



Evaluating the potential of direct RNA nanopore sequencing: Metatranscriptomics highlights possible seasonal differences in a marine pelagic crustacean zooplankton community

Ilias Semmouri^{a,*}, Karel A.C. De Schampelaere^a, Jan Mees^{b,c}, Colin R. Janssen^a, Jana Asselman^{a,d}

^a Ghent University, Laboratory of Environmental Toxicology and Aquatic Ecology, Faculty of Bioscience Engineering, 9000, Ghent, Belgium

^b Ghent University, Marine Biology Research Group, Faculty of Sciences, 9000, Ghent, Belgium

^c Flanders Marine Institute VLIZ, InnovOcean Site, Wandelaarkaai 7, 8400, Ostend, Belgium

^d Ghent University, Greenbridge, Wetenschapspark 1, 8400, Ostend, Belgium

ARTICLE INFO

Keywords:

Community transcriptomics
Metabarcoding
North sea
Copepods
Temora longicornis
Minlon

ABSTRACT

The implementation of cost-effective monitoring programs for zooplankton remains challenging due to the requirements of taxonomical expertise and the high costs of sampling and species identification. To reduce costs, molecular methods have been proposed as alternatives to morphology-based monitoring. Metatranscriptomics can contribute to promote both cost-effectiveness and accuracy of biological assessments of aquatic ecosystems. Here, we describe and evaluate the construction of a metatranscriptome dataset from a pelagic crustacean zooplankton community. We sampled zooplankton in one marine station, named LW02, in the North Sea, in both winter and summer, and generated transcripts using Oxford Nanopore Technology (ONT), a third-generation nanopore-based sequencing technology. ONT is, uniquely, capable of sequencing RNA directly, rather than depending on reverse transcription and PCR, and applicable to be used directly in the field. We found that metatranscriptomics is capable of species detection, including screening for the presence of endoparasites, hence competing with morphological identification. Taxonomic analysis based on ribosomal 18S transcripts identified calanoid copepods, particularly *Temora longicornis* and *Acartia clausi*, as the most abundant community members. Moreover, up to 40.4% and 50.5% of all sequences could be assigned to predicted genes in the winter and summer sample, respectively. The most abundant mRNA transcripts with known function coded for essential metabolic processes. GO term annotation revealed that genes involved in glycolytic and translation-related processes were most expressed in the community. Although small in scale, our study provides the basis for future efforts to characterize the metatranscriptome of marine zooplankton communities and its application in biomonitoring programs.

1. Introduction

Zooplankton plays an important role in mediating the transfer of energy in food webs from lower to higher trophic levels in marine ecosystems (i.e. other zooplankton, several pelagic fish species, several marine birds and mammals), but also to the benthic zone through sedimentation of faecal pellets (Hovinen et al., 2014; Nielsen et al., 1993; Richardson, 2008; Van Ginderdeuren et al., 2014). Zooplankton also contributes substantially to ecosystem services, such as biogeochemical cycling (Roemmich and McGowan, 1995), and influences the

extent of climate change through carbon fixation via the biological pump (Richardson, 2008). However, they are very susceptible to a changing climate and to other changing environmental conditions as well, such as eutrophication and pollution (e.g. Abdulwahab and Rabee, 2015; Drira et al., 2017; Richardson, 2008). Furthermore, due to their rapid responses to environmental variation (i.e. changes in zooplankton biomass, community structure and trophic linkages in marine food webs), planktonic organisms have been used as bio-indicators of ecosystem changes (Chiba et al., 2018; Ferdous and Mukhtadir, 2009).

With the growing need to better understand the impacts of

* Corresponding author.

E-mail address: ilias.semmouri@ugent.be (I. Semmouri).

<https://doi.org/10.1016/j.marenvres.2019.104836>

Received 15 April 2019; Received in revised form 29 October 2019; Accepted 3 November 2019

Available online 5 November 2019

0141-1136/© 2019 Elsevier Ltd. All rights reserved.

environmental change on zooplankton communities, zooplankton monitoring programs (e.g. Continuous Plankton Recorder) have been carried out in the marine environment globally since the early 20th century (e.g. Batten et al., 2003; O'Brien et al., 2017). In the Belgian Part of the North Sea (BPNS), most recent mesoscale zooplankton surveys were carried out by Van Ginderdeuren et al. (2014) between 2009 and 2010 and by Deschutter et al. (2017) between 2016 and 2017, investigating differences in zooplankton dynamics. With 70 species (51% of all species recorded), Crustacea is the most found taxon in the BPNS (Van Ginderdeuren et al., 2014). Most zooplankton monitoring studies focus mainly on variability in biodiversity and biomass (Chiba et al., 2018). However, monitoring of these organisms, especially crustaceans, is hindered by challenges in their identification, which is time-consuming, complicated and requires biological expertise. Advances in practical, cost-effective molecular approaches, such as (meta)barcoding, can help overcome the issues with morphology-based biomonitoring (Trivedi et al., 2016; Yang et al., 2017). Metabarcoding extends DNA-based species identification to the community level using bulk samples (total DNA) or potentially degraded environmental DNA (eDNA) for which species identification is not practical (Bohmann et al., 2014; Ji et al., 2013). However, the potential for taxonomic biases to be introduced by PCR primers is a limitation of metabarcoding (Zhou et al., 2013). However, Krehenwinkel et al. (2017) recently showed that amplification bias can be considerably reduced, by degenerate primers or by targeting amplicons with conserved priming sites, and that PCR-based barcoding might yield quantitative results.

A more comprehensive molecular data set would be able to identify and assess the impact of the main drivers of changes in the marine ecosystem, rather than only determining species richness. Studying the functional activities of a community - *in situ* and without *a priori* knowledge of genes - has been facilitated by metatranscriptomics, i.e. the study of community gene transcription. In community transcriptomics, the randomly sequenced mRNAs from an entire community at a single time point permit the identification of their active cellular and biological processes in its given ecological context (Helbling et al., 2012; Lavelle and Sokol, 2018). Therefore, metatranscriptomics offers a more informative perspective compared to metagenomics, as it reveals the transcriptionally active processes in addition to the genetic content (Bashiardes et al., 2016). RNA extractions – at least those without an mRNA purification step – also often contain >90% ribosomal RNAs of the total reads, including the taxonomically informative barcode regions, allowing the user to address the species composition of a community at the same time from the same sample (McGrath et al., 2008). Hence, metatranscriptomics is generally employed to assess both the functional and the taxonomic components of a community, hereby avoiding the taxonomic biases introduced by PCR primers (Wilson et al., 2018). Moreover, studies have demonstrated that eggs, dead carcasses and feces and several sources of (e)DNA are picked up by metabarcoding, potentially resulting in a biased view on which members are actively present in a community (Barnes and Turner, 2016; Merkes et al., 2014). Metatranscriptomics is to date already widely used to study microbial community functions, including marine bacterioplankton communities (e.g. Helbling et al., 2012; Hewson et al., 2014; Kopf et al., 2015; Miller et al., 2017; Wermheuer et al., 2014) and phytoplankton communities (e.g. Alexander et al., 2015). However, no community transcriptomic studies have been reported so far for zooplankton communities specifically.

Current RNA sequencing-based transcriptomic analyses (RNA-seq) were until recently only based on the high-throughput sequencing of complementary DNA (cDNA) (Galarde et al., 2018). These RNA-seq methods detect the products of a synthesis reaction rather than directly detecting the RNA molecule. Major disadvantages of this approach are the occurrence of reverse transcription biases and PCR biases (such as a distortion of relative cDNA numbers), the loss of some RNA molecules and/or their modification during PCR amplification (Galarde et al., 2018; Sanders et al., 2014). In contrast, using nanopore

sequencing technology, RNA can be sequenced directly, without introducing reverse transcription or PCR amplification-related biases (Galarde et al., 2018). This improves the ease and speed of RNA analysis, while yielding richer biological information. Here, we sequenced the metatranscriptome of a pelagic crustacean community in the BPNS, using the MinION™, a nanopore-based DNA/RNA sequencing platform (Oxford Nanopore Technologies). Unlike other sequencing technologies, the MinION is a portable platform and therefore not restricted to laboratories and usable in the field (Quick et al., 2016). Our objective was to establish a method for capturing the metatranscriptome of the zooplankton community in marine samples. We aimed to evaluate the benefits and disadvantages of nanopore RNA sequencing of complex environmental samples. We wanted to address the composition of this community by analysis of transcripts from a commonly used DNA barcoding gene, 18S rRNA, a component of the small ribosomal subunit (SSU), and to study the community's activity, too. From such functional data, we aimed to identify the most active metabolic pathways in the metatranscriptome and we tried to place these results into a broader context of physiological activities within the marine realm. Our study also provides the first insights into changes in structure and activity of these communities as a response to seasonal differences in environmental conditions, hence making metatranscriptomics an interesting tool for future monitoring purposes and environmental studies in general (Wermheuer et al., 2014).

2. Material & methods

2.1. Sampling and RNA extraction

Zooplankton were collected with the research vessel (RV) Simon Stevin during two sampling campaigns at the off-shore station LW02 (Lat.: 51° 48' N, Long. 2° 33' E; whole column about 35 m deep) in the North Sea, NO Atlantic Ocean, approximately at 65 km of Sea Bruges. During the first sampling campaign, held on 21 February 2018 (mid-winter period; below referred to as the “winter sample”), a water temperature of 7.6 °C, a salinity of 34.7 PSU and a conductivity of 3.59 S/m were being recorded. The second sample (below referred to as the “summer sample”), was taken on 26 September 2018, with a water temperature of 17.5 °C, a salinity of 34.9 PSU and a conductivity of 4.56 S/m being measured. Zooplankton were collected using a WP2 net (200 µm mesh size) and specimens were immediately stored in Invitrogen™ RNeasy™ Stabilization Solution until they could be transported to the laboratory. Here, holo- and meroplanktonic crustaceans were picked out immediately with a Pasteur pipette and used for RNA extractions. Crustacean taxa as mysids, amphipods and cumaceans were left out of the sample for RNA extractions, as they are being referred to as hyperbenthic or tychoplanktonic (Van Ginderdeuren et al., 2014). Zooplankton identification was carried out with the aid of a stereoscopic microscope (Kyowa Optical SD-2PL), but densities were not counted.

Total RNA was extracted with the RNeasy kit and Qiashredder (Qiagen, Venlo, Netherlands) following the manufacturer's protocol. DNA contamination was removed by a DNase treatment (Qiagen, Venlo, Netherlands). RNA Quality was assessed with the NanoDrop™ spectrophotometer (Version, 2000c; Thermo Scientific™, Thermo Fisher Scientific) with the quality standards being based on the 260/230 and 260/280 absorbance ratios. The amount of RNA in the sample was determined by means of a Qubit™ fluorometer (Version 2.0, Invitrogen™, Thermo Fisher Scientific). The sample was stored at –80 °C until library preparation.

2.2. Nanopore library preparation and direct RNA sequencing by ONT

The MinION is a pocket-sized, USB-powered nanopore sequencing platform, weighing less than 100 g, yet capable of directly sequencing full-length DNA or RNA sequences (Russell et al., 2018; Urban et al., 2015). Direct RNA sequencing was performed using the Direct RNA

Sequencing protocol with the SQK-RNA001 kit (ONT, USA, 2018a), with 700 ng of RNA (to ensure sufficient material remained for sequencing) used as input for library preparation. The RNA was ligated to 1.0 µL reverse transcription (RT) adapter using T4 DNA ligase (NEB 0202). A reverse transcription reaction was performed on the RT ligated RNA, using SuperScript™ III Reverse Transcriptase (Invitrogen™), and left to incubate in a thermal cycler at 50 °C for 50 min and 70 °C for 10 min. The products were then purified using 72 µL of CleanNGS beads (CleanNA), followed by a washing step in 70% ethanol. Sequencing adapters preloaded with motor proteins were then ligated onto the cDNA and the final library was cleaned up once more using 40 µL of the CleanNGS beads. Next, the library was loaded onto a FLO-MIN106 flow cell and the RNA library was sequenced on a MinION device for 48 h. For base calling, the locally-installed software Guppy v3.2.4 (ONT, USA, 2019) was used with standard settings. Two and three technical replicates were sequenced for the winter and summer sample, respectively, to provide information about process variability and reproducibility.

2.3. Bioinformatics and data analysis

First, the raw reads, obtained from direct RNA sequencing, were trimmed using Poreplex V0.2 (<https://github.com/hyeshik/poreplex>) to remove any 3' adapter sequences. Using the “-headcrop 15” command, Nanofilt V2.2 (De Coster et al., 2018) trimmed 30 nucleotides from the start of each read. Additionally, Nanofilt V2.2 filtered these reads as well, using a quality score cutoff of 8 (cf. Jenjaroenpun et al., 2018) and a minimum length of 200 bp to obtain high quality reads. Quality scores and read lengths of both samples were visualised and comprehensively summarised using NanoPlot V1.15 and NanoStat V1.1.2 respectively (De Coster et al., 2018). All runs were additionally compared and visualised using the NanoComp (V1.1.0) package (De Coster et al., 2018). The high quality RNA reads are deposited in the NCBI Short Read Archive (SRA) database under accession number BioProject: PRJNA577375.

For functional annotation of each generated read, all sequences were first searched against the non-redundant protein database (nr) using BLASTx with a cutoff e-value < 1E-05 and a similarity cut-off of 75%. Gene ontology (GO) terms were then assigned by BLAST2GO v5.2.0, using default filtering parameters (Annotation cutoff 55, GO Weight 5) (Conesa et al., 2005; Gotz et al., 2008). GO terms were subsequently assigned to metabolic pathways according to KEGG (Kyoto Encyclopedia of Genes and Genomes) mapping (Kanehisa et al., 2012).

Taxonomic composition of the samples was determined using Metaxa2 (v2.2), which allows for the extraction of SSU rRNA (18S) (Bengtsson-Palme et al., 2015). The extracted 18S reads were blasted against the NCBI nr database (BLASTn, accessed on 13.08.2018) at the e-value cutoff of 1E-05 and a similarity cut-off of 90% to obtain a taxonomic annotation at the genus/species level. Both zooplankton samples were characterized using alpha-diversity indices (i.e., Species richness, Shannon diversity index and Simpson diversity index) at the species level (Shannon, 1948; Simpson, 1949). Seasonal difference in diversity was examined statistically by either an unpaired two-sample *t*-test or a Welch *t*-statistic. All data were checked for normality and homogeneity of variance using Shapiro-Wilk normality test and Levene's test, respectively, with a *p* = 0.05.

Next, we employed Minimap2 software version 2.12 (Li, 2018) to align the high-quality reads of the metatranscriptome on (1) the *Temora longicornis* Müller O.F. transcriptome, published by Semmouri et al. (2019), (2) a draft *Semibalanus balanoides* L. genome (Nunez et al., 2018) and (3) an Expressed Sequence Tags (EST) library of *Calanus helgolandicus* Claus (Carotenuto et al., 2014), as these were the only three species in our samples for which transcriptome/genome/EST data is currently available in GenBank (May 2019). The SAM alignment files generated via Minimap2 mapping were imported into SAMtools version 1.6 (Li et al., 2009) for summarizing purposes of the reads mapping to the transcriptomes.

Next, we evaluated which remaining unassigned transcripts – i.e.

those that could not be annotated by Blast2GO, were not recognised by Metaxa2, or that did not map to the *Temora* transcriptome – had evidence of being functional non-coding RNA (ncRNA). We searched for four types of ncRNA, namely: microRNAs (miRNAs), transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and small nuclear RNAs (snRNAs). Using the default significance threshold e-value of 1E-03, the cmscan option of INFERNAL v1.1.2 (Nawrocki et al., 2015; Nawrocki and Eddy, 2013) was implicated in searching for putative sequences against the Rfam database, a collection of multiple sequence alignments and secondary structure profiles representing non-coding RNA families (Griffiths-Jones et al., 2003). A schematic flow-chart of the performed analyses is presented in Fig. 1.

3. Results

Successful base calling could be performed for more reads of the summer samples when compared to the winter sample (Table 1), with the number of bases being higher as well, indicating more and longer reads in the summer samples than the winter samples (Table 1; Fig. 2). Reads with a Q-score lower than 8 or of a length smaller than 200 bp were filtered out, for the winter sample resulting in 274,744 (46.7%) and 172,423 (50.7%) remaining RNA sequences for use in downstream analyses (Table 1; Fig. 2). The N50, defined as the length of the longest contig such that all contigs of at least that length compose at least 50% of the bases of the assembly, was 758 bp and 894 bp long for the winter samples, respectively, while the average transcript length was 678 bp (Figs. 3 and S1, Table 1). The summer samples consisted of 266,210 (69%), 535,139 (66.9%) and 499,560 (68.8%) remaining trimmed RNA sequences with an average N50 and an average contig length of 928 and 775 bp, respectively (Fig. 3, Tables 1 and S1). Summary statistics were obtained from NanoStat. 132,532 of the 274,744 (48.2%) and 73,488 of the 172,423 (42.6%) reads of the winter sample could be mapped to the *Temora* transcriptome, while on average 4.25% of these reads mapped to the *Semibalanus* genome (Table 1). In contrast less than 20% and 2% of the reads of the summer replicates could be mapped to the *Temora* transcriptome and the *Semibalanus* genome, respectively (Table 1). 4,517 (1.6%) and 4,855 (2.8%) transcripts of the winter sample replicates and 12,395 (4.7%), 45,045 (8.4%), 41,684 (8.3%) of the summer sample replicates mapped to the 770 ESTs of the *Calanus helgolandicus* library.

Additionally, a total of 111,102 (40.4%) and 59,418 (34.4%) transcripts of the samples ‘winter 1’ and ‘winter 2’ respectively, had similarities to known proteins in the NCBI nr database, with a cutoff of *e* < 1E-05 (Tables 1 and S2). 68,852 (42%) of the 163,642 and 44,065 (38.9%) of the 113,005 transcripts of the winter samples which did not give a hit after blasting, could be mapped though to the *T. longicornis* transcriptome (Fig. 6). BLAST2GO assigned a functional annotation to 46,111 and 27,361 transcripts of the winter sample (16.8% and 15.9% of the total assemblies, respectively; Tables 1 and S2). Annotation was clustered into three general sections: biological processes, cellular components and molecular functions (Fig. 4). Translational processes and structural constituent of the ribosome were the most annotated GO terms in the biological processes and molecular functions sections, respectively. For the summer sample, annotation was more successful, as between 24.2% and 36.8% of the transcripts could be annotated (Table 1). 9,551 out of 131,578 (7.3%), 15,737 out of 265,909 (5.9%) and 14,568 out of 247,457 (5.9%) transcripts of all three summer samples, could be mapped to the *T. longicornis* transcriptome, even though they could not be blasted or annotated successfully (Fig. 6). Using Blast2GO, all transcripts were mapped against the KEGG database, resulting in average 3,888 and 9,220 transcripts grouped into 58/58 and 57/67/83 KEGG pathways for the winter and summer samples respectively (Table 1, Table S1). Glycolysis/Gluconeogenesis, biosynthesis of antibiotics, purine metabolism, thiamine metabolism and oxidative phosphorylation pathways were the most represented pathways (Table S1).

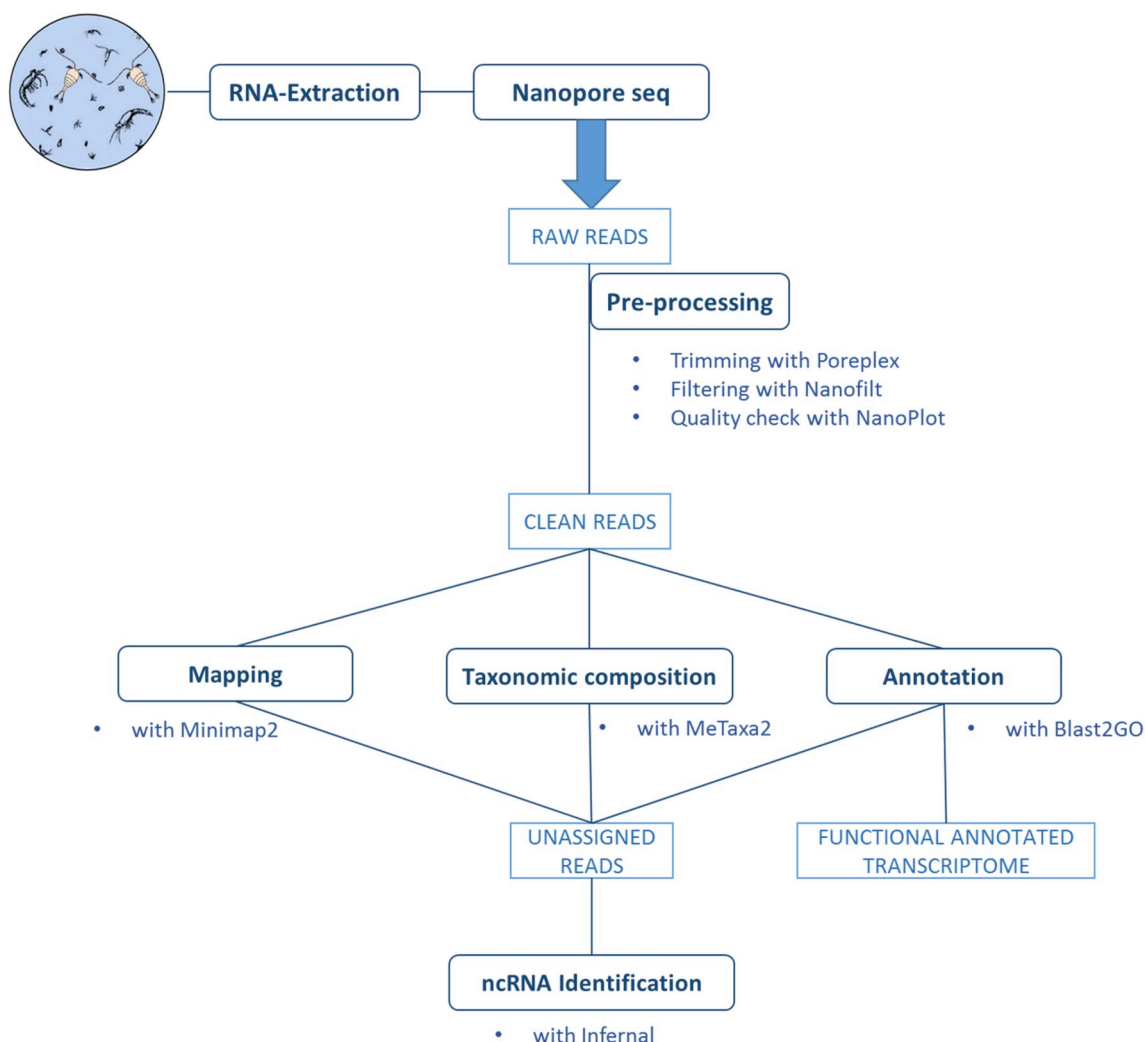


Fig. 1. Flow chart of the metatranscriptomic analyses, presenting the workflow and analysis pipeline described in the text. The details of each step are elaborated in section 2.3 Bioinformatics and data analysis.

Table 1

Summary statistics for the de novo zooplankton community transcriptome for each technical replicate. %GC is the percent of nucleotide bases in sequences that are either G or C. N50 and N90 bp are the mean number of basepairs in all transcripts that, ordered by length, make up respectively 50% and 90% of the assembly. Percentages are reported within brackets.

	Winter Sample		Summer Sample		
	Replicate 1	Replicate 2	Replicate 1	Replicate 2	Replicate 3
Total number of base called reads	587,747	339,804	385,274	799,461	725,369
Total number of trimmed and quality reads	274,744	172,423	266,210	535,139	499,560
Total number of bases (bp)	173,029,846	118,093,589	195,432,478	425,422,810	398,643,745
Minimum read length (bp)	200	200	200	200	200
Average read contig length (bp)	629.79	727.07	734.40	794.98	797.28
Median contig length (bp)	535	622	674	713	714
Maximum contig length (bp)	12,949	6,173	6,628	7,301	7,877
N50 assembled transcripts length (bp)	758	894	878	952	953
N90 assembled transcripts length (bp)	347	401	430	456	456
Total GC count (bp)	76,711,814	53,258,610	87,007,167	184,962,204	173,547,163
GC count for the metatranscriptome (%)	44.3	45.1	44.5	43.5	43.5
Number of reads mapped to <i>Temora</i> transcriptome (%)	132,532 (48.2)	73,488 (42.6)	38,260 (14.4)	100,946 (18.8)	95,413 (19.1)
Number of reads mapped to draft <i>Semibalanus</i> genome (%)	9,936 (3.6)	8,507 (4.9)	3,690 (1.3)	9,803 (1.8)	9,235 (1.8)
Number of reads mapped to 770 <i>Calanus</i> ESTs	4,517 (1.6)	4,855 (2.8)	12,395 (4.7)	45,045 (8.4)	41,684 (8.3)
Transcripts with BLAST hits against nr database	111,102 (40.4)	59,418 (34.4)	134,632 (50.5)	269,230 (50.3)	252,103 (50.4)
Transcripts with Blast2GO annotation	46,111 (16.8)	27,361 (15.9)	98,006 (36.8)	129,873 (24.2)	120,105 (24.4)
Transcripts with KEGG annotation	4,636	3,141	11,345	8,714	7,600
Number of unassigned reads (%)	96,182 (35.0)	43,911 (25.5)	96,472 (36.2)	249,771 (46.6)	231,791 (46.4)

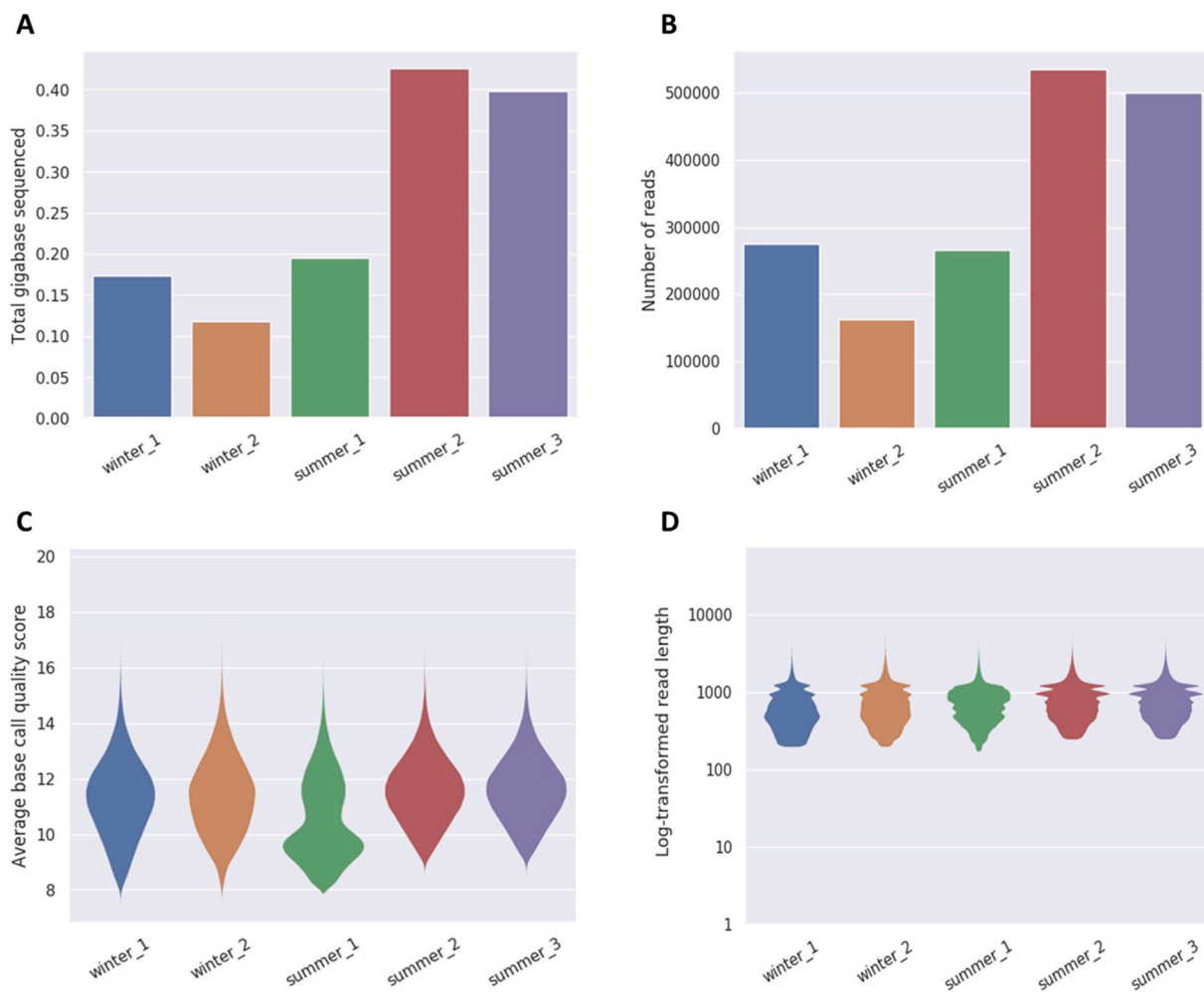


Fig. 2. Sequencing performance metrics for both runs: A) bar graph of total throughput in gigabases; B) bar graph of total number of reads; C) violin plot of base call quality scores; D) violin plot of log-transformed read lengths. Plots were generated by NanoComp (De Coster et al., 2018).

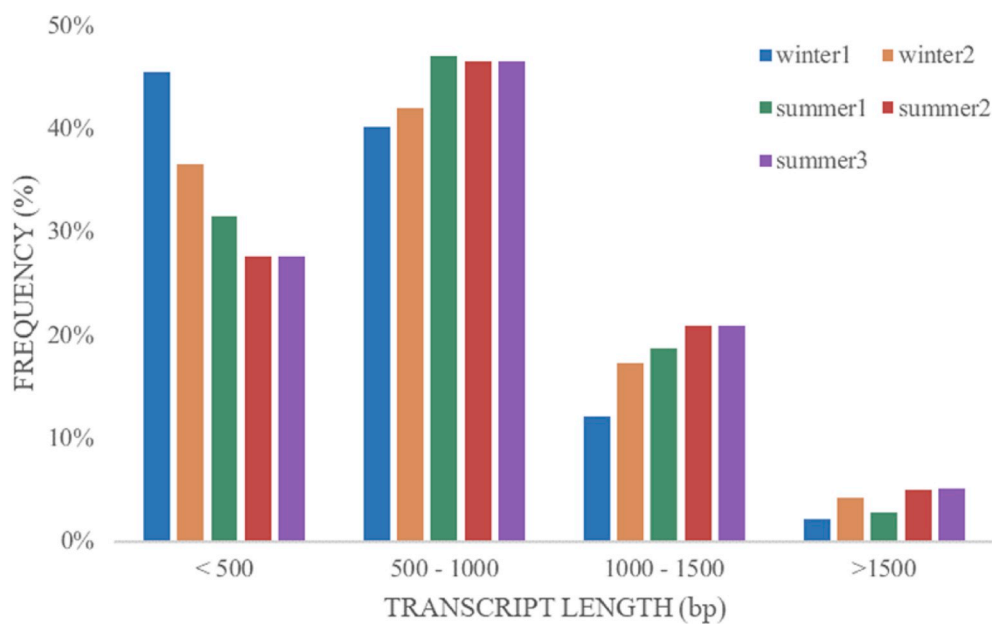


Fig. 3. Histogram of filtered read lengths.

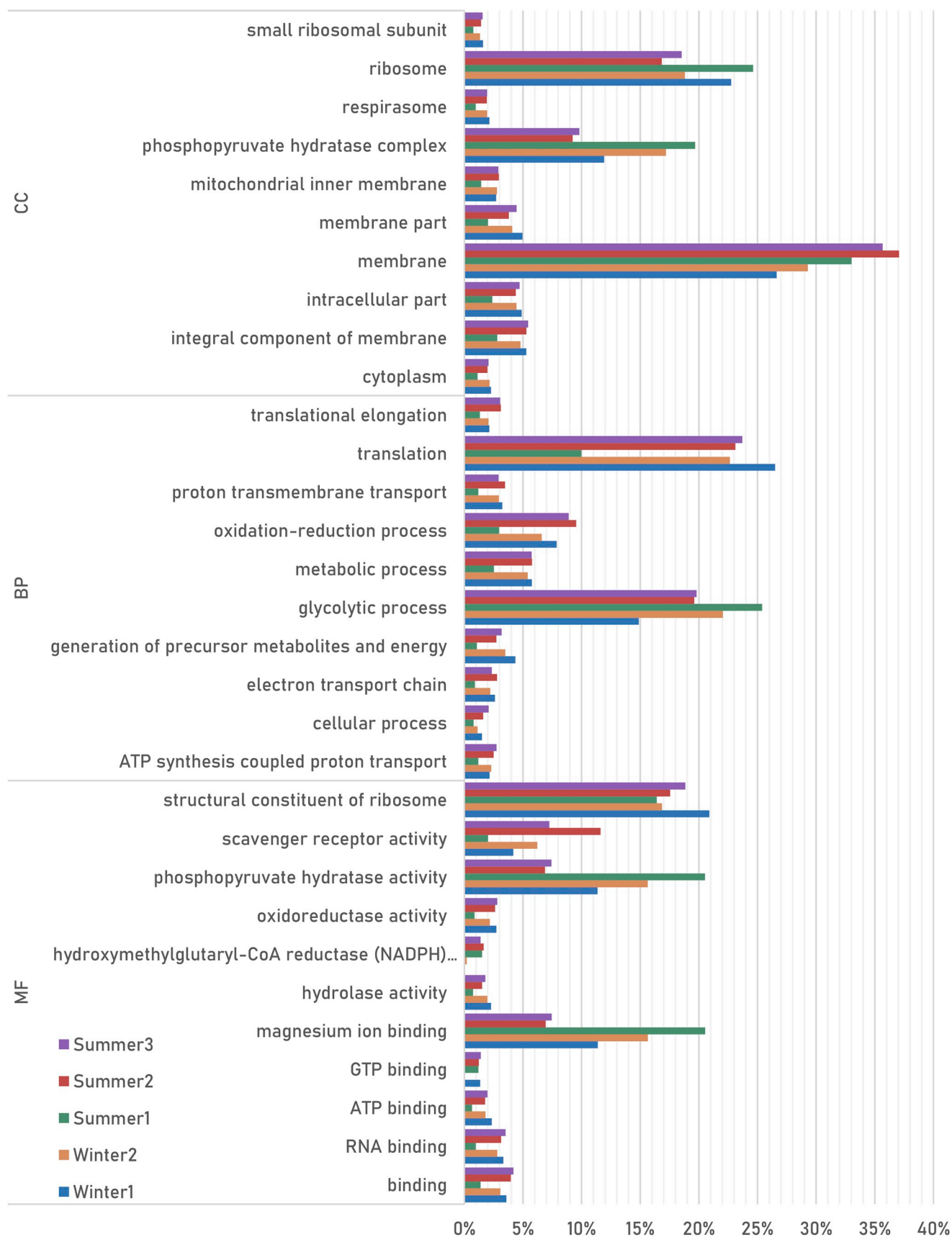


Fig. 4. Comparison between the distributions of the top 10 GO annotations (gene ontology level 2) obtained from the winter and summer zooplankton samples for biological process (BP), molecular function (MF) and cellular component (CC). Percentages were calculated as the number of sequences annotated to a given GO term divided by the total number of GO annotated transcripts of the respective class it belongs to (x100).

Taxonomic analysis based on expressed 18S rRNA revealed that the core of the community consists of calanoid copepods. The majority of the 18S reads could be assigned to calanoid copepods in both analyses, i. e. *Temora longicornis*, *Acartia clausi* and *Paracalanus parvus* (Fig. 5, Table S2). In the summer sample, up to 90% of the reads can be assigned to copepods (Fig. 5). However, in the SSU analysis of the winter sample (Fig. 5, Table S2), about 40% of the reads was detected for the infraclass Cirripedia (barnacles). Up to 5% of all the expressed 18S rRNA reads of the winter sample could be assigned to protists, namely *Paradinium*. (Fig. 5, Table S2). Additional information regarding diversity and richness of the samples is provided in Table 2.

Finally, we evaluated which non-protein coding transcripts showed evidence of being functional non-coding RNA transcripts. Using the Infernal tools, we identified 459 and 10 transcripts of the winter sample replicates and 455, 401 and 1,400 transcripts for the summer sample replicates respectively, that were matched to an Rfam profile (Supplementary file2). About half of the ncRNAs in both samples belonged to rRNA families, such as 5.8S_rRNA. Hence, in total, 20,711 and 9,350 transcripts of the winter sample replicates and 26,355, 46,772 and 44,696 reads from the summer sample, which account for 7.5, 5.4, 9.9, 8.5 and 8.9% of the datasets respectively, represent rRNA (Fig. 6). We also found representatives of miRNA families (e.g. mir-186, mir-598, ...) and small nucleolar RNA (snoR) families in both samples (Supplementary file2).

4. Discussion

Expression monitoring, in which the transcript levels of genes are measured in different environmental and/or physiological conditions to search for regulatory expression patterns (Celis et al., 2000), is potentially useful for fast and cost-effective biomonitoring purposes. Here, we applied direct RNA sequencing to get a first insight into the genomic seasonal response of the pelagic crustacean zooplankton community from the North Sea to changing environmental conditions. While it is now no longer obligatory to synthesise a cDNA strand opposite the RNA template during library preparation (as is the case in second generation sequencing), it is still advised in the protocol to synthesise a complementary DNA strand (ONT, USA, 2018a). Garalde et al. (2018) found that the duplex improves throughput, possibly by reducing intra-molecular secondary structure of the RNA. However, since the MinION does not sequence the cDNA strands (only the RNA strand of a duplex), incomplete cDNA synthesis does not prevent sequencing of RNA strands any longer (Garalde et al., 2018). In this study, we have demonstrated the feasibility of direct RNA - nanopore - sequencing of a crustacean

community transcriptome. Despite the mentioned advantages of using ONT nanopore sequencing, a high read level error rate and high input material requirements are to date the primary limitations of this technology (cf. Tyler et al., 2018). For both samples, about half of the base-called reads were of a sufficient quality to be used in subsequent analyses (Table 1). Hence, reducing the error rate would enable a higher output. While we used the same protocol between the runs, we were unable to maintain technical consistency, with a smaller RNA sequence yield generated in some of the runs (Table 1; Fig. 2). Possible factors which may have an impact on output data is the quality of the input library and the quality of the flowcell (available number of active pores). Hence, the high level of inter-run variability, already described in literature (e.g. by Tyler et al., 2018), is a major drawback of this technology. It is not yet possible to predict run output simply based upon the quantity of RNA input into the library preparation. Finally, the ONT library preparation protocol is still prone to bias, as it includes the ligation of adapters via a T4 ligase, which might introduce artificial chimera. Because this bias is random, it will not affect the results substantially (Viehweger et al., 2018).

First, we focused on mapping the metatranscriptome data to representative data. Using minimap2, on average 45.4% of the reads from the winter sample replicates mapped to the transcriptome of *T. longicornis*, while less than 20% of the summer sample replicates mapped to this transcriptome. On average, only 4.25% of the winter sample replicates could be mapped to the draft genome of *S. balanoides*, compared to less than 2% for the summer sample replicates. This might be explained by the absence of barnacle nauplii in summer sample. Species specific reads might not map onto the transcriptome of another taxon – even not of congeners – with the same efficiency, probably explaining the lower mapping rates of the diverse summer sample (Fig. 5, Table 2), compared to the winter sample. In contrast, the BLAST results were obtained for 34–40% (winter) and 50% (summer) of all transcripts. The remaining unannotated transcripts may be (1) uncharacterized genes or (2) species specific reads, which are not known up to now. Most sequences we identified are indeed derived from crustaceans whose particular genome has not been characterized. It is important to note that for many crustacean genes, sequences are very similar among distantly related organisms, such as insects. The GenBank database appears to be heavily weighted in favor of model organisms that are frequently studied for molecular genetics or of organisms whose genomes have been sequenced (Semmouri et al., 2019).

Reads in the metatranscriptome reveal a diverse assemblage of cellular activities. As would be expected, transcripts of both samples were observed to be related to a number of general cellular processes

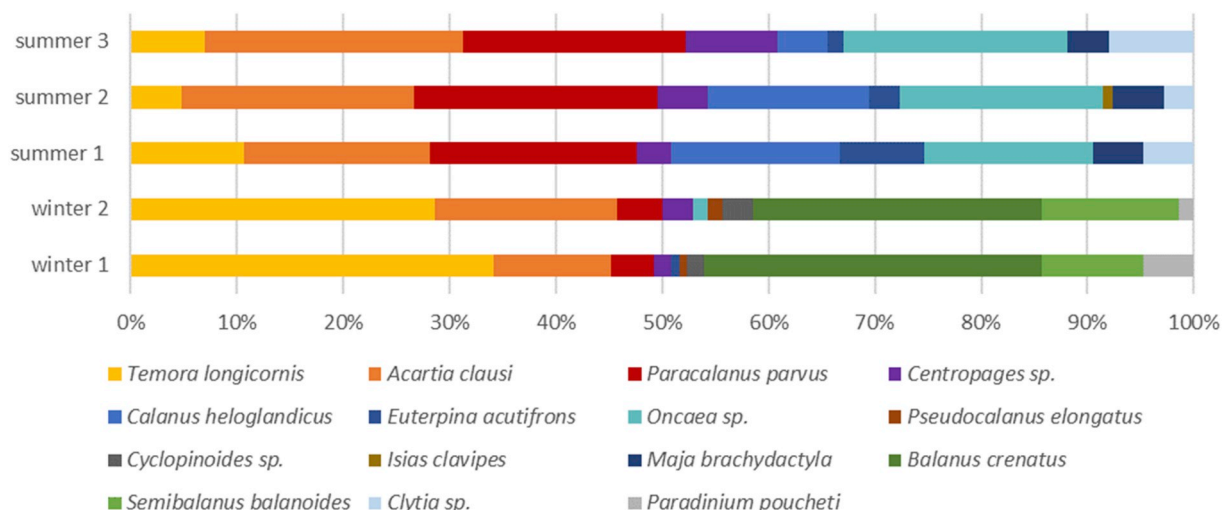


Fig. 5. Summary of taxonomic representation across the metatranscriptome based on the SSU (18S) reads extracted from the Metaxa2 output.

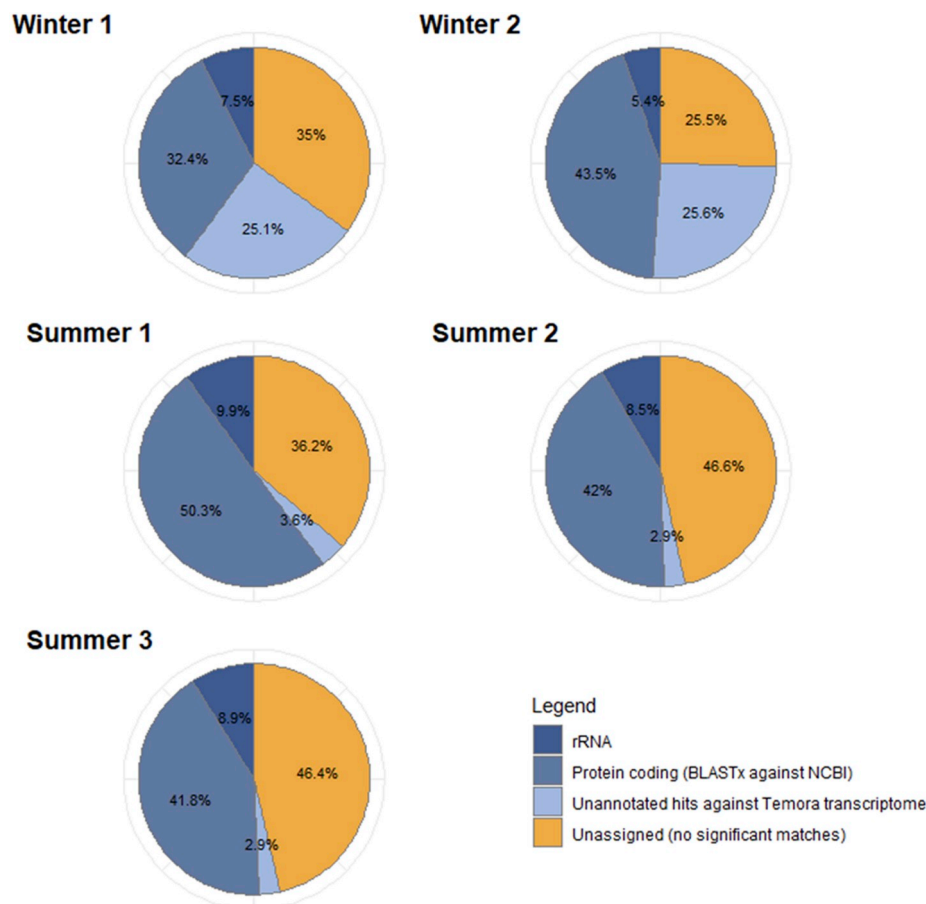


Fig. 6. Inventory of RNAs from each sample in the zooplankton metatranscriptomic datasets. The ‘unassigned’ slices represent reads that have no significant matches with (1) rRNA, (2) known protein-coding genes, (3) ncRNA or (4) gave a hit when mapping against the *Temora* transcriptome. Numbers represent the percentage of each category.

Table 2

Biodiversity and richness estimators of the crustacean community. Species richness, Shannon’s index of biodiversity and Simpson’s index of biodiversity are reported for the barcode reads. Formulas were taken from Shannon (1948) and Simpson (1949). P_i is the proportion of individuals belonging to species i . P-values are given (unpaired two-sample t -test/Welch t -statistic) (* is significant).

METRIC	FORMULA	SAMPLE					P-VALUE
		WINTER SAMPLE		SUMMER SAMPLE			
		REPLICATE 1	REPLICATE 2	REPLICATE 1	REPLICATE 2	REPLICATE 3	
Species richness (<i>S</i>)	# species	10	10	9	10	9	0.1835
Shannon's diversity (<i>H</i>)	$-\sum P_i \ln(P_i)$	1.681	1.798	2.056	1.955	1.955	0.02716*
Simpson's diversity (<i>D</i>)	$1 - \sum p_i^2$	0.757	0.795	0.861	0.832	0.836	0.0356*

such as protein, amino acid, and nucleic acid synthesis and energy metabolism, indicating active cells (Fig. 4). The most abundant mRNA transcripts with known function coded for essential metabolic processes (Fig. 4). Particularly, the summer data set yielded an outspoken number of glycolytic enzymes (Fig. 4). While these enzymes were also abundantly expressed in the winter sample (Fig. 4), our study also revealed expression of transcripts related to (1) protein synthesis processes and (2) aerobic respiration processes (Fig. 4). Classified among those most abundant GO terms, were genes encoding for enolase (phosphopyruvate hydratase; EC 4.2.1.11) as well as aldolase (fructose-bisphosphate aldolase; EC 4.1.2.13). Enolase is a highly conserved metalloenzyme, in both pro- and eukaryotes, with an absolute requirement for magnesium as the natural cofactor (cf. GO term magnesium ion binding, GO:0000287) (Kikuchi et al., 2017). Enolase has at least two important functions in arthropods: (1) gluconeogenesis, the generation of glucose

molecules and (2) glycolysis, the first fundamental metabolic process of aerobic respiration, which is responsible for the majority of energy (ATP) production in all organisms (Kikuchi et al., 2017). In addition, arthropod (especially insect) enolases can have species-specific roles as well (Kikuchi et al., 2017). Fructose-bisphosphate aldolase is an enzyme encoded by just one gene and is involved in functional gluconeogenic and glycolytic pathways as well (Guo et al., 2012). This result was further confirmed by KEGG analysis, as the “Glycolysis/Gluconeogenesis” pathway was one of the most abundant pathways. Additionally, “Oxidative phosphorylation” was found among the top represented KEGG pathways. A possible explanation for the abundantly expressed glycolytic enzymes in summer samples is the additional requirement for energy of the zooplankton because of an increase in temperature (7.6 °C vs 17.5 °C). Respiration rates of several marine crustaceans, including *T. longicornis* and several *Centropages* and *Calanus* species, were indeed

found to increase with temperature (Castellani and Altunbas, 2013; Gaudy, 1975; Vidal, 1980). Moreover, Semmouri et al. (2019) found that an increase in temperature indeed triggered the gene expression of energy metabolism related pathways in *Temora longicornis*. All these findings might indicate an increased demand in energy in higher water temperatures. Other possible actors triggering a difference in energy demand can be salinity, conductivity, phytoplankton densities, (abiotic stressors, etc. While this study presents a first effort to reveal the impact on environmental conditions on community gene expression, more detailed research (across spatial and temporal scales) is necessary to better understand the triggers of changing environmental factors on this community.

Our results represent one of the first analyses of ncRNAs in marine zooplankton communities (Table S3). Small ncRNAs play key roles (1) in the modification and processing of ribosomal RNA and (2) post-transcriptional gene regulation of eukaryotes, with at least 50% of all eukaryote genes being the targets of small RNAs (He et al., 2015; Scott and Ono, 2011; Zamore, 2007). For crustaceans, studies have shown that some unique small RNAs are involved in the immune response as well, leading to lower or higher immune responses to infections and diseases (He et al., 2015). Hence, the identified ncRNAs in this study might provide an important resource for future studies.

Even though transcriptome data does not reflect gene functions or species diversity proportionally, it is thought that the frequencies of expressed genes in a sample reflect the activities of functional genes, or for barcode reads specifically the frequencies of the detected taxa (Ogura et al., 2011). We conducted taxonomy profiling on the metatranscriptome to identify the active members in the community, using the commonly used barcode region 18S. The taxa we identified by microscopy (Table S3) were all detected by metatranscriptomics as well, but with a higher resolution (i.e. to the species/genus level) compared to microscopy (Table S2, Table S3). This is due to the capacity of molecular methods to not discriminate between different developmental stages, such as larvae of many crustaceans (e.g. of barnacles, crabs or copepods) of which morphological identification to the species level is more complicated compared to the adult life stage. In addition, metatranscriptomics is capable of taking into account endosymbionts, such as endoparasites. Based on the OTUs, we were able to calculate diversity indices, such as the Shannon diversity (H') index (Table 2). Van Ginderdeuren et al. (2014) reported a Shannon H' , ranging from 1.5 to 2.5 for the zooplankton community in the BPNS. Our H' richness estimations for the pelagic crustacean community (Table 2) are in this range. Approximately 50–55% (winter) and 90% (summer) of the identified reads were related to copepod sequences. On average, Van Ginderdeuren et al. (2014) found that Copepoda comprise up to 66% of total zooplankton abundances in the BPNS (including non-crustacean groups, such as cnidarians), supporting our finding as copepods being the most abundant taxon in the zooplankton community of the BPNS. Moreover, the most abundant pelagic species in the North Sea, i.e. the omnipresent calanoid copepods *Acartia clausi* Giesbrecht and *T. longicornis* (Deschutter et al., 2017; Van Ginderdeuren et al., 2014), were also represented by the most SSU and COI reads in the metatranscriptome. Comparing the two samples, we observed a low number of *T. longicornis* individuals in the summer sample (pers. obs.; Table S3), which is also translated into a lower number of expressed OTUs. This is in line with the lower mapping percentage of the summer sample replicates against the *T. longicornis* transcriptome (Table 1). This decline in abundance of *Temora* numbers, at the end of the summer, was already described by Deschutter et al. (2017), Fransz et al. (1989) and Van Ginderdeuren et al. (2014) and is the result of both a reduced fecundity (low egg production rates) and an increase of the mortality rate (Fransz et al., 1989). Other detected copepod species in our metatranscriptome were calanoids *Paracalanus parvus* Claus, *Centropages* spp., *Pseudocalanus elongatus* Boeck, *Calanus helgolandicus* Claus and the harpacticoid *Euterpina acutifrons* Dana. These findings correspond with earlier observations in the BPNS made by Daro et al. (2006), Van Ginderdeuren

et al. (2014) and Deschutter et al. (2017). Also, in Dutch coastal waters, calanoids *T. longicornis*, *A. clausi* and *Centropages hamatus* have been ranked as the stock-forming copepod species (Fransz, 1975). Bonnet et al., 2005 reported a peak in autumn for *C. helgolandicus*. This is supported in the taxonomy analysis, as almost 15% of the barcode transcripts could be aligned to *C. helgolandicus*, while being absent in February. It also supports the findings of (Reid et al., 2003), who described the progressively northward shift of *Calanus finmarchicus* and its replacement by *C. helgolandicus* in the southern North Sea due to global warming. These *Calanus* congeners are morphologically almost indistinguishable and a lot of confusions exist in older literature between the two species, making morphobased identification of these species challenging (Van Ginderdeuren et al., 2014). Hence, monitoring the whereabouts of these two species via molecular techniques, such as DNA-barcoding (e.g. in Hill et al., 2001), could be a better alternative. Also, barnacle nauplii and cyprid larvae (Cirripedia) were relatively abundant in the winter sample (Table S3) and could be detected by the taxonomic analysis of the metatranscriptome as well. SSU reads were detected by Metaxa2 for *Balanus crenatus* Bruguière and *Semibalanus balanoides* L., two common barnacle species in the North Sea. In the summer zooplankton sample, we also found one megalopa larva (Table S3), later assigned to the Common spider crab, *Maja brachydactyla* Balss, by the BLAST analysis. No other zoea and/or megalopa larvae of Decapoda (meroplankton) were found in both zooplankton samples. Notably absent in our plankton samples were Branchiopoda (holoplanktonic, i.e. *Podon* and *Evadne* spp.). Consistent with our findings, Van Ginderdeuren et al. (2014) reported their absence in winter periods, yet we did not find any specimens in the summer sample as well. Next to crustaceans, SSU profiling also revealed the presence of the rhizarian *Paradinium* (Rhizaria) in our winter sample, a genus of parasitic protists, known to be infesting both calanoids (*Acartia clausi*) and harpacticoids (*Euterpina acutifrons*) (Skovgaard et al., 2005, 2008). These copepod species were also detected by the taxonomy analysis, explaining the presence of these non-crustacean barcodes in our transcriptome. The parasites exist as 8–12 µm long amoeboid cells in the body cavity of its host, developing a multicellular plasmodium (Skovgaard et al., 2005, 2008). Eventually, the plasmodium passes from the body cavity of the host into the lumen of its intestine, whereupon it is expelled through the anus, eventually followed by the formation of flagellated free-swimming zoospores (Skovgaard et al., 2005, 2008).

Biomonitoring programs evaluating environmental status over time usually rely on morphological taxonomy. However, such evaluations are expensive, time-consuming and require taxonomic expertise. Yet, the Marine Strategy Framework Directive (MSFD, 2008/56/EC) and the European Water Framework Directive (WFD, 2000/60/EC) have emphasized the need to develop more cost-effective and reliable tools to assess the marine environmental status faster (Aylagas et al., 2018; Heiskanen et al., 2016). This is of major importance to allow for an effective conservation strategy and/or management of marine resources. Rapid biological monitoring, using molecular techniques, gives the opportunity to overcome the limitations of taxonomy-based biomonitoring. Given the ability to perform sequencing and data analyses rapidly, cost-effective biomonitoring solutions could be developed in matters of days instead of months: Quinn et al. (2016) for example was able to develop a novel integrated omics pipeline for the analysis of human and environmental samples in less than 48 h. Several attempts have also been made to use metabarcoding for rapid biomonitoring of marine ecosystems (e.g. Aylagas et al., 2018; Pawlowski et al., 2014). However, until recently, molecular analyses in such studies have always been performed in the lab with a delay after sampling in the field, and this with varying time intervals. Even though molecular studies are growing in popularity, one fundamental challenge for rapid biomonitoring remains the transport of the biological material to a laboratory for DNA/RNA extractions and sequencing. Performing genetic analyses on site could help to avoid delays and decrease sample quality decline. The MinION weighs less than 100 g, is easily transportable and

is powered to sequence DNA/RNA using the USB port on a laptop, hence making it suitable for mobile research setups and real time monitoring campaigns onsite. Due to the MinION's real-time sequencing, reads can be analysed as they are generated, significantly speeding up analyses and detection of species. Several *in situ* studies, using the MinION, have been published already: e.g. real-time DNA barcoding in a rainforest (Pomerantz et al., 2018), *in situ* field DNA sequencing in Canadian arctic permafrost ice microbial communities (Goordial et al., 2017), in-field metagenomic detection of arboviruses (Batovska et al., 2017), etc. These studies have focused on DNA sequencing techniques, however due to the complex library preparation protocol for direct RNA sequencing, onsite metatranscriptomic studies without specialised equipment are not yet feasible *in situ*. Nevertheless, metatranscriptomics data sets are able to capture species richness, and how key taxa consistently respond to specific environmental variation as well. If current issues can be overcome, *in situ* and real time biomonitoring using transcriptomics, becomes realistic. The metatranscriptome data described in this study provides a first, valuable sequence resource for scientists investigating the characteristics of pelagic zooplankton, including the study of its biodiversity. Future efforts will hopefully enable more complete descriptions of the zooplankton metatranscriptome from different conditions and marine environments.

5. Conclusion

This study reports a pilot metatranscriptomic dataset of a marine, pelagic crustacean community. We have demonstrated the feasibility of characterizing gene transcripts using nanopore direct RNA sequencing. Taxonomic analysis based on expressed 18S ribosomal RNA genes identified calanoid copepods as the most abundant/active community members. GO and KEGG term annotation revealed that mostly genes involved in glycolytic and respiration-related processes were expressed in the community. This is a first insight into the genomic response of the zooplankton community on changing environmental conditions.

Declaration of competing interest

There are no interests to declare.

Acknowledgements

The authors are grateful to the Flanders Marine Institute (VLIZ) for the assistance, in particular André Cattrijsse, Jonas Mortelmans, Jolien Goossens and the crew of the research vessel Simon Stevin. The authors also thank Klaas De Baerdemaeker, Nicolas de Fooz, Goedele Luyten, Hanne Windels, Jan-Lucas Benoit, Maxime Deprince and Jolien Depecker for the technical assistance. Jana Asselman was awarded an FWO (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen) post-doctoral fellowship. Funding was received from FWO (FWO KAN 1521617N awarded to J.A.). Finally, we would like to thank the anonymous referees for their comments that have greatly improved the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.marenvres.2019.104836>.

References

- Abdulwahab, S., Rabee, A.M., 2015. Ecological factors affecting the distribution of the zooplankton community in the Tigris River at Baghdad region, Iraq. *Egyptian J. of Aquatic Research* 41, 187–196.
- Alexander, H., Jenkins, B.D., Rynearson, T.A., Dyhrman, S.T., 2015. Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2182–E2190.
- Aylagas, E., Borja, Á., Muxika, I., Rodríguez-Ezpeleta, N., 2018. Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecol. Indic.* 95, 194–202.
- Barnes, M.A., Turner, C.R., 2016. The ecology of environmental DNA and implications for conservation genetics. *Conserv. Genet.* 17 (1), 1–17. <https://doi.org/10.1007/s10592-015-0775-4>.
- Bashirades, S., Zilberman-Schapira, G., Elinav, E., 2016. Use of metatranscriptomics in microbiome research. *Bioinf. Biol. Insights* 10, 19–25.
- Batovska, J., Lynch, S.E., Rodoni, B.C., Sawbridge, T.I., Cogan, N.O., 2017. Metagenomic arbovirus detection using MinION nanopore sequencing. *J. Virol. Methods* 249, 79–84.
- Batten, S., Clark, R., Flinkman, J., Hays, G., John, E., John, A.W., Jonas, T., Lindley, J.A., 2003. CPR sampling: the technical background, materials and methods, consistency and comparability. *Prog. Oceanogr.* 58, 193–215. <https://doi.org/10.1016/j.pocean.2003.08.004>.
- Bengtsson-Palme, J., Hartmann, M., Eriksson, K.M., Pandal, C., Thorell, K., Göran, D., Larsson, J., Nilsson, R.H., 2015. METAXA2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Mol. Ecol. Resour.* 15, 1403–1414.
- Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W., de Bruyn, M., 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol. Evol.* 29 (6), 358–367.
- Bonnet, D., Richardson, A., Harris, R., Hirst, A., Beaugrand, G., Edwards, M., et al., 2005. An overview of *Calanus helgolandicus* ecology in European waters. *Progress In Oceanography* 65 (1), 1–53. <https://doi.org/10.1016/j.pocean.2005.02.002>.
- Carotenuto, Y., Dattolo, E., Lauritano, C., Pisano, F., Sanges, R., Miralto, A., Procaccini, G., Ianora, A., 2014. Insights into the transcriptome of the marine copepod *Calanus helgolandicus* feeding on the oxyllipin-producing diatom *Skeletonema marinoi*. *Harmful Algae* 31, 153–162.
- Castellani, C., Altunbaş, Y., 2013. Seasonal change in acclimatised respiration rate of *Temora longicornis* (Müller). *Mar. Ecol. Prog. Ser.* 500 <https://doi.org/10.3354/meps10661>.
- Celis, J.E., Kruhöffer, M., Gromova, I., Frederiksen, C., Ostergaard, M., Thykjaer, T., Gromov, P., Yu, J., Palsdottir, H., Magnusson, N., Orntoft, T.F., 2000. Gene expression profiling: monitoring transcription and translation products using DNA micro-arrays and proteomics. *FEBS Lett.* 480, 2–16.
- Chiba, S., Batten, S., Martin, C.S., Ivory, S., Miloslavich, P., Weatherdon, L.V., 2018. Zooplankton monitoring to contribute towards addressing global biodiversity conservation challenges. *J. Plankton Res., fby030*.
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21 (18), 3674–3676.
- Daro, M.H., Breton, E., Antajan, E., Gasparini, S., Rousseau, V., 2006. Do *Phaeocystis* colony blooms affect zooplankton in the Belgian coastal zone? In: Rousseau, V., et al. (Eds.), *Current Status of Eutrophication in the Belgian Coastal Zone*, pp. 61–72.
- De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C., 2018. NanoPack: visualizing and processing long read sequencing data. *Bioinformatics* 34 (15), 2666–2669.
- Deschutter, Y., Everaert, G., De Schampelaere, K.A.C., De Troch, M., 2017. Relative contribution of multiple stressors on copepod density and diversity dynamics in the Belgian part of the North Sea. *Mar. Pollut. Bull.* 125 (1–2), 350–359.
- Drira, Z., Kmiha-Megdiche, S., Sahnoun, H., Tedetti, M., Pagano, M., Ayadi, H., 2017. Copepod assemblages as a bioindicator of environmental quality in three coastal areas under contrasted anthropogenic inputs (Gulf of Gabes, Tunisia). *J. Mar. Biol. Assoc. U. K.* 1–17. <https://doi.org/10.1017/S0025315417001515>.
- Ferdous, Z., Mukhtadir, A., 2009. A review: potentiality of zooplankton as bioindicator. *Am. J. Appl. Sci.* 6 (10), 1815–1819.
- Franz, H.G., 1975. The spring development of calanoid copepod populations in the Dutch coastal waters as related to primary production. In: Persoone, G., Jaspers, E. (Eds.), *10th Eur. Mar. Biol. Symp. Universal Press, Wetteren*, pp. 247–269.
- Franz, H.G., Gonzalez, S.R., Klein Breteler, W.C.M., 1989. Fecundity as a factor controlling the seasonal population cycle in *Temora longicornis* (Copepoda, Calanoida). In: Ryland, J.S., et al. (Eds.), *Reproduction, Genetics and Distributions of Marine Organisms: 23rd European Marine Biology Symposium, School of Biological Sciences, University of Wales, Swansea*, pp. 5–9. September 1988. *International Symposium Series*, pp. 83–90.
- Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., et al., 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* 15, 201–206.
- Gaudy, R., 1975. Etude de la respiration chez des copépodes pélagiques méditerranéens (bassin Occidental et mer Ionienne) et de ses variations en fonction de la bathymétrie des espèces et de leur origine géographique. *Mar. Biol.* 29, 109–118.
- Goordial, J., Althuler, I., Hindson, K., Chan-Yam, K., Marcollef, E., Whyte, L.G., 2017. *In Situ* field sequencing and life detection in remote (79 degrees 26' N) Canadian high arctic permafrost ice wedge microbial communities. *Front. Microbiol.* 8.
- Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36 (10), 3420–3435.
- Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., Eddy, S.R., 2003. Rfam: an RNA family database. *Nucleic Acids Res.* 31, 439–441.
- Guo, W., Zou, L.F., Li, Y.R., Cui, Y.P., Ji, Z.Y., Cai, L.L., Zou, H.S., Hutchins, W.C., Yang, C.H., Chen, G.Y., 2012. Fructose-bisphosphate aldolase exhibits functional roles between carbon metabolism and the hrp system in rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. *PLoS One* 7 (2), e31855. <https://doi.org/10.1371/journal.pone.0031855>.

- He, Y., Ju, C., Zhang, X., 2015. Roles of small RNAs in the immune defense mechanisms of crustaceans. *Mol. Immunol.* 68, 399–403.
- Heiskanen, A.-S., Berg, T., Uusitalo, L., Teixeira, H., Bruhn, A., Krause-Jensen, D., Lynam, C.P., Rossberg, A.G., Korpinen, S., Uyarra, M.C., Borja, A., 2016. Biodiversity in marine ecosystems—European developments toward robust assessments. *Front. Mar. Sci.* 3.
- Helbling, D.E., Ackermann, M., Fenner, K., Kohler, H.P., Johnson, D.R., 2012. The activity level of a microbial community function can be predicted from its metatranscriptome. *ISME J.* 6, 902–904.
- Hewson, I., Eggleston, E.M., Doherty, M., Lee, D.Y., Owens, M., Shapleigh, J.P., Cornwell, J.C., Crump, B.C., 2014. Metatranscriptomic analyses of plankton communities inhabiting surface and subpycnocline waters of the Chesapeake Bay during oxic-anoxic-oxic transitions. *Appl. Environ. Microbiol.* 80 (1), 328–338.
- Hill, R.S., Allen, L.D., Bucklin, A., 2001. Multiplexed species-specific PCR protocol to discriminate four *N. Atlantic Calanus* species, with an mtCOI gene tree for ten *Calanus* species. *Mar. Biol.* 139, 279–287.
- Hovinen, J.E.H., Welcker, J., Rabindranath, A., Brown, Z.W., Hop, H., Berge, J., Steen, H., 2014. At-sea distribution of foraging little auks relative to physical factors and food supply. *Mar. Ecol. Prog. Ser.* 503, 263.
- Jenjaroenpun, P., Wongsurawat, T., Pereira, R., Patumcharoenpol, P., Ussery, D.W., Nielsen, J., Nookaew, I., 2018. Complete genomic and transcriptional landscape analysis using third-generation sequencing: a case study of *Saccharomyces cerevisiae* CEN.PK113-7D. *Nucleic Acids Res.* 46 (7), e38.
- Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Yu, D.W., 2013. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecol. Lett.* 16 (10), 1245–1257. <https://doi.org/10.1111/ele.12162>.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109–D114.
- Kikuchi, A., Nakazato, T., Ito, K., Nojima, Y., Yokoyama, T., Iwabuchi, K., et al., 2017. Identification of functional enzyme genes of the silkworm *Bombyx mori* from public databases with a combination of dry and wet bench processes. *BMC Genomics* 18 (1), 83. <https://doi.org/10.1186/s12864-016-3455-y>.
- Kopf, A., Kostadinov, I., Wichels, A., Quast, C., Glöckner, F.O., 2015. Metatranscriptome of marine bacterioplankton during winter time in the North Sea assessed by total RNA sequencing. *Marine Genomics* 19, 45–46.
- Krehenwinkel, H., Wolf, M., Lim, J.Y., Rominger, A.J., Simison, W.B., Gillespie, R.G., 2017. Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Sci. Rep.* 7 (1), 17668.
- Lavelle, A., Sokol, H., 2018. Gut microbiota: beyond metagenomics, metatranscriptomics illuminates microbiome functionality in IBD. *Nat. Rev. Gastroenterol. Hepatol.* 15 (4), 193–194. <https://doi.org/10.1038/nrgastro.2018.15>.
- Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics.* <https://doi.org/10.1093/bioinformatics/bty191>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25, 2078–2079.
- McGrath, K.C., Thomas-Hall, S.R., Cheng, C.T., Leo, L., Alexa, A., Schmidt, S., Schenk, P. M., 2008. Isolation and analysis of mRNA from environmental microbial communities. *J. Microbiol. Methods* 75, 172–176.
- Merkes, C.M., McCalla, S.G., Jensen, N.R., Gaiowski, M.P., Amberg, J.J., 2014. Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. *PLoS One* 9, e113346. <https://doi.org/10.1371/journal.pone.0113346>.
- Miller, D., Pfreundt, U., Hou, S., Lott, S.C., Hess, W.R., Berman-Frank, I., 2017. Winter mixing impacts gene expression in marine microbial populations in the Gulf of Aqaba. *Aquat. Microb. Ecol.* 80, 223–242.
- Nawrocki, E.P., Eddy, S.R., 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29, 2933–2935.
- Nawrocki, E.P., Burge, S.W., Bateman, A., Daub, J., Eberhardt, R.Y., Eddy, S.R., Floden, E.W., et al., 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* 43, D130–D137.
- Nielsen, T.G., Løkkegaard, B., Richardson, K., Pedersen, F.B., Hansen, L., 1993. Structure of plankton communities in the Dogger Bank area (North Sea) during a stratified situation. *Mar. Ecol. Prog. Ser.* 95 (1–2), 115–131.
- Nunez, J.C.B., Elyanow, R.G., Ferranti, D.A., Rand, D.M., Oleksiak, M.F., Rajora, O.P., 2018. Population genomics and biogeography of the northern acorn barnacle (*Semibalanus balanoides*) using pooled sequencing approaches. *Population Genomics: Marine Organisms*. Springer, Cham. https://doi.org/10.1007/13836_2018.58.
- Ogura, A., Lin, M., Shigenobu, Y., Fujiwara, A., Ikeo, K., Nagai, S., 2011. Effective gene collection from the metatranscriptome of marine microorganisms. *BMC Genomics* 12. <https://doi.org/10.1186/1471-2164-12-S3-S15>.
- Oxford Nanopore Technologies, 2018. September 20) direct RNA sequencing (SQK-RNA001). Available at: https://community.nanoporetech.com/protocols/direct-rna-sequencing/v/drs_9026_v1_rev_p15dec2016.
- Oxford Nanopore Technologies, 2019. September 20) Guppy on windows. Available at: <https://community.nanoporetech.com/posts/guppy-3-2-4-release>.
- O'Brien, T.D., Lorenzoni, L., Isensee, K., Valdes, L., 2017. What are Marine Ecological Time Series telling us about the ocean? A status report. *IOC Tech. Ser.* 129, 1–297.
- Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T., Wilding, T.A., 2014. Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. *Mol. Ecol. Resour.* 14 (6), 1129–1140.
- Pomerantz, A., Peñañiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L.A., et al., 2018. Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. *GigaScience* 7 (4), giv033.
- Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530 (7589), 228–232.
- Quinn, R.A., Navas-molina, J.A., Hyde, E.R., Song, J., Vázquez-baeza, Y., Humphrey, G., Gaffney, J., Minich, J.J., Melnik, A.V., Herschend, J., 2016. From sample to multi-omics conclusions in under 48 h. *mSystems* 1 (2). <https://doi.org/10.1128/mSystems.00038-16.e00038-16>.
- Reid, P., Edwards, M., Beaugrand, G., Skogen, M., Stevens, D., 2003. Periodic changes in the zooplankton of the North Sea during the twentieth century linked to oceanic inflow. *Fisheries Oceanography* 12 (4–5), 260–269. <https://doi.org/10.1046/j.1365-2419.2003.00252.x>.
- Richardson, A.J., 2008. In hot water: zooplankton and climate change. *ICES (Int. Council. Explor. Sea) J. Mar. Sci.* 65 (3), 279–295.
- Roemmich, D., McGowan, J., 1995. Climatic warming and the decline of zooplankton in the California current. *Science* 267 (5202), 1324.
- Russell, J.A., Campos, B., Stone, J., Blosser, E.M., Burkett-Cadena, N., Jonathan, L., 2018. Unbiased strain-typing of arbovirus directly from mosquitoes using nanopore sequencing: a field-forward biosurveillance protocol. *Jacobs Scientific Reports* 8. Article number: 5417.
- Sanders, R., Mason, D.J., Foy, C.A., Huggett, J.F., 2014. Considerations for accurate gene expression measurement by reverse transcription quantitative PCR when analysing clinical samples. *Anal. Bioanal. Chem.* 406 (26), 6471–6483.
- Scott, M.S., Ono, M., 2011. From snoRNA to miRNA: dual function regulatory non-coding RNAs. *Biochimie* 93 (11), 1987–1992.
- Semmouri, I., Asselman, J., Van Nieuwerburgh, F., Deforce, D., Janssen, C., De Schampelaere, K., 2019. The transcriptome of the marine calanoid copepod *Temora longicornis* under heat stress and recovery. *Mar. Environ. Res.* 143, 10–23. <https://doi.org/10.1016/j.marenvres.2018.10.017>.
- 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.
- Skovgaard, A., Daugbjerg, N., 2008. Identity and systematic position of *Paradinium poucheti* and other *Paradinium*-like parasites of marine copepods based on morphology and nuclear-encoded SSU rDNA. *Protist* 159, 401–413.
- Skovgaard, A., Massana, R., Balagué, V., Saiz, E., 2005. Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist* 156, 413–423.
- Trivedi, S., Aloufi, A.A., Ansari, A.A., Ghosh, S.K., 2016. Role of DNA barcoding in marine biodiversity assessment and conservation: an update. *Saudi J. Biol. Sci.* 23 (2), 161–171.
- Tyler, A.D., Mataseje, L., Urfano, C.J., Schmidt, L., Antonation, K.S., Mulvey, M.R., Corbett, C.R., 2018. Evaluation of Oxford nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Sci. Rep.* 8, 10931.
- Urban, J.M., Bliss, J., Lawrence, C.E., Gerbi, S.A., 2015. Sequencing Ultra-long DNA Molecules with the Oxford Nanopore MinION. *bioRxiv* 19281.
- Van Ginderdeuren, K., Van Hoey, G., Vincx, M., Hostens, K., 2014. The mesozooplankton community of the Belgian shelf (North Sea). *J. Sea Res.* 85, 48–58.
- Vidal, J., 1980. Physioecology of zooplankton. III. Effects of phytoplankton concentration, temperature, and body size on the metabolic rate of *Calanus pacificus*. *Mar. Biol.* 56 (3), 195–202.
- Viehweiger, A., Krautwurst, S., Lamkiewicz, K., Madhugiri, R., Ziebuhr, J., Hölzer, M., Marz, M., 2018. Nanopore direct RNA sequencing reveals modification in full-length coronavirus genomes. *Biorxiv*. <https://doi.org/10.1101/483693>.
- Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H., Voget, S., Simon, M.L., Daniel, R., 2014. Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.* 87 2, 378–389.
- Wilson, J.-J., Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., 2018. High-throughput terrestrial biodiversity assessments: mitochondrial metabarcoding, metagenomics or metatranscriptomics? Mitochondrial DNA Part A. <https://doi.org/10.1080/24701394.2018.1455189>.
- Yang, J., Zhang, X., Zhang, W., Sun, J., Xie, Y., Zhang, Y., Burton, G.J., Yu, H., 2017. Indigenous species barcode database improves the identification of zooplankton. *PLoS One* 12 (10), e0185697. <https://doi.org/10.1371/journal.pone.0185697>.
- Zamore, P.D., 2007. RNA silencing: genomic defence with a slice of pi. *Nature* 446, 864–865.
- 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, 4.