ICES Journal of Marine Science

ICES Journal of Marine Science (2021), 78(9), 3411–3427. https://doi.org/10.1093/icesjms/fsaa256

Contribution to the Themed Section: *'Patterns of biodiversity of marine zooplankton based on molecular analysis'*

Metabarcoding reveals hidden species and improves identification of marine zooplankton communities in the North Sea

Ilias Semmouri [1]*, Karel A. C. De Schamphelaere[1], Stijn Willemse[1], Michiel B. Vandegehuchte[2], Colin R. Janssen[1,3], and Jana Asselman[1,3]

[1]*Laboratory of Environmental Toxicology and Aquatic Ecology, Faculty of Bioscience Engineering, Ghent University, Ghent 9000, Belgium*
[2]*Flanders Marine Institute VLIZ, InnovOcean Site, Wandelaarkaai 7, Ostend 8400, Belgium*
[3]*Blue Growth Research Lab, Ghent University, Bluebridge Ostend Science Park, Wetenschapspark 1, Ostend 8400, Belgium*

**Corresponding author: tel: + 32 9 264 89 19; e-mail:* ilias.semmouri@ugent.be.

Semmouri, I., De Schamphelaere, K. A. C., Willemse, S., Vandegehuchte, M. B., Janssen, C. R., and Asselman, J. Metabarcoding reveals hidden species and improves identification of marine zooplankton communities in the North Sea. – ICES Journal of Marine Science, 78: 3411–3427.

Received 2 September 2020; revised 6 November 2020; accepted 29 December 2020; advance access publication 28 February 2021.

Although easily collected in large numbers, the subsequent processing and identification of zooplankton have usually been a barrier to large-scale biodiversity assessments. Therefore, DNA barcoding has been increasingly used by non-taxonomists to identify specimens. Here, we studied the community composition of zooplankton in the Belgian part of the North Sea over the course of 1 year. We identified zooplankton using both a traditional approach based on morphological characteristics and by metabarcoding of a 650 bp fragment of the V4-V5 region of the 18S rRNA gene using nanopore sequencing. Using long rDNA sequences, we were able to identify several taxa at the species level, across a broad taxonomic scale. Using both methods, we compared community composition and obtained diversity metrics. Diversity indices were not significantly correlated. Metabarcoding allowed for comparisons of diversity and community composition, but not all groups were successfully sequenced. Additionally, some disparities existed between relative abundances of the most abundant taxa based on traditional counts and those based on sequence reads. Overall, we conclude that for zooplankton samples, metabarcoding is capable of detecting taxa with a higher resolution, regardless of developmental stage of the organism. Combination of molecular and morphological methods results in the highest detection and identification levels of zooplankton.

Keywords: 18S barcode, biomonitoring, copepods, Minlon, nanopore sequencing, *Temora longicornis*

Introduction

Growing threats to biodiversity such as pollution and climate change (Sánchez-Bayo and Wyckhuys, 2019; Reddin *et al.*, 2020) highlight the importance of obtaining accurate baseline measurements of the current biodiversity. However, detailed studies (including spatial and temporal data) on zooplankton dynamics and/or biodiversity declines are limited, compared to other groups such as vertebrates or insects. Zooplankton constitutes an important link between phytoplankton and higher trophic levels in the pelagic zone (e.g. several pelagic fish species and marine mammals) and to the benthic zone through sedimentation of faecal pellets (Wexels Riser *et al.* 2002; Van Ginderdeuren *et al.*, 2014). Moreover, it contributes to the role of marine systems as sources or sinks of CO_2 and other greenhouse gasses through the formation of marine snow, among others (Kiørboe, 2001). Despite their ecological importance, spatiotemporal variations in

the composition of zooplankton assemblages is not well described, primarily due to challenging taxonomic identification (Djurhuus *et al.*, 2018). As planktonic assemblages are known to rapidly respond to environmental variation (i.e. by changes in community structure, biomass, and trophic linkages in marine food webs), they are applicable as bio-indicators of ecosystem change (Ferdous and Muktadir, 2009; Chiba *et al.*, 2018).

Studying the spatial and temporal dynamics of zooplankton diversity can be an important method for assessing the marine environmental status and assisting mitigation and restoration efforts. Zooplankton biomass and diversity is an Essential Ocean Variable in the Global Ocean Observing System (Batten *et al.*, 2016). Several strategies for biodiversity assessments exist, each with their own strengths and weaknesses. Classical microscopy methods require a high degree of taxonomic expertise and are time-consuming. Correctly identifying taxonomically similar (cryptic) species, or juvenile life stages, can be challenging using solely morphological characteristics and is, hence, highly dependent on the expertise of the investigator (Pomerantz *et al.*, 2018). Classical zooplankton diversity studies in the marine environment have been carried out already on a global scale since the early 20th century (e.g. Batten *et al.*, 2003; O'Brien *et al.*, 2017). Digital image capturing/processing (such as the ZOOSCAN digital imaging) and machine learning methods are automated techniques for zooplankton identification, allowing both a high throughput and quantitative results (abundance, biomass, size ranges), yet the taxonomic resolution of these techniques is still considerably lower than that of a trained taxonomist (Grosjean *et al.*, 2004; Leow *et al.*, 2015). With the advent of molecular biology, DNA-based species identification methods and protein profiling (proteome fingerprinting) have been developed and used. The latter is established on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (in short MALDI-TOF MS) techniques, which requires the development of an accessible proteome profile reference library. Such a reference library is currently not yet available for Metazoa (Riccardi *et al.* 2012; Laakmann *et al.*, 2013). In contrast, DNA-based methods relying on reference libraries have been successfully developed using molecular markers, both mitochondrial and nuclear. Metabarcoding (i.e. the large-scale taxonomic identification of a complex sample of multiple specimens via the analysis of one or more orthologous DNA regions, the so-called barcodes) has the significant advantage of detecting the "hidden diversity" of communities (Bucklin *et al.*, 2011; Taberlet *et al.*, 2012; Trivedi *et al.*, 2016). Metabarcoding may also be useful to discover new species, to reveal cryptic species, to detect non-native and/or invasive species, and to assess taxonomically significant variation within species with broad or disjoined distributions (Bucklin *et al.*, 2011; Darling and Mahon 2011; Lawson Handley, 2015). However, metabarcoding, as implemented today, still has several limitations as well: the inability to discriminate between life stages (if necessary, this is a double-edged sword), the possibility of incorrect species identification in a taxonomic reference database and its requirement for PCR (DNA amplification), which can introduce noise into the data, as polymerases do not replicate DNA perfectly, and chimeric sequences (Gaspar and Thomas 2013; Brown et al., 2015). The mtDNA cytochrome c oxidase subunit I (COI) marker is contemplated the universal animal barcode, while the nuclear small subunit (18S) marker is often a standard regarding marine microbial eukaryotic diversity (Hebert *et al.*, 2003; Amaral-Zettler *et al.*, 2009; Huang *et al.*, 2018) and has been used for analysis of zooplankton samples as well (Bucklin *et al.*, 2016). Some recent freshwater and marine studies have already demonstrated the suitability of metabarcoding in macroinvertebrate and zooplankton communities as an alternative method for the identification of zooplankton samples (Carew *et al.*, 2013; Lindeque *et al.*, 2013; Pearman and Irigoien, 2015; Carroll *et al.*, 2019). However, limited work has been done on applying metabarcoding to study seasonal variation in marine zooplankton (e.g. Chain *et al.*, 2016). Using metabarcoding data, and imaging or flow cytometry abundances, along the already other available data of the *Tara* Oceans project, Ibarbalz *et al.* (2019) studied latitudinal gradients and global predictors of plankton diversity. They found that planktonic diversity—across archaea, bacteria, eukaryotes, and major virus clades—is more important around the equator, and decreases towards the poles. The existence of such a latitudinal diversity gradient is well established for most terrestrial organisms, being described by Alexander von Humboldt 200 years ago and is now being proven valid for most planktonic groups in the global ocean as well (Tara Expeditions, 2020). Only recently, a few studies (e.g. Berry *et al.*, 2019; Bucklin *et al.*, 2019) have explored whether metabarcoding is capable of detecting both temporal and spatial variation in the diversity of marine zooplankton on a local/regional scale, even though this type of information is crucial to understand biodiversity processes. By doing so, Berry *et al.* (2019) were able to identify consistent seasonal assemblages of zooplankton species and, moreover, they were able to detect clear departures from these observed regular seasonal patterns, during a marine heatwave.

The Belgian part of the North Sea (BPNS) is a relatively small (3454 $km^{2)}$ and shallow area (maximum 35 m deep) with a coastal length of ~66 km (Mortelmans *et al.*, 2019). The area is characterized by several harbours and a complex of sandbanks and includes several valuable habitats, such as biogenic reefs formed by aggregations of the bristle worm, *Lanice conchilega* (Rabaut *et al.*, 2007). Furthermore, the BPNS is one of the most intensively used seas in the world and is exploited by a diversity of activities, including aggregate extraction, fisheries, aquaculture, tourism, wind energy development, shipping, and dredging (Pecceu *et al.*, 2016). This makes the BPNS a relevant case study on how anthropogenic activities affect biodiversity. In the BPNS "classical" mesoscale zooplankton surveys were carried out by Van Ginderdeuren *et al.* (2014) between 2009 and 2010 and the LifeWatch zooplankton surveys between 2012 and 2019 (Flanders Marine Institute, 2019). However, due to the chosen methods in these studies, a detailed overview of the distribution and richness of the zooplankton community in the BPNS is still lacking, in particular for morphologically challenging taxa: larval stages, parasites, and cryptic species have so far been lumped together. Moreover, due to time constraints and limitations of the ZOOSCAN, the LifeWatch surveys are only capable of monitoring zooplankton at the phylum (and for some taxa at the order) level.

Therefore, in this study, we examined the spatial and temporal distribution of the zooplankton assemblage of the BPNS during a 1-year period. A 650 bp fragment of the V4 and V5 region of the 18S rDNA barcode was characterized using the MinION^{TM}, a nanopore-based DNA/RNA sequencing platform (Oxford Nanopore Technologies, 2020a, b). This technology has multiple advantages over traditional sequencing technologies, including long-read output, rapid real-time analysis, and low initial start-up costs, relative to other commercial sequencers (Pomerantz

et al., 2018). Previous studies were able to demonstrate that data obtained through nanopore technology—by sequencing nearly full-length 16S rRNA gene amplicons—are adequate for the study of microbial communities (e.g. Benitez-Paez *et al.*, 2016; Benitez-Paez and Sans, 2017; Cuscó et al., 2019). Metabarcoding studies of eukaryote communities using nanopore sequencing, as far as we know, are not available. The aim of the present work is to establish a method for capturing the biodiversity of the zooplankton community across space and time in the BPNS using nanopore sequencing. We aim to identify the dominant zooplankters in the community as they present higher biomass, higher abundance, and higher occurrence rates and reflect the characteristics of a zooplankton community structure better than other species (Xu *et al.*, 2016).

Material and methods

Sampling

Zooplankton were collected with the research vessel (RV) Simon Stevin during one sampling campaign in 2018 (15th of March), and four subsequent campaigns in 2019 (20th of February, 23rd of April, 24th of June, and 21st of August) at three stations in the BPNS, NE Atlantic Ocean. Within the BPNS one nearshore and two offshore stations were selected, of which the nearshore station (130) is located close (>5 km) to the harbour of Ostend, Belgium (Figure 1). Exact GPS locations for each station and physical seawater properties per collection date are provided in Table 1. At each sampling station, temperature and salinity data were collected (Table 1). Zooplankton were collected by means of a vertical WP2 plankton net, consisting of a 57-cm diameter steel ring, a 2.6-m long net with a 200 µm mesh size, a plastic bucket to collect the sample, a heavyweight to prevent uplifting, and a mechanical flowmeter. The entire sea column is sampled as the WP2 net is brought down to just above the sea floor, before being hauled up again slowly (max. speed of 1 m/s, total duration ~3 min.). The sample is concentrated in the plastic net bucket where zooplankton remains while most of the water escapes through a 200 µm mesh window. One-third of each sample (after determination of the total volume of the sample, using a measuring cylinder) was used for DNA extraction, while the remaining aliquots of the sample were used for morphological identification. Within 5 min, specimens were stored in freezers at −16°C on the RV and subsequently at −85°C in the laboratory. In the laboratory, samples for species identification were thawed and zooplankton were identified to the lowest taxonomic level possible using a stereomicroscope (Leica MZ 10) and counted by only one person (the first author of this study) to avoid possible bias. Based on the measured flow and the counted abundances, densities (ind.m^3) could be calculated for each sample. *Oikopleura* sp. (tunicate) and *Noctiluca scintillans* (dinoflagellate) specimens, if present in the samples, were not counted and were omitted in the data analyses (cf. Van Ginderdeuren *et al.*, 2014) because of two reasons. First, due to the gelatinous nature of their body, preservation and transport proved to be more difficult than expected and many specimens were shrunk or ruptured when identifying, making it challenging to assess true densities for these taxa. Second, in particular for *Noctiluca*, this taxon is very rare (absent) throughout the year, followed by a massive bloom, reaching peak densities of 39 800 ind.m^{-3} in the BPNS, hence skewing data analyses (Van Ginderdeuren *et al.*, 2014).

In order to evaluate the usage of nanopore sequencing technology and analysis tools for the identification of zooplankton samples, a mock community was constructed as a quality control. Artificial communities are a particularly efficient tool to reveal mismatches between the presence, read abundance, and the actual proportion of species in a community (Braukmann *et al.*, 2019). Nine pooled taxa from different taxonomic levels (i.e. at the family and order level in the phyla Arthropoda and Annelida) were added at different densities to disentangle potential biases in density-to-read ratio. Organisms were taken from one single sample. Location and corresponding environmental data are listed in Table 1. A list of species that were included in the artificial community is provided in Supplementary Table S2. The taxa were selected based on (i) their dominance in the Belgian zooplankton community (Van Ginderdeuren *et al.*, 2014; Semmouri *et al.*, 2020) and (ii) the observed mismatch between metabarcoding identifications and microscopy identification in preliminary analyses.

DNA extraction and amplification

Total genomic DNA of all samples was extracted using the Cetyltrimethylammonium Bromide (CTAB) DNA isolation protocol (Chen and Ronald, 1999) with modifications. CTAB buffer was pre-warmed to 65°C and 300 µl was added to the thawed samples. Next, the samples were grinded using an autoclaved pestle and incubated for 1 h at 65°C with vortexing every 20 min. After adding 300 µl of phenol: choroform:isoamyl alcohol (25:24:1) to the samples, the mixture was mixed and centrifuged for 20 min at 15 000 rcf and 4°C after which the organic phase was removed. Then 2 µl of RNase A (5 µg µl^{-1}) was added to each sample and incubated for 20 min at room temperature, to reduce RNA contamination. To this solution 500 µl of chloroform: isoamyl alcohol (24:1) was added to separate additional contaminants. The liquids were mixed and centrifuged at 15 000 rcf for 15 min. This chloroform: isoamyl step was repeated twice. The aqueous phase in the supernatant was transferred into a new tube. DNA was precipitated by adding 27 µl of 3 M Na acetate and 500 µl of 2-propanol. The mixture was gently mixed and incubated for 10 min at room temperature and DNA was pelleted by spinning the tube at 4°C, 10 000 rcf for 15 min. Next, the supernatant was removed and the pellet was washed with 70% ethanol, air dried, and resuspended in 50 µl of TE buffer.

Before amplification, primers were selected based on the available literature. The primer pair F-566/R-1200 [sequences (5′–3′) CAG CAG CCG CGG TAA TTC C and CCC GTG TTG AGT CAA ATT AAG C, respectively] was selected based on its wide taxonomic range (Hadziavdic *et al.*, 2014). Theoretical coverage reached 80% and empirical data showed amplification for 71% of all eukaryotic phyla (Hadziavdic *et al.*, 2014). Nevertheless, we evaluated *in silico* the compatibility between barcodes of marine eukaryote phyla for the 18S rRNA gene with the primer pair as well. The primer set was tested for matches to sequences in the SILVA database (SSU 138; Quast *et al.*, 2012) using TestPrime under the following parameters: maximum number of mismatches of three bases and a length of 0 mismatches at the 3′ end (Klindworth *et al.*, 2013). TestPrime computes coverages for each taxonomic group by running *in silico* PCR on the SILVA database by categorizing the available database sequences into "match", "mismatch", and "no data (sequences not covering the primer match position)" (Klindworth *et al.*, 2013).

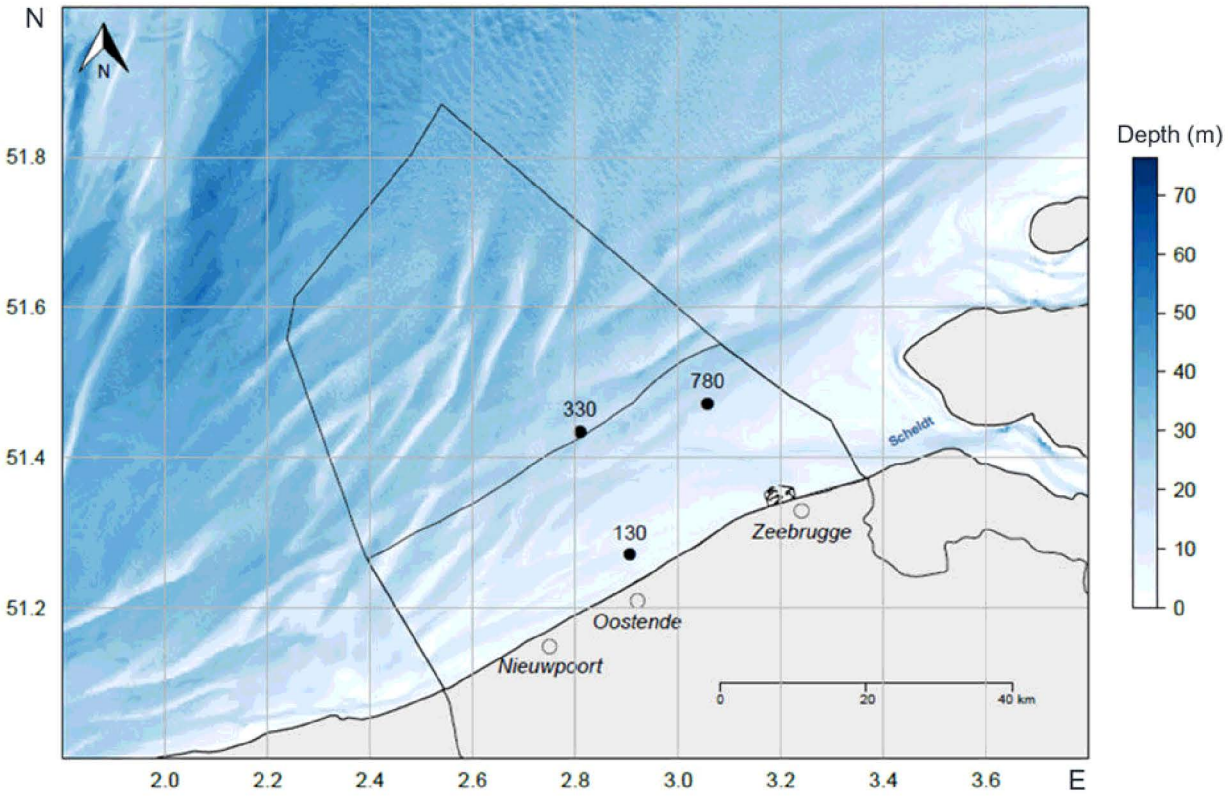

Figure 1. Map showing sampling locations and names of stations in the Belgian Part of the North Sea. Bathymetry in metre (m). Modified from Mortelmans *et al.* (2019) with permission from the authors.

Table 1. Summary of zooplankton sample collections in 2018 and 2019.

Station	*Latitude N*	*Longitude E*	*Date*	*Depth (m)*	*Temperature (°C)*	*Salinity (PSU)*
130	51.271	2.905	15 March 2018	10.81	5.2	32.2
			20 February 2019	9.46	7.9	33.1
			23 April 2019	9.33	13.3	31.4
			24 June 2019	11.74	20.4	32.9
			21 August 2019	9.78	18.9	34.8
330	51.433	2.808	15 March 2018	25.15	6.4	33.9
			20 February 2019	25.71	8.1	34.2
			23 April 2019	25.19	13.6	32.3
			24 June 2019	22.03	20.8	33.3
			21 August 2019	22.45	19.5	35
780	51.440	3.058	20 February 2019	25.48	8.3	34.4
			23 April 2019	21.27	12.4	32.5
			24 June 2019	20.34	19.8	33.0
			21 August 2019	22.92	19.3	34.9
Mock Community	51.433	2.808	21 August 2019	22.45	19.5	35

Additionally, a mock community was constructed as a quality control.

Amplifications were performed in volumes of 50 μl containing 25 μl of ThermoFisher Scientific DreamTaq Hot Start PCR Master Mix (2X), 1 μl of 20 μM forward and reverse primer stock solution, 2 μl of 1/300 diluted high molecular weight DNA template, and 21 μl of DNase free water. PCR amplification on the nuclear 18S gene was carried out with primer pair F-566/R-1200 (Hadziavdic *et al.*, 2014). For the 18S amplification, a Bio-Rad C1000^{TM} Thermal Cycler was set to 95°C for 4 min to activate DNA polymerase. The actual PCR reaction consisted of 35 cycles of melting at 95°C for 30 s, annealing at 44°C for 30 s and extension at 72°C for 1 min, followed by a 10 min final extension at 72°C. PCR amplification results (2 µl/sample) were checked with electrophoresis on 1.5% agarose gel stained with ThermoFisher Scientific DNA gel loading dye (1 µl/sample). PCR success was indicated by the presence of a band for each sample on an agarose gel. Amplicon length was assessed by comparing with 6 μl of ThermoFisher Scientific GeneRuler 100 bp DNA Ladder in the first well. DNA quality was assessed with the NanoDrop^{TM}

spectrophotometer (Version 2000c, Thermo Scientific™, Thermo Fisher Scientific) with the quality standards being based on the 260/230 and 260/280 absorbance ratios. DNA quantity was measured with a Qubit® double-stranded DNA Broad Range (BR) Assay Kit (Life Technologies, Eugene, OR, USA) and read by Qubit® 2.0 Fluorometer (Life Technologies, Eugene, OR, USA). The sample was stored at −80°C until library preparation.

Library preparation and DNA sequencing using the MinION platform

Library preparation and DNA sequencing was performed using the 1D PCR barcoding genomic DNA Sequencing protocol with the SQK-LSK109 kit (Oxford Nanopore Technologies, 2020a), with 1500 ng of genomic DNA used as input for library preparation. Library preparation was carried out according to the manufacturer's protocol with slight adaptations. Briefly, amplicons were end-repaired using 3 µl NEBNext Ultra II End-Repair/dA-tailing enzyme mix (New England Biolabs, USA) at 20°C for 5 min and 65°C for 5 min and then purified with 60 µl of resuspended CleanNGS beads (CleanNA, The Netherlands). Next, adapter ligation was carried out with 25 µl of Ligation Buffer (ONT), 5 µl of Adapter Mix (ONT), and 10 µl of NEBNext Quick T4 DNA Ligase (New England Biolabs, USA) and followed by another purification step using the magnetic beads. The washing step, using 250 µl of Short Fragment Buffer (ONT), was performed twice before the DNA was eluted by incubation in 15 µl of Elution Buffer (ONT) for 10 min. Each flow cell was primed with 1000 µl of priming mixture (ONT). Then, 12 µl of the amplicon library was diluted in 75 µl of running buffer according to protocol. Next, each of the libraries was loaded onto an FLO-MIN106 flow cell via the SpotON sample port and the DNA was sequenced on a MinION device for about 6 h using standard MinIon settings. For base calling, the locally installed software Guppy version v3.4.4 (Oxford Nanopore Technologies, 2020b) was used with standard settings.

Bioinformatics and data analysis

A schematic flow chart of the data processing is presented in Figure 2. After sequencing on the MinION, quality scores and read lengths of all samples were visualized and summarized using NanoPlot V1.15 and NanoStat V1.1.2, respectively (De Coster *et al.*, 2018). Next, the raw reads were trimmed using PoreChop V0.2.3 (Wick *et al.*, 2017) to remove any adapters. Chimeric reads, containing internal adapters, smaller than 600 bp were omitted using the –min_split_read_size command (barcode itself around 650 bp long). We then filtered the reads for quality (Phred score of >12) and read length (>500 bp) using Nanofilt V2.2 (De Coster *et al.*, 2018). The high-quality DNA reads are deposited in the NCBI Short Read Archive (SRA) database under accession number BioProject: PRJNA622488.

To assign taxonomy to the trimmed and filtered reads, we used a local BLAST alignment (Altschul *et al.*, 1990) to compare those reads with a reference database. A local database of over 52 000 18S sequences was compiled, by downloading the records of relevant taxa (i.e. marine species) that are publicly available in GenBank at the National Center for Biotechnology Information database. Relevant taxa were chosen by consulting the WoRMS database (WoRMS Editorial Board, 2020). The database itself can be found in Supplementary data S1. The best BLAST hit against our local database was used to classify each sequence, and a

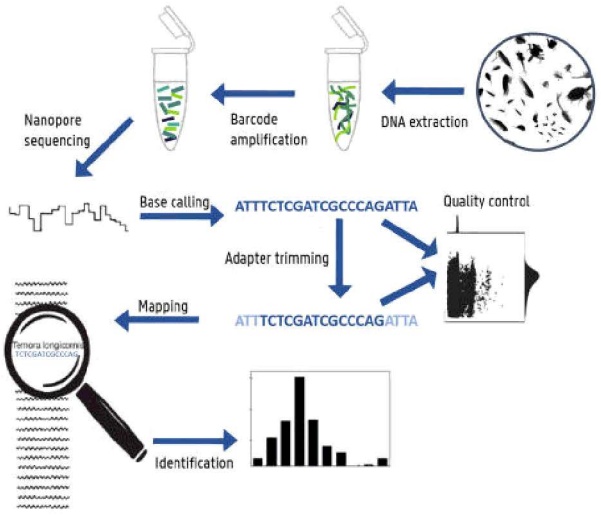

Figure 2. Work flow from sampling to data processing and taxon determination. The details of each step are elaborated in the Material and Methods section.

positive identification was defined as a hit with at least 97% identity, an *e*-value cut-off of 1E−50 and an alignment length of at least 500 nucleotides with a database sequence. To check for non-indigenous taxa and synonyms, we compared all the obtained names to taxa recorded in the Belgian Register of Marine Species (Vandepitte *et al.*, 2010). Names were manually checked and taxonomy was determined based on the following set of rules: (i) synonyms were resolved, (ii) sub-species level identifications were set to species level, and (iii) non-indigenous species were set to genus level. Consequently, we merged the reads with the same taxonomic annotation to compute read abundances.

Prior to downstream analyses, sequence data were normalized by rarefaction (lowest sequence depth = 10 177), to account for the compositional nature of this data type and the large variation in total numbers of reads per sample (Weiss *et al.*, 2017). Next, diversity indices for each sample were calculated, using species richness, the Shannon–Weaver index (Shannon, 1948), the Simpson index, and the inverse Simpson Index (Simpson, 1949) as measures of α diversity for both strategies. The Bray–Curtis distance on square root transformed morphological data was used to estimate the degree of sample dissimilarity, an aspect of β diversity. Similarly, we applied the Bray–Curtis index on the rarefied metabarcoding data set to establish beta diversity. To test the hypothesis that zooplankton communities differ in time and/or space, the Bray–Curtis dissimilarity matrices were ordinated by Multidimensional Scaling (MDS) and the stress level, or measure of the goodness of fit, was calculated for each ordination. Differences among community states were tested by Permutational Multivariate Analysis of Variance (PERMANOVA) with 999 permutations (Anderson, 2001), implemented by the "adonis" function in the vegan package (Oksanen *et al.*, 2015). Here, a Bray–Curtis dissimilarity matrix, computed from square-root transformed abundance or rarefied sequence data, was used to represent community composition. Time of sampling and sampling location were used as predictor variables. Pearson's correlation analysis was used to calculate the correlation coefficient (R^2) between variables. Prior to applying

correlation statistics, data were analysed to verify their normal distributions, using the Shapiro–Wilk test at $p \leq 0.05$. All data were analysed using the statistical computing environment R (v. 3.5.1), using the packages Vegan v.2.5.2 (Oksanen *et al.*, 2015), FactoMineR v.2.0 (Lê *et al.*, 2008), and pheatmap v.1.0.12 (Kolde, 2019).

Finally, for the morphological data set, the spatial distribution and the seasonal variability in the community composition were visualized using a principal component analysis (PCA) and the corresponding correlation circle, after log-transformation. Prior to PCA plot visualization, sequence data were center log-ratio (clr) transformed using the acomp (Aitchison composition) function in the compositions v.2.0.0 package (van den Boogaart and Tolosana-Delgado, 2008) in R. In every correlation circle, each recorded phylum is represented as a vector, which signals the combined strength of the relationships between the taxon and the two PCs (vector length) and whether these relationships are positive or negative (vector direction). The angle between two vectors signals the degree of correlation between two taxa [0° = indicates a complete positive correlation, an adjacent angle indicates a highly positive correlation, an orthogonal angle (90°) indicates that two variables are completely uncorrelated, and opposite angles (180°) indicate a complete negative correlation].

Results and discussion

Nanopore sequencing is applicable for metabarcoding plankton communities

This study used nanopore sequencing technology to assess the species richness and seasonal changes in marine zooplankton assemblages via metabarcoding at three locations of the North Sea through an annual cycle. PCR was successful for 13 out of 15 samples (samples taken at 130—21 August 2019 and 330—24 June 2019 could not be amplified successfully and were omitted in further analyses). After nanopore sequencing, processing and stringent quality filtering of the sequenced reads, we obtained on average 40% of the initial reads, which were of high quality for species identification (Table 2). While this seems significantly lower in comparison to standard sequencing technologies, this is comparable to what has been reported by other recent nanopore studies (e.g. 37% in Menegon *et al.*, 2017; 50% in Semmouri *et al.*, 2020). Furthermore, we set more stringent quality cut-offs in this study to reduce possible bias due to base-calling errors, hence ensuring high-quality species identification.

Errors observed in the output sequences can have three origins: degradation of template DNA, errors during amplification and errors during sequencing, the extent of the latter varying between the different sequencing technologies. Current error rates range reportedly from 5 to 20% (as compared to 1% for Illumina), dependent on the type of molecules and library preparation methods, including both insertions and deletions (Rang *et al.*, 2018; Krehenwinkel *et al.*, 2019). While the lower per base accuracy of the MinIon generates bioinformatic challenges, this problem is being progressively mitigated with advances in technology and bioinformatic tools (Hargreaves and Mulley, 2015; Srivathsan *et al.*, 2018). Additionally, the lower accuracy is partially compensated for in this study by, on the one hand, longer fragments (with a total average length of 622.5 bp; Table 2), which improves the resolution and accuracy of the taxonomy assessment and, on the other hand, the high amount of reads that can be produced per run. In total (for the 13 samples), the quality-controlled data set included 3 239 255 reads, which is considerably larger than the average data set produced by next-generation sequencing (NGS, e.g. 2.0 million reads for 154 samples in Casas *et al.*, 2017; 2.4 million reads for 24 samples in Carroll *et al.*, 2019). In comparison, DNA barcodes obtained with NGS are currently cost-effective (high throughput, high accuracy, and reduced costs) but require long processing times and the sequences have a reduced length, ranging from 50 to 300 bp (Hebert *et al.*, 2018; Santos *et al.*, 2020). On the other hand, dideoxy chain-termination sequencing (also known as Sanger sequencing) is more expensive, but full-length barcodes (up to ~1000) can be obtained fairly quickly (Sanger *et al.*, 1977; Ivanova, *et al.*, 2009). However, Sanger sequencing can only process a single DNA template per sample and multiple templates can cause uninterpretable results (Batovska *et al.*, 2016). A final advantage over the traditional sequencing technologies is the MinION's portability, allowing rapid real-time analysis. Real-time data generation is particularly appealing when rapid (species level) identification is required (Børsting and Morling, 2015; Greninger *et al.*, 2015; Srivathsan *et al.*, 2018). Hence, this technique has already been used at

Table 2. Results of metabarcoding of 18S rRNA loci from all sampling locations.

Station	Sampling time	Base-called reads	Quality-filtered reads	Mean read length	Read length N50	Total bases	Annotated reads	Detected number of taxa
130	15 March 2018	330 328	219 900 (66.6%)	619	617	136 144 263	95 085 (43.2%)	33
	20 February 2019	798 209	150 000 (18.8%)	618	616	94 548 728	48 962 (32.6%)	18
	23 April 2019	746 818	152 802 (20.5%)	619	617	92 621 774	49 085 (32.1%)	13
	24 June 2019	649 040	274 695 (42.3%)	622	621	170 738 689	50 855 (18.5%)	17
	21 August 2019	–	–	–	–	–	–	–
330	15 March 2018	268 595	181 920 (67.7%)	625	624	113 756 269	48 993 (26.9%)	20
	20 February 2019	702 690	153 131 (21.8%)	620	624	94 940 882	10 177 (6.6%)	10
	23 April 2019	603 107	305 911 (50.7%)	616	615	188 368 815	74 650 (24.4%)	15
	24 June 2019	–	–	–	–	–	–	–
	21 August 2019	604 596	187 048 (30.9%)	617	616	115 399 906	17 035 (9.1%)	23
780	20 February 2019	885 208	599 998 (67.8%)	619	618	371 242 935	160 046 (26.7%)	18
	23 April 2019	670 723	150 890 (22.5%)	619	618	93 383 135	48 504 (32.1%)	11
	24 June 2019	927 412	407 545 (43.9%)	651	624	265 395 229	64 214 (15.8%)	15
	21 August 2019	618 296	171 334 (27.7%)	626	623	107 178 763	10 236 (6.0%)	17
MC	Mock Community	604 209	284 081 (47.0%)	624	619	171 229 291	56 418 (19.9%)	6

remote sites outside of conventional labs, including the Antarctic to sequence microbial communities (Quick *et al.*, 2016; Johnson *et al.*, 2017; Pomerantz *et al.*, 2018). It also opens opportunities for use during remote sampling campaigns on RVs where data can be generated almost immediately after sampling, which also avoids DNA degradation during long sample storage on board. For example, Truelove *et al.* (2019) were able to identify the presence of white sharks at sea, offshore of the Hawaiian Islands, using the MinIon sequencer to allow for rapid eDNA sequencing.

Assessment of the potential of the 18S marker for metabarcoding using an artificial mock community

To assess whether (i) nanopore sequencing is capable of sequencing barcodes and (ii) the 18S amplicon is informative in zooplankton studies, we constructed a mock community as a quality control. Metabarcoding successfully identified six out of the nine taxa in the standard (Supplementary Tables S1 and S2). Two dominating taxa in this sample, namely comma shrimps (cumaceans) and the harpacticoid copepod *Euterpina acutifrons* (Dana, 1847), were consistently not identified by metabarcoding, although being present in all samples. Hence, this is not due to sampling biases. Moreover, primer mismatches during the PCR step are unlikely as well, as this was *a priori* checked *in silico*. According to a review by Raupach and Radulovici (2015), no metabarcoding study has focused so far specifically on Cumacea (1500 species), which makes it hard to compare these results to available literature. One more species, the mysid *Schistomysis spiritus* (Norman, 1860) was not detected through metabarcoding in the mock community sample either, although one individual was provided. In retrospect, it seems possible that insufficient amount of DNA of this organism in the total sample was outcompeted by DNA molecules of the other species in the sample. Another possibility is that—even though the overall DNA sample, i.e. the collection of all DNA molecules of all species in the mock community, was in good condition—the DNA molecules from this source were too degraded to be recovered through PCR, as this taxon could be successfully identified in the environmental samples. Examination of the relationship between the read count for each species and its frequency in the mock community did not reveal a significant, positive relationship after normalization of the data, i.e. log-transformation ($R^2 = 0.1006$, $p = 0.4$; Supplementary Figure S1).

The 18S barcode is less variable than other barcoding genes and, therefore, has not received as much attention as the mitochondrial gene COI. This might be due to its lower evolutionary rate, which makes it harder to distinguish between closely related taxa or because of a more difficult alignment of 18S, due to a higher occurrence of indels (Machida and Tsuda, 2010; Machida and Knowlton, 2012). Conversely, a lower mutation rate means that universal primers are easier to design. In our study, the use of the 18S rRNA gene allowed successful amplification of barcodes for evolutionary diverse metazoan phyla, ranging from cnidarians to chordates. However, notable exceptions were cumaceans and harpacticoid copepods in this study (as established analysing the results from the mock community). In contrast, while the use of the COI barcode provides relatively poor resolution at higher taxonomic levels in some cases, it has a high resolution at species level, is useful for revealing intraspecific variation and has, hence, been widely used as the barcode for species identification (Lefebure *et al.*, 2006; Tang *et al.*, 2012; Wu *et al.*, 2015; Huang *et al.*, 2018). Although some studies do not support using the 18S rRNA gene for species delimitation (e.g. Tang *et al.*, 2012), Laakmann *et al.* (2013) considered the complete 18S rRNA gene effective for the species-level identification of calanoid copepods, the dominant group of zooplankters in the BPNS. However, in our entire molecular 18S dataset, there were two copepod taxa exhibiting mixing in identification (*Calanus glacialis* with *Calanus helgolandicus*, *Centropages abdominalis* with *Centropages hamatus/Centropages typicus*). These varying degrees of agreement of the *Centropages* congeners has also been described by Young *et al.* (2017) for the COI barcode. The genus *Centropages* has been noted as having discordant molecular clustering as compared to morphological identifications (Blanco-Bercial *et al.*, 2014; Young *et al.*, 2017). Moreover, the calanoid family Centropagidae has been reported to have a plastic response to changing environments, thereby making morphological identifications more difficult as well (Blanco-Bercial *et al.*, 2011). Also, *Calanus* has been described as problematical for barcoding (e.g. Kozol *et al.*, 2012; Blanco-Bercial *et al.*, 2014). Similarly, for a limited number of non-copepod taxa (*Crangon*, *Ophiotrix*, *Sagitta*) a non-indigenous congener was identified as the best-fitting BLAST hit in our data set, even though indigenous congeners were present in the database. In such cases, identification was set to the genus level. The identification of these non-indigenous congeners might be a consequence from the higher error rates of base-called nanopore reads, possibly affecting taxonomic assignment. As a safety principle, the annotation of those reads was raised to the genus level. In doing so, we are possibly overlooking the presence of alien species in our community, but we avoid false positives. Using nanopore sequencing, it is possible, under both relaxed and strict filtering conditions, to retrieve annotations at the species level if the species in a sample are sufficiently genetically distant from another. However, as is shown here, the high error rate of raw reads may cause problems during read clustering and taxon assignments for closely related species (e.g. as in some congeneric species). One possible way to circumvent these challenges in future studies is by increasing the query sequence length. The F-566/R-1200 primer set only yields a ~650 bp region of the gene. A second method to improve identification is linked to the quality of the reads (reducing base-calling errors).

Finally, accurate taxonomic assignment in metabarcoding studies relies heavily on the quality of the reference database to compare query sequences against. Currently, the NCBI GenBank database hosts the greatest overall sequence data, hence, the reason being the source of our own database. For most of these barcodes, however, the associated specimen collection metadata is often missing. Moreover, also non-specialists are able to upload sequences, hereby possibly creating assignation errors, often leading to a snowball effect in later studies. As a consequence, there are several examples of barcodes being annotated with the incorrect species, or multiple morpho-species are assigned to the same DNA barcode, ignoring the existence of species complexes (Piper *et al.*, 2019). While we did check our database for obvious assignation errors (e.g. cnidarian barcodes being annotated as copepod sequences), more subtle errors cannot be checked for, which, by consequence, may possibly affect our results. Here lies an opportunity for future barcoding studies in improving the quality and curation of barcode databases, possibly applying ONT technology in doing so.

Metabarcoding identified more taxa than morphological observations

For the environmental samples, DNA metabarcoding resulted in the identification of 10–33 taxa based on 10 236 to 160 046 reads (Table 2). Morphological observations identified 8–14 morpho-taxa. All detected genera/species by barcoding (Supplementary Table S1) are listed in the Belgian Register of Marine Species (Vandepitte *et al.*, 2010), except for the recently invaded calanoid copepod *Pseudodiaptomus marinus* Sato, as described by Deschutter *et al.* (2018). Over 75% of the identified barcodes were related to copepod sequences (Figure 3). Copepoda, Cirripedia (barnacles), and Mysida (opossum shrimps) represented the three major taxa in terms of both read numbers and frequency of morphologically determined specimens (Supplementary Table S1, Figure 3). However, branchiopods (*Podon* and *Evadne* spp.) were not picked up by the metabarcoding approach, even though they were also abundant in some samples. The number of taxa, identified by metabarcoding, always exceeded the number of taxa identified by microscopy (Figure 4a, S2). Hence, the present study suggests that metabarcoding has a higher sensitivity than that of microscopy. This is likely due to the capacity to take into account individuals at any life stage, such as eggs or (nauplius) larvae (of which morphological identification is very complicated or impossible, e.g. Cirripedia, Polychaeta), confirming previous studies (Pochon *et al.*, 2013; Zhan *et al.*, 2013). While Polychaeta larvae were lumped into one taxon due to inconsistent identification using morphology, metabarcoding was able to annotate these reads to the genus/species level, revealing a relatively large diversity. As in other metabarcoding studies (Brandon-Mong *et al.*, 2015; Ficetola *et al.*, 2015), some species were only detected through metabarcoding (e.g. some Actinopterygii and Bryozoa), likely reflecting environmental DNA (eDNA) or possibly contamination during sample processing. However, if the latter were the case, we argue that contaminations should also have been found in the mock community (quality step) and this was not the case. The number of taxa detected by metabarcoding in each sample did not have a significant correlation ($R^2 = 0.05$, $p = 0.48$) with that obtained through morphological identification (Figure 4a).

However, we found a significant positive correlation ($R^2 = 0.37$, $p < 0.04$) comparing the relative number of reads in a sample with the abundance of the corresponding taxon obtained through morphological identification (Figure 4b). Other studies also report significant positive correlations between total abundance of morpho-taxa and read abundance. Serrana *et al.* (2019), for example, found an $R^2 = 0.48$ ($p = 0.02$) in a study on macro-invertebrate community in streams, while a zooplankton study in lake Tai, China reports an $R^2 = 0.43$ ($p < 0.001$) for the correlation between read and specimen numbers (Yang *et al.*, 2017). Based on these results, it is tempting to presume that barcode abundance reflects the relative biomass of organisms. Some authors have indeed found read numbers to be related to biomass

Figure 3. Stacked bar plot of major groups recovered in the samples (through metabarcoding and morphological identification). Number of detected taxa (lowest annotation of the reads possible) is reported at the top of each bar. Metabarcoding identified consistently more taxa with a higher taxonomic resolution (genus/species level) compared to morphological identification. At a higher taxonomy level, the bulk of barcodes was assigned to copepods.

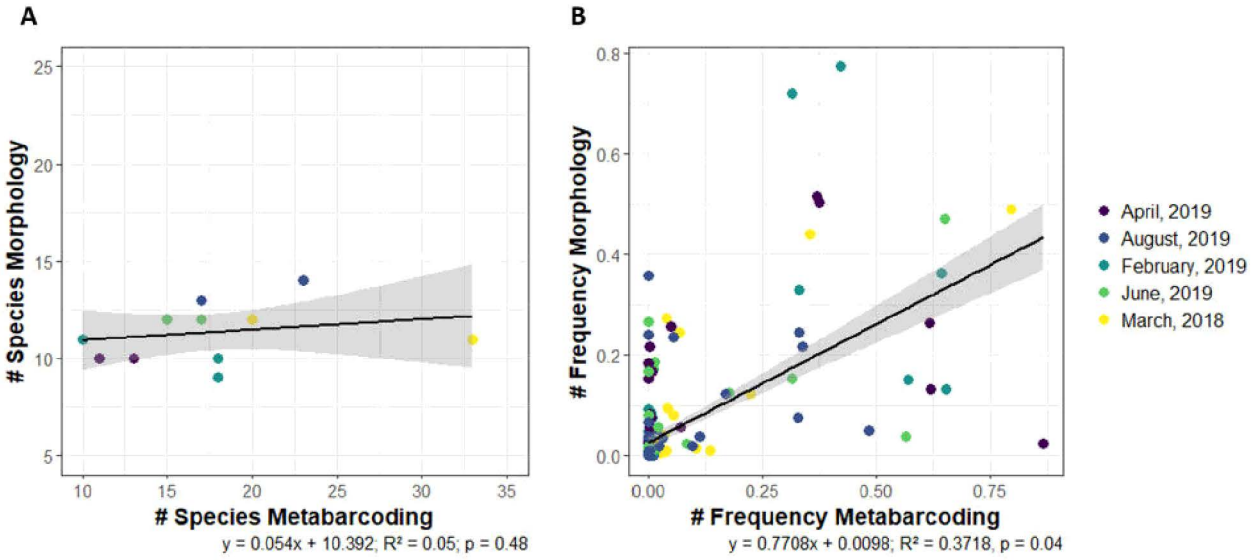

Figure 4. Comparison of zooplankton identification in water samples between metabarcoding and the morphology approach. (a) Species number. (b) The frequency of each taxon in both the barcode and microscopy approach. The R^2 and *p*-value are indicated for each regression axis, as is the 95% confidence interval. In Supplementary Figure S3, graph 4B is also depicted, but here the points are coloured by taxonomic identity.

density, but most studies report this is not the case (Barnes and Turner, 2016; Thomas *et al.*, 2016; Deagle *et al.*, 2018; Carroll *et al.*, 2019). Mainly Mysida and Acartiidae are overrepresented in read numbers, comparing to abundances, while Temoridae are generally underrepresented (Supplementary Figure S3). For this study, we did consider to study the relation between biomass and number of reads (as compared to abundance versus number of reads), yet estimations would definitely be inaccurate due to size differences caused by differences in age, sex, or developmental stage, as we did not measure all specimens when identifying. It is important to affirm that no current metabarcoding protocol allows the estimation of taxon abundances. Differences in total abundance between metabarcoding and morphological assessments are presumably due to differences in the total sampled biomass by each technique. Moreover, the relationship between biomass and number of sequences is not linear as well, since the barcode generation is subject to biases that can be introduced at several steps during the processing of the sample such as DNA extraction, PCR, and bioinformatic processing (Bik *et al.* 2012; Kebschull and Zador, 2015). In addition, the sequence abundance relative to biomass ratio could be inflated by a high gene copy number for different taxa or markers (Klappenbach *et al.*, 2000). Hence, the variable 18S rRNA gene copy numbers between different species (up to four orders of magnitude in eukaryotes, as established by Prokopowich *et al.*, 2003) is another important reason to explain the incapacity to reconstruct relative taxon abundances using barcodes. Morphological analyses measure numerical abundance of organisms including their life stages, whereas the number of reads in metabarcoding analyses depends more on biomass than on numerical abundance and does not yield any information on life stages. The current presence/absence data, generated by metabarcoding, are less prone to potential biases caused by laboratory handling, but are at risk of exaggerating the ecological role of organisms with low abundance or vice versa (Deagle *et al.*, 2018; Carroll *et al.*, 2019). Hence, there is an urgent need to move metabarcoding applications from identification and detection of taxa to their quantification in terms of abundance and/or biomass (Bucklin *et al.*, 2016).

Both metabarcoding and morphology identify the same five major targets

The majority of all reads could be assigned to five taxa, i.e. to the calanoid copepods *Temora longicornis*, *Acartia clausi*, *Centropages* sp., *C. helgolandicus* and *Paracalanus parvus* (Supplementary Table S1, Figure 5). Agreeing with these data, morphology confirmed that these calanoid copepods, along with barnacle larvae, were the dominant zooplankton taxa in the BPNS (Figures 3 and 5, Supplementary Table S1). Our findings correspond with earlier observations in the BPNS made by Daro *et al.* (2006), Deschutter *et al.* (2017), Van Ginderdeuren *et al.* (2014), and Van Meel (1975). Also, in Dutch coastal waters, calanoids *T. longicornis*, *A. clausi*, and *C. hamatus* have been ranked as the stock-forming copepod species (Fransz, 1975). Globally, most marine copepod assemblages are dominated by the same reoccurring genera (e.g. *Calanus*, *Acartia*, *Oithona*, *Paracalanus*), which were all detected in this study as well, making metabarcoding a suitable method for the analysis of copepod communities at other locations as well (Djurhuus *et al.*, 2018). Additionally, in the nearshore sample of March 2018 (station 130), large numbers of the copepod *P. marinus* were found, while being completely absent in other samples (Supplementary Table S2). This was also confirmed by the metabarcoding approach (Supplementary Table S1). On average, Copepoda comprise up to 66% of total zooplankton abundances in the BPNS (Van Ginderdeuren *et al.*, 2014), supporting our finding that copepods are the most abundant taxon in the zooplankton community of the BPNS. Also, barnacle larvae (Cirripedia, Balanomorpha) were relatively abundant in the samples and several species were detected by the barcoding approach. Barcodes were found for *Balanus crenatus* and *Semibalanus balanoides*, two common barnacle species in

Figure 5. Stacked bar plot of five most abundantly detected taxa (calanoid copepods) in all samples through metabarcoding and morphological identification. Percentage refers to normalized number of reads (Metabarcoding) or number of individuals per m^3 (Morphology).

the North Sea, next to barcodes of *Elminius modestus*, an Australasian barnacle species of which its introduction in Europe is well-documented (Kerckhof, 2002). Specimens of *E. modestus* are competing with all the other barnacle species and occurring in a variety of habitats in the littoral and intertidal zone, except for the subtidal offshore region, making it by far the most common barnacle of the Belgian fauna (Kerckhof, 2002). *Balanus crenatus* is probably the only common subtidal barnacle in Belgian waters (Kerckhof, 2002). Conspicuously absent in our barcoding results were Branchiopoda (i.e. *Podon* sp.), while being present in the spring/summer samples (established through morphological identification, Supplementary Table S2). Agreeing with our findings, Van Ginderdeuren *et al.* (2014) reported their absence in winter periods. The detected Brachyura, Mollusca, and Polychaeta reads represent juvenile pelagic stages (Supplementary Table S2, cf. Ellien *et al.*, 2004; Andresen *et al.*, 2014).

Biodiversity estimates based on morphology and metabarcoding are overlapping in range

Biodiversity of the different samples was estimated by calculating Species richness, Shannon index, and Simpson indices (Figure 6). The diversity values for all samples are shown in the Supplementary data (Supplementary Table S3). The individual Shannon indices ranged from ~0.78 to 2.12 and 0.55 to 1.90 for the morphology and metabarcoding approach, respectively (Figure 6). The Shannon diversity index (H) relies on both relative species abundance and species richness in its calculation, having a larger value for samples with a high number of species persisting with a similar abundance (or biomass). For the Simpson index, we found estimates ranging between 0.15 and 0.54 and between 0.24 and 0.79 for microscopy and metabarcoding, respectively. Overall, the biodiversity estimates, as measured by these indices, had an overlapping range when comparing both approaches, with morphology underestimating true species richness and metabarcoding incorrectly quantifying evenness due to dominating read numbers of a few taxa. Despite the comparable ranges, the Shannon and Simpson indices calculated for the metabarcoding data set did not have a significant correlation ($R^2 = 0.06$; $p = 0.45$ and $R^2 = 0.0078$; $p = 0.79$, respectively) with those obtained for the morphological dataset (Supplementary Figure S4). Van Ginderdeuren *et al.* (2014) reported a species richness ranging from 10 to 45 species and a Shannon H, ranging from 0.5 to 2.5 for the zooplankton community in the BPNS. Ibarbalz *et al.* (2019), studying global trends in marine plankton

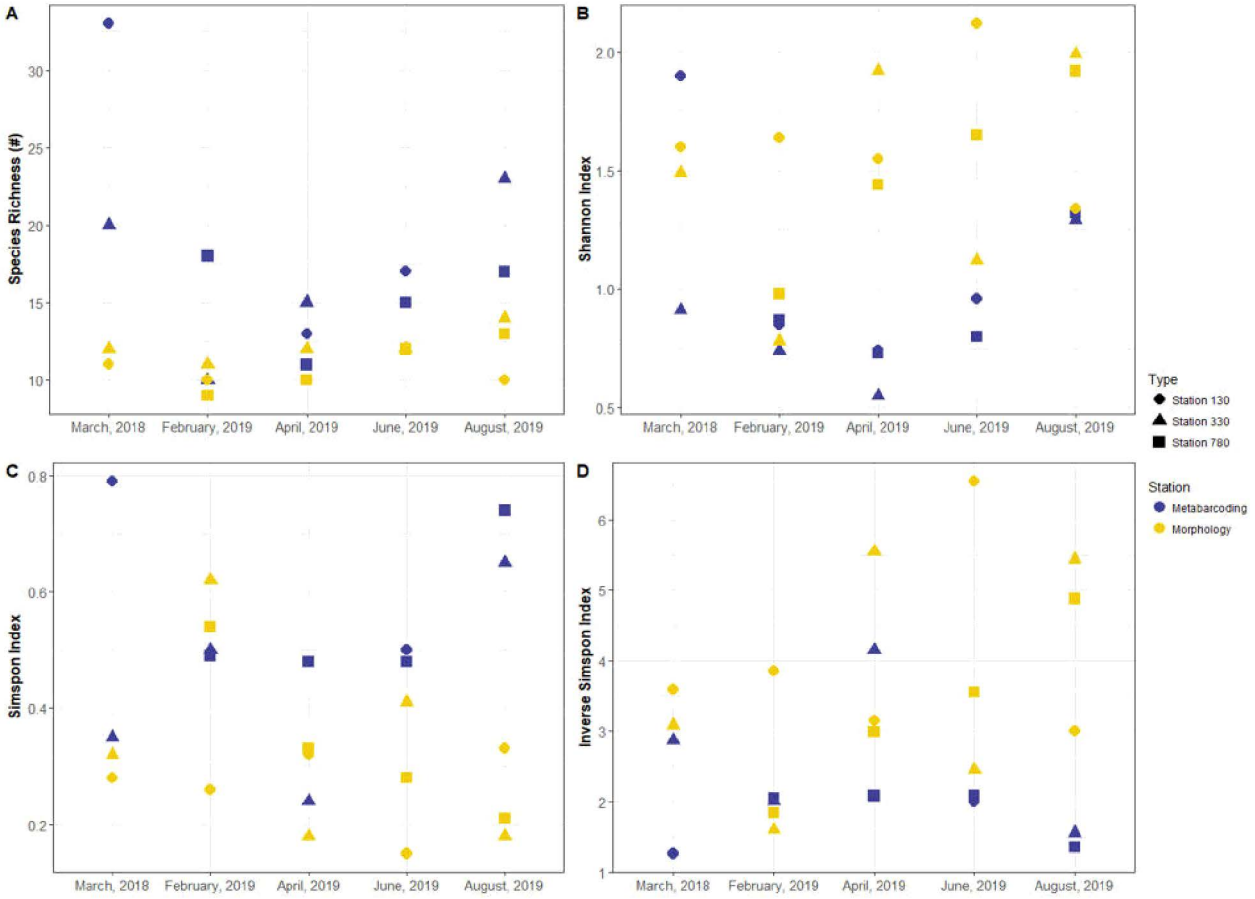

Figure 6. Comparison of four alpha diversity measures in a time series for both the metabarcoding and the morphological approach. (a) Species Richness. (b) Shannon Index. (c) Simpson Index. (d) Inverse Simpson Index.

diversity using data obtained from 12 different marine areas across the world from *Tara* Oceans, reported Shannon Indices ranging between 0.3 and 3.081 for copepods and 0.11 and 7.001 across all kingdoms of life. Our richness estimations for the pelagic crustacean community (Supplementary Table S3) are situated in the mentioned ranges.

Temporal/spatial variation in the zooplankton community

In order to understand how the structure of the zooplankton community changes over time/location, we applied MDS to the Bray–Curtis dissimilarity matrix, as a measure of beta diversity. The Nonmetric MDS ordination indicated no clear separation in the samples (Stress metabarcoding = 0.123; Stress morphology = 0.191; Figure 7). In contrast, using data from the morphology approach, a PERMANOVA analysis revealed that the community composition of the stations is significantly different over time (p-value = 0.026, $R^2 = 0.460$), and indicated that 46% of the temporal variation was explained by the different communities' membership (Supplementary Table S3). Some taxa were only found in the summer samples (e.g. the branchiopod *Podon*) compared to the winter samples. However, we did not find such a pattern in the metabarcoding dataset. Due to the ubiquitous presence in time and space of multiple dominant zooplankton species, the (meso)zooplankton in the BPNS is typified as a single neritic (coastal) zooplankton community, with the addition of some oceanic species, such as *C. helgolandicus* and *C. typicus* in offshore locations (Van Ginderdeuren *et al.*, 2014). Importantly, the analyses of Van Ginderdeuren *et al.* (2014) showed *T. longicornis*, *A. clausi*, and to lesser extent *C. hamatus* and polychaete larvae as the most important contributors to the similarity among the plankton samples in the BPNS. Yet, Van Ginderdeuren *et al.* (2014) found a temporal structure in the zooplankton community of the BPNS, with highest diversity found in spring and summer based on a 2-year sampling campaign. Temperature proved to be the most important variable explaining seasonality in the community structure of the BPNS (Van Ginderdeuren *et al.*, 2014). Similarly, in the global scale study of Ibarbalz *et al.* (2019), temperature emerged as the best predictor of epipelagic plankton diversity with a poleward decline of diversity. In our study, we did find a small positive correlation between species richness and temperature, but only for the morphological data set ($R^2 = 0.38$; $p = 0.04$; Supplementary Figure S5). No significant correlations were found between species richness and salinity (Supplementary Figure S5).

Finally, PCAs were performed and correlation circles were generated for the centre log-ratio transformed read numbers and

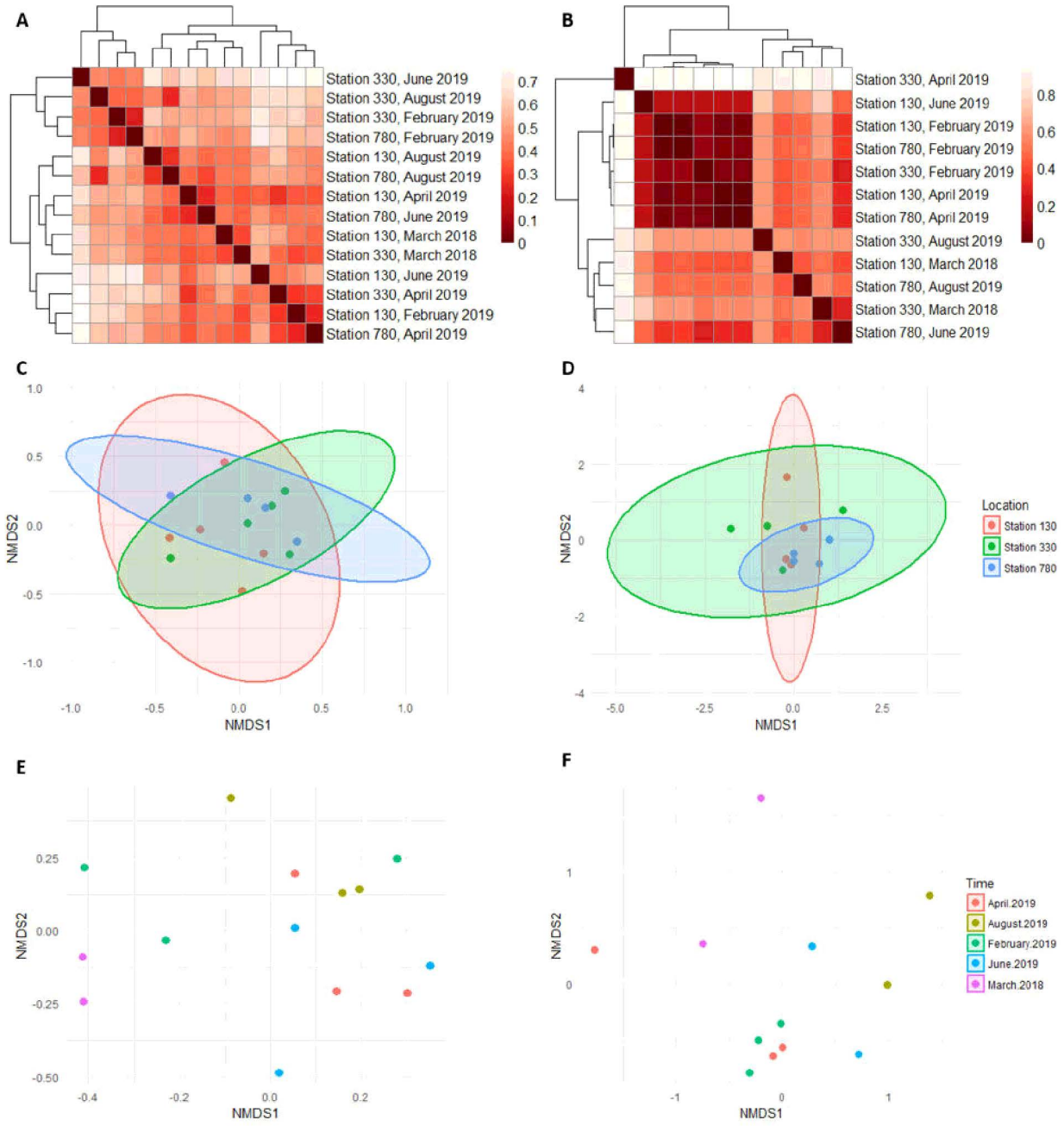

Figure 7. Heatmaps depicting the clustering of samples based on Bray–Curtis dissimilarity for (a) morphological approach and (b) the metabarcoding approach. The degree of similarity of the samples is also shown by the dendrograms on the *x* and *y* axes. (c and d) Beta diversity (Bray–Curtis dissimilarity) ordinated with MDS of zooplankton communities for the morphology and metabarcoding approach, respectively. Each symbol represents one individual sample and each colour represents a station. (e and f) These figures depict the same MDS plots, but here the samples are grouped per date. The samples collected on the same dates did not group together, nor did samples taken at the same location.

the log-transformed densities obtained through the metabarcoding and the morphology approach, respectively (Figure 8). The angles between vectors in the correlation circles differed widely between both approaches indicating that possible relationships among taxa were generally inconsistent among the approaches, as this might be due to the limited data set. For example, Copepoda and Amphipoda were uncorrelated in the metabarcoding approach and highly correlated in the morphology

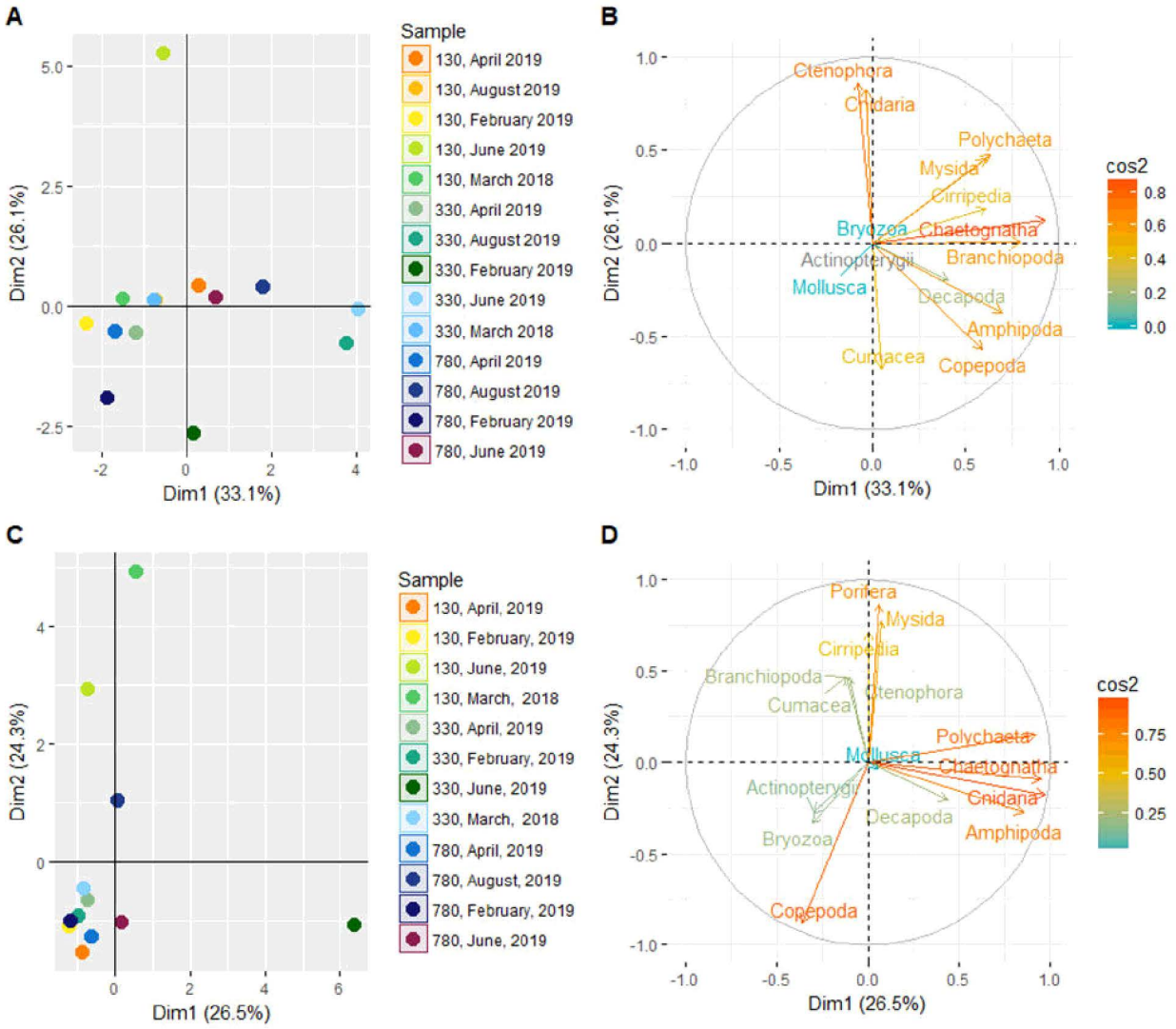

Figure 8. PCA and correlation circle plots generated for the log-transformed zooplankton densities obtained through morphological identification (a and b) and centred log-ratio transformed read numbers obtained through metabarcoding (c and d). In figures A and C, a single point refers to an individual sample taken at a specific time and location, as indicated in the legend. In figures b and d, vectors are the loadings on PC1 (*x*-axis) and PC2 (*y*-axis). The quality of representation of the variables (taxa) is indicated by the squared cosine (cos2). A high cos2 value indicates a good representation of the taxon on the principal component. In this case, the taxon is positioned close the circumference of the correlation circle. A low cos2 indicates the variable is not well represented by the PCs and will be positioned close to the centre of the circle. Vector length indicates the strength of the relationship and the angle between two vectors gives the degree of correlation [adjacent = highly correlated taxa, orthogonal (90°) = uncorrelated taxa, and opposite (180°) = negatively correlated taxa].

approach. This can be explained by the existing discrepancy between read numbers and densities, of which the extent differs for each taxon.

Conclusions

Metabarcoding using the 18S barcode shows promise as a way to study marine biodiversity, as metabarcoding proved to identify taxa at a higher resolution, especially when morphological identification is particularly challenging. At the same time, we observed several limitations: not all taxa were identified by metabarcoding and read numbers did not always reflect counted densities. As a consequence, we did not find significant correlations between the diversity indices obtained for the two data types. Therefore, we conclude that a complementary approach of both metabarcoding and morphological techniques is the best way forward to study marine biodiversity if the goal is to identify as many as taxa as possible at the highest resolution. If the focus lies on only identifying the dominant taxa in a rapid and cost-efficient way, metabarcoding can be used on its own. Although the promise of metabarcoding in acquiring a new and improved understanding for global patterns of plankton diversity, it remains critically important to continue to maintain expertise in morphological taxonomic identification of zooplankton to ensure that metabarcoding approaches can be further validated and

improved. An integrative morphological and molecular approach is proposed for an in-depth assessment and science-based management of the diversity in marine ecosystems.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Data availability

The high-quality DNA reads are deposited in the NCBI SRA database under accession number BioProject: PRJNA622488. All other datasets generated and analysed for this study are included in the article/Supplementary material.

Author contributions

IS, SW, and JA contributed conception and design of the study. Funding acquisition was provided by KDS, CJ, JA, and MV. IS performed the statistical analyses and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Acknowledgements

The authors are grateful to the Flanders Marine Institute (VLIZ) and the LifeWatch project for the assistance, in particular Jonas Mortelmans and the crew of the RV Simon Stevin. The authors also thank Jolien Depecker for the technical assistance. The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO and the Flemish Government—department EWI. Funding was received from FWO (FWO KAN 1521617N awarded to JA).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. Journal of Molecular Biology, 215: 403–410.

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., and Huse, S. M. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS One, 4: e6372.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance: non-parametric manova for ecology. Austral Ecology, 26: 32–46.

Andresen, H., Strasser, M., and van der Meer, J. 2014. Estimation of density-dependent mortality of juvenile bivalves in the Wadden Sea. PLoS One, 9: e102491.

Barnes, M. A., and Turner, C. R. 2016. The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics, 17: 1–17.

Batovska, J., Cogan, N. O., Lynch, S. E., and Blacket, M. J. 2016. Using next-generation sequencing for DNA barcoding: capturing allelic variation in ITS2. G3 (Bethesda, MD), 7: 19–29.

Batten, S., Clark, R., Flinkman, J., Hays, G., John, E., John, A. W., Jonas, T. *et al.* 2003. CPR sampling: the technical background, materials and methods, consistency and comparability. Progress in Oceanography, 58: 193–215.

Batten, S., Chiba, S., and Checkley, D. 2016. Essential Ocean Variables (EOV) for Biology and Ecosystems: Zooplankton Biomass and Diversity. https://goosocean.org/components/com_oe/oe.php?task=download&id=38718&version=1.0&lang=1&format=1 (last accessed 21 August 2020).

Benitez-Paez, A., Portune, K. J., and Sanz, Y. 2016. Species-level resolution of 16S rRNA gene amplicons sequenced through the MinION portable nanopore sequencer. GigaScience, 5: 4.

Benitez-Paez, A., and Sanz, Y. 2017. Multi-locus and long amplicon sequencing approach to study microbial diversity at species level using the MinION portable nanopore sequencer. GigaScience, 6: 1–12.

Berry, T. E., Saunders, B. J., Coghlan, M. L., Stat, M., Jarman, S., Richardson, A. J., Davies, C. H. *et al.* 2019. Marine environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies ecosystem responses to anomalous climatic events. PLOS Genetics, 15: e1007943.

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

Blanco-Bercial, L., Bradford-Grieve, J., and Bucklin, A. 2011. Molecular phylogeny of the Calanoida (Crustacea: Copepoda). Molecular Phylogenetics and Evolution, 59: 103–113.

Blanco-Bercial, L., Cornils, A., Copley, N., and Bucklin, A. 2014. DNA barcoding of marine copepods: assessment of analytical approaches to species identification. PLoS Currents, 6: 1–21.

Børsting, C., and Morling, N. 2015. Next generation sequencing and its applications in forensic genetics. Forensic Science International: Genetics, 18: 78–89.

Brandon-Mong, G. J., Gan, H. M., Sing, K. W., Lee, P. S., Lim, P. E., and Wilson, J. J. 2015. DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. Bulletin of Entomological Research, 105: 717–727.

Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Elbrecht, V., Steinke, D., Ratnasingham, S., de Waard, J. R. *et al.* 2019. Metabarcoding a diverse arthropod mock community. Molecular Ecology Resources, 19: 711–727.

Brown, E. A., Chain, F. J. J., Crease, T. J., MacIsaac, H. J., and Cristescu, M. E. 2015. Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? Ecology and Evolution, 5: 2234–2251.

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., and Lehtiniemi, M. 2016. Metabarcoding of marine zooplankton: prospects, progress and pitfalls. Journal of Plankton Research, 38: 393–400.

Bucklin, A., Steinke, D., and Blanco-Bercial, L. 2011. DNA barcoding of marine metazoa. Annual Review of Marine Science, 3: 471–508.

Bucklin, A., Yeh, H., Questel, J., Richardson, D., Reese, B., Copley, N., and Wiebe, P. 2019. Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. ICES Journal of Marine Science, 76: 1162–1176.

Carew, M. E., Pettigrove, V. J., Metzeling, L., and Hoffmann, A. A. 2013. Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species. Frontiers in Zoology, 10: 45.

Carroll, E. L., Gallego, R., Sewell, M. A., Zeldis, J., Ranjard, L., Ross, H. A., Tooman, L. K. *et al.* 2019. Multi-locus DNA metabarcoding of zooplankton communities and scat reveal trophic interactions of a generalist predator. Scientific Reports, 9: 281.

Casas, L., Pearman, J., and Irigoien, X. 2017. Metabarcoding reveals seasonal and temperature-dependent succession of zooplankton communities in the Red Sea. Frontiers in Marine Science, 4: 241. doi:10.3389/fmars.2017.00241

Chain, F., Brown, E., MacIsaac, H., and Cristescu, M. 2016. Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. Diversity and Distributions, 22: 493–504.

Chen, D. H., and Ronald, P. C. 1999. A rapid DNA mini preparation method suitable for AFLP and other PCR applications. Plant Molecular Biology Reporter, 17: 53–57.

Chiba, S., Batten, S., Martin, C., Ivory, S., Miloslavich, P., and Weatherdon, L. 2018. Zooplankton monitoring to contribute towards addressing global biodiversity conservation challenges. Journal of Plankton Research, 40: 509–518.

Cuscó, A., Catozzi, C., Viñes, J., Sanchez, A., and Francino, O. 2019. Microbiota profiling with long amplicons using Nanopore sequencing: full-length 16S rRNA gene and the 16S-ITS-23S of the RRN operon. F1000research, 7: 1755.

Darling, J. A., and Mahon, A. R. 2011. From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environmental Research, 111: 978–988.

Daro, M-H., Breton, E., Antajan, E., Gasparini, S., and Rousseau, V. 2006. Do Phaeocystis colony blooms affect zooplankton in the Belgian coastal zone? *In* Current Status of Eutrophication in the Belgian Coastal Zone, pp. 61–72. Ed. by V. Rousseau Presses universitaires de Bruxelles, Brussels.

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and Van Broeckhoven, C. 2018. NanoPack: visualizing and processing long read sequencing data. Bioinformatics, 34: 2666–2669.

Deagle, B. E., Division, A. A., Thomas, A. C., and Mcinnes, J. 2018. Counting with DNA in metabarcoding studies: how should we convert sequence reads to dietary data? Molecular Ecology, 28: 391–406. http://doi.org/10.1111/mec.14734.

Deschutter, Y., Everaert, G., De Schamphelaere, K. A. C., and De Troch, M. 2017. Relative contribution of multiple stressors on copepod density and diversity dynamics in the Belgian part of the North Sea. Marine Pollution Bulletin, 125: 350–359.

Deschutter, Y., Vergara, G., Mortelmans, J., Deneudt, K., De Schamphelaere, K. A. C., and De Troch, M. 2018. Distribution of the invasive calanoid copepod *Pseudodiaptomus marinus* Sato, 1913) in the Belgian part of the North Sea. Bioinvasions Records, 7: 33–41.

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., Muller-Karger, F. *et al.* 2018. Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. Limnology and Oceanography: Methods, 16: 209–221.

Ellien, C., Thiébaut, E., Dumas, F., Salomon, J.-C., and Nival, P. 2004. A modelling study of the respective role of hydrodynamic processes and larval mortality on larval dispersal and recruitment of benthic invertebrates: example of *Pectinaria koreni* (Annelida: Polychaeta) in the Bay of Seine (English Channel). Journal of Plankton Research, 26: 117–132.

Ferdous, Z., and Muktadir, A. 2009. A review: potentiality of zooplankton as bioindicator. American Journal of Applied Sciences, 6: 1815–1819.

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., Gielly, L. *et al.* 2015. Replication levels, false presences, and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources, 15: 543–556.

Flanders Marine Institute (VLIZ). 2019. Belgium LifeWatch Observatory Data: CTD, Temperature and Salinity Measurements in the Belgian Part of the North Sea. http://www.vliz.be/nl/imis?module=dataset&dasid=46856 (last accessed 15 October 2019).

Fransz, H. G. 1975. The spring development of calanoid copepod populations in the Dutch coastal waters as related to primary production. *In* 10th European Marine Biology Symposium, pp. 247–269. Ed. by G. Persoone and E. Jaspers. Universal Press, Wetteren.

Gaspar, J. M., and Thomas, W. K. 2013. Assessing the consequences of denoising marker-based metagenomic data. PLoS One, 8: e60458.

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., Stryke, D. *et al.* 2015. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Medicine, 7: 99.

Grosjean, P., Picheral, M., Warembourg, C., and Gorsky, G. 2004. Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. ICES Journal of Marine Science, 61: 518–525.

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., and Troedsson, C. 2014. Characterization of the 18S RRNA gene for designing universal eukaryote specific primers. PLoS One, 9: e87624.

Hargreaves, A. D., and Mulley, J. F. 2015. Assessing the utility of the Oxford Nanopore MinION for snake venom gland cDNA sequencing. PeerJ, 3: e1441.

Hebert, P. D. N., Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., Janzen, D. H. *et al.* 2018. A sequel to Sanger; amplicon sequencing that scales. BMC Genomics, 19: 219.

Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B: Biological Sciences, 270: 313–321.

Huang, D., Meier, R., Todd, P. A., and Chou, L. M. 2018. Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. Journal of Molecular Evolution, 66: 167–174.

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P. *et al.* 2019. Global trends in marine plankton diversity across kingdoms of life. Cell, 179: 1084–1097.e21.

Ivanova, N. V., Borisenko, A. V., and Hebert, P. D. 2009. Express bar-codes: racing from specimen to identification. Molecular Ecology Resources, 9: 35–41.

Johnson, S., Zaikova, E., Goerlitz, D., Bai, Y., and Tighe, S. 2017. Real-time DNA sequencing in the Antarctic Dry Valleys using the Oxford Nanopore sequencer. Journal of Biomolecular Techniques: JBT, 28: 2–7.

Kebschull, J. M., and Zador, A. M. 2015. Sources of PCR-induced distortions in high-throughput sequencing data sets. Nucleic Acids Research, 43: e143.

Kerckhof, F. 2002. Barnacles (Cirripedia, Balanomorpha) in Belgian waters, an overview of the species and recent evolutions, with emphasis on exotic species. Bulletin de l'Institut Royal des Sciences Naturelles de Belgique, 72(Suppl.): 93–104.

Kiørboe, T. 2001. Formation and fate of marine snow: small-scale processes with large-scale implications. Scientia Marine, 65: 57–71.

Klappenbach, J. A., Dunbar, J. M., and Schmidt, T. M. 2000. rRNA operon copy number reflects ecological strategies of bacteria. Applied and Environmental Microbiology, 66: 1328–1333.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research, 41: e1.

Kolde, R. 2019. Pheatmap: pretty heatmaps. R Package Version, 61: 1–7.

Kozol, R., Blanco-Bercial, L., and Bucklin, A. 2012. Multi-gene analysis reveals a lack of genetic divergence between *Calanus agulhensis* and *C. sinicus* (Copepoda; Calanoida). PLoS One, 7: e45710.

Krehenwinkel, H., Pomerantz, A., Henderson, J. B., Kennedy, S. R., Lim, J. Y., Swamy, V., Shoobridge, J. D. *et al.* 2019. Nanopore sequencing of long ribosomal DNA amplicons enables portable and simple biodiversity assessments with high phylogenetic resolution across broad taxonomic scale. Gigascience, 8: giz006. doi: 10.1093/gigascience/giz006

Laakmann, S., Gerdts, G., Erler, R., Knebelsberger, T., Arbizu, P. M., and Raupach, M. J. 2013. Comparison of molecular species identification for North Sea calanoid copepods (Crustacea) using proteome fingerprints and DNA sequences. Molecular Ecology Resources, 13: 862–876.

Lawson Handley, L. 2015. How will the "molecular revolution" contribute to biological recording? Biological Journal of the Linnean Society, 115: 750–766.

Lê, S., Josse, J., and Husson, F. 2008. FactoMineR: an R package for multivariate analysis. Journal of Statistical Software, 25: 1–18.

Lefebure, T., Douady, C. J., Gouy, M., and Gibert, J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Molecular Phylogenetics and Evolution, 40: 435–447.

Leow, L., Chew, L., Chong, V., and Dhillon, S. 2015. Automated identification of copepods using digital image processing and artificial neural network. BMC Bioinformatics, 16: S4.

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., and Atkinson, A. 2013. Next Generation Sequencing reveals the hidden diversity of zooplankton assemblages. PLoS One, 8: e81327.

Machida, R. J., and Knowlton, N. 2012. PCR primers for metazoan nuclear 18S and 28S ribosomal DNA sequences. PLoS One, 7: e46180.

Machida, R. J., and Tsuda, A. 2010. Dissimilarity of species and forms of planktonic neocalanus copepods using mitochondrial COI, 12S, nuclear ITS, and 28S gene sequences. PLoS One, 5: e10278.

Menegon, M., Cantaloni, C., Rodriguez- Prieto, A., Centomo, C., Abdelfattah, A., Rossato, M., Bernardi, M. *et al.* 2017. On site DNA barcoding by nanopore sequencing. PLoS One, 12: e0184741.

Mortelmans, J., Goossens, J., Amadei Martínez, L., Deneudt, K., Cattrijsse, A., and Hernandez, F. 2019. LifeWatch observatory date: zooplankton observations in the Belgian part of the North Sea. Geoscience Data Journal, 6: 76–84. doi:10.14284/329

O'Brien, T. D., Lorenzoni, L., Isensee, K., and Valdes, L. 2017. What are Marine Ecological Time Series telling us about the ocean? A status report. IOC Technical Series, 129: 1–297.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L. *et al.* 2015. Vegan: Community Ecology Package. R Package Version 2.2-0. http://CRAN.Rproject.org/package=vegan (last accessed 21 August 2020).

Oxford Nanopore Technologies. 2020a. Amplicons by Ligation (SQK-LSK109). https://community.nanoporetech.com/protocols/pcr-barcoding-amplicons/v/PBAC12_9067_v109_revL_14Aug2019 (last accessed 21 August 2020).

Oxford Nanopore Technologies. 2020b. Guppy on Windows. https://community.nanoporetech.com/downloads/guppy/release_notes (last accessed 21 August 2020).

Pearman, J. K., and Irigoien, X. 2015. Assessment of zooplankton community composition along a depth profile in the central Red Sea. PLoS One, 10: e0133487.

Pecceu, E., Hostens, K., and Maes, F. 2016. Governance analysis of MPAs in the Belgian part of the North Sea. Marine Policy, 71: 265–274.

Piper, A., Batovska, J., Cogan, N., Weiss, J., Cunningham, J., Rodoni, B., and Blacket, M. 2019. Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. Gigascience, 8: giz092. doi:10.1093/gigascience/giz092

Pochon, X., Bott, N. J., Smith, K. F., and Wood, S. A. 2013. Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. PLoS One, 8: e73935.

Pomerantz, A., Penafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., Barrio-Amoros, C. L., Salazar-Valenzuela D. *et al.* 2018. Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. GigaScience, 7: 1–14.

Prokopowich, C. D., Gregory, T. R., and Crease, T. J. 2003. The correlation between rDNA copy number and genome size in eukaryotes. Genome, 46: 48–50.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. *et al.* 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41: D590–D596.

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., Bore, J. A. *et al.* 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature, 530: 228–232.

Rabaut, M., Guilini, K., Van Hoey, G., Vincx, M., and Degraer, S. 2007. A bio-engineered soft-bottom environment: the impact of *Lanice conchilega* on the benthic species-specific densities and community structure. Estuarine, Coastal and Shelf Science, 75: 525–536.

Rang, F., Kloosterman, W., and de Ridder, J. 2018. From squiggle to basepair: computational approaches for improving Nanopore sequencing read accuracy. Genome Biology, 19: 90. doi: 10.1186/s13059-018-1462-9

Raupach, M., and Radulovici, A. 2015. Looking back on a decade of barcoding crustaceans. Zookeys, 539: 53–81.

Reddin, C., Nätscher, P., Kocsis, Á., Pörtner, H., and Kiessling, W. 2020. Marine clade sensitivities to climate change conform across timescales. Nature Climate Change, 10: 249–253.

Riccardi, N., Lucini, L., Benagli, C., Welker, M., Wicht, B., and Tonolla, M. 2012. Potential of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) for the identification of freshwater zooplankton: a pilot study with three *Eudiaptomus* (Copepoda: Diaptomidae) species. Journal of Plankton Research, 34: 484–492.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America, 74: 5463–5467.

Santos, A., van Aerle, R., Barrientos, L., and Martinez-Urtaza, J. 2020. Computational methods for 16S metabarcoding studies using Nanopore sequencing data. Computational and Structural Biotechnology Journal, 18: 296–305.

Semmouri, I., De Schamphelaere, K., Mees, J., Janssen, C., and Asselman, J. 2020. Evaluating the potential of direct RNA nanopore sequencing: metatranscriptomics highlights possible seasonal differences in a marine pelagic crustacean zooplankton community. Marine Environmental Research, 153: 104836.

Serrana, J., Miyake, Y., Gamboa, M., and Watanabe, K. 2019. Comparison of DNA metabarcoding and morphological identification for stream macroinvertebrate biodiversity assessment and monitoring. Ecological Indicators, 101: 963–972.

Shannon, C. E. 1948. A mathematical theory of communication. Bell System Technical Journal, 27: 379–423.

Simpson, E. H. 1949. Measurement of diversity. Nature, 163: 688–688.

Srivathsan, A., Baloğlu, B., Wang, W., Tan, W. X., Bertrand, D., Ng, A. H. Q., Boey, E. J. H. *et al.* 2018. A MinION-based pipeline for fast and cost-effective DNA barcoding. Molecular Ecology Resources, 18: 1035–1049.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., and Willerslev, E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology, 21: 2045–2050.

Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T. G., and Fontaneto, D. 2012. The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. Proceedings of the National Academy of Sciences of the United States of America, 109: 16208–16212.

Tara Expeditions. 2020. Variations in Plankton Biodiversity and Activity from the Equator to the Poles. https://oceans.taraexpeditions.org/en/m/environment/ocean-biodiversity/cp-taraoceanscell/ (last accessed 6 August 2020).

Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., and Trites, A. W. 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Molecular Ecology Resources, 16: 714–726.

Trivedi, S., Aloufi, A. A., Ansari, A. A., and Ghosh, S. K. 2016. Role of DNA barcoding in marine biodiversity assessment and conservation: an update. Saudi Journal of Biological Sciences, 23: 161–171.

Truelove, N., Andruszkiewicz, E., and Block, B. 2019. A rapid environmental DNA method for detecting white sharks in the open ocean. Methods in Ecology and Evolution, 10: 1128–1135.

van den Boogaart, K., and Tolosana-Delgado, R. 2008. "compositions": a unified R package to analyze compositional data. Computers & Geosciences, 34: 320–338.

Van Ginderdeuren, K., Van Hoey, G., Vincx, M., and Hostens, K. 2014. The mesozooplankton community of the Belgian shelf (North Sea). Journal of Sea Research, 85: 48–58.

Van Meel, L. I. J. 1975. La mer du Nord méridionale, le Pas de Calais et la Manche. Etude d'écologie marine, principalement en ce qui concerne le microplankton. Vol. II Etude planctonique. Bruxelles, Institut Royal des Sciences Naturelles de Belgique. 655 pp.

Vandepitte, L., Decock, W. and Mees, J. (Eds). 2010. Belgian Register of Marine Species, compiled and validated by the VLIZ Belgian Marine Species Consortium. *In* VLIZ Special Publication, p. 46. Vlaams Instituut voor de Zee (VLIZ), Oostende, Belgium. 78 pp.

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C. *et al.* 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome, 5: 27. doi:10.1186/s40168-017-0237

Wexels Riser, C., Wassmann, P., Olli, K., Pasternak, A., and Arashkevich, E. 2002. Seasonal variation in production, retention and export of zooplankton faecal pellets in the marginal ice zone and the central Barents Sea. Journal of Marine Systems, 38: 175–188.

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microbial Genomics, 3: e000132.

WoRMS Editorial Board. 2020. World Register of Marine Species. VLIZ. http://www.marinespecies.org (last accessed 1 April 2020). doi:10.14284/170.

Wu, S., Xiong, J., and Yu, Y. 2015. Taxonomic resolutions based on 18S rRNA genes: a case study of subclass copepoda. PLoS One, 10: e0131498.

Xu, J., Xu, Z., and Gao, Q. 2016. Distribution of dominant zooplankton species along a latitudinal gradient in China Sea during spring. Journal of Ocean University of China, 15: 502–508.

Yang, J., Zhang, X., Zhang, W., Sun, J., Xie, Y., Zhang, Y., Burton, G. A. Jr *et al.* 2017. Indigenous species barcode database improves the identification of zooplankton. PLoS One, 12: e0185697.

Young, R. G., Abbott, C., Therriault, T., and Adamowicz, S. J. 2017. Barcode-based species delimitation in the marine realm: a test using Hexanauplia (Multicrustacea: Thecostraca and Copepoda). Genome, 60: 169–182.

Zhan, A., Hulák, M., Sylvester, F., Huang, X., Adebayo, A. A., Abbott, C. L., Adamowicz, S. J. *et al.* 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution, 4: 558–565.

Handling editor: David Fields