Top-down and bottom-up controls on microeukaryotic diversity (i.e. amplicon analyses of SAR lineages) and function (i.e. metatranscriptome analyses) assessed in microcosm experiments.

Jean-David Grattepanche[1,+] and Laura A. Katz[1,2,*]
[1] Department of Biological Sciences, Smith College, Northampton, Massachusetts 01063, USA
[2] Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, Massachusetts 01003, USA
[+] current address: Department of Biology, Temple University, Philadelphia, Pennsylvania 19122, USA

[*] **Correspondence:**
Laura A. KATZ
Department of Biological Sciences, Smith College,
44 College Lane, Northampton, Massachusetts 01063, USA
Phone: +1 (413) 585-3825
Fax: +1 (413) 585-3786
Email: lkatz@smith.edu

Running title: microeukaryotic diversity and function

Keywords: protists, phytoplankton bloom, copepods, gene expression, nanoplankton, microplankton.

Abstract

The availability of high-throughput sequencing (HTS) has transformed our understanding of the diversity of microbial eukaryotes (i.e. protists) across diverse habitats. Yet relating this biodiversity to function remains a challenge, particularly in the context of microbial food webs. Here we perform a set of microcosm experiments to evaluate the impact of changing predator and prey concentrations on a marine protist community, focusing on SAR (Stramenopila, Alveolata, and Rhizaria) lineages. We combine an estimate of taxonomic diversity through analysis of SSU-rDNA amplicons with metatranscriptomics, a proxy for function. We assess changes in a community sampled from New England waters with varying concentrations of predators (copepods) and prey (phytoplankton less than 15μm in size). The greatest impact observed is on the diversity and function of the small plankton (2-10μm) community in the presence of high prey abundance (i.e. bloom conditions). Many SAR taxa in the nanosized fraction decrease with increasing phytoplankton abundance, while ciliates (from both nano- and microsized fraction) increase. A large number of transcripts and function estimates in the nanoplankton are downregulated during our simulated phytoplankton bloom. We also find evidence of an interaction between increasing phytoplankton and copepods on the microsized planktonic community, consistent with the hypothesis that phytoplankton and copepods exert bottom-up control and top-down control on the microsized protists, respectively. Together our analyses suggest community function (i.e. diversity of gene families) remains relatively stable, while the function at the species level (i.e. transcript diversity within gene families) show a substantial reduction of function under bloom conditions. Our study demonstrated that interactions within plankton food webs are complex, and that the relationships between diversity and function for marine microeukaryotes remain poorly understood.

Contribution to the field

The number of publications on plankton diversity has increased dramatically over the last decade, but gaps remain in our understanding of why there is such great diversity, and how it relates planktonic food webs. Our study analyzes community turnover within the eukaryotic microbial communities in response to changing food resources (phytoplankton) and predators (copepods). Using microcosms to assess the impact of these trophic pressures on ciliates and other microbial eukaryotes, we characterize biodiversity through amplicon high-throughput sequencing of SAR (Stramenopila, Alveolata, and Rhizaria) clade and function through metatranscriptomics. One of the main results of our study is the surprising and strong impact of simulated phytoplankton bloom on nanosize (2-10μm) eukaryotes, and the absence of substantial change among microplankton (10-80μm). Our analyses also reveal a reduction of expression and diversity of cytoskeletal genes for nanosize eukaryotes and an increase of photosynthesis with high predation pressure. We also believe that the methods we developed will be of interest to community ecologists as well as evolutionary biologists**.**

1 Introduction

Microbes dominate biodiversity and are responsible for key ecosystem functions. They can function as heterotrophs, autotrophs or mixotrophs, and they can be free-living, parasitic (detrimental for host) or symbiotic (beneficial for host). Thus, microbes have critical roles for other organisms including humans (e.g. microbiome including gut flora, human health), other animals (e.g. coral), and plants (e.g. mycorrhiza). While we know that microbes are important, we still have only limited knowledge about eukaryotic microbes and their functions (e.g., what are the main factors driving their diversity). This contrasts to the many tools that have been developed and used to explore the diversity and function of bacteria. Here, we use some of these tools, including metatranscriptomics, and amplicon analyses, to look at the marine planktonic food web.

In marine food webs, the importance of microbes for transfer of energy from bacteria to higher trophic levels is well-established (reviewed in Azam et al., 1983; Edgcomb, 2016; Fenchel, 1988; Menden-Deuer and Kiørboe, 2016; Worden et al., 2015). Many recent studies have characterized the marine plankton diversity on a large scale (e.g. Tara Ocean for plankton in the photic zone and Malaspina for the aphotic zone and others; de Vargas et al., 2015; Irigoien et al., 2004; Pernice et al., 2016) and show a geographical distribution related to oceanic basins. Other studies have looked at patterns on a smaller scale (from meter to kilometer; e.g. Grattepanche et al., 2014, 2016b, 2016a; Mousing et al., 2016) and find that distance alone does not explain the patterns. One of the likely reasons for the mismatch is because these studies assess different processes: studies at smaller scales tend to be more directly impacted by species-specific interactions and functions expressed by each member of the community, while the larger scale studies are designed to look at overall ecosystem processes such as the impact of climate change.

Other studies have looked at species-specific relationships between prey and predators, such as between microzooplankton and phytoplankton or between copepods and microzooplankton. These studies show that microzooplankton such as ciliates and dinoflagellates are efficient grazers of phytoplankton of small size (<10 μm for ciliates and <100μm for dinoflagellates; e.g. Calbet, 2008; Grattepanche et al., 2011b; Martínez et al., 2017; Schoener and McManus, 2017). Microzooplankton have been described as a better food source for copepods than phytoplankton (Berk et al., 1977; Calbet and Saiz, 2005). The fact that microzooplankton are strongly controlled by copepods (and other mesozooplankton) is a possible explanations for phytoplankton spring blooms as preferential grazing of microzooplankton by copepods may allow dramatic increase in phytoplankton abundance (e.g. dilution-recoupling hypothesis; Behrenfeld, 2010; Irigoien et al., 2005; Kuhn et al., 2015).

The impact of phytoplankton on microheterotrophs and other parts of the community has been studied particularly during bloom events (Grattepanche et al., 2011a; Monchy et al., 2012; Rosetta and McManus, 2003). Based on these analyses, microheterotrophs are assumed to be the primary consumer of phytoplankton of small size, with bacteria playing a negligible role (García-Martín et al., 2017). While still in discussion, the phytoplankton spring blooms have been assumed to be linked to an excess in nutrients at the end of the winter (remineralization by bacteria and upwelling), combined with a lack of control by predators including heterotrophic protists and higher trophic levels. On the other hand, microheterotrophs are efficient grazers when blooms start to decline (Archer et al., 2000; Laws et al., 1988). As such, questions remain concerning how the community and function change in relation to increasing phytoplankton abundance.

Metatranscriptomics has mainly been deployed to study bacterial functions such as the microbiomes of organisms as diverse as humans (Franzosa et al., 2014) and termites (Tartar et al., 2009), or in marine systems (Gifford et al., 2011) and in response to environmental changes such as oil spills (Mason et al., 2012; Rivers et al., 2013). The studies looking at eukaryotic metatranscriptome are rare. One reason is the paucity of databases of eukaryotic gene functions; current databases contain mainly bacterial functions or functions related to diseases, which result in the majority of the environmental eukaryotic transcripts not being annotated (sometimes more than 80% of the transcripts cannot be annotated; Cooper et al., 2014; Damon et al., 2012; Lesniewski et al., 2012). To study eukaryotic metatranscriptomes, given the paucity in available microbial eukaryote genomes (reviewed in del Campo et al., 2014), authors have focused on clades with reference genomes such as diatoms (Alexander et al., 2015) or fungi (Bailly et al., 2007).

Here, we use a microcosm approach and high-throughput molecular tools, combining amplicon and metatranscriptomic analyses, to assess the impact of phytoplankton (prey) and copepods (predators) on the rest of the planktonic community. To study protist diversity, we use primers designed to amplify the SAR (Stramenopila, Alveolata, Rhizaria; Grattepanche et al., 2018; Sisson et al., 2018) community as this includes many major marine clades such as dinoflagellates, diatoms, and ciliates. We added phytoplankton (i.e. prey) at three concentrations ($5x10^2$, $5x10^3$ and $5x10^4$ cell. mL^{-1}, which mimics marine bloom conditions) and copepods (i.e. predators) at two abundances (5 and 10 copepod. L^{-1}). We predict that the microsized heterotrophic plankton (microheterotrophs thereafter) will consume added phytoplankton, resulting an increase of their contribution, and that copepods will regulate this increase. One part of the community generally ignored in the phytoplankton-microheterophs-copepods link are the nanosized plankton. We therefore specifically address the impact on small plankton (2-10µm) and expect (1) strong competition as we add phytoplankton and (2) absence of impact of copepods as nanosized plankton are too small to be captured.

2 Materials and methods

2.1 Starting materials: community, phytoplankton, copepods

The experiments conducted here used a community sample of microbes sampled in the open water of the Long Island sound in front of University of Connecticut's Avery Point campus (41.30 ºN, 72.06º W) on 15 March 2016. We collected *in situ* water in four 20L carboys, filtering first through an 80µm mesh to remove predators and other large organisms, enabling us to focus on the protist community.

To perform our microcosm experiment, we used four cultures: three phytoplankton (prey) cultures *Tetraselmis chui*, *Isochrysis galbana*, and *Phaeodactylum tricornutum*; and a culture of *Acartia tonsa* (copepod, predators). The species of phytoplankton used are common in aquaculture application, represented various size, shape, pigment composition and lineages, and has been used successfully in the past (Grattepanche et al., 2019; McManus et al., 2004, 2012).. All phytoplankton were from the culture collection of the National Marine Fisheries Service Laboratory in Milford CT (USA) and were grown in F/2 medium prior to experiments. The copepods were picked from cultures maintained at the University of Connecticut Department of Marine Sciences. In order to achieve the final concentration of phytoplankton for our microcosm experiment, we concentrated the combined phytoplankton by centrifugation at 3,9g to create a

stock concentration of $5x10^7$ cells. mL^{-1} with the three algae at the same order of magnitude (*P. tricornutum* was added at 1/5 of the concentration of the two other species).

2.2 Experimental set-up

We used dialysis tubing to perform our microcosm experiment and incubated samples in a sea table with circulating water from collections sites to minimize impact of isolation, following the approach of Grattepanche et al. (2019). The dialysis tubing is composed of a cellulose membrane that is pervious to molecules <12,000 molecular weight (product D9402, Sigma), allowing for exchange of nutrients during incubation in sea tables. Each microcosm (tied-off dialysis tubing) contained one liter and was floated in a sea table with continuous *in situ* seawater circulation as recommended by Capriulo (1982).

To test impact of varying levels of predators and food resources we pooled our four water samples together, and then divided back into four carboys of 20L to which we added: 1) nothing (control; p0), phytoplankton at 2) $5x10^2$ cell. mL^{-1} (p1), 3) $5x10^3$ cell. mL^{-1} (p2), and 4) $5x10^4$ cell. mL^{-1} (p3) in final concentration. For each of the carboys, we filled 9 microcosms with 1L of the seawater with (or without) phytoplankton added. For each of the phytoplankton abundance conditions, including controls, we added zero (z0), 5 (z1) or 10 copepods (z2) to triplicate microcosm for each treatment (3 sets of dialysis tubing per phytoplankton/zooplankton combination). Based on insights from a preliminary study (Grattepanche et al., 2019), the microcosms were incubated for 3 days and then the total contents of each bag were collected for nucleic acids extractions. We used a three-day incubation to minimize cascade effect and maximize nano- and microsized plankton impact based on insights from Grattepanche et al. (2019). We also collected duplicates of the initial *in situ* water sample such as the *in situ* sample plus high phytoplankton abundance added (t0p3) at the beginning of the experiment.

2.3 Sampling and nucleic acids extraction

We isolated organisms of whole content of each dialysis bag on polycarbonate filters after size-selection, cut filters in two and isolated DNA and RNA using appropriate kits. Prior to isolation of nucleic acids, each sample passed through serial filtration at 80µm (to remove predators and larger organisms to focus on small protists and avoid some PCR inhibitors) and collected on 10μm (i.e. microsized fraction) and 2μm (i.e. nanosized fraction). The filters do not exactly line up with current definitions of nano- and microplankton as some have argued picoplankton are up to 3μm in diameter (Vaulot et al., 2008) while nanoplankton can include species up to 20μm (Sieburth et al., 1978). We use the terms 'microsize' and 'nanosize' as the bulk of the lineages caught on our filters likely fall in these sizes classes, though the variation in the filters, the differential flexibility of species and the irregularity of some body plans all confound the efficiency of filtering.

For the nucleic acid extraction, we cut the resulting filters in half, storing one half for DNA (DNA prep buffer; 100mM NaCl, Tris-EDTA at pH 8, and 0.5% of SDS) and the other half for RNA (RLT lysis buffer [Qiagen, Germany] plus beta mercaptoethanol). DNA and RNA were extracted using Zymo Research soil extraction kit (Zymo Research, CA) and Qiagen RNeasy (Qiagen, Germany) kits following the manufacturer's instruction. For the RNA prep, residual DNA was removed using the Turbo DNA-free kit (Invitrogen, CA). This approach (both filtration and extraction) was tested on previous samples and resulted in good DNA/RNA quality (Grattepanche et al., 2019; Sisson et al., 2018; Tucker et al., 2017).

2.4 Amplicon sequencing for community composition

The amplicon analyses followed protocols from Grattepanche et al. (2016b) and Sisson et al. (2018). In sum, we used a primer set specific of the SSU-rDNA gene of the SAR lineages, which amplified a 150 bp fragment of the hypervariable region V3, to amplify DNAs extracted from filters. As discussed in Grattepanche et al. (2019), these primers amplify almost all known SAR lineages, excepted the highly divergent Foraminifera, which are also removed by our prescreening on 80µm). The amplification was done with the Q5 polymerase enzyme (NEB, MA) following manufacture's protocol. PCR products were cleaned using Agencourt AMPure XP beads (Beckam Coulter, CA) and sent to University of Rhode Island Genomics and Sequencing Center for sequencing on an Illumina MiSeq sequencer (2x150 cycles). This resulted in a dataset of 7,207,909 reads. The raw reads are available from NCBI under the BioProject PRJNA550423 and the Sequence Read Archive SRP212194.

The Amplicon dataset was analyzed using a pipeline combining third party tools and custom python scripts (Grattepanche et al., 2019; Sisson et al., 2018). We first generated paired-end reads using Paired-End reAd mergeR (Zhang et al., 2013), refined the sequences (removed unpaired sequences; sequences without primers), created an OTU library using SWARM (v 2.1.9; Mahe et al., 2015), refined the OTUs (removed chimeric OTUs, outgroup OTUs, and OTUs with less than 10 reads), and then assigned taxonomy by tree using a curated SAR SSU-rDNA gene database. The final step corresponds to the rarefaction i.e., subsampling a fixed number of reads for each sample to enable comparison. This resulted in a total of 1,148 OTUs corresponding to 3,400,000 rarefied reads (50,000 reads per sample).

To compare the effect of experimental conditions on diversity (i.e. taxonomy), only the OTUs present in two replicates of the same treatment and the same size fraction, and with more than 5 reads were kept. The reported number of reads for each treatment and size corresponds to the average of the read number among the three replicates, resulting in a total of 277 OTUs. We used this set of OTUs to assess changes between the control and the treatments.

2.5 Metatranscriptome for analyses of community function

To characterize expressed eukaryotic genes (i.e., those with polyA tails) we constructed mRNA libraries using TruSeq Stranded mRNA sample prep (Illumina, CA) following manufacturer protocol. In summary, we isolated RNA using oligo-dT beads, fragmented the transcripts and synthesized first and second strands of DNA (cDNA). The 3' end of the fragments was adenylated to avoid ligation to another fragment, the adapters were ligated, and the fragments were amplified using specific PCR (see TruSeq protocol). DNA purifications were performed between each of these steps using Agencourt(r) AMPure XP beads (Beckam Coulter, CA). The remainder of the protocol was performed at the University of Maryland, Baltimore Institute for Genome Sciences. This included library quantification and quality checking with an Agilent Technologies 2100 Bioanalyzer, pooling, and sequencing with an Illumina HiSeq 4000. We used a similar depth of sequencing for all our samples. While this depth of sequencing does not allow to access the lowly expressed transcripts, the use of similar depth of sequencing across samples allowed us to compare these different samples. This resulted in a dataset of 783,472,526 reads, which are available from NCBI under the BioProject PRJNA550423 and the Sequence Read Archive SRP212194.

The metatranscriptome libraries were assembled using rnaSPAdes (version 3.10.1 with default parameters; Bankevich et al., 2012). We assembled the libraries in 3 ways: (1) each library independently (replicate set), (2) each triplicate by size fraction (nano- and microsized

fractions, separately; treatment set); and (3) all libraries together in order to create a reference database for read number and RPKM calculation (Reads Per Kilobase Million = Reads Per Kilobase of transcript, per Million mapped reads, which is a proxy for gene expression).

As the reference genomes of microbial eukaryotes are still rare (del Campo et al., 2014), we looks at the pattern of conserved gene families using both PhyloToL (Cerón-Romero et al., 2019), which relies on the classification of gene families determined in OrthoMCL (Chen et al., 2006; Li et al., 2003). The resulting 1,485,323 transcripts were then refined using the "gene family assessment for taxa" part of the PhyloTOL pipeline (Cerón-Romero et al., 2018, 2019; Maurer-Alcala et al., 2018). In sum, we removed transcripts smaller than 200bp (77,230 transcripts smaller than 200 bp), ribosomal DNA (5,699 transcripts matching rDNA gene references) and bacterial (38,305 transcripts matching reference bacterial genomes) contamination, bin the remaining sequences in orthologous groups (hereafter OG; proxy for gene families) using OrthoMCL (Chen et al., 2006; Li et al., 2003) as reference (990,632 transcripts did not match our reference gene families).

At this stage, the dataset was composed of 373,457 transcripts. We consider only the transcripts matching an OG with an E-value cutoff of $1e^{-50}$ and a coverage of 10 (coverage as calculated by rnaSPAdes; 18,599 transcripts were discarded). This results in a reference dataset composed of 86,951 transcripts representing 5,258 conserved eukaryotic gene families (i.e. OGs from OrthoMCL; Chen et al., 2006) plus a large number of lineages-specific genes. We then decided to remove transcripts from the phytoplankton we added. For this purpose, we isolated transcripts from the added phytoplankton by removing transcripts absent from the *in situ* samples without added phytoplankton (t0) from the *in situ* sample plus high phytoplankton abundance (t0p3). By this way, we identified 5,159 transcripts from the phytoplankton we added. We also removed the transcripts that were 10 times more expressed in t0p3 than in t0, leading to an additional 158 transcripts. Of the original reference transcripts (generated by combining mRNAs from all treatments), we removed 5,317 transcripts that represent genes expressed by the phytoplankton added resulting in a finally reference dataset of 81,634 transcripts. We then use BLAST2Go to assign Gene Ontology (GO; Conesa et al., 2005) and eggNOG (Huerta-Cepas et al., 2018) to assign Clusters of Orthologous Groups (COG; Tatusov et al., 2000) and KEGG (Kyoto Encyclopedia of Genes and Genomes) orthology (KO; Kanehisa et al., 2015).

To assess the gene expression for the two sets (replicate and treatment sets), we mapped back the reads of each metatranscriptome to the refined reference (all libraries pooled together and refine using PhyloTOL) using Seal (Sequence Expression AnaLyzer, version 35.92) implemented in BBmap (Bushnell, 2016). In short, each read is mapped to our reference using a Kmer of 31, and the count is transformed in Reads Per Kilobase Million (RPKM) to take in account the depth of sequencing for each sample. This software and a custom python script produce a table with the RPKM (read number or gene expression) of each reference transcript present in each of our samples (similar to an 'OTU table').

2.6 Statistics

To assess the effect of each treatment on the community composition, we used the Unifrac dissimilarity index (Hamady et al., 2009; Lozupone et al., 2011) and Principal Coordinates Analysis (PCoA). The analyses were performed in R using the Phyloseq (McMurdie and Holmes, 2013) and vegan packages (Oksanen et al., 2007, 2016). We tested the significance of the PCoA axis and factor (phytoplankton or predators) using envfit implemented in vegan, in which the data are randomly permuted 999 times (total of 1,000 datasets) and the random data

are compared to the data generated through a fitted regression model. Under the null hypothesis of no relationship between the ordination "axis" scores and the environmental variable, the observed R^2 value should be a value among the permuted R^2 values. However, if the observed R^2 is extreme compared to the permutation distribution of R^2 then we can reject the null hypothesis. The proportion of times a randomized R^2 from the distribution is equal to or greater than the observed R^2 is a value known as the permutation p value. The same analysis is performed for the Metatranscriptomics data using Canberra dissimilarity index (Lance and Williams, 1967).

Differential gene expression was calculated as (1) difference of read number between control and treatment or (2) ratio of RPKM between control and treatment. If the ratio was above 1, the transcript was upregulated in the treatment compared to the control. If the ratio was below 1, the transcript was downregulated, and we inverted the ratio (e.g. a ratio of 0.5 becomes -2). To be conservative, we only consider up/downregulation when the ratio was larger (or smaller) than 2 (or -2 for downregulation). In addition, we also assessed the OTUs and transcripts significantly up- or down-regulated (log2 fold change with $p<0.05$) using DESeq2 package (Love et al., 2014). The transcripts significantly up- and downregulated assessed with DESeq2 matched with our estimates.

3 Results

3.1 Top-down and bottom-up impact on community composition

The diversity of SAR lineages remained relatively constant despite increasing predation pressure and food availability (Figure 1, Supplementary Figure S1). Because the three replicates were overall similar (Supplementary Figure S1), we pooled replicates to evaluate the impact of treatment and considered only the OTUs (i.e. species) as described in the methods section. We observed that abundant OTUs (i.e. OTUs with more than 10 reads) varied between 63 and 125 OTUs (Figure 1a), and up to 239 OTUs when including the rare OTUs (Figure 2). The richness of abundant OTUs showed no clear pattern in response to phytoplankton abundance or copepods density (Supplementary Figure S1). In other words, the perturbations (adding prey and/or predators) did not have a strong impact on the overall diversity of SAR lineages.

While the overall diversity did not change within our microcosm, the community composition did respond to the various treatments (i.e. food availability and predation pressure), and size. Overall, up to 10% of the nanosized OTUs significantly decreased in read number with increasing phytoplankton abundance, and up to 5% of the OTUs significantly increased in read number with increasing predation pressure (copepods, Figure 2). By comparison, the microsized plankton community showed a more neutral/stochastic response, with no clear relation with increasing phytoplankton and/or copepods (Figure 2). We also observed that 81% of the community variability is explained by differences between the nanosized and microsized plankton ($R^2=0.7894$, $p<0.0001$; Supplementary Figure S2).

Among the nanosized community, the contribution of ciliates relative to stramenopiles increased with increasing phytoplankton (i.e. p0z0 vs. p1z0 vs. p2z0 vs. p3z0 showed an increasing contribution of ciliates; Figure 1a and b). The pattern among microsized species was more complex. For instance, ciliates, especially within Spirotrichea, increased with increasing phytoplankton in the absence of predators (Figure 1). However, this effect appears buffered by the increasing copepods as the ciliate contribution was almost constant when incubated with high predation pressure (i.e. 10 copepods. L^{-1}) regardless of the concentration of phytoplankton (Figure 1). This suggests that while microsized ciliates responded to the increase of food, the

copepods were able to keep them at a constant abundance. By comparison, we only observed a slight impact of copepods on nanociliates (Figure 1). In addition, we observed a slight decrease of phytoplankton (mainly Dictyochophyceae and Coscinodiscophyceae; Figure 1b) with increasing phytoplankton and a more mixed response with increasing copepods (Figure 1, Supplementary Material).

Only a few OTUs, our proxy for species, responded repeatedly to all treatments. The control (no copepod and no phytoplankton added [p0z0]), the high predation pressure (high copepod abundance and no phytoplankton added [p0z2]) and the phytoplankton bloom incubations (high phytoplankton abundance and no copepods added [p3z0]) shared a quarter of the total OTUs (56 of the 212 OTUs observed in p0z0, in p3z0 and in p0z2 ; Supplementary Figure S3). The dominant OTUs did not show clear response to our treatment (see Supplementary Material; Supplementary Figure S2). A few OTUs increased or decreased significantly with food availability including OTU302 (closely related to the stramenopile *Rhizochromulina*) and OTU74 (likely an oligotrich ciliates) or with both food availability and predation pressure such as OTU94 (closely related to the euglyphid *Cyphoderia major*; Supplementary Figure S4). In the same way, only a third of the OTUs are specific to a size fraction and treatment (70 OTUs; Supplementary Figure S3), and these OTUs represented less than 1% of the community based on read number.

To assess changes in the overall community, we used principal coordinate analysis (PCoA) with UniFrac index, which considers phylogenetic relationship among OTUs. Overall, these analyses indicate that predators impacted the microsized SAR community while the phytoplankton changed the nanosized SAR community. Adding copepods significantly impacted the microsized community (phytoplankton R^2=0.3689, p=0.2238; copepods R^2=0.3726, $p < 0.05$ by envfit test, Figures 2B and 2D), while phytoplankton addition had a significant impact on nanoplankton (phytoplankton R^2=0.6446 $p<0.05$; copepods R^2=0.2976 p=0.1638; Figures 2A and 2C). Looking at lower taxonomic levels, we observed that Ciliophora, Stramenopila and within Stramenopila, Dictyochophyceae in the nanosized fraction showed a significant response to phytoplankton treatments, while none of the microsized fraction showed significant response with increasing phytoplankton (Table 1). The copepods seemed to have only impacted the Bacillariophyceae from both nano- and microsized fractions (Table 1).

3.2 Community function by metatranscriptomics

Our metatranscriptomics analysis focused on a dataset composed of 81,634 transcripts that match conserved eukaryotic gene families, and represent 13,029 GO terms, 5,176 conserved eukaryotic gene families, and 3,699 KO terms. We evaluated the impact of our treatment on the community function using the number of conserved gene families (GFs) and transcripts at three levels of expression: present, expressed and highly expressed (>0, >10 and >1,000 as gene expression estimated by RPKM, respectively). The nano- and microsized fractions are composed of a similar number of expressed transcripts (40,667 and 41,369 transcripts, respectively) with half shared in both sizes (21,940 transcripts; see Supplementary Material).

Differential expression profiles showed a clear difference between the two size fractions, and more intriguingly, a large number of transcripts are downregulated in the nanosized community incubated with phytoplankton at bloom conditions (Figure 4). The nanosized plankton showed up to 7 times more downregulated transcripts when incubated in phytoplankton bloom conditions (Figure 4). In addition to transcript number, we also evaluated the impact of our treatments on the community function using three metrics: the number of conserved gene families (GFs), Gene Ontology (GO), and KEGG orthology (KO) terms associated with up- or

down-regulated transcripts (Figure 2 and Supplementary Figure S5). Using all of these measures, we again saw a substantial change in nanosized plankton functions (up to 70% of transcripts and conserved gene families are downregulated in bloom conditions; Figures 2, 4 and Supplementary Figure S5), while the impact on microsized plankton is less marked (Figures 2 and 4 and Supplementary Figure S5).

In comparison to the impact of prey (i.e. phytoplankton), the impact of predators on function is in agreement with our expectations: negligible for the nanosized fraction and leading to an increase in upregulated function in the microsized fraction (Figures 2 and 4; Supplementary Figure S5). Looking in detail at the interactions between phytoplankton bloom and high predation pressure, the increase in copepods density tends to increase the proportion of differentially expressed transcripts (Figures 2 and 4; Supplementary Figure S5). The copepods also increased the number of upregulated functions in the microsized plankton (Figure 2; Supplementary Figure S5). For both sizes, the copepods tend to reduce the proportion of downregulated transcripts and GFs during the phytoplankton bloom incubations (Figures 4 and 2; Supplementary Figure S5). In other words, the impact of copepod counterbalanced the effect of phytoplankton bloom on the community function and *vice-versa*.

Overall patterns of transcripts assessed by PCoA with a Canberra distance metric show the same pattern as the SAR amplicon sequencing, i.e. samples group firstly by size fraction (nano- and microsized) and then by phytoplankton abundance for the nanosized community (35% of the variance expression; Figure 3). The expression within the microsized community is similarly impacted by phytoplankton and copepods (18% vs 16% of the variance, Figure 3). This again suggests that while the nanosized plankton is primarily impacted by the change in phytoplankton abundance, the microsized plankton expression profile is a result of interactions between phytoplankton and copepod abundances.

To further assess community function, we assigned transcripts into functional categories using COG (Cluster of Orthologous Gene). One third of the gene expression (as measured by RPKM) and one fifth of the number of transcripts are involved in translation (category COG J). Another 40% of the transcript expression and transcript number is represented by transcripts from post-translational modification and protein turnover (COG O), unknown function (COG S), cytoskeleton (COG Z), energy production and conversion (COG C) and carbohydrate metabolism and transport (COG G). The COG S includes transcripts related to Fucoxanthin Chlorophyll a/c (55% of the reads for this study), Chlorophyll a/b binding proteins (6% of the read) by similarity.

To evaluate the main functions impacted during our experiments, we pooled the transcripts by COGs and estimate the differential expression in each treatment relative to the control. We observed a downregulation in cytoskeleton mRNA (COG Z in Figure 5) and a slight upregulation in mRNAs involved in translation in the nanosized communities incubated in phytoplankton bloom condition (COG J in Figure 5). The microsized community incubated with many copepods experienced an upregulation of transcripts of “unknown” function (mainly transcripts related to Fucoxanthin chlorophyll a/c by similarity; COG S in Figure 5) and a downregulation of transcripts involved in translation (COG J in Figure 5). The other treatments showed only slight differences and no pattern was discernable between COG and incubation condition.

3.3 Function and taxonomy

Phylogenomic analyses indicate that transcripts strongly impacted in our microcosm are from SAR lineages and other phytoplankton (e.g. chlorophyte, cryptophyte, haptophyte). We

generated phylogenies for eleven exemplar genes that showed high diversity and/or high expression (e.g. three ribosomal proteins, α and β tubulins, actin, HSP70, chlorophyll a/b binding proteins, and glyceraldehyde 3-phosphate dehydrogenase; Supplementary Figure S6). For each of these genes, SAR represents almost 70% of the transcripts; Stramenopila, Alveolata and Rhizaria representing approximately 45%, 40%, and 10% of the highly expressed transcripts, respectively. Other algae (e.g. chlorophytes, haptophytes, crytpophytes) comprise ~25% of the highly expressed transcripts and Opisthokonta, Amoebozoa and Excavata are represented by less than 6% of the highly expressed genes (Supplementary Figure S6).

While ciliates clearly responded to increased phytoplankton in amplicon analyses (Figure 1), our taxonomic analyses of some gene families did not find evidence of their impact on the community function. Ciliates contributed up to 30% to some genes including eEF1α (Supplementary Figure S6). However, ciliates did not show any strong response to treatments for this set of genes except a more important contribution to the actin expression in our phytoplankton bloom treatment (Supplementary Figure S6). Other cytoskeleton proteins, such as α-tubulin and β-tubulin, did not show this pattern. We looked at other genes involved in phagosome formation such as ARP complex and RAC as a proxy of phagotrophy (Yutin et al., 2009) and the protein identify by Burns et al (2018), but again we did not observe a clear pattern across treatments (Supplementary Figure S7). While this is surprising, it is important to remember that many transcripts (almost 1 million) are not considered here as they are not annotated in databases and instead might represent lineage-specific genes. In other words, the impact of ciliates on the community may be through lineage-specific genes that lack annotated homologs in current databases.

4 Discussion

Contrary to our expectation that increasing phytoplankton would have the greatest impact on the microsized heterotrophs (i.e. the potential predators of added phytoplankton), the nanosized community showed the greatest response to bloom levels: we observed an increase of ciliate community members with increasing phytoplankton abundance (Figure 1) and a decrease in community function (i.e. transcript number and expression levels) among the nanosized plankton incubated in bloom condition (Figure 4). The microsized community did respond as predicted, ciliates increased with increasing phytoplankton abundance (i.e. bottom-up control) and the microsized plankton were impacted by increasing copepods (i.e. top-down control). We also demonstrate the power of combining amplicon and metatranscriptomics in microcosm experiments to characterize changes in microeukaryotic communities in response to environmental changes.

4.1 Bloom of phytoplankton and impact on other small plankton

The nanosized plankton (i.e. 2-10 μm) response to increasing phytoplankton (three species ranging in size from 4-15 μm) was unexpected. We added small sized phytoplankton expecting to see an impact on their predators within the microsized (i.e. 10-80 μm) heterotrophs and saw a small increase of microsized ciliates with increasing phytoplankton (Figure 1). In other *in situ* studies, microheterotroph biomass increased with phytoplankton and nanoplankton were responding to bacteria (Grattepanche et al., 2011a; Irigoien et al., 2005). So, we did not expect to observe an increase of nanosized lineages, assessed by SAR amplicon analyses, incubated in bloom conditions (Figure 1). This suggests that either nanosized lineages (i.e. ciliates) were able to consume prey of the same size or that their response is indirect. For example, it is possible that

phytoplankton exudate caused increased growth of bacteria, which are in turn grazed by nanosized ciliates (i.e. more food, more abundant) (Fenchel, 1987). Another possibility is that nanosized ciliate predators fed on the same size phytoplankton, and therefore relieved predation pressure on these small ciliates (i.e. less predators, more abundant).

We evaluate the possibility of lineage-specific changes in predation among nanosized community members by looking at genes involved in phagocytosis. While we observed an upregulation of actin for the nanosized ciliates in bloom condition (Supplementary Figure S6), we did not see an increase in expression for phagotrophic genes such as the ARP complex and RAC gene family identified by Yutin et al (2009) and a longer list of genes identified in two other studies (Burns et al., 2018; McKie-Krisberg et al., 2018; Supplementary FigureS7). These analyses are consistent with an indirect response of nanosized ciliates rather than an increase in grazing in response to added phytoplankton. However, this conclusion needs to be taken with caution as the mechanism of phagotrophy for microbial eukaryotes is still poorly understood (Yutin et al., 2009) and many lineage-specific gene families are likely missing from our database (Cerón-Romero et al., 2019).

4.2 Phytoplankton blooms reduce community function

The transcript number and expression levels for communities incubated in the bloom conditions strongly decreased for the nanosized fraction (Figure 4 and Supplementary Figure S5). A decrease of functional diversity is expected when just a few species dominate the ecosystem (i.e. in bloom conditions), but the observed impact on nanosized and not on the microsized plankton is surprising. One possibility is that changes in ecological niche (abiotic, and biotic factors) in the phytoplankton bloom conditions might have selected for only a subset of the nanosized species present at the beginning of incubation, resulting in a decrease of expressed function. A decrease of diversity and function has been documented for bacterioplankton (Teeling et al., 2012; Wemheuer et al., 2014), but not yet for eukaryotes.

We did see patterns in changes of genes related to photosynthesis and stress. For example, we saw a reduction expression of transcripts related to photosynthesis (COG S, Figure 5; Chlorophyll a/b binding protein; Supplementary Figure S6). In addition, stress-related proteins (e.g. HPS90, HSP70), particularly of phytoplankton groups (e.g. Archaeplastida, Supplementary Figure S6), did increased with increased phytoplankton. This suggests a stress among phytoplankton in response to bloom conditions and is consistent with the observed decrease of SAR phytoplankton (mainly Dictyochophyceae and Coscinodiscophyceae) observed with increasing phytoplankton (Figure 1).

The addition of varying levels of phytoplankton impacted the diversity and function of both the nanoplankton and, to a lesser extent, the microsized plankton. As expected for bloom conditions, we observed an amplification of changes in both up- and downregulated transcripts with increasing phytoplankton (Figure 2). However, none of these genes are directly related to phytoplankton blooms based on findings in previous studies (e.g. proteins involved in nutrient metabolism or carbohydrate metabolism; Zhang et al., 2019). We hypothesize that the change in community function is related to changes in metabolism during the extreme conditions generated by the bloom.

4.3 Phytoplankton bloom is enhanced by copepods

The composition (i.e. OTUs) of the microsized community was impacted by top-down control but not the functions of the microsized community. We had expected microsized ciliates to increase with phytoplankton concentration as we selected three phytoplankton species

(*Tetraselmis chui*, *Isochrysis galbana*, and *Phaeodactylum tricornutum*) considered to be 'good' food resources for ciliates (Christaki and Van Wambeke, 1995; McManus et al., 2012; Schoener and McManus, 2012; Stoecker et al., 1988; Verity and Villareal, 1986). Instead, we only observed top-down control of ciliates by copepods, as ciliates represent a superior food resource for copepods and may be selectively grazed (Calbet and Saiz, 2005). Our data are a direct illustration that copepod grazing on microheterotrophs (i.e. ciliates) reduced top-down controls, allowing phytoplankton to increase (Behrenfeld, 2010; Irigoien et al., 2005; Kuhn et al., 2015). While this has been hypothesized from microscopic observations of *in situ* samples (Irigoien et al., 2005; Leising et al., 2005), here we showed, in our closed system, that ciliates were not able to control the amplitude of a phytoplankton bloom because of the predation pressure from zooplankton (copepods).

Copepods did have a significant impact on microsized phytoplankton community function as photosynthetic activities (e.g. chlorophyll a/b binding proteins) increased with increasing copepod abundance (Figure 5; Supplementary Figure S6). Copepods have been reported to release dissolved organic carbon and inorganic nutrients through sloppy feeding and fecal pellets (Saba et al., 2011), which may increase phytoplankton production. Together these factors (increase in nutrients and release of predation pressure by feeding on microciliates) may explain the functional changes in microsized community members observed across our experiments.

5 Conclusions

This work combines experimental microcosms and 'omics approaches (both amplicon and metatranscriptomics) to reveal: (1) phytoplankton blooms strongly decrease gene expression within the nanosized community; and (2) copepods control microsized heterotrophs when phytoplankton abundance is low, reducing predation pressure on phytoplankton. We acknowledge a major caveat in interpreting function within these communities: of the 1.5 million transcripts, we were able to assign taxonomy and function to less than 25% because of the lack of knowledge about eukaryotic microorganisms. Hence, we anticipate that future studies will be able to further refine insights on functional responses through analyses of specific genes. Nevertheless, the strong decrease of transcript diversity showed the impact of phytoplankton blooms and zooplankton grazing at the species level, while the overall community functions (i.e. number of GFs, KO, and GO) remained unchanged. This illustrates the resilience of the community in maintaining ecosystem functions. Finally, these analyses show the power of combining amplicon and metatranscriptomics approaches to better understand processes driving microeukaryotic diversity and function in marine systems.

6 Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author contributions

Conceived and designed the experiments: JDG LAK. Performed the experiments: JDG. Analyzed the data: JDG LAK. Wrote the paper: JDG LAK.

8 Acknowledgements

We thank the University of Rhode Island Genomics and Sequencing and the IGS Genome Resource Center at the University of Maryland for MiSeq and HiSeq4000 sequencing; the National Marine Fisheries Service Lab (NOAA) in Milford CT, for algal cultures, and the Dam lab at UCONN for copepods. We also thank George B. McManus (University of Connecticut) for the help with the experimental setup. This work was supported by the National Science Foundation [OCE-1436003 and DEB-1541511 to LAK], Blakeslee funds at Smith College.

9 References

Alexander, H., Jenkins, B. D., Rynearson, T. A., and Dyhrman, S. T. (2015). Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *Proc. Natl. Acad. Sci. U. S. A.* 112, E2182–E2190. doi:10.1073/pnas.1421993112.

Archer, S. D., Verity, P. G., and Stefels, J. (2000). Impact of microzooplankton on the progression and fate of the spring bloom in fjords of northern Norway. *Aquat. Microb. Ecol.* 22, 27–42.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyerreil, L. A., and Thingstad, F. (1983). The Ecological Role of Water-Column Microbes in the Sea. *Mar. Ecol. Prog. Ser.* 10, 257–263. doi:Doi 10.3354/Meps010257.

Bailly, J., Fraissinet-Tachet, L., Verner, M.-C. C., Debaud, J.-C. C., Lemaire, M., Wésolowski-Louvel, M., et al. (2007). Soil eukaryotic functional diversity, a metatranscriptomic approach. *Isme J.* 1, 632–642. doi:10.1038/ismej.2007.68.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477.

Behrenfeld, M. J. (2010). Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. *Ecology* 91, 977–989. doi:10.1890/09-1207.1.

Berk, S. G., Brownlee, D. C., Heinle, D. R., Kling, H. J., and Colwell, R. R. (1977). Ciliates as a food source for marine planktonic copepods. *Microb. Ecol.* 4, 27–40. doi:10.1007/bf02010427.

Burns, J. A., Pittis, A. A., and Kim, E. (2018). Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic. *Nat. Ecol. Evol.* 2, 697.

Bushnell, B. (2016). BBMap short read aligner. Available at: http://sourceforge.net/projects/bbmap.

Calbet, A. (2008). The trophic roles of microzooplankton in marine systems. *Ices J. Mar. Sci.* 65, 325–331. doi:10.1093/icesjms/fsn013.

Calbet, A., and Saiz, E. (2005). The ciliate-copepod link in marine ecosystems. *Aquat. Microb. Ecol.* 38, 157–167.

Capriulo, G. M. (1982). Feeding of field collected tintinnid micro-zooplankton on natural food. *Mar. Bio.* 71, 73–86.

Cerón-Romero, M. A., Maurer-Alcalá, X. X., Grattepanche, J.-D., Yan, Y., Fonseca, M. M., and Katz, L. A. (2019). PhyloToL: A taxon/gene-rich phylogenomic pipeline to explore genome evolution of diverse eukaryotes. *Mol. Biol. Evol.*, msz103. Available at: http://doi.org/10.1093/molbev/msz103.

Cerón-Romero, M. A., Nwaka, E., Owoade, Z., and Katz, L. A. (2018). PhyloChromoMap, a tool for mapping phylogenomic history along chromosomes, reveals the dynamic nature of karyotype evolution in Plasmodium falciparum. *Genome Biol Evol* 10, 553–561. doi:10.1093/gbe/evy017.

599 Chen, F., Mackey, A. J., Stoeckert, C. J., and Roos, D. S. (2006). OrthoMCL-DB: querying a
600 comprehensive multi-species collection of ortholog groups. *NAR* 34, D363–D368.
601 doi:10.1093/nar/gkj123.

602 Christaki, U., and Van Wambeke, F. (1995). Simulated phytoplankton bloom input in top-down
603 manipulated microcosms: comparative effect of zooflagellates, ciliates and copepods.
604 *Aquat. Microb. Ecol.* 9, 137–147.

605 Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005).
606 Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics
607 research. *Bioinformatics* 21, 3674–3676. doi:Doi 10.1093/Bioinformatics/Bti610.

608 Cooper, E. D., Bentlage, B., Gibbons, T. R., Bachvaroff, T. R., and Delwiche, C. F. (2014).
609 Metatranscriptome profiling of a harmful algal bloom. *Harmful Algae* 37, 75–83.

610 Damon, C., Lehembre, F., Oger-Desfeux, C., Luis, P., Ranger, J., Fraissinet-Tachet, L., et al.
611 (2012). Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest
612 soils. *PLoS One* 7, e28967.

613 de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., et al. (2015). Eukaryotic
614 plankton diversity in the sunlit ocean. *Science (80-.).* 348. doi:10.1126/science.1261605.

615 del Campo, J., Sieracki, M. E., Molestina, R., Keeling, P., Massana, R., and Ruiz-Trillo, I.
616 (2014). The others: our biased perspective of eukaryotic genomes. *Trends Ecol. Evol.* 29,
617 252–259.

618 Edgcomb, V. P. (2016). Marine protist associations and environmental impacts across trophic
619 levels in the twilight zone and below. *Curr. Opin. Microbiol.* 31, 169–175.
620 doi:https://doi.org/10.1016/j.mib.2016.04.001.

621 Fenchel, T. (1987). *Ecology of Protozoa*. Berlin: Springer-Verlag.

622 Fenchel, T. (1988). Marine plankton food chains. *Annu. Rev. Ecol. Syst. Biodivers.* 19, 19–38.

623 Franzosa, E. A., Morgan, X. C., Segata, N., Waldron, L., Reyes, J., Earl, A. M., et al. (2014).
624 Relating the metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci.*
625 *U. S. A.* 111, E2329–E2338. doi:10.1073/pnas.1319284111.

626 García-Martín, E. E., Daniels, C. J., Davidson, K., Lozano, J., Mayers, K. M. J., McNeill, S., et
627 al. (2017). Plankton community respiration and bacterial metabolism in a North Atlantic
628 Shelf Sea during spring bloom development (April 2015). *Prog. Oceanogr.*

629 Gifford, S. M., Sharma, S., Rinta-Kanto, J. M., and Moran, M. A. (2011). Quantitative analysis
630 of a deeply sequenced marine microbial metatranscriptome. *ISME J.* 5, 461.

631 Grattepanche, J.-D., Breton, E., Brylinski, J.-M., Lecuyer, E., and Christaki, U. (2011a).
632 Succession of primary producers and micrograzers in a coastal ecosystem dominated by
633 Phaeocystis globosa blooms. *J. Plankton Res.* 33. doi:10.1093/plankt/fbq097.

634 Grattepanche, J.-D., Juarez, D. L., Wood, C. C., McManus, G. B., and Katz, L. A. (2019).
635 Incubation and grazing effects on spirotrich ciliate diversity inferred from molecular
636 analyses of microcosm experiments. *PLoS One* 14, e0215872. Available at:
637 https://doi.org/10.1371/journal.pone.0215872.

638 Grattepanche, J.-D., McManus, G. B., and Katz, L. A. (2016a). Patchiness of ciliate communities

sampled at varying spatial scales along the New England shelf. *PLoS One* 11. doi:10.1371/journal.pone.0167659.

Grattepanche, J.-D., Santoferrara, L. F., Andrade, J., Oliverio, A. M., McManus, G. B., and Katz, L. A. (2014). Distribution and diversity of Oligotrich and choreotrich ciliates assessed by morphology and DGGE in temperate coastal waters. *Aquat. Microb. Ecol.* 71. doi:10.3354/ame01675.

Grattepanche, J.-D., Santoferrara, L. F., McManus, G. B., and Katz, L. A. (2016b). Unexpected biodiversity of ciliates in marine samples from below the photic zone. *Mol. Ecol.* 25. doi:10.1111/mec.13745.

Grattepanche, J.-D., Vincent, D., Breton, E., and Christaki, U. (2011b). Microzooplankton herbivory during the diatom-Phaeocystis spring succession in the eastern English Channel. *J. Exp. Mar. Bio. Ecol.* 404, 87–97. doi:10.1016/j.jembe.2011.04.004.

Grattepanche, J.-D., Walker, L. M., Ott, B. M., Paim Pinto, D. L., Delwiche, C. F., Lane, C. E., et al. (2018). Microbial Diversity in the Eukaryotic SAR Clade: Illuminating the Darkness Between Morphology and Molecular Data. *BioEssays* 40. doi:10.1002/bies.201700198.

Hamady, M., Lozupone, C., and Knight, R. (2009). Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J.* 4, 17–27. doi:10.1038/ismej.2009.97.

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2018). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.*

Irigoien, X., Flynn, K. J., and Harris, R. P. (2005). Phytoplankton blooms: a 'loophole'in microzooplankton grazing impact? *J. Plankton Res.* 27, 313–321.

Irigoien, X., Huisman, J., and Harris, R. P. (2004). Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature* 429, 863–867. doi:Doi 10.1038/Nature02593.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2015). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462.

Kuhn, A. M., Fennel, K., and Mattern, J. P. (2015). Model investigations of the North Atlantic spring bloom initiation. *Prog. Oceanogr.* 138, 176–193. doi:https://doi.org/10.1016/j.pocean.2015.07.004.

Lance, G. N., and Williams, W. T. (1967). Mixed-Data Classificatory Programs I - Agglomerative Systems. *Aust. Comput. J.* 1, 15–20.

Laws, E. A., Bienfang, P. K., Ziemann, D. A., and Conquest, L. D. (1988). Phytoplankton population dynamics and the fate of production during the spring bloom in Auke Bay, Alaska 1. *Limnol. Oceanogr.* 33, 57–67.

Leising, A. W., Pierson, J. J., Halsband-Lenk, C., Horner, R., and Postel, J. (2005). Copepod grazing during spring blooms: can Pseudocalanus newmani induce trophic cascades? *Prog. Oceanogr.* 67, 406–421.

Lesniewski, R. A., Jain, S., Anantharaman, K., Schloss, P. D., and Dick, G. J. (2012). The metatranscriptome of a deep-sea hydrothermal plume is dominated by water column

methanotrophs and lithotrophs. *ISME J.* 6, 2257.

Li, L., Stoeckert Jr., C. J., and Roos, D. S. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Res. 10.1101/gr.1224503* 13, 2178–2189. Available at: http://www.genome.org/cgi/content/abstract/13/9/2178.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.

Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., and Knight, R. (2011). UniFrac: an effective distance metric for microbial community comparison. *ISME J.* 5, 169.

Mahe, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2015). Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3, e1420. doi:10.7717/peerj.1420.

Martínez, R. A., Calbet, A., and Saiz, E. (2017). Effects of small-scale turbulence on growth and grazing of marine microzooplankton. *Aquat. Sci.* 80, 2. doi:10.1007/s00027-017-0558-8.

Mason, O. U., Hazen, T. C., Borglin, S., Chain, P. S. G., Dubinsky, E. A., Fortney, J. L., et al. (2012). Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. *ISME J.* 6, 1715. doi:10.1038/ismej.2012.59.

Maurer-Alcala, X. X., Yan, Y., Pilling, O., Knight, R., and Katz, L. A. (2018). "Twisted Tales: Insights into Genome Diversity of Ciliates Using Single-Cell 'omics." *Genome Biol Evol*. doi:10.1093/gbe/evy133.

McKie-Krisberg, Z. M., Sanders, R. W., and Gast, R. J. (2018). Evaluation of Mixotrophy-Associated Gene Expression in Two Species of Polar Marine Algae. *Front. Mar. Sci.* 5, 1–12. doi:10.3389/fmars.2018.00273.

McManus, G. B., Schoener, D., and Haberlandt, K. (2012). Chloroplast symbiosis in a marine ciliate: ecophysiology and the risks and rewards of hosting foreign organelles. *Front. Microbiol.* 3, 321.

McManus, G. B., Zhang, H., and Lin, S. J. (2004). Marine planktonic ciliates that prey on macroalgae and enslave their chloroplasts. *Limnol. Oceanogr.* 49, 308–313.

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8, e61217.

Menden-Deuer, S., and Kiørboe, T. (2016). Small bugs with a big impact: linking plankton ecology with ecosystem processes. *J. Plankton Res.* 38, 1036–1043.

Monchy, S., Grattepanche, J.-D., Breton, E., Meloni, D., Sanciu, G., Chabé, M., et al. (2012). Microplanktonic community structure in a coastal system relative to a Phaeocystis bloom inferred from morphological and tag pyrosequencing methods. *PLoS One* 7. doi:10.1371/journal.pone.0039924.

Mousing, E. A., Richardson, K., Bendtsen, J., Cetinić, I., and Perry, M. J. (2016). Evidence of small-scale spatial structuring of phytoplankton alpha- and beta-diversity in the open ocean. *J. Ecol.* 104, 1682–1695. doi:10.1111/1365-2745.12634.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., and O'hara, R. B. (2016). vegan: Community Ecology Package. R package version 2.3-4 [WWW document].

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., et al.

(2007). The vegan package. *Community Ecol. Packag.* 10, 631–637.

Pernice, M. C., Giner, C. R., Logares, R., Perera-Bel, J., Acinas, S. G., Duarte, C. M., et al. (2016). Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *ISME J.* 10, 945–958.

Rivers, A. R., Sharma, S., Tringe, S. G., Martin, J., Joye, S. B., and Moran, M. A. (2013). Transcriptional response of bathypelagic marine bacterioplankton to the Deepwater Horizon oil spill. *ISME J.* 7, 2315.

Rosetta, C. H., and McManus, G. B. (2003). Feeding by ciliates on two harmful algal bloom species, Prymnesium parvum and Prorocentrum minimum. *Harmful Algae* 2, 109–126. doi:Doi 10.1016/S1568-9883(03)00019-2.

Saba, G. K., Steinberg, D. K., and Bronk, D. A. (2011). The relative importance of sloppy feeding, excretion, and fecal pellet leaching in the release of dissolved carbon and nitrogen by Acartia tonsa copepods. *J. Exp. Mar. Bio. Ecol.* 404, 47–56.

Schoener, D. M., and McManus, G. B. (2012). Plastid retention, use, and replacement in a kleptoplastidic ciliate. *Aquat. Microb. Ecol.* 67, 177–187.

Schoener, D. M., and McManus, G. B. (2017). Growth, grazing, and inorganic C and N uptake in a mixotrophic and a heterotrophic ciliate. *J. Plankton Res.* 39, 379–391. doi:10.1093/plankt/fbx014.

Sieburth, J. M., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1. *Limnol. Oceanogr.* 23, 1256–1263.

Sisson, C., Gulla-Devaney, B., Katz, L. A., and Grattepanche, J.-D. (2018). Seed bank and seasonal patterns of the eukaryotic SAR (Stramenopila, Alveolata and Rhizaria) clade in a New England vernal pool. *J. Plankton Res.* 40, 376–390. doi:10.1093/plankt/fby020.

Stoecker, D. K., Silver, M. W., Michaels, A. E., and Davis, L. H. (1988). Obligate mixotrophy in Laboea strobila, a ciliate which retains chloroplasts. *Mar. Biol.* 99, 415–423.

Tartar, A., Wheeler, M. M., Zhou, X., Coy, M. R., Boucias, D. G., and Scharf, M. E. (2009). Parallel metatranscriptome analyses of host and symbiont gene expression in the gut of the termite Reticulitermes flavipes. *Biotechnol. Biofuels* 2, 25.

Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28, 33–36. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10592175.

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science (80-.).* 336, 608–611.

Tucker, S. J., McManus, G. B., Katz, L. A., and Grattepanche, J.-D. (2017). Distribution of abundant and active planktonic ciliates in coastal and slope waters offNew England. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.02178.

759 Vaulot, D., Eikrem, W., Viprey, M., and Moreau, H. (2008). The diversity of small eukaryotic
760 phytoplankton (≤ 3 μm) in marine ecosystems. *FEMS Microbiol. Rev.* 32, 795–820.

761 Verity, P. G., and Villareal, T. A. (1986). The relative food value of diatoms, dinoflagellates,
762 flagellates, and cyanobacteria for tintinnid ciliates. *Arch. für Protistenkd.* 131, 71–84.

763 Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H.-A., Voget, S., Simon, M., et al. (2014).
764 Impact of a phytoplankton bloom on the diversity of the active bacterial community in the
765 southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.*
766 87, 378–389. Available at: http://dx.doi.org/10.1111/1574-6941.12230.

767 Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., and Keeling, P.
768 J. (2015). Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of
769 microbes. *Science (80-.).* 347, 1257594. doi:10.1126/science.1257594 %J Science.

770 Yutin, N., Wolf, M. Y., Wolf, Y. I., and Koonin, E. V. (2009). The origins of phagocytosis and
771 eukaryogenesis. *Biol. Direct* 4, 9. doi:10.1186/1745-6150-4-9.

772 Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2013). PEAR: a fast and accurate Illumina
773 Paired-End reAd mergeR. *Bioinformatics* 30, 614–620.

774 Zhang, Y., Lin, X., Shi, X., Lin, L., Luo, H., Li, L., et al. (2019). Metatranscriptomic signatures
775 associated with phytoplankton regime shift from diatom dominance to a dinoflagellate
776 bloom. *Front. Microbiol.* 10, 590.

777

10 Tables

Table 1. Correlation analyses show a significant impact of phytoplankton on nanosized community, while Copepods impacted microsized plankton, and the Bacillariophyceae (significant correlation are in bold: ***** $p < 0.05$; ****** $p < 0.01$):

	phytoplankton		**Copepods**	
	nanosized	microsized	nanosized	microsized
SAR	**0.65***	0.37	0.3	**0.37***
Alveolata	0.46	0.29	0.36	0.33
Ciliophora	**0.62***	0.25	0.19	0.16
Stramenopila	**0.60***	0.37	0.21	0.27
Bacillariophyceae	0.15	0.17	**0.54****	**0.70****
Dictyochophyceae	**0.71****	0.45	0.16	0.27

11 Figure legends

Figure 1 Comparison of community composition based on amplicon analyses (a, b) across treatments within our microcosms. a) reads and abundant OTUs (>10 reads) distributions of SAR (Stramenopila, Alveolata, and Rhizaria) lineages within each treatment show a dominance in different stramenopiles lineages by size fraction (Dictyochophyceae in the nanosize and Bacillaryophyceae in the microsize) and ciliates in both sizes and . b) and increase of spirotrich ciliates with phytoplankton bloom. For all panels, the letter 'p' in the sample label represent the abundance of phytoplankton added: none for p0 [white] , 5.10^2 cell. mL^{-1} for p1 [yellow-green]; 5.10^3 cell. mL^{-1} for p2 [light green], and 5.10^4 cell. mL^{-1} in final concentration for p3 [green]. The letter 'z' in the label represent the number of copepods added: none for z0 [white], 5 per liter for z1 [light brown] and 10 per liter for z2 [dark brown]. For example, p0z0 has no copepods and no phytoplankton added and represents our control; p3z0 was incubated without copepods but in phytoplankton bloom condition.

Figure 2 Proportion of OTUs, transcripts, and conserved gene families (GFs) significantly increasing (up) or decreasing (down) in response to the incubation conditions show a large proportion of transcripts and GFs downregulated (up to 70% of the transcripts and GFs) in the nanoplankton with phytoplankton bloom treatment. OTU, transcripts and GFs were identified with DESeq2 (Love et al., 2014). The number represent the total number of OTUs, transcripts, and GFs in the samples.

Figure 3 Principal coordinates analysis by size fractions of the impact of phytoplankton and copepods abundance on community composition (i.e. OTUs) and on function (i.e. transcripts) show the strongest impact of phytoplankton added on nanosized community. An exception is the microsize community composition, which shows an absence of clear response to the abundance of the phytoplankton we added. UniFrac dissimilarity index was used for the OTUs and the Canberra dissimilarity index for the metatranscriptome data.

Figure 4 Heatmap of differential gene expression for each of the 81,634 transcripts matching gene families (rows) across treatments (columns) shows highest downregulation in nanosized communities in phytoplankton bloom conditions. The clustering of the samples is based on Euclidean distance and WardD algorithm. The bar graph at the bottom show the number of transcripts up or down-regulated (absolute difference to the control higher than log2). See Figure 1 for details about sample labels.

Figure 5 Heatmap of differential gene expression of the transcripts grouped by Clusters of Orthologous Groups (COG) indicates downregulation of transcript expression related to translation (GOG J) in the microsized plankton incubated with copepods, while nanosized plankton show a downregulation of cytoskeleton (COG Z) function incubated with phytoplankton bloom. The microsized plankton show an upregulation of protein with unknown function (COG S, almost all involved in fucoxanthin chlorophyll a/c pathway in our data). The clustering of the samples is based on Euclidean distance and WardD algorithm. See Figure 1 for details about sample labels.