# Dynamics of finely resolved, abundant symbiotic marine plankton and other interacting microbes via automated high-frequency sampling

David M. Needham[1,3], Erin B. Fichot[1], Ellice Wang[1], Lyria Berdjeb[1], Jacob A. Cram[1], Cedric G. Fichot[2], Jed A. Fuhrman[1]

[1]Department of Biological Sciences, University of Southern California

[2]Department of Earth and Environment, Boston University

Contacts

David M. Needham, dmneedha@gmail.com

Jed A. Fuhrman, fuhrman@usc.edu

Conflicts of Interest: The authors declare no conflict of interest.

**Abstract**

Short time-scale observations are valuable for understanding microbial ecological processes. We assessed dynamics in relative abundance and potential activities by sequencing the small sub-unit ribosomal RNA gene (rDNA) and rRNA molecules (rRNA), of Bacteria, Archaea and Eukaryotes once to twice-daily between March and May in the surface ocean off Catalina Island, California. Typically *Ostreococcus, Braarudosphaera, Teleaulax, and Synechococcus* dominated phytoplankton sequences while SAR11, *Sulfitobacter* and *Fluvicola* dominated non-phytoplankton prokaryotes. We observed short-lived increases of diatoms, mostly *Pseudo-nitzschia* and *Chaetoceros,* with quickly-responding prokaryotes including Flavobacteria (*Polaribacter*, *Formosa*), *Roseovarius,* and Euryarchaea (MGII), which were the exact sequence variants we also observed as temporally most-abundant in another diatom bloom at a nearby location, 3 years prior. We observed positive correlations representing known interactions among abundant taxa in chloroplastic rRNA sequences, demonstrating the ecological relevance of such interactions and their influence on the environment: 1) The kleptochlorplastidic ciliate *Myrionecta* 18S and *Teleaulax* chloroplasts (16S) were correlated (Spearman r =0.83) yet uncorrelated to *Teleaulax* nuclear 18S, nor any other taxon and 2) the photosynthetic prymnesiophyte *Braarudosphaera bigelowii* and 2 strains of diazotrophic cyanobacterium UCYN-A were correlated and each was correlated to multiple other taxa, including *Braarudosphaera* to a Verrucomicrobium and a Dictyophyte phytoplankter (all $r > 0.8$). We also report strong correlations ($r > 0.7$) between ciliates and bacteria and phytoplankton, possibly representing mutually beneficial interactions. These data reiterate the utility of high-frequency time-series to show rapid microbial reactions to stimuli, and provide new information about *in-situ* dynamics of previously recognized and hypothesized interactions.

## Introduction

Natural marine microbial communities, consisting of both prokaryotes and eukaryotes, are diverse and dynamic. The interactions between microbial species and their environment and between microbial species dictate how energy and nutrients flow through the ocean (Fuhrman et al. 2015). Microbial communities are known to be seasonally variable (Gilbert et al. 2012; Fuhrman et al. 2006; Cram et al. 2014) and can show rapid responses to environmental variation, such as stratification and pulses of nutrients (Teeling et al. 2012; Needham & Fuhrman 2016). Time-series studies with sampling at various temporal-scales contribute to our understanding of different processes in the sea, such as seasonal variation or climate change. Daily or diel-scale high-resolution time-series are particularly useful for observing ecological responses to short-term perturbations, such as phytoplankton blooms and interactions of organisms, because whole microbial community turnover time is on the scale of a few days (Fuhrman & Azam 1982; Cram et al. 2014). During phytoplankton blooms, microbial communities can vary in pronounced, succession-like ways with dominant taxa shifting quickly (Teeling et al. 2016, 2012), even on time scales of one to several days (Needham and Fuhrman 2016; Needham et al. 2017).

Ecological interactions between microorganisms are of great importance in the ocean (Worden et al. 2015). Such interactions can be general, such as lineages of bacteria that repeatedly respond to increases in phytoplankton biomass and the organic matter produced by blooms (Buchan et al. 2014). However, many interactions appear to be species specific, including direct microbe-microbe interactions and can be observed at short temporal scales (Fuhrman et al. 2015). Such interactions include grazing, cross-feeding, mutualism, parasitism, symbiosis, or kleptochloroplasty (i.e., where a heterotrophic protist captures chloroplasts from another species and the chloroplast continues to function inside the grazer)(Mitra et al. 2016). Many of these interactions occur beween organisms of different domains or trophic states; e.g., , between bacteria and eukaryotes, or between phototrophs and heterotrophs. Studying all of these organisms together allows a more complete view of components in the “microbial loop” (Azam et al. 1983).

The dynamics and ecology of microbial organisms via time-series is often assessed via sequencing of the small subunit ribosomal DNA gene (rDNA) of cellular organisms, which is conserved across all three domains of life. With current sequencing outputs from the Illumina MiSeq and HiSeq platform (paired end 2x250 or 2x300 high quality reads), it is possible to confidently discriminate taxa by as little as a single base difference in this conserved gene, which can resolve taxa at the "strain" or species level (Eren et al. 2014; Callahan et al. 2016; Tikhonov et al. 2015). We have recently shown that a single rRNA gene primer set can simulataneously assess bacteria, archaea, and eukaryotic phytoplankton via their chloroplasts (with the exception of dinoflagellates due to aberrant chloroplast sequnces) via 16S (Needham & Fuhrman 2016; Parada et al. 2016). Using the same rRNA gene primer set, the full eukaryotic community can be assessed via 18S (Needham & Fuhrman 2016; Parada et al. 2016).

A complementary approach to sequencing the rDNA is reverse transcribing and sequencing of the small sub-unit of the rRNA molecule itself (rRNA) which provides the same identity information as DNA, but the number of sequences is considered a proxy for the cumulative number of ribosomes from that taxon. This approach may reveal information about the biomasses and potential activities of taxa across the full community (Campbell et al. 2011; Hunt et al. 2013; Blazewicz et al. 2013; Lankiewicz et al. 2016). The rRNA and rDNA approaches each have benefits and uncertainties. For the rDNA, while the gene copy number varies between taxa, it is consistent within individuals of a given taxon and across time. The large majority of free living planktonic marine prokaryotes have 1-2 copies per cell (Brown & Fuhrman 2005). For chloroplasts, copy number is usually between 1-2 per chloroplast, and the number of chloroplasts per cell can vary from 1 to hundreds (depending largely on cell size; Needham and Fuhrman 2016). However, for small phytoplankton most commonly found at our location, the variation is typically low (2-4 chloroplasts for common taxa) (Needham and Fuhrman 2016). The 18S of eukaryotes, on the other hand, has a larger range in copy number, from 2 to 50,000 (de Vargas et al. 2015). Thus, comparing relative abundances for these taxa via 18S

is tenuous, but the copy number relates very roughly to cellular biomass, when compared over many orders of magnitude on a log-log plot (de Vargas et al. 2015). rRNA in contrast may reflect variation of "potential activity" between and within taxa over time. However, the number of ribosomes per cell does not consistently reflect growth rate across taxa, since the relationship is very noisy and irregular between taxa, and it is not anything like a linear measure of growth rate (Blazewicz et al. 2013; Lankiewicz et al. 2016). Previous work has assessed the ratio between the rRNA and rDNA of individual taxa. This work results in an "index" that aims to examine the relative activities across taxa and describe patterns across all taxa. Such an analysis, with all its inherent complexities and complicating factors, is outside the scope of this paper.

Here we apply rRNA and rDNA sequencing to study the full cellular microbial community -- bacteria, archaea, and eukaryotes -- from seawater samples collected from the photic zone once to twice per day over about 1.5 months via an Environmental Sample Processor (ESP), which also provided continuous physical and chemical measurements. During the course of the sampling, a short-lived bloom of phytoplankton occurred, allowing us examine the dynamics and activities, before, during, and after the bloom. Additionally, we found that the members of two well-known symbioses were commonly found in high abundance during our time-series: 1. ) the ciliate *Myrionecta* and the chloroplasts of cryptophyte *Teleaulax* (Johnson et al. 2016) and 2.) the diazotrophic cyanobacterium UCYN-A and haptophyte alga *Braarudosphaera bigelowii* (Zehr et al. 2016). This allowed us to assess the *in-situ* abundances and physiological dynamics of these relationships, which provides insight into the nature of these associations. We also explore other strong co-occurrence patterns between phytoplankton, potential eukaryotic grazers, and prokaryotes to examine potential new interactions.

156 **Methods**

157 **Sampling**

158 An Environmental Sample Processor (ESP) (Scholin et al. 2009), which
159 autonomously draws seawater samples and filters them sequentially while also
160 recording depth, temperature, conductivity, and chlorophyll fluorescence was
161 deployed about 1 km offshore of Santa Catalina Island, California, USA in about 200
162 m of water (33 28.990 N, 118 30.470 W) from 13 March to 1 May, 2014. The ESP
163 was tethered to the bottom and thus sampled via Eulerian sampling. Due to tides
164 and tidal currents, the depth systematically varied slightly over the course of a day.
165 Over the first 5 days, the depth of sampling was between 5 and 10 m. The ESP
166 malfunctioned on day 6. After the instrument was restored two weeks later, samples
167 were collected between 7 m and 15 m for the remainder of the time series (Figure
168 1). One L water samples for molecular analysis were drawn once (at 10 AM, 18
169 March to 23 March) or twice per day after the interruption of sampling (10 AM, 10
170 PM, 9 April to 1 May). The samples were pre-screened with 300 μm mesh and then
171 sequentially collected on a 1 μm AE filter (Pall Gelman) and a 0.22 μm Durapore
172 filter (Millipore) with an HA backing filter. All filters were stored in RNAlater at
173 ambient seawater temperature until ESP retrieval (1 May), upon which the filters
174 were stored at -80 C until processing. The ESP recorded depth, temperature,
175 conductivity, and chlorophyll fluorescence measurements every 5 minutes.

176 **Satellite imagery**

177 Level-2 remote-sensing reflectances (*Rrs*) from *Aqua* MODIS (MODerate Resolution
178 Imaging Spectrometer) were used to produce daily maps of surface chlorophyll-*a*
179 concentrations over the 01 March 2014 - 30 April 2014 time period. About one third
180 of all the images were discarded because of cloud coverage. The surface chlorophyll-
181 *a* concentrations were derived by applying a local empirical algorithm to the
182 *Rrs*(488)/*Rrs*(547) remote-sensing reflectance ratio (Trinh et al. 2017). This local
183 empirical algorithm was parameterized specifically for MODIS using *in situ*
184 measurements made in the coastal waters off the Los Angeles, CA area (i.e., our
185 region of study).
186

**DNA/RNA extraction**

Each AE and Durapore filter was aseptically cut in half, one half for DNA extraction, the other half for RNA extraction. DNA was extracted and purified from the Durapore filters using a hot SDS extraction protocol (Fuhrman et al. 1988) and from the AE filters using a NaCl/cetyl trimethylammonium bromide (CTAB) bead-beating extraction (Countway et al. 2005). Each of these methods was modified to include lysozyme and proteinase K lysis steps (30 minutes at 37 C and 30 minutes at 50 C, respectively). Supernatant from either method was subjected to phenol/chloroform/isoamyl alcohol purification, precipitated overnight in ammonium acetate and ethanol, then resuspended in TE buffer. Each half-filter underwent two sequential lysis steps, and the extracted DNA was combined.

RNA was extracted from the Durapore and AE filter via RNeasy kit (Qiagen), as per manufacturer's instructions, including an on-column DNAse step. For the AE filters, a second DNA-removal step was performed on 10 ng of RNA with Invitrogen DNAse I, Amplification Grade (Cat. Number: 18068015).

**Reverse transcription and PCR**

RNA was reverse transcribed to cDNA using SuperScript III from Invitrogen using random hexamers, with 0.1 ng for the Durapore size fraction and all of the RNA from the Invitrogen DNAse treated AE RNA (input 10 ng). cDNA was cleaned up with 2x Ampure beads. Cleaned cDNA was then amplified for 30 cycles via PCR with 5' Hot Start master mix with 515F primer (5'- GTGYCAGCMGCCGCGGTAA -3') and 926R primer (5' – CCGYCAATTYMTTTRAGTTT -3'), which amplifies bacterial, archaeal, and chloroplast 16S, as well as eukaryotic 18S (Parada et al. 2016). We confirmed that RNA extracts were devoid of significant DNA by performing no-RT PCR reactions and observing an absence of amplification in an agarose gel. DNA from Durapore (0.5 ng) and AE (0.05 ng) was amplified 30 and 35 cycles, respectively. For the AE DNA PCR reactions, 5 extra PCR cycles and 10-fold reduced DNA template were necessary because of an inhibitory affect of the RNALater on the extracted DNA. After PCR, products were cleaned and concentrated with Ampure beads and pooled. All samples were sequenced in one MiSeq 2x300 run at University of California, Davis.

**Sequence Analysis**

All commands run during data analysis and figure generation are available via Figshare (10.6084/m9.figshare.5373916). Sequences are available via EMBL study accession number PRJEB22356. Demultiplexed samples were trimmed via a sliding window of 5, trimming where the window dropped below q20 via Trimmomatic. Sequences less than 200bp were removed. For 16S analysis, forward and reverse reads were then merged with a minimum overlap of 20, minimum merged length of 200 and maximum differences (in overlap region) of 3 using USEARCH (Edgar 2013). 18S forward and reverse reads did not overlap so this merging step retains only 16S. A separate analysis is necessary for the 18S (see below). Primers were removed from the sequences with cutadapt (Martin 2011). Chimeras were detected de novo and reference based searching with QIIME *identify_chimeric seqs.py* and with the SILVA gold database (Pruesse et al. 2007) as the reference (Caporaso et al. 2010). Merged 16S reads were then padded to make them all the same length with *o-pad-with-gaps* via the *Oligotyping* pipeline (Eren et al. 2013). Then the sequences were "decomposed" with Minimum Entropy Decomposition (MED) default settings (Eren et al. 2014). MED decomposes the sequences into types that are distinguished by as little as a single base, based on an assessment of the underlying sequence variability and positions of high variability. We recently confirmed such an approach to be suitable for our assays via custom made marine mock communities (Needham et al. 2017). We refer to these highly resolved sequences as Amplicon Sequence Variants (ASVs).

Sequencing classification was performed on representative sequences from the 16S ASVs via the SILVA (Pruesse et al. 2007), Greengenes (McDonald et al. 2012), and PhytoREF (Decelle et al. 2015) (for classification of Chloroplast 16S) using QIIME *assign_taxnomy.py* with UCLUST. Additionally we classified against the NCBI database of cultured organisms (see Needham and Fuhrman 2016) with BLASTn (Altschul et al. 1990). All classifications and representative sequences are available via Figshare (Public Project Link: https://goo.gl/nM1cwe). Often classifications vary slightly between database sources, with the NCBI matches providing information

"closest cultured relative" which may not be updated in the more curated databases since they may lack the most up-to-date sequences available (e.g., see UCYN-A below). For the 16S non-phytoplankton sequences, we generally display the SILVA and Phytoref classification of the 16S ASVs of non-phytoplankton and non-phytoplankton, respectively.

Some classifications were noted to be conflicting or lacked satisfactory resolution for a given database. The reasons vary but were most often due to sequences missing from databases or misleading annotations of sequences within a database. Therefore in the following cases, we performed further manual curation of our classification: 1.) Prasinophyte sequences were all manually curated because an abundant prasinophyte sequence was initially annotated as *Ostreococcus* and *Bathycoccus* via searches against the NCBI and Phytoref databases, respectively. By manual inspection, we found that the ASV that had the discrepancy perfectly matched an *Ostreococcus genome sequence*, and made the change throughout the manuscript (Supplementary Figure 1). 2.). We found that UCYN-A sequences are generically classified by the SILVA database as Cyanobacterial Subsection I, Family I (i.e., same groups as to *Prochlorococcus* and *Synechoccocus*). To resolve the UCYN-A sequences to their respective sub-groups we downloaded the sequences of UCYN-A1 (Zehr et al. 2008) (gi|284809060) and UCYN-A2 (Bombar et al. 2014)(gi|671395793) from NCBI and compared them to our representative ASV sequence. UCYN-A1 and UCYN-A2 are 3 bp different in the V4 to V5 amplified region we used (i.e., 99.2% similar) and each of our 2 UCYN-A ASVs matched one of each of the representative genomes at 100% (Supplementary Figure 2), thus we label them accordingly.

We split the 16S data into two datasets, the "phytoplankton" and "non-phytoplankton." "Phytoplankton" included those sequences determined by the Greengenes taxonomy to be of chloroplastic orgin and cyanobacteria. "Non-phytoplankton" included the remaining bacteria and archaeal 16S sequences. After this step, samples that did not have greater than 100 reads in a given dataset were

removed from further analysis. The low value of 100 increased sensitivity for to samples with a low number of reads but, the average was 16,576 ± 9,302 SD for non-phytoplankton and 13,426 ± 11,478 SD for phytoplankton.

The 18S amplicon sequencing products are too long to overlap given the MiSeq 2x300 forward and reverse sequencing that we used. The following steps were taken to process these data. 1.) The data were quality filtered the same as for the 16S analysis via Trimmomatic. 2.) The resulting quality-filtered forward and reverse reads were trimmed to 290 and 250, respectively. Reads that were shorter than those thresholds were discarded. The sequences were trimmed to the different lengths due to the difference in read qualities between the forward and reverse reads (forward is higher quality). Given these trimmed sequence lengths, the 16S reads will overlap but the 18S reads will not. 3.) We collected the 18S reads, by running all the reads through PEAR merging software, using default settings, and retained the unassembled reads. 4.) The forward and reverse reads of the unassembled reads were then joined with a degenerate base, "N", between two reads. This approach is suitable for the k-mer based classifier we used (Jeraldo et al. 2014). Due to relatively low numbers of 18S sequences 646 ±403 SD reads per sample), we did not perform MED, but clustered the sequences into OTUs at 99% sequence similarity via QIIME *pick_otus.py*, using the UCLUST option. 18S OTU representative sequences were classified with *assign_taxnomy.py* via RDP against the PR2 database and SILVA database, and against the NCBI database as previously described using BLASTn. We generally use the PR2 classifications. All classifications and representative sequences are available via Figshare (https://goo.gl/nM1cwe). For the 18S data, we generally report the data as proportions of non-metazoan 18S sequences, except where specified.

Phylogenetic trees were generated for the most abundant unique sequence from the ASVs (16S) and OTUs (18S) with MUSCLE default settings with a maximum of 100 iterations (Edgar 2004). Phylogenies were reconstructed using PhyML default settings (Guindon et al. 2010). Notably, in the 18S tree, *Myrionecta* is divergent from

311 the rest of the ciliates due to a very aberrant 18S sequence which has been
312 previously reported (Johnson et al. 2004).
313
314 **Statistical Analysis**
315 Pairwise correlations between parameters were performed using eLSA (Xia et al.
316 2011, 2013). Missing data were interpolated linearly (typically only a few dates per
317 dataset, except for rRNA of 18S which had 12 (of 53) days missing), and p and q
318 values were determined via a theoretical calculation (Xia et al. 2013). Due to the two
319 weeks of missing data and number we did not consider time-lagged correlations for
320 this time-series. Only correlations that had p and $q < 0.005$ were considered
321 significant. Mantel tests were performed in R (R Core Team 2015) with "mantel" as
322 part of the vegan package (Oksanen et al. 2015), on a fully overlapping datasets of
323 41 days. We excluded the rRNA 18S dataset from this analysis since it was the most
324 limited in number of total number of days appropriate for analysis.
325 **Results and Discussion**
326 Over the initial 6 days, conditions at the sample location, 1 km off Catalina Island,
327 CA, USA, were relatively stable, with chlorophyll concentrations of 0.5-1.5 µg/L
328 (Figure 1). After the 6$^{th}$ sample, the ESP malfunctioned. During the 15-day period
329 when it was non-operational, satellite data indicated that a modest increase in
330 phytoplankton chlorophyll occurred throughout the Southern California region. The
331 increase in phytoplankton biomass reached closest to the location of the ESP 4 days
332 before the instrument was repaired and sampling continued (Figure 1,
333 Supplementary Figure 3). When sampling resumed, the chlorophyll concentrations
334 were still elevated (though below peak levels according to satellite data) and
335 remained between 1-5 µg/L for the remainder of the time-series. We also noted
336 cyclical patterns within the chlorophyll data, apparently reflecting a combination of
337 diel phytoplankton physiological variations (Dandonneau & Neveux 1997) and
338 effects of tidal height and probably tidal current related depth variations of the
339 instrument (Figure 1).
340
341

342 *Overall community dynamics*

343 Small sub-unit sequencing of rRNA and rDNA retrieved 16S sequences of archaea,
344 bacteria and chloroplasts, as well as 18S of eukaryotes. Almost all of the non-
345 phytoplankton taxa (bacteria and archaea) that we observed in these near-surface
346 are presumed to have largely heterotrophic lifestyles, though many have
347 phototrophic capability via proteorhodopsin (Fuhrman et al. 2008) or
348 bacteriochlorophyll (Béjà et al. 2002; Schwalbach & Fuhrman 2005), and some are
349 chemoautotrophs (e.g., Thaumarchaea (Könneke et al. 2005)). Although we
350 recognize that divisions between classically defined trophic levels are increasingly
351 recognized as being "fuzzy," including the common occurrence of various kinds of
352 mixotrophs (Worden et al. 2015), we aimed to separate the 16S data into functional
353 guilds of classically-defined chlorophyll-*a* based phytoplankton and everything else
354 (presumably largely heterotrophs). This enables a first approximation of the
355 influence of the classic "base of the food web" ("phytoplankton") on the rest of the
356 food web. Due to the difficulty of accurately predicting the primary lifestyle of many
357 eukaryotic taxa determined by 18S, because of the unknown presence of
358 chloroplasts in some lineages and the ability of many protists to be mixotrophic
359 (Mitra et al. 2016), we generally analyzed them as one group. For most analyses, we
360 excluded metazoan sequences (e.g., copepods) that appeared in the data
361 sporadically (and unintentionally, when metazoans or their fragments passed the
362 300 µm prefilter), and their very high 18S copy number would alter the
363 interpretation of a primary focus, the microbial eukaryotes.
364
365 At the broadest level, the 16S rDNA sequences tended to be from non-
366 phytoplankton taxa, especially in the smaller size fraction, averaging 58% of the
367 total in the large size-fraction (1-300 µm) and 85% in the smaller size-fraction (0.22
368 – 1 µm)(Figure 2). Phytoplankton (via 16S) made up the majority of the rest of the
369 rDNA sequences, 39% and 15%, respectively, of the large and small size fractions.
370 Correspondingly, 18S made up 3.6% and 0.1% in the large and small fractions,
371 respectively. In the large fraction rDNA, relative proportion of phytoplankton and
372 non-phytoplankton varied considerably over the time-series, with the non-

373 photosynthetic prokaryotes being more abundant following the increase in
374 chlorophyll concentrations (Figure 2). This suggests increased heterotrophic
375 biomass in response to the increase in algal biomass. 18S sequences reached
376 maximum of 20% after the region-wide increase in chlorophyll (Figure 2).
377
378 In contrast to the rDNA, the relative proportion of phytoplankton rRNA was higher
379 than non-phytoplankton rRNA in the larger size fraction (65% and 35%,
380 respectively) including up to 75% rRNA sequences from 19 to 22 March and for
381 most samples between 14 to 17 April in the larger size fraction. An exception was
382 following the small phytoplankton bloom, when the 16S rRNA sequences from non-
383 photosynthetic taxa (like the rDNA) were up to 70% in the large size fraction. In the
384 smaller size fraction, the proportions of non-phytoplankton and phytoplankton
385 rRNA were of roughly equal proportion (averages of 53% and 43%, respectively),
386 with the exception of following the phytoplankton bloom when non-phytoplankton
387 made up >95% of the rRNA sequences in the small size fraction for several sampling
388 dates (Figure 2). In both size fractions, 18S rRNA always constituted less than 10%
389 of the total reads, and were almost always negligible in the small size fraction
390 (Figure 2).
391
392 *Dynamics of individual phytoplankton taxa*
393 Within the phytoplankton community, *Synechococcus* was typically the dominant
394 taxon in both rRNA and rDNA, in both size fractions (Figure 3). In the larger size
395 fraction, 2 different *Synechococcus* ASVs were the most abundant taxon in 24 and 44
396 of 50 days in rDNA and rRNA, respectively. In the smaller size fraction, a single
397 *Synechococcus* taxon was dominant in all 47 rDNA samples, and in 52 of 53 of the
398 rRNA samples, with *Prochlorococcus* exceeding it on a single date in rRNA.
399
400 Besides *Synechococcus*, in the larger size fraction, a variety of eukaryotic
401 phytoplankton (chloroplasts) were found to be periodically most abundant,
402 including *Ostreococcus* (14 days), *Teleaulax* (6 days), *Imantonia* (5 days),
403 *Braarudosphaera* (3 days), and *Pseudo-nitzschia* (1 day) (Figure 3). Several other

diatom ASVs, mostly *Chaetoceros sp.* and *Pseudo-nitzschia sp.*, peaked in abundance for a few days following the small phytoplankton biomass increase (potentially already decreasing by the time we resumed sampling) (Supplementary Figure 4). Pico-eukaryotic phytoplankton taxa (*Micromonas*, *Ostreococcus*) increased in numbers steadily over the second half of the time-series, and ultimately were the second and third most represented taxa in the phytoplankton rDNA on average, overall (Supplementary Figure 4). In addition, 2 ASVs of the diazotrophic, symbiotic unicellular cyanobacterium UCYN-A were cumulatively 1.1% and 5.6% on average in the large size-fraction rDNA and rRNA, respectively. UCYN-A constituted up to 25% of all rRNA phytoplankton sequences in that size fraction (more detail below) (Supplementary Figure 4). This observation of high rRNA and rDNA presence of UCYN-A a productive upwelling region is highly significant from an oceanographic standpoint since they are diazotrophs, and may be an important source of bio-available nitrogen in these surface waters even during spring and proceeding increases in phytoplankton biomass. These short-term dynamics and observations of high prevalence complements their previously documented activity at this location throughout the year where they were reported as particularly active in summer and winter (Hamersley et al. 2011).

In the smaller size fraction, besides Cyanobacteria, *Ostreococcus*, *Micromonas*, *Bathycoccous*, and *Pelagomonas* were commonly abundant (Figure 3, Supplementary Figure 4). It appears that these taxa tended to be equally split between both size fractions, with the exception of *Pelagomonas*, which was primarily found in the small size fraction. Generally, Cyanobacteria tended to be a higher proportion in the rRNA than rDNA, while the opposite was the case for the eukaryotic phytoplankton in the small size fraction. This suggests that cyanobacteria are likely among the more active members of the phytoplankton community.

435 *Dynamics of individual non-phytoplankton taxa*

436 A single SAR11 generally dominated ASV the rDNA of the non-phytoplankton
437 prokaryotic communities in the small size fraction (most abundant on 44 of 47
438 samples). The larger size fraction was dominated by a more diverse set of taxa.
439 However, in both the smaller and larger size fractions, the rRNA dominance shifted
440 among 11 and 10 different taxa, respectively, with particularly rapid dynamics
441 following the increase in phytoplankton biomass (Figure 3). For the large size
442 fraction, the same taxa tended to dominate in both the rRNA and rDNA. In the rDNA,
443 the dominance shifted between *Fluviicola* (24 days), *Roseovarius* (12), *Polaribacter*
444 (3), *Roseibacillus* (3), Verrucomicrobia (1), and Marine Group II Archaea (1).
445
446 Previously, we a reported on a larger diatom bloom that occurred 3 years earlier at
447 a location about 20 km away (Needham & Fuhrman 2016; Needham et al. 2017). We
448 also had daily-resolved data for this time-series. For that study we generated 99%
449 OTUs and then discriminated ASVs within the abundant OTUs (i.e., > 2.5% relative
450 abundance on any given day, or 0.4% on average). Overall, 119 of the 279 bacterial
451 and archaeal ASVs that we detected here were also reported in the previous study.
452 15 of the 20 ASVs that became most abundant during the present study were also
453 among the ASVs in the previous study. Several of the ASVs became most abundant in
454 both time-series: members of Flavobacteria (Polaribacter and Formosa),
455 Verrucomicrobia (*Roseibacillus* and *Puniceicoccaceae*) Marine Group II Archaea, and
456 *Roseovarius*, and SAR11. The rapid day-to-day variation 8 April – 12 April is similar
457 to what we observed previously and in both time-series, and the same ASVs of
458 *Polaribacter*, *Roseibacillus*, and Marine Group II Archaea became most abundant in
459 response to increases in chlorophyll, while *Roseovarius*, *Puniceicoccaceae*, and
460 SAR11 peaked during more stable conditions. However, the response here is not as
461 pronounced as in 2011, probably because that was a larger bloom, with presumably
462 a larger release of organic material. The consistency between years of
463 phytoplankton bloom response, even among exact sequence variants, is similar to
464 those reported from the North Sea (Chafee et al. 2017).
465

Often, particular ASVs were observed within both size fractions, but in the smaller size fraction, their temporal variation and overall relative abundances were reduced due to the sustained high relative abundance of SAR11 ASVs (cumulatively 23% and 30% in the rRNA and rDNA in 0.2-1 µM, respectively versus 2% and 6% in the 1-300 µm size fraction). Besides SAR11, other non-photosynthetic taxa that were higher in the smaller fraction were the gammaproteobacteria SAR92 and SAR86, and the alphaproteobacterium, OCS116 (Figure 3, Supplementary Figure 5). Notably a *Vibrio* ASV peaked up to 30% for one date prior to the bloom (up to 30% in rRNA and 2% in rDNA). This is surprising considering that *Vibrios* are typically thought to be "bloom-responders" but here were very active before the bloom, instead.

*Dynamics of individual eukaryotic taxa via 18S*

The whole eukaryotic community (1-300 µm) via 18S was often dominated by metazoans, such as herbivorous copepods (*Paracalanus*) and larvaceans (*Oikopleura*, which can graze particles as small as bacteria), with a copepod OTU (*Paracalanus* sp.) being most represented in the rDNA on 34 of 50 samples and larvacean OTU (*Oikopleura dioica)* being most represented on 16 of 44 dates in the rRNA (Supplementary Figure 6). Excluding metazoans, we observed 20 different 'most abundant' organisms via rDNA over the 51 sample points, including 21 days by Ciliates (10 days by *Myrionecta*), 11 days Chlorophytes (*Ostreococcus* (4), *Bathycoccus* (5), *Micromonas* (2)), and 9 dates Dinoflagellates (primarily *Gyrodinium* and *Gymnodinium* 4 and 2 dates, respectively) (Figure 3, Supplementary Figure 7). Similarly, ciliates were typically the most represented taxon in the rRNA (29 of 44 days), but in contrast to the rDNA, Stramenopiles were commonly most represented (14 of 44) dates, including a MAST-3 relative of *Solenicola* (99% match to clone FGII (Accession: HM163289) which is usually found associated with chain-forming diatoms (Padmakumar et al. 2012; Gómez et al. 2011) (6 dates), *Pseudo-nitzschia* (4), and MAST-1 (Massana et al. 2004), a distant 87% best match to *Rhizidiomyces* (Hypochytrids) (4). As suggested by the higher number of most dominant taxa, the Bray-Curtis community similarity metric showed that the eukaryotic community via 18S was much more variable than the 16S based prokaryotes and phytoplankton

(Supplementary Figure 8). The reasons that the dominance patterns vary between rRNA and rDNA are probably a combination of copy number differences and levels of activity, even given that dormant cells have a baseline level of rRNA (Blazewicz et al. 2013).

**Correlations between taxa**

Previously almost all correlation analyses between taxa have been between the abundance of organisms (DNA- or count- based), irrespective of activity. However for many types of interactions, it would be valuable to consider some indicator of activity level of the organisms as well. We aim to do so here by including rRNA in addition to the rDNA relative abundances in the co-occurrence patterns between taxa. We first examine known 2-organism positive interactions that occur among abundant taxa within our samples. This allows assessment of the nature of known associations in the environment. Additionally, it allows identification of organisms that may be previously unrecognized members of these associations, including those that may replace a known member under some circumstances while retaining similar function. Additionally, we examine the strong correlations across all taxa to identify possible interactions among and between domains, such as syntrophy, symbiosis, or grazing.

*UCYN-A and Braarudosphaera*

Researchers studying marine nitrogen fixation by molecular genetic analysis of *(nifH)* genes discovered a widely distributed and important group of nitrogen fixers that for several years went unidentified, but recently were found to be part of a symbiotic association (Zehr et al. 2016; Farnelid et al. 2016; Thompson et al. 2012). The organism, recently named *Candidatus Atelocyanobacterium thalassa,* but still commonly known as UCYN-A, is a marine unicellular nitrogen-fixing cyanobacterium with greatly reduced genome for a cyanobacterium, lacking the ability to generate oxygen (which inhibits nitrogen fixation) and possessing incomplete TCA cycle pathways (Tripp et al. 2010). Its metabolic deficiencies are evidently met by having a symbiotic relationship with algae (Thompson et al. 2012). At least four strains of UCYN-A have been reported (denoted UCYN-A1, A2, A3, and

A4), based on phylogeny of the *nifH* sequence. These strains vary in their global distribution, size, symbiotic hosts, in addition to other genetic and probably physiological differences (Farnelid et al. 2016). The most well-supported UCYN-A symbiosis is a relationship between UCYN-A2 and the haptophyte alga *Braarudosphaera bigelowii* (Zehr et al. 2016; Thompson et al. 2012; Farnelid et al. 2016), in which fixed nitrogen is exchanged for organic substrates UCYN-A needs for growth. Other UCYN-A types are thought to be associated with different phytoplankton, including with species closely related to *Braarudosphaera (*Zehr et al. 2016).

We observed two 16S ASVs of UCYN-A, each an exact match to genomic sequences from UCYN-A types. One ASV was a perfect match to a genome sequence of UCYN-A1 (gi|284809060) and another with a perfect match to a genome scaffold of UCYN-A2 (gi|671395793) (Figure 4, Supplementary Figure 7). These two ASVs differed by 3 base pairs over the 375 base pairs 16S amplicon sequence we analyzed. The dynamics of the rDNA relative abundance of the UCYN-A1 and UCYN-A2 were similar over the full time-series (Spearman r= 0.64). There was a pronounced increase in both types from 18 April to 25 April when UCYN-A1 increased from about 0.5% to about 3% in relative abundance of all phytoplankton, while the increase in UCYN-A2 was less pronounced (it peaked to about 1.5% on 25 April). Both UCYN-A types also peaked in early March -- though the peaks were offset slightly (by 1 day via rDNA, 2 days via rRNA). Both were relatively low in early and late April.

Overall, the rRNA levels were similar in dynamics to the rDNA and to one another (Figure 4), though the mid-to-late April peaks were more similar in amplitude and timing in the rRNA than the rDNA when UCYN-A1 was about 2x as relatively abundant. The average rRNA relative abundance of UCYN-A1 and UCYN-A2 was 2.4% and 3.1%, respectively, which is was about 3 and 10x the relative concentration of their rDNA.

559 Likewise, *Braarudosphaera bigelowii* (1 bp different over 368 bp (99.7%) to an NCBI
560 16S chloroplast sequence from *Braarudosphaera, Accession:* AB847986.2, (Hagino et
561 al. 2013) was high during March, low in early April, peaked during the middle of
562 April and decreased after April 24 (Figure 4). The rRNA and rDNA of
563 *Braarudosphaera* chloroplasts were correlated (0.64, $p < 0.001$), with an average
564 rRNA to rDNA ratio of 0.93.
565
566 In general, *Braarudosphaera* and UCYN-A were highly positively correlated, and the
567 best correlations were between the *Braarudosphaera rDNA* and UCYN-A1 rDNA (r=
568 0.85, Figure 4C), while the correlation to UCYN-A2 was not quite as strong (r = 0.76).
569 *Braarudosphaera* rRNA was correlated to both UCYN-A1 and UCYN-A2 rRNA (r=0.81
570 and 0.83, respectively). UCYN-A1 rDNA was also significantly correlated to
571 *Braarudospheara* rRNA (r = 0.63), but the other combinations of rRNA to rDNA and
572 vice-versa between the taxa were not as significantly correlated (i.e., $p < 0.005$).
573 Given that the literature reports a specific relationship between UCYN-A2 and
574 *Braarudospheara* and between UCYN-A1 and a closely related taxon (Zehr et al.
575 2016), it may be that the 16S amplicon sequence does not discriminate between
576 distinct but closely related haptophyte species that may be present with similar
577 dynamics. An alternative speculation is that the *Braarudosphaera* strain(s) present
578 at this site can be a host to UCYN-A1 and UCYN-A2.
579
580 We found that there were several other taxa highly correlated to *Braarudosphaera*,
581 which implies the possibility of additional members of this known symbiosis, or at
582 least strong affiliations. A Dictyochophyte alga (rRNA) had a particularly strong
583 correlation to *Braarudosphaera* (r = 0.86), with a slightly weaker correlation
584 between the rDNA of the two UCYN-A taxa (Figure 4). A Puniceicoccaceae
585 (Verrucomicrobium) rRNA and rDNA was very strongly correlated to
586 *Braarudosphaera* (all r >0.81) and UCYN-A1. Verrucomicrobia are often found to be
587 particle associated (Crespo et al. 2013; Rath et al. 1998; Mestre et al. 2017;
588 Needham & Fuhrman 2016), and were indeed enriched in the larger size fraction in
589 our samples, suggesting possible physical attachment in an association. Other ASVs

strongly correlated (r > 0.8) to *Braarudosphaera* were two photosynthetic algae, both *Chrysochromulina*, a genus that previous evidence suggests also may be a host to UCYN-A, particularly UCYN-A1 (Thompson et al. 2012; Zehr et al. 2016) (Figure 4).

*Myrionecta and Teleaulax*

Another known interaction between abundant taxa in our dataset is that of the ciliate *Myrionecta rubra (=Mesodinium rubrum)* with the photosynthetic cryptophyte, *Teleaulax.* In this interaction, *Myrionecta* is thought to phagocytize *Teleaulax* and retain functioning *Telaulax* chloroplast within the *Myrionecta* cells, becoming functionally phototrophic (Gustafson et al. 2000). *Myrionecta* is not capable of photosynthesis without this association, but, when possessing the chloroplasts, it can perform high rates of photosynthesis (Stoecker et al. 1991) and can form massive blooms ("red tides")(Taylor et al. 1971). *Myrionecta* is capable of consuming many strains of cryptophytic algae from the Teleaulax/ Plagioselmis/ Geminigera clade (Peterson et al. 2012; Hansen et al. 2012; Park et al. 2007), but it appears most often associated with chloroplasts from *Teleaulax amphioxae* (Johnson et al. 2016; Herfort et al. 2011; Hansen et al. 2006). The exact nature and mechanisms of the interaction is unclear, but the *Teleaulax* chloroplasts can remain intact and harbored for days to weeks within the *Myrionecta* (Johnson et al. 2007; Herfort et al. 2011). The chloroplasts reportedly can be replicated within the *Myrionecta* with the assistance of captured *Teleaulax* nuclei (Johnson et al. 2007). It is unclear to what extent the relationship is most similar to *kleptochloroplastic* relationships, whereby chloroplasts are consumed and used until they lose function without nuclear assistance; *karyoplastic* ones, whereby chloroplasts can be maintained by consuming and retention of the nucleus of grazed *Teleaulax*, or an *endosymbiosis* where *Myrionecta* harbors *Teleaulax* chloroplasts permanently or nearly so. Alternatively, *Myrionecta* may be "farming" whole *Teleaulax* cells (Qiu et al. 2016), a conclusion met with firm skepticism (Johnson et al. 2016). Additionally, *Myrionecta* has been observed to harbor extensive microdiversity, with various strains co-existing within a location (Herfort et al. 2011) and variability in dominant types between locations (Herfort et al. 2011; Johnson et al. 2016); the nature of

interactions between *Myrionecta* and *Teleaulax* may be variable between different strains. A further aspect is that the toxic dinoflagellate *Dinophysis* is reported to obtain its chloroplasts by feeding on *Myrionecta*, which in that case would be an intermediate source from *Telaulax* (Garcia-Cuetos et al. 2010; Sjöqvist & Lindholm 2011).

We found that *Myrionecta* and *Teleaulax* were among the most abundant and common taxa found in the eukaryotic community (18S) and phytoplankton communities (via 16S chloroplasts), respectively (Figure 3, Supplementary Figure 4 and 7). On average, the most abundant *Teleaulax* ASV (an exact match over the full 374 base pairs to *Teleaulax amphioxeia,* Supplementary Figure 9*)* made up 5.5% and 12.4% of chloroplast rDNA and rRNA, respectively and the most abundant *Myrionecta* OTU (an exact match to *Myrionecta major* strain LGC-2011, which was described from coastal Denmark, Supplementary Figure 10) made up 2.0% and 2.5% of rDNA and rRNA 18S sequences (Figure 5). The rDNA of these taxa increased in abundance between April 15 and April 20, and again between 24 April and 26 April. The Spearman correlation between the rDNA of these taxa (*Myrionecta* and *Teleaulax_amphioxea_1* chloroplasts) was 0.86 and the relationship seemed highly exclusive, i.e., no significant correlations to anything else (Figure 5D).

A second abundant *Teleaulax* chloroplast sequence (3 base pairs different from the best match, *Teleaulax amphioxeia*) was also commonly detected with an average abundance of 2.0% rDNA and 1.6% rRNA. This *Teleaulax* chloroplast ASV was not significantly correlated with *Myrionecta;* however, it was significantly correlated with a *Teleaulax* 18S (nuclear) OTU ($r = 0.67$, $p < 0.005$, Figure 5D). Unlike the *Myrionecta-Teleaulax* association, these *Teleaulax* taxa were positively correlated with many other taxa, rDNA of *Synechococcus*, Alphaproteobacteria (OCS116 and *Defluuivicoccus*) NS5 genus of Bacteroidetes, Marine Group II Archaea, and a Sphingobacterium (all $r > 0.7$).

Thus, we saw a strong relationship between one ASV of one *Teleaulax_1* chloroplast ASV, but not nuclei, and *Myrionecta* over the 1.5 month of study. Additionally, based on correlation between *Teleaulax* (nuclear) 18S and a second *Teleaulax* chloroplast, it appears the second strain of free-living *Teleaulax* present may not be associated with *Myrionecta* cells. We did observe *Dinophysis* in our samples (Supplementary Figure 4-5), without significant correlations to support a *Myrionecta* or *Teleaulax* interaction (i.e., secondary kleptoplastidy) specifically or consistently with these taxa at our location and timeframe; however such a statistical relationship may not be expected if the abundance of *Dinophysis* is not dependent on contemporaneous availability of *Myrionecta-Teleaulax* via a specific grazing dependency. In our study, we found the *Teleaulax* nuclei were regularly present at the sample site, but they did not correlate with either *Myrionecta-Teleaulax* chloroplast association.

Our observations of strong, consistent relationship over about 1.5 months between specific types of *Myrionecta* and chloroplasts from *Teleaulax* lends some support either the hypothesis that the relationship is an endosymbiosis (Herfort et al. 2011; Hansen et al. 2006) or that *Myrionecta* can maintain chloroplasts a long time with the periodic help of *Teleaulax* nuclei (Johnson et al. 2007). Because a single *Teleaulax* nucleus in a *Myrionecta* cell might support replication of many more captured chloroplasts than would be found in a single *Teleaulax* cell (Johnson et al. 2007), the possible need for *Teleaulax* nuclei might be hard to discern via correlations between 18S nuclear genes in our system (i.e., if the ratio of *Teleaulax* nuclei and chloroplasts is highly variable within *Myrionecta*), in contrast to the strong correlative relationship between the relative abundances of *Teleaulax* chloroplasts and *Myrionecta* 18S. Alternatively, other *Teleaulax* nuclei may be present but in lesser abundance (and 18S copies per cell), reducing the ability to regularly detect them in strong co-occurrence with the *Teleaulax* and *Myrionecta*.

*Other correlations between taxa*

To gain an understanding for how the communities (i.e. phytoplankton, non-phytoplankton 16S, 18S, DNA or RNA, different size fractions) changed, overall, in

relation to one another, we performed Mantel tests. While all the different communities were significantly correlated ($p < 0.001$), the strongest correlations were between phytoplankton via 16S and non-phytoplankton 16S (Supplementary Figure 11). The best correlation was between the rDNA of the phytoplankton and the large or particle-attached non-phytoplankton 16S, and the slightly less so to the free-living and small non-phytoplankton 16S (Figure 6, Supplementary Figure 11). This is a similar result as we previously reported (Needham and Fuhrman 2016) where there was a stronger correlation of eukaryotic phytoplankton via chloroplasts to attached bacteria and archaea than to free-living bacteria and archaea. The rDNA of phytoplankton 16S was more related to both the rRNA and rDNA of non-phytoplankton 16S communities, than the rRNA of phytoplankton 16S was to either. These strong affiliations of phytoplankton and attached bacteria and archaea are perhaps a result of symbiotic interactions where both phytoplankton and prokaryotes benefit from the association (Amin et al. 2012, 2015). Another hypothesis is that different phytoplankton communities generate different suspended and sinking marine aggregates that in turn harbor different bacterial communities. There may be stronger relationships between rDNA versus rRNA because physical attachment, potentially limited by cell numbers (rDNA), is relatively more important than the potential activity (rRNA) of cells.

The observation of relatively weak correlation between eukaryotes by 18S to the other communities may be because grazing may be less species-specific, as has been previously concluded based on co-occurrence patterns (Chow et al. 2014). Thus, the strength of correlation between particular grazers and bacteria may not be strong since grazers may be more likely to graze bacteria of similar size than in a species-specific manner. This could result in an increase in the abundance of the grazer in response to an increase in any of a variety of similarly sized bacteria.

For pairwise correlations, several of the phytoplankton-to-heterotrophic bacteria correlations are the same as those that we previously reported, including those between Rhodobacteraceae, *Polaribacter* (Flavobacteria), and SAR92 to diatoms

713 *Pseudo-nitzschia* and *Chaetoceros* (Figure 6), suggesting that these correlations are
714 specific and repeatable between different time-series even though they were
715 separated by 3 years and about 20 km. The associations of these prokaryotic groups
716 with phytoplankton, especially in diatom blooms, have been reported previously,
717 with responses at time-scales from weeks to months (Teeling et al. 2012 & 2016,
718 Buchan et al. 2012). These interactions involve finely-resolved taxa (ASVs)
719 consistent with the patterns uncovered in the North Sea (Chafee et al. 2017).
720
721 We also observed a group of highly positively correlated *Prochlorococcus* ASV to
722 various Flavobacteria and Verrucomicrobia, indicating the shared ecosystem
723 preferences or interactions. These taxa all steadily increased in rRNA and rDNA
724 from about 11 April until to the end of the time-series (Figure 6).
725
726 In addition to those types of interactions that were previously described, we also
727 observed many strong correlations between heterotrophic or mixotrophic
728 eukaryotic taxa and potential symbionts or prey (Figure 7). Of these, only 5 taxa had
729 strong correlations ($|\text{Spearman } r| > 0.7$, $p < 0.005$) to bacteria or phytoplankton; of
730 these 4 were ciliates. In addition to the relationship between *Myrionecta* and
731 *Teleaulax* described previously, the ciliate OTUs of *Strombidium* were shown to have
732 correlations to a variety of bacteria, including Flavobacteria, Gammaproteobacteria,
733 relatively rare ASVs classified as mycoplasma, and *Pseudo-nitzschia*. Thus it appears
734 that these particular ciliates may have specific interactions with these bacteria, and
735 may be good targets for future analyses to determine the nature of these
736 interactions. We observed no strong negative correlations (Spearman $r < -0.7$), in
737 contrast to many strong positive ones >0.7 or even 0.8, suggesting that the specific
738 interactions we observe were much more likely to represent mutually beneficial
739 interactions, without strong antagonistic relationships. This observation of strong
740 positive without strong negative correlations in fact applies to all the associations
741 we observed in this study, generally.
742
743

## Conclusions

Our results show a rapid, day-to-day response of particular microbial taxa to changes in phytoplankton. In this case we saw only a small increase in chlorophyll (especially compared to previous studies) and yet the patterns we have previously observed persisted. Observations of microbial dynamics via rRNA and rDNA yielded somewhat similar results, though the overall proportions of taxa could change between the rRNA and rDNA, with phytoplankton often dominating the rRNA sequences. Via co-occurrence patterns, our results provide new *in-situ* characterizations of previously observed interactions among the most abundant taxa. They suggested that the *Myrionecta*-to-*Teleaulax* chloroplast association appeared to occur independently of other microbial associations, while UCYN-A-to-*Braarudosphaera* co-occurred strongly with several taxa. Overall, the study reiterates the utility of short-term time-series for understanding environmental responses and microbe-to-microbe interactions where turnover times can be very fast.

## Acknowledgements

We thank the Gordon and Betty Moore Foundation for funding (Grant Number: 3779) and their support of this project. We thank the Scholin Lab at the Monterey Bay Aquarium Research Institute for deployment, data acquisition, quality control, and help with experimental design. In particular we thank Christina Preston, Roman Marin III, James Birch, Scott Jensen, Brent Roman, Bill Ussler, and Kevan Yamahara for their help with the ESP. We thank the Wrigley Institute of Environmental Studies for logical support, especially Gordon Boivin. We thank the National Science Foundation for financial support (Grant Number: 1136818).

## Conflicts of Interest

The authors declare no conflict of interest

775 **References**

776 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment
777 search tool. J Mol Biol 215:403–10.
778 Amin SA, Hmelo LR, van Tol HM, Durham BP, Carlson LT, Heal KR, et al. (2015).
779 Interaction and signalling between a cosmopolitan phytoplankton and associated
780 bacteria. Nature 522:98–101.
781 Amin SA, Parker MS, Armbrust EV. (2012). Interactions between diatoms and
782 bacteria. Microbiol Mol Biol Rev 76:667–84.
783 Azam F, Fenchel T, Field JG, Gray JS, Thingstad TF. (1983). The Ecological Role of
784 Water-Column Microbes in the Sea. Mar Ecol Prog Ser 10:257–263.
785 Béjà O, Suzuki MT, Heidelberg JF, Nelson WC, Preston CM, Hamada T, et al. (2002).
786 Unsuspected diversity among marine aerobic anoxygenic phototrophs. Nature
787 415:630–633.
788 Blazewicz SJ, Barnard RL, Daly RA, Firestone MK. (2013). Evaluating rRNA as an
789 indicator of microbial activity in environmental communities: limitations and uses.
790 ISME J 7:2061–2068.
791 Bombar D, Heller P, Sanchez-Baracaldo P, Carter BJ, Zehr JP. (2014). Comparative
792 genomics reveals surprising divergence of two closely related strains of
793 uncultivated UCYN-A cyanobacteria. ISME J 8:2530–2542.
794 Brown M V., Fuhrman JA. (2005). Marine bacterial microdiversity as revealed by
795 internal transcribed spacer analysis. Aquat Microb Ecol 41:15–23.
796 Buchan A, LeCleir GR, Gulvik CA, González JM. (2014). Master recyclers: features and
797 functions of bacteria associated with phytoplankton blooms. Nat Rev Microbiol
798 12:686–698.
799 Callahan BJ, Mcmurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. (2016).
800 DADA2 : High resolution sample inference from amplicon data. Nat Methods
801 13:581–583.
802 Campbell BJ, Yu L, Heidelberg JF, Kirchman DL. (2011). Activity of abundant and
803 rare bacteria in a coastal ocean. Proc Natl Acad Sci U S A 108:12776–12781.
804 Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al.
805 (2010). QIIME allows analysis of high-throughput community sequencing data. Nat

806 Methods 7:335–336.
807 Chafee M, Fernàndez-Guerra A, Buttigieg PL, Gerdts G, Eren AM, Teeling H, et al.
808 (2017). Recurrent patterns of microdiversity in a temperate coastal marine
809 environment. ISME J 1–16.
810 Chow C-ET, Kim DY, Sachdeva R, Caron DA, Fuhrman JA. (2014). Top-down controls
811 on bacterial community structure: microbial network analysis of bacteria, T4-like
812 viruses and protists. ISME J 8:816–29.
813 Countway PD, Gast RJ, Savai P, Caron DA. (2005). Protistan Diversity Estimates
814 Based on 18S rDNA from Seawater Incubations in the Western North Atlantic. J
815 Eukaryot Microbiol 52:95–106.
816 Cram JA, Chow C-ET, Sachdeva R, Needham DM, Parada AE, Steele JA, et al. (2014).
817 Seasonal and interannual variability of the marine bacterioplankton community
818 throughout the water column over ten years. ISME J 1–18.
819 Crespo BG, Pommier T, Fernández-Gómez B, Pedrós-Alió C. (2013). Taxonomic
820 composition of the particle-attached and free-living bacterial assemblages in the
821 Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA.
822 Microbiologyopen 2:541–52.
823 Dandonneau Y, Neveux J. (1997). Diel variations of in vivo fluorescence in the
824 eastern equatorial Pacific: an unvarying pattern. Deep Res Part II Top Stud Oceanogr
825 44:1869–1880.
826 Decelle J, Romac S, Stern RF, Bendif EM, Zingone A, Audic S, et al. (2015). PhytoREF:
827 a reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes
828 with curated taxonomy. Mol Ecol Resour 15:1435–1445.
829 Edgar RC. (2004). MUSCLE: Multiple sequence alignment with high accuracy and
830 high throughput. Nucleic Acids Res 32:1792–1797.
831 Edgar RC. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon
832 reads. Nat Methods 10:996–8.
833 Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, et al. (2013).
834 Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA
835 gene data. Methods Ecol Evol 4:1111–1119.
836 Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis JH, Sogin ML. (2014).

837 Minimum entropy decomposition: Unsupervised oligotyping for sensitive
838 partitioning of high-throughput marker gene sequences. ISME J 9:968–979.
839 Farnelid H, Turk-Kubo K, Muñoz-Marín M, Zehr JP. (2016). New insights into the
840 ecology of the globally significant uncultured nitrogen-fixing symbiont UCYN-A.
841 Aquat Microb Ecol 77:125–138.
842 Fuhrman JA, Azam F. (1982). Thymidine incorporation as a measure of
843 heterotrophic bacterioplankton production in marine surface waters: Evaluation
844 and field results. Mar Biol 66:109–120.
845 Fuhrman JA, Comeau DE, Hagstrom A, Chan AM. (1988). Extraction from natural
846 planktonic microorganisms of DNA suitable for molecular biological studies. Appl
847 Environ Microbiol 54:1426–1429.
848 Fuhrman JA, Cram JA, Needham DM. (2015). Marine microbial community dynamics
849 and their ecological interpretation. Nat Rev Microbiol 13:133–146.
850 Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown M V., Naeem S. (2006).
851 Annually reoccurring bacterial communities are predictable from ocean conditions.
852 Proc Natl Acad Sci U S A 103:13104–13109.
853 Fuhrman JA, Schwalbach MS, Stingl U. (2008). Proteorhodopsins: an array of
854 physiological roles? Nat Rev Micro 6:488–494.
855 Garcia-Cuetos L, Moestrup Ø, Hansen PJ, Daugbjerg N. (2010). The toxic
856 dinoflagellate Dinophysis acuminata harbors permanent chloroplasts of
857 cryptomonad origin, not kleptochloroplasts. Harmful Algae 9:25–38.
858 Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, et al. (2012).
859 Defining seasonal marine microbial community dynamics. ISME J 6:298–308.
860 Gómez F, Moreira D, Benzerara K, López-garcía P. (2011). Solenicola setigera is the
861 first characterized member of the abundant and cosmopolitan uncultured marine
862 stramenopile group MAST-3. Environ Microbiol 13:193–202.
863 Guindon S, Dufayard J, Lefort V, Anisimova M, Hordijk W, Gascuel O. (2010). New
864 algorithms and methods to estimate maximum-likelihood phylogenies: Assessing
865 the performance of PhyML 3.0. Syst Biol 59:307–21.
866 Gustafson DE, Stoecker DK, Johnson MD, Van Heukelem WF, Sneider K. (2000).
867 Cryptophyte algae are robbed of their organelles by the marine ciliate Mesodinium

868 rubrum. Nature 405:1049–1052.
869 Hagino K, Onuma R, Kawachi M, Horiguchi T. (2013). Discovery of an endosymbiotic
870 nitrogen-fixing cyanobacterium UCYN-A in Braarudosphaera bigelowii
871 (Prymnesiophyceae). PLoS One 8:e81749.
872 Hamersley M, Turk K, Leinweber A, Gruber N, Zehr JP, Gunderson T, et al. (2011).
873 Nitrogen fixation within the water column associated with two hypoxic basins
874 within the Southern California Bight. Aquat Microb Ecol 63:193–205.
875 Hansen PJ, Fenchel TOM, Juel Hansen P. (2006). The bloom-forming ciliate
876 Mesodinium rubrum harbours a single permanent endosymbiont. Mar Biol Res
877 2:169–177.
878 Hansen PJ, Moldrup M, Tarangkoon W, Garcia-Cuetos L, Moestrup. (2012). Direct
879 evidence for symbiont sequestration in the marine red tide ciliate Mesodinium
880 rubrum. Aquat Microb Ecol 66:63–75.
881 Herfort L, Peterson TD, McCue LA, Crump BC, Prahl FG, Baptista AM, et al. (2011).
882 Myrionecta rubra population genetic diversity and its cryptophyte chloroplast
883 specificity in recurrent red tides in the Columbia River estuary. Aquat Microb Ecol
884 62:85–97.
885 Hunt DE, Lin Y, Church MJ, Karl DM, Tringe SG, Izzo LK, et al. (2013). Relationship
886 between abundance and specific activity of bacterioplankton in open ocean surface
887 waters. Appl Environ Microbiol 79:177–184.
888 Jeraldo P, Kalari K, Chen X, Bhavsar J, Mangalam A, White B, et al. (2014). IM-
889 TORNADO: a tool for comparison of 16S reads from paired-end libraries. PLoS One
890 9:e114804.
891 Johnson MD, Beaudoin DJ, Laza-martinez A, Dyhrman ST, Fensin E, Lin S, et al.
892 (2016). The Genetic Diversity of Mesodinium and Associated Cryptophytes. Front
893 Aquat Microbiol 7.
894 Johnson MD, Oldach D, Delwiche CF, Stoecker DK. (2007). Retention of
895 transcriptionally active cryptophyte nuclei by the ciliate Myrionecta rubra. Nature
896 445:426–8.
897 Johnson MD, Tengs T, Oldach DW, Delwiche CF, Stoecker DK. (2004). Highly
898 Divergent SSU rRNA Genes Found in the Marine Ciliates Myrionecta rubra and

899 Mesodinium pulex. Protist 155:347–359.
900 Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. (2005).
901 Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature 437:543–
902 546.
903 Lankiewicz TS, Cottrell MT, Kirchman DL. (2016). Growth rates and rRNA content of
904 four marine bacteria in pure cultures and in the Delaware estuary. ISME J 10:823–
905 832.
906 Martin M. (2011). Cutadapt removes adapter sequences from high-throughput
907 sequencing reads. EMBnet.journal 17.
908 Massana R, Castresana J, Balague V, Guillou L, Romari K, Groisillier A, et al. (2004).
909 Phylogenetic and Ecological Analysis of Novel Marine Stramenopiles. Appl Environ
910 Microbiol 70:3528–3534.
911 McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. (2012).
912 An improved Greengenes taxonomy with explicit ranks for ecological and
913 evolutionary analyses of bacteria and archaea. ISME J 6:610–8.
914 Mestre M, Borrull E, Sala Mm, Gasol JM. (2017). Patterns of bacterial diversity in the
915 marine planktonic particulate matter continuum. ISME J 1–12.
916 Mitra A, Flynn KJ, Tillmann U, Raven JA, Caron DA, Stoecker DK, et al. (2016).
917 Defining Planktonic Protist Functional Groups on Mechanisms for Energy and
918 Nutrient Acquisition: Incorporation of Diverse Mixotrophic Strategies. Protist
919 167:106–120.
920 Needham DM, Fuhrman JA. (2016). Pronounced daily succession of phytoplankton,
921 archaea and bacteria following a spring bloom. Nat Microbiol 1:16005.
922 Needham DM, Sachdeva R, Fuhrman JA. (2017). Ecological dynamics and co-
923 occurrence among marine phytoplankton, bacteria and myoviruses shows
924 microdiversity matters. ISME J 11:1614–1629.
925 Oksanen AJ, Blanchet FG, Kindt R, Legendre P, Minchin PR, Hara RBO, et al. (2015).
926 vegan: Community Ecology Package.
927 Padmakumar KB, Cicily L, Shaji A, Maneesh TP, Sanjeevan VN. (2012). Symbiosis
928 between the stramenopile protist Solenicola setigera and the diatom Leptocylindrus
929 mediterraneus in the North Eastern Arabian Sea. Symbiosis 56:97–101.

930 Parada AE, Needham DM, Fuhrman JA. (2016). Every base matters: assessing small
931 subunit rRNA primers for marine microbiomes with mock communities, time-series
932 and global field samples. Environ Microbiol 18:1403–1414.
933 Park JS, Myung G, Kim HS, Cho BC, Yih W. (2007). Growth responses of the marine
934 photosynthetic ciliate Myrionecta rubra to different cryptomonad strains. Aquat
935 Microb Ecol 48:83–90.
936 Peterson TD, Golda RL, Garcia ML, Li B, Maier MA, Needoba JA, et al. (2012).
937 Associations between Mesodinium rubrum and cryptophyte algae in the Columbia
938 River estuary. Aquat Microb Ecol 68:117–130.
939 Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. (2007). SILVA: a
940 comprehensive online resource for quality checked and aligned ribosomal RNA
941 sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196.
942 Qiu D, Huang L, Lin S. (2016). Cryptophyte farming by symbiotic ciliate host
943 detected in situ. Proc Natl Acad Sci U S A 199:201612483.
944 R Core Team. (2015). R: A Language and Environment for Statistical Computing.
945 Rath J, Wu KY, Herndl GJ, DeLong EF. (1998). High phylogenetic diversity in a
946 marine snow-associated bacterial assemblage. Aquat. Microb. Ecol. 14:261–269.
947 Scholin C, Doucette G, Jensen S, Roman B, Pargett D, Marin RI, et al. (2009). Remote
948 detection of marine microbes, small invertebrates, harmful algae, and biotoxins
949 using the Environmental Sample Processor (ESP). Oceanography 22:158–167.
950 Schwalbach MS, Fuhrman JA. (2005). Wide-ranging abundances of aerobic
951 anoxygenic phototrophic bacteria in the world ocean revealed by epifluorescence
952 microscopy and quantitative PCR. Limnol Oceanogr 50:620–628.
953 Sjöqvist CO, Lindholm TJ. (2011). Natural co-occurrence of dinophysis acuminata
954 (Dinoflagellata) and mesodinium rubrum (Ciliophora) in thin layers in a coastal
955 inlet. J Eukaryot Microbiol 58:365–372.
956 Stoecker DK, Putt M, Davis LH, Michaels AE. (1991). Photosynthesis in Mesodinium
957 rubrum: species specific measurements and comparison to community rates. Mar
958 Ecol Prog Ser 73:245–252.
959 Taylor FJR, Blackbourn DJ, Blackbourn J. (1971). The Red-Water Ciliate Mesodinium
960 rubrum and its “Incomplete Symbionts”: A Review Including New Ultrastructural

961 Observations. J Fish Res Board Canada 28:391–407.
962 Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, et al. (2012).
963 Substrate-controlled succession of marine bacterioplankton populations induced by
964 a phytoplankton bloom. Science 336:608–611.
965 Teeling H, Fuchs BM, Bennke CM, Krüger K, Chafee M, Kappelmann L, et al. (2016).
966 Recurring patterns in bacterioplankton dynamics during coastal spring algae
967 blooms. Elife 5:1–31.
968 Thompson AW, Foster RA, Krupke A, Carter BJ, Musat N, Vaulot D, et al. (2012).
969 Unicellular Cyanobacterium Symbiotic with a Single-Celled Eukaryotic Alga. Science
970 337:1546–1550.
971 Tikhonov M, Leach RW, Wingreen NS. (2015). Interpreting 16S metagenomic data
972 without clustering to achieve sub-OTU resolution. ISME J 9:68–80.
973 Trinh RC, Fichot CG, Gierach MM, Holt B, Malakar NK, Hulley G, et al. (2017).
974 Application of Landsat 8 for monitoring impacts of wastewater discharge on coastal
975 water quality. Front Mar Sci 4:329.
976 Tripp HJ, Bench SR, Turk K a, Foster R a, DeSantis TZ, Niazi F, et al. (2010). Metabolic
977 streamlining in an open-ocean nitrogen-fixing cyanobacterium. Nature 464:90–4.
978 de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, et al. (2015). Eukaryotic
979 plankton diversity in the sunlit ocean. Science 348:1261605-1–11.
980 Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ. (2015).
981 Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of
982 microbes. Science 347:1257594–1257594.
983 Xia LC, Ai D, Cram JA, Fuhrman JA, Sun F. (2013). Efficient statistical significance
984 approximation for local similarity analysis of high-throughput time series data.
985 Bioinformatics 29:230–7.
986 Xia LC, Steele JA, Cram JA, Cardon ZG, Simmons SL, Vallino JJ, et al. (2011). Extended
987 local similarity analysis (eLSA) of microbial community and other time series data
988 with replicates. BMC Syst Biol 5:1–12.
989 Zehr JP, Bench SR, Carter BJ, Hewson I, Niazi F, Shi T, et al. (2008). Globally
990 Distributed Uncultivated oceanic N2-Fixing Cyanobacteria Lack Oxygenic
991 Photosystem II. Nature 322:1110–1112.

992 Zehr JP, Shilova IN, Farnelid HM, Muñoz-Marín M del C, Turk-Kubo KA. (2016).
993 Unusual marine unicellular symbiosis with the nitrogen-fixing cyanobacterium
994 UCYN-A. Nat Microbiol 2:1:10.
995
996

# Figure 1

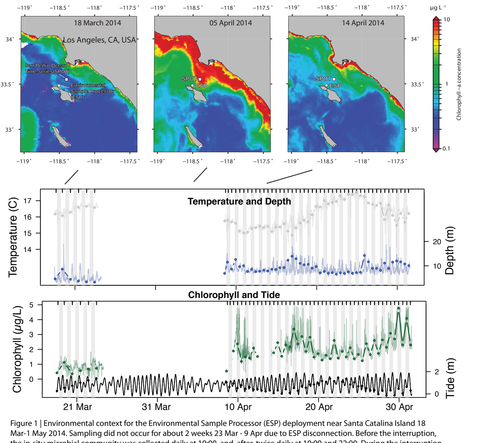


Figure 1 | Environmental context for the Environmental Sample Processor (ESP) deployment near Santa Catalina Island 18 Mar–1 May 2014. Sampling did not occur for about 2 weeks 23 Mar - 9 Apr due to ESP disconnection. Before the interruption, the in-situ microbial community was collected daily at 10:00, and, after, twice daily at 10:00 and 22:00. During the interruption, A.) satellite chlorophyll measurements indicated a small increase in chlorophyll occurred throughout the San Pedro Channel, peaking 4 days before resumption of sampling. B.) Temperature and depth, and C.) Chlorophyll and Tide measurements were taken from the ESP every 5 minutes (the thin lines); the circles represent the average value during sample collection (usually about 30 minutes). Sampling times are indicated by ticks at the top of B.) and C.).

# Figure 2

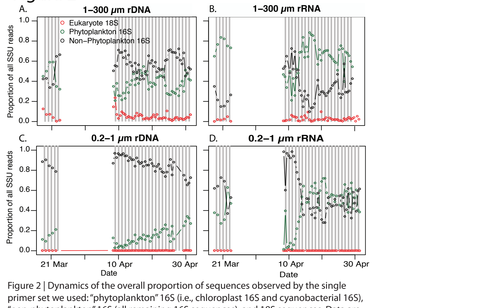


Figure 2 | Dynamics of the overall proportion of sequences observed by the single primer set we used: "phytoplankton" 16S (i.e., chloroplast 16S and cyanobacterial 16S), "non-phytoplankton" 16S (all remaining 16S sequences), and 18S sequences. Data are shown for the A.) 1-300 μm size fraction rDNA, B.) 1-300 μm rRNA, C.) 0.2-1 μm rDNA, and D.) 0.2-1 μm rRNA.

# Figure 3

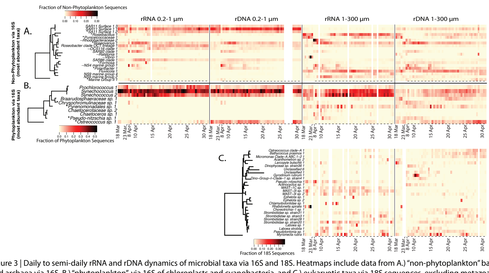


Figure 3 | Daily to semi-daily rRNA and rDNA dynamics of microbial taxa via 16S and 18S. Heatmaps include data from A.) "non-phytoplankton" bacteria and archaea via 16S, B.) "phytoplankton", via 16S of chloroplasts and cyanobacteria, and C.) eukaryotic taxa via 18S sequences, excluding metazoan sequences. Only ASVs (16S) or OTUs (18S) that ever became most abundant in either dataset (rRNA or rDNA) as a fraction within the respective datasets are shown. The tree shows the phylogenetic relatedness of the ASV or OTU according to the amplicon sequence region. Note that *Myrionecta* is known to have a very aberrant 18S sequence (Johnson et al. 2004). For the dates where two samples were taken per day (10:00 AM and 10:00 PM, 10 April - 1 May), a dash underneath a given sample indicates the sample was taken at night. Note, all 16S ASVs shown here were also detected during the 2011 diatom bloom study (Needham and Fuhrman 2016, Needham et al. 2017), except where "--" is found next to the ASV name; asterisks next to taxon names indicate that ASV was also found to most abundant during the 2011 study.

# Figure 4

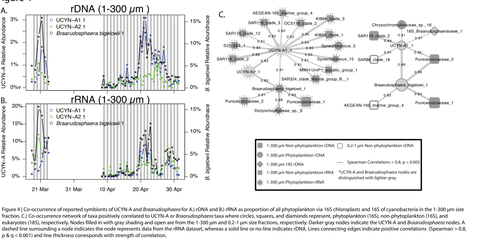


Figure 4 | Co-occurrence of reported symbionts of UCYN-A and *Braarudosphaera* for A.) rDNA and B.) rRNA as proportion of all phytoplankton via 16S chloroplasts and 16S of cyanobacteria in the 1-300 μm size fraction. C.) Co-occurrence network of taxa positively correlated to UCYN-A or *Braarudosphaera* taxa where circles, squares, and diamonds represent, phytoplankton (16S), non-phytoplankton (16S), and eukaryotes (18S), respectively. Nodes filled in with gray shading and open are from the 1-300 μm and 0.2-1 μm size fractions, respectively. Darker gray nodes indicate the UCYN-A and *Braarudosphaera* nodes. A dashed line surrounding a node indicates the node represents data from the rRNA dataset, whereas a solid line or no-line indicates rDNA. Lines connecting edges indicate positive correlations (Spearman > 0.8, p & q < 0.001) and line thickness corresponds with strength of correlation.

# Figure 5

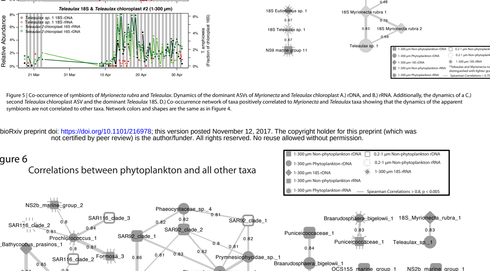


Figure 5 | Co-occurrence of symbionts of *Myrionecta rubra* and *Teleaulax*. Dynamics of the dominant ASVs of *Myrionecta* and *Teleaulax* chloroplast A.) rDNA, and B.) rRNA. Additionally, the dynamics of a C.) second *Teleaulax* chloroplast ASV and the dominant *Teleaulax* 18S. D.) Co-occurrence network of taxa positively correlated to *Myrionecta* and *Teleaulax* taxa showing that the dynamics of the apparent symbionts are not correlated to other taxa. Network colors and shapes are the same as in Figure 4.



# Figure 6

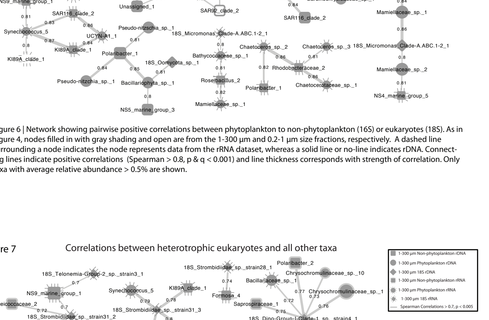


Figure 6 | Network showing pairwise positive correlations between phytoplankton to non-phytoplankton (16S) or eukaryotes (18S). As in Figure 4, nodes filled in with gray shading and open are from the 1-300 μm and 0.2-1 μm size fractions, respectively. A dashed line surrounding a node indicates the node represents data from the rRNA dataset, whereas a solid line or no-line indicates rDNA. Connecting lines indicate positive correlations (Spearman > 0.8, p & q < 0.001) and line thickness corresponds with strength of correlation. Only taxa with average relative abundance > 0.5% are shown.

# Figure 7




Figure 7 | Network showing pairwise positive correlations between heterotrophic eukaryotes to prokaryotes and phytoplankton. Vertical lines surrounding a node indicates the node represents data from the rRNA dataset, whereas no-line indicates rDNA. Lines connecting edges indicate correlations (Spearman > 0.7, p & q < 0.001; no correlations were observed < -0.7 ) and line thickness corresponds with strength of correlation.