ICES Journal of Marine Science, 2026, Vol. 83, Issue 3, fsag035
https://doi.org/10.1093/icesjms/fsag035
Original Article

Significance of rare species in evaluating marine conservation effects on fish community using hill numbers

Rongjie Zhao [1], Bin Kang [2], Yip Hung Yeung [3], Louise Wai Hung Li [3], Yifang Chen [1], Jingying Sun [4], Jianghua Yang [5], Xiaowei Zhang [6], Kenneth Mei Yee Leung [1,7,*], Meng Yan [1,7,*]

[1]State Key Laboratory of Marine Environmental Health and Department of Chemistry, City University of Hong Kong, Hong Kong 999077, China
[2]Key Laboratory of Mariculture, Ocean University of China, Ministry of Education, Qingdao 266003, China
[3]Agriculture, Fisheries and Conservation Department, the Government of the Hong Kong Special Administrative Region, Hong Kong 999077, China
[4]Nanjing E-genomics Technology Co., Ltd, Nanjing 210019, China
[5]School of Marine Science and Engineering, Nanjing Normal University, Nanjing 210046, China
[6]State Key Laboratory of Pollution Control & Resource Reuse, School of the Environment, Nanjing University, Nanjing 210023, China
[7]Research Centre for the Oceans and Human Health, City University of Hong Kong Shenzhen Research Institute, Shenzhen 518000, China
*Corresponding author. Kenneth M.Y. Leung, State Key Laboratory of Marine Environmental Health, Yeung Kin Man Academic Building, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong. E-mail: kmyleung@cityu.edu.hk; Meng Yan, State Key Laboratory of Marine Environmental Health, Yeung Kin Man Academic Building, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong. Kenneth M.Y. Leung. E-mail: mengyan@cityu.edu.hk

Abstract

Marine protected areas (MPAs) have been established worldwide to protect biodiversity and enhance ecosystem services. Rare species play vital roles in marine conservation as they exhibit unique ecosystem functions and are vulnerable to extinction. However, determining their importance to conservation effects remains challenging due to the lack of feasible and standardized evaluation methods. The noninvasive environmental DNA (eDNA) method is increasingly used in marine biomonitoring because of its efficiency and sensitivity. Hill numbers provide a uniform evaluation scheme for assessing biodiversity. In this study, we calculated Hill numbers from eDNA data to determine marine fish diversity in MPAs. Traditional survey data and different bioinformatic clustering pipelines were also included to decide the impact of survey and analysis methods on ecological interpretation. The results showed that eDNA detected significantly more species than contemporaneous traditional survey data. We further found less significant inter-group differences for different seasons and pipelines when rare species diminished, indicating that excluding rare species in evaluations could affect the ecological interpretation of marine conservation effects. Protection time of MPAs demonstrated positive correlations with all diversity indices, albeit not statistically significant. In conclusion, we demonstrated the importance of rare species in marine biodiversity assessment and recommend appropriate usage of eDNA and downstream diversity indices to achieve a comprehensive evaluation of the marine conservation effect of MPAs.

Keywords subtropical, biomonitoring, phylogenetic diversity, path model

Introduction

Under the scenario of global climate change, anthropogenic activities have disturbed marine coastal ecosystems, caused severe habitat loss and resulted in declining marine biodiversity (Blowes et al. 2019, Chaudhary et al. 2021). Consequently, it is imperative to adopt effective actions to prevent biodiversity from declining and protect ecosystem functions. Marine protected areas (MPAs) are established worldwide to protect biodiversity within designated areas (Roberts et al. 2017, Sala et al. 2021). Currently, there are over 18 000 MPAs which cover approximately 8.2% of the whole ocean area according to the Marine Protected Planet (UNEP-WCMC 2024). Hong Kong, located in the subtropical region, is surrounded by the South China Sea and influenced by freshwater input from the Pearl River. Hydrological conditions are varied and nourish tremendous marine biodiversity in the region (Morton et al. 1983, Ng et al. 2017). To protect local biodiversity, especially fishery resources that serve as important food for humans and play a vital role in the marine ecosystems, eight marine parks (MPs) and one marine reserve (MR) have been established in succession in Hong Kong since 1996. These MPAs are regulated by the Agriculture, Fisheries and Conservation Department (AFCD), the Government of the Hong Kong Special Administrative Region (SAR) (Wong 1996). MPAs have been shown to significantly enhance local biodiversity in Hong Kong (Liu et al. 1997). The fishery survey conducted by the AFCD demonstrated a positive conservation effect of MPs and the MR on fish species richness, biomass, and body length (AFCD 2021). Nevertheless, studies of the status of fish biodiversity

Received: 1 October 2025. **Revised:** 27 February 2026. **Accepted:** 2 March 2026

within MPAs and its influencing mechanism are still lacking. Biodiversity is jointly shaped by intrinsic environmental conditions, anthropogenic factors, and protection activities, including protection level, duration, and size (Frid et al. 2023, Clausius et al. 2025). However, the current evaluation scheme for marine conservation effects rarely disentangles the relative importance of protection activity and spatiotemporal factors. It is imperative to conduct the investigation and provide fundamental proof for marine conservation management.

Protection activities usually focus on diversity hotspots with high species richness, while rare species, i.e. species with small population sizes and/or limited distribution range may not occur in these hotspots (Hepper & Synge 1982, Prendergast et al. 1993). Thus, conservation efforts mainly exert on the dominant species with wide distribution ranges in the community, while rare species with more restricted ranges were less likely to be protected (Lawler et al. 2003). However, rare species with low abundance are more vulnerable to environmental changes and more susceptible to extinction (Davies et al. 2004). Moreover, fish species that perform unique functions and have low functional redundancy are usually rare in an ecosystem and are important for ecosystem functioning (Mouillot et al. 2013). Extinction of rare species can lead to functional richness reductions and further affect multiple ecosystem processes (Leitão et al. 2016). Increasing the representation of these range-restricted species in conservation planning could contribute to constructing effective protection networks (Larsen et al. 2007). Despite their high vulnerability and critical functional roles, rare species are often overlooked in conventional hotspot-based conservation, making targeted assessment a necessary step to address current protection gaps (Lin et al. 2024). It is important to determine the status of the local fish community, including rare species, to further advise the ecological management in the area.

Fishery surveys within MPAs in Hong Kong currently rely on traditional survey methods including hand lining, long lining, and gill netting methods. These methods depend largely on fishers' experience and can lead to sampling bias, inadequate sampling coverage, and more importantly, neglecting rare species. Thus, finding supplementary methods for biomonitoring in the MPAs is necessary. Environmental DNA (eDNA) refers to genetic materials obtained from environmental samples such as water, sediment, and air (Taberlet et al. 2012). With the combination of metabarcoding and downstream bioinformatic analyses, it provides easy access to biodiversity at local and regional scales. The eDNA method has been widely used in biomonitoring in various environments as it is efficient and low-cost (Bohmann et al. 2014, Djurhuus et al. 2020) and more sensitive than traditional methods at detecting species richness as shown in meta-analyses (McElroy et al. 2020, Keck et al. 2022). Furthermore, the eDNA approach is sensitive to biotic signals with low concentrations, making it an efficient method for detecting rare species and suitable for marine conservation research (Jerde et al. 2011, Thomsen et al. 2015, Kvalheim et al. 2024). Several studies have demonstrated the feasibility of eDNA method in assessing biodiversity within MPAs (Gold et al. 2021, Capurso et al. 2023, Polinski et al. 2025). Nevertheless, the downstream bioinformatic analysis may severely impact the ecological interpretation of the data (Beng et al. 2020). Different analysis methods, e.g. operational taxonomic unit (OTU)-level and amplicon sequence variant (ASV)-level clustering pipelines, can produce varying interpretation of the same metabarcoding dataset (Prodan et al. 2020). OTU-level pipelines cluster sequences at a typical identity of 97%, while ASV-level pipelines provide higher resolution by reconstructing the exact biological sequences present in the sample (Callahan et al. 2017). In studies where OTU- and ASV-level methods were both used and yielded different ecological interpretations of diversity, the differences in ecological outcomes were found to be the result of differences in how rare calculating units were interpreted (Chiarello et al. 2022). To achieve an unbiased understanding of target ecosystems, adopting appropriate sequencing and analysis strategies is both necessary and challenging (Zinger et al. 2019, Mächler et al. 2021).

Another key component for ecological evaluation is the bioindicator selected to assess how diverse the ecosystem is. Commonly used bioindicators including species richness, Shannon, Simpson, Rao's quadratic entropy, and Pielou's evenness are one-sided diversity indices (Shannon 1948, Simpson 1949, Pielou 1966, Rao 1982). Usage of these indices requires thorough considerations such as calculating unit, abundance relativity, and representativeness to obtain meaningful diversity measures within an ecosystem (Alberdi & Gilbert 2019). The Hill numbers address the need for comprehensive bioindicators by incorporating an array of diversity indices, including the effective number of species, variations of Shannon diversity, and Simpson diversity, forming a unified framework to provide intuitive and meaningful estimates of diversity (Hill 1973). The commonly used indices mentioned above are highly sample-dependent and are therefore challenging to compare across different studies as sampling efforts can be inadequate and uneven (Coddington et al. 2009, Marathe et al. 2021). Hill numbers eliminate the effect of under-sampling by quantifying diversity with equal sample coverage and leveraging the sensitivity of common and rare species (Roswell et al. 2021). In addition, relative abundance and under-sampling can have a large effect when determining the regional variation of a biological community, i.e. beta diversity. Hill numbers resolve the above discrepancy by calculating the effective number of species, where beta diversity represents the effective number of distinct assemblages (Chao et al. 2012, Chao et al. 2014). Therefore, Hill numbers can serve as a robust tool for biodiversity assessment.

In this study, we collected water samples from MPAs in Hong Kong across all four seasons and utilized the eDNA method to investigate fish diversity and assess the effectiveness of marine protection on the fish communities, especially the rare species. Traditional survey data collected over 21 years by the AFCD was included to verify the reliability of the eDNA data. By using different clustering methods and multidimensional biodiversity indicators, including the Hill number reflecting relative abundance and the rarity index reflecting species occurrence, we aim to (i) verify the feasibility of eDNA metabarcoding by comparing the results to those inferred from traditional survey datasets, (ii) determine whether rare species affect ecological evaluation outcomes, (iii) evaluate the advantage of Hill numbers over traditional bioindicators, and (iv) assess the impact of marine protection measures on local fish communities.

Materials and methods

Sampling sites and sample collection

From October 2022 to September 2023, sample collection was conducted in sites with three different protection levels, i.e. the no-

take MR, the partially protected marine park (MP) and the non-protected area (NPA) (Fig. 1a, Table S1). The Cape D'Aguilar Marine Reserve (CDMR), a no-take area, is in the southern waters of Hong Kong. The Tung Ping Chau Marine Park (TPCMP), Hoi Ha Wan Marine Park (HHWMP), and Yan Chau Tong Marine Park (YCTMP) in the eastern waters were partially protected, and established to protect the diverse coral community. The Brothers Marine Park (BMP), Sha Chau and Lung Kwu Chau Marine Park (SCLKCMP), South Lantau Marine Park (SLMP), Southwest Lantau Marine Park (SWLMP) in the western waters were partially protected and aimed to protect the prey species of Chinese White Dolphins (*Sousa chinensis*). The CDMR (20 hectares), YCTMP (680 hectares), HHWMP (260 hectares) and SCLKCMP (1,200 hectares) were designated as MPAs in July 1996. The TPCMP (270 hectares), BMP (970 hectares), SWLMP (650 hectares) and SLMP (2,067 hectares) were enacted from 2001, 2016, 2020, and 2022, respectively. The proposed North Lantau MP was situated in the western water and nominated for the conservation of the habitat of Chinese White Dolphins and the marine environment. Recently, it has been designated as an official MP since 1st November 2024, but it was not under regulation before and during the sampling time in this study. Sample collection within the MPs and the MR was conducted with permits (#000 780 (28) and #000 781 (30) in AF MPD/09/3 Pt.27) issued by AFCD.

The sampling was implemented seasonally, i.e. Transitional II (T2, October–November 2022), Dry (December 2022–March 2023), Transitional I (T1, April–May 2023), and Wet (June–September 2023). At each site, water depth was measured using a sonar meter (HONDEX PS-7). Two liters of seawater were collected using a two-liter Van Dorn water sampler from the surface (1 m below the sea surface), middle (half of the water depth) and bottom (1 m above the sea bottom) water layers, respectively, to have complete biodiversity information at the greatest efficiency (Fig. 1b). Seawater from each depth was pooled, mixed, and subdivided into three identical replicates containing the same amount of seawater from three different layers. Two liters of ultrapure water (Milli-Q®) were loaded into a sterile bag at the laboratory and taken to the field at each site and exposed to the air during the field sampling to be used as a field blank. To avoid cross-site contamination, the water sampler was first decontaminated with 10% bleach, then rinsed with on-site seawater of the upcoming site three times between different sites. All samples were stored in a light-protected 4°C cold room and each sample was filtered through a 0.22 μm mixed cellulose esters (MCE) membrane (MF-Millipore™) within 24 h of the collection.

In total, 108 membrane samples (4 seasons × sites × 3 replicates) and 36 field blank controls (4 seasons × 9 sites) were obtained. The DNeasy Blood & Tissue Kit (Qiagen, Germany) was used for eDNA extraction from all membranes according to the manufacturer's instructions with slight modifications, including 4 times the amount of Buffer ATL, Proteinase K, Buffer AL, and ethanol to fit our samples. The extracted DNA was diluted in TE buffer (10 mM Tris·Cl, 0.5 mM EDTA, pH 9.0) and stored in a −80°C freezer until further amplification.

The traditional survey was conducted using gill netting, hand lining, and long lining at the same sampling sites and on the same date as eDNA sampling (AFCD 2025). On the sampling day, traditional sampling was conducted after eDNA sampling to avoid influence on eDNA samples. The sampling range of the traditional gears could cover the entire water column. As the number of specimens collected by traditional methods was limited in the western sites, we only included the samples collected in the eastern MPAs, i.e. YCTMP, TPCMP, HHWMP, and CDMR, for the comparison between the two surveys.

PCR, high-throughput sequencing, and bioinformatic analysis

The primer set 12S_V5 (F: 5'-ACTGGGATTAGATACCCC-3', R: 5'-TAGAACAGGCTCCTCTAG-3') targeting the fish mitochondrial *12S ribosomal RNA* gene was used for the amplification (Riaz et al. 2011). A unique 8-bp barcode was added at the 5' end of the forward and reverse primers to distinguish different samples. After amplification, samples were purified using TaKaRa MiniBEST DNA Fragment Purification Kit (TaKaRa, Japan), quantified using a Qubit Flex Fluorometer (Thermo Fisher Scientific Inc., USA) and pooled with equimolar concentrations to make the sequencing library. Samples were selected for library preparation if they had a final volume $\geq$ 20 μL and concentration $\geq$ 60 ng/μL. The PCR products from the SCLKCMP T2, CDMR Wet, and YCTMP Wet samples did not qualify for constructing sequencing libraries and were excluded from sequencing. Four sequencing libraries were generated and one PCR blank control was added to each library to measure and account for any potential contamination. In total, 97 samples, 36 field blank controls, and 4 PCR controls were sequenced. The pooled samples and 1 PCR control from T2 season were sequenced using the DNBSEQ PE150 platform at BGI. The pooled samples from T1, Wet, and Dry seasons and three PCR controls were sequenced using the Illumina NovaSeq PE150 platform at Novogene.

We used two clustering pipelines for the bioinformatic analysis. The first one generated ASV using DADA2 package with Ribosomal Database Project (RDP) Classifier (hereafter, RDP-based). The raw high-throughput sequencing data were demultiplexed and the barcodes and primer sequences were trimmed using *Cutadapt* in Qiime2 (Martin 2011, Bolyen et al. 2019) to get paired-end data. We used two methods to assign sequencing reads to taxa. The DADA2 package in R v4.3.1 was used following the standard tutorial for quality control, including trimming, filtering, denoising, merging paired-end reads, dereplication and chimera removal, to generate ASVs (Callahan et al. 2016). Taxonomic annotation of ASVs was based on the RDP Classifier and the resulting ASV table reported the abundance of cleaned reads for each ASV (Wang et al. 2007). The second method used Usearch v.11 to generate OTUs following the standard tutorial (Edgar 2010) (hereafter, Usearch-based). Paired-end fastq files were merged using *fastq_mergepairs* to generate a single fasta file for each sample. The *cluster_otu* function in Usearch v.11 was used to cluster OTUs with 97% similarity based on fasta files. Taxonomy was assigned using *usearch_global* and the abundance of each OTU was calculated using *otutab* (Edgar 2010, 2016). We used a strict data curation standard to exclude potential sequencing errors. Specifically, the ASVs and OTUs that could be detected by sampling and PCR blank controls, with less than three occurrences in all samples, or with a total abundance of less than 100 reads were regarded as contamination and excluded from downstream analysis.

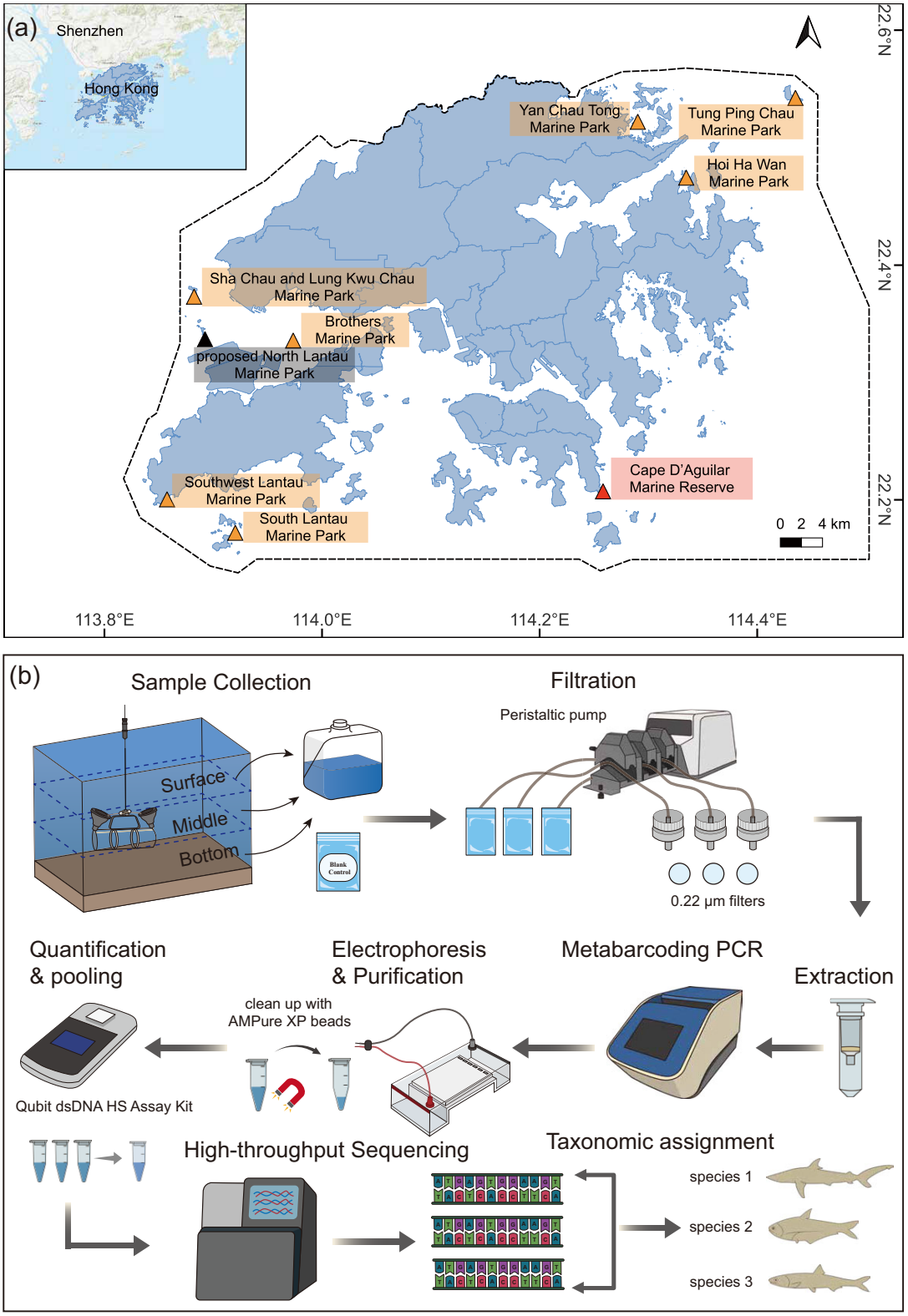

Figure 1 (a) Locations of Hong Kong SAR and sample sites in this study. Fully protected MR, partially protected MPs and NPA were shown in orange, red, and blue, respectively. (b) eDNA sample collection procedure used in this study.

Table 1 Data type and data source for five datasets used to calculate Hill numbers

Dataset	Data type	Data source
I	All ASVs with precise taxonomic annotation	eDNA (2022–2023)
II	All OTUs with precise taxonomic annotation	eDNA (2022–2023)
III	ASVs assigned to species detected in historical traditional surveys	eDNA (2022–2023), historical traditional surveys (2000–2023)
IV	OTUs assigned to species detected in historical traditional surveys	eDNA (2022–2023), historical traditional surveys (2000–2023)
V	Species detected by traditional survey	Contemporaneous traditional survey (2022–2023)

For the taxonomic annotation, we assigned the ASV and OTU sequences using the MIDORI2 database (v. 4.06) (Leray *et al*. 2022). Only species with a similarity threshold > 98% and match length > 100 were regarded as a precise match and retained for further analyses. All species names, taxonomic status, and distribution range were manually checked using *Match taxa* in WoRMS online (WoRMS Editorial Board. 2026). Species that had no local distribution were excluded from downstream analysis (Table S2).

Establishment of five datasets

In this study, five datasets were generated to compare different survey methods and bioinformatic pipelines for exploring fish diversity in MPAs (Table 1). Apart from the contemporaneous eDNA and traditional survey during 2022–2023, we also included the 21-year (2000–2023, except 2019 and 2020) historical traditional survey data. The data covers fish diversity from the four eastern MPAs, i.e. CDMR, HHWMP, TPCMP, and YCTMP (Fig. 2a). During the historical surveys, samples were collected twice per wet and dry seasons. A combination of sampling gears included cage trap, hand line, long line, and gill net was used for each survey. For the comparison between eDNA and traditional survey, we only included four eastern MPAs, i.e. CDMR, HHWMP, TPCMP, and YCTMP (Fig. 2a).

Hill number calculations

The R package iNEXT.beta3D was used to calculate the TD and decompose the gamma, beta, and alpha components at three diversity orders with abundance-based data (Chao et al. 2023, Chao & Hu 2023, Chao, Thorn et al. 2023). Three diversity orders, including ${}^{0}\mathrm{D}$ (species richness, $q = 0$), ${}^{1}\mathrm{D}$ (Shannon diversity, $q = 1$) and ${}^{2}\mathrm{D}$ (Simpson diversity, $q = 2$) were determined based on the different priorities of abundance. Specifically, when $q = 0$, Hill numbers count species equally without regard to their relative abundance; when $q = 1$ ($q \rightarrow 1$), individuals are counted equally and thus species are counted in proportion to their abundances; when $q = 2$, abundant species are favored and rare species are discounted and can be interpreted as the effective number of dominant or very abundant species in the assemblage. As the sequencing effort for each sample was usually not equal in high-throughput sequencing, we adopted the Hill numbers of each sample at the 100% level of the Sample-coverage-based (or coverage-based) rarefaction and extrapolation (R/E) sampling curves.

Hill number I: TD was calculated based only on the relative abundance of all ASVs.

Hill number II: TD was calculated based only on the relative abundance of all OTUs.

Hill number III: TD was calculated based on the overlapped historical traditional data and annotated ASV data.

Hill number IV: TD was calculated based on the overlapped historical traditional data and annotated OTU data.

Hill number V: TD was calculated based only on the contemporaneous traditional survey data (Dataset V).

We also calculated the TD and phylogenetic diversity (PD, here we calculated meanPD, effective number of equally divergent lineages) of all ASVs and OTUs based on samples collected from all eight MPAs and one NPA. To generate a comprehensive evaluation scheme of the marine protection activity, TD and PD were included and calculated using R package iNEXT.beta3D. TD was calculated based on the relative abundance of ASVs/OTUs. For the PD, the final RDP/Usearch-based sequences were aligned using MAFFT (Katoh et al. 2002) and used to construct a phylogenetic tree using FastTree (Price et al. 2009) in R.

Statistic analyses

To determine the difference levels between different Hill numbers, a one-way analysis of variance (ANOVA) analysis was implemented using package rstatix (Kassambara 2025) in R. The Kruskal–Wallis test and Pairwise Wilcoxon Rank Sum test were used to determine the significance level of differences between different Hill number datasets as the data were not normally distributed (Kruskal et al. 1952, Corder et al. 2011).

The rarity weights for all species obtained by eDNA and traditional survey methods were calculated based on the rarity cut-off points of the first quartile of species occurrences (Gaston 1994) using the Rarity package in R (Leroy et al. 2012). The Index of Relative Rarity (Irr) indicating the average rarity weights of all species in an assemblage, was then calculated for each sample of the RDP-based, Usearch-based, and contemporaneous traditional survey datasets. Finally, rare species were selected based on the rarity cut-off point of each dataset.

To determine the influence of region, season, and protection activity on the taxonomic and phylogenetic diversity variation of fish communities, the Partial Least Squares Path Modeling (PLS-PM) was completed using the PLSPM package (Sanchez 2013) in R v.4.3.1. The Hill numbers calculated based on eDNA samples from all MPAs and NPA were used for analysis. Reflective mode A was selected to construct the model. Protection, season, and region were set as latent variables and diversity orders (${}^{0}\mathrm{D}$, ${}^{1}\mathrm{D}$ and ${}^{2}\mathrm{D}$) were set as manifest variables. Specifically, the protection level, MPA size,

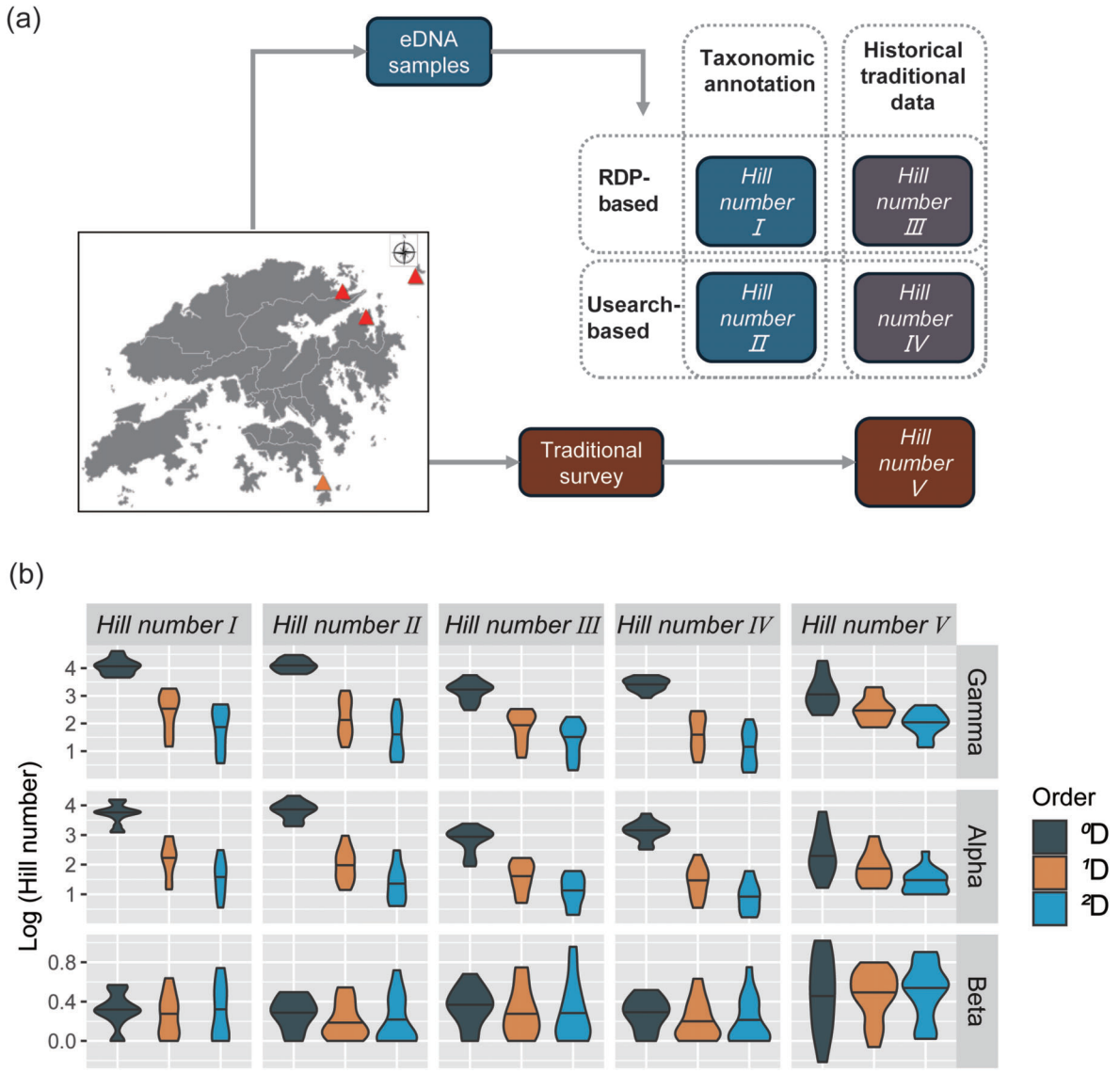

Figure 2 (a) The five datasets based on eDNA and traditional survey samples in MPAs of eastern waters of Hong Kong. (b) taxonomic diversity (TD) calculated based on the above five datasets. Line inside each box indicates median value within each group. Different diversity orders were illustrated with different colors (^{0}D: species richness; ^{1}D: Shannon diversity, and ^{2}D: Simpson diversity).

and MPA age were included in the protection group to generate a comprehensive evaluation. Three diversity aspects (gamma, alpha, and beta) were combined to form a calculation unit of diversity. As the PLS-PM only considers numerical factors, we converted the season name from character factors, i.e. T2, Dry, T1, and Wet to 1, 2, 3, and 4, respectively. Similarly, protection level was renamed from fully protected, partially protected, and nonprotected to 2, 1, and 0, respectively.

Results

Fish composition and richness

The historical data of traditional surveys during 2000–2023 included 284 species from 67 families (Table S3). The number of fish species collected by traditional methods during 2022–2023 in the eastern MPAs was 119 species from 46 families (Table S3).

For RDP-based and Usearch-based pipelines, 35.9% and 35.7% of the total sequences remained after the bioinformatic analysis, respectively. After taxonomic annotation, 722 ASVs were successfully assigned to 234 species from 162 genera, 76 families and 32 orders, and 314 OTUs were assigned to 234 species from 161 genera, 76 families and 30 orders. In total, 319 species were detected by RDP-based and Usearch-based methods, and 149 species (46.7%) were detected by both methods. The top 20 abundant fish species detected by RDP-based and Usearch-based datasets are illustrated as Fig. S1 and **S2**, respectively.

Compared to the historical data of traditional surveys during 2000–2023, 82 (29% of 284 species collected) fish species were detected by both traditional survey and RDP-based dataset and 93 (33% of 284 species collected) fish species for Usearch-based dataset (Table S4, S5, Fig. S3).

TD calculated by different assessing pipelines

The Hill numbers for each of the five datasets showed inter-group differences (Fig. 2b). After overlapping the taxonomically annotated eDNA data with 21-year historical traditional survey data, the seasonal patterns remained highly consistent, but the gamma and alpha diversity decreased by half. Compared with the eDNA data and overlapped historical data, the seasonal pattern of traditional survey data during 2022–2023 had large deviations in terms of all three aspects of Hill numbers.

When analyzing different Hill number datasets at multiple diversity orders, pipelines for generating statistical datasets caused significant differences (Table 2). For the gamma and alpha aspect, the *Hill number I* and *II* datasets were significantly higher than *Hill number V* at ^{0}D order, but showed no significant differences at ^{1}D and ^{2}D orders. *Hill number III* was significantly lower than *Hill number I* at ^{0}D order but not at ^{1}D and ^{2}D orders. The same pattern was found for *Hill number II* and *Hill number IV*. Specifically, the beta diversity showed no significant differences across diversity orders.

Fish TD and phylogenetic diversity detected by eDNA

The TD of RDP-based (ASVTD) and Usearch-based (OTUTD) datasets were similarly distributed, while the PD of RDP-based (ASVPD) and Usearch-based (OTUPD) datasets showed inconsistent distributions (Fig. 3a, Table S6). According to the ANOVA results (Table S7), the gamma, alpha, and beta diversity of ASVTD and OTUTD showed no significant difference across diversity orders. For the phylogenetic aspect, at diversity order $q = 0$, ASVPD was significantly higher than OTUPD.

Across the sampling duration, ASVTD, ASVPD, OTUTD, and OTUPD showed an apparent seasonal variation (Fig. 3a). At ^{0}D order, there were significant differences among different seasons for ASVTD, ASVPD, OTUTD, and OTUPD, with the dry season showing higher values than the other seasons (Table S8). At ^{1}D and ^{2}D orders, ASVTD and OTUTD showed no significant differences among seasons, except for beta diversity.

The ANOVA results indicated that there was no significant difference among different protection levels according to the Hill numbers, except for ${}^1D_\gamma$, ${}^2D_\gamma$, and ${}^2D_\beta$ of OTUTD (Fig. 3b, Table S9).

Impact of marine protection and spatial-temporal factors at different diversity orders

The PLS-PM results showed that sampling region, sampling season, and protection activity, including MPA level, MPA age, and MPA size, had varied impacts on the diversity metrics (Fig. 4). The goodness of fit values indicated comparable or better fitting models for the RDP-based datasets compared to Usearch-based. MPA age, i.e. the protection duration, showed positive correlations with all diversity indices, though no significant relationship was found. MPA level showed positive relationships with all RDP-based diversity indices, but was negative for Usearch-based diversity indices except ^{0}D of OTUTD. MPA size showed different relationships with different diversity orders. For RDP-based diversity, positive relationships with ^{1}D and ^{2}D, and negative with ^{0}D were detected for ASVTD, but were reversed for ASVPD (Fig. 4a, b). MPA size showed positive relationships with Usearch-based diversity except for ^{2}D of OTUPD (Fig. 4c, d).

Sampling region and season exerted opposite effects on ASVTD (Fig. 4a). Additionally, sampling region had negative correlations with ^{2}D of ASVPD, but sampling season showed significant positive correlations with ^{0}D, ^{1}D, and ^{2}D of ASVPD (Fig. 4b). For the OTUTD, sampling season showed a significant positive relationship with ^{0}D, a positive correlation with ^{1}D, and negative with ^{2}D (Fig. 4**c**). For the ASVPD, sampling region showed negative correlations with all diversity orders, while season showed positive correlation with ^{0}D, and negative with ^{1}D and ^{2}D (Fig. 4**d**).

Correlation between Hill number and rarity index

Among all the species detected, 69 (29%), 71 (30%), and 55 (46%) were recognized as rare species by RDP-based, Usearch-based and contemporaneous traditional survey, respectively, according to their occurrence in the samples (Fig. S4, Table S10). The samples with Irr ranking top three were collected from SLMP, YCTMP, and CDMR in T2 for RDP-based dataset, CDMR and HHWMP in T1 for Usearch-based, and HHW in T1, T2, Wet, and YCT in Wet for the traditional survey.

The linear regression model showed that there was no significant relationship between the alpha aspect of Hill number and rarity index for both eDNA and traditional surveys (Fig. 5). Notably, as the diversity order increased, the correlations decreased, although all models showed weak fitting levels.

Discussion

By combining different clustering pipelines and calculating Hill numbers, we verified the importance of rare species in the biodiversity assessments used to quantify the conservation effect brought by MPAs on fish diversity. We also detected significantly higher fish diversity estimates from eDNA compared to traditional surveys, likely impacted by the presence of rare species that weren't detectable using traditional methods. Our findings show that the diversity of fish species was positively correlated with the duration of MPAs, and that there were significant seasonal differences in fish diversity using eDNA with the lowest richness occurring during October-December.

eDNA metabarcoding detected higher fish diversity than traditional surveys

The noninvasive eDNA method has been increasingly used in marine biodiversity monitoring in recent years and is especially suitable for investigating how biodiversity responds to marine protection activities (Takahashi et al. 2023). The eDNA method proved to be a good supplementary choice for marine fish biomonitoring in the Pearl River Estuary (Zou et al. 2020). In this study, we compared the eDNA method with historical traditional surveys and found that eDNA detected less fish species than the 21-year historical traditional survey data, indicating deficiency in reference databases and potential species loss over the last two decades in Hong Kong.

Table 2 Results of Kruskal–Wallis test and Pairwise Wilcoxon Rank Sum test for different Hill number datasets on rank of different diversity orders

Diversity order	Kruskal-Wallis test				Pairwise Wilcoxon Rank Sum test[a]				
	χ^2	*df*	*n*	*p*	*Hill number*				
					I	*II*	*III*	*IV*	*V*
${}^0D_\gamma$	47.8	4	74	<0.001	A	a	b	b	b
${}^0D_\alpha$	48.39	4	74	<0.001	A	a	bc	b	c
${}^0D_\beta$	4.42	4	69	>0.05	A	a	a	a	a
${}^1D_\gamma$	16.41	4	74	0.003	A	ab	ab	b	a
${}^1D_\alpha$	13.98	4	74	0.007	A	ab	ab	b	ab
${}^1D_\beta$	9.83	4	69	0.043	A	a	a	a	a
${}^2D_\gamma$	15.88	4	74	0.003	Ab	ab	ab	b	a
${}^2D_\alpha$	12.6	4	74	0.013	A	ab	ab	b	ab
${}^2D_\beta$	10.52	4	69	0.033	A	a	a	a	a

[a]: Groups share with the same letter indicating an insignificant difference.

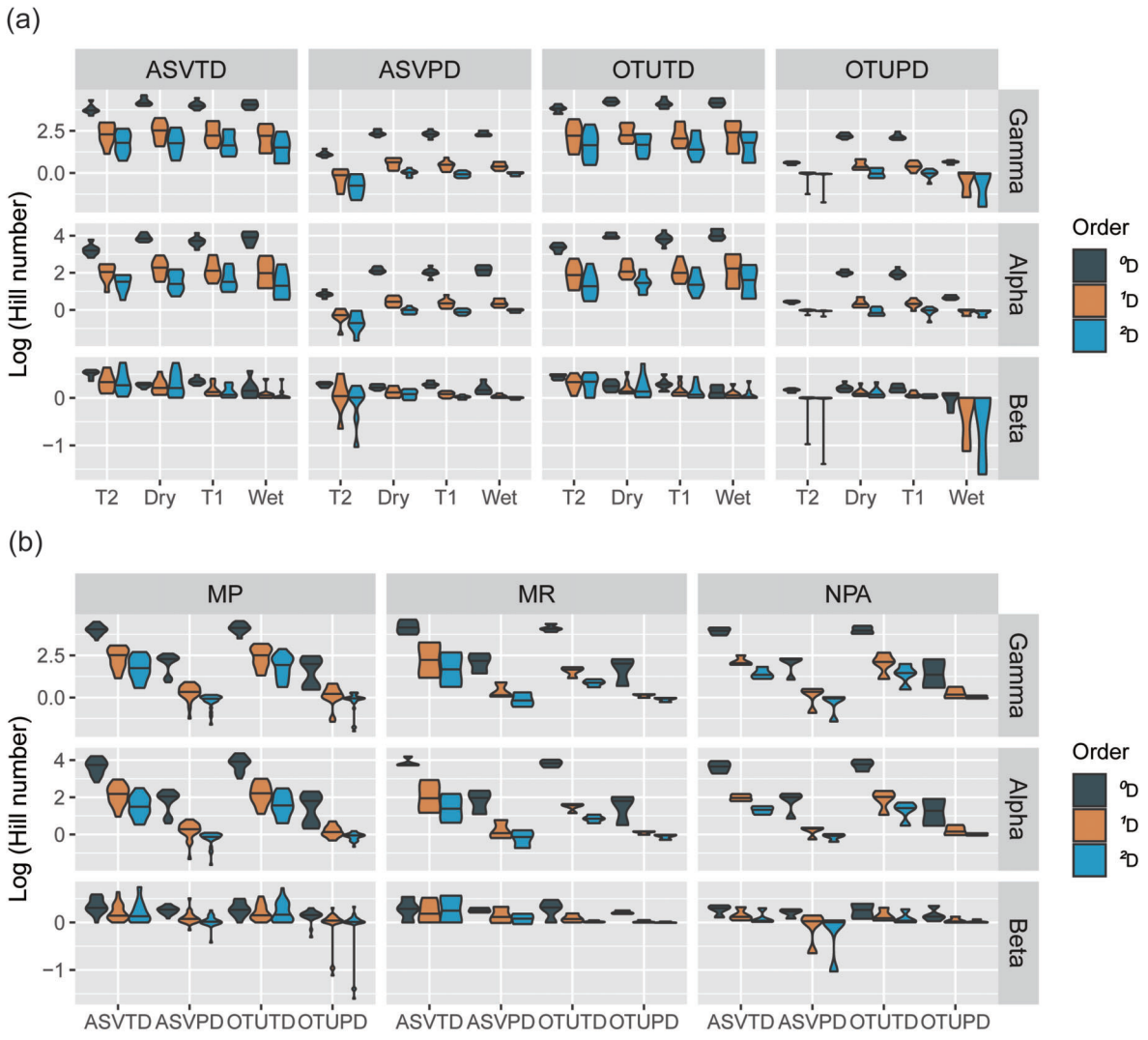

Figure 3 (a) Seasonal variation of gamma, alpha, and beta components of taxonomic and phylogenetic diversity inferred from RDP-based and Usearch-based datasets. (b) Fish diversity of different protection levels for ASVTD, ASVPD, OTUTD, and OTUPD. Different orders of Hill numbers were annotated with different colors. The horizontal line inside of the violin represents the median.

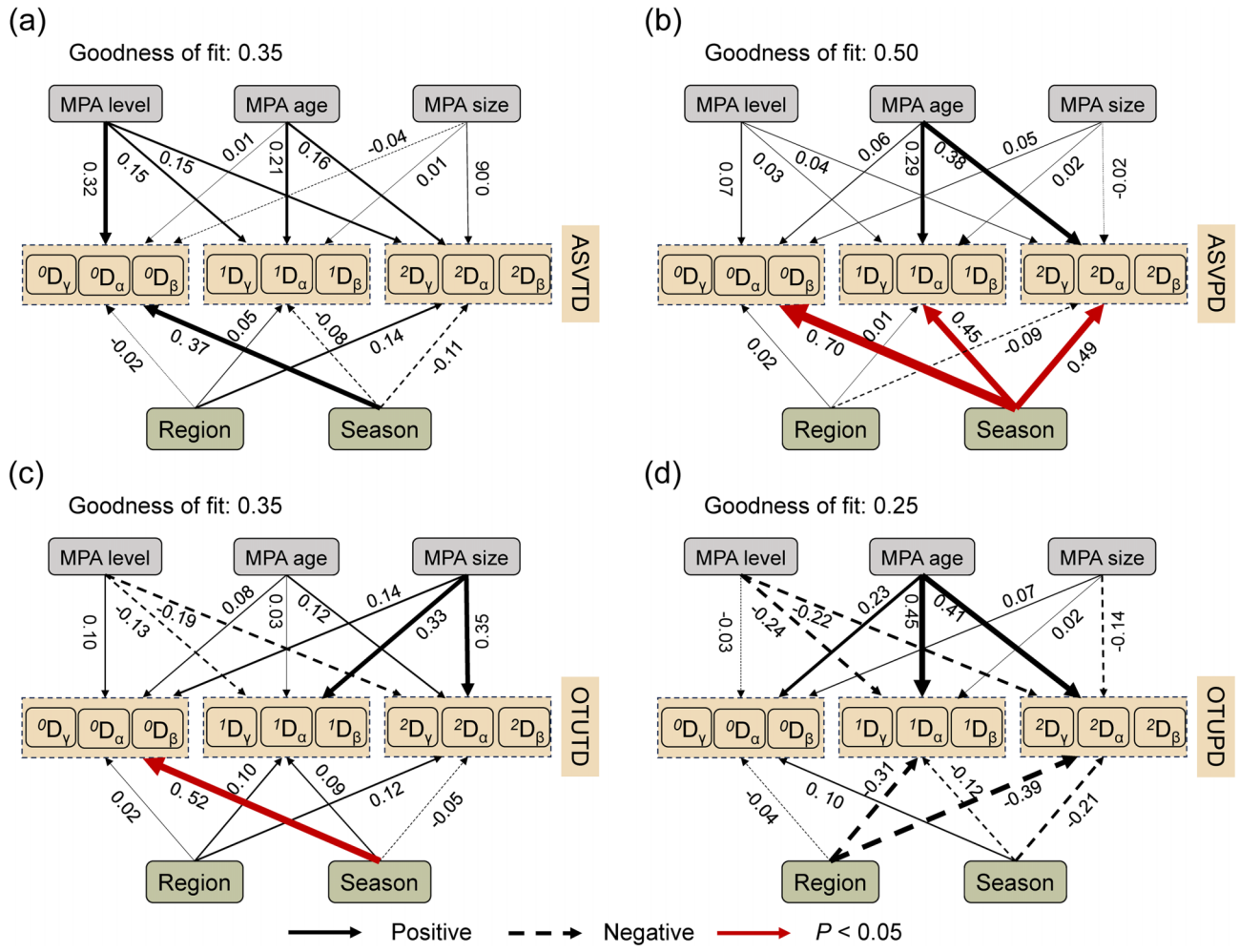

Figure 4 PLS-PM results showing the impact of protection activity and spatial-temporal factors on the variation of (a) ASVTD, (b) ASVPD, (c) OTUTD, and (d) OTUPD. The lines indicate direct pairwise effects with a solid line representing a positive effect, and a dashed line as negative. The width of the line corresponds with the level of influencing effect. The coefficient of each pairwise effect was annotated next to each line. The effects with $P < 0.05$ were highlighted with red.

However, the fish diversity obtained by contemporaneous traditional surveys during 2022–2023 was significantly lower than that detected by eDNA, potentially due to the sampling bias and low sampling efficiency of traditional sampling gears (Hallam et al. 2021). This kind of data deficiency can introduce biases in interpreting the true fish diversity and ecological status of the local ecosystem and further hinder the management of the protection activities (Rassweiler et al. 2012). Taken together, eDNA detected greater fish diversity than the traditional survey methods, while inconsistent overlapping results with historical traditional data highlighted the need for robust local reference databases. Future research should prioritize integrating multi-year eDNA monitoring with expanded local reference libraries to accurately assess biodiversity and validate methodological efficacy.

The reliability of downstream taxonomic annotation can largely influence the species-level information of eDNA data. In this study, we used the RDP Classifier for ASV assignment and the USEARCH algorithm for OTU assignment. The underlying principles of RDP and USEARCH are Bayesian probability and searching for high-identity hits, respectively, which have proven effective for taxonomic annotation in metabarcoding data. Comparisons between various assignment algorithms, including Least Common Ancestor (LCA)-based, Top-hit-based, Naive Bayesian-based, etc. indicated comparable accuracy in conjunction with parameter tuning (Hleap et al. 2021). Some studies suggested that LCA-based algorithms performed better in species assignment and should be adopted for metabarcoding and metagenomic data analysis (Cheng et al. 2025). In addition, the precise taxonomy assignment also depends on accurate reference databases (Bourret et al. 2023). Future studies should consider verifying different assignment algorithms and constructing reliable local reference databases to obtain precise species-level taxonomic annotation for eDNA data.

Rare fish groups influence interpretation of conservation effects on fish diversity

Notably, the presence of rare species in communities resulted in significant inter-group differences. Rare species also contributed to the seasonal variation of taxonomic and phylogenetic fish diversity detected by eDNA methods, as significantly lower diversity

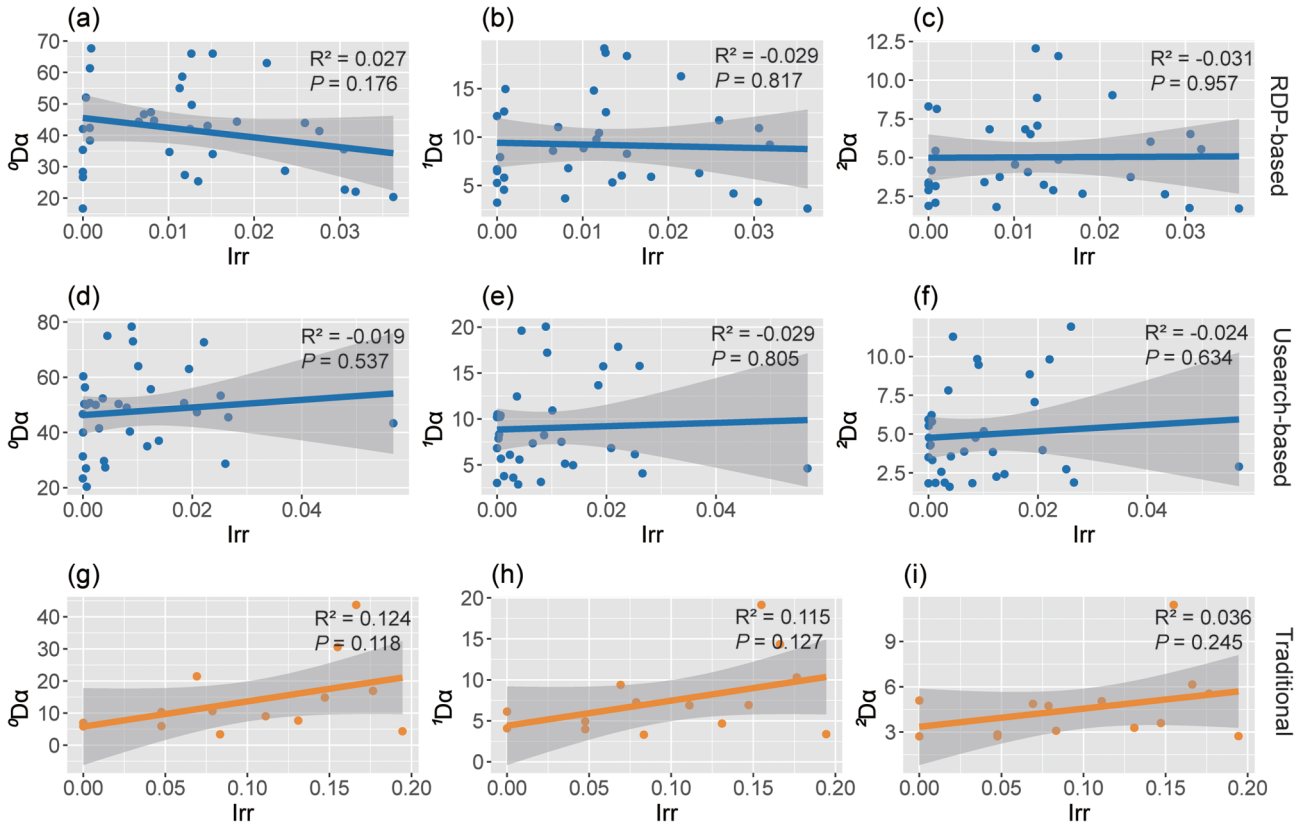

Figure 5 Linear regression model calculated based on different diversity orders of Hill number and Irr of all samples obtained by RDP-based (a–c), Usearch-based (d–f), and traditional survey (g–i). R^2 indicates adjusted R^2 of the model. All $P > 0.05$.

was detected in T2 when using lower diversity orders instead of higher orders. Consistently, Huang *et al*. (2016) demonstrated rare species were important in coral conservation as species rarity and richness differed largely among reef areas in the South China Sea, and the anthropogenic influence on coral could not be estimated only by species diversity. Moreover, rarity was verified to be an effective surrogate for marine fish biodiversity and further proved the potential use of it in site prioritization for biodiversity conservation planning (Astudillo-Scalia et al. 2019).

Under the global climate change scenario, rare species are more prone to population reductions and range shifts as they usually have already limited distributions. Top-down control from predators can also influence the population size of rare species and structure local community (Lynam et al. 2017). According to seven-decade worth of historical data of marine benthic biodiversity along the Swedish West coast, species richness has decreased, especially impacting rare species, and resulting in higher homogeneity of the community in the region (Obst et al. 2018). On a global scale, environmental conditions including temperature and precipitation can also shape the distribution gradients of species rarity for vertebrates (Albuquerque et al. 2019). We detected a positive, though not significant, relationship between protection activities and TD for Usearch-based data in local MPAs, and a corresponding negative correlation between protection activities and the diversity of abundant species. This indicates potential enhancement of local rarity under the current protection scheme and should be considered for future conservation planning.

Hill numbers showed high utility in marine conservation

Instead of using the more commonly implemented species richness and evenness indices (such as Shannon, Simpson, and Pielou's evenness), we applied Hill numbers with varying diversity orders, decomposed gamma diversity into beta and alpha diversities, and further proved the utility of Hill numbers in this study. By assigning different priority levels to rare and dominant species, we uncovered the important role of rare species in ecological interpretation. Understanding the compositional differences between local species assemblages (beta diversity) is essential for conservation planning (McKnight et al. 2007). In conservation schemes, high beta diversity may reveal the spatial scaling of diversity loss or high heterogeneity between local communities (Socolar et al. 2016). Depending on the decomposition of gamma diversity into alpha or beta diversities, we found that the beta diversity of MPAs was higher than in the NPA, although no significant difference was observed. Among all MPAs, no significant difference was found in any diversity order. Nevertheless, the gamma and alpha diversity of MPAs were also higher than in the NPA. This suggested high heterogeneity of local communities in the protected areas and the importance of existing MPAs in Hong Kong for biodiversity conservation.

Compared with Hill numbers, which highlighted rare species based on relative abundance, rarity index Irr inferred rare species according to their occurrences. Hill numbers utilize a continuous parameter to scale their importance, whereas Irr employs discrete

thresholds or predefined weights. As a result, Hill numbers provide an effective count of species, while Irr presents a proportion or ratio. In this study, we found high inconsistency between the two indicators, indicating that the rarity index may uncover hidden rarity hotspots with low fish diversity. Hill numbers are particularly useful for comparing diversity across varying levels of abundance, whereas Irr is more effective for assessing the role of rare species in ecosystem health and function. Future studies should consider using both bioindicators to monitor rare species for conservation as they provide complementary functions.

Conservation outcomes determined by protection activities

MPAs have proven to be effective in protecting specific fish taxa (Marraffini et al. 2024) and enhancing overall fish diversity (Lester et al. 2008). In this study, different eDNA pipelines showed positive correlations between fish diversity and protection duration, although the correlations were not significant. According to the RDP-based dataset, protection level showed positive relationships with taxonomic and phylogenetic diversity, while negative for Usearch-based data except ^{0}D. The contrary conclusions may be due to different fish assemblages generated by different pipelines. Previous studies have stated this controversy by examining the effect of MPAs on different organisms and found a positive effect on fisheries targeted species but negative for prey of these species (Giakoumi et al. 2017). In addition, the impact of protection level has long been debated, including positive effects for enhancing fish species richness (Sciberras et al. 2013, Loiseau et al. 2021) and biomass (Sciberras et al. 2015), and the reverse effect on the functional diversity of echinoderms (Rojas-Montiel et al. 2020). This reminds us of the importance of selecting suitable target species when determining the effect of certain MPAs. Regarding the MPA size, positive correlations were found for species richness detected by Usearch-based dataset, albeit not significant. Numerous studies have demonstrated the importance of protection size in conserving fishery resources (Friedlander et al. 2017) and spillover effects (Vandeperre et al. 2011). Our results supported the findings and further suggested an integrated scheme of no-take protection level, large protection size, and long duration to form an effective protection system.

The PLS-PM results further indicated that only season had a significant correlation with diversity indices, though the seasonal variation might be due to the sequencing bias introduced by different sequencing platforms we used for different seasons (T2: DNBSEQ; T1, Wet, and Dry: NovaSeq). However, Kim et al. (2021) demonstrated that these sequencing platforms showed comparable levels of sequencing quality, uniformity of coverage, % GC coverage, and variant accuracy. The previous study focusing only on the samples collected in transitional II season revealed a significant positive effect of MPAs in enhancing fish taxonomic and phylogenetic diversity (Zhao et al. 2024). Other studies of marine protection effectiveness have stressed the importance of protection level and connectivity between MPAs at local scales (Sève et al. 2023, Gill et al. 2024), while our results highlight the additional significance of local environmental conditions in structuring the fish community in the protected areas.

The evaluation of effectiveness of MPAs depends on multiple aspects, including the biodiversity, ecosystem health, and economic values provided (Grorud-Colvert et al. 2021). Understanding how individual species respond to protection is vital for conservation planning, as protection activities exert different impacts on different species. MPAs generally benefit top-predators and small herbivores the most, while they benefit mid-trophic level species more than rare large herbivores, based on a comprehensive metadata analysis in tropical regions (Sanchez et al. 2024). Therefore, we suggest addressing the importance of rare species groups to achieve a more integral understanding of biodiversity in MPAs to further facilitate conservation management and sustainable fishery development.

Conclusion

By calculating Hill numbers from eDNA metabarcoding data, we demonstrated the importance of accounting for rare species in evaluating conservation effects on fish communities. This strategy showed high feasibility in evaluating the fish biodiversity in MPAs, however, different bioinformatic clustering pipelines introduced significant differences in estimates of richness and evenness. Different biases were introduced with each pipeline in regard to interpreting environmental influences on fish diversity, but regardless of the chosen pipeline, the influence of protection of activity was less than that of the sampling season. In conclusion, we recommend carefully refining and comparing different clustering pipelines before adopting one to calculate multiple bioindicators, ensuring a comprehensive assessment of conservation effects on fish diversity. Future studies should consider improving and updating the local list of rare species including endemic and threatened ones to further enhance the taxonomy assignment and decision making of marine conservation strategy.

Acknowledgements

Any opinions, findings, conclusions, or recommendations expressed in this publication do not reflect the views of the AFCD nor the ITC. The authors thank two AFCD officers, Ms Yip Hung Yeung and Ms Louise Wai Hung Li, for their great support in this study. The authors also thank Ms Xiya Wang, Mr Daniel Chun Seng Hong, Mr Momo Long Chun Mou, and Mr Gabriel Yeung Lee for their technical support in this study.

Author contributions

R.Z. (Conceptualization [lead], Data curation [lead], Formal Analysis [lead], Methodology [lead], Software [lead], Visualization [lead], Writing – original draft [lead]), B.K. (Validation [supporting], Writing – review & editing [supporting]), Y.H.Y. (Resources [supporting], Writing – review & editing [supporting]), L.W.H.L. (Resources [supporting], Writing – review & editing [supporting]), Y.C. (Investigation [equal]), J.S. (Resources [supporting], Writing – review & editing [supporting]), J.Y. (Resources [supporting], Writing – review & editing [supporting]), X.Z. (Resources [supporting], Writing – review & editing [supporting]), K.M.Y.L. (Funding acquisition [equal], Resources [equal], Writing – review & editing [equal]), M.Y. (Funding acquisition [equal], Project administration [lead], Resources [equal], Supervision [equal], Writing – review & editing [supporting]).

Supplementary material

Supplementary material is available at *ICES Journal of Marine Science* online.

Conflicts of interest

The authors declare that they have no known competing interests.

Funding

This work was supported by two contract studies funded by the Agriculture, Fisheries and Conservation Department (AFCD), the Government of Hong Kong SAR, China (AFCD/SQ/104/21/C and AFCD/SQ/90/21/C to KMY Leung), Shenzhen Science and Technology Program (Project No. JCYJ20220530140813030 to M Yan), the Innovation and Technology Commission (ITC) of the Hong Kong SAR Government (9448002), and Jiangsu Provincial Science and Technology Program Special Fund (BZ2024053).

Data availability

Data is available on Sequence Read Archive NCBI with the BioProject numbers PRJNA1106654 and PRJNA1167905.

References

AFCD. Fisheries Resources Surveys at Marine Parks and the Marine Reserve. In, 2021.

AFCD. Fisheries Resources Surveys at Marine Parks and the Marine Reserve. In, 2025.

Alberdi A, Gilbert MTP. A guide to the application of Hill numbers to DNA-based diversity analyses. *Mol Ecol Resour* 2019;**19**:804–17. https://doi.org/10.1111/1755-0998.13014

Albuquerque F, Astudillo-Scalia Y, Loyola R *et al.* Towards an understanding of the drivers of broad-scale patterns of rarity-weighted richness for vertebrates. *Biodivers Conserv* 2019;**28**:3733–47. https://doi.org/10.1007/s10531-019-01847-z

Astudillo-Scalia Y, de Albuquerque FS. Evaluating the performance of rarity as a surrogate in site prioritization for biodiversity conservation. *Global Ecology and Conservation* 2019;**18**:e00639. https://doi.org/10.1016/j.gecco.2019.e00639

Beng KC, Corlett RT. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodivers Conserv* 2020;**29**:2089–121. https://doi.org/10.1007/s10531-020-01980-0

Blowes SA, Supp SR, Antao LH *et al.* The geography of biodiversity change in marine and terrestrial assemblages. *Science* 2019;**366**:339–45. https://doi.org/10.1126/science.aaw1620

Bohmann K, Evans A, Gilbert MTP *et al.* Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol Evol* 2014;**29**:358–67. https://doi.org/10.1016/j.tree.2014.04.003

Bolyen E, Rideout JR, Dillon MR *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;**37**:852–7. https://doi.org/10.1038/s41587-019-0209-9

Bourret A, Nozères C, Parent E *et al.* Maximizing the reliability and the number of species assignments in metabarcoding studies using a curated regional library and a public repository. *Metabarcoding and Metagenomics* 2023;**7**:e98539. https://doi.org/10.3897/mbmg.7.98539

Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* 2017;**11**:2639–43. https://doi.org/10.1038/ismej.2017.119

Callahan BJ, McMurdie PJ, Rosen MJ *et al.* DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3. https://doi.org/10.1038/nmeth.3869

Capurso G, Carroll B, Stewart KA. Transforming marine monitoring: using eDNA metabarcoding to improve the monitoring of the Mediterranean Marine Protected Areas network. *Mar Policy* 2023;**156**:105807. https://doi.org/10.1016/j.marpol.2023.105807

Chao A, Chiu C-H, Hsieh TC. Proposing a resolution to debates on diversity partitioning. *Ecology* 2012;**93**:2037–51. https://doi.org/10.1890/11-1817.1

Chao A, Chiu C-H, Hu K-H *et al.* Revisiting Alwyn H. Gentry's forest transect data: latitudinal beta diversity patterns are revealed using a statistical sampling-model-based approach. *Japanese J Stat Data Sci* 2023;**6**:861–84. https://doi.org/10.1007/s42081-023-00214-1

Chao A, Chiu C-H, Jost L. Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. *Annu Rev Ecol Evol Syst* 2014;**45**:297–324. https://doi.org/10.1146/annurev-ecolsys-120213-091540

Chao A, Hu K. The iNEXT.beta3D package: interpolation and extrapolation with beta diversity for three dimensions of biodiversity. In, 2023.

Chao A, Thorn S, Chiu C-H *et al.* Rarefaction and extrapolation with beta diversity under a framework of Hill numbers: the iNEXT.beta3D standardization. *Ecol Monogr* 2023;**93**:e1588. https://doi.org/10.1002/ecm.1588

Chaudhary C, Richardson AJ, Schoeman DS *et al.* Global warming is causing a more pronounced dip in marine species richness around the equator. *Proc Natl Acad Sci* 2021;**118**:e2015094118. https://doi.org/10.1073/pnas.2015094118

Cheng Y, Walsh D, Gauthier J *et al.* Improved COI taxonomic assignment workflow enhances detection of invertebrates in environmental DNA. *Limnol Oceanogr: Methods* 2025. https://doi.org/10.1002/lom3.70021

Chiarello M, McCauley M, Villéger S *et al.* Ranking the biases: the choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. *PLoS One* 2022;**17**:e0264443. https://doi.org/10.1371/journal.pone.0264443

Clausius E, Edgar GJ, Phillips GAC *et al.* Habitat and local factors influence fish biomass recovery in marine protected areas. *Proc R Soc B Biol Sci* 2025;**292**:20242708. https://doi.org/10.1098/rspb.2024.2708

Coddington JA, Agnarsson I, Miller JA *et al.* Undersampling bias: the null hypothesis for singleton species in tropical arthropod surveys. *J Anim Ecol* 2009;**78**:, 573–84. https://doi.org/10.1111/j.1365-2656.2009.01525.x

Corder GW, Foreman DI. *Nonparametric Statistics for Non-Statisticians*. In: Hoboken: John Wiley & Sons, Inc., 2011.

Davies KF, Margules CR, Lawrence JF. A synergistic effect puts rare, specialized species at greater risk of extinction. *Ecology* 2004;**85**:265–71. https://doi.org/10.1890/03-0110

Djurhuus A, Closek CJ, Kelly RP *et al.* Environmental DNA reveals seasonal shifts and potential interactions in a marine community. *Nat Commun* 2020;**11**:254. https://doi.org/10.1038/s41467-019-14105-1

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1. https://doi.org/10.1093/bioinformatics/btq461

Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. In: *Biorxiv*, 2016.

Frid O, Malamud S, Di Franco A *et al.* Marine protected areas' positive effect on fish biomass persists across the steep climatic gradient of the Mediterranean Sea. *J Appl Ecol* 2023;**60**:638–49. https://doi.org/10.1111/1365-2664.14352

Friedlander AM, Golbuu Y, Ballesteros E *et al.* Size, age, and habitat determine effectiveness of Palau's Marine Protected Areas. *PLoS One* 2017;**12**:e0174787. https://doi.org/10.1371/journal.pone.0174787

Gaston KJ. Abundances and range sizes: measuring rarity. In: *Rarity*, Dordrecht: Springer Netherlands, 1994, 22–56. https://doi.org/10.1007/978-94-011-0701-3

Giakoumi S, Scianna C, Plass-Johnson J *et al.* Ecological effects of full and partial protection in the crowded Mediterranean Sea: a regional meta-analysis. *Sci Rep* 2017;**7**:8940. https://doi.org/10.1038/s41598-017-08850-w

Gill DA, Lester SE, Free CM *et al.* A diverse portfolio of marine protected areas can better advance global conservation and equity. *Proc Natl Acad Sci* 2024;**121**:e2313205121. https://doi.org/10.1073/pnas.2313205121

Gold Z, Sprague J, Kushner DJ *et al.* eDNA metabarcoding as a biomonitoring tool for marine protected areas. *PLoS One* 2021;**16**:e0238557. https://doi.org/10.1371/journal.pone.0238557

Grorud-Colvert K, Sullivan-Stack J, Roberts C *et al.* The MPA Guide: a framework to achieve global goals for the ocean. *Science* 2021;**373**:eabf0861. https://doi.org/10.1126/science.abf0861

Hallam J, Clare EL, Jones JI *et al.* Biodiversity assessment across a dynamic riverine system: a comparison of eDNA metabarcoding versus traditional fish surveying methods. *Environmental DNA* 2021;**3**:1247–66. https://doi.org/10.1002/edn3.241

Hepper F, Synge H. The biological aspects of rare plant conservation. *Kew Bulletin* 1982;**37**:507. https://doi.org/10.2307/4110060

Hill MO. Diversity and evenness: a unifying notation and its consequences. *Ecology* 1973;**54**:427–32. https://doi.org/10.2307/1934352

Hleap JS, Littlefair JE, Steinke D *et al.* Assessment of current taxonomic assignment strategies for metabarcoding eukaryotes. *Mol Ecol Resour* 2021;**21**:2190–203. https://doi.org/10.1111/1755-0998.13407

Huang D, Hoeksema BW, Affendi YA *et al.* Conservation of reef corals in the South China Sea based on species and evolutionary diversity. *Biodivers Conserv* 2016;**25**:331–44. https://doi.org/10.1007/s10531-016-1052-7

Jerde CL, Mahon AR, Chadderton WL *et al.* Sight-unseen" detection of rare aquatic species using environmental DNA. *Conserv Lett* 2011;**4**:150–7. https://doi.org/10.1111/j.1755-263X.2010.00158.x

Kassambara A. rstatix: pipe-friendly framework for basic statistical tests. *R package version 0.7.3*, 2025. https://rpkgs.datanovia.com/rstatix/

Katoh K, Misawa K, Kuma K *et al.,* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;**30**:3059–66. https://doi.org/10.1093/nar/gkf436

Keck F, Blackman RC, Bossart R *et al.* Meta-analysis shows both congruence and complementarity of DNA and eDNA metabarcoding to traditional methods for biological community assessment. *Mol Ecol* 2022;**31**:1820–35. https://doi.org/10.1111/mec.16364

Kim H-M, Jeon S, Chung O *et al.* Comparative analysis of 7 short-read sequencing platforms using the Korean Reference Genome: MGI and Illumina sequencing benchmark for whole-genome sequencing. *GigaScience* 2021;**10**:giab014. https://doi.org/10.1093/gigascience/giab014

Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. *J Am Statist Assoc* 1952;**47**:583–621. https://doi.org/10.1080/01621459.1952.10483441

Kvalheim L, Stensrud E, Knutsen H *et al.* Integration of citizen science and eDNA reveals novel ecological insights for marine fish conservation. *Environmental DNA* 2024;**6**:e584. https://doi.org/10.1002/edn3.584

Larsen FW, Bladt J, Rahbek C. Improving the performance of indicator groups for the identification of important areas for species conservation. *Conserv Biol* 2007;**21**:731–40. https://doi.org/10.1111/j.1523-1739.2007.00658.x

Lawler JJ, White D, Sifneos JC *et al.* Rare species and the use of indicator groups for conservation planning. *Conserv Biol* 2003;**17**:875–82. https://doi.org/10.1046/j.1523-1739.2003.01638.x

Leitão RP, Zuanon J, Villéger S *et al.* Rare species contribute disproportionately to the functional structure of species assemblages. *Proc R Soc B Biol Sci* 2016;**283**:20160084. https://doi.org/10.1098/rspb.2016.0084

Leray M, Knowlton N, Machida RJ. MIDORI2: a collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences. *Environmental DNA* 2022;**4**:894–907. https://doi.org/10.1002/edn3.303

Leroy B, Petillon J, Gallon R *et al.* Improving occurrence-based rarity metrics in conservation studies by including multiple rarity cut-off points. *Insect Conservation and Diversity* 2012;**5**:159–68. https://doi.org/10.1111/j.1752-4598.2011.00148.x

Lester SE, Halpern BS. Biological responses in marine no-take reserves versus partially protected areas. *Mar Ecol: Prog Ser* 2008;**367**:49–56. https://doi.org/10.3354/meps07599

Lin L, Liu Y, Lin H *et al.* Considering species functional and phylogenetic rarity in the conservation of fish biodiversity. *Diversity and Distributions* 2024;**30**:e13804. https://doi.org/10.1111/ddi.13804

Liu JH, Hills P. Marine protected areas and local coastal conservation and management in Hong Kong. *Local Environment* 1997;**2**:275–97. https://doi.org/10.1080/13549839708725533

Loiseau N, Thuiller W, Stuart-Smith RD *et al.* Maximizing regional biodiversity requires a mosaic of protection levels. *PLoS Biol* 2021;**19**:e3001195. https://doi.org/10.1371/journal.pbio.3001195

Lynam CP, Llope M, Möllmann C *et al.* Interaction between top-down and bottom-up control in marine food webs. *Proc Natl Acad Sci* 2017;**114**:1952–7. https://doi.org/10.1073/pnas.1621037114

Mächler E, Walser JC, Altermatt F. Decision-making and best practices for taxonomy-free environmental DNA metabarcoding in biomonitoring using Hill numbers. *Mol Ecol* 2021;**30**:3326–39. https://doi.org/10.1111/mec.15725

Marathe A, Priyadarsanan DR, Krishnaswamy J *et al.* Gamma diversity and under-sampling together generate patterns in beta-diversity. *Sci Rep* 2021;**11**:21420. https://doi.org/10.1038/s41598-021-99830-8

Marraffini ML, Hamilton SL, Marin Jarrin JR *et al.* Evaluating the influence of marine protected areas on surf zone fish. *Conserv Biol* 2024;**38**:e14296. https://doi.org/10.1111/cobi.14296

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011;**17**:10–2. https://doi.org/10.14806/ej.17.1.200

McElroy ME, Dressler TL, Titcomb GC *et al.* Calibrating environmental DNA metabarcoding to conventional surveys for measuring fish species richness. *Front Ecol Evol* 2020;**8**:276. https://doi.org/10.3389/fevo.2020.00276

McKnight MW, White PS, McDonald RI *et al.* Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biol* 2007;**5**:e272. https://doi.org/10.1371/journal.pbio.0050272

Morton B, Morton J. *The sea shore ecology of Hong Kong*. Hong Kong: Hong Kong University Press, 1983.

Mouillot D, Bellwood DR, Baraloto C *et al.* Rare species support vulnerable functions in high-diversity ecosystems. *PLoS Biol* 2013;**11**:e1001569. https://doi.org/10.1371/journal.pbio.1001569

Ng TPT, Cheng MCF, Ho KKY *et al.* Hong Kong's rich marine biodiversity: the unseen wealth of South China's megalopolis. *Biodivers Conserv* 2017;**26**:23–36. https://doi.org/10.1007/s10531-016-1224-5

Obst M, Vicario S, Lundin K *et al.* Marine long-term biodiversity assessment suggests loss of rare species in the Skagerrak and Kattegat region. *Mar Biodivers* 2018;**48**:2165–76. https://doi.org/10.1007/s12526-017-0749-5

Pielou EC. The measurement of diversity in different types of biological collections. *J Theor Biol* 1966;**13**:131–44. https://doi.org/10.1016/0022-5193(66)90013-0

Polinski JM, Strand EL, O'Donnell TP *et al.* Environmental DNA documents ecosystem-wide biodiversity within the marine protected area Stellwagen Bank National Marine Sanctuary. *Environmental DNA* 2025;**7**:e70138. https://doi.org/10.1002/edn3.70138

Prendergast JR, Quinn RM, Lawton JH *et al.* Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* 1993;**365**:335–7. https://doi.org/10.1038/365335a0

Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;**26**:1641–50. https://doi.org/10.1093/molbev/msp077

Prodan A, Tremaroli V, Brolin H *et al.* Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS One* 2020;**15**:e0227434. https://doi.org/10.1371/journal.pone.0227434

Rao CR. Diversity and dissimilarity coefficients: a unified approach. *Theor Popul Biol* 1982;**21**:24–43. https://doi.org/10.1016/0040-5809(82)90004-1

Rassweiler A, Costello C, Siegel DA. Marine protected areas and the value of spatially optimized fishery management. *Proc Natl Acad Sci* 2012;**109**:11884–9. https://doi.org/10.1073/pnas.1116193109

Riaz T, Shehzad W, Viari A *et al.* ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Res* 2011;**39**:e145–. https://doi.org/10.1093/nar/gkr732

Roberts CM, O'Leary BC, McCauley DJ *et al.* Marine reserves can mitigate and promote adaptation to climate change. *Proc Natl Acad Sci* 2017;**114**:6167–75. https://doi.org/10.1073/pnas.1701262114

Rojas-Montiel B, Reyes-Bonilla H, Calderon-Aguilera LE *et al.* Echinoderm functional diversity does not correlate with the protection level of marine protected areas in the Mexican Pacific. *Biodivers Conserv* 2020;**29**:1871–96. https://doi.org/10.1007/s10531-020-01952-4

Roswell M, Dushoff J, Winfree R. A conceptual guide to measuring species diversity. *Oikos* 2021;**130**:321–38. https://doi.org/10.1111/oik.07202

Sala E, Mayorga J, Bradley D *et al.* Protecting the global ocean for biodiversity, food and climate. *Nature* 2021;**592**:397–402. https://doi.org/10.1038/s41586-021-03371-z

Sanchez G. PLSPM package. In, 2013

Sanchez L, Loiseau N, Edgar GJ *et al.* Rarity mediates species-specific responses of tropical reef fishes to protection. *Ecol Lett* 2024;**27**:e14418. https://doi.org/10.1111/ele.14418

Sciberras M, Jenkins SR, Kaiser MJ *et al.* Evaluating the biological effectiveness of fully and partially protected marine areas. *Environmental Evidence* 2013;**2**:4. https://doi.org/10.1186/2047-2382-2-4

Sciberras M, Jenkins SR, Mant R *et al.* Evaluating the relative conservation value of fully and partially protected marine areas. *Fish and Fisheries* 2015;**16**:58–77. https://doi.org/10.1111/faf.12044

Sève C, Belharet M, Melià P *et al.* Fisheries outcomes of marine protected area networks: levels of protection, connectivity, and time matter. *Conserv Lett* 2023;**16**:e12983. https://doi.org/10.1111/conl.12983

Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J* 1948;**27**:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Simpson EH. Measurement of Diversity. *Nature* 1949;**163**:688–. https://doi.org/10.1038/163688a0

Socolar JB, Gilroy JJ, Kunin WE *et al.* How should beta-diversity inform biodiversity conservation? *Trends Ecol Evol* 2016;**31**:67–80. https://doi.org/10.1016/j.tree.2015.11.005

Taberlet P, Coissac E, Hajibabaei M *et al.* Environmental DNA. *Mol Ecol* 2012;**21**:1789–93. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Takahashi M, Sacco M, Kestel JH *et al.* Aquatic environmental DNA: a review of the macro-organismal biomonitoring revolution. *Sci Total Environ* 2023;**873**:162322. https://doi.org/10.1016/j.scitotenv.2023.162322

Thomsen PF, Willerslev E. Environmental DNA—An emerging tool in conservation for monitoring past and present biodiversity.

Biol Conserv 2015;**183**:4–18. https://doi.org/10.1016/j.biocon.2014.11.019

UNEP-WCMC I. Marine Protected Planet. In, 2024.

Vandeperre F, Higgins RM, Sánchez-Meca J *et al.* Effects of no-take area size and age of marine protected areas on fisheries yields: a meta-analytical approach. *Fish and Fisheries* 2011;**12**:412–26. https://doi.org/10.1111/j.1467-2979.2010.00401.x

Wang Q, Garrity GM, Tiedje JM *et al.* Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;**73**:5261–7. https://doi.org/10.1128/AEM.00062-07

Wong ECK. Establishment of marine protected areas in Hong Kong. In. *The Marine biology of the South China Sea III : proceedings of the Third International Conference on the Marine Biology of the South China Sea, Hong Kong*. Hong Kong: Hong Kong University Press, 1996, 509–518.

WoRMS Editorial Board, *World Register of Marine Species (WoRMS)*. In: Belgium: VLIZ, 2026. https://doi.org/10.14284/170

Zhao R, Kang B, Chen Y *et al.* Multifaceted biological indicators reveal an effective conservation scheme for marine protected areas. *Ecol Indic* 2024;**166**:112389. https://doi.org/10.1016/j.ecolind.2024.112389

Zinger L, Bonin A, Alsos IG *et al.* DNA metabarcoding-Need for robust experimental designs to draw sound ecological conclusions. *Mol Ecol* 2019;**28**:1857–62. https://doi.org/10.1111/mec.15060

Zou KS, Chen JW, Ruan HT *et al.* eDNA metabarcoding as a promising conservation tool for monitoring fish diversity in a coastal wetland of the Pearl River Estuary compared to bottom trawling. *Sci Total Environ* 2020;**702**:134704. https://doi.org/10.1016/j.scitotenv.2019.134704

Handling editor: John Horne

Received: 1 October 2025. **Revised:** 27 February 2026. **Accepted:** 2 March 2026