all subsequent analyses.

A total of 97 fish species were identified across all samples covering 74 genera and 43 families, with on average 30 fish species (±9) per sample (figure 2). Almost half of the species detected belong to four families, i.e. sparids (Sparidae, n = 16), gobies (Gobiidae, n = 12), wrasses (Labridae, n = 7), and combtooth blennies (Blenniidae, n = 7), whereas 30 families were represented by only one species (electronic supplementary material, figure S2). Detections included common Mediterranean taxa such as the damselfish (Chromis chromis), the salma porgy (Sarpa salpa), and the white seabream (Diplodus sargus), as well as rare species like the grey triggerfish (Balistes capriscus) and the blue shark (Prionace glauca).

(b) Species richness paradox

Overall, fish species richness seemingly increased outside the reserve, with on average, 28 (±6) fish species within the reserve, 30 (±10) species at 5 km outside, and 32 (±11) species at 10 km outside, albeit no significant difference was detected (Kruskal–Wallis \( \chi^2 = 5.1, d.f. = 2, p = 0.078 \)) (figure 2a).

To consider the confounding effects of environmental conditions (habitat and climate) on the protection level effect, we created 500 m buffer zones around each transect and extracted the coverage of each substrate type (electronic supplementary material, Methods), mean bottom depth, and the mean benthic and surface chlorophyll a. We recorded the sea surface temperature (SST) during sampling and calculated the mean distance of each transect to the closest point on land. To avoid collinearity between all these covariates, we performed a principal component analysis (PCA). The first four orthogonal PCA axes explained 74.2% of the total variance among sites and were retained as explanatory variables in the next analyses to control for environmental confounding factors (electronic supplementary material, figure S3).

We used generalized linear models (GLMs) to investigate the effect of protection on species richness while accounting for environmental differences represented by the PCA axes. The influence of protection on fish richness was highly significant (GLM, \( R^2 = 0.40, p < 0.01; \) figure 2b; electronic supplementary material, table S2). We detected no effect of
the region on model residuals (Kruskal–Wallis \( \chi^2 = 5.76, \) d.f. = 5, \( p = 0.33 \)). The model revealed a significant 45% increase in overall species richness at 10 km outside the reserve compared to inside (\( p < 0.01 \); electronic supplementary material, table S3).

(c) Species dissimilarity between assemblages

We estimated species dissimilarity or \( \beta \)-diversity between fish assemblages using the Jaccard distance. Two independent patterns may occur when measuring \( \beta \)-diversity: turnover and nestedness [42,43]. Turnover occurs when species present at one site are absent at another site but are replaced by other species absent from the first. Nestedness occurs when species present at one site are absent at another but are not replaced by new species.

Species dissimilarity between protection levels was high with an average \( \beta \)-diversity of 58.8% between the reserve and 5 km outside (figure 2c), 57.6% between the reserve and 10 km outside (figure 2d), and 57.5% between 5 km and 10 km outside (figure 2e). On average, 42.3% (±15.5%) of fish species were replaced between sites under different protection levels. This turnover represented 74% (±22.4%) of the pairwise dissimilarities, whereas nestedness represented the remaining 26% (±22.4%). Eight species were only recorded inside a reserve while 18 species were only detected outside a reserve across the six regions (electronic supplementary material, figure S4).

Distance-based redundancy analysis (dbRDA) on Jaccard distances showed that both protection and environmental variables significantly explained the dissimilarity in species composition (\( F \)-test = 2.53, \( p < 0.001 \), \( R^2 = 0.20 \)). The turnover component was also significantly explained by both the protection and the environment (\( F \)-test = 2.66, \( p < 0.001 \), \( R^2 = 0.32 \)). The nestedness, however, is not significantly explained by any of the variables (\( p > 0.05 \)) (electronic supplementary material, tables S4–S5).

Partial dbRDA revealed that the protection level, after accounting for environmental conditions, significantly explained 4.8% of fish assemblages (\( F \)-test = 1.80, \( p < 0.001 \), \( R^2 = 0.05 \)) (figure 3a). Fish assemblages inside the reserves were mostly characterized by pelagic fishes, whereas assemblages outside reserves were predominantly characterized by cryptobenthic fish species as shown by the strongest contribution of species scores on the partial dbRDA axis 1 (CAP1, figure 3b; electronic supplementary material, figure S5).

(d) Unpacking the paradox by species traits

We then analysed all species scores along the first axis of the partial dbRDA to determine which traits characterize the species’ associations to the reserves or to the fished areas (left versus right side on figure 3a). Species scores were significantly correlated with their trophic level (Kendall tau = −0.20, \( p < 0.01 \)), common length (Kendall tau = −0.18, \( p = 0.01 \)), and vulnerability to fishing (Kendall tau = −0.21, \( p < 0.01 \)) (figure 4). We also found significant differences in species scores according to their vertical position in the water column (ANOVA \( F \)-value = 7.64, \( p < 0.001 \)), with pelagic species significantly differing from cryptobenthic species.
Species richness of the different fish categories (cryptobenthic, pelagic, and rare) were significantly explained by protection and environment (GLM, $R^2 = 0.40, 0.48, \text{and } 0.45$, respectively, all $p < 0.01$) except for highly vulnerable fishes ($p = 0.09$). Each model accounted for unmeasured variations among regions with model residuals being not significant (Kruskal–Wallis test = 9.37, 0.70, 7.54, and 0.55, respectively, all $p > 0.05$). Cryptobenthic species richness increased by 66% ($p < 0.01$) at 5 km outside compared to inside the reserves and by 136% at 10 km outside ($p < 0.001$) (figure 5a).

Although pelagic species richness decreased with increasing distance to the reserve, this was mainly driven by environmental differences with no significant marginal effect of protection ($p > 0.05$, figure 5b). Rare fish species richness significantly increased by 53% at 5 km outside ($p < 0.01$) and 69% at 10 km outside the reserve ($p = 0.01$) (figure 5c). Vulnerable fish species richness was homogeneous across protection levels with, on average, one species per site (figure 5d).

3. Discussion

(a) Less but more vulnerable species in marine reserves

This study, showing less but different species inside reserves compared to fishing grounds nearby, does not negate the key role of reserves in protecting biodiversity but sheds new light on how under-represented species in classical visual surveys—cryptobenthic, pelagic, and rare—can react counterintuitively to fishing pressure. We reveal, through the sampling of six no-take reserves using a standardized eDNA protocol while accounting for environmental differences, that fish species richness decreases with protection. This paradox could only emerge with a reliable eDNA metabarcoding approach and an extensive genetic reference database. Besides, we highlight a marked species turnover along the protection gradient, indicating a strong difference
**Figure 3.** Partial dbRDA plot of sites (a) and species scores (b). In (a), the dbRDA biplot shows the variation in fish assemblages between sites (purple shapes) explained by the protection level while accounting for environmental variables. Shaded areas: convex hulls grouping sites under the same protection level. Grey points (a) and arrows (b): species scores. Black points (a) and arrows (b): species whose projected length on the first axis (CAP1) belongs to the top 25% of absolute species scores along CAP1. These species are mostly associated with the reserves (left on CAP1) or the fished areas outside (right on CAP1). Blue species: pelagic, red species: cryptobenthic, dark green: benthic, light green: demersal. (Online version in colour.)

\[ \text{Kendall tau} = -0.19 \quad p = 0.007 \]

**Figure 4.** Correlation between species scores along the partial dbRDA axis CAP1 (figure 3), which quantifies its association to the reserve (negative) or areas outside (positive), and the different species traits: (a) trophic level, (b) common length, and (c) vulnerability to fishing. (d) Variation in species scores between cryptobenthic, benthic, demersal, and pelagic species. Groups with different symbols (downward or upward triangle) indicate significant differences between species scores at \( \alpha = 0.05 \) (Tukey test). (Online version in colour.)

\[ \text{Kendall tau} = -0.2 \quad p = 0.003 \]
of fish assemblages within and outside reserves. Reserves offer protection to fishes that are characterized by a high trophic level and large body size, so vulnerable to fishing, as previously shown [44,45]. By contrast, fished areas host less vulnerable species with a lower trophic level and smaller body size (figure 4).

The disparity in trophic level could explain the paradoxical changes in species richness, as fishing pressure and protection can change trophic interactions. Predation is one of the key processes influencing the richness and composition of ecological assemblages [46–48]. Since fishing pressure often targets large predatory fishes [15,49], their removal in fished areas can induce a decrease of top-down control on prey species increasing their richness. Conversely, marine reserves increase predator populations and thus can restore trophic cascades within their boundaries [29], thereby affecting biodiversity at lower trophic levels and potentially causing the local extinction of prey species inside MPAs [50].

This hypothesis is supported by the increasing diversity of cryptobenthic fishes with increasing distance from the reserve (figure 5a). Cryptobenthic fishes represent a large but overlooked dimension of fish biomass and diversity on reef ecosystems [31]. Combined with their rapid growth, high productivity, and high mortality due to predation [51,52], cryptobenthic fishes represent almost 60% of consumed reef fish biomass [53]. As they are not targeted by fisheries but predated by almost every other fish, cryptobenthic fishes could find refuge from predators outside marine reserves and form highly diverse assemblages coexisting in human-dominated areas. However, because of their small body size and cryptic lifestyle, they are easily overlooked by conventional survey methods that do not target them specifically [54].

Our results also show more occurrences of rare species outside the reserves. This pattern can be explained by the ‘oddity effect’ where predators focus on conspicuous prey, in this case rare species, to optimize foraging success [48]. This strategy is especially beneficial when prey tend to form aggregations making it harder for predators to single out individual prey, called the ‘confusion effect’ [48]. So, predators could preferentially target low-abundance prey species inside reserves and remove them. Similar disproportionate effects of predation have previously been demonstrated in field and laboratory experiments. Stier et al. [55] found that predation by the peacock grouper Cephalopholis argus removed 64% of rare species from experimental reef patches but only 36% of common species. In the same vein, Almany et al. [48] observed that the brown dottyback Pseudochromis fuscus, a small generalist predator, targets rare prey in mixed assemblages irrespective of colouration or visual marks. They hypothesize that the odd behaviour of rare species compared to the common ones sets them apart and makes them an easier target for predators.

The strong turnover in fish assemblages under different protection levels suggests that different ecological processes and ecosystem services operate within and outside reserves [7]. Although less diverse, assemblages within reserves are characterized by larger and higher trophic-level species which typically have higher commercial and touristic values. Fish assemblages inside reserves are characterized by pelagic species, especially Sarda sarda, Engraulis encrasicolus, Chelon auratus, Xiphias gladius, and Trachurus mediterraneus. These species have a commercial interest but also contribute to reef productivity through water nutrient enrichment and could play important trophic roles inside reserves [56]. Outside reserves, we find higher biodiversity and assemblages dominated by smaller, (crypto)benthic and demersal species. The increased diversity of cryptobenthic species in impacted areas is hopeful for conservation as well. Cryptobenthic fishes fuel reef trophodynamics and provide crucial food resources for carnivorous fishes [53]. In doing so, they contribute to sustain fish populations in exploited areas. They also provide a reservoir of available biomass for exploited predator populations to recover if fishing pressure is alleviated or suspended in currently fished areas.

**Figure 5.** Results of the generalized linear models testing the effect of protection on species richness while accounting for environmental variables. The plots represent the conditional predicted species richness (+/− confidence intervals) under the different protection levels for each species category: (a) cryptobenthic, (b) pelagic, (c) rare, and (d) vulnerable. Asterisks indicate significant changes in species richness at 5 km or 10 km outside in reference to the reserve (marginal effects test, p < 0.05 *, less than 0.01 **). (Online version in colour.)

(b) Potential and limitations of eDNA metabarcoding

The development of eDNA as a reliable method to monitor biodiversity and evaluate anthropogenic impacts is crucial...
because important changes in biodiversity might currently occur under our radar [39,57]. Depending on the organisms of interest, eDNA can be used to sample whole eukaryote assemblages [58] or more specific taxonomic groups ranging from sponges and corals [59] to larger taxa such as sharks [30]. In our case, we potentially missed some native and common species in the Mediterranean Sea since the teleo marker cannot distinguish the wrasses Symphodus rostratus, S. cinnereus, S. mediterraneus, and S. roissali from each other, or Labrus merula from L. viridis, as well as the rarer pipefishes Syngnathus abaster and S. sp cf taenionotus. Most regional and global reference databases for the teleo metabarcode also still need to be completed to avoid limited species assignments [41]. In our case, we enriched the online genetic database (European Nucleotide Archive) which covered only 31% of all Mediterranean fish species by sampling and sequencing additional 115 species to reach 75% coverage of the regional species pool (see Methods). This unprecedented effort allowed the detection of 97 species ranging from the very small Liechtenstein’s goby Corygobius liechtensteini (2.7 cm) to the large blue shark Prionace glauca (250 cm). Cryptobenthic, pelagic, and elusvie species are often ignored in MPA assessments. Without their detection, we would not be able to uncover the hidden biodiversity patterns between marine reserves and their proximate outskirts. We also show that our genetic reference database was not biased towards some species groups (electronic supplementary material, figure S6), so we are confident that a more exhaustive database would provide the same patterns.

The detection of eDNA in seawater is partly due to its persistence in the environment, which depends on biotic and abiotic factors driving eDNA production, degradation, and transport [60–62]. Much is still unknown about the spatial and temporal resolution of eDNA in the marine coastal environment. Mesocosm experiments report variable decay rates of eDNA in seawater, with half-lives ranging from 1 up to 71 h [60,63]. However, decay and dilution happen faster in natural environments. A field experiment in coastal seawater finds that eDNA becomes undetectable only 1 h after introduction [64]. Coupling decay rates to dispersal distances, estimated by particle tracking models, suggests that suspended eDNA can on average be transported for only 1 km before 50% has decayed [36]. In our system, the average current velocity during sampling was 0.04 m s\(^{-1}\) or approximately 140 m h\(^{-1}\). Combined with relatively short half-lives, it is unlikely that sufficient detectable eDNA could be transported between our sites 5 km and 10 km apart confirming the independence of our sites and the local origin of our signal. These estimates are corroborated by the growing body of empirical studies finding strong spatial fidelity of eDNA signals, differentiating sites only hundreds of metres apart despite tidal and oceanic movements [65–68]. Together, these results demonstrate the applicability of eDNA for local monitoring studies [69]. Yet, accurate particle transport models which directly take into account eDNA concentration, advection, dilution, coastline morphology, and ground-truthed decay rates would allow a better understanding of eDNA transport and detectability patterns across the seascape.

(c) Diversifying management for diversifying regional fish assemblages

The higher species richness found outside reserves does not imply that marine reserves fail to protect biodiversity. It rather tells us that species richness and site-level diversity metrics cannot be considered as reliable indicators of human pressure since they miss important species compositional changes and traits [6,58]. Our results shed light on how conservation, like fisheries management [70], can shape biodiversity patterns at a regional scale. Marine reserves do not necessarily increase species diversity. Rather, a mosaic of protection levels, that creates heterogeneous fishing pressures, can promote heterogeneous ecological processes at various intensities, thus increasing biotic dissimilarity between adjacent areas and the overall level of regional diversity (or \(\gamma\)-diversity). Since the sustainability of ecosystem functioning and the continuous delivery of ecosystem services at the regional scale is positively related to the number of species comprising the regional pool (\(\gamma\)-diversity) [71,72], which depends on both local or \(\alpha\)-diversity and the dissimilarity in species composition between sites (\(\beta\)-diversity) [73], our study suggests that diversifying management options could better sustain ecosystem functioning and limit the ongoing biotic homogenization [74].

4. Methods

(a) Study area and eDNA sampling

Six Mediterranean marine reserves with strict no-take policy and established at least six years prior to sampling were selected for our study (figure 1). The Mediterranean Sea is a hotspot of marine biodiversity [75] but all its ecoregions and territorial waters are under high human pressure [76], except for the very few fully protected marine reserves [77]. Four replicates of 301 of seawater were collected along a 2 km transect inside each site within each reserve, 5 km, and 10 km outside, i.e. in impacted areas, and filtered using a 0.20-\(\mu\)m filtration capsule (electronic supplementary material, Methods).

(b) eDNA extraction and sequencing

eDNA was extracted and amplified by PCR with the fish-specific primer pair teleo targeting a 70 bp fragment at the end of the mitochondrial DNA 12S rRNA gene [78,79] (electronic supplementary material, Methods). PCR reactions were carried out in 12 replicates per sample and unique tags were given to each sample. Libraries were prepared using the MetaFast protocol. A paired-end sequencing (2 x 125 bp) was carried out on a MiSeq (Illumina, San Diego, CA, USA) using the MiSeq Flow Cell Kit Version3 (Illumina, San Diego, CA, USA) following the manufacturer’s instructions. Library preparation and sequencing were performed at Fasteris (Geneva, Switzerland).

(c) Reference database

At the onset of our study only 31% of all Mediterranean fish species [80] were referenced in the European Nucleotide Archive (ENA) [81] (release 138) for the 12S rRNA fragment targeted by the teleo primers. To supplement this reference database, fin clips of 115 fish species from the North-Western Mediterranean were collected from fisheries landings and added to the database. DNA was extracted from tissue samples and a 12S rRNA gene fragment of ca 675 bp encompassing the teleo metabarcoding fragment was targeted using the forward primer V05F_898 and the reverse teleo primer [82]. A 340 bp fragment was additionally targeted for 16 species using the newly designed forward primer MF12S_F (5'−CTAGAGGAGCCTGTYVT) and the reverse primer MF12S_R (5'−GRHAAGTCGTAACATGGTA) (electronic supplementary material, Methods).

The final reference database used in this study contained sequences of 320 species corresponding to 41% of all...
Mediterranean fish species but 75% of the regional North-Western species pool. The remaining gap did not bias the biodiversity assessment made in this study (electronic supplementary material, Methods, table S6, figure S6).

(d) Taxonomic assignment of reads
The sequence reads were analysed using the OBITools package [78,83]. Taxonomic assignment of reads was performed using the program ECOTAG, with both the new regional fish reference database and the public reference database of sequences extracted from ENA (release 140) using the ECOPCR program [84,85]. Reads showing less than 98% similarity were removed. Taxa were preferentially assigned based on the local reference database, except if the similarity was higher for the public reference database. The resulting dataset was manually checked to correct erroneous identifications and remove foreign species (electronic supplementary material, Methods).

(e) Diversity indices
We compared total fish species richness among protection levels as well as the richness of the cryptobenthic, pelagic, rare, and vulnerable species. Cryptobenthic species were selected based on their families [31]. Pelagic species were those defined by the ‘Vertical Distribution’ parameter in the FishMed database [86]. Rare species were those detected in two samples or less within each region. The vulnerability of species to fishing was obtained from FishBase [87]. The vulnerable species are those with a vulnerability higher than 70, which corresponds to ‘high’ or ‘very high’ vulnerability to fishing on a scale from 1 to 100 [87].

We estimated species dissimilarity or β-diversity between assemblages using the Jaccard distance. To determine the relative contribution of species turnover and nestedness to total β-diversity, we used the additive partitioning of the pairwise Jaccard dissimilarity [42]. This framework teases apart the variation in species composition from species turnover only, which is independent of richness, and from nested patterns [43]. We calculated the total dissimilarity, turnover, and nestedness between all samples using the R packages vegan and betapart [88,89].

(f) Modelling reserve effect
We used GLMs to investigate the effect of protection on species richness while accounting for environmental differences represented by the PCA axes. After checking for their distribution, total, cryptobenthic, pelagic, and rare species richness were modelled using a Gaussian distribution, whereas vulnerable species richness was modelled using a Poisson distribution.

We determined model fit by calculating the $R^2$ for each model, and tested the conditional effect of coefficients by calculating the marginal effects with the R package margins [90]. We tested the effects of potentially missed important factors by comparing residuals of our models between regions using a Kruskal–Wallis test.

We used a dbRDA (function capscale, package vegan) to analyse changes in assemblage composition measured by Jaccard distance between samples in relation to the protection level and the four environmental PCA axes. We computed additional dbRDA on the species turnover and nestedness with each time the protection category and four environmental PCs as explanatory variables. Significance of the models as well as the significance of each axis and of the marginal effect of each variable were tested using ANOVA-like permutation tests with 9999 permutations as implemented in the vegan’s anova.cca function [89,91].

Next, we computed a partial dbRDA using the Jaccard distance to isolate the effect of protection after accounting for environment [92]. From this partial dbRDA, we extracted the species scores along the axis that explains most of the variance to infer which species contribute most to the differences in assemblage composition between protection levels. We focused on species whose projected length on the first axis (CAP1) belongs to the top 25% of absolute species scores. We then used the Kendall rank correlation coefficient to test the correlation between the species scores and the species’ trophic level, common length, and vulnerability to fishing. We used ANOVA and Tukey post hoc to test for the differences in species scores according to their vertical position (cryptobenthic, benthic, demersal, and pelagic). All analyses were carried out in R v. 3.6.1 [93].

Data accessibility. The data and R codes to replicate analyses and figures are available at https://github.com/eboulanger/MEDeDNA-reserves. The new teleo reference sequences and Illumnia raw sequences are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.189312xc1 [94] and https://doi.org/10.5061/dryad.9k51cr [95].

Authors’ contributions. E.B., N.L., P.B., T.D., S.M., and D.M. designed the study. E.B., P.L., J.B.J., J.D., F.H., and N.G. conducted fieldwork. A.V. and T.D. supervised the biomolecular analysis and A.V. performed the bioinformatic analyses. E.B. and D.M. performed the statistical analysis. E.B. and D.M. designed the statistical analyses and wrote the first draft. E.B., N.L., J.B.J., S.M., and D.M. interpreted data and wrote the paper. All authors offered revision suggestions and approved the final version of the manuscript.

Competing interests. We declare we have no competing interests.

Funding. Funding was provided by the Region Occitania, the Agence de l’Eau RMC and the RESERVEBENEFIT project through the 2015-2016 BiodivERsA COFUND call with the ANR (France).

Acknowledgements. The authors thank all staff and students involved in the field and laboratory work, as well as the managers of the studied MPAs.

References


