

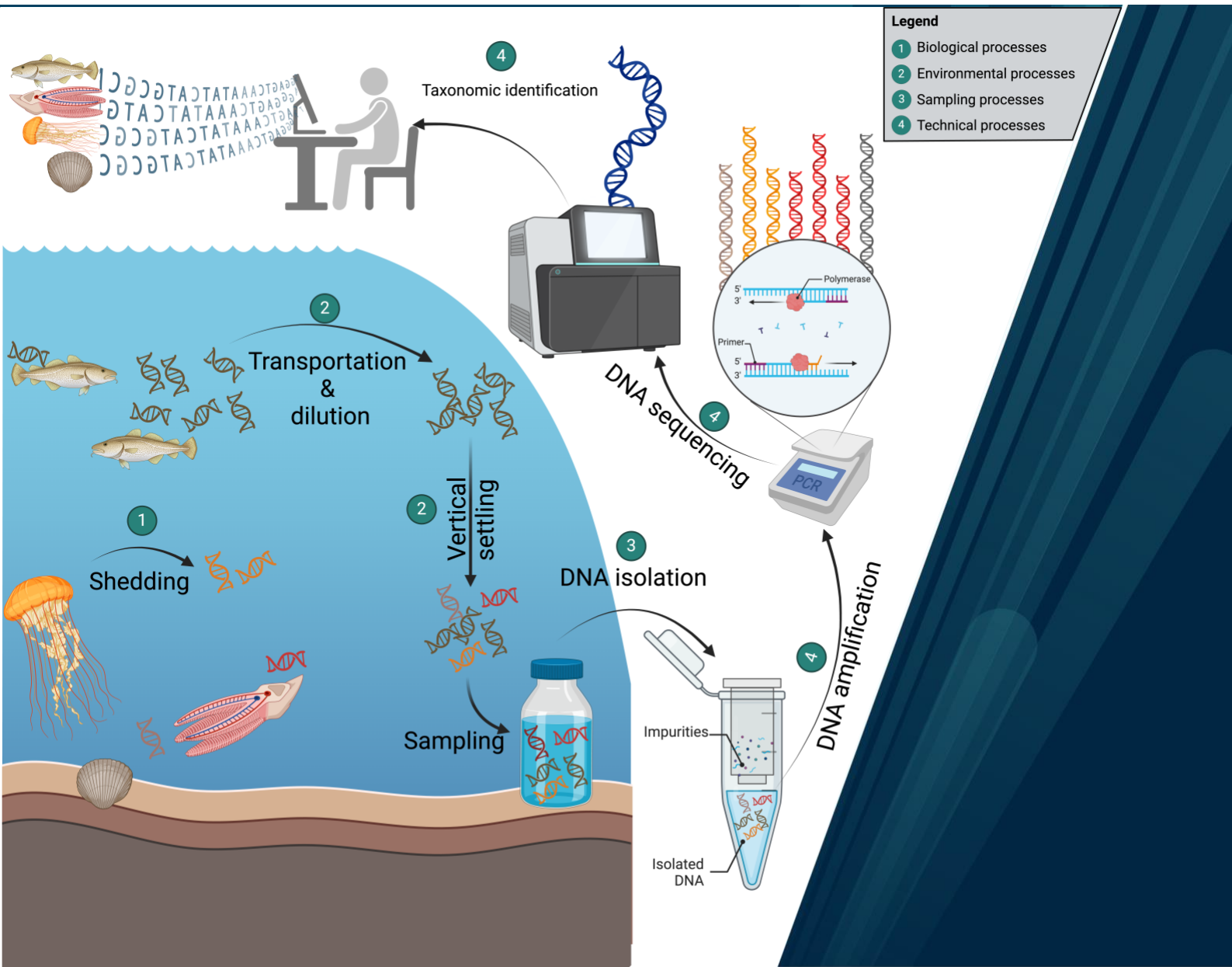
Faculty of Bioscience, Fisheries and Economics  
Department of Norwegian College of Fishery Science

## Integrating environmental DNA into marine management

Non-invasive surveillance tools for biomonitoring

Gledis Guri

A dissertation for the degree of Philosophiae Doctor January 2024





*“On the dark great sea, in the midst of javelins and arrows, in sleep,  
in confusion, in the depths of shame, **the good deeds a man has done  
before defend him.**”*

~ The Mahabharata

# Table of Contents

1	Introduction .....	3
1.1	Marine conservation .....	4
1.2	eDNA marine monitoring approaches .....	7
1.3	The need for this study .....	11
1.4	Aims and objectives.....	12
2	Materials and methods.....	14
2.1	Environmental DNA collection and isolation.....	14
2.1.1	Metabarcoding.....	15
2.1.2	qPCR .....	15
2.1.3	ddPCR .....	15
2.2	Trawling.....	17
2.3	Short description of analysis.....	17
3	Main findings .....	19
4	Discussion .....	24
4.1	Scale representation of eDNA .....	24
4.2	How to treat eDNA data for efficient biomonitoring .....	26
4.3	Single-species quantitative biomonitoring .....	28
4.4	Multi-species quantitative eDNA biomonitoring .....	29
4.5	Additional features of eDNA adaptable to marine management.....	31
5	Future consideration of eDNA for marine monitoring.....	34
6	Conclusion.....	36
7	References .....	38

## List of Tables

Table 1. Sequences for qPCR and ddPCR assays targeting 103-bp region of the ATPase gene of Atlantic cod ( <i>Gadus morhua</i> ) and saithe ( <i>Pollachius virens</i> ), and cytochrome b sequence of Atlantic herring ( <i>Clupea harengus</i> ). All gene regions belong to the mitochondrial DNA. Taken from <b>Paper II</b> . .....	16
---	----

## List of Figures

Figure 1. Cumulative number of efforts attributed to biological conservation programs since 1900 from Miloslavich et al. (2018) .....	5
Figure 2. Biodiversity index trends and Global policy commitments affecting the trends shown with dashed lines. Modified from Mace et al. (2020) .....	6

Figure 3. Schematic workflow of environmental DNA metabarcoding. Created using BioRender.....	9
Figure 4. Schematic workflow of quantitative PCR. Created using BioRender.....	10
Figure 5. Schematic workflow of Droplet Digital PCR. Created using BioRender.....	11
Figure 6. Latent processes (biological, environmental, sampling, technical and quantitation processes) of eDNA metabarcoding, qPCR and ddPCR investigated in this study (together with their workflow). This workflow is modified from Gold et al. (2023) and Shelton et al. (2016). Created using BioRender.....	12
Figure 7. Sample utilization among all the papers and objectives.....	14
Figure 8. Map over the study area and sampling stations' distribution.....	17
Figure 9. Simplified schematic overview used in <b>Paper III</b> of the joint Bayesian model workflow including all the inferences and processes. Note that the model specifies $\theta$ in the natural-log scale indicating the magnitude of conversion between fish densities and eDNA concentration. Extracted from <b>Paper III</b> .....	18
Figure 10. Dissimilarity of community composition between ecosystems measured as the PERMANOVA pseudo-F ratio using qualitative (blue; Jaccard dissimilarity matrix based on the presence/absence of species) and semiquantitative (red; Bray-Curtis dissimilarity matrix based on the inverse Wisconsin double-standardization of amplicon reads) approaches. The analyses were performed on progressively removed sampling stations, with each dot representing the mean pseudo-F of the sample combination (iterated 999 times) for the given number of remaining sampling stations. Regression lines for each approach are shown with their colours respectively. The horizontal black dashed line, representing an approximation of pseudo-F value ( $f = 2.01$ ) equal to $p = 0.05$ , indicated that minimum three and thirteen sampling stations are required for detecting significant differences between the three ecosystems respectively for semiquantitative and qualitative approach. Extracted from <b>Paper I</b> .....	19
Figure 11. Correlation plot of samples' pairwise dissimilarity (a) and the ratio of pairwise dissimilarity over the average dissimilarity per each depth (b) along samples' pairwise horizontal distance. The dissimilarity measured using Bary-Curtis index based on inverse Wisconsin double-standardization of amplicon reads indicated higher differences on surface communities (a) and average dissimilarity was achieved after 16 km of horizontal distance between samples (b). The standard errors are presented for both plots. Extracted from <b>Paper I</b> .....	20
Figure 12. Sensitivity of (a) ddPCR and qPCR for three assays (cod = red, herring = blue, and saithe = orange) shown as modelled detection probability as a function of nominal DNA concentration. Difference of detection probability (b) between ddPCR and qPCR is also shown over the eDNA concentration where positive values indicate higher ddPCR detection probability and zero indicates similar detection probabilities between the two employed methods. Extracted from <b>Paper II</b> .....	21
Figure 13. Comparison of modelled quantities between qPCR and ddPCR and their credible interval (grey bars) for three assays (cod = red, herring = blue, and saithe = orange; a). Concentration below 10 <sup>-3</sup> for both methods are considered non-detect. The difference between ddPCR and qPCR credible intervals is also shown (b) and grey shade for its variance. Extracted from <b>Paper II</b> .....	21
Figure 14. The joint Bayesian model workflow translating metabarcoding reads (a) and ddPCR droplets (b) into fish density (expressed as DNA concentration * and conversion factor; c) and their correlation with the fish density estimated from trawl observations in the x-axis. C = DNA concentration and $\theta$ = conversion factor between fish density to DNA concentrations. Metabarcoding and ddPCR models indicated here are model compartments within the joint Bayesian model. The plot (c) indicates model fit and reliable parameter estimation for linking trawl and eDNA observations. Extracted from <b>Paper III</b> .....	22



# Supervisors

**Dr. Kim Præbel**

Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø,  
Norway

**Dr. Jon-Ivar Westgaard**

Norwegian Institute of Marine Research, Framsenteret, Tromsø, Norway

**Dr. Torild Johansen**

Norwegian Institute of Marine Research, Framsenteret, Tromsø, Norway

# Acknowledgment

This research is dedicated to **my parents**, my ever-inspiring heroes, who have always encouraged me to push the limits within me, and to **my nephews**, to whom I hold a fervent hope that this research will contribute to shaping a better world for their generation to inherit.

I would like to express my gratitude to all of my direct and indirect supervisors that supported me in every step of the process and for being my friends beside supervision.

I would like to thank everyone who helped in field work, laboratory work and has given me any emotional support.



## List of papers

**Paper I:** Guri G., Westgaard. J-I., Yoccoz N., Wangensteen O. S., Præbel K., Ray J. L., Kelly R. P., Shelton A. O., Hanebrette T., Johansen T., 2023. Maximizing sampling efficiency to detect differences in fish community composition using eDNA metabarcoding in subarctic fjords. *Environmental DNA*. <https://doi.org/10.1002/edn3.409>

**Paper II:** Guri G., Ray J. L., Shelton A. O., Kelly R. P., Præbel K., Allan A., Yoccoz N., Johansen T., Wangensteen O. S., Hanebrette T., Westgaard J-I. (in revision). Quantifying the variance and precision between qPCR and ddPCR mechanisms for eDNA samples. *Methods in Ecology and Evolution*.

**Paper III:** Guri G., Shelton A. O., Kelly R. P., Yoccoz N., Johansen T., Præbel K., Hanebrette T., Ray J. L., Fall J., Westgaard J-I. eDNA quantification of fish abundance: Moving forward by coupling ddPCR, metabarcoding and trawl surveys.

### Author Contributions

Task	Paper I	Paper II	Paper III
Study concept	GG, TJ, NY, JIW, OW, RK, JR, AS, TH, and KP	GG, TJ, NY, JIW, OW, RK, JR, AS, TH, EA, and KP	GG, TJ, NY, JIW, RK, JR, AS, TH, JF, and KP
Study design	GG	GG	GG
Fieldwork	GG, TH, and JIW	GG, TH, and JIW	GG, TH, and JIW
Laboratory work	GG and TH	GG and OW	GG and TH
Statistical analysis	GG	GG, and AS	GG, RK, and AS
Ecological interpretations	GG, TJ, RK, OW, AS, JIW, JR, NY, and KP	GG, JR, RK, EA, TJ, JIW, NY, AS, and KP	GG, AS, TJ, JIW, NY, JR, RK, JF, and KP
Manuscript writing	GG	GG	GG

GG = Gledis Guri; TJ = Torild Johansen; JIW = Jon-Ivar Westgaard; NY = Nigel Yoccoz; OSW = Owen S. Wangensteen; KP = Kim Præbel; RK = Ryan Kelly; JLR = Jessica L. Ray; AOS = Andrew O. Shelton; and TH = Tanja Hanebrette; JF = Johanna Fall

## Summary

Human-induced factors such as pollution, overfishing, habitat destruction, and climate change have negatively impacted marine ecosystems, causing declines in biodiversity, and threatening the resilience of marine life. Biodiversity is vital for ecological sustainability, providing ecosystem services and supporting human economies. To address the biodiversity crisis, international frameworks like the CBD, UNCLOS, SDG 14, and GOA have set goals for marine conservation and management, including protected areas and sustainable management practices. With countries, including Norway, shifting from single-species management to Ecosystem Based Management to maintain healthy ecosystems, the need for biotic data has increased. To manage ecosystem health and prevent the deterioration of biological resources, there is a growing dependence on effective data collection.

Traditional and commercial trawling, mark-recapture, telemetry, hydro-acoustics, and electrofishing are tried and tested techniques for collecting ecological data with respect to fish surveys, but their invasive nature, high costs, and labor intensity often limit their application, and potentially harming marine ecosystems. Environmental DNA (eDNA), on the other hand, has emerged as a non-invasive, cost-efficient, and highly accurate alternative, revolutionizing marine biodiversity assessments, species detection and distribution, and community composition studies. However, eDNA faces its own challenges, including species-specific DNA shedding rates, Polymerase Chain Reaction (PCR) biases, and environmental conditions that affect DNA persistence and quantitative interpretation. While advanced PCR techniques like quantitative PCR (qPCR) and droplet digital PCR (ddPCR) can help quantify DNA in samples and improve accuracy, biological and environmental variables still pose significant obstacles. Enhancing and refining eDNA methods is crucial for gathering more precise and detailed information about marine ecosystems. This will lead to better-informed management decisions and assist in meeting international biodiversity conservation objectives.

In this thesis I utilized eDNA metabarcoding samples partially coupled with bottom trawl surveys, qPCR and ddPCR analysis to understand the eDNA dynamics by estimating the processes affecting the metabarcoding data output which will help resource managers designing better action plans for marine management.

In this thesis (composed of three research papers) I thoroughly explored the complex dynamics of eDNA metabarcoding (multi-species) and how various factors shape its persistence and

distribution with regards to marine observation and ecosystem management. These factors include DNA shedding rates, transportation and sedimentation rate, degradation rate, eDNA sampling and isolation efficiency, as well as technical methods used for detection. Given the influence of these factors, the eDNA persistence in the marine water is very short lived (days to weeks) and its spread is set to several kilometers range (depending on the water conditions). Subsequently, eDNA metabarcoding data (multi-species observation) was shown to hold some approximate quantities (semi-quantitative) regarding species abundances, thus serving as a valuable tool for rapid monitoring of marine community structures. Subsequently, I developed a novel approach for empirically assessing the factors affecting eDNA dynamics which can thereafter enable eDNA observations to be translated into fish abundances and densities, similar to those derived from trawl catch observations. By quantifying these dynamics, managers and policymakers can deepen their understanding of eDNA surveys thus enabling them to draw meaningful monitoring plans regarding anthropogenic impacts and climate change. With regards to early detection of invasive species or quantification of commercially important species, ddPCR was shown to have higher sensitivity and precision compared to the alternative single-species quantitation method (qPCR). This finding becomes particularly beneficial for resource managers who heavily rely on high sensitivity and precision for making informed decisions regarding the early detection of pathogens, introduced species, and quantitative assessment of marine biota.

eDNA also opens new horizons for comprehensive community and health assessments, facilitating the identification of crucial habitats, like spawning grounds, with minimal ecological disruption. However, despite its advantages, current eDNA approaches fall short in determining individual traits such as size or age. Overcoming this challenge lies in combining eDNA with supplementary genetic markers, forging a path toward more sophisticated fisheries and conservation tactics.

# 1 Introduction

The world's oceans encompass over 70% of our planet's surface and host an astounding array of life. However, in recent decades, human activities have significantly influenced marine ecology and biodiversity (O'Hara et al., 2021; Pan, 2023). Pollution, overfishing, habitat destruction, and climate change are among the key anthropogenic factors contributing to these impacts (O'Hara et al., 2021). Industrial and agricultural pollution introduce harmful substances into marine environments, causing various physiological and reproductive disorders in organisms. Overfishing has disrupted marine food webs and depleted fish stocks, leading to imbalances within ecosystems (Madin et al., 2016). Habitat destruction, whether from bottom trawling, coastal development, or coral reef degradation, poses a significant threat to marine biodiversity by eliminating crucial breeding and shelter sites (Madin et al., 2016). Furthermore, climate change, driven by human-induced greenhouse gas emissions, is causing rising sea temperatures, sea level rise, and ocean acidification, profoundly affecting marine life (Mieszkowska et al., 2014). These cumulative anthropogenic impacts have resulted in a decline in overall biodiversity, impairing the resilience and recovery abilities of marine ecosystems (Mace et al., 2020). Understanding the consequences of these impacts is paramount in developing effective conservation strategies for the protection of biodiversity.

The biologic diversity (biodiversity) is the variety of life and it is a well-established barometer of environmental sustainability jointly with the wellbeing of our biological resources (Vačkář et al., 2012). The biodiversity spans from microscopic life forms to whales and includes genetic, population, species, and ecosystem diversity. The importance of this diversity goes beyond just the number of species, as each level of diversity contributes to the overall functionality and resilience of the ecosystems and our oceans (Zou et al., 2020). It plays a significant role in maintaining the vast majority of our biological resources (McCann, 2000), providing essential ecosystem services, and contributing to the wide variety of human economic development (Lotze, 2021; Pimentel et al., 2005), starting from food security and medicines to cultural and common heritage (Cardinale et al., 2012). In addition to key anthropogenic factors, escalating threats such as introduction of invasive species, and maritime activities (Warner, 2014) are putting biodiversity at risk (Marco et al., 2015; McCann, 2000). For instance, introduction of a single invasive species can trigger ecosystem collapse, as seen in the case of the Nile perch in Lake Victoria (McCann, 2000) which contains highly ecologically diverse cichlid assemblages (Meier et al., 2023).

The marine ecosystem in northern Norway is uniquely distinct, hosting the world's largest populations of cod and herring (Jakobsen, 1987), ancient coral reefs (Freiwald et al., 2002), abundant kelp forests (Kvile et al., 2022), and vast seabird colonies, accompanied by six species of seals and 30 species of whales and dolphins (Bevanger, 2018). The northeast Atlantic is rich in marine biodiversity with thousands of species documented where fish make up about 1100 of species (Merrett, 1994). Aside from the seas, thousands of untouched lakes, rivers, and streams spanning across Norway are the ideal habitats for wild salmon, Arctic char, and trout.

## **1.1 Marine conservation**

With the expansion of the blue economy - economic growth derived from oceans - anthropogenic threats will continue to increase, bolstering the urgency of conservation efforts. The ongoing biodiversity crisis has prompted enormous efforts (Figure 1) and attempts to preserve all elements of ocean biodiversity at genetic, population, species, habitat, community, and ecosystem levels (Mace et al., 2020). Within the context of marine resources, Convention for Biological Diversity (CBD) aimed to protect 10% of marine areas (Figure 2; Pimm et al., 2014). The United Nations Convention on the Law of the Sea (UNCLOS) set out the legal framework for marine and maritime activities requiring states to protect and preserve the marine environment and sustainably use the marine biological resources (Figure 2; Nguyen, 2023). Mirroring and supporting the objectives of the CBD, Sustainable Development Goal 14 (SDG 14) addressed issues like overfishing, and marine pollution reduction (Figure 2; Kirkfeldt & Frazão Santos, 2021). Adding to these endeavors, the Global Ocean Alliance (GOA) is aiming to protect at least 30% of the world's ocean by 2030 (Figure 2), which was furthermore adopted by The Kunming-Montreal Global Biodiversity Framework as part of CBD (Stephens, 2023).

For achieving these strategic visionary goals, policy makers from each country often design action plans in form of programs, measures, regulations, and controls (Cormier et al., 2017). In practice, these action plans are turned into management plans where governmental institutions or independent entities evaluate the outcomes through measuring the performance by comparing indicators against a benchmark. In case of adaptive management, depending on the performance, such action plans can furthermore be improved (Cormier et al., 2017). Up until recently the management plans have often been focused on single species without explicitly accounting for the complex interactions that species have with other species or the ecosystems. An example of this approach is shown off the coast of Newfoundland in Canada where ignoring

ecosystem interactions brought Atlantic cod (*Gadus morhua*) fishery stock to a collapse obliging the Canadian government to impose a moratorium in 1992 (Bavington, 2010). Multiple countries including Norway have transitioned from a single-species management approach to an Ecosystem Based Management model (i.e., community and ecosystem level managements) where more focus is laid on complex interaction with other species and the ecosystem to sustain healthier marine biological resources (Fogarty, 2014). This approach involves maintaining long-term socioeconomic benefits, generating knowledge of ecosystem processes and mitigating risks of irreversible changes to the marine ecosystem (Gullestad et al., 2017).

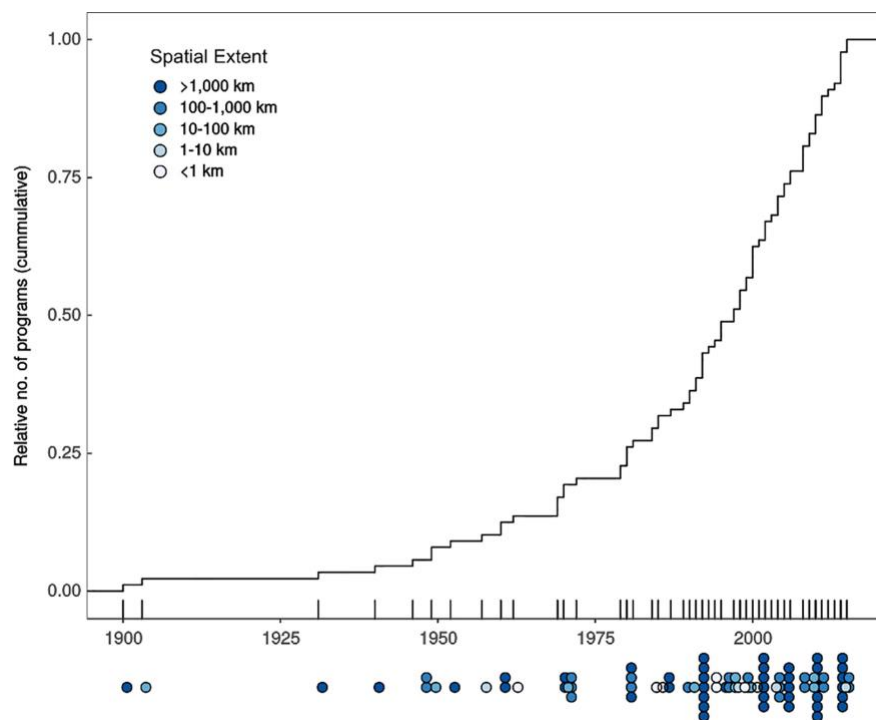


Figure 1. Cumulative number of efforts attributed to biological conservation programs since 1900 from Miloslavich et al. (2018)

Despite the international extensive efforts for marine conservation, a multitude of research reporting worsening trends on the majority of the Aichi targets - a set of 20 global targets set by CBD aimed at halting biodiversity loss (Mace et al., 2020; Rounsevell et al., 2020). For instance, four of the benchmarks highly relevant to marine conservation have not been reached by the end of the decade (Carr et al., 2020), indicating a significant failure to meet the international targets (Marco et al., 2015; Rounsevell et al., 2020). Researchers attribute this lack of fulfilment to the overly ambitious nature of these marine conservation targets, which were not grounded in measurable, scalable, and realistic terms (Green et al., 2020). Additionally, the absence of comprehensive biotic data for assessing marine biodiversity and the efficacy of

Marine Protected Areas (MPAs) also poses significant barriers to meeting these international marine conservation goals (Carr et al., 2020). Moreover, Saeedi et al. (2019) recommended continuous effort on field biotic observations coupled with environmental variables where the anthropogenic impacts are higher (e.g., fishing areas) to better achieve Aichi goals. O’Leary et al. (2016) in their summary review of 144 studies, indicated that 37% of the marine area should be protected to meet the stipulated marine conservation goals. Such contradiction indicates that there remains a divergence of opinion among studies regarding what is feasible, and what needs to be conveyed to conserve marine ecosystems effectively. This discrepancy can be attributed also to a lack of in-depth understanding of marine ecology due to data insufficiency (Fogarty, 2014). Hence, it’s apparent that an intricate balance needs to be struck in order to navigate the path between anthropogenic impact, resource biology, and marine biodiversity conservation. The collection and quantification of marine species data are fundamental to understanding ecosystem resilience and diversity (Magurran, 2013) which is a vital aspect for a sustainable blue economy (Claudet et al., 2020; Saeedi et al., 2019). In addition, marine monitoring is an integral component of marine management and is conducted under the scope of EBM. Moreover, the EBM approach calls for enhanced data collection on inferring ecosystem interactions (Pikitch et al., 2004) due to the multi-dimensionality and complexity when multi-species management is implemented (Fogarty, 2014).

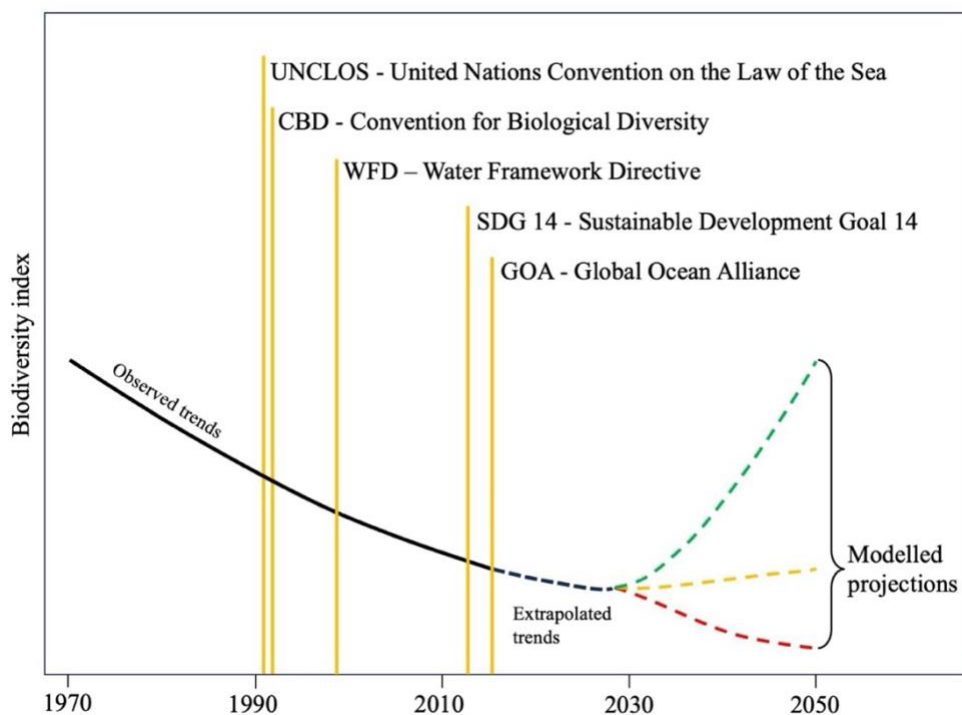


Figure 2. Biodiversity index trends and Global policy commitments affecting the trends shown with dashed lines. Modified from Mace et al. (2020)

## 1.2 eDNA marine monitoring approaches

Traditional methods such as scientific and commercial trawling, mark-recapture techniques, telemetry, hydro-acoustics, and electrofishing have been fundamental in collecting ecological data (with respect to fish) for ecosystem management and conservation (Crossin et al., 2017; Fraser et al., 2007; Lees et al., 2021; Neebling & Quist, 2011). Despite their extensive use and reliability in gathering biotic data, these methods are often invasive, potentially impacting marine ecosystems negatively (Afzali et al., 2021; Biju Kumar & Deepthi, 2006; Eigaard et al., 2017). They can also be expensive, laborious, and necessitate specialized knowledge, thus limiting their suitability for effective marine surveillance. Additionally, due to their high costs, these methods are not employed as frequently as researchers recommend, which hinders the ability to fully comprehend the marine ecology and refine management objectives (Andres et al., 2022; Burian et al., 2021; Lacoursière-Roussel et al., 2016).

A multitude of research articles have continuously advocated for less destructive methods in the monitoring of marine ecosystems (Barros et al., 2021; Fraija-Fernández et al., 2020; Gilbey et al., 2021; Knudsen et al., 2019). Recent advances in DNA sequencing and bioinformatics have enabled the use of environmental DNA (eDNA; Box 1) for non-invasive marine surveillance approach (Afzali et al., 2021; Fraija-Fernández et al., 2020). eDNA metabarcoding (Box 1) has emerged as a revolutionary method in ecology, combining cost-efficiency (Cantera et al., 2019) and high accuracy to inventory and locate marine species (Günther et al., 2018; Taberlet et al., 2012). In the field of marine science, its utility extends to marine biotic monitoring (hereafter biomonitoring; Bohmann et al., 2014; Taberlet et al., 2012), detection of invasive species (Takahara et al., 2013), and assessments of community compositions (Cilleros et al., 2019). A key advancement within these fields is the ability of eDNA to simultaneously identify multiple species, thereby contributing invaluable data regarding biodiversity patterns (Cilleros et al., 2019) and community structure (Lacoursière-Roussel et al., 2016) as required by EBM. Studies have also recognized the ability of eDNA to detect subtle biotic changes and community dynamics, often affected by anthropogenic factors (Atienza et al., 2020; Guri et al., 2023; Hansen et al., 2018; Jeunen et al., 2019; Larson et al., 2022; Turon et al., 2022). Given its numerous advantages, eDNA metabarcoding represents a pivotal advancement in marine science and conservation biology as it offers an appealing alternative to traditional methods (Bohmann et al., 2014; Cantera et al., 2019; Gilbey et al., 2021). As such, its wider applicability could foster management policies and enrich conservation strategies, eventually steering towards more robust action plans for EBM (Gilbey et al., 2021).



While eDNA offers several advantages, it also encounters challenges arising from biological (Yates et al., 2022), environmental (Collins et al., 2018) and technical factors (Hansen et al., 2018; Shelton et al., 2023; Figure 6) which will be denoted as latent processes hereafter. In nature, different species shed DNA differently. Such species-specific DNA shedding rates (latent process 1 in Figure 6), differ also based on factors such as size, metabolic rate, and life stage. These differences can result in some individuals, particularly those that are smaller and have higher metabolic rates, releasing more DNA relative to larger individuals (Yates et al., 2022). Consequently, their eDNA might be detected more frequently, giving an erroneous impression of higher abundance relative to other species, and infusing considerable bias into the observational data. Additionally, due to the stochastic nature of the Polymerase Chain Reaction (PCR; i.e., PCR stochasticity), some species' DNA tends to be amplified more efficiently than others (latent process 4 in Figure 6), causing amplification bias which leads to misrepresentations in the final observational (metabarcoding) outputs (Shelton et al., 2023). Additional to biologically risen biases, environmental factors like lateral water movement, temperature, pH, and microbial activity further convolute the challenges (Andruszkiewicz et al., 2019; Collins et al., 2018; Harrison et al., 2019; latent process 2 in Figure 6). Such factors can complicate the source of eDNA thus increasing the ambiguity of species distribution and abundances. For instance, in aquatic environments, strong water currents can transport and dilute DNA (Andruszkiewicz et al., 2019), causing potential discrepancies between the detected DNA and the actual distribution of organisms. Furthermore, vertical settling of DNA (as particle sedimentation; Mauvisseau et al., 2021) can confound analyses because surface dwelling species areas can be detected at deeper water masses (latent process 2 in Figure 6; Canals et al., 2021; Turner et al., 2015). Equally, conditions like high temperature and acidity can accelerate DNA degradation, potentially reducing the eDNA detection rate and leading to false negatives. High microbial activity, which typically leads to faster eDNA degradation (Mächler et al., 2018), can add another layer of complexity to the interpretation of eDNA metabarcoding data (Collins et al., 2018). Such multipartite biases (Figure 6) lead to discrepancies when comparing eDNA metabarcoding outputs with traditional methods such as trawling, presenting potential inaccuracies. These latent processes, while not directly observable, significantly influence the result outcomes and are paramount to account for when working with eDNA surveys. Given these limitations, scholars often utilize metabarcoding data through presence/absence metrics as a more robust approach towards more reliable analysis (Guri et al., 2023).

**Box 1: Environmental DNA (eDNA).** Free DNA molecules in environment originating from sources like mucus, sweat, skin, urine, sperm, pollen and rotting of cells can be captured (Bohmann et al., 2014) using dedicated filters (Figure 3). After capturing the free DNA, it gets isolated by using reagents to remove impurities (i.e., proteins, polysaccharides, and other cellular debris) that can prohibit the following steps. The isolated DNA is thereafter amplified using polymerase chain reaction (PCR). This is done using a set of primers (short, single-stranded DNA sequences, typically artificially synthesized) that serve as the starting point for DNA amplification. Primers enables the amplification (selection) only of the specific regions of DNA (organisms) that we are interested (e.g., 12S region for detecting vertebrate organisms, COI for eukaryotes, 16S for bacteria, and trnL for plants). The amplification process (typically ca. 40 cycles), makes millions of “photocopies” of the captured DNA (Figure 3), an essential procedure due to the small amounts of DNA present in samples (Abath et al., 2002). Subsequently, the captured and amplified DNA - referred to as amplicon (Bohmann et al., 2014) - is sequenced using Next Generation Sequencing (NGS) machines (Figure 3). Bioinformatics tools (software and algorithms) are then employed to identify the sequences to their species of origin (also differently called as Molecular Taxonomic Operational Units - MOTU; Ji et al., 2013). In this workflow, multiple species (MOTUs) can be identified and detected simultaneously, a process known as eDNA metabarcoding.

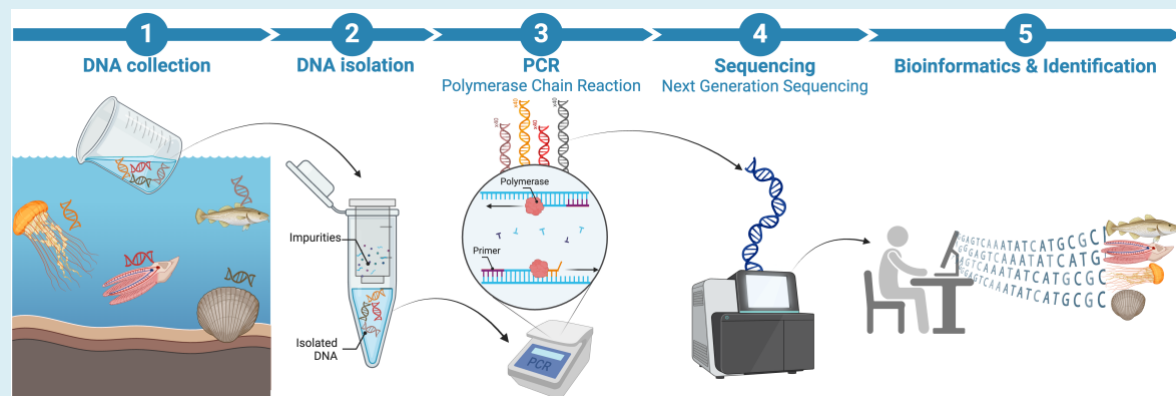


Figure 3. Schematic workflow of environmental DNA metabarcoding. Created using BioRender.

Furthermore, some of the aforementioned biases, particularly those associated with the PCR stochasticity, can be tackled by employing quantitative PCR (qPCR; Box 2) or droplet digital PCR (ddPCR; Box 3) instead of conventional PCR (Box 1). These approaches enable precise measurement of the DNA concentrations in environmental samples, which can enable the researchers to draw conclusions for distribution patterns and abundances of various marine species in their natural habitats. The application of quantitative methods provides invaluable insights for conservation efforts, allowing for a more accurate monitoring of biodiversity and

ecosystem health. They are also instrumental in the early detection of invasive species, and they enhance our capabilities in pathogen surveillance, offering the potential for more timely and effective responses to outbreaks. Nevertheless, despite these benefits, the other biological and environmental latent processes (Figure 6) are yet unresolved which hampers the applicability of quantitative eDNA methods to comprehensively explain the biology and ecology of marine environments. Moreover, quantitative methods are generally designed to target individual species, not suitable for whole ecosystems survey and management. When scaled for multi-species, quantitative methods can be more expensive and less efficient compared to metabarcoding (multi-species) or other traditional methods (Schneider et al., 2016). These limitations underscore the need for continued research and development in eDNA methodologies. Taking these factors into account, the field of eDNA research needs further refinement and understandings to better succeed as a tool for biomonitoring and EBM.

**Box 2: Quantitative PCR.** Similar to metabarcoding (Box 1), environmental DNA, after being collected and isolated get amplified during a PCR process (Figure 4). The latter PCR is defined as quantitative PCR (qPCR also called Real-Time PCR), which uses fluorescent probes that attach to the targeted DNA (Figure 4). As the DNA is amplified during the PCR process, the amount of fluorescence released is proportionally related to the amount of DNA present (Figure 4). The number of cycles required to pass the fluorescence threshold ( $C_t$ ) indicates the amount of target DNA. To accurately quantify the amount of DNA in environmental samples, qPCR compares the fluorescence from the latter samples to standard samples (samples with known concentrations constructed in laboratory; Taylor et al., 2015).

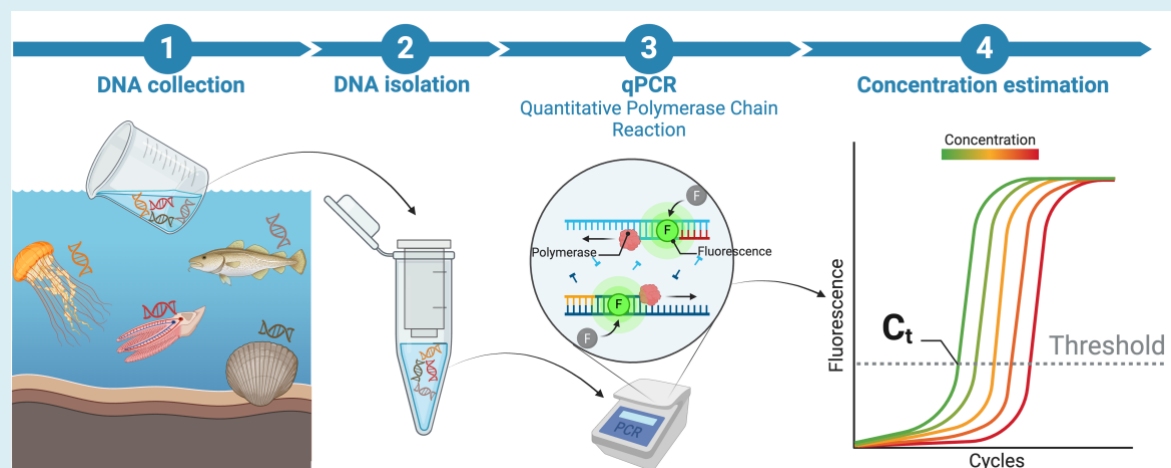


Figure 4. Schematic workflow of quantitative PCR. Created using BioRender.

**Box 3: Droplet digital PCR.** Droplet digital PCR (ddPCR) is a method used for precise quantification of DNA in a sample. In ddPCR, the sample is partitioned into thousands of nano droplets (Figure 5). The amplification of the target DNA occurs within each individual droplet during the PCR process (using fluorescent probes), and the fluorescence signal in each droplet is counted (Figure 5). By analyzing the fluorescence, each droplet is categorized as either positive or negative for the target DNA (Figure 5). The ratio of positive droplets to the total droplets is used to calculate the quantity of the DNA target, using a statistical approach known as Poisson statistics. This digital approach enables the accurate calculation of the absolute quantity of the target DNA without the need for standard curves (like in qPCR). It is also highly precise and resistant to certain inhibitors, making it a valuable tool in various research and diagnostic applications.

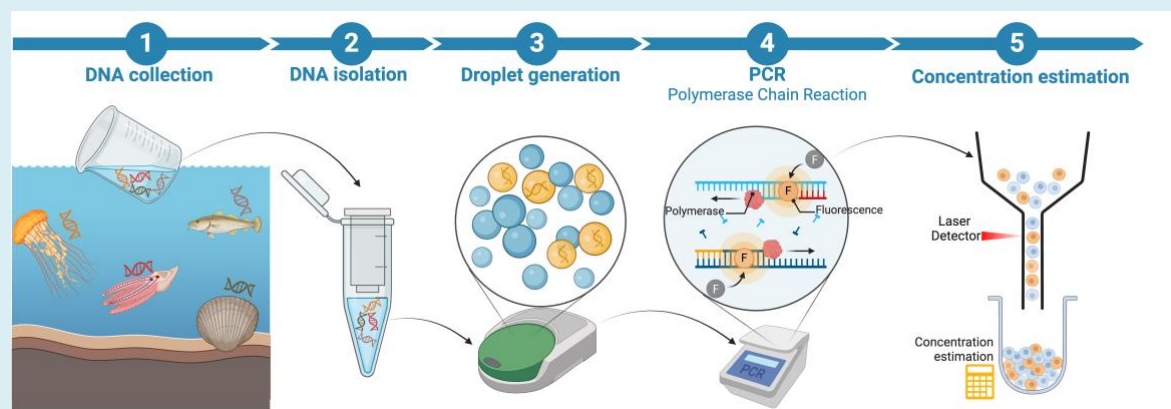


Figure 5. Schematic workflow of Droplet Digital PCR. Created using BioRender.

### 1.3 The need for this study

The current rate of species extinction, largely due to human activities, has led to a significant increase in the demand for data in biodiversity and ecosystem assessments. This urgency reflects the scientific community's consensus on the importance of intensifying our sampling and biomonitoring approaches (Grey et al., 2018; Saedi et al., 2019). Given this context, biological resource management is becoming increasingly reliant on continuous streams of data to manage ecosystem health while inducing little stress to ecosystems. Additionally, there is an acknowledged appeal for the adoption of non-invasive methods to complement current ecosystem-based management strategies. This is where enhancing our understanding and refining environmental DNA (eDNA) techniques becomes imperative. By addressing and overcoming the limitations of eDNA methods, we can acquire more detailed and accurate information about marine ecosystems. In turn, this can inform better, cheaper, and more

sustainably management decisions. Improvements in observational capabilities through refined eDNA methods are likely to contribute significantly to shaping effective action plans, ultimately supporting the attainment of international biodiversity conservation goals.

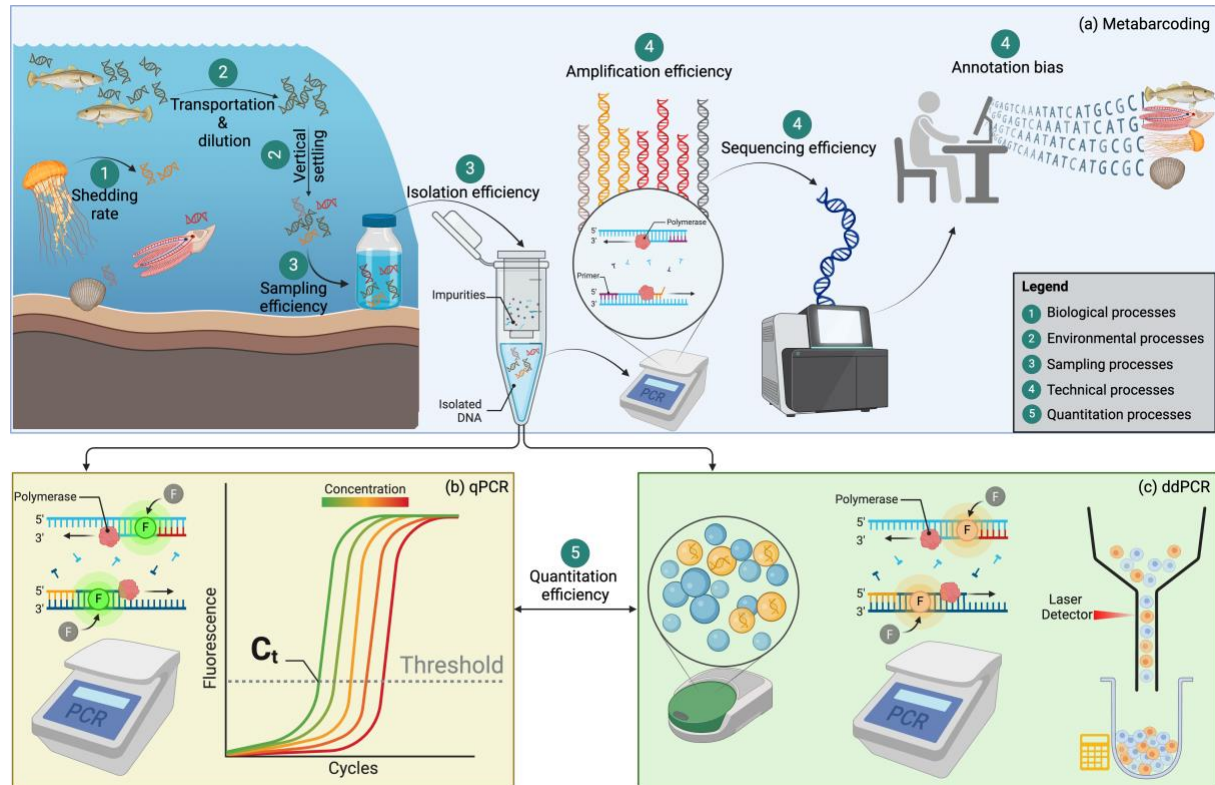


Figure 6. Latent processes (biological, environmental, sampling, technical and quantitation processes) of eDNA metabarcoding, qPCR and ddPCR investigated in this study (together with their workflow). This workflow is modified from Gold et al. (2023) and Shelton et al. (2016). Created using BioRender.

## 1.4 Aims and objectives

The overall aim of this study was to shed light on the latent processes (Figure 6) and mitigate limitations of eDNA methods in use of biomonitoring and quantitative ecology. Specifically, this study aimed to empirically evaluate the latent processes by developing novel approaches and models focusing only on fish as the organism of observation.

This study sets out to accomplish three integrated objectives (translated in three research papers) for achieving this aim. The first was to investigate and establish the spatial scale representation of eDNA (latent processes 2 in Figure 6) for ecological community inference and to investigate the data treatment approaches which minimizes biases related to technical interpretation of metabarcoding data (latent process 4 in Figure 6) with respect to maximizing the efficiency for biomonitoring and biodiversity assessments (**Paper I**). Second, to determine and estimate the quantification efficiency (latent processes 5 in Figure 6) of the two well-

established quantitation methods (qPCR and ddPCR; **Paper II**). Third, to empirically estimate the metabarcoding latent processes (latent processes 1, 2, and 4 in Figure 6) and provide frameworks for quantitative ecology from eDNA (**Paper III**). I note that latent processes 3 are kept constant throughout the entire study thus circumvented in the parametrization of the processes.



## 2 Materials and methods

This project took place in northern Norway, where three fjords were selected, namely, Balsfjord, Olderfjord, and Frakkfjord (Figure 8). The study's coastal area is located at the 70° latitude and is characterised by a sub-Arctic climate and water regime. Multiple biomonitoring methods were used across the fjords (Figure 7) to answer different research questions. **Paper I** utilized eDNA samples collected in March 2021 from all the three fjords, while **Paper II** utilized eDNA samples collected in October 2019, 2020, and 2021 in Balsfjord and in March 2021 and 2022 in Frakkfjord employing qPCR and ddPCR methods. Subsequently, **Paper III** utilized only the Balsfjord samples included in **Paper II**. Additionally in **Paper III**, trawl samples of fish catch count coupled with eDNA samples were used for comparison between the two methods and estimation of latent processes.

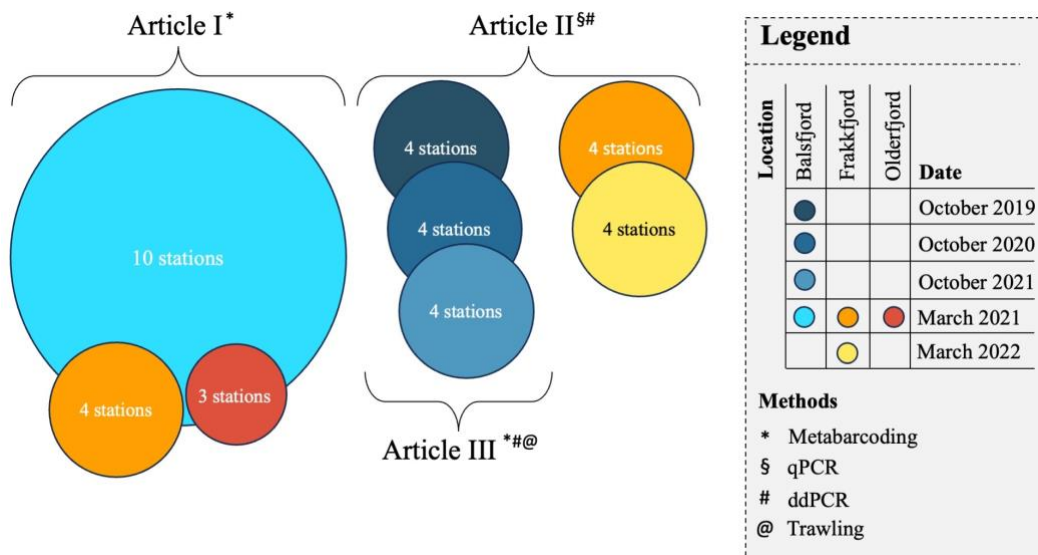


Figure 7. Sample utilization among all the papers and objectives.

### 2.1 Environmental DNA collection and isolation

At each eDNA sampling station (Figure 8), water samples each consisting of three biological replicates (using different Niskin bottles) were gathered at three distinct levels - surface level at about 10 meters depth, pycnocline level at about 50 meters depth, and bottom layer at about 10 meters above the sea floor. The water collected was filtered onboard through a 0.22 µm Sterivex filters and after the filtration was complete, the filters were stored until DNA isolation. The DNA was isolated using DNeasy PowerWater Sterivex Kit (Qiagen GmbH, Hilden Germany) following the manufacturer's protocol with slight modifications (all steps involving PowerBead Tubes were bypassed). Subsequently all samples were eluted at 100 µL volume.

### 2.1.1 Metabarcoding

The eDNA samples subjected to metabarcoding workflow, after being isolated underwent the PCR process for DNA amplification following the PCR program including an initial denaturing step of 95°C for 10 min, 40 cycles of 95°C for 30s, 60°C for 30s, and 72°C for 30s and a final extension step of 72°C for 5 min. MiFish primers (Forward: 5'-GTCGGTAAACTCGTGCCAGC-3'; Reverse: 5'-CATAGTGGGGTATCTAATCCCAGTTTG-3'; Miya et al., 2015) targeting a fragment of the mitochondrial gene region of the 12S rRNA gene was used during the amplification of eDNA. After the amplification, the PCR product was sequenced using the Ion 530™-sequencing chip with the 200 bp protocol on a Ion GeneStudio™ S5 system (Thermo Fisher Scientific). The sequences were thereafter filtered for erroneous sequences and were thereafter clustered into Molecular Operational Taxonomic Units (MOTU) using SWARM v2 (Mahé et al., 2014) with a distance of  $d = 3$ . MOTUs were taxonomically annotated using the *ecotag* algorithm (Boyer et al., 2016) and by performing BLAST function to match the sequences against the NCBI nucleotide (nt) database.

### 2.1.2 qPCR

The eDNA samples subjected to qPCR analysis, after being isolated, were run in Applied Biosystems 7500 Fast real-time PCR System, with duplexed reactions for cod assays alongside herring and saithe assays (Table 1) following a thermocycler protocol of 10 min at 95 °C for denaturation followed by the cycling stage of 42 – 52 cycles of 15 sec at 95 °C and 1 min at 58 °C. Different channels and fluorescence were used for each assay (6-FAM for cod, JOE for herring, and Cy3 for saithe; Table 1), and all thermocycler reactions were run in 20 µL volume utilizing 2 µL of DNA template. Standard samples of known concentration from  $10^{-1} - 10^6$  copies/µL were used in all qPCR runs plates to calibrate the quantities. The assays were optimized before running samples.

### 2.1.3 ddPCR

The eDNA samples subjected to ddPCR analysis, after being isolated, were run in QX200 ddPCR system (Bio-Rad) where ca. 20,000 droplets were generated on 20 µL reaction volumes.



Similar to qPCR, the samples were run in duplexed reactions for cod assays alongside herring and saithe assays with different channels and fluorescence for each assay (VIC for cod, 6-FAM for both herring and saithe). Thereafter the samples were amplified through a PCR program as follows: 10 min at 95°C for enzyme activation, followed by 44 cycles of denaturation for 1 min at 95°C and primer annealing and elongation for 2 min at 55.6°C, with a ramp rate of 2°C per s, followed by 10 min at 98°C and stored at 4°C. Alongside the environmental samples we ran standard samples of known concentration from  $10^{-2}$  –  $10^4$  copies/ $\mu$ L for each species to construct the relationship between positive droplets and known concentration.

Table 1. Sequences for qPCR and ddPCR assays targeting 103-bp region of the ATPase gene of Atlantic cod (*Gadus morhua*) and saithe (*Pollachius virens*), and cytochrome b sequence of Atlantic herring (*Clupea harengus*). All gene regions belong to the mitochondrial DNA. Taken from **Paper II**.

Target	Primers and probes	Sequence	Reference	Dye	
<i>Gadus morhua</i>	Forward	GAD-FII	GCAATCGAGTYGTA TCYCTHCAAGGAT	(Taylor <i>et al.</i> , 2002)	
	Reverse	GAD-R III	GCAAGWAGYGGHG CRCADTTGTG	(Nash <i>et al.</i> , 2012)	
	qPCR probe	Custom	CTTTTTACCTCTAAA TGTGGGAGG		FAM
	ddPCR probe	Custom	CTTTTTACCTCTAAA TGTGGGAGG		VIC
<i>Clupea harengus</i>	Forward	Cluhar_CYBF 14928	CCCATTTGTGATTG CAGGGG	(Knudsen <i>et al.</i> , 2019)	
	Reverse	Cluhar_CYBR 15013	CTGAGTTAAGTCCT GCCGGG	(Knudsen <i>et al.</i> , 2019)	
	qPCR probe	Cluhar_CYBP 14949	TACTATTCTCCACCT TCTGTTCCTC		JOE
	ddPCR probe	Cluhar_CYBP 14949	TACTATTCTCCACCT TCTGTTCCTC		FAM
<i>Pollachius virens</i>	Forward	Saithe-F	GAATCCCAATAATT TTAATAGCCT	Unpublished/Joha nsen et al. 2018	
	Reverse	Saithe-R	TCGATTGCTTAGTC ATCGAGA	Unpublished/Joha nsen et al. 2018	
	qPCR probe	Custom	TGATTACTCATCCCT ACG		Cy3
	ddPCR probe	Custom	TGATTACTCATCCCT ACG		FAM

## 2.2 Trawling

Bottom trawls study samples were established during Norwegian coastal annual surveys in October 2019, 2020, and 2021. Bottom trawls were conducted in four sampling stations in Balsfjord alongside eDNA samples (Figure 8). During the surveys, catch count and weight of all fish species were recorded. The trawl surveys utilized a standard sampling trawl known as Campelen 1800 with an 80 mm (stretched) mesh size in the front section and 22 mm in the codend and trawl sweeps were ca. 40 meters in length.

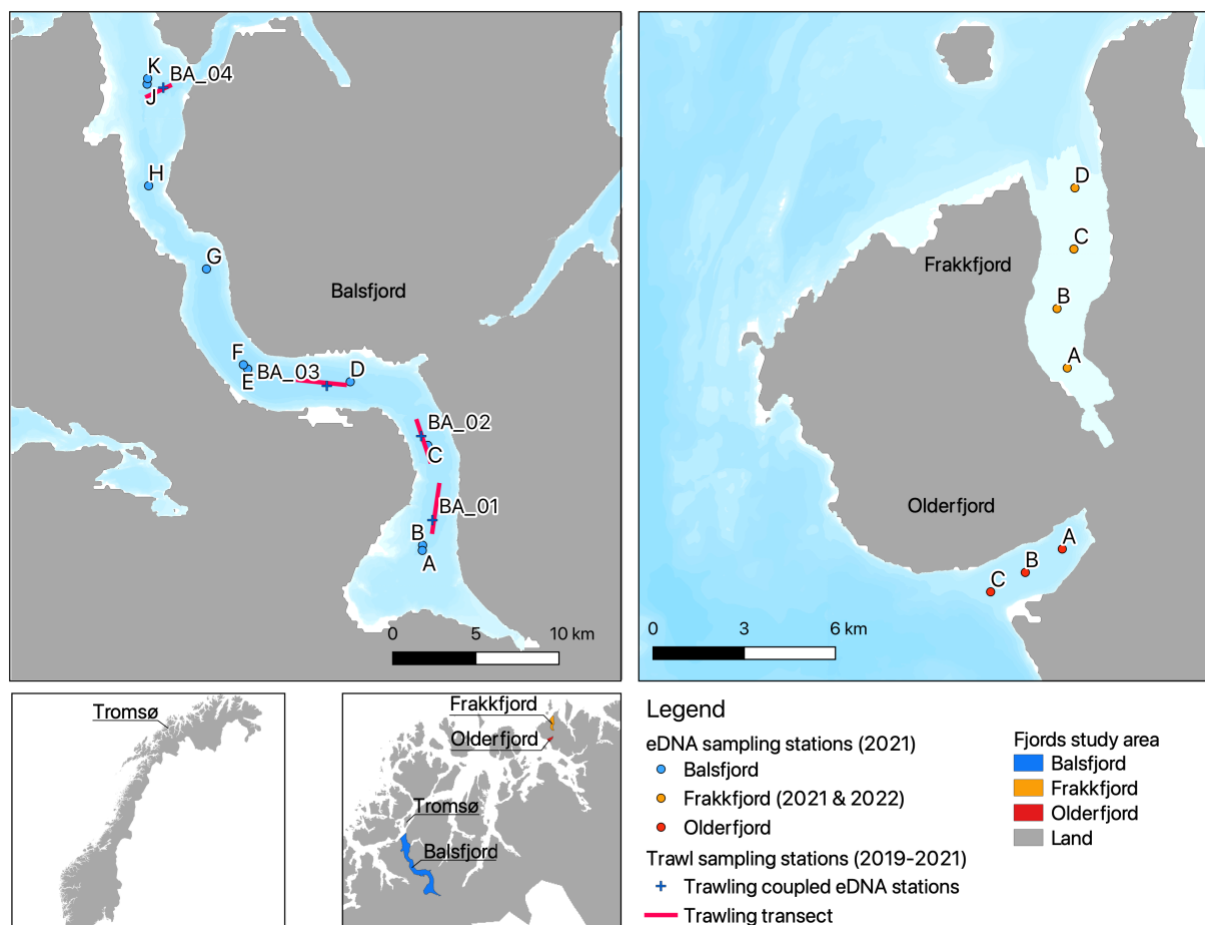


Figure 8. Map over the study area and sampling stations' distribution.

## 2.3 Short description of analysis

All analyses and visualizations were conducted using R (R Development Core Team, 2011) and Stan (Stan Development Team, 2023). In **Paper I**, my focus was on the efficiency of two different data treatment approaches (i.e., presence/absence and eDNA index; Kelly et al., 2019) for yielding the highest community dissimilarity. To assess the representativeness of eDNA for

community structure, I analyzed dissimilarity in relation to spatial trends both horizontally and vertically.

Furthermore, in **Paper II**, I conducted an examination to determine whether qPCR or ddPCR offers less bias and consequently, is more reliable for biomonitoring applications based on their robustness of accurately detecting presence of organisms. Within this paper, I compared the state-of-the-art knowledge on qPCR and ddPCR techniques and estimated the measurement error associated with each method.

With the preferred quantitation method identified, my research in **Paper III** revolved around the capability of quantifying metabarcoding samples. By applying Bayesian inference, I modelled the intricate relationship among the various methodologies used in marine ecosystem surveys (i.e., metabarcoding, traditional sampling such as trawls, and quantitation methods like ddPCR; Figure 9). Through the Bayesian model (Figure 9) I aimed to establish the conversion parameters that would enable the accurate translation of eDNA metabarcoding data into quantitative ecological information ( $\theta$  parameter in Figure 9) that are on par with traditional survey outputs.

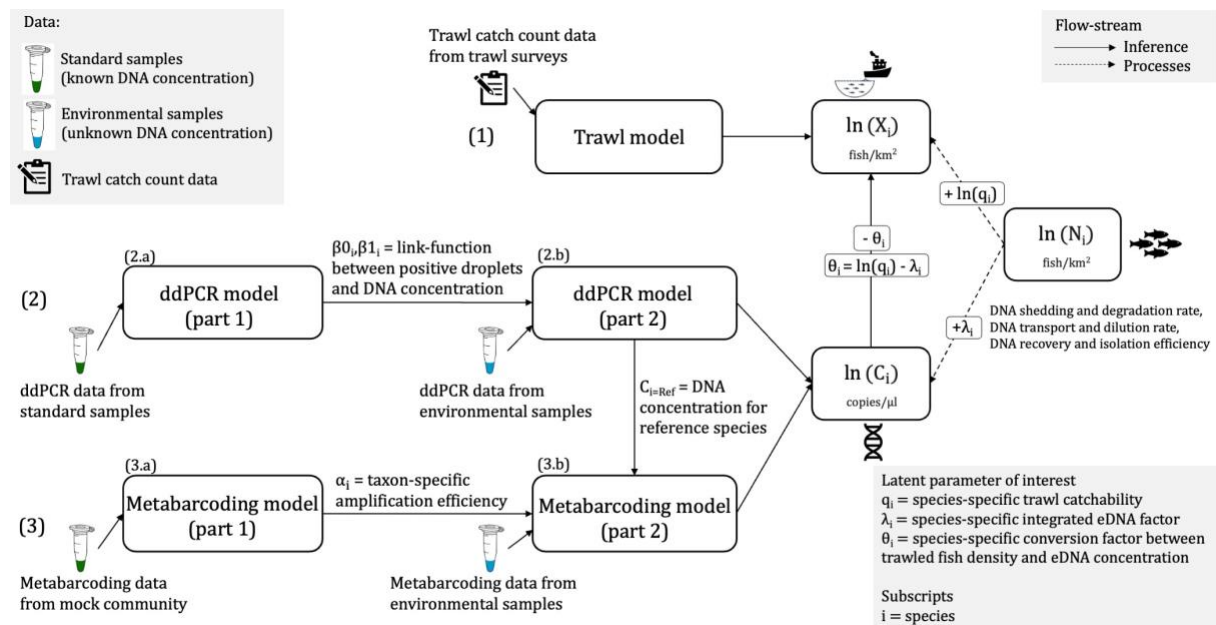


Figure 9. Simplified schematic overview used in **Paper III** of the joint Bayesian model workflow including all the inferences and processes. Note that the model specifies  $\theta$  in the natural-log scale indicating the magnitude of conversion between fish densities and eDNA concentration. Extracted from **Paper III**.

### 3 Main findings

The comparison of two different data treatment approaches, qualitative (presence/absence) data and the semiquantitative (eDNA index), found distinct differences in the results on the community composition across the studied fjords. The semiquantitative approach (Kelly et al., 2019) was more effective at discriminating the fish community composition between the fjords compared to the qualitative approach (**Paper I**). This indicated that metabarcoding provides some level of information regarding marine organisms' quantities. An additional analysis revealed that when using semiquantitative approach, a smaller number of samples (namely one sampling station per each fjord) are required to significantly detect differences in fish community composition (Figure 10; **Paper I**).

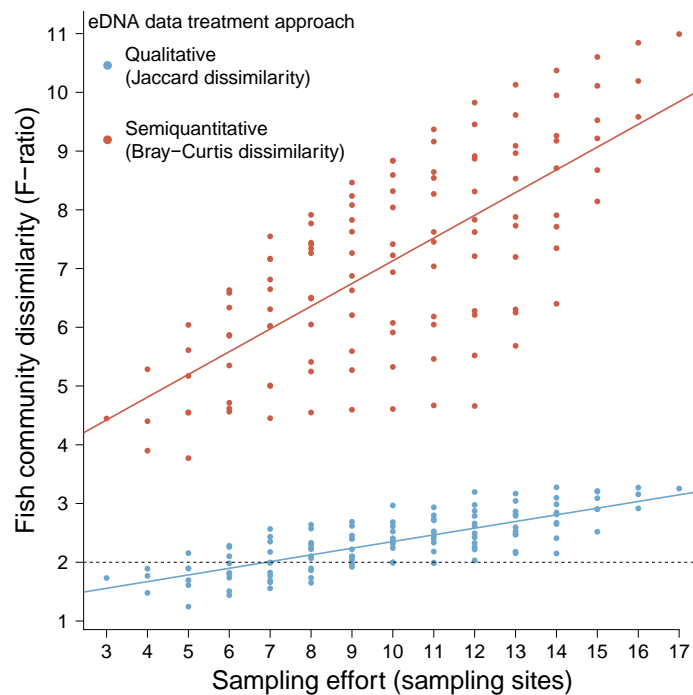


Figure 10. Dissimilarity of community composition between ecosystems measured as the PERMANOVA pseudo-F ratio using qualitative (blue; Jaccard dissimilarity matrix based on the presence/absence of species) and semiquantitative (red; Bray-Curtis dissimilarity matrix based on the inverse Wisconsin double-standardization of amplicon reads) approaches. The analyses were performed on progressively removed sampling stations, with each dot representing the mean pseudo-F of the sample combination (iterated 999 times) for the given number of remaining sampling stations. Regression lines for each approach are shown with their colours respectively. The horizontal black dashed line, representing an approximation of pseudo-F value ( $f = 2.01$ ) equal to  $p = 0.05$ , indicated that minimum three and thirteen sampling stations are required for detecting significant differences between the three ecosystems respectively for semiquantitative and qualitative approach. Extracted from **Paper I**.

Furthermore, when using semiquantitative approach, the community composition differed significantly at distances between 26 to 40 kilometers on the horizontal spread (Figure 11) implying that eDNA metabarcoding has a spatial representation scale to ca. 30 km in the study area (Figure 11; **Paper I**). With respect to vertical representation, **Paper I** identified that only

the surface and bottom communities changed significantly, indicating that mid-water (pycnocline) communities were a combination of the other two community compositions. This also indicated that the communities change gradually throughout the vertical water column (**Paper I**). This gradual change suggested a vertical DNA settling process, supported by the fact that the surface communities contained fewer species than the bottom ones, and while all surface species were found at the bottom, the reverse was not true (**Paper I**).

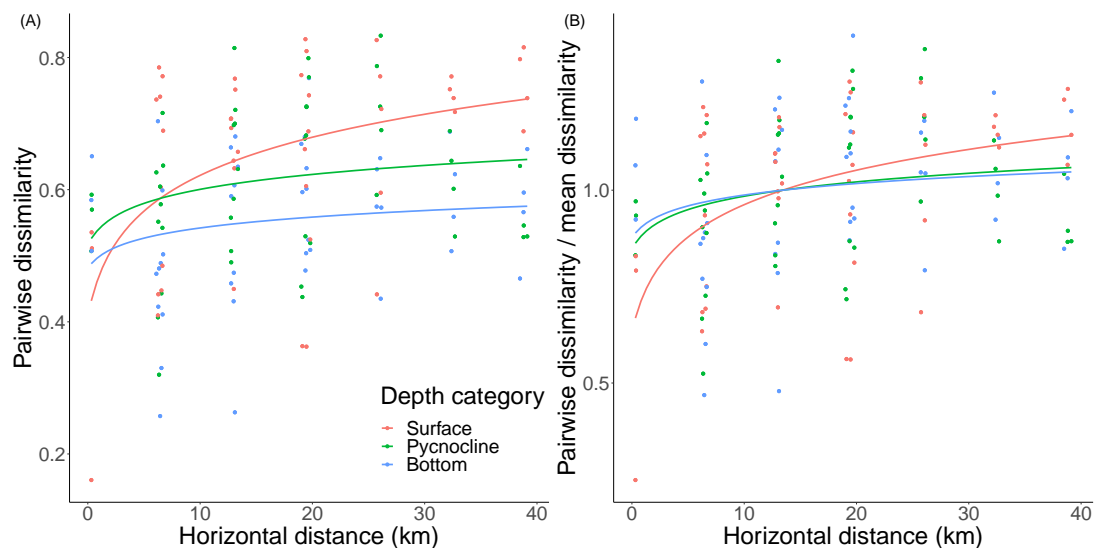


Figure 11. Correlation plot of samples' pairwise dissimilarity (a) and the ratio of pairwise dissimilarity over the average dissimilarity per each depth (b) along samples' pairwise horizontal distance. The dissimilarity measured using Bary-Curtis index based on inverse Wisconsin double-standardization of amplicon reads indicated higher differences on surface communities (a) and average dissimilarity was achieved after 16 km of horizontal distance between samples (b). The standard errors are presented for both plots. Extracted from **Paper I**.

Moreover, the comparison of the two molecular methods, quantitative PCR (qPCR) and droplet digital PCR (ddPCR) revealed that both methods, overall, estimated equivalent eDNA concentrations (Figure 13a) across the analyzed environmental samples (**Paper II**). However, the sensitivity analysis indicated that ddPCR demonstrated higher detection rates at the lower end of DNA concentrations ( $10^{-2} - 10^0$  copies/ $\mu\text{L}$ ; Figure 12) which are commonly found in marine environments (**Paper II**). Additionally, the analysis revealed that the precision of ddPCR outperformed that of qPCR (**Paper II**). This was evident in the reduced variation by 0.5 to 1 order of magnitude between technical replicates of ddPCR compared to that of qPCR specifically on low DNA concentrations (as mentioned above; Figure 13b; **Paper II**). This increased precision of ddPCR can be attributed to its nature as an endpoint PCR method, making it less susceptible to issues of PCR inhibition. Consequently, the results of **Paper II** highlight the greater reliability of ddPCR, particularly for low concentration eDNA samples,

and additionally advocate for the use of standard samples to achieve more interpretable ddPCR results.

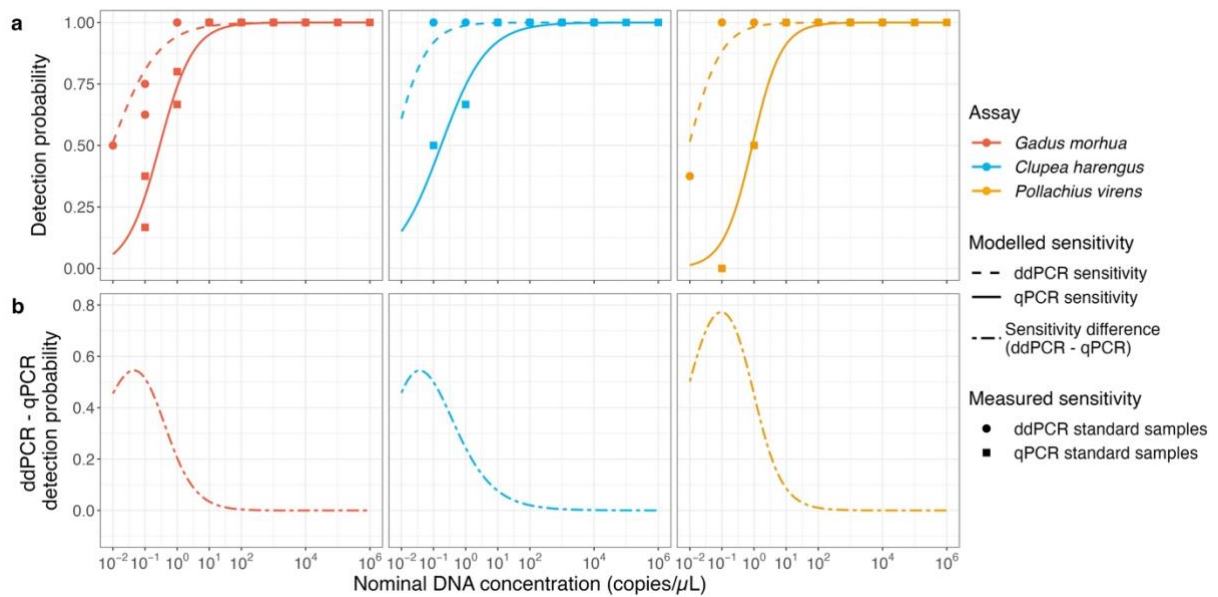


Figure 12. Sensitivity of (a) ddPCR and qPCR for three assays (cod = red, herring = blue, and saithe = orange) shown as modelled detection probability as a function of nominal DNA concentration. Difference of detection probability (b) between ddPCR and qPCR is also shown over the eDNA concentration where positive values indicate higher ddPCR detection probability and zero indicates similar detection probabilities between the two employed methods. Extracted from **Paper II**.

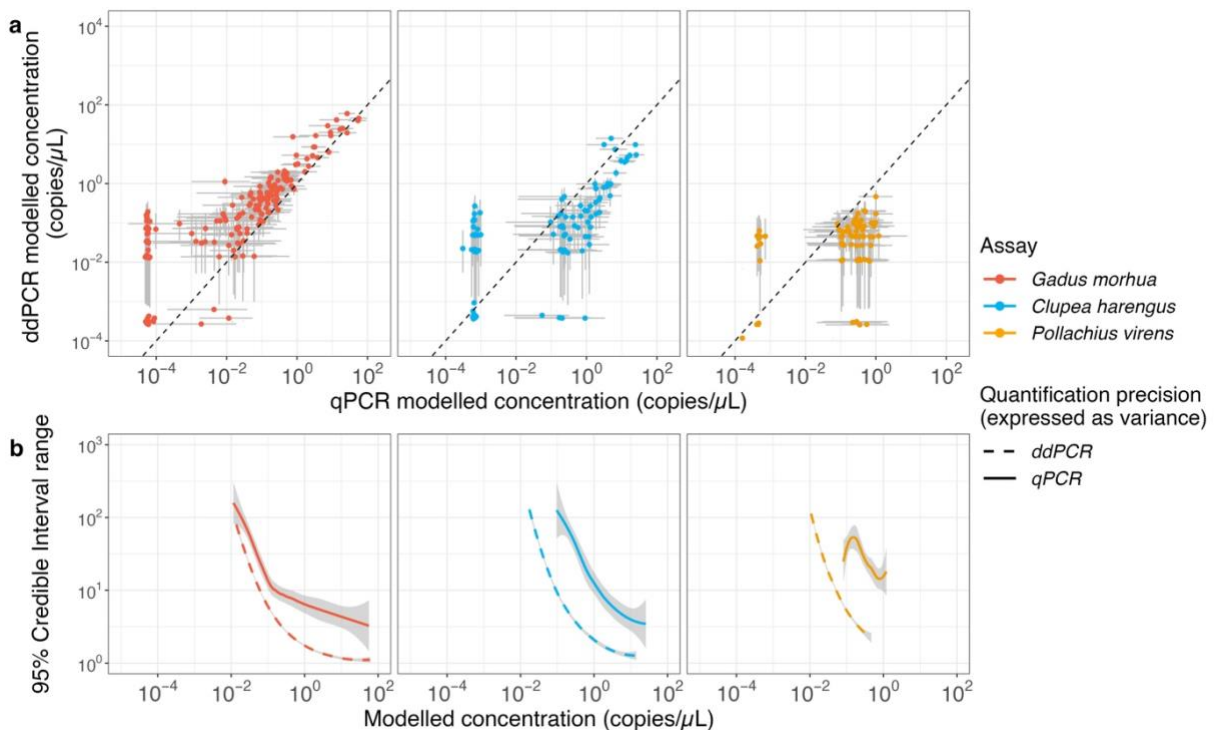


Figure 13. Comparison of modelled quantities between qPCR and ddPCR and their credible interval (grey bars) for three assays (cod = red, herring = blue, and saithe = orange; a). Concentration below 10<sup>-3</sup> for both methods are considered non-detect. The difference between ddPCR and qPCR credible intervals is also shown (b) and grey shade for its variance. Extracted from **Paper II**.

Lastly, **Paper III** introduced a statistical framework that leveraged the power of eDNA metabarcoding for abundance estimation by combining the latter method with ddPCR, standard samples (in form of a mock community - multiple species of known DNA concentration analyzed together for assessment of PCR bias), and trawl catches (Figure 14).

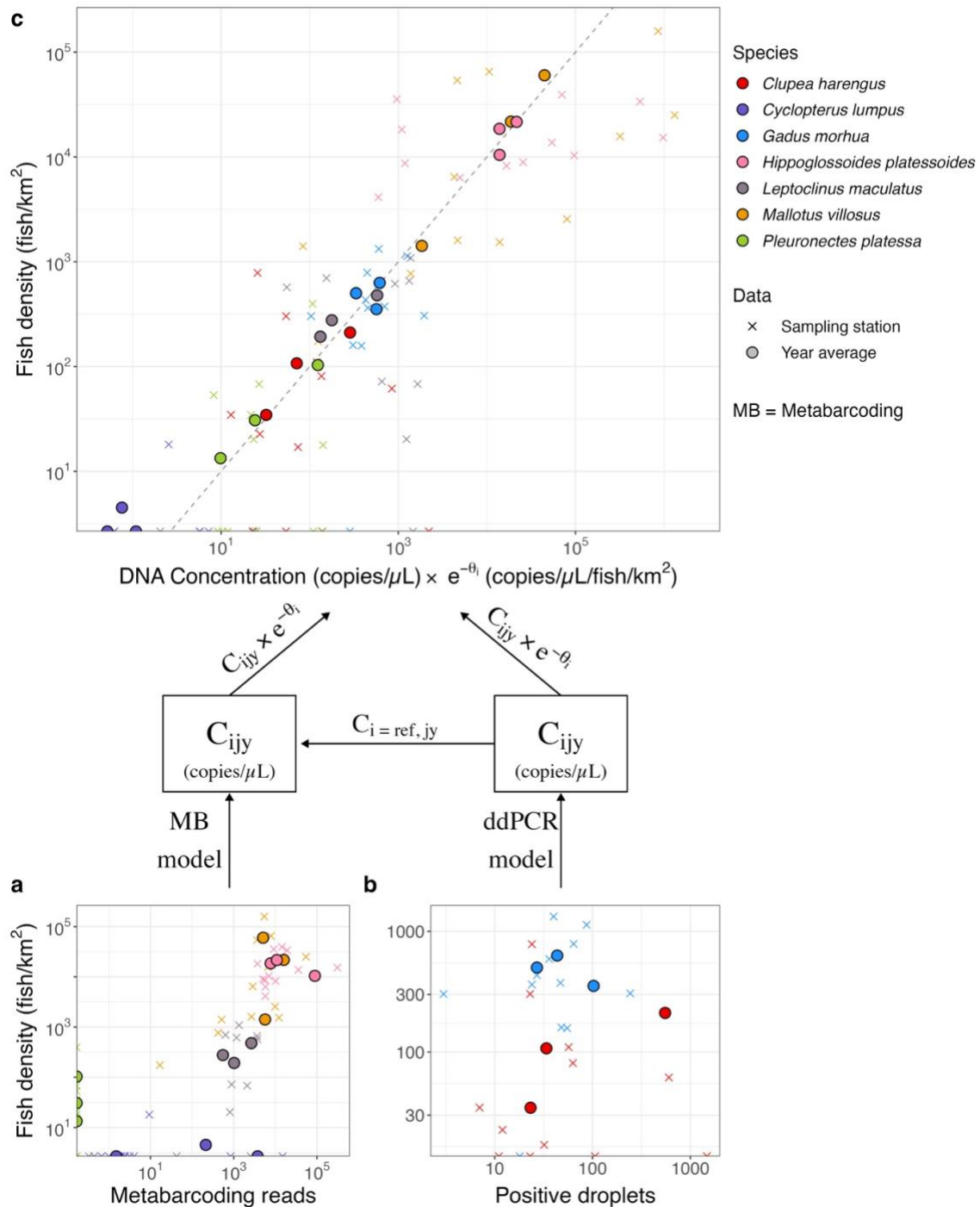


Figure 14. The joint Bayesian model workflow translating metabarcoding reads (a) and ddPCR droplets (b) into fish density (expressed as eDNA concentration \* and conversion factor; c) and their correlation with the fish density estimated from trawl observations in the y-axis.  $C$  = eDNA concentration and  $\theta$  = conversion factor between fish density to eDNA concentrations. Metabarcoding and ddPCR models indicated here are model compartments within the joint Bayesian model. The plot (c) indicates model fit and reliable parameter estimation for linking trawl and eDNA observations. Extracted from **Paper III**.



This framework utilized Bayesian statistics incorporating an “integrated DNA factor –  $\lambda$ ” (Figure 9) which account for the combined processes of DNA shedding, degradation, transport, dilution, recovery, and isolation (latent processes 1, 2, and 4 in Figure 5) and trawl catchability ( $q$ ) which account for the gear selectivity, to estimate the DNA to trawl catch count “conversion parameter -  $\theta$ ” (Figure 9). The estimated conversion parameters demonstrated high consistency across most species (Figure 14), establishing the proposed method as a robust approach for using eDNA metabarcoding as a quantitative molecular tool with respect to biomonitoring and conservation biology (**Paper III**). This framework can be tailored to incorporate different sampling methods and different marine environment (**Paper III**). Lastly, **Paper III** yielded unique conversion parameter ( $\theta$  parameter in Figure 9) for each examined taxon, implying that the parameters addressed (latent processes 1, 2, and 4 in Figure 5) manifest at species level and thus species-specific parameters should be incorporated in the future analysis of eDNA.



## 4 Discussion

Here I discuss how the findings of this thesis (three research papers) can foster the applicability of eDNA surveys in marine biomonitoring with respect to data collection for marine management either as an alternative method to traditional ones or as a complementary one. Additionally, I touch upon advantages and future consideration of eDNA into marine monitoring and management.

### 4.1 Scale representation of eDNA

Environmental DNA (eDNA) provides a novel and promising approach for monitoring marine biodiversity, including the detection of fish presence. Typically, fish eDNA originates from intestinal cells, sloughed skin, scales, or mucus and can exist as free DNA unlike the eDNA of microbial organisms, which comprises also DNA from living cells (Bohmann et al., 2014). The free DNA exists in intracellular or extracellular form (Barnes & Turner, 2016) with a typical size of 10  $\mu\text{m}$  (Sassoubre et al., 2016). The persistence and detectability of eDNA in aquatic environments is influenced by several factors (Lamb et al., 2022; Salter, 2018) and may vary drastically depending on whether it is found in marine, freshwater or sediments (Collins et al., 2018). For instance, eDNA is more persistent in the marine compared to freshwater environments due to more stable temperatures and higher pH, salinity, and ionic content in the former one (Collins et al., 2018; Harrison et al., 2019). Additionally, sediment (benthic) environments have been displaying significantly longer persistence compared to aqueous eDNA, most likely due to higher affinity of eDNA to sediments (Mauvisseau et al., 2021) and additionally due to the eDNA sedimentation processes, leading to higher accumulations at the bottom (Salter, 2018; Turner et al., 2015). In ice-free fjords (typical of northern Norwegian fjords), DNA particles can also sink through lateral advection (Canals et al., 2021). The process of eDNA sedimentation was also observed in **Paper I** as the fish community composition changed gradually through the water column with the significant community difference between surface and bottom samples. These depth community differences are apparent due to different water regimes as surface waters are heavily affected by light exposure and high water movement (in form of waves) thus leading to higher eDNA degradation (**Paper I**). Conversely, as bottom waters are less prone to such factors, alongside the sedimentation process, they serve as eDNA repository, hence longer snapshot representation of the community (**Paper I**).

The level of eDNA preservation in marine environment leads eDNA fragments to persist only a few days above detection threshold in seawater, and the decay of eDNA beyond detectability happens at a scale of days or weeks (Thomsen et al., 2012). Turner et al. (2015) observed that the eDNA in sediments persisted for over 90 days, compared to maximum persistence times of 25 days previously recorded for the water column. In summary, these findings indicate that bottom water or even sediment samples can enable researchers to prolong the timescale of eDNA representation, allowing for longer retrospective analyses of past communities and potentially offering a more comprehensive perspective on marine biodiversity over time (Suter et al., 2023). This is of particular importance when water-based methods are insufficient and when historical occupancy data is desired.

Abiotic factors, such as pH and temperature, as well as biotic factors, including microbial activity, contribute to eDNA degradation rates (Mächler et al., 2018), with studies noting faster decay in acidic and warm conditions (Collins et al., 2018). Collins et al. (2018) identified that microbes play a vital role in eDNA persistence due to high degradation rates caused by enzymes produced by heterotrophic microbes (extracellular DNases). Research also shows that eDNA decay rates are not uniform and are influenced by the DNA fragment length, where longer fragments decaying more quickly (Jo et al., 2017). This suggests that fragment length could potentially provide information about the timescale of eDNA's environmental presence. The eDNA degradation rates can vary due to seasons as organic phosphorous or carbon concentrations fluctuates (Collins et al., 2018). As a summary, employing different fragment lengths and water depth can reveal deeper insights into the community dynamics with respect to timescale representation.

Essential for the effectiveness of eDNA in biomonitoring, is the need to accurately represent the current condition of an ecosystem. The persistence rates of eDNA (days to weeks) generally allows for the effective mapping and monitoring of local fish community structures (Yamamoto et al., 2017). Despite the local spatial scale representation of eDNA, it can be affected by eDNA transport from other locations additional to degradation rates (Andruszkiewicz et al., 2019). Multiple studies have noted a fish community spatial autocorrelation of 1 – 2 km (Harrison et al., 2019; Yamamoto et al., 2017; Zhang et al., 2020) supporting the local scale eDNA inference. However, these studies are a representation of lentic systems and low latitude marine systems where eDNA transport is low due to low water exchange (lentic) and eDNA degradation is high due to high light activity (low latitude; Salter, 2018). Although a couple of studies suggest that light intensity has negligible effect on detection probabilities (Mächler et

al., 2018), it is important to recognize that detection probability alone may not adequately reflect the influence of ultraviolet (UV) on eDNA degradation. Conversely, other research highlights that UV light impacts eDNA degradation (Strickler et al., 2015).

In alternative marine conditions represented by high water exchange, eDNA was found to spread about 10 km in seven days, and its source can be less than 20 km under moderate decay scenarios (Andruszkiewicz et al., 2019). Additionally, in low light activity such as the case for northern Norwegian fjords representing low decay ratios (Salter, 2018), a spatial heterogeneity was achieved at distances 30 km or more (**Paper I**). Moreover, this heterogeneity is achieved differently throughout the water column, with surface water reaching faster heterogeneity than bottom water (Harrison et al., 2019), supporting a longer timescale representation in deeper areas (**Paper I**). The ability to accurately map and monitor local fish community structures using eDNA reduces the need for extensive and time-consuming field surveys or invasive sampling techniques. Instead, taking these findings into account, researchers and resource managers can strategically select sampling locations based on the knowledge of eDNA persistence rates and timescale representation for monitoring communities in marine ecosystems.

## 4.2 How to treat eDNA data for efficient biomonitoring

The detectability of the eDNA depends on a variety of factors (Figure 6) including species density, DNA excretion, dilution, and degradation rates, among others. This multitude of factors are significant limitations of metabarcoding methods is their ability to accurately estimate the biomass (the total mass of organisms in a particular ecosystem) or abundance (the number of individuals per species within an ecosystem) of the detected species. Additional to the biological issues mentioned, technical issues, such as PCR bias, eDNA sampling efficiency, and DNA extraction complicate the matter further leaving metabarcoding studies often to rely on non-quantitative (qualitative) data (i.e., presence/absence), rather than providing a detailed picture of the relative or absolute abundance of each species present. This means that while metabarcoding can provide an extensive list of the species present in an environmental sample, it cannot reliably indicate how many individuals of each species are present or the total biomass they represent within that ecosystem. This limitation can sometimes lead to a lack of information or misinterpretations when metabarcoding data is used for biomonitoring (**Paper I**).

Despite such limitations, different eDNA metabarcoding data treatment approaches such as rank abundances (Wangensteen et al., 2018), log of relative abundances (Stoeckle et al., 2021), and eDNA index (Kelly et al., 2019) have revealed reliable estimates of species abundances in a semiquantitative manner. As eDNA metabarcoding can offer some sort of information regarding species quantities (**Paper I**), transforming the amplicon (sequence) read abundances into some semiquantitative indices can yield better community segregation between ecosystems (Figure 10) compared to qualitative approaches (**Paper I**). This semiquantitative superiority in differentiation ecosystems is a combination of multiple factors. Initially, the qualitative approach assigns equal weight to all species, regardless of their abundance. This, coupled with potential biases introduced during the metabarcoding process (latent process 4 in Figure 6), can cast doubt on the accuracy of the present taxa, particularly those that may be artifacts of PCR errors. Complicating matters further, the transport and dilution of DNA (latent process 2 in Figure 6), can indicate spurious eDNA traces (Thomsen et al., 2012) in the collected samples, thereby masking the relationship between the origin of DNA and the sample collection site. Such limitations from qualitative data treatment approaches are mitigated to an extent by semiquantitative methods due to the typically low read abundances associated with these erroneous DNA sequences raised from the bias processes above.

The primary factor driving the disparities in ecological approaches is the ubiquitousness of the species in the northern Norwegian fjords. These fjords, characterized by their high-latitude and cold-water environments, typical host low biodiversity but high biomass (Tittensor et al., 2010). This paradigm suggests that any observable variations in the ecological communities with these water bodies are more prone to biomass change rather than compositional change. To assess and monitor these water bodies, a semi-quantitative approach when using eDNA would serve as the most effective course of action (**Paper I**). This approach allows research to approximate the abundance and distribution of organisms, providing a more robust insights into the accurate health of ecosystems (especially north latitude ones). Consequently, semiquantitative data treatment approach of eDNA results offers a more efficient way for marine biomonitoring as significantly fewer samples are required to distinguish the resolution desired for drawing meaningful conclusions compared to qualitative approaches (Figure 10). These findings promote the use of eDNA surveys as a non-invasive, faster, and more cost-effective sampling method that can enhance management decisions.

### 4.3 Single-species quantitative biomonitoring

Beyond species identification through metabarcoding, quantitative eDNA techniques such as qPCR and ddPCR play a critical role in detecting and tracking rare, exotic, and invasive species. These techniques inform conservation efforts and ecological management by providing precise data for rapid species detection and marine monitoring (Rees et al., 2014). Quantitative eDNA efficacy in the early discovery of invasive species (Jerde et al., 2011), in the application of occupancy modeling (Keller et al., 2022), and in mapping out the spatial distribution of specific species (Salter et al., 2019; Shelton et al., 2022) has been well documented. One notable example is the study by Ficetola et al. (2008), which employed qPCR methods to detect the spread of the American bullfrog in France - a species native to America. In another example, quantitative eDNA detections indicated that two invasive carp species (silver carp and bighead carp) had surpassed the electric barriers of Lake Michigan eight months prior to traditional observation (Jerde et al., 2011) and had already started to establish a population. The rapid detection and assessment capabilities through quantitative eDNA techniques not only can streamline decision-making in environmental stewardship but also can reduce overall costs of sampling and monitoring efforts.

An additional aspect to the detection/non-detection feature of quantitative eDNA methods is the quantification of eDNA traces to infer organism abundance and biomass (Salter et al., 2019). However, factors like DNA shedding, transport, and degradation (Figure 5) in marine environments can affect eDNA concentrations, potentially complicating the relationship between DNA quantities and fish density and biomass (Yates et al., 2019). Nonetheless, research has demonstrated a direct and positive correlation between DNA concentration and the abundance or the total biomass present (Salter et al., 2019). Reinforcing this, Shelton et al. (2022) found that eDNA concentrations reflect fish densities observed in other quantitative assessments, such as acoustic surveys conducted for natural resource management.

Quantitative eDNA methods illustrate a greater sensitivity of eDNA to early detect invasive species when compared with traditional fisheries surveillance tools, such as netting or electrofishing for aquatic species (Jerde et al., 2011). While both qPCR and ddPCR (quantitative methods) generate consistent results in quantifying eDNA (**Paper II**), their physical approach to quantification is substantially different (Box 2 and 3). The ddPCR quantitation mechanism is less prone to PCR-inhibition due to their reaction happening independent in each droplet and furthermore due to endpoint detection mechanism. Higher

sensitivity and precision for eDNA quantification revealed by ddPCR indicated that newer technologies are more suited for eDNA detection and quantification (**Paper II**) with respect to invasive species, species distribution and abundances. More importantly, the ddPCR outperformance of qPCR was significantly higher at low concentration (**Paper II**) making the former method a valuable tool for rapid and accurate environmental monitoring when using eDNA sampling methods. Moreover, ddPCR method does not rely on the creation of standard curves, typically seen in qPCR, where known DNA concentrations from the targeted species are required (**Paper II**). Such non-reliance towards standard curves greatly simplifies the eDNA quantification process and broadens the applicability of ddPCR for marine monitoring scenarios, where obtaining accurate DNA concentrations might be crucial (**Paper II**). Despite the enhanced sensitivity and precision of ddPCR, both quantification methods are limited to the reaction volume when conducting PCR (typically 20  $\mu\text{L}$ ) as having 1 DNA copy per reaction would allow quantification of up to 0.05 copies/ $\mu\text{L}$  as a lower boundary (**Paper II**). DNA concentration below these levels would be hard to quantify due to randomness of the DNA being present in the reaction thus yielding stochastic results (**Paper II**). However, as the ddPCR mechanism relies on independent nano-droplets reactions (Box 3), thus increasing the number PCR reactions (technical replicates) would directly increase the number of droplets, hence lower the absolute boundary of quantification dependent on the number of technical replicates. Such advantages in the ddPCR technology, differently from qPCR, can catalyze faster and more informed decision-making processes in environmental management strategies as typical eDNA samples lies around the quantification limit (0.05 copies/ $\mu\text{L}$ ). The advanced technique of ddPCR simplifies the quantification process, as it doesn't heavily rely on standard curves, and can lower the limit of detection through increased technical replicates, producing highly accurate biotic data that can enhance the robustness in decision-making on conservation actions and biomonitoring.

#### **4.4 Multi-species quantitative eDNA biomonitoring**

Although a growing interest has been laid on multi-species quantification through the use of eDNA metabarcoding, very little substantial advancement has been recorded. The multi-facet processes (Figure 6) can skew the accurate translation of eDNA metabarcoding results into species abundance thus making the former data un-interpretable for thorough biodiversity estimates which hinder further detailed analysis of conservation biology. However, research efforts are concentrating on uncovering the processes of DNA dynamics, such as shedding rates

(Andruszkiewicz Allan et al., 2021), DNA transport (Andruszkiewicz et al., 2019), DNA degradation (Strickler et al., 2015), sampling efficiency (Bessey et al., 2020), and extraction efficiency (Cantera et al., 2019) among others in order to understand the intricate relationship between eDNA and biology.

Empirically estimating all the processes involved in the eDNA dynamic, can allow scientists to construct a more cohesive understanding of how eDNA relates to the underlying biological reality, thus enabling the extrapolation of metabarcoding data into quantitative biology. Such understanding can increase the applicability of non-invasive sampling such as eDNA into biodiversity and marine management. Yet, individual studies of these processes conducted in isolated parameters and settings or in mesocosms may not fully capture their complex interactions in the wild (Yates et al., 2019). Studying the processes in conjunction to one another builds an accurate understanding on how these processes interplay hence yielding more accurate results for linking the eDNA metabarcoding observational data to species abundances (**Paper III**) or biomass. This approach (integrated multi-facet processes; Figure 9) allows for a more authentic capture of how DNA behavior and dynamics in natural environments affects metabarcoding data interpretation.

Consistent with earlier research, it is recognized that both the dilution and transport of DNA through the environment and the species-specific DNA shedding rates are pivotal factors to consider when applying metabarcoding analysis for quantitative purposes (**Paper III**). When external factors that influence DNA dynamics, such as abiotic conditions that affect degradation, and technical processes that involve the isolation and amplification of DNA (Figure 6), are either controlled or assessed, the remaining biological processes can be encapsulated into a single, species-specific parameter (denoted as  $\theta$  in **Paper III**) and could thereafter be empirically estimated with the help of traditional surveys (i.e., trawling in **Paper III**). This parameter could act as a crucial conversion factor enabling the interpretation of eDNA metabarcoding data as a reliable reflection of multi-species abundance, thus bridging the gaps of eDNA applicability biomonitoring and marine management and conservation. Additionally, having empirical estimation of the biological processes, researchers can furthermore easily estimate other processes such as those related to environmental and technical processes.

The workflow presented in **Paper III** could be adjusted to incorporate sampling observation from different traditional monitoring techniques such as acoustic, visual, mark-recapture surveys, and fishers' landings among others, enhancing our understanding of the biological

processes that shape eDNA signals observed through metabarcoding. This progressive approach of multiple assessment methods could lead to better integration of eDNA analysis into EBM or stock assessments. From a management perspective, traditional surveys are valuable but face challenges including high costs, restricted areas, and dependability on weather conditions. Integrating eDNA surveys (given their known species-specific conversion factors) with established methods in **Paper III** offers a promising solution to enhance the frequency and efficiency of biotic data collection for marine ecosystem management.

#### **4.5 Additional features of eDNA adaptable to marine management**

The urgency for new insights into marine ecology and the health of ecosystems is heightened by external pressures including climate change, human influence, and limited survey durations. Traditional surveys often struggle to encompass the vast diversity of species across phyla, but alternatively eDNA metabarcoding offers a cost-effective, non-invasive option for assessing multi-species distribution (**Paper I**) and abundance (**Paper III**). The shift of paradigm towards EBM mirrors this need, integrating considerations of habitat, predator-prey dynamics, ocean conditions and stakeholder engagement for effective management (Cucuzza et al., 2021). The addition of eDNA alongside the other ecological survey data promises to create comprehensive, multi-faceted views of ecosystems (**Paper III**).

The strength of eDNA lies in its ability to conduct extensive analyses without the need for direct sampling of organisms and additionally by using multiple genetic markers it can reveal a holistic overview (i.e., multiple kingdoms) of the biotic abundance and distribution in small and large ecosystems. Taking into account that recent advancements have shown that eDNA from organisms - known as intraorganismal eDNA - can reflect the abundance of microorganisms (Bourque et al., 2023), and coupling it with trace eDNA quantification as outlined in **Paper III**, can effectively bridge the gap between eDNA research and quantitative biology. These discoveries allow for description and quantification of whole communities such as bacteria, and viruses as intraorganismal eDNA, and eukaryotes as trace eDNA. Bridging such gap can enhance our understanding and mapping the whole ecosystem and its dynamics (Barnes & Turner, 2016). By quantifying the ecosystem as a whole together with its dynamics, researchers and policymakers can gain a more profound comprehension of anthropogenic impact, subsequently leading to the creation of more effective strategies to reduce and manage



these impacts on ecosystems (Mieszkowska et al., 2014). Subsequently, changes in the community composition over time (typical impact of climate change or anthropogenic impact) can be studied using eDNA metabarcoding (**Paper I**) which can indicate shifts in ecosystem health or anthropogenic stressor.

Understanding and accurately measuring the interactions within food webs is also critical for resource managers and policy makers to grasp the complexities of ecosystems. Changes in predator-prey relationships due to changes of population abundance are typical responses to environmental or anthropogenic pressures (Polis & Strong, 1996). As the use of eDNA surpasses conventional gut-content surveys in efficiency, it can provide richer insights into the health of ecosystems and the interactions within food webs (Novotny et al., 2023). Although eDNA gut-content analysis may be intrusive (non-invasive), it requires smaller sample sizes than conventional methods (Pompanon et al., 2012). Combining eDNA analysis from dietary studies in conjunction with eDNA collected from environment can allow researchers to assess species and community plasticity thus indicating ecosystem resilience (Novotny et al., 2023). This integrated approach can enhance our understanding on how marine biotic communities function which can provide a tool for empirical assessments of the factors that influence the dynamics of the ecosystems, thus enabling resource managers to take informed decisions about the conservation efforts (Voelker et al., 2020).

Spawning grounds serve as vital habitats for the reproductive processes of aquatic species, playing a key role in sustaining fish populations and preserving genetic variation (Erisman et al., 2017). Consequently, resource managers often implement spawning closures to ensure that reproductive processes are well maintained (Van Overzee & Rijnsdorp, 2015). The effective management relies heavily on animal behavior and their lifecycle aspects (Schadewell & Adams, 2021), thus the identification and protection of these spawning sites is crucial (Van Overzee & Rijnsdorp, 2015). However, the identification of the spawning grounds is challenging, particularly since they are often situated in remote or hard-to-access locations where conventional survey methods often struggle to access. Furthermore, the intrusive nature of traditional methods often disrupt the species they aim to protect (Tsuji & Shibata, 2021). With the recent advancements, studies have showed that eDNA metabarcoding can promise a less intrusive and potentially more effective method for identification of spawning locations. The applicability of eDNA in such scenarios is threefold as it can (i) identify important lifecycles of diverse species (Schadewell & Adams, 2021), (ii) enable rapid responses from

managers in case of environmental stressors (Di Muri et al., 2023), and (iii) assess the effectiveness of their conservation actions (Mieszkowska et al., 2014).

Another important aspect of conservation biology is the assessment of functional diversity in ecosystem (Cadotte et al., 2011). Functional diversity offers a more comprehensive understanding of biodiversity, aiming at crafting more impactful conservation tactics, not just to safeguard individual species but also their ecological roles (Petchey & Gaston, 2006). An ecosystem with a high level of functional diversity is often more resilient to disturbances such as disease, climate change, and human impact (Cadotte et al., 2011; Petchey & Gaston, 2006). Additionally, high functional diversity contributes to ecosystem productivity, which includes the growth of biomass and the efficient cycling of nutrients (Condachou et al., 2023). The use of eDNA metabarcoding in the field of functional diversity has shown promising results as the detection and quantification of a wide array of species is now easier due to increased effort in completion of reference databases (Condachou et al., 2023). Due to eDNA being easily adaptable to a large variety of species (also after the collection of environmental samples has been conducted), by changing the target gene region, thus species (i.e., mammals, vertebrate, eukaryotes, and bacteria among others), it can facilitate the workflow for assessing the breadth of functions that organisms in a community contribute to the ecosystem and understanding how this biodiversity supports ecosystem health and resilience (Barnes & Turner, 2016). Moreover, the use of highly sensitive eDNA methods (**Paper II**), can accurately detect species at low abundances. This is crucial for recognition of potential hazards, including invasive species, emerging diseases, or the systematic quantification of keystone species (Suter et al., 2023) - all of which have the potential to alter ecosystem functionality - thus allowing resource managers and stakeholders to take action on mitigating such impacts in ecosystem with low plasticity.

With respect to fisheries management, a crucial facet is the size, age, and sex identification of species. These metrics provide insights into the maturity levels, reproductive capacities, and overall health and sustainability of the fish populations (Cadrin et al., 2014). While the field of eDNA has experienced significant technological breakthroughs, it still faces limitations when it comes to providing the detailed level of information required (i.e., age, size structure and sex of individuals) from metabarcoding data alone. However, workarounds can be implemented by using the workflow described in **Paper III** coupled with traditional monitoring that are size or age selective (sampling selectivity;  $q$ ). By knowing the sampling selectivity eDNA quantitative signals can be translated to infer size distribution of fish species in marine ecosystems. In addition, there is an emerging wave of research focused on the development of specific genetic

markers or SNP-s (single nucleotide polymorphisms), which could eventually allow for the determination of a population sex ratio from eDNA samples (Nichols & Spong, 2017). Such innovations would mark a significant leap forward in the non-invasive monitoring for stock assessment and marine conservation, providing a more holistic view of the biological resources.

The rise of automated eDNA sampling has also garnered attention, especially with robotic technology now aiding standardized sample collection at regular intervals. This reduces sampling handling time, which is particularly advantageous for repeated long-term biomonitoring. Autonomous vehicles could further streamline data gathering, facilitating robust measurement of species abundances and ecosystem dynamics. While currently, high acquisition and operational costs may limit usage, further advancements in this field could revolutionize the data gathering and thereafter marine management.

## **5 Future consideration of eDNA for marine monitoring**

While the field of eDNA continues to advance and show promise for future applications, there are still technical hurdles that need to be addressed in the future. The effectiveness of eDNA in marine management depends highly on the quality of the reference database used to identify species. Although many authors confirm that steady progress is made on updating the reference database (Ruppert et al., 2019), the lack of specificity can be a problem for the differentiation of sub-species or different populations (Chain et al., 2016) which can infer challenges on effective management. Further refinement on the reference database would foster the applicability of eDNA into biomonitoring.

Moreover, eDNA metabarcoding surveys have been reporting absence of species known to be present, also known as false-negatives (Hansen et al., 2018; Kelly et al., 2014). For instance, Kelly et al. (2014) were not able to detect sharks or turtles eDNA from sampling the Monterey Bay Aquarium mesocosm despite their confirmed presence. Although such technical challenges can be reverted by altering the genetic markers, the issue of a persistently low false-negative detection rate, particularly among rare or hard-to-find species, remains a challenge (Kelly et al., 2014). Consequently, the development of primers targeting such species can facilitate the molecular tools to be used by resource managers for effective ecosystem management.

Since the eDNA field is relatively new with many unknown factors, scholars tend to rely on controlled experiments settings for understanding the factors and parameters involved in eDNA dynamics (Harrison et al., 2019). Although mesocosm experiments seems as a straightforward method forward (Harrison et al., 2019), replicating the real biological complexity and its interactions observed in natural ecosystems within mesocosms (which suffer from limited realism) is challenging (Stewart et al., 2013). Consequently, the empirical findings derived from such controlled experiments might not accurately reflect the realities of natural environments, where the complexity is magnified (Stewart et al., 2013). In natural settings, the interactions and fluctuations are diverse, and the direct extrapolation of mesocosm-derived insights could lead to inaccuracies. Therefore, it is crucial to recognize the inherent disparities between mesocosms and wild ecosystems when applying laboratory findings to field studies. To circumvent such challenges, studies should measure such parameters (when possible) in the natural ecosystems or integrated into other parameters as conducted in **Paper III**.

Despite the comprehensive exploration of various latent processes (Figure 6), the sampling processes (latent process 3 in Figure 6) remained unchanged in this study (see material and methods) thus their effects on eDNA dynamics remains unknown. To enhance our understanding of this aspect, future studies should incorporate different sampling protocols (latent process 3 in Figure 6) to assess their influence on eDNA collection and analysis alongside the other latent processes (as conducted in **Paper III**). By doing so, valuable insights can be gained to optimize environmental monitoring and biodiversity assessments for increased precision and accuracy (**Paper III**).

Lastly, this thesis has not accounted for microbial activity or seasonality into the biological or environmental processes; thus, it is important to exercise caution when interpreting the findings, as multiple studies have suggested that microbial activity and difference environmental conditions during seasons are in pair with eDNA degradation and persistence (Collins et al., 2018; Mächler et al., 2018). Further investigation of the influence of microbial activity and seasonality, jointly with other latent processes (Figure 6) is necessary to gain a better understanding of this topic and provide more insights for effective marine management.

## 6 Conclusion

Environmental DNA (eDNA) is a promising tool for monitoring marine biodiversity and identifying the presence of organisms and their community structures. The persistence of eDNA varies greatly in aquatic environments, depending on biological, environmental, and technical processes (Figure 6). The persistence of eDNA spans from days up to weeks within marine environment with longer eDNA fragments degrading faster than the short ones. Also, depending on water exchange rate and sunlight regime, eDNA can spread up to several kilometers. An in-depth understanding of processes involved in molecular ecology, such as eDNA shedding, transport, degradation rates, sampling, and isolation efficiencies, proves crucial for scholars adopting molecular tools into marine management. Accurate integration of these processes which also consider seasonal variations, water chemistry changes, and biological factors will enable eDNA analysis to provide a sensitive and informative measure of marine resources. Ultimately, appropriately leveraging the power of eDNA has the potential to significantly enhance our understanding of marine ecosystems, leading to more efficient and effective conservation and management protocols.

Metabarcoding studies often rely on qualitative data, such as species presence or absence, rather than providing detailed information on species abundance. However, this study (**Paper I**) shows that eDNA metabarcoding data holds some information of species abundances thus semiquantitative data treatment approaches can be developed to estimate approximate species abundances. These approaches are significantly more effective in scenarios on high latitude marine environments where species ubiquitousness is high, but their abundance or dominance varies. Additionally, factor such as DNA dilution, transport, and species-specific shedding rates are important considerations for accurate quantitative analysis. By controlling or assessing these factors, eDNA metabarcoding data can be interpreted as a reliable reflection of multi-species abundance (**Paper III**). Integrating eDNA analysis with traditional monitoring techniques or citizen science offers a promising solution for enhancing biotic data and improving marine ecosystem management.

Additionally, the use of quantitative environmental DNA (qPCR and ddPCR) techniques allows for the early detection and monitoring of rare, elusive, and invasive species, providing managers with the ability to implement rapid response measures. This is particularly valuable as controlling invasive species before they become widely established is easier and less costly, minimizing both ecological and economic impacts. Although trace eDNA (typically low eDNA

concentration) often creates problem for quantitative eDNA methods, ddPCR offers advantages by being more sensitive and precise biomonitoring method, aiding in more effective conservation and management efforts (**Paper II**).

The specificity and sensitivity of eDNA methodologies allow for detailed community profiling, including microorganisms, enhancing our ability to track ecosystem health and respond to ecological shifts. These advanced techniques offer valuable insights into the complexity of food web dynamics, essential for informed resource management and conservation efforts. Additionally, eDNA metabarcoding presents a less intrusive method for locating spawning and other critical habitats. The ability of eDNA technology to detect a broad spectrum of species and thus estimating the functional diversity of ecosystem is fundamental to maintaining ecosystem resilience and productivity. Moreover, eDNA is particularly effective in identifying potential threats such as invasive species, diseases, and the influence of keystone species on ecosystem function. Although eDNA offers numerous advantages, it still has limitations in assessing individual characteristics such as size, age, and sex. Nevertheless, combining eDNA techniques with traditional monitoring and emerging genetic markers holds promise for overcoming these challenges, thereby contributing to more targeted and effective fisheries management and conservation strategies.

## 7 References

- Afzali, S. F., Bourdages, H., Laporte, M., Mérot, C., Normandeau, E., Audet, C., & Bernatchez, L. (2021). Comparing environmental metabarcoding and trawling survey of demersal fish communities in the Gulf of St. Lawrence, Canada. *Environmental DNA*, 3(1), 22–42. <https://doi.org/10.1002/edn3.111>
- Andres, K. J., Lambert, T. D., Lodge, D. M., Andrés, J., & Jackson, J. R. (2022). Combining sampling gear to optimally inventory species highlights the efficiency of eDNA metabarcoding. *Environmental DNA*, edn3.366. <https://doi.org/10.1002/edn3.366>
- Andruszkiewicz Allan, E., Zhang, W. G., C. Lavery, A., & F. Govindarajan, A. (2021). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *Environmental DNA*, 3(2), 492–514. <https://doi.org/10.1002/edn3.141>
- Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., & Boehm, A. B. (2019). Modeling Environmental DNA Transport in the Coastal Ocean Using Lagrangian Particle Tracking. *Frontiers in Marine Science*, 6, 477. <https://doi.org/10.3389/fmars.2019.00477>
- Atienza, S., Guardiola, M., Præbel, K., Antich, A., Turon, X., & Wangenstein, O. S. (2020). DNA metabarcoding of deep-sea sediment communities using COI: community assessment, spatio-temporal patterns and comparison with the 18S rDNA marker. *Diversity*, 12(4), 123. <https://doi.org/10.3390/d12040123>
- Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, 17(1), 1–17. <https://doi.org/10.1007/s10592-015-0775-4>
- Barros, M. J. G. de, Eduardo, L. N., Bertrand, A., Lucena-Frédou, F., Frédou, T., Lira, A. S., & Ferreira, B. P. (2021). Bottom trawling on a carbonate shelf: Do we get what we see? *Continental Shelf Research*, 213, 104314. <https://doi.org/10.1016/j.csr.2020.104314>
- Bavington, D. (2010). From Hunting Fish to Managing Populations: Fisheries Science and the Destruction of Newfoundland Cod Fisheries. *Science as Culture*, 19(4), 509–528. <https://doi.org/10.1080/09505431.2010.519615>
- Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., Power, M., McLaughlin, J., Edgar, G. J., & Keesing, J. (2020). Maximizing fish detection with eDNA metabarcoding. *Environmental DNA*, 2(4), 493–504. <https://doi.org/10.1002/edn3.74>
- Bevanger, Kjetil. (2018). Biodiversity in Norway. In Pullaiah, T. (Ed.), *Global Biodiversity: Selected Countries in Europe* (1st ed., pp. 223–260). Apple Academic Press.
- Biju Kumar, A., & Deepthi, G. R. (2006). Trawling and by-catch: Implications on marine ecosystem. *Current Science*, 90(7), 922–931.
- 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 and Evolution*, 29(6), 358–367. <https://doi.org/10.1016/j.tree.2014.04.003>
- Bourque, D. A., Wang, X. S., Fryxell, J. M., Betini, G. S., McCann, K. S., & Hanner, R. (2023). Environmental DNA tracks zooplankton population dynamics in experimental mesocosms. *Environmental DNA*, 5(5), 880–893. <https://doi.org/10.1002/edn3.347>
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1), 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Burian, A., Mauvisseau, Q., Bulling, M., Domisch, S., Qian, S., & Sweet, M. (2021). Improving the reliability of eDNA data interpretation. *Molecular Ecology Resources*, 21(5), 1422–1433. <https://doi.org/10.1111/1755-0998.13367>

- Cadotte, M. W., Carscadden, K., & Mirotnick, N. (2011). Beyond species: Functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology*, 48(5), 1079–1087. <https://doi.org/10.1111/j.1365-2664.2011.02048.x>
- Cadrin, S. X., Karr, L. A., & Mariani, S. (2014). Stock Identification Methods. In *Stock Identification Methods* (pp. 1–5). Elsevier. <https://doi.org/10.1016/B978-0-12-397003-9.00001-1>
- Canals, O., Mendibil, I., Santos, M., Irigoien, X., & Rodríguez-Ezpeleta, N. (2021). Vertical stratification of environmental DNA in the open ocean captures ecological patterns and behavior of deep-sea fishes. *Limnology And Oceanography Letters*, 1–9. <https://doi.org/10.1002/lol2.10213>
- Cantera, I., Cilleros, K., Valentini, A., Cerdan, A., Dejean, T., Iribar, A., Taberlet, P., Vigouroux, R., & Brosse, S. (2019). Optimizing environmental DNA sampling effort for fish inventories in tropical streams and rivers. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-39399-5>
- Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., MacE, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486(7401), 59–67. <https://doi.org/10.1038/nature11148>
- Carr, H., Abas, M., Boutahar, L., Caretti, O. N., Chan, W. Y., Chapman, A. S. A., De Mendonça, S. N., Engleman, A., Ferrario, F., Simmons, K. R., Verdura, J., & Zivian, A. (2020). The Aichi Biodiversity Targets: Achievements for marine conservation and priorities beyond 2020. *PeerJ*, 8, 1–35. <https://doi.org/10.7717/peerj.9743>
- Chain, F. J. J., Brown, E. A., MacIsaac, H. J., & Cristescu, M. E. (2016). Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. *Diversity and Distributions*, 22(5), 493–504. <https://doi.org/10.1111/ddi.12427>
- Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., Iribar, A., Taberlet, P., Vigouroux, R., & Brosse, S. (2019). Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. *Molecular Ecology Resources*, 19(1), 27–46. <https://doi.org/10.1111/1755-0998.12900>
- Claudet, J., Loiseau, C., Sostres, M., & Zupan, M. (2020). Underprotected Marine Protected Areas in a Global Biodiversity Hotspot. *One Earth*, 2(4), 380–384. <https://doi.org/10.1016/j.oneear.2020.03.008>
- Collins, R. A., Wangensteen, O. S., O’Gorman, E. J., Mariani, S., Sims, D. W., & Genner, M. J. (2018). Persistence of environmental DNA in marine systems. *Communications Biology*, 1(1), 185. <https://doi.org/10.1038/s42003-018-0192-6>
- Condachou, C., Milhau, T., Murienne, J., Brosse, S., Villéger, S., Valentini, A., Dejean, T., & Mouillot, D. (2023). Inferring functional diversity from environmental DNA metabarcoding. *Environmental DNA*, 5(5), 934–944. <https://doi.org/10.1002/edn3.391>
- Cormier, R., Kelble, C. R., Anderson, M. R., Allen, J. I., Grehan, A., & Gregersen, Ó. (2017). Moving from ecosystem-based policy objectives to operational implementation of ecosystem-based management measures. *ICES Journal of Marine Science*, 74(1), 406–413. <https://doi.org/10.1093/icesjms/fsw181>
- Crossin, G. T., Heupel, M. R., Holbrook, C. M., Hussey, N. E., Lowerre-Barbieri, S. K., Nguyen, V. M., Raby, G. D., & Cooke, S. J. (2017). Acoustic telemetry and fisheries management. *Ecological Applications*, 27(4), 1031–1049. <https://doi.org/10.1002/eap.1533>
- Cucuzza, M., Stoll, J. S., & Leslie, H. M. (2021). Evaluating the theoretical and practical linkages between ecosystem-based fisheries management and fisheries co-management. *Marine Policy*, 126, 104390. <https://doi.org/10.1016/j.marpol.2020.104390>



- Di Muri, C., Lawson Handley, L., Bean, C. W., Benucci, M., Harper, L. R., James, B., Li, J., Winfield, I. J., & Hänfling, B. (2023). Spatio-temporal monitoring of lake fish spawning activity using environmental DNA metabarcoding. *Environmental DNA*, 5(5), 849–860. <https://doi.org/10.1002/edn3.343>
- Eigaard, O. R., Bastardie, F., Hintzen, N. T., Buhl-Mortensen, L., Buhl-Mortensen, P., Catarino, R., Dinesen, G. E., Egekvist, J., Fock, H. O., Geitner, K., Gerritsen, H. D., González, M. M., Jonsson, P., Kavadas, S., Laffargue, P., Lundy, M., Gonzalez-Mirelis, G., Nielsen, J. R., Papadopoulou, N., ... Rijnsdorp, A. D. (2017). The footprint of bottom trawling in European waters: Distribution, intensity, and seabed integrity. *ICES Journal of Marine Science*, 74(3), 847–865. <https://doi.org/10.1093/icesjms/fsw194>
- Erisman, B., Heyman, W., Kobara, S., Ezer, T., Pittman, S., Aburto-Oropeza, O., & Nemeth, R. S. (2017). Fish spawning aggregations: Where well-placed management actions can yield big benefits for fisheries and conservation. *Fish and Fisheries*, 18(1), 128–144. <https://doi.org/10.1111/faf.12132>
- Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4(4), 423–425. <https://doi.org/10.1098/rsbl.2008.0118>
- Fogarty, M. J. (2014). The art of ecosystem-based fishery management. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(3), 479–490. <https://doi.org/10.1139/cjfas-2013-0203>
- Fraija-Fernández, N., Bouquieaux, M. C., Rey, A., Mendibil, I., Cotano, U., Irigoien, X., Santos, M., & Rodríguez-Ezpeleta, N. (2020). Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution*, 10(14), 7560–7584. <https://doi.org/10.1002/ece3.6482>
- Fraser, H. M., Greenstreet, S. P. R., & Piet, G. J. (2007). Taking account of catchability in groundfish survey trawls: Implications for estimating demersal fish biomass. *ICES Journal of Marine Science*, 64(9), 1800–1819. <https://doi.org/10.1093/icesjms/fsm145>
- Freiwald, A., Hühnerbach, V., Lindberg, B., Wilson, J. B., & Campbell, J. (2002). The Sula Reef Complex, Norwegian shelf. *Facies*, 47(1), 179–200. <https://doi.org/10.1007/BF02667712>
- Gilbey, J., Carvalho, G., Castilho, R., Coscia, I., Coulson, M. W., Dahle, G., Derycke, S., Francisco, S. M., Helyar, S. J., Johansen, T., Junge, C., Layton, K. K. S., Martinsohn, J., Matejusova, I., Robalo, J. I., Rodríguez-Ezpeleta, N., Silva, G., Strammer, I., Vasemägi, A., & Volckaert, F. A. M. (2021). Life in a drop: Sampling environmental DNA for marine fishery management and ecosystem monitoring. *Marine Policy*, 124, 104331. <https://doi.org/10.1016/j.marpol.2020.104331>
- Gold, Z., Shelton, A. O., Casendino, H. R., Duprey, J., Gallego, R., Van Cise, A., Fisher, M., Jensen, A. J., D'Agnesse, E., Andruszkiewicz Allan, E., Ramón-Laca, A., Garber-Yonts, M., Labare, M., Parsons, K. M., & Kelly, R. P. (2023). Signal and noise in metabarcoding data. *PLOS ONE*, 18(5), e0285674. <https://doi.org/10.1371/journal.pone.0285674>
- Green, E. J., Buchanan, G. M., Butchart, S. H. M., Chandler, G. M., Burgess, N. D., Hill, S. L. L., & Gregory, R. D. (2020). Relating characteristics of global biodiversity targets to reported progress. *Conservation Biology*, 33(6), 1360–1369. <https://doi.org/10.1111/cobi.13322>
- Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., Lacoursière-Roussel, A., Leong, S. C. Y., Li, Y., Olds, B., Pfrender, M. E., Prowse, T. A. A., Renshaw, M. A., & Lodge, D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports*, 8, 8843. <https://doi.org/10.1038/s41598-018-27048-2>
- Gullestad, P., Abotnes, A. M., Bakke, G., Skern-Mauritzen, M., Nedreaas, K., & Søvik, G. (2017). Towards ecosystem-based fisheries management in Norway – Practical tools for keeping track

- of relevant issues and prioritising management efforts. *Marine Policy*, 77, 104–110. <https://doi.org/10.1016/j.marpol.2016.11.032>
- Günther, B., Kneibelsberger, T., Neumann, H., Laakmann, S., & Martínez Arbizu, P. (2018). Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes. *Scientific Reports*, 8(1), 1–13. <https://doi.org/10.1038/s41598-018-32917-x>
- Guri, G., Westgaard, J., Yoccoz, N., Wangensteen, O. S., Præbel, K., Ray, J. L., Kelly, R. P., Shelton, A. O., Hanebrekke, T., & Johansen, T. (2023). Maximizing sampling efficiency to detect differences in fish community composition using environmental DNA metabarcoding in subarctic fjords. *Environmental DNA*, edn3.409. <https://doi.org/10.1002/edn3.409>
- Hansen, B. K., Bekkevold, D., Clausen, L. W., & Nielsen, E. E. (2018). The sceptical optimist: Challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, 19(5), 751–768. <https://doi.org/10.1111/faf.12286>
- Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, 286(1915). <https://doi.org/10.1098/rspb.2019.1409>
- Jakobsen, T. (1987). Coastal cod in Northern Norway. *Fisheries Research*, 5(2–3), 223–234. [https://doi.org/10.1016/0165-7836\(87\)90042-7](https://doi.org/10.1016/0165-7836(87)90042-7)
- Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). “Sight-unseen” detection of rare aquatic species using environmental DNA: eDNA surveillance of rare aquatic species. *Conservation Letters*, 4(2), 150–157. <https://doi.org/10.1111/j.1755-263X.2010.00158.x>
- Jeunen, G. J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., Bunce, M., & Gemmell, N. J. (2019). Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Molecular Ecology Resources*, 19(2), 426–438. <https://doi.org/10.1111/1755-0998.12982>
- Jo, T., Murakami, H., Masuda, R., Sakata, M. K., Yamamoto, S., & Minamoto, T. (2017). Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. *Molecular Ecology Resources*, 17(6), e25–e33. <https://doi.org/10.1111/1755-0998.12685>
- Keller, A. G., Grason, E. W., McDonald, P. S., Ramón-Laca, A., & Kelly, R. P. (2022). Tracking an invasion front with environmental DNA. *Ecological Applications*, 32(4), e2561. <https://doi.org/10.1002/eap.2561>
- Kelly, R. P., Port, J. A., Yamahara, K. M., & Crowder, L. B. (2014). Using Environmental DNA to Census Marine Fishes in a Large Mesocosm. *PLoS ONE*, 9(1), e86175. <https://doi.org/10.1371/journal.pone.0086175>
- Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. *Scientific Reports*, 9(1), 1–14. <https://doi.org/10.1038/s41598-019-48546-x>
- Kirkfeldt, T. S., & Frazão Santos, C. (2021). A Review of Sustainability Concepts in Marine Spatial Planning and the Potential to Supporting the UN Sustainable Development Goal 14. *Frontiers in Marine Science*, 8, 713980. <https://doi.org/10.3389/fmars.2021.713980>
- Knudsen, S. W., Ebert, R. B., Hesselsøe, M., Kuntke, F., Hassingboe, J., Mortensen, P. B., Thomsen, P. F., Sigsgaard, E. E., Hansen, B. K., Nielsen, E. E., & Møller, P. R. (2019). Species-specific detection and quantification of environmental DNA from marine fishes in the Baltic Sea. *Journal of Experimental Marine Biology and Ecology*, 510(May 2018), 31–45. <https://doi.org/10.1016/j.jembe.2018.09.004>
- Kvile, K. Ø., Andersen, G. S., Baden, S. P., Bekkby, T., Bruhn, A., Geertz-Hansen, O., Hancke, K., Hansen, J. L. S., Krause-Jensen, D., Rinde, E., Steen, H., Wegeberg, S., & Gundersen, H.

- (2022). Kelp Forest Distribution in the Nordic Region. *Frontiers in Marine Science*, 9, 850359. <https://doi.org/10.3389/fmars.2022.850359>
- Lacoursière-Roussel, A., Rosabal, M., & Bernatchez, L. (2016). Estimating fish abundance and biomass from eDNA concentrations: Variability among capture methods and environmental conditions. *Molecular Ecology Resources*, 16(6), 1401–1414. <https://doi.org/10.1111/1755-0998.12522>
- Lamb, P. D., Fonseca, V. G., Maxwell, D. L., & Nnanatu, C. C. (2022). Systematic review and meta-analysis: Water type and temperature affect environmental DNA decay. *Molecular Ecology Resources*, 22(7), 2494–2505. <https://doi.org/10.1111/1755-0998.13627>
- Larson, W. A., Barry, P., Dokai, W., Maselko, J., Olson, J., & Baetscher, D. (2022). Leveraging eDNA metabarcoding to characterize nearshore fish communities in Southeast Alaska: Do habitat and tide matter? *Environmental DNA*, November 2021, 1–13. <https://doi.org/10.1002/edn3.297>
- Lees, K. J., MacNeil, M. A., Hedges, K. J., & Hussey, N. E. (2021). Estimating demographic parameters for fisheries management using acoustic telemetry. *Reviews in Fish Biology and Fisheries*, 31(1), 25–51. <https://doi.org/10.1007/s11160-020-09626-8>
- Lotze, H. K. (2021). Marine biodiversity conservation. *Current Biology*, 31(19), R1190–R1195. <https://doi.org/10.1016/j.cub.2021.06.084>
- Mace, G. M., Barrett, M., Burgess, N. D., Cornell, S. E., Freeman, R., Grooten, M., & Purvis, A. (2020). Aiming higher to bend the curve of biodiversity loss. *Nature Sustainability*, 1(September 2018), 448–451. <https://doi.org/10.1038/s41893-018-0130-0>
- Mächler, E., Osathanukul, M., & Altermatt, F. (2018). Shedding light on eDNA: Neither natural levels of UV radiation nor the presence of a filter feeder affect eDNA-based detection of aquatic organisms. *PLOS ONE*, 13(4), e0195529. <https://doi.org/10.1371/journal.pone.0195529>
- Madin, E. M. P., Dill, L. M., Ridlon, A. D., Heithaus, M. R., & Warner, R. R. (2016). Human activities change marine ecosystems by altering predation risk. *Global Change Biology*, 22(1), 44–60. <https://doi.org/10.1111/gcb.13083>
- Magurran, A. E. (2013). Open questions: Some unresolved issues in biodiversity. *BMC Biology*, 11(1), 118. <https://doi.org/10.1186/1741-7007-11-118>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: Robust and fast clustering method for amplicon-based studies. *PeerJ*, 2014(1), 1–13. <https://doi.org/10.7717/peerj.593>
- Marco, M. D., Butchart, S. H. M., Visconti, P., Buchanan, G. M., Ficetola, G. F., & Rondinini, C. (2015). *Synergies and trade-offs in achieving global biodiversity targets*. 30(1), 189–195. <https://doi.org/10.1111/cobi.12559>
- Mauvisseau, Q., Harper, L., Sander, M., Hanner, R. H., Kleyer, H., & Deiner, K. (2021). The multiple states of environmental DNA and what is known about their persistence in aquatic environments. *Authorea Preprints*, preprint(preprint), 1–17. <https://doi.org/10.1021/acs.est.1c07638>
- McCann, K. S. (2000). The diversity–stability debate. *Nature*, 405(6783), 228–233. <https://doi.org/10.1038/35012234>
- Meier, J. I., McGee, M. D., Marques, D. A., Mwaiko, S., Kische, M., Wandera, S., Neumann, D., Mrosso, H., Chapman, L. J., Chapman, C. A., Kaufman, L., Taabu-Munyaho, A., Wagner, C. E., Bruggmann, R., Excoffier, L., & Seehausen, O. (2023). Cycles of fusion and fission enabled rapid parallel adaptive radiations in African cichlids. *Science*, 381(6665), eade2833. <https://doi.org/10.1126/science.ade2833>

- Merrett, N. R. (1994). Reproduction in the North Atlantic oceanic ichthyofauna and the relationship between fecundity and species' sizes. *Environmental Biology of Fishes*, 41(1–4), 207–245. <https://doi.org/10.1007/BF02197846>
- Mieszowska, N., Sugden, H., Firth, L. B., & Hawkins, S. J. (2014). The role of sustained observations in tracking impacts of environmental change on marine biodiversity and ecosystems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2025), 20130339. <https://doi.org/10.1098/rsta.2013.0339>
- Miloslavich, P., Bax, N. J., Simmons, S. E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S. D., Benedetti-Cecchi, L., Checkley, D. M., Chiba, S., Duffy, J. E., Dunn, D. C., Fischer, A., Gunn, J., Kudela, R., Marsac, F., Muller-Karger, F. E., Obura, D., & Shin, Y. J. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. *Global Change Biology*, 24(6), 2416–2433. <https://doi.org/10.1111/gcb.14108>
- 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*. <https://doi.org/10.1098/rsos.150088>
- Neebling, T. E., & Quist, M. C. (2011). Comparison of Boat Electrofishing, Trawling, and Seining for Sampling Fish Assemblages in Iowa's Nonwadeable Rivers. *North American Journal of Fisheries Management*, 31(2), 390–402. <https://doi.org/10.1080/02755947.2011.576198>
- Nguyen, L. N. (2023). *The Development of the Law of the Sea by UNCLOS Dispute Settlement Bodies* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108980296>
- Nichols, R., & Spong, G. (2017). An eDNA-Based SNP Assay for Ungulate Species and Sex Identification. *Diversity*, 9(3), 33. <https://doi.org/10.3390/d9030033>
- Novotny, A., Serandour, B., Kortsch, S., Gauzens, B., Jan, K. M. G., & Winder, M. (2023). DNA metabarcoding highlights cyanobacteria as the main source of primary production in a pelagic food web model. *Science Advances*, 9(17), eadg1096. <https://doi.org/10.1126/sciadv.adg1096>
- O'Hara, C. C., Frazier, M., & Halpern, B. S. (2021). At-risk marine biodiversity faces extensive, expanding, and intensifying human impacts. *Science*, 372(6537), 84–87. <https://doi.org/10.1126/science.abe6731>
- O'Leary, B. C., Winther-Janson, M., Bainbridge, J. M., Aitken, J., Hawkins, J. P., & Roberts, C. M. (2016). Effective coverage targets for ocean protection. *Conservation Letters*, 9(6), 398–404. <https://doi.org/10.1111/conl.12247>
- Pan, B. (2023). The effect of human activities on marine biodiversity. *Second International Conference on Biological Engineering and Medical Science (ICBioMed 2022)*, 130. <https://doi.org/10.1117/12.2669385>
- Petchey, O. L., & Gaston, K. J. (2006). Functional diversity: Back to basics and looking forward. *Ecology Letters*, 9(6), 741–758. <https://doi.org/10.1111/j.1461-0248.2006.00924.x>
- Pikitch, E. K., Santora, C., Babcock, E. A., Bakun, A., Bonfil, R., Conover, D. O., Dayton, P., Doukakis, P., Fluharty, D., Heneman, B., Houde, E. D., Link, J., Livingston, P. A., Mangel, M., McAllister, M. K., Pope, J., & Sainsbury, K. J. (2004). Ecosystem-Based Fishery Management. *Science*, 305(5682), 346–347. <https://doi.org/10.1126/science.1098222>
- Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52(3), 273–288. <https://doi.org/10.1016/j.ecolecon.2004.10.002>
- 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), 1246752–1246752.  
<https://doi.org/10.1126/science.1246752>
- Polis, G. A., & Strong, D. R. (1996). Food Web Complexity and Community Dynamics. *The American Naturalist*, *147*(5), 813–846. <https://doi.org/10.1086/285880>
- Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology*, *21*(8), 1931–1950. <https://doi.org/10.1111/j.1365-294X.2011.05403.x>
- Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, *51*(5), 1450–1459. <https://doi.org/10.1111/1365-2664.12306>
- Rounsevell, M. D. A., Harfoot, M., Harrison, P. A., Newbold, T., Gregory, R. D., & Mace, G. M. (2020). A biodiversity target based on species extinctions. *Science*, *368*(6496), 1193–1195. <https://doi.org/10.1126/science.aba6592>
- 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>
- Saeedi, H., Reimer, J. D., Brandt, M. I., Dumais, P. O., Jażdżewska, A. M., Jeffery, N. W., Thielen, P. M., & Costello, M. J. (2019). Global marine biodiversity in the context of achieving the Aichi Targets: Ways forward and addressing data gaps. *PeerJ*, *2019*(10), 1–17. <https://doi.org/10.7717/peerj.7221>
- Salter, I. (2018). Seasonal variability in the persistence of dissolved environmental DNA (eDNA) in a marine system: The role of microbial nutrient limitation. *PLOS ONE*, *13*(2), e0192409. <https://doi.org/10.1371/journal.pone.0192409>
- Salter, I., Joensen, M., Kristiansen, R., Steingrund, P., & Vestergaard, P. (2019). Environmental DNA concentrations are correlated with regional biomass of Atlantic cod in oceanic waters. *Communications Biology*, *2*(1), 1–9. <https://doi.org/10.1038/s42003-019-0696-8>
- Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., & Boehm, A. B. (2016). Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine Fish. *Environmental Science & Technology*, *50*(19), 10456–10464. <https://doi.org/10.1021/acs.est.6b03114>
- Schadewell, Y., & Adams, C. I. M. (2021). Forensics Meets Ecology – Environmental DNA Offers New Capabilities for Marine Ecosystem and Fisheries Research. *Frontiers in Marine Science*, *8*, 668822. <https://doi.org/10.3389/fmars.2021.668822>
- Schneider, J., Valentini, A., Dejean, T., Montarsi, F., Taberlet, P., Glairot, O., & Fumagalli, L. (2016). Detection of Invasive Mosquito Vectors Using Environmental DNA (eDNA) from Water Samples. *PLOS ONE*, *11*(9), e0162493. <https://doi.org/10.1371/journal.pone.0162493>
- Shelton, A. O., Gold, Z. J., Jensen, A. J., D’Agnese, E., Andruszkiewicz Allan, E., Van Cise, A., Gallego, R., Ramón-Laca, A., Garber-Yonts, M., Parsons, K., & Kelly, R. P. (2023). Toward quantitative metabarcoding. *Ecology*, *104*(2). <https://doi.org/10.1002/ecy.3906>
- Shelton, A. O., O’Donnell, J. L., Samhour, J. F., Lowell, N., Williams, G. D., & Kelly, R. P. (2016). A framework for inferring biological communities from environmental DNA. *Ecological Applications*, *26*(6), 1645–1659. <https://doi.org/10.1890/15-1733.1>
- Shelton, A. O., Ramón-Laca, A., Wells, A., Clemons, J., Chu, D., Feist, B. E., Kelly, R. P., Parker-Stetter, S. L., Thomas, R., Nichols, K. M., & Park, L. (2022). Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean.

- Proceedings of the Royal Society B: Biological Sciences*, 289(1971), 20212613.  
<https://doi.org/10.1098/rspb.2021.2613>
- Stephens, T. (2023). The Kunming–Montreal Global Biodiversity Framework. *International Legal Materials*, 62(5), 868–887. <https://doi.org/10.1017/ilm.2023.16>
- Stewart, R. I. A., Dossena, M., Bohan, D. A., Jeppesen, E., Kordas, R. L., Ledger, M. E., Meerhoff, M., Moss, B., Mulder, C., Shurin, J. B., Suttle, B., Thompson, R., Trimmer, M., & Woodward, G. (2013). Mesocosm Experiments as a Tool for Ecological Climate-Change Research. In *Advances in Ecological Research* (Vol. 48, pp. 71–181). Elsevier.  
<https://doi.org/10.1016/B978-0-12-417199-2.00002-1>
- Stoeckle, M. Y., Adolf, J., Charlop-Powers, Z., Dunton, K. J., Hinks, G., & VanMorter, S. M. (2021). Trawl and eDNA assessment of marine fish diversity, seasonality, and relative abundance in coastal New Jersey, USA. *ICES Journal of Marine Science*, 78(1), 293–304.  
<https://doi.org/10.1093/icesjms/fsaa225>
- Strickler, K. M., Fremier, A. K., & Goldberg, C. S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, 183, 85–92.  
<https://doi.org/10.1016/j.biocon.2014.11.038>
- Suter, L., Wotherspoon, S., Kawaguchi, S., King, R., MacDonald, A. J., Nester, G. M., Polanowski, A. M., Raymond, B., & Deagle, B. E. (2023). Environmental DNA of Antarctic krill (*Euphausia superba*): Measuring DNA fragmentation adds a temporal aspect to quantitative surveys. *Environmental DNA*, 5(5), 945–959. <https://doi.org/10.1002/edn3.394>
- Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, 21(8), 1789–1793. <https://doi.org/10.1111/j.1365-294X.2012.05542.x>
- Takahara, T., Minamoto, T., & Doi, H. (2013). Using Environmental DNA to Estimate the Distribution of an Invasive Fish Using Environmental DNA to Estimate the Distribution of an Invasive Fish Species in Ponds. *PLoS ONE*, 8(2), e56584.  
<https://doi.org/10.1371/journal.pone.0056584>
- 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), 1–9. <https://doi.org/10.1371/journal.pone.0041732>
- Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466(7310), 1098–1101. <https://doi.org/10.1038/nature09329>
- Tsuji, S., & Shibata, N. (2021). Identifying spawning events in fish by observing a spike in environmental DNA concentration after spawning. *Environmental DNA*, 3(1), 190–199.  
<https://doi.org/10.1002/edn3.153>
- Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, 183, 93–102.  
<https://doi.org/10.1016/j.biocon.2014.11.017>
- Turon, M., Nygaard, M., Guri, G., Wangensteen, O. S., & Præbel, K. (2022). Fine-scale differences in eukaryotic communities inside and outside salmon aquaculture cages revealed by eDNA metabarcoding. *Frontiers in Genetics*, 13, 957251. <https://doi.org/10.3389/fgene.2022.957251>
- Vačkář, D., Ten Brink, B., Loh, J., Baillie, J. E. M., & Reyers, B. (2012). Review of multispecies indices for monitoring human impacts on biodiversity. *Ecological Indicators*, 17, 58–67.  
<https://doi.org/10.1016/j.ecolind.2011.04.024>
- Van Overzee, H. M. J., & Rijnsdorp, A. D. (2015). Effects of fishing during the spawning period: Implications for sustainable management. *Reviews in Fish Biology and Fisheries*, 25(1), 65–83. <https://doi.org/10.1007/s11160-014-9370-x>

- Voelker, M. R., Schwarz, D., Thomas, A., Nelson, B. W., & Acevedo-Gutiérrez, A. (2020). Large-scale molecular barcoding of prey DNA reveals predictors of intrapopulation feeding diversity in a marine predator. *Ecology and Evolution*, *10*(18), 9867–9885. <https://doi.org/10.1002/ece3.6638>
- Wangensteen, O. S., Cebrian, E., Palacín, C., & Turon, X. (2018). Under the canopy: Community-wide effects of invasive algae in Marine Protected Areas revealed by metabarcoding. *Marine Pollution Bulletin*, *127*, 54–66. <https://doi.org/10.1016/j.marpolbul.2017.11.033>
- Warner, R. M. (2014). Conserving marine biodiversity in areas beyond national jurisdiction: Co-evolution and interaction with the law of the sea. *Frontiers in Marine Science*, *1*. <https://doi.org/10.3389/fmars.2014.00006>
- 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>
- Yates, M. C., Fraser, D. J., & Derry, A. M. (2019). Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. *Environmental DNA*, *1*(1), 5–13. <https://doi.org/10.1002/edn3.7>
- Yates, M. C., Wilcox, T. M., Stoeckle, M. Y., & Heath, D. D. (2022). Interspecific allometric scaling in eDNA production among northwestern Atlantic bony fishes reflects physiological allometric scaling. *Environmental DNA*, edn3.381. <https://doi.org/10.1002/edn3.381>
- Zhang, S., Lu, Q., Wang, Y., Wang, X., Zhao, J., & Yao, M. (2020). Assessment of fish communities using environmental DNA: Effect of spatial sampling design in lentic systems of different sizes. *Molecular Ecology Resources*, *20*(1), 242–255. <https://doi.org/10.1111/1755-0998.13105>
- Zou, K., Chen, J., Ruan, H., Li, Z., Guo, W., Li, M., & Liu, L. (2020). eDNA metabarcoding as a promising conservation tool for monitoring fish diversity in a coastal wetland of the Pearl River Estuary compared to bottom trawling. *Science of The Total Environment*, *702*. <https://doi.org/10.1016/j.scitotenv.2019.134704>





# Paper 1

## **Maximizing sampling efficiency to detect differences in fish community composition using environmental DNA metabarcoding in subarctic fjords.**

Gledis Guri<sup>1,2\*</sup>, Jon-Ivar Westgaard<sup>1</sup>, Nigel Yoccoz<sup>3</sup>, Owen S. Wangensteen<sup>2,4</sup>, Kim Præbel<sup>2</sup>, Jessica Louise Ray<sup>5</sup>, Ryan P. Kelly<sup>6</sup>, Andrew Olaf Shelton<sup>7</sup>, Tanja Hanebrekke<sup>1</sup>, Torild Johansen<sup>1</sup>









1. Norwegian Institute of Marine Research, Framsenteret, Tromsø, Norway
2. Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway
3. Dept. of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway
4. Dept. of Evolutionary Biology, Ecology & Environmental Sciences, Universitat de Barcelona, Barcelona, Spain
5. NORCE Norwegian Research Centre AS, Uni Research Environment, N-5020 Bergen, Norway
6. School of Marine and Environmental Affairs, University of Washington, Seattle, Washington, USA
7. Northwest Fisheries Science Center, NOAA Fisheries, Seattle, Washington, USA

\*Corresponding author: Gledis Guri, [gledis.guri@hi.no](mailto:gledis.guri@hi.no), +46 700658104

*Environmental DNA*, edn3.409 (2023). <https://doi.org/10.1002/edn3.409>



# Maximizing sampling efficiency to detect differences in fish community composition using environmental DNA metabarcoding in subarctic fjords

Gledis Guri<sup>1,2</sup>  | Jon-Ivar Westgaard<sup>1</sup>  | Nigel Yoccoz<sup>3</sup>  | Owen S. Wangensteen<sup>2,4</sup>  |  
Kim Præbel<sup>2</sup>  | Jessica Louise Ray<sup>5</sup>  | Ryan P. Kelly<sup>6</sup>  | Andrew Olaf Shelton<sup>7</sup>  |  
Tanja Hanebrekke<sup>1</sup>  | Torild Johansen<sup>1</sup> 

<sup>1</sup>Norwegian Institute of Marine Research, Framsenderet, Tromsø, Norway

<sup>2</sup>Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway

<sup>3</sup>Dept. of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway

<sup>4</sup>Dept. of Evolutionary Biology, Ecology & Environmental Sciences, Universitat de Barcelona, Barcelona, Spain

<sup>5</sup>NORCE Norwegian Research Centre AS, Climate & Environment Division, NO-5838, Bergen, Norway

<sup>6</sup>School of Marine and Environmental Affairs, University of Washington, Seattle, Washington, USA

<sup>7</sup>Northwest Fisheries Science Center, NOAA Fisheries, Seattle, Washington, USA

## Correspondence

Gledis Guri, Norwegian Institute of Marine Research, Framsenderet, Tromsø, Norway.  
Email: [gledis.guri@hi.no](mailto:gledis.guri@hi.no)

## Funding information

Norges Forskningsråd, Grant/Award Number: NRC301691

## Abstract

Environmental DNA (eDNA) has gained popularity as a tool for ecosystem biomonitoring and biodiversity assessment. Although much progress has been made regarding laboratory and fieldwork protocols, the issue of sampling efficiency requires further investigation, particularly in three-dimensional marine systems. This study focuses on fish community composition in marine ecosystems and aims to analyze the efficiency of sampling design given the sampling effort for distinguishing between different communities. We sampled three fjords in Northern Norway, taking samples along fjord transects and at three different depths, and amplified a fragment of the mitochondrial 12S rRNA gene of bony fishes using the MiFish primers. We evaluated the effect of (i) the number of sampling stations, (ii) samples' spatial distribution, and (iii) the data treatment approach (presence/absence versus semiquantitative) for maximizing the efficiency of eDNA metabarcoding sampling when inferring differences of fish community compositions between fjords. We found that the manner of data treatment strongly affected the minimum number of sampling stations required to detect differences among communities; because the semiquantitative approach retained some information about abundance of the underlying reads, it was the most efficient. Furthermore, we found little-to-no difference of fish communities in samples from intermediate depths when comparing vertical fish communities. Lastly, we found that the differences between fish communities at the surface were the highest across the horizontal distance and overall, samples ~30 km apart showed the highest variation in the horizontal distribution. Boosting sampling efficiency (reducing sampling effort without compromising ecological inferences) can significantly contribute to enhanced biodiversity management and efficient biomonitoring plans.

## KEYWORDS

eDNA, fish, heterogeneity, metabarcoding, sampling design, spatial distribution

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Environmental DNA* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Biological monitoring, as a key component of the assessment of ecological resources (Gold et al., 2021), allows decision makers to adjust policies and management plans to achieve environmental targets and sustainable use of marine resources. Many studies have attested to the importance of biological monitoring for a robust decision-making (Borja & Elliott, 2013; Bourlat et al., 2013; Salter et al., 2019; Stat et al., 2017). Moreover, recent changes toward ecosystem-based fisheries management rely heavily on biodiversity assessments and multispecies identification where quantification of community metrics is crucial for reliable monitoring (Gullestad et al., 2017).

Multiple studies have encouraged the use of environmental DNA (eDNA) in biomonitoring for management policies as the DNA expelled from organisms can be captured and sequenced (Taberlet et al., 2012), enabling the detection of the community composition present at the study site (Ji et al., 2013). As a noninvasive sampling technique (Cilleros et al., 2019), the reliability of eDNA metabarcoding results is an attractive alternative to traditional capture sampling for biomonitoring (Barnes & Turner, 2016; Cantera et al., 2019; Gilbey et al., 2021). However, metabarcoding approaches for biomonitoring have only recently been developed and face several obstacles (e.g., amplification bias, DNA transport) that hinder the direct use of metabarcoding data for inference of ecological communities (Banerjee et al., 2021). An essential component of any eDNA metabarcoding study is to provide an adequate sampling design to address specific ecological questions. Moreover, minimizing the DNA sampling effort can enable stakeholders and policymakers to increase sampling events and thus broaden the monitoring scale of marine biodiversity assessments.

Optimizing biological and technical replication as a trade-off between decreasing false negatives and costs has been investigated (Buxton et al., 2021; Cantera et al., 2019). Furthermore, optimization of eDNA sampling design has been laid on water volume filtered, filter types, and effort distribution among methods for maximizing univariate measures of diversity (such as  $\alpha$  diversity) and single/rare species detection while decreasing sampling effort (Andres et al., 2022; Bessey et al., 2020; Sanches & Schreier, 2020; Wood et al., 2021). One additional study focused on the spatial distribution of eDNA samples in lentic systems to evaluate sampling design strategies for inferring patterns of fish biodiversity ( $\alpha$  and  $\beta$  diversity; Zhang et al., 2020), where they suggested a systematic spatial sampling protocol and concluded that eDNA samples autocorrelated up to 2 km on the horizontal distance. However, the latter study circumvented issues related to sampling effort; thus, at present, no research has been conducted on sampling design accounting for community compositions and heterogeneity as a function of the sampling effort. Additionally, in contrast to lentic and lotic systems, marine systems are more complex in DNA transport due to multidirectional and three-dimensional water flow (Hansen et al., 2018).

Obtaining an efficient sampling design for eDNA-based studies, especially in marine systems, is a challenging task (convoluted question due to the multitude of research questions, hence every context

has a specific sampling design; Yoccoz et al., 2001) that still requires attention. In this study, we focus maximizing the sampling efficiency explicitly on detecting ecosystems' fish community composition (hereafter communities); thus, circumventing univariate diversity metrics and single/rare species detection as such questions have been resolved. We study this problem in two different levels, that is, by investigating how well the sampled communities segregate given the unit of sampling effort, thus how much can we reduce the sampling effort to still obtain significant differences between these fjord communities, and second, which samples contribute the least to the dissimilarity metrics, thus can potentially be removed without affecting the variability of the ecosystem.

To achieve these objectives, we formulated the following questions: (i) What is the metabarcoding data treatment approach, presence/absence versus semiquantitative, that most efficiently represents the differences between fjords ecosystems, and what is the minimal number of sampling sites for each approach; and (ii) where to effectively collect the samples (in three-dimensional space) to maximize our efficiency to detect differences in ecological communities while minimizing sampling effort.

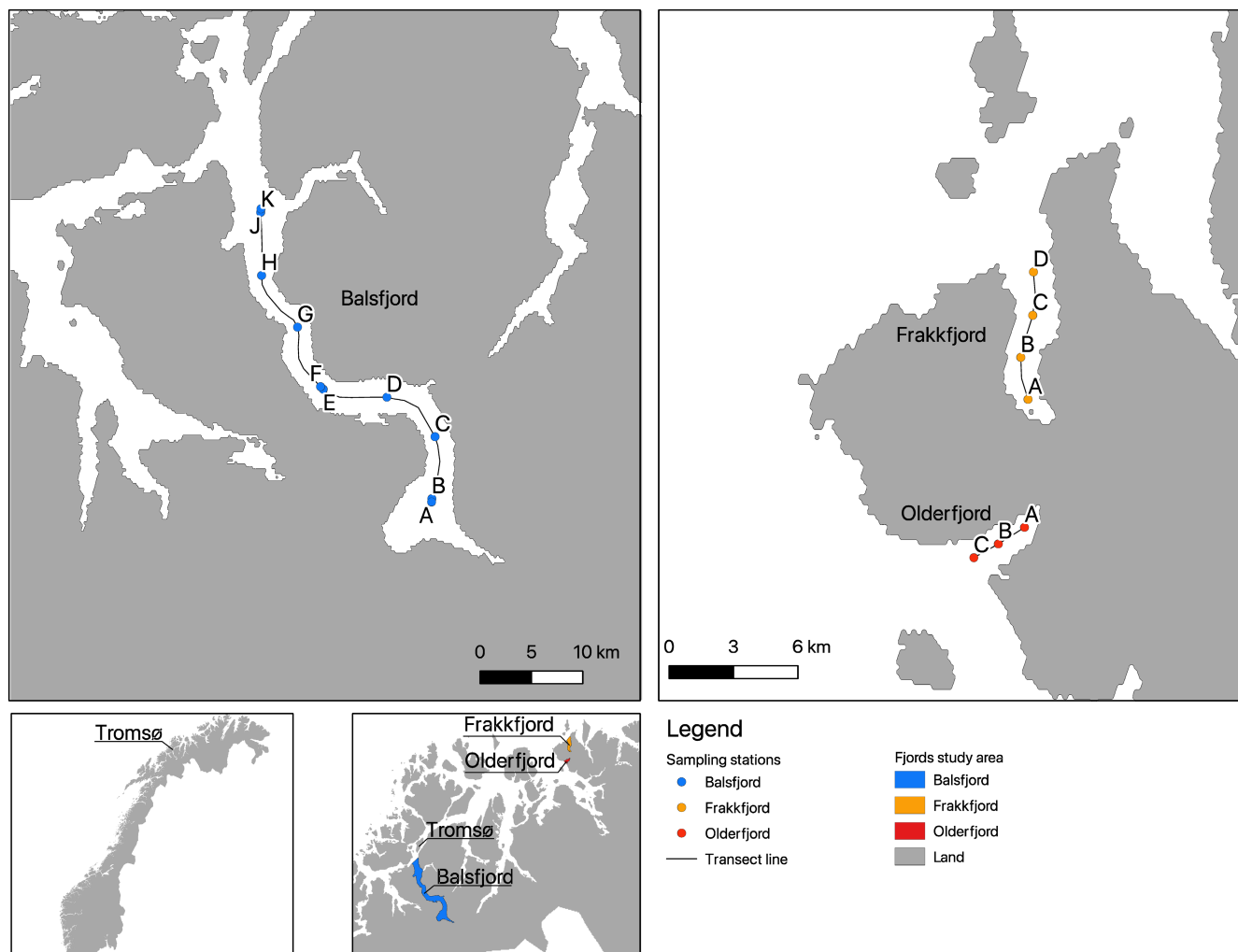
## 2 | METHODS

### 2.1 | Study sites

Study samples were collected on research cruises on March 2021 by R/V Kristine Bonnevie. Three Northern Norwegian fjords: Balsfjord, Frakkfjord, and Olderfjord, were selected based on their size, bathymetry, and distance from the open ocean. Balsfjord has a length of 40 km and a maximum depth of 190 m in the middle of the fjord. There is a sill at the fjord's entrance, and the archipelago limits the water exchange and categorizes it as protected from the open ocean. Frakkfjord and Olderfjord are relatively similar regarding their size and shape, with a fjord length of 8 and 4 km and a maximum depth is 86 and 74 m, respectively. These two fjords are more exposed to the open ocean compared with Balsfjord; however, Frakkfjord has the highest degree of exposure to open water. Nearly all high latitude Norwegian fjords are ice-free and depicted with an Arctic light regime (Reigstad & Wassmann, 1996). Although a subarctic water climate characterizes all fjords in the area, the degree of openness to the open sea can create differences regarding community composition. In total 17 eDNA sampling stations were deployed for the three fjords, of which 10 were located in Balsfjord, four in Frakkfjord and three in Olderfjord (Figure 1). GPS coordinates and other metadata of sampling stations are provided in Table S1.

### 2.2 | Water sampling, eDNA filtration, and extraction

To investigate the spatial heterogeneity of the eDNA signals for inference of fish communities, we established three sampling transects



**FIGURE 1** Map of the eDNA sampling station distributed among three localities (Balsfjord—blue, Frakkfjord—orange, and Olderfjord—red) and the radius of the area represented by the sampling station.

in three localities (i.e., three fjord ecosystems) where samples were assigned using fixed-distance nonoverlapping radii scaled by the size of the locality (i.e., 3000m in Balsfjord, 1000m in Frakkfjord, and 750m in Olderfjord). In alphabetical order, sampling stations were designated by letters starting from the innermost station (Figure 1). To constrain the background level of spatial heterogeneity from metabarcoding, we double-sampled three locations in Balsfjord using “twin” sampling stations next to each other (i.e., A/B, E/F, and J/K). At each sampling station, 12 5-L Niskin bottles mounted on a stainless-steel frame with CTD were deployed to desired sampling depths for water collection. The three depths investigated were surface (10m), pycnocline (depth of highest density, ~50m), and bottom (10m above bottom). At each sampling depth, triplicate 2-L water samples, derived from three distinct Niskin bottles, were filtered onboard the research vessel through 0.22 $\mu$ m Sterivex filters (MerckMillipore) using a peristaltic pump (multichannel flow Heidolph™ Hei-Flow Advantage 01). After removing the remaining water drops by pumping air, filters were transferred to sterile 50-mL Falcon centrifuge tubes and immediately stored at  $-20^{\circ}\text{C}$  until they were transported to the laboratory for  $-80^{\circ}\text{C}$  storage until DNA extraction. DNA was

extracted using DNeasy PowerWater Sterivex Kit (Qiagen GmbH) following the manufacturer's protocol with slight modifications (all steps involving PowerBead Tubes were omitted). DNA extraction from water samples was conducted randomly to minimize the possible biases from contamination during laboratory workflow.

### 2.3 | Contamination control

To control for contamination occurring at each step of the workflow, we included field-negative controls (one air and one water blank per sampling station), laboratory-negative controls (one extraction blank per block of 12 extracted samples; three PCR blanks per block of 96 extracted samples), and PCR-positive controls (one positive control PCR sample per block of 96 extracted samples) as described by Shu et al. (2020). To reduce the risk of cross-contamination during the sampling event, all sampling equipment was decontaminated with 20% (v/v) sodium hypochlorite solution (household bleach) and then rinsed with Milli-Q water onboard before and after fieldwork. Sterile nitrile gloves were used when in contact with water samples.

To minimize the risk of sample cross-contamination in the laboratory, before and after each round of DNA extraction, all the workbenches (extraction and PCR hood) were decontaminated through UV exposure for 30 min, as suggested by Goldberg et al. (2016). Additionally, the air in the laboratory was decontaminated regularly (once a week) with hydrogen peroxide. Lastly, pre- and post-PCR tasks were performed in different laboratories to reduce the risk of cross-contamination.

## 2.4 | Library preparation and sequencing

A total of 192 samples (divided into two libraries), including eDNA samples ( $n=150$ ), PCR blanks ( $n=6$ ), positive controls ( $n=2$ ), extraction blanks ( $n=6$ ), fieldwork water and air blank ( $n=16$  and  $12$  respectively), were amplified using the MiFish-U universal primer set (Forward: 5'-GTCGGTAAACTCGTGCCAGC-3'; Reverse: 5'-CATAGTGGGTATCTAATCCCAGTTTG-3'; Miya et al., 2015) targeting the mitochondrial 12S rRNA gene region (169–172 bp fragment). Fusion primers containing adaptor, index, and primer-specific sequences were used to allow one-step PCR amplification, where each 20  $\mu$ L PCR reaction consisted of 3  $\mu$ L eDNA template, 1  $\mu$ L of primer mix (0.5  $\mu$ L each of 5  $\mu$ M stock solutions of MiFish-U-forward and MiFish-U-reverse primers), 10  $\mu$ L of QIAGEN Multiplex PCR Master Mix, 0.16  $\mu$ L of BSA (concentration 20  $\mu$ g/mL), and lastly 5.84  $\mu$ L of RNase-free water. All samples were amplified in triplicate. The thermocycler program included an initial denaturing step of 95°C for 10 min, 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s and a final extension step of 72°C for 5 min.

PCR amplification success and product size were assessed using capillary electrophoresis (QIAxcel; Qiagen GmbH). Products from triplicate PCR reactions were pooled into a single tube before 100  $\mu$ L from each PCR product pool were combined to generate two sequencing libraries, each consisting of 96 samples/library. Sequencing libraries were electrophoresed on a 2% (w/v) agarose gel in 1 $\times$  TAE buffer and stained with SYBR safe (Qiagen GmbH). Bands of the expected size (300 bp) were excised and purified using GeneJet Gel Extraction and DNA Cleanup Micro Kit (Thermo Fisher Scientific). DNA concentration of the purified sequencing libraries was measured using Qubit dsDNA HS kit (Thermo Fisher Scientific), diluted to a final concentration of 50 pM and spiked with 4  $\mu$ L of Ion S5 Calibration Standard prior to loading onto the Ion Chef instrument (Thermo Fisher Scientific). Libraries were sequenced using the Ion 530-sequencing chip with the 200 bp protocol (Thermo Fisher Scientific).

## 2.5 | Bioinformatics

Sequences were automatically demultiplexed and quality filtered after the sequencing process using Torrent Suite™ inbuilt in the sequencer following their inbuilt standard settings. The sampled

sequence dataset was thereafter filtered for chimeric sequences using a *uchime-denovo* algorithm in VSEARCH (Rognes et al., 2016). Second, we clustered the sequences into Molecular Operational Taxonomic Units (MOTU) using SWARM v2 (Mahé et al., 2014) with a distance of  $d=3$ . Third, singletons were removed, and the remaining MOTUs were taxonomically annotated using the *ecotag* algorithm (Boyer et al., 2016) and a well-curated local database assembled from data available from EMBL and NCBI (access date: 14 March 2022).

To select only high-quality sequences and remove pseudogenes and artifacts, we arbitrarily customized a low-quality dataset filtering process where MOTUs with an identity match  $\geq 97\%$  (Li et al., 2018; Nakagawa et al., 2018; Sales et al., 2021) or higher number of reads than 1/10,000 of the total library reads were retained (Coguiec et al., 2021). MOTUs unassigned to species taxonomic rank after *ecotag* algorithm were manually checked by blasting in NCBI online database (date: March 14, 2022) where the same selection criteria were used. For downstream biodiversity analysis, we removed biological replicates with sequencing depth lower than 500 reads as an indication of low-quality samples. Amplicon sequences whose occurrence in negative controls exceeded 10% of their total read abundance in the entire dataset were removed from the dataset as an indication of contamination. Additionally, we summed (pooled) all MOTU reads across the biological replicates of the same depth taken in the same sampling station. Lastly, we removed all taxa that were not assigned to fish (class: Actinopterygii or Chondrichthyes).

## 2.6 | Statistical analysis

All analyses and plots were performed in R (R Core Team, 2022). To see that our sequencing depth was not a limiting factor for representing the taxonomic diversity of the study area, we examined the DNA sequencing effort through species accumulation curves (hereafter sequencing effort curves), using *rarecurve* function in the *vegan* package. The curves plotted the average number of species identified as a function of sequencing depth, indicating if the latter was sufficient for covering the biodiversity. Additionally, species accumulation curves (hereafter DNA collection curves), the number of species against the number of samples, were plotted to examine whether more samples would yield more species, thus determining whether our DNA collection effort was sufficient for covering the taxonomic diversity of the sampled area. The DNA collection curves were drawn using *specaccum* function in *vegan* package.

## 2.7 | Efficiency of data treatment approaches and minimal number of sampling stations for detecting community differences

To explore the strength of two commonly used metabarcoding data treatment approaches on the ability to detect differences between

fish communities, we selected and compared qualitative (presence/absence) and semiquantitative approaches. We outline the qualitative approach as the binary transformation of MOTU reads (after quality filtering) into the presence and absence where MOTUs with reads  $\geq 1$  were defined as a presence. For the semiquantitative approach, we selected the eDNA index proportion model developed by Kelly et al. (2019), also called inverse “Wisconsin double-standardization,” a simplified way to account for species-specific differences in amplification efficiency. The following analyses were performed on both qualitative and semiquantitative treatment of the eDNA reads of pooled replicates. We used the Jaccard and Bray–Curtis distance matrices, respectively. The dissimilarity matrices between ecosystems were initially represented visually through nonmetric multidimensional scaling (nMDS) ordination using *metaMDS* in *vegan* package (Oksanen et al., 2022), with two dimensions ( $k=2$ ) while using 20 random starts in search of a stable solution. Differences in communities between ecosystems were computed through permutational analysis of variances (PERMANOVA) using *adonis* function in *vegan* package (Oksanen et al., 2022) with factor location (Balsfjord, Frakkfjord, Olderfjord, and positive control) and 999 permutations. For significant PERMANOVA results, we conducted permutational multivariate dispersion test (PERMDISP) to check whether the significance was due to the differences in centroids or due to the differences in dispersion of samples in principal coordinate space of dissimilarity (PCoA). Additionally, we used permutational pair-wise comparisons with the Benjamini–Yekutieli FDR correction—False Discovery Rate (Benjamini & Yekutieli, 2001) to indicate the paired location comparisons that were significantly different. We used *betadisper* and *pairwise.adonis2* functions for both tests, respectively, supplied by *vegan* package (Oksanen et al., 2022). To indicate the species driving the differences between localities, we used indicator species analysis (Dufrene & Legendre, 1997) on both approaches using *indval* function in *labdsv* package (Roberts, 2019) with 999 permutations.

Furthermore, to investigate the efficiency of each data treatment approach (qualitative and semiquantitative), we conducted a linear regression of dissimilarity (measured in pseudo- $F$ -value) as a function of sampling effort (number of sampling stations) for each approach using *lm* function in R with the formula  $\text{dissimilarity} \sim \text{sampling effort} \times \text{method}$ . We measured the dissimilarity as the pseudo- $F$  and  $p$  values of a series of PERMANOVA tests for detecting ecosystem differences (factor locality) across a range of sampling effort. The range of sampling effort was conducted through progressively removing one sampling station from each locality (excluding positive controls) until one sampling station remained in each locality. During each removal of stations, the dataset was randomly subsampled 999 times and we estimated the mean of  $p$  and pseudo- $F$ -values from PERMANOVA tests for each combination. Lastly, the minimal number of sampling stations required to infer differences between ecosystems was concluded as the mean plus two standard deviation of PERMANOVA  $p$ -values lower than 0.05. Subsequently, the data treatment approach with the highest efficiency (highest pseudo- $F$  value) was selected for all analyses below.

## 2.8 | Sampling efficiency for maximum vertical and horizontal community difference

To investigate the spatial relationship of eDNA samples and maximum dissimilarity (sampling efficiency), we selected only samples taken in Balsfjord due to the linear shape of the fjord and the extensive number of sampling stations investigated ( $n=10$ ).

For exploring the sampling efficiency among depth layers, we explored whether the community composition differed significantly between each depth category using PERMANOVA analysis with Bray–Curtis dissimilarity index with factor depth category. Thereafter a post hoc test was conducted for identifying pairs of samples that differed using permutational pair-wise comparisons with the Benjamini–Yekutieli FDR correction. Subsequently, for significant PERMANOVA results, we conducted PERMDISP test. Samples of depth categories that did not differ significantly were defined uninformative, thus can be removed for reducing sampling effort.

Additionally, we explored the dissimilarity of the community among the horizontal distance of the fjord. We fitted a beta regression using *betareg* function in *betareg* package with the pairwise Bray–Curtis dissimilarity index (measured through *vegdist* function in *vegan* package) as response variable, and the log-transformed pairwise horizontal distance (measured in QGIS) and pairwise depth categories as predictor using the formula  $\text{dissimilarity} \sim \log_{10}(\text{distance} + 1) \times \text{depth category}$ . Note that only pairwise comparisons of the similar depth categories were included (i.e., bottom–bottom, pycnocline–pycnocline, and surface–surface). We concluded that the most efficient horizontal distance for sampling was the minimum distance with the maximized community dissimilarity among samples.

## 3 | RESULTS

### 3.1 | Sequencing and eDNA sampling effort

After the standard Ion GeneStudio quality and sequence-length filtering, the run yielded 13,579,483 sequence reads from 150 eDNA samples (4,984,499 and 8,594,984, respectively, for each of the two sample pools), two positive controls, and 40 negative controls (Figure S1a). Filtering of chimeric reads and singletons resulted in the removal of 64,581 reads. After our customized low-quality removal and blank treatment subtraction, the final dataset (152 samples) used for analysis contained 13,398,370 reads. We observed very few reads in all negative controls (mean =  $32.2 \pm 29.3$ ; Figure S1b), indicating a low level of contamination. As a result, no “contaminating” MOTUs were removed from the dataset. From the quality-filtered and nonchimeric reads, we detected 33 marine taxa, of which 31 were assigned to the species level, and the remaining two (*Sebastes* spp. and *Anarhichas* spp.) were assigned to the genus level. Around 99.99% of MOTU reads belonged to nine different orders of bony



fishes (Actinopterygii), while the remaining reads were assigned to *Amblyraja radiata* (Chondrichthyes).

Sequencing effort curves revealed that samples in Frakkfjord and Olderfjord had higher sequencing depth than the Balsfjord samples (Figure 2a). Overall, most curves from samples with high sequencing depth saturated, while the remaining samples achieved a slope change but did not reach a plateau. However, visual inspection indicated that over ca. 500,000 reads per sampling station (samples pooled) would be needed to approximate the full diversity of the samples. This indicated that our sequencing effort (>1,000,000 reads per sampling station) in Frakkfjord and Olderfjord successfully covered the taxonomic complexity, while in Balsfjord (~300,000 reads per sampling station), such coverage was not achieved.

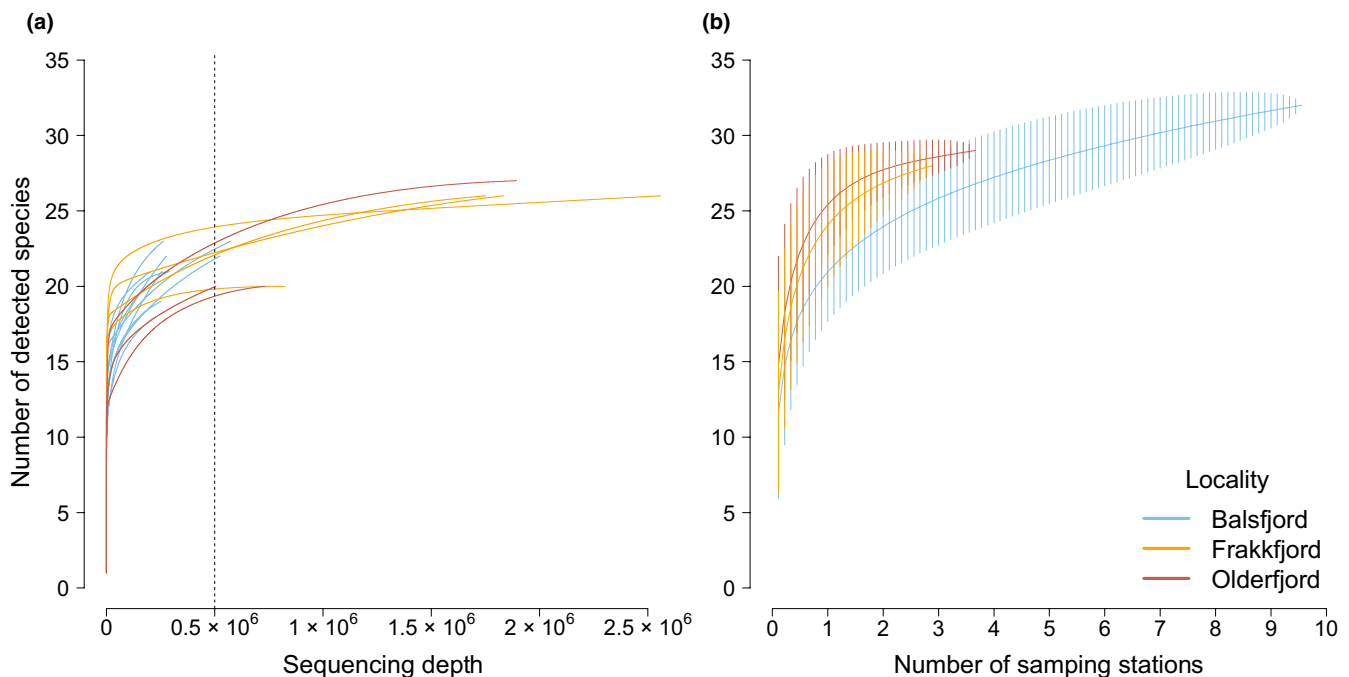
DNA collection curves (Figure 2b) signaled that samples in Frakkfjord and Olderfjord detected most species present as the curve saturated around two sampling stations, stating that the taxonomic diversity of the fjord was successfully covered from the samples deployed. In contrast, the Balsfjord curve did not plateau, indicating that deploying more samples could potentially detect more species in the fjord.

### 3.2 | Efficiency of data treatment approaches and minimal number of sampling stations for detecting community differences

Two nonmetric multidimensional scaling ordination (nMDS) plots displayed dissimilarities of species composition among the three localities

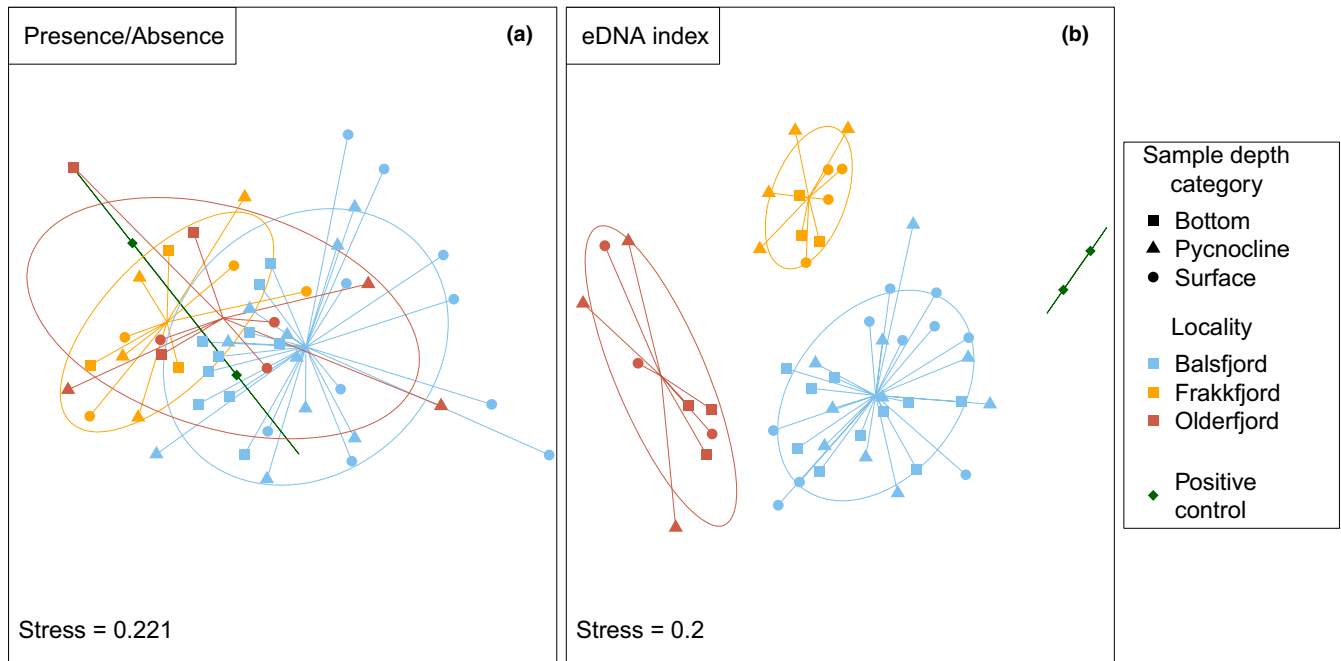
(biological replicates pooled), based on Jaccard (Figure 3a) and Bray-Curtis (Figure 3b) dissimilarity index for qualitative (presence/absence) and semiquantitative (eDNA index) representation of the metabarcoding data, respectively. The qualitative-based nMDS indicated separation of the centroids of the samples based on their locality, with a strong fish community variation and overlap among localities. Furthermore, positive control samples were embedded within 70% confidence interval ellipsoids of all locality centroids. Conversely, the semiquantitative-based nMDS revealed a strong separation of clusters by locality, stating that the semiquantitative community composition in each locality is distinct. Additionally, positive control samples are confidently separated from all localities' community composition.

Both PERMANOVAs (Table 1) indicated significant differences in the composition between localities ( $p=0.005$  and  $p<0.001$ ). Pseudo- $F$  and  $R^2$  values differed considerably between the tests (Table 1). Semiquantitative-based pair-wise tests indicated significant differences between each pair-wise comparison of locality (including the positive control samples; Table S2b). Conversely, qualitative tests failed to identify significant differences between Frakkfjord and Olderfjord (Table S2a). Additionally, all pair-wise differences between positive control and each locality were found to be nonsignificant (Table S2a,b). Semiquantitative-based PERMDISP was found significant ( $p<0.001$ ; Table 1), indicating a difference in dispersion of samples (average distance to its centroid). Pair-wise comparison of dispersion revealed a significant difference between Balsfjord and Frakkfjord and all pair-wise comparisons of positive control (Table S2b). Conversely, qualitative-based PERMDISP indicated a homogeneous dispersion among all localities (Table 1).



**FIGURE 2** Species accumulation curves as a function of sequencing depth, thus sequencing effort curves (a; samples and their biological replicates pooled together) and as a function of number of the sampling stations, thus DNA collection effort curve (b; biological replicates not pooled) colored for each locality. The vertical dashed line indicates the sequencing depth (500,000 reads) where sample curves start to plateau.





**FIGURE 3** Nonmetric multidimensional scaling (nMDS) ordination of samples (biological replicates pooled) obtained using qualitative approach using Jaccard dissimilarity index based on presence/absence of species (a) and semiquantitative using Bray–Curtis dissimilarity index based on inverse Wisconsin double-standardization of amplicon reads (b). The centroids (differed by colors) for the localities (positive control included as locality) and their ellipses (95% confidence interval of the group centroid dispersion) are indicated. Higher dissimilarity of ecosystems was observed when using semiquantitative approach.

**TABLE 1** PERMANOVA tests results using qualitative (Jaccard dissimilarity matrix) and semiquantitative (Bray–Curtis dissimilarity matrix) approaches comparing community composition among the factor locality (positive control included).

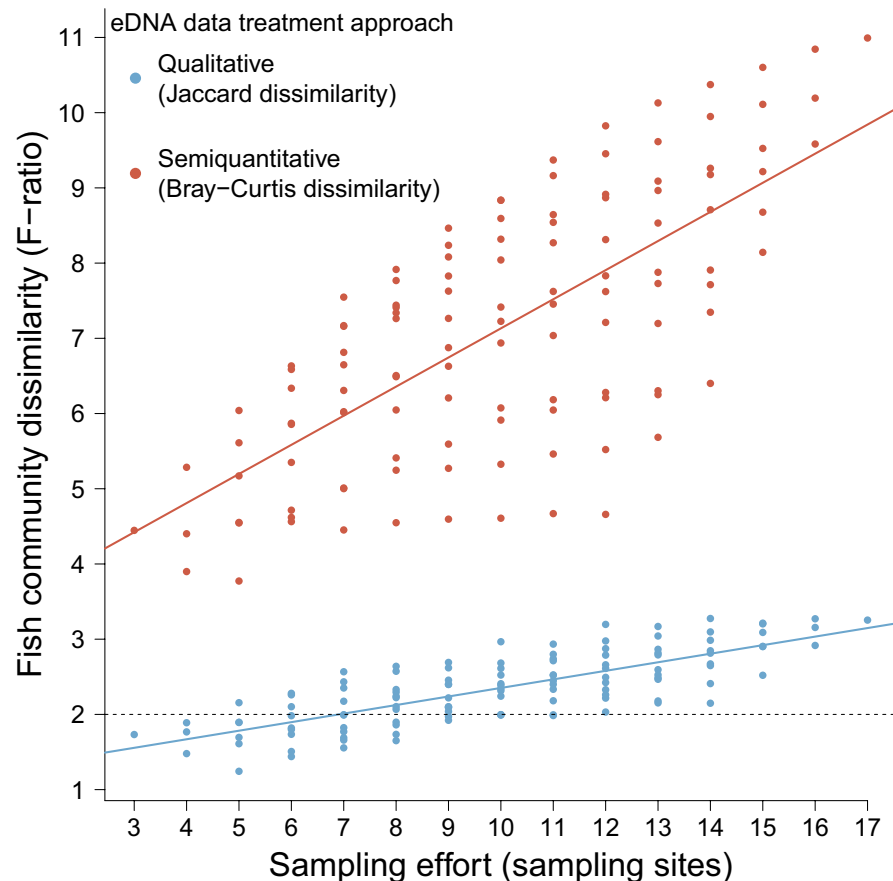
Approach	Factor	df	SS	MS	Pseudo-F	$R^2$	$p$ -Value	PERMDISP
Qualitative	Locality (fjords)	3	0.423	0.141	2.525	0.1363	0.005	0.618
	Residuals	48	2.683	0.055		0.8637		
	Total	51	3.107			1.0000		
Semi-quantitative	Locality (fjords)	3	5.026	1.675	9.011	0.3603	0.001	0.001
	Residuals	48	8.923	0.185		0.6397		
	Total	51	13.949			1.0000		

Note: PERMDISP probabilities for homogeneity of dispersion are also shown. Significant values are indicated in bold.

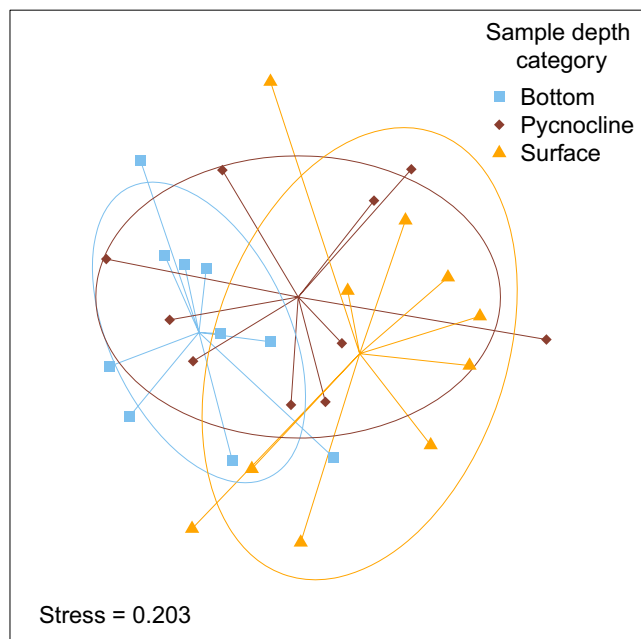
Comparing the two dissimilarity matrices (qualitative and semi-quantitative based), the Mantel test revealed a significant correlation ( $r=0.2$ ,  $p<0.001$ ), stating that the two data treatment approaches were weakly positively associated.

The linear regression analysis (Figure 4) indicated a significant positive linear relation of ecosystems' dissimilarity with sampling effort for each data treatment approach ( $p<0.001$  for both,  $t=4.64$  and  $t=7.51$  respectively and  $R^2=0.89$ ; Table S4). Moreover, the analysis indicated a significant different intercept ( $p<0.001$ ; number of stations=3) for qualitative (pseudo- $F=1.52$ ) compared with semi-quantitative (pseudo- $F=4.42$ ; Table S4) and significant slope (0.12 and 0.39, respectively). Reflecting the greater information content in the semiquantitative data transformation, three semiquantitative samples produced a degree of resolution equivalent to 13 presence/absence samples.

Indicator species analysis differed considerably between the two approaches (Figure S2). The semiquantitative approach indicated 13 species as significant drivers of community compositional differences between the localities. Moreover, all three localities contained some species indicators (relative frequency of occurrence and relative average abundance of species was the highest within the indicative locality), whereas Balsfjord and Frakkfjord included five species each, and Olderfjord three species. Conversely, the qualitative approach only found five species as significant drivers of the community differences and no species driver defined for Balsfjord. Further investigation of the commonness of species in each community showed six species as common, three as semicommon and four rare species in the semiquantitative data (Table S3), while the qualitative data included three rare species and two semicommon species (Table S3).



**FIGURE 4** Dissimilarity of community composition between ecosystems measured as the PERMANOVA pseudo-F ratio using qualitative (blue; Jaccard dissimilarity matrix based on the presence/absence of species) and semiquantitative (red; Bray-Curtis dissimilarity matrix based on the inverse Wisconsin double-standardization of amplicon reads) approaches. The analyses were performed on progressively removed sampling stations, with each dot representing the mean pseudo-F of the sample combination (iterated 999 times) for the given number of remaining sampling stations. Regression lines for each approach are shown with their colors, respectively. The horizontal black dashed line, representing an approximation of pseudo-F value ( $f=2.01$ ) equal to  $p=0.05$ , indicated that minimum three and 13 sampling stations are required for detecting significant differences between the three ecosystems respectively for semiquantitative and qualitative approach.



**FIGURE 5** Nonmetric multidimensional scaling (nMDS) ordination of Balsfjord samples (biological replicates pooled) obtained using the Bray-Curtis dissimilarity index based on the inverse Wisconsin double-standardization of amplicon reads. The centroids for each depth category (distinguished by color) and their 95% confidence intervals (shown as ellipses) indicate a gradual shift of the vertical communities in the multivariate space.

### 3.3 | Sampling efficiency for maximum vertical and horizontal community difference

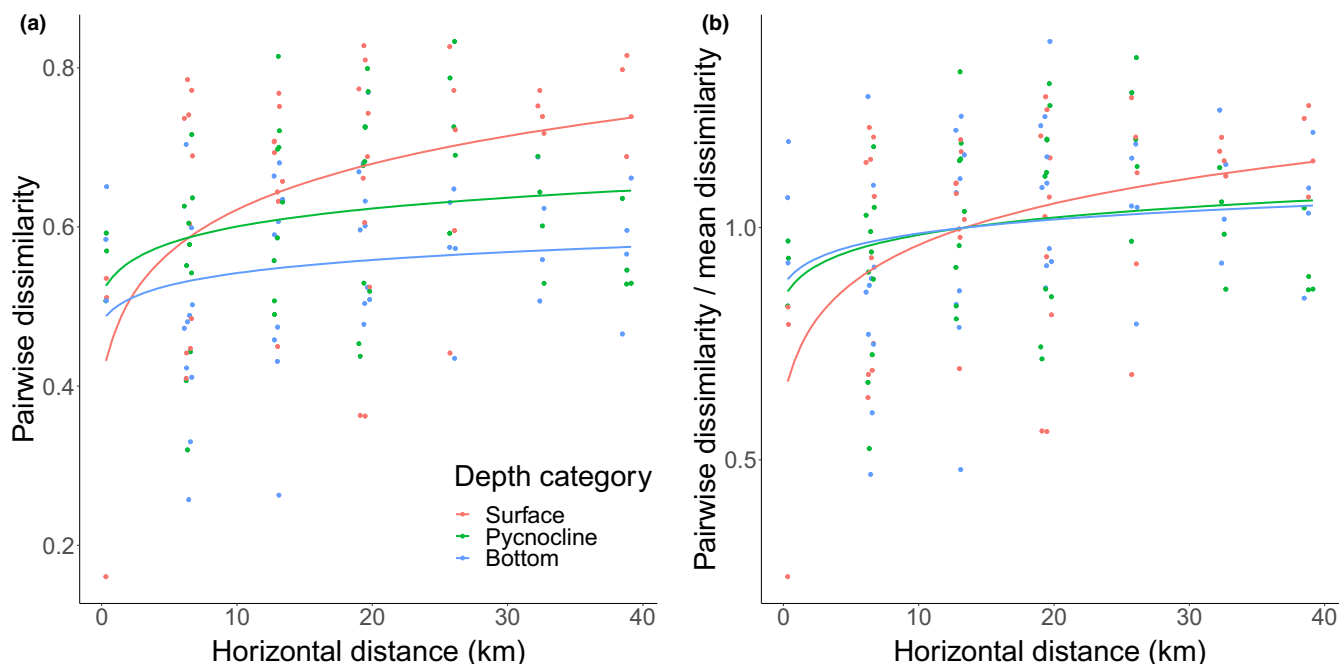
The nMDS plot showed a gradual change of the communities along the depth as the centroids shifted aligned with the depth vector (Figure 5). PERMANOVA test indicated significant differences in communities between at least two depth categories in both approaches ( $p=0.014$ ; Table 2). Pair-wise comparison of communities showed significant differences between bottom and surface samples ( $p<0.001$ ; Table S2c). However, no significant differences were found between bottom - pycnocline and surface - pycnocline pair comparisons ( $p=0.118$  and  $p=0.396$ , respectively). PERMDISP indicated homogeneous dispersion of fish communities among depth categories, which suggests that community differences are solely due to centroid differences and not due to differences in heterogeneity.

Beta regression indicated that dissimilarity of pycnocline and surface samples differed significantly across the horizontal distance ( $p=0.024$ , and  $p<0.001$ ) but these changes were insignificant on bottom samples ( $p=0.237$ ; Table S5). Changes in dissimilarity across distance were significantly different for surface samples compared to pycnocline and bottom samples ( $p<0.001$  for both; Table S5). Conversely, changes in pycnocline samples dissimilarity index across the horizontal distance compared to bottom samples resulted insignificant ( $p=0.439$ ; Table S5). The overall maximum dissimilarity for pycnocline and bottom samples was achieved at ca. 26km while

**TABLE 2** PERMANOVA tests results using semiquantitative (Bray–Curtis dissimilarity matrix) approaches comparing community composition among the factor depth (only samples in Balsfjord selected).

Factor	df	SS	MS	Pseudo-F	R <sup>2</sup>	p-Value	PERMDISP
Depth (categorical)	2	0.793	0.396	<b>2.090</b>	0.134	0.014	0.127
Residuals	27	5.124	0.189		0.865		
Total	29	5.917			1.000		

Note: PERMDISP probabilities for homogeneity of dispersion are also shown. Significant values are indicated in bold.



**FIGURE 6** Correlation plot of samples' pairwise dissimilarity (a) and the ratio of pairwise dissimilarity over the average dissimilarity per each depth (b) along samples' pairwise horizontal distance. The dissimilarity measured using the Bray–Curtis index based on inverse Wisconsin double-standardization of amplicon reads indicated higher differences on surface communities (a) and average dissimilarity was achieved after 16 km of horizontal distance between samples (b). The standard errors are presented for both plots.

for the surface samples at 40 km (Figure 6a). Subsequently, samples' pairwise dissimilarity were divided with the average dissimilarity per each depth category to empirically measure the relative dissimilarity of samples across the horizontal distance (Figure 6b).

Additionally, the intrinsic and extrinsic variability was diagnosed via violin plots (Figure 7) on two levels, fjord ecosystem (Balsfjord) and regional ecosystem (all fjords together) using the distribution of dissimilarity between samples. The plot indicated that slightly more than half of the samples had intrinsic variability lower than the extrinsic variability within one ecosystem and the intrinsic variability of remaining samples was equal to extrinsic variability of that ecosystem (Figure 7). When looking at all ecosystems together, the intrinsic variability (within each fjord variability) was distinguishably lower than the extrinsic variability (between fjords variability; Figure 7).

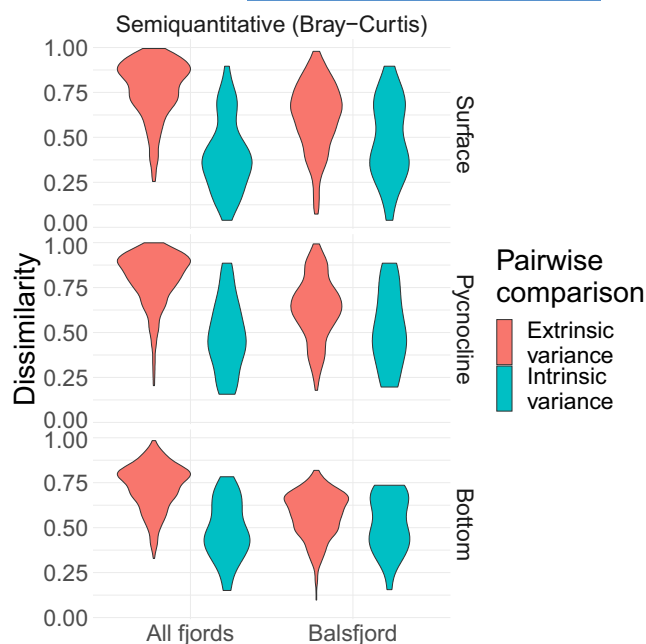
## 4 | DISCUSSION

The sampling design distinguishing fish communities eDNA metabarcoding still poses challenges regarding the optimization of

sampling depth, distance between samples and data treatment approaches. This study aimed to investigate the optimal data treatment approach, sampling effort, and spatial distribution of eDNA metabarcoding samples allowing to effectively detect differences in communities within and between fjords. Our study found that using two sampling stations (two depths, and three biological replicates), 26–40 km apart in Balsfjord, one in Frakkfjord and one in Olderfjord, when using a semiquantitative approach and omitting pycnocline depth samples, was the most efficient design to detect differences between ecosystems given the sampling effort and the fish community dissimilarity within the ecosystem.

### 4.1 | Sequencing and eDNA sampling effort

Exploring the unbalanced sequencing effort curves (Figure 2a) revealed that samples with higher read abundances of *Gadus morhua* and *Mallotus villosus* also had a high sequencing depth. Cumulatively, the sequence read abundance of these two species accounted for 76.6% of the total reads in all samples. Metabarcoding sequence read



**FIGURE 7** Violin plots comparing the intrinsic and extrinsic variability of dissimilarity index (using the Bray–Curtis index based on inverse Wisconsin double-standardization of amplicon reads) on two instances, fjord level (i; Balsfjord, right) and regional level (ii; all fjords, left). In fjord level (i.e., Balsfjord), variability between biological replicates was indicated as intrinsic (blue) and between all samples within the fjord as extrinsic (red). In all fjords, the variability within each fjord was indicated as intrinsic (blue) and between fjords as extrinsic (red).

abundance is a function of species DNA abundance in the samples, together with primer bias (Elbrecht & Leese, 2015). For instance, species with high abundance, but also high amplification efficiency, can result in a skewed distribution (i.e., over-represented high proportion) of sequence read abundances among samples. We consider a similar process to have caused the unbalanced sequencing depth of our localities, as *G. morhua* and *M. villosus* had high sequence read abundance predominantly in Frakkfjord and Olderfjord (the localities with higher sequencing depth; Figure S3). An additional explanation of the unbalanced sequencing effort could be the difference in the aggregation of nontarget taxa (prokaryotic and non-target eukaryotic DNAs) when competing for oligonucleotides during PCR (Díaz et al., 2020; Miya et al., 2020). Despite the nonuniform sequencing depth, the curves have similar shapes among the localities (Figure 2a), suggesting similar species detection rates among distinct fjords. Furthermore, our sequencing depth among distinct localities can be considered adequate for conducting our analysis but not adequate for covering the full diversity of the studied localities. Such limitation of sequencing depth can critically affect the quality of qualitative approach analysis as this approach strongly relies on detecting presence of non-abundant taxa.

The DNA collection curves (Figure S2b) suggested that increasing the number of sampling stations in Balsfjord might uncover more species, indicating that the number of sampling stations deployed in Balsfjord was insufficient for representing the total biodiversity of

this ecosystem. Given the excessive area engaged in Balsfjord, the spatial heterogeneity can be greater than in the other two localities; thus, increasing the sample density can potentially uncover this ecosystem's additional taxonomic complexity.

## 4.2 | Differences between qualitative and semiquantitative approaches

Although we expected the localities to differ in fish community complexity (with respect to species abundances), evidence for differences among locations differed drastically between the qualitative and semiquantitative approaches. Overall, both methods indicated differences in community composition between at least two ecosystems (Table 1). But the semiquantitative approach discriminated samples between all localities (including the positive control; Table S2a). In contrast, the qualitative approach failed to distinguish the samples between Frakkfjord and Olderfjord and to distinguish the positive control samples from the remaining samples in all localities (Table S2b). Focusing on only the presence/absence of species when using eDNA metabarcoding can lead to inconclusive results due to the limitations of this approach (Wang et al., 2021). Below, we discuss how multiple reasons can lead to such limitations and make true biodiversity differences indistinguishable when using a qualitative approach to eDNA metabarcoding.

First, communities with ubiquitous species make it impossible for a qualitative approach (without any quantitative information) to distinguish spatial differences in communities. In our survey, we observed that almost one-third of species were ubiquitous (Table S3). Being present in all the sampling stations, such species cannot contribute to the dissimilarity between localities measured by presence/absence metrics. Hence, less frequent species have higher contributions to the differences between localities. Moreover, rare taxa are more prone to PCR biases than common taxa (Shirazi et al., 2021). Due to the stochasticity of the PCR amplification, rare taxa are detected sporadically within samples, making their reliability low. Such an issue was observed in our positive control samples, where seven taxa were found present only in one sample despite the source of the samples being the same (Figure S4), explaining the high dissimilarity between positive control samples. Larson et al. (2022) found similar results regarding rare taxa and stated that they occurred sporadically among PCR replicates. An additional cause to such patterns could also be explained due to amplification efficiency. Sequence abundance output is subjected to amplification biases due to primer efficiency (Kelly et al., 2019), especially when using universal primers, where selected taxa have a higher match with the primer than others (Banerjee et al., 2021). Such biases influence the commonness of the species, which can lead to divergence from the true species as drivers of differences between ecosystems (since low amplification species will be encouraged to be the species indicators).

Analogously, Cilleros et al. (2019) found metabarcoding results to be less discriminant when using the qualitative approach compared to traditional methods in an attempt to segregate the

faunal composition between large rivers and nearby streams. This is explained due to DNA transport as eDNA metabarcoding samples detected species found distant from the sampling source, thus increasing the homogeneity of the ecosystems (Cilleros et al., 2019). Using a qualitative approach can uprise biases from DNA transport, especially when the survey area is small compared with the DNA transport area (i.e., sampling distance is smaller than eDNA transport) as the DNA of the same origin can be captured in multiple samples simultaneously, thus less discrimination between localities. Lastly, errors arising from tag jumping (index hopping) and potential contamination are nontrivial these issues increase the occurrence of false-positive taxa, contributing to the aforementioned biases when using a qualitative approach. Although we included negative controls during all steps and stated a relatively clean workflows, the average sequence abundance in blanks was 32.2 reads, indicating that false-positive taxa can potentially be present in samples. In an additional exploratory analysis, we accounted for such false-positive by increasing the threshold of occurrence to the highest sequence abundance in negative control samples (minimum number of reads for defining amplicon presence = 81 reads) and found a substantial increase in disparity between ecosystems (Figure S5).

Notwithstanding, the semiquantitative approach does not entirely eliminate all the aforementioned obstacles. Issues such as DNA transport are prevalent regardless of the approach used for data inference. Nevertheless, using a semiquantitative approach can minimize the biases deriving from such matters. The concentration of transported eDNA decreases with increasing distance from the source due to dilution and degradation (Goldberg et al., 2018). Having a high density right at the source and exponentially lower density at neighboring sampling stations can result in relatively lower sequence abundance in the latter samples (Zou et al., 2020). Although many biotic and abiotic factors can potentially affect the displacement and the decay rate of DNA in the water (Harrison et al., 2019; Holman et al., 2021). Andruszkiewicz et al. (2019) predicted an average DNA transport of 5 km with 30% decay rate. Moreover, semiquantitative approaches lower the importance of rare taxa in detecting differences among communities compared with a qualitative approach, as their relative abundances are generally low. This simply alleviates the contribution of stochastic processes (such as PCR amplification of rare taxa) to determine the drivers of community differences. Semiquantitative approaches also minimize the contribution of false positives from tag jumping (typically in low abundance levels) by simply devaluing their importance due to low sequence abundance. Additionally, using the eDNA index as a semiquantitative approach offers an archaic way to take amplification efficiency into consideration (Kelly et al., 2019). Issues regarding ubiquitous taxa can be easily accounted for by using a semiquantitative approach, as differences in abundance will be reflected as contributions to quantitative dissimilarity indices. Although many studies have confirmed, to some extent, the correlation of eDNA metabarcoding sequence abundance with species biomass, more studies are required to come to sound conclusions. Nonetheless, our positive control samples revealed similar sequence

abundance for common species, indicating a fairly robust PCR process with relatively comparable outcomes.

The difference in the two approaches is also reflected during species indicator analysis (Figure S2) as using the semiquantitative approach recognized considerably more species as indicators. Such analysis relies on relative abundances and frequencies of occurrences; hence, by removing the former layer of information (i.e., qualitative approach), it can result in reluctant conclusions. Therefore, having relative abundances as an additional layer of information contributes to finding more species as significant drivers and offers a more realistic and accurate detection of diversity patterns between localities.

Additionally, PERMANOVA on sample removal analysis indicated that the signals of dissimilarity are maintained using the semiquantitative approach (Figure 4), even when the number of sampling stations is reduced. Although reducing the number of sampling stations can have other indirect impacts on additional information, such as the heterogeneity of the ecosystem, we conclude that strong community dissimilarity can be efficiently recognized by implementing a semiquantitative approach. In contrast, it might be masked when a presence/absence approach is used. These findings could result from the small number of fish taxa ( $n=33$ ) present in the relatively low-diversity subarctic fjords and applying similar analyses in more diverse ecosystems might lead to different results. Moreover, having observations from only one season could impair our conclusions, as the seasonal variability might alter the differences between the ecosystems, especially in subarctic ecosystems where the seasonal variability is high mainly due to light conditions. These factors, together with the different lifecycles of fish species (i.e., spawning or feeding cycles), could reshape our conclusions. Moreover, clustering the MOTUs into species (as we did in this study) can diminish the heterogeneity of the study, therefore, enhancing the similarities between ecosystems. Jeunen, Knapp, et al. (2019) found higher dissimilarity when using MOTUs compared with species. They indicated that higher proportion of rare MOTUs can contribute to discrepancies between ecosystems when using a qualitative approach. Although the latter study used a COI marker, such properties can also be inherited by our marker choice.

However, we cross-validated our results with previous studies and management reports conducted on these localities. We found that the fish communities detected in our study matched the species composition recorded in Artsdatabanken (access date: March 2022; <https://artsdatabanken.no/>). Despite that the most abundant species were ubiquitous, similarly to eDNA data, records of *Ciliata mustela* and *Scophthalmus maximus* were only found in Frakkfjord. Moreover, our eDNA index results found similar outcomes, most abundant species as trawl catches in Balsfjord (as part of annual coastal surveys by IMR, unpublished data). Similar matches between eDNA and trawling have been previously documented (Fraija-Fernández et al., 2020). Kiærbech (2017) mentioned a predominant aggregation of *Melanogrammus aeglefinus* and *Pleuronectes platessa* in Balsfjord. Meanwhile *G.morhua* and *Hippoglossoides platessoides* are more abundant in the Olderfjord area. Such differences in



assemblages could be explained by the physical and chemical characteristics of the three fjords (Jo et al., 2019), as Balsfjord has limited water exchange due to closure by a large sill at the entrance of the fjord, whereas Olderfjord and Frakkfjord are more open and have a higher water exchange. Our eDNA results matched those traditional surveys as the eDNA index for *M. aeglefinnus* and *P. platessa* was predominantly aggregated in Balsfjord (Figure S6). Additionally, exceptionally high signals of *M. villosus* were found in Frakkfjord which is highly likely due to the Barents sea population of this species having one of their major spawning grounds on the Fugløy bank just outside the fjord (Alrabeei et al., 2021). Lastly, the innermost part of Balsfjord is a known and regionally important spawning area for Atlantic cod (Aglen et al., 2020), corresponding to high eDNA index in this part of the fjord (Figure S7).

Such consistency indicates that the eDNA index can robustly be used as a semiquantitative assessment of fish communities. Many studies have now confirmed the strengths of eDNA metabarcoding for local detection of species and, thus community composition inference (Hansen et al., 2018). Similarly to analogous studies (Cilleros et al., 2019; Fraija-Fernández et al., 2020; Jeunen, Knapp, et al., 2019; Jo et al., 2019; Larson et al., 2022; Li et al., 2021; Turon et al., 2022), our eDNA metabarcoding study confirm that fish community composition and biodiversity patterns can be reliably estimated using this approach.

### 4.3 | Sampling efficiency for vertical discrepancy and horizontal community dissimilarity

We showed that communities were segregated according to depth (Table 3), and this result confirms previous studies showing that eDNA metabarcoding can resolve vertical assemblages (Closek et al., 2019; Jeunen, Lamare, et al., 2019). However, in our results, the pycnocline community changed insignificantly when compared with bottom and surface samples (Table S2c), indicating a gradual transition of the communities along the depth where the significant difference could be detected only when sampling on the extremes of water column (surface and bottom samples) for our study area. Our findings are utterly aligned with those found in Closek et al. (2019), as they stated no differences in communities above and below the pycnocline samples, and these differences could only start being observed at a minimum depth distance of 80m. This result suggests that removing pycnocline samples would not affect the robustness of eDNA metabarcoding studies when used for monitoring communities in different ecosystems. Although removing pycnocline samples might relieve the sampling effort, we note that such signal sensitivity might differ among various ecosystems and different seasons.

Research on DNA transport has repeatedly found that eDNA is deposited towards the bottom as part of the downwards transport (i.e., vertical settling; Andruszkiewicz et al., 2019; Canals et al., 2021; Turner et al., 2015). Additionally to regular settling, DNA particles can sink through lateral advection dynamics in ice-free fjords (Canals et al., 2021; Wiedmann et al., 2016). Although sedimentation rate

depends on the DNA state (Mauvisseau et al., 2021), studies have found that eDNA is 8–1800 times more concentrated in the bottom sediments than in surface samples (Turner et al., 2015). Canals et al. (2021) observed epipelagic fish eDNA in deep water samples, but not vice versa, concluding that eDNA had been transported downwards. Although it can be argued how environmental factors might affect this phenomenon (Turner et al., 2015), in our study, we observed higher species detectability in samples from the bottom waters, potentially indicating that bottom samples have a wider time span to record species occurrence at the site compared with surface samples which are mainly described as a snapshot of the present community (Díaz et al., 2020).

Among the three surveyed localities, Balsfjord had the highest heterogeneity (Figure 3). We noted that, in Balsfjord, the dissimilarity of communities increased with the distance between samples, indicating a gradual increase in fish species segregation. Coherently with other studies, our findings support the concept that eDNA can be used to measure heterogeneity of ecosystems (Wood et al., 2021). For instance, Fraija-Fernández et al. (2020) found that both methods (i.e., eDNA and trawling) tended to be more different when sampling stations were further apart.

We found significant difference of community dissimilarity in horizontal distance (Figure 6a) for surface and pycnocline communities. This indicated that with increase in spatial spread of sampling stations the dissimilarity increased for the latter communities. However, for bottom communities, these differences were not found to be significant, thus indicating that such communities remained spatially homogeneous although the spatial spreading of sampling stations increased. Such results demonstrate that the bottom layers can deposit DNA and thereby inferring a wider timespan of biological occurrence as mentioned above (Mestre et al., 2018). Subsequently, we noted that the most efficient sampling design—the highest dissimilarity within the ecosystem given the total dissimilarity of the ecosystem—was observed for samples obtained 26 km apart from each other when deploying bottom and pycnocline samples, and 40 km when deploying surface samples.

As no previous studies have been conducted in marine environments regarding spatial changes of dissimilarity, it is challenging to cross-validate our conclusions. However, Zhang et al. (2020) surveyed eDNA sampling design in a lentic system and found a spatial autocorrelation on samples up to 2 km apart. Although marine environments are more complex compared with lentic ones, their finding aligns with the concept of increased heterogeneity with increasing distances between samples. Moreover, Andruszkiewicz et al. (2019) reported that eDNA, on average, spreads until about 10 km after 7 days (when including the settling process) and furthermore estimated that, when in a moderate decay scenario, its source of origin can be <20 km. Incorporating those findings into our research, we have reasons to think that eDNA transport can influence the detection of heterogeneity in metabarcoding studies in ways that we have aforementioned. Nevertheless, we encourage follow-up studies on this topic as it can advance state of the art for eDNA metabarcoding applications in biomonitoring.

In short, we indicated that deploying sampling stations at such distance range from each other would yield the highest community dissimilarity within an ecosystem while avoiding unnecessary sampling. Note that we did not have high statistical power to precisely conclude the homogeneity of samples very close to each other ( $n=3$ ) or very far apart from each other ( $n=4$ ); therefore, we express the need for similar studies with an increased number of sampling stations. Moreover, our study design is limited to only three subarctic fjords with no repetition over the years and no changes in seasonality. Considering our findings, future studies should account for different ecosystems and include different seasons.

With the current rate of biodiversity loss in a climate crisis scenario, the need for more frequent sampling is increasing, therefore maximizing the efficiency of eDNA metabarcoding (i.e., reducing the unnecessary sampling while retaining the ability to detect differences in fish compositions) can enable monitoring agencies to increase their biomonitoring frequency. Reducing the amount of unnecessary sampling and the costs associated while still maintaining robust fish ecological inferences can foster the adaptation of eDNA metabarcoding into biomonitoring of coastal fish and ecosystem-based management.

## 5 | CONCLUSIONS

Qualitative and semiquantitative approaches to analyze eDNA metabarcoding data drastically differ in their efficiency to detect differences in communities between fjords. Thus, recommendations on a suitable number of samples and their spatial distribution to efficiently detect such differences is heavily dependent on the data treatment approach. For this reason, we stress the differences between the two approaches. Our results suggest that a semiquantitative approach has significantly higher efficiency to detect community differences from eDNA metabarcoding compared with a qualitative approach. Finally, our results showed that spreading the sampling stations ~30km apart and avoiding sampling of the pycnocline can be sufficient for capturing fish community differences or changes in subarctic fjords when using a semiquantitative approach.

### AUTHOR CONTRIBUTIONS

TJ, GG, NY, JIW, OSW, RK, JLR, AOS, TH, and KP designed the study. GG, TH, and JIW conducted the fieldwork. GG and TH did the laboratory work. GG, TH, and JIW did the bioinformatic analyses. GG did the statistical analyses. GG with support from TJ, RK, JLR, OSW, AOS, JIW, NY, and KP interpreted the statistical and ecological results. GG wrote the manuscript draft with contributions from all co-authors and all authors approved the submitted version.

### ACKNOWLEDGMENTS

We thank Agneta Hansen for the significant contribution during the field and laboratory work and the crew onboard "F/F Kristine Bonnevie" for technical help with sampling instruments. This research was funded by the project FISHDIV (NRC301691) funded by

the Norwegian Research Council (NRC) and the program "Marine Ecosystem processes" at Institute of Marine Research.

### CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

### DATA AVAILABILITY STATEMENT

The original raw sequences have been deposited in the NCBI SRA archive with accession number PRJNA878662.

### ORCID

Gledis Guri  <https://orcid.org/0000-0002-3711-6758>

Jon-Ivar Westgaard  <https://orcid.org/0000-0002-9627-8124>

Nigel Yoccoz  <https://orcid.org/0000-0003-2192-1039>

Owen S. Wangensteen  <https://orcid.org/0000-0001-5593-348X>

Kim Præbel  <https://orcid.org/0000-0002-0681-1854>

Jessica Louise Ray  <https://orcid.org/0000-0002-7305-5737>

Ryan P. Kelly  <https://orcid.org/0000-0001-5037-2441>

Andrew Olaf Shelton  <https://orcid.org/0000-0002-8045-6141>

Tanja Hanebrekke  <https://orcid.org/0000-0003-3248-1681>

Torild Johansen  <https://orcid.org/0000-0002-9858-052X>

### REFERENCES

- Aglen, A., Nedreaas, K., Knutsen, J. A., & Huse, G. (2020). *Kysttorsk nord for 62-grader nord (Fisken og havet 2020-2)*.
- Alrabeei, S., Subbey, S., Gundersen, S., & Gjøsaeter, H. (2021). Spatial and temporal patterns of capelin (*Mallotus villosus*) spawning sites in the Barents Sea. *Fisheries Research*, 244, 106117. <https://doi.org/10.1016/j.fishres.2021.106117>
- Andres, K. J., Lambert, T. D., Lodge, D. M., Andrés, J., & Jackson, J. R. (2022). Combining sampling gear to optimally inventory species highlights the efficiency of eDNA metabarcoding. *Environmental DNA*, 5(1) 146–157. <https://doi.org/10.1002/edn3.366>
- Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., & Boehm, A. B. (2019). Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. *Frontiers in Marine Science*, 6, 1–14. <https://doi.org/10.3389/fmars.2019.00477>
- Banerjee, P., Dey, G., Antognazza, C. M., Sharma, R. K., Maity, J. P., Chan, M. W., Huang, Y. H., Lin, P. Y., Chao, H. C., Lu, C. M., & Chen, C. Y. (2021). Reinforcement of environmental dna based methods (*Sensu stricto*) in biodiversity monitoring and conservation: A review. *Biology*, 10(12), 1223. <https://doi.org/10.3390/biology10121223>
- Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, 17, 1–17. <https://doi.org/10.1007/s10592-015-0775-4>
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4), 1165–1188. <https://doi.org/10.1214/aos/1013699998>
- Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., Power, M., McLaughlin, J., Edgar, G. J., & Keesing, J. (2020). Maximizing fish detection with eDNA metabarcoding. *Environmental DNA*, 2(4), 493–504. <https://doi.org/10.1002/edn3.74>
- Borja, Á., & Elliott, M. (2013). Marine monitoring during an economic crisis: The cure is worse than the disease. *Marine Pollution Bulletin*, 68(1–2), 1–3. <https://doi.org/10.1016/j.marpolbul.2013.01.041>
- Bourlat, S. J., Borja, A., Gilbert, J., Taylor, M. I., Davies, N., Weisberg, S. B., Griffith, J. F., Lettieri, T., Field, D., Benzie, J., Glöckner, F. O.,

- Rodríguez-Ezpeleta, N., Faith, D. P., Bean, T. P., & Obst, M. (2013). Genomics in marine monitoring: New opportunities for assessing marine health status. *Marine Pollution Bulletin*, 74(1), 19–31. <https://doi.org/10.1016/j.marpolbul.2013.05.042>
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16, 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Buxton, A., Matechou, E., Griffin, J., Diana, A., & Griffiths, R. A. (2021). Optimising sampling and analysis protocols in environmental DNA studies. *Scientific Reports*, 11, 11637. <https://doi.org/10.1038/s41598-021-91166-7>
- Canals, O., Mendibil, I., Santos, M., Irigoien, X., & Rodríguez-Ezpeleta, N. (2021). Vertical stratification of environmental DNA in the open ocean captures ecological patterns and behavior of deep-sea fishes. *Limnology and Oceanography Letters*, 6(6), 339–347. <https://doi.org/10.1002/lol2.10213>
- Cantera, I., Cilleros, K., Valentini, A., Cerdan, A., Dejean, T., Iribar, A., Taberlet, P., Vigouroux, R., & Brosse, S. (2019). Optimizing environmental DNA sampling effort for fish inventories in tropical streams and rivers. *Scientific Reports*, 9, 3085. <https://doi.org/10.1038/s41598-019-39399-5>
- Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., Iribar, A., Taberlet, P., Vigouroux, R., & Brosse, S. (2019). Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. *Molecular Ecology Resources*, 19, 27–46. <https://doi.org/10.1111/1755-0998.12900>
- Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., Bograd, S. J., Hazen, E. L., Field, J. C., & Boehm, A. B. (2019). Marine vertebrate biodiversity and distribution within the central California current using environmental DNA (eDNA) Metabarcoding and ecosystem surveys. *Frontiers in Marine Science*, 6, 1–17. <https://doi.org/10.3389/fmars.2019.00732>
- Cogúic, E., Ershova, E. A., Daase, M., Vonnahme, T. R., Wangensteen, O. S., Gradinger, R., Præbel, K., & Berge, J. (2021). Seasonal variability in the zooplankton community structure in a sub-Arctic Fjord as revealed by morphological and molecular approaches. *Frontiers in Marine Science*, 8, 705042. <https://doi.org/10.3389/fmars.2021.705042>
- Díaz, C., Wege, F. F., Tang, C. Q., Crampton-Platt, A., Rüdél, H., Eilebrecht, E., & Koschorreck, J. (2020). Aquatic suspended particulate matter as source of eDNA for fish metabarcoding. *Scientific Reports*, 10, 14352. <https://doi.org/10.1038/s41598-020-71238-w>
- Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67(3), 345–366. <https://doi.org/10.2307/2963459>
- 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>
- Fraija-Fernández, N., Bouquieaux, M. C., Rey, A., Mendibil, I., Cotano, U., Irigoien, X., Santos, M., & Rodríguez-Ezpeleta, N. (2020). Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution*, 10(14), 7560–7584. <https://doi.org/10.1002/ece3.6482>
- Gilbey, J., Carvalho, G., Castilho, R., Coscia, I., Coulson, M. W., Dahle, G., Derycke, S., Francisco, S. M., & Volckaert, F. A. M. (2021). Life in a drop: Sampling environmental DNA for marine fishery management and ecosystem monitoring. *Marine Policy*, 124, 104331. <https://doi.org/10.1016/j.marpol.2020.104331>
- Gold, Z., Sprague, J., Kushner, D. J., Marin, E. Z., & Barber, P. H. (2021). eDNA metabarcoding as a biomonitoring tool for marine protected areas. *PLoS One*, 16(2), e0238557. <https://doi.org/10.1371/journal.pone.0238557>
- Goldberg, C. S., Strickler, K. M., & Fremier, A. K. (2018). Degradation and dispersion limit environmental DNA detection of rare amphibians in wetlands: Increasing efficacy of sampling designs. *Science of the Total Environment*, 633, 695–703. <https://doi.org/10.1016/j.scitotenv.2018.02.295>
- Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., Spear, S. F., McKee, A., Oyster-McCance, S. J., Cornman, R. S., Laramie, M. B., Mahon, A. R., Lance, R. F., Pilliod, D. S., Strickler, K. M., Waits, L. P., Fremier, A. K., Takahara, T., Herder, J. E., & Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, 7(11), 1299–1307. <https://doi.org/10.1111/2041-210X.12595>
- Gullestad, P., Abotnes, A. M., Bakke, G., Skern-Mauritzen, M., Nedreaas, K., & Søvik, G. (2017). Towards ecosystem-based fisheries management in Norway – Practical tools for keeping track of relevant issues and prioritising management efforts. *Marine Policy*, 77, 104–110. <https://doi.org/10.1016/j.marpol.2016.11.032>
- Hansen, B. K., Bekkevold, D., Clausen, L. W., & Nielsen, E. E. (2018). The sceptical optimist: Challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, 19(5), 751–768. <https://doi.org/10.1111/faf.12286>
- Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, 286(1915), 20191409. <https://doi.org/10.1098/rspb.2019.1409>
- Holman, L. E., Chng, Y., & Rius, M. (2021). How does eDNA decay affect metabarcoding experiments? *Environmental DNA*, 4(1), 108–116. <https://doi.org/10.1002/edn3.201>
- Jeunen, G. J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., Bunce, M., & Gemmill, N. J. (2019). Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Molecular Ecology Resources*, 19(2), 426–438. <https://doi.org/10.1111/1755-0998.12982>
- Jeunen, G. J., Lamare, M. D., Knapp, M., Spencer, H. G., Taylor, H. R., Stat, M., Bunce, M., & Gemmill, N. J. (2019). Water stratification in the marine biome restricts vertical environmental DNA (eDNA) signal dispersal. *Environmental DNA*, 2(1), 99–111. <https://doi.org/10.1002/edn3.49>
- Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., ... Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, 16(10), 1245–1257. <https://doi.org/10.1111/ele.12162>
- Jo, H., Kim, D. K., Park, K., & Kwak, I. S. (2019). Discrimination of spatial distribution of aquatic organisms in a coastal ecosystem using eDNA. *Applied Sciences*, 9(17), 3450. <https://doi.org/10.3390/app9173450>
- Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific Reports*, 9(1), 1–14. <https://doi.org/10.1038/s41598-019-48546-x>
- Kiærbech, K. (2017). *Do diet of long rough dab (Hippoglossoides platessoides) differ between fjords with and without red king crab (Paralithodes camtschaticus)?* MSc Thesis, UiT The Arctic University of Norway, <https://hdl.handle.net/10037/11340>
- Larson, W. A., Barry, P., Dokai, W., Maselko, J., Olson, J., & Baetscher, D. (2022). Leveraging eDNA metabarcoding to characterize nearshore fish communities in Southeast Alaska: Do habitat and tide matter? *Environmental DNA*, 4(4), 868–880. <https://doi.org/10.1002/edn3.297>
- Li, W. P., Liu, Z. F., Guo, T., Chen, H., & Xie, X. (2021). Using optimal environmental dna method to improve the fish diversity survey—From laboratory to aquatic life reserve. *Water*, 13(11), 1468. <https://doi.org/10.3390/w13111468>
- Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., Deiner, K., Lodge, D. M., Lamberti, G. A., & Pfender, M. E. (2018).



- Estimating fish alpha- and beta-diversity along a small stream with environmental DNA metabarcoding. *Metabarcoding and Metagenomics*, 2, e24262. <https://doi.org/10.3897/mbmg.2.24262>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: Robust and fast clustering method for amplicon-based studies. *PeerJ*, 2, e593. <https://doi.org/10.7717/peerj.593>
- Mauvisseau, Q., Harper, L., Sander, M., Hanner, R. H., Kleyer, H., & Deiner, K. (2021). The multiple states of environmental DNA and what is known about their persistence in aquatic environments. *Environmental Science & Technology*, 56(9), 5322–5333. <https://doi.org/10.1021/acs.est.1c07638>
- Mestre, M., Ruiz-González, C., Logares, R., Duarte, C. M., Gasol, J. M., & Sala, M. M. (2018). Sinking particles promote vertical connectivity in the ocean microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, 115(29), E6799–E6807. <https://doi.org/10.1073/pnas.1802470115>
- Miya, M., Gotoh, R. O., & Sado, T. (2020). MiFish metabarcoding: A high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fisheries Science*, 86(6), 939–970. <https://doi.org/10.1007/s12562-020-01461-x>
- 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>
- Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., & Miya, M. (2018). Comparing local- and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshwater Biology*, 63(6), 569–580. <https://doi.org/10.1111/fwb.13094>
- Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Weedon, J. (2022). *Vegan: Community ecology package*. <https://cran.r-project.org/package=vegan>
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reigstad, M., & Wassmann, P. (1996). Importance of advection for pelagic-benthic coupling in north Norwegian fjords. *Sarsia*, 80(4), 245–257. <https://doi.org/10.1080/00364827.1996.10413599>
- Roberts, D. W. (2019). *labdsv: Ordination and multivariate analysis for ecology*. <https://cran.r-project.org/package=labdsv>
- 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>
- Sales, N. G., Wangensteen, O. S., Carvalho, D. C., Deiner, K., Præbel, K., Coscia, I., McDevitt, A. D., & Mariani, S. (2021). Space-time dynamics in monitoring neotropical fish communities using eDNA metabarcoding. *Science of the Total Environment*, 754, 142096. <https://doi.org/10.1016/j.scitotenv.2020.142096>
- Salter, I., Joensen, M., Kristiansen, R., Steingrund, P., & Vestergaard, P. (2019). Environmental DNA concentrations are correlated with regional biomass of Atlantic cod in oceanic waters. *Communications Biology*, 2(1), 1–9. <https://doi.org/10.1038/s42003-019-0696-8>
- Sanches, T. M., & Schreier, A. D. (2020). Optimizing an eDNA protocol for estuarine environments: Balancing sensitivity, cost and time. *PLoS One*, 15(5), 1–18. <https://doi.org/10.1371/journal.pone.0233522>
- Shirazi, S., Meyer, R. S., & Shapiro, B. (2021). Revisiting the effect of PCR replication and sequencing depth on biodiversity metrics in environmental DNA metabarcoding. *Ecology and Evolution*, 11(22), 15766–15779. <https://doi.org/10.1002/ece3.8239>
- Shu, L., Ludwig, A., & Peng, Z. (2020). Standards for methods utilizing environmental dna for detection of fish species. *Genes*, 11(3), 296. <https://doi.org/10.3390/genes11030296>
- Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S., & Bunce, M. (2017). Ecosystem bio-monitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, 7, 12240. <https://doi.org/10.1038/s41598-017-12501-5>
- Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, 21(8), 1789–1793. <https://doi.org/10.1111/j.1365-294X.2012.05542.x>
- Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, 183, 93–102. <https://doi.org/10.1016/j.biocon.2014.11.017>
- Turon, M., Nygaard, M., Guri, G., Wangensteen, O. S., & Præbel, K. (2022). Fine-scale differences in eukaryotic communities inside and outside salmon aquaculture cages revealed by eDNA metabarcoding. *Frontiers in Genetics*, 13, 957251. <https://doi.org/10.3389/fgene.2022.957251>
- Wang, S., Yan, Z., Hänfling, B., Zheng, X., Wang, P., Fan, J., & Li, J. (2021). Methodology of fish eDNA and its applications in ecology and environment. *Science of the Total Environment*, 755, 142622. <https://doi.org/10.1016/j.scitotenv.2020.142622>
- Wiedmann, I., Reigstad, M., Marquardt, M., Vader, A., & Gabrielsen, T. M. (2016). Seasonality of vertical flux and sinking particle characteristics in an ice-free high arctic fjord-different from subarctic fjords? *Journal of Marine Systems*, 154, 192–205. <https://doi.org/10.1016/j.jmarsys.2015.10.003>
- Wood, Z. T., Lacoursière-Roussel, A., LeBlanc, F., Trudel, M., Kinnison, M. T., Garry McBrine, C., Pavey, S. A., & Gagné, N. (2021). Spatial heterogeneity of eDNA transport improves stream assessment of threatened Salmon presence, abundance, and location. *Frontiers in Ecology and Evolution*, 9, 650717. <https://doi.org/10.3389/fevo.2021.650717>
- Yoccoz, N. G., Nichols, J. D., & Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution*, 16(8), 446–453. [https://doi.org/10.1016/S0169-5347\(01\)02205-4](https://doi.org/10.1016/S0169-5347(01)02205-4)
- Zhang, S., Lu, Q., Wang, Y., Wang, X., Zhao, J., & Yao, M. (2020). Assessment of fish communities using environmental DNA: Effect of spatial sampling design in lentic systems of different sizes. *Molecular Ecology Resources*, 20(1), 242–255. <https://doi.org/10.1111/1755-0998.13105>
- Zou, K., Chen, J., Ruan, H., Li, Z., Guo, W., Li, M., & Liu, L. (2020). eDNA metabarcoding as a promising conservation tool for monitoring fish diversity in a coastal wetland of the Pearl River estuary compared to bottom trawling. *Science of the Total Environment*, 702, 134704. <https://doi.org/10.1016/j.scitotenv.2019.134704>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Guri, G., Westgaard, J.-I., Yoccoz, N., Wangensteen, O. S., Præbel, K., Ray, J. L., Kelly, R. P., Shelton, A. O., Hanebrekke, T., & Johansen, T. (2024). Maximizing sampling efficiency to detect differences in fish community composition using environmental DNA metabarcoding in subarctic fjords. *Environmental DNA*, 6, e409. <https://doi.org/10.1002/edn3.409>



## Paper 2

**Quantifying the detection sensitivity and precision of qPCR and ddPCR mechanisms for eDNA samples (under review).**

Gledis Guri<sup>1,2\*</sup>, Jessica Louise Ray<sup>3</sup>, Andrew Olaf Shelton<sup>4</sup>, Ryan P. Kelly<sup>5</sup>, Kim Præbel<sup>2</sup>, Elizabeth Andruszkiewicz Allan<sup>5</sup>, Nigel Yoccoz<sup>6</sup>, Torild Johansen<sup>1</sup>, Owen S. Wangensteen<sup>2,7</sup>, Tanja Hanebrekke<sup>1</sup>, Jon-Ivar Westgaard<sup>1</sup>

1. Norwegian Institute of Marine Research, Framsenteret, Tromsø, Norway
2. Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway
3. NORCE Norwegian Research Centre AS, Uni Research Environment, N-5020 Bergen, Norway
4. Northwest Fisheries Science Center, NOAA Fisheries, Seattle, Washington, USA
5. School of Marine and Environmental Affairs, University of Washington, Seattle, Washington, USA
6. Dept. of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway
7. Dept. of Evolutionary Biology, Ecology & Environmental Sciences, Universitat de Barcelona, Barcelona, Spain

\*Corresponding author: Gledis Guri, gledis.guri@hi.no, +46 700658104



1 **Abstract**

- 2 1. Environmental DNA (eDNA) detection employing quantitative PCR (qPCR) and droplet digital  
3 PCR (ddPCR) offers a non-invasive and efficient approach for monitoring aquatic organisms.  
4 Accurate and sensitive quantification of eDNA is crucial for tracking rare and invasive species,  
5 and for understanding the biodiversity, abundance, and distribution of aquatic organisms.
- 6 2. This study assesses the sensitivity and quantification precision of quantitative PCR (qPCR) and  
7 droplet digital PCR (ddPCR) for eDNA surveys through Bayesian inference, using latent  
8 parameters from both standards of known concentration and environmental samples across  
9 three teleost fish species assays.
- 10 3. The results showed that ddPCR offers higher sensitivity (detection probability) and  
11 quantification precision (lower variance) than qPCR particularly at low DNA concentrations  
12 ( $< 1$  copy/ $\mu$ L). These findings highlight the superior performance of ddPCR for eDNA  
13 detection at low concentrations and guide researchers toward more reliable methods for  
14 effective species monitoring.
- 15 4. We found that a two-step (detection and concentration) model increased the precision of qPCR  
16 results.

17 **Introduction**

18 Environmental DNA (eDNA) - DNA that is released into the environment by living organisms through  
19 various means, such as skin shedding, faeces, urine, and mucus secretion - has emerged as a powerful  
20 tool for detecting and monitoring aquatic organisms (Taberlet et al., 2012). The detection of eDNA in  
21 water or soil samples can provide a non-invasive and efficient way to detect the presence of aquatic  
22 organisms (Wood et al., 2021), including rare, cryptic, or invasive species (Keller et al., 2022). Accurate  
23 quantification of eDNA is essential for reliable interpretation of ecological data such as population  
24 monitoring, biodiversity assessment and species detection for invasive species monitoring (Sassoubre  
25 et al., 2016). The most common used methods for eDNA quantification in marine and lentic and lotic  
26 systems are quantitative PCR (qPCR) and droplet digital PCR (ddPCR). These quantitation techniques  
27 have diverse applications such as clinical microbiology and environmental DNA studies. In clinical  
28 microbiology, qPCR and ddPCR can be used to detect and quantify specific pathogens, aiding in the  
29 diagnosis and treatment of diseases. In environmental DNA studies, these techniques can assist in  
30 detecting biodiversity changes, tracking the spread of invasive species or monitoring the abundance of  
31 endangered species. Here, we compare these two methods with regard to (1) detection sensitivity, and  
32 (2) quantification precision using laboratory standards and environmental samples.

33  
34 Quantitative PCR (qPCR) with hydrolysis probes uses both template-specific oligonucleotide primers  
35 and a fluorescently labelled oligonucleotide probe, all of which bind specifically to the DNA target.  
36 The amount of fluorescence that accumulates in real-time during the PCR amplification process is  
37 proportional to the amount of DNA target present in the sample. During PCR, the cycle threshold value  
38 (Ct; cycle number at which the fluorescence signal generated by the probe in the PCR reaction reaches  
39 a threshold above background levels) is determined, providing an indirect measure of the amount of  
40 target DNA in the sample. For quantification of a DNA target from unknown samples, qPCR relies on  
41 parallel analysis of standard samples with known concentrations of DNA fragments as reference points  
42 (S. C. Taylor et al., 2015) typically consisting of synthetic DNA fragments that match the target DNA  
43 sequence. A standard curve relating Ct values to incremental changes in DNA concentration is  
44 calculated, and quantification of unknown samples is achieved by extrapolating the amount of  
45 fluorescence emitted by the unknown samples (or indirectly measured as Ct values) to the  
46 corresponding nominal concentration on the standard curve. qPCR is highly sensitive and specific, but  
47 heavily dependent on the standard curves which are themselves subject to substantial technical variation  
48 due to pipetting error.

49  
50 Droplet digital PCR (ddPCR) is a DNA quantitation technique that involves partitioning a sample into  
51 oil-encapsulated nano droplets, each of which constitutes of independent PCR reaction. Partitioning  
52 aims to create a large population of nano droplets with a predictable distribution of droplets containing  
53 zero, one or several copies of the target DNA. Amplification of the target DNA fragment within each

54 individual droplet during PCR results in fluorescence signal accumulation at the end of PCR, which can  
55 then be measured in all individual droplets using microfluidic droplet detection. Dependent upon the  
56 magnitude of fluorescence signal (measured at the end of the PCR), each droplet is then counted as  
57 either a positive or a negative detection event for the target DNA. Template DNA quantity is estimated  
58 based on the proportion of the positive droplets relative to the total number of droplets (B. J. Hindson  
59 et al., 2011), with no standard curve necessary. Through this mechanism ddPCR is absolute and non-  
60 reliant on standard curves as the quantification relies on DNA distribution in droplets (S. C. Taylor et  
61 al., 2015).

62

63 Both methods measure concentration, and so results of the two methods are strongly and linearly  
64 correlated (Campomenosi et al., 2016; Jerde et al., 2016; Nathan et al., 2014; Verhaegen et al., 2016).  
65 However, ddPCR has shown higher sensitivity (detection probability) in clinical microbiology  
66 (Campomenosi et al., 2016; S. C. Taylor et al., 2015) and environmental molecular ecology (Jerde et  
67 al., 2016), and ddPCR consistently estimates template concentration more precisely than qPCR (C. M.  
68 Hindson et al., 2013; S. C. Taylor et al., 2015). This pronounced difference has been attributed primarily  
69 to ddPCR being less susceptible to inhibition than qPCR (Mahendran et al., 2020). Several studies have  
70 attempted to circumvent inhibition of qPCR by modifying reaction chemistry. For example, C. M.  
71 Hindson et al. (2013) used ABI and Bio-Rad Master Mix for qPCR, while Doi et al. (2015) used  
72 Environmental and Universal Master Mix containing AmpliTaq DNA polymerase variants that are less  
73 prone to inhibitors to improve sensitivity and concentration accuracy. Additionally, S. C. Taylor et al.  
74 (2015) delved into the differences within and between qPCR reactions as part of assay optimization  
75 effort. Although substantial attention has been given to laboratory techniques and chemical solutions to  
76 enhance sensitivity and accuracy by these methods, particularly for qPCR, little emphasis has been  
77 placed on developing mathematical approaches to provide higher sensitivity and quantitative precision  
78 and accuracy. Although technical replication is often used to gauge performance sensitivity and  
79 precision, it is often under-utilized for enhancing measurement accuracy (C. M. Hindson et al., 2013).

80

81 The aim of this study is to empirically assess the sensitivity, and quantification precision, of qPCR and  
82 ddPCR quantification methods using improved concentration estimation formulas applied to eDNA  
83 samples. In the case of qPCR, we adopt the framework used in McCall et al. (2014) and Shelton et al.  
84 (2022), while for ddPCR, we refine quantity estimations via a model that makes use of the relationship  
85 between the binomial and Poisson distributions (B. J. Hindson et al., 2011). We present a Bayesian  
86 inference approach designed to independently model DNA quantities for qPCR and ddPCR using  
87 samples with known DNA concentrations (standards, ranging from  $10^{-3}$  -  $10^4$  copies/ $\mu$ L) alongside  
88 environmental samples for three different teleost fish assays. In addition, this study aims to identify  
89 limitations or biases associated with new formulas designed for enhanced concentration estimation, and  
90 to provide recommendations for optimizing quantitative eDNA protocols in future studies.

## 91 **Methods**

### 92 **eDNA water sampling, filtration, extraction.**

93 We surveyed a total of twelve and eight distinct sampling locations over the span of three and two years  
94 in Balsfjord and Frakkfjord, respectively, on the R/V Kristine Bonnevie, as part of Norwegian coastal  
95 surveys. For each sample, we filtered 2 L of seawater. The filters were sealed into sterile 50 ml Falcon  
96 centrifuge tubes and stored at -20 °C until transported to laboratory facilities and preserved at -80 °C  
97 until extraction. For more details regarding these workflows see Guri et al. (2023). Negative controls  
98 were taken at every station to indicate potential contamination. We used DNeasy PowerWater Sterivex  
99 Kit (Qiagen GmbH) to extract the eDNA as described in (Guri et al., 2023)

100

### 101 **Standard samples**

102 The DNA from tissue samples of *Gadus morhua* (Atlantic cod, hereafter “cod”), *Clupea harengus*  
103 (Atlantic herring, hereafter “herring”), and *Pollachius virens* (saithe) was extracted using the DNeasy  
104 Blood & Tissue Kit (Qiagen GmbH, Hilden Germany). To create standard samples, we amplified the  
105 103-bp region of the ATPase and cytochrome b regions for cod and saithe, respectively, using specific  
106 primers (Table 1) with the thermocycler program: an initial 5 min at 95 °C, followed with 40 cycles of  
107 1 min at 95 °C, 1 min at 62 °C, and 1 min at 72 °C, and a final phase at 72 °C for 10 min. The reactions  
108 were run on SimpliAmp Thermal Cycler machine (Thermo Fisher Scientific). We performed a total of  
109 24 reaction (8 per each species), each in 25 µL volume including 12.5 µL of TaqMan Environmental  
110 Master Mix 2.0, 1.25 µL of each of forward and reverse primers (10 nM concentration each), 8 µL of  
111 RNA-free water and 2 µL of genomic DNA template for the above-listed species. The PCR products  
112 were thereafter examined by gel electrophoresis, and replicate reactions with the expected product size  
113 were pooled for subsequent purification, wherein primer-dimers were eliminated using the MinElute  
114 PCR Purification Kit (Qiagen GmbH, Hilden Germany). The DNA concentration in the pooled samples  
115 was measured using Qubit quantification system using the High Sensitivity dsDNA assay. The samples  
116 were converted from mass per volume (ng/µL) to copies per volume (copies/µL) using the fragment  
117 size of each marker. Subsequently, samples underwent a series of 10-fold dilutions to achieve final  
118 concentrations ranging from  $10^{-3}$  –  $10^6$  copies/µL (nine orders of magnitude).

119

### 120 **qPCR**

121 Prior to running all the samples (environmental n = 1010 and standard samples n = 247), assay  
122 optimization was conducted to achieve satisfactory amplification efficiency (see Fig. S1). All qPCR  
123 samples were analysed on an Applied Biosystems 7500 Fast real-time PCR System machine (Thermo  
124 Fisher Scientific) using the thermocycler protocol of 10 min at 95 °C for denaturation followed by the  
125 cycling stage of 42 – 52 cycles of 15 sec at 95 °C and 1 min at 58 °C. Cod assays were duplexed with  
126 either herring or saithe assays for samples in Balsfjord or Frakkfjord, respectively. 5'-hydrolysis probes



127 were labelled with 6-FAM for cod detection, JOE for herring detection or Cy3 for saithe detection  
 128 (Table 1), and all probes were modified at their 3'-end with the quencher moiety BHQ1. All  
 129 thermocycler reactions were run in 20  $\mu\text{L}$  volume consisting of 10  $\mu\text{L}$  of TaqMan Environmental Master  
 130 Mix 2.0, 1  $\mu\text{L}$  of each primer (forward and reverse, 10 nM concentration each; see Table 1), 0.5  $\mu\text{L}$  of  
 131 probe (10nM concentration), 3  $\mu\text{L}$  of dH<sub>2</sub>O and 2  $\mu\text{L}$  of DNA template. Negative controls were run  
 132 together with the samples. Duplexed standard dilution series containing  $10^{-1} - 10^5$  copies  $\mu\text{L}^{-1}$  of  
 133 purified target fragments were included in all qPCR plates to generate standard curves and verify  
 134 performance consistency between qPCR runs.

135

### 136 ddPCR

137 Prior to running the environmental samples, ddPCR assays were optimized by testing different  
 138 primer/probe concentrations and annealing/elongation temperature gradients to identify the conditions  
 139 that resulted in the highest fluorescence difference between positive and negative droplets. Herring (6-  
 140 FAM) or saithe (6-FAM) assays were run in duplexed reactions with the cod (VIC) assay on samples  
 141 from Balsfjord or Frakkfjord, respectively. Duplex cod/saithe ddPCR reactions consisted of 11  $\mu\text{L}$  of  
 142 ddPCR Supermix with no dUTP (Bio-Rad), 11.9 pmol of each primer, 3.5 pmol of probe, 0.04  $\mu\text{L}$  of  
 143 RNA-free water, and 5.5  $\mu\text{L}$  of DNA template. For duplex cod/herring assays we used similar volumes  
 144 of Supermix, DNA template, and cod primers and probe, and for herring we used 4.4 pmol of forward  
 145 primer, 1.32 pmol of reverse primer, 4.4 pmol of probe and 2.35  $\mu\text{L}$  of RNA-free water. Total volume  
 146 of prepared for all ddPCR reactions was 22  $\mu\text{L}$ , from which 20  $\mu\text{L}$  were pipetted to the reaction well to  
 147 ensure volume precision. Droplets were generated according to manufacturer instructions, aiming for  
 148 20,000 droplets per reaction. Emulsion PCR was performed in a C1000 Touch Thermal Cycler with 96-  
 149 Deep Well Reaction Module (Bio-Rad) with a program of 10 min at 95°C, 44 cycles for 1 min at 95°C  
 150 and 2 min at 55.6°C, with a ramp rate of 2°C per s, followed by 10 min at 98°C and stored at 4°C. Room  
 151 temperature-equilibrated ddPCR plates were then analysed using a QX200 droplet reader (Bio-Rad).  
 152 Additional runs consisted of duplexed standard samples of nominal concentration ranging from  $10^{-3} -$   
 153  $10^4$  copies/ $\mu\text{L}$  using the same protocol and amplification program.

154

155 *Table 1. Sequences for qPCR and ddPCR assays targeting 103-bp region of the ATPase gene of Atlantic*  
 156 *cod (Gadus morhua) and saithe (Pollachius virens), and cytochrome b sequence of Atlantic herring*  
 157 *(Clupea harengus). All gene regions belong to the mitochondrial DNA.*

Target	Primers	Sequence	Dye
<i>Gadus morhua</i>	Forward	GAD-FII	GCAATCGAGTYGTATCYCTHCAAGGAT
	Reverse	GAD-R III	GCAAGWAGYGGHGCRCADTTGTG
	qPCR probe	Custom	CTTTTTACCTCTAAATGTGGGAGG
			(M. I. Taylor et al., 2002) (Nash et al., 2012) FAM

	ddPCR probe	Custom	CTTTTACCTCTAAATGTGGGAGG	VIC
	Forward	Cluhar_CYBF14928	CCCATTGTGATTGCAGGGG	(Knudsen et al., 2019)
	Reverse	Cluhar_CYBR15013	CTGAGTTAAGTCCTGCCGGG	(Knudsen et al., 2019)
<i>Clupea harengus</i>	qPCR probe	Cluhar_CYBP14949	TACTATTCTCCACCTTCTGTTTCCTC	JOE
	ddPCR probe	Cluhar_CYBP14949	TACTATTCTCCACCTTCTGTTTCCTC	FAM
	Forward	Saithe-F	GAATCCCAATAATTTAATAGCCT	Unpublished/Johansen et al. 2018
<i>Pollachius virens</i>	Reverse	Saithe-R	TCGATTGCTTAGTCATCGAGA	Unpublished/Johansen et al. 2018
	qPCR probe	Custom	TGATTACTCATCCCTACG	Cy3
	ddPCR probe	Custom	TGATTACTCATCCCTACG	FAM

158

## 159 Bayesian approach and statistics

### 160 *Detection probability (i.e., sensitivity) between methods*

161 We established the relationship between the starting DNA concentration and the probability of positive  
 162 target detection for both methods (qPCR and ddPCR) independently using standard samples of known  
 163 concentration by using the logistic regression model:

$$164 Z_{ik} \sim \text{Bernoulli}(\theta_{ik}) \quad (1.1)$$

$$165 \text{logit}(\theta_{ik}) = \varphi_0 + (\varphi_1 \times S_{ik}) \quad (1.2)$$

166 where Z is the binary outcome for sample aliquot (k) of a target (i) being present or absent following a  
 167 Bernoulli distribution given the probability of detection  $\theta$ . The parameters  $\varphi_0$  and  $\varphi$  indicate the  
 168 intercept and the slope respectively in the logistic function with the  $\log_{10}$  nominal DNA concentration  
 169 in copies/ $\mu\text{L}$  (S) as the independent variable. Subsequently, we compared the differences in  
 170 probabilities of detection ( $\theta_{ik}$ ) between the two methods across the range of tested concentrations.

171

### 172 *Quantity precision estimation between methods*

173 To estimate qPCR-modelled concentrations, we employed the approach formulated by McCall et al.  
 174 (2014) and Shelton et al. (2022). In short, the model combines the use of a detection probability model  
 175 (based on the technical replicate of samples, equation 1.1 – 1.2) and the continuous model (equation  
 176 1.3 – 1.5) jointly, to evaluate the starting DNA concentration. To estimate ddPCR-modelled  
 177 concentration, we used Poisson statistics (Hindson et al., 2013; equation 3) in the form of cloglog  
 178 transformation (equation 4) as follows:

$$179 C = \frac{-\ln(1 - \omega)}{V} \quad (3)$$

$$180 \ln(C) = \text{cloglog}(\omega) + \ln(V) \quad (4)$$

181 where C is the starting DNA concentration in the sample in copies/ $\mu$ L,  $\omega$  is the proportion of positive  
 182 droplets, and V is the ddPCR oil droplet volume. We assume that the number of positive droplets (W)  
 183 follows a binomial distribution with probability of success ( $\omega$ ) given the total number of droplets  
 184 generated (U; equation 2.1). Following equation 4 we establish a linear relationship between  $\omega$  and  
 185 starting DNA concentration ( $C_d$ ; or S for nominal DNA concentration) using an intercept and a slope  
 186  $\kappa_0$  and  $\kappa_1$  respectively (equation 2.2). In our formulation, we aggregate the ratio of positive droplets  
 187 from all the technical replicates of a sample (see more in discussion).  
 188

	Known concentrations (standard samples)	Unknown concentrations (environmental samples)
	$Z_{ik} \sim \text{Bernoulli}(\theta_{ik})(1.1.1)$	$Z_{ijk} \sim \text{Bernoulli}(\theta_{ij})(1.1.2)$
	$\text{logit}(\theta_{ik}) = \varphi_0_i + (\varphi_1_i \times S_{ik})(1.2.1)$	$\text{logit}(\theta_{ij}) = \varphi_0_i + (\varphi_1_i \times C_{qij})(1.2.2)$
qPCR model	$Y_{ik} \sim \text{Normal}(\mu_{ik}, \sigma_i) \text{ if } Z_{ik} = 1(1.3.1)$	$Y_{ijk} \sim \text{Normal}(\mu_{ij}, \sigma_i) \text{ if } Z_{ijk} = 1(1.3.2)$
	$\mu_{ik} = \beta_0_i + (\beta_1_i \times S_{ik})(1.4.1)$	$\mu_{ij} = \beta_0_i + (\beta_1_i \times C_{qij})(1.4.2)$
	$\sigma_i = e^{(\gamma_0 + (\gamma_1 \times S_{ik}))}(1.5.1)$	$\sigma_i = e^{(\gamma_0 + (\gamma_1 \times C_{qij}))}(1.5.2)$
ddPCR model	$W_{ik} \sim \text{Binomial}(\omega_{ik}, U_{ik})(2.1.1)$	$W_{ijk} \sim \text{Binomial}(\omega_{ik}, U_{ik})(2.1.2)$
	$\text{cloglog}(\omega_{ik}) = \kappa_0_i + (\kappa_1_i \times S_{ik})(2.2.1)$	$\text{cloglog}(\omega_{ij}) = \kappa_0_i + (\kappa_1_i \times C_{dij})(2.2.2)$

189 We denote S and C as the  $\log_{10}$  nominal concentration and unknown concentration from standard  
 190 samples and environmental samples respectively where i, j and k indicate species (target assay), eDNA  
 191 sample (eDNA filter), and sample aliquot (technical replicate).  
 192

193  
 194 The two parts of the qPCR model (equation 1.1 – 1.5) were run jointly, but independently from the  
 195 ddPCR model (equation 2.1 – 2.2). To assess the robustness of the methods studied, we calculated and  
 196 compared the uncertainty of each model for eDNA quantification. We determined the uncertainty  
 197 (quantification precision) as the difference between the 2.5 and 97.5 % credible intervals of the  
 198 modelled concentration ( $C_{qij}$  and  $C_{dij}$  for qPCR and ddPCR, respectively).  
 199

200 Furthermore, we explored the potential benefits of incorporating the detection probability model  
 201 (equation 1.1 - 1.2) alongside the continuous model (equation 1.3 - 1.5) for improving the quantification  
 202 precision and accuracy (the latter it is not estimated in this article) of qPCR. To achieve this, we ran the  
 203 continuous model independently and in conjunction with the detection probability model, measuring

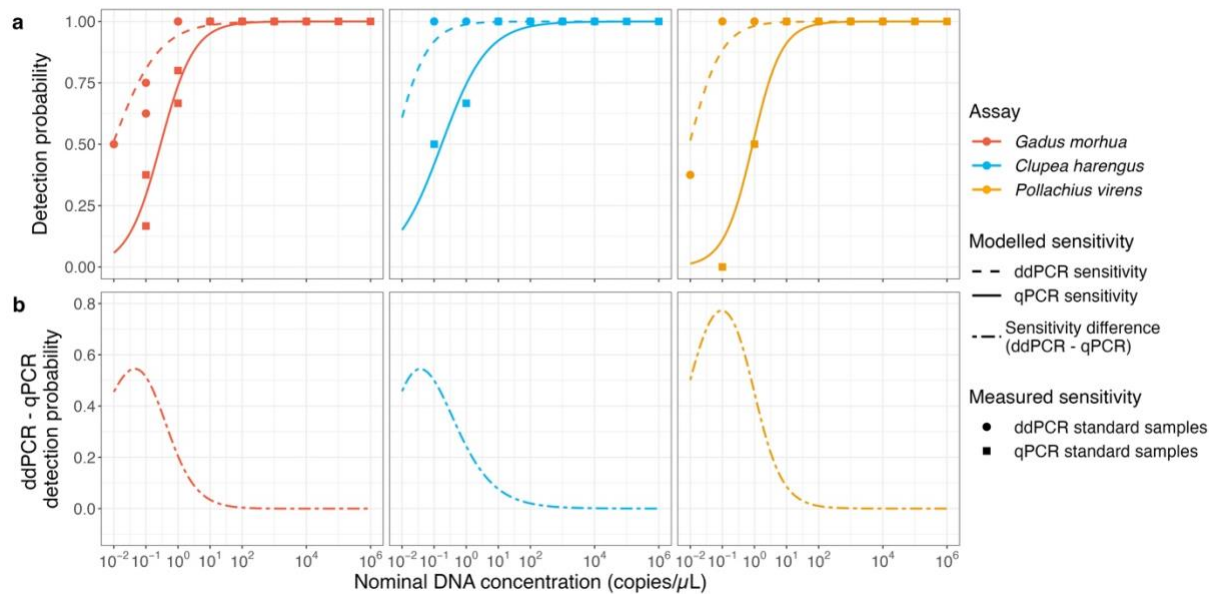
204 the variance in each scenario. Subsequently, we compared the precision variances to assess the  
 205 enhanced value introduced by the detection probability model.

206

207 **Results**

208 **Sensitivity of methods**

209 All the DNA concentrations hereafter are expressed per 20  $\mu\text{L}$  reaction volume, thus 1 DNA copy in a  
 210 20  $\mu\text{L}$  reaction = 0.05 copies/ $\mu\text{L}$  reaction volume (ca.  $10^{-1.3}$  copies/ $\mu\text{L}$ ; see also Fig. 4). The sensitivity  
 211 for ddPCR standard samples showed on average a 50% detection probability at  $10^{-2}$  copies/ $\mu\text{L}$  for all  
 212 assays (Fig. 1a). Conversely, the qPCR results indicated a lower probability of detection compared to  
 213 ddPCR. Specifically, for the cod and herring assays, the 50% detection probability was observed at a  
 214 concentration of ca.  $10^{-0.7}$  copies/ $\mu\text{L}$ , while for the saithe assay, it was indicated at  $10^0$  copies/ $\mu\text{L}$  for  
 215 qPCR assays (Fig. 1a). A comparison of detection probabilities for the two methods revealed that  
 216 ddPCR had a higher probability of detection for concentrations ranging from  $10^{-2}$  to  $10^0$  copies/ $\mu\text{L}$  for  
 217 all target species (Fig. 1b). Notably, when deducting the qPCR detection probability from that of  
 218 ddPCR, cod and herring assay of ddPCR exhibited a 60% higher detection probability at DNA  
 219 concentrations of  $10^{-1}$ , while saithe showed 80% higher detection probability when compared to qPCR  
 220 (Fig. 1b).

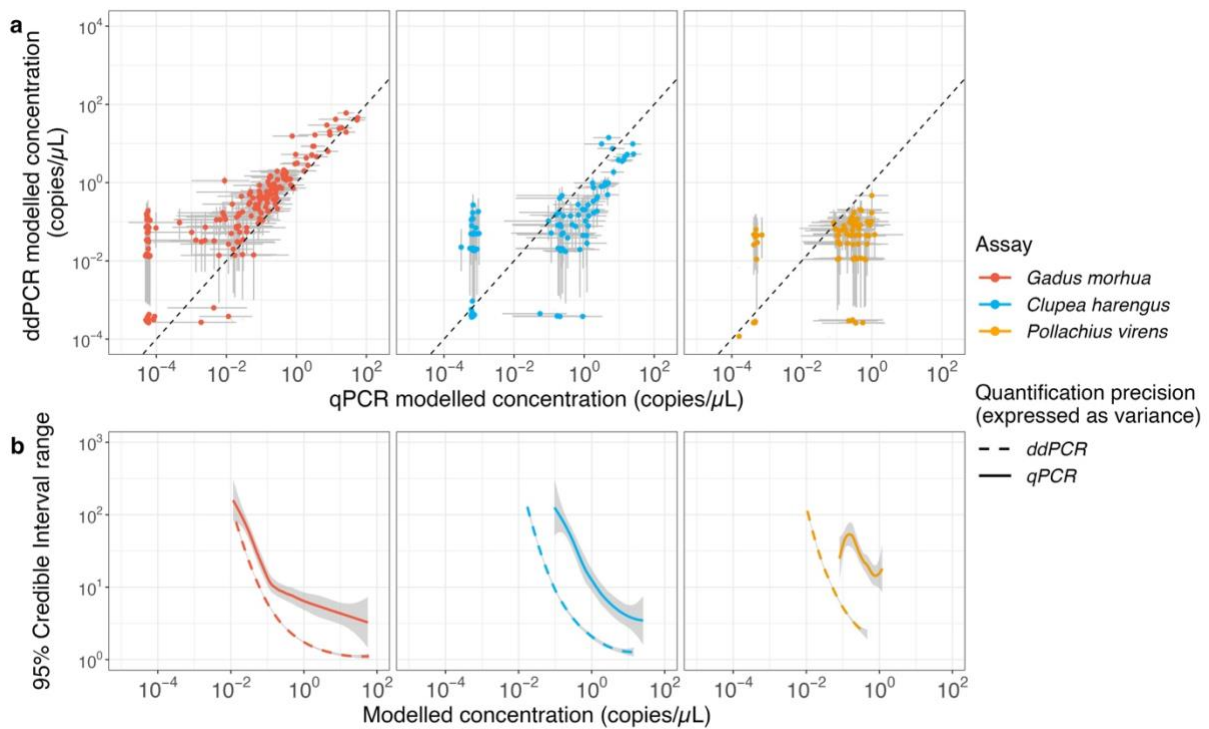


221

222 *Figure 1. Sensitivity of (a) ddPCR and qPCR for three assays (cod = red, herring = blue, and saithe =*  
 223 *orange) shown as modelled detection probability as a function of nominal DNA concentration.*  
 224 *Difference of detection probability (b) between ddPCR and qPCR is also shown over the eDNA*  
 225 *concentration where positive values indicate higher ddPCR detection probability and zero indicates*  
 226 *similar detection probabilities between the two employed methods.*

227 **Quantification precision of methods**

228 qPCR and ddPCR model estimates of starting DNA concentration in environmental samples were  
 229 positively correlated (Fig. 2a). The qPCR model slightly underestimated the starting DNA  
 230 concentration ( $C_{qij}$  from equation 1.2 and 1.4) for cod detection relative to ddPCR model, while it  
 231 overestimated starting DNA concentrations for herring and saithe detection. Consistent with the  
 232 sensitivity analysis above, ddPCR yielded fewer negative detections from environmental samples  
 233 ( $n=38$ ) compared to qPCR ( $n=86$ ). Both methods estimated fewer than  $10^{-3}$  target copies/ $\mu\text{L}$  for all  
 234 negative samples. Furthermore, the qPCR model exhibited a wider credible interval range (lower  
 235 precision for estimating starting DNA concentration) for all assays and template concentrations (Fig.  
 236 2b). The qPCR variability in estimated starting DNA concentration was inversely correlated with DNA  
 237 concentration, indicating a positive relationship between precision and target concentration. For the cod  
 238 assay in particular, the variability was indicated to be one order of magnitude, while for herring and  
 239 saithe, it was approximately two orders of magnitude (see the difference between dashed line and dotted  
 240 line in Fig. 2b).



241  
 242 *Figure 2. Comparison of modelled quantities between qPCR and ddPCR and their credible interval*  
 243 *(grey bars) for three assays (cod = red, herring = blue, and saithe = orange; a). Concentration below*  
 244  *$10^{-3}$  for both methods are considered non-detect. The difference between ddPCR and qPCR credible*  
 245 *intervals is also shown (b) and grey shade for its variance.*

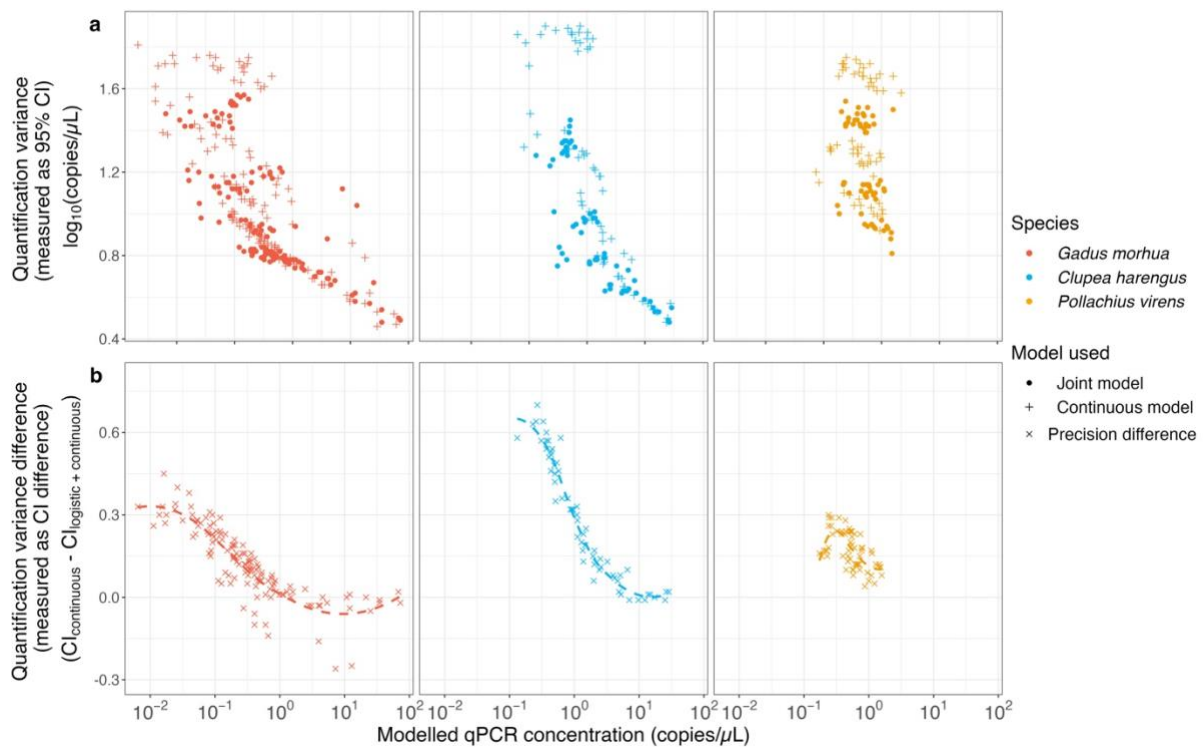
246  
 247 The modelled cod and saithe ddPCR parameters ( $\kappa_0 = -7.07$  and  $\kappa_1 = 2.3$  from equation 2.2, Table S1)  
 248 mirrored default parameters for ddPCR (default Poisson statistics in QuantSoft software; Fig. S3).

249 Conversely, the modelled parameters for herring detection were higher than the aforementioned assays  
 250 ( $\kappa_0 = -7.54$  and  $\kappa_1 = 2.24$ , Table S1).

251

## 252 Two-step model of qPCR quantification precision

253 When estimating the additive value of the detection probability model with the continuous model for  
 254 qPCR, results indicate that the quantification precision (expressed as variance) for continuous model  
 255 alone was lower (i.e., higher variance) than the precision when models were run jointly (detection  
 256 probability + continuous; Fig. 3). We observed an increase in quantification precision by an average  
 257 0.5 orders of magnitude for samples with starting DNA concentrations lower than  $10^2$  copies/ $\mu\text{L}$  (Fig.  
 258 3b).



259

260 *Figure 3. Quantitation precision (expressed as the quantification variance) of continuous model (plus*  
 261 *symbols) and joint model (continuous and detection probability model for qPCR; dots symbols) for*  
 262 *three assays (cod = red, herring = blue, and saithe = orange; a) and their difference in precision (b;*  
 263 *crosses symbol).*

264

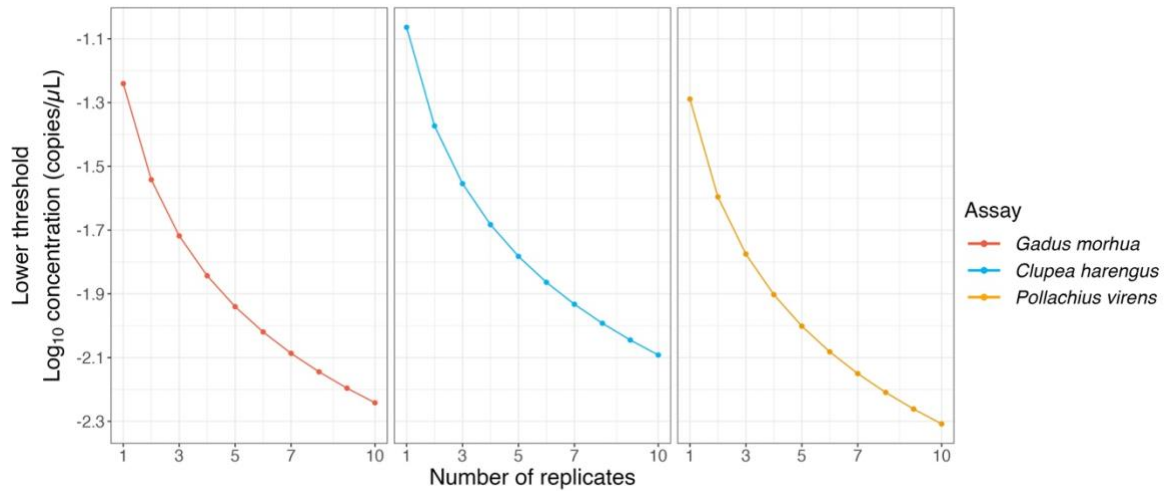
## 265 Lower and upper limit of ddPCR detection and quantification

266 In contrast to qPCR, ddPCR has estimable lower limit of detection and quantification, which is defined  
 267 as a single positive droplet from the pool of generated droplets ( $U$ ). The lower limit of detection is  
 268 calculated in following equation derived from equation 2.2:

$$269 \log_{10}(C_{\text{threshold}}) = \frac{\text{cloglog}(1/U) - \kappa_0}{\kappa_1}$$

270 where  $C_{\text{threshold}}$  is absolute concentration (copies/ $\mu\text{L}$ ) lower limit of detection and quantification. Given  
 271 that many of the parameters involved in ddPCR workflow are constrained, the most efficient way to

272 decrease the limit of quantification is by increasing the number of droplets generated and analysed.  
273 Since each droplet in ddPCR acts as an independent end-point PCR reaction, increasing the number of  
274 “amplification replicates” of the sample by increasing the number of droplets expands the ranges of  
275 both upper and lower limits of detection and quantification (Fig. 4).



276  
277 *Figure 4. The lower limit of ddPCR quantitation as a factor of number of replicates for three assays*  
278 *(cod = red, herring = blue, and saithe = orange).*

279

## 280 Discussion

281 Sensitive and precise quantification of eDNA is essential for understanding the presence, distribution,  
282 and abundance of aquatic organisms, and for informing management decisions related to biodiversity  
283 conservation, invasive species control, and ecosystem health. Estimating the sensitivity and precision  
284 of quantitation methods can provide guidance to researchers in choosing the most effective techniques  
285 for quantifying organisms of interest. The current study compared the sensitivity and precision of qPCR  
286 and ddPCR for estimating starting DNA concentration using three detection assays for important teleost  
287 fish species. This study showed that ddPCR outperformed qPCR by showing higher sensitivity and  
288 precision, especially at low DNA concentration ( $10^{-2} - 10^0$  copies/μL), typical of environmental DNA  
289 concentrations.

290

### 291 Comparison between qPCR and ddPCR (i.e., sensitivity and quantification precision)

292 This study represents a first comparison of qPCR and ddPCR detection platforms with regard to  
293 empirical and modelled performance indicators. While our findings align with those of prior research,  
294 we independently quantify the sensitivity and the quantification precision of these methods through  
295 empirical measurements. Although the results found in this study are consistent with previous findings,  
296 here we empirically measure methods’ sensitivity and quantification precision. One study concluded  
297 that qPCR detection loses precision at starting DNA concentration lower than  $10^0$  copies/μL  
298 (Mahendran et al., 2020). Concentrations lower than this lead to an increased coefficient of variation

299 (CV) that surpasses the limit of detection for qPCR (C. M. Hindson et al., 2013), thus introducing the  
300 risk of false negative detections (McCall et al., 2014). Our results provide empirical evidence that  
301 ddPCR has higher sensitivity than qPCR for concentrations lower than  $10^0$  copies/ $\mu$ L. This may be  
302 attributed to the fact that ddPCR is based on end-point detection, which circumvents issues related to  
303 PCR inhibition and variation in amplification efficiencies prior to reaching amplification plateaus (S.  
304 C. Taylor et al., 2015; Yang et al., 2014). Verhaegen et al. (2016) have shown that ddPCR and qPCR  
305 (when using Environmental Master Mix) are not affected by the PCR inhibition (induced as bile salt).  
306 However, their standard sample target concentration was relatively high ( $10^3$  copies/ $\mu$ L) and thus may  
307 not accurately reflect patterns at the low concentration ( $10^{-2} - 10^0$  copies/ $\mu$ L) which are often observed  
308 in eDNA-based surveys.

309  
310 Despite higher ddPCR detection sensitivity at low target concentrations, we observed a positive  
311 correlation between starting DNA concentration estimates for qPCR and ddPCR, in agreement with  
312 findings from previous studies (Campomenosi et al., 2016; Doi et al., 2015; Jerde et al., 2016; Nathan  
313 et al., 2014; Tang et al., 2016). While we did notice slight differences between assays, both approaches  
314 provided similar overall estimates of concentration. However, when considering the precision of each  
315 method, we found that ddPCR outperformed qPCR in estimating quantities by half to one order of  
316 magnitude across all concentrations. This may again be attributed to the endpoint nature of ddPCR,  
317 which avoids the stochastic variation that can occur between cycles of DNA amplification. One of the  
318 explanations for low detectability and higher variance of qPCR could be the low amount of DNA  
319 template used in such reactions. Although Takahara et al. (2015) concluded from their experiments that  
320 using 2  $\mu$ L of DNA template showed higher detection rate due to less PCR inhibition (Doi et al., 2015;  
321 Tang et al., 2016). As eDNA extracts vary greatly in content of PCR inhibitors, it is difficult to make  
322 general statements about template volume and potential negative impact on PCR amplification  
323 efficiency. In the present study, the extent of observed platform-dependent variance depended on the  
324 assay in question, with the cod assay showing lower variation between the two methods compared to  
325 the herring assay.

326

### 327 **ddPCR (non)reliance external calibration**

328 Although studies have stated that the advantage of ddPCR is freedom from the standard curves  
329 (typically required for qPCR; Verhaegen et al., 2016), our ddPCR results indicate an assay-specific  
330 difference in the relationship between positive droplets and nominal concentrations of different assays  
331 (Fig. S3). We can assume from equation 2 and 4 that  $\kappa_0 \approx -\ln(V)$  and  $\kappa_1 \approx \ln(10)$ . Given our estimates  
332 for the latent parameters (Table S1) we conclude that a single droplet volume for cod and saithe assay  
333 (as well as for default ddPCR calculations) is  $\sim 0.00085 \mu$ L. Parameters of herring assay indicated a  
334 droplet volume 1.6 times smaller than those of cod and saithe ( $\sim 0.00054 \mu$ L). Such values would, by  
335 extrapolation, indicate that fewer DNA molecules are being encapsulated into a single droplet, or the



336 droplets generated for herring detection were smaller. However, droplet metrics indicated similar  
337 numbers of accepted (positive + negative) droplets for all three assays (data now shown), which  
338 challenges the possibility that herring droplets were smaller. Additionally, we noticed that fluorescence  
339 amplitude for herring was significantly higher compared to the other two assays (Fig. S4) yet unable to  
340 explain how that can affect our parameters. Subsequently, DNA degradation of herring standard  
341 samples (see the difference between nominal and measured herring standard concentration in Fig. S2)  
342 could potentially explain the difference between nominal and known concentrations, thus altering the  
343 parameters of our model (Fig. S3). We suggest that standard curve inclusion during ddPCR assay  
344 optimization may be a useful supportive tool to increase understanding of ddPCR assay behaviour and  
345 allow targeted calibration measures to improve application precision and accuracy.

346

347 Our analysis showed that the default calculations in the QuantaSoft ddPCR software tend to  
348 overestimate low DNA concentrations from standard samples (Fig. S2). We denote that in scenarios  
349 when quantitation accuracy is paramount, the inclusion of standard samples may become instrumental  
350 for attaining the necessary level of accuracy. This can be done using the models described here with  
351 adoption of  $xlogx$  function in equation 2.2 which can additionally account for the low concentration  
352 tilt. Conversely, in situations where accuracy is not the primary focus of the study, resorting to the  
353 default ddPCR software calculations, or employing the  $cloglog$  transformation of the proportions of  
354 positive droplets, remain acceptable alternative approaches. It is essential to recognize that, in the latter  
355 case, any overestimation by ddPCR will be uniform across all samples.

356

### 357 **Conclusion**

358 This study provides valuable insights into the strengths and limitations of qPCR and ddPCR methods  
359 and help to inform best practices for eDNA research and monitoring in the future. We find that ddPCR  
360 has higher sensitivity and precision at low DNA concentrations, which are of particular relevance for  
361 eDNA-based detection surveys. Furthermore, we recommend the use of standard samples when  
362 optimizing ddPCR assays as it increases understanding of assay technical performance and facilitates  
363 troubleshooting when assay efficiency is unsatisfactory. We show that implementation of a qPCR  
364 detection probability model, when used jointly with the continuous model, improves quantification  
365 precision. We strongly encourage the use of technical replicates not only as a tool to increase the  
366 precision and accuracy of measurement but also as means to reduce statistical uncertainty of detection  
367 at low target concentrations.

368 **References**

- 369 Campomenosi, P., Gini, E., Noonan, D. M., Poli, A., D'Antona, P., Rotolo, N., Dominionioni, L., &  
370 Imperatori, A. (2016). A comparison between quantitative PCR and droplet digital PCR  
371 technologies for circulating microRNA quantification in human lung cancer. *BMC*  
372 *Biotechnology*, *16*(1), 60. <https://doi.org/10.1186/s12896-016-0292-7>
- 373 Doi, H., Takahara, T., Minamoto, T., Matsushashi, S., Uchii, K., & Yamanaka, H. (2015). Droplet  
374 Digital Polymerase Chain Reaction (PCR) Outperforms Real-Time PCR in the Detection of  
375 Environmental DNA from an Invasive Fish Species. *Environmental Science & Technology*,  
376 *49*(9), 5601–5608. <https://doi.org/10.1021/acs.est.5b00253>
- 377 Guri, G., Westgaard, J., Yoccoz, N., Wangensteen, O. S., Præbel, K., Ray, J. L., Kelly, R. P., Shelton,  
378 A. O., Hanebrekke, T., & Johansen, T. (2023). Maximizing sampling efficiency to detect  
379 differences in fish community composition using environmental DNA metabarcoding in  
380 subarctic fjords. *Environmental DNA*, edn3.409. <https://doi.org/10.1002/edn3.409>
- 381 Hindson, B. J., Ness, K. D., Masquelier, D. A., Belgrader, P., Heredia, N. J., Makarewicz, A. J., Bright,  
382 I. J., Lucero, M. Y., Hiddessen, A. L., Legler, T. C., Kitano, T. K., Hodel, M. R., Petersen, J.  
383 F., Wyatt, P. W., Steenblock, E. R., Shah, P. H., Bousse, L. J., Troup, C. B., Mellen, J. C., ...  
384 Colston, B. W. (2011). High-Throughput Droplet Digital PCR System for Absolute  
385 Quantitation of DNA Copy Number. *Analytical Chemistry*, *83*(22), 8604–8610.  
386 <https://doi.org/10.1021/ac202028g>
- 387 Hindson, C. M., Chevillet, J. R., Briggs, H. A., Gallichotte, E. N., Ruf, I. K., Hindson, B. J., Vessella,  
388 R. L., & Tewari, M. (2013). Absolute quantification by droplet digital PCR versus analog real-  
389 time PCR. *Nature Methods*, *10*(10), 1003–1005. <https://doi.org/10.1038/nmeth.2633>
- 390 Jerde, C. L., Olds, B. P., Shogren, A. J., Andruszkiewicz, E. A., Mahon, A. R., Bolster, D., & Tank, J.  
391 L. (2016). Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate  
392 Environmental DNA. *Environmental Science & Technology*, *50*(16), 8770–8779.  
393 <https://doi.org/10.1021/acs.est.6b01761>
- 394 Keller, A. G., Grason, E. W., McDonald, P. S., Ramón-Laca, A., & Kelly, R. P. (2022). Tracking an  
395 invasion front with environmental DNA. *Ecological Applications*, *32*(4), e2561.  
396 <https://doi.org/10.1002/eap.2561>
- 397 Knudsen, S. W., Ebert, R. B., Hesselsøe, M., Kuntke, F., Hassingboe, J., Mortensen, P. B., Thomsen,  
398 P. F., Sigsgaard, E. E., Hansen, B. K., Nielsen, E. E., & Møller, P. R. (2019). Species-specific  
399 detection and quantification of environmental DNA from marine fishes in the Baltic Sea.  
400 *Journal of Experimental Marine Biology and Ecology*, *510*, 31–45.  
401 <https://doi.org/10.1016/j.jembe.2018.09.004>
- 402 Mahendran, P., Liew, J. W. K., Amir, A., Ching, X.-T., & Lau, Y.-L. (2020). Droplet digital polymerase  
403 chain reaction (ddPCR) for the detection of Plasmodium knowlesi and Plasmodium vivax.  
404 *Malaria Journal*, *19*(1), 241. <https://doi.org/10.1186/s12936-020-03314-5>
- 405 McCall, M. N., McMurray, H. R., Land, H., & Almudevar, A. (2014). On non-detects in qPCR data.  
406 *Bioinformatics*, *30*(16), 2310–2316. <https://doi.org/10.1093/bioinformatics/btu239>
- 407 Nash, R. D. M., Wright, P. J., Matejusova, I., Dimitrov, S. P., O'Sullivan, M., Augley, J., & Höffle, H.  
408 (2012). Spawning location of Norway pout (*Trisopterus esmarkii* Nilsson) in the North Sea.  
409 *ICES Journal of Marine Science*, *69*(8), 1338–1346. <https://doi.org/10.1093/icesjms/fss130>
- 410 Nathan, L. M., Simmons, M., Wegleitner, B. J., Jerde, C. L., & Mahon, A. R. (2014). Quantifying  
411 Environmental DNA Signals for Aquatic Invasive Species Across Multiple Detection  
412 Platforms. *Environmental Science & Technology*, *48*(21), 12800–12806.  
413 <https://doi.org/10.1021/es5034052>

414 Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., & Boehm, A. B. (2016).  
415 Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine  
416 Fish. *Environmental Science & Technology*, *50*(19), 10456–10464.  
417 <https://doi.org/10.1021/acs.est.6b03114>

418 Shelton, A. O., Ramón-Laca, A., Wells, A., Clemons, J., Chu, D., Feist, B. E., Kelly, R. P., Parker-  
419 Stetter, S. L., Thomas, R., Nichols, K. M., & Park, L. (2022). Environmental DNA provides  
420 quantitative estimates of Pacific hake abundance and distribution in the open ocean.  
421 *Proceedings of the Royal Society B: Biological Sciences*, *289*(1971), 20212613.  
422 <https://doi.org/10.1098/rspb.2021.2613>

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

425 Takahara, T., Minamoto, T., & Doi, H. (2015). Effects of sample processing on the detection rate of  
426 environmental DNA from the Common Carp (*Cyprinus carpio*). *Biological Conservation*, *183*,  
427 64–69. <https://doi.org/10.1016/j.biocon.2014.11.014>

428 Tang, H., Cai, Q., Li, H., & Hu, P. (2016). Comparison of droplet digital PCR to real-time PCR for  
429 quantification of hepatitis B virus DNA. *Bioscience, Biotechnology, and Biochemistry*, *80*(11),  
430 2159–2164. <https://doi.org/10.1080/09168451.2016.1196576>

431 Taylor, M. I., Fox, C., Rico, I., & Rico, C. (2002). Species-specific TaqMan probes for simultaneous  
432 identification of (*Gadus morhua* L.), haddock (*Melanogrammus aeglefinus* L.) and whiting  
433 (*Merlangius merlangus* L.). *Molecular Ecology Notes*, *2*(4), 599–601.  
434 <https://doi.org/10.1046/j.1471-8286.2002.00269.x>

435 Taylor, S. C., Carbonneau, J., Shelton, D. N., & Boivin, G. (2015). Optimization of Droplet Digital  
436 PCR from RNA and DNA extracts with direct comparison to RT-qPCR: Clinical implications  
437 for quantification of Oseltamivir-resistant subpopulations. *Journal of Virological Methods*,  
438 *224*, 58–66. <https://doi.org/10.1016/j.jviromet.2015.08.014>

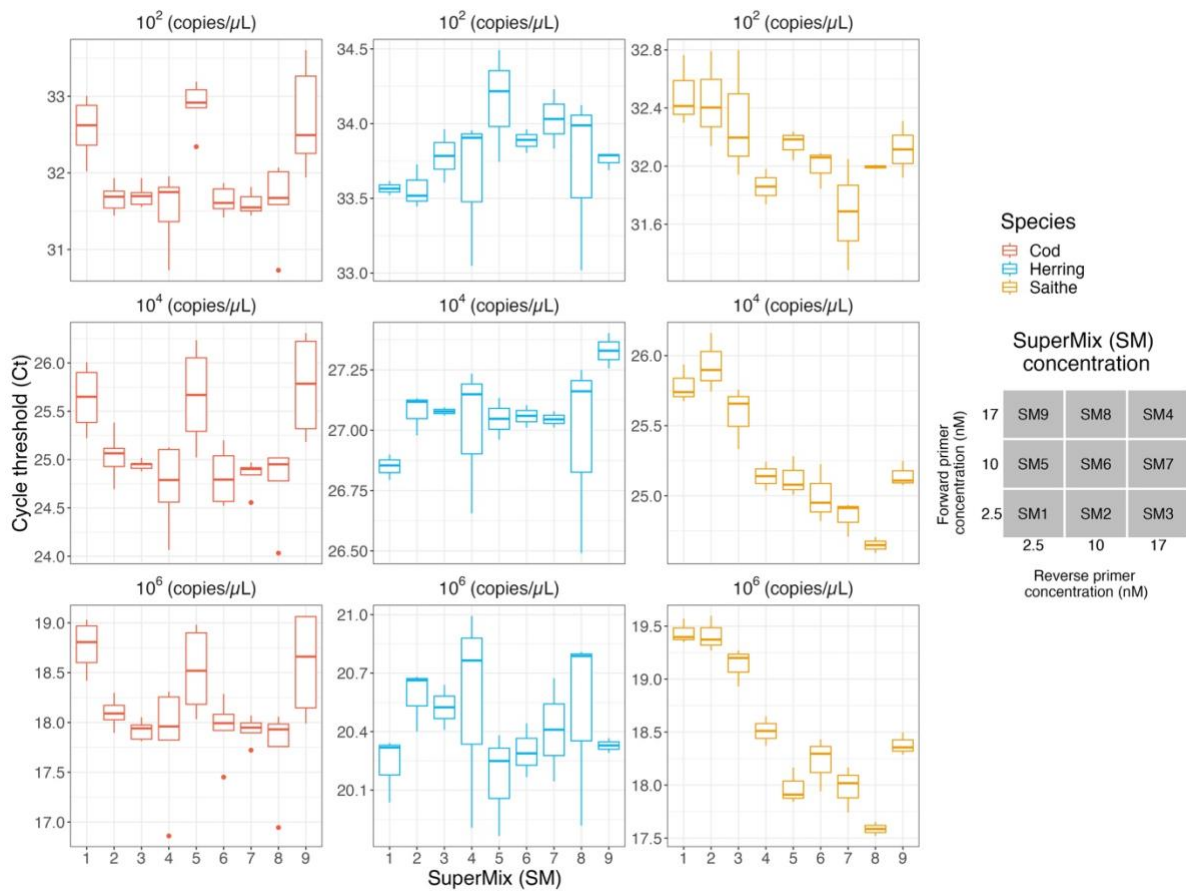
439 Verhaegen, B., De Reu, K., De Zutter, L., Verstraete, K., Heyndrickx, M., & Van Coillie, E. (2016).  
440 Comparison of Droplet Digital PCR and qPCR for the Quantification of Shiga Toxin-Producing  
441 *Escherichia coli* in Bovine Feces. *Toxins*, *8*(5), 157. <https://doi.org/10.3390/toxins8050157>

442 Wood, Z. T., Lacoursière-Roussel, A., LeBlanc, F., Trudel, M., Kinnison, M. T., Garry McBrine, C.,  
443 Pavey, S. A., & Gagné, N. (2021). Spatial Heterogeneity of eDNA Transport Improves Stream  
444 Assessment of Threatened Salmon Presence, Abundance, and Location. *Frontiers in Ecology*  
445 *and Evolution*, *9*, 650717. <https://doi.org/10.3389/fevo.2021.650717>

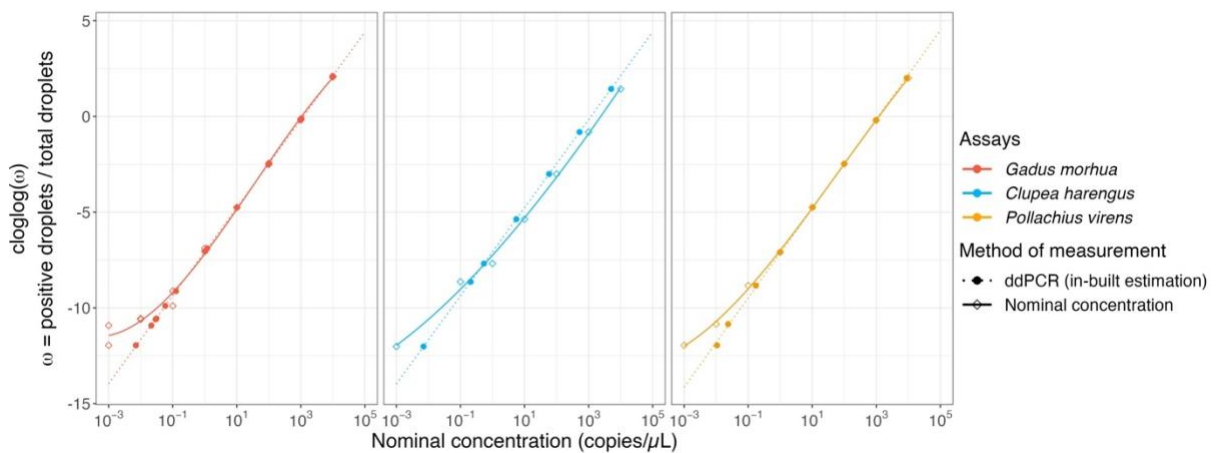
446 Yang, R., Papparini, A., Monis, P., & Ryan, U. (2014). Comparison of next-generation droplet digital  
447 PCR (ddPCR) with quantitative PCR (qPCR) for enumeration of *Cryptosporidium* oocysts in  
448 faecal samples. *International Journal for Parasitology*, *44*(14), 1105–1113.  
449 <https://doi.org/10.1016/j.ijpara.2014.08.004>

450  
451

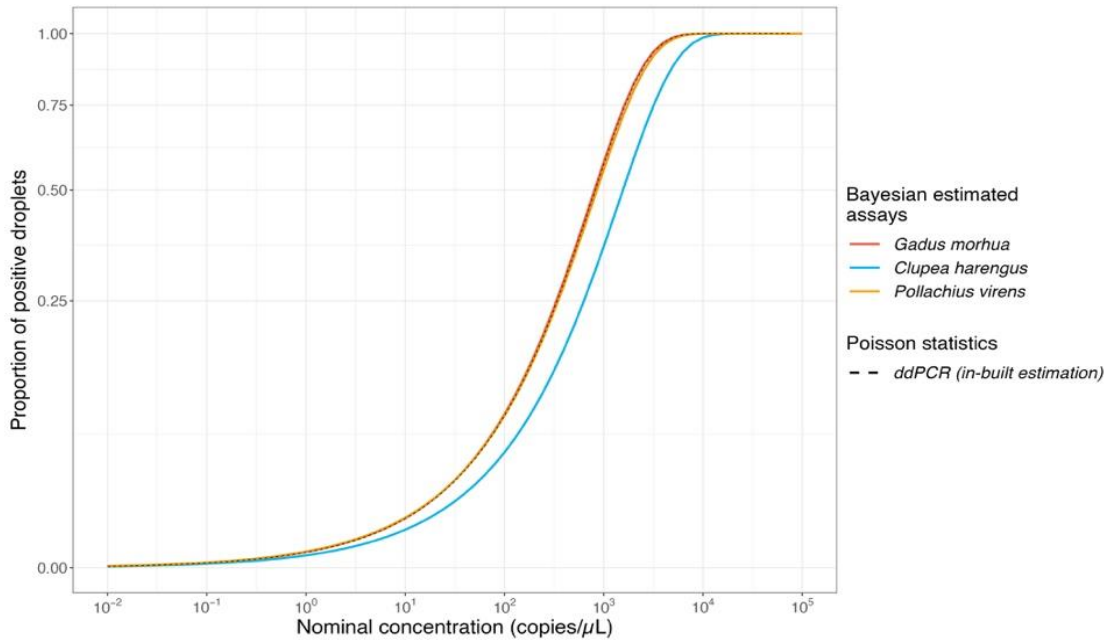
452 **Supplementary material**



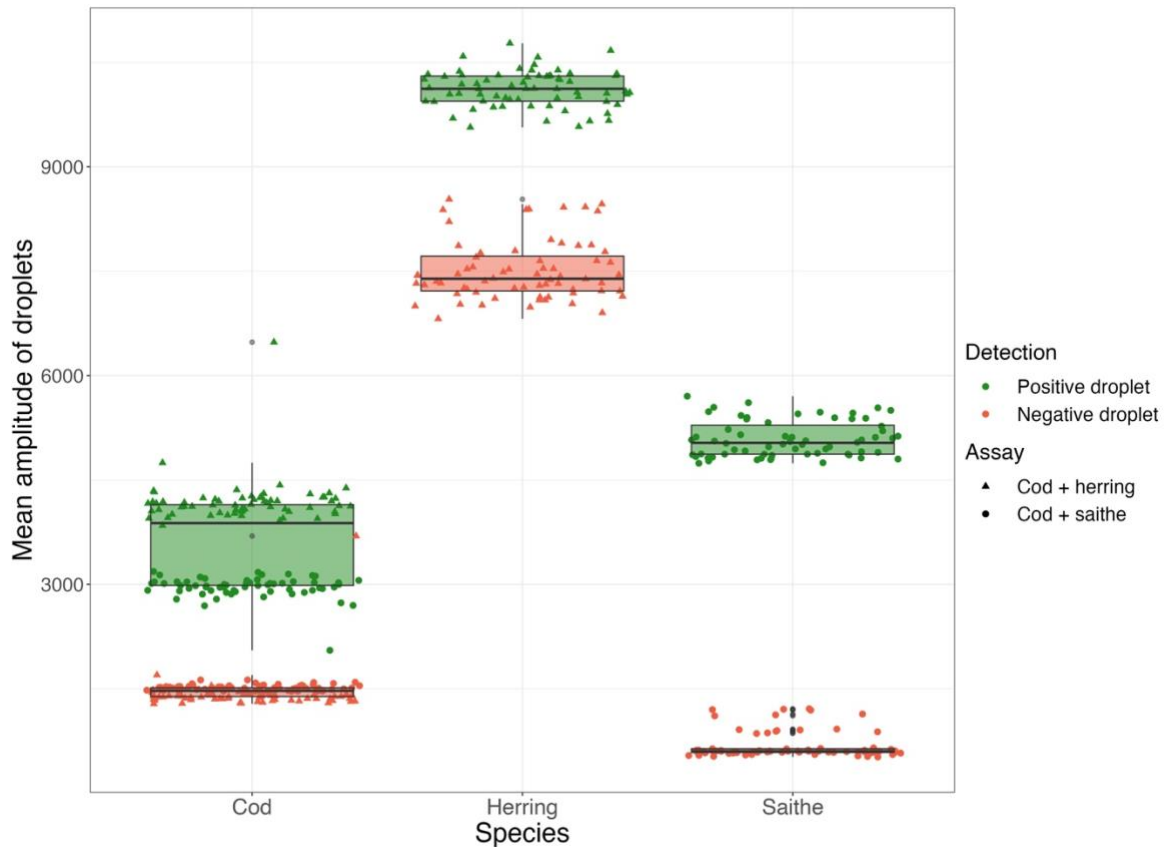
453  
 454 *Figure S1. Optimisation runs for all three assays (cod = red, herring = blue, and saithe = orange) by*  
 455 *using 9 different mixtures of Forward and Reverse primers (SM 1 to 9) and three nominal*  
 456 *concentrations ( $10^2$ ,  $10^4$ , and  $10^6$  copies/ $\mu\text{L}$ ). Lowest Ct indicates the highest efficiency.*  
 457



458  
 459 *Figure S2. The relationship between proportion of positive droplets and nominal DNA concentration*  
 460 *(line) and ddPCR default software estimation (dotted line) for three assays (cod = red, herring = blue,*  
 461 *and saithe = orange) using ddPCR mechanism of measurement.*  
 462



463  
 464 *Figure S3. The relationship between proportion of positive droplets over the total number of droplets*  
 465 *and nominal eDNA concentrations for three assays (cod = red, herring = blue, and saithe = orange)*  
 466 *and the in-built ddPCR Poisson statistics.*



467  
 468 *Figure S4. Mean amplitude of positive (green) and negative (red) droplet for all standard samples for*  
 469 *two multiplexed assays (i) cod and herring (triangle symbol) and (ii) cod and saithe (circle symbol)*  
 470 *indicating high amplitude (both negative and positive droplets) for herring target.*  
 471

472 *Table S1. Latent parameters estimate for three assays (cod, herring, and saithe) for both models (joint*  
 473 *qPCR model and ddPCR model) implemented through Bayesian statistics (equation 1 and 2).*

Mean parameters estimates  $\pm$  standard deviation

	qPCR				ddPCR	
	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$\kappa_0$	$\kappa_1$
Cod	$2.22 \pm 0.19$	$1.35 \pm 0.13$	$38.01 \pm 0.12$	$-3.38 \pm 0.02$	$-7.05 \pm 0.01$	$2.30 \pm 0.00$
Herring	$0.87 \pm 0.20$	$2.09 \pm 0.29$	$40.32 \pm 0.13$	$-3.23 \pm 0.02$	$-7.52 \pm 0.02$	$2.24 \pm 0.00$
Saithe	$0.42 \pm 0.20$	$1.31 \pm 0.24$	$38.58 \pm 0.24$	$-3.09 \pm 0.05$	$-6.99 \pm 0.02$	$2.26 \pm 0.01$

474

## Paper 3

### **eDNA quantification of fish abundance: A novel approach integrating ddPCR, metabarcoding and trawl surveys**

Gledis Guri<sup>1,5\*</sup>, Andrew Olaf Shelton<sup>2</sup>, Ryan P. Kelly<sup>3</sup>, Nigel Yoccoz<sup>4</sup>, Torild Johansen<sup>1</sup>, Kim Præbel<sup>5</sup>, Tanja Hanebrekke<sup>1</sup>, Jessica Louise Ray<sup>6</sup>, Johanna Fall<sup>7</sup>, Jon-Ivar Westgaard<sup>1</sup>

1. Norwegian Institute of Marine Research, Framsenteret, Tromsø, Norway
2. Northwest Fisheries Science Center, NOAA Fisheries, Seattle, Washington, USA
3. School of Marine and Environmental Affairs, University of Washington, Seattle, Washington, USA
4. Dept. of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway
5. Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway
6. NORCE Norwegian Research Centre AS, Climate & Environment Department, Uni Research Environment, N-5020 Bergen, Norway
7. Norwegian Institute of Marine Research, Bergen, Norway

\*Corresponding author: Gledis Guri, [gledis.guri@hi.no](mailto:gledis.guri@hi.no), +46 700658104





1 **Abstract**

2 Quantifying the abundance (i.e., biomass or number of individuals), diversity, and distribution of marine  
3 species is a critical aspect of understanding and managing marine ecosystems. In recent years, there has  
4 been growing interest in using environmental DNA (eDNA) for marine ecosystem management and  
5 biodiversity assessment. However, the main challenge for hindering its applicability towards ecosystem  
6 management has been the inability to infer absolute species abundances from eDNA metabarcoding. In  
7 this study, we demonstrate a way forward by estimating the abundance of commercially important fish  
8 species in a Norwegian fjord using a joint Bayesian statistical model of traditional trawl-catch data and  
9 molecular data derived from eDNA. Using this model, we accurately predicted out-of-sample trawl  
10 catches using eDNA alone. Moreover, our model provides empirical estimates for key processes linking  
11 marine eDNA concentration to the fish population abundance estimated from trawl observations (i.e.,  
12 trawl catchability, DNA shedding, degradation, dilution, transport, recovery rate, and isolation  
13 efficiency). These findings have broad implications for the use of eDNA in marine ecosystem  
14 management and conservation efforts.

15 **Introduction**

16 Quantifying the abundance, diversity, and distribution of species is a fundamental component of natural  
17 resource management and conservation (Preston, 1948; Jetz *et al.*, 2012; Hill *et al.*, 2014; Thomsen *et*  
18 *al.*, 2016; Callaghan *et al.*, 2021; Farr *et al.*, 2022). Traditional survey methods (i.e., scientific, and  
19 commercial trawl catches, close kin mark-recapture, telemetry, hydro-acoustics, and electrofishing  
20 surveys) have been widely used to gather ecological data and form the foundation for most fisheries  
21 management and conservation (Fraser *et al.*, 2007; Neebling and Quist, 2011; Crossin *et al.*, 2017; Lees  
22 *et al.*, 2021). However, these methods are invasive, and a few have a net negative ecological impact  
23 (Biju Kumar and Deepthi, 2006; Eigaard *et al.*, 2017). Furthermore, several of these methods are cost-  
24 intensive, labour intensive, typically require specialized expertise, and may be unsuitable for specific  
25 habitats (Lacoursière-Roussel *et al.*, 2016).

26

27 In recent years, there has been an increasing focus on the application of molecular tools, such as  
28 environmental DNA (eDNA), for quantifying species abundance and conducting biodiversity  
29 assessments (Thomsen *et al.*, 2012). Despite the recent technological advancements, eDNA has not yet  
30 reached the level to fully replace traditional surveys for fish ecosystem monitoring and stock  
31 assessments. However, the integration of eDNA surveys with traditional methods can increase our  
32 understanding of marine ecosystems (Valdivia-Carrillo *et al.*, 2021; Pont *et al.*, 2022), fisheries  
33 management (Stoeckle *et al.*, 2021), and conservation (Hill *et al.*, 2014; He *et al.*, 2023). By combining  
34 these two survey types, we can obtain more comprehensive and accurate information on the species  
35 distribution, including elusive species that may not be captured by traditional trawl surveys (Schadewell  
36 and Adams, 2021; Veron *et al.*, 2023). This becomes especially pivotal when considering the ecological  
37 repercussions of trawl surveys, including adverse effects on non-target species and vulnerable benthic  
38 habitats (Lacoursière-Roussel *et al.*, 2016). The integration of eDNA and trawl data can help to address  
39 some of the drawbacks of traditional survey methods, such as their high costs, invasive nature, and  
40 reliance on taxonomic expertise (Closek *et al.*, 2019; Pont *et al.*, 2022; He *et al.*, 2023; Veron *et al.*,  
41 2023), thereby mitigating the dependence on the latter survey method (Valdivia-Carrillo *et al.*, 2021).

42

43 The “eDNA metabarcoding” – molecular identification of multiple species simultaneously – is now  
44 used to detect species, infer community composition, biodiversity changes, and ecological shifts due to  
45 anthropogenic impacts (Hansen *et al.*, 2018; Jeunen *et al.*, 2019; Atienza *et al.*, 2020; Larson *et al.*,  
46 2022). Recent studies have highlighted the power of eDNA to reveal fine-scale changes and differences  
47 in eukaryotic communities (Turon *et al.*, 2022; Guri *et al.*, 2023).

48

49 Despite the potential of eDNA as a method for understanding marine ecosystems, there are still several  
50 challenges that need to be addressed (Ramírez-Amaro *et al.*, 2022). Due to PCR stochasticity and  
51 species-specific amplification efficiencies, especially when using universal markers (i.e., 12S, 16S or,

52 COI), metabarcoding methods generally do not reflect absolute concentration of eDNA, typically due  
53 to the data being compositional in nature thus providing only inferable proportions (Kelly *et al.*, 2019).  
54 Consequently, to date, the quantitative assessment of biomass or the abundance of targeted organisms  
55 remains elusive (Rourke *et al.*, 2022).

56  
57 Such constraints have significantly hindered the broader adoption and integration of eDNA  
58 metabarcoding methods in cross-disciplinary fields such as fisheries, conservation biology, and  
59 ecosystem-based management (Ramírez-Amaro *et al.*, 2022). Species-specific quantitative PCR  
60 methods (i.e., real-time PCR - qPCR or digital droplet PCR – ddPCR; Pont *et al.*, 2022) offer an  
61 attractive alternative to metabarcoding because they are both quantitative and reflect absolute eDNA  
62 concentrations. For example, Salter *et al.* (2019) showed a high correlation ( $R^2 = 0.66$ ;  $p = 0.008$ )  
63 between absolute eDNA quantities and the biomass of Atlantic cod, and Shelton *et al.* (2022) showed  
64 reliable hake distributions from eDNA on a continental scale. More recently, Maes *et al.* (2023)  
65 employed ddPCR and found eDNA concentrations can represent fish abundance and biomass for two  
66 commercial flatfishes (*Solea solea* and *Pleuronectes platessa*). In context of ecosystem based  
67 management (multiple species), species-specific methods can be more expensive and less efficient  
68 compared to metabarcoding (multi-species), primarily due to the requirement for prior knowledge or  
69 isolation of the target species (Schneider *et al.*, 2016).

70  
71 With respect to fish communities (multi-species data), eDNA metabarcoding has advanced rapidly.  
72 Several studies have shown positive relationships between eDNA metabarcoding data and trawl data  
73 (Thomsen *et al.*, 2016; Afzali *et al.*, 2021; Stoeckle *et al.*, 2021; Yates *et al.*, 2022). Despite variation  
74 in methods, nearly all studies rely on high-throughput sequencing reads of some sort for comparison  
75 with trawl data. It is important to acknowledge that such comparisons inherit considerable biases  
76 embedded at multiple levels (i.e., PCR bias, species-specific DNA shedding rates, and environmental  
77 factors), potentially leading to inaccuracies in the results. Even when accounting for species-specific  
78 DNA shedding rates (in the form of allometric scaling; Yates *et al.*, 2022), neglecting species-specific  
79 amplification biases and trawl catchability can result in substantial unexplained variation between  
80 metabarcoding reads and trawled biomass (see also Veron *et al.*, 2023). Stoeckle *et al.* (2021) compared  
81 proportional metabarcoding reads to proportional abundances from bottom trawls and observed a  
82 positive correlation when comparing across different months (mean  $R^2 = 0.565$ ). Such evidence  
83 indicates that metabarcoding reads (without any type of transformation accounting for PCR biases)  
84 cannot yield absolute abundance but can reveal information of rank or semiquantitative abundances  
85 (Salter *et al.*, 2019; Stoeckle *et al.*, 2021; Guri *et al.*, 2023; Veron *et al.*, 2023), which may be an inherent  
86 characteristic of the data. Pont *et al.* (2022) and Allan *et al.* (2023) addressed this issue by combining  
87 eDNA metabarcoding with qPCR (multiplying metabarcoding relative read-abundances with total  
88 qPCR fish DNA concentration) and found strong correlation between total DNA concentrations and

89 total fish biomass measured through traditional electrofishing in river systems. However, when  
90 comparing metabarcoding with the species-specific DNA concentration, they found roughly two orders  
91 of magnitude variation, which may have resulted from differences in species-specific amplification  
92 efficiencies. Shelton *et al.* (2023) developed a model to estimate initial DNA proportions derived from  
93 metabarcoding while accounting for species-specific amplification biases by sequencing mock  
94 community samples (known concentration of DNA extracts for a given list of taxa) alongside  
95 environmental samples. To further improve the application of these models, additional information is  
96 required that can go beyond proportions and deliver absolute concentration from metabarcoding data.  
97 Addressing all aforementioned challenges will be critical to the effective application of eDNA  
98 metabarcoding in ecological research.

99  
100 In this study, we developed a comprehensive and reliable framework for analysing eDNA data in  
101 conjunction with other ecological data sources (i.e., trawl catch data) to improve the quantitative  
102 accuracy of metabarcoding analyses and provide a more holistic understanding of aquatic ecosystems.  
103 Our Bayesian joint model framework integrated eDNA concentrations (from eDNA metabarcoding and  
104 ddPCR) with fish density from trawl surveys (Figure 1) to provide estimates of biological parameters  
105 of interest (e.g. biomass) and parameters describing the links between eDNA and traditional sampling  
106 data. Additionally, using these links, we show how we can predict out-of-sample trawl catches solely  
107 using eDNA observations.

## 108 109 **Methods**

110 We developed and joined five distinct datasets – trawl-catch counts, ddPCR observations of known  
111 concentration from standard samples for a reference species, ddPCR observations of unknown  
112 concentration from environmental samples for the reference species, metabarcoding reads of mock  
113 communities (a known mixture of DNA extracts from multiple species of interest), and metabarcoding  
114 reads of environmental samples for target fish species – into a common analysis. Below, we describe  
115 how each dataset was produced before providing detailed information on the joint statistical model and  
116 subsequent analyses.

117  
118 In brief, the model uses species-specific ddPCR and a mock community to both calibrate metabarcoding  
119 results and to derive eDNA concentrations in environmental samples. We use this molecular dataset in  
120 combination with traditional trawl data to estimate the true abundances of several species, and  
121 rigorously test the resulting joint model against out-of-sample data to show its reliability.

## 122 123 **Study area and samples**

124 Balsfjord, northern Norway (Figure S1), has a length of 40 km and an average depth of 150 m with a  
125 sill at the fjord's entrance, and the archipelago limits the water exchange between the fjord and the

126 Norwegian Sea (open sea). Nearly all high-latitude Norwegian fjords, including Balsfjord, are ice-free  
127 and characterized by an Arctic light regime (Reigstad and Wassmann, 1996). GPS coordinates and other  
128 metadata of sampling stations including GPS coordinates are provided in Table S1.

129

### 130 **Trawl surveys**

131 Study samples were collected on research cruises in October 2019, 2020, and 2021 on the R/V Kristine  
132 Bonnevie as part of the Norwegian coastal annual surveys. The main aim of the survey is to collect data  
133 for abundance estimation of Norwegian coastal cod (*Gadus morhua*) and Northeast Arctic saithe  
134 (*Pollachius virens*) and catch count and weight of all fish species are recorded. Bottom trawl surveys  
135 were conducted annually in four distinct stations in Balsfjord alongside eDNA samples (Figure S1).  
136 Trawl surveys used a standard sampling trawl known as Campelen 1800 with an 80 mm (stretched)  
137 mesh size in the front section and 22 mm in the codend. The trawl sweeps were 40 meters in length,  
138 and rockhopper gear was employed. Sorting, weighing, measuring, and sampling of the catch were  
139 performed according to the applicable instructions (Mjanger *et al.*, 2019) by the Institute of Marine  
140 Research crew.

141

### 142 **eDNA water sampling, filtration, extraction**

143 Water was collected prior to trawling in 5 L Niskin bottles attached to a rosette which was raised and  
144 lowered using a deck-mounted winch at the established sampling station (Figure S1). At each station,  
145 we filtered triplicates of 2 L of seawater through 0.22 µm Sterivex filters (MerckMillipore) from three  
146 depths, surface (10 m), pycnocline (depth of highest density, ~ 50 m) and bottom (10 m above bottom).  
147 Each filter were transferred to sterile 50 ml Falcon centrifuge tubes and directly stored at -20 °C. DNA  
148 extraction of water samples followed the manufacturer's protocol using the DNeasy PowerWater  
149 Sterivex Kit (Qiagen GmbH, Hilden Germany), with minor adjustments (excluding PowerBead Tubes  
150 steps). For more details regarding these workflows see [Guri \*et al.\* \(2023\)](#). Negative controls such as  
151 field water and air blanks, laboratory blanks, and PCR blanks were taken throughout the workflow as  
152 described in [Shu \*et al.\* \(2020\)](#).

153

### 154 **Droplet digital PCR**

155 Quantitative molecular assays targeting 103-bp region of the ATPase gene of Atlantic cod (*Gadus*  
156 *morhua*, see Table 1) the cytochrome b sequence of the mitochondrial DNA (mtDNA) of Atlantic  
157 herring (*Clupea harengus*, see table 1) were performed using droplet digital PCR (ddPCR) on seawater  
158 samples (Table 1). Assays were optimized using different primer/probe concentrations and temperature  
159 gradient to give the highest fluorescence contrast between positive and negative droplets. Herring and  
160 cod ddPCR assays were run in duplex reactions using 6-FAM and VIC dedicated channels, respectively.  
161 All ddPCR runs were conducted using a QX200 ddPCR system (Bio-Rad) where ca. 20,000 droplets  
162 were generated on 20 µL reaction volumes. Each reaction used 5 µL of DNA template from

163 environmental samples, 10  $\mu$ L of ddPCR Supermix with no dUTP (Bio-Rad), 900 nM of final  
 164 concentration of forward and reverse primers, and 250 nM final concentration of TaqMan probe for  
 165 each assay (cod and herring respectively, Table 1). The thermocycler reactions were run in C1000  
 166 Touch Thermal Cycler with 96-Deep Well Reaction Module (Bio-Rad) using the PCR program as  
 167 follows: 10 min at 95°C for enzyme activation, followed by 44 cycles of denaturation for 1 min at 95°C  
 168 and primer annealing and elongation for 2 min at 55.6°C, with a ramp rate of 2°C per s, followed by 10  
 169 min at 98°C and stored at 4°C. Alongside the environmental samples we ran standard samples of known  
 170 concentration from  $10^{-2}$  –  $10^4$  copies/ $\mu$ L for each species to construct the relationship between positive  
 171 droplets and known concentration (see Table S2).

172

173 *Table 1. Sequences for ddPCR assays targeting Gadus morhua and Clupea harengus (targeting 103-*  
 174 *bp region of the ATPase gene and cytochrome b sequence of the mitochondrial DNA respectively)*

Target	Primers and probe	Sequence	5' Dye	Size
<i>Gadus morhua</i>	Forward	GAD-FII	GCAATCGAGTYGTATCYCTHCAAGGAT	103 bp (Taylor <i>et al.</i> , 2002)
	Reverse	GAD-R III	GCAAGWAGYGGHGRCADTTGTG	(Nash <i>et al.</i> , 2012)
	Probe	Custom MGBNFQ	CTTTTACCTCTAAATGTGGGAGG	VIC
<i>Clupea harengus</i>	Forward	Cluhar_CYBF14928	CCCATTTGTGATTGCAGGGG	86 bp (Knudsen <i>et al.</i> , 2019)
	Reverse	Cluhar_CYBR15013	CTGAGTTAAGTCTGCCGGG	(Knudsen <i>et al.</i> , 2019)
	Probe	Cluhar_CYBP14949	TACTATTCTCCACCTTCTGTTCCTC-BHQ1	6FAM

175

## 176 Mock communities

177 To calibrate metabarcoding observations and account for amplification bias (Gold *et al.*, 2023; Shelton  
 178 *et al.*, 2023a) we selected ten fish species that were common in previous metabarcoding or trawl surveys  
 179 in the study area (Guri *et al.*, 2023). For the species *Maurolicus muelleri* (silvery lightfish), *Gadus*  
 180 *morhua* (cod), *Leptoclinus maculatus* (daubed shanny), *Hippoglossoides platessoides* (long rough dab),  
 181 *Myoxocephalus scorpius* (shorthorn sculpin), *Cyclopterus lumpus* (lumpsucker), *Pholis gunnellus* (rock  
 182 gunnel), *Brosme brosme* (cusk), we used samples collected by and extracted at the University of Tromsø  
 183 for improving the taxonomical resolution of DUFA - DNA Universal database for Fisheries and  
 184 Aquaculture, used for taxonomic assignment of metabarcoding data. The remaining species, *Mallotus*  
 185 *villosus* (capelin) and *Pleuronectes platessa* (European plaice), were purchased from the local fish  
 186 market in Tromsø (Dragøy Fisk AS). Genomic DNA from fish muscle tissue samples was thereafter  
 187 extracted using the DNeasy Blood & Tissue Kit (Qiagen GmbH, Hilden Germany) following the  
 188 manufacturer's protocol. A mock community of representative fish species was constructed from these  
 189 genomic DNA samples by amplifying the mitochondrial 12S ribosomal RNA gene region (169-172 bp  
 190 fragment) from each sample using MiFish-U universal primer set (Forward: 5'-  
 191 GTCGGTAAACTCGTGCCAGC-3'; Reverse: 5'-CATAGTGGGGTATCTAATCCCAGTTTG-3';  
 192 Miya *et al.*, 2015). Reactions were run in 20  $\mu$ L volume containing 12.5  $\mu$ L of TaqMan Environmental  
 193 Master Mix 2.0, 1.25  $\mu$ L of each forward and reverse primers (10 nM concentration each), 8  $\mu$ L of

194 dH<sub>2</sub>O and 2 μL of DNA template of above-listed species. The thermocycler program consisted of an  
195 initial polymerase activation for 10 min at 95°C, followed by 40 cycles of denaturation for 1 min at  
196 95°C, annealing for 1 min at 60°C and primer extension for 1 min at 72°C, followed by a final extension  
197 for 10 min at 72°C and lastly storage at 4°C. PCR products were purified using the MinElute PCR  
198 Purification Kit (Qiagen GmbH, Hilden Germany) to ensure complete removal of primer-dimers. All  
199 samples were measured for Concentration of dsDNA in purified PCR products was measured using the  
200 High Sensitivity dsDNA assay with Qubit quantification system. PCR products were diluted in 10-fold  
201 steps to achieve a final concentration of ~10<sup>5</sup> copies/μL.

202

203 In total six separate (but similar in concentration) mock community samples (250μL each) were created  
204 by combining 25μL of amplicon DNA from all ten species in equimolar concentrations resulting in ca.  
205 10<sup>4</sup> copies/μL for each species (Table S4). Mock community samples were amplified and sequenced  
206 alongside the environmental samples following the same workflow and protocol. Only six species in  
207 the mock community were also caught in trawl surveys (Table S5).

208

### 209 **Library preparation, and NGS sequencing**

210 In total, 109 samples (divided into two libraries), including eDNA samples (n = 84), PCR blanks (n =  
211 2), positive controls (n = 6), extraction blanks (n = 5), fieldwork water and air blank (n = 9 and 3  
212 respectively), were amplified using the MiFish-U universal primer set (the same used for mock  
213 community samples). We used fusion primers (primers containing adaptor and index), allowing for a  
214 one-step PCR protocol, and followed a thermocycler program as described in Guri *et al.* (2023).  
215 Libraries were sequenced on GeneStudio S5 sequencer (Thermo Fisher Scientific) using the Ion 540-  
216 sequencing chip with the 200 bp protocol (Thermo Fisher Scientific).

217

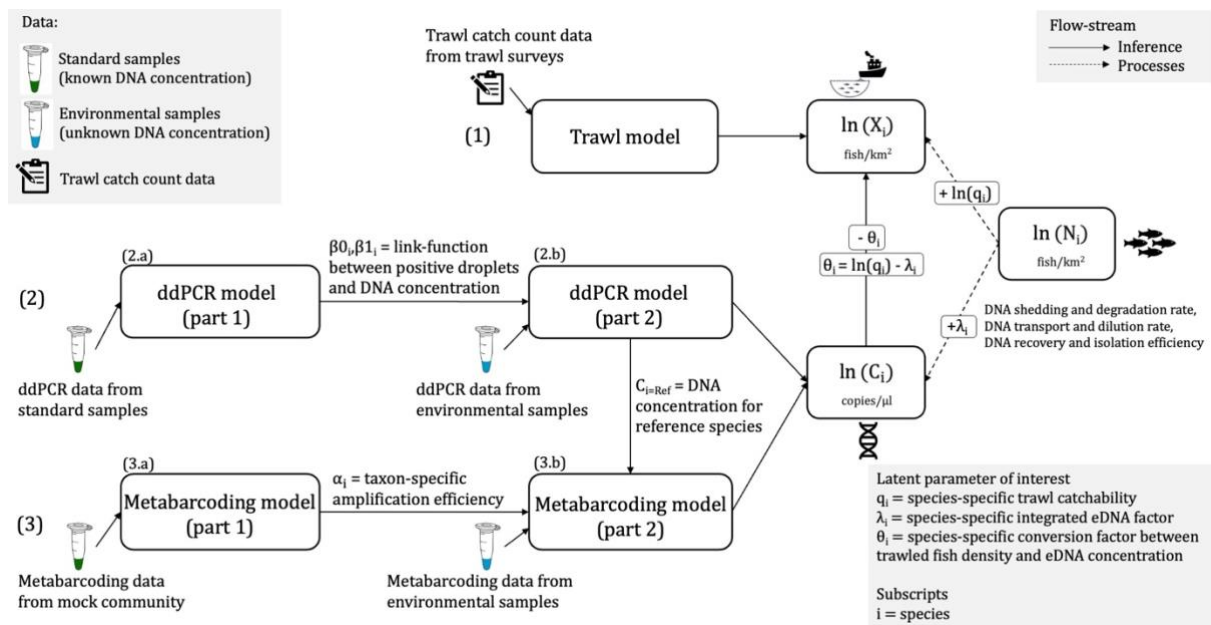
### 218 **Bioinformatics**

219 Sequences were automatically demultiplexed and quality filtered after the sequencing process using  
220 Torrent Suite™ software inbuilt in the sequencer following their default settings. The sampled sequence  
221 dataset was thereafter filtered for chimeric sequences using a *uchime-denovo* algorithm in VSEARCH  
222 (Rognes *et al.*, 2016). The sequences were clustered into Molecular Operational Taxonomic Units  
223 (MOTU) using SWARM v2 (Mahé *et al.*, 2014) with a distance of d = 3 and subsequently the singletons  
224 were removed. MOTUs with higher number of reads than 50 of the total library reads were retained and  
225 annotated by performing BLAST search against NCBI nucleotide (nt) database (date: 23 August 2023)  
226 using the BLASTn algorithm. We used an E-value threshold of 1e-30, maximum target sequences of  
227 50 and a minimal percentage identity of 90. Lastly, we removed all taxa that were not assigned to fish  
228 (class: Actinopterygii or Chondrichthyes) and sequences with read abundance higher than 10% in  
229 negative controls over the environmental samples.

230

231 **Bayesian model**

232 We developed a Bayesian joint model to synthesize information from trawl catches, ddPCR, and  
 233 metabarcoding, to estimate fish quantities. We used mock community samples (box 3a in Figure 1; only  
 234 six species used) with known initial DNA concentrations to estimate species-specific amplification  
 235 efficiencies. Knowing the amplification efficiencies enables the model to determine the species' initial  
 236 proportion thus their relative abundance in environmental samples of metabarcoding data (box 3b in  
 237 Figure 1; see Shelton *et al.*, 2023). Simultaneously, the model determines a link between known  
 238 concentrations and positive droplets observed by the ddPCR system (box 2a in Figure 1). Using this  
 239 link-function, the ddPCR model estimates the absolute concentration of the reference species (one  
 240 species arbitrarily chosen that is present in ddPCR and mock community data) in the environmental  
 241 samples (box 2b in Figure 1). Having both the relative abundance of species and the absolute  
 242 concentration of one species in the environmental samples, the model expands proportion into absolute  
 243 concentration for all species (or at least all species for which amplification efficiency can be estimated).  
 244 Lastly, the model assumes a species-specific correlation between trawl catches and absolute DNA  
 245 concentration which is a conversion parameter (and primary parameter of interest;  $\theta$ ) linking the two  
 246 methods. See Figure 1 for a schematic representation of the joint Bayesian model.



247  
 248 *Figure 1. Simplified schematic overview of the joint Bayesian model workflow including all the*  
 249 *inferences and processes. Note that the model specifies  $\theta$  in the natural-log scale indicating the*  
 250 *magnitude of conversion between fish densities and eDNA concentrations ( $e^\theta$ , with units:*  
 251 *copies/ $\mu$ L/fish/km<sup>2</sup>). Processes (dashed lines) are shown for schematic understanding, thus are*  
 252 *circumvented (not measured) in this model.*

253  
 254 In detail, in this model, we define true fish density as  $N$  (fish/km<sup>2</sup>) and  $q$  for the catchability parameter  
 255 on a scale of 0 to 1 (i.e.,  $\ln(q) < 0$ ; Zhang *et al.*, 2020). We also use subscripts  $i, j, y, d, b, r, m$ , and  $p$  to  
 256 indicate species, station, year, depth category, biological replicate, technical replicate, observation in



257 mock community samples (aliquote samples), and observation in ddPCR standard samples (aliquote  
 258 samples), respectively.

259

260 ***The process (unmeasured biological parameters, Figure 1)***

261 Let  $X$  be the catch per unit effort (CPUE; observed fish density) sampled by trawl catches (units:  
 262 fish/km<sup>2</sup>). Based on equations 1 and 3 in Mahévas *et al.* (2011), we derive CPUE as a function of  
 263 catchability ( $q$ ) and true fish density ( $N$ ; units: fish/km<sup>2</sup>):

264 
$$\ln(X_{ijy}) = \ln(N_{ijy}) + \ln(q_i) \quad (1.1)$$

265 Let  $C$  be the DNA concentration (copies/μL) from the same sampling stations and  $\lambda$  be the “integrated  
 266 eDNA factor” (the conversion factor between the true fish density and observed DNA concentration).  
 267 By analogy to catchability in the context of trawl data, we can assume we model the concentration of  
 268 environmental samples is as a function of the true fish density ( $N$ ) and an integrated eDNA factor ( $\lambda$ )  
 269 expressed as:

270 
$$\ln(C_{ijy}) = \ln(N_{ijy}) + \lambda_i \quad (1.2)$$

271

272 ***The inference (measured biological parameters, Figure 1)***

273 *Trawl model compartment (MC 1)*

274 To model species-specific catch count trawl observations ( $Z$ ) as random draws from an annual mean for  
 275 a given species, we assume a negative binomial distribution of fish density ( $V$ ) and fishing effort ( $E$ ;  
 276 Fraser *et al.*, 2007) with an dispersion parameter  $\phi$  where  $V$  is the mean  $X$  across all stations ( $j$ ) for a  
 277 given species ( $i$ ) and year ( $y$ ):

278 
$$Z_{ijy} \sim \text{NegativeBinomial}(e^{V_{iy}} E_{jy}, \phi_{iy}) \quad (2.1)$$

279 
$$V_{iy} = \frac{\sum_j^J X_{ijy}}{J} \quad (2.2)$$

280

281 *ddPCR model compartment (MC 2)*

282 Let  $S$  be the DNA concentration (copies/μL) from standards. The number of positive droplets ( $W$ ) in  
 283 ddPCR follows a binomial distribution (Guri *et al.*, 2024, in prep), where the probability of a droplet  
 284 being positive ( $\omega$ ) is a linear function of  $\log_e$  DNA concentration ( $S$  or  $C$ ) with an intercept ( $\beta_0$ ) and  
 285 slope ( $\beta_1$ ) using clog-log (link-function), given the total number of droplets generated ( $U$ ) expressed  
 286 as:

287 
$$W_{ip} \sim \text{Binomial}(\omega_{ip}, U_{ip}) \quad (3.1.1)$$

288 
$$\text{cloglog}(\omega_{ip}) = \beta_0_i + (\beta_1_i \ln(S_{ip})) \quad (3.1.2)$$

289 
$$W_{ijy\text{dbr}} \sim \text{Binomial}(\omega_{ijy}, U_{ijy\text{dbr}}) \quad (3.1.3)$$

290 
$$\text{cloglog}(\omega_{ijy}) = \beta_0_i + (\beta_1_i \ln(C_{ijy})) \quad (3.1.4)$$

291 *Metabarcoding model compartment (MC 3)*

292 Based on Shelton *et al.* (2023) we established that the number of observed reads ( $Y$ ) from  
293 metabarcoding is a draw from a multinomial distribution given the proportions for each species  $\psi$  and  
294 the total number of reads per sample  $R$ :

$$295 Y_{im} \sim \text{Multinomial}(\psi_{im}, R_m) \quad (4.1.1)$$

$$296 Y_{ijydb} \sim \text{Multinomial}(\psi_{ijydb}, R_{jydb}) \quad (4.1.2)$$

297

298 These lines of equations (4.2.1 and 4.2.2) are the softmax transformation which produces proportions  
299  $\psi$  for each species ( $i$ ) from the ratios of species abundance  $\gamma$ .

$$300 \vec{\psi}_{im} = \frac{e^{\vec{\gamma}_{im}}}{\sum_i e^{\gamma_{im}}} \quad (4.2.1)$$

$$301 \vec{\psi}_{ijydb} = \frac{e^{\vec{\gamma}_{ijydb}}}{\sum_i e^{\gamma_{ijydb}}} \quad (4.2.2)$$

302 The known initial concentration from mock community samples is converted into additive log ratios  
303 ( $alr$ ) given a reference species of choice (i.e., *Gadus morhua* in this study, see [Shelton \*et al.\*, 2023](#)  
304 regarding selecting a reference species), we map the following equation:

$$305 \gamma_{im} = alr_i + (NPCR \times \alpha_i) + \eta_{im} \quad (4.3.1)$$

306 where NPCR is the number of PCR cycles used to amplify the amplicons,  $\alpha$  is the species-specific  
307 amplification efficiency and  $\eta$  is the parameter that allows for overdispersion in the counts beyond the  
308 variability provided by the multinomial distribution, capturing the substantial variance among replicates  
309 often observed in metabarcoding data.

$$310 \eta_{im} \sim \text{Normal}(0, \tau_i) \quad (4.4.1)$$

$$311 \eta_{ijydb} \sim \text{Normal}(0, \tau_i) \quad (4.4.2)$$

312

313 Having estimated the species-specific amplification efficiencies in equation 4.3.1 we can derive the  
314 absolute DNA quantities for each species, station, and year ( $C_{ijy}$ ), given the known concentration of the  
315 reference species estimated in equation 3.2.2 for the same samples.

$$316 Y_{ijydb} = \ln(C_{ijy}) - \ln(C_{i=Ref,jy}) + (NPCR \times \alpha_i) + \eta_{ijydb} \quad (4.3.2)$$

317 Note that because we have information about the absolute concentration of DNA from some species the  
318 term  $alr$  in equation 4.3.1 can be replaced by the explicit ratio (difference in log-space) of DNA  
319 concentrations (eq. 4.3.2).

320

321 *Joining of the model compartments*

322 Using equations 1.1 and 1.2, we can derive the expression for  $C_{ijy}$  as follows:

$$323 \ln(C_{ijy}) = \ln(X_{ijy}) + \lambda_i - \ln(q_i) \quad (5.1)$$

324

325 Without additional information we cannot estimate neither the catchability ( $q$ ) nor the integrated eDNA  
326 factor in the marine environment ( $\lambda$ ; these two parameters are unidentifiable). Therefore, we introduce  
327 a new parameter ( $\theta$ ) that encapsulates both the integrated eDNA factor and catchability as follows:

$$328 \theta_i = \lambda_i - \ln(q_i) \quad (5.2)$$

329

330 Summarizing equations 5.1 and 5.2, for species detected by both trawl and eDNA we can establish that  
331 eDNA concentration is a function of trawled fish density ( $X$ ) and the new parameter ( $\theta$ ) which is  
332 interpreted as the conversion factor between fish density observed by trawl towards eDNA  
333 concentrations (units: fish/km<sup>2</sup> \* (copies/ $\mu$ L)<sup>-1</sup>) in the following formulation:

$$334 \theta_i = \ln(C_{ijy}) - \ln(X_{ijy}) \quad (5.3)$$

335 By extension of equation 5.2 and 5.3, where a species is only observed in one data stream, indicating  
336 no information on  $\lambda$  or  $q$ ,  $\theta$  inferences are highly uncertain (see Results, Figure 3 b and d).

337

338 Here we denote that the conversion parameter ( $\theta$ ) is in natural logarithmic space thus intuitively  $-\theta$   
339 means the opposite conversion, from eDNA to trawled fish density (units: copies/ $\mu$ L \* (fish/km<sup>2</sup>)<sup>-1</sup>).  
340 The integrated eDNA factor ( $\lambda$ ) is a latent parameter that encapsulates multiple biological latent  
341 variables such as DNA shedding and degradation rate, DNA transport and dilution, and DNA isolation  
342 and extraction efficiency (for more detailed information see  $\theta$ ,  $\phi$ ,  $\psi$ , and  $\xi$  parameters in [Shelton \*et al.\*,  
343 2016](#)).

344

345 The joint model (Figure S6) was implemented using the Stan language as implemented in R (package:  
346 Rstan) running four independent MCMC chains using 4000 warm-up and 6000 sampling iterations (see  
347 Table S6 for all the parameters and their priors). The posterior predictions were diagnosed using  $\hat{R}$   
348 statistics (Gelman and Rubin, 1992) and considered convergence for  $\hat{R}$  values lower than 1.1 and  
349 effective sample size (ESS) greater than 1000 for all parameters.

350

351 To test the robustness of the parameters and the ability to use eDNA observations to estimate fish  
352 density we conducted an out-of-sample analysis using two years of data as a training subset (part I) and  
353 the remaining year as a test subset (part II). In part I, the model estimated the latent parameters ( $\theta$ ,  $\phi$ ,  
354  $\beta_0$ ,  $\beta_1$ ,  $\alpha$ ,  $\eta$ ,  $\psi$ ,  $\gamma$  and  $\tau$ ) using observations (trawl, ddPCR, and metabarcoding) from only two years.  
355 Subsequently, for part II, we predicted fish density in the third year using the estimated parameters from  
356 part I in the trawl model (MC 1) and the eDNA model (MC 2 and 3). We defined delta ( $\delta$ ) as the  
357 difference between log fish density estimated by eDNA models and trawl models thus the log-fold  
358 difference for the out-of-sample year. Delta ( $\delta$ ) values closer to 0 indicate that trawl observations were  
359 well predicted using the eDNA model (or *vice versa*). This process was iterated three times separately  
360 with a different year data left out each time.

361

362 We included the seven species in the Bayesian model for which we had joint presence between trawl  
363 surveys or eDNA observations and ddPCR standard or mock community samples (Table S5)

364

## 365 **Results**

366 The joint model successfully linked disparate datasets in a comprehensive and quantitative framework,  
367 yielding both meaningful estimates of fish abundances and derived parameters of substantial value. We  
368 show the underlying datasets in turn below, before focusing on the synthetic model outputs.

369

### 370 **Trawl data**

371 The average area swept by each trawl was calculated to be  $0.057 \pm 0.01$  km<sup>2</sup>. In total, our bottom trawl  
372 catches encompassed 31 identified species. Among these, 23 species were classified as bony fishes,  
373 four as crustaceans, three as jellyfish, and one ray. Among the catch counts, bony fishes comprised the  
374 most caught species with a total count of 59,381 individuals, with *Mallotus villosus* and  
375 *Hippoglossoides platessoides* having the highest catch number within that class (Table S7). Among the  
376 selected species for the Bayesian model, *Cyclopterus lumpus* was the species with the least catch counts  
377 (n=1).

378

### 379 **ddPCR**

380 The ddPCR protocol consistently yielded an average of  $18,568 \pm 1,281$  droplets for environmental  
381 samples and slightly more for standard samples, ( $19,676 \pm 1,011$  droplets) indicating acceptable levels  
382 of droplet formation for statistical calculation of target gene concentration. Concentrations within the  
383 environmental samples ranged from 0 to 8.8 copies/ $\mu$ L for *Clupea harengus* and 0-2.5 copies/ $\mu$ L for  
384 *Gadus morhua*, with an overall mean concentration of 0.06 copies/ $\mu$ L for both assays. Using the  
385 standard samples, we observed that the cloglog relationship demonstrated a one-to-one correspondence  
386 between the proportion of positive droplets and the measured DNA concentration (Figure S2).  
387 Furthermore, the ddPCR measurements underestimated the concentration of *Clupea harengus* target  
388 assay in standard samples by a factor of two compared to the initial concentrations. We denote that the  
389 proportion of positive droplets for standard samples of *Clupea harengus* with a known concentration of  
390 0.1 copies/ $\mu$ L were higher than expected (explained in detail in [Guri et al., 2024](#), in prep). These  
391 samples were therefore considered outliers and removed from the analysis.

392

### 393 **Metabarcoding**

394 The metabarcoding analysis yielded a total of 29,997,240 reads (Table S3) prior to undergoing  
395 bioinformatic filtering. Among these, just under 14 million reads were attributed to *Poecilia sphenops*  
396 (artificial DNA spiked into the samples for a separate research project). The remaining reads,  
397 encompassing over 12 million reads, were allocated to the intended target sequences. Of this subset, the

398 mock community samples yielded 4,771,331 reads, while the negative control samples, both from field  
399 and laboratory, contributed 26,906 reads (Table S3). The remaining 7,635,648 reads were derived from  
400 the field environmental samples (Table S3). Among these, nearly all reads, totalling ca. 99.2 % of all  
401 reads were assigned to bony fishes (Actinopterygii) while the remaining were distributed among  
402 Mammals (0.6%), Chondrichthyes (0.027%), Bacteria (0.2%) and Aves (<0.01%). After the filtration  
403 and taxonomic annotation process, we identified a total of 50 unique Molecular Operational Taxonomic  
404 Units (MOTUs) across the environmental and negative control samples. Among these MOTUs, 39 were  
405 attributed to bony fishes, seven to Mammalians, two to sharks and rays, and one each to birds and  
406 bacteria (all bacterial sequences consolidated into a single MOTU). Subsequently, the study's  
407 sequencing depth was sufficient in capturing the taxonomic complexity of the study site, with an  
408 average of 90,000 reads per sample surpassing the rarefaction curve's saturation point of ca. 30,000  
409 reads (Figure S3). Species such as *Homo sapiens*, *Crystallogobius linearis*, *Gasterosteus aculeatus*,  
410 *Zeugopterus norvegicus*, *Rangifer tarandus*, *Sus scrofa* were removed from metabarcoding dataset as  
411 their read abundance exceeded the 10 % proportion limit between negative controls and environmental  
412 samples.

413

#### 414 **Mock community**

415 The mock community samples encompassed approximately 4.8 million reads after the filtering process,  
416 with 98.6% of the reads originating from the ten fish species belonging to the constructed mock  
417 community. In seven of the ten species in the mock community, we observed minimal differences in  
418 proportions were observed between initial concentrations (pre-PCR) and sequencing read abundances  
419 (post-PCR), signifying similar PCR amplification efficiencies between them. Conversely, amplicons  
420 attributed to *Leptoclinus maculatus*, *Cyclopterus lumpus*, and *Mallotus villosus* had significantly  
421 different PCR amplification efficiency with *Leptoclinus maculatus* being the most efficiently amplified  
422 sequence increasing the abundance from 5.5% in the pre-PCR pool to 24% post-PCR (Figure S4). In  
423 contrast, the proportional abundance of *Cyclopterus lumpus* decreased from 18% pre-PCR to 8% post-  
424 PCR, indicating low amplification efficiency of this species (Figure S4).

425

#### 426 **Joint model results**

427 The joint Bayesian model resulted in  $\hat{R}$  values lower than 1.005 indicating that the chains converged to  
428 a common distribution, ensuring the stability of the outcomes. Additionally, the model had low  
429 autocorrelation thus indicating efficient mixing as the effective sample sizes (ESS) were higher than ca.  
430 2000. Posterior summaries of parameters together with  $\hat{R}$  and ESS can be found in Table S8.

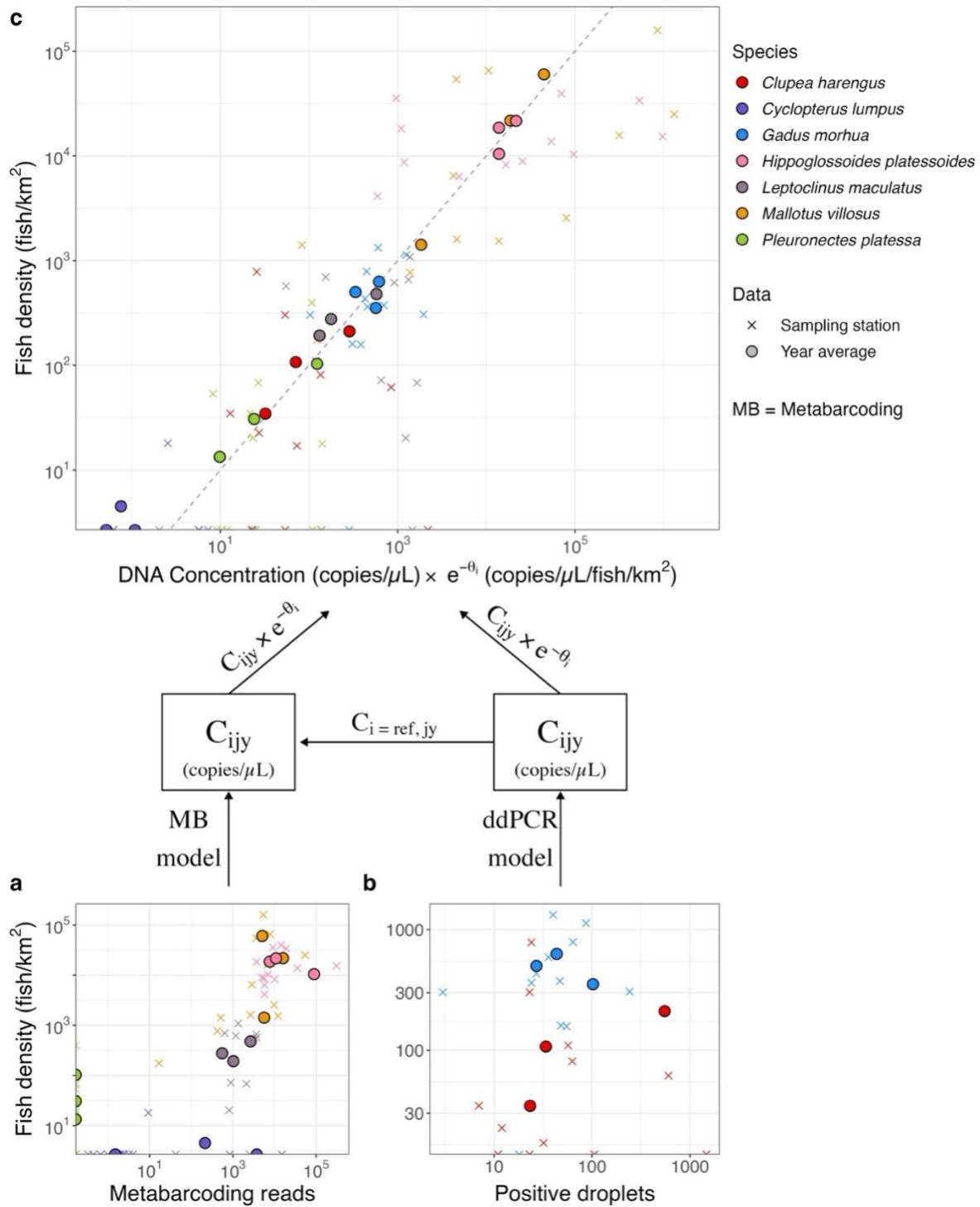
431

432 Posterior predictive checks consistently indicate substantial agreement between predicted and observed  
433 data across all three observation methods: ddPCR, metabarcoding, and trawl surveys (Figure S5 a, b,  
434 and c respectively). This result indicates the model's capacity to accurately capture the characteristics

435 of the observed data (relationship and uncertainty) and its ability to extrapolate meaningful predictions  
436 forecasting future outcomes and deepening the understanding of the connections between these three  
437 observation methods.

438

439 The DNA concentration data generated by the eDNA models (ddPCR and metabarcoding) indicated a  
440 strong correlation with the observed fish density on the trawl data (Figure 2 c) given the species-specific  
441 conversion factor ( $\theta$ ). Such correlation indicates the reliability of the latent parameters ( $\theta$ ,  $\beta_0$ ,  $\beta_1$ ,  $\alpha$ ,  $\eta$ ,  
442 and  $\tau$ ) in accurately representing the intricate relationships between the underlying processes and the  
443 observed data. This empirical evidence reinforces the dependency of the species-specific conversion  
444 factor ( $\theta$ ; Table S7) for linking eDNA observation with other conventional methods (i.e., trawl catches)  
445 as the explanatory variable for underlying biological mechanisms between organisms and their shed  
446 and captured DNA.



447

448

449 *Figure 2. The joint Bayesian model workflow translating metabarcoding reads (a) and ddPCR droplets*  
 450 *(b) into fish density (expressed as DNA concentration \* conversion factor; c) and their correlation with*  
 451 *the fish density estimated from trawl observations in the y-axis. C = DNA concentration and θ =*  
 452 *conversion factor between fish density to DNA concentrations. Metabarcoding and ddPCR models*  
 453 *indicated here are model compartments within the joint Bayesian model. The plot (c) indicates model*  
 454 *fit and reliable parameter estimation for linking trawl and eDNA observations.*

455

456 The model output of the conversion factor ( $\theta$ ) ranged considerably (Table 2) from -30 for *Pleuronectes*  
 457 *platessa* to -5.95 for *Clupea harengus* (with  $e^{-\theta}$  copies/ $\mu$ L equalling 1 fish/km<sup>2</sup>). The posterior  
 458 distribution of the conversion factor differed significantly between the remaining species and varied  
 459 between  $e^{-8}$  and  $e^{-14}$  copies/ $\mu$ L per fish/km<sup>2</sup>. We note that we observed no reads of *Pleuronectes platessa*  
 460 in the metabarcoding data (Table S7) thus the  $\theta$  values for that species is being estimated at the lowest  
 461 boundary possible by the model. Similarly, due to extremely low catch counts, the conversion parameter  
 462 of *Cyclopterus lumpus* is very hard for the model to estimate, thus a very high standard deviation of the  
 463 parameter's posterior distribution was observed. These species represent lack of data scenarios and thus  
 464 including such results reveals some of the pitfalls and model prerequisites.

466 *Table 2. Species-specific  $\theta$  parameter estimated by the joint model (10,000 iterations), with the mean,*  
 467 *standard deviation (SD), and 95% credibility intervals (CI).*

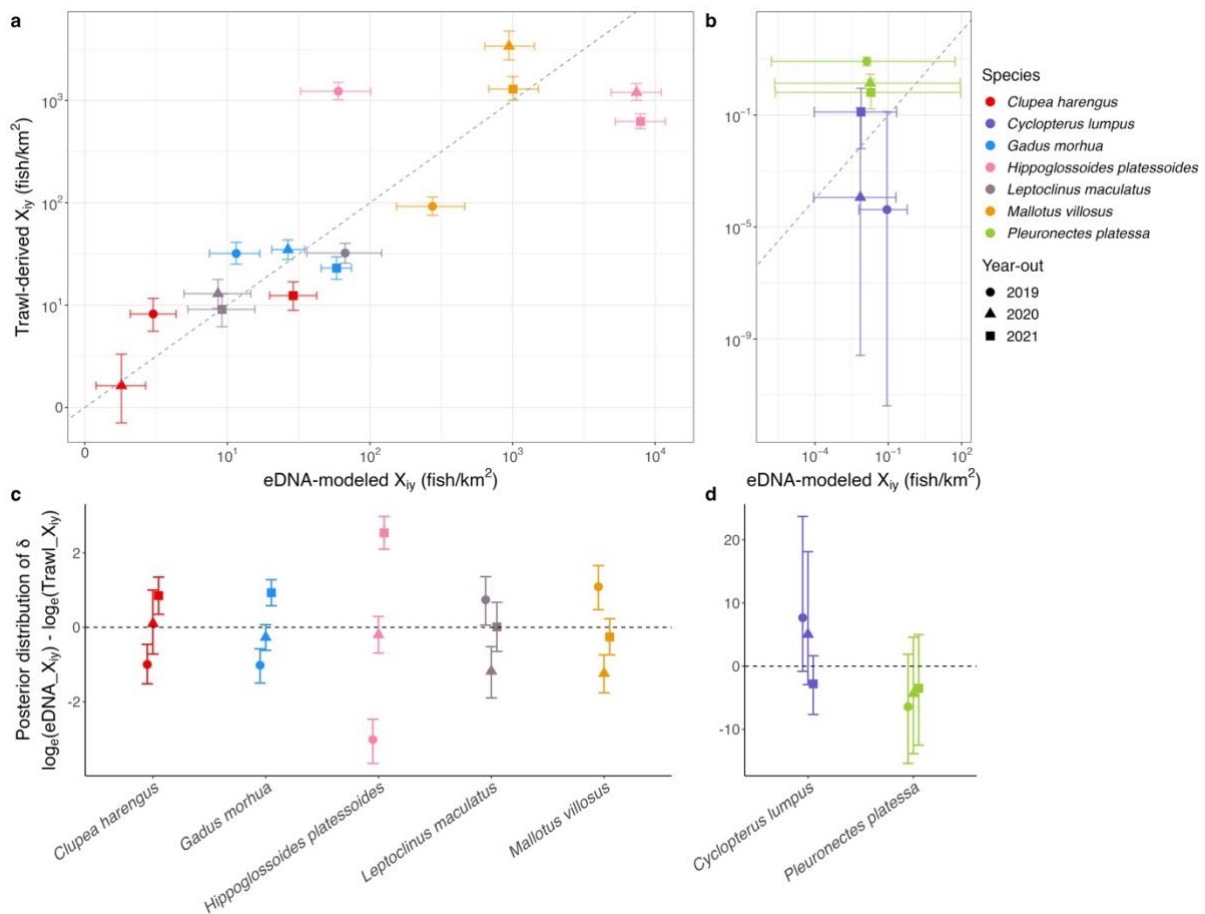
Species	Posterior distribution of $\theta$		95% CI	
	Estimate (mean)	SD	lower	upper
<i>Gadus morhua</i>	-8.09	0.09	-8.40	-8.06
<i>Clupea harengus</i>	-5.95	0.13	-6.22	-5.69
<i>Mallotus villosus</i>	-12.96	0.14	-13.24	-12.67
<i>Cyclopterus lumpus</i>	-8.93	1.13	-10.72	-6.33
<i>Leptoclinus maculatus</i>	-14.56	0.17	-14.89	-14.22
<i>Hippoglossoides platessoides</i>	-12.28	0.15	-12.57	-12.00
<i>Pleuronectes platessa</i>	-30.84	3.9	-39.70	-24.43

468  
 469 The out-of-sample analysis showed that molecular data alone reliably predicted trawl catches (Figure  
 470 3). Five out of seven species demonstrated a strong correlation between the predicted and observed fish  
 471 densities, thus resulting in low  $\delta$ -values (Figure 3 a and c). Among these, four species (i.e., *Clupea*  
 472 *harengus*, *Gadus morhua*, *Leptoclinus maculatus*, and *Mallotus villosus*) resulted in  $\delta$ -values of less  
 473 than or equal to one for each predicted year, underscoring the robustness of the parameters governing  
 474 the underlying relationships. Notably, *Hippoglossoides platessoides* displayed higher variability  
 475 compared to the other four species, suggesting slightly less robust and reproducible parameters among  
 476 different years. This variation may be attributed to an outlier year, although it is challenging to  
 477 definitively ascertain this due to the limited dataset of n=3 years in total (and therefore n = 2 years in  
 478 training-data subsets). Increased availability of coupled data (trawl and eDNA) could enhance the  
 479 robustness of parameter estimation.

480



481 The remaining two species (i.e., *Cyclopterus lumpus* and *Pleuronectes platessa*) offer a useful  
 482 illustration: where insufficient data are available, no meaningful inferences can be made. These species  
 483 showed a notable average discrepancy of two orders of magnitude between the eDNA model-predicted  
 484 fish densities and those derived from trawl observations (Figure 3 b and d). Furthermore, the mean  
 485 within-year-out variability spanned over eight orders of magnitude, signifying the data insufficiency in  
 486 either the trawl or the eDNA modules yields challenges in estimating parameters. Important to note that  
 487  $\delta$  values indicate the log-fold change in fish density, with  $\delta = 0$  denoting similar estimations between  
 488 eDNA and trawl, negative  $\delta$  indicating overestimation by trawl, and vice versa.  
 489



490  
 491 *Figure 3. Leave-one-year-out analysis indicating the correlation between the predicted (x-axis) and*  
 492 *observed fish densities (y-axis; a and b). The species-specific posterior distribution of  $\delta$  variable ( $\pm$*   
 493 *95% confidence interval of the posterior distribution) indicates the difference between predicted and*  
 494 *observed fish densities (c and d). Due to the infrequent presence in the observed data species*  
 495 **Cyclopterus lumpus* and *Pleuronectes platessa* are considered outliers and thus presented separately*  
 496 *(b and d).*

497  
 498 **Discussion**

499 The current study has successfully bridged the gap between eDNA and trawl surveys, quantifying the  
 500 biological relationships between fish abundance and DNA concentration in the marine environment.

501 Even more significantly, this bridge has been extended to multiple species, thereby enhancing the  
502 versatility and applicability of eDNA methods for ecosystem-based management.

503

#### 504 **Biological interpretation of the model**

505 Given the inextricable nature of DNA shedding and degradation, DNA transport and dilution, and DNA  
506 recovery and isolation processes, we unified them under the comprehensive term “integrated DNA  
507 factor –  $\lambda$ ” in this study ([see the framework in Shelton \*et al.\*, 2016](#)). We emphasize that the processes  
508 of catchability ( $q$ ) and the integrated DNA factor ( $\lambda$ ) are not assessed through this model but have been  
509 circumvented using the trawl to DNA conversion parameter ( $\theta$ ). Below, we discuss how these processes  
510 biologically interplay with the latent parameters estimated in this model.

511

512 The high replicability of  $\theta$  for most species - as indicated by small values of  $\delta$  in leave-one-out analysis  
513 - demonstrates that trawl, metabarcoding, and ddPCR observations yielded measurable and reliable  
514 results using the established model. Replicable results support the assumptions that trawl catchability  
515 ( $q$ ) and the integrated DNA factor ( $\lambda$ ) have remained constant throughout the survey period. Our  
516 rationale for assuming constant trawl catchability is rooted in the standardized nature of the annual trawl  
517 surveys conducted by the IMR (Institute of Marine Research) and similar environmental and biological  
518 factors (Zhang *et al.*, 2020).

519

520 The joint modelling of trawl and eDNA observations allows us to enhance the spatial resolution of fish  
521 densities. Following established research protocols, we model fish density as the annual average ( $V$ ),  
522 based on four trawling catch count observations (assuming homogenous fish densities among stations),  
523 given the trawling effort. This approach becomes necessary because trawls at a specific location lack  
524 replication thus catches are commonly averaged over larger survey areas (Beare *et al.*, 2005).  
525 Conversely, when we complement trawl observations using eDNA data, we gain the ability to develop  
526 models at the site level ( $X$ ; and potentially future models at the depth level), rather than restricting  
527 eDNA analysis to the large aggregated (i.e., fjord level in our study). We denote that this flexibility  
528 offers a twofold advantage. First, it enables the ability to incorporate replicates of observation at  
529 different unit levels (i.e., site, area, station, and depth) thus increasing the accuracy of measurements.  
530 Second, it enables the model to establish the link between eDNA concentrations and fish densities  
531 without requiring a fixed relationship at every individual station, a factor that has been the downfall of  
532 previous studies attempting to establish this connection (Pont *et al.*, 2022; Veron *et al.*, 2023). This  
533 matches our understanding of the ecology of eDNA, in which DNA is transported among stations  
534 (Rourke *et al.*, 2022), thus this model advantage is particularly important as it mitigates issues related  
535 to lateral transport and dilution of eDNA (Andruszkiewicz *et al.*, 2019), corresponding in maintaining  
536 the consistency of  $\lambda$ . Additionally, factors of eDNA recovery and isolation efficiency must have  
537 remained the same as the sampling technique and the laboratory workflow have remained unchanged

538 among the entire set of samples. Lastly, since all observations were conducted in October of different  
539 years, we posit that the behaviours of the fish population must have remained largely consistent, as did  
540 the relevant abiotic factors, both responsible for DNA shedding and degradation (Goldberg *et al.*, 2011;  
541 Strickler *et al.*, 2015; Jo *et al.*, 2019; Mauvisseau *et al.*, 2021). These results lead us to conclude that  
542 the integrated DNA factor ( $\lambda$ ) has likely remained unchanged, particularly for species with low  $\delta$  values.

543  
544 For species with high  $\delta$  values, several factors may explain variations of the  $\theta$  parameter over different  
545 years, which in turn could indicate either changes in  $q$  or  $\lambda$  between those years of observation. In the  
546 case of *Cyclopterus lumpus*, the metabarcoding data suggest relatively low quantities, but trawl surveys  
547 observed minimal to no catches, suggesting low fish densities, and consequently pointing to an  
548 exceptionally low catchability term. This catchability term has been previously observed and indicated  
549 that bottom trawl surveys rarely catch *Cyclopterus lumpus*, especially at depths of more than 60 m  
550 (Kennedy *et al.*, 2016). Conversely pelagic trawls have been indicating higher catch rates of such  
551 species (Eriksen *et al.*, 2014). Consequently, having unidentifiable catchability term can hinder the  
552 parametrization of  $\theta$ . Conversely, flatfish species like *Hippoglossoides platessoides* and *Pleuronectes*  
553 *platessa* exhibit high catchability due to their preference for sandy and muddy bottoms. Trawl  
554 observations revealed relatively high fish densities for these species. However, in the case of  
555 *Hippoglossoides platessoides*, no reads were observed through metabarcoding data, indicating an  
556 apparent absence of eDNA in the water and rendering the  $\lambda$  term unidentifiable. An additional cause to  
557 this could be our filtering pipeline as we removed amplicons with lower than 50 reads, thus considered  
558 rare species (Jerde *et al.*, 2016). Nevertheless, *Pleuronectes platessa* exhibited high read numbers and  
559 relatively high eDNA concentrations, but its  $\delta$  values were substantial, indicating variation in the  $\theta$   
560 parameter between each year of observation. This presents a more complex problem, and additional  
561 observations may be necessary to identify the source of this issue. We hypothesize that two potential  
562 errors could contribute to this problem. First, there may be an abnormality in the catchability terms in  
563 one of the years, making the outlier year unidentifiable given only three years of observations. Second,  
564 using water samples instead of sediment samples could introduce biases for benthic species in the  
565 processes encompassed by the integrated DNA factor, such as DNA vertical transport and DNA  
566 recovery.

567  
568 The variability among species highlights a general point that the details of biology matter for the  
569 interpretation and use of eDNA observations, and that the use of eDNA for biomass estimation is likely  
570 species- and context- specific. Additional applications using eDNA information need to be developed  
571 to understand the kinds and magnitudes of uncertainty associated with using eDNA information in  
572 different contexts and with different species. Having a deeper understanding of these parameters and  
573 estimating their empirical value can enhance our biological understanding of the molecular sampling  
574 tools, thus boosting the applicability of eDNA to fisheries management and stock assessments (Jo *et*

575 *al.*, 2019). However, our work provides a roadmap for moving away from simple sample-to-sample  
576 correlations among sampling methods and toward a method that can inform and improved management  
577 and conservation actions. In this study, the specific conversion parameters are associated with the  
578 catchability of bottom trawling (i.e., the method's inherent bias), and therefore, their applicability cannot  
579 be broadly extended to other methods or study areas. However, the statistical framework developed is  
580 very general and can be adapted to other sampling methods.

581

### 582 **The future of eDNA and trawl joint model**

583 To the best of our knowledge, this is the first empirical demonstration of species-specific conversion  
584 factors between eDNA metabarcoding and bottom trawl datasets. Future refinements will expand upon  
585 the existing model to include alternative sampling techniques and further generalize the framework we  
586 have presented here.

587

588 One potential area for improvement is the external estimation of trawl catchability ( $q$ ; Fraser *et al.*,  
589 2007), which can not only lead to a more robust conversion parameter  $\theta$  but also provide a direct  
590 estimation of the integrated DNA factor ( $\lambda$ ). Additionally, by incorporating environmental covariates  
591 such as temperature and salinity into the model, we can quantify how these abiotic factors interact with  
592 DNA shedding and degradation processes (Strickler *et al.*, 2015; Jo *et al.*, 2019), thus further  
593 disaggregating the components of  $\lambda$ . Including depth information in the model can allow for a more  
594 comprehensive understanding of species distribution in the water column thus depth- and species-  
595 specific DNA recovery, providing insights that extend beyond what can be captured through trawling  
596 alone (Jeunen *et al.*, 2019; Canals *et al.*, 2021). Furthermore, considering the influence of allometric  
597 scaling on DNA shedding, metabolic rate, consumption, and excretion, in conjunction with  
598 environmental factors and ecosystem characteristics, can offer a more holistic understanding of the  
599 relationships between biology and the latent parameters (Urban *et al.*, 2023). Having a deeper  
600 understanding of allometric scaling can ultimately enable molecular tools to estimate the size  
601 distribution of populations in marine environments (Yates *et al.*, 2022), a crucial step for population  
602 assessments.

603

604 The model can also be adapted to accommodate alternative monitoring techniques and data inputs. For  
605 example, catch-release, acoustic surveys, and trawl biomass data could either replace or be used in  
606 conjunction with the trawl model. Incorporating these additional methods would expand the range of  
607 available data collection techniques and provide further insight into the biological processes that  
608 underpin the molecular tools used in ecological assessments. Although this may alter the interpretation  
609 of parameters such as  $\theta$  or  $\lambda$ , it would contribute to a more comprehensive understanding of their  
610 underlying biology. Our work shows how eDNA can inform and be included in the class of integrated

611 analyses which are the foundation for modern stock assessments and fisheries management (Maunder  
612 and Punt, 2013).

613

614 In summary, the future of this model holds significant potential for development and improvement. By  
615 meticulously documenting and mapping all parameters used in the model, along with the specific  
616 biological processes they represent, we can increase transparency and facilitate comparisons with other  
617 studies. This, in turn, enhances the replicability and reliability of surveys, a capability that conventional  
618 surveys alone cannot offer. Ultimately, these advances will contribute to the development of molecular  
619 sampling tools and their applicability in fisheries management and stock assessments.

620

### 621 **Management applications**

622 Trawl surveys provide consistent time-series data essential for stock and ecosystem assessments.  
623 However, the reliability and continuity of these time series can be influenced by various factors. Area  
624 accessibility (i.e., closed areas for conservation or energy transition), the necessity to explore new areas  
625 as fish stocks migrate, funding limitations, vessel availability, and adverse weather conditions can all  
626 impact the consistency of data collection. In response to these challenges, there is a need to adapt and  
627 innovate in the field of survey methodologies. Embracing innovative molecular tools and techniques  
628 and coupling them with conventional methods can provide more flexibility and resilience to changing  
629 circumstances while significantly reducing the ecological impact and maximizing the efficiency of  
630 surveys (Schneider *et al.*, 2016; Di Muri *et al.*, 2020; Veron *et al.*, 2023). This strategic fusion can  
631 enhance the robustness and adaptability of survey programs, thereby strengthening our capacity for  
632 effective stock and ecosystem assessments in an ever-changing marine environment.

633

### 634 **Conclusions**

635 This study establishes the framework for species quantification and ecological inference of eDNA  
636 metabarcoding. We stress the need for species-specific conversion parameters to accurately estimate  
637 species abundances from eDNA metabarcoding. Furthermore, accounting for amplification efficiencies  
638 is necessary for ensuring the accuracy and replicability of metabarcoding data. These insights hold  
639 considerable implications for the application of eDNA in the management and conservation of marine  
640 ecosystems.

641 **References**

- 642 Afzali, S. F., Bourdages, H., Laporte, M., Mérot, C., Normandeau, E., Audet, C., and Bernatchez, L.  
643 2021. Comparing environmental metabarcoding and trawling survey of demersal fish  
644 communities in the Gulf of St. Lawrence, Canada. *Environmental DNA*, 3: 22–42.
- 645 Allan, E. A., Kelly, R. P., D’Agnese, E. R., Garber-Yonts, M. N., Shaffer, M. R., Gold, Z. J., and  
646 Shelton, A. O. 2023. Quantifying impacts of an environmental intervention using  
647 environmental DNA. *Ecological Applications*, 33: e2914.
- 648 Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., and  
649 Boehm, A. B. 2019. Modeling environmental DNA transport in the coastal ocean using  
650 Lagrangian particle tracking. *Frontiers in Marine Science*, 6: 1–14.
- 651 Atienza, S., Guardiola, M., Præbel, K., Antich, A., Turon, X., and Wangenstein, O. S. 2020. DNA  
652 metabarcoding of deep-sea sediment communities using COI: community assessment, spatio-  
653 temporal patterns and comparison with the 18S rDNA marker. *Diversity*, 12: 123.
- 654 Beare, D. J., Needle, C. L., Burns, F., and Reid, D. G. 2005. Using survey data independently from  
655 commercial data in stock assessment: an example using haddock in ICES Division VIa. *ICES*  
656 *Journal of Marine Science*, 62: 996–1005.
- 657 Biju Kumar, A., and Deepthi, G. R. 2006. Trawling and by-catch: Implications on marine ecosystem.  
658 *Current Science*, 90: 922–931.
- 659 Callaghan, C. T., Nakagawa, S., and Cornwell, W. K. 2021. Global abundance estimates for 9,700 bird  
660 species. *Proceedings of the National Academy of Sciences*, 118: e2023170118.
- 661 Canals, O., Mendibil, I., Santos, M., Irigoien, X., and Rodríguez-Ezpeleta, N. 2021. Vertical  
662 stratification of environmental DNA in the open ocean captures ecological patterns and  
663 behavior of deep-sea fishes. *Limnology And Oceanography Letters*: 1–9.
- 664 Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M.,  
665 Bograd, S. J., *et al.* 2019. Marine Vertebrate Biodiversity and Distribution Within the Central  
666 California Current Using Environmental DNA (eDNA) Metabarcoding and Ecosystem  
667 Surveys. *Frontiers in Marine Science*, 6: 1–17.
- 668 Crossin, G. T., Heupel, M. R., Holbrook, C. M., Hussey, N. E., Lowerre-Barbieri, S. K., Nguyen, V.  
669 M., Raby, G. D., *et al.* 2017. Acoustic telemetry and fisheries management. *Ecological*  
670 *Applications*, 27: 1031–1049.
- 671 Di Muri, C., Handley, L. L., Bean, C. W., Li, J., Peirson, G., Sellers, G. S., Walsh, K., *et al.* 2020. Read  
672 counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass  
673 in drained ponds. *Metabarcoding and Metagenomics*, 4: 97–112.
- 674 Eigaard, O. R., Bastardie, F., Hintzen, N. T., Buhl-Mortensen, L., Buhl-Mortensen, P., Catarino, R.,  
675 Dinesen, G. E., *et al.* 2017. The footprint of bottom trawling in European waters: distribution,  
676 intensity, and seabed integrity. *ICES Journal of Marine Science*, 74: 847–865.

677 Eriksen, E., Durif, C. M. F., and Prozorkevich, D. 2014. Lumpfish (*Cyclopterus lumpus*) in the Barents  
678 Sea: development of biomass and abundance indices, and spatial distribution. *ICES Journal of*  
679 *Marine Science*, 71: 2398–2402.

680 Farr, M. T., O'Brien, T., Yackulic, C. B., and Zipkin, E. F. 2022. Quantifying the conservation status  
681 and abundance trends of wildlife communities with detection–nondetection data. *Conservation*  
682 *Biology*, 36: e13934.

683 Fraser, H. M., Greenstreet, S. P. R., and Piet, G. J. 2007. Taking account of catchability in groundfish  
684 survey trawls: implications for estimating demersal fish biomass. *ICES Journal of Marine*  
685 *Science*, 64: 1800–1819.

686 Gelman, A., and Rubin, D. B. 1992. Inference from Iterative Simulation Using Multiple Sequences.  
687 *Statistical Science*, 7. [https://projecteuclid.org/journals/statistical-science/volume-7/issue-](https://projecteuclid.org/journals/statistical-science/volume-7/issue-4/Inference-from-Iterative-Simulation-Using-Multiple-Sequences/10.1214/ss/1177011136.full)  
688 [4/Inference-from-Iterative-Simulation-Using-Multiple-](https://projecteuclid.org/journals/statistical-science/volume-7/issue-4/Inference-from-Iterative-Simulation-Using-Multiple-Sequences/10.1214/ss/1177011136.full)  
689 [Sequences/10.1214/ss/1177011136.full](https://projecteuclid.org/journals/statistical-science/volume-7/issue-4/Inference-from-Iterative-Simulation-Using-Multiple-Sequences/10.1214/ss/1177011136.full) (Accessed 1 December 2023).

690 Gold, Z., Shelton, A. O., Casendino, H. R., Duprey, J., Gallego, R., Van Cise, A., Fisher, M., *et al.*  
691 2023. Signal and noise in metabarcoding data. *PLOS ONE*, 18: e0285674.

692 Goldberg, C. S., Pilliod, D. S., Arkle, R. S., and Waits, L. P. 2011. Molecular Detection of Vertebrates  
693 in Stream Water: A Demonstration Using Rocky Mountain Tailed Frogs and Idaho Giant  
694 Salamanders. *PLoS ONE*, 6: e22746.

695 Guri, G., Westgaard, J., Yoccoz, N., Wangensteen, O. S., Præbel, K., Ray, J. L., Kelly, R. P., *et al.*  
696 2023. Maximizing sampling efficiency to detect differences in fish community composition  
697 using environmental DNA metabarcoding in subarctic fjords. *Environmental DNA*: edn3.409.

698 Guri, G., Shelton, A. O., Kelly, R. P., Ray, J., Louise Rayica Louise, Yoccoz, N., Torild, J., Præbel, K.,  
699 *et al.* 2024. Quantifying the sensitivity between qPCR and ddPCR mechanisms for eDNA  
700 samples. *Methods in Ecology and Evolution*.

701 Hansen, B. K., Bekkevold, D., Clausen, L. W., and Nielsen, E. E. 2018. The sceptical optimist:  
702 challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish*  
703 *and Fisheries*, 19: 751–768.

704 He, X., Jeffery, N. W., Stanley, R. R. E., Hamilton, L. C., Rubidge, E. M., and Abbott, C. L. 2023.  
705 eDNA metabarcoding enriches traditional trawl survey data for monitoring biodiversity in the  
706 marine environment. *ICES Journal of Marine Science*, 80: 1529–1538.

707 Hill, N. A., Barrett, N., Lawrence, E., Hulls, J., Dambacher, J. M., Nichol, S., Williams, A., *et al.* 2014.  
708 Quantifying Fish Assemblages in Large, Offshore Marine Protected Areas: An Australian Case  
709 Study. *PLoS ONE*, 9: e110831.

710 Jerde, C. L., Olds, B. P., Shogren, A. J., Andruszkiewicz, E. A., Mahon, A. R., Bolster, D., and Tank,  
711 J. L. 2016. Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate  
712 Environmental DNA. *Environmental Science & Technology*, 50: 8770–8779.

713 Jetz, W., McPherson, J. M., and Guralnick, R. P. 2012. Integrating biodiversity distribution knowledge:  
714 toward a global map of life. *Trends in Ecology & Evolution*, 27: 151–159.

715 Jeunen, G. J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., Bunce, M., *et al.*  
716 2019. Environmental DNA (eDNA) metabarcoding reveals strong discrimination among  
717 diverse marine habitats connected by water movement. *Molecular Ecology Resources*, 19: 426–  
718 438.

719 Jo, T., Murakami, H., Yamamoto, S., Masuda, R., and Minamoto, T. 2019. Effect of water temperature  
720 and fish biomass on environmental DNA shedding, degradation, and size distribution. *Ecology  
721 and Evolution*, 9: 1135–1146.

722 Kelly, R. P., Shelton, A. O., and Gallego, R. 2019. Understanding PCR Processes to Draw Meaningful  
723 Conclusions from Environmental DNA Studies. *Scientific Reports*, 9: 1–14.

724 Kennedy, J., Jónsson, S. Þ., Ólafsson, H. G., and Kasper, J. M. 2016. Observations of vertical  
725 movements and depth distribution of migrating female lumpfish (*Cyclopterus lumpus*) in  
726 Iceland from data storage tags and trawl surveys. *ICES Journal of Marine Science*, 73: 1160–  
727 1169.

728 Knudsen, S. W., Ebert, R. B., Hesselsøe, M., Kuntke, F., Hassingboe, J., Mortensen, P. B., Thomsen,  
729 P. F., *et al.* 2019. Species-specific detection and quantification of environmental DNA from  
730 marine fishes in the Baltic Sea. *Journal of Experimental Marine Biology and Ecology*, 510: 31–  
731 45. Elsevier.

732 Lacoursière-Roussel, A., Côté, G., Leclerc, V., and Bernatchez, L. 2016. Quantifying relative fish  
733 abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology*,  
734 53: 1148–1157.

735 Larson, W. A., Barry, P., Dokai, W., Maselko, J., Olson, J., and Baetscher, D. 2022. Leveraging eDNA  
736 metabarcoding to characterize nearshore fish communities in Southeast Alaska: Do habitat and  
737 tide matter? *Environmental DNA*: 1–13.

738 Lees, K. J., MacNeil, M. A., Hedges, K. J., and Hussey, N. E. 2021. Estimating demographic parameters  
739 for fisheries management using acoustic telemetry. *Reviews in Fish Biology and Fisheries*, 31:  
740 25–51.

741 Maes, S. M., Desmet, S., Brys, R., Sys, K., Ruttink, T., Maes, S., Hostens, K., *et al.* 2023. Detection  
742 and quantification of two commercial flatfishes *Solea solea* and *Pleuronectes platessa* in the  
743 North Sea using environmental DNA. *Environmental DNA*: edn3.426.

744 Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. 2014. Swarm: Robust and fast  
745 clustering method for amplicon-based studies. *PeerJ*, 2014: 1–13.

746 Mahévas, S., Trenkel, V. M., Doray, M., and Peyronnet, A. 2011. Hake catchability by the French  
747 trawler fleet in the Bay of Biscay: estimating technical and biological components. *ICES  
748 Journal of Marine Science*, 68: 107–118.



749 Maunder, M. N., and Punt, A. E. 2013. A review of integrated analysis in fisheries stock assessment.  
750 Fisheries Research, 142: 61–74.

751 Mauvisseau, Q., Harper, L., Sander, M., Hanner, R. H., Kleyer, H., and Deiner, K. 2021. The multiple  
752 states of environmental DNA and what is known about their persistence in aquatic  
753 environments. Authorea Preprints, preprint: 1–17.

754 Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., *et al.* 2015. MiFish,  
755 a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection  
756 of more than 230 subtropical marine species. Royal Society Open Science.

757 Mjanger, H., Svendsen, B. V., Senneset, H., Fotland, Å., Mehl, S., Fuglebakk, E., Gulbrandsen, M. L.,  
758 *et al.* 2019, January. Håndbok for prøvetaking av fisk, krepsdyr og andre evertebrater. Versjon  
759 5.0. (In Norwegian).

760 Nash, R. D. M., Wright, P. J., Matejusova, I., Dimitrov, S. P., O’Sullivan, M., Augley, J., and Höffle,  
761 H. 2012. Spawning location of Norway pout (*Trisopterus esmarkii* Nilsson) in the North Sea.  
762 ICES Journal of Marine Science, 69: 1338–1346.

763 Neebling, T. E., and Quist, M. C. 2011. Comparison of Boat Electrofishing, Trawling, and Seining for  
764 Sampling Fish Assemblages in Iowa’s Nonwadeable Rivers. North American Journal of  
765 Fisheries Management, 31: 390–402.

766 Pont, D., Meulenbroek, P., Bammer, V., Dejean, T., Erős, T., Jean, P., Lenhardt, M., *et al.* 2022.  
767 Quantitative monitoring of diverse fish communities on a large scale combining eDNA  
768 metabarcoding and qPCR. Molecular Ecology Resources: 1755–0998.13715.

769 Preston, F. W. 1948. The Commonness, And Rarity, of Species. Ecology, 29: 254–283.

770 Ramírez-Amaro, S., Bassitta, M., Picornell, A., Ramon, C., and Terrasa, B. 2022. Environmental DNA:  
771 State-of-the-art of its application for fisheries assessment in marine environments. Frontiers in  
772 Marine Science, 9: 1004674.

773 Reigstad, M., and Wassmann, P. 1996. Importance of advection for pelagic-benthic coupling in north  
774 Norwegian fjords. Sarsia, 80: 245–257.

775 Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. 2016. VSEARCH: A versatile open source  
776 tool for metagenomics. PeerJ, 2016: 1–22.

777 Rourke, M. L., Fowler, A. M., Hughes, J. M., Broadhurst, M. K., DiBattista, J. D., Fielder, S., Wilkes  
778 Walburn, J., *et al.* 2022. Environmental DNA (eDNA) as a tool for assessing fish biomass: A  
779 review of approaches and future considerations for resource surveys. Environmental DNA, 4:  
780 9–33.

781 Salter, I., Joensen, M., Kristiansen, R., Steingrund, P., and Vestergaard, P. 2019. Environmental DNA  
782 concentrations are correlated with regional biomass of Atlantic cod in oceanic waters.  
783 Communications Biology, 2: 1–9.

784 Schadewell, Y., and Adams, C. I. M. 2021. Forensics Meets Ecology – Environmental DNA Offers  
785 New Capabilities for Marine Ecosystem and Fisheries Research. *Frontiers in Marine Science*,  
786 8: 668822.

787 Schneider, J., Valentini, A., Dejean, T., Montarsi, F., Taberlet, P., Glazot, O., and Fumagalli, L. 2016.  
788 Detection of Invasive Mosquito Vectors Using Environmental DNA (eDNA) from Water  
789 Samples. *PLOS ONE*, 11: e0162493.

790 Shelton, A. O., O'Donnell, J. L., Samhouri, J. F., Lowell, N., Williams, G. D., and Kelly, R. P. 2016.  
791 A framework for inferring biological communities from environmental DNA. *Ecological*  
792 *Applications*, 26: 1645–1659.

793 Shelton, A. O., Ramón-Laca, A., Wells, A., Clemons, J., Chu, D., Feist, B. E., Kelly, R. P., *et al.* 2022.  
794 Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution  
795 in the open ocean. *Proceedings of the Royal Society B: Biological Sciences*, 289: 20212613.

796 Shelton, A. O., Gold, Z. J., Jensen, A. J., D'Agnese, E., Andruszkiewicz Allan, E., Van Cise, A.,  
797 Gallego, R., *et al.* 2023a. Toward quantitative metabarcoding. *Ecology*, 104.  
798 <https://onlinelibrary.wiley.com/doi/10.1002/ecy.3906> (Accessed 24 April 2023).

799 Shelton, A. O., Gold, Z. J., Jensen, A. J., D'Agnese, E., Andruszkiewicz Allan, E., Van Cise, A.,  
800 Gallego, R., *et al.* 2023b. Toward quantitative metabarcoding. *Ecology*, 104.  
801 <https://onlinelibrary.wiley.com/doi/10.1002/ecy.3906> (Accessed 2 April 2023).

802 Shu, L., Ludwig, A., and Peng, Z. 2020. Standards for methods utilizing environmental dna for detection  
803 of fish species. *Genes*, 11.

804 Stoeckle, M. Y., Adolf, J., Charlop-Powers, Z., Dunton, K. J., Hinks, G., and VanMorter, S. M. 2021.  
805 Trawl and eDNA assessment of marine fish diversity, seasonality, and relative abundance in  
806 coastal New Jersey, USA. *ICES Journal of Marine Science*, 78: 293–304.

807 Strickler, K. M., Fremier, A. K., and Goldberg, C. S. 2015. Quantifying effects of UV-B, temperature,  
808 and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, 183: 85–92.

809 Taylor, M. I., Fox, C., Rico, I., and Rico, C. 2002. Species-specific TaqMan probes for simultaneous  
810 identification of (*Gadus morhua* L.), haddock (*Melanogrammus aeglefinus* L.) and whiting  
811 (*Merlangius merlangus* L.). *Molecular Ecology Notes*, 2: 599–601.

812 Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., and Willerslev, E. 2012.  
813 Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples.  
814 *PLoS ONE*, 7: 1–9.

815 Thomsen, P. F., Møller, P. R., Sigsgaard, E. E., Knudsen, S. W., Jørgensen, O. A., and Willerslev, E.  
816 2016. Environmental DNA from Seawater Samples Correlate with Trawl Catches of Subarctic,  
817 Deepwater Fishes. *PLOS ONE*, 11: e0165252.

818 Turon, M., Nygaard, M., Guri, G., Wangensteen, O. S., and Præbel, K. 2022. Fine-scale differences in  
819 eukaryotic communities inside and outside salmon aquaculture cages revealed by eDNA  
820 metabarcoding. *Frontiers in Genetics*, 13: 957251.

821 Urban, P., Bekkevold, D., Degel, H., Hansen, B. K., Jacobsen, M. W., Nielsen, A., and Nielsen, E. E.  
822 2023. Scaling from eDNA to biomass: controlling allometric relationships improves precision  
823 in bycatch estimation. *ICES Journal of Marine Science*, 80: 1066–1078.

824 Valdivia-Carrillo, T., Rocha-Olivares, A., Reyes-Bonilla, H., Domínguez-Contreras, J. F., and  
825 Munguia-Vega, A. 2021. Integrating eDNA metabarcoding and simultaneous underwater  
826 visual surveys to describe complex fish communities in a marine biodiversity hotspot.  
827 *Molecular Ecology Resources*, 21: 1558–1574.

828 Veron, P., Rozanski, R., Marques, V., Joost, S., Deschez, M. E., Trenkel, V. M., Lorange, P., *et al.*  
829 2023. Environmental DNA complements scientific trawling in surveys of marine fish  
830 biodiversity. *ICES Journal of Marine Science*: fsad139.

831 Yates, M. C., Wilcox, T. M., Stoeckle, M. Y., and Heath, D. D. 2022. Interspecific allometric scaling  
832 in eDNA production among northwestern Atlantic bony fishes reflects physiological allometric  
833 scaling. *Environmental DNA*: edn3.381.

834 Zhang, C., Chen, Y., Xu, B., Xue, Y., and Ren, Y. 2020. Evaluating the influence of spatially varying  
835 catchability on multispecies distribution modelling. *ICES Journal of Marine Science*, 77: 1841–  
836 1853.

837  
838

839 **Extensive discussion**

840

841 **Metabarcoding**

842 The metabarcoding effort yielded sufficient sequencing depth to capture the taxonomic complexity of  
843 the study area, as the rarefaction curves revealed no indication of sequencing depth limitation despite  
844 the exogenous addition of a non-target internal control DNA (*Poecilia sphenops*) to most of the  
845 environmental samples for reasons not related to this study. Additionally, the metabarcoding model can  
846 be applied to a subset of taxa; thus, removing non-target taxa, for example *Poecilia sphenops*, would  
847 not affect the model outcomes (Shelton *et al.*, 2023). Moreover, the number of reads in negative control  
848 samples was on average less than 10% of the reads in field environmental samples, indicating low signs  
849 of field and laboratory contamination.

850

851 Reads deriving from mock community samples indicated a successful (clean) laboratory workflow as  
852 ca. 99% of the reads belonged to the species artificially put into the mixture - comprising the mock  
853 community. Moreover, the replicability between the six biological replicates of the mock community  
854 samples was very high. The relative abundance of the amplicons in each biological replicate sample  
855 after PCR had a difference of less than 2% compared with the average relative abundance among all  
856 replicates thus indicating that the PCR process was robust and replicable. Amplification efficiency ( $\alpha$ )  
857 ranged from -0.02 to 0.04 units, which were comparable to the Pacific Ocean fish study in Shelton *et*  
858 *al.*, 2023.

859

860 **ddPCR**

861 The connection between the proportion of positive droplets ( $\omega$ ) and the measured concentration  
862 highlighted the biases of QX200 ddPCR system formulas for measuring DNA concentrations. For low  
863 concentration standard samples, measured concentrations did not overlap with known concentrations,  
864 indicating a ddPCR system overestimation of low quantity samples (Figure S2). In addition to the low  
865 concentration overestimation, the FAM channel assay for *Clupea harengus* bias was indicated  
866 throughout the whole range of quantities ( $10^{-2} - 10^4$  copies/ $\mu$ L) of standard samples as the measured  
867 concentrations were half of the known concentrations (Figure S2). A potential reason for this issue  
868 could be the difference in primer/probe concentrations and/or duplexing of VIC and FAM assays  
869 simultaneously. Nevertheless, using known concentrations from standard samples alleviates this bias.

870

871 Although the ddPCR system is designed as an absolute quantification method, implying no need for  
872 standard samples and standard curves, the users often rely on standard samples for assay optimization  
873 (i.e., temperature selection and primer/probe concentration). Given these reasons, our result shows that  
874 including the standard samples can significantly elevate the accuracy of quantification and detection  
875 probability (Guri *et al.*, 2024) without significantly increasing the effort or costs.

876 **Supplementary tables**

877 *Table S2. Overview of ddPCR standard samples and their average number of positive droplets given*  
 878 *their nominal concentration. Samples of Clupea harengus with nominal concentration 0.1 copies/μL*  
 879 *were not included in this article due to being outliers.*

Species	Nominal concentration (copies/μL)	Number of sample replicates	Average proportion of positive droplets
<i>Clupea harengus</i>	0.01	8	0.000112
	0.10	8	0.000528
	1.00	8	0.000460
	10.00	8	0.00469
	100.00	8	0.0485
	1000.00	8	0.358
	10000.00	8	0.985
<i>Gadus morhua</i>	0.01	16	0.0000256
	0.10	16	0.0000786
	1.00	16	0.000945
	10.00	16	0.00862
	100.00	16	0.0807
	1000.00	16	0.577
	10000.00	16	1.00

880  
 881 *Table S3. Overview of the metabarcoding number of reads yielded and the number of samples analyzed*  
 882 *prior to the bioinformatic filtering process.*

Sample type	Number of samples	Number of raw reads	Number of total reads (blasted)	Spike DNA reads	Amplicon reads without spike DNA	Fish and shark (only) amplicon reads
Air field blank	3	649,466	643,072	642,892	180	38
Water field blank	9	1,552,126	1,469,292	1,436,494	32,798	20,844
PCR blank	2	345,276	341,788	341,524	264	256
Extraction blank	4	745,682	654,932	649,136	5,796	5,746
Laboratory blank	1	228,410	224,820	224,790	30	22
Bottom	28	7,749,155	6,374,325	3,141,004	3,233,321	3,220,209
Pycnocline	28	7,480,395	6,699,173	4,080,680	2,618,493	2,609,732
Surface	28	5,614,166	5,134,038	3,311,000	1,823,038	1,805,707
Mock community	6	5,781,694	4,771,360	0	4,771,360	4,771,331
Total	109	29,997,240	26,312,800	13,827,520	12,485,280	12,433,885

883 Table S4. Initial DNA concentration of species in mock community samples. Concentration measured  
 884 through Qubit quantification system using products for dsDNA concentrations. The measurement was  
 885 conducted three times, and the average value is shown in this table.

Species in the mock community	DNA concentration (copies/ $\mu$ L)
<i>Maurolicus muelleri</i>	7087
<i>Gadus morhua</i>	5942
<i>Mallotus villosus</i>	9816
<i>Leptoclinus maculatus</i>	3725
<i>Hippoglossoides platessoides</i>	6812
<i>Myoxocephalus scorpius</i>	8908
<i>Cyclopterus lumpus</i>	12061
<i>Pholis gunnellus</i>	4477
<i>Brosme brosme</i>	6022
<i>Pleuronectes platessa</i>	2637

886  
 887 Table S5 Overview of species in this study with their presence marked with “X” across the data source.  
 888 Only species that were included in the mock community or ddPCR standards samples are shown for  
 889 metabarcoding and trawl observation.

Species	Mock community	Metabarcoding	ddPCR standards	ddPCR	Trawl	Used in Bayesian Model
<i>Gadus morhua</i>	x	x	x	x	x	x
<i>Mallotus villosus</i>	x	x			x	x
<i>Cyclopterus lumpus</i>	x	x			x	x
<i>Leptoclinus maculatus</i>	x	x			x	x
<i>Hippoglossoides platessoides</i>	x	x			x	x
<i>Pleuronectes platessa</i>	x				x	x
<i>Clupea harengus</i>		x	x	x	x	x
<i>Maurolicus muelleri</i>	x	x				
<i>Myoxocephalus scorpius</i>	x	x				
<i>Pholis gunnellus</i>	x	x				
<i>Brosme brosme</i>	x	x				

890  
 891

892 *Table S6. Data, state processes, parameters, and subscripts employed in the joint Bayesian model and*  
 893 *their prior distributions (N stands for normal distribution with mean and standard deviation;  $\Gamma$  stands*  
 894 *for gamma distribution with shape and rate). N/A stands for data or processes that do not require prior*  
 895 *distributions.*

896  
 897

	Description	Prior
<b>Data</b>		
Z	observed number trawl catch count	N/A
E	trawling effort estimated in km <sup>2</sup>	N/A
W	positive droplets observed through ddPCR	N/A
P	total number of droplets accepted by ddPCR	N/A
S	known DNA concentration in c/μL	N/A
Y	observed number of reads through metabarcoding	N/A
N	total number of reads for all I species	N/A
alr	additive-log-ratio of initial concentration for all I species relative to the reference species prior to sequencing	N/A
NPCR	number of PCR cycles run	N/A
<b>State processes</b>		
N	true fish density in fish count/km <sup>2</sup>	N/A
X	trawled fish density in fish count/km <sup>2</sup>	N(0,10)
C	estimated eDNA concentration in copies/μL	N/A
V	trawled fish density in fish count/km <sup>2</sup> (averaged across stations)	N/A
<b>Parameters</b>		
$\theta$	conversion factor between trawled fish and eDNA concentration	N(0,10)
$\phi$	negative binomial distribution overdispersion parameter	$\Gamma(50,1)$
$\beta_0$	intercept of the linear relation between positive droplets and DNA concentration	N(0,10)
$\beta_1$	regression slope of the linear relation between positive droplets and DNA concentration	N(0,10)
$\alpha$	amplification efficiency	N(0,0.01)
$\eta$	multinomial distribution overdispersion parameter	N(0, $\tau$ )
$\tau$	standard deviation parameter for $\eta$	$\Gamma(100,1000)$
$\psi$	vector of probabilities for multinomial distribution	N/A
$\gamma$	vector the log-concentration of all I species relative to the reference species after sequencing	N/A
$\lambda$	integrated eDNA factor	N/A
q	trawl catchability	N/A
<b>Subscripts</b>		
i	species (n=7)	
j	station (n=4)	
y	year (n=3)	
d	depth (n=3)	
b	biological replicate (n=2 for year 2019 and 2020, n=3 for 2021)	
r	technical replicate (n=3 for ddPCR runs)	
p	standard aliquote sample (n=160)	
m	mock community aliquote sample (n=6)	

898

899

900 *Table S7: The species-specific average number of (i) reads in metabarcoding data (MB) and their*  
 901 *proportion (ii) catches in trawl data (iii) positive droplets in ddPCR data (iv) estimated concentration*  
 902 *(C) and (v) fish density (X) across all samples through the joint model. Note that  $\theta$  is in natural*  
 903 *logarithmic space thus  $X \approx C * e^{-\theta}$ .*

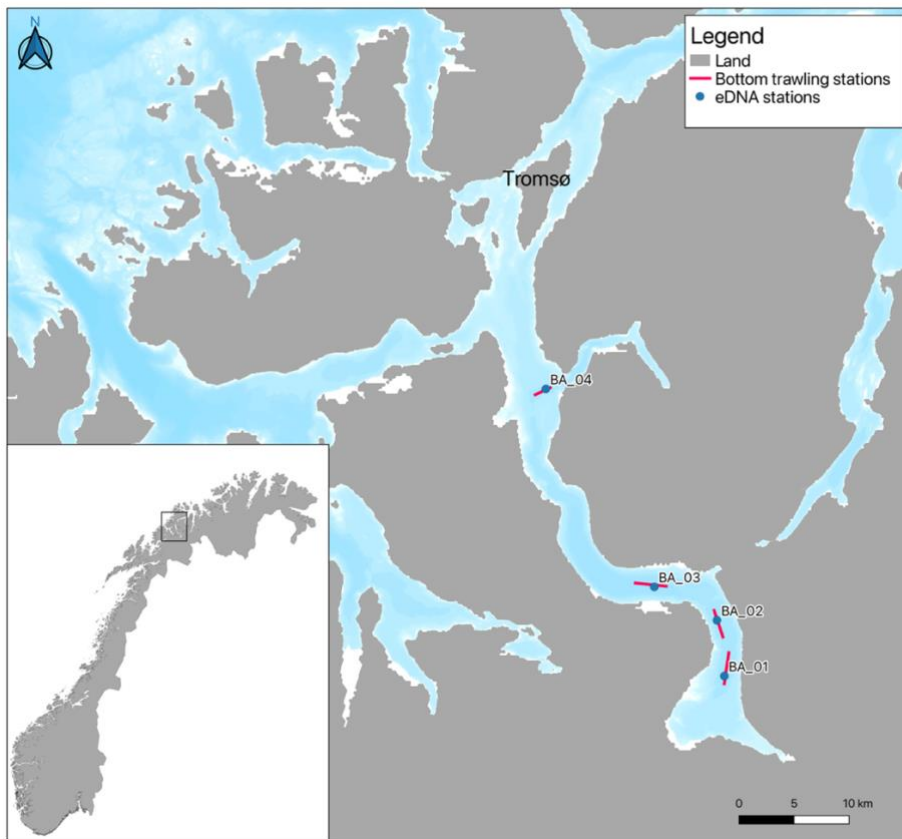
Species	MB reads	MB %	Trawl catches	ddPCR positive droplets	C (copies/ $\mu$ L)	$\theta$ (mean)	X (fish/km <sup>2</sup> )
<i>Gadus morhua</i>	116462	0.22	30	2.8	$e^{-2.04}$	-8.09	$e^{6.20}$
<i>Clupea harengus</i>			7.5	9.8	$e^{-1.49}$	-5.95	$e^{4.76}$
<i>Mallotus villosus</i>	69314	0.14	1861		$e^{-3.61}$	-12.96	$e^{10.20}$
<i>Cyclopterus lumpus</i>	8159	0.02	.083		$e^{-9.21}$	-8.93	$e^{0.40}$
<i>Leptoclinus maculatus</i>	9474	0.02	18		$e^{-9.08}$	-14.56	$e^{5.76}$
<i>Hippoglossoides platessoides</i>	305022	0.60	990		$e^{-2.58}$	-12.28	$e^{9.73}$
<i>Pleuronectes platessa</i>	0	0.00	3.7		$e^{-27.40}$	-30.84	$e^{3.89}$

904

905

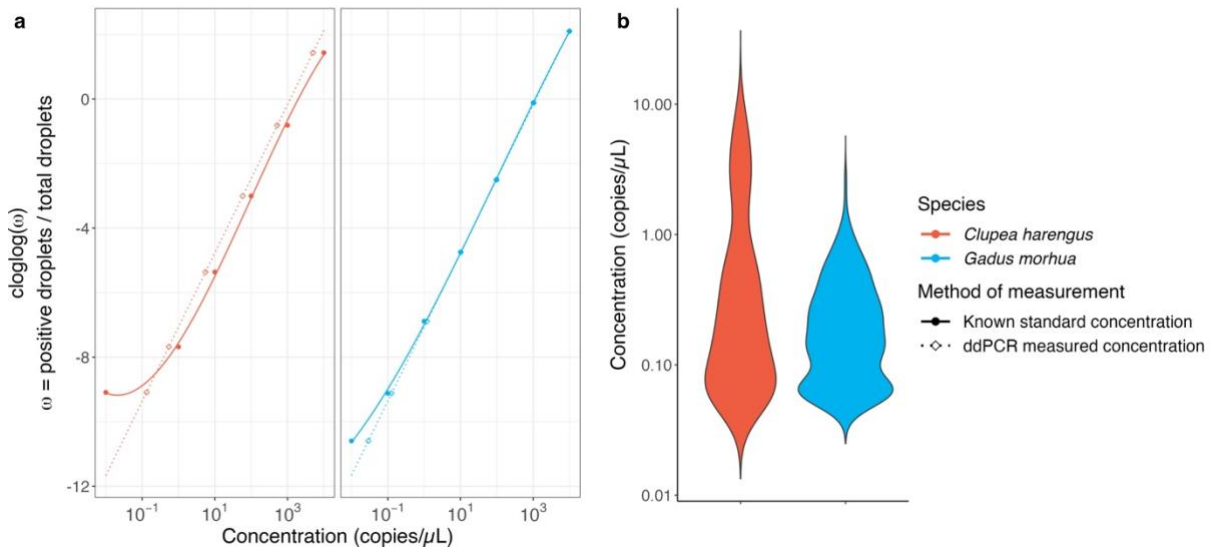


906 **Supplementary figures**

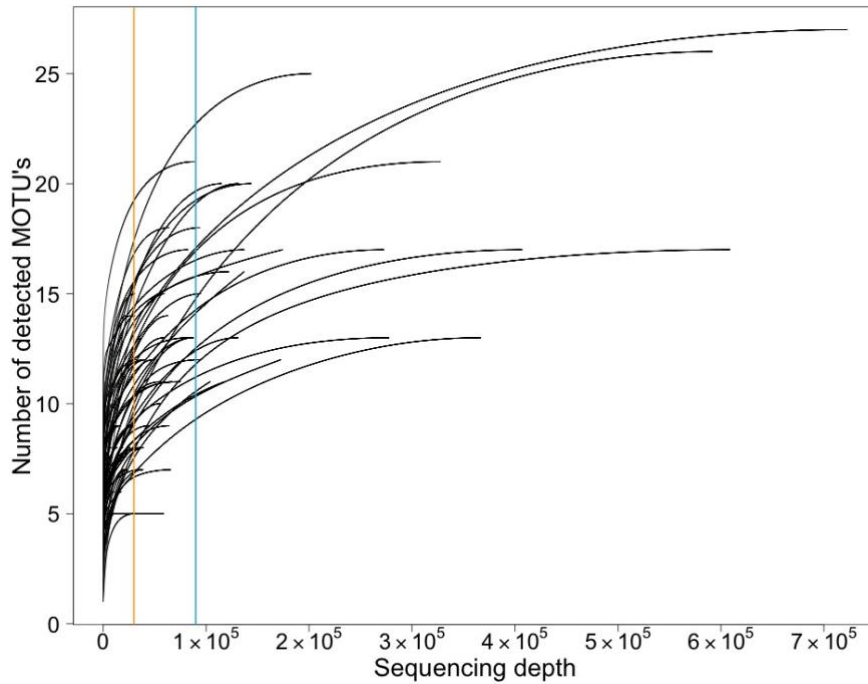


907  
908 *Figure S1. Map over trawl transects and eDNA sampling stations.*

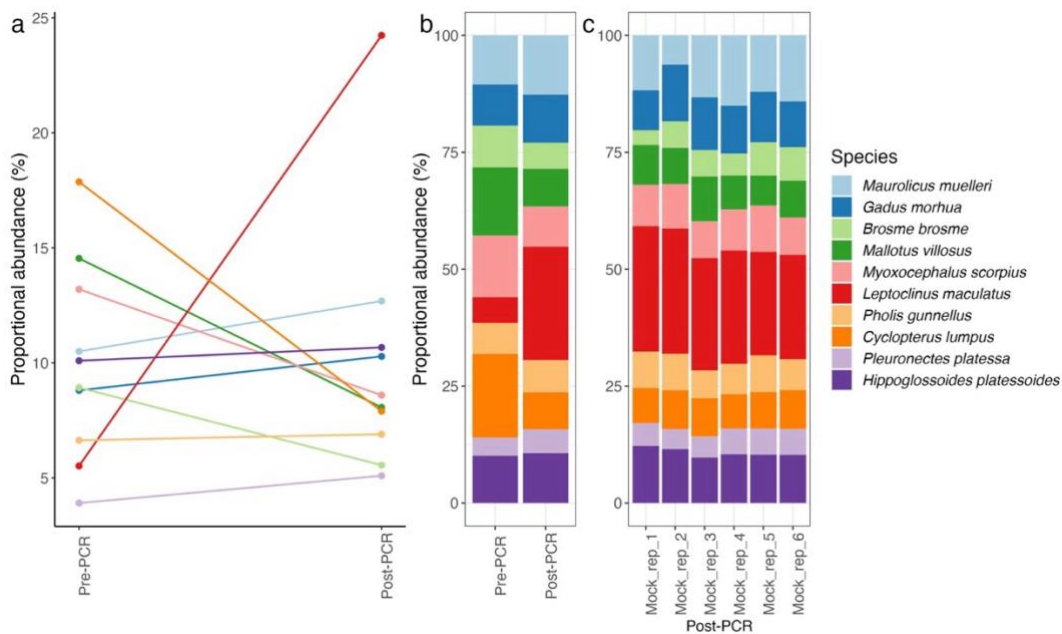
909



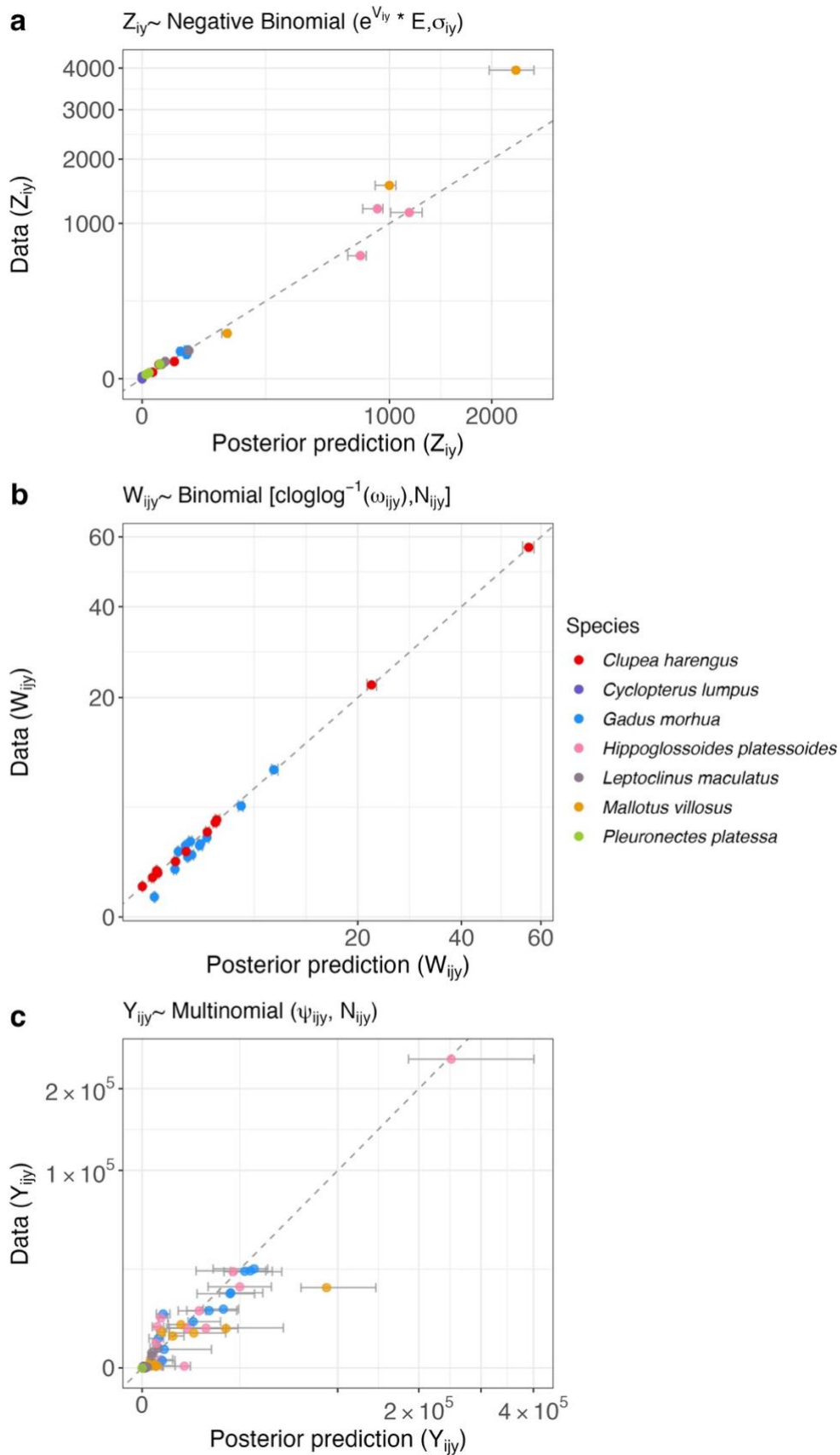
910  
911 *Figure S2. The cloglog link-function between the proportion of positive droplets, ddPCR measured*  
912 *concentrations (dotted line with diamond dots for samples) and known standard concentration (solid*  
913 *line with circles for samples; a) for *Clupea harengus* and *Gadus morhua*. The distribution of*  
914 *environmental samples concentration for each species (b).*  
915



916  
 917 *Figure S3. Rarefaction curves among all the environmental samples excluding the amplicons belonging*  
 918 *to Poecilia sphenops. The lines orange indicates the curves' saturation point of ca. 30,000 reads and*  
 919 *the blue line indicates the average number of reads per sample of 90,000 reads.*  
 920



921  
 922 *Figure S4. The change in the relative abundance of species composition in mock community before*  
 923 *(measured through Qubit quantification system) and after the PCR (measured as the average amplicon*  
 924 *abundance of six bio-replicate samples; a and b). Species composition shown in relative abundance*  
 925 *(measured as amplicon abundance) for each of the six bio-replicates of mock community samples (c).*  
 926



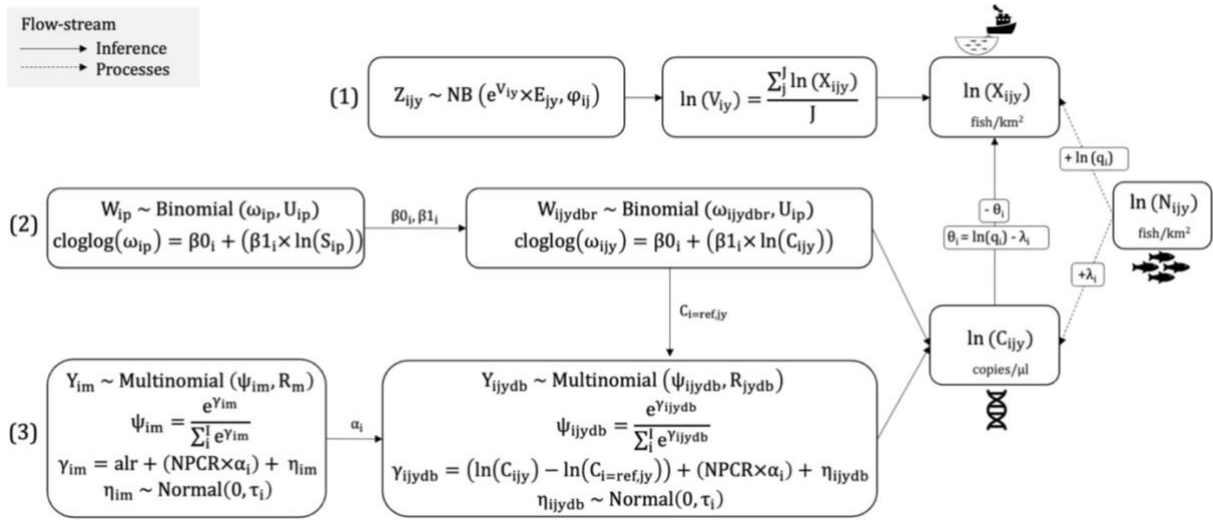
927

928 *Figure S5. Posterior predictive checks for ddPCR model compartment (a), metabarcoding model*

929 *compartment (b), and trawl model compartment (c).*

930

931  
 932  
 933  
 934



935  
 936  
 937  
 938

Figure S6. Full schematic overview of the joint Bayesian model workflow including all the inferences and processes.

939 **References**

- 940 Guri, G., Shelton, A. O., Kelly, R. P., Ray, J., Louise Rayica Louise, Yoccoz, N., Torild, J., Præbel, K.,  
941 *et al.* 2024. Quantifying the sensitivity between qPCR and ddPCR mechanisms for eDNA  
942 samples. *Methods in Ecology and Evolution*.
- 943 Shelton, A. O., Gold, Z. J., Jensen, A. J., D'Agnese, E., Andruszkiewicz Allan, E., Van Cise, A.,  
944 Gallego, R., *et al.* 2023. Toward quantitative metabarcoding. *Ecology*, 104.  
945 <https://onlinelibrary.wiley.com/doi/10.1002/ecy.3906> (Accessed 24 April 2023).  
946

