



# Advancing molecular macrobenthos biodiversity monitoring: a comparison between Oxford Nanopore and Illumina based metabarcoding and metagenomics

Karlijn Doorenspleet<sup>1,\*</sup>, Amalia Aikaterini Mailli<sup>2,\*</sup>, Berry B. van der Hoorn<sup>3</sup>, Kevin K. Beentjes<sup>3</sup>, Annelies De Backer<sup>4</sup>, Sofie Derycke<sup>4</sup>, Albertinka J. Murk<sup>1</sup>, Henning Reiss<sup>2</sup> and Reindert Nijland<sup>1</sup>

<sup>1</sup> Marine Animal Ecology, Wageningen University and Research, Wageningen, Netherlands

<sup>2</sup> Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway

<sup>3</sup> Naturalis Biodiversity Center, Leiden, Netherlands

<sup>4</sup> ILVO Marine Research, Flanders Research Institute for Agriculture, Fisheries and Food, Oostende, Belgium

\* These authors contributed equally to this work.

## ABSTRACT

DNA-based methods and developments of sequencing technologies are integral to macrobenthos biodiversity studies, and their implementation as standardized monitoring methods is approaching. Evaluating the efficacy and reliability of these technological developments is crucial for macrobenthos biodiversity assessments. In this study, we compared three DNA-based techniques for assessing the diversity of bulk macrobenthos samples from the Belgian North Sea. Specifically, we compared amplicon sequencing using Illumina MiSeq and portable real-time sequencing of Oxford Nanopore *versus* shotgun sequencing using Illumina NovaSeq sequencing. The 313 bp mitochondrial cytochrome c oxidase subunit I (COI) metabarcoding fragment served as the target region for the metabarcoding analysis. Our results indicate that Oxford Nanopore and MiSeq metabarcoding had similar performances in terms of alpha and beta diversity, revealing highly similar location-specific community compositions. The NovaSeq metagenomics method also resulted in similar alpha diversity, but slightly different community compositions compared to the metabarcoding approach. Despite these differences, location-specific community compositions were maintained across all platforms. Notably, read counts from the NovaSeq metagenomic analysis showed the weakest correlation to size corrected morphological abundance and there were mismatches between morphological identification and all DNA based findings which are likely caused by a combination of factors such as primer efficiency and an incomplete reference database. Our findings underscore the critical importance of database completeness prior to implementing DNA-based techniques as standardized monitoring method, especially for metagenomics. Nevertheless, our findings emphasize that Oxford Nanopore metabarcoding proves to be a viable alternative to the conventional Illumina MiSeq metabarcoding platform for macrobenthos biodiversity monitoring.

Submitted 13 June 2024

Accepted 21 February 2025

Published 14 April 2025

Corresponding author  
Reindert Nijland,  
reindert.nijland@wur.nl

Academic editor  
Federica Costantini

Additional Information and  
Declarations can be found on  
page 21

DOI 10.7717/peerj.19158

© Copyright

2025 Doorenspleet et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Biodiversity, Genetics, Genomics, Marine Biology, Molecular Biology

**Keywords** North Sea, Benthic communities, DNA, Metabarcoding, Metagenomics, Nanopore, NovaSeq, Illumina, Miseq

**How to cite this article** Doorenspleet K, Mailli AA, van der Hoorn BB, Beentjes KK, De Backer A, Derycke S, Murk AJ, Reiss H, Nijland R. 2025. Advancing molecular macrobenthos biodiversity monitoring: a comparison between Oxford Nanopore and Illumina based metabarcoding and metagenomics. *PeerJ* 13:e19158 <http://doi.org/10.7717/peerj.19158>

## INTRODUCTION

The European Union established the Marine Strategy Framework Directive (MSFD, 2008/56/EC), alongside the European Water Framework Directive (WFD, 2000/60/EC) and the European network for protected areas for conserving the most valuable species and habitats (Natura 2000 network, Habitat Directive). These directives form the basis of ecosystem management and use a variety of monitoring-based assessments (e.g., *Borja et al., 2010*) that aim to inform (non-) governmental decisions on marine environmental health. However, taxonomy-based monitoring practices rely on great taxonomic expertise, are time-consuming and limited in taxonomic resolution (*Danovaro et al., 2016; Pawlowski et al., 2018*), particularly when identifying understudied taxonomic groups and species at different life stages (*von Ammon et al., 2018; Gleason, Hanner & Cottenie, 2023*). Consequently, DNA-based technological developments have gained particular interest for their potential in studying marine biodiversity (*Bucklin, Steinke & Blanco-Bercial, 2011; Cordier et al., 2019; Elbrecht et al., 2017; Lanzén et al., 2017; Leray & Knowlton, 2016*). A reduction in labour time and the growing reference sequence libraries demonstrate DNA based methods to be a promising alternative for monitoring (*Baird & Hajibabaei, 2012; DeSalle & Goldstein, 2019; Gostel & Kress, 2022; Leray & Knowlton, 2015*). Different DNA-based monitoring studies have demonstrated correspondence with morphological findings (*Derycke et al., 2021; Mauffrey et al., 2021*). As a result, DNA-based methods are repeatedly proposed as a suitable routine biodiversity assessment strategy to inform policy (*Aylagas et al., 2020; Goodwin et al., 2017; Hering et al., 2018; Hinz et al., 2022; Pawlowski et al., 2018; Thalinger et al., 2019*).

Generally, DNA-based biodiversity assessments rely on DNA metabarcoding, which enables the identification of species from environmental samples by amplifying a short DNA fragment using universal primer pairs. This amplified DNA is then sequenced with next-generation sequencing platforms to identify the taxa found in samples (*Taberlet et al., 2012*). As biases can be introduced at each stage (*Van der Loos & Nijland, 2021*), decisions need to be made on the sampling method and the appropriate use of preservation techniques (*Gaither et al., 2011; Ransome et al., 2017*), DNA extraction methods, using sufficient replicates (*Van den Bulcke et al., 2021*), and using appropriate primer pair(s) (*Creer et al., 2016; Devloo-Delva et al., 2018; Leray & Knowlton, 2016*). Bioinformatics pipelines are also crucial to take into consideration as the pipeline choice, and the processing settings greatly influence the output (*Antich et al., 2021*).

Current developments in DNA-based methods have resulted in the availability of several different third-generation sequencing platforms, such as Illumina, Ion Torrent, Oxford Nanopore sequencing and Pacific Biosciences (*Hu et al., 2021*). These platforms provide exciting opportunities to study the environment in convenient ways that include obtaining abundance data (*Klunder et al., 2022*), epigenetic modifications (*Zhao, Van Bodegom & Trimpos, 2023*) and population genetics (*Jahnke et al., 2022*). For metabarcoding of gene fragments shorter than 500 bp, Illumina technologies are currently the standard platform because of its high accuracy (*Meyer & Kircher, 2010*). In comparison to Illumina MiSeq, the Oxford Nanopore sequencing platform measures an electrical current that is produced

when the nucleotides of a sequence pass through a transmembrane nanopore, allowing for real-time sequencing and base calling (Bleidorn, 2016; Wang et al., 2021). The advantages of Oxford Nanopore sequencing include lower costs, the sequencing of long fragments and its suitability for real-time in-field experiments (Krehenwinkel, Pomerantz & Prost, 2019). However, the error rates of raw Oxford Nanopore sequences are currently higher (87–99% accuracy) compared to Illumina platforms, therefore different bioinformatics processing pipelines have been used to circumvent this problem (Baloğlu et al., 2021; Doorenspleet et al., 2025; Egeter et al., 2020). Short read Oxford Nanopore sequencing has previously been shown to be consistent with Illumina MiSeq in low-diversity samples (Egeter et al., 2020; Van Der Reis et al., 2023). However, comparisons of the sequencing platforms with high diversity samples were only performed using bacterial communities, and show a lack of several taxa (Heikema et al., 2020; Stevens et al., 2023). Thus, it remains unclear to what extent short read Oxford Nanopore sequencing is directly comparable to Illumina data.

Recently, metagenomics sequencing have gained interest as alternatives to metabarcoding for community analysis (Bernatchez et al., 2024; Theissinger et al., 2023). Shotgun sequencing can bypass some methodological disadvantages that are inherent to metabarcoding such as PCR (Zhou et al., 2013) and primer amplification bias (Leray & Knowlton, 2015), because DNA is directly processed for sequencing. Moreover, shotgun sequencing has been suggested to cover the full spectrum of biota in a sample and provide a correlation with morphological biodiversity studies (Monchamp et al., 2022). This method is seen as a viable contestant to metabarcoding methods to monitoring genetic diversity (Bista et al., 2018; Lopez et al., 2022; Monchamp et al., 2022). However, metagenomics can be limited by (i) the sequencing depth, which has become more cost-efficient with the advent of current sequencing platforms such as NovaSeq, and (ii) the availability and quality state of reference sequences within widely used databases (Weigand et al., 2019), given that the shotgun sequencing process is non-selective. Thus, it is inconclusive whether shotgun sequencing is currently more useful for both diversity detection and relative abundance data for macrobenthos studies.

In this study, we compared different DNA-based methods: paired end Illumina MiSeq metabarcoding, Oxford Nanopore MinION metabarcoding of the 313 bp cytochrome c oxidase subunit I (COI) mitochondrial fragment and Illumina NovaSeq Metagenomics shotgun sequencing. We used bulk macrobenthos community samples collected from different soft-bottom habitats along the Belgium North Sea. By using different DNA-based methods and sequencing platforms for metabarcoding, we assessed the suitability of these methods for monitoring benthic community composition and diversity. For the metabarcoding methods, we used a curated COI North Sea invertebrate reference database and compared it to a metagenomics method using the complete NCBI-nt database. We hypothesized that the metabarcoding data from both sequencing platforms and the shotgun metagenomics method are generally robust and resemble the morphologically identified community both in alpha and beta diversity.

## MATERIALS AND METHODS

### Sample collection

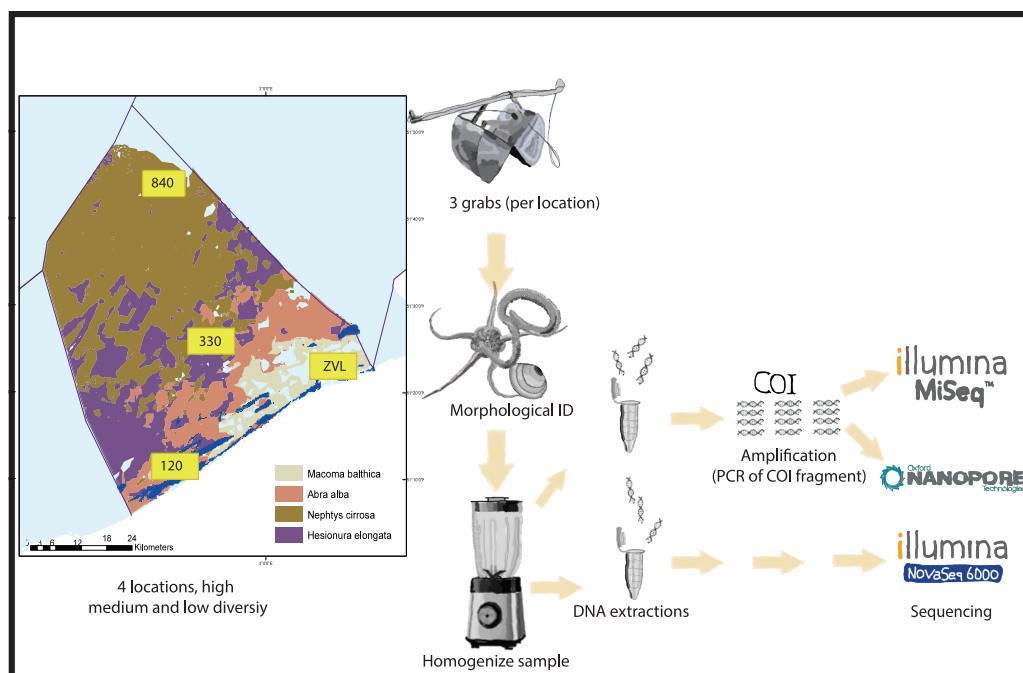
Sampling was conducted at four locations in the Belgian North Sea that contained different macrobenthos communities with high, medium, and low diversity (see [Breine et al., 2018](#)) ([Fig. 1](#)). The bulk samples have been previously used to optimize the metabarcoding protocols and to test the method reproducibility ([Derycke et al., 2021](#); [Van den Bulcke et al., 2021](#); [Van den Bulcke et al., 2023](#)). Samples were collected from a coastal muddy fine-sand habitat with a high taxa diversity of sessile tube-forming organisms and high bioturbation (location 120—*Abra alba* community). Samples were also collected from a medium sand habitat with a medium taxa diversity of mobile organisms (location 330—*Nephtys cirrosa* community), a coarse sand habitat with high taxa diversity for sessile interstitial species (location 840—*Hesionura elongata* community) and lastly, a muddy habitat with low taxa diversity (location ZVL—*Macoma balthica* community) ([Breine et al., 2018](#)). A Van Veen grab (with a sampling surface area of 0.1 m<sup>2</sup>) was used to collect three biological replicates per location (A, B, C). On average, each sample consisted of 9.3 Liters of material. All sediment samples were sieved using a one mm sieve, and the remaining material (for example, shells and rocks) was fixed using absolute ethanol and stored at −20 °C prior to further processing.

### Morphological identification

The morphological identification followed the protocols described by [Derycke et al. \(2021\)](#) and [Van den Bulcke et al. \(2021\)](#); [Van den Bulcke et al. \(2023\)](#). Organisms from one replicate per location (120-B, 840-C, 330-C, ZVL-A) were identified to species level and juveniles to genus level, except for specimens belonging to Nemertea, Anthozoa and Oligochaeta, which were identified up to phylum, subphylum, and order level, respectively. The complete list of species identified in each location is available ([Table S1](#)). Species were recorded per individual hence no biomass information was available for this dataset. To correct for the lack of biomass data, the count data were multiplied by the average size (from each size class). This was obtained to correlate morphology abundance data and the read count of each DNA based method for the identified species.

### DNA extraction

For molecular comparison, all specimens isolated from each field replicate were retained to obtain a bulk sample. Bulk samples were homogenized with a blender or, if the sample was <100 ml, with a mortar and pestle. Subsamples of six ml were taken and stored in Eppendorf tubes at −20 °C before processing in different institutes ([Table S2](#)). Samples used for Illumina MiSeq and Oxford Nanopore sequencing were extracted at Naturalis Biodiversity Centre (Leiden, The Netherlands) and used by Wageningen University and research (Wageningen, The Netherlands) for amplification and sequencing with Oxford Nanopore. Samples used for Illumina NovaSeq were processed at Nord University (Bodø, Norway). For every location, three subsamples of two ml were used for DNA extraction. For location 840C, ZVLA and ZVLC, DNA extract from ILVO was used ([Table S2](#), Oostende, Belgium) due to low recovery of sample material. All institutes (Nord, Naturalis and ILVO)



**Figure 1** **Experimental set-up.** Graphical presentation of methods used in this study. Samples were taken in the Belgian part of the North Sea with known macrobenthos communities (see [Breine et al., 2018](#) for details). The samples were taken in triplicates using a van Veen grab on 4 different locations with high medium and low biodiversity. The animals found in the samples were either morphologically identified or homogenized for DNA extraction. For both amplicon based and metagenomics, different subsamples were used for DNA extraction. Depending on DNA based method, either a COI 313 bp fragment was amplified (Nanopore or Illumina MiSeq) or DNA was directly sequenced with Illumina Novaseq shotgun sequencing for metagenomics.

Full-size DOI: [10.7717/peerj.19158/fig-1](https://doi.org/10.7717/peerj.19158/fig-1)

used the same extraction protocol according to [Van den Bulcke et al. \(2023\)](#). In short, the Eppendorf tubes were centrifuged for 3 min at 10,000 RPM, and the supernatant was removed. Samples were incubated at 50 °C for 1 h to evaporate the remaining ethanol. Three subsamples from each biological replicate (3X3) were incubated overnight at 56 °C in the power-beads tube of the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany), supplemented with 10 µL of proteinase K (20 mg/ml).

The DNA extracted from each subsample was pooled and cleaned using the Wizard DNA clean-up system (Promega, Madison, WI, USA) and eluted in 50 µL TE buffer. After processing, samples were stored at −20 °C before amplification or shotgun sequencing ([Fig. 1](#)).

### PCR amplification

For amplification, a 313 bp COI fragment was amplified using the forward mlCOIintF primer ([Leray et al., 2013](#)) with the reverse jgHCO2198 primer ([Geller et al., 2013](#)) in combination with Oxford Nanopore and Illumina MiSeq extension sequence (printed red in [Table 1](#)). However, a modified version of the primers was used by replacing deoxy-inosines (I) with degenerated bases (N) ([Table 1](#)). Amplification was performed

**Table 1 Primers used for metabarcoding.** Primers used for the Illumina MiSeq and Nanopore metabarcoding method. Sequences indicated in bold are the platform specific adapters to the primers.

Oxford nanopore		
Forward	miCOLintF: 5'- <b>TTTCTGTGGTGTGATATTGCGGWACWGGWTGAA</b> CWGTWTAYCCYCC-3'	Leray et al. (2013)
reverse	hgHCO2198:5'- <b>ACTTGCCTGTCGCTCTATCTTCTANACYTCNGGRTGNCCRA</b> ARAYCA-3'	Geller et al. (2013)
Illumina Miseq		
Forward	miCOLintF: 5'- <b>TCGTCGGCAGCGTCAGATGTGTATAAAGAGACAGGGWACWGG</b> WTGAACWGTWTAYCCYCC-3'	Leray et al. (2013)
reverse	hgHCO2198:5'- <b>GTCTCGTGGGCTCGGAGATGTGTATAAAGAGACAGTANA</b> CYTCNGGRTGNCCRAARAYCA-3'	Geller et al. (2013)

on each sample in triplicate. Each reaction contained 8.5  $\mu$ L nuclease-free water, 12.5  $\mu$ L 2x KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland), 0.75  $\mu$ L (10  $\mu$ M) forward and 0.75  $\mu$ L (10  $\mu$ M) reverse primer and 2.5  $\mu$ L of DNA template. For Oxford Nanopore sequencing, DNA template was diluted 10x prior to amplification. PCR conditions were 3 min at 95  $^{\circ}$ C, 35 cycles of 30 s at 98  $^{\circ}$ C, 30 s at 57  $^{\circ}$ C, 30 s at 72  $^{\circ}$ C and a final extension for 1 min at 72  $^{\circ}$ C. PCR replicates were pooled, and a clean-up was performed using a 2:1 mixture with AMPURE beads (Beckman Coulter Inc., Brea, CA, USA) and >70% ethanol. Amplification was confirmed using gel electrophoresis (1% gel, ethidium bromide).

## Illumina MiSeq metabarcoding

### Index PCR

For the index PCR, five  $\mu$ L nuclease-free water, 12.5  $\mu$ L 2X KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland) and 2.5  $\mu$ L of each index primer (Nextera XT primer 1 and 2) without modifications was used with 2.5  $\mu$ L of initial pooled PCR product. The PCR program was 3 min at 95  $^{\circ}$ C, 8 cycles of 30 s at 95  $^{\circ}$ C, 30 s at 55  $^{\circ}$ C, 30 s at 72  $^{\circ}$ C and a final extension for 3 min at 72  $^{\circ}$ C. Amplification was confirmed using gel-electrophoresis (1% gel, ethidium bromide). The purified Index PCR products were equimolarly pooled and sequenced using the Illumina MiSeq 2X300bp platform (v2), with an addition of 20% of PhiX (sequenced by Baseclear BV).

### Bioinformatics of Illumina MiSeq reads

After Illumina MiSeq sequencing, the quality of the demultiplexed reads was checked using MultiQC (Ewels et al., 2016), and primers were removed using Trimmomatic (Bolger, Lohse & Usadel, 2014). Amplicon sequence variants (ASV) were generated using the DADA2 pipeline in the Dada2 v1.17.0 package (Callahan et al., 2016) in R Studio v4.0.2 (R Core Team, 2021). Standard settings were used and an error rate of three mismatches was allowed. Reads with a quality score lower than 30 were removed. Unique paired-end reads were determined, merged, and filtered for chimeras for each sample with the removeBimeraDenovo function using the Dada2 v1.17.0 package (Callahan et al., 2016). Taxonomy was assigned using the assignTaxonomy function in the Dada2 package (Wang et al., 2007) using the Ribosomal Database Project (RDP) (Wang et al., 2007) with a minimum bootstrap confidence parameter of 80.

## Reference database for metabarcoding methods

A morphologically curated database that contains 1,992 COI sequences of 565 North Sea invertebrate species was used for taxonomic identification (<http://dx.doi.org/10.5883/DS-GEANS1>). This database was assembled from public and an in-house marine macrobenthos COI sequences that were obtained during multiple diversity monitoring campaigns throughout the North Sea (<https://northsearegion.eu/geans/about/>). During these campaigns, missing species were actively collected, vouchered and barcodes when possible.

## Oxford Nanopore metabarcoding

### *Oxford Nanopore sequencing*

The PCR barcoding kit 96 (EXP-PCB096) was used for the barcoding PCR, and the sequence library was prepared with the SQK-LSK114 kit (Oxford Nanopore Technologies, UK). Several adaptations deviated from the manufacturer's instructions: barcoding PCR was achieved in a total volume of 10  $\mu$ L using 0.3  $\mu$ L 10  $\mu$ M PCR barcode primer pair and 10–50 ng amplicon. The following PCR program was used: initial denaturation at 95 °C for 3 min, 15 cycles of 95 °C for 10 s, 62 °C for 15 s, 65 °C for 90 s, followed by a final extension at 65 °C for 180 s.

The concentration of the barcoded PCR products was measured using the Qubit HS kit (Thermo Fisher Scientific, Waltham, MA, USA) on the non-purified products, after which barcoded PCR products were pooled in equimolar ratios. The pooled amplicon sequence library was cleaned twice using AMPURE beads (Beckman Coulter Inc., Brea, CA, USA). The first clean-up step used 70% ethanol and the second used Short Fragment Buffer (SFB) to enrich for the target size fragments. After end prep and adapter ligation, the library was again washed with SFB during the final clean-up of the protocol. A total of five  $\mu$ L library containing 98.5ng DNA was loaded onto an R10.4.1 flow cell (Oxford Nanopore Technologies, UK) mounted on a Minion Mk1C device.

### *Oxford Nanopore sequence read processing*

Oxford Nanopore sequence read processing was performed according to the post-processing protocol as described by *Doorenspleet et al. (2025)*. Basecalling of the fast5 pass files was performed using *Guppy* (Version 6.5.7, Oxford Nanopore Technologies, UK) in super high accuracy (SUP) mode. The *Decona* pipeline was used (<https://github.com/Saskia-Oosterbroek/decona>) for trimming, clustering, and taxonomic assignment of the reads. Raw base-called reads were trimmed to between 250–400 bases for each sequence. A cluster similarity of 85% was set as the clustering threshold of the sequences. Medaka consensus sequences were generated from each cluster larger than five reads (Decona -f -q 10 -T 18 -l 300 -m 320 -c 0.85 -g “GGWACWGGWTGAACWGTWTAYCCYCC;max\_error\_rate=0.1;min\_overlap=20... TANACYTCNGGRTGNCCRAARAAYCA;max\_error\_rate=0.1;min\_overlap=19” -n 5 -r 0.99 -R 500 -k 6 -M).

## Taxonomic assignment

The consensus sequences were classified using *BLASTn* (NCBI, version 2.11.0), prior to further processing, the classifications were manually checked on inconsistent classifications.

For taxonomic identification, the Top five hits were considered for each consensus sequence. Of these, the best match was used for taxonomic assignment based on hits with the highest e-value. The best match was assigned to species level when there was a minimal alignment length of 250 nucleotides with <4 bp mismatches and >98% identity. The best match was assigned to genus level with >97% identity, family level with >95% identity, >93% for order level and >90% for phylum level identification. The North Sea invertebrate species reference database was used for taxonomic identification (<http://dx.doi.org/10.5883/DS-GEANS1>). This database was the same as used for the Illumina MiSeq metabarcoding method, but not for the metagenomics method.

### Read abundance correction

After taxonomic assignments, a tag correction was performed on the Oxford Nanopore data to correct the tag jumping that had occurred. Tag jumping had occurred when both forward and reverse barcode tags were on sequences that did not belong to that barcode. After troubleshooting, most of the contamination could be alleviated by removing 1% of the total read count of each species from each barcode. This correction is intended to correct proportionally for the total read abundance of each taxon. Although no positive control sample was used for this study, PCR negative controls did not show any positive band nor higher DNA concentrations after barcode PCR. However, the raw data contained a low amount of reads of the most commonly occurring species in these. As the same species also occurred in the ZVL samples with low read abundance where we did not expect them based on the Illumina results, we decided to remove these reads by a minimum 1% read abundance correction. Therefore, we decided to remove these reads by a 1% read abundance correction and this appeared to reduce these observations.

First, the total read count was calculated for all species in all barcodes using the `mutate()` function in the `dplyr` v1.1.0 package. Second, 1% of the total read count was calculated and rounded to a whole number of reads. Finally, 1% of the total read count of each species was subtracted from each barcode.

### NovaSeq shotgun sequencing

#### *Preparations for sequencing*

After DNA extraction and clean-up, libraries were directly prepared using NEBNext<sup>®</sup> Ultra<sup>™</sup> II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA). Samples were indexed using NEBNext<sup>®</sup> Multiplex Oligos for Illumina (New England Biolabs, Ipswich, MA, USA). The final quality check was conducted using the Agilent TapeStation system using hds1000 screen tape (Agilent Technologies, Santa Clara, CA, USA). Samples were pooled in equimolar concentrations and sent to the Norwegian Sequencing Center in Oslo to be sequenced using the NovaSeq S4 quarter flow cell (2x150).

### Sequence read processing

The NovaSeq reads were trimmed using `cutadapt` (`-minimum-length = 100, q = 30`), and merged with `PEAR v 1.7.2` ([Zhang et al., 2014](#)) using the default parameters. Reads were then assembled using `idba_ud` ([Peng et al., 2012](#)) using default parameters. Reads were mapped back to the contigs using `BBMap` (<http://sourceforge.net/projects/bbmap/>) with

the parameter `-minid` set to 90. BLASTn was used to align the contigs with settings set at >97% identity, *e*-value set at  $-10$ , and alignment length equal and over 100 bp. Taxonomic assignments to more than one phylum per query were removed and then the match with the highest bit-score was kept for the data analysis. Single alignments were discarded for the data analysis.

### Reference databases for metagenomics method

To compare the performance of different databases, contigs were locally aligned to four sets of databases. One, the curated COI database for North Sea specific marine invertebrates (GEANS v4, <http://dx.doi.org/10.5883/DS-GEANS1>). Two, the metagenomics data were also aligned a manually curated NCBI-nt database (original complete NCBI-nt downloaded on 22.05.2022) that mainly contained marine eukaryotes. Three, the nuclear entries of the curated NCBI-nt database only, and four, the mitochondrial entries of the curated NCBI-nt database only. The complete NCBI-nt database including both nuclear and mitochondrial entries gave the most optimal result, and therefore this database was used for the comparison with the metabarcoding methods.

### Data analysis

Data analysis was carried out in RStudio v4.2.2 (R Studio v4.2.2). A rarefaction curve (*Vegan* v2.6.4) was plotted to understand the effect of differential sequencing depths between samples. Each sample showed a flattening curve (Fig. S1), which indicated that for each DNA-based method, an appropriate sequencing depth was achieved. Therefore, the data were not rarefied but transformed using a  $\log_{10}$  transformation. All reads from the metabarcoding method (Illumina MiSeq and Oxford Nanopore) were compared on ASV /consensus sequence level to verify the initial assignment between sequences (>97% percentage identity). Phylum level comparisons were made to determine the differences between general taxonomic composition between the different DNA based methods (Miseq, Nanopore, NovaSeq). For visualisation of the database comparison with the Novaseq results, a Pie chart was used (*ggplot2*, v3.4.0). Species level richness, and Shannon diversity index (*H*,  $\log_{10}$  transformed) of the read counts were calculated using the `diversity()` function (*Vegan* v2.6.4) and visualized using `boxplot` (*ggplot2*, v3.4.0). Normality of the data was tested using Shapiro–Wilk for normal distribution and Q-Q plots. Based on these results a 2-way ANOVA using the `aov()` function (*stats*, v3.6.2) was carried out to determine the differences between sampling locations and the DNA-based method used and whether an interaction effect could be observed. For a pairwise comparison, a *post-hoc* analysis was performed using the Tukey test (HSD) using the `TukeyHSD()` function (*stats*, v3.6.2). For beta diversity, non-metric multidimensional scaling ('bray') was performed on each dataset separately (Oxford Nanopore metabarcoding, Illumina MiSeq metabarcoding and Shotgun metagenomics sequencing) in combination with *betadisper* to check for homogeneity of variance. A PERMANOVA was used along with a pairwise *post-hoc* analysis using a Bonferroni correction (both available in `adonis2()`, *Vegan* v2.6.4) to analyse which locations were significantly different from each other, within each dataset. A Spearman rank correlation was used (*stats* v3.6.2), to compare the size class corrected morphological abundance findings with each DNA-based method.

## RESULTS

### Read processing comparison between DNA-based methods

From the Illumina MiSeq metabarcoding samples, a total of 10,206,087 sequences were obtained with 8,692,409 used after processing. Of these, 8,330,994 reads were used as ASVs for taxonomy assignment (39 species). An overview is available of the sequencing output of each DNA-based method after sequencing and processing (Table S3). From the Oxford Nanopore metabarcoding dataset, 2,426,017 reads were basecalled. This is a lower sequencing depth than commonly used for Oxford Nanopore sequencing (Van Der Reis *et al.*, 2023). However as indicated, the lower sequencing depth was not reflected in the rarefaction curves indicating that it was not informative to sequence any further. Of the basecalled reads, 1,841,385 remained after clustering and consensus building. A total of 1,191,853 of the remaining reads were used as consensus sequences for taxonomy (48 species). From the 3,060,417,120 shotgun reads obtained from the Illumina NovaSeq metagenomics run, 2,425,520,473 reads passed the quality values for direct taxonomic assignment of the reads. For the species assignment using the COI curated invertebrate database, 42,262 reads (39 species) could be assigned to species level whereas the complete NCBI database with marine eukaryotes was used, a total of 8,512,483 (36 species) reads could be assigned to species level, which represented 0.35% of the total filtered reads. For direct comparison of each species for each dataset see further.

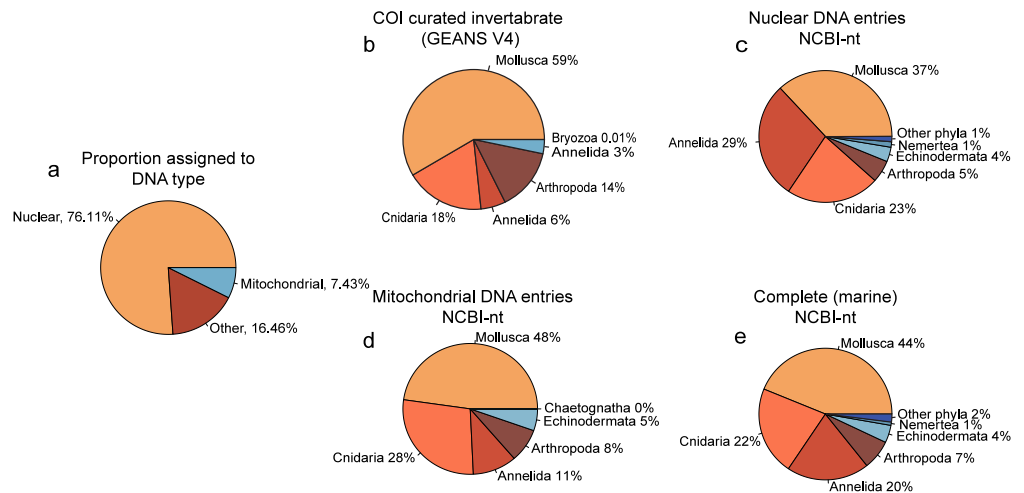
### Initial comparison and Novaseq metagenomics reference database choice

Direct comparison to the processed metabarcoding sequences between Oxford Nanopore (consensus sequences) and Illumina MiSeq (ASVs), indicated that most of the sequences were similar between methods (Fig. S2). In addition, phylum level assignment showed that both the number of phyla as well as the proportion of phyla is very similar between metabarcoding methods (Fig. S3).

Of the reads that were used for taxonomic classification of the NovaSeq metagenomics dataset, 67.85% (5,775,720 reads) could be assigned to the nuclear fraction of the DNA whereas 7.08% (602,684 reads) was assigned to mitochondrial DNA (Fig. 2A). Between the taxonomic assignment of the nuclear (Fig. 2B) and mitochondrial fraction (Fig. 2C), the nuclear fraction represents more Annelida whereas the mitochondrial fraction represents more Mollusca. In addition, the nuclear fraction represents more phyla that are present with <1%. Within the comparison of alignments with the curated COI marine invertebrate database, the amount of Annelida was 14.32% lower with a total of 5.08% (Fig. 2E). The proportions between phyla obtained with the complete marine NCBI-nt data (Fig. 2D), was used for the diversity comparison between methods, as this covered the most diversity. However, the proportions between phyla obtained through NovaSeq were very dissimilar to the proportions observed in the metabarcoding methods (Fig. S3).

### Alpha diversity obtained with the three DNA-based methods

A two-way ANOVA showed a significant interaction effect between sequencing techniques (Illumina MiSeq, Nanopore, NovaSeq), and location (120, 330, 840 and ZVL) for species



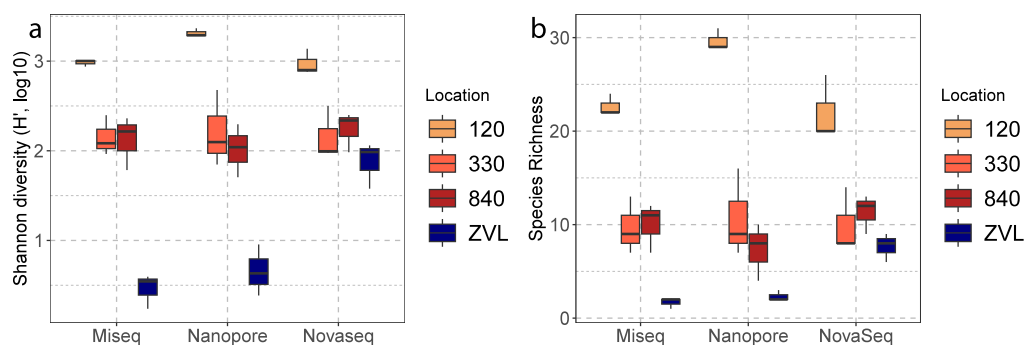
**Figure 2** Proportion phyla found with each database using NovaSeq metagenomics. Pie charts showing (A) the proportions of types of DNA within the Novaseq metagenomics method. (B) Pie chart showing the phyla found using Illumina Shotgun metagenomics in combination with the COI curated marine invertebrate database, (C) the marine Nuclear DNA NCBI-nt database (D) the marine mitochondrial NCBI-nt database and (E) the complete marine NCBI-nt database.

Full-size DOI: [10.7717/peerj.19158/fig-2](https://doi.org/10.7717/peerj.19158/fig-2)

richness ( $F = 4.34$ ,  $p < 0.01$ ). A main effect was also observed for location ( $F = 98.79$ ,  $p < 0.01$ ) and no significant difference was found between methods ( $p = 0.23$ ). *Post-hoc* analysis using a Tukey's HSD test showed a significant difference between location 120 and all other locations (Table S4), where 120 had the highest richness (Fig. 3A). A significant difference was also observed between ZVL and all other locations (Table S4), where ZVL had the lowest richness. No significant difference was found between location 330 and 840 ( $p = 0.97$ ). As for the interaction effect, the NovaSeq showed no significant different effect between ZVL and location 330 and 840. A two-way ANOVA showed a significant interaction of Shannon indices between sequencing techniques and locations ( $F = 7.67$ ,  $p < 0.01$ , Table S4). The main effects were only significant for the factor location ( $F = 104.49$ ,  $p < 0.01$ , Table S4) and for the factor method ( $F = 7.31$ ,  $p \leq 0.05$ ). *Post-hoc* analysis using a Tukey's HSD test showed significant differences in Shannon indices between location 120, and all other locations (Table S5), where 120 had the highest Shannon index (Fig. 3B). With the exception of the NovaSeq analysis, a significant difference was also found between ZVL and all other locations (Table S4), where ZVL had the lowest Shannon index (Fig. 3B). Overall, these results highlight that all sequencing techniques similarly observe alpha diversity between locations, except for the richness and Shannon of the Novaseq shotgun metagenomics in location ZVL as more species were observed than the metabarcoding approaches. Shannon indices were also higher for NovaSeq for the low diversity location ZVL.

### Beta diversity obtained from the three DNA-based methods

The PERMANOVA demonstrated significant differences in macrobenthic community compositions between locations ( $F = 16.05$ ,  $p < 0.01$ ) and between techniques used



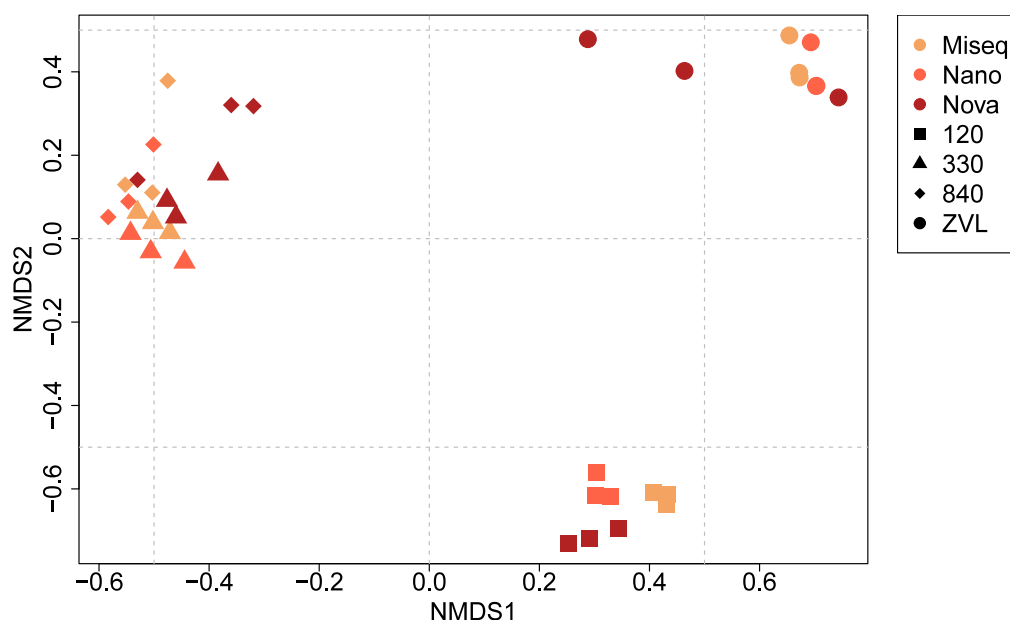
**Figure 3** Boxplots of alpha diversities between locations of each DNA based method. The color represents the different location 120, 330, 840 and ZVL. The facets represent each DNA-based method, Amplicon, Illumina MiSeq and Nanopore. Panel (A) represents the number of species (species richness) and panel (B) represents the Shannon diversity for each location and DNA-based method.

Full-size [DOI: 10.7717/peerj.19158/fig-3](https://doi.org/10.7717/peerj.19158/fig-3)

( $F = 3.25$ ,  $p < 0.01$ ) (Table S5, Fig. 4). In addition, a significant interaction effect was observed between location and the DNA-based approach used ( $F = 2.48$ ,  $p < 0.01$ ), indicating that benthic community composition and location depend on the DNA-based method used and the other way around (Table S5). *Post-hoc* pairwise PERMANOVA tests showed a significant difference in community composition between all locations ( $p < 0.01$ ). Surprisingly, *post-hoc* analysis showed no significant difference in community composition between either sequencing technique. The NMDS plot showed a clear clustering for each location, except for locations 330 and 840 (Fig. 4). The plot also illustrates that location was responsible for the biggest contrasts in community composition, indicating that the location explained most of the variation between community compositions. Similarly, the NMDS plot also indicated that NovaSeq has a different community composition in some locations, particularly within location ZVL. An interaction effect was also observed, indicating that the effect of sequencing platform choice on species composition depends on the community present. Sequencing platforms choice did not affect the species composition for Illumina MiSeq and Oxford Nanopore metabarcoding, as these approaches clustered together within each location.

### Comparison of DNA-based methods for assessing location-specific species composition

All three DNA-based methods (Fig. 5) shared 25 species and an additional 29 species were shared between just the metabarcoding methods (Fig. 5, Fig. S4). At location 120, 23 species were shared between the two metabarcoding methods whereas only 12 species were shared between all three DNA-based methods (Fig. 5, Fig. S4). The Oxford Nanopore metabarcoding data contained seven unique species for all locations (Fig. 5B, Fig. S4). Interestingly, NovaSeq contained 38 unique species for all locations, and most of these unique species were found in all locations (Fig. 5B, Fig. S4). Several species such as, *Crepidula fornicata*, *Eumida mackiei* and *Magelona mirabilis*, occurred at location 330 using the Oxford Nanopore dataset but with less reads than in location 120 (Fig. 5B). However,



**Figure 4** NMDS ordination plot (Bray–Curtis dissimilarity) of similarities of communities between samples and DNA based method. Colors indicate the different DNA based methods whereas the shapes indicate the different locations on species level (log10 transformation).

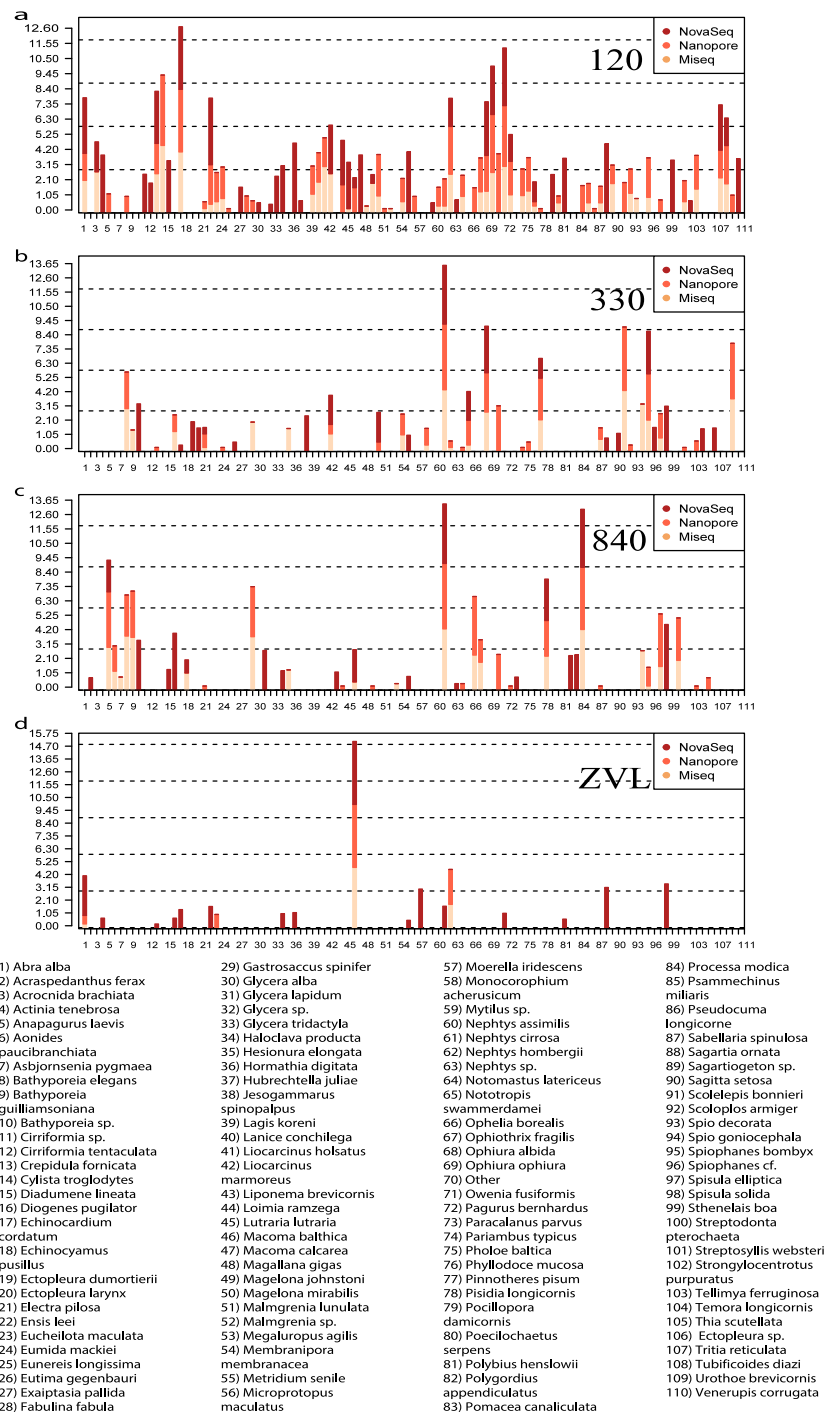
Full-size [DOI: 10.7717/peerj.19158/fig-4](https://doi.org/10.7717/peerj.19158/fig-4)

these species were only present in location 120 using the MiSeq dataset. Nevertheless, all DNA-based methods detected unique species with each method.

### Comparison of species presence and abundance between DNA-based methods and morphology

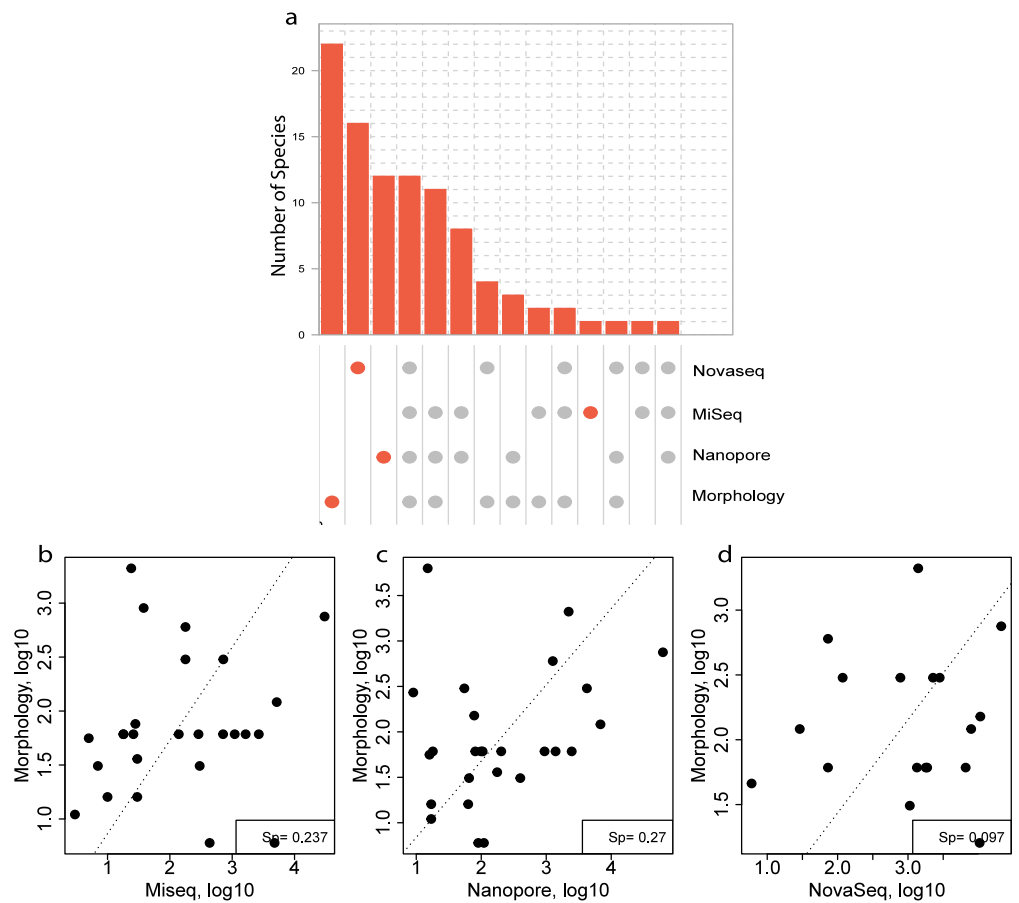
For the morphological identification, one biological replicate was available (120B, 330C, 840C and ZVLA) for each location. This resulted in the identification of 56 species. A total of 39, 13, 10 and three species were identified at locations 120, 330, 840 and ZVL, respectively (Table S1). Of the 56 species identified, 22 species were identified using morphology only (Fig. 6A, Table S7). Of these 22 species, 13 did not have a reference sequence available in the chosen reference database (Table S7, coloured in red). In addition, nine of the species that did not have a reference sequence available belonged to the phylum Annelida (Table S7), which are known to have low primer efficiency and are therefore harder to include in COI reference databases (Carr *et al.*, 2011). Of the 32 species that were identified using both DNA-based methods and morphology, 12 species were identified using all three DNA-based methods (Fig. 6A), 11 species were identified using the metabarcoding-based methods only (Fig. 6A). Lastly, eight species that were identified morphologically were identified only with the NovaSeq metagenomics method whereas an additional three (Nanopore) and two (MiSeq) were unique for the metabarcoding method (Fig. 6A).

Three species were identified with all three DNA-based methods (Fig. 6A), that were not identified with morphology and an additional eight were identified using only the metabarcoding methods. These eight species included *Scolecipois bonnierii* and *Cylista*



**Figure 5** Relative read abundance of each species found between locations and DNA-based method. Bar plot showing the relative read abundance ( $\log_{10}$ ) of all species found. Colors represent the different DNA-based methods. (A) The relative read abundance of location 120, (B) the relative read abundance of location 330, (C) the relative read abundance of location 840 and (D) the relative read abundance of location ZVL.

Full-size DOI: 10.7717/peerj.19158/fig-5



**Figure 6** Comparison of species overlap and relative abundance between morphology and DNA-based methods. (A) UpSet plot showing the overlapping and unique species found using each DNA-based method in all the morphologically identified samples. Where the orange dots refer to unique species per method, and the gray dots shared among the respective methods. A Spearman correlation of the overlapping species abundance between morphology and (B) Illumina MiSeq metabarcoding (C) Oxford Nanopore metabarcoding and (D) NovaSeq shotgun metagenomics.

Full-size DOI: 10.7717/peerj.19158/fig-6

*troglydites* which represented a substantial amount of the identified reads (Table S6). This indicated that the DNA-based methods can identify species that are missed using morphological identification. Oxford Nanopore identified an additional nine species that were unique (Fig. 6A) but found as singletons (Table S6).

A Spearman rank correlation between morphological size class corrected counts and relative read abundance between each DNA-based method across all locations showed the highest positive correlation with both the metabarcoding methods ( $Sp = 0.237$  (MiSeq),  $Sp = 0.27$  (Nanopore) (Figs. 6B–6D). However, the metagenomics method only showed a weak correlation ( $Sp = 0.097$ ) with the size class corrected morphological counts (Fig. 6D).

## DISCUSSION

We compared three different DNA-based methods: two amplicon-based metabarcoding approaches (MiSeq, Nanopore) and a shotgun metagenomics approach (NovaSeq). We assessed the robustness between techniques in terms of alpha and beta diversity to understand the suitability for macrobenthos monitoring. We have demonstrated that the two metabarcoding methods showed a similar diversity richness and Shannon as well as location-specific species compositions. The NovaSeq metagenomics method also showed similar alpha diversities, but the community composition was slightly different, mostly due to taxonomic assignments that were uniquely found with this method. Interestingly, 22/56 species found in the morphological dataset were not identified with molecular techniques and shows that more than half of the species were identified. Nevertheless, all methods showed that most species were shared within each location.

### **Macrobenthos communities are highly similar using metabarcoding despite using different sequencing platforms, platform specific considerations remain**

No difference was found between alpha and beta diversity when comparing a standardized lab protocol with Illumina MiSeq sequencing ([Van den Bulcke et al., 2023](#)) to Oxford Nanopore metabarcoding. Furthermore, both methods identified similar species composition and community structure at each location and generally, shared similar ASVs/consensus sequences and proportions of phyla. Therefore, this study clearly demonstrated that Oxford Nanopore and Illumina MiSeq metabarcoding are equally suited for monitoring macrobenthos biodiversity. This is in line with recent comparisons ([Chang et al., 2023](#); [Van Der Reis et al., 2023](#); [Voorhuijzen-Harink et al., 2019](#)) showing that the Oxford Nanopore datasets are overall comparable with Illumina MiSeq results. This study demonstrates that, even for high-diversity samples, both sequencing platforms perform comparably in terms of macrobenthos diversity assessments, however platform specific considerations remain, including bioinformatics pipelines and sequencing error rates, that may impact species detection and read count estimations. Oxford Nanopore sequencing identified some species that were uniquely present in this dataset. These exclusive taxa were found mostly in very low relative read abundance and did not influence the alpha and beta diversity. These detections could be explained as an effect of the stochastic nature of PCR that is observed in each dataset ([Van der Loos & Nijland, 2021](#)). Indeed, previously reported assessments of reproducibility of laboratories showed minor variations between detected species ([Van den Bulcke et al., 2021](#)).

Another potential explanation for these differences is that the Oxford Nanopore chemistry and its protocols might be more prone to 'barcode leakage' ([Framst et al., 2024](#)). The Oxford Nanopore dataset contained some species with only several reads in one of the replicates, while in another location these species were abundant as is the case for example *Crepidula fornicata* and *Eumidia mackei*. Even though a simple correction of barcode leakage was used for this study (as presented in the materials and methods), it is possible that tag jumping, or barcode leakage was still a problem for this dataset. Although this could not be confirmed as no positive control was considered, several independent

projects within the same lab that used the same barcoding protocol had similar issues. After routine testing, this problem was found to be related to the temperature at which barcoded samples were pooled for the Oxford Nanopore sequencing library. Barcode leakage problems have been reported often in metabarcoding studies ([Beentjes et al., 2019](#); [Van der Loos & Nijland, 2021](#)). Therefore, it is important to consider protocols that minimize the possibility of barcode leakage. This includes using negative and positive controls at each stage, minimizing the handling and amplification of tagged products, or correcting for a crossover of tags between samples ([Beentjes et al., 2019](#)). Similar to other platforms, Oxford Nanopore uses two cycles of PCR, one for the amplification of the region of interest and a second PCR for the barcode attachment using Oxford Nanopore-specific protocols and kits for barcoding. The protocol used for this study could be optimized in such a way that individual samples are amplified with already tagged initial barcodes to circumvent the second PCR step ([Srivathsan et al., 2021](#)) or by using amplicon-free barcode kits ([Toxqui Rodríguez, Vanhollebeke & Derycke, 2023](#); [Van Der Reis et al., 2023](#)).

The bioinformatics pipeline used in this study could also be responsible for the small differences observed between the metabarcoding results. The 10.4.1 flow cells were used, which greatly improved the quality of the metabarcoding data ([Zhang et al., 2023](#)). Nevertheless, clustering and consensus building of the data with bioinformatics processing might also result in overlooking some elements of the biodiversity ([Brandt et al., 2021](#)). Therefore, the reads that were not included in a cluster for consensus building were also considered when they met the threshold for taxonomy assignment. This inclusion may have led to the detection of additional species within the individual samples, although this is unlikely since the singletons did not contain any additional species ([Table S8](#)).

Despite these minor differences between metabarcoding methods, this did not lead to significant differences in commonly used biodiversity indices. Therefore, this study suggests that Oxford Nanopore and Illumina MiSeq metabarcoding are highly similar for macrobenthos monitoring. However, platform specific challenges need to be considered, as Illumina generally provides high-accuracy short read data effective for fine scale taxa identification. In contrast, Nanopore offers sequencing real-time and long-read reads albeit typically with a higher error rate resulting in more intensive error correction during downstream analysis.

### **NovaSeq Shotgun sequencing identified the most method specific species**

The primer and PCR-free shotgun metagenomics method yielded similar alpha diversities compared to the metabarcoding methods but resulted in a slightly different community composition between methods within location. These different community compositions between methods were mostly due to additional species that were only found using the metagenomics method. Nevertheless, most species that were detected by NovaSeq metagenomics were also detected by the two metabarcoding methods suggesting that either method can describe the general community equally well.

This study, therefore, reflects current environmental metagenomics studies that show equal or higher levels of biodiversity compared to amplicon-based approaches ([Bista et](#)

*al.*, 2018; *Garlapati et al.*, 2019; *Monchamp et al.*, 2022; *Paula et al.*, 2022). However, these studies, were not targeting a certain taxonomic group and therefore, reflected a wider spectrum of biota. In contrast, this study aimed to test whether metagenomics is suitable for specifically species/level macrobenthos biodiversity monitoring and therefore the data presented are not directly comparable to other metagenomics studies.

The highest number of uniquely found species was found with the metagenomics method, which was especially apparent in the low diversity samples (ZVL). The most straightforward explanation would be contamination between the high and low diversity samples, as these species were not found in the morphological analysis. Although molecular methods are always prone to cross contaminations in the lab (*Van der Loos & Nijland, 2021*), it is in this case less likely since part of the extractions used in this study for ZVL previously were also used in other studies (*Van den Bulcke et al.*, 2023) and did not experience similar problems.

Alternative explanations for these findings are (1) that misidentification could have occurred due to the use of a database (NCBI) with mislabelled references (*Sinniger et al.*, 2016) or (2) false positive detection of species that are genetically similar but do not occur in the North Sea as is the case for *Actinia tenebrosa* (*Ayre, 1985*). Furthermore, (3) the species is present but were not picked up with the metabarcoding method because these species did not have a curated COI sequence. As most of the additionally found species with the metagenomics approach belonged to cnidarians, it is possible that during morphological identifications only part of the animal was present in the initial sample and was therefore also never recorded for the morphological analysis (*Lobo et al.*, 2017).

Although the complete NCBI reference database with relevant species was used, still only 0.35% of the metagenomics reads were assigned to a species. Reference databases that contain environmental sequences are mainly focused on genetic regions that are popular for metabarcoding (*Weigand et al.*, 2019) and references of full (mito)genomes are still in their infancy (*Blasiak et al.*, 2020; *Leray, Knowlton & Machida, 2022*), it is not surprising that such a low percentage of reads are assigned. Thus, the present findings align with previous research that suggest that shotgun sequencing metagenomics is hardly feasible for environmental studies (*Ficetola & Taberlet, 2023; Quince et al.*, 2017). These results also indicate the necessity to improve reference databases, particularly for full (mito)genomes of macrobenthos, as the limited availability of references largely contributed to the low number of reads that are assigned.

Similarly, there was a weak correlation between species that were identified with the metagenomics approach and the morphologically identified species when correcting for size class. Unfortunately, no biomass data was obtained from the morphologically identified samples, which would be a more appropriate representation of the actual abundance. Nevertheless, a better correlation was found with the metabarcoding data, which is in contrast with findings that suggest a better correlation between found metagenomics reads and biomass (*Bista et al.*, 2018). Since in the study by Bista et al., the complete mitogenomes of all species in the mock community was available, the lack of correlations in this study further suggests that the metagenomics method is presently not feasible for reliable DNA-based monitoring of macrobenthos biodiversity. In the future, focusing on

reference databases and especially databases that contain complete (mito)genomes will greatly improve detection rates and improve DNA-based monitoring.

### **Mismatches between morphology and DNA based methods not only linked to reference sequence availability**

As typical for metabarcoding studies, several species were found with metabarcoding that were not reported during the morphological analysis. This is not surprising as these species are either missed during morphological sorting or can only be identified to a higher taxonomic level (*Lobo et al., 2017; Robinson et al., 2022*). For this study this was apparent for the Anthozoa *Cylista* that was found with high read count (>13,000 reads), whose tissue becomes hard or impossible to recognize when not intact.

Most taxa were detected using both morphology and a DNA-based technique, but there were 22 species that were not identified using any DNA-based method. This was especially apparent in the high diversity samples. Of these, 13 species could not be identified due to incompleteness of the curated database, which especially in relation to marine invertebrates, has often been reported as a limiting factor (*Aylagas et al., 2016; Günther et al., 2018; Steyaert et al., 2020; Willassen et al., 2022*). Several efforts have been made to improve (*Lavrador et al., 2023*) and increase the coverage of these databases (*Leray et al., 2018; Radulovici et al., 2021*). Approximately 22–43% of all European marine species have reference sequences available in the widely used reference database BOLD (*Weigand et al., 2019*). Multiple studies in the North Sea have emphasized the importance of enhancing these databases (*Christodoulou et al., 2021*), indicating a need for sustained long-term efforts. To improve the reference database used here, much effort has been put in vouchering as many species as possible by collaborating internationally with molecular scientists (<https://northsearegion.eu/geans/about/>). However, after morphological sorting of samples of this study, samples were preserved in formaldehyde which is known to hamper DNA extraction from these specimens therefore a one-on-one comparison was not possible. In addition, of the 13 unidentified species due to incompleteness of the reference database, 7 could be found in the COI NCBI database (manual search, 2025), suggesting that the curated database may be less complete in terms of species presence, but was on the other hand referenced with a voucher species and could therefore be confirmed as the correct identification.

However, as nine species did have the potential to be picked up with the used reference database, more factors should explain mismatches between methods. One explanation is the procedure in the field as small differences may occur in species composition between the grabs. 12 out of the 22 species that were not identified with any DNA based method, had an abundance of only one specimen in the morphologically assigned grab sample. Although others with only one morphological assignment were identified (see [Table S6](#)), four species that were not found in the comparison, (*Aonide paucibranciata*, *Nototrips Swammerdamei*, *Phyllodoce mucosa* and *Streptodonta pterochaeta*) did occur in other replicas of at least one of the DNA based methods. This suggesting that small differences between grabs explain the mismatch to a certain extend.

During lab procedures biases can also occur, for example primer efficiency have also been often reported to influence metabarcoding results (*Van der Loos & Nijland, 2021*). Of the 13 species that were missing in the reference database, nine species belonged to Annelida. Annelida, and especially certain Polychaeta, are known to be difficult to sequence using the conventional markers because this group has high variation within the COI region and therefore has lower primer binding efficiency (*Carr et al., 2011*). This have caused the COI reference databases to be biased towards a lower representation of Annelida. Contrariwise, as the primer efficiency is lower for Annelida, it is likely that the chosen COI primer is not representing the taxonomic group well. This is emphasized by the fact that an additional five annelids were not detected with a DNA based method while they were represented in the reference database. Using multiple markers that are more specific to certain taxonomic groups might therefore greatly improve the capability to detect species from metabarcoding methods. However, consequently this may increase laboratory time and costs (*Cordier et al., 2019; Gielings et al., 2021; Meyer et al., 2021*). Metagenomic methods can also improve the detection of annelids as this technique does not rely on the amplification of specific regions. Nevertheless, the metagenomics method in this study did not result in the retrieval of more annelids.

Other factors such as the taxonomic assignment of the species using the algorithms of RDP or selecting the best scoring match per query can also be of influence as the size and training set of especially RDP influences the sensitivity of detecting certain species (*Ritari et al., 2015*). Finally, the taxonomic assignment of the species can be wrong. In this case however, this is less likely as this survey is conducted by experts and samples were quality controlled (see *Derycke et al., 2021*) suggesting that this had little influence on the results.

In conclusion, although DNA-based methods missed 22 species compared to morphological analysis, the DNA-based methods detected additional species that otherwise remained undetected. This is in line with previous reports, also confirming that DNA-based methods and morphology should remain complementary (*Cahill et al., 2018; Kelly et al., 2017; Meyer et al., 2021*). However, improvements to the databases, the use of multiple replicates and the choice of multi-marker approach for specific taxonomic groups might further improve detection sensitivity and at some point, may surpass morphological identification.

## CONCLUSIONS

In recent years, there has been an increased interest in implementing DNA-based tools into routine biodiversity monitoring practices. To achieve this, standardized protocols are necessary to ensure reproducibility and data robustness across studies and regions, particularly as sequencing technology is evolving rapidly (*Van den Bulcke et al., 2023*).

In this study, we showed that similar alpha and beta diversity patterns were found regardless of the metabarcoding platform used. Thus, Illumina MiSeq and Oxford Nanopore sequencing results are, at this stage, highly similar and can both be used to monitor macrobenthos biodiversity, although platform specific challenges remain. In addition, NovaSeq metagenomics has the potential for environmental monitoring and

showed similar communities. However, metagenomics at this stage does not surpass metabarcoding methods, as it did not improve representation of abundance data, nor did it identify more species that are not efficiently amplified by the COI primer. Incomplete reference databases still hamper detection of metagenomics sequences but mismatches between morphological and DNA based findings are also influenced by fieldwork choices, primer biases and bioinformatics choices. These findings demonstrate that portable, real-time Oxford Nanopore sequencing is ready to be integrated for standard monitoring practices and emphasize the importance of improving sequence reference databases to implement and enhance metagenomics methods for robust and harmonized monitoring practices.

## ACKNOWLEDGEMENTS

We would like to thank the following individuals for their valuable comments during the research project and on the manuscript: Laure Van Den Bulcke, Pedro Martinez Arbizu, Magdalini Christodoulou, Truls Borg Moum, Rumakanta Sapkota, Anne Winding, Sahar Khodami, Silvana Birchenough, Peter Anton Upadhyay Stæhr, Phil Davison, Matthias Obst, Ingrid Kroencke, Klaas Deneudt, Per Sundberg, Jørgen L.S. Hansen, Pascal Hablutzel, Rune Lagaisse, Carolin Uhler, Kris Hostens, Sofie Vandendriessche, Herman Verstuyft. Sampling took place on board RV Belgica and ship time was provided by BELSPO and RBINS OD Nature. We thank the crew of RV Belgica for the logistic support during the sampling campaign.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

Funding for this research was received through GEANS–Genetic tools for Ecosystem health Assessment in the North Sea region, an Interreg project supported by the North Sea Programme of the European Regional Development Fund of the European Union. Sofie Derycke and Annelies De Backer received funding through the Sand Fund of the Federal Public Service Economy. There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

GEANS–Genetic tools for Ecosystem health Assessment in the North Sea.

North Sea Programme of the European Regional Development Fund of the European Union.

Sand Fund of the Federal Public Service Economy.

### Competing Interests

The authors declare there are no competing interests.

## Author Contributions

- Karlijn Doorenspleet performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Amalia Aikaterini Mailli performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Berry B. van der Hoorn performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Kevin K. Beentjes performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Annelies De Backer conceived and designed the experiments, authored or reviewed drafts of the article, project administration, funding acquisition, and approved the final draft.
- Sofie Derycke conceived and designed the experiments, authored or reviewed drafts of the article, project administration, funding acquisition, and approved the final draft.
- Albertinka J. Murk analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Henning Reiss analyzed the data, authored or reviewed drafts of the article, supervision, and approved the final draft.
- Reindert Nijland conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, supervision, Project administration, and approved the final draft.

## DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

Marine data archive

## Data Availability

The following information was supplied regarding data availability:

The raw data, processed data and the used scripts including descriptive metadata are available in the Marine data Archive:

– MiSeq: Doorenspleet, K.; Mailli, A. A.; van der Hoorn, B.; Beentjes, K.; De Backer, A.; Derycke, S.; Murk, A.; Reiss, H.; Nijland, R.; Wageningen University and Research; Nord University; Naturalis Biodiversity Center; ILVO Marine Research; (2024): Illumina MiSeq dataset for molecular macrobenthos biodiversity monitoring comparing Oxford Nanopore and Illumina based metabarcoding and metagenomics. Marine Data Archive. <https://doi.org/10.14284/682>.

– Nanopore: Doorenspleet, K.; Mailli, A. A.; van der Hoorn, B.; Beentjes, K.; De Backer, A.; Derycke, S.; Murk, A.; Reiss, H.; Nijland, R.; Wageningen University and Research; Nord University; Naturalis Biodiversity Center; ILVO Marine Research; (2024): Oxford Nanopore dataset for molecular macrobenthos biodiversity monitoring comparing Oxford Nanopore and Illumina based metabarcoding and metagenomics. Marine Data Archive. <https://doi.org/10.14284/683>.

– NovaSeq: Doorenspleet, K.; Mailli, A. A.; van der Hoorn, B.; Beentjes, K.; De Backer, A.; Derycke, S.; Murk, A.; Reiss, H.; Nijland, R.; Wageningen University and Research;

Nord University; Naturalis Biodiversity Center; ILVO Marine Research; (2024): Illumina Novaseq dataset for molecular macrobenthos biodiversity monitoring comparing Oxford Nanopore and Illumina based metabarcoding and metagenomics. Marine Data Archive. <https://doi.org/10.14284/684>.

The custom R scripts used for the Data Analysis are available at GitHub and Zenodo:

– [https://github.com/amalia03/GEANS\\_Platform\\_Comparison\\_Scripts/blob/main/tax\\_geans.R](https://github.com/amalia03/GEANS_Platform_Comparison_Scripts/blob/main/tax_geans.R).

– Amalia M. (2025). amalia03/GEANS\_Platform\_Comparison\_Scripts: GEANS\_Platform\_Comparison\_Scripts (v.1.0). Zenodo. <https://doi.org/10.5281/zenodo.15000726>.

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.19158#supplemental-information>.

## REFERENCES

- Antich A, Palacin C, Wangensteen OS, Turon X. 2021.** To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography. *BMC Bioinformatics* **22**(1):177 DOI [10.1186/s12859-021-04115-6](https://doi.org/10.1186/s12859-021-04115-6).
- Aylagas E, Borja A, Pochon X, Zaiko A, Keeley N, Bruce K, Hong P, Ruiz GM, Stein ED, Theroux S, Geraldi N, Ortega A, Gajdzik L, Coker DJ, Katan Y, Hikmawan T, Saleem A, Alamer S, Jones BH, Duarte CM, Pearman J, Carvalho S. 2020.** Translational molecular ecology in practice: linking DNA-based methods to actionable marine environmental management. *Science of the Total Environment* **744**:140780 DOI [10.1016/j.scitotenv.2020.140780](https://doi.org/10.1016/j.scitotenv.2020.140780).
- Aylagas E, Mendibil I, Borja A, Rodríguez-Ezpeleta N. 2016.** Marine sediment sample pre-processing for macroinvertebrates metabarcoding: mechanical enrichment and homogenization. *Frontiers in Marine Science* **3**:203 DOI [10.3389/fmars.2016.00203](https://doi.org/10.3389/fmars.2016.00203).
- Ayre DJ. 1985.** Localized adaptation of clones of the sea anemone *Actinia tenebrosa*. *Evolution* **39**(6):1250–1260 DOI [10.2307/2408782](https://doi.org/10.2307/2408782).
- Baird DJ, Hajibabaei M. 2012.** Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology* **21**(8):2039–2044 DOI [10.1111/j.1365-294X.2012.05519.x](https://doi.org/10.1111/j.1365-294X.2012.05519.x).
- Baloğlu B, Chen Z, Elbrecht V, Braukmann T, MacDonald S, Steinke D. 2021.** A workflow for accurate metabarcoding using nanopore MinION sequencing. *Methods in Ecology and Evolution* **12**(5):794–804 DOI [10.1111/2041-210X.13561](https://doi.org/10.1111/2041-210X.13561).
- Beentjes KK, Speksnijder AGCL, Schilthuizen M, Hoogeveen M, vander Hoorn BB. 2019.** The effects of spatial and temporal replicate sampling on eDNA metabarcoding. *PeerJ* **7**:e7335 DOI [10.7717/peerj.7335](https://doi.org/10.7717/peerj.7335).
- Bernatchez L, Ferchaud A-L, Berger CS, Venney CJ, Xuereb A. 2024.** Genomics for monitoring and understanding species responses to global climate change. *Nature Reviews Genetics* **25**(3):165–183 DOI [10.1038/s41576-023-00657-y](https://doi.org/10.1038/s41576-023-00657-y).

- Bista I, Carvalho GR, Tang M, Walsh K, Zhou X, Hajibabaei M, Shokralla S, Seymour M, Bradley D, Liu S, Christmas M, Creer S. 2018.** Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources* **18**(5):1020–1034 DOI [10.1111/1755-0998.12888](https://doi.org/10.1111/1755-0998.12888).
- Blasiak R, Wynberg R, Grorud-Colvert K, Thambisetty S, Bandarra NM, Canário AVM, Da Silva J, Duarte CM, Jaspars M, Rogers A, Sink K, Wabnitz CCC. 2020.** The ocean genome and future prospects for conservation and equity. *Nature Sustainability* **3**(8):588–596 DOI [10.1038/s41893-020-0522-9](https://doi.org/10.1038/s41893-020-0522-9).
- Bleidorn C. 2016.** Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity* **14**(1):1–8 DOI [10.1080/14772000.2015.1099575](https://doi.org/10.1080/14772000.2015.1099575).
- Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15):2114–2120 DOI [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- Borja Á, 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**(12):2175–2186 DOI [10.1016/j.marpolbul.2010.09.026](https://doi.org/10.1016/j.marpolbul.2010.09.026).
- Brandt MI, Trouche B, Quintric L, Günther B, Wincker P, Poulain J, Arnaud-Haond S. 2021.** Bioinformatic pipelines combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic metabarcoding. *Molecular Ecology Resources* **21**(6):1904–1921 DOI [10.1111/1755-0998.13398](https://doi.org/10.1111/1755-0998.13398).
- Breine NT, DeBacker A, Van Colen C, Moens T, Hostens K, Van Hoey G. 2018.** Structural and functional diversity of soft-bottom macrobenthic communities in the Southern North Sea. *Estuarine, Coastal and Shelf Science* **214**:173–184 DOI [10.1016/j.ecss.2018.09.012](https://doi.org/10.1016/j.ecss.2018.09.012).
- Bucklin A, Steinke D, Blanco-Bercial L. 2011.** DNA barcoding of marine metazoa. *Annual Review of Marine Science* **3**(1):471–508 DOI [10.1146/annurev-marine-120308-080950](https://doi.org/10.1146/annurev-marine-120308-080950).
- Cahill AE, Pearman JK, Borja A, Carugati L, Carvalho S, Danovaro R, Dashfield S, David R, Féral J-P, Olenin S, Šiaulys A, Somerfield PJ, Trayanova A, Uyarra MC, Chenuil A. 2018.** A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. *Ecology and Evolution* **8**(17):8908–8920 DOI [10.1002/ece3.4283](https://doi.org/10.1002/ece3.4283).
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016.** DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**(7):581–583 DOI [10.1038/nmeth.3869](https://doi.org/10.1038/nmeth.3869).
- Carr CM, Hardy SM, Brown TM, Macdonald TA, Hebert PDN. 2011.** A tri-oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian polychaetes. *PLOS ONE* **6**(7):e22232 DOI [10.1371/journal.pone.0022232](https://doi.org/10.1371/journal.pone.0022232).
- Chang JJM, Ip YCA, Neo WL, Mowe MAD, Jaafar Z, Huang D. 2023.** Primed and ready: Nanopore metabarcoding can now recover highly accurate consensus barcodes that are generally indel-free. *bioRxiv* DOI [10.1101/2023.08.04.552069](https://doi.org/10.1101/2023.08.04.552069).

- Christodoulou M, Van der Hoorn B, Van den Bulcke L, Derycke S, De Backer A, Kröncke I, Arbizu PM. 2021.** A reliable DNA barcode reference library for the identification of benthic invertebrates: essential for biomonitoring of the North Sea. In: *ARPHA conference abstracts. Vol. 4.* e65063 DOI [10.3897/aca.4.e65063](https://doi.org/10.3897/aca.4.e65063).
- Cordier T, Lanzén A, Apothéoz-Perret-Gentil L, Stoeck T, Pawlowski J. 2019.** Embracing environmental genomics and machine learning for routine biomonitoring. *Trends in Microbiology* **27**(5):387–397 DOI [10.1016/j.tim.2018.10.012](https://doi.org/10.1016/j.tim.2018.10.012).
- Creer S, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas WK, Potter C, Bik HM. 2016.** The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution* **7**(9):1008–1018 DOI [10.1111/2041-210X.12574](https://doi.org/10.1111/2041-210X.12574).
- Danovaro R, Carugati L, Berzano M, Cahill AE, Carvalho S, Chenuil A, Corinaldesi C, Cristina S, David R, Dell'Anno A, Dzhenbekova N, Garcés E, Gasol JM, Goela P, Féral J-P, Ferrera I, Forster RM, Kurekin AA, Rastelli E, Marinova V, Miller PI, Moncheva S, Newton A, Pearman JK, Pitois SG, Reñé A, Rodríguez-Ezpeleta N, Saggiomo V, Simis SGH, Stefanova K, Wilson C, Lo Martire M, Greco C, Cochrane SKJ, Mangoni O, Borja A. 2016.** Implementing and innovating marine monitoring approaches for assessing marine environmental status. *Frontiers in Marine Science* **3**:213 DOI [10.3389/fmars.2016.00213](https://doi.org/10.3389/fmars.2016.00213).
- Derycke S, Maes S, Van den Bulcke L, Vanhollebeke J, Wittoeck J, Hillewaert H, Ampe B, Haegeman A, Hostens K, De Backer A. 2021.** Detection of macrobenthos species with metabarcoding is consistent in bulk DNA but dependent on body size and sclerotization in eDNA from the ethanol preservative. *Frontiers in Marine Science* **8**:637858 DOI [10.3389/fmars.2021.637858](https://doi.org/10.3389/fmars.2021.637858).
- DeSalle R, Goldstein P. 2019.** Review and interpretation of trends in DNA barcoding. *Frontiers in Ecology and Evolution* **7**:302 DOI [10.3389/fevo.2019.00302](https://doi.org/10.3389/fevo.2019.00302).
- Devloo-Delva F, Huerlimann R, Chua G, Matley JK, Heupel MR, Simpfendorfer CA, Maes GE. 2018.** How does marker choice affect your diet analysis: comparing genetic markers and digestion levels for diet metabarcoding of tropical-reef piscivores. *Marine and Freshwater Research* **70**(1):8–18 DOI [10.1071/MF17209](https://doi.org/10.1071/MF17209).
- Doorenspleet K, Jansen L, Oosterbroek S, Kamermans P, Bos O, Wurz E, Murk A, Nijland R. 2025.** The long and the short of it: nanopore based eDNA metabarcoding of marine vertebrates works; sensitivity and specificity depend on amplicon lengths. *Molecular Ecology Resources* **10**:e14079 DOI [10.1111/1755-0998.14079](https://doi.org/10.1111/1755-0998.14079).
- Egter B, Veríssimo J, Lopes-Lima M, Chaves C, Pinto J, Riccardi N, Beja P, Fonseca NA. 2020.** Speeding up the detection of invasive aquatic species using environmental DNA and nanopore sequencing. *bioRxiv* DOI [10.1101/2020.06.09.142521](https://doi.org/10.1101/2020.06.09.142521).
- Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F. 2017.** Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution* **8**(10):1265–1275 DOI [10.1111/2041-210X.12789](https://doi.org/10.1111/2041-210X.12789).
- Ewels P, Magnusson M, Lundin S, Käller M. 2016.** MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**(19):3047–3048 DOI [10.1093/bioinformatics/btw354](https://doi.org/10.1093/bioinformatics/btw354).

- Ficetola GF, Taberlet P. 2023.** Towards exhaustive community ecology via DNA metabarcoding. *Molecular Ecology* 32(23):6320–6329 DOI 10.1111/mec.16881.
- Framst I, Wolking RM, Schonfeld J, Ricker N, Beeler-Marfisi J, Chalmers G, Kamath PL, Maboni G. 2024.** High-throughput rapid amplicon sequencing for multilocus sequence typing of *Mycoplasma ovipneumoniae* from archived clinical DNA samples. *Frontiers in Veterinary Science* 11:1443855 DOI 10.3389/fvets.2024.1443855.
- Gaither MR, Szabó Z, Crepeau MW, Bird CE, Toonen RJ. 2011.** Preservation of corals in salt-saturated DMSO buffer is superior to ethanol for PCR experiments. *Coral Reefs* 30(2):329–333 DOI 10.1007/s00338-010-0687-1.
- Garlapati D, Charankumar B, Ramu K, Madeswaran P, Ramana Murthy MV. 2019.** A review on the applications and recent advances in environmental DNA (eDNA) metagenomics. *Reviews in Environmental Science and Bio/Technology* 18(3):389–411 DOI 10.1007/s11157-019-09501-4.
- 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 DOI 10.1111/1755-0998.12138.
- Gielings R, Fais M, Fontaneto D, Creer S, Costa FO, Renema W, Macher J-N. 2021.** DNA metabarcoding methods for the study of marine benthic meiofauna: a review. *Frontiers in Marine Science* 8:730063 DOI 10.3389/fmars.2021.730063.
- Gleason JE, Hanner RH, Cottenie K. 2023.** Hidden diversity: DNA metabarcoding reveals hyper-diverse benthic invertebrate communities. *BMC Ecology and Evolution* 23(1):19 DOI 10.1186/s12862-023-02118-w.
- Goodwin KD, Thompson LR, Duarte B, Kahlke T, Thompson AR, Marques JC, Caçador I. 2017.** DNA sequencing as a tool to monitor marine ecological status. *Frontiers in Marine Science* 4:107 DOI 10.3389/fmars.2017.00107.
- Gostel MR, Kress WJ. 2022.** The expanding role of DNA barcodes: indispensable tools for ecology, evolution, and conservation. *Diversity* 14(3):213 DOI 10.3390/d14030213.
- Günther B, Knebelsberger T, Neumann H, Laakmann S, Martínez Arbizu P. 2018.** Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes. *Scientific Reports* 8(1):14822 DOI 10.1038/s41598-018-32917-x.
- Heikema AP, Horst-Kreft D, Boers SA, Jansen R, Hiltmann SD, De Koning W, Kraaij R, De Ridder MAJ, Van Houten CB, Bont LJ, Stubbs AP, Hays JP. 2020.** Comparison of Illumina versus nanopore 16S rRNA gene sequencing of the human nasal microbiota. *Genes* 11(9):1105 DOI 10.3390/genes11091105.
- Hering D, Borja A, Pont D, Jones JI, 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 DOI 10.1016/j.watres.2018.03.003.
- Hinz S, Coston-Guarini J, Marnane M, Guarini J-M. 2022.** Evaluating eDNA for use within marine environmental impact assessments. *Journal of Marine Science and Engineering* 10(3):375 DOI 10.3390/jmse10030375.

- Hu T, Chitnis N, Monos D, Dinh A. 2021.** Next-generation sequencing technologies: an overview. *Human Immunology* **82(11)**:801–811 DOI [10.1016/j.humimm.2021.02.012](https://doi.org/10.1016/j.humimm.2021.02.012).
- Jahnke M, Moknes P-O, Le Moan A, Martens GA, Jonsson PR. 2022.** Seascape genomics identify adaptive barriers correlated to tidal amplitude in the shore crab *Carcinus maenas*. *Molecular Ecology* **31(7)**:1980–1994 DOI [10.1111/mec.16371](https://doi.org/10.1111/mec.16371).
- Kelly RP, Closek CJ, O'Donnell JL, Kralj JE, Shelton AO, Samhuri JF. 2017.** Genetic and manual survey methods yield different and complementary views of an ecosystem. *Frontiers in Marine Science* **3**:283 DOI [10.3389/fmars.2016.00283](https://doi.org/10.3389/fmars.2016.00283).
- Klunder L, Van Bleijswijk JDL, Kleine Schaars L, Van der Veer HW, Luttikhuisen PC, Bijleveld AI. 2022.** Quantification of marine benthic communities with metabarcoding. *Molecular Ecology Resources* **22(3)**:1043–1054 DOI [10.1111/1755-0998.13536](https://doi.org/10.1111/1755-0998.13536).
- Krehenwinkel H, Pomerantz A, Prost S. 2019.** Genetic biomonitoring and biodiversity assessment using portable sequencing technologies: current uses and future directions. *Genes* **10(11)**:858 DOI [10.3390/genes10110858](https://doi.org/10.3390/genes10110858).
- Lanzén A, Lekang K, Jonassen I, Thompson EM, Troedsson C. 2017.** DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. *PLOS ONE* **12(6)**:e0179443 DOI [10.1371/journal.pone.0179443](https://doi.org/10.1371/journal.pone.0179443).
- Lavrador AS, Fontes JT, Vieira PE, Costa FO, Duarte S. 2023.** Compilation, revision, and annotation of DNA barcodes of marine Invertebrate Non-Indigenous Species (NIS) occurring in European coastal regions. *Diversity* **15(2)**:174 DOI [10.3390/d15020174](https://doi.org/10.3390/d15020174).
- Leray M, Ho S-L, Lin I-J, Machida RJ. 2018.** MIDORI server: a webserver for taxonomic assignment of unknown metazoan mitochondrial-encoded sequences using a curated database. *Bioinformatics* **34(21)**:3753–3754 DOI [10.1093/bioinformatics/bty454](https://doi.org/10.1093/bioinformatics/bty454).
- 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 of the United States of America* **112(7)**:2076–2081 DOI [10.1073/pnas.1424997112](https://doi.org/10.1073/pnas.1424997112).
- 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 DOI [10.1098/rstb.2015.0331](https://doi.org/10.1098/rstb.2015.0331).
- Leray M, Knowlton N, Machida RJ. 2022.** MIDORI2: a collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences. *Environmental DNA* **4(4)**:894–907 DOI [10.1002/edn3.303](https://doi.org/10.1002/edn3.303).
- Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ. 2013.** A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* **10(1)**:34 DOI [10.1186/1742-9994-10-34](https://doi.org/10.1186/1742-9994-10-34).
- Lobo J, Shokralla S, Costa MH, Hajibabaei M, Costa FO. 2017.** DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. *Scientific Reports* **7(1)**:15618 DOI [10.1038/s41598-017-15823-6](https://doi.org/10.1038/s41598-017-15823-6).

- Lopez MLD, Lin Y, Sato M, Hsieh C, Shiah F-K, Machida RJ. 2022. Using metatranscriptomics to estimate the diversity and composition of zooplankton communities. *Molecular Ecology Resources* 22(2):638–652 DOI 10.1111/1755-0998.13506.
- Mauffrey F, Cordier T, Apothéloz-Perret-Gentil L, Cermakova K, Merzi T, Delefosse M, Blanc P, Pawlowski J. 2021. Benthic monitoring of oil and gas offshore platforms in the North Sea using environmental DNA metabarcoding. *Molecular Ecology* 30(13):3007–3022 DOI 10.1111/mec.15698.
- Meyer A, Boyer F, Valentini A, Bonin A, Ficetola GF, Beisel J-N, Bouquerel J, Wagner P, Gaboriaud C, Leese F, Dejean T, Taberlet P, Usseglio-Polatera P. 2021. Morphological vs. DNA metabarcoding approaches for the evaluation of stream ecological status with benthic invertebrates: testing different combinations of markers and strategies of data filtering. *Molecular Ecology* 30(13):3203–3220 DOI 10.1111/mec.15723.
- Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols* 2010(6):pdb-prot5448 DOI 10.1101/pdb.prot5448.
- Monchamp M-È, Walsh DA, Garner RE, Kraemer SA, Beisner BE, Cristescu ME, Gregory-Eaves I. 2022. Comparative analysis of zooplankton diversity in freshwaters: what can we gain from metagenomic analysis? *Environmental DNA* 4(6):1250–1264 DOI 10.1002/edn3.335.
- Paula DP, Barros SKA, Pitta RM, Barreto MR, Togawa RC, Andow DA. 2022. Metabarcoding versus mapping unassembled shotgun reads for identification of prey consumed by arthropod epigeal predators. *GigaScience* 11:giac020 DOI 10.1093/gigascience/giac020.
- Pawlowski J, Kelly-Quinn M, Altermatt F, Apothéloz-Perret-Gentil L, Beja P, Boggero A, Borja A, Bouchez A, Cordier T, Domaizon I, Feio MJ, Filipe AF, Fornaroli R, Graf W, Herder J, Van der Hoorn B, Iwan Jones J, Sagova-Mareckova M, Moritz C, Barquín J, Piggott JJ, 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 DOI 10.1016/j.scitotenv.2018.05.002.
- Peng Y, Leung HC, Yiu SM, Chin FY. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28(11):1420–1428 DOI 10.1093/bioinformatics/bts174.
- Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. 2017. Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology* 35(9):833–844 DOI 10.1038/nbt.3935.
- Radulovici AE, Vieira PE, Duarte S, Teixeira MAL, Borges LMS, Deagle BE, Majaneva S, Redmond N, Schultz JA, Costa FO. 2021. Revision and annotation of DNA barcode records for marine invertebrates: report of the 8th iBOL conference hackathon. *Metabarcoding and Metagenomics* 5:e67862 DOI 10.3897/mbmg.5.67862.

- Ransome E, Timmers M, Geller JB, Leray M, Mahardini A, Sembiring A, Collins AG, Meyer CP. 2017. The importance of standardization for biodiversity comparisons: a case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia. *PLOS ONE* 12(4):e0175066 DOI 10.1371/journal.pone.0175066.
- R Core Team. 2021. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at <https://www.R-project.org/>.
- Ritari J, Salojärvi J, Lahti L, De Vos WM. 2015. Improved taxonomic assignment of human intestinal 16S rRNA sequences by a dedicated reference database. *BMC Genomics* 16(1):1056 DOI 10.1186/s12864-015-2265-y.
- Robinson CV, Porter TM, Maitland VC, Wright MTG, Hajibabaei M. 2022. Multi-marker metabarcoding resolves subtle variations in freshwater condition: bioindicators, ecological traits, and trophic interactions. *Ecological Indicators* 145:109603 DOI 10.1016/j.ecolind.2022.109603.
- Sinniger F, Pawlowski J, Harii S, Gooday AJ, Yamamoto H, Chevalloné 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 DOI 10.3389/fmars.2016.00092.
- Srivathsan A, Lee L, Katoh K, Hartop E, Kutty SN, Wong J, Yeo D, Meier R. 2021. ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone. *BMC Biology* 19(1):217 DOI 10.1186/s12915-021-01141-x.
- Stevens BM, Creed TB, Reardon CL, Manter DK. 2023. Comparison of Oxford Nanopore Technologies and Illumina MiSeq sequencing with mock communities and agricultural soil. *Scientific Reports* 13(1):9323 DOI 10.1038/s41598-023-36101-8.
- Steyaert M, Priestley V, Osborne O, Herraiz A, Arnold R, Savolainen V. 2020. Advances in metabarcoding techniques bring us closer to reliable monitoring of the marine benthos. *Journal of Applied Ecology* 57(11):2234–2245 DOI 10.1111/1365-2664.13729.
- Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* 21(8):2045–2050 DOI 10.1111/j.1365-294X.2012.05470.x.
- Thalinger B, Wolf E, Traugott M, Wanzenböck J. 2019. Monitoring spawning migrations of potamodromous fish species via eDNA. *Scientific Reports* 9(1):15388 DOI 10.1038/s41598-019-51398-0.
- Theissinger K, Fernandes C, Formenti G, Bista I, Berg PR, Bleidorn C, Bombarely A, Crottini A, Gallo GR, Godoy JA, Jentoft S, Malukiewicz J, Mouton A, Oomen RA, Paez S, Palsbøll PJ, Pampoulie C, Ruiz-López MJ, Secomandi S, Svardal H, Theofanopoulou C, de Vries J, Waldvogel A-M, Zhang G, Jarvis ED, Bálint M, Ciofi C, Waterhouse RM, Mazzoni CJ, Höglund J, Aghayan S, Alioto TS, Almudi I, Alvarez N, Alves PC, Amorim do Rosario, IR, Antunes A, Arribas P, Baldrian P, Bertorelle G, Böhne A, Bonisoli-Alquati A, Boštjančić L, Boussau B, Breton CM, Buzan E, Campos, Paula F. AU - Carreras, Carlos AU - Castro LFC, Chueca

- LJ, Čiampor F, Conti E, Cook-Deegan R, Croll D, Cunha MV, Delsuc F, Dennis AB, Dimitrov D, Faria R, Favre A, Fedrigo OD, Fernández R, Ficetola GF, Flot J-F, Gabaldón T, Agius DR, Giani AM, Gilbert MTP, Grebenc T, Guschanski K, Guyot R, Hausdorf B, Hawlitschek O, Heintzman PD, Heinze B, Hiller M, Husemann M, Iannucci A, Irisarri I, Jakobsen KS, Klinga P, Kloch A, Kratochwil CF, Kusche H, Layton KKS, Leonard JA, Lerat E, Liti G, Manousaki T, Marques-Bonet T, Matos-Maraví P, Matschiner M, Maumus F, Mc Cartney AM, Meiri S, Melo-Ferreira J, Mengual X, Monaghan MT, Montagna M, Mysłajek RW, Neiber MT, Nicolas V, Novo M, Ozretić P, Palero F, Pârvulescu L, Pascual M, Paulo OS, Pavlek M, Pegueroles C, Pellissier L, Pesole G, Primmer CR, Riesgo A, Rüber L, Rubolini D, Salvi D, Seehausen O, Seidel M, Studer B, Theodoridis S, Thines M, Urban L, Vasemägi A, Vella A, Vella N, Vernes SC, Vernesi C, Vieites DR, Wheat CW, Wörheide G, Wurm Y, Zammit G. 2023. How genomics can help biodiversity conservation. *Trends in Genetics* 39(7):545–559 DOI 10.1016/j.tig.2023.01.005.
- Toxqui Rodríguez M del S, Vanhollebeke J, Derycke S. 2023. Evaluation of DNA metabarcoding using Oxford Nanopore sequencing for authentication of mixed seafood products. *Food Control* 145:109388 DOI 10.1016/j.foodcont.2022.109388.
- Van den Bulcke L, De Backer A, Ampe B, Maes S, Wittoeck J, Waegeman W, Hostens K, Derycke S. 2021. Towards harmonization of DNA metabarcoding for monitoring marine macrobenthos: the effect of technical replicates and pooled DNA extractions on species detection. *Metabarcoding and Metagenomics* 134:e71107 DOI 10.3897/mbmg.5.71107.
- Van den Bulcke L, De Backer A, Wittoeck J, Beentjes K, Maes S, Christodoulou M, Martinez Arbizu P, Sapkota R, Van der Hoorn B, Winding A, Hostens K, Derycke S. 2023. DNA metabarcoding on repeat: Sequencing data of marine macrobenthos are reproducible and robust across labs and protocols. *Ecological Indicators* 150:110207 DOI 10.1016/j.ecolind.2023.110207.
- Van der Loos LM, Nijland R. 2021. Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology* 30(13):3270–3288 DOI 10.1111/mec.15592.
- Van Der Reis AL, Beckley LE, Olivar MP, Jeffs AG. 2023. Nanopore short-read sequencing: a quick, cost-effective and accurate method for DNA metabarcoding. *Environmental DNA* 5(2):282–296 DOI 10.1002/edn3.374.
- Voorhuijzen-Harink MM, Hagelaar R, Van Dijk JP, Prins TW, Kok EJ, Staats M. 2019. Toward on-site food authentication using nanopore sequencing. *Food Chemistry: X* 2:100035 DOI 10.1016/j.fochx.2019.100035.
- von Ammon U, Wood SA, Laroche O, Zaiko A, Tait L, Lavery S, Inglis GJ, Pochon X. 2018. Combining morpho-taxonomy and metabarcoding enhances the detection of non-indigenous marine pests in biofouling communities. *Scientific Reports* 8(1):16290 DOI 10.1038/s41598-018-34541-1.
- Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* 73(16):5261–5267 DOI 10.1128/AEM.00062-07.

- Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. 2021. Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology* 39(11):1348–1365 DOI 10.1038/s41587-021-01108-x.
- Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MF, Grabowski M, Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaovičová Z, Ferreira S, Dijkstra K-DB, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, van der Hoorn BB, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher JN, Tomasz M, Paz G, Pešić V, Pfannkuchen DM, Pfannkuchen MA, Price BW, Rinkevich B, Teixeira MAL, Várbíró G, Ekrem T. 2019. DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. *Science of The Total Environment* 678:499–524 DOI 10.1016/j.scitotenv.2019.04.247.
- Willassen E, Westgaard J-I, Kongsrud JA, Hanebrekke T, Buhl-Mortensen P, Holte B. 2022. Benthic invertebrates in Svalbard fjords—when metabarcoding does not outperform traditional biodiversity assessment. *PeerJ* 10:e14321 DOI 10.7717/peerj.14321.
- Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30(5):614–620 DOI 10.1093/bioinformatics/btt593.
- Zhang T, Li H, Ma S, Cao J, Liao H, Huang Q, Chen W. 2023. The newest Oxford Nanopore R10.4.1 full-length 16S rRNA sequencing enables the accurate resolution of species-level microbial community profiling. *Applied and Environmental Microbiology* 89(10):e00605-23 DOI 10.1128/aem.00605-23.
- Zhao B, Van Bodegom PM, Trimpos KB. 2023. Environmental DNA methylation of *Lymnaea stagnalis* varies with age and is hypermethylated compared to tissue DNA. *Molecular Ecology Resources* 23(1):81–91 DOI 10.1111/1755-0998.13691.
- Zhou X, Li Y, Liu S, Yang Q, Su X, Zhou L, Tang M, Fu R, Li J, Huang Q. 2013. Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *GigaScience* 2(1):2047–217X–2–4 DOI 10.1186/2047-217X-2-4.