Effects of monsoons and storms on the structuring and diversity of picoeukaryotic microbial communities in a tropical coastal environment

Maria Anna Michaela de la Cruz, Brian William Hingpit, Laure Guillou, Deo Florence L. Onda

Effects of monsoons and storms on the structuring and diversity of picoeukaryotic microbial communities in a tropical coastal environment

Maria Anna Michaela De La Cruz[a], Brian William Hingpit[a], Laure Guillou[b], Deo Florence L. Onda[a,*]

[a] Microbial Oceanography Laboratory, The Marine Science Institute, Velasquez St., University of the Philippines, Diliman, 1101, Quezon City, Philippines

[b] Sorbonne Université, CNRS, UMR7144 Adaptation et Diversité en Milieu Marin, Ecology of Marine Plankton (ECOMAP),Station Biologique de Roscoff, 29680, Roscoff, France

*Corresponding author: Deo Florence L. Onda (dfonda@msi.upd.edu.ph)

Abstract

Picoeukaryotes are key components in marine ecosystems that play crucial roles in food webs and biogeochemical cycles. Despite their significance, many aspects of their community ecology and diversity remain understudied. Here, we investigated the taxonomic and functional diversity of picoeukaryotic communities in response to monsoonal patterns and weather disturbances brought about by storms, characterizing tropical coastal regions. To do this, water samples were collected almost weekly or bi-weekly at a single location in a tropical coastal environment covering the late northeast (NE) and southwest (SW) monsoons. We then performed high-throughput amplicon sequencing of the V4 region of the 18S rRNA gene to generate taxonomic profiles of the communities across time. Clustering based on environmental parameters grouped our samples into months associated with NE monsoon, SW monsoon, and stormy SW monsoon, demonstrating seasonality influenced by monsoons and storms, typically observed in tropical coastal waters. In comparison, clustering based on abundance only grouped the samples into NE and SW monsoon, with most communities during storm period joining the NE monsoon samples. These samples exhibited greater diversity, with smaller taxa such as Syndiniales, Prymnesiophyceae, Picozoa, Cercozoa, Stramenopiles, and Chlorophytes being the most abundant groups present. In contrast, SW monsoon samples have lower diversity but have become generally dominated by large-celled taxa, mostly diatoms. Multivariate and correlation analyses both revealed nitrate as the strongest environmental driver of the picoeukaryotic community structuring. Meanwhile, network analysis grouped the taxa into three modules, more consistent with the clustering based on environmental parameters, implying that although storms may not significantly change the community composition, they may however influence the dominating taxa. Each module was composed of a unique set of co-occurring taxa, highlighting high turnover of picoeukaryotic communities between each season. In addition, our results showed that SW monsoon-associated module had higher interconnectivity than other modules, suggesting that the interactions during this period may be less species-specific, thus, more adaptable than during NE monsoon. However, we observed that extreme fluctuations caused by storms could have possibly allowed for selection of dominant taxa. Shotgun metagenomic sequencing of representative samples from each monsoon period also revealed that differently abundant functional genes, particularly genes associated to nitrogen metabolism, might have also helped in adaptation to the changing

nutrient conditions. Our observations provide new insights on the potential trajectory of microbial communities under environmental stresses, which are important in understanding the implications of emerging threats such as coastal eutrophication and climate change.

Keywords: picoeukaryotic communities, seasonal monsoons, tropical coastal ecosystems, 18S rRNA gene, shotgun metagenomics, co-occurrence network

1. INTRODUCTION

Marine picoeukaryotes are unicellular eukaryotes belonging to the size fraction of less than 3.0 µm (Worden & Not, 2008). They are ubiquitous in marine ecosystems and play crucial roles in the biogeochemical cycles (Massana et al., 2004; Vaulot et al., 2008) with their functional role in the ecosystem being defined by their trophic modes (Worden & Not, 2008). Traditionally, picoeukaryotes can be classified as photoautotrophs, which are considered part of the phytoplankton that can perform photosynthesis, or as heterotrophs, which play a key role in remineralizing organic matter in the ocean (Massana et al., 2004; Not et al., 2009; Xie et al., 2021). Some picoeukaryotes may also have both phototrophic and heterotrophic capabilities and are considered as mixotrophs (Sanders & Gast, 2012). However, studying picoeukaryotes have been challenging owing to their very small size and lack of morphological landmarks (Massana et al., 2004), resulting in the limited understanding of their ecology and diversity, despite contributing significantly to overall biomass and carbon standing stock (Worden et al., 2004). The advent of molecular methods (e.g., cloning, high throughput sequencing) have consistently confirmed the high diversity of picoeukaryotes, with majority of the recovered operational taxonomic units (OTUs) belonging to novel lineages, highlighting the need for more in-depth studies on their taxonomic and even functional diversity (Massana et al., 2004; Romari & Vaulot, 2004).

Previous works demonstrated that picoeukaryotes also exhibit community patterns and structuring akin to their larger counterparts, where changes in the environmental conditions can significantly influence their community structure (Lambert et al., 2019; Logares et al., 2020). Specifically, a few studies demonstrated that environmental conditions linked to seasonal variability could potentially be more significant driver than differences caused by spatial variability. This is particularly true for regions with pronounced seasonal shifts such as in the temperate and polar waters. In a temperate upwelling region, for example, multivariate analysis grouped the samples according to season, but not by location (Rocke et al., 2020). Similarly,

in a temperate coastal basin, the community structure of pico- and nanoeukaryotes significantly differed between summer and winter, but not between sampling sites (Wang et al., 2020). However, most of these studies were done in temperate or polar waters or in freshwater bodies, which cannot represent the same conditions as in marine tropical waters that experience less pronounced seasonality. Initial work on marine picoeukaryotes in the tropics revealed that temperature may potentially have the most significant influence on their community structuring (dela Peña et al., 2021; Liu et al., 2021; Logares et al., 2020). For example, a recent study in the Western Pacific Ocean, encompassing the regions influenced by the North Equatorial Current, Kuroshio Current, and Mindanao Current, showed that environmental variation, particularly changes in temperature, dissolved oxygen (DO), phosphate, and silicate concentrations, was the main driver of picoeukaryotic community structuring in this area (Liu et al., 2021). However, in a study using the data from both *Malaspina-2010* and *Tara Oceans* expeditions covering large or 'planetary' geographic scale, it was reported that dispersal limitation was more influential in shaping picoeukaryotic communities in tropical and subtropical surface waters than heterogeneous (environmental) selection (Logares et al., 2020). In addition, other than environmental factors, co-occurrence analyses derived from the *Tara Oceans* data consistently reported that biotic interactions and positive associations dominate over environmental factors in driving the structuring of planktonic community (Lima-Mendez et al., 2015; Sunagawa et al., 2020; Vincent & Bowler, 2020). Open oceans however exhibit less heterogeneity over large spatial and time scales compared to coastal environments, which are more vulnerable to short-term changes (Lambert et al., 2019).

In the tropics, the high seasonal variability along the coast is primarily brought about by the periodic reversal of the winds during southwest (SW) and northeast (NE) monsoon (Rivera, 1997; Wyrtki, 1961). The NE monsoon, also known as the 'dry season', is typically observed from November to March, and characterized by strong northeasterly winds and relatively less precipitation (Matsumoto et al., 2020). Meanwhile, the SW monsoon or 'wet season', from June to September, is characterized by heavier precipitation and strong, southwesterly winds (Matsumoto et al., 2020). During this season, tropical cyclones are more frequently experienced, which may cause disturbances in water conditions (Cayanan et al., 2011; Rivera, 1997). In particular, external pollution loads including silt, nutrients, and organic matter brought about by runoffs during the typhoons from terrestrial sources can have significant impacts on environmental conditions, and thus, also influence the microbial communities (Lambert et al., 2019; Rivera, 1997). Previous studies investigating the phytoplankton

dynamics in such coastal setting have provided insights on how the community changes on a seasonal basis and suggested possible environmental drivers during each season (Azanza & Benico, 2013; Canini et al., 2013; dela Peña et al., 2021; Escobar et al., 2013; Yap et al., 2004). These studies reported that phytoplankton community composition varied with changing monsoons and inferred that the changes were associated with conditions associated with monsoon periods. However, the same studies focused only on the microphytoplankton communities and were limited using traditional methods such as microscopy, except for De La Peña et al. (2021), where they observed the previously unexplored vast diversity of picoeukaryotes in Philippine waters using high throughput amplicon sequencing. Recent studies have also consistently reported the presence of picoeukaryotic groups in tropical coasts (Amin et al., 2021; dela Peña et al., 2021; Mitbavkar & D'souza, 2021; Qiu et al., 2019; Rajaneesh et al., 2018). In particular, Amin et al. (2021) reported that the distribution of picoeukaryotes progressively decreased from the coast of Peninsular Malaysia towards the open ocean during SW monsoon. These reports highlighted the contribution of picoeukaryotes to the productivity and biogeochemical cycles in a tropical coastal habitat. Understanding how picoeukaryotic communities change with monsoonal shifts will provide a more accurate picture of the overall productivity and functioning of the tropical coastal system. Further, as the base of the food web, changes in picoeukaryotic communities may also cascade to higher trophic levels. Thus, there is a need to understand how picoeukaryotes are influenced by conditions brought about by monsoons and storms, especially in tropical coasts that are more vulnerable to effects of intensifying storms and other impacts of climate change (Kurtay et al., 2021).

Here, our goal was to provide a more in-depth understanding of how picoeukaryotic communities respond to environmental changes brought about by monsoonal shifts in coastal environments, including disturbances caused by tropical storms, by looking at the changes in their taxonomic and functional diversity. We hypothesize that shifts in monsoons will result in significant changes in the physical and chemical regimes of the water column in coastal habitats, also greatly affecting the proliferating microbial communities. To do this, we collected samples at a single location on an almost weekly or bi-weekly basis from January to October (16 samples in total), capturing the communities from late NE to SW monsoon. We then performed high-throughput amplicon sequencing of the V4 region of the 18S rRNA gene of the entire microbial eukaryotic community to obtain their taxonomic profiles across time. The change in environmental conditions along with the monsoon is expected to be accompanied by the dominance of taxa with specific adaptive mechanisms to the new conditions. This would

also be reflected in the abundant functional genes that may be differentially expressed between monsoonal periods. For this, we performed shotgun metagenomic sequencing on select samples representing each monsoonal period and investigated genes that may indicate adaptation to prevailing environmental conditions. Finally, we also performed co-occurrence network analysis, which extends community analysis beyond the diversity indices by visualizing interactions between taxa including those previously uncharacterized (Vincent & Bowler, 2020). This allowed gaining more insights on the underlying processes that structure the community and predict possible changes of food web systems in response to monsoonal shifts and storms.

2. MATERIALS AND METHODS

2.1. Sample collection and processing

To monitor the changes in the community across monsoon periods, a single sampling site (16°22'56.7"N 119°54'46.4"E) was maintained at a tropical marine coastal area in front of the Bolinao Marine Laboratory located in Bolinao, Pangasinan at the northwestern side of the Philippines (Supplementary Figure S1). It is a relatively well-studied site where majority of phytoplankton work has been done, including spatial and temporal investigations, evolving into becoming a model site for the Philippines (Azanza et al., 2004; Yap et al., 2004; Azanza & Benico, 2013; Escobar et al., 2013; Albelda et al., 2019; dela Peña et al., 2021). Surface seawater samples (ca. 1 m below) were collected at an approximately weekly to biweekly basis from January to October 2019 to investigate temporal variability of the entire picoeukaryotic microbial community.

Water samples were collected using a Niskin-type bottle sampler (General Oceanics, Inc.). A total of 2 L of water was collected each time and pre-filtered immediately through a 120 µm sieve to remove zooplankton and other large debris. Samples were collected in sterilized Nalgene bottles (acid and sterile water washed) and immersed in ambient seawater in an icebox during transport. Separate water samples for Chlorophyll *a* (Chl *a*) analysis were also collected from the same Niskin bottle and kept in an icebox until they reached the laboratory. Environmental parameters including temperature, salinity, and dissolved oxygen (DO) were measured every sampling using a multiparameter (Hanna HI9829 and YSI Pro 2030).

In the laboratory, the pre-filtered seawater samples for DNA extraction were serially filtered through a 42 mm polycarbonate (PC) 3 µm-pore size filter (Whatman), which represents the large micro-nanoeukaryotes, and through a 0.2 µm Sterivex filter (Millipore) or a 42 mm 0.2 µm-pore size filter (Whatman and Membrane Solutions) corresponding to the picoeukaryotes. A total of 5x10 mL of the serially filtered seawater was also collected and stored in separate conical tubes used for nutrient analyses including nitrite, nitrate, phosphate, and silicate using spectrophotometry following the methods based on Parsons et al. (1984). The filters were transferred to 2 mL sterile microcentrifuge tubes and DNA was preserved by adding at least 1.6 mL RNALater (Ambio). Samples were stored at -20°C until extraction. For Chl *a* analysis, 500 mL of seawater samples were filtered through Whatman GF/F filters and extracted with 90% acetone and subsequently analyzed using a Trilogy laboratory fluorometer following the method of Parsons et al. (1984).

2.2. DNA extraction and high throughput sequencing

Total genomic DNA was only extracted from the 0.2 µm-pore size filters to focus on the picoeukaryotic communities using the DNeasy PowerSoil Kit (Qiagen, Germany) following the manufacturer's protocol with some modifications. Briefly, the filters were cut into smaller pieces before being transferred to PowerBead tubes provided by the kit. An additional 0.5 mL of Buffer AL lysis buffer (Qiagen, Germany) was pipetted inside the tube to ensure complete immersion of the filters to enhance cell lysis. Subsequently, 60 µL of Solution C1 was added, then vortexed at maximum speed for 10 minutes. This was followed by the procedures prescribed by the kit's manufacturer. DNA quality and yield were measured using spectrophotometry (Nanodrop, Thermo Fisher Scientific). Integrity of the extracted DNA was also checked by agarose gel electrophoresis. In addition, PCR using universal eukaryotic primer pair 4616F and 4618R (Logares et al., 2007) targeting near full-length 18S rRNA gene was performed before sending the samples for sequencing.

Here, two sequencing approaches were used, namely 18S rRNA gene amplicon sequencing to probe on community shifts with monsoonal change and their co-occurrences (January to October), and shotgun metagenome sequencing to further look at potential changes in dominant functional genes with monsoons. For the targeted sequencing, all DNA extracts (n=16) that passed the quality check were sent to the Integrated Microbiome Resource (IMR) housed in Dalhousie University, Canada for Illumina multiplex sequencing using E572F/E1009R forward and reverse primers (Comeau et al., 2011) targeting the V4 region of the 18S rRNA

gene. Raw amplicon sequences are available in the Sequence Read Archive (SRA) database of NCBI under Accession No. PRJNA656691.

For shotgun metagenome sequencing, samples representing the intermonsoon (April) where community diversity is expected to be greatest (dela Peña et al., 2021) and towards the end of the southwest monsoon (September), were sent to Macrogen Genome Center (Macrogen NGS Division, Seoul, Korea). Libraries were prepared using the Nextera XT DNA Library Preparation Kit. Sequencing was performed on the Illumina NovaSeq 6000 platform with 150 bp paired-end setting. Raw shotgun sequencing data are also available in the SRA database of NCBI with Accession No. PRJNA788625.

2.3. Post-sequencing processing and analysis

Bioinformatics processing and quality filtering of the resulting 18S rRNA reads were done in Quantitative Insights in Microbial Ecology v.2 (QIIME 2; Bolyen et al., 2019). Raw reads were first imported as a QIIME 2 artifact using the qiime tools 'import' method. After quality-checking, the reads were assembled with a minimum overlap of 12 bp using the q2-vsearch plugin with the 'join-pairs' method (Rognes et al., 2016). Low-quality sequences were removed, and the resulting data set was subsequently processed for Operational Taxonomic Unit (OTU) clustering using the q2-vsearch plugin with the 'cluster-features-open-reference' method. Under this method, OTUs were clustered against an inhouse curated reference database (Philippine Reference 18S rRNA Gene Database, *unpublished*), which uses a combination of the SILVA v138.1, Nordicana Reference database v2.0 (Lovejoy et al., 2016), and additional sequences amplified from Philippine isolates curated from available GenBank-deposited sequences at 98% similarity, and any sequences that had no match were then clustered *de novo* (Rideout et al., 2014). Singletons and chimeras were removed using the q2-vsearch plugin with the 'uchime-denovo' method (Rognes et al., 2016). Assignment of taxonomic identity was performed using the sci-kit learn plugin against the inhouse curated reference database. Fungi, and metazoa-related reads were excluded from the analysis. The final OTU dataset used for community analysis was then rarefied at 1,559 reads per sample using single rarefaction based on the sample (Jun15BML) with the lowest number of reads to make the diversity indices comparable.

For the shotgun sequencing data, both samples were first subjected to Kaiju (web server version) taxonomic classification (Menzel et al., 2016) using the maximum exact matches

(MEM) mode with a minimum match length of 11. Additionally, the nr+euk was used as the reference database and the SEG low-complexity region filtering was enabled. Afterwards, reads that were unclassified and were inferred as belonging to the Eukaryota superkingdom (NCBI taxonomy ID: 2759) were extracted from both samples. The extracted reads were then quality checked using FastQC v0.11.9 (Andrews, 2010) and trimmed and filtered using Trimmomatic v0.39 (Bolger et al., 2014) (Parameters: SLIDINGWINDOW:5:18, MINLEN:75, LEADING:3, TRAILING:3, AVGQUAL:20). Subsequently, the remaining read pairs from each of the samples were merged using PEAR v0.9.11 (Zhang et al., 2014) and the merged and unmerged reads were concatenated together into one file. To functionally annotate the shotgun data, the SUPER-FOCUS pipeline v0.35 (Silva et al., 2016) was utilized with DIAMOND (Buchfink et al., 2021) as the aligner and DB_100 as the reference database (Parameters: e-value=$1e^{-5}$, minimum percent identity=60%, minimum alignment length=15 amino acids, normalization=0, FOCUS=0). Finally, the generated functional abundance table of SUPER-FOCUS was imported into Statistical Analysis of Metagenomic Profiles (STAMP v. 2.1.3) (Parks et al., 2014) for differential abundance analysis. Statistical significance was calculated using Fisher's exact test with Benjamini-Hochberg false discovery rate method of multiple test correction. Additionally, the Newcombe-Wilson method was applied to calculate for the confidence intervals and features with a *q*-value of less than 0.05 were considered significant. Lastly, a combination of basic local alignment search tool (BLAST) and phylogenetic analysis of the four most abundant photosynthetic genes was performed to validate if they belong to the dominant species identified by the amplicon sequencing. Briefly, quality filtered reads from the two samples were assembled individually using MEGAHIT (Li et al., 2015). Afterwards, all *PsaA, PsbA, PsbB,* and *PsbC* amino acid sequences were retrieved from the manually curated subset of UniProt (UniProt Consortium, 2019), the SWISS-PROT database (Boeckmann et al., 2003). The assembled metagenomes were then queried against the retrieved sequences using MetaEuk v5.34c21f2 (Levy et al., 2020) to predict the protein sequences in the contigs respective to the reference sequences used. Next, the predicted sequences and the reference sequences were concatenated together and aligned using MAFFT v7.475 (Katoh et al., 2013) applying the local pair method with a maximum iteration of 1000. Finally, the phylogenetic tree was constructed from the aligned sequences using FastTree v2.1.10 (Price et al., 2010) with default parameters.

2.4. Statistical and ecological analyses

Alpha diversity indices such as Observed OTUs, Chao1, and Shannon were estimated from the rarefied OTU table using the 'alpha' function in the package 'microbiome' in R (Lahti et al., 2019). Phylogenetic diversity (Faith's PD) was also calculated using the 'pd' function in the R package 'picante' (Kembel et al., 2010). Boxplots showing the alpha diversity indices were produced using the 'ggboxplot' function in the R package 'ggpubr' (Kassambara, 2020). Beta diversity was also assessed by calculating for the weighted UniFrac distances using the 'GUniFrac' R package (Chen et al., 2022). Non-metric multidimensional scaling (NMDS) ordination plot was then performed based on the weighted UniFrac distances using the R package 'vegan' (Oksanen et al., 2020). Correlation between environmental variables and community composition was explored using canonical correspondence analysis (CCA) also using 'vegan'. In addition to CCA, possible drivers of clustering were also investigated using linear regression, with principal coordinates as x-axis and measured environmental parameters as y-axis. Most of the statistical tests in this study were performed in the R environment v.3.6.2 (R Core Team, 2019).

2.5. Weighted co-occurrence network analysis

To further investigate the possible co-occurrence vis-à-vis interactions between taxa, network analysis was performed using the R package 'Weighted Gene Co-expression Network Analysis (WGCNA)' (Langfelder & Horvath, 2008). An OTU matrix without the low abundance OTUs (<10 counts) across all time points or 0 values for at least 11 time points to yield more robust seasonal patterns (Bryant et al., 2016) was used for WGCNA. The resulting matrix was checked for any missing values using the 'GoodSamplesGenes' function included in the package. Subsequently, the OTU matrix was used to identify the modules and construct the network manually following the method of Zhang & Horvath (2005). Briefly, we first chose to set the soft-thresholding power to 6 by using the 'pickSoftThreshold' function. This soft-thresholding power was used to encode the connection strength (weight) between each pair of OTU, resulting to an adjacency matrix. To reduce noise and spurious associations, the adjacency matrix was transformed into the Topological Overlap Matrix (TOM), which reflects the relative interconnectedness between OTUs. Then, average linkage hierarchical clustering using the TOM was generated as data input to obtain a clustering dendrogram of OTUs. The clusters in the resulting dendrogram correspond to the modules, which are subgroups of highly correlated OTUs with similar distribution patterns across the samples. Module identification based on the dendrogram was done using the 'cutreeDynamic' function. Lastly, the modules were used to

calculate for eigengene values to correlate with the environmental parameters and elucidate potential drivers of the module clustering. The modules were exported to Cytoscape v.3.8.2 (Csardi & Nepusz, 2006) for visualization but only included edges with weight of $\geq$ 0.5 to filter out less significant associations. Key topological features of the network, such as the degree (number of edges), shortest path distribution, and betweenness centrality, were also calculated for each node using the plugin 'Network Analyzer' in Cytoscape to assess significant associations between taxa.

3. RESULTS

3.1. Environmental context

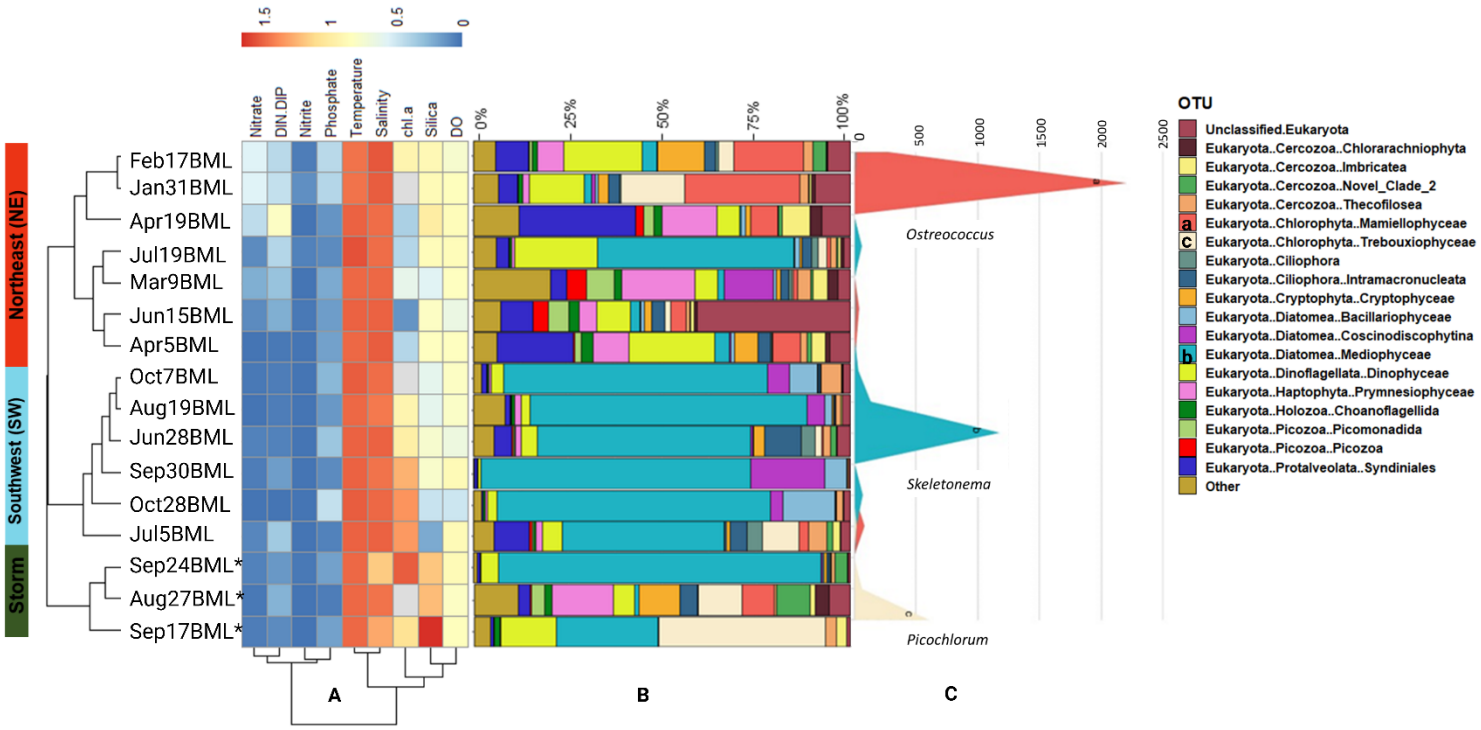


Fig. 1. A) UPGMA clustering of samples based on log values of measured environmental parameters. The samples were grouped into NE, SW and storm clusters; B) Barplots showing the relative abundance of binned OTUs at the class level present in the picoplankton fraction. An in-house curated reference database including entries from both SILVA v.138.1 and Northern Reference Database v.1.0 was used as the reference for the barplot. The OTU dataset used was rarefied at 1,559 reads; C) Shifts in the most abundant taxa (represented by

315 the OTU with the highest read count) per module driven by changes in local environmental
316 conditions influenced by monsoons and by storms.

317

318 Results showed that the surface waters were marked by strong seasonal variation, where
319 measured physico-chemical parameters clustered the samples into three main groups (Figure
320 1A). Groupings include January to mid-June, which were months associated with the late
321 northeast (NE) monsoon to intermonsoon, and late June-October associated with southwest
322 (SW) monsoon. Notably, some samples collected in August (Aug27BML) and September
323 (Sep17BML and Sep24BML) also separated into another group outside of the clusters
324 associated with the monsoon periods. These samples coincided with episodic disturbance when
325 tropical storms were reported near the study site (PAGASA, 2019) during the collection period,
326 herein referred to as the 'storm' cluster. Significant differences on the dissimilarities based on
327 environmental conditions (p<0.05) were found using ANOSIM (Supplementary Figure S3A).

328 The shift in monsoon as well as the disturbance caused by storms were accompanied by drastic
329 changes in environmental conditions. Generally, months during NE season had lower
330 temperature, lower Chl *a* concentrations, and higher salinity and nitrate values compared to
331 months during SW season (Table 1, Supplementary Figure S2). Temperature generally
332 increased in the surface waters from NE (27.8℃ - 30.6℃) to SW monsoon (30.2℃ - 32.8℃;
333 t-test, p<0.05). Meanwhile, salinity values steadily declined from an average of 30.8 $\pm$ 0.7
334 during the NE monsoon to 29.2 $\pm$ 0.7 during the SW monsoon (t-test, p<0.05) in the surface
335 waters. A similar trend was observed for nitrate concentrations, which decreased from an
336 average of 1.17 $\pm$ 0.44 µM to 0.09 $\pm$ 0.03 µM (t-test, p<0.05) during the SW monsoon. Chl *a*
337 concentrations fluctuated from a minima of 0.31 mg m^{-3} during the NE monsoon to a maxima
338 of 21.46 mg m^{-3} in the SW monsoon (t-test, p<0.05). In comparison, samples associated with
339 the storm were observed to have the lowest salinity values, averaging only 19.7 $\pm$ 3.97, lower
340 by 10 units when compared to both monsoon periods. The observed decrease in salinity could
341 be attributed to the higher amount of rainfall and runoff during those sampling dates (Table 2).
342 In addition, the lowest phosphate concentrations were recorded, from approximately 0.8 µM
343 for both monsoon periods to 0.36 $\pm$ 0.13 µM during stormy conditions. However, the highest
344 dissolved silicate concentrations (25.03 $\pm$ 10.07 µM) were recorded for those same samples,
345 compared to the average silicate values during NE and non-stormy SW monsoon samples

(Table 1), respectively. Meanwhile, nitrite concentrations (<1.0 µM) were observed to be relatively low for the entire duration of sampling.

3.2. Community diversity and composition

We observed a clear shift in community composition both at the binned reads (relative abundance; Figure 1B) and OTU levels (Figure 1C). In the samples associated with the NE monsoon, the most abundant picoplanktonic taxa present (>1% of the total reads per sample) were Chlorophyta, Syndiniales, Cryptophyta, Prymnesiophyceae, Picozoa, and Cercozoa. Larger taxa such as ciliates and dinoflagellates were also detected at more than 1% of the total reads of the picoplankton fractions but could be related to cell breakage during collection or filtration (Onda et al., 2017). In contrast, larger phytoplankton taxa such as diatoms dominated the communities during SW monsoon, making up 44% (July) to as much as 97% (September) of the entire community, consistent with previous reports on microphytoplankton composition in Bolinao (Escobar et al., 2013; Yap et al., 2004). However, the dominance of diatoms was less pronounced during the dates when storms occurred. Instead, the most abundant picoplanktonic taxa present in the storm samples were Chlorophyta and Prymnesiophyceae.

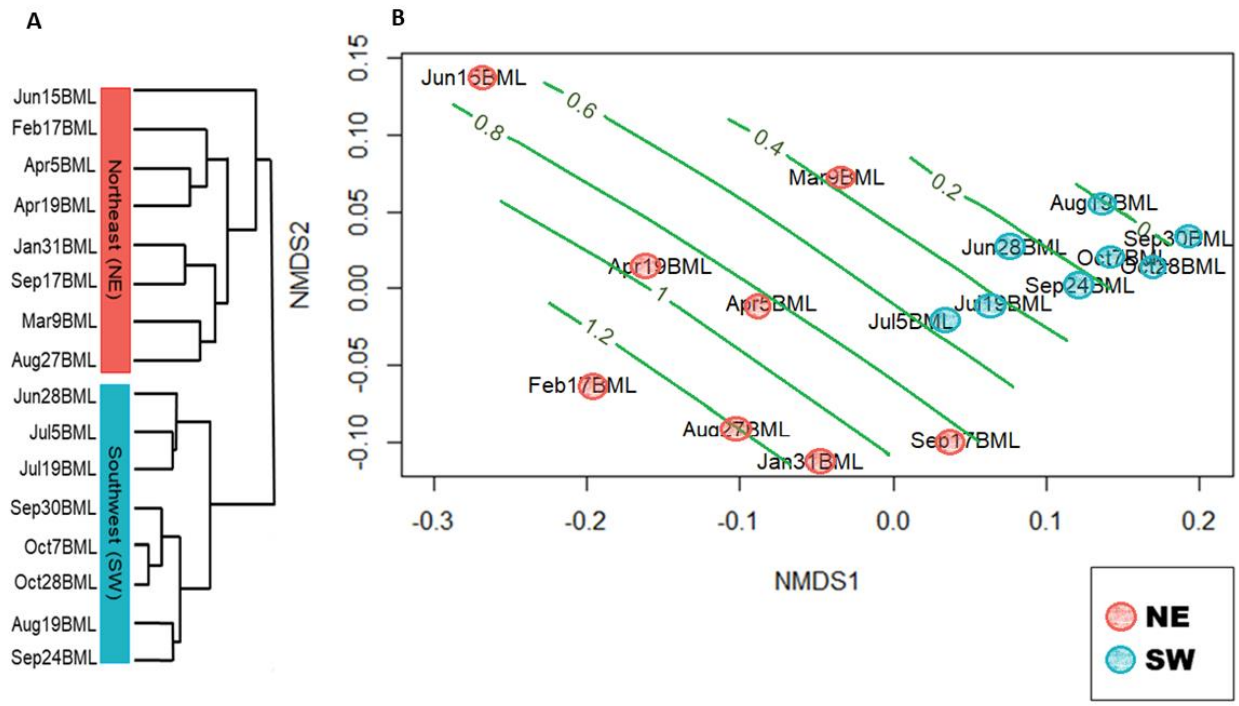


Fig. 2. A) Hierarchical clustering of OTUs based on weighted UniFrac distances. Communities were only separated into two distinct clusters; B) NMDS ordination plot also

based on weighted UniFrac distances of all OTUs (stress=0.078) showing temporal shifts influenced by nitrate (contour lines).

In contrast to the clustering based on environmental data (Figure 1A), clustering of beta diversity based on weighted UniFrac distances (Figure 2A) showed only two significant community clusters based on their phylogenetic composition (PERMANOVA, $p<0.01$). The communities during the NE and SW monsoon periods still separated into two significantly distinct clusters (r = 0.70, $p<0.001$, Supplementary Figure S3B), but no observed separate or distinct cluster for the storm samples. Instead, Aug27BML and Sep17BML samples grouped with the NE monsoon cluster, while Sep24BML sample grouped with the SW monsoon cluster.

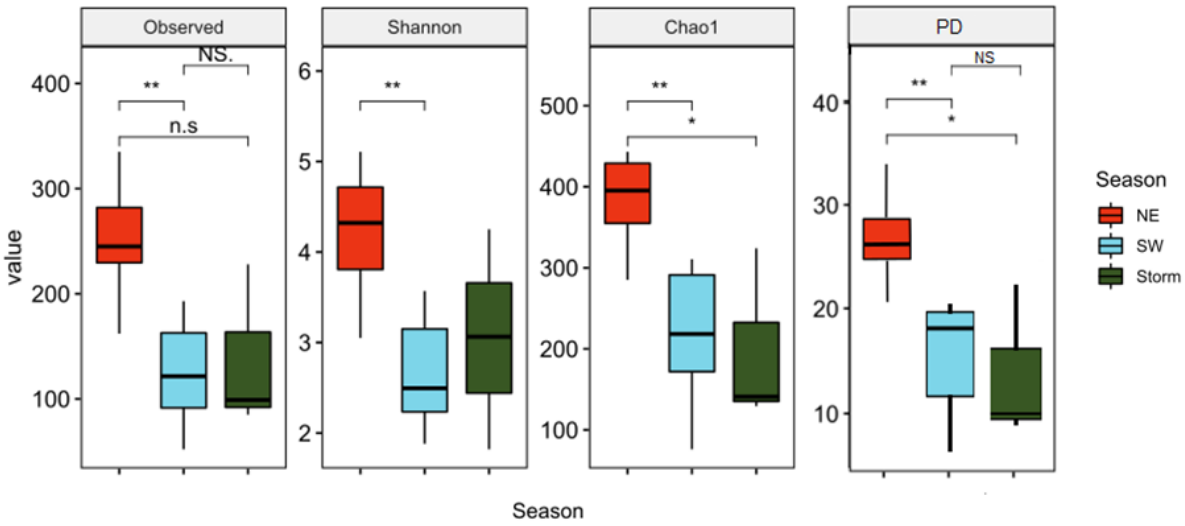
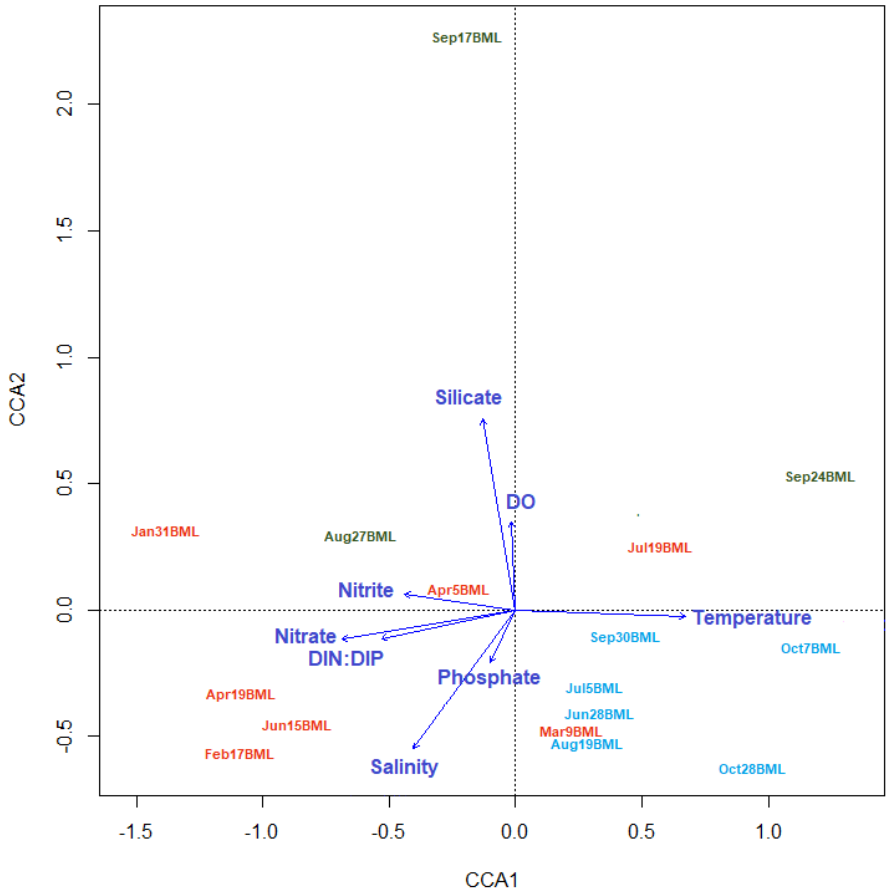


Fig. 3. Non-phylogenetic alpha diversity indices (Observed OTUs, Chao and Shannon) and phylogenetic diversity (Faith's PD) of the total eukaryotic communities present in the picoplankton fraction. The ** symbol denotes a significant p-value<0.01, * means a significant p-value<0.05, and NS means non-significant p-value>0.05.

A decrease in alpha diversity of the total community was further observed from NE to SW monsoon (Figure 3). The different alpha diversity indices we used consistently demonstrated that the mean diversity of the whole community was higher during NE monsoon (mean observed OTUs = 252 $\pm$ 59; Chao1 = 379.95 $\pm$ 57.34; Shannon = 4.22 $\pm$ 0.74; PD = 27.06 $\pm$

4.39) than during SW monsoon (mean observed OTUs = 124 ± 53; Chao1 = 209.94 ± 89.65; Shannon = 2.66 ± 0.66; PD = 15.44 ± 5.91). The drastic decrease in diversity during SW monsoon was due to the dominance of diatoms observed during that period. SIMPER analysis (Table 3) further showed that Chlorophyta and diatoms primarily contributed to the significant differences observed between monsoons. Chlorophyta dominated during NE monsoon, particularly during the colder months of January and February, and steadily declined in abundance until it was almost undetectable during SW monsoon when diatoms started to dominate. Furthermore, we also observed a decrease in alpha diversity during the stormy days (mean observed OTUs = 137 ± 79; Chao1 = 194.77 ± 109.33; Shannon = 3.05 ± 1.22; PD = 14.21 ± 7.79) compared to the NE monsoon. However, we only observed significant differences between the alpha diversities of communities during the NE monsoon and storm period for Chao1 (p<0.05) and Faith's PD (p<0.05). Meanwhile, no significant difference was observed for all alpha diversity measures between communities during the SW monsoon and the storm period (p<1.0).

3.3. Potential drivers of community structuring



399 **Fig. 4.** Canonical correlation analysis (CCA) plot based on total picoplanktonic communities
400 and measured environmental parameters where the environmental factors explained 21%
401 (sum of CCA1 and CCA2 axes) of the variance. Forward selection with function 'ordistep'
402 revealed that the best explanatory factor was nitrate. Sample names are color-coded
403 according to monsoon season, where red is for samples during NE monsoon, blue is for
404 samples during non-stormy SW monsoon, and green is for samples during stormy SW
405 monsoon.

406

407 Canonical correspondence analysis (CCA) and multiple linear regression analysis further
408 provided insights on the potential drivers influencing the temporal variability of the
409 picoplanktonic communities. At the total community level, both analyses identified nitrate as
410 the strongest environmental factor driving the seasonal community patterns. However, only a
411 total of 21% can be accounted for the variance observed in the total picoplankton communities

by the measured environmental parameters based on CCA Axes 1 and 2 (Figure 4). Forward selection in CCA only identified nitrate as the best explanatory variable. This was further explored using NMDS (Figure 2B) overlaid with the contour plots of nitrate, which revealed a relatively more scattered clustering of NE samples collected when nitrate was from 0.4 to 1.2 µM. Meanwhile, the plot showed a tighter clustering of SW samples collected when nitrate concentrations were lower from <0.1 to 0.6 µM.

In addition, while nitrate was identified to be the strongest predictor, CCA triplot also showed that temperature, dissolved silicate, and salinity could also be significantly influencing the variability of the communities (Figure 4). Interestingly, samples belonging to the storm cluster, particularly Sep17BML, correlated with high dissolved silicate levels while Sep24BML sample with low salinity. We also observed that NE monsoon samples were associated with high nitrate levels while samples collected from SW monsoon with high temperature based on the triplot. To further support our CCA results, we then performed multiple linear regression analysis with PC1 (49.6% of the variance observed) which showed temperature (p=0.010) significantly positively correlating with it, while nitrate (p=0.09) and DIN:DIP ratio (p=0.022) to be significantly negatively correlated. In contrast, silicate (p=0.081) and dissolved oxygen levels (p=0.062) positively correlated with PC2 (16.7% of the variance observed), but only slightly significant. PC1 and PC2 were based on the PCoA of weighted UniFrac distances.

3.4. Shifts in community interactions

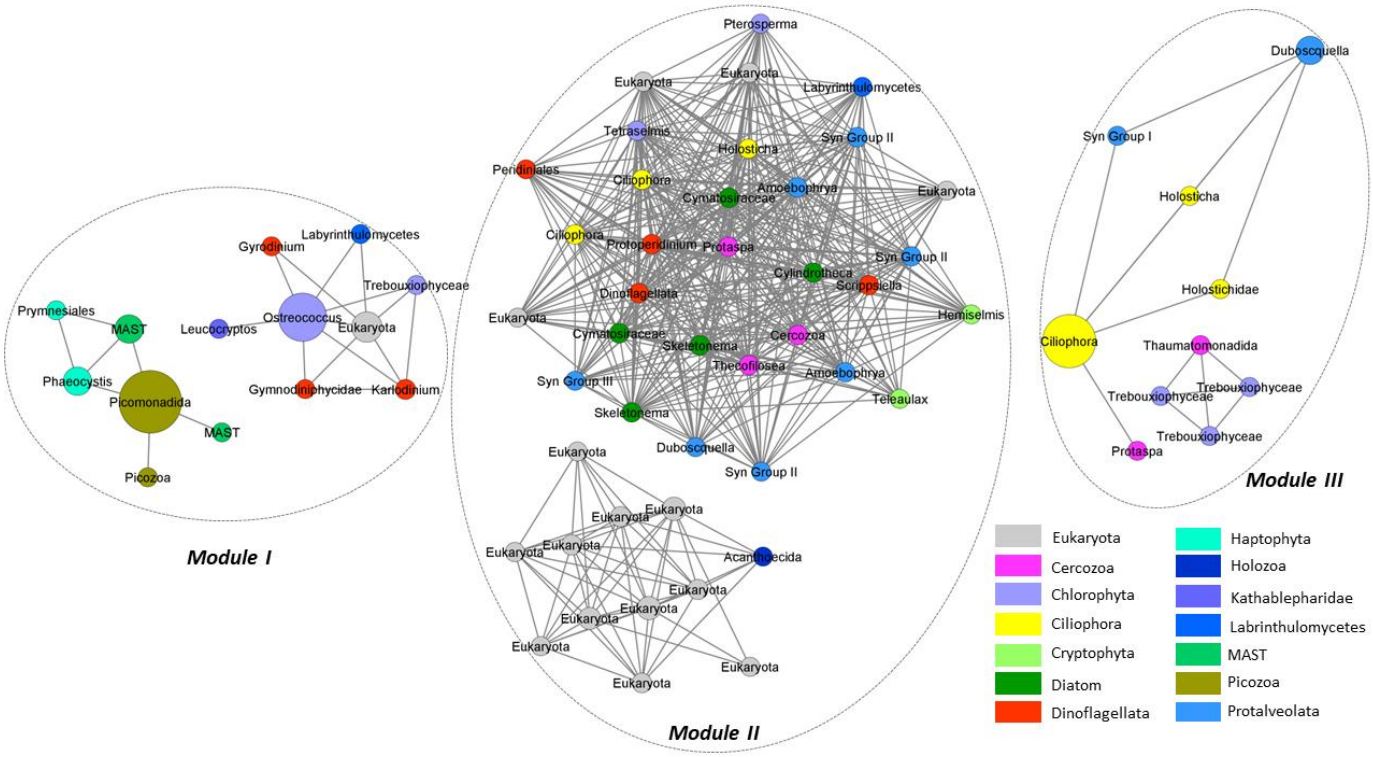


Fig. 5. Networks of the microbial eukaryotic communities during the Northeast (NE) monsoon (Module I), Southwest (SW) monsoon (Module 2) and during storms (Module 3). The colors of the nodes represent the phylum of the OTUs while the sizes represent betweenness centrality.

Shifts in the abundant taxa driven by changes in the environment also affected the possible trophic structure. In our resulting network (Figure 5), the 238 most abundant OTUs grouped only into three modules, namely those OTUs abundant during the NE monsoon (Module 1), SW monsoon (Module 2), and during storms (Module 3), similar to the clustering based on environmental parameters. Module 1 was dominated by a chlorophyte taxa, which was observed to peak in January (Figure 1C), and had a total of 2,754 reads, accounting for around 35% of the total OTU reads belonging to Module 1. We validated the phylogenetic placement of the chlorophyte sequences, which revealed that the most abundant sequence in our dataset with barcode "3f5ff" was more closely related to *Ostreococcus* Clade E (Supplementary Figure S5). Meanwhile, Module 2 was dominated by the diatom *Skeletenoma*, with a total read count of 1,561, which was approximately 21% of the total OTU reads. It was observed to have highest abundance during late June at the onset of SW monsoon but remained present the entire season

(Figure 1C). Lastly, the most abundant OTU in Module 3 was the chlorophyte *Picochlorum* with 740 reads, accounting for approximately 32% of the total OTU reads. This taxon was observed to be the most abundant during dates associated with the storm cluster (Figure 1C). Spearman's rank correlation of each module (denoted as ρ) with the measured environmental variables further highlighted the seasonality of the modules (Supplementary Figure S4). For example, Module 1 correlated with lower temperatures ($\rho = 0.34$, $p>0.10$), low Chl *a* concentrations ($\rho = 0.58$, $p<0.05$), higher nitrate concentrations ($\rho = 0.46$, $p>0.05$), and higher DIN: DIP values ($\rho = 0.74$, $p<0.01$), which marked the NE monsoon, although only the correlations for Chl *a* and DIN: DIP were significant. Low temperature, low Chl *a* concentrations, and higher nitrate values were environmental conditions observed during the NE monsoon period.

The turnover in seasonal conditions coincided with the changes in the co-occurring taxa and dominant OTUs. Based on Figure 5, Modules 1 and 3, which correspond to NE and storm clusters, respectively, had lower number of edges (42 and 28) and nodes (14 and 10) compared to the separate Module 2 (SW cluster) that had 43 nodes and 918 edges (Figure 5). In Module 1, the chlorophytes *Ostreococcus* and *Trebouxiophyceae* dominated, which also coincided with their peak abundances during the late NE monsoon. *Ostreococcus* specifically was the most connected (7 edges), mostly with dinoflagellates. Interestingly, at the start of the intermonsoon, the chlorophytes decreased and were replaced by picozoa as the most abundant taxa. Picozoa OTUs accounted for 56% of the total read counts during March and formed a sub-network together with prymnesiophytes. In contrast, most of the abundant OTUs in Module 2 (SW monsoon) were diatoms, accounting for 51% of the total reads and with 16% of all the total edges in this module. The diatoms were connected to almost all other taxa present in Module 2. Parasitic Syndiniales were the second most abundant in this module accounting for 9.5% of the total reads, and significantly co-occurring with their potential hosts, such as the dinoflagellates (12% of module total) and ciliates (8.8% of the module total). OTUs that were unclassified were also abundant particularly in June, peaking at 9.2% of the module total. Notably, the unclassified OTUs only formed connections with each other, and formed a separate sub-network that was not significantly affiliated with the diatom OTUs. Finally, Module 3 had the lowest number of OTUs with chlorophyte OTUs being the most abundant taxa, making up around 46% of the module total. Specifically, the chlorophyte *Picochlorum* (32% of the module total) was the most abundant taxon, which was significantly associated only with other chlorophytes and a Cercozoa. Similar to Module 2, parasitic Syndiniales were

also abundant in this module (14% of the module total), forming significant connections only with ciliates (16% of the module total), which could potentially be their hosts.

3.5. Changes in functional community profiles

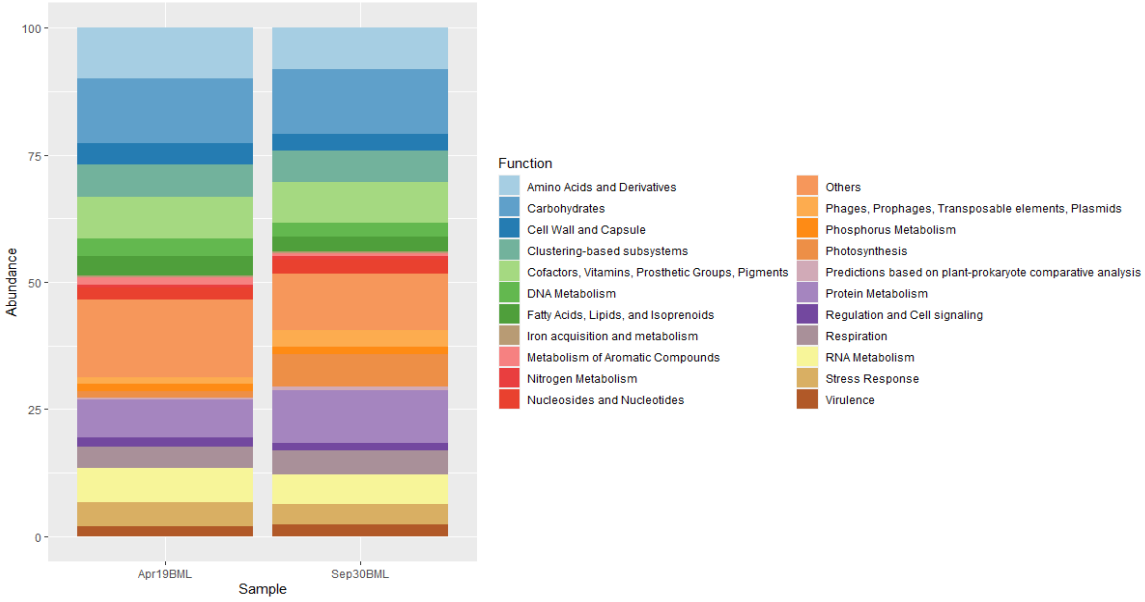


Fig. 6. Relative abundance of the "Level 1" functional categories based on eukaryotic and unclassified sequences retrieved from shotgun metagenomic analysis of the representative samples for NE and SW monsoon. "Others" category is composed of 12 other low abundance "Level 1" functional categories.

For the shotgun data, our two representative samples, namely April (Apr19BML) for late NE to early intermonsoon and September (Sept30BML) for late SW to early intermonsoon, each had a total of 42,948,874 and 42,677,963 sequence reads, respectively. We first extracted sequences inferred as Eukaryota or unclassified from both datasets, and only these were used for further processing to focus on picoeukaryotes. Using these filters, around 11% of the total sequences were extracted from the April sample while ca. 24% from September. We also observed significant differences between the resulting profiles in each sample based on the relative abundance of 'Level 1' functional categories (Figure 6). These differences were particularly visible in genes related to photosynthesis, being 5 times lower in NE monsoon (1.28%) than in SW monsoon (6.38%). Genes related to protein metabolism along with phage- and prophage-related functions were also significantly lower during NE monsoon than in SW

monsoon. In contrast, genes for the metabolism of amino acids and derivatives were slightly higher during NE monsoon (10.05% vs. 8.11%).

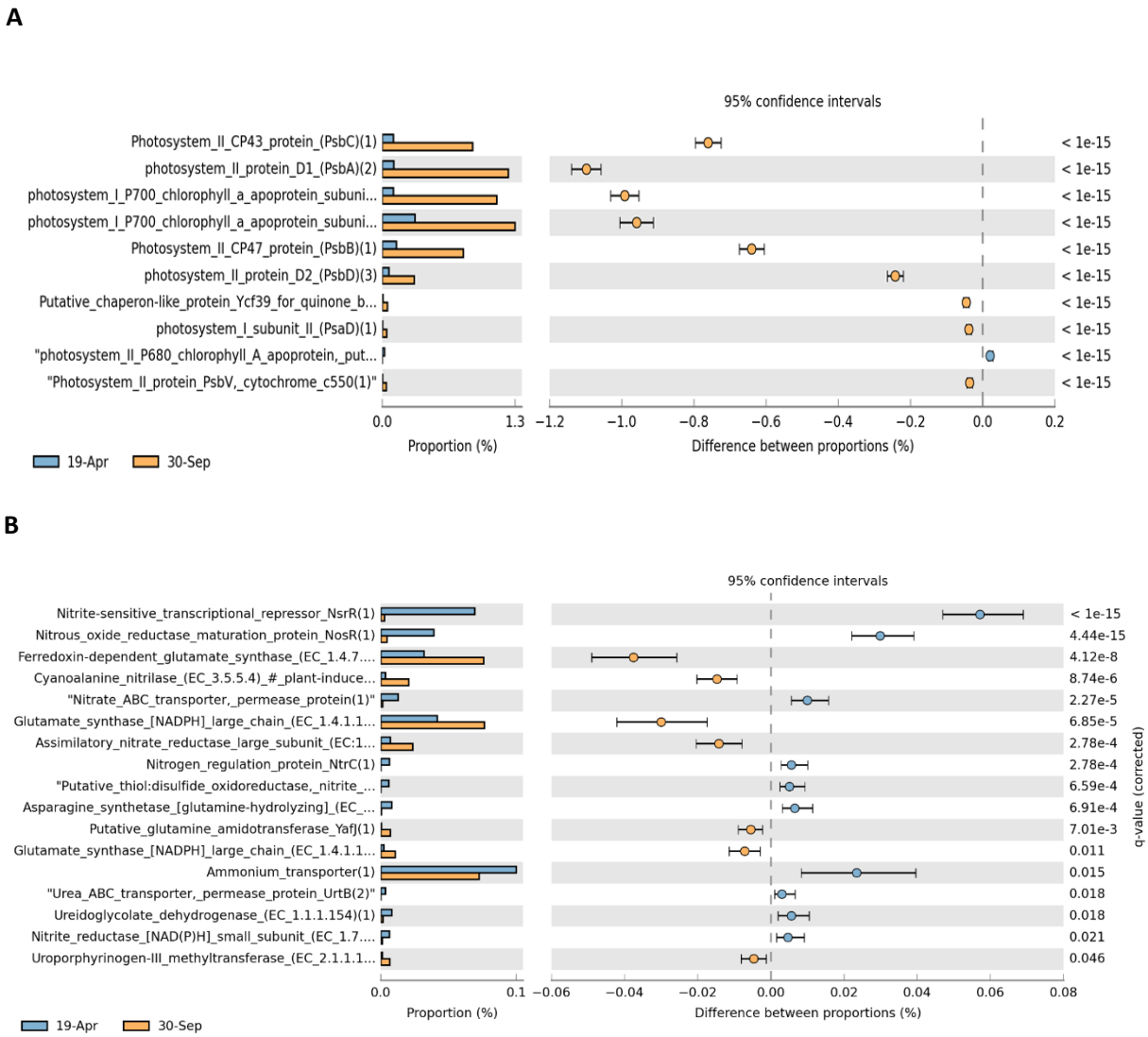


Fig. 7. Metagenomic profile comparisons of genes associated with A) photosynthesis and B) nitrogen metabolism created using STAMP. Positive differences between proportions imply that Apr19BML (blue) dataset have greater abundance for the given genes, whereas negative differences show that Sept30BML (orange) dataset have higher abundance. Only the most significant genes with corrected *p*-values (*q*-values)<$1.0e^{-15}$ were included in the photosynthesis gene profile while significant genes with corrected *p*-values (*q*-values)<0.05 were included in the nitrogen metabolism gene profile.

We then compared the abundance profiles of photosynthetic-related genes at a finer level and noted that almost all genes related to both photosystems I and II were higher during SW monsoon when diatoms dominated (Figure 7A). Most of the significant photosynthetic-related genes detected (corrected $p<0.05$) were related to the proteins that form the photosystem II core complex, including CP43 (*PsbC*) and CP47 (*PsbB*) core antenna proteins as well as the D1 (*PsbA*) and D2 (*PsbD*) reaction center core proteins. These genes were all detected at an abundance of almost 10 times lower in NE monsoon than in SW monsoon. In particular, the proportions of D1 (*PsbA*) and D2 (*PsbD*) during NE monsoon were only around 0.11% and 0.07% respectively, lower than 1.2% and 0.31% proportions in SW monsoon. Similarly, the proportions of CP43 (*PsbC*) and CP47 (*PsbB*) increased from 0.11% and 0.14% during NE to 0.87% and 0.78% respectively in SW samples. Furthermore, genes important to the electron transport activity in photosystem I, specifically P700 (*PsaA*) and P700 (*PsaB*), were also lower in abundance during NE monsoon at 0.11% and 0.31%, respectively, and increased to 1.10% and 1.27% during SW monsoon. We further identified the taxonomic identity of the reads that were associated with the most differentially abundant photosynthetic genes (*PsbC, PsbA, PsaA, PsbB),* and observed that the April sample was composed of around 51% chlorophytes (Figure 8A) while the September sample was composed of approximately 90% diatoms (Figure 8B), consistent with the amplicon datasets. The phylogenetic trees using the predicted protein sequences further agree with our previous results (Supplementary Figure S6). In general, genes from the April sample were associated with chlorophytes particularly *Ostreococcus,* while genes from the September sample with diatoms such as *Thalassiosira*.

Another significant change observed was those related to nitrogen metabolism, consistent with the correlations showing that nitrate was the most significant driver of community structuring (Figure 7B). Results suggest that genes associated with ammonia assimilation were the most dominant for both seasons, although lower for NE monsoon than in SW monsoon (0.16% vs. 0.20%). Majority were ammonium transporter or regulator during NE monsoon (0.09%) and were slightly more abundant than during SW monsoon (0.06%). Meanwhile, genes associated with ammonia assimilation mostly were glutamate synthase during the SW monsoon (0.14%) and were higher than NE monsoon (0.06%).

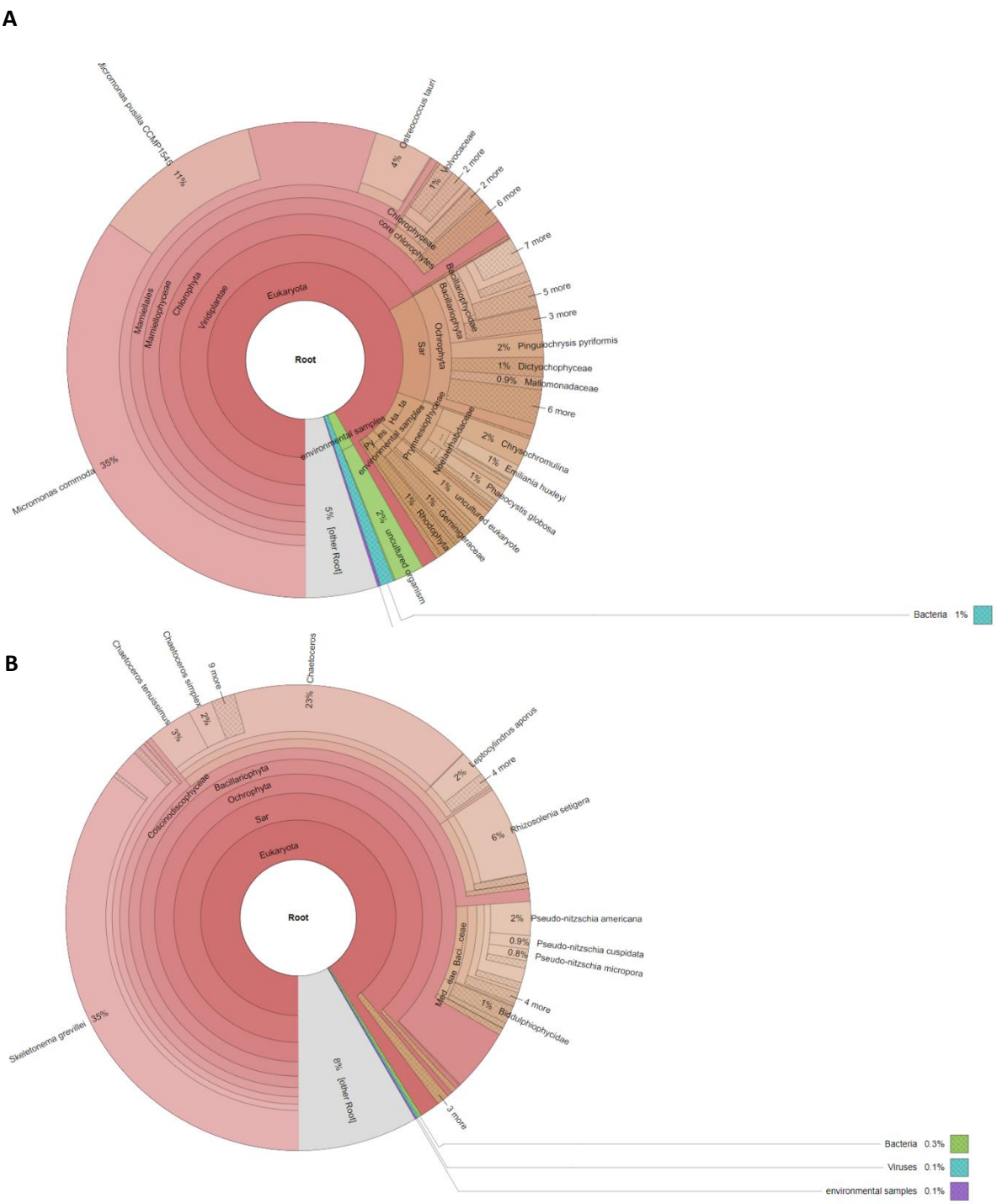


Fig. 8. Potential taxonomic identity based on BLAST against the nt database of the reads from A) Apr19BML and B) Sep30BML samples, that were associated with the top four most differentially abundant photosynthetic genes (*PsbC*, *PsbA*, *PsaA*, *PsbB)*.

4. DISCUSSION

4.1. Shifts in environmental conditions

Clustering of the samples based on measured environmental parameters resulted in three major groups, namely from January to early June, from late June to October, and some samples from August to September. The first two groups are the months historically associated with NE and SW monsoons, respectively (Wyrtki, 1961), while the last group, although being within the SW monsoon, were the days associated with the occurrence of storms or heavy rainfall events. Such observations are typical of tropical coastal settings, such as in Bolinao, where changes in conditions with season are mainly brought about by the reversal of the monsoon winds (Wyrtki, 1961; Rivera, 1997). The ensuing changes in the conditions due to monsoon-driven shifts as well as episodic disturbances (i.e., storms) have significant influences on the local environmental conditions and thus, on the productivity of the different marine ecosystems including the coastal habitats (Canini et al., 2013; Mitbavkar et al., 2015).

Changes in temperature were observed between the identified monsoon periods but were very small when compared to temperate areas. The highest temperature was observed around July, consistent with previous observations that water temperature slightly lags with the change in air temperature, with maximum values observed around June and July (Escobar et al., 2013; Rivera, 1997). These warm conditions may result in thermal stratification, resulting in shallower mixed layer depth reducing nutrient transport from deeper waters to the surface (Gittings et al., 2018). Salinity was also observed to be generally higher during NE period possibly due to the higher evaporation rates and minimum freshwater input (Rivera, 1997), except during La Nina years (Ramos et al., 2017), although inputs from submarine groundwater discharge have also been reported in the area (Cardenas et al., 2010). In contrast, heavy precipitation during SW monsoon with strong freshwater input from runoff and riverine discharge significantly dilutes the surface waters lowering the salinity (Canini et al., 2013; Ravelo et al., 2022). Such discharges from land and rivers however also bring in the needed nutrients, causing a net influx of resources for phytoplankton growth and primary productivity (Ferrera et al., 2016; Yap et al., 2004). Here, a significant increase in Chl *a* was observed with increased phytoplankton biomass during the SW monsoon, similar to previous reports in the area (Ferrera et al., 2016; Yap et al., 2004).

Occurrence of storms characterized by strong winds (> 62 kph) and heavy rainfall (> 9 mm) may also reduce the stratification and cause turnover of the water column, also replenishing the nutrients in the surface waters (Gittings et al., 2018; Rivera, 1997). Interestingly, some of the samples collected in August and September formed a separate group, which were

incidentally collected when gale or tropical storm warnings were up. Specifically, Aug27BML sample was collected when a tropical cyclone wind signal was raised over the area due to tropical storm 'Jenny' (PAGASA, 2019). Meanwhile, although a storm signal was not raised during the collection day for Sep17BML sample, rains and heavy winds were still present brought by the trough or extension of tropical depression 'Nimfa' (PAGASA, 2019), which passed by the area a few days before. The lowest salinity values within the SW monsoon period were indeed recorded during these stormy days. We also observed low nutrient concentrations (nitrate, nitrite, and phosphate) but high Chl *a* concentrations, possibly indicating active uptake by phytoplankton (Ferrera et al., 2016). Notably, unlike the other periods within the SW monsoon, the storm cluster was observed to have higher dissolved silicate values (Figure 1A; Figure 3). Silicate generally during SW monsoon should be high due to freshwater forcing; however, diatoms were likely to have utilized the silicate, thus, possibly explaining the low silicate measurements (Escobar et al., 2013). During storms however, diatoms were the most affected as characterized by a significant decline in their abundance, resulting also in lower uptake and thus, buildup of silicate in the sea surface. These fluctuating conditions caused by monsoonal shifts, accompanied by increased sediment transport and riverine input during SW monsoon make establishing predictable patterns of microbial communities difficult (Lambert et al., 2019). These changes in local conditions caused by the monsoon-driven shift as well as the episodic disturbances could select organisms that could adapt, and thus, would have significant effects in the picoeukaryotic communities.

4.2. Effects of monsoons and storms on community patterns

In Bolinao, our results revealed two distinct picoeukaryotic communities similar to our monsoon-based grouping, indicating the significant effect of monsoonal shifts on the structuring of microbial communities (dela Peña et al., 2021). Community and phylogenetic diversities among communities during NE monsoon were higher compared to the ones during SW monsoon, when diatoms started to dominate. No distinct storm cluster was observed, instead most communities during stormy period grouped together with NE communities, likely because disturbance caused by the storm events disrupted the dominance of diatoms, causing an increase in community diversity. For example, during Aug19BML, a week before a typhoon was recorded in the area, diatoms accounted for approximately 80.2% of the total community reads. A week after, on Aug27BML, during typhoon 'Podul' (local name 'Jenny' in the Philippines), the diatoms only accounted for around 1.2% of the total community reads. Instead

of diatoms, we observed that there was an increase in the abundance of chlorophytes specifically, *Picochlorum,* which are organisms known to thrive in freshwater brought in by the storms (Foflonker et al., 2016). Further, abundance of mixotrophic or wholly heterotrophic species such as Cercozoa and Cryptophyta increased, perhaps due to the decrease of light intensity related to the increased turbulence and terrigenous inputs that accompanied the storms. The increase in community diversity during short-term disturbance is consistent with the intermediate disturbance hypothesis (Fox, 1979), which postulates that increase in disturbance results to an increase in diversity up to a certain level of strength until diversity decreases or stabilizes. In addition, we also observed an increase in the phylogenetic diversity from pre-storm (Aug19BML: PD = 19.74) to during a typhoon (Aug27BML: PD = 23.19), rising to above the average PD observed during the SW monsoon. This may indicate that additional niches were created, and that more phylogenetically distinct species may co-occur during storms than during the SW monsoon (Galand et al., 2016).

Multivariate approach further revealed that nutrients, specifically nitrate, which was observed to have decreased from NE to SW monsoon, was the strongest predictor of shifts in picoeukaryotic community structuring. This is consistent with previous microphytoplankton studies, and was likely because our site was at the opening of the Guiguiwanen channel, where most of the mariculture activities take place making the area N-limited (Ferrera et al., 2016). This shift in the community was also accompanied by the shift in the general taxonomic composition especially in the most abundant phytoplankton groups. This is consistent with previous reports, which observed opposite patterns between small and large-celled taxa in coastal environments (Collado-Fabbri et al., 2011). This shift in the abundant taxa to the larger diatoms during SW monsoon also resulted in lower diversity (Shannon), rareness (Chao1), and richness (no. of observed OTUs) compared to NE monsoon. The detection of larger taxa such as diatoms even though our study was limited to pico-fraction could be due to cell breakage during filtration (dela Peña et al., 2021; Onda et al., 2017) or some pennate diatoms maybe passing through the filter membranes such as *Pseudo-nitzschia* which was relatively abundant in our study (Vaulot et al., 2008). Nevertheless, such abundance of reads could still indicate a sudden and significant shift in the cells being filtered when the samples were taken. Generally, we observed that communities associated with NE monsoon experienced relatively colder and more saline environmental conditions relative to those associated with SW monsoon. Productivity (using Chl *a* as the proxy) was also lower during the NE monsoon, despite recording higher mean nitrate concentration compared to SW monsoon. In contrast, during SW,

644 the lower nitrate concentration despite potential inputs from runoff and riverine/groundwater
645 discharge (Ferrera et al., 2016) may be due to active uptake by opportunistic chain-forming
646 diatoms such as *Skeletonema* and *Thalassiosira*, which overwhelmingly dominated during this
647 period, accompanied by a drastic increase in Chl *a*. Aside from nitrate, turbulence, here
648 represented by strong winds and heavy rainfall events during tropical storms, was not measured
649 directly but may also be a significant driver of community structuring since previous studies
650 have reported that turbulence could select for particular species and regulate phytoplankton
651 community structure in aquatic environments (Trigueros & Orive, 2001; Zhou et al., 2016).
652 Margalef's Mandala hypothesis predicted that in high nutrient and high turbulent conditions,
653 r-strategists such as diatoms will be more favored over K-strategists such as dinoflagellates
654 (Margalef, 1978). In the Philippines, based on rainfall and wind speed data recorded in 2019
655 from a weather station approximately 90 km east of our study site (Table 2), rainfall was higher
656 during SW monsoon. Thus, it can be deduced that disturbance was more frequent during this
657 period. Moreover, there were 21 typhoons in 2019, with 13 of them making landfall and passed
658 through our study site during the same period (PAGASA, 2019). The higher frequency of
659 typhoons and greater disturbance would have caused more frequent high turbulent mixing
660 conditions, which have been reported to favor diatoms possibly due to their high adaptability
661 to fluctuating light conditions and to reduce grazing pressure during turbulent conditions
662 (Margalef, 1978; Zhou et al., 2016). However, we were not able to uncover how much force
663 was optimal for diatoms to continue to thrive in this area as studies have also noted that
664 excessive turbulent force might have negative effects such as mechanical damage to
665 phytoplankton (Fraisse et al., 2015; Thomas & Gibson, 1990). In addition, as previously noted,
666 turbulence associated with storm events may likely disrupt proliferation of diatoms due to
667 decrease of light availability.

668 Smaller phytoplankton taxa have been suggested to be more adapted to low-nutrients
669 conditions due to their higher volume-to-surface area ratio (Li et al., 2009). Aside from this
670 however, genes related to nutrient assimilation, particularly nitrogen may also help. Indeed,
671 changes in taxonomic profiles were accompanied by changes in the abundant functional genes
672 especially those related to nutrient assimilation and photosynthesis. Our shotgun sequencing
673 data for example revealed that ammonium transporter (AMT) genes were abundant for both
674 seasons, but slightly higher during NE monsoon. It would appear that uptake of ammonium is
675 prioritized over other forms of inorganic nitrogen. This is likely because ammonium is less
676 costly to assimilate since it can be directly used for amino acid synthesis via the glutamate

synthetase/glutamine oxoglutarate aminotransferase (GS/GOGAT) cycle in photosynthetic eukaryotes (Hopes & Mock, 2015; McDonald et al., 2010; Sanz-Luque et al., 2015). AMT genes have also been reported for chlorophytes, prymnesiophytes, and diatoms (McDonald et al., 2010; Sanz-Luque et al., 2015), which were the major taxa identified to comprise our representative NE monsoon sample (Figure 8A). In contrast, genes for both NADPH-dependent and ferredoxin-dependent glutamate synthase increased during SW monsoon. The presence of both functional genes gives a competitive advantage in assimilating ammonia since the former allows synthesis even during conditions of low irradiance while the latter performs during conditions of high or continuous irradiance (Suzuki & Rothstein, 1997). The increased presence of these genes highlighted the importance of increased capacity for ammonia assimilation during this period. We also observed that genes related to transport of nitrate and genes related to reduction of nitrite decreased in SW monsoon. The lower abundance of genes for the transport and reduction of nitrogen coupled with increase for ammonia assimilation concurs with previous culture studies on the response of the diatoms *Thalassiosira pseudonana* and *Skeletonema costatum* to nitrogen starvation (Hockin et al., 2012; Takabayashi et al., 2005). The increase in biomass of photosynthetic taxa, especially diatoms, could also explain the increase of genes responsible for photosynthesis from NE to SW monsoon. This is consistent with previous studies where increase in photosynthetic taxa likely resulted in higher relative abundance of photosynthesis-related genes (Warden et al., 2016; Yergeau et al., 2017).

CCA however only accounted for 21% of the variance observed for axes 1 and 2 (Figure 4), indicating that other factors not measured may be contributing to the observed community structures. These include current speed and tides, total suspended solids (TSS), and light availability (Bryant et al., 2016; Canini et al., 2013; Yap et al., 2004), noting that the site was also near a mariculture area (Albelda et al., 2019; dela Peña et al., 2021). Yap et al. (2001) for example reported that TSS was higher during NE monsoon due to the lower current speed during this period, resulting in longer residence time of suspended solids in the water column. The high TSS might have caused lower light availability during the NE monsoon (Canini et al., 2013; Yap et al., 2004). During these low-light conditions, picoeukaryotes with smaller biomass compete better than diatoms which might explain the abundance of Chlorophyta during NE monsoon. These environmental trends were similar to those reported in previous studies in this site (Escobar et al., 2013; Ferrera et al., 2016). Overall, our results demonstrated that community composition and structure were significantly distinct for each monsoon season indicating the high turnover of communities caused by the accompanying conditions in each

monsoon period. Our results further showed that disturbances such as storm events may disrupt dominance of a species, thus, increasing community diversity.

4.3. Co-occurrence and implications

Microbial network analysis is an explorative technique used in microbiome studies to predict significant biotic interactions as well as to explore the contribution of specific taxa to global biogeochemical processes (Faust, 2021; Jones et al., 2018; Lima-Mendez et al., 2015). Using this approach, our resulting network (Figure 5) had a low average clustering coefficient (CI) of 0.824, which implies that our system is highly interconnected (Jones et al., 2018). In addition, its average shortest path length (AL) was 1.26, further indicating that the response of the microbial picoeukaryotes in this ecosystem to disturbance is fast (Faust & Raes, 2012). These network properties are typical for microbial communities and are similar to those reported in freshwater lakes (Debroas et al., 2017; Jones et al., 2018) and in marine open ocean systems (Berdjeb et al., 2018; Lima-Mendez et al., 2015), where seasonal turnover seems to be significant or environmental heterogeneity is high.

We filtered our network to only show strong co-occurrence patterns (correlation threshold = >0.5) between taxa, which could be interpreted as a possible direct relationship (*e.g.*, grazing) or as taxa potentially preferring similar environmental conditions (Faust & Raes, 2012; Lima-Mendez et al., 2015). Our results were also consistent with our monsoon-based clustering, implying that although community composition did not shift drastically during storms, dominance of certain taxa appeared to have been significantly altered. Further, similar patterns were also reported in other tropical coastal sites, where stratified conditions due to increasing temperature and lower nutrients during the intermonsoon likely favored the growth of the mixotrophic prymnesiophytes and heterotrophic picozoa (dela Peña et al., 2021). We observed that the NE monsoon-associated module involved two dominant chlorophytes, identified as *Ostreococcus* belonging to Clade E, and *Trebouxiophyceae*. Interestingly, *Ostreococcus* was more connected than *Trebouxiophyceae*, co-occurring with mixotrophic dinoflagellates and heterotrophic eukaryotes, despite both taxa accounting for similar percentage of module total. Although reports of grazing activity on *Ostreococcus* have been very limited, larger microplankton taxa such as dinoflagellates have been known to feed on chlorophytes (Jeong et al., 2010). It also had a higher betweenness centrality (0.468), which suggests that *Ostreococcus* had higher control over the interactions of other nodes in the module than *Trebouxiophyceae*. This observation may hint of a food preference for *Ostreococcus* in both

742 microplankton and nanoplankton predators. This potential predator-prey relationship however
743 needs further exploration to determine its validity and explore the value of *Ostreococcus* as the
744 base of the tropical trophic food web.

745 During SW monsoon, the dominant taxa in terms of abundance shifted from chlorophytes to
746 diatoms. However, despite their abundance, we observed that diatom OTUs were not the most
747 connected. This is consistent with observations for co-occurrence networks using the global
748 data of *Tara* Oceans where the most abundant OTUs did not have the most connections
749 (Vincent & Bowler, 2020). We also observed that in contrast to the NE-associated module, the
750 SW-associated module was highly interconnected, suggesting that the interactions were not
751 species-specific and possibly that the predators have no problem shifting from one prey to
752 another (Käse et al., 2021; Vincent & Bowler, 2020). The flexibility of the possible predators
753 and parasites might indicate higher stability and adaptability of the trophic food web system
754 during SW monsoon than during NE monsoon (Käse et al., 2021). Interestingly, we also
755 observed that, instead of diatoms, the parasitic Syndiniales, which are mostly obligate marine
756 parasites (Guillou et al., 2008), had the highest number of connections for this module similar
757 to oceanic interactome (Lima-Mendez et al., 2015), demonstrating the ubiquity of these
758 parasitic groups. The abundance of Syndiniales associations in our network further highlights
759 the possible significant role of parasitism in this coastal ecosystem. Previous studies have
760 already suggested the significant ecological impacts of parasites like Syndiniales on the
761 population dynamics of the phytoplankton by promoting mortality of their hosts (Chambouvet
762 et al., 2008). However, host-parasite associations are difficult to determine just based on the
763 network alone since the interactions during this period within and between genera can also
764 become less species-specific (Vincent & Bowler, 2020). Most of the Syndiniales in this module
765 were identified as Syndiniales Group II, with two OTUs further assigned to the genus
766 *Amoebophrya*. It is a genus classified as a "species complex" composed of strains with varying
767 degrees of host specificity, ranging from extremely species specific to nonspecific
768 (Chambouvet et al., 2008; Chen et al., 2018; Coats & Park, 2002; Kim, 2006). Recently, a
769 polyphasic approach combining genetic and phenotypic characters delineated eight individual
770 cryptic *Amoebophrya* species, with different host ranges (Cai et al., 2020). It is possible that
771 the Syndiniales present are generalists and can infect various hosts, but it is also highly likely
772 that the Syndiniales were in the free-living dinospore stage and acted as food for other protists
773 taxa such as ciliates. Further studies should be done to investigate and establish the role of
774 parasitism in shaping phytoplankton community dynamics in this ecosystem.

In contrast to the non-stormy SW period, the dominating taxa during stormy events shifted back to the smaller chlorophyta from diatoms, particularly being dominated by *Picochlorum*. It is a fast-growing organism that can adapt to extreme environmental fluctuations particularly salinity (Foflonker et al., 2016; Goswami et al., 2022), which was very evident during the stormy days in Bolinao. Notably, *Picochlorum* did not form any connections with other taxa and instead formed a sub-network consisting almost exclusively of OTUs belonging to the same taxa. This demonstrated that these chlorophyte OTUs occurred sporadically in high abundance, possibly brought about by the large volume of freshwater discharge during the storm (Käse et al., 2021). The main source of freshwater discharge in our study site are rivers and submarine groundwater around the area (Rivera, 1997). The other sub-network in this module consisted mostly of ciliates co-occurring with parasitic Syndiniales classified as Group I. One of the Syndiniales OTU was further assigned to the genus *Euduboscquella* (Coats et al., 2012), known as a parasite infecting mainly tintinnid ciliates (Horiguchi, 2015), thus highlighting a potentially significant parasitic interaction during this period.

Overall, our results revealed how specific taxa respond to changes in conditions brought about by changes in monsoons. Our networks demonstrated a potentially more stable food web system during the SW monsoon than during the NE monsoon due to the presence of more adaptable predators and parasites, however, fluctuations caused by storms appeared to select for dominant taxa. The changes we have observed here to the base of the food web system may have impacts on higher trophic levels, and thus, should be considered in future studies.

5. CONCLUSIONS

Our combination of community structuring and network analysis indicates a high turnover of picoeukaryotic communities associated with monsoon changes. Network analysis further revealed that communities during the SW monsoon might have higher stability and adaptability to natural environmental fluctuations than during the NE monsoon, however, drastic environmental changes due to storm events could select for dominant taxa. This resulted in higher diversity of communities when storms occurred compared to non-stormy SW monsoon conditions. Stability of the SW communities might be due to the presence of less selective predators and parasites as well as higher variability of prey options. In addition, functional genes, particularly genes related to nitrogen assimilation, might have also enabled them to

adapt to changing nutrient conditions. This study provided insights on how picoeukaryotic communities and their key taxa respond to changes caused by monsoonal shifts and by storms in a tropical coastal ecosystem. The implications in our study should help in further understanding picoeukaryotic community adaptation and resilience in a tropical coastal ecosystem.

TABLES

Table 1. Sampling dates in 2019 (Date) and measured environmental parameters of the samples used in this study. *NS* denotes no sample, *nr* denotes no replicate, and *nd* denotes that the actual concentration is below the detected limit of the instrument. The mean values ± standard deviation of two replicates were reported for nitrate, nitrite, phosphate and silicate. DIN: DIP ratio was calculated using the sum of the average nitrate and nitrite values divided by the average phosphate concentration.

Date	**T (°C)**	**Salinity**	**Chl *a* (mg $m^{-3)}$**	**NO_3^- (µM)**	**NO_2^- (µM)**	**PO_4^{3-} (µM)**	**DIN: DIP**	**SiO_4^{4-} (µM)**	**DO (mg L^{-1})**
31-Jan-19	27.8	32.4	*NS*	2.67 ± 0.03	0.29 ± 0.01	1.55 ± 0.05	1.91	6.36 ± 0.12	6.0
17-Feb-19	27.9	33.3	7.24	2.59 ± 0.22	0.11 ± 0.00	1.63 ± 0.06	1.66	6.69 ± 0.02	4.7
9-Mar-19	30.3	29.3	3.16	0.66 ± 0.05	*nd*	0.59 ± 0.09	1.06	2.73 ± 0.04	5.7
5-Apr-19	29.7	31.7	1.51	*nd*	0.02 ± 0.14	0.47 ± 0.12	0.05	5.57 ± 0.52	5.6

19-Apr-19	30.5	29.1	1.41	1.79 ± 0.41	*nd*	0.33 ± 0.01	5.48	8.72 ± 2.03	6.2
15-Jun-19	30.6	31.2	0.31	0.25 ± 0.27	0.07 ± 0.06	0.45 ± 0.27	0.71	5.45 ± 0.07	3.6
28-Jun-19	31.3	30.9	8.41	0.14 ± 0.14	*nd*	1.11 ± 0.04	0.12	4.81 ± 0.22	3.7
5-Jul-19	31.2	30.8	20.5	0.19 ± 0.19	*nd*	0.16 ± 0.02	1.18	0.59 ± 0.00	6.5
19-Jul-19	34.3	28.3	1.60	0.24 ± 0.13	0.13 ± 0.08	0.24 ± 0.14	1.54	6.15 ± 0.28	6.2
19-Aug-19	30.2	26.6	7.32	0.04 ± 0.01	*nd*	0.38 ± 0.02	0.10	3.09 ± 0.08	5.1
27-Aug-19	29.8	27.2	*NS*	0.08 ± 0.08	*nd*	0.11 ± 0.02	0.75	15.74 ± 0.02	5.3
17-Sep-19	30.3	18.2	10.67	0.11 ± 0.00	*nd*	0.49 ± 0.01	0.22	45.16 ± 5.34	6.0
24-Sep-19	30.4	13.7	32.71	0.15 ± 0.08	0.05 ± 0.20	0.49 ± 0.10	0.40	14.20 ± 0.95	6.6
30-Sep-19	30.6	27.5	17.18	0.11 ± 0.05	*nd*	0.23 ± 0.02	0.47	4.84 ± 0.95	6.4

7-Oct-19	32.8	29.0	*NS*	0.05 *(nr)*	0.02 *(nr)*	0.83 ± 0.02	0.08	3.06 ± 0.34	5.4
28-Oct-19	32.3	30.1	21.5	0.03 ± 0.02	*nd*	1.92 ± 0.07	0.02	1.93 ± 0.10	2.1

820

821 **Table 2.** Rainfall (mm) and wind speed (m/s) data of each sampling date were recorded from
822 a station ~90 km east of Bolinao (Dagupan monitoring site of the Philippine Atmospheric,
823 Geophysical and Astronomical Services Administration-PAGASA).

824

SAMPLING DATE	**RAINFALL, mm**	**WIND SPEED, m/s**
January 31, 2019	*nil*	2
February 17, 2019	*nil*	3
March 9, 2019	*nil*	3
April 5, 2019	*nil*	3
April 19, 2019	0.1	3
June 15, 2019	4	3
June 28, 2019	54.4	2
July 5, 2019	*nil*	2
July 19, 2019	33	2
August 19, 2019	*nil*	2
August 27, 2019	1.6	3
September 17, 2019	3.4	3

September 24, 2019	*nil*	3
September 30, 2019	*nil*	3
October 7, 2019	*nil*	3
October 28, 2019	*nil*	2

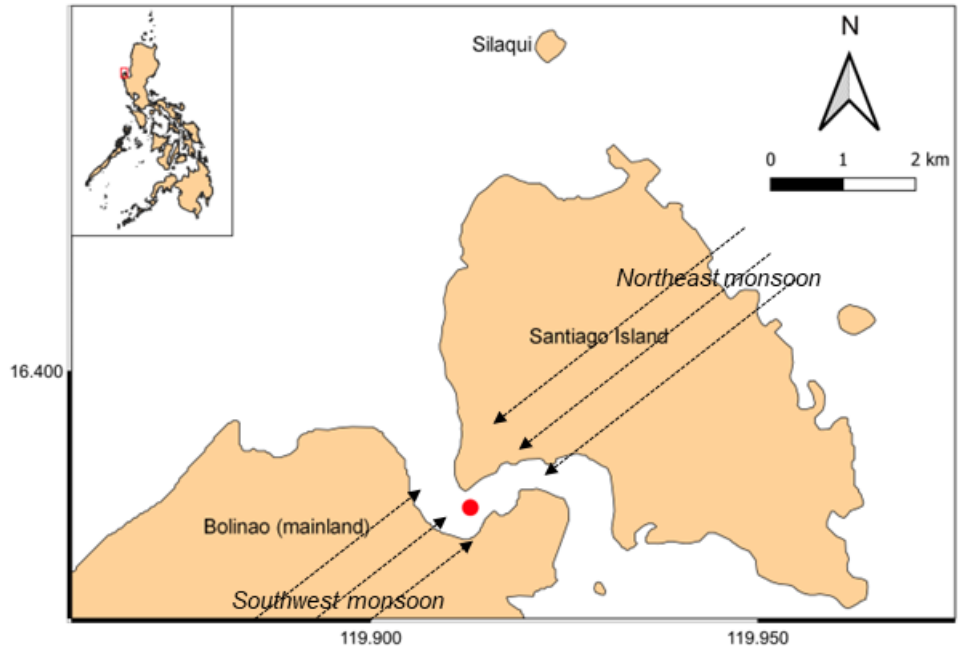
825

826 **Table 3.** SIMPER analysis showing the most influential OTUs and their respective percent
827 contribution to the pairwise dissimilarity between NE, SW and Storm samples.

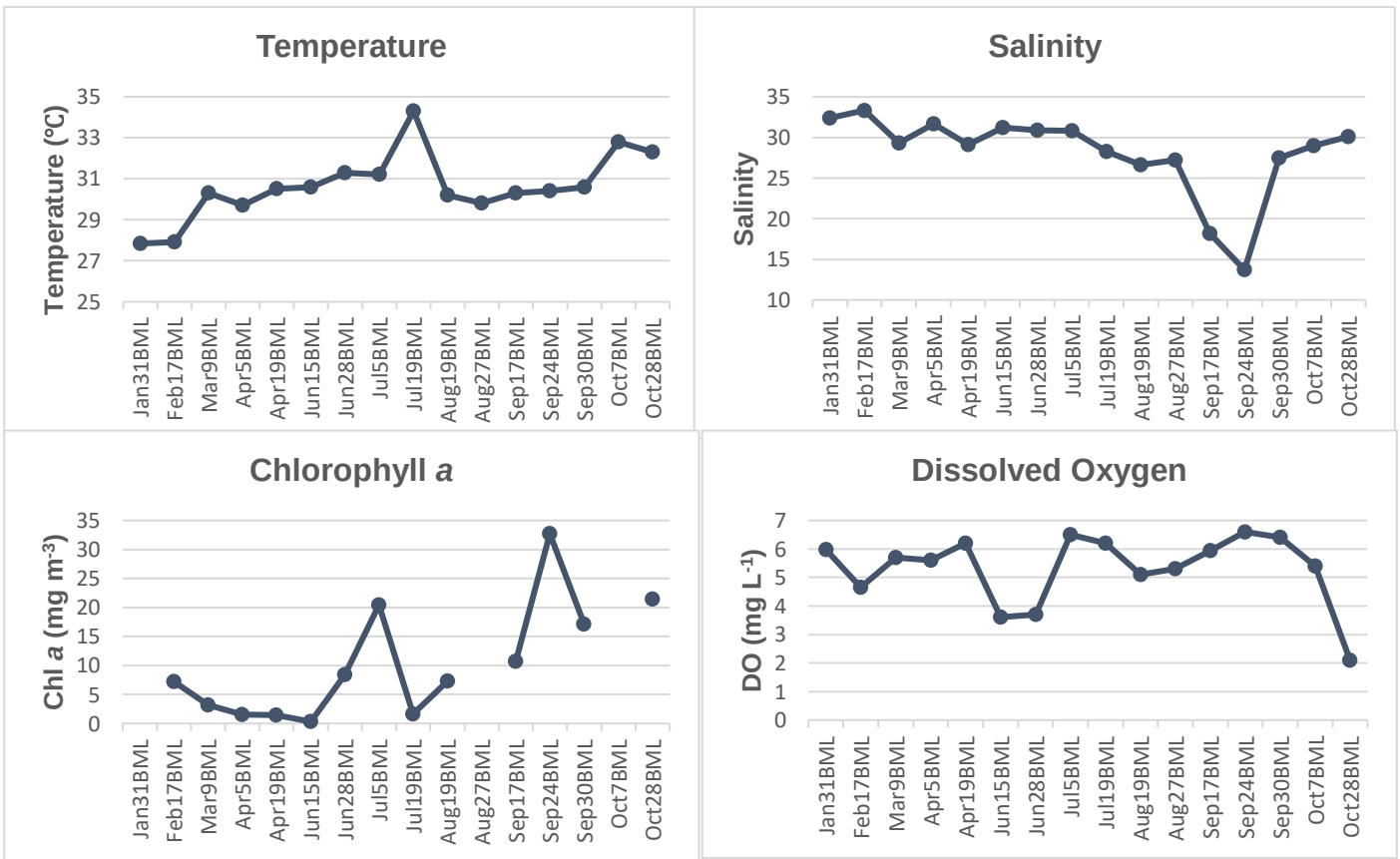
Taxon	**Most resolved taxonomy**	**Percent contribution to pairwise dissimilarity**			**Average abundance (reads)**		
		NE/ Storm (27.46%)	NE/ SW (40.09%)	SW/ Storm (53.81%)	NE	SW	Storm
Diatomea	*Chaetoceros*	0.55	10.55	17.32	53	1817	114
Diatomea	*Skeletonema*	*nil*	7.72	13.14	7	1298	6
Diatomea	*Skeletonema*	2.69	8.33	11.15	1	1395	299
Diatomea	*Thalassiosira*	9.05	4.20	3.05	8	711	1011
Chlorophyta	*Ostreococcus*	5.10	3.27	*nil*	588	41	23
Chlorophyta	*Picochlorum*	2.70	*nil*	3.34	37	8	336
Chlorophyta	*Micromonas*	3.22	2.61	0.81	439	2	82
Diatomea	*Rhizosolenia*	*nil*	1.44	2.55	14	255	4
Dinoflagellata	Dinophyceae	2.47	1.64	*nil*	274	0	0
Diatomea	*Leptocylindrus*	1.68	0.33	2.45	186	241	0

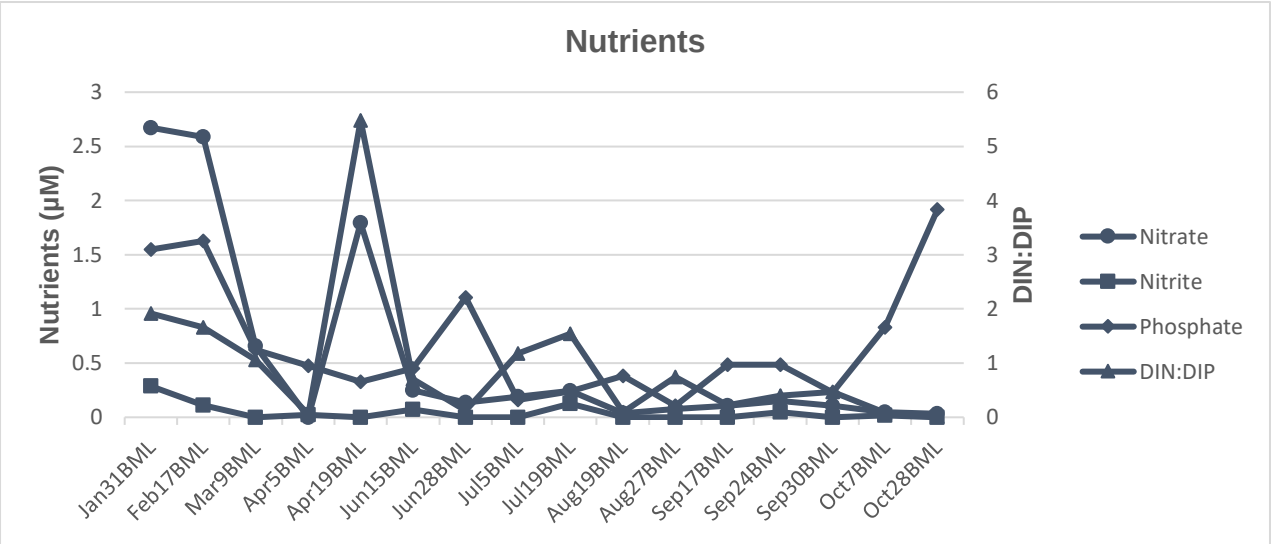
828

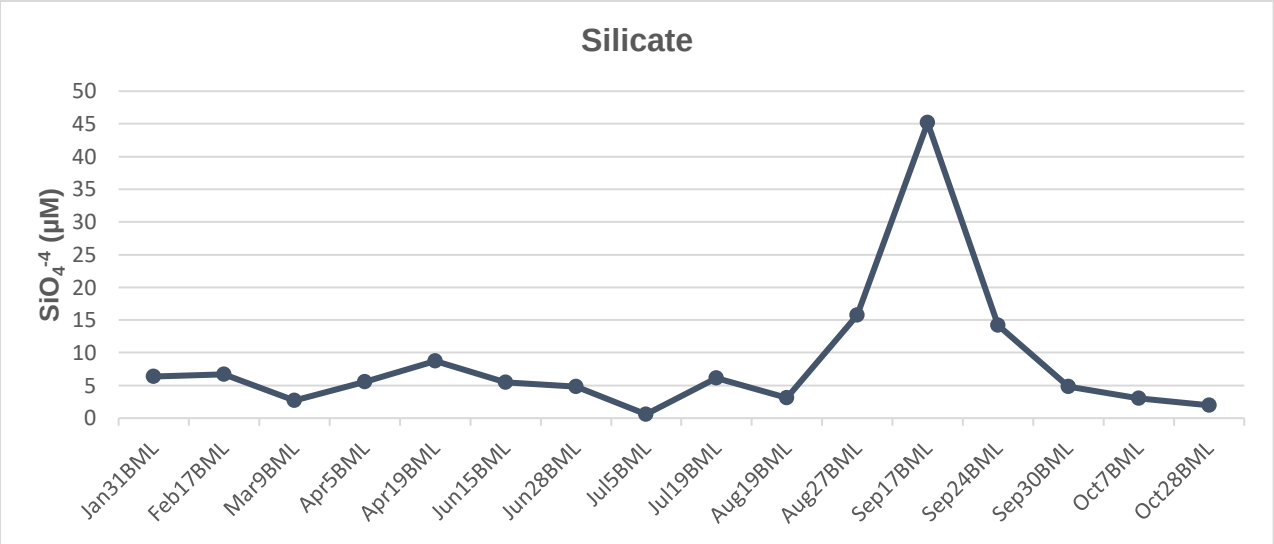
SUPPLEMENTARY FIGURES



Supplementary Fig. S1. Map of Bolinao, Pangasinan, Philippines. Red dot represents the sampling station. Arrows represent wind direction during each monsoon season.



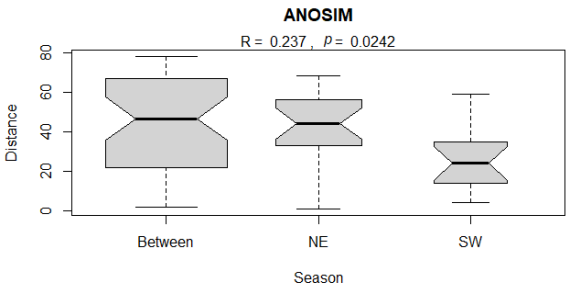


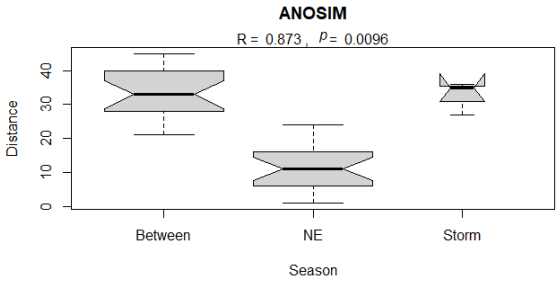


833

834 **Supplementary Fig. S2.** Line graphs of the measured environmental parameters showing
835 their trends over time.

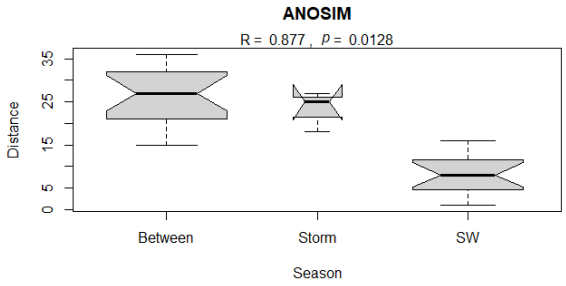
836 **A**



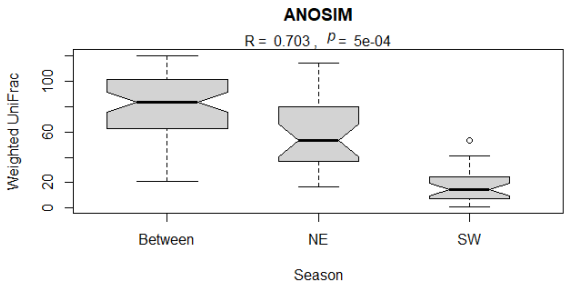


837

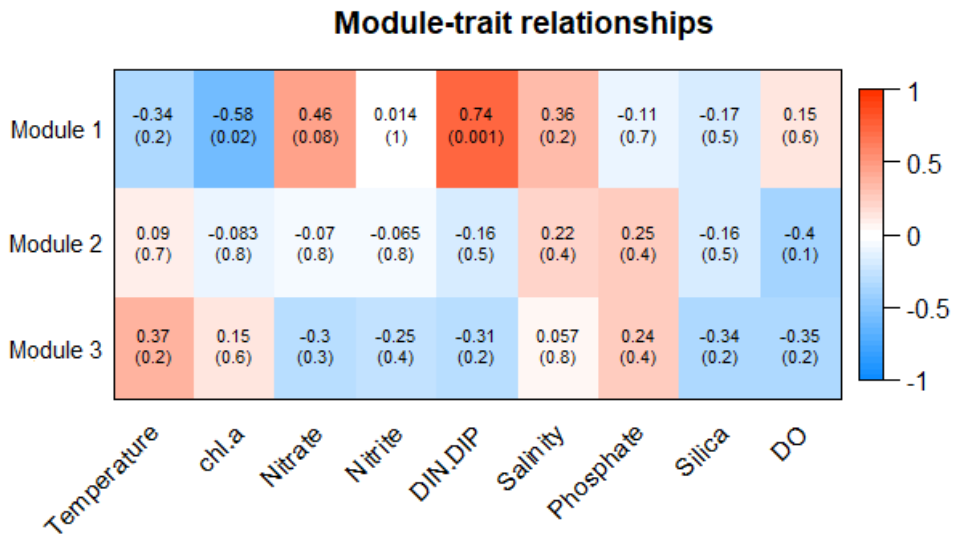
838



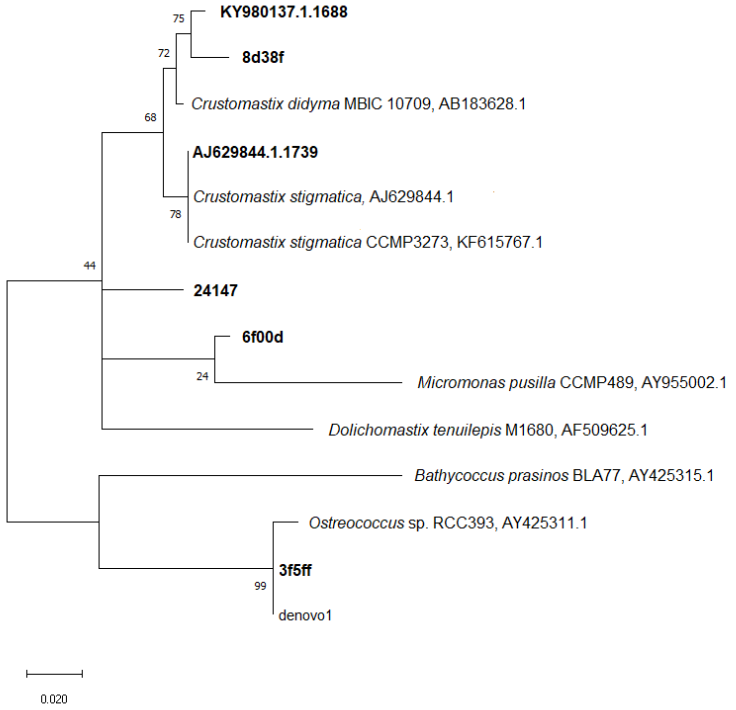
B



Supplementary Fig. S3. Analysis of similarity (ANOSIM) between clusters based on (A) log values of measured environmental parameters and (B) normalized abundance of all OTUs.

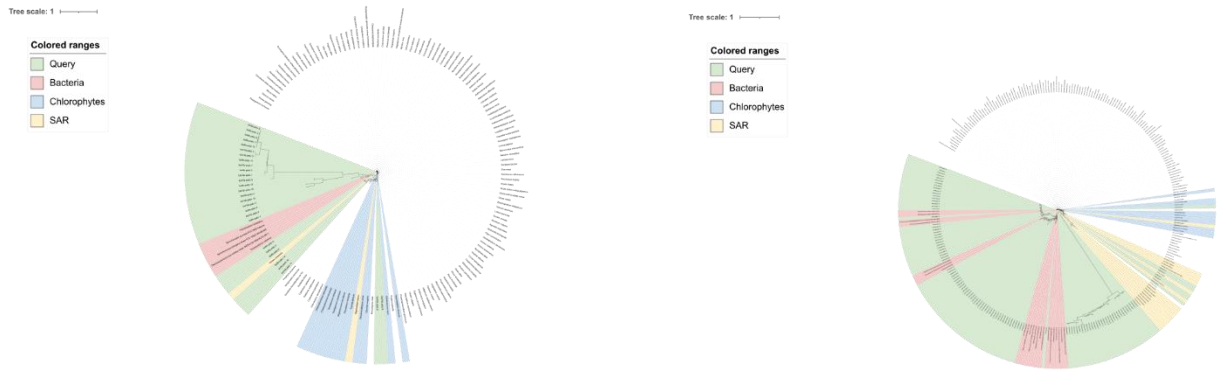


Supplementary Fig. S4. Correlation of each module eigengene to every environmental variable. The y-axis corresponds to the modules while the x-axis corresponds to the environmental variables. The values are the Spearman's rank correlation coefficient (ρ) while the values in the parentheses indicate the respective *p*-values. The colors in the matrix indicate direction and magnitude of each coefficient ρ as given by the scale.

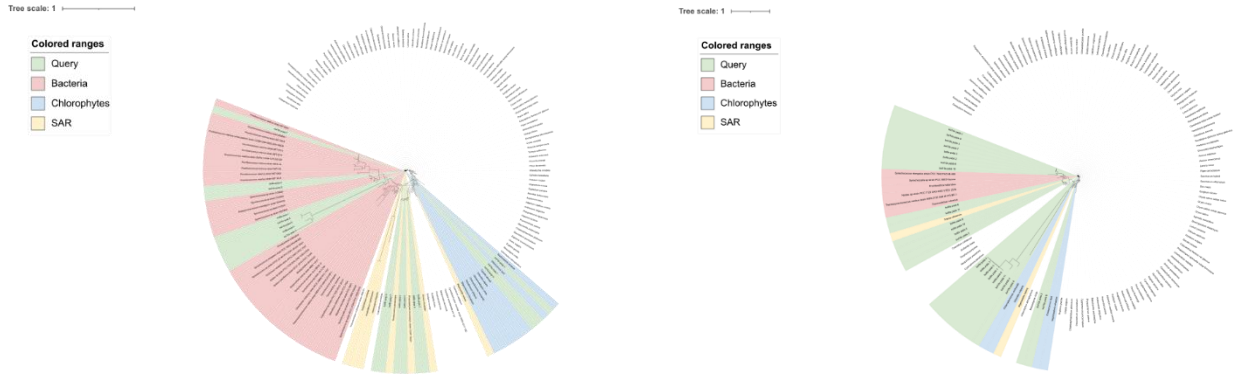


Supplementary Fig. S5. Maximum likelihood analysis of chlorophyte sequences retrieved from our study (shown in boldface) as well as reference sequences from Viprey et al., 2008 and Monier et al., 2016 using MEGA v10.2.6. The most abundant *Ostreococcus* sequence in a similar tropical coastal site reported by dela Peña et al., 2021 (shown as denovo1) was also included in the analysis.

A **B**



C **D**



Supplementary Fig. S6. Phylogenetic trees constructed from the predicted protein sequences [(A) *psbC*, (B) *psbA*, (C) *psaA*, (D) *psaB*] as well as reference sequences.

ACKNOWLEDGEMENTS

The authors thank the University of the Philippines (UP) Office of the Vice President for Academic Affairs' Balik-PhD Grant as well as the UP MSI In-house Grant for funding this study. Also, to the members of the Microbial Oceanography Laboratory for sample collection, pre-processing and extraction, and bioinformatics assistance. We also thank the administration and the whole staff of the Bolinao Marine Laboratory (BML) for their assistance during sampling and fieldwork. We also acknowledge the Continuous Operational and Outcomes-based Partnership for Excellence in Research and Academic Training Enhancement (COOPERATE) grant from UP Office of International Linkages (UP OIL) that allowed us to visit the research laboratory of our co-author Dr. Laure Guillou in Roscoff, France. Lastly, we thank the Department of Science and Technology – Philippine Atmospheric, Geophysical and Astronomical Services Administration (DOST-PAGASA) for providing climatological data, particularly rainfall and wind speed data.

LITERATURE CITED

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Albelda, R., Purganan, D. J., Gomez, N. C., Narvarte, B. C., Calalang, P. C., Genovia, T. G., Gernato, E. G., Bondoc, K. G., San Diego-McGlone, M. L., & Onda, D. F. (2019). Summer phytoplankton community structure and distribution in a mariculture-affected coastal environment. *Philippine Science Letters*, *12*(2), 157–166.

Amin, R. M., Idris, M. S., Mudiman, N. A., Azmi, N. H. M., & Siang, H. L. (2021). Spatial distribution of picophytoplankton in southeastern coast of peninsular malaysia using flow cytometry. *Pertanika Journal of Science and Technology*, *29*(3), 2103–2123. https://doi.org/10.47836/pjst.29.3.18

Azanza, R. V., Fukuyo, Y., Yap, L. G. & Takayama, H. (2004). *Prorocentrum minimum* bloom and its possible link to a massive fish kill in Bolinao, Pangasinan, Northern Philippines. *Harmful Algae*, *4*, 519-524. https://doi.org/10.1016/j.hal.2004.08.006

Azanza, R. v, & Benico, G. (2013). Toxic Alexandrium blooms in Fish Farming Sites in Bolinao, Pangasinan Harmful Algal Bloom Mitigation in the Tropics View project Sustainable Aquaculture Water Treatment View project. *Journal of Environmental Science and Management*, 44–49. https://www.researchgate.net/publication/259867455

Berdjeb, L., Parada, A., Needham, D. M., & Fuhrman, J. A. (2018). Short-term dynamics and interactions of marine protist communities during the spring-summer transition. *ISME Journal*, *12*(8), 1907–1917. https://doi.org/10.1038/s41396-018-0097-x

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., ... & Schneider, M. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic acids research, 31(1),* 365-370.

Bolger, A. M., Lohse, M. & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30(15):2114-2120.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. In *Nature Biotechnology* (Vol. 37, Issue 8, pp. 852–857). Nature Publishing Group. https://doi.org/10.1038/s41587-019-0209-9

Bryant, J. A., Aylward, F. O., Eppley, J. M., Karl, D. M., Church, M. J., & DeLong, E. F. (2016). Wind and sunlight shape microbial diversity in surface waters of the North Pacific Subtropical Gyre. *ISME Journal*, *10*(6), 1308–1322. https://doi.org/10.1038/ismej.2015.221

Buchfink, B., Reuter, K., & Drost, H. G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature methods*, *18*(4), 366-368.

Canini, N. D., Metillo, E. B., & Azanza, R. v. (2013). Monsoon-influenced phytoplankton community structure in a Philippine mangrove estuary. *Tropical Ecology*, *54*(3), 331–343. www.tropecol.com

Cardenas, M. B., Zamora, P. B., Siringan, F. P., Lapus, M. R., Rodolfo, R. S., Jacinto, G. S., San Diego-Mcglone, M. L., Villanoy, C. L., Cabrera, O., & Senal, M. I. (2010). Linking regional sources and pathways for submarine groundwater discharge at a reef by electrical resistivity tomography, 222Rn, and salinity measurements. *Geophysical Research Letters*, *37*(16). https://doi.org/10.1029/2010GL044066

Cayanan, E. O., Chen, T. C., Argete, J. C., Yen, M. C., & Nilo, P. D. (2011). The effect of tropical cyclones on southwest monsoon rainfall in the philippines. *Journal of the Meteorological Society of Japan*, *89*(A), 123–139. https://doi.org/10.2151/jmsj.2011-A08

Chambouvet, A., Morin, P., Marie, D., & Guillou, L. (2008). Control of toxic marine dinoflagellate blooms by serial parasitic killers. *Science*, *322*(5905), 1254–1257. https://doi.org/10.1126/science.1163971

Chen, J., Zhang, X., & Yang, L. (2022). GUniFrac: Generalized UniFrac Distances, Distance-Based Multivariate Methods and Feature-Based Univariate Methods for Microbiome Data Analysis. R package version 1.5. https://CRAN.R-project.org/package=GUniFrac

Chen, T., Liu, Y., Song, S., & Li, C. (2018). Characterization of the parasitic dinoflagellate *Amoebophrya* sp. infecting *Akashiwo sanguinea* in coastal waters of China. *J. Eukaryot. Microbiol.*, 0:1-10.

Coats, D. W., Bachvaroff, T. R., & Delwiche, C. F. (2012). Revision of the family Duboscquellidae with description of Eduboscquella crenulata n. gen., n. sp.

(Dinoflagellata, Syndinea), an intracellular parasite of the ciliate Favella panamensis Kofoid & Campbell. *Journal of Eukaryotic Microbiology*, *59*(1), 1–11. https://doi.org/10.1111/j.1550-7408.2011.00588.x

Coats, D. W., & Park, M.G. (2002). Parasitism of photosynthetic dinoflagellates by three strains of *Amoebophrya* (Dinophyta): parasite survival, infectivity, generation time and host specificity. *J. Phycol.*, 38:520–528.

Collado-Fabbri, S., Vaulot, D., & Ulloa, O. (2011). Structure and seasonal dynamics of the eukaryotic picophytoplankton community in a wind-driven coastal upwelling ecosystem. *Limnology and Oceanography*, *56*(6), 2334–2346. https://doi.org/10.4319/lo.2011.56.6.2334

Comeau, A. M., Li, W. K. W., Tremblay, J. É., Carmack, E. C., & Lovejoy, C. (2011). Arctic ocean microbial community structure before and after the 2007 record sea ice minimum. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027492

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Syst* 1695.

Debroas, D., Domaizon, I., Humbert, J. F., Jardillier, L., Lepére, C., Oudart, A., & Taib, N. (2017). Overview of freshwater microbial eukaryotes diversity: A first analysis of publicly available metabarcoding data. *FEMS Microbiology Ecology*, *93*(4). https://doi.org/10.1093/femsec/fix023

dela Peña, L. B. R. O., Tejada, A. J. P., Quijano, J. B., Alonzo, K. H., Gernato, E. G., Caril, A., dela Cruz, M. A. M., & Onda, D. F. L. (2021). Diversity of Marine Eukaryotic Picophytoplankton Communities with Emphasis on Mamiellophyceae in Northwestern Philippines. In *Philippine Journal of Science* (Vol. 150, Issue 1).

Escobar, Ma. T. L., Sotto, L. P. A., Jacinto, G. S., Benico, G. A., San Diego-McGlone, M. L., & Azanza, R. v. (2013). *Eutrophic conditions during the 2010 fish kill in Bolinao and Anda, Pangasinan*. https://www.researchgate.net/publication/259867727

Faust, K. (2021). Open challenges for microbial network construction and analysis. In *ISME Journal* (Vol. 15, Issue 11, pp. 3111–3118). Springer Nature. https://doi.org/10.1038/s41396-021-01027-4

Faust, K., & Raes, J. (2012). Microbial interactions: From networks to models. In *Nature Reviews Microbiology* (Vol. 10, Issue 8, pp. 538–550). https://doi.org/10.1038/nrmicro2832

Ferrera, C. M., Watanabe, A., Miyajima, T., San Diego-McGlone, M. L., Morimoto, N., Umezawa, Y., Herrera, E., Tsuchiya, T., Yoshikai, M., & Nadaoka, K. (2016). Phosphorus as a driver of nitrogen limitation and sustained eutrophic conditions in Bolinao and Anda, Philippines, a mariculture-impacted tropical coastal area. *Marine Pollution Bulletin*, *105*(1), 237–248. https://doi.org/10.1016/j.marpolbul.2016.02.025

Foflonker, F., Ananyev, G., Qiu, H., Morrison, A., Palenik, B., Dismukes, G. C., & Bhattacharya, D. (2016). The unexpected extremophile: Tolerance to fluctuating salinity in the green alga Picochlorum. *Algal Research*, *16*, 465–472. https://doi.org/10.1016/j.algal.2016.04.003

Fox, J. F. (1979). Intermediate-disturbance hypothesis. *Science*, *204*(4399), 1344–1345. https://doi.org/10.1126/science.204.4399.1344

Fraisse, S., Bormans, M., & Lagadeuc, Y. (2015). Turbulence effects on phytoplankton morphofunctional traits selection. *Limnology and Oceanography*, *60*(3), 872–884. https://doi.org/10.1002/lno.10066

Galand, P. E., Lucas, S., Fagervold, S. K., Peru, E., Pruski, A. M., Vétion, G., Dupuy, C., & Guizien, K. (2016). Disturbance increases microbial community diversity and production in marine sediments. *Frontiers in Microbiology*, *7*(DEC). https://doi.org/10.3389/fmicb.2016.01950

Gittings, J. A., Raitsos, D. E., Krokos, G., & Hoteit, I. (2018). Impacts of warming on phytoplankton abundance and phenology in a typical tropical marine ecosystem. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-20560-5

Goswami, R. K., Agrawal, K., & Verma, P. (2022). Phycoremediation of nitrogen and phosphate from wastewater using Picochlorum sp.: A tenable approach. In *Journal of Basic Microbiology* (Vol. 62, Issues 3–4, pp. 279–295). John Wiley and Sons Inc. https://doi.org/10.1002/jobm.202100277

995 Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R.,
996 Scanlan, D. J., & Worden, A. Z. (2008). Widespread occurrence and genetic diversity of
997 marine parasitoids belonging to Syndiniales (Alveolata). *Environmental Microbiology*,
998 *10*(12), 3349–3365. https://doi.org/10.1111/j.1462-2920.2008.01731.x

999 Hockin, N. L., Mock, T., Mulholland, F., Kopriva, S., & Malin, G. (2012). The response
1000 of diatom central carbon metabolism to nitrogen starvation is different from that of green
1001 algae and higher plants. *Plant Physiology*, *158*(1), 299–312.
1002 https://doi.org/10.1104/pp.111.184333

1003 Horiguchi, T. (2015). Diversity and phylogeny of marine parasitic dinoflagellates. In
1004 *Marine Protists: Diversity and Dynamics* (pp. 397–419). Springer Japan.
1005 https://doi.org/10.1007/978-4-431-55130-0_16

1006 Hopes, A., & Mock, T. (2015). Evolution of Microalgae and Their Adaptations in
1007 Different Marine Ecosystems. In *eLS* (pp. 1–9). Wiley.
1008 https://doi.org/10.1002/9780470015902.a0023744

1009 Jeong, H. J., du Yoo, Y., Kim, J. S., Seong, K. A., Kang, N. S., & Kim, T. H. (2010).
1010 Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates
1011 in marine planktonic food webs. *Ocean Science Journal*, *45*(2), 65–91.
1012 https://doi.org/10.1007/s12601-010-0007-2

1013 Jones, A. C., Hambright, K. D., & Caron, D. A. (2018). Ecological Patterns Among
1014 Bacteria and Microbial Eukaryotes Derived from Network Analyses in a Low-Salinity
1015 Lake. *Microbial Ecology*, *75*(4), 917–929. https://doi.org/10.1007/s00248-017-1087-7

1016 Käse, L., Metfies, K., Kraberg, A. C., Neuhaus, S., Meunier, C. L., Wiltshire, K. H., &
1017 Boersma, M. (2021). Metabarcoding analysis suggests that flexible food web interactions
1018 in the eukaryotic plankton community are more common than specific predator-prey
1019 relationships at Helgoland Roads, North Sea. *ICES Journal of Marine Science*, *78*(9),
1020 3372–3386. https://doi.org/10.1093/icesjms/fsab058

1021 Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package
1022 version 0.4.0. https://CRAN.R-project.org/package=ggpubr

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772-780.

Kembel, S., Cowan, P., Helmus, M., Cornwell, W., Morlon, H., Ackerly, D., Blomberg, S., Webb, C. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics, 26*, 1463–1464.

Kim, S. (2006). Patterns in host range for two strains of *Amoebophrya* (Dinophyta) infecting thecate dinoflagellates: *Amoebophrya* spp. ex *Alexandrium affine* and ex *Gonyaulax polygramma*. *J. Phycol*., 42:1170–1173.

Kurtay, G., Prevost, H. J., & Stauffer, B. A. (2021). Pico- and nanoplankton communities on a near to offshore transect along the continental shelf of the northwestern Gulf of Mexico in the aftermath of Hurricane Harvey. *Limnology and Oceanography*, *66*(7), 2779–2796. https://doi.org/10.1002/lno.11788

Lahti, L., Shetty, S. et al. (2017). Tools for microbiome analysis in R. Microbiome package version 1.17.42. *Bioconductor*. URL: http://microbiome.github.com/microbiome.

Lambert, S., Tragin, M., Lozano, J. C., Ghiglione, J. F., Vaulot, D., Bouget, F. Y., & Galand, P. E. (2019). Rhythmicity of coastal marine picoeukaryotes, bacteria and archaea despite irregular environmental perturbations. *ISME Journal*, *13*(2), 388–401. https://doi.org/10.1038/s41396-018-0281-z

Langfelder P, & Horvath S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559.

Levy Karin, E., Mirdita, M., & Söding, J. (2020). MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. *Microbiome*, *8*(1), 1-15.

Li, W. K. W., McLaughlin, F. A., Lovejoy, C., & Carmack, E. C. (2009). Smallest algae thrive as the arctic ocean freshens. In *Science* (Vol. 326, Issue 5952, p. 539). https://doi.org/10.1126/science.1179798

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674-1676.

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S., Ignacio-Espinosa, J. C., Roux, S., Vincent, F., Bittner, L., Darzi, Y., Wang, J., Audic, S., Berline, L., Bontempi, G., Cabello, A. M., Coppola, L., Cornejo-Castillo, F. M., … Velayoudon, D. (2015). Determinants of community structure in the global plankton interactome. *Science*, *348*(6237). https://doi.org/10.1126/science.1262073

Liu, Q., Zhao, Q., Jiang, Y., Li, Y., Zhang, C., Li, X., Yu, X., Huang, L., Wang, M., Yang, G., Chen, H., & Tian, J. (2021). Diversity and co-occurrence networks of picoeukaryotes as a tool for indicating underlying environmental heterogeneity in the Western Pacific Ocean. *Marine Environmental Research*, *170*. https://doi.org/10.1016/j.marenvres.2021.105376

Logares, R., Deutschmann, I. M., Junger, P. C., Giner, C. R., Krabberød, A. K., Schmidt, T. S. B., Rubinat-Ripoll, L., Mestre, M., Salazar, G., Ruiz-González, C., Sebastián, M., de Vargas, C., Acinas, S. G., Duarte, C. M., Gasol, J. M., & Massana, R. (2020). Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome*, *8*(1). https://doi.org/10.1186/s40168-020-00827-8

Logares, R., Shalchian-Tabrizi, K., Boltovskoy, A., & Rengefors, K. (2007). Extensive dinoflagellate phylogenies indicate infrequent marine-freshwater transitions. *Molecular Phylogenetics and Evolution*, *45*(3), 887–903. https://doi.org/10.1016/j.ympev.2007.08.005

Lovejoy, C., Comeau, A., Thaler, M. (2016). Curated reference database of SSU rRNA for Northern marine and freshwater communities of Archaea, Bacteria and microbial eukaryotes, v. 1.1 (2002-2008). *Nordicana* D23. doi: 10.5885/45409XD-79A199B76BCC4110.

Margalef, R. (1978). *Life-forms of phytoplankton as survival alternatives in an unstable environment*.

Massana, R., Balagué, V., Guillou, L., & Pedrós-Alió, C. (2004). Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiology Ecology*, *50*(3), 231–243. https://doi.org/10.1016/j.femsec.2004.07.001

Matsumoto, J., Olaguera, L. M. P., Nguyen-Le, D., Kubota, H., & Villafuerte, M. Q. (2020). Climatological seasonal changes of wind and rainfall in the Philippines. *International Journal of Climatology*, *40*(11), 4843–4857. https://doi.org/10.1002/joc.6492

McDonald, S. M., Plant, J. N., & Worden, A. Z. (2010). The mixed lineage nature of nitrogen transport and assimilation in marine eukaryotic phytoplankton: A case study of Micromonas. *Molecular Biology and Evolution*, *27*(10), 2268–2283. https://doi.org/10.1093/molbev/msq113

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature communications*, *7*(1), 1-9.

Mitbavkar, S., Patil, J. S., & Rajaneesh, K. M. (2015). Picophytoplankton as Tracers of Environmental Forcing in a Tropical Monsoonal Bay. *Microbial Ecology*, *70*(3), 659–676. https://doi.org/10.1007/s00248-015-0599-2

Mitbavkar, S., & D'souza, S. (2021). Eco-Biology of Photosynthetic Picoeukaryotes in a Monsoon Influenced Tropical Bay: A Flow Cytometric and Chemotaxonomic Approach. https://doi.org/10.21203/rs.3.rs-860143/v1

Nakayama, T., Kawachi, M., & Inouye, I. (2000). Taxonomy and the phylogenetic position of a new prasinophycean alga, Crustomastix didyma gen. & sp. nov. (Chlorophyta). *Phycologia*, *39*(4), 337–348.

Not, F., del Campo, J., Balagué, V., de Vargas, C., & Massana, R. (2009). New insights into the diversity of marine picoeukaryotes. *PLoS ONE*, *4*(9). https://doi.org/10.1371/journal.pone.0007143

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R.B., Solymos, P., Stevens, H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H. B. A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.

O., Lahti, L., McGlinn, D., Ouellette, M., Cunha, E. R., Smith, T., Stier, A., Ter Braak, C. J. F., & Weedon, J. (2022). vegan: Community Ecology Package. R package version 2.6-2. https://CRAN.R-project.org/package=vegan

Onda, D. F. L., Medrinal, E., Comeau, A. M., Thaler, M., Babin, M., & Lovejoy, C. (2017). Seasonal and interannual changes in ciliate and dinoflagellate species assemblages in the Arctic Ocean (Amundsen Gulf, Beaufort Sea, Canada). *Frontiers in Marine Science*, *4*(FEB). https://doi.org/10.3389/FMARS.2017.00016

Parks, D. H., Tyson, G. W., Hugenholtz, P., & Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics*, *30*(21), 3123-3124.

Parsons, T. R., Maita, Y., Lalli, C. M. I, I. (1984). A manual of chemical and biological methods for seawater analysis. *Pergamon Press, Oxford, UK* 1, 73.

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one*, *5*(3), e9490.

Qiu, D., Zhong, Y., Chen, Y., Tan, Y., Song, X., & Huang, L. (2019). Short-Term Phytoplankton Dynamics During Typhoon Season in and Near the Pearl River Estuary, South China Sea. *Journal of Geophysical Research: Biogeosciences*, *124*(2), 274–292. https://doi.org/10.1029/2018JG004672

Rajaneesh, K. M., Mitbavkar, S., & Anil, A. C. (2018). Dynamics of size-fractionated phytoplankton biomass in a monsoonal estuary: Patterns and drivers for seasonal and spatial variability. *Estuarine, Coastal and Shelf Science*, *207*, 325–337. https://doi.org/10.1016/j.ecss.2018.04.026

Ramos, R. D., Goodkin, N. F., Siringan, F. P., & Hughen, K. A. (2017). Diploastrea heliopora Sr/Ca and δ18O records from northeast Luzon, Philippines: An assessment of interspecies coral proxy calibrations and climate controls of sea surface temperature and salinity. *Paleoceanography*, *32*(4), 424–438. https://doi.org/10.1002/2017PA003098

Ravelo, S. F., Yap-Dejeto, L. G., Silaras, M. L. S., Amparado, M. L. L., Ocampo, J. A., Abria, E. G., & Albina, M. B. (2022). A Snapshot on the Distribution of Coastal Phytoplankton Communities in Five HAB-Affected Bays in Eastern Visayas, Philippines. *Frontiers in Marine Science*, *9*. https://doi.org/10.3389/fmars.2022.730518

Rideout, J. R., He, Y., Navas-Molina, J. A., Walters, W. A., Ursell, L. K., Gibbons, S. M., Chase, J., McDonald, D., Gonzalez, A., Robbins-Pianka, A., Clemente, J. C., Gilbert, J. A., Huse, S. M., Zhou, H. W., Knight, R., & Gregory Caporaso, J. (2014). Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ*, *2014*(1). https://doi.org/10.7717/peerj.545

Rivera, P. C. (1997). *HYDRODYNAMICS, SEDIMENT TRANSPORT AND LIGHT EXTINCTION OFF CAPE BOLINAO, PHILIPPINES*. http://www.balkema.nl

Rocke, E., Cheung, S., Gebe, Z., Dames, N. R., Liu, H., & Moloney, C. L. (2020). Marine Microbial Community Composition During the Upwelling Season in the Southern Benguela. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00255

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, *18*(4), e2584. https://doi.org/10.7717/peerj.2584. PMID: 27781170; PMCID: PMC5075697.

Romari, K., & Vaulot, D. (2004). Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. In *Limnol. Oceanogr* (Vol. 49, Issue 3). http://www.aslo.org/lo/toc/vol49/issue3/0784a1.pdf.

Sanders, R. W., & Gast, R. J. (2012). Bacterivory by phototrophic picoplankton and nanoplankton in Arctic waters. *FEMS Microbiology Ecology*, *82*(2), 242–253. https://doi.org/10.1111/j.1574-6941.2011.01253.x

Sanz-Luque, E., Chamizo-Ampudia, A., Llamas, A., Galvan, A., & Fernandez, E. (2015). Understanding nitrate assimilation and its regulation in microalgae. In *Frontiers in Plant Science* (Vol. 6, Issue OCTOBER). Frontiers Research Foundation. https://doi.org/10.3389/fpls.2015.00899

Silva, G. G. Z., Green, K. T., Dutilh, B. E., & Edwards, R. A. (2016). SUPER-FOCUS: a tool for agile functional analysis of shotgun metagenomic data. *Bioinformatics*, *32*(3), 354-361.

Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Acinas, S. G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., de Vargas, C., Follows, M., Gorsky, G., Grimsley,

N., Guidi, L., Hingamp, P., Iudicone, D., Jaillon, O., Kandels, S., … de Vargas, C. (2020). Tara Oceans: towards global ocean ecosystems biology. In *Nature Reviews Microbiology* (Vol. 18, Issue 8, pp. 428–445). Nature Research. https://doi.org/10.1038/s41579-020-0364-5

Suzuki', A., & Rothstein', S. (1997). Structure and regulation of ferredoxin-dependent glutamase synthase from Arabidopsis thaliana Cloning of cDNA, expression in different tissues of wild-type and gZtS mutant strains, and light induction. In *Eur. J. Biochem* (Vol. 243).

Takabayashi, M., Wilkerson, F. P., & Robertson, D. (2005). Response of glutamine synthetase gene transcription and enzyme activity to external nitrogen sources in the diatom Skeletonema costatum (Bacillariophyceae). *Journal of Phycology*, *41*(1), 84–94. https://doi.org/10.1111/j.1529-8817.2005.04115.x

Thomas, W. H., & Gibson, C. H. (1990). Effects of small-scale turbulence on microalgae. *Journal of Applied Phycology*, *2*(1), 71–77. https://doi.org/10.1007/BF02179771

Trigueros, J. M., & Orive, E. (2001). Seasonal variations of diatoms and dinoflagellates in a shallow, temperate estuary, with emphasis on neritic assemblages. *Hydrobiologia*, *444*, 119–133. https://doi.org/10.1023/A:1017563031810

UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. Nucleic acids research, 47(D1), D506-D515.

Vaulot, D., Eikrem, W., Viprey, M., & Moreau, H. (2008). The diversity of small eukaryotic phytoplankton (≤3 μm) in marine ecosystems. In *FEMS Microbiology Reviews* (Vol. 32, Issue 5, pp. 795–820). https://doi.org/10.1111/j.1574-6976.2008.00121.x

Vincent, F., & Bowler, C. (2020). Diatoms Are Selective Segregators in Global Ocean Planktonic Communities. *MSystems*, *5*(1). https://doi.org/10.1128/msystems.00444-19

Viprey, M., Guillou, L., Ferréol, M., & Vaulot, D. (2008). Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environmental Microbiology*, *10*(7), 1804–1822. https://doi.org/10.1111/j.1462-2920.2008.01602.x

Wang, Y., Li, G., Shi, F., Dong, J., Gentekaki, E., Zou, S., Zhu, P., Zhang, X., & Gong, J. (2020). Taxonomic Diversity of Pico-/Nanoeukaryotes Is Related to Dissolved Oxygen and Productivity, but Functional Composition Is Shaped by Limiting Nutrients in Eutrophic Coastal Oceans. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.601037

Warden, J. G., Casaburi, G., Omelon, C. R., Bennett, P. C., Breecker, D. O., & Foster, J. S. (2016). Characterization of microbial mat microbiomes in the modern thrombolite ecosystem of lake clifton, western australia using shotgun metagenomics. *Frontiers in Microbiology*, *7*(JUL). https://doi.org/10.3389/fmicb.2016.01064

Worden, A. Z., Nolan, J. K., & Palenik, B. (2004). Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. In *Limnol. Oceanogr* (Vol. 49, Issue 1).

Worden, A. Z., & Not, F. (2008). Ecology and Diversity of Picoeukaryotes. In D. L. Kirchman (Ed.), *Microbial Ecology of the Oceans* (2nd Edition, pp. 159–205). John Wiley & Sons, Inc.

Wyrtki, K. (1961). *Physical oceanography of the Southeast Asian waters.* Scripps Institution of Oceanography, University of California, La Jolla.

Xie, N., Hunt, D. E., Johnson, Z. I., He, Y., & Wang, G. (2021). Annual partitioning patterns of Labyrinthulomycetes protists reveal their multifaceted role in marine microbial food webs. *Applied and Environmental Microbiology*, *87*(2). https://doi.org/10.1128/AEM.01652-20

Yap, L. G., Azanza, R. v., & Talaue-McManus, L. (2004). The community composition and production of phytoplankton in fish pens of Cape Bolinao, Pangasinan: A field study. *Marine Pollution Bulletin*, *49*(9–10), 819–832. https://doi.org/10.1016/j.marpolbul.2004.06.030

Yergeau, E., Michel, C., Tremblay, J., Niemi, A., King, T. L., Wyglinski, J., Lee, K., & Greer, C. W. (2017). Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Scientific Reports*, *7*. https://doi.org/10.1038/srep42242

1224 Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate
1225 Illumina Paired-End reAd mergeR. *Bioinformatics*, *30*(5), 614-620.

1226 Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression
1227 network analysis. *Stat. Appl. Genet. Mol. Biol.* 4(1):Article 17.

1228 Zhou, J., Qin, B., & Han, X. (2016). Effects of the magnitude and persistence of
1229 turbulence on phytoplankton in Lake Taihu during a summer cyanobacterial bloom.
1230 *Aquatic Ecology*, *50*(2), 197–208. https://doi.org/10.1007/s10452-016-9568-1