Combining environmental DNA and visual surveys can inform conservation planning for coral reefs

Author list: Dominic Muenzel[1 2] (ORCID: 0000-0002-7158-7748)*, Alessia Bani[1 3 4] (ORCID: 0000-0001-7195-6446), Maarten De Brauwer[1 5] (ORCID: 0000-0002-9592-979X), Eleanor Stewart[2] (ORCID:) , Cilun Djakiman[1 6] (ORCID: 0000-0002-6535-9482), Halwi[7], Ray Purnama[6], Syafyuddin Yusuf[8] (ORCID: 0000-0003-1158-7938), Prakas Santoso[9], Frensly D Hukom[10 11], Matthew Struebig[2] (ORCID: 0000-0003-2058-8502), Jamaluddin Jompa[8] (ORCID: 0000-0001-9740-333X), Gino Limmon[6 11] (ORCID: 0000-0003-3030-7022), Alex Dumbrell[3] (ORCID: 0000-0001-6282-3043) *,**, Maria Beger[1 12] (ORCID: 0000-0003-1363-3571) *,**

[1] School of Biology, Faculty of Biological Sciences, University of Leeds, LS2 9JT, UK

[2] Durrell Institute of Conservation and Ecology, University of Kent, CT2 7NR, UK

[3] School of Life Sciences, University of Essex, CO4 3SQ, UK

[4] School of Built and Natural Environment, University of Derby, DE22 1GB, UK

[5] CSIRO Oceans & Atmosphere, Battery Point, Hobart, TAS 7004, Australia

[6] Maritime and Marine Science Center of Excellence, Pattimura University, Kampus Poka, Ambon, Indonesia

[7] Graduate School, Universitas Hasanuddin, Makassar, South Sulawesi, Indonesia

[8] Faculty of Marine Science and Fisheries, Universitas Hasanuddin, Makassar, South Sulawesi, Indonesia

[9] Department of Marine Science and Technology, IPB University, Bogor 16680, Indonesia

[10] Research Centre for Oceanography, BRIN, Jakarta 14430, Indonesia

[11] The Center For Collaborative Research On Aquatic Ecosystem in Eastern Indonesia, Pattimura University, Kampus Poka, Ambon 97234, Indonesia

[12] Centre for Biodiversity and Conservation Science, School of Biological Sciences, University of Queensland, Brisbane, QLD 4072, Australia

* Corresponding authors: Dominic Muenzel, Alex Dumbrell, Maria Beger

**Joint last authorship

Keywords: eDNA, spatial prioritisation, Wallacea, Indonesia, protected areas, marine spatial planning, high-resolution fisheries cost, coral reef biodiversity

Abstract

Environmental DNA (eDNA) metabarcoding has the potential to revolutionise conservation planning by providing spatially and taxonomically comprehensive data on biodiversity and ecosystem conditions, but its utility to inform the design of protected areas remains untested. Here, we quantify whether and how identifying conservation priority areas within coral reef ecosystems differs when biodiversity information is collected via eDNA analyses or traditional visual census records. We focus on 147 coral reefs in Indonesia's hyper-diverse Wallacea region and show large discrepancies in the allocation and spatial design of conservation priority areas when coral reef species were surveyed with underwater visual techniques (fishes, corals, algae) or eDNA metabarcoding (eukaryotes and metazoans). Specifically, incidental protection occurred for 55% of eDNA species when targets were set for species detected by visual surveys, and 71% vice versa. This finding is supported by generally low overlap in detection between visual census and eDNA methods at species level, with more overlap at higher taxonomic ranks. Incomplete taxonomic reference

databases for the highly diverse Wallacea reefs, and the complementary detection of species by the two methods, underscore the current need to combine different biodiversity data sources to maximise species representation in conservation planning.

Significance statement

Environmental DNA (eDNA) is emerging as a popular tool for biodiversity monitoring, as it allows organisms to be detected from environmental samples. We compare the use of eDNA and underwater visual census surveys in informing priority areas for coral reef biodiversity conservation in Indonesia's hyper-diverse Wallacea region. We find that different areas are identified when planning is informed by either method in isolation. The two survey methods show low overlap in species detection and identify some different taxonomic groups, suggesting that both methods should be deployed in a complementary assessment of biodiversity. Our analysis emphasizes the urgency for more collaborations in the region to address deficient taxonomic reference information, which hampers application of eDNA.

Main text

Introduction

Monitoring biodiversity via sampling environmental DNA (eDNA) has the potential to revolutionise conservation management (1–3). The DNA which organisms shed into their surroundings via skin cells, saliva, urine, faeces, or other pathways can be detected non-invasively in samples taken from the environment (4), the fragments of which are then matched to reference databases to obtain taxonomic identities (e.g., species). As extra-

cellular DNA is generally quick to break down *in situ* (ranging from hours to days in open water, but can last for thousands of years when preserved in sediments), the detection of DNA is interpreted as a spatiotemporally explicit signal of an organisms' presence (2, 5). Detection is not limited to single species, as samples can provide records of entire communities using metabarcoding, whereby universal primers bind to regions of genes that are conserved across taxa (6). Ongoing research efforts are addressing some of the limitations of eDNA metabarcoding (e.g., establishing universally accepted best practice protocols and improving reference databases) in order to generate highly comprehensive and spatially explicit data over wide geographic areas to help identify areas of high conservation priority (2).

eDNA analysis is particularly suited to study and manage hyper-diverse ecosystems, including coral reefs (6). Coral reefs host between one quarter and one third of marine biodiversity, yet traditional methods of surveying reef diversity often focus on a subset of large and well-studied taxonomic groups as surrogates (7). For example, underwater visual census is conducted by a group of experts whilst diving, typically for fishes (8). However, individual taxonomic expertise and detectability of species limit which taxa can be recorded, with a bias against certain groups, such as cryptic or shy species (9). As visual census is also time and resource-intensive, the geographic area covered tends to be limited, resulting in patchy data. Given the ongoing loss and degradation of coral reefs worldwide (10), eDNA metabarcoding surveys can help address the urgent need for detailed, extensive, and rapid biodiversity surveys to effectively allocate conservation resources (3, 6, 11).

Amongst the world's coral reefs, the Wallacea region in Indonesia and Timor-Leste stands out. Wallacea is renowned for its unparalleled levels of endemism and biodiversity, and is therefore a region of high conservation concern (Fig. 1) (12–14). Complex geological processes and island effects have led to widespread speciation and ecological diversification, with new species still being discovered (15). At the same time, economic development centred on natural resource exploitation is widespread in both marine and terrestrial realms (16, 17). Given the ecological importance of the region, eDNA could greatly facilitate the documentation and monitoring of Wallacea's unique and threatened biodiversity (12). In Indonesia, eDNA metabarcoding has experienced some success as a means to increase the number of fish species recorded and by revealing community structure patterns in coral reefs (11, 18), as well as for other taxa such as echinoderms, molluscs, and chordates (19).

Coral reef biodiversity data are a prerequisite in conservation planning to design protected areas. Spatially explicit data on species distributions, for example from visual census, can identify areas that will return the greatest conservation benefits if protected (20). Spatial conservation planning often utilises spatial prioritisation software that uses transparent, reproducible algorithms to balance ecological and socioeconomic objectives (21). Complementary sites that capture regional biodiversity at the lowest combined cost are identified as potential conservation areas. However, there is currently no consistent framework for translating eDNA data into spatial prioritisation plans (1). As eDNA metabarcoding can provide much higher information content than traditional survey techniques, it is unclear whether similar areas would be prioritised if the conservation objective was to protect regional biodiversity.

In this study, we compare conservation priority sites arising from visual census and eDNA metabarcoding biodiversity surveys of coral reefs in the Wallacea region. First, we explore similarities in the detection of taxonomic groups at reefs surveyed by both methods. Next, we develop a framework of how to use eDNA data in conservation planning. We model species prevalence data across space and design protected area systems that protect 30% target of each species' distribution across Wallacea, in line with global 30% by 2030 targets (22). We compare three separate objectives, where we identify priority areas for protecting only species recorded by visual census, eDNA, or both. For each objective, we determine the extent that species only recorded by a single method are captured. Given the exponential rise of eDNA monitoring and its untested potential to inform conservation, we benchmark the yet unrealised opportunity to use big eDNA datasets in conservation planning.

Results

Comparison of species detection

We surveyed 147 coral reef sites across the Wallacea archipelago in Indonesia between June 2019 and April 2021 (Fig. 1A). At 46 sites, we conducted visual census with experts counting fish, coral, and algae species. At 36 of the 46 sites, eDNA metabarcoding was conducted for water samples using universal primers that targeted the 18S and COI genes to capture most eukaryotes and metazoans respectively. The additional 101 sites across Wallacea were surveyed using eDNA methods only. Across sites surveyed by both methods, visual census and eDNA identified 993 and 2,073 unique species, respectively, of which 191 were identified by both (Table S1). eDNA metabarcoding data was clustered into Operational Taxonomic Units (OTUs), a method of grouping together DNA sequences from taxonomically

142 similar organisms to identify them to a given taxonomic level based on DNA sequence
143 similarity, which in this case (at 97% similarity) is approximately equivalent to a species.
144 Here, 17% of OTUs were matched to species in existing databases. eDNA methods generally
145 identified a much greater taxonomic breadth including fungi, protists, and animals which
146 were not visually recorded (Fig. 2A).
147
148 At all sites, co-detection by both methods was relatively low at species level, but increased
149 with higher taxonomic ranks (Fig. 2B, C). Species observed visually at a given site were only
150 detected by eDNA an average of 5% (±2% SD), 12% (±6% SD), and 6% (±11% SD) of the time
151 for coral, fish, and macroalgae (Fig. 2B). In part, this is caused by species whose
152 representative OTU could not be identified due to missing or incomplete records in
153 taxonomic databases (23). Detection of co-detected or shared taxonomic groups, for
154 example a species or genus which both survey methods were capable of detecting in at least
155 one of the 36 sites, was also low at species level, increasing with taxonomic rank (Fig. 2C).
156 Species which were detected by both eDNA and visually were co-detected at the same site
157 only 24% (±7 SD) of the time. Averaged across all sites, 53% (±12 SD) of shared species were
158 detected by visual census only, compared to 23% (±8 SD) by eDNA only.
159
160 Detection of fishes by either visual census or eDNA metabarcoding was related to their
161 position in the water column (Fig. S1). Pelagic and demersal species were detected more
162 often by eDNA than visually, whilst cnidarian-associated species were detected more often
163 by both methods than would be expected by chance ($\chi^2_{d.f.=16} = 515.39$, $p<0.001$). Detection
164 of a species by both methods at a given site did not guarantee co-detection in other sites
165 (Fig. S1A).

166

167 **Species distribution modelling and spatial prioritisation**

168 We built species distribution models (SDMs) for recorded species across 1 km^2 coral reef
169 habitat pixels (24) using environmental and human population covariates. SDMs were
170 successfully generated for 116 and 185 species for visual census and eDNA, respectively,
171 with an overlap of nine species (Table S2).

172

173 To assess how conservation priorities differ when planning with information from either
174 survey method, we used the SDMs in a spatial prioritisation analysis. We used a
175 conservation planning tool to design a cost-efficient system of protected areas that
176 accounted for existing levels of protection. We assumed implementation of strict no-take
177 protected areas that excluded fishing and calculated a unitless metric of fishery
178 displacement to account for spatial differences in the opportunity cost of protection. The
179 prioritisation analysis minimised this cost whilst meeting targets of 30% protection for
180 species identified by either visual census, eDNA, or both survey methods.

181

182 Overall, the agreement between solutions was considered 'slight' to 'moderate' when
183 comparing how often areas were selected across 100 repeat protected area systems (i.e.
184 selection frequency). Cohen's Kappa values (25), where a value of 1 indicates full agreement
185 and a value of 0 indicates no agreement beyond chance, were 0.12, 0.34, and 0.43 between
186 solutions of visually detected and eDNA, visually detected and both, and eDNA and both,
187 respectively. The fishery displacement cost of the top ten solutions with the lowest scores
188 was lowest if targets were set for visually detected species only (11,753 ±17 SD), 2% higher
189 for eDNA species only (11,988 ±4 SD), and 4% higher for both (12,277 ±9 SD). The area of

reef covered by these scenarios was 31.95% ±0.04 SD, 31.51% ±0.03 SD, and 32.73% ±0.04 SD. Meanwhile, the number of species for which targets were set were 116, 185, and 301, respectively, meaning that the cost and area increase did not directly scale with the increase in number of species.

Solutions were partially successful in protecting species even if targets were not set for them specifically (Fig. 3). If spatial prioritisation targets were set for visually detected species only, 55% of eDNA species also met or exceeded the target level of protection. The most frequent taxonomic classes of eDNA species for which targets were unmet were ray-finned fish (Actinopteri 11 spp), copepods (Hexanauplia 10 spp), brown algae (Phaeophyceae 8 spp), and gastropods (Gastropoda 8 spp). If spatial prioritisation targets were set for eDNA species only, 71% of visually detected species also met or exceeded the target level of protection. Visually detected species below the target level of protection belonged mainly to the fish families of wrasses (Labridae 7 spp), damselfishes and clownfishes (Pomacentridae 7 spp), and snappers (Lutjanidae 4 spp) and coral families of Merulinidae (3 spp) and Acroporidae (2 spp) (Table S3) which are all poorly represented in the DNA sequence taxonomy databases.

Spatial prioritisation identified some overlapping conservation priorities when targets were set either for visually detected or eDNA species (Fig. 1B). Areas including north of Muna Island in Southeast Sulawesi and the southern side of East Nusa Tenggara were higher priorities for the visual census scenario, whilst areas including south of Seram Island were higher priorities for the eDNA scenario.

Discussion

Here, we demonstrate how eDNA metabarcoding can complement traditional coral reef biodiversity survey techniques to inform protected area design in hyper-diverse marine regions such as Wallacea. We identified a greater overall taxonomic diversity across coral reef sites with eDNA targeting the COI and 18S genes compared to visual census, yet both methods identified unique taxonomic groups not detected by the other. By spatially extrapolating survey data with species distribution models and identifying priority areas for conservation, we found a low overlap in areas identified depending on whether conservation targets were set for species identified by visual census or eDNA. A greater proportion of visual census species were incidentally protected when targets were set for eDNA species than vice versa, at 71% compared to 55%. If only one survey method were to be used to inform priority areas, then eDNA provides a more comprehensive choice for greater overall protection of biodiversity. However, genera important for fisheries in Indonesia, such as Lutjanus and Scarus (26), were inadequately protected if conservation priorities were set by eDNA records alone (Fig. 3, Table S3). Meanwhile groups such as gastropods recorded in eDNA surveys were inadequately protected if priorities were set by visual census records only (Table S3). Taken together, the difference in identified taxa, low probability of co-detection, and moderate incidental protection suggest that both visual census and eDNA survey data should be used in combination to inform protected area design.

Our spatial prioritisation scenarios had the objective to protect 30% of the distribution of each identified species, assuming that it is desirable to protect the entire breadth of biodiversity (27). This is more conservative than the 30% by 2030 target which calls for 30%

238 of terrestrial and marine areas to be protected (22), rather than 30% of all species
239 distributions. Despite this, our solutions had a similar spatial coverage by selecting between
240 32-33% of the available reef areas. There is value in protecting wider biodiversity, as species
241 interactions and 'hidden' diversity (e.g., microbial diversity) sustain ecosystem resilience,
242 functioning, and integrity (28, 29). As visual census and eDNA detected different taxonomic
243 groups, the greatest protection of regional biodiversity would be achieved by combining the
244 two datasets to set conservation targets. This approach could also protect more varied
245 ecological niches and a greater functional trait space, the phenotypic space occupied by a
246 set of species that determines their effect on processes and responses to environmental
247 factors (30), since different survey techniques may be biased towards different functional
248 groups. Setting conservation targets for species surveyed by both techniques only increased
249 the cost of protected area solutions by 4%, suggesting that protected areas need not be
250 substantially more expensive to protect greater levels of biodiversity.
251
252 If sufficient information about species' ecologies and conservation status are available,
253 targets in conservation planning may be modified accordingly. Not all taxa identified in
254 eDNA samples are equally important to protect. Different taxa contribute to ecosystem
255 functioning in different ways. For example, keystone species are important as they can have
256 a disproportionately large role with many downstream effects (31), whereas other species
257 may be less important if there is high functional redundancy and multiple species fulfil
258 similar functions (32). Prevalence and extinction risk will also determine the importance of
259 protecting a species. Given the wealth of information eDNA metabarcoding generates,
260 managers must consider which groups are important and why, as well as what they indicate.

Some taxa may also be indicators of areas undesirable for protection, such as certain bacteria found in sewage pollution (33).

Apart from conserving biodiversity, marine protected areas are often designed with additional goals. These include supporting sustainable fisheries by providing spawning and nursery grounds, granting exclusive access rights to local users, generating income from tourism, restricting extractive activities to allow ecosystem restoration, and enhancing ecosystem services such as carbon sequestration and erosion control (14, 34, 35). Although goals may complement each other, they may also come into conflict. For example, criteria for long-term population persistence within protected areas often conflict with criteria for fishery spillover (i.e., the movement of individuals from protected to fished areas) (36). Indonesian marine protected areas have a dual purpose of conserving biodiversity and supporting fisheries, with not all zoning categories being strictly no-take as our cost calculation assumes. By minimising fishery displacement cost, our analysis reflects this need to mitigate conflict. Additionally, our approach of protecting overall biodiversity may also indirectly support fisheries as biodiversity is amongst the strongest predictors of reef fish biomass (37), assuming sufficient spillover.

eDNA sampling and SCUBA-based visual surveys differ in some major respects which have implications for their use. Costs for eDNA sampling can be lower than for visual surveys (38), although this greatly depends on available infrastructure and equipment for either biomonitoring method. Visual survey costs remain relatively unchanging across time, but eDNA costs are expected to decrease as more commercial laboratories which process collected samples are established (38). SCUBA visual surveys are more constrained by

weather, ocean conditions, and personnel, and the remoteness of many of the world's coral reef may favour methods requiring fewer equipment and personnel (39). eDNA metabarcoding samples have the advantage that they can be preserved, archived, and reanalysed in the future when methods and databases are updated, allowing data to be used dynamically. Although eDNA captures a large amount of biodiversity which can reveal large scale ecological patterns, much of this diversity is for unnamed species until databases improve. In contrast, visual survey data is generally species or genus-specific but less taxonomically comprehensive.

One obstacle we encountered in using survey data to identify priority conservation areas was that single species distribution models were successful for few of the recorded species (12% and 9% of visually detected and eDNA species, respectively). Rare or threatened species have low prevalence, which can result in sample sizes too small to build reliable statistical models. Apart from increasing sampling effort, one solution could be to use joint species distribution models (40). These methods model species responses to both the environment and to other species, recognising effects of interspecific interactions such as competition, predation, and facilitation. Using such community models can improve predictions of rare species compared to single species models (41), making them suitable to analyse the big community data that eDNA generates (42).

Given the low co-detection of shared taxonomic groups by visual census and eDNA, some thought should be given as to why this is the case and how detection could be improved. Compared to terrestrial sampling, the marine environment poses additional challenges to dispersion and degradation of eDNA. Abiotic factors such as temperature, salinity, and

ultraviolet radiation lead to eDNA breakdown (5). Differences in the time of day or strength of wind, currents, and tides during our surveys may explain some of the variability in co-detection. eDNA dispersion in the sea can be as short as 30 m (43), or up to several kilometres (44), depending on local conditions. Collecting eDNA samples at different depths or across a grid may improve co-detection with visual census if vertical or horizontal water mixing is limited. Additionally, less abundant species may be more difficult to detect at populous sites, as co-amplification of DNA from many taxa lowers the sampling depth for rare species (45). Co-detection may therefore be higher in less diverse systems, where DNA fragments belong to comparatively fewer unique species. Additional research into the ecology of eDNA in tropical marine environments will be necessary to refine future study designs and sampling efforts.

As only 17% of the OTUs were matched to species, our study echoes the need for more complete reference databases of marine fauna and flora (46, 47). Expanded barcoding efforts are particularly needed in areas such as the Coral Triangle, where comparatively little research focus is given despite high levels of biodiversity and human resource dependence (48, 49). Barcoding corals can be challenging as their mitochondrial DNA, where COI is encoded, is highly conserved (50). Solving this challenge may require genome-wide sequencing to develop nuclear markers of variable genomic regions which can be used in eDNA metabarcoding (51). In the case of fish and corals, genomic introgression from hybridisation between species can impede species assignment (52). Developing custom genetic databases of reference species to supplement genetic repositories and using taxon-specific primers will greatly improve species assignment.

eDNA will play a growing role in future coral reef conservation efforts to provide taxonomically comprehensive data, including for previously understudied taxa. This study explores how these data can be used in conservation planning to protect greater taxonomic space. Corroborating other research comparing eDNA and other techniques (47, 53, 54), we show that eDNA metabarcoding can complement traditional survey techniques to give a more comprehensive picture of biodiversity and its distribution across space.

Methods

Field surveys

Underwater visual census surveys of coral, fishes, and macroalgae were carried out by a team of four taxonomists on SCUBA at 8m depth, covering four replicate 50m belt transects at each site. For each transect, two observers identified, counted, and sized non-cryptic fish at species level across a 50m x 5m belt and laid out a 50m tape. This transect was followed by one observer counting algae to species or genus level across a 2m x 30m belt and one observer counting coral colonies to species level across a 0.5m x 20m belt.

eDNA sampling was carried out by collecting replicate 1L seawater samples on SCUBA just above the reef at 8m depth. Where sampling overlapped with visual census, three samples were collected at the beginning, middle and end of each 50m transect at the same time as visual surveys, creating a total of 12 samples collected along 200m of reef per site. For other sites, 6 replicate samples were collected by swimming a similar approximate distance. Bottles for collection were first sterilised for 30 minutes with chlorine 12% (780 mg of NaDCC) and rinsed with surface seawater. eDNA samples were then filtered using Merck Sterivex 0.22 µm (Merck, United Kingdom) and filled with 2 ml of Longmire for preservation

(55). As controls, blank samples consisting of PCR water (ThermoFisher, United Kingdom) were also filtered in the same conditions during the survey period.

eDNA analysis

Our metabarcoding followed standard approaches (56), and assessed community compositions with the two universal primers, targeting the 18S gene for eukaryotes and the COI gene for metazoans. Eukaryotes are organisms whose cells contain a nucleus and mitochondria and include animals, plants, fungi, and protists. Metazoans are a subset of eukaryotes and refer to multicellular animals with differentiated cells. The primers overlap with the taxonomic groups of fishes, corals, and algae surveyed by visual census, but include additional groups such as arthropods, molluscs, sponges, and fungi. We extracted Sterivex using the Qiagen DNeasy Power Water Sterivex kit (Qiagen, Germany). Longmire, removed from the Sterivex, was centrifuged for 40 minutes at 6000 g (55). We discarded the supernatant and resuspended the pellet in the first solution of the DNeasy PowerWater Sterivex kit. The rest of the extraction was performed following the user manual and extracted DNA was stored at -20°C until library preparation. We extracted field controls in the same way as the samples and included additional extraction blanks in the extraction procedure (57). Library preparations followed the standard Illumina protocol of two stage PCR and index using dual indexing (Environmental DNA Sequencing | Biomonitoring using eDNA (illumina.com). We used a well-established protocol for data cleaning steps that have been successfully applied in multiple environments (58) (full details in Supporting Information).

Comparison of species detection

In the 36 sites surveyed using both visual census and eDNA, we explored how often species or higher taxonomic classifications were detected by both methods. This method provided an estimate on the reliability of eDNA detection, based on how often species observed visually are present or absent from eDNA samples. We also matched fish species with their position in the water column based on functional trait data (59) to investigate whether detection by either method was influenced by where the fish are generally found (60).

Species distribution modelling

We used species distribution models (SDMs) to relate observed visual census and eDNA records from surveyed sites to environmental and human population covariates to predict probabilities of observation in non-surveyed areas (61). We created ensemble SDMs that combined the different models Random Forest, Generalised Linear Model, and Generalised Additive Model, weighted by the performance of each model, to improve predictive accuracy by reducing uncertainty caused by differences amongst modelling techniques (62).

Species counts from UVC and eDNA presence-absence data were modelled assuming a Poisson and binomial distribution, respectively. For eDNA data, SDMs were only built for Operational Taxonomic Units that were matched to species, as multiple unassigned Operational Taxonomic Units could belong to the same species. Covariates were selected as those known to drive coral and reef fish distributions from a list of candidates: sea surface temperature, sea surface temperature anomaly, pH, salinity, chlorophyll a, dissolved oxygen, photosynthetically available radiation, wave exposure, and human pressure (Table S4) (63, 64). Models with the greatest explanatory power were selected from a set of preliminary models consisting of all possible covariate combinations with the restriction of

having no more than one predictor variable per 10 datapoints to avoid model overfitting (65). Variables with a variance inflation factor >10 were removed to avoid collinearity (66).

We cross-validated the predictive accuracy of models by dividing data into 80% training and 20% testing splits a total of 1000 times. We evaluated count models using Root Mean Square Error (RMSE) and Pearson's correlation, where only models with an average RMSE smaller than half the range of the data and an average correlation >0.25 were retained in the ensembles (67). We evaluated presence-absence models using the Area Under the Curve (AUC), where only models with AUC >0.7 were retained in the ensembles (68). The final model ensembles were used to predict species distributions across 40,922 $1km^2$ coral reef pixels in the Wallacea region, building off the resolution of a previously published fine-scale sea surface temperature dataset (24). We selected thresholds for classifying probabilities into binary presences and absences to give the maximum value of Kappa, a measure which compares model predictive accuracy to accuracy expected to occur by chance (69). All analyses outside bioinformatics were run in R v4.2.0 (70) using the *randomForest* v4.7-1.1 (71), *mgcv* v1.8-39 (72), and *base* packages.

Spatial prioritisation scenarios

We identified priority areas for conservation based on the extrapolated visual census and eDNA data with Marxan (73). Marxan is a spatial prioritisation tool that selects management areas (termed 'planning units') to meet user-specified conservation targets at least cost. We used $1km^2$ reef pixels (24) as planning units and set a target to represent 30% of pixels containing each species using three different scenarios. In scenario 1, targets were set for species recorded by visual census only. In scenario 2, targets were set for species recorded

by eDNA only. In scenario 3, targets were set for species recorded by both survey techniques. Planning units that are frequently selected across 100 iterations are considered important areas for conservation. Planning units occurring within existing Marine Protected Areas of IUCN category I and II (highly protected) were locked in with a protected status, meaning they are forcibly kept in solutions.

The cost of each planning unit was quantified in terms of an opportunity cost based on three proxies related to displaced access by different types of fishers.

$$Cost = \frac{F + DistS + G}{3} * A \quad \text{(Eqn. 1)}$$

F is a binary value indicating whether the area is adjacent to a village that has artisanal fishing as the main livelihood, signifying local fishers with limited boat access (fishing within 1km of village/ coast). Livelihood information was extracted from the 2018 Potensi Desa census from the Indonesian Government (74). *DistS* is the overwater distance to nearest coastal village ('desa' category in (74)) to signify fishing off small boats with engines (fishing within tens of kms off village/coast), where the distance was scaled by

$$DistS = 1 - \frac{\log(distance)}{\max(\log(distance))} \quad \text{(Eqn. 2)}$$

G is the gravity or human impact metric calculated as

$$G = \frac{P}{T^2} \quad \text{(Eqn. 3)}$$

where *P* is the population of the nearest city from WorldPop data in 2020 (75) and *T* is the travel time over sea to the nearest city with a constant boat speed, signifying the sea-scape level fishing pressure exerted by large population centres on their surrounding reef (fishing to feed a city applied at 100's of km, (76)). *A* is the habitat area in each planning unit based on Allen Coral Atlas maps excluding the category of rubble (77).

We assessed how well solutions from our three spatial prioritisation scenario selected similar planning units and captured the distributions of species. First, we determined the degree of overlap in priority areas using the Cohen's Kappa statistic on the selection frequency of planning units (25). Second, we evaluated how well setting targets for species detected by either visual census or eDNA could incidentally protect species that were only detected by the method for which targets were not yet set.

Acknowledgments and funding sources

The UK Natural Environment Research Council (NERC) grants, NE/S006931/1 (Maria Beger), NE/S006958/1 (Alex Dumbrell), NE/S007067/1 (Matthew Struebig), D.M. was funded by the University of Leeds School of Biology Doctoral Training Studentship, National Geographic Early Career Grant EC-66640R-20 (Maarten De Brauwer), Funding from Ristekdikti (Indonesia) number7/AMD/E1/KP.PTNBH/2020 (Jamaluddin Jompa). Portions of the paper were developed from the PhD thesis of D.M.

References

1. A. Bani, *et al.*, *Informing marine spatial planning decisions with environmental DNA*, 1st Ed. (Elsevier Ltd., 2020).
2. K. M. Ruppert, R. J. Kline, M. S. Rahman, Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* **17**, e00547 (2019).
3. K. Bohmann, *et al.*, Environmental DNA for wildlife biology and biodiversity

monitoring. *Trends Ecol. Evol.* **29**, 358–367 (2014).

4. G. F. Ficetola, C. Miaud, F. Pompanon, P. Taberlet, Species detection using environmental DNA from water samples. *Biol. Lett.* **4**, 423–425 (2008).

5. J. B. Harrison, J. M. Sunday, S. M. Rogers, Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proc. R. Soc. B Biol. Sci.* **286** (2019).

6. M. Stat, *et al.*, Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep.* **7**, 1–11 (2017).

7. L. Plaisance, M. J. Caley, R. E. Brainard, N. Knowlton, The diversity of coral reefs: What are we missing? *PLoS One* **6** (2011).

8. Z. R. Caldwell, B. J. Zgliczynski, G. J. Williams, S. A. Sandin, Reef fish survey techniques: Assessing the potential for standardizing methodologies. *PLoS One* **11**, 1–14 (2016).

9. G. J. Edgar, N. S. Barrett, A. J. Morton, Biases associated with the use of underwater visual census techniques to quantify the density and size-structure of fish populations. *J. Exp. Mar. Bio. Ecol.* **308**, 269–290 (2004).

10. T. D. Eddy, *et al.*, Global decline in capacity of coral reefs to provide ecosystem services. *One Earth* **4**, 1278–1285 (2021).

11. O. N. Marwayana, Z. Gold, C. P. Meyer, P. H. Barber, Environmental DNA in a global biodiversity hotspot: Lessons from coral reef fish diversity across the Indonesian archipelago. *Environ. DNA* **4**, 222–238 (2022).

12. M. J. Struebig, *et al.*, Safeguarding imperiled biodiversity and evolutionary processes in the Wallacea center of endemism. *Bioscience* **72**, 1118–1130 (2022).

13. G. V. Limmon, *et al.*, A cause for hope: largely intact coral-reef communities with high reef-fish biomass in a remote Indonesian island group. *Mar. Freshw. Res.* (2023)

https:/doi.org/10.1071/mf22075.

14. M. Beger, *et al.*, Integrating regional conservation priorities for multiple objectives into national policy. *Nat. Commun.* **6**, 1–8 (2015).

15. F. E. Rheindt, *et al.*, A lost world in Wallacea: Description of a montane archipelagic avifauna. *Science (80-.).* **367**, 167–170 (2020).

16. M. Voigt, *et al.*, Emerging threats from deforestation and forest fragmentation in the Wallacea centre of endemism. *Environ. Res. Lett.* **16** (2021).

17. S. Yusuf, A. M. Moore, Hunting in the seas: Population status and community perspectives on giant clams (Tridacnidae) and Napoleon wrasse (Cheilinus undulatus), endangered marine taxa of the Wallacea Region, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **473** (2020).

18. E. R. E. Gelis, *et al.*, Environmental biomonitoring of reef fish community structure with eDNA metabarcoding in the Coral Triangle. *Environ. Biol. Fishes* **104**, 887–903 (2021).

19. H. Madduppa, *et al.*, eDNA metabarcoding illuminates species diversity and composition of three phyla (chordata, mollusca and echinodermata) across Indonesian coral reefs. *Biodivers. Conserv.* **30**, 3087–3114 (2021).

20. A. D. Olds, *et al.*, Incorporating surrogate species and seascape connectivity to improve marine conservation outcomes. *Conserv. Biol.* **28**, 982–991 (2014).

21. C. R. Margules, R. L. Pressey, Systematic conservation planning. *Nature* **405**, 243–253 (2000).

22. Convention on Biological Diversity, “Kunming-Montreal Global Biodiversity Framework” (2022).

23. V. Marques, *et al.*, GAPeDNA: Assessing and mapping global species gaps in genetic

databases for eDNA metabarcoding. *Divers. Distrib.* **27**, 1880–1892 (2021).

24. A. M. Dixon, P. M. Forster, S. F. Heron, A. M. K. Stoner, M. Beger, Future loss of local-scale thermal refugia in coral reef ecosystems. *PLOS Clim.* **1**, e0000004 (2022).

25. K. A. Wilson, M. I. Westphal, H. P. Possingham, J. Elith, Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biol. Conserv.* **122**, 99–112 (2005).

26. A. T. Humphries, *et al.*, Catch composition and selectivity of fishing gears in a multi-species Indonesian coral reef fishery. *Front. Mar. Sci.* **6**, 1–9 (2019).

27. L. J. Pollock, *et al.*, Protecting Biodiversity (in All Its Complexity): New Models and Methods. *Trends Ecol. Evol.* **35**, 1119–1128 (2020).

28. E. Harvey, I. Gounand, C. L. Ward, F. Altermatt, Bridging ecology and conservation: from ecological networks to ecosystem function. *J. Appl. Ecol.* **54**, 371–379 (2017).

29. M. Delgado-Baquerizo, *et al.*, Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat. Commun.* **7**, 1–8 (2016).

30. G. Aglieri, *et al.*, Environmental DNA effectively captures functional diversity of coastal fish communities. *Mol. Ecol.* **30**, 3127–3139 (2021).

31. R. D. Davic, Linking keystone species and functional groups: A new operational definition of the keystone species concept. *Ecol. Soc.* **7** (2003).

32. D. Mouillot, *et al.*, Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13757–13762 (2014).

33. S. L. Wear, R. V. Thurber, Sewage pollution: Mitigation is key for coral reef stewardship. *Ann. N. Y. Acad. Sci.* **1355**, 15–30 (2015).

34. C. M. Roberts, *et al.*, Marine reserves can mitigate and promote adaptation to climate

544 change. *PNAS* **114**, 6167–6175 (2017).
545 35. D. Muenzel, *et al.*, Integrating larval connectivity into the marine conservation
546 decision-making process across spatial scales. *Conserv. Biol.* **37**, 1–11 (2023).
547 36. I. Chollett, *et al.*, A genuine win-win: resolving the “conserve or catch” conflict in
548 marine reserve network design. *Conserv. Lett.* **10**, 555–563 (2017).
549 37. J. E. Duffy, J. S. Lefcheck, R. D. Stuart-Smith, S. A. Navarrete, G. J. Edgar, Biodiversity
550 enhances reef fish biomass and resistance to climate change. *Proc. Natl. Acad. Sci. U.*
551 *S. A.* **113**, 6230–6235 (2016).
552 38. Z. Gold, J. Sprague, D. J. Kushner, E. Z. Marin, P. H. Barber, eDNA metabarcoding as a
553 biomonitoring tool for marine protected areas. *PLoS One* **16**, 1–19 (2021).
554 39. K. M. West, *et al.*, The applicability of eDNA metabarcoding approaches for sessile
555 benthic surveying in the Kimberley region, north-western Australia. *Environ. DNA* **4**,
556 34–49 (2022).
557 40. O. Ovaskainen, *et al.*, How to make more out of community data? A conceptual
558 framework and its implementation as models and software. *Ecol. Lett.* **20**, 561–576
559 (2017).
560 41. C. Zhang, Y. Chen, B. Xu, Y. Xue, Y. Ren, Improving prediction of rare species’
561 distribution from community data. *Sci. Rep.* **10**, 1–9 (2020).
562 42. M. Pichler, F. Hartig, A new joint species distribution model for faster and more
563 accurate inference of species associations from big community data. *Methods Ecol.*
564 *Evol.* **12**, 2159–2173 (2021).
565 43. H. Murakami, *et al.*, Dispersion and degradation of environmental DNA from caged
566 fish in a marine environment. *Fish. Sci.* **85**, 327–337 (2019).
567 44. G. J. Jeunen, *et al.*, Environmental DNA (eDNA) metabarcoding reveals strong

discrimination among diverse marine habitats connected by water movement. *Mol. Ecol. Resour.* **19**, 426–438 (2019).

45. B. N. Nguyen, *et al.*, Environmental DNA survey captures patterns of fish and invertebrate diversity across a tropical seascape. *Sci. Rep.* **10**, 1–14 (2020).

46. J. D. DiBattista, *et al.*, Assessing the utility of eDNA as a tool to survey reef-fish communities in the Red Sea. *Coral Reefs* **36**, 1245–1252 (2017).

47. A. Polanco Fernández, *et al.*, Comparing environmental DNA metabarcoding and underwater visual census to monitor tropical reef fishes. *Environ. DNA* **3**, 142–156 (2021).

48. P. H. Barber, *et al.*, Advancing biodiversity research in developing countries: The need for changing paradigms. *Bull. Mar. Sci.* **90**, 187–210 (2014).

49. G. V. Limmon, F. Rijoly, O. T. S. Ongkers, S. R. Loupatty, J. A. Pattikawa, Reef fish in the southern coastal waters of ambon Island, Maluku province, Indonesia. *AACL Bioflux* **10**, 234–240 (2017).

50. J. Neigel, A. Domingo, J. Stake, DNA barcoding as a tool for coral reef conservation. *Coral Reefs* **26**, 487–499 (2007).

51. C. I. M. Adams, *et al.*, Beyond biodiversity: Can environmental DNA (eDNA) cut it as a population genetics tool? *Genes (Basel).* **10** (2019).

52. D. W. Cruz, P. L. Harrison, Enhanced larval supply and recruitment can replenish reef corals on degraded reefs. *Sci. Rep.* **7**, 1–13 (2017).

53. L. Mathon, *et al.*, Cross-ocean patterns and processes in fish biodiversity on coral reefs through the lens of eDNA metabarcoding. *Proc. R. Soc. B Biol. Sci.* **289** (2022).

54. S. ichiro Oka, *et al.*, Environmental DNA metabarcoding for biodiversity monitoring of a highly diverse tropical fish community in a coral reef lagoon: Estimation of species

592 richness and detection of habitat segregation. *Environ. DNA* **3**, 55–69 (2021).
593 55. J. Spens, *et al.*, Comparison of capture and storage methods for aqueous macrobial
594 eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods*
595 *Ecol. Evol.* **8**, 635–645 (2017).
596 56. K. Bohmann, *et al.*, Strategies for sample labelling and library preparation in DNA
597 metabarcoding studies. *Mol. Ecol. Resour.* **22**, 1231–1246 (2022).
598 57. Illumina, "16S Metagenomic Sequencing Library Preparation" (2013).
599 58. D. A. Bohan, *et al.*, Next-Generation Global Biomonitoring: Large-scale, Automated
600 Reconstruction of Ecological Networks. *Trends Ecol. Evol.* **32**, 477–487 (2017).
601 59. M. Miller, *et al.*, Temperate functional niche availability not resident-invader
602 competition shapes tropicalisation in reef fishes. *Nat. Commun.* **14**, 2181 (2023).
603 60. I. Adrian-Kalchhauser, P. Burkhardt-Holm, An eDNA assay to monitor a globally
604 invasive fish species from flowing freshwater. *PLoS One* **11**, 1–22 (2016).
605 61. J. Elith, J. R. Leathwick, Species distribution models: Ecological explanation and
606 prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* **40**, 677–697 (2009).
607 62. C. Ramirez-Reyes, *et al.*, Embracing ensemble species distribution models to inform
608 at-risk species status assessments. *J. Fish Wildl. Manag.* **12**, 98–111 (2021).
609 63. E. Ottimofiore, *et al.*, Responses of coral reef fishes to past climate changes are
610 related to life-history traits. *Ecol. Evol.* **7**, 1996–2005 (2017).
611 64. L. A. Freeman, J. A. Kleypas, A. J. Miller, Coral reef habitat response to climate change
612 scenarios. *PLoS One* **8**, 1–14 (2013).
613 65. P. Peduzzi, J. Concato, E. Kemper, T. R. Holford, A. R. Feinstein, A simulation study of
614 the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **49**,
615 1373–1379 (1996).

66. R. Salmerón, C. B. García, J. García, Variance Inflation Factor and Condition Number in multiple linear regression. *J. Stat. Comput. Simul.* **88**, 2365–2384 (2018).

67. L. Zhang, *et al.*, The use of classification and regression algorithms using the random forests method with presence-only data to model species' distribution. *MethodsX* **6**, 2281–2292 (2019).

68. A. H. Fielding, J. F. Bell, A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* **24**, 38–49 (1997).

69. O. Allouche, A. Tsoar, R. Kadmon, Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232 (2006).

70. R Core Team, R: A language and environment for statistical computing (2021).

71. A. Liaw, M. Wiener, Classification and Regression by randomForest. *R News* **2**, 18–22 (2002).

72. S. N. Wood, Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73**, 3–36 (2011).

73. I. Ball, H. P. Possingham, M. E. Watts, "Marxan and relatives: Software for spatial conservation prioritisation" in *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools*, A. Moilanen, K. A. Wilson, H. P. Possingham, Eds. (Oxford University Press, 2009), pp. 185–195.

74. Badan Pusat Statistik, Village Potential Statistics (PODES) 2020 (2020).

75. A. E. Gaughan, F. R. Stevens, C. Linard, P. Jia, A. J. Tatem, High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. *PLoS One* **8** (2013).

76. J. E. Cinner, *et al.*, Gravity of human impacts mediates coral reef conservation gains.

Proc. Natl. Acad. Sci. **115**, E6116–E6125 (2018).

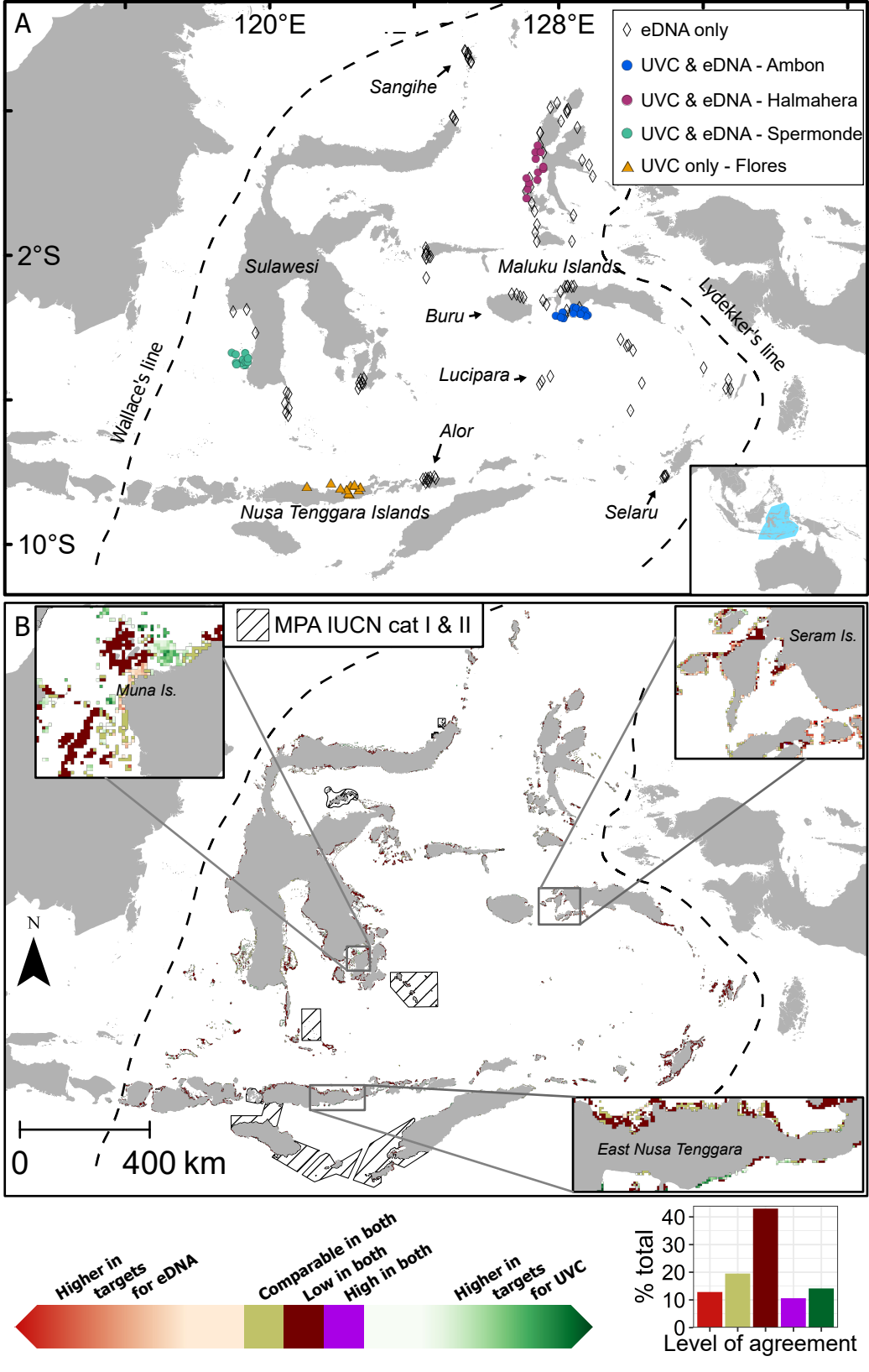
77. Allen Coral Atlas, Imagery, maps and monitoring of the world's tropical coral reefs. (2022) https:/doi.org/doi.org/10.5281/zenodo.3833242k.

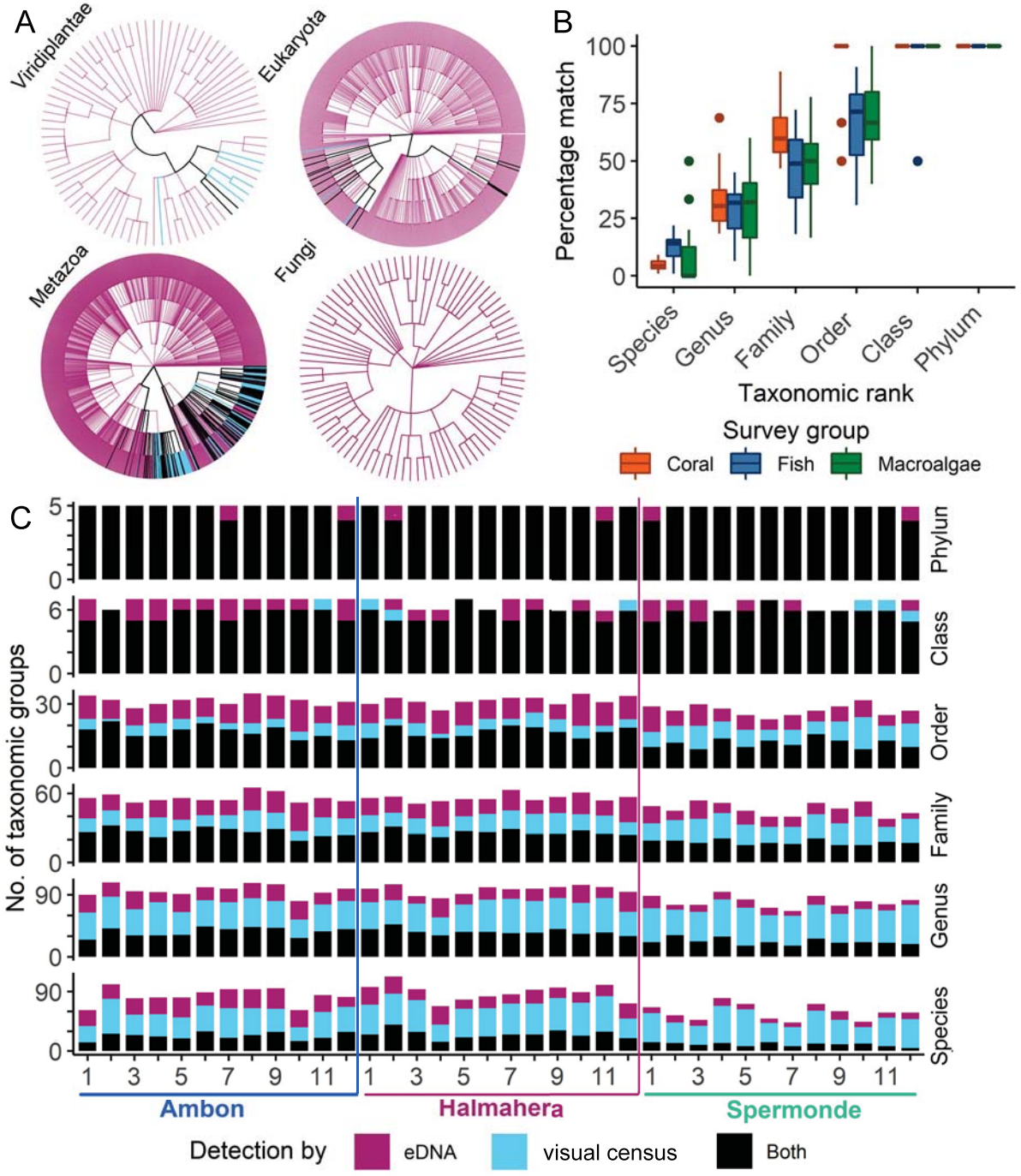
Figure legends

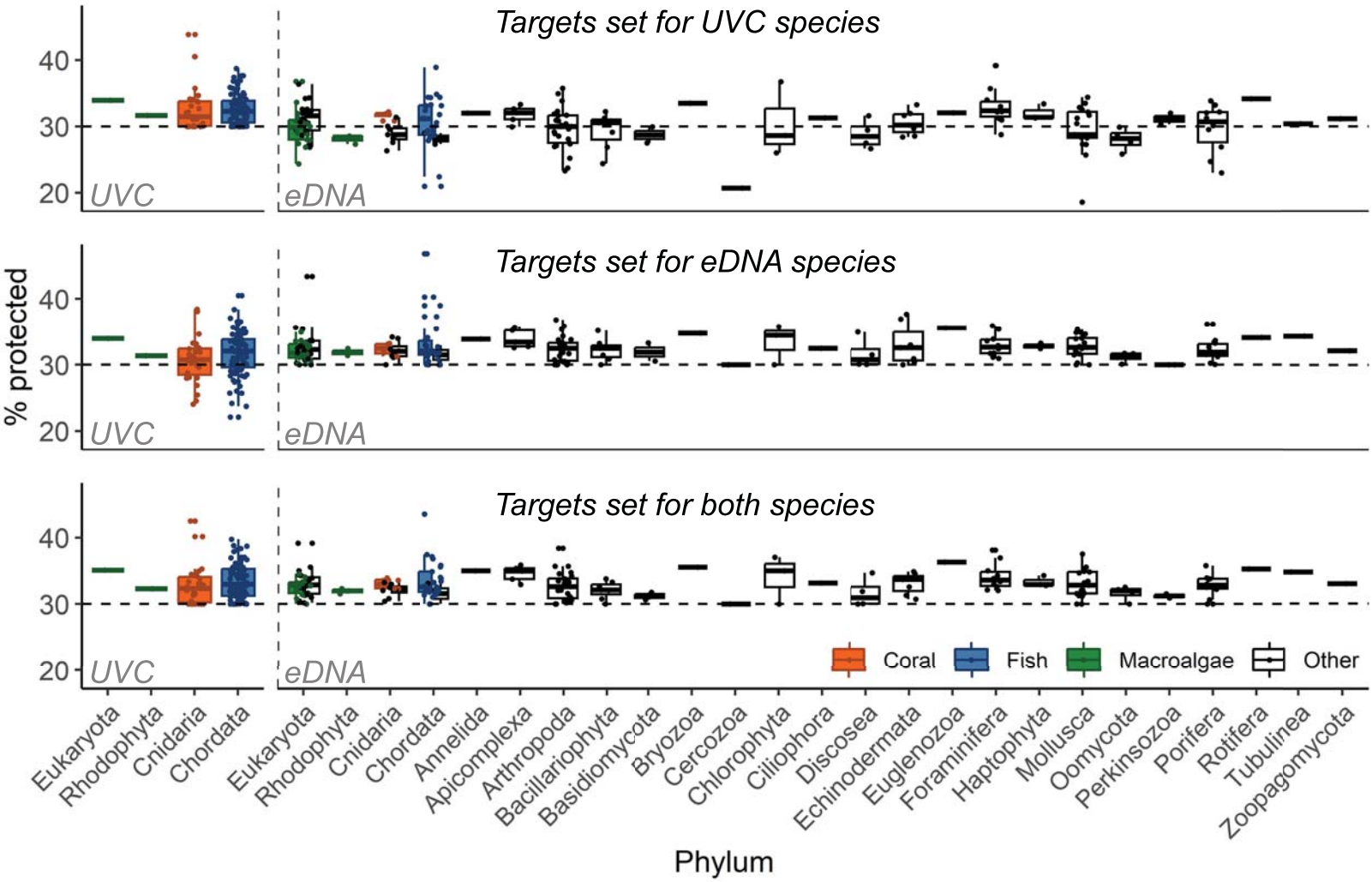
FIG 1. Map of Wallacea biogeographical region. A) Coral reef sites surveyed using either underwater visual census (UVC), environmental DNA (eDNA) metabarcoding, or both. B) Differences in conservation priorities when targets are set for species recorded by UVC or eDNA with a histogram showing the distribution of the changes.

FIG 2. Comparison of the marine taxonomic diversity identified by visual census and eDNA metabarcoding of the COI and 18S genes across 36 surveyed sites in the Wallacea region. A) Phylogenetic tree pruned at genus level showing genera identified by either eDNA (violet), visual census (turquoise), or both methods (black) across all sites . B) Percentage of taxonomic groups identified by visual census that were also identified by eDNA at individual sites by corals (orange), fish (blue), and macroalgae (green). C) Detection of shared taxonomic groups (Table S1) by either one or both methods at individual sites.

FIG 3. Summary of three spatial prioritisation analyses where targets were set for either species recorded by underwater visual census (UVC), eDNA, or both. The y-axis shows the percentage of protection for visually detected species (left column) and eDNA species (right column). The dashed horizontal line indicates the conservation target set at 30%. Boxplots are coloured by groups of coral, fish, macroalgae, and other.


A
120°E
128°E
2°S
10°S
eDNA only
UVC & eDNA - Ambon
UVC & eDNA - Halmahera
UVC & eDNA - Spermonde
UVC only - Flores
Sangihe
Sulawesi
Maluku Islands
Buru
Lucipara
Alor
Nusa Tenggara Islands
Selaru
Wallace's line
Lydekker's line
B
MPA IUCN cat I & II
Muna Is.
Seram Is.
East Nusa Tenggara
N
0
400 km
Higher in targets for eDNA
Comparable in both
Low in both
High in both
Higher in targets for UVC
% total
40
30
20
10
0
Level of agreement


A
Viridiplantae
Eukaryota
Metazoa
Fungi
B
Percentage match
100
75
50
25
0
Species
Genus
Family
Order
Class
Phylum
Taxonomic rank
Survey group
Coral
Fish
Macroalgae
C
No. of taxonomic groups
Phylun
Class
Order
Family
Genus
Species
1 3 5 7 9 11
Ambon
Halmahera
Spermonde
Detection by
eDNA
visual census
Both


Targets set for UVC species
Targets set for eDNA species
Targets set for both species
% protected
40
30
20
UVC
eDNA
Coral
Fish
Macroalgae
Other
Eukaryota
Rhodophyta
Cnidaria
Chordata
Annelida
Apicomplexa
Arthropoda
Bacillariophyta
Basidiomycota
Bryozoa
Cercozoa
Chlorophyta
Ciliophora
Discosea
Echinodermata
Euglenozoa
Foraminifera
Haptophyta
Mollusca
Oomycota
Perkinsozoa
Porifera
Rotifera
Tubulinea
Zoopagomycota
Phylum