Prospects for integrating environmental DNA metabarcoding into the conservation of marine megafauna

Elizabeth Anne Boyse

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Biology

May 2023

The candidate confirms that the work submitted is her own, except where work has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

- The work in Chapter 2 has been prepared as a manuscript for publication as follows:

Boyse, E., Robinson, K.P., Beger, M., Carr, I., Raynor, M., Valsecchi, E., Goodman, S.J. 2023. **Environmental DNA reveals fine-scale spatial and temporal variation of prey species for marine mammals in a Scottish marine protected area.** *In preparation*.

Author contributions: *EB conceived the study with support from KPR, MB and SJG, conducted all laboratory work, performed all data analysis and data visualisation, wrote the initial manuscript draft and revised the manuscript. EB and KPR conducted the fieldwork. MR contributed to the sequencing. IC contributed towards the bioinformatics analyses. KPR, MB, EV and SJG provided feedback on and edited the manuscript.*

- The work in Chapter 3 has been prepared as a manuscript for publication as follows:

Boyse, E., Robinson, K.P., Beger, M., Goodman, S.J. 2023. **Inferring species interactions from co-occurrence networks with environmental DNA metabarcoding data.** *In preparation*.

Author contributions: EB, MB and SJG conceived the study. EB and KPR conducted the fieldwork. EB performed all data analysis and data visualisation, wrote the initial manuscript draft and revised the manuscript. KPR, MB and SJG provided feedback on and edited the manuscript. MB and SJG contributed equally to this work.

- The work in Chapter 4 has appeared in publication as follows:

Boyse, E., Beger, M., Valsecchi, E., Goodman, S.J. 2023. **Sampling from commercial vessel routes can capture marine biodiversity distributions effectively.** *Ecology and Evolution*. 13(2): e9810.

Author contributions: EB, MB and SJG conceived the study. EB curated the dataset, developed the methodology, performed all data analysis and data visualisation, wrote the initial manuscript draft and revised the manuscript. MB, EV and SJG provided feedback on and edited the manuscript.

- The work in Chapter 5 has been submitted for publication as follows:

Boyse, E., Goodman, S.J., Beger, M. 2023. **Multi-taxon versus taxon specific spatial conservation priorities for marine predators.** *In review. Conservation Biology.*

Author contributions: EB and MB conceived the study, with support from SJG. EB curated the dataset, performed all data analysis and data visualisation, wrote the initial manuscript draft and revised the manuscript. MB and SJG provided feedback on and edited the manuscript.

Acknowledgements

To all my supervisors, Simon Goodman, Maria Beger, Kevin Robinson, Ian Carr and Elena Valsecchi, thank you so much for your support, encouragement and expertise which has allowed me to grow so much as a scientist. Firstly, Simon, thank you so much for giving me the opportunity to enter the crazy, ever-growing world of environmental DNA, and for taking a gamble on me despite my complete lack of genetics background. I am thoroughly grateful for this experience, the wonderful places I have had the opportunity to go to, and for all the skills I have gained. Maria, thank you for always pushing me to aim higher and for giving me confidence in myself. You have been a massive inspiration to me, and it has been a privilege to work with you. Ian, thank you for your never-ending patience while teaching me bioinformatics, answering my stupid questions without laughing at me (mostly!) and for all the great hedgehog videos. Elena, thank you for kick-starting me on my eDNA journey in the lab and for all your interesting ideas and expert knowledge along the way. One day, I'll make it onto a ferry in Italy!

Kev, I will never be able to repay for you for single-handedly saving my PhD when I was losing all hope of ever being able to collect samples thanks to that pesky pandemic. I will treasure my memories of the Moray Firth forever, especially my first fin whale encounter, getting stuck in the sand towing a dead Sowerby's beaked whale and the northern lights on one of the last nights of the field season. You are stuck with me for life I am afraid!

A very special thank you from 'Betty Bucket' to all the Cetacean Research and Rescue Unit volunteers who helped me to collect and filter my samples. To Kev & Jen, thank you very much for letting me turn your cellar into my makeshift laboratory. Jen, thank you so much for all your pearls of wisdom, and for the most incredible food I have ever eaten!

Thank you to all my friends that I have met in Leeds and been a part of this crazy journey with me, even tackling a pandemic together. To my housemate Izzy, thank you for always being there on my lowest days to make mug cakes, or to celebrate the highs with wonderful vegan dinners. To my lockdown buddy, Rachel, thank you for converting me into a fully-fledged birder and

keeping me sane with trashy TV and countless takeaways. To Meg, thank you for all the gym therapy sessions and garden centre trips during our house plant craze.

To all of the Goodman and Beger lab group members, Julia, Laura, Harri, Katie, Dom, Fiona, Louise, Adele, James, Victoria, Mark, Morven, Hannah, Jenni and Tim, thank you for being such supportive colleagues and friends. A special thanks to Julia for embarking on this journey with me from start to the end, it wouldn't have been the same without you.

Thank you to my parents for always believing in me, even when I haven't believed in myself. I told you when I was around 3 years old after watching 'Free Willy' for the millionth time that I was going to save the whales, and thanks to your unwavering support, I am living my dreams. Thank you to my grandma, I will always smile picturing you saying, "This is my granddaughter, she's going to be a doctor of fish." Last but definitely not least, thank you to Roger the Norfolk Terrier, my smallest but biggest supporter, for sitting on my lap whilst I wrote up my thesis, and for chasing squirrels with me when I needed to clear my head.

Abstract

Marine megafauna include some of the most threatened taxa globally, risking the loss of key ecosystem functions and services that they provide unless conservation action can reverse their declining population statuses. Marine protected areas (MPAs) are the most popular tool implemented to conserve marine megafauna, but it remains unclear how effectively current MPA systems are capturing the requirements of marine megafauna due to incomplete knowledge of their distributions and limited long-term monitoring post-introduction of conservation measures.

Environmental DNA (eDNA) offers a novel way to survey the marine environment, improving the detectability of rare species and observing changes across whole ecosystems simultaneously. The aim of my thesis was to explore how eDNA metabarcoding could improve the conservation of marine megafauna. I addressed this aim across four main research questions:

1) Can simultaneous monitoring of marine megafauna and their prey with eDNA metabarcoding improve understanding of their fine-scale habitat use?
2) Does the structure and complexity of cetaceans' ecosystems change spatially and temporally, and how does this effect vulnerability to climate change?
3) Can we retrieve high quality species distributions from commercial vessel surveys, permitting eDNA surveys to be upscaled across greater spatiotemporal resolutions?
4) How does improved spatiotemporal and taxonomic coverage offered by eDNA impact spatial priority area designations within a marine spatial planning framework?

eDNA revealed previously unknown patterns in the spatiotemporal availability of key prey species for cetaceans that improved understanding of their habitat use within a newly designated marine protected area. These prey species are highly connected and important ecosystem components but are particularly vulnerable to climate change so future changes to the ecosystem structure can be expected.

Accessing spatiotemporal scales relevant to marine megafauna distributions often relies on expensive dedicated research vessels, but, using a simulation approach, I demonstrate that commercial vessels offer viable alternative survey platforms. Sampling along commercial routes that cover environmental variability and community composition adequately can predict species distributions as accurately as comparative non-spatially biased samples. Commercial vessels coupled with advances in automating eDNA sample processing offers a promising avenue to increase regular monitoring of the marine environment, especially in areas that are often under sampled and otherwise inaccessible.

Marine spatial planning often implements coarse ecological data, such as habitat types or bioregions, to identify priority areas for conservation management. eDNA surveys can improve taxonomic coverage and understanding of species distributions. I showed that considering taxonomic groups (marine mammals, elasmobranchs, teleost fishes) separately more effectively identified spatial priority areas for the different groups which were missed in spatial priority areas designed for all taxa simultaneously. Including conservation objectives specific to different marine megafauna taxa will achieve more effective conservation that incorporates their contrasting and specialised life histories and habitat requirements.

Table of Contents

List of Figures

List of Tables

Chapter One – General Introduction

The Anthropocene epoch has exerted human influence on every part of the ocean, more and more necessitating well-implemented effective conservation actions (Halpern et al., 2019). On average, threatened marine species are exposed to cumulative human impacts in over half of their ranges (O'Hara et al., 2021). Overexploitation, habitat destruction, climate change and pollution have resulted in increased losses of marine biodiversity, including 38% and 71% declines in the global abundance of bony fishes and sharks and rays respectively since the 1970s (Hutchings et al., 2010; WWF, 2022). These losses jeopardise vital ecosystem services including carbon uptake and storage, mediating weather patterns, food for over a billion people and coastal defences. Reversing or reducing the loss of marine biodiversity is therefore essential for long-term food security, as well as increased resilience to climate change (Sumaila & Tai, 2020).

1.1. The state of marine conservation in the 21st century

Marine protected areas (MPAs), *i.e.*, geographically defined areas with regulations or management to achieve conservation objectives, are the main tool to conserve marine ecosystems from human impacts (Maestro et al., 2019; Trouillet & Jay, 2021). The Convention on Biological Diversity's (CBD) Aichi Targets objective to protect 10% of oceans by 2020, and more recently, the 2030 Agenda to increase protection to 30%, have been key legislations in the expansion of MPA coverage (Maestro et al., 2019). As of 2023, 18, 445 MPAs have been established, covering 29.6 million km^2 equating to 8.16% of the ocean's surface (Figure 1.1)(UNEP-WCMC & IUCN, 2023). Effective MPAs have resulted in increased richness and biomass of large fish species, and rely on five key features; no-take regulations, enforcement, large (>100 km^2), old (>10 years) and isolated (Edgar et al., 2014). MPAs covering these criteria are a small minority, with most failing due to a lack of effective, equitable management resulting from monetary and workforce constraints (Edgar et al., 2014; Gill et al., 2017). Instead, less than 3% of the ocean is fully or highly protected because other MPAs still permit significant human activities that prevent conservation objectives being met or are 'paper parks' which are legally designated but no management actions are implemented, resulting in no ecosystem benefits

(Grorud-Colvert et al., 2021; Rilov et al., 2020). Further, global coverage of MPAs is biased, with only 1.44% of protected area coverage in areas beyond natural jurisdiction (ABNJ) despite constituting over 60% of the ocean (Maestro et al., 2019). Whilst large size contributes towards effective MPAs, they are also the most challenging to enforce and manage, conflicting with the lack of available resources (Wilhelm et al., 2014). There has also been a tendency to implement large MPAs in remote areas with low human impacts and economic value for ease of designation, contributing marginally to protecting threatened species and habitats (Devillers et al., 2015; Devillers et al., 2020). Looking ahead, achieving protected area coverage targets will not automatically accomplish marine biodiversity conservation objectives, without improved management and compliance (Edgar et al., 2014).

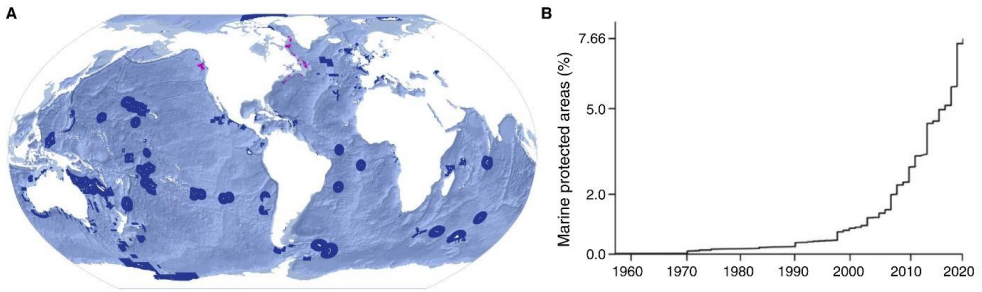


Figure 1.1. a) The global distribution of marine protected areas (blue) and other effective area-based conservation measures (pink). b) The temporal expansion of global marine protected area coverage (%). Data from UNEP-WCMC and IUCN (2021). Figure from Lotze (2021).

Marine spatial planning (MSP) is the process of spatially and temporally allocating resource use and management areas to a seascape to achieve ecological, economic and social objectives whilst minimising stakeholder conflict (Ehler, 2021). This process has expanded massively over the past 30 years, with 70 countries covering six continents and four ocean basins now developing MSP (Santos et al., 2021). In theory, MPAs integrated within an MSP framework offer a logical means for allocating areas to achieve ecological goals (Trouillet & Jay, 2021). However, real-world MSP is generally driven by blue growth, such as the expansion of renewable energy projects or oil and gas infrastructure, whilst MPA designations are handled by separate institutional processes (Jones et al., 2016; Trouillet & Jay, 2021; Young, 2015). For example, in Europe, two different approaches for marine

spatial planning have been adopted through the Marine Strategy Framework Directive (MSFD) and the Integrated Maritime Policy (IMP) (Qiu & Jones, 2013). The MSFD's main objective is to achieve good environmental status with MPAs as an integral tool, while IMP places far greater emphasis on promoting maritime economic development and conservation and MPAs form one of several uses of marine space (Jones et al., 2016; Qiu & Jones, 2013). Regardless, the effectiveness of achieving conservation through either MPAs, MSP, or the convergence of both, will depend on financial constraints, and adequate data to inform these processes (Trouillet & Jay, 2021). As biodiversity data availability increases, more emphasis should be placed on separate conservation requirements by different groups of species or habitats, which are typically considered synergistically in MSP decision-making tools (Augé et al., 2018).

1.2. The conservation of marine megafauna

Marine megafauna are an ecologically important but highly threatened group that are not being sufficiently protected in global MPA systems (Conners et al., 2022). They encompass a large and diverse group of organisms including marine mammals, elasmobranchs, large teleost fishes and seabirds, which typically occupy high trophic levels in marine ecosystems. Most perform key roles in ecosystem functioning such as top-down control of food webs, nutrient cycling and redistribution, ecosystem engineering and climate change mitigation (Hammerschlag et al., 2019). They are also highly valuable and heavily targeted commodities by global fisheries and eco-tourism, both multi-billion-dollar industries creating millions of job opportunities (Cisneros-Montemayor et al., 2010; Juan-Jordá et al., 2011; Myers & Worm, 2005). Marine megafauna can themselves be useful conservation tools by acting as indicators of broader ecosystem changes (Hazen et al., 2019). For example, population trends and dietary or distribution changes can indicate changes in the abundance or distribution of their prey species which are often more challenging ecosystem components to monitor (Hazen et al., 2019). Our understanding of the ecological influence of marine megafauna is based on a limited number of well-studied species, but different species will serve different functions thus necessitating further research to argue their importance for increased conservation measures (Estes et al., 2016).

Marine megafauna constitute some of the most threatened taxa globally, with over a third of marine mammals and elasmobranch species classified as threatened in IUCN red list assessments (Davidson et al., 2012; Dulvy et al., 2021). Overfishing is the leading threat to all taxa, with targeted fisheries resulting in the loss of 90% of large predatory fishes (Myers & Worm, 2003). In contrast, elasmobranchs and marine mammals are more threatened by incidental catches, *i.e.*, bycatch or entanglement in ghost fishing gear (Dulvy et al., 2021; Estes et al., 2016). Other shared threats include habitat degradation, pollution, and climate change (Arthington et al., 2016; Avila et al., 2018; Dulvy et al., 2021). Noise pollution and ship strikes are well established threats for marine mammals, but are less understood, although potentially significant threats, for teleost fishes and elasmobranchs (Avila et al., 2018; Mickle & Higgs, 2022; Pirotta et al., 2019). Climate change is likely to supersede other threats over the coming decades as the oceans continue warming and could also reduce the effectiveness of conservation measures, such as MPAs, to mitigate against other threats (Albouy et al., 2020). Tropical sharks and teleost fishes who are already close to their thermal tolerances, as well as marine mammals with specialised diets, small geographical ranges or dependency on ice are most susceptible (Albouy et al., 2020; Diaz-Carballido et al., 2022). Global mismatches between MPAs and marine megafauna ranges suggest that their conservation needs are not currently being well captured with MSP (Conners et al., 2022).

Marine megafauna are challenging to protect with MPAs as we frequently lack the population-level distributional data required for MPA implementation, and to set realistic population growth or recovery objectives, exacerbated because significant human impacts precede baseline data collection (Jorgensen et al., 2022). Generally, marine megafauna are migratory species with large spatial ranges, subjected to different legislations and protections as they pass through Exclusive Economic Zones or ABNJ, necessitating cross-country collaborations (Lascelles et al., 2014). To encompass their ranges within MPAs would require protecting impossibly large areas (Wilhelm et al., 2014), and thus marine megafauna often remain exposed to threats when travelling outside of MPAs (Notarbartolo di Sciara et al., 2016). MPAs designated for marine megafauna often focus on habitats

deemed critical such as breeding or foraging grounds, which are constrained to relatively small spatial scales compared to their overall ranges (Nelms et al., 2021). However, this method will only be effective if the life stages which confer highest population growth rates are protected, requiring data on specific life stages (Conners et al., 2022). Further, these smaller areas, especially foraging grounds, are expected to shift temporally resulting from both megafauna and prey species responding to global warming. Therefore, incorporating more dynamic management into MPA objectives is encouraged to ensure long-term effectiveness (Maxwell et al., 2020). Enhanced understanding of marine megafauna distributions, and climate-induced distributional shifts, through technological advances will support MPA expansion to better cover marine megafauna requirements (Grémillet et al., 2022).

1.3. Enhancing marine conservation through new technological innovations

Technological advances, such as remote sensing, telemetry, and environmental DNA (eDNA), have enabled retrieval of higher quality data across larger spatiotemporal scales, enhancing our knowledge of species distributions and areas where mitigation against human impacts are required (Berger-Tal & Lahoz-Monfort, 2018; Dutton et al., 2019). Combined, these technologies provide new ways of identifying marine species and their whereabouts without needing to visually sight them, including tracking of long-range migrations, and the environmental variables responsible for driving their movements (Grémillet et al., 2022). These advances can support marine research where difficult accessibility and monetary constraints have resulted in a lack of knowledge and delayed conservation initiatives relative to the terrestrial environment (Maestro et al., 2019). Even so, only 5% of the ocean's surface has been explored, with pervasive biases in record collection in coastal environments (< 5 km from shore) or along busy shipping routes (Hughes et al., 2021). Modern technologies to reduce key knowledge gaps in the marine environment are essential to assess the health of ecosystems, mitigate human impacts and climate change, and to optimise MPA placement (Edgar et al., 2016).

Key technological innovations have revolutionised marine megafauna research in two ways; firstly, by directly improving our knowledge of species habitat use, distributions and behaviours, and secondly, by increasing the availability of high-resolution environmental data to determine abiotic drivers (Grémillet et al., 2022)(Table 1.1). Telemetry technologies, such as satellite or acoustic tags, have helped better understand the spatial ecology of marine megafauna, providing new insights into home ranges and critical habitats, trans-oceanic migrations and vertical distributions (Grémillet et al., 2022; Hussey et al., 2015). Previously, species ranges were evaluated using mark-recapture methods which relied on individuals being recaptured and only provided very coarse data, *i.e*., mark location and recapture location (Eberhardt et al., 1979; Hussey et al., 2015). Unmanned aerial vehicles (UAVs) have improved population abundance estimates, especially for highly dispersed, far-ranging species such as polar bears, and health assessments based on body condition indexes which can inform the conservation status of populations (Fiori et al., 2017; Rees et al., 2018). UAVs can also identify behaviours which may indicate critical habitats, *i.e.*, foraging or breeding grounds, for protection (Schofield et al., 2019). However, UAVs are restricted to species which surface regularly, *i.e.,* marine mammals, rely on land for part of their life history strategy, or inhabit shallow, coastal environments, such as reef sharks (Kiszka et al., 2016). Alongside improved understanding of fine-scale species distributions and habitat use, remote sensing has provided data on important environmental parameters, often at global scales with increasingly fine-scale resolution, to identify abiotic and biotic drivers of species occurrences (Rose et al., 2015). For example, satellite remote sensing of ocean colour provides data on phytoplankton biomass and composition, whilst infrared or microwave radiometers provide data on sea surface temperature (Kavanaugh et al., 2021; Minnett et al., 2019). The challenge is now to utilise these tools in complementary ways to continue expanding spatiotemporal coverage of species distributions, whilst incurring minimum costs.

Table 1.1. Summary of modern technologies that have been incorporated into marine megafauna research and have potential to improve their conservation.

Technology	Data gained	Conservation applications	Limitations
Telemetry – animal borne sensors	- Horizontal and vertical movements of individuals - Ambient physical variables *i.e.,* conductivity, temperature, depth - Heart rate/dissolved oxygen/blood chemistry	- Improved knowledge of species distributions, especially long-range migrations - Filling important observational gaps in key physical environmental variables	- Small sample sizes - Difficult and invasive attachment techniques - GPS tags only transmit location at the surface - Acoustic tags dependent on location of receivers
Remote sensing	- Identification of large species/species aggregations - Environmental variables *i.e.,* primary productivity/depth that drive species distributions	- Understand biotic and abiotic drivers of species habitat use - Detect changes in important environmental variables	- Erroneous species identification/abundance estimates from user error/automated methods - Incomplete coverage/reliance on cloud free conditions - Trade-off between spatial and temporal resolution
Molecular techniques including environmental DNA	- Identification of species and delineation between populations - Effective population size - Species diversity - Inbreeding depressions - Adaption potential	- Broader taxonomic coverage in line with ecosystem-based management - Distinguish different management stocks more accurately - Non-invasive technique for threatened species	- Incomplete reference libraries for species identification - No exact abundance measures - Different marker genes can show conflicting patterns/not represent the whole genome
Unmanned aerial vehicles	- Species identification - Quantify behavioural states - Collect genetic material - Assess body condition	- Population assessments - Identification of critical habitats - Improved understanding of individual and population health	- Restricted to air-breathing species or species in shallow/clear waters - Constrained by battery life/weather

Coupled advances in techniques, such as telemetry and remote sensing, have increased the widespread use of species distribution models, which are important to support conservation decision making (Hussey et al., 2015; Zurell, Franklin, et al., 2020). These models relate presence or abundance data from the species of interest with environmental covariates to predict species distributions across space and time. Past or future distributions can be predicted as well as overall community composition, and biotic and abiotic drivers of distributions can be identified (Randin et al., 2020). These models are instrumental for evaluating whether MPAs are adequately covering species distributions and whether they will continue to do so in the future (Wilson et al., 2020). They are also frequently used as conservation features within spatial planning decision making tools, such as Marxan, to select planning units that cover specific distribution coverage targets whilst incurring the least costs, often represented as opportunity costs, *i.e.*, displaced fishing effort (Smith et al., 2009). Near real-time species distribution modelling predictions have enabled dynamic oceanic management for bycatch mitigation and dynamic fisheries closures based on daily oceanographic parameters (Hazen et al., 2018). This approach has been successfully implemented in Eastern Australia's longline fisheries to reduce catches of southern bluefin tuna (Hobday & Hartmann, 2006). Smaller spatial closures are required compared to static management areas, thus benefitting the fisheries whilst adequately protecting the species of conservation concern (Hazen et al., 2018; Hobday & Hartmann, 2006).

Molecular techniques can complement other technological advances by providing unique data that enhances our ability to conserve biodiversity at different levels depending on genetic variation, *i.e.*, between individuals, populations or species (Nielsen et al., 2022). Central to conservation genetics is assessing the short-term and long-term viability of populations through levels of inbreeding and adaptive potential respectively (Fuentes-Pardo & Ruzzante, 2017; Ouborg et al., 2010). Genetics tools also provide essential understanding of population structure, including delineation of separate population units, their phylogeographic history and connectivity between populations (Hohenlohe et al., 2021; Kershaw et al., 2021). High-throughput sequencing has been pivotal for increasing the resolution and accuracy of

these analyses through greater coverage of sequenced loci (thousands to millions) across the whole genome compared to 10-20 neutral markers previously (Hunter et al., 2018). For example, runs of homozygosity (ROH) mapped to a reference genome can reveal more precise individual inbreeding estimates without parental analysis, and the length of ROH regions can indicate how recent or old inbreeding events were (Fuentes-Pardo & Ruzzante, 2017; Hohenlohe et al., 2021). High-throughput sequencing has also enabled us to answer new questions, such as which regions of the genome are involved in adaptation to local environmental change, facilitated by the identification of functional genetic variation, *i.e.*, portions of the genome under selection which effect fitness (Kardos et al., 2021; van Oppen & Coleman, 2022). While genomic techniques typically rely on high quality DNA samples, high-throughput sequencing has also increased the applicability of low-quality samples from which shorter fragments at fewer loci are retrieved, mainly as a tool to detect species presence (Hohenlohe et al., 2021). This advance led to the emergence of eDNA approaches, *i.e.*, DNA sourced directly from the environment, to detect species and make population genetics inferences, which are rapidly becoming popular tools for monitoring marine biodiversity (Takahashi et al., 2023).

1.4. Incorporating environmental DNA into conservation

eDNA coupled with technological advances in molecular approaches have provided novel methods to characterise ecosystem community profiles with metabarcoding or to monitor specific species with more sensitive assays (Deiner et al., 2017). eDNA refers to DNA extracted directly from environmental samples, such as air, soil or water, without needing to isolate the target organism(s) first (Taberlet et al., 2012). eDNA originates from a wide variety of sources including whole living microbes, or cellular remains (*i.e.*, faeces, urine, mucus, sloughed skin cells, gametes) and extracellular DNA for macrofauna (Deiner et al., 2017). eDNA analyses are comprised of three different approaches; single-species identification, metabarcoding where multiple taxa are identified, and metagenomics where functionality or genome assemblies are investigated (Figure 1.2). The first steps in all approaches includes sample collection from the field followed by DNA extraction. To detect a single species or to investigate population genetics,

sensitive PCR techniques, such as qPCR or digital-droplet PCR, are employed to quantify low levels of species-specific DNA (Williams et al., 2019). The metabarcoding approach instead uses conventional PCRs with conservative primers that amplify barcoding regions from multiple taxa of interest. High-throughput sequencing then allows for sequences from multiple templates to be sequenced in parallel, contributing to the expansion of sequencing data incorporation into ecological research (Shokralla et al., 2012). Bioinformatic pipelines are employed to de-noise the data by demultiplexing and removing errors (*i.e.,* from contamination or chimeras through PCR), and to identify taxa from sequencing reads through molecular operational taxonomic units (OTUs) or amplicon specific variants (ASVs) using reference sequencing libraries (Calderón-Sanou et al., 2020; Callahan et al., 2017). The metagenomics approach forgoes any DNA amplification, and instead indiscriminately sequences DNA extracts, which can be useful for population genomic studies to identify SNPs when full reference genomes are available (Cammen et al., 2016).

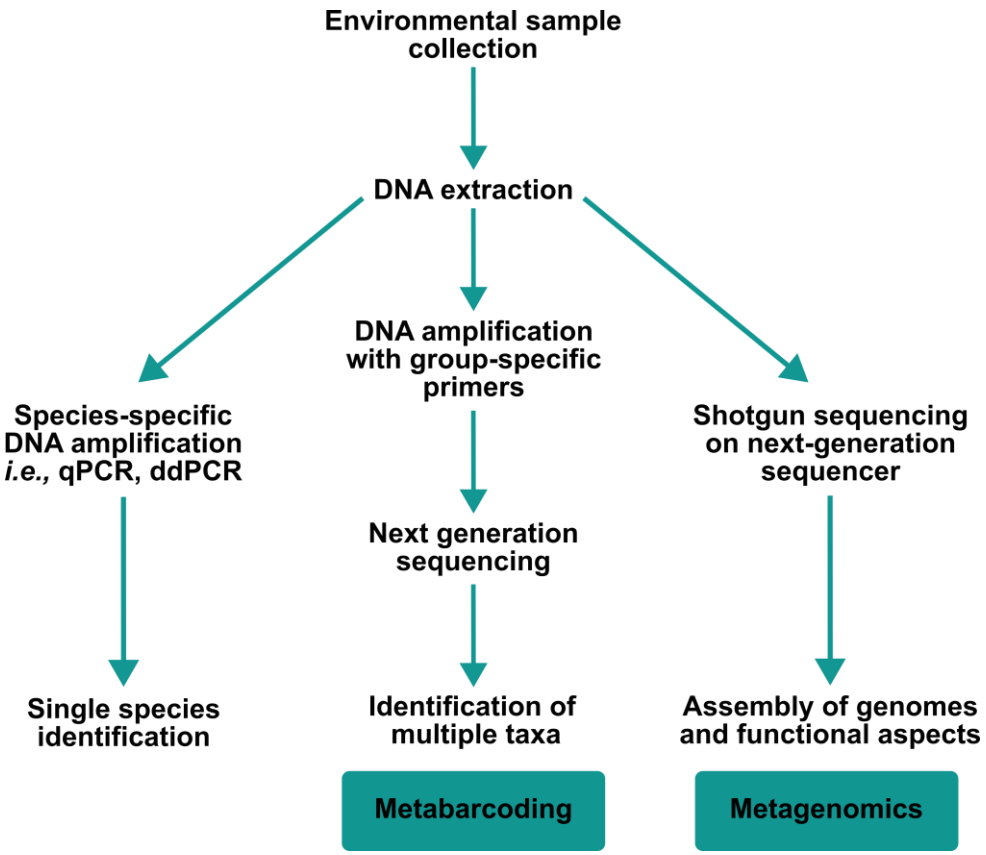


Figure 1.2. The different workflows for the three main approaches (single species, metabarcoding and metagenomics) employed in environmental DNA studies. Adapted from Taberlet et al. (2018).

eDNA approaches were initially incorporated into microbiology to predict community composition of microbes which were unable to be cultured successfully (Thomsen & Willerslev, 2015). Throughout the early 2000s, this method was applied to investigate the diversity and functionality of microbial communities, and assemble genomes for uncultivable microbes in a range of environments including soil, marine, freshwater and permafrost (Riesenfeld et al., 2004). This was followed by the first study on macro-organisms in 2003 from permafrost cores and sediment deposits where 300 to 400 ky old DNA from plants, and 20-30 ky old DNA from vertebrates were detected, including both extinct and extant species (Willerslev et al., 2003). Examining paleoenvironments with eDNA has permitted the reconstruction of high-latitude communities and altered views about paleocommunities of certain epochs (Willerslev et al., 2007). In 2008, Ficetola et al. (2008) utilised eDNA in the freshwater environment for the first time, demonstrating that eDNA could accurately predict the presence and absence of an important invasive

species, the American bullfrog (*Rana catesbeiana*). 2012 marked a pivotal year in establishing eDNA as a biodiversity monitoring tool (Takahashi et al., 2023), with the first comprehensive eDNA reviews published (Taberlet et al., 2012; Yoccoz, 2012), and the first studies on fish and marine mammals in the marine environment (Foote et al., 2012; Thomsen et al., 2012). eDNA detected the same amount or more fish species than nine methods commonly used for fish surveys, showing that eDNA is an important tool for investigating community composition in the marine environment (Thomsen et al., 2012). Meanwhile, detections of harbour porpoises (*Phocoena phocoena*) were less reliable with eDNA compared to acoustic methods (Foote et al., 2012). Since then, the number of aquatic eDNA papers has steadily grown, with metabarcoding publications exceeding single-species approaches for the first time in 2021 (Takahashi et al., 2023), and the method is continuously being applied across more environments, to answer more questions, *i.e.*, community composition characterisation (Sawaya et al., 2019), threatened or invasive species detections (Juhel et al., 2021; Whitaker et al., 2021), population genetics (Parsons et al., 2018), response of communities to human impacts or climate anomalies (Bakker et al., 2017; Djurhuus et al., 2020), and trophic interactions (D'Alessandro & Mariani, 2021).

eDNA has been proposed as a useful biomonitoring tool in multiple applied conservation settings, for example, in fisheries management, invasive species monitoring, and MPA management to assess the success or failure of conservation efforts (Schadewell & Adams, 2021). eDNA can complement commercial fish stock assessments to ensure that fishing practices are sustainable (Hansen et al., 2018). Multiple studies comparing fish with commercial fishing gear have shown that eDNA reliably detects the same number or more species (Afzali et al., 2021; Fraija-Fernández et al., 2020; Thomsen et al., 2012; Thomsen et al., 2016). Spatiotemporal patterns in relative eDNA abundances could expand knowledge of migration patterns and life-history events, and improve fisheries genetics methods, but eDNA cannot provide information on the size, age or number of fish which are well established parameters in fisheries for estimating productivity and reproductive outputs (Hansen et al., 2018; Schadewell & Adams, 2021).

Invasive species can lead to native species extinctions, with propagating effects throughout ecosystems (Giakoumi et al., 2019). High connectivity in the marine environment makes eradicating invasive species challenging, and efforts have only been successful with early detections (Giakoumi et al., 2019). eDNA's sensitivity for detecting rare species improves the likelihood of early detections, and has already been incorporated into invasive species monitoring for Asian carps (*Hypophthalmichthys nobils* and *Hypophthalmichthys molitrix*) in the Great Lakes (Sepulveda et al., 2020). eDNA also presents a cost-effective tool for facilitating broad ecosystem-based management within MPAs, which often fail to achieve biodiversity objectives due to lack of adequate monitoring (Dunham et al., 2020). eDNA sampling within MPAs has revealed previously unidentified threatened species, such as the critically endangered European eel (*Anguilla Anguilla*)(Barco et al., 2022), and distinguished community composition between MPAs and adjacent areas (Boulanger et al., 2021; Gold et al., 2021). Enhanced knowledge of species distributions derived from eDNA surveys may inform improved selection of MPA sites in MSP (Bani et al., 2020). However, limited guidelines exist for best practices in MPA monitoring and MSP (Bani et al., 2020; Gold et al., 2021). Targeted guidelines will be necessary to address uncertainties related to specific objectives, *i.e.*, quantifying the likelihood of false positives to mitigate against MPAs being placed in areas without the species of conservation concern (Bani et al., 2020; De Brauwer et al., 2023).

Below I highlight some main areas where eDNA can improve our understanding of marine megafauna ecology, current limitations to these approaches, and how this can feed into more effective conservation for marine megafauna.

A) Expanding knowledge of species distributions across greater spatiotemporal scales

Both single-species and metabarcoding eDNA approaches can benefit conservation by providing better knowledge of marine megafauna distributions. Currently, population-level distributions are frequently lacking for marine megafauna, but these are essential to accurately designate marine

protected areas within MSP (Jorgensen et al., 2022). Further, climate change is altering marine megafauna distributions, necessitating monitoring to update distributions to adapt current MPA systems to ensure long-term effectiveness (van Weelden et al., 2021). eDNA is especially good at detecting species in low abundance, which has led to marine megafauna being detected in previously unknown areas despite being surveyed with other technologies, such as underwater or boat-based visual censuses and baited remote underwater videos (BRUVs) (Albonetti et al., 2023; Boussarie et al., 2018; Ip et al., 2021). eDNA methods have led to the first detections of critically endangered scalloped hammerheads (*Sphyrna lewini*) in Guam for five decades (Budd et al., 2021), and detections of the data deficient dwarf sperm whale (*Kogia sima*) in new areas (Juhel et al., 2021). Enhanced detection detectability has indicated habitat use in areas that overlap with high human impacts, higher occurrences in areas with management regulations and seasonal distributional changes within MPAs (Budd et al., 2023; Stewart et al., 2017; Valsecchi et al., 2023). As well as providing occurrence data, eDNA read counts can indicate spatiotemporal trends in the relative abundance of species (Jensen et al., 2023; Mariani et al., 2021), or life history stages, *i.e.*, putative spawning events for fishes (Di Muri et al., 2022).

eDNA studies in the marine environment have largely been constrained to small spatial scales, *i.e.*, a few square kilometres in coastal areas or ports (Fraija-Fernández et al., 2020; Jerde et al., 2019), but to inform us about marine megafauna distributions, it will be necessary to upscale surveys to cover the extent of their large distribution ranges (Sequeira et al., 2019). Accessibility to survey large spatial (hundreds of square kilometres) marine areas typically relies on the use of expensive dedicated research vessels (Truelove et al., 2022). However, these platforms are unlikely to be feasible for implementing widescale, long-term monitoring given that limited finances are a major barrier to effective conservation management (Gill et al., 2017). Method standardisation across different eDNA studies would enable comparisons between different study areas, and therefore increase the spatiotemporal scope of eDNA surveys (Takahashi et al., 2023). However, the opposite trend is currently happening, with increasingly more techniques being applied across different stages of the workflow (Takahashi et al., 2023).

Commercial vessels, such as ferries, navy and fishing vessels, represent alternative, cost-effective platforms to achieve wider spatiotemporal coverage. Typically, eDNA studies utilising these platforms have investigated spatial patterns in community composition along the route or temporal community changes by repeatedly surveying the same route (Jensen et al., 2023; Maiello et al., 2023; Valsecchi et al., 2021). However, inferring species distributions from these sampling platforms will require statistical modelling approaches to project occurrences into unsurveyed areas. Given that commercial vessels usually cover set routes, this could bias the species distributions recovered if the commercial vessels do not adequately cover the environmental range utilised by the species of interest (Tessarolo et al., 2014).

Commercial vessels have been widely utilised as opportunistic sampling platforms for visual surveys resulting in over 50% of marine occurrence records being collected along shipping routes (Hughes et al., 2021). Despite their widespread application, there is still limited guidance to execute effective sampling strategies to ensure that inherent biases associated with restricted sampling frames are limited (Tessarolo et al., 2014). Spatially biased sampling decreases the predictive accuracy of species distribution models, where species distributions are projected across space based on relationships between species occurrences and environmental predictors (Santini et al., 2021). These biases can result in negative consequences when models are implemented to prioritise spatial conservation areas. Omission errors can result in critical habitats for species not being included in reserve systems, while commission errors can lead to limited resources being delegated to protect areas without the species of conservation concern (Kramer-Schadt et al., 2013). Spatially restricted sampling can still produce species distribution models with comparable accuracy to unbiased strategies, *i.e.*, random or systematic sampling, as long as environmental variability is accurately captured (Tessarolo et al., 2014). However, there is no precedent for selecting networks of commercial vessel routes to ensure that environmental gradients are adequately covered. Further, eDNA surveys also require guidelines for effectively allocating point-based sampling locations across routes (Bani et al., 2020). Deducing sampling strategies to quantify and reduce biases associated with commercial vessels will be necessary to

leverage these platforms of opportunity effectively to sample at the spatiotemporal scales relevant for assessing marine megafauna distributions.

B) Improved understanding of biotic interactions to assess whole ecosystem-level responses to disturbance or management

Species are often considered in isolation within conservation management objectives, but failure to account for biotic interactions could be detrimental to the species of concern if important biotic links are not sustained, or have unintended consequences on other ecosystem components, *i.e.*, increase in predatory species may drive prey species to low biomass (Cashion et al., 2020; McDonald-Madden et al., 2016). However, biotic interactions are challenging to establish in speciose systems, or in the marine environment where direct observations are limited (Albouy et al., 2019; Bonebrake et al., 2018). The unique snap shots of community characterisations derived from eDNA metabarcoding enables further exploration of interactions and co-occurrences of species, enhancing understanding of how whole ecosystems respond to disturbances or management actions (DiBattista et al., 2020; Djurhuus et al., 2020). Significant spatial co-occurrence between species can infer biotic interactions whereby two species impacting each other's presence or abundance will share a non-random co-occurrence, although this can also be due to shared environmental requirements (Freilich et al., 2018). Diet metabarcoding studies have improved knowledge of marine megafauna trophic interactions, which traditionally rely on stomach content analyses or stable isotopes (Sonsthagen et al., 2020; Zhang et al., 2023). We can now couple this with eDNA metabarcoding to construct local community food webs (D'Alessandro & Mariani, 2021), and improve understanding of predator-prey dynamics (Visser et al., 2021), which is normally limited due to a lack of reliable prey data (Robertson et al., 2022; van Weelden et al., 2021). For example, Risso's dolphins and Cuvier's beaked whales exhibit spatial partitioning in the Azores, assumed to be driven by niche separation of prey species, but eDNA metabarcoding showed that preferred prey of both species were available across different depths and foraging zones (Visser et al., 2021). Foraging grounds are a focal point for marine megafauna MPA designation (Nelms et al., 2021), but climate change may impact their long-term effectiveness resulting from decoupling of trophic interactions due to

asynchronous range shifts (Bonebrake et al., 2018). Therefore, increased understanding of predator-prey dynamics is essential to understand how marine megafauna will respond to distribution shifts of prey, *i.e.* by switching to new prey species or by changing their distributions (van Weelden et al., 2021).

C) Non-invasive population genetics

Population genetics metrics are useful in MSP to identify priority managements units, through delineation of different populations and to incorporate connectivity and migrations into reserve networks (Kershaw et al., 2021). Obtaining these metrics for marine megafauna typically relies on collecting tissue samples using biopsy darts, which can be challenging for small, fast species and difficult to acquire permits for protected or endangered species due to the potential disturbance caused (Parsons et al., 2018; Sigsgaard et al., 2016). eDNA offers a non-invasive and cost-effective alternative for making intraspecific inferences, potentially expanding population genetics data across larger spatial scales and covering more species (Tsuji et al., 2020). So far, population genetics inferences from eDNA have been retrieved from mitochondrial haplotypes due to higher copy numbers in the environment and greater coverage in reference libraries (Sigsgaard et al., 2020). This approach has successfully delineated harbour porpoise populations which were previously treated as a single management unit, with one unit likely to be endangered and thus needing to be prioritised in further conservation efforts (Parsons et al., 2018). Advancements in high-throughput sequencing technologies may soon permit whole mitochondrial genome sequencing from environmental samples, potentially allowing unique individuals to be identified, providing more robust abundance estimates than read counts, and greater phylogenetic signals (Dugal et al., 2022; Jamy et al., 2020; Sigsgaard et al., 2020). Abundance metrics are critical for biodiversity monitoring to assess the conservation status of populations and to respond to population declines with conservation action prior to a species extinction (Hammond et al., 2021). The mitochondrial genome only represents a limited view of genetic variation as recombination is rare and only maternally inherited (Sigsgaard et al., 2020). Improved sampling techniques and technologies such as target capture may allow for increased use of nuclear

DNA as a source of eDNA in the future, thus allowing for more robust genetic diversity inferences (Adams et al., 2019; Sigsgaard et al., 2020).

1.5. Study Areas

In this thesis, I aim to explore spatiotemporal trends in marine community composition and biotic interactions, derived from eDNA metabarcoding. In particular, I explore whether knowledge of prey availability improves understanding of the fine-scale habitat use of marine mammals. The North Sea was chosen as a case study to investigate predator-prey dynamics and biotic interactions as waters surrounding the British Isles are subjected to above average warming (Belkin, 2009), increasing the likelihood of key prey species shifting distributions or declining, and increased asynchrony in the phenology of interacting species (Perry et al., 2005), necessitating long-term monitoring to understand and respond to ecosystem changes. I also aim to evaluate the effectiveness of commercial vessels as sampling platforms to upscale the spatiotemporal resolution of eDNA surveys, and how the resulting species distributions could improve marine spatial planning initiatives. Here, I selected the Mediterranean Sea as a study area as it hosts disproportionately high numbers of threatened marine predatory megafauna, and it is unclear whether the current Mediterranean MPA system is effectively capturing their requirements (Dulvy et al., 2021; Giménez et al., 2020).

1.5.1. Marine megafauna around the warming British Isles

Mid-latitude oceanic regions in the Northern hemisphere are predicted to experience climate change impacts imminently, especially in the form of species range shifts (Evans & Waggitt, 2020). The North Sea is a climate change hotspot that is currently warming at least three times faster than the average global rate, exacerbated by high population and industry densities (Belkin, 2009). Warming has also facilitated successful colonisation of invasive species, resulting in the highest ecological and economic costs across European Seas (Vilà et al., 2010). Over 24 species of cetaceans have been sighted around the British Isles, with the harbour porpoise (*Phocoena phocoena*) being the most abundant and widely distributed species, while other species are regionally or seasonally common such as minke whales (*Balaenoptera acutorostrata*), killer whales (*Orcinus orca*), white-beaked

dolphins (*Lagenorhynchus albirostris*), bottlenose dolphins (*Tursiops truncatus*), Risso's dolphins (*Grampus griseus*) and short-beaked common dolphins (*Delphinus delphis*) (Evans & Waggitt, 2020). The British Isles also host 38% of the global grey seal population and 30% of the eastern Atlantic harbour seal population (Carter et al., 2020). 22% of these marine mammal species are threatened, with accidental mortality in fishing gear or ship strikes, pollution and directed hunting being identified as leading threats (Temple & Terry, 2009). Over 50 elasmobranch species have been recorded off of the British Isles (Ellis et al., 2005), although local extinctions of the common skate complex (*Diputurus batis/D. intermedius*) and angel sharks (*Squatina squatina*) (Moore, 2023; Shephard et al., 2019), along with notable declines in larger bodied species such thornback rays (*Raja clavata*), tope sharks (*Galeorhinus galeus*) and spurdogs (*Squalus acanthias*) have occurred, following the widespread commencement of beam trawling in the 1960-70s (Sguotti et al., 2016). Similar trends have been observed in large predatory fish populations, with Jennings and Blanchard (2004) estimating that the biomass of 16-66 kg fish is 99% lower than predicted in the absence of fisheries. Bluefin tuna disappeared from the North Sea and the fishery collapsed, while Atlantic cod (*Gadus morhua*) stocks have declined by as much as 80% (Griffin, 2008).

Evidence of range shifts across the entire North Sea ecosystem have been observed, with important implications for conservation management. The dominant copepod species, *Calanus finmarchicus*, has been replaced with a warm-water associated species, *C. helgolandicus*, coupled with a 70% decline in copepod biomass (Alvarez-Fernandez et al., 2012; Reid, Edwards, et al., 2003). Two thirds of North Sea fishes have shifted their distributions northwards or to deeper depths, including commercially important species such as Atlantic cod (Perry et al., 2005). Common dolphins may be outcompeting white-beaked dolphins at their northern boundary, contracting the range of white-beaked dolphins who are a cold-water associated species (MacLeod et al., 2008). Range shifts could result in threatened species moving outside of MPA boundaries, requiring continuous monitoring to allow for adaptive responses, *i.e.*, moving or extending MPA boundaries as species move (Wilson et al., 2020).

Other species, including dominant forage fishes such as sandeels (*Ammodytes* sp.), herring (*Clupea harengus*) and sprat (*Sprattus sprattus*), have limited range shifting capabilities due to reliance on specific substrates for certain life stages, *i.e.*, overwintering or spawning, leaving them particularly vulnerable to warming temperatures (Frederiksen et al., 2011; Petitgas et al., 2013). These species represent the main food source for many North Sea predators, and previous declines have been linked to breeding failures in seabird populations (MacDonald et al., 2019; Wanless et al., 2005). Furthermore, temperature changes will affect seasonal timings of life history events which could result in mismatches across trophic levels, *i.e.*, peak productivity of copepods will occur earlier while sandeel larvae will emerge later which could result in limited prey availability for the larvae (Régnier et al., 2019). Shifting ranges or timings of abundance peaks, along with decreased abundance of species will all have repercussions across the entire ecosystem. Therefore, ecosystem-based management will be essential to detect early changes in species presence and abundance, and to monitor cascading effects throughout the ecosystem (Johnson et al., 2011). Given eDNA metabarcoding's unique ability to sample whole ecosystems simultaneously, this presents an effective method for analysing how North Sea ecosystems will respond to continued warming (Djurhuus et al., 2020).

1.5.2. The Mediterranean Sea – a hotspot of biodiversity and human impacts

Despite only constituting 0.82% of the global oceanic surface, the Mediterranean hosts up to 18% of all described marine species, including high numbers of endemic species and marine megafauna (Coll et al., 2010). Mediterranean marine ecosystems are at high risk of collapsing due to exposure to some of the highest cumulative human impacts globally (Halpern et al., 2019). There are twelve regularly occurring marine mammal species, seven of which are classified as threatened, comprising Mediterranean monk seals (*Monachus monachus*), short-beaked common dolphins (*Delphinus delphis*), sperm whales (*Physeter macrocephalus*), bottlenose dolphins (*Tursiops truncatus*), striped dolphins (*Stenella coeruleoalba*), Black Sea harbour porpoises (*Phocoena phocoena relicta*) and fin whales (*Balaenoptera physalus*) (Notarbartolo Di Sciara, 2016). The five remaining

species, Risso's dolphins (*Grampus griseus*), killer whales (*Orcinus orca*), rough-toothed dolphin (*Steno bredanensis*), long-finned pilot whales (*Globicephala melas*), and Cuvier's beaked whales (*Ziphius cavirostris*), are either data deficient or not assessed (Notarbartolo Di Sciara, 2016). 85 elasmobranch species also inhabit the Mediterranean Sea, including 20 species of top predatory sharks, with the three most commonly observed species being blue sharks (*Prionace glauca*), shortfin mako sharks (*Isurus oxyrinchus*) and thresher sharks (*Alopias vulpinas*) (Bradai et al., 2012; Ferretti et al., 2008). 40% of these elasmobranchs are classified as threatened, leading to the Mediterranean being acknowledged as an extinction risk hotspot for elasmobranchs, with populations experiencing 96-99% declines in abundance (Bradai et al., 2012; Dulvy et al., 2014; Ferretti et al., 2008). The Mediterranean Sea includes important spawning areas for large, commercially important predatory fish such as bluefin tuna (*Thunnus thynnus*), swordfish (*Xiphias gladius*) and albacore (*Thunnus alalunga*), which are of high conservation priority following overexploitation (De Juan & Lleonart, 2010).

The western basin is assumed to have the highest diversity of marine megafauna, corresponding with highest productivity (Coll et al., 2012). The Alboran Sea and offshore waters in the western basin are important for marine mammals while coastal waters off Tunisia and Libya contain highest diversity of elasmobranchs (Coll et al., 2010). These areas overlap with some of the highest human impacted areas, with climate change, fishing, shipping traffic, pollution and habitat modification and destruction all potentially contributing to the threatened statuses of these megafaunal species (Micheli et al., 2013). Robust predictions of species distributions are currently hampered by biases in occurrence data, with far greater sampling effort in the north-western Mediterranean Sea and coastal regions (Levin et al., 2014; Ramírez et al., 2022). Given the highly threatened statuses of Mediterranean marine megafauna, increased research effort to address these data gaps is needed to ensure that conservation efforts are concentrated in the most effective areas.

Expanding MPA coverage in the Mediterranean Sea has been the main strategy employed to conserve biodiversity, although Mediterranean

coverage remains lower (6.01%) than global coverage (8.16%) (Claudet et al., 2020). Geographical coverage of protected areas is strongly biased towards the north-western basin and coastal areas, mirroring the areas with highest sampling effort (Claudet et al., 2020). There are also pervasive taxonomic biases, with the two largest MPAs, the Pelagos Sanctuary and the Cetacean Migration Corridor, both being designated for marine mammals (Abalo-Morla et al., 2022). Even so, this does not offer marine mammals adequate protection due to spatially heterogenous distributions and diverse habitat requirements. For example, the Hellenic Trench in the Aegean Sea has been identified as an 'Important Marine Mammal Area' (IMMA) for deep diving species, sperm whales and Cuvier's beaked whales, but remains unprotected due to conflicts with the oil and gas industry (Notarbartolo di Sciara & Hoyt, 2020). The Tunisian plateau is the least protected ecoregion in the Mediterranean Sea despite having the highest diversity of elasmobranch species as well as constituting important breeding and nursery grounds for them (Bradai et al., 2012; Claudet et al., 2020). Perhaps of even greater concern is that 95% of areas covered by MPAs have no regulations in place which distinguish them from outside of MPAs (Claudet et al., 2020). These shortcomings stem from limited financial resources to implement monitoring within MPAs (Amengual & Alvarez-Berastegui, 2018). Therefore, further work to expand MPA coverage and reduce biases to meet global targets is required, alongside increased monitoring efforts utilising novel, cost-effective tools such as eDNA metabarcoding.

1.6. Thesis aims and outline

This thesis contributes to understanding of how eDNA approaches can be incorporated into marine megafauna monitoring to facilitate improved conservation outcomes across four papers (Figure 1.3). Firstly, in Chapters 2 and 3, I aim to identify previously unknown spatiotemporal trends in community composition and interactions in a newly designated MPA in the Moray Firth, North-east Scotland, employing eDNA metabarcoding. I focus on cetacean habitat use in relation to key prey species, as inadequate prey data currently limits our understanding of predator-prey dynamics (NatureScot, 2020). Improved understanding is necessary to predict how marine megafauna will respond to changes in the distribution and abundance of their

prey species stemming from fishing or climate pressures (Young et al., 2015). In Chapter 4, I establish effective sampling strategies to reduce biases associated with spatially constrained sampling frames, such as commercial vessel routes. Commercial vessels allow eDNA surveying across greater spatiotemporal scales, relevant to the large ranges of marine megafauna (Valsecchi et al., 2021), but will only effectively improve our knowledge of species distributions if biases are limited (Santini et al., 2021). Accurate species distribution models are required in MSP decision making tools to effectively prioritise spatial conservation areas, such as MPAs (Sofaer et al., 2019). Finally, in Chapter 5, I investigate whether priority areas for taxonomic groups with heterogenous distributions are more effectively captured in prioritisation solutions including all taxa simultaneously or when taxonomic groups are considered separately. eDNA approaches will improve taxonomic coverage of species distributions and permit increased use of taxon-specific objectives in MSP (Bani et al., 2020).

Objective 1: Investigate fine-scale habitat use of cetaceans in relation to key prey species in a newly designated MPA.

MPAs protecting marine megafauna foraging grounds may be changeable over time as prey species respond to climate change by shifting distributions, but megafauna and prey are rarely studied synergistically to investigate how predatory megafauna respond to changing prey distributions (Young et al., 2015). In Chapter 2, I aim to improve understanding of spatiotemporal availability and abundance of key prey species in relation to cetacean habitat use, in a newly designated MPA in the Moray Firth to protect their foraging grounds. Baseline community characteristics prior to the implementation of conservation actions are essential to evaluate the effectiveness of future management actions by comparing changes in community composition (Giakoumi et al., 2018). Therefore, I also assess overall spatiotemporal trends across species from all taxonomic levels of the community and identify other species of conservation concern.

Objective 2: Evaluate drivers of species interactions and identify key ecosystem components in the Moray Firth community.

Understanding how communities interact is essential to evaluate the vulnerability of species and ecosystems to disturbances from anthropogenic pressures and climate change (Bonebrake et al., 2018). In Chapter 3, I evaluate spatial and temporal changes in species interactions, and identify highly connected species to assess the vulnerability of the Moray Firth ecosystem to future climate change impacts.

Objective 3: Assess the effectiveness of different sampling strategies to reduce biases associated with spatially constrained sampling platforms.

Commercial vessels represent cost-effective sampling platforms that enable coverage of greater spatiotemporal scales and access to hard-to-reach pelagic regions (Valsecchi et al., 2021). Sampling is usually constrained to set routes which could bias the species distributions retrieved, but limited guidance exists to implement sampling strategies to reduce these inherent biases. In Chapter 4, I aim to quantify biases associated with spatially restricted sampling frames, such as commercial vessel routes and consider how different coverage of environmental variability and species composition impacts the retrieval of accurate species distributions.

Objective 4: Compare multi-taxon and taxon specific spatial conservation priorities for marine megafauna.

Marine megafauna include diverse taxa with different life histories, threats and resource needs (Jorgensen et al., 2022; Nelms et al., 2021). Typically, conservation practitioners incorporate all biodiversity data across taxa into spatial planning prioritisations, but this may fail to account for taxon-specific requirements adequately. In Chapter 5, I evaluate differences between spatial conservation areas from prioritisation solutions for individual taxa and all taxa, to identify whether each taxon requires specific conservation areas. I compare how my prioritisation solutions differ from the current Mediterranean Sea MPA system.

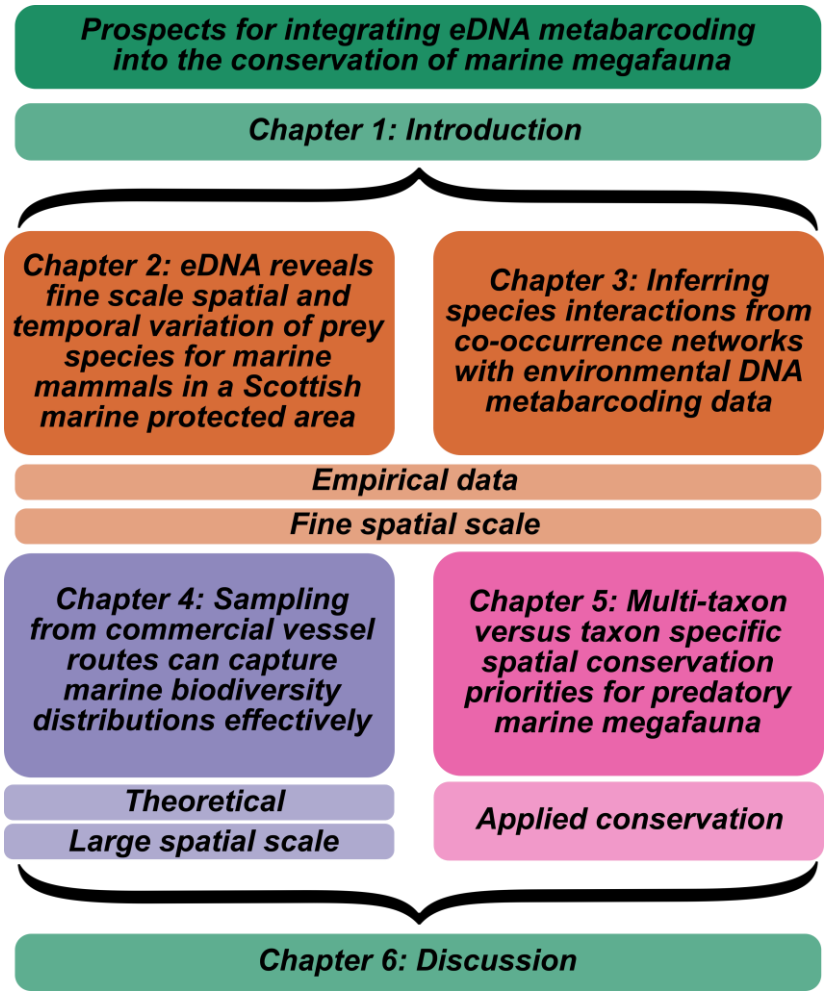


Figure 1.3. Schematic diagram of the thesis, highlighting the structure and main themes.

Chapter Two - Environmental DNA reveals fine scale spatial and temporal variation of prey species for marine mammals in a Scottish marine protected area

Elizabeth Boyse[1], Kevin P. Robinson[2], Maria Beger[1], Ian M. Carr[3], Morag Raynor[3], Elena Valsecchi[4], Simon J. Goodman[1]

[1]School of Biology, University of Leeds, Leeds, United Kingdom
[2]Cetacean Research & Rescue Unit (CRRU), PO Box 11307, Banff, AB45 3WB, Aberdeenshire, United Kingdom
[3]Leeds Institute of Medical Research at St James's, St James's University Hospital, Leeds, United Kingdom
[4]Department of Environmental and Earth Sciences, University of Milano-Bicocca, Milan, Italy

2.0. Abstract

Foraging grounds for marine mammals are popular focal points for marine protected area (MPA) implementation, but may be temporally dynamic, requiring continuous monitoring to infer prey species availability and abundance. Marine mammal distributions are assumed to be driven by their prey in these foraging areas, but limited understanding of prey distributions often prevents us from exploring how shifting prey availability impacts both seasonal and long-term distributions of marine mammals. Environmental DNA (eDNA) metabarcoding could enhance our understanding of marine mammal habitat use in relation to their prey through simultaneous monitoring of both. However, eDNA applications focused on marine mammals or predator-prey dynamics have been limited so far. In this study, we assess spatial and temporal changes in the availability and abundance of minke whale (*Balaenoptera acutorostrata*) prey species in a newly established MPA to protect their foraging grounds, employing eDNA metabarcoding. We recovered 105 molecular operational taxonomic units (OTUs) from marine vertebrates using two primer sets targeting 12S and 16S genes, along with 112 OTUs from a broader eukaryotic primer set targeting 18S rRNA. Overall, key forage fish prey species, sandeels and clupeids, were the most abundant teleost fishes detected, although their availability varied temporally, and with distance from shore. We also found clear spatial partitioning between coastal bottlenose dolphins and the more pelagic minke whales and harbour porpoises, which aligned with the availability of their key prey species. This study demonstrates the utility of eDNA to detect spatiotemporal trends in the occurrence and abundance of cetaceans and their prey species, furthering our

understanding of their fine-scale habitat use within MPAs. Future long-term monitoring of predator-prey dynamics with eDNA could improve our ability to predict climate-induced shifts in foraging grounds and enhance rapid responses with appropriate management actions.

2.1. Introduction

Over a quarter of European marine mammals are threatened as a result of human activities including overfishing (both targeted and bycatch), pollution, changes in prey dynamics and habitat degradation (Temple & Terry, 2009). These threats are jeopardising important ecosystem functions that marine mammals provide, such as top-down control, nutrient cycling and ecosystem engineering (Estes et al., 2016). Marine protected areas (MPAs) are the main tool implemented to protect marine mammals from human impacts, but European MPAs are currently too small and disjointed to provide adequate protection for such far ranging mammals (Bearzi & Reeves, 2021). Marine mammal foraging grounds are often concentrated in small spatial areas relative to their whole distribution ranges, making them a popular focal point for MPA implementation (Notarbartolo di Sciara et al., 2016). However, these foraging grounds are likely to be dynamic across time as common prey species are sensitive to seasonal and long term environmental change, requiring an adaptive, responsive approach, or risk MPAs becoming redundant (Cashion et al., 2020). Therefore, coupled monitoring of marine mammals and their prey species will be essential to ensure long-term effectiveness of management measures.

Drivers of cetacean distributions are assumed to have a hierarchical structure, whereby distributions at fine spatiotemporal resolutions (10s of kilometres) are best described with prey availability, although at broader scales (100s of kilometres), ocean currents or persistent water masses become more important (Mannocci et al., 2017). However, we rarely have prey data at complementary spatial and temporal scales available to incorporate into predictive distribution models of marine mammals, so environmental proxies are often used as an alternative (Mannocci et al., 2020; Pendleton et al., 2020). Including prey occurrence or multispecies prey hotspots to model cetacean distributions has improved understanding of their seasonal dynamics and

heterogenous habitat use (Pendleton et al., 2020; Szesciorka et al., 2023). However, prey may also be more widely available than areas covered by cetaceans due to their reliance on environmental features to form dense prey aggregations, interspecific competition, or targeted prey selection, *i.e.*, mature individuals (Cox et al., 2018; Visser et al., 2021). Uncertainty over climate-induced changes in the abundance, distributions and temporal availability of prey limits the predictive accuracy of future marine mammal distributions, especially at finer spatial scales (Silber et al., 2017). Marine mammals are expected to respond by prey switching or shifting their distributions to track prey species, which could result in increased overlap with threats such as ship strikes or fisheries bycatch (Agardy et al., 2019). Recent advances in environmental DNA metabarcoding offers a unique opportunity to simultaneously monitor cetaceans and their prey which could enhance our understanding of their dynamics and permit long-term monitoring to track changes in their distributions (Székely et al., 2021).

The uptake of eDNA as a monitoring tool for marine mammals has been comparatively slow, with only 21 papers published between 2012-2021, compared to over 50 papers focused on marine teleost fishes (Suarez-Bregua et al., 2022; Takahashi et al., 2023). This delay is likely due to early mixed success in detection rates of cetaceans (Pinfield et al., 2019), although new primer sets targeting marine vertebrates are now helping to resolve this problem (Valsecchi et al., 2020). eDNA can expand the spatiotemporal scope of marine mammal monitoring where visual surveys or passive acoustic monitoring are infeasible, *i.e.*, no visual detections at night time, or in adverse weather conditions and no acoustic detections for cetaceans who do not vocalise or have unknown vocalisations (Baumgartner et al., 2019; Valsecchi et al., 2021). So far, single species eDNA approaches have improved understanding about distributions of rare and threatened species, *i.e.*, dwarf sperm whales (*Kogia sima*) and Mediterranean monk seals (*Monachus monachus*) (Juhel et al., 2021; Valsecchi et al., 2023), and species that are challenging to detect with conventional survey techniques, *i.e.*, beaked whales (Hooker et al., 2019). They have also provided insights into population genetics with important management consequences (Parsons et al., 2018). Metabarcoding studies have uncovered previously unknown

marine mammal diets, and can be utilised to detect future dietary changes (Sonsthagen et al., 2020; Zhang et al., 2023). eDNA metabarcoding can also reveal spatiotemporal availability of cetacean prey species or detect co-occurrences between cetaceans and their prey (Djurhuus et al., 2020; Visser et al., 2021), but few studies have harnessed eDNA to elucidate marine mammal trophic interactions to date (Székely et al., 2021).

The Southern Trench MPA in the Moray Firth, north-east Scotland, has recently been designated to protect important seasonal foraging grounds for minke whales (*Balaenoptera acutorostrata*) (NatureScot, 2020). Their diets within the Moray Firth consist predominantly of sandeels (*Ammodytes* sp.), sprat (*Sprattus sprattus*) and herring (*Clupea harengus*), none of which are commercially fished in the area resulting in limited knowledge on their spatiotemporal availability and abundance (Pierce et al., 2004). The area also overlaps with bottlenose dolphin (*Tursiops truncatus*) and harbour porpoise (*Phocoena phocoena*) distributions, both requiring protection under Annex II of the European Habitats Directive. Harbour porpoises also rely predominantly on sandeels as well as whiting (*Merlangius merlangus*), whilst bottlenose dolphins target Gadidae species including cod (*Gadus morhua*), saithe (*Pollachius virens*) and whiting, and salmonids (*Salmo* sp.) (Santos et al., 2004; Santos et al., 2001). In this study, we aim to characterise the overall vertebrate and broader eukaryotic community composition within this MPA and identify the availability of key prey species for cetaceans. We assess temporal changes in community composition during the minke whale foraging season (June to October) to evaluate the overall strength of seasonality, and to detect changes in the availability of suitable prey species. We also examine how the community changes with distance from the shore, as spatial partitioning between different age-classes of minke whales has been observed previously which could indicate dietary differences (Robinson et al., 2023). We expect that changes in minke whale habitat use relate to spatiotemporal changes in the availability of different prey species, and potentially signal prey switching. This work will provide an essential monitoring baseline in a newly established MPA (Dunham et al., 2020), with the potential to develop longer-term monitoring protocols to improve ecological knowledge and contribute to management actions.

2.2. Methods

2.2.1. Study Area

The Moray Firth is the largest estuarine embayment in North-East Scotland, covering 5,230 km^2, and opens into the wider North Sea basin (Figure 2.1a)(Tetley et al., 2008). It is an internationally recognised area of outstanding biological importance, with the 'inner' section designated as a Special Area of Conservation (SAC) under the Habitats Directive (92/43/EEC), as well as the newly designated Southern Trench MPA in the south-eastern 'outer' firth (Cheney et al., 2013; NatureScot, 2020). The Southern Trench is an enclosed seabed basin that reaches depths in excess of 250 m, the deepest portion of the firth (Holmes et al., 2004). Two oceanographic features govern water movement. Firstly, cold water is transported into the firth from the northern North Sea via the Dooley current, whilst a warm plume ebbs out from the inner firth and rivers discharging into the firth which is associated with greater primary productivity levels (Tetley et al., 2008).

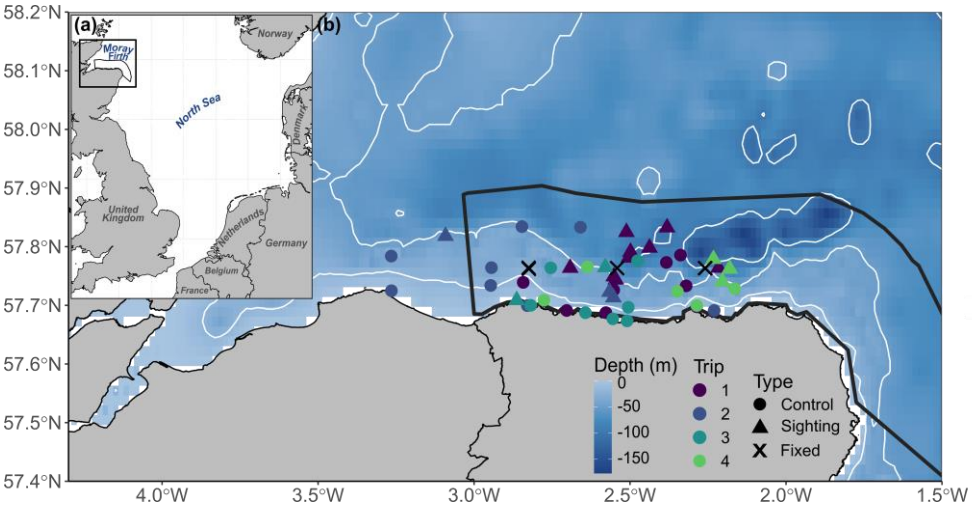


Figure 2.1. a) Map displaying the location of the Southern Trench MPA and the Moray Firth within the North Sea basin, indicated by the black rectangle. b) Bathymetric map showing the three fixed sampling points (crosses) where samples were collected monthly, and control (circle) and sighting (triangle) samples, coloured by sampling month.

2.2.2. Sample Collection

We collected seawater samples between June and October 2021, corresponding to the months when minke whales are most abundant within

the firth (Robinson et al., 2009). Samples were collected each month over four separate sampling trips to capture spatiotemporal trends in the availability and abundance of the main prey species, and the overall community trends (Appendix Table A2.1). To quantify temporal differences in community composition, we collected samples from three fixed points located five nautical miles offshore on each trip (total 12 fixed samples) (Figure 2.1b). The bathymetry of the westerly fixed point was the shallowest at 33 m, whilst the easterly fixed point was on the edge of the Southern Trench at 118 m. The middle-fixed point was at 104 m depth, and a known hotspot for foraging minke whales. Samples were also collected when minke whales were visually sighted to assess whether minke whales needed to be in close proximity to detect their DNA and as an indicator to where their prey species are also suspected to be in high abundance (total 18 sighting samples). Finally, in the absence of minke whales, random samples were collected across the entire study area to capture as much environmental variability as possible to determine environmental drivers of community composition (total 30 control samples) (Figure 2.1c). This design resulted in a total of 60 11-litre seawater samples. A further three field-controls (blanks) were collected employing the same sampling equipment but replacing seawater with tap water to detect potential sources of contamination. All sampling was carried out from an 8-metre rigid hulled vessel using weighted buckets deployed at four metres depth with the water subsequently transferred to 'Bags in the Boxes' for storage and transport to the laboratory (Valsecchi et al., 2021). Reusable field equipment was cleaned with 50% bleach, left to soak for 30 minutes and then washed thoroughly with tap water in between sample collection.

2.2.3. Sample Filtration

We filtered samples between one and six days after collection, with an average delay of 1.8 ± 1.1 s.d. days. 49% of samples were filtered the day after collection, and 89% of samples were filtered within three days of collection. Each 11-litre seawater sample was split into three replicates for filtering: two 4-litre replicates and one 3-litre replicate. For ten of the samples, between 10 and 10.8 litres was filtered as a result of filters being saturated before 11 litres had been reached. Samples were filtered using either the

BioSart® 100 filtration system (Sartorius) or Nalgene™ reusable analytical funnels (Thermo Scientific) with either the Fisherbrand™ FB70155 Pump or Welch™ WOB-L Piston Dry Vacuum Pump. Samples were filtered through cellulose nitrate filters of 0.45 µm pore size. Immediately after filtration, filter papers were wrapped and stored in aluminium foil at -20°C until DNA extraction.

2.2.4. DNA extraction, amplification and sequencing

DNA extractions were carried out in a dedicated molecular laboratory, and bench surfaces were cleaned with 50% bleach followed by deionised water. Prior to extraction, filters were cut into small pieces to increase their surface area using sterile scalpels and tweezers. Extractions were carried out using the Qiagen DNeasy PowerSoil Pro Kit following the manufacturer's protocol.

Pre- and post-PCR analysis were carried out in separate rooms. All PCRs were prepared in a class II microbiology safety cabinet that was cleaned with 50% bleach and illuminated with ultraviolet light for 20 minutes between sample preparations. We amplified marine vertebrate DNA with two primer sets; MarVer1 which amplifies a ~202 bp sequence from the mitochondrial 12S rRNA gene, and MarVer3 which amplifies a ~245 bp sequence from the mitochondrial 16S rRNA gene (Valsecchi et al., 2020). These primers can resolve most taxa to species level across all marine vertebrates, inclusive of marine mammals, elasmobranchs and teleost fishes. Our primers were designed with six to eight random nucleotides, an eight-base pair Illumina barcode tag and the amplification primer sequence from 5' to 3' (Bohmann et al., 2022). PCR reactions were 20 µL volume containing 0.025 u/µL GoTaq® Hot Start Polymerase (Promega), 5X Green GoTaq® Flexi Buffer (Promega), 1mM or 2mM $MgCl_2$ (Promega) for MarVer1 or MarVer3 respectively, 0.2 mM dNTPs (Promega), 0.2 µM each of the reverse and forward primer, and UltraPure™ distilled water (Invitrogen). Annealing temperatures for MarVer1 were 54/55/56°C for 10/10/18 cycles, and 58/57/56/55°C for 8/10/10/10 cycles for MarVer3. Both had an initial denaturation time of four minutes at 95°C, and a final elongation of five minutes at 72°C, then per cycle, 30 seconds at 95°C, 10 s at annealing temperature and 20 s at 72°C for MarVer1, and 30 s at 95°C, 2 s for first 8 cycles and 10 s for remaining cycles

at annealing temperature, and 20 s at 72°C for MarVer3. The three PCR replicates per 11-litre water sample were then pooled for individual primer sets. Samples were cleaned up and primer dimers removed with AMPure beads (Beckman Coulter). We then checked that the fragment size distributions were as expected with an Agilent TapeStation, followed by quantification with a Qubit fluorometer. Samples for each primer set were pooled in equimolar ratios to create two Illumina NEBNext Ultra DNA libraries, for MarVer1 and MarVer3 respectively. The libraries were sequenced separately on an Illumina MiSeq Sequencer with 150 bp paired-end lanes at the University of Leeds Genomics Facility, St James' Hospital.

We also amplified DNA using a general eukaryotic primer set, 1391F and EukBr, targeting a ~260 bp region from 18S rRNA to detect zooplankton and other invertebrates (Amaral-Zettler et al., 2009). Primers included sequences homologous to Illumina sequencing adapters appended to the 5' end. PCR reactions were 25 µL consisting of 0.025 u/µL GoTaq® Hot Start Polymerase (Promega), 5X Green GoTaq® Flexi Buffer (Promega), 2 mM $MgCl_2$ (Promega), 0.2 mM dNTPs (Promega), 0.2 µM each of the reverse and forward primer, 1.6 µM mammal blocking primer and UltraPure™ distilled water (Invitrogen). Thermocycling conditions included an initial denaturation at 94°C for three minutes, 35 cycles at 94°C for 45 s, 65°C for 15 s, 57°C for 30 s, 72°C at 90 s, and a final elongation at 72°C for ten minutes (Sawaya et al., 2019). The three PCR replicates per 11-litre water sample were then pooled and cleaned up as above. The final sequencing libraries were generated using a second PCR in which Nextera XT indexed adaptor sequences were used as primers, such that each sample was uniquely indexed. The PCR reaction consisted of 5 µL of the pooled amplicons, 25 µL of the NEBNext Q5 Hot Start HiFi PCR Master Mix, 10 µL of water and 5 µL each of the appropriate Nextera XT Index Primer 1 and Primer 2. Thermocycling conditions included an initial denaturation at 95°C for three minutes followed by 8 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s and a final hold at 72°C for five minutes. The products were then cleaned again with AMPure beads to remove adapter dimers and free adaptor oligos and checked for the presence of the correctly formed libraries by running on a D1000 tape of a TapeStation followed by quantification with Qubit fluorometry. The libraries

were then combined to create an equimolar pool that was sequenced on a MiSeq 250 bp paired end lane with V2.0 chemistry and 15% PhiX control library to aid base calling.

2.2.5. Bioinformatics

A description of the full bioinformatics workflow is available in Valsecchi et al. (2020) and at http://www.dna-leeds.co.uk/eDNA/. Firstly, sample de-multiplexing was performed with 'Deconvolute' by screening for expected index pair combinations. Read pairs were then quality checked to remove spurious primer combinations and trimmed to remove low quality sequences until the average PHRED score for a contiguous 10 bp region was greater than 25, before being combined to form a single sequence. Only one occurrence of each unique sequence per sample was retained to reduce the likelihood of PCR duplicates or chimaeras. A counts matrix was created with Aggregator by counting the number of instances for each unique amplicon sequence per sample. Amplicon sequences were blasted against the Genbank Nucleotide Database to identify the taxonomic origin of sequences, and the top ten hits were linked to the sequence if they were more than 70% homologous. Full taxonomic hierarchy for species names was obtained for the GenBank hit sequences from a Microsoft SQL server instance of the ITIS taxonomy database. When taxonomic information was found, the name and taxonomy of the best hit was retained. Finally, FilterrRNAData clustered sequences into molecular operational taxonomic units (OTUs) using a 98% threshold of homology to the GenBank hit sequence at species level for our vertebrate primer sets, and a 95% threshold of homology at family level for our 18S primer set.

2.2.6. Contamination control

For the purpose of this study, non-marine OTUs and off target OTUs, *i.e.,* invertebrates detected with our vertebrate primer sets, were removed. Non-marine contaminate OTUs were primarily comprised of *Homo sapiens* and agricultural species such as *Bos taurus, Canis lupus* and *Sus scrofa,* although we also detected local terrestrial wildlife including *Capreolus capreolus* (roe dear), *Erinaceus europaeus* (hedgehog) and *Myodes glareolus* (voles), potentially originating from river inflow. For our vertebrate primers, we

identified amplicon sequences which had been assigned to marine species not previously recorded in the North Sea. We established whether native congenerics or family members were known to be present in the North Sea, and if so, compared amplicon sequences to assess whether there was enough differentiation in our amplicon regions to distinguish between the non-indigenous and native congenerics or family members(Valsecchi et al., 2021). Non-native species reads were either confirmed as a potential invasive species, reassigned to a native congeneric or family group, or occurred in low abundance (<10 reads) so were removed from the dataset. A full list of species that were removed or merged is provided in the supplementary appendix (Appendix Tables A2.2 and A2.3).

The most likely source of contamination in this study was from crossover contamination between samples (Calderón-Sanou et al., 2020). To reduce this background contamination, we calculated two times the standard deviation for the proportion of each OTU in the field blanks, and then subtracted this from the specific OTU proportion in each sample, following a similar approach to Kelly et al. (2018). This method produced the most sensible results with our dataset, *i.e.*, removed known contaminate OTUs such as *Homo sapiens*, among different methods trialled. To facilitate abundance comparisons, we standardised read counts using an OTU-specific index under the assumption that amplification efficiency is consistent for each OTU regardless of community composition. This scaling allows for the comparison of within OTU abundance across samples (Kelly et al., 2019). The OTU-specific index was made by converting read counts into proportions then dividing the maximum proportion for every OTU from that OTU's proportion per sample resulting in an index between 0-1 for each OTU (Kelly et al., 2019). For our vertebrate primer sets, we created an ensemble OTU index by averaging across the indices per sample for individual primer sets (Kelly et al., 2019).

2.2.7. Community analysis

Initial descriptions of community composition and visualisations of the data were conducted using the 'Phyloseq' R package version 1.38.0 (McMurdie & Holmes, 2013). We analysed temporal trends in the community through

partitioning the data by sampling month, and spatial trends by dividing samples into categories based on their distance from the shore; near (<1.2 km), middle-near (between 2.5 and 7 km), middle-far (between 7 and 10 km), and far (> 10 km up to 16.1 km) (Drummond et al., 2021; Fraija-Fernández et al., 2020). These categories were defined based on previous observations that juvenile minke whales were most frequently sighted near shore (<2.5 km), whilst adults more frequently occur offshore (>10 km) (Robinson et al., 2023). Community statistics were calculated using the 'Vegan' R package version 2.5-7, and using abundance data with the OTU-specific index applied (Dixon, 2003). Alpha diversity of samples was estimated with the Shannon-Weiner index and compared across the sampling months and with distance from shore using Kruskal-Wallis tests followed by pairwise Wilcoxon rank sum tests. We transformed our abundance matrix into a Bray-Curtis dissimilarity matrix to compare beta-diversity between sampling months and distance from shore categories. We assessed differences between communities with non-metric multidimensional scaling (NMDS) and tested for significant differences between communities using permutational multivariate analysis of variance (PERMANOVA). Homogeneity of group dispersion is an assumption of PERMANOVA, so this was first assessed using the 'betadisper' function. We also performed pairwise multilevel comparisons to further evaluate where differences between communities existed with the 'PairwiseAdonis' R package version 0.4 (Martinez Arbizu, 2017). We identified indicator species for the different months and distance from shore categories using the 'multipatt' function, with 999 permutations, from the 'Indicspecies' R package version 1.7.12 (Cáceres & Legendre, 2009). This includes two components; A quantifies the specificity of the species as an indicator for the group where A=1 means that species only appears in that group, and B quantifies the sensitivity of the species as an indicator for that group where B = 1 means that all sites within that group include the species (Cáceres & Legendre, 2009).

We evaluated environmental covariates associated with changes in community composition using multiple regression in distance matrices (MRM), employing Spearman correlation ranked distances and 10,000 permutations with the 'MRM' function from the 'Ecodist' R package version

2.0.9 (Goslee & Urban, 2007). We used four environmental predictors, bathymetry (m), sea surface temperature (°C), chlorophyll *a* (mg/m^3) and distance from shore (m). Bathymetry was obtained from the General Bathymetric Chart of the Oceans (GEBCO) at a 0.004x0.004° resolution (GEBCO Compilation Group, 2020), and SST (https://doi.org/10.48670/moi-00156) and chlorophyll *a* (https://doi.org/10.48670/moi-00289) were downloaded from the E.U. Copernicus Marine Service Information at 0.02°x0.02° and 0.01°x0.01° resolution respectively (Høyer & Karagali, 2016). Values for each predictor were then extracted from individual sampling points. Distance from shore was calculated using the 'dist2Line' function from the 'Geosphere' R package version 1.5-14 (Hijmans, 2021). Prior to running MRM, we tested for collinearity among predictor variables with Spearman's rho rank correlation coefficient from the 'Hmisc' R package version 4.7-1 (Harrell Jr, 2022). Bathymetry and distance from shore were highly correlated (rho = -0.81, $p < 0.001$) so only distance from shore was retained for MRM. We created distance matrices for each environmental predictor using Euclidean distance, as well as a distance matrix for the distance between sites, and the number of days between sample collection. Our response variable was a Bray-Curtis dissimilarity matrix of community composition derived from either our vertebrate ensemble OTU index or eukaryotic OTU index. Initially, maximal models were created using all terms, and then reduced to a minimal model with only significant terms retained.

2.3. Results

2.3.1. Composition of vertebrate taxa detected

Following contamination removal procedures, our final datasets for vertebrate primers MarVer1 and MarVer3 comprised of 2, 880, 775 and 4,013,997 sequences which were assigned to 56 and 80 operational taxonomic units (OTUs) respectively, clustered at species level where possible (Table 2.1). 31 OTUs were detected by both markers, whilst 25 were unique to MarVer1, and 49 to MarVer3 (Figure 2.2). Over 90% of reads from both markers were assigned to teleost fishes from 22 families for MarVer1 and 31 families for MarVer3 (20 shared families, 2 unique to MarVer1 and 11 unique to MarVer3) (Figure 2.3a). This list included species of potential conservation

interest, such as rare taxa (bluefin tuna, *Thunnus thynnus*), critically endangered species (European eel, *Anguilla anguilla*) and invasive species (pink salmon, *Oncorhynchus gorbuscha*). Mammalia had the second highest proportion of reads, including all four marine mammals common in the study area, minke whales (*Balaenoptera acutorostrata*), grey seals (*Halichoerus grypus*), harbour porpoises (*Phocoena phocoena*), and bottlenose dolphins (*Tursiops truncatus*), as well as some less commonly sighted vagrants, fin whales (*Balaenoptera physalus* – MarVer1 only), white-beaked dolphins (*Lagenorhynchus albirostris*- MarVer3 only) and Sowerby's beaked whales (*Mesoplodon bidens*).

Table 2.1. Comparison of the taxa composition across two vertebrate primer markers, MarVer1 and MarVer3.

Class	MarVer1 – total reads (% of reads)	MarVer1 – OTUs detected	MarVer3 – total reads (% of reads)	MarVer3 – OTUs detected
All	2,880,775	56	4,013,997	80
Teleosts	2,763,655 (96 %)	41	3,653,206 (91 %)	63
Mammals	89,835 (3 %)	5	294,778 (7 %)	6
Chondrichthyes	5726 (<1 %)	3	62,980 (1.5 %)	5
Birds	21,113 (<1 %)	6	3033 (<1 %)	6
Cephalaspidomorphi (lamprey)	447 (<1%)	1	0	0

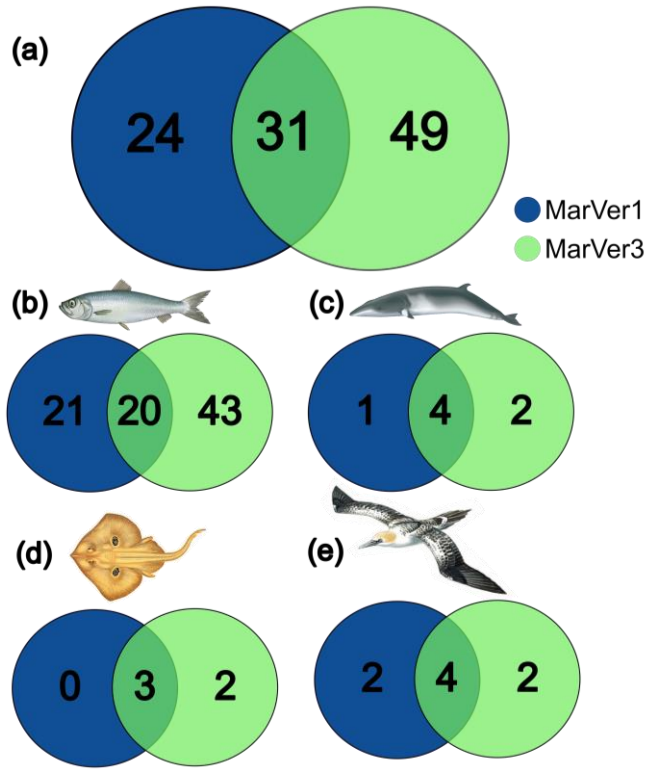


Figure 2.2. Venn diagrams showing overlap between (a) all OTUs, (b) teleosts, (c) mammals, (d) chondrichthyes and (e) birds detected by both vertebrate primer sets, MarVer1 (blue) and MarVer3 (green).

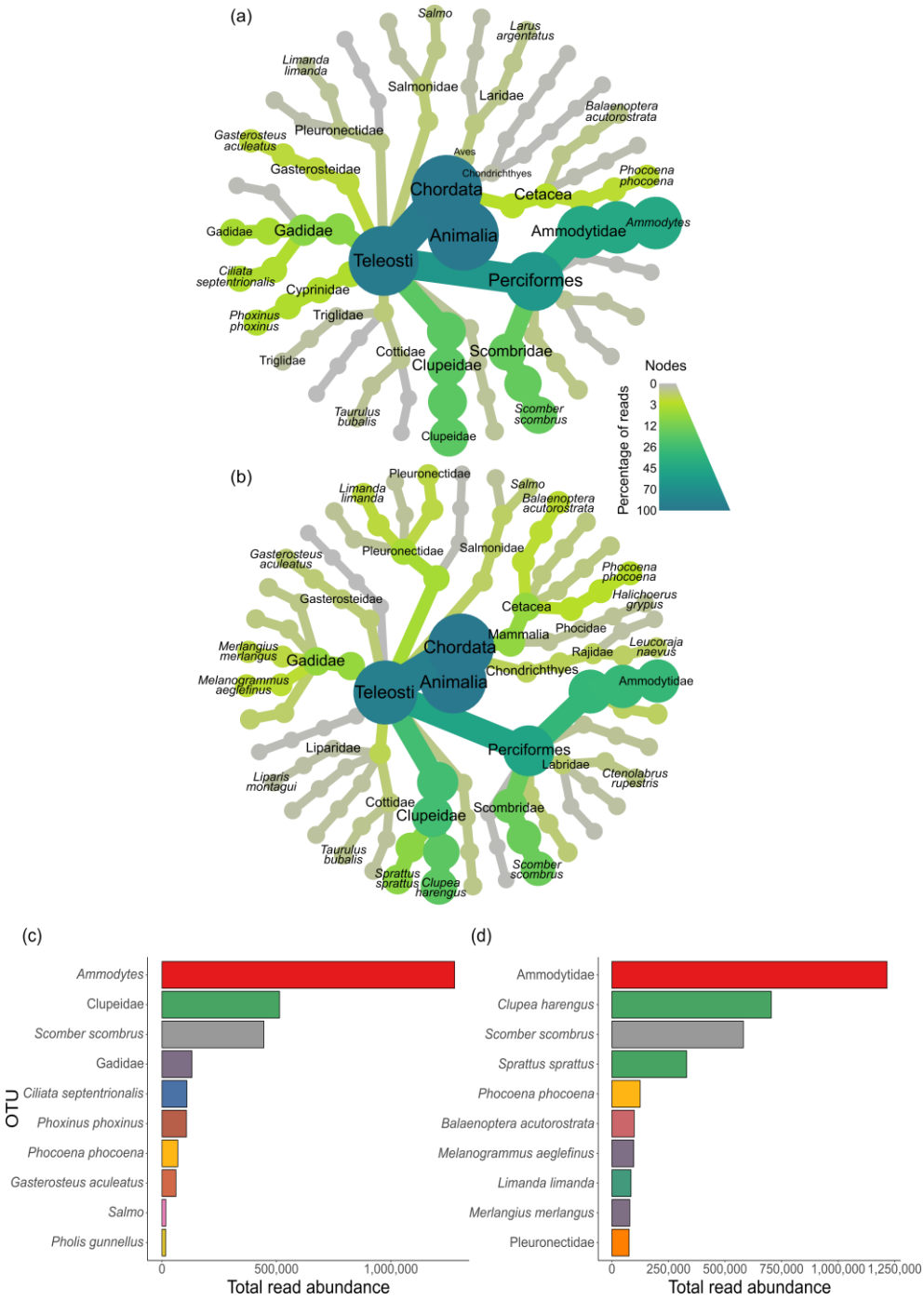


Figure 2.3. Heat trees showing taxa detected with more than 2000 reads for the (a) MarVer1 primer set, and taxa with more than 3000 reads for the (b) MarVer3 primer set. The size and colour of the nodes represents the proportion of reads that a taxa contributes too. Bar charts displaying the ten OTUs with the most abundant read counts for (c) MarVer1, and (d) MarVer3. The colour highlights OTUs belonging to the same family.

The top three most abundant OTUs across both markers were forage fish species, with sandeels having the highest read counts, accounting for 30% and 44% of the reads for MarVer1 and MarVer3 respectively (Figure 2.3). This

was followed by the Clupeidae family, which could only be distinguished at species level to herring and sprat with MarVer3, and then mackerel (*Scomber scombrus*). The Gadidae family had the fourth highest reads for MarVer1, whilst species from the Gadidae family, haddock (*Melanogrammus aeglefinus*) and whiting (*Merlangius merlangus*) had seventh and ninth highest read counts respectively for MarVer3. Two cetacean species, minke whales and harbour porpoises appeared in the ten most abundant reads for MarVer3, but only the harbour porpoise was present among the ten most abundant taxa for MarVer1, although minke whales had the eleventh most abundant reads. The three-spined stickleback (*Gasterosteus aculeatus*) and rock gunnel (*Pholis gunnellus*) have the eighth and tenth most abundant reads compared to fourteenth and fifteenth with MarVer3, likely resulting from more OTUs being grouped at family level with the MarVer1 primer set.

2.3.2. Composition of broader eukaryotic taxa detected

We retrieved 1,769,650 reads in total with our 18S primer set, belonging to 122 different eukaryotic families (Table 2.2). These sequences largely comprised of families belonging to either the Animalia (48% total reads) or Chromista (41% total reads) kingdoms (Figure 2.4a). Reads in the Animalia kingdom were dominated by Calanidae and Acartiidae copepod families from the Maxillopoda class, which together accounted for 42% of the total reads (Figure 2.4). The most abundant families, Leptocylindraceae and Calciodinellaceae, from the diatom (Bacillariophyceae) and dinoflagellate (Dinophyceae) classes respectively accounted for 18% of total reads. The next most abundant classes represent haptophytes (Prymnesiophyceae), also from the Chromista kingdom, and fungi (Phycomycota). Didiniidae and Strombidiidae belong to the ciliates class, but Ciliata was removed as potential contamination when analysed at class level due to other ciliate families being abundant in field blanks.

Table 2.2. Composition of sequencing reads within different taxonomic kingdoms retrieved from the V9 region of 18S rRNA, and the number of OTUs detected at family level.

Kingdom	18S - Total Reads (% of reads)	18S – OTUs detected
All	1,769,650	122
Animalia	845,096 (48%)	47
Chromista	727,542 (41%)	48
Protozoa	130,151 (7%)	14
Fungi	53,767 (3 %)	5
Plantae	13,092 (<1%)	8

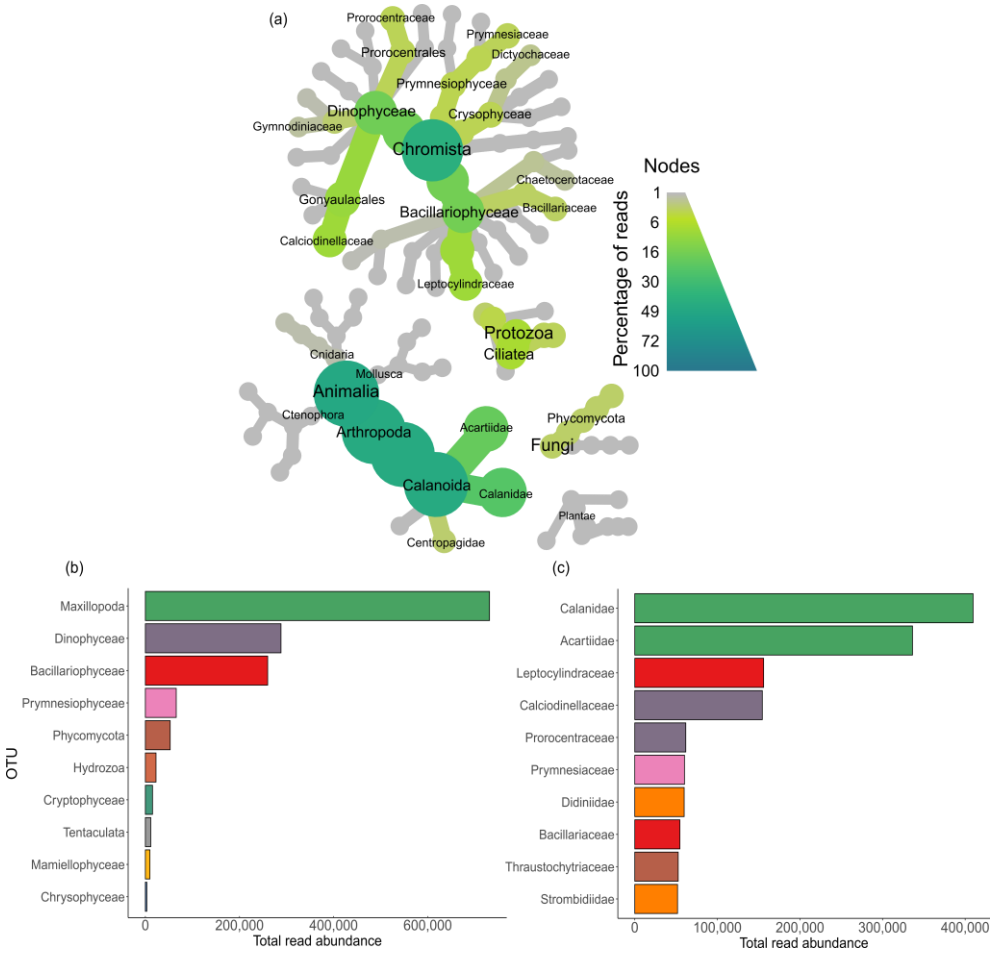


Figure 2.4. (a) Heat tree showing taxa detected with more than 500 reads by the 18S primer set. The size and colour of the node represents the proportion of reads contributed by the taxa. Bar charts with the ten most abundant OTUs for 18S at (b) Class taxonomic groups, and (c) Family taxonomic groups. Colours indicate which class the families belong in. Didiniidae and Strombidiidae families belong to the Ciliatea class which is removed as contamination at class level.

2.3.3. Temporal trends in community composition

Sandeels (Ammodytidae) formed higher proportions of read counts between June and August with MarVer1 and June and July with MarVer3 (Figure 2.5). Using our ensemble OTU index, sandeels were identified as an indicator OTU for June, July and August (specificity = 0.98, sensitivity = 0.82, stat = 0.9, p = 0.005) (Appendix Table A2.4). Minke whales (Balaenopteridae) were most prevalent in June and July, and an indicator for these months (specificity = 0.93, sensitivity = 0.88, stat = 0.91, p = 0.02), whilst harbour porpoises (Phocoenidae) had higher read proportions between June and August, and were an indicator for these months (specificity = 0.9999, sensitivity = 0.87, stat = 0.93, p = 0.005). Mackerel (Scombridae) were detected across our full temporal scale but made up a greater proportion of reads between August and October. Similarly, herring and sprat (Clupeidae) were also detected across all months but had the highest read proportions in June and September/October.

The Maxillopoda class made up a greater proportion of reads in June and September/October, with Acartiidae being the most prominent copepod family in early sampling months and Calanidae in the latter months (Figure 2.5). Calanidae is an indicator for August and September/October (specificity=0.93, sensitivity=0.78, stat=0.85, p=0.001) (Appendix Table A2.5). Diatoms (Bacillariophyceae) were most abundant in June and July, predominantly composed of Leptocylindraceae and Bacillariaceae families, whilst diatoms in the latter sampling months composed of the Chaetocerotaceae family. Dinoflagellates were most prevalent in July and August, with the Calciodinellaceae family contributing the greatest proportion, followed by the Prorocentraceae and Gymnodiniaceae families. Calciodinellaceae was an indicator species for July and August as a result of this family being present at all sites from these months (specificity=0.87, sensitivity=1, stat=0.93, p=0.001). Ciliate families, Didiniidae and Strombidiidae, were more abundant in June and July, and Didiniidae (specificity=0.98, sensitivity=1, stat=0.989, p=0.001) was an indicator for these months, while Strombidiidae was an indicator for June, July and September/October (specificity=0.996, sensitivity=0.93, stat=0.964, p=0.001). The haptophyte family Prymnesiaceae was also most abundant in

June and July, although an indicator species for June, July and August (specificity=0.998, sensitivity=1, stat=0.999, p=0.001). Hydrozoa were most frequent in August, represented by the Bougainvilliidae family, which was also an indicator OTU for this month (specificity=0.98, sensitivity=0.58, stat=0.76, p=0.001).

Alpha diversity, with the Shannon Index, did not significantly differ between sampling months for our ensemble vertebrate OTU index or our eukaryote OTU index, although all primer sets had lowest alpha diversity in September/October (Figure 2.6a). PERMANOVA shows that community composition differs between sampling months for both our vertebrate OTU index (*adonis*; $df = 3$, $F = 2.35$, $R^2 = 0.12$, $p = 0.001$) and our eukaryote OTU index (*adonis*; $df = 3$, $F = 9.6$, $R^2 = 0.35$, $p = 0.001$). For eukaryotes, community composition was significantly different between all months (pairwise adonis, $P < 0.01$), whereas for vertebrate communities, the communities in June and July differed from the communities in August and September/October (pairwise adonis, $P < 0.05$). NMDS analysis also revealed distinct communities per month for both vertebrates and eukaryotes (Figure 2.6).

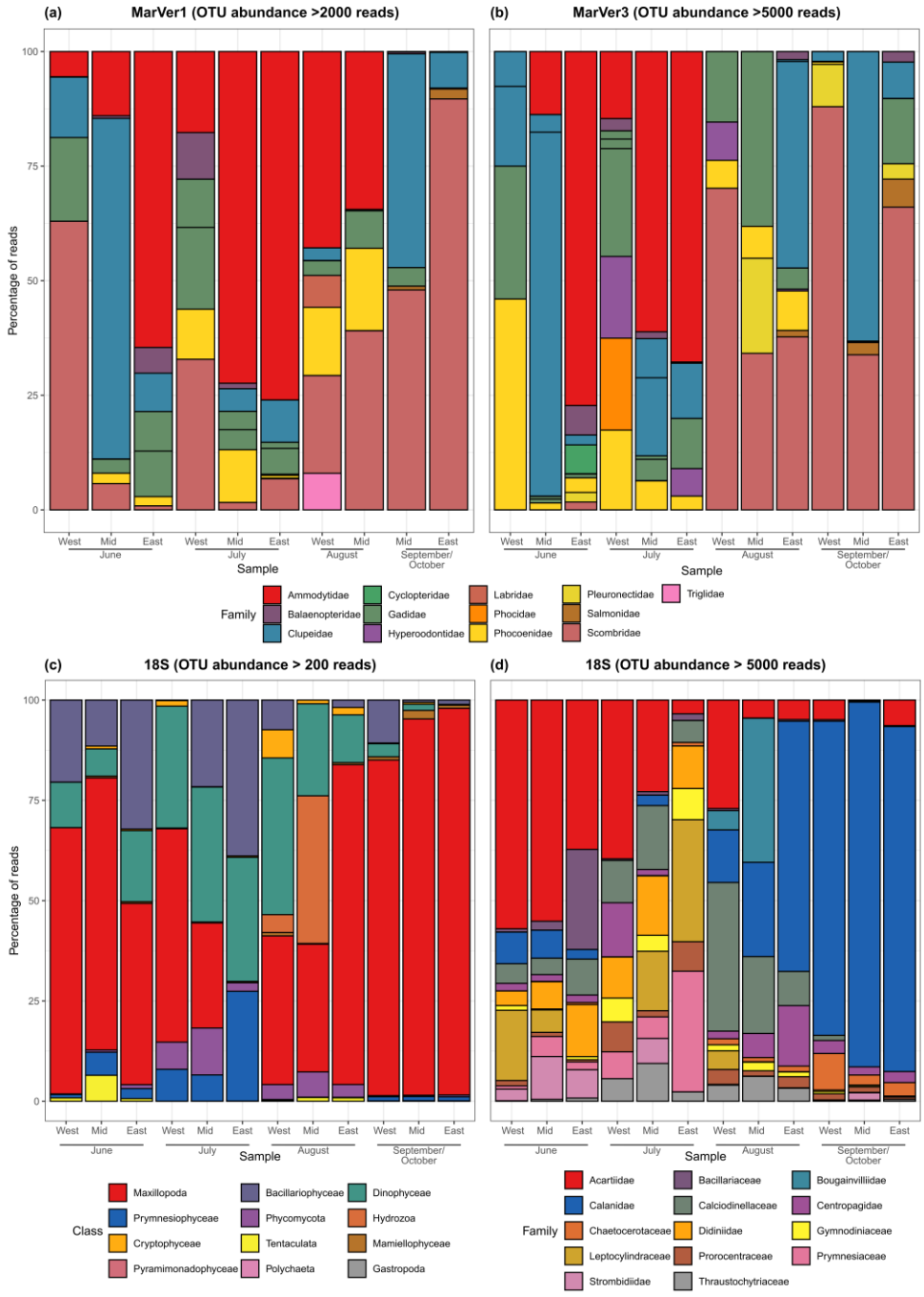


Figure 2.5. Bar charts showing the most abundant vertebrate families for the primer sets (a) MarVer1 and (b) MarVer3, the (c) most abundant 18S classes and the (d) most abundant 18S families from twelve fixed sampling points. Two fixed samples are missing for MarVer1 due to failed amplification.

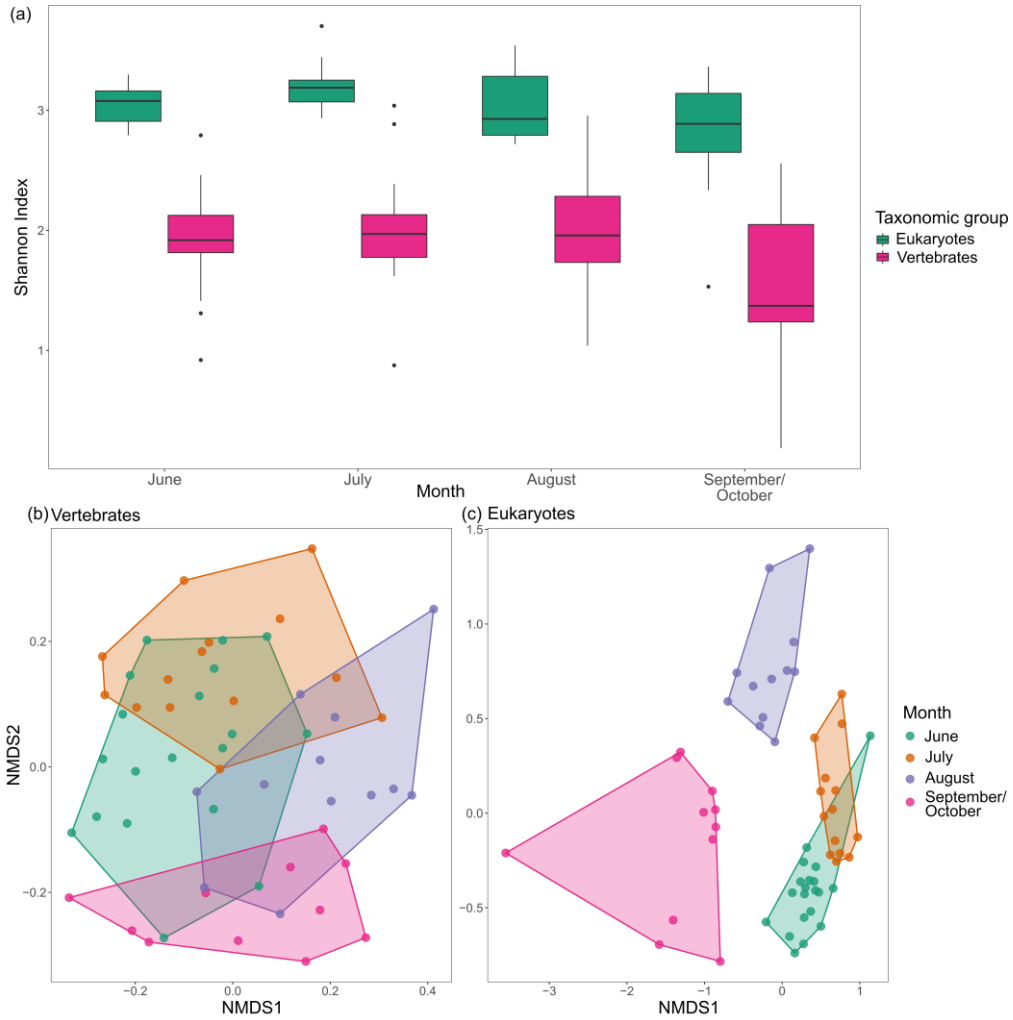


Figure 2.6. (a) Box plots showing alpha diversity for eukaryotes and vertebrates across different sampling months, and NMDS plots for (b) our ensemble vertebrate OTU index (k = 2, stress = 0.262) and (c) eukaryotic OTU index (k = 2, stress = 0.204) partitioned by sampling month.

2.3.4. Spatial trends in community composition

We detected spatial partitioning between cetaceans commonly found in the Moray Firth, with bottlenose dolphins occurring in greatest abundance closest to shore, whilst minke whales and harbour porpoises were more abundant in samples greater than 2.5 km from shore (Figure 2.7). Bottlenose dolphins were an indicator for the near and middle near categories (specificity = 0.996, sensitivity = 0.82, stat = 0.91, p = 0.005) whilst the harbour porpoise was an indicator for the middle near, middle far and far categories (specificity = 0.98, sensitivity = 0.84, stat = 0.9, p = 0.005). Known dominant prey species from stomach analyses for these species shared similar distributions with their predators across distance from shore. Minke whales and harbour porpoises

both prey on sandeels, which were found across all depths, whilst minke whales' also prey on clupeids, which were most prominent between 7 and 10 km from the coast. The Gadidae family, including whiting, targeted by harbour porpoises, and Atlantic cod (*Gadus morhua*) targeted by bottlenose dolphins, were detected across all distances from shore. Conversely, salmonids (*Salmo* spp.), the other main prey of bottlenose dolphins, were most abundant in the nearshore environment.

Alpha diversity, described with the Shannon Index, significantly differed with distance from shore for both vertebrates (Kruskal–Wallis $\chi^2 = 15.68$, df = 3, p <0.001) and eukaryotes (Kruskal–Wallis $\chi^2 = 18.12$, df = 3, p <0.001). In both cases, the nearshore community had significantly higher alpha diversity compared to all groups further offshore (Pairwise Wilcoxon Rank Sum test <0.05) (Figure 2.8a). PERMANOVA analyses found beta diversity to significantly differ with distance from shore for both vertebrate (*adonis*: *df* = 3, $F = 2.03$, $R^2 = 0.10484$, $p = 0.001$) and eukaryote communities (*adonis*: *df* = 3, $F = 1.8274$, $R^2 = 0.09$, $p = 0.002$), although betadisper indicates that within-community variance is not homogenous ($F = 2.9$, p<0.05) for eukaryotes. The vertebrate nearshore community differed significantly from all offshore categories, whilst the eukaryotic nearshore community differed from the middle-far and far categories (pairwise adonis, $p < 0.05$). NMDS showed that the vertebrate nearshore community is distinct from communities further offshore but displayed great overlap between distance from shore categories for eukaryotes (Figure 2.8). The nearshore community had the highest number of OTU indicators for both eukaryotic and vertebrate communities (Figure 2.9). 20 eukaryotic OTUs were identified as indicators for the nearshore community, including 6 families that were only found in nearshore samples; red (Rhodomelaceae) and brown (Dictyotaceae) algae families, calcareous sponges (Leucosoleniidae), tunicates (Molgulidae), bryozoans (Membraniporidae) and cyclopoid copepods (Archinotodelphyidae) (Appendix Table A2.6). Ten vertebrate OTUs were recognised as indicators, most of which were species that commonly reside in shallow depths less than 50 metres, such as the rock gunnel, ballan wrasse (*Labrus bergylta*), eelpout (*Zoarces*), Montagu's seasnail (*Liparis montagui*),

painted goby (*Pomatoschistus pictus*) and the leopard spotted goby (*Thorogobius ephippiatus*) (Appendix Table A2.7).

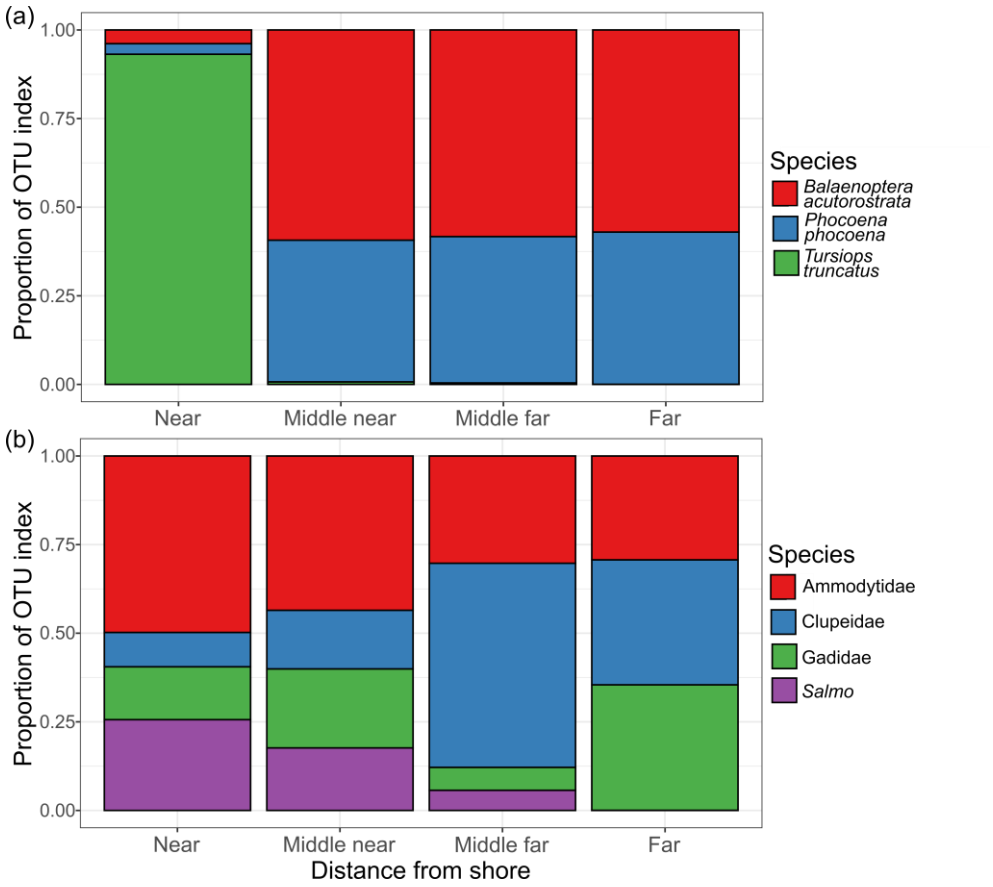


Figure 2.7. Bar charts showing proportion of the OTU index for (a) common cetacean species in the study area, minke whales (*Balaenoptera acutorostrata*), harbour porpoises (*Phocoena phocoena*) and bottlenose dolphins (*Tursiops truncatus*) and (b) their dominant prey species across different distances from shore, near <1.2 km, middle near between 2.5 and 7 km, middle far between 7 and 10 km, and far > 10 km (up to 16.1 km) offshore.

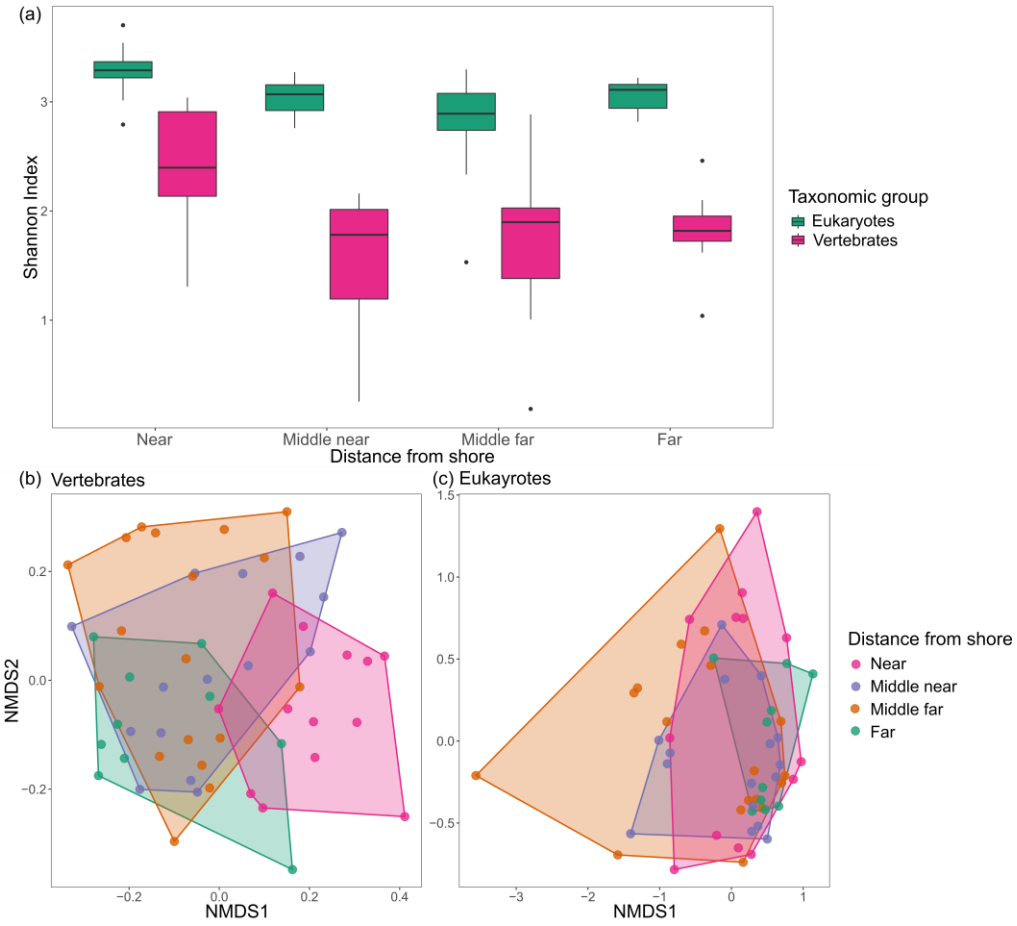


Figure 2.8. (a) Box plots showing alpha diversity for eukaryotes and vertebrates across different distances from shore, and NMDS plots for the (b) ensemble vertebrate OTU index (k = 2, stress = 0.2) and the (c) eukaryotic OTU index (k = 2, stress = 0.18) partitioned by distance from shore.

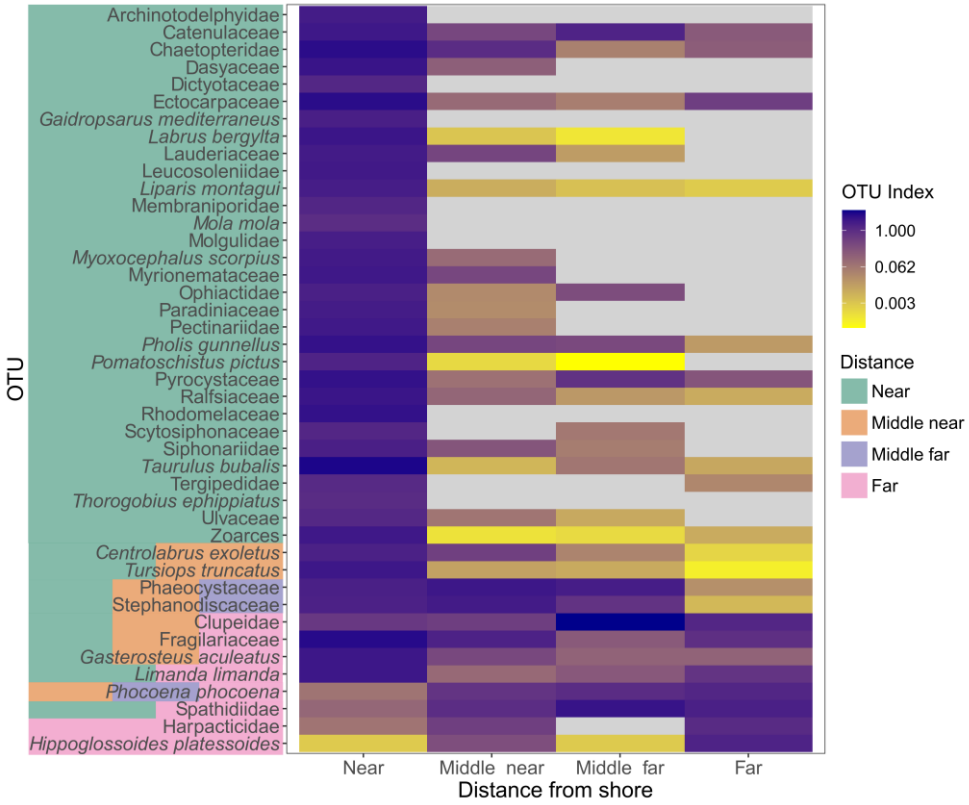


Figure 2.9. Heat map showing indicator OTUs for distance from shore categories across vertebrate and broader eukaryotes. OTUs are coloured to show which distance from shore category or combination of categories they are an indicator species for.

2.3.5. Environmental drivers of community composition

MRM revealed that both temporal, *i.e.,* distance between days, and spatial drivers, *i.e.*, geographical distance, are positively correlated with vertebrate beta diversity (Table 2.3). For eukaryotic beta diversity, difference between days was also positively correlated, along with sea surface temperature. For both vertebrates (coeff = 0.32) and eukaryotes (coeff = 0.8), difference between days had the greatest influence on beta diversity.

Table 2.3. Results of minimal multiple regression distance models for eukaryote and vertebrate OTU indexes.

Eukaryote minimal MRM model			Vertebrate minimal MRM model		
Variable	Coefficient	P value	Variable	Coefficient	P value
Intercept	68.16	1.00	Intercept	443.15	1.00
Sea Surface Temperature	0.08	0.02	Geographic distance	0.1	0.02
Difference in calendar days	0.8	0.0001	Difference in calendar days	0.32	0.0001
Full model statistics					
$R^2 = 0.66$			$R^2 = 0.12$		
P value = 0.0001			P value = 0.0001		

2.4. Discussion

Limited knowledge of prey availability and distributions frequently prevents us from fully understanding heterogeneity and seasonality in marine mammal distributions (Pendleton et al., 2020; Szesciorka et al., 2023). Given that MPAs for marine mammals commonly focus on foraging grounds, it is important to understand their seasonal habitat use within these areas and predict the long-term stability or potential shifts in their location (Notarbartolo di Sciara et al., 2016). Here, we discovered that the availability and abundance of key forage fish prey species varied across the season and with distance from shore, providing important insights into the habitat use of marine mammal species present in a newly implemented MPA in the southern Moray Firth, Northeast Scotland. Further, we recovered spatiotemporal trends in overall community composition which were mirrored in both vertebrate and broader eukaryotic communities, potentially indicating a highly connected ecosystem. As the North Sea is a global warming hotspot (Holt et al., 2012), and forage fish present in the area are particularly vulnerable (Frederiksen et al., 2011; Petitgas et al., 2013), continued monitoring will be essential to ensure continued prey availability for marine mammals, and to detect asynchronous timings in predator-prey presence

which could have negative cascading effects throughout the ecosystem (Silber et al., 2017).

2.4.1. Seasonality in community composition

Temporal drivers had the strongest effects on both eukaryotic and vertebrate communities (Table 2.3), despite sampling being constrained to a relatively small temporal scale. Communities in June and July were more similar compared to communities in latter sampling months, which was reflected in shifts in the abundances of our most abundantly detected OTUs (Figures 2.5). For example, Acartiidae was the most prevalent copepod family in June and July but switched to Calanidae from August onwards. Similarly, sandeels were the most abundant vertebrate in June and July, but mackerel from August onwards. The similarity in temporal patterns between eukaryotic and vertebrate communities could be indicative of strong connectivity between the taxonomic groups, and tight coupling of interactions. Previous declines in the North Sea zooplankton biomass due to warming temperatures have been linked to declines in forage fish biomass, and prevented forage fish populations from recovering following the enforcement of stricter fishing regulations (Clausen et al., 2018; Lindegren et al., 2018). Conversely, forage fish can also exert top-down control on zooplankton, although this is not the dominate mechanism (Fauchald et al., 2011; Lynam et al., 2017). Given that SST is a driver of our eukaryotic community (Table 2.3), and the North Sea is warming three times quicker than the global average, we can expect future changes in the composition of zooplankton (Belkin, 2009; Emeis et al., 2015). Whole ecosystem-based monitoring will therefore be essential to detect early changes and track the cascading effects throughout the ecosystem.

2.4.2. Community composition changes with distance from shore

The nearshore community, within 1.2 km of the shore and less than 25 metres depth, significantly differed from samples collected further offshore. The nearshore community had higher species richness which could partially be a methodological artefact resulting from all samples being collected at four metres depth, potentially increasing the likelihood of detecting benthic species at shallower depths (Figure 2.8). However, OTUs detected in higher abundance or only in nearshore samples represent species and families that

are constrained to shallower depths. For example, two algae families, Dictyotaceae and Rhodomelaceae, were only found in nearshore samples which corresponds to the depth limits of algae growth in the North Sea (Pehlke & Bartsch, 2008; van der Stap et al., 2016). Similarly, American plaice (*Hippoglossoides platessoides*), a demersal species, was most prevalent in samples collected furthest from shore, between 80 and 100 metres depth, suggesting strong mixing of eDNA within the water column, or detections of their pelagic larval phase (Walsh, 1994). Clear spatial partitioning between common cetacean species, which corroborates with distributions based on visual surveys (Robinson et al., 2007), was also detected with eDNA. Bottlenose dolphins were far more abundant in the nearshore community, and previous visual surveys have found this population to occupy depths less than 25 metres (Culloch & Robinson, 2008). Both minke whales and harbour porpoises were detected across all distances from shore, but in greater abundance offshore (>2.5 km), corroborating visual survey efforts (Robinson et al., 2007). Further, minke whale sighting samples had significantly higher abundance compared to control samples (Appendix Figure A2.1). Transport of eDNA from its source by currents and tides is an ongoing concern for the incorporation of eDNA into monitoring in the marine environment (Andruszkiewicz et al., 2019; Hansen et al., 2018). Our results contribute to a growing body of work demonstrating that eDNA provides a relatively local signal from species in the marine water column, suggesting that eDNA degrades rapidly or becomes diluted beyond detectable limits quickly as it is transported away from its source (Hansen et al., 2018).

2.4.3. Minke whale habitat use in relation to prey species

Minke whales are the only species included as a biodiversity feature in the Southern Trench MPA, as a result of the area being an important foraging ground attracting above average abundances of minke whales (Robinson et al., 2009). Current knowledge on spatiotemporal availability of different prey species is hampering identification of important focal areas within the MPA for minke whales. Here, we discover that dominant minke whale prey species, e.g., sandeels and clupeids, are the most abundant species detected with our primer sets (Figure 2.3), although abundances vary throughout the foraging season which would account for the dietary plasticity apparently shown by

these whales (Robinson et al., 2023). Sandeels were most abundant during June and July when they are foraging in the water column, and far less abundant from August onwards when they remain buried in the sediment (Henriksen et al., 2021) (Figure 2.5). Meanwhile, clupeids were most abundant in June and September. The early peak in abundance coincides with the peak spawning period of sprat, whilst the latter peak likely represents the arrival of juvenile sprat and herring to overwinter in the firth (Thompson et al., 1991). Juvenile minke whales target yearling sandeels throughout the foraging season whilst adults target larger sandeels before switching to sprat and juvenile herring as they become more abundant (Robinson et al., 2023). Juvenile minkes are also found at shallower depths while adults prefer deeper, offshore waters, corresponding with the distance from shore that their targeted prey species are found (Robinson et al., 2023). Accordingly, while sandeels were detected across all depths, clupeids were detected in greatest abundance between 7-10 km offshore, at depths between 50 and 120 metres (Figure 2.7). Both sandeels and clupeids are reliant on specific substrates, making them vulnerable to climate-induced depletions as they are restricted in their ability to find new habitat patches facilitating northward migration (Frederiksen et al., 2011; Petitgas et al., 2013). Marine mammals generally respond to prey depletions by switching to alternative prey species or moving to new foraging grounds (Agardy et al., 2019). Elsewhere, mackerel and gadoids such as cod, haddock and whiting have become more important components of minke whale diets as their preferred prey such as krill species and capelin (*Mallotus villosus*) have reduced in abundance (Víkingsson et al., 2014; Windsland et al., 2007). We detected high abundances of mackerel and gadoids, suggesting that alternative prey sources are available in the Moray Firth in case of similar climate-induced declines in the abundance of their preferred prey species. However, alternative species have lower energy values compared with sandeels (Ransijn et al., 2019; Van Pelt et al., 1997), which can result in reduced body conditions and population declines (Österblom et al., 2008; Spitz et al., 2012).

2.4.4. Species of conservation interest

eDNA also detected other species of conservation interest within the Southern Trench MPA. For example, such high detections for Sowerby's beaked

whales were unexpected as limited strandings and no definitive sightings have been recorded in the North Sea despite high survey efforts (MacLeod et al., 2004). In view of the long periods spent beneath the surface, beaked whales are notoriously difficult to detect visually, so eDNA could be an important tool to improve our understanding of this species distribution and conservation status, given that they are listed as data deficient in the IUCN 'Red List' (Hooker et al., 2019). We also cannot rule out sample contamination as the first author EB, who processed the samples, attended a Sowerby's beaked whale necropsy which coincided with highest read proportions in our samples. However, precautions to reduce sample contamination were taken and Sowerby's were also detected in several samples collected and filtered prior to the necropsy, suggesting a true presence signal. Likewise, the North Sea Atlantic bluefin tuna fishery collapsed in the 1960s and detection records have been sparse ever since, although in recent years, evidence suggests that they are making a return to UK waters (Horton et al., 2021). This detection in the Moray Firth therefore provides an important record of where bluefin tuna are potentially expanding their migration routes in between foraging and overwintering/spawning grounds (Horton et al., 2021). We also detected the critically endangered European eel, for which the timing of migration and movement patterns are currently poorly understood around Scotland (Malcolm et al., 2010). Detections of European eel peaked in August which could be indicative of adult eels leaving rivers to return to their breeding grounds in the Sargasso sea, or the arrival of juvenile glass eels (Malcolm et al., 2010). eDNA also detected invasive pink salmon which have been found at low abundance in Scottish rivers for over 50 years, particularly in the River Spey which flows into the Moray Firth at the western boundary of the study area (Armstrong et al., 2018). It is speculated that pink salmon fry enter the sea at the onset of winter leading to low survival rates, but there is currently no evidence to support this (Armstrong et al., 2018). eDNA could provide an additional tool to monitor temporal dynamics of pink salmon in Scottish rivers in relation to the development of management strategies.

2.4.5. Limitations and future work

One of the biggest limitations of using eDNA metabarcoding to explore predator-prey dynamics is the inability to distinguish between different age classes (Hansen et al., 2018), as predators often preferentially select prey of certain sizes (Robinson et al., 2023; Visser et al., 2021). In particular, spawning events have been observed to increase the abundance of read counts retrieved in eDNA studies, but larval life forms are unlikely to represent suitable prey for piscivorous marine mammals (Di Muri et al., 2022; Ratcliffe et al., 2021). However, eDNA can provide information about where and when to carry out more intensive surveys to retrieve parameters such as the age-class structure of fishes present. Similarly, abundance estimates of cetaceans are important to monitor trends in population sizes (Hammond et al., 2021), but we are unable to conclude from eDNA how many individual cetaceans were using the MPA and whether these individuals stayed throughout the season or changed. We only collected samples during one foraging season but the number of minke whales visiting the area is known to vary interannually so it would be interesting to investigate broader interannual community changes and relate these to the number of minkes and environmental drivers (Robinson et al., 2009). Changeable weather conditions limited where samples could be collected, with less samples collected offshore from August onwards, and more westerly samples collected in July compared to easterly samples in September/October (Figure 2.1). These spatial differences could therefore have influenced the temporal signals retrieved, although samples collected <2.5 km offshore all had similar community composition, so we believe this had minimal impact on temporal trends. Future work should extend sampling to areas with similar environmental conditions outside of MPA protection in order to further evaluate the effectiveness of management actions within the MPA (Dunham et al., 2020), and to other MPAs designated for cetaceans around the British Isles to compare community composition.

2.4.6. Conclusion

Our results demonstrate that eDNA metabarcoding can serve as a powerful tool to monitor marine mammals and their prey species simultaneously, improving the understanding of marine mammal habitat use in their foraging grounds. eDNA approaches could support monitoring of MPAs focused on

these foraging areas by informing us about seasonal distribution changes and heterogenous habitat use, as well as contributing to long-term dynamic management of foraging areas as prey species shift distributions in response to global warming (Notarbartolo di Sciara et al., 2016). To this end, we provide an important baseline characterisation of community composition within the Southern Trench MPA in the outer Moray Firth against which future changes in the community composition can be evaluated, for example, due to the implementation of management measures or global warming. Key forage fish species, sandeels and clupeids, account for 86% of the total fish biomass in the Moray Firth and are the main prey species for many piscivorous fishes, seabirds and marine mammals (Greenstreet et al., 1998). These species are especially vulnerable to global warming so monitoring will be essential to inform potential changes in their abundances and to assess how predators respond (Frederiksen et al., 2011; Petitgas et al., 2013).

Chapter 3 - Inferring species interactions from co-occurrence networks with environmental DNA metabarcoding data

Elizabeth Boyse[1], Kevin P. Robinson[2], Ian Carr[3], Morag Raynor[3], Maria Beger[1]*, Simon J. Goodman[1]*

[1]School of Biology, University of Leeds, Leeds, United Kingdom
[2] Cetacean Research & Rescue Unit (CRRU), PO Box 11307, Banff, AB45 3WB, Aberdeenshire, United Kingdom
[3]Leeds Institute of Medical Research at St James's, St James's University Hospital, Leeds, United Kingdom

*M. Beger and S.J. Goodman contributed equally to this work

3.0. Abstract

Improved understanding of biotic interactions is necessary to accurately predict the vulnerability of ecosystems to climate change. Recently, co-occurrence networks built from environmental DNA (eDNA) metabarcoding data have been advocated as a means to explore interspecific interactions in ecological communities exposed to different human and environmental pressures. Co-occurrence networks have been widely used to characterise microbial communities but is unclear if they are effective for characterising eukaryotic ecosystems, or whether biotic interactions drive inferred co-occurrences. Here, we assess spatiotemporal variability in the structuring and complexity of a North Sea coastal ecosystem derived from co-occurrence networks and food webs with 60 eDNA samples covering vertebrates and other eukaryotes. We compare topological characteristics and identify potential keystone species, *i.e.*, highly connected species, across spatial and temporal subsets to evaluate variance in community composition and structure. We find consistent trends in topological characteristics across co-occurrence networks and food webs, despite trophic interactions forming a minority of significant co-occurrences. Significant trophic interactions were more frequently detected when both species were less prevalent. Known keystone species in food webs were not highly connected in co-occurrence networks. The lack of significant trophic interactions detected in co-occurrence networks may result from ecological complexities such as generalist predators having flexible interactions across multiple prey or behavioural partitioning, as well as methodological limitations such as the inability to distinguish age class with eDNA or co-occurrences being driven by other interaction types or shared environmental requirements. We suggest

that inferring biotic interactions with co-occurrence networks constructed from eDNA surveys requires further validation in well-understood ecosystems, and improved reporting of methodological uncertainties, *i.e.*, potential missed species, which could affect ecosystem complexity.

3.1. Introduction

Quantifying biotic interactions is key to predicting how vulnerable species and their ecosystems are to perturbations from climate change (Foden et al., 2019; Paquette & Hargreaves, 2021). Species distributions are typically assumed to be bound by physiological limits to abiotic conditions at broad scales, while biotic interactions play a secondary role at finer spatial scales (Araújo & Rozenfeld, 2014; Cazelles et al., 2016). However, increasing evidence suggests that certain interactions, *i.e.*, mutualistic or commensalism, can impact species distributions at broad scales and are especially important for determining species warm-edge range limits (Araújo & Rozenfeld, 2014; Paquette & Hargreaves, 2021). Climate change is disrupting species interactions due to contrasting rates of individual species range shifts and altered phenologies, affecting the timing of interactions and potentially accelerating the loss of species and linked ecosystem functions (Foden et al., 2019; Valiente-Banuet et al., 2015). Assessing the vulnerability of a species to climate change, or predicting their future distributions without accounting for biotic interactions is therefore likely to yield inaccurate outcomes, particularly for specialist species who strongly depend on a another species (Foden et al., 2019; Tylianakis et al., 2010). Species interactions are numerous making it challenging to gather evidence across whole ecosystems so species co-occurrences have been used instead to elucidate potential interactions (Morales-Castilla et al., 2015).

The construction of co-occurrence networks is data intensive, requiring community composition data across spatiotemporal gradients (Russo et al., 2022). Environmental DNA (eDNA) metabarcoding data, which provides snapshots of whole community composition within a single sample, can therefore enhance the use of co-occurrence networks (Barroso-Bergada et al., 2021). Interspecific co-occurrences may be indicative of a biotic interaction whereby two interacting species will affect the presence and/or abundance of

each other resulting in non-random co-occurrence across space and time (Freilich et al., 2018). Topological characteristics of networks can influence the relative stability or resilience of an ecosystem to perturbations (Barroso-Bergada et al., 2021). For example, high modularity, *i.e.*, separate compartments of non-overlapping, strongly interacting species within the network, may enhance community stability, as the effects of perturbation will be contained within a single module (Delmas et al., 2019). Co-occurrence networks can also identify possible keystone species which have disproportionate negative effects on the whole ecosystem when removed. These species are generally highly connected, with high closeness centrality, *i.e.*, short path between the node and all other nodes in the network, and betweenness centrality, *i.e.*, node forms shortest path to connect other nodes (Zamkovaya et al., 2021). However, the proportion of biotic interactions, and which type of interactions contribute to co-occurrence networks is still uncertain, and the interactions detected can vary greatly among replicate networks even from the same environment (Barroso-Bergada et al., 2021; Russo et al., 2022).

Detecting biotic interactions in co-occurrence networks depends on the type and strength of the interaction, as well as the spatial scale of sampling (Blanchet et al., 2020; Morales-Castilla et al., 2015). Correlation coefficients assume symmetric relationships but lots of biotic interactions are asymmetric, such as predation (positive for the predator, negative for prey), or the association strength for each species involved in a symmetric interaction, such as mutualism (positive for both species), differs, potentially obscuring any co-occurrence signal (Blanchet et al., 2020; Goberna & Verdú, 2022). The more interactions that a species is involved in, the weaker the interaction strength, as the species is less dependent on any one other species (Cazelles et al., 2016). Further, co-occurrences may also stem from shared environmental requirements, dispersal limitations or response to another species, *i.e.*, predator avoidance (Freilich et al., 2018). Indirect interactions can also generate interspecies co-occurrence whereby a predator may co-occur indirectly with a primary producer as a result of the predator's prey depending on the primary producer (Blanchet et al., 2020). It is impossible to distinguish between co-occurrences derived from biotic interactions and other

factors without comparing co-occurrences to previously established biotic interactions. However, to date, co-occurrence networks have largely been assembled for microbial communities where limited knowledge of functional roles limits validation of the nature of co-occurrences (Berry & Widder, 2014).

Trophic interactions are generally more well understood than other interaction types so have been used most frequently to identify biotic interactions in co-occurrence networks (Ford & Roberts, 2019). Current estimates suggest between a quarter and half of co-occurrences in networks could be trophic in origin. However, these have been estimated using putative trophic interactions for plankton data, which are not well resolved, or using presence-absence datasets (Freilich et al., 2018; Russo et al., 2023; Russo et al., 2022). Detecting trophic interactions in co-occurrence networks is particularly challenging as they exhibit strong spatial dependency resulting in either positive or negative co-occurrences (Cazelles et al., 2016; Russo et al., 2023; Thurman et al., 2019). Negative co-occurrences are expected at finer scales where prey are successfully avoiding predators whilst positive co-occurrences over greater scales indicate predators tracking prey (Thurman et al., 2019). Consequently, further exploration comparing known trophic interactions with co-occurrences is needed to validate whether trophic interactions are likely to be detected and to determine the spatial influence on this relationship. If co-occurrence networks can successfully detect trophic interactions, this could enhance our knowledge of the spatiotemporal variability of trophic interactions which is often poorly described relative to overall trophic relationships (Young et al., 2015).

Species interactions in the North Sea ecosystem are well characterised as a result of concern over negative impacts from fishing pressure and climate change, and the need to understand the consequences of these top-down and bottom-up pressures. The North Sea represents a “wasp-waist” system, where a few key forage fish species, *i.e.*, sandeels (*Ammodytes* spp.), herring and sprat (clupeids), exert control over the abundance of predators, *i.e.*, marine mammals, predatory fishes, seabirds, through bottom-up interactions, and control the abundance of zooplankton prey through top-down interactions (Fauchald et al., 2011; Lynam et al., 2017). The abundance and quality of

sandeels, the dominant forage bait fish in the North Sea, has been declining in recent decades, contributing to failures in seabird breeding (MacDonald et al., 2019; Wanless et al., 2005). Forage-fish predators will respond differently depending on diet and foraging specialisations, as well spatial restrictions, *i.e.*, central placed foragers are more sensitive to prey depletions (Engelhard et al., 2014). Forage fish species are also vulnerable to climate change due to their reliance on particular substrates limiting their ability to redistribute northwards, with temperature driving key lifecycle stages, *i.e.,* spawning (Frederiksen et al., 2011; Petitgas et al., 2013). These changes could lead to further declines and reduced temporal synchrony between predators and prey in the future, where prey are not available when required by predators, for example, during seasonal foraging periods or when raising young (Edwards & Richardson, 2004). Monitoring temporal changes in forage fish availability is therefore a priority for understanding how this effects interactions in the North Sea ecosystem.

Within the southern Moray Firth, an inlet of the North Sea, the Southern Trench marine protected area (MPA) has recently been designated to protect minke whale foraging grounds (NatureScot, 2020). Here, we investigate spatiotemporal changes in the complexity and structuring of this ecosystem, employing co-occurrence networks and food webs. We explore changes in complexity between inshore and offshore environments, and temporally during the minke whale foraging season. We quantify differences in co-occurrences detected between spatial and temporal networks and assess the proportion of significant co-occurrences indicative of trophic interactions by comparing co-occurrences with known trophic interactions from the literature. We identify potential keystone species, *i.e.*, species with the most interactions, and evaluate whether key components of food webs are detected in co-occurrence networks. We expect interactions in co-occurrence networks to be more volatile as they can stem from different biotic interactions and shared environmental requirements, potentially resulting in key food web components remaining undetected.

3.2. Methods

3.2.1. Sample collection and analysis

We employed an eDNA metabarcoding dataset derived from 60 samples collected on four sampling trips during June to October 2021 to assess temporal community change, from the southern Moray Firth, north-east Scotland (Chapter 2). Marine vertebrate DNA was amplified using two primer sets, MarVer1 and MarVer3, targeting 12S and 16S rRNA respectively (Valsecchi et al., 2021; Valsecchi et al., 2020), as well as eukaryotic DNA with 18S rRNA to capture zooplankton and other invertebrate taxa (Amaral-Zettler et al., 2009; Sawaya et al., 2019). Sequencing libraries were prepared and sequenced separately at the University of Leeds Genomics Facility, St James Hospital, using an Illumina MiSeq Sequencer with a 150 bp paired-end lane for both vertebrate primer sets, and 250 bp paired-end lane for 18S rRNA. The bioinformatics pipeline is described fully in Valsecchi et al. (2020) and at http://www.dna-leeds.co.uk/eDNA/. Following the removal of low-quality sequences, PCR duplicates and chimaeras, we clustered sequences into molecular operational taxonomic units (OTUs) using a 98% threshold of homology to the GenBank sequence at species level for the two vertebrate primers, and a 95% threshold at class level for the 18S primer set. We converted read counts into an OTU-specific index, allowing comparison of within-OTU abundances across all samples. Firstly, we converted read counts into proportions, and divided the maximum proportion for each OTU from the proportion at individual sites resulting in an index between 0 and 1 for each OTU (Kelly et al., 2019). For vertebrate OTUs that were present across both primer sets, we built an ensemble OTU index by taking the average across both indices at each site (Djurhuus et al., 2020).

3.2.2. Co-occurrence network construction

We subset our dataset into groups to account for spatial and temporal trends in community composition for co-occurrence analyses (Chapter 2), as more similar communities produce more specific co-occurrence networks (Berry & Widder, 2014). We will refer to early season, June to July (34 sites), and late season subsets, August to October (23 sites), as ‘temporal subsets’ and will retain spatial signals in the dataset. Nearshore (13 sites) and offshore (47

sites) subsets will be called ‘spatial subsets’ and preserve temporal patterns in the dataset. Small sample sizes (<20) can affect the reliability of co-occurrence networks to accurately predict interactions, so extra caution must be applied to networks with sample sizes below this threshold (DiBattista et al., 2020; Hirano & Takemoto, 2019). For each subset, we only retained OTUs that appeared in at least 25% of the sites to remove rare species and reduce erroneous correlations (Berry & Widder, 2014).

We assembled individual co-occurrence networks by calculating pairwise co-occurrences between species OTU indexes, a measure of relative abundance, with five different metrics (Pearson and Spearman correlations, Bray-Curtis and Kullback-Leibler dissimilarities, and mutual information) in Cytoscape’s Conet plugin (Faust & Raes, 2016). Edges, *i.e.*, significant co-occurrences between two species, were represented in the network, if they were supported by at least two metrics, reducing the likelihood of false positives, and the highest and lowest scoring 500 edges were retained to capture both positive and negative interactions (Faust & Raes, 2016). *P* values were calculated using the ReBoot method which compares the null distribution of correlations, accounting for compositionality using 100 iterations of method and edge specific renormalised permutations, and 100 iterations of bootstrapped confidence intervals of observed correlations (Faust et al., 2012). *P* values across different metrics were then merged using Brown’s method and corrected for multiple testing with the Benjamini-Hochberg method (Benjamini & Hochberg, 1995; Brown, 1975).

3.2.3. Food web construction

We used a meta-web approach to construct food webs for the Moray Firth ecosystem by determining all potential trophic interactions between consumers and resources detected by our eDNA metabarcoding dataset (D’Alessandro & Mariani, 2021). We downloaded diet items for fishes, marine mammals and seabirds from FishBase (https://www.fishbase.se) and SeaLifeBase (https://www.sealifebase.ca), through the ‘rfishbase’ R package version 4.0.0 (Boettiger et al., 2012). We complemented these data with information from the primary literature, including invertebrate species, using the Google Scholar search engine and search terms “Latin species name” or

"common species name" with "feeding", "diet" or "stomach contents". For some well-studied species, such as marine mammals or seabirds, we restricted the search to dietary studies within the North Sea. We constructed an edge list describing all possible consumer-resource interactions, and subset the data as described above for co-occurrence networks, removing rare species present in fewer than 25% of samples.

Topological properties for both co-occurrence networks and food webs were calculated using the Cytoscape NetworkAnalyzer plugin (Assenov et al., 2008). Food webs and networks were visualised with the iGraph R package version 1.2.1 (Csardi & Nepusz, 2006). Trophic levels for vertebrate species were assigned from FishBase and SeaLifeBase records. We assigned primary producers (*i.e.,* algae) and fungi to trophic level 1, and all other invertebrate classes to trophic level 2 for the purpose of this study. Significant co-occurrences that represented trophic interactions were inferred from the literature used to build food webs. We identified potential keystone species as those with the highest degree of co-occurrences or interactions from co-occurrence networks and food webs respectively (Berry & Widder, 2014).

3.3. Results

3.3.1. Moray Firth community composition

We retrieved 6,894,772 sequences assigned to 88 OTUs across both vertebrate primer sets, and 1,469,355 sequences assigned to 36 OTUs for 18S rRNA (Chapter 2). Over 90% of vertebrate reads belonged to teleost fishes, although mammals, chondrichthyes and birds were also detected. OTUs with the most abundant read counts included forage fish such as sandeels, clupeids and mackerel (Scomber scombrus). Classes from Animalia (48% total reads) and Chromista (41% total reads) contributed relatively equally to overall eukaryotic reads. Copepods from the Maxillopoda class comprised most of the Animalia reads, while dinoflagellates (Dinophyceae) and diatoms (Bacillariophyceae) both dominated Chromista read counts. There was a clear temporal signal in community composition, differing significantly between early and late season samples for vertebrates, and across all four sampling months for eukaryotes (Chapter 2). Sandeels were more prevalent in the early season while mackerel were more abundant in the late season. Maxillopoda

accounted for most reads in the first and last sampling months, while Dinophyceae were more prevalent in the middle sampling months (July and August). For both vertebrates and broader eukaryotes, the nearshore community (< 1000 m from shore) had higher alpha diversity and significantly different beta diversity from communities composed of samples collected further offshore (Chapter 2). Numerous fish species and eukaryote classes were found exclusively in the nearshore community that are known to be associated with shallow depths.

3.3.2. Temporal food webs and co-occurrence network subsets

We found six and nine fewer nodes in the early season (June and July) compared to the late season (August to October), and 108 and 172 fewer edges in the early season for co-occurrence networks and food webs respectively (Table 3.1). Edges in the co-occurrence networks were dominated by positive interactions, representing 85.6% of edges in the early season, and 77.9% edges in the late season, resulting from OTU copresences (Appendix Figures A3.1 – A3.4). Only 53 (17%) edges and 104 (25%) edges in the co-occurrence networks represented known trophic interactions from our literature search. Potential keystone OTUs, *i.e.*, highly connected OTUs, did not overlap between temporal subsets for the co-occurrence networks and food webs, apart from for sandeels, which were a keystone OTU for the late season co-occurrence network and both food web subsets (Table 3.2). We found a negative correlation between average OTU abundance and degree for the early season co-occurrence network (Pearson, $r = -0.41$, $p<0.05$), but no pattern for the late season network. For example, sandeels only showed a high degree of edge interactions in the late season when they were less abundant (Figure 3.1). OTUs with the most edge interactions in food webs were dominated by species which occupied mid trophic levels (2-3), as both consumers and prey within the ecosystem (Figure 3.1). This included some of the most abundant OTUs detected, such as copepods, mackerel, sandeels, and clupeids.

Table 3.1. Topological characteristics of temporal co-occurrence networks (undirected) and food webs (directed) for early season (June to July) and late season (August to October) Moray Firth subsets.

	Co-occurrence networks		Food webs	
	Early season	Late season	Early season	Late season
Nodes[i]	54	60	53	62
Edges[ii]	305	413	290	462
Avg. neighbours[iii]	11.3	13.767	10.49	14.52
Diameter[iv]	4	4	5	6
Radius[v]	3	2	1	1
Characteristic path length[vi]	2.03	1.94	1.87	1.88
Clustering coefficient[vii]	0.43	0.44	0.34	0.34
Network density[viii]	0.21	0.23	0.11	0.12

[i]Nodes represent OTUS.
[ii]Edges represent interspecific co-occurrences in co-occurrence networks, or trophic interactions in food webs.
[iii]Average number of neighbours refers to the average number of edges per node.
[iv]The diameter is the maximum length of the shortest path between two nodes.
[v]The radius is the minimum length of the shortest path between two nodes.
[vi]The characteristic path length is the average shortest path length between any two nodes in the network.
[vii]The clustering coefficient represents the average clustering coefficient across all nodes and is a value between 0-1. It represents the proportion of co-occurrences among the neighbours of a node.
[viii]Network density describes the proportion of realised co-occurrences from all potential co-occurrences.

Table 3.2. Potential keystone OTUs, *i.e.*, the ten molecular operational taxonomic units (OTUs) with the highest degree (number of edges), identified in co-occurrence networks and food webs from the early season (June to July) and late season (August to October) Moray Firth subsets. Gradients of blue indicate trophic levels, from light blue (trophic level 1) to dark blue (trophic level 4).

Co-occurrence networks		Food webs	
Early season	Late season	Early season	Late season
Phaeophyceae	*Oncorhynchus*	Polychaeta	Polychaeta
Spinachia spinachia	Dinophyceae	Gadidae	Gadidae
Dicentrarchus	*Taurulus bubalis*	Maxillopoda	Maxillopoda
Symphodus melops	Ammodytidae	Clupeidae	Bivalvia
Centrolabrus exoletus	*Gasterosteus aculeatus*	Gastropoda	*Scomber scombrus*
Gasterosteus aculeatus	*Zoarces*	Ammodytidae	Clupeidae
Taurulus bubalis	*Larus argentatus*	*Scomber scombrus*	Gastropoda
Pholis gunnellus	*Anguilla anguilla*	*Pomatoschistus minutus*	Ammodytidae
Chirolophis ascanii	*Pholis gunnellus*	Pleuronectidae	Pleuronectidae
Salmo	*Phocoena phocoena*	*Trisopterus esmarkii*	*Pomatoschistus minutus*

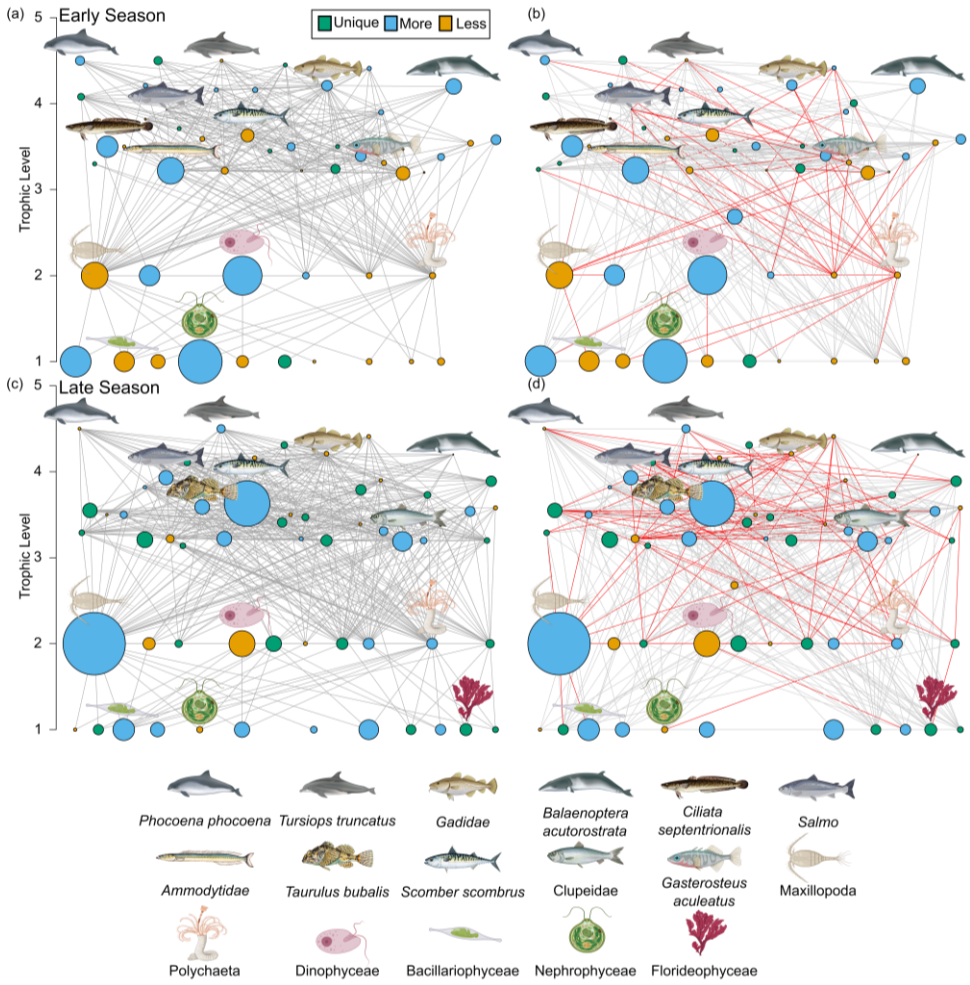


Figure 3.1. Food webs from (a) early season (June to July) and (c) late season (August to October), determined by environmental DNA metabarcoding detections and known trophic interactions, and respective co-occurrence networks for the (b) early season and (d) late season. The size of the node represents the scaled average abundance of the molecular operational taxonomic units (OTU) across samples, and the colour indicates whether the OTU is unique to that time period (green) or more (blue) or less (yellow) abundant. Red edges in co-occurrence networks signify known trophic interactions while grey edges represent all other significant co-occurrences. Individual OTUs are plotted in the same location between graphs.

3.3.3. Spatial food webs and co-occurrence network subsets

We detected 22 more OTUs in the nearshore community compared to offshore, despite the nearshore community including only 13 samples compared to 47 offshore samples (Table 3.3). This included 27 OTUs which were only found in the nearshore community (Figure 3.2). More edges were formed between nodes for the nearshore community, with 139 more edges for the co-occurrence networks, and 206 more edges for the food webs. The

spatial subsets also detected a greater proportion of co-presences (nearshore 73.6%, offshore 77.9%), compared to mutual exclusions, and a small proportion of trophic interactions (nearshore 20.9%, offshore 22.5%). Similar to the temporal subsets above, we discovered little overlap between potential keystone OTUs in the co-occurrence networks and food webs, with Gastropoda being the only keystone OTU in both the nearshore co-occurrence network and food web, and sandeels for offshore (Table 3.4). Two keystone OTUs, the three-spined stickleback (*Gasterosteus aculeatus*), and the long-spined bullhead (*Taurulus bubalis*), were shared across all four co-occurrence networks, whilst five keystone OTUs were shared by one spatial subset and one temporal subset, dinoflagellates (Dinophyceae), brown algae (Phaephyceae), harbour porpoise (*Phocoena phocoena*) and salmon/trout genuses (*Salmo*, *Oncorhynchus*). We found a positive correlation between average OTU abundance and degree for the nearshore network (Pearson, $r = 0.31$, $p<0.05$), but no correlation was detected for the offshore network.

Table 3.3. Topological characteristics of spatial co-occurrence networks (undirected) and food webs (directed) for nearshore (<1000 km) and offshore communities.

	Co-occurrence networks		Food webs	
	Nearshore	Offshore	Nearshore	Offshore
Nodes[i]	68	46	67	46
Edges[ii]	383	244	465	259
Avg. neighbours[iii]	11.27	10.61	13.55	10.826
Diameter[iv]	4	3	5	4
Radius[v]	3	2	1	1
Characteristic path length[vi]	2.13	1.88	1.86	1.8
Clustering coefficient[vii]	0.38	0.4	0.31	0.36
Network density[viii]	0.17	0.24	0.11	0.13

[i]Nodes represent OTUS.
[ii]Edges represent interspecific co-occurrences in co-occurrence networks, or trophic interactions in food webs.
[iii]Average number of neighbours refers to the average number of edges per node.
[iv]The diameter is the maximum length of the shortest path between two nodes.
[v]The radius is the minimum length of the shortest path between two nodes.
[vi]The characteristic path length is the average shortest path length between any two nodes in the network.
[vii]The clustering coefficient represents the average clustering coefficient across all nodes and is a value between 0-1. It represents the proportion of co-occurrences among the neighbours of a node.
[viii]Network density describes the proportion of realised co-occurrences from all potential co-occurrences.

Table 3.4. Potential keystone OTUs, *i.e.*, the ten molecular operational taxonomic units (OTUs) with the highest degree (number of edges), from co-occurrence networks and food webs from the nearshore and offshore communities. Gradients of blue indicate trophic levels, from light blue (trophic level 1) to dark blue (trophic level 4).

Co-occurrence networks		Food webs	
Nearshore	Offshore	Nearshore	Offshore
Zoarces	Dinophyceae	Polychaeta	Polychaeta
Calcarea	*Taurulus bubalis*	Maxillopoda	Gadidae
Ascidiacea	*Limanda limanda*	Gadidae	Maxillopoda
Granuloreticulosea	Ammodytidae	Bivalvia	Clupeidae
Taurulus bubalis	*Salmo*	Gastropoda	Ammodytidae
Phocoena phocoena	*Gasterosteus aculeatus*	*Scomber scombrus*	*Scomber scombrus*
Phaeophyceae	*Oncorhynchus*	Clupeidae	Gastropoda
Gasterosteus aculeatus	*Tursiops truncatus*	Pleuronectidae	*Pomatoschistus minutus*
Gastropoda	Bacillariophyceae	Ammodytidae	Pleuronectidae
Chlorodendrophyceae	*Ctenolabrus rupestris*	Ophiuroidea	*Trisopterus esmarkii*

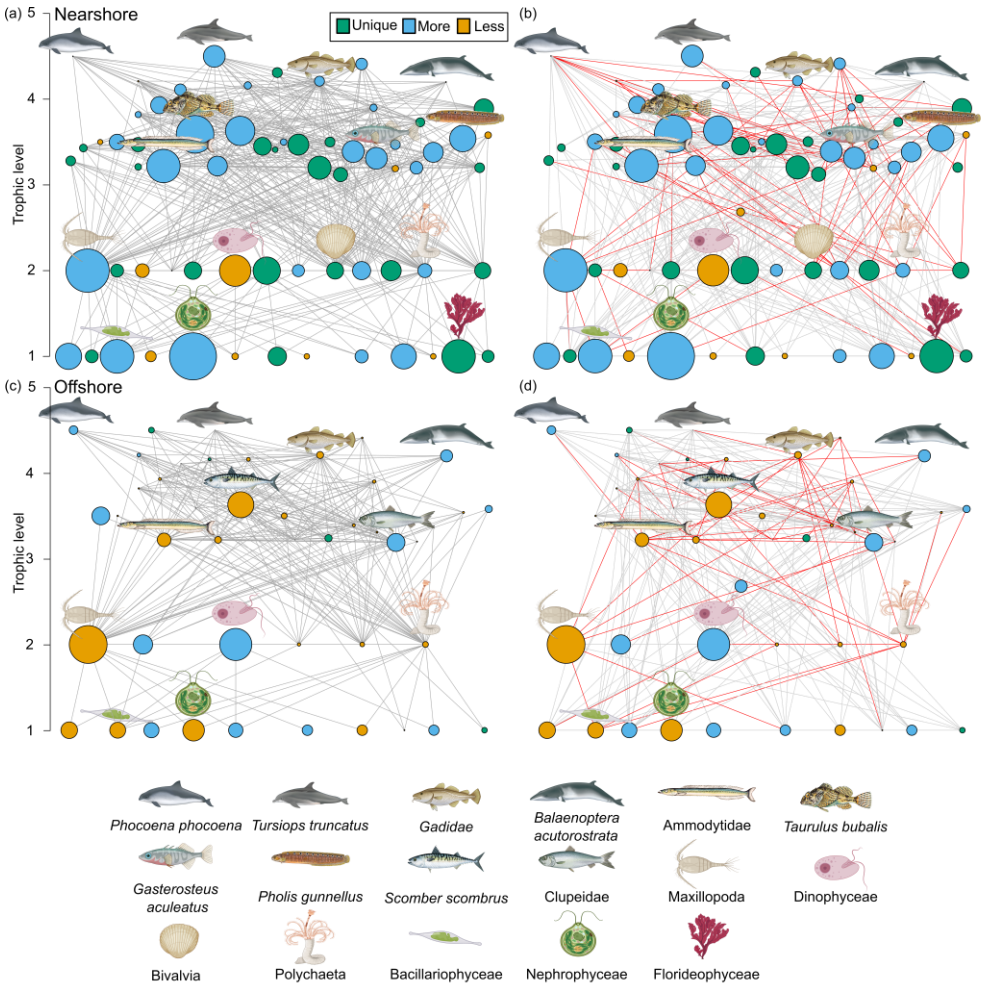


Figure 3.2. Food webs from the (a) nearshore community with samples collected <1000 m from shore, and (c) offshore community determined by environmental DNA metabarcoding detections and known trophic interactions, and respective co-occurrence networks for the (b) nearshore and (d) offshore communities. The size of the node represents the scaled average abundance of the molecular operational taxonomic units (OTU) across samples, and the colour indicates whether the OTU is unique to that community (green) or more (blue) or less (yellow) abundant. Red edges in co-occurrence networks signify known trophic interactions while grey edges represent all other significant co-occurrences. Individual OTUs are plotted in the same location between graphs.

3.3.4. Stability of edges between different co-occurrence network subsets

Within the temporal and spatial co-occurrence network subsets, there were a high number of unique edges, with only 61 (9%) and 37 (6%) shared edges in the temporal and spatial subsets respectively (Figure 3.3). Despite the temporal and spatial subsets being assembled with the same data, there was a high number of edges only detected in either the temporal or spatial networks. Only 268 edges (27%) were found in both subsets, with a similar number of edges unique to the temporal or spatial subsets, 394 and 325 edges respectively. We investigated edge stability further with cetacean trophic interactions and found very few overlapping edges between subsets (Figure 3.3). Only two trophic interactions between bottlenose dolphins (*Tursiops truncatus*) and European seabass (*Dicentrachus*), and between harbour porpoises and sandeels were detected in more than one subset. Only 21% of the detected trophic interactions were with known dominant prey species, and these interactions were all unique to one subset, apart from between harbour porpoises and sandeels which were detected in both nearshore and offshore networks.

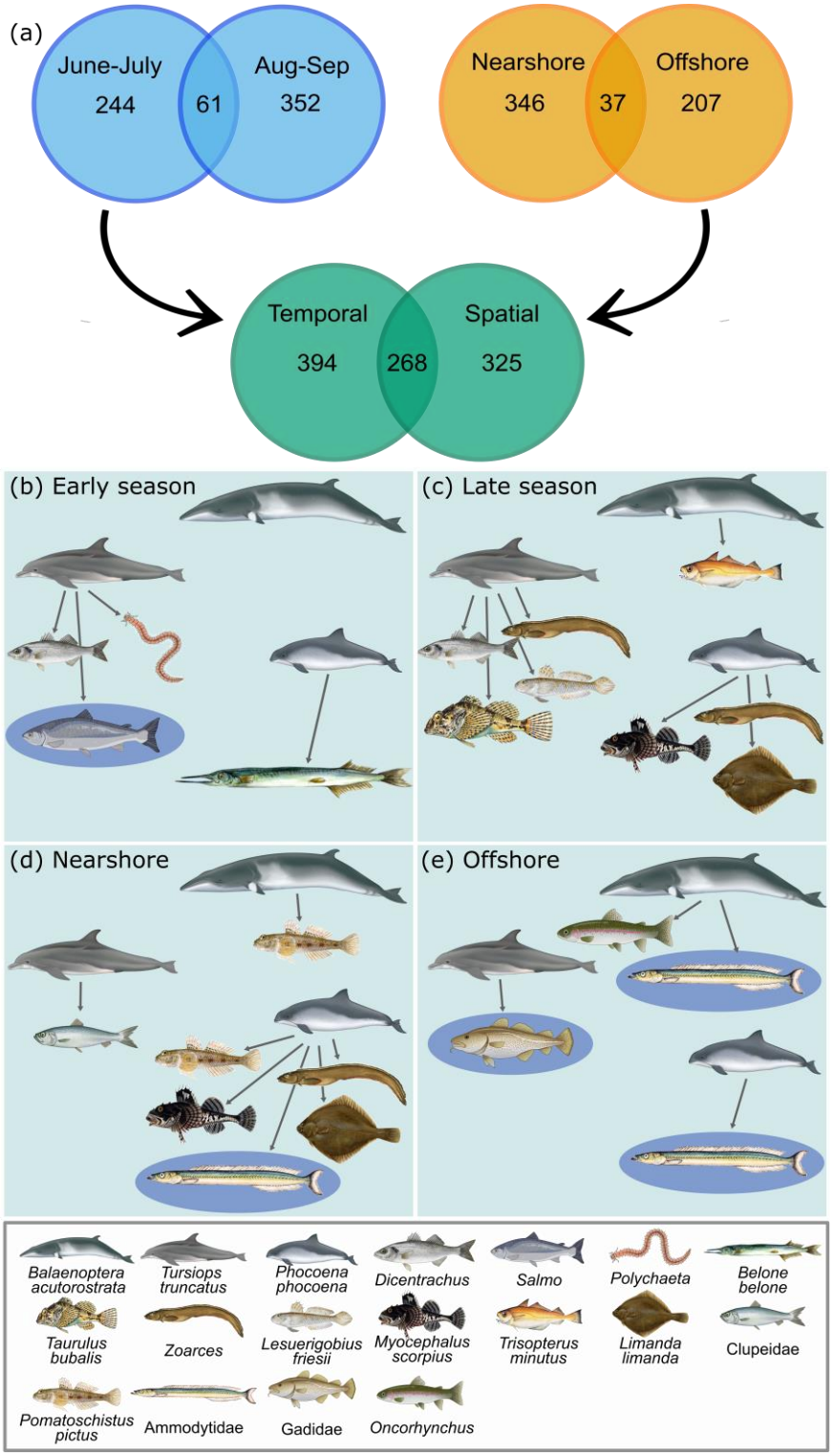


Figure 3.3. (a) Venn diagrams showing the number of overlapping edges detected in the different co-occurrence network subsets. 'Temporal' represents all edges in the monthly samples combined and 'Spatial' represents all edges in the distance from shore communities combined, with duplicates removed. (b) to (d) show known trophic interactions detected between cetaceans and prey in temporal (early season, late season) and spatial (nearshore, offshore) co-occurrence network subsets. Dark blue ellipses indicate known dominate prey species.

3.4. Discussion

In this study, we assessed spatiotemporal variability in the structure and complexity of the Moray Firth ecosystem and identified potential keystone species derived from food webs and co-occurrence networks. We found consistent trends in topological characteristics used to describe overall co-occurrence networks and food webs, with higher interaction diversity found in the nearshore and late season subsets. However, the interactions contributing to different co-occurrence networks were highly changeable, with only 27% of interactions shared between the spatial and temporal subsets. As hypothesised, trophic interactions only formed a small proportion (<25%) of co-occurrences detected across both our spatial and temporal networks, and significant co-occurrences between cetaceans and their known dominant prey species were rare. As a result, we discovered that keystone OTUs, *i.e.*, highly connected nodes, in food webs were not comparable with those detected in co-occurrence networks.

Topological characteristics did not vary greatly between co-occurrence networks and food webs, or between spatial and temporal subsets (Tables 3.1 and 3.3), which is expected given the small spatial scale (tens of kilometres) of sampling. Despite the nearshore network being built from fewer samples (13 samples), it shared similar topological characteristics with other networks suggesting small sample size had limited impact on retrieving significant interactions (Hirano & Takemoto, 2019). The most notable differences were found in the number of edges, with higher interaction diversity detected in the late season and nearshore networks. Networks with more interactions could indicate higher ecosystem functionality and greater redundancy of interactions, thus increasing the community's resilience to disturbance (Tylianakis et al., 2010; Valiente-Banuet et al., 2015). Higher interaction richness in the nearshore subset was likely driven by greater species richness, which has been found to increase closer to shore in previous eDNA metabarcoding studies (Jiménez et al., 2018; O'Donnell et al., 2017; Ríos Castro et al., 2022). Despite the widespread application of eDNA metabarcoding data for building co-occurrence analyses, methodological limitations impacting topological characteristics have not been addressed sufficiently. For example, higher species richness nearshore could be a

methodological artefact, since the water samples were collected from surface water (at 4 metres depth; Chapter 2), such that more benthic species might be detected in the shallower nearshore samples. Previous studies employing eDNA metabarcoding have detected benthic species at deeper depths (>200 m) due the presence of eddies (O'Donnell et al., 2017), although stratification in the southern Moray firth may prevent DNA mixing in the water column (Adams & Martin, 1986). Further, our 18S primer set targeting eukaryotes notably failed to amplify DNA from organisms in the Malacostraca class (Crustacea), such as amphipods, carideans, decapods and isopods. Whilst copepods dominate North Sea plankton assemblages and are the major component of forage fishes diets, these other crustacean groups still form part of the diet of most planktivorous forage fish, and therefore would be reasonably well connected in food webs, potentially becoming more important as copepod abundance declines with warming (Garzke et al., 2015; Mortelmans et al., 2021; Segers et al., 2007). Further, other potentially important OTUs may have been removed due to incomplete reference libraries, affecting the overall ecosystem complexity (Sawaya et al., 2019; Zamkovaya et al., 2021). Extracting abundance data from eDNA metabarcoding is debated, especially methods to tackle bias stemming from differing amplification efficiencies (Hansen et al., 2018; Shelton et al., 2023). Here, we transformed our read counts into an OTU-specific index, which assumes the amplification efficiency is constant across all samples regardless of the community composition (Kelly *et al*., 2019). However, species composition will likely affect the read numbers retrieved, with more accurate abundance estimates recovered for dominant taxa compared with rarer taxa (Skelton et al., 2022). All these effects could impact interaction diversity and network complexity so must be acknowledged as shortfalls when employing eDNA metabarcoding data with co-occurrence network analyses.

Only a small proportion of co-occurrences were attributable to trophic interactions in the present analyses. Our food webs showed that most predators within the study area feed on multiple prey species (Figures 3.1 and 3.2), thus reducing the likelihood of detecting trophic interactions as their associations with prey are likely to be flexible and interchangeable (Thurman et al., 2019). Some trophic interactions may not be described in the literature

and are subsequently categorised as non-trophic interactions in our analyses. For example, dinoflagellates often make up a significant component of plankton biomass and are both important consumers and a food resource but these organisms were not well connected in our food webs, likely owing to difficulties identifying planktonic species in stomach content analyses (Pethybridge et al., 2014; Sherr & Sherr, 2007). In this case, co-occurrence network analyses can potentially contribute to our current understanding, as dinoflagellates are an abundant and well-connected OTU in the offshore network suggesting they play an important role in this community. Increased diet metabarcoding studies will also improve our knowledge of trophic interactions in plankton communities in the future (Zamora-Terol et al., 2020). Detecting trophic relationships in co-occurrence networks relies on the assumption that predators are tracking their prey (co-presence) or prey are avoiding their predators (mutual exclusion) (Thurman et al., 2019). In reality, species will partition their time between different behaviours or, at the extreme end of the scale, show seasonality in their foraging and breeding grounds, which will affect interspecific interactions over different spatiotemporal scales (Risch et al., 2014). An important limitation to eDNA sampling, often overlooked in co-occurrence networks, is that the method cannot distinguish between age classes (Hansen et al., 2018). However, age substantially influences what a species eats and who it is eaten by (Bossier et al., 2020). For example, herring will consume juvenile cod, but adult cod will also consume herring, which may skew correlation trends when age-class cannot be distinguished (Lynam et al., 2017). Spawning events may additionally result in increased peaks in eDNA abundances which could further bias interactions towards those incorporating early life stages, since the presence of DNA derived from gametes may not reflect the occurrence of trophic interactions in the same way as for older age classes (Di Muri et al., 2022; Valsecchi et al., 2021).

We discovered high numbers of edges unique to either the spatial or temporal subsets, despite the same data being used to create these subsets (Figure 3.3). Designating subsets prior to co-occurrence analyses is often carried-out arbitrarily, but samples must come from similar environments to avoid habitat filtering dominating potential interactions (Berry & Widder, 2014). In this

study, we focused on trophic interactions with cetacean species to investigate edge stability, as we know that cetacean diets are typically dominated by a few target prey species, thus increasing the likelihood of detecting the trophic interactions present (Thurman et al., 2019). For example, sandeels and clupeids comprise over 80% by weight of minke whale (*Balaenoptera acutorostrata*) diets, whilst sandeels and whiting (Gadidae) make up 80% of harbour porpoise diets (Pierce et al., 2004; Santos et al., 2004). However, significant co-occurrences between sandeels and minke whales or harbour porpoises were only detected in the spatial co-occurrence subsets (Figure 3.3). This may be due to the spatial subsets retaining a temporal signal, and the abundance of sandeels displays strong temporal variability with higher abundances in June and July when they are actively feeding within the water column (Henriksen et al., 2021). Similarly, we also detected higher abundances of both minke whales and harbour porpoises in June and July. Gadidae species contribute up to 84% of the bottlenose dolphin diet, and salmon are also suspected to be a dominate prey species although their otoliths are completely digested so are difficult to detect in stomach content analyses (Santos et al., 2001). We found co-occurrences between bottlenose dolphins and Gadidae in the offshore network, and with salmon in the early season network, when species occurred at very low abundances. Zurell et al. (2018) previously observed that predator-prey co-occurrences were more likely to be detected when both species were rare, and detectability decreased as one of the species became more abundant. However, rare species in eDNA analyses may represent false positive detections stemming from transport of DNA in tides and currents (Hansen et al., 2018). Finally, the three species further displayed trophic interactions with species not known to form large portions of their diet, which could indicate cetacean's prey switching although seems unlikely as their preferred prey species were available in higher relative abundances. Without baseline information about these species, it is not possible to interpret these interactions.

Potential keystone OTUs in co-occurrence networks and food webs were largely different (Tables 3.2 and 3.4). We identified keystones based on degree centrality, but OTUs with high closeness and between centralities were largely very similar (Appendix Tables A3.1 and A3.2). The three-spined

stickleback and long-spined bullhead were the only two OTUs that had high degree centrality across all four co-occurrence networks (Tables 3.2 & 3.4). The functional role of these species in ecosystems is not well understood, especially non-trophic interactions which make up the majority in our networks. However, both OTUs were most abundant in the nearshore environment and co-occurred within all networks, suggesting similar habitat use could be driving their co-occurrence. The three-spined stickleback spawns in very shallow coastal environments (<3 m depth) during the spring and summer but spends most of its life cycle in open seas, whilst the long-spined bullhead is a permanent resident of the intertidal zone (Bergstrom *et al*., 2015, Barrett *et al*., 2016). They also share lots of edges with other predominantly coastal species, such as the European herring gull (*Larus argentatus*), corkwing wrasse (*Symphodus melops*), rock gunnel (*Pholis gunnellus*), salmon/trout (*Salmo*), and benthic invertebrate species (bivalves, gastropods, polychaetes) (Figure 3.2), enhancing the likelihood that co-occurrence resulted from shared habitat requirements. In this instance, they are therefore unlikely to be keystone species with disproportionate negative effects across the ecosystem if removed (Faust et al., 2012). Other OTUs such as bottlenose dolphins in the offshore network and sandeels in the late season network, only have high degree centrality in the networks where they are least abundant. We sighted no bottlenose dolphins offshore whilst collecting samples, and previous research corroborates that they are regionally coastal, found in depths up to 25 meters (Culloch & Robinson, 2008; Robinson et al., 2007). Therefore, it is likely that the small quantities of eDNA detected for bottlenose dolphins in the offshore environment resulted from movement of DNA particles in the water column as opposed to the bottlenose dolphins being present (Andruszkiewicz et al., 2019). Similarly, sandeels are only detected as a keystone OTU in the late season network when they occur in low abundance as they are buried in the sand and are unavailable to interact with other species in the water column (Henriksen et al., 2021). Conversely, OTUs that exert large influences in North Sea food webs, were not identified as keystone OTUs in co-occurrence networks. For example, copepods contribute up to 90% of the zooplankton biomass in the North Sea and are responsible for transferring energy from primary producers to commercially important fish species, as well as nutrient recycling and carbon export (Kürten

et al., 2013; Mortelmans et al., 2021). We suspect that they fail to have significant correlations with other OTUs as a result of being present in such high abundance across all samples (Zurell et al., 2018). These examples highlight potential flaws associated with both eDNA metabarcoding and correlative relationships, highlighting the need for more robust validation of co-occurrence networks.

In conclusion, ecological interpretations from co-occurrence analyses to explore ecosystem functioning and interspecific interactions can be challenging and potentially misleading. Co-occurrence networks and food webs revealed similar trends in ecosystem complexity, despite interactions forming the networks being largely different. However, it will be important to explore whether this trend is consistent for networks with more varied complexity across greater spatiotemporal scales. Trophic interactions formed a small proportion of co-occurrences leading to key food web components such as forage fish and copepods not being highly connected in co-occurrence networks. Therefore, we strongly recommend that co-occurrence networks should be employed alongside validation methods, such as ground truthing within well-studied ecosystems. In this scenario, food webs and co-occurrence networks may be complementary, and co-occurrence networks could highlight key taxa to focus future diet analyses to overcome current limitations, especially in planktonic communities. Further, limitations of both eDNA metabarcoding and correlation methods need to be explicitly accounted for in these analyses, as both may impact upon which interactions are detected as demonstrated in the present examination. Rapid characterisations of whole ecosystems are necessary to understand their resilience to climate change, but these must be used with caution to avoid drawing incorrect conclusions about the states of ecosystems.

Chapter 4 - Sampling from commercial vessel routes can capture marine biodiversity distributions effectively

Authors: Elizabeth Boyse[1], Maria Beger[1,] Elena Valsecchi[2], Simon J. Goodman[1]

[1]School of Biology, University of Leeds, Leeds, United Kingdom
[2]Department of Environmental and Earth Sciences, University of Milano-Bicocca, Milan, Italy

4.0. Abstract

Collecting fine-scale occurrence data for marine species across large spatial scales is logistically challenging, but is important to determine species distributions and for conservation planning. Inaccurate descriptions of species ranges could result in designating protected areas with inappropriate locations or boundaries. Optimising sampling strategies therefore is a priority for scaling up survey approaches using tools such as environmental DNA (eDNA) to capture species distributions. In a marine context, commercial vessels, such as ferries, could provide sampling platforms allowing access to under-sampled areas and repeatable sampling over time to track community changes. However, sample collection from commercial vessels could be biased and may not represent biological and environmental variability. Here, we evaluate whether sampling along Mediterranean ferry routes can yield unbiased biodiversity survey outcomes, based on perfect knowledge from a stacked species distribution model (SSDM) of marine megafauna derived from online data repositories. Simulations to allocate sampling point locations were carried out representing different sampling strategies (random vs regular), frames (ferry routes vs unconstrained) and number of sampling points. SSDMs were remade from different sampling simulations and compared to the 'perfect knowledge' SSDM to quantify the bias associated with different sampling strategies. Ferry routes detected more species and were able to recover known patterns in species richness at smaller sample sizes better than unconstrained sampling points. However, to minimise potential bias, ferry routes should be chosen to cover the variability in species composition and its environmental predictors in the SSDMs. The workflow presented here can be used to design effective sampling strategies using commercial vessel routes globally for eDNA and other biodiversity survey techniques. This approach has potential to provide a cost-effective method to

access remote oceanic areas on a regular basis, and can recover meaningful data on spatiotemporal biodiversity patterns.

4.1. Introduction

Knowledge of species' ranges is essential for assessments of conservation status, to detect changes in distributions, and to inform spatial planning decisions (Wetzel et al., 2018). Initiatives to aggregate biodiversity data, including the Global Biodiversity Information Facility (GBIF) and the Ocean Biodiversity Information System (OBIS), have increased access to global standardised datasets (Grassle, 2000; Telenius, 2011). However, these datasets are limited by data quality issues, such as positional accuracy or duplicates of records, and spatial, temporal and taxonomic biases (Moudrý & Devillers, 2020). Marine species and habitats are underrepresented due to the monetary and logistical challenges of collecting data, with up to 50% of records for marine taxa being collected from coastal regions, or are classified as Data Deficient in IUCN Red List assessments (Dulvy et al., 2014; Hughes et al., 2021). Data limitations increase uncertainty in marine spatial planning prioritisations and could lead to less efficient marine reserve systems (Bani et al., 2020; Foley et al., 2010). Novel methods that provide high quality biodiversity data are needed for remote areas to improve our knowledge of species distributions, and their conservation. This paper presents a novel framework to design sampling strategies using commercial vessels as data collection platforms that could help to scale up surveys to record species communities more accurately and comprehensively.

In biodiversity surveys, it is usually infeasible to collect samples at very high coverage across large geographical scales, so sampling strategies target the collection of non-biased data at resolutions relevant to study aims. Design-based sampling methods, including random, regular, and stratified random sampling, ensure that every sampling unit has a non-zero probability of being sampled (Wang et al., 2012). Model-based sampling designs aim to avoid bias by considering spatial autocorrelation and heterogeneity in the sampling frame, the area to which sampling is restricted (Zhang et al., 2020). The choice of sampling design is dependent on the study objectives and study area

characteristics as no method consistently outperforms others (Zhang et al., 2020). These sampling designs assume that it is possible to access the entire sampling frame for sample collection. However, in the marine environment this is often impossible to achieve, especially when considering the large spatial scales relevant for marine spatial planning, or the conservation of highly mobile species (Notarbartolo di Sciara et al., 2016).

Commercial vessels, such as ferries, typically follow specific shipping routes covering large spatial scales comprehensively, making them effective platforms for replicable sampling transects. Ferry-based sampling is a similar concept to collecting samples close to road networks, which is commonly employed in terrestrial biodiversity surveys due to greater accessibility (Kadmon et al., 2004). The data collected can be biased because the presence of roads directly affects species distributions or because they do not represent the environmental gradients in the whole sampling frame (Kadmon et al., 2004). We therefore need to explore sampling methods which can best capture variability in species distributions from restricted sampling frames, as these often offer us low-cost sampling and accessibility to hard-to-reach areas. Samples from a restricted area (*i.e.,* road networks or shipping routes) can still produce species distribution model predictions similar to samples collected from an unconstrained area if the environmental gradients are adequately captured (Tessarolo et al., 2014). For commercial vessel surveys to be effective, we need a framework for selecting networks of individual routes to accurately capture species composition, for which there is no precedent despite their frequent implementation in visual cetacean surveys and continuous marine plankton recorder surveys (Arcangeli et al., 2017; Reid, Colebrook, et al., 2003). Furthermore, other survey technologies, such as environmental DNA (eDNA) or trawl deployment for fishery surveys, also require effective methods for allocating sample points along ferry routes (Aubert et al., 2018; Valsecchi et al., 2021). Understanding which sampling strategies will reduce the inherent bias of restricted sampling frames will allow us to best leverage these low-cost sampling opportunities.

Species distribution models can serve as sampling backgrounds for simulating sampling strategies (Tessarolo et al., 2014). Individual species distribution models can be summed using probability or binary predictions to

create a stacked-species distribution model (SSDM) that predicts species richness (Calabrese et al., 2014). Species distribution models only consider environmental constraints on species distributions which can lead to overprediction of species richness when combining multiple models, as biotic mechanisms such as dispersal limitations or resource competition are not accounted for (Gavish et al., 2017). However, using stacking methods based on occurrence probabilities instead of thresholding occurrence probabilities leads to SSDMs which predict species richness similarly to macroecological models, whilst also retaining information on individual species (Calabrese et al., 2014; Distler et al., 2015; Grenié et al., 2020). The use of empirical versus simulated communities allows for complex community "organisation" to be included in sampling simulations and can highlight areas of important conservation interest, i.e. rare species distribution ranges or gradients of diversity (Miller, 2014). We can use the outputs from SSDMs as a benchmark to assess sampling biases associated with different sampling strategies (Braunisch & Suchant, 2010).

This study develops a novel approach for assessing the suitability of different sampling strategies to reduce biases associated with spatially constrained sampling platforms, such as commercial vessel routes. Such a strategy could be used to gain high quality data from pelagic areas that are currently under-sampled due to accessibility and monetary constraints (Hughes et al., 2021). Firstly, we quantify the magnitude of bias of a spatially constrained network of ferry routes, relative to unconstrained sampling across the Mediterranean Sea, employing different sampling strategies to allocate sampling points. Second, we consider how environmental variability or species composition impact the effectiveness of ferry routes as a sampling frame with different subsets of ferry routes. Finally, we evaluate the impact of taxonomic sampling biases on correctly predicting gradients in biodiversity as these biases are pervasive in sampling methods such as eDNA metabarcoding. We use ferry routes in the Mediterranean Sea, but the workflow could be applied to shipping networks anywhere, with any kind of vessel.

4.2. Methods

4.2.1. Building stacked species distribution models

We assembled a SSDM to represent true species distributions based on observational data from online biodiversity repositories and environmental data. An initial literature search identified 171 species of predatory marine megafauna (elasmobranchs, mammals, teleost fishes and turtles) with known occurrences in the Mediterranean. We defined predatory megafauna based on two criteria; maximum length greater than or equal to 100 cm and a trophic level greater than or equal to four as reported in FishBase (https://www.fishbase.se/) or SeaLifeBase (https://www.sealifebase.ca/). Nine species were retained that only satisfied one of the criteria (Appendix Table A4.1). Occurrence records for species were downloaded from GBIF (https://www.gbif.org, June 2020, GBIF Occurrence Download https://doi.org/10.15468/dd.tqx2he), OBIS (https://obis.org/) and EurOBIS (https://www.eurobis.org/) and supplemented by the Mediterranean Large Elasmobranchs Monitoring (Medlem) database and ACCOBAMS dataset for elasmobranchs and cetaceans respectively (ACCOBAMS Survey Initiative, 2020; Mancusi et al., 2020). We subset occurrences to include records from 2000 onwards to correspond with the years that environmental variables were available. We removed occurrence records where GPS coordinates had fewer than three decimal places to improve positional accuracy, and duplicates between the datasets based on species, coordinates, year and month (Moudrý & Devillers, 2020). Records which had the same species, year and month but different coordinates as a result of potentially rounding between databases were also assumed to be duplicates and removed manually. After quality checking, we only retained species with 40 or more occurrence records to improve model accuracy, leading to 43 species in the final presence only dataset, with records for individual species ranging from 41 to 7,822 occurrences (Meynard et al., 2019). The selected species were representative of all marine vertebrates including teleost fishes (n=20), elasmobranchs (n=13), marine mammals (n=9) and one sea turtle species (Appendix Table A4.1) To account for sampling bias in data repositories, occurrences were spatially thinned with a nearest neighbour distance of 10 km using the spThin R package (Aiello-Lammens et al., 2015). This approach prevents clusters of

occurrences although does not account for large scale spatial biases. This method resulted in less than 40 occurrences for ten species, in which case the original data was used instead. We downloaded six environmental predictors, bathymetry, sea surface temperature mean, sea surface temperature range and chlorophyll *a* mean from Bio-ORACLE, and bathymetric slope and distance from shore from Marspec, in WGS84 projection at a resolution of 0.83x0.83 degrees (Assis et al., 2018; Sbrocco & Barber, 2013). These environmental variables are of known importance to marine predatory megafauna, or their prey species (Azzellino et al., 2012; Klippel et al., 2016; Lambert et al., 2017). These variables were normalised to between 0 and 1 to account for units differing by orders of magnitude.

We modelled individual species distributions with three different approaches, maximum entropy (MAXENT), multiple adaptive regression splices (MARS) and random forest (RF). MAXENT was run with 10, 000 random background points using the dismo R package (Hijmans et al., 2017). We selected presence-absence algorithms MARS and RF despite having a presence-only dataset as they perform better than presence-only models, when employed with pseudo-absence data (Barbet-Massin et al., 2012; Zhang et al., 2019). We generated 1000 pseudo absences for MARS and an equal number of pseudo absences as presences for RF, both randomly selected within a restricted sampling frame using the two-degree method as recommended by Barbet-Massin et al. (2012). We allowed first order interactions to be fitted for MARS (Wisz et al., 2008). RF was run with 5000 regression trees and a terminal node of 5 (Zhang et al., 2019). We randomly assigned the data set into training (70%) and testing (30%) sets three times for cross validation (Arenas-Castro et al., 2022; Sundaram & Leslie, 2021). We assembled the model projections across the three modelling methods using weighted AUC scores for each species. Probabilities of occurrence were translated to binary occurrences using the sensitivity (i.e. true positive rate) equals specificity (i.e. true negative rate) threshold (Liu et al., 2005). The individual species binary ensemble models were then summed to show species richness in the final binary SSDM (Figure 4.1). We selected a binary SSDM as binary data was required for sampling simulations. This initial SSDM created with occurrence data from online repositories will be referred to as the ‘perfect knowledge’

SSDM for sampling simulation comparisons. All species distribution modelling was carried out using the SSDM R Package using R version 4.1.0 (Schmitt et al., 2017).

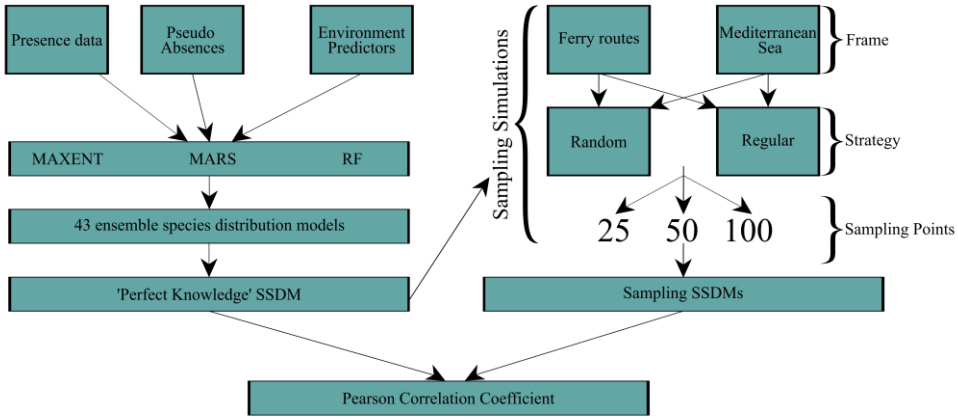


Figure 4.1. Schematic diagram showing the workflow to create the 'perfect knowledge' SSDM using occurrence data from online repositories and extracting occurrence data from the 'perfect knowledge' SSDM to build the sampling SSDMs. Sampling SSDMs were compared to the 'perfect knowledge' SSDM to evaluate their predictive capacity.

4.2.2. Sampling strategy simulations

To enable comparisons of different sampling strategies relative to the 'perfect knowledge' SSDM, we selected fifteen operational ferry routes of varying lengths (both intra/inter country tracks) to represent the distribution of ferry routes in the Mediterranean (Figure 4.2A). We simulated two sampling strategies (random and regular) across different sample sizes (25, 50, 100 sampling points) with either the ferry route network or the Mediterranean as a sampling frame to compare differences between biodiversity detected by a restricted sampling frame versus unconstrained sampling (Figure 4.1). Random sampling allocates sampling points anywhere within the sampling frame, whilst regular sampling places sampling points at uniform intervals but introduces randomness with a varied starting point. We explored different combinations of ferry routes, referred to as 'ferry subnetworks', to consider the importance of environmental and species data coverage by the ferry routes. We simulated each sampling strategy combination 1000 times to calculate the mean number of species sampled, and the mean number of occurrences per species in the simulations. All sampling simulations were

carried out using the spsample() function from the sp R package to allocate sampling point locations according to the defined sampling frame, strategy and sampling size (Bivand et al., 2008).

For each sampling strategy simulation, species occurrences were extracted from the 'perfect knowledge' SSDM to regenerate new SSDMs from the simulated sampling data, referred to as 'sampling SSDMs'. We created these SSDMs as before, except that species were not spatially thinned prior to modelling and species with >20 occurrences were retained. We chose this threshold to evaluate the effect of small sample sizes on model prediction accuracy. To compare species richness across the Mediterranean and ferry route network as sampling frames, 40 replicate SSDMs were built for each combination of sampling frame, size and strategy, using 40 different sampling simulations. We assessed correlation of species richness between the 'perfect knowledge' SSDM and the sampling SSDMs based on Pearson's correlation coefficient to evaluate the effectiveness of different sampling strategies. A three-way Analysis of Variance (ANOVA) was performed to evaluate the effects of sampling strategy, frame and number of sampling points on the correlation coefficient.

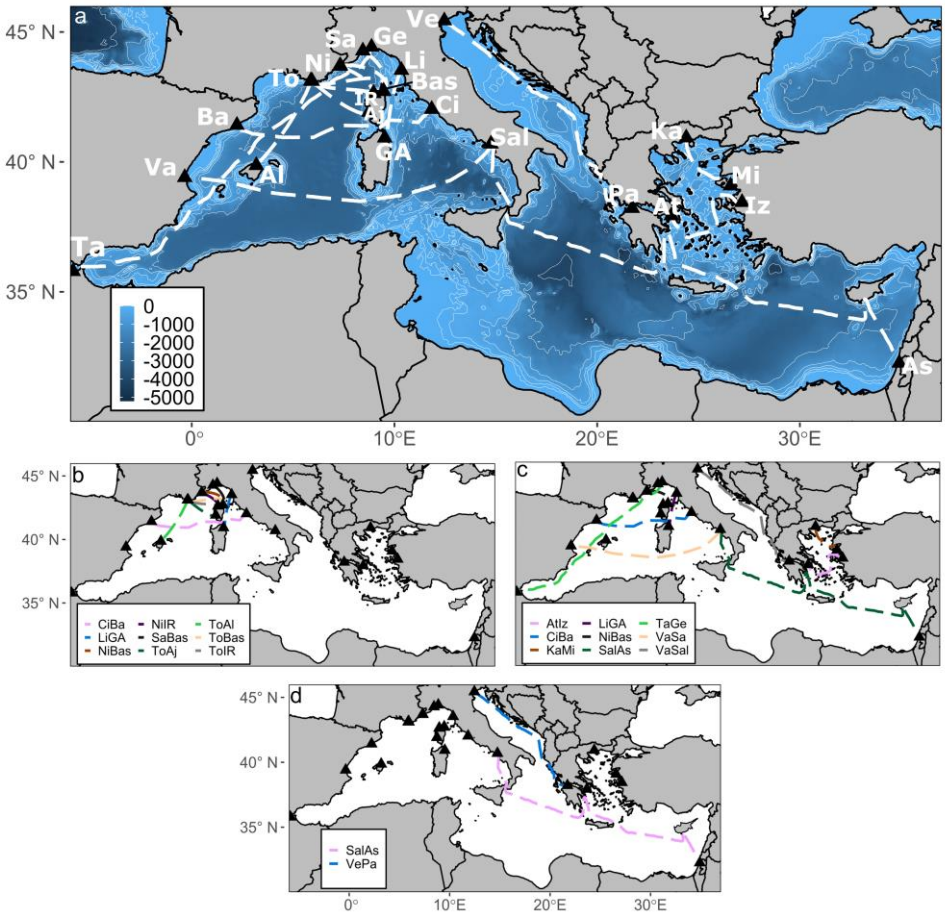


Figure 4.2. Maps showing the layout of the a) whole ferry route network consisting of 15 individual ferry routes. Abbreviations for Ports: Al=Alcudia, Aj=Ajaccio, As=Ashdod, At=Athens, Ba=Barcelona, Bas=Bastia, Ci=Civitavecchia, GA=Golfo Aranci, Ge=Genoa, IR=Ille Rousse, Iz=Izmir, Ka=Kavala, Li=Livorno, Mi=Mitilini, Ni=Nice, Pa=Patras, Sa=Savona, Sal=Salerno, Ta=Tangier, To=Toulon, Va=Valencia, Ve=Venice. b) biased ferry route subnetwork, c) community ferry route subnetwork, d) environment ferry subnetwork.

4.2.3. Ferry route subnetworks

We built different ferry subnetworks to evaluate how different coverage of environmental variability and community composition affected the predictive capacity of ferry routes as a sampling frame. The environmental predictors were collapsed into a single index of environmental variability using principal component analysis to quantify the main gradients of environmental variability in the study area (Appendix Methods A4.1). The first four principal components explained >80 % of the variability in the environmental predictors. Therefore, we collapsed these principal components by summing

the site scores of each principal component weighted according to its contribution (Long & Fisher, 2006; Maina et al., 2008). The resulting environmental variability map was normalised between zero and one, where zero and one represent the most different environments. We quantified climatic bias for different ferry subnetworks by comparing the difference in density functions between environmental variability over the whole study area and those covered by the ferry routes. We split the density functions into 5 equal bins of 0.2 to calculate the climatic bias index. We define our climatic bias index as the sum of the differences of density functions of environmental variability. Salerno-Ashdod was the only ferry route that covered the eastern basin and environmental variability between 0-0.2. Venice-Patras was the only ferry route encompassing the Adriatic Sea and environmental variability 0.6-1. These two ferry routes were therefore used to create the environmental subnetwork as they covered all environmental variability in the study area (Figure 4.2D).

We also considered how community composition differed between the ferry routes. For each ferry route, species occurrences were extracted from each grid cell of the 'perfect knowledge' SSDM that overlapped with the ferry route. The number of grid cells that a species occurred in per route was treated as an abundance estimate. We applied a Hellinger transformation to the resulting species abundance x ferry route matrix to dampen the inflated abundances from longer ferry routes (Legendre & Gallagher, 2001). This transformed matrix was then used to create a Bray-Curtis dissimilarity matrix and differences in species composition between ferry subnetworks were quantified by Nonmetric Multidimensional Scaling (NMDS). The NMDS analysis confirmed, as expected, that ferry routes closer together had more similar species composition, with the main cluster formed from routes in the north-western basin (Appendix Figure A4.1). This cluster was used to create a deliberately biased ferry route subnetwork (Figure 4.2B). We also used the NMDS analysis to reduce the number of ferry routes from the original ferry route network by randomly selecting one ferry route from each cluster on the NMDS plot to create a subnetwork representing community composition. This reduced the number of ferry routes in the original network from 15 to 9 (Figure 4.2C).

We also produced ferry subnetworks with differing numbers of ferry routes, including 2, 4, 6, 8, 10 and 12 ferry routes by randomly selecting routes from the original ferry route network to evaluate the importance of the number of ferry routes. We built 10 sampling SSDMs using 50 regular sampling points per ferry route subset and compared to the 'perfect knowledge' SSDM with Pearson's correlation coefficient. We assessed the difference between biased, community and environmental subnetworks, and the difference between subnetworks with differing numbers of ferry routes with one-way ANOVAs. We performed post-hoc pairwise comparisons with the Tukey's Test.

4.2.4. Taxonomic biases in data collection

The sampling SSDMs were constructed with occurrence data from every sampling point that overlapped with the species distribution. Realistically, no methods for collecting biodiversity data have perfect rates of detectability, so understanding how imperfect detection affects predictions of biodiversity patterns or gradients in biodiversity is important. All biodiversity monitoring techniques, including eDNA metabarcoding, suffer from taxonomic biases (Balint et al., 2018). However, it is unclear how such uncertainty can in turn bias SSDM predictions. To quantify the effect of taxonomic bias, we either removed taxa (Chondrichthyes or Mammalia) or a random subset of species before individual species distribution models were stacked. The random species subset removed the same number of species as the equivalent taxonomically biased model. The models were then compared to the 'perfect knowledge' SSDM using Pearson's correlation coefficient. We analysed the effect of removing specific taxa with a three-way ANOVA and post-hoc pairwise comparisons with the Tukey's Test.

4.3. Results

4.3.1. Stacked species distribution model

The SSDM of marine predatory megafauna in the Mediterranean revealed two main gradients in species richness (Figure 4.3). There was higher species richness in the north-western basin compared to the south-eastern basin, and higher species richness nearer to shore. The environmental variable with the greatest influence on model predictions was mean sea surface temperature, whilst the variable with the least influence was bathymetric slope (Appendix

Table A4.2). The remaining variables, mean bathymetry, mean chlorophyll concentration, mean temperature range and distance from shore, contributed equally to model predictions. The model tended to overpredict species richness although the extent varied greatly (species richness error mean = 19.06 ± 7.23 s.d). The proportion of presences that were correctly predicted (sensitivity = 0.98 ± 0.12 s.d) was much higher than the proportion of absences correctly predicted (specificity = 0.54 ± 0.17 s.d; Appendix Table A4.3).

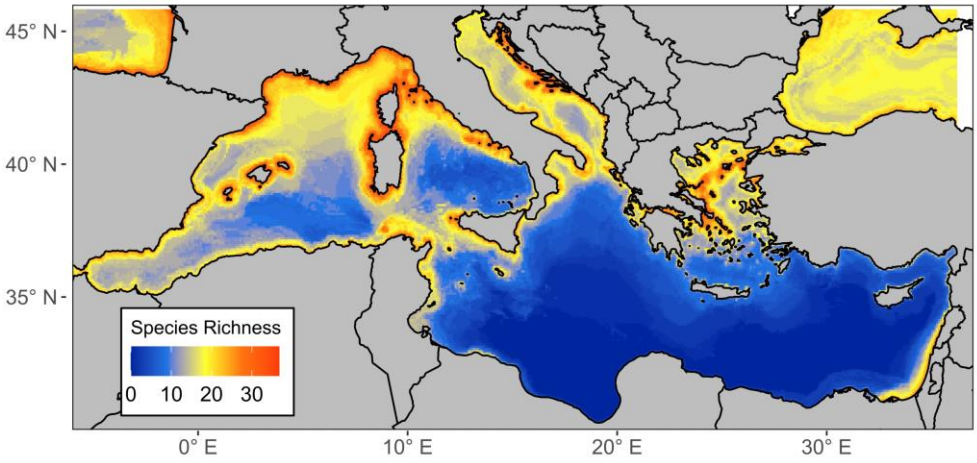


Figure 4.3. Original binary stacked species distribution model of 43 marine predatory megafauna in the Mediterranean using occurrence data obtained from online repositories.

4.3.2. Comparison of ferry route sampling frame to whole Mediterranean

The number of species with enough occurrences for modelling (>20) was consistently higher for samples collected along the ferry route network compared to unconstrained sampling across the Mediterranean Sea (Figure 4.4A). For the smallest number of sampling points (25), only the ferry routes could detect any species with enough occurrence points for modelling (random = 6.3 ± 1.27 s.d, regular = 6.17 ± 0.91 s.d). With 50 sampling points, the ferry routes (random = 18.56 ± 1.53, regular = 18.69 ± 1.99 s.d) detected double the amount of species compared with the Mediterranean (random = 9.42 ± 2.37, regular = 9.67 ± 1.74). The sampling strategy, random vs regular, had no effect on the number of species detected in both the ferry route and whole Mediterranean simulated sampling. The number of species detected increased quickly at small sample sizes but asymptotes between 200 and 500

sampling points where only 5 new species were detected using the ferry route network, and 7 species using the Mediterranean.

Sampling SSDMs created from 100 regular sampling points across the Mediterranean were most correlated to the 'perfect knowledge' SSDM (85.2% ± 4 sd) (Figure 4.4B). Sampling SSDMs produced from 100 sampling points collected regularly (82.2% ± 3 sd) or randomly (78.3% ± 6 sd) across the ferry route network also produced SSDMs highly correlated with the 'perfect knowledge' SSDM. Sample size and sampling strategy had less effect on the predictive capacity of sampling SSDMs produced with the ferry route network compared with the Mediterranean Sea. Sampling SSDMs created with either 25 or 50 sampling points along the ferry routes correlated more with the 'perfect knowledge' SSDM compared with 50 sampling points across the Mediterranean Sea regardless of sampling strategy ($F_{(1,373)}$=15.8, p <0.001; Table 4.1). Sampling SSDMs created from randomly allocated sampling points correlated less with the 'perfect knowledge SSDM compared to sampling SSDMs with regularly spaced sampling points. The difference in predictive capacity between the two sampling strategies was greater for samples collected using the Mediterranean instead of the ferry route network ($F_{(1,373)}$=3.91. p = 0.05).

Table 4.1. Three-way ANOVA table to evaluate impact of sampling strategy, sampling frame and number of sampling points on correlation coefficients between the sampling SSDMs and the 'perfect knowledge' SSDM.

Factor	Df	Sum Sq	Mean Sq	F value	P value
Strategy	1	0.4564	0.45651	37.4637	<0.001*
Size	2	1.1520	0.57601	47.2701	<0.001*
Sampling frame	1	0.0960	0.09603	7.8804	0.005*
Strategy:Size	2	0.0471	0.02355	1.9324	0.15
Strategy:Sampling frame	1	0.0477	0.04770	3.9147	0.049*
Size:Sampling frame	1	0.1925	0.19249	15.7964	<0.001*
Strategy:Size:Sampling frame	1	0.0097	0.00966	0.7927	0.37

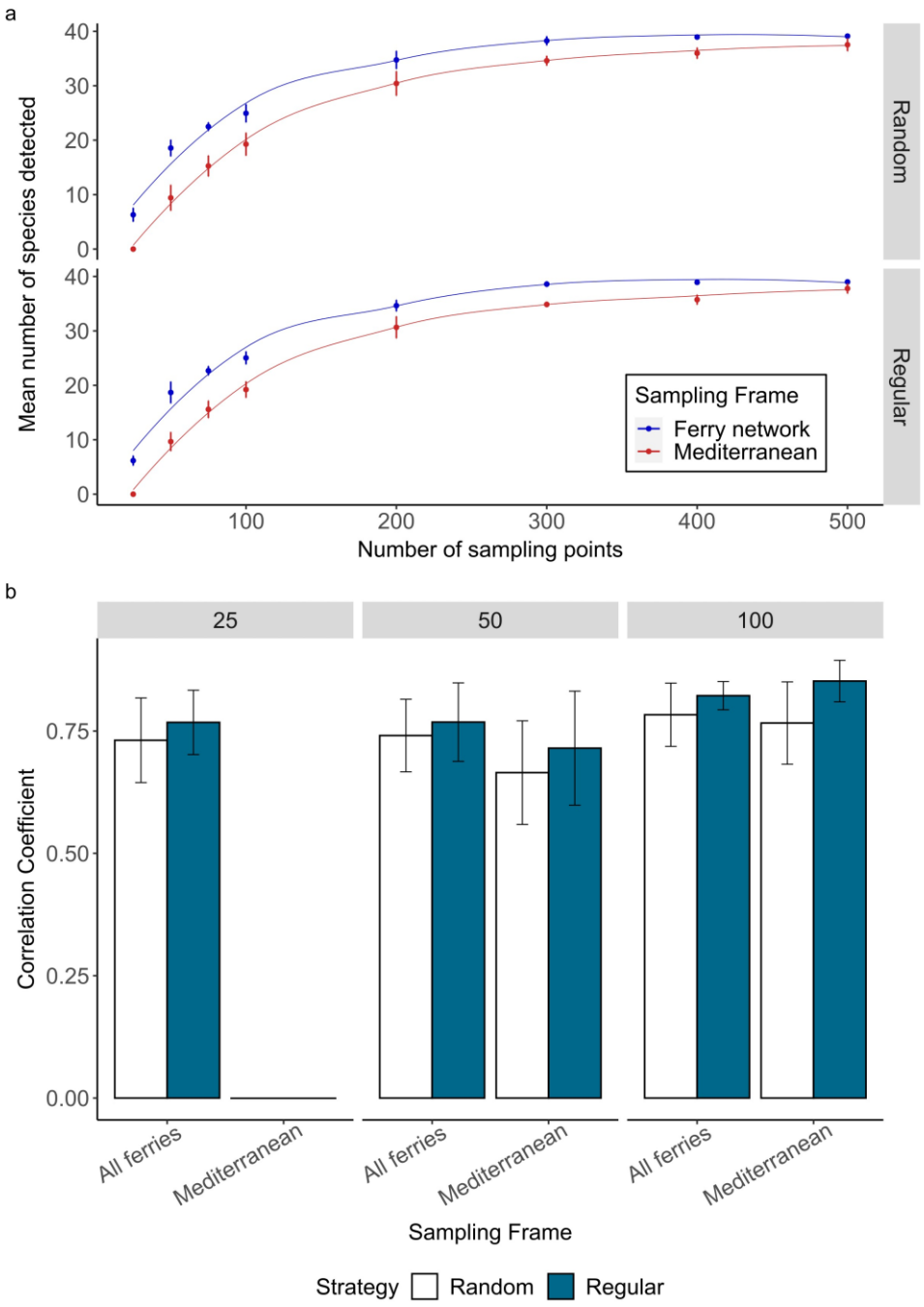


Figure 4.4. a) The mean number of species detected, with standard deviation bars, across different number of sampling points using either the ferry network or Mediterranean as a sampling frame and either random or regular sampling strategy. B) Mean Pearson correlation coefficient between the original SSDM and sampling SSDMs for 40 replicate simulations across the ferry network and Mediterranean for 2 sampling strategies (random and regular) across 3 different sample sizes (25, 50 and 100 sampling points). There was not enough occurrence data with 25 sampling points and the Mediterranean as a sampling frame to remake SSDMs.

4.3.3. Ferry route subnetworks

Different ferry subnetworks varied in their ability to accurately capture community composition in the 'perfect knowledge' SSDM ($F_{(1, 26)}$=342.96, $p < 0.001$) (Figure 4.5A; Appendix Table A4.4). The community subnetwork was able to predict the original community composition ~40% better than either the biased or environment subnetworks (Tukey's, $p<0.05$). The community subnetwork also had a similar climatic bias index to the network with all ferry routes included (Appendix Table A4.5). The environment subnetwork predicted community composition (34.3% ±0.6) ~9% worse than the deliberately biased subnetwork (43.8% ±0.4) (Tukey's, $p<0.05$). The deliberately biased sampling strategy had the highest climatic bias index, whilst the environment subnetwork performed similarly to the original ferry network (Appendix Table A4.5). The number of ferry routes included in a network affected its predictive capacity ($F_{(1, 54)}$=15.286, $p<0.001$), with correlation to the 'perfect knowledge' SSDM increasing from 32.8% ±13 for networks with 2 ferry routes, to 69.3% ±16 with 8 ferry routes (Tukey's, $p<0.001$)(Figure 4.5B). Increasing beyond 8 routes does not improve the predictive capacity of the sampling frame but reduces the variability related to which ferry routes are selected in the sub-network (Tukey's, $p>0.05$).

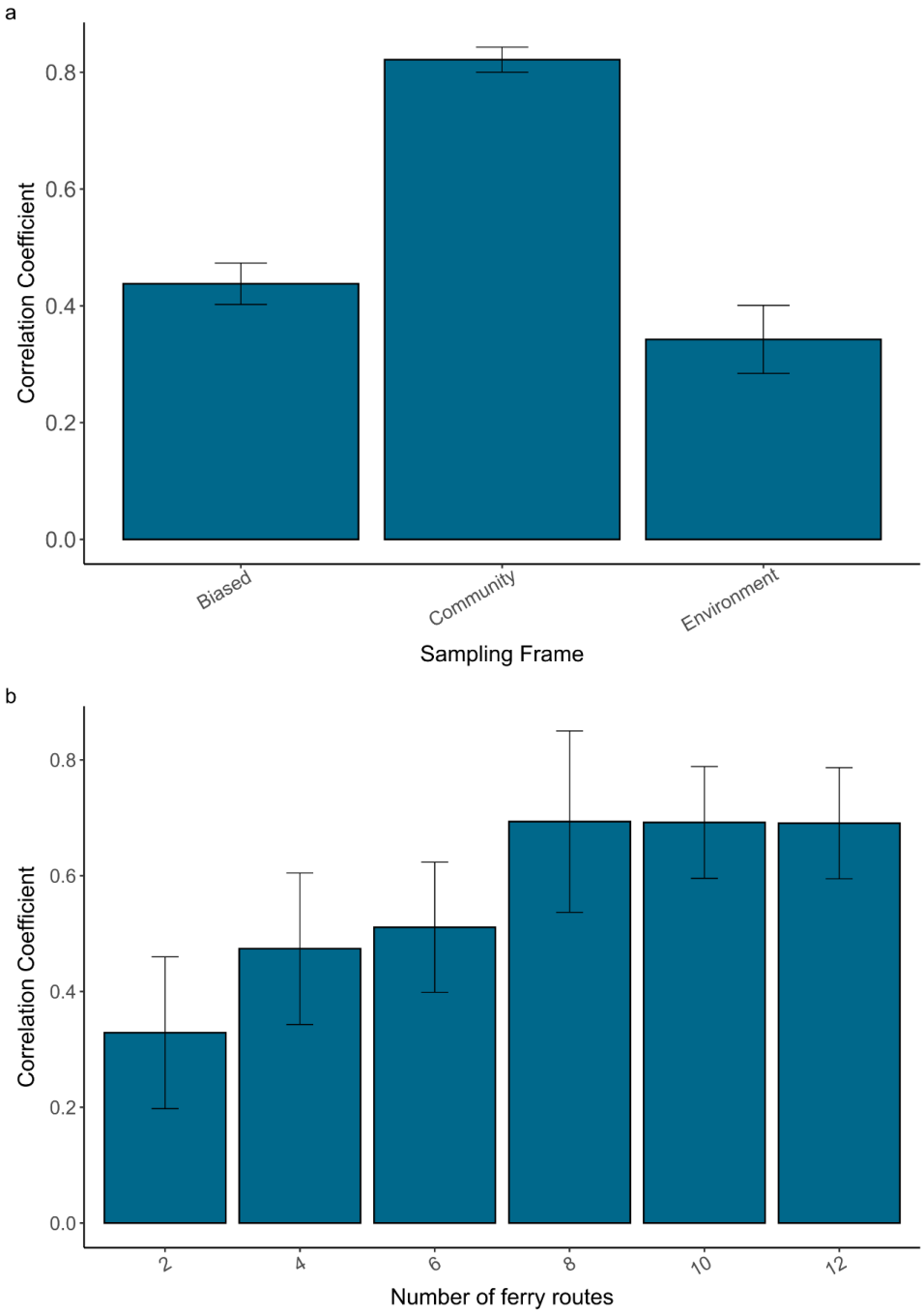


Figure 4.5. a) Mean Pearson correlation coefficient between the original SSDM and sampling SSDMs for 10 replicate simulations across ferry route subnetworks using 50 regular sampling points. B) Mean Pearson correlation coefficient between original SSDM and sampling SSDMs for 10 replicate simulations across subnetworks with differing numbers of ferry routes.

4.3.4. Taxonomic biases in data collection

When species in the same class were stacked together and compared to the 'perfect knowledge' SSDM, Actinopterygii (91.9%) and Chondrichthyes

(90.9%) both had similar species richness patterns to the 'perfect knowledge' SSDM (Figure 4.6). The Mammalia only SSDM (67.96%) showed weaker correlation with the 'perfect knowledge' SSDM. Sampling SSDMs with different taxa removed affected the predicted community composition ($F_{(1,116)}$=8.72, p<0.001)(Figure 4.7). Sampling SSDMs with Mammalia species removed improved the predictive capacity by 10% compared to sampling SSDMs with Chondrichthyes removed, or by 7% compared to random subset of species removed (Tukey's, p<0.001). This pattern was consistent across a range of sampling sizes and strategies.

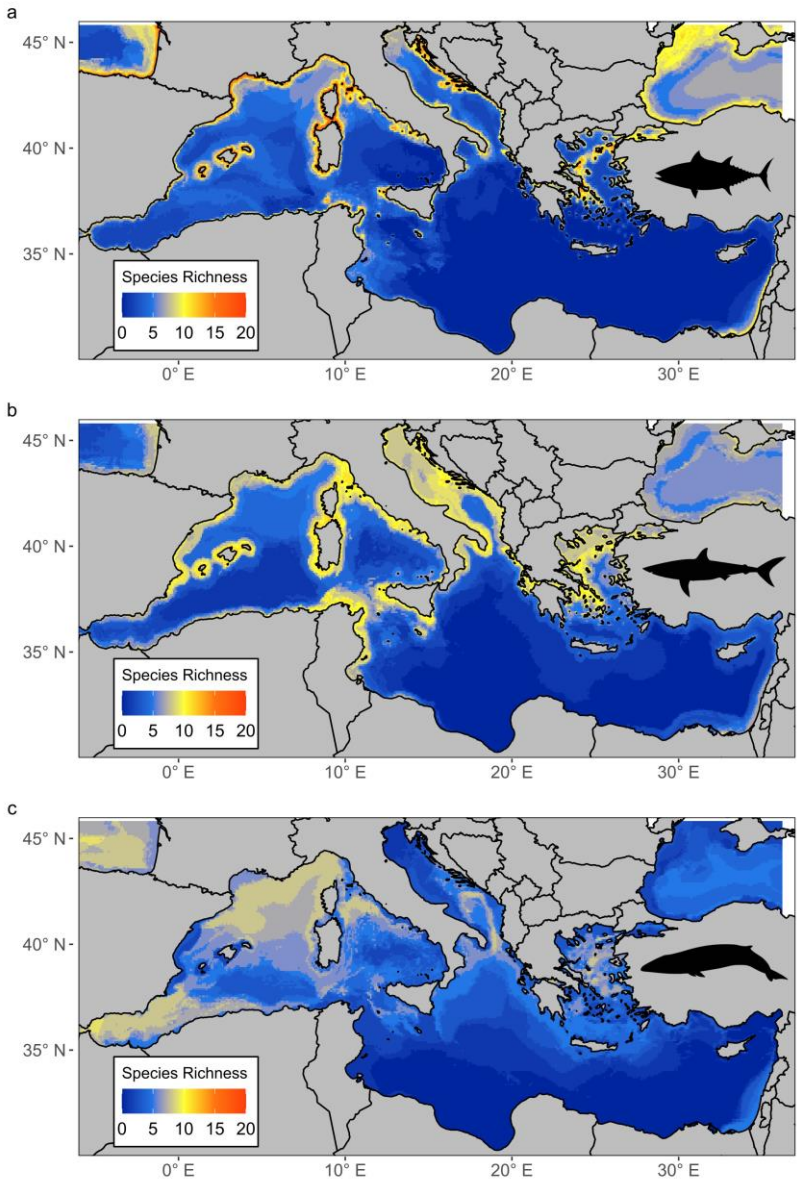


Figure 4.6. Stacked species distribution models for Class a) Actinopterygii, b) Chondrichthyes, c) Mammalia.

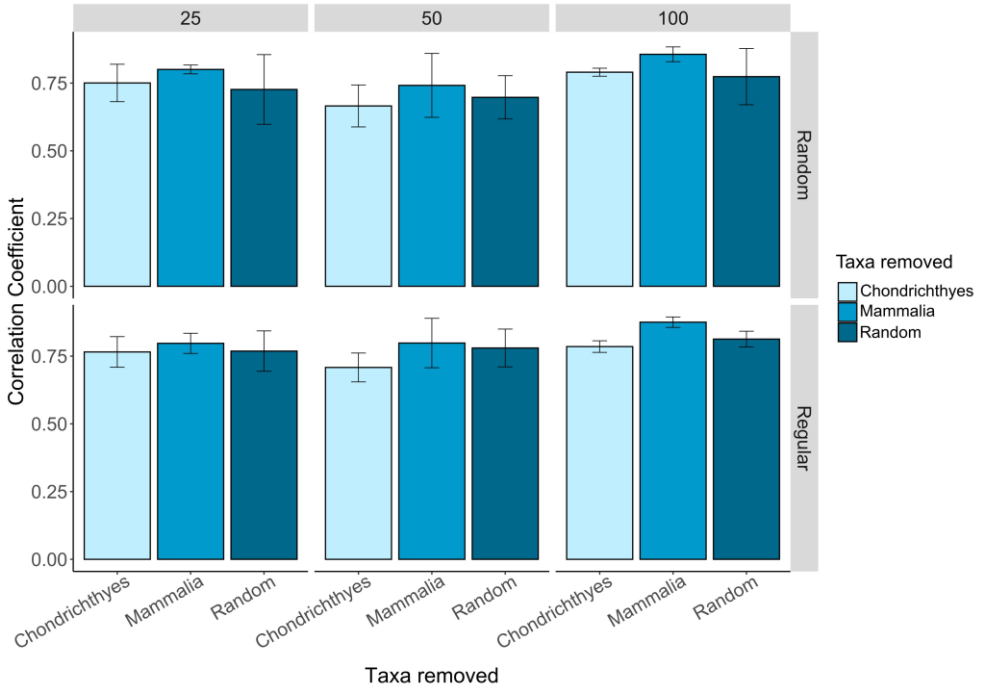


Figure 4.7. Mean Pearson correlation coefficient between the original SSDM and sampling SSDMs with different taxonomic biases for samples collected using the ferry network sampling frame.

4.4. Discussion

Biased sampling remains a key hurdle to predicting biodiversity patterns (Hughes et al., 2021; Moussy et al., 2022; Tydecks et al., 2018). We evaluated the feasibility of using biased sampling frames (in this case commercial vessels) as sampling platforms for collecting species occurrence data for marine species distribution modelling. In this study, we test ferry routes that could offer low-cost access to vessels (compared to dedicated research cruises) for hard-to-reach pelagic regions, but introduce biases because spatial sampling is restricted to the routes covered. We found that the inherent bias associated with restricted sampling frames did not lead to a loss in predictive capacity. In fact, for our case study, sampling simulations with ferry routes recovered species richness gradients more accurately than unconstrained sampling at small (25 sampling points) and medium (50 sampling points) sample sizes as a result of ferry routes constraining sampling to areas with higher biodiversity. This result further highlights the cost-effectiveness of ferry routes as sampling platforms and demonstrates that high quality biodiversity data can be recovered from restricted sampling frames.

Implementing this workflow to design surveys across the global shipping network, including from other vessel types (e.g. container ships), could vastly expand our knowledge of marine biodiversity in inaccessible areas, and is especially applicable for expanding the spatiotemporal scale of emerging techniques, such as automated environmental DNA sampling (Valsecchi et al., 2021).

4.4.1. Marine predatory megafauna SSDM

The SSDM shows that predatory megafauna species richness is much higher in the north-western basin (Figure 4.3). This result is unsurprising due to the Strait of Gibraltar linking the western basin to the Atlantic Ocean allowing migration of megafauna into the Mediterranean (Coll et al., 2010). Critical habitat, including breeding and foraging grounds, for marine megafauna have been recognised in the north western basin through Ecologically or Biologically Significant Areas (EBSAs), and the implementation of the Pelagos Sanctuary for Marine Mammals (Notarbartolo-di-Sciara et al., 2008; UNEP/CBD/EBSA/WS/2014/3/4, 2014). However, there was also a greater density of occurrence points used to create the 'perfect knowledge' SSDM in the northwestern region compared with offshore and in the southern basin (Appendix Figure A4.2). This sampling bias is driven by greater economic resources in northern basin countries which benefit from European Union (EU) funding for survey and conservation initiatives (Amengual & Alvarez-Berastegui, 2018; Coll et al., 2010). The binary SSDM tended to overpredict species richness, as has been previously reported (Pottier et al., 2013). Combining SSDMs with macroecological constraints may reduce overprediction by accounting for biotic interactions (d'Amen et al., 2015; Guisan & Rahbek, 2011). However, SSDMs can provide similar predictions to macroecological models or joint species distribution models when using a probabilistic stacking approach (Calabrese et al., 2014; Zurell, Zimmermann, et al., 2020). Despite its limitations, we chose to use a binary stacking procedure as we required presence data to re-run species distribution models from the simulated sampling strategies and the model represents realistic community patterns as a base for sampling simulations.

4.4.2. Comparison of ferry route sampling frame to whole Mediterranean

Our selected 15 operational ferry routes are assumed to be representative of the spatial extent of the Mediterranean wide ferry network (Figure 4.2A). Using this ferry route network as a sampling frame achieved species distribution models that predicted the known community from the 'perfect knowledge' SSDM as well as or better than samples collected across the whole Mediterranean. Ideally, occurrence data for species distribution modelling would represent a random sample from the population of interest across the entire study area (Araújo & Guisan, 2006). However, geographically biased sampling strategies, *i.e.*, samples only collected close to road networks, can still produce accurate models as long as the environmental predictors are not also biased, as is the case with the ferry route network (Kadmon et al., 2004; Tessarolo et al., 2014). Here, we demonstrate that with smaller sample sizes, samples collected from the biased sampling frame produced more accurate models than samples collected from across the whole Mediterranean Sea (Figure 4.4B). It is more feasible to routinely collect samples on board ferries than to implement dedicated research surveys over large spatial scales comparable to the Mediterranean Sea. Therefore, we show that routine sampling on ferries can serve as an important approach to conduct representative biodiversity sampling.

Fewer samples are required to produce models with similar accuracy from ferry routes compared with the whole Mediterranean, but smaller sample sizes result is less species being detected. For the ferry route network, there is no cost benefit to doubling the sample size as this does not improve the SSDM community composition prediction (Figure 4.4). However, the SSDM made with 25 sampling points only detected between 5-8 species (11-18%) whereas SSDMs with 50 sampling points detected 16-21 (37-48%) species, and SSDMs with 100 sampling points detected 19-26 species (44-60%). If the aim of the study is to look at patterns in species richness, such as gradients in diversity, then a small sample size is adequate. However, if individual species distributions, or the detection of rare species is also important, then larger sample sizes will be required. These sample sizes are based on 100% detection rate of the species when they are present which is unrealistic for any

sampling method. However, we expect that the patterns observed between sample sizes and sampling frames should hold true as long as the detection probabilities are constant across sampling frames. Sampling SSDMs from the ferry networks were less affected by sampling strategy than sampling SSDMs from the whole Mediterranean, where random sampling consistently produced more poorly performing SSDMs. By limiting the available sampling frame to such an extent, this potentially reduces the impact of sampling strategy, and prevents random sampling from forming clusters which do not cover the study area's environmental variability (Zhang et al., 2020). These results suggest that ferries, or other commercial shipping routes, represent a promising sampling platform to alleviate constraints on access to pelagic environments that currently limit marine biodiversity surveys.

4.4.3. Differences between ferry routes and subnetworks

Environmental variability and species composition were compared between individual ferry routes to understand which ferry routes were important when building a subnetwork. The routes between Salerno-Ashdod and Venice-Patras were the only two routes that covered the extremities of environmental variability and so were required in any ferry subnetwork to achieve full coverage of the environmental parameter space. Previous research suggests that sampling frames can be geographically biased as long as the full range of environmental variability in the whole sampling area is covered (Kadmon et al., 2004; Tessarolo et al., 2014). However, our results demonstrate that the environment subnetwork was not able to accurately predict community composition despite covering environmental variability, and in fact performed similarly to the deliberately biased subnetwork (Figure 4.5A). This highlights that considering environmental variability alone may not reduce the biases associated with restricted sampling frames. The NMDS analysis showed that the routes covering Salerno-Ashdod and Venice-Patras do not cluster with any other routes suggesting that different species compositions occur on these routes (Appendix Figure A4.1). Meanwhile, the community subset, which covered both community composition and environmental variability, predicted species richness in the 'perfection knowledge' SSDM most accurately. This result highlights that community composition as well

as environmental variability must be considered when selecting ferry routes to be representative sampling frames.

The original ferry route network had a high density of shipping routes in the northwest Mediterranean, coinciding with the region with most biodiversity data available, which we expected to bias the predictive capacity of the sampling SSDMs using this network (Figure 4.2). However, the community subnetwork, with six routes removed from the northwest basin, was still able to accurately predict community composition suggesting this was not a driving factor in the effectiveness of the ferry routes as a sampling frame (Figure 4.5). A limitation of using existing community composition knowledge to select ferry routes for sampling is that it requires reliable occurrence data to model a 'perfect knowledge' SSDM. Here, the NMDS analysis shows that community composition along the ferry routes is related to the geographical location of the routes, with routes closer together having more similar community composition (Appendix Figure A4.1). We also demonstrated that increasing the number of routes within the network and having fewer sampling points along more routes, will lead to improved predictions of community composition. Therefore, we would recommend implementing a large number of ferry routes, at least 8, that cover as many different regions of a study area as possible if pre-existing occurrence data is unreliable or limited.

4.4.4. Random and systematic biases in data collection

Reports identifying taxonomic biases in biodiversity surveys are pervasive in the literature, but little is known about how taxonomic biases can affect downstream analyses such as species distribution modelling or spatial planning (Di Marco et al., 2017; Donaldson et al., 2016; Troudet et al., 2017). Instead, efforts to reduce bias in species distribution models have largely been directed at spatial and temporal biases in data collection (Beck et al., 2014; Inman et al., 2021; Kramer-Schadt et al., 2013). We demonstrate that different taxa have varying species richness gradients, thus removing different taxonomic groups affected which species richness gradients were revealed. The classes Actinopterygii (fishes) and Chondrichthyes (sharks and rays) both showed highest species richness closest to shore whereas marine

mammals were more prevalent offshore. This is unsurprising as Actinopterygii and Chondrichthyes are more closely related, and are largely ectothermic so more constrained by temperature requirements than marine mammals (Grady et al., 2019; Losos, 2008). However, this may have been exaggerated by greater availability of marine mammal data offshore from visual ferry surveys compared to Actinopterygii and Chondrichthyes data which is largely collected by coastal fisheries (Aïssi et al., 2015; Mancusi et al., 2020). Models with marine mammals removed were more correlated to the 'perfect knowledge' SSDM as a result of more species belonging to the Actinopterygii and Chondrichthyes classes than marine mammals (Figure 4.7). This highlights that the proportion of species representing each class has an important influence on the overall species richness gradients captured. If biases lead to certain taxonomic groups being underrepresented then it is unlikely that their species richness gradients would be adequately captured, unless they follow similar distributions to another taxa. To utilise novel methods for biodiversity data collection most effectively, it is important to understand the effect taxonomic bias can have, and how new methods can best reduce current biases.

4.4.5. Conclusion/Future research

Our study demonstrates that high quality biodiversity data can be collected from biased sampling frames, providing they cover wide areas and diversified habitats. Utilising these biased sampling frames, such as ferries, allows data collection from challenging and remote areas which are often inaccessible to researchers due to logistical and financial constraints. This is particularly relevant for upscaling sampling for emerging biodiversity monitoring techniques, such as automated eDNA sampling, to reduce current spatial, temporal and taxonomic biases (Pawlowski et al., 2020). This study focused on the ferry routes in the Mediterranean to carry out simulated sampling strategies, but sampling from ferry routes, as well as other commercial vessel types, could be carried out across the global shipping network. The efficiency for ferry routes as sampling platforms will depend on the concentration of ferry routes in the study area or region of interest. Global cargo routes are largely concentrated in the North Atlantic, North Pacific and Indian Oceans, linking Europe, North America, East and Southeast Asia. High traffic routes

crossing the South Atlantic and South Pacific also connect with Southern Africa, South America and Australasia. These represent key areas where commercial vessels could contribute to closing gaps in biodiversity data (Wang & Wang, 2011). These areas also coincide with those most affected by human impacts emphasising the need for regular monitoring to understand effects on biodiversity (Halpern et al., 2008; Pirotta et al., 2019). The workflow presented here can be used as a template to evaluate the efficiency of a shipping route network in a study area of interest before undertaking sampling. This study focused on the impact of sampling strategies on species distribution models, which are frequently used as conservation features in marine spatial planning to designate protected areas. Therefore, our findings confirm that biased sampling, if designed adequately, can provide a useful data basis for marine species and management of marine environments.

Chapter 5 - Multi-taxon versus taxon specific spatial conservation priorities for predatory marine megafauna

Authors: Elizabeth Boyse[1], Simon J. Goodman[1], Maria Beger[1]

[1]School of Biology, University of Leeds, Leeds, United Kingdom

5.0. Abstract

Predatory marine megafauna are globally threatened by anthropogenic stressors, but are key for ecosystem functioning. Their worsening conservation statuses indicate that current management is failing, requiring us to urgently reimagine their conservation needs to ensure their survival. Their life histories, threats, and resource needs are diverse. Consequently, spatial conservation areas targeting all species will overlook such heterogeneity, contributing to the problem. Here, we model 42 predatory marine megafauna species distributions (marine mammals, elasmobranchs, teleost fishes) in the Mediterranean Sea using available biodiversity data to highlight diversity among species richness gradients for separate taxonomic groups. Secondly, we employ the Marxan spatial planning decision making tool to identify priority conservation areas for the different taxonomic groups and quantify overlap with the current marine protected area (MPA) system. Different marine megafauna taxonomic groups had heterogenous distributions, resulting in drastically different spatial conservation priority areas. None of the marine megafauna are sufficiently covered by Mediterranean MPAs (< 30% coverage), with marine mammals being the least protected despite having the greatest designated MPA extent, highlighting disconnects between conservation goals and current management outcomes. To save marine megafauna, taxon specific ecological requirements and resulting spatial heterogeneity need to be accounted for in marine spatial planning.

5.1. Introduction

Predatory marine megafauna (mammals, elasmobranchs and teleost fishes) are globally declining due to anthropogenic pressures including fishing, climate change and habitat degradation (Avila et al., 2018; Dulvy et al., 2021). Their loss reduces important ecosystem functions such as top-down control, redistribution of nutrients, habitat engineering, and carbon sequestration (Hammerschlag et al., 2019). They also create millions of job opportunities through being heavily targeted commodities in multi-billion dollar industries, global fisheries and eco-tourism (Cisneros-Montemayor et al., 2010; Juan-Jordá et al., 2011). Marine megafauna are typically either protected within marine protected areas (MPAs) planned across multiple species and habitats, or focused on popular taxa, *i.e.,* marine mammals (Notarbartolo di Sciara et al., 2016), but both approaches neglect the heterogeneous habitat requirements of megafauna stemming from their complex life histories, contributing towards their ineffectiveness (Klein et al., 2015). Further, marine megafauna home ranges typically exceed plausible sizes for MPAs, allowing exposure to threats outside MPA boundaries (Conners et al., 2022).

Marine megafauna includes some of the most globally threatened taxa, with over a third of marine mammals and elasmobranchs classified as threatened in IUCN Red List assessments, and targeted fisheries causing the loss of 90% biomass of large predatory fishes (Avila et al., 2018; Dulvy et al., 2021; Myers & Worm, 2003). Fishing is the largest threat to all marine megafauna through resource exploitation, direct harvesting and bycatch. Further, climate-induced habitat degradation and range shifts are increasingly prominent threats (Avila et al., 2018; Dulvy et al., 2021). The impacts of shipping traffic are well established for marine mammals, but less certain, although potentially significant, for sharks or teleost fishes (Schoeman et al., 2020). The loss of marine megafauna from ecosystems causes trophic cascades and reduced resilience to climate change (Estes et al., 2016). They also have high cultural and economic significance which can be both a benefit (*i.e.*, high conservation interest) and a detriment (*i.e.*, drives high demand) (Estes et al., 2016). Therefore, we urgently need to improve understanding of

the conservation requirements for predatory marine megafauna to prevent further declines.

Marine spatial planning (Directive 2014/89/EU) offers a coordinated and transparent approach to managing different stakeholders using the marine space, whilst minimising impacts to the environment. Incorporating marine megafauna into spatial planning is important to indicate ecologically significant areas, *i.e*., with high productivity, species diversity or biomass of prey species (Augé et al., 2018). Incorporating multi-taxa approaches to identify priority areas has been advocated (Augé et al., 2018), but this oversimplifies the diverse habitat requirements of different taxonomic groups (Heupel et al., 2019). In contrast, megafauna-specific MPAs focus on narrower objectives, *i.e*., protecting specific life stages such as breeding or feeding grounds, but this strategy will only be effective if the protected life history stages maximise population growth rates (Conners et al., 2022). Different marine megafauna taxa have distinct spatial requirements, affecting their susceptibility and exposure to different threats (Avila et al., 2018). For example, divergent thermoregulatory strategies mean that marine mammals represent highest megafaunal richness in temperate and polar waters, and in pelagic zones, while sharks and teleost fishes dominate tropical and coastal waters (Grady et al., 2019). At finer-scales, spatial partitioning between species is driven by mechanisms such as competitive exclusion and varying life history strategies (Heupel et al., 2019). Attempting to maximise conservation benefits for taxonomic groups offers a suitable balance between species-specific approaches, which will not afford protection to unstudied species, and broad (all taxa) biodiversity objectives that fail to account for taxon-specific requirements.

The Mediterranean Sea hosts a high diversity of predatory marine megafauna that are exposed to some of the highest human impacts globally (Coll et al., 2010). Consequently, the Mediterranean Sea is an extinction risk hotspot for elasmobranchs (Dulvy et al., 2021), and local extinctions of marine mammal and teleost fish populations have already occurred (Bearzi et al., 2008; MacKenzie et al., 2009). Only 6% of the Mediterranean Sea is covered by MPAs, and of these, 95% have no regulations in place, owing to most being coastal and coinciding with high vessel density areas resulting in stakeholder

conflicts (Claudet et al., 2020). Transboundary marine spatial planning has been encouraged given the large number of relatively small countries bordering the Mediterranean Sea (Li & Jay, 2020), yet marine spatial planning for marine megafauna has so far focused on single species or small spatial scales (Carlucci et al., 2021; Mazor et al., 2016). To ensure that spatial planning results in the best conservation benefits for marine megafauna, it needs to be executed at scales relevant to the expansive spatial ranges of marine megafauna (Conners et al., 2022; Estes et al., 2016).

In practice, conservation practitioners use all available biodiversity information for spatial planning prioritisations across taxa, and typically omit taxon-specific requirements. In this paper, we test whether taxa with different spatial ranges and habitat requirements need different conservation priority areas in the Mediterranean Sea. We firstly model the distributions of marine mammals, elasmobranchs, and large teleost fishes. Second, we identify separate and joint reserve prioritisation solutions for different taxa to test our expectation that each taxon requires specific conservation areas. Finally, we evaluate how different our reserve networks are compared to currently designated MPAs. We highlight discrepancies in the realised and required conservation efforts for marine megafauna and develop recommendations to facilitate the implementation of improved management measures for each group.

5.2. Methods

5.2.1. Species distribution modelling

We classify a predatory marine megafauna as having a total length ≥100 cm and a trophic level ≥4 based on FishBase (https://www.fishbase.se/) or SeaLifeBase (https://www.sealifebase.ca/) records (Boyse et al., 2023). Setting a threshold of ≥40 occurrence records from GBIF (https://www.gbif.org, June 2020, GBIF Occurrence Download https://doi.org/10.15468/dd.tqx2he), OBIS (https://obis.org/), EurOBIS (https://www.eurobis.org/), the Mediterranean Large Elasmobranchs Monitoring (Medlem) database and Accobams (ACCOBAMS Survey Initiative, 2020; Mancusi et al., 2020), data were available to model distributions of 42 marine megafauna species, covering three taxonomic

groups (20 teleost fishes, 13 elasmobranchs, 9 marine mammals) (Appendix Table A4.1). We obtained data for six environmental predictors (bathymetry, bathymetric slope, chlorophyll *a*, distance from shore, sea surface temperature mean and sea surface temperature range) from Bio-ORACLE and Marspec in WGS84 projection and 0.83° x 0.83° resolution (Assis et al., 2018; Sbrocco & Barber, 2013). Prior to modelling, we spatially thinned occurrence records with a nearest neighbour distance of 10 km (Aiello-Lammens et al., 2015). We modelled species distributions using maximum entropy (MAXENT), multiple adaptive regression splices (MARS) and random forest (RF) algorithms with the SSDM R package (Phillips et al., 2006; Schmitt et al., 2017). We generated 10, 000 background points for MAXENT. MARS and RF require pseudo-absence data, which we created randomly using the two-degree method, with 1000 points for MARS and an equal number of pseudo-absences as presence data for RF (Barbet-Massin et al., 2012). We made ensemble models across the different algorithms using weighted AUC scores. We converted ensemble habitat suitability models for each species to presence-absence models using the sensitivity equals specificity threshold (Schmitt et al., 2017). Binary ensemble models were summed to produce stacked species distribution models to visualise patterns in species richness.

5.2.2. Priority areas for marine megafauna in the Mediterranean

We divided the Mediterranean Sea into planning units of 10 km x 10 km, in line with European Union guidelines on spatial planning (European Commission, 2007), resulting in 25,141 planning units in total, in the Lambert Azimuthal Equal Area projection (EPSG:3035). We assigned each planning unit an opportunity cost of displaced vessel traffic, represented by annual vessel density (hours per square kilometre) at a 1 km x 1 km resolution (European Marine Observation and Data Network, EMODnet; https://www.emodnet.eu/). We averaged annual vessel density across the available four years (2017-2020) and summed the data to a 10 km x 10 km resolution. Vessel density is a suitable surrogate for opportunity cost as this incorporates multiple sectors, including fishing, cargo, passenger and tankers which represent important threats for marine megafauna (Avila et al., 2018; Dulvy et al., 2021).

We employed the decision support software Marxan to identify priority areas for marine megafauna conservation. Marxan provides near-optimal solutions to the minimum set problem where conservation features (*i.e.*, species) are adequately represented for the least possible cost (Ball et al., 2009). Our conservation features consisted of individual marine megafauna binary species distributions for 42 species. We set a target of 30% protection for each species following the post-2020 Global Biodiversity Framework guidance (CBD, 2021). We ran Marxan using the simulated annealing algorithm and a boundary length modifier of 0.01 after calibration. We performed 100 iterations for each of the conservation feature scenarios: 1) all marine megafauna species, 2) marine mammals, 3) teleost fishes and 4) elasmobranchs.

5.2.3. Comparing conservation planning scenarios with different taxonomic information

We used both the selection frequency, *i.e.*, how many times each planning unit was selected in 100 iterations, and the ten solutions with the lowest objective scores to compare conservation priority differences across marine megafauna taxa. First, we quantified the overlap between planning unit selection frequencies of different taxa with Cohen's Kappa coefficient (McHugh, 2012). The Kappa statistic requires categorical data so we classified the selection frequencies into five groups, 0, <25, 26-50, 51-75, >75 (Ruiz-Frau et al., 2015). Second, we performed hierarchical clustering with Jaccard dissimilarities from the 10 best solutions across the conservation feature scenarios (Brumm et al., 2021). We also compared the average cost and number of planning units required across the different taxa.

We downloaded the most current database for MPAs in the Mediterranean from MAPAMED, which includes 1,126 designated MPAs (MedPAN & SPA/RAC, 2022)(Figure 5.1B). We included MPAs with a national statute, Natura 2000 sites, and Specially Protected Areas of Mediterranean Importance (SPAMI) (MedPAN & SPA/RAC, 2022). We calculated the overlapping area between species distributions and MPAs and our ten best spatial prioritisation solutions to quantify which taxa are currently receiving most protection, and differences among taxa-specific prioritisation solutions.

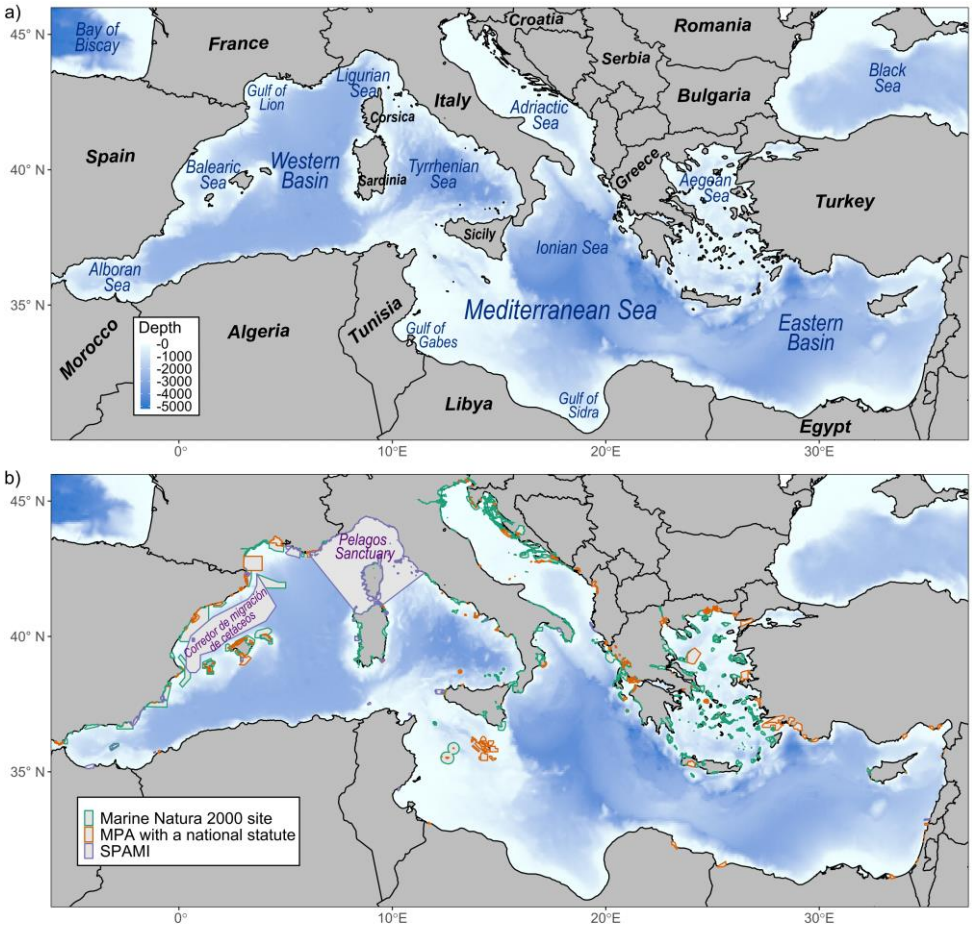


Figure 5.1. Bathymetric maps of the Mediterranean Sea showing a) marine regions, and b) designated marine protected areas (MPAs) including Marine Natura 2000 sites, MPAs with a national statute and Specially Protected Areas of Mediterranean Importance (SPAMIs).

5.3. Results

Different taxa of marine megafauna have distinct distribution patterns (Figure 5.2). High species richness of elasmobranchs and teleost fishes is found along the coastlines of the north-western basin as well as the Balearic Islands, Corsica and Sardinia. Elasmobranchs have wider distributions in the Adriatic Sea, Aegean Sea and along the coastlines of Tunisia and Sicily, compared to teleosts. Highest species richness of marine mammals occurs in the Alboran Sea and between the Balearic Islands and Corsica/Sardinia. Overall, there is clear decrease in species richness with distance from shore, with highest species richness occurring in the north-western basin as well as the Adriatic and Aegean Seas. The marine megafauna SSDM overpredicted species richness, with a high proportion of true presences predicted correctly

(sensitivity = 0.98 ± 0.12 s.d) but a lower proportion of absences predicted correctly (specificity = 0.54 ± 0.17 SD) (Appendix Table A4.3).

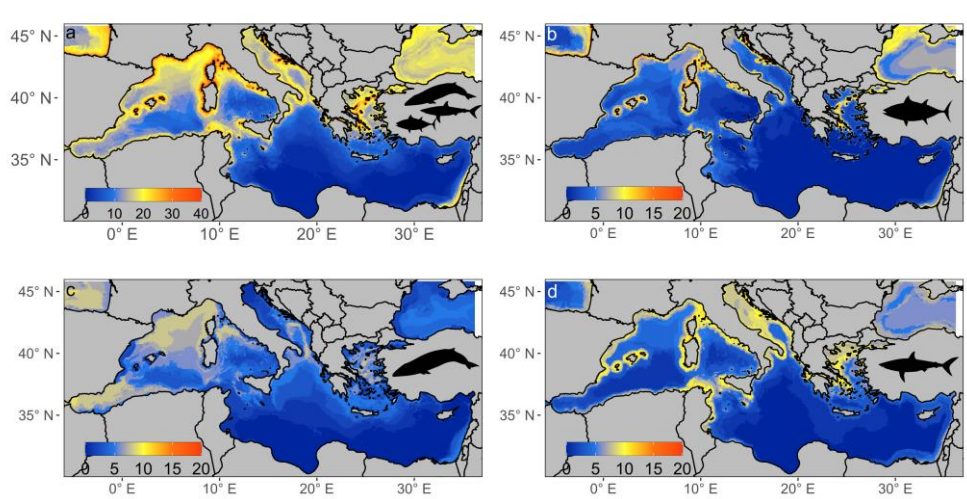


Figure 5.2. Maps of the Mediterranean Sea showing species richness from stacked species distribution models of a) all taxa, b) teleost fishes, c) marine mammals, and d) elasmobranchs.

Conservation feature scenarios considering taxa separately resulted in vastly different spatial prioritisation solutions (Figure 5.3). The Kappa statistic and hierarchical cluster analysis show highest similarity between selection frequencies for elasmobranchs and all taxa (Figure 5.4). Visually, elasmobranchs and all taxa scenarios share similar high selection frequency areas occurring along the coastlines of Tunisia and Egypt, the southern Alboran Sea, and the northern Aegean Sea. The Kappa statistic shows mammals and teleosts to have similar disagreement to the all taxa scenario, while cluster analysis reveals greatest dissimilarity between mammals and all taxa.

Including all taxa resulted in solutions with highest costs and greatest number of planning units (Appendix Figure A5.1). Marine mammal prioritisations required the lowest costs (81,939 ±2206) despite requiring the greatest number of planning units (5644 ±178) whilst elasmobranchs have the greatest costs (104,217 ±2043) despite needing a relatively similar number of planning units to mammals (5526 ±180).

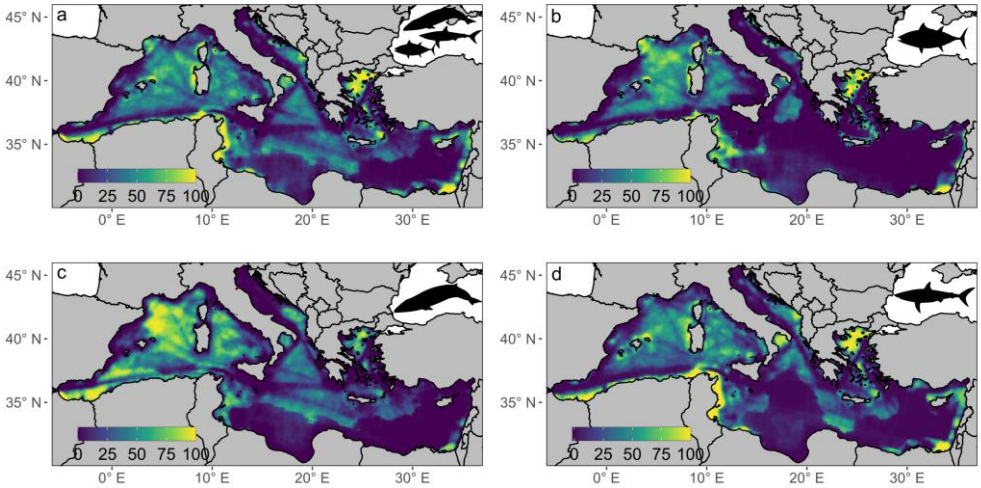


Figure 5.3. Maps of the Mediterranean Sea showing the selection frequency, *i.e.,* how many times each planning unit was selected in 100 iterations, from Marxan outputs for each of the taxa scenarios; a) all species, b) teleost fishes, c) marine mammals, and d) elasmobranchs.

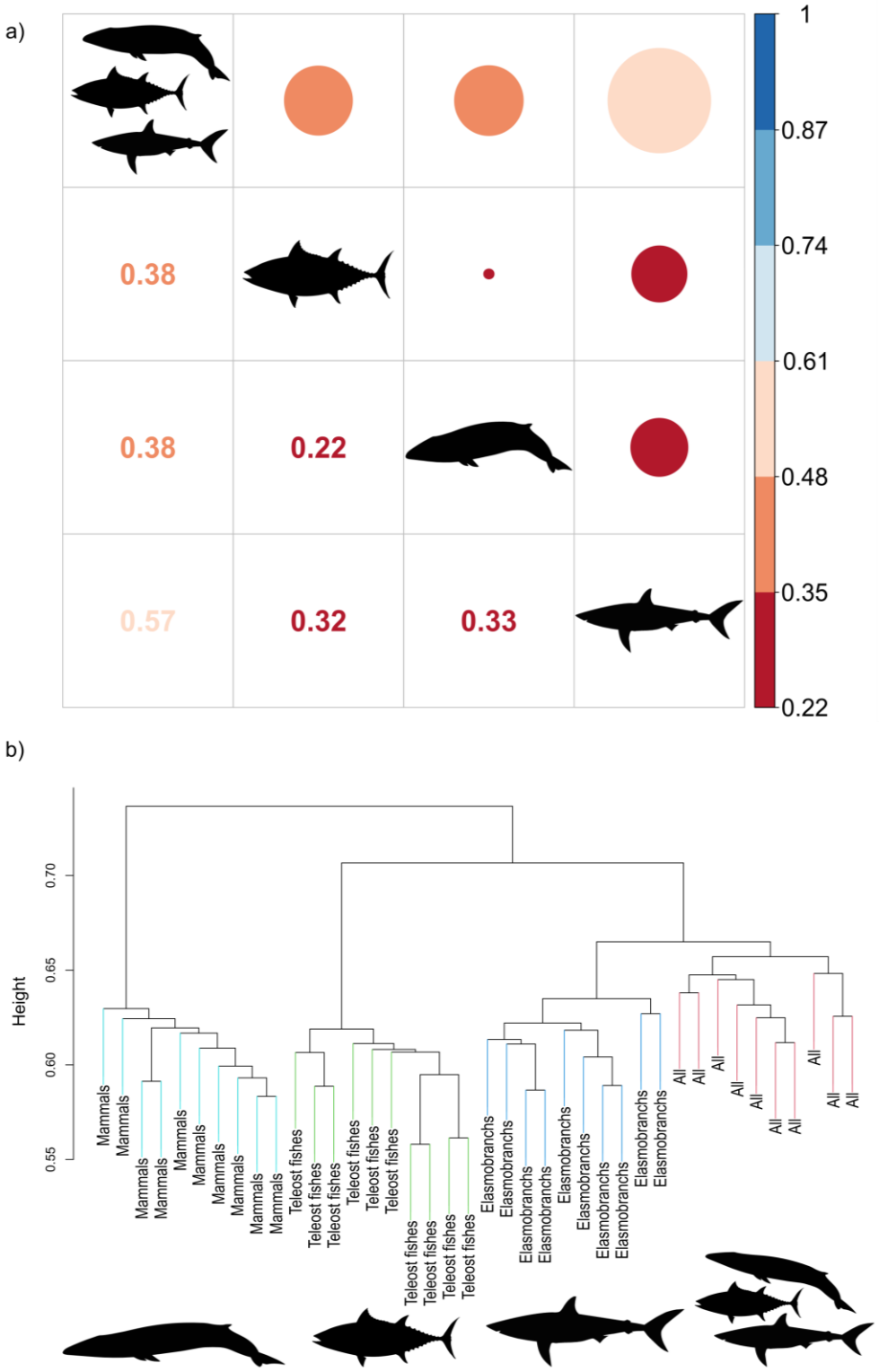


Figure 5.4. a) Cohen's Kappa coefficient showing similarity between selection frequency classes across the different conservation feature scenarios. b) Dendrogram displaying the average Jaccard distances between the ten best solutions across the different taxa conservation feature scenarios.

The Mediterranean MPA network does not fulfil the 30% coverage target for any of the marine megafauna taxa distributions (Figure 5.5). Teleost fish distributions overlap most with the current MPA network (24.12% ±13.27 s.d), whilst marine mammals (16.54% ±6.49) and elasmobranchs (18.58% ±10.81) share similar lower levels of protection. Overlap with the MPA network varied greatly for species within a taxonomic group, so the overall differences between groups were not significant (Kruskal-Wallis, $p>0.05$). Scenarios including a single taxonomic group resulted in greatest overlap with species distributions from that group (Kruskal-Wallis, $p < 0.05$) (Appendix Table A5.1). Our spatial prioritisation solution for marine mammals afforded the least co-protection to other taxa, with teleost fish distributions overlapping 18.42% ±11.08 and sharks 19.92% ±11.06.

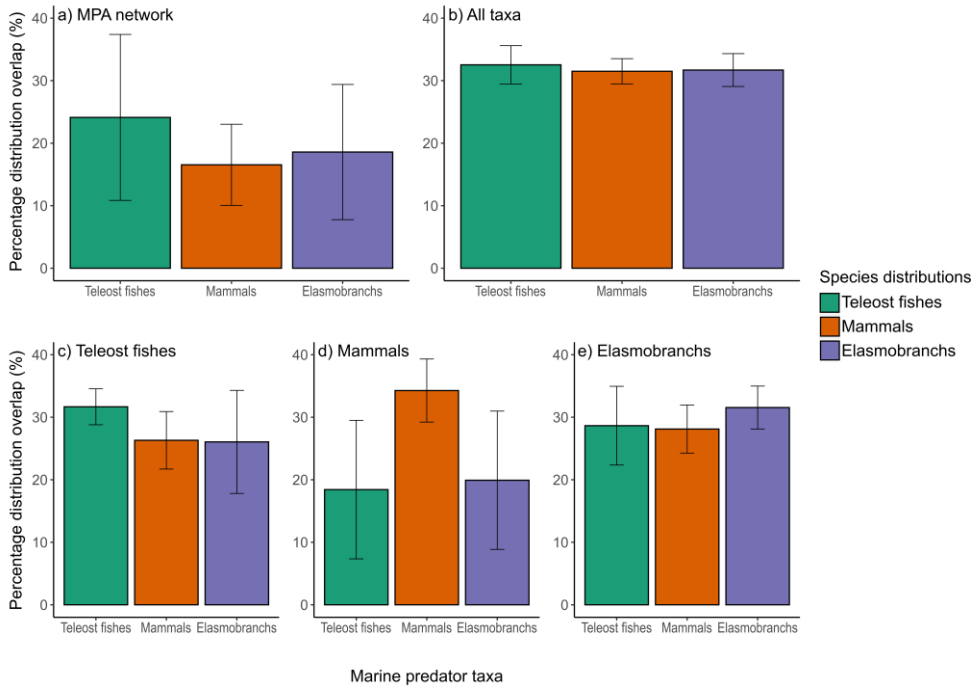


Figure 5.5. Average percentage distribution overlap of fish, mammal and elasmobranch taxa with a) the current network of Mediterranean MPAs, or conservation feature scenarios with b) all taxa, c) teleost fishes, d) mammals or e) elasmobranchs. Standard error bars shown.

5.4. Discussion

We discover that existing MPA systems in the Mediterranean Sea only afford limited protection to marine megafauna, with highly variable coverage within and between different taxonomic groups. Instead, marine megafauna taxa require different conservation priority areas, due to their specific habitat

requirements and life histories, as indicated by their heterogeneous distributions. Focussing spatial planning on all species simultaneously, as is common practice, captured the conservation needs well for elasmobranchs, but excluded sites that would gain highest conservation benefits for marine mammals and teleost fishes. Hence, where spatial planning aims to capture all (*i.e.*, as many as possible) taxa in conservation management areas, there is a risk of missing conservation needs of important taxa. We advocate that including conservation objectives and actions specific to taxonomic groups will better achieve effective conservation of taxa with contrasting or specialised life histories and habitat needs.

We find a striking contrast between the coastal distributions of sharks and teleost fishes, with the offshore ranges of marine mammals, consistent with globally observed patterns in predatory megafaunal richness (Grady et al., 2019) (Figure 5.2). These differences may be exaggerated by greater data availability offshore for marine mammals from ferry-based visual surveys, whilst data for elasmobranchs and teleosts largely comes from coastal fisheries (Mancusi et al., 2020; Mannocci et al., 2018). We also find higher megafauna richness in the north-western basin, due to its proximity to the Strait of Gibraltar which is an important migration corridor connecting the Mediterranean Sea with the Atlantic Ocean (Coll et al., 2010). This result may be inflated by higher observation effort in the north-western basin (Coll et al., 2010). Our SSDM was good at correctly predicting presences but had a high false positive rate which could result in larger areas being protected than necessary (Appendix Table A4.3). Further, marine megafauna are often migratory with seasonal breeding and foraging grounds (Lascelles et al., 2014), but we had inadequate data to consider seasonal variations in distributions for most species. Increasing research efforts to reduce biases in data, and thus ensuring the most effective conservation actions should remain a priority. For example, incorporating seasonal distributions could allow for a dynamic MPA approach where management measures are implemented seasonally to correspond with marine megafauna presence (Maxwell et al., 2020). Similarly, the teleost group includes species with highly divergent life histories, for example, pelagic, migratory species such as tunas and swordfish, as well has more benthic species or reef-associated species with often smaller

spatial ranges. Therefore, the teleost group could benefit from being further subcategorised to permit more targeted management measures, but adequate data only exists for few pelagic teleosts presently. For example, gillnets or beach seines capture more pelagic fish species so restricting these specific fishing gears will have a greater conservation benefit for pelagic species (Mbaru et al., 2020). However, spatial planning should be viewed as a continuous process whereby priority areas are designated based on the best available data, especially given the threatened statuses of Mediterranean marine megafauna, and then iteratively updated and adapted as better data become available (Smith et al., 2009).

Contrasting distributions of marine megafauna taxa translated into significant differences in the spatial arrangement of priority areas (Figure 5.3), highlighting that omitting any of these taxonomic groups from conservation planning will locate MPAs in the wrong areas. Prioritisation solutions including all taxa and elasmobranchs were most similar, driven by elasmobranch species occupying species-poor habitats, or costlier areas which are avoided in the teleost fish or mammal solutions (Kujala et al., 2018). Including representative species across different taxa granted protection to rare species within the taxonomic groups considered. For example, currently recognised important habitats for critically endangered angel sharks were covered in the elasmobranch priority areas despite the species not being included in the current analysis (Giovos et al., 2022). Most importantly, encompassing all taxa simultaneously provides no information about which species or taxa are covered by which priority areas, making it difficult to implement targeted management measures. The requirement for taxa-specific conservation actions to be incorporated into spatial planning has been acknowledged through 'Important Marine Mammal Areas' (IMMAs) and 'Ecologically and Biologically Significant Marine Areas' (EBSAs) (Corrigan et al., 2014). However, obligations to act in response to IMMAs/EBSAs are unclear, and it is debatable how they will specifically contribute to area-based conservation (Corrigan et al., 2014). Specific objectives for separate taxa will necessitate less severe restrictions, *i.e.*, banning of fishing gear which affects the target taxon, instead of a complete fishing ban (Tixier et al., 2021). As Marxan only incorporates a single cost

layer, it was not possible to consider different susceptibilities of marine megafauna to threats, but some threats are likely to be more important for certain species, *i.e.*, time spent at the surface will affect the likelihood of vessel strikes being a large threat (Schoeman et al., 2020).

Current Mediterranean MPAs fail to achieve the 30% coverage target for any of the marine megafauna taxa (Figure 5.5). Teleost fish distributions overlap most with Mediterranean MPAs, despite the two largest MPAs, 'the Pelagos Sanctuary for Marine Mammals' and 'El Corredor de Migración de Cetaceos del Mediterraneo', being designated for cetaceans (MedPAN & SPA/RAC, 2022) (Figure 5.1B). Instead, marine mammal distributions overlapped the least with Mediterranean MPAs, showing misalignment between conservation objectives and outcomes. Currently, the majority of MPAs in the Mediterranean Sea are within European Union waters (Claudet et al., 2020), but our prioritisation solutions highlight important areas for marine megafauna in the southern Mediterranean Sea, including the Alboran Sea, and along the Tunisian and Egyptian coastlines (Figure 5.3). The Alboran Sea has been classified as an IMMA with important habitat for threatened cetaceans and overall high cetacean diversity (IUCN-MMPATF)(Appendix Figure A5.2). The Tunisian and Egyptian coastlines are both included in EBSAs, and encompass feeding and spawning grounds for fin whales (*Balaenoptera physalus*) and bluefin tuna (*Thunnus thynnus*) respectively (UNEP/CBD/EBSA/WS/2014/3/4, 2014). Since priority areas are not shared equally across countries, cross-country collaborations are required to support those with the highest burden, which will be challenging in the dynamic political environment of the Mediterranean (Mazor et al., 2013). However, this is the most cost-effective method to prioritise key habitats for marine megafauna, and will improve the likelihood of successful compliance given that stakeholders and conservation features have been considered synergistically (Mazor et al., 2013).

Despite the rapid expansion in the global extent of MPAs, marine megafauna distributions are not being sufficiently protected, leading to increasing proportions of threatened species. They are notoriously difficult to conserve through MPAs as their vast distributions cannot be encompassed completely, resulting in debate over which key habitats or life history stages should be

prioritised in MPA systems. Here, we advocate spatial planning considering marine megafauna taxonomic groups separately to incorporate their heterogenous distributions arising from divergent life history strategies and habitat use. This approach allows key priority areas for individual taxa to be identified, which are otherwise excluded when considering all taxa concurrently. Designating taxa specific MPAs will enhance the development of more guided management actions to meet the requirements of different taxonomic groups.

Chapter 6 - General Discussion

6.1. Research Summary

Marine megafauna are vital for ecosystem functioning and provide important ecosystem services (Hammerschlag et al., 2019). Currently, they include some of the most threatened species and taxa globally, requiring us to urgently reconsider effective conservation measures to ensure their long-term survival (Avila et al., 2018; Dulvy et al., 2021). Spatial management approaches, such as marine protected areas (MPAs), are most commonly employed to conserve marine megafauna, but their effectiveness depends on accurate knowledge of species distributions (MacKeracher et al., 2019; Notarbartolo di Sciara et al., 2016). Therefore, improved understanding of marine megafauna habitat use in relation to environmental and biological drivers is essential to highlight areas of overlap between marine megafauna and key threats, ensure correct placement of MPAs and facilitate dynamic management strategies in response to climate change (Nelms et al., 2021). Environmental DNA (eDNA) is an important new tool which could enhance our understanding of marine megafauna distributions, and support conservation decision making regarding when and where to implement management actions. In this thesis, I demonstrate that eDNA metabarcoding can improve our understanding of seasonality and heterogeneity in the fine-scale habitat use of marine megafauna in relation to key prey species within their foraging grounds, which are frequently the focus of MPAs. I also show that biodiversity surveys utilising commercial vessels can produce unbiased species distributions, and therefore represent promising platforms for upscaling eDNA surveys to increase spatiotemporal coverage in under surveyed areas. Enhanced understanding of species distributions resulting from greater taxonomic and spatiotemporal coverage offered by eDNA approaches can improve the effectiveness of spatial priority area designations within marine spatial planning (MSP) and permit more dynamic management in the future.

6.2. Chapter Overview

Marine megafauna foraging grounds often cover small spatial areas, relative to their overall distributions, making them popular areas for MPA

implementation (Notarbartolo di Sciara et al., 2016). However, our understanding of marine megafauna habitat use within these areas is often limited by a lack of accurate prey data, especially when prey species are not of commercial interest (Sadykova et al., 2020). In Chapters 2 and 3, I demonstrate the utility of eDNA to enhance our understanding of predator-prey dynamics and to elucidate trophic interactions, focusing on marine mammals in a newly implemented MPA in the Moray Firth. In Chapter 2, I captured spatial partitioning between cetacean species and spatiotemporal patterns in the availability and abundance of key prey species, contributing evidence to a limited pool of eDNA studies focused on marine mammal monitoring (Suarez-Bregua et al., 2022). These data also reduce knowledge gaps in our understanding of prey availability within the MPA and provide an important biodiversity characterisation of the Moray Firth marine ecosystem, which will be essential to evaluate the effectiveness of future management measures. In Chapter 3, I highlight that key prey species, such as forage fish and copepods, which are the most abundant and highly connected taxa in Moray Firth food webs, are vulnerable to climate change. Given that the North Sea is warming three times faster than the global average, future changes in the availability of prey species can be expected, with cascading effects throughout the ecosystem (Alvarez-Fernandez et al., 2012; Belkin, 2009). In contrast, these key food web components were not highly connected in co-occurrence networks, and interactions were volatile across spatial and temporal gradients. Therefore, I caution the continued application of this approach to evaluate eDNA derived communities in unknown systems, as important ecosystem components may be overlooked (Goberna & Verdú, 2022).

In Chapter 4, I provide recommendations for selecting networks of commercial vessel routes and point-based sampling strategies to reduce biases associated with spatially constrained sampling frames, allowing the retrieval of high-quality biodiversity data utilising these cost-effective sampling platforms. I demonstrate that at low sample sizes, sampling along ferry routes can detect more species and recover known patterns in community composition better than unrestricted sampling across the entire Mediterranean Sea, therefore representing suitable platforms to scale up the

spatial resolution of eDNA surveys. This approach could be employed using the global shipping network to provide regular ocean-scale surveying with eDNA, increasing understanding of species distributions from areas that are often expensive or inaccessible (Valsecchi et al., 2021).

Given the substantial increases in available eDNA survey data, it is important to consider how these can be effectively incorporated into MSP (Bani et al., 2020). In chapter 5, I show how improved taxonomic coverage, potentially offered by eDNA approaches, can more effectively identify spatial priority areas for different taxonomic groups. Considering all marine megafauna taxa simultaneously, as is common practice, allocated priority areas for elasmobranchs well, but excluded important sites for marine mammals and teleost fishes. However, assessing taxonomic groups separately resulted in very different spatial configurations of priority areas, stemming from heterogenous distributions among the taxonomic groups. Further, I discover that the current MPA system in the Mediterranean Sea fails to provide adequate coverage (< 30%) for any megafauna taxonomic group, with marine mammals being the least protected despite having the largest designated extent, highlighting the urgent need for MPAs in the Mediterranean to be adapted to effectively conserve marine megafauna (MedPAN & SPA/RAC, 2022).

6.3. Biomonitoring with environmental DNA

eDNA approaches have been widely promoted as important tools to improve our ability to survey biodiversity, covering more taxa and greater spatiotemporal scales than previously possible (Deiner et al., 2017). The advantages of eDNA as a surveying tool are highlighted throughout this thesis, including detections at low abundance, studying predator-prey dynamics simultaneously and whole-ecosystem based monitoring with ecological networks. In Chapter 2, despite sample collection within the Scottish MPA only being implemented across a five-month period, I detected a range of threatened (*i.e.,* European eel *Anguilla anguilla*, bluefin tuna *Thunnus thynnus*) and invasive (*i.e.*, pink salmon *Oncorhynchus gorbuscha*) species, and species that are difficult to survey with other conventional methods (*i.e.,* Sowerby's beaked whales). Detections of both invasive and

threatened species at low abundance levels are essential. In the case of invasive species, low abundance indicates recent settlement or few individuals which improves the likelihood of successful eradication (Giakoumi et al., 2019). Meanwhile, highly threatened species generally have small populations, but detections are vital to ensure mitigation against threats, or to detect range expansions as populations recover, as is the case for bluefin tuna in the Moray Firth (Boussarie et al., 2018; Horton et al., 2021).

Marine megafauna are advocated as important indicators of wider ecosystem changes, given that they are generally easier to monitor compared to other ecosystem components (Hazen et al., 2019). However, not all megafauna will make good indicators, and lag times in the responses of megafauna to changes at lower trophic levels could result in delayed management actions (Hazen et al., 2019). In chapters 2 & 3, I demonstrate the capacity of eDNA to instead conduct monitoring across the entire ecosystem, increasing the likelihood of detecting early changes in usually unobserved ecosystem components. Relatively few eDNA metabarcoding studies to date have investigated interactions across trophic levels, but whole community responses to climatic anomalies such as El Niño or human pressure levels have already been detected (DiBattista et al., 2020; Djurhuus et al., 2020). Here, I detected clear seasonal differences in the Moray Firth community, particularly in the abundances of key prey species for minke whales, with sandeels being the dominant species in June and July, whereas clupeids are more abundant from August onwards. These forage fish species are highly connected, important components of food webs, and are also vulnerable to climate change owing to reliance on particular substrates, so may experience future population declines (Wright et al., 2020). Further, warming is expected to delay sandeel spawning which could result in less prey being available when minke whales first arrive at their foraging grounds (Wright et al., 2020). Therefore, continued monitoring of this ecosystem will be essential to detect changes in key forage fish availability and their impacts on North Sea predators.

Whilst eDNA offers many advantages as a biodiversity monitoring tool, some limitations of eDNA have also become apparent throughout this thesis. In Chapters 2 and 3, I conducted eDNA surveys without comparison to other survey techniques, as is becoming more common practice (Hansen et al.,

2018). However, due to time and monetary constraints, only one primer set targeting 18S rRNA was selected to target broader eukaryotic diversity. I successfully detected Copepoda that dominate North Sea zooplankton communities, but failed to detect other orders of Crustacea, including amphipods, carideans, decapods and isopods which are also important prey groups for planktivorous fishes (Raupach et al., 2015). The North Sea ecosystem has been extensively studied (Raupach et al., 2015), but caution should be applied when using eDNA as the sole surveying tool in less well known environments as it is impossible to know which important ecosystem components may be missing. In Chapter 2, I demonstrated that higher eDNA read proportions for coastal bottlenose dolphins were found in nearshore samples, and higher read proportions for minke whales and harbour porpoises offshore, accurately representing their known habitat use (Robinson et al., 2007). However, as these cetacean species can travel great distances over short time periods, it is difficult to establish whether repeated detections are from the same individuals, or new animals entering the area, making abundance estimates extremely challenging. Abundance estimates are essential to make population-level assessments of threats such as bycatch, and to monitor the stability or recovery of populations following the introduction of conservation measures (Hammond et al., 2021; Magera et al., 2013). Here, I traced the availability of different key prey species temporally using the eDNA index described by Kelly et al. (2019), which reflects the relative abundance of individual species. Changes in abundances of the key forage fish species related to known stages of their life histories, although not all abundance peaks represented suitable prey for megafaunal predators. Therefore, alternative methods to confirm age-class and size would need to be consolidated with eDNA data to confirm whether suitable prey are present. Given the advantages and disadvantages of eDNA, I foresee eDNA being most effective as a monitoring tool for marine megafauna as complementary to other technologies. eDNA will be particularly useful to inform where and when megafauna and their prey species are present, and direct more intensive surveying in these areas to gain further population parameters such as abundance and age structure.

6.4. Scaling up environmental DNA surveys

Marine megafauna typically cover vast ranges, but eDNA studies have largely been conducted at small spatial scales, *i.e.,* a few square kilometres, or onboard expensive dedicated research vessels (Closek et al., 2019; Fraija-Fernández et al., 2020; Jerde et al., 2019). Commercial vessels, such as ferries, container, navy or fishing vessels, offer alternative cost-effective platforms for achieving spatiotemporal coverage necessary to cover marine megafauna ranges (Jensen et al., 2023; Maiello et al., 2023; Valsecchi et al., 2021). However, care must be taken when utilising these platforms for sample collection, as sampling is spatially constrained to the set routes covered, which could result in biased species distributions (Tessarolo et al., 2014). In Chapter 4, I demonstrate that high quality biodiversity data can be retrieved from biased sampling frames if environmental variability and species composition are adequately covered, encouraging the increased use of these opportunistic sampling platforms. However, the global density of shipping traffic is heterogenous, with the Northern hemisphere, particularly East Asia, Southeast Asia, Northwest Europe and North America being covered more comprehensively than other areas (Wang & Wang, 2011) (Figure 6.1). These areas do correspond with those experiencing highest cumulative anthropogenic pressures thus permitting surveying in areas where conservation action is most urgent (Halpern et al., 2015).

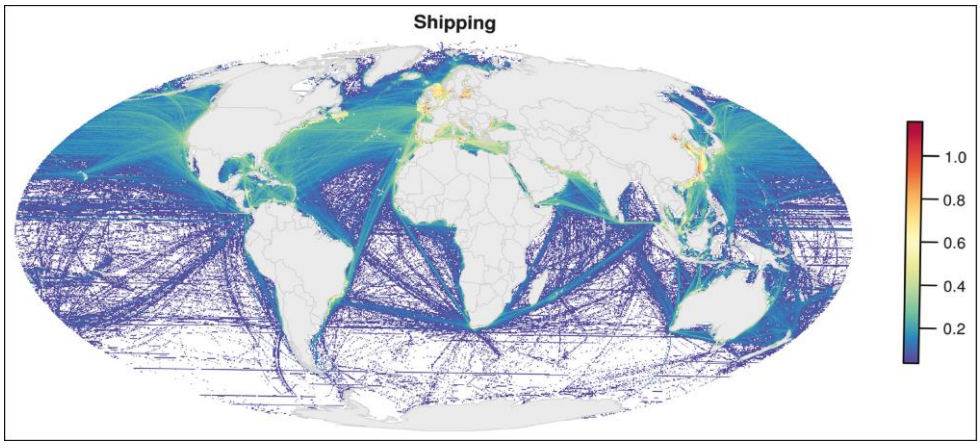


Figure 6.1. A map showing the spatial reach of major global shipping routes, which could expand the spatiotemporal scope of eDNA surveys, especially when combined with automated sampling units. The colour scale indicates change in cumulative human impacts from 2008-2013. From Halpern et al. (2015).

Further, in Chapter 4, I illustrated that ferry routes as sampling platforms could characterize marine megafauna community composition more accurately at small sample sizes than unconstrained sampling across the Mediterranean Sea. This likely stemmed from coincidental overlap between the highest densities of ferry routes and highest species richness concentrating sampling in areas where species detections were more probable. Elsewhere, commercial vessel routes may not cover species distributions as effectively, resulting in missed detections or biased distributions (Figure 6.1). I would therefore recommend some prior knowledge about species distributions to ensure that routes cover desired areas or to implement this strategy in areas with exceptionally good commercial vessel coverage and increased sampling efforts (Boyse et al., 2023). Further, some marine megafauna may actively avoid high density shipping areas which could result in lower species detections than is characteristic of the whole study area of interest (Pirotta et al., 2019). Future work will include implementing the sampling strategies discussed in Chapter 4, to conduct eDNA surveys across the entire Mediterranean Sea. This work was initially meant to form part of this thesis, but sampling was postponed due to the Covid-19 pandemic. The biases and gaps currently present in biodiversity records in the Mediterranean, as highlighted in chapter 4, could be reduced through the increased implementation of eDNA surveys as a result of greater taxonomic and spatial coverage. I predict that Mediterranean-wide eDNA metabarcoding surveys will result in higher predatory megaufana diversity in the eastern basin, which is currently under sampled, and improve fine-scale distributional data for rare species, such as the critically endangered angel shark (Giovos et al., 2022).

6.5. Incorporating eDNA into MPA monitoring and MSP initiatives

Associations of scientists globally have produced best practices guidelines to execute robust and repeatable eDNA studies, with transparent reporting of uncertainties, across government, industry and academia (De Brauwer et al., 2023). These rules have largely been broadscale, *i.e.*, general best practices across the eDNA workflow, or focused on specific environments, *i.e.*, freshwater, and conservation applications, *i.e.*, invasive species monitoring (Abbott et al., 2021; Bruce et al., 2021; De Brauwer et al., 2023; Nicholson et al., 2020). Different applications will require prioritisation of different

workflow elements depending on specific aims, but guidelines are currently limited for MPA biomonitoring and MSP initiatives (Bani et al., 2020). Through advancing understanding of species distributions, eDNA approaches could improve the accuracy of MPA placement as MPA systems continue to expand to fulfil 30% coverage by 2030 (Visalli et al., 2020). In Chapter 5, I demonstrate that well-resolved taxonomic coverage permitted taxonomic groups to be considered separately in MSP resulting in more effective priority area designations for predatory marine megafauna. Often, MSP initiatives rely on coarser ecological data, such as habitat type or bioregions, to generate spatial conservation priority areas, but the comprehensive taxonomic coverage offered by eDNA metabarcoding could permit the inclusion of more high-resolution species data (Astudillo-Scalia et al., 2021). Inclusion of eDNA data in MSP decision making tools will require the transformation of point-based data into spatial data, through methods such as predictive distribution modelling (Bani et al., 2020). However, the spatial bounds of a point-based eDNA sample in marine environments is currently unclear, and will change in different areas depending on factors influencing the persistence of eDNA and transport in currents and tides, with DNA potentially travelling 10s of kilometres and thus increasing the spatial bounds of a sampling point (Andruszkiewicz et al., 2019; Harrison et al., 2019).

As well as improving MPA placement, eDNA approaches also offer the potential to increase biomonitoring in MPAs to evaluate the effectiveness of conservation actions (Bani et al., 2020). Data limitations usually impose the use of static maps in MPA designation, but these are unlikely to be the most effective way to protect highly mobile marine megafauna, especially as they respond to climate-induced prey shifts. Instead, dynamic ocean management (DOM) has been advocated, but this requires the use of near real-time, fine resolution spatiotemporal data of species to respond to changes or to forecast future responses (Cashion et al., 2020; Crespo et al., 2020). In chapters 2 and 3, I show temporal changes in the Moray Firth ecosystem are driven by changes in the abundances of dominant forage fishes and copepods, which are the most highly connected taxa in food webs and particularly vulnerable to climate change. The North Sea is warming three times faster than the global average, signifying likely changes in key ecosystem components, with

rippling effects across the ecosystem, therefore representing a good candidate for implementing DOM (Alvarez-Fernandez et al., 2012; Belkin, 2009). Future work to incorporate eDNA as a biomonitoring tool for DOM will require standardisation across monitoring in different MPAs. For example, Barco et al. (2022) only found 21 fish species in the North Sea Doggerbank MPA and 24 in the Sylt Outer Reef MPA compared to 68 fish species and families detected in Southern Trench MPA in Chapter 2. These differences could stem from true differences in species richness, or methodological differences in water volume and primer sets chosen (Barco et al., 2022). Extending eDNA surveys outside of MPAs will improve evaluation of MPA protection measures and detect potential species range shifts to areas outside of protection. For example, Gold et al. (2021) and Boulanger et al. (2021) both found lower species richness in MPAs compared to unprotected areas, with eDNA metabarcoding. Boulanger et al. (2021) suggested that greater numbers of high trophic level species found in the MPA had driven lower species richness through trophic cascades. However, Gold et al. (2021) cautions that when MPAs cover specific habitats such as kelp beds, resident fishes congregate and dominate sequences retrieved, potentially reducing the likelihood of detecting rarer species and hence lowering species diversity indexes, compared to other habitats made up of more transient fish species, with generally lower detection probabilities. Therefore, validating eDNA diversity indexes with other techniques such as visual SCUBA surveys are recommended (Gold et al., 2021).

Delineating ecological networks may allow protection coverage targets to be set for different networks in MSP, moving away from species centric approaches. Ecological networks can also identify important ecosystem components to protect, *i.e.*, highly connected species within a network, or species which connect multiple networks together, as well as important interactions (Berry & Widder, 2014). Similarly, MPA conservation actions can be evaluated for whole ecosystems, instead of measuring MPA success based on species metrics, *i.e.*, changes in population sizes (DiBattista et al., 2020). In Chapter 3, food webs constructed from eDNA metabarcoding and known diets distinguished spatial and temporal variation in highly connected nodes, some of which are particularly vulnerable to climate change, such as

copepods and forage fishes. More complex food webs can be used simulate how food web components respond to human or climate pressures, or MPA protection levels, but require accurate biomass estimates (Dahood et al., 2020; Heymans et al., 2016). Biomass estimates from eDNA are contentious as PCR amplification biases can distort the starting composition of sequences per species, but methods to quantify these biases are increasingly improving, so future spatiotemporal simulations of food webs derived from eDNA may be possible (Shelton et al., 2023). The food webs in Chapter 3 were constructed using diet information derived from literature, which is less well resolved for lower taxonomic levels, and may be impossible for less well studied ecosystems. Therefore, I explored co-occurrence networks as an alternative method to detect spatiotemporal differences in the Moray Firth community. I found that methodological choices had a large impact on what interactions were detected and discovered that highly connected nodes included species only when they were least abundant. This could have negative impacts if used to identify important ecosystem components to protect, for example, bottlenose dolphins are highly connected in the offshore co-occurrence network, but detections of dolphins here likely stem from eDNA dispersal which could result in protection measures being incorrectly placed. Therefore, I caution the use of co-occurrence networks without further validation in well-known ecosystems.

6.6. Future Directions

eDNA has the potential to revolutionise how we survey biodiversity, particularly through offering more comprehensive taxonomic coverage and expanding the spatial and temporal range of sampling beyond what was previously possible (Pascher et al., 2022). The non-invasiveness and high sensitivity of the technique are particularly beneficial for continued monitoring of rare, threatened species or to detect early signs of invasive species in new habitats. However, for eDNA to be operationally ready as a biodiversity survey tool on a global scale, there are still multiple important uncertainties that need to be addressed. Firstly, identifying which species a DNA sequence originates from is contingent on adequate and accurate coverage in online sequence data repositories. However, even for comparatively well-studied European fish species, the largest reference

database, GenBank, only covers 50% of species for the preferred target regions, 12S or 16S rRNA (Claver et al., 2023). Multiple countries have launched national initiatives to complete reference libraries for native species or individual eDNA surveys have built custom, localised reference libraries, which will be particularly important in regions with high endemic species richness (Weigand et al., 2019; Yao et al., 2022). Similarly, the spatial and temporal resolution of an individual eDNA sample is still unclear, although this will vary substantially depending on the environment, *i.e.*, lotic versus lentic ecosystems (Bani et al., 2020). Improved understanding of environmental and biological factors contributing to eDNA production and degradation, combined with increased use of oceanic models and particle tracking to predict eDNA transport will be vital to refine the spatiotemporal scope of samples (Hansen et al., 2018; Ramírez et al., 2022). Improved understanding in turn will increase confidence in species distributions or abundance described from eDNA surveys (Shelton et al., 2022). The full eDNA workflow requires expertise from multiple different disciplines including ecologists to design sampling strategies, molecular biologists to conduct laboratory work and bioinformaticians to handle and analyse the large sequencing datasets produced (Pascher et al., 2022). The laboratory work and bioinformatics steps are commonly outsourced to external providers, but it is essential that any end-user interpreting eDNA surveys understands the decisions made during these stages as they can have significant impacts on the results achieved (Pascher et al., 2022; Rodriguez-Ezpeleta et al., 2021).

Technological advances and novel methods will also play an important role in increasing the applicability of eDNA surveys globally. Currently, eDNA sampling is typically labour intensive, including manual collection of samples followed by filtration and preservation (Yao et al., 2022). This process prohibits eDNA surveys in remote areas or where eDNA samples need to be taken regularly. Passive eDNA sampling is emerging as a cost-effective and rapid method that forgoes filtration by directly immersing filters in water or utilising species that naturally filter water, such as marine sponges (Bessey et al., 2021; Mariani et al., 2019). These methods are still in preliminary stages with appropriate submersion times or filter type remaining ambiguous and

further validation compared with active filtering still required (Takahashi et al., 2023). Autonomous eDNA samplers are now also being developed to collect, filter and store multiple samples which will improve the spatial and temporal scope of eDNA surveys, for example, through long-term deployment of autonomous eDNA samplers in remote areas (Hendricks et al., 2023; Truelove et al., 2022). In the future, hopefully these technologies can be expanded further to include in-situ real-time PCR and sequencing as well, rapidly reducing overall processing time (Yao et al., 2022). However, it is important to find a balance between standardising techniques so that different surveys can be compared, whilst integrating modern technologies and method developments as they occur (Stepien et al., 2023). At the moment, the portable Oxford Nanopore MinIon sequencer permits in-situ sequencing, but higher error rates (6%) compared to the commonly implemented Illumina sequencers currently prevents widescale usage for eDNA surveys which rely on small percentage sequence differences (< 3%) to distinguish between species (Ames et al., 2021; Truelove et al., 2019). Employing automatic eDNA sampling and sequencing onboard commercial vessels could vastly increase our capacity to monitor marine species at ocean wide scales regularly (Truelove et al., 2022).

6.7. Conclusions

It is currently unclear whether MPA coverage and conservation measures implemented are protecting marine megafauna effectively. Accurate species distributions are required to ensure that marine megafauna are effectively protected in MPAs as network coverage expands, and increased long-term monitoring within MPAs is necessary to evaluate the impacts of conservation actions. In this thesis, I demonstrated how eDNA can improve our understanding of marine megafauna habitat use across greater spatiotemporal scales, in relation to key prey species, and to assess how vulnerable the ecosystems within which they exist are to human impacts or climate change. Rapid advances in eDNA approaches are likely to further increase the utility of the method, *i.e.,* with more robust abundance estimates and increased certainty in the spatial bounds of eDNA samples. Ultimately, the increased spatiotemporal and taxonomic coverage offered by eDNA approaches can prioritise spatial conservation areas for marine megafauna more effectively.

Supplementary Appendix

Chapter 2 - Environmental DNA reveals fine scale spatial and temporal variation of prey species for marine mammals in a Scottish marine protected area

Appendix Table A2.1. Metadata for each eDNA sampling point. Samples were named such that the first letter corresponded to the type of sample (C = Control, F = Fixed, S = Sighting), the first number to the monthly trip (1 = June, 2 = July, 3 = August, 4 = September/October) and the second number to the sample site (1 to 3 for fixed samples from west to east, or the chronological order of samples collected for that trip for control and sighing samples.)

Name	Date	Time	Longitude	Latitude	Bathymetry (m)	Sea surface temperature (ºC)	Chlorophyll *a* (mg/m^3)	Distance from shore (m)	Distance from shore category
C1.1	23/06/2021	16:17	-2.38349	57.77308	-134	13.4	0.8069	9,840	Middle far
C1.2	24/06/2021	10:47	-2.55324	57.72608	-72	13.55	0.5582	4,896	Middle near
C1.3	24/06/2021	12:14	-2.3384	57.7856	-99	13.81	0.8025	10,188	Far
C1.4	24/06/2021	12:39	-2.32011	57.73321	-70	13.12	1.2222	4,278	Middle near
C1.5	27/06/2021	13:22	-2.28656	57.69763	-20	NA	1.63	352	Near
C1.6	30/06/2021	12:23	-2.7032	57.69119	-7	NA	NA	393	Near
C1.7	30/06/2021	14:37	-2.84249	57.73908	-27	14.21	1.2632	3,535	Middle near
C1.8	30/06/2021	16:50	-2.57757	57.68753	-21	13.79	2.5682	458	Near
F1.1	30/06/2021	14:55	-2.82517	57.76298	-33	14.28	1.0743	6,333	Middle near
F1.2	23/06/2021	13:48	-2.54138	57.76318	-104	13.46	0.6049	9,065	Middle far
F1.3	27/06/2021	15:33	-2.26	57.7637	-97	13.01	1.2358	7,765	Middle far
S1.1	23/06/2021	12:56	-2.54576	57.74366	-91	13.36	0.5796	6,889	Middle near
S1.2	23/06/2021	14:25	-2.51157	57.82474	-85	13.74	0.6064	16,144	Far
S1.3	23/06/2021	15:25	-2.38127	57.83258	-91	13.56	0.8228	15,928	Far
S1.4	24/06/2021	11:25	-2.49849	57.79097	-77	14.07	0.6134	12,755	Far

S1.5	24/06/2021	11:48	-2.43766	57.79806	-90	14.07	0.6742	13,794	Far
S1.6	27/06/2021	15:21	-2.21621	57.76389	-91	12.25	1.5723	8,837	Middle far
S1.7	28/06/2021	12:12	-2.55214	57.74915	-87	13.36	0.9526	7,407	Middle far
S1.8	28/06/2021	15:59	-2.50392	57.78123	-82	12.94	1.148	11,625	Far
S1.9	30/06/2021	15:18	-2.69178	57.76373	-69	13.96	1.4653	8,023	Middle far
C2.1	14/07/2021	14:50	-2.82913	57.69946	-5	NA	NA	521	Near
C2.2	14/07/2021	16:08	-2.94661	57.73386	-24	14.93	1.3747	4,212	Middle near
C2.3	14/07/2021	16:25	-2.94428	57.76418	-34	15.02	1.1862	7,370	Middle far
C2.4	15/07/2021	14:46	-2.2303	57.69117	-19	14.32	9.4317	1,115	Near
C2.5	19/07/2021	13:47	-2.26489	57.72487	-7	15.79	1.733	721	Near
C2.6	19/07/2021	14:44	-3.26542	57.78394	-54	14.67	0.9296	6,571	Middle near
C2.7	19/07/2021	16:25	-2.84649	57.83421	-82	14.82	1.4408	14,112	Far
C2.8	19/07/2021	16:50	-2.65892	57.83317	-85	14.26	1.9699	15,871	Far
F2.1	14/07/2021	17:08	-2.82509	57.76437	-33	14.98	1.0744	6,486	Middle near
F2.2	12/07/2021	15:20	-2.54151	57.76329	-104	13.49	1.0481	9,075	Middle far
F2.3	15/07/2021	16:13	-2.25834	57.76334	-118	14.39	1.562	7,756	Middle far
S2.1	12/07/2021	14:44	-2.5535	57.71339	-50	13.25	1.0458	3,521	Middle near
S2.2	14/07/2021	10:57	-2.56572	57.72279	-64	14.54	0.9909	4,383	Middle near
S2.3	19/07/2021	15:32	-3.09144	57.81815	-71	15.23	1.0563	14,999	Far
C3.1	08/08/2021	13:41	-2.50945	57.67346	-8	NA	0.5912	488	Near
C3.2	10/08/2021	17:13	-2.75408	57.76392	-52	15.47	0.6513	7,301	Middle far
C3.3	10/08/2021	18:49	-2.5058	57.69665	-33	14.99	0.7197	2,554	Middle near
C3.4	11/08/2021	12:16	-2.4759	57.77603	-97	15.27	2.0819	11,334	Far
C3.5	15/08/2021	10:04	-2.81838	57.7	-12	NA	NA	468	Near

C3.6	15/08/2021	13:08	-2.64304	57.68734	-5	NA	NA	236	Near
C3.7	15/08/2021	16:31	-2.55517	57.67698	-5	NA	1.2377	373	Near
F3.1	10/08/2021	16:57	-2.82418	57.76188	-32	15.7	0.5867	6,228	Middle near
F3.2	10/08/2021	18:18	-2.54113	57.76343	-104	15.84	0.6061	9,095	Middle far
F3.3	11/08/2021	13:03	-2.25863	57.76345	-118	14.9	1.9636	7,762	Middle far
S3.1	10/08/2021	17:56	-2.57948	57.76481	-88	15.86	0.6755	9,040	Middle far
S3.2	15/08/2021	10:23	-2.86333	57.70873	-16	NA	1.4816	305	Near
C4.1	25/09/2021	13:51	-2.63648	57.76547	-82	13.39	1.5558	8,508	Middle far
C4.2	28/09/2021	11:48	-2.28578	57.69962	-20	NA	4.7475	568	Near
C4.3	28/09/2021	12:34	-2.16478	57.72858	-56	13.62	4.5537	4,042	Middle near
C4.4	28/09/2021	13:08	-2.34827	57.72362	-53	13.69	4.2259	3,958	Middle near
C4.5	28/09/2021	16:46	-2.77605	57.70878	-25	13.79	4.0918	1,109	Near
F4.1	25/09/2021	13:17	-2.82406	57.76314	-33	13.53	1.9329	6,366	Middle near
F4.2	25/09/2021	14:22	-2.54144	57.76383	-104	13.17	1.4	9,135	Middle far
F4.3	28/09/2021	12:08	-2.26	57.76317	-118	13.34	2.0601	7,709	Middle far
S4.1	04/10/2021	14:41	-2.2043	57.7405	-97	13.08	2.4202	6,579	Middle near
S4.2	04/10/2021	15:40	-2.2327	57.7788	-108	13.03	1.4482	9,898	Middle far
S4.3	08/10/2021	15:46	-2.1808	57.7623	-49	13.03	0.9055	7,735	Middle far

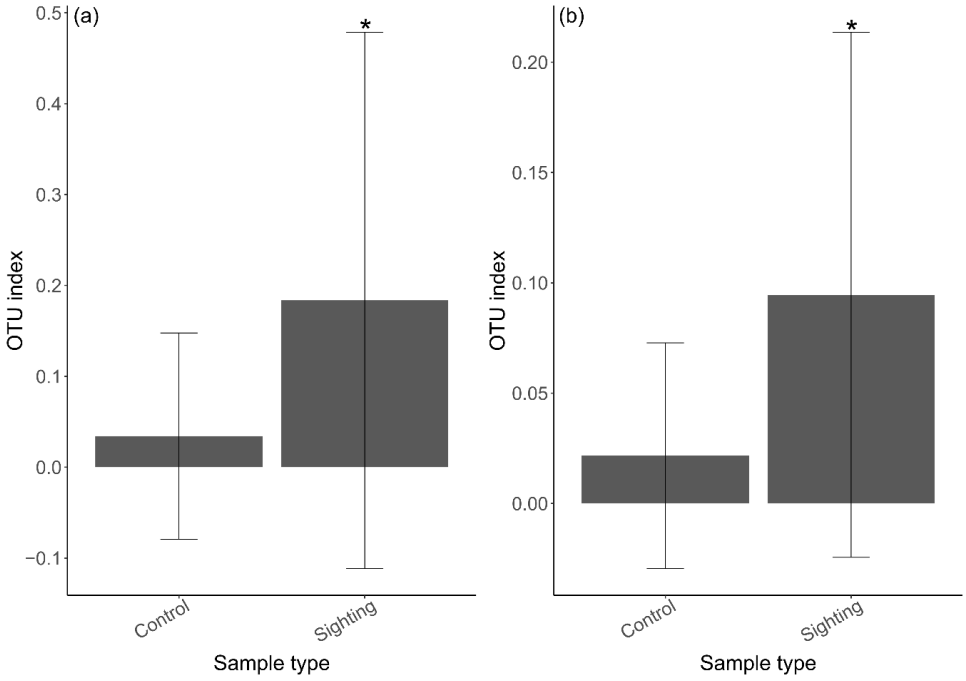
Appendix Table A2.2. Foreign OTUs detected with MarVer1 were compared to native relatives from the same genus or family. Reads were either reassigned to the native relative, removed due to low abundance or identified as a potential invasive.

Foreign OTU	North Sea relative (shared genus or family)	Action
Alectrias benjamini	*Chirolophis ascanii*	Reassign to native relative
Ammodytes dubius *Ammodytes hexapterus* *Ammodytes personatus*	*Ammodytes marinus* *Ammodytes tobianus*	All reads combined as *Ammodytes* genus
Anguilla japonica	*Anguilla anguilla*	Remove as singleton
Chelidonichthys spinosus	*Eutriglia gurnardus* *Trigloporus lastoviza*	All reads combined as Triglidae family
Clupea pallasii	*Clupea harengus*	Reassign to native relative
Diplodus puntazzo		No congeneric in North Sea so potential invasive
Gadus chalcogrammus	*Gadus morhua* *Melanogrammus aeglefinus* *Merlangius merlangus* *Pollachius pollachius* *Trisopterus minutus*	All reads combined as Gadidae family
Hippoglossus stenolepsis	*Hippoglossoides platessoides*	Reassign to native relative
Larus dominicanus	*Larus argentatus*	Reassign to native relative
Myoxocephalus jaok	*Myoxocephalus scorpius*	Reassign to native relative
Oncorhynchus clarkii *Oncorhynchus mykiss* *Oncorhynchus nerka*		Collapsed to *Oncorhynchus* genus, known invasive
Parahucho perryi *Salmo ischchan* *Salmo labrax*	*Salmo salar* *Salmo trutta*	Collapsed to *Salmo* genus
Pholis picta	*Pholis gunnellus*	Reassign to native relative
Sardinella longiceps	*Clupea harengus* *Sprattus sprattus*	Distinct from native family members, potential native
Sousa teuszii	*Tursiops truncatus*	Reassign to native relative
Thunnus maccoyii	*Thunnus thynnus*	Reassign to native relative

Appendix Table A2.3. Foreign OTUs detected with MarVer3 were compared to native relatives from the same genus or family. Reads were either reassigned to the native relative, removed due to low abundance or identified as a potential invasive.

Foreign OTU	North Sea relative (shared genus or family)	Action
Ammodytes americanus *Ammodytes hexapterus* *Ammodytes personatus* *Hyperoplus lanceolatus*	*Ammodytes marinus* *Ammodytes Tobianus*	Collapse to ammodytidae family
Anarhichas orientalis	*Anarhichas lupus*	Remove as <10 reads
Anguilla rostrata	*Anguilla Anguilla*	Reassign to native relative
Arctogadus glacialis *Boreogadus saida*	*Gadus morhua* *Melanogrammus aeglefinus* *Merlangius merlangus* *Pollachius pollachius* *Pollachius virens* *Trisopterus esmarkii* *Trisopterus minutus*	Reassign to native relative, *Gadus morhua*
Clidoderma asperrimum		Remove as singleton
Clupea pallasii	*Clupea harengus*	Reassign to native relative
Dipturus innominatus *Dipturus trachyderma*	*Dipturus batis* *Dipturus oxyrinchus*	Collapse to *Dipturus* genus
Eleginus gracilis *Gadus chalcogrammus* *Gadus macrocephalus*	*Gadus morhua* *Melanogrammus aeglefinus* *Merlangius merlangus* *Pollachius pollachius* *Pollachius virens* *Trisopterus esmarkii* *Trisopterus minutus*	Remove as <25 reads Remove as <3 reads Remove as <3 reads
Glyptocephalus zachirus *Hippoglossoides dubius* *Hippoglossoides elassodon* *Hippoglossoides stenolepsis* *Isoptetta isolepsis* *Lepidopsetta bilineata* *Lepidopsetta mochigarei* *Liopsetta pinnifasciata* *Parophyrs vetulus* *Platichthys stellatus* *Prosopium cylindraceum* *Psettichthys melanostictus* *Pseudopleuronectes americanus*	*Glytocephalus cynoglossus* *Hippoglossoides platessoides* *Limanda limanda* *Microstomus kitt* *Pleuronectes flesus* *Pleuronectes platessa*	Collapse all foreign species and *G. cynoglossus, H. platessoides, P. flesus* and *P. platella* into Pleuronectidae family. Keep *L. limanda* and *M. kitt* separate

Pseudopleuronectes yokohamae		
Limanda aspera *Limanda proboscidea* *Limanda sakhalinensis*	*Limanda limanda*	Reassign to native relative
Microgadus proximus	*Gadus morhua*	Reassign to native relative
Myoxocephalus brandtii *Myoxocephalus ochotensis* *Myoxocephalus polyacanthocephalus*	*Myoxocephalus scorpius*	Reassign to native relative
Oncorhynchus clarkii *Oncorhynchus gilae* *Oncorchynchus gorbuscha* *Oncorhynchus mykiss*		Collapse *O. clarkii* and *O. gilae* to *Oncorhynchus* genus but keep *O. gorbuscha* and *O. mykiss* separate. Known invasives.
Phocoena sinus	*Phocoena phocoena*	Remove as singleton
Pholis laeta	*Pholis gunnellus*	Reassign to native relative
Pungitius sinensis	*Pungitius pungitius*	Reassign to native relative
Pusa hispida	*Halichoerus grypus*	Reassign to native relative
Salmo ischchan *Salmo obtusirostris*	*Salmo salar* *Salmo trutta*	Reassign to native relative, *S. trutta*
Symphodus melamocercus	*Symphodus melops*	Remove as singleton
Tursiop aduncus	*Tursiop truncatus*	Reassign to native relative



Appendix Figure A2.1. OTU index for minke whales in control versus sighting samples for (a) MarVer1 and (b) MarVer3. Minke whales were detected in 53 % of sighting samples and 18 % of control samples with MarVer 1 whilst they were detected in 88 % of sighting samples compared to 76 % of control samples with MarVer3. Sighting samples had significantly higher abundances of minke whales compared to control samples for both primer sets (Wilcoxen test, $p < 0.05$).

Appendix Table A2.4. Indicator species analysis for vertebrates per trip

Trip	OTU	A	B	Stat	P value
1	*Thunnus thynnus*	0.9946	0.3684	0.605	0.025
2	*Mesoploden bidens*	0.8244	0.8571	0.841	0.015
2	*Leuciscus*	0.9835	0.5714	0.75	0.015
2	*Spinachia spinachia*	0.9751	0.5714	0.746	0.02
2	*Diplodus puntazzo*	1	0.4286	0.655	0.005
2	*Fulmarus glacialis*	1	0.2143	0.463	0.025
3	*Oncorhynchus*	0.9664	0.9167	0.941	0.005
3	*Anguilla anguilla*	0.7431	0.75	0.747	0.005
3	*Lesuerigobius friesii*	0.9983	0.4167	0.645	0.015
3	*Mola mola*	1	0.3333	0.577	0.005
3	*Pomatoschistus pictus*	0.9993	0.3333	0.577	0.02
4	*Micrenophrys lilljeborgii*	0.9947	0.6364	0.796	0.01
4	*Mullus barbatus*	0.9924	0.6364	0.795	0.005
4	*Pomatoschistus microps*	0.9838	0.6364	0.791	0.005
4	*Gobiusculus flavescens*	0.9519	0.2727	0.51	0.02
4	*Lophius piscatorius*	1	0.1818	0.426	0.02
1+2	*Balaenoptera acutorostrata*	0.9328	0.8788	0.905	0.025
1+2	*Ciliata septentrionalis*	0.9167	0.7879	0.85	0.005
2+3	*Leucoraja naevus*	0.974	0.7692	0.866	0.025
2+3	*Pleuronectidae*	0.8742	0.8462	0.86	0.005
2+3	*Centrolabrus exoletus*	0.9824	0.7308	0.847	0.005
2+3	*Taurulus bubalis*	0.8242	0.6923	0.755	0.025
2+3	*Dicentrarchus*	0.994	0.5385	0.732	0.01
2+3	*Liparis montagui*	0.9989	0.5	0.707	0.005
2+3	*Zoarces*	0.9982	0.4615	0.679	0.03
2+3	*Chirolophis ascanii*	0.9996	0.3846	0.62	0.02
2+4	*Scyliorhinus canicula*	0.9587	0.68	0.807	0.015
3+4	*Salmo*	0.9518	0.8261	0.887	0.005
3+4	*Trisopterus minutus*	0.8445	0.4783	0.636	0.01
1+2+3	*Gymnammodytes semisquamatus*	0.996	0.9556	0.76	0.005
1+2+3	*Phocoena phocoena*	0.9999	0.8667	0.931	0.005
1+2+3	Ammodytidae	0.9764	0.8222	0.896	0.005
1+2+3	*Uria aalge*	0.9895	0.7778	0.877	0.03
1+2+3	*Ctenolabrus rupestris*	1	0.7556	0.869	0.005
1+2+4	*Ciliata mustela*	0.9416	0.8409	0.89	0.045
2+3+4	*Phoxinus phoxinus*	0.9998	0.6216	0.788	0.005

Appendix Table A2.5. Indicator species analysis for eukaryotes per trip

Trip	OTU	A	B	Stat	P value
1	Tintinnidae	1.00	0.25	0.5	0.018
2	Attheyaceae	0.9465	0.9286	0.938	0.001
2	Rhizophlyctidaceae	0.9796	0.7143	0.837	0.001
2	Metacyclididae	0.7385	0.9286	0.828	0.002
2	Codonellopsidae	0.9420	0.7143	0.820	0.001
2	Campanulariidae	1	0.3571	0.598	0.002
2	Harpacticidae	1	0.3571	0.598	0.002
2	Canthocamptidae	1	0.2143	0.463	0.036
3	Pandeidae	0.9671	0.6667	0.803	0.001
3	Bougainvilliidae	0.9797	0.5833	0.756	0.001
3	Trochamminidae	0.9263	0.5833	0.735	0.001
3	Cylichnidae	0.9601	0.5	0.693	0.002
3	Undellidae	0.7876	0.5833	0.678	0.002
3	Cladocorynidae	0.9640	0.3333	0.567	0.008
3	Dotoidae	0.9468	0.3333	0.562	0.02
3	Molgulidae	0.8521	0.3333	0.533	0.016
3	Corycaeidae	1	0.25	0.5	0.01
3	Dictyotaceae	1	0.25	0.5	0.013
4	Naviculaceae	0.911	1	0.954	0.001
4	Phaeocystaceae	0.9396	0.9091	0.924	0.001
4	Rhaphoneidaceae	0.9914	0.5455	0.735	0.001
4	Veneridae	0.7542	0.3636	0.524	0.025
4	Corethraceae	0.7495	0.3636	0.522	0.023
4	Euchaetidae	0.9885	0.2727	0.519	0.021
4	Eucalanidae	0.9804	0.1818	0.422	0.047
1&2	Didiniidae	0.9778	1	0.989	0.001
1&2	Noctilucaceae	0.9989	0.9706	0.985	0.001
1&2	Spathidiidae	1	0.5882	0.767	0.001
1&2	Xystonellidae	1	0.4706	0.686	0.002
1&3	Bolinopsidae	0.9274	0.9062	0.917	0.001
2&3	Calciodinellaceae	0.8654	1	0.930	0.001
2&3	Pyrocystaceae	1	0.4615	0.679	0.001
2&3	Halosphaeraceae	0.9906	0.4615	0.676	0.002
3&4	Stephanodiscaceae	0.9308	1	0.965	0.001
3&4	Syndiniaceae	0.9719	0.8261	0.896	0.001
3&4	Calanidae	0.9279	0.7826	0.852	0.001
3&4	Lauderiaceae	0.9941	0.4348	0.657	0.002
3&4	Paradiniaceae	0.89	0.3478	0.556	0.05
1&2&3	Prymnesiaceae	0.9978	1	0.999	0.001
1&2&3	Pedinellaceae	0.9921	1	0.996	0.001
1&2&3	Dinophysiaceae	0.9879	1	0.994	0.001
1&2&3	Warnowiaceae	0.9839	1	0.992	0.001
1&2&3	Goniodomataceae	0.9999	0.9565	0.978	0.001
1&2&3	Gonyaulacaceae	0.999	0.913	0.955	0.001
1&2&3	Coccolithaceae	0.9937	0.8913	0.941	0.001
1&2&4	Strombidiidae	0.9959	0.9333	0.964	0.001
1&2&4	Hemiaulaceae	0.9923	0.7556	0.866	0.002
1&2&4	Nephroselmidaceae	0.9844	0.6889	0.823	0.001
1&3&4	Cymatosiraceae	0.9846	0.9535	0.969	0.002
1&3&4	Hemiaulaceae	0.9475	0.9535	0.951	0.001
1&3&4	Euplokamididae	0.9885	0.6279	0.788	0.004
1&3&4	Bodinidae	0.9653	0.5349	0.719	0.01
2&3&4	Fragilariaceae	0.9778	0.5405	0.727	0.017
2&3&4	Ectocaraceae	0.9738	0.5135	0.707	0.03

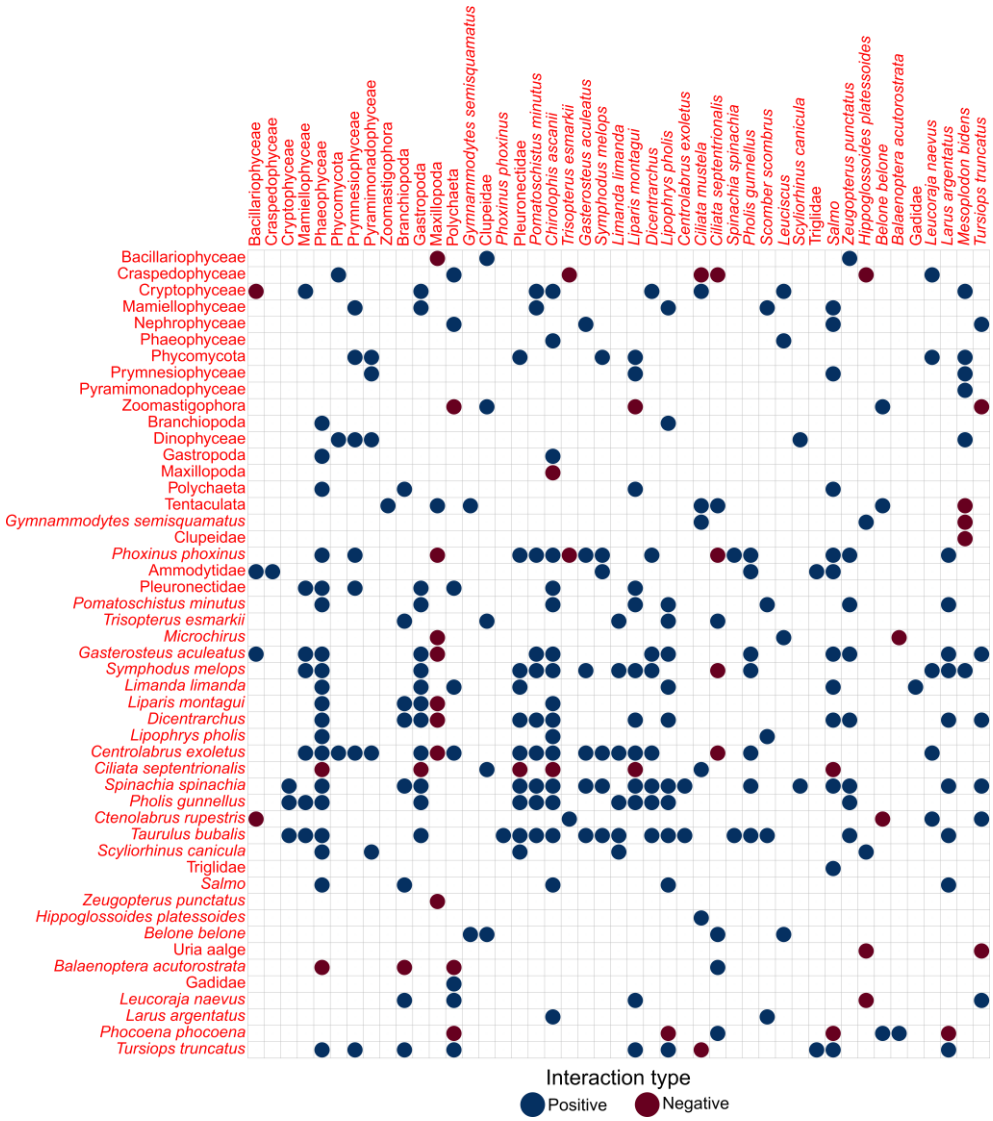
Appendix Table A2.6. Indicator species analysis for eukaryotes for distance from shore categories.

Trip	OTU	A	B	Stat	P value
Far	Harpacticidae	0.7433	0.3	0.472	0.035
Near	Ectocarpaceae	0.80293	1	0.896	0.005
Near	Paradiniaceae	0.98717	0.76923	0.871	0.005
Near	Ralfsiaceae	0.95006	0.76923	0.855	0.005
Near	Rhodomelaceae	1	0.69231	0.832	0.005
Near	Dasyaceae	0.99224	0.69231	0.829	0.005
Near	Ulvaceae	0.95142	0.61538	0.765	0.005
Near	Leucosoleniidae	1	0.53846	0.734	0.005
Near	Molgulidae	1	0.46154	0.679	0.001
Near	Chaetopteridae	0.74645	0.61538	0.678	0.01
Near	Siphonariidae	0.88262	0.46154	0.638	0.015
Near	Catenulaceae	0.58006	0.69231	0.634	0.015
Near	Ophiactidae	0.86935	0.46154	0.633	0.005
Near	Lauderiaceae	0.84958	0.46154	0.626	0.02
Near	Membraniporidae	1	0.38462	0.62	0.01
Near	Pyrochystaceae	0.74297	0.46154	0.586	0.01
Near	Archinotodelphyidae	1	0.30769	0.555	0.015
Near	Scystosiphonaceae	0.96654	0.30769	0.545	0.015
Near	Myrionemataceae	0.87352	0.30769	0.518	0.025
Near	Dictyotaceae	1	0.23077	0.48	0.035
Near	Pectinariidae	0.98235	0.23077	0.476	0.025
Far + Middle far	Spathidiidae	0.8088	0.5357	0.658	0.02
Far+middle near+near	Fragilariaceae	0.977	0.5641	0.742	0.02
Middle far+middle near+near	Stephanodiscaceae	0.9977	0.766	0.874	0.005
Middle far+middle near+near	Phaeocystaceae	0.9929	0.7021	0.835	045

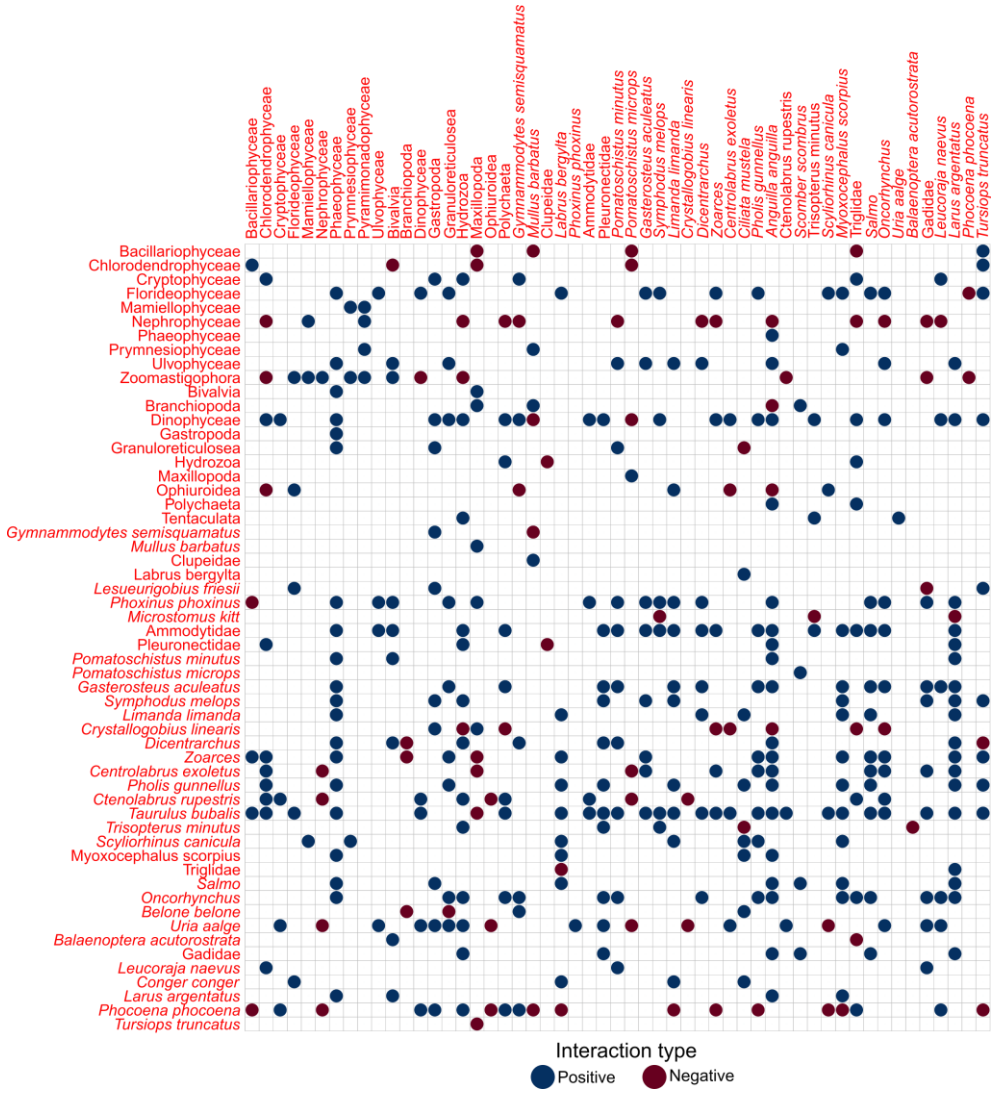
Appendix Table A2.7. Indicator species analysis for vertebrates for distance from shore categories.

Trip	OTU	A	B	Stat	P value
Far	*Hippoglossoides platessoides*	0.8783	0.5	0.663	0.005
Near	*Pholis gunnellus*	0.8521	0.9231	0.887	0.005
Near	*Taurulus bubalis*	0.9825	0.6923	0.825	0.005
Near	*Labrus bergylta*	0.9984	0.6154	0.784	0.005
Near	*Zoarces*	0.9926	0.5385	0.731	0.005
Near	*Liparis montagui*	0.9913	0.5385	0.731	0.01
Near	*Myoxocephalus scorpius*	0.9589	0.4615	0.665	0.005
Near	*Mola mola*	1	0.3077	0.555	0.005
Near	*Pomatoschistus pictus*	0.9995	0.3077	0.555	0.015
Near	*Gaidropsarus mediterraneus*	1	0.2308	0.48	0.025
Near	*Thorogobius ephippiatus*	1	0.2308	0.48	0.015
Far+near	*Limanda limanda*	0.9326	0.8261	0.878	0.02
Middle near + near	*Tursiops truncatus*	0.9961	0.8214	0.905	0.005
Middle near +near	*Centrolabrus exoletus*	0.9843	0.5357	0.726	0.045
Far+middle far+middle near	*Phocoena phocoena*	0.9759	0.8372	0.904	0.005
Far+middle far+near	Clupeidae	0.9396	0.878	0.908	0.03
Far+middle near+near	*Gasterosteus aculeatus*	0.9654	0.8421	0.902	0.03

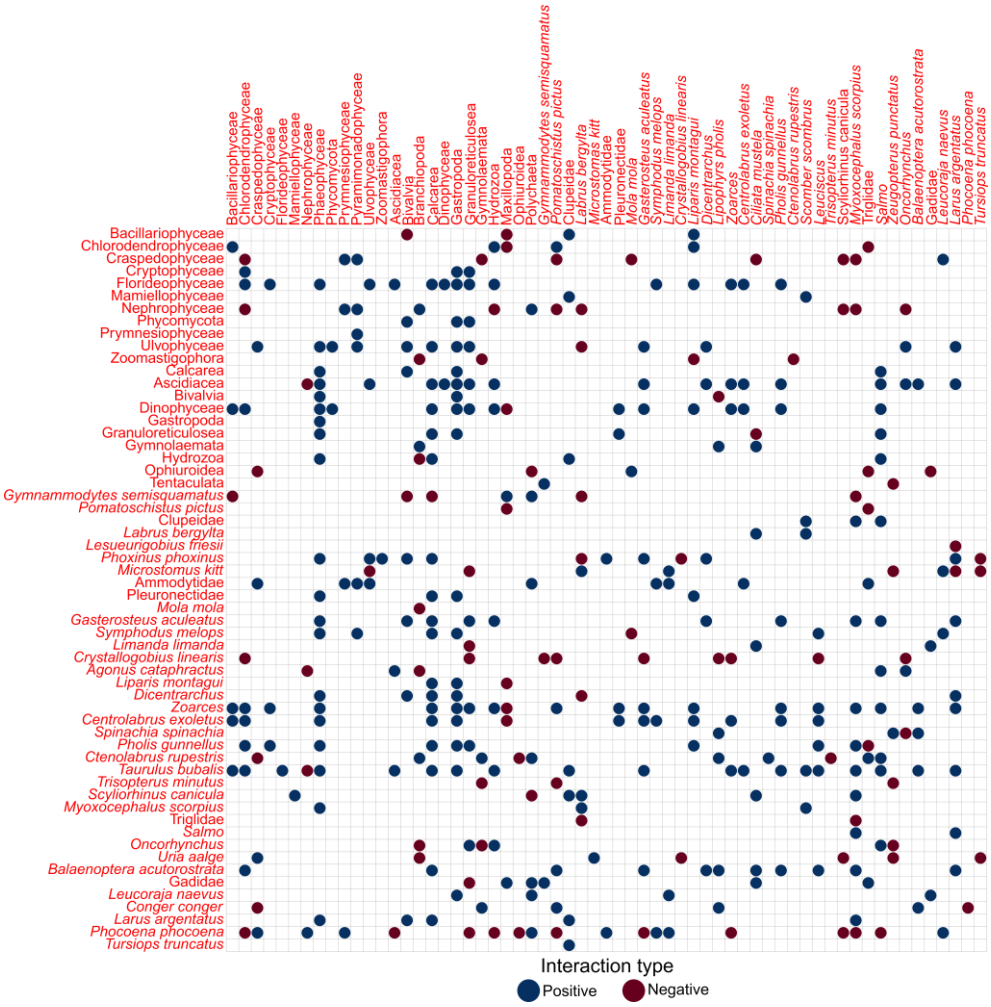
Chapter 3 - Inferring species interactions from co-occurrence networks with environmental DNA metabarcoding data



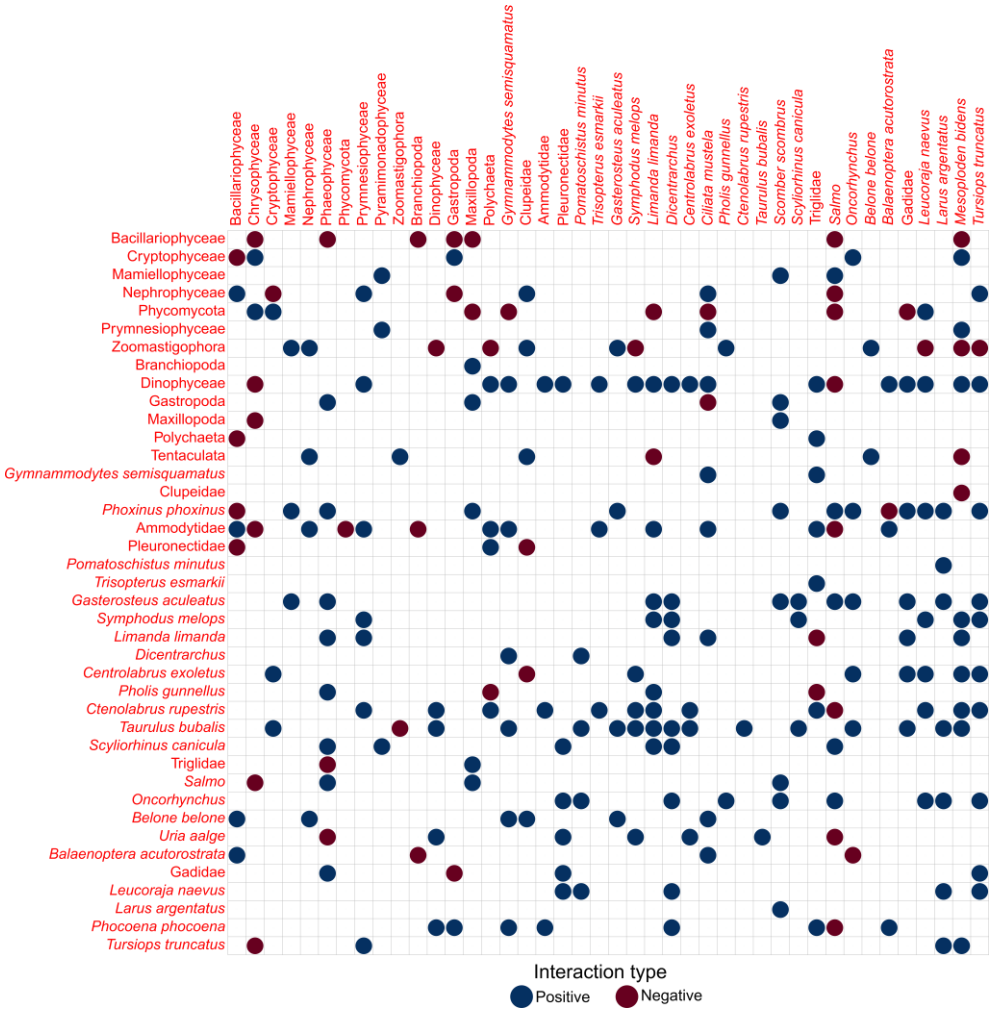
Appendix Figure A3.1. Correlation matrix showing positive and negative interactions detected in the early season (June-July) co-occurrence network.



Appendix Figure A3.2. Correlation matrix showing positive and negative interactions detected in the late season (August-October) co-occurrence network.



Appendix Figure A3.3. Correlation matrix showing positive and negative interactions detected in the nearshore co-occurrence network.



Appendix Figure A3.4. Correlation matrix showing positive and negative interactions detected in the offshore co-occurrence network.

Appendix Table A3.1. Ten OTUs with the highest closeness centrality in co-occurrence networks.

Co-occurrence networks		Co-occurrence networks	
Nearshore	Offshore	Jun-Jul	Aug-Oct
Zoarces	Dinophyceae	Phaeophyceae	*Oncorhynchus*
Calcarea	*Limanda limanda*	*Symphodus melops*	Dinophyceae
Ascidacea	*Salmo*	*Centrolabrus exoletus*	*Taurulus bubalis*
Granuloreticulosea	*Taurulus bubalis*	*Liparis montagui*	Ammodytidae
Phocoena phocoena	Ammodytidae	*Tursiops truncatus*	Zoarces
Gasterosteus aculeatus	*Tursiops truncatus*	*Salmo*	*Anguilla anguilla*
Ulvophyceae	*Gasterosteus aculeatus*	Dicentrachus	*Larus argentatus*
Myoxocephalus scorpius	*Oncorhynchus*	Spinachia spinachia	*Pholis gunnellus*
Chlorodendrophyceae	Bacillariophyceae	*Gasterosteus aculeatus*	*Phocoena phocoena*
Craspedophyceae	*Ctenolabrus rupestris*	*Chirolophis ascanii*	*Gasterosteus aculeatus*

Appendix Table A3.2. Ten OTUs with the highest betweenness centrality in co-occurrence networks.

Co-occurrence networks		Co-occurrence networks	
Nearshore	Offshore	Jun-Jul	Aug-Oct
Craspedophyceae	Dinophyceae	*Tursiops truncatus*	Dinophyceae
Phocoena phocoena	*Salmo*	*Ciliata septentrionalis*	Nephrophyceae
Granuloreticulosea	Ammodytidae	Phaeophyceae	Hydrozoa
Zoarces	*Limanda limanda*	*Salmo*	*Phocoena phocoena*
Larus argentatus	Bacillariophyceae	Polychaeta	*Uria aalge*
Balaenoptera acutorostrata	*Taurulus bubalis*	*Centrolabrus exoletus*	Florideophyceae
Ulvophyceae	*Oncorhynchus*	*Symphodus melops*	*Oncorhynchus*
Myoxocephalus scorpius	Zoomastigophora	*Liparis montagui*	*Taurulus bubalis*
Nephrophyceae	*Gasterosteus aculeatus*	*Gasterosteus aculeatus*	Gadidae
Gymnammodytes semisquatamus	*Tursiops truncatus*	Maxillopoda	*Larus argentatus*

Chapter 4 - Sampling from commercial vessel routes can capture marine biodiversity distributions effectively

Appendix Table A4.1. List of 43 predatory marine megafauna species that had > 40 occurrence points from combined data from online repositories, GBIF, OBIS and EurOBIS, Accobams and the Medlem database. Total length and trophic level as reported by FishBase (Actinopterygii and Chondrichthyes) or SeaLifeBase (Mammalia and Reptilia). Predatory marine megafauna defined as having total length ≥ 1m and trophic level ≥ 4. Species that do not meet these criteria but were retained are denoted with an Asterix.

Species	Common name	Class	Total length (cm)	Trophic Level	Number of occurrences
Conger conger	European Conger	Actinopterygii	573.3	4.3	649
Dentex dentex	Common dentex	Actinopterygii	100	4.5	372
Echelus myrus	Painted eel	Actinopterygii	100	4.3	45
Epinephelus aeneus	White grouper	Actinopterygii	120	4	49
Epinephelus marginatus	Dusky grouper	Actinopterygii	150	4.4	969
Fistularia commersonii	Bluespotted cornetfish	Actinopterygii	160	4.3	63
Lophius piscatorius	Angler	Actinopterygii	200 (SL)	4.5	602
Merluccius merluccius	European hake	Actinopterygii	140	4.4	1179
Mola mola	Ocean sunfish	Actinopterygii	333	3.3	3530
Molva dypterygia	Blue ling	Actinopterygii	155	4.5	119
Muraena helena	Mediterranean moray	Actinopterygii	150	4.2	1061
Ophisurus serpens	Serpent eel	Actinopterygii	250	4.1	47
Pomatomus saltatrix	Bluefish	Actinopterygii	130	4.5	82
Seriola dumerili	Greater amberjack	Actinopterygii	190	4.5	118
Sphyraena sphyraena	European barracuda	Actinopterygii	165	4	159
Sphyraena viridensis	Yellowmouth barracuda	Actinopterygii	128 (FL)	4.3	54
Thunnus alalunga	Albacore	Actinopterygii	140 (FL)	4.3	270
Thunnus thynnus	Bluefin tuna	Actinopterygii	458	4.5	177
Xiphias gladius	Swordfish	Actinopterygii	455	4.5	48
Zu cristatus	Scalloped ribbonfish	Actinopterygii	118 (SL)	4.5	204
Alopias vulpinas	Common thresher	Chondrichthyes	573.3	4.5	114
Carcharhinus longimanus	Oceanic whitetip shark	Chondrichthyes	400	4.2	77

*Cetorhinus maximus**	Basking shark	Chondrichthyes	1520	3.2	142
*Dasyatis pastinaca**	Common stingray	Chondrichthyes	64 (WD)	4.1	163
Echinorhinus brucus	Bramble shark	Chondrichthyes	310	4.4	41
Hexanchus griseus	Bluntnose sixgill shark	Chondrichthyes	482	4.5	140
Isurus oxyrinchus	Short-fin mako shark	Chondrichthyes	445	4.5	81
*Mobula mobular**	Giant devil ray	Chondrichthyes	520 (WD)	3.7	874
*Myliobatis aquila**	Common eagle ray	Chondrichthyes	183 (WD)	3.6	119
Prionace glauca	Blue shark	Chondrichthyes	400	4.4	324
*Raja clavata**	Thornback ray	Chondrichthyes	139	3.8	431
*Squalus acanthias**	Spiny dogfish	Chondrichthyes	95	4.4	78
Torpedo marmorata	Marbled electric ray	Chondrichthyes	100	4.5	116
Balaenoptera physalus	Fin whale	Mammalia	2700	3.2-4.3	1245
Delphinus delphis	Short-beaked common dolphin	Mammalia	260	4.5	1693
Globicephala melas	Long-finned pilot whale	Mammalia	670	4.5	1147
Grampus griseus	Risso's dolphin	Mammalia	380	4.36-4.54	410
Orcinus orca	Killer whale	Mammalia	980	4.5-4.6	115
Physeter macrocephalus	Sperm whale	Mammalia	2400	4.5-4.7	2307
Stenella coeruleoalba	Striped dolphin	Mammalia	260	4.5	7822
Tursiops truncatus	Bottlenose dolphin	Mammalia	380	4.5	3991
Ziphius cavirostris	Cuvier's beaked whale	Mammalia	750	4.5	113
*Caretta caretta**	Loggerhead turtle	Reptilia	125 (CL)	3.5-3.6	1557

Ray species *D. pastinaca* and *Torpedo torpedo* were retained due to having trophic levels greater than 4. Total length is not reported for these species so retained despite their width being less than 1 metre.

R. clavata and *S. acanthias* were retained as they were so close to threshold.

Mola mola, Cethorhinus maximus, Mobular mobular, Myliobatis aquila and *Caretta caretta* were retained as they are classed as megafauna species which can exert top down effects on ecosystems, similar to that of predators (Pimiento et al., 2020).

Pimiento, C., Leprieur, F., Silvestro, D., Lefcheck, J., Albouy, C., Rasher, D., Davis, M., Svenning, J.-C. and Griffin, J. 2020. Functional diversity of marine megafauna in the Anthropocene. *Science Advances.* **6**(16), peaay7650.

Appendix Methods A4.1

Principal component analysis (PCA) was used to create an environmental variability map. Initially, PCA was conducted on a correlation matrix of standardized environmental predictors used to create the SSDM using the prcomp() function in the 'stats' R package. The first four principal components were retained for downstream analysis as they explained >80 % of the variability in the environmental predictors. Site scores, *i.e.*, weighted linear combinations of the environmental predictors, were used to produce surface maps of each of the first four principal components to visualize the main gradients of environmental variability in the study area. The principal components were collapsed into one surface map of environmental variability by summing the site scores of each principal component weighted according to its contribution as following the equation:

EV = (0.3665*PC1) + inv(0.2435*PC2) + (0.164*PC3) + (0.1121*PC4)

Mean bathymetry made an important contribution to PC1 and PC2 (factor loading >0.4) but a high score in PC1 related to shallow bathymetry while a high score in PC2 related to greater depths. Therefore, PC2 was inverted otherwise the site scores offset each other and the variability was lost. The final 'environmental variability map' is unitless and shows the main gradients in environmental variability in the study area.

The first principal component explained 36.65 % of the variability in the environmental predictors where the highest values correspond to shallow bathymetry and low sea surface temperatures but high values of chlorophyll concentration and sea surface temperature range. The second principal component explained 24.35 % of the variability and represents variability related to distance from shore, where bathymetry is deepest further from shore. The third principal component explained 16.4 % of the variation where larger values correlated to the largest bathymetric slope. The fourth principal component explained 11.21 % of the variation with the northern Adriatic clearly being the most different area due to having a high chlorophyll concentration and lower sea surface temperature.

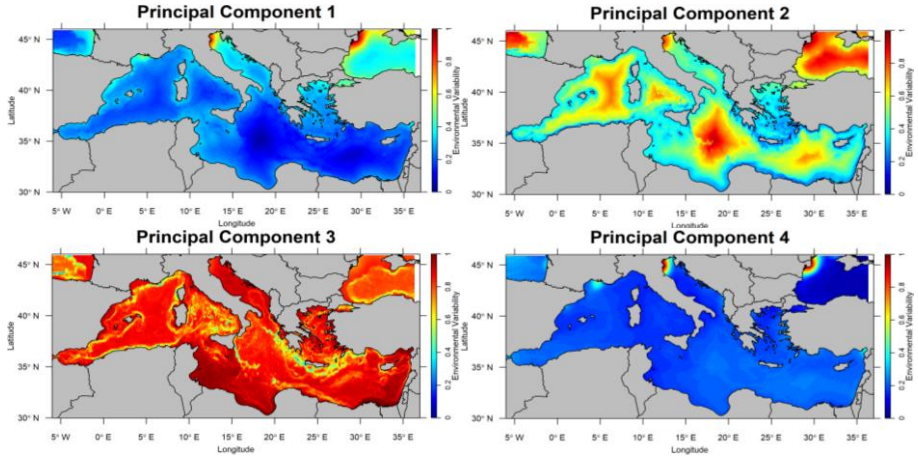
The weighted overlay of the principal components shows the overall trends in environmental variability where the values are unitless but the larger the range between values represents areas with the most different environmental conditions. There are two main trends in environmental variability, 1) a difference between the north-western and south-eastern basin, 2) a gradient with distance from shore. The northern tip of the Adriatic Sea is clearly most different from the rest of the Mediterranean.

Principal components respective contribution ratios.

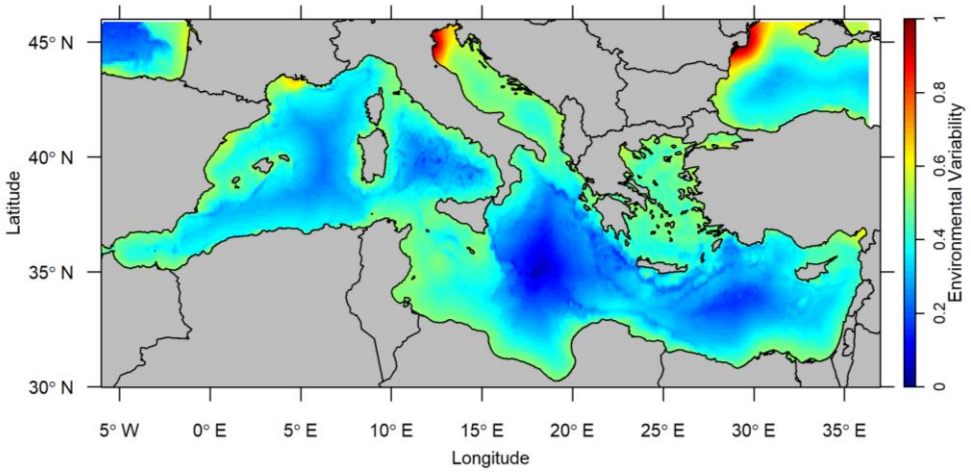
	PC1	PC2	PC3	PC4	PC5	PC6
Eigenvalue	1.4828	1.2088	0.992	0.8203	0.65213	0.50760
Contribution ratio (%)	0.3665	0.2435	0.164	0.1121	0.07088	0.04294
Cumulative contribution (%)	0.3665	0.6100	0.774	0.8862	0.95706	1.0000

Eigenvectors. Factor loadings >0.4 are highlighted in bold.

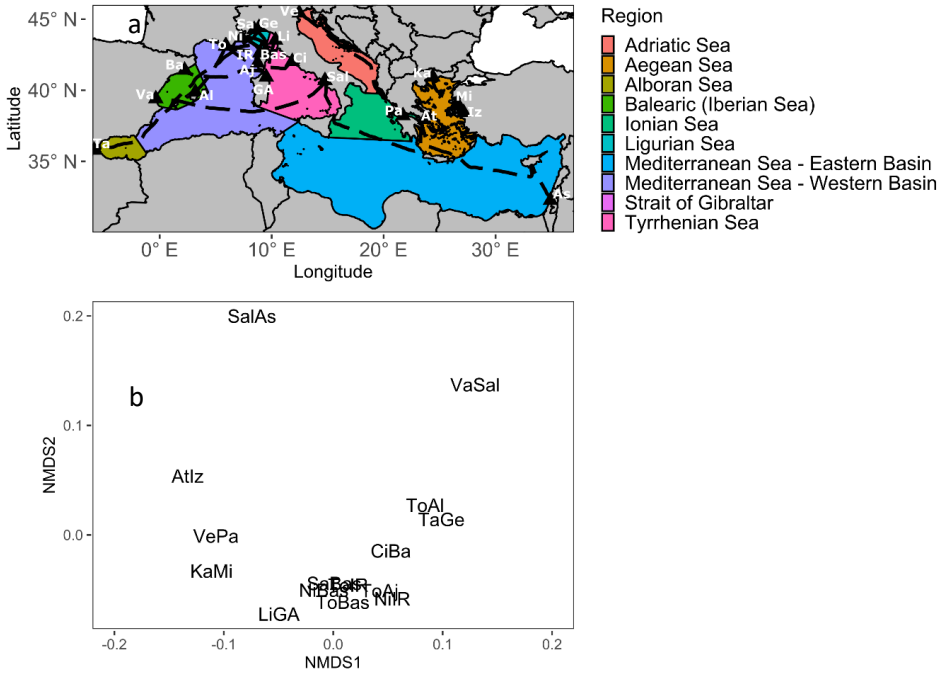
Variable (units)	PC1	PC2	PC3	PC4
Mean bathymetry (m)	**-0.4281**	**0.5227**	-0.2578	0.0677
Mean sea surface temperature (°C)	**-0.4773**	-0.3344	0.3005	0.2529
Mean chlorophyll concentration (mg/m^3)	**0.4436**	0.1910	-0.0693	**0.8660**
Sea surface temperature range (°C)	**0.4783**	0.3533	-0.1049	**-0.4166**
Bathymetric slope (°)	-0.1490	-0.2761	**-0.9079**	0.0511
Distance to shore (km)	-0.3756	**0.6143**	0.0568	0.0735



Principal component scores projected onto the study area.



Environmental variability map using weighted overlay of the first four principal components.



Appendix Figure A4.1. a) Map of the Mediterranean showing regions covered by different ferry routes. b) Non-metric multidimensional scaling plot based on Bray-Curtis dissimilarity matrix for species composition between different ferry routes.

Appendix Table A4.2. Pearson correlation coefficient calculated from the difference between a full model and one with each environmental variable omitted in turn for individual species models then averaged across species.

	Mean bathymetry	Mean sea surface temperature	Mean chlorophyll concentration	Mean temperature range	Bathymetric slope	Distance from shore
Mean	14.59	31.76	13.17	16.7	8.31	15.47
SD	6.56	13.74	5.1	5.53	1.96	8.32

Appendix Table A4.3. Results from six metrics used to evaluate the prediction accuracy of species assemblage predictions in the 'perfect knowledge' SSDM.

	Species Richness Error	Prediction Success	Kappa	Specificity	Sensitivity	Jaccard
Mean	19.06	5.98	0.995029	0.544447	0.983425	0.065596
SD	7.229589	1.797052	0.000441	0.172317	0.124641	0.059542

Appendix Table A4.4. Analysis of Variance Tables

One-way ANOVA to evaluate the impact of different ferry route subnetworks on correlation coefficients.

Factor	Df	Sum Sq	Mean Sq	F value	P value
Sampling frame	2	1.20122	0.60061	342.96	<0.001*

Tukey's Post Hoc Test to evaluate which ferry route subnetworks differed from each other.

Group 1	Group 2	Estimate	Conf. low	Conf. high	P value
Biased	Community	0.384	0.336	0.432	<0.001*
Biased	Environment	-0.0952	-0.142	-0.0487	<0.001*
Community	Environment	-0.479	-0.527	-0.431	<0.001*

One-way ANOVA to evaluate the effect of the number of ferry routes in a ferry subnetwork. Dependent variable was square transformed prior to analysis.

Factor	Df	Sum Sq	Mean Sq	F value	P value
Number of ferries	5	1.2987	0.259745	15.286	<0.001*

Tukey's Post Hoc Test to evaluate which ferry subnetworks differed depending on the number of ferry routes.

Group 1	Group 2	Estimate	Conf. low	Conf. high	P Value
2	4	0.116	-0.0558	0.289	ns
2	6	0.149	-0.0233	0.321	ns
2	8	0.379	0.207	0.552	<0.001*
2	10	0.364	0.191	0.536	<0.001*
2	12	0.362	0.189	0.534	<0.001*
4	6	0.0325	-0.140	0.205	ns
4	8	0.263	0.0907	0.435	<0.001*
4	10	0.247	0.0750	0.419	<0.001*
4	12	0.245	0.0730	0.418	<0.001*
6	8	0.230	0.0582	0.403	<0.001*
6	10	0.215	0.0425	0.387	<0.001*
6	12	0.213	0.0405	0.385	<0.001*
8	10	-0.0157	-0.188	0.157	ns
8	12	-0.0177	-0.190	0.155	ns
10	12	-0.00194	-0.174	0.170	ns

Three-way ANOVA to evaluate the impact of removing specific taxa from stacked species distribution models across sampling strategies (random vs regular) and sampling sizes (25, 50, 100 sampling points). Dependent variable (the correlation coefficient) was square transformed prior to analysis.

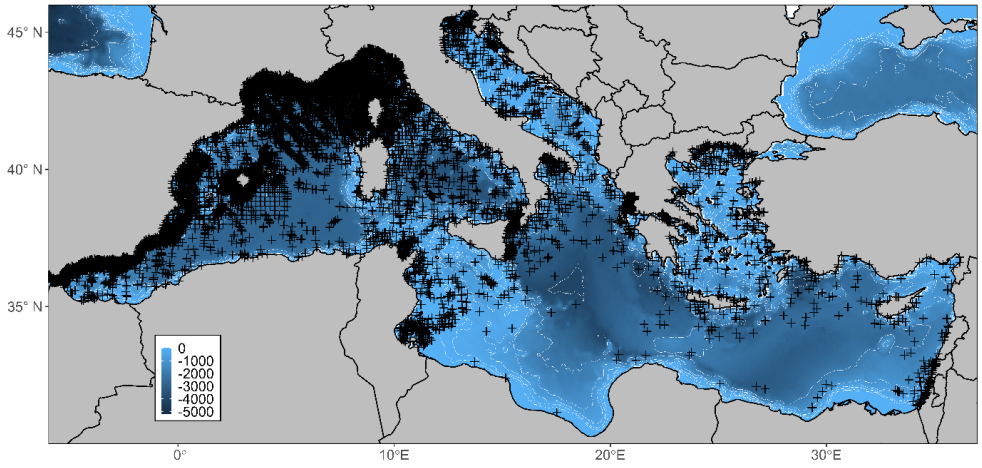
Factor	Df	Sum Sq	Mean Sq	F value	P value
Strategy	1	0.08385	0.083854	7.9500	<0.05*
Size	2	0.26295	0.131475	12.4648	<0.001*
Taxa removed	2	0.18388	0.091941	8.7167	<0.001*

Tukey's Post Hoc Test to look at pairwise differences between taxa removed, sampling strategy and sampling size.

Term	Group 1	Group 2	Estimate	Conf. low	Conf. high	P value
Strategy	Random	Regular	0.0522	0.0155	0.0889	<0.05*
Size	25	50	-0.0400	-0.0936	0.0136	ns
Size	25	100	0.0715	0.0170	0.126	<0.05*
Size	50	100	0.112	0.0580	0.165	<0.001*
Taxa removed	Chondrichthyes	Mammalia	0.106	0.0428	0.169	<0.001*
Taxa removed	Chondrichthyes	Random	0.0299	-0.0241	0.0840	ns
Taxa removed	Mammalia	Random	-0.0758	-0.130	-0.0217	<0.05*

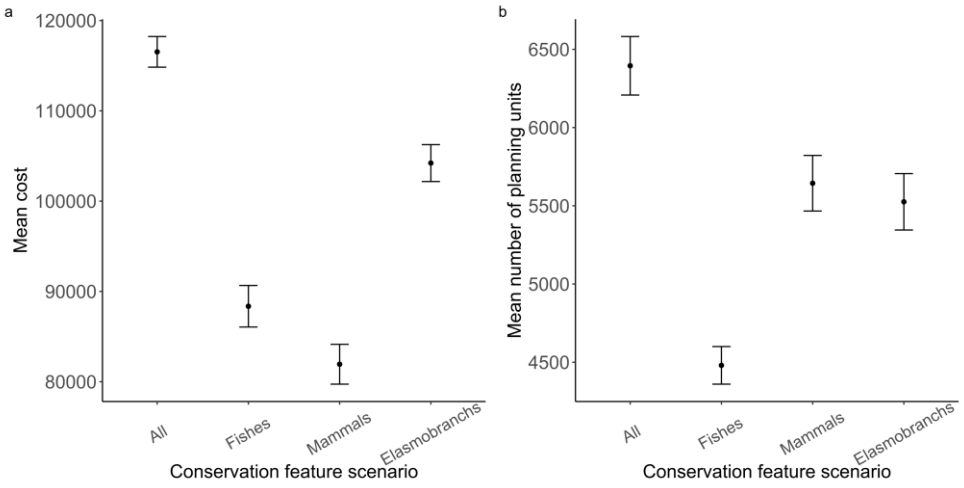
Appendix Table A4.5. The ferry route or ferry route subnetwork length, the number of species distributions overlapping with the ferry route or subnetwork, and the climatic bias index.

Ferry Route	Ferry length (no. grid cells covered)	Number of species	Climatic Bias Index
SaAs	404	36	0.3944031
TaGe	278	41	0.17952261
VaSa	218	36	0.42306601
VePa	194	38	1.0530864
CiBa	144	39	0.38805581
AtIz	81	34	1.1377996
ToAl	69	39	0.6424208
ToBas	55	38	0.43982341
KaMi	52	36	1.1377996
ToAj	48	39	0.77694461
NiBas	43	39	0.1230776
LiGA	42	40	1.1377996
ToIR	42	39	0.60432561
SaBas	35	38	0.39494239
NiIR	32	37	0.797778
All ferries	1744	42	0.06914759
Biased subnetwork	513	41	0.3147127
Community subnetwork	1462	42	0.07962026
Environment subnetwork	598	39	0.08448943



Appendix Figure A4.2. Bathymetric map of the Mediterranean Sea showing raw occurrence data from the online data repositories GBIF, OBIS and EurOBIS as well as the ACCOBAMS and Medlem datasets for cetaceans and elasmobranchs respectively.

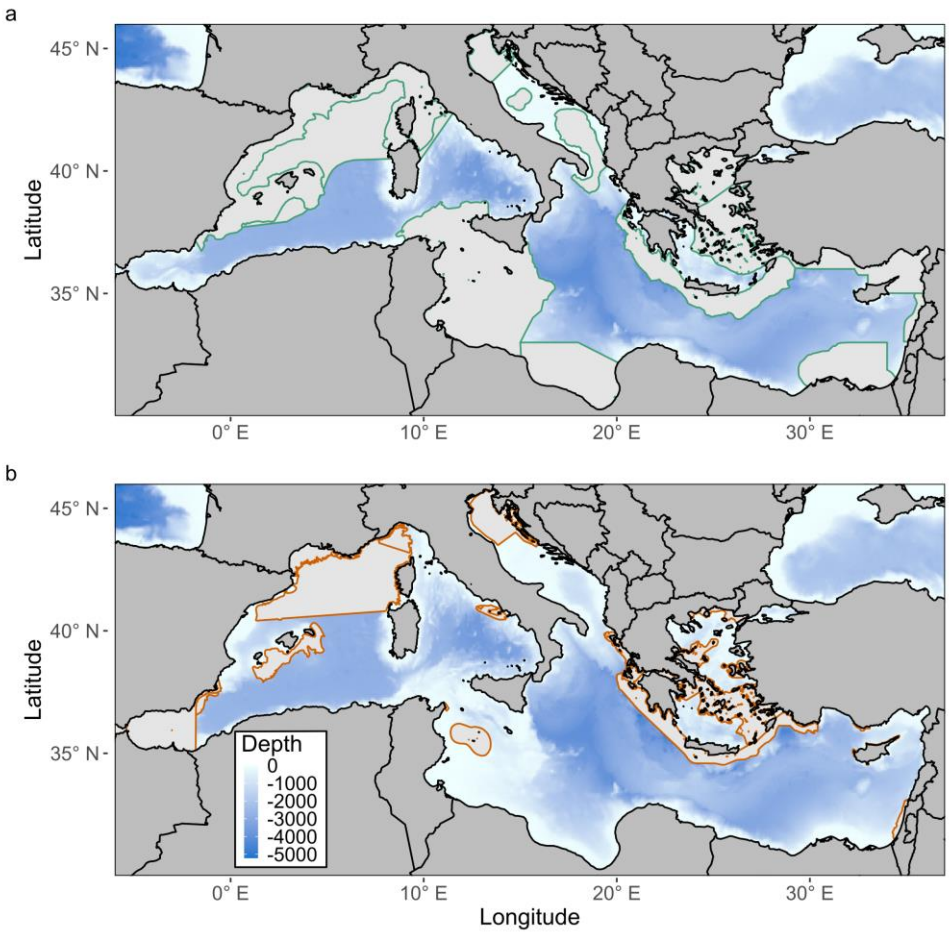
Chapter 5 - Multi-taxon versus taxon specific spatial conservation priorities for predatory marine megafauna



Appendix Figure A5.1. a) Mean cost across 10 best Marxan solutions for each of the conservation feature scenarios. b) Mean number of planning units in solutions from 10 best Marxan runs for each of the conservation feature scenarios.

Appendix Table A5.1. Kruskal Wallis tests to compare the overlap between MPAs or conservation feature scenarios for mammals, fishes and sharks and species distributions of individual taxa (mammals, fishes, sharks).

	Chi-squared	Degrees of freedom	P value
MPAs	3.2742	2	0.1945
Mammals	108.43	2	<0.001*
Fishes	142.49	2	<0.001*
Sharks	55.629	2	<0.001*



Appendix Figure A5.2. Bathymetric map of the Mediterranean Sea displaying a) Ecologically or Biological Significant areas (EBSAs) and b) Important Marine Mammal Areas (IMMAs).

References

Abalo-Morla, S., Belda, E. J., March, D., Revuelta, O., Cardona, L., Giralt, S., Crespo-Picazo, J. L., Hochscheid, S., Marco, A., & Merchán, M. (2022). Assessing the use of marine protected areas by loggerhead sea turtles (Caretta caretta) tracked from the western Mediterranean. *Global Ecology and Conservation*, *38*, e02196.

Abbott, C., Coulson, M., Gagné, N., Lacoursière-Roussel, A., Parent, G. J., Bajno, R., Dietrich, C., & May-McNally, S. (2021). *Guidance on the use of targeted environmental DNA (eDNA) analysis for the management of aquatic invasive species and species at risk*. Canadian Science Advisory Secretariat (CSAS) Ottawa, ON, Canada.

ACCOBAMS Survey Initiative. (2020). Thirteenth Meeting of the Accobams Scientific Committee. Technical Reports of the Mediterranean and Black Sea Surveys. Monaco.

Adams, C. I., Knapp, M., Gemmell, N. J., Jeunen, G.-J., Bunce, M., Lamare, M. D., & Taylor, H. R. (2019). Beyond biodiversity: Can environmental DNA (eDNA) cut it as a population genetics tool? *Genes*, *10*(3), 192.

Adams, J., & Martin, J. (1986). The hydrography and plankton of the Moray Firth. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, *91*, 37-56.

Afzali, S. F., Bourdages, H., Laporte, M., Mérot, C., Normandeau, E., Audet, C., & Bernatchez, L. (2021). Comparing environmental metabarcoding and trawling survey of demersal fish communities in the Gulf of St. Lawrence, Canada. *Environmental DNA*, *3*(1), 22-42.

Agardy, T., Cody, M., Hastings, S., Hoyt, E., Nelson, A., Tetley, M., & Notarbartolo di Sciara, G. (2019). Looking beyond the horizon: an early warning system to keep marine mammal information relevant for conservation. *Aquatic conservation: marine and freshwater ecosystems*, *29*, 71-83.

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*(5), 541-545.

Aïssi, M., Arcangeli, A., Crosti, R., Daly Yahia, M., Loussaief, B., Moulins, A., Pellegrino, G., Rosso, M., Tringali, L., & Tepsich, P. (2015). Cetacean occurrence and spatial distribution in the central Mediterranean Sea using ferries as platform of observation. *Russian Journal of Marine Biology*, *41*, 343-350.

Albonetti, L., Maiello, G., Cariani, A., Carpentieri, P., Ferrari, A., Sbrana, A., Shum, P., Talarico, L., Russo, T., & Mariani, S. (2023). DNA metabarcoding of trawling bycatch reveals diversity and distribution patterns of sharks and rays in the central Tyrrhenian Sea. *ICES Journal of Marine Science*, fsad022.

Albouy, C., Archambault, P., Appeltans, W., Araújo, M. B., Beauchesne, D., Cazelles, K., Cirtwill, A. R., Fortin, M.-J., Galiana, N., & Leroux, S. J. (2019). The marine fish food web is globally connected. *Nature Ecology & Evolution*, *3*(8), 1153-1161.

Albouy, C., Delattre, V., Donati, G., Frölicher, T. L., Albouy-Boyer, S., Rufino, M., Pellissier, L., Mouillot, D., & Leprieur, F. (2020). Global vulnerability of marine mammals to global warming. *Scientific reports*, *10*(1), 1-12.

Alvarez-Fernandez, S., Lindeboom, H., & Meesters, E. (2012). Temporal changes in plankton of the North Sea: community shifts and environmental drivers. *Marine Ecology Progress Series*, *462*, 21-38.

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., & Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PloS one*, *4*(7), e6372.

Amengual, J., & Alvarez-Berastegui, D. (2018). A critical evaluation of the Aichi Biodiversity Target 11 and the Mediterranean MPA network, two years ahead of its deadline. *Biological Conservation*, *225*, 187-196.

Ames, C. L., Ohdera, A. H., Colston, S. M., Collins, A. G., Fitt, W. K., Morandini, A. C., Erickson, J. S., & Vora, G. J. (2021). Fieldable environmental DNA sequencing to assess jellyfish biodiversity in nearshore waters of the Florida Keys, United States. *Frontiers in Marine Science*, *8*, 640527.

Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., & Boehm, A. B. (2019). Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. *Frontiers in Marine Science*, *6*, 477.

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*(10), 1677-1688.

Araújo, M. B., & Rozenfeld, A. (2014). The geographic scaling of biotic interactions. *Ecography*, *37*(5), 406-415.

Arcangeli, A., Campana, I., & Bologna, M. A. (2017). Influence of seasonality on cetacean diversity, abundance, distribution and habitat use in the western Mediterranean Sea: implications for conservation. *Aquatic conservation: marine and freshwater ecosystems*, *27*(5), 995-1010.

Arenas-Castro, S., Regos, A., Martins, I., Honrado, J., & Alonso, J. (2022). Effects of input data sources on species distribution model predictions across species with different distributional ranges. *Journal of Biogeography*, *49*(7), 1299-1312.

Armstrong, J. D., Bean, C. W., & Wells, A. (2018). The Scottish invasion of pink salmon in 2017. *Journal of fish biology*, *93*(1), 8-11.

Arthington, A. H., Dulvy, N. K., Gladstone, W., & Winfield, I. J. (2016). Fish conservation in freshwater and marine realms: status, threats and management. *Aquatic conservation: marine and freshwater ecosystems*, *26*(5), 838-857.

Assenov, Y., Ramírez, F., Schelhorn, S.-E., Lengauer, T., & Albrecht, M. (2008). Computing topological parameters of biological networks. *Bioinformatics*, *24*(2), 282-284.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., & De Clerck, O. (2018). Bio-ORACLE v2. 0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, *27*(3), 277-284.

Astudillo-Scalia, Y., Albuquerque, F., Polidoro, B., & Beier, P. (2021). Environmental diversity as a reliable surrogacy strategy of marine biodiversity: A case study of marine mammals. *Perspectives in Ecology and Conservation*, *19*(4), 429-434.

Aubert, A., Antajan, E., Lynam, C., Pitois, S., Pliru, A., Vaz, S., & Thibault, D. (2018). No more reason for ignoring gelatinous zooplankton in ecosystem assessment and marine management: concrete cost-effective methodology during routine fishery trawl surveys. *Marine Policy*, *89*, 100-108.

Augé, A. A., Dias, M. P., Lascelles, B., Baylis, A. M., Black, A., Boersma, P. D., Catry, P., Crofts, S., Galimberti, F., & Granadeiro, J. P. (2018). Framework for mapping key areas for marine megafauna to inform Marine Spatial Planning: The Falkland Islands case study. *Marine Policy*, *92*, 61-72.

Avila, I. C., Kaschner, K., & Dormann, C. F. (2018). Current global risks to marine mammals: taking stock of the threats. *Biological Conservation*, *221*, 44-58.

Azzellino, A., Panigada, S., Lanfredi, C., Zanardelli, M., Airoldi, S., & di Sciara, G. N. (2012). Predictive habitat models for managing marine areas: spatial and temporal distribution of marine mammals within the Pelagos Sanctuary (Northwestern Mediterranean sea). *Ocean & Coastal Management*, *67*, 63-74.

Bakker, J., Wangensteen, O. S., Chapman, D. D., Boussarie, G., Buddo, D., Guttridge, T. L., Hertler, H., Mouillot, D., Vigliola, L., & Mariani, S. (2017).

Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. *Scientific reports*, *7*(1), 1-11.

Balint, M., Pfenninger, M., Grossart, H.-P., Taberlet, P., Vellend, M., Leibold, M. A., Englund, G., & Bowler, D. (2018). Environmental DNA time series in ecology. *Trends in ecology & evolution*, *33*(12), 945-957.

Ball, I. R., Possingham, H. P., & Watts, M. (2009). Marxan and relatives: software for spatial conservation prioritisation. *Spatial conservation prioritisation: Quantitative methods and computational tools*, 185-195.

Bani, A., De Brauwer, M., Creer, S., Dumbrell, A. J., Limmon, G., Jompa, J., von der Heyden, S., & Beger, M. (2020). Informing marine spatial planning decisions with environmental DNA. In *Advances in ecological research* (Vol. 62, pp. 375-407). Elsevier.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327-338.

Barco, A., Kullmann, B., Knebelsberger, T., Sarrazin, V., Kuhs, V., Kreutle, A., Pusch, C., & Thiel, R. (2022). Detection of fish species from marine protected areas of the North Sea using environmental DNA. *Journal of fish biology*, *101*(3), 722-727.

Barroso-Bergada, D., Pauvert, C., Vallance, J., Deliere, L., Bohan, D. A., Buee, M., & Vacher, C. (2021). Microbial networks inferred from environmental DNA data for biomonitoring ecosystem change: Strengths and pitfalls. *Molecular ecology resources*, *21*(3), 762-780.

Baumgartner, M. F., Bonnell, J., Van Parijs, S. M., Corkeron, P. J., Hotchkin, C., Ball, K., Pelletier, L. P., Partan, J., Peters, D., & Kemp, J. (2019). Persistent near real-time passive acoustic monitoring for baleen whales from a moored buoy: System description and evaluation. *Methods in Ecology and Evolution*, *10*(9), 1476-1489.

Bearzi, G., Agazzi, S., Gonzalvo, J., Costa, M., Bonizzoni, S., Politi, E., Piroddi, C., & Reeves, R. R. (2008). Overfishing and the disappearance of short-beaked common dolphins from western Greece. *Endangered Species Research*, *5*(1), 1-12.

Bearzi, G., & Reeves, R. R. (2021). Shifting baselines of cetacean conservation in Europe. *ICES Journal of Marine Science*, *78*(7), 2337-2341.

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, *19*, 10-15.

Belkin, I. M. (2009). Rapid warming of large marine ecosystems. *Progress in Oceanography*, *81*(1-4), 207-213.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289-300.

Berger-Tal, O., & Lahoz-Monfort, J. J. (2018). Conservation technology: The next generation. *Conservation Letters*, *11*(6), e12458.

Berry, D., & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in microbiology*, *5*, 219.

Bessey, C., Neil Jarman, S., Simpson, T., Miller, H., Stewart, T., Kenneth Keesing, J., & Berry, O. (2021). Passive eDNA collection enhances aquatic biodiversity analysis. *Communications Biology*, *4*(1), 236.

Bivand, R. S., Pebesma, E. J., & Gomez-Rubio, V. (2008). Applied spatial data analysis with R. Springer. In: Springer.

Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, *23*(7), 1050-1063.

Boettiger, C., Lang, D. T., & Wainwright, P. (2012). rfishbase: exploring, manipulating and visualizing FishBase data from R. *Journal of fish biology*, *81*(6), 2030-2039.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2022). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular ecology resources*, *22*(4), 1231-1246.

Bonebrake, T. C., Brown, C. J., Bell, J. D., Blanchard, J. L., Chauvenet, A., Champion, C., Chen, I. C., Clark, T. D., Colwell, R. K., & Danielsen, F. (2018). Managing consequences of climate-driven species redistribution requires integration of ecology, conservation and social science. *Biological Reviews*, *93*(1), 284-305.

Bossier, S., Nielsen, J. R., & Neuenfeldt, S. (2020). Exploring trophic interactions and cascades in the Baltic Sea using a complex end-to-end ecosystem model with extensive food web integration. *Ecological modelling*, *436*, 109281.

Boulanger, E., Loiseau, N., Valentini, A., Arnal, V., Boissery, P., Dejean, T., Deter, J., Guellati, N., Holon, F., & Juhel, J.-B. (2021). Environmental DNA metabarcoding reveals and unpacks a biodiversity conservation paradox in Mediterranean marine reserves. *Proceedings of the Royal Society B*, *288*(1949), 20210112.

Boussarie, G., Bakker, J., Wangensteen, O. S., Mariani, S., Bonnin, L., Juhel, J.-B., Kiszka, J. J., Kulbicki, M., Manel, S., & Robbins, W. D. (2018). Environmental DNA illuminates the dark diversity of sharks. *Science Advances*, *4*(5), eaap9661.

Boyse, E., Beger, M., Valsecchi, E., & Goodman, S. J. (2023). Sampling from commercial vessel routes can capture marine biodiversity distributions effectively. *Ecology and Evolution*, *13*(2), e9810.

Bradai, M. N., Saidi, B., & Enajjar, S. (2012). Elasmobranchs of the Mediterranean and Black Sea: status, ecology and biology bibliographic analysis. *Studies and Reviews-General Fisheries Commission for the Mediterranean*(91).

Braunisch, V., & Suchant, R. (2010). Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. *Ecography*, *33*(5), 826-840.

Brown, M. B. (1975). 400: A method for combining non-independent, one-sided tests of significance. *Biometrics*, 987-992.

Bruce, K., Blackman, R. C., Bourlat, S. J., Hellström, M., Bakker, J., Bista, I., Bohmann, K., Bouchez, A., Brys, R., & Clark, K. (2021). *A practical guide to DNA-based methods for biodiversity assessment*. Pensoft Advanced Books.

Brumm, K. J., Hanks, R. D., Baldwin, R. F., & Peoples, B. K. (2021). Accounting for multiple dimensions of biodiversity to assess surrogate performance in a freshwater conservation prioritization. *Ecological Indicators*, *122*, 107320.

Budd, A. M., Cooper, M. K., Le Port, A., Schils, T., Mills, M. S., Deinhart, M. E., Huerlimann, R., & Strugnell, J. M. (2021). First detection of critically endangered scalloped hammerhead sharks (Sphyrna lewini) in Guam, Micronesia, in five decades using environmental DNA. *Ecological Indicators*, *127*, 107649.

Budd, A. M., Schils, T., Cooper, M. K., Lyons, M. B., Mills, M. S., Deinhart, M. E., Le Port, A., Huerlimann, R., & Strugnell, J. M. (2023). Monitoring threatened species with environmental DNA and open ecological data: Local distribution and habitat preferences of scalloped hammerhead sharks (Sphyrna lewini). *Biological Conservation*, *278*, 109881.

Cáceres, M. D., & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology*, *90*(12), 3566-3574.

Calabrese, J. M., Certain, G., Kraan, C., & Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, *23*(1), 99-112.

Calderón-Sanou, I., Münkemüller, T., Boyer, F., Zinger, L., & Thuiller, W. (2020). From environmental DNA sequences to ecological conclusions: How strong is the influence of methodological choices? *Journal of Biogeography*, *47*(1), 193-206.

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal*, *11*(12), 2639-2643.

Cammen, K. M., Andrews, K. R., Carroll, E. L., Foote, A. D., Humble, E., Khudyakov, J. I., Louis, M., McGowen, M. R., Olsen, M. T., & Van Cise, A. M. (2016). Genomic methods take the plunge: recent advances in high-throughput sequencing of marine mammals. *Journal of Heredity*, *107*(6), 481-495.

Carlucci, R., Manea, E., Ricci, P., Cipriano, G., Fanizza, C., Maglietta, R., & Gissi, E. (2021). Managing multiple pressures for cetaceans' conservation with an Ecosystem-Based Marine Spatial Planning approach. *Journal of Environmental Management*, *287*, 112240.

Carter, M. I., Boehme, L., Duck, C. D., Grecian, J., Hastie, G. D., McConnell, B. J., Miller, D. L., Morris, C., Moss, S., & Thompson, D. (2020). Habitat-based predictions of at-sea distribution for grey and harbour seals in the British Isles: Report to BEIS, OESEA-16-76, OESEA-17-78.

Cashion, T., Nguyen, T., Ten Brink, T., Mook, A., Palacios-Abrantes, J., & Roberts, S. M. (2020). Shifting seas, shifting boundaries: dynamic marine protected area designs for a changing climate. *PloS one*, *15*(11), e0241771.

Cazelles, K., Araújo, M. B., Mouquet, N., & Gravel, D. (2016). A theory for species co-occurrence in interaction networks. *Theoretical Ecology*, *9*, 39-48.

CBD. (2021). First draft of the post-2020 global biodiversity framework.

Cheney, B., Thompson, P. M., Ingram, S. N., Hammond, P. S., Stevick, P. T., Durban, J. W., Culloch, R. M., Elwen, S. H., Mandleberg, L., & Janik, V. M. (2013). Integrating multiple data sources to assess the distribution and abundance of bottlenose dolphins T ursiops truncatus in Scottish waters. *Mammal Review*, *43*(1), 71-88.

Cisneros-Montemayor, A. M., Sumaila, U. R., Kaschner, K., & Pauly, D. (2010). The global potential for whale watching. *Marine Policy*, *34*(6), 1273-1278.

Claudet, J., Loiseau, C., Sostres, M., & Zupan, M. (2020). Underprotected marine protected areas in a global biodiversity hotspot. *One Earth*, *2*(4), 380-384.

Clausen, L. W., Rindorf, A., van Deurs, M., Dickey-Collas, M., & Hintzen, N. T. (2018). Shifts in North Sea forage fish productivity and potential fisheries yield. *Journal of Applied Ecology*, *55*(3), 1092-1101.

Claver, C., Canals, O., de Amézaga, L. G., Mendibil, I., & Rodriguez-Ezpeleta, N. (2023). An automated workflow to assess completeness and curate GenBank for environmental DNA metabarcoding: The marine fish assemblage as case study. *Environmental DNA*.

Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., Bograd, S. J., Hazen, E. L., Field, J. C., & Boehm, A. B. (2019). Marine vertebrate biodiversity and distribution within the central California Current using environmental DNA (eDNA) metabarcoding and ecosystem surveys. *Frontiers in Marine Science*, *6*, 732.

Coll, M., Piroddi, C., Albouy, C., Ben Rais Lasram, F., Cheung, W. W., Christensen, V., Karpouzi, V. S., Guilhaumon, F., Mouillot, D., & Paleczny, M. (2012). The Mediterranean Sea under siege: spatial overlap between marine biodiversity, cumulative threats and marine reserves. *Global Ecology and Biogeography*, *21*(4), 465-480.

Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Ben Rais Lasram, F., Aguzzi, J., Ballesteros, E., Bianchi, C. N., Corbera, J., & Dailianis, T. (2010). The biodiversity of the Mediterranean Sea: estimates, patterns, and threats. *PloS one*, *5*(8), e11842.

Conners, M. G., Sisson, N. B., Agamboue, P. D., Atkinson, P. W., Baylis, A. M., Benson, S. R., Block, B. A., Bograd, S. J., Bordino, P., & Bowen, W. (2022). Mismatches in scale between highly mobile marine megafauna and marine protected areas. *Frontiers in Marine Science*, 1236.

Corrigan, C. M., Ardron, J. A., Comeros-Raynal, M. T., Hoyt, E., Notarbartolo Di Sciara, G., & Carpenter, K. E. (2014). Developing important marine mammal area criteria: learning from ecologically or biologically significant areas and key biodiversity areas. *Aquatic conservation: marine and freshwater ecosystems*, *24*(S2), 166-183.

Cox, S., Embling, C., Hosegood, P., Votier, S., & Ingram, S. (2018). Oceanographic drivers of marine mammal and seabird habitat-use across shelf-seas: a guide to key features and recommendations for future research and conservation management. *Estuarine, Coastal and Shelf Science*, *212*, 294-310.

Crespo, G. O., Mossop, J., Dunn, D., Gjerde, K., Hazen, E., Reygondeau, G., Warner, R., Tittensor, D., & Halpin, P. (2020). Beyond static spatial management: Scientific and legal considerations for dynamic management in the high seas. *Marine Policy*, *122*, 104102.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, complex systems*, *1695*(5), 1-9.

Culloch, R. M., & Robinson, K. P. (2008). Bottlenose dolphins using coastal regions adjacent to a Special Area of Conservation in north-east Scotland. *Journal of the Marine Biological Association of the United Kingdom*, *88*(6), 1237-1243.

d'Amen, M., Dubuis, A., Fernandes, R. F., Pottier, J., Pellissier, L., & Guisan, A. (2015). Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography*, *42*(7), 1255-1266.

D'Alessandro, S., & Mariani, S. (2021). Sifting environmental DNA metabarcoding data sets for rapid reconstruction of marine food webs. *Fish and Fisheries*, *22*(4), 822-833.

Dahood, A., de Mutsert, K., & Watters, G. M. (2020). Evaluating Antarctic marine protected area scenarios using a dynamic food web model. *Biological Conservation*, *251*, 108766.

Davidson, A. D., Boyer, A. G., Kim, H., Pompa-Mansilla, S., Hamilton, M. J., Costa, D. P., Ceballos, G., & Brown, J. H. (2012). Drivers and hotspots of extinction risk in marine mammals. *Proceedings of the National Academy of Sciences*, *109*(9), 3395-3400.

De Brauwer, M., Clarke, L. J., Chariton, A., Cooper, M. K., De Bruyn, M., Furlan, E., MacDonald, A. J., Rourke, M. L., Sherman, C. D., & Suter, L. (2023). Best practice guidelines for environmental DNA biomonitoring in Australia and New Zealand. *Environmental DNA*.

De Juan, S., & Lleonart, J. (2010). A conceptual framework for the protection of vulnerable habitats impacted by fishing activities in the Mediterranean high seas. *Ocean & Coastal Management*, *53*(11), 717-723.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., & De Vere, N. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular ecology*, *26*(21), 5872-5895.

Delmas, E., Besson, M., Brice, M. H., Burkle, L. A., Dalla Riva, G. V., Fortin, M. J., Gravel, D., Guimarães Jr, P. R., Hembry, D. H., & Newman, E. A. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, *94*(1), 16-36.

Devillers, R., Pressey, R. L., Grech, A., Kittinger, J. N., Edgar, G. J., Ward, T., & Watson, R. (2015). Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquatic conservation: marine and freshwater ecosystems*, *25*(4), 480-504.

Devillers, R., Pressey, R. L., Ward, T. J., Grech, A., Kittinger, J. N., Edgar, G. J., & Watson, R. A. (2020). Residual marine protected areas five years on: Are we still favouring ease of establishment over need for protection? *Aquatic conservation: marine and freshwater ecosystems*, *30*(9), 1758-1764.

Di Marco, M., Chapman, S., Althor, G., Kearney, S., Besancon, C., Butt, N., Maina, J. M., Possingham, H. P., von Bieberstein, K. R., & Venter, O. (2017). Changing trends and persisting biases in three decades of conservation science. *Global Ecology and Conservation*, *10*, 32-42.

Di Muri, C., Lawson Handley, L., Bean, C. W., Benucci, M., Harper, L. R., James, B., Li, J., Winfield, I. J., & Hänfling, B. (2022). Spatio-temporal monitoring of lake fish spawning activity using environmental DNA metabarcoding. *Environmental DNA*.

Diaz-Carballido, P. L., Mendoza-Gonzalez, G., Yanez-Arenas, C. A., & Chiappa-Carrara, X. (2022). Evaluation of Shifts in the Potential Future Distributions of Carcharhinid Sharks Under Different Climate Change Scenarios. *Frontiers in Marine Science*, *8*, 2039.

DiBattista, J. D., Reimer, J. D., Stat, M., Masucci, G. D., Biondi, P., De Brauwer, M., Wilkinson, S. P., Chariton, A. A., & Bunce, M. (2020). Environmental DNA can act as a biodiversity barometer of anthropogenic pressures in coastal ecosystems. *Scientific reports*, *10*(1), 1-15.

Distler, T., Schuetz, J. G., Velásquez-Tibatá, J., & Langham, G. M. (2015). Stacked species distribution models and macroecological models provide congruent projections of avian species richness under climate change. *Journal of Biogeography*, *42*(5), 976-988.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927-930.

Djurhuus, A., Closek, C. J., Kelly, R. P., Pitz, K. J., Michisaki, R. P., Starks, H. A., Walz, K. R., Andruszkiewicz, E. A., Olesin, E., & Hubbard, K. (2020). Environmental DNA reveals seasonal shifts and potential interactions in a marine community. *Nature communications*, *11*(1), 1-9.

Donaldson, M., Burnett, N., Braun, D., Suski, C., Hinch, S., Cooke, S., Kerr, J., & Hutchings, J. (2016). Taxonomic bias and international biodiversity conservation research. Facets, 1 (1), 105–113. In.

Drummond, J. A., Larson, E. R., Li, Y., Lodge, D. M., Gantz, C. A., Pfrender, M. E., Renshaw, M. A., Correa, A. M., & Egan, S. P. (2021). Diversity metrics are robust to differences in sampling location and depth for environmental DNA of plants in small temperate lakes. *Frontiers in Environmental Science*, *9*, 617924.

Dugal, L., Thomas, L., Jensen, M. R., Sigsgaard, E. E., Simpson, T., Jarman, S., Thomsen, P. F., & Meekan, M. (2022). Individual haplotyping of whale sharks from seawater environmental DNA. *Molecular ecology resources*, *22*(1), 56-65.

Dulvy, N. K., Fowler, S. L., Musick, J. A., Cavanagh, R. D., Kyne, P. M., Harrison, L. R., Carlson, J. K., Davidson, L. N., Fordham, S. V., & Francis, M. P. (2014). Extinction risk and conservation of the world's sharks and rays. *elife*, *3*, e00590.

Dulvy, N. K., Pacoureau, N., Rigby, C. L., Pollom, R. A., Jabado, R. W., Ebert, D. A., Finucci, B., Pollock, C. M., Cheok, J., & Derrick, D. H. (2021). Overfishing drives over one-third of all sharks and rays toward a global extinction crisis. *Current Biology*, *31*(21), 4773-4787. e4778.

Dunham, A., Dunham, J. S., Rubidge, E., Iacarella, J. C., & Metaxas, A. (2020). Contextualizing ecological performance: Rethinking monitoring in marine

protected areas. *Aquatic conservation: marine and freshwater ecosystems*, *30*(10), 2004-2011.

Dutton, P. H., Komoroske, L., Bejder, L., & Meekan, M. (2019). Integrating emerging technologies into marine megafauna conservation management. *Frontiers in Marine Science*, 693.

Eberhardt, L. L., Chapman, D. G., & Gilbert, J. R. (1979). A review of marine mammal census methods. *Wildlife Monographs*(63), 3-46.

Edgar, G. J., Bates, A. E., Bird, T. J., Jones, A. H., Kininmonth, S., Stuart-Smith, R. D., & Webb, T. J. (2016). New approaches to marine conservation through the scaling up of ecological data. *Annual Review of Marine Science*, *8*, 435-461.

Edgar, G. J., Stuart-Smith, R. D., Willis, T. J., Kininmonth, S., Baker, S. C., Banks, S., Barrett, N. S., Becerro, M. A., Bernard, A. T., & Berkhout, J. (2014). Global conservation outcomes depend on marine protected areas with five key features. *Nature*, *506*(7487), 216-220.

Edwards, M., & Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, *430*(7002), 881-884.

Ehler, C. N. (2021). Two decades of progress in Marine Spatial Planning. *Marine Policy*, *132*, 104134.

Ellis, J., Cruz-Martinez, A., Rackham, B., & Rogers, S. (2005). The distribution of chondrichthyan fishes around the British Isles and implications for conservation. *Journal of Northwest Atlantic fishery science*, *35*(195-213), 113.

Emeis, K.-C., van Beusekom, J., Callies, U., Ebinghaus, R., Kannen, A., Kraus, G., Kröncke, I., Lenhart, H., Lorkowski, I., & Matthias, V. (2015). The North Sea—A shelf sea in the Anthropocene. *Journal of Marine Systems*, *141*, 18-33.

Engelhard, G. H., Peck, M. A., Rindorf, A., C. Smout, S., van Deurs, M., Raab, K., Andersen, K. H., Garthe, S., Lauerburg, R. A., & Scott, F. (2014). Forage fish, their fisheries, and their predators: who drives whom? *ICES Journal of Marine Science*, *71*(1), 90-104.

Estes, J. A., Heithaus, M., McCauley, D. J., Rasher, D. B., & Worm, B. (2016). Megafaunal impacts on structure and function of ocean ecosystems. *Annual Review of Environment and Resources*, *41*, 83-116.

European Commission. (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). *Official Journal of the European Union*, *50*, 1-14.

Evans, P., & Waggitt, J. (2020). Impacts of climate change on marine mammals, relevant to the coastal and marine environment around the UK.

Fauchald, P., Skov, H., Skern-Mauritzen, M., Johns, D., & Tveraa, T. (2011). Wasp-waist interactions in the North Sea ecosystem. *PloS one*, *6*(7), e22729.

Faust, K., & Raes, J. (2016). CoNet app: inference of biological association networks using Cytoscape. *F1000Research*, *5*, 1519.

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., & Huttenhower, C. (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS computational biology*, *8*(7), e1002606.

Ferretti, F., Myers, R. A., Serena, F., & Lotze, H. K. (2008). Loss of large predatory sharks from the Mediterranean Sea. *Conservation biology*, *22*(4), 952-964.

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology letters*, *4*(4), 423-425.

Fiori, L., Doshi, A., Martinez, E., Orams, M. B., & Bollard-Breen, B. (2017). The use of unmanned aerial systems in marine mammal research. *Remote Sensing*, *9*(6), 543.

Foden, W. B., Young, B. E., Akçakaya, H. R., Garcia, R. A., Hoffmann, A. A., Stein, B. A., Thomas, C. D., Wheatley, C. J., Bickford, D., & Carr, J. A. (2019). Climate change vulnerability assessment of species. *Wiley interdisciplinary reviews: climate change*, *10*(1), e551.

Foley, M. M., Halpern, B. S., Micheli, F., Armsby, M. H., Caldwell, M. R., Crain, C. M., Prahler, E., Rohr, N., Sivas, D., & Beck, M. W. (2010). Guiding ecological principles for marine spatial planning. *Marine Policy*, *34*(5), 955-966.

Foote, A. D., Thomsen, P. F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L. A., Salling, A. B., Galatius, A., Orlando, L., & Gilbert, M. T. P. (2012). Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PloS one*, *7*(8), e41781.

Ford, B. M., & Roberts, J. D. (2019). Evolutionary histories impart structure into marine fish heterospecific co-occurrence networks. *Global Ecology and Biogeography*, *28*(9), 1310-1324.

Fraija-Fernández, N., Bouquieaux, M. C., Rey, A., Mendibil, I., Cotano, U., Irigoien, X., Santos, M., & Rodríguez-Ezpeleta, N. (2020). Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution*, *10*(14), 7560-7584.

Frederiksen, M., Elston, D. A., Edwards, M., Mann, A. D., & Wanless, S. (2011). Mechanisms of long-term decline in size of lesser sandeels in the North Sea explored using a growth and phenology model. *Marine Ecology Progress Series*, *432*, 137-147.

Freilich, M. A., Wieters, E., Broitman, B. R., Marquet, P. A., & Navarrete, S. A. (2018). Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? *Ecology*, *99*(3), 690-699.

Fuentes-Pardo, A. P., & Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular ecology*, *26*(20), 5369-5406.

Garzke, J., Ismar, S. M., & Sommer, U. (2015). Climate change affects low trophic level marine consumers: warming decreases copepod size and abundance. *Oecologia*, *177*, 849-860.

Gavish, Y., Marsh, C. J., Kuemmerlen, M., Stoll, S., Haase, P., & Kunin, W. E. (2017). Accounting for biotic interactions through alpha-diversity constraints in stacked species distribution models. *Methods in Ecology and Evolution*, *8*(9), 1092-1102.

GEBCO Compilation Group. (2020). GEBCO 2020 grid. (doi:10.5285/e0f0bb80-ab44-2739-e053-6c86abc0289c). In: British Oceanographic Data Centre, National Oceanography Centre, NERC ….

Giakoumi, S., Katsanevakis, S., Albano, P. G., Azzurro, E., Cardoso, A. C., Cebrian, E., Deidun, A., Edelist, D., Francour, P., & Jimenez, C. (2019). Management priorities for marine invasive species. *Science of The Total Environment*, *688*, 976-982.

Giakoumi, S., McGowan, J., Mills, M., Beger, M., Bustamante, R. H., Charles, A., Christie, P., Fox, M., Garcia-Borboroglu, P., & Gelcich, S. (2018). Revisiting "success" and "failure" of marine protected areas: a conservation scientist perspective. *Frontiers in Marine Science*, 223.

Gill, D. A., Mascia, M. B., Ahmadia, G. N., Glew, L., Lester, S. E., Barnes, M., Craigie, I., Darling, E. S., Free, C. M., & Geldmann, J. (2017). Capacity shortfalls hinder the performance of marine protected areas globally. *Nature*, *543*(7647), 665-669.

Giménez, J., Cardador, L., Mazor, T., Kark, S., Bellido, J. M., Coll, M., & Navarro, J. (2020). Marine protected areas for demersal elasmobranchs in highly exploited Mediterranean ecosystems. *Marine Environmental Research*, *160*, 105033.

Giovos, I., Katsada, D., Spyridopoulou, R. N. A., Poursanidis, D., Doxa, A., Katsanevakis, S., Kleitou, P., Oikonomou, V., Minasidis, V., & Ozturk, A. A. (2022). Strengthening Angel Shark Conservation in the Northeastern Mediterranean Sea. *Journal of Marine Science and Engineering*, *10*(2), 269.

Goberna, M., & Verdú, M. (2022). Cautionary notes on the use of co-occurrence networks in soil ecology. *Soil Biology and Biochemistry*, *166*, 108534.

Gold, Z., Sprague, J., Kushner, D. J., Zerecero Marin, E., & Barber, P. H. (2021). eDNA metabarcoding as a biomonitoring tool for marine protected areas. *PloS one*, *16*(2), e0238557.

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*, 1-19.

Grady, J. M., Maitner, B. S., Winter, A. S., Kaschner, K., Tittensor, D. P., Record, S., Smith, F. A., Wilson, A. M., Dell, A. I., & Zarnetske, P. L. (2019). Metabolic asymmetry and the global diversity of marine predators. *Science*, *363*(6425), eaat4220.

Grassle, J. F. (2000). The Ocean Biogeographic Information System (OBIS): an on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multidimensional geographic context. *Oceanography*, *13*(3), 5-7.

Greenstreet, S. P., McMillan, J. A., & Armstrong, E. (1998). Seasonal variation in the importance of pelagic fish in the diet of piscivorous fish in the Moray Firth, NE Scotland: a response to variation in prey abundance? *ICES Journal of Marine Science*, *55*(1), 121-133.

Grémillet, D., Chevallier, D., & Guinet, C. (2022). Big data approaches to the spatial ecology and conservation of marine megafauna. *ICES Journal of Marine Science*, *79*(4), 975-986.

Grenié, M., Violle, C., & Munoz, F. (2020). Is prediction of species richness from stacked species distribution models biased by habitat saturation? *Ecological Indicators*, *111*, 105970.

Griffin, L. (2008). The North Sea fisheries crisis and good governance. *Geography Compass*, *2*(2), 452-475.

Grorud-Colvert, K., Sullivan-Stack, J., Roberts, C., Constant, V., Horta e Costa, B., Pike, E. P., Kingston, N., Laffoley, D., Sala, E., & Claudet, J. (2021). The MPA Guide: A framework to achieve global goals for the ocean. *Science*, *373*(6560), eabf0861.

Guisan, A., & Rahbek, C. (2011). SESAM–a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. In (Vol. 38, pp. 1433-1444): Wiley Online Library.

Halpern, B. S., Frazier, M., Afflerbach, J., Lowndes, J. S., Micheli, F., O'Hara, C., Scarborough, C., & Selkoe, K. A. (2019). Recent pace of change in human impact on the world's ocean. *Scientific reports*, *9*(1), 11609.

Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., Lowndes, J. S., Rockwood, R. C., Selig, E. R., & Selkoe, K. A. (2015). Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature communications*, *6*(1), 1-7.

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., d'Agrosa, C., Bruno, J. F., Casey, K. S., Ebert, C., & Fox, H. E. (2008). A global map of human impact on marine ecosystems. *Science*, *319*(5865), 948-952.

Hammerschlag, N., Schmitz, O. J., Flecker, A. S., Lafferty, K. D., Sih, A., Atwood, T. B., Gallagher, A. J., Irschick, D. J., Skubel, R., & Cooke, S. J. (2019). Ecosystem function and services of aquatic predators in the Anthropocene. *Trends in ecology & evolution*, *34*(4), 369-383.

Hammond, P. S., Francis, T. B., Heinemann, D., Long, K. J., Moore, J. E., Punt, A. E., Reeves, R. R., Sepúlveda, M., Sigurðsson, G. M., & Siple, M. C. (2021).

Estimating the abundance of marine mammal populations. *Frontiers in Marine Science*, *8*, 1316.

Hansen, B. K., Bekkevold, D., Clausen, L. W., & Nielsen, E. E. (2018). The sceptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, *19*(5), 751-768.

Harrell Jr, F. E. (2022). Hmisc: Harrell Miscellaneous. R package version 4.7-1. https://CRAN.R-project.org/package=Hmisc.

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B*, *286*(1915), 20191409.

Hazen, E. L., Abrahms, B., Brodie, S., Carroll, G., Jacox, M. G., Savoca, M. S., Scales, K. L., Sydeman, W. J., & Bograd, S. J. (2019). Marine top predators as climate and ecosystem sentinels. *Frontiers in Ecology and the Environment*, *17*(10), 565-574.

Hazen, E. L., Scales, K. L., Maxwell, S. M., Briscoe, D. K., Welch, H., Bograd, S. J., Bailey, H., Benson, S. R., Eguchi, T., & Dewar, H. (2018). A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. *Science Advances*, *4*(5), eaar3001.

Hendricks, A., Mackie, C. M., Luy, E., Sonnichsen, C., Smith, J., Grundke, I., Tavasoli, M., Furlong, A., Beiko, R. G., & LaRoche, J. (2023). Compact and automated eDNA sampler for in situ monitoring of marine environments. *Scientific reports*, *13*(1), 5210.

Henriksen, O., Rindorf, A., Brooks, M. E., Lindegren, M., & van Deurs, M. (2021). Temperature and body size affect recruitment and survival of sandeel across the North Sea. *ICES Journal of Marine Science*, *78*(4), 1409-1420.

Heupel, M. R., Munroe, S. E., Lédée, E. J., Chin, A., & Simpfendorfer, C. A. (2019). Interspecific interactions, movement patterns and habitat use in a diverse coastal shark assemblage. *Marine Biology*, *166*(6), 1-17.

Heymans, J. J., Coll, M., Link, J. S., Mackinson, S., Steenbeek, J., Walters, C., & Christensen, V. (2016). Best practice in Ecopath with Ecosim food-web models for ecosystem-based management. *Ecological modelling*, *331*, 173-184.

Hijmans, R. J. (2021). geosphere: Spherical Trigonometry. R package version 1.5-14. https://CRAN.R-project.org/package=geosphere.

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package 'dismo'. *Circles*, *9*(1), 1-68.

Hirano, H., & Takemoto, K. (2019). Difficulty in inferring microbial community structure based on co-occurrence network approaches. *BMC bioinformatics*, *20*(1), 1-14.

Hobday, A., & Hartmann, K. (2006). Near real-time spatial management based on habitat predictions for a longline bycatch species. *Fisheries Management and Ecology*, *13*(6), 365-380.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular ecology*, *30*(1), 62-82.

Holmes, R., Bulat, J., Henni, P., Holt, J., James, C., Kenyon, N., Leslie, A., Long, D., Morri, C., & Musson, R. (2004). DTI Strategic Environmental Assessment Area 5 (SEA5): Seabed and superficial geology and processes.

Holt, J., Hughes, S., Hopkins, J., Wakelin, S. L., Holliday, N. P., Dye, S., González-Pola, C., Hjøllo, S. S., Mork, K. A., & Nolan, G. (2012). Multi-decadal variability and trends in the temperature of the northwest European continental shelf: a model-data synthesis. *Progress in Oceanography*, *106*, 96-117.

Hooker, S. K., De Soto, N. A., Baird, R. W., Carroll, E. L., Claridge, D., Feyrer, L., Miller, P. J., Onoufriou, A., Schorr, G., & Siegal, E. (2019). Future directions in research on beaked whales. *Frontiers in Marine Science*, 514.

Horton, T. W., Block, B. A., Davies, R., Hawkes, L. A., Jones, D., Jones, H., Leeves, K., Maoiléidigh, N. Ó., Righton, D., & van der Kooij, J. (2021). Evidence of increased occurrence of Atlantic bluefin tuna in territorial waters of the United Kingdom and Ireland. *ICES Journal of Marine Science*, *78*(5), 1672-1683.

Høyer, J. L., & Karagali, I. (2016). Sea surface temperature climate data record for the North Sea and Baltic Sea. *Journal of Climate*, *29*(7), 2529-2541.

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, *44*(9), 1259-1269.

Hunter, M. E., Hoban, S. M., Bruford, M. W., Segelbacher, G., & Bernatchez, L. (2018). Next-generation conservation genetics and biodiversity monitoring. *Evolutionary applications*, *11*(7), 1029-1034.

Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., Harcourt, R. G., Holland, K. N., Iverson, S. J., & Kocik, J. F. (2015). Aquatic animal telemetry: a panoramic window into the underwater world. *Science*, *348*(6240), 1255642.

Hutchings, J. A., Minto, C., Ricard, D., Baum, J. K., & Jensen, O. P. (2010). Trends in the abundance of marine fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(8), 1205-1210.

Inman, R., Franklin, J., Esque, T., & Nussear, K. (2021). Comparing sample bias correction methods for species distribution modeling using virtual species. Ecosphere 12, e03422. In.

Ip, Y. C. A., Chang, J. J. M., Lim, K. K., Jaafar, Z., Wainwright, B. J., & Huang, D. (2021). Seeing through sedimented waters: environmental DNA reduces the phantom diversity of sharks and rays in turbid marine habitats. *BMC Ecology and Evolution*, *21*(1), 1-14.

IUCN-MMPATF. 'Alborán Sea IMMA', Marine Mammal Protected Areas Task Force (MMPATF) Website. Available at: https://www.marinemammalhabitat.org/portfolio-item/alboran-sea/. Accessed October 2022.

Jamy, M., Foster, R., Barbera, P., Czech, L., Kozlov, A., Stamatakis, A., Bending, G., Hilton, S., Bass, D., & Burki, F. (2020). Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. *Molecular ecology resources*, *20*(2), 429-443.

Jennings, S., & Blanchard, J. L. (2004). Fish abundance with no fishing: predictions based on macroecological theory. *Journal of Animal Ecology*, 632-642.

Jensen, M. R., Høgslund, S., Knudsen, S. W., Nielsen, J., Møller, P. R., Rysgaard, S., & Thomsen, P. F. (2023). Distinct latitudinal community patterns of Arctic marine vertebrates along the East Greenlandic coast detected by environmental DNA. *Diversity and Distributions*, *29*(2), 316-334.

Jerde, C. L., Wilson, E. A., & Dressler, T. L. (2019). Measuring global fish species richness with eDNA metabarcoding. In: Wiley Online Library.

Jiménez, M. F. S.-S., Cerqueda-García, D., Montero-Muñoz, J. L., Aguirre-Macedo, M. L., & García-Maldonado, J. Q. (2018). Assessment of the bacterial community structure in shallow and deep sediments of the Perdido Fold Belt region in the Gulf of Mexico. *PeerJ*, *6*, e5583.

Johnson, C. R., Banks, S. C., Barrett, N. S., Cazassus, F., Dunstan, P. K., Edgar, G. J., Frusher, S. D., Gardner, C., Haddon, M., & Helidoniotis, F. (2011). Climate change cascades: Shifts in oceanography, species' ranges and subtidal marine community dynamics in eastern Tasmania. *Journal of Experimental Marine Biology and Ecology*, *400*(1-2), 17-32.

Jones, P. J., Lieberknecht, L. M., & Qiu, W. (2016). Marine spatial planning in reality: Introduction to case studies and discussion of findings. *Marine Policy*, *71*, 256-264.

Jorgensen, S. J., Micheli, F., White, T. D., Van Houtan, K. S., Alfaro-Shigueto, J., Andrzejaczek, S., Arnoldi, N. S., Baum, J. K., Block, B., & Britten, G. L. (2022). Emergent research and priorities for shark and ray conservation. *Endangered Species Research*, *47*, 171-203.

Juan-Jordá, M. J., Mosqueira, I., Cooper, A. B., Freire, J., & Dulvy, N. K. (2011). Global population trajectories of tunas and their relatives. *Proceedings of the National Academy of Sciences*, *108*(51), 20650-20655.

Juhel, J. B., Marques, V., Fernández, A. P., Borrero-Pérez, G. H., Martinezguerra, M. M., Valentini, A., Dejean, T., Manel, S., Loiseau, N., & Velez, L. (2021). Detection of the elusive Dwarf sperm whale (Kogia sima) using environmental DNA at Malpelo island (Eastern Pacific, Colombia). *Ecology and Evolution*, *11*(7), 2956.

Kadmon, R., Farber, O., & Danin, A. (2004). Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, *14*(2), 401-413.

Kardos, M., Armstrong, E. E., Fitzpatrick, S. W., Hauser, S., Hedrick, P. W., Miller, J. M., Tallmon, D. A., & Funk, W. C. (2021). The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences*, *118*(48), e2104642118.

Kavanaugh, M. T., Bell, T., Catlett, D., Cimino, M. A., Doney, S. C., Klajbor, W., Messié, M., Montes, E., Muller-Karger, F. E., & Otis, D. (2021). Satellite remote sensing and the marine biodiversity observation network. *Oceanography*, *34*(2), 62-79.

Kelly, R. P., Gallego, R., & Jacobs-Palmer, E. (2018). The effect of tides on nearshore environmental DNA. *PeerJ*, *6*, e4521.

Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific reports*, *9*(1), 1-14.

Kershaw, F., McClintock, W., Andrews, K. R., Riet-Sapriza, F. G., Caballero, S., Tetley, M. J., Notarbartolo di Sciara, G., Hoyt, E., Goldberg, G., & Chou, E. (2021). Geospatial genetics: Integrating genetics into marine protection and spatial planning. *Aquatic conservation: marine and freshwater ecosystems*, *31*(9), 2440-2458.

Kiszka, J. J., Mourier, J., Gastrich, K., & Heithaus, M. R. (2016). Using unmanned aerial vehicles (UAVs) to investigate shark and ray densities in a shallow coral lagoon. *Marine Ecology Progress Series*, *560*, 237-242.

Klein, C. J., Brown, C. J., Halpern, B. S., Segan, D. B., McGowan, J., Beger, M., & Watson, J. E. (2015). Shortfalls in the global protected area network at representing marine biodiversity. *Scientific reports*, *5*(1), 17539.

Klippel, S., Amaral, S., & Vinhas, L. (2016). Development and evaluation of species distribution models for five endangered elasmobranchs in southwestern Atlantic. *Hydrobiologia*, *779*, 11-33.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A. K., & Augeri, D. M. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, *19*(11), 1366-1379.

Kujala, H., Moilanen, A., & Gordon, A. (2018). Spatial characteristics of species distributions as drivers in conservation prioritization. *Methods in Ecology and Evolution*, *9*(4), 1121-1132.

Kürten, B., Painting, S. J., Struck, U., Polunin, N. V., & Middelburg, J. J. (2013). Tracking seasonal changes in North Sea zooplankton trophic dynamics using stable isotopes. *Biogeochemistry*, *113*, 167-187.

Lambert, C., Laran, S., David, L., Dorémus, G., Pettex, E., Van Canneyt, O., & Ridoux, V. (2017). How does ocean seasonality drive habitat preferences of highly mobile top predators? Part I: the north-western Mediterranean Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, *141*, 115-132.

Lascelles, B., Notarbartolo Di Sciara, G., Agardy, T., Cuttelod, A., Eckert, S., Glowka, L., Hoyt, E., Llewellyn, F., Louzao, M., & Ridoux, V. (2014). Migratory marine species: their status, threats and conservation management needs. *Aquatic conservation: marine and freshwater ecosystems*, *24*(S2), 111-127.

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*, 271-280.

Levin, N., Coll, M., Fraschetti, S., Gal, G., Giakoumi, S., Göke, C., Heymans, J. J., Katsanevakis, S., Mazor, T., & Öztürk, B. (2014). Biodiversity data requirements for systematic conservation planning in the Mediterranean Sea. *Marine Ecology Progress Series*, *508*, 261-281.

Li, S., & Jay, S. (2020). Transboundary marine spatial planning across Europe: Trends and priorities in nearly two decades of project work. *Marine Policy*, *118*, 104012.

Lindegren, M., Van Deurs, M., MacKenzie, B. R., Worsoe Clausen, L., Christensen, A., & Rindorf, A. (2018). Productivity and recovery of forage fish under climate change and fishing: North Sea sandeel as a case study. *Fisheries Oceanography*, *27*(3), 212-221.

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*(3), 385-393.

Long, J. M., & Fisher, W. L. (2006). Analysis of environmental variation in a Great Plains reservoir using principal components analysis and geographic information systems. *Lake and Reservoir management*, *22*(2), 132-140.

Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, *11*(10), 995-1003.

Lynam, C. P., Llope, M., Möllmann, C., Helaouët, P., Bayliss-Brown, G. A., & Stenseth, N. C. (2017). Interaction between top-down and bottom-up control in marine food webs. *Proceedings of the National Academy of Sciences*, *114*(8), 1952-1957.

MacDonald, A., Speirs, D. C., Greenstreet, S. P., Boulcott, P., & Heath, M. R. (2019). Trends in sandeel growth and abundance off the east coast of Scotland. *Frontiers in Marine Science*, *6*, 201.

MacKenzie, B. R., Mosegaard, H., & Rosenberg, A. A. (2009). Impending collapse of bluefin tuna in the northeast Atlantic and Mediterranean. *Conservation Letters*, *2*(1), 26-35.

MacKeracher, T., Diedrich, A., & Simpfendorfer, C. A. (2019). Sharks, rays and marine protected areas: A critical evaluation of current perspectives. *Fish and Fisheries*, *20*(2), 255-267.

MacLeod, C. D., Pierce, G. J., & Santos, M. B. (2004). Geographic and temporal variations in strandings of beaked whales (Ziphiidae) on the coasts of the UK and the Republic of Ireland from 1800-2002. *Journal of Cetacean Research and Management*, *6*(1), 79-86.

MacLeod, C. D., Weir, C. R., Santos, M. B., & Dunn, T. E. (2008). Temperature-based summer habitat partitioning between white-beaked and common dolphins around the United Kingdom and Republic of Ireland. *Journal of the Marine Biological Association of the United Kingdom*, *88*(6), 1193-1198.

Maestro, M., Pérez-Cayeiro, M. L., Chica-Ruiz, J. A., & Reyes, H. (2019). Marine protected areas in the 21st century: Current situation and trends. *Ocean & Coastal Management*, *171*, 28-36.

Magera, A. M., Mills Flemming, J. E., Kaschner, K., Christensen, L. B., & Lotze, H. K. (2013). Recovery trends in marine mammal populations. *PloS one*, *8*(10), e77908.

Maiello, G., Talarico, L., Brodie, C., Carpentieri, P., Sbrana, A., Shum, P., Mariani, S., & Russo, T. (2023). Net gain: Low-cost, trawl-associated eDNA samplers upscale ecological assessment of marine demersal communities. *Environmental DNA*.

Maina, J., Venus, V., McClanahan, T. R., & Ateweberhan, M. (2008). Modelling susceptibility of coral reefs to environmental stress using remote sensing data and GIS models. *Ecological modelling*, *212*(3-4), 180-199.

Malcolm, I., Godfrey, J., & Youngson, A. (2010). *Review of migratory routes and behaviour of Atlantic salmon, sea trout and European eel in Scotland's coastal environment: implications for the development of marine renewables*. Marine Scotland Science.

Mancusi, C., Baino, R., Fortuna, C., De Sola, L. G., Morey, G., Bradai, M. N., Kallianotis, A., Soldo, A., Hemida, F., & Saad, A. A. (2020). MEDLEM database, a data collection on large Elasmobranchs in the Mediterranean and Black seas. *Mediterranean Marine Science*, *21*(2), 276-288.

Mannocci, L., Boustany, A. M., Roberts, J. J., Palacios, D. M., Dunn, D. C., Halpin, P. N., Viehman, S., Moxley, J., Cleary, J., & Bailey, H. (2017). Temporal resolutions in species distribution models of highly mobile marine animals: Recommendations for ecologists and managers. *Diversity and Distributions*, *23*(10), 1098-1109.

Mannocci, L., Roberts, J. J., Halpin, P. N., Authier, M., Boisseau, O., Bradai, M. N., Cañadas, A., Chicote, C., David, L., & Di-Méglio, N. (2018). Assessing cetacean surveys throughout the Mediterranean Sea: A gap analysis in environmental space. *Scientific reports*, *8*(1), 1-14.

Mannocci, L., Roberts, J. J., Pedersen, E. J., & Halpin, P. N. (2020). Geographical differences in habitat relationships of cetaceans across an ocean basin. *Ecography*, *43*(8), 1250-1259.

Mariani, S., Baillie, C., Colosimo, G., & Riesgo, A. (2019). Sponges as natural environmental DNA samplers. *Current Biology*, *29*(11), R401-R402.

Mariani, S., Fernandez, C., Baillie, C., Magalon, H., & Jaquemet, S. (2021). Shark and ray diversity, abundance and temporal variation around an Indian Ocean Island, inferred by eDNA metabarcoding. *Conservation Science and Practice*, *3*(6), e407.

Martinez Arbizu, P. (2017). PairwiseAdonis: pairwise multilevel comparison using Adonis. R Package. In.

Maxwell, S. M., Gjerde, K. M., Conners, M. G., & Crowder, L. B. (2020). Mobile protected areas for biodiversity on the high seas. *Science*, *367*(6475), 252-254.

Mazor, T., Beger, M., McGowan, J., Possingham, H. P., & Kark, S. (2016). The value of migration information for conservation prioritization of sea turtles in the Mediterranean. *Global Ecology and Biogeography*, *25*(5), 540-552.

Mazor, T., Possingham, H. P., & Kark, S. (2013). Collaboration among countries in marine conservation can achieve substantial efficiencies. *Diversity and Distributions*, *19*(11), 1380-1393.

Mbaru, E. K., Graham, N. A., McClanahan, T. R., & Cinner, J. E. (2020). Functional traits illuminate the selective impacts of different fishing gears on coral reefs. *Journal of Applied Ecology*, *57*(2), 241-252.

McDonald-Madden, E., Sabbadin, R., Game, E. T., Baxter, P. W., Chades, I., & Possingham, H. P. (2016). Using food-web theory to conserve ecosystems. *Nature communications*, *7*(1), 10245.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276-282.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4), e61217.

MedPAN, & SPA/RAC. (2022). Marine protected areas in the Mediterranean https://www.mapamed.org/

Meynard, C. N., Leroy, B., & Kaplan, D. M. (2019). Testing methods in species distribution modelling using virtual species: what have we learnt and what are we missing? *Ecography*, *42*(12), 2021-2036.

Micheli, F., Halpern, B. S., Walbridge, S., Ciriaco, S., Ferretti, F., Fraschetti, S., Lewison, R., Nykjaer, L., & Rosenberg, A. A. (2013). Cumulative human impacts on Mediterranean and Black Sea marine ecosystems: assessing current pressures and opportunities. *PloS one*, *8*(12), e79889.

Mickle, M. F., & Higgs, D. M. (2022). Towards a new understanding of elasmobranch hearing. *Marine Biology*, *169*(1), 1-13.

Miller, J. A. (2014). Virtual species distribution models: using simulated data to evaluate aspects of model performance. *Progress in Physical Geography*, *38*(1), 117-128.

Minnett, P., Alvera-Azcárate, A., Chin, T., Corlett, G., Gentemann, C., Karagali, I., Li, X., Marsouin, A., Marullo, S., & Maturi, E. (2019). Half a century of satellite remote sensing of sea-surface temperature. *Remote Sensing of Environment*, *233*, 111366.

Moore, A. B. (2023). A ray of hope? The re-appearance of Irish Sea skate decades after local extinction. *Journal of fish biology*.

Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from proxies. *Trends in ecology & evolution*, *30*(6), 347-356.

Mortelmans, J., Aubert, A., Reubens, J., Otero, V., Deneudt, K., & Mees, J. (2021). Copepods (Crustacea: Copepoda) in the Belgian part of the North Sea: Trends, dynamics and anomalies. *Journal of Marine Systems*, *220*, 103558.

Moudrý, V., & Devillers, R. (2020). Quality and usability challenges of global marine biodiversity databases: An example for marine mammal data. *Ecological Informatics*, *56*, 101051.

Moussy, C., Burfield, I. J., Stephenson, P., Newton, A. F., Butchart, S. H., Sutherland, W. J., Gregory, R. D., McRae, L., Bubb, P., & Roesler, I. (2022). A quantitative global review of species population monitoring. *Conservation biology*, *36*(1), e13721.

Myers, R. A., & Worm, B. (2003). Rapid worldwide depletion of predatory fish communities. *Nature*, *423*(6937), 280-283.

Myers, R. A., & Worm, B. (2005). Extinction, survival or recovery of large predatory fishes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1453), 13-20.

NatureScot. (2020). Conservation and Management Advice: Southern Trench MPA.

Nelms, S. E., Alfaro-Shigueto, J., Arnould, J. P., Avila, I. C., Nash, S. B., Campbell, E., Carter, M. I., Collins, T., Currey, R. J., & Domit, C. (2021). Marine mammal conservation: over the horizon. *Endangered Species Research*, *44*, 291-325.

Nicholson, A., McIsaac, D., MacDonald, C., Gec, P., Mason, B. E., Rein, W., Wrobel, J., de Boer, M., Milián-García, Y., & Hanner, R. H. (2020). An analysis of metadata reporting in freshwater environmental DNA research calls for the development of best practice guidelines. *Environmental DNA*, *2*(3), 343-349.

Nielsen, E. S., Hanson, J. O., Carvalho, S. B., Beger, M., Henriques, R., Kershaw, F., & von der Heyden, S. (2022). Molecular ecology meets systematic conservation planning. *Trends in ecology & evolution*.

Notarbartolo-di-Sciara, G., Agardy, T., Hyrenbach, D., Scovazzi, T., & Van Klaveren, P. (2008). The Pelagos sanctuary for Mediterranean marine mammals. *Aquatic conservation: marine and freshwater ecosystems*, *18*(4), 367-391.

Notarbartolo Di Sciara, G. (2016). Marine mammals in the Mediterranean Sea: An overview. *Advances in Marine Biology*, *75*, 1-36.

Notarbartolo di Sciara, G., & Hoyt, E. (2020). Healing the wounds of marine mammals by protecting their habitat. *Ethics in Science and Environmental Politics*, *20*, 15-23.

Notarbartolo di Sciara, G., Hoyt, E., Reeves, R., Ardron, J., Marsh, H., Vongraven, D., & Barr, B. (2016). Place-based approaches to marine mammal conservation. *Aquatic conservation: marine and freshwater ecosystems*, *26*, 85-100.

O'Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhouri, J. F., Lowell, N. C., & Williams, G. D. (2017). Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ*, *5*, e3044.

O'Hara, C. C., Frazier, M., & Halpern, B. S. (2021). At-risk marine biodiversity faces extensive, expanding, and intensifying human impacts. *Science*, *372*(6537), 84-87.

Österblom, H., Olsson, O., Blenckner, T., & Furness, R. W. (2008). Junk-food in marine ecosystems. *Oikos*, *117*(7), 967-977.

Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. K., & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in genetics*, *26*(4), 177-187.

Paquette, A., & Hargreaves, A. L. (2021). Biotic interactions are more often important at species' warm versus cool range edges. *Ecology Letters*, *24*(11), 2427-2438.

Parsons, K. M., Everett, M., Dahlheim, M., & Park, L. (2018). Water, water everywhere: Environmental DNA can unlock population structure in elusive marine species. *Royal Society open science*, *5*(8), 180537.

Pascher, K., Švara, V., & Jungmeier, M. (2022). Environmental DNA-based methods in biodiversity monitoring of protected areas: Application range, limitations, and needs. *Diversity*, *14*(6), 463.

Pawlowski, J., Apothéloz-Perret-Gentil, L., & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular ecology*, *29*(22), 4258-4264.

Pehlke, C., & Bartsch, I. (2008). Changes in depth distribution and biomass of sublittoral seaweeds at Helgoland (North Sea) between 1970 and 2005. *Climate research*, *37*(2-3), 135-147.

Pendleton, D. E., Holmes, E. E., Redfern, J., & Zhang, J. (2020). Using modelled prey to predict the distribution of a highly mobile marine mammal. *Diversity and Distributions*, *26*(11), 1612-1626.

Perry, A. L., Low, P. J., Ellis, J. R., & Reynolds, J. D. (2005). Climate change and distribution shifts in marine fishes. *Science*, *308*(5730), 1912-1915.

Pethybridge, H., Bodin, N., Arsenault-Pernet, E.-J., Bourdeix, J.-H., Brisset, B., Bigot, J.-L., Roos, D., & Peter, M. (2014). Temporal and inter-specific variations in forage fish feeding conditions in the NW Mediterranean: lipid content and fatty acid compositional changes. *Marine Ecology Progress Series*, *512*, 39-54.

Petitgas, P., Rijnsdorp, A. D., Dickey-Collas, M., Engelhard, G. H., Peck, M. A., Pinnegar, J. K., Drinkwater, K., Huret, M., & Nash, R. D. (2013). Impacts of climate change on the complex life cycles of fish. *Fisheries Oceanography*, *22*(2), 121-139.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Pierce, G. J., Santos, M., Reid, R., Patterson, I., & Ross, H. (2004). Diet of minke whales Balaenoptera acutorostrata in Scottish (UK) waters with notes on

strandings of this species in Scotland 1992–2002. *Journal of the Marine Biological Association of the United Kingdom*, *84*(6), 1241-1244.

Pimiento, C., Leprieur, F., Silvestro, D., Lefcheck, J., Albouy, C., Rasher, D., Davis, M., Svenning, J.-C., & Griffin, J. (2020). Functional diversity of marine megafauna in the Anthropocene. *Science Advances*, *6*(16), eaay7650.

Pinfield, R., Dillane, E., Runge, A. K. W., Evans, A., Mirimin, L., Niemann, J., Reed, T. E., Reid, D. G., Rogan, E., & Samarra, F. I. (2019). False-negative detections from environmental DNA collected in the presence of large numbers of killer whales (Orcinus orca). *Environmental DNA*, *1*(4), 316-328.

Pirotta, V., Grech, A., Jonsen, I. D., Laurance, W. F., & Harcourt, R. G. (2019). Consequences of global shipping traffic for marine giants. *Frontiers in Ecology and the Environment*, *17*(1), 39-47.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C. F., Vittoz, P., & Guisan, A. (2013). The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography*, *22*(1), 52-63.

Qiu, W., & Jones, P. J. (2013). The emerging policy landscape for marine spatial planning in Europe. *Marine Policy*, *39*, 182-190.

Ramírez, F., Sbragaglia, V., Soacha, K., Coll, M., & Piera, J. (2022). Challenges for Marine Ecological Assessments: Completeness of Findable, Accessible, Interoperable, and Reusable Biodiversity Data in European Seas. *Frontiers in Marine Science*, *8*, 2038.

Randin, C. F., Ashcroft, M. B., Bolliger, J., Cavender-Bares, J., Coops, N. C., Dullinger, S., Dirnböck, T., Eckert, S., Ellis, E., & Fernández, N. (2020). Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sensing of Environment*, *239*, 111626.

Ransijn, J., Booth, C. G., & Coffen-Smout, S. (2019). *A calorific map of harbour porpoise prey in the North Sea*. JNCC.

Ratcliffe, F. C., Uren Webster, T. M., Garcia de Leaniz, C., & Consuegra, S. (2021). A drop in the ocean: Monitoring fish communities in spawning areas using environmental DNA. *Environmental DNA*, *3*(1), 43-54.

Raupach, M. J., Barco, A., Steinke, D., Beermann, J., Laakmann, S., Mohrbeck, I., Neumann, H., Kihara, T. C., Pointner, K., & Radulovici, A. (2015). The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. *PloS one*, *10*(9), e0139421.

Rees, A. F., Avens, L., Ballorain, K., Bevan, E., Broderick, A. C., Carthy, R. R., Christianen, M. J., Duclos, G., Heithaus, M. R., & Johnston, D. W. (2018). The potential of unmanned aerial systems for sea turtle research and conservation: a review and future directions. *Endangered Species Research*, *35*, 81-100.

Régnier, T., Gibb, F., & Wright, P. (2019). Understanding temperature effects on recruitment in the context of trophic mismatch. *Scientific reports*, *9*(1), 1-13.

Reid, P. C., Colebrook, J., Matthews, J., Aiken, J., & Team, C. P. R. (2003). The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. *Progress in Oceanography*, *58*(2-4), 117-173.

Reid, P. C., Edwards, M., Beaugrand, G., Skogen, M., & Stevens, D. (2003). Periodic changes in the zooplankton of the North Sea during the twentieth century linked to oceanic inflow. *Fisheries Oceanography*, *12*(4-5), 260-269.

Riesenfeld, C. S., Schloss, P. D., & Handelsman, J. (2004). Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.*, *38*, 525-552.

Rilov, G., Fraschetti, S., Gissi, E., Pipitone, C., Badalamenti, F., Tamburello, L., Menini, E., Goriup, P., Mazaris, A. D., & Garrabou, J. (2020). A fast-moving

target: achieving marine conservation goals under shifting climate and policies. *Ecological Applications*, *30*(1), e02009.

Ríos Castro, R., Costas Selas, C., Pallavicini, A., Vezzulli, L., Novoa, B., Teira Gonzalez, E. M., & Figueras, A. (2022). Co-occurrence and diversity patterns of benthonic and planktonic communities in a shallow marine ecosystem. *Frontiers in Marine Science*.

Risch, D., Castellote, M., Clark, C. W., Davis, G. E., Dugan, P. J., Hodge, L. E., Kumar, A., Lucke, K., Mellinger, D. K., & Nieukirk, S. L. (2014). Seasonal migrations of North Atlantic minke whales: novel insights from large-scale passive acoustic monitoring networks. *Movement Ecology*, *2*(1), 1-17.

Robertson, M. D., Koen-Alonso, M., Regular, P. M., Cadigan, N., & Zhang, F. (2022). Accounting for a nonlinear functional response when estimating prey dynamics using predator diet data. *Methods in Ecology and Evolution*, *13*(4), 880-893.

Robinson, K. P., Baumgartner, N., Eisfeld, S. M., Clark, N. M., Culloch, R. M., Haskins, G. N., Zapponi, L., Whaley, A. R., Weare, J. S., & Tetley, M. J. (2007). The summer distribution and occurrence of cetaceans in the coastal waters of the outer southern Moray Firth in northeast Scotland (UK). *Lutra*, *50*(1), 19.

Robinson, K. P., MacDougall, D., Bamford, C. C. G., Brown, W. J., Dolan, C. J., Hall, R., Haskins, G. N., Russell, G., Sidiropoulos, T., Sim, T. M. C., Spinou, E., Stroud, E., Williams, G., & Culloch, R. M. (2023). Ecological partitioning and feeding specialisations of coastal minke whales (*Balaenoptera acutorostrata*) using a recently designated MPA in northeast Scotland. *PLoS ONE. In Press.*

Robinson, K. P., Tetley, M. J., & Mitchelson-Jacob, E. G. (2009). The distribution and habitat preference of coastally occurring minke whales (Balaenoptera acutorostrata) in the outer southern Moray Firth, northeast Scotland. *Journal of Coastal Conservation*, *13*(1), 39-48.

Rodriguez-Ezpeleta, N., Zinger, L., Kinziger, A., Bik, H. M., Bonin, A., Coissac, E., Emerson, B. C., Martins Lopes, C., Pelletier, T. A., & Taberlet, P. (2021). Biodiversity monitoring using environmental DNA. In: Wiley-VCH.

Rose, R. A., Byler, D., Eastman, J. R., Fleishman, E., Geller, G., Goetz, S., Guild, L., Hamilton, H., Hansen, M., & Headley, R. (2015). Ten ways remote sensing can contribute to conservation. *Conservation biology*, *29*(2), 350-359.

Ruiz-Frau, A., Possingham, H. P., Edwards-Jones, G., Klein, C. J., Segan, D., & Kaiser, M. J. (2015). A multidisciplinary approach in the design of marine protected areas: Integration of science and stakeholder based methods. *Ocean & Coastal Management*, *103*, 86-93.

Russo, L., Bellardini, D., Zampicinini, G., Jordán, F., Congestri, R., & D'Alelio, D. (2023). From metabarcoding time series to plankton food webs: The hidden role of trophic hierarchy in providing ecological resilience. *Marine Ecology*, e12733.

Russo, L., Casella, V., Marabotti, A., Jordán, F., Congestri, R., & D'Alelio, D. (2022). Trophic hierarchy in a marine community revealed by network analysis on co-occurrence data. *Food Webs*, *32*, e00246.

Sadykova, D., Scott, B. E., De Dominicis, M., Wakelin, S. L., Wolf, J., & Sadykov, A. (2020). Ecological costs of climate change on marine predator–prey population distributions by 2050. *Ecology and Evolution*, *10*(2), 1069-1086.

Santini, L., Benítez-López, A., Maiorano, L., Čengić, M., & Huijbregts, M. A. (2021). Assessing the reliability of species distribution projections in climate change research. *Diversity and Distributions*, *27*(6), 1035-1050.

Santos, C. F., Agardy, T., Andrade, F., Crowder, L. B., Ehler, C. N., & Orbach, M. K. (2021). Major challenges in developing marine spatial planning. *Marine Policy*, *132*, 103248.

Santos, M., Pierce, G. J., Learmonth, J. A., Reid, R., Ross, H., Patterson, I., Reid, D., & Beare, D. (2004). Variability in the diet of harbor porpoises (Phocoena phocoena) in Scottish waters 1992–2003. *Marine Mammal Science*, *20*(1), 1-27.

Santos, M., Pierce, G. J., Reid, R., Patterson, I., Ross, H., & Mente, E. (2001). Stomach contents of bottlenose dolphins (Tursiops truncatus) in Scottish waters. *Journal of the Marine Biological Association of the United Kingdom*, *81*(5), 873-878.

Sawaya, N. A., Djurhuus, A., Closek, C. J., Hepner, M., Olesin, E., Visser, L., Kelble, C., Hubbard, K., & Breitbart, M. (2019). Assessing eukaryotic biodiversity in the Florida Keys National Marine Sanctuary through environmental DNA metabarcoding. *Ecology and Evolution*, *9*(3), 1029-1040.

Sbrocco, E. J., & Barber, P. H. (2013). MARSPEC: ocean climate layers for marine spatial ecology: Ecological Archives E094-086. *Ecology*, *94*(4), 979-979.

Schadewell, Y., & Adams, C. I. (2021). Forensics meets ecology–environmental DNA offers new capabilities for marine ecosystem and fisheries research. *Frontiers in Marine Science*, *8*, 668822.

Schmitt, S., Pouteau, R., Justeau, D., De Boissieu, F., & Birnbaum, P. (2017). ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, *8*(12), 1795-1803.

Schoeman, R. P., Patterson-Abrolat, C., & Plön, S. (2020). A global review of vessel collisions with marine animals. *Frontiers in Marine Science*, *7*, 292.

Schofield, G., Esteban, N., Katselidis, K. A., & Hays, G. C. (2019). Drones for research on sea turtles and other marine vertebrates–A review. *Biological Conservation*, *238*, 108214.

Segers, F., Dickey-Collas, M., & Rijnsdorp, A. (2007). Prey selection by North Sea herring (Clupea harengus), with special reference to fish eggs. *ICES Journal of Marine Science*, *64*(1), 60-68.

Sepulveda, A. J., Nelson, N. M., Jerde, C. L., & Luikart, G. (2020). Are environmental DNA methods ready for aquatic invasive species management? *Trends in ecology & evolution*, *35*(8), 668-678.

Sequeira, A. M. M., Hays, G. C., Sims, D. W., Eguíluz, V. M., Rodríguez, J. P., Heupel, M. R., Harcourt, R., Calich, H., Queiroz, N., & Costa, D. P. (2019). Overhauling ocean spatial planning to improve marine megafauna conservation. *Frontiers in Marine Science*, *6*, 639.

Sguotti, C., Lynam, C. P., García-Carreras, B., Ellis, J. R., & Engelhard, G. H. (2016). Distribution of skates and sharks in the North Sea: 112 years of change. *Global change biology*, *22*(8), 2729-2743.

Shelton, A. O., Gold, Z. J., Jensen, A. J., D′ Agnese, E., Andruszkiewicz Allan, E., Van Cise, A., Gallego, R., Ramón-Laca, A., Garber-Yonts, M., & Parsons, K. (2023). Toward quantitative metabarcoding. *Ecology*, *104*(2), e3906.

Shelton, A. O., Ramón-Laca, A., Wells, A., Clemons, J., Chu, D., Feist, B. E., Kelly, R. P., Parker-Stetter, S. L., Thomas, R., & Nichols, K. M. (2022). Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. *Proceedings of the Royal Society B*, *289*(1971), 20212613.

Shephard, S., Wögerbauer, C., Green, P., Ellis, J. R., & Roche, W. K. (2019). Angling records track the near extirpation of angel shark Squatina squatina from two Irish hotspots. *Endangered Species Research*, *38*, 153-158.

Sherr, E. B., & Sherr, B. F. (2007). Heterotrophic dinoflagellates: a significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Marine Ecology Progress Series*, *352*, 187-197.

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Molecular ecology*, *21*(8), 1794-1805.

Sigsgaard, E. E., Jensen, M. R., Winkelmann, I. E., Møller, P. R., Hansen, M. M., & Thomsen, P. F. (2020). Population-level inferences from environmental DNA—Current status and future perspectives. *Evolutionary applications*, *13*(2), 245-262.

Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., Pedersen, M. W., Jaidah, M. A., Orlando, L., & Willerslev, E. (2016). Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature Ecology & Evolution*, *1*(1), 0004.

Silber, G. K., Lettrich, M. D., Thomas, P. O., Baker, J. D., Baumgartner, M., Becker, E. A., Boveng, P., Dick, D. M., Fiechter, J., & Forcada, J. (2017). Projecting marine mammal distribution in a changing climate. *Frontiers in Marine Science*, 413.

Skelton, J., Cauvin, A., & Hunter, M. E. (2022). Environmental DNA metabarcoding read numbers and their variability predict species abundance, but weakly in non-dominant species. *Environmental DNA*.

Smith, R. J., Eastwood, P. D., Ota, Y., & Rogers, S. I. (2009). Developing best practice for using Marxan to locate Marine Protected Areas in European waters. *ICES Journal of Marine Science*, *66*(1), 188-194.

Sofaer, H. R., Jarnevich, C. S., Pearse, I. S., Smyth, R. L., Auer, S., Cook, G. L., Edwards Jr, T. C., Guala, G. F., Howard, T. G., & Morisette, J. T. (2019). Development and delivery of species distribution models to inform decision-making. *BioScience*, *69*(7), 544-557.

Sonsthagen, S. A., Jay, C. V., Cornman, R. S., Fischbach, A. S., Grebmeier, J. M., & Talbot, S. L. (2020). DNA metabarcoding of feces to infer summer diet of Pacific walruses. *Marine Mammal Science*, *36*(4), 1196-1211.

Spitz, J., Trites, A. W., Becquet, V., Brind'Amour, A., Cherel, Y., Galois, R., & Ridoux, V. (2012). Cost of living dictates what whales, dolphins and porpoises eat: the importance of prey quality on predator foraging strategies. *PloS one*, *7*(11), e50096.

Stepien, C. A., Lance, R. F., Klymus, K. E., & Hunter, M. E. (2023). The Government eDNA Working Group 6th Annual eDNA Technical Exchange Workshop. *Environmental DNA*.

Stewart, K., Ma, H., Zheng, J., & Zhao, J. (2017). Using environmental DNA to assess population-wide spatiotemporal reserve use. *Conservation biology*, *31*(5), 1173-1182.

Suarez-Bregua, P., Álvarez-González, M., Parsons, K. M., Rotllant, J., Pierce, G. J., & Saavedra, C. (2022). Environmental DNA (eDNA) for monitoring marine mammals: Challenges and opportunities. *Frontiers in Marine Science*, 1886.

Sumaila, U. R., & Tai, T. C. (2020). End overfishing and increase the resilience of the ocean to climate change. *Frontiers in Marine Science*, *7*, 523.

Sundaram, M., & Leslie, A. B. (2021). The influence of climate and palaeoclimate on distributions of global conifer clades depends on geographical range size. *Journal of Biogeography*, *48*(9), 2286-2297.

Székely, D., Cammen, K. M., & Tange Olsen, M. (2021). Needles in an ocean haystack: using environmental DNA to study marine mammals in the North Atlantic. *NAMMCO Scientific Publications*, *12*.

Szesciorka, A. R., Demer, D. A., Santora, J. A., Forney, K. A., & Moore, J. E. (2023). Multiscale relationships between humpback whales and forage species hotspots within a large marine ecosystem. *Ecological Applications*, *33*(2), e2794.

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring*. Oxford University Press.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. In (Vol. 21, pp. 1789-1793): Wiley Online Library.

Takahashi, M., Saccò, M., Kestel, J. H., Nester, G., Campbell, M. A., Van Der Heyde, M., Heydenrych, M. J., Juszkiewicz, D. J., Nevill, P., & Dawkins, K. L. (2023). Aquatic environmental DNA: A review of the macro-organismal biomonitoring revolution. *Science of The Total Environment*, 162322.

Telenius, A. (2011). Biodiversity information goes public: GBIF at your service. *Nordic Journal of Botany*, *29*(3), 378-381.

Temple, H. J., & Terry, A. (2009). European mammals: Red List status, trends, and conservation priorities. *Folia zoologica*, *58*(3), 248.

Tessarolo, G., Rangel, T. F., Araújo, M. B., & Hortal, J. (2014). Uncertainty associated with survey design in Species Distribution Models. *Diversity and Distributions*, *20*(11), 1258-1269.

Tetley, M., Mitchelson-Jacob, E., & Robinson, K. (2008). The summer distribution of coastal minke whales (Balaenoptera acutorostrata) in the southern outer Moray Firth, NE Scotland, in relation to co-occurring mesoscale oceanographic features. *Remote Sensing of Environment*, *112*(8), 3449-3454.

Thompson, P. M., Pierce, G. J., Hislop, J., Miller, D., & Diack, J. (1991). Winter foraging by common seals (Phoca vitulina) in relation to food availability in the inner Moray Firth, NE Scotland. *The Journal of Animal Ecology*, 283-294.

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples.

Thomsen, P. F., Møller, P. R., Sigsgaard, E. E., Knudsen, S. W., Jørgensen, O. A., & Willerslev, E. (2016). Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. *PloS one*, *11*(11), e0165252.

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4-18.

Thurman, L. L., Barner, A. K., Garcia, T. S., & Chestnut, T. (2019). Testing the link between species interactions and species co-occurrence in a trophic network. *Ecography*, *42*(10), 1658-1670.

Tixier, P., Lea, M. A., Hindell, M. A., Welsford, D., Mazé, C., Gourguet, S., & Arnould, J. P. (2021). When large marine predators feed on fisheries catches: global patterns of the depredation conflict and directions for coexistence. *Fish and Fisheries*, *22*(1), 31-53.

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences. *Scientific reports*, *7*(1), 1-14.

Trouillet, B., & Jay, S. (2021). The complex relationships between marine protected areas and marine spatial planning: Towards an analytical framework. *Marine Policy*, *127*, 104441.

Truelove, N. K., Andruszkiewicz, E. A., & Block, B. A. (2019). A rapid environmental DNA method for detecting white sharks in the open ocean. *Methods in Ecology and Evolution*, *10*(8), 1128-1135.

Truelove, N. K., Patin, N. V., Min, M., Pitz, K. J., Preston, C. M., Yamahara, K. M., Zhang, Y., Raanan, B. Y., Kieft, B., & Hobson, B. (2022). Expanding the temporal and spatial scales of environmental DNA research with autonomous sampling. *Environmental DNA*.

Tsuji, S., Miya, M., Ushio, M., Sato, H., Minamoto, T., & Yamanaka, H. (2020). Evaluating intraspecific genetic diversity using environmental DNA and

denoising approach: A case study using tank water. *Environmental DNA*, *2*(1), 42-52.

Tydecks, L., Jeschke, J. M., Wolf, M., Singer, G., & Tockner, K. (2018). Spatial and topical imbalances in biodiversity research. *PloS one*, *13*(7), e0199327.

Tylianakis, J. M., Laliberté, E., Nielsen, A., & Bascompte, J. (2010). Conservation of species interaction networks. *Biological Conservation*, *143*(10), 2270-2279.

UNEP-WCMC, & IUCN. (2023). Marine Protected Planet. Available at: http://www.protectedplanet.net/. Accessed March 2023.

UNEP/CBD/EBSA/WS/2014/3/4. (2014). Report of the Mediterranean regional workshop to facilitate the description of ecologically or biologically significant marine areas. Available from: https://www.cbd.int/ebsa/. Accessed October 2022.

Valiente-Banuet, A., Aizen, M. A., Alcántara, J. M., Arroyo, J., Cocucci, A., Galetti, M., García, M. B., García, D., Gómez, J. M., & Jordano, P. (2015). Beyond species loss: the extinction of ecological interactions in a changing world. *Functional Ecology*, *29*(3), 299-307.

Valsecchi, E., Arcangeli, A., Lombardi, R., Boyse, E., Carr, I. M., Galli, P., & Goodman, S. J. (2021). Ferries and environmental DNA: underway sampling from commercial vessels provides new opportunities for systematic genetic surveys of marine biodiversity. *Frontiers in Marine Science*, 1136.

Valsecchi, E., Bylemans, J., Goodman, S. J., Lombardi, R., Carr, I., Castellano, L., Galimberti, A., & Galli, P. (2020). Novel universal primers for metabarcoding environmental DNA surveys of marine mammals and other marine vertebrates. *Environmental DNA*, *2*(4), 460-476.

Valsecchi, E., Tavecchia, G., Boldrocchi, G., Coppola, E., Ramella, D., Conte, L., Blasi, M., Bruno, A., & Galli, P. (2023). Playing "hide and seek" with the Mediterranean monk seal: a citizen science dataset reveals its distribution from molecular traces (eDNA). *Scientific reports*, *13*(1), 2610.

van der Stap, T., Coolen, J. W., & Lindeboom, H. J. (2016). Marine fouling assemblages on offshore gas platforms in the southern North Sea: effects of depth and distance from shore on biodiversity. *PloS one*, *11*(1), e0146324.

van Oppen, M. J., & Coleman, M. A. (2022). Advancing the protection of marine life through genomics. *PLoS biology*, *20*(10), e3001801.

Van Pelt, T. I., Piatt, J. F., Lance, B. K., & Roby, D. D. (1997). Proximate composition and energy density of some North Pacific forage fishes. *Comparative Biochemistry and Physiology Part A: Physiology*, *118*(4), 1393-1398.

van Weelden, C., Towers, J. R., & Bosker, T. (2021). Impacts of climate change on cetacean distribution, habitat and migration. *Climate Change Ecology*, *1*, 100009.

Víkingsson, G. A., Elvarsson, B. Þ., Ólafsdóttir, D., Sigurjónsson, J., Chosson, V., & Galan, A. (2014). Recent changes in the diet composition of common minke whales (Balaenoptera acutorostrata) in Icelandic waters. A consequence of climate change? *Marine Biology Research*, *10*(2), 138-152.

Vilà, M., Basnou, C., Pyšek, P., Josefsson, M., Genovesi, P., Gollasch, S., Nentwig, W., Olenin, S., Roques, A., & Roy, D. (2010). How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. *Frontiers in Ecology and the Environment*, *8*(3), 135-144.

Visalli, M. E., Best, B. D., Cabral, R. B., Cheung, W. W., Clark, N. A., Garilao, C., Kaschner, K., Kesner-Reyes, K., Lam, V. W., & Maxwell, S. M. (2020). Data-driven approach for highlighting priority areas for protection in marine areas beyond national jurisdiction. *Marine Policy*, *122*, 103927.

Visser, F., Merten, V., Bayer, T., Oudejans, M., De Jonge, D., Puebla, O., Reusch, T. B., Fuss, J., & Hoving, H. (2021). Deep-sea predator niche segregation

revealed by combined cetacean biologging and eDNA analysis of cephalopod prey. *Science Advances*, *7*(14), eabf5908.

Walsh, S. J. (1994). Life history traits and spawning characteristics in populations of long rough dab (American plaice) Hippoglossoides platessoides (Fabricus) in the North Atlantic. *Netherlands Journal of Sea Research*, *32*(3-4), 241-254.

Wang, C., & Wang, J. (2011). Spatial pattern of the global shipping network and its hub-and-spoke system. *Research in Transportation Economics*, *32*(1), 54-63.

Wang, J.-F., Stein, A., Gao, B.-B., & Ge, Y. (2012). A review of spatial sampling. *Spatial Statistics*, *2*, 1-14.

Wanless, S., Harris, M., Redman, P., & Speakman, J. (2005). Low energy values of fish as a probable cause of a major seabird breeding failure in the North Sea. *Marine Ecology Progress Series*, *294*, 1-8.

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., & Rulik, B. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of The Total Environment*, *678*, 499-524.

Wetzel, F. T., Bingham, H. C., Groom, Q., Haase, P., Kõljalg, U., Kuhlmann, M., Martin, C. S., Penev, L., Robertson, T., & Saarenmaa, H. (2018). Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological Conservation*, *221*, 78-85.

Whitaker, J. M., Brower, A. L., & Janosik, A. M. (2021). Invasive lionfish detected in estuaries in the northern Gulf of Mexico using environmental DNA. *Environmental Biology of Fishes*, *104*, 1475-1485.

Wilhelm, T. A., Sheppard, C. R., Sheppard, A. L., Gaymer, C. F., Parks, J., Wagner, D., & Lewis, N. a. (2014). Large marine protected areas–advantages and challenges of going big. *Aquatic conservation: marine and freshwater ecosystems*, *24*(S2), 24-30.

Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M. B., Brand, T. B., Hofreiter, M., Bunce, M., Poinar, H. N., & Dahl-Jensen, D. (2007). Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Science*, *317*(5834), 111-114.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A., & Cooper, A. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, *300*(5620), 791-795.

Williams, M. A., O'Grady, J., Ball, B., Carlsson, J., de Eyto, E., McGinnity, P., Jennings, E., Regan, F., & Parle-McDermott, A. (2019). The application of CRISPR-Cas for single species identification from environmental DNA. *Molecular ecology resources*, *19*(5), 1106-1114.

Wilson, K. L., Tittensor, D. P., Worm, B., & Lotze, H. K. (2020). Incorporating climate change adaptation into marine protected area planning. *Global change biology*, *26*(6), 3251-3267.

Windsland, K., LINDSTRØM, U., NILSSEN, K. T., & HAUG, T. (2007). 167 Relative abundance and size composition of prey in the common minke whale diet in selected areas of the northeastern Atlantic during 2000-04. *J. Cetacean Res. Manage*, *9*(3), 167-178.

Wisz, M. S., Hijmans, R., Li, J., Peterson, A. T., Graham, C., Guisan, A., & Group, N. P. S. D. W. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763-773.

Wright, P., Pinnegar, J., & Fox, C. (2020). Impacts of climate change on fish, relevant to the coastal and marine environment around the UK. *MCCIP Sci. Rev*, *2020*, 354-381.

WWF. (2022). *Living Planet Report 2022 - Building a nature-positive society.* Almond, R.E.A., Grooten, M, Juffe Bignoli, D., & Petersen, T. (Eds). WWF, Gland, Switzerland.

Yao, M., Zhang, S., Lu, Q., Chen, X., Zhang, S. Y., Kong, Y., & Zhao, J. (2022). Fishing for fish environmental DNA: Ecological applications, methodological considerations, surveying designs, and ways forward. *Molecular ecology*, *31*(20), 5132-5164.

Yoccoz, N. G. (2012). The future of environmental DNA in ecology. *Molecular ecology*, *21*(8), 2031-2038.

Young, J. W., Hunt, B. P., Cook, T. R., Llopiz, J. K., Hazen, E. L., Pethybridge, H. R., Ceccarelli, D., Lorrain, A., Olson, R. J., & Allain, V. (2015). The trophodynamics of marine top predators: current knowledge, recent advances and challenges. *Deep Sea Research Part II: Topical Studies in Oceanography*, *113*, 170-187.

Young, M. (2015). Building the blue economy: the role of marine spatial planning in facilitating offshore renewable energy development. *The International Journal of Marine and Coastal Law*, *30*(1), 148-174.

Zamkovaya, T., Foster, J. S., de Crécy-Lagard, V., & Conesa, A. (2021). A network approach to elucidate and prioritize microbial dark matter in microbial communities. *The ISME journal*, *15*(1), 228-244.

Zamora-Terol, S., Novotny, A., & Winder, M. (2020). Reconstructing marine plankton food web interactions using DNA metabarcoding. *Molecular ecology*, *29*(17), 3380-3395.

Zhang, C., Xu, B., Xue, Y., & Ren, Y. (2020). Evaluating multispecies survey designs using a joint species distribution model. *Aquaculture and Fisheries*, *5*(3), 156-162.

Zhang, L., Huettmann, F., Zhang, X., Liu, S., Sun, P., Yu, Z., & Mi, C. (2019). The use of classification and regression algorithms using the random forests method with presence-only data to model species' distribution. *MethodsX*, *6*, 2281-2292.

Zhang, X., Luo, D., YU, R. Q., & Wu, Y. (2023). Multilocus DNA metabarcoding diet analyses of small cetaceans: a case study on highly vulnerable humpback dolphins and finless porpoises from the Pearl River Estuary, China. *Integrative Zoology*, *18*(1), 183-198.

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., & Guisan, A. (2020). A standard protocol for reporting species distribution models. *Ecography*, *43*(9), 1261-1277.

Zurell, D., Pollock, L. J., & Thuiller, W. (2018). Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? *Ecography*, *41*(11), 1812-1819.

Zurell, D., Zimmermann, N. E., Gross, H., Baltensweiler, A., Sattler, T., & Wüest, R. O. (2020). Testing species assemblage predictions from stacked and joint species distribution models. *Journal of Biogeography*, *47*(1), 101-113.