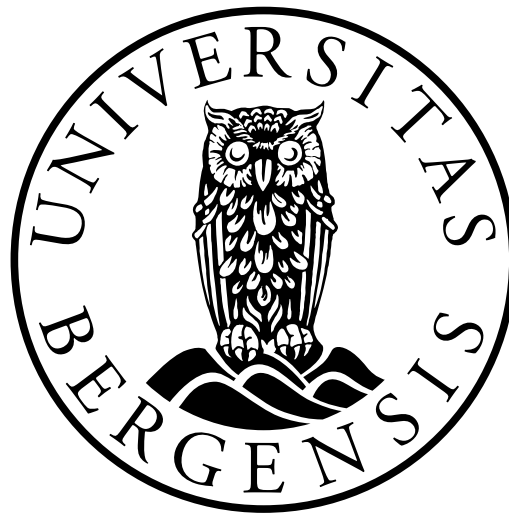




# Environmental DNA metabarcoding for diversity-based assessment of a landlocked fjord system: a comparison with traditional methods

Torill Synnøve Fjørtoft Johansen



Integrated Teacher Programme in Science and Mathematics  
Biodiversity, Evolution and Ecology  
Department of Biological Sciences  
University of Bergen

Supervisor: Jon T. Hestetun (NORCE)  
Co-supervisors: Henrik Glenner and Thorolf Magnesen  
(Department of Biological Sciences, University of Bergen)

November 28, 2022

# Acknowledgements

The last few months of writing this thesis have been challenging, overwhelming and fun. DNA-based methods are currently revolutionizing the field of biomonitoring, and it has been incredibly interesting to learn about the research in this field. First, thank you to my supervisor Jon Hestetun for thorough, helpful feedback and guidance while writing this thesis. I have always felt welcome to share thoughts and ask questions, and I have greatly appreciated the opportunity to learn from you.

I also want to thank co-supervisor, Henrik Glenner, for piquing my interest in marine biodiversity and allowing me to do this project. I appreciate your help in the field and lab and your feedback throughout the process of writing this thesis.

Next, I want to thank co-supervisor Thorolf Magnesen for helpful feedback on my thesis and Tomas Sørli for help during the fieldwork. Thank you to Tom Alvestad, David Rees, and Kenneth Meland for assistance and support in identifying our sampled species in the lab. Furthermore, I would like to thank Sigrid Mugu at NORCE for assisting me and answering all my questions during the work in the DNA lab. And to my field- and lab partner Jonette Eckholdt, I had a blast spending endless hours in the lab with you.

To my sister Marte and my very patient friend Ruben, thank you for lending your expertise when R, Excel, or Latex did not cooperate with me. You've saved me so much frustration. Further, I want to extend all my love to the friends I have made here in Bergen. I am forever grateful for all the good times and happy moments. And to my roomie, Johanne, thank you for putting up with me for the last few months. You're an absolute treasure.

A huge thanks to my family for supporting me throughout my studies. I am so lucky to have an engaged and loving support system of aunts, uncles, cousins, and siblings always cheering me on and, most importantly, making sure I am properly fed.

Lastly, I want to dedicate this thesis to my parents. The last five years have been challenging, and I could not have done this without your love and support.

Torill Synnøve Fjørtoft Johansen  
Bergen, 28 November





# Abstract

Marine ecosystems are under significant pressure from human activity, causing a loss of biodiversity and ecosystem functions. As a response to these threats, governmental authorities require monitoring of water bodies to identify threats and assess the ecological status. The current method for monitoring ecological status of water bodies is based on biological quality elements, such as the composition of benthic invertebrates, and biotic indices. However, morphological identification of benthic infauna is labor-intensive and costly. In recent years, increased interest in DNA-based methods has promoted environmental DNA metabarcoding coupled with high-throughput sequencing as an alternative approach for monitoring marine ecosystems. This method has the potential for a more rapid and cost-efficient assessment using environmental samples, but further validation in different habitats must be performed. Here, the applicability of eDNA metabarcoding for diversity-based assessment of a landlocked fjord was assessed by comparing the benthic invertebrate community described by metabarcoding and the morphological method. The Norwegian word “poll” is often used in studies of this habitat and will be used throughout this study. Invertebrates were identified based on morphology, and metabarcoding-based taxonomy was performed using universal barcodes cytochrome oxidase I gene (COI) and V1-V2 region of the small ribosomal subunit 18S. Taxonomic composition, alpha- and beta diversity and biotic indices were described and calculated using the data from both methods. The results show that eDNA metabarcoding can detect a significant fraction of biodiversity and provide complementary information about the taxonomic composition. The taxonomic resolution of 18S was higher than expected, illustrating the potential of 18S with higher local database coverage. Taxonomic groups are not equally represented by metabarcoding, but similar beta diversity patterns across all datasets indicate congruent community differentiation in the study area. However, diversity estimates and biotic index values from morphological and metabarcoding data were inconsistent, and discrepancies in the description of heavily impacted sites call into question the ability of metabarcoding to accurately describe the local diversity. The existing biotic indices could not be applied to molecular data retrieved from the poll, but the results show potential for the development of de novo indices. In conclusion, eDNA metabarcoding can be a valuable tool offering new possibilities in diversity assessments. Still, further effort to overcome the current limitations is necessary before implementing eDNA metabarcoding in standard monitoring programs.



# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and motivation . . . . .	1
1.2 Anthropogenic impacts on marine coastal ecosystems . . . . .	1
1.3 Environmental policies regarding water and biodiversity protection . . . . .	2
1.4 Current methods in marine biomonitoring . . . . .	4
1.5 The need for new methods in biomonitoring . . . . .	6
1.5.1 eDNA metabarcoding for biodiversity assessments . . . . .	7
1.5.2 Challenges in implementing eDNA metabarcoding in biomonitoring . . . . .	8
1.6 Landlocked fjords . . . . .	11
1.6.1 Kviturspollen . . . . .	12
1.7 Research Question and Objective of this Study . . . . .	12
<b>2 Material and methods</b>	<b>15</b>
2.1 Study area . . . . .	15
2.2 Sample collection and morphology-based taxonomic assignment . . . . .	16
2.3 Metabarcoding laboratory processing . . . . .	18
2.3.1 DNA extraction . . . . .	18
2.3.2 PCR amplification and high-throughput sequencing . . . . .	19
2.4 Bioinformatic processing of metabarcoding sequence data . . . . .	21
2.5 Statistical analysis . . . . .	22
2.5.1 Alpha diversity . . . . .	22
2.5.2 Beta diversity . . . . .	26
<b>3 Results</b>	<b>27</b>
3.1 CTD results . . . . .	27
3.1.1 Salinity . . . . .	27
3.1.2 Oxygen . . . . .	28

3.1.3	Temperature . . . . .	28
3.2	Morphological taxonomy . . . . .	29
3.3	High-throughput sequencing output . . . . .	31
3.3.1	Sequencing output for COI . . . . .	31
3.3.2	Sequencing output for 18S . . . . .	33
3.4	Comparison of taxonomic composition . . . . .	35
3.5	Alpha diversity patterns . . . . .	41
3.6	Beta-diversity patterns . . . . .	42
3.7	Biotic index calculation . . . . .	47
<b>4</b>	<b>Discussion</b>	<b>53</b>
4.1	Main findings from the morphology-based assessment . . . . .	53
4.2	Comparison of taxonomic composition . . . . .	55
4.2.1	The effect of primer bias on taxonomic composition . . . . .	56
4.2.2	COI and 18S can detect taxa undetected by morphology . . . . .	57
4.3	Performance of COI and 18S markers . . . . .	58
4.3.1	Taxonomic resolution of 18S and COI . . . . .	59
4.4	DNA signals from heavily impacted sites . . . . .	60
4.4.1	Challenges of using extracellular DNA in sediment samples . . . . .	61
4.5	Highly variable diversity estimates . . . . .	64
4.6	Beta diversity patterns are similar across datasets, but reveal incomplete sampling for certain stations . . . . .	66
4.7	Macrofauna indices do not perform well on molecular data retrieved from a poll	67
<b>5</b>	<b>Conclusions and Future Work</b>	<b>73</b>
5.1	Future outlook/research . . . . .	74
	<b>Appendix I. Bray-Curtis similarity matrices</b>	<b>101</b>
5.2	Bray-Curtis similarity matrix from morphological data . . . . .	101
5.3	Bray-Curtis similarity matrix from COI data . . . . .	101
5.4	Bray-Curtis similarity matrix from 18S (Metazoa) data . . . . .	102
5.5	Bray-Curtis similarity matrix from 18S (Eukaryota) data . . . . .	103
	<b>Appendix II. Taxa lists</b>	<b>105</b>
5.6	Morphological taxa list . . . . .	105
5.7	Unique metazoan taxonomic groups from COI OTU table . . . . .	109
5.8	Unique metazoan taxonomic groups from 18S OTU table . . . . .	114

# Chapter 1

## Introduction

### 1.1 Background and motivation

The ocean covers approximately 71 percent of the Earth's surface, and the life and processes found in the ocean are essential to our planet (National Geographic Society 2022<sup>a,b</sup>). These marine systems host an enormous variety of life in many habitats, described with the term *biodiversity* (Convention of Biological Diveristy 2022). Biodiversity is closely related to ecosystem functioning, which De Groot defines as “the capacity of natural processes and components to provide goods and services that satisfy human needs, directly and indirectly” (1992, referred in de Groot et al. (2002)). These functions result from biological processes by living organisms interacting with each other and the abiotic components of the environment (de Groot et al. 2002). Further, these interactions regulate the climate and nutrient cycling, and marine systems provide benefits such as food resources, clean air, and recreational experiences (Beaumont et al. 2007, Convention of Biological Diveristy 2022, de Groot et al. 2002). The flow of energy and matter is affected by the diversity of traits in a local community, and with greater diversity follows greater resource use (O'Connor & Byrnes 2014, p.111). Thus, the loss of species from a community ultimately affects the outcome of ecosystem processes (O'Connor & Byrnes 2014, p.109). These processes underline the significance of marine ecosystems and maintaining their biodiversity. However, despite being so important to us, marine ecosystems are under significant pressure from human activity.

### 1.2 Anthropogenic impacts on marine coastal ecosystems

Due to proximity to human settlement, coastal marine ecosystems are under a significant amount of pressure from human activity (Breitburg et al. 2018, Harley et al. 2006, He & Silliman 2019). The continued urbanization of the world's population will naturally affect coastal ecosystems, and studies have identified several sources of impact (Islam & Tanaka 2004, Todd

et al. 2019). These pressures come from nutrient and heavy metals inputs, wastewater, land-based activity (agriculture) pollutants, and natural input to marine systems (He & Silliman 2019, Tuholske et al. 2021). Pollution affects most coastal areas of the world, and the greatest volume of waste to marine systems is sewage (Islam & Tanaka 2004). Highly populated cities generate large amounts of waste daily, which wash out in nearby aquatic systems (Islam & Tanaka 2004). These anthropogenic pressures cause changes in water quality due to nutrient loading causing eutrophication, which has become an enormous problem worldwide (Breitburg et al. 2018, Levinton 2009, p.576). Eutrophication increases the amount of organic matter in bottom waters where microbial decomposition occurs (Breitburg et al. 2018). A consequence of this is oxygen depletion (hypoxia and anoxia) caused by increased microbial activity (Breitburg et al. 2018, Diaz & Rosenberg 1995, Lenihan & Micheli 2001). In the absence of oxygen, microorganisms utilize other electron acceptors and generate the highly toxic compound  $H_2S$  during metabolism (Diaz & Rosenberg 1995). Oxygen depletion significantly affects geochemical and ecological processes (Breitburg et al. 2018), and eutrophication accelerates these conditions (Diaz & Rosenberg 1995). Oxygen is essential for marine organisms, and the declining oxygen levels significantly limit the abundance and distribution of animals (Breitburg et al. 2018).

Another human-induced impact on marine ecosystems is climate change, which the ocean is highly affected by (Breitburg et al. 2018, Harley et al. 2006, He & Silliman 2019). Climate change directly influences the performance of individuals via changes in physiology, morphology, and behavior and causes changes at the community level through altering species interactions (Harley et al. 2006). Climate stressors for coastal soft-bottom ecosystems include warming, sea level rise, and ocean acidification (He & Silliman 2019). Ocean warming contributes to a further decrease in oxygen content in the ocean (Breitburg et al. 2018), and ocean acidification will have severe consequences for marine invertebrates that build carbonate structures (Harley et al. 2006). The overall impact on the ecosystem is alterations in species distribution, biodiversity, productivity, and microevolutionary processes (Harley et al. 2006). If these trends continue, we could eventually reach a tipping point in which reversing the trends and restoring ecosystems will become very challenging (Convention of Biological Diversity 2022). These findings emphasize the need to monitor the changes in the marine environment.

### **1.3 Environmental policies regarding water and biodiversity protection**

In light of these concerns, authorities have introduced laws, regulations, and international agreements to better monitor and care for ecosystems (Borja et al. 2010). The European Water Framework Directive (WFD) came into force in 2000, and this directive aims to achieve good status for all waters, including surface and groundwater (European Commission n.d.). The

Water Framework Directive is included in the European Economic Area (EEA) agreement, and it requires that all of its members assess the ecological quality of water bodies (European Commission n.d., Rygg 2006). Norway has issued its own implementation of WFD (Vannforskriften 2006), which entered into force in Norway in 2007 (Direktoratsgruppen vanndirektivet 2018*b*, NIVA 2017, Petersen et al. 2009). The purpose of the regulation is to provide a framework for environmental goals that ensure the protection and sustainable use of water bodies (Vannforskriften 2006, §1). The regulation gives orders to monitor water bodies in terms of ecological quality based on several quality elements (Direktoratsgruppen vanndirektivet 2018*b*, Vannforskriften 2006).

In addition to the European Water Framework Directive (WFD) addressing the ecological status of water bodies, the Convention of Biological Diversity (CBD) targets the conservation of biological diversity (Convention of Biological Diversity 2012). CBD is a global agreement that Norway is also signatory to that entered into force in 1993. Based on this agreement, the Norwegian government passed the Nature Diversity Act ("Naturmangfoldlova"), which aims to protect biodiversity (Regjeringen 2009, 2021). As of 2018, the Norwegian implementation of the Water Framework Directive was given authority in the Nature Diversity Act (Trøen 2018). This authority will ensure that the regulations also apply to water bodies affected by impacts other than pollution, such as invasive species (Prop. 93 L (2017-2018) 2017). This approach aims to ensure policymakers access to the information necessary to make informed decisions about environment monitoring (Meld.St.14 2015).

In the Norwegian water management plans, characterization of environmental impact for different water bodies is essential in building a knowledge base (Vannportalen 2020a). This process involves a characterization of the water body category, e.g., river, coastal water, and the anthropogenic impacts in these areas (Direktoratsgruppen vanndirektivet 2018*b*). Every water body category has different quality elements, and one of the biological quality elements for all categories is the composition and amount of soft-bottom benthic invertebrates (>1mm) (Rygg 2006, Vannportalen 2020b). The classification system uses five state values with defined limit values. These categories are "excellent", "good", "moderate", "poor", and "very poor" as illustrated in Figure 1.1. The values indicating the different categories are comparable with similar water bodies of neighboring countries through a legally binding inter-calibration (Direktoratsgruppen vanndirektivet 2018*b*). The condition category "excellent" is a reference condition defined as the condition of a biological quality element with no or only minor impact from human activity (Borja et al. 2004, Direktoratsgruppen vanndirektivet 2018*b*).

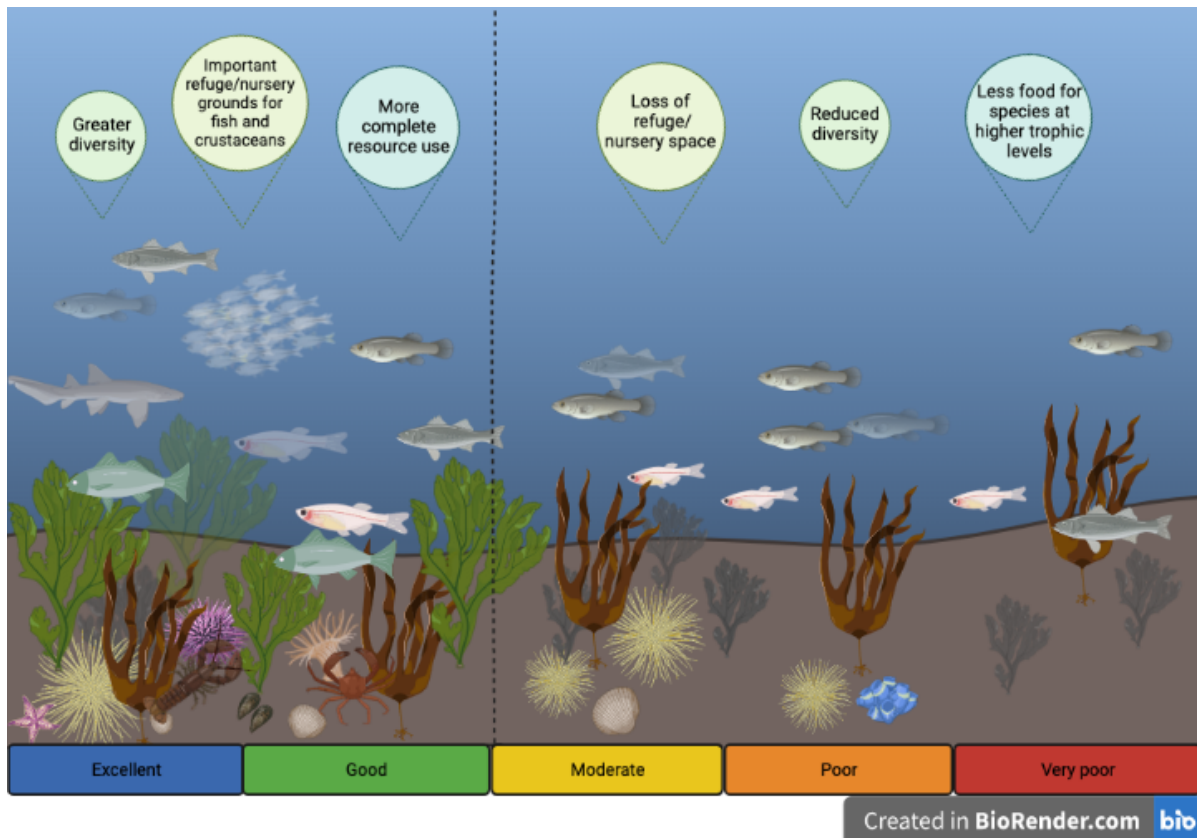


Figure 1.1: Illustration of the five state values of ecological condition for coastal waters (Adapted from Miljødirektoratet (n.d.), created in Biorender). The state value “excellent” is defined as the condition of a biological quality element with no or only minor impact from human activity and is illustrated with greater diversity. The environmental conditions to the right of the dotted line indicate the impact of human activity, leading to the loss of species and reduced diversity (Miljødirektoratet n.d.).

The European countries bound to the WFD follow the same principles and schedule (Vannportalen n.d.a). The current plan period is 2022-2027, and the goal is to achieve good ecological status for all water bodies by 2027. To achieve this goal, we need insight into the ecological status of water bodies and assess whether the measures must be adjusted (Vannportalen n.d.b). The plans in Western Norway underline further development of a knowledge basis and specifically refer to fjords systems and water bodies with poor bottom-water replacement, such as landlocked fjords and fjords with thresholds (Vannportalen n.d.c).

## 1.4 Current methods in marine biomonitoring

As described above, the Water Framework Directive, the law of biodiversity, and other international agreements require the implementation of marine biomonitoring programs. As mentioned above, the biological quality elements of coastal waters is the composition and amount of soft-bottom benthic invertebrates (>1mm) (Direktoratsgruppen vanndirektivet 2018b, Rygg 2006). These are often long-lived sedentary species that cannot avoid local water and sediment conditions, and the benthic community consists of species with different tolerances (Dauer



1993). Due to their rapid response to environmental changes, benthic invertebrates can reflect the local conditions and trends (Pearson & Rosenberg 1978, Salas et al. 2006). This characteristic means we expect to see more tolerant and opportunistic species in more polluted areas, providing insight into the local environment (Pearson & Rosenberg 1978). In practice, the morphological approach in biomonitoring is based on a quantitative method where the number of benthic species and individuals are related to a known sample area (ISO 2014). A local-site sampling of the seafloor is followed by sample processing in the laboratory, and the biological indicators are identified using morphological taxonomy (Aylagas et al. 2018, ISO 2014). Sample processing includes sorting the benthic fauna into higher taxonomic groups (e.g., Polychaeta, Mollusca), and identification to the lowest possible taxonomic level. The identified species are counted and represented in a taxon list for each locality (ISO 2014), and this data is used to describe ecological quality status using biotic indices (Direktoratsgruppen vanndirektivet 2018*b*, Pawlowski et al. 2018).

Diversity indices have been an essential tool for classifying the state of the environment for several decades (Molvær et al. 1997). Such indices include the Shannon index  $H'$  (alternatively called Shannon-Wiener index) and the Hurlbert index ( $ES_n$ ), which are still crucial in the current classification system (Direktoratsgruppen vanndirektivet 2018*b*, Molvær et al. 1997). In addition to diversity indices that only account for the number and relative abundance of organisms, several sensitivity indices have been developed for assessing the ecological quality of water bodies (Rygg 2002, 2006, Rygg & Norling 2013). Sensitivity indices are based on specific taxonomic groups referred to as biological quality elements and their response to environmental change (Cordier et al. 2019). For coastal waters, the biological quality element is the composition of benthic invertebrates (Direktoratsgruppen vanndirektivet 2018*b*), and the indices are based on the different sensitivity of benthic invertebrates to disturbance and changes in environmental conditions (Cordier et al. 2019, Rygg 2002, Rygg & Norling 2013). The current monitoring programs rely on biotic indices to assess ecological status (Cordier et al. 2021), and the indices used in the Norwegian monitoring program are the Indicator Species Index (ISI) (Rygg 2002), Norwegian Sensitivity Index (NSI) (Rygg & Norling 2013) and Norwegian Quality Index (NQI1) (Rygg 2006). The Azti Marine Biotic Index (AMBI) is extensively used in Europe, but AMBI has largely been replaced with NSI for Norwegian waters (Rygg 2006, Rygg & Norling 2013).

To make good decisions and implement the necessary measures for ecosystems, it is of great importance that the information about the ecosystem's state is readily available to policymakers (Lanzén et al. 2021). Concerning this, an obvious limitation of the traditional biomonitoring method is how time-consuming it is (Bourlat et al. 2013, Danovaro et al. 2016). The laboratory processing of the sediment may be particularly labor-intensive when dealing with many samples. These time-consuming processes put further limitations on the scale of biomonitoring project (Bourlat et al. 2013). In addition, the morphological method requires specialized

knowledge, and taxonomic precision is unverifiable and often varies across studies (Baird & Hajibabaei 2012, Bourlat et al. 2013). Identification of cryptic species complexes is essential in conservation, but they are challenging for taxonomists as they are morphologically indistinguishable and represent a taxonomic bias in biodiversity estimates (Bickford et al. 2007). A similar challenge is also posed by juvenile stages of certain species (Danovaro et al. 2016) as well as phenotypic plasticity and genomic variability in characters used for species recognition (Hebert et al. 2003). Further, species sensitivity to stressors is the foundation of biotic indices, but we still need more knowledge on these stressor-impact relationships (Pawlowski et al. 2018). In conclusion, the current method may restrict possibilities in large-scale monitoring as a result of a trade-off between resources (e.g., time) and scale of monitoring (e.g., sample size), and it is arguably insufficient as a basis for informed ecosystem management (Baird & Hajibabaei 2012, Bourlat et al. 2013).

## 1.5 The need for new methods in biomonitoring

As the need for monitoring ecosystems increases, the need for more objective, cost-effective, and faster methods in biomonitoring is pressing (Cordier et al. 2021). Due to a time lag between the survey and final results, the current approach fails to provide continuous up-to-date information about water bodies (Lanzén et al. 2021). Therefore, it is appropriate to explore alternative methods, and recent progress in molecular biology and DNA sequencing techniques presents opportunities for this objective. Hebert et al. (2003) presented the idea of using DNA for biological identifications and referred to the technique as DNA barcoding. In this method, an identification system based upon short DNA sequences known as DNA barcodes can distinguish between different species due to genetic variation (Hajibabaei et al. 2007, Hebert et al. 2003). These are stretches of DNA that are close to identical between members of a taxon but may vary between different taxa and can therefore be used to identify species (Blaxter 2004). The selected barcode is retrieved from the specimen in question, compared to a reference barcode database, and assigned to a species if there is a match (Hajibabaei et al. 2007).

As the number of large-scale DNA-sequencing projects increased, the need for more efficient sequencing methods became evident, and further efforts gave rise to high-throughput sequencing (HTS) in 2005 (Margulies et al. 2005, Taberlet et al. 2012a). High-throughput sequencing allows for the simultaneous analysis of numerous sequences (Deiner et al. 2017) and enables DNA-based identifications on a much grander scale (Shokralla et al. 2012, Taberlet et al. 2012a). Studies on microbial diversity introduced the concept of environmental DNA, which later expanded to detect multicellular animals after it was first demonstrated by Ficetola et al. (2008) (Pawlowski et al. 2020). Taberlet et al. (2012b) defines the term eDNA as a “complex mixture of genomic DNA from many different organisms found in an environmental sample”. The sample can, for example, be soil, water, sediments, gut content, or feces (Shokralla et al.

2012, Taberlet et al. 2012a). In response to the emergence of high-throughput sequencing, the term “DNA metabarcoding” was introduced by Taberlet et al. (2012a). Metabarcoding refers to the simultaneous identification of multiple species from high-throughput sequencing of DNA from an environmental sample (Taberlet et al. 2012a). The technique also includes bulk samples of entire organisms separated from the environmental sample before analyses, thus providing higher-quality DNA in the tissue samples (Aylagas et al. 2018, Macher et al. 2018, Taberlet et al. 2012a).

### 1.5.1 eDNA metabarcoding for biodiversity assessments

Since its emergence, eDNA metabarcoding has been applied to many biomonitoring projects (Aylagas et al. 2018, Taberlet et al. 2018, 2012b, p.1). In this methodology, a DNA metabarcode sequence is amplified from the environmental sample (Taberlet et al. 2018, p. 2), and it has shown great potential for high-throughput species identification (Taberlet et al. 2012a). Standardized barcoding is limited to the identification of single species with high-quality DNA, thus making metabarcoding more suitable for large-scale monitoring projects (Taberlet et al. 2012a).

The different steps of a standard eDNA metabarcoding study are illustrated in Figure 1.2 with steps presented by Pawlowski et al. (2018) and Zinger et al. (2019). First, eDNA samples (water, soil, sediment, or bulk samples) are obtained in the field and stored in a freezer until further handling in the laboratory. The DNA is subsequently extracted before PCR amplification of the chosen marker gene using a universal primer, e.g., COI or 18S (Leray et al. 2013, Pawlowski et al. 2018, Zinger et al. 2019). The resulting amplicons are then sequenced using high-throughput sequencing. The data is processed using bioinformatic pipelines, ultimately producing amplicon sequence variants (ASVs) (Callahan et al. 2017) or operational taxonomic units (OTUs), which are defined by clustering closely related sequences (Blaxter et al. 2005). The obtained OTUs are assigned to a taxon through comparison with a sequence database, resulting in an OTU table representing the detected taxa, which is conceptually analogous to a morphological species list. Such OTU tables can be used to calculate biotic indices and assess the ecological status of the water body (Pawlowski et al. 2018, Zinger et al. 2019).

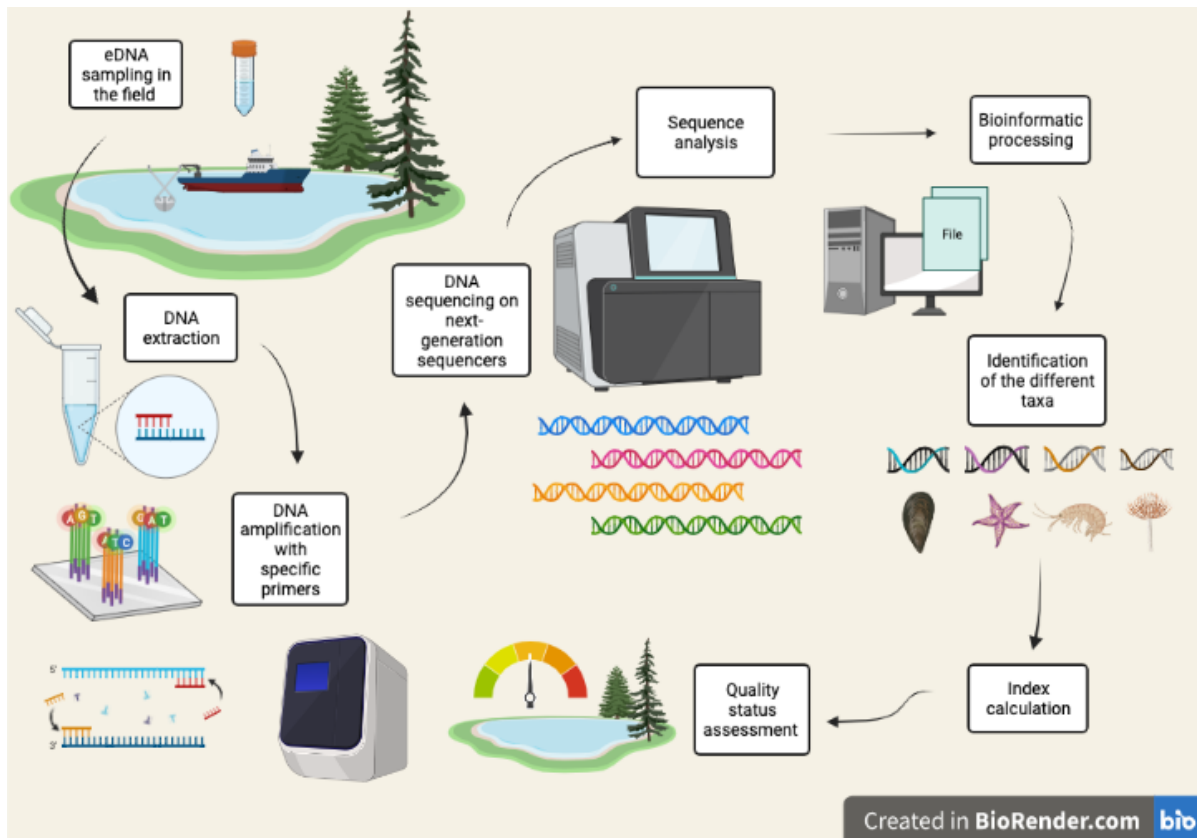


Figure 1.2: A general workflow in eDNA metabarcoding for biomonitoring (adapted from Pawlowski et al. (2018) and Zinger et al. (2019); made in BioRender). The first step is sampling eDNA (e.g., sediment or water sample), followed by DNA extraction in the laboratory. The extracted DNA is amplified in PCR using an appropriate primers, and amplicons are sequenced using high-throughput sequencing techniques. The obtained sequences are bioinformatically processed, resulting in an OTU table with assigned taxonomy. The resulting taxon list can be used to calculate biotic indices and infer ecological status (Pawlowski et al. 2018, Zinger et al. 2019)

## 1.5.2 Challenges in implementing eDNA metabarcoding in biomonitoring

To successfully implement eDNA metabarcoding in biomonitoring, we need a degree of consensus about the protocol and routines, including laboratory steps and bioinformatic analysis (Andújar et al. 2018, Jeunen et al. 2019). The relevant information regarding the steps of eDNA metabarcoding is currently scattered in the scientific literature in a context where standards are lacking (Taberlet et al. 2018, p. 150). This lack of standardized protocols may cause discrepancies and differences in results, making it difficult to compare studies (Clarke et al. 2017, Cristescu & Hebert 2018, Jeunen et al. 2019). To successfully integrate eDNA metabarcoding into monitoring programs, we need to resolve a number of technical and conceptual issues (Aylagas et al. 2018, Cordier et al. 2021, Cristescu & Hebert 2018, Hering et al. 2018).

As illustrated in Figure 1.2, eDNA metabarcoding studies work across different disciplines

requiring multiple skills, and may appear as a minefield when delving into it (Murray et al. 2015, Taberlet et al. 2018, p.150). Each of the steps described in Figure 1.2 contains sources of possible bias we must consider when designing the experimental protocol and drawing conclusions based on the results (Zinger et al. 2019). These steps include work performed in the field, laboratory, and bioinformatics (Zinger et al. 2019). One of the main issues with DNA metabarcoding is its dependency on PCR and sources of bias related to the method (Taberlet et al. 2012a). First, amplification with PCR can introduce errors to the sequences (Taberlet et al. 2012a). PCR errors and contamination during PCR may produce false positives, which will affect the results (Taberlet et al. 2018). To address this, Taberlet et al. (2018) emphasizes having good laboratory practice to limit contamination risk, and Yu et al. (2012) specifies avoidance of PCR as desired future improvement to metabarcoding.

The second challenge with the technique is finding a suitable metabarcode to achieve the desired accuracy of metabarcoding data (Deagle et al. 2014, Taberlet et al. 2012a). When choosing a DNA marker, several elements should be considered. The taxonomic group of interest and the level of taxonomic resolution (i.e., taxonomic level of identification) necessary to answer the study's objective must be clearly defined beforehand (Taberlet et al. 2018, p.7). For example, COI targets metazoans (Deagle et al. 2014), 16S rRNA targets prokaryotes (Aylagas et al. 2017) and 18S rRNA targets eukaryotes (Tang et al. 2012). Further, the expected DNA quality must be considered since shorter DNA metabarcodes are beneficial when dealing with degraded DNA (Taberlet et al. 2018, p.7). In general, (Taberlet et al. 2018, p.8) describes the perfect metabarcode as a short and highly-variable DNA sequence with conserved regions at both ends to anchor the primers. The barcode should be in a genetic region represented by all the species in a DNA barcode reference database (Taberlet et al. 2018, p.8).

For years, the standard barcode used for biological identifications of individual metazoans was the 658 bp cytochrome c oxidase I gene (COI), a region of the gene often referred to as the "Folmer fragment" (Folmer et al. 1994, Hebert et al. 2003, Taberlet et al. 2012a). Several required attributes make this gene region applicable for biological identifications. Studies show that more than 95 % of species possess unique COI barcode sequences (Hajibabaei et al. 2007), and the genetic variation makes species-level identification possible (Deagle et al. 2014). The COI gene can be amplified using PCR, and DNA barcode reference libraries, such as the Barcode of Life Data System or Midori, cover COI sequences of many different taxa (Deagle et al. 2014, Leray et al. 2022, Ratnasingham & Hebert 2007). In recent years, new primers targeting a shorter 313 bp region within the Folmer region have been designed, the so-called Leray primers (Leray et al. 2013). The shorter primer set meets the read-length limitations of HTS (Andújar et al. 2018) and has shown great promise for targeting metazoan diversity (Leray et al. 2013). In the shorter 313 bp primers, increased degeneracy (i.e., ambiguous positions) of the most variable positions can improve universality and accommodate variation in primer sites (Leray et al. 2013, Wangenstein et al. 2018).

However, in an opinion piece by Deagle et al. (2014), questions have been raised about whether COI is an appropriate marker for metabarcoding. For instance, primer binding sites in the COI gene are not highly conserved, causing a mismatch between primer and template DNA and unreliable PCR amplification across taxa (Deagle et al. 2014, Elbrecht & Leese 2015, Zhan et al. 2014). Consequently, this taxonomic bias results in unequal recovery of the taxa present in a sample and can result in false negatives (i.e., failing to detect present species) (Deagle et al. 2014, Elbrecht & Leese 2015, Hajibabaei et al. 2011). Furthermore, challenges of COI amplifying non-target bacteria are also reported in several studies using eDNA (Andújar et al. 2018, Yang et al. 2013).

Despite challenges, COI is still a widely used barcode for metazoans due to extensive representation in reference databases and high nucleotide variability allowing species-level identification (Andújar et al. 2018). To address taxonomic bias related to primers and increase the taxonomic coverage, researchers often use multiple markers (Alberdi et al. 2018, Clarke et al. 2017, Duarte et al. 2021, Kelly et al. 2017, Leray et al. 2013). For this objective, the nuclear subunit 18S ribosomal RNA gene is widely used in biodiversity studies targeting eukaryotes (Tang et al. 2012). Different regions of the 18S genetic marker are used for metabarcoding, such as V1-V2 (Hestetun, Lanzén & Dahlgren 2021), V4 (Lejzerowicz et al. 2015) and V8 (Mauffrey et al. 2020). This marker has more conserved primer binding sites (Deagle et al. 2014) and has shown to have broader taxonomic coverage (Zhan et al. 2014). However, the taxonomic resolution (i.e., species level identification) for 18S is lower than COI (Tang et al. 2012, Zhan et al. 2014). The two markers have been applied to several biodiversity and impact assessments with reports of important and valuable findings (Kelly et al. 2017, Lanzén et al. 2021, Leduc et al. 2019, Mauffrey et al. 2020).

Another prerequisite for species identifications using DNA metabarcoding is sufficient coverage in barcode reference libraries (i.e., BOLD and Midori for COI and Silva for 18S) (Bourlat et al. 2013, Leray et al. 2022, Quast et al. 2013, Ratnasingham & Hebert 2007), and dependence upon the databases is a significant limitation of DNA metabarcoding (Taberlet et al. 2012a). The current gaps in the reference databases affect the taxonomic assignment of eDNA sequences (Pawłowski et al. 2018). Consequently, the incompleteness of databases reduces the accuracy of biodiversity estimates and limits the amount of OTUs that can be used for biotic index calculation (Pawłowski et al. 2018). As previously mentioned, the COI marker is linked to a growing reference library, but more effort is necessary to expand it (Bourlat et al. 2013). The North Sea, the heaviest sampled marine area worldwide, shows relatively high database coverage for COI and 18S (Hestetun et al. 2020). However, the substantial gaps uncovered underline the need to expand the worldwide coverage to improve the accuracy of eDNA metabarcoding studies (Hestetun et al. 2020). The identification of species requires taxonomic expertise and staff to perform sequencing of the specimen for a chosen marker gene (Bourlat et al. 2013). To address this issue, international research alliances such as the Consortium for the Barcode

of Life (2004-2011) and International Barcode of Life (2008-) have launched and overseen international projects aiming at increasing our understanding of biodiversity through barcoding (International Barcode of Life n.d.).

Despite these challenges, studies show that eDNA metabarcoding can be a valuable tool for ecological research and impact assessments. The method has proved successful in assessing impacts of oil drilling, extraction, or spills (Lanzén et al. 2021), the pressures posed by offshore platforms (Mauffrey et al. 2020), detection of invasive invertebrate species (Klymus et al. 2017) and analysis of diet and trophic relationships (Günther et al. 2021). Compared to the current monitoring method, metabarcoding is a cheaper and faster approach for monitoring ecosystems (Mauffrey et al. 2020). The sampling protocol is less invasive to the ecosystem, especially when using water samples (Leese et al. 2016), which is a strong argument to stakeholders (Blancher et al. 2022). Furthermore, metabarcoding offers a solution to the issue presented by cryptic species and can thus provide more accurate estimations of biodiversity (Pereira et al. 2021, Tang et al. 2012). The method has been used for diversity-based monitoring in several habitats (Aylagas et al. 2018, Fernández et al. 2019, Lejzerowicz et al. 2015), with promising results. First, studies show how eDNA metabarcoding can reveal aspects of biodiversity not covered by the morphological method (Keck et al. 2022). Second, eDNA can provide valuable information about community composition and diversity patterns (Cahill et al. 2018, Ji et al. 2022, Lejzerowicz et al. 2015). Metabarcoding is a promising alternative to traditional methods as it can assign taxonomy to many samples more rapidly, making it possible for policymakers to make informed decisions within reasonable time (Aylagas et al. 2018, Lanzén et al. 2021). Before implementing eDNA metabarcoding to standard monitoring programs, further testing of the method in various habitats is necessary (Blancher et al. 2022). While metabarcoding has been applied to estuaries and other habitats (Aylagas et al. 2018), the approach has not yet been tested in landlocked fjords, which have been part of standard monitoring programs in Norway for years (Vannportalen n.d.c).

## 1.6 Landlocked fjords

Landlocked fjords are marine/brackish water pools with limited water exchange with surrounding coastal waters due to their narrow openings and a shallow threshold (Direktoratet for naturforvaltning 2007). Landlocked fjords differ from typical fjords, so the Norwegian word “poll” is often used in studies about this habitat (Wassmann 1985), and will also be used throughout this study. The physical and biological characteristics of polls are affected by the depth of sill threshold, freshwater input, and anthropogenic load (Direktoratet for naturforvaltning 2007), which makes them different from the surrounding fjords (Wassmann 1985). The hydrography is usually relatively stable, and stagnant bottom water frequently occurs (Wassmann 1985). These differences often make polls very different from the surrounding areas in

terms of physical characteristics and biological diversity in the benthic fauna (Direktoratet for naturforvaltning 2007). Due to the limited water exchange and proximity to land, polls are assumed to be particularly sensitive to eutrophication (Wassmann 1985). Thus, monitoring the biodiversity in this habitat is essential to protect the unique fauna that resides in these habitats (Direktoratet for naturforvaltning 2007).

### 1.6.1 Kviturspollen

Kviturspollen is a poll located in Fanafjorden outside of Bergen. The poll has been monitored extensively in the last few years, and changes in both physical characteristics and benthic fauna have been registered throughout this time period (Kvalø et al. 2015). Kviturspollen is experiencing limited groundwater renewal and has previously been a recipient of wastewater and other runoff, as well as being affected by other human activities (Kvalø et al. 2015). In the report series, “Recipient monitoring of the fjord systems around Bergen”, Kvalø et al. (2015, 2014) assessed the poll in terms of physical characteristics and benthic fauna. The two stations included in the surveys were Kv1 at 14 m depth and Kv5 at 10 m depth, as shown in Figure 2.1. The physical characteristics included nutrition levels, chlorophyll-a values, bacteria and oxygen levels, and sediment composition (Kvalø et al. 2015). Due to insufficient water exchange, the oxygen levels at Kv1 were classified as “poor”, and the investigation of sediment composition at Kv5 (10 m) from 2014 indicated some sedimentation of organic matter (Kvalø et al. 2015). Based on the calculation of benthic diversity ( $H'$ ) and NSI, the station was classified as good in 2015 and had shown improvements from previous years (Kvalø et al. 2015). In addition, Lundberg (2015) assessed the poll based on the occurrence of eelgrass *Zostera marina*. The eelgrass east of Langeneset was classified as good, while the eelgrass in the northern parts of the bay was classified as in poor condition (Lundberg 2015).

## 1.7 Research Question and Objective of this Study

This study aims to assess the performance of eDNA metabarcoding in comparison to the current traditional method for diversity-based assessment of a poll. Research and legislation clearly state the need to monitor water bodies and biodiversity as a part of knowledge-based management. Due to the limitations of the current morphological method, the interest in DNA-based approaches for monitoring is increasing. Does metabarcoding meet quality standards as a tool for biomonitoring? Here, I will assess the suitability of eDNA metabarcoding for diversity-based monitoring of a poll system by answering the three following research questions:

- i. Is the taxonomic composition (genus/family/species) obtained by eDNA metabarcoding consistent with taxonomic composition from the morphological method?



ii. How well does eDNA metabarcoding describe changes in community composition (beta diversity)?

iii. Can molecular data retrieved from a poll be used to infer ecological status by applying current diversity and sensitivity indices?



# Chapter 2

## Material and methods

### 2.1 Study area

The study area Kviturspollen is located northwest of Fanafjorden in Bergen (Kvalø et al. 2015). It is an s-shaped poll consisting of an inner and outer part, with a total length of 1800 m and a width of 250 m at the biggest (Dybern 1967). The inner part of the fjord has depths of around 4-6 m, while the outer can reach a depth of over 10 m (Dybern 1967). The poll is divided into a series of small basins with mixed bottoms consisting of mud, sand, and a high organic matter content (Dybern 1967). In the deepest, stagnant basins, organic matter settles and breaks down, leading to low oxygen conditions and H<sub>2</sub>S production (Dybern 1967). Allochthonous material from surrounding land affects the topography and hydrography of Kviturspollen, and this makes its way to the polls basins, which act as sedimentation pans for organic material (Dybern 1967). The main challenges to water bodies of the region Hordaland were presented in a consultation document addressing the current plan period of 2022-2027 (Vannportalen 2019). Due to increased urbanization around Bergen, wastewater constitutes the most significant impact on Kviturspollen. The poll is close to Flesland airport, which may also impact the area due to the release of chemicals (Vannportalen 2019).

In this study, sediment samples were collected from six stations inside the poll and one reference station outside the poll (St.9). Two stations, 7 and 8, were sampled to map the occurrence of the holothurian *Leptosynapta* sp. as a part of a separate study, and these are not included here. An overview of the study area and positions of the sampled stations are shown in Figure 2.1. During the sampling of each station, the coordinates, depths, and a general description, including color, smell, and sediment type, were noted. An overview of the stations with coordinates and depth info is shown in Table 2.1.

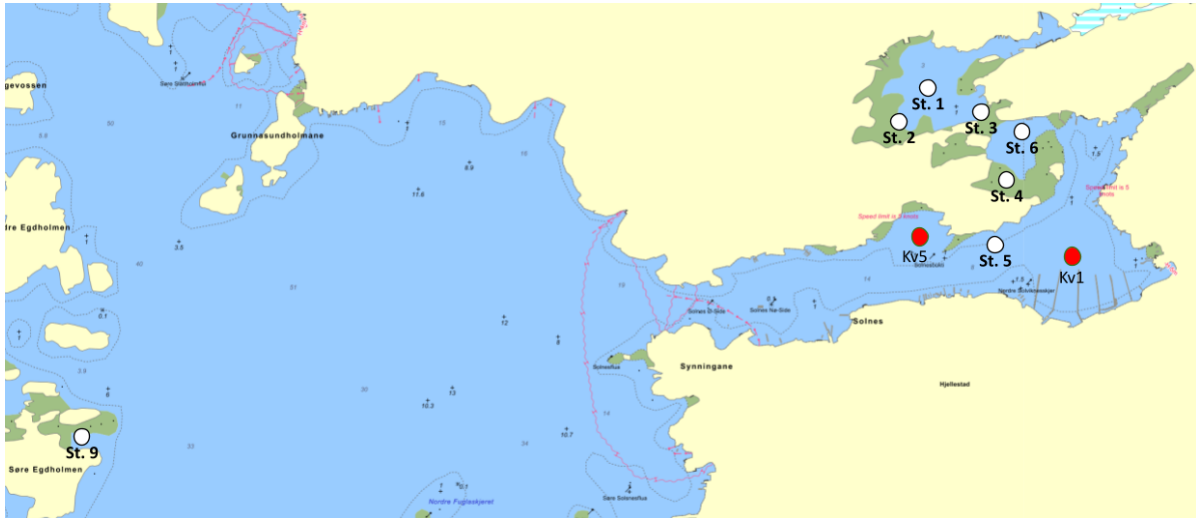


Figure 2.1: A geographical overview of the study area and sampling stations. The study area Kvitur-spollen is an s-shaped landlocked fjord with stations 1-6 distributed in the inner and outer part of the poll, as indicated by white circles. Station 9 is located in Raunefjorden and indicated by a white circle in the far left of the map. The stations Kv1 and Kv5 from previous investigations of Kvitur-spollen are indicated by red circles.

## 2.2 Sample collection and morphology-based taxonomic assignment

### Fieldwork

Benthic samples were collected from seven stations in Kvitur-spollen between December 2021 - February 2022 from a small research vessel provided by the University of Bergen. Sample collection was done in accordance with standard methodology (ISO 2014) using a van-Veen grab (0.1 m<sup>2</sup>) with two sample replicates per station. One subsample for eDNA was collected from the surface layer of each grab sample into a 50 ml falcon tube as shown in Figure 2.2. For station 9, the grab samples used for the morphological identification and grab samples for eDNA were not the same, and the grabs were separated as 9A/9B for morphology and 9C/9D for eDNA samples. For remaining stations, the samples for morphological identification and eDNA were from the same grabs.

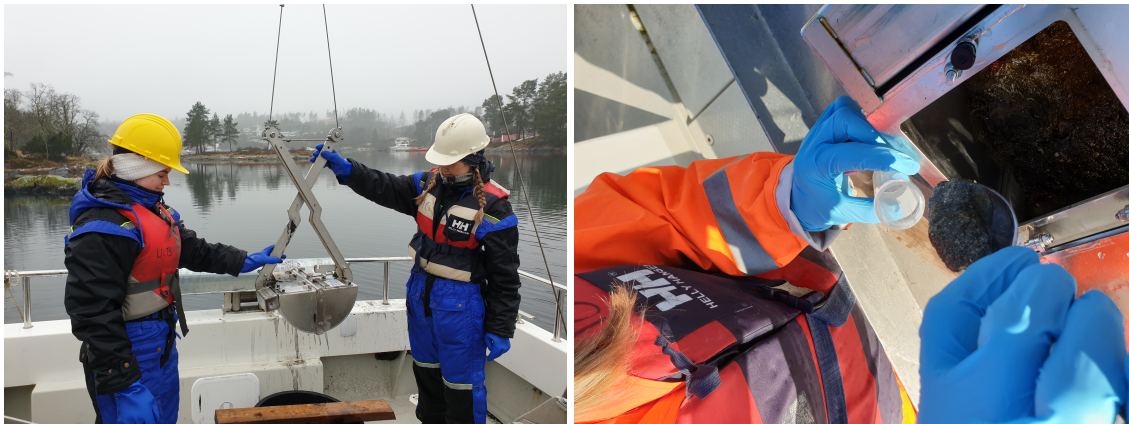
Furthermore, an overview of the samples collected from each station is given in Table 2.1. The eDNA samples were stored at -20 degrees until laboratory processing. Samples collected for morphological taxonomy were sieved in Espegrend Marine Research Field Station facilities using a 1 mm sieve and fixed in 96% ethanol for short term storage (ISO 2014). The physical properties of the water column were measured using a CTD sensor. These included oxygen concentration, salinity and temperature. Tomas Sørli recorded the CTD measurements on May 20th, September 22nd, October 7th, and December 15th in 2021, as well as January 21st and February 3rd in 2022. The measurements were made in the inner, shallow part of

Kviturspollen, the outer part by the sailing association, and outside of the poll (Mynteviken).

×

*Table 2.1: Overview of coordinates, depth and the type of sample collected for each station. Depth was not measured at station 2. \* At station 1, eDNA sample was only collected for 1B. \*\* At station 1 and 2, only a sub-sample of sediment was investigated for morphological identification.*

Station no.	Coordinates		Depth (m)	Samples collected	
	N	S		Morphological	eDNA
1	60°15'59.8"	5°14'46.8"	7.9	x**	x*
2	60°15'59.8"	5°14'41.6"	N/A	x**	x
3	60°15'58.7"	5°14'52.7"	3.8	x	x
4	60°15'55.5"	5°15'00.3"	4.2	x	x
5	60°15'49.4"	5°14'52.5"	4.2	x	x
6	60°15'57.6"	5°15'01.7"	4.9	x	x
9	60°15'35.5"	5°12'52.4"	6.3	x	x

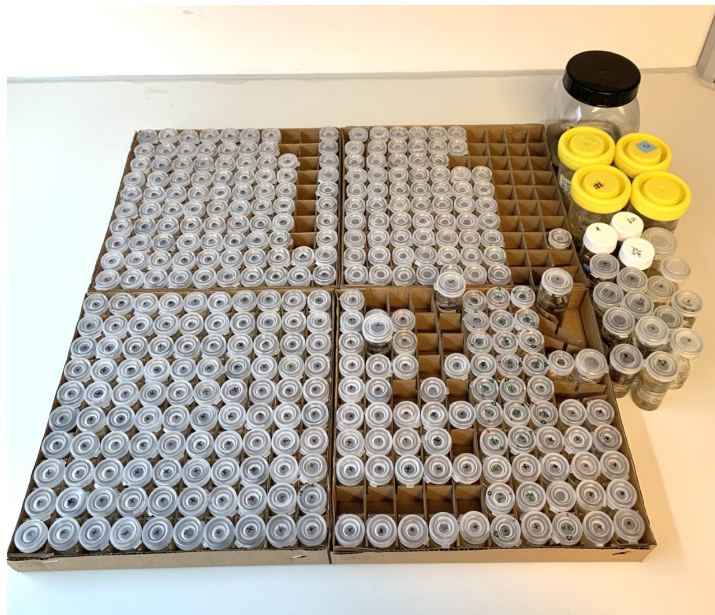


*Figure 2.2: Jonette Eckholdt and Torill Fjørtoft Johansen sampling in the field (left) and T. Johansen collecting eDNA sample (right) (Photo: Jon T. Hestetun)*

### Laboratory work

In the laboratory, morpho-taxa samples were separated from the sediment and sorted into higher taxonomic groups (Echinodermata, Mollusca, Polychaeta/Annelida, Crustacea) using a stereo microscope. The samples from stations 1 and 2 were high in organic matter content, and no animals were observed during sieving. The grab samples from these stations were therefore only sub-sampled to investigate whether any macroinvertebrates were present. For stations 3-6 and 9, the two sample replicates were sorted separately to account for differences between each grab sample for the same stations. The specimens were placed in separate containers containing 96% ethanol and stored until further identification. Containers were labeled with date of sampling, location, sample number, and taxonomic group using alcohol-proof paper.

Each individual in the samples was identified to the lowest possible taxonomic level (family, genus, or species), and subsequently quantified by T. Johansen and J. Eckholdt. During the identification, taxonomic experts were consulted to increase the accuracy of the taxonomic identification. Experts consulted were Henrik Glenner, David Rees, and Kenneth Meland for crustaceans, Tom Alvestad for polychaetes, and Jon Hestetun for mollusks. For quantification, only individuals with intact anterior parts (head) were included (ISO 2014). Specimens of the phyla Nematoda, Nemertea, and Phoronida were also sorted into their respective groups, but these were not quantified. A “+” symbol in the taxon list (see Section 5.6) indicates the relative occurrence of these species, as permissible by standard protocols (Direktoratsgruppen vanndirektivet 2018b, ISO 2014).



*Figure 2.3: Samples of morphologically identified species sorted into vials at the UiB lab.*

## 2.3 Metabarcoding laboratory processing

### 2.3.1 DNA extraction

DNA extraction was performed using the DNeasy PowerMax Soil Kit (Qiagen) and associated protocol (DNeasy PowerMax Soil Kit Handbook). Sediment samples were thawed at 4 °C, and approximately 10 g of each sample was mixed with 15 ml PowerBead Solution in a PowerMax Bead Tube. A negative control was included to control for contamination during the extraction, and was treated similarly as the samples. Solution C1 was added to the PowerMax Bead Tube before the solution was vortexed and placed on an orbital shaker, and remaining steps were carried out, strictly following the manufacturer's protocol. DNA extracts were stored at -20 degrees until further use. DNA quantity was determined using the Qubit dsDNA HS Assay Kit and a Qubit 4 fluorometer.

### 2.3.2 PCR amplification and high-throughput sequencing

PCR amplification was performed using duplicates of extracted DNA where the target gene was amplified. Two markers were targeted for amplification: the V1-V2 region of the 18S rRNA gene (~350-400 bp) and the mitochondrial cytochrome oxidase I gene (COI, 313 bp). The 18S rRNA gene was amplified using universal primer pair NSSU\_F04mod and NSSU\_R22mod (Sinniger et al. 2016), and the COI gene was amplified using universal primer pair mlCOIintFXT (Wangensteen et al. 2018) and jgHCO2198 (Geller et al. 2013). An overview of the primer pairs used in this study is given in Table 2.2. Two negative controls were included to account for cross-contamination in the different steps of PCR preparation. For amplification of the 18S gene, the PCR-mix contained 0.6  $\mu$ l of each primer (10  $\mu$ M), 12.5  $\mu$ l KAPA buffer, 0.3  $\mu$ l BSA, 0.2  $\mu$ l 3G Taq DNA Polymerase, 8.8  $\mu$ l water and 2  $\mu$ l of extracted DNA. For amplification of the COI gene, PCR-mix contained 2.5  $\mu$ l of each primer (10  $\mu$ M), 12.5  $\mu$ l KAPA buffer, 0.25 BSA, 0.2  $\mu$ l 3G Taq DNA Polymerase and 2  $\mu$ l of extracted DNA. This resulted in a reaction volume of 25  $\mu$ l for PCR amplification, and a total volume of 50  $\mu$ l after pooling samples for both markers.

×

Table 2.2: Overview of the 18S and COI primer pairs used in this study.

Genetic marker	Primers	Sequence (5'-3')	Source
COI	mlCOIintFXT	GGWACQRGWTGRAC- WITITAYCCYCC	Wangensteen et al. (2018)
	jgHCO2198	TAIACYTCIGGRTGICC- RAARAAYCA	Geller et al. (2013)
18S (V1-V2)	NSSU_F04mod	GCTTGWCTCAAAGA- TTAAGCC	Sinniger et al. (2016)
	NSSU_R22mod	CCTGCTGCCTTC- CTTRGA	Sinniger et al. (2016)

#### PCR protocol conditions for the first PCR

PCR protocol conditions for 18S consisted of initial 3 min step at 95 °C, 30 cycles with 30 s at 95 °C, 30 s at 57 °C and 30 s at 72 °C, followed by a 10 min step at 72 °C and a final step at 12 °C.

PCR protocol for COI consisted of initial 3 min step at 95 °C, 30 cycles with 20 s at 95 °C, 30 s at 57 °C and 30 s at 72 °C, followed by a 10 min step at 72 °C and a final step at 4 °C.

## Gel electrophoresis and purification

The pooled samples from the first round of PCR amplification were assessed using gel electrophoresis with 4.1  $\mu\text{l}$  GelRed on an 1.5% agarose gel. Pooled samples were subsequently purified by binding DNA to magnetic beads using Roche KAPA Pure Beads, according to the manufacturer's protocol. Two rounds of wash with 80% ethanol and elution with a low salt elution buffer (Tris) was performed as described in KAPA Pure Beads protocol. This procedure is essential for separating and removing primers, salts, primer-dimers, and dNTPs from the PCR products. Purified amplicons were quantified using the Qubit dsDNA HS Assay Kit and a Qubit 4 fluorometer, and subsequently diluted 1:200 for 18S and 1:100 for COI before library preparation. Control samples were diluted 1:10.

## Library preparation

For the library preparation, Illumina TruSeq dual index adapters i7/i5 and equimolar concentration of PCR product was used. First, a total of five mastermixes were prepared with primer 7 (Mix 1-5). The different mixes contained 225  $\mu\text{l}$  of KAPA buffer, 11.25  $\mu\text{l}$  of primer 7, 3.6  $\mu\text{l}$  KAPA 3G Taq DNA Polymerase and 142.65  $\mu\text{l}$  water. Mastermixes were allocated to PCR strips of eight wells as indicated by Table 2.3. Mix 1 to strip A, mix 2 to strip B and mix 3 to strip C and so on. This was prepared in pre-PCR lab. In the wells containing primer 7, 2.5  $\mu\text{l}$  of primer 5 (D501-8) was added as indicated by Table 2.3, giving each well a different combination primer 5 and 7. Finally, 5  $\mu\text{l}$  of DNA extract template was added.

×

*Table 2.3: Overview of primer pairs in PCR wells. Five master mixes prepared with primer 7 were added in each rows A-E (representing five PCR strips). Primer 5 was added to columns 1-8 giving each well (sample) a different combination of primer 5 and primer 7.*

	1	2	3	4	5	6	7	8
<b>A</b>	Mix 1 + D501	Mix 1 + D502	Mix 1 + D503	Mix 1 + D504	Mix 1 + D505	Mix 1 + D506	Mix 1 + D507	Mix 1 + D508
<b>B</b>	Mix 2 + D501	Mix 2 + D502	Mix 2 + D503	Mix 2 + D504	Mix 2 + D505	Mix 2 + D506	Mix 2 + D507	Mix 2 + D508
<b>C</b>	Mix 3 + D501	Mix 3 + D502	Mix 3 + D503	Mix 3 + D504	Mix 3 + D505	Mix 3 + D506	Mix 3 + D507	Mix 3 + D508
<b>D</b>	Mix 4 + D501	Mix 4 + D502	Mix 4 + D503	Mix 4 + D504	Mix 4 + D505	Mix 4 + D506	Mix 4 + D507	Mix 4 + D508
<b>E</b>	Mix 5 + D501	Mix 5 + D502	Mix 5 + D503	Mix 5 + D504	Mix 5 + D505	Mix 5 + D506	Mix 5 + D507	Mix 5 + D508



### PCR protocol conditions for the second PCR

PCR protocol conditions for the second PCR included a 3 min step at 95 °C, 15 cycles of 30 s at 95 °C, 30 s at 62 °C and 30 s at 72 °C, followed by a 5 min step at 72 °C and a final step at 4 °C.

### Gel electrophoresis and purification

After library preparation, the products were analysed using gel electrophoresis with 4.1  $\mu$ l GelRed on an 1.5% agarose gel before the products were purified using Roche KAPA Pure Beads, as previously described. DNA quantity was determined using Qubit dsDNA HS Assay Kit using a Qubit 4 fluorometer. The equimolar concentrations of each PCR product were pooled into a single sample, and DNA quantity of the pooled sample was measured. The pooled sample was analyzed by gel electrophoresis and NanoDrop before sequencing at the Norwegian Sequencing Centre (NSC) using an Illumina MiSeq instrument (University of Oslo, Norway).

## 2.4 Bioinformatic processing of metabarcoding sequence data

Data processing was conducted following standard protocols at NORCE on the NORCE in-house server, as described in Hestetun, Lanzén & Dahlgren (2021). Raw sequences from the NSC sequence run were automatically demultiplexed, and FASTQC files were quality-checked using FastQC v0.11.9 (Andrews 2010). Read pair merging, removal of singleton and chimera, and filtering were done in vsearch v2.15.2 (Rognes et al. 2016), and primers were subsequently trimmed with cutadapt v3.4 (Martin 2017). OTU clustering was performed using SWARM v.2.2.1 (Mahé et al. 2015), and post-clustering curation was done in LULU (Frøslev et al. 2017). Taxonomic assignment of OTUs was done against the Midori Reference 2 database (Leray et al. 2022) for COI and SilvaMod v138 (Quast et al. 2013) for 18S using CREST 4 (Lanzén et al. 2012). Low abundance sequences (<1%) were filtered to account for potential sequencing errors.

In standard monitoring, the biological quality element for coastal ecosystems is benthic invertebrates (Direktoratsgruppen vanddirektivet 2018b). Metazoans were, therefore, the target group in datasets. For the COI dataset, all non-metazoan OTUs and assumed terrestrial (Arachnida and Insecta) OTUs were removed from the datasets. All OTUs assigned to the same species were combined. For 18S, the entire 18S eukaryotic dataset was analyzed to get an overview of the complete eukaryotic diversity (available in Supplemental attachment). To specifically target metazoans, all metazoan OTUs were filtered from the total eukaryotic to make an additional 18S metazoan dataset. All species assigned to the same species were

combined before alpha and beta diversity analysis.

A list of unique metazoan taxonomic groups from the COI OTU table is in Appendix section 5.7, and unique metazoan taxonomic groups from the 18S OTU table in Appendix section 5.8. A list of unique eukaryotic taxonomic groups from the entire 18S dataset is available in the supplemental attachment.

## 2.5 Statistical analysis

To answer the research questions, alpha and beta diversity patterns across all datasets were analysed and compared using Primer v7 (Clarke & Gorley 2015) and R (R Core Team 2022). The figures presented in the results were made using the ggplot2 package in R (Wickham 2016).

### 2.5.1 Alpha diversity

#### Diversity indices

Diversity occurs at different scales, and patterns of diversity are often explained through alpha, beta and gamma diversity (Whittaker et al. 2001). Alpha diversity is defined as the local within-habitat species diversity, often measured in richness (Whittaker 1972).

In accordance with standard monitoring, the Shannon diversity index was calculated from each dataset (morphological, COI and both 18S datasets) in Primer v7 (Clarke & Gorley 2015). State values for diversity is only calibrated for morphological data, so these are only applicable for the morphological dataset.

#### *Shannon Diversity Index $H'$*

The Shannon diversity index (Shannon & Weaver 1963) describes diversity in terms of richness and evenness, and uses the number of species in the sample (Rygg 2006, Whittaker et al. 2001). The formula is as follows:

$$H' = - \sum (p_i)(\log_2 p_i) \quad (2.1)$$

where  $p_i$  is the proportion of individuals of the total sample belonging to species  $i$  (Rygg 2006). In accordance with Direktoratgruppen vanndirektivet (2018b),  $\log_2$  is used when calculating the index and the calculation was performed using Primer7.

#### *Hurlbert Diversity Index ( $ES_n$ )*

The Hurlbert diversity index was not calculated directly for the datasets in this study, but is described since it is a part of several of the biotic indices. This index describes the expected

number of species selected randomly from a sample of  $N$  individuals,  $S$  number of species and  $N_i$  individuals of species  $i$  (Hurlbert 1971, Pedersen et al. 2016, Rygg & Norling 2013). The diversity measure  $ES_{100}$  is used to calculate both ISI and NSI (Rygg & Norling 2013).  $ES_{100}$  is the expected number of species among 100 individuals (Rygg 2002). The formula of  $ES_n$  is:

$$ES_n = \sum_i^S \left[ 1 - \frac{\binom{N-N_i}{n}}{\binom{N}{n}} \right] \quad (2.2)$$

### Biotic indices

The sensitivity indices are based the different sensitivities of benthic invertebrates to environmental disturbance (Borja et al. 2000, Rygg 2002). As described by Grall & Glémarec (1997), species can be divided into groups (I-V) based on their tolerance to stress from organic matter enrichment. Species in group I are sensitive to stress by organic enrichment, while group V are referred to as first-order opportunistic species that may proliferate in organic matter enriched conditions (Grall & Glémarec 1997). The ecological groups are distributed according to their sensitivity to organic enrichment, and this is the basis of biotic indices which infer the different stages of community degradation (Borja et al. 2000, Grall & Glémarec 1997).

The biotic indices highlighted in *Direktoratsgruppen vanndirektivet (2018b)* are the Indicator Species Index ( $ISI_{2012}$ ) and Norwegian Sensitivity Index ( $NSI_{2012}$ ) for sensitivity and Norwegian Quality Index 1 (NQI1) which is a composite index. ISI and NSI have been revised in recent years, and new versions were developed in 2018 ( $ISI_{2018}$  and  $NSI_{2018}$ ), which include a new species list based on new sensitivity values for soft-bottom invertebrates and new class boundaries (Borgersen et al. 2020). However, in accordance with previous analysis of Kviturspollen (Kvalø et al. 2015, 2014), this study used the 2012-versions of NSI and ISI. As a commonly used index in European waters, the AZTI Marine Biotic Index (AMBI) was also calculated (Borja et al. 2000).

The indices were calculated from the morphological, COI and the 18S metazoan dataset using the BBI package version 0.3.0 in R (Cordier & Pawlowski 2018, R Core Team 2022), and compared to the class values as described below.

#### *Indicator Species Index (ISI)*

The Indicator Species Index is a Norwegian presence/absence sensitivity index for assessing ecological quality based on benthic invertebrates (Rygg 2002, Rygg & Norling 2013). The index is based on the observation of species' different sensitivities to environmental conditions, so the presence or absence of a species indicating a given condition can be used to calculate the Indicator Species Index (Rygg 2002). ISI is based on the  $ES_{100}$  values where sensitivity values are derived from diversity of the samples from which the species occurred (Rygg 2002), and ISI value of a sample is defined by Rygg (2002) as “the average of the sensitivity values

( $ES_{100min}$ ) of the taxa occurring in the sample.” ISI is a qualitative index (Rygg & Norling 2013).

$$ISI = \sum_i^S \left[ \frac{ISI_i}{S_{ISI}} \right] \quad (2.3)$$

where  $ISI_i$  is the sensitivity value of species  $i$  and  $S_{ISI}$  is the number of species with assigned value (Rygg & Norling 2013).

#### *Norwegian Sensitivity Index (NSI)*

The Norwegian Sensitivity Index is a quantitative index that includes abundance in the calculation. Species sensitivity status ( $ES_{100min}$ ) is calculated ( $NSI_i$ ) and weighted in the NSI calculations using species abundance ( $N_i$ ) (Pedersen et al. 2016, Rygg & Norling 2013).

$$NSI = \sum_i^S \left[ \frac{N_i * NSI_i}{N_{NSI}} \right] \quad (2.4)$$

where  $N_i$  is the number of individuals of species  $i$ ,  $NSI_i$  is the sensitivity value of species  $i$  and  $N_{NSI}$  is the number of individuals of assigned sensitivity values (Rygg & Norling 2013).

#### *AZTI Marine Biotic Index (AMBI)*

Based on the work of Pearson & Rosenberg (1978), Glémarec & Hily (1981) and Grall & Glémarec (1997), a benthic biotic index (sensitivity index) was proposed by Borja et al. (2000), who inferred an index with eight levels from 0 (unpolluted) to 7 (extremely polluted) (Borja et al. 2000). This sensitivity index is called AMBI (AZTI's Marine Biotic Index) and was developed to assess the benthic ecological quality of European estuarine and coastal environment (Borja et al. 2000, Borja & Muxika 2005, Borja et al. 2003).

AMBI is a sensitivity index used in Spain, the UK, Ireland, and Denmark (Rygg 2006), where species are assigned to ecological groups (EG) according to sensitivity (Pedersen et al. 2016, Rygg & Norling 2013). More recent reports recommend using NSI as a sensitivity component rather than AMBI in the Norwegian classification system (Rygg & Norling 2013). This is because NSI is based on Norwegian macroinvertebrate data while AMBI is based on South European data, and species' sensitivity may differ between geographical regions (Rygg & Norling 2013). The index is a component of the Norwegian Quality Index 1 (NQI1) (Pedersen et al. 2016).

$$AMBI = \sum_i^S \left[ \frac{N_i * AMBI_i}{N_{AMBI}} \right] \quad (2.5)$$

where  $AMBI_i$  represents the assigned tolerance value and  $N_{AMBI}$  is the number of individuals

with assigned tolerance value (Rygg & Norling 2013)

### Norwegian Quality Index 1, NQI1

NQI1 is a composite index that accounts for both sensitivity and diversity, and it includes AMBI in the calculations. NQI1 displays values between 0 and 1 (Pedersen et al. 2016). The formula is:

$$NQI1 = [0,5 * (1 - \frac{AMBI}{7}) + 0,5 * (\frac{\ln(S)}{2,7}) * (\frac{N}{N+5})] \quad (2.6)$$

where  $N$  is the number of individuals and  $S$  is the number of species (Pedersen et al. 2016).

### Class values for alpha diversity metrics

Kviturspollen is registered as a sheltered fjord and water type M3 in Vann-nett.no (Vann-nett n.d.). To assess the ecological state of a water body by using the above-mentioned indices, the values were compared to class values as indicated in Table 2.4 and 2.5. These class limits are taken from Veileder 02:2018 according to its water type (Direktoratsgruppen vanndirektivet 2018b, p.167) and AMBI values are taken from Borja et al. (2003).

×

Table 2.4: Class values for alpha diversity metrics of soft-bottom invertebrates including  $H'$ ,  $ES_{100}$ ,  $ISI$ ,  $NSI$  and  $NQI1$  (Direktoratsgruppen vanndirektivet 2018b)

Index	Type	Excellent	Good	Moderate	Poor	Very poor
$H'$	Diversity	5.9-3.9	3.9-3.1	3.1-2	2-0.9	0.9-0
$ES_{100}$	Diversity	52-26	26-18	18-10	10-5	5-0
$ISI$	Sensitivity	13.1-8.5	8.5-7.6	7.6-6.3	6.3-4.5	4.5-0
$NSI$	Sensitivity	29-24	24-19	19-14	14-10	10-0
$NQI1$	Composite	0.90-0.72	0.72-0.63	0.63-0.49	0.49-0.31	0.31-0

×

Table 2.5: Site pollution classification for AMBI based on values from Borja et al. (2003).

Index	Type	Unpolluted	Slightly polluted	Meanly polluted	Heavily polluted	Extremely polluted
AMBI	Sensitivity	0-1.2	1.2-3-3	3.3-5.0	5.0-6.0	>6

## 2.5.2 Beta diversity

According to Whittaker (1972), beta diversity is the differentiation between communities along a habitat gradient. In the case of this study, this translates to the difference in species composition between stations, and this will be expressed using multivariate analysis based on Bray-Curtis pairwise similarity (Bray & Curtis 1957). This index deals with abundance data, and looks at the pairwise similarity/dissimilarity in abundance between samples (Chao et al. 2005, Ricotta et al. 2021).

$$BC_{UV} = \frac{\sum_{(j=1)^N} |X_{Uj} - X_{Vj}|}{\sum_{(j=1)^N} |X_{Uj} + X_{Vj}|} \quad (2.7)$$

(Ricotta et al. 2021)

Bray-Curtis similarity (0-100) for the different datasets was calculated using Primer version 7.0.21 (Clarke & Gorley 2015). The data was transformed to fourth-root to reduce contribution of the most abundant species (Clarke & Gorley 2015), and the resulting distance matrix is visually expressed/analysed using non-metric multidimensional scaling. Ordination, such as nMDS, is used to visualize trends between sites in terms of species composition. The method shows the pairwise similarities from the distance matrix as points in the ordination (Hui et al. 2014). Stress values related to the nMDS plots indicate how well the represented the patterns are, where low stress indicate better representation (Clarke & Gorley 2015).

The correlation between the Bray-Curtis distance matrices for the different datasets was tested with a non-parametric form of a Mantel test using the relate function in primer (Clarke & Gorley 2015, Mantel 1967). This is a permutation test for rank correlation between the Bray-Curtis similarity matrices based on Spearman's rank correlation rho ( $\rho$ ). A set number of permutations (randomization) is used to test the null hypothesis of “no agreement in multivariate pattern” between the matrices. The  $\rho$  generated from the permutations is compared to the real  $\rho$ , and the null hypothesis can be rejected at  $p < 1\%$  (Clarke & Gorley 2015).

# Chapter 3

## Results

### 3.1 CTD results

CTD measurements including salinity, dissolved oxygen saturation and temperatures were measured for the inner and outer part of Kviturspollen and a reference station outside of the study area (Mynteviken).

#### 3.1.1 Salinity

Results from salinity measurements are illustrated in Figure 3.1, and show an overall increase in salinity with increasing depth for all stations. Salinity was highest in Mynteviken with surface values starting at 29-30‰ closing up to 35‰ at the sea floor. The inner part of Kviturspollen had the lowest recorded salinity at 27‰ in December 2021 at the surface water, and highest salinity at approximately 32‰. The sailing association also had lower salinity for December 2021 at 28‰, but a clear increase towards the ocean floor around 33-34‰.

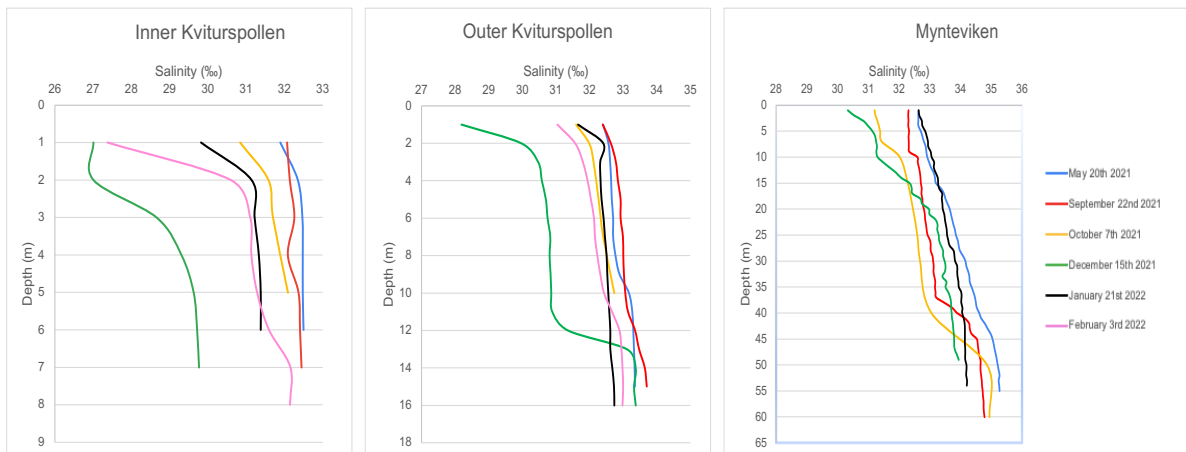


Figure 3.1: Salinity (‰) in the inner and outer part of Kviturspollen and Mynteviken.

### 3.1.2 Oxygen

Dissolved oxygen saturation at the three different locations is shown in Figure 3.2. The oxygen saturation in the surface water ranged from 70-90% in the months from September 2021 to February 2022, and a significantly higher value of approximately 130% in May 2021 across all locations. The figure illustrating dissolved oxygen saturation in Mynteviken shows a clear decrease in dissolved oxygen with depth. For the inner part of Kviturspollen, the dissolved oxygen saturation varied more with depth for the different measurements with a slight increase in December 2021 and February 2022 before decreasing closer to the sea floor. The dissolved oxygen saturation in the outer part of Kviturspollen started at around 90% for the months from September 2021 to February 2022, and saturation decreased close to the sea floor. For May 2021, the outer part shows the highest value at approximately 130% close to the surface, and decreasing reaching the sea floor.

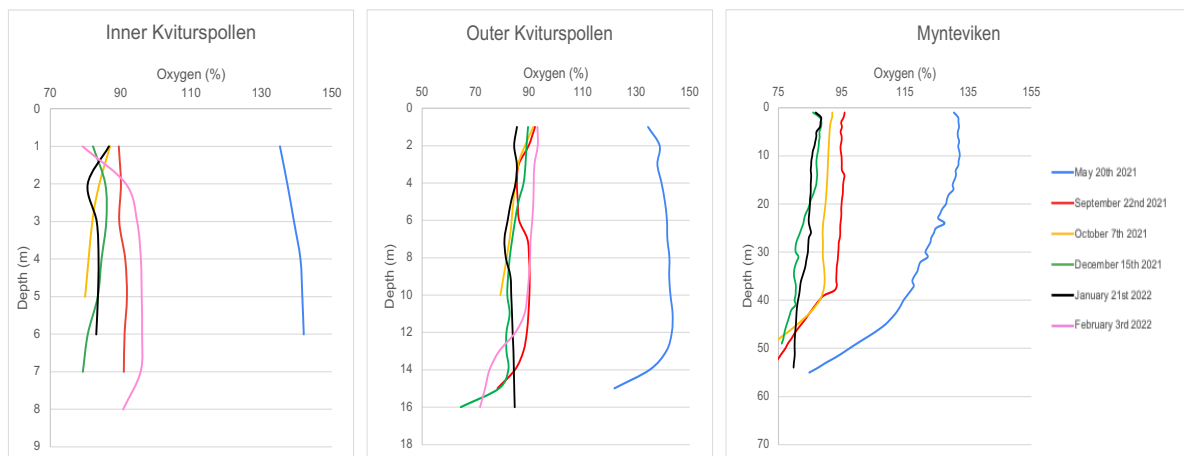


Figure 3.2: Dissolved oxygen saturation (%) in the inner and outer part of Kviturspollen and in Mynteviken.

### 3.1.3 Temperature

The temperature at the three locations from May 2021 to February 2022 is shown in Figure 3.3. December 2021, January and February 2022 had the lowest temperatures across all locations ranging from around 5-10 °C. October and September were the warmest months with surface temperatures ranging from 13-15 °C. Mynteviken displays the most evident changes with depth, while the inner and outer parts of Kviturspollen are more stable.



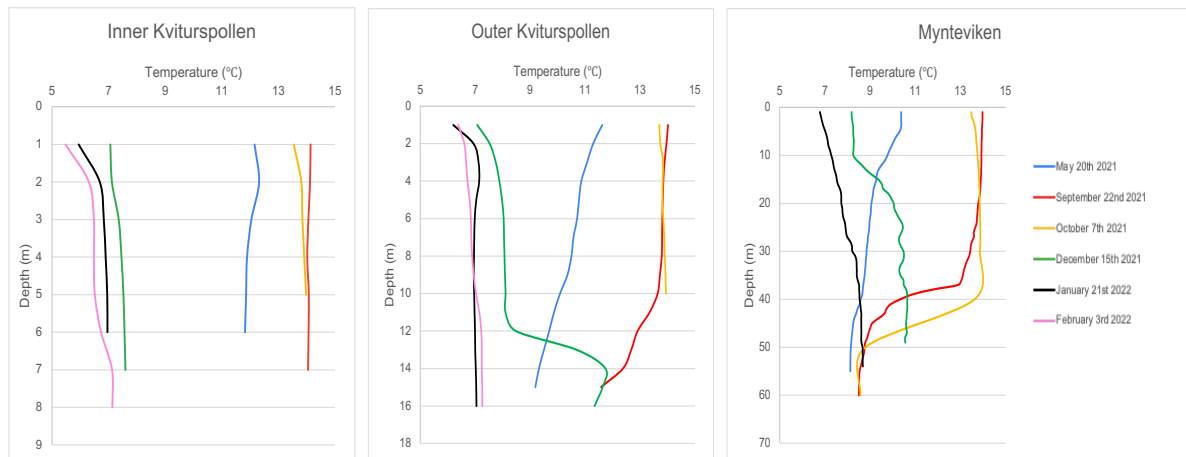


Figure 3.3: Temperature ( $^{\circ}\text{C}$ ) in the inner and outer part of Kviturspollen and in Mynteviken.

## 3.2 Morphological taxonomy

During sampling in the field, sediment type, smell and other observations were noted. A strong smell of  $\text{H}_2\text{S}$  was noted for stations 1 and 2 with a predominance of black mud and organic matter for station 1, and eelgrass on muddy sediments at station 2. A weak smell of  $\text{H}_2\text{S}$  was noted for station 4 and 5, which both consisted mixed sediment of mud and gravel. No smell was noted for stations 3, 6 and 9. An overview of sediment type is given in Table 3.1.

×

Table 3.1: An overview of observations made in the field including sediment type and smell at each station.

Station no.	Observations in the field	
	Sediment type	Smell
1	Black mud with organic material	Strong $\text{H}_2\text{S}$ smell
2	Brown eelgrass on dark mud	Strong $\text{H}_2\text{S}$ smell
3	Shell sand with larger grain	No smell
4	Muddy sediment with large rocks.	Weak $\text{H}_2\text{S}$ smell.
5	Muddy sediment with gravel and rocks.	Weak $\text{H}_2\text{S}$ smell.
6	Shell sand mixed with sand and mud.	No smell
9	Sand and shell sand. Large pieces of green algae in grab 9B.	No smell

A total of 2779 benthic invertebrates made up a list of 93 taxa covering four phyla, 26 orders, 50 families and 62 genera, and 43 species (See section 5.6). Individuals from Nematoda, Nemertea, and Phoronida were also noted in the taxon list but were not quantified. Thus, the number of phyla quantified by morphology was four, but the total number detected was seven.

From the assemblage of 93 morphologically assigned taxa, 46% were assigned to species level, 28% to genus level, and 26% to higher taxonomic levels. From the 2779 metazoan individuals, 73% were assigned to species, 20% to genus, 6.6% to family and 0.4% to class (See section 5.6). The most abundant phylum was Annelida (64%), and the second most abundant was Mollusca (28%). Abundance at the class level for the morphological dataset is shown in Figure 3.4, with grab samples being dominated by individuals from either Polychaeta or Bivalvia. The most abundant species from morphological taxonomy were the polychaete *Protodorvillea kefersteini* (869 specimens), the mollusk *Kurtiella bidentata* (630 specimens), and a polychaete species from genus *Platynereis* (181 specimens).

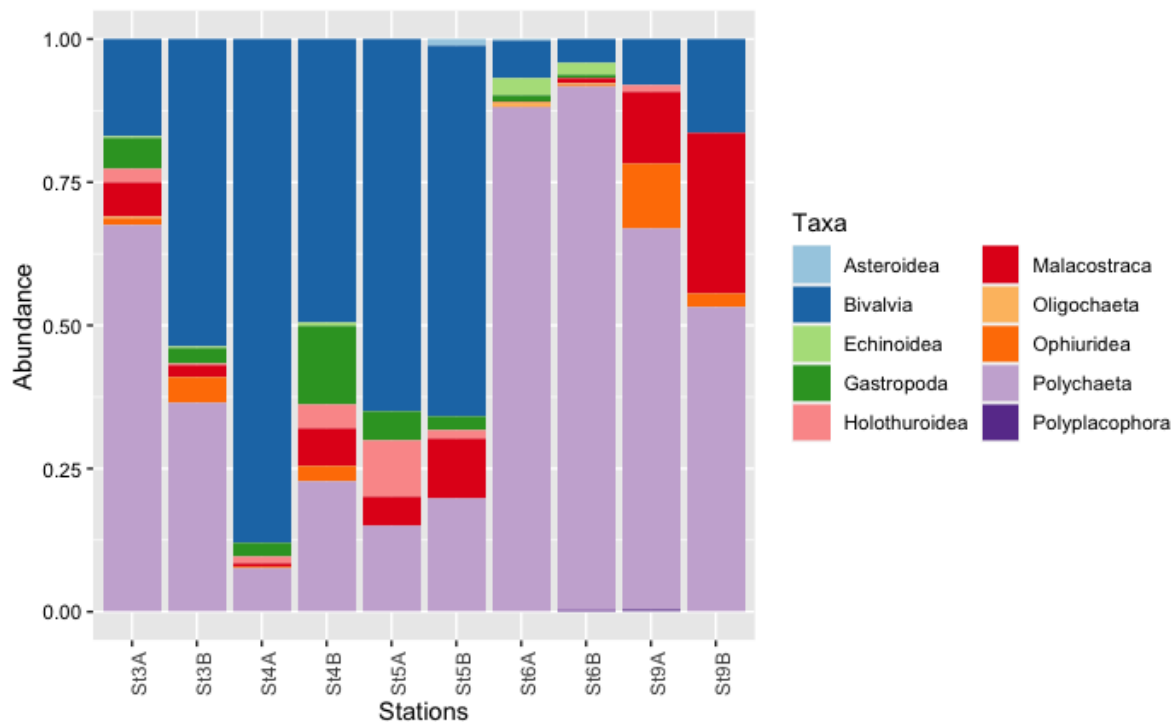


Figure 3.4: Abundance at class level from the morphological dataset.

×

Table 3.2: Number of species per phylum and total from 1 mm sieve samples.

	3A	3B	4A	4B	5A	5B	6A	6B	9A	9B
Annelida	189	136	17	43	3	15	678	511	146	42
Crustacea	16	7	1	12	1	8	1	5	28	22
Echinodermata	11	19	4	14	2	2	24	13	27	2
Mollusca	62	209	205	119	14	51	61	28	19	13
Total	278	371	227	188	20	76	764	557	220	79
Number of taxa	33	33	10	19	7	10	38	41	41	19

A summary of the primary data from the morphology-based taxonomy is given in Table 3.2. The table summarizes the number of individuals of each phylum, the number of individuals, and the number of different taxa found in each grab replicate for stations 3-6 and 9. No animals were found in samples collected from stations 1 and 2. The highest number of individuals were found at station 6A, with 764 distributed among 38 taxa. The most abundant species from this grab sample/station was the polychaete *Protodorvillea kefersteini* (424 specimens, 55.5%), followed by the polychaete *Platynereis* sp. (66 specimens, 8.6%) and mollusk *Kurtiella bidentata* (42 specimens, 5.5%). Station 6B had the highest number of different taxa and the second highest number of individuals, with 557 distributed among 41 species. The most abundant species was the polychaete *Protodorvillea kefersteini* (353 specimens, 63.4%), followed by the polychaete *Cirriiformia tentaculata* (44 specimens, 7.9%) and polychaetes identified as *Cirratulidae* indet. (23 specimens, 4.1 %). Station 5A had the lowest abundance, with 20 individuals distributed among seven species (Table 3.2, See Section 5.6).

### 3.3 High-throughput sequencing output

A total of 13 samples for each genetic marker (COI and 18S) were sent for sequencing representing the two sample replicates for each station (stations 1-6 and 9). The exception is station 1, where only one sample (1B) was sent for sequencing. In addition, one extraction control and two negative PCR controls were included to account for contamination during extraction and PCR. The sequence reads resulting from sequencing were bioinformatically processed, ultimately producing operational taxonomic units (OTUs) representing the taxa detected by each genetic marker.

#### 3.3.1 Sequencing output for COI

Environmental DNA metabarcoding using the COI marker yielded a total of 5,075,018 raw reads, of which 4,485,794 remained after the removal of singletons and chimeras and quality filtering. The sequences were filtered based on an abundance threshold (0.0001 abundance threshold, reads <25), and 4,279,740 remained. Table 3.3 gives a summary of COI sequence reads per station. The taxonomic assignment was done using the MIDORI v2 database. A total of 6298 OTUs were obtained, but 2504 were not assigned to a taxon ("no hits"). From the total number of reads, 28.6% of the reads (1,225,263 reads) could be assigned to metazoan taxa (incl. negative controls). The remaining reads consisted of non-metazoan taxa such as bacteria (1,064,971 reads, 24.9%) and protists Stramenopiles, Alveolata, and Rhizaria (SAR group) (632,758 reads, 14.8%) and unassigned reads (999,941 reads, 23%). In the negative controls, 76 reads were detected. This corresponds to 0.0018% of the total number of reads, indicating no contamination. Relative abundances for the COI dataset after abundance filtering are shown in Figure 3.5 (not including "no hits" reads or controls). The most abundant OTUs in

the COI dataset were the oligochaete *Tubificoides benedii*, proteobacterium *Rubrivivax gelatinosus*, and an unknown species from phylum Annelida. The four metazoan OTUs present in extraction control represented *Tubificoides benedii*, *Macrochaeta clavicornis*, an unknown Arthropoda, and an unknown Annelida. The OTUs present in the negative controls were an unknown species of Porifera, and the two OTUs in the other negative control were an unknown Annelida and *Pholoe baltica*.

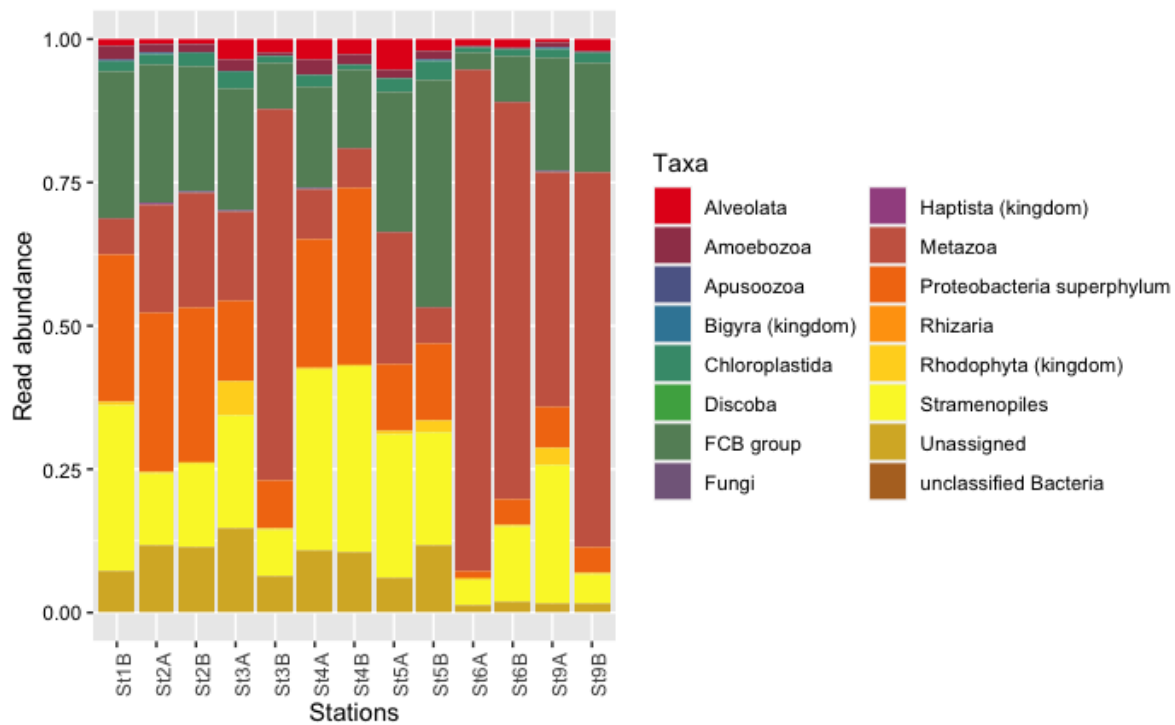


Figure 3.5: Read abundance for higher taxonomic level for entire COI dataset after abundance filtering, but not including “no hits” reads. “Unassigned” refers to all reads not assigned to the taxonomic level included in the figure.

COI is used for targeting metazoans, making other taxonomic groups non-target OTUs, and these were removed from the dataset. After the removal of non-target OTUs, 789 OTUs remained. From the COI metazoan dataset, all OTUs with probable terrestrial origin (Insecta and Arachnida) were removed, and all OTUs assigned to the same species were combined. Excluding reads from control samples and all terrestrial reads, the total metazoan reads were reduced to 1,214,914, and the total number of OTUs was reduced to 753. From this number of reads, 40.7% were assigned to species level (494,864), of which Annelida made up the majority of reads (477,210) represented by 11 species-level OTUs. The most abundant species was annelid *Tubificoides benedii* with 356,892 reads, making up most of the reads assigned to species level (72%). Furthermore, 5.4% of the reads were assigned to genus level, 0.3% to family, 0.2% to order, 7.4% to class, 36.6% to phylum (444,628 reads) and 9.5% to kingdom Metazoa.

From the total number of OTUs, 34 OTUs were assigned to species, 27 OTUs to genera, 5

OTUs to family, 9 OTUs to order, 52 OTUs to class, 341 to phylum, and 285 to kingdom level (Metazoa) (See Section 5.7).

×

Table 3.3: Table giving the per station and total number of COI sequence reads at different stages of data processing; raw, after quality filtering and after abundance filtering. Total number of OTUs in entire COI dataset and total number of metazoan OTU are summarized. “Exc” and “NTC” refers to extraction control and PCR negative controls, respectively.

Station no.	COI sequence reads			Total number of OTUs	Metazoan OTUs
	Raw	After quality filtering	After abundance filtering		
1B	528,299	481,734	451,498	4100	381
2A	421,716	382,550	366,315	2851	237
2B	398,952	360,797	344,032	3033	264
3A	359,247	306,964	290,425	3303	348
3B	385,570	327,510	313,624	2959	310
4A	317,045	253,031	242,131	2196	199
4B	347,707	282,744	271,373	2804	263
5A	301,322	251,935	240,549	2282	188
5B	337,188	305,403	291,872	3088	252
6A	479,895	422,184	414,535	2165	231
6B	411,785	381,795	370,036	2716	256
9A	360,925	335,752	303,909	2919	324
9B	424,975	393,315	377,391	2390	211
Exc	162	59	55	28	4
NTC	107	10	10	8	1
NTC	123	11	11	10	2
<b>Total</b>	<b>5,075,018</b>	<b>4,485,794</b>	<b>4,279,740</b>	<b>6298</b>	<b>789</b>

### 3.3.2 Sequencing output for 18S

For the 18S marker, eDNA metabarcoding yielded 6,898,910 raw reads, of which 6,399,596 remained after singletons and chimera removal and quality filtering. After read-abundance filtering (abundance <37), 6,118,033 sequences remained. A total of 3790 OTUs were obtained, and 15 of these were not assigned to a taxon ("no hits"). A summary of sequence reads, and OTUs per station is given in Table 3.4. From the 18S reads, 13.4% could be assigned to metazoan taxa (804,146 reads). The remaining reads consisted of non-metazoan taxa such as fungi (180,784 reads, 3%) and the SAR group (4,513,236, 74%). "No hits" reads only made up 0.17% of total reads. Relative abundances the 18S dataset after abundance filtering (exclud-

ing control samples and "No hits" reads) are shown in Figure 3.6, and the dominance of the SAR group is evident. The most abundant OTUs were a stramenopile in the phyla Ochrophyta, a species of Ochrophyta (Araphid-pennate), and a dinoflagellate of genus *Heterocapsa*. The most common metazoan OTUs were a nematode from the genus *Bathylaimus*, the ostracod *Cytheromorpha acupunctata*, and the nematode *Spirinia parasitifera*. For the negative controls, 640 reads were detected, corresponding to 0.01% of the total number of reads.

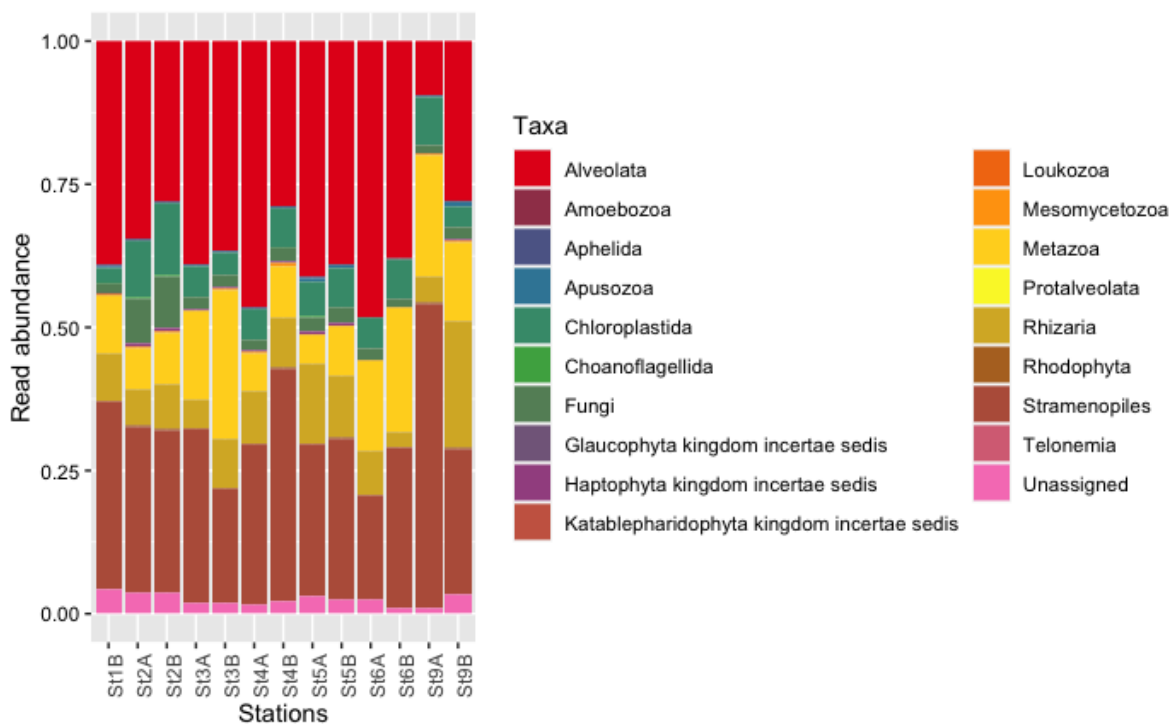


Figure 3.6: Read abundance for higher taxonomic level for entire 18S dataset after abundance filtering, but not including "no hits" reads. "Unassigned" refers to all reads not assigned to the taxonomic level included in the figure.

Two separate datasets were made from the 18S data, one including all Eukaryota (See supplemental attachment) and one with metazoans only (See section 5.8). In further processing of the Eukaryota dataset, all "no hits" and terrestrial OTUs were removed from the dataset, and all non-species OTUs placed in the species column (e.g., *Askenasia* sp.) were moved to the genus column. Of the remaining species, all OTUs assigned to the same species were combined, resulting in a dataset of 3713 OTUs. In further processing of the 18S Metazoan dataset, all OTUs with terrestrial origin (e.g., Arachnida - 8 OTUs) were removed, and all OTUs assigned to the same species were combined (e.g., all 16 *Calomicrolaimus parahonestus* OTUs were combined to one). These steps resulted in the 18S metazoan dataset of 766,528 reads with 298 OTUs (not including reads from control samples). From the total number of reads, 33% were assigned to species level (253,231 reads), of which Arthropoda made up the majority of reads (147,968 reads, 58 %) represented by 11 OTUs. The most abundant species detected by 18S was the ostracod *Cytheromorpha acupunctata* (57,193 reads), making up 22.5% of species.

Furthermore, 17.2% of the reads were assigned to genus level, 7.2% to family, 39.8% to order, 1.2% to class, 0.2% to phylum and 1.3% to kingdom Metazoa). On the level of OTUs, 36 OTUs were assigned to species, 41 OTUs to genus, 48 OTUs to family, 131 OTUs to order, 28 OTUs to class, 7 OTUs to phylum, and 6 to kingdom level (Metazoa) (Section 5.8).

×

*Table 3.4: Table giving the per station and total number of 18S sequence reads at different stages of data processing; raw, after quality filtering and after abundance filtering. Total number of OTUs in entire 18S dataset and total number of metazoan OTU are summarized. “Exc” and “NTC” refers to extraction control and PCR negative controls, respectively.*

Station no.	18S sequence reads			Total number of OTUs	Metazoan OTUs
	Raw	After quality filtering	After abundance filtering		
1B	506,476	472,628	448,475	2286	106
2A	517,492	479,385	461,105	1588	77
2B	523,560	486,971	465,843	1892	90
3A	555,217	512,638	491,940	2026	135
3B	606,829	562,489	544,560	2121	136
4A	570,179	535,095	520,324	1833	73
4B	600,425	558,045	375,049	2177	71
5A	517,217	481,697	483,456	1578	44
5B	463,343	430,354	535,712	2043	76
6A	476,033	438,525	463,617	1843	131
6B	535,222	495,850	414,111	1742	100
9A	515,769	473,325	458,771	1523	158
9B	511,148	472,594	454,430	1654	79
Exc	248	128	488	84	17
NTC	1230	505	123	73	11
NTC	202	30	29	26	2
<b>Total</b>	<b>6,900,590</b>	<b>6,400,259</b>	<b>6,118,033</b>	<b>3790</b>	<b>339</b>

### 3.4 Comparison of taxonomic composition

The morphological results identified and quantified a total of 93 taxa. These taxa covered four phyla, 22 orders, 49 families, 62 genera, and 43 species. When including the non-quantified phyla Nematoda, Nemertea, and Phoronida, the total number of phyla observed was 7. The most abundant phyla in the morphological dataset were Annelida, making up 64% of the data, and the remaining were Mollusca (28%), Echinodermata (4%) and Arthropoda (4%). Most taxa assigned to species level belonged to phylum Mollusca (most to Bivalvia). The number

of taxa and species detected per site is given in Figure 3.7 and 3.8, showing no findings at stations 1 and 2. The highest number of species was found at stations 6B and 3A (Figure 3.8).

The metazoan community described by the COI dataset, a total of 789 OTUs, covered 18 phyla, 28 classes, 34 orders, 49 families, 48 genera, and 34 species. The per-site detection of taxa was highest at station 1B, while per-site detection of species was highest at stations 3A and 9A (Figure 3.7 and 3.8). The most abundant phylum in the COI dataset was Annelida which constituted 79% of the data, similar to the morphological data, but 37% of the reads was only assigned to the phylum level only (i.e., OTUs assigned “Annelida”). Furthermore, in contrast to 18S results, large proportions of reads were unassigned to phylum level (i.e. only assigned to “Metazoa”) for COI, as shown in Figure 3.9. Most OTUs assigned to species level for COI belonged to phylum Annelida (most to class Polychaeta). However, although Annelida made up most of the COI data, there are differences between stations when looking at Figure 3.9. At station 6, Annelida made up 97% of the reads at 6A and 92% at 6B. In contrast, at station 5B, Annelida only made up 2% of the reads.

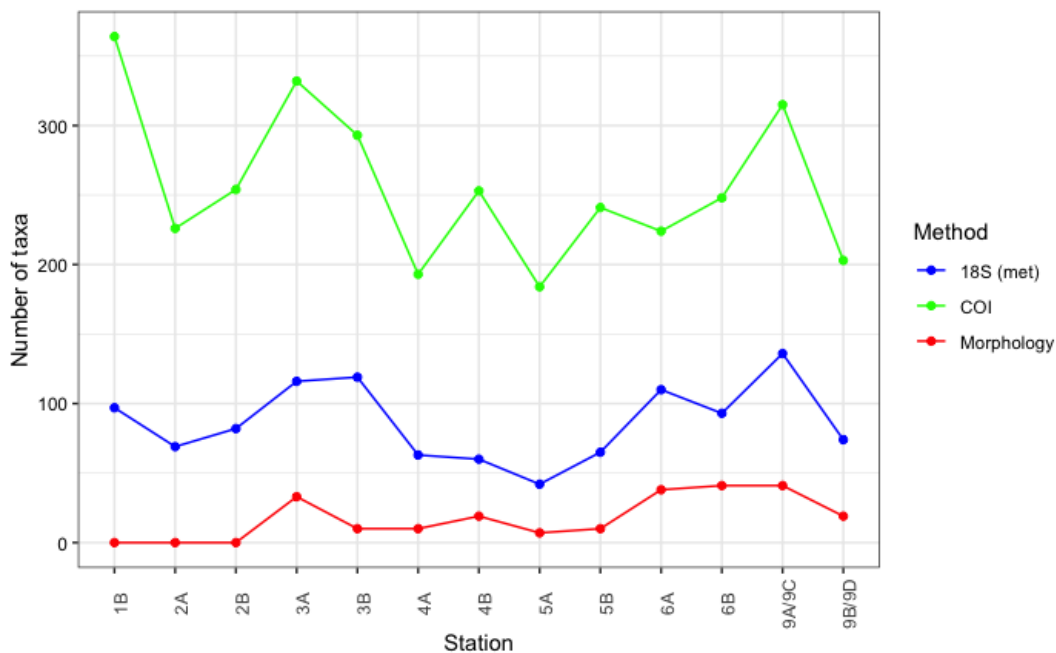


Figure 3.7: Number of metazoan taxa per grab replicate/station detected by eDNA metabarcoding using 18S and COI and by the morphological identification. 9A and 9B refers to grab samples for morphology, while 9C and 9D refers to grab samples for metabarcoding.



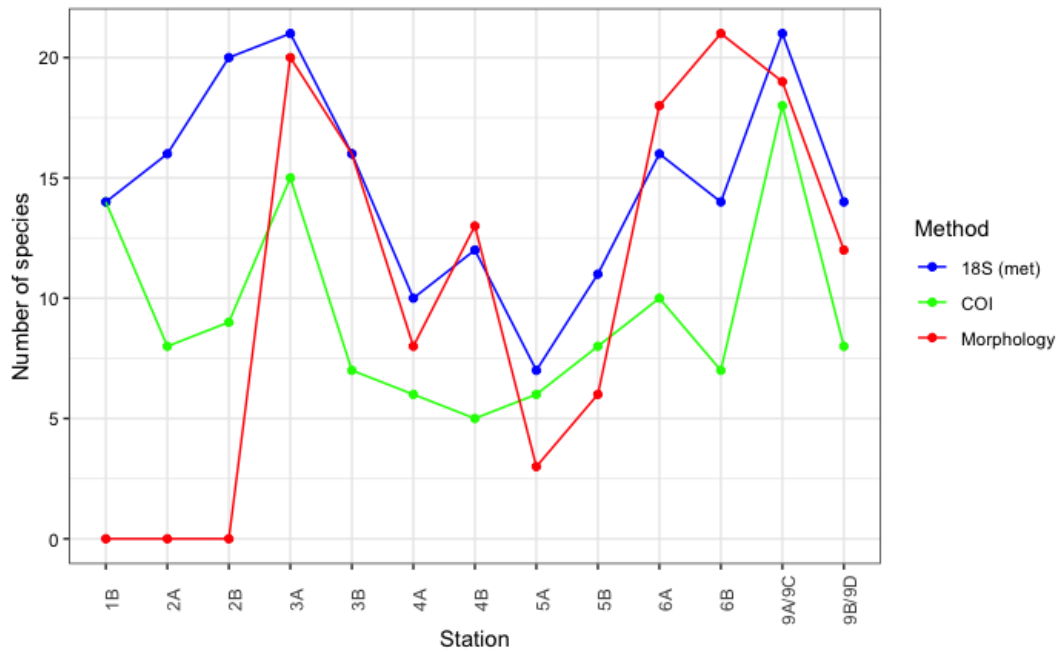


Figure 3.8: Number of metazoan species per grab replicate/station detected by eDNA metabarcoding using 18S and COI and by the morphological identification. 9A and 9B refers to grab samples for morphology, while 9C and 9D refers to grab samples for metabarcoding.

The eukaryotic community described by the entire 18S dataset was dominated by Stramenopiles, Alveolates, and Rhizaria, making up 74% of the dataset, while Metazoa made up 13.4% of the dataset in number of reads (Figure 3.6). Metazoa made up a smaller part of the 18S eukaryotic dataset compared to the complete COI dataset. The metazoan community as described by 18S is represented by 766,528 reads distributed amongst 298 OTUs covering 15 phyla, 29 classes, 47 orders, 69 families, 68 genera, and 36 species. Compared to COI, taxonomic coverage is greater in 18S (more taxonomic groups). The most abundant phylum in the 18S metazoan dataset was Nematoda, making up 55% of the total number of reads (419,801) representing 157 OTUs. Nematoda was the most abundant phylum in 11 out of 13 samples, while Arthropoda dominated the remaining two (Figure 3.9). The majority of OTUs assigned to species level belonged to phylum Arthropoda (most to class Ostracoda). The per station detection by 18S shows that the highest number of taxa were detected from 3B and 9A, and the highest number of species were detected at 3A and 9A (Figure 3.7 and 3.8).

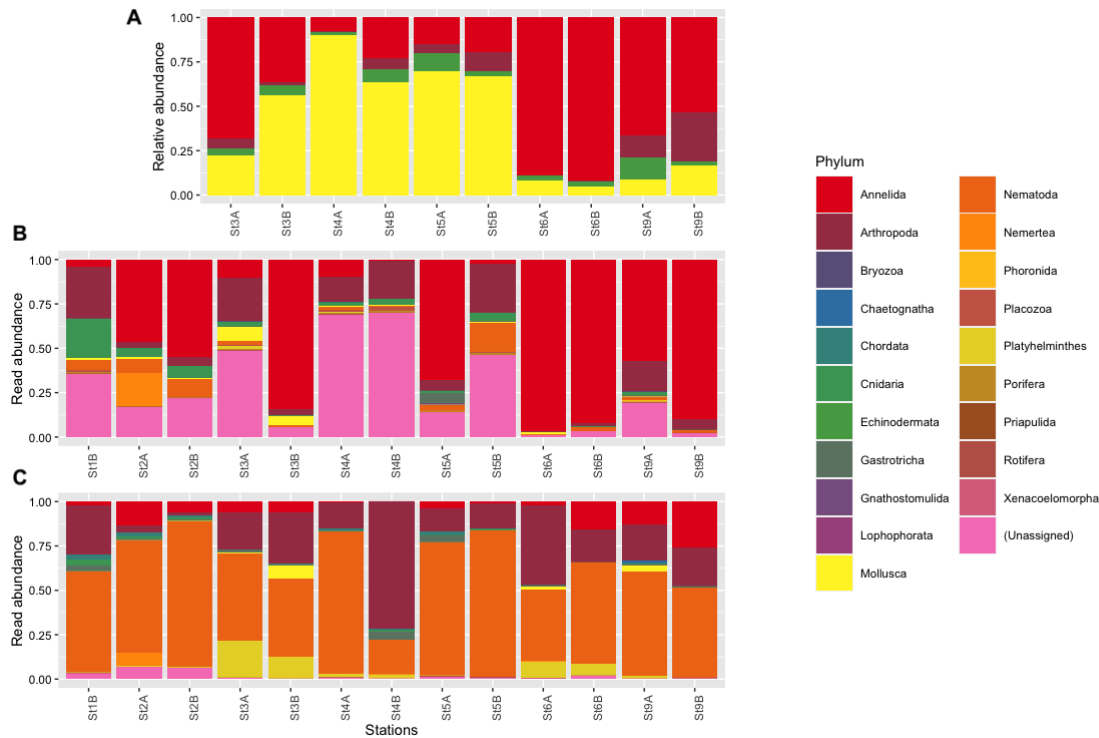


Figure 3.9: Taxonomic composition of Metazoa phylum level obtained by (A) morphology, (B) COI and (C) 18S. "Unassigned" refers to all reads not assigned to phylum level (i.e., reads only assigned to Metazoa).

The number of overlapping taxa at different taxonomic levels is shown in Figure 3.10. It clearly shows a decreasing overlap between the datasets from higher to lower taxonomic levels. An overlap of 6 phyla was detected (including Nematoda and Nemertea from the morphological dataset) (Figure 3.10). However, the relative abundance of the overlapping phyla did vary between the datasets. Annelida has the highest abundance in both the COI (79%) and the morphological (64%) dataset, while Nematoda has the highest abundance in the 18S metazoan dataset (55%). The pattern of relative abundance of Annelida per station is congruent across the COI and morphological datasets, particularly at stations 6 and 9 (Figure 3.9). Nematoda was not quantified using the morphological method and only made up 2% of COI data (25,550 reads).

For Annelida, despite being the most abundant phylum in the morphological and COI datasets, the richness at the family level differs, with 13 families detected by COI and 19 families detected by morphology. However, despite making up much less of the 18S dataset (8%), the richness at the family level nearly matches that of COI, with 11 annelid families being detected by 18S. The most abundant polychaete in the morphological results was *Protodorvillea kefersteini*, which COI did not detect. The species was detected to genus level by 18S at stations 6A and 6B, where abundance was highest also in the morphological method. Despite not detecting this highly abundant species, congruence was high between COI and morphology for other abundant annelids. For instance, *Pholoe* sp., *Cirriformia tentaculata* and *Macrochaeta*

*clavicornis* were amongst the most common species in the morphological dataset (see section 5.6), and *Pholoe baltica*, *Cirriiformia* sp. and *Macrochaeta clavicornis* were amongst the ten most abundant OTUs in the COI dataset (see section 5.7). These findings show that COI and morphological results are more congruent when looking at abundance, rather than only presence/absence of taxa. Furthermore, by comparing the relative abundance of Annelida across stations from Figure 3.9, similar patterns of annelid dominance at station 6 and 9 were observed for both COI and morphology.

Noteworthy differences were observed in the phylum Mollusca, which made up 28% of the morphological dataset covering 15 families. Meanwhile, Mollusca only constituted 1% of the COI data covering two families and 2% of the 18S data covering five families. Only two of these families were not detected by morphological investigations (Plakobranchidae and Mopaliidae). The most abundant mollusk, *Kurtiella bidentata*, was only detected by COI, with the highest abundance at stations 3A and 3B (see section 5.7), compared to the highest abundance at 3B and 4A in the morphological results (see section 5.6). 18S detected the family Montacutidae (now Lasaeidae) with highest abundance at stations 3A and 3B (see section 5.8). For the phylum Echinodermata, the relative abundance was low in each dataset, making up 4% of the morphological, 0.05% of the COI data, and 0.1 % of 18S data. The richness at the family level was low, with only four families detected by morphology, three by COI, and two families detected by 18S. The only family not detected by morphology was Schizasteridae, belonging to class Echinoidea. More obvious discrepancies were observed for the phylum Arthropoda, which made up 4% of the morphological dataset, 6% of COI, and 25 % of the 18S dataset. The Malacostraca species identified by morphology were not detected by 18S or COI, but several unique species of mainly pelagic classes Hexanauplia (COI), Maxillopoda, and Ostracoda (18S) were detected in the metabarcoding datasets.

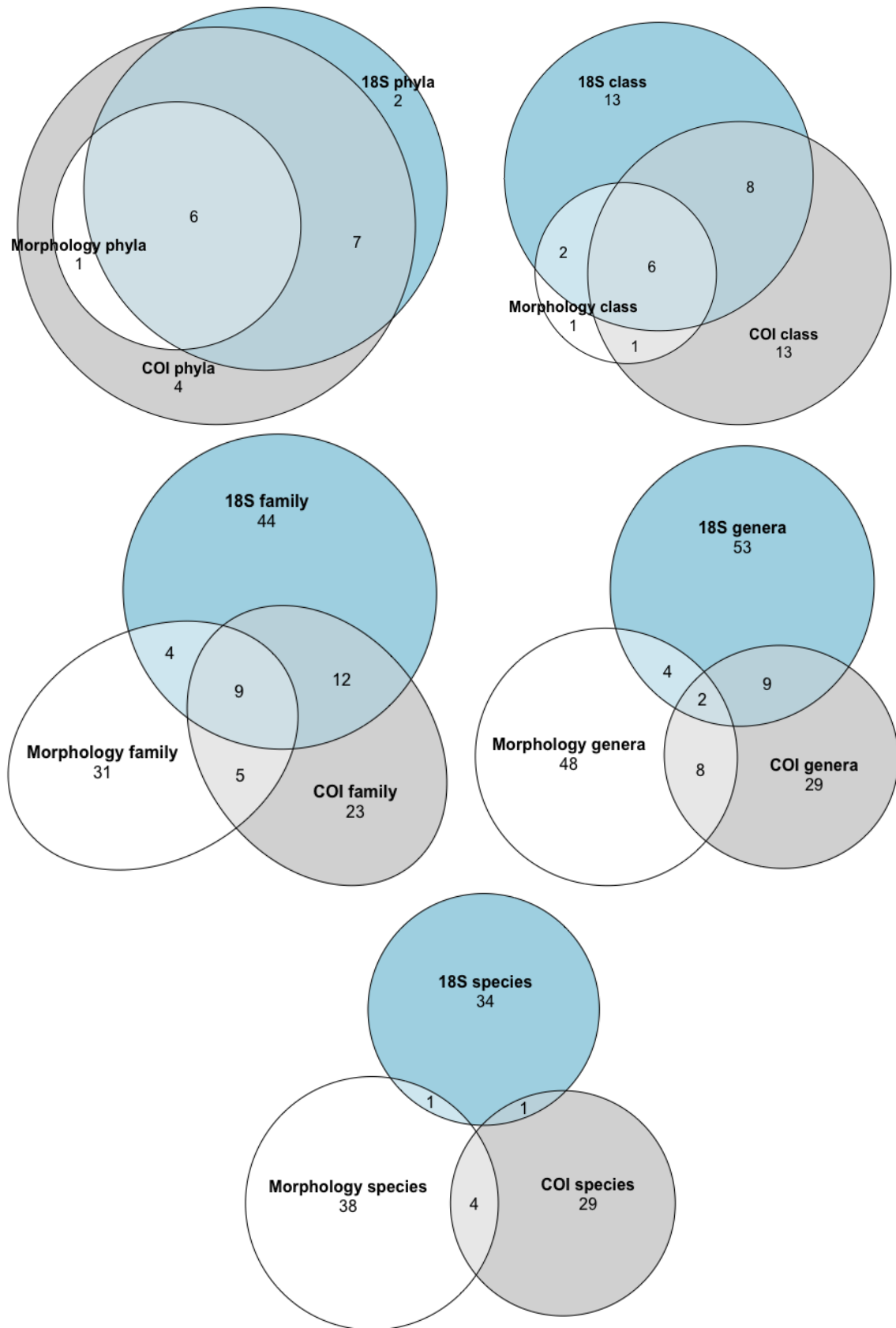


Figure 3.10: Euler diagram showing the number of shared and unique metazoan phyla, classes, families, genera and species between morphological data and molecular COI and 18S data.

Furthermore, both COI and 18S detected an additional seven phyla, including Chordata, Cnidaria, Gastrotricha, Platyhelminthes, Porifera, Rotifera, and Xenacoelomorpha. In addition to these phyla, COI detected Bryozoa, Placozoa, Gnatostomulida, Priapulida, and 18S detected the phyla Chaetognatha and Lophophorata (group representing Phoronida, Brachiopoda and Ectoprocta). In both COI and 18S datasets, these phyla were present in low abundance and each phylum made up less than 1% of the dataset. The exception is the phylum Platyhelminthes, which constituted 6% of the 18S metazoan dataset. Several of these are pelagic, for example, the chordate *Mugilogobius platynotus* detected by 18S. The number of unique and shared taxa illustrated in Figure 3.10 shows a high number of unique taxa at each taxonomic level. However, a closer look at the unique taxa from metabarcoding datasets reveals that many unique families belong to the low-abundance phyla described previously, thus constituting small proportions of the total datasets.

The most evident discrepancies were observed for stations 1 and 2, where no animals were detected in the morphological survey, and many OTUs were obtained by molecular analysis. For COI, the most abundant OTUs for station 1 were an unclassified metazoan (2202 reads) and the hydrozoans *Melicertum octocostatum* (1381 reads) and *Leuckartiara octona* (1260 reads) (See section 5.7). For 18S, the most abundant OTUs at station 1 were the ostracod *Cytheromorpha acupunctata* (7972 reads), nematodes from order Chromadorida, and the copepod *Centropages hamatus* (3350 reads) (See section 5.8). At station 2, eelgrass was observed during sampling, but no animals were found in the morphological investigation. In the COI dataset, *Tubificoides benedii* was the most abundant OTU for both 2A and 2B (See section 5.7). In the 18S dataset, the most abundant OTUs were by nematodes from genus *Triploides* (2A), order Enoplida (2B) and the species *Molgolaimus demani* (2A and 2B) (see section 5.8).

### 3.5 Alpha diversity patterns

The Shannon diversity index ( $H' \log_2$ ) was calculated for each dataset and is shown in Table 3.5. The calculated values from the morphological data were compared to class values from Table 2.4 and marked with color indicating state values. As state values are based on morphological data, molecular datasets have not been calibrated to a state value, and comparison between datasets is not valid. The grab replicates from morphological data show highly variable values covering all state values from “very poor” at 4A to “excellent” at 3A and 9A. Only stations 5 and 6 have the same state values for both grab replicates. Station 9, the reference station, shows high values corresponding to “excellent” for 9A and “good” for 9B based on morphological data. Diversity values calculated for COI data varied greatly, ranging from 1.05 (9B) to 6.27 (1B). The difference between grab replicates from the same station is evident, particularly for stations 3, 5, and 9. The diversity values calculated from the 18S metazoan dataset show more consistency between grabs, ranging from 2.44 (4A) to 4.66 (3A). The values cal-

culated from the 18S eukaryotic dataset show high overall values, with grab 2B showing the highest value of 7.99 and 6B having the lowest value of 5.81 (Table 3.5).

×

Table 3.5: Shannon diversity index values calculated from morphological, COI, 18S (Metazoa) and 18S (Eukaryota) data with colors indicating state value as defined by class values. Blue = "Excellent", green = "Good", yellow = "Moderate", orange = "Poor" and red = "Very poor". 9A and 9B refers to grab samples for morphology, while 9C and 9D refers to grab samples for metabarcoding.

Grab no.	Shannon diversity index ( $H'(\log_2)$ )			
	Morphological	COI	18S (Met)	18S (Euk)
1B	N/A	6.27	4.42	7.24
2A	N/A	3.99	4.46	7.93
2B	N/A	3.43	3.81	7.99
3A	4.19	5.67	4.66	7.17
3B	3.11	2.20	4.13	7.61
4A	0.89	3.75	2.44	6.70
4B	2.89	3.81	3.51	7.39
5A	1.82	2.65	3.38	7.71
5B	1.93	5.67	3.44	7.81
6A	2.84	1.87	4.06	7.11
6B	2.52	2.38	3.59	5.81
9A/9C	4.40	4.81	4.55	5.91
9B/9D	3.52	1.05	3.41	7.39

### 3.6 Beta-diversity patterns

Bray-Curtis similarity (0-100) was calculated for the morphological dataset, COI (Metazoa), 18S (Metazoa), and 18S (Eukaryota) using Primer7. All datasets were fourth root transformed prior to calculation, and Bray-Curtis distance matrices for each dataset is attached in Appendix section 5.1.

In the morphological data, the similarity was highest between grab samples 5B and 4A (68.7), second highest between 3A and 3B (64.6), and third highest between 6A and 6B (63.5). The similarity was lowest between 9A and 5A (12.7) (Section 5.1). An nMDS plot based on Bray-Curtis similarity display similarity as points representing grab samples in the ordination. Grabs are distinctly separated and well represented in the nMDS plot, as indicated by a low-stress value (stress = 0.04), except for 5B and 4A (Figure 3.11). Grabs belonging to stations 3 and 6 are clustered closely together in the nMDS plot. Stations 3 and 6 are located closely together in a more current-exposed area of the fjord and have similar sediment of shell sand and larger

grains. 9A and 9B have low similarity (38,7), considering that the grabs belong to the same station.



Figure 3.11: Non-metric multidimensional scaling (nMDS) based on Bray-Curtis similarity (0-100) from the morphological dataset. Stress value = 0.04.

For COI, Bray-Curtis similarities were highest between grabs 2A and 2B (68.9), second highest between 4A and 4B (57.1), and third highest between station 5A and 5B (55.8) (Section 5.2). The lowest similarity was between 9C and 2A (19.7), but similar to the morphological results, the similarity was also low between 9C and 5A (20.5) as illustrated in the nMDS plot for COI (Figure 3.12). Further, the grab samples are easily separated as indicated by a low-stress value (stress value = 0.13). Most plots are located on the right-hand side of the ordination, with a plot representing grab 9C on the far left. Grabs belonging to the same station have relatively high similarity, as indicated by the three highest similarity values previously mentioned. In contrast, grabs from station 9 (9C and 9D) are far apart in the ordination (similarity of 27.5) (Figure 3.12). This is similar to morphological results, showing more distance between grab replicates from station 9 compared to grab replicates from other stations. Similar to the morphological results, grabs belonging to stations 3 and 6 are clustered relatively closely together. Similarities between these grabs range from 41.1-47.7.

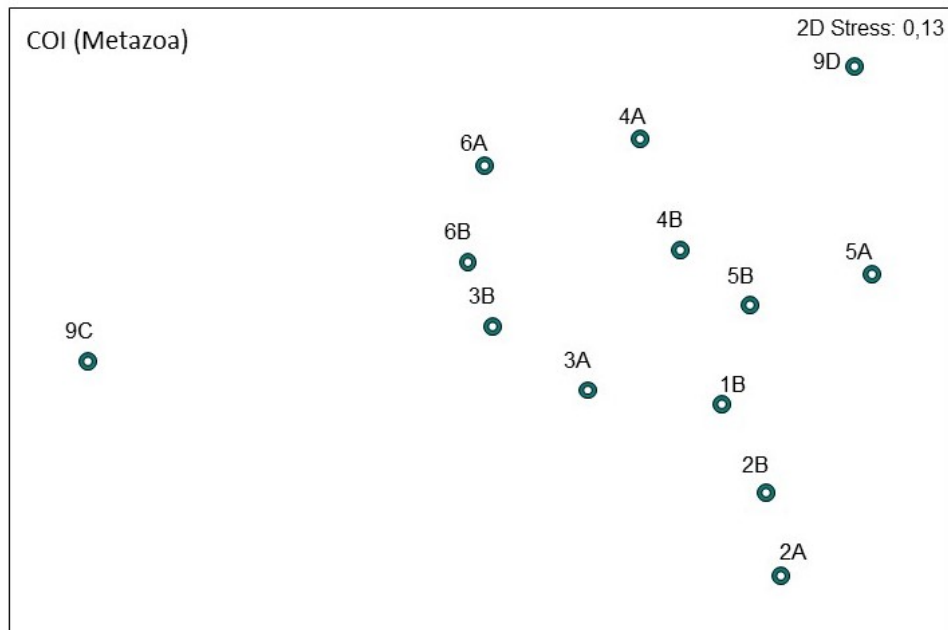


Figure 3.12: Non-metric multidimensional scaling (nMDS) based on Bray-Curtis similarity (0-100) from the COI dataset. Stress value = 0.13.

Bray-Curtis similarity (0-100) for the 18S metazoan dataset shows the highest similarity between grabs 2A and 2B (63.9) and the second highest similarity between 4A and 4B (58.3), similar to COI results. The third highest similarity was between 6A and 6B (55.3). The lowest similarity was between 9C and 5A (17.9) (Section 5.3), consistent with results for morphological data. In the nMDS plot based on Bray-Curtis similarity for the 18S metazoa (Figure 3.13), the grab samples are easily separated, as indicated by a low-stress value (stress value = 0.09). Grabs 2A and 2B are far from the other plots in the ordination, with similarities below 35 relative to all other grabs. Similar to COI results, grab 5A and 9C have low similarity. Similarities between stations 3 and 6 are relatively high, ranging from 39.9 to 46.8.



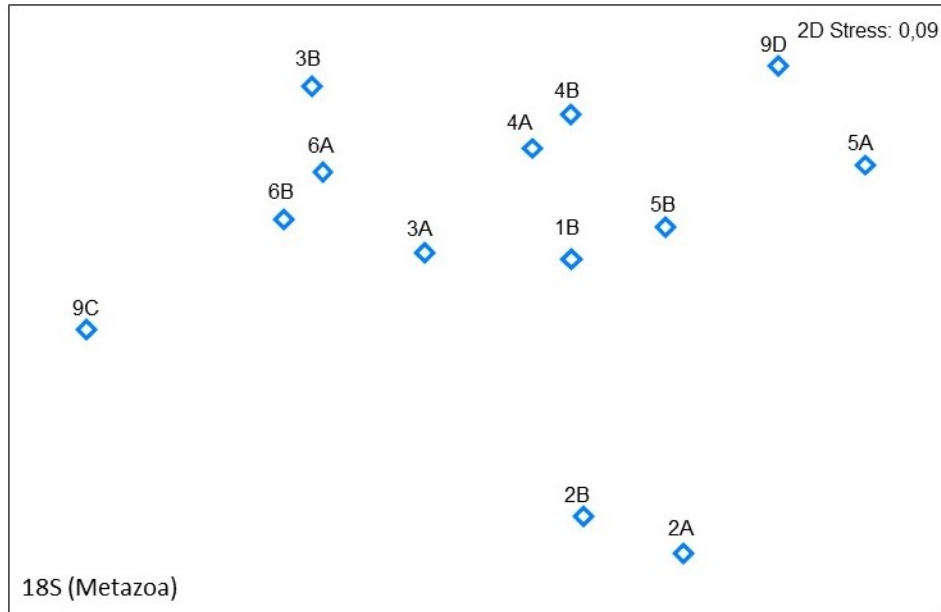


Figure 3.13: Non-metric multidimensional scaling (nMDS) based on Bray-Curtis similarity (0-100) from the 18S Metazoan dataset. Stress value = 0.09.

Bray-Curtis similarity (0-100) for the entire 18S Eukaryota dataset shows the highest similarity between grabs 4A and 4B (72.9), second highest for 6A and 6B (71.0), and third highest between 2A and 2B (70.9) (Section 5.4). These results are similar to 18S Metazoa. The lowest similarities were between 9C and 2A (38.8), similar to COI results. Similarities were also low between 9C and 2B (42.4) and 9C and 5A (43.7). These results are similar to the COI and 18S metazoa data, where similarities are lowest between station 2 and grab 5A. Stations 9 and 2 are the stations furthest apart, as shown in the map of the poll (Figure 2.1). Community structure as indicated by Bray-Curtis similarity from the 18S Eukaryota dataset is illustrated in the nMDS plot in Figure 3.14 with low stress value (stress value = 0.08). Similar to the nMDS plots for morphology, COI, and 18S Metazoa, 9A/9C is in the far right of the ordination, relatively far apart from 9B/9D.

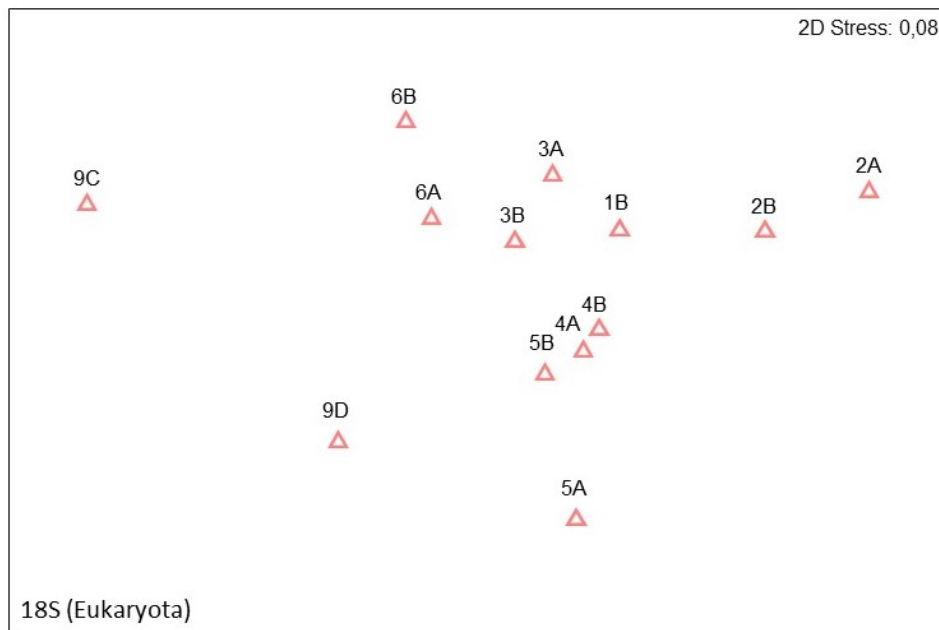


Figure 3.14: Non-metric multidimensional scaling (nMDS) based on Bray-Curtis similarity (0-100) from the 18S Eukaryota dataset. Stress value = 0.08.

The four datasets describe similar community structures, and similar patterns are observed across nMDS plots and distance matrices. Across all nMDS plots, station 9 (9A/9C) stands out as being distanced from other points in the ordination. Grab replicates 9A/9C is least similar to 5A in both the morphological and 18S metazoan datasets and 2A in the COI and 18S eukaryotic datasets. For COI and eukaryotic 18S, the similarity between 9C and 5A is also low. Another pattern observed between datasets is that grab replicates from the same station have the highest similarity. For both 18S datasets, grab replicates from stations 2, 4, and 6 are the grabs with the highest similarities (different order), and grab replicates from stations 2 and 4 are also the most similar in the COI dataset. In contrast to these patterns, grab replicates from station 9 are placed far from each other in all nMDS plots to varying extents. Grab 1B has the highest similarity to grab 3A across all metabarcoding datasets. A Mantel test was performed to test the correlation between the Bray-Curtis distance matrices of morphological and metabarcoding datasets (Table 3.6). Comparing Bray-Curtis matrices from the molecular datasets showed that the correlation is strongest between the 18S eukaryotic and COI metazoan matrices (Spearman's  $\rho=0.875$ ,  $p < 0.0001$ ). However, the correlation was strong between all molecular datasets. For the morphological dataset, correlation was strongest to the 18S metazoan dataset (Spearman's  $\rho=0.69$ ,  $p < 0.001$ ) (Table 3.6).

×

Table 3.6: The results from mantel testing the similarity matrices from all four datasets based on Spearman rank correlation using RELATE function in Primer7. The primary and secondary data shows the similarity matrices that were tested. Rho is the sample statistic ( $\rho$ ), and the significance level is given in percentage. The number of permutations was 9999. The number of permuted statistics greater or equal to Rho is as indicated.

Primary data	Secondary data	Sample statistic ( $\rho$ )	Significance level of sample statistic	Number of permutations	Number of permuted statistics greater or equal to Rho
18S (Euk)	18S (Met)	0.834	0.1 %	9999	0
18S (Euk)	COI	0.875	0.01 %	9999	0
18S (Met)	COI	0.821	0.01 %	9999	0
Morphological	COI	0.637	0.05 %	9999	4
Morphological	18S (Met)	0.69	0.04 %	9999	3
Morphological	18S (Euk)	0.649	0.03 %	9999	2

### 3.7 Biotic index calculation

For biomonitoring purposes, biotic indices are calculated to assess the ecological status of marine water bodies. NSI, ISI, NQI1 and AMBI values inferred from the morphological, COI and 18S (Metazoa) datasets is shown in Figure 3.15. NSI values from the molecular datasets are much higher compared to the morphological dataset. The NSI values for different grabs show a similar pattern for COI and 18S, although COI values are higher than 18S (Figure 3.15). Most stations/grabs from the morphological dataset are classified as “moderate” and grabs 3A, 9A, and 9B were classified as “good” (Table 3.7). Based on COI data, most of the grabs are classified as “excellent” (8), while three are classified as “good”, and two are classified as “moderate” (Figure 3.8). NSI values for 18S show a overall higher calculation compared to morphology, with five grabs are classified as “excellent”, six as “good”, and two as “moderate” (Figure 3.9).

As illustrated in Figure 3.15, ISI values are much higher based on COI data compared to 18S and morphology. A similar pattern is observed between the datasets, with higher values for grabs 3B to grab 5B. The morphological is consistent between the grabs, and the grabs are classified as “moderate”, “good” and “excellent” (Table 3.7). The high values from the COI data classified all grabs as “excellent” (Table 3.8). ISI values from 18S showed more variation (Table 3.9), but was more similar to morphology for grabs 5B-9B.

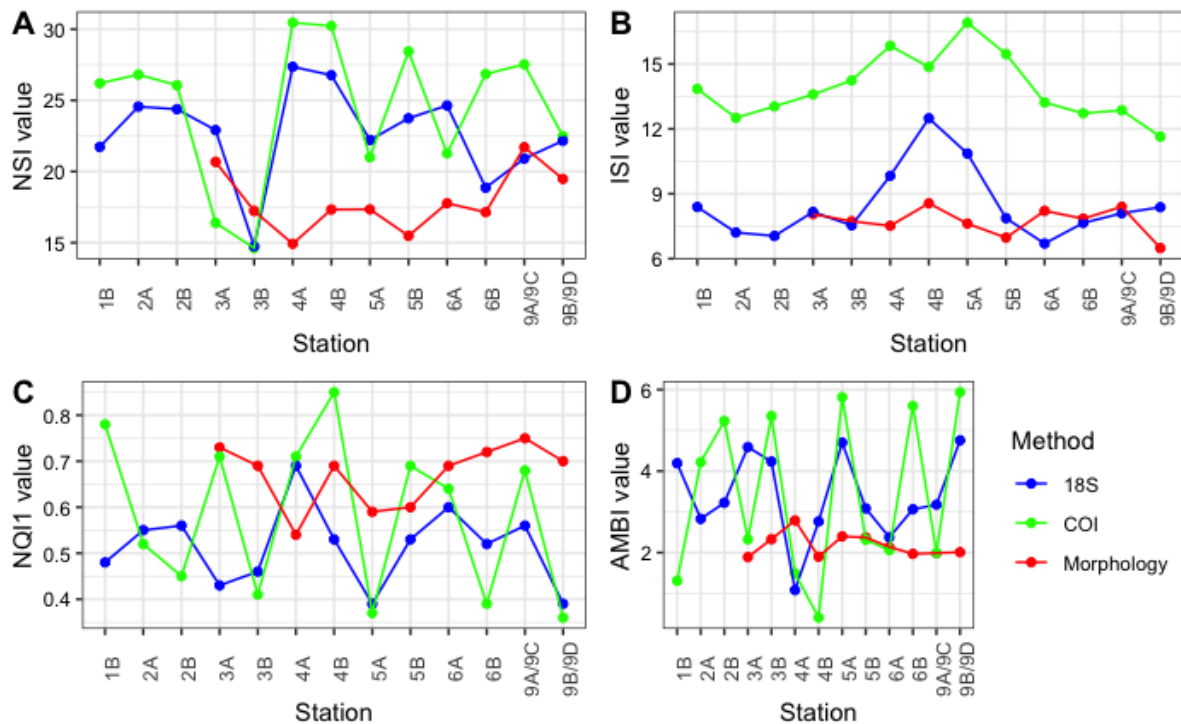


Figure 3.15: Biotic indices (A) NSI, (B) ISI, (C) NQI1 and (D) AMBI calculated from the COI, 18S (Metazoa) and morphological dataset. 9A and 9B refers to grab samples for morphology, while 9C and 9D refers to grab samples for metabarcoding.

NQI1 values were inconsistent between the datasets, with overall higher values based on the morphological data (Figure 3.15). Most grabs from the morphology are classified as “good”, two stations were classified as “moderate” (4A and 5B), and two were classified as “excellent” (3A and 9A) (Table 3.7). NQI1 values from COI show high variation between grabs, with classification ranging from “moderate” to “excellent” (Table 3.8). From 18S data, the NQI1 values show less variation between grabs compared to COI (Figure 3.15), and most grabs are classified as “moderate” or “poor”, with only one grab (4A) classified as “good” (Table 3.9).

AMBI values based on molecular data was overall higher compared to morphology, with greater variation between grab replicates for COI and 18S (Figure 3.15). From the morphological data, AMBI values range from 1.89 (3A) to 2.79 (4A), and all grabs are classified as “slightly polluted” (Table 3.7). For COI, AMBI values are highly variable and the classification of grabs range from “heavily polluted” to “unpolluted” (Table 3.8). The AMBI values from 18S is more consistent between grabs compared to COI, and most grabs are classified as either “meanly polluted” or “slightly polluted”, with station 4A being classified as “unpolluted” (Table 3.9).

×

Table 3.7: Biotic indices NSI, ISI, NQI1 and AMBI calculated from the morphological dataset with each value marked with colour indicating state value. Blue = "Excellent", green = "Good", yellow = "Moderate", orange = "Poor" and red = "Very poor". For AMBI, colours refer to Blue = "Unpolluted", green = "Slightly polluted", yellow = "Meanly polluted", orange = "Heavily polluted" and red = "Extremely polluted".

Grab no.	Morphological			
	NSI	ISI	NQI1	AMBI
1A	N/A	N/A	N/A	N/A
1B	N/A	N/A	N/A	N/A
2A	N/A	N/A	N/A	N/A
2B	N/A	N/A	N/A	N/A
3A	20.68	8.06	0.73	1.89
3B	17.23	7.73	0.69	2.33
4A	14.92	7.52	0.54	2.79
4B	17.33	8.56	0.69	1.90
5A	17.34	7.62	0.59	2.40
5B	15.48	6.97	0.60	2.37
6A	17.77	8.21	0.69	2.14
6B	17.14	7.85	0.72	1.97
9A	21.71	8.40	0.75	1.99
9B	19.48	6.49	0.70	2.01

×

Table 3.8: Biotic indices NSI, ISI, NQI1 and AMBI calculated from the COI dataset with each value marked with colour indicating state value. Blue = "Excellent", green = "Good", yellow = "Moderate", orange = "Poor" and red = "Very poor". For AMBI, colours refer to Blue = "Unpolluted", green = "Slightly polluted", yellow = "Meanly polluted", orange = "Heavily polluted" and red = "Extremely polluted".

Station no.	COI			
	NSI	ISI	NQI1	AMBI
1B	26,20	13,85	0,78	1,31
2A	26,81	12,51	0,52	4,22
2B	26,06	13,04	0,45	5,23
3A	16,39	13,59	0,71	2,33
3B	14,62	14,24	0,41	5,35
4A	30,46	15,84	0,71	1,49
4B	30,23	14,86	0,85	0,41
5A	21,00	16,91	0,37	5,81
5B	28,44	15,46	0,69	2,32
6A	21,28	13,22	0,64	2,06
6B	26,85	12,72	0,39	5,60
9C	27,53	12,86	0,68	1,99
9D	22,49	11,64	0,36	5,94

×

Table 3.9: Biotic indices NSI, ISI, NQI1 and AMBI calculated from the 18S metazoan dataset with each value marked with colour indicating state value. Blue = "Excellent", green = "Good", yellow = "Moderate", orange = "Poor" and red = "Very poor". For AMBI, colours refer to Blue = "Unpolluted", green = "Slightly polluted", yellow = "Meanly polluted", orange = "Heavily polluted" and red = "Extremely polluted".

Grab no.	18S			
	NSI	ISI	NQI1	AMBI
1B	21,73	8,39	0,48	4,20
2A	24,56	7,21	0,55	2,83
2B	24,38	7,05	0,56	3,22
3A	22,90	8,16	0,43	4,59
3B	14,75	7,54	0,46	4,23
4A	27,36	9,83	0,69	1,08
4B	26,77	12,49	0,53	2,77
5A	22,21	10,86	0,39	4,70
5B	23,74	7,87	0,53	3,08
6A	24,64	6,70	0,60	2,38
6B	18,86	7,65	0,52	3,06
9C	20,91	8,10	0,56	3,18
9D	22,16	8,38	0,39	4,75





# Chapter 4

## Discussion

This study has investigated the benthic invertebrate community in a poll using molecular and morphological methods to assess the performance of eDNA metabarcoding for diversity-based assessment of this habitat. For this objective, taxonomic composition, alpha and beta diversity, and biotic index values obtained by eDNA metabarcoding and the morphological taxonomy approach were compared.

### 4.1 Main findings from the morphology-based assessment

During fieldwork, we noted H<sub>2</sub>S smell and the sediment consistency and characteristics between the stations (Table 3.1). Stations 1 and 2 were located in the innermost part of the poll, which is highly affected by input of organic material from surrounding land (Dybern 1967). The sediment at these stations consisted of mud and organic matter with a distinct smell of H<sub>2</sub>S (Table 3.1). Station 1 was in the deepest basin of the inner part of the poll (7.9 m), where bottom water is more stagnant. Station 2 was located in an eelgrass meadow, and Lundberg (2015) previously recorded depth of this meadow at 2-3 meters. Stations 3 and 6 were located in the narrow and shallow area between the inner and outer poll, where the current is stronger (Figure 2.1) (Dybern 1967). This was reflected in the higher fraction of sand, shell and larger grain in the sediment at stations 3 and 6 (Table 3.1). Furthermore, the sediment was similar at station 9, the reference station, located outside the poll in a more typical marine environment. Stations 4 and 5 were located in more protected areas, with sediment consisting of mud and large rocks (Table 3.1). These findings are similar to the sediment characterization and hydrographic patterns previously reported by Dybern (1967). This suggests that the conditions in Kviturspollen has largely remained the same since 1962 (Dybern 1967), where input of organic matter still affects the local conditions, particularly in the deepest basins.

In previous investigations of Kviturspollen, Dybern (1967) recorded a homogeneous oxygen concentration along the water column at the shallower stations and a strong oxygen deficit

at the deeper stations (10-15 m) (Dybern 1967). These findings are similar to measurements from 2013 where Kvalø et al. (2014) report “very poor” oxygen conditions at station Kv1 (14 m depth) (Figure 2.1). In the present study, the CTD measurements show high oxygen concentration in the entire water column recorded through multiple measurements, which does not indicate anoxic conditions (Figure 3.2). According to Dybern (1967), the oxygen content is affected by seasonal variation in circulation patterns, and greater winter circulation alleviates total loss in the deepest basins (Dybern 1967). Moreover, significant input of organic material to the bottom sediment combined with limited water exchange will likely affect the oxygen conditions. At station Kv5 (Figure 2.1), the total organic content (TOC) was moderate in 2013, indicating sedimentation of organic matter at this station (Kvalø et al. 2014). Given the observation of organic material and smell of H<sub>2</sub>S in this study, it is reasonable to assume that certain stations may experience periodic anoxic conditions, despite normally high oxygen concentration in the water column. Salinity measurements show higher values in Mynteviken, located outside the poll. Inside the poll, the salinity decreases going from the outer to the inner part (Figure 3.1), with the inner part of Kviturspollen having the lowest salinity at the surface water. These findings indicate that the inner part of the poll experience a higher freshwater input, similar to results by Dybern (1967). Further, temperature differences between surface and bottom layers can contribute to stagnation of the bottom water, which a typical feature in polls (Dybern 1967). In this study, the temperature in the inner and outer parts of the poll was fairly stable with depth (Figure 3.3).

The morphological analysis of Kviturspollen yielded results of high taxonomic resolution resulting in the identification of 43 species from an assemblage of 2779 individuals. From the total of 93 taxa, 46% were assigned to species level, constituting 73% of the identified individuals. The findings made during the morphological identification are congruent with observations of sediment composition and smell made during fieldwork. We expected to find a higher number of individuals and species at the more current-exposed stations higher up and fewer species at the sheltered and deeper stations. We found no animals at stations 1 and 2, which is consistent with observations of organic matter and a strong H<sub>2</sub>S smell. However, at station 2, eelgrass constituted large parts of the grab sample. The eelgrass meadow at station 2 was assessed in 2014 and described as brown, overgrown, and in bad condition (Lundberg 2015). These findings are similar to our observations, indicating that the state of the eelgrass meadow has remained the same. For the remaining stations, the expected pattern was reflected in the results, with a higher number of taxa detected at the more current-exposed stations (3, 6, and 9) compared to more protected stations (4 and 5) (Table 3.2).

Compared to the taxon list from previous analyses of Kviturspollen, we made similar findings in this study, particularly for the phyla Polychaeta and Mollusca (Kvalø et al. 2014). However, the taxonomic precision between the current and previous surveys has a few differences. In the present study, *Pholoe* sp. was only identified to genus level, while previous studies identified

*Pholoe baltica* and *Pholoe inornata*. This comparison underlines the difference in results obtained by different taxonomists. High taxonomic resolution in large-scale projects solely based on the morphological method requires a lot of time and effort, which are resources monitoring programs generally do not have (Aylagas et al. 2018).

In summary, the findings from the morphological analysis of Kviturspollen were largely consistent with previous findings. Local variation based on topography, depth, and hydrographical patterns affects the degree of sedimentation of organic material, which was more evident in the inner part of the poll, particularly in the deepest basin (station 1). Here, organic input and limited water exchange may lead to anoxic conditions in the sediment. These differences are reflected in observations of sediment and smell, and community composition at the stations.

## 4.2 Comparison of taxonomic composition

Comparisons of taxonomic composition across morphological and molecular datasets in the present study illustrate that each method gives insight into different parts of the benthic community rather than replicating, which is similar to other studies (Keck et al. 2022). The metazoan taxa detected greatly varies with the method (morphological vs. eDNA metabarcoding) and marker choice (COI vs. 18S metazoa), as seen in Figure 3.9. The number of unique taxa was especially high in each dataset compared to the number of shared taxa (Figure 3.10), which suggests that each dataset provides complementary taxonomic findings. However, when looking at abundance congruence is higher: the phylum Annelida dominates the COI (79%) and the morphological (64%) datasets, and congruence is high for the most abundant annelid species (e.g., *Pholoe* sp.). Furthermore, Figure 3.9 shows that Annelida dominates stations 6 and 9 in both the COI and morphological dataset, while Annelida makes up a significantly smaller part of the 18S metazoan dataset. For 18S, the most abundant metazoan phylum was Nematoda. While dominating the 18S dataset, Nematoda was not quantified in the morphological survey and only made up 2% of the COI data. These findings suggest that congruence is higher between COI and morphology, whereas 18S detects more meiofaunal species.

Furthermore, the results show a decreasing overlap going from higher to lower taxonomic levels (Figure 3.10). While each method detected a significant fraction of biodiversity, the partial overlap at lower taxonomic levels also indicates the occurrence of false negatives (i.e., species present in the morphological dataset but not detected by metabarcoding). This is especially true for the polychaete *Protodorvillea kefersteini*, which is highly abundant in the morphological dataset but undetected by COI and only detected at the genus level by 18S. Mollusca is also largely undetected in molecular datasets compared to morphological results. An example is *Kurtiella bidentata*, the second most abundant species in the morphological dataset. This species was detected by COI but not by 18S. A closer look at higher taxonomic levels shows the family of *K. bidentata*, Montacutidae, was present in the 18S dataset. These findings

can be explained by gaps in the reference databases limiting the species-level identification of species such as *P. kefersteini* and *K. bidentata*. The accuracy of taxonomic assignment relies on the completeness of reference databases (Aylagas et al. 2016), where species-specific reference sequences are necessary for species-level identification of OTUs (Eckert et al. 2018). Consequently, this results in a significant overlap at higher taxonomic levels and incongruence at the species level, which is clearly illustrated in Figure 3.10.

### 4.2.1 The effect of primer bias on taxonomic composition

Primer bias is a highly acknowledged issue in metabarcoding (Deagle et al. 2014, Elbrecht & Leese 2015), resulting in a taxonomic selectivity among primers (Kelly et al. 2017), which can explain the difference in taxonomic composition in this study. Studies show that the detection of taxa is highly dependent on its amplification efficiency for a given set of primers, which is highly species-specific (Elbrecht & Leese 2015, Kelly et al. 2019). The results from this study show a strong taxonomic selectivity towards Annelida for COI and meiofaunal species for 18S. These findings are similar to Mauffrey et al. (2020), where Nematodes and Arthropods dominated the 18S V1-V2 dataset, and Annelids and Arthropods dominated the COI dataset. The taxonomic selectivity observed is likely a result of the affinity between template DNA and primers (Kelly et al. 2017). These interactions play a significant role in detecting species and the consequent occurrence of false negatives (Deagle et al. 2014, Elbrecht & Leese 2015, Kelly et al. 2017, Piñol et al. 2015). The dominance of Nematoda in the 18S dataset is not surprising, considering that the 18S primers used here were originally developed for Nematoda, specifically (Blaxter et al. 1998, Lanzén et al. 2021, Sinniger et al. 2016).

Moreover, in this study, Mollusca is largely overlooked by both COI and 18S. This undersequencing of Mollusca is similar to findings by Cahill et al. (2018) using Leray COI primers, and Lejzerowicz et al. (2015) using the 18S V4 region. According to Fernández et al. (2019), the COI primer pair mlCOIintF (Leray et al. 2013) and jgHCO2198 (Geller et al. 2013) are known for poor performance for mollusks. However, the COI primer pair used in this study was modified to improve the universality of the primers (Wangensteen et al. 2018). A potential effect of increased universality is a higher probability of non-target amplification (Leray et al. 2013), which we observed for COI in this study. Non-target amplification can reduce the primer's sensitivity for target taxa (Hering et al. 2018), which presents additional challenges for detecting target groups such as mollusks. In addition, limited DNA release to the surroundings due to the protective hard shell can reduce the detectability of Mollusca (Ji et al. 2022). The combination of potentially low concentration of mollusk DNA and reduced sensitivity may be important factors for the detection of this group. Nonetheless, the poor amplification of mollusks supports the general assertion that truly universal primers are lacking and are difficult to design (Deagle et al. 2014, Hestetun et al. 2020, Leray et al. 2013). The two genetic markers, COI and 18S, detect different taxonomic groups, and the high number of unique taxa

at each taxonomic level highlights the importance of a multigene approach (Günther et al. 2018, Kelly et al. 2017).

### 4.2.2 COI and 18S can detect taxa undetected by morphology

In contrast to the occurrence of false negatives in the molecular datasets, COI and 18S also detected several taxa that were undetected by morphology. These were, for example, oligochaetes *Tubificoides benedii* (AMBI ecological group: V) detected by COI and *Tubifex tubifex* detected by 18S, both present in high abundance (See section 5.7 and 5.8). We did find oligochaetes during the morphological assessment, but these were only identified as “Oligochaeta indet.” (NSI EG: V). Other species undetected by morphology were *Nereimyra punctata* (NSI EG: IV) by COI and *Scoloplos armiger* (NSI EG: III) by 18S (Direktoratsgruppen vanndirektivet 2018a). Furthermore, several meiofaunal species are found in the eDNA metabarcoding datasets. This finding is not surprising, considering their size and restricted mobility (Lejzerowicz et al. 2015, Pawlowski et al. 2022), and that the V1-V2 region of 18S is commonly used in studies for marine meiofauna (Blaxter et al. 1998, Mauffrey et al. 2020, Sinniger et al. 2016). Routine monitoring surveys traditionally use benthic macroinvertebrates (>1mm) as a biological quality element due to their ability to reflect environmental conditions (Direktoratsgruppen vanndirektivet 2018b), whereas meiofaunal species are not included. This is due to the challenge of identifying meiofaunal species based on morphology (Direktoratsgruppen vanndirektivet 2018b, Rzeznik-Orignac et al. 2017). The detection of inconspicuous meiofaunal species by eDNA metabarcoding may represent an opportunity for ecological quality assessment, which is not possible by the morphological approach (Mauffrey et al. 2020, Moens et al. 2022, Rzeznik-Orignac et al. 2017). Studies show that meiofaunal species such as nematodes can be a promising alternative to macrofauna in biomonitoring studies (Franzo et al. 2022, Rzeznik-Orignac et al. 2017), which will be discussed further below (see section 4.7).

In addition to meiofauna, COI and 18S detect taxa with a possible pelagic origin, such as taxa from the phyla Cnidaria, Chordata, and Class Copepoda. In biomonitoring, benthic invertebrates are the target group. Thus, pelagic species may qualify as a non-target group when the purpose of the study is benthic monitoring, but the degree to which pelagic species are filtered varies between studies. In Lanzén et al. (2021), OTUs of pelagic origin were removed from the dataset since the target group was benthic invertebrates. In contrast, in a study on benthic monitoring, Lejzerowicz et al. (2015) does not address the potential presence of taxa with a pelagic origin, and also detected the phyla Hydrozoa and Copepoda possibly representing pelagic taxa. Such discrepancies are likely due to the fact that correct recognition of pelagic OTUs is challenging: A phylum may contain pelagic and benthic species, such as

Chordata with Actinopterygii and Ascidiacea, represented in the 18S dataset. Copepoda can also occur in both the pelagic and benthos (e.g., Harpacticoida). For these reasons, Lanzén et al. (2021) acknowledges the possibility of having removed benthic species along with the pelagic ones. First, these different approaches make comparisons across studies complex (Leese et al. 2016) and underline the need for more consensus about the protocol. Moreover, the presence of pelagic DNA signals may affect diversity estimates, particularly for sites with low benthic diversity, such as stations 1 and 2 in this study (See section 4.4.1).

### 4.3 Performance of COI and 18S markers

In standard monitoring, ecological status assessment is based on the composition of benthic invertebrates (Direktoratsgruppen vanndirektivet 2018b). The ability and performance of the markers for describing the metazoan community is, therefore, highly relevant. COI is used as the standard genetic marker targeting animals, while 18S is a universal eukaryotic marker (Andújar et al. 2018, Hebert et al. 2003, Zhan et al. 2014). These features are reflected in the results with COI having a higher number of reads (28.6%) and OTUs (753) assigned to Metazoa than 18S (13.4% and 298 metazoan OTUs). Although COI detected more metazoan OTUs than 18S, challenges remain with COI for targeting specific groups in environmental samples (Collins et al. 2019, Deagle et al. 2014), which is related to non-target amplification. For example, COI amplified a substantial amount of non-target bacteria (24.9% of the total number of reads), which is also documented in other studies using the 313 bp COI primers (Aylagas et al. 2016, Collins et al. 2019, Yang et al. 2013). The amplification of bacteria is not surprising considering the origin of the mitochondrion from an  $\alpha$ -proteobacteria (Gray 2015), and that COI is a mitochondrial gene. Further, the concentration of prokaryotic and eukaryotic microbial DNA in the sediment eDNA pool is higher than metazoan DNA (Andújar et al. 2018, Stat et al. 2017), which is also illustrated here by the high abundance of the protist group SAR in the full 18S dataset (74% of the reads).

The consequence of non-target amplification, particularly bacteria for COI, is reduced primer sensitivity and reproducibility for target taxa (Collins et al. 2019, Hering et al. 2018). Thus, for monitoring benthic invertebrates specifically, bulk sample metabarcoding is a more suitable option than eDNA from sediment samples (Alberdi et al. 2018, Collins et al. 2019). This is due to the mixture of DNA in eDNA samples, while bulk samples represent mainly target taxa (Macher et al. 2018, Pawlowski et al. 2018). However, processing and separation of the specimens from large sediment samples can be labor-intensive, which makes bulk-sample metabarcoding less suitable for large-scale monitoring (Pawlowski et al. 2022). Alternatively, access to microbial diversity through sediment eDNA presents opportunities for using new bioindicators (Pawlowski et al. 2022). Studies have illustrated the potential of bacteria (Aylagas et al. 2017) and protists (Lanzén et al. 2021, Pawlowski et al. 2014) as indicators of

impacts. Still, the current monitoring programs have not yet included these groups in ecological quality assessment (Pawlowski et al. 2022).

Moreover, the number of "no hits" OTUs was far greater in the COI dataset with 23%, compared to 0.17% in the entire eukaryotic 18S dataset (Figure 3.9). In addition, a higher number of metazoan OTUs were unassigned to phylum level for COI compared to 18S (i.e., only assigned to "Metazoa") (Figure 3.9). These differences between COI and 18S are similar to results obtained by Mauffrey et al. (2020) and Hestetun, Lanzén & Dahlgren (2021). For the COI gene, faster evolution (i.e., higher nucleotide substitution rate) can lead to lower genetic distance between distantly related groups (homoplasy) (Leray & Knowlton 2016, Rach et al. 2017, Yu et al. 2012). Consequently, genetic differentiation of OTUs at higher taxonomic levels is reduced, particularly in the absence of reference sequences at lower taxonomic levels (Collins et al. 2019, Leray & Knowlton 2016, Rach et al. 2017). As a result, COI tends to saturate at higher taxonomic levels (Hestetun et al. 2020, Leray & Knowlton 2016), as illustrated by the high number of OTUs only assigned "Metazoan" by COI in this study. Further, high substitution rate and high variability can limit the use of COI for phylogenetic assignment (Leray & Knowlton 2015), leading to unidentified OTUs observed as "no hits" in this study (Leray & Knowlton 2015, 2016). Due to slower evolution in the 18S gene (Hillis & Dixon 1991), this marker performs better at assigning sequences to higher taxonomic levels (i.e., order and class) (Tang et al. 2012, Wu et al. 2015). However, a consequence of slower evolution is low species-level resolution, which is previously documented for 18S (Tang et al. 2012).

### 4.3.1 Taxonomic resolution of 18S and COI

In this study, we found a higher taxonomic resolution for 18S (Metazoa) compared to COI, which is unexpected compared to other studies using COI and 18S (Clarke et al. 2017, Duarte et al. 2021, Lanzén et al. 2021, Zhan et al. 2014): In the 18S metazoan dataset, 33% of the reads (12.4% of OTUs) were assigned to species level. For COI, 40.7% of the reads (but only 4.5% of the OTUs) were assigned to species level. At higher taxonomic levels, such as genus, family, and order, 18S outperforms COI in taxonomic resolution. At the genus level, the percentage of reads was 5.4% for COI and 17.2% for 18S. In contrast, Lanzén et al. (2021) used the same primers as this study, but 52% of the reads were assigned to genus level for COI, compared to only 5% for 18S. When looking at read abundance, COI had a higher number of reads assigned to species level; however, 72% of the reads belonged to *Tubificoides benedii*. Elbrecht & Leese (2015) demonstrate that the number of sequences obtained is highly species-specific, indicating high amplification efficiency for *Tubificoides benedii* using COI primers in this study. Other abundant polychaetes in the COI dataset, such as *Pholoe baltica* and *Macrochaeta clavicornis*, further illustrate successful species-level assignment for the most abundant groups using COI. For 18S, 33% of the reads to species level represents the most abundant OTUs. These are, for example, the ostracod *Cythermorpha acupunctata*, the

nematode *Spirinia parasitifera*, and a nematode from the genus *Bathylaimus*, the two latter illustrating the higher affinity towards meiofauna for the 18S primers.

A prerequisite for sufficient taxonomic resolution is adequate reference sequence coverage (Alberdi et al. 2018, Bourlat et al. 2013), which has been lower for 18S compared to COI on a global scale (Collins et al. 2019, Hestetun et al. 2020). Several studies have therefore relied on an *ad hoc* construction of 18S reference library (Collins et al. 2019, Mohrbeck et al. 2015). However, a recent study on DNA database coverage of the most common species in the North Sea found a relatively high coverage for 18S compared to the global coverage (Hestetun et al. 2020). These findings indicate that the relatively higher barcode coverage for common species of the Norwegian coast may explain the higher resolution for 18S in this study. However, as underlined by the mollusk *Kurtiella bidentata* only identified to family Montacutidae by 18S, gaps in the reference library still affect the results.

Furthermore, with more conserved markers such as 18S (Hillis & Dixon 1991, Wu et al. 2015), reduced genetic differentiation of closely related species may cause misidentification of OTUs (Tang et al. 2012, Wangensteen et al. 2018). An example is the fish *Mugilogobius platynotus*, which 18S detected at stations in the poll and the reference station (station 9). This species is endemic to Australia (Hammer 2006) and does not occur in the study area. A likely explanation is that the detected fish does not have a reference sequence in the database and is linked to the most closely related species with an 18S reference sequence in the database (Wangensteen et al. 2018). The “presence” of *Mugilogobius platynotus* underlines the potential incidence of false positives due to misidentification during the taxonomic assignment, similar to a finding by Wangensteen et al. (2018) using the 18S marker. Taxonomy is assigned using CREST, which classifies OTUs based on alignment and the lowest common ancestor method (Lanzén et al. 2012). The detected species is more likely to be the Norway goby *Pomatoschistus norvegicus* (Forsgren & Næss 2019), which does not have a reference sequence in the Silva database (Quast et al. 2013). Thus, while 18S has previously not been considered informative at the species level due to less genetic variability (Tang et al. 2012), this study demonstrates that the resolving power of 18S may have been underestimated due to low database coverage and that greater local database coverage enables higher taxonomic resolution for 18S.

#### 4.4 DNA signals from heavily impacted sites

The most evident discrepancies between morphological and molecular results were found at stations 1 and 2. During sampling in the field, we noted a strong smell of H<sub>2</sub>S at these locations. The sediment was high in organic matter and mud, and we observed no animals during sieving or when investigating a subsample of the sediment. We therefore suspected that these stations were in poor condition given that H<sub>2</sub>S is toxic to metazoan life and linked to anoxic sediments (Diaz & Rosenberg 1995). In contrast to these observations, a substantial amount



of DNA signals were retrieved from stations 1 and 2 (Figure 3.7 and 3.8.) At station 1B in particular, a high number of species were detected by metabarcoding compared to other stations. These conflicting findings call into question the ability of eDNA to accurately assess biodiversity and ecological status in a complex poll-like habitat where pockets of anoxic sediments are common (Dybern 1967, Wassmann 1985). A possible explanation for these findings is the presence of extracellular DNA originating from other areas of the poll or DNA from past assemblages, thus not accurately describing the local diversity at these stations.

#### **4.4.1 Challenges of using extracellular DNA in sediment samples**

Environmental DNA metabarcoding from sediment samples is primarily based on extracellular DNA (or extDNA) since most of the DNA in sediments is extracellular (Aylagas et al. 2016, Corinaldesi et al. 2011, Lejzerowicz et al. 2015, Pawlowski et al. 2022). Further support that eDNA mainly consists of extDNA was provided by Pansu et al. (2021), finding no significant difference in richness estimates between total DNA (intra- and extracellular) and extracellular DNA. However, when dealing with extracellular DNA, it is difficult to differentiate between DNA signals from living and dead organisms and local and allochthonous DNA (Collins et al. 2018, Pawlowski et al. 2022). According to Pansu et al. (2021), extracellular DNA will reflect the active community by integrating extDNA continuously released through biomass turnover, thus making exDNA appropriate for monitoring. However, studies have indicated that extracellular DNA may also represent past assemblages (Corinaldesi et al. 2011, 2018), which result in false positive detection and an inaccurate representation of the present community (Collins et al. 2018). Thus, recognizing that total eDNA consists of both intra- and extracellular DNA will be important when interpreting results and drawing conclusions based on the molecular findings (Nagler et al. 2022). Additionally, information on eDNA origin, transport, and persistence in a given environment is essential to describe the present community (Barnes & Turner 2016, Collins et al. 2018, Turner et al. 2015). However, with the limited information we have for the study area, it is not possible to make absolute conclusions regarding these processes.

First, due to limited water exchange and relatively stable hydrography (Wassmann 1985), polls may not face the same challenges as highly dynamic open marine systems regarding the transport of eDNA signals (Jeunen et al. 2019). However, a study by Turner et al. (2015) found that the DNA concentration of the target fish was much higher in sediments than in water, indicating a high degree of settlement from the water column. In polls, organic matter from sewage and other sources contributes to a high degree of sedimentation, particularly in the deepest basins (Dybern 1967). Station 1 was located in the deepest basin of the inner part of the poll, possibly acting as a sedimentation pan for DNA from the water column and other nearby sites (Dybern 1967, Turner et al. 2015). Considering the pelagic species detected by 18S and COI

in this study, settling of pelagic and allochthonous benthic extracellular DNA to stations 1 and 2 likely explains the substantial amount of DNA signals. Among the most abundant species at 1B in the COI dataset are the pelagic hydrozoans *Melicertum octocostatum* and *Leuckartiara octona*, which further supports the claim that a substantial amount of DNA from the water column accumulates in the deep basin at station 1. These findings further highlight the issue of retaining OTUs of pelagic origin in the molecular datasets, resulting in an incorrect description of the local diversity. More stringent filtering of pelagic taxa would provide a different view of the richness and diversity. However, the issue of accurate pelagic filtering is a large and complex subject beyond the scope of this project.

Moreover, understanding the factors affecting the persistence of DNA is important to make accurate inferences about the present diversity (Barnes & Turner 2016, Collins et al. 2018). Considering this, Nagler et al. (2022) proposes that two constraints characterize environments: cell lysis and extracellular DNA degradation rates. Further, in environmental conditions where cell lysis is fast and DNA degradation is slow, eDNA may primarily represent extracellular DNA of past and allochthonous assemblages (Nagler et al. 2022). Such environments are linked to high microbial activity and environmental factors preserving the extracellular DNA (Nagler et al. 2022). These processes are determined by biotic and abiotic factors, such as light, oxygen, pH, and salinity (Barnes & Turner 2016), and DNA can bind to organic or inorganic particles, possibly protecting the DNA from degradation (Lorenz & Wackernagel 1987, Nagler et al. 2022). In a study on degradation rates of exDNA in anoxic deep-sea sediments by Corinaldesi et al. (2011), results indicated that anoxic sediments with high organic matter appear to favor the preservation of extracellular DNA, possibly representing past assemblages (Corinaldesi et al. 2011). Preservation and accumulation of eDNA in sediments is further demonstrated by Turner et al. (2015), where eDNA from the target fish were detected in the sediment 132 days after removing the fish from the pond. These findings are similar to Nevers et al. (2020), showing that eDNA accumulating in the sediment persisted longer than eDNA in water. Consequently, the persistence of eDNA in the sediment influence the accuracy of spatiotemporal inferences made about diversity (Barnes & Turner 2016, Turner et al. 2015), especially at sites with no DNA representing an extant metazoan community, which is assumed for stations 1 and 2. In the context of this study, making any well-founded claims about the transport and persistence of DNA is difficult. Nevertheless, based on the contradictory observations, it is conceivable that the communities detected at the heavily impacted sites mainly represent allochthonous extracellular DNA preserved within the sediment matrix.

In addition, we should note the occurrence of eelgrass at station 2. Lundberg (2015) assessed the ecological status of this eelgrass meadow in 2014 and classified it as poor. This is likely related to the high organic matter content in the bottom sediments of the inner part of the poll, which was observed in this study and previously observed by Dybern (1967). Kviturspollen has previously been a recipient of wastewater (Kvalø et al. 2015), which is recognized as an

environmental disturbance to seagrass communities (Cabaço et al. 2008), related to anoxic sediment caused by nutrient input (Cabaço et al. 2008). As a marine foundation species, eelgrass creates a habitat for the associated metazoan community (Duffy et al. 2014) and can make it more habitable for other organisms (Lundberg 2015). Thus, station 2 likely hosts an invertebrate community different from other stations, also indicated by its relative position in the spatial analysis (e.g., Figure 3.12 and 3.13). An example of this is the polychaete *Nicolea Zostericola*, which was found in grab replicate 2A by COI (Section 5.7), a species closely associated with eelgrass meadows (Nygren 2022). Additionally, the oligochaete *Tubificoides benedii* was highly abundant in grabs 2A and 2B in the COI dataset (see section 5.7). This species is a pollution indicator (AMBI EG: V) (Rygg & Norling 2013) with high tolerance to hypoxia and high levels of hydrogen sulfide (Thiermann et al. 1996). Based on the observations in the field, station 2 was assumed to be in poor condition, unlikely to support a diverse metazoan community. However, in contrast to the deeper basin of station 1, parts of the DNA signals obtained from the eelgrass meadow at station 2 may represent living species tolerant to the poor local conditions, as indicated by the presence of pollution indicator *Tubificoides benedii*. Still, the conditions are unlikely to support the diverse community suggested by the high number of DNA signals obtained by metabarcoding from this station.

There are several possible solutions to these issues at the heavily impacted sites. First, for the problem of false positives related to past assemblages, environmental RNA is a possible solution (Cristescu & Hebert 2018, Nagler et al. 2022, Pawlowski et al. 2018). This approach is based on the assumption that concentrations of RNA correlate to metabolic activity and that RNA is more unstable and degrades much more rapidly than DNA when released to the environment (Pawlowski et al. 2018, Torti et al. 2015), thus providing a higher temporal resolution. However, in a study by Orsi et al. (2013), ancient RNA signals were detected in marine sediments, indicating that RNA may persist longer in the environment than previously considered. The use of RNA may also contain additional sources of bias, making it less suitable for solving the problem of describing the active community (Brandt et al. 2020). Further, RNA is less practical than DNA for routine monitoring programs due to several technical challenges related to handling samples and extraction (Pawlowski et al. 2018), and Brandt et al. (2020) therefore favors using eDNA over eRNA for biodiversity surveys. Another alternative could be using bulk samples instead of sediment samples. This way, only animals found in the sediment sample will be detected. However, bulk samples only deal with larger, multicellular organisms (Macher et al. 2018), and processing samples can be labor-intensive (Pawlowski et al. 2022). A third possibility is to use other potential bioindicators provided by eDNA metabarcoding, such as bacteria (Aylagas et al. 2017, Pawlowski et al. 2022). In a study on a bacterial community, Aylagas et al. (2017) found a significant correlation between the increase of organic matter content and the relative abundance of sulfate-reducing bacteria, which are favoured in anoxic conditions (Aylagas et al. 2017, Diao et al. 2018). As illustrated in Figure 3.5, the bac-

terial community detected at station 1B by COI greatly exceeds the metazoan community in the number of reads. Thus, for anoxic and H<sub>2</sub>S-rich environments such as station 1 and 2, bacteria might be a possible alternative to benthic invertebrates for assessing ecological status based on the extant community.

## 4.5 Highly variable diversity estimates

In monitoring, diversity indices such as the Shannon diversity index (Direktoratsgruppen vandedirektivet 2018b) are used to classify ecological status. Due to uncertainties in richness and abundance estimates, we must interpret alpha diversity metrics from metabarcoding data with caution (Andújar et al. 2018, Bik et al. 2012, Kelly et al. 2019, Pawlowski et al. 2018). These uncertainties are related to species-specific amplification efficiencies of universal primers, resulting in considerable variation in abundance retrieved from different taxa (Elbrecht & Leese 2015). Moreover, biological factors such as intragenomic variability can affect the estimation of richness, and variation in gene copy number and primer-bias complicates using sequence reads as a proxy for abundance (Andújar et al. 2018, Bik et al. 2012, Pawlowski et al. 2018). Further, the choices made during bioinformatic handling of the data, such as threshold for filtering and OTU clustering, also affects estimates of richness (Pawlowski et al. 2018, Zinger et al. 2019). There is currently no consensus on how the data should be bioinformatically processed (Murray et al. 2015), and the results will therefore vary between studies. Lastly, molecular data is not calibrated to a state value, making classification to state values valid for morphological data only.

Keeping this in mind, it is still interesting to investigate the consistency of the molecular values in light of morphological findings. First, the values obtained from morphological data show higher diversity at stations 3 and 9 (classified to “good”, and “excellent”) and lower values for the remaining stations (“moderate”) (Table 3.5). Station 3 was located in the more current-exposed area of the poll, while station 9 is the reference station located outside of the poll. The most remarkable difference is observed at station 4, with 4A being classified as “very poor” (0.89) and 4B as “moderate” (2.89) (Figure 3.5). Station 4 was located in a more sheltered area of the poll. Here, the dominance of mollusks can explain the low value in 4A, which reduces the evenness in this sample (Figure 3.9). Furthermore, the abundance of mollusks based on morphological taxonomy may be overvalued and consequently affect diversity estimates (Vaughn 2009, referred in Ji et al. (2022)).

Shannon values obtained from the entire 18S eukaryotic dataset were very high, ranging from 5.80-7.99 (Figure 3.5), resulting from a higher number of OTUs using the entire dataset (Figure 3.9). The values obtained by both the full eukaryotic and metazoa 18S datasets were relatively consistent across grabs. The values obtained by COI were less consistent with highly different values obtained from grab replicates from certain stations, for example, 3A (5.67) and 3B

(2.20) and 9A (4.80) and 9B (1.05) (Figure 3.15 and Table 3.5). Higher inconsistency in diversity values for COI compared to 18S is also observed in Hestetun, Lanzén & Dahlgren (2021). In light of observation and findings from morphological investigations, the high values from stations 1 and 2 are particularly conspicuous. In the deep basin of station 1, I suspect that the diversity does not represent local DNA extracellular but possibly DNA from past assemblages or elsewhere in the pool. For station 2, the poor condition of the eelgrass meadow is unlikely to host a diverse metazoan community. Therefore, I assume that DNA from past assemblages or other sites also causes inaccurate diversity estimates at this station (See section 4.4.1).

Furthermore, the high value for 1B (6.27) compared to reference station 9B (1.05) (Figure 3.9) for COI points to limitations of using PCR, leading to inaccurate diversity estimates. As shown in the results, OTU richness is higher at station 1B than 9B (Figures 3.7-3.8), presumably due to the presence of extracellular DNA stored in the sediment. However, as illustrated in Figure 3.5, the number of metazoan sequence reads retrieved from 1B is much lower than 9B. Thus, fewer reads distributed among a high number of OTUs for 1B results in many low-abundance OTUs, making the sample more homogeneous. For reference, the most abundant OTU (an unclassified Metazoan) in 1B had a read abundance of 2202 reads. In contrast, station 9B is likely a more heterogeneous sample, indicated by the high number of reads from *Tubificoides benedii* of 168 636 reads (see section 5.7). According to Kelly et al. (2019), alpha diversity and Shannon values depend on the distribution of amplification efficiencies across taxa and the number of PCR cycles. Thus, high amplification efficiency for oligochaete *Tubificoides benedii* combined with the exponential increase in PCR cycles may lead to significant differences in the obtained amplicons relative to the DNA concentration in the sample (Kelly et al. 2019). As Shannon considers the relative abundance across samples (evenness), paradoxically, many equal low-abundance OTUs at 1B results in higher Shannon value than the lopsided 9B, where high abundance of *Tubificoides benedii* reduces the sample's evenness.

Observations during the fieldwork and morphological assignment indicate that station 9 is in better condition than station 1. Still, diversity values do not reflect this for 1B and 9B. First, these findings emphasize the importance of grab replicates to account for spatial heterogeneity, particularly regarding station 9. Second, the difference between grab 1B and 9B illustrate the weakness of using PCR when making diversity estimates. This is related to inaccurate abundance estimates, one of the major challenges in metabarcoding (Deiner et al. 2017, Elbrecht & Leese 2015). The highly species-specific amplification efficiency causes the number of obtained sequences to vary across taxa (Elbrecht & Leese 2015, Kelly et al. 2019), which further affects the diversity metrics (Kelly et al. 2019). Third, the findings from stations 1 and 2 underline the challenge of making reliable diversity estimates for heavily contaminated sites of mainly extracellular DNA.

Several studies have investigated the effect of extracellular DNA on diversity estimates (Carini

et al. 2017, Lennon et al. 2018), with opposite conclusions. First, Carini et al. (2017) found that removing extracellular DNA (referred to as “relic DNA”) significantly reduced soil microbial diversity estimates. In contrast, Lennon et al. (2018) found that relic DNA had a minimal effect on microbial diversity. However, based on the observations during field and sample processing, extracellular DNA originating from other sites is assumed to constitute most of the total DNA pool at stations 1 and 2, thus significantly obscuring the diversity estimates. For the remaining stations, the impact of extracellular DNA is less evident. First, the morphological investigations confirm the presence of living species at stations 3-6 and 9. Therefore, it is more likely that parts of the extracellular DNA in the eDNA pool at these stations represent the active community releasing extracellular DNA in biomass turnover, as assumed by Pansu et al. (2021).

#### **4.6 Beta diversity patterns are similar across datasets, but reveal incomplete sampling for certain stations**

Despite obtaining very different taxonomic compositions from each method and marker, similar beta diversity patterns were observed across the datasets. First, the spatial analyses illustrate the ability of eDNA metabarcoding to separate stations successfully. As shown in nMDS plots (Figures 3.11-3.14), grab replicates from the same stations are most similar to each other, thus accounting for spatial heterogeneity at the meter scale at most stations (Hestetun, Lanzén & Dahlgren 2021). However, there are some exceptions to this conclusion. In the nMDS plot based on the morphological dataset, grabs 4A and 5B are more similar to each other than their corresponding grab replicates. At station 4, sampling was challenging due to rocky bottoms, indicating a patchy local topography that may affect the spatial distribution of benthic species.

The grab replicates from station 9 are also clearly separated in nMDS plots from all datasets (Figures 3.11-3.14). This finding is interesting, considering that station 9 in the morphological and molecular dataset was not based on the same grab samples. First, station 9 is a reference station located in a more typical marine environment, less affected by freshwater and organic matter input than Kviturspollen. Still, habitat heterogeneity in the near-shore sub-littoral position of station 9 likely influences the community composition. For the morphological data, the difference between grab replicates can be explained by sample 9B being more heterogeneous, containing a large piece of macroalgae (Table 3.1). The benthic community associated with the macroalgae is likely to be different than the benthic community collected from grab 9A, which consisted only of sediment. The number of taxa found by the morphological investigation reflects these differences, with half the number found in 9B compared to 9A (Table 3.2). For the molecular datasets, small-scale spatial heterogeneity may have caused the separation of 9C and 9D. Due to the scale of this project, we only collected two sample replicates per station, despite the recommendations in the standard methodology of three or more repli-

## **4.7 Macrofauna indices do not perform well on molecular data retrieved from a poll**

cate (ISO 2014). For eDNA metabarcoding, we only collected one subsample per grab. Thus, the separation of grab replicates from station 9 illustrates the importance of additional replicates to account for spatial heterogeneity on the meter scale for morphology and small-scale for eDNA.

Despite the separation of replicates for station 9, the spatial analyses show that molecular data can reflect the sampled stations' geographical positions (Figure 2.1). Stations 3 and 6 are plotted relatively close together in nMDS plots across all datasets (Figures 3.11-3.14). These stations are located in a more current-exposed area of the poll and had similar sediment consisting of shell sand (Table 3.1). Likewise, grab replicates from station 4 and grab 5B are plotted relatively closely together, with grab 5A being more distanced. These stations are more sheltered with sediments consisting of mud with a weak smell of H<sub>2</sub>S. Station 1 is plotted in the middle of all other grabs, consistent with its geographical position between stations 2 and 3. Station 2 is distanced from the other grabs, which is most evident in the nMDS plot by 18S (Metazoa) (Figure 3.13). Station 2 was in an eelgrass meadow, likely to host a community different from the other stations. Station 9 is located outside the study area, as reflected by each nMDS plot (Figures 3.11-3.14).

Moreover, in addition to similar patterns in the spatial analyses, Mantel testing the Bray-Curtis distance matrices showed a significant positive correlation across all datasets (Table 3.6). Thus, despite the unequal recovery of taxonomic groups across datasets, the methods are equally informative on changes in community composition in the study area (Serrana et al. 2019). These findings are similar to Cahill et al. (2018) and Grey et al. (2018), showing that beta diversity patterns are robust to the choice of method and marker. As previously highlighted, the stations differ from each other in terms of physical and biological characteristics. This differentiation between stations is reflected in nMDS plots from each dataset, illustrating the promise of eDNA metabarcoding for community analyses. In contrast to alpha diversity, we find a higher degree of congruence for beta diversity patterns across the datasets.

## **4.7 Macrofauna indices do not perform well on molecular data retrieved from a poll**

In standard monitoring, several biotic indices are used to assess ecological status based on morphological taxonomy. Biotic indices for Norwegian waters are NSI, ISI and NQI1 (Direktoratsgruppen vanndirektivet 2018b), but I also included AMBI here since it is a commonly used index in Europe (Lejzerowicz et al. 2015, Rygg & Norling 2013). Accurate calculation of these indices requires sufficient taxonomic resolution since species within the same genus may have highly different sensitivity values or belong to different ecological groups (e.g., *Pholoe* sp. vs. *Pholoe baltica*) (Direktoratsgruppen vanndirektivet 2018b, Rygg & Norling 2013).

Therefore, 18S has not previously been considered applicable due to the low taxonomic resolution reported for metazoans (Lanzén et al. 2021). In this study, a higher number of OTUs were assigned to the species level for 18S compared to COI. Despite being based on the traditional morpho-taxa approach (Direktoratsgruppen vanndirektivet 2018b, Rygg 2002, Rygg & Norling 2013), the biotic indices have been applied to several studies using metabarcoding for status assessments (Hestetun, Lanzén & Dahlgren 2021, Lanzén et al. 2021, Lejzerowicz et al. 2015): In a metabarcoding study from the Norwegian continental shelf, Lanzén et al. (2021) successfully calculated NSI based on COI data with a significant correlation with morpho-taxonomic data. Hestetun, Lanzén & Dahlgren (2021) made similar results when calculating AMBI, ISI2012, NSI2012, and NQI1 with COI data from the Norwegian continental shelf. Here, the COI data gave similar results to the standard monitoring survey of "no disturbed fauna," indicating that biotic indices could be used on COI molecular data from the continental shelf. Similarly, Lejzerowicz et al. (2015) found that applying the Infaunal Trophic Index and AMBI to 18S data (V4 region) reflected the results obtained by the morphological method.

In contrast, biotic indices based on morphological and molecular datasets were largely inconsistent in this study (Figure 3.15). NSI based on morphological data classified most stations as moderate, whereas NSI values based on molecular data were higher overall. NSI values based on COI and 18S classify stations 1 and 2 as "good" and "excellent", and the two markers displayed similar patterns across the grabs (Figure 3.15). For ISI, COI values differ from both 18S and morphology by distinctly higher values, with all grabs classified as "excellent". However, ISI based on 18S data is more similar to morphology ranging from "moderate" to "excellent". For morphological data, NQI1 values are more consistent across stations and overall higher than molecular NQI1 values. The NQI1 index values based on COI data are highly inconsistent across grabs, while 18S is slightly less variable. For AMBI, the values are consistent across grabs based on morphological data, with all classified as "slightly polluted". In contrast, AMBI values based on COI and 18S are highly variable between grabs and inconsistent with morphological values (Figure 3.15). In conclusion, the overall pattern from biotic index calculation shows that molecular data is inconsistent with morphology.

These inconsistent findings across datasets contrast with Lanzén et al. (2021), Hestetun, Lanzén & Dahlgren (2021), and Lejzerowicz et al. (2015) and indicate challenges in applying macrofauna-based biotic indices to the data retrieved for the poll. The characteristics of the studied habitat combined with methodological limitations may present challenges in using eDNA to assess ecological status with existing biotic indices. As previously discussed, the transport and preservation of extracellular DNA may give an inaccurate description of the diversity at some stations. First, ISI is based on the presence/absence of species with specific sensitivities (Rygg 2002). Consequently, OTUs with only a single sequence read is enough to affect the ISI value. As an example, the oligochaete *Tubificoides benedii*, which is highly abundant at several stations in the COI dataset, is present at station 1B with only one sequence



#### **4.7 Macrofauna indices do not perform well on molecular data retrieved from a p619**

read (see Section 5.7). This is presumably extracellular DNA from other stations or possible contamination, thus representing species that do not currently reside in that location. Although it is assumed that extracellular DNA is less evident in stations 3-9, the high ISI values for COI indicate the potential contribution of pelagic and extracellular DNA.

Additionally, extracellular DNA may also affect NSI, NQI1, and AMBI resulting in inconsistent values at most stations compared to the morphological values (Figure 3.15), but these indices are also affected by inaccurate abundance estimates. Further, the calculation of biotic indices was performed using the package “BBI” in R (Cordier & Pawlowski 2018), which is based on ISI2012 and NSI2012, which is the same version of the indexes used in previous analyses of Kviturspollen (Kvalø et al. 2015, 2014). NSI2012 and ISI2012 were revised in 2018, resulting in a new species list with new sensitivity values and class boundaries (Borgersen et al. 2020). For example, in contrast to the 2012 versions, the 2018 version of these indices include the pollution indicator *Tubificoides benedii*, which is placed in ISI and NSI ecological group 4 (Borgersen et al. 2020). According to Borgersen et al. (2020), the revised indices may provide more accurate assessments since they include more species with associated sensitivity value, and NIVA, therefore, recommends the replacement of the 2012-versions with ISI2018 and NSI2018 (Borgersen et al. 2020). Considering the conspicuously high values of NSI and ISI using the 2012 versions in this study, it would be interesting to see the performance using the new revised versions, which include more species with assigned sensitivity values (Borgersen et al. 2020).

Although 18S displays relatively high taxonomic resolution, most of the species from the 18S metazoan dataset belonged to meiofaunal phyla, which are not included in routine monitoring surveys or indices used (Rygg & Norling 2013). The inference of biotic indices requires that the obtained taxa are assigned a sensitivity value and to an ecological group (Mauffrey et al. 2020). Furthermore, the calculation is limited to the fraction of taxa represented in a reference database (Cordier et al. 2021, Pawlowski et al. 2018). The observed differences between index values calculated from 18S and COI can be explained by using different databases, with Silva for 18S and Midori (curated GenBank repository) for COI. Also, primer-bias related to each marker resulted in different taxonomic compositions on which the calculation is based. Thus, despite the high taxonomic resolution of 18S obtained in this study, 18S may have limited informational value when applied to the existing macrofauna-based indices. However, biotic indices for meiofauna and nematodes do exist (Moens et al. 2022). Research on the potential use of meiofauna as a tool for ecological assessment has shown that meiofauna-based indices perform equally well compared to macrofauna-based indices (Moens et al. 2022). These findings demonstrate the potential of eDNA metabarcoding to extend the use of biological quality elements to meiofauna and protists (Lejzerowicz et al. 2015). Finally, in the context of biomonitoring, integrating metabarcoding data to existing biotic indices is proposed as a short-term solution for the implementation of metabarcoding to standard monitoring (Pawlowski et al.

2018). This study illustrates the challenges of using the existing biotic indices in a poll and highlight the need for developing a new indices based on molecular data.

### **An alternative to the taxonomy-based method using a de novo index**

As an alternative, several studies have developed a de novo index for impact and ecological assessments using benthic invertebrates (Lanzén et al. 2021, Mauffrey et al. 2020) and bacteria (Aylagas et al. 2017). The taxonomy-based approach used in this study is limited due to several factors. First, the metabarcoding data used for index calculation is significantly reduced by only using macroinvertebrates. Second, the number of OTUs used for biotic index calculation is limited to the number of macroinvertebrates represented in reference libraries and assigned a sensitivity value (Cordier et al. 2021, Mauffrey et al. 2020). By identifying new bioindicators from the total number of retrieved OTUs, the de novo approach is not affected by these limitations (Cordier et al. 2021). In an impact assessment using benthic eDNA, Lanzén et al. (2021) developed a de novo biotic index based on the COI and 18S datasets. Here, Lanzén et al. (2021) used a pressure index based on several physicochemical indicators of impact (e.g., total hydrocarbons, metals) to identify new bioindicators from the dataset (Lanzén et al. 2021). The pressure index was developed with defined class values, and the identified bioindicator taxa were assigned to an ecological group (Lanzén et al. 2021). The results from Lanzén et al. (2021) demonstrated that the performance of de novo biotic indices was comparable to that of NSI. Although Lanzén et al. (2021) assigned OTUs to a taxon, OTUs can be assigned an autoecological value without being identified to a taxon, thus avoiding the limitation associated with incomplete databases (Cordier et al. 2021). The approach can offer a more holistic view of the ecosystem by using additional bioindicators, such as protists (Aylagas et al. 2016, Cordier et al. 2021). However, developing a de novo index requires a large dataset and is still in the early phase of development (Hestetun, Lanzén, Bagi, Ray, Larsen & Dahlgren 2021).

The number of stations and datasets in this study is inadequate for de novo index development, but this alternative approach can offer relevant future monitoring opportunities in poll-systems. In this study, the nMDS plot based on the full 18S eukaryotic dataset illustrates the ability of eDNA metabarcoding to successfully separate stations (Figure 3.14), which is promising for the potential use of de novo indices. The identification of bioindicators from the entire dataset is based on local indicators of impact (basis of the pressure index) (Lanzén et al. 2021), thus making the distinction of sites a requirement. The stations are different in terms of sediment, hydrodynamics, and other environmental variables not accounted for in this study, and the spatial analyses clearly separate stations accordingly. Lanzén et al. (2021) identified 118 potential bioindicators based on 18S, including many protists. In this study, the protists Stramenopiles, Alveolata, and Rhizaria, constituted 74% of the total 18S dataset, representing potential bioindicators. In a study on benthic community composition in estuaries, Chariton et al. (2015) used a similar approach and found that the community composition was corre-

#### **4.7 Macrofauna indices do not perform well on molecular data retrieved from a pool**

lated with several environmental variables such as nutrients, pH, and turbidity. These findings further underline the potential of using metabarcoding for increased insight into the relationship between community composition and environmental variables (Chariton et al. 2015). In the long term, Pawlowski et al. (2018) recommends developing new molecular indices, which must be intercalibrated against the existing ones (Aylagas et al. 2018, Pawlowski et al. 2018).



# Chapter 5

## Conclusions and Future Work

This study's objective was to assess the applicability of eDNA metabarcoding for diversity-based assessment of a marine poll based on marine benthos, in comparison to the current morphological approach. The results show the ability of eDNA metabarcoding to capture a wide range of taxonomic groups. However, the partial taxonomic overlap between morphological and molecular datasets illustrates how each method gives insight into different aspects of local diversity. These findings highlight the importance of using multiple markers to achieve greater taxonomic coverage and reduce the effect of primer-specific bias. Metabarcoding increased the taxonomic coverage by including meiofaunal taxa, which are promising alternatives to macrofauna for ecological status assessment. Additionally, the taxonomic resolution of 18S was high compared to previous reports, providing species-level information on several meiofaunal taxa, thus illustrating the potential of 18S with greater local database coverage. Beta diversity proved robust to method and marker choice, and the results show that the methods are equally informative on changes in community composition.

From the morphological analysis, the spatial variability in the distribution of marine benthos was consistent with local variation in hydrodynamic conditions, sediment type, and depth. However, DNA signals retrieved from the highly impacted stations in the innermost part of the poll do not conform with the observed conditions of high organic matter content and H<sub>2</sub>S. The characteristics of the studied environment affect the transport and persistence of eDNA, with the accumulation of allochthonous extracellular DNA obscuring diversity estimates in the deepest basins. Consequently, these findings create uncertainty regarding the description of the present local biodiversity and complicate the inference of ecological status based on species presence. Furthermore, the inference of the current macrofauna biotic indices based on metabarcoding data was not feasible. Again, this was likely due to the presence of extracellular DNA, causing inconsistent values. These findings underline our need for understanding the ecology of eDNA, meaning the origin, transport, and preservation (Barnes & Turner 2016). The potential outcome could be a directly incorrect description of biodiversity, leading to

ecosystem mismanagement.

In conclusion, this study illustrates that eDNA is not *yet* suitable as a replacement for the traditional morphological method, especially not in poll-like environments. We need more insight into the origin, transport, and persistence of DNA in the environment to make accurate estimates of the present diversity. Despite the unsuccessful inference of biotic indices, the beta diversity patterns show that it is possible to use the *de novo* approach in the future. Thus, with increased efforts from interdisciplinary research networks, DNA-based methods can be a valuable monitoring tool in the near future. To achieve this goal, I recommend some essential steps to include in further research.

## 5.1 Future outlook/research

Environmental DNA metabarcoding show promising results, but before the method can be fully incorporated into standard monitoring programs, a few issues must be resolved. In this context, I propose more efforts to the following steps:

### *Expand barcode reference libraries*

Similar to other studies, gaps in the reference databases continues to be a limitation to eDNA metabarcoding (Cowart et al. 2015, Duarte et al. 2021, Dunshea et al. 2021). Thus, increased efforts to building a reference database will reduce the bottleneck that these databases currently represent.

### *Increased knowledge on the ecology of eDNA*

In this study, the presence of extracellular DNA from other sites or past assemblages resulted in inaccurate estimates of local diversity at the highly impacted sites. This is related to the characteristics of the habitat, where complex topography and partly anoxic conditions make poll-like environments challenging for eDNA monitoring. The potential outcome could be a directly incorrect description of biodiversity, leading to ecosystem mismanagement. In order to make accurate spatiotemporal inferences about species presence based on DNA signals, we need more insight into the origin, transport, and persistence of eDNA in different systems (Barnes & Turner 2016).

### *Further validation of the method in different habitats*

Consequently, before implementing DNA-based methods to standard monitoring programs, eDNA metabarcoding must be further tested in different habitats where this has not been done before (Blancher et al. 2022). This is crucial for understanding habitat-specific limitations and challenges. In this study, it is evident that the complex topography and substantial input of organic matter in polls introduce some challenges for molecular tools. For this objective, pilot studies can provide insight into appropriate choices for the experimental protocol (e.g.,

sampling and laboratory protocol) in different habitats (Dunshea et al. 2021). In this process, the performance of eDNA metabarcoding should be compared to morphological results to identify challenges (Bourlat et al. 2013, Dunshea et al. 2021).

#### *Standardized protocols in accordance with WFD*

Standardized and intercalibrated protocols for monitoring water bodies are currently anchored in the Water Framework Directive. Further, DNA-based approaches show potential as an alternative method for diversity-based monitoring, but different approaches are developed across institutions (Leese et al. 2016). Thus, if DNA-based methods are to be implemented in standard monitoring, consensus and standardization of protocols are necessary (Aylagas et al. 2016, Blancher et al. 2022, Bourlat et al. 2013, Leese et al. 2016).

#### *Development of molecular indices*

A short-term solution regarding biotic indices proposed by Pawlowski et al. (2018) is integrating molecular data into existing indices. However, this proved difficult in this study, and the results further underline the need to develop indices based on metabarcoding data (Bourlat et al. 2013, Leese et al. 2016, Pawlowski et al. 2018). As previously mentioned, the de novo approach has shown great potential for identifying new bioindicators from the total number of obtained OTUs (Lanzén et al. 2021). A challenge with this method is to making the identified bioindicators useful in a new environmental context (Cordier et al. 2021). For this objective, (Cordier et al. 2017) has explored opportunities to use supervised machine learning (SML) to infer biotic indices from eDNA metabarcoding data, and findings demonstrate the potential use of SML for biomonitoring (Tristan Cordier 2018). Further exploration and validation of these tools is required before they can be implemented in standard monitoring (Cordier et al. 2017).





# Bibliography

- Alberdi, A., Aizpurua, O., Gilbert, M. T. P. & Bohmann, K. (2018), 'Scrutinizing key steps for reliable metabarcoding of environmental samples', *Methods in Ecology and Evolution* **9**(1), 134–147.  
**URL:** <https://doi.org/10.1111/2041-210X.12849>
- Andrews, S. (2010), 'Fastqc: A quality control tool for high throughput sequence data', Available at <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Andújar, C., Arribas, P., Yu, D. W., Voegler, A. P. & Emerson, B. C. (2018), 'Why the coi barcode should be the community dna metabarcode for the metazoa', *Molecular Ecology* **27**, 3968–3975.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.14844>
- Aylagas, E., Borja, A., Irigoien, X. & Rodríguez-Ezpeleta, N. (2016), 'Benchmarking DNA Metabarcoding for Biodiversity-Based Monitoring and Assessment', *Frontiers in Marine Science* **3**.  
**URL:** <https://doi.org/10.3389/fmars.2016.00096>
- Aylagas, E., Borja, A., Muxika, I. & Rodríguez-Ezpeleta, N. (2018), 'Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks', *Ecological Indicators* **95**, 194–202.  
**URL:** <https://doi.org/10.1016/j.ecolind.2018.07.044>
- Aylagas, E., Borja, A., Tangherlini, M., Dell'Anno, A., Corinaldesi, C., Mitchell, C. T., Irigoien, X., Danovaro, R. & Rodríguez-Ezpeleta, N. (2017), 'A bacterial community-based index to assess the ecological status of estuarine and coastal environments', *Marine Pollution Bulletin* **114**(2), 679–688.  
**URL:** <https://doi.org/10.1016/j.marpolbul.2016.10.050>
- Baird, D. J. & Hajibabaei, M. (2012), 'Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing', *Molecular Ecology* **21**, 2039–2044.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05519.x>

Barnes, M. A. & Turner, C. R. (2016), 'The ecology of environmental DNA and implications for conservation genetics', *Conservation genetics* **17**(1), 117.

**URL:** <https://doi.org/10.1007/s10592-015-0775-4>

Beaumont, N. J., Austen, M. C., Atkins, J. P., Burdon, D., Degraer, S., Dentinho, T. P., Derous, S., Holm, P., Horton, T., van Ierland, E., Marboe, A. H., Starkey, D. J., Townsend, M. & Zarzycki, T. (2007), 'Identification, definition and quantification of goods and services provided by marine biodiversity: Implications for the ecosystem approach', *Marine Pollution Bulletin* **54**, 253–265.

**URL:** <https://www.sciencedirect.com/science/article/pii/S0025326X06005200>

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K. & Das, I. (2007), 'Cryptic species as a window on diversity and conservation', *Trends in Ecology Evolution* **22**, 148–155.

**URL:** <https://www.sciencedirect.com/science/article/pii/S0169534706003703>

Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, J. G., Knight, R. & Thomas, W. K. (2012), 'Sequencing our way towards understanding global eukaryotic biodiversity', *Trends in Ecology and Evolution* **27**, 233–243.

**URL:** <https://www.sciencedirect.com/science/article/pii/S0169534711003338>

Blancher, P., Lefrançois, E., Rimet, F., Vasselon, V., Argillier, C., Jens Arle, P. B., Boets, P., Boughaba, J., Chauvin, C., Deacon, M., Duncan, W., Ejdung, G., Erba, S., Ferrari, B., Fischer, H., Hänfling, B., Haldin, M., Hering, D., Hette-Tronquart, N., Hiley, A., Järvinen, M., Jeannot, B., Kahlert, M., Kelly, M., Kleinteich, J., Koyuncuolu, S., Krenek, S., Langhein-Winther, S., Leese, F., Mann, D., Marcel, R., Marcheggiani, S., Meissner, K., Mergen, P., Monnier, O., Narendja, F., Neu, D., Pinto, V. O., Pawlowska, A., Pawlowski, J., Petersen, M., Poikane, S., Pont, D., Renevier, M.-S., Sandoy, S., Svensson, J., Trobajo, R., Zagyva, A. T., Tziortzis, I., van der Hoorn, B., Vasquez, M. I., Walsh, K., Weigand, A. & Bouchez, A. (2022), 'A strategy for successful integration of DNA-based methods in aquatic monitoring', *Metabarcoding and Metagenomics* **6**, e85652.

**URL:** <https://doi.org/10.3897/mbmg.6.85652>

Blaxter, M. L. (2004), 'The Promise of a DNA Taxonomy', *Philosophical Transactions: Biological Sciences* **359**, 669–679.

**URL:** <http://www.jstor.org/stable/4142260>

Blaxter, M. L., Ley, P. D., Garey, J. R., Liu, L. X., Scheldeman, P., Vierstraete, A., Vanfleteren, J. R., Mackey, L. Y., Dorris, M., Frisse, L. M., Vida, J. T. & Thomas, W. K. (1998), 'A molecular evolutionary framework for the phylum Nematoda', *Nature* **392**, 7175.

**URL:** <https://doi.org/10.1038/32160>

- Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R. & Abebe, E. (2005), 'Defining Operational Taxonomic Units Using DNA Barcode Data', *Philosophical Transactions: Biological Sciences* **360**, 1935–1943.  
**URL:** <http://www.jstor.org/stable/30040940>
- Borgersen, G., Hektoen, M., Melsom, F. & Todt, C. (2020), Uttesting av sensitivitetsindeksene isi2018 og nsi2018, og en revidert artsliste med sensitivitetsverdier for bløtbunnsfauna, Technical report. NIVA-rapport ISSN 1894-7948.  
**URL:** <https://www.miljodirektoratet.no/globalassets/publikasjoner/m1696/m1696.pdf>
- Borja, A., Elliott, M., Carstensen, J., Heiskanen, A.-S. & van de Bund, W. (2010), 'Marine management Towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives', *Marine Pollution Bulletin* **60**, 2175–2186.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0025326X1000442X>
- Borja, A., Franco, J. & Pérez, V. (2000), 'A Marine Biotic Index to Establish the Ecological Quality of Soft-Bottom Benthos Within European Estuarine and Coastal Environments', *Marine Pollution Bulletin* **40**(12), 1100–1114.  
**URL:** [https://doi.org/10.1016/S0025-326X\(00\)00061-8](https://doi.org/10.1016/S0025-326X(00)00061-8)
- Borja, A., Franco, J., Valencia, V., Bald, J., Muxika, I., Belzunce, M. J. & Solaun, O. (2004), 'Implementation of the European water framework directive from the Basque country (northern Spain): a methodological approach', *Marine Pollution Bulletin* **48**, 209–218.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0025326X03005423>
- Borja, A. & Muxika, I. (2005), 'Guidelines for the use of AMBI (AZTIs Marine Biotic Index) in the assessment of the benthic ecological quality', *Marine Pollution Bulletin* **50**(7), 787–789.  
**URL:** <https://doi.org/10.1016/j.marpolbul.2005.04.040>
- Borja, A., Muxika, I. & JFranco (2003), 'The application of a Marine Biotic Index to different impact sources affecting soft-bottom benthic communities along European coasts', *Marine Pollution Bulletin* **46**(7), 835–845.  
**URL:** [https://doi.org/10.1016/S0025-326X\(03\)00090-0](https://doi.org/10.1016/S0025-326X(03)00090-0)
- 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**, 19–31.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0025326X13002890>

- Brandt, M. I., Trouche, B., Henry, N., Liautard-Haag, C., Maignien, L., de Vargas, C., Wincker, P., Poulain, J., Zeppilli, D. & Arnaud-Haond, S. (2020), 'An assessment of environmental metabarcoding protocols aiming at favoring contemporary biodiversity in inventories of deep-sea communities', *Frontiers in Marine Science* **7**, 234.  
**URL:** <https://doi.org/10.3389/fmars.2020.00234>
- Bray, J. R. & Curtis, J. T. (1957), 'An ordination of the upland forest communities of southern wisconsin', *Ecological Monographs* **27**(4), 325–349.  
**URL:** <https://doi.org/10.2307/1942268>
- Breitbart, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., Telszewski, M., Yasuhara, M. & Zhang, J. (2018), 'Declining oxygen in the global ocean and coastal waters', *Science* **359**, eaam7240.  
**URL:** <https://www.science.org/doi/abs/10.1126/science.aam7240>
- Cabaço, S., Machás, R., Vieira, V. & Santos, R. (2008), 'Impacts of urban wastewater discharge on seagrass meadows (*Zostera noltii*)', *Estuarine, Coastal and Shelf Science* **78**(1), 1–13.  
**URL:** <https://doi.org/10.1016/j.ecss.2007.11.005>
- Cahill, A. E., Pearman, J. K., Borja, A., Carugati, L., Carvalho, S., Danovaro, R., Dashfield, S., David, R., Féral, J.-P., Olenin, S. & iaulys, A. (2018), 'A comparative analysis of metabarcoding and morphologybased identification of benthic communities across different regional seas', *Ecology and Evolution* **8**(17).  
**URL:** <https://doi.org/10.1002/ece3.4283>
- Callahan, B. J., McMurdie, P. J. & Holmes, S. P. (2017), 'Exact sequence variants should replace operational taxonomic units in marker-gene data analysis', *The ISME Journal volume* **11**(12), 26392643.  
**URL:** <https://doi.org/10.1038/ismej.2017.119>
- Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S. & Fierer, N. (2017), 'Relic DNA is abundant in soil and obscures estimates of soil microbial diversity', *Nature microbiology* **2**(3), 16242–16242.  
**URL:** [10.1038/nmicrobiol.2016.242](https://doi.org/10.1038/nmicrobiol.2016.242)
- Chao, A., Chazdon, R. L., Colwell, R. K. & Shen, T.-J. (2005), 'A new statistical approach for assessing similarityof species composition with incidence and abundance data', *Ecological Letters* **8**(2), 148–159.

- Chariton, A. A., Stephenson, S., Morgan, M. J., Steven, A. D. L., Colloff, M. J., Court, L. N. & Hardy, C. M. (2015), 'Metabarcoding of benthic eukaryote communities predicts the ecological condition of estuaries', *Environmental Pollution* **203**, 165–174.  
**URL:** <https://doi.org/10.1016/j.envpol.2015.03.047>
- Clarke, K. R. & Gorley, R. N. (2015), *PRIMER v7: User Manual/Tutorial*, PRIMER-E: Plymouth.  
**URL:** [http://updates.primer-e.com/primer7/manuals/User\\_manual\\_v7a.pdf](http://updates.primer-e.com/primer7/manuals/User_manual_v7a.pdf)
- Clarke, L. J., Beard, J. M., Swadling, K. M. & Deagle, B. E. (2017), 'Effect of marker choice and thermal cycling protocol on zooplankton dna metabarcoding studies', *Ecology and Evolution* **7**, 873–883.  
**URL:** <https://doi.org/10.1002/ece3.2667>
- Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., Genner, M. J. & Mariani, S. (2019), 'Nonspecific amplification compromises environmental dna metabarcoding with coi', *Methods in Ecology and Evolution* **10**(11), 1985–2001.  
**URL:** <https://doi.org/10.1111/2041-210X.13276>
- Collins, R. A., Wangensteen, O. S., O’Gormann, E. J., Mariani, S., Sims, D. W. & Genner, M. J. (2018), 'Persistence of environmental dna in marine systems', *Communications Biology* **1**(1), 185–185.  
**URL:** <https://doi.org/10.1038/s42003-018-0192-6>
- Convention of Biological Diversity (2022), 'What’s the problem?', Available at <https://www.cbd.int/marine/problem.shtml>. (Accessed: 3 May 2022).
- Convention of Biological Diversity (2012), 'Introduction', Available at <https://www.cbd.int/intro/>. (Accessed: 10 May 2022).
- Cordier, T., Alonso-Sáez, L., Apothéloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., Chariton, A., Creer, S., Frühe, L., Keck, F., Keeley, N., Laroche, O., Leese, F., Pochon, X., Stoeck, T., Pawlowski, J. & Lanzén, A. (2021), 'Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap', *Molecular Ecology* **30**, 2937–2958.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.15472>
- Cordier, T., Esling, P., Lejzerowicz, F., Visco, J., Ouadahi, A., Martins, C., Cedhagen, T. & Pawlowski, J. (2017), 'Predicting the Ecological Quality Status of Marine Environments from eDNA Metabarcoding Data Using Supervised Machine Learning', *Environmental science technology* **51**(16), 91189126.  
**URL:** <https://doi.org/10.1021/acs.est.7b01518>

- Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T. & Pawlowski, J. (2019), 'Embracing environmental genomics and machine learning for routine biomonitoring', *Trends in Microbiology* **27**(5), 387–397.  
**URL:** <https://doi.org/10.1016/j.tim.2018.10.012>
- Cordier, T. & Pawlowski, J. (2018), 'BBI: an R package for the computation of Benthic Biotic Indices from composition data', *Metabarcoding and Metagenomics* **2**, 1–4.  
**URL:** <https://doi.org/10.3897/mbmg.2.25649>
- Corinaldesi, C., Barucca, M., Luna, G. M. & Dell'Anno, A. (2011), 'Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments', *Molecular Ecology* **20**(2), 642–654.  
**URL:** <https://doi.org/10.1111/j.1365-294X.2010.04958.x>
- Corinaldesi, C., Tangherlini, M., Manea, E. & Dell'Anno, A. (2018), 'Extracellular DNA as a genetic recorder of microbial diversity in benthic deep-sea ecosystems', *Scientific reports* **8**, 1839.  
**URL:** <https://doi.org/10.1038/s41598-018-20302-7>
- Cowart, D. A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Miné, J. & Arnaud-Haond, S. (2015), 'Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities', *PloS one* **10**(2), e0117562–e0117562.  
**URL:** <https://doi.org/10.1371/journal.pone.0117562>
- Cristescu, M. E. & Hebert, P. D. N. (2018), 'Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation', *Annual review of Ecology, Evolution and Systematics* **49**, 209–230.  
**URL:** <https://doi.org/10.1146/annurev-ecolsys-110617-062306>
- Danovaro, R., Carugati, L., Berzano, M., Cahill, A. E., Carvalho, S., Chenuil, A., Corinaldesi, C., Cristina, S., David, R., Dell'Anno, A., Dzhembekova, N., Garcés, E., Gasol, J. M., Goela, P., Féral, J.-P., Ferrera, I., Forster, R. M., Kurekin, A. A., Rastelli, E., Marinova, V., Miller, P. I., Moncheva, S., Newton, A., Pearman, J. K., Pitois, S. G., Reñé, A., Rodríguez-Ezpeleta, N., Saggiomo, V., Simis, S. G. H., Stefanova, K., Wilson, C., Martire, M. L., Greco, S., Cochrane, S. K. J., Mangoni, O. & Borja, A. (2016), 'Implementing and Innovating Marine Monitoring Approaches for Assessing Marine Environmental Status', *Frontiers in Marine Science* .  
**URL:** <https://doi.org/10.3389/fmars.2016.00213>
- Dauer, D. M. (1993), 'Biological criteria, environmental health and estuarine macrobenthic community structure', *Marine Pollution Bulletin* **26**, 249–257.  
**URL:** <https://www.sciencedirect.com/science/article/pii/0025326X9390063P>

- de Groot, R. S., Wilson, M. A. & Boumans, R. M. J. (2002), 'A typology for the classification, description and valuation of ecosystem functions, goods and services', *Ecological Economics* **41**, 393–408.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0921800902000897>
- Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F. & Taberlet, P. (2014), 'Dna metabarcoding and the cytochrome c oxidase subunit i marker: not a perfect match', *Biology Letters* **10**, 20140562.  
**URL:** <https://royalsocietypublishing.org/doi/abs/10.1098/rsbl.2014.0562>
- Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E. & Bernatchez, L. (2017), 'Environmental dna metabarcoding: Transforming how we survey animal and plant communities', *Molecular Ecology* **26**, 5872–5895.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.14350>
- Diao, M., Huisman, J. & Muyzer, G. (2018), 'Spatio-temporal dynamics of sulfur bacteria during oxic–anoxic regime shifts in a seasonally stratified lake', *FEMS microbiology ecology* **94**(4), 1.  
**URL:** <https://doi.org/10.1093/femsec/fiy040>
- Diaz, R. & Rosenberg, R. (1995), 'Marine benthic hypoxia: A review of its ecological effects and the behavioural responses of benthic macrofauna', *Oceanography and Marine Biology: an Annual Review* **33**, 245–303.
- Direktoratet for naturforvaltning (2007), 'Kartlegging av marint biologisk mangfold. dn-håndbok 19-2001. revidert 2007'. (Accessed: 14 May 2022).  
**URL:** [https://www.miljodirektoratet.no/globalassets/publikasjoner/dirnat2/attachment/69/handbok-19-2001rev-2007\\_marin\\_nettt.pdf](https://www.miljodirektoratet.no/globalassets/publikasjoner/dirnat2/attachment/69/handbok-19-2001rev-2007_marin_nettt.pdf)
- Direktoratsgruppen vanndirektivet (2018a), 'Vedlegg til veileder 02:2018. klassifisering av miljøtilstand i vann', Online. Oslo: Direktoratgruppen for gjennomføringen av vannforskriften.
- Direktoratsgruppen vanndirektivet (2018b), 'Veileder 02:2018. klassifisering av miljøtilstand i vann.'. Oslo: Direktoratgruppen for gjennomføringen av vannforskriften.
- Duarte, S., Leite, B. R., Feio, M. J., Costa, F. O. & Filipe, A. F. (2021), 'Integration of dna-based approaches in aquatic ecological assessment using benthic macroinvertebrates', *Water* **13**(3), 331.  
**URL:** <https://doi.org/10.3390/w13030331>

- Duffy, J. E., Hughes, A. R. & Moksnes, P. (2014), Ecology of seagrass communities, in M. D. Bertness, J. F. Bruno, B. R. Silliman & J. J. Stachowicz, eds, 'Marine Community Ecology and Conservation', Sinauer Associates, Inc., pp. 271–297.
- Dunsha, G., Martell, L., Bakken, T., Budaeva, N., Ekrem, T., Tandberg, A. H. S., Baussant, T., Boer, H. D., Hestetun, J. T., Hobæk, A., Kallioniemi, E. P., Larsen, A., Markussen, S. S., Mauvisseau, Q., Ray, J. L., Yoccoz, N. & Willassen, E. (2021), 'Kunnskapsstatus for bruk av molekylære verktøy i kartlegging og overvåking av biologisk mangfold i marine miljø'. Rapport M-2062. p. 76. Miljødirektoratet.  
**URL:** <https://www.miljodirektoratet.no/publikasjoner/2021/juni-2021/kunnskapsstatus-for-bruk-av-molekylare-verktoy-i-kartlegging-og-overvakning-av-biologisk-mangfold-i-marine-miljo/>
- Dybern, B. I. (1967), 'Topography and hydrography of Kviturdvickollen and Vågsbøpollen on the west coast of Norway', *Sarsia* **30**, 1–27.
- Eckert, I. M. K., Littlefair, J. E., Zhang, G. K., Chain, F. J. J., Crease, T. J. & Cristescu, M. E. (2018), Bioinformatics for Biomonitoring: Species Detection and Diversity Estimates Across Next-Generation Sequencing Platforms, in G. W. M. J. David A. Bohan, Alex J. Dumbrell, ed., 'Advances in Ecological Research', Academic Press, pp. 1–32.  
**URL:** <https://doi.org/10.1016/bs.aecr.2018.06.002>
- 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–e0130324.  
**URL:** <https://doi.org/10.1371/journal.pone.0130324>
- European Commission (n.d.), 'Introduction to the EU Water Framework Directive', Available at [https://ec.europa.eu/environment/water/water-framework/info/intro\\_en.htm](https://ec.europa.eu/environment/water/water-framework/info/intro_en.htm). (Accessed: 5 May 2022).
- Fernández, S., Rodríguez-Martínez, S., Martínez, J. L., Garcia-Vasquez, E. & Ardura, A. (2019), 'How can eDNA contribute in riverine macroinvertebrate assessment? A metabarcoding approach in the Nalón River (Asturias, Northern Spain)', *Environmental DNA* **1**(4), 385–401.  
**URL:** <https://doi.org/10.1002/edn3.40>
- Ficetola, G. F., Miaud, C., Pompanon, F. & Taberlet, P. (2008), 'Species detection using environmental dna from water samples', *Biology Letters* **4**(4), 423–425.  
**URL:** <https://doi.org/10.1098/rsbl.2008.0118>



- Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994), 'DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates', *Molecular Marine Biology and Biotechnology* **3**(5), 294–299.
- Forsgren, E. & Næss, C. (2019), 'Mudderkutling *Pomatoschistus norvegicus* (Collett, 1902)', Available at [https://www.artsdatabanken.no/Pages/237672/Pomatoschistus\\_norvegicus](https://www.artsdatabanken.no/Pages/237672/Pomatoschistus_norvegicus). (Accessed: 6 November 2022).
- Franzo, A., Baldrighi, E., Grassi, E., Grego, M., Balsamo, M., Basili, M. & Semprucci, F. (2022), 'Free-living nematodes of mediterranean ports: A mandatory contribution for their use in ecological quality assessment', *Marine Pollution Bulletin* **180**, 113814.  
**URL:** <https://doi.org/10.1016/j.marpolbul.2022.113814>
- Frøslev, T. G., Kjølner, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C. & Hansen, A. J. (2017), 'Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates', *Nature Communications* **8**, 1188.  
**URL:** <https://doi.org/10.1038/s41467-017-01312-x>
- Geller, J., Meyer, C., Parker, M. & Hawk, H. (2013), 'Redesign of PCR Primers for Mitochondrial Cytochrome C Oxidase Subunit I for Marine Invertebrates and Application in All-taxa Biotic Surveys', *Molecular Ecology Resources* **13**(5), 851–861.  
**URL:** <https://doi.org/10.1111/1755-0998.12138>
- Glémarec, M. & Hily, C. (1981), 'Perturbations apportées à la macrofaune benthique de la baie de concarneau par les effluents urbains et portuaires', *Acta Oecologica Applicata* **2**, 139–150.
- Grall, J. & Glémarec, M. (1997), 'Using biotic indices to estimate macrobenthic community perturbations in the Bay of Brest', *Estuarine, Coastal and Shelf Science* **44**, 43–53.  
**URL:** [https://doi.org/10.1016/S0272-7714\(97\)80006-6](https://doi.org/10.1016/S0272-7714(97)80006-6)
- Gray, M. W. (2015), 'Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of mitochondria', *Proceedings of the National Academy of Sciences - PNAS* **112**(33), 1013310138.  
**URL:** <https://doi.org/10.1073/pnas.1421379112>
- 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**(1), 8843.  
**URL:** <https://doi.org/10.1038/s41598-018-27048-2>

- Günther, B., Fromentin, J.-M., Metral, L. & Arnaud-Haond, S. (2021), 'Metabarcoding confirms the opportunistic foraging behaviour of Atlantic bluefin tuna and reveals the importance of gelatinous prey', *PeerJ* **9**, e11757–e11757.  
**URL:** <https://doi.org/10.7717/peerj.11757>
- Günther, B., Knebelsberger, T., Hermann Neumann, S. L. & Arbizu, P. M. (2018), 'Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes', *Scientific reports* **8**(1), 14822.  
**URL:** <https://doi.org/10.1038/s41598-018-32917-x>
- Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C. & Baird, D. J. (2011), 'Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos', *PLoS One* **6**(4), e17497.  
**URL:** <https://doi.org/10.1371/journal.pone.0017497>
- Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N. & Hickey, D. A. (2007), 'DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics', *Trends in Genetics* **23**, 167–172.  
**URL:** <https://doi.org/10.1016/j.tig.2007.02.001>
- Hammer, M. P. (2006), 'Range Extensions for Four Estuarine Gobies (Pisces: Gobiidae) in Southern Australia: Historically Overlooked Native Taxa or Recent Arrivals?', *Transactions of the Royal Society of South Australia* **130**(2), 187–196.  
**URL:** <https://doi.org/10.1080/3721426.2006.10887057>
- Harley, C. D. G., Hughes, A. R., Hultgren, K. M., Miner, B. G., Sorte, C. J. B., Thornber, C. S., Rodriguez, L. F., Tomanek, L. & Williams, S. L. (2006), 'The impacts of climate change in coastal marine systems', *Ecology letters* **9**(2), 228–241.  
**URL:** <https://doi.org/10.1111/j.1461-0248.2005.00871.x>
- He, Q. & Silliman, B. R. (2019), 'Climate Change, Human Impacts, and Coastal Ecosystems in the Anthropocene', *Current Biology* **29**, R1021–R1035.  
**URL:** <https://doi.org/10.1016/j.cub.2019.08.042>
- Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003), 'Biological Identifications through DNA Barcodes', *Proceedings: Biological Sciences* **270**, 313–321.  
**URL:** <http://www.jstor.org/stable/3558697>
- Hering, D., Borja, A., Jones, J., Pont, D., Boets, P., Bouchez, A., Bruce, K., Drakare, S., Hänfling, B., Kahlert, M., Leese, F., Meissner, K., Mergen, P., Reyjol, Y., Segurado, P., Vogler, A. & Kelly, M. (2018), 'Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive', *Water Research*

138, 192–205.

URL: <https://doi.org/10.1016/j.watres.2018.03.003>

Hestetun, J. T., Bye-Ingebrigtsen, E., Nilsson, R. H., Glover, A. G., Johansen, P.-O. & Dahlgren, T. G. (2020), 'Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding', *Marine biodiversity* **50**(70).

URL: <https://doi.org/10.1007/s12526-020-01093-5>

Hestetun, J. T., Lanzén, A., Bagi, A., Ray, J., Larsen, A. & Dahlgren, T. G. (2021), Metamon final project report., Norce environment report no. 8-2021. 61 p. ISBN 978-82-8408-181-6.

Hestetun, J. T., Lanzén, A. & Dahlgren, T. G. (2021), 'Grab what you can—an evaluation of spatial replication to decrease heterogeneity in sediment eDNA metabarcoding', *PeerJ* **9**, e11619–e11619.

URL: <https://doi.org/10.7717/peerj.11619>

Hillis, D. M. & Dixon, M. T. (1991), 'Ribosomal DNA: Molecular Evolution and Phylogenetic Inference', *The Quarterly Review of Biology* **66**(4), 411–453.

URL: <https://www.jstor.org/stable/2831326>

Hui, F. K., Taskinen, S., Pledger, S., Foster, S. D. & Warton, D. I. (2014), 'Model-based approaches to unconstrained ordination', *Methods in Ecology and Evolution* **6**(4), 399–411.

URL: <https://doi.org/10.1111/2041-210X.12236>

Hurlbert, S. H. (1971), 'The nonconcept of species diversity: A critique and alternative parameters', *Ecology* **52**, 577–586.

URL: <http://www.jstor.org/stable/1934145>

International Barcode of Life (n.d.), 'Our vision and story', Available at <https://ibol.org/about/our-vision/>. (Accessed: 27 October 2022).

Islam, M. S. & Tanaka, M. (2004), 'Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis', *Marine Pollution Bulletin* **48**, 624–649.

URL: <https://doi.org/10.1016/j.marpolbul.2003.12.004>

ISO (2014), 'ISO 16665:2014. Water quality—Guidelines for quantitative sampling and sample processing of marine soft-bottom macrofauna'. International Organization for Standardization, p. 33.

Jeunen, G., Knapp, M., Spencer, H. G., Taylor, H. R., Lamare, M. D., Stat, M., Bunce, M. & Gemmell, N. J. (2019), 'Species-level biodiversity assessment using marine environmental

- DNA metabarcoding requires protocol optimization and standardization', *Ecology and Evolution* **9**(3), 1323–1335.  
**URL:** <https://doi.org/10.1002/ece3.4843>
- Ji, F., Han, D., Yan, L., Yan, S., Zha, J. & Shen, J. (2022), 'Assessment of Benthic Invertebrate Diversity and River Ecological Status along an Urbanized Gradient Using Environmental DNA Metabarcoding and a Traditional Survey Method', *The Science of the Total Environment* **806**(2), 150587.  
**URL:** <https://doi.org/10.1016/j.scitotenv.2021.150587>
- Keck, F., Blackman, R. C., Bossart, R., Brantschen, J., Couton, M., Hürlemann, S., Kirschner, D., Locher, N., Zhang, H. & Altermatt, F. (2022), 'Meta-analysis shows both congruence and complementarity of DNA and eDNA metabarcoding to traditional methods for biological community assessment', *Molecular Ecology* **31**(6), 1820–1835.  
**URL:** <https://doi.org/10.1111/mec.16364>
- Kelly, R. P., Closek, C. J., O'Donnell, J. L., Kralj, J. E., Shelton, A. O. & Samhouri, J. F. (2017), 'Genetic and Manual Survey Methods Yield Different and Complementary Views of an Ecosystem', *Frontiers in Marine Science* **3**.  
**URL:** <https://doi.org/10.3389/fmars.2016.00283>
- Kelly, R. P., Shelton, A. O. & Gallego, R. (2019), 'Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies', *Scientific reports* **9**(1), 12133–14.  
**URL:** <https://doi.org/10.1038/s41598-019-48546-x>
- Klymus, K. E., Marshall, N. T. & Stepien, C. A. (2017), 'Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes', *PLoS One* **12**.  
**URL:** <https://doi.org/10.1371/journal.pone.0177643>
- Kvalø, S. E., Torvanger, R., Haave, M., Jacobsen, S. H., Lode, T., Johannessen, P. & Øydis Alme (2015), 'Resipientovervåking av fjordsystemene rundt Bergen 2011-2015. Årsrapport 2014.'. SAM e-rapport nr. 4-2015. Uni Research SAM-Marin.
- Kvalø, S. E., Torvanger, R., Haave, M., Johannessen, P. & Øydis Alme (2014), 'Resipientovervåking av fjordsystemene rundt Bergen 2011-2015.'. SAM e-rapport nr. 27-2014. Uni Research SAM-Marin.
- Lanzén, A., Dahlgren, T. G., Bagi, A. & Hestetun, J. T. (2021), 'Benthic eDNA metabarcoding provides accurate assessments of impact from oil extraction, and ecological insights', *Ecological Indicators* **130**, 108064.  
**URL:** <https://doi.org/10.1016/j.ecolind.2021.108064>

Lanzén, A., Jørgensen, S. L., Huson, D. H., Gorfer, M., Grindhaug, S. H., Jonassen, I., Øvreås, L. & Urich, T. (2012), 'Crest-classification resources for environmental sequence tags', *PLoS One* **7**(11), e49334.

**URL:** <https://doi.org/10.1371/journal.pone.0049334>

Leduc, N., Lacoursière-Roussel, A., Howland, K. L., Archambault, P., Sevellec, M., Normandeau, E., Dispas, A., Winkler, G., McKindsey, C. W., Simard, N. & Bernatchez, L. (2019), 'Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity', *Environmental DNA* **1**(4), 342–358.

**URL:** [10.1002/edn3.35](https://doi.org/10.1002/edn3.35)

Leese, F., Altermatt, F., Bouchez, A., Ekrem, T., Hering, D., Meissner, K., Mergen, P., Pawlowski, J., Piggott, J. J., Rimet, F., Steinke, D., Taberlet, P., Weigand, A. M., Abarenkov, K., Beja, P., Bervoets, L., Björnsdóttir, S., Boets, P., Boggero, A., Bones, A. M., Ángel Borja, Bruce, K., Bursi, V., Carlsson, J., iampor, F., iamporová Zatoviová, Z., and-Filipe Costa, E. C., Costache, M., Creer, S., Csabai, Z., Deiner, K., Ángel DelValls, Drakare, S., Duarte, S., Elerek, T., Fazi, S., Fier, C., Flot, J.-F., Fonseca, V., Fontaneto, D., Grabowski, M., Graf, W., Guðbrandsson, J., Hellström, M., Hershkovitz, Y., Hollingsworth, P., Japoshvili, B., Jones, J. I., Kahlert, M., Stroil, B. K., Kasapidis, P., Kelly, M. G., Kelly-Quinn, M., Keskin, E., Kõljalg, U., Ljubei, Z., Maek, I., Mächler, E., Mahon, A., Mareková, M., Mejdandzic, M., Mircheva, G., Montagna, M., Moritz, C., Mulk, V., Nausmoski, A., Navodaru, I., Padisák, J., Pálsson, S., Panksep, K., Penev, L., Petrussek, A., Pfannkuchen, M. A., Primmer, C. R., Rinkevich, B., Rotter, A., Schmidt-Kloiber, A., Segurado, P., Speksnijder, A., Stoev, P., Strand, M., ulius, S., Sundberg, P., Traugott, M., Tsigenopoulos, C., Turon, X., Valentini, A., van der Hoorn, B., Várбірó, G., Hadjilyra, M. I. V., Viguri, J., Vitonyt, I., Vogler, A., Vrålstad, T., Wägele, W., Wenne, R., Winding, A., Woodward, G., Zegura, B. & Zimmermann, J. (2016), 'DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe', *Research Ideas and Outcomes* **2**, e11321–24.

**URL:** <https://doi.org/10.3897/rio.2.e11321>

Lejzerowicz, F., Esling, P., Pillet, L., T A Wilding, K. D. B. & Pawlowski, J. (2015), 'High-throughput Sequencing and Morphology Perform Equally Well for Benthic Monitoring of Marine Ecosystems', *Scientific reports* **5**(13932).

**URL:** <https://doi.org/10.1038/srep13932>

Lenihan, H. S. & Micheli, F. (2001), Soft-sediment communities, in M. D. Bertness, S. D. Gaines & M. E. Hay, eds, 'Marine Community Ecology', Sinauer Associates Inc., pp. 253–287.

Lennon, J. T., Muscarella, M. E., Placella, S. A. & Lehmkuhl, B. K. (2018), 'How, When, and

- Where Relic DNA Affects Microbial Diversity’, *mBio* **9**(3), e00637–18.  
**URL:** <https://doi.org/10.1128/mBio.00637-18>
- Leray, M. & Knowlton, N. (2015), ‘DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity’, *Proceedings of the National Academy of Sciences - PNAS* **112**(7), 2076–2081.
- Leray, M. & Knowlton, N. (2016), ‘Censusing marine eukaryotic diversity in the twenty-first century’, *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**(1702), 20150331–20150331.
- Leray, M., Knowlton, N. & Machida, R. J. (2022), ‘MIDORI2: A collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences’, *Environmental DNA* **4**, 894907.  
**URL:** <https://doi.org/10.1002/edn3.303>
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudela, N., Ranwez, V., Boehm, J. T. & Machida, R. J. (2013), ‘A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents’, *Frontiers in Zoology* **10**(1), 14.  
**URL:** <https://doi.org/10.1186/1742-9994-10-34>
- Levinton, J. S. (2009), *Marine Biology: Function, Biodiversity, Ecology*, 3rd ed. edn, Oxford University Press.
- Lorenz, M. G. & Wackernagel, W. (1987), ‘Adsorption of DNA to sand and variable degradation rates of adsorbed DNA’, *Applied and Environmental Microbiology* **53**(12), 2948–2952.  
**URL:** <https://doi.org/10.1128/aem.53.12.2948-2952.1987>
- Lundberg, A. (2015), ‘Ålegrasenger i Bergen. Utbredelse, tilstand, verdi og trusler’.
- Macher, J.-N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D. & Leese, F. (2018), ‘Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers’, *Mol Ecol Resour.* **18**, 1456–1468.  
**URL:** <https://doi.org/10.1111/1755-0998.12940>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. (2015), ‘Swarm v2: highly-scalable and high-resolution amplicon clustering’, *PeerJ* **3**, e1420.  
**URL:** <https://doi.org/10.7717/peerj.1420>
- Mantel, N. (1967), ‘The detection of disease clustering and a generalized regression approach’, *Cancer research* **27**(2), 209–20.

- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. & Rothberg, J. M. (2005), 'Genome sequencing in microfabricated high-density picolitre reactors', *Nature* **437**, 376–380.  
**URL:** <https://doi.org/10.1038/nature03959>
- Martin, M. (2017), 'Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates', *EMBnet Journal* **17**(1).  
**URL:** <https://doi.org/10.14806/ej.17.1.200>
- Mauffrey, F., Cordier, T., Apothéloz-Perret-Gentil, L., Cermakova, K., Merzi, T., Delefosse, M., Blanc, P. & Pawlowski, J. (2020), 'Benthic monitoring of oil and gas offshore platforms in the North Sea using environmental DNA metabarcoding', *Molecular Ecology* **30**(13), 3007–3022.  
**URL:** <https://doi.org/10.1111/mec.15698>
- Meld.St.14 (2015), 'Meld. St. 14 (2015-2016) Nature for life —Norways national biodiversity action plan (Chapter 49)', Available at: <https://www.regjeringen.no/en/dokumenter/meld.-st.-14-20152016/id2468099/?ch=1>.
- Miljødirektoratet (n.d.), 'Figurar innan tema - vassforvaltning (nynorsk)', Available at: <https://www.vannportalen.no/kunnskapsgrunnlaget/figurer-og-illustrasjoner---tema-innen-vannforvaltning/figurar-innan-tema---vassforvaltning-nynorsk/>.
- Moens, T., Sroczynska, K. & Adão, H. (2022), 'Meiofauna in a changing world', *Ecological Indicators* **138**, 108769.  
**URL:** <https://doi.org/10.1016/j.ecolind.2022.108769>
- Mohrbeck, I., Raupach, M. J., Arbizu, P. M., Knebelsberger, T. & Laakman, S. (2015), 'High-throughput sequencing - the key to rapid biodiversity assessment of marine metazoa', *Plos ONE* **10**(10), e0140342.  
**URL:** <https://doi.org/10.1371/journal.pone.0140342>
- Molvær, J., Knutzen, J., Magnusson, J., Rygg, B., Skei, J. & Sørensen, J. (1997), 'Veiledning 97:03. klassifisering av miljøkvalitet i fjorder og kystvann', Online.

**URL:** <https://www.miljodirektoratet.no/globalassets/publikasjoner/klif2/publikasjoner/vann/1467/ta1467.pdf>

Murray, D. C., Coghlan, M. L. & Bunce, M. (2015), 'From Benchtop to Desktop: Important Considerations when Designing Amplicon Sequencing Workflows', *PloS one* **10**(4), e0124671–e0124671.

**URL:** [10.1371/journal.pone.0124671](https://doi.org/10.1371/journal.pone.0124671)

Nagler, M., Podmirseg, S. M., Archer-Jenull, J., Sint, D. & Traugott, M. (2022), 'Why eDNA fractions need consideration in biomonitoring', *Molecular ecology resources* **22**(7), 2458–2470.

**URL:** <https://doi.org/10.1111/1755-0998.13658>

National Geographic Society (2022a), 'Ocean'. (Accessed: 7 November 2022).

National Geographic Society (2022b), 'Why the ocean matters'. (Accessed: 7 November 2022).

Nevers, M. B., Przybyla-Kelly, K., Shively, D., Morris, C. C., Dickey, J. & Byappanahalli, M. N. (2020), 'Influence of sediment and stream transport on detecting a source of environmental DNA', *PloS one* **15**(12), e0244086.

**URL:** <https://doi.org/10.1371/journal.pone.0244086>

NIVA (2017), 'EU Water Framework Directive, ecological water quality, classification, environmental goals', Available at: <https://www.niva.no/en/services/eu-water-framework-directive>. Accessed: 5 May 2022.

Nygren, A. (2022), '*Nicolea zostericola* (Ørsted, 1844 in Grube, 1860)', Available at: [https://www.artsdatabanken.no/Pages/313102/Nicolea\\_zostericola](https://www.artsdatabanken.no/Pages/313102/Nicolea_zostericola). (Accessed: 7 November 2022).

O'Connor, M. I. & Byrnes, J. E. K. (2014), Biodiversity and Ecosystem Function. Does Pattern Influence the Process?, in B. R. S. Mark D. Bertness, John F. Bruno & J. J. Stachowicz, eds, 'Marine Community Ecology and Conservation', Sinauer Associates, Inc., pp. 109–130.

Orsi, W., Biddle, J. F. & Edgcomb, V. (2013), 'Deep Sequencing of Subseafloor Eukaryotic rRNA Reveals Active Fungi across Marine Subsurface Provinces', *PLoS ONE* **8**(2), e56335.

**URL:** <https://doi.org/10.1371/journal.pone.0056335>

Pansu, J., Chapman, M. B., Hose, G. C. & Chariton, A. A. (2021), 'Comparison of an extracellular v. total DNA extraction approach for environmental DNA-based monitoring of sediment biota', *Marine and Freshwater Research* .

**URL:** <https://doi.org/10.1071/MF20269>



- Pawłowski, J., Apothéloz-Perret-Gentil, L. & Altermatt, F. (2020), 'Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring', *Molecular Ecology* **29**, 4258–4264.
- Pawłowski, J., Bruce, K., Panksep, K., Aguirre, F. I., Amalfitano, S., Apothéloz-Perret-Gentil, L., Baussanti, T., Bouchez, A., Carugati, L., Cermakova, K., Cordier, T., Corinaldesi, C., Costa, F. O., Danovaro, R., Dell'Anno, A., Duarte, S., Eisendle, U., Ferrari, B. J. D., Frontalini, F., Frühe, L., Haegerbaeumer, A., Kisand, V., Krolicka, A., Lanzén, A., Leese, F., Lejzerowicz, F., Lyautey, E., Maek, I., Sagova-Mareková, M., Pearman, J. K., Pochon, X., Stoeck, T., Vivien, R., Weigand, A. & Fazi, S. (2022), 'Environmental DNA metabarcoding for benthic monitoring: A review of sediment sampling and DNA extraction methods', *Science of the Total Environment* **818**, 151783–151783.  
**URL:** <https://doi.org/10.1016/j.scitotenv.2021.151783>
- Pawłowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T. & Wilding, T. A. (2014), 'Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities', *Molecular ecology resources* **14**(6), 1129–1140.  
**URL:** <https://doi.org/10.1111/1755-0998.12261>
- Pawłowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., Feio, M. J., Filipe, A. F., Fornaroli, R., Graf, W., Herder, J., van der Hoorn, B., Jones, J. I., Sagova-Mareckova, M., Moritz, C., Barquín, J., Piggott, J. J., Pinna, M., Rimet, F., Rinkevich, B., Sousa-Santos, C., Specchia, V., Trobajo, R., Vasselon, V., Vitecek, S., Zimmerman, J., Weigand, A., Leese, F. & Kahlert, M. (2018), 'The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems', *Science of The Total Environment* **637-638**, 1295–1310.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0048969718316322>
- Pearson, T. H. & Rosenberg, R. (1978), 'Macrobenthic succession in relation to organic enrichment and pollution of the marine environment', *Oceanogr. Mar. Bio. Ann. Rev.* **16**, 229–311.
- Pedersen, A., Alve, E., Alvestad, T., Borgersen, G., Dolven, J. K., Gundersen, H., Hess, S., Kutti, T., Rygg, B., Velvin, R. & Vedal, J. (2016), Bløtbunnsfauna som indikator for miljøtilstand i kystvann. Ekspertvurderinger og forslag til nye klassegrenser og metodikk. Rapport M-633., Technical report.  
**URL:** <https://www.miljodirektoratet.no/globalassets/publikasjoner/m633/m633.pdf>
- Pereira, C. L., Gilbert, M. T. P., Araújo, M. B. & Matias, M. G. (2021), 'Fine-tuning biodiversity assessments: A framework to pair eDNA metabarcoding and morphological ap-

- proaches', *Methods in Ecology and Evolution* **12**, 2397–2409.  
**URL:** <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13718>
- Petersen, T., Klauer, B. & Manstetten, R. (2009), 'The environment as a challenge for governmental responsibility The case of the European Water Framework Directive', *Ecological Economics* **68**, 2058–2065. Methodological Advancements in the Footprint Analysis.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0921800909000329>
- Piñol, J., Mir, G., Gomez-Polo, P. & Agustí, N. (2015), 'Universal and blocking primer mismatches limit the use of highthroughput DNA sequencing for the quantitative metabarcoding of arthropods', *Molecular ecology resources* **15**(4), 819–830.  
**URL:** <https://doi.org/10.1111/1755-0998.12355>
- Prop. 93 L (2017-2018) (2017), 'Prop. 93 L (2017-2018) Endringer i naturmangfoldloven (hjemmel for gjennomføring av vanddirektivet)', Online.  
**URL:** <https://www.regjeringen.no/no/dokumenter/prop.-93-l-20172018/id2602022/?ch=1>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner, F. O. (2013), 'The SILVA ribosomal RNA gene database project: improved data processing and web-based tools', *Nucleic Acids Research* **41**(D1), D590D596.  
**URL:** <https://doi.org/10.1093/nar/gks1219>
- R Core Team (2022), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.  
**URL:** <https://www.R-project.org/>
- Rach, J., Bergmann, T., Paknia, O., DeSalle, R., Schierwater, B. & Hadrys, H. (2017), 'The marker choice: Unexpected resolving power of an unexplored CO1 region for layered DNA barcoding approaches', *PloS One* **12**(4), e0174842.  
**URL:** [10.1371/journal.pone.0174842](https://doi.org/10.1371/journal.pone.0174842)
- Ratnasingham, S. & Hebert, P. D. N. (2007), 'bold: The Barcode of Life Data System (<http://www.barcodinglife.org>)', *Molecular Ecology Notes* **7**, 355–364.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1471-8286.2007.01678.x>
- Regjeringen (2009), 'Nature diversity act'. Accessed: 23 May 2022.
- Regjeringen (2021), 'Naturmangfoldlova', Available at: <https://www.regjeringen.no/no/tema/klima-og-miljo/naturmangfold/innsiktsartikler-naturmangfold/naturmangfoldlova/id2339663/>. Accessed: 23 May 2022.
- Ricotta, C., Szeidl, L. & Pavoine, S. (2021), 'Towards a unifying framework for diversity and dissimilarity coefficients', *Ecological Indicators* **129**, 107971.  
**URL:** <https://doi.org/10.1016/j.ecolind.2021.107971>

- Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. (2016), 'Vsearch: a versatile open source tool for metagenomics', *PeerJ* **4**, e2584.  
**URL:** <https://doi.org/10.7717/peerj.2584>
- Rygg, B. (2002), 'NIVA 4548-2002 Indicator species index for assessing benthic ecological quality in marine waters of Norway'. NIVA report: Norwegian Institute for Water Research. p. 32.
- Rygg, B. (2006), 'NIVA Report 5208-2006. Developing indices for quality-status classification of marine soft-bottom fauna in Norway'. NIVA Report: Norwegian Institute for Water Research. p. 33.
- Rygg, B. & Norling, K. (2013), 'NIVA Report 6475-2013. Norwegian Sensitivity Index (NSI) for marine macroinvertebrates, and an update of Indicator Species Index (ISI)'. NIVA Report: Norwegian Institute for Water Research. p. 47.  
**URL:** <http://hdl.handle.net/11250/216238>
- Rzeznik-Orignac, J., Kalenitchenko, D., Mariette, J., Bodiou, J.-Y., Bris, N. L. & Derelle, E. (2017), 'Comparison of meiofaunal diversity by combined morphological and molecular approaches in a shallow mediterranean sediment', *Marine Biology* **164**(40), 1.  
**URL:** <https://doi.org/10.1007/s00227-017-3074-4>
- Salas, F., Marcos, C., Neto, J. M., Patrício, J., Pérez-Ruzafa, A. & Marques, J. C. (2006), 'User-friendly guide for using benthic ecological indicators in coastal and marine quality assessment', *Ocean Coastal Management* **49**, 308–331.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0964569106000378>
- Serrana, J. M., Miyake, Y., Gamboa, M. & Watanabe, K. (2019), 'Comparison of DNA metabarcoding and morphological identification for stream macroinvertebrate biodiversity and monitoring', *Ecological Indicators* **101**, 963–972.  
**URL:** <https://doi.org/10.1016/j.ecolind.2019.02.008>
- Shannon, C. E. & Weaver, W. (1963), *The Mathematical Theory of Communication*, University of Illinois Press, Urbana. 117 p.
- Shokralla, S., Spall, J. L., Gibson, J. F. & Hajibabaei, M. D. (2012), 'Next-generation sequencing technologies for environmental DNA research', *Molecular Ecology* **21**, 1794–1805.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05538.x>
- Sinniger, F., Pawlowski, J., Harii, S., Gooday, A. J., Yamamoto, H., Chevaldonné, P., Cedhagen, T., Carvalho, G. & Creer, S. (2016), 'Worldwide Analysis of Sedimentary DNA Reveals Major Gaps in Taxonomic Knowledge of Deep-Sea Benthos', *Frontiers in Marine Science* **3**(92).  
**URL:** <https://doi.org/10.3389/fmars.2016.00092>

- Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S. & Bunce, M. (2017), 'Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment', *Scientific reports* **7**(1), 12240–11.  
**URL:** <https://doi.org/10.1038/s41598-017-12501-5>
- Taberlet, P., Bonin, A., Zinger, L. & Coissac, E. (2018), *Environmental DNA: For Biodiversity Research and Monitoring*, Oxford University Press, Incorporated.
- Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L. H. (2012b), 'Environmental DNA', *Molecular Ecology* **21**, 1789–1793.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05542.x>
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. (2012a), 'Towards next-generation biodiversity assessment using DNA metabarcoding', *Molecular Ecology* **21**, 2045–2050.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05470.x>
- Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barradough, T. G. & Fontaneto, D. (2012), 'The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna', *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16208–16212.  
**URL:** <http://www.jstor.org/stable/41763227>
- Thiermann, F., Niemeyer, A. S. & Giere, O. (1996), 'Variations in the sulfide regime and the distribution of macrofauna in an intertidal flat in the North Sea', *Helgolander Meeresunters* **50**(1), 87104.  
**URL:** <https://doi.org/10.1007/BF02367138>
- Todd, P. A., Heery, E. C., Loke, L. H. L., Thurstan, R. H., Kotze, D. J. & Swan, C. (2019), 'Towards an urban marine ecology: characterizing the drivers, patterns and processes of marine ecosystems in coastal cities', *Oikos* **128**, 1215–1242.  
**URL:** <https://doi.org/10.1111/oik.05946>
- Torti, A., Lever, M. A. & Jørgensen, B. B. (2015), 'Origin, dynamics, and implications of extracellular dna pools in marine sediments', *Marine Genomics* **24**(3), 185–196.  
**URL:** <https://doi.org/10.1016/j.margen.2015.08.007>
- Tristan Cordier, Dominik Forster, Y. D. C. I. M. M. T. S. J. P. (2018), 'Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring', *Molecular ecology resources* **18**(6), 1381–1391.  
**URL:** <https://doi.org/10.1111/1755-0998.12926>

- Trøen, T. W. (2018), 'Vedtak til lov om endringer i naturmangfoldloven (hjemmel for gjennomføring av vanndirektivet)', Online.  
**URL:** <https://www.stortinget.no/no/Saker-og-publikasjoner/Vedtak/Beslutninger/Lovvedtak/2018-2019/vedtak-201819-008/>
- Tuholske, C., Halpern, B. S., Blasco, G., Villasenor, J. C., Frazier, M. & Caylor, K. (2021), 'Mapping global inputs and impacts from of human sewage in coastal ecosystems', *PLoS One* **16**(11), e0258898.  
**URL:** <https://www.proquest.com/scholarly-journals/mapping-global-inputs-impacts-human-sewage/docview/2595989895/se-2>
- 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.  
**URL:** <https://doi.org/10.1016/j.biocon.2014.11.017>
- Vann-nett (n.d.), 'Vann-nett: Kviturspollen', Available at: <https://vann-nett.no/portal/#/waterbody/0261010500-4-C>. Accessed: 10 May 2022.
- Vannforskriften (2006), 'Forskrift om rammer for vannforvaltningen'.  
**URL:** <https://lovdata.no/dokument/LTI/forskrift/2006-12-15-1446>
- Vannportalen (2020a), 'Karakterisering', Available at: <https://www.vannportalen.no/kunnskapsgrunnlaget/karakterisering2/>. (Accessed: 10 May 2022).
- Vannportalen (2020b), 'Klassifisering', Available at: <https://www.vannportalen.no/kunnskapsgrunnlaget/klassifisering/>. (Accessed: 10 May 2022).
- Vannportalen (n.d.a), 'Helhetlig vannforvaltning', Available at: <https://www.vannportalen.no/organisering2/helhetlig-vannforvaltning/>, note = (Accessed: 21 October 2022).
- Vannportalen (n.d.b), 'Plandokumenter, planperiode og plansyklus.', Available at: <https://www.vannportalen.no/plansyklus/>. (Accessed: 21 October 2022).
- Vannportalen, ed. (2019), *Høyringsdokument 2: Hovedutfordringer i vassregionen Hordaland*, Vannportalen.  
**URL:** <https://www.vannportalen.no/plansyklus/planperioden-2022-2027/hovedutfordringer-2022-2021/>
- Vannportalen, ed. (n.d.c), *Høyringsdokument 1: Planprogram for regional vassforvaltningsplan 2022-2027 for Sogn og Fjordane vassregion og Hordaland vassregion*, Vannportalen.  
**URL:** <https://www.vannportalen.no/plansyklus/planperioden-2022-2027/planprogram-2022-2027/>

- Wangenstein, O. S., Palacín, C., Guardiola, M. & Turon, X. (2018), 'Dna metabarcoding of littoral hard-bottom communities: High diversity and database gaps revealed by two molecular markers', *PeerJ* **6**, E4705.  
**URL:** [10.7717/peerj.4705](https://doi.org/10.7717/peerj.4705)
- Wassmann, P. (1985), 'Sedimentation of particulate material in nordåsvannet, a hypertrophic, land-locked fjord in western norway', *Marine ecology Progress series (Halstenbek)* **22**(3), 259271.  
**URL:** <https://www.jstor.org/stable/24816011>
- Whittaker, R. H. (1972), 'Evolution and measurement of species diversity', *Taxon* **21**, 213–251.  
**URL:** <http://www.jstor.org/stable/1218190>
- Whittaker, R. H., Willis, K. J. & Field, R. (2001), 'Scale and species richness: towards a general, hierarchical theory of species diversity', *Journal of Biogeography* **28**(4), 453–470.  
**URL:** <https://doi.org/10.1046/j.1365-2699.2001.00563.x>
- Wickham, H. (2016), *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.  
**URL:** <https://ggplot2.tidyverse.org>
- Wu, S., Xiong, J. & Yu, Y. (2015), 'Taxonomic Resolutions Based on 18S rRNA Genes: A Case Study of Subclass Copepoda', *PloS one* **10**(6), e0131498–e0131498.  
**URL:** <https://doi.org/10.1371/journal.pone.0131498>
- Yang, C., Ji, Y., Wang, X., Yang, C. & Yu, D. W. (2013), 'Testing three pipelines for 18s rdna-based metabarcoding of soil faunal diversity', *Science China. Life Sciences* **56**, 73–81.  
**URL:** <https://www.proquest.com/scholarly-journals/testing-three-pipelines-18s-rdna-based/docview/1328385590/se-2>
- Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C. & Ding, Z. (2012), 'Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring', *Methods in Ecology and Evolution* **3**, 613–623.  
**URL:** <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-210X.2012.00198.x>
- Zhan, A., Bailey, S. A., Heath, D. D. & Macisaac, H. J. (2014), 'Performance comparison of genetic markers for high-throughput sequencing-based biodiversity assessment in complex communities', *Molecular Ecology Resources* **14**, 1049–1059.  
**URL:** <https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.12254>

Zinger, L., Bonin, A. & et al., I. G. A. (2019), 'DNA metabarcoding - Need for robust experimental designs to draw sound ecological conclusions', *Molecular Ecology* **28**(8), 1857–1862.

**URL:** <https://doi.org/10.1111/mec.15060>





# Appendix I. Bray-Curtis similarity matrices

## 5.2 Bray-Curtis similarity matrix from morphological data

×

Table 5.1: The Bray-Curtis similarity (0-100) for samples 3A-6B and 9A-9B from the morphological dataset.

		Samples									
		3A	3B	4A	4B	5A	5B	6A	6B	9A	9B
Samples	3A										
	3B	64.6									
	4A	35.6	40.0								
	4B	57.8	54.5	54.9							
	5A	17.5	14.7	35.5	28.1						
	5B	39.3	37.6	68.7	55.8	37.4					
	6A	56.2	50.4	25.1	41.5	15.0	27.0				
	6B	59.7	51.2	29.9	40.7	15.6	29.8	63.5			
	9A	42.0	41.8	22.7	32.0	12.7	22.7	32.4	43.8		
	9B	29.2	32.2	30.8	32.1	21.3	29.9	27.4	29.3	38.7	

## 5.3 Bray-Curtis similarity matrix from COI data

×

Table 5.2: The Bray-Curtis similarity (0-100) for samples 1B-6B and 9C-9D from COI dataset.

		Samples												
Samples		1B	2A	2B	3A	3B	4A	4B	5A	5B	6A	6B	9C	9D
	1B													
	2A	44.7												
	2B	48.5	68.9											
	3A	51.6	40.4	45.1										
	3B	40.3	30.7	33.7	50.5									
	4A	39.3	28.3	30.5	40.5	42.4								
	4B	47.2	34.6	37.2	43.1	45.5	57.1							
	5A	42.4	33.9	37.2	37.8	32.2	38.3	40.3						
	5B	50.5	44.6	45.5	47.3	39.9	45.7	47.8	55.8					
	6A	35.5	29.4	35.1	41.9	43.6	40.7	36.4	32.8	40.1				
	6B	38.8	32.2	35.7	41.1	47.7	39.3	39.1	32.8	37.4	52.7			
	9C	25.1	19.7	23.1	32.7	33.7	21.6	24.3	20.5	26.1	28.3	34.9		
	9D	33.9	28.4	32.3	30.7	32.7	32.4	31.6	36.8	42.2	31.3	35.7	27.5	

## 5.4 Bray-Curtis similarity matrix from 18S (Metazoa) data

×

Table 5.3: The Bray-Curtis similarity (0-100) for samples 1B-6B and 9C-9D from 18S metazoan dataset.

		Samples												
Samples		1B	2A	2B	3A	3B	4A	4B	5A	5B	6A	6B	9C	9D
	1B													
	2A	33.2												
	2B	37.9	63.9											
	3A	47.9	32.8	37.2										
	3B	38.8	24.5	26.1	51.5									
	4A	38.9	31.6	34.3	41.1	36.5								
	4B	39.7	31.6	32.4	43.2	35.5	58.3							
	5A	37.1	30.9	27.9	32.2	22.6	33.7	37.8						
	5B	39.1	36.0	38.6	43.0	33.5	44.3	44.4	51.4					
	6A	36.8	23.9	30.8	45.6	39.9	41.1	38.4	24.9	37.9				
	6B	36.9	26.2	32.3	46.8	45.1	37.8	34.1	24.0	33.0	55.3			
	9C	25.7	22.2	25.6	37.4	38.7	25.0	27.3	17.9	23.2	35.0	37.4		
	9D	38.2	28.2	28.7	31.5	33.1	34.6	35.6	40.2	48.3	30.3	32.3	21.7	

## 5.5 Bray-Curtis similarity matrix from 18S (Eukaryota) data

×

Table 5.4: The Bray-Curtis similarity (0-100) for samples 1B-6B and 9C-9D from 18S Eukaryotic dataset.

		Samples												
		1B	2A	2B	3A	3B	4A	4B	5A	5B	6A	6B	9C	9D
Samples	1B													
	2A	55.7												
	2B	59.2	70.9											
	3A	64.6	54.6	58.8										
	3B	63.7	51.3	56.3	69.7									
	4A	60.4	51.2	55.9	61.8	64.3								
	4B	64.1	51.9	57.5	63.1	65.9	72.9							
	5A	55.9	47.9	51.9	53.4	54.6	59.8	58.7						
	5B	62.6	55.6	61.4	62.2	62.2	64.7	65.7	67.9					
	6A	58.8	47.9	53.9	64.1	65.6	61.0	58.4	55.0	60.5				
	6B	56.3	47.2	52.2	62.8	63.3	56.5	54.9	49.6	56.4	71.0			
	9C	44.9	38.8	42.4	50.7	50.9	46.4	46.8	43.7	49.6	52.5	54.4		
	9D	54.7	45.8	50.5	54.3	56.1	55.5	54.4	55.8	61.1	55.9	53.4	52.9	



## **Appendix II. Taxa lists**

### **5.6 Morphological taxa list**









## **5.7 Unique metazoan taxonomic groups from COI OTU table**

Taxa	1B	2A	2B	3A	3B	4A	4B
Acartia	1806	403	63	1171	248	348	300
Acartia margalefi	12	0	0	0	0	0	0
Acoela (order)	0	0	0	0	0	0	0
Aequorea	125	0	0	0	0	0	0
Amathia sp. n. 1 AW	0	55	0	0	0	0	0
Amphiuridae	0	0	0	0	0	0	0
Annelida	37	8958	2901	2585	203	0	2
Archaphanostoma agile	6	0	0	0	0	0	0
Arthropoda	4040	855	2080	6127	4729	2044	3006
Ascidiella aspersa	20	0	0	0	0	22	2
Aurelia	0	0	3	0	0	0	0
Bougainvillia	168	0	0	0	0	0	0
Bulbamphiascus imus	0	0	138	9	377	1	0
Bursovaginoidea (order)	0	0	0	0	46	0	0
Capitella sp. TD53	314	0	0	0	0	0	0
Centropages hamatus	5	0	0	26	1	0	9
Chaetonotida	0	0	0	0	125	0	15
Chaetozone sp. NTNUVM 7453	0	0	0	1	0	0	0
Chromadorea	17	968	165	62	687	72	11
Chromadorida	3	0	0	76	5	0	0
Chromadoridae sp. 09PROBE	4	0	0	0	23	0	13
Ciona intestinalis	23	5	1	7	1	0	0
Cirriformia	0	0	0	1	44050	9	0
Cladonematidae	0	1382	1626	20	0	0	0
Cletodes tenuipes	0	0	0	1	0	0	0
Clitellata	2	19	88	32	7	7	62
Cnidaria	1047	723	879	533	532	220	351
Demospongiae	56	92	78	54	28	2	16
Dinophilus vorticoides	0	0	0	0	0	0	0
Diplostraca	6	4	5	2	1	0	5
Echinocyamus pusillus	0	0	0	0	0	0	0
Echinodermata	50	73	8	36	22	42	16
Ectinosomatidae	0	0	0	0	0	0	0
Elysia	0	0	0	0	0	0	0
Enhydrosoma	0	0	0	0	0	0	0
Enoplida	0	0	0	9	213	0	0
Eumida	111	0	0	0	0	0	0
Euphausiacea	9	0	0	8	4	9	0
Eurotatoria	62	4	5	83	22	0	20
Gastropoda	0	0	0	0	0	0	0
Glycera alba	0	0	0	0	0	0	0
Gnathostomula armata	0	0	0	0	0	0	0
Harpacticoida sp. 11AlgonqNJ0	0	0	0	0	0	0	0
Hydrozoa	335	174	150	173	66	59	109
Juxtacribrilina	1	0	13	125	16	0	0
Kurtiella bidentata	13	2	19	2096	7770	15	12
Leptosynapta sp. Echin 6667V	0	0	0	10	10	32	11
Leuckartiara octona	1260	0	152	9	0	0	0
Macrochaeta clavicornis	0	0	0	4	0	1628	0
Macrodasysida	3	35	3	0	0	0	12
Macrodasys	25	0	105	114	0	0	0

Taxa	1B	2A	2B	3A	3B	4A	4B
Malacostraca	0	0	0	10	12	0	0
Melicertum octocostatum	1381	0	0	0	0	0	0
Metazoa	7136	7724	10033	14579	9382	11827	10609
Microtopopus maculatus	0	220	0	0	0	0	0
Mollusca	201	307	189	118	161	62	75
Montacutidae	0	0	0	0	122	0	0
Nematoda	1161	2708	4587	586	596	220	54
Nemertea	0	0	0	0	14	0	6
Nereimyra punctata	0	0	0	0	0	0	0
Nicolea zostericola	23	636	40	0	0	0	0
Normanella obscura	21	0	108	0	0	0	0
Obelia dichotoma	0	0	26	6	0	0	0
Oerstedtia	0	8635	0	0	0	0	0
Parougia eliasoni	0	0	0	35	0	0	0
Phallusia	0	0	0	0	4	70	2
Philactinoposthia rhamnifera	0	0	0	0	0	0	0
Philactinoposthia sp. UJ	0	0	0	51	0	0	0
Pholoe	0	0	0	0	0	0	0
Pholoe baltica	9	1	0	0	0	0	5
Phoronis hippocrepia	0	0	0	0	0	0	0
Placozoa	155	218	184	146	58	231	284
Placozoon sp.	14	26	37	0	19	0	0
Platyhelminthes	0	0	37	9	0	20	7
Polychaeta	360	43	33	407	417	17	1
Polychaeta sp. 09PROBE (famil	0	0	0	0	0	0	0
Polyplacotoma	5	0	0	34	0	0	7
Porifera	149	29	24	126	70	168	142
Priapulid caudatus	0	0	0	0	0	0	0
Prionospio sp. BOLD	0	0	0	0	0	0	0
Proameira sp. n. 2 SR	0	0	0	0	0	0	0
Rathkea octopunctata	26	0	0	2	0	0	0
Rhabditophora	0	0	0	510	0	0	0
Rotifera	0	0	0	3	15	91	0
Sabellida	0	31	9	0	0	0	0
Sarsia tubulosa	0	36	97	1	26	0	0
Terschellingia	1	0	0	13	0	0	0
Tubificoides benedii	1	11801	22336	3	93703	9	0
Xenacoelomorpha	0	0	0	0	0	0	0

Taxa	5A	5B	6A	6B	9C	9D
Acartia	101	649	288	1283	0	0
Acartia margalefi	0	23	0	0	0	0
Acoela (order)	0	0	0	0	53	0
Aequorea	0	0	0	0	0	0
Amathia sp. n. 1 AW	0	0	0	0	0	0
Amphiuridae	0	0	0	62	0	0
Annelida	0	1	235167	86086	9513	4718
Archaphanostoma agile	0	0	0	0	38	33
Arthropoda	1766	2178	1236	2353	16561	11531
Ascidiella aspersa	7	7	2	0	0	3
Aurelia	15	25	1	26	0	70
Bougainvillia	0	0	0	0	0	0
Bulbamphiascus imus	0	551	26	28	3	100
Bursovaginoidea (order)	0	0	3	110	488	0
Capitella sp. TD53	25322	0	25	0	0	0
Centropages hamatus	0	0	0	0	0	0
Chaetonotida	0	0	0	15	30	9
Chaetozone sp. NTNUVM 7453	0	0	0	0	7102	0
Chromadorea	150	194	321	767	628	15
Chromadorida	17	4	0	11	499	0
Chromadoridae sp. 09PROBE	0	0	0	0	35	5
Ciona intestinalis	3	2	6	1	0	0
Cirriformia	0	0	103	341	0	1
Cladonematidae	0	0	0	0	0	0
Cletodes tenuipes	0	0	0	0	559	0
Clitellata	24	133	6	883	49	127
Cnidaria	243	400	177	287	562	204
Demospongiae	10	41	8	19	95	21
Dinophilus vorticoides	0	0	0	0	7345	0
Diplostraca	132	12	0	15	0	0
Echinocyamus pusillus	0	0	1	33	0	0
Echinodermata	0	48	4	12	6	19
Ectinosomatidae	0	0	0	0	59	0
Elysia	0	0	108	0	0	0
Enhydrosoma	33	0	0	0	0	0
Enoplida	0	0	13	0	55	0
Eumida	0	0	0	0	0	0
Euphausiacea	12	0	4	0	0	0
Eurotatoria	1	6	3	19	29	0
Gastropoda	0	0	0	0	48	0
Glycera alba	0	0	0	0	67	0
Gnathostomula armata	422	12	0	0	0	0
Harpacticoida sp. 11AlgonqNJ0	0	0	1	0	62	0
Hydrozoa	17	44	50	81	277	34
Juxtacribrilina	0	0	0	0	204	0
Kurtiella bidentata	9	23	44	20	37	10
Leptosynapta sp. Echin 6667V	0	54	0	11	43	0
Leuckartiara octona	0	0	41	0	9	0
Macrochaeta clavicornis	0	0	0	0	26295	2
Macrodasysida	0	32	0	0	0	0
Macrodasys	2206	0	19	0	642	7

Taxa	5A	5B	6A	6B	9C	9D
Malacostraca	39	0	4	0	0	0
Melicertum octocostatum	0	0	0	0	0	0
Metazoa	5264	5707	3538	6861	19083	3926
Microtopopus maculatus	0	0	0	0	0	0
Mollusca	56	93	2988	37	253	5
Montacutidae	0	0	0	0	0	0
Nematoda	1053	1776	590	4068	184	2706
Nemertea	0	10	654	0	0	0
Nereimyra punctata	0	0	0	0	1176	0
Nicolea zostericola	0	0	0	0	0	0
Normanella obscura	272	0	0	0	0	0
Obelia dichotoma	0	0	0	0	0	0
Oerstedtia	0	0	0	0	0	0
Parougia eliasoni	0	0	0	0	0	0
Phallusia	79	0	0	26	0	48
Philactinoposthia rhamnifera	0	0	0	0	47	0
Philactinoposthia sp. UJ	0	0	0	0	30	0
Pholoe	0	0	0	0	146	0
Pholoe baltica	0	0	49715	2	130	0
Phoronis hippocrepia	0	0	0	0	429	0
Placozoa	8	49	25	13	0	6
Placozoon sp.	14	38	6	2	0	17
Platyhelminthes	26	0	15	28	821	6
Polychaeta	0	130	33	73886	5024	20
Polychaeta sp. 09PROBE (famil	0	0	0	0	91	0
Polyplacotoma	0	2	3	0	0	0
Porifera	128	72	8	19	122	25
Priapulus caudatus	0	0	0	0	0	169
Prionospio sp. BOLD	0	0	0	0	0	441
Proameira sp. n. 2 SR	0	0	0	0	208	0
Rathkea octopunctata	0	0	0	0	3	0
Rhabditophora	16	0	0	0	39	26
Rotifera	0	0	57	10	37	149
Sabellida	0	0	0	0	0	0
Sarsia tubulosa	0	0	2	0	629	0
Terschellingia	0	0	136	3	66	0
Tubificoides benedii	0	1	25784	34612	6	168636
Xenacoelomorpha	0	0	0	0	149	0

## **5.8 Unique metazoan taxonomic groups from 18S OTU table**

Taxa	1B	2A	2B	3A	3B	4A	4B
Abatus cavernosus	0	0	0	0	0	0	0
Abra nitida	0	0	52	0	0	0	0
Actinaria	0	0	0	0	0	0	0
Ameira scotti	0	0	0	40	41	0	0
Annelida	0	0	0	0	0	0	0
Anthoathecata	1313	0	147	4	0	0	0
Araeolaimida	0	0	15	76	318	0	0
Archaphanostoma agile	52	0	0	0	0	0	0
Arthropoda	0	0	0	38	7	0	0
Asciidiidae	1084	465	320	230	197	180	109
Aspidiophorus	48	0	0	0	0	0	0
Aspidiophorus tentaculatus	569	0	0	0	40	0	0
Aurelia	0	0	11	0	0	0	0
Axonolaimidae	75	59	55	465	49	100	146
Bathylaimus	93	0	2	4580	26716	1	0
Bivalvulida	3	0	0	2	1	29	8
Botryllus	0	0	49	13	0	0	0
Bythoceratina hanejiensis	0	0	0	0	0	0	0
Caenogastropoda	0	4	13	0	28	6	2
Calanoida	0	0	0	0	0	0	0
Calcarea	0	0	0	0	84	0	0
Calomicrolaimus parahonestus	1046	518	166	3539	1498	1411	615
Capitellida	990	14	22	1262	903	39	7
Cauleriella	0	0	0	0	0	0	0
Centropages hamatus	3358	1	124	2167	117	711	8639
Cephalothrix filiformis	0	0	10	0	0	9	33
Cephalothrix rufifrons	0	18	24	0	15	17	12
Chaetonotida	553	149	390	434	0	0	1290
Chaetonotidae	266	113	0	0	135	14	1103
Chromadorea	0	50	163	515	6	0	0
Chromadorida	15879	7051	9966	10397	5718	2713	2986
Chromadoridae	847	0	0	18	178	0	0
Chromadorita	2955	0	0	3910	558	0	0
Cicerinidae	0	0	0	0	51	0	0
Cirratulidae	3	0	0	10	5157	1	0
Cladonema californicum	0	484	275	23	0	0	0
Corella inflata	50	0	20	2	0	0	5
Cyatholaimidae	0	0	0	32	0	0	0
Cyclopoida	0	20	0	0	7	0	0
Cythere lutea	0	0	0	0	0	0	0
Cytheromorpha acupunctata	7972	2	0	4898	14810	4244	11756
Desmodora ovigera	0	0	0	220	806	0	44
Desmodorida	630	1429	1317	154	812	92	136
Desmodoridae	527	0	0	2	326	0	0
Desmoscolecida	0	0	0	164	194	212	0
Desmoscolecidae	0	0	0	0	0	0	0
Enoplea	0	0	0	285	2859	0	0
Enoplida	0	2615	11772	27	382	215	0
Enoplodes	0	0	0	14	738	0	0
Enoplolaimus	0	360	613	0	0	0	0
Eteone	0	0	1	0	0	0	0

Taxa	1B	2A	2B	3A	3B	4A	4B
Eunicida	0	0	0	61	0	0	0
Euthyneura family incertae sed	0	0	0	0	0	0	0
Fabriciola	2	24	3	26	17	5	25
Harpacticidae	0	0	0	0	0	0	0
Harpacticoida	163	87	128	135	54	70	41
Heterobranchia	0	0	0	0	0	0	0
Howeina	0	0	0	0	12	0	141
Hydrozoa	182	0	0	23	117	0	7
Kuma	13	0	0	6	0	0	0
Leptocythere lacertosa	0	866	5	6	0	0	7783
Leptostraca	0	0	0	0	0	0	0
Leptosynapta	0	33	26	21	0	16	278
Linhomoeidae	130	24	24	118	422	0	0
Litinium	0	0	0	0	0	0	0
Macrodasys	3	0	15	18	0	0	0
Maxillopoda	5	0	57	21	0	0	0
Metachromadora	0	0	0	0	0	0	0
Metazoa	1437	2367	2578	462	162	167	133
Microstomidae	0	0	40	0	0	0	0
Molgolaimus demani	2625	3022	6238	200	25	0	10
Molgula	185	0	24	21	478	238	381
Molgulidae	14	32	0	0	0	0	0
Monhysterida	785	733	1436	4811	716	21408	3833
Monostilifera	0	2492	0	0	0	0	0
Montacutidae	0	0	0	346	8775	3	0
Mopaliidae	12	0	47	52	0	0	0
Mugilogobius platynotus (flat-l	6	52	37	7	9	0	0
Mytilus edulis	0	32	0	0	0	10	0
Nicolea	0	141	13	0	0	0	0
Nudora ilhabelae	0	0	15	2	0	0	0
Obelia	9	22	80	0	9	0	0
Odontorhynchus aculeatus	0	0	0	226	179	0	0
Oikopleuridae	0	37	0	0	0	20	0
Oxystomina	0	0	0	112	0	0	0
Paracanthonchus	0	0	0	65	0	0	0
Paramphiascella fulvofasciata	5	120	337	0	0	44	0
Pareurystomina	0	0	0	0	1500	0	0
Phallusia nigra	11	2	19	11	9	89	45
Pholoe	0	0	0	0	3	19	0
Phoroniformea	0	0	0	0	4	0	0
Phyllococida	0	2674	966	0	1	0	0
Placobranchidae	0	0	0	0	0	0	0
Placorhynchidae	42	0	0	1605	0	0	39
Platyhelminthes	0	0	0	0	64	0	0
Ploimida	45	0	25	85	61	119	19
Podocopida	286	174	0	5280	11529	3	348
Polychaeta	0	34	7	0	0	0	1
Polycope	0	174	0	20	101	0	5442
Polycystididae	80	0	0	35	12069	1	0
Pontocypris mytiloides	0	0	0	1	0	0	0
Pontonema vulgare	2	0	1	0	0	0	0



Taxa	1B	2A	2B	3A	3B	4A	4B
Porifera	0	0	0	0	0	53	1
Protodorvillea	0	0	0	0	0	0	0
Rhabdozoela	72	64	4	12985	1425	729	1077
Sagittozoela	0	0	0	0	0	0	0
Sarsia lovenii	0	0	13	0	0	0	0
Scoloplos armiger	7	73	25	328	276	0	0
Semicytherura striata	790	0	0	2203	5763	2	10
Spionida	0	12	0	0	1	0	0
Spionidae	0	0	0	0	0	0	0
Spirinia parasitifera	0	305	499	3029	3031	1251	911
Styelidae	0	43	0	0	0	0	0
Syllidae	0	827	1073	2759	32	0	0
Syllidia	0	406	57	0	0	0	0
Symplocostoma	0	299	0	0	0	0	0
Syringolaimus	0	0	0	0	0	0	0
Temora longicornis	0	0	0	0	0	0	0
Terebellidae	0	0	0	0	0	0	0
Thalassoalaimus	0	0	0	41	0	0	0
Thaumastoderma ramuliferum	0	0	0	54	0	0	0
Theristus	0	22	273	0	0	0	0
Tonicella lineata	35	19	12	99	0	0	0
Trigonostomum venenosum	0	32	81	14	0	0	0
Tripyloides	186	4585	7	1283	0	0	1
Tubifex tubifex (sludge worm)	0	226	213	0	790	0	0
Uncinocythere occidentalis	0	0	0	0	0	0	0
unclassified Oncholaimidae	0	83	415	236	23	3	517
unclassified Thoracostomopsid	0	0	0	0	3186	1	0
Xyalidae	0	0	0	185	2	0	0
Calomicrolaimus parahonestus	0	0	0	36	33	0	0

Taxa	5A	5B	6A	6B	9C	9D
Abatus cavernosus	0	0	0	0	100	0
Abra nitida	17	11	0	0	0	0
Actinaria	0	0	0	0	89	0
Ameira scotti	0	0	0	0	2	0
Annelida	680	0	0	0	0	0
Anthoathecata	0	0	12	0	0	0
Araeolaimida	0	0	87	0	236	0
Archaphanostoma agile	0	0	0	0	131	305
Arthropoda	0	0	448	0	160	0
Asciidiidae	488	122	134	60	110	98
Aspidiophorus	0	0	31	0	0	0
Aspidiophorus tentaculatus	0	0	0	1	147	5
Aurelia	10	6	0	0	0	12
Axonolaimidae	0	206	229	38	21	122
Bathylaimus	0	0	1	33581	4	936
Bivalvulida	0	3	4	0	0	0
Botryllus	0	0	0	0	0	0
Bythoceratina hanejiensis	0	0	0	0	190	0
Caenogastropoda	0	0	10	0	2975	0
Calanoida	18	0	0	0	24	7
Calcarea	0	0	3	0	10	0
Calomicrolaimus parahonestus	0	106	435	40	1648	0
Capitellida	206	279	119	634	858	16
Cauteriella	0	0	0	0	69	0
Centropages hamatus	59	15	16224	1312	0	0
Cephalothrix filiformis	38	2	0	0	0	0
Cephalothrix rufifrons	0	0	410	0	0	18
Chaetonotida	108	0	0	6	25	12
Chaetonotidae	0	0	0	5	19	242
Chromadorea	0	36	37	0	0	0
Chromadorida	7343	17784	4426	3302	24185	2883
Chromadoridae	0	0	39	0	710	1
Chromadorita	0	0	1498	954	61	0
Cicerinidae	0	0	0	0	15	0
Cirratulidae	0	0	40	72	0	2
Cladonema californicum	0	0	0	0	0	0
Corella inflata	0	0	0	2	70	0
Cyatholaimidae	0	0	8	0	0	0
Cyclopoida	0	0	0	6	225	0
Cythere lutea	0	0	0	0	0	180
Cytheromorpha acupunctata	2898	0	5896	2125	2371	259
Desmodora ovigera	0	0	0	0	291	0
Desmodorida	6950	5293	221	491	436	17776
Desmodoridae	0	0	0	0	869	0
Desmoscolecida	0	7	314	0	0	0
Desmoscolecidae	0	0	0	0	86	0
Enoplea	0	347	7	0	8	0
Enoplida	0	2	170	6976	165	18
Enoplodes	0	0	0	0	24	0
Enoplolaimus	0	0	0	0	0	0
Eteone	0	0	1	15352	7039	0

Taxa	5A	5B	6A	6B	9C	9D
Eunicida	0	0	0	0	0	0
Euthyneura family incertae sed	0	0	830	0	0	0
Fabriciola	0	0	0	0	0	0
Harpacticidae	0	0	0	0	1642	4
Harpacticoida	71	85	385	239	463	441
Heterobranchia	0	0	100	0	0	0
Howeina	0	0	611	185	0	0
Hydrozoa	0	14	0	4	0	148
Kuma	0	0	0	28	5	0
Leptocythere lacertosa	0	0	0	0	3853	1
Leptostraca	0	0	0	0	0	1694
Leptosynapta	37	249	4	0	47	0
Linhomoeidae	615	2	1311	272	151	639
Litinium	0	0	0	0	676	0
Macrodasys	679	0	15	0	221	0
Maxillopoda	0	180	113	223	1894	87
Metachromadora	0	0	0	0	3778	0
Metazoa	358	353	156	1791	123	42
Microstomidae	0	0	0	0	0	0
Molgolaimus demani	359	561	1	0	0	760
Molgula	0	0	5	5	0	0
Molgulidae	0	0	0	0	11	0
Monhysterida	840	2032	8617	1493	2395	3109
Monostilifera	0	0	0	0	0	0
Montacutidae	0	0	0	52	0	0
Mopaliidae	0	2	0	0	0	0
Mugilogobius platynotus (flat-l	8	24	0	0	0	29
Mytilus edulis	0	0	0	0	0	0
Nicolea	0	0	0	0	0	0
Nudora ilhabelae	0	0	0	0	25	0
Obelia	0	0	0	0	0	0
Odontorhynchus aculeatus	0	0	5175	0	5	0
Oikopleuridae	0	0	0	0	0	0
Oxystomina	0	0	0	0	0	0
Paracanthonchus	0	0	10	0	0	0
Paramphiascella fulvofasciata	0	375	464	0	617	31
Pareurystomina	0	0	0	0	331	0
Phallusia nigra	0	10	33	18	0	22
Pholoe	0	0	1000	0	103	0
Phoroniformea	0	0	0	0	135	0
Phyllodocida	0	0	2	1	0	0
Placobranchidae	0	0	305	0	0	0
Placorhynchidae	0	0	0	0	0	0
Platyhelminthes	0	0	0	0	0	0
Ploimida	28	119	129	180	49	83
Podocopida	11	128	325	5482	1173	1
Polychaeta	0	0	0	0	0	0
Polycope	0	0	0	0	0	0
Polycystididae	0	0	0	0	42	2
Pontocypris mytiloides	0	0	0	0	6116	0
Pontonema vulgare	0	0	330	9910	0	0

Taxa	5A	5B	6A	6B	9C	9D
Porifera	6	0	0	0	0	0
Protodorvillea	0	0	40	116	0	0
Rhabdozoa	52	3	1	7105	1163	1
Sagittozoa	0	0	0	0	1818	0
Sarsia lovenii	0	0	3	0	117	0
Scoloplos armiger	0	10	17	3	53	0
Semicytherura striata	0	4435	127	9001	46	11141
Spionida	0	0	0	0	0	4563
Spionidae	0	0	0	0	176	0
Spirinia parasitifera	176	3104	4673	2367	20652	5399
Styelidae	0	0	0	0	0	0
Syllidae	0	0	0	7	3932	0
Syllidia	0	0	0	43	0	0
Symplocostoma	0	0	0	0	0	0
Syringolaimus	0	0	0	0	95	0
Temora longicornis	0	0	751	42	0	0
Terebellidae	0	0	0	0	41	0
Thalassoalaimus	0	0	0	0	0	0
Thaumastoderma ramuliferum	0	0	0	0	11	0
Theristus	0	0	0	0	12	0
Tonicella lineata	0	0	0	5	0	0
Trigonostomum venenosum	0	0	0	0	0	0
Tripyloides	1090	0	5	506	49	0
Tubifex tubifex (sludge worm)	0	0	2	35	0	11430
Uncinocythere occidentalis	0	0	162	1329	1087	44
unclassified Oncholaimidae	4	0	116	27	312	0
unclassified Thoracostomopsid	0	0	0	0	2	0
Xyalidae	33	131	0	0	288	0
Calomicrolaimus parahonestus	0	0	0	0	5	0