COMPARING ENVIRONMENTAL DNA METABARCODING AND UNDERWATER VISUAL CENSUS IN ASSESSING FISH BIODIVERSITY IN A MARINE PROTECTED AREA: KAS-KEKOVA MPA

by

Merjan Beyza Koşucu

B.Sc. in Molecular Biology and Genetics, Boğaziçi University, 2020

Submitted to the Institute of Environmental Sciences
in partial fulfillment of the requirements for the degree of
Master of Science
in
Environmental Sciences

Boğaziçi University
2024

COMPARING ENVIRONMENTAL DNA METABARCODING AND UNDERWATER VISUAL CENSUS IN ASSESSING FISH BIODIVERSITY IN A MARINE PROTECTED AREA: KAS-KEKOVA MPA

APPROVED BY:

Prof. Dr. Raşit Bilgin . . . . . .
Thesis Advisor

Prof. Dr. Andrzej Furman . . . . . .

Assoc. Prof. Dr. İsmail K. Sağlam . . . . . .

DATE OF APPROVAL: …/02/2024

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my advisor, Prof. Dr. Raşit Bilgin, for his supervision, insightful guidance, and unwavering support both for my studies and the laboratory work.

I am grateful to my committee members, Prof. Dr. Andrzej Furman and Assoc. Prof. İsmail K. Sağlam, for their time, invaluable feedback, and thoughtful contributions.

This project was made possible through the generosity of Dr. Baki Yokeş, who let us join their visual survey field in Kaş-Kekova MPA and shared their valuable visual count findings with us, as well as Dr. Ünsal Karhan for his guidance and collaboration in sampling, and Açelya Eren for her contributions in the fieldwork and filtration of the samples.

I would like to express my gratitude to Aygün Karaçay for her help in the laboratory experiments, Onur Doğan for his help during the bioinformatic analyses, as well as to the members of our laboratory Maha Salameh, Merve Sıvacı, Nehir Sevim, Ayşe Mergenci, and Anıl Örün for their companionship and support.

My deepest gratitude extends to my mother, father and dearest sister. As a girl growing up in the countryside, I am profoundly grateful for their selfless sacrifices during my education. In a society where many parents may not support their daughters to pursue higher education, their endless and unwavering support ultimately led me to this day.

Last but not least, I would like to convey my sincerest appreciation for my dearest friend İdil Çağla Ertaş for her encouragement, her immeasurable impact, and the precious moments we shouldered together. I am very grateful for her presence in my life, her camaraderie in every shared laughter and tears, and for resisting and standing together through thick and thin.

ABSTRACT

COMPARING ENVIRONMENTAL DNA METABARCODING AND UNDERWATER VISUAL CENSUS IN ASSESSING FISH BIODIVERSITY IN A MARINE PROTECTED AREA: KAS-KEKOVA MPA

The impacts of anthropological activities are one of the major causes leading to biodiversity loss. Marine biodiversity, especially fish, suffers from human activities the most through overexploitation, pollution, and habitat loss and fragmentation; therefore, Marine Protected Areas play a significant role in protecting and conserving marine biodiversity and habitat. However, biomonitoring of species diversity in such areas is usually dependent on conventional surveys, which can be costly, invasive, require taxonomic expertise, and are potentially prone to subjectivity and error. As an alternative to traditional surveys, eDNA metabarcoding can be utilized to assess and monitor diversity in biodiversity spots. In this study, 21 seawater samples were collected in a marine protected area (Kaş-Kekova) in Antalya, Turkey, from three different sampling depths (surface, 12m, and 24m) in 16 stations. Two sets of universal primers targeting COI and 12S rRNA markers were used in the eDNA metabarcoding. This study aimed to assess the fish biodiversity in Kaş-Kekova MPA by eDNA metabarcoding and compare it with the simultaneous underwater visual census (UVC) findings and whether eDNA metabarcoding cross-validates the fish biodiversity recorded in the area and to detect the presence of any invasive or elusive species where the conventional methods might fall short of their detection. The results of the analyses demonstrated that the two methods were complementary to each other and with better sampling effort, utilizing eDNA metabarcoding could be promising in terms of species diversity assessment and forewarning the possible invasive species.

ÖZET

KAŞ-KEKOVA DEKA'DA BALIK BİYOÇEŞİTLİLİĞİNİ DEĞERLENDIRMEDE ÇEVRESEL DNA METABARKODLAMA VE SU ALTI GÖRSEL SAYIMIN KARŞILAŞTIRILMASI

Antropolojik faaliyetlerin etkileri biyolojik çeşitlilik kaybına yol açan başlıca nedenlerden biridir. Deniz biyoçeşitliliği, özellikle de balıklar, en çok zararı, aşırı avlanma, kirlilik, habitat kaybı ve fragmantasyonu nedeniyle insan faaliyetlerinden görmektedir. Bu nedenle Deniz Koruma Alanları, deniz biyolojik çeşitliliğinin ve habitatının korunması ve muhafaza edilmesinde önemli bir rol oynamaktadır. Bununla birlikte, bu tür alanlardaki tür çeşitliliğinin biyoizlemesi genellikle maliyetli, invaziv olabilen, yüksek taksonomik uzmanlık gerektiren ve potansiyel olarak sübjektifliğe ve hataya açık olan geleneksel araştırmalara bağlıdır. Geleneksel araştırmalara alternatif olarak eDNA metabarkodlaması, biyolojik çeşitlilik noktalarındaki çeşitliliği değerlendirmek ve izlemek için kullanılabilir. Bu çalışmada, Antalya, Türkiye'de Kaş-Kekova deniz koruma alanından 16 istasyonda üç farklı örnekleme derinliğinden (yüzey, 12m ve 24m) 21 deniz suyu örneği toplandı. eDNA metabarkodlamasında COI ve 12S rRNA işaretleyicilerini hedef alan iki set evrensel primer kullanıldı. Bu çalışma, Kaş-Kekova DEKA'daki balık biyoçeşitliliğini eDNA metabarkodlama ile değerlendirmek ve bunu eş zamanlı sualtı görsel sayım (UVC) bulgularıyla karşılaştırmak ve eDNA metabarkodlamanın alanda kaydedilen balık biyoçeşitliliğini çapraz doğrulayıp doğrulamadığını ve geleneksel yöntemlerin tespit etmekte yetersiz kalabileceği istilacı veya yakalanması zor türlerin varlığını tespit etmeyi amaçlamıştır. Analiz sonuçları, iki yöntemin birbirini tamamladığını ve daha iyi örnekleme çalışmasıyla eDNA metabarkodlamanın kullanılmasının tür çeşitliliğinin değerlendirilmesi ve olası istilacı türlerin varlığının önceden habercisi olması açısından umut verici olabileceğini gösterdi.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....iii
ABSTRACT.....iv
ÖZET.....v
TABLE OF CONTENTS.....vi
LIST OF FIGURES.....viii
LIST OF TABLES.....xi
LIST OF SYMBOLS/ABBREVIATIONS.....xii
1. INTRODUCTION.....1
1.1. Marine Biodiversity as an Indicator of Ecosystem Health.....1
1.2. Marine Protected Areas: Their Role in Conservation.....2
1.3. Biomonitoring in the Marine Protected Areas.....2
1.4. Biodiversity Monitoring: Conventional vs eDNA Metabarcoding.....3
1.5. Objective of the Thesis.....3
2. LITERATURE REVIEW.....5
2.1. Environmental DNA Approaches.....5
2.2. Environmental DNA Advantages and Disadvantages.....6
2.3. Environmental DNA Metabarcoding Compared with Traditional Methods.....7
2.4. Case Studies of eDNA Metabarcoding in Fish.....7
3. MATERIALS AND METHODS.....11
3.1. Sample Collection.....11
3.2. DNA Extraction.....12
3.3. DNA Amplification and Library Preparation (Two-Step PCR).....12
3.4. Library Pooling and DNA Sequencing.....14
3.5. Bioinformatics.....14
3.6. Statistics.....15
4. RESULTS.....17
4.1. The Validity of eDNA Metabarcoding.....17
4.2. Complementarity of eDNA Metabarcoding and Underwater Visual Census.....19
4.3. Species Identification and Taxonomic Assignment Through Analysis of the 12S rRNA Gene Marker.....23
4.4. Species Identification and Taxonomic Assignment Through Analysis of the COI Gene Marker.....27

4.5. Comparison of the Fish Biodiversity Characterized by Both Primers……31
5. DISCUSSION.....36
6. CONCLUSIONS AND RECOMMENDATIONS.....42
REFERENCES…….....44
APPENDIX A: DNA EXTRACTION PROTOCOLS……...55
APPENDIX B: THE LIST OF FISH SPECIES DETECTED BY BOTH METHODS……..56
APPENDIX C: THE REPLICATE ANALYSIS……...……..……...58

LIST OF FIGURES

Figure 3.1. The map of the study area Kaş-Kekova MPA……...11

Figure 4.1. The number of taxa detected by both methods that were on the UVC checklist. A) at species, B) at genus, C) at family level……..……..…17

Figure 4.2. The distribution of species detected by the two methods normalized to the species of the UVC checklist……..……19

Figure 4.3. The total number of A) fish genera B) fish species detected in terms of species observed in each transect based on all observations in sampling locations. C) The fish species detected in terms of species observed at each transect with the same sampling locations and depth as the eDNA sampling.....…….20

Figure 4.4. The pairwise comparison of the fish species that were identified by both methods.....21

Figure 4.5. Principal Coordinates Analysis (PCoA) graph incorporating methodology as a factor A) at the genus level, B) at the species level.....22

Figure 4.6. Principal Coordinates Analysis (PCoA) graph of UVC incorporating A) location as a factor and B) depth as a factor, C) depth as a factor for all the sampling depths, with categories interconnected.....23

Figure 4.7. A) The number of reads per sample, B) the number of OTUs per sample and C) the number of OTUs assigned at the species level per sample, after data filtration (sample S02 excluded) and read normalization.....24

Figure 4.8. A) Taxonomic richness and B) taxonomic resolution for each taxon assigned to the OTUs generated from 12S rRNA primers after S02 is excluded and the reads are normalized to the minimum number of reads…….…….25

Figure 4.9. Alpha diversity box plot illustrating number of species assigned by 12S primers per each sampling depth at the Kaş-Kekova MPA ……...25

Figure 4.10. Principal Coordinates Analysis (PCoA) graph incorporating A) location as a factor and B) depth as a factor, with categories interconnected.……26

Figure 4.11. The bar chart representing the class proportions.……26

Figure 4.12. The pie chart demonstrating the species proportions.……27

Figure 4.13. A) The number of reads per sample, B) the number of OTUs per sample and C) the number of OTUs at the species level per sample, after data filtration and read normalization.…….....28

Figure 4.14. A) Taxonomic richness and B) Taxonomic resolution for each taxon assigned to the OTUs generated from COI primers. The x-axis represents each taxonomic level from phylum to species. ……29

Figure 4.15. The alpha diversity box plot illustrates the number of species assigned by COI primers to each sampling depth at the Kaş-Kekova MPA.……29

Figure 4.16. Principal Coordinates Analysis (PCoA) graph incorporating A) location as a factor and B) depth as a factor, with categories interconnected.……30

Figure 4.17. The bar chart demonstrating the phylum proportions (cut-off 1% abundance)…….31

Figure 4.18. Alpha diversity boxplot demonstrates species diversity as the number of species in the Chordata phylum assigned by both primers……...31

Figure 4.19. The number of species in the Chordata phylum that were detected by COI primers (red), 12S rRNA primers (green), and both primer sets (orange)……32

Figure 4.20. The number of fish species that were detected by 12S rRNA primers (green), COI primers (red), and both primer sets (orange).….....32

Figure 4.21. The pairwise sample comparison of the fish species that were identified by both primers……...35

LIST OF TABLES

Table 3.1. The sampling dates, the locations and the renamed samples after the biological replicates were merged.……12

Table 3.2. The list of primers that are used in the two-step PCR protocol. (Green: Overhang sequences used in 16S rRNA library preparation, Blue: Illumina adapter sequences, Red: Index sequences)……13

Table 4.1. The species that were detected by the UVCs exclusively and their records in the NCBI database……...17

Table 4.2. The species found in the Chordata phylum detected by primers COI, 12S and both.....30

Table 4.3. The fish species detected by primers targeting 12S rRNA and COI regions……...31

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	Explanation
g	Gram
L	Liter
ml	Milliliter
mM	Millimolar
ng	Nanogram
ºC	Celcius
µl	Microliter

Abbreviation	Explanation
A	Adenine
ANOSIM	Analysis Of Similarities
BLAST	Basic Local Alignment Search Tool
BOLD	Barcode of Life Data
bp	Base Pair
C	Cytosine
COI	Cytochrome C Oxidase Subunit I
DNA	Deoxyribonucleic Acid
dsDNA	Double-Stranded Deoxyribonucleic Acid
eDNA	Environmental DNA
G	Guanine
HS	High Sensitivity
IUCN	International Union for Conservation of Nature
MPA	Marine Protected Area
N	Adenine / Guanine / Cytosine / Thymine
n	Number of Samples
NCBI	The National Center for Biotechnology Information
NGS	Next-Generation Sequencing
nM	Nanomolar
OTU	Operational Taxonomic Unit
PBS	Phosphate-Buffered Saline

PCoA	Principal Coordinate Analysis
RNA	Ribonucleic Acid
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
SEPA	Specially Protected Area
sp.	Species
T	Thymine
TAE	Tris-acetate-EDTA
UVC	Underwater Visual Census
W	Adenine / Thymine
Y	Cytosine / Thymine (Pyrimidine)

1. INTRODUCTION

The decline in global biodiversity is an undeniable fact. Compared to the average species extinction rates of the last 10 million years, today's global extinction rates are estimated to be ten to a hundred times higher (IPBES, 2019; Pimm et al., 2014). It has been reported that at least 1000 plant and animal species have gone extinct since 1600 (Gaston & Spicer, 2004). IUCN has so far assessed 147,500 species, and at least 28% of the species included in the Red List are under the threat of extinction (IUCN, 2022). However, some groups of taxa are more susceptible to extinction and have higher percentages of threatened species on the Red List; 41% of the listed species are amphibians, 33% are corals, and 38% constitute sharks (IUCN, 2022). Even though speciation and extinction are natural events through the course of evolution, today's extinction rates and biodiversity loss are inevitably alarmingly higher than the natural rate of extinction (Primack & Morrison, 2013). The reason for the increase in biodiversity loss globally is due to anthropological activities resulting in habitat loss, fragmentation and degradation, overexploitation of species and overharvesting, the introduction of alien/ exotic species, and finally, extinction cascades, which are known as the "evil quartet" (Diamond,1984 as cited in Gaston & Spicer, 2004).

1.1. Marine Biodiversity As An Indicator of Ecosystem Health Waters

Marine biodiversity is one of the indicators of the earth's health and its durability against any fluctuations in environmental conditions. 72% of the earth's surface is covered by oceans, and many organisms reside in 90% of the diverse habitats provided by the oceans and the seas (Goulletquer et al., 2014). Besides global nutrient cycling, marine biodiversity provides many ecosystem services and resources, providing oxygen to the atmosphere, carbon sequestration, and storage, as well as ecosystem services and many resources for the coastal communities (Goulletquer et al., 2014). Conservation of marine biodiversity is also one of the Sustainable Development Goals (United Nations, 2022). Therefore, the conservation and protection of marine biodiversity, their management, and monitoring in exceptionally high biodiversity spots are crucial. Fish biodiversity plays a significant role in the health of the earth's oceans and seas; since fish are the representatives of megafauna in marine ecosystems. They are also the most susceptible group to overexploitation, and their loss can have detrimental impacts on ecosystem functioning (Norris et al., 2020). Besides their cultural and use-value, fish populations provide a range of ecosystem services, such as regulating food web processes, nutrient recycling, carbon flux (Mariani et al., 2020), impacting sediment processes, and regulating ecosystem resilience (Holmlund & Hammer, 1999).

1.2. Marine Protected Areas: Their Role in Conservation

As marine habitats and ecosystems have essential roles in the health of the planet, their protection, conservation, and sustainability must be ensured. One of the solutions for the protection and conservation of marine ecosystems and their biodiversity is the establishment of Marine Protected Areas (MPA) (Mcleod, 2013). There are many definitions and implications of Marine Protected Areas, and the definitions are prone to confusion. In order to overcome this problem, IUCN generalized a "Protected Area" regardless of the environment the protected area is in, whether terrestrial or marine, and defined it as: "A protected area is a clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values." (Dudley, 2008).

MPAs help to protect, conserve, and restore marine biodiversity as well as marine ecosystem structure and function by supporting the preservation of nursery and spawning sites for marine organisms, maintaining the populations of keystone species, conserving the populations of threatened endemic species, and preventing overharvesting of species that have economic and use values (Laffoley et al., 2018; Mcleod, 2013). Moreover, MPAs can provide economic and cultural benefits to the local communities through marine tourism, education, and training about the cultural heritage and the importance of the marine protected areas (Hoagland et al., 2019; Laffoley et al., 2018; Mcleod, 2013) However, MPAs solely are not adequate to conserve ecosystems and biodiversity since, they are still prone to the impacts of pollution, climate change, and other anthropological activities (e.g. industrial and agricultural) taking place in the proximal terrestrial ecosystems (Laffoley et al., 2018). Therefore, a multi-faceted and scientific approach is essential in the establishment, management, and monitoring of the MPAs.

1.3. Biomonitoring in the Marine Protected Areas

Biomonitoring of the biodiversity in the MPAs is critical to assess the biomass and the abundance of the native species in the ecosystem, as well as to track the status of the threatened or invasive species (Ushio et al., 2018). The methods used for the biodiversity assessment in protected areas generally depend on observations carried out with fishing gears such as catch and release, trawling, or underwater surveys such as underwater visual census (UVC) or via remote cameras (Whitmarsh et al., 2017) or underwater traps, depending on the species (e.g. benthic traps for benthic species or light traps for light-sensitive species) (Costello et al., 2017). The conventional methods are cost- and labour-intensive and require higher expertise compared to a relatively new approach called

environmental DNA (eDNA) metabarcoding (Bohmann et al., 2014; Deiner et al., 2017). eDNA metabarcoding is based on the characterization of biodiversity in an environmental sample by total DNA isolation, amplifying a barcode or a marker gene of interest by generic or species-specific primers, sequencing via high-throughput sequencing (NGS), and comparing the sequenced data with a reference library.

1.4. Biodiversity Monitoring: Conventional versus eDNA Metabarcoding

The conventional biodiversity surveying methods are generally prone to subjectivity as it depends on the taxonomic expertise of the species and the habitat that is monitored (Sanchez et al., 2022; Thomsen & Willerslev, 2015). Therefore, eDNA metabarcoding can be a more standardized method for the characterization of the biodiversity of a given study area when compared to the visual identification methods (Thomsen et al., 2012; Thomsen & Willerslev, 2015) and has better sensitivity in the detection of species that are elusive/ cryptic or invasive (Bohmann et al., 2014; Dejean et al., 2012a). eDNA metabarcoding can detect an assemblage of many different species by just taking water samples from the area of the study, making it a preferable approach in the case of the biodiversity assessment of the protected and vulnerable areas, which makes this method non-invasive compared to the previous methods such as trawling which cannot be used in shallow waters (Thomsen et al., 2012) and can be destructive on the structure of the habitat of the vulnerable and threatened species (Bohmann et al., 2014).

1.5. Objective of the Thesis

In this study, the fish biodiversity of the different zones in the Kas-Kekova MPA will be assessed by using eDNA metabarcoding while comparing the findings of this method with the underwater visual census (UVC). The eDNA will be compared to the physical observations and will be tested for its utilization as a novel approach in bio-monitoring and biodiversity assessment of fish in the area. Kaş-Kekova Special Environmental Protected Area consists of both marine and terrestrial areas, with the former and latter covering 166km2 and 92km2, respectively (Huseyinoglu et al., 2021). Previously the PA consisted of the Kekova region; however, in 2006, the area was recognized and specified as Kaş-Kekova Specially Protected Area (SEPA) by the Council of Ministers (WWF-Turkey, 2012). Kaş-Kekova Marine Protected Area is a protected area that is home to many endemic fish species in the Mediterranean. Besides being a biodiversity hotspot for many different marine species, it also has economic, cultural, and archeological significance (Demir, 2011). Therefore, this study aims to assess the fish biodiversity in the area and compare it with the findings of the UVC and

check whether eDNA metabarcoding approach cross-validates the fish biodiversity recorded in the area and whether it can detect the presence of any invasive or elusive species where the conventional methods might fall short of their detection.

2. LITERATURE REVIEW

Even though eDNA metabarcoding is a promising approach in spatial and temporal biodiversity assessments and monitoring in terrestrial, aquatic, and marine habitats, possible challenges and shortcomings remain. This section will review the definition and the concept of eDNA. The advantages and disadvantages of eDNA metabarcoding will be presented, and as a relatively novel method, the success of eDNA compared to the traditional methods will be included. Since this study aims to assess the fish biodiversity and compare it to the findings of the underwater visual census, the studies done to determine the fish biodiversity by eDNA and traditional methods and their effectiveness in the assessment will be included in the review. Finally, to have a reference base for the taxonomy of the area, the previous biodiversity detected at the study site (Kaş, Antalya, Turkey) will be included.

2.1. Environmental DNA Approaches

Environmental DNA can be defined as the genetic material which is present in different types of habitats and mediums, such as soil, water, sediments, biofilms, and air (Ruppert et al., 2019). The sources of eDNA could be tissues, i.e., skin, blood, larvae, eggs, sperm, decaying organisms, bodily fluids or secretions such as mucus, saliva, feces, urine, and the organism itself in case of microorganisms (Taberlet et al., 2012). From samples obtained from soil, water, and or air, among others, eDNA can be preserved, extracted, and amplified. The amplicons of eDNA can then be sequenced, and the taxonomic analyses and assignment of the organisms can be achieved, resulting in the species characterization of the environmental sample.

eDNA samples can be and are utilized in many ecological studies, whether focusing on single-species detection in a habitat or multispecies approach where different sets of universal primers are selected to amplify the eDNA, and taxonomic analyses are performed after high-throughput sequencing. Besides biodiversity monitoring in ecosystems such as marine, estuarine and terrestrial, eDNA metabarcoding can be used for different purposes, including paleogenetic/ancient DNA, plant-pollinator interactions, diet analyses, invasive species detection, pollination response and ecological quality assessment, and monitoring of air quality. (Bohmann et al., 2014; Ruppert et al., 2019). Moreover, in the case of microorganisms, wastewater samples or any other medium can be used for pathogen analyses (metagenomics), which microbiologists have historically employed longer than it has been used in biodiversity assessments (Taberlet et al., 2012).

2.2. Environmental DNA Advantages and Disadvantages

eDNA metabarcoding, despite being a promising methodology in the detection of multiple species in a single sample, has its own intrinsic/inherent disadvantages. eDNA is mostly a cost-effective and rapid approach, can detect a myriad of different species with increased taxonomic resolution, and is a non-invasive method; however, several considerations and limitations are needed to be taken into account. For instance, the DNA amount in the environmental samples exists in relatively low quantities (Turner et al., 2014), and to represent the taxa of interest accurately, the sample preservation should be ensured by having replicates and diminish the possibility of contamination and false positives (Ruppert et al., 2019). Primer bias is another issue in eDNA metabarcoding due to the limitations that arise from the employment of universal primers, which may not accurately and equally amplify the DNA fragments of all the members in a taxonomic class or subclass in the characterization of biodiversity (Eble et al., 2020; Ruppert et al., 2019). Another limitation in the study of eDNA metabarcoding is the lack of a high-quality reference database, especially in the areas where the universal public databases might not be adequate, especially if the species are not historically well-studied. In this case, creating a local reference database might be necessary to be able to characterize the biodiversity with eDNA metabarcoding with high accuracy. Interspecific sequence homology is another limitation of eDNA metabarcoding, resulting in a lower resolution of the taxonomy of the samples. One way to overcome the primer bias and interspecific sequence homology is by using multiple primers specific to the taxonomic groups of interest (Stat et al., 2017). Besides the technical limitations of eDNA metabarcoding, in the case of conservation and biodiversity monitoring, eDNA cannot be used as a quantitative method. Multiple species detection and characterization can be achieved; however, eDNA metabarcoding fails to inform on the abundance of each species or their population structure and demographics (Valentini et al., 2016). However, overall, eDNA metabarcoding is still a promising method in the characterization of biodiversity and has the potential to be utilized in the conservation ecology and biomonitoring since it has detected invasive (Dejean et al., 2012b), elusive, and cryptic species and in most cases was on par with the traditional surveying methods or detected even higher levels of biodiversity in comparison (Thomsen et al., 2012; Thomsen & Willerslev, 2015).

2.3. Environmental DNA Metabarcoding Compared with Traditional Methods

A recent meta-analysis shows that eDNA metabarcoding is a more sensitive and precise method outperforming traditional methods of surveying macro-organisms in the case of biodiversity assessments (Fediajevaite et al., 2021). However, as mentioned above, eDNA metabarcoding in most cases depends on qualitative analyses (i.e., the presence/absence data), and cannot draw conclusions for quantitative data such that no information can be derived on the structure of the populations or the abundance and biomass of the species in the habitat (Ruppert et al., 2019; Valentini et al., 2016). Despite its limitations, eDNA can be used as a complementary methodology in biodiversity assessments where traditional methods might fall short in the detection of invasive, elusive, and cryptic species. Moreover, eDNA metabarcoding can be improved not only for species detection but also for the estimation of abundance by increasing amplification efficiency through the use of taxa-specific primers (which can overcome primer bias), applying consistent PCR protocols (cycles lower than 35, using mock communities to test the correlation between the number of reads and the species abundance) accordingly, and regular spatial and temporal sampling to prevent the overrepresentation of taxa which originally have lower abundance (Kelly et al., 2019). However, primer bias cannot be overcome in metabarcoding studies because, by itself, it can vary among the taxa with up to four orders of magnitude despite the taxa having similar abundances (Elbrecht & Leese, 2015).

2.4. Case Studies of eDNA Metabarcoding in Fish

In this section, a review of the studies which compare eDNA metabarcoding and traditional methods of fish biodiversity is included. The first study from 2012, which employed eDNA metabarcoding to detect marine fish biodiversity by Thomsen et al., showed that eDNA metabarcoding by detecting 15 temperate coastal species outperformed or were equaled to nine conventional surveying methods which were conducted once every year between 2009 and 2011 such as fish pots, angling, push net, fyke nets, multi-mesh gill net, day snorkeling, beach seine, bottom trawl and night-snorkeling represented in increasing order in performance (2012). The bottom trawling and night-snorkeling had the best performance, detecting 13.3 and 14.7 species, respectively, among the other traditional surveys compared to eDNA metabarcoding, which detected 15 species and was conducted in 2011 (Thomsen et al., 2012). In the study, they chose two generic primers and four species-specific primers targeting fragments shorter than 100 bp in the cytochrome b gene. In order to ensure that the eDNA samples were not dispersed from somewhere else via ocean currents, they designed an experiment by sampling 50L of seawater and sub-sampled it for 15 days and checked the degradation rates of eDNA from two fish species by qPCR (Thomsen et al., 2012). The

degradation rates ([DNA] pr 400 ml of seawater/day) of eDNA for two fish species were 0.9 and 6.7 days (reaching beyond detectable levels by qPCR), disregarding environmental conditions such as salinity, water temperature, wind, and water current, which indicated that the identified species by eDNA metabarcoding was from local populations (Thomsen et al., 2012).

In another study, the deep seawater from the sub-arctic region was sampled at depths of 188±918 m to assess the fish biodiversity and compare to the findings of the trawling that was carried out simultaneously with the water sampling (Thomsen et al., 2016). In this study, the researchers used tagged teleo primers targeting the 12S rRNA gene that was proved to be successful in a previous study (Thomsen et al., 2016; Valentini et al., 2016). They compared the eDNA and bottom trawling results and found that twenty-six families of fish were detected in both surveys (Thomsen et al., 2016). Three families were only present in eDNA, and two families were detected only by trawling survey (Thomsen et al., 2016). The species which were only detected in eDNA metabarcoding were pelagic or preferred shallow waters which cannot be detected by bottom trawling (Thomsen et al., 2016). The ones that were detected in trawling only were very rare species that were detected as one individual in a single sampling only (Thomsen et al., 2016). Therefore, their DNA might probably have been below detectable levels or was nonexistent in the samples (Thomsen et al., 2016). Despite being detected in several eDNA samples, the Greenland shark species was only present in one trawling sample, which indicates that these species can escape the trawls and makes them and similar species with larger body size less detectable in trawling (Thomsen et al., 2016).

Yamamoto et al., in a species-rich coastal line in Maizuru Bay, Japan, conducted an eDNA metabarcoding-based biodiversity assessment of fish species (2017). The results of eDNA metabarcoding were compared to a 14-year period of underwater visual census (UVC) records (Yamamoto et al., 2017). The chosen primers were MiFish-U/E targeting a hyper-variable region in the 12S rRNA gene in fish, and they were able to detect 112 marine fish species (Miya et al., 2015; Yamamoto et al., 2017). Compared to eDNA metabarcoding, 140 censuses conducted in 14-year period had detected 80 marine fish species (Yamamoto et al., 2017). However, eDNA metabarcoding detected 40 species among the 80 species detected by UVC, with only 65 species having barcodes in the reference database (Yamamoto et al., 2017). They argued that seasonal migratory species may not have been present during their sampling in the bay, which might explain the absence of species DNA in the water samples (Yamamoto et al., 2017). Moreover, eDNA metabarcoding detected more than 20 fish species that had not been recorded by visual surveys before. Overall, eDNA metabarcoding detected a higher number of species in the area (Yamamoto et al., 2017). Even though there were still species that were only present in visual surveys, spatial and temporal water sampling could provide

a better census in detecting and monitoring fish biodiversity (Yamamoto et al., 2017). Simultaneous comparison of eDNA sampling and visual surveys might have demonstrated a better comparison since the visual survey data had seasonal depth.

Another study comparing eDNA metabarcoding and UVC in coral reefs conducted in two species-rich sites in Colombia used three primer sets that target the 12S rRNA gene (Polanco Fernández et al., 2020). Two of the primers were designed for teleosteans and chondrichthyans teleo/Teleo1 and Chon01, respectively, and one generic primer that targeted vertebrates Vert01 (Polanco Fernández et al., 2020). They compared the species detected by eDNA metabarcoding to UVC surveys that were conducted multiple times in previous years (Polanco Fernández et al., 2020). UVCs previously detected 113 species in one of the sites, Providencia; however, only 35 of them were also detected by eDNA metabarcoding (Polanco Fernández et al., 2020). In the other study site, Gayraca Bay, out of 57 species observed by UVCs, 16 of them were also found in eDNA samples (Polanco Fernández et al., 2020). However, 72 and 85 species were present in eDNA but absent in UVCs in the regions, respectively (Polanco Fernández et al., 2020). Moreover, most of these species were known as cryptobenthic species, which are hard to detect by visual surveys (Polanco Fernández et al., 2020). The absence of species detected by UVC in eDNA samples was mainly due to the absence of 60% of the species detected by UVC from the reference database (Polanco Fernández et al., 2020). Furthermore, eDNA had higher taxonomic coverage than UVC, which was taxonomically clustered, indicating an ecological signal (Polanco Fernández et al., 2020).

Another recent study that compares eDNA metabarcoding by using 12S rRNA targeting MiFish primers (Miya et al., 2015) and UVC of coastal fish communities in Southeast Korea, which included two study sites, found a similar correlation with the study of Polanco Fernández et al. (2020), in terms of the species detected by both methods (Lee et al., 2022). In the study, among the 18 species detected only by visual surveys in both study sites, 13 of them could not be detected by eDNA metabarcoding due to the lack of reference sequences of these species in the GenBank database (Lee et al., 2022). The absence of the other five species in the eDNA metabarcoding was suggested to be the result of the sampling method's inefficiency because the visual surveys are conducted in a continuous manner, whereas eDNA samples are taken at 50-meter intervals (Lee et al., 2022). However, the species identified solely by eDNA metabarcoding were pelagic, quick-swimming, or species that dwelled in rock crevices and cracks, explaining the inability of UVC surveys to detect cryptic and elusive species (Lee et al., 2022).

Lamy et al. (2021), compared the fish community beta diversity in a kelp forest in California, as detected by eDNA and UVC and found that the eDNA approach demonstrated a more detailed spatial structure of fish communities. For eDNA metabarcoding, the primer of choice was MiFish/U, targeting the 12S rRNA hypervariable region (Lamy et al., 2021; Miya et al., 2015). Even though at the species level, the two methods had similar success, at the genus and family levels, eDNA metabarcoding comprised more taxa (Lamy et al., 2021). eDNA metabarcoding was not able to detect 11 taxa that UVC was able to detect. Five of these species were absent from the reference database, and one was eliminated due to its low number of reads (Lamy et al., 2021). The other five taxa, despite being present in the reference database, were not detected, most probably due to their low biomass contribution in the habitat except for one abundant fish species (Lamy et al., 2021). The taxa that were detected by only eDNA were mostly pelagic and cryptobenthic species, some of which had never been observed by visual surveys before (Lamy et al., 2021).

Overall, both methods have their inherent disadvantages when compared to another. However, in most of the studies, the traditional surveys and eDNA metabarcoding do not have a high overlap in detected taxa indicating that conventional methods and eDNA can be used as complementary methods. However, when analyzed at genus and family levels, eDNA metabarcoding usually has a higher taxonomic coverage at all phylogenic levels and is not concentrated on specific taxa. If multiple markers are combined, and the reference databases are improved, eDNA metabarcoding is a promising approach in spatial and temporal biomonitoring of species richness.

3. MATERIALS AND METHODS

3.1. Sample Collection

Samples were taken from 16 zones in Kaş-Kekova MPA: Kekova Island coasts (n=2), Kaş coasts (n=14) (Table 3.1). For each sample, approximately 4.5 liters of water was filtered on site through Sterivex™ (Merck Millipore) filters by a peristaltic pump. Five samples were taken from the surface, 12 samples from 12m depth, and three samples from 24m depth. A total of 7 blank samples were taken throughout the sampling. Before the sampling, the equipment was sterilized using a 10% bleach solution and rinsed with 70% ethanol. After each filtration, 2 ml of PBS solution was passed through the filter cartridge using a syringe to rinse and prepare the sample. Subsequently, 2 ml of DNA/RNA Shield™ solution (Zymo Research) was introduced into the filter cartridge using a syringe to stabilize the genetic material. The prepared samples were then stored at -80°C until further processing. The simultaneous UVCs were conducted on 500m transects at 16 stations and three depths i.e. surface, 12m, and 24m. The visual counts were performed based on a predetermined checklist which includes 66 taxa, 62 of them at the species, two of them at the genus and two of them at the family level. The seawater samples were collected at the end of each transect.

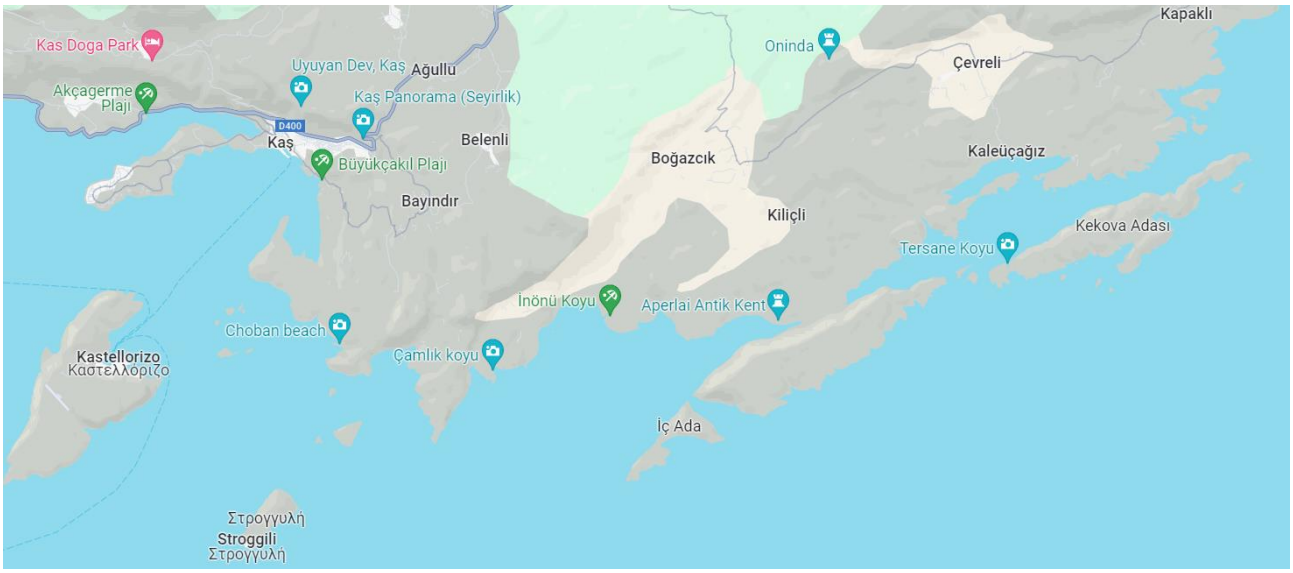


Figure 3.1. The map of the study area Kaş-Kekova MPA.

Table 3.1. The sampling dates, the locations and the renamed samples after the biological replicates were merged.

Sample No	Date	Zone	Depth	Samples Renamed Sample #-Z#-Depth
1	9.07.2022	13	surface	S1-Z13-S
2	9.07.2022	13	12	S2-Z13-12
3	9.07.2022	11	24	S3-Z11-24
4	10.07.2022	16	12	S4-Z16-12*
5	10.07.2022	61	24	S5-Z61-24
6	11.07.2022	23	24	S6-Z23-24
7	11.07.2022	39	12	S7-Z39-12
8	11.07.2022	39	surface	S8-Z39-S
9	12.07.2022	10	12	S9-Z10-12
10	12.07.2022	6	12	S10-Z6-12
11	14.07.2022	98	12	S11-Z98-12
12	14.07.2022	97	12	S12-Z97-12
13	15.07.2022	48	surface	S13-Z48-S
14	15.07.2022	49	12	S14-Z49-12
15	15.07.2022	48	12	S15-Z48-12
16	16.07.2022	15	12	S16-Z15-12
17	16.07.2022	16	12	S4-Z16-12*
18	17.07.2022	63	surface	S17-Z63-S
19	17.07.2022	63	12	S18-Z63-12
20	18.07.2022	4	surface	S19-Z4-S
21	18.07.2022	3	12	S20-Z3-12

*Biological replicates

3.2. DNA Extraction

The DNA extractions were carried out in the UV hood and all the equipment and surfaces are sterilized by 10% bleach solution and rinsed with 70% ethanol. The protocol to open up the Sterivex™ (Merck Millipore) filter casing was adapted from Cruaud et al. (2017). The DNA was extracted from the Sterivex™ (Merck Millipore) filters with DNeasy Blood & Tissue Kit (Qiagen), with the protocol adapted from Spens et al. (2016). The DNA concentrations were measured by Qubit dsDNA HS Assay (Invitrogen, CA, USA), and the extracted DNA was stored at -20°C. The adapted step-by-step protocols for the DNA extraction are further detailed in the Appendix.

3.3. DNA Amplification and Library Preparation (Two-Step PCR)

Two sets of universal markers, COI (Leray et al., 2013) and 12S rRNA (Riaz et al., 2011), were used. A two-step PCR protocol was used for both regions for library preparation (adapted from 16S

Metagenomic Sequencing Library Preparation (© 2019 Illumina, Inc)). The primers that were used for first- and second-step PCRs are shown in Table 2.

Table 3.2. The list of primers that are used in the two-step PCR protocol. (Green: Overhang sequences used in 16S rRNA library preparation, Blue: Illumina adapter sequences, Red: Index sequences)

mlCOIintF (Leray et al., 2013)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGWACWGGWT GAACWGTWTAYCCYCC
dgHCO2198 (Leray et al., 2013)	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTAAACTTCAGG GTGACCAAARAAYCA
12S-V5F (Riaz et al., 2011)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTAGAACAGGCT CCTCTAG
12S-V5R (Riaz et al., 2011)	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTAGATACCCC ACTATGC
Forward Primer for the Index PCR	AATGATACGGCGACCACCGAGATCTACACNNNNNNNNTCGTCGG CAGCGTCAGATGTGTATAAGAGACAG
Reverse Primer for the Index PCR	CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTCTCGTGGGCT CGGAGATGTGTATAAGAGACAG

The first round of PCRs was performed in a total volume of 25 µl reaction tubes for each primer set. For the primers targeting the COI region (mlCOIintF & dgHCO2198R), the PCR mix consisted of 12.5 µl of GeneDireX OnePCR™, 1.2 µl of each primer (10 µM), 7.6 µl of Invitrogen™ UltraPure™ water, and 2.5 µl of DNA isolates. The PCR conditions for the reaction were as follows: initial denaturation at 95°C for 3 minutes followed by 35 cycles of denaturation at 95°C for 10 seconds, annealing at 48°C for 30 seconds, and extension at 72°C for 45 seconds and final extension at 72°C for 5 minutes.

The reaction mix for the 12S primers (12S-V5-F & 12S-V5-R) included 12.5 µl of GeneDireX OnePCR™, 0.5 µl of each primer (10 µM), 6.5 µl of Invitrogen™ UltraPure™ water, and 5 µl of DNA isolate. The PCR protocol comprised an initial denaturation at 95°C for 2 minutes, 35 cycles of 1 minute of denaturation at 95°C, 30 seconds of annealing at 55°C, 1 minute of elongation at 72°C and 7 minutes of final extension at 72°C.

After the 1st step of PCR, amplification was validated by running the samples on a 1% agarose gel prepared with 40 ml TAE buffer, 0.4 g agarose, and 4 µl of RedSafe™ (90V for 35 minutes). After the validation of the PCR products, the index PCR was performed in a total volume of 25 µl

consisting of 12.5 µl of GeneDireX OnePCR™ (GeneDireX, Taiwan), 0.5 µl of each i7 and i5 index primers (10 µM), 11 µl of Invitrogen™ UltraPure™ water, and 0.5 µl of the first PCR products. The PCR conditions for the reaction were as follows: the initial denaturation at 95°C for 3 minutes followed by 10 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds, extension at 72°C for 30 seconds, and the final 5-minute extension at 72°C. The second PCR products were purified with Beckman Coulter™ Agencourt AMPure XP beads according to the manufacturer's protocol.

3.4. Library Pooling and DNA Sequencing

The concentrations of the purified 2nd PCR products were measured by Qubit dsDNA HS Assay (Invitrogen, CA, USA). The concentrations of each library were calculated in nM and normalized to 30nM by dilution. Next, the libraries were pooled together and sent to sequencing to Gen-Era Diagnostik (İstanbul, Türkiye). A paired-end (2x150) sequencing was carried out on an Illumina NovaSeq sequencer, targeting 1,200,000 reads per sample.

3.5. Bioinformatics

The sequences were analyzed by APSCALE-GUI (version 1.2.1) (Buchner et al., 2022) and TaxonTableTools (version 1.4.8) (Macher et al., 2021). The bioinformatics pipeline is detailed further in this section, see below. The APSCALE-GUI pipeline incorporates VSEARCH (version 2.22.1) (Rognes et al., 2016) to merge the overlapping regions, if any, of the demultiplexed paired-end reads. For 12S rRNA metabarcoding the paired-end reads were merged. However, for COI reads, only R1 reads (150 bp length) were used as the amplicon size was more than twice 150 bp. After the paired-end merging step, the primers are trimmed by cutadapt (version 4.4) (Martin, 2011). The reads are filtered according to their quality and expected length via VSEARCH. The reads for each sample are dereplicated, and according to their abundance, reads with counts below four are eliminated. Then, the dereplicated reads for each sample are pooled together and dereplicated again. The reads are clustered together into individual OTUs, and chimeras are removed. The erroneous OTUs are detected and filtered by the LULU algorithm (Frøslev et al., 2017).

At the end of the LULU filtering step, a fasta file and a read table file are generated. From the fasta files, the OTUs were assigned to their taxonomies by using local BLAST (BLAST+ version 2.14.1) (Camacho et al., 2009), NCBI BLAST and BOLDigger software (Buchner & Leese, 2020). In this analysis, the three databases were combined for the COI marker, and NCBI BLAST was used

for the 12S rRNA marker. TaxonTableTools (version 1.4.8) (Macher et al., 2021) was used for the downstream analyses to create a taxonomy table and visualize the data. The taxonomic assignments from the three databases were generated by the JAMP Pipeline (Elbrecht et al., 2018), by the following thresholds to find the best hit, (98%: species level, 95%: genus level, 90%: family level, 85%: order level, <85%: class level). The algorithm to find the best hit works by finding the maximum similarity among the top 20 hits, and with the maximum similarity as the cutoff threshold, the hits with a similarity score below the threshold are eliminated. All individual classifications are counted and sorted by their abundance. The ones with missing values for taxonomic levels such as Order, Family are dropped. The most common hit with no missing values is found and the algorithm returns that hit. If there is no common hit, the threshold is set to the next higher level and the process is repeated until there is a hit.

3.6. Statistics

The statistical methods used in this study were executed in TaxonTableTools (version 1.4.8) (Macher et al., 2021). The alpha diversities were calculated as taxon richness (the number of species). The Principal Coordinate Analyses (PCoA) were performed based on Jaccard distance (Jaccard, 1912) in the TaxonTableTools, which incorporates the scikit-bio package and ANOSIM (Clarke, 1993). The Jaccard distance matrices were incorporated with PCoA, which is an ordination analysis that operates by using distance matrices to assign coordinates into two or three-dimensional space and help visualize the dissimilarities between groups (Gower, 2014). Jaccard Similarity Index (J) is calculated according to the formulae below:

$$J(A, B) = |A \cap B| / |AUB| \quad (3.1)$$

Here, A and B represent the two binary (presence/absence) datasets, $|A \cap B|$ represents the intersection, whereas |AUB| represents the union of the two datasets. The Jaccard similarities take values between 0 and 1, 0 indicating perfect similarity. The Jaccard distances are derived from the below equation:

$$D_J(A, B) = 1 - J(A, B) \quad (3.2)$$

The Jaccard distances take values $0 \leq D_J \leq 1$; in our case, the distances are calculated as matrices. $D_J = 0$ indicates perfect similarity, whereas $D_J = 1$ indicates perfect dissimilarity.

Another statistical method used in this study was the Analysis of similarities (ANOSIM), which is a non-parametric ANOVA-like test that uses permutations (n = 999), incorporating a ranked matrix of dissimilarities (Jaccard distance matrix) to test the difference between two or more groups. The null hypothesis of ANOSIM is that there are no dissimilarities between the groups. The R-value for the test statistics is calculated as follows:

$$R = (\bar{r}_B - \bar{r}_W)/(M/2) \quad (3.3)$$

$\bar{r}_B$ represents the average of rank similarities between the groups and $\bar{r}_W$ represents the average rank of similarities within the groups. The M value in the equation represents the number of sample pairs ($M = n(n-1)/2$). The p-value is calculated by the percentage that the R' values derived by the permutations exceed the actual R (Clarke, 1993).

4. RESULTS

4.1. The Validity of eDNA Metabarcoding

In order to check whether eDNA metabarcoding validates the species observed by the Underwater Visual Censuses (UVC), the species detected by the two gene markers were mapped to the 66 taxa that were in the UVC checklist. The Venn diagram below shows the species observed by UVCs and the eDNA approach in the 20 transects that the eDNA samples were collected from. Among the species in the checklist, 35 species were detected by the visual surveys, whereas 19 species were detected by eDNA. In general, 84% of the species detected by the eDNA metabarcoding was shared with the visual counts, and 46% of the species detected by the visual counts overlapped with the eDNA dataset. Three species were exclusively detected by eDNA metabarcoding even though there were in the checklist. Since some of the taxa were counted at genus and family level, the two Venn diagrams below demonstrate the taxa shared by both datasets at these levels. 80% of the genera detected by eDNA metabarcoding was shared by the UVC dataset. On the other hand, 93% of the families detected by eDNA metabarcoding was shared by the UVC dataset.

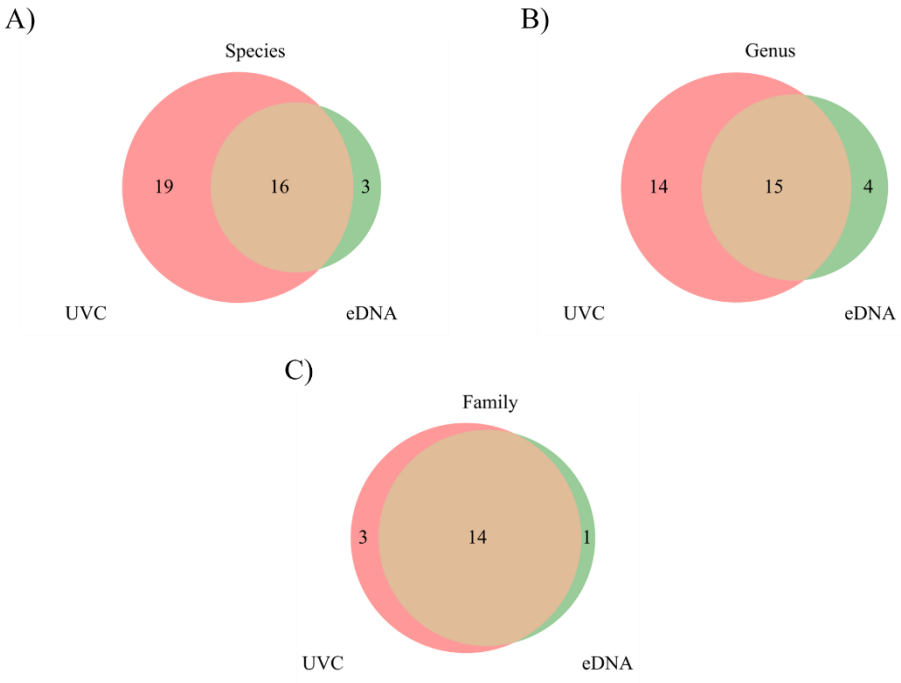


Figure 4.1. The number of taxa detected by both methods that were on the UVC checklist. A) at species, B) at genus, C) at family level.

Table 4.1. The species that were detected by the UVCs exclusively and their records in the NCBI database.

Species detected exclusively by the UVCs	**Records**
Caranx crysos	Complete Mitochondrion
Dasyatis pastinaca	Complete Mitochondrion
Diplodus vulgaris	COI & Partial 12S
Diplodus puntazzo	Complete Mitochondrion
Epinephelus marginatus	COI & Partial 12S
Lithognathus mormyrus	COI
Mullus surmuletus	Complete Mitochondrion
Mycteroperca rubra	COI & Partial 12S
Pagrus pagrus	COI & 12S
Pempheris rhomboidea	COI
Pteragogus trispilus	COI
Pterois miles	Complete Mitochondrion
Sarpa salpa	COI & Partial 12S
Scarus ghobban	Complete Mitochondrion
Seriola dumerili	Complete Mitochondrion
Serranus cabrilla	COI & Partial 12S
Serranus scriba	COI & Partial 12S
Spicara flexuosa	COI
Spicara smaris	COI
Symphodus tinca	COI
Thalassoma pavo	COI & Partial 12S
Upeneus pori	COI & Partial 12S
Vanderhorstia mertensi	COI

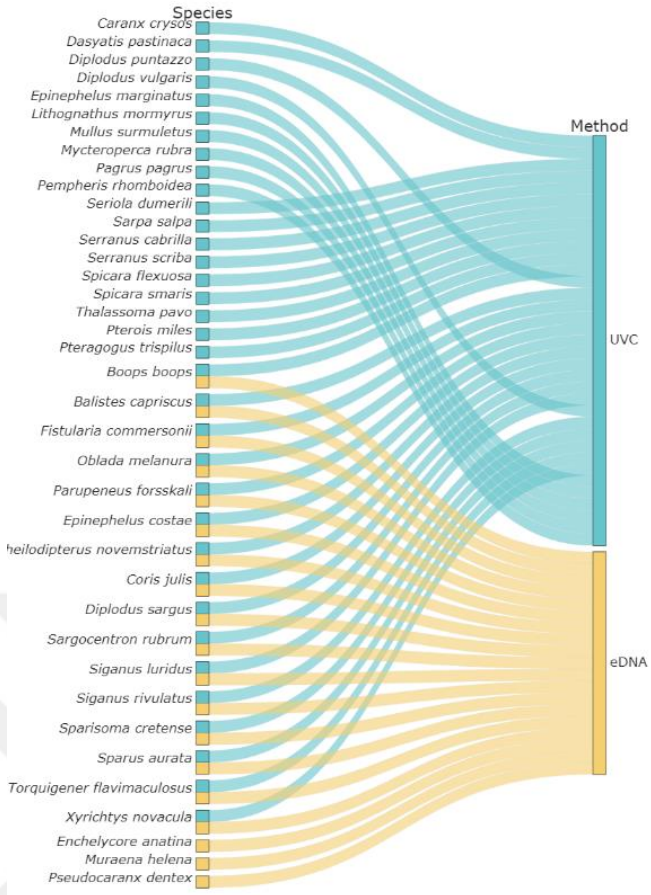


Figure 4.2. The distribution of species detected by the two methods normalized to the species of the UVC checklist.

4.2. Complementarity of eDNA Metabarcoding and Underwater Visual Census

The comparison between the fish species detected with the eDNA metabarcoding approach and the simultaneous Underwater visual censuses (UVCs) could shed light on whether they complement each other. Overall, visual surveys in the sampling stations spanning 48 transects detected a total of 42 taxa. Among these, five taxa were counted at the genus level and grouped together as *Atherina* & *Atherinomorus*, and the other group *Mugil* sp., *Liza* sp., *Chelon* sp. In comparison, eDNA metabarcoding detected 55 species at the sampling stations but did not scan all the depths and paths that were studied by UVCs. The Venn diagram below demonstrates the total number of fish genera detected by each method (Figure 4.3). In total, 39 genera were detected by UVCs and 48 by eDNA metabarcoding; both methods shared 20 genera. At the species level, 23 were detected exclusively by visual censuses, 38 by eDNA metabarcoding and 17 species were detected by both. When the species composition was compared based on location and depth, 16 species were detected by both methods, and 19 by UVC and 39 by eDNA exclusively.

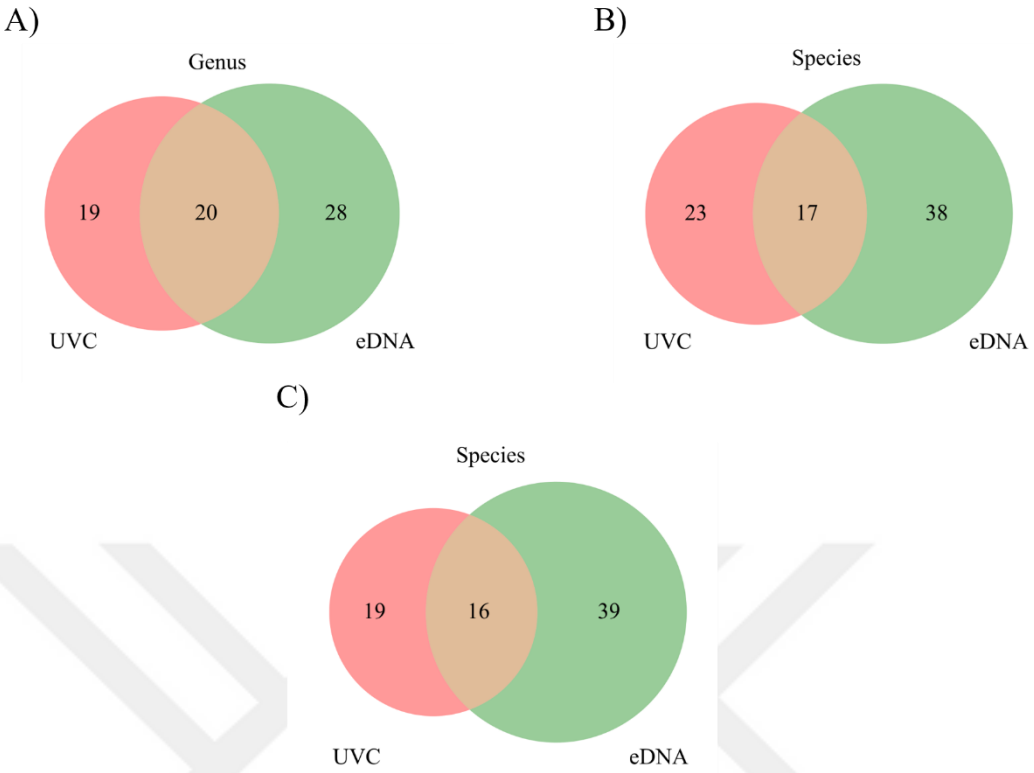


Figure 4.3. The total number of A) fish genera B) fish species detected in terms of species observed in each transect based on all observations in sampling locations. C) The fish species detected in terms of species observed at each transect with the same sampling locations and depth as the eDNA sampling. UVCs (red), eDNA (green), and both (orange).

When the number of species detected by both methods was compared pairwise based on the exact location and the depth of the eDNA samples (*i.e.* 20 transects for each method), the percentage of species shared by both methods ranged between 3% and 25% (Figure 4.4). On average, 11% of the species were shared by both methods. No fish species were detected at the station 3 depth 12m, by the eDNA metabarcoding.

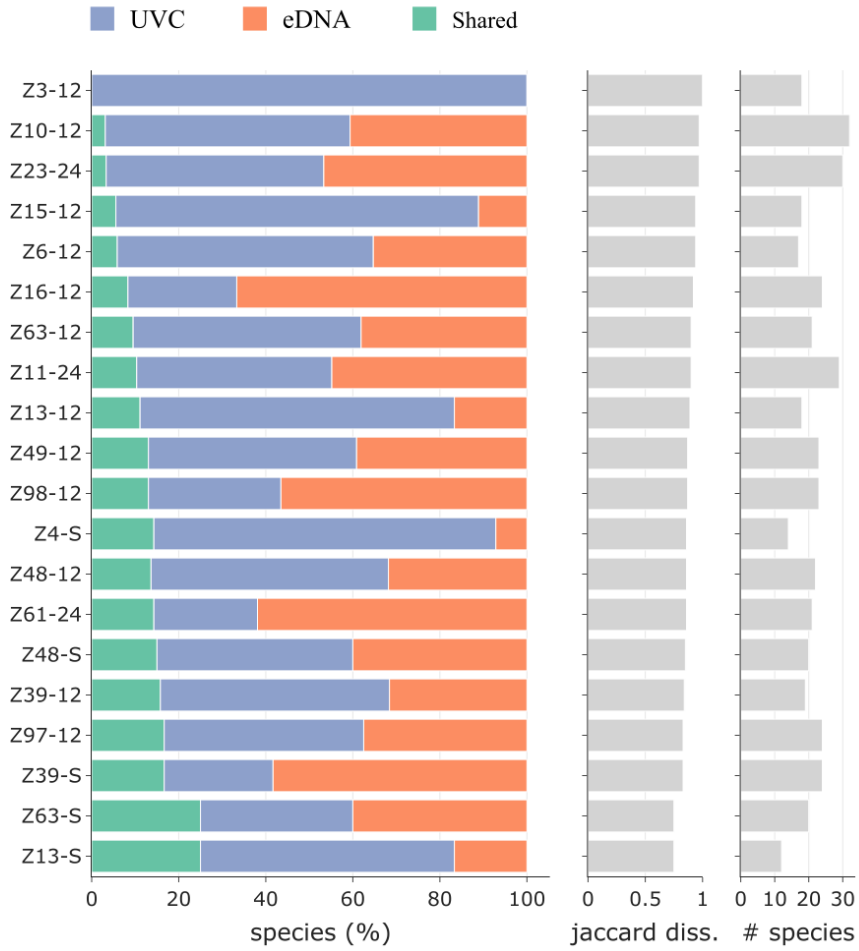
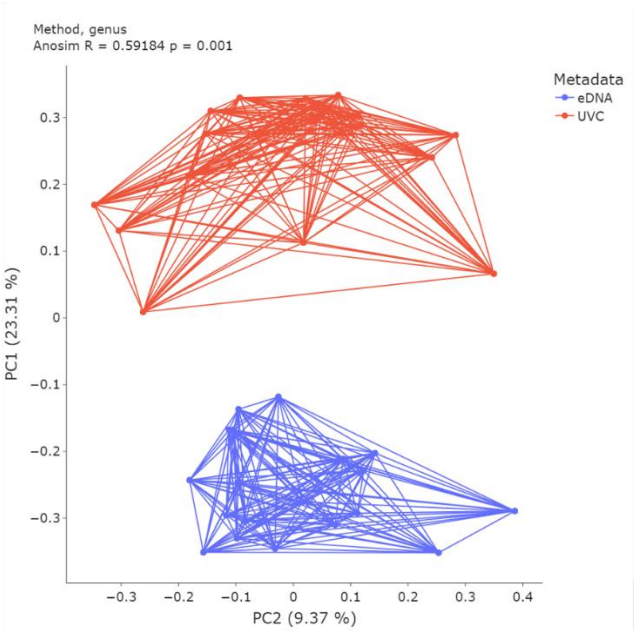


Figure 4.4. The pairwise comparison of the fish species that were identified by both methods. (The fish species identified by Blue: UVC, Red: eDNA, Green: Both)

The PCoA graph below demonstrates the distribution of coordinates accounting methodology as the grouping variable (Figure 4.5). The ANOSIM, when categories were interconnected, resulted in an R-value of 0.59 and p-value of 0.001 at the genus level and an R-value of 0.56 and p-value of 0.001 at the species level, indicating a significant and relatively strong dissimilarity in terms of fish species diversity detected by both methods.

A)



B)

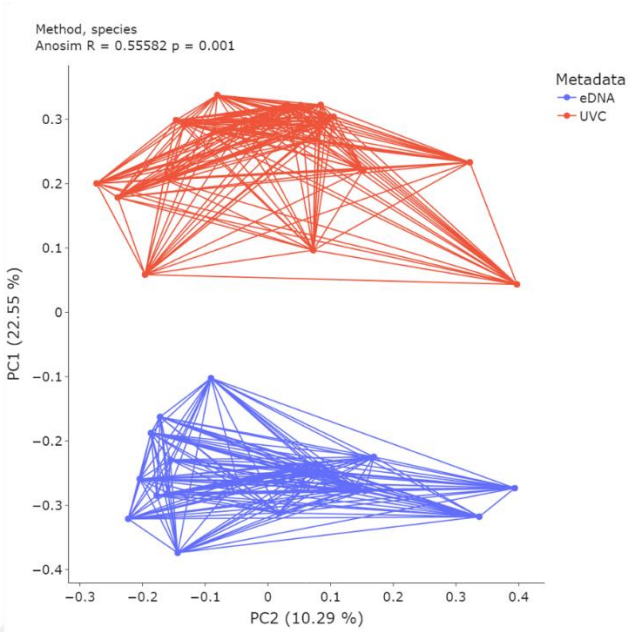


Figure 4.5. Principal Coordinates Analysis (PCoA) graph incorporating methodology as a factor A) at the genus level, B) at the species level.

To compare whether the species biodiversity was different in the UVCs conducted in terms of sampling location (Figure 4.6A) and depth (Figure 4.6B & Figure 4.6C), PCoA ordination analysis was used together with ANOSIM at the species level for location (Figure 4.6A) and depth (Figure 4.6B) -the same location and depth with the eDNA samples- and all sampling depths in the UVCs (Figure 4.6C). Similar to the PCoA with the COI dataset, depth as a factor resulted in a significant and moderate level of dissimilarity in both.

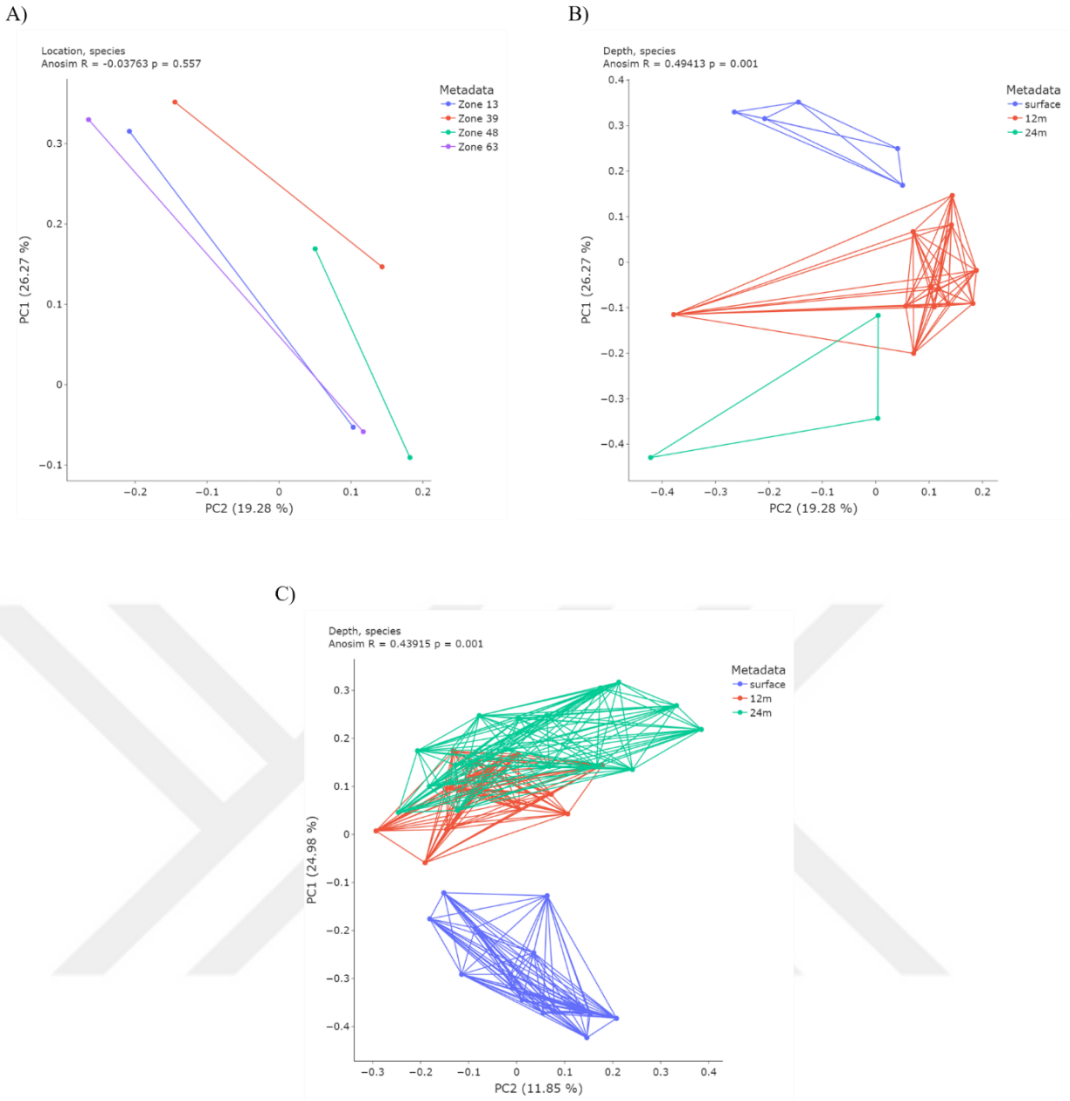


Figure 4.6. Principal Coordinates Analysis (PCoA) graph of UVC incorporating A) location as a factor and B) depth as a factor, C) depth as a factor for all the sampling depths, with categories interconnected.

4.3. Species Identification and Taxonomic Assignment Through Analysis of the 12S rRNA Gene Marker

At the end of the metabarcoding analysis pipeline, 217 OTUs were generated. After the filtration, 85% similarity cutoff and both filter blanks and negative controls being subtracted, the number of OTUs decreased to 134. The biological replicates, S4 and S17, were merged together, and OTUs assigned to *Homo sapiens* were subtracted. After the negative control subtraction and taxonomic filtration of human hits, 116 OTUs were left. Before the reads were normalized, the read counts per sample were checked and S02 had the lowest number of reads (n=8). Therefore, S02 was removed from the downstream analyses. Next, the reads were normalized based on minimum reads of 3588 per sample. Out of 1,321,123 total reads, the mean reads per sample was 69,532. However, the

samples below the normalization threshold were 14. After the removal of S02 the total number of OTUs left in the dataset were 115, and following the normalization of the reads based on the minimum reads per sample, the number of OTUs were 108.

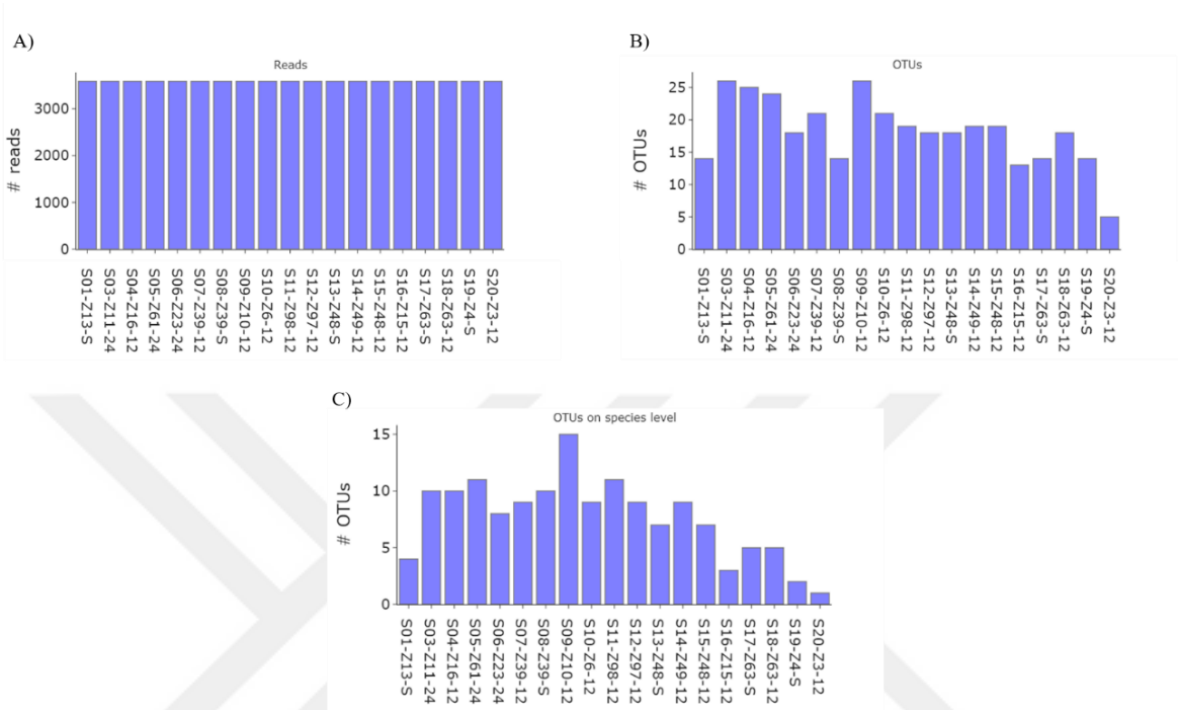


Figure 4.7. A) The number of reads per sample, B) the number of OTUs per sample and C) the number of OTUs assigned at the species level per sample, after data filtration (sample S02 excluded) and read normalization.

The number of taxa for the OTUs obtained from 12S rRNA gene markers targeting vertebrates is shown in Figure 4.8A. The taxonomic richness encompasses four classes, 29 orders, 40 families, 59 genera, and 43 species, with the primary classes identified as *Actinopteri*, *Aves*, *Mammalia*, and *Amphibia*. Figure 4.2B illustrates the number of OTUs obtained per each taxonomic level. The total number of OTUs assigned to taxonomy and which passed the filtration (cutoff = 85% similarity) was 108 (Figure 4.8B). Out of these, 45 OTUs reached species-level resolution, representing a taxonomic resolution of 42% (Figure 4.8B).

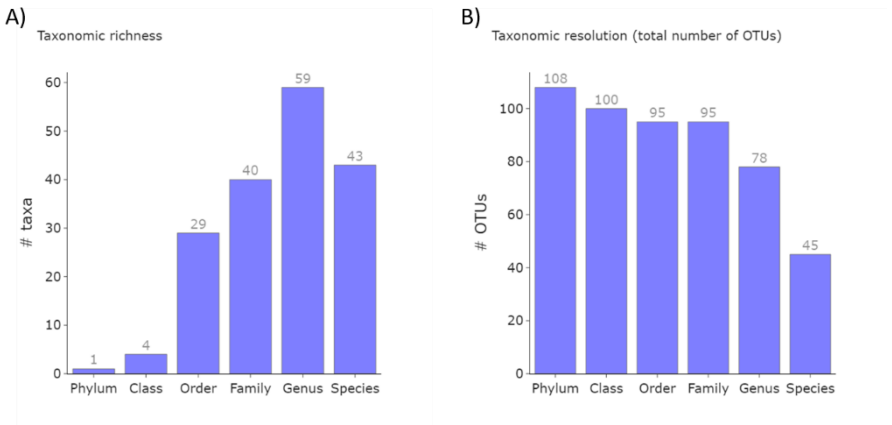


Figure 4.8. A) Taxonomic richness and B) taxonomic resolution for each taxon assigned to the OTUs generated from 12S rRNA primers after S02 is excluded and the reads are normalized to the minimum number of reads. The x-axis represents each taxonomic level from phylum to species. The y-axis shows the number of taxa for A), whereas for B), it represents the number of OTUs.

Figure 4.9 demonstrates the number of species assigned per each depth that the samples were taken in Kaş-Kekova MPA. The samples from 24 m had the highest diversity, whereas the samples taken from the surface had the lowest diversity.

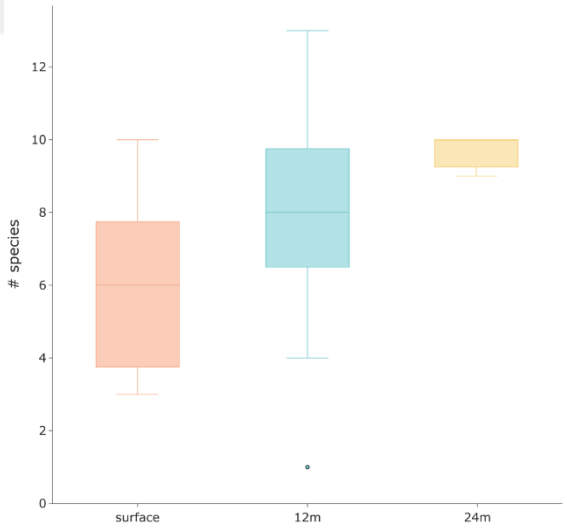


Figure 4.9. Alpha diversity box plot illustrating number of species assigned by 12S primers per each sampling depth at the Kaş-Kekova MPA.

The null hypothesis of ANOSIM for our study design is that there are no differences in the species diversity among different groups after accounting for depth. Therefore, finding higher similarity or dissimilarity in locations with two different sampling depths is expected. However, for the three groups of samples in Zone 39, Zone 48, and Zone 63, the similarities or dissimilarities demonstrated in the PCoA graph in Figure 4.10, either the depth along the water column has varying

biodiversity or there is no effect of depth in species biodiversity, which might be the case for Zone 63. When depth was considered as the grouping variable, the R-value was closer to zero, indicating highly similar species composition among the groups, and there was not any significant dissimilarity ($p = 0.37$) among the samples with sampling depths of surface, 12m and 24m.

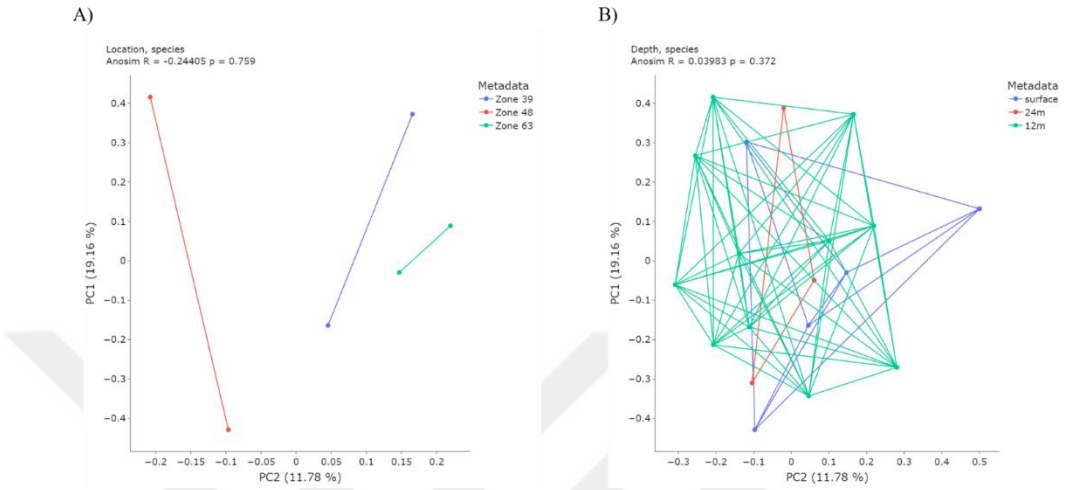


Figure 4.10. Principal Coordinates Analysis (PCoA) graph incorporating A) location as a factor and B) depth as a factor, with categories interconnected.

The taxonomic composition of each sample at the class level is shown in the bar chart below (Figure 4.11). The predominant class in all samples was *Actinopteri* except S20, in which the predominant class was *Amphibia*. Besides *Actinopteri* and *Amphibia*, *Mammalia* and *Aves* classes were also present in the samples. *Aves* was only present in S13. In most of the samples the read proportions of *Actinopteri* were above 50%.

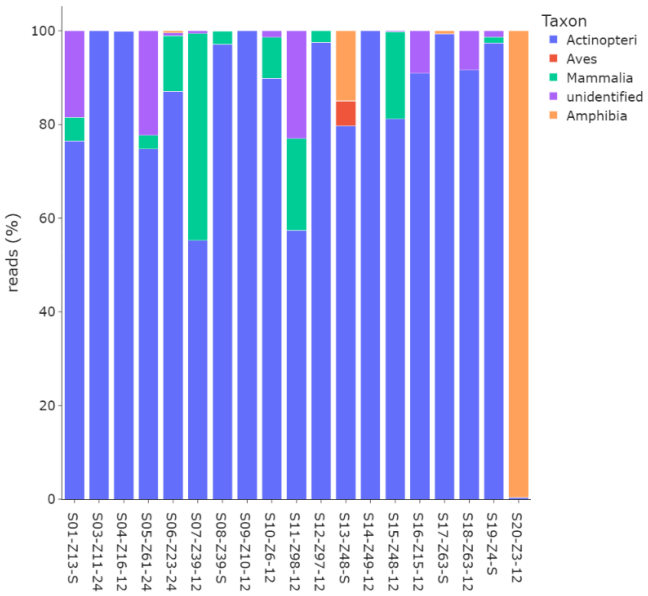


Figure 4.11. The bar chart representing the class proportions.

When the species level is considered, the species with the highest abundance percentage among the samples were *Atherina boyeri* (14.9 %), followed by *Sparus aurata* (12.9 %), *Parupeneus margaritatus* (9.37 %), *Dicentrachus labrax* (6.58 %), *Gobius niger* (5.9 %), *Sardina pilchardus* (5.21 %). Among 45 species, 24 of them were below the 1% abundance threshold. As for the taxonomic composition, 42 species were in the class *Actinopteri*, ray-finned fishes.

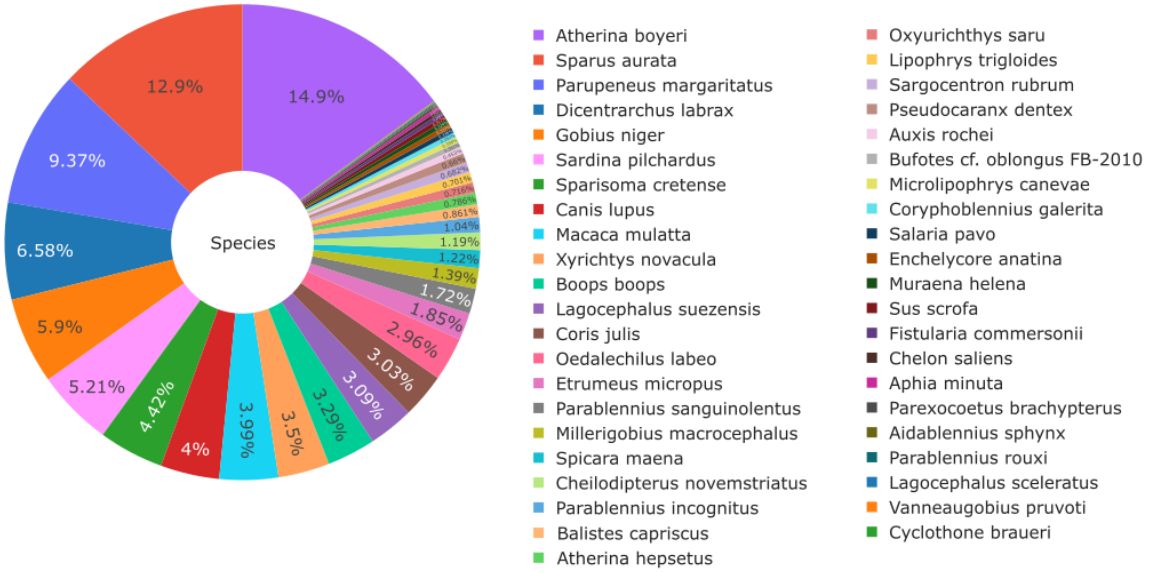


Figure 4.12. The pie chart demonstrating the species proportions.

4.4. Species Identification and Taxonomic Assignment Through Analysis of the COI Gene Marker

Following the completion of the metabarcoding analysis pipeline, an initial count of 8,359 operational taxonomic units (OTUs) was obtained from the databases NCBI BLASTN, Midori2, and BOLD. Subsequently, after applying an 85% similarity cutoff and filtering out *Homo sapiens* hits and OTUs with no matches, the OTU count was reduced to 3,086. The biological replicates, S4 and S17, were then combined, and the dataset was then normalized based on a minimum of 52,678 (Figure 4.13A) reads per sample out of a total of 6,665,166 reads, with a mean of 333,258 reads per sample. The average percentage of discarded OTUs during normalization was 15.69%. None of the samples fell below the normalization threshold, and after normalization, the final count of OTUs was 2,848. The samples with the fewest OTUs were identified in S19 and S20, as depicted in Figure 4.13B and Figure 4.13C.

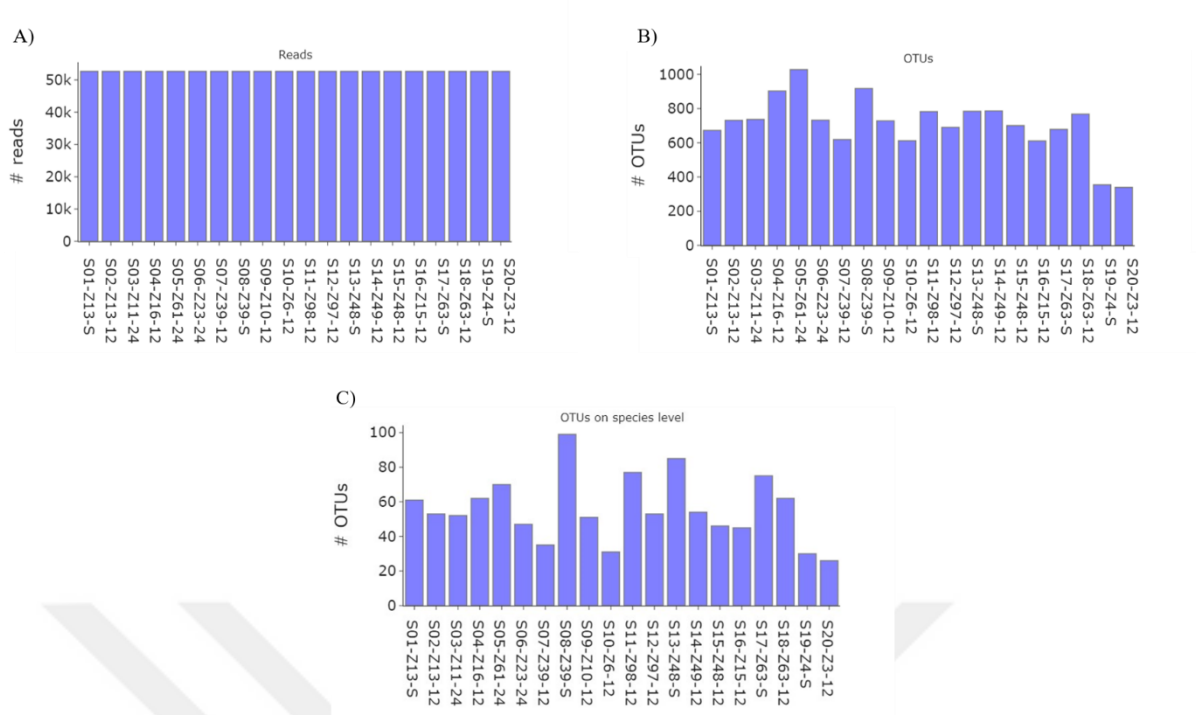


Figure 4.13. A) The number of reads per sample, B) the number of OTUs per sample and C) the number of OTUs at the species level per sample, after data filtration and read normalization.

The number of taxa for the OTUs obtained from COI primers targeting eukaryotes is shown in Figure 4.14A. The taxonomic richness encompassed 62 phylum, 113 classes, 338 orders, 520 families, 339 genera, and 255 species. The prominent phyla identified include *Arthropoda* (46.3 %), *Mollusca* (9.53 %), *Rhodophyta* (9.48 %), *Pseudomonadota* (5.51 %), and *Chordata* (5.11 %). Figure 4.14B illustrates the distribution of OTUs across different taxonomic levels. The total number of OTUs that were assigned to a specified taxonomic level and passed the filtration (cutoff = 85% similarity, uncultured phyla discarded) was 2668 (Figure 4.14B). Out of these, 290 OTUs reached species-level resolution, representing a taxonomic resolution of 11% (Figure 4.14B).

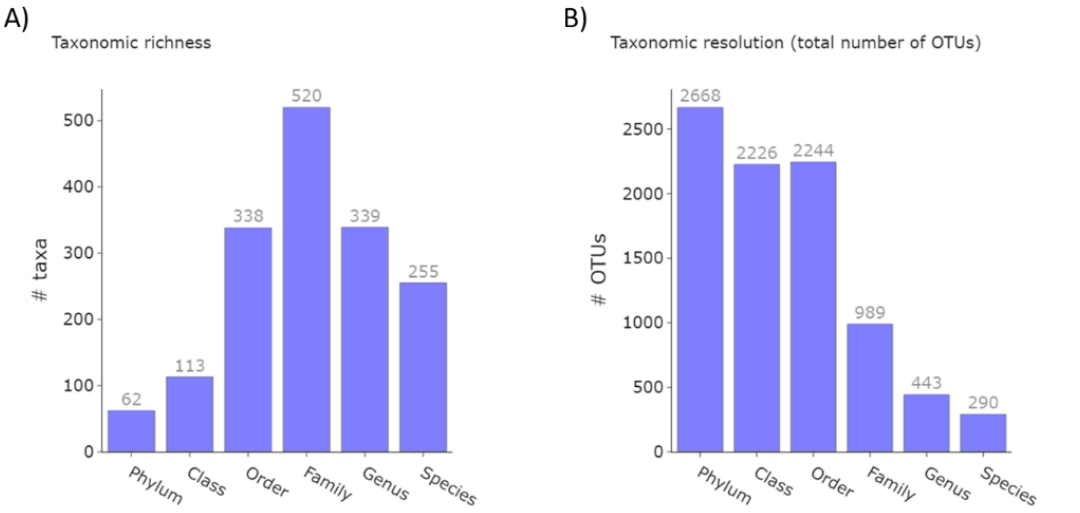


Figure 4.14. A) Taxonomic richness and B) Taxonomic resolution for each taxon assigned to the OTUs generated from COI primers. The x-axis represents each taxonomic level from phylum to species. The y-axis shows the number of taxa for A), whereas for B), it represents the number of OTUs.

Figure 4.15 demonstrates the alpha diversity measured as taxon richness, which, in this case, is the number of species assigned to each depth where the samples were taken in the Kaş-Kekova MPA. The highest diversity was found at the samples taken from the surface with the highest variation.

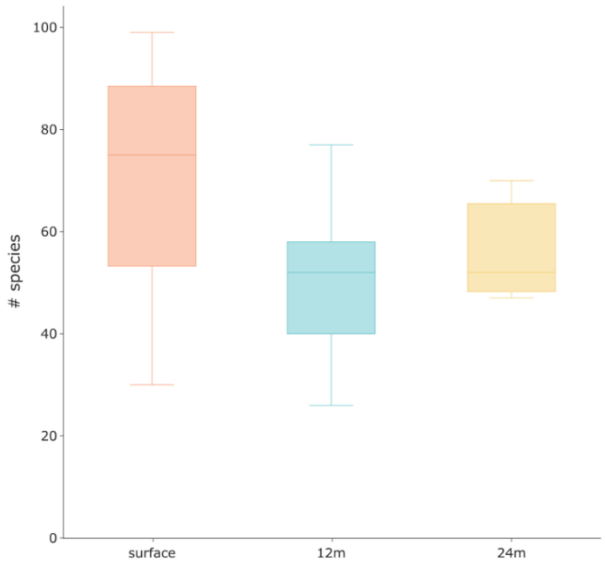


Figure 4.15. The alpha diversity box plot illustrates the number of species assigned by COI primers to each sampling depth at the Kaş-Kekova MPA.

The ANOSIM based on the PCoA, when accounting for location as a factor and when the groups are interconnected, resulted in an R-value of 0.11 and p-value of 0.33, meaning that the samples are relatively similar and have no significant dissimilarity. However, the samples with the same sapling depth seem to have been clustered together. In Figure 4.16A, the coordinates on the upper left corner are of the surface, whereas the coordinates on the bottom right corner are from 12m, indicating dissimilarity between different sampling depths. In order to check whether there was a difference in the species composition among different sampling depths, the PCoA was performed while taking depth as the grouping variable. Testing depth as a factor, the ANOSIM R-value was 0.42 with a p-value of 0.003, suggesting moderate and statistically significant dissimilarity (Figure 4.16B).

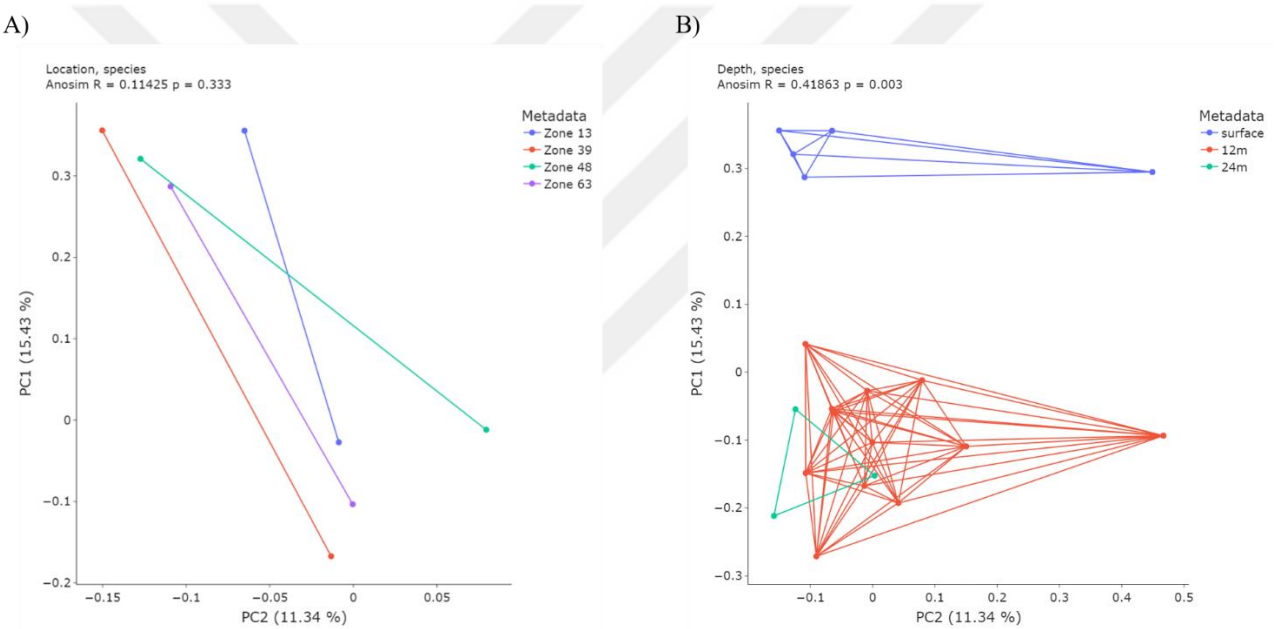


Figure 4.16. Principal Coordinates Analysis (PCoA) graph incorporating A) location as a factor and B) depth as a factor, with categories interconnected.

The phylum proportions per sample with a cutoff of 1% abundance of the COI dataset are given in Figure 4.17 below. The number of phyla with an abundance higher than 1% was 11 among the 63 total phyla detected, indicating that 17% of the phyla were above the 1% threshold, whereas 83% of the phyla had abundance below 1%. Other than sample S19, the phylum with the highest abundance was Arthropoda in all of the samples. In S19, the highest abundance was found in the phylum Echinodermata.

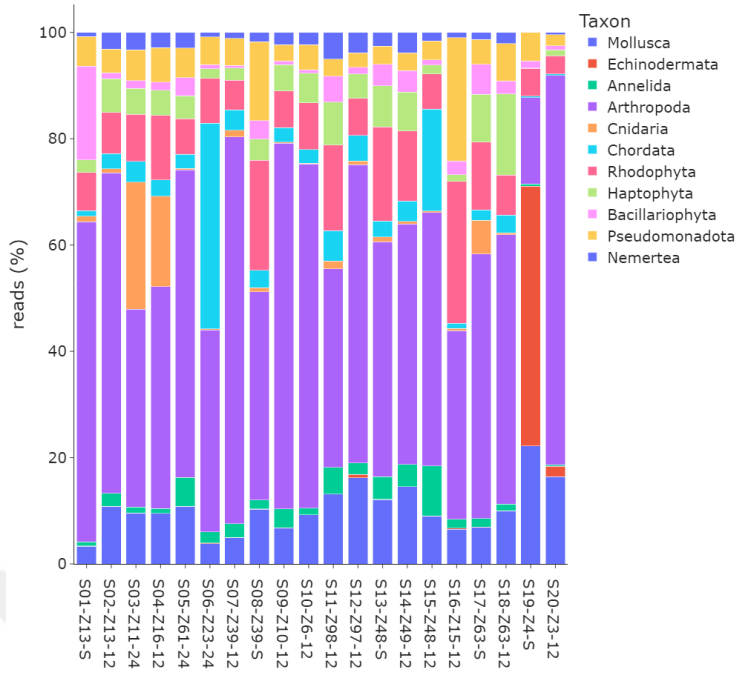


Figure 4.17. The bar chart demonstrating the phylum proportions (cut-off 1% abundance).

4.5. Comparison of the Fish Biodiversity Characterized by Both Primers

In order to compare the fish species diversity assigned by both primers, we first checked the species diversity found at the *Chordata* level to have an overall overview of the differences between the two datasets. The alpha diversity was calculated as the taxon richness, in this case, the number of species in each dataset at the *Chordata* level, and it is depicted in Figure 4.18 as a boxplot. The median for the 12S rRNA primer sets, which was around 11 species, was higher than that for the COI primer sets, which was around 7 species.

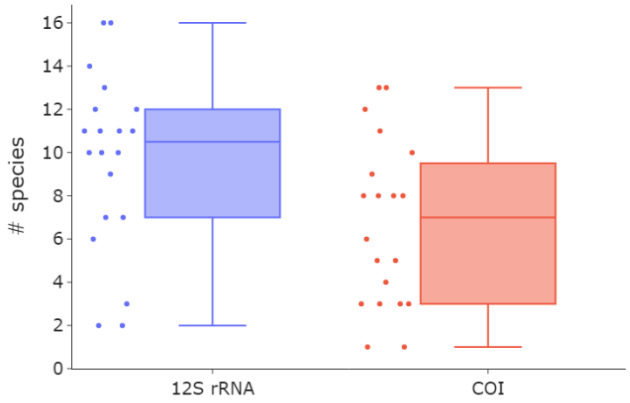


Figure 4.18. Alpha diversity boxplot demonstrates species diversity as the number of species in the *Chordata* phylum assigned by both primers.

Even though the generic primers used in this study target different taxa (Chordata and Eukaryota), we can still compare the two based on the common phylum *Chordata* and their accuracy in detecting fish species. This helps to understand whether the taxonomic resolution of each primer is complementary to each other in case of fish species diversity. The Venn diagram below (Figure 4.19) demonstrates the species-level biodiversity assigned by COI and 12S rRNA primers. Overall, the unique taxa at the species level that we could assign by COI primers were 36, and for 12S rRNA, it was 44. 11 of these species were shared by both datasets.



Figure 4.19. The number of species in the *Chordata* phylum that were detected by COI primers (red), 12S rRNA primers (green), and both primer sets (orange).

When we filter the species diversity down to fish species only, the Venn diagram below (Figure 4.20) depicts the overall picture. The number of fish species detected by the 12S rRNA primers was 39, whereas for the COI primers the value was 26. The number of fish species shared by both primer sets was 10. In general, 38% of the species detected by the COI primers overlapped with the 12S rRNA primers and 26% of the species detected by the 12S rRNA primers were shared.



Figure 4.20. The number of fish species that were detected by 12S rRNA primers (green), COI primers (red), and both primer sets (orange).

The species in the *Chordata* phylum identified by the generic primer sets for COI and 12S rRNA are given in Table 4.1.

Table 4.1. The species found in the *Chordata* phylum detected by primers COI, 12S and both.

Species 12S Only	**Species Shared**	**Species COI Only**
Fistularia commersonii	*Millerigobius macrocephalus*	*Poecilia wingei*
Lagocephalus sceleratus	*Lagocephalus suezensis*	*Noctilio leporinus*
Oxyurichthys saru	*Sparus aurata*	*Mullus barbatus*
Spicara maena	*Sparisoma cretense*	*Siganus luridus*
Aphia minuta	*Coris julis*	*Gobio gobio*
Vanneaugobius pruvoti	*Enchelycore anatina*	*Acomys russatus*
Atherina hepsetus	*Lipophrys trigloides*	*Diplodus sargus*
Xyrichtys novacula	*Muraena helena*	*Tachymenoides affinis*
Coryphoblennius galerita	*Canis lupus*	*Nyctimystes infrafrenatus*
Parablennius rouxi	*Sargocentron rubrum*	*Callionymus filamentosus*
Gobius niger	*Balistes capriscus*	*Herdmania momus*
Salaria pavo		*Chromis chromis*
Auxis rochei		*Hygophum benoiti*
Parablennius incognitus		*Cynoglossus sinusarabici*
Microlipophrys canevae		*Ascidia ahodori*
Bufotes cf. oblongus FB-2010		*Torquigener flavimaculosus*
Chelon saliens		*Heteromycteris hartzfeldii*
Macaca mulatta		*Pipistrellus pipistrellus*
Sardina pilchardus		*Siganus rivulatus*
Etrumeus micropus		*Parupeneus forsskali*
Boops boops		*Epinephelus costae*
Atherina boyeri		*Alsodes vanzolinii*
Oedalechilus labeo		*Chelon labrosus*
Aidablennius sphynx		*Oblada melanura*
Pseudocaranx dentex		*Odorrana hosii*
Sus scrofa		
Cheilodipterus novemstriatus		
Parexocoetus brachypterus		
Dicentrarchus labrax		
Parupeneus margaritatus		
Cyclothone braueri		
Parablennius sanguinolentus		

The fish species that were represented in the Venn diagram above are given in detail in the Table 4.2 below. The possible invasive species are shown by the asterisk, and the species distribution data was checked by the FishBase database. Among the 55 species detected by both primers, 16 are possibly invasive with already established populations or were introduced recently.

Table 4.2. The fish species detected by primers targeting 12S rRNA and COI regions.

Fish Species 12S Only	**Fish Species Shared**	**Fish Species COI Only**
Aidablennius sphynx	*Balistes capriscus*	*Callionymus filamentosus*** (Oral, 2010)
Aphia minuta	*Coris julis*	*Chelon labrosus*
Atherina boyeri	*Enchelycore anatina*	*Chromis chromis*
Atherina hepsetus	*Lagocephalus suezensis*** (Oral, 2010)	*Cynoglossus sinusarabici*** (Oral, 2010)
Auxis rochei	*Lipophrys trigloides*	*Diplodus sargus*
Boops boops	*Millerigobius macrocephalus*	*Epinephelus costae*
*Cheilodipterus novemstriatus*** (Turan et al., 2015)	*Muraena helena*	*Gobio gobio*
Chelon saliens	*Sargocentron rubrum* (Oral, 2010)**	*Heteromycteris hartzfeldii**
Coryphoblennius galerita	*Sparisoma cretense*	*Hygophum benoiti*
Cyclothone braueri	*Sparus aurata*	*Mullus barbatus*
Dicentrarchus labrax		*Oblada melanura*
*Etrumeus micropus**		*Parupeneus forsskali*** (Ergüden et al., 2018)
*Fistularia commersonii*** (Golani, 2000)		*Poecilia wingei*
Gobius niger		*Siganus luridus*** (Sala et al., 2011)
*Lagocephalus sceleratus*** (Akyol et al., 2005)		*Siganus rivulatus ** (Sala et al., 2011)
Microlipophrys canevae		*Torquigener flavimaculosus*** (Bilecenoğlu, 2005)
Oedalechilus labeo		
*Oxyurichthys auchenolepis **		
Parablennius incognitus		
Parablennius rouxi		
Parablennius sanguinolentus		
*Parexocoetus brachypterus** (Oral, 2010) *(excluded)*		
*Parupeneus margaritatus**		
Pseudocaranx dentex		
Salaria pavo		
Sardina pilchardus		
Spicara maena		
Vanneaugobius pruvoti		
Xyrichtys novacula		

*Possible alien species

**Alien species recorded in the literature, some with established populations

The pairwise sample comparison for fish species detected by both primers is given in Figure 4.21. In samples S10 and S16, the species were detected by the 12S rRNA metabarcoding datasets only. Since S02 was excluded from the 12S rRNA dataset, the only species detected in that sample were from the COI metabarcoding dataset. In S20, none of the datasets comprised any fish species. In samples S04, S05, S06, S11, S12, and S14, species were detected by both primers, and their proportions were below 20 %.

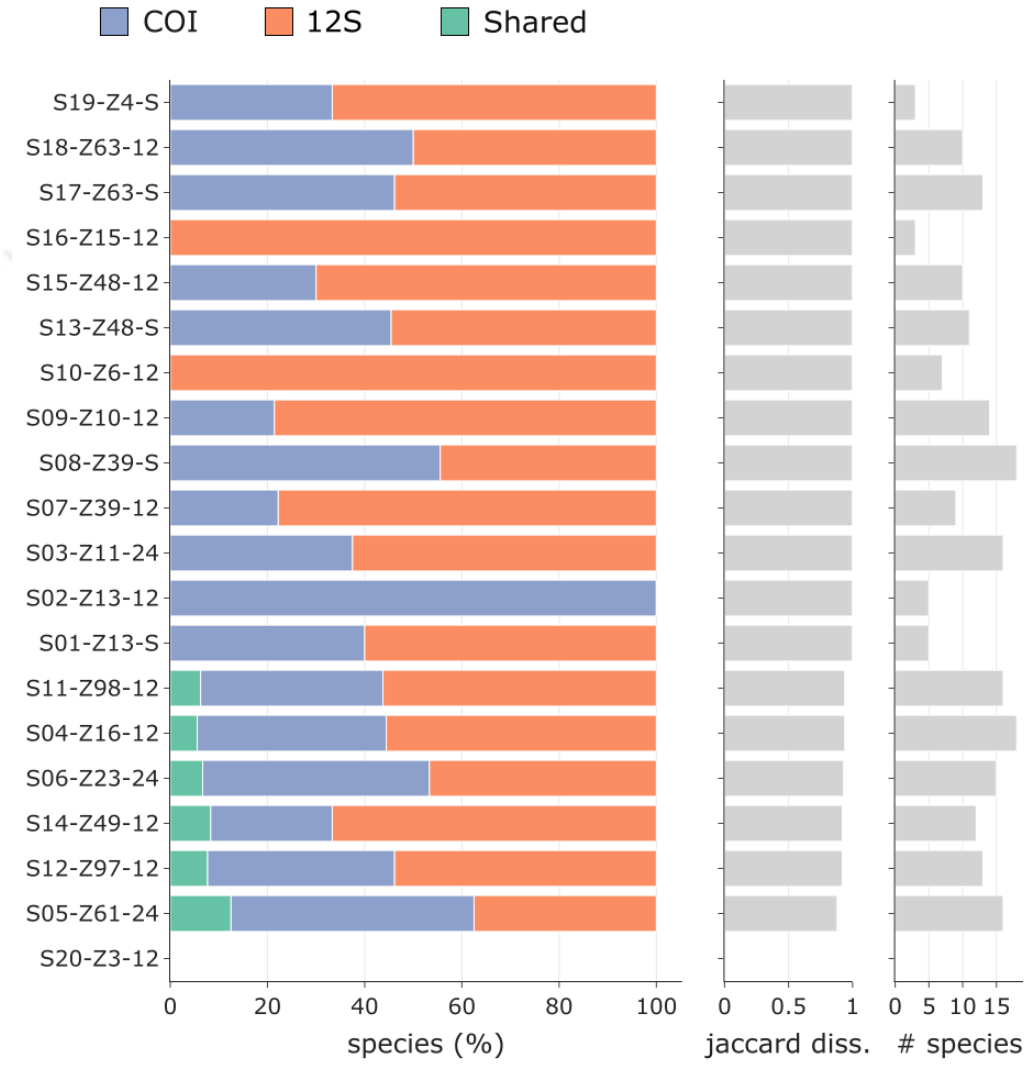


Figure 4.21. The pairwise sample comparison of the fish species that were identified by both primers. (The fish species identified by Red :12S rRNA primers, Blue: COI primers, Green: Both primers)

5. DISCUSSION

The main objective of this study was to compare the fish biodiversity assessed by eDNA metabarcoding and Underwater visual census (UVC) in the Kaş-Kekova Marine Protected Area in order to check that the two approaches cross-validate or complement each other. In addition, we aimed to evaluate whether eDNA metabarcoding could detect elusive/ cryptic or invasive species, where the visual surveys might fall short in their detection. We also used two different primer sets for the eDNA metabarcoding analysis; therefore, another objective was to compare whether the two primers would complement each other regarding species assignment.

In order to compare the two primer sets in terms of their power to assess the fish biodiversity, the first approach was to compare the number of reads and OTUs obtained by both primers. In terms of raw reads, the number of raw reads in the COI dataset was 38,772,663, whereas for 12S, this value was 32,364,379. When comparing the number of OTUs obtained before the normalization and data filtration (i.e. negative control subtraction, no match subtraction and 85% similarity cutoff), the number of OTUs for 12S and CO1 were 247 and 8359, respectively. For 12S out of 247 OTUs, 82 of them did not get a hit in the NCBI BLAST database, and overall, 25% of the raw reads were *Homo sapiens* hits. 4% of the reads could not be assigned to a taxonomy in the database; however, these reads comprised 82 OTUs. As for the COI dataset, since we used three different databases and combined them according to the OTUs that had the highest hits, there were no OTUs that did not have any matches in the databases. For the *Homo sapiens* hits, it consisted of 24 OTUs which was about 0.08 % of the total raw reads.

After the filtration and the normalization, the number of OTUs decreased to 108 and 2848, respectively. Since the normalization was performed according to the minimum number of reads, for 12S, the number of reads was normalized to 3588 reads and 52,678 for COI. Even though there was a disproportion in the number of reads left after the normalization and for the number of OTUs, the number of OTUs at the species level for Chordata did not change in 12S primers. Moreover, for the COI primers, the number of species at the OTUs level before and after normalization the number of OTUs at species level decreased from 362 to 290 and for the Chordata phylum, it decreased from 66 to 36. In general, the number of reads varied considerably between the samples in the 12S dataset, whereas in COI, the variance was not that drastic. This could be one of the drawbacks of the 12S primer set and one of the limitations in our sampling effort and sequencing depth.

The taxonomic resolution, however, shows that even though we had to eliminate one of the samples (S02) and the read numbers were normalized to 3,588 reads in the 12S dataset, it had a higher taxonomic resolution at the species level of 42%. Meanwhile, even though it was normalized to a higher number of reads (52,678) after the read normalization for the COI dataset, the taxonomic resolution could only reach 11%. It is important to note that decision on which primer set to be utilized is based on a trade-off between taxonomic coverage and taxonomic resolution (van der Loos & Nijland, 2021). Therefore, COI primers fit a general species characterization, whereas 12S primers are fit for fine-tuned characterizations in terms of our study design trying to achieve the balance between greater taxonomic coverage and better taxonomic resolution.

As another methodological point, when we consider cross-contamination, especially OTUs that are assigned to *Homo sapiens*, 12S had a higher percentage of reads than COI, but the number of OTUs was one and 24, respectively. This shows that 12S rRNA primers might be more prone to primer bias, considering the negative controls and blank samples did not have high reads that get a hit at human DNA. In the literature, human-specific blocking primers are usually used to overcome this and shown to reduce the read proportions of human DNA amplicon sequences (Zhang et al., 2020). However, they can also inhibit the other sequences during the amplification by PCR. In a review that was assessing the performance of 22 eDNA metabarcoding primer sets in the detection of fish diversity, blocking oligos decreased the number of fish taxa detected for two primers 12SV05 and Tele01 (Zhang et al., 2020). Therefore, increasing sequencing depth rather than addition of blocking oligos seemed a better approach then addition of oligos. Another approach we employed to increase the sequencing depth was to use the NovaSeq instrument to target 1,200,000 reads instead of the most commonly used MiSeq, in which the sequencing depth per sample in most of the studies varied but had a median of 60,000 reads per sample (Singer et al., 2019). In a study using both instruments, MiSeq with sequencing depth of 100,000 reads and NovaSeq with 7,000,000 reads, they found that NovaSeq could detect many more taxa. Even with the same sequencing depth in NovaSeq and MiSeq, they also found that NovaSeq still had higher detection of taxa (Singer et al., 2019). Besides the contamination by human DNA, 12S primers have another drawback: the lack of reference sequences in the databases (Miya, 2022). In the 12S dataset, 82 OTUs did not hit any sequences in the database, which might indicate the lack of sequences in the databases for 12S rRNA.

Regarding species diversity assigned to the Chordata phylum by both primers and at the fish species level, 11 and 10 species are shared by both primers, respectively. Since COI databases are diverse and have been used extensively in the DNA barcoding of the species (van der Loos & Nijland,

2021), incorporating more than one primer set could overcome the lack of reference sequences. As it could be pointed out, the number of species that were detected by both primers did not completely overlap. In total, the number of fish species detected by both primers was ten; however, when compared at the sample level, only six samples had the same species detected in the 12S and COI primer datasets. The species that were common in both primer sets for these six samples were *Sparus aurata*, *Sparisoma cretense*, and *Coris julis*. The remaining species that were shared by both primer sets were not at the sample level but overall, meaning that one of the species among the seven species was present in one sample in only one primer dataset and in another sample in the other primer dataset. If we have had PCR replicates for each primer set, we may not have seen this instance considering the sequences for these species were present in the samples, and were amplified in only one primer set, even though they were detectable and amplifiable by both primer sets and had reference sequences in the databases for both primer sets. However, in the literature, rather than PCR replicates, biological replicates were found to be more promising and more comprehensive in terms of species detection (Czachur et al., 2022) which was the approach we employed. This means that biological replicates varied substantially in the number of unique taxa detected by eDNA metabarcoding compared to their PCR replicate counterparts (Macher et al., 2021). In our study, even though only one biological replicate was taken from the same location and depth, the percentage of the shared OTUs was approximately 45%, indicating that more than half of the OTUs in both samples were unique (replicate stochasticity).

When compared in terms of sampling location and depth, the 12S and COI primers seemed to differ in terms of species biodiversity detected with respect to sampling depth. Both did not show dissimilarities in species diversity when the sampling location was taken as the factor. However, for the COI dataset, the biodiversity in each sampling depth i.e. surface, 12m, and 24m showed significant and moderate dissimilarities. Even though in another study conducted with 12S rRNA primers (MiFish), the eDNA metabarcoding was able to detect a dissimilarity in terms of marine vertebrate biodiversity (Andruszkiewicz et al., 2017) there was no significant dissimilarity in terms of depth in our 12S dataset. The reason might be due to 12S rRNA primers targeting *Chordata* phylum only, whereas COI primers target broader diversity within *Eukaryota*. The difference in the read numbers might also been another reason why there was no effect of sampling depth in terms of species diversity, with 12S having 3,588 reads whereas COI having 52,678 reads, the latter has almost 15 times more reads. It is also important to note that the number of OTUs that did not have matching sequences in the NCBI database in the 12S rRNA metabarcoding dataset, might be another cause for the lack of dissimilarity in the species diversity in terms of sampling depth, since it indicates the

possibility of data gap and lack of representative power of the reference sequences (Andruszkiewicz et al., 2017). Another reason might be *Homo sapiens* reads having dominated the PCR reactions, as mentioned above, indicating the possibility of primer bias (Eble et al., 2020; Ruppert et al., 2019).

As for the visual survey data, similar to COI, the sampling depth demonstrated a dissimilarity in the species diversity. In a study comparing eDNA and UVCs, in terms of ecological traits, UVCs were found to have a more focused phylogenetic signal in contrast to eDNA with a wider phylogenetic distribution (Polanco Fernández et al., 2020). Since there is a predetermined checklist in the underwater visual surveys, the reason for the dissimilarities might be the bias introduced by this approach or a phylogenetic signal with the taxonomic expertise, since the checklist was created with respect to prior visual surveys conducted in the area.

The consistency of the taxa detected by both eDNA metabarcoding and UVCs when the species detected by the eDNA metabarcoding were mapped to the UVC checklist of 66 taxa, indicates that eDNA metabarcoding failed to detect 23 species that UVCs were able to detect at all transects and 19 species at the 20 transects the eDNA samples were taken. However, among the 55 species that were detected by eDNA metabarcoding, only five species were not recorded in the Mediterranean before indicating that 91% of the species detected by eDNA metabarcoding exist in the area. Since the UVCs have been conducted in Kaş-Kekova MPA for 20 years, the checklist includes specific fish species only. In case of biomonitoring, the eDNA could not replace the UVCs since, it was unable to detect 23 species. However, this indicates that the two methods could be used to complement each other while reducing the labour and costs of biomonitoring by visual surveys.

The comparison of the results obtained from the eDNA metabarcoding and UVCs without mapping the taxa to the UVC checklist, indicates that the two methods differ significantly in terms of taxonomic detection at the genus and the species level. This indicates that both methods could be used to complement each other in case of biomonitoring of fish species in important areas, such as marine protected areas as in our study. In most of the studies, eDNA detected a higher number of species compared to the visual surveys (Lamy et al., 2021; Polanco Fernández et al., 2020; Thomsen et al., 2012; Valdivia-Carrillo et al., 2021; Yamamoto et al., 2017). This could be explained by the nature of the visual surveys where factors such as taxonomic expertise and the challenges introduced by the restricted vision underwater, as well as presence of elusive and or cryptic species (Brock, 1982), come into play. For instance, in the checklist of the taxa screened by the UVCs in this study, some taxa were counted at genus level such as *Chelon* sp. *Liza* sp. *Mugil* sp. in the family of

Mugilidae, *Atherina* sp. and *Atherinomorus* sp. in the family of *Atherinidae*, and *Lagocephalus sp*. In the eDNA dataset two *Atherina* species, two *Lagocephalus* species, and one *Chelon* species were detected at the species level demonstrating the complementarity of the two methods. Besides introducing higher resolution, eDNA was able to detect five species i.e. *Pseudocaranx dentex, Enchelycore anatina, Lagocephalus sceleratus, Lagocephalus suezensis, Muraena helena* that were in the UVC checklist but were not observed in the UVCs. Among these, *Enchelycore anatina* (FishBase) and *Muraena helena* (nocturnal) (Matić-Skoko et al., 2011) are both species that are found in rocky bottoms and known to conceal themselves or lurk in holes, indicating a possible reason why they might not have been observed by the UVCs. We cannot exclude these species as non-present, because their presence was known prior to the UVCs and were in the species checklist, but either at the time of surveys they were not observed due to their elusiveness or the other challenges of UVCs.

Regarding fish species detected by eDNA and whether they have been observed in the Mediterranean before and were previously reported as invasive species, eleven species in the eDNA dataset were observed and recorded in the Mediterranean and the Levant coasts as invasive species. Seven of these species were also detected in the visual surveys. Five species detected by eDNA metabarcoding were neither native nor recorded in the Mediterranean, with no evidence in the literature. However, some could be assigned to these species due to the lack of reference sequences in the NCBI database, especially for 12S rRNA reference sequences. One of these species was *Etrumeus micropus*. Its native distribution is in the Western Pacific (FishBase). However, in the checklist of invasive species in the Mediterranean, two other species in the same genus were present, i.e. *Etrumeus teres* (Oral, 2010) and *Etrumeus golanii* (Turan et al., 2018). *Etrumeus teres* has the reference of complete mitochondrion genome in the NCBI database; however, *Etrumeus golanii* does not have a 12S RNA record. Thus, *Etrumeus micropus* might have been misassigned instead of *Etrumeus golanii*.

Similarly, though detected in the 12S rRNA dataset, *Oxyurichthys auchenolepis* (*Oxyurichthys saru unaccepted* name) was not previously recorded in the Mediterranean or the Levantine Sea. It is distributed in the Central Pacific, however, in the same genus *Oxyurichthys petersii* (no 12S record) (Turan et al., 2018), *Oxyurichthys papuensis* (partial 12S) (Oral, 2010), *Oxyurichtys keiensis* (no 12S, COI only) (Özden et al., 2021) had been recorded as non-indigenous or alien species in the coasts of Turkey. In the NCBI dataset, the species was assigned as *Oxyurichthys saru,* but the name was not accepted; instead, the accepted name is *Oxyurichthys auchenolepis* in the literature (WoRMS). Therefore, *Oxyurichthys auchenolepis* taxonomic assignment in the database might have been

inaccurate because *Oxyurichthys petersii* and *Oxyurichtys keiensis* do not have 12S rRNA records in the database, and *Oxyurichthys papuensis* has only a partial 12S rRNA record.

Another similar species, *Parexocoetus brachypterus* has a complete mitochondrion genome in the database, and even though reported as an invasive species, was excluded from the CIESM Atlas due to lack of confirmation (Oral, 2010). *Parexocoetus mento* is recorded as an alien fish species in the Mediterranean and has partial COI and 12S rRNA records in the database, indicating either that *Parexocoetus brachypterus* might exist in the Mediterranean coasts or the taxonomic assignment was wrong due to a lack of reference sequences in the database.

Yet another case for tentative assignment could be *Parupeneus margaritatus,* again assigned from the 12S rRNA dataset. In this case, this assignment could be interpreted as either the sequence belonging to another invasive species, *Parupeneus forsskali,* which was detected in the COI dataset since *Parupeneus forsskali* does not have a 12S rRNA reference sequence in the NCBI dataset.

Finally, *Heteromycteris hartzfeldii* was the only species assigned in the COI dataset that was not recorded in the Mediterranean, and its native distribution is in the Western Pacific, in the family of *Soleidae*. At least 11 fish species in the Soleidae family are on the checklist of marine fishes in Turkey (Bilecenoğlu et al., 2014). Since for the COI datasets, we only used the R1 reads, which corresponds to 150 bp length of the original barcode (which is 313 bp long), even though we obtained a higher number of reads and OTUs per sample that reached a resolution at the species level, some of the species like *Hetereomycteris hartzfeldii* might have been assigned incorrectly.

6. CONCLUSIONS AND RECOMMENDATIONS

This study aimed to assess fish biodiversity in Kaş-Kekova MPA by using the eDNA metabarcoding approach while comparing the findings with the simultaneous underwater visual census. The eDNA metabarcoding was tested for its utilization as a novel approach to biomonitoring and biodiversity characterization of the fish species. The comparison of the two methods demonstrated the complementarity of each method such that 24 species were exclusively detected by UVCs whereas 42 by eDNA. Even though there were limitations to eDNA metabarcoding in terms of lack of reference sequences in 12S rRNA, as well as low resolution of the COI primers, five species were detected, which UVCs fell short in their detection of. Two of these species, *Enchelycore anatina*, known to hide among the rocks, and *Muraena helena*, a nocturnal species known to inhabit rocky bottoms and burrow in holes, were species that could be considered elusive, which could be challenging to detect during visual surveys due to their behaviour—demonstrating the possibility of incorporation of eDNA metabarcoding with other visual methods to better predict and assess the species presence in an area. Besides, by eDNA metabarcoding, five other species were detected with no data in the literature of their recordings on the Mediterranean coasts. However, due to the lack of reference sequences in the NCBI database, whether they were indeed present during the sampling cannot be concluded. However, one of them, *Parexocoetus brachypterus*, was reported to be found in the Mediterranean but, due to lack of information, specimen, or evidence, was excluded in the CIESM Atlas, indicating a higher possibility of its existence. However, without any visual confirmation, we cannot draw any conclusions. This study demonstrated the possible power of eDNA metabarcoding to aid and enhance species detection when paired with visual surveys, where the latter could be cost- and labour-intensive.

Besides comparing the two methodologies, in this study, we also compared the two primer sets and their resolution and complementarity to each other. Even though more species were detected by 12S rRNA primers, about 33 % of the OTUs did not have any corresponding reference sequences in the database. In contrast, when three databases were combined, all of the OTUs in the COI dataset hit a match in the reference libraries. However, the resolution at the species level was considerably lower. Therefore, using multiple gene markers can help overcome the disadvantages of eDNA metabarcoding. Although, we used two gene markers, some of the samples had a considerably lower number of reads than the others either due to sequencing errors or experimental errors during cleanup

or library pooling steps. This can be overcome by use of technical replicates, the lack of which points out to one of the limitations of this study.

In the literature, using multiple markers with three PCR replicates and two or more biological replicates showed to provide enough coverage to dismiss false positive taxonomic assignments and had higher representative power to detect rare taxa and eliminate primer bias. In this study, even though we used multiple gene markers, the lack of technical replicates probably resulted in uncertainties in the taxonomic assignment to the OTUs. Even though there was one biological replicate in this study, the number of OTUs shared was 45 %, pointing out the significance of replicate stochasticity. Therefore, for further research, incorporating more gene markers, with both technical and biological replicates, could result in better taxonomic assignment with higher representative power of the species composition in the study area.

REFERENCES

Ahyong, S., Boyko, C. B., Bailly, N., Bernot, J., Bieler, R., Brandão, S. N., Daly, M., De Grave, S., Gofas, S., Hernandez, F., Hughes, L., Neubauer, T. A., Paulay, G., Boydens, B., Decock, W., Dekeyzer, S., Vandepitte, L., Vanhoorne, B., Adlard, R., … Zullini, A. (2024). World Register of Marine Species (WoRMS). WoRMS Editorial Board. https://www.marinespecies.org

Akyol, O., Ünal, V., Ceyhan, T., & Bilecenoglu, M. (2005). First confirmed record of Lagocephalus sceleratus (Gmelin, 1789) in the Mediterranean Sea. *Journal of Fish Biology*, 66(4), 1183–1186. https://doi.org/10.1111/j.0022-1112.2005.00667.x

Andruszkiewicz, E. A., Starks, H. A., Chavez, F. P., Sassoubre, L. M., Block, B. A., & Boehm, A. B. (2017). Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLOS ONE*, 12(4), e0176343. https://doi.org/10.1371/journal.pone.0176343

Bariche, M., Bilecenoglu, M., & Azzurro, E. (2013). Confirmed presence of the Red Sea goatfish Parupeneus forsskali (Fourmanoir & Guézé, 1976) in the Mediterranean Sea. *BioInvasions Records*, 2(2), 173–175. https://doi.org/10.3391/bir.2013.2.2.15

Bilecenoğlu, M. (2005). Observations on the burrowing behaviour of the Dwarf Blaasop, Torquigener flavimaculosus (Osteichthyes: Tetraodontidae) along the coast of Fethiye, Turkey. *Zoology in the Middle East*, 35(1), 29–34. https://doi.org/10.1080/09397140.2005.10638100

Bilecenoglu, M. (2008). First record of Vanderhorstia mertensi Klausewitz, 1974 (Pisces, Gobiidae) in the Mediterranean Sea. *Aquatic Invasions*, 3(4), 487–490. https://doi.org/10.3391/ai.2008.3.4.22

Bilecenoglu, M. (2010). Alien marine fishes of Turkey – an updated review (pp. 189–217).

Bilecenoğlu, M., & Çınar, M. E. (2021). Alien Species Threat across Marine Protected Areas of Turkey—An Updated Inventory. *Journal of Marine Science and Engineering*, *9*(10), 1077. https://doi.org/10.3390/jmse9101077

Bilecenoğlu, M., Kaya, M., Cihangir, B., & Çiçek, E. (2014). An updated checklist of the marine fishes of Turkey. *Turkish Journal of Zoology*, *38*, 901–929. https://doi.org/10.3906/zoo-1405-60

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W., & de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, *29*(6), 358–367. https://doi.org/10.1016/j.tree.2014.04.003

Brock, R. E. (1982). A critique of the visual census method for assessing coral reef fish populations. *Bulletin of Marine Science*, 32(1), 269–276.

Brondízio, E. S., Settele, J., Díaz, S., & Ngo, H. T. (2019). *Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*.

Buchner, D., & Leese, F. (2020). BOLDigger – a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding and Metagenomics*, *4*. https://doi.org/10.3897/mbmg.4.53535

Buchner, D., Macher, T.-H., & Leese, F. (2022). APSCALE: advanced pipeline for simple yet comprehensive analyses of DNA metabarcoding data. *Bioinformatics*, *38*(20), 4817–4819. https://doi.org/10.1093/bioinformatics/btac588

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Çiftçi, N., & Ayas, D. (2022). First Record of Pempheris rhomboidea (Kossmann & Räuber, 1877) Harvested from Aquaculture Fish Cages. *Natural and Engineering Sciences*, *7*(2), 182–189. https://doi.org/10.28978/nesciences.1159280

Çinar, M. E., Bilecenoğlu, M., Yokeş, M. B., Öztürk, B., Taşkin, E., Bakir, K., Doğan, A., & Açik, Ş. (2021). Current status (as of end of 2020) of marine alien species in Turkey. *PLOS ONE*, *16*(5), e0251086. https://doi.org/10.1371/journal.pone.0251086

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, *18*(1), 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Costello, M. J., Basher, Z., McLeod, L., Asaad, I., Claus, S., Vandepitte, L., Yasuhara, M., Gislason, H., Edwards, M., Appeltans, W., Enevoldsen, H., Edgar, G. J., Miloslavich, P., De Monte, S., Pinto, I. S., Obura, D., & Bates, A. E. (2017). Methods for the Study of Marine Biodiversity. In *The GEO Handbook on Biodiversity Observation Networks* (pp. 129–163). Springer International Publishing. https://doi.org/10.1007/978-3-319-27288-7_6

Cruaud, P., Vigneron, A., Fradette, M.-S., Charette, S. J., Rodriguez, M. J., Dorea, C. C., & Culley, A. I. (2017). Open the SterivexTM casing: An easy and effective way to improve DNA extraction yields. *Limnology and Oceanography: Methods*, *15*(12), 1015–1020. https://doi.org/https://doi.org/10.1002/lom3.10221

Czachur, M. V., Seymour, M., Creer, S., & von der Heyden, S. (2022). Novel insights into marine fish biodiversity across a pronounced environmental gradient using replicated environmental DNA analyses. *Environmental DNA*, *4*(1), 181–190. https://doi.org/10.1002/edn3.238

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., & Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: The example of the American bullfrog Lithobates catesbeianus. *Journal of Applied Ecology*, *49*(4), 953–959. https://doi.org/10.1111/J.1365-2664.2012.02171.X

Demir, V. (2011). *Marine Biological Diversity Assessment for Marine Conservation Planning in Antalya-Kas Using Decision Support Systems* [Doctoral Dissertation]. Istanbul University.

Dudley, N. (2008). *Edited by Nigel Dudley Including IUCN WCPA Best Practice Guidance on Recognising Protected Areas and Assigning Management Categories and Governance Types Guidelines for Applying Protected Area Management Categories*. www.iucn.org/pa_guidelines

Eble, J. A., Daly-Engel, T. S., DiBattista, J. D., Koziol, A., & Gaither, M. R. (2020). Marine environmental DNA: Approaches, applications, and opportunities. *Advances in Marine Biology*, *86*(1), 141–169. https://doi.org/10.1016/BS.AMB.2020.01.001

Elbrecht, V., & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. *PLOS ONE*, *10*(7), e0130324. https://doi.org/10.1371/journal.pone.0130324

Ergüden, D., Bayhan, Y. K., Ergüden, S. A., & Altun, A. (2018). Occurrence of the Red Sea Goatfish Parupeneus forsskali (Fourmanoir & Guézé, 1976) from the Western Mediterranean Coast of Turkey. *Turkish Journal of Maritime and Marine Sciences*, *4*(1), 68–72. https://dergipark.org.tr/en/pub/trjmms/issue/40088/477215

Fediajevaite, J., Priestley, V., Arnold, R., & Savolainen, V. (2021). Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. *Ecology and Evolution*, *11*(9), 4803–4815. https://doi.org/10.1002/ece3.7382

FishBase. (2024). Oxyurichthys saru Tomiyama, 1936. In R. Froese & D. Pauly (Eds.), FishBase. Retrieved from World Register of Marine Species website: https://www.marinespecies.org/aphia.php?p=taxdetails&id=273464

Froese, R., & Pauly, D. (2023). Fishbase. World Wide Web electronic publication. Version (01/2023). https://www.fishbase.org.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, *8*(1), 1188. https://doi.org/10.1038/s41467-017-01312-x

Gaston, K. J., & Spicer, J. I. (2004). Human Impacts. In *Biodiversity: An Introduction* (2nd ed.). Blackwell Publishing.

Golani, D. (2000). First record of the bluespotted cornetfish from the Mediterranean Sea. *Journal of Fish Biology*, *56*(6), 1545–1547. https://doi.org/10.1111/j.1095-8649.2000.tb02163.x

Golani, D., & Sonin, O. (1992). New records of the Red Sea fishes,Pterois miles (Scorpaenidae) andPteragogus pelycus (Labridae) from the eastern Mediterranean Sea. *Japanese Journal of Ichthyology*, *39*(2), 167–169. https://doi.org/10.1007/BF02906001

Goulletquer, P., Gros, P., Boeuf, G., & Weber, J. (2014). The Importance of Marine Biodiversity. In *Biodiversity in the Marine Environment* (pp. 1–13). Springer Netherlands. https://doi.org/10.1007/978-94-017-8566-2_1

Gower, J. C. (2014). Principal Coordinates Analysis. In *Wiley StatsRef: Statistics Reference Online*. Wiley. https://doi.org/10.1002/9781118445112.stat05670

Hoagland, P., Sumaila, U. R., Farrow, S., & Bohorquez, J. J. (2019). Marine protected areas. In *Encyclopedia of Ocean Sciences* (pp. 456–552). Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.11390-9

Holmlund, C. M., & Hammer, M. (1999). Ecosystem services generated by fish populations. *Ecological Economics*, *29*(2), 253–268. https://doi.org/10.1016/S0921-8009(99)00015-4

Huseyinoglu, M. F., Demir, V., Arda, Y., Draman, M., & Yokes, M. B. (2021). Spatio-temporal distribution of lionfish, Pterois miles (Bennett, 1828) in Kas-Kekova Special Environmental Protected Area, Turkey. *Estuarine, Coastal and Shelf Science*, *254*, 107331. https://doi.org/10.1016/j.ecss.2021.107331

IUCN. (2022). The IUCN Red List of Threatened Species. Version 2023-1. https://www.iucnredlist.org. Accessed on [4 January 2022].

Jaccard, P. (1912). The Distribution Of The Flora In The Alpine Zone. *New Phytologist*, *11*(2). https://doi.org/10.1111/j.1469-8137.1912.tb05611.x

Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. *Scientific Reports*, *9*(1), 12133. https://doi.org/10.1038/s41598-019-48546-x

Laffoley, D., Baxter, J. M., Day, J. C., Wenzel, L., Bueno, P., & Zischka, K. (2018). Marine protected areas. In *World Seas: An Environmental Evaluation Volume III: Ecological Issues and Environmental Impacts* (pp. 549–569). Elsevier. https://doi.org/10.1016/B978-0-12-805052-1.00027-9

Lamy, T., Pitz, K. J., Chavez, F. P., Yorke, C. E., & Miller, R. J. (2021). Environmental DNA reveals the fine-grained and hierarchical spatial structure of kelp forest fish communities. *Scientific Reports*, *11*(1), 14439. https://doi.org/10.1038/s41598-021-93859-5

Lee, Y.-D., Lee, G.-M., Park, J.-Y., & Gwak, W.-S. (2022). Comparing Environmental DNA Metabarcoding and Underwater Visual Census to Monitor Korean Coastal Fish Community. *Ocean Science Journal*. https://doi.org/10.1007/s12601-022-00082-y

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 34. https://doi.org/10.1186/1742-9994-10-34

Macher, T., Beermann, A. J., & Leese, F. (2021). TaxonTableTools: A comprehensive, platform-independent graphical user interface software to explore and visualise DNA metabarcoding data. *Molecular Ecology Resources*, *21*(5), 1705–1714. https://doi.org/10.1111/1755-0998.13358

Macher, T.-H., Schütz, R., Arle, J., Beermann, A. J., Koschorreck, J., & Leese, F. (2021). Beyond fish eDNA metabarcoding: Field replicates disproportionately improve the detection of stream associated vertebrate species. *Metabarcoding and Metagenomics*, *5*. https://doi.org/10.3897/mbmg.5.66557

Mariani, G., Cheung, W. W. L., Lyet, A., Sala, E., Mayorga, J., Velez, L., Gaines, S. D., Dejean, T., Troussellier, M., & Mouillot, D. (2020). Let more big fish sink: Fisheries prevent blue carbon sequestration—half in unprofitable areas. *Science Advances*, *6*(44). https://doi.org/10.1126/sciadv.abb4848

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Matić-Skoko, S., Tutman, P., Petrić, M., Skaramuca, D., Đikić, D., Lisičić, D., & Skaramuca, B. (2011). Mediterranean moray eel Muraena helena (Pisces: Muraenidae): biological indices for life history. *Aquatic Biology*, *13*(3), 275–284. https://doi.org/10.3354/ab00373

Mcleod, E. (2013). Marine Protected Areas: Static Boundaries in a Changing World. In *Encyclopedia of Biodiversity: Second Edition* (pp. 94–104). Elsevier Inc. https://doi.org/10.1016/B978-0-12-384719-5.00347-6

Minamoto, T., Miya, M., Sado, T., Seino, S., Doi, H., Kondoh, M., Nakamura, K., Takahara, T., Yamamoto, S., Yamanaka, H., Araki, H., Iwasaki, W., Kasai, A., Masuda, R., & Uchii, K. (2021). An illustrated manual for environmental DNA research: Water sampling guidelines and experimental protocols. *Environmental DNA*, *3*(1), 8–13. https://doi.org/10.1002/edn3.121

Miya, M. (2022). Environmental DNA Metabarcoding: A Novel Method for Biodiversity Monitoring of Marine Fish Communities. *Annual Review of Marine Science*, *14*(1), 161–185. https://doi.org/10.1146/annurev-marine-041421-082251

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., Kondoh, M., & Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, *2*(7), 150088. https://doi.org/10.1098/rsos.150088

Norris, K., Terry, A., Hansford, J. P., & Turvey, S. T. (2020). Biodiversity Conservation and the Earth System: Mind the Gap. In *Trends in Ecology and Evolution* (Vol. 35, Issue 10, pp. 919–926). Elsevier Ltd. https://doi.org/10.1016/j.tree.2020.06.010

Oral, M. (2010). Alien fish species in the Mediterranean – Black Sea Basin. *Journal of Black Sea/Mediterranean Environment*, *16*(1), 87–132. https://dergipark.org.tr/en/pub/jbme/issue/9841/121857

Özden, U., Seyhan Öztürk, D., Irmak, E., Çağlar Oruç, A., & Engin, S. (2021). A new alien fish in the Mediterranean Sea; Oxyurichtys keiensis (Smith, 1938) (Gobiiformes: Gobiidae). *Mediterranean Marine Science*. https://doi.org/10.12681/mms.28081

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187). https://doi.org/10.1126/science.1246752

Polanco Fernández, A., Marques, V., Fopp, F., Juhel, J., Borrero-Pérez, G. H., Cheutin, M., Dejean, T., González Corredor, J. D., Acosta-Chaparro, A., Hocdé, R., Eme, D., Maire, E., Spescha, M., Valentini, A., Manel, S., Mouillot, D., Albouy, C., & Pellissier, L. (2021). Comparing environmental DNA metabarcoding and underwater visual census to monitor tropical reef fishes. *Environmental DNA*, *3*(1), 142–156. https://doi.org/10.1002/edn3.140

Pörtner, H.-O., Scholes, R. J., Agard, J., Archer, E., Arneth, A., Bai, X., Barnes, D., Burrows, M., Chan, L., Cheung, W. L. (William), Diamond, S., Donatti, C., Duarte, C., Eisenhauer, N., Foden, W., Gasalla, M. A., Handa, C., Hickler, T., Hoegh-Guldberg, O., … Ngo, H. (2021). *Scientific outcome of the IPBES-IPCC co-sponsored workshop on biodiversity and climate change*. Zenodo. https://doi.org/10.5281/zenodo.5101125

Primack, R. B., & Morrison, R. A. (2013). Extinction, Causes of. *Encyclopedia of Biodiversity: Second Edition*, 401–412. https://doi.org/10.1016/B978-0-12-384719-5.00050-2

Ratnasingham, S., & Hebert, P. D. N. (2007). bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*(3), 355–364. https://doi.org/https://doi.org/10.1111/j.1471-8286.2007.01678.x

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*(21), e145–e145. https://doi.org/10.1093/nar/gkr732

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, *17*, e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Sala, E., Kizilkaya, Z., Yildirim, D., & Ballesteros, E. (2011). Alien Marine Fishes Deplete Algal Biomass in the Eastern Mediterranean. *PLoS ONE*, *6*(2), e17356. https://doi.org/10.1371/journal.pone.0017356

Sanchez, L., Boulanger, E., Arnal, V., Boissery, P., Dalongeville, A., Dejean, T., Deter, J., Guellati, N., Holon, F., Juhel, J.-B., Lenfant, P., Leprieur, F., Valentini, A., Manel, S., & Mouillot, D. (2022). Ecological indicators based on quantitative eDNA metabarcoding: the case of marine reserves. *Ecological Indicators*, *140*, 108966. https://doi.org/10.1016/j.ecolind.2022.108966

Singer, G. A. C., Fahner, N. A., Barnes, J. G., McCarthy, A., & Hajibabaei, M. (2019). Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific Reports*, *9*(1), 5991. https://doi.org/10.1038/s41598-019-42455-9

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., Sigsgaard, E. E., & Hellström, M. (2017). Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, *8*(5), 635–645. https://doi.org/10.1111/2041-210X.12683

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S., & Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, *7*(1), 12240. https://doi.org/10.1038/s41598-017-12501-5

Taberlet, P., Coissac, E., Hajibabaeiaji, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Taskavak, E., Bilecenoglu, M., Başusta, N., & Mater, S. (2000). Occurrence of Pteragogus pelycus Randall, 1981 (Teleostei: Labridae) and Petroscirtes ancylodon Rüppell, 1838 (Teleostei: Blennidae) in *Turkish Mediterranean waters*. *Acta Adriatica*, *41*, 53–58.

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

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. *PLoS ONE*, *7*(8), e41732. https://doi.org/10.1371/journal.pone.0041732

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. https://doi.org/10.1371/journal.pone.0165252

Turan, C., Erguden, D., Gürlek, M., Yaglioglu, D., & Uygur, N. (2014). First record of the blue-barred parrotfish, Scarus ghobban Forsskål, 1775 from Turkish coastal waters. *Journal of Applied Ichthyology*, *30*(2), 424–425. https://doi.org/10.1111/jai.12402

Turan, C., Erguden, D., Uygur, N., Gurlek, M., Erdogan, Z. A., Sonmez, B., Uyan, A., Karan, S., & Dogdu, S. A. (2015). First record of the Indian Ocean twospot cardinalfish, Cheilodipterus novemstriatus (Actinopterygii: Perciformes: Apogonidae), from Turkish marine waters. *Acta Ichthyologica et Piscatoria 45(3): 319-322*, *45*(3), 319–322. https://doi.org/10.3750/AIP2015.45.3.12

Turan, C., Gürlek, M., Başusta, N., Uyan, A., Doğdu, S. A., & Karan, S. (2018). A Checklist of the Non-indigenous Fishes in Turkish Marine Waters. *Natural and Engineering Sciences*, *3*(3), 333–358. https://doi.org/10.28978/nesciences.468995

Turner, C. R., Barnes, M. A., Xu, C. C. Y., Jones, S. E., Jerde, C. L., & Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution*, *5*(7), 676–684. https://doi.org/10.1111/2041-210X.12206

United Nations. (2022). The Sustainable Development Goals Report.

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., Yamanaka, H., Minamoto, T., & Kondoh, M. (2018). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *Metabarcoding and Metagenomics*, *2*. https://doi.org/10.3897/mbmg.2.23297

Valdivia-Carrillo, T., Rocha-Olivares, A., Reyes-Bonilla, H., Domínguez-Contreras, J. F., & Munguia-Vega, A. (2021). Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. *Molecular Ecology Resources*, *21*(5), 1558–1574. https://doi.org/10.1111/1755-0998.13375

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., … Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, *25*(4), 929–942. https://doi.org/10.1111/mec.13428

van der Loos, L. M., & Nijland, R. (2021). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology*, *30*(13), 3270–3288. https://doi.org/10.1111/mec.15592

Whitmarsh, S. K., Fairweather, P. G., & Huveneers, C. (2017). What is Big BRUVver up to? Methods and uses of baited underwater video. *Reviews in Fish Biology and Fisheries*, *27*(1), 53–73. https://doi.org/10.1007/s11160-016-9450-1

WWF-Turkey. (2012). Kaş-Kekova Marine Protected Area Project Report. www.wwf.org.tr

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., Minamoto, T., & Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, *7*(1), 40368. https://doi.org/10.1038/srep40368

Zhang, S., Zhao, J., & Yao, M. (2020). A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. *Methods in Ecology and Evolution*, *11*(12), 1609–1625. https://doi.org/10.1111/2041-210X.13485

APPENDIX A: DNA EXTRACTION PROTOCOLS

The pieces of equipment that are used to open up the Sterivex filter are cleaned with ethanol. The Sterivex filter is opened up aseptically with a pipe-cutter in the hood. The filter unit inside the cartridge is taken out carefully with a college plier without touching any of the outsides of the surface of the cartridge. The solution inside the filters is poured into a sterile petri dish. The filter paper is cut into two pieces with a sterile scalpel blade first, and then each equally sized piece is cut into many small pieces, and with sterile tweezers, the pieces are distributed into 2 ml tubes. The solution in the petri dish is equally distributed into two tubes of 500 µl each, and the remaining solution, 500 µl, is put into another 2 ml tube, assuming that the solution inside is 1.5 ml. Two of the tubes, the one with the filter (backup duplicate) and the one with the remaining solution, are kept at -20^{o}C. For DNA extraction, DNeasy Blood & Tissue Kit (Qiagen, USA) is used. Into the sample tube, 360 µl of ATL buffer, 40 µl of proteinase K, and 600 µl of pure distilled water are added and then incubated at 56^{o}C and 300rpm for 24 hours. After the incubation, the samples in the tubes are distributed in equal volumes into three 2 ml tubes. With the same volumes as the samples in the tube, cold ethanol (99%) and AL buffer are added into the tubes. 700 µl of the sample volumes are added into DNeasy spin columns. The columns are centrifuged at 8000rpm for one minute. After the centrifugation, the eluted liquid is discarded, and this step is repeated until all the samples in the tube have been used up. The DNeasy spin columns are then transferred into new 2 ml collection tubes. 500 µl of AW1 buffer is added, and the columns are centrifuged at 8000 rpm for one minute. The collection tubes are discarded, the spin columns are transferred into new 2 ml collection tubes, and 500 µl AW2 buffer is added. The tubes are centrifuged at 14000 rpm for 3 minutes. The collection tubes are discarded, and the spin columns are transferred into new 2 ml collection tubes and centrifuged at 13000 rpm for one minute. Spin columns are transferred into new collection tubes, and 50 µl of AE buffer is added. The columns are incubated at room temperature for 10 minutes. The columns are centrifuged at 8000 rpm for a minute. The spin columns are discarded, and the eluted DNA is transferred into new 2 ml tubes. The DNA concentrations are measured by QuBit dsDNA HS Assay (Invitrogen, CA, USA), and the extracted DNA is stored at -20°C until further processing.

APPENDIX B: THE LIST OF FISH SPECIES DETECTED BY BOTH METHODS

Table B.1: The fish genera detected by UVC, eDNA, and both.

Genus UVC Only	Genus Shared	Genus eDNA Only
Atherinomorus	*Atherina*	*Aidablennius*
Caranx	*Balistes*	*Aphia*
Dasyatis	*Boops*	*Auxis*
Lithognathus	*Cheilodipterus*	*Callionymus*
Liza	*Chelon*	*Chromis*
Mugil	*Coris*	*Coryphoblennius*
Mycteroperca	*Diplodus*	*Cyclothone*
Pagrus	*Epinephelus*	*Cynoglossus*
Pempheris	*Fistularia*	*Dicentrarchus*
Pteragogus	*Mullus*	*Enchelycore*
Pterois	*Oblada*	*Etrumeus*
Sarpa	*Parupeneus*	*Gobio*
Scarus	*Pseudocaranx*	*Gobius*
Seriola	*Sargocentron*	*Heteromycteris*
Serranus	*Siganus*	*Hygophum*
Symphodus	*Sparisoma*	*Lagocephalus*
Thalassoma	*Sparus*	*Lipophrys*
Upeneus	*Spicara*	*Microlipophrys*
Vanderhorstia	*Torquigener*	*Millerigobius*
	Xyrichtys	*Muraena*
		Oedalechilus
		Oxyurichthys
		Parablennius
		Parexocoetus
		Poecilia
		Salaria
		Sardina
		Vanneaugobius

Table B.2: The fish species detected by UVC, eDNA, and both. (Orange : The alien species in the list and their records in the literature)

Species UVC Only	Species Shared	Species eDNA Only
Caranx crysos	*Balistes capriscus*	*Aidablennius sphynx*
Dasyatis pastinaca	*Boops boops*	*Aphia minuta*
Diplodus vulgaris	*Cheilodipterus novemstriatus* (Çinar et al., 2021; Turan et al., 2015)	*Atherina boyeri*
Diplodus puntazzo	*Coris julis*	*Atherina hepsetus*
Epinephelus marginatus	*Diplodus sargus*	*Auxis rochei*
Lithognathus mormyrus	*Epinephelus costae*	*Callionymus filamentosus* (Çinar et al., 2021)(Oral, 2010)
Mullus surmuletus	*Fistularia commersonii* (Golani, 2000)	*Chelon labrosus*
Mycteroperca rubra	*Oblada melanura*	*Chelon saliens*
Pagrus pagrus	*Parupeneus forsskali* (Bariche et al., 2013; Ergüden et al., 2018)	*Chromis chromis*
Pempheris rhomboidea (Çiftçi & Ayas, 2022; Çinar et al., 2021)	*Pseudocaranx dentex*	*Coryphoblennius galerita*
Pteragogus trispilus (Çinar et al., 2021)	*Sargocentron rubrum* (Çinar et al., 2021)	*Cyclothone braueri*
Pterois miles (Golani & Sonin, 1992)	*Siganus luridus* (Çinar et al., 2021; Sala et al., 2011)	*Cynoglossus sinusarabici* (Çinar et al., 2021; Oral, 2010)
Sarpa salpa	*Siganus rivulatus* (Çinar et al., 2021; Sala et al., 2011)	*Dicentrarchus labrax*
Scarus ghobban	*Sparisoma cretense*	*Enchelycore anatina*
Seriola dumerili	*Sparus aurata*	*Epinephelus costae*
Serranus cabrilla	*Torquigener flavimaculosus*** (Bilecenoğlu, 2005)	*Etrumeus micropus*
Serranus scriba	*Xyrichtys novacula*	*Gobio gobio*
Spicara flexuosa		*Gobius niger*
Spicara smaris		*Heteromycteris hartzfeldii*
Symphodus tinca		*Hygophum benoiti*
Thalassoma pavo		*Lagocephalus sceleratus* (Akyol et al., 2005)
Upeneus pori(Bilecenoğlu & Çınar, 2021; Oral, 2010)		*Lagocephalus suezensis* (Bilecenoglu, 2010; Oral, 2010)
Vanderhorstia mertensi (Bilecenoglu, 2008)		*Lipophrys trigloides*
		Microlipophrys canevae
		Millerigobius macrocephalus
		Mullus barbatus
		Muraena helena
		Oedalechilus labeo
		Oxyurichthys saru
		Parablennius incognitus
		Parablennius rouxi
		Parablennius sanguinolentus
		Parexocoetus brachypterus
		Parupeneus margaritatus
		Poecilia wingei
		Salaria pavo
		Sardina pilchardus
		Spicara maena
		Vanneaugobius pruvoti

APPENDIX C: REPLICATE ANALYSIS

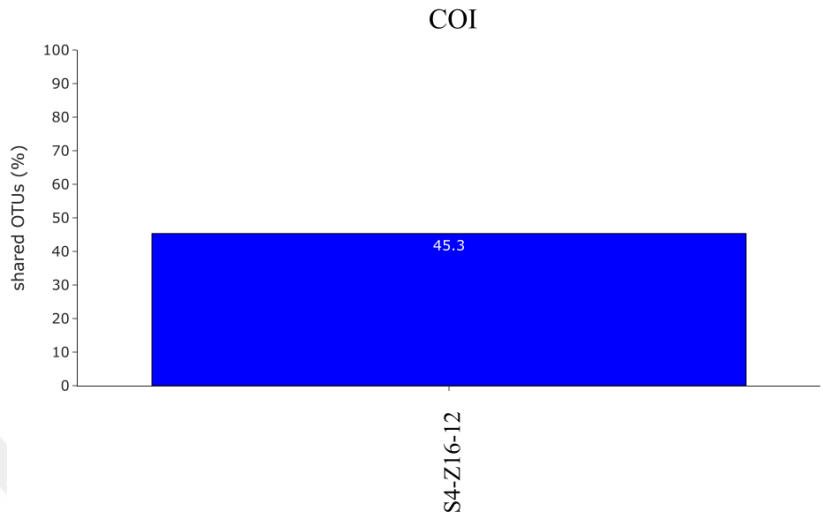


Figure C.1. The percentage of shared OTUs between biological replicates in COI dataset.

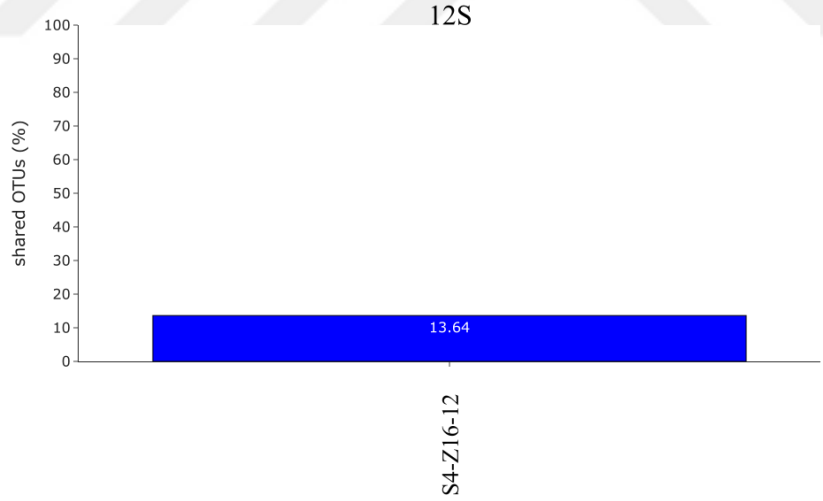


Figure C.2. The percentage of shared OTUs between biological replicates in 12S dataset.