Bowman Jeff (Orcid ID: 0000-0002-8811-6280)
Wilson Jesse (Orcid ID: 0000-0003-2837-2060)

Recurrent microbial community types driven by nearshore and seasonal processes in coastal Southern California

Jesse M. Wilson[1], Emelia J. Chamberlain[1], Natalia Erazo[1], Melissa L. Carter[1], Jeff S. Bowman[1,2,3]

[1]Scripps Institution of Oceanography, UCSD

[2]Center for Microbiome Innovation, UCSD

[3]Center for Marine Biodiversity and Conservation, UCSD

Contact: Jesse Wilson

8622 Kennel Way, La Jolla, CA 92037

+1.805.403.6464

jmwilson@ucsd.edu

Running Title: Scales of variation in coastal microbial community

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1462-2920.15548

Originality-Significance Statement

Time-series studies provide insight into how various elements of an ecosystem are connected and the scales on which they naturally vary. We used a sub-weekly, 18-month time-series of the microbial community (assessed via flow cytometry and 16S rRNA gene sequences) alongside many other ecological data in nearshore Southern California to assess drivers of variability in the microbial community utilizing multiple bioinformatic tools. We clearly demonstrated that nearshore Southern California possesses cohesive units of microorganisms that vary seasonally—something that has been demonstrated previously in other ecosystems. However, one particular method showed that a minority group of microorganisms that did not follow a seasonal pattern responded quickly to all influxes of photosynthetic matter, demonstrating that ecologically important groups can turnover rapidly and may never be a dominant cohort.

Summary

A multitude of concurrent biological and physical processes contribute to microbial community turnover, especially in highly dynamic coastal environments. Characterizing what factors contribute most to shifts in microbial community structure and the specific organisms that correlate with changes in the products of photosynthesis improves our understanding of nearshore microbial ecosystem functions. We conducted high frequency sampling in nearshore Southern California in order to capture sub-weekly microbial community dynamics. Microbial communities were characterized by flow cytometry and 16S rRNA gene sequencing, and placed in the context of physicochemical parameters. Within our time-series season and nutrient availability corresponded to changes in dominant microbial community members. Concurrent aseasonal drivers with overlapping scales of variability were also apparent when we used network analysis to assess the microbial community as subsets of the whole. Our analyses revealed the microbial community as a mosaic, with overlapping groups of taxa that varied on different timescales and correlated with unique abiotic and biotic factors. Specifically, a subnetwork associated with chlorophyll *a* exhibited rapid turnover, indicating that ecologically important subsets of the microbial community can change on timescales different than and in response to factors other than those that govern turnover of most members of the assemblage.

Introduction

Marine microbes respond to high and low-frequency variations in the marine environment with shifts in abundance and community structure that have implications for ecosystem function. Shifts in the microbial community are particularly pronounced in highly dynamic nearshore environments, with large variations in both abiotic and biotic variables that are a function of concurrent processes acting across many different spatial and temporal scales (Blain et al., 2004; Martin-Platero et al., 2018; Cocquempot et al., 2019). These include purely physical processes such as tides, advection, and vertical mixing, as well as biological processes such as primary production and grazing. The microbial community is part of this changing seascape, with microbes both responding to and affecting change in their surroundings (Needham et al., 2018; Wilson et al., 2018; Wilson et al., 2020).

Improvements in the temporal frequency, extent, and analysis of microbial time-series allow us to assess variations in the microbial community that result from the balance of concurrent drivers at a range of temporal scales. These include climate trends, seasonality, interactions, and various high-frequency drivers of change. Seasonal environmental variables have been shown to be important in shaping the microbial community during long-term monthly time-series in many marine environments (Gilbert et al., 2012; Giovannoni and Vergin, 2012; Chow et al., 2013). Higher frequency regular (*e.g.*, tide) and sporadic fluctuations (*e.g.*, storm surge and phytoplankton blooms), however, drive substantial variability in coastal systems leading to shifts in microbial community structure and ecological function (Martin-Platero et al., 2018; Needham et al., 2018). Indeed, Hatosy et al. (2013) found that at least 73% of variability in bacterial community structure occurred at intraseasonal temporal scales and both Lindh et al. (2015) and Lambert et al. (2019) showed that greater temporal variability and succession exists

in addition to a slower seasonal shift in the microbial community. These studies demonstrate that both short-term episodic events and long-term seasonal changes are important in structuring the microbial community.

The Ellen Browning Scripps Pier (Scripps Pier) is located within the well-studied and highly variable Southern California Bight ecosystem (Di Lorenzo, 2003; Tai and Palenik, 2009; Tai et al., 2011; McGowan et al., 2017; Nagarkar et al., 2018) and home to the Scripps Ecological Observatory (SEO). This dynamic environment is often influenced by the adjacent California Current ecosystem, with offshore upwelling typically occurring in summer months (Bakun, 1973). The combination of environmental and biological variation in the area provides an ideal framework in which to assess the scales of variation that affect and are affected by the microbial community. Specifically, we combined twice-weekly observations of the microbial community with environmental and biological parameters measured by the National Oceanic and Atmospheric Administration (NOAA) and the Southern California Coastal Ocean Observing System (SCCOOS) to assess seasonal vs. intraseasonal variation in the microbial community according to key ecological drivers. Microbial community observations consisted of flow cytometry-based abundances of autofluorescent (AF) and SYBR-Green I (SG)-stained populations and 16S rRNA gene sequences to evaluate microbial community structure. Microbial community structure characterization at the community level (whole community segmentation) and community subsets (subnetworks) allowed us to assess how different components of the microbial community were affected by various drivers. We expected to observe recurrent seasonal communities in addition to short-term variations driven by both regular (*e.g.*, tide) and sporadic (*e.g.*, winds, advection, phytoplankton, etc.) processes.

Results

Environmental variation

Daily averages of temperature, salinity, water level (measured via hydrostatic pressure and indicative of tidal height), wind-speed, wind direction (Fig. S1), and nutrients (Fig. 1A and B) displayed seasonal and daily variation. A 29-day running temperature average was used to eliminate tidal effects and is indicative of season while temperature, salinity, and water depth at sample time (10:00 AM) were used to indicate short-term environmental drivers. Temperature varied seasonally and with water level (tide), reaching a daily average maximum on 9 August 2018 of 25.9 °C and a minimum on 23 March 2018 of 13.8 °C (Fig. S1A and C). Salinity and water level also each expressed seasonality (reaching maxima in summer for salinity and winter for water level) in addition to monthly and daily variations (Fig. S1B and C). Wind-speed was generally more variable during winter months while wind direction primarily originated from the west during late spring, summer, and fall, and occasionally originated from the east during winter and early spring months (Fig. S1D and E). All nutrients peaked in the spring of 2018 except for ammonium, which peaked in late-winter 2019 (Fig. 1A and 1B). Finally, chlorophyll *a*, phaeophytin, and the microscopic counts of diatoms and dinoflagellates showed periodic and often overlapping peaks (Fig. 1C and D). Chlorophyll *a* and phaeophytin followed the spring and fall bloom pattern (Kim et al., 2009) in 2018 while a bioluminescent red-tide caused a massive chlorophyll *a* and dinoflagellate peak in June of 2019 (Fig. 1C and D).

Variations and correlations in autofluorescent (AF) and SYBR Green 1 (SG)-stained populations

All AF and SG-stained populations showed variation over the 14 months during which flow cytometry data were collected (Fig. 1E and F). In total, five distinct AF populations were observed, representing large eukaryotes, small eukaryotes, two Cyanobacterial groups (1 and 2), and an unknown/detrital group (Fig. 1E). SG-stained populations were classified into three

populations based on cell size (large, small, and other) (Fig. 1F). Smaller AF cells possessed higher overall abundances and peaked in fall for Cyanobacteria 1 and spring and early summer for Cyanobacteria 2, with a dip in both populations in February/March when eukaryotic AF size classes (which peaked during various spring months) were elevated. Many AF and SG-stained populations positively correlated with one another according to Pearson's correlation coefficient (at a Holm-Bonferroni corrected significance of $p \leq 0.00027$; Fig. S2A). For example, all AF eukaryotic and Cyanobacteria populations were positively correlated with one another ($r = 0.36 – 0.55$) and all AF populations except Cyanobacteria 1 were positively correlated with the SG-stained small population ($r = 0.44 – 0.64$). Meanwhile, the AF Cyanobacteria 1 population was positively correlated with the SG-stained large population ($r = 0.48$).

Many of the AF and SG-stained populations correlated with ecological variables according to the same corrected p-value (Fig. S2B). For example, season (the 29-day running temperature average) negatively correlated with the AF small eukaryotic and SG-stained small populations ($r = -0.34 – -0.53$). Additionally, phosphate concentration positively correlated with the AF small eukaryotic population, the natural log of chlorophyll *a* positively correlated with the AF large eukaryotic population, and the SG-stained small population positively correlated with silicate, the natural log of chlorophyll *a*, and dinoflagellates ($r = 0.38 – 0.54$).

Microbial community variation and correlations

We next assessed the microbial community via analysis of 16S rRNA gene sequences. Sequencing produced 2,750,003 copy number corrected reads (with a mean library size of 19,928 ± 12,368) across 138 sampling days. These 16S rRNA gene sequences include both prokaryotes and chloroplasts as both are functionally important parts of the microbial community. We refer to both as photosynthesizers. Following Bowman et al. (2017), bi-weekly sequence data were

first assessed at the whole community level using a self-organizing map (SOM; Fig 2A). The SOM reduced the multidimensional taxonomic dataset to a single categorical variable termed "taxonomic mode" (TM) that represented the community type for a given date. The 5 TMs demonstrated clear seasonality: the "fall/winter" TM appeared during fall and winter (from October through January), the "transition" TM appeared for a brief period in winter and early-spring (from January to early-March), the "spring" TM appeared in the spring (from late-February or April through June), the "summer" TM appeared in summer and early-fall (from late-June through early-October), and the "winter/spring" TM appeared in winter and spring (from January or February through April) (Table 1; Fig. 3A).

Tukey's honest significant difference (HSD) also demonstrated that modes differed with respect to several environmental, biochemical, and biological variables. These included seasonal environmental variables such as temperature at sample time (the "summer" TM correlated with highest temperatures and the "transition" and "winter/spring" TMs correlated with lowest temperatures), hydrostatic pressure (tide or water level) at sample time (the "spring" and "winter/spring" TMs correlated with lower tides), and wind direction (the "spring" and "summer" TMs were associated with westerly winds) (Fig. 2 B-E). Tukey's HSD also noted significant differences between modes according to several nutrients and metrics of the phytoplankton community (Table 1; Fig. S3 and S4). In particular, the "spring" TM had a significantly higher mean than most other TMs for many of the variables tested (nitrite, phosphate, silicate, chlorophyll *a*, phaeophytin, total phytoplankton, diversity [inverse Simpson Index], most AF size classes, diatom abundance, dinoflagellate abundance, and the SG-stained small population). It is important to note, however, that the "spring" TM did not always have the highest concentration/abundance for a given variable. For example, the "winter/spring" TM had

a significantly higher phosphate concentration and the “fall/winter” and “transition” TMs had a significantly higher diversity than the “spring” TM (Table 1; Fig. S3 and S4).

Following Wilson et al. (2018), we next used weighted gene correlation network analysis (WGCNA) to relate subsets (or subnetworks) of co-occurring microbial taxa to environmental, biogeochemical, and biological variables. The two distinct methods used (SOM and WGCNA) work to assess the microbial community in different but complementary ways and reflect sample-centric and taxon-centric approaches. While a SOM sorts entire samples into categorical community types that can be related to environmental data, WGCNA identifies co-occurring microbes throughout the dataset allowing for the coexistence of multiple subnetworks. If minority taxa follow the trends of dominant taxa (which will drive a new SOM type) then the two methods will give the same result. Additionally, WGCNA subnetworks are related to environmental data via the first principal component of a subnetwork’s expression (abundance) matrix, meaning that WGCNA is not dependent on categorical statistical methods. WGCNA identified a total of 7 different subnetworks that varied on both seasonal and non-seasonal timescales (Fig. 3B). Subnetworks also showed varying degrees of correlation with ecological variables (at a Holm-Bonferroni corrected significance of $p \leq 0.00030$; Fig. 4) and taxa within each subnetwork were sorted according to importance or module membership (MM; an MM value close to 1 means a taxon was present and abundant when the subnetwork was prevalent) for further investigation (Table S1).

The dominant subnetwork during summer 2018 transitioned from an “early-summer” to a “late-summer” subnetwork in August. Both subnetworks positively correlated with water temperature at sample time ($\rho = 0.77$) and the SG-stained large population ($\rho = 0.41 – 0.42$). They also both negatively correlated with nitrate, ammonium, phosphate, silicate, phaeophytin,

AF detrital/unknown cells, AF small eukaryotes, and SG-stained small cells ($\rho = -0.77 - -0.35$). In addition to timing, the two summer subnetworks differed in the strength of correlations with the above variables and based on differences with respect to key functional taxonomic groups. For example, using the paprica pipeline (Bowman and Ducklow, 2015; https://github.com/bowmanjeffs/paprica) and the Ribosomal Database Project (RDP) online classifier (Wang et al., 2007) we identified 28% (paprica) to 33% (RDP) of amplicon sequence variants (ASVs) within the "early-summer" subnetwork as photosynthetic (Table S1). However, when we restricted ourselves to top ASVs (MM > 0.50) this increased to 56% (both paprica and RDP). By comparison only 16% (paprica and RDP) of ASVs within the "late-summer" subnetwork were identified as photosynthetic, with 11% (paprica and RDP) of top ASVs representing photosynthetic groups (Table S1). Other microbial groups differed between the two subnetworks as well. For examples, the "late-summer" subnetwork possessed more Flavobacteriales both overall (18% vs. 10% according to paprica and 15% vs. 7% according to RDP) and as top ASVs (22% vs. 0 according to both paprica and RDP).

Meanwhile, there were two main subnetworks observed throughout the non-summer months that overlapped with the "fall/winter" and "winter/spring" TMs (Fig. 3B), and we thus gave these subnetworks the same name as their corresponding modes. The "fall/winter" subnetwork negatively correlated with chlorophyll *a*, phaeophytin, diatoms, and dinoflagellates ($\rho = -0.66 - -0.44$) and positively correlated with taxon richness and diversity (inverse Simpson Index) ($\rho = 0.60 - 0.64$). This subnetwork possessed 90% (paprica) to 86% (RDP) of all Pelagibacterales/SAR11 in the entire dataset and 72% (paprica) to 68% (RDP) of Thermoplasmata in the entire dataset. Meanwhile, only 15% (paprica) to 17% (RDP) of ASVs were photosynthetic (Table S1). Conversely, the "winter/spring" subnetwork positively

correlated with nitrate, phosphate, and silicate ($\rho = 0.40 – 0.57$), and negatively correlated with temperature ($\rho = -0.67$) (Fig. 4). The "winter/spring" subnetwork also positively correlated with chlorophyll *a*, phaeophytin, diatoms, dinoflagellates, small eukaryotic and detrital/unknown AF populations, and SG-stained small cells ($\rho = 0.33 – 0.53$). Photosynthesizers increased relative to the "fall/winter" subnetwork as 31% (paprica) to 33% (RDP) of ASVs were classified as photosynthetic (Table S1). 16% (paprica) to 15% (RDP) of ASVs were identified as Flavobacteriales, but this number increased when only top ASVs were assessed (25% for paprica and 23% for RDP). Another subnetwork appeared during the period that the "winter/spring" subnetwork dominated (Fig. 3B). The "spring" subnetwork positively correlated with a subset of the variables that the "winter/spring" subnetwork also correlated with: silicate, phaeophytin, dinoflagellates, the large eukaryotic and Cyanobacteria 2 AF populations, and taxon richness ($\rho = 0.30 – 0.54$) (Fig. 4). However, only 18% (paprica) to 23% (RDP) of ASVs, or 17% of top ASVs (for both paprica and RDP) were identified as photosynthetic (Table S1). Other important top ASVs included Flavobacteriales (26% via paprica and 23% via RDP) and Rhodobacterales (6% via both paprica and RDP).

Finally, the "chlorophyll *a* associated" and "other" subnetworks did not follow clear seasonal patterns like the other subnetworks (Fig. 3B). Despite only possessing 19% (paprica) to 20% (RDP) photosynthetic ASVs (9% for both paprica and RDP when restricted to top taxa), the "chlorophyll *a* associated" subnetwork correlated with chlorophyll *a*, dinoflagellates, the large eukaryotic and Cyanobacteria 2 AF populations, and the SG-stained small population ($\rho = 0.38 – 0.57$) (Fig. 4). This subnetwork possessed a high proportion of Flavobacteriales (43% for both paprica and RDP) and Puniceicoccales (14% for both paprica and RDP) as top ASVs. The "chlorophyll *a* associated" subnetwork also synced with minor non-seasonal peaks that the

"spring" subnetwork possessed and displayed a spike immediately following the seasonal "spring" subnetwork peak in 2019 when there was a bioluminescent red-tide (consisting of the dinoflagellate, *Lingulodinium polyedrum*; M. Carter, personal communication) (Fig. 3B). Meanwhile, the "other" subnetwork was typically one of the least abundant subnetworks with the exception of a brief period in the winter of 2018 (Fig. 3B). This brief spike led to it negatively correlating with temperature ($\rho = -0.43$). Due to the fact that this time period overlapped with the "fall/winter" subnetwork, the "other" subnetwork positively correlated with taxon richness and diversity ($\rho = 0.55$) (Fig. 4).

Discussion

This work represents a unique high-frequency, long-term time-series within the nearshore environment of the California Bight. Through this time-series we were able to assess timescales of variation of the microbial community based on different flow cytometric populations (AF and SG-stained populations) and taxonomic groupings (segmenting based on sample similarity at the whole community level and grouping subsets of taxa according to co-occurrence). Similar to other time-series work, we noted cohesive units of the microbial community that varied both seasonally and intraseasonally (Lindh et al., 2015; Martin-Platero et al., 2018; Needham et al., 2018; Lambert et al., 2019). Additionally, we were able to 1) relate microbial flow cytometric populations and 16S rRNA gene-based subnetworks to ecological variables, and 2) assess the importance of individual taxa to changes in community structure in response to environmental, biochemical, and biological variables. We note that relatively few flow cytometric populations varied with season or in response to environmental factors that varied on other timescales, though many correlated with one another (Fig. S2). Community segmentation with a SOM revealed seasonality between community types. Potential drivers for this ecological variation

were often unclear until we assessed subsets of the entire community using WGCNA, which reinforced the pattern of the SOM-based TMs and also revealed less abundant groups that were changing on different timescales than the most abundant community members (Fig. 3).

Relatively few flow cytometric populations correlated with seasonal and intraseasonal environmental drivers because the phytoplankton community followed different patterns in 2018 and 2019. Kim et al. (2009) noted that sea surface temperature within the Southern California Bight follows clear seasonality while phytoplankton blooms (measured via chlorophyll *a*) are irregular. Specifically, Kim et al. (2009) identified three possible bloom patterns (spring, spring and fall, or summer) and our data followed the spring and fall pattern in 2018. Meanwhile, 2019 was dominated by a bioluminescent dinoflagellate bloom in June. SG-stained small heterotrophs were extremely prevalent before and during the 2019 bioluminescent bloom as various AF size groups peaked, suggesting that they responded rapidly to increased substrates.

Expanding on these relationships using 16S rRNA gene-derived community structure revealed seasonal variation at the whole community level (community segmentation) and both seasonal and intraseasonal variation in subsets of the community (WGCNA). Both methods have been successful at describing temporal variations in the microbial community (Bowman et al., 2017; Wilson et al., 2018). In our dataset, the two methods agreed surprisingly well with persistent assemblages lasting weeks to months during particular seasons (Fig. 3). However, WGCNA demonstrated that many subnetworks persisted at lower abundances while the dominant subnetwork shifted with a distinct seasonal pattern (which caused transitions between TMs). Additionally, the "chlorophyll *a* associated" subnetwork exhibited rapid turnover in response to photosynthetic groups. This indicates that ecologically important subsets of the microbial community can change on different timescales and are driven by different factors than

those governing turnover in the majority of the assemblage. Additionally, because WGCNA correlates each taxon to its subnetwork (with MM, see Results for description) we identified taxonomic groups that drove shifts in subnetwork abundance over time.

The strongly seasonal summer community/communities identified by community segmentation and WGCNA were distinct from all other communities and were most likely driven by a shift in conditions and available substrates. Growth rates tend to increase with increased substrate availability and temperature (Huete-Stauffer et al., 2015), something that was apparent in our dataset with increased cell abundances during summer months (Fig. 1E and F) and the correlation between both summer subnetworks and the SG-stained large population (Fig. 4). The fact that diversity decreased for the "summer" TM (Table 1; S3I) indicates that these ASVs were generalists. Mou et al. (2008) noted a large number of generalist species present in a coastal environment and attributed this to the heterogeneity found in the composition of organic matter and Chen et al. (2020) found that metabolic flexibility allowed generalists to dominate coastal sediments that were frequently disturbed. The summer community/communities dominated after the spring and early-summer spike in chlorophyll *a* and AF populations, which may have provided extremely varied but labile substrates that metabolic generalists could take advantage of. Additionally, the bioluminescent red-tide in 2019 may have provided ample substrates for generalist species. Therefore, while the highest degree of physical disturbance occurs during winter, the variable phytoplankton blooms and grazing pressures present in the summer are forms of ecosystem variability that generalists would have an easier time acclimating to.

WGCNA resolved the singular summer mode into two different summer subnetworks that co-occurred but dominated at different times and separated based on basic life history traits.

Members of the "early-summer" subnetwork reflected the more readily available resources/niches, which were utilized by the high proportion of photosynthesizers in this subnetwork (Table S1). Meanwhile, the "late-summer" subnetwork was characterized by organisms that could process the influx of organic matter from the "early-summer" subnetwork. In addition to possessing fewer photosynthesizers likely from different clades as the "early-summer" subnetwork (Tai and Palenik, 2009; Tai et al., 2011; Nagarkar et al., 2020), the "late-summer" subnetwork had a higher proportion of Flavobacteriales (Table S1). Flavobacteriales have been found to be important to processing polysaccharides derived from phytoplankton blooms (Kirchman 2002; Avci et al., 2020; Ferrer-Gonzalez et al., 2020).

Seasonality was also apparent in many of the non-summer TMs and subnetworks, which we attribute to the availability of substrates and nutrients. For example, the "fall-winter" TM and subnetwork dominated during winter, which is generally a less productive time (Fargion et al., 1993; Kim et al., 2009; Nagarkar et al., 2018; Wilson et al., 2020). The decreased resources (*e.g.*, nutrients) translated to lower chlorophyll *a*, eukaryotic phytoplankton populations, and an absence of large AF populations (Fig. 1A-E, 3, and 4). The fact that Pelagibacterales were most prevalent in the "fall/winter" TM and subnetwork reflects this limitation as Pelagibacterales are characterized by their small genomes, an adaptation that enables them to persist in low nutrient regimes (Giovannoni et al., 2005; Mende et al. 2017; Shenhav and Zeevi, 2020).

The increase in nutrients in the spring marked a shift to the "winter/spring" TM and subnetwork, with the "transition" TM marking a transitional phase when both the "fall/winter" and "winter/spring" subnetworks were both fairly abundant. The "winter/spring" and "transition" TMs and "winter/spring" subnetwork may reflect open niches that fast-growing species were able to exploit. This is evident in the rapid increase in abundance of many AF populations and

smaller SG-stained cells (Fig. 1E, 1F, 4, and S4). The "winter/spring" subnetwork also had high rates of photosynthesizers and—as found during the productive early summer period—we observed a high proportion of Flavobacteriales and other groups commonly associated with processing primary production products. Multiple ASVs were also associated with the gammaproteobacterial family Halieaceae (MM = 0.73 – 0.78), for which one ASV had 100% similarity to a sequence (OM-RGC.v1.001592647) that was associated with high productivity in the Tara Ocean's database (Villar et al., 2018).

Towards the end of the period when the "winter/spring" subnetwork dominated, the highly seasonal "spring" subnetwork peaked, though it never became dominant and appears to have been responding to the photosynthetic products or environmental conditions present at that time. This is because fewer ASVs were identified as photosynthesizers than the "winter/spring" or "early summer" subnetworks despite increasing light levels. However, many top members belonged to the orders Flavobacteriales and Rhodobacterales, which have both been identified as important to the processing of phytoplankton biomass and exudates (Ferrer-Gonzalez et al., 2020).

WGCNA also identified intraseasonal variation in the microbial community, which corresponded with opportunistic non-seasonal variation in the photosynthetic community. This variation was not apparent in the SOM-based modes because it consisted of minor portions of the microbial community. Despite often low or infrequent abundances of individual microbes, these microbes may be extremely important to processing photosynthetically derived organic matter as this subnetwork correlated strongly with dinoflagellate counts and multiple AF populations (large eukaryotes and Cyanobacteria 2) (Fig. 4). Reinforcing this idea is the fact that this subnetwork had its greatest peak in June 2019 when there was a bioluminescent red-tide (Fig.

1C, 1D, and 3B) explaining why the "chlorophyll *a* associated" subnetwork correlated more strongly with dinoflagellates and the large eukaryotic AF population than any other subnetwork. The idea that a few rapidly growing heterotrophs may be responsible for processing a majority of the labile organic matter in an area is not new (Mou et al., 2008; Gómez-Consarnau et al., 2012; Nelson and Weir, 2014; Pedler et al., 2014), but the idea that a cohesive group of microorganisms remain poised to respond to influxes of phytoplankton derived organic matter against a backdrop of seasonal succession in the majority fraction of the microbial community has not been widely demonstrated. Examples of rapidly growing copiotrophs within the "chlorophyll *a* associated" subnetwork included several taxa identified as belonging to the family Puniceicoccaceae (MM = 0.51 – 0.64) and many different important Flavobacteriales taxa (MM = 0.50 – 0.69). Puniceicoccaceae are aerobic chemoorganotrophic heterotrophs that have been found in a range of marine habitats (Cho et al., 2011) and, as stated before, members of Flavobacteriales have been found to be important to processing polysaccharides derived from phytoplankton blooms.

There were some broad-scale taxonomic distinctions between the seasonal subnetworks and the aseasonal "chlorophyll *a* associated" subnetwork that provide insight into the taxa most important to responding to influxes of photosynthetic derived organic matter in this system. All of the seasonal subnetworks but the "early-summer" subnetwork had multiple instances of both Rhodobacterales and Flavobacteriales as top members (MM > 0.50; according to both paprica and RDP), with the "early-summer" subnetwork not possessing any Flavobacteriales (Table S1). Conversely, the "chlorophyll *a* associated" subnetwork did not possess any Rhodobacterales as top members, but had many Flavobacteriales represented (Table S1). While both orders are commonly seen following phytoplankton blooms, many studies have reported a succession

between major bacterial groups (Teeling et al., 2012; Buchan et al., 2014), with members of Flavobacteria identified as r-growth strategists often observed immediately following phytoplankton blooms and specializing in the consumption of complex organic matter (Pinhassi et al., 2004; Edwards et al., 2010; Gómez-Pereira et al., 2011; Thomas et al., 2011; Teeling et al., 2012). Thus, their association with pulses of organic matter (and the "chlorophyll *a* associated" subnetwork) makes sense and provides insight into the specific taxa important to degrading sporadic influxes of organic matter in coastal Southern California. Meanwhile, Rhodobacterales are also frequently observed following phytoplankton blooms, but are thought to be important to the consumption of low molecular weight phytoplankton metabolites (Landa et al., 2017; Ferrer-Gonzalez et al., 2020). Thus, they appear to be associated with seasonal primary producers that were present for prolonged periods of time and were part of the clear seasonal succession between various taxa. The specialization of various taxa furthers the idea that niche portioning (Teeling et al., 2012; Buchan et al., 2014; Ferrer-Gonzalez et al., 2020) is important in the processing of phytoplankton derived organic matter in this system.

Overall, the variety of microbial flow cytometric populations, SOM-based community types, and WGCNA-based subnetworks that were present in our time-series demonstrate that many different scales of variation drive shifts in the microbial community within nearshore coastal California. Our overall findings are in general agreement with previous work that found both seasonal and intraseasonal variation in cohesive units of the microbial community in dynamic marine environments (Lindh et al., 2015; Martin-Platero et al., 2018; Needham et al., 2018; Lambert et al., 2019). Within our time-series, season was the dominant factor that drove changes in interacting groups of microorganisms whose members often represented the most abundant taxa. Both whole community segmentation and subnetwork analysis picked up on this

level of variation. It should also be noted that while these dominant microbial cohorts follow a clear seasonal pattern, it remains to be shown if there are additional patterns according to interdecadal or multi-year modulation. For example, the timing of transitions may shift or additional subnetworks/community types may appear during El Niño years—something that will be resolved in the future from the ongoing SEO time-series. Regardless, associations between the microbial community and certain ecological factors (*i.e.* those involving primary producers), and the succession in the summer community, were not as apparent at the whole community level. The correlation between a minority subnetwork (the "chlorophyll *a* associated" subnetwork) and opportunistic primary producers shows that there are differential drivers with overlapping scales of variability for the microbial community. It also shows that the microbial community can display rapid turnover in response to ecological stimuli, but that not all members of the community will respond similarly. These results reinforce the use of a SOM for identifying large scale seasonal changes in the microbial community in dynamic environments, but also indicate that potentially ecologically important subsets of the microbial community can change on different scales.

Experimental Procedures

Sample Collection

All data were collected from the Ellen Browning Scripps Pier in San Diego, CA (32.8663° N, 117.2546° W) from 4 January 2018 through 6 June 2019. SEO (www.ecoobs.ucsd.edu) conducts bi-weekly analysis of the microbial community in conjunction with the SCCOOS Scripps Pier Shore Station sampling effort. For each sampling, surface water was collected using a bucket off the Scripps Pier at approximately 10:00 AM and transferred to acid washed containers that were brought back to the laboratory for immediate analysis. Bi-

weekly concentrations of nutrients, chlorophyll *a*, and phaeophytin, and weekly counts of diatoms, dinoflagellates, and total phytoplankton were provided by SCCOOS (https://sccoos.org/) and the McGowan Plankton and Chlorophyll Program (see, Wilson et al., 2020 for methods). Continuous temperature, water depth (hydrostatic pressure), and salinity at Scripps Pier were provided by the SCCOOS Automated Shore Station Program (https://erddap.sccoos.org/erddap/tabledap/autoss.html). Meteorological data were provided by the NOAA station LJAC1 (https://www.ndbc.noaa.gov). Bacterial and archaeal community structure was determined through analysis of the 16S rRNA gene, while flow cytometry was used to quantify cells belonging to different AF and SG-stained populations.

DNA Collection & Bioinformatics

Seawater was filtered through a sterile 0.2 µm Supor membrane disc filter (Pall Corporation, Port Washington, NY, USA) and stored at −80 ºC until extraction. Through 8 January 2019, filters were extracted manually using the MoBio DNEASY PowerWater Kit (Qiagen, Venlo, Netherlands), after which point filters were extracted using the KingFisher™ Flex Purification System and MagMax Microbiome Ultra Nucleic Acid Extraction kit (ThermoFisher Scientific, Waltham, Massachusetts, USA). Extracted DNA was sent to Argonne National Laboratory for amplicon library preparation and sequencing using the Illumina MiSeq platform, universal primers 515F and (modified) 806R (Walters et al., 2016), and 2 x 151 bp library architecture. Reads were filtered, denoised, and merged with dada2 (Callahan et al., 2016). Merged reads were analyzed with paprica v0.7.0 (Bowman and Ducklow, 2015; https://github.com/bowmanjeffs/paprica) and Ribosomal Database Project (RDP) online classifier (Wang et al., 2007; version 11) to determine the community structure. Paprica utilizes phylogenetic placement (Barbera et al., 2019) to place query reads on a reference tree

constructed from the full-length 16S or 18S rRNA genes from all completed genomes in GenBank. All unique reads were assigned to internal branches or terminal branches on the reference tree. Sequences were submitted to NCBI SRA at BioProject PRJNA662174. Diversity was calculated for each sample using the inverse Simpson Index.

Flow cytometry

Flow cytometry samples were aliquoted into 3 mL sample tubes and fixed (final concentration 2% HCHO before 7 January 2019, or 0.25% glutaraldehyde after) before running on a CyFlow Space (Sysmex America Inc., Lincolnshire, IL, USA). Fluorescent signals of AF (unstained) samples were measured for forward scatter, side scatter, FL 1 (em. 536 nm), FL 2 (em. 590 nm), FL 3 (em. 675 nm), and FL 4 (em. 748 nm) for excitation with a 488 nm laser, FL 5 (em. 455 nm) for excitation with the 405 nm laser, and FL 6 (em. 675 nm) for excitation with the 638 nm laser. Duplicate samples were stained with SG (Molecular Probes Inc., Eugene, OR, USA) prior to analysis. Absolute cell counts were determined by spiking a standard volume of 1:10 diluted 123-count beads (ThermoFisher, Waltham, Massachusetts, USA) to each sample and blank.

Subgroups were identified using a SOM from forward scatter, side scatter, and FL 2 – FL 6 for AF communities, and from forward scatter, side scatter, FL 1, and FL 5 for SG-stained communities following Bowman et al. (2017) (Fig. 1E and F) This analysis was conducted using the 'kohonen' package in R (Wehrens and Kruisselbrink, 2018). In brief, a training set was constructed by randomly sampling 50 events from each sample (10,600 total events). These data were trained using a toroidal map grid of size 41 x 41. Populations were identified using k-means clustering (AF k = 6, SG-stained k=3) with k chosen through the visual evaluation of a within-cluster sum of squares scree plot and *a priori* knowledge of populations.

Determination of taxonomic modes and subnetworks

Following Bowman et al. (2017), a SOM was used to reduce the multidimensional taxonomic dataset to a single categorical variable for each date ('kohonen' package in R version 3.0.8; Wehrens and Kruisselbrink, 2018). We input a Hellinger-transformed relative abundance matrix of all 6,102 unique reads across 130 of the 138 days (days that had > 5,000 reads). During SOM training we varied the SOM map units between grids sized 5x5 to 9x9 and settled on a 7x7 toroidal grid based on the distribution of the number of samples (130 days) assigned to each map unit 1000 times (Fig. 2A). K-means clustering was used to segment map units into TMs, with final k selected based on a within-clusters sum of squares scree plot and experimentally varying k around the perceived optimum before settling on the final segments. The resulting TMs were related to environmental variables and different AF and SG-stained populations using analysis of variance (ANOVA) and Tukey's HSD test ('stats' package in R version 3.5.1; R Core Team, 2018) to assess whether the means for each variable were statistically different for various TMs.

Following Wilson et al. (2018), WGCNA was used to relate subnetworks of microbial taxa to environmental and ecological data and to different AF and SG-stained populations ('WGCNA' package in R version 1.68; Langfelder and Horvath, 2007, 2008). WGCNA assesses the co-occurrence of taxa via an adjacency function that factors in the degree of shared neighbors between two taxa magnified by a power-law function so that the topology of the graph becomes scale-free. After taxa are sorted into subnetworks, the co-occurrence profiles for each subnetwork (*i.e.* the first principal component of that subnetwork) are then related to variables. We again used a Hellinger-transformed relative abundance matrix (that factored in all 6,102 unique taxa across all 138 days) and then limited the matrix to the 1,541 most abundant taxa (1,512 Bacteria and 29 Archaea) so that the topology overlap measure (TOM) would fit. We

utilized a signed adjacency matrix and selected a soft thresholding power of 7 which gave an r^2 of 0.959 for the TOM model fit and a mean connectivity between taxa of 22.6. We set our minimum module (subnetwork) size to 50. The co-occurrence expression profiles for each subnetwork were then related to other variables using the first principal component of that subnetwork by way of a Spearman Rank Correlation.

Acknowledgments

The authors would like to thank Elizabeth Connors, Avishek Dutta, Melissa Hopkins, Gabriella Berman, and Mackenzie Davey for their assistance with microbial and flow cytometry sample collection and processing. We would also like to thank Kristi Seech for her assistance with SCCOOS phytoplankton cell identification and enumeration and John A. McGowan for support and leadership over the decades with the McGowan Plankton and Chlorophyll Program (Funding provided by private donors and the MacArthur Foundation) and Southern California Coastal Ocean Observing Harmful Algal Bloom Monitoring Program (NOAA NA16NOS0120022, NA11NOS120029, and NA17RJ1231). Finally, the authors would like to thank Sarah Abboud for her assistance with the manuscript. This work was supported by a Simons Foundation Early Career Marine Microbial Ecology and Evolution award to JSB.

References

Avci, B., Kruger, K., Fuchs, B.M., Teeling, H., and Amann, R.I. (2020) Polysaccharide niche partitioning of distinct Polaribacter clades during North Sea spring algal blooms. *ISME J* 13699–1383.

Bakun, A. (1973) Coastal upwelling indices, west coast of North America, 1946-71. U.S. Dep. Commer. NOAA Tech. Rep. NMFS SSRF-671, 103 p.

Barbera, P., Kozlov, A.M., Czech, L., Morel, B., Darriba, D., Flouri, T., and Stamatakis, A. (2019) EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. *Syst Biol* **68**: 365–369.

Blain, S., Guieu, C., Claustre, H., Leblanc, K., Moutin, T., Quèguiner, B., et al. (2004) Availability of iron and major nutrients for phytoplankton in the northeast Atlantic Ocean. *Limnol Oceanogr* **49**: 2095–2104.

Bowman, J.S., Amaral-Zettler, L.A., Rich, J.J., Luria, C.M., and Ducklow, H.W. (2017) Bacterial community segmentation facilitates the prediction of ecosystem function along the coast of the western Antarctic Peninsula. *ISME J* **11**: 1460–1471.

Bowman, J.S. and Ducklow, H.W. (2015) Microbial communities can be described by metabolic structure: A general framework and application to a seasonally variable, depth-stratified microbial community from the coastal West Antarctic Peninsula. *PLoS One* **10**.

Buchan, A., LeCleir, G.R., Gulvik, C.A., and Gonzalez, J.M. (2014) Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol* **12**: 686–698.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581–583.

Chen, Y.-J., Leung, P.M., Bay, S., Hugenholtz, P., Kessler, A., Shelley, G., et al. (2020) Metabolic flexibility allows generalist bacteria to become dominant in a frequently disturbed ecosystem. *bioRxiv* 2020.02.12.945220.

Cho, J.-C., Yoon, J., and Hedlund, B. (2011) Family I. Puniceicoccaceae. *Bergey's Man Syst Bacteriol* **4**.

Chow, C.E.T., Sachdeva, R., Cram, J.A., Steele, J.A., Needham, D.M., Patel, A., et al. (2013) Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. *ISME J* **7**: 2259–2273.

Cocquempot, L., Delacourt, C., Paillet, J., Riou, P., Aucan, J., Castelle, B., et al. (2019) Coastal ocean and nearshore observation: A French case study. *Front Mar Sci* **6**.

Di Lorenzo, E. (2003) Seasonal dynamics of the surface circulation in the Southern California Current System. *Deep Res Part II Top Stud Oceanogr* **50**: 2371–2388.

Edwards, J.L., Smith, D.L., Connolly, J., McDonald, J.E., Cox, M.J., Joint, I., et al. (2010) Identification of Carbohydrate Metabolism Genes in the Metagenome of a Marine Biofilm Community Shown to Be Dominated by Gammaproteobacteria and Bacteroidetes. *Genes (Basel)* **1**: 371–384.

Fargion, G.S., Mcgowan, J.A., and Stewart, R.H. (1993) Seasonality of chlorophyll concentrations in the California Current: a comparison of two methods. *CalCOFI Rep* **34**: 35–50.

Ferrer-Gonzalez, F.X., Widner, B., Holderman, N.R., Glushka, J., Edison, A.S., Kujawinski, E.B., and Moran, M.A. (2020) Resource partitioning of phytoplankton metabolites that support bacterial heterotrophy. *ISME J*.

Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbruck, L., Reeder, J., Temperton, B., et al. (2012) Defining seasonal marine microbial community dynamics. *ISME J* **6**: 298–308.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., et al. (2005) Genetics: Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–1245.

Giovannoni, S.J. and Vergin, K.L. (2012) Seasonality in ocean microbial communities. *Science* **335**: 671–676.

Giovannoni, S.J., Thrash, J.C., and Temperton, B. (2014) Implications of streamlining theory for microbial ecology. *ISME J* **8**: 1553–1565.

Gómez-Consarnau, L., Lindh, M. V., Gasol, J.M., and Pinhassi, J. (2012) Structuring of bacterioplankton communities by specific dissolved organic carbon compounds. *Environ Microbiol* **14**: 2361–2378.

Gómez-Pereira, P.R., Schüler, M., Fuchs, B.M., Bennke, C., Teeling, H., Waldmann, J., et al. (2012) Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. *Environ Microbiol* **14**: 52–66.

Hatosy, S.M., Martiny, J.B.H., Sachdeva, R., Steele, J., Fuhrman, J.A., and Martiny, A.C. (2013) Beta diversity of marine bacteria depends on temporal scale. *Ecology* **94**: 1898–1904.

Huete-Stauffer, T. M., Arandia-Gorostidi, N., Díaz-Pérez, L., & Moran, X. A. G. (2015) Temperature dependences of growth rates and carrying capacities of marine bacteria depart from metabolic theoretical predictions. *FEMS microbiology ecology* **91**: 1-10.

Kim, H.J., Miller, A.J., McGowan, J., and Carter, M.L. (2009) Coastal phytoplankton blooms in the Southern California Bight. *Prog Oceanogr* **82**: 137–147.

Kirchman, D.L. (2002) The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol Ecol* **39**: 91–100.

Lambert, S., Tragin, M., Lozano, J.C., Ghiglione, J.F., Vaulot, D., Bouget, F.Y., and Galand, P.E. (2019) Rhythmicity of coastal marine picoeukaryotes, bacteria and archaea despite irregular environmental perturbations. *ISME J* **13**: 388–401.

Landa, M., Burns, A.S., Roth, S.J., and Moran, M.A. (2017) Bacterial transcriptome remodeling during sequential co-culture with a marine dinoflagellate and diatom. *ISME J* **11**: 2677–2690.

Langfelder, P. and Horvath, S. (2007) Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol* **1**.

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

Lindh, M. V., Sjöstedt, J., Andersson, A.F., Baltar, F., Hugerth, L.W., Lundin, D., et al. (2015) Disentangling seasonal bacterioplankton population dynamics by high-frequency sampling. *Environ Microbiol* **17**: 2459–2476.

Martin-Platero, A.M., Cleary, B., Kauffman, K., Preheim, S.P., Mcgillicuddy, D.J., Alm, E.J., and Polz, M.F. (2018) High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nat Commun* **9**: 1–11.

McGowan, J.A., Deyle, E.R., Ye, H., Carter, M.L., Perretti, C.T., Seger, K.D., et al. (2017) Predicting coastal algal blooms in southern California. *Ecology* **98**: 1419–1433.

Mende, D.R., Bryant, J.A., Aylward, F.O., Eppley, J.M., Nielsen, T., Karl, D.M., and DeLong, E.F. (2017) Environmental drivers of a microbial genomic transition zone in the ocean's interior. *Nat Microbiol* **2**: 1367-1373.

Mou, X., Sun, S., Edwards, R.A., Hodson, R.E., and Moran, M.A. (2008) Bacterial carbon processing by generalist species in the coastal ocean. *Nature* **451**: 708–713.

Nagarkar, M., Countway, P.D., Du Yoo, Y., Daniels, E., Poulton, N.J., and Palenik, B. (2018) Temporal dynamics of eukaryotic microbial diversity at a coastal Pacific site. *ISME J* **12**: 2278–2291.

Nagarkar, M., Wang, M., Valencia, B., and Palenik, B. (2020) Spatial and temporal variations in *Synechococcus* microdiversity in the Southern California coastal ecosystem. *Environ Microbiol* 1462-2920.15307.

Needham, D.M., Fichot, E.B., Wang, E., Berdjeb, L., Cram, J.A., Fichot, C.G., and Fuhrman, J.A. (2018) Dynamics and interactions of highly resolved marine plankton via automated high-frequency sampling. *ISME J* **12**: 2417–2432.

Nelson, C.E. and Wear, E.K. (2014) Microbial diversity and the lability of dissolved organic carbon. *Proc Natl Acad Sci U S A* **111**: 7166–7167.

Park, S., Yoshizawa, S., Inomata, K., Kogure, K., and Yokota, A. (2012) *Halioglobus japonicus* gen. nov., sp. nov. and *halioglobus pacificus* sp. nov., members of the class gammaproteobacteria isolated from seawater. *Int J Syst Evol Microbiol* **62**: 1784–1789.

Pedler, B.E., Aluwihare, L.I., and Azam, F. (2014) Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proc Natl Acad Sci U S A* **111**: 7202–7207.

Pinhassi, J., Sala, M.M., Havskum, H., Peters, F., Guadayol, Ò., Malits, A., and Marrasé, C. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. *Appl Environ Microbiol* **70**: 6753–6766.

R Studio Team. 2018. RStudio: Integrated development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/

Shenhav, L. and Zeevi, D. (2020) Resource conservation manifests in the genetic code. *Science* **370**: 683-687.

Tai, V., Burton, R., and Palenik, B. (2011) Temporal and spatial distributions of marine Synechococcus in the Southern California Bight assessed by hybridization to bead-arrays. *Mar Ecol Prog Ser* **426**: 133–147.

Tai, V. and Palenik, B. (2009) Temporal variation of Synechococcus clades at a coastal Pacific Ocean monitoring site. *ISME J* **3**: 903–915.

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

Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M., and Michel, G. (2011) Environmental and Gut Bacteroidetes: The Food Connection. *Front Microbiol* **2**: 93.

Villar, E., Vannier, T., Vernette, C., Lescot, M., Cuenca, M., Alexandre, A., et al. (2018) The Ocean Gene Atlas: Exploring the biogeography of plankton genes online. *Nucleic Acids Res* **46**: W289–W295.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., et al. (2016) Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **1**.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Wehrens, R. and Kruisselbrink, J. (2018) Flexible Self-Organizing Maps in kohonen 3.0. *J Stat Softw* **87**.

Wilson, J.M., Carter, M.L., Mühle, J., and Bowman, J.S. (2020) Using empirical dynamic modeling to assess relationships between atmospheric trace gases and eukaryotic phytoplankton populations in coastal Southern California. *Mar Chem* **227**: 103896.

Wilson, J.M., Litvin, S.Y., and Beman, J.M. (2018) Microbial community networks associated with variations in community respiration rates during upwelling in nearshore Monterey Bay, California. *Environ Microbiol Rep* **10**: 272–282.

Zhao, Z., Shen, X., Chen, W., Yu, X.Y., Fu, G.Y., Sun, C., and Wu, M. (2018) *Emcibacter congregatus* sp. Nov., isolated from sediment cultured in situ. *Int J Syst Evol Microbiol* **68**: 2846–2850.

Table 1) Environmental and biological characteristics of self-organizing map (SOM)-based taxonomic modes. See Fig. 2 for SOM grid and Fig. S3 and 4 for visualization of these characteristics.

Fig. 1) Time-series of bi-weekly sampled concentrations for A) nitrate (dark green triangles), nitrite (green circles), B) ammonium (orange squares), phosphate (purple circles), silicate (red triangles), C) chlorophyll *a* (light green diamonds), and phaeophytin (dark green circles). D) Time-series of the abundances of once weekly measured diatoms (blue triangles) and dinoflagellates (green circles). Timeseries of bi-weekly sampled E) autofluorescent (AF) populations (detrital/unknown in red, large eukaryotes in orange, small eukaryotes in green, Cyanobacteria 1 in blue, and Cyanobacteria 2 in purple) and F) SYBR-Green 1 (SG)-stained large (black), small (brown), and other (grey) populations. A and E insets show flow cytometric populations (selected via a self-organizing map—see methods) on 2 August 2018 for AF populations and on 13 December 2018 for SG-stained populations. Flow cytometry analyses began in May of 2018 while all other time-series began in January 2018.

Fig. 2) Self-organizing map (SOM) analysis development and comparison of taxonomic modes (TMs) with environmental variables. A) SOM grid (7x7) with samples (dots) sorted into different map units. K-means clustering partitioned map units into like TMs, shown by both the color of the map unit and the thick lines surrounding groups of map units. The map was arranged as a toroid, so units on opposite sides of the grid are touching. B) Temperature at sample time for each TM (seasonal temperature was similar). C) Hydrostatic pressure at sample time for each TM. D) Daily average wind speed for each TM. E) Daily average wind direction for each TM. For each boxplot, the horizontal line represents the median (and is colored according to the TM it belongs to), the grey box spans the first quartile to the third quartile, the whiskers show the maximum and minimum, and any dots represent outliers. Tukey's HSD found differences in the means for temperature at sample time between all TMs except for the "fall/winter" vs. the "spring" TMs and between the "transition" vs. the "winter/spring" TMs; hydrostatic pressure at sample time between the "fall/winter" vs. the "spring" and "winter/spring" TMs, between the "transition" vs. the "spring" and "winter/spring" TMs, and between the "spring" vs. the "summer" TMs; and for daily average wind direction between the "spring" and "summer" vs. all other TMs. There were no significant differences between the means for wind speed.

Fig. 3) Time-series of the microbial community using different methods. A) Time-series of taxonomic modes (TMs) identified using a self-organizing map (SOM) on all 6,102 Hellinger Transformed 16S rRNA sequences (whole community segmentation). B) Time-series of abundances of subnetworks selected via weighted gene correlation network analysis (WGCNA) on Hellinger Transformed relative abundances of the 1,541 most abundant 16S rRNA sequences when the miniature module size was set at 50. Only taxa that had a p-value of ≤ 0.05 were summed to create the WGCNA time-series.

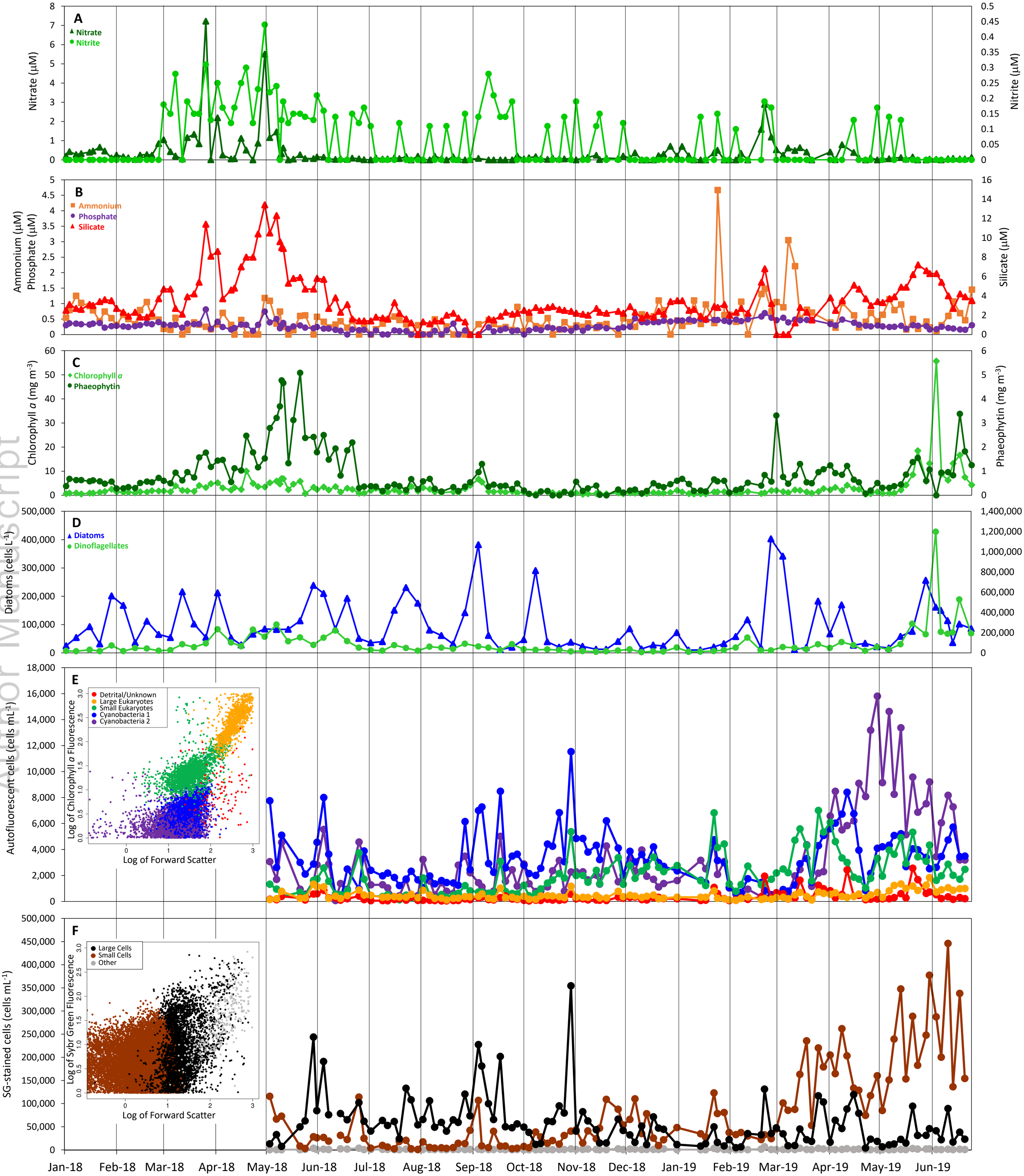
Fig. 4) Spearman Rank correlations relating untransformed ecological variables to the "eigengene" (first principal component of the expression matrix) of each weighted gene correlation network analysis (WGCNA) subnetwork found at a miniature module size of 50. Positive relationships are in red and negative relationships are in blue. The top number is the ρ value and the number in parentheses is the p-value for each relationship. Daily average temperature, hydrostatic pressure (tidal height), and salinity behaved similarly to variables at 10 AM and so were not included. Autofluorescent (AF) and SYBR-Green I (SG)-stained populations were identified using flow cytometry and a self-organizing map.

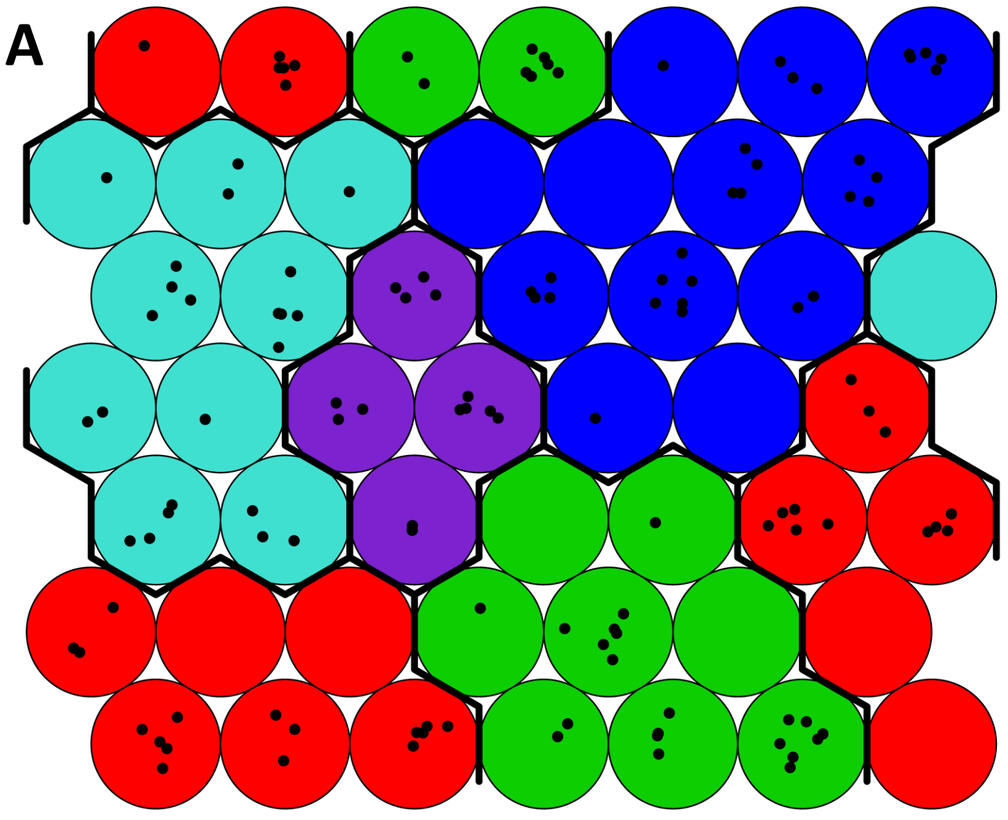
Fig. S1) A) Time-series of the daily average temperature (blue), temperature at 10:00 AM (red), and the 29-day average temperature (indicative of season; black). B) Time-series of the daily average salinity (blue), salinity at 10:00 AM (red), and the daily average temperature (black). C) Time-series of the daily average hydrostatic pressure (blue), pressure at 10:00 AM (red), and the daily average temperature (black). D) Time-series of the daily average wind speed (blue) and the daily average temperature (black). E) Time-series of the daily average wind direction (blue) and the daily average temperature (black).

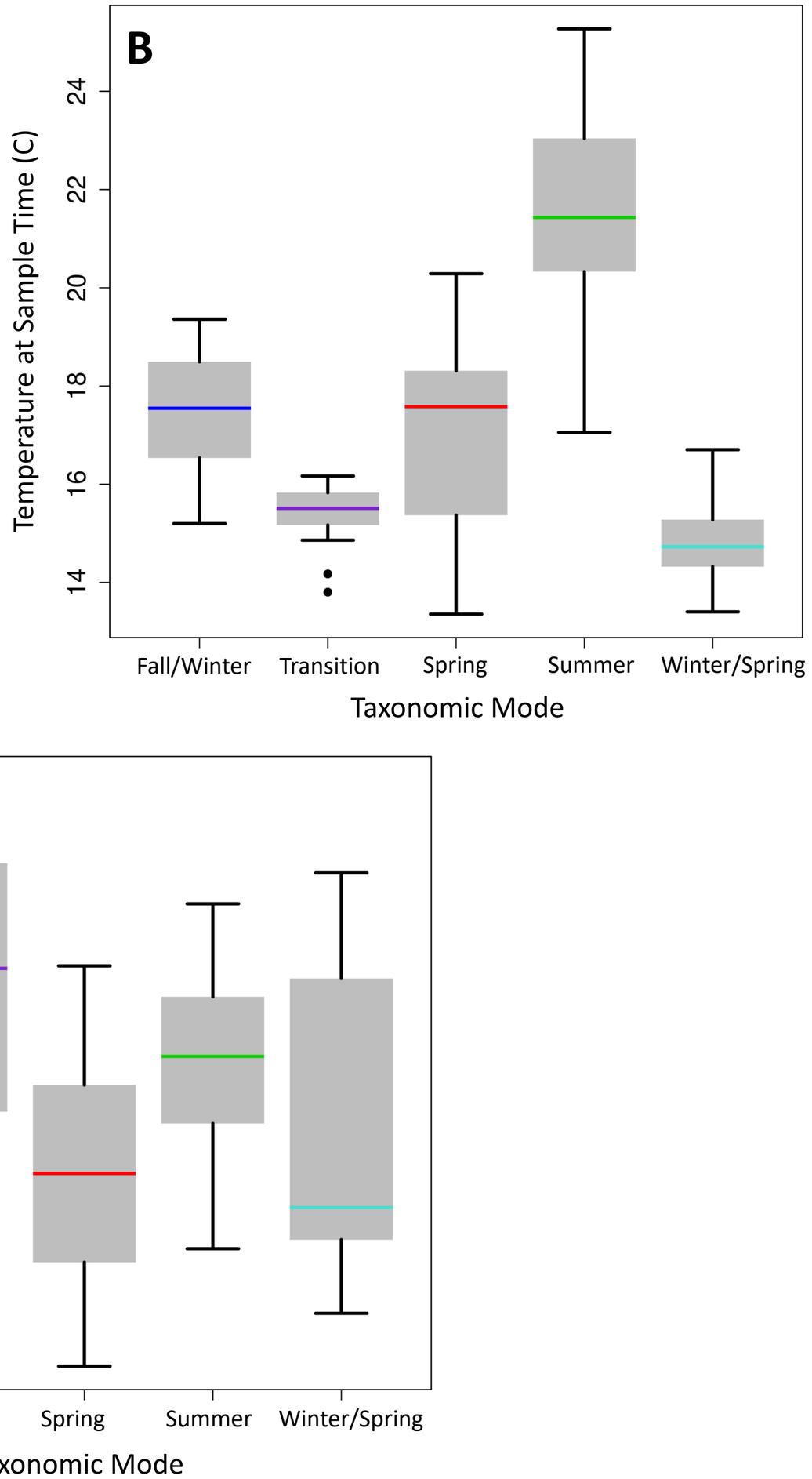
Fig. S2) Pearson Rank correlations between flow cytometry autofluorescent (AF) populations and SYBR-Green I (SG)-stained populations vs. A) other flow cytometry populations and B) abiotic environmental data, nutrients, ln chlorophyll *a*, ln phaeophytin, diatoms, and dinoflagellates. Daily average temperature behaved similarly to the 29-day average (seasonal) and 10 AM sampled water temperature and so was not included. Positive relationships are in red and negative relationships are in blue. The top number is ρ and the number in parentheses is the p-value for each relationship. Only significant relationships are shown, with a significance threshold of $p \leq 0.00027$ selected based on a Bonferroni correction.

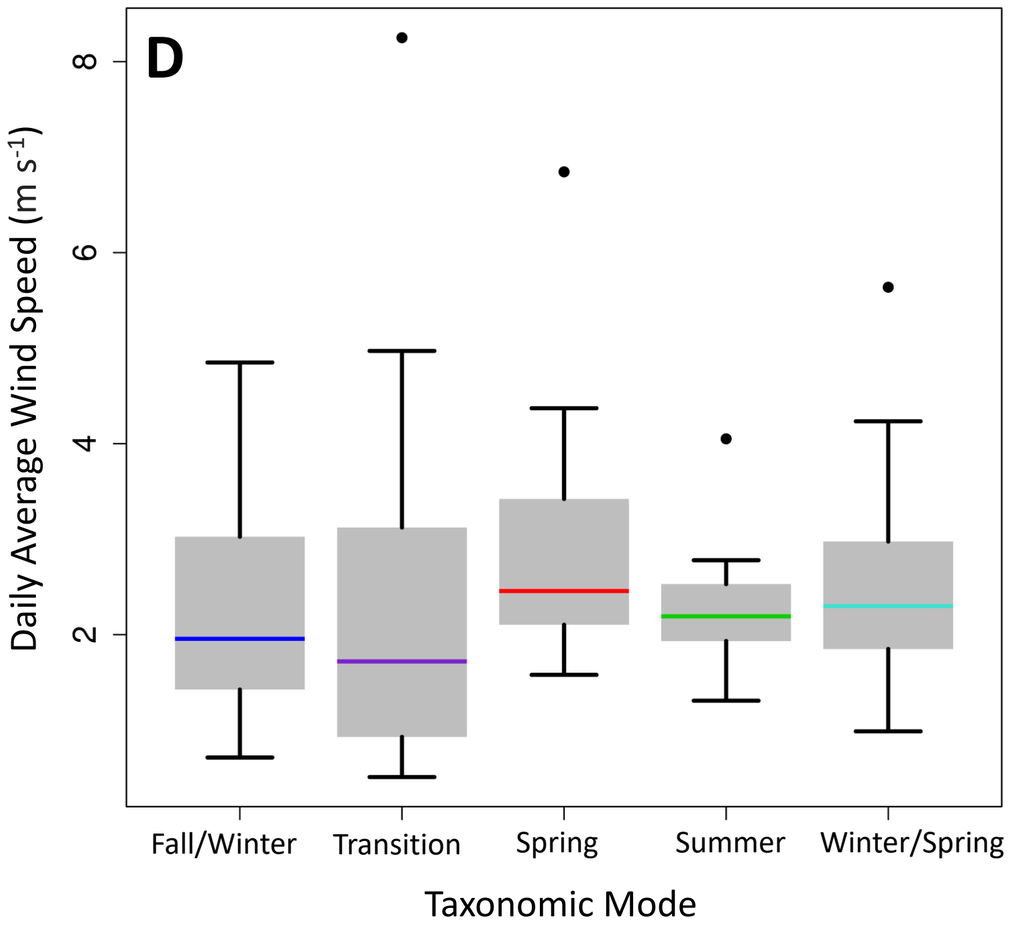
Fig. S3) Concentration/count ranges for the different taxonomic modes (TM) identified using a self-organizing map (7x7; see Fig 2A) according to A) nitrate, B) nitrite, C) ammonium, D) phosphate, E) silicate, F) ln chlorophyll *a*, G) ln phaeophytin, H) total phytoplankton, and I) the inverse Simpson's diversity index. All variables were measured twice a week except for total phytoplankton, which was enumerated weekly. For each boxplot, the horizontal line represents the median (and is colored according to the TM it belongs to), the grey box spans the first quartile to the third quartile, the whiskers show the maximum and minimum, and any dots represent outliers. Tukey's HSD found differences in the means for nitrate between the "summer" vs. the "winter/spring" TMs; nitrite between the "spring" vs. the "fall/winter" and "transition" TMs; ammonium between the "summer" vs. the "winter/spring" TMs; phosphate between the "summer" vs. all other TMs and also between the "winter/spring" vs. the "fall/winter" and "spring" TMs; silicate between the "spring" vs. all other TMs and also between the "summer" vs. the "winter/spring" TMs; ln chlorophyll *a* concentration between the "spring" vs. all other TMs and also between the "fall/winter" vs. the "summer" and "winter/spring" TMs; ln phaeophytin concentration between the "spring" vs. all other TMs except the "winter/spring" TM and also between the "winter/spring" vs. the "fall/winter" and "summer" TMs; total phytoplankton count between the "fall/winter" vs. the "spring" TMs; and diversity between the "fall/winter" vs. all other TMs except the "transition" TM, between the "transition" vs. the "spring" and "summer" TMs, and also between the "summer" vs. the "winter/spring" TMs.

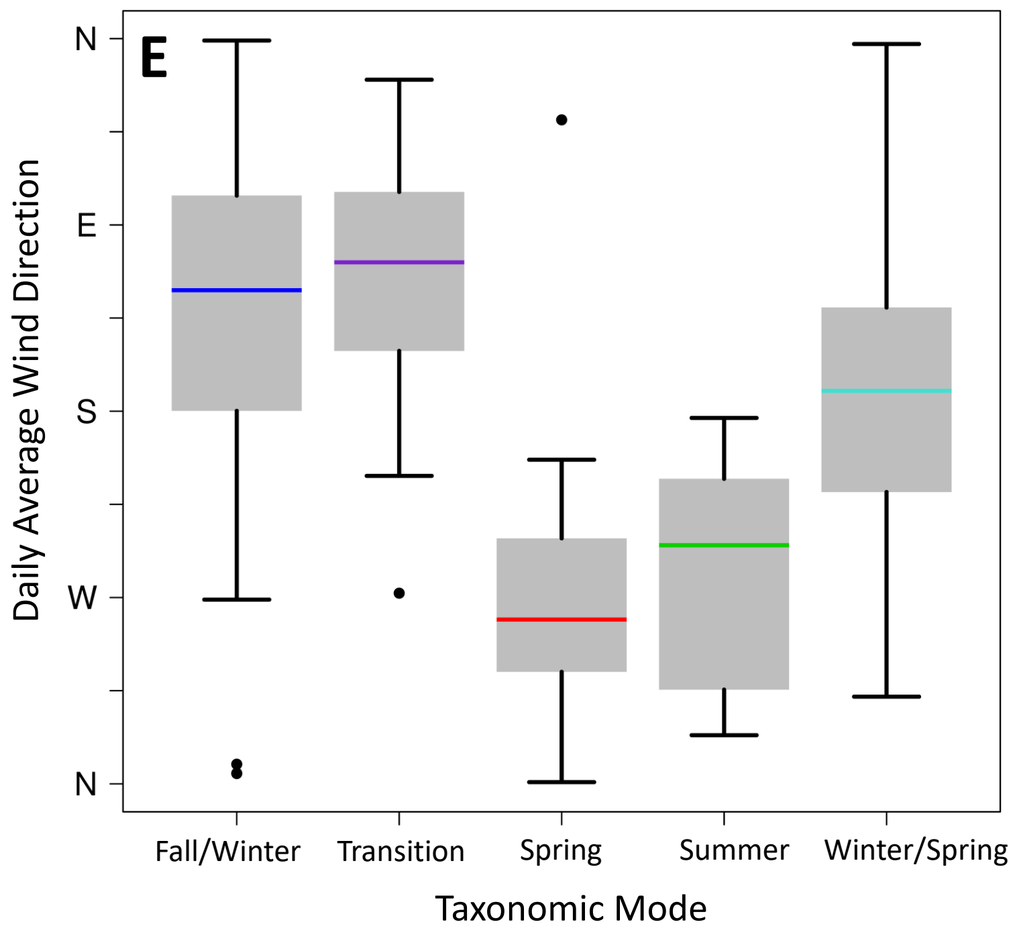
Fig. S4) Count ranges for flow cytometric determined autofluorescent (AF) populations (A-E), microscopically determined F) total diatoms, G) total dinoflagellates, and flow cytometric determined SYBR-Green I (SG)-stained cells (H, I) for the different taxonomic modes identified using a self-organizing map (7x7; see Fig 2A). Total diatoms and dinoflagellates were enumerated weekly while all flow cytometry counts took place twice a week. For each boxplot, the horizontal line represents the median (and is colored according to the TM it belongs to), the grey box spans the first quartile to the third quartile, the whiskers show the maximum and minimum, and any dots represent outliers. Tukey's HSD found differences in the means for the AF detrital/unknown population between the "fall/winter" vs. the "winter/spring" TMs and also between the "summer" vs. the "spring" and "winter/spring" TMs; the AF large eukaryote population between the "spring" vs. all other TMs; the AF small eukaryote population between the "summer" vs. all other TMs except the "transition" TM and also between the "spring" vs. the "winter/spring" TMs; the AF Cyanobacteria 2 population between the "spring" vs. all other TMs; the total diatom abundance between the "fall/winter" vs. the "winter/spring" TMs; the total dinoflagellate abundance between the "spring" vs. the "fall/winter" and "summer" TMs; and the SG-stained small population between the "spring" vs. the "fall/winter" and "summer" TMs and also between the "winter/spring" vs. the "fall/winter" and "summer" TMs. There were no significant differences between the means for the other variables. Unshown are the TMs according to the SG-stained other population, which had no significant differences between the TM means.

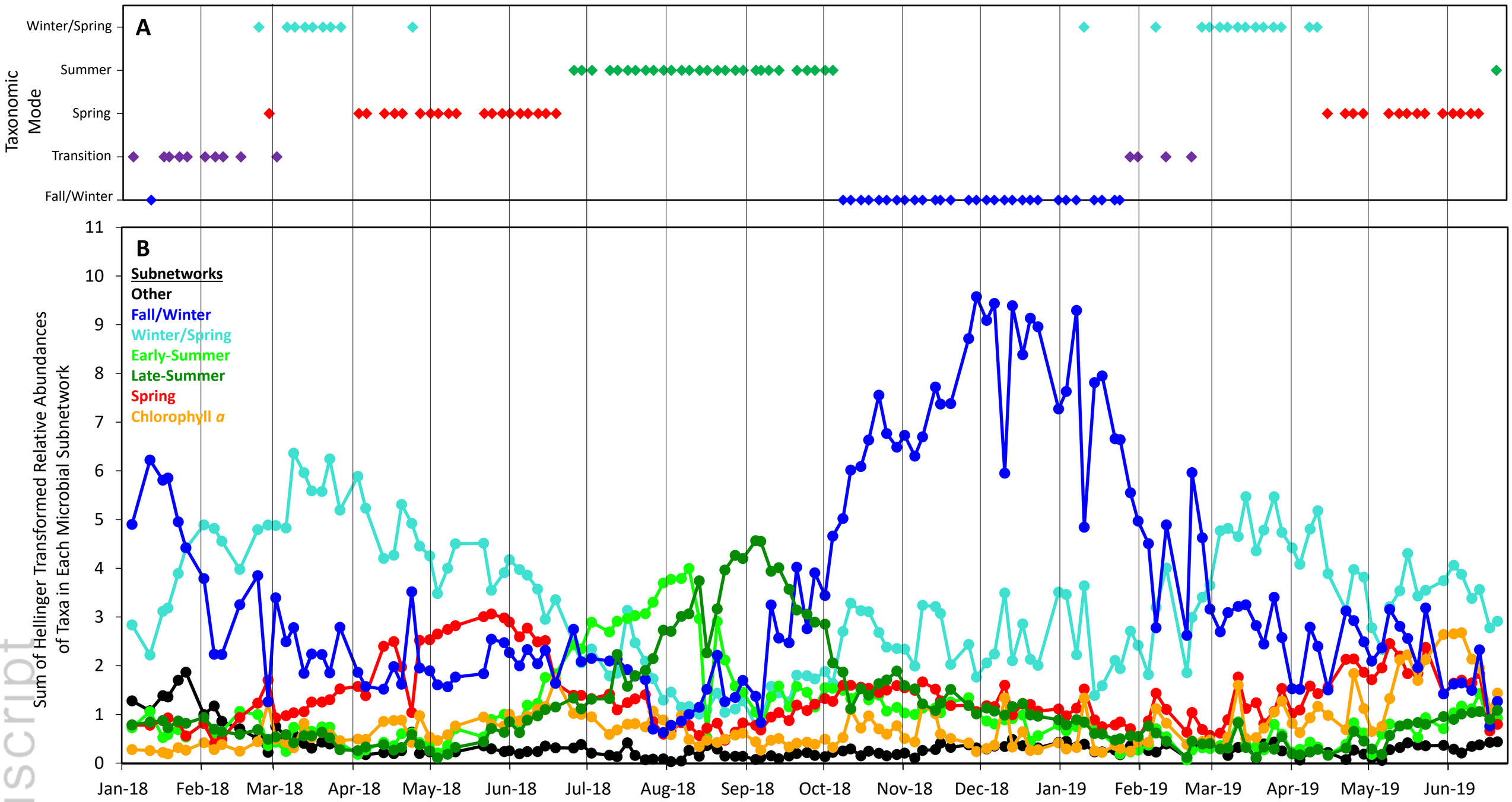


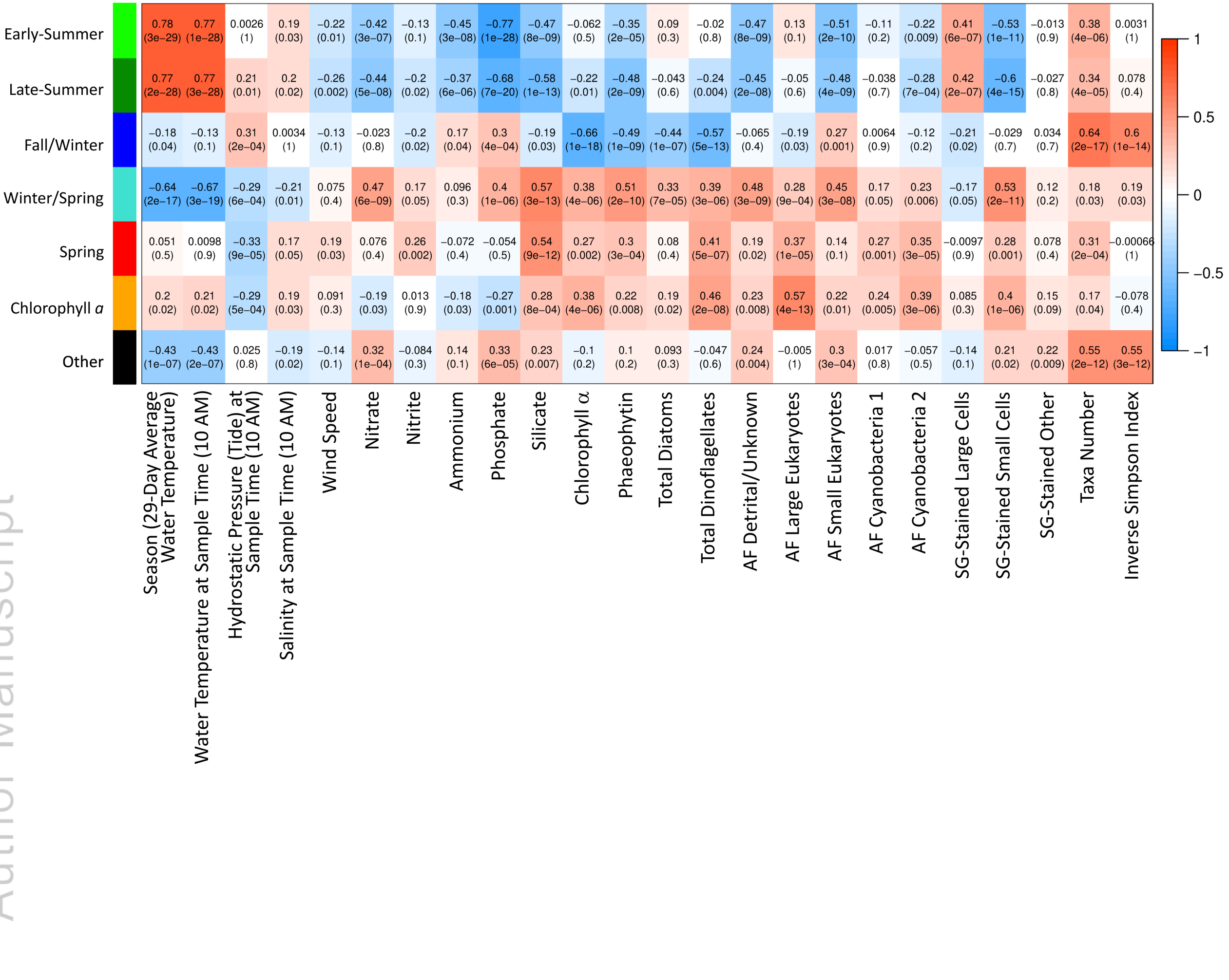












	Taxonomic Mode				
	Fall/Winter	**Transition**	**Spring**	**Summer**	**Winter/Spring**
Observed	Fall/Winter	Winter	Spring	Summer/Early Fall	Winter/Spring
Nutrient Characteristics	Low nitrite	Low nitrite and high phosphate	High nitrite and silicate	Low nitrate, ammonium, phosphate, and silicate	High nitrate, ammonium, and phosphate
Phytoplankton Characteristics	Low chlorophyll *a*, phaeophytin, diatoms, and dinoflagellates	Low chlorophyll *a*, phaeophytin, and dinoflagellates	High chlorophyll *a*, phaeophytin, and dinoflagellates	Low phaeophytin and dinoflagellates	High diatoms and low dinoflagellates
Diversity Characteristics	High diversity	High diversity	Medium diversity	Low diversity	Medium diversity
AF Population Characteristics		High unknown/detrital and low large eukaryotic populations	High large eukaryotic and Cyanobacteria 2 populations	Low small eukaryotic population	High unknown/detrital and small eukaryotic populations
SG-Stained Population Characteristics			High small cell population		High small cell population
ID of Average Top 16S rRNA Gene Sequences (Hellinger Transformed Relative Abundance)*	*Candidatus Pelagibacter* (0.31 ± 0.04) *Candidatus Pelagibacter* (0.20 ± 0.03) *Candidatus Pelagibacter* (0.19 ± 0.03) *Candidatus Pelagibacter* (0.18 ± 0.02) Cyanobacteria/Chloroplast (0.16 ± 0.03)	Cyanobacteria/Chloroplast (0.23 ± 0.06) *Candidatus Pelagibacter* (0.22 ± 0.05) Unidentified Rhodobacteraceae (0.20 ± 0.04) Unidentified Acidimicrobiales (0.19 ± 0.04) Unidentified Rhodobacteraceae (0.18 ± 0.03)	*Candidatus Pelagibacter* (0.32 ± 0.07) Cyanobacteria/Chloroplast (0.21 ± 0.06) Cyanobacteria/Chloroplast (0.20 ± 0.05) Unidentified Acidimicrobiales (0.20 ± 0.08) Unidentified Alphaproteobacterium (0.19 ± 0.03)	Unidentified Rhodobacteraceae (0.35 ± 0.09) *Candidatus Pelagibacter* (0.23 ± 0.07) Cyanobacteria/Chloroplast (0.16 ± 0.07) Cyanobacteria/Chloroplast (0.15± 0.05) Unidentified Rhodobacteraceae (0.15 ± 0.02)	*Candidatus Pelagibacter* (0.27 ± 0.06) Unidentified Gammaproteobacteria (0.21 ± 0.04) Unidentified Rhodobacteraceae (0.20 ± 0.06) Unidentified Rhodobacteraceae (0.19 ± 0.05) Cyanobacteria/Chloroplast (0.17 ± 0.06)

*The classification listed is the finest scale that agreed between paprica and RDP